Parasitology Research
https://doi.org/10.1007/s00436-021-07159-y

GENETICS, EVOLUTION, AND PHYLOGENY - ORIGINAL PAPER

First next-generation sequencing data for Haploporidae (Digenea: Haploporata): characterization of complete mitochondrial genome and ribosomal operon for *Parasaccocoelium mugili* Zhukov, 1971

Dmitry M. Atopkin[1,2] · Alexander A. Semenchenko[3] · Daria A. Solodovnik[1] · Yana I. Ivashko[1,2] · Kirill A. Vinnikov[3,4]

Received: 12 February 2021 / Accepted: 7 April 2021

Abstract

The first data on a whole mitochondrial genome of Haploporidae, *Parasaccocoelium mugili* (Digenea: Haploporata: Haploporidae) was generated using the next-generation sequencing (NGS) approach. We sequenced the complete mitochondrial DNA (mtDNA) and ribosomal operon of *Parasaccocoelium mugili*, intestine parasite of mullet fish. The mtDNA of *P. mugili* contained 14,021 bp, including 12 protein-coding genes, two ribosomal genes, 22 tRNA genes, and non-coding region. The ribosomal operon of *P. mugili* was 8308 bp in length, including 18S rRNA gene (1981 bp), ITS1 rDNA (955 bp), 5.8S rRNA gene (157 bp), ITS2 rDNA (268 bp), 28S rRNA gene (4180 bp), and ETS (767 bp). We used the mtDNA protein-coding regions to make phylogenetic reconstructions of Haploporidae. Additionally, we performed the sequence cluster analysis based on codon usage bias of most of currently available mitochondrial genome data for trematodes. The observed gene arrangement in mtDNA sequence of *P. mugili* is identical to those of *Plagiorchis maculosus* (Rudolphi, 1802). Results of maximum likelihood (ML) phylogenetic analysis showed that *P. mugili* was closely related to *Paragonimus* species from the suborder Xiphidiata. The results of sequence cluster analysis based on codon usage bias showed that *P. mugili* has the highest similarity with *Plagiorchis maculosus* (Xiphidiata). Our results do not contradict to proposing a new suborder for Haploporoidea–Haploporata. On the basis of obtained results, the relationship between mitochondrial protein-coding gene rearrangements and synonymous nucleotide substitutions in mitochondrial genomes has been suggested.

Keywords Haploporidae · NGS · Digenea · Mitochondrial genome · Codon usage

Introduction

Family Haploporidae Nicoll, 1914 represents a group of intestine trematodes infecting estuarine, marine, and freshwater fish species. Representatives of Haploporidae are characterized by small body size, the presence of an armed (Pseudohaploporinae Atopkin, Besprozvannykh, Ha, Nguyen, Nguyen, Chalenko, 2019, some of Haploporinae Nicoll, 1914) or unarmed (other haploporids) hermaphroditic sac that encloses the male and female terminal genitalia and one or two (Megasoleninae Manter, 1935, Pseudohaploporinae) testes. Also, haploporid cercariae are unarmed with a stylet. For this reason, the phylogenetic position of Haploporidae within the suborder Xiphidiata Olson, Cribb, Tkach, Bray, Littlewood, 2003, raised by Olson et al. (2003), was recognized as possible phylogenetic misplacement. However, Pérez-Ponce de León and Hernández-Mena (2019), based on phylogenetic analysis of the 28S rDNA sequence data, proposed a new suborder, Haploporata Pérez-Ponce de León and Hernández-Mena 2019 for representatives of Haploporoidea Nicoll, 1914, which includes the Haploporidae and Atractotrematidae Yamaguti, 1939 families. The authors argued their decision, stating that members of the new suborder were characterized by a lack of stylet in cercariae as a unique feature, the presence of a hermaphroditic sac

Section Editor: Christoph G. Grevelding

✉ Dmitry M. Atopkin
atop82@gmail.com

[1] Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the RAS, Vladivostok, Russia

[2] Department of Cell Biology and Genetics, Far Eastern Federal University, Vladivostok, Russia

[3] Laboratory of Ecology and Evolutionary Biology of Aquatic Organisms, Far Eastern Federal University, Vladivostok, Russia

[4] Laboratory of Genetics, National Scientific Center of Marine Biology, Far Eastern Branch of the RAS, Vladivostok, Russia

Published online: 24 April 2021

considered synapomorphy for this group and resolving position of Haploporoidea on the 28S rDNA-based phylogenetic tree. Presumably, mitochondrial genome based phylogenetic analysis could be more informative in resolving phylogenetic relationships of haploporid trematodes that was a main scope of the present study. Pérez-Ponce de León and Hernández-Mena 2019 reported that with the still limited number of digenean species for which complete mitochondrial genomes have been sequenced, it is not possible to determine the power of the phylogenetic signal of mitogenomes to resolve phylogenetic relationships at deeper levels of the classification of Digenea, in comparison with the resolution power of single nuclear rDNA sequences. Nevertheless, the digenean mitogenome dataset has to be increased with new data for taxonomically problematic groups, such as Haploporidae, to resolve its relationships with a greater number of molecular characters. In the present study, we provide the first data on the whole mitochondrial genome for a representative of Haploporidae, *Parasaccocoelium mugili* Zhukov, 1971, a parasite of mullet fish, sampled in the Russian Far East, generated using next-generation sequencing (NGS) approach, with phylogenetic relationship reconstructions of Haploporidae on the basis of proteins encoding part of the mtDNA. Additionally, we provide sequence data and annotation for the complete ribosomal operon of *P. mugili*.

Material and methods

Sample collection and DNA extraction

Adult worms were collected from intestine of single fish specimen of *Planiliza haematochelia* (Temminik & Schlegel, 1845) (Mugilidae Jarocki, 1822) from the estuary of River Kievka, Primorksy Region, south of the Russian Far East, during parasitological field work in July 2018. Trematodes were killed with hot water and then fixed in 96% ethanol. Total DNA was extracted from 40 worms simultaneously with Qiamp Investigator Kit, Qiagen, according to manufacturer's protocol. Amount of total DNA was measured with Qubit Fluorometer 3.0, Invitrogen, and then used for NGS sequencing in final 2 ng/μl.

