Comparative mitogenomics supports synonymy of the genera *Ligula* and *Digramma* (Cestoda: Diphyllobothriidae)

Wen X. Li, Pei P. Fu, Dong Zhang, Kellyanne Boyce, Bing W. Xi, Hong Zou, Ming Li, Shan G. Wu and Gui T. Wang

**Abstract**

**Background:** After observing differences in the number of reproductive complexes per proglottid within the genus *Ligula*, the genus *Digramma* was erected. However, the validity of *Digramma* has been previously questioned due to a low variability in the *cox1*, *nad1* and ITS rDNA sequences between the two genera. We undertook a study to greatly increase the amount of sequence data available for resolution of this question by sequencing and characterizing the complete mitogenomes of *Digramma interrupta* and *Ligula intestinalis*.

**Results:** The circular mtDNA molecules of *Digramma interrupta* and *Ligula intestinalis* are 13,685 bp and 13,621 bp in size, respectively, both comprising 12 PCGs, 22 tRNA genes, two rRNA genes, and two mNCRs. Both mitogenomes exhibit the same gene order and share 92.7% nucleotide identity, compared with 85.8–86.5% to the most closely related genus *Dibothriocephalus*. Each gene from *D. interrupta* and *L. intestinalis* is almost of the same size, and the sequence identity ranges from 87.5% (*trnD*) to 100% (*tmh*, *tmQ* and *trnV*). NCR2 sequences of *D. interrupta* and *L. intestinalis* are 249 bp and 183 bp in length, respectively, which contributes to the main difference in length between their complete mitogenomes. A sliding window analysis of the 12 PCGs and two rRNAs indicated nucleotide diversity to be higher in *nad5*, *nad6*, *nad2*, *nad4* and *cox3*, whereas the most conserved genes were *rrnL* and *rrnS*. Lower sequence identity was also found in *nad2*, *nad4*, *nad5*, *nad6* and *cox3* genes between the two diphyllobothriids. Within the Diphyllobothriidae, phylogenetic analysis indicated *Ligula* and *Digramma* to be most closely related to one another, forming a sister group with *Dibothriocephalus*.

**Conclusions:** Owing to higher nucleotide diversity, the genes *nad2*, *nad4*, *nad5*, *nad6* and *cox3* should be considered optimal candidates to use as molecular markers for population genetics and species identification between the two closely related species. The phylogenetic results in combination with the comparative analysis of the two mitogenomes, consistently support the congeneric status of *L. intestinalis* and *D. interrupta*.

**Keywords:** Mitogenome, Eucestoda, Diphyllobothriidea, *Digramma*, *Ligula*

**Background**

Based upon its paraphyly, differences in the position of the genital pore, the presence of an external seminal vesicle and the absence of a uterine sac in the Diphyllobothriidea Kuchta, Scholz, Brabec & Bray, 2008, the order Pseudo-phyllidea van Beneden in Carus, 1863 was suppressed and the Diphyllobothriidea and Bothriocephalidea Kuchta, Scholz, Brabec & Bray, 2008 were proposed [1–3].

The order Diphyllobothriidea includes 70 species considered valid, classified into 18 genera across three families [2, 4]. Adult diphyllobothriideans are found only in tetrapods, never having been recorded in fish [2], and the plerocercoids of groups such as *Spirometra*, *Diphyllobothrium* Cobbold, 1858 (syn. Diplogonoporus Lönnberg, 1892) and *Dibothriocephalus* Lühe, 1899 (a recently resurrected genus including some species from *Diphyllobothrium*, i.e. *Dib. dendriticus* (Nitzsch, 1824), *Dib. nihonkaiensis* (Yamane, Kamo, Bylund & Wikgren, 1992)) have been recorded in fish [2].
1986) and *Dib. latus* (Linnaeus, 1758) belonging to the family Diphyllobothriidae are often the principal agents of food-borne cestodosis [5]. *Ligula* spp. also belong to the Diphyllobothriidae; these species use copepods as first intermediate hosts, freshwater fish as second intermediate hosts and birds as definitive hosts [6]. *Ligula intestinalis* (Linnaeus, 1758) is a tapeworm of veterinary importance worldwide that reduces the fecundity of the cyprinid fishes by parasitic castration [6, 7] and induces mass mortalities of the carp *Chanodichthys erythropterus* [8].

Cholodkovsky [9] erected the genus *Digramma* Cholodkovsky, 1915 after observing differences in the number of reproductive complexes contained within each proglottid when studying the genus *Ligula* Linnaeus, 1758. In China, *Ligula* spp. are distributed in the Qinghai-Tibet Plateau with Schizothoracinae fishes serving as the primary second intermediate host. *Digramma* spp. are found across the rest of China where the goldfish *Carassius auratus* acts as their common second intermediate host [10]. However, the validity of *Digramma* has been questioned due to a low level of difference between the species of two genera in the *cox1*, *nad1* and ITS rDNA sequences [11–13]. Thus *Digramma* is considered to be synonymous with *Ligula* [2]. Owing to the fact that only one gene and a limited number of isolates were included in that study, more sequence data and a greater range of taxa from different genera are required for a more comprehensive phylogenetic analysis of the Diphyllobothriidea [13].

This study, therefore, aimed to sequence and characterize the complete mitogenomes of *Digramma interrupta* and *Ligula intestinalis* and to perform phylogenetic analysis to investigate whether or not these two diphyllobothriids are congeneric using mitogenomic data. Differences within the mitochondrial genes were also compared to determine which genes would be suitable for the design of molecular markers as a means to differentiate *D. interrupta* from *L. intestinalis*.

