A novel phylogeny for the genus *Echinococcus*, based on nuclear data, challenges relationships based on mitochondrial evidence

U. SAARMA[1]*, I. JÕGISALU[1,2], E. MOKS[1], A. VARCASIA[3], A. LAVIKAINEN[4], A. OKSANEN[5], S. SIMSEK[6], V. ANDRESIUK[7], G. DENEGRI[7], L. M. GONZÁLEZ[8], E. FERRER[8,9], T. GÁRATE[8], L. RINALDI[10] *and* P. MARAVILLA[11]

[1] *Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Vanemuise 46, 51014, Tartu, Estonia*
[2] *Department of Game Monitoring, Centre of Forest Protection and Silviculture, Rõõmu tee 2, 51013, Tartu, Estonia*
[3] *Dipartimento di Biologia Animale Sezione di Parassitologia e Malattie Parassitarie, Università degli Studi di Sassari, Via Vienna 2, 07100, Sassari*
[4] *Department of Bacteriology and Immunology (FINPAR), Haartman Institute, P.O. Box 21, FI-00014 University of Helsinki, Finland*
[5] *Finnish Food Safety Authority Evira, Fish and Wildlife Health Research Unit (FINPAR), P.O. Box 517, FI-90101 Oulu, Finland*
[6] *Department of Parasitology, Faculty of Veterinary Medicine, University of Firat, 23119, Elazig, Turkey*
[7] *Laboratory of Parasitic Zoonoses, Department of Biology, FCEyN, Mar del Plata National University, Funes 3350, CP 7600, Mar del Plata, Buenos Aires, Argentina (CONICET)*
[8] *Instituto de Salud Carlos III, Centro Nacional de Microbiología, Ctra. Majadahonda Pozuelo Km 2,2, 28220 Majadahonda, Madrid, Spain*
[9] *Departamento de Parasitología and Instituto de Investigaciones Biomédicas (BIOMED), Universidad de Carabobo, Maracay, Venezuela*
[10] *Department of Pathology and Animal Health, Faculty of Veterinary Medicine, University of Naples Federico II, Naples, Italy*
[11] *Direccion de Investigacion, Hospital General "Dr. Manuel Gea Gonzalez", SSA, Mexico DF, Mexico*

(*Received 2 October 2008; revised 11 November 2008; accepted 12 November 2008; first published online 21 January 2009*)

SUMMARY

The taxonomic status of *Echinococcus*, an important zoonotic cestode genus, has remained controversial, despite numerous attempts to revise it. Although mitochondrial DNA (mtDNA) has been the source of markers of choice for reconstructing the phylogeny of the genus, results derived from mtDNA have led to significant inconsistencies with earlier species classifications based on phenotypic analysis. Here, we used nuclear DNA markers to test the phylogenic relationships of members of the genus *Echinococcus*. The analysis of sequence data for 5 nuclear genes revealed a significantly different phylogeny for *Echinococcus* from that proposed on the basis of mitochondrial DNA sequence data, but was in agreement with earlier species classifications. The most notable results from the nuclear phylogeny were (1) *E. multilocularis* was placed as basal taxon, (2) all genotypes of *Echinococcus granulosus* grouped as a monophyletic entity, and (3) genotypes G8 and G10 clustered together. We conclude that the analysis of nuclear DNA data provides a more reliable means of inferring phylogenetic relationships within *Echinococcus* than mtDNA and suggest that mtDNA should not be used as the sole source of markers in future studies where the goal is to reconstruct a phylogeny that does not only reflect a maternal lineage, but aims to describe the evolutionary history at species level or higher.

Key words: *Echinococcus*, nuclear genes, mitochondrial genes, phylogeny.

INTRODUCTION

Tapeworm parasites of the genus *Echinococcus* Rudolphi, 1801 (Cestoda: Taeniidae) are causative agents of echinococcosis, which is one of the most important zoonoses worldwide. Larval stages (metacestodes) of *Echinococcus* species establish as cysts in the internal organs (mainly liver and lung) of intermediate hosts – frequently wild and domestic ungulates, but also humans – leading to severe health problems and even death. Adult parasites are found in the small intestines of wild and domestic canids and/or felids. Their effects on the health of these definitive hosts are less severe compared with those experienced by intermediate hosts (McManus and Thompson, 2003; Jenkins *et al.* 2005; Thompson, 2008). In order to understand the biology, life cycle/transmission, distribution and evolution of *Echinococcus* parasites, numerous studies have been conducted in recent times. Despite the major medical and veterinary importance of this genus, efforts to

* Corresponding author: Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Vanemuise 46, 51014, Tartu, Estonia. Tel: +37 27375099. Fax: +37 27375830. E-mail: Urmas.Saarma@ut.ee

 © 2009 Cambridge University Press
doi:10.1017/S0031182008005453 Printed in the United Kingdom

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

establish a widely-accepted taxonomy have not yet been successful. Based on existing data, 4 species have received general acceptance, namely *E. granulosus*, *E. multilocularis*, *E. vogeli* and *E. oligarthrus*. However, significant taxonomic uncertainties in the genus are yet to be resolved, particularly within *E. granulosus*. On the basis of considerable intraspecific variation, *E. granulosus* has been regarded as a collection of 10 different genotypes, which have been numbered consecutively G1–G10, reflecting the order in which they were described (Bowles *et al*. 1992, 1994; Scott *et al*. 1997; Lavikainen *et al*. 2003). It has been proposed that some of these genotypes or groups of genotypes may deserve species status: G1–G3 as *E. granulosus sensu stricto*, G4 as *E. equinus*, G5 as *E. ortleppi* and G6–G10 as *E. canadensis* (Thompson and McManus, 2002; Nakao *et al*. 2007; Thompson, 2008). Furthermore, a new species, *E. shiquicus*, has been described recently (Xiao *et al*. 2005), and it has been argued that *E. felidis* Ortlepp, 1937 should be reaccepted as a distinct species, rather than falling within *E. granulosus sensu stricto* (see Hüttner *et al*. 2008). The existence of a human-specific genotype G9 of *E. granulosus*, as proposed by Scott *et al*. (1997), remains controversial. Analysis of mitochondrial sequence data revealed no significant differences between isolates from pigs (G7) and humans (G9) (Kedra *et al*. 1999).

Although phylogenetic studies have contributed significantly toward a better understanding of the taxonomy of the genus *Echinococcus*, even the use of complete mitochondrial (mt) genome sequences (e.g. Nakao *et al*. 2007) has failed to produce results in agreement with earlier phenotypic classifications (Williams and Sweatman, 1963; Rausch, 1967). Specifically, the phylogeny of *Echinococcus* based on analyses of complete mtDNA sequence data has confirmed earlier studies using partial mtDNA sequences (Lavikainen *et al*. 2003, 2006; Xiao *et al*. 2005; Thompson *et al*. 2006). These indicated that different genotypes of *E. granulosus* do not form a monophyletic cluster, but include also *E. multilocularis* and *E. shiquicus* (see Nakao *et al*. 2007), which are both phenotypically distinct from *E. granulosus*. Moreover, it has been demonstrated that the cervid genotypes of *E. granulosus*, G8 and G10, are not positioned on the phylogenetic tree as sister taxa; instead, the camel- and pig-specific genotypes G6/G7 were shown to form a clade with G10 (Moks *et al*. 2008). Taken together, these data indicate the need for additional studies to resolve the taxonomy of *Echinococcus*.

