New mitogenome and nuclear evidence on the phylogeny and taxonomy of the highly zoonotic tapeworm Echinococcus granulosus sensu stricto

Kinkar, Liina

2017-08

Kinkar , L , Laurimae , T , Sharbatkhori , M , Mirhendi , H , Kia , E B , Ponce-Gordo , F , Andresiuk , V , Simsek , S , Lavikainen , A , Irshadullah , M , Umhang , G , Oudni-M'rad , M , Acosta-Jamett , G , Rehbein , S & Saarma , U 2017 , ' New mitogenome and nuclear evidence on the phylogeny and taxonomy of the highly zoonotic tapeworm Echinococcus granulosus sensu stricto ' , Infection, genetics and evolution , vol. 52 , pp. 52-58 . https://doi.org/10.1016/j.meegid.20

http://hdl.handle.net/10138/237086
https://doi.org/10.1016/j.meegid.2017.04.023

publishedVersion

Infection, Genetics and Evolution 52 (2017) 52–58

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

New mitogenome and nuclear evidence on the phylogeny and taxonomy of the highly zoonotic tapeworm *Echinococcus granulosus* sensu stricto

Liina Kinkar [a], Teivi Laurimäe [a], Mitra Sharbatkhori [b], Hossein Mirhendi [c], Eshrat Beigom Kia [d], Francisco Ponce-Gordo [e], Vanessa Andresiuk [f], Sami Simsek [g], Antti Lavikainen [h], Malik Irshadullah [i], Gérald Umhang [j], Myriam Oudni-M'rad [k], Gerardo Acosta-Jamett [l], Steffen Rehbein [m], Urmas Saarma [a,*]

[a] Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Vanemuise 46, 50410 Tartu, Estonia
[b] Laboratory Sciences Research Center, Golestan University of Medical Sciences, Gorgan, Iran
[c] Department of Medical Mycology and Parasitology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
[d] Department of Medical Parasitology and Mycology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[e] Departmento de Parasitología, Facultad de Farmacia, Plaza Ramón y Cajal s/n, UCM, Madrid, Spain
[f] Laboratorio de Zoonosis Parasitarias, FCEyN, UNMdP, Funes 3350, CP: 7600, Mar del Plata, Buenos Aires, Argentina
[g] Department of Parasitology, Faculty of Veterinary Medicine, University of Firat, 23119 Elazig, Turkey
[h] Department of Bacteriology and Immunology/Immunobiology Program, Faculty of Medicine, P.O. Box 21, FIN-00014, University of Helsinki, Finland
[i] Section of Parasitology, Department of Zoology, Aligarh Muslim University, Aligarh 202002, India
[j] Anses, Wildlife Surveillance and Eco-epidemiology Unit, National Reference Laboratory for Echinococcus spp., Nancy Laboratory for Rabies and Wildlife, 54220 Malzéville, France
[k] LP3M: Laboratory of Medical and Molecular Parasitology-Mycology, LR12ES08, Faculty of Pharmacy, University of Monastir, 5000 Monastir, Tunisia
[l] Instituto de Medicina Preventiva Veterinaria y Programa de Investgación Aplicada en Fauna Silvestre, Facultad de Ciencias Veterinarias, Universidad Austral de Chile, Casilla 567, Valdivia, Chile
[m] Merial GmbH, Kathrinenhof Research Center, Walchenseestr. 8-12, 83101 Rohrdorf, Germany

ARTICLE INFO

Article history:
Received 28 November 2016
Received in revised form 20 March 2017
Accepted 26 April 2017
Available online 27 April 2017

Keywords:
Cystic echinococcosis
Echinococcus granulosus sensu stricto
Echinococcus taxonomy
Hydatidosis
Mitochondrial genome
Nuclear genes

ABSTRACT

Cystic echinococcosis, a zoonotic disease caused by *Echinococcus granulosus* sensu lato (s. l.), is a significant global public health concern. *Echinococcus granulosus* s. l. is currently divided into numerous genotypes (G1–G8 and G10) of which G1–G3 are the most frequently implicated genotypes in human infections. Although it has been suggested that G1–G3 could be regarded as a distinct species *E. granulosus* sensu stricto (s. s.), the evidence to support this is inconclusive. Most importantly, data from nuclear DNA that provide means to investigate the exchange of genetic material between G1–G3 is lacking as none of the published nuclear DNA studies have explicitly included G2 or G3. Moreover, the commonly used relatively short mtDNA sequences, including the complete *cox1* gene, have not allowed unequivocal differentiation of genotypes G1–G3. Therefore, significantly longer mtDNA sequences are required to distinguish these genotypes with confidence. The main aim of this study was to evaluate the phylogenetic relations and taxonomy of genotypes G1–G3 using sequences of nearly complete mitogenomes (11,443 bp) and three nuclear loci (2984 bp). A total of 23 G1–G3 samples were analysed, originating from 5 intermediate host species in 10 countries. The mtDNA data demonstrate that genotypes G1 and G3 are distinct mitochondrial genotypes (separated by 37 mutations), whereas G2 is not a separate genotype or even a monophyletic cluster, but belongs to G3. Nuclear data revealed no genetic separation of G1 and G3, suggesting that these genotypes form a single species due to ongoing gene flow. We conclude that: (a) in the taxonomic sense, genotypes G1 and G3 can be treated as a single species *E. granulosus* s. s.; (b) genotypes G1 and G3 should be regarded as distinct genotypes only in the context of mitochondrial data; (c) we recommend excluding G2 from the genotype list.

1. Introduction

Cystic echinococcosis (CE), a zoonotic disease caused by the larval stage of tapeworm *Echinococcus granulosus* sensu lato (s. l.), is a significant global public health concern (Eckert et al. 2001; Alvarez Rojas et al. 2014; Marcinkute et al. 2015). The life cycle of the parasite involves mainly dogs and wild carnivores as definitive hosts and a wide range of domestic and wild mammals as intermediate hosts, but also humans as aberrant intermediate host (e.g. Eckert et al. 2001; Moks et al. 2006, 2008; Deplazes et al. 2011; Laurimaa et al. 2015a). CE is listed among the most severe parasitic diseases in humans, ranking second in the list of food-borne parasites globally (FAO/WHO report,

* Corresponding author at: Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu 51014, Estonia.
E-mail address: Urmas.Saarma@ut.ee (U. Saarma).

http://dx.doi.org/10.1016/j.meegid.2017.04.023
1567-1348/

2014) and representing one of the 17 Neglected Tropical Diseases prioritised by the World Health Organization (Daumerie et al. 2010).

