Parasitology

cambridge.org/par

Comparative mitogenomics of the zoonotic parasite *Echinostoma revolutum* resolves taxonomic relationships within the '*E. revolutum*' species group and the Echinostomata (Platyhelminthes: Digenea)

Thanh Hoa Le[1,2], Linh Thi Khanh Pham[1], Huong Thi Thanh Doan[1], Xuyen Thi Kim Le[1], Weerachai Saijuntha[3], R.P.V. Jayanthe Rajapakse[4] and Scott P. Lawton[5]

[1]Institute of Biotechnology (IBT); Vietnam Academy of Science and Technology (VAST), 18. Hoang Quoc Viet Rd, Cau Giay, Hanoi, Vietnam; [2]Graduate University of Science and Technology (GUST), Vietnam Academy of Science and Technology (VAST), 18. Hoang Quoc Viet Rd, Cau Giay, Hanoi, Vietnam; [3]Walai Rukhavej Botanical Research Institute (WRBRI), Biodiversity and Conservation Research Unit, Mahasarakham University, Mahasarakham 44150, Thailand; [4]Department of Veterinary Pathobiology, Faculty of Veterinary Medicine and Animal Science, University of Peradeniya, Peradeniya, Sri Lanka and [5]Molecular Parasitology Laboratory, School of Life Sciences, Pharmacy and Chemistry, Kingston University London, Kingston Upon Thames, Surrey, KT1 2EE, UK

Research Article

Cite this article: Le TH, Pham LTK, Doan HTT, Le XTK, Saijuntha W, Rajapakse RPVJ, Lawton SP (2020). Comparative mitogenomics of the zoonotic parasite *Echinostoma revolutum* resolves taxonomic relationships within the '*E. revolutum*' species group and the Echinostomata (Platyhelminthes: Digenea). *Parasitology* **147**, 566–576. https://doi.org/10.1017/S0031182020000128

Received: 3 December 2019
Revised: 1 January 2020
Accepted: 15 January 2020
First published online: 29 January 2020

Key words:
'37-collar-spined'; *Echinostoma revolutum*; Echinostomatidae; Echinostomatoidea; mitochondrial genome; phylogenetic analysis; repeats; skewness value

Author for correspondence:
Thanh Hoa Le, E-mail: imibtvn@gmail.com

Abstract

The complete mitochondrial sequence of 17,030 bp was obtained from *Echinostoma revolutum* and characterized with those of previously reported members of the superfamily Echinostomatoidea, i.e. six echinostomatids, one echinochasmid, five fasciolids, one himasthlid, and two cyclocoelids. Relationship within suborders and between superfamilies, such as Echinostomata, Pronocephalata, Troglotremata, Opisthorchiata, and Xiphiditata, are also considered. It contained 12 protein-coding, two ribosomal RNA, 22 transfer RNA genes and a tandem repetitive consisting non-coding region (NCR). The gene order, one way-positive transcription, the absence of *atp*8 and the overlapped region by 40 bp between *nad*4L and *nad*4 genes were similar as in common trematodes. The NCR located between $tRNA^{Glu}$ (*trn*E) and *cox*3 contained 11 long (LRUs) and short repeat units (SRUs) (seven LRUs of 317 bp, four SRUs of 207 bp each), and an internal spacer sequence between LRU7 and SRU4 specifying high-level polymorphism. Special DHU-arm missing tRNAs for Serine were found for both $tRNA^{S1(AGN)}$ and $tRNA^{S2(UCN)}$. *Echinostoma revolutum* indicated the lowest divergence rate to *E. miyagawai* and the highest to *Tracheophilus cymbius* and *Echinochasmus japonicus*. The usage of ATG/GTG start and TAG/TAA stop codons, the AT composition bias, the negative AT-skewness, and the most for Phe/Leu/Val and the least for Arg/Asn/Asp codons were noted. Topology indicated the monophyletic position of *E. revolutum* to *E. miyagawai*. Monophyly of Echinostomatidae and Fasciolidae was clearly solved with respect to Echinochasmidae, Himasthlidae, and Cyclocoelidae which were rendered paraphyletic in the suborder Echinostomata.

Introduction

Human echinostomiasis is a global zoonotic foodborne trematodiasis caused by flukes within the *Echinostoma revolutum* group, and despite its worldwide distribution, it is a particular public health problem in South East Asia (Chai, 2009; Toledo and Esteban, 2016). *Echinostoma revolutum* (Fröhlich, 1802) Rudolphi, 1809, is a member of the family Echinostomatidae (Platyhelminthes: Echinostomata), and the '*E. revolutum*' group is characterized by the '37-collar-spines' found on the cercariae (Kostadinova, 2005; Georgieva *et al.*, 2014). There are nine *Echinostoma* species within the *E. revolutum* group including *Echinostoma caproni*, *Echinostoma echinatum*, *Echinostoma friedi*, *Echinostoma jurini*, *Echinostoma miyagawai*, *Echinostoma paraensei*, *Echinostoma parvocirrus*, *E. revolutum* and *Echinostoma trivolvis*; while in other Echinostomatidae species the number of collar spines may vary, such as 25–29 on *Echinostoma hortense*, 43 on *Echinostoma malayanum*, 41–45 on *Hypoderaeum conoideum* and 43–50 on *Echinoparyphium recurvatum* (Chai, 2009; Saijuntha *et al.*, 2011*a*). The similarity of these species within the *E. revolutum* complex usually required additional identification approaches for their discrimination, mostly enzymatic and molecular techniques (Saijuntha *et al.*, 2011*a*, 2011*c*; Georgieva *et al.*, 2014; Tkach *et al.*, 2016).

The taxonomic status of *E. revolutum* is still controversial although recently a number of molecular studies have identified the parasite to be a highly cosmopolitan species comprising of several distinct geographical lineages corresponding to parasite populations with European, American, and Southeast Asian origins (Saijuntha *et al.*, 2011*a*; Georgieva *et al.*, 2014; Faltýnková *et al.*, 2015; Nagataki *et al.*, 2015). The taxonomic identification and the

CAMBRIDGE UNIVERSITY PRESS

phylogenetic assessment of each species within the '*E. revolutum*' group and as well between member taxa in the family Echinostomatidae require accurate genomic data. Many attempts of interspecific clarification for the echinostomatids, particularly for those within the '37-collar-spined' taxa have relied predominantly on tenuous morphological features (Georgieva *et al.*, 2014; Faltýnková *et al.*, 2015; Nagataki *et al.*, 2015; Tkach *et al.*, 2016). However, by using single 28S ribosomal DNA, limited short mitochondrial DNA sequences (mtDNA) or a combination of both, new cryptic echinostome species and the systematic relationships within and between members within the Echinostomatidae have been revealed as well as their association with the other families in the superfamily Echinostomatoidea (Platyhelminthes: Echinostomata) (Olson *et al.*, 2003; Georgieva *et al.*, 2013, 2014; Nagataki *et al.*, 2015; Tkach *et al.*, 2016). However, in order to provide a detailed account of current species and to taxonomically validate echinostomes more effectively, it has been argued that genomic analyses could provide insights into the fine scale inter-relationships between echinostome species (Detwiler *et al.*, 2010; Faltýnková *et al.*, 2015; Gordy and Hanington, 2019). In fact, the analyses of complete mitochondrial genomes to perform taxonomic and phylogenetic analyses of other members of the Echinostomata, as well as other trematode species, have been widely used and have provided not only a deeper understanding of the evolutionary relationships within and between trematode families but have also provided essential molecular markers for population genetics and diagnostics, crucial for modern epidemiological studies (Wey-Fabrizius *et al.*, 2013; Georgieva *et al.*, 2014; Faltýnková *et al.*, 2015).

However, many morphologically similar species, and particularly, for those of the 'collar-spined' *Echinostoma* spp. in the Echinostomatoidea lack complete mitochondrial genomic data. Currently, only four of the nine species of the '*E. revolutum*' group, including *E. caproni*, *E. paraensei*, *E. miyagawai*, *E. hortense* (Saijuntha *et al.*, 2011*c*), and a few species within the Echinostomata suborder have complete mitochondrial genomes available (Yang *et al.*, 2015; Fu *et al.*, 2019; Suleman *et al.*, 2019; Li *et al.*, 2019*a*, 2019*b*).