Contemporary phylogenetic hypotheses have been based largely on data derived from mitochondrial sequences, whereas an alternative source of phylogenetic characters from the nuclear genome has not received much attention. Although mtDNA provides useful genetic markers – it is, generally, haploid, multicopy, neutral and provides access to a set of genes with no recombination (Brown, 1983) and 'rapid' evolution (Brown *et al*. 1979) – it has some significant limitations. First, mtDNA is usually inherited through the maternal lineage (Giles *et al*. 1980; Avise *et al*. 1987). This is presumably valid also for parasitic flatworms, although experiments have suggested that intraspecific paternal inheritance of mtDNA may occur across multiple generations in *Schistosoma mansoni* (Jannotti-Passos *et al*. 2001). Should female evolutionary history differ from that of the species, then mtDNA would not reflect the evolutionary history of the species as a whole. Furthermore, insertions of mitochondrial sequences into the nuclear genome and *vice versa* can occur (see Leister, 2005). As genes in the mitochondrial genome are linked (Anderson *et al*. 1981; Le *et al*. 2000), information from different mtDNA genes does not fulfil the criteria of statistical independence, which is a requirement if a phylogeny is to represent the evolutionary history of genomes and not just of single loci (Ballard and Rand, 2005). If the goal is to reconstruct a phylogeny that involves species and not just maternal lineages, the use of nuclear DNA is frequently recommended (for recent reviews dealing with the problems associated with using mtDNA data for phylogenetic analysis, see Funk and Omland, 2003; Hurst and Jiggins, 2005; Ballard and Rand, 2005). Although nuclear genes show a number of favourable characteristics, one should be aware that paralogous copies of nuclear genes can occur in the genome and may lead to erroneous phylogenies. In the present study, the aim was to employ data from a range of nuclear genes to reconstruct the phylogeny of the genus *Echinococcus* and describe relevant properties of these genes.

MATERIALS AND METHODS

Parasite material

Samples of *E. granulosus* used in this study originated from various intermediate hosts from Eurasia (Italy, Greece, Turkey, Iraq, Finland and Estonia) and South-Central America (Argentina and Mexico) (Table 1). In addition, one adult specimen of *E. multilocularis* from Estonia (described by Moks *et al*. 2005) was used.

DNA extraction, enzymatic amplification and sequencing

Total genomic DNA was isolated using the High Pure PCR Template Preparation Kit (Roche Diagnostics, Germany), following the manufacturer's protocols.

Partial sequences of the mitochondrial NADH dehydrogenase subunit 1 (*nad1*) gene were obtained for 3 previously uncharacterized samples to verify their respective mitochondrial genotypes (Table 1).

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

Table 1. Details of samples used in this study

	Nuclear genotype	Host	Geographical origin	Mitochondrial genotype	Reference for mitochondrial sequence
1.	G1A	Sheep	Turkey	G1	Utuk *et al*. (2008)
2.	G1B	Cattle	Turkey	G1	Utuk *et al*. (2008)
3.	G1C/G3	Sheep	Greece	G1	Varcasia *et al*. (2007)
4.	G1C/G3	Sheep	Italy	G3	Varcasia *et al*. (2006)
5.	G1D	Human	Iraq	G1	This study, FJ226756
6.	G4A	Horse	Italy	G4	Varcasia *et al*. (2008)
7.	G4B	Horse	Finland*	G4	This study, FJ226757
8.	G6/G7A	Cattle	Mexico	G7	Maravilla *et al*. (2004)
9.	G6/G7B	Pig	Argentina	G7	This study, FJ231470
10.	G8	Moose	Estonia	G8	Moks *et al*. (2008)
11.	G10A	Moose	Finland	G10	Lavikainen *et al*. (2006)
12.	G10B	Moose	Estonia	G10	Moks *et al*. (2006, 2008)
13.	EmulA	Fox	Estonia	Emul	Moks *et al*. (2005)

* The host imported to Finland, origin not documented.

Regions of *nad1* were amplified and sequenced as described previously (Lavikainen *et al*. 2006; Moks *et al*. 2006). Sequences were submitted to the GenBank database and are publicly available, with the Accession numbers being listed in Table 1. The mitochondrial sequences of other samples have been published previously, and the corresponding references can be found in Table 1. Thus, mitochondrial genotypes were known for all specimens used for further analyses.

To amplify 4 individual nuclear genes elongation factor 1 alpha (*ef1*), transforming growth factor beta receptor kinase (*tgf*), thioredoxin peroxidase (*th*) and calreticulin (*cal*), we developed primers specific to *Echinococcus* (Table 2). For the nuclear gene coding for ezrin-radixin-moesin-like protein (*elp*), previously described primers were used (Xiao *et al*. 2005; Table 2). We sequenced the 5 nuclear genes, yielding an alignment of 5086 base-pairs (bp) in total (Table 2). Conditions for the polymerase chain reaction (PCR) for all nuclear gene fragments were as follows: 20–80 ng of purified genomic DNA and 5 pmol of primers were used in the PCRs, which were performed in a total volume of 20 μl containing 1× BD Advantage 2 PCR buffer, 1U BD Advantage 2 Polymerase mix (BD Biosciences, USA), 0·2 mM dNTP (Fermentas) and 1·5 mM $MgCl_2$. Cycling was performed using a touchdown protocol: a 1 min-denaturing step at 95 °C, followed by 10 cycles of 20 sec at 95 °C, 30 sec at 59 °C and 2 min at 68 °C, but with the annealing temperature reduced by 0·5 °C in each step (touchdown), followed by 27 cycles of 20 sec at 95 °C, 30 sec at 55 °C and 2 min at 68 °C. The PCR was concluded with 5 min at 68 °C. PCR products were purified with shrimp alkaline phosphatase/exonuclease I treatment (Fermentas). One unit of each enzyme was added to 10 μl of PCR reaction mix and incubated for 30 min at 37 °C, followed by an inactivation at 80 °C for 15 min.

DNA cycle sequencing was performed in a total volume of 10 μl using the Big Dye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems), following the manufacturer's protocols. Initial denaturing was at 96 °C for 1 min, followed by 27 cycles at 96 °C for 10 sec (denaturation), 50 °C for 15 sec (annealing) and 60 °C for 4 min (extension). Sequences were resolved on the ABI3130xl automated DNA sequencer (Applied Biosystems). Both DNA strands were sequenced using the respective primers used in the primary amplification (Table 2). Consensus sequences were created using the Consed program (Gordon *et al*. 1998).

Sequences were aligned using Clustal W (Thompson *et al*. 1994) and checked and corrected in BioEdit (Hall, 1999). All sequences were deposited in GenBank and are available under Accession numbers EU834886–EU834940.

