Isolation and characterization of microsatellites from the tapeworm *Echinococcus granulosus*

M. L. BARTHOLOMEI-SANTOS[1,2], L. S. HEINZELMANN[1], R. P. OLIVEIRA[3], G. CHEMALE[1], A. M. GUTIERREZ[4], L. KAMENETZKY[4], K. L. HAAG[5] *and* A. ZAHA[1]*

[1] *Centro de Biotecnologia e Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15.005, Porto Alegre 91.501-970, RS, Brazil*
[2] *Departamento de Biologia, Centro de Ciências Naturais e Exatas, Universidade Federal de Santa Maria, Santa Maria 97.105-900, RS, Brazil*
[3] *Departamento de Ciências Biológicas, Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, Ouro Preto 34.500-000, MG, Brazil*
[4] *Departamento de Parasitologia, Instituto Nacional de Enfermedades Infecciosas, Administración Nacional de Laboratorios e Institutos de Salud 'Dr. Carlos G. Malbrán', Vélez Sarsfield 563, 1281 Buenos Aires, Argentina*
[5] *Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15.053, Porto Alegre 91.501-970, RS, Brazil*

(*Received 10 September 2002; revised 2 January 2003; accepted 2 January 2003*)

SUMMARY

The *Echinococcus granulosus* genome was searched for microsatellites using 8 different repeated oligonucleotides as probes (GT_{15}, CT_{15}, AT_{15}, CG_{15}, CAT_{10}, CAA_{10}, CGG_{10} and $CATA_{10}$). Southern blot experiments revealed that DNA regions containing GT, CAA, CATA and CT repeats are the most frequent in the *E. granulosus* genome. AT and CG probes showed no hybridization signal. Two loci containing CA/GT (Egmsca1 and Egmsca2) and 1 locus containing GA/CT (Egmsga1) repeats were cloned and sequenced. The locus Egmsca1 was analysed in 73 isolates from Brazil and Argentina whose strains were previously characterized. Brazilian isolates from cattle strain and Argentinean isolates from camel strain were monomorphic and shared the allele $(CA)_7$. Argentinean isolates of sheep and Tasmanian sheep strains shared 2 alleles [$(CA)_8$ and $(CA)_{10}$] with Brazilian isolates of sheep strain. The allele $(CA)_{11}$ was found only in Brazilian isolates of sheep strain at a low frequency. The Brazilian and the Argentinean sheep strain populations were tested for the Hardy-Weinberg equilibrium, and only the former was in agreement with the expectations. No polymorphism was found among individual protoscoleces from a single hydatid cyst, validating the utilization of pooled protoscoleces from 1 cyst, grouped as an isolate, in population studies. This work describes for the first time the isolation and characterization of microsatellites from *E. granulosus*.

Key words: *Echinococcus granulosus*, Cestoda, microsatellite markers.

INTRODUCTION

Echinococcus granulosus is a small endoparasitic flatworm causative of hydatid disease in intermediate hosts (wild and domestic herbivores and also man) and of echinococcosis in definitive hosts (canids). The hydatid disease is one of the most important zoonosis (Thompson & Limbery, 1995). Although control programmes against human hydatid disease have been established in some countries and effective strategie controls are available, the parasite has still a wide geographical distribution affecting many countries of all continents (Eckert, Conraths & Tackmann, 2000).

Microsatellites are tandemly repeated DNA segments, with short repeat units ($\leqslant$6 bp), which are usually polymorphic in number (Tautz, 1989; Weber & May, 1989; Litt & Lutty, 1989). Because of their high abundance, widespread distribution and high heterozygosities in the genomes of various organisms (Stallings *et al*. 1991), microsatellites are useful in studies of population structure, molecular epidemiology and genomic mapping.

Although a variety of microsatellite loci have been described in parasites, we noticed only 2 such studies in Cestoda. Pentanucleotide repeats were found within the *E. multilocularis* U1 snRNA gene complexes, which discriminated parasites from different localities (Bretagne *et al*. 1996). Also, 5 polymorphic loci were isolated from the tapeworm *Schistocephalus solidus* (Binz *et al*. 2000). No microsatellite markers have been described for *E. granulosus* to date. Since the species has a high degree of differentiation, and strains are supposed to differ in developmental, transmission and other features, microsatellites would be valuable markers for epidemiological and clinical studies. In this work the presence and abundance of microsatellites in *E. granulosus* genome was analysed

* Corresponding author: Centro de Biotecnologia e Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15.005, Porto Alegre 91.501-970, RS, Brazil. Tel: +55 51 3316 6054. Fax: + 55 51 3316 7309. E-mail: zaha@dna.cbiot.ufrgs.br (A. Zaha).

