Parasitology Research (2019) 118:2203–2211
https://doi.org/10.1007/s00436-019-06358-y

GENETICS, EVOLUTION, AND PHYLOGENY - ORIGINAL PAPER

# Characterization of the complete mitochondrial genome of *Uvitellina* sp., representative of the family Cyclocoelidae and phylogenetic implications

Suleman[1] · Mian Sayed Khan[2] · Petr Heneberg[3] · Cheng-Yan Zhou[4] · Nehaz Muhammad[1] · Xing-Quan Zhu[1,4] · Jun Ma[1]

Received: 2 February 2019 / Accepted: 16 May 2019 / Published online: 1 June 2019


## Abstract

Mitochondrial (mt) DNA has been useful in revealing the phylogenetic relationship of eukaryotic organisms including flatworms. Therefore, the use of mitogenomic data for the comparative and phylogenetic purposes is needed for those families of digenetic trematodes for which the mitogenomic data are still missing. Molecular data with sufficiently rich informative characters that can better resolve species identification, discrimination, and membership in different genera is also required for members of some morphologically difficult families of trematodes bearing few autapomorphic characters among its members. Here, the internal transcribed spacer (ITS) region of nuclear ribosomal DNA (rDNA) and the complete mt genome of the trematode *Uvitellina* sp. (Cyclocoelidae: Haematotrephinae) was determined and annotated. The mt genome of this avian trematode is 14,217 bp in length, containing 36 genes plus a single non-coding region. The ITS rDNA sequences were used for the pairwise sequence comparison of *Uvitellina* sp. with European cyclocoelid species, and the mitochondrial 12 protein-coding genes (PCGs) and two ribosomal RNA genes were used to evaluate the position of the family within selected trematodes. The ITS rDNA analysis of *Uvitellina* sp. showed less nucleotide differences with *Hyptiasmus oculeus* (16.77%) than with other European cyclocoelids (18.63–23.58%). The Bayesian inference (BI) analysis using the 12 mt PCGs and two rRNA genes supported the placement of the family Cyclocoelidae within the superfamily Echinostomatoidea (Plagiorchiida: Echinostmata). The availability of the mt genome sequences of *Uvitellina* sp. provides a novel resource of molecular markers for phylogenetic studies of Cyclocoelidae and other trematodes.

**Keywords** Cyclocoelidae · *Uvitellina* sp. · Nuclear ribosomal DNA · Mitochondrial genome · Phylogenetic implications

Section Editor: Neil Bruce Chilton

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00436-019-06358-y) contains supplementary material, which is available to authorized users.

✉ Jun Ma
dreamerjm@163.com

[1] State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu Province 730046, People's Republic of China

[2] Department of Zoology, University of Swabi, Khyber Pakhtunkhwa, Pakistan

[3] Third Faculty of Medicine, Charles University, 100 00 Prague, Czech Republic

[4] Jiangsu Co-innovation Center for the Prevention and Control of Important Animal Infectious Diseases and Zoonoses, Yangzhou University College of Veterinary Medicine, Yangzhou, Jiangsu Province 225009, People's Republic of China

## Introduction

The family Cyclocoelidae Stossich, 1902, contains medium-sized to large trematodes, which parasitize the airways, air sacs, nasal and infraorbital sinuses, hypothalamus, liver and body cavity of their avian hosts that are feeding on molluscs (Kanev et al. 2002; Dronen and Blend 2015). For the life cycle of these avian trematodes, freshwater pulmonate snails or xerothermic snails serve as both first and second intermediate hosts (Poulin and Cribb 2002; Dronen and Blend 2015). The tailless cercariae do not emerge from the snails; instead, they encyst either in rediae or in the tissue of the snail, develop to metacercariae and await ingestion by the avian host (Poulin and Cribb 2002; Dronen and Blend 2015). Thus, birds feeding on molluscs, including waders, get infected upon ingestion of an infected snail (Kanev et al. 2002; Shutler et al. 2016).

The classification of Cyclocoelidae underwent many changes over time with the proposal of additional subfamilies, genera

and species. Dronen and Blend (2015) published an extensive monograph with keys to the six subfamilies, 22 genera and 128 species. The subfamily Haematotrephinae was characterized by having the ovary ranging from pretesticular in position to opposite the anterior testis forming a triangle with the testes (Dronen and Blend 2015). Within Haematotrephinae, the genus *Uvitellina* Witenberg, 1923, was distinguished from other four genera by having a distinctly pretesticular ovary forming a triangle with the testes, postpharyngeal genital pore and posteriorly confluent vitelline fields (Dronen et al. 2017). Dronen and Blend (2015) assigned nine species to *Uvitellina* i.e. *Uvitellina pseudocotylea* Witenberg, 1923, *Uvitellina macroisophaga* Hannum and Wilson, 1934, *Uvitellina vanelli* Rudolphi, 1819, *Uvitellina titiri* Chatterji, 1958, *Uvitellina simile* Stossich, 1902, *Uvitellina kaniharensis* Gupta, 1958, *Uvitellina adelphus* Johnston, 1917, *Uvitellina iraquensis* Dronen, Ali and Al-Amura, 2013, and *Uvitellina himantopi* Dronen and Tkach, 2013.

All the hitherto proposed classification systems of Cyclocoelidae are primarily based on morphological characters, but the family has been recognized as a "morphologically difficult group" (Dronen et al. 2017), "unsettled taxonomic group" (Dronen and Blend 2015; Sitko et al. 2017) and "controversial family" (Kalani et al. 2015). Among the main reasons which lead to misidentification are the presence of their low numbers in definitive hosts, low chance of regaining specimens studied previously, morphological similarity between genera, for example, *Cyclocoelum*, *Harrahium* and *Hyptiasmus*, and higher similarities in their developmental cycles (Dronen and Blend 2015; Sitko et al. 2006, 2017). Thus, molecular studies could benefit more for the family Cyclocoelidae than probably any other family of flukes (Dronen et al. 2017). Surprisingly, few studies have been done on the molecular phylogenetics, positioning the Cyclocoelidae within the higher taxa based on nuclear ribosomal DNA (rDNA) (Littlewood and Bray 2001; Olson et al. 2003; Libert et al. 2012; Tkach et al. 2016). Sitko et al. (2017) performed the phylogenetic analyses of central European cyclocoelids based on two nuclear (ITS2, 18S rDNA) and two mitochondrial (*cox*1, *nad*1) DNA loci. However, for the genus *Uvitellina* (Haematotrephinae), there are no sequences available, neither from nuclear nor from mitochondrial DNA.

