Title
The Echinococcus granulosus antigen B shows a high degree of genetic variability

Journal
Experimental Parasitology, 108(1-2)

ISSN
0014-4894

Authors
Arend, AC
Zaha, A
Ayala, FJ
et al.

Publication Date
2004-09-01

DOI
10.1016/j.exppara.2004.07.009

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Experimental Parasitology 108 (2004) 76–80
Experimental Parasitology
www.elsevier.com/locate/yexpr

Research Brief

The *Echinococcus granulosus* antigen B shows a high degree of genetic variability

A.C. Arend[a], A. Zaha[b], F.J. Ayala[c], K.L. Haag[c,d,*]

[a] *Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, CEP 91501-970, Porto Alegre, RS, Brazil*
[b] *Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Caixa Postal 15005, CEP 91501-970, Porto Alegre, RS, Brazil*
[c] *Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA*
[d] *Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, CEP 91501-970, Porto Alegre, RS, Brazil*

Received 8 April 2004; received in revised form 22 July 2004; accepted 22 July 2004
Available online 2 September 2004

Abstract

Echinococcus granulosus larvae secret a polymeric lipoprotein known as antigen B (AgB) into the metacestode hydatid fluid. Three similar AgB subunits have been previously identified (AgB1, AgB2, and AgB3), and their respective genes isolated, but the actual number of genes encoding AgB subunits remains uncertain. In this study, we characterize the variability of genes encoding the AgB2 subunit, using PCR and RT-PCR followed by cloning and sequencing. We have analyzed 32 cDNA and 34 genomic sequences from a single metacestode, showing a high degree of sequence polymorphism. In addition, we have identified a possibly new AgB subunit, which we call AgB4. Additionally, we describe an *AgB2* genomic clone lacking (i) a segment corresponding to the intron and (ii) a short, 45 bp sequence within exon II. The 45 bp segment encompasses the conserved splicing signals and corresponds to a highly conserved insect promoter motif.

Keywords: *Echinococcus granulosus*; Antigen B; Genetic variability; Cloning; Sequencing

Infection by the metacestode larval stage of the tapeworm *Echinococcus granulosus* causes an important zoonosis, distributed worldwide, known as cystic hydatid disease. The adult stage develops mainly in the small intestine of dogs and others canids. The metacestode occurs in viscera (mainly liver and lung) of ungulates macropods and humans (as accidental hosts).

Antigen B (AgB) is an extremely abundant component of the hydatid cyst (metacestode) fluid and is a polymeric thermostable lipoprotein of 120–160 KDa (Oriol and Oriol, 1975). The role of AgB in the parasite biology is not completely elucidated, but it seems to be related to the evasion of the host's immune response. In vitro studies have shown that it inhibits neutrophil recruitment, has protease inhibitor activity (Shepherd et al., 1991) and exploits the activation of T-helper cells by eliciting a non-protective Th2 cell response (Riganò et al., 2001). Recently, it has been demonstrated that AgB induces apoptosis in polymorphonuclear cells from patients with active disease (Riganò et al., 2002). It is a highly immunogenic antigen used in immunodiagnostic assays. Recombinant antigens derived from distinct AgB subunits possess variable diagnostic values, with AgB2 showing the highest sensitivity and specificity in immunodiagnostic tests (Virginio et al., 2003).

The first isolated *AgB* nucleotide sequence (*AgB1*) was a cDNA cloned by Shepherd et al. (1991), and a second sequence (*AgB2*) was described by Fernández et al. (1996). AgB1 and AgB2 show a hydrophobic stretch from amino acid positions 1–16 and 1–20, respectively, followed by one potential eukaryotic signal for peptidase cleavage (Von Heijini, 1986), indicating that such peptides are secreted. Chemale et al. (2001), using primers specific to the nucleotide sequence for *AgB1*, have isolated a third AgB gene (*AgB3*), and raised the possibility that the gene family actually includes even more than three copies.

