



Research paper

The complete mitochondrial genome of *Echinostoma miyagawai*: Comparisons with closely related species and phylogenetic implications

Ye Li^a, Yang-Yuan Qiu^a, Min-Hao Zeng^a, Pei-Wen Diao^a, Qiao-Cheng Chang^a, Yuan Gao^a, Yan Zhang^a, Chun-Ren Wang^{a,b,*}

^a College of Animal Science and Veterinary Medicine, Heilongjiang Bayi Agricultural University, Daqing, Heilongjiang Province 163319, PR China

^b College of Life Science and Biotechnology, Heilongjiang Bayi Agricultural University, Daqing, Heilongjiang Province 163319, PR China

ARTICLE INFO

Keywords:

Echinostoma miyagawai
Echinostomatidae
Mitochondrial genome
Comparative analysis
Phylogenetic analysis

ABSTRACT

Echinostoma miyagawai (Trematoda: Echinostomatidae) is a common parasite of poultry that also infects humans. *Es. miyagawai* belongs to the “37 collar-spined” or “revolutum” group, which is very difficult to identify and classify based only on morphological characters. Molecular techniques can resolve this problem. The present study, for the first time, determined, and presented the complete *Es. miyagawai* mitochondrial genome. A comparative analysis of closely related species, and a reconstruction of Echinostomatidae phylogeny among the trematodes, is also presented. The *Es. miyagawai* mitochondrial genome is 14,416 bp in size, and contains 12 protein-coding genes (*cox1–3*, *nad1–6*, *nad4L*, *cytb*, and *atp6*), 22 transfer RNA genes (tRNAs), two ribosomal RNA genes (rRNAs), and one non-coding region (NCR). All *Es. miyagawai* genes are transcribed in the same direction, and gene arrangement in *Es. miyagawai* is identical to six other Echinostomatidae and Echinochasmidae species. The complete *Es. miyagawai* mitochondrial genome A + T content is 65.3%, and full-length, pair-wise nucleotide sequence identity between the six species within the two families range from 64.2–84.6%. The *Es. miyagawai* sequences is most similar to *Echinostoma caproni*. Sequence difference are 15.0–33.5% at the nucleotide level, and 8.6–44.2% at the amino acid level, among the six species, for the 12 protein-coding genes. ATG and TAG are the most common initiation and termination codons, respectively. Twenty of the *Es. miyagawai* transfer RNA genes transcribe products of the conventional cloverleaf structure, while two of the transfer RNA genes, namely *trnS1*^(AGC) and *trnS2*^(UGA), have unpaired D-arms. Phylogenetic analyses using our mitochondrial data indicate that *Es. miyagawai* is closely related to other Echinostomatidae species, except for *Echinostoma hortense*, which forms a distinct paraphyletic branch, and *Echinochasmus japonicus*, which is outside the clade containing all other Echinostomatidae species. These phylogenetic results support the elevation of subfamily Echinostomatidae. Our dataset also provides a significant resource of molecular markers to study the taxonomy, population genetics, and systematics of the echinostomatids.

1. Introduction

While *Echinostoma miyagawai* (Digenea: Echinostomatidae) is often a neglected species, further understanding its biology is nonetheless essential for the control of echinostomiasis. *Es. miyagawai* is a food-borne pathogen, found in waterfowl, including ducks and geese, in New Zealand, Thailand, China, and other countries (Georgieva et al., 2017; Nagataki et al., 2015; Faltýnková et al., 2015; Tkach et al., 2016; Wang and Yu, 1993). It has also been reported in humans in China, and is, therefore, a zoonotic parasite (Chen, 2002; Wang and Yu, 1993). *Es. miyagawai* is a member of the so-called “37 collar-spined” or “revolutum” group, which has long been problematic, having confusing

and contradictory taxonomic status within the echinostomes (Fried and Toledo, 2004; Nagataki et al., 2015; Faltýnková et al., 2015). In 1994, Kanev placed *Es. miyagawai* as synonymous with *Echinostoma echinatum* (Kanev, 1994), but that has now been discredited, and *Es. echinatum* is now considered a *species inquirenda* (Faltýnková et al., 2015). *Echinostoma friedi* has also been considered a synonym of *Es. miyagawai*, because both exhibit largely overlapping body and internal organ size ranges (Faltýnková et al., 2015; Toledo et al., 2000).

Molecular tools comprise useful alternative methods for studying taxonomy, population genetics, and systematics. *Es. miyagawai* data available in GenBank was limited, only containing partial sequences of 28S ribosomal DNA (rDNA), cytochrome c oxidase subunit 1 (*cox1*), and

* Corresponding author at: No. 5 Xinfeng Road, High-tech Development Zone, Daqing, Heilongjiang Province, PR China.

E-mail address: chunrenwang@sohu.com (C.-R. Wang).

<https://doi.org/10.1016/j.meegid.2019.103961>

Received 6 June 2018; Received in revised form 30 June 2019; Accepted 9 July 2019

Available online 10 July 2019

1567-1348/ © 2019 Elsevier B.V. All rights reserved.

nicotinamide adenine dinucleotide dehydrogenase subunit 1 (*nad1*). Mitochondrial (mt) genome sequences can provide valuable genetic markers for classification. Wang et al. (2011) reported that *Orientobilharzia turkestanicum* belongs to *Schistosoma* based on mtDNA genome data, and in nematodes, the hypothesis that *Triodontophorus* species belong to Cyathostominae has also been supported with mtDNA data (Wang et al., 2011; Gao et al., 2017). However, the complete mitochondrial genome of this fluke has not been sequenced yet. Therefore, we determined the complete *Es. miyagawai* mt genome sequence, compared it with those of other species from the families Echinostomatidae and Echinochasmidae, and reconstructed the phylogeny of Echinostomatidae.

