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.

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).