This current study determined the complete mitochondrial genome sequence of *E. revolutum* and correlatively characterized its genomic features and compared them with those previously reported in the superfamily Echinostomatoidea. A phylogeny for members of families in the suborders Echinostomata, Opisthorchiata, Troglotremata, Pronocephalata, and Xiphidiata is provided.

Materials and methods

Samples, DNA extraction and species identification

Adult *E. revolutum* flukes were obtained from the intestines of the naturally infected domestic ducks from abattoirs in Khon Kaen province, Thailand. The flukes were thoroughly washed in physiological saline and morphologically identified based on the size of the body and circumoral disc, the appearance of testes and the presence of '37-collar spines' around the head (Miliotis and Bier, 2003; Georgieva *et al.*, 2014). The worms were individually fixed in 70% (v/v) ethanol and stored at −20°C until use. Subsequently, species were confirmed by molecular phylogenetic analyses using nuclear ITS-1, mitochondrial *cox*1 and *nad*1 markers (Saijuntha *et al.*, 2011*a*, 2011*b*; Nagataki *et al.*, 2015).

Total genomic DNA was extracted from individual worms using the DNA extraction kit (QIAGEN, Hilden, Germany) following the manufacturer's protocol. The *E. revolutum*-species used for mitochondrial sequencing in this study belonged to the *nad*1-based *E. revolutum*-Eurasian lineage (Nagataki *et al.*, 2015).

PCR strategies for obtaining the complete mitochondrial genome

The first, initial specific primer pairs (ERE1F/ERE2R; ERE3F/ERE4R; ERE5F/ERE6R) designed based on the conserved nucleotide sequences aligned by those *E. revolutum*-mt sequences, *cox*1, *nad*1, *rrn*S (12S), respectively, available in GenBank and others, namely platyhelminth-universal primers (TRECOBF; TRECOBR; GLYF; GLYR) previously described in Le *et al.* (2019) were used. They were paired to bind on the target regions for amplification of long PCR of 4.0–7.5 kb or short of <4.0 kb overlapping fragments. The sequence data obtained were used to design further *E. revolutum*-specific primers (Table 1).

All reagents and kits used in this study were from Thermo Fisher Scientific Inc. (Waltham, MA, USA), including Phusion for long, and Dream Taq PCR Master Kits for short PCRs. PCRs were prepared in 50 μL volume with the addition of DMSO to 1.5%, and performed in an MJ PTC-100 Thermal Cycler. Long PCRs were conducted with initial denaturation at 98°C for 30 s, followed by 35 cycles, each consisting of denaturation step for 30 s at 98°C, annealing/extension step at 72°C for 6–8 min and final extension at 72°C for 10 min (in some cases, at 68°C). Short PCRs were started at 95°C for 5 min, followed by 35 cycles consisting of denaturation for 1 min at 94°C, annealing at 52°C for 1 min, extension at 72°C for 2–5 min and a final extension at 72°C for 7 or 10 min. A negative (no-DNA) control was included in some cases. The PCR products (5 − 10 μL of each) were examined on a 1% agarose gel, stained with ethidium bromide and visualized under UV light (Wealtec, Sparks, NV, USA). The primer-walking sequencing was applied until the complete sequence for the whole fragment, and the overlapping assembly was used to complete the mitochondrial genome.

Characterization of mitogenomic features

Protein-encoding genes (PCGs) were identified by alignment with the available mt genomes of other *Echinostoma* trematode species and ATG/GTG as start and TAA/TAG as stop codons were used to define gene boundaries. PCGs were translated using the echinoderm/flatworm mitochondrial genetic code: translation Table 9 in GenBank. Nucleotide and codon composition were analysed with MEGA 7.0 (Kumar *et al.*, 2016) and codon usage for all PCGs was determined with the program GENE INFINITY (Codon Usage: http://www.geneinfinity.org/sms/sms_codonusage.html). Nucleotide percentage (%) for comparison of individual/concatenated PCGs and mitochondrial ribosomal genes (MRGs) between *E. revolutum* and 14 representative members of the superfamily Echinostomatoidea (Table 2) was determined by using GENEDOC 2.7 for alignment, Gblocks 0.91b (Castresana, 2000) (online accession at http://molevol.cmima.csic.es/castresana/Gblocks_server.html) for picking the best quality block (10,112 bp) and MEGA 7.0 for percentage estimation.

The transfer RNA genes (tRNA or *trn*) were identified using tRNAscan-SE 1.21 program (www.genetics.wustl.edu/eddy/tRNAscan-SE/) (Lowe and Eddy, 1997); ARWEN at http://mbio-serv2.mbioekol.lu.se/ARWEN/ (Laslett and Canback, 2008) for finding the best final tRNA sequences and secondary structures. Any tRNAs not detected by these programs were found by inspection of the sequences, based on the alignment with sequences of other trematode and by their potential formation of tRNA configuration. The ribosomal 16S (*rrn*L) and 12S (*rrn*S) RNA genes were recognized as described in Le *et al.* (2019) in the region located between $tRNA^{Thr}$ (*trn*T) and *cox*2 separated by $tRNA^{Cys}$ (*trn*C), respectively.

The nucleotide composition, AT and GC content for concatenated 12 PCGs (not excluding the overlapped sequences between

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

Table1. Primers for amplification and sequencing fragments of the mitochondrial genome of *Echinostoma revolutum*

Primer	Sequence (5′ to 3′)	Location
ERE1F	GGTCTTATTCTKGCTATGGCTGC	*cox*1
ERE2R	AGCCGACTACGAGTTCCAC	*cox*1
ERE3F	TGCTTAGTTGTGTTCGTTCTGC	*nad*1
ERE4R	CCTAAGACCACACAATAACCGC	*nad*1
ERE5F	CTATGTGCTGCTGATGTTGGG	*rrn*S
ERE6R	GATGCTGGCACTGTGTATCC	*rrn*S
ERE7F	TTTCAGCCCATGTTTGTTTAGC	*cyt*b
ERE8R	ACAAAGAGGGGATTGTTTGAACC	*cyt*b
ERE9F	ATCTGGTTTTGGGTTTCGGG	*nad*5
ERE10R	AACCAAAGCCGCAAAAGAGG	*nad*5
ERE11F	AGATGCTATACCCGGACGTC	*cox*2
ERE12R	ACCACCTCACACACCAATCA	*cox*1
ERE13R	CACAAAGAGTGGCAAGCTCC	*nad*2
ERE16F	AGAATTTTGGCTTGTCGTGCC	*trn*D
ERE17R	CTAACACCCCCTATAAACCCAG	*nad*4
ERE18R	ACTCTGATGTTGGGGTGTTGG	*cox*1
ERE19F	GTGTGGTTTCATTTTATCGTTGGGAGG	*nad*5
ERE20R	CAACCCAAGCTTTATACATAGGCAACC	*cox*3
ERE21R	AGGAACAACAAACTCCTCCTC	*cox*3
ECH3F	ATGAKTTGRTTGCCWATRTATAAAGC	*cox*3
ERE22F	AATGGGCAATTAAATTTGATGTGG	NCR
ERE23R	CATTGCCATACAGCAAATGCCAATC	NCR

NCR, non-coding region.

*nad*4L and *nad*4), two MRGs and complete mt genome for 15 members of the Echinostomatoidea were determined by MEGA7.0, and the AT and GC skewness values (from −1 to +1) calculated according to the formula by Perna and Kocher (1995) [AT skew = (A + T)/(A–T) and GC skew = (G + C)/(G–C)].

The non-coding region (NCR) was determined by the recognition of boundaries between $tRNA^{Glu}$ (*trn*E) and *cox*3. Tandem Repeat Finder v3.01 (Benson, 1999) was used to detect repeat units (RUs) in the NCR of mitogenome of *E. revolutum* in this study and other *Echinostoma* spp. and digeneans which were not available in GenBank or not previously reported.

