Journal of Helminthology

cambridge.org/jhl

Characterization and phylogenomics of the complete mitochondrial genome of the polyzoic cestode *Gangesia oligonchis* (Platyhelminthes: Onchoproteocephalidea)

W.X. Li[1], D. Zhang[1,2], P.P. Fu[1,2], R. Song[3], H. Zou[1], M. Li[1], S.G. Wu[1] and G.T. Wang[1]

[1]Key Laboratory of Aquaculture Disease Control, Ministry of Agriculture, and State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China; [2]University of Chinese Academy of Sciences, Beijing 100049, China and [3]Hunan Fisheries Science Institute, Changsha 410153, China

Research Paper

Cite this article: Li WX, Zhang D, Fu PP, Song R, Zou H, Li M, Wu SG, Wang GT (2020). Characterization and phylogenomics of the complete mitochondrial genome of the polyzoic cestode *Gangesia oligonchis* (Platyhelminthes: Onchoproteocephalidea). *Journal of Helminthology* **94**, e58, 1–10. https://doi.org/10.1017/S0022149X19000452

Received: 26 February 2019
Accepted: 30 April 2019

Key words:
Mitogenome; Onchoproteocephalidea; Eucestoda; parasitic platyhelminths; *Gangesia oligonchis*

Author for correspondence:
G.T. Wang, E-mail: gtwang@ihb.ac.cn

Abstract

The order Onchoproteocephalidea (Eucestoda) was recently erected to accommodate the hook-bearing tetraphyllideans and the proteocephalideans, which are characterized by internal proglottization and a tetra-acetabulate scolex. The recognized subfamilies in the Proteocephalidae appeared to be non-monophyletic based on 28S recombinant DNA (rDNA) sequence data. Other molecular markers with higher phylogenetic resolution, such as large mitochondrial DNA fragments and multiple genes, are obviously needed. Thus the mitochondrial genome of *Gangesia oligonchis*, belonging to the putative earliest diverging group of the Proteocephalidae, was sequenced. The circular mitogenome of *G. oligonchis* was 13,958 bp in size, and contained the standard 36 genes: 22 transfer RNA genes, two rRNA genes and 12 protein-coding genes, as well as two major non-coding regions. A short NCR and a large NCR (lNCR) region were 216 bp and 419 bp in size, respectively. Highly repetitive regions in the lNCR region were detected with that of 11 repeat units. The mitogenome of *G. oligonchis* shared 71.1% nucleotide identity with *Testudotaenia* sp. WL-2016. Phylogenetic analyses of the complete mitochondrial genomes with Bayesian inference and maximum likelihood methods indicated that *G. oligonchis* formed a sister clade with *Testudotaenia* sp. WL-2016 with maximum support. The ordinal topology is (Caryophyllidea, (Diphyllobothriidea, (Bothriocephalidea, (Onchoproteocephalidea, Cyclophyllidea)))). The mitogenomic gene arrangement of *G. oligonchis* was identical to that of *Testudotaenia* sp. WL-2016. Both mitogenomic and nuclear sequence data for many more taxa are required to effectively explore the inter-relationships among the Onchoproteocephalidea.

Introduction

Based on comparative morphology, the eucestodes Tetraphyllidea, Lecanicephalidea, Onchoproteocephalidea (syn. Proteocephalidea), Nippotaeniidea, Tetrabothriidea and Cyclophyllidea are closely related taxa (Hoberg *et al.*, 1997). Evidence from 28S recombinant DNA (rDNA) and 18S rDNA (Olson *et al.*, 2001, 2008; Waeschenbach *et al.*, 2007; Caira *et al.*, 2014) and the large fragments of mitochondrial DNA (mtDNA) (Waeschenbach *et al.*, 2012) suggest that the acetabulate lineages (Litobothriidea, Lecanicephalidea, Nippotaeniidea, Cyclophyllidea and Tetrabothriidea) form a monophyletic group. Owing to the non-monophyly of the order Tetraphyllidea, the new order Rhinebothriidea was established to house the tetraphyllideans with stalked acetabula (Healy *et al.*, 2009), and the new order Onchoproteocephalidea was established for ten described genera of hook-bearing tetraphyllideans and the members of the order Proteocephalidea (Caira *et al.*, 2014).

Onchoproteocephalidean tapeworms, with a cosmopolitan distribution, represent a diverse group of parasites with 316 valid species in bony fish, lizard, snake and amphibian hosts (Caira *et al.*, 2017), and 246 valid species in elasmobranch hosts (Caira *et al.*, 2017). The traditionally accepted families Proteocephalidae and Monticelliidae have been abandoned, and the only family Protecephalidae has been split into a number of subfamilies and genera (de Chambrier *et al.*, 2009). 28S rDNA-based phylogeny suggests that most of the presently recognized subfamilies (and genera) appear to be non-monophyletic, and a deep systematic reorganization of the order is thus needed (de Chambrier *et al.*, 2015). Regardless of the validity of subfamilies, the Gangesiinae and the Acanthotaeniinae appear to be the most primitive, and the Old World (Palaearctic Region) origin of onchoproteocephalideans in freshwater fish is confirmed by the phylogenetic analysis of the 28S rDNA sequence (de Chambrier *et al.*, 2004, 2009, 2015) and the internal transcripted spacer sequence and 18S rRNA (Hypša *et al.*, 2005). Owing to the non-monophyly of subfamilies and the polytomy of the phylogenetic tree

CAMBRIDGE UNIVERSITY PRESS

of siluriform parasites from the Neotropics, large mtDNA fragments and multiple genes are obviously needed (de Chambrier *et al.*, 2015).

Owing to its maternal inheritance, a lack of recombination and a fast rate of evolution, the haploid mitochondrial genome has proven to be a useful marker for population studies, species identification and phylogenetics (Huyse *et al.*, 2008). Using gene sequences and gene arrangements from the complete mitochondrial genome, the phylogenies of some parasitic platyhelminths have been reconstructed (Littlewood *et al.*, 2006). Although mitogenomes for more than 40 cyclophyllideans are available in GenBank, the only onchoproteocephalidean mitogenome available in GenBank is that of *Testudotaenia* sp. WL-2016, belonging to the newly erected subfamily Testudotaeniinae (de Chambrier *et al.*, 2009).

Gangesia oligonchis (Gangesiinae) parasitizes in the intestine of the bullhead catfish, *Tachysurus fulvidraco* (Siluriformes: Bagridae), distributed in Russia (Ash *et al.*, 2015) and China (Fu *et al.*, 2019). Thus the mitogenome of *G. oligonchis* was sequenced and characterized to provide mitogenomic data for future studies on inter-relationships of onchoproteocephalideans.

Materials and methods

Specimen collection and DNA extraction

Tapeworms were collected from the intestine of the bullhead catfish (*T. fulvidraco*), which was anesthetized with 0.02% MS-222, from Liangzi Lake in Hubei Province, China (30°11′05″N, 114°37′34″E). Their identity with *Gangesia oligonchis* was confirmed using morphology and partial 28S rDNA data (Fu *et al.*, 2019). Voucher specimen (accession number: IHB-Gangesia001) was deposited in the museum at the Institute of Hydrobiology, Wuhan, China.

