

Research Article

Cite this article: Solórzano-García B, Hernández-Mena DI, Choudhury A, Pérez-Ponce de León G (2024). The complete mitochondrial genome of 3 species of allocreadiids (Digenea, Allocreadiidae): characterization and phylogenetic position within the order Plagiorchiida. *Parasitology* **151**, 309–318. <https://doi.org/10.1017/S0031182024000064>

Received: 19 October 2023
Revised: 19 December 2023
Accepted: 9 January 2024
First published online: 15 January 2024



Keywords:

Allocreadium; *Creptotrematina*; DNA; mitogenome; Trematoda; *Wallinia*

Corresponding author:

Gerardo Pérez-Ponce de León;
Email: ppdleon@enesmerida.unam.mx

The complete mitochondrial genome of 3 species of allocreadiids (Digenea, Allocreadiidae): characterization and phylogenetic position within the order Plagiorchiida

Brenda Solórzano-García¹ , David I. Hernández-Mena¹ , Anindo Choudhury²  and Gerardo Pérez-Ponce de León¹ 

¹Escuela Nacional de Estudios Superiores unidad Mérida (ENES-Mérida, UNAM), Mérida, Yucatán, Mexico and

²St. Norbert College, De Pere, WI, USA

Abstract

Trematodes of the family Allocreadiidae are primarily found in the intestines of freshwater fishes around the world. The family includes 15 genera and c. 130 species. The last 2 decades have witnessed an increase in the genetic library of its species. Molecular data have been crucial for species delimitation and species description within Allocreadiidae and for understanding their evolutionary and biogeographical history and classification. Here, the mitogenomes of 3 species of allocreadiids were obtained using high throughput sequencing methods. Mitogenomes were compared with other members of the order Plagiorchiida to determine their molecular composition, gene rearrangement and phylogenetic interrelationships. The complete circular mitogenomes of *Allocreadium lobatum*, *Creptotrematina aguirrequeñoi* and *Wallinia mexicana* were 14 424, 13 769 and 13 924 bp long respectively, comprising 12 protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes and 2 non-coding regions. Gene arrangements were identical to other Xiphidiatan trematodes. Phylogenetic analyses using the mitogenomes revealed Allocreadiidae as a monophyletic group closely related to other members of the suborder Xiphidiata; *A. lobatum* was yielded as the sister taxon of *C. aguirrequeñoi* + *W. mexicana*. Our study increases the complete mitochondrial genome library of trematodes and strengthens our understanding of the phylogenetic relationships and classification of this parasite group.

Introduction

The family Allocreadiidae contains trematodes commonly found, as adults, in the intestine of freshwater fishes, and occasionally in snakes, salamanders and frogs (Caira and Bogéa, 2005). Adult allocreadiids typically possess an unspined tegument, a well-developed cirrus-sac, gonads in tandem and relatively large eggs. Allocreadiid metacercariae encyst in the haemocoel of aquatic arthropods; the cercariae are ophthalmoxiphidocercariae and develop in sphaeriid clams rather than in gastropods (Caira and Bogéa, 2005; Choudhury *et al.*, 2007). *Allocreadium* is the type-genus and the most speciose among allocreadiids. The somewhat complex taxonomic history of the family was discussed in detail by Caira and Bogéa (2005). These authors also revised the genus composition within the family and, after discussing the doubtful validity of 13 of them, concluded at the time that Allocreadiidae comprised 14 genera. The phylogenetic position of the family within the trematode order Plagiorchiida was first tested using molecular data by Curran *et al.* (2006) and Choudhury *et al.* (2007) using 28S rDNA sequences or a combined analysis of 28S and 18S rDNA, respectively. In both analyses, allocreadiids formed a highly supported monophyletic clade with species of Callodistomidae and Gorgoderidae, which, in turn appeared as sister taxa to a clade containing representative species of Encyclometridae, Brachycoeliidae, Dicrocoelidae and Orchipidae.

Traditional approaches based on morphology have not been useful to assess the interrelationships among members of the family, although a significant progress has been made unravelling the species delimitation, evolutionary history, classification and historical biogeography of this group of trematodes by using DNA sequences. For instance, 2 genera validated by Caira and Bogéa (2005), *Paracreptotrematina* and *Polylekithum*, were proved not to belong in Allocreadiidae (Curran *et al.*, 2006, 2011; Choudhury *et al.*, 2007); the genera *Margotrema* and *Wallinia* were validated as allocreadiids and not macroderoidids (Pérez-Ponce de León *et al.*, 2007); furthermore, 3 of the genera considered invalid by Caira and Bogéa (2005) have been resurrected, i.e., *Acrolichanus*, *Stephanophiala* and *Megalogonia* (Atopkin *et al.*, 2020; Vainutis *et al.*, 2021); finally, 4 new genera have been described (*Paracreptotrema*, *Paracreptotrematoides*, *Pseudoparacreptotrema* and *Mesoamericatrema*) (Choudhury *et al.*, 2006; Pérez-Ponce de León *et al.*, 2016, 2020; Mendoza-Garfias *et al.*, 2022). Our own account on the taxonomic composition of Allocreadiidae shows that the family currently contains

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

15 genera and approximately 130 species. This provides but a glimpse on the progress made on the classification scheme and species composition within the family Allocreadiidae after the revisionary work by Caira and Bogéa (2005).

The application of molecular tools has provided useful information that expands our knowledge about the species diversity of the family and their interrelationships (e.g., Faltýnková *et al.*, 2020; Pérez-Ponce de León *et al.*, 2020; Franceschini *et al.*, 2021; Petkevičiūtė *et al.*, 2023; Sokolov *et al.*, 2023; Vainutis *et al.*, 2023). Since the year 2000, c. 40 new species have been described following an integrative taxonomy approach, and molecular phylogenetic hypotheses have addressed aspects of the classification scheme of the group, providing more robust species delimitation criteria upon which biogeographical interpretations are made and nomenclatural changes proposed (see Atopkin *et al.*, 2020; Franceschini *et al.*, 2021; Vainutis *et al.*, 2021). In the same period, about 30 papers generated information to enhance the genetic library of ribosomal and mitochondrial genes for allocreadiid species, particularly the 28S rDNA and the cytochrome oxidase subunit 1, *cox1*.

Complete mitochondrial genomes (mt genomes) obtained through high throughput sequencing methods have been used in the last 20 years for reconstructing the phylogenetic history of trematodes and to test the higher-level classification scheme. The use of mt genomes is based on the premise that the number of sequenced base pairs, gene order and genome arrangement may provide evidence of shared ancestry (see Littlewood *et al.*, 2006), although their power to resolve deep levels of phylogenetic trees remains controversial (Pérez-Ponce de León and Hernández-Mena, 2019). Still, mt genomes have not been generated for species in the family Allocreadiidae to test their interrelationships with other trematodes. Thus, the main objective of this paper was to sequence the mt genome of 3 species of allocreadiids, namely *Allocreadium lobatum*, *Creptotrematina aguirrepequeno* and *Wallinia mexicana*, parasites of freshwater fishes, to determine the gene content, arrangement, and composition, and to assess the phylogenetic relationships with other trematode species belonging in the order Plagiorchiida.

