

The nuclear 18S ribosomal RNA gene as a source of phylogenetic information in the genus *Taenia*

Hongbin Yan · Zhongzi Lou · Li Li · Xingwei Ni ·
Aijiang Guo · Hongmin Li · Yadong Zheng ·
Viktor Dyachenko · Wanzhong Jia

Received: 8 March 2012 / Accepted: 9 November 2012 / Published online: 25 November 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Most species of the genus *Taenia* are of considerable medical and veterinary significance. In this study, complete nuclear 18S rRNA gene sequences were obtained from seven members of genus *Taenia* [*Taenia multiceps*, *Taenia saginata*, *Taenia asiatica*, *Taenia solium*, *Taenia pisiformis*, *Taenia hydatigena*, and *Taenia taeniaeformis*] and a phylogeny inferred using these sequences. Most of the variable sites fall within the variable regions, V1–V5. We show that sequences from the nuclear 18S ribosomal RNA gene have considerable promise as sources of phylogenetic information within the genus *Taenia*. Furthermore, given that almost all the variable sites lie within defined variable portions of that gene, it will be appropriate and economical to sequence only those regions for additional species of *Taenia*.

Introduction

Tapeworms of the genus *Taenia* (Cyclophyllidea: Taeniidae) are parasites of mammals, with carnivores or humans as definitive hosts and mostly herbivores and humans as intermediate hosts (Carabin et al. 2005). Since the 1700s, nearly 100 nominal species have been proposed or described based on either adult or larval forms (Abuladze 1964). The actual number of valid species now appears to be close to 42 (and four subspecies) (Loos-Frank 2000; Hoberg 2006). Many of these have a global socio-economic impact by causing morbidity in humans and domestic livestock (Hoberg 2002; Murell 2005). Because of its medical and veterinary significance, the genus *Taenia* has become a focus of intensive epidemiological, ecological, and taxonomic studies.

Traditionally, identification of taeniids to species has been based on morphological criteria together with ecological and biological aspects such as host specificity (Abuladze 1964). The development of molecular genetic techniques has provided alternative tools for the characterization of taeniid species and for the investigation of their relationships.

The most comprehensive hypotheses regarding the phylogenetic relationships within *Taenia* were presented by Hoberg et al. (2000) and Hoberg (2006). These were necessarily based on morphological and biological characteristics: of the 42 valid species, most remain genetically uncharacterized. Several molecular phylogenies that include some *Taenia* species have been proposed (e.g., Bowles and McManus 1994; Okamoto et al. 1995; de Queiroz and Alkire 1998; von Nickisch-Roseneck et al. 1999; Lavikainen et al. 2008; Jia et al. 2010; Liu et al. 2011; Knapp et al. 2011). Data derived from mitochondrial (mt) gene sequences were

H. Yan · Z. Lou · L. Li · X. Ni · A. Guo · H. Li · Y. Zheng · W. Jia
State Key Laboratory of Veterinary Etiological Biology,
Key Laboratory of Veterinary Parasitology of Gansu Province,
Key Laboratory of Veterinary Public Health of Agriculture Ministry,
Lanzhou Veterinary Research Institute,
Chinese Academy of Agricultural Sciences,
Lanzhou 730046, Gansu Province,
People's Republic of China

V. Dyachenko
Institute for Infectious Diseases and Zoonoses,
Ludwig-Maximilians-University of Munich,
Veterinaerstr. 13,
80539 Munich, Germany

