Available online at www.sciencedirect.com
SCIENCE DIRECT
Parasitology International 55 (2006) S31 – S37
www.elsevier.com/locate/parint

Molecular discrimination of taeniid cestodes

Donald P. McManus *

Molecular Parasitology Laboratory, Division of Infectious Diseases and Immunology, Australian Centre for International and Tropical Health and Nutrition, The Queensland Institute of Medical Research and The University of Queensland, Post Office Royal Brisbane Hospital, Queensland 4029, Australia

Available online 5 December 2005

Abstract

DNA approaches are now being used routinely for accurate identification of *Echinococcus* and *Taenia* species, subspecies and strains, and in molecular epidemiological surveys of echinococcosis/taeniasis in different geographical settings and host assemblages. The publication of the complete sequences of the mitochondrial (mt) genomes of *E. granulosus*, *E. multilocularis*, *T. solium* and Asian *Taenia*, and the availability of mtDNA sequences for a number of other taeniid genotypes, has provided additional genetic information that can be used for more in depth phylogenetic and taxonomic studies of these parasites. This very rich sequence information has provided a solid molecular basis, along with a range of different biological, epidemiological, biochemical and other molecular-genetic criteria, for revising the taxonomy of the genus *Echinococcus* and for estimating the evolutionary time of divergence of the various taxa. Furthermore, the accumulating genetic data has allowed the development of PCR-based tests for unambiguous identification of *Echinococcus* eggs in the faeces of definitive hosts and in the environment. Molecular phylogenies derived from mtDNA sequence comparisons of geographically distributed samples of *T. solium* provide molecular evidence for two genotypes, one being restricted to Asia, with the other occurring in Africa and America. Whether the two genetic forms of *T. solium* differ in important phenotypic characteristics remains to be determined. As well, minor DNA sequence differences have been reported between isolates of *T. saginata* and Asian *Taenia*. There has been considerable discussion over a number of years regarding the taxonomic position of Asian *Taenia* and whether it should be regarded as a genotype, strain, subspecies or sister species of *T. saginata*. The available molecular genetic data do not support independent species status for Asian *Taenia* and *T. saginata*. What is in agreement is that both taxa are closely related to each other but distantly related to *T. solium*. This is important in public health terms as it predicts that cysticercosis in humans attributable to Asian *Taenia* does not occur, because cysticercosis is unknown in *T. saginata*.

Keywords: *Echinococcus*; *Taenia*; Mitochondrial genome; *cox1* gene; *nad1* gene; DNA identification; PCR-based tests; Genotype; Strain; Asian *Taenia*; *Taenia saginata*; *Taenia solium*

1. Introduction

The most accurate ways to identify *Taenia* taxa infecting humans are based on molecular and immunological diagnostic approaches combined with comparative morphology [1]. Molecular techniques are proving of value in the study of both *Echinococcus* and *Taenia*, in particular for investigating genetic variation, phylogenetic relationships and molecular epidemiology, for the detection of parasite nucleic acids in clinical samples and in the identification of taeniid eggs. Indeed, detecting minute amounts of taeniid cestode DNA and mRNA, not only to identify but also to characterize the biological status of the parasite material, is becoming recognized as a powerful complementary method to conventional identification methods.

A number of techniques have been employed for the DNA identification of *Echinococcus* and *Taenia* (reviewed in [1–8]) These approaches include restriction fragment length polymorphism (RFLP) analysis, RFLP-polymerase chain reaction (PCR) or PCR-linked RFLP analysis, direct comparison of PCR-amplified DNA sequences, random amplified polymorphic DNA-PCR (RAPD-PCR), single-strand conformation polymorphism (SSCP) analysis and microsatellite analysis. Some of the more recent studies undertaken on *Echinococcus* and *Taenia*, where some of these techniques have been applied, are described later.

2. The mitochondrial genome of taeniid cestodes

DNA identification of taeniid cestodes has used specific sequences from both the nuclear and mitochondrial genomes. The mitochondrial genome is located in mitochondria and the

* Tel.: +61 7 3362 0401; fax: +61 7 3362 0104.
E-mail address: donM@qimr.edu.au.

1383-5769/$ - see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.parint.2005.11.004

Table 1
Complete mitochondrial genomes of taeniid cestodes

Species	Size (base pairs)	GenBank accession number	Reference
Echinococcus granulosus (G1 genotype; sheep–dog strain)	13,588	AF297617	[9]
Echinococcus granulosus (G4 genotype; horse–dog strain)	13,598	AF346403	[9]
E. multilocularis	13,788	AB018440	[10]
Taenia crassiceps	13,503	AF216699	[11]
Taenia solium	13,709	AB086256	[12]
Taenia saginata asiatica	13,703	AF445798	[13]
Taenia saginata	ca. 13,720	AF031285, B066495, AB066581, AY195858	K.S. Eom (pers. comm.)

much larger nuclear genome distributed on the chromosomes in the nucleus. Here emphasis is placed on the utility of mitochondrial DNA (mtDNA) sequences for molecular discrimination of taeniid taxa. Mitochondrial genomes are small (generally less than 20,000 bp in metazoans), multi-copied in the cell and circular. Mitochondrial DNA is useful for the discrimination of closely related organisms because of its relatively rapid rate of evolution. Furthermore, as mtDNA is haploid, allele haplotypes can be determined unambiguously. Mitochondrial DNA has the additional advantage that, as far as is known, it is maternally inherited and does not recombine, thus simplifying analysis. The advent of the polymerase chain reaction (PCR) has provided a highly sensitive approach that is now widely used to target mtDNA sequences for *Echinococcus* and *Taenia* identification purposes, including discrimination of eggs.