Materials and methods

Sample collection, morphological identification, and DNA extraction

Specimens of allocreadiids were sampled from the intestines of their host as follows: Specimens of *C. aguirrepequeno*, *W. mexicana*, and *A. lobatum* were isolated from *Astyanax mexicanus* (Jalpan, Queretaro, Mexico), *Astyanax aeneus* (Matías Romero, Oaxaca, Mexico) and *Luxilus cornutus/Semotilus atromaculatus* (West Twin River, Wisconsin, USA), respectively. Some specimens were killed and relaxed with nearly boiling tap water and stored in vials with 96% ethanol for further processing (staining and mounting) and morphological identification. Trematodes were identified to species level following previous bibliographic accounts. Some other specimens were thoroughly washed in physiological saline solution and fixed in 100% ethanol for molecular analysis. To confirm the taxonomic identity, DNA was extracted from 1 specimen of each morphologically identified species and a fragment of the 28S ribosomal RNA gene was amplified as described in previous studies, using the primers 502 (5' - CAAGTACCGTGAGGGAAAGTTGC- 3') and 536 (5' - CAGCTATCCTGAGGGAAAC-3') (García-Varela and Nadler, 2005). PCR products were sequenced at the LANBIO-IBUNAM, Mexico City and a BLAST-N search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was performed.

For mt genome sequencing, genomic DNA was extracted from a pool of 5–9 individuals of each trematode species using QIAamp

Fast DNA Tissue kit (Qiagen) according to the manufacturer's protocol. The dosage of the obtained DNA was determined in a microvolume spectrophotometer (NanoDrop_ND-1000).

High throughput sequencing and assembly

Extracted genomic DNA of each trematode species was shotgun sequenced on an Illumina HiSeq 4000 with a 100x depth, and 150-bp paired-end libraries were built with Nextera adaptors at Azenta GENEWIZ (NJ, USA). Bioinformatic analysis and de-novo genome assembly of the obtained reads were also conducted by Azenta. SOAPDeNovo2 (Luo *et al.*, 2012) was used on each of the samples with a minimum contig length of 1000 bp. EMBOSS tools GetOrf was then used to find the open reading frames (ORF) within the *de novo* assembled genome. The protein sequences from ORF were then annotated using Diamond BLASTp. To extract the mt genome, scaffolds were mapped onto a barcode sequence of the *cox1* gene from each allocreadiid species in this study using the software Geneious Prime v.11 with the highest sensitivity and up to 5 iterations. This initial assembly was then used to seed iterative extensions until even and thorough coverage was obtained. The final assembly was annotated with MITOS Webserver (refseq = 63, translation table = 5) (Bernt *et al.*, 2013). The ORF of each protein gene was verified using Geneious, employing the invertebrate mitochondrial genetic code. Putative tRNA genes and their secondary structure were identified using ARWEN (Laslett and Canbäck, 2008) and MITOS. Circular mitochondrial genome maps were drawn using GenomeVx tool (Conant and Wolfe, 2008).

Comparative analysis among allocreadiid mt genomes.

Contents of A + T and G + C were calculated in Geneious. AT-skew and GC-skew values were calculated using the equations $AT\text{-skew} = (A - T)/(A + T)$ and $GC\text{-skew} = (G - C)/(G + C)$ in each of the 12 protein-coding genes (PCGs) and the non-coding regions (NCRs). The PCG sequences were translated into their corresponding amino acid sequences using MEGA X (Kumar *et al.*, 2018) and the invertebrate mitochondrial genetic code. Amino acid composition and relative synonymous codon usage (RSCU) was also estimated in MEGA X. Mutation rate (non-synonymous/synonymous, dN/dS) ratios among the 12 PCGs of the 3 newly sequenced allocreadiid mt genomes were calculated using DnaSP v.6.12 (Rozas *et al.*, 2017). To estimate nucleotide diversity (π) across genes, a sliding window analysis was implemented in DnaSP using a window size of 300 bp and a step size of 30 bp, for the 12 PCGs plus the 2 rRNAs (*rrnS* and *rrnL*); values of nucleotide diversity were then plotted according to the midpoint position of each window. Nucleotide and amino acid differences were calculated using MEGA X.

Phylogenetic analyses

The newly sequenced mt genomes were used along with those of other plagiorchiid trematodes downloaded from the GenBank database to reconstruct phylogenetic trees. Analyses included species allocated in the suborders Echinostomata, Opisthorchiata and Xiphidiata. Based on previous molecular phylogenetic analyses (Pérez-Ponce de León and Hernández-Mena, 2019), *Fasciola hepatica* (AF216697) was used as an outgroup. A concatenated matrix of the 12 PCGs was built in Geneious. Previously, each gene was independently aligned in ClustalW through the GenomeNet web interface (www.genome.jp). Phylogenetic relationships were inferred using 2 datasets: the first one with the nucleotide alignment of 12 PCGs plus the 2 ribosomal genes (*rrnS* and *rrnL*), and the second with the amino acid alignment.

Phylogenetic trees were constructed using the maximum likelihood (ML) method in RaxML (Stamatakis, 2014). Empirical substitution models GTR + CAT and PROTGAMMA were employed for nucleotide and amino acid data sets, respectively; bootstrap support values for nodes were estimated with 1000 replicates.

Results and discussion

General features of allocreadiid mitogenomes

The complete circular mt genomes of *A. lobatum*, *C. aguirrepequeno* and *W. mexicana* were 14,427, 13,798 and 13 926 pb long, respectively (GenBank accession numbers: OR987847, OR987848, OR987849, respectively). These mt genomes were composed of 12 PCGs (*cox1* – 3, *nad 1*–6, *cytb* and *atp6*; *atp8* gene is lacking), 2 rRNA genes (*rrnS* and *rrnL*), 22tRNAs (1 for each amino acid and 2 each for leucine and serine), and 2 non-coding regions (NCR) (Fig. 1 and Table 1). All genes were transcribed in the same direction. Total length and gene order of all 3 newly sequenced allocreadiid mt genomes corresponds to the typical arrangement observed in other species allocated in the order Plagiorchiida (Liu *et al.*, 2014; Briscoe *et al.*, 2016; Chang *et al.*, 2016; Qian *et al.*, 2018; Le *et al.*, 2019; Suleman *et al.*, 2019, 2020, 2021; Guo *et al.*, 2022; Gao *et al.*, 2023). The order of the 12PCGs seems to be highly conserved among trematodes, with the location of tRNAs as variable among species (Suleman *et al.*, 2019; Gao *et al.*, 2023).

Intergenic spacers ranged from 1–34 bp in *A. lobatum* and *C. aguirrepequeno*, and from 1–63 bp in *W. mexicana*. A 40 bp overlap between *nad4L* and *nad4* was present in the 3 mt genomes. This is consistent with some species of xiphidiatan trematodes such as *Paragonimus heterotremus*, *P. ohirai*, *Prosthogonimus* spp., *Tamerlania zarudnyi* and *Plagiorchis maculosus* (Qian *et al.*, 2018; Le *et al.*, 2019; Suleman *et al.*, 2019, 2021; Guo *et al.*, 2022). However, other xiphidiatans (e.g., *Dicrocoelium dendriticum*, *D. chinensis* and *Paragonimus westermani*) show an overlap of just 1– 14 bp (Biswal *et al.*, 2014; Liu *et al.*, 2014).

