University of Science and Technology of Hanoi

Doctoral School

Study on the genomics of the mitochondrial genome and ribosomal transcription units of some intestinal flukes in the family Echinostomatidae of the suborder Echinostomata

PhD student: Pham Thi Khanh Linh

Department: Pharmacological, Medical, and Agronomical Biotechnology

The thesis has been successfully defended on April 26, 2025 in front of jury composed of:

Pro. Nguyen Linh Toan, 103 Military Hospital, Jury chairman

Assoc. Prof. Truong Dinh Hoai, Vietnam National University of Agriculture, Reviewer

Prof. Nicolas Papon, University of Angers, Reviewer

Assoc. Prof. Nguyen Manh Hung, Institute of Biology, Reviewer

Assoc. Prof. Kim Thi Phuong Oanh, Institute of Biology, Jury Member

Dr. Nguyen Quang Huy, USTH, Jury Member

Dr. Tran Thi Thanh Tam, USTH, Jury Member, Secretary

Prof. Le Thanh Hoa, Institute of Biology, Supervisor

Assoc. Prof. Dong Van Quyen, Institute of Biology, Supervisor

Approval of supervisors:

Prof. Le Thanh Hoa Dong Van Quyen

Hanoi, July 30, 2025

STATEMENT OF ORIGINALITY

To the best of my knowledge and belief, the work presented in this thesis is original except as acknowledged in the text. The work is entirely of my own contribution under the direction of my supervisors and due reference is given for any assistance made. I hereby declare that I have not submitted this material either in whole or in part, for a degree at this or any other institution.

Hanoi, July 30, 2025

Pham Thi Khanh Linh

ACKNOWLEDGEMENTS

The successful and highly productive research work presented in this thesis, could not have been done without the exceptional supervision of my supervisors. First of all and formost, I am very grateful to my principal supervisor, Prof. Dr. Le Thanh Hoa, Institute of Biology (IB, formerly, Institute of Biotechnology), Vietnam Academy of Science and Technology (VAST), for his encouragement, guidance, suggestion, support, advice and wholehearted assistance, and Associate Prof. Dr. Dong Van Quyen, my co-supervisor, Institute of Biology (IB) and University of Science and Technology of Hanoi (USTH), Vietnam Academy of Science and Technology, for his intellectual support and wholehearted assistance during my research work and during the completion of my thesis.

Great gratitude is expressed to all members of the Immunology Department, Institute of Biotechnology (currently, Biochemistry and Immunology Department Institute of Biology), Vietnam Academy of Science and Technology, especially Dr. Doan Thi Thanh Huong, Dr. Nguyen Thi Khue, and MSc. PhD student Do Thi Roan for their close laboratory collaboration. I am grateful for the international collaborators, Assoc. Prof. Weerachai Saijuntha, Faculty of Medicine, Mahasarakham University (Thailand), Prof. David Blair, College of Science and Engineering, James Cook University, Townsville, Queensland (Australia), and Dr. Scott P. Lawton, Department of Veterinary and Animal Sciences, Northern Faculty, Scotland's Rural College (United Kingdom) for providing echinostome materials for my studies and for their invaluable supervision of my work and my publications.
I am also grateful to the professors at the University of Science and Technology of Hanoi. Being a student over the years, I gained not only knowledge but also passion to investigate. My gratude goes to the Doctoral School for their excellent advice and assistance throughout my PhD program.

Special thanks are due to the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant number 108.02–2020.07 and the Supporting Program for the Principal Research Scientists from the Vietnam Academy of Science and Technology Grant No NVCC08-05/22–23 (to Prof. LE TH) for providing financial support.

This research has taken away from my time with my family. My family's understanding and support for me during this time cannot be expressed in a few words, but it is felt very deeply. Thank you for always being with me.

Pham Thi Khanh Linh

LIST OF PUBLICATIONS

A. Direct publications of the Thesis

1. **Pham LTK**, Quyen DV, Saijuntha W, Doan HTT, Le TH, Scott P. Lawton* (2024). Mitogenomics of the zoonotic parasite *Echinostoma miyagawai* and insights into the evolution of tandem repeat regions within the mitochondrial non-coding control region**.** Parasitology 151:1543–1554. (ISSN: 0031-1820; E-ISSN: 1469-8161; SCI/IF2023=2.1). https://doi.org/10.1017/S0031182024001422.
2. Le TH*, **Pham LTK**, Quyen DV, Nguyen KT, Doan HTT, Saijuntha W, Blair D (2024). The ribosomal transcription units of five echinostomes and their taxonomic implications for the suborder Echinostomata (Trematoda: Platyhelminthes). Parasitology Research 123(1):103. (ISSN: 0932-0113; E-ISSN: 1432-1955; SCI/IF2023=2.383). https://doi.org/10.1007/s00436-023-08110-z
3. **Pham KLT**, Saijuntha W, Lawton SP, Le TH* (2022). Mitophylogenomics of the zoonotic fluke *Echinostoma malayanum* confirms it as a member of the genus *Artyfechinostomum* Lane, 1915 and illustrates the complexity of Echinostomatidae systematics. Parasitology Research 121:899–913. (ISSN: 0932-0113; E-ISSN: 1432-1955; SCI/IF2022=2.11). https://doi.org/10.1007/s00436-022-07449-z
4. 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 Echinostmata (Platyhelminthes: Digenea). Parasitology 147:566–576. (ISSN: 0031-1820; E-ISSN: 1469-8161; SCI/IF2020=3.81). https://doi.org/10.1017/S0031182020000128
5. Le Thanh Hoa*, **Pham Thi Khanh Linh**, Nguyen Thi Khue, Do Thi Roan, Le Thi Kim Xuyen, Doan Thi Thanh Huong (2022). Genetic distance and phylogenetic relationships of some *Echinostoma* species (*E. malayanum*, *E. revolutum*, *E. miyagawai*) and *Hypoderaeum conoideum* (family Echinostomatidae) inferred from partial 28S rDNA sequence analysis. Vietnam Journal of Biotechnology 20(2):253–263 (ISSN: 2815-5955; E-ISSN: 2815-5912). https://doi.org/10.15625/1811-4989/16903
6. **Pham Thi Khanh Linh**, Nguyen Thi Khue, Doan Thi Thanh Huong, Le Thanh Hoa* (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). Vietnam Journal of Biotechnology 20(1):15–23. (ISSN: 2815-5955; E-ISSN: 2815-5912). https://doi.org/10.15625/1811-4989/16138

B. Additional publications related to the topics of the thesis

7. Nguyen KT, Doan HTT, **Pham TKL**, Do RT, Agatsuma T, Doanh PN, Le TH* (2024). Nuclear ribosomal transcription units in Asian *Paragonimus* species (Paragonimidae: Platyhelminthes): genetic characteristics, polymorphism, and implications for interfamilial phylogeny. Parasitology Research 123(11):368 (ISSN: 0932-0113; E-ISSN: 1432-1955; SCI/IF2023=1.8). https://doi.org/10.1007/s00436-024-08391-y.
8. Le TH*, Nguyen KT, **Pham LTK**, Doan HTT, Do RT, Le XTK, Agatsuma T, Blair D (2023). Mitogenomic and nuclear ribosomal transcription unit datasets support the synonymy of *Paragonimus iloktsuenensis* and *P. ohirai* (Paragonimidae: Platyhelminthes). Parasitology Research 122(7):1531–1544. (ISSN: 0932-0113; E-ISSN: 1432-1955; SCI/IF2023=2.383). https://doi.org/10.1007/s00436-023-07854-y
9. Le TH*, Nguyen KT, **Pham LTK**, Doan HTT, Agatsuma T, Blair D (2022). The complete mitogenome of the Asian lung fluke *Paragonimus skrjabini miyazakii* and its implications for the family Paragonimidae (Trematoda: Platyhelminthes). Parasitology 149(13):1709–1719. (ISSN: 0031-1820; E-ISSN: 1469-8161; SCI/IF2021=3.243). https://doi.org/10.1017/S0031182022001184
10. 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). Journal of Helminthology 94:e136. (ISSN: 0022-149X; E-ISSN: 1475-2697; SCI/IF2020=2.170). https://doi.org/10.1017/S0022149X20000164
11. Rajapakse RPVJ, **Pham KLT**, Karunathilake KJK, Lawton SP, Le TH* (2020). Characterization and phylogenetic properties of the complete mitochondrial genome of *Fascioloides jacksoni* (syn. *Fasciola jacksoni*) support the suggested intergeneric change from *Fasciola* to *Fascioloides* (Platyhelminthes: Trematoda: Plagiorchiida). Infection, Genetics and Evolution 82:104281 (ISSN: 1567-1348; E-ISSN: 1567-7257; SCI/IF2020=3.342) (https://doi.org/10.1016/j.meegid.2020.104281)
12. Nguyen Thi Khue, **Pham Thi Khanh Linh**, Do Thi Roan, Doan Thi Thanh Huong, Pham Ngoc Doanh, Le Thanh Hoa* (2020) Molecular evolutionary relationships of Vietnamese and global pulmonary *Paragonimus* species in the family Paragonimidae and suborder Xiphidiata (Platyhelminthes: Trematoda). Vietnam Journal of Biotechnology 18(4):653–662. (ISSN: 2815-5955; E-ISSN: 2815-5912). DOI: 10.15625/1811-4989/18/4/15673
13. Nguyen Thi Khue, Doan Thi Thanh Huong, **Pham Thi Khanh Linh**, Do Thi Roan, Le Thi Kim Xuyen, Le Thanh Hoa* (2021) Mitogenomic characteristics of the zoonotic lung fluke *Paragonimus heterotremus* (family Paragonimidae) of Vietnam. Proceedings of Vietnam National Conference on Biotechnology 2021. Thai Nguyen University Publishing House (ISBN: 978-604-9987-88-5), pp. 72–78.

TABLE OF CONTENTS

Page

STATEMENT OF ORIGINALITY……. i
ACKNOWLEDGEMENTS ……..……..…… ii
LIST OF PUBLICATIONS ……...…….. iii
TABLE OF CONTENTS …….…... v
ABBREVIATIONS …….……... viii
LIST OF TABLES…… x
LIST OF FIGURES……..…… xii

Chapter 1: Literature review ……...…… 1
1.1 Echinostomes of the Echinostomatidae family ……..…… 2
1.1.1 Classification of the echinostomid flukes of the Echinostomatidae family ……... 2
1.1.2 Lifecycle, epidemiology, and geographical distribution ……. 5
1.2 Mitochondria and mitochondrial genomes in animals and parasites …….. 7
1.2.1 General features of mitochondria …… 7
1.2.2 Mitochondrial genes and genomes ……..…… 9
1.2.2.1 Gene order and nomenclature of genes ……..…….. 9
1.2.2.2 Transfer and mitoribosomal RNAs …….. 10
1.2.2.3 Mitochondrial protein-encoding genes and genetic codes …….. 12
1.2.2.4 Mitochondrial non-coding regions …….. 13
1.3 Ribosomal transcription units in animals and parasites …….……. 14
1.3.1 Ribosomes and ribosomal transcription units (rTUs) …….. 14
1.3.2 Organization and genetic location of ribosomal transcription units …… 16
1.3.3 Characteristics of the ribosomal genes and intergenic regions ……... 16
1.3.4 Secondary structure of ribosomal genes and intergenic regions ……. 17
1. 4 Research on the mitogenomes and ribosomal transcription units ……..… 18
1.4.1 Research on the mitogenomes of trematodes (mitogenomics)…….. 18
1.4.2 Research on ribosomal transcription units of trematodes (ribosomal genomics) …….……..…….. 20
1.5 The importance of studying mitogenomes and ribosomal transcription units and particular aims for taxonomic resolution of echinostomatid species in this study ……..…… 22
1.5.1 The importance of studying mitogenomes and ribosomal transcription units of echinostomes…….……..…… 22
1.5.2 The particular aims for taxonomic resolution of echinostomatid species in this study…….……..…… 23
For the Echinostoma revolutum species 23
For the Echinostoma/Artyfechinostomum malayanum species 24

For the Echinostoma miyagawai and Hypoderaeum conoideum species 26
For the taxonomy of the Echinostomatidae family and the Echinostomata suborder……. 27

Chapter 2: Materials and Methods…….. 29
2.1 Parasite samples and species identification…….. 29
2.2 Total genomic DNA extraction…… 31
2.3 Primer design and PCR strategies…….. 31
2.4 Approach for next-generation sequencing (NGS) …….. 32
2.4.1 Targeted enrichment of the mtDNAs by long-range polymerase chain reaction……. 32
2.4.2 Library preparation, long-read sequencing and *de novo* assembly of the mtDNA sequences…… 33
2.5 Mitogenomic annotation……. 33
2.6 Identification of structural features and promotor sequences in the non-coding region…….. 34
2.7 Comparative mitogenomic analysis of the Echinostomatidae family……. 34
2.7.1 Gene identity comparisons…….. 34
2.7.2 Base composition and skewness……. 35
2.7.3 Codon usage and bias in protein-coding genes…… 35
2.7.4 Pairwise genetic distances among echinostomes……. 35
2.8 Annotation and sequence analysis of ribosomal transcription units…….. 35
2.8.1 Annotation of the rTU sequences……. 35
2.8.2 Modeling the *de novo* structure of the 28S rRNA gene……... 36
2.9 Mitophylogenetic analysis and tree construction…….. 36
2.10 Ribosomal phylogenetic analyses and tree reconstruction…….. 37

Chapter 3: Results ……..……...……... 39
3.1 Mitogenomic genes and gene order of *Echinostoma revolutum*, *Echinostoma/ Artyfechinostomum malayanum, Echinostoma miyagawai* and *Hypoderaeum conoideum*…… 39
3.1.1 *Echinostoma revolutum*…… 39
3.1.2 *Echinostoma/Artyfechinostomum malayanum*……. 40
3.1.3 *Echinostoma miyagawai and Hypoderaeum conoideum*……. 43
3.2 Comparative mitogenomic analyses of the echinostomes in the Echinostomatidae family…….. 47
3.2.1 Gene identity comparisons among echinostomes…… 47
3.2.2 Base composition and skew values in the Echinostomatidae species….. 47
3.2.3 Codon usage in the protein-coding genes of the Echinostomatidae species…….. 50
3.2.4 Pairwise genetic distance among the Echinostomatidae species…….. 51
3.3 Polymorphismic features in non-coding regions (NCR) of echinostomes….. 52
3.3.1 Structural polymorphism in the NCRs of the Echinostomatidae species. 52
3.3.2 The NCRs featured by the promotor and regulatory sequences in the

NCR's long and short repeat units in echinostomes……..…….... 54
3.3.3 The NCRs featured by the palindromic sequences in the *Echinostoma* species…….. 55
3.3.4 The NCRs featured by the evolution of mitochondrial control region in the Echinostomatidae species…….. 56
3.4 The ribosomal transcription units of five echinostomes and their implications 57
3.4.1 Ribosomal transcription unit features of the echinostomatid and echinochasmid species……. 57
3.4.2 *De novo* structure of the 28S rRNA gene of echinostomes……. 59
3.5 Mitophylogenetic relationships within Echinostomatidae and Echinostomata…… 60
3.6 Phylogenetic analyses utilizing the ribosomal transcription unit sequences…… 62
3.6.1 Interfamilial phylogenetic relationships within Echinostomata and among other suborders……. 62
3.6.2 Phylogenetic relationships within families Echinostomatidae and Echinochasmidae…… 65

Chapter 4: Discussion, Conclusions, Contributions, and Future Prospects…. 68
4.1 The achievements of the mitogenomic investigations…… 68
4.1.1 Comparative mitogenomic datasets for the suborder Echinostomata….. 68
4.1.2 The comparative mitogenomic characteristics of echinostomes…….. 71
4.1.3 The polymorphism in the mtDNA's non-coding regions of the echinostomes…… 73
4.1.4 The mitogenomic phylogeny implications for echinostomes…….. 74
4.2 The achievements of the ribosomal transcription unit investigations……. 77
4.2.1 Comparative rTU's datasets for the suborder Echinostomata…….. 77
4.2.2 The ribosomal phylogeny implications of the suborder Echinostomata 78
4.2.2.1 Phylogenetic relationships above the level of suborder Echinostomata…… 78
4.2.2.2 Phylogenetic relationships among and within echinostomatan taxa…….. 78
4.3 An overall assessment of the thesis's contributions…….. 80
4.4 Conclusions…… 82
4.5 Future prospects…….. 83

References……. 85

APPENDICES…….. 104

ABBREVIATIONS

Abbreviation	Full phrase
aa	amino acid
ATP	Adenosine triphosphate
bp	base pair
CR	Control region
DHU	dihydrouridine
DNA	Deoxyribonucleic acid
Eca.	*Echinostoma*
Ecs.	*Echinochasmus*
ETS	External Transcribed Spacer
GD	Genetic distance
IGS	Non-transcribed intergenic spacer
ITIS	Integrated Taxonomic Information System
ITS	Internal Transcribed Spacer
kb (= kbp)	kilo base pair
LNR	Long non-coding region
LPCR	Long- Polymerase chain reaction
LRU	Long repeat unit
LRUPd	Long repeat unit palindrome
MFE	Minimum free energy
min	minute
ML	maximum likelihood method
MRG	Mitoribosomal gene
mRNA	Messenger RNA
mtDNA	Mitochondrial genome or mitogenome
mtDNA*	Coding mitogenome (5’ *cox*1 to 3’ *nad*5)
mt-LSU	Large mitoribosomal unit
mt-SSU	Small mitoribosomal unit
NCBI	National Center for Biotechnology Information
NCR	Non-coding region
NGS	Next-generation sequencing
NOR	Nucleolar organizer region
PCG	Protein coding gene

PCR	Polymerase chain reaction
rDNA	Ribosomal DNA
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
rTU	Ribosomal transcription unit
SMRT	single-molecule real-time
SRU	Short repeat unit
SRUPd	Short repeat unit palindrome
tRNA/*trn*	Transfer RNA

LIST OF TABLES

		Page
Table 1.1	Species of Echinostomatidae infecting humans with their collar spine numbers and geographical distribution ……	4
Table 1.2	Nomenclature for mitochondrial genes of animals used in the thesis ……..	10
Table 2.1	List of *Echinostoma* and *Echinochasmus* samples that were used to obtain the entire mitogenome and ribosomal transcription units…….....	29
Table 3.1	Locations of genes and other features in the complete mitochondrial genome of *Echinostoma revolutum*…….....	39
Table 3.2	Locations of genes and other features in the mitochondrial genome of *Echinostoma malayanum* (17,175 bp) and *Artyfechinostomum sufrartyfex* (14,567 bp) ……...	42
Table 3.3	Locations of genes and other features in the mitochondrial genomes of *Echinostoma miyagawai* (Emiya-RED11-TH, 19,417 bp, GenBank: OP326312); and *Hypoderaeum conoideum* (Hcono-RED42-TH, 18,011 bp, GenBank: PP110501) ……	43
Table 3.4	Nucleotide comparison for divergence rate (%) of individual and concatenated protein-coding (PCGs) and mitoribosomal genes (MRGs) between *Echinostoma miyagawai* (strain RED11, Thailand) and members of the family Echinostomatidae (Platyhelminthes: Echinostomata) ……	48
Table 3.5	Base composition and skew/skewness value for AT and GC of the protein-coding genes (PCGs), mito-ribosomal genes (MRGs), and the coding region (abbreviated as mtDNA*) of the mitogenomes of *Echinostoma miyagawai* and other echinostomatid members of the family Echinostomatidae…….	49
Table 3.6	Codon usage for 12 protein-coding genes in the mitogenomes of four strains of four species (*Eca. revolutum*; *Eca./A. malayanum*; *Eca. miyagawai*, and *H. conoideum*) of the family Echinostomatidae…….	51
Table 3.7	Pairwise genetic distance (%) among 15 strains of 12 species in the family Echinostomatidae estimated based on the analysis of concatenated nucleotide sequences of protein-coding genes……..	52
Table 3.8	An updated summary of the mitogenomic datasets of all 15 strains of 12 species of the family Echinostomatidae including four echinostomes from this study regarding their length, the non-coding regions (NCR), repeats, PCGs, and references. ……	53
Table 3.9	List of palindromic sequences found in the long and short repeat unit in the mitochondrial control region of *Echinostoma miyagawai*……	56
Table 3.10	Positions of ribosomal RNA genes, internal transcribed spacers (ITS), external transcribed spacer (ETS), and non-transcribed intergenic spacers (IGS) in the ribosomal transcriptional unit of *Echinostoma* species (Echinostomatidae) and *Echinochasmus* (Echinochasmidae) species in this study.……	59
Table S2.1	Primers for amplification and sequencing fragments of mitochondrial	

genome of *Echinostoma revolutum*…… 104

Table S2.2 Primers for amplification and sequencing fragments of mitochondrial genome of *Echinostoma malayanum*…… 104

Table S2.3 List of trematode-universal and specific primers for long-range PCR used for enrichment of the mitogenomes of *Echinostoma miyagawai* (Emiya) and *Hypoderaeum conoideum* (Hcono) to obtain amplicons for next-generation sequencing…….. 105

Table S2.4 List of 57 trematode strains of 41 species providing information for the available mitochondrial genome in the suborder Echinostomata and other suborders used in this study for sequence comparative and phylogenetic analyses……. 105

Table S2.5 List and information on strains and species that provide complete sequences of the ribosomal 28S and 18S rRNA genes for phylogenetic analysis and tree reconstruction to assess the intra- and interfamilial relationships in the class Trematoda (Digenea: Platyhelminthes)…… 108

Table S2.6 List and information on strains and species that provide partial nuclear ribosomal 28S rDNA D1−D3 sequences for phylogenetic analysis and tree reconstruction for the assessment of the taxonomic relationships of the suborder Echinostomata (Trematoda: Platyhelminthes)……. 110

Table S3.1 Codon usage for 12 protein-coding genes in the mitochondrial genomes of 15 strains of 12 species of the family Echinostomatidae in this study…….. 116

LIST OF FIGURES

		Page
Figure 1.1	Taxonomic hierarchy of the family Echinostomatidae and its genera within family according to the Integrated Taxonomic Information System (ITIS)….	1
Figure 1.2	*Echinostoma revolutum* specimens recovered from school children in Pursat Province, Cambodia, which had 2 testes in the postequatorial region	3
Figure 1.3	Global distribution of *Echinostoma* and *Echinochasmus* species and members of the family Echinostomatidae based on the presence of their life cycles. ……	6
Figure 1.4	A multi-host (indirect) life cycle of echinostomid flukes. ……	6
Figure 1.5	Animal mitochondria and functions. ……	8
Figure 1.6	Mitochondrial genomes and their transfer RNAs in animals and parasites...	11
Figure 1.7	Ribosomal transcription units and their transcription in animals and parasites. …….	15
Figure 2.1	Layout and steps to implement research on two main subjectives	30
Figure 3.1	A schematic drawing of circular map of the mitogenome of four echinostomes…….	41
Figure 3.2	Drawings of predicted structure models of 22 transfer RNAs in the mitogenome of *Echinostoma* species……	45
Figure 3.3	Schematic of the position of predicted promoter regions within the tandem repeat units repeat units of *Echinostoma miyagawai* mitochondrial control region…….	55
Figure 3.4	Palindromic repeat regions identified in the long repeat unit and short repeat unit in non-coding region of *Echinostoma miyagawai*……..	55
Figure 3.5	Structural organization of the near-complete ribosomal transcription units for echinostomatids and echinochasmids…….	58
Figure 3.6	The *de novo* secondary structure model of the 28S rRNA gene (domains D1–D3) …….…….	60
Figure 3.7	A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatidae and among other superfamilies and suborders…….	61
Figure 3.8	A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatoidea (suborder: Echinostomata) and among other superfamilies and suborders (Opisthorchiata, Xiphidiata, Pronocephalata, and Diplostomata)…….	63
Figure 3.9	A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatoidea (suborder: Echinostomata) and between the other superfamilies and suborders (Opisthorchiata, Xiphidiata, Pronocephalata, and Diplostomata) …….	64
Figure 3.10	Phylogeny based on the analysis of the D1−D3 sequences (1.1–1.3 kb) of the 28S rRNA genes showing the detailed relationships of the families Echinostomatidae and Echinochasmidae and other families in the suborder Echinostomata) …….……	66

CHAPTER 1

Literature review

1.1. Echinostomes of the Echinostomatidae family

1.1.1. Classification of the echinostomid flukes of the Echinostomatidae family

The Echinostomatidae are parasitic intestinal trematodes, sometimes known as echinostomid intestinal flukes (or echinostomes), that infect a wide range of animals, including humans. Human echinostomiasis is a zoonotic foodborne trematodiasis that, despite its global prevalence, is primarily a public health issue in Southeast Asia [1, 2]. It comprises several echinostomid species that play an important epidemiological function and have a characteristic infection cycle in humans and animals. [2, 3]. Previously, scientists classified *Echinochasmus* Dietz, 1909, as trematodes belonging to the subfamily Echinochasminae Odhner, 1910, of the family Echinostomatidae (Trematoda: Platyhelminthes). However, contemporary sequencing and phylogenetic studies using ribosomal and mitochondrial genetic markers indicated upgrading this subfamily to family rank, constituting a new family Echinochasmidae and removing it from the Echinostomatidae family [4−6]. Both families are the two principal families in the suborder Echinostomata, with significant genera including essential species implicated in human infections worldwide [1, 2, 7], see: https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/.

The family Echinostomatidae Looss, 1899, includes seventeen species infecting humans from at least seven genera, including *Acanthoparyphium* Dietz, 1909; *Echinoparyphium* Dietz, 1909; *Echinostoma* Rudolphi, 1809; *Himasthla* Dietz, 1909; *Hypoderaeum* Dietz, 1909; *Isthmiophora* Lühe, 1909; and *Artyfechinostomum* Lane, 1915 [1, 3]. The suborder Echinostomata, which comprises these two key families, is a vast, diversified, and extensively dispersed group of parasitic flatworms in the Plagiorchiida order. The family Echinochasmidae Odhner, 1910 includes the genus *Echinochasmus* Dietz, 1909, which harbors at least six human pathogen species, whereas the family Echinostomatidae contains nearly tens of echinostosome human pathogens [1, 3, 5, 7–9]. In this thesis, we aimed to focus on the molecular investigations on some zoonotic echinostomid species in the family Echinostomatidae, with their taxonomic hierarchy shown in **Fig. 1.1**. The species studied are members of the Echinostomatidae family, specifically those of the genera *Echinostoma*, *Artyfechinostomum*, and *Hypoderaeum*.

Kingdom	Animalia – Animals
Subkingdom	Bilateria – triploblasts
Infrakingdom	Protostomia
Superphylum	Spiralia
Phylum	Platyhelminthes Minot, 1876 – flatworms, plathelminthes
Subphylum	Rhabditophora Ehlers, 1985
Infraphylum	Trepaxonemata
Superclass	Euneoophora
Class	Trematoda Rudolphi, 1808
Subclass	Bothrioneodermata
Infraclass	Neodermata
Superorder	Digenea Carus, 1863
Order	Plagiorchiida La Rue, 1057
Suborder	Echinostomata
Family	Echinostomatidae Looss, 1899
	Direct Children:
Genus	● Artyfechinostomum (Leiper, 1911) Mendheim. 1943
Genus	Aporchis Stossich, 1905
Genus	Baschkirovitrema Skrjabin, 1944
Genus	Drepanocephalus Dietz, 1909
Genus	Echinochasmus Dietz, 1909
Genus	Echinoparyphium Dietz, 1909
Genus	● Echinostoma Rudolphi, 1809
Genus	Euparyphium Dietz, 1909
Genus	Himasthla Dietz, 1909
Genus	● Hypoderaeum Dietz, 1909
Genus	Ignavia Freitas, 1948
Genus	Isthmiophora Luhe, 1909
Genus	Longicollia Bykovskaia-paviovskaia, 1954
Genus	Patagifer Dietz, 1909
Genus	Pelmatostomum Dietz, 1909
Genus	Petasiger Dietz, 1909
Genus	Prionosoma Dietz, 1909
Genus	Protechinostoma Beaver, 1943
Genus	Stephanoprora Odhner, 1902

Figure 1.1. Taxonomic hierarchy of the family Echinostomatidae and its genera within this family according to the Integrated Taxonomic Information System (ITIS). Source: ITIS at (https://www.itis.gov/). The solid circles indicate the genera that were directly investigated in this study.

The family Echinostomatidae Looss, 1899, was founded by the first species of the type genus *Echinostoma* Rudolphi, 1809. The echinostome species of this genus predominantly parasitize in birds, commonly found in mammals and humans as well, but are rarely in fish and reptiles. Morphologically, echinostomes are characterized by collar-spines around the oral sucker around the head. The varying number and unique arrangement of collar spines for each group of echinostome flukes is a significant key in taxonomic evaluation [1]. The Echinostomatidae currently includes 44 genera and 10 subfamilies, with seven genera (*Acanthoparyphium*, *Artyfechinostomum*, *Echinoparyphium*, *Echinostoma*, *Himasthla*, *Hypoderaeum*, and *Isthmiophora*) comprising 20 species that are regarded medically significant [1]. *Echinochasmus*, a genus previously listed in the subfamily Echinochaminae within the Echinostomatidae family, has been transferred to the newly elevated family Echinochasmidae [4, 5].

Morphologically, according to the number of rows present on the head of the echinostomes, and the way of the collar spine-arrangement, medically important echinotomes are classified into two groups [1]:

i) The first group is *Echinostoma*, *Isthmiophora*, *Echinoparyphium*, and *Hypoderaeum*, which have collar-spines arranged in two rows, that are interrupted ventrally but not dorsally.

ii) The second group includes the taxa *Artyfechinostomum*, *Acanthoparyphium*, and *Himasthla*, which have collar-spines organized usually in a single row, interrupted ventrally but not dorsally.

The *Echinochasmus* Dietz, 1909 group with collar-spines arranged in a single or alternative rows, interrupted both ventrally and dorsally, was listed as the third group in [1], but in the current classification of class Trematoda, it is now as one group in the family Echinochasmidae (**Table 1.1**).

According to the Taxonomic Browser available in the National Center for Biotechnology Information (NCBI Taxonomy: https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/), the classification system of the family Echinostomatidae and its relative genera are arranged in the Lineage, as follows:

Lineage (full): cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Protostomia; Spiralia; Lophotrochozoa; Platyhelminthes; Trematoda; Digenea; Plagiorchiida; Echinostomata; Echinostomatoidea; Echinostomatidae; *Echinostoma*/ *Artyfechinostomum*/ *Hypoderaeum*

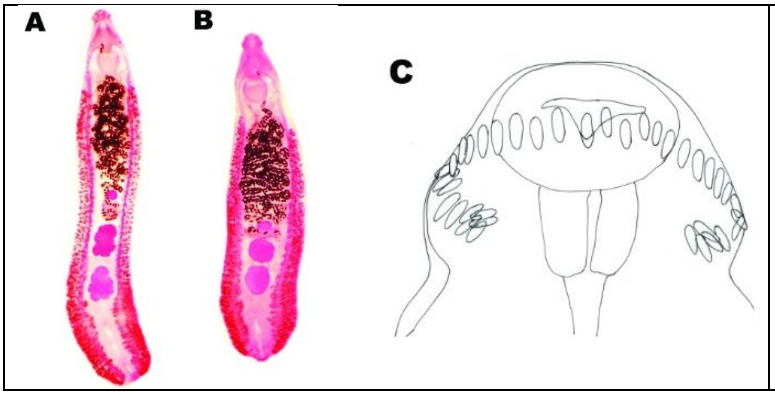


Figure 1.2. *Echinostoma revolutum* specimens recovered from school children in Pursat Province, Cambodia, which had 2 testes in the postequatorial region. A) An adult worm (8 mm long) showing lobulated testes; B) Another adult worm showing globular testes; C) Head collar of an adult specimen armed with 37 collar-spines arranged in a single row, including 5 end-group spines on each side. (Source: Sohn et al. [10]).

The family Echinostomatidae exhibits substantial taxonomic diversity and wide geographical distribution. The type genus, *Echinostoma*, contains many species and is distributed worldwide in all four continents [1] (**Table 1.1**). Morphological characters used to distinguish species include the presence of “collar-spines” and their numbers, structure, and arrangement around the oral sucker [11–13]. The genus *Echinostoma* (abbreviated as *Eca*. in this thesis) is divided into five groups [1, 2. 14, 15]. The most important group is the namely "*Echinostoma* revolutum" or shortly "revolutum" group, which was formed with the type *Echinostoma* species, and shares the most related morphological and molecular characteristics. The members of this "revolutum" group are characterized by 37 collar-spines and based on the type species, found on their cercariae [11, 12]. Nine *Echinostoma* species within the

*"*revolutum*"* group, including *Echinostoma caproni*, *Echinostoma echinatum*, *Echinostoma friedi*, *Echinostoma jurini*, *Echinostoma miyagawai*, *Echinostoma paraensei*, *Echinostoma parvocirrus*, *Echinostoma revolutum*, and *Echinostoma trivolvis*, are all of medical and zoonotic importance [1, 3, 7].

Other groups and genera have a variable number of "collar"-spines, such as 25–29 (*Eca. hortense* or *Isthmiophora hortense*), 31 (*Eca. anseries*), and 43–45 (*Eca. aegyptiacum*), while *Artyfechinostomum malayanum* has 43, *Hypoderaeum conoideum* has 41–45, and *Echinoparyphium recurvatum* has 43–50 collar-spines [1, 3, 16]. *Echinochasmus* genus (referred to as *Ecs.* in this thesis), which includes *Echinochasmus coaxatus*, *Ecs. japonicus*, *Ecs. beleocephalus*, and *Ecs. perfoliatus*, have 24 collar-spines, but *Ecs. mordax*, *Ecs. milvi*, and *Ecs. suifunensis* have 20 to 22 spines. These spines are arranged in a single row around the oral sucker [1, 17]. Members of Himasthlidae, *Acanthoparyphium tyosenense* has 23, while *Himasthla muehlensi* has 31 spines (**Table 1.1; Fig. 1.2**). The similarity of these echinostome species within the *Eca. revolutum* complex or in the Echinostomatidae family usually necessitated the use of additional identification methods, primarily enzymatic and molecular techniques for their discrimination [4, 6, 9, 12, 18, 19].

1.1.2 Lifecycle, epidemiology, and geographical distribution

Table 1.1 lists 24 human infecting echinostome species of Echinostomatidae, Echinochasmidae, and Himasthlidae with their collar numbers and geographical distribution. The zoonotic echinostomes are distributed mainly in Southeast Asia and the Far East [1, 3]. The majority of these parasite infections are known as zoonotic and can be found all over the world, although they are most common in Asian communities such as India, Indonesia, the Philippines, China, Malaysia, Singapore, Korea, Japan, Thailand, Myanmar, Laos, Cambodia, and Vietnam (**Table 1.1**). It is estimated that tens of millions of people have become infected, with hundreds of millions more at danger (**Fig. 1.3**). Humans become infected after eating raw or undercooked mollusks, fish, crustaceans, or amphibians [3, 7].

Table 1.1. Species of Echinostomatidae infecting humans with their collar spine numbers and geographical distribution (according to Chai [1]; Chai and Jung [3]; Toledo et al [7]

	Family/Species	No of collar-spines	Geographical distribution (country)
	Family Echinostomatidae		
1	*Artyfechinostomum oraoni* Bandyopadhyay, Manna and Nandy, 1989	41	India
2	*Artyfechinostomum malayanum*[a] (Leiper, 1911) Mendheim, 1943	43	Cambodia, China, India, Indonesia, Lao PDR, Malaysia, Philippines, Singapore, Thailand
3	*Artyfechinostomum sufrartyfex* Lane, 1915	43	India, Vietnam?

4	*Echinoparyphium recurvatum* (von Linstow, 1873) Lühe, 1909	45	Bangladesh, Bulgaria, Canada, China, Croatia, Czech Republic, Egypt, England, India, Indonesia, Japan, South Korea, Mexico, Philippines, New Zealand, Poland, Russia, Spain, Taiwan, Thailand, USA
5	*Echinostoma aegyptica* Khalil and Abaza, 1924	43-45	China, Egypt, Japan, Taiwan, Turkey, Lao PDR, Vietnam
6	*Echinostoma angustitestis* Wang, 1977	41	China
7	*Echinostoma cinetorchis* Ando and Ozaki, 1923	37	China, Japan, Korea, Taiwan, Vietnam?
8	*Echinostoma ilocanum* (Garrison, 1908) Odhner, 1911	51	Cambodia, China, India, Indonesia, Lao PDR, Malaysia, The Philippines, Thailand
9	*Echinostoma lindoense*[d] Sandground and Bonne, 1940	37	Indonesia, Lao PDR, Malaysia, Thailand
10	*Echinostoma macrorchis* Ando and Ozaki, 1923	45	Japan, South Korea, Lao PDR, Taiwan
11	*Echinostoma mekongi* Cho, Jung, Chang, Sohn, Sinuon and Chai, 2020		Cambodia, (probably also in Lao PDR, Thailand, Vietnam)
12	*Echinostoma paraensei* Lie and Basch, 1967	37	Australia, Brazil
13	*Echinostoma revolutum* (Fröhlich, 1802) Looss, 1899	37	**Asia** (Bangladesh, Cambodia, China, India, Indonesia, Japan, Korea, Lao PDR, Malaysia, Taiwan, Thailand, Vietnam); **Europe** (Austria, Belarus, Bulgaria, Czech Rep., England, Finland, France, Germany, Greece, Hungary, Iceland, The Netherlands, Poland, Russia, Slovak Rep., Yugoslavia), The Middle East (Iran); **Oceania** (New Zealand); **North America** (USA); **South America** (Brazil)
14	*Hypoderaeum conoideum* (Bloch, 1872) Dietz, 1909	49	Bangladesh, China, Indonesia, Japan, Mexico, North America, Russia, Spain, Taiwan, Thailand
15	*Isthmiophora hortensis*[f] (Asada, 1926) Kostadinova and Gibson, 2002	27	China, Japan, South Korea
16	*Isthmiophora melis* (Schrank, 1788) Lühe, 1909	27	China, Taiwan, Belarus, Bulgaria, Canada, Czech Republic, England, France, Germany, Hungary, Lithuania, Poland, Romania, Russia, Ukraine, USA
	Family Echinochasmidae		
17	*Echinochasmus caninus*[b] (Verma, 1935) n. comb.	24	Thailand
18	*Echinochasmus fujianensis* Cheng et al., 1992	24	China
19	*Echinochasmus japonicus* Tanabe, 1926	24	China, Japan[c], South Korea, Kuwait, Lao PDR, Russia, Taiwan, Thailand, Vietnam
20	*Echinochasmus jiufoensis* Yu and Mott, 1994	24	China
21	*Echinochasmus liliputanus* (Looss, 1896) Odhner, 1910	24	China
22	*Echinochasmus perfoliatus* (Ratz, 1908) Gedoelst, 1911	24	Bulgaria, China, Croatia, Denmark, Egypt, England, Hungary, India, Italy, Japan, South Korea, Poland, Romania, Russia, Serbia, Taiwan, Thailand, Ukraine, Vietnam
	Family Himasthlidae		
23	*Acanthoparyphium tyosenense* Yamaguti, 1939	23	South Korea
24	*Himasthla muehlensi*[e] Vogel, 1933	31	United States[e] , North America

[a]Syn. *Echinostoma malayanum*; [b]Syn. *Episthmium caninum*; [c]Experimental infection; [d]Syn. *Echinostoma echinatum*; [e]Imported infection; [f]Syn. *Echinostoma hortense*

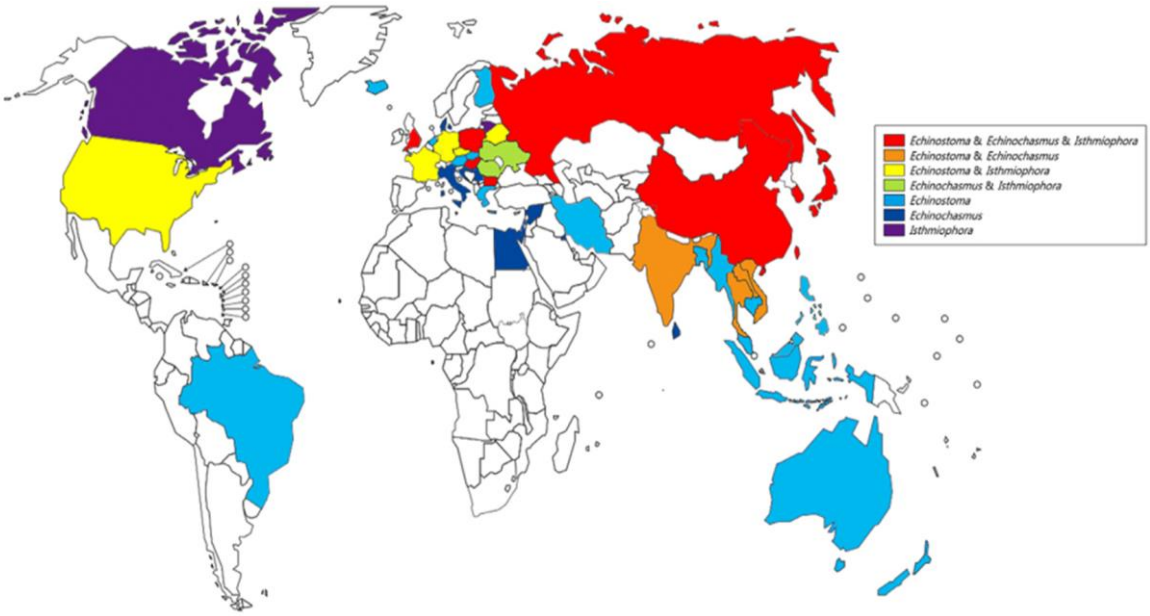


Figure 1.3. Global distribution of *Echinostoma* and *Echinochasmus* species and members of the family Echinostomatidae (*Eca. revolutum*, *Eca. cinetorchis*, *Eca. lindoense*, *Eca. paraensei*, *Eca. ilocanum*, *Eca. macrorchis*, *Eca. aegyptica*, and *E. angustitestis*), *Echinochasmus* spp. (*Ecs. japonicus*, *Ecs. perforliatus*, *Ecs. liliputanus*, and *Ecs. caninus*), and *Isthmiophora* spp. (*I. hortensis* and *I. melis*) based on the presence of their life cycles (Source: Chai and Jung [3]).

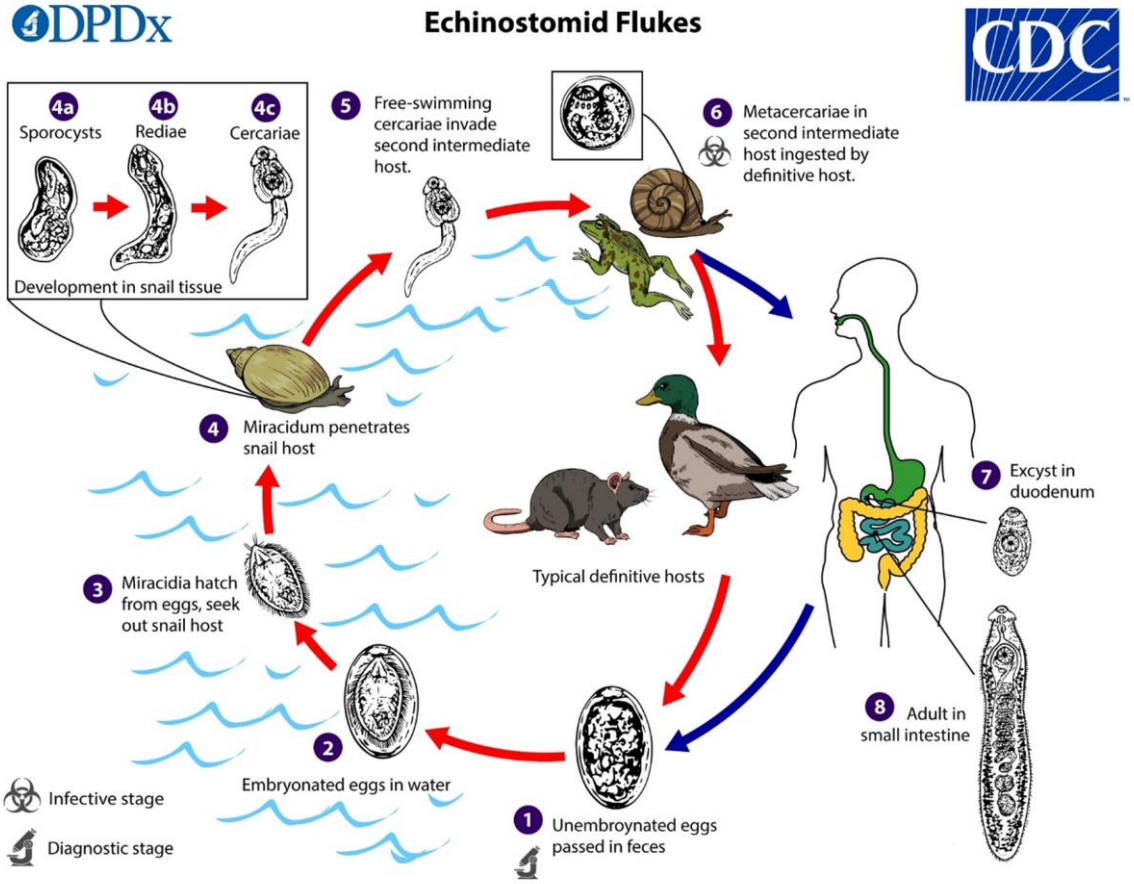


Figure 1.4. A multi-host (indirect) life cycle of echinostomid flukes. Unembryonated eggs are passed in feces of infected definitive hosts (**1**) and develop in water (**2**). Miracidia usually take about three weeks to mature before hatching (**3**) after which they swim freely and penetrate the first intermediate host, a snail (**4**) The intramolluscan stages include a sporocyst stage (**4a**) one or two generations of rediae (**4b**) and cercariae (**4c**) which are released from the snail. The cercariae may encyst as metacercariae within the same first intermediate host or leave the host and penetrate a new second intermediate host (**5**). The definitive host becomes infected after eating metacercariae

in infected second intermediate hosts (**6**). Metacercariae excyst in the duodenum (**7**) and adults reside in the small intestine (for some species, occasionally in the bile ducts or large intestine) (**8**). (Source: https://www.cdc.gov/dpdx/echinostomiasis/index.html).

The lifecycle of echinostomes is indirect and multiple staged (**Fig. 1.4**). The echinostome eggs are immature when laid, but they mature after leaving the host, and hatch in about three weeks in the environment. Miracidia enter the snail host, where they develop into mother rediae and in turn, daughter rediae, and cercariae. The cercariae have well-developed tails and usually bear collar-spines around the oral sucker similar to that of the adults. The encysted metacercariae are round or oval, and show two branches of the excretory bladder filled with coarse granules and a head collar with collar-spines varying species by species. Humans or animals are infected through ingestion of metacercariae encysted in the second intermediate host. Eating raw snails, clams, fishes, or vegetation harboring metacercariae is the main practical mode of infection in humans [1].

The infected definitive hosts (humans) released unembryonated eggs, they passed in feces and develop in water. Echinostomes have several developmental stages, including the eggs, miracidia, sporocysts, rediae, daughter rediae, cercariae, metacercariae, and adults during their life cycle (**Fig. 1.4**). In the water, the eggs develop into miracidia, which normally take approximately 3 weeks to hatch, swimming freely before penetrating into the aquatic snail, the first intermediate host. A first sporocyst and second redial generations and finally the cercariae are then developed in the snail tissue which are liberated from the snail host and begin to swim. Free-swimming cercariae invade second intermediate host (frogs, snails, clams, fishes, amphibian, and reptiles), and the definitive host (fishes, reptiles, birds, and mammals, including humans). The definitive host is mainly infected by consuming the second intermediate hosts harboring the metacercariae. When infected in the definitive host, including humans, the main habitat of the flukes is the small intestines.

1.2 Mitochondria and mitochondrial genomes in animals and parasites

1.2.1 General features of mitochondria

Mitochondria (single name, mitochondrion) are membrane-bound organelles present in the cytoplasm of all eukaryotic cells, including multicellular parasites, and also in all trematodes, which play a central role in provision of cellular energy and contain their own genome with a modified genetic code [20, 21]. Within a cell, each mitochondrion is enclosed by an outer membrane and also has an inner membrane, which has several folds in a special structure called cristae and the area within the inner membrane is called matrix (**Fig. 1.5**). The matrix also contains a host of enzymes, as well as ribosomes for protein synthesis. Mitochondria are thought to be descended from bacteria that formed an endosymbiotic relationship with the earliest

eukaryotic cells [22]. Over the millions of years that have elapsed since then, most of their genes, even those vital for the functions of the mitochondrion itself, have been transferred to the nuclear genome. Products of these translocated genes (the regulatory proteins) now have to be imported into the organelle by specific transport systems [23].

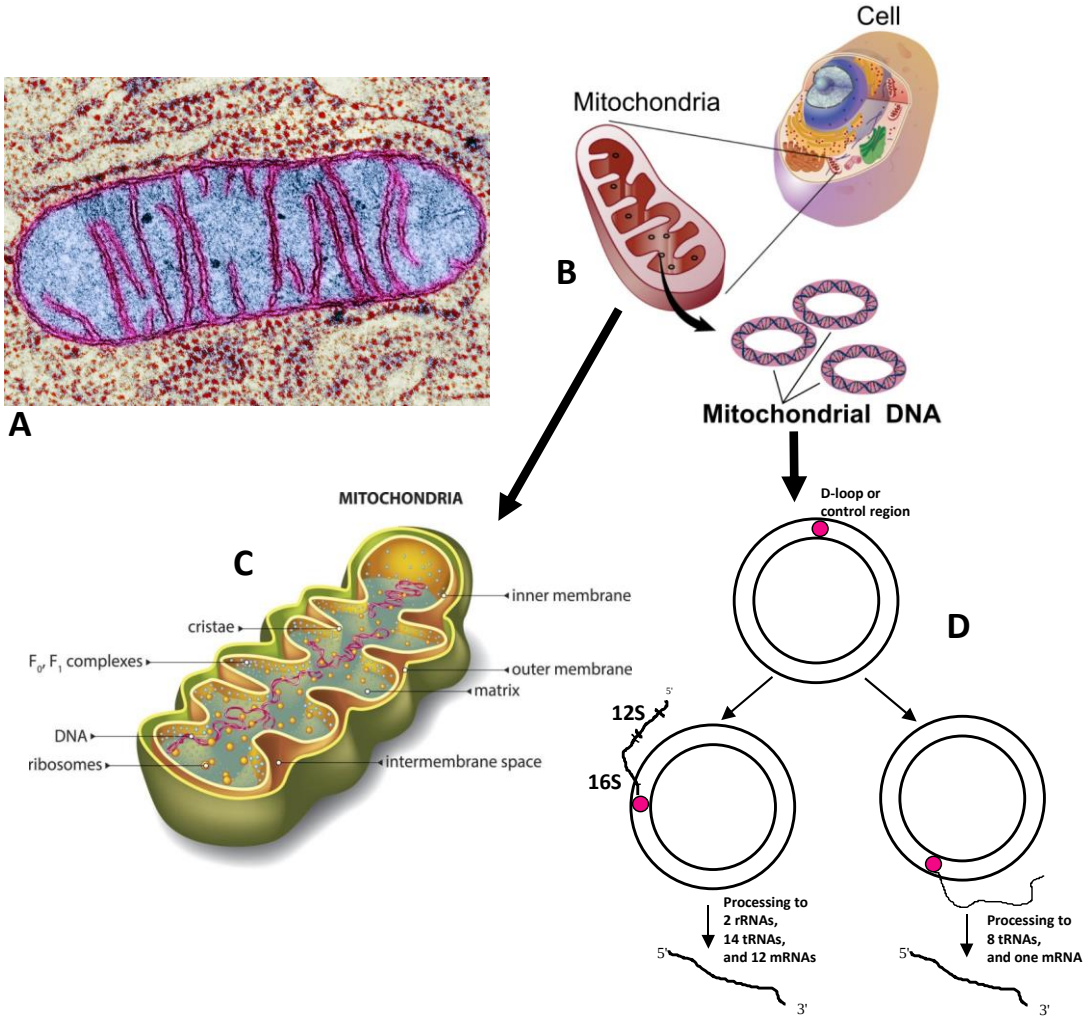


Figure 1.5. Animal mitochondria and functions. A. Electron micrograph; B. Location in a cell; C. Organelle structure; D. Transcription of human mitochondrial DNA and processing of primary transcripts. (Source: https://www.sciencephoto.com/media/215002/view/mitochondrion-tem; https://www.news-medical.net/life-sciences/Mitochondria-Overview.aspx; https://en.wikipedia.org/wiki/Mitochondrial_DNA) (Le [26]).

In eukaryotic cells, there are two genomes functioning together in close interaction to maintain and regulate cellular events, and they are the nuclear and the mitochondrial (mt) genomes or mitogenomes. Within a mitochodrion, with some exceptions of linear forms for several eukaryotic species, the mitogenomes are circular DNA molecules (mtDNA), which are localized in the matrix. In animals and human, there are two strands comprising the mtDNA molecule and can be separated into heavy (H) and light (L) strands by ultracentrifugation. The mtDNA transcription in vertebrates are started in two ways with bipartite mitochondrial H- and L-strand promoters located in a core region (control region or D-loop). A very special feature of vertebrates' mitochondrial transcription is that genes tend to co-transcribe giving long multi-

cistronic primary transcripts, which are subsequently cleaved into individual products at the location of tRNAs [24] (**Fig. 1.5**). In platyhelminths (including trematode parasites) the mtDNA transcription occurs in one way, on the positive strand [25].

In contrast to the nuclear genome, which contains only two copies per cell, the mt genome is present in multiple copies per cell (from hundreds to thousands), depending on cell type and its function. The nuclear genome, or in other words, all the chromosomes in a cell, contains most of the genetic material of the cell, while the mitogenome comprises their own genes to program the expression of 12–13 proteins, which are vital respiratory-chain enzyme complexes [20, 23]. Cellular chemical energy produced by the mitochondria is stored in a form of a small molecule called adenosine triphosphate (ATP) and on the surface of the inner membrane. In animals, high ATP requirement cells have as high as ~7,000–10,000 mitogenome copies per cell, while low energy requirement cells have as low as ~100 copies per cell. Mitochondria are important parts of the cell, exist in haploid form, are not recombinant, are maternally inherited, have their own genome and ribosomes, function independently, and interact with other compartments of the cells. Due to its small size and maximum containment of protein-producing genes necessary for cell life, it can be considered an ideal object to investigate genetic/genome changes for research. Mitochondrial genes have evolved 10–15 times faster than the nuclear genome, which is very convenient for research on evolution and population genetics [20]. Mitochondrial genes within the same variants of a species, within the same species, even within biologically closely related species, have a very high closeness, so any small change is a valuable sign in genetic assessment and taxonomic classification [27].

1.2.2 Mitochondrial genes and genomes

1.2.2.1 Gene order and nomenclature of mitogenes

In the mitochondria of almost all animals, there is a compact, circular genome, of 13.5–25 kb in length, which typically contains 36–37 genes and some tracts necessary for replication and transcription. In vertebrates, of the genes present, there are 13 PCGs, including *atp*8 but in platyhelminths/trematodes, including echinostomes, the mtDNA contains only 12 PCGs with the absence of the *atp*8 gene[25, 28, 29]. Each trematode mtDNA contains 12 protein-coding genes, PCGs (*atp*6, *cob*, *cox*1–3, *nad*1–6 and *nad*4L), 2 mitochondrial ribosomal RNA genes, or mitoribosomal genes, MRGs (*rrn*L/(16S) and *rrn*S/(12S)), 22 transfer RNA genes (tRNA or *trn*) and a non-coding region (NCR) rich in multiple repeat units of variable length [20, 22, 25, 28, 30–32] (**Fig. 1.6A**). The linear map of a trematode mtDNA (including the mtDNA map of *Echinostoma* spp.) is: 5'-*cox*3-H-*cob*-*nad*4L-*nad*4-QFM-*atp*6-*nad*2-VAD-*nad*1-NPIK-*nad*3-S_1W-*cox*1-T-*rrn*L-C-*rrn*S-*cox*2-*nad*6-Y$L_1S_2L_2$R-*nad*5-GE/(or EG)-NCR (RUs/ or none)]-3',

except the downstream region after *nad*5 and the NCR, where the $tRNA^{Gly}$ (G) and $tRNA^{Glu}$ (E) positions are often interchanged in some species (**Fig. 1.6B**).

