Breeding systems in *Echinococcus granulosus* (Cestoda; Taeniidae): selfing or outcrossing?

K. L. HAAG[1]*, A. M. ARAÚJO[1], B. GOTTSTEIN[3], M. SILES-LUCAS[3], R. C. A. THOMPSON[4] *and* A. ZAHA[2]

[1] *Departamento de Genética, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, Porto Alegre, 91501-970, Brazil*
[2] *Departamento de Biotecnologia, Universidade Federal do Rio Grande do Sul, Caixa Postal 15005, Porto Alegre, 91501-970, Brazil*
[3] *Institute of Parasitology, University of Berne, Laengass Strasse 122, PO Box 8466, Berne, CH-3001, Switzerland*
[4] *World Health Organization Collaborating Centre for the Molecular Epidemiology of Parasitic Infections, Division of Veterinary and Biomedical Sciences, Murdoch University, Murdoch, Western Australia 6150, Australia*

(*Received 10 April 1998; revised 20 June 1998; accepted 20 June 1998*)

SUMMARY

We used the PCR–SSCP method followed by sequencing in order to assess the genetic variability of coding and non-coding parts of the genome of *Echinococcus granulosus* (Cestoda; Taeniidae) and to test whether or not the parasite populations are mainly self-fertilizing. For this, we analysed a sample of 110 *E. granulosus* metacestode isolates collected from different geographical regions (Southern Brazil, Europe and Australia) and from different intermediate hosts (ovine, bovine, human, macropod, swine and equine). Using appropriate controls, we were able to identify 4 strains in that sample (sheep, cattle, pig and horse strains). The high degree of genetic differentiation between strains, but not within, and the monomorphism found in most loci (EgAg4, EgActII, EgHbx2 and EgAg6 – non-coding – EgAgB/1 and EgND1 – coding) indicated that they are largely selfed. On the other hand, outcrossing was also shown to occur, since 5 potential hybrid genotypes between cattle and sheep strains were found in populations of Southern Brazil, but absent in other geographical areas. We suggest that both processes are adaptive. The article also reports, for the first time, the occurrence of the *E. granulosus* cattle strain in South America.

Key words: *Echinococcus granulosus*, SSCP, strain, genetic variability, selfing.

INTRODUCTION

The origin and the degree of genetic differentiation in the eucestode parasite *Echinococcus granulosus* have been controversial subjects (Lymbery, 1995; Lymbery & Thompson, 1996; Thompson, Lymbery & Constantine, 1995). The species normally uses domestic and wild dogs as definitive hosts and domestic and wild ungulates as intermediate hosts, but a range of different herbivore species can be infected by the larval stage (metacestode); man can also be infected. *E. granulosus* is maintained in several biological cycles involving a series of apparently host-adapted 'strains', some with well-defined geographical distributions (see Rausch, 1995, Schantz *et al*. 1995 and Thompson, 1995 for a review).

During the last 5 years a number of techniques have been used to assess *E. granulosus* genetic variability (see references in Eckert & Thompson, 1997). However, most studies have focused on strain identification and characterization. Little effort has been made to investigate evolutionary problems by population genetic studies, probably due to difficulties with obtaining adequate biological samples.

To date, only 3 studies have been published using population genetic approaches in *E. granulosus* (Lymbery & Thompson, 1988; Lymbery, Thompson & Hobbs, 1990; Lymbery, Constantine & Thompson, 1997). All were based on multilocus enzyme electrophoresis data and attempted to elucidate genetic variability patterns and processes in Australian populations of *E. granulosus*. The authors addressed the question of strain origin, referring to 2 hypothetical models (1) strains would arise via selection by the host on the metacestode, which is able to amplify new mutations by asexual reproduction, and the adult worms would be mainly self-fertilizing (Smyth & Smyth, 1964); (2) strains would originate by adaptation to different host species or by geographical isolation, and the adult worms would be cross-fertilizing (Rausch, 1985). The authors did not find evidence of linkage disequilibrium among loci, indicating cross-fertilization of adults. Also, using F-statistics they did not find significant differences (F_{ST}) among strains or geographical

* Corresponding author: Departamento de Genética, UFRGS, Caixa Postal 15053, 91501-970, Porto Alegre, Brazil. Tel: 0055 51 316 6717. Fax: 0055 51 319 2011. E-mail: haag@dna.cbiot.ufrgs.br

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

populations. Indeed, more variability was found within than between those groups. Nevertheless, the absence of heterozygotes suggested self-fertilization. Their interpretation of these data was that although self-fertilization predominates, at least in colonizing populations of *Echinococcus*, it occurs principally through cross-insemination by genetically identical clone-mates in the intestine of the definitive host (Thompson & Lymbery, 1996; Lymbery *et al*. 1997). Further, they considered that widespread dispersal via intermediate host movements and occasional outcrossing prevent both genetic differentiation between populations and substantial linkage disequilibrium, both of which normally accompany self-fertilization.