 © 2003 Cambridge University Press
DOI: 10.1017/S0031182003003184 Printed in the United Kingdom

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

Table 1. Oligonucleotide sequences used as probes and conditions used in the hybridization experiments

Oligo sequence	Oligo Tm (°C)	Hybridization temperature (°C)	Membrane washing conditions
$(GT)_{15}$	90	65	2 × SSC; 0·1% SDS, 50 °C, 10 min (twice) 1 × SSC; 0·1% SDS, 50 °C, 10 min (once)
$(CT)_{15}$	90	65	2 × SSC; 0·1% SDS, 50 °C, 10 min (twice) 1 × SSC; 0·1% SDS, 50 °C, 10 min (once)
$(AT)_{15}$	60	40	5 × SSC; 0·1% SDS, room temperature, 5 min (twice)
$(CG)_{15}$	Undetermined	65	2 × SSC; 0·1% SDS, 50 °C, 10 min (twice) 1 × SSC; 0·1% SDS, 50 °C, 10 min (twice)
$(CAT)_{10}$	80	60	2 × SSC; 0·1% SDS, 50 °C, 10 min (3 times)
$(CAA)_{10}$	80	60	2 × SSC; 0·1% SDS, 50 °C, 10 min (3 times)
$(CGG)_{10}$	Undetermined	65	2 × SSC; 0·1% SDS, 50 °C, 10 min (twice) 1 × SSC; 0·1% SDS, 50 °C, 10 min (once)
$(CATA)_{10}$	Undetermined	65	2 × SSC; 0·1% SDS, 50 °C, 10 min (3 times)

by Southern hybridizations. We also isolated and characterized 2 CA and 1 GA repeat loci. We used one of the CA loci for a preliminary population analysis.

MATERIALS AND METHODS

Southern hybridization

Total DNA obtained from protoscoleces of a single hydatid cyst was digested with *Alu*I, *Taq*I, *Rsa*I, *Eco*RI or *Hin*dIII, electrophoresed in a 0·8% agarose gel and transferred to nylon membranes (Hybond N^+, Amersham-Pharmacia). The hybridization was carried out according to Sambrook & Russell (2001). Eight different ^{32}P-labelled oligoprobes (Table 1) were used for hybridization with DNA of *E. granulosus*. The temperature of hybridization and the stringency of washing conditions varied according to the T_m (melting temperature) of each oligonucleotide (Table 1).

Microsatellite isolation

$(CA)_n$, $(GA)_n$, and (TA)n repeats were isolated as described earlier (Refseth, Fangan & Jakobsen, 1997; Oliveira *et al*. 1998), with some modifications. Genomic DNA was digested with *Taq*I (Gibco BRL) and 500 ng of fragments were ligated to 50 pmoles of an adapter at 16 °C overnight using 1 unit of T4 DNA ligase (Gibco BRL). Then 100 ng of the ligation product were mixed with 20 pmoles of one of the biotinylated oligonucleotides GATGATCCGACGCAT$(CA)_{12}$, GATGATCCGACGCAT$(GA)_{12}$ or GATGATCCGACGCAT$(TA)_{12}$ in 50 μl of TE buffer (10 mM Tris–HCl, pH 8·0; 1 mM EDTA, pH 8·0) containing 2 μM of the *Taq*I-20 mer oligonucleotide to avoid hairpin formation. The mixture was heated to 95 °C for 10 min and then annealed at 60 °C (or 40 °C for AT-repeats) for 1 min. The fragments containing the repeats were captured by streptavidin-coated magnetic beads (Dynabeads M280, Dynal) and amplified by PCR (polymerase chain reaction) in a 50 μl reaction volume containing 10 pmoles of the adapter oligonucleotide, 2·5 units of *Taq* DNA polymerase (Cenbiot), 100 μM of each dNTP and *Taq* buffer (10 mM Tris–HCl, 50 mM KCl, 4 mM $MgCl_2$). The samples were incubated at 94 °C for 3 min and subjected to 23 cycles consisting of 1 min at 94 °C for denaturation and 3 min at 72 °C for annealing and extension and then incubated for a final extension for 5 min at 72 °C. PCR products were purified in GFX columns (Amersham-Pharmacia) and cloned in the pGEM-T Easy Vector (Promega). Plasmid inserts were sequenced using the DYEnamic Dye Terminator Cycle Sequencing kit for MegaBACE (Amersham-Pharmacia).

E. granulosus *isolates*

A total of 73 *E. granulosus* metacestode isolates (Table 2) from the Southern region of Brazil and from different Argentinean provinces were analysed.

