International Journal for Parasitology 42 (2012) 259–267

Contents lists available at SciVerse ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Substitution saturation and nuclear paralogs of commonly employed phylogenetic markers in the Caryophyllidea, an unusual group of non-segmented tapeworms (Platyhelminthes) ☆

Jan Brabec [a,*], Tomáš Scholz [a], Ivica Králová-Hromadová [b], Eva Bazsalovicsová [b], Peter D. Olson [c]

[a] Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic and Faculty of Science, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic
[b] Institute of Parasitology, Slovak Academy of Sciences, Hlinkova 3, 040 01 Košice, Slovakia
[c] Department of Zoology, The Natural History Museum, London SW7 5BD, United Kingdom

ARTICLE INFO

Article history:
Received 21 November 2011
Received in revised form 10 January 2012
Accepted 11 January 2012
Available online 16 February 2012

Keywords:
Caryophyllidea
Cestoda
Split-decomposition
Numts
rDNA
mtDNA
cox1
nad3

ABSTRACT

Caryophyllidean cestodes (Platyhelminthes) represent an unusual group of tapeworms lacking serially repeated body parts that potentially diverged from the common ancestor of the Eucestoda prior to the evolution of segmentation. Here we evaluate the utility of two nuclear and two mitochondrial molecular markers (ssrDNA and lsrDNA, *nad*3 and *cox*1) for use in circumscribing generic boundaries and estimating interrelationships in the group. We show that these commonly employed markers do not contain sufficient signal to infer well-supported phylogenetic estimates due to substitution saturation. Moreover, we detected multiple *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1 haplotypes within individuals, indicating a history of gene exchange between the mitochondrial and nuclear genomes. The presence of such nuclear paralogs (i.e. numts), to our knowledge described here in cestodes for the first time, together with the results of phylogenetic, saturation and split-decomposition analyses all suggest that finding informative markers for estimating caryophyllidean evolution is unusually problematic in comparison to other major lineages of tapeworms.

© 2012 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The Caryophyllidea (Platyhelminthes: Cestoda) is an unusual group of tapeworms lacking serially-repeated body structures. Whether their unsegmented condition represents divergence prior to the evolution of segmentation in tapeworms, or alternatively, secondary loss of segmentation, remains controversial. Molecular phylogenetic analyses have identified the group as either the sister group to the other true tapeworms (i.e. Eucestoda), or in a more derived position as a sister lineage to the segmented Diphyllobothriidea (Olson et al., 2001, 2008; Brabec et al., 2006; Waeschenbach et al., 2007; for reviews see Mackiewicz, 2003 and Olson and Tkach, 2005). In addition to lacking the hallmark feature of tapeworms, they are also unique among extant groups in parasitising oligochaetes, rather than arthropods, as first intermediate hosts. Adult worms infect predominately benthic feeding siluriform and cypriniform fishes, are modestly diverse in comparison to other cestode orders (41 genera, 150 species) and are nearly cosmopolitan in distribution (Mackiewicz, 1972, 1994; Oros et al., 2008, 2010).

Although their phylogenetic position has been assessed using both molecular and morphological data (e.g. Hoberg et al., 1997, 2001; Olson et al., 2001), they remain one of the few tapeworm groups to date whose interrelationships have been largely unexplored, with the first morphological cladistic-based estimate published only recently by Oros et al. (2008). Their assessment, based on 30 morphological characters of all known genera, differed considerably from the most recent classification of the group (i.e. Mackiewicz, 1994) and demonstrated that most characters commonly used to circumscribe taxa exhibit high levels of homoplasy. Moreover, there appeared to be little geographic structure related to their phylogenetic history, suggesting that either the extensive movement of taxa through time has obscured their centre of origin, or that morphology-based phylogenetic estimates are misleading.

Here we evaluate the suitability of four commonly used nuclear and mitochondrial (mt) markers for phylogenetic inference and taxonomic circumscription: the large and small nuclear ribosomal RNA subunits (lsrDNA and ssrDNA) and the mt cytochrome c

☆ *Note:* Nucleotide sequence data newly reported in this paper are available in the GenBank™, EMBL, DDBJ databases under the Accession Nos. JQ033989–JQ034148 and JN004224–JN004265.

* Corresponding author. Tel.: +420 38 7775428; fax: +420 38 5310388.
E-mail address: brabcak@paru.cas.cz (J. Brabec).

0020-7519/$36.00 © 2012 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2012.01.005

oxidase subunit 1 (barcoding region) and nicotinamid dehydrogenase subunit 3 (*cox*1 and *nad*3) genes. Of these, the ribosomal genes have been used most extensively for estimating interrelationships within and among tapeworm orders, having been found to provide informative characters across a broad range of divergences (e.g. Lockyer et al., 2003; de Chambrier et al., 2004; Brabec et al., 2006; Healy et al., 2009; Olson et al., 2010). In contrast, mt genes have been employed with less frequency, particularly for studies aimed at resolving interrelationships above the level of genus, and have been used most extensively for studies of the highly derived cyclophyllidean cestodes of medical importance (e.g. Taeniidae; see Olson and Tkach, 2005 for review). In a study of these markers in caryophyllidean cestodes we found intra-individual variation in the primary sequence of both mt genes tested and were able to demonstrate the existence of multiple mt haplotypes that most likely represent paralogs that have become incorporated into the nuclear genome (i.e. numts; Bensasson et al., 2001).

2. Materials and methods

2.1. DNA amplification and sequencing workflow

Table 1 lists 25 specimens representing 19 species sequenced. Genomic DNA was extracted using a standard phenol chloroform extraction method (Sambrook and Russell, 2001). The following individual genes were amplified by PCR using the four primer pairs: nearly complete ssrDNA with WormA and WormB (Littlewood and Olson, 2001), the D1–D3 region of lsrDNA with LSU5 and 1500R (Littlewood et al., 2000; Olson et al., 2003) and the mt region comprising *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1 with CFCYT1 (5′ GCA GGT TAC TTT GAT ATA G 3′) and CRCYT2 (5′ CCA AAA AAC CAA AAC AT 3′) specifically designed for caryophyllidean cestodes; see Bazsalovicsová et al. (2011) and Scholz et al. (2011) for the details. Cycling conditions for both rDNA regions were as follows: denaturation for 5 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 55 °C, 2 min at 72 °C and completed by 7 min at 72 °C. Cycling conditions for the mt region were exactly the same except for a lower annealing temperature of 50 °C. All products were verified on a 1% agarose gel and purified either using exonuclease I and shrimp alkaline phosphatase (Werle et al., 1994) or with the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Where the presence of multiple mt variants was indicated, the PCR amplicons of the *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1 fragment were cloned using the pGEM®-T Easy Vector System (Promega, Madison, USA). For each amplicon cloned, 10 colonies were selected randomly and sequenced. BigDye® Terminator v3.1 cycle sequencing reagents and a PRISM 3130xl automatic sequencer (Applied Biosystems, Foster City, USA) were used for bidirectional sequencing of all products using the PCR primers together with internal primers 600F, 600R, 1600F, 1600R (ssrDNA), 300F, 400R, 900F, ECD2 (lsrDNA), and CFCYT2 (5′ ACT AAG TGT TTT CAA AA 3′) and CRCYT0 (5′ GTG TTT TGA AAA CAC YYA GT 3′); mtDNA (Littlewood and Olson, 2001; Olson et al., 2003; Bazsalovicsová et al., 2011). Sequences were assembled and inspected for errors using Geneious Pro ver. 5.1.6 (Drummond et al., 2010. Geneious v5.1. Available from http://www.geneious.com) and aligned using the program MAFFT (Katoh et al., 2005) using either the E-INS-*i* algorithm (in the case of rDNA data) or the G-INS-*i* and translational alignment (using the echinoderm translation code – translation table No. 9) in the case of mt data. The resulting

Table 1
List of the specimens and genes sequenced within the scope of this study.

