



Comparative Mitochondrial Genomic Analysis Robustly Supported That Cat Tapeworm *Hydatigera taeniaeformis* (Platyhelminthes: Cestoda) Represents a Species Complex

Hui-Mei Wang, Rong Li, Yuan-Ping Deng, Guo-Hua Liu* and Yi-Tian Fu*

Research Center for Parasites and Vectors, College of Veterinary Medicine, Hunan Agricultural University, Changsha, China

OPEN ACCESS

Edited by:

Urusa Thaenkham,
Mahidol University, Thailand

Reviewed by:

Chun-Ren Wang,
Heilongjiang Bayi Agricultural
University, China
Dandan Hu,
Guangxi University, China

*Correspondence:

Yi-Tian Fu
fyt1997118@163.com
Guo-Hua Liu
liuguohua5202008@163.com

Specialty section:

This article was submitted to
Parasitology,
a section of the journal
Frontiers in Veterinary Science

Received: 28 April 2022

Accepted: 23 May 2022

Published: 22 June 2022

Citation:

Wang H-M, Li R, Deng Y-P, Liu G-H
and Fu Y-T (2022) Comparative
Mitochondrial Genomic Analysis
Robustly Supported That Cat
Tapeworm *Hydatigera taeniaeformis*
(Platyhelminthes: Cestoda)
Represents a Species Complex.
Front. Vet. Sci. 9:931137.
doi: 10.3389/fvets.2022.931137

Hydatigera taeniaeformis is one of the most common intestinal tapeworms that has a worldwide distribution. In this study, we sequenced the complete mitochondrial (mt) genome of *H. taeniaeformis* from the leopard cat (designated HTLC) and compared it with those of *H. taeniaeformis* from the cat in China (designated HTCC) and Germany (designated HTCG). The complete mt genome sequence of HTLC is 13,814 bp in size, which is 167 bp longer than that of HTCC and is 74 bp longer than that of HTCG. Across the entire mt genome (except for the two non-coding regions), the sequence difference was 3.3% between HTLC and HTCC, 12.0% between HTLC and HTCG, and 12.1% between HTCC and HTCG. The difference across both nucleotide and amino acid sequences of the 12 protein-coding genes was 4.1 and 2.3% between the HTLC and HTCC, 13.3 and 10.0% between the HTLC and HTCG, and 13.8 and 10.6% between the HTCC and HTCG, respectively. Phylogenetic analysis based on concatenated amino acid sequences of 12 protein-coding genes showed the separation of *H. taeniaeformis* from different hosts and geographical regions into two distinct clades. Our analysis showed that the cat tapeworm *H. taeniaeformis* represents a species complex. The novel mt genomic datasets provide useful markers for further studies of the taxonomy and systematics of cat tapeworm *H. taeniaeformis*.

Keywords: cat tapeworm, *Hydatigera taeniaeformis*, species complex, mitochondrial genome, comparative analysis

HIGHLIGHTS

- The complete mitochondrial (mt) genome of *Hydatigera taeniaeformis* from the leopard cat in China is 13,814 bp in size.
- Phylogenetic analysis showed the separation of *H. taeniaeformis* from different hosts and geographical regions into two distinct clades.
- The molecular evidence presented here supports the hypothesis that *H. taeniaeformis* represents a species complex.

INTRODUCTION

Tapeworms (Platyhelminthes: Cestoda) are one of the major groups of parasitic flatworms that are passively transmitted between hosts and parasitize virtually every vertebrate species (1). The adults of tapeworms live in the digestive tract of all groups of vertebrates, whereas larvae occur in different organs and body cavities of invertebrates and vertebrates (2). Tapeworms cause neglected diseases that can be fatal and are a major impediment to socioeconomic development (3, 4). Cestoda is a large animal class that is currently classified into 19 orders with ~5,000 valid species (5).

The family Taeniidae is a medically important group of tapeworms that generally consists of four valid genera, *Taenia*, *Echinococcus*, *Hydatigera*, and *Versteria*. Based on morphological features, some authors have recognized that the genus *Hydatigera* is valid, whereas others have treated this genus as a junior synonym of the genus *Taenia*. In a recent revision, the resurrection of the genus *Hydatigera* was proposed based on molecular evidence (6). The cat tapeworm *H. taeniaeformis* (formerly referred to as *T. taeniaeformis*), the type species of the genus *Hydatigera*, is the most common tapeworm of cats and rats (7). Although the zoonotic significance of this tapeworm remains unclear, it has also been recorded in humans in unusual circumstances (8). The correct identification and differentiation of cat tapeworm *H. taeniaeformis* have important implications for studying its epidemiology, systematics, and population genetics. Many studies have shown that several criteria (e.g., infectivity, development, morphology, and biochemistry) differ between *H. taeniaeformis* from different geographical origins (9–12). Based on morphological analysis, *H. taeniaeformis* was considered a complex of three cryptic species (13). Phylogenetic analyses of the *H. taeniaeformis* based on sequences of nuclear 18S rRNA and mitochondrial (mt) *cox1* DNA also strongly support the presence of three distinct clades (13).

The mitochondrial genome (containing 36–37 genes) has been considered a useful marker for the species identification and differentiation of many tapeworms (14–16). The complete mt genomes of *H. taeniaeformis* from cats in China and Germany have been sequenced and supported that they represent different species (17), but it is yet unknown whether *H. taeniaeformis* from leopard cats represents third cryptic species. Therefore, in this study, we (i) sequenced the full mt genome sequence of *H. taeniaeformis* from the leopard cat (HTLC), (ii) compared it with that of *H. taeniaeformis* from the cat in China (HTCC) and Germany (HTCG), and (iii) tested the hypothesis that three isolates of *H. taeniaeformis* are genetically distinct *via* phylogenetic analyses of amino acid sequence datasets.