Tapeworm specimens were preserved in 100% ethanol and stored at 4 °C. Total genomic DNA was extracted from a single worm using a TIANamp Micro DNA Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's protocol, and stored at −20 °C.

PCR and DNA sequencing

Sequences from GenBank were used to design five primer pairs (see supplementary material table S1). These primers were used to amplify partial sequences of the *rrnS*, *cox1*, *cytb*, *nad2* and *nad4* genes. Based on these fragments, seven specific primers were designed for subsequent PCR amplification. PCR reactions were conducted in a 20 µl reaction mixture, containing 7.4 µl dd H_2O, 10 µl 2×PCR buffer (Mg^{2+}, dNTP plus, Takara), 0.6 µl of each primer, 0.4 µl rTaq polymerase (250 U, Takara, Dalian, China) and 1 µl DNA template. Amplification was performed under the following conditions: initial denaturation for 2 min at 98 °C, followed by 40 cycles: 10 s at 98 °C, 15 s at 48–60 °C, 1 min at 68 °C and a final extension for 10 min at 68 °C. PCR products were sequenced bidirectionally at Sangon Company (Shanghai, China) using the primer walking strategy.

Sequence annotation analyses

After checking the quality of the sequences, amplification of the complete mitochondrial genomic sequence of *G. oligonchis* was assembled using the DNAstar program (Burland, 2000) and confirmed by BLAST (Altschul *et al.*, 1990). Mitogenome annotation followed the procedure described by Li *et al.* (2017, 2018). Protein-coding genes (PCGs) were detected by searching for open reading frames (employing genetic code 9, invertebrate mitochondrial) and by checking the nucleotide alignments against homologues. All of the transfer RNAs (tRNAs) were predicted and confirmed using the ARWEN program (Laslett & Canback, 2008) and MITOS web server (Bernt *et al.*, 2013). Similarly, the positions of *rrnL* and *rrnS* were preliminarily located using the MITOS program (Bernt *et al.*, 2013), and their ends were assumed to extend to the boundaries of their flanking genes. Tandem repeats in the non-coding region were identified using the Tandem Repeats Finder program (Benson, 1999), and the secondary structure was predicted using Mfold software (Zuker, 2003). Codon usage and relative synonymous codon usage (RSCU) of the 12 PCGs were computed and sorted using the PhyloSuite program (Zhang *et al.*, 2018), and the RSCU figures were finally drawn using ggplot2 plugin (Hadley, 2009). The circular map of the *G. oligonchis* mitogenome was drawn with the mitochondrial visualization tool MTVIZ (http://pacosy.informatik.uni-leipzig.de/mtviz/).

Phylogenetic analyses

Phylogenetic analyses were carried out using the newly sequenced mitogenome of *G. oligonchis* and the 35 cestode mitogenomes available in GenBank (table 1 and supplementary material table S2). Two species of trematode, *Dicrocoelium dendriticum* (Rudolphi, 1819) (NC 025280) and *Dicrocoelium chinensis* (Sudarikov and Ryjikov, 1951) (NC 025279), were used as outgroups. The PhyloSuite program was used to generate the AT content and the GC skew (see supplementary material table S2) and the *.sqn file for GenBank submission. A Fasta file with the nucleotide sequences for all 36 genes (12 PCGs, 2 rRNAs and 22 tRNAs) for the 35 cestodes was downloaded from GenBank using PhyloSuite. All genes were aligned with the MAFFT program (Katoh & Standley, 2013) integrated in PhyloSuite, wherein codon-alignment mode was used for the 12 PCGs, and normal-alignment mode was used for the remaining RNAs (two rRNAs and 22 tRNAs). PhyloSuite was then used to concatenate these alignments and generate input files for the phylogenetic analyses, conducted using maximum likelihood (ML) and Bayesian inference (BI) methods. The most appropriate evolutionary models for the dataset were determined using ModelGenerator v0.8527 (Keane *et al.*, 2006). Based on the Akaike information criterion, GTR + I + G was chosen as the optimal model of nucleotide evolution. ML analysis was performed using the RaxML GUI (Silvestro & Michalak, 2011) and the ML + rapid bootstrap (BP) algorithm with 1000 replicates. BI analysis was performed with MrBayes 3.2.1 (Ronquist *et al.*, 2012) with default settings, and 6×10^6 metropolis-coupled Markov chain Monte Carlo generations. Bayesian posterior probability (BPP) values were calculated in a consensus tree after discarding the first 25% samples as burn-in.

Results

Genome organization and base composition

The typical circular duplex molecule mitogenome of *G. oligonchis* was 13,958 bp in length (GenBank accession number: MF314173). Apart from lacking the *Atp8* gene, which is typical of parasitic flatworms (Le *et al.*, 2002), the mitogenome contained the standard 36 genes: 22 tRNA genes, two rRNA genes and 12 PCGs, as well as two major non-coding regions (mNCRs) (fig. 1). All genes were transcribed from the same strand. Five overlapping regions were found in the genome (table 2).

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

Table 1. The list of cestode species used for comparative analyses mitogenomes.