In order to answer similar questions in *E. granulosus* populations of Southern Brazil, Europe and Australia (here defined as groups of isolates from the same strain, collected from different hosts, in the same geographical area) we used the PCR–SSCP method (Single Strand Conformation Polymorphism of Polymerase Chain Reaction products) followed by sequencing. In our view, the method has 3 advantages (1) it generates allele frequency and sequencing data sets and both may be statistically tested using appropriate population genetic models; (2) it is possible to analyse coding and non-coding parts of the genome; (3) it is much more sensitive than the multilocus enzyme electrophoresis technique and therefore reveals more genetic variation.

MATERIALS AND METHODS

Molecular analyses

A total of 110 *E. granulosus* (Eg) isolates from Southern Brazil, Europe (Germany, Ireland, Poland, Spain and Switzerland) and Australia (mainland and Tasmania) were used for genomic DNA extraction and further analyses. Within each region the isolates came from different localities. All procedures were done using metacestode tissue as starting material, collected from ovine, bovine, macropod, human, swine and equine hosts. DNA extraction was done by standard procedures (McManus & Simpson, 1985).

For each isolate, 6 different targets were amplified by PCR (Table 1). Two targets (ActII and Hbx2) are non-coding introns from the respective genes characterized in *E. granulosus* (Silva *et al*. 1993; Vispo & Ehrlich, 1994). Two are coding regions: part of the nuclear antigen B gene (AgB/1 – Frosch *et al*. 1994), which codes for a protein with inhibition of elastase and neutrophil chemotaxis activity (Shepherd, Aitken & McManus, 1991), and the other is part of the mitochondrial NADH dehydrogenase 1 gene (ND1 – Bowles & McManus, 1993). The last 2 targets (Ag4 and Ag6) were non-coding 5′ and 3′ flanking regions, containing regulatory sequences of the respective genes (Jaqueline Rodrigues, personal communication). Ag4 contains also a small (44 nt) intron immediately after the first methionine codon. The gene codes for a cytosolic malate dehydrogenase (Rodrigues, Ferreira & Zaha, 1993), while Ag6 codes for a calcium-binding protein (Rodrigues *et al*. 1997).

The designed primers were shown to be specific for *Echinococcus* DNA, since no amplification occurred using host DNA as template. The primer sequences and amplicon size are shown in Table 1. Subsequent to PCR, the denatured PCR products from each test run were used for SSCP screening. PCR and SSCP conditions were described previously, in a similar study with *E. multilocularis* (Haag *et al*. 1997). For visualization of the SSCP electrophoretic resolution, we used conventional silver staining techniques.

Genotype and strain identification

The secondary structures of single-stranded PCR fragments run slower than the double strands (Fig. 1). Heterozygotes were readily identified by the multiple single-strand banding patterns and by the heteroduplexes of renatured DNA. Allele identification was done cutting out the single-strand bands from the fresh, stained SSCP gels, washed several times in 1 ml of distilled water and eluted for 20 min at 94 °C in 50 μl 1 × PCR buffer (Gibco); 0·1 % Triton X-100. One μl of the eluted single strands was used for re-amplification with the corresponding primers, following the same procedures described by Haag *et al*. (1997). The purified PCR products (Qiagen) were used for direct fluorescence sequencing of double-stranded PCR products using a 373A system (Applied Biosystems).

At least 4 isolates from each SSCP pattern were chosen for sequencing. Heterozygote genotypes were assigned only to the patterns for which 2 different sequences were obtained. In most cases the respective homozygote patterns were also found. All SSCP patterns obtained for each target were interpreted in the same manner. We are aware that the sensitivity of the technique for point mutations in fragments with less than 400 bp is around 90–100 % (Lessa & Applebaum, 1993; Sheffield *et al*. 1993), but for simplicity, we assumed that phenotype identity was due to genotype identity. Indeed, no sequencing differences were found among identical SSCP bands from different individuals.

After genotype identification, isolates collected from ovine, bovine, macropod, human, swine and equine hosts were grouped in strains (Table 2). For this, we included in the analyses some isolates that were previously characterized using RAPD and isoenzyme markers (Siles-Lucas, Benito & Cuestra-Bandera, 1996). Taking them as references, we were

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

Table 1. Primer sequences and amplicon size of the six *Echinococcus granulosus* genomic targets analysed in this study

Target	Region*	Size (bp)	Primers
EgAg4	1, 3	106	5′-TGACAAACTCTGGGGTA-3′ 5′-AGAACCCTAAGAGGACC-3′
EgAgB/1	2	102	5′-CGTGATCCGTTGGGTCAG-3′ 5′-GGCACCTCTATTCACCTTCA-3′
EgAg6	2, 4	403–405	5′-ACCCTCGGTTCTACGTC-3′ 5′-TCAGCAGAACGGCATGAGAG-3′
EgHbx2	3	329–331	5′-TTCTCCTCTAGCCAGGTCCA-3′ 5′-TATAGCGCCGATTCTGGAAC-3′
EgActII	3	268	5′-GTCTTCCCCTCTATCGTGGG-3′ 5′-CTAATGAAATTAGTGCTTGTGCGC-3′
EgND1	5	141	5′-TTCTAGGTATTCTTTGTTGTG-3′ 5′-CAAGCTTCATCAACAACTATAA-3′

* 1 = 5′ flanking region; 2 = nuclear coding region; 3 = intron; 4 = 3′ flanking region; 5 = mitochondrial coding region.