Family	Host species	Locality	Collection No.[a]	Nuclear[c]		Mitochondrial[c]	
Species			(Field sample No.)	ssrDNA	lsrDNA	*nad*3	*cox*1
Lytocestidae							
Atractolytocestus huronensis	*Cyprinus carpio*	Hungary	C-472 (AH/HU/22)	JQ034132	JQ034115	JQ033989[d]	*HM480476*
Atractolytocestus sagittatus	*Cyprinus carpio*	Japan	C-340 (JP8b)	JQ034133	JQ034116	JQ033998	*JF424669*[d]
Caryophyllaeides fennica	*Rutilus rutilus*	Slovakia	C-1 (CF/SK/135/06)	JQ034136	JQ034119	JQ033990 JQ033995–7	JQ034060–3
Caryophyllaeides fennica	*Leuciscus leuciscus*	Finland	C-1 (TS06/123)	JQ034135	JQ034118	JQ033991–4	JQ034052 JQ034057–9
Djombangia penetrans	*Clarias batrachus*	India	C-542 (TS09/50)	JQ034142	JQ034125	JQ034021–4	JQ034084–7[d]
Khawia armeniaca	*Coregonus lavaretus*	Armenia	C-48 (KA/AR/89)	*JN004246*	*JN004257*	*JN004235*[d]	*JN004224*
Khawia japonensis	*Cyprinus carpio*	Japan	C-348 (TS04/148)	*JN004247*	*JN004258*	*JN004236*	*JN004225*
Khawia rossittensis	*Carassius auratus*	Slovakia	C-214 (KR/SK/231/07)	*JN004249*	*JN004260*	*JN004238*	*JN004227*
Khawia rossittensis	*Carassius auratus*	Japan	C-214 (JP226a)	*JN004248*	*JN004259*	*JN004237*	*JN004226*
Khawia saurogobii	*Saurogobio dabryi*	China	C-537 (TS09/135)	*JN004251*	*JN004262*	*JN004240*	*JN004229*
Khawia sinensis	*Cyprinus carpio*	China	C-46 (TS09/104)	*JN004254*	*JN004265*	*JN004243*	*JN004232*
Khawia sinensis	*Cyprinus carpio*	Japan	C-46 (JP16)	*JN004253*	*JN004264*	*JN004242*[d]	*JN004231*
Khawia sinensis	*Cyprinus carpio*	Slovakia	C-46 (KS/SK/488/05)	*JN004250*	*JN004261*	*JN004239*	*JN004228*
Lytocestus indicus	*Clarias batrachus*	India	C-539 (TS09/67)	JQ034145	JQ034128	JQ034034–5	JQ034097–8[d]
Monobothrioides sp.	*Auchenoglanis* sp.	Sudan	C-504 (TS06/74)	JQ034146	JQ034129	JQ034036–43	JQ034099–106[d]
Capingentidae							
Breviscolex orientalis	*Hemibarbus barbus*	Japan	BMNH-2001.1.30.1–4[b]	*AF286978*	*AF286910*	JQ033999–4000[d]	JQ034055–6
Caryophyllaeidae							
Caryophyllaeus brachycollis sk35	*Barbus meridionalis*	Slovakia	C-51 (CB/SK/365/05)	JQ034137	JQ034120	JQ034001	JQ034064
Caryophyllaeus sp. sk96	*Abramis brama*	Slovakia	C-2 (Csp/SK/41/07B)	JQ034141	JQ034124	JQ034015–20	JQ034078–83
Caryophyllaeus laticeps sk36	*Abramis sapa*	Slovakia	C-2 (CL/SK/131/05)	JQ034139	JQ034122	JQ034014[d]	JQ034077
Caryophyllaeus laticeps sk97	*Abramis brama*	Slovakia	C-2 (CL/SK/41/07A)	JQ034140	JQ034123	JQ034006–13	JQ034069–76
Caryophyllaeus laticeps 04/112	*Cyprinus carpio*	Slovakia	C-2 (TS04/112)	JQ034138	JQ034121	JQ034002–5	JQ034065–8
Glaridacris catostomi	*Catostomus commersoni*	USA	C-5 (TS08/51)	JQ034143	JQ034126	JQ034025–7[d]	JQ034088–90
Hunterella nodulosa	*Catostomus commersoni*	USA	C-321 (TS08/53)	JQ034144	JQ034127	JQ034028–33[d]	JQ034091–6
Monobothrium hunteri	*Catostomus commersoni*	USA	C-505 (TS08/54)	JQ034147	JQ034130	JQ034044–7[d]	JQ034107–10
Wenyonia virilis	*Synodontis schall*	Sudan	C-503 (SUD382)	JQ034148	JQ034131	JQ034048–51	JQ034111–4

[a] Helminthological collection of the Institute of Parasitology, České Budějovice, Czech Republic.
[b] Collection of The Natural History Museum, London, England.
[c] GenBank accession numbers. Accession numbers in italics were already published by the authors (Olson et al., 2001; Bazsalovicsová et al., 2011, 2012; Scholz et al., 2011).
[d] Indicates presence of GTG initiation codon (see Section 3.1).

automated alignments were checked in MacClade ver. 4.08 (Maddison, D.R., and Maddison, W.P., 2005. MacClade 4: Analysis of phylogeny and character evolution. Version 4.08a. Analysis of phylogeny and character evolution. Version 4.08a. http://www.macclade.org) and ambiguously aligned positions were manually excluded prior to phylogenetic analyses. This resulted in the removal of 16.6%, 23.1%, 4.4%, and 0% of positions from the ssrDNA, lsrDNA, *cox*1 and *nad*3 alignments, respectively.

