*Parasitology*

**cambridge.org/par**

# Haplotype comparisons of *Echinococcus granulosus sensu lato via* mitochondrial gene sequences (*co*1, *cytb*, *nadh*1) among Pakistan and its neighbouring countries

Aisha Khan[1,2,3], Haroon Ahmed[3], Sami Simsek[4], Khuram Shahzad[3], Figen Celik[4], Muhammad Sohail Afzal[5], Mobushir Riaz Khan[6], Hua Liu[1,2], Yujuan Shen[1,2] and Jianping Cao[1,2]

[1]National Institute of Parasitic Diseases, Chinese Centre for Disease Control and Prevention (Chinese Center for Tropical Diseases Research); Key Laboratory of Parasite and Vector Biology; National Health Commission of the People's Republic of China; WHO Collaborating Centre for Tropical Diseases; National Center for International Research on Tropical Diseases; Shanghai; China; [2]The School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [3]Department of Biosciences, COMSATS University Islamabad, Islamabad, Pakistan; [4]Department of Parasitology, Faculty of Veterinary Medicine, University of Firat, Elazig, Turkey; [5]Department of Lifesciences, School of Science, University of Management & Technology (UMT), Lahore, Pakistan and [6]School of Environmental Science, Charles Sturt University, Albury, NSW, Australia

## Research Article

**Cite this article:** Khan A *et al* (2021). Haplotype comparisons of *Echinococcus granulosus sensu lato via* mitochondrial gene sequences (*co*1, *cytb*, *nadh*1) among Pakistan and its neighbouring countries. *Parasitology* 1–11. https://doi.org/10.1017/S0031182021000688

Received: 11 February 2021
Revised: 8 April 2021
Accepted: 15 April 2021

**Key words:**
Bioinformatics; *co*1; *cytb*; *Echinococcus granulosus sensu lato*; haplotype; *nadh*1

**Author for correspondence:**
Haroon Ahmed,
E-mail: haroonahmad12@yahoo.com;
Jianping Cao,
E-mail: caojp@yahoo.com

## Abstract

*Echinococcus granulosus sensu lato* (*s.l.*) is a zoonotic parasite that causes cystic echinococcosis (CE) in humans. However, *E. granulosus sensu stricto* (*s.s.*) is considered the predominant species in CE infections worldwide. According to the population genetic diversity and structure of *E. granulosus s.l.*, gene flow can explain the parasite drift among the neighbouring countries of Pakistan. The mitochondrial (mt) *co*1 ($n = 47$), *nadh*1 ($n = 37$) and *cytb* ($n = 35$) nucleotide sequences of *E. granulosus s.l.* isolates from Pakistan, Iran, China and India were retrieved from the National Centre for Biotechnology Information database to determine the genealogical relationships. The sequences were grouped as the mt-*co*1 (genotypes G1 and G3, G6-G7), mt-*cytb* (genotypes G1 and G3), and mt-*nadh*1(genotypes G1 and G3). The data were analysed using bioinformatic tools. A total of 19 polymorphic sites for the mt-*co*1 sequence (374 bp) were observed of which 31.6% (6/19) were parsimony-informative sites. Unique singleton haplotypes within the *E. granulosus s.s.* haplotype network based on the mt-*co*1 gene were highly prevalent (68.4%; 13/19) in Pakistani isolates followed by Chinese, Indian and Iranian isolates; four polymorphic sites were detected in the *E. canadensis* (G6/G7). In *E. canadensis* mt-*co*1 haplotype network, 75% (3/4) unique singleton haplotypes were from the Iranian isolates. Twelve polymorphic sites were found using the mt-*cytb* sequence (547 bp); 25% (3/12) were parsimony-informative and there were 66.7% (8/12) unique singleton haplotypes within the mt-*cytb* haplotype network in *E. granulosus s.s.* with the most reported from Pakistan followed by Iran and China. 20 polymorphic sites were detected in *E. granulosus s.s.* mt-*nadh*1 sequences (743 bp); 20% (4/20) were parsimony-informative. There were 66.7% (8/12) main single haplotypes within the mt-*nadh*1 haplotype network, with the most reported from Pakistan followed by that from India, Iran and China. The sequence analyses show low nucleotide diversity and high haplotype diversity in general.

## Introduction

*Echinococcus granulosus sensu lato* (*s.l.*) is a tapeworm that causes cystic echinococcosis (CE). It is distributed in both livestock and humans worldwide (Addy *et al.*, 2017). It is quite widespread in countries where pastoral activities are more common. *E. granulosus s.l.* consists of five cryptic species on the basis of mitochondrial (mt) and nuclear DNA sequences: *E. granulosus sensu stricto* (*s.s.*) (genotypes G1 and G3), *E. equinus* (genotype G4), *E. ortleppi* (genotype G5), *E. canadensis* (genotype G6–G10) and *E. felidis* (Wen *et al.*, 2019).

A report on food-borne parasites in humans listed CE as the second most serious disease from such parasites [EFSA Panel on Biological Hazards (BIOHAZ) *et al.*, 2018]. Wild carnivores and dogs are the definitive host, and the intermediate host are livestock species while accidental intermediate hosts are humans in the parasite's life cycle. Contaminated water and food with parasite eggs are responsible for human infection after ingestion. Typically, the development of the hydatid cyst affects the lungs and liver of the intermediate host. The cycle continues until a carnivore ingests the infected host's fertile cyst (Kinkar *et al.*, 2016). The rate of spread of *E. granulosus s.s.* (genotypes G1 and G3) has increased in domestic and wild animals of South America and Central Asia. It is also involved in 88% of human infection worldwide (Kinkar *et al.*, 2017).

At least 270 million people (58% of the total population) in Central Asian countries such as Pakistan, Iran, Afghanistan, Turkmenistan, Mongolia, Tajikistan, Kazakhistan, Kyrgyzstan and China are at risk of CE (Zhang *et al.*, 2015). Southwestern part of the China is one of main affected regions of CE infections, where it has been reported in 20 provinces. Climate change and geologic events may be responsible for the high level of species biodiversity in that area (Wang *et al.*, 2014*a*, *b*). In Iran, the prevalence of CE have been reported in 5.1–74% of sheep, 1.7–20% of goats, 11.4–70% of camels and 3.5–38.3% of cattle (Sharbatkhori *et al.*, 2010). Further, humans are greatly affected by CE (Shahnazi *et al.*, 2011).

Hydatid cysts were found in 12.2% (28/230) of sheep and 10.7% (21/197) of goats in Kangra Valley of the northwestern Himalayas, India (Moudgil *et al.*, 2019). Moreover, in northern India 4130 animals were examined, 66 (1.6%) were positive for hydatid cysts. The occurrence of CE was 5.39%, 4.36%, 3.09%, 2.23% and 0.41% in cattle, buffaloes, pigs, sheep and goats, respectively (Singh *et al.*, 2014). Out of 1429 human samples from Kashmir, North India, 72 (5.03%) were immunoglobulin G-positive by enzyme-linked immunosorbent assay (Fomda *et al.*, 2015). Hydatidosis causes US$212.35 million economic losses annualy in India (Singh *et al.*, 2013).

CE in humans has been documented all over the Pakistan, including the Lahore, Karachi, Hyderabad, Peshawar and Punjab Sindh provinces (Khan *et al.*, 2020). In Karachi, more than 225 CE cases were reported over a 10-year period (Butt and Khan, 2020). In Pakistan, CE cases increases day by the day and cause health issues, and the annual economic losses of the livestock sector are about US$26.5 million (Haleem *et al.*, 2018).

Currently, livestock import and export has increased between Pakistan and its neighbouring countries such as Iran, India and China. This is the main reason for the genetic variation and gene flow among *Echinococcus* spp. isolates, as in Iran (Spotin *et al.*, 2017). However, there is currently no comparative study on the population structure and genetic variability of *Echinococcus* spp. in the above geographic line.

The objective of the current study was to compare the mt gene sequences (*co*1, *cytb*, and *nadh*1) belonging to *E. granulosus s.l.* human isolates in Pakistan with published sequences from China, Iran and India to track the evolutionary history of CE isolates that have drifted among these regions.

## Materials and methods

### Data collection

The dataset was created from *E. granulosus s.s.* mt-*co*1 ($n = 19$), mt-*nadh*1 ($n = 11$) and mt-*cytb* ($n = 9$) nucleotide sequences and two mt-*co*1 nucleotide sequences of *E. canadensis* (G6/G7) isolates of human origin from Pakistan that had been previously submitted by our team to the National Center for Biotechnology Information (NCBI) database. From the NCBI database, we also retrieved 28 mt-*co*1 (374 bp), 26 mt-*nadh*1 (743 bp) and 26 mt-*cytb* (547 bp) nucleotide sequences of *E. granulosus s.s.* and 12 mt-*co*1 (366 bp) nucleotide sequences of *E. canadensis* (G6/G7) isolates based on neighboring countries like Iran, China and India to determine the genealogical relationships (Table 1). The sequence data for each gene fragment from PubMed were searched for the three countries. Based on the search results, sequences with a length close to that of our gene region were selected, and only one of the group sequences was selected for the analysis. The published sequence data were retrieved from NCBI PubMed during May 2020.

