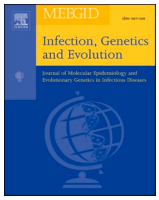
Infection, Genetics and Evolution 98 (2022) 105223

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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



Genetic structure and phylogeography of *Echinococcus granulosus sensu stricto* genotypes G1 and G3 in Pakistan and other regions of the world based on *nad5* gene

Naunain Mehmood [a,*], Hira Muqaddas [a,b], Muhammad Irfan Ullah [c], Urmas Saarma [d], Antonio Varcasia [e]

[a] Department of Zoology, University of Sargodha, Sargodha, Pakistan
[b] Department of Zoology, The Women University Multan, Multan, Pakistan
[c] Department of Entomology, University of Sargodha, Sargodha, Pakistan
[d] Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia
[e] Department of Veterinary Medicine, University of Sassari, Sassari, Italy

ARTICLE INFO

Keywords:
Echinococcus granulosus s.s.
G3 genotype
G1 genotype
Haplotypes
nad5
CE

ABSTRACT

Pakistan is a neglected endemic focus for *Echinococcus granulosus sensu lato*, a zoonotic parasite species complex with the ability to infect wide spectrum of hosts. Wide gaps exist in literature for etiological agents of cystic echinococcosis (CE) in Pakistan due to a very low number of studies on identifying the exact genotypes involved in epidemiological manifestation of this disease. Focusing on transmission patterns and epidemiological dynamics, this study aimed at investigating infective genotypes among the cattle population of south Punjab, Pakistan, employing a mitochondrial marker *nad5* (680 bp). Nucleotide sequences retrieved from 28 hydatid cyst isolates displayed considerable intraspecific variation revealing the existence of G3 and G1 strains of *Echinococcus granulosus sensu stricto*. The G3 genotype emerged as the predominant cause (78.57%) of hydatidosis in cattle. Apart from this, to understand phylogeographical relations, homologous nucleotide sequences of the partial *nad5* gene from six major regions of the world were employed in the population genetics analysis to have an insight into genetic variability and demographics of G3 genotype in particular. Diversification of G3 and its haplotypes in Pakistan ($n = 11$) and other regions of the world (India, Iran, Turkey, Italy and France) was demonstrated. It was further demonstrated that the South Asian population (Pakistan and India) was highly differentiated from the other regions. It could, therefore, be speculated that G3 is diverging and expanding its population with South Asia as the main focal point.

1. Introduction

Cystic echinococcosis (CE) caused by *Echinococcus granulosus sensu lato* (*s.l.*) is listed among the most severe parasitic infections of humans (Deplazes et al., 2017; Kinkar et al., 2018a). Having cyclozoonotic pattern between different mammalian hosts, this cestode utilizes canids (usually dogs) as definitive host and a wide variety of omnivorous and herbivorous ungulates as intermediate host (Eckert et al., 2001). The adult worm lives inside the dog intestines, releasing eggs, which have the capability of causing serious infections in the intermediate hosts, that ingest these eggs while grazing (Thompson, 2017). In intermediate host, the parasite develops fluid filled cysts called hydatids in the visceral organs, most frequently in liver and lungs. Humans can also be infected *via* contact with infected definitive host or by ingesting eggs of contaminated fruits, vegetables or drinking water. However, humans are usually considered the dead-end hosts and do not contribute in perpetuation of the natural life cycle of the tapeworm (Eckert et al., 2001; Rojas et al., 2014; Thompson, 2017).

An assemblage of different mitochondrial genotypes, *E. granulosus s.l.* differs considerably in its pathogenicity, infectivity, host adaptability, life cycle patterns and other biological attributes displaying significant intraspecific variability (Eckert et al., 2001). Different molecular approaches have outlined five major taxa of *E. granulosus s.l.* with some significant taxonomic uncertainties. G1 and G3 strains are regarded as

* Corresponding author at: Department of Zoology, University of Sargodha, Sargodha 40100, Pakistan.
E-mail address: naunain.mahmood@uos.edu.pk (N. Mehmood).

https://doi.org/10.1016/j.meegid.2022.105223
Received 18 July 2021; Received in revised form 25 December 2021; Accepted 19 January 2022
Available online 29 January 2022

E. granulosus sensu stricto (*s.s.*), G4 as *E. equinus*, G5 as *E. ortleppi*, whereas the taxonomic status of the G6-G8/ G10 genotypic group is still controversial (Laurimäe et al., 2018). *E. felidis*, identified from lions in Africa, is also the part of this species complex (Lymbery et al., 2015; Nakao et al., 2015; Thompson, 2020). Note that the genotypes G2 (Kinkar et al., 2017) and G9 (Kedra et al., 1999) are no longer valid.

The tapeworm *E. granulosus s.s.* is the most widespread species among *E. granulosus s.l.* and is highly endemic to Asia, South America and Mediterranean basin (Deplazes et al., 2017; Thompson, 2017). The biological variants of *E. granulosus s.s.* include G1 (sheep strain) and G3 (buffalo strain) occurring sympatrically across wide geographical range. There are fixed genetic differences for these genotypes at the mitochondrial level (Kinkar et al., 2018a), however, no such pattern occurs for nuclear genes (Kinkar et al., 2017) due to which both genotypes are grouped as a single species. Despite that, these genotypes have rather different distributional patterns and host range. Moreover, the studies pertaining to morphometric measurements of the protoscolices (Cengiz and Gonenc, 2020; Pednekar et al., 2009; Shariatzadeh et al., 2015), organ predilection in the intermediate hosts (Muqaddas et al., 2020; Sharma et al., 2013) and cyst fertility (Pednekar et al., 2009) suggest biological relevance of these genotypes. A recent study has found that the large hook width of the protoscolex of these genotypes differs significantly (Cengiz and Gonenc, 2020), whereas, other morphological parameters for large and small hooks are similar in G1 and G3 (Pednekar et al., 2009; Shariatzadeh et al., 2015). Most commonly used genetic markers (*cox1* and *nad1*) aimed at CE etiology in different regions of the world provide less and sometimes insufficient number of diagnostic positions for differentiating *E. granulosus s.s.* genotypes (Kinkar et al., 2018b). Encountering intermediate haplotypes of G1/G3 is also common among these gene markers and therefore, limit our knowledge of genotype spread and epidemiology. Moreover, due to lower global burden of CE from G3, this genotype is mostly grouped together with G1 as *E. granulosus s.s.* and no genotypic distinction is made. However, recently partial mitochondrial *nad5* gene (680 bp) is proposed as a reliable genetic marker for correct identification of these genotypes as it has six potential diagnostic positions which differ among these genotypes (Kinkar et al., 2018b). Utilization of this gene marker in future studies will greatly improve our understanding of epidemiology, spatial dynamics and infectivity of G3 genotype. So far, very few reports are published on the molecular confirmation of G1 and G3 genotypes by the partial *nad5* gene (Bonelli et al., 2021; Kinkar et al., 2017, 2018a, 2018b, 2018c; Laurimäe et al., 2019; Ohiolei et al., 2019a, 2019b; Soba et al., 2020) but with more incoming data, the phylogeographical patterns for G3 genotype would be greatly enhanced.