2.2. Estimating saturation in the data partitions

To examine the degree of substitution saturation in our datasets, we adapted two distinct methodologies. For the rDNA, we made use of the methodology specifically designed for this type of data described by Struck et al. (2008) in which we first examined the sequence variation across both rDNA alignments by performing sliding window analyses, and based on those results, grouped the alignment positions into classes of identity (at 10% increments in degree of variation). For each class, we determined its *C* factor (i.e. the ratio of the standard deviation of the transition/transversion distribution to that of the uncorrected *p* (i.e. patristic distance) distribution), and *O*/*E* ratio, which is calculated from the ratio of the consistency index (*CI*) of the complete dataset to the *CI* of the reduced datasets (i.e. each identity class). To determine *CI* values, equally weighted heuristic searches with 1,000 repetitions of tree-bisection-reconnection (TBR) branch-swapping were performed using PAUP* ver. 4.0b10 (Swofford, 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, MA, USA) which was also used to obtain transition/transversion ratios and uncorrected *p* distances.

To test for substitution saturation within the individual codon positions of the mt genes we employed the entropy-based index of substitution saturation approach of Xia et al. (2003). In short, this approach tests whether the observed entropy in the sequences is significantly smaller than the entropy of "full substitution saturation". The ratio of observed entropy to the entropy of full substitution saturation is the index of substitution saturation (I_{SS}) and when this is not significantly smaller than the critical I_{SS} value (the value at which the sequences start to fail to recover the correct tree; $I_{SS.C}$), one can conclude that the sequences have experienced severe substitution saturation (see Xia and Lemey, 2009). All of the information entropy-based index analyses were done using the software DAMBE (Xia and Xie, 2001), following the practice of Xia and Lemey (2009).

2.3. Phylogenetic analyses

To produce phylogenetic estimates from individual datasets and to compare their contributions with the overall phylogenetic evidence, we ran a set of analyses based on maximum likelihood (ML) and Bayesian inference (BI) criteria. ML and BI were conducted on each individual dataset, both complete and with saturated positions excluded, and in case of mt genes, also including analyses of the amino acid alignments. To root the trees as well as to compare divergences within the Caryophyllidea with that of another group, we included representatives of the possible sister lineage Diphyllobothriidea (Waeschenbach et al., 2007; Olson et al., 2008). ML analyses were conducted using the program RAxML ver. 7.2.8-ALPHA (Stamatakis, 2006; Stamatakis et al., 2008), employing the GTR + Γ substitution model in case of nucleotide alignments, and the MtZoa model (Rota-Stabelli et al., 2009) with Γ rate heterogeneity in the case of amino acid alignments. All model parameters and bootstrap nodal support values (100 repetitions) were estimated using RAxML and, in the case of mt data, each codon position was used as an independent partition with independent optimisation of ML model parameters. BI trees were constructed using MrBayes ver. 3.1.2 (Ronquist and Huelsenbeck, 2003), running two independent MCMCMC runs of four chains for 5 million generations and sampling tree topologies every 1,000th generation. As in ML analyses, the GTR + Γ model was employed. Burn-in periods were set to 1.25 million generations according to the standard deviation of split frequencies values (<0.01) and by checking for MCMC convergence using AWTY (Nylander et al., 2008).

2.4. Estimating the strength of the phylogenetic signal

To visualise the number of positions that support bipartitions within the ML trees constructed, and to identify other strongly supported splits contradicting those found in the phylogenetic trees (i.e. to test the level of the background noise in the alignments), we performed split decomposition analysis (SDA) using the software Splits Analyses Methods (SAMS; Wägele and Mayer, 2007). SAMS analyses were run under default settings with the following exceptions: gaps were evaluated as a new state and the search for the total number of supporting characters of each split was determined with splits = search option. Each time, 50 highly supported splits were plotted and only those actually present in the tree with highest likelihood were depicted on corresponding phylograms. To make the SDA as well as the phylogenetic analyses of mt data more straightforward, we included only one randomly chosen "functional" haplotype representative of each of the caryophyllidean species that bear multiple mt variants, so that the final datasets to be analysed consisted of 28 sequences including three outgroup taxa. In order to choose a single haplotype for each taxon, we first excluded haplotypes with coding regions interrupted by one or more in-frame stop codons, and then chose randomly among the remaining presumed "functional" haplotypes for each taxon, which always grouped together in phylogenetic analyses (Supplementary Fig. S1) and exhibited little intra-specific divergence (Supplementary Table S1).

3. Results and discussion

3.1. Primary sequence characteristics

The combined data (nuclear ssrDNA, D1–D3 lsrDNA, and mt *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1) produced 4,277 unambiguously aligned coding positions for analysis. The length of the amplified fragments varied from 2,074 to 2,190 bp in ssrDNA, 1,539 to 1,665 bp in lsrDNA and 1,198 to 1,324 bp in the *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1 mt region. Comparison of the latter region with available tapeworm mt genomes (Lupi et al., 2010) showed the codon GTG to serve as an additional initiation codon to the universal ATG in the *nad*3 and *cox*1 genes of 12 of the 25 samples. Moreover, based on the inferred termination codons of our *nad*3 data, we found no evidence to support the idea that TAA codes for tyrosine in caryophyllidean tapeworms, rather than the standard stop codon in the universal genetic code. These features of the flatworm mt genetic code support the previous findings of Nakao et al. (2000) in cestodes and Telford et al. (2000) in the rhabditophoran platyhelminths. Thus, the "echinoderm mt translation code" (GenBank's translation table 9) should be used to translate cestode mt nucleotide sequences as opposed to the "flatworm translation code" (translation table 14) available in common bioinformatic resources.

3.2. Nuclear copies of mitochondrial genes (numts)

Cloned sequences showed a total of 59 distinct mt haplotypes in 13 of the 25 samples, with 2–8 unique mt haplotypes per individ-

Table 2
Number and sequence characteristics of *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1 mtDNA haplotypes observed within the scope of this study.

Specimen	All[a]	Numts[b]	Indels[c]	Stops[d]	Mutations[e]	Substitutions (non-numts)[f]	Substitutions (non-numts/numts)[g]	Identical[h]	Non-numt length/AT%[i]	Numt length/AT%[j]
Caryophyllaeides fennica Slovakia	4	3	1–2	1–7	0–98/1–46	NA	2–141	1,151	1,299/60.8%	1,298–1,303/60.2%
Caryophyllaeides fennica Finland	4	1	1	1	0–2/1–2	2–4	1–3	1,294	1,299/60.4%	1,300/60.5%
Djombangia penetrans	4	3	1–10	1–12	4–85/12–111	NA	24–211	943	1,303/59.3%	1,198–1,302/60.4%
Monobothrioides sp.	8	0	0	0	0–30/0–14	2–51	NA	1,257	1,320–1,321/57.7%	NA
Lytocestus indicus	2	0	0	0	0/3	6	NA	1,296	1,301–1,302/58.6%	NA
Breviscolex orientalis	2	1	2	11	4/8	NA	18	1,258	1,290/56.1%	1,275/56.2%
Caryophyllaeus sp. sk96	6	0	0	0	0–1/0–4	1–5	NA	1,300	1,308/60.6%	NA
Caryophyllaeus laticeps sk97	8	6	1–4	2–11	0–87/1–39	8	3–127	1,097	1,297–1,298/59.2%	1,288–1,296/59.3%
Caryophyllaeus laticeps 04/112	4	1	2	14	2–113/0–49	4–36	173–176	1,107	1,298/58.4%	1,296/60.5%
Glaridacris catostomi	3	1	1	9	0–1/0–1	1	1–2	1,319	1,322/58.7%	1,321/58.8%
Hunterella nodulosa	6	0	0	0	0–2/0–4	3–9	NA	1,284	1,300/58.9%	NA
Monobothrium hunteri	4	2	1	3–5	1–40/1–17	4	6–68	1,245	1,324/58.7%	1,322–1,323/59.4%
Wenyonia virilis	4	3	1–2	4–6	0–3/2–5	NA	4–7	1,273	1,288/56.6%	1,287–1,288/56.6%