The *rrnL* gene was located between *trnT* and *trnC*, while the *rrnS* gene was located between *trnC* and *cox2* in the 3 allocreadiid species. The length of the 22 tRNA genes was 1403 bp in *A. lobatum*, 1404 bp in *C. aguirrepequeno*, and 1,427 in *W. mexicana*,

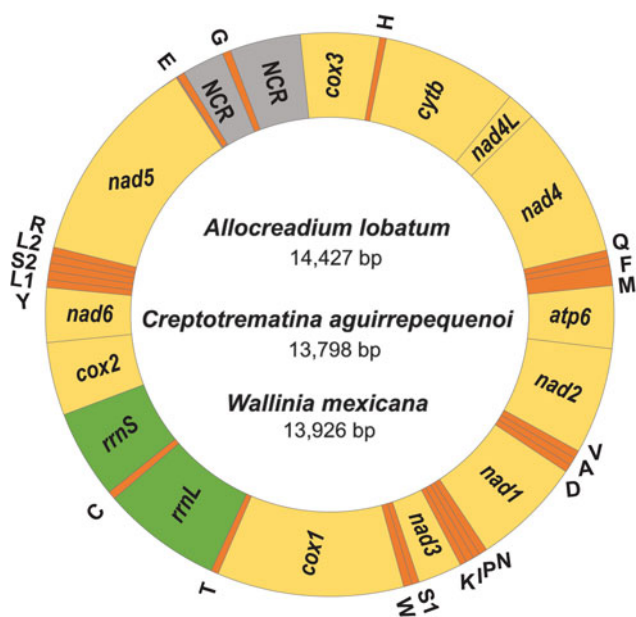


Figure 1. Gen map of the 3 allocreadiid mitochondrial genomes, *Allocreadium lobatum*, *Creptotrematina aguirrepequeno*, *Wallinia mexicana*. NCR, non-coding regions.

Table 1. Comparison of the mitochondrial genome arrangement of *Allocreadium lobatum* (Al), *Creptotrematina aguirrepequeno* (Cp.) and *Wallinia mexicana* (Wa)

Genes	Star position	End position	Length	Initiation codon	Stop codon	Anticodon	NN similarity	AA similarity
<i>cox3</i>	1/1/1	645/645/645	645/645/645	ATG/ATG/ATG	TAA/TAG/TAA	GTG	82.4	81.2
<i>trnH</i>	661/663/663	717/716/716	57/54/54	Al/Cp/Wa	Al/Cp/Wa			
<i>cytb</i>	724/725/725	1833/1834/1834	1110/1110/1110	ATG/ATG/ATG	TAG/TAG/TAG		90.1	92.8
<i>nad4L</i>	1867/1868/1868	2103/2104/2104	237/237/237	ATG/ATG/GTG	TAG/TAG/TAG		82.1	79.7
<i>nad4</i>	2064/2065/2065	3320/3321/3345	1257/1257/1281	ATG/ATG/GTG	TAA/TAA/TAA		84.2	80.9
<i>trnQ</i>	3321/3323/3346	3381/3383/3406	61/61/61			TTG		
<i>trnF</i>	3383/3385/3408	3445/3447/3470	63/63/63			GAA		
<i>trnM</i>	3346/3448/3471	3512/3512/3535	67/72/65			CAT		
<i>atp6</i>	3513/3513/3536	4028/4028/4051	516/516/516	ATG/ATG/ATG	TAG/TAA/TAG		75.6	80.8
<i>nad2</i>	4035/4035/4052	4913/4901/4903	879/867/852	ATG/ATG/ATG	TAG/TAG/TAA		79.2	75.3
<i>trnV</i>	4914/4914/4918	4978/4979/4980	65/66/63			TAC		
<i>trnA</i>	4988/4982/4983	5051/5055/5044	64/63/62			TGC		
<i>trnD</i>	5056/5052/5053	5120/5116/5117	65/65/65			GTC		

(Continued)

Table 1. (Continued.)

Genes	Star position Al/Cp/Wa	End position Al/Cp/Wa	Length Al/Cp/Wa	Initiation codon Al/Cp/Wa	Stop codon Al/Cp/Wa	Anticodon	NN similarity	AA similarity
<i>nad1</i>	5121/5117/5118	6026/6022/6023	906/906/906	ATG/ATG/ATG	TAG/TAG/TAG		91.5	90.1
<i>trnN</i>	6036/6032/6033	6102/6098/6099	67/67/67			GTT		
<i>trnP</i>	6103/6099/6100	6157/6164/6166	56/66/66			TGG		
<i>trnI</i>	6166/6165/6167	6229/6228/6230	64/64/64			GAT		
<i>trnK</i>	6237/6232/6234	6300/6299/6299	64/68/66			CTT		
<i>nad3</i>	6304/6301/6303	6663/6660/6662	360/360/360	ATG/ATG/ATG	TAG/TAA/TAG		75.8	78.5
<i>trnS1</i>	6669/6666/6666	6729/6724/6724	61/59/59			GCT		
<i>trnW</i>	6731/6726/6726	6791/6791/6791	61/66/66			TCA		
<i>cox1</i>	6810/6798/6798	8336/8336/8336	1527/1539/1539	TGA/ATG/ATG	TAG/TAG/TAG		79.7	85.1
<i>trnT</i>	8337/8337/8337	8402/8402/8402	66/66/66			TGT		
<i>rrnL</i>	8404/8403/8403	9377/9369/9383	974/966/980					
<i>trnC</i>	9381/9386/9387	9446/9443/9444	66/58/57			GCA		
<i>rrnS</i>	9450/9448/9448	10178/10174/10180	729/727/755					
<i>cox2</i>	10202/10197/10203	10801/10796/10802	600/600/600	ATG/ATG/ATG	TAA/TAA/TAG		81.8	80.7
<i>nad6</i>	10803/10798/10804	11252/11247/11253	450/450/453	ATG/GTG/GTG	TAA/TAG/TAA		83.0	85.0
<i>trnY</i>	11256/11253/11257	11319/11315/11320	65/63/64			GTA		
<i>trnL1</i>	11325/11321/11342	11388/11383/11412	64/63/71			TAG		
<i>trnS2</i>	11396/11381/11413	11453/11445/11478	58/65/66			TGA		
<i>trnL2</i>	11463/11450/11480	11528/11515/11547	66/64/68			TAA		
<i>trnR</i>	11529/11519/11549	11592/11577/11614	64/59/66			TCG		
<i>nad5</i>	11625/11586/11677	13199/13160/13251	1596/1575/1599	ATT/ATT/ATT	TAA/TAA/TAA		84.2	87.1
<i>trnE</i>	13225/13187/13276	13294/13250/13347	70/64/72			TTC		
NCR	13295/13251/13348	13783/13427/13514	489/177/167					
<i>trnG</i>	13784/13429/13515	13852/13497/13583	69/70/69			TCC		
NCR	13853/13498/13584	14427/13798/13926	572/301/397					

Similarity values are expressed in percentage. NN, nucleotides; AA, amino acids.

whereas individual tRNA gene lengths ranged from 54 to 72 bp (Table 1). The secondary structures predicted for the tRNA genes were similar among the 3 allocreadiids. Twenty of the tRNAs in *C. aguirrepequeno* and *W. mexicana*, along with 18 tRNAs in *A. lobatum* can be folded into the conventional cloverleaf secondary structure. The D-arm is absent in the tRNA for Serine 1 (*trnS1*) in the 3 species, in the *trnS2* only in *A. lobatum*, and in the tRNA for Cysteine in *C. aguirrepequeno* and *W. mexicana* (Fig S1). The tRNAs for Proline (*trnP*) and Lysine (*trnK*) are lacking the T-arm in *A. lobatum* (Fig S1). The lack of D-arm in the Serine 1 tRNAs is consistent with previous reports on mt genomes of other trematodes (Chang *et al.*, 2016; Le *et al.*, 2019; Suleman *et al.*, 2021; Guo *et al.*, 2022). Likewise, tRNAs with non-canonical structures have been observed in different organisms including mammals, nematodes and arthropods, where either T-arm, D-arm or both could be missing (Krahn *et al.*, 2020).

