Vietnam Journal of Biotechnology **20**(1): 15-23, 2022

CHARACTERISTICS OF THE NUCLEOTIDE COMPOSITION AND *DE NOVO* STRUCTURE OF 28S rRNA GENE OF INTESTINAL FLUKES (ECHINOSTOMATIDAE) AND GENETIC DISTANCE AMONG SPECIES (SUBORDER ECHINOSTOMATA)

Pham Thi Khanh Linh[1,2], Nguyen Thi Khue[1,2], Doan Thi Thanh Huong[1,2], Le Thanh Hoa[1,2,✉]

[1]*Institute of Biotechnology, Vietnam Academy of Science and Technology , 18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam*
[2]*Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam*

✉To whom correspondence should be addressed. E-mail: imibtvn@gmail.com

Received: 09.6.2021
Accepted: 17.9.2021

SUMMARY

The family Echinostomatidae includes four genera, *Echinostoma*, *Hypoderaeum*, *Echinoparyphium*, and *Artyfechinostomum*, which are of public epidemiological and pathogenic importance. The morphological similarity of species requires differentiating methods, and the use of molecular genetic markers is mainly encountered. The 28S rRNA gene sequence in the family Echinostomatidae is an important marker for the analysis of the interspecies and phylogenetic relationships. However, information about the structure and composition of the complete 28S rRNA gene in Echinostomatidae is limited. In this study, the complete 28S rRNA sequences (3,881–3,883 bp) of four species, including *Echinostoma revolutum* (designated Erev-MSD15-TH), *E. malayanum* (Emal-EMI3-TH), *E. miyagawai* (Emiy-RED11-TH), and *Hypoderaeum conoideum* (Hcon-RED42-TH) were obtained and analyzed for their genetic characteristics. All had complex secondary structures, containing numerous branches, "hairpins", and "loops" that provide stability of the rRNA gene. The *de novo* structural pattern and the "*skew*" value in the nucleotide use of *E. revolutum*, *E. malayanum*, *E. miyagawai*, and *H. conoideum* were analyzed. The nucleotide usage of A, T, G, C and A+T and G+C of 28S rRNA in all species of the family Echinostomatidae was similar, in which A and T usage were almost equal, and G and C were slightly different in the range of 7–8%, so the skewness for AT was very low negative value (–0.029 to –0.045), and for GC was not very high (0.163 to 0.172). The genetic distance estimated among strains within a species (intra-specific level) was very low in both the Echinostomatidae (0.05%) and the Fasciolidae families (0–0.03%), and that among different species (inter-specific level) was higher. However, among species in different families (inter-familial level) the genetic distance was highest, at 3.80–4.86%, as seen between *Echinostoma* and *Fasciola* species. The 28S rRNA gene dataset obtained from four species of the intestinal flukes (family Echinostomatidae) in this study contribute to the understandings of genetic markers in the suborder Echinostomata for use in diagnostic, phylogenetic, taxonomic, epidemiological, and population genetics studies.

Keywords: 28S rRNA, *Echinostoma*, genetic distance, intra-specific, inter-specific, inter-familial level, secondary structure, skew/skewness

INTRODUCTION

The family Echinostomatidae includes four genera, ie., *Echinostoma*, *Hypoderaeum*, *Echinoparyphium*, and *Artyfechinostomum*, which are of widespread pathogenicity and

epidemiological importance in the world and in Vietnam. The five groups of the genus *Echinostoma* have characteristic morphology, which can be distinguished by examination of the structure, arrangement and number of "collar-spines" around the pharynx (Saijuntha *et al*., 2011a, b; Toledo, Esteban, 2016). The most important group is the "*revolutum*" group, characterized by the arrangement of 37-collar-spines found on cercariae (Kostadinova, 2005; Georgieva *et al*., 2014; Chai, 2019), typically, *Echinostoma revolutum* (Fröhlich, 1802) and eight other species, *E. caproni, E. echinatum, E. friedi, E. judini, E. miyagawai, E. paraensei, E. parvocirrus* and *E. trivolvis*. Other groups have variable number of "collar" spines, such as 25–29 on *E. hortense*, 43 on *E. malayanum*, 41–45 on *Hypoderaeum conoideum* and 43–50 on *Echinoparyphium recurvatum* (Chai, 2009; 2019; Saijuntha *et al*., 2011a). However, the morphological similarity of the species requires additional methods for discrimination, mainly molecular techniques using genetic markers (Saijuntha *et al*., 2011c; 2011a; Georgieva *et al*., 2014; Tkach *et al*., 2016).