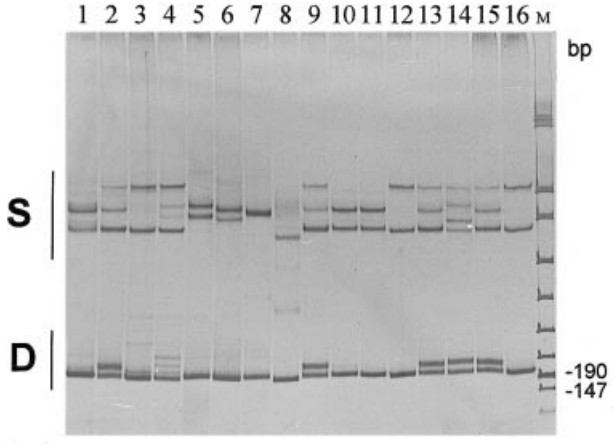

Fig. 1. *Echinococcus* Ag4 SSCP patterns. Bands indicated by S are secondary structures of single-stranded DNA, and those indicated by D are double strands. M is marker VIII (Boehringer). Patterns 1, 10 and 11 are homozygotes A1/A1; 3, 12 and 16 are A2/A2; 3, 9, 13 and 15 are heterozygotes A1/A2; 5 is A3/A3; 4 is A2/A3; 14 is A6/A2; 6, 7 and 8 are A5/A5, A4/A4 and an *E. vogeli* pattern, respectively. The heterozygote A2/A3 (lane no. 4) indicated crossing between sheep and cattle strains (see text).

able to recognize 4 strains within our sample (1) the sheep strain, which uses a broader range of host species (genotypes s1–s12); (2) the cattle strain (genotypes c1–c3); (3) the pig strain, which was also found in human hosts (genotype p1) and (4) the horse strain (genotype h1).

Statistical analyses

E. granulosus sequences were aligned using the GCG Package (version 8, 1994). All statistical analyses were performed with Arlequin (version 1.0), a program developed by Schneider *et al.* (1997).

Non-random association of nucleotides among loci (linkage disequilibrium) was tested by an exact test on contingency tables (Raymond & Rousset, 1995). The test consists of obtaining the probability of finding a table with the same marginal totals, which has a probability equal to or less than the observed table, and a Markov chain is used to explore adequately the space of possible tables. Linkage disequilibrium based on allele frequency data was tested using a likelihood-ratio test, whose empirical distribution is obtained by a permutation procedure (Slatkin & Excoffier, 1996).

The Analysis of Molecular Variance (AMOVA – Michalakis & Excoffier, 1996) within the sheep strain was performed according to the following hierarchical levels (1) individuals or isolates; (2) meta-populations of individuals collected from the same host species in the same geographical area (Brazil, Europe or Australia); (3) supra-populations of isolates belonging to different meta-populations of the same geographical area.

RESULTS

The most polymorphic loci found were Ag4 (5′ flanking region and intron) and Ag6 (3′ flanking region), both with 6 alleles; followed by Act II (intron) and ND1 (mitochondrial coding region), with 4 alleles; and Hbx2 (intron) and AgB/1 (nuclear coding region) with 3 alleles (Fig. 2 and Table 2). From the 1351 nucleotide sites analysed, 117 were polymorphic, 5 of which were sites with indels.

A highly significant non-random association of nucleotides ($P < 0{\cdot}01$) was found among all 6 loci using the Markov chain exact test on the whole *E. granulosus* sample. A simple visual inspection of the data in Table 2 would also lead to that conclusion, since only 17 out of the 1 587 600 possible genotypes occurred (given the actual number of alleles and loci). It is clear, from our data, that particular gene

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

Table 2. Genotypes of the six loci analysed in this study and their respective hosts and geographical localizations (nucleotide differences among alleles are shown in Fig. 2)