Preparing genome library for NGS

Libraries were prepared using an Ion Plus Fragment Library Kit and unique adapters (Ion Xpress, Waltham, MA, USA) with pre-fragmentation on a Covaris M220 Focused ultrasonicator. The emulsion PCR and template preparation were obtained on an Ion One Touch2 System (Thermo Fisher Scientific) followed by sequence on an Ion S5 sequencing platform using Ion 540 chip at the Far Eastern Federal University (Vladivostok, Russia). Ambiguous parts of genome sequence were reexamined with Sanger's sequencing using highly specific oligonucleotide primers, developed for this study.

The quality of raw reads were checked using FastQC 0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and then reads were assemble using SPAdes 3.14.1 (Nurk et al. 2013) with correction of IonTorrent data using the IonHammer tool available in the software SPAdes. The scaffolds, containing mitochondrial and ribosomal operon DNA data were manually assembled in the MEGA X software (Kumar et al. 2018).

Mitochondrial genome annotation was performed with MITOS on-line software, available at http://mitos2.bioinf.unileipzig.de/index.py n.d.. Searching of tandem repeats completed with Tandem Repeat Finders software (Benson 1999). The nucleotide sequences of both ribosomal operon and mitochondrial complete genome were deposited in GenBank under accession numbers MW813991 and MW846232, respectively.

Codon usage and phylogenetic analyses

Nucleotide and amino-acid sequence alignments were performed with ClustalW algorithm in the MEGA X software. The poorly aligned regions were removed using Gblocks Server v. 0.91b (http://molevol.cmima.csic.es/castresana/Gblocks_server.html n.d.).

Codon usage statistic was calculated for concatenated protein-coding gene sequence data with MEGA X software. Sequence cluster analysis based on codon usage bias was performed with Statistica 13 software (TIBCO Software Inc. 2017) using weighted pair group average joining (tree clustering) method with Euclidian distances calculation. Phylogenetic analysis was performed on the basis of concatenated amino-acid sequences with Maximum likelihood (ML) algorithm, realized using PhyML 3.1 (Guindon and Gascuel 2003) software. The ML algorithm was performed using LG evolutionary model (Le and Gascuel 2008), SPRs tree topology search and random sequence addition. Statistical support for maximum likelihood algorithm was estimated with Bayesian posterior probabilities calculation (Huelsenbeck et al. 2001). The phylogenetic relationships were inferred using our sample and trematode species from the NCBI GenBank database (Table 1).

Results

Parasaccocoelium mugili nuclear ribosomal operon annotation

The ribosomal operon of *P. mugili* was 8308 bp in length, including the 18S rRNA gene (1981 bp), ITS1 rDNA (955 bp), 5.8S rRNA gene (157 bp), ITS2 rDNA (268 bp), 28S rRNA gene (4180 bp) and ETS (767 bp). The nucleotide composition of ribosomal operon of *P. mugili* was as follows: A—

Table 1 List of taxa, incorporated into analysis

Species	GenBank accession number	Reference
Xiphidiata		
Brachycladium goliath	KR703278	Briscoe et al. (2016)
Dicroroelium chinensis	KF318786	Liu et al. (2014a)
Dicrocoelium dendriticum	KF318787	Liu et al. (2014a)
Eurytrema pancreaticum	KP241855	Chang et al. (2016)
Paragonimus heterotremus	MH059809	Qian et al. (2018)
Paragonimus kellicotti	MH322000	Wang et al. (2018)
Paragonimus ohirai	KX765277	Le et al. (2019)
Paragonimus westermani	KX943544	Biswal et al. (2014)
Plagiorchis maculosus	MK641809	Suleman et al. (2019a)
Echinostomata		
Echinostoma caproni	AP017706	Holroyd et al. (2016), unpublished
Echinostoma hortense	KR062182	Liu et al. (2016)
Echinostoma miyagawai	MH393928	Fu et al. (2019a)
Echinochasmus japonicus	KP844722	Le et al. (2016)
Fasciola hepatica	AF216697	Le et al. (2000)
Fasciola gigantica	KF543342	Liu et al. (2014b)
Fasciola sp.	KF543343	Liu et al. (2014b)
Fasciolopsis buski	KX169163	Ma et al. (2016), unpublished
Fascioloides magna	KU060148	Ma et al. (2016), unpublished
Hypoderaeum conoideum	KM111525	Yang et al. (2015)
Pronocephalata		
Acanthoparyphium sp.	MG792058	Kandari et al. (2018), Unpublished
Tracheophilus cymbius	MK355447	Li et al. 2019
Uvitellina sp.	MK227160	Suleman et al. (2019b)
Calicophoron microbothrioides	KR337555	Ma et al. (2015), unpublished
Explanatum explanatum	KT198989	Ma et al. (2015), unpublished
Fischoederius cobboldi	KX169164	Ma et al. (2016), unpublished
Fischoederius elongatus	KM397348	Fang (2014), unpublished
Gastrothylax crumenifer	KM400624	Yang et al. (2016)
Homalogaster paloniae	KX169165	Ma et al. (2016), unpublished
Ogmocotyle sikae	KR006934	Ma et al. (2015), unpublished
Orthocoelium streptocoelium	KM659177	Yang (2014), unpublished
Paramphistomum cervi	KF475773	Yan et al. (2013)
Hemiurata		
Azygia hwangtsiyui	MN844889	Wu et al. (2020)
Opisthorchiata		
Amphimerus sp.	MK238506	Ma et al. (2019)
Clonorchis sinensis	FJ381664	Shekhovtsov et al. (2010)
Haplorchis taichui	KF214770	Lee et al. (2013)
Metagonimus yokogawai	KC330755	Jeon et al. (2012), unpublished
Metorchis orientalis	KT239342	Na et al. (2016)
Opisthorchis felineus	EU921260	Shekhovtsov et al. (2010)
Diplostomata		
Clinostomum complanatum	KM923964	Chen (2015), unpublished
Cyathocotyle prussica	MH536510	Locke et al. (2018)

Table 1 (continued)

Species	GenBank accession number	Reference
Postharmostomum commutatum	MN200359	Fu et al. (2019b)
Schistosoma bovis	CM014335	Oey et al. (2019)
Schistosoma haematobium	DQ157222	Littlewood et al. (2006)
Schistosoma japonicum	AF215860	Le et al. (2000)
Schistosoma mekongi	AF217449	Le et al. (2000)
Schistosoma spindale	DQ157223	Littlewood et al. (2006)
Trichobilharzia regenti	DQ859919	Webster et al. (2007)
Trichobilharzia szidati	MF136777	Semyenova et al. (2017)
Outgroup (Cestoda)		
Diphyllobothrium latum	DQ985706	Park et al. (2007)

23%; T (U)—26.9%; C—21.4%; and G—28.7%. These data represents just additional information for rDNA of Haploporidae and will not be analyzed and discussed here.