Echinococcus granulosus s. l. exhibits considerable variability in terms of genetic diversity, morphology, host range, infectivity to humans, pathogenicity, antigenicity, developing rate and other aspects (e.g. Eckert et al. 2001; Thompson, 2008; Gholami et al. 2011; Romig et al. 2015). Molecular studies have identified a number of genotypes/species within the *E. granulosus* complex (Bowles et al. 1992; Bowles et al. 1994; Thompson and McManus, 2002; Lavikainen et al. 2003; Nakao et al. 2007; Hüttner et al. 2008; Thompson, 2008; Saarma et al. 2009; Knapp et al. 2011) that are closely related to other species in the genus *Echinococcus* (Knapp et al. 2015). Traditionally, the complex is considered to consist of genotypes G1–G8, G10 and *E. felidis* (Bowles et al. 1992; Bowles et al. 1994; Lavikainen et al. 2003; Hüttner et al. 2008). A novel genotype recently discovered from Ethiopia by Wassermann et al. (2016) adds even more complexity. Suggestions have been made to split the genotypes into distinct species: *E. granulosus* sensu stricto (s. s.; genotypes G1–G3), *E. equinus* (G4), *E. ortleppi* (G5) and *E. canadensis* (G6–G8 and G10), or *E. intermedius* (G6, G7) and *E. canadensis* (G8, G10) (Thompson and McManus, 2002; Nakao et al. 2007; Thompson, 2008; Saarma et al. 2009; Knapp et al. 2011; Thompson, 2016). However, the evidence is still inconclusive, especially for genotype groups G1–G3 and G6–G10 (e.g. Saarma et al. 2009; Nakao et al. 2015; Romig et al. 2015; Lymbery, 2016).

Echinococcus granulosus s. s. (genotypes G1–G3) is spread globally among wild and domestic animals, with highly endemic foci in South America, the Mediterranean basin and Central Asia (Dakkak et al. 2010; Hajialilo et al. 2012; Rostami et al. 2015; Boufana et al. 2015; Zhang et al. 2015; Cusher et al. 2016; Laurimäe et al. 2016). *E. granulosus* s. s. is also the most frequently implicated *Echinococcus* species in human infections, 88% worldwide, according to a recent estimate by Alvarez Rojas et al. (2014), therefore deserving particularly close attention. Dogs may present a considerable risk factor in the spread of CE due to their close contact with humans. Although generally the infection rates for humans are higher in rural areas (Possenti et al. 2016), there is increasing potential for exposure in urban environment. Indeed, using non-invasive molecular diagnostics (Laurimaa et al. 2015b), genotype G1 has been recently found in urban dogs (Laurimaa et al. 2015a).

The original molecular definitions of genotypes G1–G3 were based on 366 bp of the *cox1* gene (Bowles et al. 1992). It became immediately apparent that genotypes G1–G3 are genetically more closely related to each other than to any other known genotype and a proposal was made to treat G1–G3 as a single species (Thompson et al. 1995). To evaluate the taxonomy of G1–G3, phylogenetic analysis of nuclear loci is absolutely crucial, as it would provide means to investigate the exchange of genetic material between G1–G3. However, nuclear evidence on the inter- and intragenotypic variation of G1–G3 is still missing. Moreover, previous taxonomic studies based on multiple nuclear loci have never explicitly included G2 or G3 (Hüttner et al. 2008; Saarma et al. 2009; Knapp et al. 2011). Thus, despite the assumptions that mitochondrial genotypes G1–G3 can be regarded as a distinct species *E. granulosus* s. s., the evidence is still inconclusive.

An important prerequisite prior to the taxonomic evaluation of *E. granulosus* s. s. is the correct allocation of samples into genotypes. To date, the commonly used relatively short mtDNA sequences in analysis have not been able to clearly differentiate genotypes G1–G3 due to low phylogenetic resolution (Busi et al. 2007; Vural et al. 2008; Casulli et al. 2012; Yanagida et al. 2012; Andresiuk et al. 2013; Romig et al. 2015). As a result, allocation to genotypes G1–G3 has been dubious and without a clear definition. This problem is also highlighted in a recently published network of 137 haplotypes in Romig et al. (2015). Using the 1609 bp sequence of the *cox1* gene, the G1–G3 phylogenetic network revealed a low level of differentiation into G1 and G2/G3, but without clear differentiation into separate haplogroups. Furthermore, a large proportion of the haplotypes described in Romig et al. (2015) were not homologous with any of the sequences of G1, G2 or G3 originally described in Bowles et al. (1992). However, using significantly longer mtDNA sequences (Kinkar et al. 2016; Laurimäe et al. 2016) could provide means to differentiate G1–G3 from each other with confidence. The correct genotyping of this highly zoonotic cluster would also be of great epidemiological importance as it forms the basis for further analysis regarding the possible differences between these genotypes (e.g. infectivity to humans).

Delimiting species is generally a stepwise process (see also Lymbery, 2016). First step is to reconstruct a phylogeny (a phylogenetic tree or network) to identify monophyletic groupings, i.e. organisms that share a common ancestor and where taxa are more closely related to each other than to any other groupings. Second step is to analyse gene flow between these monophyletic groupings in order to identify: (i) groupings for which cohesion is maintained by gene flow, (ii) for which there is no gene flow. If, for example, G1 and G3 represent different mtDNA lineages, whereas on the basis of nDNA data there is no separation of G1 and G3 due to genetic exchange, then one can consider G1 and G3 as a single species. From the taxonomic point of view, the analysis should ideally involve genotypes from various geographic locations, including sympatric and geographically distant ones. The results based on genetic data should also be evaluated in the context of other scientific evidence (e.g. morphological and ecological data, host range, infectivity to humans, developing rate), if such data is available.

The aim of this study was to analyse the phylogenetic relations and taxonomic status of *E. granulosus* s. s. genotypes G1–G3 using 11,443 bp of mtDNA (~85% of the whole mitogenome) and 2984 bp of three nuclear genes (*cal*, *tgf*, *ef1*) analysing samples from different host species covering a wide geographical range, but also from areas where *E. granulosus* s. s. genotypes occur in sympatry.

2. Materials and methods

2.1. Parasite material

A total of 23 *E. granulosus* s. s. samples were analysed in this study, originating from 5 intermediate host species (sheep, cattle, buffalo, camel and human) in 10 countries: India (n = 2), Iran (n = 4), Turkey (n = 4), Spain (n = 5), France (n = 3), Finland (patient from Algeria; n = 1), Chile (n = 1), Argentina (n = 1), Albania (n = 1) and Tunisia (n = 1; Fig. 1; Table 1). In addition, one *E. equinus* (G4) sample from a Turkish donkey and three *E. ortleppi* (G5) samples from Indian buffaloes were analysed to evaluate their genetic distance from genotypes G1–G3. Samples were obtained during routine meat inspections or from hospital cases and were ethanol-preserved at −20 °C until further use.