The mitochondrial DNA (mtDNA) is an effective and rich source of genetic markers (Smith 2016; Locke et al. 2018) for molecular identification and phylogenetics of many organisms. However, mt genomes of only a limited number of trematodes have been sequenced. The mitogenomic data is still missing for some morphologically difficult groups of trematodes including Cyclocoelidae. Here, we determined the internal transcribed spacer region of nuclear ribosomal DNA (ITS1-5.8S-ITS2 rDNA) and the complete mt genome of *Uvitellina* sp. as the representative of the family Cyclocoelidae, examined the gene arrangement and composition and conducted a phylogenetic analysis with other selected trematodes using the amino acid sequences of 12 mitochondrial protein-coding genes (PCGs).

## Materials and methods

### Sampling, identification and DNA extraction

Three cyclocoelid specimens were collected from the air sacs (lungs) of their definitive host, the red-wattled Lapwing (*Vanellus indicus*) (Boddaert, 1783) (Charadriiformes: Charadriidae) in Swabi, Khyber Pakhtunkhwa (KP), Pakistan. The live worms were killed and relaxed in hot water and stored in 80% ethanol (Lutz et al. 2017). For light microscopy, a specimen was stained, dehydrated and mounted permanently according to the recommended protocol (Lutz et al. 2017) and identified to the genus level following the existing identification keys and descriptions of cyclocoelids (Kanev et al. 2002; Dronen 2007; Dronen and Blend 2015; Gupta 1958; Siddiqi and Jairajpuri 1962). For molecular work, the total genomic DNA was extracted from 80% ethanol-preserved specimens according to a published protocol (Gasser et al. 2007).

### ITS rDNA sequencing and comparison with European cyclocoelids

The nuclear ribosomal DNA region spanning the 3′ end of the 18S rDNA gene, the complete ITS (ITS1, 5.8S, ITS2) and the 5′ end of the 28S rDNA gene was amplified by PCR using two sets of primers: BD1 (forward; 5′-GTCGTAACAAGGTTTCCGTA-3′) and BD2 (reverse; 5′-ATGCTTAAATTCAGCGGGT-3′) (Morgan and Blair 1995) and NC13 (ITS2)/F (5′-ATCGATGAAGAACGCAGC-3′) (Jacobs et al. 1997) and Dd28SR1 (5′-ACAAACAACCCGACTCCAAG-3′) (Otranto et al. 2007). The positive amplicons were purified and subjected to Sanger sequencing by GENEWIZ sequencing company (Beijing, China). While the complete ITS rDNA sequences were not available for all cyclocoelids, the nuclear ribosomal DNA region spanning partial 5.8S, the complete ITS2 and the partial 28S rDNA has been sequenced for European cyclocoelids (Sitko et al. 2017). Therefore, this nuclear rDNA region was used for comparative purposes with other European cyclocoelids. The nucleotide differences in percent (including gaps within alignment) across the rDNA region (mentioned above) were estimated using pairwise alignment in ClustalX 1.83 (Thompson et al. 1997) and BioEdit 7.0.9.0 (Hall 1999).

### Long-PCR and mt genome sequencing

Universal PCR primers for flatworms were used for the amplification of partial *cyt*b, *nad*1, *nad*4, *cox*1, *rrn*L-*rrn*S and

*nad*5 regions in order to obtain mt gene sequence data for further design of primers (Yang et al. 2015; Le et al. 2016). The entire mt genome was amplified and sequenced in six medium-sized overlapping fragments using six pairs of newly designed primers (Additional file 1: Table S1). Long-PCR reactions were conducted in a total volume of 28 μl, using 12.5 μl PrimeSTAR Max DNA polymerase (Takara, Dalian, China), 12.5 μl dd$H_2O$, 1 μl of each primer (25 pmol) and 1 μl of total genomic DNA. The thermocycling profile included an initial denaturation step at 98 °C for 2 min, followed by 10 cycles of denaturation at 92 °C for 15 s, annealing at 50–55 °C for 30 s, extension at 60 °C for 1–3 min, followed by 2 min denaturation at 92 °C, then 22 cycles of 92 °C for 10 s, 50–55 °C for 30 s and 66 °C for 1–3 min with 10 min of final extension at 68 °C. PCR products yielding desired bands on a 1.0% (*w*/*v*) agarose gel were considered as positive. The positive amplicons were sent to Sangon Company (Shanghai, China) for sequencing.

### Sequence assembly and mitogenome characterization

Contiguous sequences were manually assembled and aligned against each other using DNAstar v7.1 program (Burland 2000), and then aligned against the entire mt genomes of relevant trematode species (echinochasmid and fasciolids) using MAFFT (Katoh and Standley 2013) to infer the approximate gene boundaries. The 12 mitochondrial protein-coding genes (PCGs) were identified by searching for ORFs or by comparison with the entire mt genome sequences of *Fasciola gigantica* (KF543342), *F. hepatica* (AF216697) and *Echinochasmus japonicus* (KP844722). The two rRNA genes were determined by the same technique, via comparison with homologues in MAFFT. The length and putative secondary structures of 22 tRNA genes were identified using ARWEN (Laslett and Canback 2008) and MITOS (Bernt et al. 2013) or detected by the scrutiny of the sequences aligned with that of the other relevant trematodes. PhyloSuite v1.1.14 (Zhang et al. 2018), a GUI-based software, was used to translate the PCGs into their corresponding amino acids and to calculate codon usage and relative synonymous codon usage (RSCU) of 12 PCGs. The RSCU figure was drawn using the ggplot2 (Wickham 2016), plugin of PhyloSuite. The nucleotide content of *Uvitellina* sp. mitogenomes was compared with other selected trematodes using PhyloSuite. The entire mt genome size, A + T/G + C contents and codon usage (for 12 PCGs) were also used for comparative purposes. A single non-coding region (NCR) was identified between *trn*G and *cox*3. The secondary structures of NCR were predicted using Mfold software (Zuker 2003).