**Methods**

**Specimen collection and DNA extraction**
Plerocercoids of *D. interrupta* and *L. intestinalis* were isolated from the body cavity of *Carassius auratus* collected from Liangzi Lake in Hubei Province and *Gymnocypris selincuoensis* from Siling Lake, Tibet, China, respectively. The tapeworms were preserved in 80% ethanol and stored at 4 °C. Total genomic DNA was extracted from the posterior region of a single tapeworm using a TIANamp Micro DNA Kit (Tiangen Biotech, Beijing, China), according to the manufacturer's instructions. DNA was stored at -20 °C for subsequent molecular analysis. The morphological identification of specimens was confirmed using sequence data from the ITS2 rDNA region [13, 14] and the *cox1* gene [11].

**Amplification and DNA sequencing**

PCR was carried out as described previously [15, 16], with minor modifications. Five degenerate primer sets (Additional file 1: Table S1) were designed to primarily amplify partial sequences of the *nad5*, *cytb*, *nad2*, *cox1* and *rrnS* genes. The sequenced fragments were subsequently used to design primers specific for the amplification and sequencing of the whole mitogenome. PCR reactions were performed in a 20 μl reaction mixture, containing 7.4 μl dd PCR grade H₂O, 10 μl 2× PCR buffer (2 mM Mg²⁺, 8 μl dNTP plus, Takara, Dalian, China), 0.6 μl of each primer (12.5 μM), 0.4 μl rTaq polymerase (250 U, Takara, Dalian, China), and 1 μl genomic DNA template. Amplification was conducted under the following conditions: initial denaturation at 98

![Fig. 1 Circular map of the mitochondrial genome of *Digramma interrupta* and *Ligula intestinalis*. Protein-coding genes (12) are shown in red, tRNAs (22) in yellow, rRNAs (2) in green and non-coding regions in grey](image_url)
| Gene       | Position | Size (bp) | Intergenic nucleotides | Codon Identity | Sequence |
|------------|----------|-----------|------------------------|----------------|----------|
|            | From     | To        |                        |                |          |
|            |          |           |                        |                |          |
| **Digramma interrupta/Ligula intestinalis** |
| cox3       | 1/1      | 643/643   | 643/643                | GTG/GTG        | T/T      |
|             |          |           |                        |                | 91.29    |
| trnH       | 644/644  | 710/710   | 67/67                  | 100            |
| GAP1       | 711/711  | 713/713   | 3/3                    | 100 gaa/gaa    |
| cytb       | 714/714  | 1820/1820 | 1107/1107              | ATG/ATG TAA/TAA |
| GAP2       | 1821/1821| 1821/1821 | 1/1                    | 100 c/c        |
| nad4L      | 1822/1822| 2082/2082 | 261/261                | ATG/ATG TAA/TAA |
| nad4       | 2043/2043| 3293/3293 | 1251/1251              | -40/-40        |
| trnQ       | 3293/3293| 3357/3357 | 65/65                  | 97.01          |
| tmF        | 3353/3353| 3419/3419 | 67/67                  | 98.51          |
| tmM        | 3416/3416| 3482/3482 | 67/67                  | -4/-4          |
| GAP3       | 3483/3483| 3485/3485 | 3/3                    | 96.67 gtt/att  |
| atp6       | 3486/3486| 3995/3995 | 510/510                | 93.33          |
| GAP4       | 3996/3996| 3997/3997 | 2/2                    | 100 tc/tc      |
| nad2       | 3998/3998| 4876/4876 | 879/879                | 91.24          |
| GAP5       | 4877/4877| 4877/4877 | 1/1                    | 100 t/t        |
| tmV        | 4878/4878| 4941/4941 | 64/64                  | 98.36          |
| GAP6       | 4942/4942| 4949/4949 | 8/8                    | 87.5 gttcttaag/gttttaag |
| tmA        | 4950/4950| 5010/5010 | 61/61                  | 95.38          |
| GAP7       | 5011/5011| 5013/5013 | 3/3                    | 100 tgg/tgg    |
| tmD        | 5014/5014| 5077/5077 | 64/64                  | 87.5           |
| nad1       | 5078/5078| 5968/5968 | 891/891                | 93.04          |
| tmN        | 5968/5968| 6033/6033 | 66/66                  | -1/-1          |
| GAP8       | 6034/6034| 6040/6040 | 7/7                    | 96.97          |
| tmP        | 6041/6041| 6105/6105 | 65/65                  | tatggt/tatggt |
| GAP9       | 6106/6106| 6112/6113 | 7/8                    | 95.38          |
| tm1        | 6113/6114| 6177/6178 | 65/65                  | 98.36          |
| GAP10      | 6178/6179| 6194/6195 | 17/17                  | 98.44          |
| tmK        | 6195/6196| 6258/6259 | 64/64                  | 93.85          |
| GAP11      | 6259/6260| 6261/6260 | 3/1                    | 94.96          |
| nad3       | 6262/6261| 6618/6617 | 357/357                | 94.96          |
| tmS1       | 6608/6607| 6666/6665 | 59/59                  | 96.61          |
| GAP12      | 6667/6666| 6668/6667 | 2/2                    | 96.61          |
| tmW        | 6669/6668| 6731/6730 | 63/63                  | 95.38          |
| GAP13      | 6732/6731| 6739/6738 | 8/8                    | 98.39          |
| cox1       | 6740/6739| 8305/8304 | 1566/1566              | 93.49          |
| tmT        | 8296/8295| 8357/8356 | 62/62                  | 96.38          |
| mlL        | 8358/8357| 9324/9323 | 967/967                | 98.44          |
| tmC        | 9325/9324| 9388/9387 | 64/64                  | 98.39          |
| msS        | 9389/9388| 10130/10130| 742/743               | 95.83          |
| cox2       | 10131/10131| 10700/10700| 570/570               | 94.39          |
| GAP14      | 10701/10701| 10701/10701| 1/1                   | 100 a/a        |
°C for 2 min, followed by 40 cycles at 98 °C for 10 s, 48–60 °C for 15 s, 68 °C for 1 min/kb, and a final extension at 68 °C for 10 min. PCR products were sequenced on an ABI 3730 automatic sequencer using the Sanger method at Sangon Company (Shanghai, China) using the primer walking strategy.