Pakistan is an endemic area for CE (Mehmood et al., 2020a) with highest prevalence rates for 8 neglected tropical diseases (NTDs; Hotez et al., 2017). Recently, new studies have reported an unusually high percentage of involvement of G3 genotype among livestock (Alvi et al., 2020; Mehmood et al., 2020b) and humans (Muqaddas et al., 2020) in Pakistan. To have better understanding of epidemiological and transmission dynamics of *E. granulosus s.s.*, current study was undertaken from southern Punjab, Pakistan employing partial mitochondrial *nad5* gene for proper discrimination of the infective genotypes among the cattle metacestodes. Moreover, the phylogeography and gene flow patterns of *E. granulosus s.s.* across Pakistan and the global population were undertaken using the *nad5* gene marker which is the best region for discrimination between G1 and G3 genotypes.

2. Methodology

2.1. Study area

Pakistan is an agrarian country and millions of people are associated with agriculture and its different sub-sectors. Livestock, one of the major sectors of agriculture, forms the backbone of rural economy of Pakistan (PES, Pakistan Economic Survey, 2020–21). Having prominent pastoral regions, terrain and climate of Pakistan support extensive livestock farming contributing to endemicity of CE (Mehmood et al., 2020a). Punjab is the largest province of Pakistan, having 65% of the total number of buffaloes in the country, followed by cattle (49%), goats (37%) and sheep (24%) (PBS, Pakistan Bureau of Statistics, 2006). The highest provincial population of cattle in Punjab is found in the southern part of province (LDD, 2018).

2.2. Genetic analysis of Pakistani isolates

2.2.1. Sampling, DNA extraction, amplification and sequencing

Sampling for hydatid cyst isolates was done from the slaughterhouses and butcher shops of Multan, south Punjab, Pakistan for a period of two months (June 2020–July 2020). A total of 28 hydatid cysts were collected from cattle and were processed for molecular characterization through the partial mitochondrial gene marker (*nad5*). Genotypic assessment of the isolates involved initial extraction of genomic DNA from the germinal layer using GeneJET genomic DNA purification kit (Thermoscientific, USA) following the manufacturer's protocol. Fragment amplification (759 bp) was carried out using EGnad5F1 and EGnad5R1 primer pair (Kinkar et al., 2018b) with few modifications in PCR cycling conditions (initial denaturation at 95 °C for 5 min followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C for 45 s, elongation at 72 °C for 1 min and a final extension at 72 °C for 3 min). PCR products were purified using GeneJET PCR purification kit (Thermoscientific, USA) and then subjected to commercial bidirectional sequencing by using the same primers as for PCR.

2.3. Data analysis

Chromatograms were viewed on FinchTV viewer (Geospiza, Seattle, WA, USA). DNA sequences for partial *nad5* gene were subjected to multiple alignment along with *E. granulosus s.s* reference sequences of G1 (AB786664; Nakao et al., 2013) and G3 genotypes (KJ559023; Wang et al., 2016). The best nucleotide substitution model was selected in MEGA X software (Kumar et al., 2018) based on Baeysian information criteria (BIC) and Akike information criteria, corrected (AICc) for computation of inter and intraspecific phylogenies. A maximum likelihood tree was generated using GTR (general time reversible) model keeping rate among sites as gamma distributed with invariant sites (G + I). *E. ortleppi* (Nakao et al., 2007) was used as an outgroup for the construction of phylogenetic tree.

In addition to this, aligned dataset for the *E. granulosus s.s.* sequences were exported to DnaSP 6 (Rozas et al., 2017) to obtain information on haplotypes, their distribution and DNA polymorphism. A haplotype network was generated using PopART software keeping statistical parsimony for the TCS network (Bandelt et al., 1999).

2.4. Population indices for partial nad5 dataset from different regions of the world

All available nucleotide sequences for *E. granulosus s.s.* (*nad5* gene) from six regions of the world (Africa, Europe, South America, Middle East, Central/East Asia and the South Asia) were retrieved from the GenBank database, NCBI (https://www.ncbi.nlm.nih.gov/genbank/). *nad5* mitochondrial nucleotide sequences of *E. granulosus s.s.* from Africa (Algeria $n = 13$, Morocco $n = 1$, Tunisia $n = 38$, Nigeria $n = 1$), Europe (Albania $n = 3$, Bulgaria $n = 1$, Finland $n = 2$, France $n = 4$, Greece $n = 5$, Italy $n = 30$, Moldova $n = 3$, Romania $n = 3$, Spain $n = 20$), South America (Argentina $n = 31$, Brazil $n = 14$,Chile $n = 6$), Middle East (Iran $n = 29$, Turkey $n = 50$), Central/East Asia (China $n = 59$, Mongolia $n = 3$) and South Asia (Pakistan $n = 28$, India $n = 5$) (Nakao et al., 2013; Wang et al., 2016; Kinkar et al., 2018a; Laurimäe et al., 2019; Ohiolei et al., 2019a, 2019b; Bonelli et al., 2021) were used to estimate genetic diversity, neutrality indices and pairwise fixation index (Fst) among the six regions of the world by employing the Arlequin package 3.5

Table 1
Alignment positions of the polymorphic sites within the partial *nad5* gene (mtDNA) haplotypes identified from Pakistan in comparison with the reference sequences for G1 (AB786664; Nakao et al., 2013) and G3 (KJ559023; Wang et al., 2016).