[a] Number of all unique haplotypes observed.
[b] Number of observed putative nuclear mitochondrial pseudogenes (“numt” haplotypes).
[c] Number of indels within *nad*3 + *cox*1 coding regions of “numt” haplotypes.
[d] Number of in-frame stop codons within *nad*3 + *cox*1 coding regions of “numt” haplotypes.
[e] Number of synonymous/non-synonymous substitutions within *nad*3 + *cox*1 coding regions.
[f] Substitution differences between all pairs of “non-numt” haplotypes.
[g] Substitution differences between all pairs of both “non-numt” and “numt” haplotypes.
[h] Number of identical sites (all “non-numt” and “numt” aligned).
[i] Ungapped length of “non-numt” haplotypes and their mean AT content.
[j] Ungapped length of “numt” haplotypes and their mean AT content.

ual that differed from each other by both synonymous and non-synonymous mutations as well as reading frame shifts (see Table 2). Several molecular processes have been described to explain the existence of multiple mt haplotypes, including (i) duplications within the mt genome, (ii) independent populations of mitochondria in cells (i.e. heteroplasmy), and (iii) nuclear mt pseudogenes, or numts (e.g. Bensasson et al., 2001; Song et al., 2008; White et al., 2008). While heteroplasmy is difficult to evaluate, none of the mt genomes of cestodes sequenced to date shows evidence of protein-coding gene duplications (see Jia et al., 2010), including the caryophyllidean, *Caryophyllaeus laticeps* (D.T.J. Littlewood, The Natural History Museum, UK, personal communication), and thus such duplications are unlikely to be the source of variation found in the present study. Approximately one-third of the 59 mt haplotypes appeared to be non-functional based on the presence of one or more indels, and often several in-frame stop codons, frequently accompanied by a large number of nucleotide substitutions compared with haplotypes without broken reading frames (Song et al., 2008; see Table 2). These characteristics, together with the multiple mt *nad*1 gene lineages detected in trematodes of the genus *Paragonimus* (van Herwerden et al., 2000a), support the notion that the source of the variation in caryophyllidean tapeworms is the presence of numts.

Phylogenetic analyses (Fig. 1) showed no clustering of the “numt” haplotypes. Instead, with just two exceptions (*Caryophyllaeus laticeps* 04/112 hpt1 and *Caryophyllaeides fennica* Slovakia hpt1), all of the identifiable “numt” haplotypes grouped with their respective “functional” ones, indicating the transfer of mtDNA to the nucleus has not been an ancient, clade-specific event in caryophyllidean evolution, but rather that information from the mt genome has been copied into the nuclear genome independently on multiple occasions. Phylogenetic analysis (Fig. 1) shows that the numts all cluster according to species and are thus conserved enough to recover the pattern of speciation in the group, albeit comparisons among paralogous genes will produce differing estimates of divergence.

3.3. Comparison of the data partitions

The aligned ssrDNA and lsrDNA data contained 2,245 and 1,828 positions of which 1,872 and 1,406 remained after the exclusion of ambiguously aligned positions, respectively. Sliding window analyses revealed that the positions could be divided into the classes of identity of 0–40% in the ssrDNA and 0–60% in the lsrDNA data parts, with the vast majority of positions belonging to the lowest 0–10% variability class in both genes. For the purposes of the *C* factor and *O*/*E* ratio calculations, lsrDNA classes of identity with variability of 30–60% were pooled to avoid estimates based on small sample sizes. Omitting the ambiguously aligned positions from our ssrDNA and lsrDNA datasets resulted in total exclusion of highly variable classes (i.e. 40–60% for ssrDNA and 70–80% for lsrDNA), as well as a significant decrease in the number of positions belonging to other classes (see Supplementary Table S2).

Our *C* factor and *O*/*E* ratio values (Supplementary Table S2) resemble those of the analysis of Struck et al. (2008). As in their study we observed a sharp drop by an order of a magnitude in *C* factor values linked to the growing genetic distances between pairs of taxa and thus growing variability of classes of identity. By analysing their datasets, Struck et al. (2008) had found the *C* factor of 20 to represent the value below which the classes of identity should be treated as saturated. Based on our data, the *C* factor drops to the boundary value of 20 (30–40% class of ssrDNA) or even slightly below it (19 of 30–60% class in lsrDNA), indicating that the highly variable classes of identity are reaching saturation levels.