Strains were identified as follows. In a previous study, Kamenetzky *et al*. (2002) sequenced the mithocondrial cytochrome c oxidase I gene and compared their results with the published sequences (Bowles, Blair & McManus, 1995). In another study, the strains from Brazilian samples were identified by SSCP of 6 different DNA segments: the non-coding introns from genes ActII and Hbx2, part of the coding region of the nuclear antigen B, part of the mithocondrial NADH dehydrogenase 1 gene, the non-coding 5′ flanking region of the cytosolic MDH (malate dehydrogenase) gene and the 3′ flanking region of Ag6 (a calcium-binding protein) gene. The obtained patterns were compared to reference isolates from each strain (Haag *et al*. 1999).

DNA amplification and analysis of microsatellite loci

Primers were designed complementary to the sequences flanking the CA/GT (Egmsca1-F: 5′ CGAAAGTGATGACAAACCAA 3′ and Egmsca1-R: 5′ GCTTGATGGAGATGAGGTCG 3′)

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

Table 2. Number of *Echinococcus granulosus* isolates from different geographical regions, strains and hosts analysed in the study

Region	Strain	Intermediate host	Number
Brazil (Southern region)	Sheep	Bovine, Ovine	38
Brazil (Southern region)	Cattle	Bovine	6
Argentina (Buenos Aires, Neuquén, Río Negro, Tucumán and Santa Cruz provinces)	Sheep	Humans, Ovine	17
Argentina (Tucumán)	Tasmanian Sheep	Humans	5
Argentina (Rio Negro and Neuquén)	Camel	Humans	7

and GA/CT repeats (Egmsga1-F: 5′ TGACGGC-GATGATGAGATAG 3′ and Egmsga1-R: 5′ CCT-TGCCACACGCTACACTG 3′). PCR was carried out in a 50 μl reaction volume containing 10–20 ng of the genomic DNA from protoscoleces of a single hydatid cyst, 20 pmoles of each primer, 10 mM Tris–HCl, pH 8·5, 50 mM KCl, 1·5 mM $MgCl_2$, 100 μM of each dNTP and 2·5U *Taq* DNA polymerase. Control reactions with bovine, ovine and human DNAs were carried out to certify that no contamination with host DNA was present in templates. PCR products were electrophoresed in 40 cm long 8% polyacrilamide gels and visualized by $AgNO_3$ staining. The different alleles found were manually sequenced to verify the number of repeats, utilizing the Thermo-Sequenase Radiolabeled Terminator Cycle Sequencing Kit (Amersham-Pharmacia).

Microsatellite analysis in individual protoscoleces from single hydatid cysts

In order to analyse possible genetic variation among individual protoscoleces from a single hydatid cyst, protoscoleces from 6 cysts were isolated under a microscope, added to 25 μl of distilled H_2O, incubated with 2 μg of proteinase K at 50 °C for 1 h and at 95 °C for 20 min, under mineral oil. PCR was carried out in a 50 μl reaction volume containing the lysed protoscolex, 10 pmoles of each Egmsca1-F and Egmsca1-R primers, 10 mM Tris–HCl, pH 8·5, 50 mM KCl, 1·5 mM $MgCl_2$, 100 μM of each dNTP and 2·5U *Taq* DNA polymerase. PCR products were electrophoresed in 40 cm long 8% polyacrilamide gels and visualized by $AgNO_3$ staining. The PCR product of 1 individual from each cyst was sequenced to determine the length of the microsatellite. The cysts used in this experiment were removed from bovine hosts from Southern region of Brazil.

Statistical analysis

All statistical analyses were performed with Genepop (v3.1b, a program developed by Raymond & Rousset (1995*a*). Only the Brazilian and Argentinean populations from sheep strain (with $n>10$) were tested. Hardy-Weinberg proportions were tested by an exact test on contingency tables (Weir, 1990). Within this test the *P*-value corresponds to the sum of probabilities of all tables (with the same allelic counts) with the same or lower probability than the probability of the observed samples, which is used to define the rejection zone. A Markov chain is used to explore adequately the space of possible tables.

Tests for genic and genotypic differentiation (concerned with the distribution of alleles and genotypes in the various populations) were performed on contingency tables by a Fisher exact test (Raymond & Rousset, 1995*b*) and a G-based test (Goudet *et al.* 1996). Fis and Fst estimations were calculated following Weir & Cockerham (1984).

RESULTS

Detection of microsatellites in the E. granulosus *genome*

The hybridization of 6 different probes [$(GT)_{15}$, $(CT)_{15}$, $(CAT)_{10}$, $(CAA)_{10}$, $(CGG)_{10}$, $(CATA)_{10}$] to *E. granulosus* genomic DNA generated a pattern of single or multiple bands ranging from approximately 0·3 to 10 kb (Fig. 1). The probes $(GT)_{15}$, $(CAA)_{10}$, $(CATA)_{10}$ and $(CT)_{15}$ showed the highest number of hybridizing bands, independently of the restriction enzyme used. Among these repeated sequences, GT and CAA were the most abundant type of microsatellites in the *E. granulosus* genome, followed by CATA and CT repeats.