One of the most important genomic datasets is the ribosomal transcription unit (rTU), present in the nuclear genome of the cell which is needed to be explored (Blair, 2006). In the nuclear genome of animal cells, hundreds of rTUs arranged in arrays/operons are located in the nuclear organizer region (NOR), and in humans in the chromosomes 13, 14, 15, 21 and 22 (McStay, 2016; Potapova, Gerton, 2019). Each unit, about 7–9 kb in length in trematodes, contains three coding genes (18S, 5.8S, and 28S rRNA) separated by two internal transcribed spacers, ITS-1 and ITS-2, respectively; and units are connected by nucleotide chains, containing many repetitive structures, called the non-transcribed intergenic spacers (IGS). The rTU-specific structure is as in the line of IGS-ETS-18S-ITS1-5.8S-ITS2-28S-IGS, one followed after another. Any of the ribosomal genes (18S rRNA, 28S rRNA) or spacer regions (ITS-1, ITS-2) can be used in taxonomic analysis and molecular evolutionary relationships of species of the same and closely related species, and from the distant origin (Weider *et al*., 2005; Blair, 2006).

Sequences of the 28S rRNA gene in many organisms, including those of the family Echinostomatidae, are important for the analysis of taxonomic and phylogenetic relationships (Olson *et al*., 2003; Tkach *et al*., 2016). Along with the 18S rRNA gene, the partial (approximately 1,200 bp, of D1–D3 domain) or full (ranging 3.7–3.9 kb) 28S rRNA sequence has been increasingly used as a powerful genetic marker (Blair, 2006; Georgieva *et al*., 2014; Tkach *et al*., 2016; Le *et al*., 2020). However, understandings about the structure and composition of the complete 28S rRNA in Echinostomatidae is still limited.

This paper presents the results from the analysis of the secondary structure and nucleotide usage, as well as the estimation of the genetic distances between *Echinostoma* spp. species, Echinostomatidae and Fasciolidae in the suborder Echinostomata.

MATERIALS AND METHODS

Samples and DNA extraction

The samples used in this study were adult intestinal flukes that were morphologically identified and validated by molecular identification. Due to the lack of Vietnamese samples of the genus *Echinostoma*, we have obtained *Echinostoma* samples from our international partner (Thailand) including *Echinostoma revolutum* (strain Erev-MSD15-TH); *E. malayanum* (Emal-EMI3-TH), *E. miyagawai* (Emiy-RED11-TH), and *Hypoderaeum conoideum* (Hcon-RED42-TH), forwarded by Assoc. Prof. Dr. Weerachai Saijuntha, Mahasarakham University, Khon Kaen, Thailand. Samples were in fresh frozen form, either in 70% alcohol or in total DNA, stored at –20°C.

DNA was extracted as total genomic DNA including nuclear and mitochondrial genomes, purified using the GeneJET™ Genomic DNA Purification Kit (Thermo Scientific Inc., MA,

USA) by instructions from the manufacturer. DNA content was quantified by a NanoDrop® ND-1000 UV-Vis Spectrophotometer to determine the DNA template for PCR reaction. An amount of 2 µL (50 ng/µL) was used in each PCR of 50 µL volume.

Primer design and PCR amplification

To obtain the complete 28S rRNA sequence of the selected species, the majority of the rTU universal primers were used as previously reported (Le *et al*., 2017; 2020). These primer pairs were designed to capture different PCR products. The PCR Master Mix kit (Thermo Scientific) was used to perform PCR. The PCR reaction mixture had a total volume of 50 µL, consisting of 25 µL PCR master mix (Thermo Scientific), 2 µL each primer (10 pmol/µL), 3 µL DNA template, 2 µL DMSO (dimethyl sulfoxide), and 16 µL DEPC deionized water. PCR amplification reaction was performed on MJ PTC-100 (USA), and the thermal cycles included 1 cycle at 95ºC/5 min, 35 cycles at [94ºC/1 min, 52ºC/3 min; 72ºC/2 min], last cycle at 72ºC/10 min. The PCR products (10 µL of each) were examined on a 1% agarose gel, stained with ethidium bromide and visualized under UV light (Wealtec, Sparks, NV, USA). Single-band PCR products were purified using the GeneJET PCR Purification Kit or isolated the PCR product DNA bands of the intended size (if multiple bands present) using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific Inc., MA, USA) and sent for sequencing at service facilities (Macrogen Inc., Seoul, Korea or Nam Khoa Biotek, Ho Chi Minh City, Vietnam).