The 3 species contain 2 NCRs between *trnE* and *cox3*, separated by *trnG* (Fig. 1). The shorter NCR was found between *trnE* and *trnG* and varies in size from almost 500 bp in *A. lobatum* to 167 bp in *W. mexicana*; the longer NCR was located between *trnG* and *cox3* and ranged from 300 bp in *C. aguirrepequeno* to 580 bp in *A. lobatum*. The pattern of 2 NCRs, one shorter than the other, has also been observed in other plagiiorchiida including *Lyperosomum longicauda*, *T. zarudnyi*, *Paragonimus ohirai*; however, these are located between *trnG* and *cox3*, separated by *trnE* (Le *et al.*, 2019; Suleman *et al.*, 2020, 2021). This arrangement differs in *Prosthogonimus* spp. and *Plagiorchis maculosus* since both species present only 1 long NCR (Suleman *et al.*, 2019; Guo *et al.*, 2022), indicating that the length, arrangement and composition of the NCRs are not ubiquitous traits in Plagiiorchiida mt genomes.

Comparative analysis of allocreadiid mitogenomes

Gene order of the mt genome was the same among the 3 species of allocreadiids (Fig. 1); likewise, the size of each gene was similar. The rank order of the 12 PCGs by length was: *nad5* > *cox1* > *nad4* > *cytb* > *nad1* > *nad2* > *cox3* > *cox2* > *atp6* > *nad6* > *nad3* > *nad4L* (Table 1). Nucleotide composition of the 3 mt genomes was biased towards A and T, with an overall AT content of 63.5% in *A. lobatum*, 64.7% in *C. aguirrepequeno*, and 65.1%

in *W. mexicana* (Fig. 2), as observed in most trematodes (Gao *et al.*, 2023). A + T contents of *cytb*, *nad4L*, *nad1*, *nad3*, *nad6* and *cox1* were the lowest in *A. lobatum*, but the highest in *W. mexicana*; the opposite was true for *cox3* and *cox2*, with the highest A + T content in *A. lobatum*. The genes *atp6* and *nad2* showed highest A + T content in *C. aguirrepequeno* and the lowest values in *W. mexicana* (Fig. 2). The AT-skew values in the *A. lobatum* mitogenome ranged from -0.16 (*cox2*) to -0.50 (*nad6*), the GC-skew values ranged from 0.23 (*cox1* and *cox2*) to 0.67 (*nad3*). In *C. aguirrepequeno* the AT-skew values ranged from -0.25 (*cox2*) to -0.50 (*nad2*), the GC-skew values ranged from 0.20 (*nad5*) to 0.31 (*nad3*, *nad4L*, *nad6*). Finally, for *W. mexicana* AT-skew values ranged from -0.30 (*cox2*) to -0.48 (*nad3*), and GC-skew values ranged from 0.15 (*atp6*) to 0.46 (*nad3*) (Fig. 2).

The most common start codon among the 12 PCGs of the 3 allocreadiid mt genomes was ATG, with a frequency of 10/12 in *A. lobatum* and *C. aguirrepequeno*, and 8/12 in *W. mexicana*. The TAG was the predominant stop codon in *A. lobatum* (8/12), as well as in *C. aguirrepequeno* and *W. mexicana* (7/12) (Table 1). There were 3341, 3342 and 3,328 amino acids in 12 PCGs of *A. lobatum*, *C. aguirrepequeno* and *W. mexicana*, respectively; the RSCU is shown in Fig. 3. In the 3 mt genomes, the most frequent amino acid was Leu (16.2–16.8%), followed by Phe (11.5–11.6%), Ser (11.1–11.4) and Val (9.2–9.6%). The least frequent amino acid was Gln (0.8–0.9%).

Overall nucleotide similarity for the 12 PCGs among the 3 allocreadiid mt genomes was 75.5%, whereas amino acids showed 61.6% overall similarity. Genes with high nucleotide diversity among allocreadiids were *atp6* > *nad3* > *nad2* > *rrnS* > *cox1*, which is consistent with previous studies using sliding window analysis in members of Plagiiorchiida (Table 1). To determine the highly conserved and variable mitochondrial genes among the 3 allocreadiid trematodes, a sliding window analysis was conducted by the concatenated nucleotide sequence of 12 PCGs plus the 2 *rrn* genes. Nucleotide diversity (π) among the 3 mt genomes ranged from 0.004 to 0.29; *atp6* ($\pi = 0.241$) was the most variable gene, while *cytb* ($\pi = 0.10$) and *nad1* ($\pi = 0.09$), showed low sequence variation (Fig. 4). *Cox 1*, *rrnS* and *rrnL* show intermediate molecular variation ($\pi = 0.19$, 0.21 and 0.15, respectively).

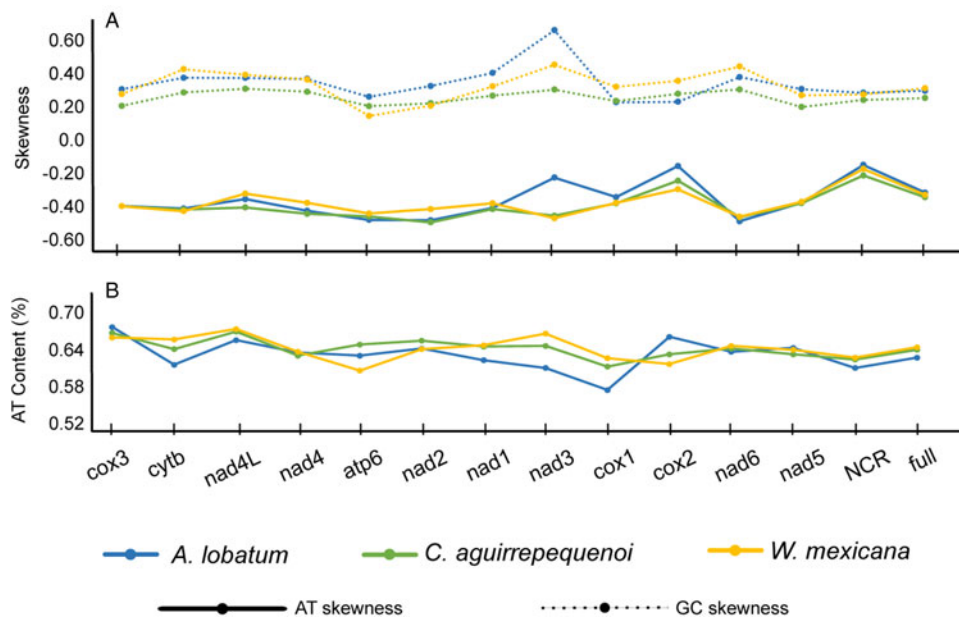


Figure 2. The nucleotide skewness (A) and AT content (B) of 3 species of Allocreadiids mitochondrial genomes (*Allocreadium lobatum*, *Creptotrematina aguirrepequeno* *Wallinia mexicana*).