**Sequence annotation and analyses**

The mitogenome was annotated broadly following the procedure described previously [15, 16]. The amplified fragments were initially checked by BLASTN [17], before being assembled manually in a stepwise manner. The annotation was recorded in a Word document with the help of the Geneious program [18], using the mitogenome of *Dibothriocephalus latus* (syn. *Diphyllobothrium latum*) (NC_008945) as the reference sequence. PCGs were found by searching for ORFs (employing genetic code 9, echinoderm mitochondrial; flatworm mitochondrial) and the nucleotide alignments against the selected reference genome in Geneious. *rrnL* and *rrnS* were annotated in a similar way, via comparison with homologs using Geneious. ARWEN [19] and MITOS [20] web servers were employed to identify and fold all tRNAs. Similarly, the NCBI submission file (*.sqn) and tables of statistics for mitogenomes were generated using a home-made GUI-based program, MitoTool [21]. MitoTool was also used to calculate codon usage and relative synonymous codon usage (RSCU) for the 12 PCGs of *D. interrupta* and *L. intestinalis* and the results were sorted and imported into ggplot2 [22] to draw the RSCU figure. A Tandem Repeats Finder [23] was used to predict tandem repeats (TR) in the major non-coding regions (mNCRs), and the secondary structures of NCR1 and TR were folded by Mfold software [24]. Non-synonymous (dN) and synonymous (dS) mutation rates among the 12 PCGs of *D. interrupta* and *L. intestinalis* were computed using KaKs_Calculator [25] utilising a modified Yang-Nielsen algorithm. DnaSP v5 [26] was adopted to conduct sliding window analyses. A sliding window of 500 bp and step size of 25 bp was implemented to estimate nucleotide divergence $\pi$ between the alignments of the mitogenomes of *D. interrupta* and *L. intestinalis*.

**Phylogenetic analyses**

The mitogenomes of 35 cestodes, covering five orders and ten families (Additional file 2: Table S2) were obtained from GenBank and were used, along with the two new mitogenomes generated in this study, to create the phylogenetic reconstruction. Two trematodes, *Dicrocoelium chinensis* (NC_025279) and *Dicrocoelium dendriticum* (NC_025280), were used as outgroups. All 36 genes (12 PCGs, 2 rRNAs and 22 tRNAs) were used for phylogenetic inference and were extracted from GenBank files using MitoTool. PCGs were aligned in batches using MAFFT and integrated into our own in-house GUI-based program, BioSuite [27], adopting codon-alignment mode.

| Table 1 Annotated mitochondrial genomes of *Diamerma interrupta* and *Ligula intestinalis* (Continued) |
|-----------------------------------------------|-------------------|-----------------|-------------------|---------------|
| **Gene** | **Position** | **Size (bp)** | **Intergenic nucleotides** | **Codon Identity (%)** | **Sequence** |
| From | To | Start | Stop |  |
| trnE | 10702/10702 | 10765/10765 | 64/64 | 98.44 | ttagc/ttagc |
| GAP15 | 10766/10766 | 10770/10770 | 5/5 | 100 | ata/ata |
| nad6 | 10771/10771 | 11229/11229 | 459/459 | ATG/ATG | TAG/TAG |
| GAP16 | 11230/11230 | 11232/11232 | 3/3 | 100 | ata/ata |
| tmY | 11233/11233 | 11298/11297 | 66/65 | 96.97 | |
| NCR1 | 11299/11298 | 11521/11521 | 223/224 | 91.07 | |
| tml1 | 11522/11522 | 11588/11588 | 67/67 | 91.04 | |
| GAP17 | 11589/11589 | 11600/11601 | 12/13 | 84.62 | tgcgagggttt/tctgtggggggttt |
| tmS2 | 11601/11602 | 11665/11667 | 65/66 | 96.97 | |
| GAP18 | 11666/11668 | 11676/11678 | 11/11 | 72.73 | tagttaaaga/crgtttaataa |
| tmL2 | 11677/11679 | 11740/11742 | 64/64 | 95.31 | |
| tmR | 11741/11743 | 11795/11797 | 55/55 | 92.73 | |
| GAP19 | 11796/11798 | 11798/11800 | 3/3 | 100 | ttt/ttt |
| nad5 | 11799/11801 | 13367/13369 | 1569/1569 | ATG/ATG | TAA/TAA |
| NCR2 | 13368/13370 | 13616/13552 | 249/183 | 67.86 | |
| tmG | 13617/13553 | 13682/13618 | 66/66 | 96.97 | |
| GAP20 | 13683/13619 | 13685/13621 | 3/3 | 100 | aag/aag |
| genome | 1/1 | 13685/13621 | 13,685/13,621 | 92.72 | |
All RNA genes (rRNA and tRNA) were aligned using a structural alignment algorithm Q-INS-i incorporated into MAFFT-with-extensions software [28]. Gaps and ambiguous sites were deleted using GBlocks [29] integrated by BioSuite with default settings. BioSuite was subsequently used to concatenate the sequences into a single alignment and generate phylip and nexus format files.