Phylogenetic analyses

Concatenated amino acid sequences of the 12 PCGs of *E. revolutum* and 44 species from 13 families [i.e., Echinostomatidae, Fasciolidae, Himasthlidae, Echinochasmidae, Cyclocoelidae, Paramphistomidae, Gastrothylacidae, Notocotylidae, Troglotrematidae/(Paragonimidae), Heterophyidae, Opisthorchiidae, Diclocoeliidae, and Schistosomatidae] in the superfamilies of Echinostomatoidea, Paramphistomoidea, Pronocephaloidea, Troglotrematoidea, Opisthorchioidea, and Gorgoderoidea were aligned for phylogenetic analysis. The sequence of *Schistosoma haematobium* (Digenea: Schistosomatidae) was chosen as an outgroup (Littlewood *et al.*, 2006). The alignment was constructed by GENEDOC2.7, confirmed by MAFFT 7.122 (Katoh and Standley, 2013) and finalized by Gblocks 0.91b. The final alignment block of 2,993–3,025 amino acids without poorly aligned regions was picked out for phylogenetic analysis. Tree was constructed using maximum likelihood by MEGA 7.0 with a bootstrap of 1000 replications. The substitution model with the best score according to the Bayesian information criterion was the Jones, Taylor and Thornton +F+G+I model, with residue frequencies estimated from the data (+F), rate variation along the length of the alignment (+G) and allowing for a proportion of invariant sites (+I).

Results

Gene organization and genomic features

The complete mitochondrial genome of *E. revolutum* was shown to be 17,030 bp in size (GenBank accession no. MN496162) (Fig. 1). As common in other trematodes, the *E. revolutum* mitogenome has one-direction transcription, similar gene organization and content with the exception of African *Schistosoma* spp. It comprises of 12 protein coding genes (*atp*6, *cox*1-3, *cyt*b, *nad*1-6, *nad*4L), two ribosomal RNA (*rrn*L and *rrn*S) and 22 transfer RNA genes (tRNA or *trn*) similar to those of common digeneans (Table 2).

Echinostoma revolutum has typical mt structural features of the platyhelminths and does not contain *atp*8 and has the overlapped region between *nad*4L and *nad*4 genes by 40 bp (Table 2). Five protein-coding genes used GTG (*nad*4L, *nad*2, *nad*1, *cox*1, *nad*5) and other seven used ATG as start codons; and seven genes used TAG and five used TAA for termination. Boundaries between *cyt*b and *nad*4L, between $tRNA^{Asp}$ and *nad*1, from $tRNA^{Thr}$ to *rrn*S *(12S)*, covering *rrn*L *(16S)*, $tRNA^{Cys}$ genes, and between repeats in the NCR are continuous whilst there are large intergenic spacers of 33 or 30 bp between other genes (*cox*1 and $tRNA^{Th}$; and $tRNA^{Val}$ and $tRNA^{Ala}$), respectively.

The mt genome of *E. revolutum* encodes 22 transfer RNAs, ranging from 60 ($tRNA^{S1(AGN)}$) to 71 nucleotides ($tRNA^{His}$). Twenty have common 'cloverleaf' folding into secondary structures with the complete four arms but two for Serine, $tRNA^{S1(AGN)}$ and $tRNA^{S2(UCN)}$, possess special forms missing DHU-arms (Table 2; SFig. 1). Two ribosomal RNA genes, *rrn*L (977 bp) and *rrn*S (756 bp long), are located between the $tRNA^{Thr}$ and *cox*2, separated by $tRNA^{Cys}$. The order of the mitochondrial DNA block of [*cox*1-$tRNA^{Thr}$-*rrn*L-$tRNA^{Cys}$-*rrn*S-*cox*2-*nad*6] is highly conserved in all the trematodes, including *E. miyagawai*, *Ech. japonicus*, *Fas. magna*, *F. hepatica*, *F. gigantica,* and Asian *Schistosoma* spp. (Le *et al.*, 2001, 2002, 2016; Liu *et al.*, 2014; Ma *et al.*, 2016; Fu *et al.*, 2019; Li *et al.*, 2019*b*).

Base composition and comparative analyses

The base composition was A (18.81%), T (47.40%), G (23.50) and C (10.29% in the mt genome of *E. revolutum* and the A + T content was 62.21% for PCGs and their skewness values were −0.46 for A + T and 0.391 for G + C, respectively). MRGs showed a similar percentage of overall A + T (62.73%) and G + C (37.27%) but their skewness values were considerably different (−0.179/A + T; and 0.275/G + C) due to the bias use of A over T in PCGs than in MRGs (Table 3).

The divergence rate (%) inferred from the nucleotide pairwise comparison of 12 individual mitochondrial protein-coding and two ribosomal genes between *E. revolutum* and 14 members of Echinostomatoidea indicated that the rate was the lowest level of divergence between *E. revolutum* and *E. miyagawai* (8.99%/*nad*4L–18.4%/*nad*4; 6.63%/*rrn*S–8.93%/*rrn*L), and in average, 14.89%/PCGs for protein-coding genes and 8.29%/MRGs for ribosomal genes, respectively.

The highest nucleotide sequence divergence between *E. revolutum* and Echinostomatoidea trematodes was 39.5% in comparison with *Tracheophilus cymbius* (Cyclocoelidae) and 38.16% for *Ech. japonicus* (Echinochasmidae) for PCGs (Table 4). Overall, the nucleotide sequence of *E. revolutum* in each gene differed from

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

Table 2. Locations of genes and other features in the complete mitochondrial genome of *Echinostoma revolutum* (17,030 bp) (GenBank: MN496162)

Gene	Position (5′>3′)	Characteristics [bp/aa(start/stop)] and regions	tRNA anti-codon	Int. seq. length (bp)
*cox*3	1–645	645/214/(ATG/TAA)		+3
$tRNA^{His}$	649–719	71	GTG	+2
*Cyt*b	722–1831	1110/369/(ATG/TAG)		0
*nad*4L	1832–2104	273/90/(GTG/TAA)		−40
*nad*4	2065–3348	1284/427/(ATG/TAA)		+4
$tRNA^{Gln}$	3353–3415	63	TTG	+12
$tRNA^{Phe}$	3428–3491	64	GAA	+26
$tRNA^{Met}$	3518–3583	66	CAT	+3
*atp*6	3587–4105	519/172/(ATG/TAA)		+12
*nad*2	4118–4987	870/289/(GTG/TAG)		+6
$tRNA^{Val}$	4994–5056	63	TAC	+30
$tRNA^{Ala}$	5087–5153	67	TGC	+1
$tRNA^{Asp}$	5155–5220	65	GTC	0
*nad*1	5221–6129	909/302/(GTG/TAG)		+13
$tRNA^{Asn}$	6143–6209	67	GTT	+4
$tRNA^{Pro}$	6214–6280	67	TGG	+1
$ttRNA^{Ile}$	6282–6343	62	GAT	+14
$tRNA^{Lys}$	6358–6425	68	CTT	+4
*nad*3	6430–6786	357/118/(ATG/TAG)		+2
$tRNA^{Ser1(AGN)a}$	6789–6848	60	GCT	+7
$tRNA^{Trp}$	6856–6921	66	TCA	+3
*cox*1	6925–8463	1539/512/(GTG/TAG)		+33
$tRNA^{Thr}$	8497–8562	66	TGT	0
rrnL (16S)	8563–9539	977		0
$tRNA^{Cys}$	9540–9605	66	GCA	0
*rrn*S (12S)	9606–10 359	756		0
*cox*2	10 360–10 968	609/201/(ATG/TAA)		+11
*nad*6	10 980–11 432	453/150/(ATG/TAG)		+3
$tRNA^{Tyr}$	11 433–11 497	65	GTA	+11
$tRNA^{Leu1(CUN)}$	11 498–11 561	64	TAG	−2
$tRNA^{Ser2(UCN)a}$	11 560–11 624	65	TGA	+10
$tRNA^{Leu2(UUR)}$	11 635–11 697	63	TAA	−2
$tRNA^{Arg}$	11 696–11 762	67	TCG	−2
*nad*5	11 761–13 326	1566/521/(GTG/TAG)		+12
$tRNA^{Gly}$	13 339–13 405	67	TCC	+11
$tRNA^{Glu}$	13 417–13 481	65	TTC	+7
Repeat units	13 489–16 912			
LRU1	13 489–13 805	317		0
LRU2	13 806–14 122	317		0
LRU3	14 123–14 439	317		0
LRU4	14 440–14 756	317		0
LRU5	14 757–15 073	317		0
LRU6	15 074–15 390	317		0
LRU7	15 391–15 707	317		0
Int. Spacer	15 708–16 341	377		0
IntS-half 1	15 708–15 895	188		0

(*Continued*)

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

Table 2. (*Continued.*)

Gene	Position (5′>3′)	Characteristics [bp/aa(start/stop)] and regions	tRNA anti-codon	Int. seq. length (bp)
IntS-half 2	15 896–16 084	189		0
SRU4	16 085–16 291	207		0
SRU3	16 292–16 498	207		0
SRU2	16 499–16 705	207		0
SRU1	16 706–16 912	207		0
unique seq	16 913–17 030	130		0

bp, base pair; aa, amino acid; start, start codon; stop, stop codon; Int. seq., intergenic sequence (+, number of nucleotides before the start of the following gene; −, number of nucleotides overlapping with the following gene); LRU, long repeat unit; SRU, short repeat unit; IntS, internal spacer sequence between LRU7 and SRU4; unique seq, nucleotide sequence between SRU1 and *cox*3.
[a]tRNAs lacking DHU-arm.