The hybridization with the probes $(CAT)_{10}$ and $(CGG)_{10}$ generated only a small number of bands. No hybridization was observed with probes $(AT)_{15}$ and $(CG)_{15}$ (data not shown).

Isolation and characterization of sequences containing dinucleotide repeats

Several clones containing perfect CA and GA-repeats were isolated, but only 3 of these presented both flanking sequences (Egmsca1, Egmsca2 and Egmsga1) (Fig. 2). The other clones presented 1 flanking sequence and 1 end with the sequence

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

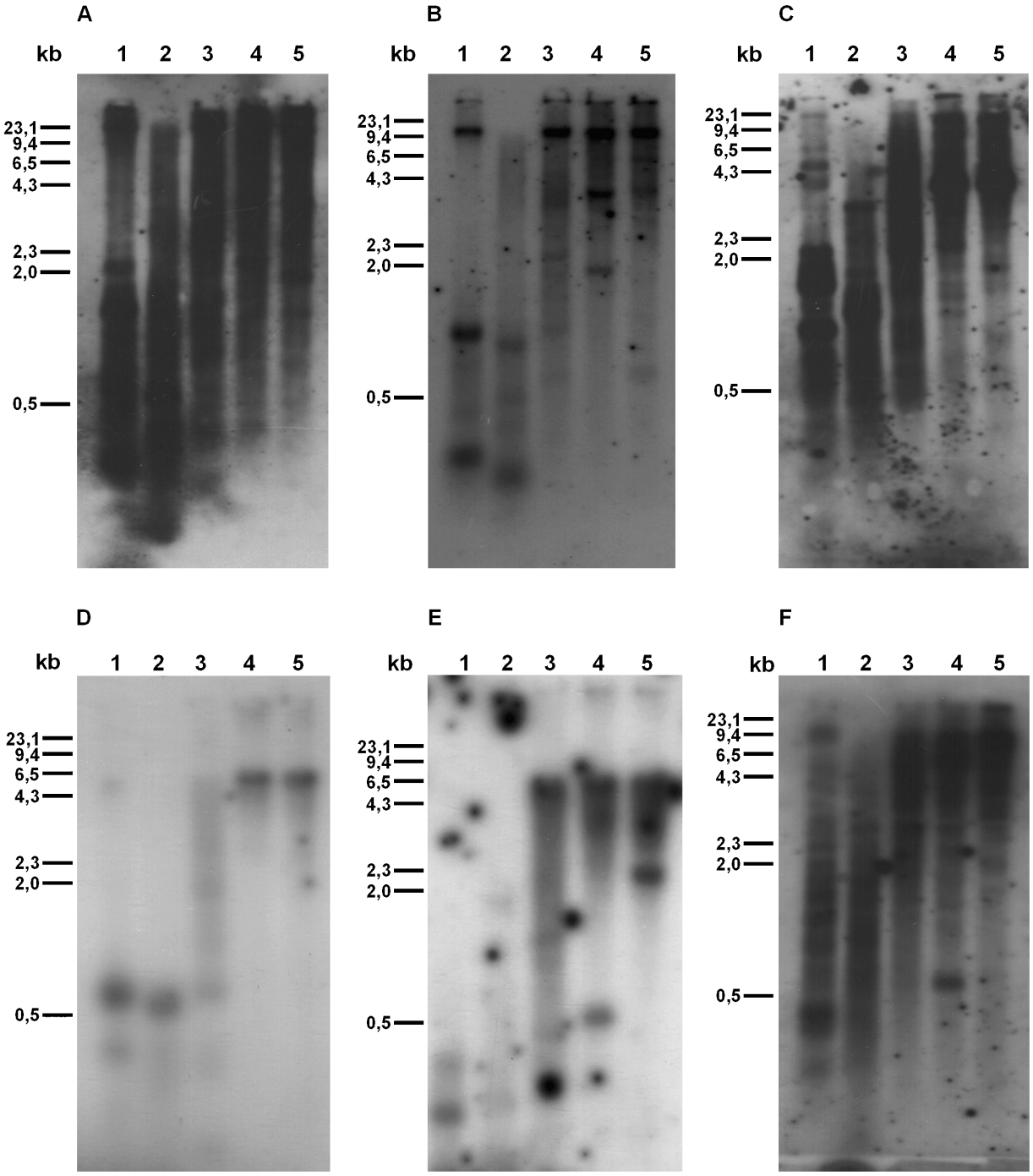


Fig. 1. Presence and abundance of microsatellites in *Echinococcus granulosus*. Total DNA obtained from protoscoleces of a single hydatid cyst was digested with *Alu*I (1), *Taq*I (2), *Rsa*I (3), *Eco*RI (4) or *Hin*dIII (5), electrophoresed in 0·8% agarose gel, transferred to nylon membranes and hybridized to the following ^{32}P-labelled oligonucleotide probes: (A) $(GT)_{15}$; (B) $(CT)_{15}$; (C) $(CAA)_{10}$; (D) $(CAT)_{10}$; (E) $(CGG)_{10}$; (F) $(CATA)_{10}$.