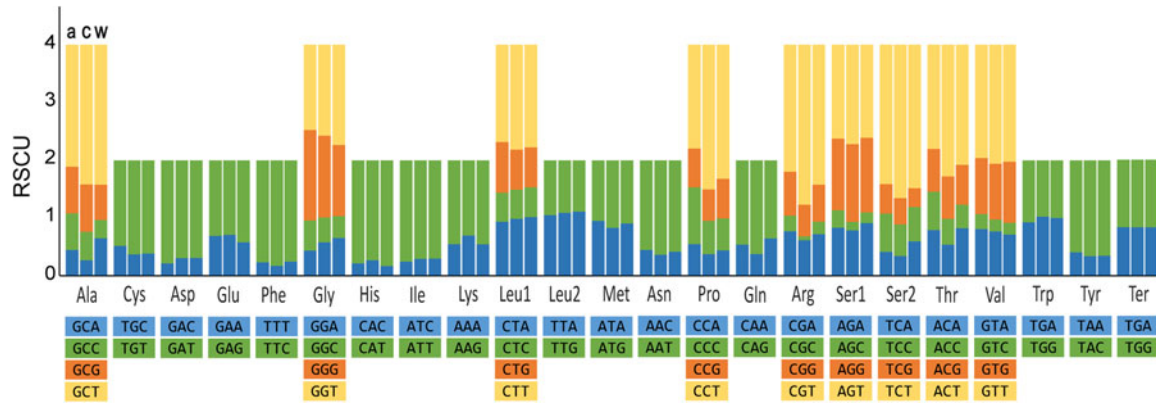


Figure 3. Relative synonymous codon usage (RSCU) of 12 protein coding genes of the 3 allocreadiid mt genomes: a) *Allocreadium lobatum*, c) *Creptotrematina aguirrepequeno*; w) *Wallinia mexicana*.

Analysis of mutation rates among the 12 PCGs of the 3 allocreadiid mt genomes showed that all PCGs were under purifying selection ($dN/dS < 1$) (Fig. 4); this coincides with previous estimations for other trematode species (Suleman et al., 2020; An et al., 2022). The highest non-synonymous substitution rate (dN) was observed in *nad4L* and *nad1*, while *cox1* gene showed the lowest

dN/dS value, followed by *atp6*, indicating a higher synonymous variation in these genes among allocreadiids.

These results plus the distribution of nucleotide diversity among genes indicate *cox1* and *atp6* as useful genetic markers for genetic lineage determination and species identification studies in the family Allocreadiidae. It has been observed in other

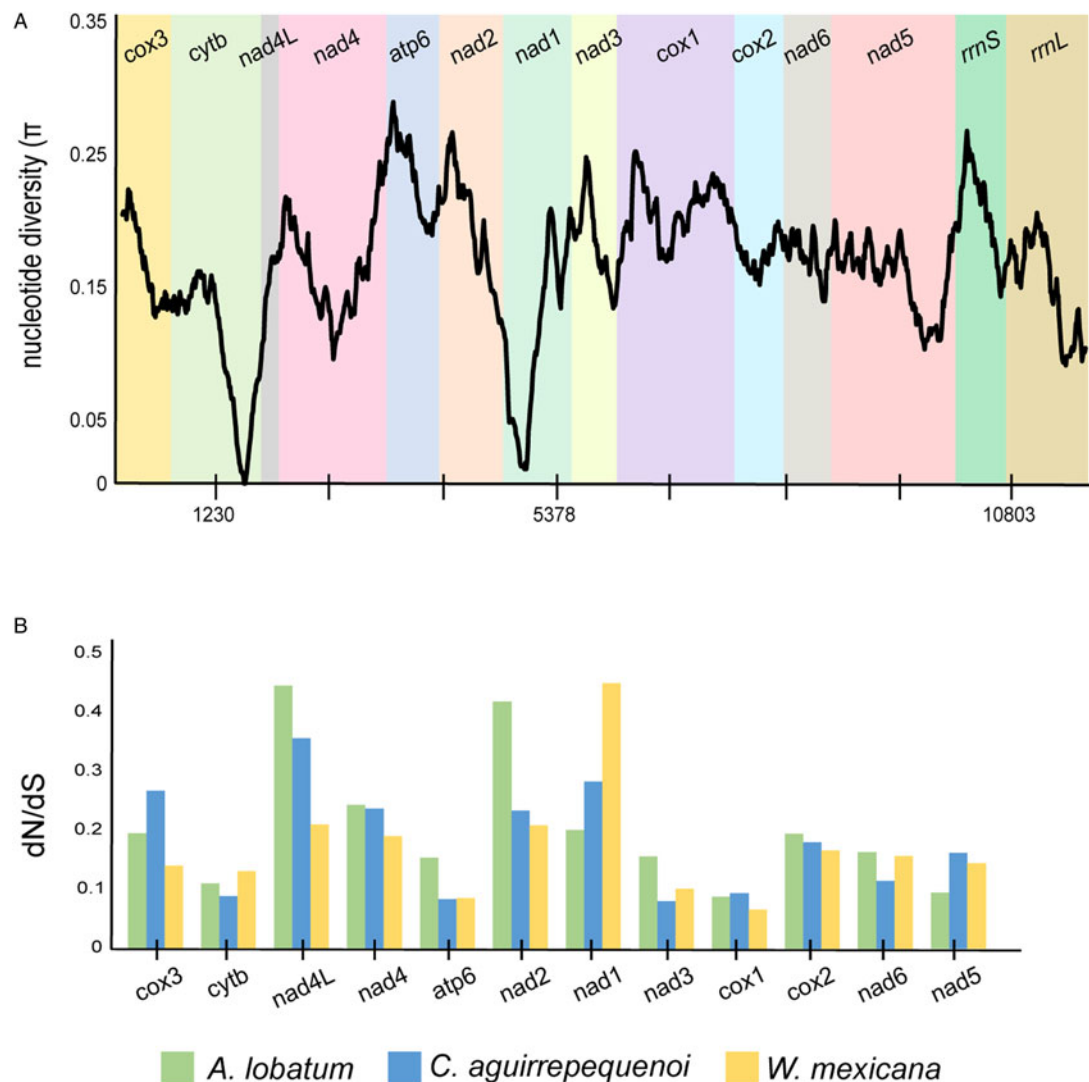


Figure 4. Nucleotide variation across *Alocreadium lobatum*, *Creptotrematina aguirrepequeno* and *Wallinia mexicana* mt genomes. (A) Sliding window analysis of the 12 PCGs plus *rns* and *rnl*. The black line represents nucleotide variation in a window of 300 bp (step size = 10 bp, with the value inserted at its mid-point). (B) Ratios of non-synonymous to synonymous (dN/dS) substitution rates calculated from individual protein-coding genes of 3 newly allocreadiid mt genomes.

trematode mt genomes that *atp6* and *nad2* genes evolve faster than *cox1*, suggesting that these genes may also serve as genetic markers to capture genetic variation among closely related trematode species or even populations (Suleman *et al.*, 2020).

Phylogenetic relationships of Allocreadiidae

The ML phylogenetic trees using either nucleotides or amino acids recovered Allocreadiidae as a highly supported monophyletic group (Figs 5, 6), with *C. aguirrepequeno* clustered together with *W. mexicana*, and these 2 with *A. lobatum*. In both analyses, allocreadiids appeared as sister taxa of other species included in the suborder Xiphidiata; this is consistent with results obtained through phylogenetic analyses based on 18S and 28S rRNA genes (Pérez-Ponce de León and Hernández-Mena, 2019), in which Allocreadiidae was recovered as a monophyletic group within the suborder Xiphidiata. Figure 5 depicts the phylogenetic tree inferred from the concatenated nucleotide sequences of 12 PCGs and 2 *rrn* for the same dataset; here the family Allocreadiidae was recovered as the sister group of a clade containing Haematoloechidae, Glythelminthidae, Plagiorchiidae and Orientocreadiidae, plus a clade comprising Eucotylidae, Prosthogonimidae and Dicrocoeliidae. In the phylogenetic tree inferred with amino acids (Fig. 6), Allocreadiidae appeared as an early divergent clade with respect to a clade containing paragonimids plus brachycladiids (both members of Xiphidiata in the current classification, see Pérez-Ponce de León and Hernández-Mena, 2019), and all of them as sister taxa of 2

families of Opisthorchiata (Fig. 6), although these relationships are poorly supported by bootstrap values.