2. Materials and methods

2.1. Sampling, specimen identity, and DNA extraction

Adult flukes were collected from the small intestines of naturally infected ducks in Daqing, Heilongjiang Province, China. The worms were thoroughly washed in physiological saline, and species were identified based on size and host predilection (Eom et al., 1984). The flukes were then fixed in 70% ethanol and stored at -20°C until use. Total genomic DNA was extracted from individual worms with the TIANamp Genomic DNA Kit (Tiagen, Beijing, China), according to the manufacturer's protocol. Species identification was performed by PCR amplification of the complete second internal transcribed spacer (ITS2) sequence (MH796365), 28S sequence (MH748722), and *cox1* sequence (MH748721), using the primers reported in a previous study and re-designed for the present study (Table S1) (Gasser et al., 2008). The 28S and *cox1* sequences were aligned against the corresponding partial sequences of *Es. miyagawai* available in GenBank (28S: KT956916, *cox1*: KP455511) with sequence identity of 99.8% and 99.5%, respectively, while the ITS2 sequence has 100% identity with the sequence of *Es. friedi* (syn. *Es. miyagawai*, KJ848450), which verifies that these worms are *Es. miyagawai*.

2.2. Amplification, sequence analysis, and comparative analysis

Using the program primer premier (v. 5.0, Biosoft International, Canada), seven pairs of primers were designed based on the mtDNA sequences of other Echinostomatidae species (Table S1). PCR (25 μL) was performed using 18.3 μL of distilled water, 2.5 μL of $10\times$ Ex *Taq* buffer, 2 μL of dNTP Mixture (2.5 mM), 0.5 μL of each primer (25 mM), 1 μL of DNA template, and 0.2 μL of Ex *Taq* DNA polymerase (5 U/ μL), in a Takara TP600 thermocycler (TKAKRA, Japan) under the following conditions: 94°C for 5 min (initial denaturation), then 94°C for 30 s (denaturation), $50\text{--}65^{\circ}\text{C}$ for 1 min (annealing), and 72°C for 1 min/kb (extension) for 40 cycles, and a final extension at 72°C for 10 min. PCR products were sent to Sangon Biotech Company (Shanghai, China) for sequencing in both directions using the same primers.

Sequences were assembled by manual inspection and aligned against complete *Echinostoma paraensei* (KT008005) and *Es. caproni* (AP017706) mtDNAs, to identify gene boundaries, using the program DNASTar v. 5.0 (Burland, 2000). Each protein-coding gene was translated into its corresponding amino acid sequence using MEGA v. 5.0 (Tamura et al., 2011). The secondary structures of tRNA genes were predicted using the online program tRNAscan-SE (<http://lowelab.ucsc.edu/tRNAscan-SE>) or by manually (Lowe and Chan, 2016). Codon usage frequencies of the 12 protein-coding genes were calculated using an online Web resource (<http://www.detaibio.com/tools/codon-usage-calculator.html>). We identified rRNA genes by comparison with mtDNAs of related species [*Es. paraensei* (KT008005), *Es. caproni* (AP017706), and *Echinostoma hortense* (KR062182)] published previously, using MEGA v. 5.0 (Tamura et al., 2011). Tandem repeats (TRs) were detected in non-coding regions (NCRs) using the Tandem Repeat Finder (<http://tandem.bu.edu/trf/trf.html>) (Benson, 1999).

Comparisons were made based on complete mtDNA size, gene arrangement, A + T content, AT/GC skew, and nucleotide and amino acid sequence difference, determined from the individual protein-coding genes among *Es. miyagawai*, *Es. caproni* (AP017706), *Es. hortense* (KR062182), *Es. paraensei* (KT008005), *Hypoderaeum conoideum* (KM111525), and *Echinochasmus japonicus* (KP844722). The nucleotide and amino acid sequence differences were calculated using MegAlign v. 5.01 (Burland, 2000). The AT-skew and GC-skew values in both coding and non-coding regions were calculated using the eqs. $\text{AT-skew} = (A - T)/(A + T)$, and $\text{GC-skew} = (G - C)/(G + C)$ (Ma et al., 2016).

2.3. Phylogenetic analyses

Conceptually translated amino acid sequences from each protein-coding gene in the *Es. miyagawai* mtDNA sequence were concatenated and aligned with those from 24 published trematode mtDNAs by MEGA v. 5.0 (Tamura et al., 2011). We used the Gblocks online server (http://molevol.cmima.csic.es/castresana/Gblocks_server.html) to exclude ambiguously aligned regions from the multiple amino acid sequence alignment, specifying the "less stringent" selection options (Castresana, 2000).

In addition to our *Es. miyagawai* data, the 25-member trematodes phylogenetic dataset contains the following: Echinostomatidae: *H. conoideum* (KM111525), *Es. caproni* (AP017706), *Es. paraensei* (KT008005), *Es. hortense* (KR062182); Echinochasmidae: *Ec. japonicus* (KP844722); Paramphistomidae: *Paramphistomum cervi* (KF475773), *Paramphistomum leydeni* (KP341657); Troglotrematidae: *Paragonimus westermani* (KX943544); Opisthorchiidae: *Opisthorchis felineus* (NC011127), *Metorchis orientalis* (NC028008), *Clonorchis sinensis* (NC012147); Notocotyliidae: *Ogmocotyle sikae* (NC027112); Heterophyidae: *Metagonimus yokogawai* (NC023249); Gastrothylacidae: *Gastrothylax crumenifer* (NC027833), *Fischoederius elongatus* (NC028001); Fasciolidae: *Fasciola gigantica* (NC024025), *Fasciola hepatica* (NC002546), *Fasciola jacksoni* (KX787886), *Fascioloides magna* (KU060148), *Fasciolopsis buski* (KX169163); Dicrocoeliidae: *Dicrocoelium chinensis* (NC025279), *Dicrocoelium dendriticum* (NC025280), *Eurytrema pancreaticum* (KP241855). Schistosomatidae: *Schistosoma turkestanicum* (syn. *Orientobilharzia turkestanicum* HQ283100) was included as an outgroup.