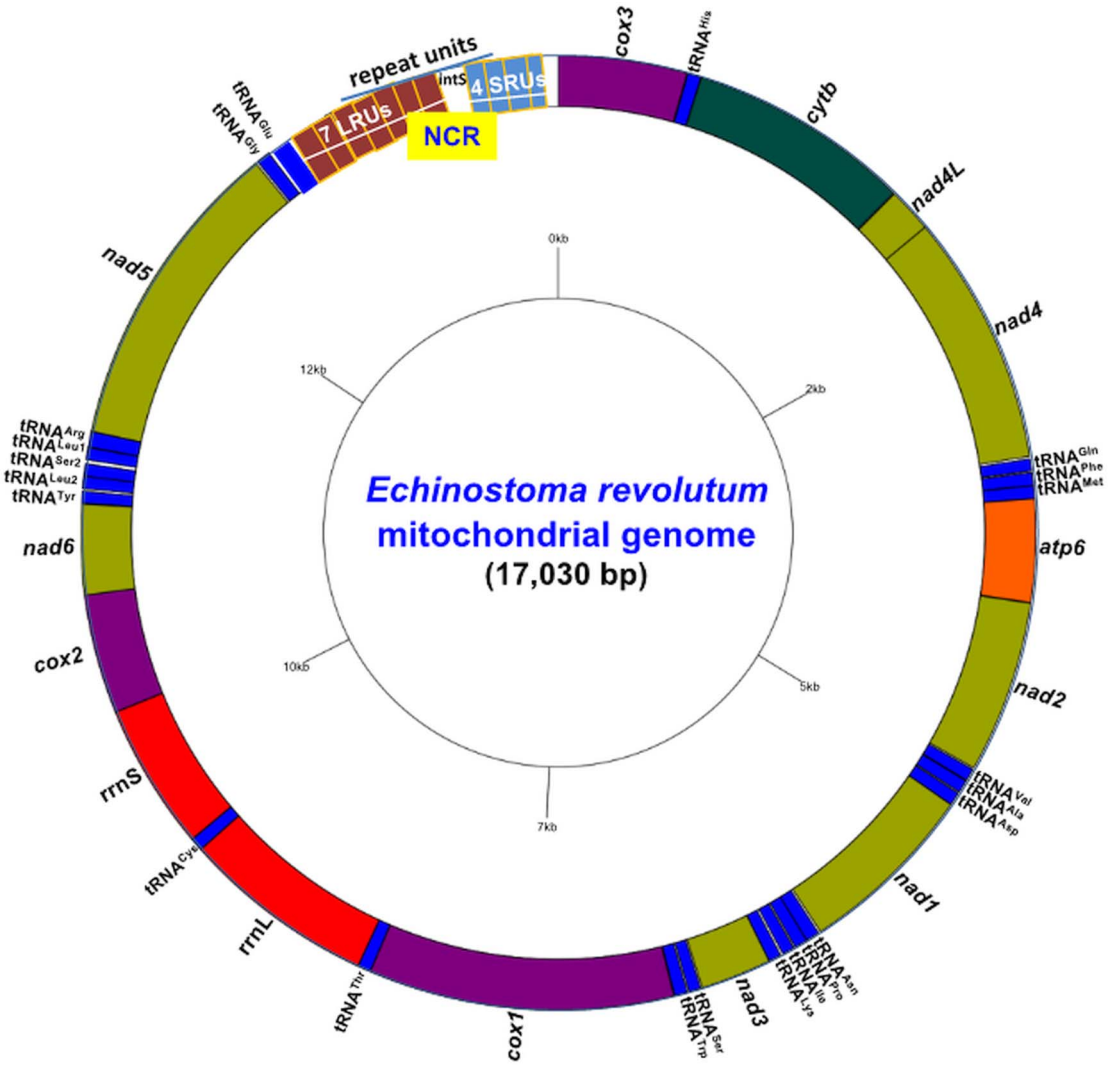

Fig. 1. A schematic drawing of a circular map of the mitochondrial genome of *Echinostoma revolutum* (GenBank: MN496162). Protein-coding and ribosomal large and small subunit genes are abbreviated according to our previous publications (Le *et al*., 2016, 2019). The transfer RNA genes (tRNA) are marked with three letter-amino acid abbreviations (see: Table 2). The non-coding region (NCR) located between $tRNA^{Glu}$ and *cox*3 consists of seven long (LRU1–7), four short repeat units (SRU1–4), and internal spacer sequence (IntS) between LRU7 and SRU4.

6.63%/*rrn*S/(*E. miyagawai*) to 59.89%/*nad*5/(*T. cymbius*). Within the Echinostomatidae, the interspecific variation does not exceed 37%, as seen between *atp*6 genes of *E. revolutum* and *H. conoideum*.

The codon usage in mtDNAs of all the Echinostomatidae trematodes (*E. revolutum*; *E. caproni*; *E. miyagawai*; *E. paraensei*; *Echinostoma* sp. JM-2019; *A. sufrartyfex*; *H. conoideum*) is biased to the use of TTT (for Phenylanine), TTG (for Leucine) and GTT (for Valine). Multiple Thymine (T) in use in these codons facilitates the mostly used frequency (from 5.96% GTT/Val in *H. conoideum* to 10.65% TTT/Phe in *E. caproni*). The least frequently used codons, comprising mostly G and C, are CGC (for Arginine), AAC (for Asparagine) and GAC (for Aspartic acid), ranging from one to two (0.03–0.06%) to six to seven (0.18–

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

Table 3. Base composition and skewness value for the mitochondrial protein-coding (PCGs) and mitoribosomal genes (MRGs) of 15 members of the superfamily Echinostomatoidea

	Species		Length (nt)	A (%)	T (%)	G (%)	C (%)	A + T (%)	AT-skew	G + C (%)	GC-skew
	Echinostomatidae										
1	*Echinostoma revolutum*	PCGs	10 134	18.81	47.40	23.50	10.29	62.21	−0.460	37.79	0.391
		MRGs	1733	25.74	36.99	23.77	13.50	62.73	−0.179	37.27	0.275
2	*Artyfechinostomum sufrartyfex*	PCGs	10 131	16.99	46.21	26.53	10.27	63.20	−0.462	36.80	0.442
		MRGs	1728	24.71	3.709	25.58	12.62	61.80	−0.20	38.20	0.339
3	*Echinostoma caproni*	PCGs	10 128	17.34	47.82	24.79	10.05	65.16	−0.468	34.84	0.423
		MRGs	1709	25.34	36.63	24.40	13.63	61.97	−0.182	38.03	0.283
4	*Echinostoma miyagawai*	PCGs	10 128	18.20	47.65	24.07	10.08	65.85	−0.447	34.15	0.410
		MRGs	1724	25.75	37.94	23.49	12.82	63.72	−0.191	36.31	0.294
5	*Echinostoma paraensei*	PCGs	10 128	18.04	47.57	24.13	10.26	65.61	−0.450	34.39	0.403
		MRGs	1748	25.92	37.76	23.68	12.64	63.68	−0.186	36.32	0.304
6	*Echinostoma* sp. JM-2019	PCGs	10 122	16.47	46.46	26.66	10.40	62.93	−0.477	37.60	0.432
		MRGs	1726	24.51	35.17	26.94	13..38	59.68	−0.179	40.32	0.336
7	*Hypoderaeum conoideum*	PCGs	10 116	16.84	45.25	26.96	10.95	62.09	−0.458	37.91	0.422
		MRGs	1730	25.14	34.68	26.59	13.58	59.82	−0.159	40.17	0.324
	Echinochasmidae										
8	*Echinochasmus japonicus*	PCGs	10 143	15.66	46.55	28.05	9.74	62.21	−0.497	37.79	0.485
		MRGs	1748	22.83	35.58	28.60	12.99	58.41	−0.218	41.65	0.375
	Fasciolidae										
9	*Fasciola hepatica*	PCGs	10 104	14.29	49.24	26.90	9.57	63.53	−0.550	36.47	0.475
		MRGs	1755	23.10	39.25	26.13	11.52	62.35	−0.259	37.65	0.388
10	*Fasciola gigantica*	PCGs	10 107	13.53	50.44	27.28	8.76	63.97	−0.577	36.04	0.514
		MRGs	1755	21.29	40.18	27.32	11.21	61.47	−0.307	38.53	0.418
11	*Fascioloides jacksoni*	PCGs	10 137	14.71	46.97	29.09	9.23	61.68	−0.523	38.52	0.516
		MRGs	1743	24.38	38.50	26.56	10.56	62.88	−0.225	37.12	0.431
12	*Fascioloides magna*	PCGs	10 131	15.41	46.59	28.29	9.71	62.0	−0.503	38.00	0.489
		MRGs	1751	23.24	38.26	26.50	11.99	61.50	−0.224	38.49	0.377
13	*Fasciolopsis buski*	PCGs	10 122	16.62	48.97	25.81	8.61	65.59	−0.493	34.42	0.500
		MRGs	1768	24.38	40.27	24.43	10.92	64.65	−0.246	35.35	0.382
	Himasthlidae										
14	*Acanthoparyphium* sp.WAK-2018	PCGs	10 119	17.20	44.93	26.83	11.05	62.13	−0.446	37.87	0.417
		MRGs	1753	24.70	36.74	25.90	12.66	61.44	−0.196	38.56	0.343
	Cyclocoelidae										
15	*Tracheophilus cymbius*	PCGs	10 152	14.83	48.78	27.17	09.22	63.61	−0.534	36.39	0.493
		MRGs	1745	22.58	38.34	27.05	12.03	60.92	−0.411	39.08	0.384