Species	GeneBank accession no.	Full length (bp)	A (%)	T (%)	C (%)	G (%)	A+T (%)	G+C (%)	AT skew	GC skew
Anoplocephala magna	KU236385	13,759	24.3	46.5	8.2	21	70.8	29.2	−0.313	0.441
Anoplocephala perfoliata	NC_028425	14,459	24.9	46	8.5	20.5	70.9	29	−0.297	0.412
Atractolytocestus huronensis	KY486754	15,130	22.4	39.9	13.3	24.4	62.3	37.7	−0.28	0.294
Breviscolex orientalis	KY486752	14,011	19.1	39.5	14.9	26.5	58.6	41.4	−0.348	0.281
Cladotaenia vulturi	KU559932	13,411	27.8	46.9	8.1	17.2	74.7	25.3	−0.256	0.36
Cloacotaenia megalops	KU641017	13,887	26.4	45.2	9.7	18.7	71.6	28.4	−0.264	0.314
Dicrocoelium chinensis	NC_025279	14,917	18.1	44	10	27.9	62.1	37.9	−0.417	0.473
Dicrocoelium dendriticum	NC_025280	14,884	18.2	44	10.1	27.7	62.2	37.8	−0.414	0.466
Dibothriocephalus latus	NC_008945	13,608	23.6	44.6	12	19.8	68.2	31.8	−0.309	0.245
Dibothriocephalus nihonkaiensis	NC_009463	13,747	23.7	44.1	12.5	19.7	67.8	32.2	−0.301	0.225
Diphyllobothrium balaenopterae	NC_017613	13,724	23.1	45.5	11.2	20.2	68.6	31.4	−0.326	0.287
Diphyllobothrium s grandis	NC_017615	13,725	23.2	45.6	11.2	20.1	68.8	31.3	−0.326	0.286
Dipylidium caninum	NC_021145	14,296	21.4	51.5	7.6	19.5	72.9	27.1	−0.413	0.441
Drepanidotaenia lanceolata	NC_028164	13,573	23.9	46	9.9	20.2	69.9	30.1	−0.317	0.342
Echinococcus equinus	NC_020374	13,605	19.9	48	7.8	24.3	67.9	32.1	−0.414	0.516
Echinococcus granulosus	KJ559023	13,605	19.1	47.9	8.1	25	67	33.1	−0.43	0.512
Echinococcus oligarthrus	NC_009461	13,791	20.8	48.5	7.4	23.3	69.3	30.7	−0.399	0.516
Echinococcus vogeli	NC_009462	13,750	18.9	48.2	7.6	25.3	67.1	32.9	−0.436	0.539
Gangesia oligonchis	MF314173	13,958	23	43.3	12.9	20.8	66.3	33.7	−0.307	0.235
Hydatigera kamiyai	AB731761	13,853	26.2	46.4	8.1	19.3	72.6	27.4	−0.278	0.407
Hydatigera krepkogorski	NC_021142	13,792	27.8	45	8.5	18.7	72.8	27.2	−0.236	0.374
Hydatigera parva	NC_021141	13,482	25.5	45.9	8.3	20.3	71.4	28.6	−0.285	0.419
Hymenolepis diminuta	AF314223	13,900	25.4	45.6	9.6	19.3	71	28.9	−0.285	0.334
Hymenolepis nana	NC_029245	13,764	27	46	8.8	18.2	73	27	−0.261	0.346
Khawia sinensis	NC_034800	13,759	21.6	38.7	16.2	23.5	60.3	39.7	−0.285	0.183
Khawia sinensis	KY486753	14,620	23.9	41.7	12.7	21.7	65.6	34.4	−0.271	0.263
Pseudanoplocephala crawfordi	NC_028334	14,192	23.9	45.8	9.2	21.1	69.7	30.3	−0.313	0.39
Schyzocotyle acheilognathi	KX589243	14,046	23.2	45.8	10.3	20.6	69	30.9	−0.327	0.333
Schyzocotyle acheilognathi	KX060595	13,849	23.2	45.9	10.3	20.6	69.1	30.9	−0.328	0.336
Schyzocotyle nayarensis	NC_030317	13,852	24	45.3	10.3	20.4	69.3	30.7	−0.307	0.327
Senga ophiocephalina	NC_034715	13,816	23.7	46.4	9.8	20.2	70.1	30	−0.324	0.348
Spirometra decipiens	NC_026852	13,641	20.3	46	11	22.6	66.3	33.6	−0.387	0.345
Spirometra erinaceieuropaei	NC_011037	13,643	20.3	46.1	11	22.6	66.4	33.6	−0.388	0.348
Taenia crassiceps	AF216699	13,503	25.4	48.6	7.6	18.3	74	25.9	−0.314	0.413
Taenia pisiformis	NC_013844	13,387	27.7	45.4	8.8	18	73.1	26.8	−0.242	0.343
Taenia saginata	NC_009938	13,670	24.2	47.3	7.9	20.6	71.5	28.5	−0.322	0.445
Testudotaenia sp. WL-2016	KU761587	13,709	23.4	42.7	13.2	20.7	66.1	33.9	−0.292	0.22
Versteria mustelae	NC_021143	13,582	23.2	48.2	8	20.7	71.4	28.7	−0.35	0.445

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

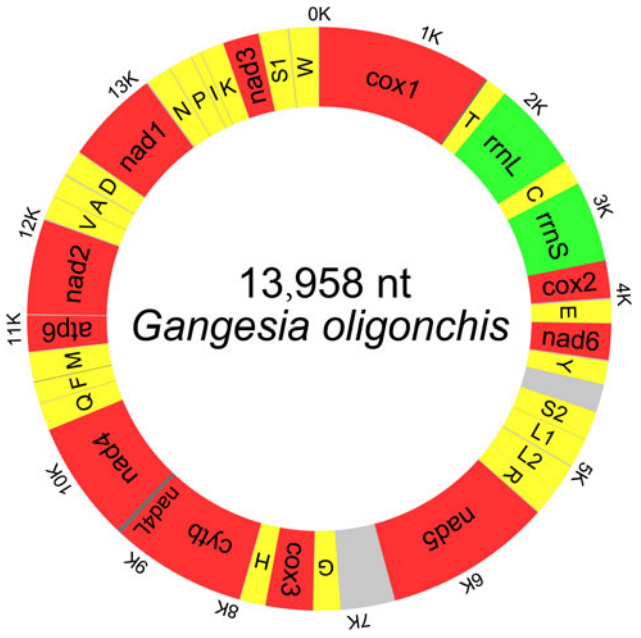

Fig. 1. Map of the complete mitochondrial genome of *Gangesia oligonchis*. All 36 genes and major non-coding regions are displayed.

An A + T bias was detected in the mitogenome of *G. oligonchis* (A = 23.0%, T = 43.3%, C = 12.9%, G = 20.8%). The nucleotide composition of the complete mitogenome of *G. oligonchis* was strongly skewed away from A, in favour of T, and was biased towards G (AT skew = −0.307, GC skew = 0.235) (table 3). The mitogenomes of *G. oligonchis* and *Testudotaenia* sp. WL-2016 shared 71.1% nucleotide identity, with 65.4–78.5% identity in PCGs and rRNA genes (table 2).

Protein-coding genes and codon usage

The total length of the concatenated 12 PCGs was 10,122 bp, with the average A + T content of 65.4%, ranging from 63.8% (*cox2*) to 71% (*nad3*) (table 3). The start codon ATG was commonly found in ten PCGs, whereas the start codon GTG was most commonly found in *cox3* and *nad6*. The most frequent stop codons were TAG (for eight PCGs), followed by TAA (four PCGs). Codon usage, relative synonymous codon usage (RSCU), and codon family proportion (corresponding to the amino acid usage) of these two onchoproteocephalideans are presented in fig. 2. Leucine (14.9%), phenylalanine (12.9%), valine (10.4%) and isoleucine (6.7%) were the most frequent amino acids in the PCGs of *G. oligonchis*, which is also observed in *Testudotaenia* sp. WL-2016 (fig. 2). In particular, all second codon positions of the codons encoding these amino acids were T, corresponding to its relatively high T skewness (AT skew = −0.465, table 3). The most frequent codons were TTT (phenylalanine, 11.3%) and TTA (leucine, 6.4%), both consisted of A and T. Codons ending in A or T were predominant (blue and green bar in fig. 2), which corresponds to the high A + T content of the third coding position of all PCGs in *G. oligonchis* (67.7%).