Phylogenetic trees were reconstructed using Bayesian inference (BI), maximum parsimony (MP), and maximum likelihood (ML) methods. BI methods were performed using the mixed model in MrBayes v. 3.1.1 and 1,000,000 metropolis-coupled Markov chain Monte Carlo generations. The first 250 trees were omitted as burn-in, and the remaining trees were used to calculate Bayesian posterior probabilities (Ronquist et al., 2012). MP methods were performed using the Fitch criterion within PAUP v. 4.0 Beta 10 (Swofford, 2002). Bootstrap support values were calculated in PAUP from 1000 bootstrap replicates with 10 random additions per replicate. ML methods (JTT + I + G + F model) using PhyML 3.0 and bootstrap values were determined from 100 replicates (Guindon and Gascuel, 2003). Building the consensus tree protocols used to phylogenetic analyses were based on those of the previous reports (Chang et al., 2016; Gao et al., 2017; Duan et al., 2015). Phylograms were viewed and drawn using TreeView 1.65 (Page, 1996).

3. Results and discussion

3.1. *Es. miyagawai* mtDNA features

The complete *Es. miyagawai* mtDNA is 14,416 bp in length, deposited in GenBank under accession number MH393928. The circular genome contains 36 genes, all transcribed in the same direction, and includes 12 protein-coding genes (*cox1-3*, *nad1-6*, *nad4L*, *cytb*, and *atp6*), two rRNA genes (*rml* and *rms*), and 22 tRNA genes, but lacks the *atp8* gene (Table 1 and Fig. S1). *Es. miyagawai* nucleotide composition is

Table 1
Mitochondrial genome organization of *Echinostoma miyagawai*.

Gene/ region	Positions		Size		Codons		
	Start	End	No. of nt	No. of aa	Initiation	Termination	Anti codons
<i>cox3</i>	1	651	651	216	ATG	TAA	
<i>trnH</i>	655	720	66				GTG
<i>cytb</i>	723	1832	1110	369	ATG	TAG	
<i>nad4L</i>	1833	2105	273	90	ATG	TAG	
<i>nad4</i>	2066	3349	1284	427	ATG	TAG	
<i>trnQ</i>	3355	3424	70				TTC
<i>trnF</i>	3425	3487	63				GAA
<i>trnM</i>	3521	3586	66				CAT
<i>atp6</i>	3590	4108	519	172	ATG	TAG	
<i>nad2</i>	4116	4985	870	289	ATG	TAG	
<i>trnV</i>	4990	5053	64				TAC
<i>trnA</i>	5078	5145	68				TGC
<i>trnD</i>	5150	5215	66				GTC
<i>nad1</i>	5216	6118	903	300	GTG	TAG	
<i>trnN</i>	6125	6190	66				GTT
<i>trnP</i>	6195	6263	69				TGG
<i>trnI</i>	6264	6327	64				GAT
<i>trnK</i>	6336	6404	69				CTT
<i>nad3</i>	6409	6765	357	118	ATG	TAG	
<i>trnS1</i>	6770	6829	60				AGC
<i>trnW</i>	6834	6899	66				TCA
<i>cox1</i>	6903	8441	1539	512	GTG	TAA	
<i>trnT</i>	8445	8512	68				TGT
<i>rrnL</i>	8530	9521	992				
<i>trnC</i>	9522	9584	63				GTA
<i>rrnS</i>	9585	10,338	754				
<i>cox2</i>	10,339	10,947	609	202	ATG	TAG	
<i>nad6</i>	10,959	11,411	453	150	ATG	TAG	
<i>trnY</i>	11,412	11,477	66				GTA
<i>trnL1</i>	11,478	11,542	65				TAG
<i>trnS2</i>	11,545	11,603	59				TGA
<i>trnL2</i>	11,635	11,697	63				TAA
<i>trnR</i>	11,700	11,759	60				TCG
<i>nad5</i>	11,762	13,327	1566	521	GTG	TAG	
<i>trnG</i>	13,347	13,412	66				TCC
<i>trnE</i>	13,419	13,485	67				TCG
NCR	13,486	14,416	931				

aa, amino acid; nt, nucleotide.

biased toward A and T, with T being the most abundant nucleotide, and C being the least used. Total *Es. miyagawai* mtDNA A + T content is 65.30% (see Table 2).

The *Es. miyagawai* mtDNA protein-coding genes quantity is consistent with the mt genomes of other reported trematodes (e.g. Echinostomatidae), Chromadorea nematodes (e.g. Strongylidae), and cestodes (e.g. Hymenolepididae), but distinct from Adenophorea nematodes (e.g. Trichuridae), which contain 13 protein-coding genes (Liu et al., 2016; von Nickisch-Roseneck et al., 2001; Duan et al., 2015; Lavrov and Brown, 2001). Except the non-coding region, *Es. miyagawai* intergenic spacer sequences range in length, 0–33 bp, with 4 bp being most frequent. The *nad4L* and *nad4* genes overlap by 40 bp in *Es. miyagawai* mtDNA, which is consistent with *H. conoideum*, *Ec. japonicus*, *Es. hortense*, *P. westermani*, *C. sinensis*, and *F. hepatica*, but is a longer overlap than those in *Schistosoma haematobium* (28 bp) and *Schistosoma japonicum* (37 bp), and shorter than those in *Schistosoma mekongi* (64 bp) (Yang et al., 2015; Liu et al., 2016; Biswal et al., 2014; Shekhovtsov et al., 2010; Le et al., 2000; Littlewood et al., 2006; Young et al., 2015). ATG and TAG are the most frequent initiation and termination codons, respectively, for the 12 *Es. miyagawai* mtDNA protein-coding genes. The most frequent codon overall is TTT (Phe), with a frequency of 10.8%. The least used codon is CGC (Arg: 0.06%) (Table 3).