			Loci						
Genotype	Region*	Host†	EgAg4	EgAgB/1	EgAg6	EgHbx2	EgActII	EgND1	*n*
s1	SB, E	O, B, H	A1/A1	B1/B1	D1/D1	E1/E1	F1/F1	G1	8
s2	A	O, M	A1/A1	B1/B1	D1/D2	E1/E1	F1/F1	G1	2
s3	A	M	A1/A1	B1/B1	D2/D2	E1/E1	F1/F1	G1	1
s4	SB	O	A1/A2	B1/B1	D1/D1	E1/E1	F1/F1	G1	12
s5	SB, E	O, B	A1/A2	B1/B1	D1/D2	E1/E1	F1/F1	G1	20
s6	SB, E, A	O, H	A1/A2	B1/B1	D2/D2	E1/E1	F1/F1	G1	9
s7	SB, E	O, B, H	A2/A2	B1/B1	D1/D1	E1/E1	F1/F1	G1	10
s8	SB, E	O, B	A2/A2	B1/B1	D1/D2	E1/E1	F1/F1	G1	6
s9	A	O, M	A2/A2	B1/B1	D2/D2	E1/E1	F1/F1	G1	2
s10	A	O	A1/A2	B1/B1	D2/D3	E1/E1	F1/F1	G1	2
s11	A	O	A2/A6	B1/B1	D2/D2	E1/E1	F1/F1	G1	1
s12	SB	B	A2/A3	B1/B1	D1/D2	E1/E1	F1/F1	G1	2
c1	SB	B	A2/A2	B2/B2	D4/D4	E2/E2	F2/F2	G2	3
c2	SB	B	A2/A3	B2/B2	D4/D4	E2/E2	F2/F2	G2	3
c3	SB, E	B	A3/A3	B2/B2	D4/D4	E2/E2	F2/F2	G2	3
p1	E	S, H	A4/A4	B2/B2	D5/D5	E2/E2	F3/F3	G3	14
h1	E	E	A5/A5	B3/B3	D6/D6	E3/E3	F4/F4	G4	12
Total									110

* SB = Southern Brazil; E = Europe; A = Australia.
† O = ovine; B = bovine; H = human; M = macropod; S = swine; E = equine.

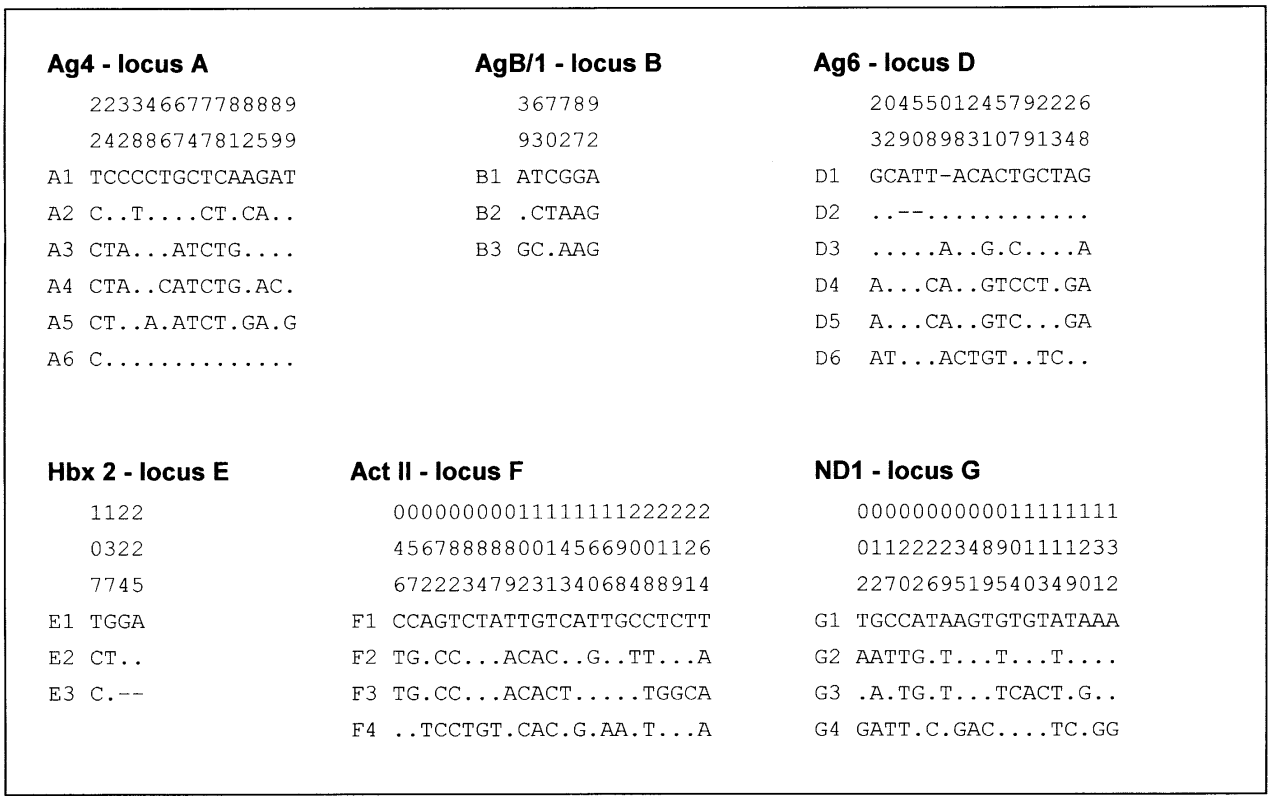

Fig. 2. Alignment of the variable nucleotide sites of alleles in the six *Echinococcus granulosus* loci analysed in this study. Site positions within each amplified target are indicated with numbers. Ag4 positions 22–48 are in the 5′ flanking region, while positions 69–99 are inside the intron.

combinations are maintained, and that these combinations are associated with groups of isolates belonging to different strains.