In the current study, phylogenetic analyses of mt genomes of 28 species of trematodes allocated in 12 families using either nucleotides or amino acids, yielded contrasting results about the monophyly of Xiphidiata, since the nucleotide tree yielded the suborder as monophyletic, whereas the amino acids tree yielded Xiphidiata paraphyletic. Our results are in contrast with Suleman *et al.* (2021) who found consistent results of the paraphyly of Xiphidiata using either nucleotide or amino acid sequences; Suleman *et al.* (2021) used the mt genomes of 30 species of trematodes allocated in 14 families. Likewise, Pérez-Ponce de León and Hernández-Mena (2019) also found Xiphidiata as paraphyletic, and the suborder Haploporata was proposed for keeping the former as monophyletic. Other mt genome phylogenetic analyses using amino acids for building the trees have consistently shown the paraphyly of Xiphidiata (Guo *et al.*, 2022; Gao *et al.*, 2023). In the most recently published study, Atopkin *et al.* (2023) used a comprehensive dataset of 61 mitogenomes of trematodes and corroborated Xiphidiata as paraphyletic, although their dataset was not analysed using nucleotides. Apparently, interrelationships between paragonimids and dicrocoeliids as members of the superfamily Gorgoderoidea is problematic even though both have been considered as members of Xiphidiata. Interestingly, Suleman *et al.* (2021) and Atopkin *et al.* (2023) found that Paragonimidae plus Brachycladiidae are nested with Opisthorchiidae and Heterophyidae. The amino acid phylogenetic tree in our study uncovered the same

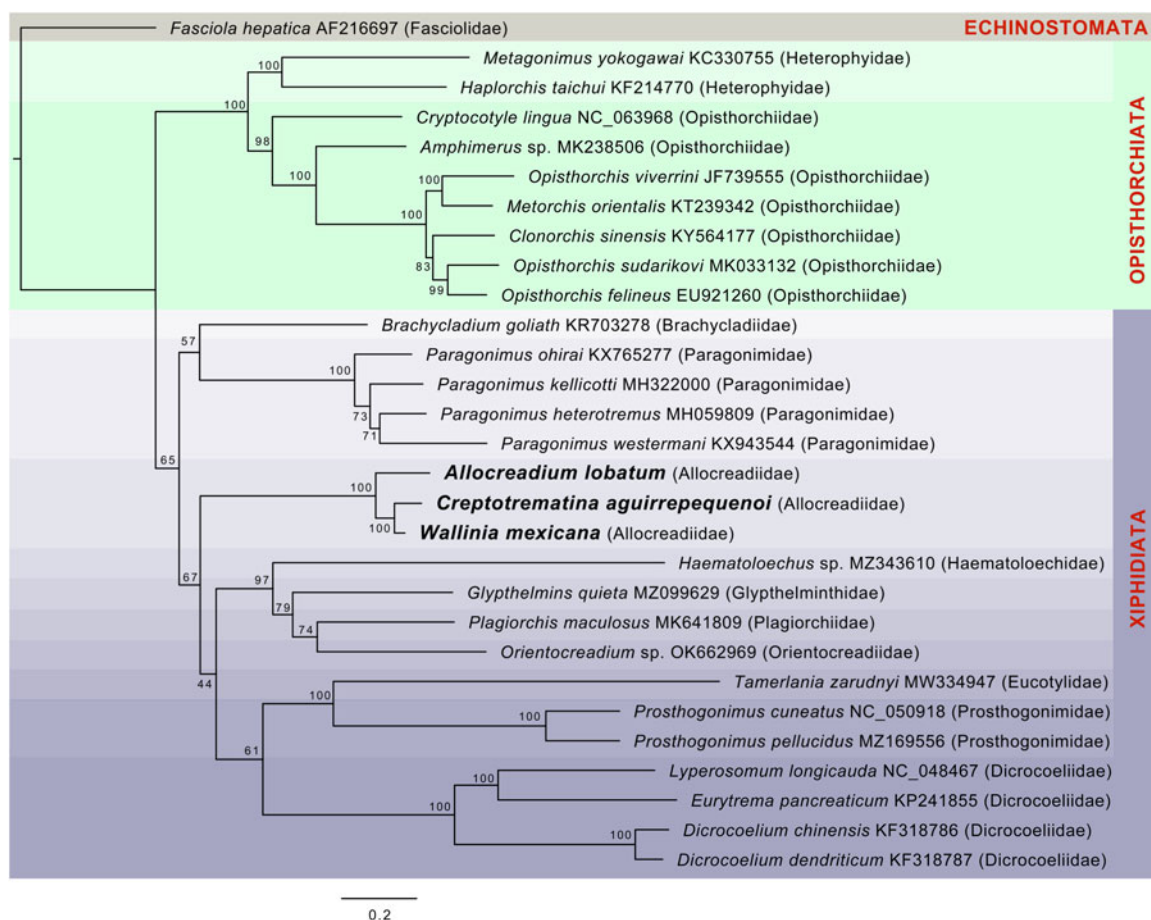


Figure 5. Maximum likelihood phylogenetic tree obtained with the nucleotide dataset of 12 protein-coding genes and 2 regions of the ribosomal RNA. Species in the suborder Xiphidiata are framed in purple shade. Species in Opisthorchiata are framed in green shade. The name of the species is followed by the mt genome accession number in the Genbank dataset, and in parentheses the family to which it belongs. The species sequenced in this study are highlighted in bold. Numbers near the nodes are the Bootstrap support values.