Phylogenetic analysis

Phylogenies were constructed for 11 different sets of alignments (hereafter referred to as 'datasets'): (1) a complete dataset of concatenated nucleotide sequences of 5 genes (5086 bp), each gene with its own best-fit nucleotide substitution model selected on the basis of Bayesian Information Criterion (BIC) scores from Modeltest 3.7 (Posada and Crandall, 1998). Concatenated nuclear sequences from this study were combined with all homologous sequences of *Echinococcus* spp. publicly available in GenBank: EF558359, EF558359, EF660053 (Hüttner *et al*. 2008), AJ012663 (Brehm *et al*. 1999), AB159141–AB159143 (Xiao *et al*. 2005) and AB306933–AB306940 (Nakao *et al*. unpublished observations). The resultant alignment included 20 taxa, comprising the entire genus *Echinococcus*; (2) a reduced dataset (13 taxa, only different genotypes of *E. granulosus* and *E. felidis* were included) of concatenated nucleotide sequences of 5 genes (5086 bp), each gene

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

Table 2. Information about the 5 nuclear genes used in this study

Nuclear gene	Length*	Primer name	Primer sequences (5′-3′)	Reference	Model**		
					DS1, Fig. 1A	DS2, Fig. 1B	DS3, Fig. 1C
1. *elp* (ezrin-radixin-moesin-like protein)	907 bp	ElpF	ATGCGCGTGAGAGTCTTCAGAAGA	Xiao *et al.* (2005)	HKY	HKY	HKY
		ElpR	ATTCTGCGAAGCTCAGCTTCA	Xiao *et al.* (2005)			
2. *ef1α* (elongation factor 1 alpha)	1055 bp	EfF	TCATTGTTATCGGTCACGTC	This study	K80 + I	TrNef	K80
		EfR	CTTCTGGGCAGATTTTTGTG	This study			
3. *tgfβ* (transforming growth factor beta receptor kinase)	1389 bp	TgfF	ATACAACCTGCTGCTTTTC	This study	K80	K80	K80
		TgfR	CAAGGTGTTACAGGGAGAAT	This study			
4. *th* (thioredoxin peroxidase)	597 bp	ThF	AAGTTTCACGTGCAAGGC	This study	K80	K80	K80
		ThR	ACGAGCTCATGAACGACTT	This study			
5. *cal* (calreticulin)	1138 bp	CalF	TTATTATCGCATTGTTATTCG	This study	K80	K80	K80
		CalR	GCAATCAAGAAGCCGTTC	This study			

* Sequence length of full nucleotide alignment (dataset 1) used in phylogenetic analysis. **Nucleotide substitution models, for a particular nuclear gene and dataset (DS), selected by Modeltest 3.7 according to the Bayesian Information Criterion.

with its own best-fit nucleotide substitution model; (3) a reduced dataset (12 taxa, only different genotypes of *E. granulosus* were included) of concatenated nucleotide sequences of 5 genes (5086 bp), each gene with its own best-fit nucleotide substitution model; (4) a nucleotide dataset based on coding regions (3512 bp), each gene with its best-fit nucleotide substitution model; (5) a nucleotide dataset based on coding regions (3512 bp), each codon position with its best-fit nucleotide substitution model; (6) a complete dataset of concatenated amino acid sequences coded by 5 genes (1169 aa), using best-fit substitution model selected on the basis of BIC-criterion by ProtTest 1.4 (Abascal *et al.* 2005); (7) – (11) nucleotide sequences of each nuclear gene fragment separately, with its own best-fit nucleotide substitution model. All alignments are available from the authors upon request. Best-fit substitution models selected by Modeltest 3.7 are shown in Table 2 for datasets 1–3 (Fig. 1A–C), models for other datasets are not shown.

To determine whether the 5 different nuclear genes had significantly different signals, the incongruence-length difference (ILD) test (Farris *et al.* 1995) alias 'partition homogeneity test' implemented in PAUP*4.0b10 (Swofford, 1998) was used. The homogeneities between all 5 nuclear gene sequences were calculated in a pairwise manner with the branch-and-bound search option, using 1000 replicates.

Bayesian phylogenetic analysis was performed in BEAST 1.4.7 (Drummond and Rambaut, 2007) using the best-fit model of sequence evolution. Posterior distributions of parameters, including the tree, were estimated by Markov Chain Monte Carlo (MCMC) sampling. Following a discard burn-in of 300 000 MCMC steps, samples were drawn every 3000 steps over a total of 3 000 000 steps. Samples were verified for acceptable mixing and convergence to a stationary distribution using Tracer 1.4 (Rambaut and Drummond, 2007). A second phylogeny was calculated using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003), with a best-fit model of sequence evolution (see Table 2). Searches were conducted with 4 simultaneous Markov Chains (3 heated and 1 cold chain) with over 2×10^6 generations, sampling every 100 generations; burn-in trees were discarded, and a 50% majority rule consensus tree was calculated. This approach resulted in phylogenetic trees with identical topology and similar nodal support (data not shown) to that produced by the BEAST analysis. Given this similarity, results are only presented for the BEAST analysis, as it has the advantage of allowing the root of the phylogenetic tree to be identified from the data, without the need for an outgroup. Phylogenetic trees were displayed using TreeView 1.6.6 (Page, 1996) and FigTree 1.1.2 (http://tree.bio.ed.ac.uk/software/figtree).

Each gene was evaluated for the presence of paralogues using a BLAST-search against contigs assembled for *E. multilocularis*, provided *via* the