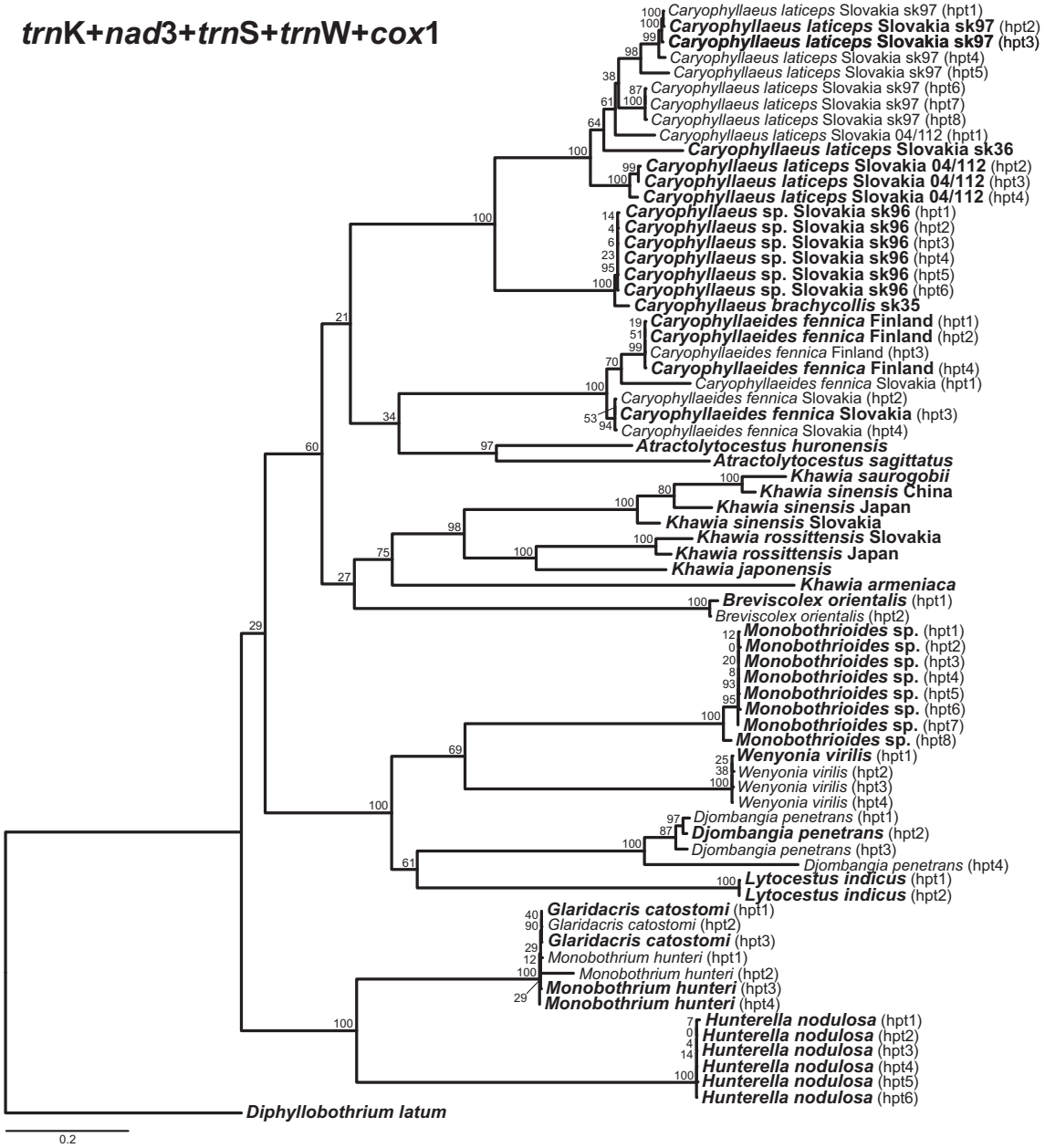

Fig. 1. Phylogenetic analysis of putative nuclear mitochondrial pseudogenes (numts) distribution in the Caryophyllidea based on the analysis of the entire *trn*K + *nad*3 + *trn*S + *trn*W + *cox*1mtDNA region by maximum likelihood. Mitochondrial haplotypes (hpt) in bold type represent the "functional" ("non-numt") haplotypes without in-frame stop codons or frame shift mutations.

Even if the statistical power of our analyses is slight due to the relatively small number of taxa and positions in the classes, we can conclude that by excluding the ambiguously aligned positions from our datasets we also remove the most variable, clearly saturated classes of identity. While any further exclusion of the most variable classes of identity remaining in each of the rDNA data partition would result in a significant decrease of parsimony-informative characters left for subsequent phylogenetic analyses, and while the *O*/*E* ratio and *C* factor values of the classes of identity remaining in the datasets (0–40% class of ssrDNA, 0–60% class of lsrDNA) indicate that homoplasy across these classes of identity does not differ substantially and saturation does not play a major role, we used those data for phylogenetic reconstruction.

Compared with the rDNA data, the mtDNA genes appear to be highly saturated as evaluated by the entropy-based index of substitution saturation values of Xia et al. (2003). Third codon positions of both mt genes and first codon positions of *nad*3 showed substantial saturation (illustrated by their I_{SS} values higher than the $I_{SS.C}$ values; see Supplementary Table S3), making them useless for phylogenetic reconstruction. The severity of the third codon positions' saturation is further underlined by their *O*/*E* ratio values, which are significantly smaller than one, indicating a high degree of homoplasy compared with the complete datasets. Moreover, according to this method, first codon positions of *cox*1 and second codon positions of *nad*3 could only have some phylogenetic value if the actual pattern of cladogenesis tended to be symmetrical rather than extremely asymmetrical (Xia et al., 2003). While extremely asymmetrical cladogenesis is generally considered unlikely (Xia and Lemey, 2009) and while phylogenetic analyses of mt amino acid datasets revealed rather symmetrical topologies