Table 1.2. Nomenclature for mitochondrial genes of animals used in the thesis (13 protein-coding, 2 mitoribosomal RNA, and 22 transfer RNA genes) (adapted from Boore [20] and Le et al.[25]).

Genes	Gene abbreviations for common use	Standardized abbreviations currently used
Cytochrome oxidase subunit I, II, III	COI, COII, COIII	*cox*1, *cox*2, *cox*3
Cytochrome b apoenzyme	Cytb or CytB	*cob* or *cyt*B
Nicotinamide dehydrogenase (NADH) subunits 1–6, and 4L	ND1-6, and ND4L	*nad*1-6, and *nad*4L
ATP synthase F_o subunit 6 and 8	A6, A8 or ATP6, ATP8	*atp*6, *atp*8*
RNA mitoribosome large subunit	LrRNA or lrRNA or 16S	*rrn*L
RNA mitoribosome small subunit	SrRNA or srRNA or 12S	*rrn*S
Transfer RNA (tRNAs) specifying for each amino acid (overall, 18 tRNAs)	Each one-letter or three-letter correspond for each amino acid that the tRNA transfers	For example: *trn*V ($tRNA^{Val}$) for Valine; *trn*H ($tRNA^{His}$) for Histidine, or one letter, V, H…
Transfer RNA specifically specifying for special amino acid leucine (2 tRNAs)	Discriminated by codons that the tRNA recognizes on Leucine 1 (CUN) and Leucine 2 (UUR)	*trn*L1; $tRNA^{Leu1(CUN)}$; *trn*L2; $tRNA^{Leu2(UUR)}$ or one letter, L_1, L_2…
Transfer RNA specifically specifying for special amino acid serine (2 tRNAs)	Discriminated by codons that the tRNA recognizes on Serine 1 (AGN) and Serine 2 (UCN)	*trn*S1; $tRNA^{Ser1(AGN)}$ *trn*S2; $tRNA^{Ser2(UCN)}$ or one letter, S_1, S_2…

Note: (*): *atp*8* is absent in the mtDNA of flatworms (platyhelminths) and nematodes (Le et al.[25]).

Genes in the mitogenomes are compactly arranged, abutting each other or separated by only a short inter-genic region, and some genes may even overlap slightly. There is no standard nomenclature in the literature for abbreviation of names of mitochondrial genes, for example, COX1, COI, *cox*1… (for cytochrome oxidase subunit I) or ATP6, A6, ATPase 6, *atp*6… (for the adenosine triphosphatase subunit 6) etc... The tRNAs normally get their names for the amino acid they are assigned to transfer, for example, $tRNA^{His}$ or *trn*H (transferring Histidine), or $tRNA^{Gly}$ or *trn*G (transferring Glycine) etc… To keep consistency throughout the thesis and the published papers, the convention for abbreviating mitochondrial genes recommended by Boore [20] available at http://biology.lsa.umich.edu/~jboore/ with slight modifications by Le et al. [25] are used, as shown in **Table 1.2**. The currently used mtDNA gene abbreviations were standardized and used in many publications and databases [5, 9, 25, 28, 31–40].

1.2.2.2 Transfer and mitoribosomal RNAs

The mitochondrial genomes of all phyla have their own 22 transfer RNAs (tRNAs). Typically, 22 of these genes are dispersed throughout the genome and are sufficient to decode the 12 or 13 protein-coding genes. Most vertebrate and invertebrate mt tRNA genes can fold into the four-armed secondary structures similar to the "*clover-leaf*" structures. This "clover leaf"-shaped structure of a tRNA is specified by four arms: i) the **acceptor** arm/AA-arm; ii) the **dihydrouridine** arm/DHU-arm (D-arm); iii) the **anticodon** arm/AC-arm containing the

anticodon site; iv) the T-arm or **TpsiC** (TΨC) arm; and a **variable** loop [41] (**Fig. 1.6C**). However, some mt tRNAs exhibit distinct structures with only three arms present, and the dihydrouridine (DHU) or the TpsiC (TΨC) arm may be missing. Instead, just a loop is formed on the missing arm. Flatworms, nematodes, insects, certain echinoderms, and some vertebrates all have mtDNA with structural alterations in tRNA arms. This variety in tRNA arms is notably prominent in the mitogenomes of platyhelminths and nematodes [25, 28, 29]. The mt tRNAs, which have a TpsiC arm missing, are called *TΨC-replacement*, which is common in tRNAs of nematode mtDNAs or dihydrouridine arm (DHU) missing are called *DHU-replacement*, which is common in tRNAs of platyhelminths' mtDNAs (**Fig. 1.6D**). In mitogenomes of species in the phylum Platyhelminthes, with some exceptions in particular families, the gene order of trematodes is highly standardized, and the one way (positive) direction of transcription, the gene order was remarkably conserved. The tRNA clusters of Q-F-M, V-A-D, N-P-I-K, and Y-L_1-S_2-L_2-R were seen conserved in all mitogenomes of trematodes, including members of the Echinostomatidae (see Le et al. [25]) (**Fig. 1.6B**).

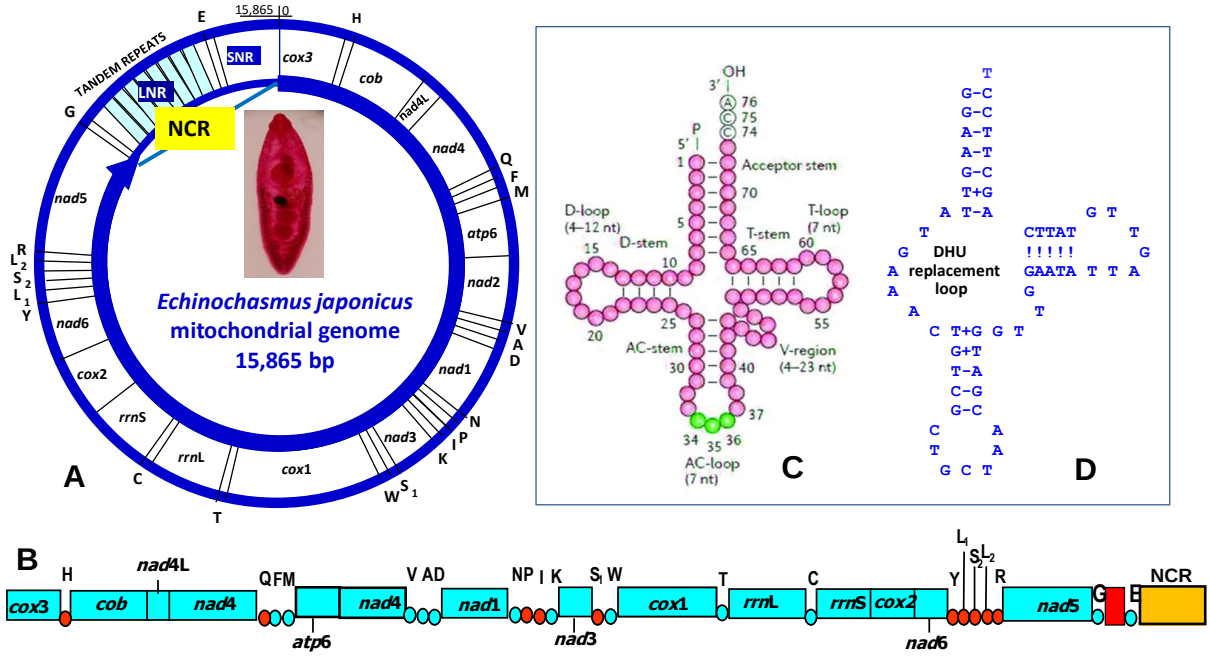


Figure 1.6. Mitochondrial genomes and their transfer RNAs in animals and parasites. A. A schematic drawing of a circular map of a mt genome (*Echinochasmus japonicus*/ family Echinochasmidae); B. A linear map of a trematode mt genome (opened at 5' terminus of the *cox*3 gene) showing the gene order, the location of genes and conserved tracts of tRNAs (QFM; VAD; NPIK; $YL_1S_2L_2R$); C. A schematic drawing of a tRNA with a normal secondary structure ("*clover leaf*" form); and D. A schematic drawing of a DHU (dihydrouridine arm)-missing form (Source: Le et al. (2016) [5]; https://www.wikiwand.com/en/Cloverleaf_model_of_tRNA).

The mitochondria have their own ribosomes (are called mitoribosomes), and these are smaller in size than those cytosolic ribosomes of the cell and function as places for polypeptide synthesis to generate enzymes responsible for the oxidative phosphorylation (OXPHOS) [23]. Two mitochondrial ribosomal RNA genes (MRGs) have been identified in all metazoan

mtDNA molecules: *rrn*L also known as large subunit, or 16S rRNA and *rrnS* (small subunit or 12S rRNA), which can be folded into specific secondary structures. These constitute the RNA components of mitoribosomes. Like the cellular ribosomes, the mitoribosomes consist of two subunits: a large subunit (mt-LSU, mitoribosomal large subunit) and a small subunit (mt-SSU, small subunit) [42]. Each subunit is a tight combination of ribosomal RNAs (12S rRNA and 16S rRNA, collectively known as mt-RNA) and ribosomal proteins (mitoribosomal proteins, mtRP or MRP). From the genome, the mRNA can be transcribed and translated to attach to small mt-SSU units and codons that interact with those encoded in tRNA, that the tRNA transfers the corresponding amino acid to synthesize the mitochondrial polypeptide enzyme [43]. The mt-LSU complex contains a peptidyl transferase center that catalyzes the formation of peptide bonds between amino acids delivered by tRNAs forming polypeptide chains [44].

In almost all species, ribosomal genes are positioned on the same strand being separated by a single tRNA, or a varying number of protein-encoding genes. In trematodes (and echinostomes) mitoribosomal RNAs (16S and 12S) are separated by tRNA for Cystein ($tRNA^{Cys}$) (**Fig. 1.6A, B**).

1.2.2.3 Mitochondrial protein-encoding genes and genetic codes

The protein-coding genes (PCGs) of mt genomes, which are 13 in animals and 12 in trematodes, cestodes and nematodes, are comprised of several complexes as follows: seven PCGs encode for the nicotinamide dehydrogenase (*nad*) complex (*nad*1–6 and *nad*4L subunits); three PCGs for the cytochrome c oxidase (*cox*) complex (*cox*1–3); one for cytochrome b (*cob*), and two for two subunits of adenosine triphosphatase (*atp*6 and *atp*8). In nematodes and all flatworms examined to date, the gene for *atp*8 is missing (see [25, 29, 30]). The mitochondrial proteins are synthesized by the mitochondria's own translation mechanism and the genetic code in mtDNA in parasites is different from the "*universal genetic code*" and between groups of each phylum, and this difference is related to initiation, termination and some specific codons for some special amino acids [45, 46].

Initiation codons of protein-encoding genes for translation are in most case ATG which codes for methionine consistent with the universal start codon. However, the other codons, ATT, ATA, and GTG are commonly used in invertebrate mt genomes and in parasitic helminths, some genes appear to use TTG or GTT as initiation codons [30, 45]. Termination codons most commonly used are TAG or TAA, but in some cases, a single T or TA is used, and the post-polyadenylation will add an additional A to complete the codon. In platyhelminths, TGA is used for specifying amino acid tryptophan and AGG and AGA for serine in invertebrates and flatworms [30, 45].

Protein-coding genes of the variants of the same species or species within the same family have a relatively high level of conserved nucleotide and amino acid sequences. The high identity has facilitated the accurate identification of genes in species classification [6]. In particular, the *cox* complex (e.g., *cox*1) and most *nad* genes (e.g., *nad*1 and *nad*3) are widely used in examination of "sister" species and closely related species. However, some other genes, such as *atp*6, *atp*8 or *nad*4L, *nad*3 and *nad*6, have less identity even in the related species [45–47]. Identification of such genes cannot be based solely on comparing similarity with known sequences in the GenBank database, but must generally rely on their biochemical properties (e.g., such as hydrophilic and hydrophobic properties and some other properties).

1.2.2.4 Mitochondrial non-coding regions

There is a region where no transcribed genes are located, which is termed a non-coding region (NCR) and known to be variable in size among species, even within the variants of a species. This is the longest intergenic spacer and contains multiple repeat units (usually tens to hundreds of nucleotides for each), and often to be arranged in tandem arrays [25, 30] (**Fig. 1.6A**). This NCR is usually a single, large region, rich in long (LRU) or short (SRU) or both types of repeat units. In vertebrates, the non-coding sequence is variable in size among taxa and is known to harbour the promoters for polycistronic transcription initiation and accelerate evolution of coding and regulatory sequences [48, 49]. Because of the presence of regulatory motifs, the NCR is sometimes called the 'control region' (CR). In humans, this region is also known as the "D-loop" (*displacement loop*), based on the unidirectional replication of the heavy chain O_H (heavy) and the light chain O_L (light) curling around each other [50]. The non-coding sequence usually forms a complicated secondary structure produced by tandem or inverted repeats that are variable in length and number and present a polymorphic characteristic for mtDNA of trematodes [49, 51].

In trematodes (class Trematoda), the NCR is divided by one or several tRNA genes generating two subregions: a short non-coding region (SNR) and a long non-coding region (LNR), but in cestode tapeworms (class Cestoda), these two subregions are separated by the *nad*5 gene and numerous tRNAs. In trematodes, the NCR is frequently a region between $tRNA^{Glu}$ (E) and *cox*3, with $tRNA^{Glu}$ is located downstream adjacent to $tRNA^{Gly}$ (G); however, in some species (for example, the *Paragonimus* genus), the NCR is between $tRNA^{Gly}$ and $tRNA^{Glu}$, and this tRNA is moved adjacent upstream of *cox*3 [25, 35, 52, 53]. Recent publications showed that the mtDNAs of most trematodes has NCRs possessing variable numbers of long (hundreds of nucleotides in length) and short (100–200 nucleotides), for example, *Eca. revolutum* (family Echinostomatidae) has an NCR containing up to 11 repeat units, including 7 long repeat (LRUs, long repeat units), each LRU is 317 bp. and 4 other short

repeat units (SRUs), each is 207 bp [31] or *Eca. miyagawai*'s NCR of 5,935 bp containing both long (15.3 LRUs of 319 bp/each) and short identical tandem repeat units (4.8 SRUs of 213 bp/each) [49].

The presence of LRUs and SRUs in the mitogenome has been reported in a range of mtDNAs, including those of species previously sequenced by Sanger-sequencing but now revealed by the next-generation sequencing (NGS) such as *Clonorchis sinensis* (Opisthorchiidae), *Paragonimus westermani* and *Paragonimus skrjabini miyazakii* (Paragonimidae), *Eca. revolutum* and *Eca. myiagawai* (Echinostomatidae), *Schistosoma bovis* (Schistosomatidae), and the cestode *Echinococcus granulosus* G1, and the lengthy NCR with multiple repeats represents a characterstic and polymorphism in trematodes [9, 31, 32, 49, 51–54].

The NCRs often have very high Adenine and Thymine (A+T) components, so it is also called the A+T rich region that were found in mitogenomes of insects (usually over 70%) [55, 56]. In trematodes, and echinostomes, NCR has a more balanced usage of A+T and G+C, about 60–65% A+T and 35–40% G+C [25, 31]. Other intergenic regions between PCGs or tRNAs are usually short, from several to tens of nucleotides. Their function is unclear, maybe, simply serving as link sequences between genes [25, 29, 30, 47].

1.3 Ribosomal transcription units in animals and parasites

1.3.1 Ribosomes and ribosomal transcription units (rTUs)

In translation, the sequence of codons on mRNA directs the synthesis of a polypeptide chain. This process takes place on the ribosomes, and the movement of tRNA and mRNA through the ribosomes is a complicated process, in which many enzymes and structures involved. The prokaryotic and eukaryotic cells have ribosomes. In eukaryotic cells, both the host cells and their mitochondria have their own ribosomes: in host cells, ribosomes are localized in cytoplasm, while in mitochondria they are located in the matrix and near the inner membrane[43]. Eukaryotic ribosomes are assembled in the nucleolus before export to the cytoplasm. The nuclear ribosomes (here referred to as ribosomes) are the intracellular particles on which proteins are assembled, are highly complex and dynamic entities, and their formation is coordinated multistep process (**Fig. 1.7**).

In ribosomes, ribosomal RNA (rRNA) molecules constitute the structural framework, the process is associated with many proteins (in Tschochner and Hurt [57]. Ribosomes are the sites of protein synthesis in all organisms. These organelles basically consist of two subunits (the large, 28S, and the small subunits, 18S) ribosomal RNA. The rDNA locus, from which ribosomal genes are transcribed, is located within a secondary constriction or NOR (nucleolar organizer region) on a number of chromosomes in animal cells (**Fig. 1.7A, B**). They are

sequentially arranged in the rDNA arrays of hundreds of units and are termed ribosomal transcription units (rTUs) or rDNA units [58] (**Fig. 1.7C**). These rDNA loci served as locations of DNA origin for transcription of the rRNA genes, and highly conserved among species suggesting that they have changed relatively little in the billion years of evolution [59, 60]. Formation of ribosomes is a fundamental and essential demand for the cell, and eukaryotic cells must assemble more than 70 ribosomal proteins with four different rRNA molecules (25S/28S, 18S, 5.8S and 5S) into two subunits. These are 60S and 40S subunits constituting the ribosomes. The 60S or large subunit (LSU) is assembled with 25S/28S and 5S/5.8S with 49 proteins and the 40S or small subunit (SSU) is assembled with 18S with 33 proteins [58] (**Fig. 1.7D**).

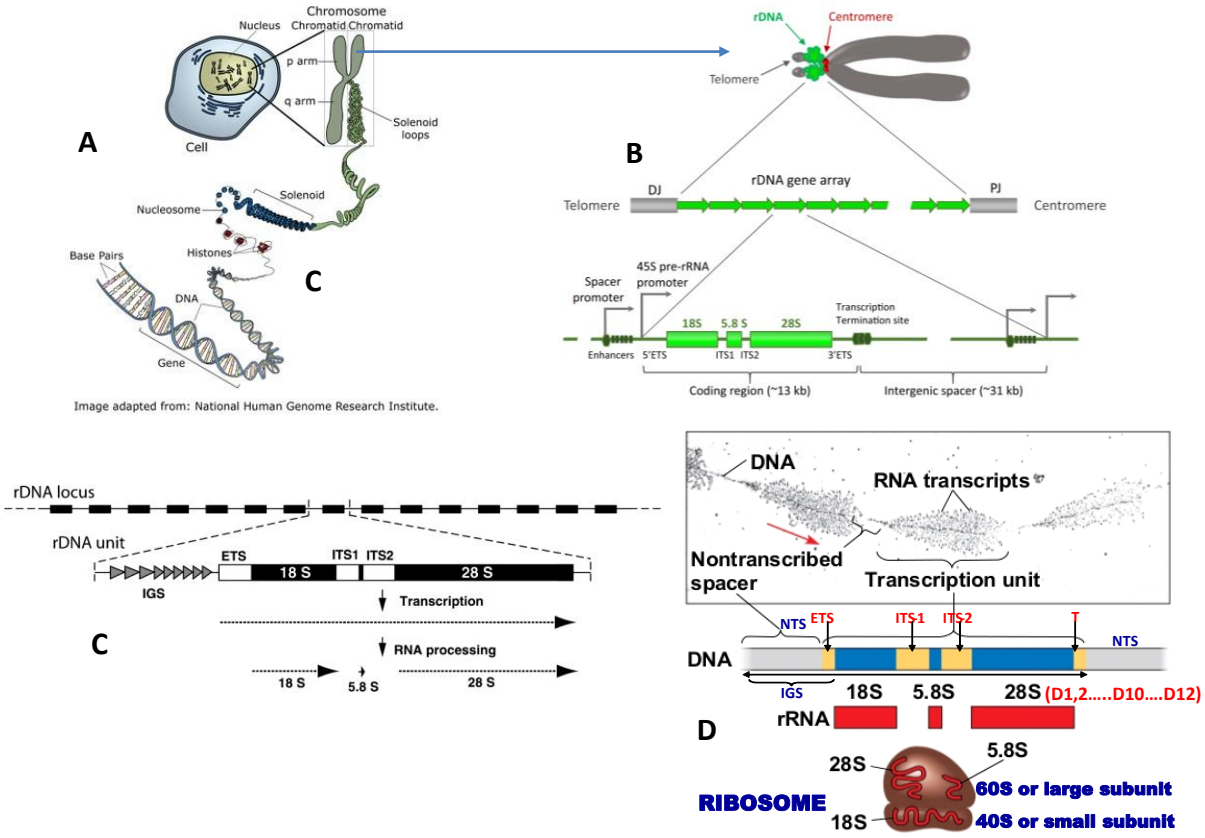


Figure 1.7. Ribosomal transcription units and their transcription in animals and parasites. A. A schematic drawing of a cell and a chromosome with DNA it contains; B. A chromosome with a secondary constriction or NOR (nucleolar organizing region) where the rDNA repeats are located (in the human genome, rDNA repeats are on the short arms of the chromosomes nos 13, 14, 15, 21, 22), and each repeat unit consists of a coding region (encoding pre-mRNA for 18S, 5.8S, and 28S ribosomal RNA subunits) and intergenic spacers. C. Schematic representation of rDNA locus and the transcription of a rDNA unit processing the 18S, 5.8S and 28S rRNA molecules; D. Transmission electron micrograph of transcription of tandemly arranged ribosomal RNA genes and each unit with their rRNA products are assembled as a framework into 60S (large) and 40S (small subunit) for the formation of a ribosome (Sources: A. National Institutes of Health, National Human Genome Research Institute; B. Potapova and Gerton (2019) [58]; C. Eickbush and Eickbush (2007) [59]; D. https://www2.le.ac.uk/projects/vgec/diagrams/36%20chromosome%20unravel.jpg/view).

1.3.2 Organization and genetic location of ribosomal transcription units

Each ribosomal transcription unit (rTU or rDNA) is a chromosomal DNA region of the nuclear genome that codes for three ribosomal genes (named 18S rRNA, 5.8S rRNA, and 28S rRNA genes), two intergenic regions, respectively, termed ITS-1 and ITS-2 (internal transcribed spacers 1 and 2). Flanking this transcribed region (5'-**18S**-ITS1-**5.8S**-ITS2-**28S**-3') is the non-coding regions, named non-transcribed intergenic spacer (IGS). The rTUs are repeated units arranged sequentially into arrays in a complex, with up to hundreds of units, called the nuclear ribosomal operon [61].

In the human nuclear genome, rTUs are located in the secondary constriction region, also known as NOR (nucleolar organizer region), on chromosomes 13, 14, 15, 21 and 22 [58, 60]. They are connected to each other by a non-coding nucleotide sequence containing many repeating structures, called the IGSs, and in some species, there is also an ETS (external transcribed sequence) region, which is merged with the IGS regions [58]. A complete rTU has a characteristic frame organization of: [5'-IGS-ETS-**18S**-ITS1-**5.8S**-ITS2-**28S**-IGS(ETS)-3'], and the rTUs are tandemly arranged in series of up to several hundred copies [62] (**Fig. 1.7**). The gene structure and gene arrangement of rTU are conserved in all species, but the characteristics of each gene and intergenic regions vary [4, 61, 63–65]. Low levels of nucleotide variation are commonly found in closely related species and higher in distant relatives, with the exception of the IGS region nucleotide sequence, which has very high polymorphic variation even within variants or subspecies in all taxa [66, 67].

1.3.3 Characteristics of the ribosomal genes and intergenic regions

The total length of a complete nucleotide sequence of rTU in trematodes, known to date, ranges between 7 and 10.3 kb, including the IGS/ETS region [6, 67–73]. The longest sequence of the complete rTU, to date in trematodes, is probably 10,221 bp found in a strain of *Paramphistomatum cervi* (family Paramphistomatidae, suborder Pronocephalata) [68]. Excluding IGS/ETS, the transcribed region of trematode rTU (reffered to as rTU*), which is the rTU-DNA region from 5' terminus of 18S to 3' terminus of 28S rRNA genes (5'-18S-ITS1-5.8S-ITS2-28S-3'), and is about 6.8 kb to 7.2 kb to-date sequenced [6, 73]. There were five complete or transcribed rTU/rTU* sequences of echinostomes (*Echinostoma revolutum*, *Eca. miyagawai*, *Artyfechinostomum malayanum*, *Hypoderaeum conoideum*, and *Echinochasmus japonicus* have been reported to date [6], and this is a part of the study for this thesis.

The length of the 18S rRNA gene ranges from 1.95 to 2 kb, for example, 1,958 bp in the large liverflukes *Fasciola* spp. (family Fasciolidae), 1,991 bp in the small liverfluke *Clonorchis sinensis* (family Opisthorchiidae), 1,992 bp in the small intestinal fluke *Haplorchis pumilio* (family Heterophyidae), and 1,974–1,977 bp in the lung flukes *Paragonimus* spp. (family

Paragonimidae). The 5.8S rRNA gene is substantially conserved in length (157–160 bp) and nucleotide composition across all taxa including distantly related species in the Trematoda class [61, 68, 71, 72]. The complete 28S rRNA gene is 3.6–4.2 kb in length in trematodes, most of which are about 3.8–3.9 kb. The skew value of nuleotide composition and the secondary structures of various trematodes have been gradually examined [6, 73, 74].

The ITS1 and ITS2 intergenic regions as well as the two ends of the IGS sequences are the least conserved DNA regions in rTU due to their high number of repetitive structures. The 5.8S rRNA gene region is small in size and has little variation among all species of the same family, even distant species [61]. The lengths of ITS1 and ITS2 regions vary greatly amongst species in the same or distinct families, ranging from a few hundred to over a thousand nucleotides. The ITS region can contain a tandemly arranged repeat units (RUs or TRUs), and the repeat numbers vary within and among species [61, 72, 73]. Many trematode species, such as the liver flukes *Fasciola* spp. (454 bp/ITS1 and 359–360 bp/ITS2) and *Eurytrema pancreaticum* (1,103 bp/ITS1 and 231 bp/ITS2), have an ITS (ITS1 and/or ITS2) that lacks RUs and so has a constant length [71, 72]. Another intergenic regions, the IGSs, which connect the previous rTU's 28S rDNA to the adjacent rTU's 18S rDNA, varies in size between strains and within the same species and is highly polymorphic due to the presence of many different structural RUs [61, 67, 68].

The two ribosomal genes, including 18S, 28S rRNA genes as well as the two intergenic regions (ITS-1 and ITS-2), and even the rTU IGS sequences, are extensively utilized in taxonomic analysis, species relationships, and phylogenetic analysis [61, 67, 75]. The molecular markers from the rTUs are also used in identification of species and “cryptic” species, discrimination and determination of independent species, or "hybrid” or “introgressive” hybridization [6, 40, 72, 76–79]. The nuclear ribosomal transcription unit (rTU) sequences, which include the 18S, ITS1, ITS2, and 28S sequences, have proved critical in resolving trematode taxonomic difficulties [4, 61, 78]. Along with the 18S rDNA, partial (approximately 1,200 bp, D1–D3 domain) or complete (range 3.7–3.9 kb) 28S rDNA sequences are increasingly being utilized as potent genetic markers [4, 6, 12, 72].

1.3.4 Secondary structure of ribosomal genes and intergenic regions

All six gene segments and non-coding regions of rTU, i.e., 18S, ITS1, 5.8S, ITS2, 28S, and IGS, have nucleotide sequences capable of creating a folded secondary structure in a three-dimensional model [61, 74, 80]. That is as well, the formation of the secondary structures also is a characteristic of rTU and is ensuring the stability of the scaffolding for the SSU and the LSU of the ribosomes [8, 43, 81]. Besides, in trematodes, there have been increasingly

evidences that the ITS (ITS1 and ITS2) and IGS regions can form *de novo* secondary structures [6, 74].

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, and creating hairpins and loops are usually found in ribosomal RNA conformation for sustaining the gene stability [66, 82]. 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 [82]. 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 rTUs in any species, such as observed in ITS-1/ITS-2 of the lung flukes, *Nanophyetus* spp. and *Paragonimus* spp. (family Paragonimidae) as recently reported [74] and also in echinostomes [9]. The secondary structural core of ITS-2 formed a "*four-finger*" structural pattern, which is highly conserved in all eukaryotes [8, 74].

Similarly, in mitoribosomes in the mitochondrial matrix, the nucleotide sequences of the 12S and 16S rRNA mitoribosomal genes can also form secondary structures but the pattern and complexity of the structures are simpler than those of the cytosolic 18S rRNA and 28S rRNA genes [25, 30].

1.4 Research on the mitogenomes and ribosomal transcription units

1.4.1 Research on the mitogenomes of trematodes (mitogenomics)

The mitogenome of the trematode flukes is a closed DNA circle, ranging in length from 14 kb to 22 kb, depending on species. To date as increasing data confirmed from the trematode mitogenomics, each mtDNA contains 12 mitoprotein-coding genes, PCGs (*atp*6, *cob*, *cox*1–3, *nad*1–6 and *nad*4L), 2 mitoribosomal RNA genes, MRGs (*rrn*L/(16S) and *rrn*S/(12S)), 22 transfer RNA genes (tRNA or *trn*) and a non-coding region (NCR), which is rich in multiple repeat units of variable length [20, 22, 25, 31]. These data have been revealed from the research work on the mitogenomes of species, including the mitogenomic isolation, sequencing, annotation and data evaluation, which is conceptualized as mitogenomics [83].

An organismal cell has several hundred to several thousands of copies of mtDNA. For the phylum Platyhelminthes (flatworms), the complete mtDNA of hundreds of species of trematodes (class Trematoda) and tapeworms (class Cestoda) has been obtained and fully annotated, providing genetic/genomic sources for multi-purpose use (Source: GOBASE: http://gobase.bcm.umontreal.ca/ and GenBank database). In recent years, the number of sequenced and annotated mitogenomes has increased considerably. It can be listed as the mtDNA of the lung flukes, *Paragonimus ohirai* (KX765277; [33]); *P. westermani* of four forms

(AF540958; AF219379; KM280646; and CM017921; [52, 84]); *P. heterotremus* of China (KY952166; [85]), and *P. kellicoti* (MH322000; [86]); of the minute intestinal flukes, *Haplorchis taichui* (GenBank: KF214770; [87]), and *Metagonimus yokogawai* (KC330755); of small liver flukes, *Opisthorchis viverrini*, *O. felineus*, and *Clonorchis sinensis* [88]; *Fasciola gigantica* and hybrids [36]. The full mtDNA of hundreds of other species from 12 other families of the Trematoda class has also been obtained.

The entire mitochondrial genomes for Trematoda are as follows:

i) From the superfamily Troglotrematoidea/suborder Xiphidiata, including the lung flukes, *Paragonimus skrjabini miyazakii* (ON782295; [32]), *P. iloktsuenensis* (ON961029; [35]), *P. ohirai* (KX765277; [33]), *P. westermani* in different existing forms in India and China (AF540958; AF219379, and KM280646; MN412705; MN412706; [52, 84]), *P. heterotremus* from China and Vietnam (MH059809 and KY952166; [85]), and *P. kellicoti* [86]);

ii) From the suborder Opisthorchiata, including the family Opisthorchiidae, the small liver flukes *Opisthorchis viverrini* from Laos (JF739555; [89]), *O. felineus* from Russia (EU921260; [88]), *O. sudarikovi* from Pakistan (MK033132; [90]), *Clonorchis sinensis* from Russia, China, South Korea (FJ381664; JF729303; JF729304; MT607652l; [51, 88, 89]), *Metorchis orientalis* (KT239342), *Metorchis bilis* from Russia (NC_079698); and the Heterophyidae family, including the small intestinal flukes *Haplorchis taichui* from South Korea and Vietnam (KF214770; MG972809l; [87]), *Metagonimus yokogawai* from South Korea (KC330755); and *Cryptocotyle lingua* from Norway (OL853496);

iii) From the suborder Echinostomata, until the project for this thesis started (in 2021) and outside of the results of this thesis, there were 14 mtDNAs of echinostomes have been sequenced and annotated. These included 13 complete mtDNA sequences from the intestinal flukes of the Echinostomatidae family as follows: *A. sufrartyfex* (KY548763), *Eca. miyagawai* of two strains from China (MH393928; MN116740; [38, 91]); *Eca. revolutum* (MN496162; [31]), *H. conoideum* (KM111525; [92]), *Eca. caproni* (AP017706), *Eca. paraensei* (LL250667; KT008005), *Echinostoma* sp. strain JM-2019, and strain GD from China (MH212284; MN116706; [93]), Echinostomatidae sp. strain CA2021 and strain MSB-A19 from the United States (MK264774; MN822299), and *Isthmiophora/Eca. hortensis* (KR062182; [94]) The only complete mtDNA sequenced in the family Echinochasmidae is from *Ecs. japonicus* (isolate EjPT, KP844722) from Vietnam [5]). The complete mtDNAs been resulted from the implementation of our project, which are presented in the thesis, are: *A. malayanum*, former name, *Eca. malayanum* from a strain of Thailand (OK509083; [9]), *Eca. miyagawai* from a strain of Thailand (OP326312; [49]), and *H. conoideum* from a strain of Thailand (PP110501).

Currently, there are about a hundred entire or nearly-complete mtDNAs available in public databases. The mitogenomes of flatworms (platyhelminths), especially echinostomes, have undoubtedly provided valuable genetic resources for species identification, diagnosis, differentiation, phylogeny, evolution, and population genetics [5, 9, 25. 31, 32, 35, 36, 39, 49, 52, 84, 87, 94]. The mtDNAs that cover as many genera in a family, the families in a suborder, and the newly discovered "cryptic," "synonymous," "polymorphic," "sister," or "hybrid" species [95–99], are absolutely important in the studies of the evolution of biological species and taxonomy, taxonomic conditions and rankings, and population genetics.

1.4.2 Research on ribosomal transcription units of trematodes (ribosomal genomics)

Animals have up to several hundred nuclear ribosomal transcription units (rTUs) grouped in arrays of 200–600 units, whereas parasites have roughly 200–300 units. Since the first use about 40 years ago, the rTUs, comprised of 18S small subunit (SSU), 28S large subunit (LSU), and various internal transcribed spacer (ITS) and intergenic spacer (IGS) regions, have continued to provide a source of nucleotide markers for taxonomic identification, diagnosis, classification, epidemiological, evolutionary, and population genetics studies [61, 100]. For molecular systematics, the 18S and 28S rDNA sequences give useful fingerprints and have been extensively utilized to identify interrelationships within and across genera, families, and suborders, as well as across the phylum for platyhelminths [65]. These include the solitary usage of individual genes, the combination of full 18S and 28S rDNA sequences, or even the highly common use of the D1-D3 variable sections of 28S alone [6, 64, 101].

To date, over 60 complete rTUs or near complete (e.g., the full transcribed region, named rTU*) have been fully characterized for a diversity of parasitic flatworms. These, from many, included:

i) From the superfamily Troglotrematoidea/(suborder Xiphidiata): *Paragonimus heterotremus* (OP081040; [73]), *P. ohirai* and *P. iloktsuenensis* from Japan (OP081041; P081042; [35]), *P. skrjabini miyazakii* from Japan (OP081043; [73]), *P. westermani* strains from South Korea and India (OP081045; OP081044; [73]), *P. kellicotti* (partial, 5,338 bp; HQ900670), *Nanophyetus salmincola* two strains from Russia (LN871822; LN871823), and *Collyriclum faba* from Czech (JQ231122; [102]).

ii) From the suborder Opisthorchiata: among 14 reported, the most interesting rTU sequences are from the family Opisthorchiidae, including *Clonorchis sinensis* from China (5 strains, 8,049–8,391 bp; MK450523–MK450527; [67]); *Metorchis orientalis* from China (5 strains, MK482051−MK482055; [67]); *Opisthorchis viverrini* from Vietnam, *O. felineus* from Russia, and *O. parageminus* from Vietnam [102]. From the family Heterophyidae, there were rTUs from *Cryptocotyle lingua* from Russia (MW361240; [103]), *Haplorchis taichui* and *H.*

pumilio from Vietnam [19, 104]), *Euryhelmis costaricensis* from Japan (7,049 bp; AB521797; [105]), and unidentified *Scaphanocephalus* species (*Scaphanocephalus* sp. ne1) from the United States (7,999 bp, PP430581; [106]). From the family Cryptogonimidae, there was only one, *Stemmatostoma cribbi* from Solomon Islands (7,431 bp, OQ968484; [107]) has been reported;

ii) From the suborder Echinostomata: until the five rTUs' sequences of the family Echinostomatidae and Echinochasmidae from this study being added, there were six rTU sequence reports from Vietnam, Australia, and Sri Lanka from the family Fasciolidae, such as *Fasciola hepatica* (7,657 bp; MN970007; [72]) and *F. gigantica* (2 strains: 6,794 bp; MN970009−MN970010); *Fasciola* sp. (hybrid) (7,966 bp; MN970008; [72]); *Fascioloides jacjsoni* (7,781 bp; MN970006; [72]); *Fasciolopsis buski* (8,361 bp; MN970005; [72]); one from the family Philophthalmidae, that is *Philophthalmus gralli* from Peru (7,194 bp, JQ627832; [108]) and one from the family Echinostomatidae, that is *Isthmiophora hortensis* from Japan (6,876 bp, AB189982; [109]).

The five complete rTU sequences of echinostomes resulted from this study have been deposited in GenBank, and are: *Artyfechinostomum malayanum* (9,499 bp, OR509026), the near-complete rTU of *Hypoderaeum conoideum* (8,076 bp, OR509029), and the transcribed regions of rTU (from 5'-terminus of 18S to 3'-terminus of 28S rRNA gene) in *Eca. revolutum* (6,856 bp, OR509028), *Eca. miyagawai* (6,854 bp, OR509027), and *Ecs. japonicus* (7,150 bp, OR509030) [6].

Besides those rTUs from the family Paragonimidae of the superfamily Troglotrematoidea (Xiphidiata), Opisthorchiata, and Echinostomata, there were a couple of tens reported from other families and suborders, such as *Eurytrema pancreaticum* from China (5 strains: 8,310 bp, 8,309 bp, 8,310 bp, 8,306 bp, and 8,309 bp; KY490000–KY490004; [71]); *Schistosoma japonicum* (8,271–8,857 bp; [110]); *Paramphistomum cervi* (5 strains: 8,493 bp, 9,908 bp, 10,056 bp, 10,167 bp, and 10,221 bp; KJ459934; [68]); *Diplostomum pseudospathaceum* (7,991 bp; KR269766; [70]); *Diplostomum spathaceum* (7,993 bp; KR269766; [70]); *Brachycladium goliath* (9,296 bp; KR703279; [69]); and *Diplostomum ardeae* (7,744 bp; MT259036; [106]), and others.

Notably, there was a cumulative number of rTUs recorded, indicating their availability for a variety of uses. However, there are still many species that do not have a complete rTU to be sequenced in the Echinostomata suborder, especially the "cryptic",'synonymous', 'polymorphic', 'sister', or 'hybrid' species [6], and if being available, they will provide valuable genetic resources for species identification, diagnosis, differentiation, molecular epidemiology,

and the studies of the evolution and phylogeny of biological species, taxonomic rankings, and population genetics.

1.5 The importance of studying mitogenomes and ribosomal transcription units and particular aims for taxonomic resolution of echinostomatid species in this study

1.5.1 The importance of studying mitogenomes and ribosomal transcription units of echinostomes

Sequences generated from the mitogenomes provide excellent molecular markers for defining population groups, for tracing the genetic history of an individual or a particular group of related individuals, for classifying the taxonomic rankings, and for constructing deep-branch taxonomic phylogenies [111]. In addition, metazoan mtDNAs exhibit an abundance of genetic novelties that include modified mitochondrial genetic codes; an unequalled variety in the secondary structures of ribosomal RNAs; variable base composition (A+T and G+C contents), which for vertebrates mostly differs from invertebrates; the characteristic replication mode of the mtDNA molecule; the codon bias in usage for protein-encoding genes; the variable and modified structural forms of mt transfer RNAs; the presence of unassigned sequence(s) known as non-coding regions that are rich in repeated sequences and mysterious functional elements within, and the link of mutations in mtDNA to apoptosis and genetic disorders [44, 112, 113].

Information from mitochondrial and ribosomal transcription unit sequences will be just as useful in studies on genetic variation in parasitic helminths, such as platyhelminths and echinostomes, as it has been in vertebrates and insects [55, 114]. Platyhelminth populations, specifically trematodes, particularly echinostomes in our study, are clearly capable of responding to selective forces and hence, genetically diverse. Conventional methods (e.g., isozyme analysis, proteomic, phenotypic, or morphological investigation) for obtaining direct evidence for variation and linking it to evolutionary responses have generally been less effective. Mitochondrial and nuclear ribosomal DNA markers provide additional optimism and next-generation sequencing technology [115] just advance obtaining the whole, realistic mtDNAs and rTUs for analyses in this direction. Nearly 200 complete or near-complete mtDNAs and over 50 complete or near-complete rTUs have been reported in public databases from different classes and orders of Platyhelminthes (as of July, 2024), but only limited information on the mtDNA and rTU complete or near-complete sequences of flatworms of the Echinostomatidae family (Phylum Platyhelminthes: Class Trematoda: Order Plagiorchiida: Suborder Echinostomata) has been available. A few mitogenomes of *Echinostoma* species (*Eca. miyagawai*, *Eca. caproni*, *Eca. paraensei*) and other echinostomes (*Artyfechinostomum sufrartyfex*, *Echinoparyphium aconiatum*, and *Hypoderaeum conoideum*), as well as several unidentified Echinostomatidae and *Echinostoma* spp., and none of complete or nearly complete

rTU sequences for any echinostome species, were available at the start of this study and over the last three years. There is a need for mtDNAs and rTUs that cover as many species in the *Echinostoma* genus and related genera in the Echinostomatidae, and related families in the Echinostomata suborder, as well as the newly identified "cryptic," "synonymous," "polymorphic," "sister," or "hybrid" species [95–99, 116].

Sequences of the nuclear ribosomal transcription unit (rTU), including 18S, ITS1, ITS2, and 28S sequences, have been crucial in resolving taxonomic issues for trematodes [4, 6, 78]. Along with the 18S rDNA, the partial (approximately 1,200 bp, D1–D3 domain) or full (ranging 3.7–3.9 kb) 28S rDNA sequences have been increasingly used as powerful genetic markers [4, 12, 61, 72]. The systematics of the superfamily Echinostomatoidea (Trematoda: Platyhelminthes) is frequently revised due to the addition of new species [5, 9, 31, 97, 98, 111, 117–120, 121]. There are still some difficulties surrounding the taxonomy and phylogenetic relationships of several species in some genera, including *Echinostoma* and *Artyfechinostomum* in the Echinostomatidae and *Echinochasmus* in the Echinochasmidae family, although some clear genus delimitations have been made [4, 5, 9, 12, 13, 78, 95, 122]. In this thesis, our study aimed to present the complete ribosomal transcription units of five echinostomatids and echinochasmids and their use for phylogenetic analyses to update the resolution within and between the families Echinostomatidae and Echinochasmidae and their positions in the superfamily Echinostomatoidea and the suborder Echinostomata. To some extent, the taxonomic and familial phylogenetic relationships within and between several suborders are also discussed. Their mtDNA and rTU datasets are critical for research into echinostome and trematode evolution and phylogeny, taxonomy rankings, and systematics.

1.5.2 The particular aims for taxonomic resolution of echinostomatid species in this study

For the Echinostoma revolutum species

The taxonomic status of *Eca. revolutum* is still controversial, although recently a number of molecular studies identified the parasite as a highly cosmopolitan species comprising of several distinct geographical lineages corresponding to parasite populations with European, American, and Southeast Asian origins [12, 13, 18, 123]. The taxonomic identification and the phylogenetic assessment of each species within the “revolutum” group and as well between member taxa in the family Echinostomatidae requires accurate genomic data. Many attempts of interspecific clarification forthe echinostomatids, particularly for those withinthe “37-collar-spined” taxa have relied predominantly on tenuous morphological features [4, 12, 13, 123]. 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) [4. 12, 63, 95, 123]. 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 [13, 124, 125]. In fact, the analyses of complete mitogenomes to perform taxonomic and phylogenetic analyses of other members of the suborder Echinostomata, as well as other trematode species, has been widely used and has provided not only a deeper understanding of the evolutionary relationships within and between trematode families but also essential molecular markers for population genetics and diagnostics, crucial for modern epidemiological studies [12, 13, 114]. However, many morphologically similar species, notably those of the "collar-spined" *Echinostoma* species of the Echinostomatoidea superfamily, lack comprehensive mitochondrial genomic data. Currently, only four of the nine species of the "*Eca. revolutum*" group, including *Eca. caproni*, *Eca. paraensei*, *Eca. miyagawai*, *Eca. hortense* [19], and a few species within the Echinostomata suborder have entire mitochondrial genomes accessible [38, 39, 90–92]. The aim of the investigation in this thesis was to complete the full mitogenome sequence of *Eca. revolutum*, a worldwide widespread and medically significant species, as well as to correlatively define its mitogenomic properties and compare them to those previously described in the Echinostomatoidea superfamily. A phylogenetic tree for families in the suborders Echinostomata, Opisthorchiata, Pronocephalata, and Xiphidiata is also presented.

For the Echinostoma/Artyfechinostomum malayanum species

Historically, echinostomes have been differentiated into five groups based on morphological characteristics, particularly the structure, position, arrangement, and number of "collar-spines" that sit around the oral sucker [2, 16, 31]. While the most important "revolutum" group has 37-collar-spines, other groups/species exhibit varying numbers of collar-spines, from 31 to 55, as observed on 25–29 on *Echinostoma hortense* or 31 on *Echinostoma anseries*, 43 on *Echinostoma/Artyfechinostomum malayanum* or 43–45 on *Echinostoma aegyptiacum*, 41–45 on *Hypoderaeum conoideum* or 43–50 on *Echinoparyphium recurvatum*, and 49–55 on *Eca. ilocanum* [1, 10–12, 16]. However, the spine collar can often be a tenuous characteristic for species differentiation as these can vary between individuals of the same species, and specimen preparation can also cause the loss of spines before identification takes place, illustrating the need for a robust molecular based approach for species identification.

Echinostoma malayanum Leiper, 1911 was first described infecting people in Malaysia in 1911 and has subsequently been identified in several countries across Asia, including China,

India, Indonesia, Laos, Malaysia, the Philippines, Singapore, Thailand, and Cambodia [10, 15, 18, 126, 127]. The discovery of *A. surfrartyfex* in India [8, 128] caused considerable taxonomic controversy, originally being synonymised with *Eca. malayanum* [129] but later being differentiated based on a few single nucleotide polymorphisms in the ribosomal ITS-1 and ITS-2 regions, suggesting that *Eca. malayanum* should, in fact, be considered as *Artyfechinostomum malayanum* and represents the type species of the genus *Artyfechinostomum* [1]. However, broadly, there is still controversy related to the generic names of echinostomatids, and this is particularly true for *Eca. malayanum*, with several studies suggesting it could also sit within the genera *Euparyphium* or *Isthmiophora* [10, 14–16, 129–131]. As highlighted previously, a major challenge in echinostome taxonomy has been the traditional use of morphologically plastic characteristics, which has led to difficulties in taxonomic classification and resolving phylogenetic relationships between species. This has been a particular issue for the genus *Echinostoma* due to multiple synonyms and the continuous addition of newly described species, leading to frequent revision of Echinostomatidae systematics [1, 12, 13, 95, 132]. The use of molecular markers has solved the interchanged generic and specific classification for particular species and genera within this family, and between families of the suborder Echinostomata [1, 4, 11, 12, 133]. For accuracy in species identification, classification, and phylogenetic relationships, markers from partial or complete mitogenomes are among the most accurate tools, and full-length mtDNA sequences provide a higher level of resolution, as in the case of the assessment of *Eca. miyagawai*, *Eca. revolutum*, *H. conoideum*, and other trematodes [25, 31, 38, 91, 92], although partial mtDNA sequences have been relatively effective in resolving relationships among echinostome taxa [14, 18, 131, 134, 135].

Over the past two decades, there have been considerable efforts to generate complete mitogenomes across the Platyhelminthes but still relatively few for the Echinostmatidae, particularly for species within the genus *Echinostoma* and *Artyfechinostomum*, increasing the challenge of resolving interspecies and intergeneric phylogenetic relationships [1, 4, 11, 13, 31]. The mitogenomes currently available include the conspecific *Artyfechinostomum sufrartyfex* from the Indian strain (GenBank: KY548763) and several from the *Echinostoma* genus (two of *Eca. miyagawai* from China; one of *Eca. revolutum* from Thailand; one of *Eca. caproni* from Egypt; and one of *Eca. paraensei* in GenBank). A detailed mitogenomic analysis and mitophylogenetic assessment of taxonomically confused echinostomes, particularly those related to *Echinostoma*/ *Artyfechinostomum malayanum* and its generic congeners, will facilitate insights into the fine-scale inter-relationships among species of the family Echinostomatidae (and Echinostomata suborder). These analyses also provide molecular markers for further genetic and molecular studies of this large family. Thus, the aim of this

study was not only to present the first full annotation and comparative features of the mitogenome of *Eca. malayanum*, but also to use an in-depth phylogenetic approach to assess the interrelationship between *Eca. malayanum* and *A. sufrartyfex* and to resolve the generic name of *Echinostoma/Artyfechinostomum*.

For the Echinostoma miyagawai and Hypoderaeum conoideum species

Echinostoma miyagawai Ishii, 1932 (Trematoda: Echinostomatidae) is a zoonotic trematode echinostomatid with a global distribution and the potential to cause human echinosomiasis [1]. *Echinostoma miyagawai* was previously considered as a synonym with numberous species such as *Echinostoma revolutum* Froelich, 1802, *Echinostoma robustum* Yamaguti, 1935, *Echinostoma friedi* Toledo et al., 2000, and *Echinostoma echinatum* Zeder, 1803 [1, 98, 136–139]. Recently, this species and *Eca. revolutum* were recognized as different and taxonomically valid species in the Echinostomatidae [12, 13, 122].

Hypoderaeum conoideum (Bloch, 1782) Dietz, 1909, is a neglected zoonotic trematode found in Bangladesh, China, Indonesia, Japan, Mexico, North America, Russia, Spain, Taiwan, and Thailand. It causes economic losses to poultry in various Asian countries [1]. This species uses snails as the first intermediate hosts; bivalves, fish or tadpoles as the second intermediate hosts, and poultry (chickens and ducks) as the definitive hosts. Although the nucleotide sequence of the mitogenome of *H. conoideum* has been reported [92], its structure and annotation require further improvement. More datasets and implications about the molecular and genetic diversity of *H. conoideum* are required, particularly data from the mitogenome, which can serve as a varied molecular source for taxonomic, epidemiological, phylogenetic, and evolutionary studies. Although many of these mitogenome sequences have been published and stored in public databases, there is evidence that some of them are incomplete and shorter in length than they should be. The length of an mitogenome in a trematode might differ between geographical isolates due to the existence of multiple repeats in the non-coding region (NCR).

Despite the continued revaluation and up dating of the phylogenetic relationships among the echinostomatid taxa, molecular studies indeed indicate a need to reinvestigate and reconsider the relationship between echinostome species, not only based on polymorphic differences between mitogenomes but also structural differences, gene content and order. However, more recently there has been a substantial increase in the interest in the non-coding regions (NCR) and their repetitive elements as has been studied in trematode species from the families Fasciolidae, Paragonimidae, Brachycladiidae, Diplostomidae, and Schistosomatidae [31, 32, 51–54, 69, 70, 84, 140–142]. The NCR is commonly referred to as the control region (CR) owing to the occurrence of promoter sites for transcription factor binding, and the origin of mtDNA replication. Thus, the NCR is crucial in the regulation of gene functionality in the

mitochondria [143]. Across many animal phyla, the CR is the most polymorphic region of the mitochondrial genome. It has a higher rate of evolution relative to the genes within the mitochondria, resulting in the accumulation of highly repetitive sequences, which in turn can increase the overall length of the mitochondrial genome [144]. This is a consequence of uniparental inheritance and lack of effective recombination, resulting in a lack of proof reading or the correction of mutations over evolutionary time [144].

The aims of our investigation in this study is to present the complete mtDNA sequences of *Eca. miyagawai* and *H. conoideum*, respectively, using long-read sequencing of the PACBIO system. Additionally, the role and function of the putative promoter sequences and regulatory elements in the LRUs and SRUs of the NCR of the *Eca. miyagawai* RED11 strain from Thailand were investigated.

For the taxonomy of the Echinostomatidae family and the Echinostomata suborder

Echinostomiasis and echinochasmiasis are neglected diseases caused by the intestinal flukes, commonly referred to as echinostomes (families Echinostomatidae and Echinochasmidae of the superfamily Echinostomatoidea and suborder Echinostomata) [1, 7]. Seventeen species from at least seven genera in the family Echinostomatidae Looss, 1899, and six species from the genus *Echinochasmus* Dietz, 1909 (family Echinochasmidae Odhner, 1910) have been implicated in human infections worldwide [1, 2, 7]. Included among these zoonotic genera are *Acanthoparyphium* Dietz, 1909; *Echinoparyphium* Dietz, 1909; *Echinostoma* Rudolphi, 1809; *Himasthla* Dietz, 1909; *Hypoderaeum* Dietz, 1909; *Isthmiophora* Lühe, 1909; *Artyfechinostomum* Lane, 1915; and *Echinochasmus* Dietz, 1909 [1, 5, 7–9]. Although regularly updated, the generic, familial, and phylogenetic relationships among the taxa in the family Echinostomatidae and the suborder Echinostomata still need to be reinvestigated, supplemented, reconsidered, and further resolved [4, 13, 63, 78, 102, 145–147]. Mitogenomes and ribosomal transcription units, particularly those from a newly identified or updated sequenced species, are the best genetic markers for assessing taxonomic, intra- and/or inter-generic and familial phylogenetic relationships [9, 31, 65, 114, 148]. In recent years, the systematics of the genus *Echinostoma*, and more broadly, the families Echinostomatidae and Echinochasmidae in the suborder Echinostomata, have been frequently revised as a result of additional analyses of the full mtDNA or its markers from newly discovered or reclassified synonymous species, such as *Eca. revolutum*, *Eca. miyagawai*, *A. malayanum*, and *H. conoideum*, as well as *Ecs. japonicus* (family: Echinochasmidae) [4].

Mitochondrial and rTU markers, including individual or concatenated gene sequences, have proven crucial in resolving taxonomic and generic difficulties for echinosomes [4, 9, 31, 38, 91, 92]. Despite regular updates, the specific, generic, and evolutionary links among

echinostomatid taxa, as well as family affiliations within the suborder Echinostomata and across closely related suborders, must be reinvestigated, supplemented, reviewed, reassessed, and resolved. Furthermore, intrinsic and intra-specific polymorphisms in mtDNA structural features among taxa need to be investigated again. The purpose of this taxonomy study is to use genetic data across 15 strains of 12 echinostome species from the family Echinostomatidae to conduct comprehensive phylogenetic analysis and construct trees for resolving the taxonomic reappraisal, evolutionary, and phylogenetic investigations in the Trematoda and digenean Plagiorchiida classes.

* * *

Based on these demands and criteria, this study was undertaken under the title "**Study on the genomics of mitochondrial genome and ribosomal transcription units of some intestinal flukes in the family Echinostomatidae of the suborder Echinostomata**". The primary goal of this study was to obtain complete or near complete mitochondrial and ribosomal transcription unit sequences, preferably by next-generation sequencing using long-read sequencing of highly-multiplexed long-amplicons for several species of echinostomes in the family Echinostomatidae of the suborder Echinostomata, that are of medical and/or veterinary importance. The annotated data was used for comparative genomic analysis to investigate the evolution and phylogeny of biological species, taxonomic rankings and classification, and provision of the mito- and ribosomal datasets for species identification, diagnosis, differentiation, and molecular epidemiology research.