MATERIALS AND METHODS

Parasites and DNA Isolation

An adult tapeworm representing *H. taeniaeformis* (designated HTLC) was collected from the small intestine of a leopard cat (*Prionailurus bengalensis*) in Beijing Zoo, China. The tapeworm was washed in physiological saline, identified primarily based on morphological characters using existing keys to species (18),

fixed in 70% (v/v) ethanol, and stored at -20°C until use. Total genomic DNA was isolated from this tapeworm using sodium dodecyl sulfate/proteinase K treatment, followed by spin-column purification (Wizard® SV Genomic DNA Purification System, Promega). Furthermore, the specimen was also identified as *H. taeniaeformis* by PCR-based sequencing of 28S rDNA gene and mt *cox1* gene and had 99.2 and 100% nucleotide homology with those of *H. taeniaeformis* from Norway rat in India and brown rat in China deposited in GenBank (GenBank accession nos. JN020350 and MF380378), respectively.

Sequencing, Assembling, and Verification

A genomic DNA library (350 bp inserts) was prepared and sequenced by Novogene Bioinformatics Technology Co. Ltd. (Tianjin, China) using the Illumina HiSeq 2500 (250 bp paired-end reads). The clean reads were obtained from raw reads by removing adaptor sequences, highly redundant sequences, and “N”-rich reads, then were assembled into contigs with Geneious v 11.1.5 (19). Preliminary *cox1* sequences of *H. taeniaeformis* were used as initial references for assembly with the suitable parameters (minimum overlap identity = 99%, minimum overlap = 150 bp, and maximal gap size = 5 bp) (19). The assembly generated a large contig ending with overlapping fragments. As this structure allowed a single circular organization of the mt genome, we assumed that the complete mt genome had been assembled completely. The completeness of the mt genome assembly was further verified by a long PCR experiment using four pairs of primers (Supplementary Figure S1 and Supplementary Table S1), which were designed in the conserved regions. When the additive sizes of amplicons are identical to the length of the assembled contig, it proves that the assembly is correct. The long PCR reaction system consisted of 25 μl : 10.5 μl ddH₂O, 0.5 μl upstream primer, 0.5 μl downstream primer, 12.5 μl Premix Taq™ (Takara, LA Taq™ V 2.0 plus dye), and 1 μl genomic DNA. Samples were tested in the C1000 Touch™ Thermal Cycler (BioRad, USA) under the following conditions: 94°C for 1 min (initial denaturation), then 98°C for 10 s (denaturation), 54–59°C for 40 s (annealing), and 68°C for 4 min for 35 cycles, with a final extension at 72°C for 8 min.

Annotation, Visualization, and Sequence Analysis

Protein-coding and rRNA genes were identified by alignment to homologous genes of the previously sequenced mt genome of *H. taeniaeformis* (17) using the MAFFT v7.122 software (20). tRNA genes were identified using ARWEN (21) and tRNAscan-SE (22). MEGA v6.0 (23) was utilized to analyze the amino acid sequences of each protein-coding gene. The map of the mt genome of the *H. taeniaeformis* was figured on the proksee online server (<https://proksee.ca/>). Pairwise comparisons of the complete mt genomes were made among China isolates (designated HTCC, GenBank accession number: FJ597547), Germany isolates (designated HTCG, GenBank accession number: JQ663994), and China isolates in the present study, such as lengths, identities, A + T content, and codons. The A + T content was computed using DNASTar (v. 5.0). Sequence identities were calculated by

ClustalW in Geneious v 11.1.5 (19): cost matrix IUB, gap open cost of 15, and gap extend cost of 6.66.

Phylogenetic Analyses

A total of 33 mt genome sequences of tapeworms within the family Taeniidae were used for phylogenetic analysis (Table S2), using *Paruterina candelabraria* (GenBank accession number NC039533) as an outgroup (24). Amino acid sequences were single aligned using MAFFT 7.122. The aligned sequences were then concatenated to form a single contig. The poor blocks were excluded from the contig using Gblocks 0.91b (http://phylogeny.lirmm.fr/phylo.cgi/one_task.cgi?task_type=gblocks) using default parameters (25). Phylogenetic analyses were conducted using two methods: Bayesian inference (BI) and maximum likelihood (ML). BI analysis was performed in MrBayes 3.1.1 as described previously (26, 27). ML analysis was performed with PhyML 3.1 as described previously (27, 28). The JTT + I + G + F [lnL = -57,363.05, gamma distribution parameter (G) = 0.857, proportion of invariable sites (I) = 0.304] model selected by ProtTest 3.4.2 (29) was used based on the Akaike information criterion (AIC). Phylograms were drawn using the program FigTree v.1.4.

Analysis of Cox1 Gene Sequence

The newly generated sequences in the present study and previously published *cox1* sequences of *H. taeniaeformis* ($n = 84$) from different hosts and regions (13), were aligned using the software MAFFT 7.122. The nucleotide sequence of these sequences was included in the present analysis, using *H. krepkogorski* (GenBank accession number NC021142) as an outgroup (6). The phylogenetic tree was constructed with Mega v 6.0 (23) using the Neighbor-Joining (NJ) method. The Kimura 2-parameter (K2P) model was the most suitable model, with 1,000 bootstrap replicates.