Frosch et al. (1994) identified the first nucleotide polymorphisms in cDNA sequences encoding the AgB1 subunit. Haag et al. (1998) showed that the rates of *AgB1* non-synonymous substitutions among isolates from different strains suggest positive selection. The aim of the present study was to unravel the amount, degree and nature of genetic variability for the *E. granulosus* AgB2 subunit within one single genome.

To isolate AgB2-related nucleotide sequences we have used DNA and mRNA derived from a single bovine cyst (a clonal organism where

* Corresponding author. Fax: +55 51 3316 7311.
E-mail addresses: anarend@terra.com.br (A.C. Arend), karen.haag@ufrgs.br (K.L. Haag).

0014-4894/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.exppara.2004.07.009

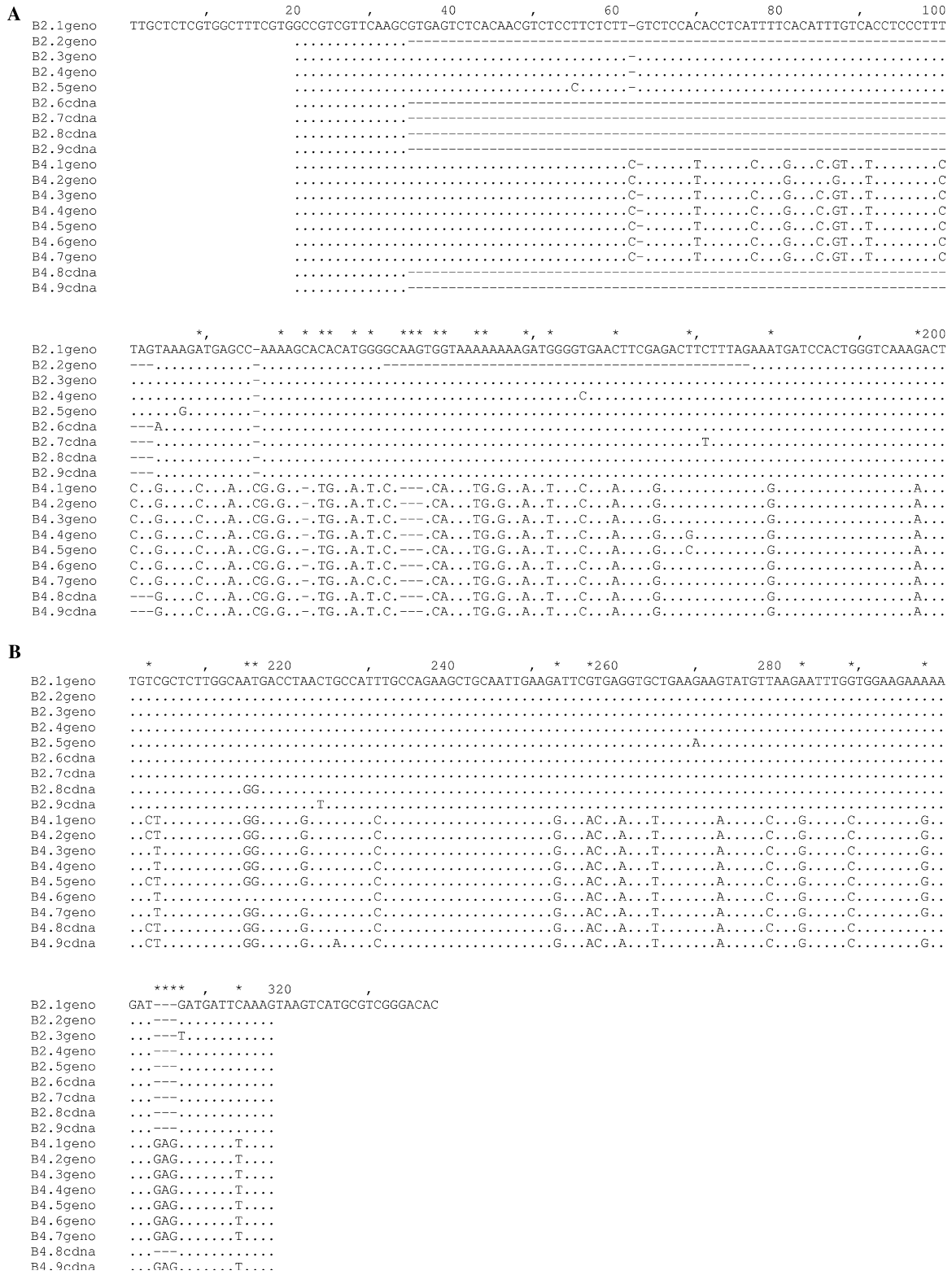

Fig. 1. Alignment of nucleotide sequences from the 18 distinct AgB clones isolated in the present study. The top nine clones, *B2.1geno* to *B2.9cdna*, are related to the previously characterized *AgB2* gene (GenBank Accession No. U15001); the other nine clones, *B4.1geno* to *B4.9cdna*, relate to a cDNA deposited in GenBank with Accession No. AF252859. Genomic and cDNA clones were obtained in independent experiments. Dots (.) indicate identity with the reference sequence. Dashes (- - -) indicate indels. Primers used to generate the inserts are shown at both ends of the reference sequence. Replacement positions are indicated with asterisks on the reference sequence. When cDNA and genomic clones were identical inside the coding region, only the genomic clone was represented.

asexual reproduction originates protoscoleces) in order to avoid intercyst polymorphism. Using DNA or mRNA-derived (cDNA) templates, low stringency polymerase chain reactions (PCR) was performed using primers AGB8/2FOR (5′-TTGCTCTCGTGGCTTTCGTG-3′) and AGB8/2REV (5′-GTGTCCCGACGCATGACTTA-3′) and the proofreading Tli DNA polymerase, or alternatively, the TaKaRa Ex Taq polymerase, to