CHAPTER 2

Materials and Methods

2.1 Parasite samples and species identification

The research subjects are intestinal flukes of the genus *Echinostoma* (family: Echinostomatidae), including *Eca. revolutum*, *Eca. malayanum* (syn. *Artyfechinostomum malayanum*), *Eca. miyagawai*, and *Hypoderaeum conoideum*, and the genus *Echinochasmus* (species: *Echinochasmus japonicum*; family: Echinochasmidae). Four echinostomatid species were subjects for obtaining the complete mtDNA and five including these species and *Ecs. japonicus* for the entire rTU (**Table 2.1**). The research imlementation on two main subjectives: i) mitochondrial genome (mtDNA); ii) ribosomal coding unit (rTU) was described in the layout and steps in **Fig. 2.1**.

Table 2.1 List of *Echinostoma* and *Echinochasmus* samples that were used to obtain the entire mitogenome and ribosomal transcription units

Species and strains	Country's origin	Mitochondrial genome	Ribosomal transcription unit
Echinostoma revolutum (MSD15)	Thailand	X	X
Echinostoma malayanum (EMI3) (syn. *Artyfechinostomum malayanum*)	Thailand	X	X
Echinostoma miyagawai (RED11)	Thailand	X	X
Hypoderaeum conoideum (RED42)	Thailand	X	X
Echniochasmus japonicus (EjPT)*	Vietnam		X

Note: The complete mitogenome of this species was obtained in previous study [4].

Adult *Eca. revolutum* flukes were obtained from the intestines of the naturally infected domestic ducks from abattoirs in Khon Kaen province, Thailand. The samples were designated as MSD15. The flukes were thoroughly washed in physiological saline and morphologically identified based on size of the body and circumoral disc, the appearance of testes and the presence of "37-collar spines" around head [12, 149]. 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 [16, 18, 123].

Adult *Eca. malayanum* (syn. *Artyfechinostomum malayanum*) flukes were recovered from the intestines of experimental hamsters fed on cysts containing metacercariae collected from the freshwater snail *Indoplanorbis exustus* in Khon Kaen province, Thailand. The collected flukes were morphologically identified by light microscopy based on the size of the body, and

the appearance of testes, and the presence of '43-collar spines' arranged in two alternating rows along the dorsal side, around the head [15]. The flukes were thoroughly washed in physiological saline and then individually fixed in 70% ethanol (v/v) and stored at −20 °C. Species of *Echinostoma* were molecularly confirmed by sequence and phylogenetic analyses using nuclear ITS-2, mitochondrial *cox*1 and *nad*1 markers [18, 131].

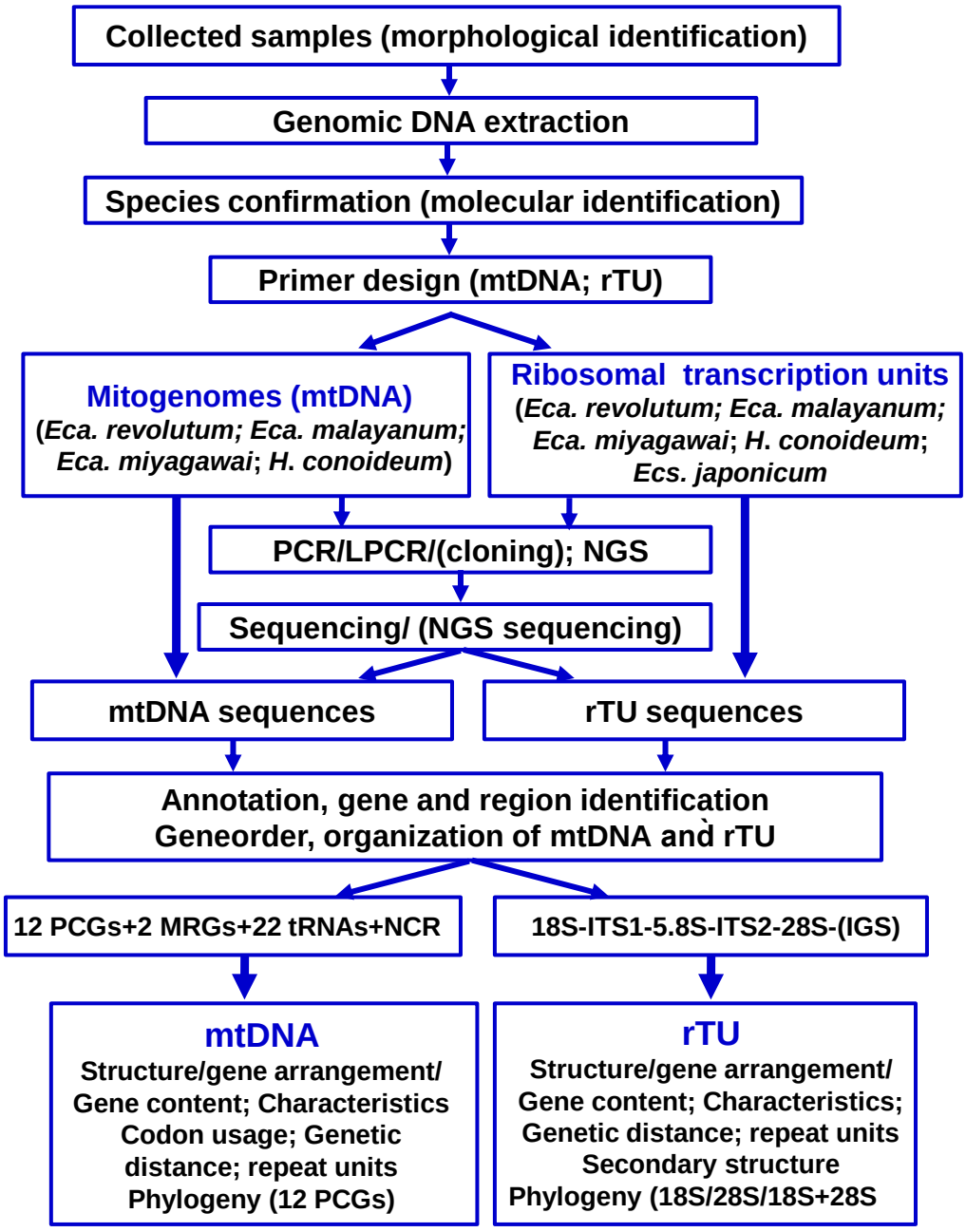


Figure 2.1. Layout and steps to implement research on two main subjectives: i) mitochondrial genome (mtDNA); ii) ribosomal coding unit (rTU). Note: Eca: *Echinostoma*; Ecs: *Echinochasmus*; PCR: polymerase chain reaction; NGS: next-generation sequencing; PCGs: protein coding genes; MRGs: mitochondrial ribosomal genes; tRNAs: amino acid transfer RNAs; 18S: 18S rRNA gene; 5.8S: 5.8S rRNA gene; 28S: 28S rRNA gene; ITS1: internal transcribed spacer 1; ITS2: internal transcribed spacer 2; IGS: non-transcribed intergenic spacer.

Adult flukes of *Eca. miyagawai* and *H. conoideum* were collected from the intestines of the naturally infected domestic ducks in Roit Et province, Thailand from abattoirs and were thoroughly washed in physiological saline. The samples were designated as RED11 for *Eca. miyagawai* and RED42 for *H. conoideum.* The flukes were morphologically examined and molecularly identified using *nad*1 and *cox*1 markers [13, 18, 123].

Adult *Echinochasmus* flukes were collected from humans including *Ecs. japonicus* in Phu Tho and Hoa Binh provinces, and *Ecs. perfoliatus* in Ha Noi City, Vietnam. Flukes were identified using morphological criteria [132], as follows: head crown with 24 collar-spines arranged in a row around the oral sucker, interrupted mid-dorsally, uterus short, and two large transversely stretched testes. Additional samples, *Patagifer bilobus* was collected from a little egret in Binh Dinh province, which was used for obtaining the ribosomal D1–D3 region for Echinostomatidae phylogeny. The species were molecularly confirmed by sequence and phylogenetic analyses using nuclear ITS-2, mitochondrial *cox*1 and *nad*1 markers.

2.2 Total genomic DNA extraction

Total genomic DNA was extracted from individual adult worms using the GeneJET™ Genomic DNA Purification Kit (Thermo Fisher Scientific Inc., MA, USA) according to the manufacturer's instructions. The genomic DNA was eluted in 50 µL of the elution buffer and kept at –20 °C until use. The DNA content was quantified using a NanoDrop® ND-1000 UV-Vis Spectrophotometer. For conventional PCR or mitogenomic DNA enrichment for NGS sequencing, a working concentration (50 ng/µL) was prepared and an amount of 2 µL was used in each long-range PCR *(*LPCR) in a 50 µL reaction volume.

2.3 Primer design and PCR strategies

The initial platyhelminth-universal primers were designed based on mitochondrial nucleotide sequences conserved among all trematode and/or Echinostomatidae mitogenomes available in GenBank or previously published [31, 33]. They were paired to bind to the target regions for amplification of long-range PCR (LPCR) of 4.0–7.5 kb overlapping fragments. The sequence data obtained was used to design further primers for sequencing or for additional PCR/LPCRs (for *Eca. revolutum*, see **Supplementary Table S2.1** and for *Eca./Artyfechinostomum malayanum*, see **Supplementary Table S2.2**). Long-range PCRs were carried out using LongAmp Taq 2X Master Mix from New England Biolabs (Ipswich, MA, USA). All LPCRs were prepared in 50 µL volumes with the addition of DMSO (dimethyl sulfoxide) to 1.5% and performed in an MJ PTC-100 or equal Thermal Cycler, with an initial denaturation at 94 °C for 3 min, followed by 30 cycles, each consisting of a denaturation step for 30 sec at 94 °C, annealing at 50 °C for 30 sec, extension at 65 °C for 8 min, and a final

extension at 65 °C for 10 min. The LPCR 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). Primer-walking sequencing was applied to each fragment and the sequence from the LPCR products spanning the NCR (around 6.5 kb) was obtained by the next generation sequencing strategy using the PacBio system at the Institute of Biotechnology, Vietnam. All of these sequences were subsequently assembled to obtain the complete mitogenome for each species.

The rTU-universal primers, including those used for amplification and sequencing of the coding region (from 5' terminal of 18S rRNA to 3' terminal of 28S rRNA genes), and the external transcribed spacer (ETS) and IGS, were described in Le et al. [72; 79]. Three primer pairs mainly used for amplification are: i) UD18SF (forward): 5' AACCTGGTTGATCCTGCCAG 3' and U3SR (reverse): 5' CGACCCTCGGACAGGCG 3'; ii) U3SF (forward): 5' GGTACCGGTGGATCACTCGGCTCGTG 3' and 28ENDR (reverse): 5' TTCTGACTTAGAGGCGTTCAGTC 3'; and iii) 28ENDF/ U28TUF (forward): 5' TTCTGACTTAGAGGCGTTCAGTC 3' /or 5' CGACGTCGCTTTTTGATCCTTCG 3' and U18TUR (reverse): 5' CGGGTCAGGGCATAGTGGC 3'. Other primers used for sequencing are listed in Le et al. [72; 79].

2.4 Approach for next-generation sequencing (NGS)

2.4.1 Targeted enrichment of the mtDNAs by long-range polymerase chain reaction

Seven primer pairs, including trematode-universal and Echinostomatidae-universal primers, were designed. These primer pairs were used for amplifying the whole mtDNA of *Eca. miyagawai* and *H. conoideum* in seven overlapping fragments, respectively, using long-range PCRs (LPCR), and these seven amplicons were used for NGS (**Supplementary Table S2.3**). LPCRs were performed in a 50-µL volume in a MJ PTC-100 Thermal Cycler. Each reaction contained 25 µL of 2X LongAmp Master Mix (New England Biolabs, Ipswich, MA, USA), 2 µL of each primer (10 pmol/µL), 2 µL of template DNA, and 19 µL DEPC-water. LPCRs were conducted with an initial denaturation at 94 °C for 1 min, followed by 30 cycles, each consisting of a denaturation step for 30 s at 94 °C, an annealing/extension step at 50 °C for 30 s, an extension at 65 °C for 8 min, and a final extension at 65 °C for 10 min. The LPCR 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).

The dsDNA products were purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific, USA), and the amplicon length was verified via 1.5% agarose gel electrophoresis. Six amplicons from the coding mtDNA and two from the NCR were pooled for NGS. The complete mitogenome of *E. miyagawai* was sequenced using the PacBio

SEQUEL system (https://www.pacb.com/) with a targeted long-read sequencing approach at the PacBio facility at the Institute of Biotechnology (Hanoi, Vietnam)

2.4.2 Library preparation, long-read sequencing and *de novo* assembly of the mtDNA sequences

Library preparation: The dsDNA products of each sample from seven overlapping amplicons were pooled into one tube and purified with AMPure XP beads (Pacific Biosciences, Menlo Park, CA, USA). Input dsDNA was quantified using the Qubit fluorometer 3.0 and Qubit dsDNA HS Assay reagents (Thermo Fisher Scientific, Waltham, MA, USA). SMRTbell Libraries were prepared using the Express Template Prep Kit 2.0 with multiplexing amplicons protocol with low DNA input (100 ng) (Pacific Biosciences, Menlo Park, CA, USA) for sequencing on the PacBio SEQUEL system according to the manufacturer's instructions. The SMRTbell templates were purified once with 1.2 volumes of AMPure PB beads, and the size and amount of the library were checked again using the Bioanalyzer 2100 system (Agilent, CA, USA) and the Qubit fluorometer 3.0 with Qubit™ dsDNA HS Assay reagents, respectively. The libraries of all amplicons were then pooled before long-read sequencing.

Sequencing and de novo assembly: The pooled library was bound to polymerase using Sequel Binding and the Internal Control Kit 3.0 (Pacific Biosciences, Menlo Park, CA, USA) and purified using AmpurePB beads. The DNA Control Complex 3.0 and the Internal Control Kit 3.0 from Sequel Binding and Internal Control Kit 3.0 were used to control thesequencing procedure. The final library was loaded onto Sample Plate (Pacific Biosciences, Menlo Park, CA, USA). The run design was created by the Sample Setup software included in the SMRTLink portal v5.1 version 9.0 with an insert size of1200 bp. The sequencing signals were processed, evaluated, and converted into raw data by the Primary Analysis Computer server. All data wasautomatically transferred to the Secondary Analysis Server system via the intranet. High quality sequence data was proofread and generated by PacBio's circular consensus sequencing (CCS), then *de novo* assembled usingCanu software v2.0[150], and the quality of the assembly was checked by using Quast software v5.0.2 [151].

2.5 Mitogenomic annotation

Protein-coding genes (PCGs) were identified by alignment with the available mt genomes of other *Echinostoma* or other trematode strains and species. For each PCG, ATG/GTG as start and TAA/TAG as stop codons were used to define individual gene boundaries. PCGs were translated using the "*Echinoderm mitochondrial genetic code*", which is the translation Table 9 in GenBank. The nucleotide composition for PCGs, MRGs, and the mtDNA coding region (5'-*cox*3 to *nad*5-3', designated as mtDNA*) were analyzed with MEGA 11 [152]. Codon usage for all concatenated 12 PCGs from each strain/species was determined with the program

GENE INFINITY (Codon Usage: http://www.geneinfinity.org/sms/sms_codonusage.html). Genetic distance was based on the pairwise amino acid comparison of concatenated 12 PCGs among members of the family Echinostomatidae and was determined using GENEDOC 2.7 (http://iubio.bio.indiana.edu/soft/molbio/ibmpc/genedoc-readme.html) for alignment and MEGA X (https://www.megasoftware.net/) for percentage estimation [153].

Transfer RNA genes (tRNA or *trn*) were identified using tRNAscan-SE 1.2.1 program (www.genetics.wustl.edu/eddy/tRNAscan-SE/) [154] as well as ARWEN at http://mbio-serv2.mbioekol.lu.se/ARWEN [155]. The ribosomal 16S (*rrn*L) and 12S (*rrn*S) RNA genes were recognized as the regions located between $tRNA^{Thr}$ (*trn*T) and $tRNA^{Cys}$ (*trn*C), and between $tRNA^{Cys}$ and *cox*2, respectively. The non-coding region (NCR) was located between $tRNA^{Gly}$ and *cox*3 or $tRNA^{Glu}$ and *cox*3, depending on each species as previously assigned in the mtDNAs of other *Echinostoma* species [9, 31, 38, 91, 93].

The non-coding region (NCR) in trematodes' mtDNAs was determined by recognition of boundaries between $tRNA^{Glu}$ (*trn*E) and *cox*3 or $tRNA^{Gly}$ (*trn*G) and *cox*3. Tandem Repeat Finder v3.01[156] was used to detect repeat units (RUs), including long (LRUs) and short repeat units (SRUs) in the NCR of mitogenome of the studied echinostomes.

The circular map and gene abbreviations on the map were created in Powerpoint based on the estimated length of each gene/region/tandem repeats. Preferably, the circular map and gene abbreviations on the map were created using the GenomeVx v2.0 drawing tool (http://conantlab.org/GenomeVx/) [157].

2.6 Identification of structural features and promotor sequences in the non-coding region

To identify putative promotor regions in LRU and SRU sequences they were submitted to SAPPHIRE.CNN (SAPPHIRE (kuleuven.be), a web based server that employs neural network algorithms to predict promoter regions in prokaryotic sequences[158]. Finding of such sequences, the *E. miyagawai* LRU and SRU sequences were submitted to the UNAFold Web Server (www.unafold.org) using default settings.

2.7 Comparative mitogenomic analysis of the Echinostomatidae family

2.7.1 Gene identity comparisons

Gene nucleotide comparison for divergence rate (%) among **15 echinostome strains of 12 species** of the Echinostomatidae was conducted (see: information of 15 strains and 12 species in **Supplementary Table S2.4**). This comparison was done between *Eca. myiagawai* strain RED11 and other 14 echinostomatid congeners, and estimated based on the alignment of individual genes, PCGs, and MRGs using MAFFT v7.407 (Katoh and Standley 2013) [159], curated using BMGE v1.12 [160] in the NGPhyogeny package (available at

https://ngphylogeny.fr) [161]. The MEGA 11 program (https://www.megasoftware.net/) [152] was used for percentage calculation.

2.7.2 Base composition and skewness

The nucleotide (base) composition for PCGs, MRGs, and the mtDNA coding region (5′-*cox*3 to *nad*5-3′, designated as mtDNA*) were analyzed using MEGA 11 [152]. Skew values (unequal representation of complementary bases on the same strand), ranging from −1 to +1, were determined by calculating the percentage of AT and GC nucleotide usage using the formula: AT skew = (A − T)/(A + T), and GC skew = (G − C)/(G + C) [162], where the letters represent the absolute usage of the corresponding nucleotides in the sequences. The AT and GC skewness values were calculated for PCGs, MRGs and mtDNA* (the coding region from *cox*3 to *nad*5).

2.7.3 Codon usage and bias in protein-coding genes

Codon usage (referred to as the number and frequency of each codon type) and codon bias (the tendency of using a certain codon instead of others that encode the same amino acid) for all concatenated aminoacids of 12 PCGs from each strain or species was determined with the online program GENE INFINITY (Codon Usage: http://www.geneinfinity.org/sms/sms_codonusage.html), and codons for the usage of each strain or species of echinostomes were summarized.

2.7.4 Pairwise genetic distances among echinostomes

Genetic distance (GD) is a measure of the genetic divergence between species or between populations within a species or closely related species. The GD is commonly used to construct phylogenetic trees and estimate divergence times for closely related populations. In this study, genetic distance was estimated based on the pairwise nucleotide comparisons of concatenated 12 PCGs among 15 strains of 12 species of the Echinostomatidae family using NGPhylogeny (at https://ngphylogeny.fr). The GD value was estimated based on the analysis of the concatenated nucleotide sequences of PCGs for alignment, and MEGA 11 [152] was used for percentage estimation.

2.8 Annotation and sequence analysis of ribosomal transcription units

2.8.1 Annotation of the rTU sequences

The entire rTU's nucleotide sequence for each echinochasmid and echinostomatid sample was obtained after editing chromatograms using Chromas 2.6.6 (http://technelysium.com.au/wp/chromas/). Each ribosomal RNA coding gene (18S, 5.8S, and 28S rRNA genes) and internal transcribed spacers (ITS1 and ITS2) were determined by using the previously published reference sequences and those available in GenBank. These reference rTU sequences were from several species, including *Isthmiophora hortensis* [109],

Paramphistomum cervi [68]; *Philophthalmus gralli* [108], *Brachycladium goliath* [69]; *Eurytrema pancreaticum* [71]; *Clonorchis sinensis* and *Metorchis orientalis* [67], *Fasciolopsis buski* and *Fasciola* species [72], and *Paragonimus* species [35]. ITS1 was located between the 18S and 5.8S genes; ITS2 between the 5.8S and 28S genes; the external transcribed spacer (ETS) (if any) is upstream of 18S; and the non-transcribed intergenic spacer (IGS) is downstream of the 28S rRNA genes. Repeat units (RUs) were detected in the ITS1, ITS2 or IGS using the Tandem Repeat Finder v3.01 [156].

2.8.2 Modeling the *de novo* structure of the 28S rRNA gene

The nucleotide sequence of RNA or DNA containing many regions rich in A and T, or G and C nucleotides that symmetrically run in opposite directions giving rise to pair up to form "*hairpin*" and "*loop*" structures. These secondary structures maintain gene stability for rRNAs in the ribosomes [82]. All three rRNA genes (18S, 5.8S, and 28S rRNA sequences) and two ITS (ITS-1 and ITS-2) can form secondary structures. Here, we pick the 28S rRNA gene sequences from rTUs of four species of the genera *Echinostoma* (*Eca. revolutum* (3,863 bp); and *Eca. miyagawai* (3,861 bp)), *Artyfechinostomum* (*A. malayanum* syn. *Eca. malayanum* (3,863 bp)); and *Hypoderaeum* (*H. conoideum* (3,863 bp)) to generate their *de novo* secondary structures. 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 [107].

2.9 Mitophylogenetic analysis and tree construction

For phylogenetic study of mitogenomes, we used the concatenated amino acid sequences from 13 PCGs. Briefly, the concatenated PCGs were imported into GENEDOC 2.7 and translated into amino acids using the "*Echinoderm and flatworm mitochondrial genetic code"* (Translation Table 9 in GenBank). For phylogenetic studies, we utilized the PhyML software package from NGPhylogeny (available at https://ngphylogeny.fr). In summary, the input concatenated amino acid sequences in FASTA format were uploaded and aligned using MAFFT v7.407, then curated with BMGE v1.12, and the tree was inferred with PhyML v3.3.1 using the maximum likelihood method with 1000 bootstrap replicates [163]. The best-quality final sequence block of 2949–3107 aa was implemented for final analysis. The resulting Newick tree (.nwk) [164] was visualized using the FigTree 1.4.4 program [165]. Phylogenetic analysis and tree reconstruction, including the outgroup sequence, were completed using the maximum likelihood (ML) analysis in the MEGA 11 program[152]. The substitution model with the best score, according to the Bayesian information criterion, was the Jones, Taylor, and Thornton + F + G + I model (JTT + F + G + I), 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) [152].

Concatenated aligned amino-acid sequences of 12 PCGs, as usually implemented for mitophylogeny of trematodes, from **57 strains of 41 trematode** species of ten families from three suborders, i.e., Echinostomata (i.e., Echinostomatidae, Cyclocoelidae, Echinochasmidae, Fasciolidae, and Himasthlidae), Opisthorchiata (Opisthorchiidae and Heterophyidae), and Xiphidiata (families Paragonimidae and Dicrocoeliidae), with *Schistosoma haematobium* species (family Schistosomatidae) as an outgroup, were used in a comparative phylogenetic analysis (listed in **Supplementary Table S2.4**). In addition to three strains of *Eca. miyagawai* (the RED11 of Thailand and the Hunan and HLJ strains of China), the eleven available echinostomid mtDNA sequences from 10 species of the family Echinostomatidae were included. Among these were four representatives of the 37 collar-spined 'revolutum' group, including *Eca. caproni*, GenBank: AP017706, from Egypt; *Eca. paraensei*, GenBank: KT008005, two artyfechinostomid species (*Artyfechinostomum malayanum*, EMI3 strain, Thailand and *Artyfechinostomum sufrartyfex*, Shillong strain, India), two strains of *Hypoderaeum conoideum* (the RED42 of Thailand and the Hubei strain of China) [92], one *Echinoparyphium aconiatum* (Chany strain, Russia) [142], and three genus- or family-level identified species (the Echinostomatidae sp. CA-2021 isolate PE4, United States (MK264774); *Echinostoma* sp. isolate JM-2019, China (MH212284); and Echinostomatidae sp. MSB para 30070 isolate A19, United States (MN822299)) [9, 31, 38, 91, 92, 142]. One echinostomatid species, the GD strain (China, MN116706) [93], was reported as "*Echinostoma revolutum*", but due to the lack of strong *Echinostoma* generic evidence, it was listed in the "*cryptic*" genus "*incertae sedis*" within the Echinostomatidae [9], was also included in the analysis.

2.10 Ribosomal phylogenetic analyses and tree reconstruction

To examine the phylogenetic and taxonomic position of the Echinochasmidae and Echinostomatidae, **three phylogenetic trees** were reconstructed from the alignment of sequences of the ribosomal rRNA genes: i) concatenated 18S +28S sequences for examining the phylogenetic and taxonomic position of the Echinochasmidae and Echinostomatidae; ii) complete 28S sevquences for examining the congruence of echinostome relationships in the Echinochasmidae and Echinostomatidae; iii) partial sequences (D1–D3 regions, about 1.1–1.3 kb) of the 28S rRNA genes for investigating the taxonomic and generic relationships among the echinochasmid and echinostomatid taxa in the suborder Echinostomata.

Sixty complete or near-complete rTUs from 42 species of 21 families of digenean trematodes, including five from Echinochasmidae and Echinostomatidae newly reported here, were used for ingroup data. The complete **18S and 28S** sequences were concatenated for each

of these 60 species (about 5,767–5,878 bp, prior to alignment) and aligned for phylogenetic analysis. The echinochasmid and echinostomatid families were Echinochasmidae (n = 1) and Echinostomatidae (n = 5), including five newly obtained rTUs. Other digenean families were of the suborders Echinostomata, Opisthorchiata, Pronocephalata, and Xiphidiata (**Supplementary Table S2.5**).

To examine congruence of relationships inferred from 28S sequences along and those inferred from concatenated 28S and 18S ribosomal sequences, another phylogenetic analysis of 70 sequences of the **complete 28S rRNA gene** (about 3,829–3,899 bp, prior to alignment) was conducted. Sixty of the 28S sequences were those mentioned above, along with a further 10 available from GenBank. For both the concatenated and single 28S sequence analyses, *Schistosoma edwardiense* from Schistosomatidae was used as an outgroup (information and author references for each are given in **Supplementary Table S2.5**).

To investigate the taxonomic and generic relationships among the echinochasmid and echinostomatid taxa, we used 169 partial sequences (from 98 species of 50 genera) of the 28S rRNA genes (**D1–D3 regions**, about 1.1–1.3 kb prior to alignment), which were downloaded from GenBank or newly sequenced here. Of these, 154 sequences from 85 species of 42 genera were from the suborder Echinostomata. A partial 28S sequence of *Schistosoma haematobium* (Schistosomatidae) was used as an outgroup (**Supplementary Table S2.6**).

We used the PhyML software package, available at https://ngphylogeny.fr, for phylogenetic analyses of the above three ribosomal datasets. In summary, the input sequences in FASTA format were uploaded and aligned using MAFFT v7.407, then curated using BMGE v1.12, and trees inferred by PhyML v3.3.1 using maximum likelihood with 1000 bootstrap replicates. The resulting Newick tree (.nwk) was visualized using the FigTree 1.4.4 program [165]. Phylogenetic analyses and tree reconstructions, including the outgroup sequences, were also completed using the maximum likelihood (ML) analysis in the MEGA 11 program [152]. The substitution model with the best score according to the Bayesian information criterion was the (GTR+G+I) model, with residue frequencies estimated from the data (GTR), rate variation along the length of the alignment (+G) and allowing for a proportion of invariant sites (+I).

CHAPTER 3

Results

3.1 Mitogenomic genes and gene order of *Echinostoma revolutum*, *Echinostoma/ Artyfechinostomum malayanum*, *Echinostoma miyagawai* and *Hypoderaeum conoideum*

3.1.1 *Echinostoma revolutum*

The complete mitochondrial genome of *Eca. revolutum* (strain MSD15) was shown to be 17,030 bp in size (GenBank accession no. MN496162) (**Fig. 3.1A**). As common in other trematodes, the *Eca. 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 3.1**).

Table 3.1 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$^{Ser1(AGN)*}$	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)	9606–10359	756		0
*cox*2	10360–10968	609/201/(ATG/TAA)		+11
*nad*6	10980–11432	453/150/(ATG/TAG)		+3
tRNA^{Tyr}	11433–11497	65	GTA	+11
tRNA$^{Leu1(CUN)}$	11498–11561	64	TAG	-2
tRNA$^{Ser2(UCN)*}$	11560–11624	65	TGA	+10
tRNA$^{Leu2(UUR)}$	11635–11697	63	TAA	-2

bp: basepair; aa: amino acid; start: start codon; stop: stop codon; Int. seq.: intergenic sequence (+. number of nucleotides before start of following gene; -, number of nucleotides overlapping with following gene); *Asterisk:

tRNA^{Arg}	11696–11762	67	TCG	-2	tRNAs lacking DHU-arm. LRU: Long repeat unit, numbered from the tRNA^{Glu} end; SRU: Short repeat unit, numbered from *cox*3 end; Int. Spacer (IntS): internal spacer sequence between LRU7 and SRU4; unique seq: nucleotide sequence between SRU1 and *cox*3.
*nad*5	11761–13326	1566/521/(GTG/TAG)		+12	
tRNA^{Gly}	13339–13405	67	TCC	+11	
tRNA^{Glu}	13417–13481	65	TTC	+7	
Repeat units	13489–16912				
LRU1	13489–13805	317		0	
LRU2	13806-14122	317		0	
LRU3	14123–14439	317		0	
LRU4	14440–14756	317		0	
LRU5	14757–15073	317		0	
LRU6	15074–15390	317		0	
LRU7	15391–15707	317		0	
Int. Spacer	15708–16084	377		0	
IntS-half 1	15708–15895	188		0	
IntS-half 2	15896–16084	189		0	
SRU4	16085–16291	207		0	
SRU3	16292–16498	207		0	
SRU2	16499–16705	207		0	
SRU1	16706–16912	207		0	
unique seq	16913–17030	130		0	

Echinostoma revolutum has typical mtstructural 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 3.1**). 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 7 genes used TAG and 5 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 tRNATh; and tRNA^{Val} and tRNA^{Ala}), respectively.

The mt genome of *Eca. revolutum* encodes twenty-two 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 3.1; Fig. 3.2**). 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 *Eca. miyagawai*, *Echinochasmus japonicus*, *Fascioloides magna*, *Fasciola hepatica*, *F. gigantica*, and Asian *Schistosoma* species [5, 30, 36−38, 91].

3.1.2 *Echinostoma/Artyfechinostomum malayanum*

The complete mitogenome of the intestinal fluke *Echinostoma malayanum* Leiper 1911, obtained from the sample designated EMI3 from Khon Kaen, Thailand, was 17,175 bp in length (GenBank accession no. OK509083) (**Table 3.2; Fig. 3.1B**). The circular mtDNA molecule comprised 12 PCGs (*cox*1−3, *cob*, *nad*1−6, *nad*4L, *atp*6), two MRGs (16S or *rrn*L and 12S or *rrn*S), and 22 tRNAs or *trn*, and a non-coding region (NCR) rich in long and short tandem repeats (LRUs and SRUs). The gene order and the length of individual genes are similar to that of *Artyfechinostomum sufrartyfex* except for the NCR (**Table 3.2**).

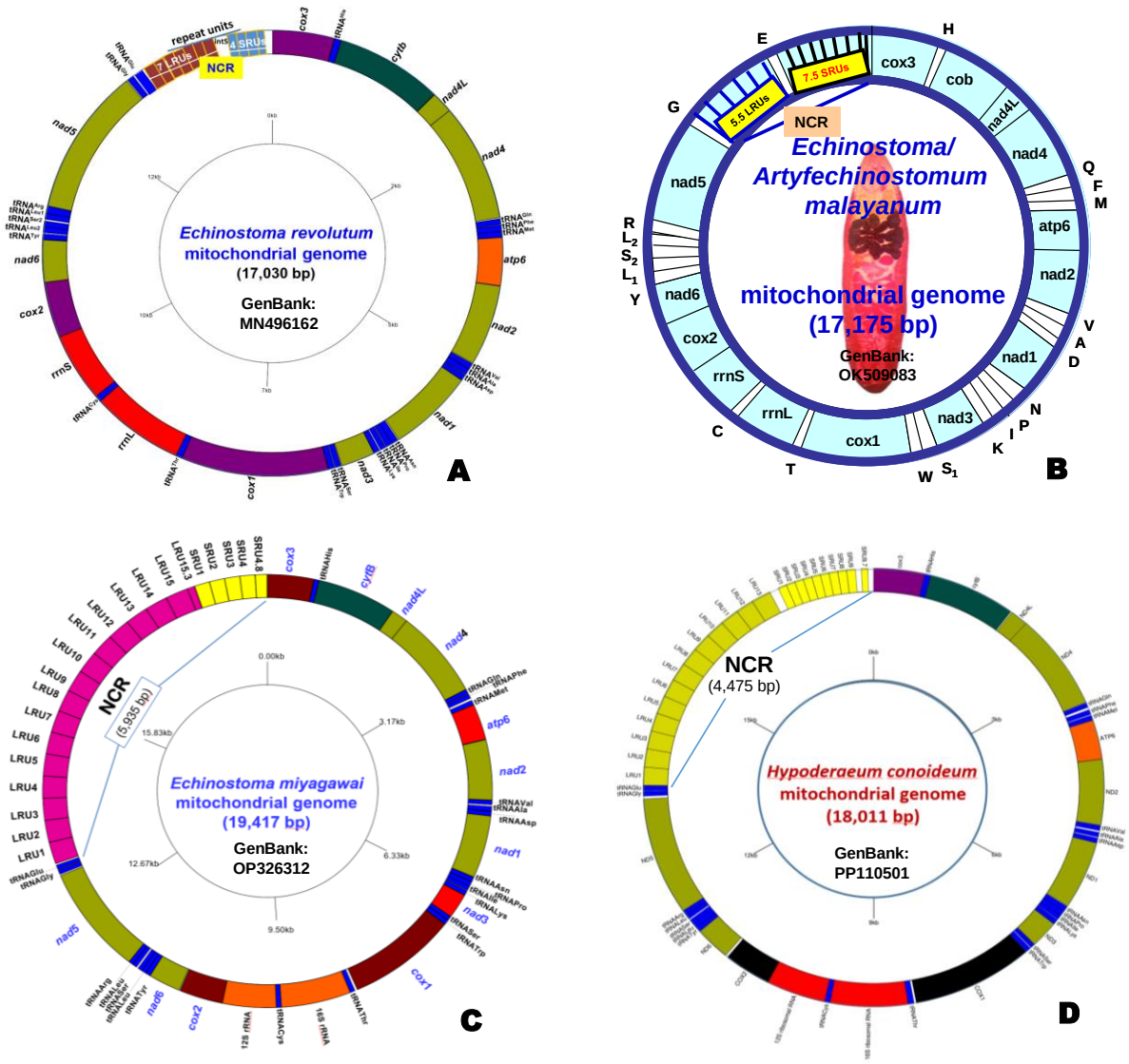


Figure 3.1. A schematic drawing of circular map of the mitogenome of four echinostomes.
A. *Echinostoma revolutum* (GenBank: MN496162); **B.** *Echinostoma/Artyfectostomum malayanum* (GenBank: OK509083); **C.** *Echinostoma miyagawai* (OP326312); and **D.** *Hypoderaeum conoideum* (GenBank: PP110501). The circular mtDNA maps were created in Powerpoint or using GenomeVx v2.0 (http://conantlab.org/GenomeVx/). The protein-coding genes (PCGs) and mitoribosomal large and small subunit genes (MRGs) are abbreviated according to those presented in our previous publications [31, 32]. The transfer RNA genes (tRNAs) are marked with three-letter amino acid abbreviations they transfer. The non-coding region (NCR) located between $tRNA^{Gly}$ or $tRNA^{Glu}$ and *cox*3, divided into two subregions, which consists of long (LRUs) and short tandem repeats (SRUs) (for more information, see text).

The mtDNA of *Eca. malayanum* was also similar to the *Eca. revolutum* MSD15 strain of Thailand (17,030 bp) but largely different in length from the other echinostome congeners, such as *Hypoderaeum conoideum* (China, 14,180 bp, KM111525), *Eca. caproni* (Egypt, 14,150 bp, AP017706), two strains of *Eca. miyagawai* from China (14,410 bp, MH393928 and 14,468 bp, MN116740), *Eca. revolutum* (strain GD, China, MN116706) and *Echinostoma* sp. (JM2019, China, 15,283 bp, MH212284), and much shorter than *Eca. paraensei* (20,298 bp, KT008005) sequenced to date (**Tables 3.1** and **3.2**). The linearized map of the *Eca. malayanum*

mtDNA is 5'-*cox*3-H-*cob*-*nad*4L-*nad*4-QFM-*atp*6-*nad*2-VAD-*nad*1-NPIK-*nad*3-S_1W-*cox*1-T-*rrn*L-C-*rrn*S-*cox*2-*nad*6-$YL_1S_2L_2$R-*nad*5-G-NCR[LRU1−5.5#]-E-[SRU1−7.5#]-3'.

The NCR of *Eca. malayanum* mtDNA was identified by the boundary of $tRNA^{Gly}$ (*trn*G) and the *cox*3 gene, and was shown to be relatively long (3,622 bp). This NCR was divided into two subregions (1,944 bp and 1,678 bp, respectively), separated by $tRNA^{Glu}$ (*trn*E). The first subregion contained 5.5 LRUs (5 perfect LRUs of 336 bp/each and a half one of 118 bp) and 7.5 SRUs (7 perfect SRUs of 207 bp/each and a half one of 103 bp) (**Table 3.2**).

Table 3.2. Locations of genes and other features in the mitochondrial genome of *Echinostoma malayanum* (17,175 bp) and *Artyfechinostomum sufrartyfex* (14,567 bp)

Gene/ Region	Position (5' > 3')	Characteristics [bp/aa(start/stop)] and regions	Int. seq. (bp)	tRNA anti-codon	Position (5' > 3')	Characteristics [bp/aa(start/stop)] and regions	Int. seq. (bp)
	***Echinostoma malayanum* (OK509083)**				***Artyfechinostomum sufrartyfex* (KY548763)**		
*cox*3	1–645	645/214/(ATG/TAG)	+7		1–645	645/214/(ATG/TAA)	+7
$tRNA^{His}$ (*trn*H)	653–720	68	+3	GTG	653–720	68	+3
cob	724–1833	1110/369/(ATG/TAG)	+6		724–1833	1110/369/(ATG/TAG)	+6
*nad*4L	1840–2112	273/90/(ATG/TAG)	–40		1840–2112	273/90/(ATG/TAG)	–40
*nad*4	2073–3356	1284/427/(ATG/TAG)	+12		2073–3356	1284/427/(ATG/TAG)	+12
$tRNA^{Gln}$ (*trn*Q)	3369–3437	69	+33	TTG	3369–3437	69	+36
$tRNA^{Phe}$ (*trn*F)	3471–3535	65	+26	GAA	3474–3535	65	+26
$tRNA^{Met}$ (*trn*M)	3562–3627	66	+4	CAT	3562–3627	66	+4
*atp*6	3632–4150	519/172/(ATG/TAG)	+33		3634–4152	519/172/(ATG/TAG)P	+31
*nad*2	4184–5056	873/290/(ATG/TAG)	+5		4184–5056	873/290/(ATG/TAG)	+6
$tRNA^{Val}$ (*trn*V)	5062–5124	63	+10	TAC	5062–5124	63	+10
$tRNA^{Ala}$ (*trn*A)	5135–5200	66	+14	TGC	5135–5200	66	+14
$tRNA^{Asp}$ (*trn*D)	5215–5279	65	0	GTC	5215–5279	65	0
*nad*1	5280–6182	909/302/(GTG/TAG)	+6		5280–6182	909/302/(ATG/TAG)	+6
$tRNA^{Asn}$ (*trn*N)	6189–6258	70	+5	GTT	6189–6258	70	+5
$tRNA^{Pro}$ (*trn*P)	6264–6329	66	+5	TGG	6264–6329	66	+5
$ttRNA^{Ile}$ (*trn*I)	6235–6396	62	+5	GAT	6235–6396	62	+5
$tRNA^{Lys}$ (*trn*K)	6402–6472	71	+1	CTT	6402–6472	71	+1
*nad*3	6474–6830	357/118/(ATG/TAA)	+6		6474–6830	357/118/(ATG/TAA)	+6
$tRNA^{Ser1(AGN)*}$ (*trn*S_1)	6837–6896	60	+23	GCT	6837–6896	60	+26
$tRNA^{Trp}$ (*trn*W)	6920–6984	65	+3	TCA	6923–6986	65	+3
*cox*1	6988–8526	1539/512/(GTG/TAG)	+35		6991–8529	1539/512/(ATG/TAG)	+35
$tRNA^{Thr}$ (*trn*T)	8562–8626	65	0	TGT	8565–8629	65	0
*rrn*L (16S)	8627–9607	981	0		8630–9610	986	0
$tRNA^{Cys}$ (*trn*C)	9608–9676	69	0	GCA	9611–9679	69	0
*rrn*S (12S)	9677–10420	744	0		9680–10422	744PP	0
*cox*2	10421–11029	609/202/(ATG/TAG)	+9		10423–11031	609/202/(ATG/TAG)	+9
*nad*6	11039–11491	453/150/(GTG/TAG)	+1		11041–11493	453/150/(ATG/TAG)	+1
$tRNA^{Tyr}$ (*trn*Y)	11493–11559	67	+11	GTA	11495–11561	67	+11
$tRNA^{Leu1(CUN)}$ (*trn*L_1)	11560–11626	67	–3	TAG	11562–11628	67	+1
$tRNA^{Ser2(UCN)*}$ (*trn*S_2)	11624–11688	65	+36	TGA	11630–11689	60	+36
$tRNA^{Leu2(UUR)}$ (*trn*L_2)	11715–11777	63	–1	TAA	11716–11778	63	–1
$tRNA^{Arg}$ (*trn*R)	11777–11844	68	–2	TCG	11778–11845	68	–2
*nad*5	11843–13408	1566/521/(GTG/TAG)	+30		11844–13409	1566/521/(ATG/TAG)	+17
$tRNA^{Gly}$ (*trn*G)	13426–13489	64	+82	TCC	13427–13490	64	+7
LRU region	13572–15369						
LRU1	13572–13907	336	0				
LRU2	13908–14243	336	0				
LRU3	14244–14579	336	0				
LRU4	14580–14915	336	0				
LRU5	14916–15251	336	0				
LRU5.5#	15252–15369	118	+64				
$tRNA^{Glu}$	15434–15497	64	+100	TTC	13498-13563	64	+78
SRU Region	15598–17046						
SRU1	15598–15804	207	0		13642–13785	144	0
SRU2	15805–16011	207	0		13786–13929	144	+191
SRU3	16012–16218	207	0		14121–14232	112 (RU2.8#)	0
SRU4	16219–16425	207	0				
SRU5	16426–16632	207	0				
SRU6	16633–16839	207	0				
SRU7	16840–17046	207	0				

SRU7.5#	17047–17149	103	0				
Uni. seq	17150–17175	26	0		14233–14567	335	0

bp: base pair; aa: amino acid; start: start codon; stop: stop codon; Int. seq.: intergenic sequence (+. number of nucleotides before start of following gene; –, number of nucleotides overlapping with following gene); Uni. seq: sequence between SRU7.5# and *cox*3; *Asterisk: tRNAs lacking DHU-arm. RU#: imperfect repeat unit.

3.1.3 *Echinostoma miyagawai* and *Hypoderaeum conoideum*

The mtDNA sequences of these two species were generated by NGS using a targeted long-read sequencing approach.

***Echinostoma miyagawai*:** The whole mitogenome of *Eca. miyagawai* was obtained by NGS. The mitogenome was assembled using two major overlapping contigs of 2,954 bp and 7,029 bp, which formed an almost complete circular genome of 19,083 bp. By comparative alignment with the available *Echinostoma* mtDNA sequences, several small gaps within the coding region were found in the first contig and subsequently filled by conventional Sanger sequencing. The second contig perfectly marched the expected genes from the 3' end of *nad*5 and the 5' end of *cyt*B. This fragment also contained the $tRNA^{Gly}$, $tRNA^{Glu}$, $tRNA^{His}$, and *cox*3 genes, and entirely bridged the tandem repeat non-coding region (5,935 bp). The whole mitogenome of *Echinostoma miyagawai*, strain RED11 from Thailand, is 19,417 bp in length (GenBank accession no. OP326312) (**Table 3.3**; **Fig. 3.1C**), which is the longest complete mtDNA to be sequenced so far among the echinosomid species.

***Hypoderaeum conoideum*:** The whole mitogenome of *Hypoderaeum conoideum* was obtained by NGS. The *denovo* assembly yielded a final contig of a single sequence, which covered the complete mtDNA sequence of 18,011 bp in length and is the whole mitogenome of the *Hypoderaeum conoideum* strain RED42 from Thailand (GenBank accession no. PP110501). The NCR is 4,475 bp long, which is located between the $tRNA^{Glu}$ and *cox*3 genes (**Table 3.3**; **Fig. 3.1D**).

Table 3.3 Locations of genes and other features in the mitochondrial genomes of *Echinostoma miyagawai* (Emiya-RED11-TH, 19,417 bp, GenBank: OP326312); and *Hypoderaeum conoideum* (Hcono-RED42-TH, 18,011 bp, GenBank: PP110501)

Gene/ Region	Position (5' > 3')	Characteristics [bp/aa(start/stop)] and regions	Int. seq. (bp)	tRNA anti-codon	Position (5' > 3')	Characteristics [bp/aa(start/stop)] and regions	Int. seq. (bp)
	Echinostoma miyagawai (Emiya-RED11-TH, Thailand, OP326312)				***Hypoderaeum conoideum*** (Hcono-RED42-TH, Thailand, PP110501)		
*cox*3	1–645	645/214/(ATG/TAA)	+3		1–645	645/214/(ATG/TAG)	+2
$tRNA^{His}$ (*trn*H)	649–713	65	+2	GTG	648–718	71	+1
cob	716–1825	1110/369/(ATG/TAG)	0		720–1829	1110/369/(ATG/TAG)	+14
*nad*4L	1826–2098	273/90/(ATG/TAG)	–40		1844–2113	270/89/(ATG/TAG)	–40
*nad*4	2059–3342	1284/427/(ATG/TAG)	+4		2074–3357	1284/427/(GTG/TAA)	+7
$tRNA^{Gln}$ (*trn*Q)	3347–3410	64	+8	TTG	3365–3429	65	+32
$tRNA^{Phe}$ (*trn*F)	3419–3484	66	+33	GAA	3462–3527	66	+12
$tRNA^{Met}$ (*trn*M)	3518–3583	66	+3	CAT	3540–3605	66	+3
*atp*6	3587–4105	519/172/(ATG/TAG)	+7		3609–4127	519/172/(ATG/TAG)	+2
*nad*2	4113–4982	870/289/(ATG/TAG)	+4		4131–4997	867/288/(ATG/TAG)	+5
$tRNA^{Val}$ (*trn*V)	4987–5050	64	+24	TAC	5003–5070	68	+23
$tRNA^{Ala}$ (*trn*A)	5075–5142	68	+4	TGC	5094–5157	64	+12
$tRNA^{Asp}$ (*trn*D)	5147–5212	66	0	GTC	5170–5235	66	0
*nad*1	5213–6115	903/300/(GTG/TAG)	+6		5236–6138	903/300/(GTG/TAG)	+7

tRNA^{Asn} (*trn*N)	6122–6188	67	+4	GTT	6146–6215	70	+3
tRNA^{Pro} (*trn*P)	6193–6261	69	+1	TGG	6219–6285	67	+1
ttRNA^{Ile} (*trn*I)	6263–6324	62	+9	GAT	6287–6350	64	+6
tRNA^{Lys} (*trn*K)	6334–6402	69	+4	CTT	6357–6428	72	0
*nad*3	6407–6763	357/118/(ATG/TAG)	+3		6429–6785	357/118/(ATG/TAA)	+3
tRNA$^{Ser1(AGN)*}$ (*trn*S$_1$)	6767–6826	60	+4	GCT	6789–6848	60	+11
tRNA^{Trp} (*trn*W)	6831–6896	66	+3	TCA	6860–6926	68	+3
*cox*1	6900–8438	1539/512/(GTG/TAA)	+35		6931–8469	1539/512/(GTG/TAG)	+29
tRNA^{Thr} (*trn*T)	8474–8543	70	0	TGT	8498–8573	75	0
*rrn*L (16S)	8544–9518	975	0		8574–9551	978	0
tRNA^{Cys} (*trn*C)	9519–9585	67	0	GCA	9552–9620	69	0
*rrn*S (12S)	9586–10335	750	0		9621–10370	750	0
*cox*2	10336–10944	609/202/(ATG/TAG)	+11		10371–10973	603/200/(ATG/TAG)	+31
*nad*6	10956–11408	453/150/(ATG/TAG)	0		11005–11457	453/150/(ATG/TAG)	–2
tRNA^{Tyr} (*trn*Y)	11409–11477	68	0	GTA	11455–11520	64	0
tRNA$^{Leu1(CUN)}$ (*trn*L$_1$)	11478–11542	65	–3	TAG	11521–11586	66	–3
tRNA$^{Ser2(UCN)*}$ (*trn*S$_2$)	11540–11604	65	+27	TGA	11584–11648	65	+17
tRNA$^{Leu2(UUR)}$ (*trn*L$_2$)	11632–11694	63	0	TAA	11666–11728	63	+1
tRNA^{Arg} (*trn*R)	11695–11758	64	0	TCG	11730–11797	68	–2
*nad*5	11759–13324	1566/521/(GTG/TAG)	+19		11796–13361	1566/521/(GTG/TAA)	+34
tRNA^{Gly} (*trn*G)	13344–13409	66	+9	TCC	13396–13460	65	+6
tRNA^{Glu}(*trn*E)	13418–13482	64	+22	TTC	13467–13536	70	+1
NCR region	**13483-19417**	**5935**			**13537–18011**	**4475**	
LRU region	**13505–18392**	**4888**			**13538–16670**	**3133**	
LRU1	13506–13824	319	0		13538–13778	241	0
LRU2	13825–14143	319	0		13779–14019	241	0
LRU3	14144–14462	319	0		14020–14260	241	0
LRU4	14463–14781	319	0		14261–14501	241	0
LRU5	14782–15100	319	0		14502–14742	241	0
LRU6	15101–15419	319	0		14743–14983	241	0
LRU7	15420–15738	319	0		14984–15224	241	0
LRU8	15739–16057	319	0		15225–15466	241	0
LRU9	16058–16376	319	0		15467–15706	241	0
LRU10	16377–16695	319	0		15707–15947	241	0
LRU11	16696–17014	319	0		15948–16188	241	0
LRU12	17015–17333	319	0		16189–16419	241	0
LRU13	17334–17652	319	0		16420–16670	241	0
LRU14	17653–17971	319	0				
LRU15	17972–18290	319	0				
LRU15.3#	18291–18392	102	0				
Junction seq.	18393–18395	3	0		16671–16785	115	0
SRU Region	**18396–19412**	1017			**16786–17939**		
SRU1	18396–18608	213	0		16786–16896	111	0
SRU2	18609–18821	213	0		16897–17007	111	0
SRU3	18822–19034	213	0		17008–17118	111	0
SRU4	19035–19247	213	0		17119–17229	111	0
SRU4.8#	19248–19412	165	0				
Uni. seq.	19413–**19417**	5	0				
SRU5					17230–17340	111	0
SRU6					17341–17451	111	0
SRU7					17452–17562	111	0
SRU8					17563–17673	111	0
SRU9					17674–17784	111	+73
SRU9.7#					17858–17939	82	0
Uni. seq.					17940–**18011**	72	0

Note: bp: base pair; aa: amino acid; start: start codon; stop: stop codon; Int. seq.: intergenic sequence (+. number of nucleotides before start of following gene; –, number of nucleotides overlapping with following gene); Junction seq.: sequence connecting the last LRU and first SRU; Uni. seq.: sequence between last SRU and *cox*3; *asterisk: tRNAs lacking DHU-arm. RU#: imperfect repeat; unit for Emiy-RED11-TH (LRU15.3# and SRU4.8#) and for Hcono-RED42-TH (SRU9.7#, position: 17858–17939).

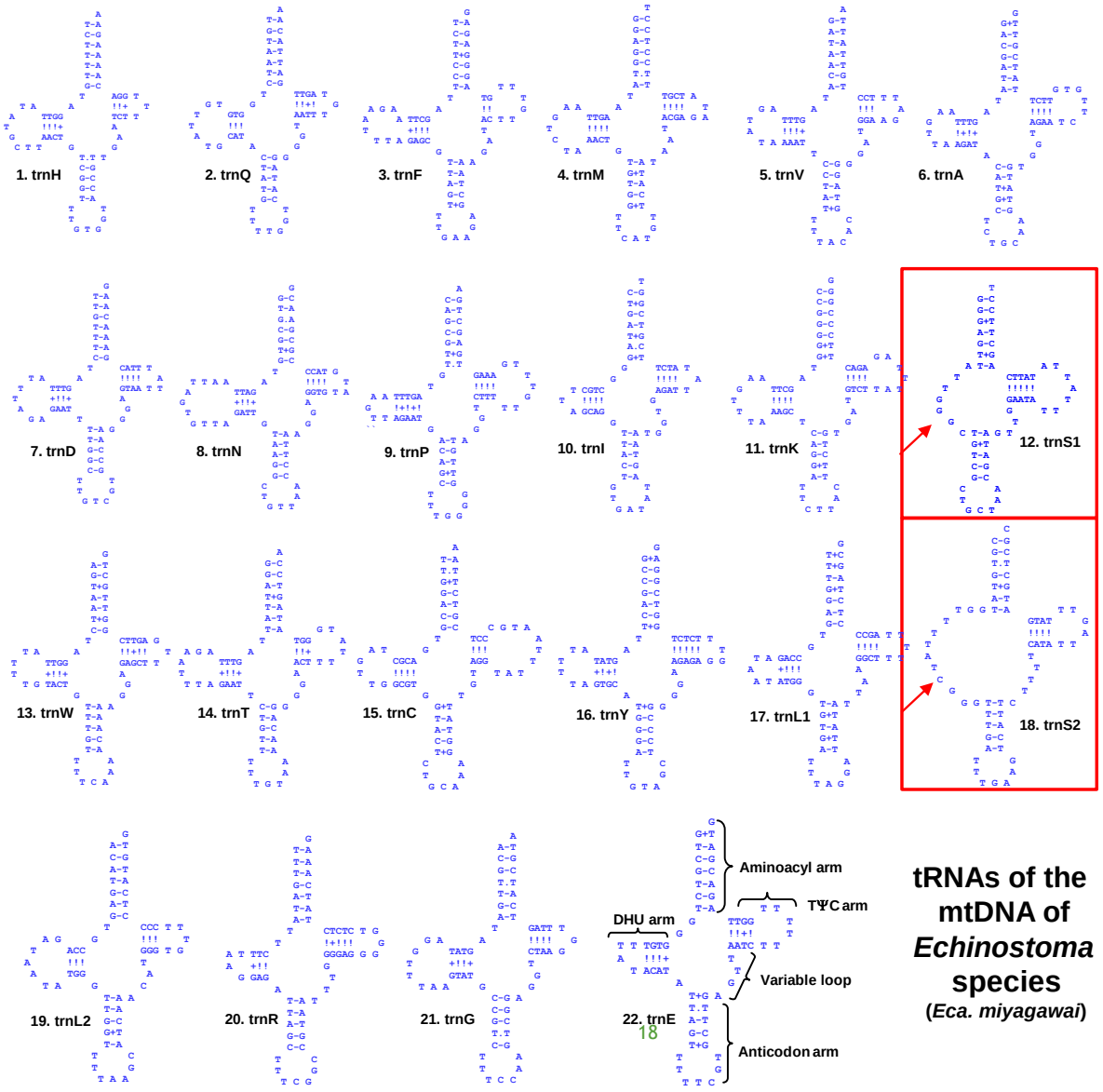


Figure 3.2. Drawings of predicted structure models of 22 transfer RNAs in the mitogenome of *Echinostoma* species (here, a representative, *Eca. miyagawai*, arranged in alphabetical order of the amino acids they specify. Each tRNA (here abbreviated as trn) gene is named according to the one-letter amino acid abbreviation, with the exception of those specifying Serine, S1 and S2 (S1, AGN; and S2, UCN); DHU arms are missing in tRNASer1$^{(AGN)}$ and in tRNASer2$^{(UCN)}$ (squared and shown by an arrow). 1. trnH (Histidine); 2. trnQ (Glutamine); 3. trnF (Phenylalanine); 4. trnM (Methionine); 5. trnV (Valine); 6. trnA (Alanine); 7. trnD (Aspartic acid); 8. trnN (Asparagine); 9. trnP (Proline); 10. trnI (Isoleucine); 11. trnK (Lysine); 12. trnS1$^{(AGN)}$ (Serine); 13. trnW (Tryptophan); 14. trnT (Threonine); 15. trnC (Cystine); 16. trnY (Tyrosine); 17. trnL1$^{(CUN)}$ (Leucine); 18. trnS2$^{(UCN)}$ (Serine); 19. trnL2$^{(UUR)}$ (Leucine); 20. trnR (Arginine); 21. trnG (Glycine); 22. trnE (Glutamic acid); Names of structural components of a tRNA gene are indicated in the trnE (Glutamic acid) structure.