### Data and phylogenetic analysis

We selected 47 mt-*co*1 (374 bp), 37 mt-*nadh*1 (743 bp) and 35 mt-*cytb* (547 bp) nucleotide sequences from *E. granulosus s.s.* and 14 mt-*co1* (366 bp) sequences of *E. canadensis* (G6/G7) for the analysis. All sequence data were downloaded from GenBank in FASTA format. The files were uploaded to MEGA-X tool for alignment (Kumar *et al.*, 2018). Sequences of several lengths were aligned using ClustalW to generate output formats such as FASTA and NEXUS for subsequent analysis. The aligned sequences with unequal lengths were trimmed at the ends for equalization. Akaike and Bayesian information criterion analyses were used to analyze the sequences in MEGA-X and most fesible model was obtained while evolutinary tree was obtained separately for each gene region (Kumar *et al.*, 2018). The statistical analysis for unique clades was run through 1,000 bootstrapping replicates.

### Haplotype analysis

The haplotype analysis of the sequences was performed in DnaSP 6 package (Rozas *et al.*, 2017). Using this package, population diversity indices [haplotype numbers ($h$), nucleotide diversity ($\pi$), haplotype diversity (Hd)], neutrality indices (Tajima's $D$ statistics, Fu's statistics), and Fu and Li's $D$ test and $F$ test were performed (Rozas *et al.*, 2017). The NEXUS format for the subsequent analysis was also generated. Later on, networks were constructed with the MSN (minimum spanning networks) (Bandelt *et al.*, 1999) method using PopART 1.7 software (Leigh *et al.*, 2015).

## Results

Neutrality and diversity indices of the sequences from mt-*co*1, mt-*cytb* and mt-*nadh*1 from *E. granulosus s.l.* isolates from Pakistan, Iran, China and India are shown in Table 2. The mt-*co*1 gene of *E. granulosus s.s.* had the highest haplotype diversity. Tajima's $D$ values were observed negative for all of the sequences, indicating purifying selection and/or population expansion. In particular, Fu's $F$ value (−14 106) was quite negative for mt-*co*1 gene of *E. granulosus s.s.*

### Phylogenetic analysis

Figure 1 illustrates the evolutionary analysis of the *E. granulosus s.s.* mt-*co*1 nucleotide sequence (374 bp). Two human (IND1 and IND2) and two buffalo (IND3 and IND4) isolates were identified as the most distant isolates. Figure 2 shows the genetic tree of the *E. granulosus s.s.* mt-*cytb* nucleotide sequence (547 bp). The Indian buffalo isolates (IND11 and IND13) were identified as the most distant sequences. The other distant isolates from the main clade were the PK30, PK35, PK28, PK36, PK31 and PK34 isolates (all human). For the *E. granulosus s.s.* mt-*nadh*1 nucleotide sequence (743 bp), two Iranian sheep isolates (IR27 and IR30) were identified as the most distant isolates, followed by isolates IR23 (camel) and IR26 (camel) (Fig. 3). Figure 4 shows the phylogenetic analysis of the mt-*co*1 nucleotide sequence (366 bp) of *E. canadensis* (G6/G7). The IR37 human isolate was a separate branch from the main clade. The others were in the main clade.

### Haplotype analysis

The analysis showed the presence of 19 polymorphic sites of the mt-*co*1 nucleotide sequences; of which 31.6% (6/19) were of parsimony-informative. The *E. granulosus s.s.* mt-*co*1 haplotype network had 20 haplotypes, organized in a star-like orientation, alongside of fundamental focal haplotype, isoated from different haplotypes by 1–5 mutational points and

**Table 1.** Pakistani (PK), Iranian (IR), Chinese (CH), and Indian (IND) mt-*co*1, mt-*cytb*, and mt-*nadh*1 gene sequences of *E. granulosus s.l.*

| Accession # | Country | Gene name | Genotype | Host | Actual gene length | Position of 120 characters of alignment |
|---|---|---|---|---|---|---|
| MK229294 | PK1 | *co*1 | G3 | Human | 424 | 43–416 |
| MK229295 | PK2 | *co*1 | G1 | Human | 445 | 54–427 |
| MK229296 | PK3 | *co*1 | G1 | Human | 448 | 57–430 |
| MK229297 | PK4 | *co*1 | G1 | Human | 446 | 55–428 |
| MK229298 | PK5 | *co*1 | G3 | Human | 449 | 57–430 |
| MK229299 | PK6 | *co*1 | G1 | Human | 443 | 51–424 |
| MK229301 | PK8 | *co*1 | G1 | Human | 448 | 57–430 |
| MK229302 | PK9 | *co*1 | G1 | Human | 446 | 55–428 |
| MK229303 | PK10 | *co*1 | G3 | Human | 449 | 57–430 |
| MK229304 | PK11 | *co*1 | G1 | Human | 443 | 51–424 |
| MK229311 | PK18 | *co*1 | G3 | Human | 447 | 56–429 |
| MK229312 | PK19 | *co*1 | G3 | Human | 449 | 57–430 |
| MK229313 | PK20 | *co*1 | G1 | Human | 447 | 56–429 |
| MK229314 | PK21 | *co*1 | G3 | Human | 448 | 56–429 |
| MK229315 | PK22 | *co*1 | G1 | Human | 448 | 56–429 |
| MK229316 | PK23 | *co*1 | G3 | Human | 446 | 55–428 |
| MK229317 | PK24 | *co*1 | G1 | Human | 447 | 55–428 |
| MK229318 | PK25 | *co*1 | G1 | Human | 446 | 55–428 |
| MK229319 | PK26 | *co*1 | G1 | Human | 446 | 55–428 |
| MG682540 | IR1 | *co*1 | G3 | Camel | 1674 | 762–1135 |
| MG682539 | IR2 | *co*1 | G3 | Camel | 1674 | 762–1135 |
| MG682538 | IR3 | *co*1 | G3 | Camel | 1674 | 762–1135 |
| MG682537 | IR4 | *co*1 | G3 | Sheep | 1674 | 762–1135 |
| MG682541 | IR5 | *co*1 | G3 | Camel | 1674 | 762–1135 |
| KY766894 | IR6 | *co*1 | G3 | Camel | 1674 | 762–1135 |
| MG672231 | IR7 | *co*1 | G1 | Sheep | 1674 | 762–1135 |
| MG672232 | IR8 | *co*1 | G1 | Sheep | 1674 | 762–1135 |
| MG672237 | IR9 | *co*1 | G1 | Sheep | 1674 | 762–1135 |
| MG672236 | IR10 | *co*1 | G1 | Sheep | 1674 | 762–1135 |
| MH050612 | CH1 | *co*1 | G1 | Human | 528 | 129–502 |
| MH050613 | CH2 | *co*1 | G1 | Human | 528 | 129–502 |
| AB688604 | CH3 | *co*1 | G1 | Human | 1609 | 762–1135 |
| AB688603 | CH4 | *co*1 | G1 | Human | 1609 | 762–1135 |
| AB688602 | CH5 | *co*1 | G1 | Human | 1609 | 762–1135 |
| MH050618 | CH6 | *co*1 | G3 | Human | 528 | 129–502 |
| MH050617 | CH7 | *co*1 | G1 | Human | 528 | 129–502 |
| MH050616 | CH8 | *co*1 | G1 | Human | 528 | 129–502 |
| MH050615 | CH9 | *co*1 | G1 | Human | 528 | 129–502 |
| MH050614 | CH10 | *co*1 | G1 | Human | 528 | 129–502 |
| KC422644 | IND1 | *co*1 | G3 | Human | 392 | 19–392 |
| KC422645 | IND2 | *co*1 | G3 | Human | 392 | 1–374 |
| MG682542 | IND3 | *co*1 | G3 | Buffalo | 1674 | 762–1135 |
| KY766902 | IND4 | *co*1 | G3 | Buffalo | 1674 | 762–1135 |
| KY766891 | IND5 | *co*1 | G1 | Buffalo | 1674 | 762–1135 |
| MG672260 | IND6 | *co*1 | G1 | Buffalo | 1674 | 762–1135 |
| MG682543 | IND7 | *co*1 | G3 | Buffalo | 1674 | 762–1135 |