Haplotype	Accession No.	Genotype	*n*	%	36[2]	59[1]	106[3]	124[3]	313[3]	395[1]	396[2]	401[1]	404[1]	454[3]	521[1]	571[3]	634[3]	639[2]	649[3]	658[3]
nd5Pak1	MZ277756	G1	6	21.42	G	A	C	T	C	C	A	G	A	T	T	T	C	G	A	G
nd5Pak2	MZ277757	G3	11	39.28	C*	G	.	.	T	.	.	A	.	.	.	.	.	.	G	A
nd5Pak3	MZ277758	G3	2	7.14	C*	G	.	.	T	.	.	A	.	.	.	.	T	.	G	A
nd5Pak4	MZ277759	G3	1	3.57	C*	G	.	.	T	.	G	A	.	.	.	.	.	.	G	A
nd5Pak5	MZ277760	G3	1	3.57	C*	G	.	.	T	T	.	A	.	.	.	.	.	.	G	A
nd5Pak6	MZ277761	G3	1	3.57	C*	G	T	.	T	.	.	A	.	.	.	C	.	.	G	A
nd5Pak7	MZ277762	G3	1	3.57	C*	G	.	.	T	.	.	A	.	.	.	.	.	.	G	A
nd5Pak8	MZ277763	G3	1	3.57	C*	G	.	.	T	.	.	A	G	.	.	.	.	.	G	A
nd5Pak9	MZ277764	G3	1	3.57	C*	G	.	.	T	.	.	A	.	C	.	.	.	.	G	A
nd5Pak10	MZ277765	G3	1	3.57	C*	G	.	C	T	.	.	A	.	.	.	.	.	.	G	A
nd5Pak11	MZ277766	G3	1	3.57	C*	G	.	.	T	.	.	A	.	.	C	.	.	.	G	A
nd5Pak12	MZ277767	G3	1	3.57	C*	G	.	C	T	.	.	A	.	.	.	.	.	A	G	A

*transversions ($n = 1$), [1] first codon position, [2] second codon position, [3] third codon position, bold letters are non-synonymous mutations.

Fig. 1. Haplotypic profile of *E. granulosus s.s.* isolates retrieved from Pakistan (south Punjab) indicating G1 and G3 genotypes. Hatch marks represent the number of mutations between each haplotype and the size of circle corresponds to the frequency of each haplotype in the population.

(Excoffier and Lischer, 2010). Due to a smaller number of sequences form the North America ($n = 1$) and Australia ($n = 3$), these regions were not included in the population genetics analysis. Moreover, a pairwise fixation index (Fst) (Reynolds et al., 1983) between G1 and G3 genotypes was also estimated for the determination of genetic differentiation. Genealogical relationships among the partial *nad5* gene haplotypes were found among the nucleotide sequences ($n = 349$) retrieved from GenBank database and the haplotype network was constructed. Moreover, the six regional haplotype networks were also created by the PopART software.

3. Results

A total of 28 specimens were successfully amplified for 680 bp fragment of partial *nad5* mitochondrial gene for the cattle isolates retrieved from southern Punjab. It was ascertained that the *E. granulosus s.s.* was prevalent among the cattle of this area with predominant involvement of G3 genotype ($n = 22$; 78.57%). G1 genotype was only characterized from six isolates (21.43%). G1 and G3 sequences had differences in 16 positions, resulting in 8 synonymous and 8 non-synonymous substitutions. Overall these substitutions manifested as one transversion and 15 transitions. There were 16 polymorphic sites in the obtained sequence variants having 8 singleton variable sites and 8 parsimony informative sites. Presence of 12 haplotypes was revealed within this dataset, among which 11 haplotypes (91.66%) were allocated to the G3 genotype (Table 1; Fig. 1). Phylogenetic assessment of the obtained sequences indicated close relation (>98% homology) of G3 haplotypes from Pakistan with that of India (MG682512), Iran (MG682515), Turkey (MG682533, KY766904), Italy (MG682517), France (MG682520) and Spain (MG682527; Fig. 2). Genetic diversity estimates for *E. granulosus s.s.* isolates from Pakistan indicated a high haplotype diversity (0.812 ± 0.061) with low nucleotide diversity (0.00433 ± 0.00073). There was a significantly negative bias from neutrality for Fu's Fs (-3.587, $p < 0.05$) whereas, negative but non-significant trend was estimated for Tajima's D (-0.97995, $p > 0.05$).

Population structure of *E. granulosus s.s.* from different areas of the world was also explored on the basis of the partial *nad5* gene to determine the differences between regional prevalence of the two genotypes. A total of 349 sequences from Africa, Europe, South America, Middle East, Central/East Asia, South Asia were included in the analysis (Table 2). A very high proportion of G1 genotype was identified from the *E. granulosus s.s.* populations of South America ($n = 51$; 100%), Africa ($n = 52$; 98.11%), Central/East Asia ($n = 58$; 93.54%) and Middle East ($n = 65$; 82.28%) followed by Europe ($n = 46$; 64.79%). G3 genotype was