RESULTS

Original data of this study including nucleotide sequence alignments (*cox1* complete haplotypes dataset) are available at Mendeley Data (<http://dx.doi.org/10.17632/b3knx84j6n.1>).

Genome Content and Organization

We sequenced the *H. taeniaeformis* genome and produced over 3 Gb of Illumina short-read sequence datasets. A total of 20,634,954 × 2 raw reads were generated and 19,826,948 × 2 clean reads were obtained for assembly of the mt genome. The complete mt genome sequence of HTLC (GenBank accession no. ON055368) was 13,814 bp in size (Figure 1) and have a high A + T content (73.2%). It contains 12 protein-coding genes (*cox1–3*, *nad1–6*, *nad4L*, *atp6*, and *cytb*), 22 tRNA genes, two rRNA genes (*rrnS* and *rrnL*), and two non-coding (control or AT-rich) regions (NCR), but lacks an *atp8* gene (Figure 1). The mt genome arrangement of this tapeworm is the same as that of *Spirometra erinaceieuropaei* and *Raillietina tetragona* (27, 30).

Annotation

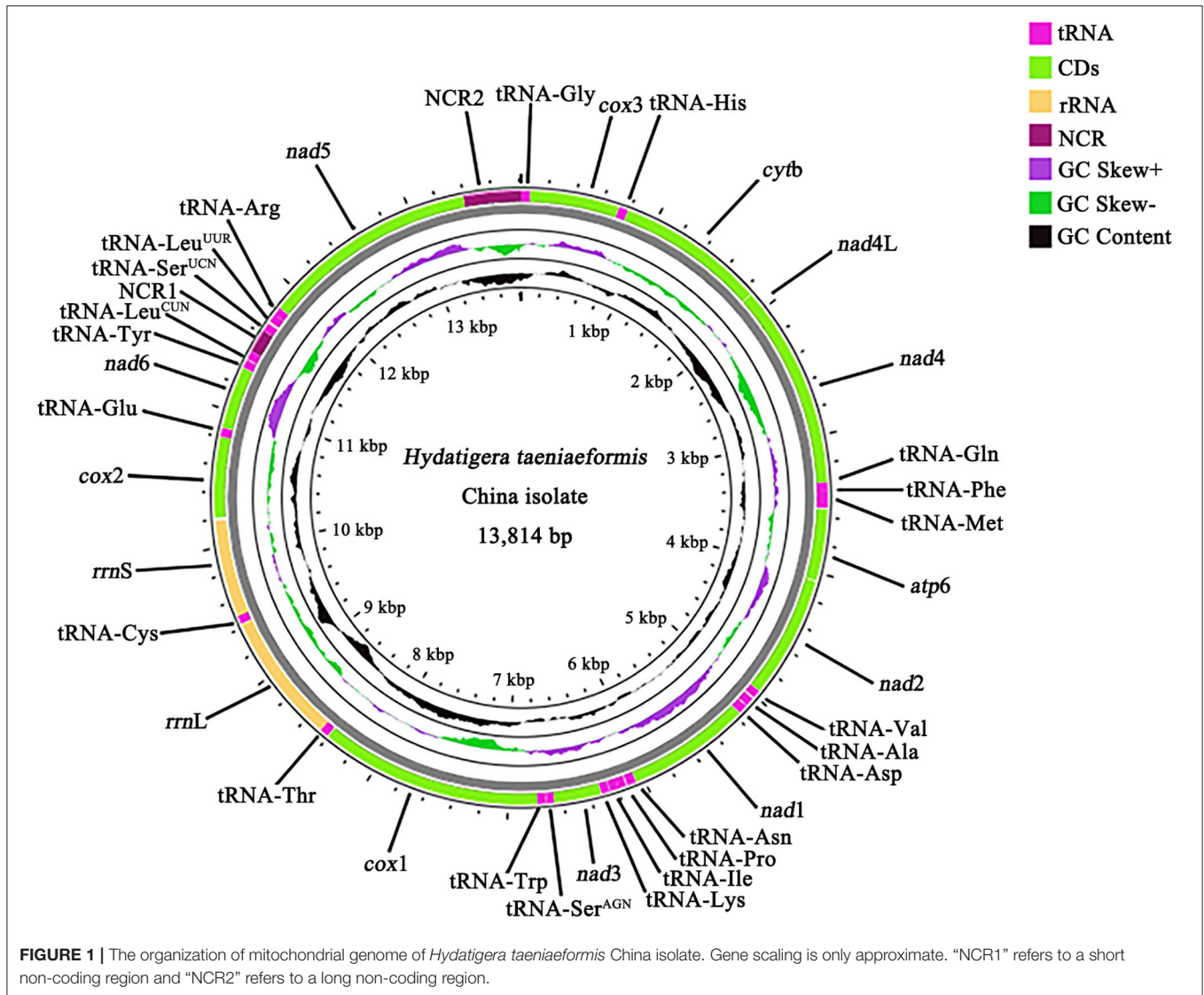
All of the 12 protein-coding genes begin with ATG (*cytb*, *atp6*, *nad4L*, *nad1–3*, *cox1*, *cox2*, and *nad5*), ATT (*cox3*), or GTG (*nad4* and *nad6*) as their start codons. Six of the 12 genes appear to use TAA (*nad1*, *nad2*, *nad5*, *nad6*, *cox3*, and *cytb*) or TAG (*atp6*, *nad4*, and *nad4L*) as the stop codons. Three genes (*cox1*, *cox2*, and *nad3*) end with incomplete stop codon T (Table 1). The 22 tRNA genes varied from 59 to 68 bp in size (Table 1). Secondary structures predicted for the 22 tRNA genes of HTLC (not shown) are similar to those of other tapeworms *T. solium* and *T. asiatica* (31). The *rrnL* gene of this mt genome is located between tRNA-Thr and tRNA-Cys, and *rrnS* gene is located between tRNA-Cys and *cox2*. The length of *rrnL* and *rrnS* genes is 953 and 724 bp, respectively. The A + T content of *rrnL* and *rrnS* is 72.8 and 71.9%, respectively. Two NCRs in the mt genome were inferred. The NCR1 (172 bp) is located between the tRNA-Leu^{CUN} (L1) and tRNA-Ser^{UCN} (S2), and has an A+T content of 79.7%. The NCR2 (418 bp) is located between the *nad5* and tRNA-Gly and has an A+T content of 79.5%.