### Phylogenetic analyses

Phylogenetic analyses were conducted using the newly sequenced *Uvitellina* sp. and 17 other digenean mitogenomes, classified in suborder Echinostomata and Pronocephalata (Olson et al. 2003; Tkach et al. 2016), with *Schistosoma japonicum* (AF215860) as an outgroup, thus adding 19 mitogenomes in total. The nucleotide sequences of 12 PCGs were extracted from GenBank files and translated to their corresponding amino acid sequences using PhyloSuite.

Phylogenetic analyses were conducted using two datasets: dataset 1 containing nucleotide sequences of 12 PCGs + two rRNAs and dataset 2 containing the amino acid sequences of 12 PCGs. The analyses were conducted in PhyloSuite as follows: nucleotide and amino acid sequences of each gene were aligned in batches followed by the deletion of ambiguously aligned regions using Gblocks (Talavera and Castresana 2007), the nucleotide (dataset 1) and amino acid (dataset 2) sequences were then concatenated and subjected to the ModelFinder (Kalyaanamoorthy et al. 2017), as implemented in PhyloSuite. Based on Akaike Information Criterion, GTR+F+I+G4 was chosen as the best-fitting model for dataset 1 while JTT+F+I+G4 was chosen as the best-fitting model for dataset 2. Bayesian inference (BI) analysis was carried out in MrBayes 3.2.6 (Ronquist et al. 2012), implemented in PhyloSuite, with default setting and $5 \times 10^6$ metropolis-coupled MCMC generations for each dataset. The analyses were considered to be completed upon finding the average standard deviation of split frequencies below 0.01. Twenty-five percent (1250) of sample trees were discarded as burn-in. The phylograms were viewed in iTOL (Letunic and Bork 2016) and annotated in Adobe Illustrator®.

## Results

### Identification and differences in ITS rDNA with European cyclocoelids

The examined flukes from *V. indicus* agree with the subfamilial description of the Haematotrephinae Dollfus, 1948, because it has an opposite to slightly pretesticular ovary forming a triangle with the testes. Moreover, the presence of the rudimentary oral sucker, postpharyngeal genital pore and uterine coils overreaching the ceca extended somewhat posterior to the posterior testis and posteriorly confluent vitellaria with a tail-like extension placed it in the genus *Uvitellina* Witenberg, 1923. The species has close morphological and morphometric similarities with *U. indica* Siddiqi and Jairajpuri, 1962, recovered from the same definitive host, which was later synonymized with *U. kaniharensis* Gupta, 1958, by Dronen and Blend (2015). As a high level of homogeneity and phenotypic plasticity found in cyclocoelids

(Dronen and Blend 2015), thus in order to avoid the misidentification, we recognized it only to generic level as *Uvitellina* sp. (supposedly *U. kaniharensis*). The nucleotide sequences obtained for the ITS rDNA (1350 bp) of *Uvitellina* sp. were deposited in GenBank (GenBank accession No. MK227162). The difference in nucleotide sequences (including internal gaps) of *Uvitellina* sp. with *Harrahium tringae* was higher (23.58%) followed by the differences with *H. obscurum* (22.84%), *Cyclocoelum mutabile* (18.64%), *Morishitium polonicum* (18.63%) and *Hyptiasmus oculeus* (16.77%).

## Mitochondrial genome features and gene content

The complete mt genome of *Uvitellina* sp. (GenBank accession No. MK227160) is circular, 14,217 bp in length, containing 36 genes, including 12 PCGs, 22 transfer RNA genes and two ribosomal RNA genes (*rrn*L and *rrn*S) plus a single non-coding region (Fig. 1, Table 1). *atp*8 is absent as commonly reported in the mt genome of other flatworms. Genes are either overlapped or abut or separated by intergenic spacers (2–70 bp). The shared gene boundaries continuously observed between five genes, *trn*T-*rrn*L-*trn*C-*rrn*S-*cox*2 without intergenic spacers. Overlap of 40 bp between *nad*4L and *nad*4 in the mt genome of *Uvitellina* sp. is consistent with that of other flatworms. The nucleotide composition in the mt genome of *Uvitellina* sp. are 49.7% (T), 16.3% (A), 24.8% (G) and 9.2% (C), with 66% overall adenine + thymine (AT) content. The A+T content of the *Uvitellina* sp. mtDNA is higher than those of echinostomatids, an echinochasmid, fasciolids and a himasthlid published to date but similar to that of