GATGATCCGACGCAT, due to internal priming of the biotinylated oligonucleotide leaking from the magnetic beads. From the 3 complete isolated sequences, 1 contained the microsatellite near to an insert end, impeding the design of a primer on this flanking extremity (Egmsca2). Primer pairs were designed to amplify Egmsca1 and Egmsga1 loci (Fig. 2). Attempts to isolate TA-repeats were made, despite the negative result in the Southern experiment, to ascertain a possible failure in hybridization with the $(TA)_{15}$ probe, related to its low Tm. However, no clones were isolated containing this type of repeat.

Polymorphism of the Egmsca1 locus

No polymorphism was found among 150 protoscoleces from a single hydatid cyst. Twenty protoscoleces from 5 additional cysts were also analysed, and no variation was found. All individuals were homozygotes for Egmsca1 locus, with a $(CA)_7$

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

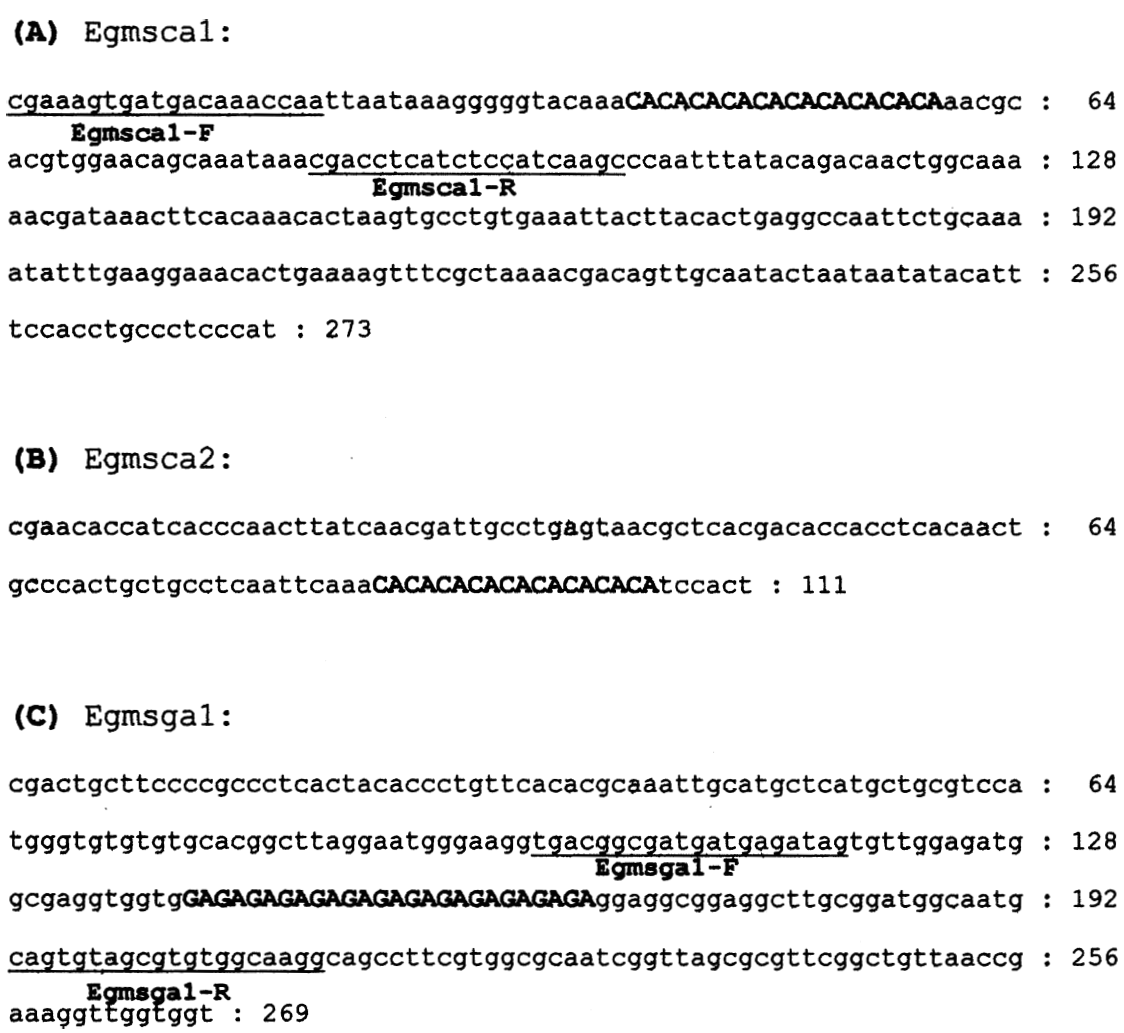


Fig. 2. Sequences of 3 *Echinococcus granulosus* microsatellite repeats and flanking regions. (A) Egmsca1 (GenBank AF459418): sequence containing a perfect stretch of 10 CA repeats; (B) Egmsca2 (GenBank AF502944): sequence containing a perfect stretch of 9 CA repeats (C) Egmsga1 (GenBank AF459419): sequence containing a perfect stretch of 13 GA repeats. Primer pairs designed for Egmsca1 and Egmsga1 loci are underlined.

(1 cyst), $(CA)_8$ (1 cyst) or $(CA)_{10}$ repeat (4 cysts), as revealed by sequencing.

Four alleles were found among the 73 studied isolates, which were confirmed by sequencing the PCR products from several homozygote and heterozygote individuals: $(CA)_7$, $(CA)_8$, $(CA)_{10}$ and $(CA)_{11}$ (Fig. 3).

The allelic frequencies are shown in Table 3. The allele $(CA)_7$ was found only in cattle and camel strains. The sheep and the Tasmanian sheep strains shared the alleles $(CA)_8$ and $(CA)_{10}$. The allele $(CA)_{11}$ was detected only in the sheep strain isolates from Brazil.

Microsatellite data from the Egmsca1 locus showed significant heterozygote deficiency in the Argentinean sheep strain population ($P<0{\cdot}05$). The Brazilian sheep strain population agreed with Hardy-Weinberg genotype expectations ($P>0{\cdot}05$). The genotype frequencies and heterozygosities for each population are presented in Table 4.