2.2. DNA extraction

DNA was extracted from protoscoleces or cyst membranes using High Pure PCR Template Preparation Kit (Roche Diagnostics, Mannheim, Germany), following the manufacturer's protocols.

2.3. PCR amplification and sequencing of mtDNA

For mtDNA sequencing we used 17 primers described in Kinkar et al. (2016) and Laurimäe et al. (2016), whereas 7 primers were newly designed (Table 2). Sequencing was performed using the same primers as for the initial PCR amplification. Cycle parameters for PCR and sequencing were as described in Kinkar et al. (2016). All mtDNA sequences were deposited in GenBank and are available under accession numbers KY766882-KY766908 (Table 1).

2.4. PCR amplification and sequencing of nuclear DNA

Amplification and sequencing of 3 nuclear genes (2984 bp in total): transforming growth factor beta receptor kinase (*tgf*; 937 bp), calreticulin (*cal*; 1272 bp) and elongation factor 1 alpha (*ef1*; 775 bp)

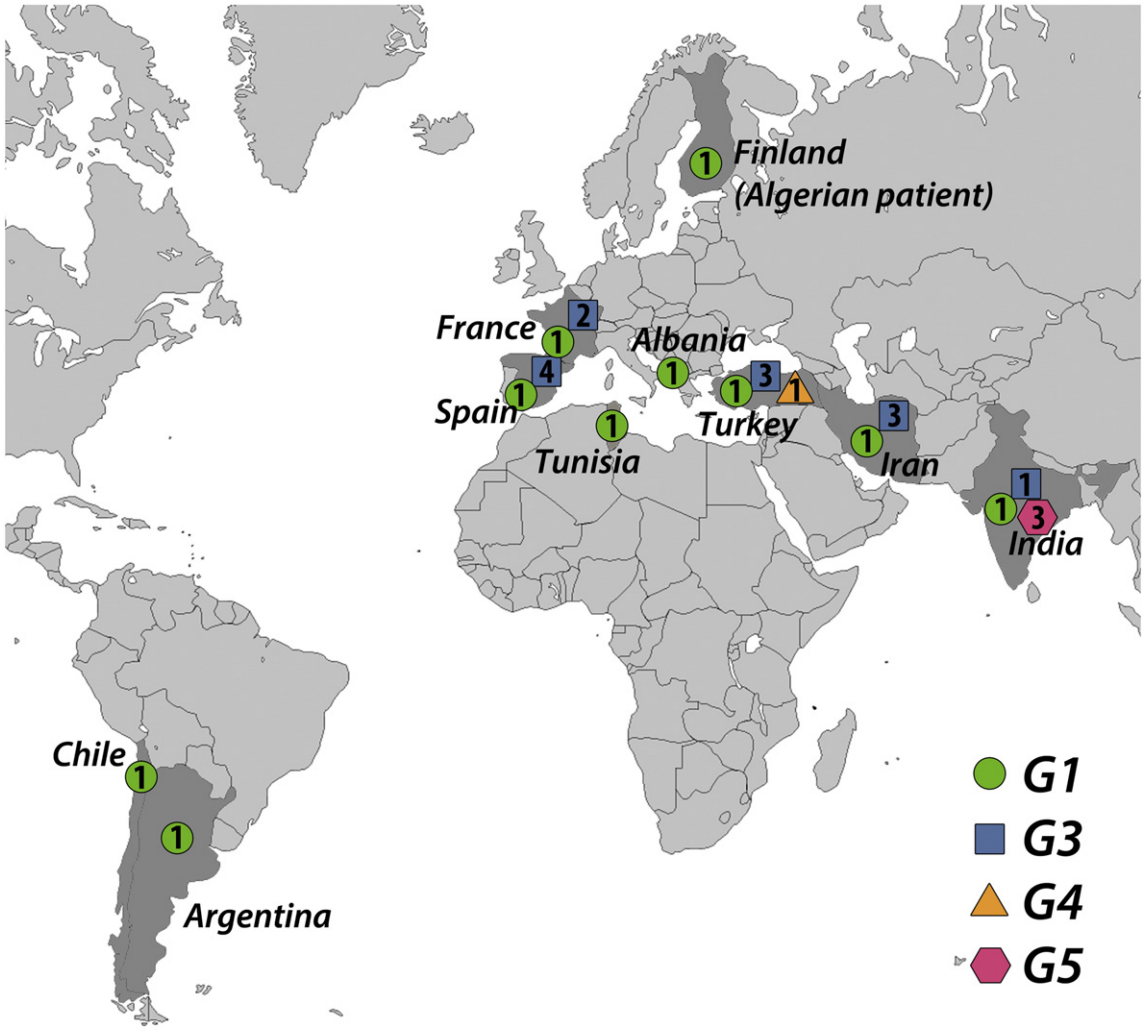

Fig. 1. Geographic locations of the analysed samples: *Echinococcus granulosus* s. s. genotype G1 (n = 10; green circles) and genotype G3 (n = 13; blue squares), *E. ortleppi* (*n* = 3; pink hexagon) and *E. equinus* (n = 1; yellow triangle).

Table 1
Data on the *E. granulosus* s. s. samples analysed in this study.

mtDNA haplotype	Mitochondrial genotype confirmed in this study	Location	Host	mtDNA accession nr in GenBank
SPA5	G1	Spain	Sheep	KY766886
FIN1	G1	Finland	Human	KY766884
TUN1	G1	Tunesia	Sheep	KY766885
ARG1	G1	Argentina	Cattle	KY766882
IND2	G1	India	Buffalo	KY766891
CHI1	G1	Chile	Cattle	KY766890
TUR4	G1	Turkey	Sheep	KY766888
FRA3	G1	France	Cattle	KY766889
ALB1	G1	Albania	Sheep	KY766883
IRA4	G1	Iran	Sheep	KY766887
TUR1	G3	Turkey	Sheep	KY766901
SPA1	G3	Spain	Sheep	KY766900
SPA2	G3	Spain	Sheep	KY766896
FRA1	G3	France	Sheep	KY766893
SPA4	G3	Spain	Sheep	KY766897
IND1	G3	India	Buffalo	KY766902
IRA3	G3	Iran	Camel	KY766899
IRA1	G3	Iran	Camel	KY766894
IRA2	G3	Iran	Camel	KY766895
FRA2	G3	France	Sheep	KY766892
SPA3	G3	Spain	Sheep	KY766903
TUR2	G3	Turkey	Cattle	KY766904
TUR3	G3	Turkey	Sheep	KY766898

was carried out according to Saarma et al. (2009). All nuclear sequences were deposited in GenBank and are available under accession numbers KY766909-KY766920.