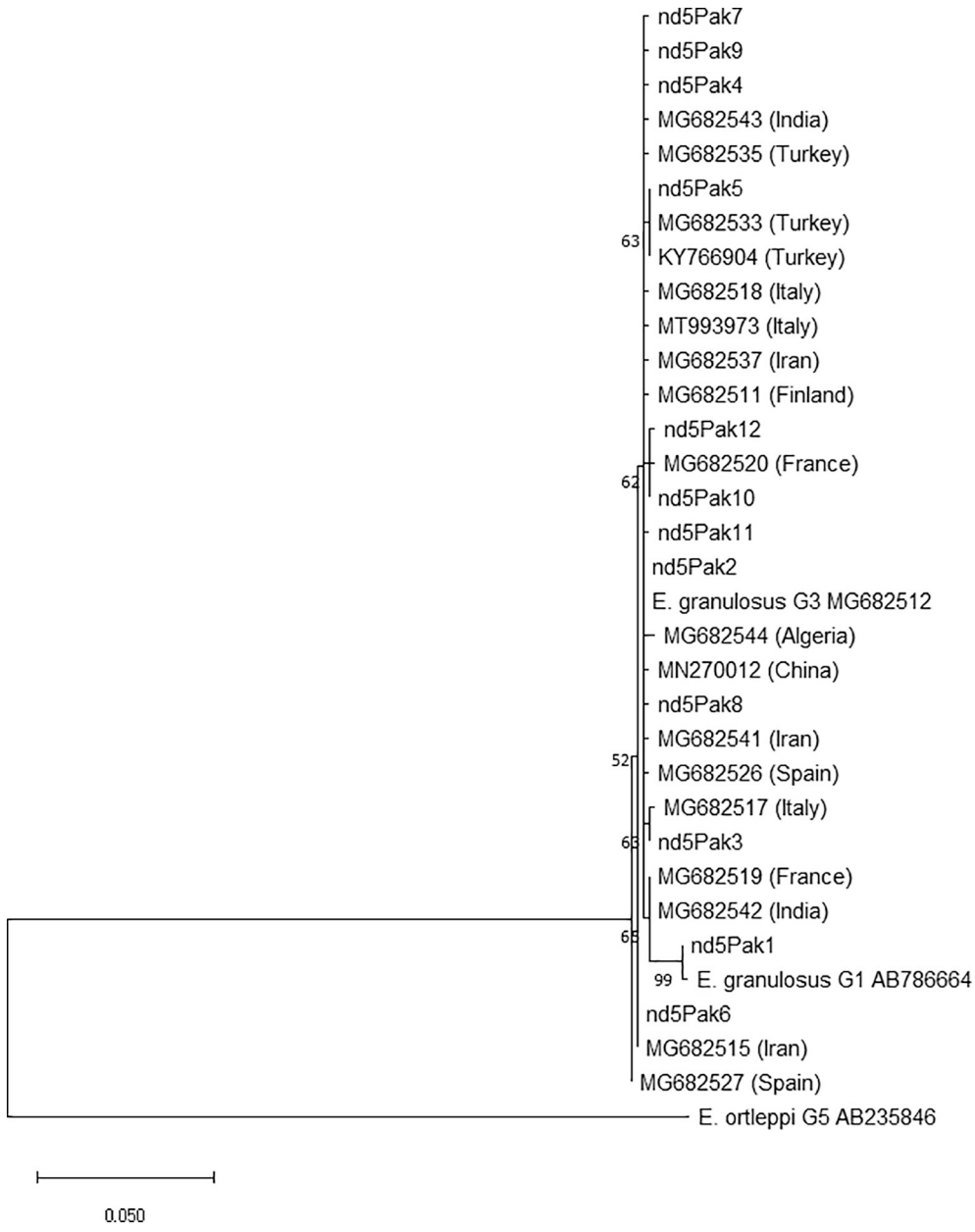


Fig. 2. Phylogenetic tree of *E. granulosus s.s.* isolates based on partial *nad5* gene (680 bp) sequences retrieved from Pakistan (south Punjab). The tree based on maximum-likelihood method indicates the positioning of the retrieved sequences alongwith the reference sequences. Bootstrap value is shown as number on each node.

Table 2
Diversity and neutrality indices for *E. granulosus s.s.* genotypes (G1 and G3) identified on the basis of partial *nad5* gene (mtDNA) among different regional populations.

Regions	No.	*E. granulosus s.s* genotypes		Diversity indices			Neutrality indices	
		G1	G3	hn	hd ± SD	nd ± SD	Tajima's *D*	Fu's Fs
Africa	53	52	1	15	0.634 ± 0.077	0.00175 ± 0.00046	−2.45741*	−11.617*
Europe	71	46	25	30	0.885 ± 0.029	0.00553 ± 0.00035	−1.37774	−18.422*
South America	51	51	0	6	0.289 ± 0.082	0.00051 ± 0.00017	−1.87656*	−4.589*
Middle East	79	65	14	26	0.771 ± 0.049	0.00409 ± 0.00052	−1.69619*	−16.052*
Central/East Asia	62	58	4	31	0.860 ± 0.043	0.00354 ± 0.00050	−2.17978*	−31.668*
South Asia	33	7	26	14	0.831 ± 0.052	0.00432 ± 0.00066	−1.00173	−5.0398*
Total	349	279	70	98	0.772 ± 0.024	0.00432 ± 0.00025	−2.34330*	−26.0367*

* Statistical significance ($p < 0.05$); hd: haplotype diversity; nd: nucleotide diversity.

Table 3
Pairwise fixation index values* between *E. granulosus s.s.* populations of different regions of the world based on partial *nad5* gene (mtDNA).

Geographical location	Africa	Europe	South America	Middle East	Central/ East Asia
Europe	0.18774	–	–	–	–
South America	0.01278	0.22759	–	–	–
Middle East	0.06630	0.04264	0.09178	–	–
Central/East Asia	0.02088	0.12945	0.03557	0.03753	–
South Asia	0.65266	0.24141	0.72977	0.43261	0.54675

* All values are statistically significant at $p < 0.05$.

mainly identified from South Asia ($n = 26$; 78.79%), Europe ($n = 25$; 35.21%) and Middle East ($n = 14$, 17.72%). Correspondingly, there were stark differences for haplotype diversities between different regions with European population manifesting the highest haplotype diversity (0.885 ± 0.029) followed by Central/East Asia (0.860 ± 0.043) and South Asia (0.831 ± 0.052). Lowest value of haplotype diversity was observed for South America (0.289 ± 0.082) having a very small number of haplotypes ($n = 6$) compared to other parts of the world. All studied populations exhibited a significantly negative bias from neutrality for both Tajima's D (−2.34330; $p < 0.05$) and Fu's Fs (−26.0367; $p < 0.05$). African population showed the lowest value for D (−2.45741; $p < 0.05$) whereas the Central/East Asian population had the lowest Fs value (−31.668; $p < 0.05$).

Analysis for the gene flow yielded contrasting situation for different regions of the world (Table 3). South Asian population showed considerably higher Fst values in comparison with South America (0.72977; $p < 0.05$), Africa (0.65266; $p < 0.05$) and Central/East Asia (0.54675; $p < 0.05$) indicating population differentiation. A lesser degree of genetic differentiation also existed between South Asia and Middle East (0.43261; $p < 0.05$), South Asia and Europe (0.24141; $p < 0.05$) and Europe and South America (0.22759; $p < 0.05$). Whereas, absence of genetic differentiation was observed between Africa and South America (0.01278; $p < 0.05$), Central/East Asia and Africa (0.02088; $p < 0.05$), South America and Central/East Asia (0.03557; $p < 0.05$) and Central/East Asia and Middle East (0.03753; $p < 0.05$). A pairwise fixation index was also computed for G1 and G3 haplogroups and the Fst value between these genotypes was 0.84214 ($p < 0.05$) indicating significant genetic differentiation at the mitochondrial level.

Haplotype analysis of the global dataset for mitochondrial *nad5* gene (680 bp) revealed the presence of 98 haplotypes; 73 haplotypes among these nucleotide sequences corresponded to the G1 genotype whereas 25 haplotypes grouped with the G3 genotype. Haplotypic profile for the sequences from different regions of the world displayed a double clustered topology with distinct G1 and G3 haplogroups (Fig. 3). It was further revealed that G3 was mostly confined to South Asia, Europe (Italy, France and Spain) and Middle East (Fig. 4).

4. Discussion

The correct identification and differentiation of *E. granulosus s.s.* genotypes G1 and G3) are of high epidemiological significance as these genotypes are most frequently implicated in human and livestock CE infections (Rojas et al., 2014). So far, most of the studies have relied on *cox1* and *nad1* genes for investigating the isolates across the endemic areas. However, previously identified diagnostic positions within the *cox1* and *nad1* gene regions have shown inconsistencies regarding the genotypic assessment, requiring an improved genetic assay for reliable discrimination of G1 and G3 genotypes. Recently, partial *nad5* mitochondrial gene (680 bp) was shown to have the discriminatory power for accurately distinguishing between G1 and G3 (Kinkar et al., 2018b). Therefore, the current study has focused on the *nad5* gene marker for correct identification of *E. granulosus s.s.* isolates.