(*Continued*)

**Table 1.** (*Continued.*)

| Accession # | Country | Gene name | Genotype | Host | Actual gene length | Position of 120 characters of alignment |
|---|---|---|---|---|---|---|
| MG682512 | IND8 | *co*1 | G3 | Buffalo | 1674 | 762–1135 |
| MK229300 | PK7 | *co*1 | G6 | Human | 436 | 27–392 |
| MK229305 | PK12 | *co*1 | G6/7 | Human | 436 | 27–392 |
| KP751430 | IR31 | *co*1 | G6 | Camel | 1183 | 677–1042 |
| KP751429 | IR32 | *co*1 | G6 | Camel | 1183 | 677–1042 |
| KP751428 | IR33 | *co*1 | G6 | Camel | 1183 | 677–1042 |
| KP751427 | IR34 | *co*1 | G6 | Camel | 1183 | 677–1042 |
| KP751426 | IR35 | *co*1 | G6 | Camel | 1183 | 677–1042 |
| HM563020 | IR36 | *co*1 | G6 | Human | 366 | 1–366 |
| KX893480 | IR37 | *co*1 | G6/7 | Human | 1608 | 745–1110 |
| KX893479 | IR38 | *co*1 | G6/7 | Human | 1608 | 745–1110 |
| KX893478 | IR39 | *co*1 | G6/7 | Human | 1608 | 745–1110 |
| KX893477 | IR40 | *co*1 | G6/7 | Camel | 1608 | 745–1110 |
| MH050619 | CH31 | *co*1 | G6/7 | Human | 528 | 112–477 |
| DQ356884 | CH32 | *co*1 | G6 | Human | 444 | 37–402 |
| MK229320 | PK27 | *cytb* | G3 | Human | 617 | 33–579 |
| MK229321 | PK28 | *cytb* | G1 | Human | 557 | 11–557 |
| MK229323 | PK30 | *cytb* | G3 | Human | 557 | 9–555 |
| MK229324 | PK31 | *cytb* | G1 | Human | 585 | 39–585 |
| MK229325 | PK32 | *cytb* | G3 | Human | 558 | 11–557 |
| MK229327 | PK34 | *cytb* | G1 | Human | 563 | 15–561 |
| MK229328 | PK35 | *cytb* | G3 | Human | 564 | 13–559 |
| MK229329 | PK36 | *cytb* | G1 | Human | 560 | 11–557 |
| K229332 | PK39 | *cytb* | G1 | Human | 563 | 14–560 |
| KC709561 | CH11 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709557 | CH12 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709559 | CH13 | *cytb* | G1 | Human | 1068 | 404–950 |
| KC709558 | CH14 | *cytb* | G1 | Human | 1068 | 404–950 |
| KC709557 | CH15 | *cytb* | G1 | Human | 1068 | 404–950 |
| KC709566 | CH16 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709565 | CH17 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709564 | CH18 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709563 | CH19 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| KC709562 | CH20 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| MG682540 | IR11 | *cytb* | G3 | Camel | 1068 | 404–950 |
| MG682539 | IR12 | *cytb* | G3 | Camel | 1068 | 404–950 |
| MG682538 | IR13 | *cytb* | G3 | Camel | 1068 | 404–950 |
| MG682537 | IR14 | *cytb* | G3 | Sheep | 1068 | 404–950 |
| MG682541 | IR15 | *cytb* | G3 | Camel | 1068 | 404–950 |
| KY766894 | IR16 | *cytb* | G3 | Camel | 1068 | 404–950 |
| MG672231 | IR17 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| MG672232 | IR18 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| MG672237 | IR19 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| MG672236 | IR20 | *cytb* | G1 | Sheep | 1068 | 404–950 |
| MG682542 | IND11 | *cytb* | G3 | Buffalo | 1068 | 404–950 |
| MG682512 | IND12 | *cytb* | G3 | Buffalo | 1068 | 404–950 |

(*Continued*)

**Table 1.** (*Continued.*)

| Accession # | Country | Gene name | Genotype | Host | Actual gene length | Position of 120 characters of alignment |
|---|---|---|---|---|---|---|
| KY766902 | IND13 | *cytb* | G3 | Buffalo | 1068 | 404–950 |
| KY766891 | IND14 | *cytb* | G1 | Buffalo | 1068 | 404–950 |
| MG672260 | IND15 | *cytb* | G1 | Buffalo | 1068 | 404–950 |
| MG682543 | IND16 | *cytb* | G3 | Buffalo | 1068 | 404–950 |
| MG682543 | IND21 | *nadh*1 | G3 | Buffalo | 894 | 140–882 |
| MG682542 | IND22 | *nadh*1 | G3 | Buffalo | 894 | 140–882 |
| MG682512 | IND23 | *nadh*1 | G3 | Buffalo | 894 | 140–882 |
| KY766902 | IND24 | *nadh*1 | G3 | Buffalo | 894 | 140–882 |
| KY766891 | IND25 | *nadh*1 | G1 | Buffalo | 894 | 140–882 |
| MG672260 | IND26 | *nadh*1 | G1 | Buffalo | 894 | 140–882 |
| MG682540 | IR21 | *nadh*1 | G3 | Camel | 894 | 140–882 |
| MG682539 | IR22 | *nadh*1 | G3 | Camel | 894 | 140–882 |
| MG682538 | IR23 | *nadh*1 | G3 | Camel | 894 | 140–882 |
| MG682537 | IR24 | *nadh*1 | G3 | Sheep | 894 | 140–882 |
| MG682541 | IR25 | *nadh*1 | G3 | Camel | 894 | 140–882 |
| KY766894 | IR26 | *nadh*1 | G3 | Camel | 894 | 140–882 |
| MG672231 | IR27 | *nadh*1 | G1 | Sheep | 894 | 140–882 |
| MG672232 | IR28 | *nadh*1 | G1 | Sheep | 894 | 140–882 |
| MG672237 | IR29 | *nadh*1 | G1 | Sheep | 894 | 140–882 |
| MG672236 | IR30 | *nadh*1 | G1 | Sheep | 894 | 140–882 |
| MH050630 | CH21 | *nadh*1 | G3 | Human | 882 | 140–882 |
| MH050629 | CH22 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050628 | CH23 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050627 | CH24 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050626 | CH25 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050625 | CH26 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050624 | CH27 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050623 | CH28 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050622 | CH29 | *nadh*1 | G1 | Human | 882 | 140–882 |
| MH050621 | CH30 | *nadh*1 | G1 | Human | 882 | 140–882 |
| | PK50 | *nadh*1 | G3 | Human | 881 | 36–778 |
| | PK51 | *nadh*1 | G1 | Human | 834 | 9–751 |
| | PK52 | *nadh*1 | G1 | Human | 900 | 18–760 |
| | PK53 | *nadh*1 | G3 | Human | 844 | 7–749 |
| | PK54 | *nadh*1 | G1 | Human | 859 | 9–751 |
| | PK55 | *nadh*1 | G3 | Human | 867 | 14–756 |
| | PK56 | *nadh*1 | G3 | Human | 830 | 5–747 |
| | PK57 | *nadh*1 | G3 | Human | 850 | 13–755 |
| | PK58 | *nadh*1 | G3 | Human | 827 | 1–743 |
| | PK59 | *nadh*1 | G1 | Human | 843 | 4–745 |
| | PK60 | *nadh*1 | G3 | Human | 844 | 6–748 |

incorporating 29.8% (14/47) of the all studied isolates. This major haplotype included five (26.3%) Pakistani and six (60%) Chinese human isolates, one (10%) Iranian sheep isolate and two (25%) Indian buffalo isolates. Next, the main haplotype included five (50%) Iranian isolates (camel, $n = 4$; sheep, $n = 1$), one (5.3%) Pakistani human isolate and one (12.5%) Indian buffalo isolate. Within the mt-*co*1 haplotype network the frequency of unique singleton haplotypes were 61.4% (13/19), with the most of them from the Pakistani isolates ($n = 5$), followed by that from the Chinese ($n = 4$), Indian ($n = 3$) and Iranian ($n = 1$) (Fig. 5).