2.5. Data analysis

Sequences were assembled in CodonCode v6.0.2, manually corrected in BioEdit v7.2.5 and aligned with corresponding sequences available in GenBank (G1: AF297617; G3: KJ559023; G4: AB786665; G5: AB235846; *cal*: EU834931; *ef1*: EU834898; *tgf*: EU834910) (Le et al. 2002; Nakao et al. 2007; Nakao et al. 2013; Saarma et al. 2009; Wang et al. 2016). Phylogenetic networks were calculated using Network v4.612 (Bandelt et al. 1999) (http://www.fluxusengineering.com, Fluxus Technology Ltd., 2004). Networks were constructed separately for mitochondrial and nuclear markers.

3. Results

In the mtDNA analysis, a total of 23 *E. granulosus* s. s., one *E. equinus* and three *E. ortleppi* samples were successfully analysed, yielding final mtDNA alignment of 11,502 bp (the sequence length was 11,442–11,443 bp for the *E. granulosus* s. s. samples, 11,465 bp for the *E. equinus* sample, and 11,466 bp for the *E. ortleppi* samples).

Nuclear markers *cal*, *ef1* and *tgf* were successfully PCR-amplified for the same set of samples, except for a putative G2 genotype from Spain

Table 2
Primers used for analysis; positions are according to AF297617 in GenBank (Le et al. 2002).

Primer	Primer sequence	Primer position	PCR product length	Reference
Ef1	TCGTTTTACACGCGATTGAACT	4924…4945		[1][a], [2][a]
Er1	ACCTGCTATGCAGCCCTATT	6147…6166	1243 bp	[1], [2]
E2fn	GATGCTGTTAACTTCAAGAAATG	6034…6056		[1], [2]
E2r2	CTCAAAGCATTCAAACGC	7053…7070	1037 bp	[1], [2]
E3fn	GTTGATTCGTGTTAATTTTTTGGAG	6873…6897		[1], [2]
E3rn	GAAAACATAGCAAACAACAACCC	7573…7595	723 bp	[1], [2]
E4f2	GTGATCCTATTTTATTTCAAC	7436…7456		[1], [2]
E4rn	GCTACCTTTGCACAGTCAATATAC	8939…8962	1527 bp	This study
E5fn	GGTACCTAGTTTTTGTTATATTGT	8712…8735		This study
E5rn	GAATCGCTCACTGCCAAACCA	9813…9833	1122 bp	This study
E6f	TAAGGGTGATGCAATTTGAG	9588…9607		[1], [2]
E6r	ACAACCATCTACAGCACGAA	10,812…10,831	1244 bp	[1], [2]
E7fn	GATGCTGTTCCTGGTCGTCTTAATC	10,623…10,647		This study
E7rn	CAATCAACTTCAACAACATAAACC	11,992…12,015	1393 bp	This study
E8fn	GCTTATGTTACGGCCATAAGA	11,716…11,736		This study
E8rn	TGCTTAGTAAAAAACACCCCA	12,764…12,784	1069 bp	This study
E10f	GATTACTGTTACTGGTTTTCA	312…332		[1], [2]
E10r	CAACTTAAAAACAAGCATCATCA	1756…1778	1467 bp	[1], [2]
E11f	TTTTATGCTATTCTTCGGTGTA	1521…1542		[1], [2]
E11r	CAAAAACACCTCATTAAACCAC	3278…3299	1779 bp	[1], [2]
E12f	TTGTGGTGTTTTTATGATG	2922…2940		[1], [2]
E12r	CACAGACGATAACCCAGA	4204…4221	1300 bp	[1], [2]
E13f	CGGGTCTTTTATTTTGATGTTG	4003…4024		[1], [2]
E13r	GATCCAAAAGCACATCGA	5507…5524	1522 bp	[1], [2]

[1] Kinkar et al. (2016).
[2] Laurimäe et al. (2016).
[a] Note that these publications report identical primers due to simultaneous publication.

