

Molecular characterization of *Echinococcus granulosus* from Peru by sequencing of the mitochondrial cytochrome C oxidase subunit 1 gene

Elizabeth Sánchez^{1,3/+}, Omar Cáceres¹, César Náquira¹, David Garcia¹, Gladys Patiño², Herrera Silvia¹, Aline C Volotão³, Octavio Fernandes³

¹Instituto Nacional de Salud, Lima, Peru ²Hospital Hipólito Unanue, Lima, Peru ³Instituto Oswaldo Cruz-Fiocruz, Rio de Janeiro, RJ, Brasil

Echinococcus granulosus, the etiologic agent of cystic echinococcosis (CE) in humans and other animal species, is distributed worldwide. Ten intra-specific variants, or genotypes (G1-G10), have been defined based on genetic diversity. To determine the genotypes present in endemic areas of Peru, samples were collected from cattle (44), sheep (41) and humans (14) from Junín, Puno Huancavelica, Cusco, Arequipa and Ayacucho. DNA was extracted from protoscolex and/or germinal layers derived from 99 *E. granulosus* isolates and used as templates to amplify the mitochondrial cytochrome C oxidase subunit 1 gene. The resulting polymerase chain reaction products were sequenced and further examined by sequence analysis. All isolates, independent of the host, exhibited the G1 genotype. Phylogenetic analysis showed that three isolates from Ayacucho shared the same cluster with microvariant G1⁴. The G1 genotype is considered the most widespread and infectious form of *E. granulosus* worldwide and our results confirm that the same patterns apply to this country. Therefore, these findings should be taken into consideration in developing prevention strategies and control programs for CE in Peru.

Key words: *Echinococcus granulosus* - cytochrome C oxidase subunit 1 - genotypes - Peru

Echinococcus granulosus (Batsch 1786) is a cestode, capable of infecting humans and several other animal species worldwide. The adult form resides mainly in dogs, however, the larval form (hydatid cyst) is found in a wide variety of mammals, including humans. The presence of cysts in an animal leads to cystic echinococcosis (CE). In humans, the disease is considered a critical public health problem. In addition to human health concerns, infections in cattle breeding areas result in economic losses (Thompson 2008). The disease is endemic in many South American countries, including Uruguay, Argentina, Chile, Brazil, Peru and Bolivia (Kamenetzky et al. 2002). The impact of CE on human health makes it one of the most important zoonotic diseases.

In Peru, CE is endemic mainly within the cattle producing areas in the central-southern regions of the country (Pérez 2007). The incidence of human cases of CE in Peru is approximately 7-11/100,000 (MS 2005), but it has been reported that in the regions of Junín, Puno, Huancavelica, Cusco, Arequipa and Ayacucho, the incidence varies between 11-39/100,000. For animal cases in these areas, the prevalence varies from 1-75% (Pérez 2007) and is primarily associated with the circulation of the parasite between sheep and dogs (Náquira 1993).

As a species, *E. granulosus* comprises a number of intraspecific variants, strains or genotypes at the genetic level (Thompson & McManus 2001, McManus & Thompson 2003, Nakao et al. 2007, 2010). Currently, 10 distinct genotypes (G1-G10) have been described worldwide, based on nucleotide sequence analysis of the mitochondrial cytochrome C oxidase subunit 1 (COI) and NADH dehydrogenase 1 genes. These different genotypes have been associated with distinct, intermediate hosts: sheep, pigs, cattle, horses, camels, goats and cervids (Bowles et al. 1992, 1995, Bowles & McManus 1993a, b, c, Scott et al. 1997, Zhang et al. 1998, McManus 2002, Lavikainen et al. 2003, Busi et al. 2007, Moks et al. 2008).

In South America, molecular studies have demonstrated the presence of several variants of *E. granulosus*. In Argentina, G1 and G6 have been isolated from human, G2 from sheep and humans and G7 from pigs (Rosenzvit et al. 1999, Kamenetzky et al. 2002). In Chile, similar studies revealed G1 and G6 in human (Manterola et al. 2008). The G1 and G6 variant have been characterized from humans in Peru (Santivañez et al. 2008). The G1 genotype has also been isolated from sheep, cattle and goats in Peru. In addition, other variants were identified, including G7 from pigs and G6 from goats (Moro et al. 2009).