GTR+I+G was chosen as the optimal model of nucleotide evolution for all datasets based on the Akaike information criterion by ModelGenerator [30]. Two analytical methods were performed: maximum likelihood (ML) and Bayesian inference (BI). The ML analysis was performed in RAxML GUI [31] using a ML+rapid bootstrap (BP) algorithm with 1000 replicates. BI analysis was performed in MrBayes 3.2.6 [32] with default settings, 6 × 10⁶ metropolis-coupled MCMC generations, and 1000 sample frequency. Stationarity was considered to have been reached when the average standard deviation of split frequencies was below 0.01, ESS (estimated sample size) was above 200, and PSRF (potential scale reduction factor) approached 1. Bayesian posterior probability (BPP) values were calculated in a consensus tree, after discarding the first 25% of samples as burn-in. Finally, the resultant trees were visualised and annotated by iTOL [33] with the help of several dataset files generated by MitoTool.

**Results**

**Genome organization and base composition**

The length of the circular mtDNA molecules of *D. interrupta* (GenBank accession number: MF671697) and *L. intestinalis* (GenBank accession number: MF671696) was 13,685 bp and 13,621 bp, respectively. Both mitogenomes were composed of 12 PCGs, 22 tRNA genes, two rRNA genes and two mNCRs (major non-coding regions), all of which were transcribed from the same strand (Fig. 1). As commonly reported for flatworms [34], the two mitogenomes lacked the *atp8* gene. The gene order of the two mitogenomes was identical, conforming to the synapomorphic gene arrangement of the order Diphyllobothriidea [35]. The A+T content of the mitogenomes of *D. interrupta* and *L. intestinalis* were 67.9% and 67%, respectively, which is in accordance with that of other cestodes (Additional file 2: Table S2).

| Region         | Size (bp) | T(U) | C   | A   | G   | AT (%) | GC (%) | AT skew | GC skew |
|----------------|-----------|------|-----|-----|-----|--------|--------|---------|---------|
| PCGs           | 10,062/10,062 | 45.6/45.1 | 12.3/13.1 | 21.9/21.5 | 20.2/20.3 | 67.5/66.6 | 32.5/33.4 | -0.351/-0.354 | 0.244/0.216 |
| 1st codon position | 3354/3354 | 41.2/41.4 | 11.2/11.3 | 23.6/23.4 | 24.0/23.9 | 64.8/64.8 | 35.2/35.2 | -0.272/-0.277 | 0.366/0.358 |
| 2nd codon position | 3354/3354 | 47.6/47.3 | 15.1/15.4 | 17.6/17.5 | 19.7/19.7 | 65.2/64.8 | 34.8/35.1 | -0.461/-0.461 | 0.132/0.122 |
| 3rd codon position | 3354/3354 | 47.9/46.5 | 10.6/12.6 | 24.5/23.6 | 17.0/17.4 | 72.4/70.1 | 27.6/30.0 | -0.323/-0.327 | 0.231/0.159 |
| *atp6*         | 510/510   | 47.6/47.3 | 13.3/13.3 | 21.6/20.2 | 17.5/19.2 | 69.2/67.5 | 30.8/32.5 | -0.377/-0.401 | 0.134/0.181 |
| *cox1*         | 1566/1566 | 45.0/44.0 | 12.6/13.8 | 22.3/21.8 | 20.1/20.4 | 67.3/65.8 | 32.7/34.2 | -0.461/-0.461 | 0.132/0.122 |
| *cox2*         | 570/570   | 40.2/38.9 | 12.8/14.2 | 23.9/24.7 | 23.2/22.1 | 64.1/63.6 | 36.0/36.3 | -0.255/-0.223 | 0.288/0.217 |
| *cox3*         | 643/643   | 46.8/47.1 | 12.4/12.4 | 20.5/19.4 | 20.1/21.0 | 67.3/65.8 | 32.7/34.2 | -0.390/-0.416 | 0.238/0.256 |
| *cytb*         | 1107/1107 | 43.8/43.9 | 13.7/14.1 | 22.1/21.6 | 20.3/20.4 | 65.9/65.5 | 34.0/34.5 | -0.329/-0.341 | 0.194/0.183 |
| *nad1*         | 891/891   | 45.3/44.9 | 11.0/11.4 | 20.4/21.1 | 23.2/22.6 | 65.7/66.0 | 34.2/34.0 | -0.379/-0.361 | 0.357/0.327 |
| *nad2*         | 879/879   | 48.9/48.9 | 10.5/11.1 | 21.0/19.8 | 19.6/20.1 | 69.9/68.7 | 30.1/31.2 | -0.398/-0.424 | 0.303/0.287 |
| *nad3*         | 357/357   | 51.3/49.9 | 7.8/9.2  | 21.8/20.7 | 19.0/20.2 | 73.1/70.6 | 26.8/29.4 | -0.402/-0.413 | 0.417/0.371 |
| *nad4*         | 1251/1251 | 47.2/46.9 | 13.7/14.2 | 19.4/19.0 | 19.6/19.8 | 66.6/65.9 | 33.3/34.0 | -0.417/-0.423 | 0.175/0.164 |
| *nad5*         | 261/261   | 48.3/49.4 | 9.2/8.8  | 27.2/27.2 | 15.3/14.6 | 75.3/76.6 | 24.3/23.4 | -0.279/-0.290 | 0.250/0.246 |
| *nad6*         | 1569/1569 | 42.4/41.4 | 13.3/15.2 | 23.8/23.6 | 20.5/19.8 | 66.2/65.0 | 33.8/35.0 | -0.281/-0.273 | 0.214/0.133 |
| *rrnL*         | 967/967   | 48.8/48.1 | 9.8/10.2 | 21.2/21.1 | 20.3/20.5 | 69.9/69.2 | 30.1/30.7 | -0.396/-0.390 | 0.348/0.333 |
| *rrnS*         | 742/743   | 40.1/39.7 | 12.0/12.2 | 28.3/28.2 | 19.5/19.9 | 68.4/67.9 | 31.5/32.1 | -0.172/-0.169 | 0.239/0.239 |
| NCR1           | 223/224   | 44.4/41.5 | 8.5/10.3 | 34.5/33.5 | 12.6/14.7 | 78.9/75.0 | 21.1/25.0 | -0.125/-0.107 | 0.191/0.179 |
| NCR2           | 249/183   | 51.8/50.3 | 7.6/6.6  | 22.1/24.6 | 18.5/18.6 | 73.9/74.9 | 26.1/25.2 | -0.402/-0.343 | 0.415/0.478 |
| tRNAs          | 1410/1410 | 38.5/37.9 | 12.4/13.0 | 30.2/29.3 | 17.9/17.5 | 68.2/67.5 | 31.8/32.4 | -0.115/-0.131 | 0.237/0.203 |
| Full genome    | 13,685/13,621 | 44.1/43.6 | 12.1/12.9 | 23.8/23.4 | 20.0/20.2 | 67.9/67.0 | 32.1/33.1 | -0.299/-0.301 | 0.245/0.221 |
nihonkaiensis, species of the most closely related genus, respectively. Nucleotide identity was 92.3% between Dib. latus and Dib. nihonkaiensis, 99.2% between Diphyllobothrium grandis (Blanchard, 1894) and D. balaenopterae (Lömberg, 1892), 99.3% between Spirometra decipiens (Diesing, 1850) and S. erinaceieuropaei (Rudolphi, 1819), respectively. Amongst all 36 genes, the majority were equal in size between D. interrupta and L. intestinalis, with the exception of rrnS, trnY and trnS2 which had only one base difference. Sequence identity ranged from 87.5% (trnD) to 100% (trnH, trnQ and trnV) (Table 1). Seven overlapping regions and 20 intergenic sequences (Gap1–20) were found in both genomes, identical in size and position, with the exception of GAP9, GAP11 and GAP17, which differed in size (Table 1).