General characteristics of the *Parasaccocoelium mugili* mitochondrial genome

The whole-genome sequence was obtained from total DNA that was extracted from 40 specimens of *P. mugili*, which possess some level of intraspecific variation of mtDNA (Atopkin et al., 2019) revealing 28 variable positions within the same reads of different parts of the genome sequence (Table 2). These variable sites were taken into account during genome assembly, annotation and phylogenetic reconstructions. In total, the mitochondrial genome of *P. mugili* contained 14,021 bp, with 12 protein-coding genes, two ribosomal genes, 22 tRNA genes, and non-coding region including two tandem repeats (TR) and one unique sequence (US) (Fig. 1, Table 3). Gene arrangements of whole mt-genome sequence of *P. mugili* were identical to those of *Plagiorchis maculosus* (Rudolphi, 1802), except for a single difference in the non-coding region for the last species. The nucleotide composition in the *P. mugili* whole mitochondrial genome was as follows: 48.1%—T (U); 10.5%—C; 19.8%—A; and 21.6%—G. Nucleotide pair frequency was 67.9% for the AT content, and 32.1% for the GC content, showing a bias towards T over A (AT skew = −0.42) and G over C (CG skew = 0.35), respectively.

Protein-coding genes

The complete sequence length of 12 protein-coding genes was 9968 bp. The order of these genes *cox*3-*cyt*b-*nad*4L-*nad*4-*atp*6-*nad*2-*nad*1-*nad*3-*cox*1-*cox*2-*nad*6-*nad*5 is identical to all Xiphidiata and representatives of other suborders of Plagiorchiida La Rue, 1957 (Biswal et al. 2014; Briscoe et al. 2016; Chang et al. 2016; Le et al. 2019; Liu et al.

Table 2 Variable positions within the same reads of different parts of the mitochondrial genome sequence of 40 *Parasaccocoelium mugili* specimens

Gene	Gene site number (in whole genome)	Substitution (UIPAC code)
*cox*3	207 (207)	C/T (Y)
*cyt*b	732 (1457)	C/T (Y)
*cyt*b	735 (1460)	A/G (R)
*cyt*b	753 (1478)	A/T (W)
*cyt*b	768 (1493)	A/G (R)
*cyt*b	774 (1499)	A/G (R)
*cyt*b	804 (1529)	C/T (Y)
*nad*4	118 (2183)	C/T (Y)
*atp*6	54 (3597)	A/G (R)
*nad*2	270 (4335)	A/G (R)
*nad*2	396 (4461)	C/T (Y)
*cox*1	34 (6848)	C/T (Y)
*cox*1	1098 (7912)	C/T (Y)
*cox*1	1182 (7996)	A/G (R)
*rrn*L	141 (8561)	C/T (Y)
*nad*6	40 (10808)	A/G (R)
tRNA-Leu (L1)	61 (11356)	G/C (S)
*nad*5	6 (11614)	C/T (Y)
*nad*5	104 (11712)	C/T (Y)
*nad*5	507 (12115)	A/G (R)
*nad*5	679 (12287)	A/G (R)
*nad*5	780 (12388)	A/G (R)
*nad*5	1320 (12928)	C/G/T(B)
*nad*5	1329 (12937)	C/G/T(B)
*nad*5	1321 (12938)	C/T (Y)
tRNA-Glu (E)	38 (13207)	C/T (Y)
TR1	102 (13420)	A/T (W)
TR2	101 (13704)	A/T (W)

2014a; Qian et al. 2018; Suleman et al. 2019a; Wang et al. 2018). Start codons for protein-coding genes were ATG. Terminal codons were TAA or TAG. The nucleotide composition of the assembled protein-coding part of the mt-genome sequence was as follows: A—17.6%; T (U)—50.6%; C—10.1%; and G—21.7%; the AT content was 68.2% (AT skew = −0.48) and the GC content was 31.8% (CG skew = 0.36). Codon usage statistics for *P. mugili* agree with the nucleotide composition ratio: the most common triplets contained T (U) and/or A bases, namely UUU (frequency = 13.8%), UUA (frequency = 4.9%) and UAU (frequency = 4.8%). A total of 3348 amino acids were encoded by the mitochondrial protein-coding genes of *P. mugili*.

Phylogenetic analysis

The maximum likelihood algorithm was used on the basis of the 2339 amino-acid alignment length, available after Gblocks processing. The ML tree topology showed that all of the used digeneans could be subdivided into two large clades (Fig. 2). The Clade I consists of seven species of the family Schistosomatidae Stiles & Hassal, 1898, while the clade II comprises 43 representatives from 17 families. The Clade II includes the family Haploporidae (*Parasaccocoelium mugili*), which was closely related to *Paragonimus* Dollfus, 1939 species. *Brachicladium goliath* (van Beneden, 1858) from the suborder Xiphidiata was the sister group to both *Paragonimus* species (Xiphidiata) and *P. mugili* (Haploporata). This subclade was a sister to the monophyletic Opisthorchiata La Rue, 1957 group, which contained representatives of Opisthorchiidae Looss, 1899 and Heterophyidae Leiper, 1909. Pronocephalata Olson, Cribb, Tkach, Bray, Littlewood, 2003 and Echinostomata La Rue, 1926, were monophyletic and sister relative to each other. Representatives of Dicrocoeliidae Looss, 1899 (Xiphidiata) were considerably separated from other Xiphidiata and formed a distinct branch, in the same way as *Azygia hwangtsiyui* Tsin, 1933, a representative of Hemiurata Skrjabin & Guschanskaja, 1954.

Sequence cluster analysis based on codon usage bias

Results of cluster analysis based on codon usage bias showed that *P. mugili* has the highest similarity to *Plagiorchis maculosus* (Rudolphi, 1802) according to the frequencies of all 64 codons, which were included in this analysis (Fig. 3a). The results of cluster analysis based on codon usage bias with only the most frequent non-synonymous codons (20 totally) indicate, as in previous analysis, the high similarity of *P. mugili* and *P. maculosus* (Fig. 3b). On the whole, the obtained results demonstrates some agreement with results of ML analysis, indicating the gathering of *Paragonimus* species, Opisthorchiata species (excluding *Opisthorchis felineus* (Rivolta, 1884) Blanchard, 1895), and *Brachycladium goliath* into the same cluster, grouping of representatives of Pronocephalata into same separate cluster and the marked differentiation of schistosomes from other digeneans.