0.21%) were noted. Clear bias was seen to the use of TAG (7–12 codons) for termination of 12 PCGs rather than TAA (0–5) in mt PCG genes of all the eight echinostomids (STable 1).

Polymorphism featured by repeat units in non-coding regions of Echinostoma spp.

The NCR of *E. revolutum* was identified by recognition of boundary of $tRNA^{Glu}$ (*trn*E) and *cox*3 gene, which is of 3,549 bp in length, perhaps the longest in the mt genomes of the echinostomid flatworm sever fully sequenced to date (Tables 2 and 5). The NCR of this species possesses seven long, identical RUs (LRU1 to LRU7, 317 bp each) and four short, identical RUs (SRU1 to SRU4, 207 bp each) tandemly arranged after each other (Tables 2 and 5; Fig. 1; GenBank: MN496162). Between LRU7 and SRU4, there is a linking region of an internal spacer sequence of 377 bp which contained 188 bp, partial of LRU (designated as IntS-half1) and 189 bp, partial of SRU (IntS-half2). A unique sequence region of 130 nucleotides continuously occurs between SRU1 and *cox*3 (Table 2; Fig. 1).

Tandem RUs were also found in *E. miyagawai* (two RUs, 319 bp each), in *E. paraensei* (at least three RUs, 206 bp each in the

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

Table 4. Nucleotide comparison for divergence rate (%) of individual and concatenated protein-coding (PCGs) and mitoribosomal genes (MRGs) between *Echinostoma revolutum* and 14 representative members of the superfamily Echinostomatoidea (Platyhelminthes: Echinostomata)

		Echinostoma revolutum															
		Protein-coding genes													Mitoribosomal genes		
Species		*atp*6	*cox*1	*cox*2	*cox*3	*cyt*b	*nad*1	*nad*2	*nad*3	*nad*4L	*nad*4	*nad*5	*nad*6	PCGs	*rrn*L	*rrn*S	MRGs
1	Ecap	19.85	15.76	14.72	14.67	15.15	15.67	18.19	13.50	16.90	19.91	19.77	20.03	17.20	12.19	6.93	10.53
2	Emiy	16.98	12.62	11.73	13.29	12.61	12.84	16.72	15.54	08.99	18.40	17.25	19.26	14.89	08.93	6.63	08.29
3	Epar	21.16	15.46	14.27	14.64	13.87	14.54	17.81	14.82	18.00	20.53	17.19	17.33	16.60	09.56	7.16	08.84
4	EJM	31.37	19.24	18.60	22.62	19.29	22.59	29.17	23.17	28.86	31.75	28.57	31.37	24.68	21.83	15.48	19.53
5	Asuf	31.60	21.75	20.65	25.57	20.53	24.27	33.03	23.39	22.10	35.33	32.66	34.36	26.74	22.82	19.15	21.47
6	Hcon	37.00	23.52	33.68	27.90	22.07	23.70	34.22	28.22	22.17	33.19	32.65	31.77	28.67	22.17	18.55	22.33
7	Ejap	43.08	30.63	49.45	48.80	28.25	30.81	48.05	32.81	35.67	45.08	41.84	41.63	38.16	30.43	34.37	32.37
8	Fhep	39.80	25.38	39.90	42.21	26.22	23.07	44.71	28.15	33.47	44.46	41.92	42.33	34.54	31.31	28.10	31.23
9	Fgig	37.35	24.17	40.36	40.31	26.54	24.17	44.46	31.39	36.92	44.49	38.49	43.07	33.92	31.86	25.95	30.77
10	Fjac	40.08	27.37	41.89	46.65	29.02	26.48	43.98	30.98	41.72	43.08	42.25	45.72	36.49	31.59	30.87	31.66
11	Fmag	39.89	25.04	42.86	46.52	28.60	26.55	42.06	31.67	41.76	42.07	41.94	53.90	35.85	33.27	26.49	30.31
12	Fbus	37.25	26.61	37.69	40.36	25.21	25.52	43.87	30.91	35.50	41.61	43.66	43.34	34.96	31.94	27.90	31.50
13	AWAK	38.07	29.92	41.64	41.46	27.07	32.50	47.07	33.53	30.57	46.85	45.17	38.88	37.43	32.78	26.99	31.32
14	Tcym	48.98	26.81	41.49	43.27	28.08	30.74	44.84	30.56	33.15	52.29	59.89	47.32	39.50	30.69	28.28	31.72

Ecap: *Echinostoma caproni*; Emiy: *E. miyagawai*; Epar: *E. paraensei*; EJM: *Echinostoma* sp. JM-2019; Asuf: *Artyfechinostomum sufrartyfex*; Hcon: *Hypoderaeum conoideum*; Ejap: *Echinochasmus japonicus*; *Fhep*: *Fasciola hepatica*; *Fgig*: *F. gigantica*; Fjac: *Fascioloides jacksoni*; Fmag: *Fas. magna*; Fbus: *Fasciolopsis buski*; AWAK: *Acanthoparyphium* sp. WAK-2018; Tcym: *Tracheophilus cymbius*. Rows of data subjected to discussion in the text were background colour shaded, including the lowest divergence between *E. revolutum* and *E. miyagawa* (family: Echinostomatidae) and the highest rate between *E. revolutum* and *Ech. japonicus* (Echinochasmidae), and between *E. revolutum* and *T. cymbius* (Cyclocoelidae).

Table 5. Number and type of the repetitive sequences in non-coding regions (NCR) of 15 representative members of the superfamily Echinostomatoidea indicating high polymorphism and interspecific/intergeneric variation

	Species	Length of NCR (bp)	Number and size of repeat units (RU)	Type of repeat units	Accession No	References
1	*Echinostoma revolutum*	3,549	7 LRUs (317 bp/each) 4 SRUs (207 bp/each)	Tandem repeat family	MN496162	This study
2	*Echinostoma caproni*	685	none	none	AP017706	GenBank
3	*Echinostoma miyagawai*	982	2 RUs (319 bp/each)	Tandem repeat family	MN116740	Fu *et al.* (2019)
4	*Echinostoma paraensei*[a]	6,798	3 RUs (206 bp/each)	Tandem repeat family	KT008005	GenBank
5	*Echinostoma* sp. JM-2019	1,877	5 LRUs (245 bp/each) 2 SRUs (166 bp/each)	Tandem repeat family	MH212284	GenBank
6	*Artyfechinostomum sufrartyfex*	1,004	2 RUs (144 bp/each)	Tandem repeat family	KY548763	GenBank
7	*Hypoderaeum conoideum*	654	none	none	KM111525	Yang *et al.* (2015)
8	*Echinochasmus japonicus*	2,001	8 RUs (240 bp/each)	Tandem repeat family	KP844722	Le *et al.* (2016)
9	*Fasciola hepatica*	817	9 RUs (85 bp/each)	Tandem repeat family	AF216697	Le *et al.* (2001)
10	*Fasciola gigantica*	841	8 RUs (86 bp/each)	Tandem repeat family	KF543342	Liu *et al.* (2014)
11	*Fascioloides jacksoni*	1,517	9 RUs (113 bp/each)	Tandem repeat family	KX787886	GenBank
12	*Fascioloides magna*	625	5 RUs (60 bp/each)	Tandem repeat family	KU060148	Ma *et al.* (2016)
13	*Fasciolopsis buski*	1,314	8 RUs (104 bp/each)	Tandem repeat family	KX169163	Ma *et al.* (2017)
14	*Acanthoparyphium* sp. WAK-2018	723	3 RUs (57 bp/each)	Tandem repeat family	MG792058	GenBank
15	*Tracheophilus cymbius*	142	none	none	MK355447	Li *et al.* (2019*a*)