W. Jia (✉)
Xujiaping 1, Yanchangbu,
Lanzhou 730046, People's Republic of China
e-mail: jwanzhong@yahoo.com.cn

used for most of these molecular studies. However, Knapp et al. (2011) used nuclear markers based on protein-coding genes. Bowles and McManus (1994), De Quieroz and Alkire (1998), and Zhang et al. (2007) used a portion of the nuclear 28S ribosomal RNA gene as well as a mitochondrial fragment. Concern has been expressed that mitochondrial data alone might yield misleading results because of the mode of inheritance of the mt genome. For example, discrepancies were noted between nuclear and mitochondrial trees for the genus *Echinococcus* (Knapp et al. 2011). In general, sequencing of organisms is time consuming or economically costly. Indeed, next-generation sequencing technologies (e.g., Illumina) permit the rapid sequencing of a large amount of sequences. However, the analysis of a huge amount of data remains after this sequencing, and this kind of apparatus is still expensive. Nuclear ribosomal genes have been found satisfactory in the past for inferring phylogenies of trematodes (reviewed in Olson and Tkach 2005; Blair 2006) and have been used in other groups of cestodes (Campos et al. 1998; Olson et al. 2001; Foronda et al. 2004; Waeschenbach et al. 2007). Here, we investigate whether full-length nuclear 18S rRNA genes are suitable for inferring phylogenies of *Taenia* species and whether these are congruent with phylogenies based on mt sequence data. We sequenced the complete nuclear 18S rRNA gene from specimens of seven *Taenia* spp., *Taenia multiceps*, *Taenia hydatigena*, *Taenia pisiformis*, *Taenia saginata*, *Taenia asiatica*, *Taenia solium*, and *Taenia taeniaeformis*. Using these sequences, we inferred phylogenetic relationships within the genus *Taenia* and compared the results with those from previously published studies that have used molecular or morphological data.

Materials and methods

Parasite specimens and DNA extraction

One specimen of *T. multiceps* was collected from a dog infected experimentally with *Coenurus cerebralis* from naturally infected sheep (Huangcheng Sheep Breeding Company, Gansu Province) in our laboratory. One specimen of *T. pisiformis* cyst was collected from the liver of a white rabbit at a slaughterhouse in Linyi City, Shandong Province. Specimens of *T. asiatica*, *T. solium*, and *T. saginata* were collected from fecal samples of patients with taeniasis after anthelmintic treatment (Dali, Yunnan Province). These species were identified primarily based on morphological traits, including the number, size, and shape of the rostellar hooks, according to Verster (1969), Loos-Frank (2000), and Huang and Sheng (2006). A fragment from each tapeworm and the protoscolex from the cyst of *T. pisiformis* were washed with cold phosphate-buffered saline and frozen in liquid nitrogen. Genomic DNA was extracted using the Genomic DNA

Purification Kit (Qiagen, Germany) according to the manufacturer's instructions and stored at -20°C . The genomic DNA of *T. taeniaeformis* was from Dr. Dyachenko's lab (Institute for Infectious Diseases and Zoonoses, Ludwig-Maximilians-University of Munich, Germany).

Amplification of 18S rDNAs

A pair of primers "full-length Cestoda 18S" (FLC18SF) (5'-tcc tgc cag tag tca tat gc-3') and FLC18SR (5'-ctt gtt acg act ttt act tcc tct-3') was designed for the amplification of taeniid 18S rDNAs based on the complete 18S rRNA gene sequences of cyclophyllid tapeworms published in GenBank (Olson et al. 2003; Foronda et al. 2004; Waeschenbach et al. 2007). Polymerase chain reactions were performed in a 50- μl volume of the following reaction mixture: 10 mM Tris-HCl (pH8.4), 50 mM KCl, 3 mM MgCl_2 , 250 mM of each dNTP, 25 pmol of each primer, and 2 U *Taq* polymerase (TaKaRa, China) under the conditions: 94 $^{\circ}\text{C}$, 4 min, then 30 cycles of 94 $^{\circ}\text{C}$, 40 s; 52 $^{\circ}\text{C}$, 30 s; 72 $^{\circ}\text{C}$, 2 min and 20 s, followed by a final step at 72 $^{\circ}\text{C}$ for 10 min. The amplicons were directly sequenced or cloned into the pGEM-T Easy vector (Promega, USA) for sequencing. Sequencing was performed in Sangon Biotech (Sangon, China) Co., Ltd. with primer walking.