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

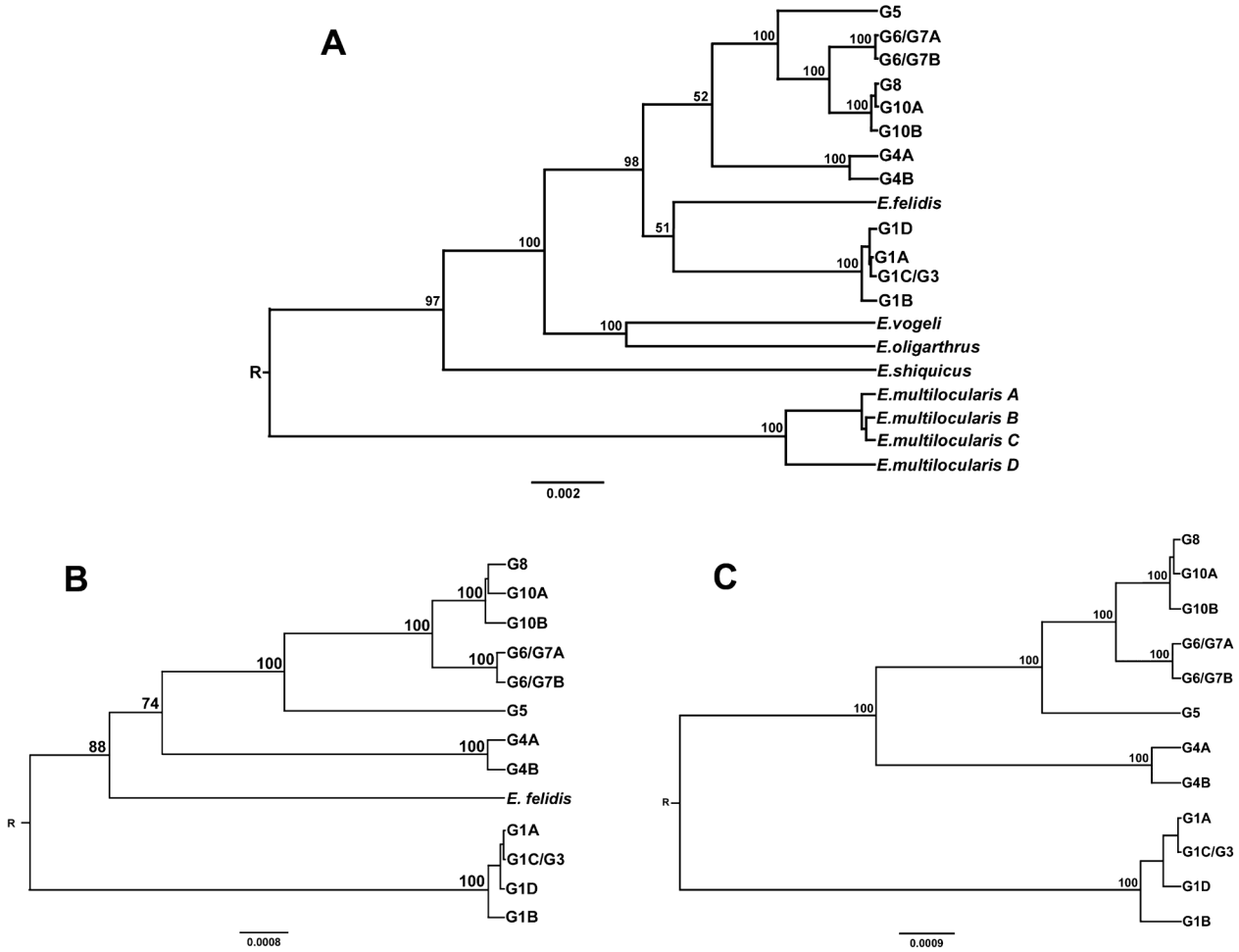


Fig. 1 (A–C). Bayesian phylograms of the genus *Echinococcus* based on sequences of 5 nuclear genes. (A) Full set of taxa (20 taxa; dataset 1, see Materials and Methods section); (B) including only 13 taxa that belong to the *E. granulosus*/*E. felidis* clade according to (A) (dataset 2); (C) including taxa used in (B), but with *E. felidis* additionally removed (12 taxa; dataset 3). Posterior probability values are indicated above branches. R, root of the tree.

Echinococcus Genome Project (http://www.sanger.ac.uk/Projects/Echinococcus). These contigs cover the vast majority of the genome (Dr M. Berriman, *Echinococcus* Genome Project Manager, personal communication).

RESULTS

Phylogenetic analysis based on a concatenated nucleotide sequence alignment of 5 nuclear genes (5086 bp, dataset 1; see also Table 3 for sequence characteristics) indicated that *E. multilocularis* is a basal taxon in the *Echinococcus* genus and that all *E. granulosus* genotypes form a monophyletic clade (Fig. 1A). Moreover, the cervid genotypes, G8 and G10, were positioned as sister taxa and formed a single clade, whereas the camel-pig genotypes G6/G7 appeared as a sister taxon to G8/G10. Although the phylogenetic analysis based on the complete set of taxa yielded a tree with a well-resolved overall topology for *Echinococcus* (all major nodes received high posterior probability values), it fell short of providing reasonable support for the internal phylogeny of the *E. granulosus*/*E. felidis* cluster. Two nodes that describe phylogenetic relations between clades G1–G3/*E. felidis* and G4–G10 did receive posterior probability values of just slightly >50% (Fig. 1A). Attempts to resolve phylogenies with a wide range of taxa can produce nucleotide data matrices that contain homoplasic characters, i.e. distant taxa may share characters that are not common by descent, but have arisen due to convergence (Arendt and Reznik, 2008). Homoplasy can lead to erroneous affinities between taxa and consequently result in a topology with low posterior probability values or even an erroneous topology (Naylor and Brown, 1998; Engstrom *et al*. 2004). As it was not possible to resolve the phylogenetic relationships between clades G1–G3/*E. felidis* and G4–G10 by manipulating the gene sets included for analysis (data not shown), we hypothesized that homoplasy was most likely the reason why 2 nodes received only moderate posterior probability values when the full set of taxa was used (Fig. 1A). To test this hypothesis, we took a

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

Table 3. Variability in the full nucleotide, coding and amino acid sequences of 5 nuclear genes

Nuclear gene	Lengths of nucleotide/coding/ amino acid sequences	Variability of nucleotide sequences*	Variability of coding sequences*	Variability of amino acid sequences*
1. *elp*	907bp/57bp/19aa	99V:65PI + 34S (7·2)	3V:2PI + 1S (3·5)	2V:1PI + 1S (5·3)
2. *ef1*	1055bp/1022bp/340aa	74V:44PI + 30S (4·2)	68V:38PI + 30S (3·7)	7V:4PI + 3S (1·2)
3. *tgf*	1389bp/939bp/313aa	67V:21PI + 46S (1·5)	33V:11PI + 22S (1·2)	14V:6PI + 8S (1·9)
4. *th*	597bp/528bp/175aa	16V:8PI + 8S (1·3)	13V:7PI + 6S (1·3)	6V:4PI + 2S (2·3)
5. *cal*	1138bp/966bp/322aa	19V:18PI + 1S (1·6)	12V:12PI + 0S (1·2)	8V:8PI + 0S (2·5)
Σ	5086bp/3512bp/1169aa	275V:156PI + 119S (3·2)	129V:70PI + 59S (2·2)	37V:23PI + 14S (2·6)

* Number of variable sites (V): parsimony informative sites (PI) + singletons (S), number of parsimony informative sites per 100 base-pairs of nucleotide sequences or 100 amino acid sequences are in parentheses.

subset of the taxa whose relationship remained questionable and performed another analysis, which gave significantly better posterior probability values (Fig. 1B). When *E. felidis* was additionally removed, very high posterior probability values were obtained for clades G1–G3 and G4–G10 (Fig. 1C), confirming the topology of the phylogenetic tree constructed using the full set of taxa (see Fig. 1A). Based on the results of these 3 phylogenies, a summarized phylogeny of the genus *Echinococcus* is presented (Fig. 2B). Datasets 4–6 (i.e., coding regions and amino acid sequences) yielded phylogenetic trees with low nodal supports (data not shown). Datasets 7–11 (i.e., nucleotide alignments for each nuclear gene separately) yielded varying results (data not shown): dataset 7 (*elp* gene) provided a tree with a similar topology to that derived from dataset 1, with the notable difference that the positions of *E. granulosus* genotypes G1 and G4 were swapped. Dataset 8 (gene for *ef1*) also gave a different topology, namely, G4 grouped together with *E. felidis* to form a monophyletic clade, whereas the relative positions of G1 and G4 were in agreement with dataset 1. However, compared with dataset 1 (Fig. 1A), all of these differences were linked to lower posterior probability values and could thus be considered as less reliable. Alignments based on sequences of other genes (*tgf*, *th* and *cal*) yielded trees that were in agreement with dataset 1, albeit with moderate posterior probability values. Using the ILD test, no incongruence was found between the 5 nuclear genes. No paralogues were detected for any of the nuclear sequences.

Variation in nuclear genes

All nuclear genes were characterized in terms of the variability of their overall nucleotide sequences, coding sequences and inferred amino acid sequences. The main characteristics for each nuclear gene are presented in Table 3; here, we summarize these characteristics.

The fragment of the *elp* gene used in this study was composed of a long intron sequence and a relatively short protein-coding region. Although indels were found at the overall nucleotide level – one indel was 6 bp, two were 3 bp and three were 1 bp long – sequences could be aligned unambiguously. The nucleotide sequence of the *elp* gene fragment provided 65 parsimony informative sites for phylogenetic analysis, whereas its coding and amino acid sequences provided only 2 and 1, respectively.

The partial *ef1* gene sequence was composed of 2, long coding sequences interrupted by a short intron. No indels were detected. The nucleotide sequence of the *ef1* gene fragment provided 44 parsimony informative sites, whereas its coding and amino acid sequences provided 38 and 4 sites, respectively.

The partial *tgf* gene sequence was composed of 2 long coding sequences interrupted by a moderate sized intron. No indels were present. The nucleotide sequence of the *tgf* gene fragment provided 21 parsimony informative sites, whereas its coding and amino acid sequences provided 11 and 6 sites, respectively.