[a]Non-coding region in *E. paraensei* not fully sequenced.

partially sequenced NCR), in *Echinostoma* sp. JM-2019 (five LRUs, 245 bp each and two SRUs, 166 bp each), in *A. sufrartyfex* (two RUs, 144 bp each) which is variable in numbers and length (Table 5). The size of the mt genome differed among echinostomes and digeneans; this is due to the variable length of their NCRs rich in multiple RUs (Table 3).

Phylogenetic analysis

The topology of the phylogenetic tree of taxonomic relationship indicated clear positions of five suborders, including Echinostomata, Pronocephalata, Troglotremata, Opisthorchiata, and Xiphidiata where *E. revolutum*, grouped in a monophyletic subclade as a sister taxa to *E. miyagawai* and paraphyletic to the other echinostomatids in the Echinostomatidae (Fig. 2). Monophyly of Echinostomatidae and Fasciolidae clearly resolved with respect to Echinochasmidae, Himasthlidae, and Cyclocoelidae; these were rendered paraphyletic in the suborder Echinostomata (Fig. 2). The high nodal bootstrap values well supported clear taxonomic relationships of the '*E. revolutum*' group in the Echinostomatoidea and this seemed to be in the paraphyletic position with all the other superfamilies, Pronocephalata, Troglotremata, Opisthorchiata, and Xiphidiata in the digenean order Plagiorchiida.

Discussion

The complete mitochondrial genome of *E. revolutum* (Fröhlich, 1802) Rudolphi, 1809, was 17,030 bp in size; the longest of all the Echinostomatoidea to date sequenced, although the mitogenome of *E. paraensei* (KT008005) was claimed longer, 20,298 bp, but some of 5,600 nucleotides were only of estimation (Liu *et al.*, 2014; Yang *et al.*, 2015; Le *et al.*, 2016; Ma *et al.*, 2016, 2017; Fu *et al.*, 2019; Suleman *et al.*, 2019; Li *et al.*, 2019*a*, 2019*b*).

The length of the mt genome of *E. revolutum* seemed to be one of the longest among trematodes fully obtained to date, shorter than the estimated, partially sequenced congener *E. paraensei*, but was slightly longer than other echinostomids, including *H. conoideum* (Yang *et al.*, 2015), *E. miyagawai* (Fu *et al.*, 2019; Li *et al.*, 2019*b*), *Ech. japonicus* (Le *et al.*, 2016) and two cyclocoelids, *Uvitellina* sp. and *T. cymbius* (Suleman *et al.*, 2019; Li *et al.*, 2019*a*). It was considerably longer than many fasciolids, such as *Fa. buski* (GenBank: KX169163) (Ma *et al.*, 2017), *F. gigantica* (KF543342), *F. hepatica* (AF216697), *Fasciola*/*Fascioloides jacksoni* (KX787886) and *Fas. magna* (KU060148) (Liu *et al.*, 2014; Ma *et al.*, 2016).

The tRNAs which were lacking DHU-arm for Serine in *E. revolutum* are usually found in many digenean mitogenomes, i.e. *Echinococcus granulosus*, *F. hepatica* (Le *et al.*, 2001, 2002), *Fas. magna* (Ma *et al.*, 2016), *E. miyagawai* (Li *et al.*, 2019*b*) and *Fas. jacksoni* (KX787886).

The gene organization, comparative description of genomic features with other members of Echinostomatidae, particularly, with *E. miyagawai* isolates (from Hunan and Helongjang of China) and those of the digenean Echinostomata were presented. In mtDNA sequence of *E. revolutum*, the nucleotide usage clearly biased to AT, and thus, constituting their negative skewness. Skewness values for A + T are consistent with those of *E. miyagawai*, *E. paraensei* (Echinostomatidae) and *Acanthoparyphium* sp. WAK-2018 (Himasthlidae), considerably higher than all of the members of Fasciolidae, slightly higher than other echinostomatids (*E. caproni*, *Echinostoma* sp. JM-2019, *A. sufrartyfex* and *H. conoideum*), echinochasmid (*Ech. japonicus*) but lower than the cyclocoelid *T. cymbius*. The G + C content and skewness of *E. revolutum* seemed to be of the lowest (GC skew = 0.391) among all species studied here (Table 3).

Echinostoma revolutum and *E. miyagawai* shared more common genomic features than others in the genus *Echinostoma* and family Echinostomatidae. The pattern of the usage of ATG/GTG start and TAG/TAA stop codons, the AT composition bias, the negative AT-skewness, and the most for Phe/Leu/Val and the least for Arg/Asn/Asp codons in *E. revolutum* were