The 22 tRNA genes range from 59 to 69 bp in length in the *Es. miyagawai* mtDNA. The secondary structures of 20 of the tRNA gene products can be folded into the conventional cloverleaf structure, while *trnS1*^(AGC) and *trnS2*^(UGA) have unpaired D-arms replaced by the loops of 13–15 bp (Fig. S2). The *rrnL* gene is 992 bp and *rrnS* is 754 bp in size. The A + T content of *rrnL* and *rrnS* is 65.02% and 61.80%, respectively. One 931 bp long NCR is found between *trnE* and *cox3* (68.53% A + T content), and contains an array of two identical TRs(319 bp each). Detailed annotation of the *Es. miyagawai* mtDNA, including the position and length of each gene, and codon usage is listed in Table 1.

3.2. Comparative analysis

The complete *Es. miyagawai* mtDNA (14,416 bp) is shorter than that of *Es. hortense* (14,994 bp) in Echinostomatidae, *Ec. japonicus* (15,865 bp) in Echinochasmidae, and *Schistosoma spindale* (16,901 bp) in Schistosomatidae; but longer than that of *Es. caproni* (14,150 bp) and *H. conoideum* (14,180 bp) in Echinostomatidae, and *C. sinensis* (13,875 bp) in Opisthorchiidae (Le et al., 2016; Littlewood et al., 2006;

Table 2
The complete nucleotide identify analyses of Echinostomatidae and Echinochasmidae species.

Genes	Nucleotide length (bp)						Nucleotide difference (%)	Amino acid numbers (aa)						Amino acid difference (%)
	<i>Ec. m.</i>	<i>Ec. c.</i>	<i>Ec. h.</i>	<i>Ec. p.</i>	<i>H. c.</i>	<i>Es. j.</i>		<i>Ec. m.</i>	<i>Ec. c.</i>	<i>Ec. h.</i>	<i>Ec. p.</i>	<i>H. c.</i>	<i>Es. j.</i>	
<i>cox3</i>	651	645	651	649	942	648	11.0–40.3	216	214	216	216	313	215	5.2–50.0
<i>cytb</i>	1110	1110	1074	1110	1056	1116	12.6–27.3	369	369	357	369	369	371	6.0–40.7
<i>nad4L</i>	273	273	288	273	279	270	14.2–31.3	90	90	95	90	92	89	8.9–33.3
<i>nad4</i>	1284	1284	1281	1284	1284	1284	17.8–37.1	427	427	126	427	427	427	10.5–40.7
<i>atp6</i>	519	519	516	519	519	519	15.2–32.8	172	172	171	172	172	172	15.1–32.6
<i>nad2</i>	870	870	870	842	867	882	15.9–38.8	289	289	289	281	288	293	12.0–48.2
<i>nad1</i>	903	903	900	902	903	906	13.4–31.5	300	300	299	300	300	301	7.7–39.9
<i>nad3</i>	357	372	357	357	357	357	14.2–29.7	118	123	118	118	118	118	16.1–38.1
<i>cox1</i>	1539	1539	1542	1548	1539	1539	12.4–25.4	512	512	513	515	512	512	2.5–23.6
<i>cox2</i>	609	609	600	609	603	597	13.6–39.3	202	202	199	202	200	198	1.5–49.0
<i>nad6</i>	453	453	450	453	453	450	15.7–37.8	150	150	149	150	150	149	11.3–45.3
<i>nad5</i>	1563	1566	1572	1609	1566	1575	14.4–36.7	521	521	523	536	521	524	9.8–41.7
Total	10,131	10,143	10,047	10,155	10,368	10,143	15.0–33.9	3366	3369	3337	3376	3462	3369	9.5–44.2
<i>rrnL</i>	992	980	993	990	979	991	10.3–32.4	–	–	–	–	–	–	–
<i>rrnS</i>	754	729	759	758	751	758	9.9–33.3	–	–	–	–	–	–	–
22 <i>trn</i>	1427	1417	1411	1434	1443	1461	–	–	–	–	–	–	–	–
overall nt	14,413	14,150	14,994	–	14,180	15,865	15.4–35.3	–	–	–	–	–	–	–
A + T %	65.30	64.47	63.05	–	61.38	61.48	–	–	–	–	–	–	–	–

Ec. m., *Echinostoma miyagawai*; *Ec. c.*, *Echinostoma caproni*; *Ec. h.*, *Echinostoma hortense*; *Ec. p.*, *Echinostoma paraensei*; *H. c.*, *Hypoderaeum conoideum*; *Es. j.*, *Echinochasmus japonicus*; aa, amino acid; nt, nucleotide.

Table 3
Codon usage for 12 protein-coding genes in the mitochondrial genome of *Echinostoma miyagawai*.