The circular mtDNA molecule for both species (*Eca. miyagawai* and *H. conoideum*) is similar in arrangement and is comprised of 12 PCGs (*cox*1−3, *cob*, *nad*1−6, *nad*4L, *atp*6), two MRGs (16S or *rrn*L and 12S or *rrn*S), 22 tRNAs or *trn*, and an NCR that possesses long and short tandem repeats. The linearized mtDNA map of the two species is 5'-*cox*3-H-*cob*-*nad*4L-*nad*4-QFM-*atp*6-*nad*2-VAD-*nad*1-NPIK-*nad*3-S_1W-*cox*1-T-*rrn*L-C-*rrn*S-*cox*2-*nad*6-$YL_1S_2L_2$R-*nad*5-G-E-((NCR[LRU1−15.3#]-[SRU1−4.8#]/*Eca. miyagawai*;] and --…*nad*5-G-

E((NCR [LRU1−13#]-[SRU1−9.7#]/*H. conoideum*))]-3'. The secondary structure drawing for 22 tRNAs for the *Eca. miyagawai* RED11 strain as a representative species is shown in **Fig. 3.2**. The representative genomic display of the fully annotated mitogenome of the *Eca. miyagawai* RED11 strain is presented in **Supplementary Fig. S3.1**.

The *Echinostoma miyagawai* (strain RED11, Thailand) and *Hypoderaeum conoideum* (strain RED42, Thailand) mitogenomes' lengths were much longer than that in the members of the Echinostomatidae, e.g., *Eca. revolutum* (17,030 bp, strain MSD15, Thailand) [31], *A. malayanum* (17,175 bp, strain EMI3, Thailand) [9], *A. sufrartyfex* (14,567 bp, strain Shillong, India; GenBank: KY548763), *Eca. caproni* (14,150 bp, strain SAMEA, Egypt; GenBank: AP017706), *H. conoideum* (14,180 bp, strain Hubei, China) [92], *Echinoparyphium aconiatum* (14,865 bp, strain Chany, Russia) [142], and several echinostomatid species to be reported to date. Similarly, in comparison between other geographical isolates of *Eca. miyagawai*, the two Chinese *Eca. miyagawai* strains (Hunan and HLJ) have mtDNA much shorter (14,468 bp and 14,410 bp, respectively) and seemed to be truncated by conventional sequencing [38, 91] than the current studied Thai strain.

Non-coding regions of the mtDNAs of echinostomes:

***Echinostoma miyagawai*:** Using the long-reading PACBIO system, the complete NCR in the mitogenome of the *Eca. miyagawai* RED11 strain was successfully obtained, which is 5,935 bp and contains two types of tandem repeat units referred to as the LRUs, and short repeat unit, the SRUs. This lengthy NCR in this *Eca. miyagawai* isolate was flanked by $tRNA^{Glu}$ (*trn*E) and the *cox*3 gene. The relatively long NCR (near 6.0 kb) was divided into two subregions: the first subregion contained 15.3 identical LRUs of 319 bp/each and a partial one of 102 bp, and the second contained 4.8 SRUs of 213 bp/each and a partial one of 165 bp). Only three nucleotides (TAA, position: 18393–18395) connected the 3' end of the last LRU15.3 and the 5' end of the first SRU1 (**Fig. 3.1; Tables 3.3**). We noted that the identical LRUs were present in the NCR of the Chinese strains as well, but in fewer numbers (2.99 LRUs in the Hunan strain and 2.3 LRUs in the HLJ strain), and no other repeats such as SRUs of the RED11 strain were found in either one (**Table 3.3**). The repetitive features in the NCR of these two Chinese strains were not stated in the original analyses by Li et al. [91] and Fu et al. [38].

***Hypoderaeum conoideum*:** The lengthy NCR (4,475 bp) in the mtDNA of the *H. conoideum* RED42 strain was successfully obtained using the long-reading PACBIO system, and possibly, of the realistic size for the mtDNA non-coding region of this species. The relatively long NCR was divided into two distinct subregions: the first subregion contained 13 identical LRUs (241 bp/each) and the second contained 9.7 SRUs (111 bp/each and a partial one of 82 bp); and there were 115 nucleotide sequence (position: 16671–16785) connecting

the 3' end of the last LRU13 and the 5' end of the first SRU1 (**Fig. 3.3; Table 3.3**). We noted that neither the identical LRUs nor SRUs were present in the NCR of the Chinese *H. conoideum* Hubei strain obtained by Yang et al. [92] (GenBank: KM111525).

3.2 Comparative mitogenomic analyses of the echinostomes in the Echinostomatidae family

3.2.1 Gene identity comparisons among echinostomes

At the nucleotide level, between *Eca. miyagawai* and echinostomatids, the *nad*6 gene showed the highest divergence over 64.68%, with some genes reaching 71.88% to 75.48% between species. A lower divergence rate is seen for all genes between the *Eca. miyagawai* RED11 strain and others in the "revolutum" group (*Eca. caproni* SAMEA strain, *Eca. paraensei*, and *Eca. revolutum* MSD15 strain) than between *Eca. miyagawai* and other echinostomatids. *Echinostoma revolutum* (MSD15 strain) is the most related species and the lowest divergence between this species and *Eca. miyagawai* is seen in *cox*2 (13.07%), *nad*4L (10.29%), PCGs (16.09%), and MRGs (12.39%) (**Table 3.4**). Similarly, the *cyt*B and *cox*1 genes showed the lowest divergence at 13.67% to 27.08% for *cyt*B; 14.14% to 28.61% for *cox*1. The PCGs and the MRGs showed almost the same moderate divergence rate for interspecific variation between *Eca. miyagawai* and other echinostomatids, for example, 16.09–17.36%/PCGs for the "*revolutum*" group and 28.39–33.42%/PCGs for others (**Table 3.4**). The three *Eca. miyagawai* geographical isolates shared intra-specific identity at less than 1% for all genes except for *cox*3 with a divergence of 1.57% and *nad*1 1.12% between the RED11 and HLJ isolates. The *nad*4 also showed a higher divergence between the RED11 and Hunan isolates at 1.18% (**Table 3.4**).

3.2.2 Base composition and skew values in the Echinostomatidae species

The base composition (nucleotide usage) of A, T, G, and C and skewness values of AT and GC content for PCGs, MRGs, and the coding mtDNA region (abbreviated as mtDNA*, from the 5' terminus of *cox*3 to the 3' terminus of *nad*5, including some short intergenic spacer sequences) of 15 strains of 12 species of Echinostomatidae are presented in **Table 3.5**. We prefer to present here the base composition and skew/skewness for four echinostome species of our study in this thesis: *A. malayanum* EMI3 strain, *Eca. miyagawai* RED11 strain, *Eca. revolutum* MSD15 strain, and *H. conoideum* RED42 strain.

The base composition was A (17.08%), T (46.32%), G (26.37%), and C (10.23%) in the mt genome of the *A. malayanum* EMI3 strain, and the A+T content was 63.40% for PCGs, and their skewness values were –0.461 for A+T and 0.391 for G+C, respectively. MRGs showed a similar percentage of overall A+T (61.51%) and G+C (38.49%), but their skewness values were considerably different (–0.201/A+T and 0.337/G+C) due to the biased use of A over T in PCGs than in MRGs. The mtDNA* (13,408 nucleotides) in *A. malayanum* is biased towards the use

of T (44.10%), then G (26.48%), and A (18.78%), and C (10.64%), giving a pattern of nucleotide usage as T > G > A > C and a skew value of –0.403/A+T and 0.427/G+C (**Table 3.5**).

Table 3.4. Nucleotide comparison for divergence rate (%) of individual and concatenated protein-coding (PCGs) and mitoribosomal genes (MRGs) between *Echinostoma miyagawai* (strain RED11, Thailand) and members of the family Echinostomatidae (Platyhelminthes: Echinostomata)

	Species	***Echinostoma miyagawai* (Emiya-RED11-TH)**															
		Individual and protein-coding genes (PCGs)												**Mitoribosomal genes (MRGs)**			
		*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	Amala-EMI3-TH	35.24	23.74	25.24	25.68	24.42	30.50	44.25	30.35	27.51	48.95	35.19	66.77	29.89	28.47	38.92	30.66
2	Asufr-Shillong-IN	34.80	23.96	25.79	25.71	24.60	31.29	45.31	29.78	26.77	48.24	35.05	67.16	30.01	27.99	38.92	30.34
3	Eacon-Chany-RU	38.61	28.61	37.09	24.65	27.08	30.64	47.14	36.24	29.32	48.48	36.93	**75.48**	32.49	28.99	35.25	30.03
4	Ecapr-SAMEA-EG	19.27	15.10	15.90	15.01	14.85	15.79	20.95	16.40	19.56	22.28	20.54	30.65	17.36	13.66	11.66	12.98
5	EcaSP-JM2019-CN	38.73	23.16	21.70	23.45	24.49	26.77	40.16	34.16	30.79	41.07	33.66	65.30	32.95	27.31	30.52	27.59
6	EchCA2021-PE4-US	48.30	26.93	46.04	28.99	25.81	30.87	45.75	30.24	32.36	44.97	39.36	**71.88**	28.39	29.09	36.96	30.11
7	EchMSB-A19-US	42.91	27.29	48.82	27.46	24.90	32.05	44.53	36.63	27.31	46.61	40.36	69.19	33.20	27.00	38.40	29.16
8	Emiya-HLJ-CN	00.78	00.39	00.33	**01.57**	00.45	**01.12**	00.81	00.85	00.73	00.78	00.90	00.31	**00.69**	00.31	00.13	00.23
9	Emiya-Hunan-CN	00.77	00.26	00.33	00.47	00.36	00.89	00.70	00.00	00.37	**01.18**	00.58	00.62	**00.58**	00.21	00.27	00.29
10	Epara	24.27	15.71	16.87	13.39	14.31	17.39	19.51	16.88	20.82	23.53	16.97	34.18	17.29	11.22	14.02	12.28
11	EcaSP-GD-CN	39.55	23.61	22.00	22.97	24.24	26.85	40.10	34.60	29.98	41.76	34.12	64.68	28.57	27.16	30.59	27.49
12	Erevo-MSD15-TH	19.29	14.14	**13.07**	13.75	13.67	15.10	18.69	18.60	**10.29**	21.04	18.68	33.36	**16.09**	12.83	11.84	**12.39**
13	Hcono-Hubei-CN	46.16	27.99	43.06	28.96	26.45	32.55	49.70	32.04	25.29	43.79	41.11	**73.96**	33.42	31.21	39.93	32.44
14	Hcono-RED42-TH	47.24	28.40	43.98	29.15	26.35	32.37	50.34	33.18	26.39	42.78	41.34	**75.24**	33.59	31.21	43.20	33.30

Note: Amala-EMI3-TH: *Artyfechinostomum malayanum* isolate EMI3, Thailand (OK509083); Asufr-Shillong-IN: *A. sufrartyfex* isolate Shillong, India (KY548763); Eacon-Chany-RU: *Echinoparyphium aconiatum* isolate Chany, Russia (ON644993); Ecapr-SAMEA-EG: *Eca. caproni* isolate SAMEA, Egypt (AP017706); EchCA2021-PE4-US: Echinostomatidae sp. CA-2021 isolate PE4, United States (MK264774); Ech-JM2019-CN: *Echinostoma* sp. isolate JM-2019, China (MH212284); EchMSB-A19-US: Echinostomatidae sp. MSB para 30070, isolate A_19, United States (MN822299); Emiya-HLJ-CN: *Eca. miyagawai* isolate Heilongjiang, China (MH393928); Emiya-Hunan-CN: *Echinostoma miyagawai* isolate Hunan, China (MN116740); Epara: *Eca. paraensei* (KT008005); Erevo-GD-CN: *Eca. revolutum* isolate Guangdong, China (MN116706); Erevo-MSD15-TH: *Eca. revolutum* isolate MSD15, Thailand (MN496162); Hcono-Hubei-CN: *Hypoderaeum conoideum* isolate Hubei, China (KM111525). The inter-specific divergence (%) between *Eca. miyagawai* isolate RED11 (Thailand) and other echinostomatids for the most conserved *cyt*B and *cox*2, and for the most divergent *nad*6 genes are highlighted. The highest value in *cox*3, *nad*1, and *nad*4, and the average value in PCGs for the intra-specific divergence within the *Eca. miyagawai* strains, are background shaded and boxed. The highest inter-specific divergence between *Eca. miyagawai* and other echinostomatids in *nad*6 are boxed. *Echinostoma revolutum* (MSD15 strain) is the most related species (background shaded) and the lowest divergence between this species and *Eca. miyagawai* in *cox*2, *nad*4L, PCGs, and MRGs is bolded and underlined.

Across all three isolates of *Eca. miyagawai* there is almost an equal use A (18.05–18.20%) and T (47.50–47.65%), G (24.07–24.22%) and C (10.08–10.21%) with A+T = 65.65–65.85% and G+C=34.15–34.35% for 12 PCGs (10,128 bp); and 19.70–20.18% A and 45.24–45.43% T (A+T = 60.23%), 23.82–24.26% G and 10.57–10.66% C (G+C = 34.39–34.92%) for its coding mtDNA* region (13,320–13,324 bp). This nucleotide usage of *Eca. miyagawai* does not vary considerably across the *Echinostoma* genus but is different in other echinostomatids with lower A+T in *Artyfechinostomum malayanum*, *A. sufrartyfex*, and *H. conoideum*. Except for *Echinoparyphium aconiatum*, Echinostomatidae sp. MSB para 30070, and *Eca. revolutum* isolate MSD15, which had the AT-skew of low negative (–0.414 to –0.432/PCGs and –0.357

to –0.379/mtDNA*), the other echinostomatids exhibit highly negative values (–0.477 to –0.483/PCGs and –0.385 to –0.420/mtDNA*), indicating that T was more frequently used than A. The data indicated that the pattern of base usage for all PCGs, MRGs, and mtDNA*s in all 14 strains/species is T > A > G > C, giving the AT-skew negative and the GC-skew positive (**Table 3.5**).

The *H. conoideum* RED42 strain used nucleotides with 16.81% A, 45.08% T (A+T = 61.89%), 26.95% G, and 11.16% C (G+C = 38.11%) for 12 PCGs; and 18.62% A, 42.74% T (A+T = 61.36%), 26.99% G, and 11.65% C (G+C = 38.64%) for its coding mtDNA* (13,361 bp). MRGs showed 59.49% A+T use (AT-skew = –0.154/MRGs) and 40.50% G+C use (GC-skew = 0.309/MRGs). The AT-skew value for the *H. conoideum* mitogenome was highly above negative (–0.457/PCGs and –0.403/mtDNA*), indicating more frequently used nucleotides of T than A. The GC-skew was also highly positive (0.441/PCGs and 0.393/mtDNA*), indicating greater numbers of C than G to be used (**Table 3.5**).

Table 3.5 Base composition and skew/skewness value for AT and GC of the protein-coding genes (PCGs), mito-ribosomal genes (MRGs), and the coding region (abbreviated as mtDNA*) of the mitogenomes of *Echinostoma miyagawai* and other echinostomatid members of the family Echinostomatidae

	Species/Strains	Sequence	Length (nt)	A (%)	T (%)	G (%)	C (%)	A+T (%)	AT-skew	G+C (%)	GC-skew
1	***Artyfechinostomum malayanum*** (Amala-EMI3-TH) (OK509083)	PCGs	10131	17.08	46.32	26.37	10.23	63.40	–0.461	36.60	0.441
		MRGs	1725	24.58	36.93	25.74	12.75	61.51	–0.201	38.49	0.337
		mtDNA*	13408	18.78	44.10	26.48	10.64	62.88	–0.403	37.12	0.427
2	***Artyfechinostomum sufrartyfex*** (Asufr-Shillong-IN) (KY548763)	PCGs	10131	16.99	46.21	26.53	10.27	63.20	–0.462	36.80	0.442
		MRGs	1728	24.71	37.09	25.58	12.62	61.80	–0.200	38.20	0.339
		mtDNA*	13409	18.73	44.03	26.57	10.66	62.76	–0.403	37.24	0.427
3	***Echinoparyphium aconiatum*** (Eacon-Chany-RU) (ON644993)	PCGs	10113	19.05	46.02	24.39	10.53	65.07	–0.414	34.93	0.397
		MRGs	1730	26.18	36.36	24.51	12.95	62.54	–0.163	37.46	0.309
		mtDNA*	13377	20.74	43.72	24.56	10.99	64.46	–0.357	35.54	0.382
4	***Echinostoma caproni*** (Ecapr-SAMEA-EG) (AP017706)	PCGs	10128	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
		mtDNA*	13293	19.05	45.45	24.81	10.69	64.50	–0.409	35.50	0.398
5	***Echinostoma* sp.** (EcaSP-JM2019-CN) (MH212284)	PCGs	10122	16.47	46.46	26.66	10.40	62.93	–0.477	37.07	0.439
		MRGs	1726	24.51	35.17	26.94	13.38	59.68	–0.179	40.32	0.336
		mtDNA*	13257	18.25	44.11	26.70	10.95	62.36	–0.395	37.89	0.416
6	**Echinostomatidae sp. CA-2021** (EchCA2021-PE4-US) (MK264774)	PCGs	10143	17.50	45.81	26.19	10.50	63.31	–0.447	36.69	0.428
		MRGs	1727	25.25	34.80	26.40	13.55	60.05	–0.159	39.95	0.322
		mtDNA*	13319	19.18	43.37	26.32	11.14	62.55	–0.387	37.45	0.405
7	**Echinostomatidae sp. MSB para 30070** (EchMSB-A19-US) (MN822299)	PCGs	10128	18.32	45.71	25.25	10.72	64.03	–0.428	35.97	0.404
		MRGs	1732	26.21	35.05	25.58	13.16	61.26	–0.144	38.74	0.321
		mtDNA*	13346	20.10	43.29	25.39	11.22	63.39	–0.366	36.61	0.387
8	***Echinostoma miyagawai*** (Emiya-HLJ-CN) (MH393928)	PCGs	10128	18.17	47.50	24.12	10.21	65.67	–0.447	34.33	0.405
		MRGs	1763	25.98	37.61	23.60	12.82	63.59	–0.183	36.41	0.296
		mtDNA*	13321	19.85	45.24	24.22	10.68	65.09	–0.390	34.91	0.388
9	***Echinostoma miyagawai*** (Emiya-Hunan-CN) (MN116740)	PCGs	10128	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
		mtDNA*	13320	20.18	45.43	23.82	10.57	65.61	–0.385	34.39	0.385
10	***Echinostoma miyagawai*** (Emiya-RED11-TH) (OP326312)	PCGs	10128	18.05	47.60	24.22	10.13	65.65	–0.450	34.35	0.410
		MRGs	1725	25.68	37.80	23.59	12.93	63.48	–0.191	36.52	0.292
		mtDNA*	13324	19.70	45.38	24.26	10.66	65.08	–0.395	34.92	0.389
11	***Echinostoma paraensei*** (Epara) (KT008005)	PCGs	10128	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
		mtDNA*	13319	19.81	45.42	24.12	10.66	65.23	–0.393	34.77	0.387

12	***Echinostoma* sp.** (*revolutum*?) (EcaSP-GD-CN) (MN116706)	PCGs	10113	16.24	46.60	26.78	10.39	62.84	–0.483	37.17	0.441
		MRGs	1754	24.57	35.12	26.91	13.40	59.69	–0.177	40.31	0.335
		mtDNA*	13282	18.06	44.19	26.78	10.97	62.25	–0.420	37.75	0.419
13	***Echinostoma revolutum*** (Erevo-MSD15-TH) (MN496162)	PCGs	10134	18.81	47.40	23.50	10.29	66.21	–0.432	33.79	0.391
		MRGs	1733	25.74	36.99	23.77	13.50	62.73	–0.179	37.27	0.276
		mtDNA*	13326	20.35	45.21	23.60	10.84	65.56	–0.379	34.44	0.371
14	***Hypoderaeum conoideum*** (Hcono-Hubei-CN) (KM111525)	PCGs	10116	16.84	45.25	26.96	10.95	62.09	–0.458	37.91	0.422
		MRGs	1727	25.13	34.57	26.64	13.67	59.70	–0.158	40.30	0.322
		mtDNA*	13361	18.64	42.92	27.00	11.44	61.56	–0.394	38.44	0.405
15	***Hypoderaeum conoideum*** (Hcono-RED42-TH) (PP110501)	PCGs	10116	16.81	45.08	26.95	11.16	61.89	–0.457	38.11	0.414
		MRGs	1728	25.17	34.32	26.50	14.00	59.49	–0.154	40.50	0.309
		mtDNA*	13361	18.62	42.74	26.99	11.65	61.36	–0.393	38.64	0.397

Note: Information for strains and/or species is given in Supplementary **Table S2.4**; their strain abbreviations and GenBank accession numbers are given in brackets after the taxonomic name; PCGs: protein-coding genes; MRGs: mitoribosomal genes; mtDNA*: mitochondrial coding nucleotide sequence (from the 5' terminus of *cox*3 to the 3' terminus of *nad*5). *Echinostoma* sp. (*revolutum*?): is a species that was reported as *Echinostoma revolutum* in Ran et al. [93], but due to the lack of strong *Echinostoma* generic evidence it was listed in the "*cryptic*" genus "*incertae sedis*" within the Echinostomatidae [9]. Four species of this study (nos 1, 10, 13, and 15) are highlighted; and *Echinostoma* strains/species of the "*revolutum*" group (no 4, 8, 9, 10, 11, and 13) are background shaded. The numbers of discussions are highlighted for notion.

The nucleotide usage does not vary much in the mtDNAs within the members of the "*revolutum*" group (*Eca. miyagawai*, *Eca. caproni*, *Eca. paraensei*, and *Eca. revolutum*) and *Echinoparyphium aconiatum* (65.07–66.21% for A+T and 33.79–34.93% for G+C)**,** but is lower for A+T and higher for G+C in *A. malayanum*, *A. sufrartyfex*, *H. conoideum*, and other four *Echinostoma* spp. and Echinostomatidae spp. echinostomatids. Except for three (*Echinoparyphium aconiatum*, Echinostomatidae sp. MSB para 30070, and *Eca. revolutum* isolate MSD15), which had the AT-skew of low negative (–0.414 to –0.432/PCGs and –0.357 to –0.379/mtDNA*), the other echinostomatids exhibit highly negative values (–0.477 to –0.483/PCGs and –0.385 to –0.420/mtDNA*), indicating that T was more frequently used than A. In overall, the data demonstrated that the pattern of base utilization for all PCGs, MRGs, and mtDNA*s in all 15 echinostomid strains/species is T > G > A > C, with the AT-skew negative and the GC-skew positive in all analyses (**Table 3.5**).

3.2.3 Codon usage in the protein-coding genes of the Echinostomatidae species

Table 3.6 indicates codon usage for 12 protein-coding genes in the mitogenomes of four strains of four species (*Eca. revolutum*; *Eca. /A. malayanum*; *Eca. miyagawai*, and *H. conoideum*) in this study. The most frequently used codons were TTT (for Phenylalanine), TTG (for Leucine), and GTT (for Valine), while the least frequently used codon was CGC (for Arginine), with only 1–2 codons being used.

Broadly, **Supplementary Table S3.1** presents the codon usage for 12 PCGs in the mitogenomes of 15 strains from 12 species of the Echinostomatidae family. There were 10,113 nucleotides (*Echinoparyphium aconiatum*, Chany strain, Russia) to 11,149 nucleotides (*Echinostoma* sp., GD strain, China) used for 3,371 to 3,383 codons in echinostomes. The most frequently used codons were TTT for Phenylalanine, TTG for Leucine, and GTT for Valine, while the least frequently used codon was CGC (for Arginine), with only 1–2 codons (0.03-

0.06%) used in all echinostomes except *Eca. paraensei*, which had five AAC/Arginine codons. *Ecchinostoma miyagawai* utilized TTT/Phenylalanine (362–363 codons/10.72–10.75%) and GTT/Valine (241–242 codons/7.14–7.17%), but *H. conoideum* used the most TTG/Leucine (285–287 codons/8.45–8.51%) (**Supplementary Table S3.1**).

Table 3.6. Codon usage for 12 protein-coding genes in the mitogenomes of four strains of four species (*Eca. revolutum*; *Eca./A. malayanum*; *Eca. miyagawai*, and *H. conoideum*) of the family Echinostomatidae

Amino acid*	Codon	Erevo (MSD15-TH) (MN496162)		Amala (EMI3-TH) (OK509083)		Emiya (Red11-TH) (OP326312)		Hcono (RED42-TH) (PP110501)	
		No	%	No	%	No	%	No	%
Ala	GCG	11	0.33	22	0.65	14	0.42	33	0.98
	GCA	20	0.59	15	0.44	20	0.59	16	0.47
	GCT	71	2.11	89	2.64	66	1.96	66	1.96
	GCC	9	0.27	9	0.27	13	0.39	15	0.45
Cys	TGT	96	2.84	89	2.64	99	2.93	92	2.73
	TGC	10	0.30	13	0.39	7	0.21	16	0.47
Asp	GAT	67	1.98	58	1.72	70	2.07	64	1.90
	GAC	7	0.21	7	0.21	5	0.15	4	0.12
Glu	GAG	51	1.51	56	1.66	53	1.57	60	1.78
	GAA	26	0.77	25	0.74	19	0.56	16	0.47
Phe	TTT	346	10.24	330	9.77	363	10.75	302	8.96
	TTC	27	0.80	21	0.62	20	0.59	40	1.19
Gly	GGG	74	2.19	115	3.41	74	2.19	96	2.85
	GGA	28	0.83	35	1.04	27	0.80	41	1.22
	GGT	175	5.18	126	3.73	163	4.83	135	4.00
	GGC	7	0.21	9	0.27	21	0.62	23	0.68
His	CAT	45	1.33	45	1.33	42	1.24	43	1.28
	CAC	9	0.27	6	0.18	13	0.39	9	0.27
Ile	ATA	92	2.72	57	1.69	85	2.52	58	1.72
	ATT	122	3.61	140	4.15	120	3.56	126	3.74
	ATC	11	0.33	16	0.47	12	0.36	23	0.68
Lys	AAG	48	1.42	48	1.42	48	1.42	47	1.39
Leu	TTG	204	6.04	267	7.91	240	7.11	287	8.51
	TTA	214	6.34	157	4.65	189	5.60	144	4.27
	CTG	15	0.44	29	0.86	20	0.59	40	1.19
	CTA	20	0.59	15	0.44	16	0.47	16	0.47
	CTT	85	2.52	65	1.93	70	2.07	63	1.87
	CTC	9	0.27	7	0.21	7	0.21	4	0.12
Met	ATG	108	3.20	110	3.26	104	3.08	110	3.26
Asn	AAA	29	0.86	30	0.89	30	0.89	19	0.56
	AAT	45	1.33	44	1.30	45	1.33	51	1.51
	AAC	6	0.18	5	0.15	8	0.24	7	0.21
Pro	CCG	10	0.30	14	0.42	9	0.27	21	0.62
	CCA	20	0.59	7	0.21	18	0.53	10	0.30
	CCT	39	1.16	56	1.66	49	1.45	42	1.25
	CCC	25	0.74	15	0.44	21	0.62	23	0.68
Gln	CAG	19	0.56	21	0.62	19	0.56	18	0.53
	CAA	8	0.24	7	0.21	7	0.21	10	0.30
Arg	CGG	8	0.24	16	0.47	13	0.39	15	0.45
	CGA	8	0.24	4	0.12	6	0.18	3	0.09
	CGT	45	1.33	41	1.21	43	1.27	42	1.25
	CGC	1	0.03	2	0.06	1	0.03	2	0.06
Ser	AGG	31	0.92	52	1.54	40	1.19	62	1.84
	AGA	29	0.86	19	0.56	27	0.80	22	0.65
	AGT	97	2.87	84	2.49	88	2.61	70	2.08
	AGC	8	0.24	11	0.33	5	0.15	11	0.33
	TCG	10	0.30	23	0.68	15	0.44	30	0.89
	TCA	31	0.92	22	0.65	24	0.71	19	0.56
	TCT	141	4.17	135	4.0	143	4.24	121	3.59
	TCC	14	0.41	14	0.42	11	0.33	24	0.71
Thr	ACG	11	0.33	20	0.59	19	0.56	24	0.71
	ACA	16	0.47	13	0.39	16	0.47	15	0.45
	ACT	50	1.48	56	1.66	39	1.16	47	1.39
	ACC	12	0.36	4	0.12	18	0.53	9	0.27
Val	GTG	79	2.34	105	3.11	69	2.04	108	3.20
	GTA	53	1.57	49	1.45	56	1.66	55	1.63
	GTT	219	6.48	221	6.54	241	7.14	196	5.81
	GTC	21	0.62	16	0.47	10	0.30	25	0.74
Trp	TGG	54	1.60	77	2.28	63	1.87	77	2.28
	TGA	52	1.54	36	1.07	45	1.33	29	0.86
Tyr	TAT	157	4.65	148	4.38	161	4.77	146	4.33
	TAC	11	0.33	17	0.50	5	0.15	18	0.53
stop	TAG	7	0.21	11	0.33	10	0.30	9	0.27
	TAA	5	0.15	1	0.03	2	0.06	3	0.09

Note: Erevo: *Echinostoma revolutum*; Amala: *Artyfechinostomum malayanum*; Emiya: *Echinostoma miyagawai*; Hcono: *Hypoderaeum conoideum*).

Interestingly, *Echinostoma* sp. strain JM-2019 and *Echinostoma* sp. strain GD still retained the two least-used CAA/Gln (Glutamine) codons (0.06%), and in rare cases, the latter species employed no codon for TGA/Tryptophan. In all 15 echinostomid strains, TAG (7–12 codons) was used to terminate 12 PCGs rather than TAA (0–5), indicating a clear bias (**Supplementary Table S3.1**).

3.2.4 Pairwise genetic distance among the Echinostomatidae species

The pairwise genetic distances (p-distances), which indicate the pairwise nucleotide differences (%), were estimated using 12 PCGs from 15 strains of 12 species in the Echinostomatidae family (**Table 3.7**). *Echinostoma* sp. strain JM-2019 (China; MH212284) had the shortest genetic distance (0.43%) to *Echinostoma* sp. strain GD (China; MN116706), implying that these two species may have an intraspecific genetic distance or are close variants within the same *Echinostoma* species. A very low pairwise nucleotide difference (0.40–0.52%) was also observed among three *Eca. miyagawai* strains. The genetic difference between the *H. conoideum* Chinese and Thai strains was extremely low (0.62%). Interestingly, between *A.*

malayanum and *A. sufrartyfex*, a low rate of genetic distance (0.89%) was observed, showing an intraspecific level between these two species, or perhaps they are synonymous taxa.

Table 3.7 Pairwise genetic distance (%) among 15 strains of 12 species in the family Echinostomatidae estimated based on the analysis of concatenated nucleotide sequences of protein-coding genes (PCGs)

	Species/strains	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Amala-(EMI3)-Thailand (OK509083)														
2	Asufr-Shillong-India (KY548763)	**0.89**													
3	Eacon-(Chany)-Russia (ON644993)	23.21	23.20												
4	Ecapr-(SAMEA)-Egypt (AP017706)	21.21	21.38	22.34											
5	EcaSP-(JM-2019)-China (MH212284)	21.14	21.09	21.34	20.61										
6	EchCA2021-(PE4)-United States (MK264774)	23.29	23.19	18.16	22.86	21.80									
7	EchMSB-(A19)-United States (MN822299)	23.77	23.75	18.32	23.50	22.37	14.09								
8	Emiya-(HLJ)-China (MH393928)	21.07	21.15	22.66	12.88	20.23	22..77	22.86							
9	Emiya-(Hunan)-China (MN116740)	21.08	21.11	22.54	12.67	20.06	22.69	22.86	**0.52**						
10	Emiya-(RED11)-Thailand (OP326312)	21.00	21.06	22.52	12.65	20.08	22.72	22.86	**0.51**	**0.40**					
11	Epara-(KT008005)	21.65	21.60	22.31	12.71	20.53	22.76	23.36	12.79	12.62	12.63				
12	EcaSP-(GD)-China (MN116706)	21.11	21.08	21.34	20.62	**0.43**	21.80	22.42	20.29	20.16	20.20	20.60			
13	Erevo-(MSD15)-Thailand (MN496162)	21.51	21.64	22.11	13.67	20.11	23.29	23.06	11.87	11.73	11.77	13.18	20.20		
14	Hcono-(Hubei)-China (KM111525)	23.72	23.69	19.58	23.83	21.85	18.76	19.41	23.01	22.99	23.01	23.60	21.82	23.23	
15	Hcono-(RED42)-Thailand (PP110501)	**23.88**	**23.83**	19.84	23.83	22.00	18.89	19.63	23.13	23.10	23.12	23.63	22.01	23.33	**0.62**

Note: information for strains and species is given in **Supplementary Table S2.4**. The lowest and highest inter-specific genetic distances (%) are bolded and highlighted; the intra-specific distance rates (%) are bolded and squared. For clarity, the distances (%) between *Artyfechinostomum* species and eight *Echinostoma* congeners are vertically double-lined squared, and the distances (%) among the "*revolutum*" (*Eca. caproni*, *Eca. revolutum*, *Eca. myagawai*, and *Eca. paraensei*) species were horizontally squared and background shaded. (See text for more details).

A significant genetic distance (e.g., 21.0–21.65%) was found between the two groups: *Artyfechinostomum* spp. and the eight *Echinostoma* congeners. Four "*cryptic*" echinostosomes (two *Echinostoma* spp. from China and two Echinostomatidae spp. from the United States) belong to the latter group. There was a higher distance between the *Artyfechinostomum* group and *Echinoparyphium aconiatum* (23.21–23.22%), and between the *Artyfechinostomum* and *H. conoideum* (23.69–23.88%). The genetic difference between species in the "*revolutum*" group (*Eca. caproni*, *Eca. miyagawai*, *Eca. paraensei*, and *Eca. revolutum*) appeared to be substantially lower (11.73–13.67%) than between these group-species and other Echinostomatidae species (20.20–23.88%) (**Table 3.7**).

3.3 Polymorphismic features in non-coding regions (NCR) of echinostomes

3.3.1 Structural polymorphism in the NCRs of the Echinostomatidae species

A summary of the numbers and types of repetitive sequences in the non-coding regions (NCR) of the updated, available mtDNAs of 15 strains of 12 species in the family Echinostomatidae (suborder Echinostomata) is presented in **Table 3.8.**

Table 3.8. An updated summary of the mitogenomic datasets of all 15 strains of 12 species of the family Echinostomatidae including four echinostomes from this study regarding their length, the non-coding regions (NCR), repeats, PCGs, and references (if any).

No	Species	mtDNA length as reported (bp)	Length of NCR (bp)	Number and size of repeat units (RUs)	Coding region (bp)	PCGs (bp/aa)	References
1	***Artyfechinostomum malayanum* (Thailand)** (Amala-EMI3-TH); GenBank: OK509083	17,175	3,622	5.5 LRUs (336 bp/each) 7.5 SRUs (207 bp/each)	13,408	10,131/ 3365	This study Pham et al. (2022)
2	***Artyfechinostomum sufrartyfex* (India)** (Asufr-Shillong-IN); GenBank: KY548763	14,567	1,004	2.8 RUs (144 bp/each)	13,409	10,131/ 3365	GenBank
3	***Echinoparyphium aconiatum* (Russia)** (Eacon-Chany-RU); GenBank: ON644993	13,377	1,388	4.7 RUs (113 bp/each)	13,377	10,113/ 3359	Gacad et al. (2023)
4	***Echinostoma caproni* (Egypt)** (Ecapr-SAMEA-EG); GenBank: AP017706)	14,150	685	none	13,293	10,113/ 3364	GenBank
5	***Echinostoma* sp. (China)** (EcaSP-JM2019-CN); GenBank: MH212284)	15,283	1,877	5.6 LRUs (245 bp/each) 2.2 SRUs (166 bp/each)	13,257	10,122/ 3362	GenBank
6	**Echinostomatidae sp. CA-2021 (United States)** (EchCA2021-PE4-US); GenBank: MK264774	14,426	963	none	13,319	10,143/ 3369	GenBank
7	**Echinostomatidae sp. MSB para 30070 (United States)** (EchMSB-A19-US); GenBank: MN822299	13,985	474	none	13,346	10,128/ 3364	GenBank
8	***Echinostoma miyagawai* (China)** (Emiya-HLJ-CN); GenBank: MH393928	14,410	931	2.3 LRUs (319 bp/each) SRUs: n/a	13,321	10,128/ 3364	Li et al. (2019b)
9	***Echinostoma miyagawai* (China)** (Emiya-Hunan-CN); GenBank: MN116740	14,460	982	2.99 LRUs (319 bp/each) SRUs: n/a	13,320	10,128/ 3364	Fu et al. (2019)
10	***Echinostoma miyagawai*** (Emiya-RED11-TH); (GenBank: OP326312)	19,417	5,935	15.3 LRUs (319 bp/each) 4.8 SRUs (213 bp/each)	13,324	10,128/ 3364	This study Pham et al. (2024)
11	***Echinostoma paraensei*** (Epar); GenBank: KT008005	20,298*	6,798*	3.2 RUs (206 bp/each)	13,319	10,128/ 3364	GenBank
12	***Echinostoma revolutum*** (Erevo-MSD15-TH); GenBank: MN496162	17,030	3,549	7.6 LRUs (317 bp/each) 5.3 SRUs (207 bp/each)	13,326	10,137/ 3366	This study Le et al. (2020b)
13	***Echinostoma* sp.** (*revolutum*?) **(China)** (EcaSP-GD-CN); GenBank: MN116706	15,714	2,283	3.6 LRUs (245 bp/each) 3.6 SRUs (208 bp/each)	13,282	10,149/ 3371	Ran et al. (2020)
14	***Hypoderaeum conoideum* (China)** (Hcono-Hubei-CN); GenBank: KM111525	14,180	644	n/a	13,361	10,116/ 3360	Yang et al. (2015)
15	***Hypoderaeum conoideum* (Thailand)** (Hcono-RED42-TH); GenBank: PP110501	18,011	4,475	13 LRUs (241 bp/each) 9.7 SRUs (111 bp/each)	13,361	10,116/ 3360	This study

Note: non-coding region (NCR): sequence between the 3' terminus of $tRNA^{Glu}$ and the 5' terminus of *cox*3; *the NCR in *E. paraensei* was not fully sequenced, and the number of repeat units in this species is incomplete. LRU: long tandem repeat; SRU: short tandem repeat. These two terms are used when two different sizes are found in the NCR. n/a: not available; none: no repeats were found.

Among the 15 echinostomatid strains investigated, the NCRs of 11 strains possess two distinct types of tandem repeat units: long (LRUs) and short tandem repeat units (SRUs) which

vary in length and numbers. The *Artyfechinostomum malayanum* EMI3 strain had 5.5 LRUs (336 bp each) and 7.5 SRUs (207 bp each) [9], while the *A. sufrartyfex* Shillong strain had only 2.8 SRUs (114 bp each). The *Eca. miyagawai* RED11 strain from Thailand possesses 15.3 identical LRUs (319 bp each) and 4.8 SRUs (213 bp each), while the two Chinese strains (the HLJ and Hunan) had only LRUs (319 bp each) to be detected, which are 2.3 LRUs for the HLJ and 2.99 LRUs for the Hunan strains, respectively. The Thai *Hypoderaeum conoideum* RED42 strain had 13 LRUs (241 bp each) and 9.7 SRUs (111 bp each), while the Chinese congener Hubei strain had no repeat units (**Table 3.8**).

The mtDNAs of the other five echinostomatid species also contained repeat units, including *Eca. paraensei* from Egypt (with at least 3.2 RUs of 206 bp each), the *Eca. revolutum* MSD15 strain from Thailand (with 7.6 LRUs of 317 bp each and 5.3 SRUs of 207 bp each), the *Echinoparyphium aconiatum* Chany strain from Russia (with 4.7 RUs of 113 bp each), the "cryptic" *Echinostoma* GD strain from China (with 3.6 LRUs of 245 bp each and 3.6 SRUs of 208 bp each), and the "cryptic" *Echinostoma* JM2019 strain from China (with 5.6 LRUs of 245 bp each and 2.2 SRUs of 166 bp each) (**Table 3.8**). The authors' previous analyses [38, 91–93, 142] did not specify these repetitive features in the NCR of all these strains.

The repetitions vary in number and length, and the majority of them were found in two subregions of the NCR, the LRU and SRU subregions, which are linked by a short nucleotide sequence. The various repeats impacted the size of the NCR, causing the length of the mtDNAs to vary between and even between strains of the echinostome species. Interestingly, *Echinostoma caproni* (from Egypt), *Hypoderaeum conoideum* (China), and two Echinostomatidae spp. (from the United States) do not have any repeats in their NCRs.

3.3.2 The NCRs featured by the promotor and regulatory sequences in the NCR's long and short repeat units in echinostomes

To identify putative promotor regions in LRU and SRU sequences of *Echinostoma miyagawai*, the SAPPHIRE.CNN ((SAPPHIRE (kuleuven.be) was used. The analyses of LRU and SRU did reveal several putative promoter regions, the likelihood of which were all significant with p-values <0.01. Both repeat units showed there to be clusters of putative promoter sequences predominantly at the at the start and end of the sequence. In total the LRU had five putative promoter sequences from 14–88 bp and then at the later end of the sequence from 204–253 bp. The SRU had a total of eight putative promoter sequences, with four overlapping promoters identified from 1–74 bp and then a further four promoters identified between 130–207 bp (**Fig. 3.3**). Conserved motifs, including TA(A)n-like sequences, TATA motifs and G(A)nT motifs, typical to the initiation sites for replication and transcription were

found across the promoters in the LRU and SRU. However, only promoter 5 in the LRU had indications of a poly T motif at the 5' end.

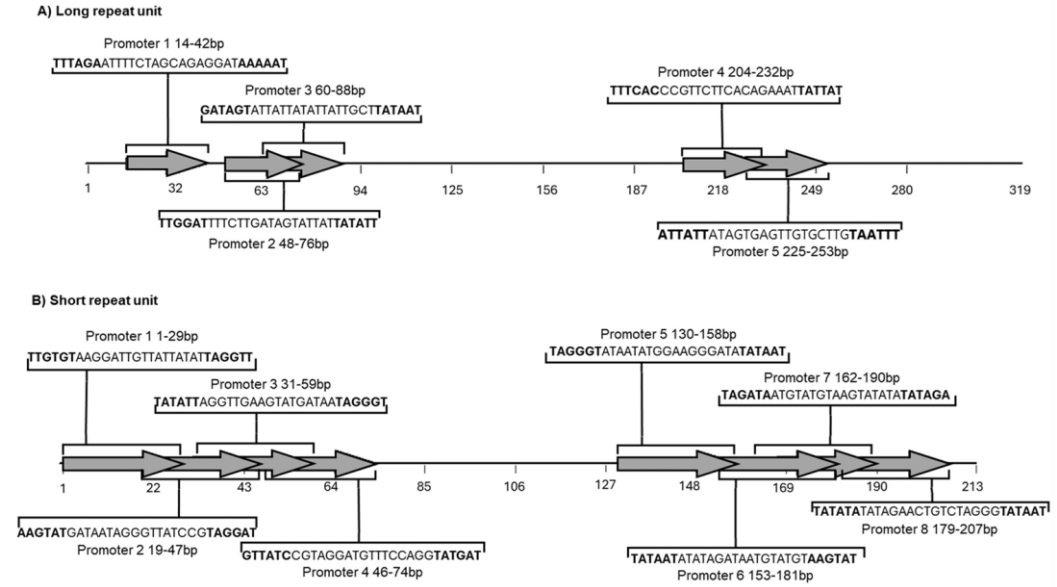


Figure 3.3. Schematic of the position of predicted promoter regions within the tandem repeat units repeat units of *Echinostoma miyagawai* mitochondrial control region. **A**) illustration of the identification of the five putative promoters within the long repeat unit (LRU); and **B**) illustration of the identification of eight putative promoter regions within the short repeat unit.

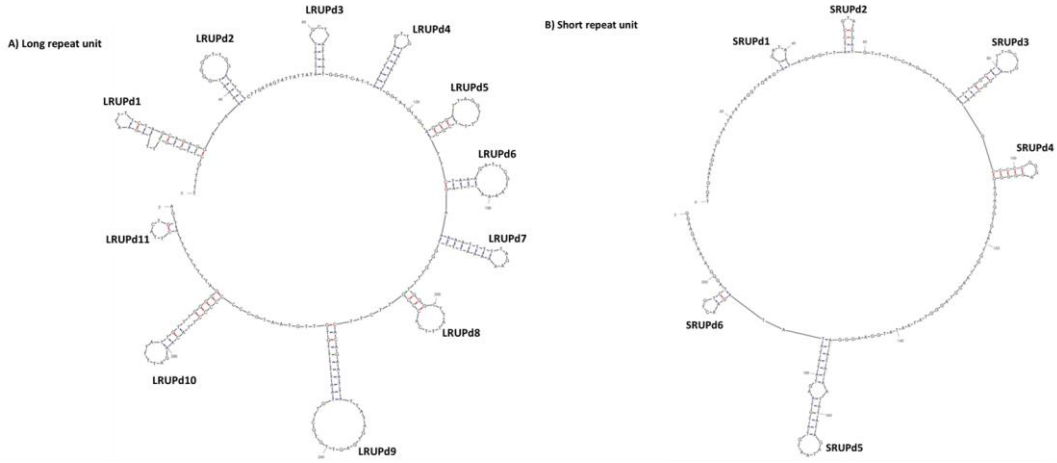


Figure 3.4. Palindromic repeat regions identified in the long repeat unit and short repeat unit in non-coding region of *Echinostoma miyagawai*. LRUPd: long repeat unit palindrome; SRUPd: short repeat unit palindrome. There are 11 LRUPd (1–11) and five SRUPd (1–5) found in the NCR of *Eca. miyagawai*.

3.3.3 The NCRs featured by the palindromic sequences in the *Echinostoma* species

A total of 11 palindromic repeat regions were identified in the long repeat units of *Eca. miyagawai* with a further six identified in the short repeat units (**Fig. 3.4**). In both cases their appeared to be a high GC or AT content with a bias to either set of nucleotides. There was consistency in palindromic hairpin length nor the size of the resultant loop domains. However, three large palindromes were identified in the LRU, these were LRUPd1, LRUPd9 and LRUPd10, each of which with had mismatched base pairs or extra nucleotides in the arm of the

hairpin region. This was also true for the large palindromic sequences identified in the SRU, denoted SRUPd3 and SRUPd5 (**Table 3.9**).

Table 3.9 List of palindromic sequences found in the long and short repeat unit in the mitochondrial control region of *Echinostoma miyagawai*

Repeat unit	Palindrome name	Position	Predicted palindrome sequence
Long repeat unit	LRUPd1	1−34	5’-CTTCTG***T***(**T$_2$**)TAGA-(N$_4$)-TCTA***G***CAGAGG-3’
	LRUPd2	38−56	5’-AAAAT-(N$_9$)-ATTTT-3’
	LRUPd3	74−88	5’-ATTAT-(N$_5$)-ATAAT-3’
	LRUPd4	96−114	5’-AAATTTA-(N$_4$)-TAAATTT-3’
	LRUPd5	125−143	5’-AGCGA-(N$_9$)-TCGCT-3’
	LRUPd6	146−166	5’-CTAAA-(N$_{11}$)-TTTAG-3’
	LRUPd7	168−188	5’-AAAATTTT-(N$_5$)-AAAATTTT-3’
	LRUPd8	196−212	5’-TGGG-(N$_9$)-CCCG-3’
	LRUPd9	128−25	5’-CACA***G***AAATTA-(N$_{19}$)-TAATTTT***T***GTG-3’
	LRUPd10	271−299	5’-CCCCC(**T$_2$**)ACAA-(N$_7$)-TTGT(**T$_2$**)GGGGG-3’
	LRUPd11	311−319	5’-AC-(N$_5$)-GT-3’
Short repeat unit	SRUPd1	35−42	5’-AT-(N$_4$)-AT-3’
	SRUPd2	49−59	5’-ATCC-(N$_3$)-GGAT-3’
	SRUPd3	73−95	5’-ATTG***G***CAT-(N$_7$)-ATG***G***CAAT-3’
	SRUPd4	97−110	5’-CCCCC-(N$_4$)-GGGGG-3’
	SRUPd5	150−185	5’-ATATATA(**A**)TATATA-(N$_7$)-TATGTA(**TG**)TATATAT-3’
	SRUPd6	188−197	5’-AGA-(N$_4$)-TCT-3’

Note: N indicates the loop domain; italicized nucleotides indicate mismatches; and bolded nucleotides indicate extra uncomplimented nucleotide(s). LRUPd: long repeat uni palindrome; SRUPd: short repeat uni palindrome.

3.3.3 The NCRs featured by the evolution of mitochondrial control region in the Echinostomatidae species

Interspecies comparisons of the LRU indicated that there were only five other species of echinostomatids with available mitochondrial sequences with similarity to the *Eca. miyagawai* LRU. These species included *Eca. revolutum*, *Eca. caproni*, *Eca. paraensei*, *A. malayanum*, and an unknown species Echinostomatidae sp. CA-2021, all of which were only represented by partial sequences. *Echinostoma revolutum* and *Eca. caproni* had the most complete LRU sequences for comparison with *Eca. miyagawai*, with *Eca. revolutum* sharing seven palindromic hairpin sequences with LRUPd4 and LRUPd6 being absent. Also, there were differences in size of palindromes in *Eca. revolutum* relative to that of *Eca. miyagawai*, with LRUPd2, LRUPd8 and LRUPd10 being highly extended owing to the occurrence of mismatched or extra base pairs, however the content of the loop domains appeared to be conserved between the two species (**Fig. 3.4**).

In comparisons between *Eca. miyagawai* and *Eca. caproni* again seven hairpin palindromes were shared, but unlike in *Eca. revolutum* LRUPd7 missing. In *Eca. caproni* LRUPd5 and LRUPd9 were substantially extended although the loop domains were homologus to those found in *Eca. miyagawai* and *Eca. caproni*. Interestingly, *Eca. caproni* also had two other unique palindromes, which were absent in the other two species. It is also important to note that the absence of palindromes LRUPd1 and LRUPd11 was the result of missing comparable sequence data from *Eca. revolutum* and *Eca. caproni*.

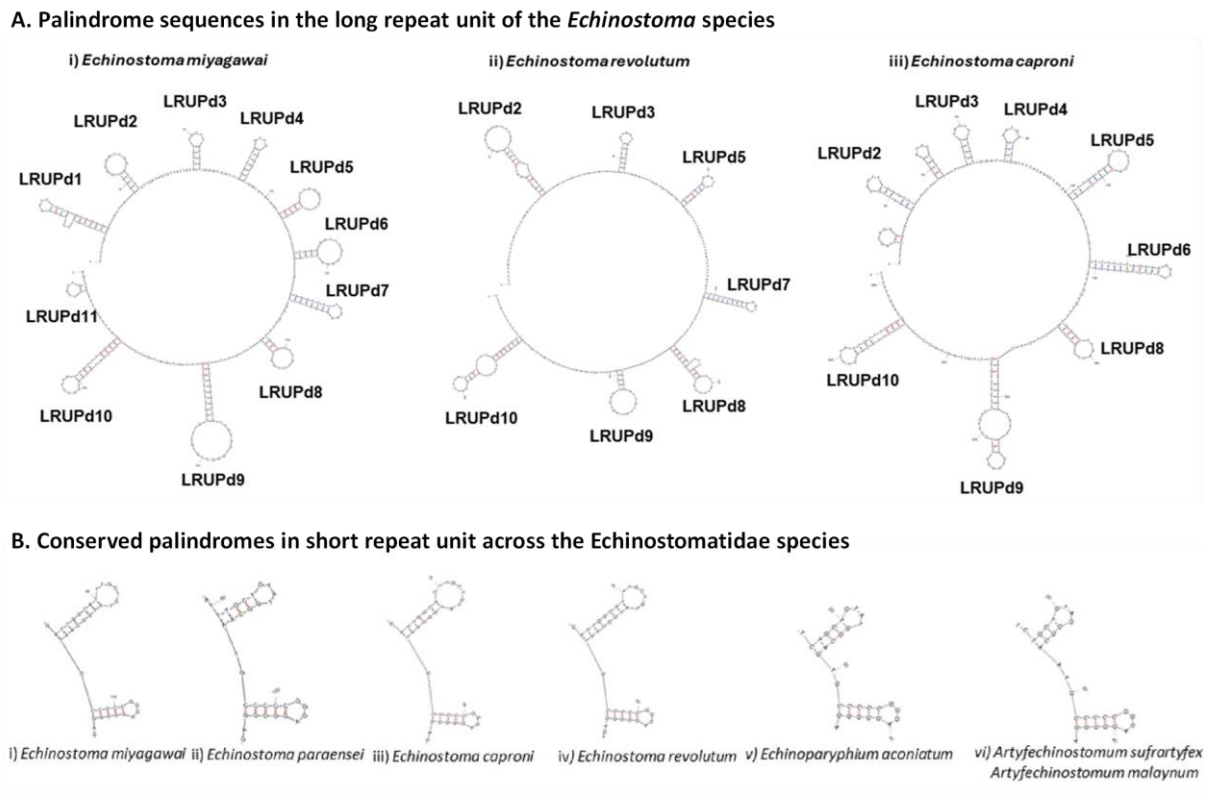


Figure 3.4. Comparative analyses of the long tandem repeat unit the *Echinostoma* species and short repeat unit containing the conserved palindromes across the Echinostomatidae species. **A**. Illustrated presentation of the number the palindromic sequences in the full LRU compared between *Eca. miyagawai*, *Eca. revolutum*, and *Eca. caproni*; **B**. Illustrated presentation of the conserved palindromes across the Echinostomatidae species. LRUPd: long repeat uni palindrome (see **Table 3.9**).

In contrast, the interspecies comparisons of SRU revealed eight other species sharing homology, including *Eca. paraensei*, *Eca. caproni*, *Eca. revolutum*, Echinostomatidae sp. MSB para 30070 isolate A19, *Echinoparyphium aconiatum*, *H. conoideum, A. malayanum*, and *A. sufrartyfex.* However, all sequences were partial and comparative palindromic hairpin analyses could only be performed on a 30-base pair region, which was denoted SRUPd3 and SRUPd4 in *Eca. miyagawai* (**Fig. 3.4**). These palindromic hairpins were consistent across all species, although SRUPd3 did appear to vary in length between species, SRUPd4 was highly conserved and identical in each of the echinostomatids.