Comparative Mt Genomic Analyses of HTLC With HTCC and HTCG

The mt genome sequence of HTLC was 13,814 bp in length, 167 bp longer than that of HTCC, and 74 bp longer than that of HTCG. The arrangement of the mt genes (i.e., 12 protein-coding genes, 22 tRNA genes, and two rRNA genes) and two NCRs was the same. Across the entire mt genome (except for the NCR), the sequence difference was 3.3% between HTLC and HTCC, 12.0% between HTLC and HTCG, and 12.1% between HTCC and HTCG. The greatest nucleotide variation among the three isolates of *H. taeniaeformis* was in the *nad3* gene, whereas the least difference was detected in the *cox3* (Table 2). Amino acid sequences inferred from individual mt protein genes of HTLC were compared with those of HTCC and HTCG. The difference across amino acid sequences of the 12 protein genes between the HTLC and HTCC was 2.3%, 10.0% between HTLC and HTCG, and 10.6% between HTCC and HTCG. The amino acid sequence differences among three isolates of *H. taeniaeformis* ranged from 0.4 to 18.3%, with COX2 being the most conserved and NAD3 the least conserved protein. For the 22 tRNAs, the sequence difference was 3.6% between HTLC and HTCC, was 5.7% between HTLC and HTCG, and was 7.5% between HTCC and HTCG. For the *rrnL* and *rrnS* genes, the sequence difference was 3.1 and 6.1% between HTLC and HTCC, 11.5 and 13.9% between HTLC and HTCG, and 10.0 and 9.4% between HTCC and HTCG, respectively (Table 2).

Phylogenetic Analyses

The topologies of the trees using BI and ML were identical (Figure 2). In the tree, four major clades were recovered within the family Taeniidae: [(*Hydatigera* + *Taenia*) + *Echinococcus*] and *Versteria* from the monophyletic groups. Within the *Hydatigera*, all *Hydatigera* isolates clustered together with high statistical support (Bpp = 1; Bf = 100), supporting that the genus *Hydatigera* is valid (32). In the present study, the separation of *H. taeniaeformis* isolates from different hosts and countries was shown to be into two distinct clades with strong support



(Bpp = 1; Bf = 100; **Figure 2**). The *H. taeniaeformis* isolates from cats in Japan, and China and leopard cat in China were clustered in the same clade, while *H. taeniaeformis* isolates from cats in Finland and Germany were clustered in the other clade, which was sister taxa with the former (**Figure 2**). The differences between the two distinct clades are slightly longer (looking at branch lengths) than between *T. asiatica* and *T. saginata* (6) and similar to *Echinococcus* spp. (33–35) (**Figure 2**). In addition, phylogenetic analyses based on the mt *cox1* sequences among *H. taeniaeformis* ($n = 84$) from different hosts and geographical regions revealed three distinct clades (**Supplementary Figure S2**), which are consistent with a previous study (13).

DISCUSSION

The broad consensus on the presence of cryptic species within cat tapeworm *H. taeniaeformis* (6, 13, 17, 36), but their

taxonomic status remains unknown. As described previously, in the morphological features, the variations were embodied in the number of proglottids, the mean number of rostellar hooks, the small hooks of adult worms and metacestodes, the direction of the small hooks, the length of the cirrus sac, and the mean number and location of testes (10). Moreover, in terms of molecular evidence, a previous study reported that the level of nucleotide variation in the *cox1* gene within the genus *Hydatigera* (formerly referred to as *Taenia*) is about 6.3–15.6% (37). Another study showed the *cox1* gene sequence of a new cryptic species *H. taeniaeformis* was distinct from those of other isolates (9.0–9.5%) (12). In the present study, the sequence variations (12.0 and 12.1%) detected in the whole mt genome were consistent with previous results mentioned above in part of the mt *cox1* gene among *H. taeniaeformis* from different regions and hosts (12, 13). These findings indicate that in terms of classification, our results support our proposal that the *H. taeniaeformis* species complex consists of at least two cryptic species. In addition, prior studies

TABLE 1 | The organization of the mitochondrial (mt) genome of *Hydatigera taeniaeformis*.