Codon	Amino acid	Number	Frequency (%)	Codon	Amino acid	Number	Frequency (%)
TTT	F	365	10.8	TAT	Y	159	4.70
TTC	F	20	0.59	TAC	Y	7	0.20
TTA	L	193	5.71	TAA	*	2	0.06
TTG	L	237	7.01	TAG	*	10	0.30
CTT	L	71	2.10	CAT	H	39	1.15
CTC	L	7	0.20	CAC	H	14	0.41
CTA	L	16	0.47	CAA	Q	8	0.27
CTG	L	19	0.56	CAG	Q	18	0.53
ATT	I	123	3.64	AAT	N	45	1.33
ATC	I	11	0.32	AAC	N	8	0.27
ATA	M	85	2.51	AAA	N	30	0.89
ATG	M	105	3.10	AAG	K	48	1.42
GTT	V	230	6.80	GAT	D	69	2.04
GTC	V	11	0.32	GAC	D	8	0.27
GTA	V	55	1.63	GAA	E	19	0.56
GTG	V	68	2.01	GAG	E	53	1.56
TCT	S	141	4.17	TGT	C	100	2.96
TCC	S	12	0.36	TGC	C	9	0.27
TCA	S	27	0.80	TGA	W	41	1.21
TCG	S	13	0.38	TGG	W	65	1.92
CCT	P	46	1.36	CGT	R	42	1.24
CCC	P	22	0.65	CGC	R	2	0.06
CCA	P	18	0.53	CGA	R	6	0.18
CCG	P	9	0.27	CGG	R	13	0.38
ACT	T	39	1.15	AGT	S	89	2.63
ACC	T	18	0.53	AGC	S	6	0.18
ACA	T	16	0.47	AGA	S	29	0.86
ACG	T	20	0.59	AGG	S	39	1.15
GCT	A	68	2.01	GGT	G	165	4.88
GCC	A	12	0.36	GGC	G	22	0.65
GCA	A	20	0.59	GGA	G	29	0.86
GCG	A	14	0.41	GGG	G	73	2.16

Shekhovtsov et al., 2010; Yang et al., 2015; Liu et al., 2016). This length discrepancy is mainly due to NCR lengths: In *Es. miyagawai*, it is 931 bp, versus *Es. hortense* (1553 bp), *Ec. japonicus* (2342 bp), *S. spindale* (2492 bp), *Es. caproni* (685 bp), *H. conoideum* (357 bp), and *C. sinensis* (220 bp) (Liu et al., 2016; Yang et al., 2015; Le et al., 2016; Littlewood et al., 2006; Shekhovtsov et al., 2010).

Pair-wise and full-length mtDNA nucleotide sequence identity are 64.2–84.6% between the six trematodes in Echinostomatidae and Echinochasmidae (*Es. miyagawai*, *Es. paraensei*, *Es. caproni*, *Es. hortense*, *H. conoideum*, and *Ec. japonicus*), which is similar to those within the Opisthorchiidae (*M. orientalis*, *C. sinensis*, *O. felineus*, and *O. viverrini*, 79.5–81.0%) (Liu et al., 2016; Yang et al., 2015; Na et al., 2016; Shekhovtsov et al., 2010; Cai et al., 2012). The highest full-length nucleotide sequence identity (84.6%) is between *Es. miyagawai* and *Es. caproni*, which is lower than that of *Fasciola* sp. and *F. hepatica* (88.2%) (Le et al., 2001; Liu et al., 2014a). Pair-wise amino acid sequence identity is 55.8–91.4% within the six Echinostomatidae and Echinochasmidae trematodes. The lower identity are primarily between *Es. hortense* and the other five. The divergence in nucleotide and amino acid between 12 protein-coding genes within six Echinostomatidae and Echinochasmidae species is not higher than 50%, seeing Table 2. The mtDNA gene arrangement in the six is identical as in *D. chinensis*, *D. dendriticum*, *E. pancreaticum*, *P. cervi*, *P. leydeni*, and *P. westermani*, but different from *O. felineus*, *M. orientalis*, *C. sinensis*, and *F. hepatica* (Liu et al., 2014b; Chang et al., 2016; Yan et al., 2013; Ma et al., 2015; Biswal et al., 2014; Shekhovtsov et al., 2010; Na et al., 2016; Le et al., 2001). The total *Es. miyagawai* mtDNA A + T content is 65.30%, which is slightly higher than that of *Es. caproni* (64.47%), *Es. hortense* (63.06%), *H. conoideum* (61.38%), and *Ec. japonicus* (61.48%), and much higher than that of *P. westermani* (51.5%), but lower than that of *S. spindale* (72.71%) and *S. haematobium* (72.35%) in the family of Schistosomatidae (Biswal et al., 2014; Littlewood et al., 2006; Yang et al., 2015; Liu et al., 2016; Le et al., 2016). *Es. miyagawai* and *Es.*

caproni have very high full-length mtDNA nucleotide sequence identity (84.6%), while nucleotide sequence identity is very low between *Es. hortense* and *Ec. japonicus* (64.2%), which is similar with that of *Schistosoma*, in which the nucleotide sequence identity between *S. spindale* and *S. haematobium* is 62.7% (Liu et al., 2016; Le et al., 2016; Littlewood et al., 2006).

This range of similar lengths in 12 protein-coding genes within six Echinostomatidae and Echinochasmidae trematode species, except for the *cox3* gene (Table 2). The *cox3* gene in *H. conoideum* is 291–312 bp longer than those of the others (*Es. miyagawai*, *Es. caproni*, *Es. hortense*, *Es. paraensei*, and *Ec. japonicus*) (Liu et al., 2016; Le et al., 2016; Yang et al., 2015). This range of similarity of mtDNA gene lengths within families is also observed in nematodes (Strongylidae), tapeworms (Hymenolepididae), and other trematodes (Opisthorchiidae) (Gao et al., 2017; Gao et al., 2015; Shekhovtsov et al., 2010; Cai et al., 2012). Pair-wise comparisons between the 12 protein-coding genes show sequence difference between the six Echinostomatidae and Echinochasmidae species at both the nucleotide (15.0–30.5%) and amino acid (10.5–41.2%) level (Table 4). The *Es. caproni* and *Es. hortense* sequences

Table 4

Pairwise differences among mt genomes of *Echinostoma miyagawai* (Ec. m), *Echinostoma caproni* (Ec. c), *Echinostoma hortense* (Ec. h), *Echinostoma paraensei* (Ec. p), *Hypoderaeum conoideum* (H. c) and *Echinochasmus japonicus* (Es. j).