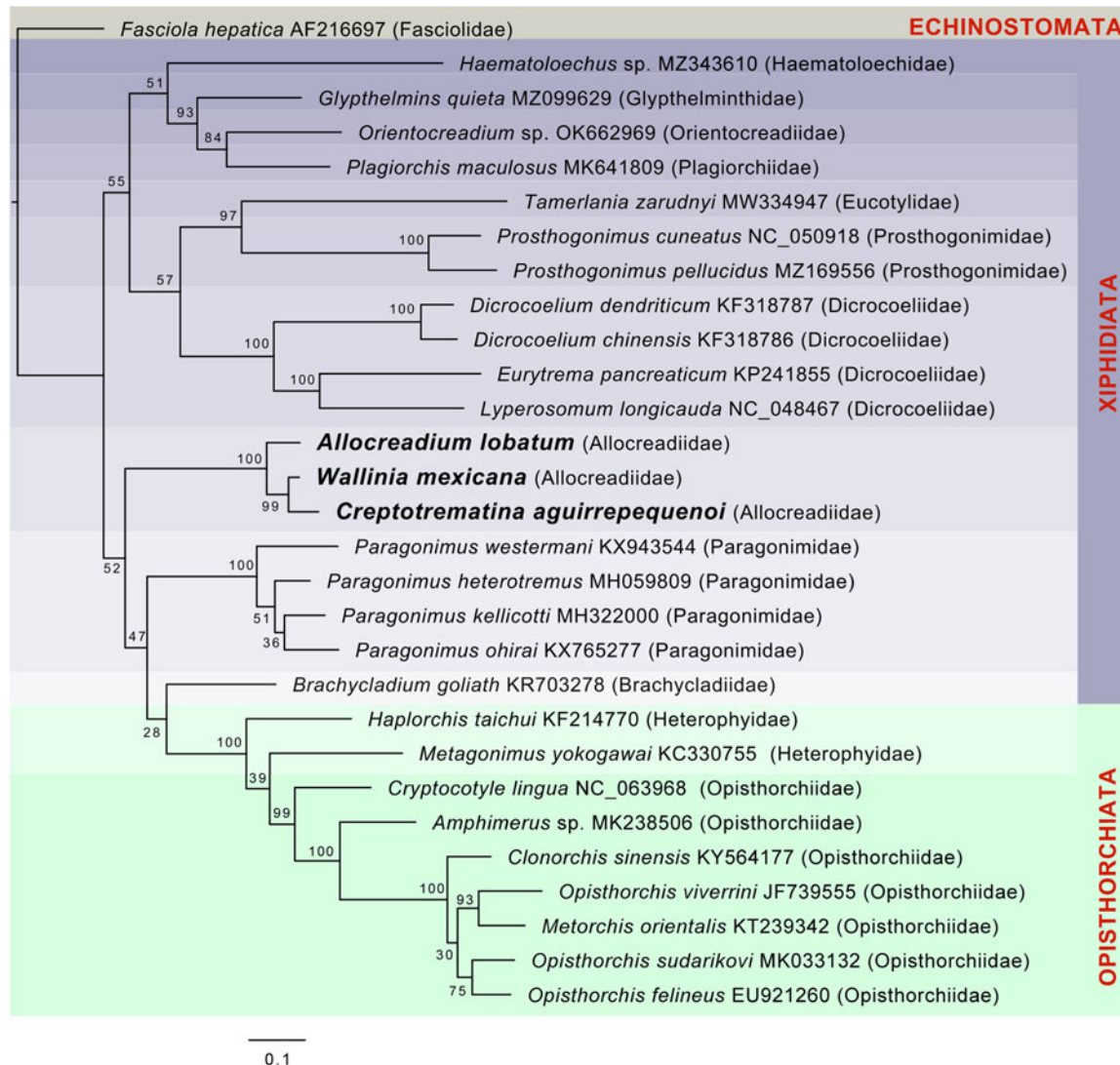


Figure 6. Maximum likelihood phylogenetic tree obtained the amino acids translated of 12 PCG's. Species in the suborder Xiphidiata are framed in purple shade. Species in Opisthorchiata are framed in green shade. The name of the species is followed by the mt genome accession number in the Genbank dataset, and in parentheses the family to which it belongs. The species sequenced in this study are highlighted in bold. Numbers near the nodes are the Bootstrap support values.

relationships, although Allocreadiidae appeared as the sister clade to these 4 families. The number and selection of taxa representative of different families, and the molecular data employed (nucleotides or amino acids), seem to exert an effect when assessing interrelationships of trematodes using mt genomes. As shown in our study, a shorter number of terminals, and a different selection of taxa may yield contrasting results.

There is no doubt that the discordant phylogenies of mt genomes obtained when using nucleotide or amino acid data, or through different inference methods, and the constant changes in the interrelationships of higher taxa when adding newly sequenced mt genomes is posing new challenges to trematode classification. In our opinion, 3 aspects must be considered to fully assess the value of mt genomes in reconstructing the phylogenetic relationships, and in elucidating higher-level interrelationships of this group of parasitic platyhelminthes. These aspects include the number of taxa sampled, outgroup selection, and use of nucleotide or amino acids. These remain key elements that need to be assessed to produce the most natural classification scheme of digenean trematodes using modern high throughput sequencing technologies.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0031182024000064>.

Data availability statement. Mitogenome sequence data is available on the NCBI GenBank database.

Acknowledgements. We sincerely thank Laura Márquez and Nelly López from the Laboratorio Nacional de Biodiversidad (LaNabio, Biology Institute, UNAM) for their help with DNA sequencing processing.

Authors' contributions. GPPL, BSG and DHM designed the study; GPPL, DHM and AC conducted field work, and processed the specimens for morphological identification. BSG conducted the experiments to obtain the mitogenomes. BSG and DHM analysed the molecular data and submitted the mitogenome sequences to GenBank. All authors discussed the results. GPPL and BSG drafted the manuscript which was reviewed and greatly improved by a DHM and AC.

Financial support. This research was supported by a grant from the Consejo Nacional de Ciencia y Tecnología (CONACYT A1-S-21694), and partially funded by the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica to GPPL (PAPIIT-UNAM IN212621).

Competing interests. The authors declare they have no competing financial interests or personal relationships that could have influenced the work reported in this manuscript.

Ethical standards. All applicable institutional, national and international guidelines for the care and use of animals were followed.