The partial *th* gene sequence was composed of 2 long coding sequences interrupted by a short intron. No indels were present. The nucleotide sequence of the *th* gene fragment provided 8 parsimony informative sites, whereas its coding and amino acid sequences provided 7 and 4 sites, respectively.

The partial *cal* gene sequence was composed of coding and intron sequences. No indels were found. The nucleotide sequence of the *cal* gene fragment provided 18 parsimony informative sites, whereas its coding and amino acid sequences provided 12 and 8 sites, respectively.

DISCUSSION

The taxonomy of the genus *Echinococcus* has remained controversial, despite numerous attempts to revise it (for a recent comprehensive review, see Thompson, 2008). Among the most difficult tasks has been to unravel relationships between *E. granulosus* and other species within the genus, and to explain the remarkable phenotypic and genetic variability between populations of *E. granulosus* from different host species and different geographical

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

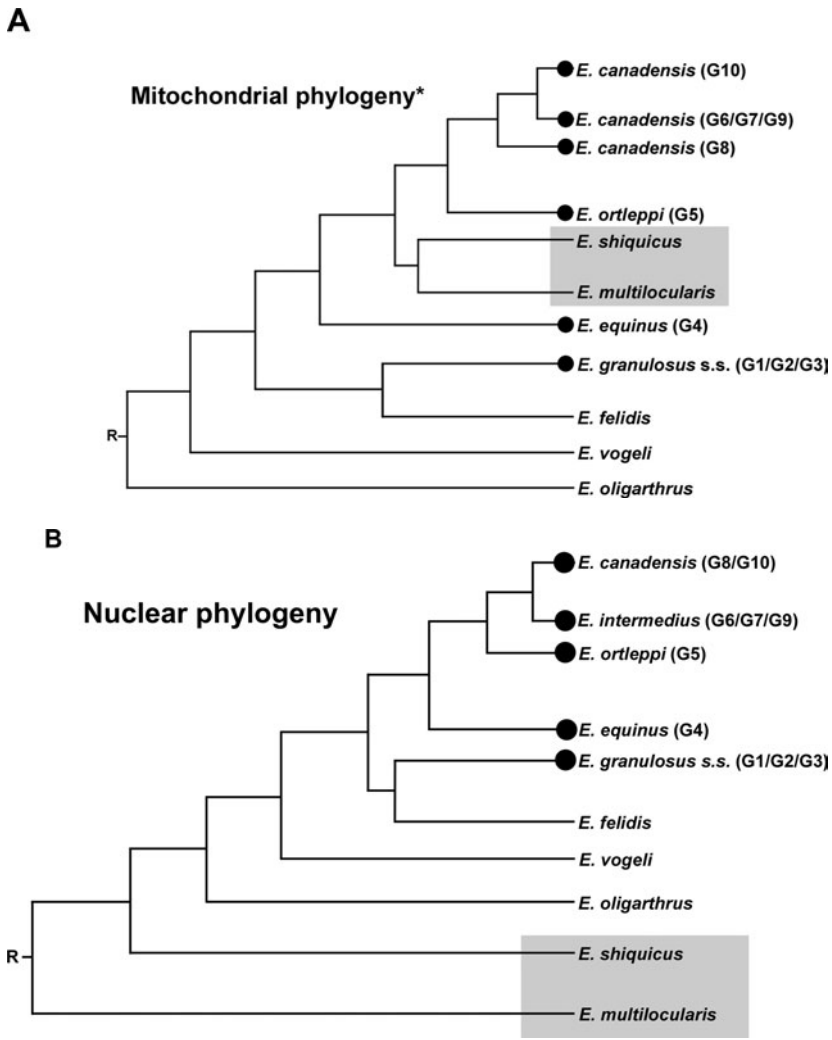


Fig. 2. Comparison of phylogenies of the genus *Echinococcus* derived from mitochondrial and nuclear data. (A) Phylogenetic tree of the genus *Echinococcus* summarizing recent data based primarily on mitochondrial sequences. Relationships between G6–G10 are summarized from Nakao *et al*. (2007) and Moks *et al*. (2008); the position of *E. felidis* is taken from Hüttner *et al*. (2008), who based their phylogeny on concatenated mitochondrial and nuclear data. (B) Phylogenetic tree of the genus *Echinococcus* from this study, based on 5 nuclear sequences. The position of *E. multilocularis* and *E. shiquicus* on both phylogenetic trees is indicated by a grey rectangle. Large dots represent the locations of *E. granulosus* genotypes on the trees.

localities. Several of these populations exhibit distinct genetic and phenotypic characteristics that have prompted the distinction of separate populations into 10 genotypes within *E. granulosus* (note that the existence of genotype G9 is not well supported; see Introduction section). The most critical issues have been to determine whether these genotypes form a monophyletic entity and whether some of the genotypes can be recognized as separate species. Results from mitochondrial data have raised doubts about the monophyly of *E. granulosus*; the conclusion from a phylogeny based on the full mitochondrial genomic data was that *E. granulosus* genotypes do not form a monophyletic entity and are divided into 2 clusters, separated from each other by 2 other species of the genus, namely *E. multilocularis* and *E. shiquicus* (see Nakao *et al*. 2007).

E. multilocularis, E. shiquicus, E. oligarthrus *and* E. vogeli

In the present study, we demonstrated that the phylogeny of *Echinococcus* inferred from nuclear genes (Figs 1A and 2B) differed significantly from the mitochondrial phylogeny (Fig. 2A). According to the nuclear phylogeny, *E. multilocularis* occupied a basal position in the genus with *E. shiquicus* positioned next to it. This result challenges the mitochondrial phylogeny, where both *E. multilocularis* and *E. shiquicus* were positioned between the clades G1–G3 and G4/G5–G10. The phylogenetic positions of *E. vogeli* and *E. oligarthrus* based on the nuclear data were also different from that of mitochondrial; whereas in the latter, the *E. oligarthrus* was shown as the basal taxon within the genus and

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

E. vogeli was positioned next to it (cf. Nakao *et al.* 2007), the nuclear phylogeny demonstrated that both of them were positioned as ingroup taxa between *E. shiquicus* and the *E. granulosus*/*E. felidis* clade.

E. granulosus *sensu stricto* (*genotypes G1, G2, G3*) *and* E. felidis

The results of this study, based on nuclear genes, are in agreement with previous mitochondrial data (Bowles *et al.* 1995; Thompson *et al.* 1995; Obwaller *et al.* 2004; Jenkins *et al.* 2005; Romig *et al.* 2006; Nakao *et al.* 2007; Busi *et al.* 2007; Rinaldi *et al.* 2008) in placing *E. granulosus* genotypes G1, G2 and G3 as very closely related taxa. Moreover, these 3 genotypes have wide and overlapping geographical distributions and host spectra (Busi *et al.* 2007; Varcasia *et al.* 2007; Rinaldi *et al.* 2008). Thus, genotypes G1–G3 can be regarded as a single taxonomic unit: *E. granulosus sensu stricto* (see Nakao *et al.* 2007; Thompson 2008).

Although *E. felidis* was positioned close to the genotypes G1–G3 in the study by Hüttner *et al.* (2008) and in the current study, its position in both cases was clearly as a sister taxon to G1–G3. Considering that contemporary samples (parasite eggs in lion faeces) yielded the same DNA sequences as obtained from a 40-year-old museum specimen of *E. felidis* (see Hüttner *et al.* 2008), it is very likely that *E. felidis* is a distinct genetic entity. Based on all available data, it can be concluded that *E. felidis* shared a recent common ancestor with G1–G3, but has evolved into a separate phylogenetic entity and most likely also into a distinct species. However, further genetic and phenotypic studies of G1–G3 and *E. felidis* from a range of host species in Africa are required to test whether *E. felidis* deserves the status of a distinct species.