prevent nucleotide misincorporation errors during PCR. PCR was carried out following the same procedures described in Haag et al. (1997). After the PCR, products were cloned into the Zero Blunt TOPO vector (Invitrogen) and, alternatively, into pGEM-T Easy (Promega), followed by sequencing. Nucleotide diversity (π) was calculated according Nei (1987) with the DnaSP 2.0 program (Rozas and Rozas, 1999).

The analysis of 66 recombinant clones obtained from a single *E. granulosus* cyst disclosed a high degree of genetic variation. We found eighteen different sequences, represented in Fig. 1. The sequence of clone *B2.1geno* corresponded to a previously deposited *AgB2* cDNA sequence (GenBank Accession No. U15001). Four other genomic clones and six cDNA clones had a high identity to this gene, with minor nucleotide differences (Fig. 1). However, eleven other clones (seven genomic and four cDNA clones), are related to a mRNA sequence previously deposited in GenBank with Accession No. AF252859 (Lu, 2000).

Table 1
Nucleotide (above diagonal) and amino acid (below diagonal) identities between the four known AgB sub-units[a]

	AgB1 (%)	*AgB2* (%)	*AgB3* (%)	*AgB4* (%)
AgB1	—	51	63	48
AgB2	41	—	43	82
AgB3	48	36	—	42
AgB4	36	70	36	—

[a] GenBank Accession Nos.: AF143813 (*AgB1*), L48620 (*AgB2*), AF362442 (*AgB3*), AY357110 (*AgB4*).

Given its low amino acid identity with AgB2 (70%, Table 1) this second set of sequences is suggested to correspond to a fourth AgB subunit, which we call AgB4. Studies employing polyclonal antisera indicate that non-denatured globular proteins differing by at least 30–40% amino acid sequence usually do not cross-react (Wilson et al., 1977). Comparison of cDNA of the sequences of the cDNA with those of the genomic clones suggested that at least four different gene variants are expressed in protoscoleces. Some cDNA clones with sequences identical to clones *B2.1geno*, *B2.2geno*, *B4.1geno* and *B4.3geno* were identified (Table 2).