No genetic differentiation was detected between the Brazilian and Argentinean sheep strain populations ($P>0{\cdot}05$). The values of Fis and Fst were, respectively, 0·48 and $-0{\cdot}02$, considering the whole sample of sheep strain isolates separated into 2 populations (Brazil and Argentina).

Despite our attempts to optimize PCR conditions, the analyses on the Egmsga1 locus resulted in a high number of amplification bands, and were not used in the population study.

DISCUSSION

The isolation of microsatellites and their flanking sequences from genomes is frequently a laborious and expensive exercise, except for those organisms whose genomes have been completely sequenced or from which a high number of sequences has been deposited in Data Banks. The analysis of microsatellites by Southern blot using repeated oligonucleotides as probes can provide valuable information before the repeat isolation.

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

Table 3. Egmsca1 allele frequencies in each population

Population	Alleles			
	$(CA)_7$	$(CA)_8$	$(CA)_{10}$	$(CA)_{11}$
Brazil – sheep strain	0,000	0,039	0,922	0,039
Brazil – cattle strain	1,000	0,000	0,000	0,000
Argentina – sheep strain	0,000	0,059	0,941	0,000
Argentina – Tasmanian sheep strain	0,000	0,200	0,800	0,000
Argentina – camel strain	1,000	0,000	0,000	0,000

Table 4. Genotype frequencies and heterozygosities for the Egmsca1 locus in each population analysed

Population*	Genotype frequencies						Hobs†	Hexp‡
	$(CA)_7/(CA)_7$	$(CA)_8/(CA)_8$	$(CA)_{10}/(CA)_{10}$	$(CA)_{11}/(CA)_{11}$	$(CA)_8/(CA)_{10}$	$(CA)_{10}/(CA)_{11}$		
SS – B	0	0	0·868	0·026	0·080	0·026	0·11	0·15
CS – B	1·000	0	0	0	0	0	0	—
SS – A	0	0·059	0·941	0	0	0	0	0·11
TSS – A	0	0	0·600	0	0·400	0	0·40	0·36
CmS – A	1·000	0	0	0	0	0	0	—

* SS – B, sheep strain from Brazil; CS – B, cattle strain from Brazil; SS – A, sheep strain from Argentina; TSS – A, Tasmanian Sheep from Argentina; CmS – A, camel strain from Argentina.
† Hobs, observed heterozygosity.
‡ Hexp, expected heterozygosity (calculated only for non-monomorphic populations).