Data processing

The entire rDNA sequence for each Echinostomatidae species/isolate was obtained after editing chromatograms using Chromas 2.6.6 (http://technelysium.com.au/wp/chromas/) and the 28S rRNA gene was determined by using the previously published reference sequences and those available in GenBank (Zheng *et al.*, 2014; Briscoe *et al.*, 2016; Qian *et al.*, 2018; Su *et al*., 2018) and from our publications (Le *et al*., 2017; 2020; Le Thanh Hoa *et al*., 2019; Le Thi Viet Ha *et al*., 2020).

Determination of the secondary structure of the 28S rRNA gene

The nucleotide sequence containing many regions rich in A and T, or G and C nucleotides that symmetrically run in opposite directions giving rise to pair up for formation of "hairpin" and "loop" structures. These secondary structures maintain gene stability for rRNAs in the cellular ribosomes (Caburet *et al*., 2005). The secondary structure of the 28S rRNA gene of four species of Echinostomatidae was modeled *de novo* using the RNAfold program available at http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/ with the minimum free energy (MFE) of –437.80 kcal/mol (Mathews *et al*., 2004).

Estimation of skewness and genetic distance

After being obtained, the total 28S rRNA sequences of 4 strains/4 species in this study were used to access the GenBank through the BLAST program to collect similar sequences of the same or different genera and the same or different families. These sequences were used for the analysis of taxonomic and phylogenetic and genetic comparison.

Twelve (12) full-length 28S rDNA sequences (~3.8–3.9 kb) (Table 1) were aligned using GENEDOC2.7, confirmed with the MAFFT 7.122 program (Katoh, Standley, 2013), then was extracted as a fasta format (.fasta). This alignment of about 3,615–3,687 nucleotides were included in the phylogenetic analysis using MEGA X using the Maximum Composite Likelihood (MCL) method with 1000 bootstrap resamplings (Kumar *et al*., 2018). The completed 28S sequences used were from strains/species of the suborder Echinostomata, including the families Echinostomatidae and Fasciolidae (Table 1). The AT and GC skewness values (from −1 to +1) were calculated according to the formula as [AT skew = (A – T)/(A + T) and GC skew = (G – C)/(G + C)] (Perna, Kocher, 1995).

Genetic distance is the estimated percentage (%) of difference between sequences representing each species or each strain within a species when pairwise comparisons were performed. The alignment block of the above 12 sequences was imported into MEGA X, then genetic distance was computed using the MCL method and pairwise distances were saved in a table. The difference rate (%) between strains/species in the same family (Echinostomatidae) and in the different families (Echinostomatidae and Fasciolidae) were estimated.

Table 1. List of 12 strains/10 species providing complete 28S rRNA nucleotide sequences used to evaluate taxonomic relationships and estimate genetic distances of species in the suborder Echinostomata (Platyhelminthes: Trematoda)

	Family/Species*	Length (bp)	Strain designation	Country	GenBank accession nos
	Echinostomatidae (5)/6				
1	*Echinostoma malayanum*	3863	Emal-EMI3-TH	Thailand	This study
2	*Echinostoma miyagawai*	3861	Emiy-RED11-TH	Thailand	This study
3	*Echinostoma miyagawai*	3861	Emiy-HLJ-CN	China	This study
4	*Echinostoma revolutum*	3863	Erev-MSD15-TH	Thailand	This study
5	*Hypoderaeum conoideum*	3863	Hcon-RED42-TH	Thailand	This study
6	*Isthmiophora hortensis*	3862	Ihor-Waka-JP	Japan	AB189982
	Fasciolidae (5)/6				
7	*Fasciola gigantica*	3863	Fgig-NB-VN	Vietnam	MN970009
8	*Fasciola gigantica*	3863	Fgig-T4V-VN	Vietnam	MN970010
9	*Fasciola hepatica*	3863	Fhep-Geelong-AU	Australia	MN970007
10	*Fasciola* sp. (hybrid)	3863	Fsp(hyb)-DL11-VN	Vietnam	MN970008
11	*Fasciola (Fascioloides) jacksoni*	3863	Fjac-Madu-LK	Sri Lanka	MN970006
12	*Fasciolopsis buski*	3862	Fbus-HT-VN	Vietnam	MN970005