It has been suggested that the extensive, intra-specific genetic variation of *E. granulosus* may be better understood within the context of variations within the life cycle pattern (Thompson & McManus 2001). In addition, other biological consequences have been suggested to be linked to the different genotypes, such as rate of development, antigenicity, transmission profiles, sensitivity to chemotherapeutic agents and pathologic patterns (Thompson & Lymbery 1988, Thompson et al. 1995, Thompson & McManus 2001, 2002).

Financial support: CAPES (to ES)

+ Corresponding author: esanchez@ioc.fiocruz.br

Received 25 February 2010

Accepted 28 July 2010

All of these aspects should be considered in developing vaccines, diagnostic tests and pharmacological therapies for CE. Therefore, due to epidemiological implications and for the design of control strategies, it is essential that circulating *E. granulosus* genotypes in a given endemic area are clearly defined (McManus & Bowles 1996, McManus & Thompson 2003). In this study, the molecular characterization of 99 *E. granulosus* isolates from humans and other animals was performed by sequence analysis of a fragment of the mitochondrial CO1 gene to define the circulating genetic variants in hyper-endemic areas of Peru.

MATERIALS AND METHODS

Study area - Hydatid cysts were collected from sacrificed animals (sheep and dairy cattle) in official abattoirs from high prevalence areas of CE in the mountain regions of Peru (Fig. 1). The areas are known as breeding zones for dairy cattle and sheep (intermediate hosts). Moreover shepherd dogs also live in this area and are definitive hosts for parasite. In addition, human isolates were collected from hydatid cysts from patients attending different hospitals in Lima. All the samples were parasitologically confirmed at the Parasitology National Reference Laboratory [National Institute of Health (NIH) Lima, Peru].

Samples collections - All hydatid cysts were obtained under aseptic conditions. Cyst contents were aspirated and examined under light microscopy to confirm the presence of protoscolex and/or germinal layers that were washed three times with phosphate buffered saline by centrifugation at 2,000 g for 15 min. The obtained pellets were kept at 95% ethanol at -20°C until DNA extraction.

DNA extraction and polymerase chain reaction (PCR) - *E. granulosus* genomic DNA was extracted from protoscolex and/or germinal layers using the QIAmp DNA Mini Kit (QIAGEN), following the manufacturer's instructions. DNA was stored at -20°C until used for PCR amplification. A region of 450 bp of the mitochondrial CO1 gene was amplified using the following primers: 5' TTTTGGGGCATCCTGAGGTTTAT 3' and 5' TAAAGAAAGAACATAATGAAAATG 3' (Bowles et al. 1992). The reactions were carried out in a final volume of 50 µL containing approximately 50 ng of genomic DNA in the presence of appropriate buffer with 200 µM of each dNTPs, 2.5 mM MgCl₂, 50 pmol of each primer and 1.5 units of *Taq* DNA polymerase (Invitrogen). The positive control was DNA from *E. granulosus* genotype G1 (kindly provided by Dr Mara Rosenzvit, Administración Nacional de Laboratorios e Institutos de Salud, Argentina) and negative controls were DNA isolated from *Giardia duodenalis* trophozoites (reference strain WB) and water. PCR reactions were performed in all samples with positive and negative controls. The thermal profile of the PCR reaction was an initial 3 min at 95°C, followed by 35 cycles of 60 sec at 95°C, 60 sec at 56°C and 90 sec at 72°C, with a final incubation at 72°C for 3 min in a Gene Amp PCR Systems 9700 (Applied Biosystems). The PCR products were analyzed with 1% agarose gel electrophoresis by ethidium bromide staining. An image of the gel was recorded under UV light using a gel documentation system (BioRad).



Fig. 1: geographic origins of *Echinococcus granulosus* isolates. A map of Peru displaying the regions in the country where samples were collected (gray regions). Inset shows location of Peru in relationships to all of South America.