Complete (Table 1) or near complete mtDNA sequences are available for a number of taeniid cestodes and these provide a very rich resource of genetic information for comparative mitochondrial genomics, systematic studies and molecular identification of unknown isolates. The mt genomes of *Echinococcus* and *Taenia* are not strikingly different from those of other metazoans and the gene order is similar for all taeniid taxa. There are 12 protein-coding genes, which fall into the following categories: nicotinamide dehydrogenase complex (*nad1-6* and *nad4L* subunits), cytochrome *c* oxidase complex (*cox1-3* subunits), cytochrome *b* (*cob*) and adenosine triphosphatase subunit 6 (*atp6*). Unlike the situation in most other metazoans, there is no *atp8*. Two genes encoding ribosomal RNA subunits are present: the large subunit (*rrnL* or 16S) and small subunit (*rrnS* or 12S). As is common in mitochondrial genomes, there are 22 tRNA genes. All genes are transcribed in the same direction. Genes lack introns and usually abut one another or are separated by only a few nucleotides. However, some genes overlap, most notably *nad4* and *nad4L*. There are two relatively long non-coding regions (NRs) that may be associated with replication of the mt genome.

In addition to the short non-coding regions (usually <30 bp in length) separating genes, there are one or two longer non-coding region(s) (NRs) in every parasitic flatworm mt genome. It is noteworthy that the mt genomes of cestodes include the shortest known to date among metazoans due to the very short NRs present separated by *nad5* and several tRNA genes.

3. Genetic variation and the case for a revised taxonomy of *Echinococcus*

The availability of the large amount of new DNA data underpins the argument for re-appraisal of the taxonomy of *Echinococcus*, particularly *E. granulosus*. It is now well recognized that *E. granulosus* comprises a number of genetic variants and, to date, analyses of mtDNA sequences have identified 10 distinct genetic types (genotypes G1–10) [5]). This categorization follows very closely the pattern of strain variation emerging based on biological characteristics. The extensive variation in nominal *E. granulosus* may influence life cycle patterns, host specificity, development rate, antigenicity, transmission dynamics, sensitivity to chemotherapeutic agents and pathology with important implications for the design and development of vaccines, diagnostic reagents and drugs. The DNA analysis has shown that the majority of the *E. granulosus* genotypes identified to date (G1, G2, G5, G6, G7, G8 and G9) are infective to humans (Table 2).

Mitochondrial DNA sequence comparisons, along with a range of different biological, epidemiological, biochemical and other molecular-genetic criteria now provide a very strong argument in favour of separate species status (*E. equinus*) for the horse–dog (G4) strain of *E. granulosus* [4]. Along similar lines, the recently acquired molecular data have provided a solid base for the proposal that the cattle–dog strain (G5) of *E. granulosus* should be reinstated as a distinct species, *E. ortleppi*, and recognized as the cattle-adapted form of *Echinococcus* [4]. As additional molecular data are generated, it is likely that several of the other currently recognized strains of *E. granulosus* will also be given separate species status.

Table 2
Percentage pairwise comparison of the divergence of fragments of the *cox1* (366 base pairs) and *nad1* (471 base pairs) genes for a number of *Echinococcus* genotypes

Genotype	G1	G2	G3	G4	G5	G6	G7	G8	M1	M2	V	O
G1	–	0.7	0.6	11.9	11.6	11.6	11.8	8.4	13.3	13.7	12.8	13.4
G2	0.35	–	0.4	11.8	11.2	11.1	11.4	8.6	13.1	13.4	12.7	13.0
G3	0.3	0.2	–	11.9	11.4	11.2	11.5	8.5	13.3	13.7	12.8	13.4
G4	5.95	5.9	5.95	–	9.7	10.8	11.0	10.0	11.6	11.6	10.4	11.6
G5	5.8	5.6	5.7	4.85	–	5.7	6.2	6.7	11.2	11.5	11.1	12.5
G6	5.8	5.55	5.6	5.4	2.85	–	0.5	4.8	11.9	12.2	12.3	12.8
G7	5.9	5.7	5.75	5.5	3.1	0.25	–	5.0	11.9	12.2	12.5	12.8
G8	4.2	4.2	4.25	5.0	3.35	2.4	2.5	–	11.0	11.5	11.1	12.8
M1	6.65	6.55	6.65	5.6	5.6	5.95	5.95	5.5	–	0.5	13.0	13.9
M2	6.85	6.7	6.85	5.8	5.75	6.1	6.1	5.75	0.25	–	13.0	13.9
V	6.4	6.35	6.4	5.2	5.55	6.15	6.25	5.55	6.5	6.5	–	11.7
O	6.7	6.5	6.7	5.8	6.25	6.4	6.4	6.4	6.95	6.95	5.85	–

Percentage nucleotide distance is shown above the diagonal and the time divergence based (in million years (MY)) on the 2% MY "mtDNA clock" [14] is shown below the diagonal (original data from Bowles et al. [16]).