The analysed mt-*cytb* gene sequence length of the *E. granulosus s.s.* isolates was 547 bp. Twelve sites were polymorphic for the

**Table 2.** DnaSP output showing diversity and neutrality indices.

| DNA | *n* | *H* | *hd* ± S.D. | $\pi d$ ± S.D. | Tajima's *D* | *P* value | Fu's *F* | *P* value | FLD | *P* value | FLF | *P* value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| E.g.s.s *co*1 | 47 | 20 | 0.883 ± 0.035 | 0.00593 ± 0.00051 | −1.64218 | 0.10 > *P* > 0.05 | −14.106 | 0.000 | −3.40365 | *P* < 0.02 | −3.31798 | *P* < 0.02 |
| E.canadensis (G6/G7) *co*1 | 14 | 4 | 0.396 ± 0.159 | 0.00157 ± 0.00075 | −1.79759 | *P* < 0.05 | −1.640 | 0.121 | −2.27380 | *P* < 0.05 | −2.44883 | *P* < 0.05 |
| E.g.s.s *cytb* | 35 | 12 | 0.793 ± 0.055 | 0.00258 ± 0.00041 | −1.62945 | 0.10 > *P* > 0.05 | −7.178 | 0.001 | −2.94627 | *P* < 0.05 | −2.96924 | *P* < 0.05 |
| E.g.s.s *nadh*1 | 37 | 12 | 0.763 ± 0.058 | 0.00283 ± 0.00052 | −1.88146 | *P* < 0.05 | −4.375 | 0.009 | −3.71937 | *P* < 0.02 | −3.67423 | *P* < 0.02 |

*n*, Number of isolates; *H*, number of haplotypes; *hd*, haplotype diversity; $\pi d$, nucleotide diversity; S.D., standard deviation; FLD, Fu and Li's *D** statistical test; FLF, Fu and Li's F* statistical test.

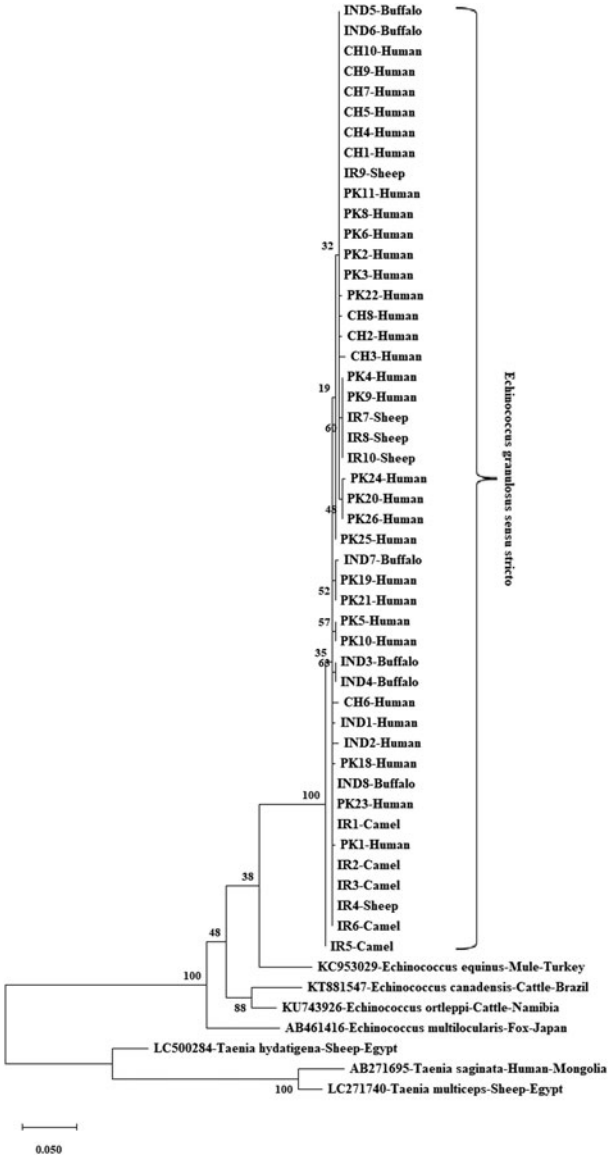


**Fig. 1.** Phylogenetic tree of *E. granulosus s.s.* isolates using the mt-*co*1 gene sequence. MEGA X was used to construct a maximum likelihood tree based on the HKY + I model. The reliability of the tree was assessed by 1000 bootstrap replications. *Echinococcus equinus* (GenBank accession no. KC953029), *E. canadensis* (GenBank accession no. KT881547), *E. ortleppi* (GenBank accession no. KU743926) and *E. multilocularis* (GenBank accession no. AB461416) were used as reference sequences, while *Taenia hydatigena* (GenBank accession no. LC500284), *T. saginata* (GenBank accession no. AB271695) and *T. multiceps* (GenBank accession no. LC271740) were used as outgroup sequences for phylogenetic tree construction.

mt-*cytb* sequence; 25% (3/12) were parsimony-informative. The haplotype network had 12 haplotypes, organized in a star-like orientattion, alongside of fundamental focal haplotype, isolated from different haplotypes by 1–6 mutational advances and incorporating 40% (14/35) of all the studied isolates. This main haplotype comprised nine (90%) Chinese isolates (sheep, $n = 6$; human, $n = 3$), three (30%) Iranian sheep isolates, and two (33.3%) Indian buffalo isolates. Next, the main haplotype included four (44.4%) Iranian isolates (camel, $n = 3$; sheep, $n = 1$), two (22.2%) Pakistani human isolates and two (33.3%) Indian buffalo isolates. There were 66.7% (8/12) unique singleton haplotypes within the mt-*cytb* haplotype network; the most were reported from Pakistan ($n = 4$), followed by Iran ($n = 3$) and China ($n = 1$) (Fig. 6).

For the *E. granulosus s.s.* mt-*nadh*1 sequence, 743-bp sequences were analysed. Twenty polymorphic sites for the mt-*nadh*1 sequence were detected; 20% (4/20) were parsimony-informative. Low nucleotide diversity and high

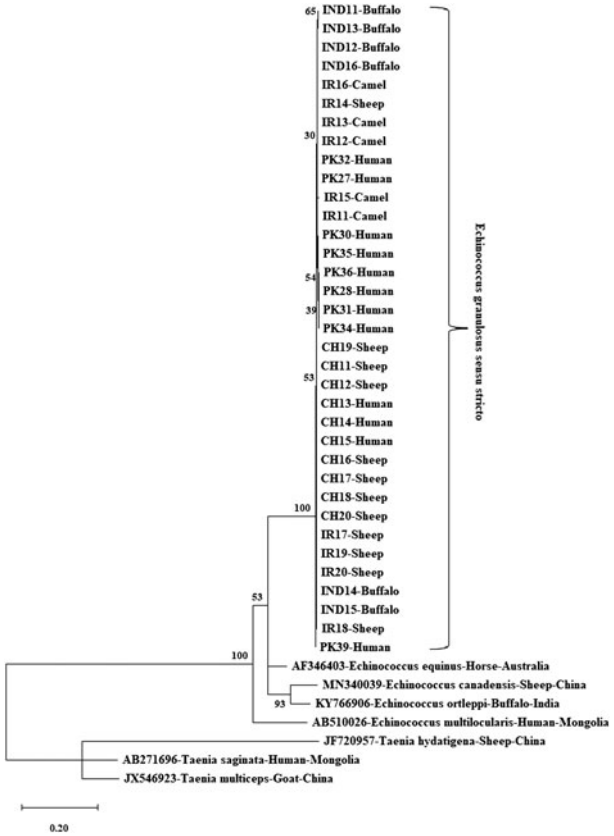


**Fig. 2.** Phylogenetic tree view of *E. granulosus s.s.* isolates using the mt-*cytb* gene sequence. MEGA X was used to construct a maximum likelihood tree based on the HKY + G + I model. The reliability of the tree was assessed by 1000 bootstrap replications. *Echinococcus equinus* (GenBank accession no. AF346403), *E. canadensis* (GenBank accession no. MN340039), *E. ortleppi* (GenBank accession no. KY766906) and *E. multilocularis* (GenBank accession no. AB510026) were used as reference sequences, while *T. hydatigena* (GenBank accession no. JF720957), *T. saginata* (GenBank accession no. AB271696) and *T. multiceps* (GenBank accession no. JX546923) were used as outgroup sequences for phylogenetic tree construction.