**Protein-coding genes and codon usage**

Concatenated PCGs of the mitogenome of D. interrupta and L. intestinalis were both 10,062 bp in size, with an A+T content of 67.5% and 66.6%, respectively (Table 2). The high A+T content was mainly concentrated on the third codon position (72.4% for D. interrupta and 70.1% for L. intestinalis). The start and termination codons within the mitogenome of D. interrupta and L. intestinalis were identical to one another. GTG was identified as the initial codon for cox3, and ATG for the rest of the 12 PCGs (Table 1). For each PCG, however, all selected Diphyllobothriidea species shared the same start codon (Additional file 3: Table S3), indicating that it may be a synapomorphy within this order. Amongst the termination codons, five out of 12 were identified as TAA, six as TAG, while cox3 used a truncated T stop codon.

Codon usage, RSCU, and codon family proportion (corresponding to the amino acid usage) among D. interrupta and L. intestinalis were investigated (Additional file 4: Figure S1). The most abundant codon families were Phe, Leu2, and Ile within the two mitogenomes, which show a preference for the A+T-rich synonymous codons (Additional file 4: Figure S1). This corresponds to the high A+T bias of the two diphyllobothriid mitogenomes.

We measured the selective pressure acted upon amino acid replacement mutations by the ratio of non-synonymous (dN) to synonymous (dS) substitution rates for all 12 PCGs of D. interrupta vs L. intestinalis. Although the values (dN/dS) of atp6 (0.113), nad5 (0.111) and nad2 (0.110) genes were higher than cox1.
(0.007) and cox2 (0.008) genes (Fig. 2a), all PCGs were under strong negative (purifying) selection (dN/dS < 0.12).

**Transfer and ribosomal RNA genes**

The sizes of *rrnL* in both the *D. interrupta* and *L. intestinalis* mitogenome was 967 bp, with 68.4% and 67.9% A+T content, respectively. Similarly, the lengths of *rrnS* were 742 and 743 bp, with an A+T content of 68.2% and 67.5%, respectively (Table 2). All 22 commonly found tRNAs were present in the mitochondrial genome of *D. interrupta* and *L. intestinalis*, adding up to a 1,410 bp total concatenated length in both mitogenomes (Table 2). All tRNAs could be folded into the conventional cloverleaf structure, with the exception of *trnS* and *trnR*, which lacked DHU arms. The absence of DHU-arms in *trnS* and *trnR* has also been reported in the Caryophyllidea [35] and the Anoplocephalidae [36].