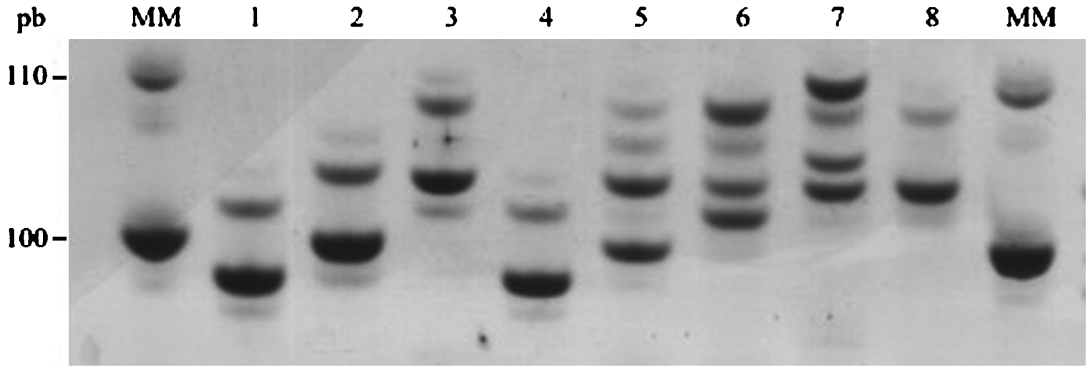


Fig. 3. Analysis of different Egmsca1 genotypes. The regions containing CA-repeats were amplified by PCR and analysed by polyacrylamide gel electrophoresis. The gel was stained with $AgNO_3$. MM, molecular marker; (1) homozygote CA_7/CA_7; (2) homozygote CA_8/CA_8; (3) homozygote CA_{10}/CA_{10}; (4) homozygote CA_7/CA_7; (5) heterozygote CA_8/CA_{10}; (6) heterozygote CA_9/CA_{10}; (7) heterozygote CA_{10}/CA_{11}; (8) control (DNA fragment amplified by PCR from the recombinant plasmid containing 10 CA repeats of Egmsca1 locus).

Among the repeated oligonucleotides we used for hybridization with DNA of *E. granulosus*, the probes $(GT)_{15}$, $(CAA)_{10}$, $(CATA)_{10}$ and $(CT)_{15}$ showed the greatest number of hybridized bands in *E. granulosus*. GT repeats (or CA) have been reported as the most frequent microsatellite motifs in several animal genomes (Tautz, 1989; Weber & May, 1989; Stallings *et al.* 1991; Katti, Ranjekar & Gupta, 2001). AAC microsatellites are among the most frequent trinucleotide repeats found in the human, *Drosophila* and yeast genomes, and CATA microsatellites are the most frequent tetranucleotide repeats found in *Drosophila* (Katti *et al.* 2001). We conclude that the $(AT)_{15}$ and $(CG)_{15}$ repeats are absent in the *E. granulosus* genome, or present at very low frequencies, because no hybridization signal was observed for these probes. A lack of CG repeats was reported in human, mouse, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana* and yeast genomes (Dokholyan *et al.* 2000; Katti *et al.* 2001). Moreover, our attempts to isolate (TA)n repeats have all failed. We were able to isolate CA/GT and GA/CT microsatellites, proving that the analysis by genomic Southern blot is helpful to study their distribution and abundance as a previous step in the isolation strategy.

Since microsatellites are unstable sequences, due to their repetitive nature (Eisen, 1999), DNA

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press

replication slippage during the asexual reproduction stage of *E. granulosus* could lead to differences among protoscoleces. The Egmsca1 locus was successfully amplified in 150 protoscoleces taken at random from 1 single cyst and no genetic differences were observed. Twenty protoscoleces isolated from each of 5 additional cysts showed no within-cyst differentiation too. These results validate the use of protoscoleces from a whole hydatid cyst in studies of *E. granulosus* population genetics, utilizing microsatellites as molecular markers.

The Egmsca1 genotype distributions confirmed that both selfing and outcrossing occur in *E. granulosus*, as reported by other authors (Lymbery, Thompson & Hobbs, 1990; Haag *et al*. 1999). The negative value found for Fst in this locus indicates that the Brazilian and Argentinean populations from the sheep strain are not structured. The analysis of a larger number of microsatellite loci will be required to enhance the sensitivity of detection of within-strain geographical differences. Indeed, only one allele [$(CA)_7$] was shared by the cattle and camel strains. The phylogenetic relatedness of these strains was previously shown in three parsimonious trees constructed from mitochondrial data sets (Bowles *et al*. 1995).

In conclusion, microsatellites are fully distributed in *E. granulosus* genome, and their polymorphism is useful to detect strain as well as some geographical differences. For this reason, we are working on the isolation and characterization of more loci, which could be used as a panel for different studies on population structure, genetic epidemiology and diagnosis of hydatid disease.