Transfer and ribosomal RNA genes

All 22 tRNAs were found in the mitogenome of *G. oligonchis*; these ranged in length from 59 bp (*trnS1*) to 67 bp (*trnN*, *trnM* and *trnG*) (table 2). In terms of secondary structure, most of the tRNA sequences could be folded into the conventional cloverleaf shape; the exceptions were *trnS1* and *trnR*, which lacked the dihydrouridine arms and loops. Standard anticodons were found in all tRNAs, except for *trnR*, which exhibited a transition from U to A. The genes *rrnL* and *rrnS* were 975 bp and 732 bp in length, with 66.7% and 67.2% A + T content, respectively (table 3). They were separated by *trnC*. The mitogenomic gene arrangement of *trnT–rrnL–trnC–rrnS* is shared by all cestodes characterized so far (fig. 3).

Non-coding regions

A total of 22 short intergenic regions (1 to 12 bp in length) were interspersed within the mitogenome of *G. oligonchis* (table 2). These included two mNCRs consisting of an sNCR located between *trnY* and *trnS2* and a lNCR located between *Nad5* and *trnG*. These non-coding regions were 216 and 419 bp in size, respectively; they had a much higher A + T content at 78.7% and 84.3%, respectively (table 3). The sNCR and lNCR of *Testudotaenia* sp. WL-2016 (108 bp and 265 bp in size, respectively) were much smaller than those of *G. oligonchis*. Highly repetitive regions (HRRs) in lNCR were detected in both *G. oligonchis* and *Testudotaenia* sp. WL-2016, with 11 and six repeat units, respectively (fig. 4). Although the predicted stem (2 bp) of the *Testudotaenia* sp. WL-2016 repeat unit was extremely short, both consensus repeat units were capable of forming stem–loop structures (fig. 4). In addition, the sNCR of the two onchoproteocephalidean mitogenomes was also capable of forming a stem–loop structure (fig. 4).

Phylogeny and gene order

Phylogenetic analyses of the concatenated 36 mitochondrial genes using BI and ML methods produced identical tree topologies in which *G. oligonchis* grouped as the sister taxon of *Testudotaenia* sp. WL-2016 with maximum support (BP = 100, BPP = 1). The ordinal topology is (Caryophyllidea, (Diphyllobothriidea, (Bothriocephalidea, (Onchoproteocephalidea, Cyclophyllidea)))). The mitogenomic gene arrangement of *G. oligonchis* was identical to that of *Testudotaenia* sp. WL-2016 and some cyclophyllideans, such as those of the families Hymenolepididae, Anoplocephalidae, Dipylidiidae and Paruterinidae (fig. 3).

Discussion

The topology of the trees resulting from phylogenetic analyses of the concatenated 36 mitochondrial genes was stable and in full agreement with those generated from studies based on complete mitogenomes (Li *et al.*, 2017) and large and small subunits of nuclear ribosomal RNA genes (28S rDNA and 18S rDNA) (Brabec *et al.*, 2006; Waeschenbach *et al.*, 2012; Kuchta & Scholz, 2017). However, this topology deviated from that recovered from the recent mitogenomic studies of Feng *et al.* (2017) and Zhang *et al.* (2017), in which a sister-group relationship of Diphyllobothriidea and Bothriocephalidea was recovered. These discrepancies may be caused by idiosyncrasies of different phylogenetic reconstruction software. To test this, another phylogenetic analysis was performed using the RaxML program (Silvestro & Michalak, 2011) (instead of MEGA) based on the same dataset and computational model as those of the studies by Feng *et al.* and Zhang *et al.* The resultant topology was identical to ours (see supplementary material fig. S1). Furthermore, a sister-group relationship of

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

Table 2. The organization of the mitochondrial genome of *Gangesia oligonchis*.

Gene	Position		Size	Intergenic nucleotides	Codon		Anti-codon	Identify (%)
	From	To			Start	Stop		
Gangesia oligonchis/ *Testudotaenia* sp. WL-2016								
cox1	1/1	1644/1617	1644/1617		ATG	TAA		73.66
trnT	1630/1608	1692/1670	63/63	−15/−10			TGT	84.13
rrnL	1693/1671	2667/2636	975/966					75.74
trnC	2668/2637	2732/2704	65/68				GCA	67.65
rrnS	2733/2705	3464/3416	732/712					74.46
cox2	3465/3423	4034/3992	570/570	-/6	ATG	TAA		75.79
trnE	4047/3999	4110/4065	64/67	12/6			TTC	74.63
nad6	4114/4070	4569/4528	456/459	3/4	GTG/ ATG	TAG		71.02
trnY	4579/4540	4644/4606	66/67	9/11			GTA	88.06
trnS2	4861/4715	4925/4780	65/66	**216/108**			TGA	83.33
trnL1	4928/4793	4992/4856	65/64	2/12			TAG	77.27
trnL2	5001/4860	5066/4924	66/65	8/3			TAA	83.33
trnR	5066/4925	5124/4984	59/60	−1/-			ACG	68.85
nad5	5131/4991	6705/6565	1575/1575	6/6	ATG	TAA		65.4
trnG	7125/6831	7191/6895	67/65	**419/265**			TCC	74.63
cox3	7195/6899	7839/7543	645/645	3/3	GTG	TAG/ TAA		69.3
trnH	7841/7554	7905/7621	65/68	1/10			GTG	85.29
cytb	7911/7627	9008/8724	1098/1098	5/5	ATG	TAA		76.05
nad4L	9009/8738	9269/8998	261/261	-/13	ATG	TAG/ TAA		68.58
nad4	9236/8965	10,480/10,209	1245/1245	−34/−34	ATG	TAG/ TAA		67.07
trnQ	10,481/10,210	10,543/10,272	63/63				TTG	84.13
trnF	10,547/10,282	10,611/10,341	65/60	3/9			GAA	65.15
trnM	10,608/10,342	10,674/10,408	67/67	−4/-			CAT	68.66
atp6	10,681/10,412	11,196/10,927	516/516	6/3	ATG	TAG		74.42
nad2	11,205/10,936	12,077/11,808	873/873	8/8	ATG	TAG		74.23
trnV	12,087/11,812	12,151/11,875	65/64	9/3			TAC	86.15
trnA	12,155/11,881	12,217/11,944	63/64	3/5			TGC	93.75
trnD	12,227/11,951	12,289/12,013	63/63	9/6			GTC	65.08
nad1	12,295/12,018	13,185/12,908	891/891	5/4	ATG	TAG		78.45
trnN	13,197/12,908	13,263/12,973	67/66	11/−1			GTT	92.54
trnP	13,269/12,980	13,332/13,041	64/62	5/6			TGG	70.31
trnI	13,338/13,056	13,403/13,120	66/65	5/14			GAT	92.42
trnK	13,408/13,132	13,470/13,195	63/64	4/11			CTT	73.85
nad3	13,473/13,200	13,820/13,547	348/348	2/4	ATG	TAG		75.57
trnS1	13,819/13,560	13,878/13,619	60/60	−2/12			GCT	78.33
trnW	13,886/13,640	13,951/13,702	66/63	7/20			TCA	77.27
Full mitogenome			13,958/13,722					71.11

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

Table 3. Nucleotide composition of the protein-coding genes, tRNAs, rRNAs and non-coding region of mitochondrial genomes of *Gangesia oligonchis* and *Testudotaenia* sp. WL-2016.