3.4 The ribosomal transcription units of five echinostomes and their implications

3.4.1 Ribosomal transcription unit features of the echinostomatid and echinochasmid species

There were four echinostomatids (*Eca. revolutum*, *A. malayanum*, *Eca. miyagawai*, and *H. conoideum*) and one echinochasmid species (*Ecs. japonicus*), that the nuclear ribosomal transcription unit (rTU) sequences were obtained. The complete rTU sequences were for *Artyfechinostomum malayanum* (9,499 bp) with the ETS and the IGS regions and near-complete for *Hypoderaeum conoideum* (8,076 bp) with the IGS but not the ETS. For the other three

species, the ribosomal sequences obtained were the coding regions (designated as rTU*, from the 5' terminus of the 18S rRNA to the 3' terminus of the 28S rRNA gene) lacking the ETS and IGS but including the internal transcribed spacers and the 5.8S rRNA gene. These three species were *Eca. miyagawai* (6,854 bp), *Eca. revolutum* (6,856 bp), and *Ecs. japonicus* (7,150 bp) (**Table 3.10**). GenBank accession numbers are OR509026–OR509030 for *A. malayanum*, *Eca. miyagawai*, *Eca. revolutum*, *H. conoideum*, and *Ecs. japonicus*, respectively.

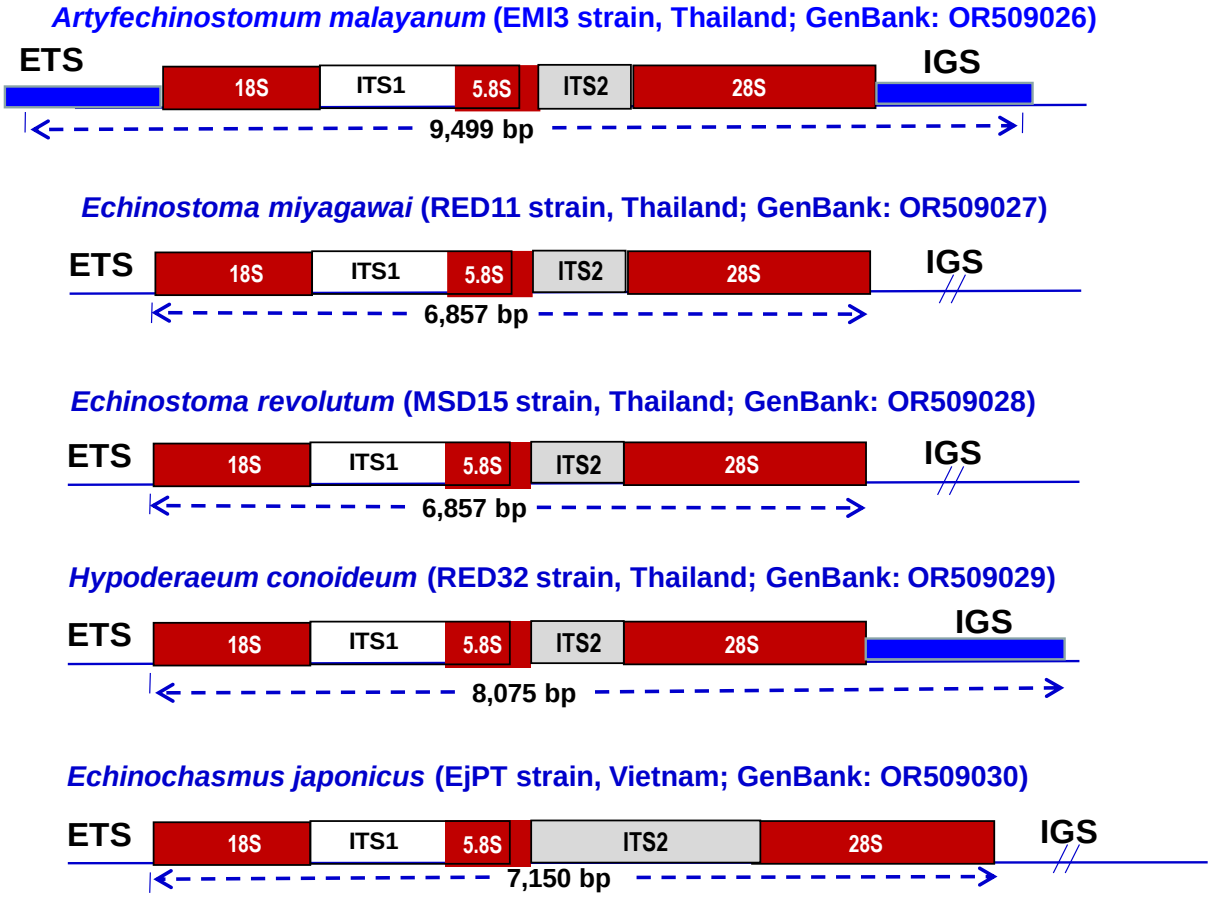


Figure 3.5. Structural organization of the near-complete ribosomal transcription units for echinostomatids and echinochasmids (*A. malayanum*, *Eca. miyagawai*, *Eca. revolutum*, *H. conoideum*, and *Ecs. japonicus*).

In all five species, the 18S rRNA was identical in length (1,988 bp), as was the 5.8S rRNA gene (160 bp). The 28S sequences were all similar in length (3,861–3,864 bp). The ITS1 regions ranged from 416 bp (*Eca. revolutum*) to 423 bp (*H. conoideum*) and from 427 bp (*A. malayanum*) to 431 bp (*H. conoideum*) for ITS2. The exception was *Ecs. japonicus*, with a longer ITS1 (436 bp) and a much longer ITS2 (705 bp). No repeat units were found in the ITS regions of the *Echinostoma* and *Echinochasmus* species. For *H. conoideum*, the complete IGS sequence (1,211 bp), and for *A. malayanum*, both the ETS (1,673 bp) and IGS (969 bp) were successfully obtained. Four tandem repeats (89 bp each) were found in the ETS of *A.*

malayanum, as well as some groups of short repeats, including three (27 bp each), 4.7 (15 bp each), 3.5 (33 bp each), and 2.5 (42 bp each) in the IGS of *H. conoideum* (**Table 3.10**).

Table 3.10 Positions of ribosomal RNA genes, internal transcribed spacers (ITS), external transcribed spacer (ETS), and non-transcribed intergenic spacers (IGS) in the ribosomal transcriptional unit of *Echinostoma* species (Echinostomatidae) and *Echinochasmus* (Echinochasmidae) species in this study

Species	The order of the ribosomal transcription unit (rTU), position and length of each region/gene						
	ETS	Coding region (rTU*)					IGS
		18S	ITS1	5.8S	ITS2	28S	
Artyfechinostomum malayanum (9,499 bp) GenBank: OR509026	1–1673 (1,673 bp) (complete) 1–89 (RU1) 90–178 (RU2) 179–267 (RU3) 268–445 (RU4)	1674–3661 (1,988 bp)	3662–4080 (419 bp)	4081–4240 (160 bp)	4241–4667 (427 bp)	4668–8530 (3,864 bp)	8531–9499 (969 bp) (complete)
Echinostoma miyagawai (6,854 bp); GenBank: OR509027	N/A	1–1988 (1,988 bp)	1989–2405 (417 bp)	2406–2565 (160 bp)	2566–2993 (428 bp)	2994–6854 (3,861 bp)	N/A
Echinostoma revolutum (6,856 bp); GenBank: OR509028	N/A	1–1988 (1,988 bp)	1989–2404 (416 bp)	2405–2564 (160 bp)	2565–2993 (429 bp)	2994–6856 (3,863 bp)	N/A
Hypoderaeum conoideum (8,075 bp); GenBank: OR509029	N/A	1–1988 (1,988 bp)	1989–2411 (423 bp)	2412–2571 (160 bp)	2572–3002 (431 bp)	3003–6864 (3,862 bp)	6865–8075 (1,211 bp); short repeats: 3 (27 bp); 4.7 (15 bp); 3.5 (33 bp); 2.5 (42 bp)
Echinochasmus japonicus (7,150 bp); GenBank: OR509030	N/A	1–1988 (1,988 bp)	1989–2424 (436 bp)	2425–2584 (160 bp)	2585–3289 (705 bp)	3290–7150 (3,861 bp)	N/A

Note: rTU: ribosomal transcription unit; rTU*: coding region of the rTU (from the 5' terminus of 18S to the 3' terminus of the 28S rRNA genes); ETS: external transcribed spacer; ITS: internal transcribed spacer; rRNA: ribosomal gene; N/A: not available for sequencing.

3.4.2 *De novo* structure of the 28S rRNA gene of echinostomes

The complete nucleotide sequence of the 28S rRNA gene was obtained from three species of the genus *Echinostoma*, including *Eca. revolutum* (3,863 bp); *Eca. malayanum* or *A. malayanum* (3,863 bp); *Eca. miyagawai* (3,861 bp); and a species of the genus *Hypoderaeum*, i.e., *H. conoideum* (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. 3.6**.

The 28S rRNA ribosomal gene of *Eca. revolutum* is also divided into two branches, but overall, it is slightly different from those of *A. malayanum*, *Eca. 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 [82]. The arrangement of GC and AT nucleotides in the 28S rRNA sequences of these three species (*A. malayanum*, *Eca. miyagawai* and *H. conoideum*) is totally

the same, although *H. conoideum* belongs to the different genus *Hypoderaeum*. Although *Eca. 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. 3.6**).

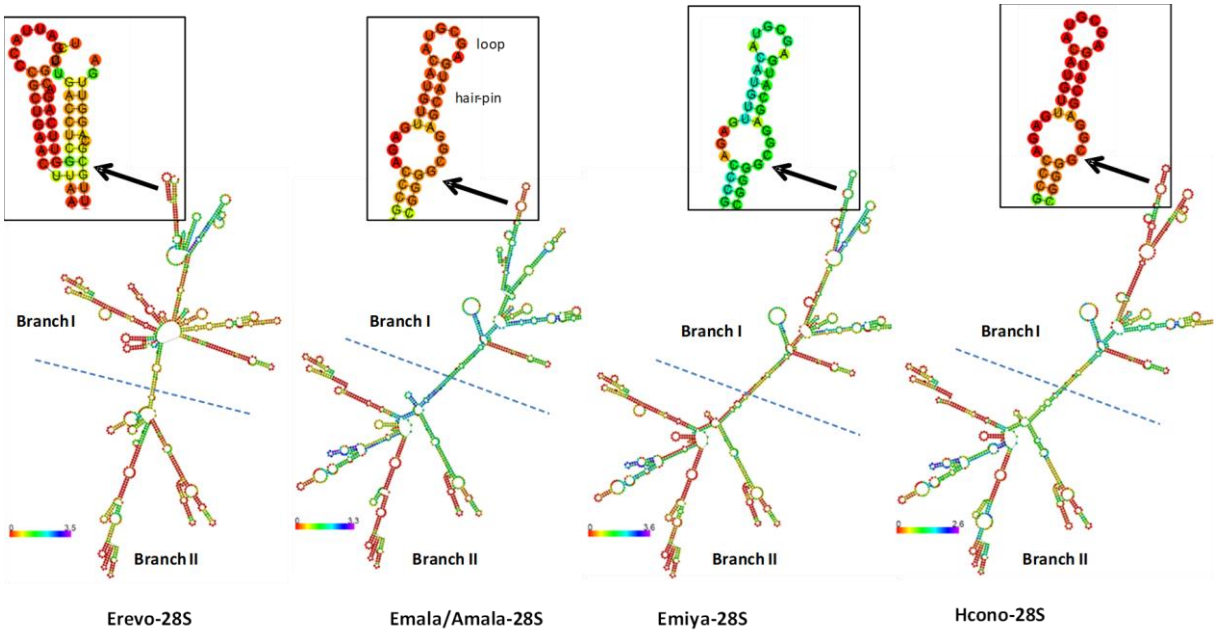


Figure 3.6. The *de novo* secondary structure model of the 28S rRNA gene (domains D1–D3) was modeled and constructed in the RNAfold program, with minimum free energy (MFE) of –437.80 kcal/mol. Structures containing "hairpin" sequences (or stem-loop) are formed from opposing nucleotide sequences (palindrome); 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.

3.5 Mitophylogenetic relationships within Echinostomatidae and Echinostomata

To assess the mitophylogenetic and taxonomic relationships within Echinostomatidae and Echinostomata, the alignment of the concatenated sequences of the amino acids inferred from the PCGs from 57 complete mitogenomes of 41 trematode species of five families from the suborder Echinostomata (i.e., Echinostomatidae, Echinochasmidae, Fasciolidae, Cyclocoelidae, and Himasthlidae), two from the suborder Opisthorchiata (Opisthorchiidae and Heterophyidae), and two families from suborder Xiphidiata (Paragonimidae and Dicrocoeliidae) was conducted. Family Schistosomatidae (*Schistosoma haematobium* species) was used as an outgroup (For information, see **Supplementary Table S2.4**).

The ML phylogenetic tree revealed four distinct groups with high bootstrap-support: Echinostomata, Opisthorchiata, and two families, Paragonimidae and Dicrocoeliidae from Xiphidiata. The Echinostomata is a well-supported monophyletic, the Opisthorchiata is a paraphyletic with two families (Opisthorchiidae and Heterophyidae), while the Xiphidiata suborder is significantly polyphyletic, with two families positioned separately (**Fig. 3.7**).

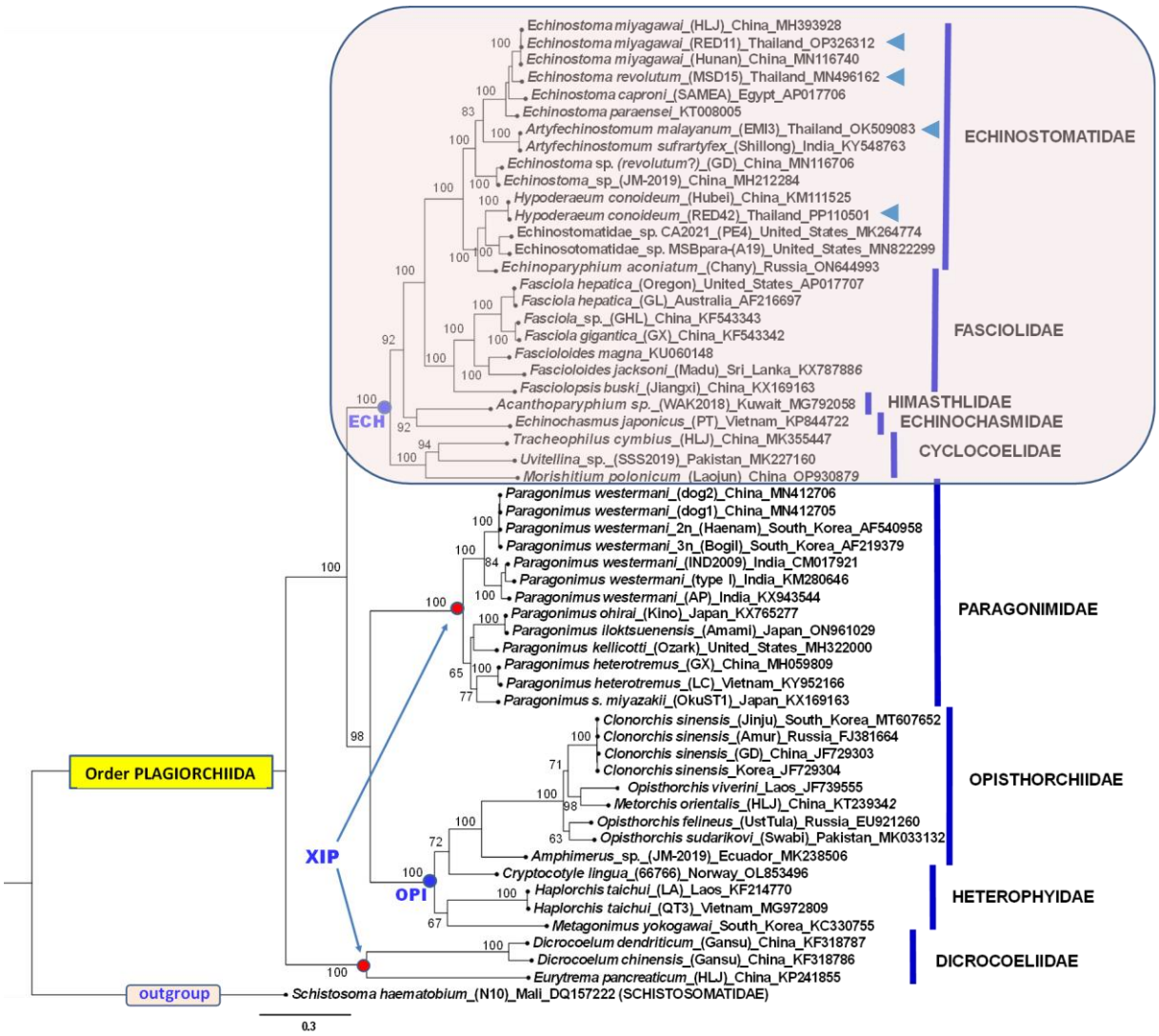


Figure 3.7. A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatidae and among other superfamilies and suborders. The tree was reconstructed based on the analysis of the concatenated sequences of the amino acids inferred from the PCGs from 57 complete mitogenomes of 41 trematode species of five families from the suborder Echinostomata (i.e., Echinostomatidae, Echinochasmidae, Fasciolidae, Cyclocoelidae, and Himasthlidae), suborder Opisthorchiata (Opisthorchiidae and Heterophyidae), and suborder Xiphidiata (Paragonimidae and Dicrocoeliidae). Family Schistosomatidae (*Schistosoma haematobium* species) was used as an outgroup (**Supplementary Table S2.4**). The alignment was performed by MAFFT v7.407, curated by BMGE v1.12, and the tree was reconstructed in PhyML v3.3.1 using a maximum likelihood method with 1000 bootstrap resamplings. The output Newick tree was extracted and visualized using FigTree v1.4.4. Nodal support values evaluated using 1000 bootstrap resamplings are shown on each branch. The suborder Echinostomata is transparently squared and designated by "ECH" at its basal node with a solid circle; likewise, the Opisthorchiata by "OPI", and the Xiphidiata by "XIP". Following the species' full name are the strain designations (where available) in brackets and the country's full name; accession numbers are given for each species or strain at the end of each sequence label. The sequences of this study are indicated by an arrow. The scale bar represents the number of substitutions per site. For clear presentation, the order Plagiorchiida in shown at the root of all the suborders as well.

The topology demonstrated the monophyletic status of Echinostomata, with the Echinostomatidae family classified as a "sister" group to the Fasciolidae. Six *Echinostoma* strains of *Eca. miyagawai*, *Eca. revolutum*, *Eca. caproni*, and *Eca. paraensei* were placed into one subcluster and are a "sister" to the *Artyfechinostomum* subclade (comprising *A. malayanum* and *A. sufrartyfex* species), whereas two strains of *Hypoderaeum conoideum*, one from

Thailand and another from China ((JM-2019/MH212284 and GD/MN116706) formed a subgroup, monophyletic to the two Echinostomatidae spp. of the United States MSB_(A19)/MN822299 and CA-2021-(PE4)/MK264774 strains). The [Echinochasmidae + Himasthlidae] form a monophyletic group with the [Echinostomatidae + Fasciolidae], however the Cyclocoelidae (with the addition of the *Morishitium polonicum*_(Laojun)_China_OP930879 strain (Liu et al., 2023) are paraphyletic (**Fig. 3.7**). The interesting phylogenetic relationships of the Echinostomata suborder, particularly the *Echinostoma* genus, and the genetic subdivision of the newly discovered "cryptic" species within, need to be further investigated.

In the phylogenetic tree, the Xiphidiata is always separated and polyphyletically positioned from the suborders mentioned above. This xiphidiatan group is divided into two subgroups: the Paragonimidae family, which is nested between the suborders Echinostomata and Opisthorchiata, and the Dicrocoeliidae family, which is marginally separated from the aforementioned clades. The clade Paragonimidae was separated into two subgroups: the solitary Paragonimus westermani complex and the other species, which include the most recently sequenced *Paragonimus* species (*Paragonimus skrjabini miyazakii*, *P. heterotremus*, *P. ohirai*, *P. iloktsuenensis*, and *P. kellicotti*) [32, 35, 85, 86]. In contrast to the tight monophyly of the Echinostomata, the Opisthorchiata appeared to be a polyphyletic clade, with the placement of the newly sequenced *Cryptocotyle lingua* (GenBank: OL853496), which was morphologically classified into Heterophyidae [166] but falls into the Opisthorchiidae subclade in these analyses (**Fig. 3.7**).

3.6 Phylogenetic analyses utilizing the ribosomal tráncription unit sequences

3.6.1 Interfamilial phylogenetic relationships within Echinostomata and among other suborders

The ML tree based on the **concatenated sequences** of the 28S and 18S rRNA genes (**Fig. 3.8)** clearly demonstrated the monophyly of Echinostomata with an absolute (100%) bootstrap support and a sister relationship to a group containing Opisthorchiata, and Xiphidiata, with a bootstrap support of 95%.

The second ML tree was based on the alignment of 70 **complete 28S sequences** for Echinostomata (families Echinostomatidae, Echinochasmidae, and Cyclocoelidae), and several families in the Xiphidiata (**Fig. 3.9**). This phylogeny was intended to test for congruence between analyses based on 28S rRNA alone and those based on concatenated 28S and 18S ribosomal sequences. The monophyletic status of the Echinostomata, Opisthorchiata, and Diplostomata was not affected, and within the Echinostomata, the family Echinochasmidae appeared as a sister to a subclade encompassing [(Echinostomatidae + Fasciolidae) +

(Philophthalmidae + Cyclocoelidae)] with bootstrap nodal support of 95–100%. The topology also agreed with that in **Fig. 3.8** in the paraphyly of the Xiphidiata with respect to Opisthorchiata (again with very low bootstrap support) (**Fig. 3.9**). In both **Figs. 3.8** and **3.9**, sequences of the xiphidiatan family Haploporidae produced a topology indicating this paraphyly.

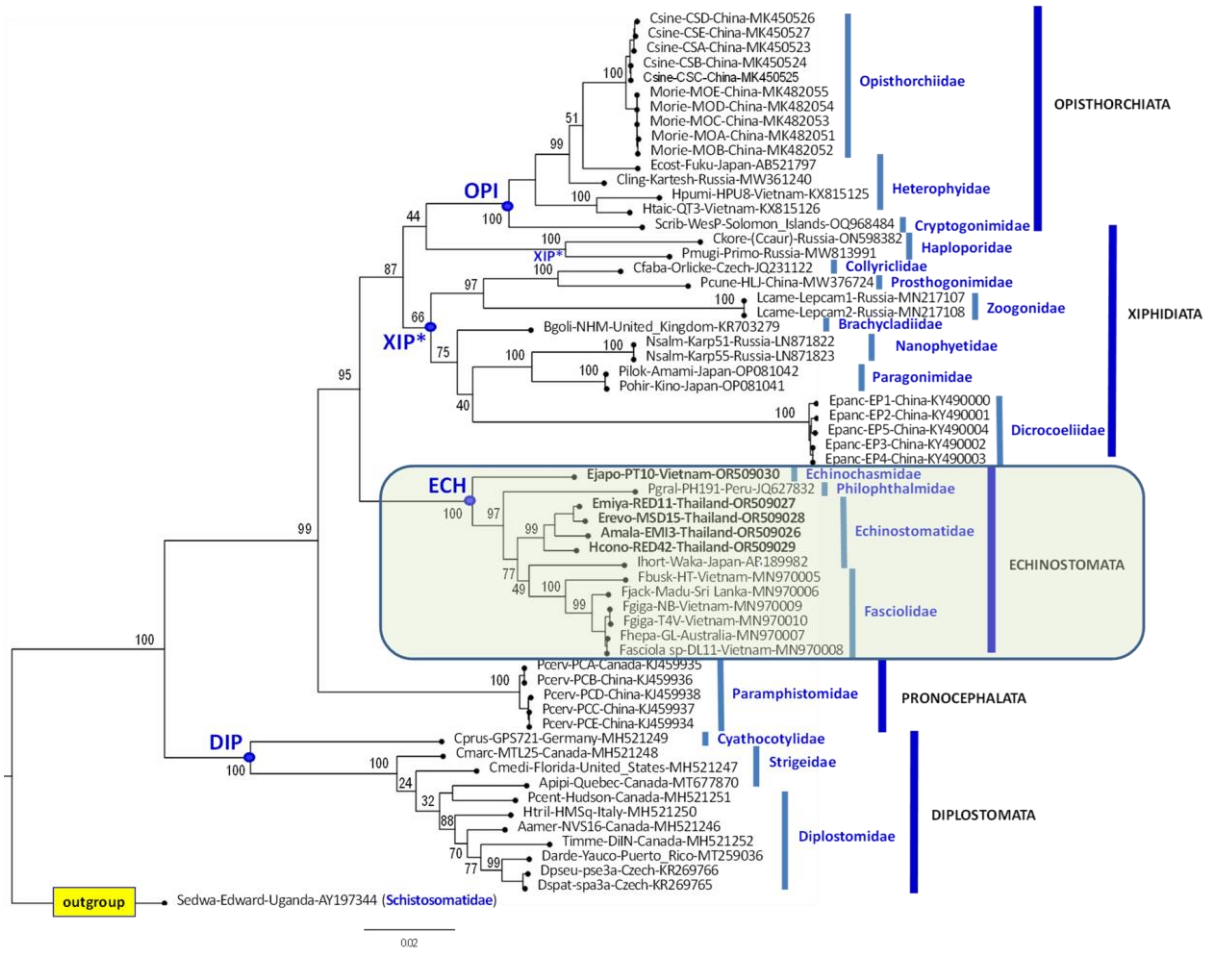


Figure 3.8. A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatoidea (suborder: Echinostomata) and among other superfamilies and suborders (Opisthorchiata, Xiphidiata, Pronocephalata, and Diplostomata). The tree was reconstructed based on the analysis of the concatenated sequences of the complete 18S rRNA and 28S rRNA genes from 60 complete ribosomal transcription units of 42 species of 21 families (**Supplementary Table S2.5**). *Schistosoma edwardiense* (Digenea: Schistosomatidae) is included as an outgroup. The alignment was performed by MAFFT v7.407, curated by BMGE v1.12, and the tree was reconstructed in PhyML v3.3.1 using a maximum likelihood method and 1000 bootstrap resamplings. The output Newick tree was extracted and visualized using FigTree v1.4.4. Nodal support values evaluated using 1000 bootstrap resamplings are shown on each branch. The superfamily Echinostomatoidea (in Echinostomata) is shown in a highlighted box and designated by "ECH" at its basal node; the superfamilies Opisthorchioidea (in Opisthorchiata) by "OPI", the suborder Xiphidiata by "XIP*" (*showing the paraphyly of the Xiphidiata; see text), and Diplostomoidea (in Diplostomata) by "DIP". Following the species' abbreviated name (five letters) are the strain designations (where available) and the country's full name; accession numbers are given for each species or strain at the end of each sequence label. The sequences of this study are in bold font. The scale bar represents the number of substitutions per site.

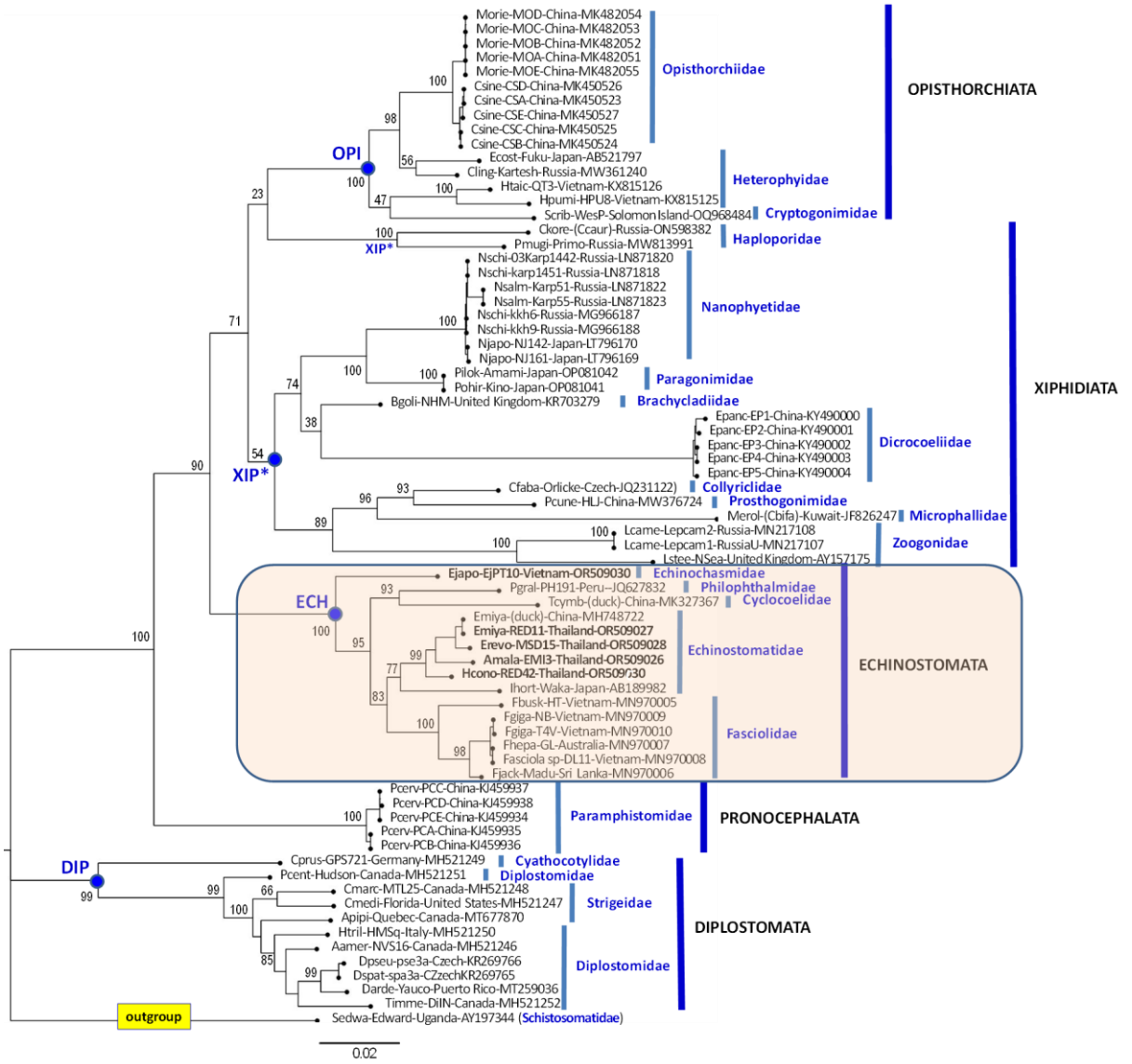


Figure 3.9. A maximum-likelihood phylogenetic tree showing the phylogenetic relationships among the taxa within the Echinostomatoidea (suborder: Echinostomata) and between the other superfamilies and suborders (Opisthorchiata, Xiphidiata, Pronocephalata, and Diplostomata). The tree reconstruction was based on the analysis of 70 sequences of the complete 28S rRNA gene of 49 species of 21 families (**Supplementary Table S2.5**). *Schistosoma edwardiense* (Digenea: Schistosomatidae) is included as an outgroup. The alignment was performed by MAFFT v7.407, curated by BMGE v1.12, and the tree was reconstructed in PhyML v3.3.1 using a maximum likelihood method and 1000 bootstrap resamplings. The output Newick tree was extracted and visualized using FigTree v1.4.4. Nodal support values, evaluated using 1000 bootstrap resamplings are shown on each branch. The superfamily Echinostomatoidea (in Echinostomata) is shown in a highlighted box and designated by "ECH" at its basal node; the superfamilies Opisthorchioidea (in Opisthorchiata) by "OPI", the suborder Xiphidiata by "XIP*" (*showing the paraphyly of the Xiphidiata; see text), and Diplostomoidea (in Diplostomata) by "DIP". Following the species' abbreviated name (five letters) are the strain designations (where available) and the country's full name; accession numbers are given for each species or strain at the end of each sequence label. The sequences of this study are shown in bold font. The scale bar represents the number of substitutions per site.

The second ML tree was based on the alignment of 70 **complete 28S sequences** for Echinostomata (families Echinostomatidae, Echinochasmidae, and Cyclocoelidae), and several families in the Xiphidiata (**Fig. 3.9**). This phylogeny was intended to test for congruence between analyses based on 28S rRNA alone and those based on concatenated 28S and 18S ribosomal sequences. The monophyletic status of the Echinostomata, Opisthorchiata, and

Diplostomata was not affected, and within the Echinostomata, the family Echinochasmidae appeared as a sister to a subclade encompassing [(Echinostomatidae + Fasciolidae) + (Philophthalmidae + Cyclocoelidae)] with bootstrap nodal support of 95–100%. The topology also agreed with that in **Fig. 3.8** in the paraphyly of the Xiphidiata with respect to Opisthorchiata (again with very low bootstrap support) (**Fig. 3.9**). In both **Figs. 3.8** and **3.9**, sequences of the xiphidiatan family Haploporidae produced a topology indicating this paraphyly.

3.6.2 Phylogenetic relationships within families Echinostomatidae and Echinochasmidae

Both 28S alone and concatenated complete 28S+18S sequence datasets recovered all *Echinostoma* species as a monophyletic group sister to *Artyfechinostomum malayanum*. In these phylogenetic analyses, the genus *Echinochasmus* (family Echinochasmidae), represented by *Ecs. japonicus*, formed the basal branch in Echinostomata with the Philophthalmidae and Cyclocoelidae diverging next. To add more taxa and explore these relationships more fully, we constructed a comprehensive phylogenetic tree using the alignments of 169 available **partial 28S sequences** (**D1–D3** regions; about 1.1–1.3 kb prior to alignment).

The detailed ML phylogenetic tree (**Fig. 3.10**) clearly demonstrated the monophyly of the suborder Echinostomata in the order Plagiorchiida with 99% bootstrap support, distinct from Xiphidiata (represented by the family Eucotylidae, which belongs to the superfamily Microphalloidea) [167] and Haplosplanchnata (with Haplosplanchnidae of the Haplospanchnoidea) [63] with 70% support.

The family Echinochasmidae is strongly supported as monophyletic and is separated from the Echinostomatidae in **Fig. 3.10** by several other families; Caballerotrematidae, Himasthlidae, Fasciolidae and Psilostomidae. Interestingly, the Echinochasmidae does not appear basal within the Echinostomata in this analysis. Instead, the Philophthalmidae occupies this position, followed by the Cyclocoelidae.

The 15 representative species from three genera of the Echinochasmidae were subdivided into two strongly supported subgroups where Subgroup 1 included the genera *Microparyphium* and some *Echinochasmus* species while Subgroup 2 contained the genus *Stephanoprora* and other members of *Echinochasmus*: this latter genus is therefore polyphyletic. Within Subgroup 1, sequences of *Ecs. japonicus* appeared in two well-separated branches: one included samples from Nam Dinh (Vietnam), and the other samples from Phu Tho and Hoa Binh Provinces (Vietnam).

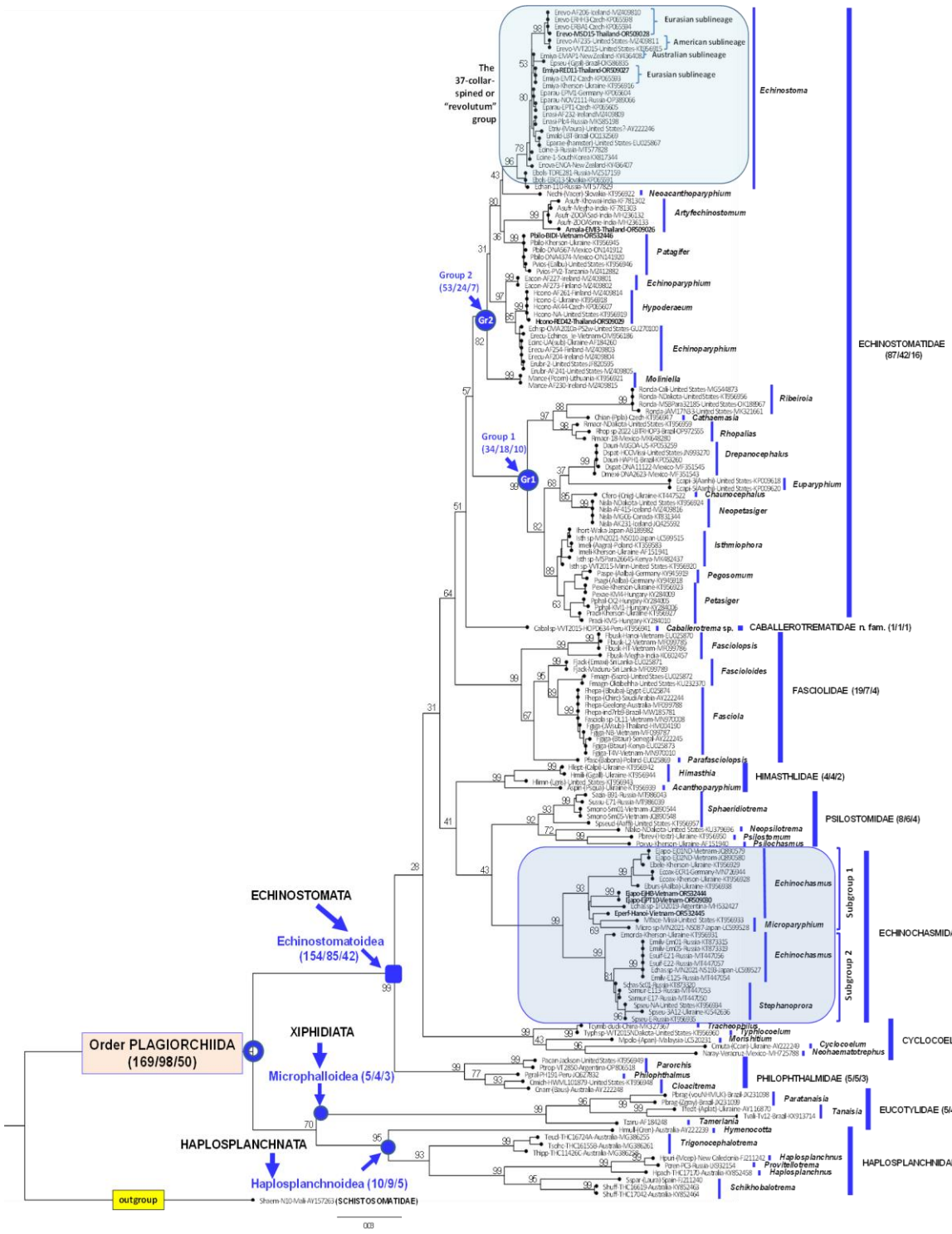


Figure 3.10. Phylogeny based on the analysis of the D1−D3 sequences (1.1–1.3 kb) of the 28S rRNA genes showing the detailed relationships of the families Echinostomatidae and Echinochasmidae and other families in the suborder Echinostomata. Sequences from other suborders in the Plagiorchiida have also been included. In total, 169 **D1−D3 28S sequences** from 98 species of 42 genera in 9 families were included (**Supplementary Table S2.6**). The numbers in brackets indicate sequences, species, and genera, accordingly. The families included in the phylogeny from Echinostomatoidea/Echinostomata (marked by a small solid square at the node and also shown by an arrow) are Caballerotrematidae, Fasciolidae, Himasthlidae, Psilostomidae, Cyclocoelidae, and Philophthalmidae; the others from Microphalloidea/Xiphidiata and from Haplospanchnoidea/ Haplosplanchnata are referred to as outgroup taxa. *Schistosoma haematobium* (Digenea: Schistosomatidae) is included as an outgroup. The alignment was performed by MAFFT v7.407, curated by BMGE v1.12, and the tree was reconstructed in PhyML v3.3 using a maximum likelihood method and 1000 bootstrap resamplings. The output Newick tree was extracted and visualized using FigTree v1.4.4. Nodal support values evaluated using 1000 bootstrap resamplings are shown on each branch. Following the species' abbreviated name (five letters) are the

strain designations (where available) and the country's full name; accession numbers are given for each species or strain at the end of each sequence label. The sequences of this study are shown in bold font. Groups 1 and 2 in the Echinostomatidae are abbreviated as Gr1 and Gr2. The "revolutum" group in the *Echinostoma* clade is shown in a highlighted box, in which the Eurasian and American lineages of *Eca. revolutum* and the Australian and Eurasian lineages of *Eca. miyagawai* are indicated. The Echinochasmidae is highlighted. A scale bar represents the number of substitutions per site.

The topology of the Echinostomatidae in **Fig. 3.10** shows two distinct groups (as named by Izrailskaia et al. [147]. Group 1 contains species of ten genera (*Cathaemasia*, *Chaunocephalus*, *Drepanocephalus Euparyphium*, *Isthmiophora*, *Neopetasiger*, *Pegosomum*, *Petasiger*, *Rhopalias*, and *Ribeiroia*), and Group 2 contains species of seven genera (*Artyfechinostomum*, *Echinoparyphium*, *Echinostoma*, *Hypoderaeum*, *Moliniella*, *Neoacanthoparyphium*, and *Patagifer*). A very tight, monophyletic *Echinostoma* clade comprising 25 sequences of 12 species was recovered with 96% bootstrap support. Within this, the *Eca*. *revolutum* cluster consisted of six sequences and was divided into two lineages: the Eurasian lineage (one from Thailand (Erevo-MSD15-Thailand-OR509028), one from Iceland, and two from the Czech Republic), and the American lineage (two from the United States). Another tight cluster, *Eca*. *miyagawai*, of four sequences, also had two subclusters: the Eurasian lineage (one from Thailand (Emiya-RED11-Thailand-OR509027), one from the Czech Republic, and one from Ukraine), and the Australasian lineage (one from New Zealand).

The partial 28S phylogenetic tree (**Fig. 3.10**) also recovered all echinostomatid genera as monophyletic with (usually) high bootstrap support. A few points should be noted. *Hypoderaeum conoideum* sequences were all near-identical despite the samples from which they came being obtained from different geographical areas (Thailand, Finland, Ukraine, the Czech Republic, and the United States). This contrasts with the geographical distinctions among sequences from *Eca. revolutum*. Some genera also formed strongly supported groups. These include *Echinoparyphium* and *Hypoderaeum*, the former appearing as paraphyletic relative to the latter. Within Group 1 of Izrailskaia et al. (2021) [147], the genera *Ribeiroia*, *Cathaemasia* and *Rhopalias* were clustered with 97% bootstrap support. The remaining seven Group 1 genera formed a separate cluster with 82% bootstrap support.

CHAPTER 4

Discussion, Conclusions, Contributions and Future Prospects

It was possible to characterize the mitogenomic features of four echinostomes, namely *Echinostoma revolutum* (strain MSD15), *Artyfechinostomum malayanum* (former name: *Echinostoma malayanum*) (strain EMI3), *Echinostoma miyagawai* (strain RED11), and *Hypoderaeum conoideum* (strain RED42), with the addition of newly sequenced mtDNA data. Similarly, the complete sequences of the transcribed region or the entire rTU from the aforementioned echinostomatids (family: Echinostomatidae) and one echinochasmid, *Echinochasmus japonicus* strain EjPT (family: Echinochasmidae), as well as the ribosomal genomic features and structural characteristics of echinostomes, were determined. The mtDNA and rTU data usefulness resulted in a new phylogenetic framework for disentangling taxonomic relationships within and between the Echinostomatidae and the suborder Echinostomata of the class Trematoda.

4.1 The achievements of the mitogenomic investigations

4.1.1 Comparative mitogenomic datasets for the suborder Echinostomata

The complete mtDNA nucleotide sequences for four echinostomes, e.g., *Eca. revolutum*, *A. malayanum* (former name: *Eca. malayanum*), *Eca. miyagawai*, and *H. conoideum* with their realistic NCR length were fully determined and annotated. The strategy of combined long-range PCR (LPCR) and next-generation sequencing technology (NGS) was applied, for which, of most strains, long amplicons were obtained by the targeted multiplexed long-range LPCRs and successfully sequenced by long-read sequencing using the PacBio SEQUEL system. Echinostomes' mitogenomes are similar to those of other trematodes in length (with some exceptions) [25], gene order and composition, and in their tRNA and rRNA structure. These included: i) the complete mitogenome of *Echinostoma revolutum* (Fröhlich, 1802) Rudolphi, 1809, strain MSD15 (Thailand), which is totally 17,030 bp in length with the fully sequenced NCR of 3,549 bp (GenBank accession no. MN496162); ii) The complete mitogenome of *Artyfechinostomum malayanum* Leiper, 1911, strain EMI3 (Thailand) (former named

Echinostoma malayanum), which is totally 17,030 bp in length with the fully sequenced NCR of 3,622 bp (GenBank accession no. OK509083); iii) The complete mitogenome of *Echinostoma miyagawai* Ishii, 1932, strain RED11 (Thailand), which is totally 19,417 bp in length with the fully sequenced NCR of 5,935 bp (GenBank accession no. OP326312); and iv) The complete mitogenome of *Hypoderaeum conoideum* Dietz, 1909, strain RED42 (Thailand), which is totally 18,011 bp in length with the fully sequenced NCR of 4,475 bp (GenBank accession no. PP110501). These completely sequenced mtDNA datasets from echinostomes have contributed to the trematode mitodatabase for a variety of exploratory purposes.

It is noteworthy that the mitogenome of *Eca. miyagawai* (19,417 bp) appeared to be substantially the longest of all fully sequenced mtDNAs of the Echinostomatidae species ever examined to date, with an eception of *Echinostoma paraensei*/ (GenBank: KT008005), which was not fully sequenced at the estimated 20,298 bp in length (summarized in **Table 3.8**). The mitogenomes' lengths of four echinostomes studied here were much longer than that in the members of the Echinostomatidae previously reported, e.g., *Artyfechinostomum sufrartyfex* (14,567 bp, strain Shillong, India; GenBank: KY548763), *Eca. caproni* (14,150 bp, strain SAMEA, Egypt; GenBank: AP017706), *Eca. miyagawai* two Chinese strains (Hunan, 14,468 bp; and HLJ) 14,410 bp) [38, 91], *H. conoideum* (14,180 bp, strain Hubei, China) [92], *Echinoparyphium aconiatum* (14,865 bp, strain Chany, Russia) [142], and several echinostomatid species (in Echinostomatidae) to be reported to date. The mitogenome size of those echinostomid species previously sequenced might be not realistic due to the limitation of using conventional PCR and sequencing. The conventional PCR and sequencing might not cross the genomic complexity (rich in GC or At content) and usually skipped making the unrealistic length for their mtDNA molecules. As technology becoming progressed, the targeted multiplexed long-read sequencing (NGS) proved to be the most advanced and effective approach for achieving the realistic extent of the mtDNA for a trematode species as currently done in our study [9, 32, 49, 51, 53]).

All the four Echinostomatidae species in this study have the circular mtDNA molecule comprising 12 protein-coding genes (PCGs) (*cox*1−3, *cob*, *nad*1−6, *nad*4L, *atp*6), two mitoribosomal RNAs (MRGs) (16S or *rrn*L and 12S or *rrn*S), and 22 transfer RNAs (tRNAs or *trn*), and a non-coding region (NCR) rich in long and short tandem repeats with numbers variable by species. The mtDNA organization (or **gene order**) is similar as that of the majority of trematodes, and highly conserved in the closely related species, as following linearized map: 5'-*cox*3-H-*cob*-*nad*4L-*nad*4-QFM-*atp*6-*nad*2-VAD-*nad*1-NPIK-*nad*3-S_1W-*cox*1-T-*rrn*L-C-*rrn*S-*cox*2-*nad*6-$YL_1S_2L_2$R-*nad*5-G-E-NCR[LRUs#]-[SRUs]-3' (for mtDNAs of *Eca. revolutum*; *Eca. miyagawai*; and *H. conoideum*), and 5'-*cox*3-H-*cob*-*nad*4L-*nad*4-QFM-*atp*6-

*nad*2-VAD-*nad*1-NPIK-*nad*3-S_1W-*cox*1-T-*rrn*L-C-*rrn*S-*cox*2-*nad*6-$YL_1S_2L_2$R-*nad*5-G-NCR[LRUs#]-E-[SRUs]-3' (for *A. malayanum*). The coding region is conserved in the majority of trematodes' mtDNA, with the exception of the variable positions of $tRNA^{Gly}$ (G) and $tRNA^{Glu}$ (E).

As is also the case in other flatworms and nematodes and a few other metazoans, ***atp8* is missing** and the 3′ end of *nad*4L overlaps the 5′ end of *nad*4 by 40 bp in mtDNAs of echinostomes [25]. The **overlapped sequence** between *nad*4L and *nad*4 is usual in the mitogenomes of trematodes sequenced to date [25, 32, 35, 38, 91].

Special **DHU-arm missing** tRNAs for Serine were found for both $tRNA^{Ser1(AGN)}$ and $tRNA^{Ser2(UCN)}$. The special DHU-arm missing tRNAs, for Serine and/or Leucine, are common found in mtDNAs of trematodes [25] as seen in *Fasciola hepatica* [30], *Fascioloides jacksoni* [40], *Fascioloides magna* [37], *Paragonimus westermani* and *P. ohirai* [32, 34]. And this DHU-arm missing tRNA-structure is a typically characteristic for trematodes' mtDNA up-to-date sequenced.

Many pairs of mitogenes are separated by **short intergenic** (often <30 nt) sequences or non-coding regions. In addition, there are **one or two longer non-coding** region(s) (which are termed **NCR**) in every mitogenome, in which stable stem–loop structures associated with genome replication and/or repeat sequences might be found [25]. Repeat sequences have been reported in the NCR of many animal mitogenomes, supposedly as a consequence of slippage-mismatching mechanism [112]. Recently there has been a substantial increase in the interest in the non-coding regions (NCR) and their repetitive elements as has been studied in trematode species from the families Fasciolidae, Paragonimidae, Brachycladiidae, Diplostomidae, and Schistosomatidae [31, 51–54, 69, 70, 72, 140–142, 167]. In our study, the lengthy NCRs of four species' mtDNAs were successfully obtained and all possess two distinct types of tandem repeat units: long (LRUs) and short tandem repeat units (SRUs) which vary in length and numbers. In humans, the repetitive elements have been discovered as a result of the DNA "*slippage*" mechanism during replication and to be connected to certain genetic disorders, and the development of some cancerous diseases [112]. In trematodes, various (tandem) repeats of the ITS polymorphisms may have an impact on infectivity, geographical and climate distribution, evolutionary status and adaptation previously reported [168], and a role of regulatory function in our study [49].

Until this project started and during implementation, there were a **mitogenomic dataset** of 11 echinostome mitogenomes available, which were not fully sequenced or not well-annotated or with truncated and/or sequence-missing full length-NCRs. With our addition of four fully sequenced and annotated echinostome mitogenomes presented in this thesis, the

mitogenomic database has up to **15 sequences from 12 Echinostomatidae species**, particularly from our four medically important echinostomes. An updated summary of the mitogenomic datasets of all 15 strains of 12 species of the family Echinostomatidae, including four echinostomes from this study regarding their length, the non-coding regions (NCR), repeats, PCGs, and references is presented (**Table 3.8 in Chapter 3**). To date, there are only approximately 100 complete/nearly complete mtDNA data for about 65–70 trematode species (GOBASE: http://gobase.bcm.umontreal.ca/; GenBank: https://www.ncbi.nlm.nih.gov/genbank/) [169]; and GenBank), 57 of which were used for assess the mitophylogenetic and taxonomic relationships within Echinostomatidae and Echinostomata (see **Supplementary Table S2.4** and **Chapter 3**). The mtDNA data of trematodes obtained is still very limited, not covering all genera in a family and important families in the suborders in Trematoda, including the Echinostomatidae family of the suborder Echinostomata. The data we produced is very timely, new and valuable for further applied research of the Echinostomatidae family, as well as the Trematoda class, and is the major contribution to the phylum Platyhelminthes' mitogenomic database.

4.1.2 The comparative mitogenomic characteristics of echinostomes

A comparative description of **mitogenomic features** with other echinostome members of the Echinostomatidae was presented, with a special emphasis on the mitogenomic relationships of *Echinostoma* species of the "revolutum" group (*Eca. caproni*, *Eca. miyagawai*, *Eca. paraensei*, and *Eca. revolutum*). The **gene identity** analysis of *Eca. miyagawai* and the other 14 echinostome strains revealed that *nad*6 was the most divergent among all the echinostomes. The nucleotide difference for the individual genes and PCGs within the *Eca. miyagawai* strains (three strains for comparison in this study) was less than 2%, which is a commonly used number as a criterion for intraspecific divergence for species delimitation [141]. The genetic properties of the mitogenomic genes and genomes were found to be virtually same across within *Eca. miyagawai*, *Eca. revolutum*, and the "*revolutum*" species (*Eca. paraensei* and *Eca. caproni*). Their genetic proximity showed that they belonged to the classified membership of the "37-collar-spined *revolutum*" group, which was based on the morphological status of these valid species [12, 31, 136, 139].

Regarding **base composition** and **nucleotide usage**, invertebrate mitogenomes tend to be AT-rich [170], a feature also noted in the mtDNAs and protein-coding genes of several parasitic flatworms. However, nucleotide composition is not uniform among the species within a family, as seen in Echinostomatidae. Echinostomes, as many trematodes have a **clear "bias" towards using A+T** over G+C. Therefore, the skewness values of four echinostome species studied also tend to be negative for A+T and positive for G+C (**Table 3.5 in Chapter 3**). Base composition

presents on each strand and skewness is a measure of the asymmetry of a distribution of nucleotides (A, T, G, and C), which indicates the unequal representation of complementary bases on the same strand. The symmetry of nucleotide usage in both protein-coding genes and entire mitogenomes (including other gene-coding and non-coding regions) could be calculated using the GC- and AT-skew index [162].

In mtDNA sequences of all 15 echinostome strains, the base composition of A, T, G, and C, as well as the skewness values of AT and GC content for PCGs, MRGs, and mtDNA*, suggested that T was used more frequently than A and G than C. These formed **the pattern** of "T > G > A > C" as the basic base use, resulting in strongly negative values for AT-skew and highly positive GC-skew in the mtDNA of all *Echinostoma* and echinostomatid species. Such patterns of "T > G > A > C", negative AT-skew, and positive GC-skew were commonly observed in most mitogenomes in a range of the other trematodes up-to-date to be analyzed. These included *Paramphistomum cervi* [171], *Echinochasmus japonicus* [5], *Fascioloides jacksoni* [40], *Paragonimus iloktsuenensis*, *P. ohirai*, *P. skrjabini miyazakii*, and *P. westermani* [32, 35], and *Morishitium polonicum* [170]. However, the G and A level could be exchanged in the pattern in some trematode species. Skew is likely to be most pronounced at third codon positions, where any mutational change is synonymous and not subject to selection pressure [45, 46, 172]. The index is normally positive for the GC-skew, and negative for the AT-skew. In general, in vertebrates, where not all genes occur on the same strand, the GC-skew becomes lower with increasing purine content, and similarly, the AT-skew increases. In platyhelminths and nematodes, all genes are predicted to be transcribed from one strand, and so there is no reciprocal pressure of composition bias from those genes located on the anti-sense (complementary) strand. Therefore, the trematodes gained highly negative AT-skew and highly positive GC-skew [162, 172].

Obviously, there is a tight correlation between **codon usage** and skewness. The bias in nucleotides of codons strongly affects codon usage across all classes in platyhelminths [45, 46, 172]. As such, subsequently, in codon usage bias, the skewness of nucleotides often plays an important role, in which the negative or positive values of AT and GC skews resulted from the use of T over A and G over C, respectively. Note that, in platyhelminths, all PCGs are encoded on the same positive strand. The use of T > A and G > C, as well as the overall "T > G > A > C" pattern, strongly interferes the **codon bias** in the PCGs of all the trematodes up-to-date available in general and the echinostomes in particular in our study. It was observed that all the AT skew values were negative, indicating that T and A were used more frequently than G and C, and this nucleotide biased usage resulted in more codons with AT than codons with GC. Understandably, AT bias in PCGs of all echinostomes affects codon usage for building proteins,

such as codons for phenyalanine (TTT), for leucine (TTG), and for valine (GTT), which were most frequently used, while codons for arginine (CGC) were least commonly used. Our study found that all 15 strains of 12 echinostome species had a significant proportion of codons with 2 or 3 Ts, consisting of 24–25% codons (~10% TTT/Phe, ~8% TTG/Leu, and ~7% GTT/Val) in their PCGs, comprising up to 40% among 3,359–3,371 codons (**Table 3.6 in Chapter 3**).