E. equinus (*genotype G4*) *and* E. ortleppi (*genotype G5*)

Extensive mitochondrial (Le *et al.* 2002; Nakao *et al.* 2007) and nuclear analyses (this study) all place *E. granulosus* genotypes G4 and G5 on the phylogenetic tree as taxa separate from others and also from each other. Williams and Sweatman (1963) originally proposed that *E. granulosus* of horse origin (nowadays called genotype G4) can be regarded as the subspecies *E. g. equinus*. However, while subspecies can potentially interbreed, no evidence of any segregating mechanism between the sympatric *E. g. equinus* and *E. g. granulosus* has been identified that could explain their separation. Thus, the subspecies status of *E. g. equinus* was rejected (Rausch, 1967). More recent investigations of *E. granulosus* have shown that the parasite of horse origin is biologically and genetically distinct from that of sheep origin. Considering all of the data, Thompson and McManus (2002) proposed that the equine genotype G4 should be regarded as a distinct species, *E. equinus*. Genotype G5, which occurs in Europe, Asia, Africa and South America, appears to be adapted to cattle as its principal intermediate host. Adult stages of the cattle isolate were originally described as *E. ortleppi* by Lopez-Neyra and Soler Planas (1943). Besides clear genetic differentiation, it is also phenotypically and developmentally distinct from *E. granulosus* of sheep or horse origin (Thompson *et al.* 1984; Bowles *et al.* 1995; Thompson and McManus, 2002; Thompson, 2008). Thus, the phylogenetic analyses of this study support earlier propositions that these genotypes should be considered as separate species, i.e., genotype G4 as *E. equinus* and G5 as *E. ortleppi*.

E. intermedius (*genotypes G6, G7, G9*) *and* E. canadensis (*genotypes G8 and G10*)

A recent study, based on mitochondrial data by Moks *et al.* (2008), demonstrated that the *E. granulosus* camel-pig genotypes G6/G7 formed a subgroup together with the cervid genotype G10, whereas the other cervid genotype G8 was shown to be a sister taxon to G6/G7–G10. Moks *et al.* (2008) concluded that their study provided support for combining G6/G7, G8 and G10 into a single species which, in terms of priority, should be *E. canadensis* (see also Nakao *et al.* 2007; Hüttner *et al.* 2008; Thompson, 2008). The results of the current study, based on nuclear data, provided a different view on the phylogeny of the camel-pig and cervid genotypes; there was no separation of the cervid genotypes G8 and G10 into distinct clades. Genotype G8 was separated from Finnish G10A by 2 point mutations, whereas the same number of mutations also separated G10A and Estonian G10B isolates. Thus, the variation within G10 was the same as that between genotypes G8 and G10. By contrast, the cervid (G8/G10A/G10B) and camel-pig (G6/G7) genotypes were separated from each other by as many as 12 point mutations. Unlike the mitochondrial phylogeny, where genotypes G6/G7 were positioned in a clade together with G10 (but separate from G8), the current nuclear phylogeny demonstrated that both the cervid genotypes, G8 and G10, form a single clade separate from the camel-pig genotypes G6/G7. Future studies should include nuclear sequence data of G8 and G10 isolates from North America to evaluate whether the genetic differentiation within *Echinococcus* from cervid hosts is associated with their geographical distribution.

From a taxonomic perspective, it has been suggested that the camel, pig and cervid genotypes (G6–G10) might belong to a single species (Thompson and McManus, 2001; Lavikainen *et al.* 2006; Moks *et al.* 2008), named *E. canadensis* (see Nakao *et al.* 2007; Hüttner *et al.* 2008). However, Thompson (2008) pointed out that this may be an

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

oversimplification of the evidence, as cervid and camel-pig isolates of *Echinococcus* differ from each other both in terms of their mitochondrial sequences and their geographical distributions (Lavikainen *et al*. 2003, 2006; Thompson *et al*. 2006; Moks *et al*. 2008). Moreover, Thompson (2008) suggested that the camel-pig genotypes G6/G7 should be regarded as a single species, *E. intermedius*, after Lopez-Neyra and Soler Planas (1943). Data derived from the nuclear genes in this study support the proposal of uniting cervid genotypes G8/G10 into a single species, *E. canadensis*, distinct from G6/G7, which could be regarded as *E. intermedius*. However, molecular data alone are not sufficient to revise the taxonomy. Thus, further biological and morphological observations, including, above all, evaluation of morphological characters and infection experiments, are required for the final revision.

Nuclear genes sequences, their variability and contribution to the phylogeny

The most informative markers at an 'overall' nucleotide level (dataset 1) were the *elp* and *ef1* sequences, which provided 7·2 and 4·2 parsimony informative sites per 100 bp, and a total of 65 and 44 parsimony informative sites, respectively. Alignments of the other 3 genes provided far less parsimony informative sites. All 5 nuclear genes used in this study were useful for phylogenetic analysis, as alignments based on both nucleotide and amino acid characters were unequivocal for all genes. Although introns were present in nuclear genes and indels were detected for the *elp* gene, these did not interfere with the alignments. Moreover, no paralogues were found for these 5 genes, and pairwise ILD tests of the nuclear genes showed no incongruence between them. However, different genes provided different signals in relation to the position of G4 (in the phylogeny derived from the *elp* gene, the positions of G4 and G1 were swapped, compared with the concatenated dataset phylogeny) and *E. felidis* (the *ef1* gene phylogeny united G4 and *E. felidis* into a monophyletic clade). Nevertheless, tree topologies of all other genes were in agreement with the concatenated dataset 1. As manipulations with sequence sets did not improve the phylogeny, slight disagreement between the phylogenies from different genes was most likely not the reason why 2 nodes (within the clade that represents the genotypes of *E. granulosus*) on the dataset 1-phylogeny received low support. Rather, homoplasy (i.e., highly divergent taxa sharing characters that are common not due to descent) may have contributed to the low support for these 2 nodes. Indeed, when distant taxa, such as *E. multilocularis*, *E. shiquicus*, *E. oligarthrus* and *E. vogeli*, were removed from the dataset, these 2 nodes in the phylogeny of *E. granulosus* received much higher support; also, when *E. felidis* was additionally removed from the dataset, the internal phylogeny between different genotypes of *E. granulosus* was fully resolved, demonstrating clearly the basal position of G1 with respect to G4. Non-coding sequences were of significance for inferring the phylogeny of *Echinococcus*, since neither nucleotide alignments based solely on coding sequences (3512 bp, datasets 4 and 5) nor alignment based on amino acids (1169 aa, dataset 6) were able to yield phylogeny with high posterior probability values. The reason for this lies in the variability found in the different datasets. Compared with the 'full' nucleotide alignment (dataset 1), which provided a total of 156 parsimony informative sites, alignments based on either coding or amino acid sequences (datasets 4–6) yielded far fewer parsimony informative sites. Although phylogenies constructed separately for each gene (datasets 7–11) were generally in agreement with the topology based on full nucleotide alignment, none of them was able to provide phylogenetic resolution equivalent to that based on the full nucleotide dataset.