*In parentheses is the species number, followed by a slash is the number of strains in a family used to provide the 28S sequence for estimation of genetic distance.

RESULTS AND DISCUSSION

Secondary structure of the 28S rRNA gene for four species in the family Echinostomatidae

The complete nucleotide sequence of the 28S rRNA gene was obtained from three species of the genus *Echinostoma*, including *Echinostoma revolutum* with 3,863 bp; *E. malayanum* with 3,863 bp; *E. miyagawai* with 3,861 bp; and a species of the genus *Hypoderaeum*, ie., *Hypoderaeum conoideum* with 3,863 bp in length. To consider the secondary structure, the 1,250 nucleotide sequence of the D1–D3 variable domain of the 28S rRNA gene from each species was modeled *de novo* using the RNAfold software. The results are shown in Fig. 1.

The secondary structure in Fig. 1 shows that the 28S rRNA ribosomal gene of *E. revolutum* is also divided into two branches, but overall, it is slightly different from those of *E. malayanum*, *E. miyagawai* and *H. conoideum* species. The latter three species have a very high level of secondary structure similarity, regarding the conformation of hairpins and loops in both branches, I and II. The 'hairpin' model is formed from nucleotide sequences with complementary symmetric sequences, also known as 'palindromic' sequences (Caburet *et al.*, 2005). The arrangement of GC and AT nucleotides in the

28S rRNA sequences of these three species (*E. malayanum, E. miyagawai* and *H. conoideum*) is totally the same, although *H. conoideum* belongs to the different genus *Hypoderaeum*. Although *E. revolutum* belongs to the genus *Echinostoma*, the 28S rRNA secondary structure of this species has a slight difference that branch I contains more hairpins (up to 15 hairpins) than 7–8 hairpins found in the rest of three species above mentioned (Fig. 1).

The secondary structure is the structural formation resulted from the pairing of complementary sequences between A and T and between G and C running on opposite directions, creating hairpins and loops usually found in ribosomal RNA conformation for sustaining the gene stability (Caburet *et al*., 2005). This is to be allowed to infer that the *de novo* formation of the secondary structure in the rRNA molecules does not depend on nucleotide composition but on the ability to pair with when the opposite strands of AT and GC allowing to use the favorable binding energy for the structures to form (Fig. 1). The formation of this folded structure of ribosomal rRNA (here, for 28S rRNA gene) is due to the fact that ribosomal RNA belongs to the structural gene family which requires a stable conformational folding position in cellular ribosomes with the most saturated binding force (Caburet *et al*., 2005). Such secondary structures are found not only in 28S rRNA or 18S rRNA but also in the ITS-1 or ITS-2 intergenic spacers in ribosomal transcription units in any species, such as observed in ITS-1/ITS-2 of the lung flukes, *Nanophyetus* spp. (family Paragonimidae) recently reported (Voronova, Chelomina, 2018).