Gene/region	Positions	Size (bp)	Number of aa ^a	Ini/Ter codons ^b	Anticodons	In ^c
tRNA-Gly (G)	1–65	65			TCC	0
cox3	66–713	648	216	ATT/TAA		+3
tRNA-His (H)	717–784	68			GTG	+2
cytb	787–1,860	1,074	358	ATG/TAA		+4
nad4L	1,865–2,125	261	87	ATG/TAG		–34
nad4	2,092–3,342	1,251	417	GTG/TAG		+0
tRNA-Gln (Q)	3,343–3,406	64			TTG	–3
tRNA-Phe (F)	3,404–3,467	64			GAA	–5
tRNA-Met (M)	3,463–3,528	66			CAT	+9
atp6	3,538–4,056	519	173	ATG/TAG		+6
nad2	4,063–4,950	888	296	ATG/TAA		+1
tRNA-Val (V)	4,952–5,012	61			TAC	+11
tRNA-Ala (A)	5,024–5,089	66			TGC	+5
tRNA-Asp (D)	5,095–5,160	66			GTC	+6
nad1	5,167–6,060	894	298	ATG/TAA		0
tRNA-Asn (N)	6,061–6,128	68			GTT	+7
tRNA-Pro (P)	6,136–6,196	61			TGG	–1
tRNA-Ile (I)	6,196–6,259	64			GAT	+2
tRNA-Lys (K)	6,262–6,322	61			CTT	+4
nad3	6,327–6,666	340	113	ATG/T		0
tRNA-Ser ^{AGN} (S1)	6,667–6,725	59			GCT	+1
tRNA-Trp (W)	6,727–6,788	62			TCA	0
cox1	6,789–8,400	1,612	537	ATG/T		+8
tRNA-Thr (T)	8,409–8,470	62			TGT	0
rrnL	8,471–9,423	953				+8
tRNA-Cys (C)	9,432–9,491	60			GCA	–21
rrnS	9,471–10,194	724				+23
cox2	10,218–10,800	583	194	ATG/T		+0
tRNA-Glu (E)	10,801–10,867	67			TTC	+3
nad6	10,871–11,320	450	150	GTG/TAA		+3
tRNA-Tyr (Y)	11,324–11,385	62			GTA	+6
tRNA-Leu ^{CUN} (L1)	11,392–11,454	63			TAG	0
Non-coding region (NCR1)	11,455–11,626	172				0
tRNA-Ser ^{UCN} (S2)	11,627–11,688	62			TCA	+17
tRNA-Leu ^{UUR} (L2)	11,706–11,767	62			TAA	–1
tRNA-Arg (R)	11,767–11,825	59			ACG	+2
nad5	11,828–13,396	1,569	523	ATG/TAA		0
Non-coding region (NCR2)	13,397–13,814	418				0

^aThe inferred length of amino acid (aa) sequence of 12 protein-coding genes.

^bIni/Ter codons, initiation and termination codons; ^cIn, intergenic nucleotides.

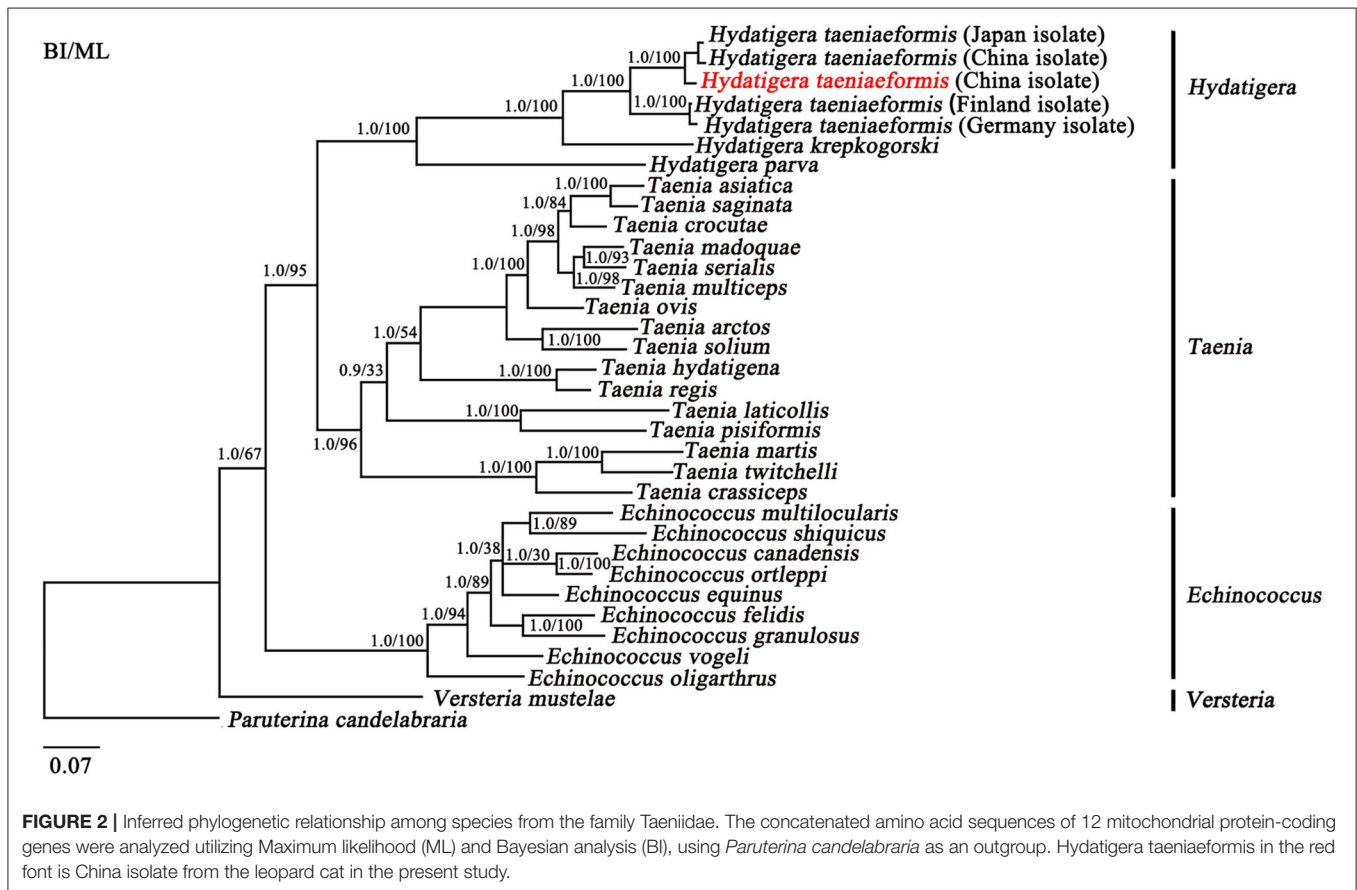
have indicated that mtDNA sequence variation between closely-related nematode species was typically in the range of 10–20% (12, 38).