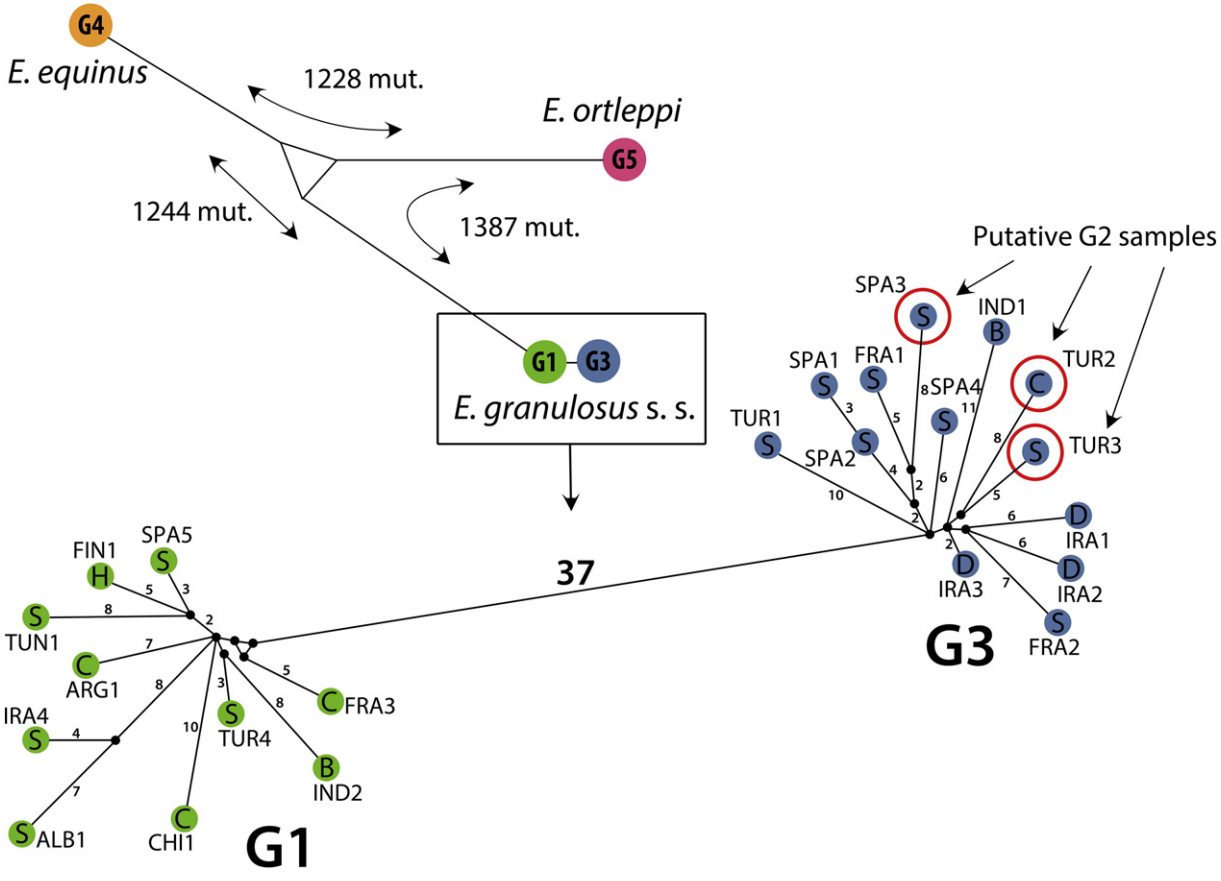

Fig. 2. Phylogenetic network of *E. equinus* (n = 1), *E. ortleppi* (n = 3) and *E. granulosus* s. s. (n = 23) based on 11,502 bp of mtDNA, portraying the *E. granulosus* s. s. genotypes G1 and G3 more closely below. Circles represent haplotypes, different colours represent different genotypes/species: *E. equinus* (G4) – orange circle, *E. ortleppi* (G5) – pink circle, *E. granulosus* s. s. G1 – green circles, *E. granulosus* s. s. G3 – blue circles. Black dots are median vectors (i.e. hypothetical haplotypes: haplotypes not sampled or extinct). The *E. ortleppi* samples were divided into three haplotypes, but were closely related and are depicted as a single circle. *E. granulosus* s. s. G1 and G3 haplotype names represent their geographical origin (TUR – Turkey, SPA – Spain, FRA – France, IND – India, IRA – Iran, CHI – Chile, ALB – Albania, ARG – Argentina, TUN – Tunisia, FIN – Finland (patient from Algeria). Numbers on the lines represent the number of mutations. Letters inside the *E. granulosus* s. s. haplotypes represent host species: S – sheep, C – cattle, B – buffalo, H – human, D – dromedary camel.

(SPA3) that did not yield positive PCR results with the nuclear markers. The final length of the nuclear genes in alignment was 2984 bp.

3.1. mtDNA networks

Echinococcus equinus and *E. ortleppi* mtDNA haplotypes were separated from *E. granulosus* s. s. by 1244 and 1387 mutations, respectively. The genetic distance between *E. equinus* and *E. ortleppi* was 1228 mutations (Fig. 2).

Echinococcus granulosus s. s. samples were divided into two haplogroups (Fig. 2). One haplogroup included 10 samples, which were highly homologous with the G1 mitogenome sequence AF297617 in GenBank (Le et al. 2002). Of these, six haplotypes (FIN1, TUN1, ARG1, TUR4, CHI1, IND2) contained the originally described G1 sequence sensu Bowles et al. (1992; 366 bp of *cox1*) (Table 1). The other haplogroup included 13 samples, which were highly homologous with the G3 mitogenome sequence KJ559023 in GenBank (Wang et al. 2016). Seven haplotypes (SPA1, SPA2, SPA4, IRA1, IRA2, IRA3, FRA2) in this haplogroup contained the originally described G3 sequences sensu 366 bp of *cox1* in Bowles et al. (1992) (Table 1). Therefore these two haplogroups corresponded to the *E. granulosus* s. s. mitochondrial genotypes G1 and G3 and were named accordingly. The G1 and G3 haplogroups were separated by 37 mutations. Three samples (SPA3, TUR2, TUR3) that corresponded to genotype G2 according to the 366 bp of *cox1* in Bowles et al. (1992), positioned inside the G3 cluster (Table 1; Fig. 2), but were not monophyletic. However, seven haplotypes (SPA5, FRA3, ALB1, IRA4, TUR1, FRA1, IND1) could not be genotyped according to the original molecular definition in Bowles et al. (1992) based on the 366 bp fragment of the *cox1* gene (Table 1).

3.2. nDNA networks

The analysed 26 sequences based on the 3 nuclear genes were divided into 4 distinct sequences (Fig. 3). *Echinococcus granulosus* s. s. samples (n = 22) comprised of 2 sequences, separated by a single mutation. One sequence was dominant, comprising 20 *E. granulosus* s. s. samples, whereas the other included only 2 samples (FRA3 and SPA5). The three analysed *E. ortleppi* samples had an identical nuclear sequence, separated from *E. granulosus* s. s. by 36 mutations. The *E. equinus* sample was separated from *E. ortleppi* and *E. granulosus* s. s. by 23 and 45 mutations, respectively.

4. Discussion

Sequencing a large portion of the mtDNA in analysis is highly recommended when the aim is to obtain high-resolution phylogeny (e.g. Keis et al. 2013; Kinkar et al. 2016; Laurimäe et al. 2016). The results in this study based on the nearly complete mitochondrial genome sequences clearly demonstrate that genotypes G1 and G3 form distinct mitochondrial haplogroups, separated from each other by 37 mutations (Fig. 2). However, not all of these are diagnostic and defining the precise set of diagnostic nucleotides requires a much larger sample size (ongoing project).

To date, samples have often been allocated to genotypes G1–G3 without a clear definition: many haplotypes cannot be unequivocally designated to any of the genotypes originally described in Bowles et al. (1992). Even the analysis of the full *cox1* gene (1609 bp) has not allowed clear differentiation of genotypes G1–G3 (see Romig et al. 2015). In this study, the considerably longer mtDNA sequences (11,443 bp) have placed G1 and G3 into distinct haplogroups, corresponding to mitochondrial genotypes G1 and G3. Thus, sequencing a significant portion of the mitochondrial genome has allowed for the first time to differentiate genotypes G1 and G3 with confidence. It is important to note that the G1 and G3 samples in this study were obtained not only from a wide geographical range, but also from countries where they exist in sympatry: India, Iran, Turkey, Spain and France (Fig. 1). Also, several host species analysed in this study were common between genotypes G1 and G3 (sheep, cattle and buffalo, see Fig. 2). Thus, the separation of these groups cannot be explained by clustering according to geographical origin or host species. Due to the relatively small number of samples analysed in this study, only a portion of the mitochondrial variation of genotypes G1 and G3 is presented. Therefore, it is possible that future studies involving significantly more samples may reveal haplotypes that position between G1 and G3 in mtDNA-based phylogenetic networks. However, these cases are probably rare since our analysis that included G1 and G3 samples from both geographically overlapping and highly distant locations demonstrate that maternal lineages of G1 and G3 are highly divergent and cluster separately (Fig. 2).