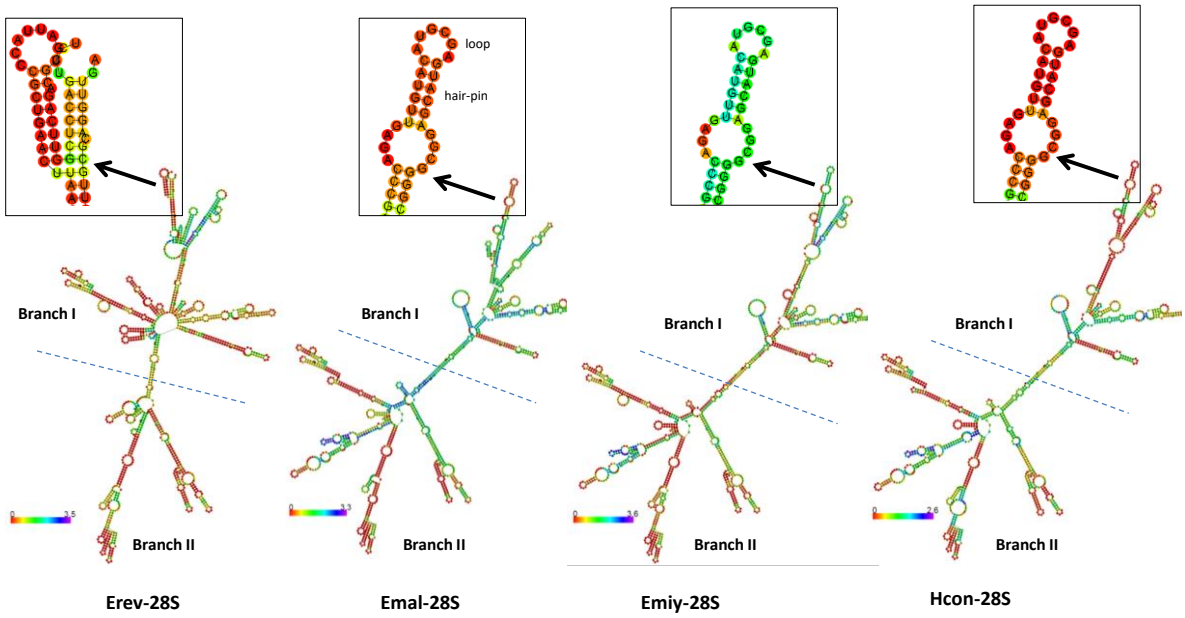

Figure 1. The *de novo* secondary structure model of the 28S rRNA gene (domains D1–D3) was modelled and constructed in the RNAfold program, with minimum free energy (MFE) of –437.80 kcal/mol. Structures containing "hairpin" sequences (or stem-loop) are formed from opposing nucleotide sequences (palindromic sequences); and "loop" sequences at the end of each branch (illustrated in the box at the top). The structures are divided into two groups: Branch I and Branch II by a dashed line.

Characteristics of the complete 28S rRNA gene of species in the family Echinostomatidae

For ribosomal transcription units, in addition to structural conformation, some features of composition and characteristics in the 28S rRNA gene have an important role in applied research, comparative analysis, and species identification. Therefore, we compared the nucleotide usage characteristics and the "skew" value of this gene

in 4 species/strains and some species in Echinostomatidae and Fasciolidae (suborder Echinostomata) from GenBank (Table 2).

Table 2 shows the usage of nucleotide composition A, T, G, C and the "skew/skewness" values of A+T and G+C in the 28S rRNA gene in 6 strains/5 species of the Echinostomatidae family, including *E. malayanum* (EMI3 strain, Thailand); *E. miyagawai* (strain RED11, Thailand); *E. miyagawai* (strain HLJ, China): *E. revolutum* (strain MSD15, Thailand); *H. conoideum* (strain RED42, Thailand); and *I. hortensis* (strain Waka, Japan).

Table 2 shows the ribosomal transcription units of 6 strains/5 species of the genera, *Echinostoma*, *Hypoderaeum,* and *Isthmiophora*, using 3,861–3,863 nucleotides to construct the 28S rRNA gene. The nucleotide usage for each component A, T, G, C, and overall, A+T and G+C, in all species of the Echinostomatidae family was similar, with no considerable difference. Specifically, the rate of using nucleotide A = 22.42–22.85%; T = 24.12–25.85%; G = 30.83–30.92%; C = 21.83–22.18%, with overall components A+T = 46.97–47.25% and G+C = 52.75–53.03%. Since T is used less than A, the skew/skewness at A+T of 28S rRNA is negative (–0.029 to –0.045), which is a very low negative deviation. The rate of using G is higher than C in the range of 7–8%, so the G+C skewness was also at a low positive value (0.163–0.172) in all 6 strains/5 species of the Echinostomatidae family in the study (Table 2).