Is a nuclear phylogeny preferable to a mitochondrial one as the basis for interpreting the evolutionary history of the genus Echinococcus*?*

This study has clearly shown that a nuclear phylogeny of the genus *Echinococcus* differed significantly from previous mitochondrial phylogenies (though a certain level of agreement was also found). This finding presents us with an important question: which phylogeny represents more accurately the true evolutionary history of *Echinococcus*? Here, we argue in favour of the nuclear phylogeny for several reasons. Firstly, the divergence of maternal lineages is a stochastic event, as mutations accumulate in the mitochondrial genome at random, except at those sites that are under selection. Once a new mutation becomes fixed in a population, the new mitochondrial lineage separates from the ancestral one. From this point onwards, mutations continue to fix progressively in an independent manner in both the new and ancestral mitochondrial lineages. In the absence of recombination, mitochondrial lineages continue to diverge once they have separated. However, this does not mean that these separate mitochondrial lineages have necessarily become separate biological entities. Separation of mitochondrial lineages without accompanying reproductive isolation has no biological meaning from the viewpoint of true separation of populations (eventually into distinct species), but merely reflects the course of evolution of maternal lineages. Although mutations also accumulate into the nuclear genome at random (except at those sites under selection), nuclear genes undergo recombination and if there is no barrier for gene-flow among parasites, then nuclear genes that are subjected to recombination during meiosis do not show any

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

separation into genetically distinct populations (for recent reviews, see Funk and Omland, 2003; Hurst and Jiggins, 2005; Ballard and Rand, 2005). Thus, even if the separation of the *E. granulosus* genotypes G8 and G10 were clearly established, as seen in the mitochondrial data reported by Moks *et al.* (2008), the nuclear data in this study showed that parasites with the mtDNA genotypes G8 and G10 have not actually separated into distinct biological entities.

Although, traditionally, mitochondrial gene sequences have been preferred over nuclear ones in phylogenetic studies, the current study illustrates the main limitation of using mitochondrial data as the sole marker for inferring phylogenies, i.e., the evolutionary history of a maternal lineage may differ from that at the species level or higher. Given the significant inconsistencies between phylogenies from mitochondrial and nuclear data, we suggest that both marker systems are used to evaluate the genetic diversity of the genus *Echinococcus* in future studies, whereas nuclear genes should be used when evaluating the phylogenetic positions of new *Echinococcus* isolates.

Despite not always being the best marker for phylogenetic studies, mitochondrial sequences are undoubtedly of great diagnostic value, particularly when samples are degraded. As cells contain relatively few copies of each nuclear gene but often thousands of mitochondrial genomes, it is clear that analysis of mitochondrial gene sequence data may often represent the best way to reveal the identity of a parasite species or genotype/strain.

We thank Jaak Kinnas, Tuuli Reisberg, Maris Roht, and John Davison for their generous help. This work was supported by grants from the Estonian Ministry of Education and Sciences (target financing grant TLOOM0122), the Estonian Science Foundation (grant 7040), the Estonian Ministry of Agriculture (grant LLOOM08015), the European Union through the European Regional Development Fund (Center of Excellence FIBIR) and from the Environmental Investment Centre.

REFERENCES

Abascal, F., Zardoya, R. and Posada, D. (2005). ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105.

Anderson, S., Bankier, A. T., Barrell, B. G., Debruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature, London* **290**, 457–465.

Arendt, J. and Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends in Ecology and Evolution* **23**, 26–32.

Avise, J. C., Arnold, R. M., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. and Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489–522.

Ballard, J. W. O. and Rand, D. M. (2005). The population biology of mitochondrial DNA and its phylogenetic implications. *Annual Review of Ecology, Evolution and Systematics* **36**, 621–642.

Bowles, J., Blair, D. and McManus, D. P. (1992). Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology* **54**, 165–173.

Bowles, J., Blair, D. and McManus, D. P. (1994). Molecular genetic characterization of the cervid strain ('northern form') of *Echinococcus granulosus*. *Parasitology* **109**, 215–221.

Bowles, J., Blair, D. and McManus, D. P. (1995). A molecular phylogeny of the genus *Echinococcus*. *Parasitology* **110**, 317–328.

Brehm, K., Jensen, K., Frosch, P. and Frosch, M. (1999). Characterization of the genomic locus expressing the ERM-like protein of *Echinococcus multilocularis*. *Molecular and Biochemical Parasitology* **100**, 147–152.

Brown, W. M. (1983). Evolution of animal mitochondrial DNA. In *Evolution of Genes and Proteins* (ed. Nei, M. and Koehn, R. K.), pp. 62–88. Sinauer, Sunderland, Mass, USA.

Brown, W. M., George, M., Jr. and Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA* **76**, 1967–1971.

Busi, M., Snabel, V., Varcasia, A., Garippa, G., Perrone, V., De Liberato, C. and Amelio, S. D. (2007). Genetic variation within and beween G1 and G3 genotypes of *Echinococcus granulosus* in Italy revealed by multilocus DNA sequencing. *Veterinary Parasitology* **150**, 75–83.

Drummond, A. J. and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.

Engstrom, T. N., Shaffer, H. B. and McCord, W. P. (2004). Multiple data sets, high homoplasy, and the phylogeny of softshell turtles (*Testudines: Trionychidae*). *Systematic Biology* **53**, 693–710.

Farris, J. S., Källersjö, M., Kluge, A. G. and Bult, C. (1995). Testing significance of incongruence. *Cladistics* **10**, 315–319.

Funk, D. J. and Omland, K. E. (2003). Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics* **34**, 397–423.

Giles, R. E., Blanc, H., Cann, M. and Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA* **77**, 6715–6719.

Gordon, D., Abajian, C. and Green, P. (1998). Consed: A graphical tool for sequence finishing. *Genome Research* **8**, 195–202.

Hall, T. A. (1999). BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series* **41**, 95–98.

Hurst, G. D. D. and Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

inherited symbionts. *Proceedings of the Royal Society, B* **272**, 1525–1534.

Hüttner, M., Nakao, M., Torsten, W., Siefert, L., Boomker, J. D. F., Dinkel, A., Sako, Y., Mackenstedt, U., Romig, T. and Ito, A. (2008). Genetic characterization and phylogenetic position of *Echinococcus felidis* (Cestoda: Taeniidae) from the African lion. *International Journal for Parasitology* **38**, 861–868.

Jannotti-Passos, L. K., Souza, C. P., Parra, J. C. and Simpson, A. J. G. (2001). Biparental mitochondrial DNA inheritance in the parasitic trematode *Schistosoma mansoni*. *Journal of Parasitology* **87**, 79–82.

Jenkins, D. J., Romig, T. and Thompson, R. C. A. (2005). Emergence/re-emergence of *Echinococcus* spp., a global update. *International Journal for Parasitology* **35**, 1205–1219.

Kedra, A. H., Swiderski, Z., Tkach, V. V., Dubinsky, P., Pawlowski, Z., Stefaniak, J. and Pawlowski, J. (1999). Genetic analysis of *Echinococcus granulosus* from humans and pigs in Poland, Slovakia and Ukraine. A multicenter study. *Acta Parasitologica* **44**, 248–254.

Lavikainen, A., Lehtinen, M. J., Laaksonen, S., Ågren, E., Oksanen, A. and Meri, S. (2006). Molecular characterization of *Echinococcus* isolates of cervid origin from Finland and Sweden. *Parasitology* **133**, 565–570.

Lavikainen, A., Lehtinen, M. J., Meri, T., Hirvelä-Koski, V. and Meri, S. (2003). Molecular genetic characterization of the Fennoscandian cervid strain, a new genotype group (G10) of *Echinococcus granulosus*. *Parasitology* **127**, 207–215.