Discussion

Previous results on molecular-based phylogenetic studies of digenetic trematodes indicate that Haploporoidea appears within Xiphidiata and that this position was interpreted as phylogenetic misposition (Olson et al. 2003). Pérez-Ponce de León and Hernández-Mena (2019) proposed a higher taxonomical status for Haploporoidea—a suborder Haploporata. This decision resolves at least two questions: (i) members of Haploporata characterized by a lack of stylet in cercariae as a unique feature, the presence of a hermaphroditic sac considered synapomorphy for this group, and (ii) resolving the phylogenetic relationships of Xiphidiata based on 28S rDNA sequence data. Our results of phylogenetic analyses, based on mitochondrial amino-acid sequences, indicate close relationships of *P. mugili* with some representatives of Xiphidiata, namely *Paragonimus* species (results of ML analysis, Fig. 2) or *Plagiorchis maculosus* (codon-usage-based cluster analysis, Fig. 3a, b). On the one hand, these results agree with the resulting phylogenetic tree from Olson et al. (2003), which showed close relationships between Haploporidae and Paragonimidae Dollfus, 1939. On the other hand, our results do not contradict to the taxonomic conclusion of Pérez-Ponce de León and Hernández-Mena (2019) relative to the

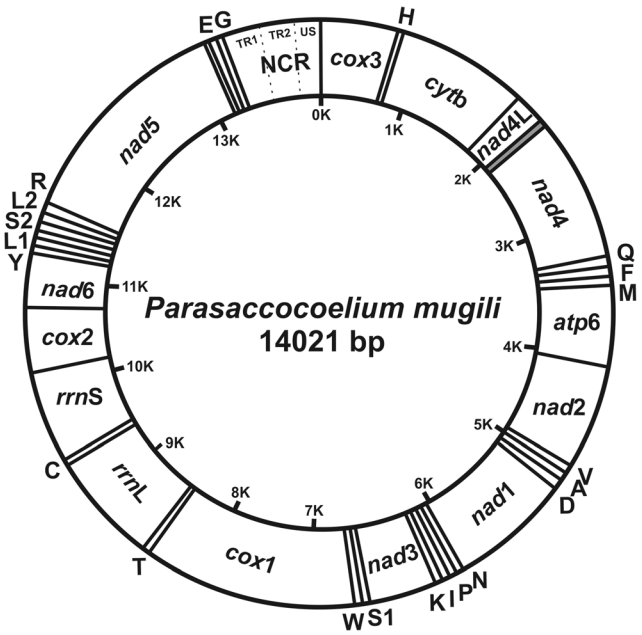

Fig. 1 Organization of the complete mitochondrial genome of *Parasaccocoelium mugili*

establishment of a new suborder for Haploporoidea, Haploporata, because Xiphidiata is obviously paraphyletic in our phylogenetic tree, as also shown in previous phylogenetic studies of trematodes based on complete mitochondrial DNA sequence data (Wang et al. 2018; Le et al. 2019; Suleman et al. 2019b; Wu et al. 2020). However, the clade, comprising *P. mugili*, Paragonimidae species and *Brachycladium goliath* can be considered distinct taxonomical unit by means of phylogenetic interpretation. Additional data on mitochondrial genome sequences for more digenean specimens are needed for more substantive conclusions.

We will not discuss the mitochondrial gene arrangements of all of the available trematode species because this has already been done in detail by Wu et al. (2020). Nevertheless, it has not

Table 3 The organization of mitochondrial genome of *Parasaccocoelium mugili* (*TR*, tandem repeat; *US*, unique sequence; *tRNA-Ser (S1) missed the paired dihydrouridine (DHU) arm)

Gene	Position 5′ to 3′	Length (bp)	Initiation codons	Termination codons	Anti-codons (tRNA)
*cox*3	1–645	645	ATG	TAG	
tRNA-His (H)	662–725	64			GTG
*cyt*b	726–1832	1107	ATG	TAA	
*nad*4L	1842–2105	264	ATG	TAG	
*nad*4	2066–3340	1275	ATG	TAA	
tRNA-Gln (Q)	3353–3415	63			TTG
tRNA-Phe (F)	3416–3478	63			GAA
tRNA-Met (M)	3478–3543	66			CAT
*atp*6	3544–4059	516	ATG	TAA	
*nad*2	4066–4935	870	ATG	TAA	
tRNA-Val (V)	4940–5003	64			TAC
tRNA-Ala (A)	5004–5064	61			TGC
tRNA-Asp (D)	5071–5134	64			GTC
*nad*1	5135–6064	930	ATG	TAA	
tRNA-Asn (N)	6050–6110	61			GTT
tRNA-Pro (P)	6111–6174	64			TGG
tRNA-Ile (I)	6176–6239	64			GAT
tRNA-Lys (K)	6243–6305	63			CTT
*nad*3	6308–6667	360	ATG	TAA	
*tRNA-Ser (S1)	6674–6734	61			GCT
tRNA-Trp (W)	6747–6809	63			TCA
cox1	6815–8356	1542	ATG	TAA	
tRNA-Thr (T)	8358–8417	60			TGT
*rrn*L	8421–9385	965			
tRNA-Cys (C)	9385–9444	60			GCA
*rrn*S	9443–10,162	720			
*cox*2	10,163–10,762	600	ATG	TAG	
*nad*6	10,769–11,218	450	ATG	TAG	
tRNA-Tyr (Y)	11,223–11,289	67			GTA
tRNA-Leu (L1)	11,296–11,360	65			TAG
tRNA-Ser (S2)	11,360–11,427	68			TGA
tRNA-Leu (L2)	11,437–11,499	63			TAA
tRNA-Arg (R)	11,499–11,562	64			ACG
*nad*5	11,609–13,114	1506	ATG	TAA	
tRNA-Glu (E)	13,170–13,235	66			TTC
tRNA-Gly (G)	13,241–13,305	65			TCC
TR1	13,319–13,603	285			
TR2	13,604–13,888	285			
US	13,889–14,021	133			