Table 2. The ratio of using nucleotide compositions A, T, G, C and skew/skewness values of the total 28S rRNA gene of intestinal flukes of the family Echinostomatidae

	Species/Strain	Gene	Length (bp)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
ECHINOSTOMATIDAE											
1	*E. malayanum* (Emal-EMI3-TH)	28S	3863	22.66	24.42	30.92	22.01	47.08	–0.037	52.92	0.168
2	*E. miyagawai* (Emiy-RED11-TH)	28S	3861	22.59	24.40	30.83	22.18	46.99	–0.039	53.01	0.163
3	*E. miyagawai* (Emiy-HLJ-CN)	28S	3861	22.59	24.46	30.83	22.12	47.05	–0.040	52.95	0.165
4	*E. revolutum* (Erev-MSD15-TH)	28S	3863	22.42	24.55	30.89	22.14	46.97	–0.045	53.03	0.165
5	*H. conoideum* (Hcon-RED42-TH)	28S	3863	22.64	24.61	30.92	21.83	47.25	–0.042	52.75	0.172
6	*I. hortensis* (Ihor-Waka-JP)	28S	3862	22.85	24.20	30.89	22.07	47.05	–0.029	52.95	0.167

Note: A: Adenine; T: Thymine; G: Guanine; C: Cytosine. Skew/skewnes: value of using nucleotides. Information of species/strains are listed in Table 1.

One of the most interesting genomic features of a gene is to consider the way in which nucleotides are used to build the gene. For ribosomal transcription units, in addition to structural conformation, some features of composition and characteristics in the 28S rRNA gene have an important role in applied research, comparative analysis, and species identification. The skewness showed that the value was not much different between A and T as well as between G and C, or in other words, the number of individual nucleotides used on both strands was almost equal. This status facilitated additional bonding between the two, A versus T and G versus C nucleotides, and thus, the possibility of complementary pairing in the

secondary structure favorably occurred in many regions of the 28S rRNA gene. The favorable AT and GC pairings resulted in numerous "hairpin" structures throughout the sequence. The nucleotide usage of A+T and G+C components in the 28S rRNA gene of the two species, *E. malayanum* and *E. miyagawai* is similar (Table 2), so overall, the secondary, folding structure model of these two species are almost exactly the same (Fig. 1).

Genetic distance estimated from species of Echinostomatidae and Fasciolidae

Twelve 28S rRNA nucleotide sequences from 12 strains of 10 trematode species (listed in Table 1) were aligned using GENEDOC 2.7, and then processed to obtain the final aligned block using the MEGA X program (Kumar *et al*., 2018). From the final alignment, pairwise genetic distance was calculated for 6 strains/5 species of family Echinostomatidae and 6 strains/5 species of family Fasciolidae according to MCL parameters included in the MEGA X program (Table 3).

The results (Table 3) showed that the intraspecific genetic distance within *Echinostoma miyagawai* (strains Emiy-RED11-TH and Emiy-HLJ-CN) is very low, only 0.05%; within "*Echinostoma*" group (strains/species of *Echinostoma* spp.) is low, ranging from 0.52–1.66%; while between *Echinostoma* spp. and *Hypoderaeum conoideum*, the inter-specific rate has a higher value (0.152–1.71%); and between *Echinostoma* spp. and *Isthmiophora hortensis* was the highest (3.33–4.02%).

The lowest genetic distance was also found within the *Fasciola gigantica* strains (0.03%, among strains Fgig-NB-VN and Fgig-T4V-VN); or within *Fasciola hepatica* (0%, among Fsp-DL11-VN and Fhep-GL-AU). The highest rate was found between *Isthmiophora hortensis* (family Echinostomatidae) and the liver flukes, *Fasciola* spp. (family Fasciolidae) and the large intestinal fluke, *Fasciolopsis buski* (family Fasciolidae), reaching 4.22–4.86% (Table 3).

Table 3. Pairwise genetic distance (%) between strains/species in suborder Echinostomata consisting of 6 strains/5 species of family Echinostomatidae and 6 strains/5 species of Fasciolidae based on analysis of 28S rRNA nucleotide sequence (~3,881 bp).