Gene flow among strains is strongly restricted, as indicated by the high and significant values of pairwise F_{ST} among populations belonging to dif-

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

Table 3. Pairwise F_{ST} (Weir & Cockerham, 1984) among *Echinococcus granulosus* populations and strains based on sequencing (above diagonal) and allele frequency data (below diagonal)

	Sheep (Europe)	Sheep (S. Brazil)	Sheep (Australia)	Cattle	Pig	Horse
Sheep (Europe)		0·0494*	0·0075$^{N.S.}$	0·8978**	0·9538**	0·9711**
Sheep (S. Brazil)	0·0914*		0·2042*	0·9020**	0·9520**	0·9706**
Sheep (Australia)	0·0912$^{N.S.}$	0·3412**		0·8984**	0·9746**	0·9836**
Cattle	0·8685**	0·8870**	0·8907**		0·9164**	0·9770**
Pig	0·9047**	0·9166**	0·9444**	0·9481**		1·0000**
Horse	0·9003**	0·9134**	0·9394**	0·9677**	1·0000**	

* $P < 0·05$; ** $P < 0·01$; S. Brazil = Southern Brazil.
$^{N.S.}$, Not significant.

Table 4. Multilocus genotype frequencies within populations of sheep, cattle, pig and horse *Echinococcus granulosus* strains

	Sheep			Cattle	Pig	Horse
	S. Brazil*	Europe	Australia	S. Brazil	Europe	Europe
n	37	28	10	9	14	12
s1	0·14	0·11	—	—	—	—
s2	—	—	0·20	—	—	—
s3	—	—	0·10	—	—	—
s4	0·32	—	—	—	—	—
s5	0·14	0·54	—	—	—	—
s6	0·02	0·21	0·20	—	—	—
s7	0·19	0·11	—	—	—	—
s8	0·14	0·03	—	—	—	—
s9	—	—	0·20	—	—	—
s10	—	—	0·20	—	—	—
s11	—	—	0·10	—	—	—
s12	0·05	—	—	—	—	—
c1	—	—	—	0·33	—	—
c2	—	—	—	0·34	—	—
c3	—	—	—	0·33	—	—
p1	—	—	—	—	1·00	—
h1	—	—	—	—	—	1·00

* S. Brazil = Southern Brazil.

ferent strains (Table 3). The only indication in our data of inter-strain fertilization was that between the cattle and the sheep strains in Southern Brazil. Genotypes c1 and c2 (Table 2) suggest that allele A2 (Ag4, see Fig. 1 for heterozygote identification) introgressed from the sheep strain into the cattle strain. Indeed, the lowest inter-strain F_{ST} value found in Table 3 is between the sheep strain populations and the cattle strain, either calculating it from the sequencing or from the allele frequency data.

The strains appear to be highly homogeneous evolutionary units. The pig and the horse strains had no variability, the cattle strain showed polymorphism in only 1 locus (Ag4) and the sheep strain was polymorphic for Ag4 and Ag6 (Tables 4 and 5). These loci showed linkage disequilibrium ($P < 0·01$) in the 3 sheep strain populations tested (Southern Brazil, Australia and Europe). Furthermore, the mean number of pairwise nucleotide differences and nucleotide diversity estimates (Table 5) were all close to zero.

However, observed heterozygosities in populations of the sheep strain were quite high, when compared to gene diversities (expected heterozygosities – Table 5). Indeed, a significant departure from Hardy–Weinberg (HW) equilibrium due to excess of heterozygotes was found for Ag4 in the sheep strain from Europe ($P < 0·05$). All other polymorphic loci, including Ag4 from the cattle strain (Southern Brazil), agreed with HW expectations.

