

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE

Isolamento e caracterização de microssatélites do genoma de *Echinococcus granulosus* (Cestoda, Taeniidae).

Marlise Ladvocat Bartholomei-Santos

Tese apresentada ao Programa de Pós-graduação em Ciências Biológicas: Bioquímica como requisito parcial para obtenção do grau de Doutor em Bioquímica

Porto Alegre

2003

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE

Isolamento e caracterização de microssatélites do genoma de *Echinococcus granulosus* (Cestoda, Taeniidae).

Marlise Ladvocat Bartholomei-Santos

Orientador: Dr. Arnaldo Zaha

Tese apresentada ao Programa de Pós-graduação em Ciências Biológicas: Bioquímica como requisito parcial para obtenção do grau de Doutor em Bioquímica

Porto Alegre

2003

Aos amores da minha vida:

Camilla e Sandro

AGRADECIMENTOS

A Deus pelo dom da vida.

Ao meu orientador, Prof. Arnaldo Zaha, pelo constante apoio, pela confiança depositada e por seu exemplo de vida acadêmica.

Ao meu esposo Sandro, por todo o incentivo, apoio e companheirismo que sempre me dedicou e também por ter sido pai e mãe de nossa Camilla de segundas a sextas-feiras durante 4 anos. E à minha filha Camilla, que tão pacientemente esperou pelo “dia em que Mamãe não vai mais precisar viajar para Porto Alegre”.

Aos meus pais Antônio e Magali, às minhas irmãs Maudy e Márcia e à minha tia Maudy, que apesar da distância de 1200 km nos separando, sempre estiveram ao meu lado.

Aos meus queridos amigos do laboratório de Biologia Molecular de Cestódeos: Larissa Schemmes Heinzelmann, Claudia Paiva Nunes, Anelise Volkweiss, Gustavo Chemale, Cristiano Bizarro, Melissa Markoski, Veridiana Virgínio, Leonardo Alves, Ana Cristina Arend, André Dannenhauer, Jeferson Badaraco, Karina Monteiro e Aline Zandonai pelos inúmeros auxílios e por todos os ótimos momentos que passamos juntos.

À Profa. Karen Luísa Haag pelos isolados de *Echinococcus granulosus* utilizados na análise populacional, por todas as sugestões e principalmente pela amizade.

Ao Prof. Henrique Bunselmayer Ferreira por todas as sugestões, discussões e sobretudo, pela amizade.

Às queridas amigas e ex-colegas do laboratório Elizabeth Cortez Herrera, Marilise Rott, Etel Gimba, Mônica Blauth e Laíze Tomazi pela ótima convivência que tivemos e pela amizade que ficou.

À Profa. Riva de Paula Oliveira por ter me ensinado a técnica de isolamento de microssatélites.

À Deise Porto Potrich pelo auxílio com o seqüenciamento automático e pela amizade.

Ao Prof. Elgion Loreto e à Prof. Lenira Sepel pela divisão de carga horária na UFSM para que eu pudesse permanecer mais tempo em Porto Alegre.

À Silvia Regina Centeno por estar sempre pronta a ajudar, por seu eterno bom-humor e sua amizade.

Ao CNPq, CAPES, FAPERGS e RTPD (Research and Training in Parasitic Diseases) Network (SIDA/SAREC, Suécia) pelo suporte financeiro.

Ao Programa de Pós-graduação em Ciências Biológicas: Bioquímica, da Universidade Federal do Rio Grande do Sul, pela oportunidade de realização deste curso.

Ao Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul pelas condições oferecidas para a execução deste projeto.

A todos aqueles que de alguma forma contribuíram para a realização deste trabalho.

Muito obrigada!