Sequence alignment and phylogenetic analysis

The complete 18S rDNA sequences were aligned using ClustalW within MEGA 5.0 (Tamura et al. 2011). Settings used were: gap opening penalty=15, gap extension penalty=6.66, and delay divergent cutoff=30 %. The resulting alignment was manually checked (Thompson et al. 1994). Percent pairwise divergences of nucleotides were calculated using MEGA 5.0. The 18S rDNA sequences of seven *Taenia* species were used for phylogenetic analysis. The 18S rDNA sequence of *Echinococcus granulosus* (GenBank accession number GQ260092) was utilized as the outgroup. Three methods, namely maximum likelihood (ML), neighbor joining (NJ), and maximum parsimony (MP), were used for phylogenetic reconstructions. The ML, NJ, and MP methods were carried out using the Tamura 3-parameter model (Tamura 1992), Maximum Composite Likelihood model (Tamura et al. 2004), and max-mini branch-and-bound algorithm (Purdum et al. 2000), respectively, and all analyses were implemented in MEGA 5.0 (Tamura et al. 2011). The consensus tree was obtained after bootstrap analysis, with 1,000 replications for NJ, MP, and ML trees, with values above 50 % reported.

Ethics statement

All experiments using rabbits, sheep, and dogs were undertaken under strict Chinese experimental animal permission