	Strain/Species	1	2	3	4	5	6	7	8	9	10	11
1	Emal-EMI3-TH											
2	Emiy-HLJ-CN-MH748722	1.58										
3	Emiy-RED11-TH	1.58	**0.05**									
4	Erev-MSD15-TH	1.66	0.52	0.52								
5	Hcon-RED42-TH	1.71	1.52	1.52	1.66							
6	Ihor-Waka-JP-AB189982	4.02	3.74	3.74	3.85	3.33						
7	Fgig-NB-VN-MN970009	3.55	3.63	3.63	3.64	3.36	4.36					
8	Fgig-T4V-VN-MN970010	3.58	3.66	3.66	3.66	3.39	4.39	**0.03**				
9	Fsp-DL11-VN-MN970008	3.55	3.64	3.64	3.64	3.36	4.36	0.10	0.13			
10	Fhep-GL-AU-MN970007	3.55	3.64	3.64	3.64	3.36	4.36	0.10	0.13	**0.0**		
11	Fjac-Madu-LK-MN970006	3.66	3.58	3.58	3.69	3.33	4.22	0.84	0.86	0.84	0.84	
12	Fbus-HT-VN-MN970005	4.05	3.94	3.91	3.94	3.80	4.86	2.65	2.68	2.65	2.65	2.33

Note: Full species names and strain information are listed in Table 1. Some intraspecific divergence (%) of *Echinostoma miyagawai*, *Fasciola gigantica* and *F. hepatica* were bolded. The boxed part represents the highest difference between *Isthmiophora hortensis* (Echinostomatidae) and members of Fasciolidae.

Thus, the intraspecific genetic distance between strains within a species is very low not only among strains of *Echinostoma* sp. in the family Echinostomatidae, but also among strains/species in the family Fasciolidae (0–0.86%), while that the interspecific and inter-familial genetic distance between Echinostomatidae and Fasciolidae is relatively high (3.80–4.86%), as well as between *Echinostoma* spp. (Echinostomatidae) and *Fasciolopsis buski* (Fasciolidae) found in these two families (Table 3).

CONCLUSION

The complete 28S rRNA gene nucleotide sequences (3,881–3,883 bp) of the ribosomal transcription unit for four intestinal flukes, *E. malayanum*, *E. miyagawai*, *E. revolutum*, and *H. conoideum*, were obtained and analyzed for their secondary structure, nucleotide composition and usage and genetic distance compared with 8 other strains/species of the families Echinostomatidae and Fasciolidae. With the exception of *E. revolutum*, the secondary structure and nucleotide skew values of *E. malayanum*, *E. miyagawai* and *H. conoideum* are relatively similar. The skewness value for AT was negative and very low (–0.029 to –0.045), for G+C was positive and not high (0.163–0.172). The intraspecific genetic distance (intra-specific level) was very low (0.05% within strains of *Echinostoma* spp. and 0–0.03% within strains of Fasciolidae (0–0.03%), the interspecific genetic distance (inter-specific level) within a genus was higher, and the inter-familial genetic distance (inter-familial level) was highest (3.80–4.86%) as seen within strains and among species of the families Echinostomatidae and Fasciolidae. Determining the nucleotide sequence, characterization of the genetic and structural formation in the 28S rRNA gene from four *Echinostoma*/*Hypoderaeum* species of intestinal flukes (Echinostomatidae) in this study contribute to providing the data for this marker in the suborder Echinostomata for use in diagnostic, phylogenetic, taxonomic, epidemiological and population genetics studies.

Acknowledgements: *This work was funded by the Supporting Program for the Principal Research Scientists from the Vietnam Academy of Science and Technology (VAST) to Prof. Dr. Thanh Hoa Le (Grant No: NVCC08.07/21-21).*

REFERENCES

Blair D (2006) Ribosomal DNA variation in parasitic flatworms. In: Maule A, editor. Parasitic Flatworms: Molecular Biology, Biochemistry, Immunology and Control. CAB International pp. 96–123.

Blair D, Nawa Y, Mitreva M, Doanh PN (2016) Gene diversity and genetic variation in lung flukes (genus *Paragonimus*). *Trans R Soc Trop Med Hyg* 110(1): 6–12.

Briscoe AG, Bray RA, Brabec J, Littlewood DTJ (2016) The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale. *Parasitol Int* 65(3): 271–275.