In comparison to worldwide association of G1 genotype with CE infections in livestock and humans (Kinkar et al., 2018a; Rojas et al., 2014), Pakistan and India (South Asia) have unusually high prevalence (>50%) of G3 genotype (Alvi et al., 2020; Mehmood et al., 2020b; Muqaddas et al., 2020; Sharma et al., 2013). Differing from G1 by 37 mutations within near-complete mitochondrial genome (Kinkar et al., 2018a), G3 genotype has a high genetic variability and is considered a distinct mitochondrial lineage (Kinkar et al., 2018c). It is commonly known that G3 has higher affinity for buffaloes (93.3% prevalence; Mehmood et al., 2020b) and that the buffaloes also harbor high number of fertile cysts having a prominent role in CE epidemiology in Pakistan

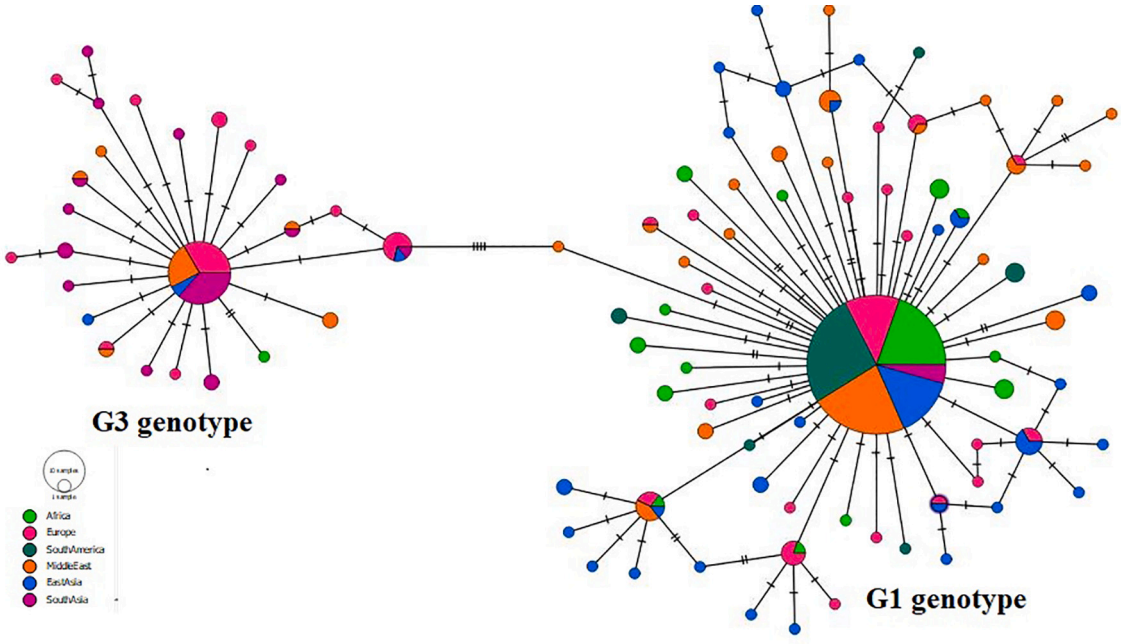


Fig. 3. Haplotype network of G1 and G3 haplotypes identified on the basis of partial *nad5* gene (680 bp) from the six regions of the world (Africa, Europe, South America, Middle East, Central/East Asia, South Asia). Hatch marks represent the number of mutations between the haplotypes and the size of circle corresponds to the frequency of each haplotype in the population.

Fig. 4. World map showing the presence of G1 and G3 genotype haplotypes based on partial *nad5* gene (680 bp) from each of the six regions of the world (Africa, Europe, South America, Middle East, Central/East Asia, South Asia).

(Mehmood et al., 2020a). However, G3 is not limited to buffaloes, it is identified from sheep (Kinkar et al., 2018c; Mehmood et al., 2020b), goats (Mehmood et al., 2020b), cattle (Alvi et al., 2020; Kinkar et al., 2018c; Mehmood et al., 2020b), camels (Kinkar et al., 2018c; Sharbatkhori et al., 2011), humans (Kinkar et al., 2018c; Marinova et al., 2017; Muqaddas et al., 2020), wild boars (Laurimäe et al., 2019) and pigs (Mehmood et al., 2021; Pednekar et al., 2009) indicating a wide host spectrum and possible expansion of this genotype. It is yet not ascertained as to why G3 genotype is particularly prevalent in South Asia. A bayesian phylogeographical analysis indicated a diffusion route of G3 from Turkey to Iran and from there to India (Kinkar et al., 2018c), however, it was drawn on the basis of a low number of samples ($n = 39$) and the scenario might change with the addition of more sequences from this region. The current study represents an overview of genetic diversity of G3 and G1 genotypes in South Asia and other regions of the world to have more insight into the distribution and genetic variability of *E. granulosus s.s.* genotypes based on *nad5* gene.

G1 genotype is the most widely distributed genotype across the world and is generally associated with CE in humans and livestock (Romig et al., 2015). Global subpopulations of G1 are not much genetically differentiated, pointing to repeated expansion *via* animal trade (Kinkar et al., 2018a). Current study delineated high gene flow among populations of Africa, South America and Central/East Asia. The South Asian population which had exceedingly high number of G3 genotype (78.79%) was genetically much differentiated from these populations. Description of G1 and G3 genotypes as distinct mitochondrial lineages has been affirmed on the basis of analysis of near-complete mitogenome sequences from the world supported by a high Fst (0.711; $p < 0.00001$) between these genotypes (Kinkar et al., 2018a). A high genetic differentiation was also estimated in the present study between these two genotypes (G1/G3) on the basis of partial mitochondrial *nad5* gene (0.84214; $p < 0.05$) with distinct regional prevalence of each genotype.

The domestication region in the Middle East has been suggested as a putative point of origin of *E. granulosus s.s.* from where this parasite has dispersed to various parts of the world (Kinkar et al., 2018a, 2018c; Nakao et al., 2010). Past domestication history of the hosts and the complex livestock trade circuits have largely shaped the phylogeographic patterns of this parasite. The *E. granulosus s.s.* genotypes occur in sympatry across all areas with considerably higher share of G1 (>90%) among the CE infections (Central/East Asia, Africa, South America). Almost negligible proportion of infections from G3 genotype have been reported from Africa, South America (Kinkar et al., 2018a, 2018c; Cucher et al., 2016) and Central/East Asia (Yang et al., 2005) whereas this genotype gradually starts showing an increased involvement in CE infections in Turkey (26.53%, Cengiz and Gonenc, 2020), Iran (39%, Abedi et al., 2019), Italy (31.25%, Capuano et al., 2006; 18.84%, Mehmood et al., 2021), India (53.1%, Sharma et al., 2013) and Pakistan (63.33%, Alvi et al., 2020; 66.94%, Mehmood et al., 2020b). Even though current data depicted the absence of G3 from South America, the genotype has been reported from this region contributing to very little disease burden (1.04%) in humans and livestock (Cucher et al., 2016) which does not significantly alter our findings based on the *nad5* gene.