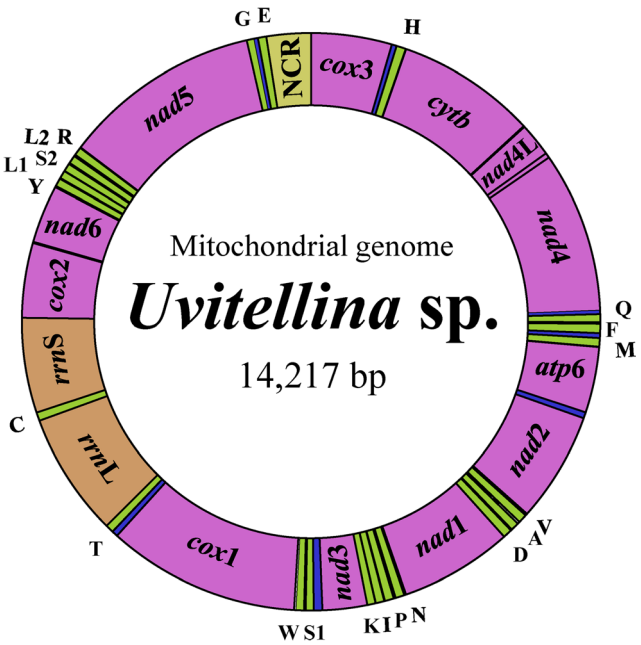


**Fig. 1** Organization of the complete mitochondrial genome of *Uvitellina* sp. Genes are differentiated by specific colours. Blue colour indicates the intergenic sequences (> 20 bp)

*O. sikae* (66.5%). The A+T content of protein-coding genes range from 63.9% (*cox*1) to 71.8% (*nad*6). In the rRNA genes, *rrn*L has higher A+T content (65.2%) than *rrn*S (62.5%) (Table 2).

In the *Uvitellina* sp. mt genome, the 12 PCGs use ATG and GTG as translation initiation codons and TAG or TAA as translation termination codons (Table 1). Incomplete or abbreviated stop codons were not used in *Uvitellina* sp. mtDNA. A total of 10, 200 bp was used to construct the 12 PCGs, encoding 3388 amino acids regardless of the overlap of nucleotides found between *cyt*b-*nad*4L (8 bp) and *nad*4L-*nad*4 (40 bp). The total number of codons was 3400 including 12 translation termination codons (TAG, eight instances, and TAA, four instances). The codon TTT was the most frequently used (12.79%, 435 instances), followed by TTG for leucine (9.94%, 338 instances), GTT for valine (8.91%, 303 instances) and TCT for serine (6.15%, 209 instances). The least commonly used codon is CGC-Arginine (one instance). The codons CTA-Leu, GCC-Ala, CAC-His, GAC-Asp, AGC-Ser and GGC-Gly are also least commonly used codons (two instances each), composed of nucleotides other than T, except CTA-Leu. The most commonly used codons were those ending with T (2,108, 62.03%), followed by those ending with G (876, 25.76%) and ending with A (342, 10.06%), but those having C at a third codon position counted only 74 (2.18%) in number. The overall codon usage and RSCU for the construction of all 12 PCGs are presented in Fig. 2. The most common amino acids were leucine ($L_1$, 2.77%, and $L_2$, 13.28%) and phenylalanine (13.37%) while glutamine (0.83%), lysine (1.51%) and histidine (1.53%) were the least used amino acids. The nucleotide composition in 12 PCGs of the mt genome of *Uvitellina* sp. in descending order is 52.2% (T), 24.4% (G), 14.7% (A) and 8.8% (C), with overall A+T content 66.9% (Table 2). The overall A+T content of the PCGs is nearly similar to the A+T content of the entire mt genome; however, the percent presence of the nucleotide T is higher in the PCGs. Among the 12 PCGs, *nad*5 is the longest (1596 bp) gene and *nad*4L is the shortest gene.

The mt genome of *Uvitellina* sp. encodes 22 tRNAs ranging from 62 to 76 bp in length, together accounting for 1751 bp. Twenty of these tRNAs can be folded into a typical four-armed "cloverleaf" structure, whereas the dihydrouridine (DHU) arm is missing in the remaining two tRNA genes, *trn*S1 (AGN) and *trn*S2 (UCN). The two ribosomal RNA genes (*rrn*L and *rrn*S) were recognized between the *trn*T and the *cox*2 genes, separated by *trn*C. The length of *rrn*L (16S) and *rrn*S (12S) in the mtDNA of *Uvitellina* sp. are 991 bp and 760 bp, respectively (Table 2). In the mt genome of *Uvitellina* sp., the single non-coding region is located between *trn*E and *cox*3 gene (Fig. 1) which is consistent with the mt genome of *H. conoideum*, *E. hortense* (Echinostomatidae) and *F. magna* and *F. buski* (Fasciolidae). The non-coding region is 355 bp in length with the A+T content of 63.94%. The