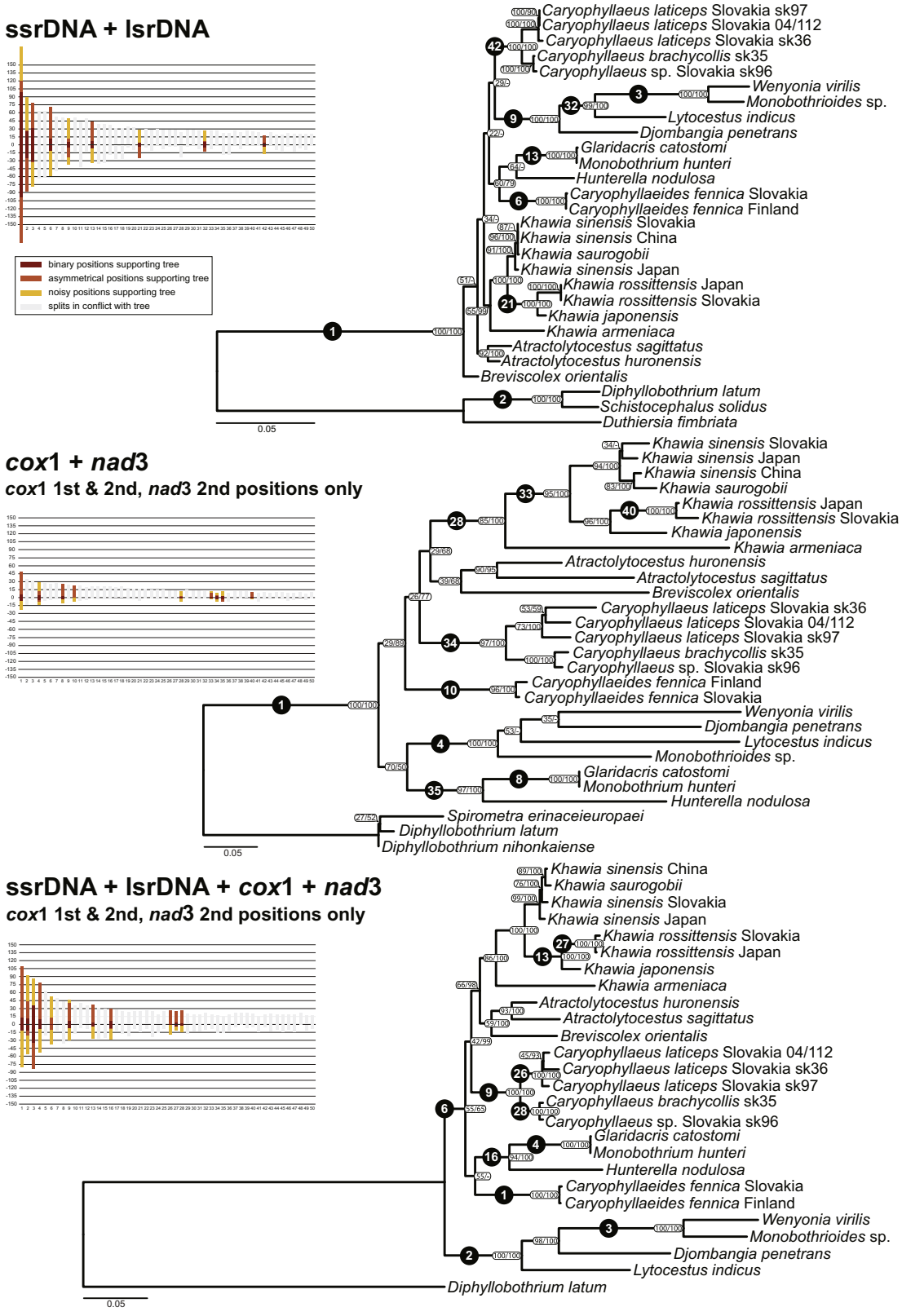

Fig. 2. Phylograms resulting from concatenated ribosomal, mitochondrial and all four genes combined analyses by maximum likelihood (ML) with bootstrap and Bayesian posterior probability nodal support values indicated at the branches (right) and the split-support spectra for the respective datasets estimated with Splits Analyses Methods program (left). Column height of a split represents the number of supporting positions counted for each partition of a split separately (in- and outgroup of a split shown above and below the horizontal axis, respectively): binary positions (two character states only); asymmetrical positions (one partition with one character state, the other with more than one state); noisy positions (more than one state in each partition, but the majority state within a group can still be identified). Compatible splits are depicted on the corresponding ML tree with their number. Results of ML and split-decomposition analyses of individual genes are provided in Supplementary Figs. S2–S4.

(Supplementary Figs. S2 and S3), we included the first codon positions of *cox*1 and the second codon positions of *nad*3 in our phylogenetic analyses.

3.4. Phylogeny reconstruction and nodal support

Phylogenetic analyses were conducted on each of the four individual genes and on the concatenated datasets (see Supplementary Figs. S2–S4 for phylogenies based on individual genes and Fig. 2 for phylogenies based on concatenated mt and rDNA data). While the resulting topologies varied amongst individual datasets and tree reconstruction methods used, we found strong agreement in the distribution of nodal support values and branch stabilities across the tree. In other words, the phylogenetic estimates based on mt and nuclear ribosomal genes differed topologically, but these differences simply reflect nodes with poor support, and thus the well-supported splits were not found to be in conflict among the different estimates. For example, all data partitions consistently recovered generic-level clades with relatively high nodal support. Similarly, none of the genes proved informative for resolving deeper phylogenetic splits, and thus none was decisive in estimating the interrelationships of genera or families. This situation is similar to that seen in analyses of members of three families of cyclophyllidean tapeworms by Littlewood et al. (2008), and thus suggests that saturation levels in these genes is similar between the earliest and latest-branching tapeworm lineages. In contrast to their conclusions, however, concatenating our data (ssrDNA + lsrDNA, *nad*3 + *cox*1, all genes together; Fig. 2) did gradually increase nodal support values, albeit we never observed significant nodal support differences when comparing the mt and rDNA data (i.e. nodes found to be weakly supported in one dataset were also weak in the other one), and we similarly never obtained strong support for basal nodes.

The lack of distinct signal in our datasets preventing resolution of caryophyllidean inter-generic relationships and basal nodes is further illustrated by the results of SDA. SDA represents one of few tree-independent methods to test the information content of alignments and to search for conflicting nucleotide patterns—the background noise (Wägele and Mayer, 2007; Wägele et al., 2009). Split support spectra on Supplementary Figs. S2–S4 (individual datasets) and Fig. 2 (combined datasets) outline the 50 most highly supported splits in each of the datasets and depict those actually present on the best tree on the corresponding phylogram. Split spectra on Supplementary Figs. S2 and S3 indicate that there is considerable conflict in the mt gene alignments. Both mt nucleotide datasets contain contradicting signal equally supporting mutually incompatible splits (represented by the columns of comparable height) and the situation does not improve when saturated third codon positions are removed from the alignments. Even if the more conservative amino acid alignments are considered, SDA reveals no authentic signal distinguishable from the background. Split support spectra of rDNA (Supplementary Fig. S4) data show slightly better signal conservation, however, it must be mentioned that the first two splits are of little significance as they relate to the outgroup taxa. Concatenation of the data gradually improves their hierarchical structure, indicating that the genuine historical signal may overcome background noise when more sequence data per specimen are collected. It is not surprising that among the 50 best SAMS splits, those compatible with the corresponding phylogenetic estimates generally show the strongest nodal support. However, many of the clades present on the ML trees never appear among the 50 best splits. Whereas split support spectra are meant to detect only the conserved patterns in the alignment relative to the background noise, the absence of splits representing some groupings does not necessarily mean that those clades are artificial. For example, closely related clades will be typically united by a small number of synapomorphies and thus fail to be recovered in the best splits (Wägele and Mayer, 2007). Nevertheless, SDA represents a useful tool for analysing phylogenetic information content in an alignment and helps reveal competing hypotheses not found in otherwise optimal tree topologies.

Mitochondrial genes are rarely utilised for phylogenetic studies of parasitic flatworms, particularly compared with nuclear ribosomal genes (see Olson and Tkach, 2005). Although mt data have proved informative for intrageneric studies, such as those on *Taenia*, *Echinococcus* and *Diphyllobothrium* (e.g. Nakao et al., 2010; Wicht et al., 2010), and have added to our understanding of interrelationships in selected groups (e.g. Hardman and Hardman, 2006; Lavikainen et al., 2008; Haukisalmi et al., 2009), for the most part they show too little conservation to have much utility in phylogenetic studies of platyhelminths. For lower level studies of parasitic flatworms (i.e. intrageneric, intrafamilial), ITS rDNA has been widely employed (particularly in Digenea; see Nolan and Cribb, 2005), albeit in cestodes intra-individual ITS sequence variation has been detected frequently: e.g. in the cyclophyllideans *Taenia solium* and *Echinococcus* sp. (van Herwerden et al., 2000b; Hancock et al., 2001), the diphyllobothriidean *Ligula intestinalis* (Olson et al., 2002; Bouzid et al., 2008) and several caryophyllidean species (for review see Bazsalovicsová et al., 2012). Analogous to the multiple mt haplotypes described in this paper, the presence of multiple ITS variants calls not just for the need of cloning and characterisation of multiple clones, but may also bias correct phylogenetic inference by detecting only a fraction of the ITS population actually present within an individual (Králová-Hromadová et al., 2010).