ÍNDICE

RESUMO	1
ABSTRACT	2
1. INTRODUÇÃO	3
1.1. O nosso organismo de estudo: <i>Echinococcus granulosus</i>	3
1.2. Caracterização de linhagens em <i>E. granulosus</i>	6
1.3. Epidemiologia de <i>E. granulosus</i>	7
1.4. O que são microssatélites?	9
1.4.1. Biologia dos microssatélites	11
1.4.2. Função dos microssatélites no genoma	17
1.4.3. Distribuição e frequência de microssatélites nos genomas	19
1.5. Microssatélites em parasitas	20
1.6. Por que estudar microssatélites em <i>E. granulosus</i> ?	21
1.7. Objetivos	23
1.7.1. Objetivos gerais	23
1.7.2. Objetivos específicos	23
2. MATERIAL E MÉTODOS	24
2.1. Cistos hidáticos	24
2.2. Linhagem bacteriana, vetor de clonagem, enzimas, reagentes, oligonucleotídeos, antibióticos, meios de cultura e soluções utilizados	24
2.3. Populações de parasitas	25
2.4. Extração de DNA genômico de isolados de <i>E. granulosus</i>	25
2.5. <i>Southern blot</i>	26
2.5.1. Clivagem do DNA com endonucleases de restrição e transferência para membranas de náilon	26
2.5.2. Marcação radioativa das sondas de oligonucleotídeos e hibridização	26
2.5.3. Marcação radioativa da sonda snRNA U1 e hibridização	27
2.6. Isolamento dos microssatélites a partir do DNA genômico de <i>E. granulosus</i>	28
2.6.1. Isolamento de seqüências contendo repetições CA, GA, AAC e TA	28

2.6.2. Isolamento do microsatélite com repetições de pentanucleotídeos contido no gene snRNA U1	30
2.7. Clonagem dos fragmentos contendo microsatélites	31
2.8. Transformação	31
2.9. Seleção de colônias recombinantes	31
2.10. Seqüenciamento dos plasmídeos recombinantes	32
2.11. Elaboração de <i>primers</i> para as regiões flanqueadoras dos microsatélites CA e GA	33
2.12. Estabelecimento das condições de PCR para a amplificação de um locus de microsatélite CA e um locus GA	33
2.13. Eletroforese dos fragmentos de DNA em gel de poliacrilamida	35
2.14. Coloração do gel de poliacrilamida	35
2.15. Teste da estabilidade de um locus de microsatélite através de PCR de protoescólices individuais	35
2.16. Amplificação do locus Egmsca1 em isolados de diferentes procedências	36
2.17. Análise estatística	36
2.18. Comparação das seqüências contendo microsatélites isoladas com seqüências depositadas em bancos de dados	37
3. RESULTADOS	38
3.1. Verificação da abundância de seqüências repetitivas no genoma de <i>E. granulosus</i> utilizando sondas de oligonucleotídeos marcadas radioativamente	38
3.2. Verificação da freqüência do gene snRNA U1 no genoma de <i>E. granulosus</i>	38
3.3. Isolamento do microsatélite contido no gene snRNA U1	41
3.4. Amplificação do microsatélite presente no gene snRNA U1	43
3.5. Isolamento de seqüências contendo repetições CA, GA, AAC e TA	43
3.6. Estabelecimento das condições de amplificação para os locus Egmsca1 e Egmsga1	49
3.7. Análise de polimorfismo no locus Egmsca1	49

4. DISCUSSÃO	58
5. CONCLUSÕES	72
6. PERSPECTIVAS	73
7. REFERÊNCIAS BIBLIOGRÁFICAS	74
8. ANEXO – ARTIGO CIENTÍFICO	85

RESUMO

A presença de microssatélites no genoma de *Echinococcus granulosus* foi verificada utilizando-se oito oligonucleotídeos com repetições como sondas (GT₁₅, CT₁₅, AT₁₅, CG₁₅, CAT₁₀, CAA₁₀, CGG₁₀ e CATA₁₀). Experimentos de hibridização revelaram que as repetições GT, CAA, CATA e CT são as mais frequentes no genoma de *E. granulosus*. As sondas AT e GC não apresentaram sinais de hibridização.

Seis loci contendo repetições CA/GT, quatro loci contendo repetições GA/CT e oito loci contendo repetições AAC/GGT foram clonados e sequenciados. O locus Egmsca1 foi analisado em 73 isolados do Brasil e da Argentina cujas linhagens haviam sido previamente caracterizadas e em 27 isolados provenientes da Etiópia cujas linhagens ainda não foram identificadas.