Sequence analysis - PCR products were purified for sequencing using GFX PCR DNA Kit and Gel Band Purification Kit (Amersham Biosciences), following the manufacturer's instructions. The purified PCR products were sequenced using the BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems) and analyzed by the genomic platform of DNA sequencing PDIIS/Fiocruz in an automatic sequencer ABI 3730 DNA Analyzer (Applied Biosystems), following the protocol described elsewhere (Otto et al. 2008). The chromatograms were analyzed and the nucleotide sequences obtained were aligned using the ClustalW method of the program MEGA 4.1 (megasoftware.net), using the sequences of the different genotypes of *E. granulosus* (G1, G2, G3, G4, G5, G6 and G7) and the microvariants of the G1 genotype (G1A, G1B, G1C, G1D, G1E and G1¹, G1², G1³, G1⁴) deposited in the GenBank as references [accessions M84661/62/63/64/65/66/67 (Bowles et al. 1992), AF458871/72/73/74/75 (Kamenetzky et al. 2002) and EF393619, EF595654, EU178103, EU178104 (Vural et al. 2008)].

Phylogenetic analysis - Phylogenetic analyses were based on alignments obtained from ClustalW of a 375 bp sequence and carried out using MEGA v4.1. The phylogenetic tree was constructed using the Neighbor-Joining algorithm (Saitou & Nei 1987) with Kimura-2 parameters. To determine the robustness of the tree, Bootstrap analysis of 1,000 replicates was applied. A correspondent nucleotide sequence of *Taenia crassiceps* (GenBank accession AB03341) was used as an outgroup.

Ethic statement - This study was reviewed and approved by the Ethical Committee of NIH.

TABLE
Geographic and host origin of Peruvian isolates of *Echinococcus granulosus* and the fertility of the cysts

Geographic origin	<i>E. granulosus</i> G1						Total
	Sheep	Fertility %	Cattle	Fertility %	Human	Fertility %	
Puno	8	62.5	7	28.6	3	100	18
Junin	18	83.3	11	27.3	5	80	34
Cusco	0	0	9	44	1	100	10
Arequipa	2	100	12	0	1	100	15
Huancavelica	1	0	5	20	2	100	8
Ayacucho ^a	12	100	0	0	2	100	14
Total	41	82.9	44	22.7	14	92.85	99

a: include G1⁴ microvariants of the G1 genotype.

RESULTS

Host and geographical origins of parasites - Samples of *E. granulosus* cysts were obtained from six cattle producing regions of southern Peru (Fig. 1). All samples collected from sheep and cattle were obtained from animals killed in abattoirs during a one-week period from each region. Isolates were prepared in the place collected to ensure proper designation of origin. Human isolates were collected from hydatid cysts of patients attending different hospitals in Lima for the treatment of the disease. The samples were subsequently sent for parasitological confirmation at the National Reference Laboratory, NIH. Every sample was examined by light microscopy to confirm the fertility of the cysts (presence of protoscolex). The Table lists the geographical and host origins of all the isolates collected for this study, as well as the fertility of the cysts.

Amplification of CO1 and sequence analysis - The CO1 gene of *E. granulosus* has been routinely utilized for variant designation, with specific primers identified for PCR amplification (Bowles et al. 1992). In our hands, these primers generated the expected 450 bp product after PCR amplification using all 99 isolates of *E. granulosus* as templates. Individual PCR products were further purified after gel electrophoresis and subjected to sequencing in duplicate.

The sequences obtained were initially analyzed by multiple alignments with reported reference sequences for the G1 genotype of *E. granulosus*. Ninety-six analyzed sequences were demonstrated to be 100% identical to the common sheep G1 genotype (Bowles et al. 1992). Three sequences isolated from Ayacucho exhibited a single nucleotide polymorphism, a thymine residue instead of a cytosine, generating a non-synonymous change in the corresponding amino acid sequence (A/V). These sequences were identical to the G1⁴ microvariant of the G1 genotype (Vural et al. 2008). Fig. 2 depicts a subset of sequences of the CO1 gene from the different isolates, compared to the G1 reference sequence. The obtained sequences were submitted to the GenBank database with the accession GU233854-GU233952.

Phylogenetic analysis - Phylogenetic analysis showed a robust tree clustering all samples as G1 genotype with a strong bootstrap (1,000 replicates). A particular group, composed of three sequences corresponding to G1⁴ microvariants, from a single geographic region is also depicted (Fig. 3). Bootstrap values are shown at relevant nodes. The evolutionary distance between the groups is very short, suggesting that the genetic divergence is recent.