The results derived from mitochondrial data do not necessarily mean that genotypes G1 and G3 are separate biological entities. MtDNA does not recombine and mutations to different mtDNA lineages accumulate at random. Once a new mitochondrial mutation becomes fixed in a population, it forms a new mitochondrial lineage that is separate from the ancestral one. From this point onwards, mutations continue to fix progressively in an independent manner in both the new and

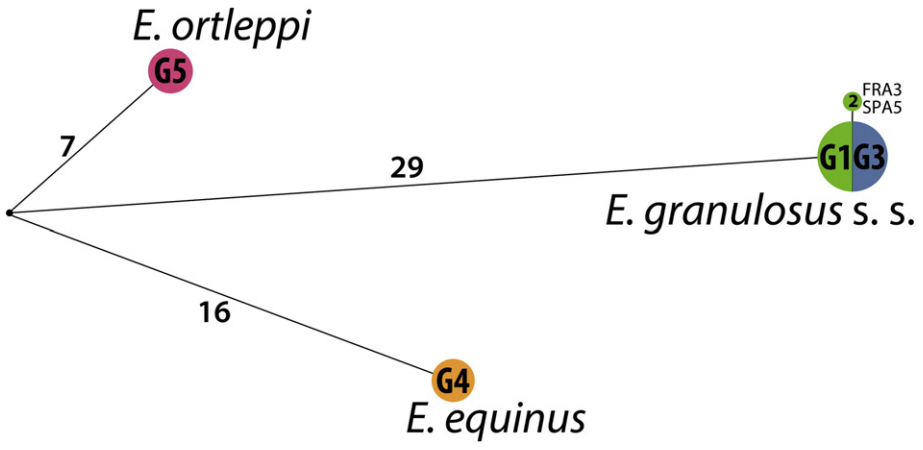

Fig. 3. Phylogenetic network of *E. equinus* (n = 1), *E. ortleppi* (n = 3) and *E. granulosus* s. s. (n = 22) based on 2984 bp of nuclear DNA. Circles represent distinct sequences, different colours represent different genotypes/species: *E. equinus* – orange, *E. ortleppi* – pink, *E. granulosus* s. s. G1 – green, *E. granulosus* s. s. G3 – blue. Since the *E. granulosus* s. s. dominant sequence contained samples from both G1 and G3 (n = 20), it is represented with half green, half blue. The black dot is a median vector. Numbers on the lines represent the number of mutations. The number inside the sequence FRA3, SPA5 represents the number of samples.

ancestral mitochondrial lineages. Although mutations also accumulate into the nuclear genome at random, nuclear genes undergo recombination and if there is no barrier for gene-flow, then nuclear genes do not show separation into genetically distinct populations (Saarma et al. 2009). Indeed, our data based on three nuclear genes enabled to distinguish *E. granulosus* s. s., *E. equinus* and *E. ortleppi* from each other with confidence, whereas there was no separation between genotypes G1 and G3 (Fig. 3). Therefore, our study confirms that in the taxonomic sense G1 and G3 can be regarded as a single species *E. granulosus* sensu stricto, which is further supported by a notion that G1 and G3 both have a wide geographical overlap and similar host spectra; hence there are limited ecological differences between these genotypes. Thus, this is the first study that confirms the species status of *E. granulosus* s. s. Also, as these results were unambiguous, we conclude that the sample size in this study was sufficient to confirm the species status of *E. granulosus* s. s. If the results had suggested that genotypes G1 and G3 are different species, contradicting earlier assumptions, then more samples would have been needed to confirm this finding.

Our data suggests that G2 is not a valid genotype even in the mitochondrial context. Three samples, matching the original molecular definition of genotype G2 sensu 366 bp of *cox1* (Bowles et al. 1992), clustered together with G3 samples based on mtDNA and with both G1 and G3 genotypes based on nuclear genes (Table 1; Fig. 2; Fig. 3). Moreover, the putative G2 samples in this study were not monophyletic. We therefore suggest excluding G2 from the genotype list. Although the G2 genotype was originally described from a sheep in Tasmania (Bowles et al. 1992), a location which was not represented in this study, all analysed G2 samples were, sensu Bowles et al. (1992), genetically identical to the originally described Tasmanian sheep samples. Therefore, the G2 samples analysed in this study were adequate for investigating the validity of this genotype.

We conclude that in the taxonomic sense, genotypes G1 and G3 should be treated as single species *E. granulosus* s. s. It is also important to note that G1 and G3 can be regarded as distinct genotypes only in the context of mitochondrial data and that G2 is not a valid genotype even in the mitochondrial context. Although possible differences in the epidemiology between genotypes G1 and G3 are largely unknown and remain to be studied in the future, applying up-to-date molecular diagnostics to separate genotypes G1 and G3 correctly is an important prerequisite to perform such studies.

Conflict of interest

Authors declare no conflict of interest.

Acknowledgements

This work was supported by institutional research funding (IUT20-32) from the Estonian Ministry of Education and Research; the European Union through the European Regional Development Fund (Centre of Excellence FIBIR); the Estonian Doctoral School of Ecology and Environmental Sciences.

References

Alvarez Rojas, C.A., Romig, T., Lightowlers, M.W., 2014. *Echinococcus granulosus* sensu lato genotypes infecting humans - review of current knowledge. Int. J. Parasitol. 44:9–18. http://dx.doi.org/10.1016/j.ijpara.2013.08.008.

Andresiuk, M.V., Gordo, F.P., Saarma, M., Elissondo, M.C., Taraborelli, A., Casalongue, C., Denegri, G., Saarma, U., 2013. *Echinococcus granulosus* genotype G1 dominated in cattle and sheep during 2003–2006 in Buenos Aires province, an endemic area for cystic echinococcosis in Argentina. Acta Trop. 127:136–142. http://dx.doi.org/10.1016/j.actatropica.2013.04.008.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Boufana, B., Lett, W.S., Lahmar, S., Buishi, I., Bodell, A.J., Varcasia, A., Casulli, A., Beeching, N.J., Campbell, F., Terlizzo, M., McManus, D.P., Craig, P.S., 2015. *Echinococcus equinus* and *Echinococcus granulosus* sensu stricto from the United Kingdom: genetic diversity and haplotypic variation. Int. J. Parasitol. 45, 161–166.

Bowles, J., Blair, D., McManus, D., 1992. Genetic variants within the genus *Echinococcus* identified by mitochondrial-DNA sequencing. Mol. Biochem. Parasitol. 54:165–174. http://dx.doi.org/10.1016/0166-6851(92)90109-W.