To date, the most taxon-rich evaluation of caryophyllidean phylogeny based on molecular data has been that of Olson et al. (2008), and although not identical in terms of the taxa sequenced, their results are highly congruent with our concatenated phylograms (Fig. 2). These molecular studies, together with the comprehensive morphological analysis of the Caryophyllidea by Oros et al. (2008), illustrate the conflict between phylogenetically-derived estimates of their evolution and traditional classification schemes that divide the group according to the position of inner longitudinal muscles relative to the reproductive organs (e.g. Mackiewicz, 1994). Using our consensus topology (Fig. 2, bottom), we searched the morphological matrix of Oros et al. (2008) for synapomorphies congruent with our molecular-based estimates and found only two characters (10 and 17; see Oros et al., 2008): the presence of scolex grooves uniting *Wenyonia* and *Monobothrioides*, and the presence of an external seminal vesicle uniting *Glaridacris*, *Hunterella* and *Monobothrium*. Although we do not have a fully resolved molecular phylogeny with which to be able to speculate on the validity of the morphological characters, the current insight foresees the need for re-evaluation of the utility of these characters in systematic studies and underlines the difficulty in evaluating this unusual and early-branching tapeworm lineage.

Acknowledgements

We thank to the following persons who kindly provided the specimens of caryophyllidean cestodes for this study: Vladimíra Hanzelová, Roman Kuchta, John Mackiewicz, Kalmán Molnár, Mikuláš Oros, Larisa Poddubnaya, Takeshi Shimazu and Csaba Székely. JB is grateful to Torsten Struck for his valuable advice and comments on the detection of substitution saturation in rRNA genes and to the members of the Laboratory of Molecular Taxonomy (Institute of Parasitology, Biology Centre of the Academy of Sciences of the Czech Republic), especially its head Miroslav Oborník, for their help over the course of his PhD studies. This study was funded by the Grant Agency of the Czech Republic (Nos. 524/08/0885, 206/09/H026 and P505/12/G112), Grant Agency of

the Slovak Republic (Project VEGA 2/0014/10) and the National Science Foundation, USA (PBI award Nos. 0818696 and 0818823).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2012.01.005.

References

Bazsalovicsová, E., Králová-Hromadová, I., Štefka, J., Scholz, T., Hanzelová, V., Vávrová, S., Szemes, T., Kirk, R., 2011. 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. 109, 125–131.

Bazsalovicsová, E., Králová-Hromadová, I., Štefka, J., Scholz, T., 2012. Molecular characterization of*Atractolytocestus sagittatus* (Cestoda: Caryophyllidea), monozoic parasite of common carp, and its differentiation from the invasive species*Atractolytocestus huronensis*. Parasitol. Res., doi:10.1007/s00436-011-2673-0.

Bensasson, D., Zhang, D.-X., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evol. 16, 314–321.

Bouzid, W., Štefka, J., Hypša, V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot, G., 2008. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). Int. J. Parasitol. 38, 1465–1479.

Brabec, J., Kuchta, R., Scholz, T., 2006. Paraphyly of the Pseudophyllidea (Platyhelminthes: Cestoda): circumscription of monophyletic clades based on phylogenetic analysis of ribosomal RNA. Int. J. Parasitol. 36, 1535–1541.

de Chambrier, A., Zehnder, M., Vaucher, C., Mariaux, J., 2004. The evolution of the Proteocephalidea (Platyhelminthes, Eucestoda) based on an enlarged molecular phylogeny, with comments on their uterine development. Syst. Parasitol. 57, 159–171.

Hancock, K., Broughel, D.E., Moura, I.N.S., Khan, A., Pieniazek, N.J., Gonzalez, A.E., Garcia, H.H., Gilman, R.H., Tsang, V.C.W., 2001. Sequence variation in the cytochrome oxidase I, internal transcribed spacer 1, and Ts14 diagnostic antigen sequences of *Taenia solium* isolates from South and Central America, India, and Asia. Int. J. Parasitol. 31, 1601–1607.

Hardman, M., Hardman, L.M., 2006. Comparison of the phylogenetic performance of neodermatan mitochondrial protein-coding genes. Zool. Scr. 35, 655–665.

Haukisalmi, V., Hardman, L.M., Henttonen, H., Laakkonen, J., Niemimaa, J., Hardman, M., Gubanyi, A., 2009. Molecular systematics and morphometrics of *Anoplocephaloides dentata* (Cestoda, Anoplocephalidae) and related species in voles and lemmings. Zool. Scr. 38, 199–220.

Healy, C.J., Caira, J.N., Jensen, K., Webster, B.L., Littlewood, D.T.J., 2009. Proposal for a new tapeworm order, Rhinebothriidea. Int. J. Parasitol. 39, 497–511.

Hoberg, E., Mariaux, J., Justine, J.-L., Brooks, D.R., Weekes, P.J., 1997. Phylogeny of the orders of the Eucestoda (Cercomeromorphae) based on comparative morphology: historical perspectives and a new working hypothesis. J. Parasitol. 83, 1128–1147.

Hoberg, E.P., Mariaux, J., Brooks, D.R., 2001. Phylogeny among orders of the Eucestoda (Cercomeromorphae): integrating morphology, molecules and total evidence. In: Littlewood, D.T.J., Bray, R.A. (Eds.), Interrelationships of the Platyhelminthes. Taylor & Francis, London, pp. 112–126.

Jia, W.Z., Yan, H.B., Guo, A.J., Zhu, X.Q., Wang, Y.C., Shi, W.G., Chen, H.T., Zhan, F., Zhang, S.H., Fu, B.Q., Littlewood, D.T.J., Cai, X.P., 2010. Complete mitochondrial genomes of *Taenia multiceps*, *T. hydatigena* and *T. pisiformis*: additional molecular markers for a tapeworm genus of human and animal health significance. BMC Genomics 11, 447.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT ver. 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Králová-Hromadová, I., Štefka, J., Špakulová, M., Orosová, M., Bombarová, M., Hanzelová, V., Bazsalovicsová, E., Scholz, T., 2010. Intra-individual internal transcribed spacer 1 (ITS1) and ITS2 ribosomal sequence variation linked with multiple rDNA loci: a case of triploid *Atractolytocestus huronensis*, the monozoic cestode of common carp. Int. J. Parasitol. 40, 175–181.

Lavikainen, A., Haukisalmi, V., Lehtinen, M.J., Henttonen, H., Oksanen, A., Meri, S., 2008. A phylogeny of members of the family Taeniidae based on the mitochondrial cox1 and nad1 gene data. Parasitology 135, 1457–1467.

Littlewood, D.T.J., Olson, P.D., 2001. Small subunit rDNA and the Platyhelminthes: signal, noise, conflict and compromise. In: Littlewood, D.T.J., Bray, R.A. (Eds.), Interrelationships of the Platyhelminthes. Taylor & Francis, London, pp. 262–278.

Littlewood, D.T.J., Curini-Galletti, M., Herniou, E.A., 2000. The interrelationships of Proseriata (Platyhelminthes: Seriata) flatworms tested with molecules and morphology. Mol. Phylogenet. Evol. 16, 449–466.

Littlewood, D.T.J., Waeschenbach, A., Nikolov, P.N., 2008. In search of mitochondrial markers for resolving the phylogeny of cyclophyllidean tapeworms (Platyhelminthes, Cestoda)—a test study with Davaineidae. Acta Parasitol. 53, 133–144.