Regions	Size (bp)	T(U)	C	A	G	AT(%)	GC(%)	AT skew	GC skew
Gangesia oligonchis									
PCGs	10,122	44.7	13.3	20.7	21.3	65.4	34.6	−0.368	0.232
1st codon position	3374	40.3	11.8	23.3	24.5	63.6	36.3	−0.267	0.349
2nd codon position	3374	47.5	15	17.3	20.2	64.8	35.2	−0.465	0.148
3rd codon position	3374	46.4	13.1	21.3	19.2	67.7	32.3	−0.37	0.19
atp6	513	47.6	14.8	17.2	20.5	64.8	35.3	−0.47	0.16
cox1	1641	42.3	13.8	22.5	21.5	64.8	35.3	−0.306	0.218
cox2	567	38.8	14.1	25	22	63.8	36.1	−0.215	0.22
cox3	642	46	14.5	18.5	21	64.5	35.5	−0.425	0.184
cytb	1095	42.5	14.4	21.7	21.4	64.2	35.8	−0.323	0.194
nad1	888	46.1	11.7	18.9	23.3	65	35	−0.418	0.331
nad2	870	50.3	10.2	17.9	21.5	68.2	31.7	−0.475	0.355
nad3	345	51.3	10.7	19.7	18.3	71	29	−0.445	0.26
nad4	1242	44.5	15.1	19.7	20.7	64.2	35.8	−0.386	0.158
nad4L	258	49.2	11.2	21.7	17.8	70.9	29	−0.388	0.227
nad5	1572	42.6	13.9	21.5	22	64.1	35.9	−0.329	0.225
nad6	453	49.4	10.6	19.6	20.3	69	30.9	−0.431	0.314
rrnL	975	39.1	13.3	27.6	20	66.7	33.3	−0.172	0.2
rrnS	732	37.6	12.7	29.6	20.1	67.2	32.8	−0.118	0.225
lNCR	419	50.4	1.7	33.9	14.1	84.3	15.8	−0.195	0.788
sNCR	216	38	10.6	40.7	10.6	78.7	21.2	0.035	0
tRNAs	1417	37.8	13.7	26.7	21.9	64.5	35.6	−0.172	0.23
Full mitogenome	13,958	43.3	12.9	23	20.8	66.3	33.7	−0.307	0.235
Testudotaenia sp. WL-2016									
PCGs	10,098	44.2	13.6	21.2	21	65.4	34.6	−0.353	0.216
1st codon position	3366	40.3	12	24.3	23.4	64.6	35.4	−0.247	0.322
2nd codon position	3366	47.7	15.3	17.1	20	64.8	35.3	−0.472	0.133
3rd codon position	3366	44.8	13.4	22.1	19.7	66.9	33.1	−0.34	0.189
atp6	513	45.4	14	19.3	21.2	64.7	35.2	−0.404	0.204
cox1	1614	41	15.3	22.5	21.2	63.5	36.5	−0.292	0.161
cox2	567	36	16.9	23.8	23.3	59.8	40.2	−0.204	0.158
cox3	642	45.5	13.1	19.3	22.1	64.8	35.2	−0.404	0.257
cytb	1095	42.9	13.6	21.9	21.6	64.8	35.2	−0.324	0.226
nad1	888	45.2	11.3	19.8	23.8	65	35.1	−0.39	0.357
nad2	870	50	10.7	19	20.3	69	31	−0.45	0.311
nad3	345	49	8.7	18.6	23.8	67.6	32.5	−0.451	0.464
nad4	1242	45.2	15.5	18.4	20.9	63.6	36.4	−0.42	0.15
nad4L	258	48.8	10.5	23.3	17.4	72.1	27.9	−0.355	0.25
nad5	1572	42.7	14.4	23.7	19.1	66.4	33.5	−0.285	0.14
nad6	456	50.7	11.6	20.2	17.5	70.9	29.1	−0.43	0.203
rrnL	966	38.5	12.7	29.4	19.4	67.9	32.1	−0.134	0.206
rrnS	712	36.8	14	29.1	20.1	65.9	34.1	−0.117	0.177

(*Continued*)

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

Table 3. (*Continued.*)

Regions	Size (bp)	T(U)	C	A	G	AT(%)	GC(%)	AT skew	GC skew
LNR	265	40	3.4	34.7	21.9	74.7	25.3	−0.071	0.731
SNR	108	36.1	9.3	44.4	10.2	80.5	19.5	0.103	0.048
tRNAs	1414	37.7	13.2	28.4	20.7	66.1	33.9	−0.14	0.219
Full mitogenome	13,709	42.7	13.2	23.4	20.7	66.1	33.9	−0.292	0.22