Little knowledge about temporal dynamics of G3 genotype restrict our understanding of the possible causes of its high occurrence in South Asian region. Large and expanding population of buffaloes in this region (Zhang et al., 2020) and affinity of G3 to this host is proposed to be a

probable hypothesis for such high prevalence. However, despite the presence of buffaloes in China, G3 genotype is not usually involved in causing infections there as evident from lower number of G3 cases (4.54%, Cao et al., 2020; 9.9%, Guo et al., 2019; 7.06%, Ohiolei et al., 2019b). One possible explanation of this differential distribution of G1 and G3 could be the presence of Himalayan mountains which form a formidable barrier between Central/East Asia and South Asia hampering animal movement and trade. However, due to lack of data on divergence of G3 from G1, such scenarios could only be considered speculative. It would be of great interest if studies are done with a prime focus on calibration of events delineating the divergence of G1 and G3. It is speculated that the *E. granulosus s.l.* genotypes have diverged since the domestication of dogs and ungulates some 3000–15,000 years ago, however, the presence of cryptic species within this complex before these domestication events cannot be ruled out (Craig et al., 2003). Another proposed hypothesis links the divergence of G1 and G3 genotypes to the last glacial maximum (LGM; 26,500–19,000 years ago) when the climatic changes largely affected the distribution and genetic structure of the species (Kinkar et al., 2018c). Formation of continental ice sheets limited the movement of species and restricted them to different refugia (Southern and Eastern Europe). Isolation of populations during these paleoclimatic events could have resulted in genetic divergence at the mitochondrial level which was retained during post-glacial expansion (Anijalg et al., 2018; Hewitt, 1999).

5. Conclusion

Being distinct mitochondrial lineages, genotypes G1 and G3 differ also significantly in their distribution across different parts of the world. It is believed that the process of domestication and livestock trade have shaped the genetic diversity and biogeographical patterns of these genotypes. Even though the molecular approaches have largely increased our understanding of diversity and distribution of *E. granulosus s.s.*, it still remains unclear as to when and where these genotypes diverged from each other and what were the key factors which have contributed to these discernible differences. Mapping the historical pathways of species diversification and dispersal requires ideally a combination of many gene markers of the mito- and nuclear genomes to be analyzed throughout the distributional range of the parasite. Therefore, despite the power to distinguish G1 and G3, future studies employing the *nad5* gene must take into account limitations of the *nad5* and other genes in phylogeographical studies.

Declaration of Competing Interest

Authors declare that there is no conflict of interest between them.

Acknowledgements

The study was supported by research funding (grant PRG1209) from the Estonian Ministry of Education and Research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2022.105223.

References

Abedi, B., Maghsood, A.H., Khansarinejad, B., Fallah, M., Matini, M., Gholami, S., Pagheh, A.S., Ghasemikhah, R., 2019. Genotyping of *Echinococcus granulosus* isolates from livestock based on mitochondrial *cox1* gene, in the Markazi province, Iran. J. Parasit. Dis. 43 (4), 592–596. https://doi.org/10.1007/s12639-019-01132-4.

Alvi, M.A., Ohiolei, J.A., Saqib, M., Li, L., Tayyab, M.H., Alvi, A.A., Wu, Y.T., Fu, B.Q., Yan, H.B., Jia, W.Z., 2020. *Echinococcus granulosus* (*sensu stricto*)(G1, G3) and *E. ortleppi* (G5) in Pakistan: phylogeny, genetic diversity and population structural analysis based on mitochondrial DNA. Parasit. Vectors 13 (1), 1–10. https://doi.org/10.1186/s13071-020-04199-8.

Anijalg, P., Ho, S.Y., Davison, J., Keis, M., Tammeleht, E., Bobowik, K., Tumanov, I.L., Saveljev, A.P., Lyapunova, E.A., Vorobiev, A.A., Markov, An I., Kryukov, A.P., Kojola, I., Swenson, J.E., Hagen, S.B., Eiken, H.G., Paule, L., Saarma, U., 2018. Large-scale migrations of brown bears in Eurasia and to North America during the Late Pleistocene. J. Biogeogr. 45 (2), 394–405. https://doi.org/10.1111/jbi.13126.

Bandelt, H., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16 (1), 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Bonelli, P., Dei Giudici, S., Peruzzu, A., Mura, L., Santucciu, C., Maestrale, C., Masala, G., 2021. Identification of *Echinococcus granulosus* genotypes G1 and G3 by SNPs genotyping assays. Pathogens 10 (2), 125. https://doi.org/10.3390/pathogens10020125.

Cao, M., Chen, K., Li, W., Ma, J., Xiao, Z., Wang, H., Gao, J., 2020. Genetic characterization of human-derived hydatid fluid based on mitochondrial gene sequencing in individuals from northern and western China. J. Helminthol. 94, e2 https://doi.org/10.1017/S0022149X18000883.

Capuano, F., Rinaldi, L., Maurelli, M.P., Perugini, A.G., Veneziano, V., Garippa, G., Genchi, C., Musella, V., Cringoli, G., 2006. Cystic echinococcosis in water buffaloes: epidemiological survey and molecular evidence of ovine (G1) and buffalo (G3) strains. Vet. Parasitol. 137 (3–4), 262–268. https://doi.org/10.1016/j.vetpar.2006.01.016.

Cengiz, G., Gonenc, B., 2020. Comparison of molecular and morphological characterization and haplotype analysis of cattle and sheep isolates of cystic echinococcosis. Vet. Parasitol. 282, 109132 https://doi.org/10.1016/j.vetpar.2020.109132.

Craig, P.S., Rogan, M.T., Campos-Ponce, M., 2003. Echinococcosis: disease, detection and transmission. Parasitology 127 (S1), S5–S20. https://doi.org/10.1017/S0031182003004384.

Cucher, M.A., Macchiaroli, N., Baldi, G., Camicia, F., Prada, L., Maldonado, L., Avila, H. G., Fox, A., Gutierrez, 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 (2), 166–175. https://doi.org/10.1111/tmi.12647.

Deplazes, P., Rinaldi, L., Rojas, C.A., Torgerson, P.R., Harandi, M.F., Romig, T., Antolova, D., Schurer, J.M., Lahmar, S., Cringoli, G., Mogambo, J., Thompson, R.C. A., Jenkins, E.J., 2017. Global distribution of alveolar and cystic echinococcosis. Adv. Parasitol. 95, 315–493. https://doi.org/10.1016/bs.apar.2016.11.001.