M. L. B.-S. was a recipient of a fellowship from CNPq (Brazil) and subsequently from CAPES (Brazil). This work was supported by CNPq, PADCT/CNPq and FAPERGS (Brazil). We would like to thank Dr Mara C. Rosenzvit for the critical reading of the manuscript.

REFERENCES

BINZ, T., REUSCH, T. B. H., WEDEKING, C., SCHÄRER, L., STERNBERG, J. M. & MILINSKI, M. (2000). Isolation and characterization of microsatellite loci from the tapeworm *Schistocephalus solidus*. *Molecular Ecology* **9**, 1926–1927.

BOWLES, J., BLAIR, D. & McMANUS, D. P. (1995). A molecular phylogeny of the genus *Echinococcus*. *Parasitology* **110**, 317–328.

BRETAGNE, S., ASSOULINE, B., VIDAUD, D., HOUIN, R. & VIDAUD, M. (1996). *Echinococcus multilocularis*: microsatellite polymorphism in U1 snRNA genes. *Experimental Parasitology* **82**, 324–328.

DOKHOLYAN, N. V., BULDYREV, S. V., HAVLIN, S. & STANLEY, H. E. (2000). Distributions of dimeric tandem repeats in non-coding and coding DNA sequences. *Journal of Theoretical Biology* **202**, 273–282.

ECKERT, J., CONRATHS, F. J. & TACKMANN, K. (2000). Echinococcosis: an emerging or re-emerging zoonosis? *International Journal of Parasitology* **30**, 1283–1294.

EISEN, J. A. (1999). Mechanistic basis for microsatellite instability. In *Microsatellites: Evolution and Applications* (ed. Goldstein, D. B. & Schlötterer, C.), pp. 34–48. Oxford University Press, Oxford.

GOUDET, J., RAYMOND, M., DE MEEÜS, T. & ROUSSET, F. (1996). Testing differentiation in diploid populations. *Genetics* **144**, 1933–1940.

HAAG, K. L., ARAÚJO, A. M., GOTTSTEIN, B., SILES-LUCAS, M., THOMPSON, R. C. A. & ZAHA, A. (1999). Breeding systems in *Echinococcus granulosus* (Cestoda; Taeniidae): selfing or outcrossing? *Parasitology* **118**, 63–71.

KAMENETZKY, L., GUTIERREZ, A. M., CANOVA, S. G., HAAG, K. L., GUARNERA, E. A., PARRA, A., GARCIA, G. E. & ROSENZVIT, M. C. (2002). Several strains of *Echinococcus granulosus* infect livestock and humans in Argentina. *Infection, Genetics and Evolution* **63**, 1–8.

KATTI, M. V., RANJEKAR, P. K. & GUPTA, V. S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution* **18**, 1161–1167.

LYMBERY, A. J., THOMPSON, R. C. A. & HOBBS, R. P. (1990). Genetic diversity and genetic differentiation in *Echinococcus granulosus* (Batsch, 1786) from domestic and sylvatic hosts on the mainland of Australia. *Parasitology* **101**, 283–289.

LITT, M. & LUTY, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**, 397–401.

OLIVEIRA, R. P., BROUDE, N. E., MACEDO, A. M., CANTOR, C. R., SMITH, C. L. & PENA, S. D. J. (1998). Probing the genetic structure of *Trypanosoma cruzi* with polymorphic microsatellites. *Proceedings of the National Academy of Sciences, USA* **95**, 3776–3780.

RAYMOND, M. & ROUSSET, F. (1995*a*). Genepop (version 1.2): population genetics software for exact tests and ecumenecism. *Journal of Heredity* **86**, 248–249.

RAYMOND, J. & ROUSSET, F. (1995*b*). An exact test for population differentiation. *Evolution* **49**, 1280–1283.

REFSETH, U. H., FANGAN, B. M. & JAKOBSEN, K. S. (1997). Hybridization capture of microsatellites directly from genomic DNA. *Electrophoresis* **18**, 1519–1523.

SAMBROOK, J. & RUSSELL, D. W. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd Edn. Cold Spring Harbor Laboratory Press, New York.

STALLINGS, R. L., FORD, A. F., NELSON, D., TORNEY, D. C., HILDEBRAND, C. E. & MOYZIZ, R. K. (1991). Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics* **10**, 807–815.

TAUTZ, D. (1989). Hypervariability of simple sequences as a general source of polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463–6471.

THOMPSON, R. C. A. & LIMBERY, A. J. (1995). *Echinococcus and Hydatid Disease*. CAB International, Wallingford.

WEBER, J. L. & MAY, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388–396.

WEIR, B. S. (1990). *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.

WEIR, B. S. & COCKERHAM, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

https://doi.org/10.1017/S0031182003003184 Published online by Cambridge University Press