Species	Overall nt difference (%)	PCGs nt difference (%)	PCGs aa difference (%)
<i>Ec.m/Ec. c</i>	15.4	15.0	11.5
<i>Ec.m/Ec. h</i>	32.6	30.5	41.2
<i>Ec.m/Ec. p</i>	15.8	15.2	10.5
<i>Ec.m/H. c</i>	25.1	23.7	21.8
<i>Ec.m/Es. j</i>	31.7	29.8	32.9

PCGs, protein coding genes; aa, amino acid; nt, nucleotide.

Table 5
AT/GC skews in mt genomes of the family Echinostomatidae and Echinochasmidae.

Gene	AT-skew						GC-skew					
	<i>Ec. m</i>	<i>Ec. c</i>	<i>Ec. h</i>	<i>Ec. p</i>	<i>H. c</i>	<i>Es. j</i>	<i>Ec. m</i>	<i>Ec. c</i>	<i>Ec. h</i>	<i>Ec. p</i>	<i>H. c</i>	<i>Es. j</i>
<i>cox3</i>	-0.392	-0.443	-0.358	-0.437	-0.336	-0.466	0.454	0.467	0.339	0.452	0.343	0.498
<i>cytb</i>	-0.467	-0.473	-0.356	-0.453	-0.408	-0.455	0.343	0.369	0.333	0.333	0.315	0.441
<i>nad4L</i>	-0.409	-0.386	-0.322	-0.444	-0.377	-0.526	0.623	0.547	0.354	0.552	0.596	0.663
<i>nad4</i>	-0.453	-0.506	-0.484	-0.457	-0.466	-0.503	0.422	0.436	0.331	0.361	0.429	0.482
<i>atp6</i>	-0.461	-0.510	-0.515	-0.456	-0.390	-0.528	0.265	0.297	0.311	0.260	0.288	0.238
<i>nad2</i>	-0.509	-0.526	-0.450	-0.528	-0.528	-0.547	0.522	0.435	0.417	0.481	0.413	0.476
<i>nad1</i>	-0.449	-0.479	-0.402	-0.448	-0.477	-0.543	0.478	0.504	0.457	0.500	0.529	0.604
<i>nad3</i>	-0.498	-0.527	-0.544	-0.526	-0.518	-0.499	0.509	0.593	0.433	0.481	0.500	0.570
<i>cox1</i>	-0.391	-0.402	-0.395	-0.376	-0.409	-0.402	0.310	0.349	0.347	0.333	0.391	0.439
<i>cox2</i>	-0.180	-0.190	-0.268	-0.167	-0.272	-0.356	0.330	0.303	0.328	0.297	0.300	0.398
<i>nad6</i>	-0.498	-0.540	-0.496	-0.489	-0.512	-0.507	0.329	0.515	0.278	0.457	0.314	0.469
<i>nad5</i>	-0.532	-0.530	-0.403	-0.539	-0.564	-0.598	0.459	0.470	0.414	0.482	0.543	0.561
<i>rrnL</i>	-0.233	-0.217	-0.238	-0.253	-0.202	-0.247	0.308	0.270	0.329	0.313	0.321	0.388
<i>rrnS</i>	-0.128	-0.132	-0.153	-0.092	-0.102	-0.178	0.277	0.285	0.337	0.292	0.326	0.361
NCR	-0.241	-0.307	-0.022	-	-0.214	-0.391	0.426	0.061	0.236	-	0.714	0.465
12PCG	-0.446	-0.468	-0.416	-0.448	-0.448	-0.496	0.406	0.424	0.367	0.403	0.413	0.484
overall	-0.379	-0.398	-0.315	-	-0.383	-0.431	0.390	0.396	0.345	-	0.393	0.464

Ec. m, *Echinostoma miyagawai*; *Ec. c*, *Echinostoma caproni*; *Ec. h*, *Echinostoma hortense*; *Ec. p*, *Echinostoma paraensei*; *H. c*, *Hypoderaeum conoideum*; *Es. j*, *Echinochasmus japonicus*.

are the most and the least different from *Es. miyagawai*, respectively. The *cox1* gene is the most conserved in both nucleotide and amino acid sequences overall. However, *nad5* is the most conserved protein-coding gene between *D. chinensis* and *D. dendriticum*, and *nad1* and *cox1* are the most highly conserved in *F. gigantica* (Liu et al., 2014b; Liu et al., 2014a). ATG and TAG are the most frequently used initiation and termination codons, respectively, within the six Echinostomatidae and Echinochasmidae trematodes (*Es. miyagawai*, *Es. caproni*, *Es. hortense*, *Es. paraensei*, *H. conoideum* and *Ec. japonicus*), which is the same situation as many other trematodes, including *F. gigantica*, *S. japonicum* and *P. westermani*, (Liu et al., 2016; Le et al., 2016; Yang et al., 2015; Biswal et al., 2014; Liu et al., 2014a; Young et al., 2015).

AT/GC skews in each mtDNA gene or region of the six Echinostomatidae and Echinochasmidae species are displayed in Table 5. AT-skew values of the six trematodes are generally negative, and the majority of GC-skew values are positive. The AT-skew and GC-skew values individually for each of the 12 protein-coding genes and rDNAs, or overall, are all quite similar; however, the AT-skew values for the *Es. hortense* NCR (-0.022) is lower than the other five, which range from -0.391 to -0.214. GC-skew values in the six species have a large range of variation. *Es. caproni* has the lowest (0.061) and *H. conoideum* has the highest (0.714). Although some genes in *Es. miyagawai* and *H. conoideum* possess different skew values, the overall pattern of nucleotide skew in these two species are quite close.