Fig. 2 shows the nucleotide diversities estimated for different gene regions (intron, exon II, and the whole sequence; exon I is excluded from the analyses, because it does not show variation among the clones isolated in this study). Values for all sequences combined are much higher than those calculated separately for the *AgB2* and *AgB4* groups, reflecting the considerable divergence between the two sets of

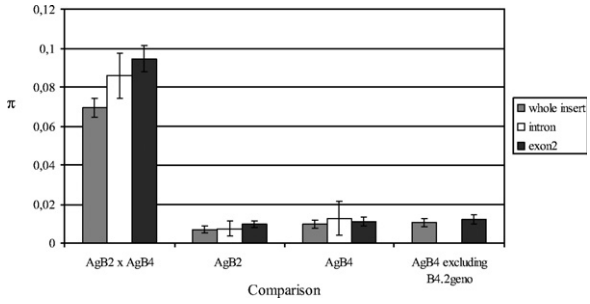

Fig. 2. Mean nucleotide diversity (π) and standard deviation (vertical bar) for the combined *AgB2* and *AgB4* clones and for clones within each group. Calculations are made for the whole fragment, for only the intron, and for only exon II. *AgB4* nucleotide diversities are also estimated without clone *B4.2geno*, because all intron substitutions occur in this clone.

Table 2
Number (N) of cDNA and genomic clones obtained for each *AgB* variant

Variants	N cDNA	N genomic	Mutation position						
			Transition		Transversion				
			C ↔ T	A ↔ G	A ↔ C	A ↔ T	C ↔ G	T ↔ G	Deletion
B2. Igeno	16	8	—	—	—	—	—	—	—
B2.2geno	1	1	—	—	—	—	—	—	132–176
B2.3geno	—	1	—	—	—	—	—	307*	—
B2.4geno	—	2	156	—	—	—	—	—	—
B2.5geno	—	1	55	107, 27	—	—	—	—	—
B2.6cdna	1	—	—	—	—	104	—	—	—
B2.7cdna	1	—	171	—	—	—	—	—	—
B2.8cdna	1	—	—	215*	—	—	—	216*	—
B2.9cdna	1	—	224*	—	—	—	—	—	—
B4.1geno	3	6	—	—	—	—	—	—	—
B4.2geno	—	1	77, 85	—	—	—	—	88	63
B4.3geno	3	8	203	—	—	—	—	—	—
B4.4geno	—	1	203	—	—	—	—	169*	—
B4.5geno	—	1	169*	—	—	—	—	—	—
B4.6geno	—	1	203	215,* 222	—	—	—	216*	—
B4.7geno	—	1	130,* 203	—	—	—	—	—	—
B4.8cdna	1	—	314	298	—	—	—	—	304–306
B4.9cdna	1	—	—	226*	—	—	—	—	—
Total	29	32	13	7	0	1	0	5	3

The alignment position of transitions, transversions, and deletions are indicated. Totals are given at the bottom. Mutations were computed by comparing *AgB2*-related sequences with *B2.1geno*, and *AgB4*-related sequences with *B4.1geno*. Amino acid replacements are indicated with asterisks.

sequences. Note that, within each group, genetic variation sometimes involves only a single non-synonymous substitution (Fig. 1 and Table 2). Also, within each group and between the two groups, the nucleotide diversity was greater in exon II than in the intron.

There is evidence that the observed nucleotide variants are not due to PCR errors. Indeed, Haag et al. (2004) have encountered the same substitutions at the same sites in separate experiments, using three independent PCR and different DNA templates. Moreover, high fidelity/proof-reading DNA polymerases were used to prevent nucleotide misincorporations. Moreover, as an external control for misincorporation artifacts, we cloned the amplification products from a nuclear but single-copy gene, encoding a malate dehydrogenase (*mdh*), using the same reagents. The *mdh* amplicons also included a small intron and were equivalent in size to the *AgB* amplicons. We sequenced 21 *mdh* clones, but no nucleotide variability was found.