escaped our notice that the results of sequence cluster analysis based on codon usage bias on the one hand indicate the closeness of *Parasaccocoelium mugili* with *Plagiorchis maculosus* that characterized by high similarity of mitochondrial gene arrangement and on the other hand the separate position of *Schistosoma* Hansen, 1916 species relatively other digeneans, which also possess gene arrangement quite different from other worms. Protein-coding gene rearrangements are not observed for all trematodes within the clade II of the ML tree (Fig. 2) in contrast to the clade I, comprising schistosomes and *Trichobilharzia* Skrjabin & Zakharov, 1920 species. *Schistosoma* species differ considerably from other digeneans, including *Trichobilharzia*, by the arrangement of protein-coding genes (see Wu et al. 2020, Table 4) and occupy a separate position in the cladogram, constructed on the basis of frequency values of all mitochondrial codons (Fig. 3a). However, *Shistosoma* and *Trichobilharzia* are closely related to each other on amino-acid based ML phylogenetic tree and on results of cluster analysis based on frequencies of most frequent non-synonymous codons (20 totally). On this basis, we propose possible relationships between gene rearrangement processes and the occurrence of synonymous nucleotide substitutions, suggested from results of codon usage analysis, within the mitochondrial genomes of Digenea. Such relationships are known for some animal groups. For example, positive correlation rates of gene rearrangement and nucleotide substitution were shown for 20 mitochondrial genomes in insects with deep statistical analysis (Shao et al. 2003). These authors also referenced other studies of different animal groups, namely marine bivalves, snakes, ascidians and some hymenopterans, for which same results have been obtained. In respect to flatworms, Lee et al. (2004) showed that codon usage bias of mitochondrial genes for trematodes and cestodes associated with the phenomenon of skew of complementary bases and proposed a hypothesis for this, developed earlier for mammals, related with asymmetric replication of mitochondrial DNA (Saccone et al. 2002). In 2019, Lamolle et al. performed compositional genome-wide analysis of 22 species of Platyhelminthes, using a set of 700 orthologous nuclear genes to demonstrate relationships of codon usage bias and GC content in genomes of different classes of flatworms to show that GC bias has a great influence on synonymous codons and amino acid usage in Platyhelminthes. In

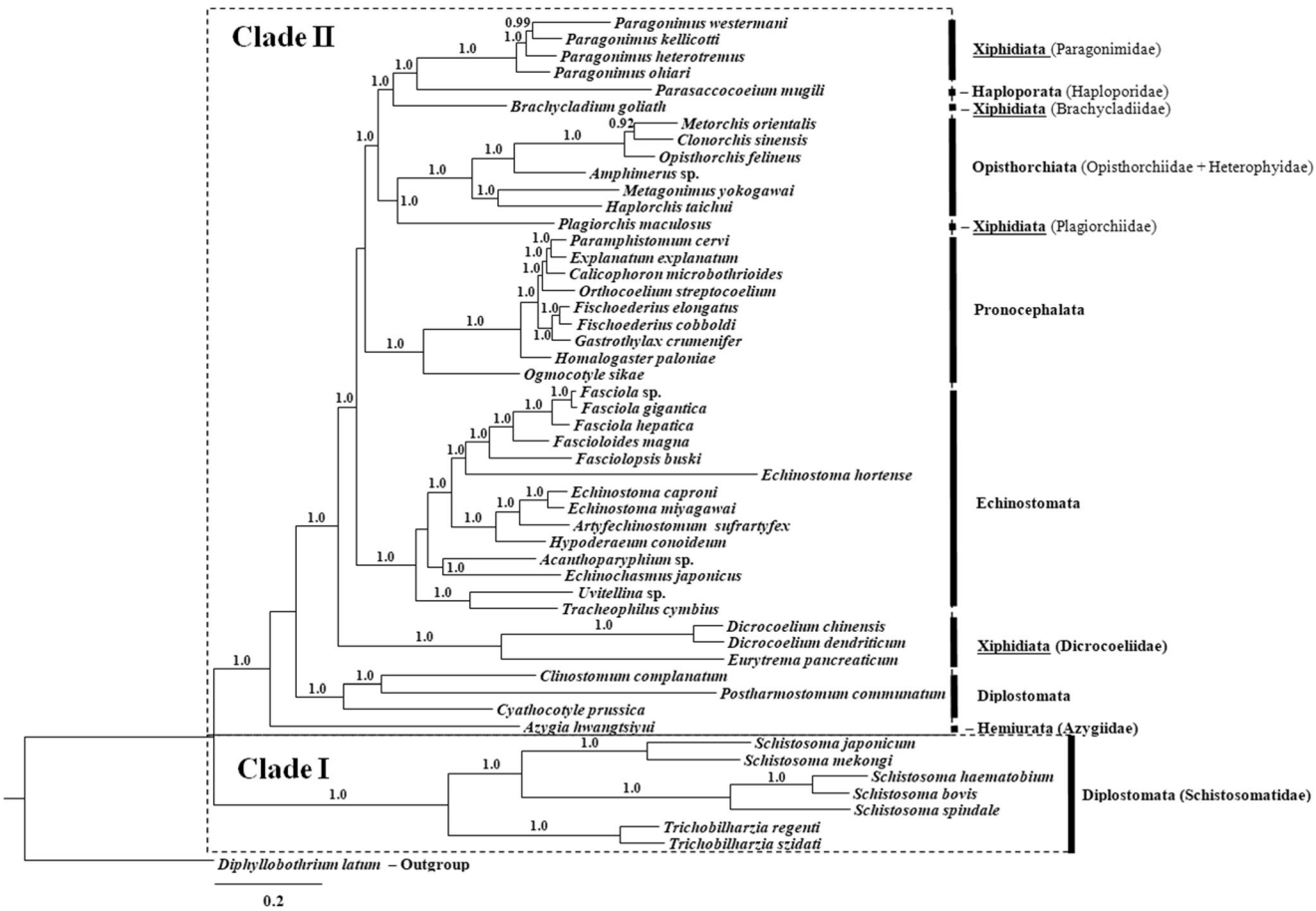

Fig. 2 Phylogenetic relationships of *Parasaccocoelium mugili* and other digenetic trematodes, reconstructed by means of maximum likelihood on the basis of the 2339 amino-acid alignment length, available after Gblock processing. Nodal support showed with posterior probabilities, calculated with Bayesian algorithm

particular, results of codon usage and AT/GC content analysis of *Fasciola hepatica* Linnaeus, 1758 and *Schistosoma mansoni* Sambon, 1907 indicate high frequency of substitutions in 3rd positions of AT-rich codons relative to GC-rich codons,