Caburet S, Conti C, Schurra C, Lebofsky R, Edelstein SJ, Bensimon A (2005) Human ribosomal RNA gene arrays display a broad range of palindromic structures. *Genome Res* 15(8): 1079–1085.

Chai JY (2009) Echinostomes in humans. In: Toledo, R., Fried, B. (Eds.), The Biology of Echinostomes: From the Molecular to the Community. Springer, New York.

Chai JY (2019) Human Intestinal Flukes. Chapter 2. Echinostomes. The Netherlands. Springer Nature B.V. pp. 169–343.

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

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772–780.

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

Kumar S, Stecher G, Li M, Knyaz C, Tamura K

(2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35: 1547–1549

Le TH, Nguyen KT, Nguyen NT, Doan HT, Dung DT, Blair D (2017) The ribosomal transcription units of *Haplorchis pumilio* and *H. taichui* and the use of 28S sequences for phylogenetic identification of common heterophyids in Vietnam. *Parasit Vectors* 10: 17.

Le TH, Pham KLT, Doan HTT, Le TKX, Nguyen KT, Lawton SP (2020) Description and phylogenetic analyses of ribosomal transcription units from species of Fasciolidae (Platyhelminthes: Digenea). *J Helminthol* 94: e136.

Le Thanh Hoa, Nguyen Thi Bich Nga, Doan Thi Thanh Huong, Le Thi Kim Xuyen, Nguyen Thi Khue (2019) Structure and characterization of the ribosomal transcription units of small liver flukes, *Opisthorchis viverrini*, *O. felineus* and *Clonorchis sinensis*. *Vietnam Journal of Biotechnology* 17(3): 561–567.

Le Thi Viet Ha, Nguyen Thi Khue, Dong Van Quyen, Thanh Hoa Le (2020) Presentation and phylogenetic analyses of ribosomal transcription units from *Haplorchis taichui* and *H. pumilio* species of Heterophyidae (Platyhelminthes: Opisthorchiata). *Vietnam Journal of Biotechnology* 18(4): 643–652.

Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci U S A* 101(19): 7287–7292.

McStay B (2016) Nucleolar organizer regions*:* genomic *'*dark matter*'* requiring illumination. *Genes and Development* 30(14): 1598–1610.

Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* 41(3): 353–358.

Potapova TA, Gerton JL (2019) Ribosomal DNA and the nucleolus in the context of genome organization. *Chromosome Res* 27(1-2): 109–127.

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

Saijuntha W, Sithithaworn P, Duenngai K, Kiatsopit N, Andrews RH, Petney TN (2011b) Genetic variation and relationships of four species of medically important echinostomes (Trematoda: Echinostomatidae) in South-East Asia. *Infect Genet Evol* 11(2): 375–381.

Saijuntha W, Tantrawatpan C, Sithithaworn P, Andrews RH, Petney TN (2011a) Spatial and temporal genetic variation of *Echinostoma revolutum* (Trematoda: Echinostomatidae) from Thailand and the Lao PDR. *Acta Trop* 118(2): 105–109.

Su X, Zhang Y, Zheng X, Wang XX, Li Y, Li Q, Wang CR (2018) Characterization of the complete nuclear ribosomal DNA sequences of *Eurytrema pancreaticum*. *J Helminthol* 92: 484–490.

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

Toledo R, Esteban JG (2016) An update on human echinostomiasis. *Trans R Soc Trop Med Hyg* 110(1): 37–45. Review.

Voronova A, Chelomina GN (2018) Genetic diversity and phylogenetic relations of salmon trematode *Nanophyetus japonensis*. *Parasitol Int* 67(3): 267–276.

Weider LJ, Elser JJ, Crease TJ, Mateos M, Cotner JB, Markow TA (2005) The functional significance of ribosomal rDNA variation: Impacts on the evolutionary ecology of organisms. *Annu Rev Ecol Evol Syst* 36: 219–242.

Zheng X, Chang QC, Zhang Y, Tian SQ, Lou Y, Duan H, Guo DH, Wang CR, Zhu XQ (2014) Characterization of the complete nuclear ribosomal DNA sequences of *Paramphistomum cervi*. *Scientific World J* 2014: 751907.