Bowles, J., Blair, D., McManus, D.P., 1994. Molecular genetic characterization of the cervid strain ('northern form') of *Echinococcus granulosus*. Parasitology 109:215–221. http://dx.doi.org/10.1017/S0031182000076332.

Busi, M., Snabel, V., Varcasia, A., Garippa, G., Perrone, V., De Liberato, C., D'Amelio, S., 2007. Genetic variation within and between G1 and G3 genotypes of *Echinococcus granulosus* in Italy revealed by multilocus DNA sequencing. Vet. Parasitol. 150: 75–83. http://dx.doi.org/10.1016/j.vetpar.2007.09.003.

Casulli, A., Interisano, M., Sreter, T., Chitimia, L., Kirkova, Z., La Rosa, G., Pozio, E., 2012. Genetic variability of *Echinococcus granulosus* sensu stricto in Europe inferred by mitochondrial DNA sequences. Infect. Genet. Evol. 12:377–383. http://dx.doi.org/10.1016/j.meegid.2011.12.014.

Cucher, M.A., Macchiaroli, N., Baldi, G., Camicia, F., Prada, L., Maldonado, L., Avila, H.G., Fox, A., Gutiérrez, A., Negro, P., Lopez, R., Jensen, O., Rosenzvit, M., Kamenetzky, L., 2016. Cystic echinococcosis in South America: systematic review of species and genotypes of *Echinococcus granulosus* sensu lato in humans and natural domestic hosts. Tropical Med. Int. Health 21:166–175. http://dx.doi.org/10.1111/tmi.12647.

Dakkak, A., 2010. Echinococcosis/hydatidosis: a severe threat in Mediterranean countries. Vet. Parasitol. 174:2–11. http://dx.doi.org/10.1016/j.vetpar.2010.08.009.

Daumerie, D., Savioli, L., Crompton, D.W.T., Peters, P., 2010. Working to Overcome the Global Impact of Neglected Tropical Diseases: First WHO Report on Neglected Tropical Diseases. World Health Organization, France.

Deplazes, P., van Knapen, F., Schweiger, A., Overgaauw, P.A., 2011. Role of pet dogs and cats in the transmission of helminthic zoonoses in Europe, with a focus on echinococcosis and toxocarosis. Vet. Parasitol. 182:41–53. http://dx.doi.org/10.1016/j.vetpar.2011.07.014.

Eckert, J., Gemmell, M.A., Meslin, F.-X., Pawlowski, Z.S., 2001. WHO/OIE Manual on Echinococcosis in Humans and Animals: A Public Health Problem of Global Concern. WHO/OIE, Paris:pp. 72–99. http://apps.who.int/iris/bitstream/10665/42427/1/929044522X.pdf.

FAO/WHO [Food and Agriculture Organization of the United Nations/World Health Organization], 2014. Multicriteria-based ranking for Risk Management of Foodborne Parasites. Microbiological Risk Assessment Series No. 23 (Rome. 302 pp.).

Gholami, S., Irshadullah, M., Mobedi, I., 2011. Rostellar hook morphology of larval *Echinococcus granulosus* isolates from the Indian buffalo and Iranian sheep, cattle and camel. J. Helminthol. 85:239–245. http://dx.doi.org/10.1017/S0022149X10000520.

Hajialilo, M., Fasihi Harandi, M., Sharbatkhori, M., Mirhendi, H., Rostami, S., 2012. Genetic characterization of *Echinococcus granulosus* in camels, cattle and sheep from the south-east of Iran indicates the presence of the G3 genotype. J. Helminthol. 86: 263–270. http://dx.doi.org/10.1017/S0022149X11000320.

Hüttner, M., Nakao, M., Wassermann, T., Siefert, L., Boomker, J.D.F., Dinkel, A., Sako, Y., Mackenstedt, U., Romig, T., Ito, A., 2008. Genetic characterization and phylogenetic position of *Echinococcus felidis* Ortlepp, 1937 (Cestoda: Taeniidae) from the African lion. Int. J. Parasitol. 38:861–868. http://dx.doi.org/10.1016/j.ijpara.2007.10.013.

Keis, M., Remm, J., Ho, S.Y.W., Davison, J., Tammeleht, E., Tumanov, I.L., Saveljev, A.P., Männil, P., Kojola, I., Abramov, A.V., Margus, T., Saarma, U., 2013. Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographic structure and migration patterns among brown bears in north-western Eurasia. J. Biogeogr. 40:915–927. http://dx.doi.org/10.1111/jbi.12043.

Kinkar, L., Laurimäe, T., Simsek, S., Balkaya, I., Casulli, A., Manfredi, M.T., Ponce-Gordo, F., Varcasia, A., Lavikainen, A., Gonzalez, L.M., Rehbein, S., van der Giessen, J., Sprong, H., Saarma, U., 2016. High-resolution phylogeography of zoonotic tapeworm *Echinococcus granulosus* sensu stricto genotype G1 with an emphasis on its distribution in Turkey, Italy and Spain. Parasitology 143:1790–1801. http://dx.doi.org/10.1017/S0031182016001530.

Knapp, J., Gottstein, B., Saarma, U., Millon, L., 2015. Taxonomy, phylogeny and molecular epidemiology of *Echinococcus multilocularis*: from fundamental knowledge to health ecology. Vet. Parasitol. 213:85–91. http://dx.doi.org/10.1016/j.vetpar.2015.07.030.

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. http://dx.doi.org/10.1016/j.ympev.2011.07.022.

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

Laurimaa, L., Davison, J., Süld, K., Plumer, L., Oja, R., Moks, E., Keis, M., Hindrikson, M., Kinkar, L., Laurimäe, T., Abner, J., Remm, J., Anijalg, P., Saarma, U., 2015a. First report of highly pathogenic *Echinococcus granulosus* genotype G1 in dogs in European urban environment. Parasit. Vectors 8:182. http://dx.doi.org/10.1186/s13071-015-0796-3.

Laurimaa, L., Davison, J., Plumer, L., Süld, K., Oja, R., Moks, E., Keis, M., Hindrikson, M., Kinkar, L., Laurimäe, T., Abner, J., Remm, J., Anijalg, P., Saarma, U., 2015b. Non-invasive molecular diagnostics identifies *Echinococcus multilocularis* spillover to an urban area in Estonia. Emerg. Infect. Dis. 21:163–164. http://dx.doi.org/10.3201/eid2101.140136.