Lockyer, A.E., Olson, P.D., Littlewood, D.T.J., 2003. Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Platyhelminthes: implications and a review of the cercomer theory. Biol. J. Linn. Soc. 78, 155–171.

Lupi, R., de Meo, P.D., Picardi, E., D'Antonio, M., Paoletti, D., Castrignanò, T., Pesole, G., Gissi, C., 2010. MitoZoa: a curated mitochondrial genome database of metazoans for comparative genomics studies. Mitochondrion 10, 192–199.

Mackiewicz, J.S., 1972. Caryophyllidea (Cestoidea): a review. Exp. Parasitol. 31, 417–512.

Mackiewicz, J.S., 1994. Order Caryophyllidea van Beneden in Carus, 1863. In: Khalil, L.F., Jones, A., Bray, R.A. (Eds.), Keys to the Cestode Parasites of Vertebrates. CAB International, Wallingford, pp. 21–43.

Mackiewicz, J.S., 2003. Caryophyllidea (Cestoidea): molecules, morphology and evolution. Acta Parasitol. 48, 143–154.

Nakao, M., Sako, Y., Yokoyama, N., Fukunaga, M., Ito, A., 2000. Mitochondrial genetic code in cestodes. Mol. Biochem. Parasitol. 111, 415–424.

Nakao, M., Yanagida, T., Okamoto, M., Knapp, J., Nkouawa, A., Sako, Y., Ito, A., 2010. State-of-the-art *Echinococcus* and *Taenia*: phylogenetic taxonomy of human-pathogenic tapeworms and its application to molecular diagnosis. Infect. Genet. Evol. 10, 444–452.

Nolan, M.J., Cribb, T.H., 2005. The use and implications of ribosomal DNA sequencing for the discrimination of digenean species. Adv. Parasitol. 60, 101–163.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.

Olson, P.D., Littlewood, D.T.J., Bray, R.A., Mariaux, J., 2001. Interrelationships and evolution of the tapeworms (Platyhelminthes: Cestoda). Mol. Phylogenet. Evol. 19, 443–467.

Olson, P.D., Littlewood, D.T.J., Griffiths, D., Kennedy, C.R., Arme, C., 2002. Evidence for the co-existence of separate strains of species of *Ligula* in Lough Neagh, Northern Ireland. J. Helminthol. 76, 171–174.

Olson, P.D., Cribb, T.H., Tkach, V.V., Bray, R.A., Littlewood, D.T.J., 2003. Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). Int. J. Parasitol. 33, 733–755.

Olson, P.D., Tkach, V.V., 2005. Advances and trends in the molecular systematics of the parasitic Platyhelminthes. Adv. Parasitol. 60, 165–243.

Olson, P.D., Poddubnaya, L.G., Littlewood, D.T.J., Scholz, T., 2008. On the position of *Archigetes* and its bearing on the early evolution of the tapeworms. J. Parasitol. 94, 898–904.

Olson, P.D., Caira, J.N., Jensen, K., Overstreet, R.M., Palm, H.W., Beveridge, I., 2010. Evolution of the trypanorhynch tapeworms: parasite phylogeny supports independent lineages of sharks and rays. Int. J. Parasitol. 40, 223–242.

Oros, M., Hanzelová, V., Scholz, T., Mackiewicz, J.S., 2008. Phylogenetic relationships of the monozoic tapeworms (Eucestoda: Caryophyllidea) inferred from morphological characters. Syst. Parasitol. 70, 1–14.

Oros, M., Scholz, T., Hanzelová, V., Mackiewicz, J.S., 2010. Scolex morphology of monozoic cestodes (Caryophyllidea) from the Palaearctic Region: a useful tool for species identification. Folia Parasitol. 57, 37–46.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rota-Stabelli, O., Yang, Z., Telford, M.J., 2009. MtZoa: a general mitochondrial amino acid substitutions model for animal evolutionary studies. Mol. Phylogenet. Evol. 52, 268–272.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Scholz, T., Brabec, J., Králová-Hromadová, I., Oros, M., Bazsalovicsová, E., Ermolenko, A., Hanzelová, V., 2011. Revision of *Khawia* spp. (Cestoda: Caryophyllidea), parasites of cyprinid fish, including a key to their identification and molecular phylogeny. Folia Parasitol. 58, 197–223.

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. USA 105, 13486–13491.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771.

Struck, T.H., Nesnidal, M.P., Purschke, G., Halanych, K.M., 2008. Detecting possibly saturated positions in 18S and 28S sequences and their influence on phylogenetic reconstruction of Annelida (Lophotrochozoa). Mol. Phylogenet. Evol. 48, 628–645.

Telford, M.J., Herniou, E.A., Russell, R.B., Littlewood, D.T.J., 2000. Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. Proc. Nat. Acad. Sci. USA 97, 11359–11364.

van Herwerden, L., Blair, D., Agatsuma, T., 2000a. Multiple lineages of the mitochondrial gene NADH dehydrogenase subunit 1 (ND1) in parasitic helminths: implications for molecular evolutionary studies of facultatively anaerobic eukaryotes. J. Mol. Evol. 51, 339–352.

van Herwerden, L., Gasser, R.B., Blair, D., 2000b. ITS-1 ribosomal DNA sequence variants are maintained in different species and strains of *Echinococcus*. Int. J. Parasitol. 30, 157–169.

Waeschenbach, A., Webster, B.L., Bray, R.A., Littlewood, D.T.J., 2007. Added resolution among ordinal level relationships of tapeworms (Platyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. Mol. Phylogenet. Evol. 45, 311–325.

Wägele, J.-W., Mayer, C., 2007. Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. BMC Evol. Biol. 7, 147.

Wägele, J.W., Letsch, H., Klussmann-Kolb, A., Mayer, C., Misof, B., Wägele, H., 2009. Phylogenetic support values are not necessarily informative: the case of the Serialia hypothesis (a mollusk phylogeny). Front. Zool. 6, 12.

Werle, E., Schneider, C., Renner, M., Völker, M., Fiehn, W., 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res. 22, 4354–4355.

White, D.J., Wolff, J.N., Pierson, M., Gemmell, N.J., 2008. Revealing the hidden complexities of mtDNA inheritance. Mol. Ecol. 17, 4925–4942.

Wicht, B., Ruggeri-Bernardi, N., Yanagida, T., Nakao, M., Peduzzi, R., Ito, A., 2010. Inter- and intra-specific characterization of tapeworms of the genus *Diphyllobothrium* (Cestoda: Diphyllobothriidea) from Switzerland, using nuclear and mitochondrial DNA targets. Parasitol. Int. 59, 35–39.

Xia, X., Xie, Z., 2001. DAMBE: data analysis in molecular biology and evolution. J. Hered. 92, 371–373.

Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. Mol. Phylogenet. Evol. 26, 1–7.

Xia, X., Lemey, P., 2009. Assessing substitution saturation with DAMBE. In: Lemey, P. (Ed.), The Phylogenetic Handbook. Cambridge University Press, Cambridge, pp. 611–626.