Genetic distances (%) within and between sequences of taxa in the Echinostomatidae provide a good basis for comparison. Moreover, the use of nucleotide sequences of the PCGs is the most reasonable and reliable target for evaluation of the closely and distantly related strains and taxa in a genus and a family, which is encompassed by the interspecific and intraspecific distances [35]. The present study explored the **pairwise genetic distance among 15 strains of 12 echinostome species** in the family Echinostomatidae from 15 different geographical locations in seven countries (China, Egypt, India, Russia, Thailand, and the United States) on four continents (**Table 3.7 in Chapter 3**). There was no doubt that a low genetic distance (~11–13%) was noted among species of the "*revolutum*" group since morphologically they are members of "37-collar spined" echinostomes distinct from the other echinostomatid flukes [12, 13, 138]. Protein-coding genes differed by 21–22% between sequences of *Artyfechinostomum* and *Echinostoma* and were substantially higher than those between *Artyfechinostomum* and *Hypoderaeum*, or *Echinostoma* and *Hypoderaeum* (23–24%). As usual, the echinostomes exhibited an intraspecific rate with limited genetic distances (approximately 0.5–0.89%), a level less than 1% as reported among closely related variants within a species [141].

4.1.3 The polymorphism in the mtDNA's non-coding regions of the echinostomes

In trematodes the long NCR region and its function has not been explored in detail and only recently using long read sequencing approaches have these regions become not only accessible but quantifiable [51]. However, there is a deficit in detailed comparisons between echinostomatid species and little known about the variation in length as a consequence of NCR repeat elements of the mitogenomes within and between echinostome species.

The repetitive sequences in the NCR limit conventional sequencing to capture the realistic, whole size of the mitogenome [31, 32, 51, 53]. To overcome this, next-generation sequencing (NGS) has become more widely used, allowing for the collection of the complete mitogenomes. As a result, the successfully sequenced entire mitogenomes have been made, including a range of species in the families Fasciolidae, Paragonimidae, Brachycladiidae, Diplostomidae, Schistosomatidae, Echinostomatidae, and others [31, 32, 51–54, 69, 70, 84, 140–142]. In actuality, Pacific Biosystems ('PacBio')'s single-molecule real-time sequencing (SMRT) offers very long read lengths (long-read sequencing) with great accuracy, overcoming errors caused

by troublesome, repetitive genomic regions [115, 173]. For mitogenomics, targeted multiplex next-generation sequencing, in which a long-range PCR was performed for mtDNA enrichment and used for NGS library preparation and sequencing. This approach proved to be an advanced technology that was successfully used for NGS sequencing [115, 174, 175]. Long-range PCR-enriched NGS technology was implied for many species, including a number of the trematodes so far, such as the Indian *Fasciolopsis buski* [130], the Indian *Paragonimus westermani* [84], the Thai *Eca. revolutum* and *A. malayanum* [9, 31], and the Japanese *Paragonimus skrjabini miyazakii* [32].

The NCRs of mtDNAs in echinostomes not only exhibited structural **polymorphism** but in our study we found they have some regulatory functions. The **tandem repeat units** found in the NCR contained **promoter sequences** containing domains typical of initiation sites for replication and transcription as well as several palindromic regions which were shared between echinostomatid species [49]. The expansive repetitive non-coding regions, which were featured by a numbers of substantial short or long repeat units (LRUs and SRUs), that possess **typical promoter** sequences and **palindrome-embedded hairpin** structures (**Figs 3.3 and 3.4 in Chapter 3**). Palindromic sequences were also identified throughout both the LRU and the SRU, these are unique inverted repeats creating a hair pin structure acting as the recognition sites for DNA binding proteins involved in gene regulation [176]. The complicated structural features, such as the promoter and regulatory elements, as well as polymorphism in the mitochondrial non-coding regions of the Echinostomata suborder, particularly the Echinostomatidae family, and the newly discovered echinostomatid species within, for the first time, to be comprehensively investigated in our study [49]. However, the specific reason why they exist in one individual but not in the others, as well as the affective function of how they interact in the tandem repeat-haboring trematode taxa, remain unknown and need to be investigated.

Mitochondrial genomes are an evolutionary paradox, exhibiting a wide divergence. They also exhibit features not seen, or not as pronounced, in nuclear genomes. The mitochondria exist as autonomous entities with their own genomes, their own structure, and their own accessory elements, thus their own genomic characteristics [22, 172, 177]. Among these, are biases in base composition (nucleotide, skew, and codon) that must have an influence on the protein subunits for which they code.

4.1.4 The mitogenomic phylogeny implications for echinostomes

The mitogenomic datasets were used to determine and resolve the intergeneric and interfamilial phylogenetic relationships in trematodes [4, 25, 28, 63, 111, 125, 178], as applied within and between the Echinostomata and other suborders at the subordinal level. Accumulative data from mitogenomes, especially from trematodes and platyhelminths, have

been extensively employed in research on animal evolution, phylogeny, biogeography, systematics, species origins, population genetics, and related fields [22, 65].

Most studies on mitogenomes of parasitic flatworms have focused on phylogenetic questions and as such, for trematodes, the concatenated aminoacids of the PCGs frequently been used with the maximum likelihood method and the substitution model with the best score according to the Bayesian information criterion (JTT + F + G + I) [152]. The Jones-Taylor-Thornton (JTT + F + G + I) model is widely used substitution models for proteins, that are based on empirical amino acid interchange matrices estimated from databases of protein alignments that incorporate the average amino acid frequencies of the data set under examination, leading to more accurate phylogenetic tree [179]. With the addition of concatenated amino acids from the newly sequenced echinostome mtDNA data from this study for alignment, a revised phylogenetic framework was created to disentangle taxonomic relationships within and between the Echinostomatidae and other families and suborders.

Mitophylogenetic relationships were characterized by the alignment of **c**oncatenated amino-acid sequences of 12 PCGs of 57 strains of 41 trematode species of five families from the suborders Echinostomata (i.e., Echinostomatidae, Cyclocoelidae, Echinochasmidae, Fasciolidae, and Himasthlidae), two from the Opisthorchiata (Opisthorchiidae and Heterophyidae), and two from the suborder Xiphidiata (Paragonimidae and Dicrocoeliidae), and one from family Schistosomatidae (*Schistosoma haematobium* species) as an outgroup (Information of species/strains and families is given in **Supplementary Table S2.4**). It should be noted that the current phylogenetic analysis, to some extent of the available mtDNAs from the families and genera within, confirmed the **monophyly of Echinostomata** and Opisthorchiata as well as the paraphyly of the Xiphidiata with the monophyletic Paragonimidae family of the Troglotrematoidea superfamily and Dicrocoeliidae separately positioned (presented in **Fig. 3.6 in Chapter 3**).

The monophyletic Echinostomatidae family included two separate echinostomatid groups, one of which were *Echinostoma* spp. and *Artyfechinostomum* spp., including the nomenclaturally retaken *Eca./A. malayanum* [9, 31, 38, 49, 91, 93], while the other contains *H. conoideum* and *Echinoparyphium aconiatum*, and all of the other Echinostomatidae spp. recently dicovered [92, 142]. The current revised analysis clearly demonstrated that the Echinostomatidae is substantially monophyletic, in which the *Echinostoma* species are grouped together in a well-supported clade, while the non-*Echinostoma* and the other "cryptic" species appeared to be in a loose position, and were paraphyletic. Also, there are two distinct subclusters formed by unidentified species from the Chinese *Echinostoma* spp. and from the American Echinostomatidae spp., which may represent novel genera with new species in them (see Table

. The mitogenome phylogeny of the echinostomatid species revealed two major “sister” relationships. The first was between *Echinostoma* and *Artyfechinostomum* at the generic level and the second between Echinostomatidae and Faciolidae at the familial level, both of which are markedly different in the phylogenetic analysis based on 28S ribosomal markers [4, 63]. The “sister” relationship for the former was interfered with by the placement of *Neoacanthoparyphium*, and for the latter by the placement of the Caballerotrematidae n. fam. Unfortunately, there is no complete mtDNA sequence for *Neoacanthoparyphium* and *Caballerotrema* and for other related echinostomatid species, which could assist in resolution of the phylogenetic position of *Echinostoma* and *Artyfechinostomum* in Echinostomatidae and this family and others in the suborder Echinostomata.

The Opisthorchiata phylogenetic status was proposed paraphyletic on the analysis of partail ribosomal sequences [180], but it appeared that this suborder is at risk of becoming polyphyletic, because *Cryptocotyle lingua* (Heterophyidae) being topologically rejected out of the Heterophyidae family and being placed in the Opisthorchiidae. The assessment of a species' taxonomic hierarchy, particularly those of a generic border rank, must be based on morphological and molecular data, with both mtDNA and ribosomal marker-based investigations being considered.

The taxonomic relationships and phylogenetic placement of the families/superfamilies in Xiphidiata are always the most frequently debated issue and are constantly revised. The debate is not only about the families Paragonimidae and Troglotrematidae but also about the subordinal and superfamilial rank, to which these families belong: whether to the suborder Xiphidiata or be reclassified as a new Troglotremata, all in the order Plagiorchiida [63, 65, 73, 78, 146, 181]. Previous studies [32, 181] proposed the subordinal level for the superfamily Troglotrematoidea (i.e., the suborder Troglotremata), which was needed to be revisited. Later research demostrated that the Paragonimidae and Troglotrematidae families are obviously sister groups within the superfamily Troglotrematoidea and are still classified as members of the suborder Xiphidiata [73]. Although the Paragonimidae and Dicrocoeliidae families are monophyletic in our phylogeny (**Fig. 3.6 in Chapter 3**), their topological placement appears to be distinct, resulting in the Xiphidiata suborder being predominantly polyphyletic.

The Echinostoma is always monophyletic, the Xiphidiata are always polyphyletic based on mtDNA or ribosomal marker-based phylogenetic studies, and the Opisthorchiata is on the edge of transitioning from paraphyletic to polyphyletic [32, 35, 49, 73]. By incorporating more xiphidiatan, opisthorchiid, and echinostomatid species and multiple sequences from species in these families/superfamilies, the generic/familial boundaries and phyletic status of most genera/families in the Echinostomata, Opisthorchiata, and Xiphidiata have been better clarified.

Despite the monophyly being revealed, the Echinostomatidae is a vast family that requires further investigation into the taxonomy and intra- and interfamilial phylogenetic relationships of echinostomatid species within and between this family and other digenean families. Fully characterized mitogenomes of additional echinosomid species, particularly those in the major Echinostomatidae family, are necessary. The mitogenomic datasets obtained using the targeted-multiplexed long-read sequencing approach presented in this study will be useful for studies of taxonomic, evolutionary, and population genetics, and applicable to other taxa in the suborder Echinostomata and the class Trematoda.

4.2 The achievements of the ribosomal transcription unit investigations

4.2.1 Comparative rTU's datasets for the suborder Echinostomata

The complete or near-complete ribosomal transcription units (the transcribed region, from the 5' terminus of 18S to the 3' terminus of 28S rRNA genes, designated as rTU*) for five echinostomatid and echinochasmid species, respectively, including, *Echinostoma/Artyfechinostomum malayanum*, referred to as *A. malayanum* [9], *Echinostoma revolutum*, *Echinostoma miyagawai*, *Hypoderaeum conoideum*, and *Echinochasmus japonicus*, were obtained and annotated. The sequences obtained are the complete rTU of *Artyfechinostomum malayanum* strain EMI3, Thailand (9,499 bp), the near-complete rTU of *Hypoderaeum conoideum* strain RED42, Thailand (8,076 bp), and the transcribed regions (rTU*) of *Eca. revolutum* strain MSD15, Thailand (6,856 bp), *Eca. miyagawai* strain RED11, Thailand (6,854 bp), and *Ecs. japonicus* strain EjPT (Phu Tho), Vietnam (7,150 bp) (family Echinochasmidae) (presented in **Table 3.10 in Chapter 3**).

The transcribed region (rTU*) is almost equal in length (6,854–6,864 bp) in all echinostomatids that we sequenced (*A. malayanum*, *Eca. revolutum*, *Eca. miyagawai*, and *H. conoideum*) but a bit longer in an echinochasmid species (*Ecs. japonicus.* 7,150 bp). The 18S, 5.8S, and 28S rRNA genes were identical in length. While the ITS1 and ITS2 regions in all the echinostomatids (Echinostomatidae) did not vary much, the ITS2 region of *Ecs. japonicus* (Echinochasmidae) was much longer than in the Echinostomatidae (**Fig. 3.5**). The whole coding sequence of the rTU of around 6.8 kb is the same length as seen in *Isthmiophora hortensis* (synonym: *Echinostoma hortense*) [109] and most fasciolids. However, that of *Fasciolopsis buski* is longer due to repetitive sequences in its ITS1 region [72]. Interestingly, no repeats were present in either ITS1 or ITS2 regions of any of the echinostomatid and echinochasmids sequenced to date [5, 68, 71, 182]. The total length of a complete nucleotide sequence of rTU in trematodes, known to date, ranges between 7 and 10.3 kb, including the IGS/ETS region [6, 67–73]. The longest sequence of the complete rTU, to date in trematodes, is probably 10,221 bp found in a strain of *Paramphistomatum cervi* (family Paramphistomatidae, suborder

Pronocephalata) [68]. Long or short repetitive units with variable lengths and frequency found in the non-coding ribosomal regions, including the ITS, ETS, and IGS regions of the rTUs, cause size variations and polymorphic features for the majority of trematodes [75, 183]. While many trematode species, particularly those from the Opisthorchiidae and Paragonimidae families, have exhibited structural polymorphism in ITS-1 or ITS-2 [73, 79, 184], no repeat units have been found in the internal transcribed spacers (ITS-1 or ITS-2) of echinostomatids and echinochasmids (Echinostomatudae and Echinochasmidae) to date.

4.2.2 The ribosomal phylogeny implications of the suborder Echinostomata

4.2.2.1 Phylogenetic relationships above the level of suborder Echinostomata

Low levels of sequence conservation, the presence of variable numbers of repeats and intra-individual polymorphism means that the ITS regions are of little value for phylogenetic reconstruction at taxonomic levels above genus or family. Thus, the single or concatenated 18S and 28S sequences or, very often, the D1−D3 28S ribosomal rDNA region, are more frequently used for taxonomic and phylogenetic studies on trematodes and cestodes [4, 63, 64, 167, 185, 186]. Our studies used the nucleotide sequences of the concatenated complete 18S and 28S, the complete 28S alone, and partial 28S rRNA genes (D1−D3 sequences) in three different phylogenetic analyses. These analyses differed in numbers of sequences that could be included, according to their availability in GenBank. Comprehensive phylogenies were constructed to clarify the specific, generic, and familial interrelationships of the taxa within and between the Echinostomatoidea (Echinostomata). Information on other superfamilies and suborders (Opisthorchiata, Pronocephalata, Xiphidiata, Haplosplanchnata, and Diplostomata) was included where appropriate.

While the suborders Opisthorchiata, Echinostomata, Pronocephalata and Diplostomata, were clearly monophyletic in our trees with some exceptions, the suborder Xiphidiata was anomalous, being paraphyletic with respect to the Opisthorchiata. The anomaly was caused by the placement of the family Haploporidae. A similar anomaly was noted by Pérez-Ponce de León and Hernández-Mena [78], who proposed to resolve it by creating the new suborder Haploporata. Recognizing this suborder and removing relevant families from the Xiphidiata is also supported by our data (but see also Sokolov et al. [187]; Nguyen et al. [73]).

4.2.2.2 Phylogenetic relationships among and within echinostomatan taxa

Maximum likelihood (ML) phylogenetic trees of the echinostomes were reconstructed, which were based on the analysis of 60 concatenated 28S + 18S rDNA sequences and of 70 complete 28S sequences only (**Supplementary Table S2.5**).

Below the subordinal level, the reconstructed ML trees indicated a sister relationship between the Fasciolidae and Echinostomatidae and placed the Echinochasmidae basal within

the Echinostomata (**Figs. 3.7** and **3.8**)**.** And in between the two, there was the Philophthalmidae (in the 60 species/strain topology) or the Philophthalmidae and Cyclocoelidae (in the 70 species/strain topology) being placed with a very high bootstrap support (97% or 95%). The placement of Echinochasmidae distinct from the Echinostomatidae demonstrated their separate, independent familial status. Their full family rank was suggested by Tkach et al. (2016) [4] based on the partial 28S rDNA sequence analysis and updated by Le et al. (2016) [5] through the systematic analysis of complete mitochondrial genomes.

To increase the density of sequences in the Echinostomatidae and Echinochasmidae, as well as the broad generic and interfamilial relationships between families in the Echinostomata, a phylogeny based on the partial 28S D1−D3 sequence analysis (1.1–1.3 kb) was also reconstructed. Sequences of the D1−D3 region are "classical" markers for trematodes and many taxa are represented in the databases.

In a previous partial 28S-based phylogenetic study, Tkach et al. [4] used 86 sequences from 82 echinostomatoid species to revise the Echinostomatoidea and divided them into eight families, of which several have been revised, redefined or raised to familial rank. In our present study, the increased coverage of 28S sequences from the Echinostomatidae and the Echinochasmidae and from other families recognized by Tkach et al. [4] has resulted in some differences in the phylogenetic reconstruction. In our partial 28S rDNA phylogeny, the Cyclocoelidae and Philophthalmidae appear as basal groups (**Fig. 3.9**), but in the complete 28S gene analysis (**Fig. 3.8**), they are sisters in our present study, as also shown by Tkach et al. [4]. This inconsistency in placement may be explained by the influence of the additional multiple sequences for the Cyclocoelidae from the genera *Tracheophilus*, *Typhlocoelum*, *Morishitium*, and *Neohaematotrephus*, which were not available to Tkach et al. [4]. The Caballerotrematidae, with a single taxon in the tree, was confirmed as sister to the Echinostomatidae. The Echinochasmidae and Psilostomidae were recovered as sisters, as seen in Tkach et al. [4].

Within the family Echinostomatidae, Izrailskaia et al. [147] recovered two well-supported groups, each containing a number of genera. We also recovered these two groups: Group 1 of 10 genera containing 18 species and Group 2 of 8 genera comprising 24 species, including 12 *Echinostoma* species (**Fig. 3.9**). By incorporating more echinostomatid species and multiple sequences from some species in our present study, the generic boundaries and phyletic status of most genera in the Echinostomatidae have been better clarified. *Echinostoma* remains monophyletic with the incorporation of several recently described species, i.e., *Eca. maldonadoi* in Brazil [121] and *Eca. pseudorobustum* in the Americas [98], *Eca. novaezealandense* in New Zealand (Georgieva et al., 2017) [188], *Eca. bolschewense* in Russia [189], and *Eca. chankensis* in Russia [147]. The *Eca. revolutum* and *Eca. miyagawai* clusters,

represented by multiple geographical samples, strongly supported the division of Eurasian and American lineages for the former and Eurasian and Australian lineages for the latter species, as shown previously primarily using mitochondrial markers [12, 95, 123]**.** The genera *Artyfechinostomum* and *Patagifer* remain clearly monophyletic after the inclusion of *A. malayanum* from Thailand and more *P. bilobus* material from Vietnam (GenBank: OR532446) and from Mexico [120] in the current analysis. These two genera appear as sisters in a weakly supported clade in our **Fig. 3.9**.

The Echinochasmidae Odhner, 1910, formed a very tight clade with a very high bootstrap support in our **Fig. 3.9** and was a sister group to the Psilostomidae. This relationship between the two families was noted in Tkach et al. [4]. Until its elevation to family rank, the taxonomy of the former subfamily Echinochasminae Odhner, 1910, was somewhat chaotic based on morphology as well as mitochondrial or ribosomal markers [1, 4, 5, 11, 133, 145, 190]. With the additional sequences included here, along with multiple sequences for *Ecs. japonicus*, the phylogenetic situation makes it clearer that *Echinochasmus* is polyphyletic. One branch of *Echinochasmus* is paraphyletic with respect to *Microparyphium*, and the other is parapyletic with respect to *Stephanoprora* (**Fig. 3.9**). This situation was noted by Tkach et al. [4] and Tatonova et al. [17]. The former group contains species with 24 collar spines (*Ecs. coaxatus*, *Ecs. japonicus*, *Ecs. beleocephalus*, and *Ecs. perfoliatus*), while species in the latter group have 20 to 22 collar spines (*Ecs. mordax*, *Ecs. milvi*, and *Ecs. csuifunensis*) [4]. Within the 24-collar-spine group, sequences of *Ecs. japonicus*, all from northern Vietnam, appear in two distinct subclusters. The sequences from Nam Dinh Province were derived from cercariae, and adults were raised experimentally in domestic chickens [191], whereas those appearing closer to *Microparyphium* species in the tree were obtained from adult worms collected from human hosts [5]. The interesting phylogenetic relationships of the *Echinochasmus* species, particularly *Ecs. japonicus*, and the genetic subdivision of the Echinochasmidae require more species constituting *Echinochasmus* and need to be further investigated.

Past studies on the taxonomy of echinostomatoids using morphological features have created much confusion and spawned the creation of many synonyms and the continuing revision of taxa. The kinds of molecular data presented here and in other recent papers [4, 17, 98, 147, 121, 191] offer a way to resolve these problems.

4.3 An overall assessment of the thesis's contributions

Currently, substantial study is required to understand the formation and re-emergence of numerous parasite species, notably zoonotic trematodes that transmit diseases from animals to people (zoonosis, zoonoses). Genomic research discoveries are necessary to develop various precise techniques for species identification, as well as rapid, sensitive, inexpensive, and time-

saving diagnoses and treatment procedures [133, 192]. Foodborne trematodes are widely found in foods including fish, crabs, shrimp, snails, and frogs, which give protein nutrients to people, particularly in developing nations. Many zoonotic trematodes, particularly the intestinal echinostome flukes of the families Echinostomatidae and Echinochasmidae of the suborder Echinostomata (Trematoda: Platyhelminthes), have common intermediate hosts in freshwater aquatic species. If not correctly treated, infectious larvae (encysted metacercariae) found in fish, shrimp, crab, and other animals will be acquired by people and grow into hazardous parasites, having devastating results [1, 7, 122]. The Echinostomatidae family contains four genera: *Echinostoma*, *Hypoderaeum*, *Echinoparyphium*, and *Artyfechinostomum* (suborder Echinostomata; order Plagiorchiida; class Trematoda; phylum Platyhelminthes), all of which are pathogenic and have global epidemiological relevance. These trematode flukes from the Echinostomatidae family infect people by food ingestion (undercooked fish, crabs, shrimp, snails, mollusks, or amphibians), and at least 23 of them may infect humans, 15 of which cause significant public health concerns [1]. These include *Eca. revolutum*, *Eca. miyagawai*, *Eca. malayanum* (*A. malayanum*), and *H. conoideum* [2, 7, 11]. Echinostomiasis caused by *Echinostoma* spp. and other Echinostomatidae species in humans is a common infection in low-income communities all over the world, but it is more prevalent in Asian countries such as India, Indonesia, the Philippines, China, Malaysia, Singapore, Korea, Japan, Thailand, Myanmar, Laos, Cambodia and Vietnam. Tens of millions are anticipated to get infected, with hundreds of millions more at risk [1, 2].

Comparison of mitogenome (mtDNA) and ribosomal transcription unit (rTU) sequences can provide insights into genetic/genomic characteristics of trematodes, including echinostomes (in the family Echinostomatidae) in this study, and can also reveal numerous contentious concerns about evolutionary and phylogenetic relationships, nomenclature, and histories among species from important groupings, such as the suborder Echinostomata, the class Trematoda, and the phylum Platyhelminthes [25, 29]. The gene order and characteristics of mitogenomes and rTUs, in particular, may be used to determine inter- and intra-specific phylogenetic relatedness across and within species, genera, and families. Furthermore, differences in nucleotide and/or (amino acid) sequences between mitomolecules/rTUs from various populations within and across species are useful for investigating population structure, determining taxonomic status, and tracing evolutionary history.

The study findings have been incorporated into the thesis structure. The thesis focuses on the Echinostomatidae family and the genetics of mitochondrial and ribosomal transcription units in certain echinostome species. As a result, the research content has been completed, meeting two set goals on studying the sequencing and characterization of mitogenomes and

ribosomal transcribed regions of the rTUs of echinostomes of the genera *Echinostoma* (*Eca. revolutum* and *Eca. miyagawai*), *Artyfechinostomum* (*Eca. /Artyfechinostomum malayanum*), and *Hypoderaeum* (*H. conoideum*), including rTU of the species *Ecs. japonicus* (family Echinochasmidae). The findings gradually address the demands of study in species identification, genetics, cytology, phylogeny, evolution, and population genetics.

Gene/genome data for mtDNA and rTU are currently scarce in the family Echinostomatidae and the suborder Echinostomata. The thesis, thus, has built, completed, and provided the research results achieved, which have contributed to genetic and genomic research by contributing mtDNA and rTU data, as well as building phylogenetic systematics and determining taxonomy of Echinostomatidae species in particular and trematodes generally.

4.4 Conclusions

1. Obtaining complete or nearly **complete mtDNA and rTU equences** from several medically Important species and strains of the Echinostomatidae family, including *Echinostoma revolutum*, *Artyfechinostomum malayanum* (former name: *Echinostoma malayanum*), *Echinostoma miyagawai*, *Hypoderaeum conoideum*, and *Echinochasmus japonicus* (in case of rTU sequencing).

2. Providing a very **comprehensive dataset** that includes the (near) complete mitogenomes and rTUs of a number of species from major genera in the Echinostomatidae, as well as the features of the individual mitogenomes and rTUs that were compared and discussed in each chapter. These characteristics included the arrangement of genes and intergenic regions; base composition and pattern of nucleotide usage for construction of sequences and genes; pairwise gene identity comparisons among protein-encoding genes and genomes; genetic distance among echinostomes; elucidation of the (secondary) structures of ribosomal and transfer RNAs; and a description of some structural and promoter features of unassigned sequences, including non-coding regions and their tandem repeat units.

3. Reevaluating the **taxonomic rankings** and comprehensively resolved **phylogenetic relationships** of the echinostomatid species (the Echinostomatidae family), including the valid generic retake of the *Artyfechinostomum malayanum* species and the Echinostomatidae's significant monophyletic status in the topological relationships of the suborders Echinostomata, Opisthorchiata, and Xiphidiata.

The thesis highlights three **scientific contributions** in the family Echinostomatidae and suborder Echinostomata, including: i) Provision of mitogenomic and nuclear ribosomal unit datasets to the database of the Echinostomata and class Trematoda in phylum Platyhelminthes; ii) Characterization of the mitogenomes and nuclear ribosomal transcription units of the Echinostomata; iii) Implementation of the mitogenomic and nuclear ribosomal unit data for

inter-generic, -familial, and -subordinal phylogenetic and systematics studies of the suborder Echinostomata.

4.5 Future prospects

Because of the very high genetic diversity of strains/species in the echinostome complexes as well as the difficulty of distinguishing by morphology such as eggs and metacercariae, molecular data must be added to clarify taxonomic conditions and positions, species, genus, family, interfamily, and suborder relationships. In addition to morphological examination, accurate diagnosis and treatment of echinostomiasis/echinochasmiasis require the use of molecular approaches, such as genetic markers derived from the mitochondrial genome (mtDNA) and the ribosomal transcription unit (rTU). Taxonomic research based on species molecular markers is required for echinostomatid intestinal flukes (Echinostomatidae) to correctly recognize the nomenclature of each species/genus/family and suborder of Echinostomata, especially for the newly discovered "cryptic," "synonymous," "polymorphic," "sister," or "hybrid" species that have emerged in this vast Echinostomatidae family [4].

The mitogenomics and ribosomal genomics of transcription units from the research implementation in this thesis have contributed to the achievement of four complete mtDNA and five rTU sequences from medically important echinostomes, resulting in the availability of molecular mtDNA and rTU datasets for up to 15 strains of 12 species of the Echinostomatidae. The mtDNA sequences from four species (*Eca. revolutum*, *Eca. miyagawai*, *Artyfechinostomum malayanum*, and *Hypoderaeum conoideum*) in this study had a long length (17−19.5 kb), indicating the size of the intrinsic mtDNA molecules in these echinostome species.

The gene arrangement (gene order) in mtDNA and rTU of all echinostomes is identical to that of trematodes and is substantially conserved [6, 25]. However, all of the echinostomes analyzed appear to have a very long non-coding region of 5−7 kb, and this is a most interesting feature to be observed that has never been detected in the mtDNA of any echinostomatid species. This is due to the previously sequenced mtDNAs from various strains having a partially missing non-coding region, which prevented significant investigation in terms of mitogenomic and NCR structural and polymorphism characterisation. A comparison of the precise size and sequence of the non-coding regions in the mtDNA of different populations of *Echinostoma* or *Hypoderaeum* is required to substantiate the findings that length heterogeneity exists among the individual geographical isolates or is due to poor sequencing.

The successful quantity and quality mtDNAs from the study in our thesis were resulted from the obvious utility of targeted multiplexed long-read sequencing and that, the conventional sequencing technique might only generate a truncated NCR-possessing mtDNA sequence [25,

51, 53]. Similar work needs to be undertaken to determine if and how much length heterogeneity occurs in the population of echinostomes as well. In the future, targeted multiplexed long-read sequencing should be considered for use when sequencing the mtDNA of the expected "cryptic," "synonymous," "polymorphic," "sister," or "hybrid" species, as well as taxonomically debated, misranked, or novel congener "sibling" species for species identification, evolutionary and population genetics research.

The genetic code and pattern of codon usage in the mtDNA of all the echinostomes studied shared the common modified mitochondrial codes for platyhelminths and similar codon usage patterns [25, 45, 46], and nucleotide and amino acid identity of the 12 PCGs could be divided into 3 groups: highly conserved, less conserved and divergent. The *nad*6 gene appeared to be the most divergent and the *cox1* and the *cob* genes share the highest level of conservation among echinostomes.

Comparisons of mtDNA and rTU sequences can help to answer and to address fundamental taxonomic questions concerning closely related taxa and populations within the family Echinostomatidae as well as across taxa in the phylum Platyhelminthes. A principal aim of this thesis was to provide the core mtDNA and rTU sequence information for a number of echinostomatid species (Echinostomatidae family) for use by the scientific community at large for genetic comparison, phylogenetic and population genetics studies.

The availability of complete mtDNA and rTU sequences (provided in the thesis) allows for better resolution of relationships among Echinostoma and echinostomatid species, as well as the construction of a more robust molecular phylogeny for the family Echinostomatidae, followed by all taxa in related suborders such as monophyletic Echinostomata, Troglotremata, Opisthorchiata, and polyphyletic Xiphidiata. The data in this thesis have once again demonstrated the complexities of Echinostomatidae systematics. The subordinal/superfamilial classification provides significant systematic and taxonomic issues that will be addressed in the future by combining morphological and molecular analyses, as well as mitochondrial and ribosomal genomic data [6, 35, 69, 167, 193, 194]. The complete rTU, or at least the transcribed region, is a useful marker for analyzing the evolutionary relationships between paragonimid and other xiphidiatan taxa. It can also be employed to resolve interspecific, interfamilial, and intersubordinal relationships inside and across families, superfamilies, and suborders.

References

1. Chai JY (2019). Echinostomes. In Chai JY (ed.), Human Intestinal Flukes, from Discovery to Treatment and Control. Springer Netherlands, 2019; pp. 169–343. https://doi.org/10.1007/978-94-024-1704-3.
2. Toledo R, Esteban JG (2016). An update on human echinostomiasis. Trans R Soc Trop Med Hyg 110(1):37–45. http://doi:10.1093/trstmh/trv099.
3. Chai JY, Jung BK (2022). General overview of the current status of human foodborne trematodiasis. Parasitology 149:1262–1285. https://doi.org/10.1017/S0031182022000725.
4. Tkach VV, Kudlai O, Kostadinova A (2016). Molecular phylogeny and systematics of the Echinostomatoidea Looss, 1899 (Platyhelminthes: Digenea). Int J Parasitol 46(3):171–185. http://doi:10.1016/j.ijpara.2015.11.001.
5. Le TH, Nguyen NTB, Nguyen KT, Doan HTT, Dung DT, Blair D (2016). A complete mitochondrial genome from *Echinochasmus japonicus* supports the elevation of Echinochasminae Odhner, 1910 to family rank (Trematoda: Platyhelminthes). Infect Genet Evol 45:369–377. http://doi:10.1016/j.meegid.2016.09.024.
6. Le TH, Pham LTK, Quyen DV, Nguyen KT, Doan HTT, Saijuntha W, Blair D (2024). The ribosomal transcription units of five echinostomes and their taxonomic implications for the suborder Echinostomata (Trematoda: Platyhelminthes). Parasitol Res 123(1):103. https://doi.org/10.1007/s00436-023-08110-z
7. Toledo R, Álvarez-Izquierdo M, Esteban JG, Muñoz-Antoli C (2022). Neglected food-borne trematodiases: echinostomiasis and gastrodiscoidiasis. Parasitology 149(10):1319–1326. https://doi.org/10.1017/S0031182022000385.
8. Prasad YK, Dahal S, Saikia B, Bordoloi B, Tandon V, Ghatani S (2019) *Artyfechinostomum sufrartyfex* Trematode Infections in Children, Bihar, India. Emerg Infect Dis 25(8):1571–1573. http://doi:10.3201/eid2508.181427.
9. Pham KLT, Saijuntha W, Lawton SP, Le TH (2022). Mitophylogenomics of the zoonotic fluke *Echinostoma malayanum* confirms it as a member of the genus *Artyfechinostomum* Lane, 1915 and illustrates the complexity of Echinostomatidae systematics. Parasitol Res 121:899–913. https://doi.org/10.1007/s00436-022-07449-z.
10. Sohn WM, Yong TS, Eom KS, Sinuon M, Jeoung WG, Chai JY (2017). *Artyfechinostomum malayanum*: metacercariae encysted in *Pila* sp. snails purchased from Phnom Penh, Cambodia. Korean J Parasitol 55:341–345. http://doi:10.3347/kjp.2017.55.3.341.
11. Kostadinova A (2005). Family Echinostomatidae Looss, 1899. In: Jones A, Bray RA, Gibson DI (eds) Keys to the Trematoda, vol 2. CAB International, Wallingford & Natural History Museum, London, UK, pp 9–64.

12. 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. http://doi:10.1186/s13071-014-0520-8.
13. Faltýnková A, Georgieva S, Soldánová M, Kostadinova A (2015). A reassessment of species diversity within the 'revolutum' group of *Echinostoma* Rudolphi, 1809 (Digenea: Echinostomatidae) in Europe. Syst Parasitol 90:1–25. https://doi.org/10.1007/s11230-014-9530-3.
14. Saijuntha W, Sithithaworn P, Andrews RH (2010a). Genetic differentiation of *Echinostoma revolutum* and *Hypodereaum conoideum* from domestic ducks in Thailand by multilocus enzyme electrophoresis. J Helminthol 84(2):143–148. http://doi:10.1017/S0022149X09990393.
15. Saijuntha W, Tapdara S, Tantrawatpan C (2010b). Multilocus enzyme electrophoresis analysis of *Echinostoma revolutum* and *Echinostoma malayanum* (Trematoda: Echinostomatidae) isolated from Khon Kaen Province, Thailand. Asian Pac J Trop Med 3:633–636. http://doi:10.1016/S1995-7645(10)60153-8
16. Saijuntha W, Tantrawatpan C, Sithithaworn P, Andrews RH, 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. Parasitol Res 108:751–755. https://doi.org/10.1007/s00436-010-2180-8.
17. Tatonova YV, Izrailskaia AV, Besprozvannykh VV (2020). *Stephanoprora amurensis* sp. nov., *Echinochasmus milvi* Yamaguti, 1939 and *E. suifunensis* Besprozvannykh, 1991 from the Russian Southern Far East and their phylogenetic relationships within the Echinochasmidae Odhner 1910. Parasitology 147(13):1469–1479. https://doi.org/10.1017/S0031182020001444.
18. Saijuntha W, Sithithaworn P, Duenngai K, Kiatsopit N, Andrews RH, Petney TN (2011a). Genetic variation and relationships of four species of medically important echinostomes (Trematoda: Echinostomatidae) in South-East Asia. Infect Genet Evol 11:375–381. https://doi.org/10.1016/j.meegid.2010.11.009.
19. Saijuntha W, Tantrawatpan C, Sithithaworn P, Andrews RH, Petney TN (2011c). Spatial and temporal genetic variation of *Echinostoma revolutum* (Trematoda: Echinostomatidae) from Thailand and the Lao PDR. Acta Trop 118:105–109. https://doi.org/10.1016/j.actatropica.2011.02.014.
20. Boore JL (1999). Animal mitochondrial genomes. Nucleic Acids Res 27:1767–1780. https://doi.org/10.1093/nar/27.8.1767.

21. Friedman JR, Nunnari J (2014). Mitochondrial form and function. Nature 505(7483):335–343. https://doi.org/10.1038/nature12985.
22. Bernt M, Braband A, Schierwater B, Stadler PF (2013). Genetic aspects of mitochondrial genome evolution. Mol Phylogenet Evol 69(2):328–338. https://doi.org/10.1016/j.ympev.2012.10.020.
23. Tang JX, Thompson K, Taylor RW, Oláhová M (2020). Mitochondrial OXPHOS Biogenesis: Co-Regulation of Protein Synthesis, Import, and Assembly Pathways. Int J Mol Sci 21:3820. https://doi.org/10.3390/ijms21113820.
24. Reichert A, Rothbauer U, Mörl M (1998). Processing and editing of overlapping tRNAs in human mitochondria. J Biol Chem 273(48):31977–31984. https://doi.org/10.1074/jbc.273.48.31977.
25. Le TH, Blair D, McManus DP (2002). Mitochondrial genomes of parasitic flatworms. Trends Parasitol 18:206–213. https://doi.org/10.1016/s1471-4922(02)02252-3.
26. Le TH (2001). "*Mitochondrial genomics of platyhelminths*". PhD thesis. The University of Queensland, Brisbane, Australia, 2001 (261 pp). https://doi.org/10.14264/93ff419.
27. Butenko A, Lukeš J, Speijer D, Wideman JG (2024). Mitochondrial genomes revisited: why do different lineages retain different genes? BMC Biol 22(1):15. https://doi.org/10.1186/s12915-024-01824-1.
28. Le TH, Blair D, Agatsuma T, Humair PF, Campbell NJH, Iwagami M, Littlewood DTJ, Peacock B, Johnston DA, Bartley J, Rollinson D, Herniou EA, Zarlenga DS and McManus DP (2000). Phylogenies inferred from mitochondrial gene orders - a cautionary tale from the parasitic flatworms. Mol Biol Evol 17(7):1123–1125. https://doi.org/10.1093/oxfordjournals.molbev.a026393.
29. Hu M, Gasser RB (2006). Mitochondrial genomes of parasitic nematodes - progress and perspectives. Trends Parasitol 22(2):78–84. https://doi.org/10.1016/j.pt.2005.12.003.
30. Le TH, Blair D, 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. https://doi.org/10.1017/s0031182001008733.
31. Le TH, Pham LTK, Doan HTT, Le XTK, Saijuntha W, Rajapakse RPVJ, Lawton SP (2020b). Comparative mitogenomics of the zoonotic parasite *Echinostoma revolutum* resolves taxonomic relationships within the '*E. revolutum*' species group and the Echinostomata (Platyhelminthes: Digenea). Parasitology 147(5):566–576. https://doi.org/10.1017/S0031182020000128.
32. Le TH, Nguyen KT, Pham LTK, Doan HTT, Agatsuma T, Blair D (2022). The complete mitogenome of the Asian lung fluke *Paragonimus skrjabini miyazakii* and its implications for the family Paragonimidae (Trematoda: Platyhelminthes). Parasitology 149(13):1709–1719. https://doi.org/10.1017/S0031182022001184.

33. Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Agatsuma T, Blair D (2019a). The complete mitochondrial genome of *Paragonimus ohirai* (Paragonimidae: Trematoda: Platyhelminthes) and its comparison with *P. westermani* congeners and other trematodes. Peer J 7:e7031. http://doi:10.7717/peerj.7031.
34. Le TH, Nguyen TBN, Doan TTH, Le TKX, Nguyen TK (2019b). 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 (Abstract in English).
35. Le TH, Nguyen KT, Pham LTK, Doan HTT, Do RT, Le XTK, Agatsuma T, Blair D (2023). Mitogenomic and nuclear ribosomal transcription unit datasets support the synonymy of *Paragonimus iloktsuenensis* and *P. ohirai* (Paragonimidae: Platyhelminthes). Parasitol Res 122(7):1531–1544. https://doi.org/10.1007/s00436-023-07854-y
36. Liu GH, Gasser RB, Young ND, Song HQ, Ai L, Zhu XQ (2014). Complete mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola gigantica*, and their comparison with *F. hepatica*. Parasit Vectors 7:150. http://doi:10.1186/1756-3305-7-150.
37. Ma J, He JJ, Liu GH, Leontovyč R, Kašný M, Zhu XQ (2016). Complete mitochondrial genome of the giant liver fluke *Fascioloides magna* (Digenea: Fasciolidae) and its comparison with selected trematodes. Parasit Vectors 9:429. http://doi:10.1186/s13071-016-1699-7.
38. Fu YT, Jin YC, Li F, Liu GH (2019). Characterization of the complete mitochondrial genome of the echinostome *Echinostoma miyagawai* and phylogenetic implications. Parasitol Res 118(10):3091–3097. http://doi:10.1007/s00436-019-06417-4.
39. Li Y, Ma XX, Lv QB, Hu Y, Qiu HY, Chang QC, Wang CR (2019a). Characterization of the complete mitochondrial genome sequence of *Tracheophilus cymbius* (Digenea), the first representative from the family Cyclocoelidae. J Helminthol 94:e101. http://doi:10.1017/S0022149X19000932.
40. Rajapakse RPVJ, Pham KLT, Karunathilake KJK, Lawton SP, Le TH (2020). Characterization and phylogenetic properties of the complete mitochondrial genome of *Fascioloides jacksoni* (syn. *Fasciola jacksoni*) support the suggested intergeneric change from *Fasciola* to *Fascioloides* (Platyhelminthes: Trematoda: Plagiorchiida). Infect Genet Evol 82:104281. http://doi:10.1016/j.meegid.2020.104281.
41. Berg MD, Brandl CJ (2021). Transfer RNAs: diversity in form and function. RNA Biol 18(3):316–339. https://doi.org/10.1080/15476286.2020.1809197.
42. Jaskolowski M, Ramrath DJF, Bieri P, Niemann M, Mattei S, Calderaro S, Leibundgut M, Horn EK, Boehringer D, Schneider A, Ban N (2020). Structural Insights into the

Mechanism of Mitoribosomal Large Subunit Biogenesis. Mol Cell 79(4):629–644.e4. https://doi.org/10.1016/j.molcel.2020.06.030.

43. Greber BJ, Ban N (2016). Structure and Function of the Mitochondrial Ribosome. Annu Rev Biochem 85:103–132. https://doi.org/10.1146/annurev-biochem-060815-014343.
44. Lopez Sanchez MIG, Krüger A, Shiriaev DI, Liu Y, Rorbach J (2021). Human Mitoribosome Biogenesis and Its Emerging Links to Disease. Int J Mol Sci 22:3827. https://doi.org/10.3390/ijms22083827.
45. Telford MJ, Herniou EA, Russell RB, Littlewood DT (2000). Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. Proc Natl Acad Sci U S A 97(21):11359–11364. https://doi.org/10.1073/pnas.97.21.11359.
46. Le TH, McManus DP, Blair D (2004). Codon usage and bias in mitochondrial genomes of parasitic platyhelminthes. Korean J Parasitol 42(4):159–167. https://doi.org/10.3347/kjp.2004.42.4.159.
47. Logan DC (2006). The mitochondrial compartment. J Exp Bot 57(6):1225–1243. https://doi.org/10.1093/jxb/erj151.
48. Gemayel R, Vinces MD, Legendre M, Verstrepen KJ (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet 44:445–477. https://doi.org/10.1146/annurev-genet-072610-155046.
49. Pham LTK, Quyen DV, Saijuntha W, Doan HTT, Le TH, Scott P. Lawton (2024). Mitogenomics of the zoonotic parasite *Echinostoma miyagawai* and insights into the evolution of tandem repeat regions within the mitochondrial non-coding control region. Parasitology 14:1–41. https://doi.org/10.1017/S0031182024001422.
50. Anderson S, Bankier AT, Barrell BG, De-Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nlerlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981). Sequence and organization of the human mitochondrial genome. Nature 290:457–465. https://doi.org/10.1038/290457a0.
51. Kinkar L, Young ND, Sohn WM, Stroehlein AJ, Korhonen PK, Gasser RB (2020). First record of a tandem-repeat region within the mitochondrial genome of *Clonorchis sinensis* using a long-read sequencing approach. PLOS Negl Trop Dis 14(8):e0008552. https://doi.org/10.1371/journal.pntd.0008552.
52. Oey H, Zakrzewski M, Narain K, Devi KR, Agatsuma T, Nawaratna S, Gobert GN, Jones MK, Ragan MA, McManus DP, Krause L (2019b). Whole-genome sequence of the oriental lung fluke *Paragonimus westermani*. Gigascience 8(1):giy146. https://doi.org/10.1093/gigascience/giy146.
53. Kinkar L, Korhonen PK, Cai H, Gauci CG, Lightowlers MW, Saarma U, Jenkins DJ, Li J, Li J, Young ND, Gasser RB (2019). Long-read sequencing reveals a 4.4 kb tandem

repeat region in the mitogenome of *Echinococcus granulosus* (sensu stricto) genotype G1. Parasit Vectors 12(1):238. http://doi:10.1186/s13071-019-3492-x.

54. Oey H, Zakrzewski M, Gravermann K, Young ND, Korhonen PK, Gobert GN, et al. (2019a). Whole-genome sequence of the bovine blood fluke *Schistosoma bovis* supports interspecific hybridization with *S. haematobium*. PLoS Pathog 15(1):e1007513. https://doi.org/10.1371/journal.ppat.1007513.
55. Zhang DX, Hewitt GM (1997). Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol 25:99–120. https://doi.org/10.1016/S0305-1978(96)00042-7
56. D'Souza AR, Minczuk M (2018). Mitochondrial transcription and translation: overview. Essays in Biochemistry 62:309–320. https://doi.org/10.1042/EBC20170102.
57. Tschochner H, Hurt E (2003). Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13(5):255–263. https://doi.org/10.1016/s0962-8924(03)00054-0.
58. Potapova TA, Gerton JL (2019). Ribosomal DNA and the nucleolus in the context of genome organization. Chromosome Res 27(1-2):109–127. https://doi.org/10.1007/s10577-018-9600-5.
59. Eickbush TH, Eickbush DG (2007). Finely orchestrated movements: evolution of the ribosomal RNA genes. Genetics 175(2):477–485. https://doi.org/10.1534/genetics.107.071399.
60. McStay B (2016). Nucleolar organizer regions: genomic 'dark matter' requiring illumination. Genes Dev 30(14):1598–1610. https://doi.org/10.1101/gad.283838.116.
61. 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.
62. Turowski T, Tollerve D (2015). Cotranscriptional events in eukaryotic ribosome synthesis. WIREs RNA 6:129–139. https://doi.org/10.1002/wrna.1263.
63. Olson PD, Cribb TH, Tkach VV, Bray RA, Littlewood DT (2003). Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int J Parasitol 33(7):733–755. https://doi.org/10.1016/s0020-7519(03)00049-3.
64. Lockyer AE, Olson PD, Littlewood DTJ (2003). Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. Biol J Linn Soc Lond 78(2):155–171. https://doi.org/10.1046/j.1095-8312.2003.00141.x
65. Littlewood DT J (2008). Platyhelminth systematics and the emergence of new characters. Parasite 15(3):333–341. https://doi.org/10.1051/parasite/2008153333.

66. Salim D, Gerton JL (2019). Ribosomal DNA instability and genome adaptability. Chromosom Res 27: 73–87. https://doi.org/10.1007/s10577-018-9599-7.
67. Qiu YY, Gao Y, Li Y, Ma XX, Lv QB, Hu Y, Qiu HY, Chang QC and Wang CR (2019). Comparative analyses of complete ribosomal DNA sequences of *Clonorchis sinensis* and *Metorchis orientalis*: IGS sequences may provide a novel genetic marker for intraspecific variation. Infect Genet Evol 78:104–125. https://doi.org/10.1016/j.meegid.2019.104125.
68. Zheng X, Chang QC, Zhang Y, Tian SQ, Lou Y, Duan H, Guo DH, Wang CR, Zhou XQ (2014). Characterization of the complete nuclear ribosomal DNA sequences of Paramphistomum cervi. Sci World J 2014:751907. https://doi.org/10.1155/2014/751907.
69. Briscoe AG, Bray RA, Brabec J, Littlewood DT (2016). The mitochondrial genome and ribosomal operon of *Brachycladium goliath* (Digenea: Brachycladiidae) recovered from a stranded minke whale. Parasitol Int 65(3):271–275. https://doi.org/10.1016/j.parint.2016.02.004.
70. Brabec J, Kostadinova A, Scholz T, Littlewood DT (2015). Complete mitochondrial genomes and nuclear ribosomal RNA operons of two species of *Diplostomum* (Platyhelminthes: Trematoda): a molecular resource for taxonomy and molecular epidemiology of important fish pathogens. Parasit Vectors 8:336. https://doi.org/10.1186/s13071-015-0949-4.
71. 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. https://doi.org/10.1017/S0022149X17000554.
72. Le TH, Pham KLT, Doan HTT, Le TKX, Nguyen KT, Lawton SP (2020a). Description and phylogenetic analyses of ribosomal transcription units from species of Fasciolidae (Platyhelminthes: Digenea). J Helminthol 94:e136. https://doi.org/10.1017/S0022149X20000164.
73. Nguyen KT, Doan HTT, Pham TKL, Do RT, Takeshi Agatsuma T, Doanh PN, Le TH (2024). Nuclear ribosomal transcription units in Asian *Paragonimus* species (Paragonimidae: Platyhelminthes): structure, polymorphism, and implications for intersubordal phylogeny. Parasitol Res 123(11):368. https://doi.org/10.1007/s00436-024-08391-y.
74. Voronova A, Chelomina GN (2018). Genetic diversity and phylogenetic relations of salmon trematode *Nanophyetus japonensis*. Parasitol Int 67(3):267–276. https://doi.org/10.1016/j.parint.2018.01.002.
75. 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. https://doi.org/10.1146/annurev.ecolsys.36.102003.152620

76. Nguyen TBN, De NV, Nguyen TKL, Quang HH, Doan HTT, Agatsuma T, Le TH (2018). Distribution of hybrid types in large liver flukes, *Fasciola* species (Digenea: Fasciolidae), from ruminants and humans in Vietnam. Korean J Parasitol 56(5):453–461. https://doi.org/10.3347/kjp.2018.56.5.453
77. Symonová R (2019). Integrative rDNAomics – Importance of the Oldest Repetitive Fraction of the Eukaryote Genome. Genes 10(5):345. https://doi.org/10.3390/genes10050345.
78. Pérez-Ponce de León G, Hernández-Mena DI (2019). Testing the higher-level phylogenetic classification of Digenea (Platyhelminthes, Trematoda) based on nuclear rDNA sequences before entering the age of the 'next-generation' Tree of Life. J Helminthol 93(3):260–276. https://doi.org/10.1017/S0022149X19000191.
79. 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. https://doi.org/10.1186/s13071-017-1968-0.
80. Michot B, Despres L, Bonhomme F, Bachellerie JP (1993). Conserved secondary structures in the ITS2 of trematode pre-rRNA. FEBS Lett 316(3):247–252. https://doi.org/10.1016/0014-5793(93)81301-f.
81. Schultz J, Maisel S, Gerlach D, Müller T, Wolf M (2005). A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA 11(4):361–364. https://doi.org/10.1261/rna.7204505.
82. 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. https://doi.org/10.1101/gr.3970105.
83. McCormick EM, Muraresku CC, Falk MJ (2018). Mitochondrial Genomics: A complex field now coming of age. Curr Genet Med Rep 6(2):52–61. https://doi.org/10.1007/s40142-018-0137-x.
84. Biswal DK, Chatterjee A, Bhattacharya A, 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. https://doi.org/10.7717/peerj.484.
85. 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. https://doi.org/10.1080/23802359.2018.1462119.
86. Wang T, Wang Y, Xu F, Li X, Qu R, Song L, Tang Y, Lin P (2018). Characterization of the complete mitochondrial genome of the lung fluke, *Paragonimus kellicotti*.

Mitochondrial DNA Part B 3(2):715–716. https://doi.org/10.1080/23802359.2018.1483763.

87. Lee D, Choe S, Park H, Jeon HK, Chai JY, Sohn WM, Yong TS, Min DY, Rim HJ, Eom KS (2013). Complete mitochondrial genome of *Haplorchis taichui* and comparative analysis with other trematodes. Korean J Parasitol 51(6):719–726. http://doi:10.3347/kjp.2013.51.6.719.
88. Shekhovtsov SV, Katokhin AV, Kolchanov NA, Mordvinov VA (2010). The complete mitochondrial genomes of the liver flukes *Opisthorchis felineus* and *Clonorchis sinensis* (Trematoda). Parasitol Int 59:100–103. http://doi:10.1016/j.parint.2009.10.012
89. Cai XQ, Liu GH, Song HQ, Wu CY, Zou FC, Yan HK, Yuan ZG, Lin RQ, Zhu XQ (2012). Sequences and gene organization of the mitochondrial genomes of the liver flukes *Opisthorchis viverrini* and *Clonorchis sinensis* (Trematoda). Parasitol Res110(1):235–243. http://doi:10.1007/s00436-011-2477-2.
90. Suleman, Khan MS, Heneberg P, Zhou CY, Muhammad N, Zhu XQ, Ma J (2019). Characterization of the complete mitochondrial genome of *Uvitellina* sp., representative of the family Cyclocoelidae and phylogenetic implications. Parasitol Res 118:2203–2211. https://doi.org/10.1007/s00436-019-06358-y.
91. Li Y, Qiu YY, Zeng MH, Diao PW, Chang QC, Gao Y, Zhang Y, Wang CR (2019b). The complete mitochondrial genome of *Echinostoma miyagawai*: Comparisons with closely related species and phylogenetic implications. Infect Genet Evol 75:103961. http://doi:10.1016/j.meegid.2019.103961.
92. Yang X, Gasser RB, Koehler AV, 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. https://doi.org/10.1186/s13071-015-0720-x.
93. Ran R, Zhao Q, Abuzeid AMI, Huang Y, Liu Y, Sun Y, He L, Li X, Liu J, Li G (2020). Mitochondrial genome sequence of *Echinostoma revolutum* from Red-Crowned Crane (*Grus japonensis*). Korean J Parasitol 58(1):73–79. https://doi.org/10.3347/kjp.2020.58.1.73.
94. Liu ZX, Zhang Y, Liu YT, Chang QC, Su X, Fu X, Yue DM, Gao Y, Wang CR (2016). Complete Mitochondrial Genome of *Echinostoma hortense* (Digenea: Echinostomatidae). Korean J Parasitol 54(2):173–179. https://doi.org/10.3347/kjp.2016.54.2.173.
95. Georgieva S, Selbach C, Faltýnková A, Soldánová M, Sures B, Skírnisson K, Kostadinova A (2013). New cryptic species of the '*revolutum*' group of *Echinostoma* (Digenea: Echinostomatidae) revealed by molecular and morphological data. Parasit Vectors 6:64. http://doi:10.1186/1756-3305-6-64.