Geographical differentiation seems to be irrelevant when compared to strain differentiation. The pairwise F_{ST} values among populations of the sheep strain were quite low compared to those obtained for different strains (Table 3). Furthermore, the AMOVA based on sequencing data revealed that

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

Table 5. Observed heterozygosity (H_o), gene diversity (H_s), proportion of polymorphic loci (P), mean number of pairwise nucleotide differences (p_i) and nucleotide diversity (π) within populations of *Echinococcus granulosus* sheep, cattle, pig and horse strains*

	Sheep			Cattle	Pig	Horse
	S. Brazil†	Europe	Australia	S. Brazil	Europe	Europe
n	37	28	10	9	14	12
H_o	0·1730	0·2643	0·1800	0·0680	0·0000	0·0000
H_s	0·1629	0·1998	0·1170	0·1000	0·0000	0·0000
P	0·3333	0·3333	0·3333	0·1667	0·0000	0·0000
p_i	3·5456 (1·8447)‡	3·6557 (1·9075)	4·2082 (2·2812)	4·2558 (2·2128)	0·0000 (0·0000)	0·0000 (0·0000)
π	0·0026 (0·0015)	0·0027 (0·0016)	0·0031 (0·0019)	0·0032 (0·0018)	0·0000 (0·0000)	0·0000 (0·0000)

* Nei (1987).
† Numbers in parentheses are standard deviations based on sample and stochastic errors.
‡ S. Brazil = Southern Brazil.

Table 6. Analysis of Molecular Variance (AMOVA)* within *Echinococcus granulosus* sheep strain based on sequencing data (see Materials and Methods section for hierarchical structure characterization)

Source of variation	D.F.	Sum of squares	Variance component	Percentage of variation
Among supra-populations	1	10·947	0·1474	6·44
Among meta-populations within supra-populations	3	5·362	−0·0183	−0·80
Within meta-populations	125	270·045	2·1604	96·36
Total	129	286·354	2·2894	

* Michalakis & Excoffier (1996).

Table 7. Analysis of Molecular Variance (AMOVA)* within *Echinococcus granulosus* sheep strain based on allele frequency data (see Materials and Methods section for hierarchical structure characterization)

Source of variation	D.F.	Sum of squares	Variance component	Percentage of variation
Among supra-populations	1	3·357	0·0485	16·37
Among meta-populations within supra-populations	3	0·920	−0·0034	−0·75
Among individuals within meta-populations	60	22·508	−0·0778	−12·13
Within individuals	65	34·500	0·5308	96·52
Total	129	61·285	0·5387	

* Michalakis & Excoffier (1996).

most variability (96·36%) occurs within meta-populations (Table 6). Using allele frequencies and adding the individual level in the hierarchical structure (see Materials and Methods section), it was shown that variation was higher (96·52%) within individuals (Table 7).

DISCUSSION

The analysis of coding and non-coding parts of *E. granulosus* genome revealed interesting biological features. First, selfing seems to be the predominant reproductive system in *Echinococcus*, associated with

https://doi.org/10.1017/S0031182098003485 Published online by Cambridge University Press

a low level of genetic variability in founding populations. There is a high degree of differentiation among strains, but the genetic variability within them is reduced (sheep and cattle strains) or non-existent (pig and horse strains). A similar situation was found in *E. multilocularis* (Haag *et al*. 1997), where even lower values of genetic diversity were obtained.

Perhaps the most striking evidence of a mainly self-fertilizing reproductive system in *E. granulosus* are the heterozygote deficiencies found by Lymbery & Thompson (1988) and Lymbery *et al*. (1997). In the first article the authors did not provide estimates of inbreeding, but they ranged from F = 0·643 (esterase locus, population of Tasmania) to F = 1 (all other polymorphic loci). In the second one, they ranged from F = 0·843 (esterase locus, population of Tasmania) to F = 1 (all other polymorphic loci). Since no multilocus associations (linkage disequilibrium) were found, Lymbery *et al*. (1997) suggested that most selfing is achieved by geitonogamy, a process in which the ova are fertilized by sperm of a separate, but clonal individual. In essence, the mechanism of geitonogamy is the same as cross-fertilization, but normally results in a higher level of homozygosis in the whole population, because only genetically identical individuals are involved.

In the present study populations of the *E. granulosus* sheep strain showed a highly significant linkage disequilibrium between the 2 polymorphic loci (Ag4 and Ag6) in 3 geographical areas (Southern Brazil, Europe and Australia). However, these loci did not show heterozygote deficiencies, most populations were in agreement with HW expectations. The mere presence of heterozygotes and the fact that most genetic variation is found within individuals clearly indicate that outcrossing does occur in *E. granulosus*. Moreover, Ag4 had an excess of heterozygotes in the sheep strain population of Europe. One explanation for this finding is that balancing selection is acting on locus Ag4 and that, if selfing occurs, it is likely to be through geitonogamy. Provided that Ag4 and Ag6 contain regulatory sequences, heterozygosity could enhance parasite plasticity. Of course, experimental approaches ought to be used to test this hypothesis. We cannot rule out the possibility that the association found between alleles in both loci is due to inbreeding or to physical linkage, but it could also be maintained by selection. Furthermore, theoretical studies of 2-locus selection models with partial selfing indicate that the heterozygosity of a population may increase as the selfing rate is increased (Holsinger & Feldman, 1985).