The five Echinostomatidae species (*Es. miyagawai*, *Es. caproni*, *Es. hortense*, *Es. paraensei*, and *H. conoideum*) all have one NCR, which is consistent with *S. spindale*; however, it is not consistent with *Ec. japonicus* in Echinochasmidae, *C. sinensis* and *M. orientalis* in Opisthorchiidae, and *F. hepatica* in Fasciolidae, which all have two mtDNA NCRs (a short region and a long region) (Littlewood et al., 2006; Shekhovtsov et al., 2010; Na et al., 2016; Le et al., 2000). The NCR location in the five Echinostomatidae species is the same, between *trnE* and *cox3*, and is also the same as *S. spindale*, but distinct from *Ec. japonicus* in Echinochasmidae (the long NCR is located between *trnG* and *trnE*, and the short NCR is located between *trnE* and *cox3*) (Liu et al., 2016; Le et al., 2016; Yang et al., 2015; Young et al., 2015; Littlewood et al., 2006). Additionally, two identical TRs, 319 bp long each, occur in the *Es. miyagawai* NCR, and also occur in *Ec. japonicus* (eight identical TRs, 240 bp long), *Trichobilharzia regenti* (three identical TRs, 184 bp long), *F. hepatica* (eight identical TRs, 85 bp long) (Le et al., 2016; Le et al., 2001; Webster et al., 2007). The function and role of this variation in the length and quantity of TRs of NCR are unknown.

Detailed comparisons are provided in Tables 2–5, including mtDNA size, A + T content, nucleotide and amino acid sequence differences, and AT/GC skews.

3.3. Phylogenetic analyses

Phylogenetic analyses using three methods (ML, MP, and BI) yielded identical tree topologies based on the concatenated amino acid sequences of 12 protein-coding genes (Fig. 1). Phylogenetic relationships among species were well resolved with very high nodal support throughout. The species in Echinostomata, Pronocephalata, Opisthorchiata, and Xiphidiata all cluster together appropriately in our phylogenetic tree. The Echinostomata clade contains three families, Echinostomatidae, Echinochasmidae, and Fasciolidae. All of the Echinostomatidae species cluster together within the Echinostomatidae clade, except *Es. hortense*. Interestingly, *Es. hortense* clusters with the five Fasciolidae species used in our study., *Ec. japonicus* forms a distinct branch basal to both Echinostomatidae and Fasciolidae. *Es. miyagawai* and *Es. paraensei* are sister taxa in our tree, which is similar to a study of ITS sequences from trematode species that from metacercariae (Cech et al., 2017).

The present phylogenetic study using combined amino acid sequences of 12 protein-coding genes shows *Es. miyagawai* to be more closely related to *Es. paraensei* than to *Es. caproni*, but nodal support is not spectacular, 58–79%. This is not consistent with a previous study based on partial *nad1* sequences, in which *Es. miyagawai* and *Es. caproni* cluster together, both being equally related to *Es. paraensei* (Georgieva et al., 2017). Three *Echinostoma* species, *Es. miyagawai*, *Es. paraensei*, and *Es. caproni*, which are held together by 100% nodal support in our tree, are more closely related to *H. conoideum*, of the same family (also with 100% nodal support), than to *Es. hortense*, of the same genus, but which is a member of the Fasciolidae clade in our tree (nodal support 83–100%). This is consistent with a previous study using the 28S rRNA gene in which *Es. miyagawai*, *Es. paraensei*, and *H. conoideum* cluster together on one branch, and *Isthmiophora hortensis* (Syn. *Es. hortense*) is a sister taxa, also considered to be within Echinostomatidae, but is in an entirely different clade (Tkach et al., 2016). Similar ambiguous results regarding the location of *I. hortensis* based on both 28S and ITS2 sequence phylogenies have been inferred (Stanevičiūtė et al., 2015). Our reconstructed Echinostomata phylogeny is consistent with another study based on mtDNA sequences, except *Es. hortense* was not included in that particular case (Le et al., 2016). Differences among studies may