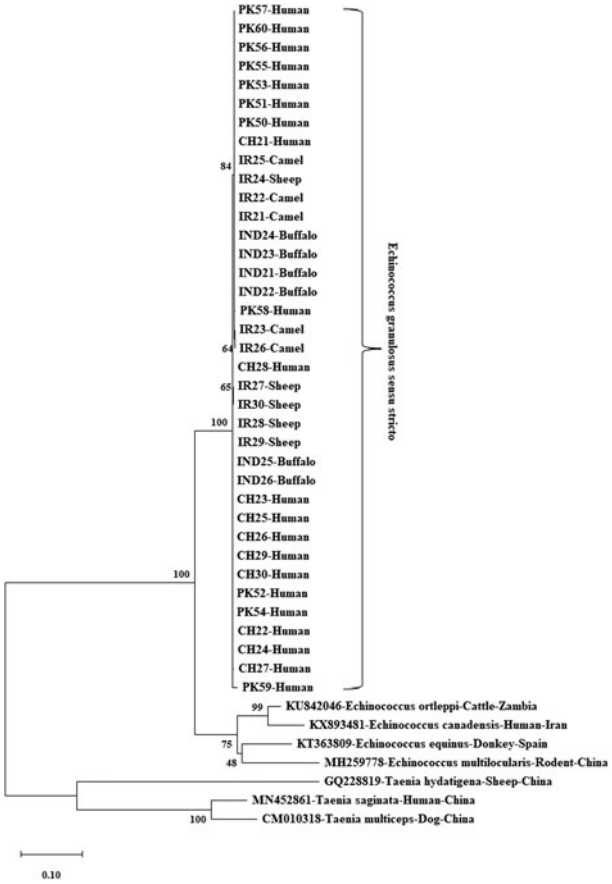


**Fig. 3.** Phylogenetic tree view of *E. granulosus s.s.* isolates using the mt-*nadh*1 gene sequence. MEGA X was used to construct a maximum likelihood tree based on the HKY + I model. The reliability of the tree was assessed by 1000 bootstrap replications. *Echinococcus equinus* (GenBank accession no. KT363809), *E. canadensis* (GenBank accession no. KX893481), *E. ortleppi* (GenBank accession no. KU842046) and *E. multilocularis* (GenBank accession no. MH259778) were used as reference sequences, while *T. hydatigena* (GenBank accession no. GQ228819), *T. saginata* (GenBank accession no. MN452861) and *T. multiceps* (GenBank accession no. CM010318) were used as outgroup sequences for phylogenetic tree construction.

haplotype diversity were observed. There were 12 haplotypes in the mt-*nadh*1 haplotype network, were organized in a star-like orientation, alongside of fundamental focal haplotype, isolated from different haplotypes by 1–11 mutational points, and comprised 29.8% (16/37) of all the studied isolates. The major haplotype have seven (43.8%) Pakistani human isolates, four (66.7%) Indian buffalo isolates, four (40%) Iranian isolates (camel, $n = 3$; sheep, $n = 1$) and one (10%) Chinese human isolate. 66.7% (8/12) singleton haplotypes were unique within the mt-*nadh*1 haplotype network, the most from China ($n = 4$), followed by Pakistan ($n = 2$) and Iran ($n = 1$) (Fig. 7).

After trimming the raw sequences, the 366-bp length of mt-*co*1 sequence of *E. canadensis* (G6/G7) was analysed. We detected four polymorphic sites without any parsimony-informative site. The mt-*co*1 haplotype network contained four haplotypes, with a central major haplotype separated by 1–3 mutational steps from the other haplotypes and which covered 78.6% (11/14) of all analyzed isolates. This haplotype have seven (26.3%) Iranian isolates (human, $n = 1$; camel, $n = 6$), two (100%) Pakistani isolates and two (100%) Chinese human isolates. There were 75% (3/4) unique singleton haplotypes within the mt-*co*1 haplotype network, with the most reported from the Iranian isolates ($n = 3$) (Fig. 8).

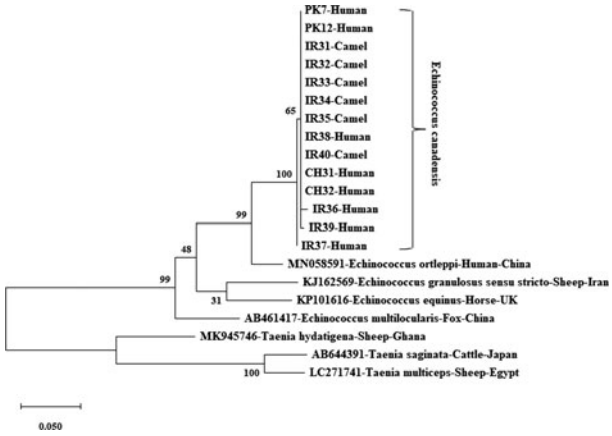


**Fig. 4.** Phylogenetic tree view of *E. canadensis* (G6/G7) isolates using the mt-*co*1 gene sequence. MEGA X was used to construct a maximum likelihood tree based on the HKY + I model. The reliability of the tree was assessed by 1000 bootstrap replications. *Echinococcus equinus* (GenBank accession no. KP101616), *E. granulosus s.s.* (GenBank accession no. KJ162569), *E. ortleppi* (GenBank accession no. MN058591) and *E. multilocularis* (GenBank accession no. AB461417) were used as reference sequences, while *T. hydatigena* (GenBank accession no. MK945746), *T. saginata* (GenBank accession no. AB644391) and *T. multiceps* (GenBank accession no. LC271741) were used as outgroup sequences for phylogenetic tree construction.

## Discussion

Generally, *E. granulosus s.s.* is one of the most common genotypes all over the world (Alvarez *et al.*, 2014; Kinkar *et al.*, 2018). The mt-*co*1 gene can be used to differentiate inter- and intra-specific

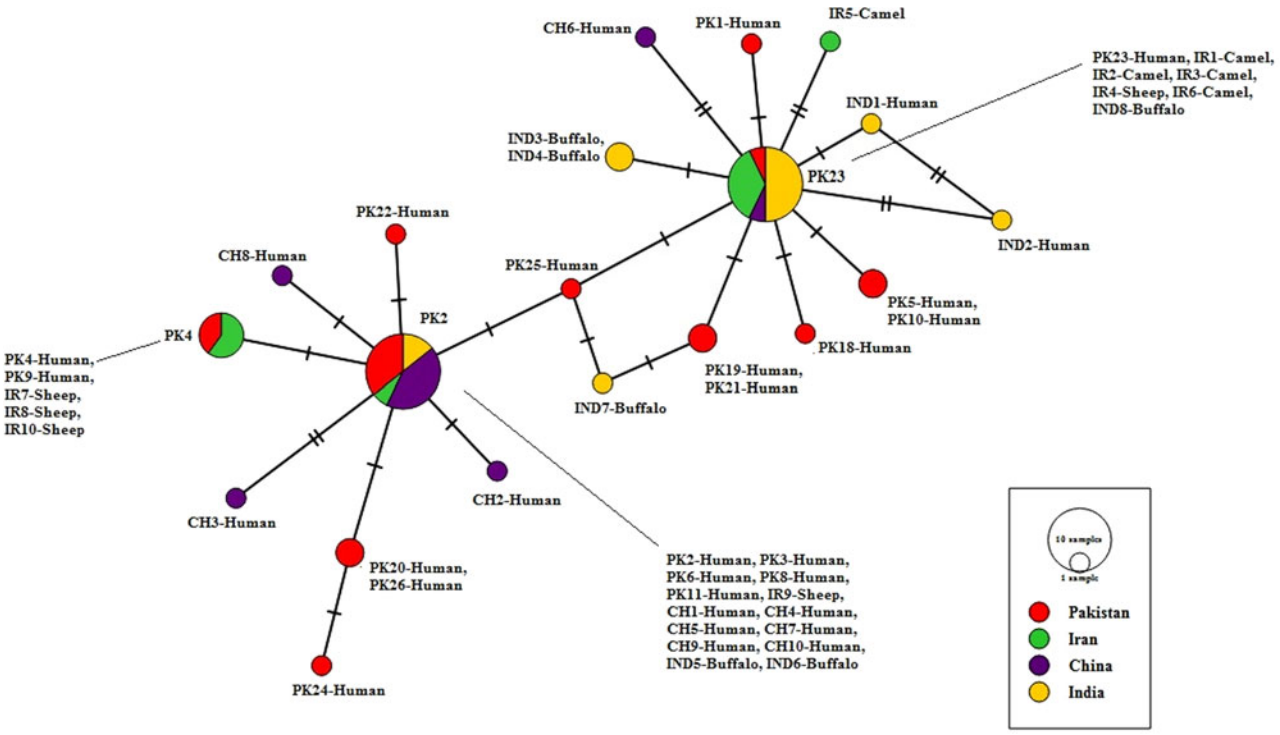


**Fig. 5.** Haplotype network of *E. granulosus s.s.* mt-*co*1 gene. Different colours indicate the geographical distribution of the haplotypes. Circle sizes are related to haplotype frequency. Hatch marks show the number of mutations distinguishing the haplotypes.