The presence of a relatively large number of different sequences (Table 2) may be due to, at least, two possibilities. (1) The *AgB* gene family has several duplicated genes within a genome, so that some duplications are relatively ancient, such as *AgB2* and *AgB4*, while other duplications are relatively recent, such as *B2.1* versus *B2.2* or *B2.3*, or *B4.1* versus *B4.2* and *B4.3*. (2) There are ancient duplicated genes (such as *AgB2* and *AgB4*), but additional variation (such as among *B2.1*, *B2.2* and *B2.3*) is being originated during asexual amplification of the metacestode, so that protoscoleces are not true clones with respect to the *AgB* genes (see Haag et al., 2004).

Five chimeric clones (three cDNA and two genomic clones) were excluded from analysis because of their possible artifactual origin. PCR-derived chimerism has been reported in previous attempts to isolate AgB genes (Chemale et al., 2001) and other multigene family studies (Bhavsar et al., 1994; Gaedigk et al., 1998). Incomplete extension products often work as primers in the following PCR cycles. Chimeras are produced by heterologous primer annealing, when mixed templates are present in the same reaction (which is the case of a multigene family with closely related gene copies, or of a mixture of DNAs from genetically related individuals). Our interpretation of the excluded clones as chimeric is supported by an additional experiment, in which a primer pair designed to specifically amplify an *AgB2*/*AgB4* chimera did not result in any amplification product from templates derived from six different *E. granulosus* metacestodes (not shown).

Considering that *AgB4* corresponds to another gene copy, the number of nucleotide substitutions within each gene group was analyzed. As expected, transitions are more frequent than transversions (20 in 26 substitutions) and the most frequent transition is C ↔ T (see Table 2). It is interesting that a synonymous C ↔ T transition at position 203 appears in four out of nine *AgB4* clones (cDNA and genomic). Since genomic and cDNA clones are obtained in independent experiments, there is a low probability that the transitions at position 203 represent PCR artifacts. We found a non-synonymous A ↔ G transition adjacent to a non-synonymous T ↔ G transversion, at positions 215 and 216, in an *AgB2* cDNA clone (*B2.8cdna*) as well as in all, but one *AgB4* sequences. One *AgB4* genomic clone (*B4.6geno*) exhibits the reversed substitutions, suggesting the possibility of gene conversion (see also Fig. 1).

Four out of 10 and 6 out of 16 nucleotide substitutions detected, respectively, in the *AgB2* and *AgB4* sequences were replacements, indicating a possible adaptive role (Table 2). Haag et al. (1998) suggested positive selection in the evolution of *AgB1*, which would not be surprising, given that AgB seems to act in immune response evasion. Endo et al. (1996), searching for candidate genes on which positive selection might be operating, obtained 17 gene groups, 9 of which encoded surface antigens of parasites or viruses.

An intriguing finding of the present study was a truncated genomic clone (*B2.2geno*) with clear characteristics of an mRNA molecule (Fig. 1). It lacked the *AgB2* region corresponding to the intron, plus a 45 bp stretch with conserved splicing junctions (GC/AG). The deleted 45 bp sequence in the *B2.2geno* clone also includes a 3′ YNYAG conserved splicing signal (Senapathy et al., 1990) and corresponds to an insect promoter motif (http://www.fruitfly.org/seq_tools/promoter.html, October 11, 2003). Although we can not rule out a PCR artifact as a possible cause for this clone, it might indicate that *AgB* genes duplicate by a mechanism requiring an intermediate mRNA. The 45 bp deletion was found both in genomic and cDNA experiments, suggesting that this *AgB2* variant is transcribed.

We have herein described a fourth *AgB* gene (*AgB4*), which is transcribed by the *E. granulosus* metacestode, and seems to encode an additional AgB subunit. We also show that the two genes, *AgB2* and *AgB4*, from a single cyst have considerable variation, which might be adaptive. The surprising discovery of a truncated *AgB2* transcribed variant, plus some additional clues, suggest a process involving an intermediate mRNA.