96. Cho J, Jung BK, Chang T, Sohn WM, Sinuon M, Chai JY (2020). *Echinostoma mekongi* n. sp. (Digenea: Echinostomatidae) from riparian people along the Mekong River in Cambodia. Korean J Parasitol 58:431–443. https://doi.org/10.3347/kjp.2020.58.4.431.
97. Tatonova YV, Izrailskaia AV, Besprozvannykh VV (21). *Stephanoprora amurensis* sp. nov., *Echinochasmus milvi* Yamaguti, 1939 and *E. suifunensis* Besprozvannykh, 1991 from the Russian Southern Far East and their phylogenetic relationships within the Echinochasmidae Odhner 1910. Parasitology 147(13):1469–1479. https://doi.org/10.1017/S0031182020001444.
98. Valadão MC, López-Hernández D, Alves PV, Pinto HA (2022). A new species of *Echinostoma* (Trematoda: Echinostomatidae) from the '*revolutum*' group found in Brazil: refuting the occurrence of *Echinostoma miyagawai* (=*E. robustum*) in the Americas. Parasitology 149(3):325–336. https://doi.org/10.1017/S0031182021001864.
99. Hong S, Shin H, Lee YH, Hong SJ, Kim SR, Kim YK, Son YJ, Song JG, Chai JY, Jung BK (2024). Rare Case of *Echinostoma cinetorchis* Infection, South Korea. Emerg Infect Dis 30(8):1726–1729. https://doi.org/10.3201/eid3008.240289.
100. Cerqueira AV, Lemos B (2019). Ribosomal DNA and the Nucleolus as Keystones of Nuclear Architecture, Organization, and Function. Trends Genet 35(10):710–723. https://doi.org/10.1016/j.tig.2019.07.011.
101. Waeschenbach A, Webster BL, Bray RA, Littlewood DTJ (2007). Added resolution among ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. Mol Phylogenet Evol 45(1):311–325. https://doi.org/10.1016/j.ympev.2007.03.019.
102. Heneberg P, Literák I (2013). Molecular phylogenetic characterization of *Collyriclum faba* with reference to its three host-specific ecotypes. Parasitol Int 62(3):262–267. https://doi.org/10.1016/j.parint.2013.01.002.
103. Tolstenkov O, Chatzigeorgiou M, Gorbushin A (2023). Neuronal gene expression in two generations of the marine parasitic worm, *Cryptocotyle lingua*. Commun Biol 6(1):1279. https://doi.org/10.1038/s42003-023-05675-4.
104. Le TVH, Nguyen TK, Dong VQ, Le TH (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.
105. Sato H, Ihara S, Inaba O, Une Y (2010). Identification of *Euryhelmis costaricensis* metacercariae in the skin of Tohoku hynobiid salamanders (*Hynobius lichenatus*), northeastern Honshu, Japan. J Wildl Dis 46(3):832–842. https://doi.org/10.7589/0090-3558-46.3.832.

106. Locke SA, Calhoun DM, Valencia Cruz JM, Ebbs ET, Díaz Pernett SC, Tkach VV, Kinsella JM, Freeman MA, Blanar CA, Johnson PTJ (2024). Expanding on expansus: A new species of *Scaphanocephalus* from North America and the Caribbean based on molecular and morphological data. Parasitology 21:1–51. https://doi.org/10.1017/S0031182024000647.
107. Mathews PD, Rabet N, L Espinoza L, Haÿ V, Bonillo C, Keith P, Lord C, Audebert F (2023). Discovery of a Digenean (Cryptogonimidae) Living in a Cleft-Lipped Goby, *Sicyopterus cynocephalus* (Teleostei: Gobiidae) from Ranongga Island, Solomon Islands: Analysis of Multiple Ribosomal DNA Regions. Pathogens 12(7):923. https://doi.org/10.3390/pathogens12070923.
108. Heneberg P, Rojas A, Bizos J, Kocková L, Malá M, Rojas D (2014). Focal *Philophthalmus gralli* infection possibly persists in *Melanoides tuberculata* over two years following the definitive hosts' removal. Parasitol Int 63(6):802–807. https://doi.org/10.1016/j.parint.2014.07.012.
109. Sato H, Suzuki K (2006). Gastrointestinal helminths of feral raccoons (*Procyon lotor*) in Wakayama Prefecture, Japan. J Vet Med Sci 68(4):311–318. https://doi.org/10.1292/jvms.68.311.
110. Zhao GH, Blair D, Li XY, Li J, Lin RQ, Zou FC, Sugiyama H, Mo XH, Yuan ZG, Song HQ, Zhu XQ (2011). The ribosomal intergenic spacer (IGS) region in *Schistosoma japonicum*: structure and comparisons with related species. Infect Genet Evol 11:610–617. https://doi.org/10.1016/j.meegid.2011.01.015.
111. Chan AHE, Chaisiri K, Saralamba S, Morand S, Thaenkham U (2021). Assessing the suitability of mitochondrial and nuclear DNA genetic markers for molecular systematics and species identification of helminths. Parasit Vectors 14:233. https://doi.org/10.1186/s13071-021-04737-y.
112. Dumbovic G, Forcales SV, Perucho M (2017). Emerging roles of macrosatellite repeats in genome organization and disease development. Epigenetics 12(7):515–526. https://doi.org/10.1080/15592294.2017.1318235.
113. Parvathy ST, Udayasuriyan V, Bhadana V (2022). Codon usage bias. Mol Biol Rep 49(1):539–565. https://doi.org/10.1007/s11033-021-06749-4.
114. Wey-Fabrizius AR, Podsiadlowski L, Herlyn H, Hankeln T (2013). Platyzoan mitochondrial genomes. Mol Phylogenet Evol 69(2):365–375. https://doi.org/10.1016/j.ympev.2012.12.015.
115. Karin BR, Arellano S, Wang L, Walzer K, Pomerantz A, Vasquez JM, Chatla K, Sudmant PH, Bach BH, Smith LL, McGuire JA (2023). Highly-multiplexed and efficient long-amplicon PacBio and Nanopore sequencing of hundreds of full mitochondrial genomes. BMC Genomics 24(1):229. https://doi.org/10.1186/s12864-023-09277-6.

116. Liu S, Liu Y, Chen B, Lu X, Jiang D, Geng L, Wang X, Peng K, Du C, Ren T, Yang X (2023). The complete mitochondrial genome of *Morishitium polonicum* (Trematoda, Cyclocoelidae) and its phylogenetic implications. Parasitol Res 122(11):2609–2620. https://doi.org/10.1007/s00436-023-07959-4.
117. Kudlai O, Kostadinova A, Pulis EE, Tkach VV (2017). The Psilostomidae Looss, 1900 (*sensu stricto*) (Digenea: Echinostomatoidea): description of three new genera and a key to the genera of the family. Syst Parasitol 94(1):21–33. https://doi.org/10.1007/s11230-016-9681-5.
118. Sitko J (2021). New Findings of Trematodes of the Superfamily Echinostomatoidea Looss, 1899 in Birds from The Czech Republic. Helminthologia 58(4):364–371. https://doi.org/10.2478/helm-2021-0040.
119. Pantoja C, Faltýnková A, O'Dwyer K, Jouet D, Skírnisson K, Kudlai O (2021). Diversity of echinostomes (Digenea: Echinostomatidae) in their snail hosts at high latitudes. Parasite 28:59. https://doi.org/10.1051/parasite/2021054.
120. Sereno-Uribe AL, González-García MT, Ortega-Olivares MP, López-Jiménez A, García-Varela M, Andrade-Gómez L (2022). First record of *Patagifer bilobus* (Rudolphi, 1819) Dietz, 1909 (Digenea: Echinostomatidae), with a morphological and molecular characterization from two threskiornithid species in Mexico. Parasitol Res 121(7):1921–1935. https://doi.org/10.1007/s00436-022-07526-3.
121. Valadão MC, Alves PV, López-Hernández D, Assis JCA, Coelho PRS, Geiger SM, Pinto HA (2023). A new cryptic species of *Echinostoma* (Trematoda: Echinostomatidae) closely related to *Echinostoma paraensei* found in Brazil. Parasitology 150(4):1–11. https://doi.org/10.1017/S003118202300001X.
122. Chai JY, Cho J, Chang T, Jung BK, Sohn WM (2020). Taxonomy of *Echinostoma revolutum* and 37-Collar-Spined *Echinostoma* spp.: A Historical Review. Korean J Parasitol 58(4):343–371. https://doi.org/10.3347/kjp.2020.58.4.343.
123. Nagataki M, Tantrawatpan C, Agatsuma T, Sugiura T, Duenngai K, Sithithaworn P, Andrews RH, Petney TN, 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. https://doi.org/10.1016/j.meegid.2015.07.022
124. Detwiler JT, Bos DH, Minchella DJ (2010). Revealing the secret lives of cryptic species: examining the phylogenetic relationships of echinostome parasites in North America. Mol Phylogenet Evol 55: 611–620. https://doi.org/10.1016/j.ympev.2012.12.015.
125. Gordy MA, Hanington PC (2019). A fine-scale phylogenetic assessment of digenean trematodes in central Alberta reveals we have yet to uncover their total diversity. Ecol Evol 9:3153–3238. https://doi.org/10.1002/ece3.4939.

126. Maji AK, Bera DK, Manna B, Nandy A, Addy M, Bandyopadhyay AK (1993). First record of human infection with *Echinostoma malayanum* in India. Trans R Soc Trop Med Hyg 87:673. http://doi:10.1016/0035-9203(93)90285-x.
127. Belizario VY, Geronilla GG, Anastacio MB, de Leon WU, Subaan AU, Sebastian AC, Bangs MJ (2007). *Echinostoma malayanum* infection, the Philippines. Emerg Infect Dis 13:1130–1131. http://doi:10.3201/eid1307.061486.
128. Lane C (1915). *Artyfechinostomum surfrartyfex*. A new parasitic echinostome of man. Indian J Med Res 2:977–983.
129. Mukherjee RP, Ghosh RK (1968). On the synonymy of the genus *Artyfechinostomum* Lane, 1915 (Trematoda: Echinostomatidae). Proc Indian Acad Sci 67:52–58.
130. Lie KJ (1963). Studies on Echinostomatidae in Malaya. III. The adult *Echinostoma malayanum* Leiper, 1911 (Trematoda) and the probable synonymy of Artyfechinostomum sufrartyfex Lane, 1915. Z Parasitenkd 23:124–135. https://doi.org/10.1007/BF00260288
131. Tantrawatpan C, Saijuntha W, Sithithaworn P, Andrews RH, Petney TN (2013). Genetic differentiation of *Artyfechinostomum malayanum* and *A. sufrartyfex* (Trematoda: Echinostomatidae) based on internal transcribed spacer sequences. Parasitol Res 112(1):437–441. http://doi:10.1007/s00436-012-3065-9.
132. Chai JY (2009). Echinostomes in humans. In Fried B and Toledo R (eds), The Biology of Echinostomes. New York, USA: Springer, pp. 147–183.
133. Chai JY, Jung BK (2020) Foodborne intestinal flukes: A brief review of epidemiology and geographical distribution. Acta Trop 201:105210. https://doi.org/10.1016/j.actatropica.2019.105210.
134. Anucherngchai S, Chontananarth T (2019). *Echinostoma revolutum*: Development of a high performance DNA-specific primer to demonstrate the epidemiological situations of their intermediate hosts. Acta Trop 189:46–53. http://doi:10.1016/j.actatropica.2018.09.014.
135. Mohanta UK, Watanabe T, Anisuzzaman, Ohari Y, Itagaki T (2019). Characterization of *Echinostoma revolutum* and *Echinostoma robustum* from ducks in Bangladesh based on morphology, nuclear ribosomal ITS2 and mitochondrial *nad*1 sequences. Parasitol Int 69:1–7. http://doi:10.1016/j.parint.2018.11.002.
136. Kostadinova A, Gibson DI, Biserkov V, Ivanova R (2000a). A quantitative approach to the evaluation of the morphological variability of two echinostomes, *Echinostoma miyagawai* Ishii, 1932 and *E. revolutum* (Frolich, 1802) from Europe. Syst Parasitol 45:1–15. https://doi.org/10.1023/A:1006232612469.
137. Kostadinova A, Gibson DI, Biserkow V, Chipev N (2000b). Re-validation of *Echinostoma miyagawai* Ishii, 1932 (Digenea: Echinostomatidae) on the basis of the

experimental completion of its life-cycle. Syst Parasitol 45:81–108. https://doi.org/10.1023/A:1006241610689.

138. Kanev I (1994). Life-cycle, delimitation and redescription of *Echinostoma revolutum* (Froelich, 1802) (Trematoda: Echinostomatidae). Syst Parasitol 28:125–144. https://doi.org/10.1007/BF00009591.
139. Chai JY, Jung BK, Chang T, Shin H, Cho J, Ryu JY, Kim HS, Park K, Jeong MH, Hoang EH, Abdullah MBM (2021). *Echinostoma miyagawai* Ishii, 1932 (Echinostomatidae) from Ducks in Aceh Province, Indonesia with Special Reference to Its Synonymy with *Echinostoma robustum* Yamaguti, 1935. Korean J Parasitol 59(1):35–45. https://doi.org/10.3347/kjp.2021.59.1.35.
140. Biswal DK, Ghatani S, Shylla JA, Sahu R, Mullapudi N, Bhattacharya A, Tandon V (2013). An integrated pipeline for next generation sequencing and annotation of the complete mitochondrial genome of the giant intestinal fluke, *Fasciolopsis buski* (Lankester, 1857) Looss, 1899. PeerJ 1: e207. https://doi.org/10.7717/peerj.207.
141. Jones BP, Norman BF, Borrett HE, Attwood SW, Mondal MMH, Walker AJ, Webster JP, Rajapakse RPVJ, Lawton SP (2020). Divergence across mitochondrial genomes of sympatric members of the *Schistosoma indicum* group and clues into the evolution of *Schistosoma spindale*. Sci Rep 10(1):2480. https://doi.org/10.1038/s41598-020-57736-x. Erratum in: Sci Rep. 2021; 11(1):1246.
142. Gacad JLJ, Tanabe-Hosoi S, Yurlova NI, Urabe M (2023). The complete mitogenome of *Echinoparyphium aconiatum* (Digenea: Echinostomatidae) and a comparison with other digenean species. Parasitol Int 92:102682. https://doi.org/10.1016/j.parint.2022.102682.
143. Tyagi K, Chakraborty R, Cameron SL, Sweet AD, Chandra K, Kumar V (2020). Rearrangement and evolution of mitochondrial genomes in *Thysanoptera* (Insecta). Sci Rep 10(1):695. https://doi:10.1038/s41598-020-57705-4.
144. Howe DK, Denver DR (2008). Muller's Ratchet and compensatory mutation in *Caenorhabditis briggsae* mitochondrial genome evolution. BMC Evol Biol 8:62. https://doi:10.1186/1471-2148-8-62.
145. Kostadinova A, Gibson DI (2000). The systematics of the echinostomes. In: Fried B, Graczyk TK, editors. Echinostomes as experimental models for biological research. Dordrecht: Kluwer Academic Publishers; pp. 31–57.
146. Olson PD (2000). New insights into platyhelminth systematics and evolution. Parasitol Today 16(1):3–5. https://doi.org/10.1016/s0169-4758(99)01555-0.
147. Izrailskaia AV, Besprozvannykh VV, Tatonova YV (2021). *Echinostoma chankensis* nom. nov., other *Echinostoma* spp. and *Isthmiophora hortensis* in East Asia: morphology, molecular data and phylogeny within Echinostomatidae. Parasitology 148(11):1366–1382. https://doi.org/10.1017/S0031182021000950.

148. Solà E, `lvarez Presas M, Frías-López C, Littlewood DTJ, Rozas J, Riutort M. (2015). Evolutionary analysis of mitogenomes from parasitic and free-living flatworms. Plos One 10:e0120081 https://doi.org/10.1371/journal.pone.0120081.
149. Miliotis MD, Bier JW (eds) (2003). International handbook of foodborne pathogens (Food Science and Technology). CRC Press, New York (860 pp.). (ISBN-13: 978-0824706852).
150. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27(5):722–736. https://doi.org/10.1101/gr.215087.116.
151. Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013). QUAST: quality assessment tool for genome assemblies. Bioinformatics 29(8):1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
152. Tamura K, Stecher G, Kumar S (2021). MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Biol Evol 38:3022–3027. https://doi.org/10.1093/molbev/msab120.
153. 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. http://doi:10.1093/molbev/msy096.
154. Lowe TM, Chan PP (2016). tRNAscan on-line: Search and contextual analysis of transfer RNA genes. Nucleic Acids Res 44:W54–57. https://doi.org/10.1093/nar/gkw413.
155. Laslett D, Canback B (2008). ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics (Oxford, England) 24:172–175. http://doi:10.1093/bioinformatics/btm573.
156. Benson G (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580. https://doi.org/10.1093/nar/27.2.573. (updated on October 20, 2022).
157. Conant GC and Wolfe KH (2008). GenomeVx: simple web-based creation of editable circular chromosome maps. Bioinformatics 24(6):861–862. https://doi.org/10.1093/bioinformatics/btm598.
158. Coppens L, Wicke L, Lavigne R (2022). SAPPHIRE.CNN: Implementation of dRNA-seq-driven, species-specific promoter prediction using convolutional neural networks. Comput Struct Biotechnol J 20:4969–4974. https://doi.org/10.1016/j.csbj.2022.09.006
159. Katoh K, Standley DM (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010

160. Criscuolo A, Gribaldo S (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol 10:210. https://doi.org/10.1186/1471-2148-10-210.
161. Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, Gascuel O (2019). NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res* 47(W1):W260–W265. https://doi.org/10.1093/nar/gkz303.
162. Perna NT, Kocher TD (1995). Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 41:353–358. http://doi:10.1007/BF00186547.
163. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59(3):307–321. https://doi.org/10.1093/sysbio/syq010.2010
164. Junier T, Zdobnov EM (2010). The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. Bioinformatics 26(13):1669–1670. https://doi.org/10.1093/bioinformatics/btq243.
165. Rambaut A (2018). FigTree, version 1.4.4. http://tree.bio.ed.ac.uk/software/figtree/.
166. Chai JY, Jung BK (2017). Fishborne zoonotic heterophyid infections: An update. Food Waterborne Parasitol 8–9:33–63. https://doi.org/10.1016/j.fawpar.2017.09.001.
167. Tkach VV, Littlewood DT, Olson PD, Kinsella JM, Swiderski Z (2003). Molecular phylogenetic analysis of the Microphalloidea Ward, 1901 (Trematoda: Digenea). Syst Parasitol 56(1):1–15. https://doi.org/10.1023/a:1025546001611.
168. Tatonova YV, Chelomina GN, Besprosvannykh VV (2012). Genetic diversity of nuclear ITS1–5.8S–ITS2 rDNA sequence in *Clonorchis sinensis* Cobbold, 1875 (Trematoda: Opisthorchidae) from the Russian Far East. Parasitol Int 61(4):664–674. https://doi.org/10.1016/j.parint.2012.07.005
169. O'Brien EA, Zhang Y, Yang L, Wang E, Marie V, Lang BF, Burger G (2009). GOBASE: an organelle genome database. Nucleic Acids Res 37(Database issue): D946–D950.
170. Wolstenholme DR (1992). Animal mitochondrial DNA, structure and evolution. Int Rev Cytol 141:173–216.
171. Yan HB, Wang XY, Lou ZZ, Li L, Blair D, Yin H, Cai JZ, Dai XL, Lei MT, Zhu XQ, Cai XP, Jia WZ (2013). The mitochondrial genome of *Paramphistomum cervi* (Digenea), the first representative for the family Paramphistomidae. PLoS One 8(8):e71300. https://doi.org/10.1371/journal.pone.0071300.
172. Lamolle G, Fontenla S, Rijo G, Tort JF, Smircich P (2019). Compositional Analysis of Flatworm Genomes Shows Strong Codon Usage Biases Across All Classes. Front Genet 10:771. https://doi.org/10.3389/fgene.2019.00771

173. Rhoads A, Au KF (2015). PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics 13(5):278–289. https://doi.org/10.1016/j.gpb.2015.08.002.
174. Hancock-Hanser BL, Frey A, Leslie M, Dutton PH, Archer FI, Morin PA (2013). Targeted multiplex next-generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics. Mol Ecol Resour 13:254–268. https://doi.org/10.1111/1755-0998.12059.
175. Tang M, Tan M, Meng G, Yang S, Su X, Liu S, Song W, Li Y, Wu Q, Zhang A, Zhou X (2014). Multiplex sequencing of pooled mitochondrial genomes-a crucial step toward biodiversity analysis using mito-metagenomics. Nucleic Acids Res 42(22):e166. https://doi.org/10.1093/nar/gku917.
176. Arunkumar KP, Nagaraju J (2006). Unusually long palindromes are abundant in mitochondrial control regions of insects and nematodes. PloS One 1:e110. https://doi.org/10.1371/journal.pone.0000110.
177. Miyazawa H, Osigus HJ, Rolfes S, Kamm K, Schierwater B, Nakano H (2021). Mitochondrial Genome Evolution of Placozoans: Gene Rearrangements and Repeat Expansions. Genome Biol Evol 13: evaa213. https://doi.org/10.1093/gbe/evaa213.
178. Bronstein O, Kroh A, Haring E (2018). Mind the gap! The mitochondrial control region and its power as a phylogenetic marker in echinoids. BMC Evol Biol 18:80. https://doi.org/10.1186/s12862-018-1198-x.
179. Wang HC, Li K, Susko E, Roger AJ (2008). A class frequency mixture model that adjusts for site-specific amino acid frequencies and improves inference of protein phylogeny. BMC Evol Biol 8:331. https://doi.org/10.1186/1471-2148-8-331.
180. Thaenkham U, Nawa Y, Blair D, Pakdee W (2011). Confirmation of the paraphyletic relationship between families Opisthorchiidae and Heterophyidae using small and large subunit ribosomal DNA sequences. Parasitol Int 60(4):521–523. https://doi.org/10.1016/j.parint.2011.07.015.
181. Berman JJ (2019) Taxonomic guide to infectious diseases: Understanding the Biologic Classes of Pathogenic Organisms. 2nd edition. Academic Press, London, UK. ISBN: 978-0-12-817576-7. https://www.elsevier.com/books-and-journals
182. van Herwerden L, Blair D, Agatsuma T (1999). Intra- and interindividual variation in ITS1 of *Paragonimus westermani* (Trematoda: Digenea) and related species: implications for phylogenetic studies. Mol Phylogenet Evol 12:67–73. https://doi.org/10.1006/mpev.1998.0572.
183. Vilas R, Criscione CD, Blouin MS (2005). A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. Parasitology 131(6):839–846. http://doi:10.1017/S0031182005008437.

184. Van VK, Dalsgaard A, Blair D, Le TH (2009). *Haplorchis pumilio* and *H. taichui* in Vietnam discriminated using ITS-2 DNA sequence data from adults and larvae. Exp Parasitol 123(2):146-51. https://doi.org/10.1016/j.exppara.2009.06.011.
185. Waeschenbach A, Webster BL, Littlewood DTJ (2012). Adding resolution to ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with large fragments of mtDNA. Mol Phylogenet Evol 63:834–847. https://doi.org/10.1016/j.ympev.2012.02.020
186. Waeschenbach A, Littlewood DTJ (2017). A molecular framework for the Cestoda. In: Caira JN and Jensen K (eds.). Planetary Biodiversity Inventory (2008–2017): Tapeworms from Vertebrate Bowels of the Earth. University of Kansas, Natural History Museum, Special Publication No. 25, Lawrence, KS, USA, pp. 431–451.
187. Sokolov SG, Shchenkov SV, Frolov EV, Gordeev II (2022) A Phylogenetic Re-Evaluation of the Stenakrine Opecoelids (Trematoda, Digenea: Opecoeloidea) with Some Taxonomic Novelties. Diversity 14:949. https://doi.org/10.3390/d14110949
188. Georgieva S, Blasco-Costa I, Kostadinova A (2017). Molecular characterisation of four echinostomes (Digenea: Echinostomatidae) from birds in New Zealand, with descriptions of *Echinostoma novaezealandense* n. sp. and *Echinoparyphium poulini* n. sp. Syst Parasitol 94(4):477–497. https://doi.org/10.1007/s11230-017-9712-x
189. Bespalaya YV, Kondakov AV, Travina OV, Khrebtova IS, Kropotin AV, Aksenova OV, Gofarov MYU, Lyubas AA, Tomilova AA, Vikhrev IV (2022). First record of metacercariae trematodes *Opisthioglyphe ranae* (Digenea: Telorchiidae) and *Echinostoma bolschewense* (Digenea: Echinostomatidae) in *Dreissena polymorpha* (Bivalvia: Dreissenidae) from the Don and Volga river basins, Russia. Ecol Montenegrina 54:57–76. https://doi.org/10.37828/em.2022.54.8
190. Sudarikov VE, Karmanova EM (1977). On validity and structure of the family Echinochasmidae (Odner, 1910). Trudy GELAN 27:129–141.
191. Besprozvannykh VV, Rozhkovan KV, Ermolenko AV (2017). *Stephanoprora chasanensis* n. sp. (Digenea: Echinochasmidae): Morphology, life cycle, and molecular data. Parasitol Int 66(1):863–870. https://doi.org/10.1016/j.parint.2016.10.005.
192. Johansen MV, Sithithaworn P, Bergquist R, Utzinger J (2010). Towards improved diagnosis of zoonotic trematode infections in Southeast Asia. Adv Parasitol 73:171–195. https://doi.org/10.1016/S0065-308X(10)73007-4.
193. Devi KR, Narain K, Mahanta J, Nirmolia T, Blair D, Saikia SP, Agatsuma T (2013). Presence of three distinct genotypes within the *Paragonimus westermani* complex in northeastern India. Parasitology 140(1):76–86. https://doi.org/10.1017/S0031182012001229.
194. Voronova AN, Vainutis KS, Tabakaeva TV, Sapotsky MV, Karaeka NN, Volkov YG, Galkina IV, Shchelkanov MY (2022). Molecular identification of the trematode *P.*

ichunensis stat. n. from lungs of siberian tigers justified reappraisal of *Paragonimus westermani* species complex. J Parasit Dis 46(3):744–753. https://doi.org/10.1007/s12639-022-01481-7

APPENDICES

Chapter 2

Supplementary Table S2.1 Primers for amplification and sequencing fragments of 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.

Supplementary Table S2.2 Primers for amplification and sequencing fragments of mitochondrial genome of *Echinostoma malayanum*

Primer	Sequence (5' to 3')	Location
URNLF*	AGCCAGGTTGGTTCTTATCTAT	*rrn*L
URNSR*	TACCWTGTTACGACTTAHCWCA	*rrnS*
TRECOBF*	CAGATGTCYTATTGGGCTGC	*cob*
TRECOBR*	GAACHRVCCACARYCCCTTAAAC	*cob*
JB3F*	TTTTTTGGGCATCCTGAGGTTTAT	*cox*1
JB4.5R*	TAAAGAAAGAACATAATGAAAATG	*cox*1
GLYF*	AGTATKYYGTCTTTCCAAGTC	*trn*G
GLYR*	ACKAGACCHCYGACTTGGAAAGAC	*trn*G
EMA1F	TTTRATTCTTGCTATGGCGGC	*cox*1
EMA2R	TCCCAATAACCATAGTCACAGACC	*cox*1
EMA3F	ACGAGTGTGACGGGGTATAG	*nad*1
EMA4R	ACCCCATAAGTACCCCCTACC	*nad*1
EMA5R	AACCCTCCCCAAACACCAAG	*cox*1
EMA6F	CTATCCATAGCCCCAACCCG	*nad*6
EMA7R	CCGCATAGCCTCCAACAATC	*nad*5
EMA8F	AGCGGTTTGAGTAGGGTATGTG	*nad*5
EMA9R	GAATGAAACAGAACCACATCACC	*cob*
EMA10F	GTGCTGCTAACTTTGTGTTGC	*cob*
EMA11R	ACCACCAGACTTTGGCAACC	*nad*2
EMA12F	CAGTGTTTGAGTTTCGTTCTTGGCTG	*nad5*
EMA13R	CCTACTGTAGCAAAACATACACCC	*cox*3
EMA14R	TCAGTACCCCCAAAACACCC	*nad*2
EMA14F	CTGGTTTTGTCGTTGTGTGG	*nad*5
EMA15F	AGGAGGCCTGTATCAATGTG	*nad*4
EMA15R	ACAGTCCCCGAAATAAACCAG	repeat

EMA16F	GGCTTAGGGTTTAAGGTAATCG	repeat
EMA17R	CCCTTTCAGAGAACACACTCAT	NCR
ECH3F	ATGAKTTGRTTGCCWATRTATAAAGC	cox3
ECHN4F	AGTTTGATTGGTATAGTTGGGG	*nad*4
ECHN2F	CTTGTTGGTGTCATATGATGC	*nad*2

*Platyhelminth-universal primers used for inital amplification of the corresponding genes/regions to get sequences to design further primers; NCR: non-coding region.

Supplementary Table S2.3 List of trematode-universal and specific primers for long-range PCR used for enrichment of the mitogenomes of *Echinostoma miyagawai* (Emiya) and *Hypoderaeum conoideum* (Hcono) to obtain amplicons for next-generation sequencing

	Primer pair	Spanning region	Length of amplicons (Emiya/ Hcono)	Position in the genome (Emiya/ Hcono)	Primer sequence (5' to 3')
1	ECH3F-JNAD1R	*cox*3–*nad*1	5.7 kb/ 5.8 kb	1–5682/ 1–5705	ECH3F: ATGAKTTGRTTGCCWATRTATAAAGC JNAD1R: ATACACATAAAACAGGCCTC
2	ECHN4F-JNAD1R	*nad*4-*nad*1	3.6 kb/ 3.6 kb	2072–5682/ 2087–5705	ECHN4F: AGTTTGATTGGTATAGTTGGGG JNAD1R: ATACACATAAAACAGGCCTC
3	ECHN2F-URNSR	*nad*2–*rrn*S	5.7 kb/ 5.7 kb	4603–10280/ 4618–10319	ECHN2F: CTTGTTGGTGTCATATGATGC URNSR: TACCATGTTACGACTTACCACA
4	ECHN2F-UNI16R	*nad*2–*rrn*L	4.6 kb/ 4.6 kb	4603–9216/ 4618–9249	ECHN2F: CTTGTTGGTGTCATATGATGC UNI16R: TCTCGGGGTCTTTCCGTCT
5	UNI16F-GLYR	*rrn*L–*trn*G	4.4 kb/ 4.4 kb	9050–13395/ 9083–13445	UNI16F: TGGCCGCAGTATHTTGACTGTGC GLYR: ACKAGACCHCYGACTTGGAAAGAC
6	URNLF-GLYR	*rrn*L–*trn*G	4.0 kb/ 4.0 kb	9455–13395/ 9488–13445	URNLF: AGCCAGGTTGGTTCTTATCTAT GLYR: ACKAGACCHCYGACTTGGAAAGAC
7	EHC5F-EHC3R*	*nad*5–*cyt*B	7.0 kb/ 5.6 kb	13200–789/ 13214–793	EHC5F: TGTTTCTTTYTATCGTTGGGAGGT EHC3R: CCCCCACACCAAAAATAACTCAA

Note: *Primer pair number 7 was used for amplifying the NCR for NGS sequencing.

Supplementary Table S2.4 List of **57 trematode strains of 41 species** providing information for the available mitochondrial genome in the suborder Echinostomata and other suborders used in this study for sequence comparative and phylogenetic analyses

	Family/Species/Strains	Strains or designed	Country of collection	GenBank	mtDNA as reported (bp)	mt DNA*	PCGs	MRGs	Reference (if any)
	Suborder Echinostomata								
	Echinostomatidae (15/12)								
1	***Artyfechinostomum malayanum***	(EMI3)	Thailand	OK509083	17,175	13408	10131	1725	Pham et al. (2022)
2	*Artyfechinostomum sufrartyfex*	(Shillong)	India	KY548763	14,567	13409	10131	1728	GenBank
3	*Echinoparyphium aconiatum*	(Chany)	Russia	ON644993	14,865	13377	10113	1730	Gacad et al. (2023)
4	*Echinostoma caproni*	(SAMEA)	Egypt	AP017706	14,150	13293	10128	1709	GenBank
5	***Echinostoma miyagawai***	(RED11)	Thailand	OP326312	19,417	13324	10128	1725	This study
6	*Echinostoma miyagawai*	(Hunan)*	China	MN116740	14,460	13320	10128	1724	Fu et al. (2019)
7	*Echinostoma miyagawai*	(HLJ)*	China	MH393928	14,410	13321	10128	1763	Li et al. (2019)
8	*Echinostoma paraensei*	n/a	n/a	KT008005	20,298	13319	10128	1748	GenBank
9	***Echinostoma revolutum***	(MSD15)	Thailand	MN496162	17,030	13326	10134	1733	Le et al. (2020)

10	*Echinostoma* sp. (*revolutum*?)	(GD)	China	MN116706	15,714	13282	10149	1754	Ran et al. (2020)
11	*Echinostoma* sp.	(JM-2019)	China	MH212284	15,283	13257	10122	1726	GenBank
12	Echinostomatidae sp. CA-2021	(PE4)	United States	MK264774	14,426	13319	10143	1727	GenBank
13	Echinostomatidae sp. MSB para 30070	(A19)	United States	MN822299	13,985	13346	10128	1732	GenBank
14	*Hypoderaeum conoideum*	(Hubei)	China	KM111525	14,180	13361	10116	1730	Yang et al. (2015)
15	***Hypoderaeum conoideum***	(RED42)	Thailand	PP110501	18,011	13361	10116	1728	This study
	Cyclocoelidae (3/3)								
16	*Morishitium polonicum*	(Laojun)	China	OP930879	14,083	13337	10137	1720	Liu et al. (2023)
17	*Tracheophilus cymbius*	(HLJ)	China	MK355447	13,760	13458	10152	1745	Li et al. (2019)
18	*Uvitellina* sp. SSS2019	(SSS2019)	Pakistan	MK227160	14,217	13705	10200	1751	Suleman et al. (2019)
	Echinochasmidae (1/1)								
19	*Echinochasmus japonicus*	(PT)	Vietnam	KP844722	15,865	13378	10143	1748	Le et al. (2016)
	Fasciolidae (7/6)								
20	*Fasciola gigantica*	(GX)	China	KF543342	14,478	13309	10107	1755	Liu et al. (2014)
21	*Fasciola* sp. (hybrid)	(GHL)	China	KF543343	14,453	13282	10107	1755	Liu et al. (2014)
22	*Fasciola hepatica*	(GL)	Australia	AF216697	14,462	13305	10104	1755	Le et al. (2001)
23	*Fasciola hepatica*	(Oregon)	United States	AP017707	14,374	13302	10110	1750	GenBank
24	*Fascioloides jacksoni*	(Madu)	Sri Lanka	KX787886	14,952	13286	10137	1743	Rajapakse et al. (2020)
25	*Fascioloides magna*	(Koko)	Czech	KU060148	14,047	13272	10131	1745	Ma et al. (2016)
26	*Fasciolopsis buski*	(Jangxi)	China	KX169163	14,833	13380	10122	1768	Ma et al. (2017)
	Himasthlidae (1/1)								
27	*Acanthoparyphium* sp.	(WAK-2018)	Kuwait	MG792058	14,191	13328	10119	1753	GenBank
	Suborder Troglotremata								
	Paragonimidae (13/6)								
28	*Paragonimus skrjabini miyazakii*	(OkuST1)	Japan	ON782295	17,591	13240	10098	1714	This study
29	*Paragonimus heterotremus*	(GX)	China	MH059809	13,927	13222	10101	1711	Qian et al. (2018)
30	*Paragonimus heterotremus*	(LC)	Vietnam	KY952166	13,526	13230	10101	1720	GenBank
31	*Paragonimus iloktsuenensis*	(Amami)	Japan	ON961029	14,827	13205	10104	1712	Le et al. (2023)
32	*Paragonimus ohirai*	(Kino)	Japan	KX765277	14,818	13222	10104	1710	Le et al. (2019)
33	*Paragonimus westermani*	(dog1)	China	MN412705	14,790	13215	10104	1731	GenBank
34	*Paragonimus westermani*	(dog2)	China	MN412706	14,774	13215	10104	1731	GenBank
35	*Paragonimus westermani* (3n)	(Bogil)	South Korea	AF219379	14,244	13213	10101	1732	GenBank
36	*Paragonimus westermani* (2n)	(Haenam)	South Korea	AF540958	14,965	13213	10101	1732	GenBank
37	*Paragonimus westermani*	(type 1)	India	KM280646	14,015	13223	10104	1729	GenBank
38	*Paragonimus westermani*	(AP)	India	KX943544	14,975*	13216	10104	1721	Biswal et al. (2014)
39	*Paragonimus westermani* (2n)	(IND2009)	India	CM017921	20,273	13230	10104	1729	Oye et al. (2019)
40	*Paragonimus kellicotti*	(Ozark)	United States	MH322000	13,786	13196	10098	1711	Wang et al. (2018)
	Suborder Opisthorchiata								
	Heterophyidae (4/3)								
41	*Cryptocotyle lingua*	(66766)	Norway	OL853496	13,983	13467	10173	1735	GenBank
42	*Haplorchis taichui*	(QT3)	Vietnam	MG972809	15,120	13277	10164	1730	GenBank
43	*Haplorchis taichui*	(LA)	Laos	KF214770	15,131	13225	10164	1730	Lee et al. (2013)
44	*Metagonimus yokogawai*	n/a	South Korea	KC330755	15,258	13364	10245	1736	GenBank
	Opisthorchiidae (9/6)								
45	*Amphimerus* sp.	(JM-2019)	Ecuador	MK238506	15,151	13387	10179	1754	Ma et al. (2019)
46	*Clonorchis sinensis*	(Amur)	Russia	FJ381664	13,875	13511	10209	1777	Shekhovtsov et al. (2010)
47	*Clonorchis sinensis*	(GD)	China	JF729303	13,879	13514	10209	1778	Cai et al. (2012)
48	*Clonorchis sinensis*	n/a	South Korea	JF729304	13,877	13512	10209	1777	Cai et al. (2012)

49	*Clonorchis sinensis*	(Jinju)	South Korea	MT607652	18,304	13512	10209	1777	Kinkar et al. (2020)
50	*Opisthorchis viverrini*	n/a	Laos	JF739555	13,510	13457	10206	1767	Cai et al. (2012)
51	*Opisthorchis felineus*	(UstTula)	Russia	EU921260	14,277	13499	10218	1769	Shekhovtsov et al. (2010)
52	*Opisthorchis sudarikovi*	(Swabi)	Pakistan	MK033132	13,641	13559	10197	1768	Suleman et al. (2019)
53	*Metorchis orientalis*	(HLJ)	China	KT239342	13,834	13422	10176	1765	Na et al. (2016)
	Suborder Xiphidiata Dicrocoeliidae (3/3)								
54	*Dicrocoelium chinensis*	(Gansu)	China	KF318786	14,917	13234	10125	1675	Liu et al. (2014)
55	*Dicrocoelium dendriticum*	(Gansu)	China	KF318787	14,884	13229	10104	1617	Liu et al. (2014)
56	*Eurytrema pancreaticum*	(HLJ)	China	KP241855	15,031	13532	10143	1746	Chang et al. (2016)
	(outgroup) Schistosomatidae (1/1)								
57	*Schistosoma haematobium*	(N10)	Mali	DQ157222	15,003	-	10116	-	Littlewood *et al.* (2006)

mtDNA: the entire mitochondrial genome; mtDNA*: the coding mitochondrial genome (5' terminus of *cox*3 to 3' terminus of *nad*5); PCGs: protein-coding genes; MRGs: mitoribosomal genes; *the length of the *Echinostoma miyagawai* Hunan (Hunan strain; MN116740) has been corrected from 14,468 bp to 14,460 bp, and HLJ (Heilongjiang strain; MH393928) from 14,416 bp to 14,410 bp. The numbers in a bracket indicate the number of isolates and species in that family used for the genetic and phylogenetic analyses.

Reading references for Supplementary Table S2.4

Biswal DK, Chatterjee A, Bhattacharya A, 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. https://doi.org/10.7717/peerj.484.

Cai XQ, Liu GH, Song HQ, Wu CY, Zou FC, Yan HK, Yuan ZG, Lin RQ, Zhu XQ (2012). Sequences and gene organization of the mitochondrial genomes of the liver flukes *Opisthorchis viverrini* and *Clonorchis sinensis* (Trematoda). Parasitol Res 110(1):235–243. https://doi.org/10.1007/s00436-011-2477-2.

Chang QC, Liu GH, Gao JF, Zheng X, Zhang Y, Duan H, Yue DM, Fu X, Su X, Gao Y, Wang CR (2016). Sequencing and characterization of the complete mitochondrial genome from the pancreatic fluke *Eurytrema pancreaticum* (Trematoda: Dicrocoeliidae). Gene 576(1):160–165. https://doi.org/10.1016/j.gene.2015.09.081.

Fu YT, Jin YC, Li F, Liu GH (2019). Characterization of the complete mitochondrial genome of the echinostome *Echinostoma miyagawai* and phylogenetic implications. Parasitol Res 118:3091–3097. https://doi.org/10.1007/s00436-019-06417-4.

Gacad JLJ, Tanabe-Hosoi S, Yurlova NI, Urabe M (2023). The complete mitogenome of Echinoparyphium aconiatum (Digenea: Echinostomatidae) and a comparison with other digenean species. Parasitol Int 92:102682. https://doi.org/10.1016/j.parint.2022.102682.

Kinkar L, Young ND, Sohn WM, Stroehlein AJ, Korhonen PK, Gasser RB (2020). First record of a tandem-repeat region within the mitochondrial genome of *Clonorchis sinensis* using a long-read sequencing approach. PLOS Negl Trop Dis 14(8):e0008552. https://doi.org/10.1371/journal.pntd.0008552.

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. https://doi.org/10.1017/s0031182001008733.

Le TH, Nguyen KT, Nguyen NTB, Doan HTT, Agatsuma T, 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. https://doi.org/10.7717/peerj.7031.

Le TH, Nguyen KT, Pham LTK, Doan HTT, DT Roan, TKX Le, Agatsuma T, Blair D (2023). Mitogenomic and nuclear ribosomal transcription unit datasets support the synonymy of *Paragonimus iloktsuenensis* and *P. ohirai* (Paragonimidae: Platyhelminthes). Parasitol Res 122(7):1531–1544. https://doi.org/10.1007/s00436-023-07854-y

Le TH, Nguyen NTB, Nguyen KT, Doan HTT, Dung DT, Blair D (2016). A complete mitochondrial genome from *Echinochasmus japonicus* supports the elevation of Echinochasminae Odhner, 1910 to family rank (Trematoda: Platyhelminthes). Infect Genet Evol 45:369–377. https://doi.org/10.1016/j.meegid.2016.09.024.

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(5):566–576. https://doi.org/10.1017/S0031182020000128.

Lee D, Choe S, Park H, Jeon HK, Chai JY, Sohn WM, Yong TS, Min DY, Rim HJ, EomKS (2023). Complete mitochondrial genome of *Haplorchis taichui* and comparative analysis with other trematodes. Korean J Parasitol 51(6):719–726. https://doi.org/10.3347/kjp.2013.51.6.719.

Li Y, Ma XX, Lv QB, Hu Y, Qiu HY, Chang QC, Wang CR (2019). Characterization of the complete mitochondrial genome sequence of *Tracheophilus cymbius* (Digenea), the first representative from the family Cyclocoelidae. J Helminthol 94:e101. https://doi.org/10.1017/S0022149X19000932.

Li Y, Qiu YY, Zeng MH, Diao PW, Chang QC, Gao Y, Zhang Y, Wang CR (2019). The complete mitochondrial genome of *Echinostoma miyagawai*: Comparisons with closely related species and phylogenetic implications. Infect Genet Evol 75:103961. https://doi.org/10.1016/j.meegid.2019.103961.

Liu GH, Gasser RB, Young ND, Song HQ, Ai L, Zhu XQ (2014). Complete mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola gigantica*, and their comparison with *F. hepatica*. Parasit Vectors. 2014; 7:150. https://doi.org/10.1186/1756-3305-7-150.

Liu GH, Yan HB, Otranto D, Wang XY, Zhao GH, Jia WZ, Zhu XQ. *Dicrocoelium chinensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochondrial and nuclear ribosomal DNA sequences. Mol Phylogenet Evol 79:325–331. https://doi.org/10.1016/j.ympev.2014.07.002.

Liu S, Liu Y, Chen B, Lu X, Jiang D, Geng L, Wang X, Peng K, Du C, Ren T, Yang X (2023). The complete mitochondrial genome of *Morishitium polonicum* (Trematoda, Cyclocoelidae) and its phylogenetic implications. Parasitol Res 122(11):2609–2620. https://doi.org/10.1007/s00436-023-07959-4.

Ma J, He JJ, Liu GH, Leontovyč R, Kašný M, Zhu XQ (2016). Complete mitochondrial genome of the giant liver fluke *Fascioloides magna* (Digenea: Fasciolidae) and its comparison with selected trematodes. Parasit Vectors 9:429. https://doi.org/10.1186/s13071-016-1699-7.

Ma J, He JJ, Zhou CY, Sun MM, Cevallos W, Sugiyama H, Zhu XQ, Calvopiña M (2019). Characterization of the mitochondrial genome sequences of the liver fluke *Amphimerus* sp. (Trematoda: Opisthorchiidae) from Ecuador and phylogenetic implications. Acta Trop 195:90–96. https://doi.org/10.1016/j.actatropica.2019.04.025.

Ma J, Sun MM, He JJ, Liu GH, Ai L, Chen MX, Zhu XQ (2017). *Fasciolopsis buski* (Digenea: Fasciolidae) from China and India may represent distinct taxa based on mitochondrial and nuclear ribosomal DNA sequences. Parasit Vectors 10:101. https://doi.org/10.1186/s13071-017-2039-2.

Na L, Gao JF, Liu GH, Fu X, Su X, Yue DM, Gao Y, Zhang Y, Wang CR (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. https://doi.org/10.1016/j.meegid.2016.01.010.

Oey H, Zakrzewski M, Narain K, Devi KR, Agatsuma T, Nawaratna S, Gobert GN, Jones MK, Ragan MA, McManus DP, Krause L (2019). Whole-genome sequence of the oriental lung fluke *Paragonimus westermani.* Gigascience 8(1): giy146. https://doi.org/10.1093/gigascience/giy146.

Pham, KLT, Saijuntha, W, Lawton, SP, Le, TH (2022). Mitophylogenomics of the zoonotic fluke *Echinostoma malayanum* confirms it as a member of the genus *Artyfechinostomum* Lane, 1915 and illustrates the complexity of Echinostomatidae systematics. Parasitol Res 121:899–913. https://doi.org/10.1007/s00436-022-07449-z.

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 B Resour 3(2):560–561. https://doi.org/10.1080/23802359.2018.1462119.

Rajapakse RPVJ, Pham KLT, Karunathilake KJK, Lawton SP, Le TH (2020). Characterization and phylogenetic properties of the complete mitochondrial genome of *Fascioloides jacksoni* (syn. *Fasciola jacksoni*) support the suggested intergeneric change from *Fasciola* to *Fascioloides* (Platyhelminthes: Trematoda: Plagiorchiida). Infect Genet Evol 82:104281. https://doi.org/10.1016/j.meegid.2020.104281.

Ran R, Zhao Q, Abuzeid AMI, Huang Y, Liu Y, Sun Y, He L, Li X, Liu J, Li G (2020). Mitochondrial genome sequence of *Echinostoma revolutum* from Red-Crowned Crane (*Grus japonensis*). Korean J Parasitol 58(1):73–79. https://doi.org/10.3347/kjp.2020.58.1.73.

Shekhovtsov SV, Katokhin AV, Kolchanov NA, Mordvinov VA (2010). The complete mitochondrial genomes of the liver flukes *Opisthorchis felineus* and *Clonorchis sinensis* (Trematoda). Parasitol Int 59:100–103. https://doi.org/10.1016/j.parint.2009.10.012.

Suleman KM, Heneberg P, Zhou CY, Muhammad N, Zhu XQ, Ma J (2019). Characterization of the complete mitochondrial genome of *Uvitellina* sp., representative of the family Cyclocoelidae and phylogenetic implications. Parasitol Res 118:2203–2211. https://doi.org/10.1007/s00436-019-06358-y.

Suleman, Ma J, Khan MS, Sun MM, Muhammad N, He JJ, Zhu XQ (2019). Mitochondrial and nuclear ribosomal DNA dataset suggests that *Hepatiarius sudarikovi* Feizullaev, 1961 is a member of the genus *Opisthorchis* Blanchard, 1895 (Digenea: Opisthorchiidae). Parasitol Res 118(3):807–815. https://doi.org/10.1007/s00436-019-06227-8.

Wang T, Wang Y, Xu F, Li X, Qu R, Song L, Tang Y, Lin P (2018). Characterization of the complete mitochondrial genome of the lung fluke *Paragonimus kellicotti*. Mitochondrial DNA Part B 3(2):715–716. https://doi.org/10.1080/23802359.2018.1483763.(retracted).

Yang X, Gasser RB, Koehler AV, 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. https://doi.org/10.1186/s13071-015-0720-x.

Supplementary Table S2.5 List and information on strains and species that provide complete sequences of the ribosomal 28S and 18S rRNA genes for phylogenetic analysis and tree reconstruction to assess the intra- and interfamilial relationships in the class Trematoda (Digenea: Platyhelminthes)

No	Suborder/Superfamily/ Family/Species	Sequence designation and GenBank Nos	Country of isolation or report	Sequences (bp)		References or sources
				Concatenated	Complete 28S	
	DIPLOSTOMATA					
	Cyathocotylidae					
1	*Cyathocotyle prussica*	Cprus-GPS721-DE-MH521249	Germany	5,873	3,885	Locke et al. (2018)
	Diplostomidae					
2	*Alaria americana*	Aamer-NVS16-CA-MH521246	Canada	5,875	3,897	Locke et al. (2018)
3	*Diplostomum ardeae*	Darde-Yauco-PR-MT259036	Puerto Rico	5,875	3,897	Locke et al. (2020)
4	*Diplostomum pseudospathaceum*	Dpseu-pse3a-CZ-KR269766	Czech	5,874	3,896	Brabec et al. (2015)
5	*Diplostomum spathaceum*	Dspat-spa3a-CZ-KR269765	Czech	5,875	3,897	Brabec et al. (2015)
6	*Hysteromorpha triloba*	Htril-HMSq-IT-MH521250	Italy	5,874	3,897	Locke et al. (2018)
7	*Posthodiplostomum centrarchi*	Pcent-Hudson-CA-MH521251	Canada	5,874	3,897	Locke et al. (2018)

8	*Tylodelphys immer*	Timme-DiIN-CA-MH521252	Canada	5,823	3,894	Locke et al. (2018)
	Strigeidae					
9	*Apharyngostrigea pipientis*	Apipi-Quebec-CA-MT677870	Canada	5,878	3,899	Locke et al. (2021)
10	*Cotylurus marcogliesei*	Cmarc-MTL25-CA-MH521248	Cananda	5,874	3,897	Locke et al. (2018)
11	*Cardiocephaloides medioconiger*	Cmedi-Florida-US-MH521247	United States	5,864	3,899	Locke et al. (2018)
	ECHINOSTOMATA					
	Echinostomatidae					
12	*Artyfechinostomum malayanum*	Amala-EMI3-TH-OR509026	Thailand	5,852	3,864	This study
13	*Echinostoma miyagawai*	Emiya-RED11-TH-OR509027	Thailand	5,849	3,861	This study
14	*Echinostoma miyagawai*	Emiya-(duck)-CN-MH748722	China	N/A	3,861	GenBank
15	*Echinostoma revolutum*	Erevo-MSD15-TH-OR509028	Thailand	5,851	3,863	This study
16	*Hypoderaeum conoideum*	Hcono-RED42-TH-OR509029	Thailand	5,850	3,862	This study
17	*Isthmiophora hortensis*	Ihort-Waka-JP-AB189982	Japan	5,846	3,862	Sato and Suzuki (2006)
	Echinochasmidae					
18	*Echinochasmus japonicus*	Ejapo-EjPT10-VN-OR509030	Vietnam	5,849	3,661	This study
	Fasciolidae					
19	*Fasciolopsis buski*	Fbusk-HT-VN-MN970005	Vietnam	5,851	3,862	Le et al. (2020)
20	*Fasciola gigantica*	Fgiga-NB-VN-MN970009	Vietnam	5,852	3,863	Le et al. (2020)
21	*Fasciola gigantica*	Fgiga-T4V-VN-MN970010	Vietnam	5,852	3,863	Le et al. (2020)
22	*Fasciola gigantica* (hybrid)	Fgiga-DL11-VN-MN970008	Vietnam	5,852	3,863	Le et al. (2020)
23	*Fasciola hepatica*	Fhepa-GL-AU-MN970007	Australia	5,852	3,863	Le et al. (2020)
24	*Fascioloides jacksoni*	Fjack-Madu-LK-MN970006	Sri Lanka	5,852	3,863	Le et al. (2020)
	Philophthalmidae					
25	*Philophthalmus gralli*	Pgral-PH#191-PE-JQ627832	Peru	5,848	3,859	Heneberg et al. (2014)
	Cyclocoelidae					
26	*Tracheophilus cymbius*	Tcymb-(duck)-CN-MK327367	China	N/A	3,832	GenBank
	OPISTHORCHIATA					
	Cryptogonimidae					
27	*Stemmatostoma cribbi*	Scrib-WesP-SB-OQ968484	Solomon Islands	5,767	3,884	Mathews et al. (2023)
	Heterophyidae					
28	*Cryptocotyle lingua*	Cling-Kartesh-RU-MW361240	Russia	5,870	3,879	GenBank
	Euryhelmis costaricensis	Ecost-Fuku-JP-AB521797	Japan	5,870	3,879	Sato et al. (2010)
30	*Haplorchis pumilio*	Hpumi-HPU8-VN-KX815125	Vietnam	5,862	3,870	Le et al. (2017)
31	*Haplorchis taichui*	Htaic-QT3-VN-KX815126	Vietnam	5,867	3,875	Le et al. (2017)
	Opisthorchiidae					
32	*Clonorchis sinensis*	Csine-CSA-CN-MK450523	China	5,851	3,860	Qiu et al. (2020)
33	*Clonorchis sinensis*	Csine-CSB-CN-MK450524	China	5,851	3,860	Qiu et al. (2020)
34	*Clonorchis sinensis*	Csine-CSC-CN-MK450525	China	5,851	3,860	Qiu et al. (2020)
35	*Clonorchis sinensis*	Csine-CSD-CN-MK450526	China	5,851	3,860	Qiu et al. (2020)
36	*Clonorchis sinensis*	Csine-CSE-CN-MK450527	China	5,851	3,860	Qiu et al. (2020)
37	*Metorchis orientalis*	Morie-MOA-CN-MK482051	China	5,869	3,876	Qiu et al. (2020)
38	*Metorchis orientalis*	Morie-MOB-CN-MK482052	China	5,869	3,876	Qiu et al. (2020)
39	*Metorchis orientalis*	Morie-MOC-CN-MK482053	China	5,869	3,876	Qiu et al. (2020)
40	*Metorchis orientalis*	Morie-MOD-CN-MK482054	China	5,869	3,876	Qiu et al. (2020)
41	*Metorchis orientalis*	Morie-MOE-CN-MK482055	China	5,869	3,876	Qiu et al. (2020)
	PRONOCEPHALATA					
	Paramphistomidae					
42	*Paramphistomum cervi*	Pcerv-PCA-CN-KJ459935	China	5,863	3,873	Zheng et al. (2014)
43	*Paramphistomum cervi*	Pcerv-PCB-CN-KJ459936	China	5,863	3,873	Zheng et al. (2014)
44	*Paramphistomum cervi*	Pcerv-PCC-KJ459937	China	5,863	3,873	Zheng et al. (2014)
45	*Paramphistomum cervi*	Pcerv-PCD-CN-KJ459938	China	5,863	3,873	Zheng et al. (2014)
46	*Paramphistomum cervi*	Pcerv-PCE-CN-KJ459934	China	5,862	3,873	Zheng et al. (2014)
	XIPHIDIATA					
	Brachycladiidae					
47	*Brachycladium goliath*	Bgoli-NHM-UK-KR703279	United Kingdom	5,860	3,867	Briscoe et al. (2016)
	Collyriclidae					
48	*Collyriclum faba*	Cfaba-Orlicke-CZ-JQ231122	Czech	5,838	3,867	Heneberg and Literák (2013)
	Dicrocoeliidae					
49	*Eurytrema pancreaticum*	Epanc-EP1-CN-KY490000	China	5,869	3,877	Su et al. (2018)
50	*Eurytrema pancreaticum*	Epanc-EP2-CN-KY490001	China	5,869	3,877	Su et al. (2018)
51	*Eurytrema pancreaticum*	Epanc-EP3-CN-KY490002	China	5,869	3,877	Su et al. (2018)
52	*Eurytrema pancreaticum*	Epanc-EP4-CN-KY490003	China	5,869	3,877	Su et al. (2018)
53	*Eurytrema pancreaticum*	Epanc-EP5-CN-KY490004	China	5,869	3,877	Su et al. (2018)
	Haploporidae					
54	*Carassotrema koreanum*	Ckore-(Ccaur)-RU-ON598382	Russia	5,829	3,841	Ivashko et al. (2022)
55	*Parasaccocoelium mugili*	Pmugi-Primo-RU-MW813991	Russia	5,847	3,866	Atopkin et al. (2021)
	Microphallidae					
56	*Maritrema eroliae*	Merol-(Cbifa)-KW-JF826247	Kuwait	N/A	3,871	Al-Kandari et al. (2011)
	Nanophyetidae					
57	*Nanophyetus japonensis*	Njapo-NJ142-JP-LT796170	Japan	N/A	3,885	GenBank
58	*Nanophyetus japonensis*	Njapo-NJ161-JP-LT796169	Japan	N/A	3,885	GenBank

59	*Nanophyetus salmincola*	Nsalm-Karp51-RU-LN871822	Russia	5,848	3,874	Voronova and Chelomina (2018)
60	*Nanophyetus salmincola*	Nsalm-Karp55-RU-LN871823	Russia	5,848	3,874	Voronova and Chelomina (2018)
61	*Nanophyetus schikhobalowi*	Nschi-03Karp1442-RU-LN871820	Russia	N/A	3,886	Voronova et al. (2017)
62	*Nanophyetus schikhobalowi*	Nschi-karp1451-RU-LN871818	Russia	N/A	3,885	Voronova et al. (2017)
63	*Nanophyetus schikhobalowi*	Nschi-kkh6-RU-MG966187	Russia	N/A	3885	Voronova et al. (2017)
64	*Nanophyetus schikhobalowi*	Nschi-kkh9-RU-MG966188	Russia	N/A	3885	Voronova et al. (2017)
	Paragonimidae					
65	*Paragonimus iloktsuenensis*	Pilok-Amami-JP-OP081042	Japan	5,858	3,881	Le et al. (2023)
66	*Paragonimus ohirai*	Pohir-Kino-JP-OP081041	Japan	5,858	3,881	Le et al. (2023)
	Prosthogonimidae					
67	*Prosthogonimus cuneatus*	Pcune-HLJ-CN-MW376724	China	5,811	3,841	GenBank
	Zoogonidae					
68	*Lepidophyllum cameroni*	Lcame-Lepcam1-RU-MN217107	Russia	5,787	3,876	GenBank
69	*Lepidophyllum cameroni*	Lcame-Lepcam2-RU-MN217108	Russia	5,787	3,876	GenBank
70	*Lepidophyllum steenstrupi*	Lstee-NSea-UK-AY157175	United Kingdom	N/A	3,829	Lockyer et al. (2003)
	Out group					
	Schistosomatidae					
71	*Schistosoma edwardiense*	Sedwa-Edward-UG-AY197344	Uganda	5,860	3,870	Morgan et al. (2003)

Note: Sequence abbreviation: five or six letters indicating the first capital letter as from the generic name and the next four or five as from the species name; the strain designation (from local, geographical, voucher, or the abbreviated host name) is given in the middle; and the country name with a two-letter abbreviation (according to the list of country codes at https://www.iban.com/country-codes). The outgroup sequence is taken from *Schistosoma edwardiense* (Schistosomatidae).