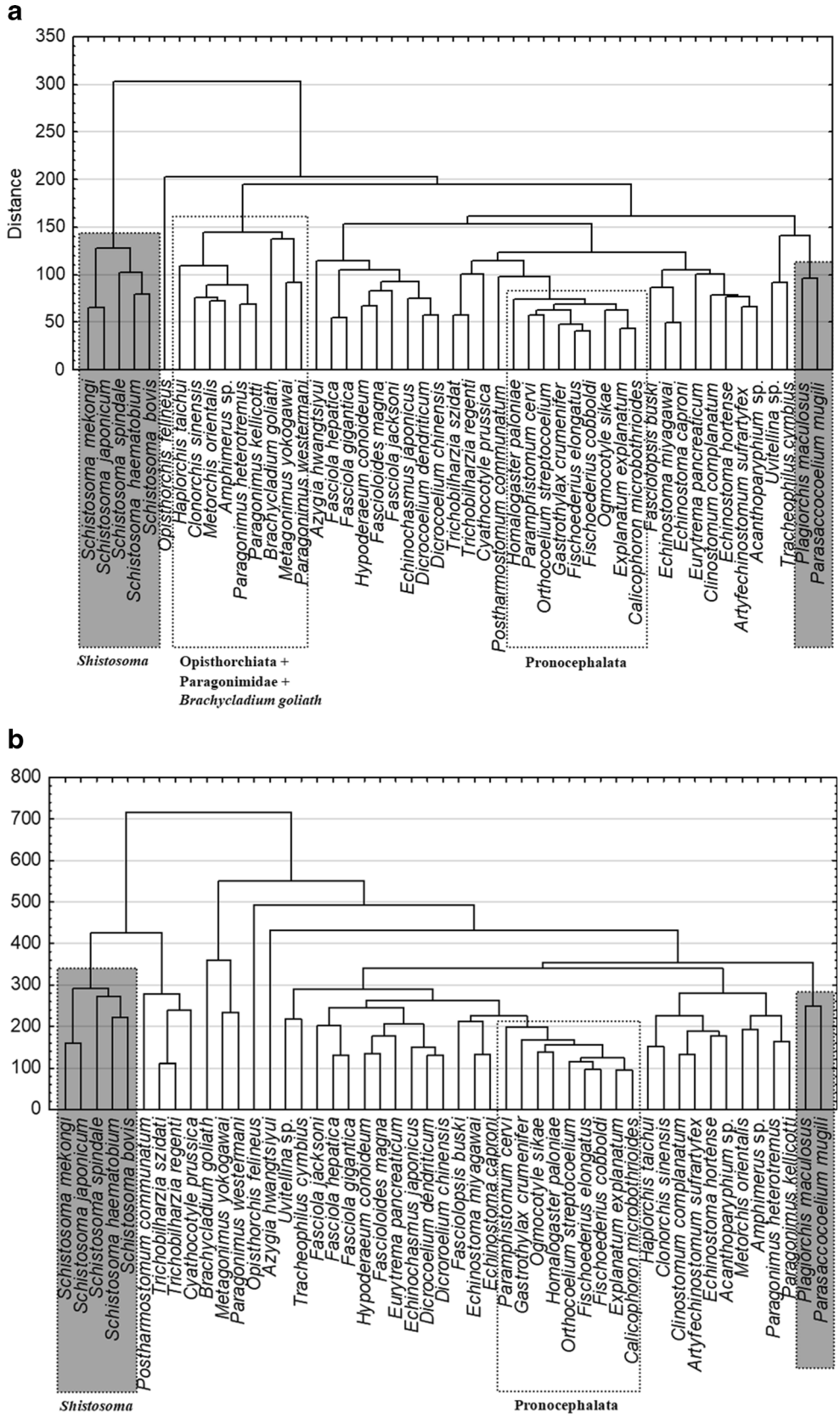

Fig. 3 **a** Results of sequence cluster analysis based on codon usage bias of Digenea, using frequencies of all 64 codons. **b** Results of sequence cluster analysis based on codon usage bias of Digenea, only the most frequent non-synonymous codons (20 totally)

Table 4 Protein-coding gene assembling of *Parasaccocoelium mugili* (*all Xiphidiata and other suborders from present study, excluding *Schistosoma*) and of *Schistosoma* species. Different gene rearrangements are colored with black and grey

	1	2	3	4	5	6	7	8	9	10	11	12
*P. mugili**	***cox*3**	***cyt*b**	***nad*4L**	***nad*4**	***atp*6**	***nad*2**	***nad*1**	***nad*3**	***cox*1**	***cox*2**	***nad*6**	***nad*5**
Shistosoma	***cox*3**	***cyt*b**	***nad*4L**	***nad*4**	***nad*3**	***nad*1**	***cox*1**	***cox*2**	***nad*6**	***atp*6**	***nad*2**	***nad*5**

suggesting maintenance of GC skew and proper amino acid encoding. Along this, a low frequency of GC-rich codons in *S. mansoni*, in contrast to *F. hepatica*, explained by markedly increase of synonymous substitutions towards to AT-rich codons, has been notified. Our calculations indicate that average GC content in mitochondrial protein-coding genes of *Schistosoma* species is also lower (28%) relative to other digeneans (31.8–44.8%), suggesting the same mechanisms of substitution processes in mitochondrial codons, showed for nuclear genes (Lamolle et al. 2019). However, studies by Lee et al. (2004) and Lamolle et al. (2019) did not considered relationships of nucleotide substitution of codons and protein-coding gene rearrangements. We clearly understand indirect nature of arguments from our study because this question is not the main aim of this work. Nevertheless, we share the opinion that the observed agreements of species clustering based on codon usage bias analysis and gene arrangement features of different groups of Digenea deserve attention and need to be studied separately in more detail with the most representative material and strong statistical methods.

Funding This study was supported by Grant of Russian Scientific Foundation, № 17-74-20074.

Declarations

Conflict of interest The authors declare no competing interests.