Phylogenetic analyses of *H. taeniaeformis* provided additional evidence that *H. taeniaeformis* represents closely but distinct taxa. Taken together, the molecular evidence presented here supports the hypothesis that *H. taeniaeformis* represents a species complex (13, 17). The taxonomy of *H. taeniaeformis* is often insufficient based on morphological features and their hosts or geographical origins; therefore, accurate identification and

differentiation of *H. taeniaeformis* can sometimes be challenging (7, 13, 36). In the present study, the characterization of the mt genome of HTLC in China also stimulates a reassessment of the phylogenetic relationships of *H. taeniaeformis* using mt genomic/proteomic datasets. Mt genomic sequences are useful molecular markers to study the species identification, genetic structure, and phylogenetic analyses of many worms, especially for searching potential cryptic species (17, 38). Therefore, in the present study, the analyses of mt genomic sequences provided insight into the phylogenetic relationships of *H. taeniaeformis*;

TABLE 2 | Nucleotide (nt) and/or predicted amino acid (aa) sequence differences in mt genes *H. taeniaeformis* from the leopard cat (HTLC) (A) and *H. taeniaeformis* from the cat in China (HTCC) (B); HTLC (A) and *H. taeniaeformis* from the cat in Germany (HTCG) (C); HTCC (B) and HTCG (C) upon pairwise comparison.

Gene/region	Nt sequence length			Nt difference (%)			Number of aa			aa difference (%)		
	A	B	C	A/B	A/C	B/C	A	B	C	A/B	A/C	B/C
<i>atp6</i>	519	519	519	3.9	12.7	13.1	172	172	172	2.3	11.1	11.1
<i>cox1</i>	1,612	1,612	1,635	3.1	10.7	10.8	537	537	544	0.4	2.2	2.8
<i>cox2</i>	583	582	583	3.8	6.4	7.0	194	193	194	1.0	1.6	1.56
<i>cox3</i>	648	645	645	2.9	11.0	11.2	216	214	214	1.4	6.9	6.5
<i>cytb</i>	1,074	1,047	1,068	5.9	14.1	16.7	358	348	355	5.0	10.9	13.5
<i>nad1</i>	894	894	894	3.5	12.5	13.3	298	297	297	2.7	12.1	12.5
<i>nad2</i>	888	891	888	4.6	14.5	14.3	296	296	295	3.0	12.1	10.8
<i>nad3</i>	340	340	348	5.6	17.8	17.8	113	113	115	7.1	17.4	18.3
<i>nad4</i>	1,251	1,251	1,251	4.2	15.0	15.4	416	416	416	2.6	14.7	14.2
<i>nad4L</i>	261	261	261	3.8	14.6	14.9	86	86	86	2.3	11.6	11.6
<i>nad5</i>	1,569	1,569	1,569	4.3	15.9	15.6	523	522	522	3.8	14.5	14.2
<i>nad6</i>	450	448	453	4.2	15.5	16.3	150	149	150	4.7	12.6	16.0
<i>rrnS</i>	724	717	725	6.1	13.9	9.4	–	–	–	–	–	–
<i>rrnL</i>	953	959	959	3.1	11.5	10.0	–	–	–	–	–	–
All 22tRNA	1,392	1,410	1,399	3.6	5.7	7.5	–	–	–	–	–	–



however, some *H. taeniaeformis* isolates from cats in Italy (clade C) (36, 39) are either not represented or underrepresented (13).

Therefore, further studies are required to resolve the taxonomic classification of *H. taeniaeformis* species complex by sequencing

the mt genomic sequences of additional *H. taeniaeformis* from different countries and hosts.

CONCLUSIONS

The present study sequenced the complete mt genome sequence of the cat tapeworm HTLC in China. The comparative mt genomic analysis provided robust genetic evidence that *H. taeniaeformis* represents a species complex. The novel mt genomic datasets provide useful markers for further studies of the taxonomy and systematics of cat tapeworm *H. taeniaeformis*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/genbank/>, ON055368.

ETHICS STATEMENT

This study was approved by the Animal Ethics Committee of Hunan Agricultural University (No. 43321503).