**Table 1** Organization and length of genes in the mitochondrial genome of *Uvitellina* sp.

| Gene/region | Position 5′ to 3′ | Length | | Ini/Ter codons | tRNA anti-codon | Int. seq. length (bp) |
|---|---|---|---|---|---|---|
| | | bp | aa | | | |
| *cox*3 | 1–645 | 645 | 214 | ATG/TAA | | + 39 |
| tRNA-His (H) | 685–760 | 76 | | | GTG | + 3 |
| *cyt*b | 764–1882 | 1119 | 372 | ATG/TAG | | − 8 |
| *nad*4L | 1875–2144 | 270 | 89 | GTG/TAG | | − 40 |
| *nad*4 | 2105–3397 | 1293 | 430 | ATG/TAG | | + 27 |
| tRNA-Gln (Q) | 3425–3487 | 63 | | | TTG | + 13 |
| tRNA-Phe (F) | 3501–3572 | 72 | | | GAA | + 39 |
| tRNA-Met (M) | 3612–3680 | 69 | | | CAT | + 3 |
| *atp*6 | 3684–4211 | 528 | 175 | ATG/TAG | | + 49 |
| *nad*2 | 4261–5142 | 882 | 293 | ATG/TAA | | + 17 |
| tRNA-Val (V) | 5160–5228 | 69 | | | TAC | + 20 |
| tRNA-Ala (A) | 5249–5314 | 66 | | | TGC | + 10 |
| tRNA-Asp (D) | 5325–5397 | 73 | | | GTC | 0 |
| *nad*1 | 5398–6300 | 903 | 300 | ATG/TAG | | + 15 |
| tRNA-Asn (N) | 6316–6386 | 71 | | | GTT | + 16 |
| tRNA-Pro (P) | 6403–6474 | 72 | | | TGG | + 5 |
| tRNA-Ile (I) | 6480–6548 | 69 | | | GAT | + 2 |
| tRNA-Lys (K) | 6551–6616 | 66 | | | CTT | 0 |
| *nad*3 | 6617–6973 | 357 | 118 | GTG/TAG | | + 70 |
| tRNA-SerAGN (S1) | 7044–7104 | 61 | | | GCT | + 14 |
| tRNA-Trp (W) | 7119–7186 | 68 | | | TCA | + 11 |
| *cox*1 | 7198–8739 | 1542 | 513 | ATG/TAG | | + 48 |
| tRNA-Thr (T) | 8788–8853 | 66 | | | TGT | 0 |
| *rrn*L | 8854–9844 | 991 | | | | 0 |
| tRNA-Cys (C) | 9845–9912 | 68 | | | GCA | 0 |
| *rrn*S | 9913–10,672 | 760 | | | | 0 |
| *cox*2 | 10,673–11,272 | 600 | 199 | ATG/TAA | | + 10 |
| *nad*6 | 11,283–11,747 | 465 | 154 | GTG/TAA | | + 10 |
| tRNA-Tyr (Y) | 11,758–11,822 | 65 | | | GTA | + 7 |
| tRNA-LeuCUN (L1) | 11,830–11,892 | 63 | | | TAG | − 2 |
| tRNA-SerUCN (S2) | 11,891–11,955 | 65 | | | TGA | + 10 |
| tRNA-LeuUUR (L2) | 11,966–12,030 | 65 | | | TAA | + 14 |
| tRNA-Arg (R) | 12,045–12,109 | 65 | | | TCG | 0 |
| *nad*5 | 12,110–13,705 | 1596 | 531 | GTG/TAG | | 0 |
| tRNA-Glu (E) | 13,706–13,767 | 62 | | | TTC | + 28 |
| tRNA-Gly (G) | 13,796–13,862 | 67 | | | TCC | |
| NCR | 13,863–14,217 | 355 | | | | |

*NCR* non-coding region, *bp* base pair, *aa* amino acid, *Ini/Ter* initial/terminal codons, *Int. seq.* intergenic sequences

NCR lacks any tandem repeats. The secondary structure of the NCR was predicted by Mfold software (Additional file 2: Fig. S1).

## Genetic relationships

Regardless of the dataset and model used, phylogenetic analyses produced phylograms with concordant branch topologies with minor changes in nodal support values; thus, only the tree based on nucleotide sequences of 12 PCGs + two rRNA genes was selected (because of strong nodal support) and shown (Fig. 3). The phylogenetic tree clearly indicates two major clades: one containing four families of the suborder Pronocephalata and the other containing five families of the suborder Echinostomata. The *Uvitellina* sp. represented the family

**Table 2** Nucleotide content and skewness of genes, individual elements and the complete mitogenome of *Uvitellina* sp.

| Regions | Size (bp) | T (U) | C | A | G | AT (%) | GC (%) | GT(%) | AT skew | GC skew |
|---|---|---|---|---|---|---|---|---|---|---|
| PCGs | 10,200 | 52.2 | 8.8 | 14.7 | 24.4 | 66.9 | 33.2 | 76.6 | −0.561 | 0.471 |
| 1st codon position | 3400 | 44.9 | 9.6 | 18.4 | 27.1 | 63.3 | 36.7 | 72 | −0.419 | 0.478 |
| 2nd codon position | 3400 | 49.8 | 14.5 | 15.6 | 20.1 | 65.4 | 34.6 | 69.9 | −0.524 | 0.163 |
| 3rd codon position | 3400 | 62 | 2.2 | 10.1 | 25.8 | 72.1 | 28 | 87.8 | −0.721 | 0.844 |
| *atp*6 | 528 | 53 | 9.1 | 15.9 | 22 | 68.9 | 31.1 | 75 | −0.538 | 0.415 |
| *cox*1 | 1542 | 47.6 | 10.6 | 16.3 | 25.5 | 63.9 | 36.1 | 73.1 | −0.489 | 0.414 |
| *cox*2 | 600 | 45 | 9 | 19.2 | 26.8 | 64.2 | 35.8 | 71.8 | −0.403 | 0.498 |
| *cox*3 | 645 | 54.4 | 8.5 | 12.4 | 24.7 | 66.8 | 33.2 | 79.1 | −0.629 | 0.486 |
| *cyt*b | 1119 | 51.7 | 8.7 | 14.5 | 25.1 | 66.2 | 33.8 | 76.8 | −0.563 | 0.487 |
| *nad*1 | 903 | 51.9 | 7.2 | 14.1 | 26.8 | 66 | 34 | 78.7 | −0.574 | 0.577 |
| *nad*2 | 882 | 55.6 | 8.3 | 13.7 | 22.4 | 69.3 | 30.7 | 78 | −0.604 | 0.461 |
| *nad*3 | 357 | 53.8 | 6.2 | 14.6 | 25.5 | 68.4 | 31.7 | 79.3 | −0.574 | 0.611 |
| *nad*4L | 270 | 52.6 | 5.2 | 14.8 | 27.4 | 67.4 | 32.6 | 80 | −0.56 | 0.682 |
| *nad*4 | 1293 | 52.5 | 8.7 | 13.5 | 25.2 | 66 | 33.9 | 77.7 | −0.59 | 0.485 |
| *nad*5 | 1596 | 54.4 | 9.8 | 14.2 | 21.6 | 68.6 | 31.4 | 76 | −0.587 | 0.377 |
| *nad*6 | 465 | 58.5 | 7.1 | 13.3 | 21.1 | 71.8 | 28.2 | 79.6 | −0.629 | 0.496 |
| *rrn*L | 991 | 42.7 | 10.9 | 22.5 | 23.9 | 65.2 | 34.8 | 66.6 | −0.31 | 0.374 |
| *rrn*S | 760 | 41.4 | 10.9 | 21.1 | 26.6 | 62.5 | 37.5 | 68 | −0.326 | 0.418 |
| tRNAs | 1481 | 41.2 | 11.6 | 21.5 | 25.7 | 62.7 | 37.3 | 66.9 | −0.315 | 0.378 |
| rRNAs | 1751 | 42.1 | 10.9 | 21.9 | 25.1 | 64 | 36 | 67.2 | −0.317 | 0.394 |
| Full genome | 14,217 | 49.7 | 9.2 | 16.3 | 24.8 | 66 | 34 | 74.5 | −0.505 | 0.459 |