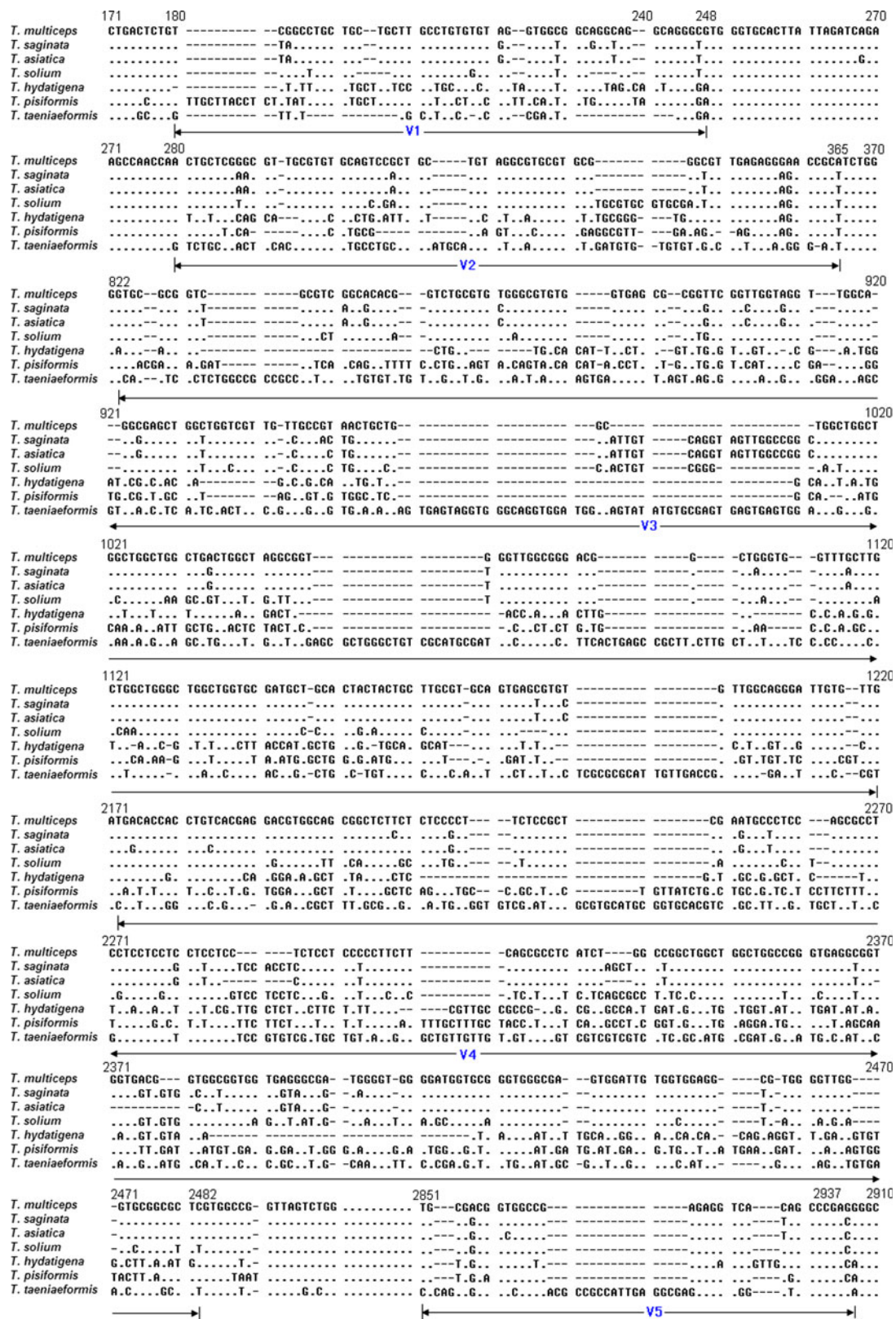


Fig. 1 Alignment of the most variable portions of the 18S RNA gene sequences of *T. multiceps*, *T. saginata*, *T. asiatica*, *T. solium*, *T. pisiformis*, *T. hydatigena*, and *T. taeniaeformis*. Dots indicate nucleotide

similarity among the seven taxa and *hyphens*, alignment gaps. The numbers refer to alignment positions

clearances, and animals at all times were treated in accordance with animal ethical procedures and guidelines for animal husbandry of the Institutional Ethics Committee of Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences.

Results and discussion

The 18S ribosomal RNA gene sequences of *T. multiceps*, *T. saginata*, *T. asiatica*, *T. solium*, *T. pisiformis*, *T. hydatigena*, and *T. taeniaeformis* were 2,574, 2,608, 2,583, 2,599, 2,642, 2,531, and 2,891 bp in size, respectively. The alignment of these seven sequences consisted of 2,910 sites (Fig. 1). Five variable regions could be recognized in the 18S gene sequences of all seven taxa (Wuyts et al. 2001). They are V1 (alignment positions 180–248), V2 (280–365), V3 (822–1,220), V4 (2,172–2,482), and V5 (2,851–2,907), respectively. Among them, V1 is quite short, whereas V3 and V4 are the longest. V1 and V4 are the most variable (Fig. 1). Almost all differences between sequences (insertions/deletions and base substitutions) were found within the variable regions: the conserved regions are therefore not shown in Fig. 1. The G+C content of sequences ranged from 52.67 to 55.79 % (*T. hydatigena* at one extreme but *T. saginata* at the other). All nucleotide sequences in this study have been deposited in the GenBank database under the accession numbers GQ260088–GQ260091 for *T. asiatica*, *T. multiceps*, *T. hydatigena*, and *T. solium* and JQ609338–JQ609340 for *T. saginata*, *T. pisiformis*, and *T. taeniaeformis*, respectively.

Pairwise distances in these 18S rDNA sequences among seven *Taenia* spp. were 0.78–14.96 %, using the Maximum Composite Likelihood model (Tamura et al. 2004). The 18S rDNAs of *T. saginata* and *T. asiatica* were highly similar to each other (with 99.2 % identity). These two species also share high similarity in all molecular studies using mitochondrial or nuclear DNA sequences (e.g., Bowles and McManus 1994; de Queiroz and Alkire 1998; Lavikainen et al. 2008; Jia et al. 2010; Knapp et al. 2011; Liu et al. 2011), reflecting their well-known sister relationship.

Elsewhere in the tree (Fig. 2), there is strong concordance with previous studies that used nuclear or mitochondrial genes. *T. taeniaeformis* lies more or less basal within *Taenia* in most studies that include it. Exceptions to this are the mt *cox1* tree in De Queiroz and Alkire (1998) and studies that include *Taenia mustelae*, which is sometimes placed basal in *Taenia* or as sister to *Echinococcus* (de Queiroz and Alkire 1998; Lavikainen et al. 2008; Knapp et al. 2011). Where the following species have been included, the relationship (((*T. asiatica*, *T. saginata*), *T. multiceps*), *T. solium*) is always seen. The clade of *T. pisiformis* and *T. hydatigena* lies between the cluster comprising *T. asiatica* and other three related species and *T. taeniaeformis*. The placement of these two relative to each other varies between studies. Here, they are sisters (Fig. 2). Jia et al. (2010) also found this relationship based on complete mt genomes. Partial mt genome sequences (Lavikainen et al. 2008; Jia et al. 2010) placed *T. pisiformis* as sister to *T. solium*, a relationship not found in any other studies.