**Non-coding regions**

The two major non-coding regions (mNCRs), NCR1 and NCR2, were located between *trnY* and *trnL* and between *nad5* and *trnG*, respectively. The mNCRs were situated in the same location as all diphyllobothriideans surveyed to date (see Additional file 3 in our recent paper [35]). The NCR1 sequences of the two mitogenomes of *D. interrupta* and *L. intestinalis* were 223 and 224 bp in length with a heightened A+T bias of 78.9% and 75%, whereas the NCR2 sequences were 249 and 183 bp in size with 73.9% and 74.9% A+T content, respectively (Table 2). The NCR2 of *D. interrupta* contained six TRs (tandem repeats). Repeat units 1–5 were identical in nucleotide composition and size (34 bp). Repeat unit 6 was truncated with 29 bp (Fig. 3). TRs were also found in the NCR2 of *L. intestinalis*, and the consensus repeat (35 bp) was almost identical to that of *D. interrupta*, with an insertion of a single nucleotide A at the 17th position. Only four repeat units, however, could be found in the mitogenome of *L. intestinalis*, which contributed to the main difference in length of the complete mitogenome between the two genera. The last repeat unit was also truncated with 22 bp. Both NCR1 and the consensus repeat sequence in NCR2 of the two mitogenomes were
capable of forming stem-loop structures (Fig. 3, predicted by Mfold web server).

**Sliding window analyses and nucleotide diversity**
A sliding window analysis of the 12 PCGs and two rRNAs of *D. interrupta* and *L. intestinalis* indicated the nucleotide diversity $\pi$ to be higher in *nad5* (0.099), *nad6* (0.088), *nad2* (0.088), *nad4* (0.087) and *cox3* (0.087), whereas the most conserved genes were *rrnL* (0.036) and *rrnS* (0.04) (Fig. 2b).

**Phylogeny**
Both methods (BI and ML) produced phylogenograms with concordant branch topologies, thus only the latter was shown (Fig. 4). The phylogenetic tree indicated the ordinal topology to be Caryophyllidea + (Diphyllobothriidea + (Bothriocephalidea + (Proteocephalidea + Cyclophyllidea))). Within the family Diphyllobothriidae, *Ligula* and *Digamma* clustered with maximum nodal support (BP = 100 and BPP = 1), which formed a sister group with the genus *Dibothriocephalus*. This clade clustered together with *Diphyllobothrium*, then forming a sister group with *Spirometra*.

**Discussion**
The ordinal topology of Caryophyllidea + (Diphyllobothriidea + (Bothriocephalidea + (Proteocephalidea + Cyclophyllidea))) was consistent with previously identified interordinal relationships of tapeworms, when reconstructed from the dataset of nucleotide and amino acid sequences from partial mitogenomes, large and small subunit rRNA genes, and a combination of the former two [37]. Additionally, this relationship identified between tapeworm groups was congruent with latest mitochondrial phylogenomics [35]. However, a sister-group relationship between the orders Diphyllobothriidea and Bothriocephalidea has been suggested based on mitochondrial phylogenomics [38, 39]. This inconsistency may be due to the different methods of phylogenetics employed.

Within the family Diphyllobothriidae, the phylogenetic relationship of the three genera *Spirometra*, *Diphyllobothrium* and *Dibothriocephalus* was congruent with that of recent studies on the phylogenetics of Eucestoda based on mitogenomes, with the topology of *Spirometra*
In the present study, however, Ligula and Digramma were closely related to one another with maximum nodal support, then forming a sister group with the genus Dibothriocephalus. Inferred by the ITS2 rDNA sequence [13] and the 18S rDNA gene [42], the complexes of Ligula and Digramma have also been shown to be closely related to Dibothriocephalus. Further studies using the concatenated nucleotide sequences of 18S rDNA + 28S rDNA + rrnL + cox1, have again demonstrated the genus Dibothriocephalus to be the sister group of Ligula [5]. These phylogenetic results suggest that Dibothriocephalus is the most closely related genus to Ligula and Digramma.

However, mitogenome sequence identity between D. interrupta and L. intestinalis is 92.7%, which is much higher than between either of these species and the represented members of Dibothriocephalus (85.8–86.5%). Furthermore, high mitogenome sequence identity was also found between the congeneric in Dibothriocephalus (92.3%), Diphyllobothrium (99.2%) and Spirometra (99.3%). These results suggest that sequence differences between D. interrupta and L. intestinalis are of a degree expected between members of the same genus.

In one study, sequence identity of the cox1 and nad1 genes between D. interrupta and L. intestinalis has been shown to be 100% and 92.6% [11]; however, identity was deemed at 93.5% and 93.0% in the present study, respectively (Table 1). This inconsistency may be due to the use of partial sequence of cox1 and nad1 genes or resulting from the use of formalin-preserved specimens [12]. The gene cox1 is considered to be a useful barcode for metazoans [43], and widely employed for cestode studies [44–47]. The two mitochondrial genes cox1 and cytb have also been used to study the population genetic structure of L. intestinalis on a local and global scale [14]. However, a lower sequence identity was found in the nad2, nad4, nad5, nad6 and cox3 genes between D. interrupta and L. intestinalis (90–92%), in comparison to the moderate variation seen between the cox1, cytb and nad1 genes (Fig. 5). Additionally, the relatively loose selection pressure of nad5 (0.111) and nad2 (0.110) may accelerate the accumulation of non-synonymous substitutions, which would increase variation of the two genes [48]. These results suggest that the nad2, nad4, nad5, nad6 and cox3 genes should be considered as optimal candidates for genetic markers to be used for population genetics and species identification studies between the two closely related species, D. interrupta and L. intestinalis.