*PCGs* protein-coding genes, *bp* base pairs

Cyclocoelidae. The tree placed the *Uvitellina* sp. on a separate branch (100% nodal support) as a distinct family in a clade containing members classified within Echinostomata including those of the family Fasciolidae, Echinostomatidae, Echinochasmidae and Himasthlidae. The phylogenetic position of *Uvitellina* sp. supports the placement of the family Cyclocoelidae within the superfamily Echinostomatoidea.

## Discussion

In the present study, the main features of the complete mt genome of *Uvitellina* sp. have been described and compared with previously published digenean mt genomes. The gene organization of the mt genome of *Uvitellina* sp. is similar to those of previously published echinostomatids (*H. conoideum* and *E. hortense*) (Yang et al. 2015; Liu et al. 2016) and

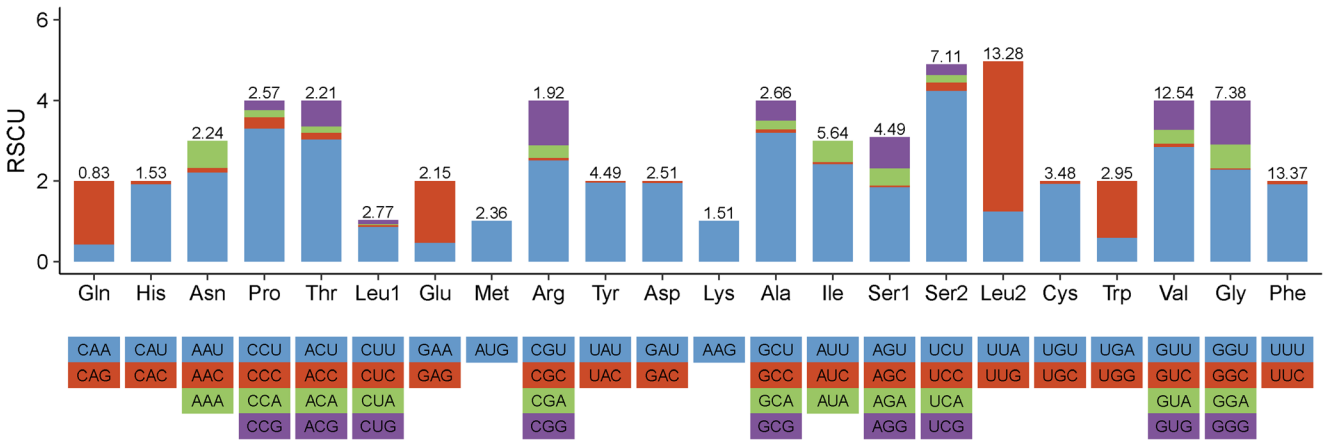


**Fig. 2** Relative Synonymous Codon Usage (RSCU) of 12 mitochondrial protein-coding genes of *Uvitellina* sp. codon families are labelled on the *x*-axis. Values on the top of each bar represent amino acid usage

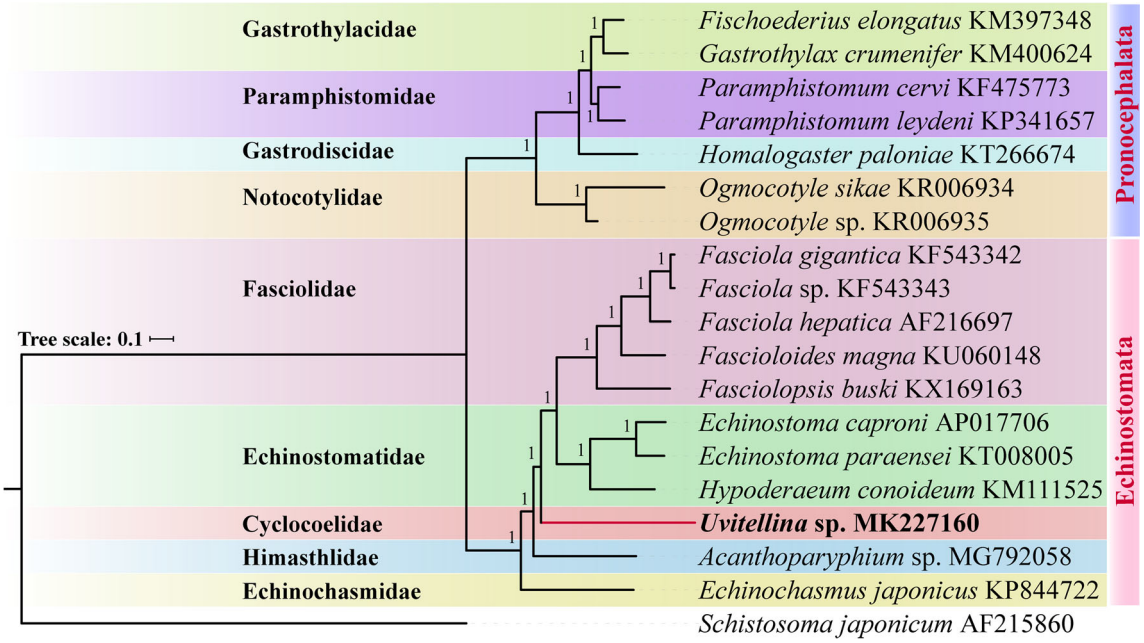


**Fig. 3** Phylogenetic relationships of *Uvitellina* sp. with selected trematodes based on Bayesian inference analysis (BI) of concatenated nucleotide sequences of 12 mitochondrial protein-coding genes and two rRNA genes. *Schistosoma japonicum* (AF215860) (Schistosomatidae: Strigeida) was included as the outgroup. GenBank accession numbers are shown to the right of each species. Families are highlighted in different colours