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

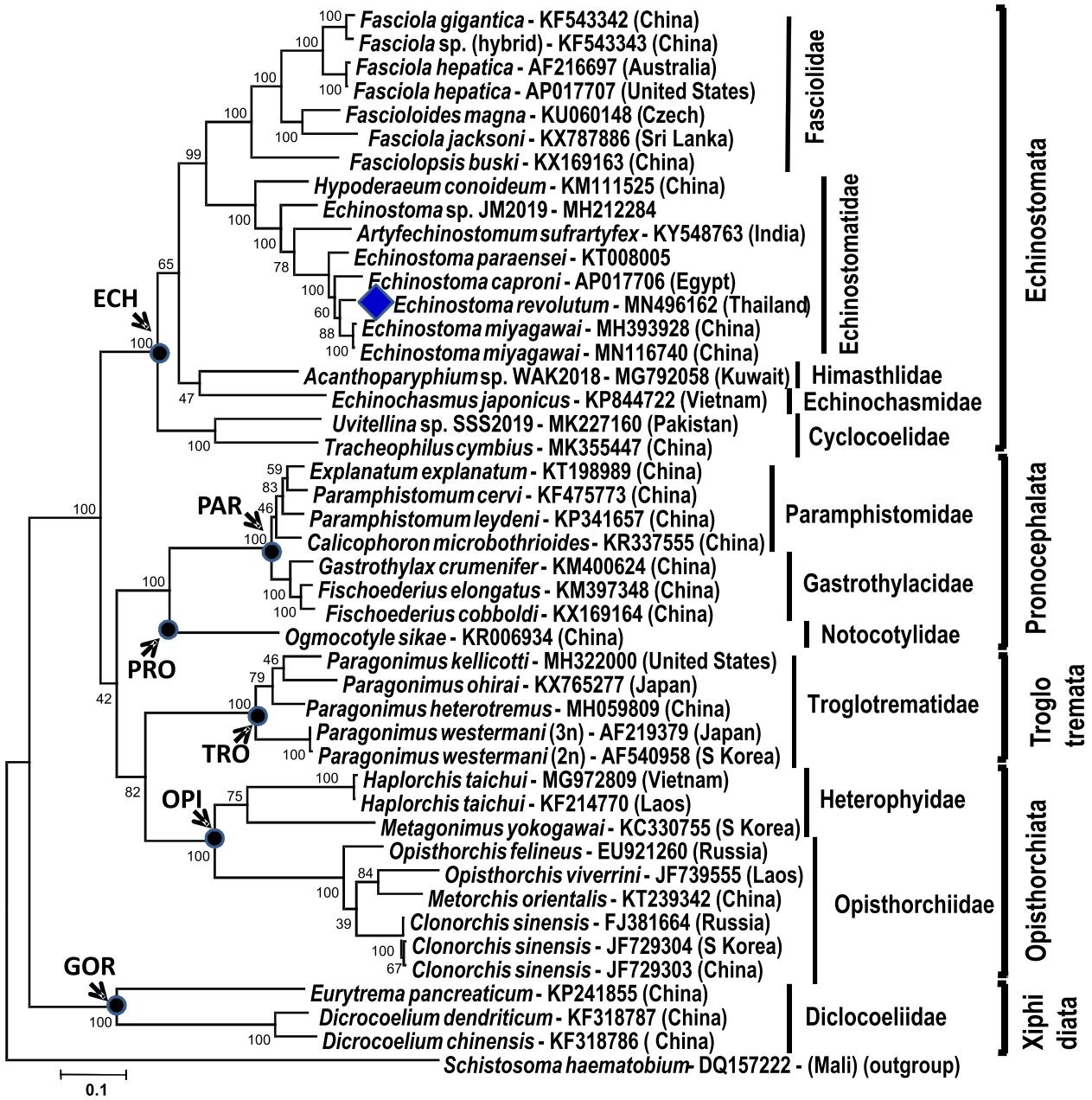

Fig. 2. A maximum likelihood phylogenetic tree showing the position of *Echinostoma revolutum* (diamond symbol) based on the analysis of concatenated amino acid sequence data for the 12 mitochondrial proteins of 45 digenean species/strains. Thirteen families [Echinostomatidae, Fasciolidae, Himasthlidae, Echinochasmidae, Cyclocoelidae, Paramphistomidae, Gastrothylacidae, Notocotylidae, Troglotrematidae/(Paragonimidae), Heterophyidae, Opisthorchiidae, Diclocoeliidae, and Schistosomatidae] belonging to six superfamilies indicated by arrows, Echinostomatoidea (ECH), Paramphistomoidea (PAR), Pronocephaloidea (PRO), Troglotrematoidea (TRO), Opisthorchioidea (OPI), and Gorgoderoidea (GOR), are represented. *Schistosoma haematobium* (Platyhelminthes: Schistosomatidae) is included as an outgroup. Nodal support values evaluated using 1000 bootstrap resamplings are shown on each branch. The scale bar represents the number of substitutions per site. Accession numbers are given for each species/strains and country name (in bracket) of their origin (where available) at the end of each sequence.

usual and similar to the members of *Echinostoma* and digenean trematodes.

The presence of 11 tandem repeats in NCR (GenBank: MN496162) made the NCR of *E. revolutum* longer and more complex relative to other echinostomatids. The repetitive sequence richness in NCR was a typical genomic feature commonly seen in a number of species, specifying high-level polymorphism in Echinostomata and other digeneans (Table 5) (Le *et al.*, 2001, 2016, 2019; Liu *et al.*, 2014; Ma *et al.*, 2016, 2017; Fu *et al.*, 2019). For some of *Echinostoma* spp. which had their complete mitogenomes fully sequenced to date, the number of RUs was fewer or absent, and the length of the NCR was less than those of *E. revolutum*.

The actual size of the mitogenomes of other echinostomids may have been an underestimation in some of the original individuals sampled as several of the repeat elements may not have been considered or incorporated in the initial analyses (Yang *et al.*, 2015; Fu *et al.*, 2019; and GenBank: MH212284; KY548763; AP017706) (Table 5), as a result of missing a part of the region containing more RUs. The missing part of the NCR may be the result of an inaccurate PCR experiment that was carried out without verification (Kinkar *et al.*, 2019; Oey *et al.*, 2019). In *E. revolutum*, the NCR was successfully amplified and accurately sequenced from a number of the verified PCR products and the RUs were confirmed to occur in the expanded NCR giving its complete mtDNA sequence as the second longest among

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press

members of the Echinostomatidae. Such repetitive regions have also occurred in the mtDNA of, for example, *E. granulosus* G1 with the addition of a 4.4 kb tandem repeat region consisting ten RUs (Kinkar *et al.*, 2019), or *Paragonimus westermani* from the Arunachal Pradesh State (India), with the full mtDNA of 20.3 kb comprising of a long repetitive region in the isolate of the East Siang district (Oey *et al.*, 2019) instead of 14,965 bp in the isolate of the Changlang District (Biswal *et al.*, 2014). However, it should be noted that the length and number of repeats are genetically variable between geographical isolates of a trematode species, as seen in *P. ohirai* and *P. westermani* (Le *et al.*, 2019; Oey *et al.*, 2019) and there are no quantity of repeats in individuals to be considered fixed. In many other taxa of trematodes reported to date, for example, *Ech. japonicus*, *F. hepatica*, *Fa. buski*, *Fas. magna*, *P. ohirai*, repetitive units either of long or short sequences and even various quantity within a species, frequently occurred and commonly found (Le *et al.*, 2001, 2016, 2019; Liu *et al.*, 2014; Ma *et al.*, 2016, 2017; Fu *et al.*, 2019; Li *et al.*, 2019*b*). Interestingly, none of the repetitive units was found in *E. caproni*, *H. conoideum* and *T. cymbius* (Yang *et al.*, 2015; Li *et al.*, 2019*a*). The occurrence of repetitive sequences in tandem order in many species certainly is one of the most interesting genomic features, specifying the high-level polymorphism in the NCRs of digenean trematodes. Also, although tandem repeats are common in eukaryotic mitochondrial genomes, their functional role is still not completely understood. However, they do appear to have an accelerated rate of evolution and some involvements in the regulation of the mtDNA coding region (Lunt *et al.*, 1998; Gemayel *et al.*, 2010).

The phylogenetic tree presented in this study indicated the precise placement of *E. revolutum* in the Echinostomatidae matched closely the relationships described in previous studies using nuclear ribosomal sequences (Olson *et al.*, 2003; Tkach *et al.*, 2016). The Fasciolidae and Echinostomatidae are always sister groups within the Echinostomata, as are the Heterophyidae and Opisthorchiidae (within the Opisthorchiata). The echinostomatid species in the tree were also clustered well in the phylogenetic studies by Li *et al.* (2019b) and Fu *et al.* (2019) using the complete mitochondrial genome sequences with one exception. The exception was, in their studies, the closeness of *E. myiagawai* and *E. paraensei* (sister groups); however, in this current study, *E. myiagawai* is closely associated with *E. revolutum*. This discrepancy of echinostomatid relationships was explained by the lack of mtDNA of *E. revolutum* for comparative analysis at that time. In our present study, the echinostomid relationship was also resolved, that the '37-collar-spine *E. revolutum*' group members, *E. revolutum*, *E. myiagawai*, *E. caproni* and *E. paraensei*, were clustered together indicating their genetically close relationships, rather than other *Echinostoma* species, *A. sufrartyfex*, *Echinostoma* sp. JM-2019 and *H. conoideum* (Fig. 2). This relatedness of the 'revolutum' group is reflected by the very low divergence rate (%) of individual and concatenated PCGs and MRGs between *E. revolutum* and *E. myiagawai/E. caproni*, which varied within the least, 6.63% and the highest, 20.03%, compared to the rate of more than 20% in all cases of other echinostomid species (Table 4).

Conclusion

The fully annotated mitogenome of *E. revolutum* and comparative description of mitogenomic features of echinostomids in the present study provide well-supported resolution of relationships of the '*revolutum*' group and the Echinostomata in the relation of other suborders in Plagiorchiida (Platyhelminthes: Digenea). The characterization revealed the taxonomic and phylogenetic relationships of *E. revolutum* to the echinostomatid species and other members in Echinostomatoidea. Molecular analyses of recently available mitogenomic sequences from Echinostomatidae, Himasthlidae and Cyclocoeliidae and comparisons of genetic features have emphasized the '*revolutum*' group to be complex, but phylogenetic analysis has confirmed monophyly Echinostomatidae and Fasciolidae. Data from this species and additional *Echinostoma* spp. will be useful for clarification and reappraisal of the complex echinostome group and for the use in the field of molecular taxonomic, diagnostic, systematic, epidemiological, phylogenetic and population studies of trematodes.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0031182020000128

Acknowledgements. We express our thanks to colleagues and technicians for contribution to our laboratory work. We would also like to express our thanks to the external reviewers for their extremely informative and constructive comments.