Two other important findings were (1) the occurrence of the cattle strain in Southern Brazil and (2) the indication of cross-fertilization between sheep and cattle strains in the same region. Alleles A2 and A3 (Ag4) seem to be exchanged, but note that allele A3 appeared only in bovine hosts, suggesting that it could be detrimental in another host (see Table 2). Considering both strains, the frequency of the hybrid genotype (A2/A3) was 11%. In Southern Brazil sheep and cattle are raised together or nearby over large areas, where parasite prevalence in dogs, which are fed with animal viscera, ranges from 20 to 28% (Ferreira, 1993). Since there may be many thousands of *E. granulosus* adults in a heavily infected host (Schantz *et al*. 1995), it is not unreasonable to imagine adult worms from both strains being in contact, within the same dog, maturing sexually at the same time and mating.

The lack of other shared polymorphisms between both strains could be explained by 2 reasons (1) high selfing rates of adult worms associated with asexual amplification in the metacestode, leading to monomorphism in most loci and (2) strong selection exerted by the host, which contributes to the elimination of detrimental alleles. We also excluded the possibility that the Ag4 shared polymorphism pre-dates the split of cattle and sheep strains, because the molecular phylogenetic analyses performed using the present set of data (Haag *et al*. manuscript in preparation), or mitochondrial COI and ND1 genes (Lymbery, 1995) suggest that both diverged before they invaded the American continent. Indeed, these analyses show that the cattle strain is much more related to the pig than to the sheep strain. Furthermore, if it would be an ancient polymorphism, it should occur also in cattle and sheep parasites from other geographical areas (Europe and Australia), which did not happen.

In conclusion, our results support those of Lymbery *et al*. (1990, 1997) suggesting that both selfing and cross-fertilization occur in *E. granulosus* populations. This means that both the Smyth & Smyth (1964) and the Rausch (1985) hypotheses about strain evolution are actually true. Selfing would be an easy and fast way to respond to host selection; in a patchy and heterogeneous environment, parasite populations could diverge genetically in a few generations. Indeed, the most variable strain (sheep strain) had the lowest intermediate host specificity. Outcrossing, on the other hand, would be necessary to prevent the elimination of adaptive heterozygous genotypes. Since both cross and self-insemination have been reported in parasitic platyhelminths (Nollen, 1983), broader studies should be designed to shed some light on the evolutionary role of these processes in their natural populations.

Thanks to Dr Z. S. Pawlowski for the genomic DNAs of *E. granulosus* isolates from Poland and to Mrs Corin Müller for her collaboration with the SSCPs. We are also indebted to Dr A. Lymbery for his critical revision of the manuscript. The support by the Swiss National Science Foundation (project no. 31-45575.95), PADCT/CNPq (Proc. 620081/95-3), EEC (DG XII CI 10284−0), the 'Jubiläumsstiftung der Schweitzerischen Lebensversicherungs- und Rentenanstalt für Volksgesundheit und

Medizinische Forschung' and the 'Sandoz-Stiftung zur Förderung der medizinisch-biologischen Wissenschaften' is gratefully acknowledged.

REFERENCES

BOWLES, J. & MCMANUS, D. P. (1993). NADH dehydrogenase 1 gene sequences compared for species and strains of the genus *Echinococcus*. *International Journal for Parasitology* **23**, 969–972.

ECKERT, J. & THOMPSON, R. C. A. (1997). Intraspecific variation of *Echinococcus granulosus* and related species with emphasis on their infectivity to humans. *Acta Tropica* **64**, 19–34.

FERREIRA, H. B. (1993). Antígenos para o imunodiagnóstico da hidatidose humana: caracterização de preparações antigênicas de líquido hidático e clonagem de genes que codificam antígenos de *Echinococcus granulosus*. Ph.D. thesis, Universidade Federal de Rio Grande do Sul.

FROSCH, P., HARTMANN, M., MÜHLSCHLEGEL, F. & FROSCH, M. (1994). Sequence heterogeneity of the echinococcal antigen B. *Molecular and Biochemical Parasitology* **64**, 171–175.

HAAG, K. L., ZAHA, A., ARAÚJO, A. M. & GOTTSTEIN, B. (1997). Reduced genetic variability within coding and non-coding regions of *Echinococcus multilocularis* genome. *Parasitology* **115**, 521–530.

HOLSINGER, K. E. & FELDMAN, M. W. (1985). Selection in complex systems. VI. Equilibrium properties of two locus selection models with partial selfing. *Theoretical Population Biology* **28**, 117–132.

LESSA, E. & APPLEBAUM, G. (1993). Screening techniques for detecting allelic variation in DNA sequences. *Molecular Ecology* **2**, 119–129.

LYMBERY, A. J. (1995). Genetic diversity, genetic differentiation and speciation in the genus *Echinococcus* Rudolphi 1801. In *Echinococcus and Hydatid Disease* (ed. Thompson, R. C. A. & Lymbery, A. J.), pp. 51–88. CAB International, Wallingford.