Acknowledgments

The work was supported by PADCT/CNPq, FAPERGS, CAPES, and CABBIO, as well as National Institutes of Health Grant GM42397 to FJA. GenBank Accession Nos. for our sequences are: AY314984–AY314991 (*AgB2*) and AY357108–AY357116 (*AgB4*).

References

Bhavsar, D., Zheng, H.D., Drysdale, J., 1994. Chimerism in PCR products from a multigene family. Biochemical and Biophysical Research Communications 205 (1), 944–947.

Chemale, G., Haag, K.L., Ferreira, H., Zaha, A., 2001. *Echinococcus granulosus* antigen B is encoded by a gene family. Molecular and Biochemical Parasitology 116, 233–237.

Endo, T., Ieko, K., Gojobori, T., 1996. Large-scale search for genes on which positive selection may operate. Molecular Biology and Evolution 13, 685–690.

Frosch, P., Hartmann, M., Muhlschlegel, F., Frosch, M., 1994. Sequence heterogeneity of the echinococcal antigen B. Molecular and Biochemical Parasitology 64, 171–175.

Fernández, V., Ferreira, H., Fernandez, C., Zaha, A., 1996. Molecular characterization of a novel 8-KDa subunit of *Echinococcus granulosus* antigen B. Molecular and Biochemical Parasitology 77, 247–250.

Gaedigk, A., Lekas, P., Berchuk, M., Grant, D.M., 1998. Novel sulfotransferases cloned by RT-PCR: real proteins or PCR artifacts? Chemico-Biological Interactions 109, 43–52.

Haag, K.L., Zaha, A., Araújo, A., Gottstein, B., 1997. Reduced genetic variability within coding and non-coding regions of the *Echinococcus multilocularis* genome. Parasitology 115 (Pt 5), 521–529.

Haag, K.L., Araújo, A., Gottstein, B., Zaha, A., 1998. Selection, recombination and history in a parasitic flatworm (*Echinococcus*) inferred from nucleotide sequence. Memórias do Instituto Oswaldo Cruz 93 (5), 695–702.

Haag, K.L., Alves-Junior, L., Zaha, A., Ayala, F.J., 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.

Lu, J.H., 2000. Direct submission. Department of Parasitology, Sun Yat-Sen University of Medical Sciences, Guangdong, China.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Oriol, R., Oriol, C., 1975. Physiochemical properties of a lipoprotein antigen of *Echinococcus granulosus*. The American Journal of Medicine and Hygiene 24, 96–100.

Riganò, R., Profumo, E., Bruschi, F., Carulli, G., Azarrà, A., Ioppolo, S., Buttari, B., Ortona, E., Margutti, P., Teggi, A., Siracusano, A., 2001. Modulation of human immune response by *Echinococcus*

granulosus antigen B and its possible role in evading host defenses. Infection and Immunity 69, 288–296.

Riganò, R., Buttari, B., Profumo, E., Ortona, E., Margutti, P., Teggi, A., Siracusano, A., 2002. Apoptosis modulation in PMBC from patients with cystic echinococcosis. The American Journal of Medicine and Hygiene 67 (2), 298.

Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.

Senapathy, P., Shapiro, M.B., Harris, N.L., 1990. Splice junctions, branch point sites and exons: sequence statistics, identification and applications to the genome project. Methods in Enzymology 183, 252–278.

Shepherd, J., Aitken, A., McManus, D., 1991. A protein secreted in vivo by *Echinococcus granulosus* inhibits elastase activity and neutrophil chemotaxis. Molecular and Biochemical Parasitology 44, 81–90.

Von Heijini, G., 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Research 14, 4683–4690.

Virginio, V., Hernandez, A., Rott, M.B., Monteiro, K.M., Zandonai, A.F., Nieto, A., Zaha, A., Ferreira, H.B., 2003. A set of recombinant antigens from *Echinococcus granulosus* with potential for use in the immunodiagnosis of human cystic hydatid disease. Clinical and Experimental Immunology 132, 309–315.

Wilson, A.C., Carlson, S.S., White, T.J., 1977. Biochemical evolution. Annual Review of Biochemistry 46, 573–639.