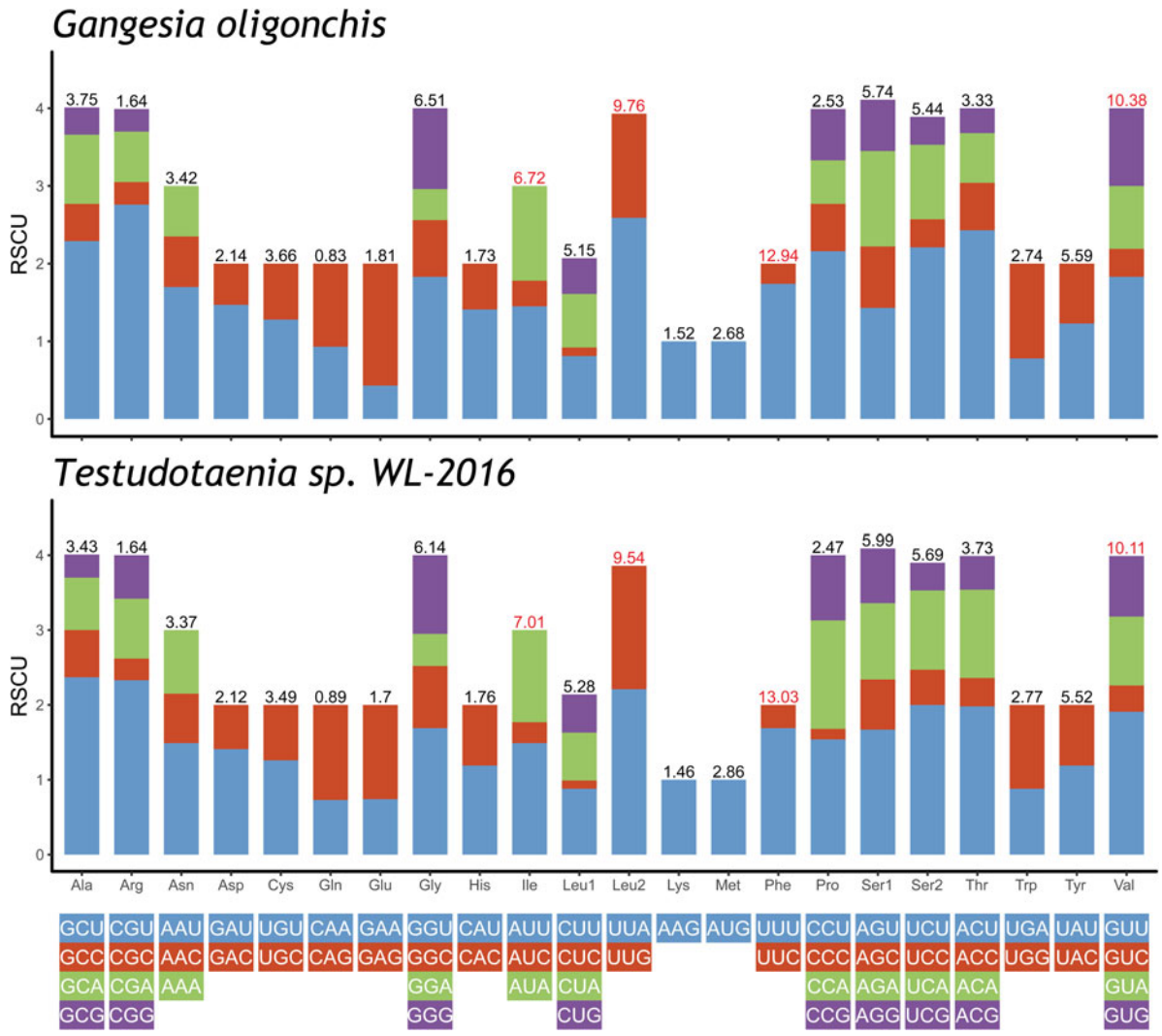

Fig. 2. Relative synonymous codon usage (RSCU) of the complete mitochondrial genome of *Gangesia oligonchis* and *Testudotaenia* sp. WL-2016. Codon families are labelled on the *x*-axis. Values on the top of the bars refer to amino acid usage.

Bothriocephalidea and 'acetabulate' was also supported using a much denser taxon by 28S rDNA + 18S rDNA (Brabec *et al.*, 2006; Kuchta *et al.*, 2008) and partial mtDNA + 28S rDNA + 18S rDNA (Waeschenbach *et al.*, 2012; Kuchta & Scholz, 2017).

The mitogenomic gene arrangement of *G. oligonchis* was identical to that of *Testudotaenia* sp. WL-2016, which belonged to gene arrangement category IV (onchoproteocephalideans and some cyclophyllideans) as summarized by Li *et al.* (2017) (fig. 3). These authors concluded that all rearrangement events in cestode mitogenomes were observed in the rearrangement hot spot–P1 (i.e. gene block between *rrnS* and *trnR*), a region that often harbours an NCR (non-coding region). The TDRL (tandem–duplication–random–loss) event contributed to the increased rate of rearrangement for genes adjacent to the origin replication because both strand slippage and imprecise termination were more likely to include the genes surrounding the origin of replication in the rearrangement hot spot (Cameron, 2014). Despite the rearrangement hot spot in Cestoda, the mitogenomic gene arrangement was too conserved to reflect the interrelationships within the order Onchoproteocephalidea.

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press

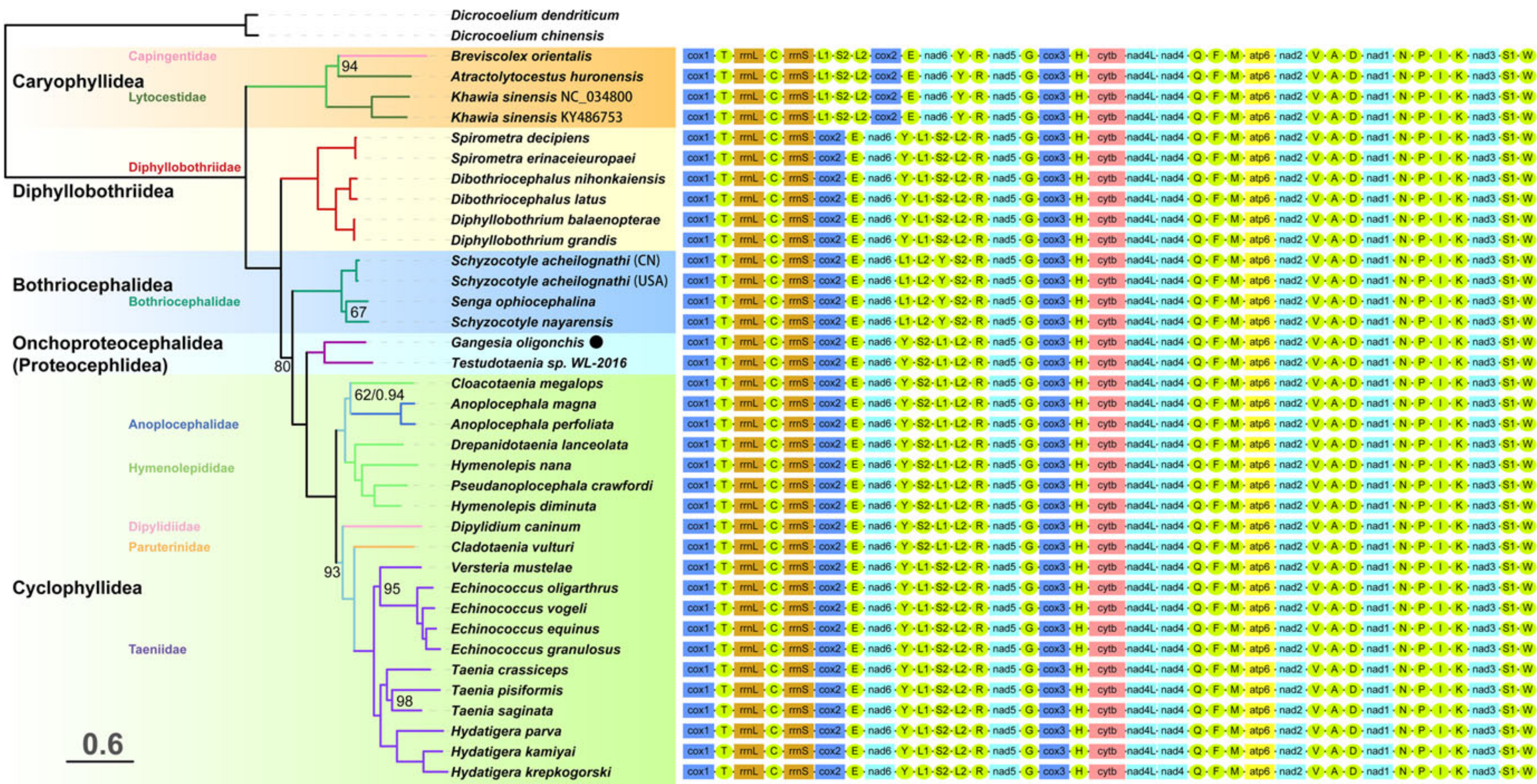

Fig. 3. The phylogenetic relationships of the five orders in Cestoda inferred from concatenated 36 genes representing almost complete mitogenomic datasets (36 genes: 12 PCGs, 2 rRNAs and 22 tRNAs), using two Trematoda species as outgroup. Scale bar represents the estimated number of substitutions per site. Bootstrap (BP)/posterior probability (BPP) support values of ML/BI analysis are shown above the nodes, only BP <100 and BPP <1 were displayed. Mitogenomic gene orders of the selected cestode species (corresponding to tip labels in the tree) were listed on the right of the tree. The order was reoriented to *cox1*.