Supplementary Table S2.6 List and information on strains and species that provide partial nuclear ribosomal 28S rDNA D1−D3 sequences for phylogenetic analysis and tree reconstruction for the assessment of the taxonomic relationships of the suborder Echinostomata (Trematoda: Platyhelminthes)

No	Suborder, Superfamily, Family and Species	Sequence designation	Country of isolation	GenBank accession No	Sources/References
	Suborder ECHINOSTOMATA (154 sequences/85 species/42 genera) **Superfamily Echinostomatoidea** **Family Echinostomatidae** (87 sequences/42 species/17 genera)				
1	*Artyfechinostomum sufrartyfex*	Asufr-Khowai-IN	India	KF781302	GenBank
2	*Artyfechinostomum sufrartyfex*	Asufr-Shillong-IN	India	KF781303	GenBank
3	*Artyfechinostomum sufrartyfex*	Asufr-ZOOASad-IN	India	MH236132	GenBank
4	*Artyfechinostomum sufrartyfex*	Asufr-ZOOASme-IN	India	MH236133	GenBank
5	*Artyfechinostomum malayanum*	Amala-EMI3-TH	Thailand	OR509026	This study
6	*Cathaemasia hians*	Chian-(Ppla)-CZ	Czech	KT956947	Tkach et al. (2016)
7	*Chaunocephalus ferox*	Cfero-(Cnig)-UA	Ukraine	KT447522	GenBank
8	*Drepanocephalus auritus*	Dauri-MJGDA-US	United States	KP053259	Pinto et al. (2016)
9	*Drepanocephalus auritus*	Dauri-HAPH1-BR	Brazil	KP053260	Pinto et al. (2016)
10	*Drepanocephalus mexicanus*	Dmexi-DNA2623-MX	Mexico	MF351543	Hernández-Cruz et al. (2018)
11	*Drepanocephalus spathans*	Dspat-DNA11122-MX	Mexico	MF351545	Hernández-Cruz et al. (2018)
12	*Drepanocephalus spathans*	Dspat-HCCMissi-US	United States	JN993270	Griffin et al. (2012)
13	*Echinostoma bolschewense*	Ebols-TDRE281-RU	Russia	MZ517159	Bespalaya et al. (2022)
14	*Echinostoma bolschewense*	Ebols-EBG13-SK	Slovakia	KP065591	Georgieva et al. (2014)
15	*Echinostoma chankensis*	Echan-110-RU	Russia	MT577829	Izrailskaia et al. (2021)
16	*Echinostoma cinetorchis*	Ecine-3-RU	Russia	MT577828	Izrailskaia et al. (2021)
17	*Echinostoma cinetorchis*	Ecine-1- KR	South Korea	KX817344	Lee et al. (1988)
18	*Echinostoma maldonadoi*	Emald-LBT-BR	Brazil	OQ132569	Valadão et al. (2023)
19	*Echinostoma miyagawai*	Emiya-EMAP1-NZ	New Zealand	KY436408	Georgieva et al. (2017)
20	*Echinostoma miyagawai*	Emiya-EMT2-CZ	Czech	KP065593	Georgieva et al. (2014)
21	*Echinostoma miyagawai*	Emiya-Kherson-UA	Ukraine	KT956916	Tkach et al. (2016)
22	*Echinostoma miyagawai*	Emiya-RED11-TH	Thailand	OR509027	This study
23	*Echinostoma nasincovae*	Enasi-AF232-IE	Ireland	MZ409809	Pantoja et al. (2021)
24	*Echinostoma nasincovae*	Enasi-Plc4-RU	Russia	MK585198	Svinin et al. (2023)
25	*Echinostoma novaezealandense*	Enova-ENCA-NZ	New Zealand	KY436407	Georgieva et al. (2017)
26	*Echinostoma paraensei*	Eparae-(hamster)-US	United States	EU025867	Lotfy et al. (2008)

27	*Echinostoma paraulum*	Eparau-EPM1-DE	Germany	KP065604	Georgieva et al. (2014)
28	*Echinostoma paraulum*	Eparau-NOV2111-RU	Russia	OP389066	Vainutis et al. (2023)
29	*Echinostoma paraulum*	Eparau-EPT1-CZ	Czech	KP065605	Georgieva et al. (2014)
30	*Echinostoma pseudorobustum*	Epseu-(Ggal)-BR	Brazil	OK586835	Valadão et al. (2022)
31	*Echinostoma revolutum*	Erevo-AF206-IS	Iceland	MZ409810	Pantoja et al. (2021)
32	*Echinostoma revolutum*	Erevo-ERHH3-CZ	Czech	KP065598	Georgieva et al. (2014)
33	*Echinostoma revolutum*	Erevo-ERBA1-CZ	Czech	KP065594	Georgieva et al. (2014)
34	*Echinostoma revolutum*	Erevo-MSD15-TH	Thailand	OR509028	This study
35	*Echinostoma revolutum*	Erevo-AF235-US	United States	MZ409811	Pantoja et al. (2021)
36	*Echinostoma revolutum*	Erevo-VVT2015-US	United States	KT956915	Tkach et al. (2016)
37	*Echinostoma trivolvis*	Etriv-(Maura)-US	United States?	AY222246	Olson et al. (2003); Tkach et al. (2016)
38	*Echinostomatidae* sp.	Ech sp-CMA2010a-US	United States	GU270100	GenBank
39	*Echinoparyphium aconiatum*	Eacon-AF227-IE	Ireland	MZ409801	Pantoja et al. (2021)
40	*Echinoparyphium aconiatum*	Eacon-AF273-FI	Finland	MZ409802	Pantoja et al. (2021)
41	*Echinoparyphium cinctum*	Ecinc-UA(sub)-UA	Ukraine	AF184260	Tkach et al. (2001)
42	*Echinoparyphium recurvatum*	Erecu-Echinos_Ie-VN	Vietnam	OM956186	GenBank
43	*Echinoparyphium recurvatum*	Erecu-AF254-FI	Finland	MZ409803	Pantoja et al. (2021)
44	*Echinoparyphium recurvatum*	Erecu-AF204-IE	Ireland	MZ409804	Pantoja et al. (2021)
45	*Echinoparyphium rubrum*	Erubr-2(Pcolc)-US	United States	JF820595	Pulis et al. (2011)
46	*Echinoparyphium rubrum*	Erubr-AF241-US	United States	MZ409805	Pantoja et al. (2021)
47	*Euparyphium capitaneum*	Ecapi-3(Aanhi)-US	United States	KP009618	Kudlai et al. (2015)
48	*Euparyphium capitaneum*	Ecapi-5(Aanhi)-US	United States	KP009620	Kudlai et al. (2015)
49	*Hypoderaeum conoideum*	Hcono-AF261-FI	Finland	MZ409814	Pantoja et al. (2021)
50	*Hypoderaeum conoideum*	Hcono-E-UA	Ukraine	KT956918	Tkach et al. (2016)
51	*Hypoderaeum conoideum*	Hcono-AK44-CZ	Czech	KP065607	Georgieva et al. (2014)
52	*Hypoderaeum conoideum*	Hcono-NA-US	United States	KT956919	Tkach et al. (2016)
53	*Hypoderaeum conoideum*	Hcono-RED42-TH	Thailand	OR509029	This study
54	*Isthmiophora hortensis*	Ihort-Waka-JP	Japan	AB189982	Sato and Suzuki (2000)
55	*Isthmiophora melis*	Imeli-(Aagra)-PL	Poland	KT359583	Hildebrand et al. (2015)
56	*Isthmiophora melis*	Imeli- UA(sub)	Ukraine	AF151941	Tkach et al. (2000)
57	*Isthmiophora* sp.	Isth sp-MN2021-NS010-JP	Japan	LC599515	Nakao and Sasaki (2021)
58	*Isthmiophora* sp.	Isth sp-MSPara26645-KE	Kenya	MK482437	Laidemitt et al. (2021)
59	*Isthmiophora* sp.	Isth sp-VVT2015-Minn-US	United States	KT956920	Tkach et al. (2016)
60	*Moliniella anceps*	Mance-(Pcorn)-LT	Lithuania	KT956921	Tkach et al. (2016)
61	*Moliniella anceps*	Mance-AF230-IE	Ireland	MZ409815	Pantoja et al. (2021)
62	*Neopetasiger islandicus*	Nisla-(Aocci)-US	United States	KT956924	Tkach et al. (2016)
63	*Neopetasiger islandicus*	Nisla-AF415-IS	Iceland	MZ409816	Pantoja et al. (2021)
64	*Neopetasiger islandicus*	Nisla-MGC6-CA	Canada	KT831344	GenBank
65	*Neopetasiger islandicus*	Nisla-AK231-IS	Iceland	JQ425592	Georgieva et al. (2012)
66	*Neoacanthoparyphium echinatoides*	Nechi-Gabci-SK	Slovakia	KT956922	Tkach et al. (2016)
67	*Patagifer bilobus*	Pbilo-BIDI-VN	Vietnam	OR532446	This study
68	*Patagifer bilobus*	Pbilo-Kherson-UA	Ukraine	KT956945	Tkach et al. (2016)
69	*Patagifer bilobus*	Pbilo-DNA567-MX	Mexico	ON141912	Sereno-Uribe et al. (2022)
70	*Patagifer bilobus*	Pbilo-DNA4374-MX	Mexico	ON141920	Sereno-Uribe et al. (2022)
71	*Patagifer vioscai*	Pvios-(Ealbu)-US	United States	KT956946	Tkach et al. (2016)
72	*Patagifer vioscai*	Pvios-PV2-TZ	Tanzania	MZ412882	Chibwana and Katandukila (2021)
73	*Pegosomum asperum*	Paspe-Biberach-DE	Germany	KY945919	GenBank
74	*Pegosomum saginatum*	Psagi-DNA4374-DE	Germany	KY945918	Sereno-Uribe et al. (2022)
75	*Petasiger exaeretus*	Pexae-Kherson-UA	Ukraine	KT956923	Tkach et al. (2016)
76	*Petasiger exaeretus*	Pexae-KM4-HU	Hungary	KY284009	Cech et al. (2017)
77	*Petasiger phalacrocoracis*	Pphal-CK2-HU	Hungary	KY284005	Cech et al. (2017)
78	*Petasiger phalacrocoracis*	Ppha-KM1-HU	Hungary	KY284006	Cech et al. (2017)
79	*Petasiger radiatus*	Pradi-KM5-HU	Hungary	KY284010	Cech et al. (2017)
80	*Petasiger radiatus*	Pradi-(Pcarb)-UA	Ukraine	KT956927	Tkach et al. (2016)
81	*Rhopalias macracanthus*	Rmacr-NDakota-US	United States	KT956959	Tkach et al. (2016)
82	*Rhopalias macracanthus*	Rmacr-18-MX	Mexico	MK648280	Pérez-Ponce de León and Hernández-Mena (2019)
83	*Rhopalias* sp.	Rhopsp-2022-LBTRHOP3-BR	Brazil	OP972555	López-Hernández et al. (2023)
84	*Ribeiroia ondatrae*	Ronda-Cali-US	United States	MG544873	Calhoun et al. (2018)
85	*Ribeiroia ondatrae*	Ronda-NDakota-US	United States	KT956956	Tkach et al. (2016)
86	*Ribeiroia ondatrae*	Ronda-MSBPara32185-US	United States	OK188967	Keller et al. (2021)
87	*Ribeiroia ondatrae*	Ronda-JAM17N33-US	United States	MK321661	GenBank
	Family Caballerotrematidae (1 sequence/1 species/1 genus)				
88	*Caballerotrema* sp.	Cabal sp-VVT2015-HCIPD634-PE	Peru	KT956941	Tkach et al. (2016)
	Family Fasciolidae (19 sequences/7 species/4 genera)				
89	*Fasciola* sp. (hybrid)	Fgiga-DL11-VN	Vietnam	MN970008	Le et al. (2020)
90	*Fasciola gigantica*	Fgiga-(Jwsub)-TH	Thailand	HM004190	Thaenkham et al. (2010)
91	*Fasciola gigantica*	Fgiga-NB-VN	Vietnam	MF099787	Dao et al. (2017)
92	*Fasciola gigantica*	Fgiga-(Btaur)-SN	Senegal	AY222245	Olson et al. (2003)
93	*Fasciola gigantica*	Fgiga-(Btaur)-KE	Kenya	EU025873	Lotfy et al. (2008)

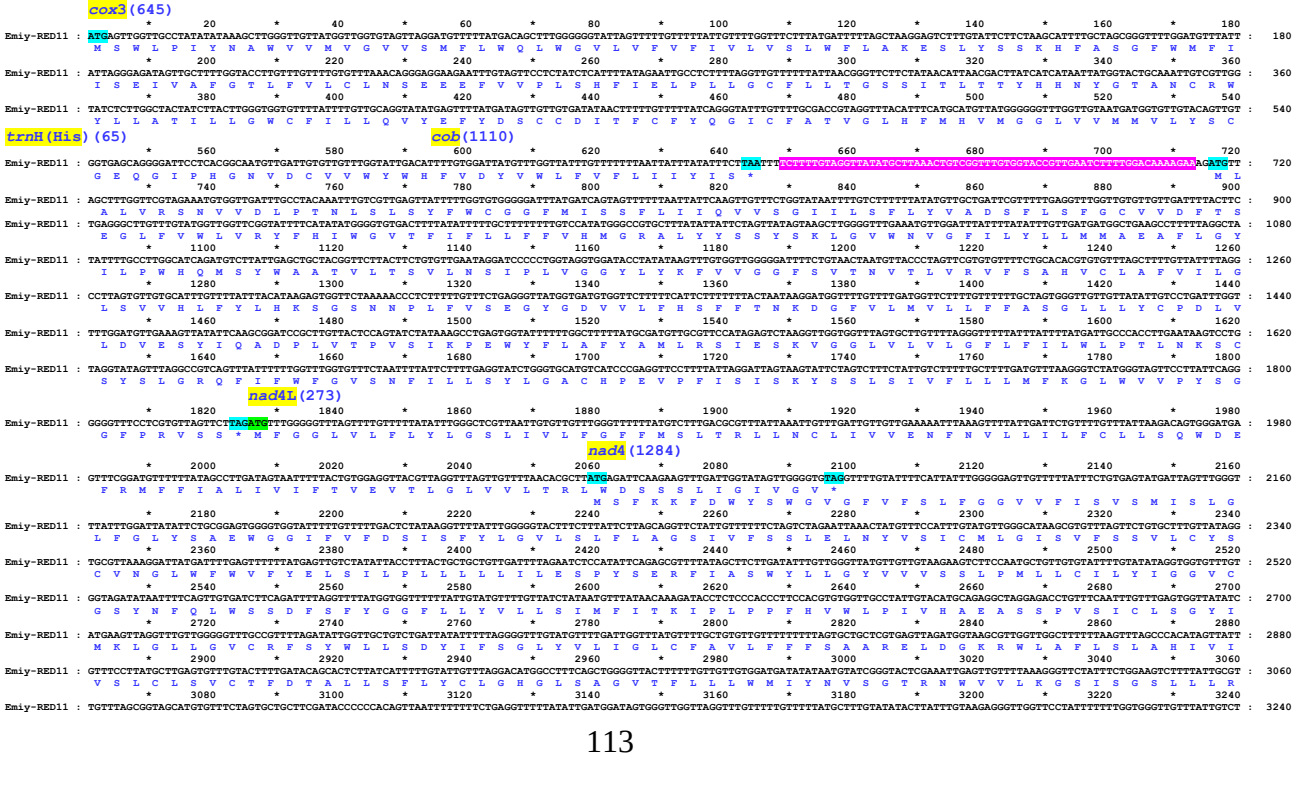
94	*Fasciola gigantica*	Fgiga-T4V-VN	Vietnam	MN970010	Le et al. (2020)
95	*Fasciola hepatica*	Fhepa-(Bbuba)-EG	Egypt	EU025874	Lotfy et al. (2008)
96	*Fasciola hepatica*	Fhepa-(Chirc)-SA	Saudi Arabia	AY222244	Olson et al. (2003)
97	*Fasciola hepatica*	Fhepa-Geelong-AU	Australia	MF099788	Dao et al. (2017)
98	*Fasciola hepatica*	Fhepa-ind7rb9-BR	Brazil	MW185781	GenBank
99	*Fascioloides jacksoni*	Fjack-(Emaxi)-LK	Sri Lanka	EU025871	Lotfy et al. (2008)
100	*Fascioloides jacksoni*	Fjack-Maduru-LK	Sri Lanka	MF099789	Dao et al. (2017)
101	*Fascioloides magna*	Fmagn-(Sscro)-US	United States	EU025872	Lotfy et al. (2008)
102	*Fascioloides magna*	Fmagn-Oktibehha-US	United States	KU232370	Lee et al. (2016)
103	*Fasciolopsis buski*	Fbusk-Hanoi-VN	Vietnam	EU025870	Lotfy et al. (2008)
104	*Fasciolopsis buski*	Fbusk-L2-VN	Vietnam	MF099785	Dao et al. (2017)
105	*Fasciolopsis buski*	Fbusk-HT-VN	Vietnam	MF099786	Dao et al. (2017)
106	*Fasciolopsis buski*	Fbusk-Megha-IN	India	KC602457	GenBank
107	*Parafasciolopsis fasciolaemorpha*	Pfasc-(Babona)-PL	Poland	EU025869	Lotfy et al. (2008)
	Family Himasthlidae (4 sequences/4 species/2 genera)				
108	*Acanthoparyphium spinulosum*	Aspin-(Psqua)-UA	Ukraine	KT956939	Tkach et al. (2016)
109	*Himasthla leptosoma*	Hlept-(Calpi)-UA	Ukraine	KT956942	Tkach et al. (2016)
110	*Himasthla limnodromi*	Hlimn-(Lgris)-US	United States	KT956943	Tkach et al. (2016)
111	*Himasthla militaris*	Hmili-(Ggall)-UA	Ukraine	KT956944	Tkach et al. (2016)
	Family Psilostomidae (8 sequences/6 species/4 genera)				
112	*Neopsilotrema lakotae*	Nlako-NDakota-US	United States	KU379696	Kudlai et al. (2016)
113	*Psilostomum brevicolle*	Pbrev-(Hostr)-UA	Ukraine	KT956950	Tkach et al. (2016)
114	*Psilochasmus oxyurus*	Poxyu-Kherson-UA	Ukraine	AF151940	Tkach et al. (2000)
115	*Sphaeridiotrema aziaticus*	Sazia-B91-RU	Russia	MT986043	Olson et al. (2003)
116	*Sphaeridiotrema monorchis*	Smono-Sm01-VN	Vietnam	JQ890544	Besprozvannykh et al. (2013)
117	*Sphaeridiotrema monorchis*	Smono-Sm05-VN	Vietnam	JQ890548	Besprozvannykh et al. (2013)
118	*Sphaeridiotrema monorchis*	Spseud-(Aaffi)-US	United States	KT956957	Tkach et al. (2016)
119	*Sphaeridiotrema ussuriensis*	Sussu-E71-RU	Russia	MT986039	Kalinina et al. (2022)
	Family Echinochasmidae (25 sequences/15 species/4 genera)				
120	*Echinochasmus beleocephalus*	Ebele-Kherson-UA	Ukraine	KT956929	Tkach et al. (2016)
121	*Echinochasmus bursicola*	Uburs-(Aalba)-UA	Ukraine	KT956938	Tkach et al. (2016)
122	*Echinochasmus coaxatus*	Ecoax-ECR1-DE	Germany	MN726944	Schwelm et al. (2020)
123	*Echinochasmus coaxatus*	Ecoax-Kherson-UA	Ukraine	KT956928	Tkach et al. (2016)
124	*Echinochasmus milvi*	Emilv-Em01-RU	Russia	KT873315	Besprozvannykh et al. (2017)
125	*Echinochasmus milvi*	Emilv-Em05-RU	Russia	KT873319	Besprozvannykh et al. (2017)
126	*Echinochasmus milvi*	Emilv-E125-RU	Russia	MT447054	Tatonova et al. (2020)
127	*Echinochasmus mordax*	Emorda-Kherson-UA	Ukraine	KT956931	Tkach et al. (2016)
128	*Echinochasmus perfoliatus*	Eperf-Hanoi-VN	Vietnam	OR532445	This study
129	*Echinochasmus suifunensis*	Esuif-E21-RU	Russia	MT447056	Tatonova et al. (2020)
130	*Echinochasmus suifunensis*	Esuif-E22-RU	Russia	MT447057	Tatonova et al. (2020)
131	*Echinochasmus japonicus*	Ejapo-Ej01ND-VN	Vietnam	JQ890579	Besprozvannykh et al. (2013)
132	*Echinochasmus japonicus*	Ejapo-Ej02ND-VN	Vietnam	JQ890580	Besprozvannykh et al. (2013)
133	*Echinochasmus japonicus*	Ejapo-EjHB-VN	Vietnam	OR532444	This study
134	*Echinochasmus japonicus*	Ejapo-EjPT10-VN	Vietnam	OR509030	This study
135	*Echinochasmus* sp.	Echas sp-MN2021-NS193-JP	Japan	LC599527	Nakao and Sasaki (2021)
136	*Echinochasmidae* sp.	Echas sp-1FD2019-AR	Argentina	MH532427	Dellagnola et al. (2019)
137	*Microparyphium facetum*	Mface-Missi-US	United States	KT956933	Tkach et al. (2016)
138	*Microparyphium* sp.	Microsp-MN2021-NS087-JP	Japan	LC599528	Nakao and Sasaki (2021)
139	*Stephanoprora amurensis*	Samur-E113-RU	Russia	MT447053	Tatonova et al. (2020)
140	*Stephanoprora amurensis*	Samur-E17-RU	Russia	MT447050	Tatonova et al. (2020)
141	*Stephanoprora chasanensis*	Schas-Sc01-RU	Russia	KT873320	Besprozvannykh et al. (2017)
142	*Stephanoprora pseudoechinata*	Spseu-NA-US	United States	KT956934	Tkach et al. (2016)
143	*Stephanoprora pseudoechinata*	Spseu-3A12-UA	Ukraine	KJ542636	Tkach et al. (2016)
144	*Stephanoprora pseudoechinata*	Spseu-E-RU	Russia	KT956935	Tkach et al. (2016)
	Family Cyclocoelidae (5 sequences/5 species/5 genera)				
145	*Cyclocoelum mutabile*	Cmuta-(Ccan)-UK	United Kingdom	AY222249	Olson et al. (2003)
146	*Morishitium polonicum malayense*	Mpolo-(Apan)-MY	Malaysia	LC520231	Urabe et al. (2020)
147	*Neohaematotrephus arayae*	Naray-Veracruz-MX	Mexico	MH725788	López-Jiménez et al. (2018)
148	*Tracheophilus cymbius*	Tcymb-duck-CN	China	MK327367	GenBank
149	*Typhlocoelum* sp.	Typh sp-VVT2015NDakota-US	United States	KT956960	Tkach et al. (2016)
	Family Philophthalmidae				

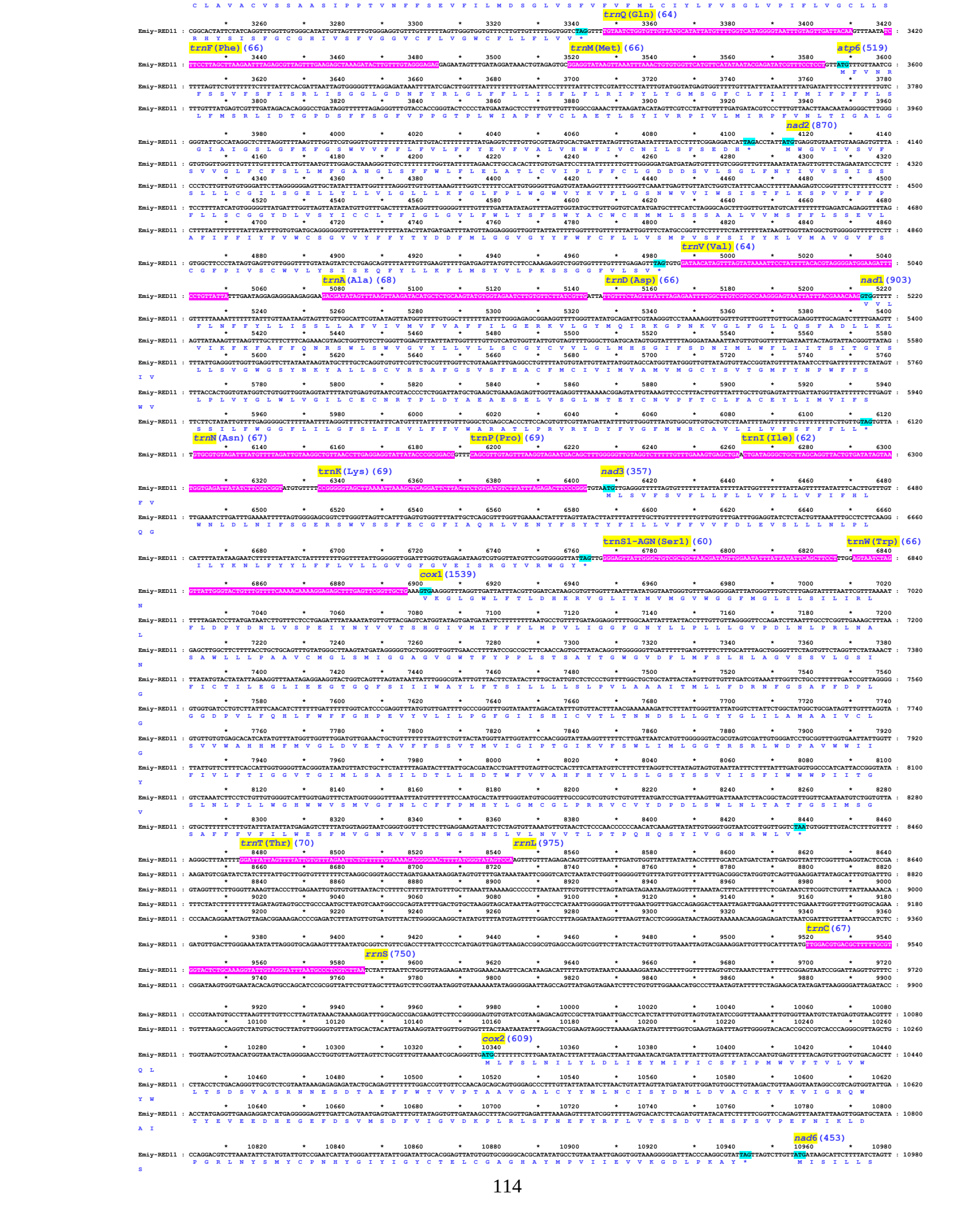
	(5 sequences/5 species/3 genera)				
150	*Cloacitrema michiganensis*	Cmich-HWML101879-US	United States	KT956948	Tkach et al. (2016)
151	*Cloacitrema narrabeenensis*	Cnarr-(Baus)-AU	Australia	AY222248	Olson et al. (2003)
152	*Parorchis acanthus*	Pacan-Jackson-US	United States	KT956949	Tkach et al. (2016)
153	*Parorchis trophoni*	Ptrop-VT2850-AR	Argentina	OP806518	Diaz et al. (2023)
154	*Philophthalmus gralli*	Pgral-PH#191-PE	Peru	JQ627832	Heneberg et al. (2014)
	Suborder XIPHIDIATA (5 sequencs/4 species/3 genera) **Superfamily Microphalloidea Family Eucotylidae** (5 sequences/4 species/3 genera)				
155	*Paratanaisia bragai*	Pbrag-(vouNHMUK)-BR	Brazil	JX231098	Unwin et al. (2013)
156	*Paratanaisia bragai*	Pbrag-(Zgray)-BR	Brazil	JX231099	Unwin et al. (2013)
157	*Tamerlania zarudnyi*	Tzaru		AF184248	Tkach et al. (2001)
158	*Tanaisia fedtschenkoi*	Tfedt-(Aplat)-UA	Ukraine	AY116870	Tkach et al. (2003)
159	*Tanaisia valida*	Tvali-Tv12-BR	Brazil	KX913714	GenBank
	Suborder HAPLOSPLANCHNATA (10 sequences/9 species/5 genera) **Superfamily Haplosplanchnoidea Family Haplosplanchnidae** (10 sequences/9 species/5 genera)				
160	*Haplosplanchnus pachysomus*	Hpach-THC17170-AU	Australia	KY852458	Huston et al. (2017)
161	*Haplosplanchnus purii*	Hpuri-(Mcep)-NC	New Caledonia	FJ211242	GenBank
162	*Hymenocotta mulli*	Hmull-(Cren)-AU	Australia	AY222239	Olson et al. (2003)
163	*Provitellotrema crenimugilis*	Pcren-PC3-RU	Russia	LK932154	Besprozvannykh et al. (2016)
164	*Trigonocephalotrema euclidi*	Teucl-THC16724A-AU	Australia	MG386255	Huston et al. (2018)
165	*Trigonocephalotrema sohcahtoa*	Tsohc-THC16155B-AU	Australia	MG386261	Huston et al. (2018)
166	*Trigonocephalotrema hipparchi*	Thipp-THC11426C-AU	Australia	MG386258	Huston et al. (2018)
167	*Schikhobalotrema huffmani*	Shuff-THC16619-AU	Australia	KY852463	Huston et al. (2017)
168	*Schikhobalotrema huffmani*	Shuff-THC17042-AU	Australia	KY852464	Huston et al. (2017)
169	*Schikhobalotrema sparisomae*	Sspar-(Laura)-ES	Spain	FJ211240	GenBank
	Outgroup: Schistosomatidae (1 sequence/1 species/1 genus)				
170	*Schistosoma haematobium*	Shaem-N10-ML	Mali	AY157263	Lockyer et al. (2003)

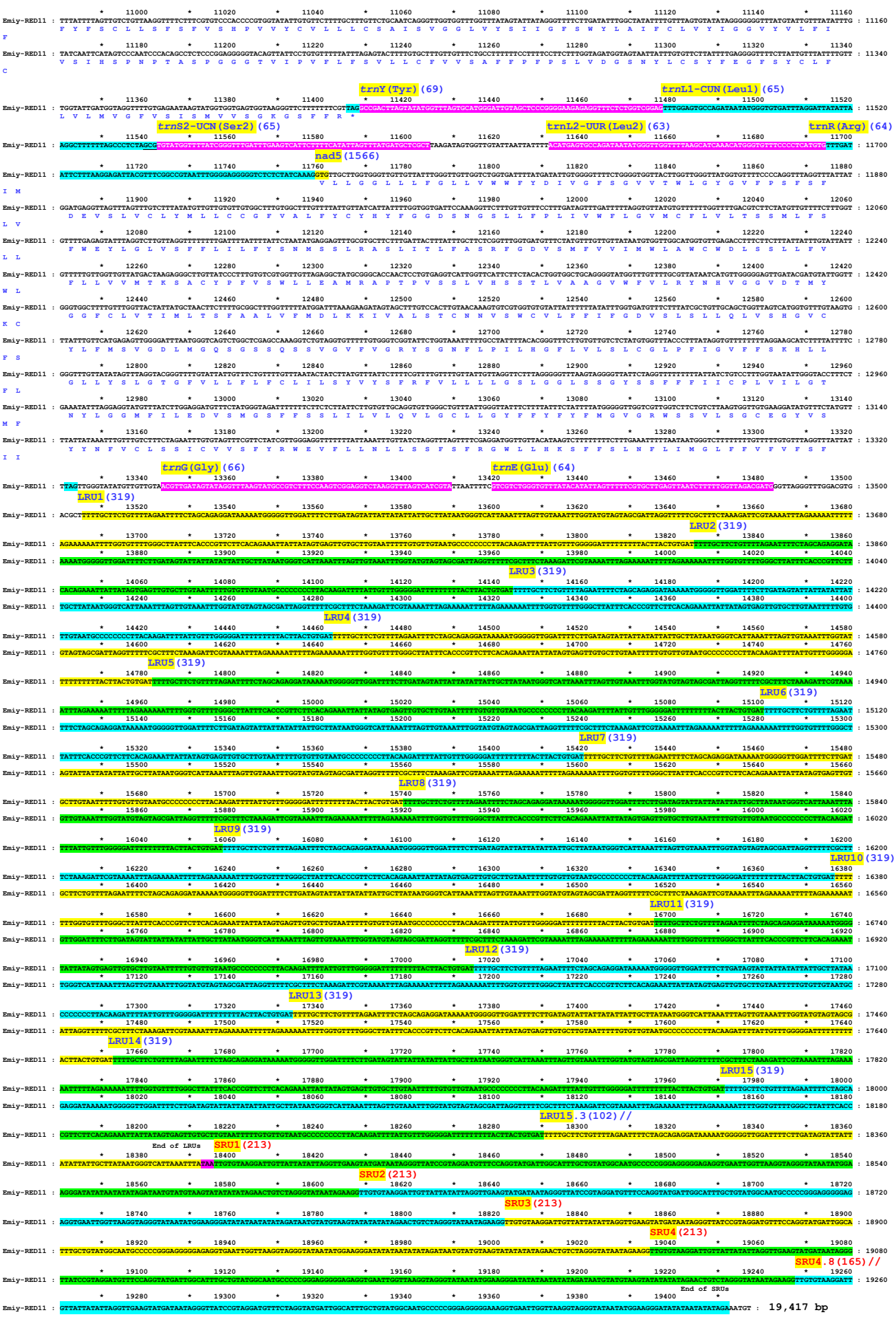
Note: Sequence abbreviation: five or six letters indicating the first capital letter as from the generic name and the next four or five as from the species name; the strain designation (from local, geographical, voucher, or the abbreviated host name) is given in the middle; and the country name with a two-letter abbreviation (according to the list of country codes at https://www.iban.com/country-codes). The outgroup sequence is taken from *Schistosoma haematobium* (Schistosomatidae).

Chapter 3

Echinostoma miyagawai complete mitochondrial genome




trnQ(Gln) (64)
trnM(Met) (66)
atp6(519)
trnF(Phe) (66)
nad2(870)
trnV(Val) (64)
trnA(Ala) (68)
trnD(Asp) (66)
nad1(903)
trnN(Asn) (67)
trnP(Pro) (69)
trnI(Ile) (62)
trnK(Lys) (69)
nad3(357)
trnS1-AGN(Ser1) (60)
trnW(Trp) (66)
cox1(1539)
trnT(Thr) (70)
rrnL(975)
trnC(67)
rrnS(750)
cox2(609)
nad6(453)


trnY(Tyr) (69)
trnL1-CUN(Leu1) (65)
trnS2-UCN(Ser2) (65)
trnL2-UUR(Leu2) (63)
trnR(Arg) (64)
nad5(1566)
trnG(Gly) (66)
trnE(Glu) (64)
LRU1(319)
LRU2(319)
LRU3(319)
LRU4(319)
LRU5(319)
LRU6(319)
LRU7(319)
LRU8(319)
LRU9(319)
LRU10(319)
LRU11(319)
LRU12(319)
LRU13(319)
LRU14(319)
LRU15.3(102)//
End of LRUs
SRU1(213)
SRU2(213)
SRU3(213)
SRU4(213)
SRU4.8(165)//
End of SRUs
19,417 bp

Supplementary Figure S3.1. Display of the ***Echinostoma miyagawai*** complete mitochondrial genome (19,417 bp, strain RED11 (Emiya-RED11-TH), GenBank: OP326312 (a representative mitogenome of echinostomes).

Supplementary Table S3.1. Codon usage for 12 protein-coding genes in the mitochondrial genomes of **15 strains of 12 species** of the family Echinostomatidae in this study

Amino acid*	Co-don	Amala (EMI3-TH) (OK509083)		Asufr (Shill-IN) (KY548763)		Ecapr (SAMEA-EG) ((AP017706)		Emiya (HLJ-CN) (MH393928)		Emiya (Hunan-CN) (MN116740)		Emiya (Red11-TH) (OP326312)		Epara (KT008005)	
		No	%	No	%	No	%	No	%	No	%	No	%	No	%
Ala	GCG	22	0.65	22	0.65	19	0.56	14	0.42	15	0.44	14	0.42	19	0.56
	GCA	15	0.44	16	0.47	11	0.33	20	0.59	19	0.56	20	0.59	16	0.47
	GCT	89	2.64	85	2.52	69	2.05	68	2.01	67	1.99	66	1.96	71	2.10
	GCC	9	0.27	10	0.30	9	0.27	12	0.36	12	0.36	13	0.39	13	0.39
Cys	TGT	89	2.64	88	2.61	107	3.17	97	2.87	99	2.93	99	2.93	98	2.90
	TGC	13	0.39	14	0.42	7	0.21	9	0.27	6	0.18	7	0.21	10	0.30
Asp	GAT	58	1.72	59	1.75	64	1.90	69	2.04	69	2.04	70	2.07	66	1.95
	GAC	7	0.21	6	0.18	5	0.15	7	0.21	6	0.18	5	0.15	6	0.18
Glu	GAG	56	1.66	58	1.72	56	1.66	53	1.57	52	1.54	53	1.57	54	1.60
	GAA	25	0.74	23	0.68	20	0.59	19	0.56	20	0.59	19	0.56	23	0.68
Phe	TTT	330	9.77	329	9.74	359	10.65	363	10.75	362	10.72	363	10.75	349	10.33
	TTC	21	0.62	22	0.65	21	0.62	20	0.59	19	0.56	20	0.59	34	1.01
Gly	GGG	115	3.41	120	3.55	78	2.31	73	2.16	74	2.19	74	2.19	62	1.83
	GGA	35	1.04	33	0.98	27	0.80	28	0.83	28	0.83	27	0.80	33	1.00
	GGT	126	3.73	123	3.64	157	4.66	162	4.80	164	4.86	163	4.83	169	5.00
	GGC	9	0.27	11	0.33	19	0.56	22	0.65	19	0.56	21	0.62	16	0.47
His	CAT	45	1.33	45	1.33	46	1.37	40	1.19	42	1.24	42	1.24	42	1.24
	CAC	6	0.18	5	0.15	8	0.24	15	0.44	13	0.39	13	0.39	12	0.35
Ile	ATA	57	1.69	56	1.66	76	2.26	85	2.52	85	2.52	85	2.52	94	2.78
	ATT	140	4.15	140	4.15	127	3.77	122	3.61	121	3.58	120	3.56	119	3.52
	ATC	16	0.47	14	0.42	11	0.33	11	0.33	11	0.33	12	0.36	8	0.24
Lys	AAG	48	1.42	48	1.42	49	1.45	48	1.42	48	1.42	48	1.42	51	1.51
Leu	TTG	267	7.91	270	7.99	258	7.65	235	6.96	231	6.84	240	7.11	244	7.22
	TTA	157	4.65	155	4.59	168	4.98	196	5.81	201	5.95	189	5.60	176	5.21
	CTG	29	0.86	30	0.89	20	0.59	19	0.56	19	0.56	20	0.59	18	0.53
	CTA	15	0.44	12	0.35	10	0.30	15	0.44	14	0.42	16	0.47	16	0.47
	CTT	65	1.93	63	1.87	88	2.61	72	2.13	73	2.16	70	2.07	83	2.46
	CTC	7	0.21	9	0.27	5	0.15	7	0.21	7	0.21	7	0.21	7	0.21
Met	ATG	110	3.26	112	3.32	102	3.03	103	3.05	104	3.08	104	3.08	106	3.14
Asn	AAA	30	0.89	29	0.86	29	0.86	30	0.90	30	0.89	30	0.89	25	0.74
	AAT	44	1.30	45	1.33	47	1.39	45	1.33	45	1.33	45	1.33	52	1.54
	AAC	5	0.15	5	0.15	6	0.18	8	0.24	9	0.27	8	0.24	5	0.15
Pro	CCG	14	0.42	15	0.44	21	0.62	9	0.27	9	0.27	9	0.27	9	0.27
	CCA	7	0.21	7	0.21	18	0.53	18	0.53	18	0.53	18	0.53	18	0.53
	CCT	56	1.66	55	1.63	43	1.28	47	1.39	48	1.42	49	1.45	49	1.45
	CCC	15	0.44	16	0.47	13	0.39	22	0.65	21	0.62	21	0.62	18	0.53
Gln	CAG	21	0.62	21	0.62	21	0.62	19	0.56	19	0.56	19	0.56	19	0.56
	CAA	7	0.21	7	0.21	6	0.18	7	0.21	7	0.21	7	0.21	7	0.21
Arg	CGG	16	0.47	16	0.47	15	0.45	13	0.39	13	0.39	13	0.39	8	0.24
	CGA	4	0.12	4	0.12	6	0.18	6	0.18	6	0.18	6	0.18	9	0.27
	CGT	41	1.21	42	1.24	40	1.19	42	1.24	43	1.27	43	1.27	40	1.18
	CGC	2	0.06	1	0.03	1	0.03	2	0.06	1	0.03	1	0.03	6	0.18
Ser	AGG	52	1.54	52	1.54	41	1.22	39	1.16	38	1.13	40	1.19	37	1.10
	AGA	19	0.56	18	0.53	23	0.68	29	0.86	29	0.86	27	0.80	23	0.68
	AGT	84	2.49	83	2.46	87	2.58	87	2.58	85	2.52	88	2.61	95	2.81
	AGC	11	0.33	11	0.33	7	0.21	7	0.21	7	0.21	5	0.15	11	0.33
	TCG	23	0.68	26	0.77	12	0.36	13	0.39	14	0.42	15	0.44	11	0.33
	TCA	22	0.65	21	0.62	15	0.45	27	0.80	25	0.74	24	0.71	17	0.50
	TCT	135	4.0	133	3.94	144	4.27	143	4.24	145	4.29	143	4.24	149	4.41
	TCC	14	0.42	13	0.39	21	0.62	10	0.30	10	0.30	11	0.33	11	0.33
Thr	ACG	20	0.59	21	0.62	20	0.59	20	0.59	19	0.56	19	0.56	17	0.50
	ACA	13	0.39	14	0.42	15	0.45	16	0.47	16	0.47	16	0.47	16	0.48
	ACT	56	1.66	56	1.66	46	1.37	39	1.16	41	1.21	39	1.16	44	1.30
	ACC	4	0.12	4	0.12	9	0.27	17	0.05	16	0.47	18	0.53	9	0.30
Val	GTG	105	3.11	101	2.99	71	2.11	68	2.01	68	2.01	69	2.04	77	2.28
	GTA	49	1.45	52	1.54	62	1.84	56	1.66	57	1.69	56	1.66	49	1.45
	GTT	221	6.54	221	6.54	236	7.00	237	7.02	242	7.17	241	7.14	235	7.00
	GTC	16	0.47	19	0.56	16	0.48	11	0.33	9	0.27	10	0.30	9	0.27
Trp	TGG	77	2.28	76	2.25	67	1.99	65	1.93	64	1.90	63	1.87	60	1.78
	TGA	36	1.07	37	1.01	42	1.25	42	1.24	44	1.30	45	1.33	50	1.48
Tyr	TAT	148	4.38	148	4.38	152	4.51	160	4.74	159	4.71	161	4.77	149	4.41
	TAC	17	0.50	18	0.53	11	0.33	6	0.18	7	0.21	5	0.15	17	0.50
stop	TAG	11	0.33	12	0.35	12	0.36	10	0.30	10	0.30	10	0.30	8	0.24
	TAA	1	0.03	0	0.00	0	0	2	0.06	2	0.06	2	0.06	4	0.12

Supplementary Table S3.1 (continued)

*aa: amino acid abbreviation according to DDBJ (http://www.ddbj.nig.ac.jp/sub/ref2-e.html); stop: stop codon.

Amino acid*	Cod-on	Eacon (Chany-RU) (ON644993)		Erevo (MSD15-TH) (MN496162)		Echinostoma sp. (GD-CN) (MN116706)		Echinostoma sp. (JM2019-CN) (MH212284)		Echino-stomatidae sp. (CA2021PE4-US) (MK264774)		Echino-stomatidae sp. (MSBA19-US) (MN822299)		Hcono (Hubei-CN) (KM111525)		Hcono (RED42-TH) (PP110501)	
		No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%
Ala	GCG	14	0.42	11	0.33	25	0.74	26	0.77	36	1.07	23	0.68	29	0.86	33	0.98
	GCA	26	0.77	20	0.59	17	0.50	16	0.47	13	0.39	22	0.65	17	0.50	16	0.47
	GCT	73	2.17	71	2.11	83	2.45	81	2.40	73	2.16	72	2.13	64	1.90	66	1.96
	GCC	7	0.21	9	0.27	11	0.33	10	0.30	12	0.36	11	0.33	16	0.47	15	0.45
Cys	TGT	102	3.03	96	2.84	114	3.37	101	2.99	102	3.02	102	3.02	98	2.91	92	2.73
	TGC	12	0.36	10	0.30	14	0.41	14	0.42	10	0.30	8	0.24	11	0.33	16	0.47
Asp	GAT	68	2.02	67	1.98	67	1.98	66	1.96	68	2.01	64	1.90	65	1.93	64	1.90
	GAC	5	0.15	7	0.21	3	0.09	3	0.09	4	0.12	8	0.24	4	0.12	4	0.12
Glu	GAG	45	1.34	51	1.51	54	1.60	54	1.60	54	1.60	44	1.30	60	1.78	60	1.78
	GAA	27	0.08	26	0.77	25	0.74	25	0.74	16	0.47	29	0.86	15	0.45	16	0.47
Phe	TTT	305	9.05	346	10.24	336	9.93	338	10.02	313	9.26	310	9.18	304	9.02	302	8.96
	TTC	38	1.13	27	0.80	27	0.80	27	0.80	31	0.92	32	0.95	39	1.16	40	1.19
Gly	GGG	87	2.58	74	2.19	108	3.19	105	3.11	108	3.19	98	2.90	98	2.91	96	2.85
	GGA	32	0.95	28	0.83	38	1.12	37	1.01	29	0.86	35	1.04	41	1.22	41	1.22
	GGT	144	4.27	175	5.18	124	3.67	130	3.85	144	4.26	133	3.94	134	3.97	135	4.00
	GGC	15	0.45	7	0.21	21	0.62	22	0.65	14	0.41	23	0.68	21	0.62	23	0.68
His	CAT	45	1.34	45	1.33	49	1.45	49	1.45	45	1.33	42	1.24	44	1.31	43	1.28
	CAC	9	0.27	9	0.27	4	0.12	4	0.12	9	0.27	9	0.27	8	0.24	9	0.27
Ile	ATA	71	2.11	92	2.72	55	1.63	57	1.69	70	2.07	61	1.81	58	1.72	58	1.72
	ATT	139	4.12	122	3.61	126	3.73	125	3.71	137	4.05	146	4.33	126	3.74	126	3.74
	ATC	18	0.53	11	0.33	13	0.38	16	0.47	13	0.39	18	0.53	20	0.59	23	0.68
Lys	AAG	48	1.42	48	1.42	47	1.39	49	1.45	46	1.36	48	1.42	47	1.39	47	1.39
Leu	TTG	214	6.35	204	6.04	273	8.07	274	8.12	253	7.48	216	6.40	285	8.45	287	8.51
	TTA	236	7.00	214	6.34	138	4.08	131	3.88	175	5.18	222	6.58	146	4.33	144	4.27
	CTG	27	0.80	15	0.44	24	0.71	23	0.68	34	1.01	26	0.77	39	1.16	40	1.19
	CTA	23	0.68	20	0.59	18	0.53	18	0.53	22	0.65	16	0.47	17	0.50	16	0.47
	CTT	47	1.39	85	2.52	77	2.28	77	2.28	51	1.51	56	1.66	62	1.84	63	1.87
	CTC	9	0.27	9	0.27	11	0.33	11	0.33	7	0.21	11	0.33	4	0.12	4	0.12
Met	ATG	115	3.41	108	3.20	105	3.10	105	3.11	100	2.96	99	2.93	111	3.29	110	3.26
Asn	AAA	17	0.5	29	0.86	25	0.74	25	0.74	15	0.44	15	0.44	18	0.53	19	0.56
	AAT	62	1.84	45	1.33	51	1.51	49	1.45	50	1.48	53	1.57	52	1.54	51	1.51
	AAC	4	0.12	6	0.18	8	0.24	7	0.21	7	0.21	6	0.18	6	0.18	7	0.21
Pro	CCG	15	0.45	10	0.30	26	0.77	24	0.71	14	0.41	22	0.65	21	0.62	21	0.62
	CCA	11	0.33	20	0.59	18	0.53	16	0.47	22	0.65	18	0.53	10	0.30	10	0.30
	CCT	51	1.51	39	1.16	48	1.42	49	1.45	38	1.12	42	1.24	42	1.25	42	1.25
	CCC	20	0.59	25	0.74	8	0.24	10	0.30	21	0.62	15	0.44	24	0.71	23	0.68
Gln	CAG	19	0.56	19	0.56	20	0.59	22	0.65	22	0.65	23	0.68	18	0.53	18	0.53
	CAA	10	0.30	8	0.24	2	0.06	2	0.06	5	0.15	5	0.15	10	0.30	10	0.30
Arg	CGG	13	0.39	8	0.24	15	0.44	15	0.44	15	0.44	16	0.47	14	0.42	15	0.45
	CGA	7	0.21	8	0.24	5	0.15	5	0.15	10	0.30	13	0.39	3	0.09	3	0.09
	CGT	41	1.22	45	1.33	41	1.21	42	1.25	36	1.07	33	0.98	44	1.31	42	1.25
	CGC	1	0.03	1	0.03	2	0.06	2	0.06	2	0.06	1	0.03	2	0.06	2	0.06
Ser	AGG	47	1.39	31	0.92	59	1.74	58	1.72	49	1.45	52	1.54	62	1.84	62	1.84
	AGA	40	1.19	29	0.86	23	0.68	22	0.65	24	0.71	33	0.98	23	0.68	22	0.65
	AGT	70	2.08	97	2.87	58	1.71	60	1.78	77	2.28	62	1.84	70	2.08	70	2.08
	AGC	10	0.30	8	0.24	7	0.21	8	0.24	14	0.41	13	0.39	13	0.39	11	0.33
	TCG	23	0.68	10	0.30	26	0.77	26	0.77	19	0.56	18	0.53	30	0.89	30	0.89
	TCA	26	0.77	31	0.92	19	0.56	19	0.56	32	0.95	30	0.90	20	0.59	19	0.56
	TCT	120	3.56	141	4.17	132	3.90	130	3.85	130	3.85	127	3.76	123	3.65	121	3.59
	TCC	19	0.56	14	0.41	15	0.44	14	0.42	14	0.41	23	0.68	20	0.59	24	0.71
Thr	ACG	11	0.33	11	0.33	23	0.68	24	0.71	17	0.50	21	0.62	23	0.68	24	0.71
	ACA	15	0.45	16	0.47	13	0.38	12	0.36	16	0.47	15	0.44	15	0.45	15	0.45
	ACT	66	1.96	50	1.48	52	1.54	53	1.57	49	1.45	49	1.45	46	1.36	47	1.39
	ACC	5	0.15	12	0.36	6	0.18	6	0.18	9	0.27	8	0.24	10	0.30	9	0.27
Val	GTG	94	2.79	79	2.34	109	3.22	108	3.20	106	3.14	99	2.93	111	3.29	108	3.20
	GTA	55	1.63	53	1.57	43	1.27	42	1.25	66	1.95	68	2.01	57	1.69	55	1.63
	GTT	191	5.67	219	6.48	240	7.09	240	7.11	196	5.80	203	6.01	201	5.96	196	5.81
	GTC	19	0.56	21	0.62	7	0.21	7	0.21	18	0.53	20	0.59	20	0.59	25	0.74
Trp	TGG	57	1.69	54	1.60	96	2.84	78	2.31	75	2.22	66	1.96	77	2.28	77	2.28
	TGA	49	1.45	52	1.54	0	0	29	0.86	32	0.95	39	1.16	29	0.86	29	0.86
Tyr	TAT	146	4.33	157	4.65	160	4.73	157	4.65	168	4.97	156	4.62	145	4.30	146	4.33
	TAC	24	0.71	11	0.33	7	0.21	7	0.21	12	0.36	16	0.47	18	0.53	18	0.53
stop	TAG	9	0.27	7	0.21	8	0.24	8	0.24	10	0.30	7	0.21	9	0.27	9	0.27
	TAA	3	0.09	5	0.15	4	0.12	4	0.12	2	0.06	5	0.15	3	0.09	3	0.09

Amala: *Artyfechinostomum malayanum* (synonym: *Echinostoma malayanum*); Asufr: *Artyfechinostomum sufrartyfex*; Ecapr: *Eca. caproni*; Emiya: *Eca. miyagawai*; Epara: *Eca. paraensei*; Eacon: *Echinoparyphium aconiatum*; Erevo: *Eca. revolutum*; Hcono: *Hypoderaeum conoideum*.

=======000========