Laurimäe, T., Kinkar, L., Andresiuk, V., Haag, K.L., Ponce-Gordo, F., Acosta-Jamett, G., Garate, T., Gonzalez, L.M., Saarma, U., 2016. Genetic diversity and phylogeography of highly zoonotic *Echinococcus granulosus* genotype G1 in the Americas (Argentina, Brazil, Chile and Mexico) based on 8279 bp of mtDNA. Infect. Genet. Evol. 45: 290–296. http://dx.doi.org/10.1016/j.meegid.2016.09.015.

Le, T.H., Pearson, M.S., Blair, D., Dai, N., Zhang, L.H., 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. http://dx.doi.org/10.1017/S0031182001008976.

Lymbery, A.J., 2016. Phylogenetic pattern, evolutionary processes and species delimitation in the genus *Echinococcus*. Adv. Parasitol. http://dx.doi.org/10.1016/bs.apar.2016.07.002.

Marcinkute, A., Šarkunas, M., Moks, E., Saarma, U., Jokelainen, P., Bagrade, G., Laivacuma, S., Strupas, K., Sokolovas, V., Deplazes, P., 2015. *Echinococcus* infections in the Baltic region. Vet. Parasitol. 213:121–131. http://dx.doi.org/10.1016/j.vetpar.2015.07.032.

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

Moks, E., Jõgisalu, I., Valdmann, H., 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. http://dx.doi.org/10.1017/S0031182008004198.

Nakao, M., Lavikainen, A., Hoberg, E., 2015. Is *Echinococcus intermedius* a valid species? Trends Parasitol. 31:342–343. http://dx.doi.org/10.1016/j.pt.2015.04.012.

Nakao, M., McManus, D.P., Schantz, P.M., Craig, P.S., Ito, A., 2007. A molecular phylogeny of the genus *Echinococcus* inferred from complete mitochondrial genomes. Parasitology 134:713–722. http://dx.doi.org/10.1017/S0031182006001934.

Nakao, M., Yanagida, T., Konyaev, S., Lavikainen, A., Odnokurtsev, V.A., Zaikov, V.A., Ito, A., 2013. Mitochondrial phylogeny of the genus Echinococcus (Cestoda: Taeniidae) with emphasis on relationships among *Echinococcus canadensis* genotypes. Parasitology 140:1625–1636. http://dx.doi.org/10.1017/S0031182013000565.

Possenti, A., Manzano-Román, R., Sánchez-Ovejero, C., Boufana, B., La Torre, G., Siles-Lucas, M., Casulli, A., 2016. Potential risk factors associated with human cystic echinococcosis: systematic review and meta-analysis. PLoS Negl. Trop. Dis. 10, e0005114. http://dx.doi.org/10.1371/journal.pntd.0005114.

Romig, T., Ebi, D., Wassermann, M., 2015. Taxonomy and molecular epidemiology of Echinococcus granulosus sensu lato. Vet. Parasitol. 213:76–84. http://dx.doi.org/10.1016/j.vetpar.2015.07.035.

Rostami, S., Shariat Torbaghan, S., Dabiri, S., Babaei, Z., Mohammadi, M.A., Sharbatkhori, M., Fasihi Harandi, M., 2015. Genetic characterization of *Echinococcus granulosus* from a large number of formalin-fixed, paraffin-embedded tissue. Am.J.Trop. Med. Hyg. 92:588–594. http://dx.doi.org/10.4269/ajtmh.14-0585.

Saarma, U., Jogisalu, I., Moks, E., Varcasia, A., Lavikainen, A., Oksanen, A., Simsek, S., Andresiuk, V., Denegri, G., Gonzalez, L.M., Ferrer, E., Garate, T., Rinaldi, L., Maravilla, P., 2009. A novel phylogeny for the genus *Echinococcus*, based on nuclear data, challenges relationships based on mitochondrial evidence. Parasitology 136:317–328. http://dx.doi.org/10.1017/S0031182008005453.

Thompson, R.C.A., 2008. The taxonomy, phylogeny and transmission of *Echinococcus*. Exp. Parasitol. 119:439–446. http://dx.doi.org/10.1016/j.exppara.2008.04.016.

Thompson, R.C.A., 2016. Biology and Systematics of *Echinococcus*. Adv. Parasitol. http://dx.doi.org/10.1016/bs.apar.2016.07.001.

Thompson, R.C., Lymbery, A.J., Constantine, C.C., 1995. Variation in *Echinococcus*: towards a taxonomic revision of the genus. Adv. Parasitol. 35, 145–176.

Thompson, R.C.A., McManus, D.P., 2002. Towards a taxonomic revision of the genus *Echinococcus*. Trends Parasitol. 18:452–457. http://dx.doi.org/10.1016/S1471-4922(02)02358-9.

Vural, G., Baca, A.U., Gauci, C.G., Bagci, O., Gicik, Y., Lightowlers, M.W., 2008. Variability in the *Echinococcus granulosus* cytochrome c oxidase 1 mitochondrial gene sequence from livestock in Turkey and a re-appraisal of the G1–3 genotype cluster. Vet. Parasitol. 154:347–350. http://dx.doi.org/10.1016/j.vetpar.2008.03.020.

Wang, N., Xie, Y., Liu, T., Zhong, X., Wang, J., Hu, D., Wang, S., Gu, X., Peng, X., Yang, G., 2016. The complete mitochondrial genome of G3 genotype of *Echinococcus granulosus* (Cestoda: Taeniidae). Mitochondrial DNA 27:1701–1702. http://dx.doi.org/10.3109/19401736.2014.961129.

Wassermann, M., Woldeyes, D., Gerbi, B.M., Ebi, D., Zeyhle, E., Mackenstedt, U., Petros, B., Tilahun, G., Kern, P., Romig, T., 2016. A novel zoonotic genotype related to *Echinococcus granulosus* sensu stricto from southern Ethiopia. Int. J. Parasitol. 46:663–668. http://dx.doi.org/10.1016/j.ijpara.2016.04.005.

Yanagida, T., Mohammadzadeh, T., Kamhawi, S., Nakao, M., Sadjjadi, S.M., Hijjawi, N., Abdel-Hafez, S.K., Sako, Y., Okamoto, M., Ito, A., 2012. Genetic polymorphisms of *Echinococcus granulosus* sensu stricto in the Middle East. Parasitol. Int. 61:599–603. http://dx.doi.org/10.1016/j.parint.2012.05.014.

Zhang, W., Zhang, Z., Wu, W., Shi, B., Li, J., Zhou, X., Wen, H., McManus, D.P., 2015. Epidemiology and control of echinococcosis in central Asia, with particular reference to the People's Republic of China. Acta Trop. 141:235–243. http://dx.doi.org/10.1016/j.actatropica.2014.03.014.