fasciolids (*F. magna* and *F. buski*) (Ma et al. 2016; Ma et al. 2017). However, it differs from some fasciolids (*F. hepatica*; *F. gigantica*; the intermediate form of *Fasciola*) and echinochasmid (*E. japonicus*) in possession of one non-coding region instead of two. The 12 PCGs exhibit a strong bias toward T and have the highest percent of the nucleotide T among the PCGs of all published trematodes species including schistosomatids (Additional file 3: Table S2). Similarly, the codon TTT-Phe is the most frequently used one with the highest percentage in all published mt genomes of trematode species.

Kanev et al. (2002) placed the family Cyclocoelidae within the superfamily Cyclocoeloidea Stossich, 1902. Later, Cribb et al. (2001) and Olson et al. (2003) classified the family Cyclocoelidae in the superfamily Echinostomatoidea based on molecular phylogenetic analyses of nuclear ribosomal DNA. Recently, the same systematic position of Cyclocoelidae was accepted in a conclusive phylogenetic analysis and systematics of the superfamily Echinostomatoidea, based on the nuclear lsrRNA (28S) gene (Tkach et al. 2016). Sitko et al. (2017) supported the phylogenetic classification of the family Cyclocoelidae within Echinostomata by observing the higher sequence similarities in four loci (included in their analyses) of cyclocoelids with members taken as outgroup from Echinostomatidae and Fasciolidae.

In the present study, the placement of the family Cyclocoelidae within Echinostomata has been well supported, based on the concatenated nucleotide sequences of 12 mitochondrial PCGs + two rRNA genes of *Uvitellina* sp., a representative of the family. The mitogenomic data from the families Fasciolidae, Echinostomatidae and Echinochasmidae (suborder Echinostomata) have been used for their systematics and taxonomic purposes (Le et al. 2001; Liu et al. 2014; Yang et al. 2015; Liu et al. 2016; Ma et al. 2016; Le et al. 2016; Ma et al. 2017). However, mtDNA data was not available for any member of the family Cyclocoelidae prior to the present study.

The cyclocoelidae, being an unsettled taxonomic group, need the molecular data from the nuclear or mtDNA from different genera and species for their correct identification and taxonomic position within this family. Sitko et al. (2017) conducted such phylogenetic analysis of central European cyclocoelids based on sequences of two nuclear (ITS2 and 28S) and two partial mitochondrial (*cox*1 and *nad*1) genes. In their analyses, they included the molecular data of *C. mutabile*, *C. obscurum*, *H. tringae*, *H. oculeus* and *M. polonicum*. Based on their results, Sitko et al. (2017) suggested the omission of the genus *Hyptiasmus* from the subfamily Hyptiasminae Dollfus, 1948, sensu Dronen and Blend (2015), and transferred the *C. obscurum* to the genus *Harrahium* as *H. obscurum* comb. n. sensu Sitko et al. (2017), being clustered with *H. tringae* in all of the four DNA loci. In the present study, the analysis of the rDNA region (partial 5.8S-ITS2-partial 28S) revealed closer genetic relatedness of *Uvitellina* sp. (Haematotrephinae) with the genus *Hyptiasmus* (83.23% similarity with

*H. oculeus*) than with the members of the genus *Harrahium* (Haematotrephinae), which is not consistent with the morphological classification of Dronen and Blend (2015). In fact, Dronen et al. (2017) also suggested the meaningful revision of the subfamilies and genera within Cyclocoelidae based on molecular studies. Here, we described the result of this study but do not assign *Uvitellina* sp. to any other existing subfamily. Therefore, to be able to refute the present taxonomic position of *Uvitellina*, nuclear rDNA and mtDNA datasets are required from more cyclocoelid species, especially from the type species of different genera including *Uvitellina*.

The lapwing has been recorded as the type host for many species of the genus *Uvitellina* and for two species of the genus *Haematotrephus* (Dronen and Blend 2015), which shows a high level of definitive host specificity of *Uvitellina* species. The posteriorly confluent vitelline fields in *Uvitellina* species distinguish it from the genus *Haematotrephus*; however, the other diagnostic characters of these two genera are almost the same (Dronen and Blend 2015). Thus, the molecular data from *Haematotrephus* species are also needed to better explore the genetic distance between these two genera of the subfamily Haemototrephinae. The mitogenomic data from other families of the suborder Echinostomata including Philphthalmidae and Psilostomidae is also required to better explore the inter- and intra-familial relationships of these trematodes.

## Conclusions

The present study determined the complete mt genome sequences of *Uvitellina* sp., the first of the genus *Uvitellina* and the family Cyclocoelidae. Phylogenetic analysis based on mitogenome sequences supported the classification of Cyclocoelidae within the superfamily Echinostomatoidea and suborder Echinostomata. The fully characterized mt genome sequences of *Uvitellina* sp. provide novel molecular sources for identification and systematic and ordinal phylogenetic studies of *Uvitellina* species and other cyclocoelids.

**Acknowledgements** We thank Dr. Jiljí Sitko (Comenius Museum, Přerov, Czechia) for the revision of the morphological identification of the examined material.

**Funding information** Project support was provided by the National Natural Science Foundation of China (Grant No. 31702225), the International Science and Technology Cooperation Project of Gansu Provincial Key Research and Development Program (Grant No. 17JR7WA031) and the Agricultural Science and Technology Innovation Program (ASTIP) (Grant No. CAAS-ASTIP-2016-LVRI-03).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no competing interests.