Phylogenies based on morphology (e.g., Fig. 2 in Hoberg et al. 2000) place *T. saginata* and *T. asiatica* as sisters. Their morphological tree contains 30 species, making comparisons difficult with molecular trees that include a much sparse sampling of taxa. However, the molecular grouping of *T. multiceps* and *T. solium* in a clade with *T. saginata* and *T. asiatica* is more or less congruent with the morphological tree. *T. taeniaeformis* lies near the base of the morphological tree, as it does in molecular trees. Morphology places *T. pisiformis* and *T. hydatigena* well apart, but within the larger clade that contains *T. asiatica*, *T. saginata*, *T. multiceps*, and *T. solium*. Molecular data from additional species of *Taenia* are required to be able to fully evaluate the relative merits of morphological and molecular data. Here, we have shown that sequences from the nuclear 18S ribosomal RNA gene have considerable promise as sources of phylogenetic information within the genus *Taenia*. Furthermore, given that almost all the variable sites lie within defined variable portions of that gene, it will be appropriate and economical to sequence only those regions for additional species of *Taenia*.

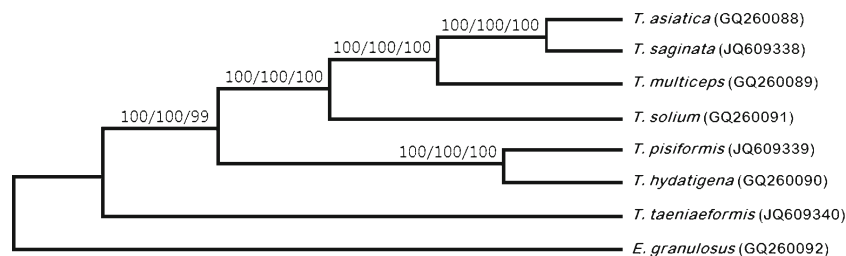


Fig. 2 Phylogenetic tree of 18S rDNA sequences of *Taenia* species inferred by the methods of maximum likelihood [using the Tamura 3-parameter model (Tamura 1992)], neighbor joining [using Maximum Composite Likelihood models (Tamura et al. 2004)], and maximum

parsimony [using max–mini branch-and-bound algorithm (Purdom et al. 2000)] with bootstrapping and 1,000 replications. Numbers at nodes represent the percentage of bootstrap replications (values >50 % reported)

Acknowledgments This study was supported by the Special Fund for Agro-scientific Research in the Public Interest (200903036-07), the People's Republic of China. The authors also thank Professor David Blair for his comments and modifications.