DISCUSSION

An initial step in controlling the lifecycle of *E. granulosus* and minimizing infections is to determine the genotype. Peru has been reported to harbor a number of different *E. granulosus* genotypes. The presence of G6 and G7 genotypes was confirmed by partial sequence analysis of samples for the mitochondrial CO1 and nuclear elongation factor 1 alpha genes (Moro et al. 2009). The G7 genotype was identified in pigs from Lima, a city considered to be a low endemic area of *E. granulosus*. The same study reported the variant G6 in goats, as well as in one human case; nevertheless, this paper reported the predominance

M84661_G1	AGTCATAITTT	GTTTGAGTAT	TAGTGCTAAT	TTTGATCCGT	TTGGGTTTTA	TGGGTTGTTG
AF458871_G1AC
AF458872_G1BC
AF458873_G1CC
AF458874_G1DC
AF458875_G1EC
EF393619_G1.1C
EF595654_G1.2C
EU178103_G1.3C
EU178104_G1.4T
PUNO_01C
PUNO_08C
JUNIN_16C
JUNIN_26C
CUSCO_45C
CUSCO_53C
AREQUIPA_54C
AREQUIPA_67C
HUANCABELICA_68C
HUANCABELICA_73C
AVACUCHO_74C
*AVACUCHO_80T
*AVACUCHO_83T
*AVACUCHO_84T
AVACUCHO_85C
HUMANO_86C
HUMANO_99C

Fig. 2: sequence alignment of cytochrome C oxidase subunit 1 (CO1) gene. Reference sequences for the CO1 gene for the genotype G1 variant are shown at the top with a random selection of isolate sequences beneath showing identity with the G1 sequence. The three isolates with a polymorphism (denoted by asterisk) show identity with the G1⁴ sequence.

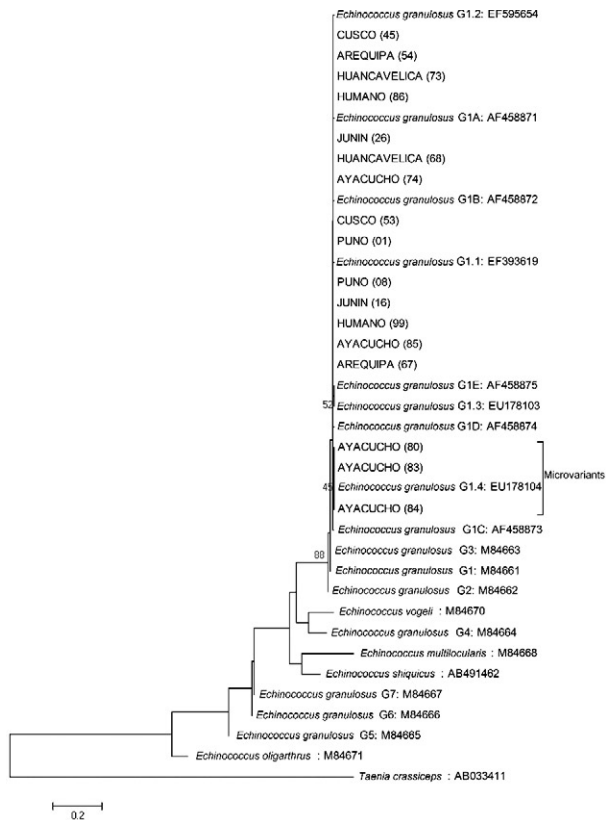


Fig. 3: phylogenetic tree of representative sequences of *Echinococcus granulosus* from Peru and reference sequences of other genotypes. The tree was obtained by neighbor-joining analysis using partial sequences (375 bp) from cytochrome *c* oxidase subunit 1 gene (COI). All isolates cluster with sequences belonging to the common sheep G1 genotype variant. The three isolates from Ayacucho with the identical one nucleotide polymorphism cluster with the G1⁴ microvariants sequence (accession EU178104).

of the common sheep genotype (G1) in animal hosts and in four human cases in other studied areas of Peru. In this study, we confirmed the higher rate of fertility of G1 genotype cysts in sheep with respect to cattle.

The results presented here represent a more exhaustive sampling of cysts in cattle producing regions of Peru. Samples were obtained from cattle and sheep soon after slaughter, while human samples were collected from cysts sent to the National Reference Laboratory, NIH. Partial sequencing of the mitochondrial COI gene, amplified by PCR, showed the presence of only the G1 genotype among all 99 isolates studied. Similar results were reported by Santivañez et al. (2008) using the same molecular approach, showing the presence of G1 genotype in 21 human isolates from other endemic regions in Peru. G1 corresponds to the most common *E. granulosus* genotype found in sheep and humans worldwide.