References

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucl Ac Res 27:573–580. https://doi.org/10.1093/nar/27.2.573

Biswal DK, Chatterjee A, Bhattacharya A, Tandon V (2014) The mitochondrial genome of *Paragonimus westermani* (Kerbert, 1878), the Indian isolate of the lung fluke representative of the family Paragonimidae (Trematoda). PeerJ 2:e484. https://doi.org/10.7717/peerj.484

Briscoe AG, Bray RA, Brabec J, Littlewood DT (2016) The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale. Parasitol Int 65:271–275. https://doi.org/10.1016/j.parint.2016.02.004

Chang Q-C, Liu G-H, Gao J-F, Zheng X, Zhang Y, Duan H, Yue D-M, Fu X, Su X, Gao Y, Wang C-R (2016) Sequencing and characterization of the complete mitochondrial genome from the pancreatic fluke *Eurytrema pancreaticum* (Trematoda, Dicroroeliidae). Gene 576:160–165. https://doi.org/10.1016/j.gene.2015.09.081

Fu Y-T, Jin Y-C, Li F, Liu G-H (2019a) Characterization of the complete mitochondrial genome of the echinostome *Echinostoma miyagawai* and phylogenetic implications. Parasitol Res 118:3091–3097. https://doi.org/10.1007/s00436-019-06417-4

Fu Y-T, Jin Y-C, Liu G-H (2019b) The complete mitochondrial genome of the caecal fluke of poultry, *Postharmostomum commutatum*, as the first representative from the superfamily Brachylaimoidea. Front Genet 10:1037. https://doi.org/10.3389/fgene.2019.01037

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704. https://doi.org/10.1080/10635150390235520

http://mitos2.bioinf.uni-leipzig.de/index.py. Accessed 19 Jun 2020

http://molevol.cmima.csic.es/castresana/Gblocks_server.html. Accessed 30 Dec 2020

http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 23 Jan 2020

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310–2314. https://doi.org/10.1126/science.1065889

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Lamolle G, Fontenla S, Rijo G, Tort JF, Smircich P (2019) Compositional analysis of flatworm genomes shows strong codon usage biases across all classes. Front Genet 10:771. https://doi.org/10.3389/fgene.2019.00771

Le S, Gascuel O (2008) An improved general amino-acid replacement matrix. Mol Biol Evol 25:1307–1320. https://doi.org/10.1093/molbev/msn067

Le TH, Blair D, Agatsuma T, Humair PF, Campbell NJ, Iwagami M, Littlewood DT, Peacock B, Johnston DA, Bartley J, Rollinson D, Herniou EA, Zarlenga DS, McManus DP (2000) Phylogenies inferred from mitochondrial gene orders–a cautionary tale from the parasitic flatworms. Mol Biol Evol 17:1123–1125. https://doi.org/10.1093/oxfordjournals.molbev.a026393

Le TH, Nguyen NTB, Nguyen KT, Doan HTT, Dung DT, Blair D (2016) A complete mitochondrial genome from *Echinochasmus japonicus* supports the elevation of Echinochasminae Odhner, 1910 to family rank (Trematoda: Platyhelminthes). Infect Genet Evol 45:369–377. https://doi.org/10.1016/j.meegid.2016.09.024

Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Agatsuma T, Blair D (2019) The complete mitochondrial genome of *Paragonimus ohirai* (Paragonimidae: Trematoda: Platyhelminthes) and its comparison with *P. westermani* congeners and other trematodes. PeerJ 7:e7031. https://doi.org/10.7717/peerj.7031

Lee D, Choe S, Park H, Jeon H-K, Chai J-Y, Sohn W-M, Yong T-S, Min D-Y, Rim H-J, Eom KS (2013) Complete mitochondrial genome of *Haplorchis taichui* and comparative analysis with other trematodes. Korean J Parasitol 51:719–726. https://doi.org/10.3347/kjp.2013.51.6.719

Li Y, Ma XX, Lv QB, Hu Y, Qiu HY, Chang QC, Wang CR (2019) Characterization of the complete mitochondrial genome sequence of *Tracheophilus cymbius* (Digenea), the first representative from the

family Cyclocoeliidae. J Helminthol 94(e101):1–7. https://doi.org/10.1017/S0022149X19000932

Littlewood DTJ, Lockyer AE, Webster BL, Johnston DA, Le TH (2006) The complete mitochondrial genomes of *Shistosoma haematobium* and *Shistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. Mol Phylogenet Evol 39:452–467. https://doi.org/10.1016/j.ympev.2005.12.012

Liu G-H, Yan H-B, Otranto D, Wang X-Y, Zhao G-H, Jia W-Z, Zhu X-Q (2014a) *Dicrocoelium chiensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochonsrial and nuclear ribosomal DNA sequences. Mol Phylogenet Evol 97:325–331. https://doi.org/10.1016/j.ympev.2014.07.002

Liu G-H, Gasser RB, Young ND, Song H-Q, Ai L, Zhu X-Q (2014b) Complete mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola gigantica*, and their comparison with *F. hepatica*. Parasit Vectors 7:150. https://doi.org/10.1186/1756-3305-7-150

Liu Z-X, Zhang Y, Liu Y-T, Chang Q-C, Su X, Fu X, Yue D-M, Gao Y, Wang C-R (2016) Complete mitochondrial genome of *Echinostoma hortense* (Digenea: Echinostomatidae). Kor J Parasit 54:173–179. https://doi.org/10.3347/kjp.2016.54.2.173

Locke SA, Dam AV, Caffara M, Pinto HA, López-Hernández D, Blanar CA (2018) Validity of the Diplostomoidea and Diplostomida (Digenea, Platyhelminthes) upheld in phylogenomic analysis. Int J Parasitol 48:1043–1059. https://doi.org/10.1016/j.ijpara.2018.07.001

Ma J, He J-J, Zhou C-Y, Sun M-M, Cevallos W, Sugiyama H, Zhu X-Q, Calvopina M (2019) Characterization of the mitochondrial genome sequences of the liver fluke *Amphimerus* sp. (Trematoda: Opisthorchiidae) from Ecuador and phylogenetic implications. Acta Trop 195:90–96. https://doi.org/10.1016/j.actatropica.2019.04.025

Na L, Gao J-F, Liu G-H, Fu X, Su X, Yue D-M, Gao Y, Zhang Y, Wang C-R (2016) The complete mitochondrial genome of *Metorchis orientalis* (Trematoda: Opisthorchiidae): comparison with other closely related species and phylogenetic implications. Infect Genet Evol 39:45–50. https://doi.org/10.1016/j.meegid.2016.01.010

Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel S, Woyke T, Tesler G, Alekseyev M, Pevzner P (2013) Assembling genomes and mini-metagenomes from highly chimeric reads. Lect Notes Comput Sci 7821:158–170. https://doi.org/10.1007/978-3-642-37195-0

Oey H, Zakrzewski M, Gravermann K, Young ND, Korhonen PK, Gobert GN, Nawaratna S, Hasan S, Martínez DM, You H, Lavin M, Jones MK, Ragan MA, Stoye J, Oleaga A, Emery AM, Webster BL, Rollinson D, Gasser RB, McManus DP, Krause L (2019) Whole-genome sequence of the bovine blood fluke *Shistosoma bovis* supports interspecific hybridization with *S. haematobium*. PLoS Pathog 15: e1007513. https://doi.org/10.1371/journal.ppat.1007513

Olson PD, Cribb TH, Tkach VV, Bray RA, Littlewood DTJ (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int J Parasitol 33:733–755. https://doi.org/10.1016/S0020-7519(03)00049-3

Park J-K, Kin K-H, Kang S, Jeon HK, Kin J-H, Littlewood DTJ (2007) Characterization of the mitochondrial genome of *Diphyllobothrium latum* (Cestoda: Pseudophyllidea)–implications for the phylogeny of eucestodes. Parasitology 134:749–759. https://doi.org/10.1017/S003118200600206X

Pérez-Ponce de León G, Hernández-Mena DI (2019) Testing the higher-level phylogenetic classification of Digenea (Platyhelminthes, Trematoda) based on nuclear rDNA sequences before entering the age of the 'next-generation' Tree of Life. J Helminthol 93:260–276. https://doi.org/10.1017/S0022149X19000191

Qian L, Zhou P, Li W, Wang H, Miao T, Hu L (2018) Characterization of the complete mitochondrial genome of the lung fluke, *Paragonimus heterotremus*. Mit DNA Part B 3(2):560–561. https://doi.org/10.1080/23802359.2018.1462119

Saccone C, Gissi C, Reyes A, Larizza A, Sbisa E, Pesole G (2002) Mitochondrial DNA in Metazoa: degree of freedom in a frozed event. Gene. 286:3–12. https://doi.org/10.1016/S0378-1119(01)00807-1

Semyenova S, Chrisanfova G, Mozharovskaya L, Guliaev A, Ryskov A (2017) The complete mitochondrial genome of the causative agent of the human cercarial dermatitis, the visceral bird shistosome species *Trichobilharzia szidati* (Platyhelminthes: Trematoda: Shistosomatidae). Mit DNA Part B 2(2):469–470. https://doi.org/10.1080/23802359.2017.1347833

Shao R, Dowton M, Murrel A, Barker SC (2003) Rates of gene rearrangements and nucleotide substitution are correlated in the mitochondrial genomes of Insects. Mol Biol Evol 20:1612–1619. https://doi.org/10.1093/molbev/msg176

Shekhovtsov SV, Katochin AV, Kolchanov NA, Mordvinov VA (2010) The complete mitochondrial genomes of the liver flukes *Opisthorchis felineus* and *Clonorchis sinensis* (Trematoda). Parasitol Int 59:100–103. https://doi.org/10.1016/j.parint.2009.10.012

Suleman S, Khan MS, Heneberg P, Zhou CY, Muhammad N, Zhu X-Q, Ma J (2019a) Characterization of the complete mitochondrial genome of *Uvitellina* sp., representative of the family Cyclocoeliidae and phylogenetic implications. Parasitol Res 118:2203–2211. https://doi.org/10.1007/s00436-017-5669-6

Suleman S, Ma J, Khan MS, Tkach VV, Muhammad N, Zhang D, Zhu X-Q (2019b) Characterization of the complete mitochondrial genome of *Plagiorchis maculosus* (Digenea, Plagiorchiidae), representative of a taxonomically complex digenean family. Parasitol Int 71:99–105. https://doi.org/10.1016/j.parint.2019.04.001

TIBCO Software Inc. (2017). Statistica (program product for data analysis), version 13. http://statistica.io.

Wang T, Wang Y, Xu F, Li X, Qu R, Song L, Tang Y, Lin P (2018) Characterization of the complete mitochondrial genome of the lung fluke, *Paragonimus kellicotti*. Mit DNA Part B 3(2):715–716. https://doi.org/10.1080/23802359.2020.1724669

Webster BL, Rudilfová J, Horák P, Littlewood DTJ (2007) The complete mitochondrial genome of the bird shistosome *Trichobilharzia regent* (Platyhelminthes: Digenea), causative agent of cercarial dermatitis. J Parasitol 93:553–561. https://doi.org/10.1645/GE-1072R.1

Wu Y-A, Gao J-W, Cheng X-F, Xie M, Yuan X-P, Liu D, Song R (2020) Characterization and comparative analysis of the complete mitochondrial genome of *Azygia hwangtsiyui* Tsin, 1933 (Digenea), the first for a member of the family Azygiidae. ZooKeys 945:1–16. https://doi.org/10.3897/zookeys.945.49681

Yan H-B, Wang X-Y, Lou Z-Z, Li L, Blair D, Yin H, Cai J-Z, Dai X-L, Lei M-T, Zhu X-Q, Cai X-P, Jia W-Z (2013) The mitochondrial genome of *Paramphistomum cervi* (Digenea), the first representative for the family Paramphistomatidae. PLoS One 8:e71300. https://doi.org/10.1371/journal.pone.0071300

Yang X, Gasser RB, Koehler AV, Wang L, Zhu K, Chen L, Feng H, Hu M, Fang R (2015) Mitochondrial genome of *Hypoderaeum conoideum*–comparison with selected trematodes. Parasit Vectors 8:97. https://doi.org/10.1186/s13071-015-0720-x

Yang X, Wang L, Chen H, Feng H, Shen B, Hu M, Fang R (2016) The complete mitochondrial genome of *Gastrothylax crumenifer* (Gastrothylacidae, Trematoda) and comparative analyses with selected trematodes. Parasitol Res 115:2489–2497. https://doi.org/10.1007/s00436-016-5019-0

Zhukov EV (1971) New trematodes of marine and freshwater fishes from the basins of the Japanese and Yellow seas. Parazitologiya 5:155–161 (In Russian).