Os isolados brasileiros da linhagem bovina e os isolados argentinos da linhagem do camelo apresentaram-se monomórficos e compartilharam o alelo (CA)₇. Isolados argentinos das linhagens da ovelha e da ovelha da Tasmânia compartilharam dois alelos [(CA)₈ e (CA)₁₀] com os isolados brasileiros da linhagem da ovelha. O alelo (CA)₁₁ foi encontrado somente em isolados brasileiros da linhagem da ovelha em uma baixa frequência. O alelo (CA)₉ ocorreu apenas em um isolado da Etiópia. As populações brasileira e argentina da linhagem da ovelha foram testadas em relação ao equilíbrio de Hardy-Weinberg e somente a primeira estava de acordo com as expectativas. Protoescólices isolados de um único cisto hidático não apresentaram polimorfismo, validando a utilização de protoescólices de um mesmo cisto, agrupados como um isolado, em estudos populacionais. Um microssatélite contendo repetições de pentanucleotídeos, contido no complexo gênico do snRNA U1, também foi isolado. Este estudo descreve pela primeira vez o isolamento e caracterização de microssatélites em *E. granulosus*.

ABSTRACT

The *Echinococcus granulosus* genome was searched for microsatellites using eight different repeated oligonucleotides as probes (GT₁₅, CT₁₅, AT₁₅, CG₁₅, CAT₁₀, CAA₁₀, CGG₁₀ and CATA₁₀). 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.

Six loci containing CA/GT repeats, four loci containing GA/CT repeats and eight loci containing AAC/TTG repeats were cloned and sequenced. The Egmsca1 locus was analyzed in 73 isolates from Brazil and Argentina whose strains were previously characterized, and in 27 isolates from Ethiopia whose strains were not identified yet. Brazilian isolates of 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 two 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 allele (CA)₉ occurred only in one isolate from Ethiopia. The Brazilian and the Argentinean populations of sheep strain 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 single hydatid cysts, validating the utilization of pooled protoscoleces from one cyst, grouped as an isolate, in population studies. We also isolated a pentanucleotide microsatellite within U1 snRNA gene complex. This work describes for the first time the isolation and characterization of microsatellites from *E. granulosus*.

1. INTRODUÇÃO

1.1. O NOSSO ORGANISMO DE ESTUDO: *Echinococcus granulosus*

E. granulosus é um pequeno verme achatado dorso-ventralmente, pertencente à família Taeniidae da Classe Cestoda. Apresenta um ciclo de vida indireto (Fig. 1.1), no qual o hospedeiro definitivo é um canídeo e o hospedeiro intermediário é, em geral, um ungulado, podendo ser também primatas, marsupiais e acidentalmente, o homem.

O verme adulto é hermafrodita e vive no intestino delgado de cães. Apresenta na sua região anterior um órgão de fixação especializado, o escólex, que possui duas fileiras de ganchos e quatro ventosas musculares. O corpo, ou estróbilo, é segmentado e consiste de unidades reprodutivas, as proglótides, que raramente ultrapassam o número de cinco (Thompson, 1995).

A fase larval, também denominada metacestódeo ou cisto hidático, desenvolve-se nas vísceras do hospedeiro intermediário e reproduz-se assexuadamente.

Os ovos liberados pelo verme adulto nas fezes dos cães infectados são capazes de permanecer viáveis no meio ambiente por períodos longos de tempo. Após serem ingeridos pelo hospedeiro intermediário, eclodem no estômago destes, liberando a oncosfera, a qual é ativada no intestino delgado, penetra na mucosa intestinal e atinge a circulação sanguínea, migrando para as vísceras, em geral, fígado e pulmões.

Nas vísceras inicia-se o desenvolvimento pós-oncosférico levando à formação do cisto hidático, no qual a oncosfera sofre uma série de eventos de reorganização durante os primeiros 14 dias, envolvendo proliferação celular, degeneração dos ganchos oncosféricos, atrofia muscular, vesicularização e formação da cavidade central, além da formação das camadas germinativa e laminar.