In three of the samples from sheep in Ayacucho, a polymorphism was observed that corresponds to the G1⁴ microvariant. Previously, this microvariant had been described solely in Turkey (Vural et al. 2008).

The G1 genotype is the more common, infectious *E. granulosus* genotype in the world with a wide range of hosts (Craig et al. 2003). Our discovery of a single variant suggests that similar mechanisms are responsible for its persistence in the endemic areas studied. A major consideration stems from the observation that close relationships between dogs and humans appear to correlate with the high prevalence of the disease in these areas (Moro et al. 2009). In these poorer regions, dogs are often fed with livestock viscera, which may be infected with the parasite. This activity alone could be sufficient to propagate the current endemic state. In conclusion, our results indicate the prominent circulation of the common sheep genotype (G1) in hyper-endemic areas of Peru, taking into account the substantial number of samples analyzed per area. Moreover, this paper presents the first report of G1⁴ microvariants of the G1 genotype in South America.

ACKNOWLEDGEMENTS

To the staff of the Regional Reference Laboratories of Puno, Junín, Cusco, Arequipa, Huancavelica and Ayacucho, Peru, for their help in samples collection, to Dr William Provance Jr, for the critical reading and English corrections of the manuscript, to Dr Mara Rosenzvit, for providing the controls genotypes, to Dr Adeilton Brandão, for their technical assistance, and to Heloisa Nogueira, for their help in designing the map.

REFERENCES

- Bowles J, Blair D, McManus DP 1992. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Mol Biochem Parasitol* 54: 165-173.
- Bowles J, Blair D, McManus DP 1995. A molecular phylogeny of the genus *Echinococcus*. *Parasitology* 110: 317-328.
- Bowles J, McManus DP 1993a. NADH dehydrogenase I gene sequences compared for species and strains of the genus *Echinococcus*. *Int J Parasitol* 23: 969-972.
- Bowles J, McManus DP 1993b. Rapid discrimination of *Echinococcus* species and strains using a polymerase chain reaction-based RFLP method. *Mol Biochem Parasitol* 57: 231-239.
- Bowles J, McManus DP 1993c. Molecular variation in *Echinococcus*. *Acta Trop* 53: 291-305.
- Busi M, Snábel V, Varcasia A, Garippa G, Perrone V, De Liberato C, D'Amelio S 2007. Genetic variation within and between G1 and G3 genotypes of *Echinococcus granulosus* in Italy revealed by multilocus DNA sequencing. *Vet Parasitol* 150: 75-83.
- Craig PS, Rogan MT, Campos-Ponce M 2003. *Echinococcosis*: disease, detection and transmission. *Parasitology* 127 (Suppl.): S5-S20.
- Kamenetzky L, Gutierrez AM, Canova SG, Haag KL, Guarnera EA, Parra A, García GE, Rosenzvit MC 2002. Several strains of *Echinococcus granulosus* infect livestock and humans in Argentina. *Infect Genet Evol* 2: 129-136.
- Lavikainen A, Lehtinen MJ, Meri T, Hirvelä-Koski V, Meri S 2003. Molecular genetic characterization of the Fennoscandian cervid strain, a new genotypic group (G10) of *Echinococcus granulosus*. *Parasitology* 127: 207-215.
- Manterola C, Benavente F, Melo A, Vial M, Roa JC 2008. Description of *Echinococcus granulosus* genotypes in human hydatidosis in a region of southern Chile. *Parasitol Int* 57: 342-346.