References

- An Q, Qiu YY, Lou Y, Jiang Y, Qiu HY, Zhang ZH, Li B, Zhang AH, Wei W, Chen YY, Gao JF and Wang CR (2022) Characterization of the complete mitochondrial genomes of *Diplodiscus japonicus* and *Diplodiscus mehari* (Trematoda: Diplodiscidae): comparison with the members of the superfamily Paramphistomoidea and phylogenetic implication. *International Journal for Parasitology – Parasites and Wildlife* **19**, 9–17.
- Atopkin DM, Sokolov SG, Vainutis KS, Voropaeva EL, Shedko MB and Choudhury A (2020) Amended diagnosis, validity and relationships of the genus *Acrolichanus* Ward, 1917 (Digenea: Allocreadiidae) based on the 28S rRNA gene, and observations on its lineage diversity. *Systematic Parasitology* **97**, 143–156.
- Atopkin DM, Semenchenko AA, Solodovnik DA and Ivashko YI (2023) A report on the complete mitochondrial genome of the trematode *Azygia robusta* Odhner, 1911, its new definitive host from the Russian Far East, and unexpected phylogeny of Azygiidae within Digenea, as inferred from mitogenome sequences. *Journal of Helminthology* **97**, e69, 1–12.
- Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsich G, Pütz J, Middendorf M and Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**, 313–319.
- Biswal DK, Chatterjee A, Bhattacharya A and 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* **12**, e484.
- Briscoe AG, Bray RA, Brabec J and Littlewood DTJ (2016) The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale. *Parasitology International* **65**, 271–275.
- Caira JN and Bogéa T (2005) Family allocreadiidae looss, 1902. In Jones A, Bray RA, Gibson DO (eds), *Keys to the Trematoda*, Vol. 2. Wallingford: CABI Publishing and the Natural History Museum, pp. 417–436.
- Chang QC, Liu GH, Gao JF, Zheng X, Zhang Y, Duan H, Yue DM, Fu X, Su X, Gao Y and Wang CR (2016) Sequencing and characterization of the complete mitochondrial genome from the pancreatic fluke *Eurytrema pancreaticum* (Trematoda: Dicrocoeliidae). *Gene* **2016**, 160–165.
- Choudhury A, Pérez-Ponce de León G, Brooks DR and Daverdin R (2006) *Paracreptotrema blancoi* n. gen. n. sp. (Digenea: Plagiorchiiformes: Allocreadiidae) in the olomina, *Priapichthys annectens* (Osteichthyes: Poeciliidae) from the Área de Conservación Guanacaste, Costa Rica. *Journal of Parasitology* **92**, 565–568.
- Choudhury A, Rosas-Valdez R, Johnson C and Pérez-Ponce de León G (2007) The phylogenetic position of Allocreadiidae (Trematoda: Digenea) from partial sequences of the 18S and 28S ribosomal RNA genes. *Journal of Parasitology* **93**, 192–196.
- Conant GC and Wolfe KH (2008) GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics (Oxford, England)* **24**, 861–862.
- Curran SS, Tkach VV and Overstreet RM (2006) A review of *Polylekithum* Arnold, 1934 and its familial affinities using morphological and molecular data, with description of *Polylekithum catahouleensis* sp. nov. *Acta Parasitologica* **51**, 238–248.
- Curran SS, Tkach VV and Overstreet RM (2011) Phylogenetic affinities of *Auriculostoma* (Digenea: Allocreadiidae), with descriptions of two new species from Peru. *Journal of Parasitology* **97**, 661–670.
- Faltýnková A, Pantoja C, Skírnisson and Kudlai O (2020) Unexpected diversity in Northern Europe: trematodes from salmonid fishes in Iceland with two new species of *Crepidostomum* Braun, 1900. *Parasitology Research* **119**, 2439–2462.
- Franceschini L, Aguiar A, Zago AC, de Oliveira F, Yamada P, Bertholdi EM and da Silva RJ (2021) Three new species of *Creptotrema* (Trematoda, Allocreadiidae) with an amended diagnosis of the genus and reassignment of *Auriculostoma* (Allocreadiidae), based on morphological and molecular evidence. *Parasite* **28**, 69.
- Gao JF, Zhang AH, Wei W, Jia B, Zhang J, Li B, Chen YY, Sun YY, Hou MR, Liu XW, Wang JW, Zhang XH and Wang CR (2023) The complete mitochondrial genome of *Ogmocotyle ailuri*: gene content, composition and rearrangement and phylogenetic implications. *Parasitology* **150**, 1–11.
- García-Varela M and Nadler SA (2005) Phylogenetic relationships of Palaecanthocephala (Acanthocephala) inferred from SSU and LSU rRNA gene sequences. *Journal of Parasitology* **91**, 1401–1409.
- Guo X, Li Y, Gao Y, Qiu Y, Jin Z, Gao Z, Zhang X, An Q, Chang Q, Gao J and Wang C (2022) The complete mitochondrial genome of *Prosthogonimus cuneatus* and *Prosthogonimus pellucidus* (Trematoda: Prosthogonimidae), their features and phylogenetic relationships in the superfamily Microphalloidea. *Acta Tropica* **232**, 106469.
- Krahn N, Fischer JT and Söll D (2020) Naturally occurring tRNAs with non-canonical structures. *Frontiers in Microbiology* **11**, 596914.
- Kumar S, Stecher G, Li M, Knyaz C and Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547.
- Laslett D and Canbäck B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics (Oxford, England)* **24**, 172–175.
- Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Agatsuma T and 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.
- Littlewood DT, Lockyer AE, Webster BL, Johnston DA and Le TH (2006) The complete mitochondrial genomes of *Schistosoma haematobium* and *Schistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. *Molecular Phylogenetics and Evolution* **39**, 452–467.
- Liu GH, Yan HB, Otranto D, Wang XY, Zhao GH, Jia WZ and Zhu XQ (2014) *Dicrocoelium chinensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochondrial and nuclear ribosomal DNA sequences. *Molecular Phylogenetics and Evolution* **79**, 325–331.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J and... Wang J (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 2047–2170.
- Mendoza-Garfias B, García-Teh JG, Caspeta-Mandujano JM, Vidal-Martínez VM and Hernández-Mena DI (2022) Discovery of a new genus and species of Allocreadiidae (Trematoda) in Mexico: *Mesoamerica-trema magnisacculus* n. gen. n. sp. *Helminthologia* **59**, 284–300.
- Pérez-Ponce de León G and 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. *Journal of Helminthology* **93**, 260–276.
- Pérez-Ponce de León G, Choudhury A, Rosas-Valdez R and Mejía-Madrid H (2007) The systematic position of *Wallinia* spp. and *Margotrema* spp. (Digenea), parasites of Middle-American and Neotropical freshwater fishes, based on the 28S ribosomal RNA gene. *Systematic Parasitology* **68**, 49–55.
- Pérez-Ponce de León G, Pinacho-Pinacho CD, Mendoza-Garfias B, Choudhury A and García-Varela M (2016) Phylogenetic analysis using the 28S rRNA gene reveals that the genus *Paracreptotrema* (Digenea: Allocreadiidae) is not monophyletic: description of two new genera and one new species. *Journal of Parasitology* **102**, 131–142.
- Pérez-Ponce de León G, Sereno-Urbe AL, García-Varela M, Mendoza-Garfias B, Hernández-Mena DI, Pinacho-Pinacho CD and Choudhury A (2020) Disentangling the evolutionary and biogeographical history of the freshwater fish trematode genus *Creptotrema* (Digenea: Allocreadiidae) using an integrative taxonomy approach: the case of *Creptotrema agonostomi* in Middle American mountain mullets. *Journal of Helminthology* **94**, 171.
- Petkevičiūtė R, Stunžėnas V and Stanevičiūtė G (2023) Hidden diversity in European *Allocreadium* spp. (Trematoda, Allocreadiidae) and the discovery of the adult stage of *Cercariaeum crassum* Wesenberg-Lund, 1934. *Diversity* **15**, 645.
- Qian L, Zhou P, Li W, Wang H, Miao T and Hu L (2018) Characterization of the complete mitochondrial genome of the lung fluke, *Paragonimus heterotremus*. *Mitochondrial Dna Part B: Resources* **3**, 560–561.
- Rozas J, Ferrer-Mata A, Sánchez-Del Barrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE and Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology Evolution* **34**, 3299–3302.
- Sokolov SG, Khasanov F and Lebedeva DI (2023) Phylogenetic assessment of some Palearctic *Allocreadium* spp. (Trematoda, Gorgoderioidea: Allocreadiidae). *Parasitology Research* **122**, 1923–1933.
- Stamatakis A (2014) RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* **30**, 1312–1313.
- Suleman MAJ, Khan MS, Tkach VV, Muhammad N, Zhang D and Zhu XQ (2019) Characterization of the complete mitochondrial genome of *Plagiorchis maculosus* (Digenea, Plagiorchiidae), Representative of a taxonomically complex digenean family. *Parasitology International* **71**, 99–105.

- Suleman KMS, Tkach VV, Muhammad N, Zhang D, Zhu XQ and Ma J** (2020) Molecular phylogenetics and mitogenomics of three avian dicrocoeliids (Digenea: Dicrocoeliidae) and comparison with mammalian dicrocoeliids. *Parasites & Vectors* **13**, 74.
- Suleman MN, Khan MS, Tkach VV, Ullah H, Ehsan M, Ma J and Zhu XQ** (2021) Mitochondrial genomes of two eucotylids as the first representatives from the superfamily Microphalloidea (Trematoda) and phylogenetic implications. *Parasites & Vectors* **14**, 48.
- Vainutis KS, Voronova AN and Urabe M** (2021) Systematics of *Crepidostomum* species from the Russian Far East and northern Japan, with description of a new species and validation of the genus *Stephanophiala*. *Parasitology International* **84**, 102412.
- Vainutis KS, Voronova AN, Urabe M and Kazarin VM** (2023) Integrative approach for discovering of the new species within the genus *Allocreadium* Looss, 1900 (Trematoda: Allocreadiidae) and framing of biogeographical hypotheses for the genus. *Systematic Parasitology* **100**, 189–213.