Conclusions

The complete mitogenomes of Ligula intestinalis and Digramma interrupta were sequenced and characterized. The mitogenomes of these two species show a higher identity to each other than to any species in closely related genera. The two mitogenomes consistently support D. interrupta to be a congeneric species with L. intestinalis. High sequence variation in the nad2, nad4, nad5, nad6 and cox3 genes between the two diphyllobothriids suggest that these five genes should be considered as optimal candidates for genetic markers when studying population genetics or looking to differentiate the two closely related species, D. interrupta and L. intestinalis.

Additional files

Additional file 1: Table S1. Primers used to amplify and sequence the mitochondrial genome of Digramma interrupta and Ligula intestinalis. (DOCX 15 kb)

Additional file 2: Table S2. The list of cestode species and outgroups used for comparative mitogenomic and phylogenetic analyses, and accession number, A+T content and skewness of different elements of each mitogenome. (XLSX 19 kb)
Abbreviations
dN: Non-synonymous substitutions; dS: Synonymous substitutions; mNCRs: major Non-coding regions; PCGs: Protein-coding genes

Acknowledgements
The authors would like to thank the editor and the two anonymous reviewers for the time they have invested into reviewing our manuscript.

Funding
This work was funded by the National Natural Science Foundation of China (31572368), the Major Scientific and Technological Innovation Project of Hubei Province (2015ABA045) and the Earmarked Fund for China Agriculture Research System (CARS-45-15).

Availability of data and materials
The datasets supporting the conclusions of this article are available in the GenBank international nucleotide sequence repository under accession numbers MF671696 and MF671697.

Authors’ contributions
WXL designed the experiments, performed the analysis and wrote the manuscript. DZ performed the laboratory work and the phylogenetic analysis. KB analysed the data. WXL, PPF, DZ, KB, BWX, HZ, ML, SGW and GTW contributed to the interpretation of the findings. All authors read and approved the final manuscript.

Ethics approval
Tapeworms were collected from fish in accordance with the recommended guidelines for animal experimentation by the Chinese Association for Laboratory Animal Sciences.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Key Laboratory of Aquaculture Disease Control, Ministry of Agriculture, and State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, People’s Republic of China. 2University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China. 3South Devon College University Centre, Long Road, Paignton TQ4 7EJ, UK. 4Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture, Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, WuXi 214081, China.

Received: 8 February 2018 Accepted: 21 May 2018
Published online: 30 May 2018

References
1. Brabec J, Kuchta R, Scholz T. Paraphyly of the Pseudophyllidea (Platyhelminthes: Cestoda): Circumscription of monophyletic clades based on phylogenetic analysis of ribosomal RNA. Int J Parasitol. 2006;36:1535–41.
2. Kuchta R, Scholz T, Brabec J, Bray RA. Suppression of the tapeworm order Pseudophyllidea (Platyhelminthes: Eucestoda) and the proposal of two new orders, Bothriocephalidea and Diphyllobothriidea. Int J Parasitol. 2008;38:49–55.
3. Waeschchenbach A, Webster BL, Bray RA, Littlewood DTJ. Added resolution among ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. Mol Phylogenet Evol. 2007;45:311–25.
4. Kuchta R, Scholz T. Diphyllobothriidea. In: Cairns JN, Jensen J, editors. Planetary Biodiversity Inventory (2008–2017): Tapeworms from vertebrate bowels of the earth. The University of Kansas Natural History Museum Special Publication No. 25, 2017.
5. Waeschchenbach A, Brabec J, Scholz T, Littlewood DTJ, Kuchta R. The catholic taste of broad tapeworms-multiple routes to human infection. Int J Parasitol. 2017;47:831–43.
6. Dubinin MN. Tapeworms (Cestoda, Ligulidae) of the Fauna of the USSR. New Delhi: Ameirin Publishing; 1988.
7. Cowx IG, Rollins D, Turneweaze R. Effect of Ligulula intestinalis on the reproductive capacity of Rastrieneoba argentea in Lake Victoria. J Fish Biol. 2008;73:2249–60.
8. Sohn WM, Na BK, Jung SK, Kim KH. Mass death of predatory carp, Chanodichthys erythrophus, induced by plerocercoids of Ligulula intestinalis (Cestoda: Diphyllobothriidae). Korean J Parasitol. 2016;54:365–8.
9. Cholodkovsky N. Notes helminthologiques. Ann Mus Zool Acad Sci Russ, Petrograd. 1915;20:164–6.
10. Liao XH, Liang ZX. Distribution of ligulid tapeworms in China. J Parasitol. 1987;73:36–48.
11. Li JJ, Liao XH. The taxonomic status of Digramma (Pseudophyllidea: Ligulidae) inferred from DNA sequences. J Parasitol. 2003;89:792–9.
12. Liu GT, Nie P, Yao WJ, Wang GT, Gao Q. Is the genus Digramma synonymous to the genus Ligula (Cestoda: Pseudophyllidea)? Evidence from ITS and 5′ end 28S rDNA sequences. Parasitol Res. 2003;89:419–21.
13. Logan FJ, Horák A, Štefka J, Aydogdu A, Scholz T. The phylogeny of diphyllobothriid tapeworms (Cestoda: Ligulidae) based on ITS-2 rDNA sequences. Parasitol Res. 2004;94:10–5.
14. Poult D, Štefka J, Hylita V, Lek S, Scholz T, Legal I, et al. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite Ligula intestinalis (Cestoda: Diphyllobothriidae). Int J Parasitol. 2006;36:1465–79.
15. Zhang D, Zou H, Wu SG, Li M, Jakovlíc I, Zhang J, et al. Sequencing, characterization and phylogenomics of the complete mitochondrial genome of Dactylogyrus larvatus (Monogenea: Dactylogyridae). J Helminthol. 2017; https://doi.org/10.1017/S0022149X17000578.
16. Zhang D, Zou H, Wu SG, Li M, Jakovlíc I, Zhang J, et al. Sequencing of the complete mitochondrial genome of a fish-parasitic flatworm Paratretochonides inermis (Platyhelminthes: Monogenea): rRNA gene arrangement reshuffling and implications for phylogeny. Parasit Vectors. 2017;10:462.
17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
18. 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.
19. Laslett D, Canback B. ARNEN: a program to detect tRNA genes in metazoa mitochondrial nucleotide sequences. Bioinformatics. 2008;24:172–5.
20. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: Improved de novo mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9.
21. Zhang D. MitoTool software: Secondary MitoTool software. Wuhan Institute of Hydrobiology, Chinese Academy of Sciences; https://github.com/dongzhang0725/MitoTool. Accessed 24 May 2018.
22. Hadley W. ggplot2: Elegant graphics for data analysis. J Stat Softw. 2010;35:65.
23. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
24. Züker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31:3406–15.
25. Zhang Z, Li J, Zhao XG, Wang J, Wang GKS, Yu J. Kaks_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics. 2006;4:259–63.
26. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data and inference of selection parameters. Bioinformatics. 2009;25:155–2.
27. Zhang D. BioSuite software: Secondary BioSuite software. https://github.com/dongzhang0725/BioSuite. Accessed 24 May 2018.
28. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–8.
29. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56:654–77.

30. Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McLnemey JD. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evol Biol. 2006;6:29.

31. Silvestro D, Michalak I. raxmlGUI: a graphical front-end for RAxML. Org Divers Evol. 2011;12:335–7.

32. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.

33. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44:242–5.

34. Le Thi, Blair D, McManus DP. Mitochondrial genomes of parasitic flatworms. Trends Parasitol. 2002;18:206–13.

35. Li WX, Zhang D, Boyce K, Xi BW, Zou H, Wu SG, et al. The complete mitochondrial DNA of three monozoic tapeworms in the Caryophyllidea: a mitogenomic perspective on the phylogeny of eucestodes. Parasit Vectors. 2017;10:314.

36. Guo A. Moniezia benedeni and Moniezia expansa are distinct cestode species based on complete mitochondrial genomes. Acta Trop. 2016;166:287–92.

37. Waeschenbach A, Webster BL, Littlewood DTJ. Adding resolution to ordinal levels of relationships of tapeworms (Platyhelminthes: Cestoda) with large fragments of mtDNA. Mol Phylogenet Evol. 2012;63:834–47.

38. Feng Y, Feng HL, Fang YH, Su YB. Characterization of the complete mitochondrial genome of Khawia sinensis belongs among platyhelminths, cestodes. Exp Parasitol. 2017;177:35–9.

39. Zhang X, Duan JY, Shi YL, Jiang P, Zeng J, Wang ZQ, et al. Comparative mitochondrial genomics among Spirometra (Cestoda: Diphyllobothriidae) and the molecular phylogeny of related tapeworms. Mol Phylogenet Evol. 2017;117:75–82.

40. Guo A. Characterization of the complete mitochondrial genome of the cloacal tapeworm Oocotzenia megaloceph (Cestoda: Hymenolepididae). Parasit Vectors. 2016;9:490.

41. Liu GH, Li C, Li JY, Zhou DH, Xiong RC, Lin RQ, et al. Characterization of the complete mitochondrial genome sequence of Spirometra enoziarieuropeae (Cestoda: Diphyllobothriidae) from China. Int J Biol Sci. 2012;8:640–9.

42. Zhang X, Duan JY, Wang ZQ, Jiang P, Liu RD, Cui J. Using the small subunit of nuclear ribosomal DNA to reveal the phylogenetic position of the plerocercoid larvae of Spirometra tapeworms. Exp Parasitol. 2017;175:1–7.

43. Hebert PDR, Cywinska A, Bull SL, deWaard JR. Biological identifications through DNA barcodes. Proc R Soc Lond B. 2003;270:313–21.

44. Aitomaa S, Sheldrick M, Yamada M, Uchikawa R, Tegoshi T, Takeda K, et al. Mitochondrial DNA divergence in populations of the tapeworm Diphyllobothrium nihonkaiense and its phylogenetic relationship with Diphyllobothrium klenovskii. Parasitol Int. 2009;58:22–8.

45. Badaraco JL, Ayala FJ, Bart JM, Gottstein B, Haag RL. Using mitochondrial and nuclear markers to evaluate the degree of genetic cohesion among Echinococcus populations. Exp Parasitol. 2008;119:453–9.

46. Bazsalovicsova E, Kralova-Hromadova I, Stefka J, Scholz T, Hanzelova V, Vavrova S, et al. Population study of Atractolytocestus huronensis (Cestoda: Caryophyllidea), an invasive parasite of common carp introduced to Europe: mitochondrial cox1 haplotypes and intragenomic ribosomal ITS2 variants. Parasitol Res. 2011;109:125–31.

47. Hu D, Song X, Wang N, Zhong X, Wang J, Liu T, et al. Molecular identification of Echinococcus granulosus on the Tibetan Plateau using mitochondrial DNA markers. Genet Mol Res. 2015;14:13915–23.

48. Huyse T, Buchmann K, Littlewood DT. The mitochondrial genome of Gyrodactylus denajovoides (Platyhelminthes: Monogenea): a mitogenomic approach for Gyrodactylus species and strain identification. Gene. 2008;417:27–34.