AUTHOR CONTRIBUTIONS

Y-TF, G-HL, and H-MW: conceptualization, design, revision critically, and analysis. H-MW:

experiment and drafted the manuscript. RL and Y-PD: study implementation, and manuscript preparation. All authors read and approved the final manuscript.

FUNDING

This study was supported in part, by the Training Program for Excellent Young Innovators of Changsha (Grant No. KQ2106044).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fvets.2022.931137/full#supplementary-material>

Supplementary Figure S1 | PCR verification of the mt genome of *Hydatigera taeniaeformis* China isolate. Lane M: DL5000 DNA marker. Lane 1–4: *cox3-nad1*, *nad1-cox1*, *nad6-cox3*, and *nad6-cox3*.

Supplementary Figure S2 | Inferred phylogenetic relationship based on *cox1* sequences of *H. taeniaeformis* ($n = 84$) from different hosts and regions. Clade A consisted of 39 haplotypes (designated as A1–A39, A37 in the red font is the *cox1* fragment of *H. taeniaeformis* in this study), clade B included 23 haplotypes (B1–B23), and clade C contained 22 haplotypes (C1–C22), using Neighbor-Joining (NJ) method, with an outgroup *H. krepkogorski*.

Supplementary Dataset | Mitochondrial DNA data. Nucleotide sequence alignments (*cox1* complete haplotypes dataset). (<http://dx.doi.org/10.17632/b3kx84j6n.1>).

REFERENCES

- Kamenetzky L, Maldonado LL, Cucher MA. Cestodes in the genomic era. *Parasitol Res.* (2022) 121:1077–89. doi: 10.1007/s00436-021-07346-x
- Webb C, Cabada MM. Intestinal cestodes. *Curr Opin Infect.* (2017) 30:504–10. doi: 10.1097/QCO.0000000000000400
- Hotez PJ, Woc-Colburn L, Bottazzi ME. Neglected tropical diseases in Central America and Panama: review of their prevalence, populations at risk and impact on regional development. *Int J Parasitol.* (2014) 44:597–603. doi: 10.1016/j.ijpara.2014.04.001
- Molyneux DH, Savioli L, Engels D. Neglected tropical diseases: progress towards addressing the chronic pandemic. *Lancet.* (2017) 389:312–25. doi: 10.1016/S0140-6736(16)30171-4
- Caira JN, Jensen K. *Planetary Biodiversity Inventory (2008–2017): Tapeworms from Vertebrate Bowels of the Earth*. University of Kansas, Natural History Museum (2017).
- Nakao M, Lavikainen A, Iwaki T, Haukisalmi V, Konyaev S, Oku Y, et al. Molecular phylogeny of the genus *Taenia* (Cestoda: Taeniidae): proposals for the resurrection of *Hydatigera* Lamarck, 1816 and the creation of a new genus *Versteria*. *Int J Parasitol.* (2013) 43:427–37. doi: 10.1016/j.ijpara.2012.11.014
- Alvi MA Li L, Ohiolei JA, Qamar W, Saqib M, Tayyab MH, et al. *Hydatigera taeniaeformis* in urban rats (*Rattus rattus*) in Faisalabad, Pakistan. *Infect Genet Evol.* (2021) 92:104873. doi: 10.1016/j.meegid.2021.104873
- Guo X. Proteomics analysis of *Hydatigera taeniaeformis* metacestode stage. *Front Vet Sci.* (2020) 7:474. doi: 10.3389/fvets.2020.00474
- Nonaka N, Iwaki T, Okamoto M, Ooi HK, Oku Y, Ohbayashi M, et al. Infectivities of four isolates of *Taenia taeniaeformis* to various rodents. *J Vet Sci.* (1994) 56:565–7. doi: 10.1292/jvms.56.565
- Iwaki T, Nonaka N, Okamoto M, Oku Y, Kamiya M. Developmental and morphological characteristics of *Taenia taeniaeformis* (Batsch, 1786) in *Clethrionomys rufocanus bedfordiae* and *Rattus norvegicus* from different geographical locations. *J Parasitol.* (1994) 80:461–7. doi: 10.2307/3283418
- Azuma H, Okamoto M, Oku Y, Kamiya M. Intraspecific variation of *Taenia taeniaeformis* as determined by various criteria. *Parasitol Res.* (1995) 81:103–8. doi: 10.1007/BF00931613
- Okamoto M, Bessho Y, Kamiya M, Kurosawa T, Horii T. Phylogenetic relationships within *Taenia taeniaeformis* variants and other taeniid cestodes inferred from the nucleotide sequence of the cytochrome *c* oxidase subunit I gene. *Parasitol Res.* (1995) 81:451–8. doi: 10.1007/BF00931785
- Lavikainen A, Iwaki T, Haukisalmi V, Konyaev SV, Casiraghi M, Dokuchaev NE, et al. Reappraisal of *Hydatigera taeniaeformis* (Batsch, 1786) (Cestoda: Taeniidae) sensu lato with description of *Hydatigera kamiyai* n. sp. *Int J Parasitol.* (2016) 46:361–74. doi: 10.1016/j.ijpara.2016.01.009
- Liu GH, Lin RQ, Li MW, Liu W, Liu Y, Yuan ZG, et al. The complete mitochondrial genomes of three cestode species of *Taenia* infecting animals and humans. *Mol Biol Rep.* (2011) 38:2249–56. doi: 10.1007/s11033-010-0355-0
- Zhao GH, Wang HB, Jia YQ, Zhao W, Hu XF, Yu SK, et al. The complete mitochondrial genome of *Pseudanoplocephala crawfordi* and a comparison with closely related cestode species. *J Helminthol.* (2016) 9:588–95. doi: 10.1017/S0022149X15000802
- Li WX, Fu PP, Zhang D, Boyce K, Xi BW, Zou H, et al. Comparative mitogenomics supports synonymy of the genera *Ligula* and *Digramma* (Cestoda: Diphyllbothriidae). *Parasit Vectors.* (2018) 11:324. doi: 10.1186/s13071-018-2910-9
- Jia W, Yan H, Lou Z, Ni X, Dyachenko V, Li H. Mitochondrial genes and genomes support a cryptic species of tapeworm within *Taenia taeniaeformis*. *Acta Trop.* (2012) 123:154–63. doi: 10.1016/j.actatropica.2012.04.006
- Schmidt GD. *Handbook of Tapeworm Identification*. Boca Raton, FL: CRC Press (1986).

19. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. (2012) 28:1647–9. doi: 10.1093/bioinformatics/bts199
20. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. (2013) 30:772–80. doi: 10.1093/molbev/mst010
21. Laslett D, Canbäck B. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*. (2008) 24:172–5. doi: 10.1093/bioinformatics/btm573
22. Chan PB, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol*. (2019) 1962:1–14. doi: 10.1007/978-1-4939-9173-0_1
23. Kumar S, Stecher G, Li M, Knyaz C, Tamura K, MEGA X. molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. (2018) 35:1547–9. doi: 10.1093/molbev/msy096
24. Guo A, Wang L, Zhang S, Zheng Y, Georgiev BB, Luo X, et al. Mitochondrial genome of *Paruterina candelabraria* (Cestoda: Paruterinidae), with implications for the relationships between the genera *Cladotaenia* and *Paruterina*. *Acta Trop*. (2019) 189:1–5. doi: 10.1016/j.actatropica.2018.09.016
25. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. (2007) 56:564–77. doi: 10.1080/10635150701472164
26. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. (2003) 19:1572–4. doi: 10.1093/bioinformatics/btg180
27. Liu GH Li C, Li JY, Zhou DH, Xiong RC, Lin RQ, et al. Characterization of the complete mitochondrial genome sequence of *Spirometra erinaceieuropaei* (Cestoda: Diphylobothriidae) from China. *Int J Biol Sci*. (2012) 8:640–9. doi: 10.7150/ijbs.4096
28. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. (2010) 59:307–21. doi: 10.1093/sysbio/syq010
29. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. (2011) 27:1164–5. doi: 10.1093/bioinformatics/btr088
30. Liang JY, Lin RQ. The full mitochondrial genome sequence of *Raillietina tetragona* from chicken (Cestoda: Davaineidae). *Mitochondrial DNA A DNA Mapp Seq Anal*. (2016) 27:4160–1. doi: 10.3109/19401736.2014.1003892
31. Jeon HK, Kim KH, Eom KS. Complete sequence of the mitochondrial genome of *Taenia saginata*: comparison with *T. solium* and *T. asiatica*. *Parasitol Int*. (2007) 56:243–6. doi: 10.1016/j.parint.2007.04.001
32. Gao Y, Yan L, Qiu D, Huang Z, Hu D, Zhang D. First mitogenome of *Moniezia sichuanensis* from forest musk deer with comparative analyses within cyclophyllidea. *Vet Parasitol*. (2021) 299:109575. doi: 10.1016/j.vetpar.2021.109575
33. Nakao M, Lavikainen A, Yanagida T, Ito A. Phylogenetic systematics of the genus *Echinococcus* (Cestoda: Taeniidae). *Int J Parasitol*. (2013) 43:1017–29. doi: 10.1016/j.ijpara.2013.06.002
34. Lymbery AJ. Phylogenetic pattern, evolutionary processes and species delimitation in the genus *Echinococcus*. *Adv Parasitol*. (2017) 95:111–45. doi: 10.1016/bs.apar.2016.07.002
35. Maldonado LL, Arrabal JP, Rosenzvit MC, Oliveira GC, Kamenetzky L. Revisiting the phylogenetic history of helminths through genomics, the case of the new *Echinococcus oligarthrus* genome. *Front Genet*. (2019) 10:708. doi: 10.3389/fgene.2019.00708
36. Galimberti A, Romano DF, Genchi M, Paoloni D, Vercillo F, Bizarri L, et al. Integrative taxonomy at work: DNA barcoding of taeniids harboured by wild and domestic cats. *Mol Ecol Resour*. (2012) 12:403–13. doi: 10.1111/j.1755-0998.2011.03110.x
37. Blouin MS. Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. *Int J Parasitol*. (2002) 32:527–31. doi: 10.1016/S0020-7519(01)00357-5
38. Jin YC Li XY, Liu JH, Zhu XQ, Liu GH. Comparative analysis of mitochondrial DNA datasets indicates that *Toxascaris leonina* represents a species complex. *Parasit Vectors*. (2019) 12:194. doi: 10.1186/s13071-019-3447-2
39. Galimberti A, Spada M, Russo D, Mucedda M, Agnelli P, Crottini A, et al. Integrated operational taxonomic units (IOTUs) in echolocating bats: a bridge between molecular and traditional taxonomy. *PLoS One*. (2012) 7:e40122. doi: 10.1371/journal.pone.0040122

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wang, Li, Deng, Liu and Fu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.