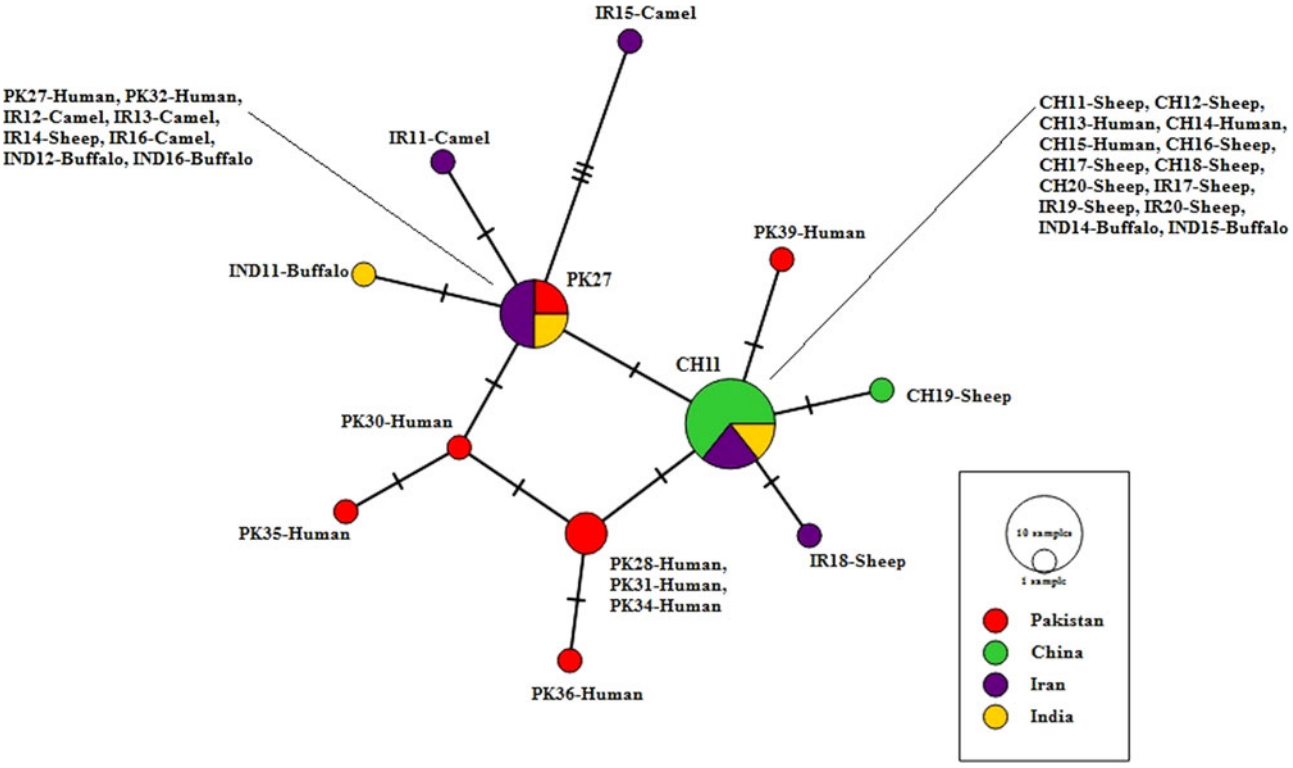


**Fig. 6.** Haplotype network of *E. granulosus s.s.* mt-*cytb* gene. Different colours indicate the geographical distribution of the haplotypes. Circle sizes are related to the haplotype frequency. Hatch marks show the number of mutations distinguishing the haplotypes.

variants in evolutinary terms (Nakao *et al.*, 2013). The *Echinococcus* spp. reproduction system is unique because adult worms are hermaphrodites, and their asexual reproduction occurs at the larval stage. This reproduction method might lead to many genetic mutations and result in variations at genus and species level (Smyth and Smyth, 1969), and is supported by the occurence of two different haplotypes in a single intermediate host (Casulli *et al.*, 2012).

Here, the *E. granulosus s.s.* haplotype network in the mt-*co*1 region revealed that most sequences were of Chinese origin

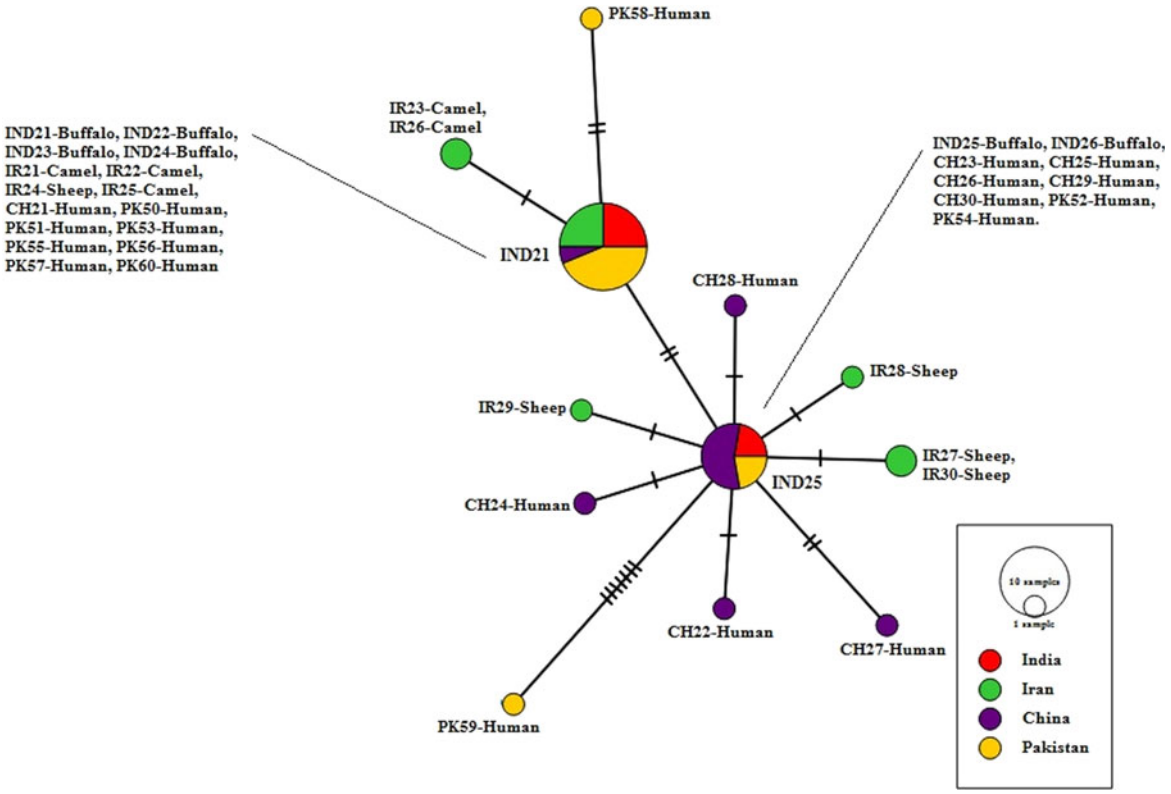


**Fig. 7.** Haplotype network of *E. granulosus s.s.* mt-*nadh*1 gene. Different colours indicate the geographical distribution of the haplotypes. Circle sizes are related to the haplotype frequency. Hatch marks show the number of mutations distinguishing the haplotypes.

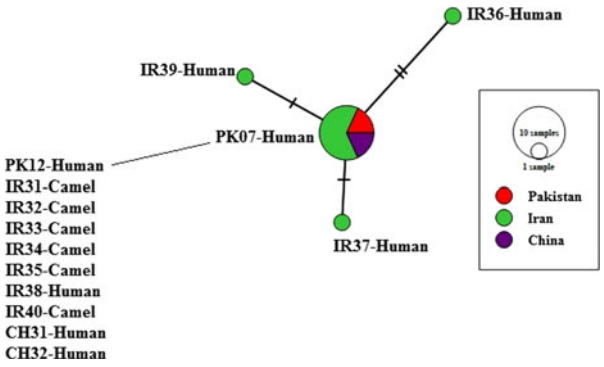


**Fig. 8.** Haplotype network of *E. canadensis* (G6/G7) mt-*co*1 gene. Different colours indicate the geographical distribution of the haplotypes. Circle sizes are related to the haplotype frequency. Hatch marks show the number of mutations distinguishing the haplotypes.

(28.6%, 6/14). Following this, the main haplotype included isolates from Pakistan ($n = 5$), India ($n = 2$) and Iran ($n = 1$). There were 65% (13/20) unique singleton haplotypes within the mt-*co*1 haplotype network, with the most reported from the Pakistani isolates ($n = 5$), followed by that of the Chinese ($n = 4$), Indian ($n = 3$) and Iranian isolates ($n = 1$), while the *E. canadensis* (G6/G7) haplotype network in the mt-*co*1 region revealed sequences that were mostly of Iranian origin (63.6%, 7/11) in the main haplotype. There were three single haplotypes in the *E. canadensis* (G6/G7) isolates, all belonging to the Iranian isolates.