- Trematoda). *Parasitology* 123 (Pt 6), 609–621.
- Le, T.H., Nguyen, N.T.B., Nguyen, K.T., Doan, H.T.T., Dung, D.T., Blair, D., 2016. A complete mitochondrial genome from *Echinochasmus japonicus*, supports the elevation of Echinochasmidae Ochner, 1910 to family rank (Trematoda: Platyhelminthes). *Infect. Genet. Evol.* 45, 369–377.
- Littlewood, D.T., Lockyer, A.E., Webster, B.L., Johnston, D.A., Le, T.H., 2006. The complete mitochondrial genomes of *Schistosoma haematobium* and *Schistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. *Mol. Phylogenet. Evol.* 39 (2), 452–467.
- Liu, G.H., Gasser, R.B., Young, N.D., Song, H.Q., Ai, L., Zhu, X.Q., 2014-03-31a. Complete mitochondrial genomes of the ‘intermediate form’ of *Fasciola gigantica*, and their comparison with *F. hepatica*. *Parasit. Vectors* 7 (1), 150.
- Liu, G.H., Yan, H.B., Otranto, D., Wang, X.Y., Zhao, G.H., Jia, W.Z., Zhu, X.Q., 2014b. *Dicrocoelium chinensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochondrial and nuclear ribosomal DNA sequences. *Mol. Phylogenet. Evol.* 79 (1), 325–331.
- Liu, Z.X., Zhang, Y., Liu, Y.T., Chang, Q.C., Su, X., Fu, X., Yue, D.M., Gao, Y., Wang, C.R., 2016. Complete mitochondrial genome of *Echinostoma hortense* (Digenea: Echinostomatidae). *Korean J. Parasitol.* 54 (2), 173–179.
- Lowe, T.M., Chan, P.P., 2016. tRNAscan-SE on-line: search and contextual analysis of transfer RNA genes. *Nucl. Acids Res.* 44, W54–W57.
- Ma, J., He, J.J., Liu, G.H., Zhou, D.H., Liu, J.Z., Liu, Y., Zhu, X.Q., 2015. Mitochondrial and nuclear ribosomal dna dataset supports that *Paramphistomum leydeni* (Trematoda: Digenea) is a distinct rumen fluke species. *Parasit. Vectors* 8 (1), 201.
- Ma, J., He, J.J., Liu, G.H., Blair, D., Liu, L.Z., Liu, Y., Zhu, X.Q., 2016. Mitochondrial genome of *Ogmocotyle sikae* and implications for phylogenetic studies of the Notocotylidae trematodes. *Infect. Genet. Evol.* 37, 208–214.
- Morgan, J.A., Blair, D., 1998. Relative merits of nuclear ribosomal internal transcribed spacers and mitochondrial CO1 and ND1 genes for distinguishing among *Echinostoma* species (Trematoda). *Parasitology* 116 (3), 289–297.
- Na, L., Gao, J.F., Liu, G.H., Fu, X., Su, X., Yue, D.M., Gao, Y., Zhang, Y., Wang, C.R., 2016. The complete mitochondrial genome of *Metorchis orientalis*, (trematoda: Opisthorchiidae): comparison with other closely related species and phylogenetic implications. *Infect. Genet. Evol.* 39, 45–50.
- Nagataki, M., Tantrawatpan, C., Agatsuma, T., Sugiura, T., Duennagai, K., Sithithaworn, P., Andrews, R.H., Petney, T.N., 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*. *Infect. Genet. Evol.* 35, 56–62.
- von Nickisch-Roseneck, M., Brown, W.M., Boore, J.L., 2001. Complete sequence of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: gene arrangements indicate that Platyhelminths are Eutrochozoans. *Mol. Biol. Evol.* 18 (5), 721–730.
- Page, R.D., 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12 (4), 357–358.
- Ronquist, F., Teslenko, M., Van, d.M.P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61 (3), 539–542.
- Shekhovtsov, S.V., Katokhin, A.V., Kolchanov, N.A., Mordvinov, V.A., 2010. The complete mitochondrial genomes of the liver flukes *Opisthorchis felineus* and *Clonorchis sinensis* (Trematoda). *Parasitol. Int.* 59 (1), 100–103.
- Sohn, W.M., Na, B.K., Shin, S.S., 2017. New definitive hosts and differential body indices of *Isthmiophora hortensis* (Digenea: Echinostomatidae). *Korean J. Parasitol.* 55 (3), 287–294.
- Son, W.Y., Huh, S., Lee, S.U., Woo, H.C., Hong, S.J., 1994. Intestinal trematode infections in the villagers in Koje-myon, Kochang-gun, Kyongsangnam-do, Korea. *Korean J. Parasitol.* 32 (3), 149–155.
- Stanevičiūtė, G., Stunžėnas, V., Petkevičiūtė, R., 2015. Phylogenetic relationships of some species of the family Echinostomatidae odner, 1910 (Trematoda), inferred from nuclear rdna sequences and karyological analysis. *Comp. Cytogenet.* 9 (2), 257–270.
- Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, MA.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28 (10), 2731–2739.
- Toledo, R., Muñoz-Antolí, C., Esteban, J.G., 2000. The life-cycle of *Echinostoma friedi* n. sp. (Trematoda: Echinostomatidae) in Spain and a discussion on the relationships within the ‘revolutum’ group based on cercarial chaetotaxy. *Syst. Parasitol.* 45 (3), 199.
- Wang, Y.Q., Yu, H.X., 1993. The discovery of *Echinostoma miyagawai* in goose in Heilongjiang province. *Chinese Vet. Sci.* 7, 44 (in Chinese).
- Wang, Y., Wang, C.R., Zhao, G.H., Gao, J.F., Li, M.W., Zhu, X.Q., 2011. The complete mitochondrial genome of *Orientobilharzia turkestanicum* supports its affinity with African *Schistosoma* spp. *Infect. Genet. Evol.* 11 (8), 1964–1970.
- Webster, B.L., Rudolfová, J., Horák, P., Littlewood, D.T., 2007. The complete mitochondrial genome of the bird schistosome *Trichobilharzia regenti* (Platyhelminthes: Digenea), causative agent of cercarial dermatitis. *J. Parasitol.* 93 (3), 553–561.
- Yan, H.B., Wang, X.Y., Lou, Z.Z., Li, L., Blair, D., Yin, H., Cai, J.Z., Dai, X.L., Lei, M.T., Zhu, X.Q., Cai, X.P., Jia, W.Z., 2013. The mitochondrial genome of *Paramphistomum cervi* (Digenea), the first representative for the family Paramphistomidae. *PLoS One* 8 (8), e71300.
- Yang, X., Gasser, R.B., Koehler, A.V., Wang, L., Zhu, K., Chen, L., Feng, H., Hu, M., Fang, R., 2015. Mitochondrial genome of *Hypoderaeum conoideum*-comparison with selected trematodes. *Parasit. Vectors* 8, 97.
- Young, N.D., Chan, K.G., Korhonen, P.K., Chong, T.M., Ee, R., Mohandas, N., Koehler, A.V., Lim, Y.L., Hofmann, A., Jex, A.R., Qian, B., Chilton, N.B., Gobert, G.N., McManus, D.P., Tan, P., Webster, B.L., Rollinson, D., Gasser, R.B., 2015. Exploring molecular variation in *Schistosoma japonicum* in China. *Sci. Rep.* 5, 17345.