Le, T. H., Blair, D., Agatsuma, T., Humair, P. F., Campbell, N. J., Iwagami, M., Littlewood, D. T., Peacock, B., Johnston, D. A., Bartley, J., Rollinson, D., Herniou, E. A., Zarlenga, D. S. and McManus, D. P. (2000). Phylogenies inferred from mitochondrial gene orders – a cautionary tale from the parasitic flatworms. *Molecular Biology and Evolution* **17**, 1123–1125.

Le, T. H., Pearson, M. S., Blair, D., Dai, N., Zhang, L. H. and McManus, D. P. (2002). Complete mitochondrial genomes confirm the distinctiveness of the horse-dog and sheep-dog strains of *Echinococcus granulosus*. *Parasitology* **124**, 97–112.

Leister, D. (2005). Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends in Genetics* **21**, 655–663.

López-Neyra, C. R. and Soler Planas, M. A. (1943). Revision del genero *Echinococcus* Rud y descripción de una especie nueva parárita intestinal del perro en Almería. *Revista Ibérica de Parasitología* **3**, 169–194 (in Spanish).

Maravilla, P., Thompson, R. C. A., Palacios-Ruiz, J. A., Estcourt, A., Ramirez-Solis, E., Mondragon-de-la-Peña, C., Moreno-Moller, M., Cardenas-Mejia, A., Mata-Miranda, P., Aguirre-Alcantara, M. T., Bonilla-Rodriguez, C. and Flisser, A. (2004). *Echinococcus granulosus* cattle strain identification in an autochthonous case of cystic echinococcosis in central Mexico. *Acta Tropica* **92**, 231–236.

McManus, D. P. and Thompson, R. C. A. (2003). Molecular epidemiology of cystic echinococcosis. *Parasitology* **127**, S37–S51.

Moks, E., Jõgisalu, I., Saarma, U., Talvik, H., Järvis, T. and Valdmann, H. (2006). Helminthologic survey of the wolf (*Canis lupus*) in Estonia, with an emphasis on *Echinococcus granulosus*. *Journal of Wildlife Diseases* **42**, 359–365.

Moks, E., Jõgisalu, I., Valdmann, H. and Saarma, U. (2008). First report of *Echinococcus granulosus* G8 in Eurasia and a reappraisal of the phylogenetic relationships of 'genotypes' G5–G10. *Parasitology* **135**, 647–654.

Moks, E., Saarma, U. and Valdmann, H. (2005). *Echinococcus multilocularis* in Estonia. *Emerging Infectious Diseases* **11**, 1973–1974.

Nakao, M., McManus, D. P., Schantz, P. M., Craig, P. S. and Ito, A. (2007). A molecular phylogeny of the genus *Echinococcus* inferred from complete mitochondrial genomes. *Parasitology* **134**, 713–722.

Naylor, G. J. P. and Brown, W. M. (1998). Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Systematic Biology* **47**, 61–76.

Obwaller, A., Schneider, R., Walochnik, J., Gollackner, B., Deutz, A., Janitschke, K., Aspöck, H. and Auer, H. (2004). *Echinococcus granulosus* strain differentiation based on sequence heterogeneity in mitochondrial genes of cytochrome *c* oxidase-1 and NADH dehydrogenase-1. *Parasitology* **128**, 569–575.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357–358.

Posada, D. and Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Rambaut, A. and Drummond, A. J. (2007). Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer

Rausch, R. L. (1967). A consideration of infraspecific categories in the genus *Echinococcus* Rudolphi, 1801 (Cestoda: Taeniidae). *The Journal of Parasitology* **53**, 484–491.

Rinaldi, L., Maurelli, M. P., Capuano, F., Perugini, A. G., Veneziano, V. and Cringoli, S. (2008). Molecular update on cystic echinococcosis in cattle and water buffaloes of southern Italy. *Zoonoses and Public Health* **55**, 119–123.

Romig, T., Dinkel, A. and Mackenstedt, U. (2006). The present situation of echinococcosis in Europe. *Parasitology International* **55**, S187–S191.

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

Scott, J. C., Stefaniak, J., Pawlowski, Z. S. and McManus, D. P. (1997). Molecular genetic analysis of human cystic hydatid cases from Poland: identification of a new genotypic group (G9) of *Echinococcus granulosus*. *Parasitology* **114**, 37–43.

Swofford, D. L. (1998). *PAUP*: Phylogenetic Analysis Using Parsimony* (**and Other Methods*), *Version 4.0g10*. Sinauer Associates, Massachusetts, USA.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press

weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Thompson, R. C. A. (2008). The taxonomy, phylogeny and transmission of *Echinococcus*. *Experimental Parasitology* **119**, 439–446.

Thompson, R. C. A., Boxell, A. C., Ralston, B. J., Constantine, C. C., Hobbs, R. P., Shury, T. and Olson, M. E. (2006). Molecular and morphological characterization of *Echinococcus* in cervids from North America. *Parasitology* **132**, 439–447.

Thompson, R. C. A., Kumaratilake, L. M. and Eckert, J. (1984). Observations on *Echinococcus granulosus* of cattle origin in Switzerland. *International Journal for Parasitology* **14**, 283–291.

Thompson, R. C. A., Lymbery, A. J. and Constantine, C. C. (1995). Variation in *Echinococcus*: towards a taxonomic revision of the genus. *Advances in Parasitology* **35**, 145–176.

Thompson, R. C. A. and McManus, D. P. (2001). Aetiology: parasites and life cycles. In *Manual on* Echinococcus *in Humans and Animals a Public Health Problem of Global Concern* (ed. Eckert, J., Gemmell, M. A., Meslin, F.-X. and Pawlowski, Z. S.), pp. 1–19. WHO/OIE, Paris. World Health Organization, Geneva, Switzerland.

Thompson, R. C. A. and McManus, D. P. (2002). Towards a taxonomic revision of the genus *Echinococcus*. *Trends in Parasitology* **18**, 452–457.

Utuk, A. E., Simsek, S., Koroglu, E. and McManus, D. P. (2008). Molecular genetic characterization of different isolates of *Echinococcus granulosus* in East and Southeast regions of Turkey. *Acta Tropica* **107**, 192–194.

Varcasia, A., Canu, S., Kogkos, A., Pipia, A. P., Scala, A., Garippa, G. and Seimenis, A. (2007). Molecular characterization of *Echinococcus granulosus* in sheep and goats of Peloponnesus, Greece. *Parasitology Research* **101**, 1135–1139.

Varcasia, A., Canu, S., Lightowlers, M. W., Scala, A. and Garippa, G. (2006). Molecular characterization of *Echinococcus granulosus* strains in Sardinia. *Parasitology Research* **98**, 273–277.

Varcasia, A., Garippa, G., Pipia, A. P., Scala, A., Brianti, E., Giannetto, S., Battelli, G., Poglayen, G. and Micagni, G. (2008). Cystic echinococcosis in equids in Italy. *Parasitology Research* **102**, 815–818.

Williams, R. J. and Sweatman, G. K. (1963). On the transmission, biology and morphology of *Echinococcus granulosus equinus*, a new subspecies of hydatid tapeworm in horses in Great Britain. *Parasitology* **53**, 391–407.

Xiao, N., Qiu, J., Nakao, M., Li, T., Yang, W., Chen, X., Schantz, P. M., Craig, P. S. and Ito, A. (2005). *Echinococcus shiquicus* n. sp., a taeniid cestode from Tibetan fox and plateau pika in China. *International Journal for Parasitology* **35**, 693–701.

https://doi.org/10.1017/S0031182008005453 Published online by Cambridge University Press