In the *E. granulosus s.s.* mt-*cytb* region, the main haplotype comprised 90% (9/10) of Chinese isolates, and three Iranian isolates and two Indian isolates. There were 66.7% (8/12) unique singleton haplotypes within the mt-*cytb* haplotype network, with the most reported from Pakistan, followed by Iran and China. Analysis of the *E. granulosus s.s.* mt-*cytb* sequence was restricted to BLAST search because sample size was small. Wang *et al.* (2014*a*, *b*) collected 45 hydatid cysts from human, sheep and yak hosts in China, sequenced the *cytb* gene, and confirmed 10 haplotypes from the 45 isolates. The results indicate that each haplotype has specificity for its intermediate host. In that study, haplotype diversity was observed as 0.626, while nucleotide diversity was noted 0.001 (Zhong *et al.*, 2014). Similarly, we detected 12 sites that were polymorphic for the mt-*cytb*, out of which 25% (3/12) sites were parsimony-informative. A total of 12 haplotypes were observed in the haplotype network, with a central main haplotype separated by 1–6 mutational steps from the other haplotypes.

The haplotype network of the *E. granulosus s.s.* mt-*nadh*1 region revealed sequences that were mostly of Pakistan origin (43.8%, 7/16). The other elements of this main haplotype were the Indian (66.7%, 4/6), Iranian (40%, 4/10) and Chinese isolates 10% (1/10). The findings show that the same haplotype contained a diversity of both geography and species such as human, buffalo, sheep and camel, and that there was a close relationship between them. On the contrary, there were 66.7% (8/12) unique singleton haplotypes in the mt-*nadh*1 haplotype network, most of which were from the human isolates of China ($n = 4$), followed by that of Pakistan ($n = 2$) and sheep isolates of Iran ($n = 2$). Genetically different *E. granulosus s.s.* variants were obtained in the same host, which is evident by all these mutations. The presence of the different variants also suggests that out-crossing between different adult worms might have occurred in the same final host.

The nucleotide diversity and total haplotype values were 0.00157 and 0.396, respectively. They were relatively low for the mt-*co*1 gene from the *E. canadensis* (G6/G7) analyzed isolates. These results show that *E. canadensis* (G6/G7) haplotypes are not genetically diverse. In the current

scenario, 7 out of 11 isolates were of Iranian origin while Pakistani and Chinese isolates were also located in the main haplotype.

The mt-*co*1 gene is one of the best options for analysis to determine the genetic variability of the different *E. granulosus s.s.* isolates (Kamenetzky *et al.*, 2002; Haag *et al.*, 2004; Badaraco *et al.*, 2008; Nakao *et al.*, 2010). The mt-*co*1 gene of the *E. granulosus s.s.* isolates had 0.883 and 0.00593 of haplotype and nucleotide diversity, respectively, which were relatively higher as compared to that of the other genes. In this case, it is not surprising that 20 haplotypes were identified within 47 isolates. The higher haplotype variations in the mt-*co*1 gene showed a star-shaped network with a centrally placed main haplotype (EgUK01), as previously reported (Boufana *et al.*, 2015). Similarly, there is high haplotype variability in the *E. granulosus s.s.* G1 genotype, and 171 haplotypes have been found in 212 samples (Kinkar *et al.*, 2018). Excessive genetic diversity of *E. granulosus s.s.* was observed by using sequences of shorter length (Casulli *et al.*, 2012; Rojas *et al.*, 2016). Previously, *E. granulosus s.s.* (G1 and G3) was reported from Italian and Eastern European populations, where 21 and four haplotypes were confirmed in Eastern Europe and Italy, respectively (Casulli *et al.*, 2012).

The overall outcomes of the present study clearly indicate low nucleotide diversity and high haplotype diversity in general. Tajima's *D* was recorded negative for all analyzed sequences, indicating population expansion and/or purifying selection. Recent population expansion or hitchhiking in *E. granulosus s.s.* can be explained by the significantly highest negative Fu's *F* values in sequence comparison of mt-*co*1, which showed the presence of rare haplotypes. Fu's *F* test was developed based on haplotype or allele distribution. The existence of unmatched single haplotypes for the mt-*co*1 (13/20) sequences of *E. granulosus s.s.* is concordant with the structure of the haplotype networks.

The negative values of both tests show the excess of rare mutations present in a population (Sharma *et al.*, 2013). On the other hand, a common founder, i.e. an *E. granulosus* haplotype (EgUK01), along with negative neutrality, show evidence of a recent history of bottleneck or founder events among populations and their expansion. The similarity in the *E. granulosus s.s.* genomes from different geographical regions indicate the presence of a single lineage globally. The most probable reason for population expansion might be the anthropogenic movement of animal hosts (Nakao *et al.*, 2010). However, the *E. granulosus s.s. co*1 gene (828 bp) shows considerable genetic variation, and the pair-wise fixation index for the *E. granulosus s.s. co*1 gene from donkeys was significantly higher compared to that of other intermediate hosts. That study also indicated that the *E. granulosus* donkey subpopulations are distinct from the subpopulations of other hosts. The statistically significant Fs values of the donkey *E. granulosus* isolates as compared to that of other hosts indicates the presence of surplus variable alleles, further indicating the limited gene flow (Boufana *et al.*, 2015).

Our findings for Pakistan are in line with that of other studies. In Tunisian isolates, the low nucleotide diversity coupled with high haplotype diversity in *E. granulosus s.s.* was comparable to other countries such as China, Europe, Iran and Jordan (Nakao *et al.*, 2010; Casulli *et al.*, 2012; Yanagida *et al.*, 2012), which indicates a rapid demographic expansion (Avise, 2000). In the present study, Tajima's *D* values were negative which indicates a bias towards the presence of nucleotide variations, a feature of recent population increase. The negative Fu's *F* values for most of the *E. granulosus s.s.* subpopulations indicate bottleneck and/or purifying selection event that might have happened in the past, with similar findings reported from China and Jordan (Nakao *et al.*, 2010; Yanagida *et al.*, 2012). Interestingly, it has been postulated that *E. granulosus* genetic diversity in the Middle East is the result of parasitic shift that might have occured from wild to domestic hosts (Yanagida *et al.*, 2012).

## Conclusion

We used three mt-DNA markers to investigate the population genetic structure of *E. granulosus s.s.* and *E. canadensis* (G6/G7) sequences from livestock and human species. The analysis of the mt-*co*1, mt-*cytb* and mt-*nadh*1 gene fragments showed that the parasite's genetic diversity in Pakistan, China, Iran and India has increased continually. However, further studies are required to confirm this conclusion. In addition, it is important to search and track sources of infection, using adult parasites obtained from different final host species, to analyse the evolutionary history and to determine the pathogenicity of *E. granulosus s.l.*

## Data

All relevant data are within the manuscript.

**Acknowledgements.** The authors would like to thank Professor Wenqin Wang at Shanghai Jiaotong University for help with the bioinformatics analysis, Dr Wang Yanjuan, Dr Jianhai Yin, and Dr Ying Wang at the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, for technical guidance.

**Author contributions.** AK, SS, FC, KS, MSA, and MRK performed the bioinformatic analyses, evaluated the results, and drafted the paper. HA and JC participated in the conception and design of the study. HL and YS contributed to the data analysis and revised the manuscript. FC and SS contributed to the data analysis. All authors have read and approved the final manuscript.

**Financial support.** This study was supported by the National Natural Science Foundation of China (grant nos. 81772225 and 81971969 to JC), the Fifth Round of Three-Year Public Health Action Plan of Shanghai (No. GWV-10.1-XK13 to JC), and the Laboratory of Parasite and Vector Biology, National Health Commission of the People's Republic of China (grant no. WSBKFKT2017-01 to AK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflict of interest.** None.

**Ethical standards.** Not applicable.

**Consent to publication.** Not applicable.

## References

**Addy F, Wassermann M, Kagendo D, Ebi D, Zeyhle E, Elmahdi IE, Umhang G, Casulli A, Harandi MF, Aschenborn O, Kern P, Mackenstedt U and Romig T** (2017) Genetic differentiation of the G6/7 cluster of *Echinococcus canadensis* based on mitochondrial marker genes. *International Journal for Parasitology* **47**, 923–931.

**Alvarez Rojas CA, Romig T and Lightowlers MW** (2014) *Echinococcus granulosus sensu lato* genotypes infecting humans – review of current knowledge. *International Journal for Parasitology* **44**, 9–18.

**Avise JC** (2000) *Phylogeography: The History and Formation of Species*. Cambridge, MA, USA: Harvard University Press.