LYMBERY, A. J., CONSTANTINE, C. C. & THOMPSON, R. C. A. (1997). Self-fertilization without genomic or population structuring in a parasitic tapeworm. *Evolution* **51**, 289–294.

LYMBERY, A. J. & THOMPSON, R. C. A. (1988). Electrophoretic analysis of genetic variation in *Echinococcus granulosus* from domestic hosts in Australia. *International Journal for Parasitology* **18**, 803–811.

LYMBERY, A. J. & THOMPSON, R. C. A. (1996). Species of *Echinococcus*: pattern and process. *Parasitology Today* **12**, 486–491.

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.

MCMANUS, D. P. & SIMPSON, A. J. G. (1985). Identification of the *Echinococcus* (hydatid disease) organisms using cloned DNA markers. *Molecular and Biochemical Parasitology* **17**, 171–178.

MICHALAKIS, Y. & EXCOFFIER, L. (1996). A genetic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. *Genetics* **142**, 1061–1064.

NEI, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.

NOLLEN, P. M. (1983). Patterns of sexual reproduction among parasitic platyhelminths. *Parasitology* **86**, 99–120.

RAUSCH, R. L. (1985). Parasitology: retrospect and prospect. *Journal of Parasitology* **71**, 139–151.

RAUSCH, R. L. (1995). Genetic diversity, genetic differentiation and speciation in the genus *Echinococcus*. In *Echinococcus and Hydatid Disease* (ed. Thompson, R. C. A. & Lymbery, A. J.), pp. 89–134. CAB International, Wallingford.

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

RODRIGUES, J. J. S., FERREIRA, H. B. & ZAHA, A. (1993). Molecular cloning and characterization of an *Echinococcus granulosus* DNA encoding malate dehydrogenase. *Molecular and Biochemical Parasitology* **60**, 157–160.

RODRIGUES, J. J. S., FERREIRA, H. B., FARIAS, S. E. & ZAHA, A. (1997). A protein with another calcium-binding domain associated with calcareous corpuscles in *Echinococcus granulosus*. *Biochemical and Biophysical Research Communications* **237**, 451–456.

SCHANTZ, P. M., CHAI, J., CRAIG, P. S., ECKERT, J., JENKINS, D. J., MACPHERSON, C. N. L. & THAKUR, A. (1995). Epidemiology and control of hydatid disease. In *Echinococcus and Hydatid Disease* (ed. Thompson, R. C. A. & Lymbery, A. J.), pp. 233–332). CAB International, Wallingford.

SCHNEIDER, S., KUEFFER, J. M., ROESSLI, D. & EXCOFFIER, L. (1997). *Arlequin*. Version 1.0. University of Geneva, Switzerland.

SHEFFIELD, V. C., BECK, J. S., KWITEK, A. E., SANDSTROM, D. W. & STONE, E. M. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* **16**, 325–332.

SHEPHERD, J. C., AITKEN, A. & MCMANUS, D. P. (1991). A protein secreted *in vivo* by *Echinococcus granulosus* inhibits elastase activity and neutrophil chemotaxis. *Molecular and Biochemical Parasitology* **44**, 81–90.

SILES-LUCAS, M., BENITO, M. C. & CUESTA-BANDERA, C. (1996). *Echinococcus granulosus*: genomic and isoenzymatic study of Spanish strains isolated from different intermediate hosts. *Veterinary Parasitology* **63**, 273–282.

SILVA, C. M. D., FERREIRA, H. B., PICON, M., GORFINKIEL, N., EHRLICH, R. & ZAHA, A. (1993). Molecular cloning and characterization of actin genes from *Echinococcus granulosus*. *Molecular and Biochemical Parasitology* **60**, 209–220.

SLATKIN, M. & EXCOFFIER, L. (1996). Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity* **76**, 377–383.

SMYTH, J. D. & SMYTH, M. M. (1964). Natural and experimental hosts of *Echinococcus granulosus* and *E. multilocularis*, with comments on the genetics of speciation in the genus *Echinococcus*. *Parasitology* **54**, 493–514.

THOMPSON, R. C. A. (1995). Biology and systematics of *Echinococcus*. In *Echinococcus and Hydatid Disease* (ed. Thompson, R. C. A. & Lymbery, A. J.), pp. 1–50. CAB International, Wallingford.

THOMPSON, R. C. A. & LYMBERY, A. J. (1966). Genetic variability in parasites and host–parasite interactions. *Parasitology* **112**, S7–S22.

THOMPSON, R. C. A., LYMBERY, A. J. & CONSTANTINE, C. C. (1995). Variation in *Echinococcus*: towards a taxonomic revision of the genus. *Advances in Parasitology* **35**, 146–176.

VISPO, M. & EHRLICH, R. (1994). Estudio de la expresion del gen con homeobox EgHbx2 del parasito *Echinococcus granulosus*. *II. Jornadas de Investigacion del Grupo Montevideo*, p. 121.

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/S0031182098003485 Published online by Cambridge University Press