References

- Abuladze KI (1964) Essentials of cestodology. In: Skrjabin KI (ed) *Taeniata of animals and man and diseases caused by them*, vol IV. Nauka, Moscow, English translation (1970), Israel Program of Scientific Translations, Jerusalem
- Blair D (2006) Ribosomal DNA variation in parasitic flatworms. In: Maule A (ed) *Parasitic flatworms: molecular biology, biochemistry, immunology and control*. CABI, London, pp 96–123
- Bowles J, McManus DP (1994) Genetic characterization of Asian *Taenia*, a newly described taeniid cestode of humans. *AmJTrop Med Hyg* 50:33–44
- Campos A, Cummings MP, Reyes JL, Laclette JP (1998) Phylogenetic relationships of plathyhelminthes based on 18S ribosomal gene sequences. *Mol Phylogenet Evol* 10:1–10
- Carabin H, Budke C, Cowan L, Willingham A, Torgerson P (2005) Methods for assessing the burden of parasitic zoonoses: echinococcosis and cysticercosis. *Trends Parasitol* 21:327–333
- de Queiroz A, Alkire NL (1998) The phylogenetic placement of *Taenia* cestodes that parasitize humans. *J Parasitol* 84:379–383
- Foronda P, Casanova JC, Valladares B, Martinez E, Feliu C (2004) Molecular systematics of several cyclophyllid families (Cestoda) based on the analysis of 18S ribosomal DNA gene sequences. *Parasitol Res* 93(4):279–282
- Hoberg EP (2002) *Taenia* tapeworms: their biology, evolution and socioeconomic significance. *Microbes Infect* 4:859–866
- Hoberg EP (2006) Phylogeny of *Taenia*: species definitions and origins of human parasites. *Parasitol Int* 55:S23–S30
- Hoberg EP, Jones A, Rausch RL, Eom KS, Gardner SL (2000) A phylogenetic hypothesis for species of the genus *Taenia* (Eucestoda: Taeniidae). *J Parasitol* 86:89–98
- Huang B, Sheng J (2006) Chinese atlas of parasites for livestock and poultry in China. China Agricultural Science and Technology Press, Beijing, pp 199–205, In Chinese
- Jia WZ, Yan HB, Guo AJ, Zhu XQ, Wang YC, Shi WG, Chen HT, Zhan F, Zhang SH, Fu BQ, Littlewood DTJ, Cai XP (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
- Knapp J, Nakao M, Yanagida T, Okamoto M, Saarma U, Lavikainen A, Ito A (2011) Phylogenetic relationships within *Echinococcus* and *Taenia* tapeworms (Cestoda: Taeniidae): an inference from nuclear protein-coding genes. *Mol Phylogenet Evol* 61:628–638
- Lavikainen A, Haukisalme V, Lehtinen MJ, 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
- Liu GH, Lin RQ, Li WM, Liu W, Liu Y, Yuan ZG, Song HQ, Zhao GH, Zhang KX, Zhu XQ (2011) The complete mitochondrial genomes of three cestode species of *Taenia* infecting animals and humans. *Mol Biol Rep* 38:2249–2256
- Loos-Frank B (2000) An up-date of Verster's (1969) 'Taxonomic revision of the genus *Taenia* Linnaeus' (Cestoda) in table format. *Sys Parasitol* 45:155–183
- Murell KD (2005) Epidemiology of taeniosis and cysticercosis. In: Murell KD (ed) *WHO/FAO/OIE guidelines for the surveillance, prevention and control of taeniosis/cysticercosis*. OIE, Paris, pp 27–44
- Okamoto M, Bessho Y, Kamiya M, Kurosawa T, Horii T (1995) Phylogenetic relationships within *Taenia taeniaeformis* variants and other taeniid cestodes inferred from the nucleotide sequence of the cytochrome oxidase subunit I gene. *Parasitol Res* 81:415–458
- Olson PD, Tkach VV (2005) Advances and trends in the molecular systematics of the parasitic plathyhelminthes. *Adv Parasitol* 60:165–243
- Olson PD, Littlewood DT, Bray RA, Mariaux J (2001) Interrelationships and evolution of the tapeworms (Plathyhelminthes: Cestoda). *Mol Phylogenet Evol* 19(3):443–467
- Olson PD, Yoder K, Fajardo L-GLF, Marty AM, van de Pas S, Olivier C, Relman DA (2003) Lethal invasive cestodiasis in immunosuppressed patients. *J Infect Dis* 187:1962–1966
- Purdom PW, Bradford PG, Tamura K, Kumar S (2000) Single column discrepancy and dynamic max-mini optimizations for quickly finding the most parsimonious evolutionary trees. *Bioinformatics* 16:140–151
- Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition–transversion and G+C-content biases. *Mol Biol Evol* 9:678–687
- Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci USA* 101:11030–11035
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Verster A (1969) A taxonomic revision of the genus *Taenia* Linnaeus, 1758 s. str. *Onderstepoort J Vet* 36:3–58
- von Nickisch-Rosenegk M, Silva-Gonzalez R, Lucius R (1999) Modification of universal 12S primers for specific amplification of contaminated *Taenia* spp. (Cestoda) gDNA enabling phylogenetic studies. *Parasitol Res* 85:819–825
- Waeschenbach A, Webster BL, Bray RA, Littlewood DTJ (2007) Added resolution among ordinal level relationships of tapeworms (Plathyhelminthes: Cestoda) with complete small and large subunit nuclear ribosomal RNA genes. *Mol Phylogenet Evol* 45:311–325
- Wuyts J, van de Peer Y, Wachter RD (2001) Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. *Nucleic Acids Res* 29:5017–5028
- Zhang L, Hu M, Jones A, Allsopp BA, Beveridge I, Schindler AR, Gasser RB (2007) Characterization of *Taenia madoquae* and *Taenia regis* from carnivores in Kenya using genetic markers in nuclear and mitochondrial DNA, and their relationships with other selected taeniids. *Mol Cell Probe* 21:379–385