**Badaraco JL, Ayala FJ, Bart JM, Gottstein B and Haag KL** (2008) Using mitochondrial and nuclear markers to evaluate the degree of genetic cohesion among *Echinococcus* populations. *Experimental Parasitology* **119**, 453–459.

**Bandelt HJ, Forster P and Röhl A** (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48.

**Boufana B, Lett WS, Lahmar S, Buishi I, Bodell AJ, Varcasia A, *et al.*** (2015) *Echinococcus equinus* and *Echinococcus granulosus sensu stricto* from the United Kingdom: genetic diversity and haplotypic variation. *International Journal for Parasitology* **45**, 161–166.

**Butt A and Khan JA** (2020) Cystic echinococcosis: a 10-year experience from a middle-income country. *Tropical Doctor* **50**, 117–121.

**Casulli A, Interisano M, Sreter T, Chitimia L, Kirkova Z, Rosa GL, *et al.*** (2012) Genetic variability of *Echinococcus granulosus sensu stricto* in Europe inferred by mitochondrial DNA sequences. *Infection, Genetics and Evolution* **12**, 377–383.

**Fomda BA, Khan A, Thokar MA, Malik AA, Fazili A, Dar RA, Sharma M and Malla N** (2015) Sero-epidemiological survey of human cystic echinococcosis in Kashmir, North India. *PLoS ONE* **10**, e0124813.

**Haag KL, Alves-Junior L, Zaha A and Ayala FJ** (2004) Contingent, non-neutral evolution in a multicellular parasite: natural selection and gene conversion in the *Echinococcus granulosus* antigen B gene family. *Gene* **333**, 157–167.

**Haleem S, Niaz S, Qureshi NA, Ullah R, Alsaid MS, Alqahtani AS, *et al.*** (2018) Incidence, risk factors, and epidemiology of cystic echinococcosis: a complex socioecological emerging infectious disease in Khyber Pakhtunkhwa, province of Pakistan. *Biomed Research International* **2018**, 5042430. doi: 10.1155/2018/5042430.

**Kamenetzky L, Gutierrez AM, Canova SG, Haag KL, Guarnera EA, Parra A, *et al.*** (2002) Several strains of *Echinococcus granulosus* infect livestock and humans in Argentina. *Infection Genetics and Evolution* **2**, 129–136.

**Khan A, Ahmed H, Khan H, Saleem S, Simsek S, Brunetti E, Afzal MS, Manciulli T and Budke CM** (2020) Cystic echinococcosis in Pakistan: a review of reported cases, diagnosis, and management. *Acta Tropica* **212**, 105709.

**Kinkar L, Laurimae T, Simsek S, Balkaya I, Casulli A, Manfredi MT, *et al.*** (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.

**Kinkar L, Laurimae T, Sharbatkhori M, Mirhendi H, Kia EB, Ponce-Gordo F, *et al.*** (2017) New mitogenome and nuclear evidence on the phylogeny and taxonomy of the highly zoonotic tapeworm *Echinococcus granulosus sensu stricto*. *Infection Genetics and Evolution* **52**, 52–58.

**Kinkar L, Laurimäe T, Acosta-Jamett G, Andresiuk V, Balkaya I, Casulli A, *et al.*** (2018) Global phylogeography and genetic diversity of the zoonotic tapeworm *Echinococcus granulosus sensu stricto* genotype G1. *International Journal for Parasitology* **48**, 729–742.

**EFSA Panel on Biological Hazards (BIOHAZ)**, Koutsoumanis K, Allende A, Alvarez-Ordóñez A, Bolton D, Bover-Cid S, Chemaly M, Davies R, De Cesare A, Herman L, Hilbert F, Lindqvist R, Nauta M, Peixe L, Ru G, Simmons M, Skandamis P, Suffredini E, Cacciò S, Chalmers R, Deplazes P, Devleesschauwer B, Innes E, Romig T, van der Giessen J, Hempen M, Van der Stede Y and Robertson L (2018) Public Health Risks Associated with Food-Borne Parasites. EFSA Journal 16, e05495.

**Kumar S, Stecher G, Li M, Knyaz C and Tamura K** (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547–1549.

**Leigh J, Bryant D and Steel M** (2015) PopART (Population Analysis with Reticulate Trees) (Technical report, University of Otago, New Zealand, http://popart.otago.ac.nz).

**Moudgil AD, Moudgil P, Asrani RK and Agnihotri RK** (2019) Hydatidosis in slaughtered sheep and goats in India: prevalence, genotypic characterization and pathological studies. *Journal of Helminthology* **94**, e27.

**Nakao M, Lavikainen A, Yanagida T and Ito A** (2013) Phylogenetic systematics of the genus *Echinococcus* (Cestoda: Taeniidae). *International Journal for Parasitology* **43**, 1017–1029.

**Nakao M, Li T, Han X, Ma X, Xiao N, Qiu J, *et al.*** (2010) Genetic polymorphisms of *Echinococcus* tapeworms in China as determined by mitochondrial and nuclear DNA sequences. *International Journal for Parasitology* **40**, 379–385.

**Rojas CAA, Fredes F, Torres M, Acosta-Jamett G, Alvarez JF, Pavletic C, *et al.*** (2016) First meeting cystic echinococcosis in Chile, update in alternatives for control and diagnostics in animals and humans. *Parasites & Vectors* **9**, 502.

**Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE and Sánchez-Gracia A** (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**, 3299–3302.

**Shahnazi M, Hejazi H, Salehi M and Andalib AR** (2011) Molecular characterization of human and animal *Echinococcus granulosus* isolates in Isfahan, Iran. *Acta Tropica* **117**, 47–50.

**Sharbatkhori M, Mirhendi H, Harandi MF, Rezaeian M, Mohebali M, Eshraghian M and Kia EB** (2010) *Echinococcus granulosus* genotypes in livestock of Iran indicating high frequency of G1 genotype in camels. *Experimental Parasitology* **124**, 373–379.

**Sharma M, Fomda BA, Mazta S, Sehgal R, Singh BB and Malla N** (2013) Genetic diversity and population genetic structure analysis of *Echinococcus granulosus sensu stricto* complex based on mitochondrial DNA signature. *PLoS ONE* **8**, e82904.

**Singh B, Dhand N, Ghatak S and Gill JPS** (2013) Economic losses due to cystic echinococcosis in India: need for urgent action to control the disease. *Preventive Veterinary Medicine* **113**, 1–12.

**Singh BB, Sharma JK, Tuli A, Sharma R, Bal MS, Aulakh RS and Gill JPS** (2014) Prevalence and morphological characterisation of *Echinococcus granulosus* from north India. *Journal of Parasitic Diseases* **38**, 36–40.

**Smyth JD and Smyth MM** (1969) Self insemination in *Echinococcus granulosus* in vivo. *Journal of Helminthology* **43**, 383–388.

**Spotin A, Mahami-Oskouei M, Harandi MF, Baratchian M, Bordbar A, Ahmadpour E and Ebrahimi S** (2017) Genetic variability of *Echinococcus granulosus* complex in various geographical populations of Iran inferred by mitochondrial DNA sequences. *Acta Tropica* **165**, 10–16.

**Wang J, Wang N, Hu D, Zhong X, Wang S, Gu X and Yang G** (2014a) Genetic diversity of *Echinococcus granulosus* in southwest China determined by the mitochondrial NADH dehydrogenase subunit 2 gene. *The Scientific World Journal*, Volume 2014, 867839. doi: 10.1155/2014/867839.

**Wang Q, Huang Y, Huang L, Yu W, He W, Zhong B, *et al.*** (2014b) Review of risk factors for human echinococcosis prevalence on the Qinghai-Tibet Plateau, China: a prospective for control options. *Infectious Diseases of Poverty* **3**, 3.

**Wen H, Vuitton L, Tuxun T, Li J, Vuitton DA, Zhang W and McManus DP** (2019) Echinococcosis: advances in the 21st century. *Clinical Microbiology Reviews* **32**, e00075–18.

**Yanagida T, Mohammadzadeh T, Kamhawi S, Nakao M, Sadjjadi SM, Hijjawi N, *et al.*** (2012) Genetic polymorphisms of *Echinococcus granulosus sensu stricto* in the Middle East. *Parasitology International* **61**, 599–603.

**Zhang W, Zhang Z, Wu W, Shi B, Li J, Zhou X, *et al.*** (2015) Epidemiology and control of echinococcosis in Central Asia, with particular reference to the People's Republic of China. *Acta Tropica* **141**, 235–243.

**Zhong X, Wang N, Hu D, Wang J, Liu T, Gu X, Wang S, Peng X and Yang G** (2014) Sequence analysis of cytb gene in *Echinococcus granulosus* from Western China. *Korean Journal of Parasitology* **52**, 205–209.