Eckert, J., Deplazes, P., Craig, P., Gemmell, M., Gottstein, B., Heath, D., Jenkins, D., Kamiya, M., Lightowlers, M., Meslin, F., 2001. Echinococcosis in animals: Clinical aspects, diagnosis and treatment. In: Eckert, J., Gemmell, M.A., Meslin, F.X., Pawlowski, Z.S. (Eds.), WHO/OIE Manual on Echinococcosis in Humans and Animals: A Public Health Problem of Global Concern. WHO/OIE.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

Guo, B., Zhang, Z., Zheng, X., Guo, Y., Guo, G., Zhao, L., Cai, R., Wang, B., Yang, M., Shou, X., Zhang, W., Jia, B., 2019. Prevalence and molecular characterization of *Echinococcus granulosus sensu stricto* in Northern Xinjiang, China. Korean J. Parasitol. 57 (2), 153. https://doi.org/10.3347/kjp.2019.57.2.153.

Herricks, J.R., Hotez, P.J., Wanga, V., Coffeng, L.E., Haagsma, J.A., Basáñez, M.G., Buckle, G., Budke, C.M., Carabin, H., Fèvre, E.M., Fürst, T., Halasa, Y.A., King, C.H., Murdoch, M.E., Ramaiah, K.D., Shepard, D.S., Stolk, W.A., Undurraga, E.A., Stanaway, J.D., Naghavi, M., Murray, C.J., 2017. The global burden of disease study 2013: what does it mean for the NTDs? PLoS Negl. Trop. Dis. 11 (8), e0005424 https://doi.org/10.1371/journal.pntd.0005424.

Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68 (1–2), 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x.

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

Kinkar, L., Laurimäe, 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. Infect. Genet. Evol. 52, 52–58. https://doi.org/10.1016/j.meegid.2017.04.023.

Kinkar, L., Laurimäe, T., Acosta-Jamett, G., Andresiuk, V., Balkaya, I., Casulli, A., Zait, H., Irshadullah, M., Jabbar, A., Jenkins, D.J., Kia, E.B., Manfredi, M.T., Mirhendi, H., M'rad, S., Rostami-Nejad, M., Oudni-M'rad, M., Pirangeli, N.B., Ponce-Gordo, F., Rehbein, S., Sharbatkhori, M., Simsek, S., Soriano, S.V., Sprong, H., Snabel, V., Umhang, G., Varcasia, A., Saarma, U., 2018a. Global phylogeography and genetic diversity of the zoonotic tapeworm *Echinococcus granulosus* sensu stricto genotype G1. Int. J. Parasitol. 48 (9–10), 729–742. https://doi.org/10.1016/j.ijpara.2018.03.006.

Kinkar, L., Laurimäe, T., Acosta-Jamett, G., Andresiuk, V., Balkaya, I., Casulli, A., Gasser, R.B., González, L.M., Haag, K.L., Zait, H., Irshadullah, M., Jabbar, A., Jenkins, D.J., Manfredi, M.T., Mirhendi, H., M'rad, S., Rostami-Nejad, M., Oudni-M'rad, M., Pirangeli, N.B., Ponce-Gordo, F., Rehbein, S., Sharbatkhori, M., Kia, E.B., Simsek, S., Soriano, S.V., Sprong, H., Snabel, V., Umhang, G., Varcasia, A., Saarma, U., 2018b. Distinguishing *Echinococcus granulosus* sensu stricto genotypes G1 and G3 with confidence: a practical guide. Infect. Genet. Evol. 64, 178–184. https://doi.org/10.1016/j.meegid.2018.06.026.

Kinkar, L., Laurimäe, T., Balkaya, I., Casulli, A., Zait, H., Irshadullah, M., Sharbatkhori, M., Mirhendi, H., Rostami-Nejad, M., Ponce-Gordo, F., Rehbein, S.,

Kia, E.B., Simsek, S., Snabel, V., Umhang, G., Varcasia, A., Saarma, U., 2018c. Genetic diversity and phylogeography of the elusive, but epidemiologically important *Echinococcus granulosus* sensu stricto genotype G3. Parasitology 145 (12), 1613–1622. https://doi.org/10.1017/S0031182018000549.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096.

Laurimäe, T., Kinkar, L., Moks, E., Romig, T., Omer, R.A., Casulli, A., Umhang, G., Bagrade, G., Irshadullah, M., Sharbatkhori, M., Mirhendi, H., Ponce-Gordo, F., Soriano, S.V., Varcasia, A., Rostami-Nejad, M., Andresiuk, V., Saarma, U., 2018. Molecular phylogeny based on six nuclear genes suggests that *Echinococcus granulosus sensu lato* genotypes G6/G7 and G8/G10 can be regarded as two distinct species. Parasitology 145, 1929–1937. https://doi.org/10.1017/S0031182018000719.

Laurimäe, T., Kinkar, L., Varcasia, A., Dessì, G., Sgroi, G., D'Alessio, N., Veneziano, V., Saarma, U., 2019. First detection of zoonotic tapeworm *Echinococcus granulosus* sensu lato genotype G7 in continental Italy. Parasitol. Res. 118 (7), 2193–2201. https://doi.org/10.1007/s00436-019-06346-2.

LDD, 2018. Livestock Census Punjab 2018: First Real Time (Door to Door) Livestock Census Big Data Analytics. Livestock and dairy Development Department, Lahore, Pakistan. http://www.9211.com.pk/LiveStockAdmin/uploads/editor_files/livestock_census_punjab_2018_sven4.pdf.

Lymbery, A.J., Jenkins, E.J., Schurer, J.M., Thompson, R.A., 2015. *Echinococcus canadensis*, *E. borealis*, and *E. intermedius*. What's in a name? Trends Parasitol. 31 (1), 23–29. https://doi.org/10.1016/j.pt.2014.11.003.

Marinova, I., Spiliotis, M., Wang, J., Muhtarov, M., Chaligiannis, I., Sotiraki, S., Rainova, I., Gottstein, B., Boubaker, G., 2017. Molecular characterization of *Echinococcus granulosus* isolates from Bulgarian human cystic echinococcosis patients. Parasitol. Res. 116 (3), 1043–1054. https://doi.org/10.1007/s00436-017-5386-1.

Mehmood, N., Arshad, M., Ahmed, H., Simsek, S., Muqaddas, H., 2020a. Comprehensive account on prevalence and characteristics of hydatid cysts in livestock from Pakistan. Korean J. Parasitol. 58 (2), 121. https://doi.org/10.3347/kjp.2020.58.2.121.