Author contributions. Thanh Hoa Le conceived the study, analysed the final data, prepared figures and tables and wrote the manuscript; Linh Thi Khanh Pham, Huong Thi Thanh Doan and Xuyen Thi Kim Le conducted laboratory works and sequence analyses. Weerachai Saijuntha collected, molecularly identified and provided specimens. R.P.V. Jayanthe Rajapakse and Weerachai Saijuntha reviewed the drafts. Scott P. Lawton and Thanh Hoa Le completed and approved the manuscript. All authors read and approved the final manuscript.

Financial support. This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 108.02-2017.09 (PI: ThanhHoa Le).

Conflict of interest. None.

Ethical standards. None.

References

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* **27**, 573–580.

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). *Peer J* **2**, e484.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540–552.

Chai JY (2009) Echinostomes in humans. In Fried B and Toledo R (eds), *The Biology of Echinostomes*. New York, USA: Springer, pp. 147–183.

Detwiler JT, Bos DH and Minchella DJ (2010) Revealing the secret lives of cryptic species: examining the phylogenetic relationships of echinostome parasites in North America. *Molecular Phylogenetics and Evolution* **55**, 611–620.

Faltýnková A, Georgieva S, Soldánová M and Kostadinova A (2015) A re-assessment of species diversity within the 'revolutum' group of Echinostoma Rudolphi, 1809 (Digenea: Echinostomatidae) in Europe. *Systematic Parasitology* **90**, 1–25.

Fu YT, Jin YC, Li F and Liu GH (2019) Characterization of the complete mitochondrial genome of the echinostome *Echinostoma miyagawai* and phylogenetic implications. *Parasitology Research* **118**, 3091–3097.

Gemayel R, Vinces MD, Legendre M and Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* **44**, 445–477.

Georgieva S, Selbach C, Faltýnková A, Soldánová M, Sures B, Skírnisson K and Kostadinova A (2013) New cryptic species of the 'revolutum' group of *Echinostoma* (Digenea: Echinostomatidae) revealed by molecular and morphological data. *Parasites & Vectors* **6**, 64.

Georgieva S, Faltýnková A, Brown R, Blasco-Costa I, Soldánová M, Sitko J, Scholz T and Kostadinova A (2014) *Echinostoma* '*revolutum*' (Digenea: Echinostomatidae) species complex revisited: species delimitation based on novel molecular and morphological data gathered in Europe. *Parasites & Vectors* **7**, 520.

Gordy MA and Hanington PC (2019) A fine-scale phylogenetic assessment of digenean trematodes in central Alberta reveals we have yet to uncover their total diversity. *Ecology and Evolution* **9**, 3153–3238.

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.

Kinkar L, Korhonen PK, Cai H, Gauci CG, Lightowlers MW, Saarma U, Jenkins DJ, Li J, Li J, Young ND and Gasser RB (2019) Long-read sequencing reveals a 4.4 kb tandem repeat region in the mitogenome of *Echinococcus granulosus* (sensu stricto) genotype G1. *Parasites & Vectors* **12**, 238.

Kostadinova A (2005) Family Echinostomatidae Looss, 1899. In Gibson DI, Jones A and Bray AR (eds), *Keys to the Trematoda*, vol. **2**, Wallingford: CAB International, pp. 9–64.

Kumar S, Stecher G and Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–1874.

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

Le TH, Blair D and McManus DP (2001) Complete DNA sequence and gene organization of the mitochondrial genome of the liver fluke, *Fasciola hepatica* L. (Platyhelminthes; Trematoda). *Parasitology* **123**, 609–621.

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

Le TH, Nguyen NTB, Nguyen KT, Doan HTT, Dung DT and Blair D (2016) A complete mitochondrial genome from *Echinochasmus japonicus* supports the elevation of Echinochasminae Odhner, 1910 to family rank (Trematoda: Platyhelminthes). *Infection, Genetics and Evolution* **45**, 369–377.

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. *Peer J* **7**, e7031.

Li Y, Ma XX, Lv QB, Hu Y, Qiu HY, Chang QC and Wang CR (2019a) Characterization of the complete mitochondrial genome sequence of *Tracheophilus cymbius* (Digenea), the first representative from the family Cyclocoelidae. *Journal of Helminthology* **94**, e101.

Li Y, Qiu YY, Zeng MH, Diao PW, Chang QC, Gao Y, Zhang Y and Wang CR (2019b) The complete mitochondrial genome of *Echinostoma miyagawai*: comparisons with closely related species and phylogenetic implications. *Infection, Genetics and Evolution* **75**, 103961.

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, Gasser RB, Young ND, Song HQ, Ai L and Zhu XQ (2014) Complete mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola gigantica*, and their comparison with *F. hepatica*. *Parasites & Vectors* **7**, 150.

Lowe TM and Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**, 955–964.

Lunt DH, Whipple LE and Hyman BC (1998) Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Molecular Ecology* **7**, 1441–1455.

Ma J, He JJ, Liu GH, Leontovyč R, Kašný M and Zhu XQ (2016) Complete mitochondrial genome of the giant liver fluke *Fascioloides magna* (Digenea: Fasciolidae) and its comparison with selected trematodes. *Parasites & Vectors* **9**, 429.

Ma J, Sun MM, He JJ, Liu GH, Ai L, Chen MX and Zhu XQ (2017) *Fasciolopsis buski* (Digenea: Fasciolidae) from China and India may represent distinct taxa based on mitochondrial and nuclear ribosomal DNA sequences. *Parasites & Vectors* **10**, 101.

Miliotis MD and Bier JW (2003) *International Handbook of Foodborne Pathogens*. New York: CRC Press.

Nagataki M, Tantrawatpan C, Agatsuma T, Sugiura T, Duenngai K, Sithithaworn P, Andrews RH, Petney TN and Saijuntha W (2015) Mitochondrial DNA sequences of 37 collar-spined echinostomes (Digenea: Echinostomatidae) in Thailand and Lao PDR reveals presence of two species: *Echinostoma revolutum* and *E. miyagawai*. Infection, *Genetics and Evolution* **35**, 56–62.

Oey H, Zakrzewski M, Narain K, Devi KR, Agatsuma T, Nawaratna S, Gobert GN, Jones MK, Ragan MA, McManus DP and Krause L (2019) Whole-genome sequence of the oriental lung fluke *Paragonimus westermani*. *Gigascience* **8**, giy146.

Olson PD, Cribb TH, Tkach VV, Bray RA and Littlewood DTJ (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). *International Journal for Parasitology* **33**, 733–755.

Perna NT and Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution* **41**, 353–358.

Saijuntha W, Sithithaworn P, Duenngai K, Kiatsopit N, Andrews RH and Petney TN (2011a) Genetic variation and relationships of four species of medically important echinostomes (Trematoda: Echinostomatidae) in South-East Asia. *Infection, Genetics and Evolution* **11**, 375–381.

Saijuntha W, Tantrawatpan C, Sithithaworn P, Andrews RH and Petney TN (2011b) Genetic characterization of *Echinostoma revolutum* and *Echinoparyphium recurvatum* (Trematoda: Echinostomatidae) in Thailand and phylogenetic relationships with other isolates inferred by ITS1 sequence. *Parasitology Research* **108**, 751–755.

Saijuntha W, Tantrawatpan C, Sithithaworn P, Andrews RH and Petney TN (2011c) Spatial and temporal genetic variation of *Echinostoma revolutum* (Trematoda: Echinostomatidae) from Thailand and the Lao PDR. *Acta Tropica* **118**, 105–109.

Suleman KM, Heneberg P, Zhou CY, Muhammad N, Zhu XQ and Ma J (2019) Characterization of the complete mitochondrial genome of *Uvitellina* sp., representative of the family Cyclocoelidae and phylogenetic implications. *Parasitology Research* **118**, 2203–2211.

Tkach VV, Kudlai O and Kostadinova A (2016) Molecular phylogeny and systematics of the Echinostomatoidea Looss, 1899 (Platyhelminthes: Digenea). *International Journal for Parasitology* **46**, 171–185.

Toledo R and Esteban JG (2016) An update on human echinostomiasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **110**, 37–45.

Wey-Fabrizius AR, Podsiadlowski L, Herlyn H and Hankeln T (2013) Platyzoan mitochondrial genomes. *Molecular Phylogenetics and Evolution* **69**, 365–375.

Yang X, Gasser RB, Koehler AV, Wang L, Zhu K, Chen L, Feng H, Hu M and Fang R (2015) Mitochondrial genome of *Hypoderaeum conoideum* – comparison with selected trematodes. *Parasites & Vectors* **8**, 97.

https://doi.org/10.1017/S0031182020000128 Published online by Cambridge University Press