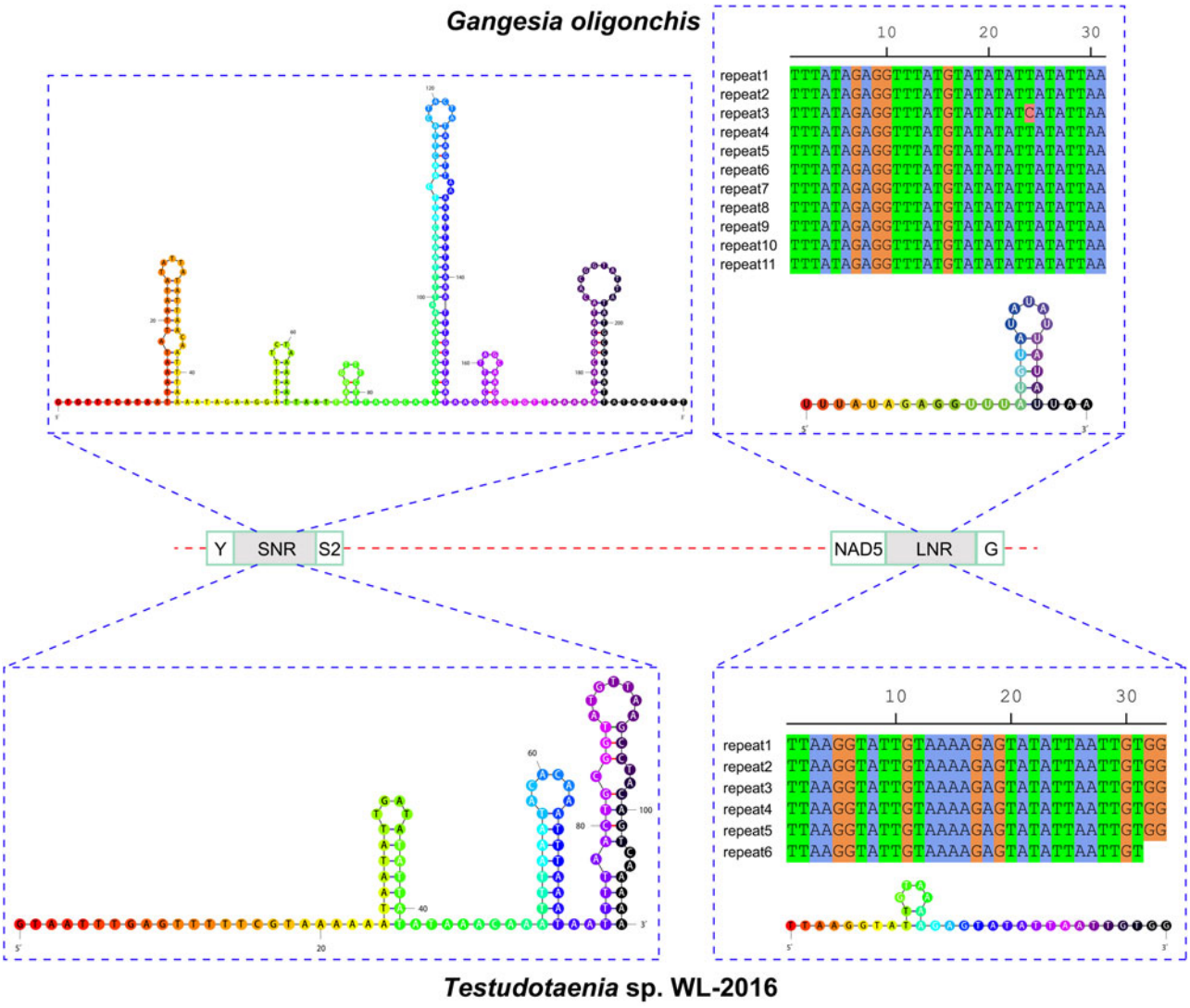

Fig. 4. Illustration of highly repetitive regions in the large major non-coding region and the predicted secondary structure of the short non-coding region of *Gangesia oligonchis* and *Testudotaenia* sp. WL-2016 mitochondrial genome.

Although mitogenomes are currently available for only two species of the order Onchoproteocephalidea, the low sequence identity (71.1%) between the mitogenomes of *G. oligonchis* and *Testudotaenia* sp. WL-2016 may provide some phylogenetic information. Within the family Diphyllobothriidae (Diphyllobothriidea), sequence identity of the mitogenome ranged from 85% to 87% between *Ligula* spp. and *Dibothriocephalus* spp. (Li *et al.*, 2018), which was much higher than that between *G. oligonchis* and *Testudotaenia* sp. WL-2016. *Gangesia* from catfishes, mostly in Indomalaya and Palearctic, is the early diverging group, while *Testudotaenia* from soft-shelled turtles in North America is the derived group (de Chambrier *et al.*, 2015). However, families and subfamilies within the Onchoproteocephalidea need to be determined based on mitogenomes of more taxa.

Conclusions

The complete mitogenome of the tapeworm *Gangesia oligonchis* from the bullhead catfish *Tachysurus fulvidraco* was sequenced and characterized. The mitogenomic gene arrangement was found to be conserved across the two members of the order Onchoproteocephalidea for which such data are available. While low nucleotide identity was found between the two onchoproteocephalideans, mitogenomes of more extensive taxa are expected to be sequenced to effectively explore the inter-relationships among the Onchoproteocephalidea.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S0022149X19000452

Author ORCIDs. W.X. Li, 0000-0001-7016-4535

Acknowledgements. The authors would like to thank Dr Jakovlić I. for his help in English language editing.

Financial support. This work was funded by the National Natural Science Foundation of China (31872604) and the earmarked fund for China Agriculture Research System (CARS-45-15).

Conflicts of interest. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional guides on the care and use of laboratory animals.

References

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410.