Mehmood, N., Muqaddas, H., Arshad, M., Ullah, M.I., Khan, Z.I., 2020b. Comprehensive study based on mtDNA signature (*nad1*) providing insights on *Echinococcus granulosus s.s.* genotypes from Pakistan and potential role of buffalo-dog cycle. Infect. Genet. Evol. 81, 104271 https://doi.org/10.1016/j.meegid.2020.104271.

Mehmood, N., Dessì, G., Ahmed, F., Joanny, G., Tamponi, C., Cappai, M.G., Varcasia, A., Scala, A., 2021. Genetic diversity and transmission patterns of *Echinococcus granulosus sensu stricto* among domestic ungulates of Sardinia, Italy. Parasitol. Res. 120 (7), 2533–2542. https://doi.org/10.1007/s00436-021-07186-9.

Muqaddas, H., Mehmood, N., Arshad, M., 2020. Genetic variability and diversity of *Echinococcus granulosus sensu lato* in human isolates of Pakistan based on *cox1* mt-DNA sequences (366bp). Acta Trop. 207, 105470 https://doi.org/10.1016/j.actatropica.2020.105470.

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 (5), 713–722. https://doi.org/10.1017/S0031182006001934.

Nakao, M., Li, T., Han, X., Ma, X., Xiao, N., Qiu, J., Wang, H., Yanagida, T., Mamuti, W., Wen, H., Moro, P.L., 2010. Genetic polymorphisms of Echinococcus tapeworms in China as determined by mitochondrial and nuclear DNA sequences. Int. J. Parasitol. 40 (3), 379–385. https://doi.org/10.1016/j.ijpara.2009.09.006.

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 (13), 1625–1636. https://doi.org/10.1017/S0031182013000565.

Nakao, M., Lavikainen, A., Hoberg, E., 2015. Is *Echinococcus intermedius* a valid species. Trends Parasitol. 31 (8), 342–343. https://doi.org/10.1016/j.pt.2015.04.012.

Ohiolei, J.A., Yan, H.B., Li, L., Magaji, A.A., Luka, J., Zhu, G.Q., Isaac, C., Odoya, M.E., Wu, Y.T., Alvi, M.A., Muku, R.J., 2019a. Cystic echinococcosis in Nigeria: first insight into the genotypes of *Echinococcus granulosus* in animals. Parasit. Vectors 12 (1), 1–10. https://doi.org/10.1186/s13071-019-3644-z.

Ohiolei, J.A., Xia, C.Y., Li, L., Liu, J.Z., Tang, W.Q., Wu, Y.T., Zhu, G.Q., Shi, B., Fu, B.Q., Yin, H., Yan, H.B., 2019b. Genetic variation of *Echinococcus* spp. in yaks and sheep in the Tibet Autonomous Region of China based on mitochondrial DNA. Parasit. Vectors 12 (1), 1–10. https://doi.org/10.1186/s13071-019-3857-1.

PBS, Pakistan Bureau of Statistics, 2006. Livestock Census 2006. Agricultural Census Organization, Statistics Division, Pakistan Bureau of Statistics, Islamabad, Pakistan. https://www.pbs.gov.pk/content/pakistan-livestock-census-2006.

Pednekar, R.P., Gatne, M.L., Thompson, R.A., Traub, R.J., 2009. Molecular and morphological characterisation of *Echinococcus* from food producing animals in India. Vet. Parasitol. 165 (1–2), 58–65. https://doi.org/10.1016/j.vetpar.2009.06.021.

PES, Pakistan Economic Survey, 2020-21. Economic Adviser's Wing. Finance Division, Government of Paistan, Islamabad, Pakistan. http://www.finance.gov.pk/survey_2021.html.

Reynolds, J., Weir, B.S., Cockerham, C.C., 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105, 767–779. https://doi.org/10.1093/genetics/105.3.767.

Rojas, C.A., Romig, T., Lightowlers, M.W., 2014. *Echinococcus granulosus sensu lato* genotypes infecting humans–review of current knowledge. Int. J. Parasitol. 44 (1), 9–18. https://doi.org/10.1016/j.ijpara.2013.08.008.

Romig, T., Ebi, D., Wassermann, M., 2015. Taxonomy and molecular epidemiology of *Echinococcus granulosus sensu lato*. Vet. Parasitol. 213 (3–4), 76–84. https://doi.org/10.1016/j.vetpar.2015.07.035.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34 (12), 3299–3302. https://doi.org/10.1093/molbev/msx248.

Sharbatkhori, M., Harandi, M.F., Mirhendi, H., Hajialilo, E., Kia, E.B., 2011. Sequence analysis of *cox1* and *nad1* genes in *Echinococcus granulosus* G3 genotype in camels (*Camelus dromedarius*) from Central Iran. Parasitol. Res. 108 (3), 521–527. https://doi.org/10.1007/s00436-010-2092-7.

Shariatzadeh, S.A., Spotin, A., Gholami, S., Fallah, E., Hazratian, T., Mahami-Oskouei, M., Montazeri, F., Moslemzadeh, H.R., Shahbazi, A., 2015. The first morphometric and phylogenetic perspective on molecular epidemiology of *Echinococcus granulosus sensu lato* in stray dogs in a hyperendemic Middle East focus, northwestern Iran. Parasit. Vectors 8 (1), 1–10. https://doi.org/10.1186/s13071-015-1025-9.

Sharma, M., Sehgal, R., Fomda, B.A., Malhotra, A., Malla, N., 2013. Molecular characterization of *Echinococcus granulosus* cysts in north Indian patients: identification of G1, G3, G5 and G6 genotypes. PLoS Negl. Trop. Dis. 7 (6), e2262 https://doi.org/10.1371/journal.pntd.0002262.

Soba, B., Gaspersic, S., Kese, D., Kotar, T., 2020. Molecular characterization of *Echinococcus granulosus sensu lato* from humans in Slovenia. Pathogens 9 (7), 562. https://doi.org/10.3390/pathogens9070562.

Thompson, R.C.A., 2017. Biology and systematics of Echinococcus. Adv. Parasitol. 95, 65–109. https://doi.org/10.1016/bs.apar.2016.07.001.

Thompson, R.C.A., 2020. The molecular epidemiology of Echinococcus infections. Pathogens 9 (6), 453. https://doi.org/10.3390/pathogens9060453.

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 A 27 (3), 1701–1702. https://doi.org/10.3109/19401736.2014.961129.

Yang, Y.R., Rosenzvit, M.C., Zhang, L.H., Zhang, J.Z., McManus, D.P., 2005. Molecular study of *Echinococcus* in west-Central China. Parasitology 131 (4), 547–555. https://doi.org/10.1017/S0031182005007973.

Zhang, Y., Colli, L., Barker, J.S.F., 2020. Asian water buffalo: domestication, history and genetics. Anim. Genet. 51 (2), 17–191. https://doi.org/10.1111/age.12911.