- McManus DP 2002. The molecular epidemiology of *Echinococcus granulosus* and cystic hydatid disease. *Trans R Soc Trop Med Hyg* 96 (Suppl. 1): S151-S157.
- McManus DP, Bowles J 1996. Molecular genetic approaches to parasite identification: their value in diagnostic parasitology and systematics. *Int J Parasitol* 26: 687-704.
- McManus DP, Thompson RC 2003. Molecular epidemiology of cystic echinococcosis. *Parasitology* 127 (Suppl.): S37-S51.
- Moks E, Jögisalu I, Valdmann H, Saarma U 2008. First report of *Echinococcus granulosus* G8 in Eurasia and a reappraisal of the phylogenetic relationships of 'genotypes' G5-G10. *Parasitology* 135: 647-654.
- Moro PL, Nakao M, Ito A, Shantz PM, Cavero C, Cabrera L 2009. Molecular identification of *Echinococcus* isolates from Peru. *Parasitol Int* 58: 184-186.
- MS - Ministério de Salud 2005. *Oficina de estadística e informática, Informe Anual Estadística MINSA*, MS, Lima.
- Nakao M, Li T, Han X, Ma X, Xiao N, Qiu J, Wang H, Yanagida T, Mamuti W, Wen H, Moro PL, Giraudoux P, Craig PS, Ito A 2010. Genetic polymorphisms of *Echinococcus* tapeworms in China as determined by mitochondrial and nuclear DNA sequences. *Int J Parasitol* 40: 379-385.
- Nakao M, McManus DP, Schantz PM, Craig PS, Ito A 2007. A molecular phylogeny of the genus *Echinococcus* inferred from complete mitochondrial genomes. *Parasitology* 134: 713-722.
- Náquira C 1993. Situación de la hidatidosis en el Perú. In *Memorias de la reunión del grupo científico sobre avances en la prevención, control y tratamiento de la hidatidosis*, Pan American Health Organization, Washington, p. 217-229.
- Otto TD, Vasconcellos EA, Gomes LH, Moreira AS, Degraive WM, Mendonça-Lima L, Alves-Ferreira M 2008. ChromaPipe: a pipeline for analysis, quality control and management for a DNA sequencing facility. *Genet Mol Res* 7: 861-871.
- Pérez LCR 2007. *Proyecto de control de hidatidosis en el Perú por vigilancia epidemiológica*, PhD Thesis, Universidad Mayor de San Marcos, Lima, 132 pp.
- Rosenzvit MC, Zhang LH, Kamenetzky L, Canova SG, Guarnera EA, McManus DP 1999. Genetic variation and epidemiology of *Echinococcus granulosus* in Argentina. *Parasitology* 118: 523-530.
- Saitou N, Nei M 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 46-425.
- Santivañez SJ, Gutierrez AM, Rosenzvit MC, Muzulin PM, Rodriguez ML, Vasquez JC, Rodriguez S, Gonzalez AE, Gilman RH, Garcia HH, Cysticercosis Working Group in Peru 2008. Human hydatid disease in Peru is basically restricted to *Echinococcus granulosus* genotypes G1. *Am J Trop Med Hyg* 79: 89-92.
- Scott JC, Stefaniak J, Pawlowski ZS, McManus DP 1997. Molecular genetic analysis of human cystic hydatid cases from Poland: identification of a new genotypic group (G9) of *Echinococcus granulosus*. *Parasitology* 114: 37-43.
- Thompson RC 2008. The taxonomy, phylogeny and transmission of *Echinococcus*. *Exp Parasitol* 119: 439-446.
- Thompson RC, Lymbery AJ 1988. The nature, extent and significance of variation within the genus *Echinococcus*. *Adv Parasitol* 27: 209-258.
- Thompson RC, Lymbery AJ, Constantine CC 1995. Variation in *Echinococcus*: towards a taxonomic revision of the genus. *Adv Parasitol* 35: 145-176.
- Thompson RC, McManus DP 2002. Towards a taxonomic revision of the genus *Echinococcus*. *Trends Parasitol* 18: 452-457.
- Thompson RCA, McManus DP 2001. Aetiology: parasites and life cycles. In J Eckert, MA Gemmel, FX Meslin, ZS Pawlowski, *Manual on Echinococcosis in humans and animals: a public health problem of global concern*, World Health Organization/World Organization for Animal Health, Paris, p. 1-9.
- Vural G, Baca AU, Gauci CG, Bagci O, Gicik Y, Lightowlers MW 2008. Variability in the *Echinococcus granulosus* cytochrome c oxidase 1 mitochondrial gene sequence from livestock in Turkey and a reappraisal of the G1-3 genotype cluster. *Vet Parasitol* 154: 347-350.
- Zhang L, Eslami A, Hosseini SH, McManus DP 1998. Indication of the presence of two distinct strains of *Echinococcus granulosus* in Iran by mitochondrial DNA markers. *Am J Trop Med Hyg* 59: 171-174.