Ash A, de Chambrier A, Shimazu T, Ermolenko A and Scholz T (2015) An annotated list of the species of *Gangesia* Woodland, 1924 (Cestoda: Proteocephalidea), parasites of catfishes in Asia, with new synonyms and a key to their identification. *Systematic Parasitology* **91**, 13–33.

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**, 573.

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M and Stadler PF (2013) MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**, 313–319.

Brabec J, Kuchta R and Scholz T (2006) Paraphyly of the Pseudophyllidea (Platyhelminthes: Cestoda): Circumscription of monophyletic clades based on phylogenetic analysis of ribosomal RNA. *International Journal for Parasitology* **36**, 1535–1541.

Burland TG (2000) DNASTAR's Lasergene sequence analysis software. In Misener S and Krawetz SA (Eds) *Bioinformatics Methods and Protocols*. Methods in Molecular Biology, 132. Totowa, NJ, Humana Press

Caira JN, Jensen K, Waeschenbach A, Olson PD and Littlewood DTJ (2014) Orders out of chaos - molecular phylogenetics reveals the complexity of shark and stingray tapeworm relationships. *International Journal for Parasitology* **44**, 55–73.

Caira JN, Jensen K and Ivanov VA (*Eds*) (2017) *Onchoproteocephalidea II*. Lawrence, KS: University of Kansas, Natural History Museum.

Cameron SL (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology* **59**, 95–117.

de Chambrier A, Zehnder M, Vaucher C and Mariaux J (2004) The evolution of the Proteocephalidea (Platyhelminthes, Eucestoda) based on an enlarged molecular phylogeny, with comments on their uterine development. *Systematic Parasitology* **57**, 159–171.

de Chambrier A, Coquille SC, Mariaux J and Tkach V (2009) Redescription of *Testudotaenia testudo* (Magath, 1924) (Eucestoda: Proteocephalidea), a parasite of *Apalone spinifera* (Le Sueur) (Reptilia: Trionychidae) and *Amia calva* L. (Pisces: Amiidae) in North America and erection of the Testudotaeniinae n. subfam. *Systematic Parasitology* **73**, 49–64.

de Chambrier A, Waeschenbach A, Fisseha M, Scholz T and Mariaux J (2015) A large 28S rDNA-based phylogeny confirms the limitations of established morphological characters for classification of proteocephalidean tapeworms (Platyhelminthes, Cestoda). *ZooKeys* **500**, 25–59.

Feng Y, Feng HL, Fang YH and Su YB (2017) Characterization of the complete mitochondrial genome of *Khawia sinensis* belongs among platyhelminths, cestodes. *Experimental Parasitology* **177**, 35–39.

Fu PP, Li WX, Zou H, Zhang D, Wu SG, Li M, Wang GT and Xi BW (2019) Identification of *Gangesia oligonchis* (Cestoda: Onchoproteocephalidea) from *Tachysurus fulvidraco* in Central China: implications for the invalidity of *Gangesia pseudobagrae*. *Systematic Parasitology* **96**, 327–335.

Hadley W (2009) ggplot2: Elegant graphics for data analysis.

Healy CJ, Caira JN, Jensen K, Webster BL and Littlewood DTJ (2009) Proposal for a new tapeworm order, Rhinebothriidea. *International Journal for Parasitology* **39**, 497–511.

Hoberg EP, Mariaux J, Justine JL, Brooks DR and Weekes PJ (1997) Phylogeny of the orders of the Eucestoda (Cercomeromorphae) based on comparative morphology: Historical perspectives and a new working hypothesis. *Journal of Parasitology* **83**, 1128–1147.

Huyse T, Buchmann K and Littlewood DTJ (2008) The mitochondrial genome of *Gyrodactylus derjavinoides* (Platyhelminthes: Monogenea) - A mitogenomic approach for *Gyrodactylus* species and strain identification. *Gene* **417**, 27–34.

Hypša V, Škerikoa A and Scholz T (2005) Phylogeny, evolution and host–parasite relationships of the order Proteocephalidea (Eucestoda) as revealed by combined analysis and secondary structure characters. *Parasitology* **130**, 359–371.

Katoh K and Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772–780.

Keane TM, Creevey CJ, Pentony MM, Naughton TJ and McInerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology* **6**, 29.

Kuchta R and Scholz T (*Eds*) (2017) *Bothriocephalidea*. Lawrence, KS: University of Kansas, Natural History Museum.

Kuchta R, Scholz T, Brabec J and Bray RA (2008) Suppression of the tapeworm order Pseudophyllidea (Platyhelminthes: Eucestoda) and the proposal of two new orders, Bothriocephalidea and Diphyllobothriidea. *International Journal for Parasitology* **38**, 49–55.

Laslett D and Canback B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**, 172–175.

Le TH, Blair D and McManus DP (2002) Mitochondrial genomes of parasitic flatworms. *Trends in Parasitology* **18**, 206–213.

Li WX, Zhang D, Boyce K, Xi BW, Zou H, Wu SG, Li M and Wang GT (2017) The complete mitochondrial DNA of three monozoic tapeworms in the Caryophyllidea: a mitogenomic perspective on the phylogeny of eucestodes. *Parasites & Vectors* **10**, 314.

Li WX, Fu PP, Zhang D, Boyce K, Xi BW, Zou H, Li M, Wu SG and Wang GT (2018) Comparative mitogenomics supports synonymy of the genera *Ligula* and *Digramma* (Cestoda: Diphyllobothriidae). *Parasites & Vectors* **11**, 324.

Littlewood DTJ, 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.

Olson PD, Littlewood DTJ, Bray RA and Mariaux J (2001) Interrelationships and evolution of the tapeworms (Platyhelminthes: Cestoda). *Molecular Phylogenetics and Evolution* **19**, 443–467.

Olson PD, Poddubnaya LG, Littlewood DTJ and Scholz T (2008) On the position of *Archigetes* and its bearing on the early evolution of the tapeworms. *Journal of Parasitology* **94**, 898–904.

Ronquist F, Teslenko M, van der Mark P, *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539–542.

Silvestro D and Michalak I (2011) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution* **12**, 335–337.

Waeschenbach A, Webster BL, Bray RA and Littlewood DTJ (2007) Added resolution among ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. *Molecular Phylogenetics and Evolution* **45**, 311–325.

Waeschenbach A, Webster BL and Littlewood DTJ (2012) Adding resolution to ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with large fragments of mtDNA. *Molecular Phylogenetics and Evolution* **63**, 834–847.

Zhang X, Duan JY, Shi YL, Jiang P, Zeng DJ, Wang ZQ and Cui J (2017) Comparative mitochondrial genomics among *Spirometra* (Cestoda: Diphyllobothriidae) and the molecular phylogeny of related tapeworms. *Molecular Phylogenetics and Evolution* **117**, 75–82.

Zhang D, Gao F, Li WX, Jakovlić I, Zou H, Zhang J and Wang GT. (2018) PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *bioRxiv*. doi: 10.1101/489088.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406–3415.

https://doi.org/10.1017/S0022149X19000452 Published online by Cambridge University Press