## References

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved *de novo* metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69:313–319

Burland TG (2000) DNASTAR's Lasergene sequence analysis software. Methods Mol Biol 132:71–91

Cribb TH, Bray RA, Littlewood DTJ, Pichelin S, Herniou EA (2001) Relationships of the Digenea – evidence from molecules and morphology. In: Littlewood DTJ, Bray RA (eds) Interrelationships of the Platyhelminthes. Taylor & Francis, London, pp 168–185

Dronen NO (2007) Revision of the family Cyclocoelidae Stossich, 1902 with the proposal of two new subfamilies and the description of a new species of *Morishitium* Witenberg, 1928 from the common snipe, *Gallinago gallinago*, from Texas, U.S.a. Zootaxa 1563:55–68

Dronen NO, Blend CK (2015) Updated keys to the genera in the subfamilies of Cyclocoelidae Stossich, 1902, including a reconsideration of species assignments, species keys and the proposal of a new genus in *Szidatitreminae* Dronen, 2007. Zootaxa 4053:1–100

Dronen NO, Al-Kassar NR, Ali AH, Abdulhameed MF, Abdullah BH, Al-Mayah SH (2017) Intraspecific variation in adult *Uvitellina iraquensis* Dronen, Ali and Al-Amura, 2013 (Cyclocoelidae: Haematotrephinae) from two collection sites of white-tailed lapwing, *Vanellus leucurus* (Lichtenstein) (Charadriiformes: Charadriidae), in Iraq. Zootaxa 4242:1–33

Gasser RB, Hu M, Chilton NB, Campbell BE, Jex AJ, Otranto D, Cafarchia C, Beveridge I, Zhu XQ (2007) Single strand conformation polymorphism (SSCP) for the analysis of genetic variation. Nat Protoc 1:3121–3128

Gupta PD (1958) On *Haematotrephus (Uvitellina) kaniharensis* n. sp. (Trematoda: Cyclocoelidae Kossack, 1911) from Allahabad. Indian J Helminthol 10:1–5

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser (Oxf) 41:95–98

Jacobs DE, Zhu X, Gasser RB, Chilton NB (1997) PCR-based methods for identification of potentially zoonotic ascaridoid parasites of the dog, fox and cat. Acta Trop 68:191–200

Kalani H, Kalani F, Faridnia R, Pestechian N, Mohaghegh MA, Mirzaei F, Sharif M, Daryani A (2015) The first report of *Morishitium vagum* (Trematoda: Cyclocoelidae) in *Turdus merula* in Iran. Helminthologia 52(3):293–296

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589

Kanev I, Radev V, Fried B (2002) Superfamily Cyclocoeloidea. In: Gibson DI, Jones A, Bray RA (eds) Keys to the Trematoda, vol 1. CABI Publishing and The Natural History Museum, Wallingford, pp 131–145

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

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

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

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

Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:242–245

Libert C, Jouet D, Ferte H, Lemberger K, Keck N (2012) Air sac fluke *Circumvitellatrema momota* in a captive blue-crowned motmot (*Momotus momota*) in France. J Zoo Wildl Med 43:689–692

Littlewood DTJ, Bray RA (2001) Interrelationships of the Platyhelminthes. CRC Press, Taylor & Francis, London, p 380

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

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 54:173–179

Locke SA, Van Dam A, Caffara M, Pinto HA, Lopez-Hernandez D, Blanar CA (2018) Validity of the Diplostomoidea and Diplostomida (Digenea, Platyhelminthes) upheld in phylogenomic analysis. Int J Parasitol 48:1043–1059

Lutz HL, Tkach VV, Weckstein JD (2017) Methods for specimen-based studies of avian symbionts. In: Webster MS (ed) The extended specimen: emerging frontiers in collections-based ornithological research. Studies in avian biology (no. 50). CRC Press, Boca Raton, pp 157–183

Ma J, He JJ, Liu GH, Leontovyc R, Kasny 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

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

Morgan JAT, Blair D (1995) Nuclear rDNA ITS sequence variation in the trematode genus *Echinostoma*: an aid to establishing relationships within the 37-collar-spine group. Parasitology 111:609–615

Olson PD, Cribb TH, Tkach VV, Bray RA, Littlewood DT (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int J Parasitol 33:733–755

Otranto D, Rehbein S, Weigl S, Cantacessi C, Parisi A, Lia RP, Olson PD (2007) Morphological and molecular differentiation between *Dicrocoelium dendriticum* (Rudolphi, 1819) and *Dicrocoelium chinensis* (Sudarikov and Ryjikov, 1951) Tang and Tang, 1978 (Platyhelminthes: Digenea). Acta Trop 104:91–98

Poulin R, Cribb TH (2002) Trematode life cycles: short is sweet? Trends Parasitol 18:176–183

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542

Shutler D, Mallory ML, Elderkin M, McLaughlin JD (2016) Cyclocoelid (*Morishitium* sp.) trematodes from an air sac of a purple sandpiper, *Calidris maritima* (Brunnich). J Parasitol 102:381–384

Siddiqi AH, Jairajpuri MS (1962) *Uvitellina indica* n. sp. (Trematoda: Cyclocoeliidae) from a redwattled lapwing, *Lobivanellus indicus* (Boddaert). Z Parasitenkd 21:212–214

Sitko J, Faltýnková A, Scholz T (2006) Checklist of the trematodes (Digenea) of birds of the Czech and Slovak Republics. Academia, Praha, p 111

Sitko J, Bizos J, Heneberg P (2017) Central European parasitic flatworms of the Cyclocoelidae Stossich, 1902 (Trematoda: Plagiorchiida): molecular and comparative morphological analysis suggests the reclassification of *Cyclocoelum obscurum* (Leidy, 1887) into the *Harrahium* Witenberg, 1926. Parasitology 144:368–383

Smith DR (2016) The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? Brief Funct Genomics 15:47–54

Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 56:564–577

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

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

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York

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

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

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