O cisto hidático de *E. granulosus* completamente desenvolvido é tipicamente unilocular e repleto de um fluido, o líquido hidático. O cisto apresenta uma camada interna germinativa, envolvida por uma camada laminar acelular, resistente e elástica. A camada externa (adventícia) é fibrosa e produzida pelo hospedeiro.

A camada germinativa prolifera assexualmente, formando cápsulas prolíferas, dentro das quais formam-se os protoescólices. Dentro de um único cisto podem se formar centenas a milhares de protoescólices, a forma infectiva para o hospedeiro definitivo. Estes permanecem

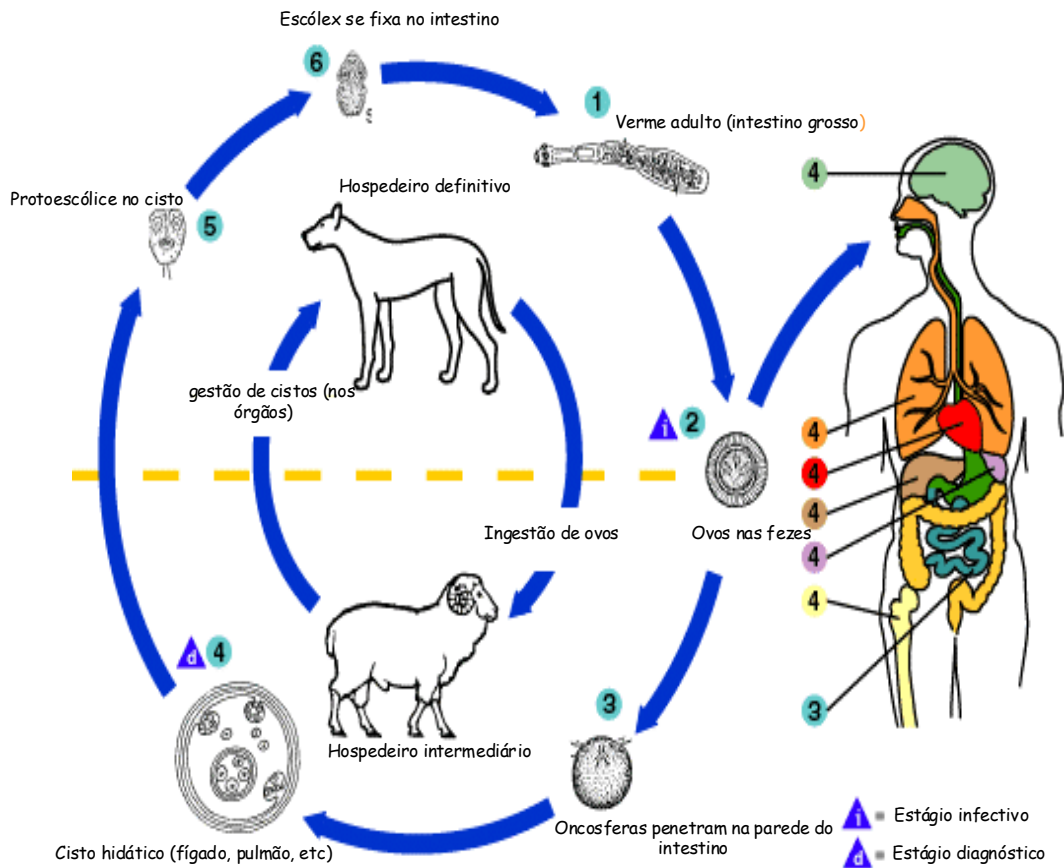


Figura 1.1. Ciclo vital do *Echinococcus granulosus*. O adulto do *E. granulosus* (2 a 11 mm de comprimento) reside no intestino delgado do hospedeiro definitivo (cães domésticos e outros canídeos). As proglótides grávidas liberam ovos que são eliminados nas fezes. Após ingestão por um hospedeiro intermediário adequado (ovinos, suínos, camelos, cavalos, etc.), os ovos eclodem no intestino delgado e liberam a oncosfera que penetra na parede do intestino e migra através do sistema circulatório para vários órgãos, especialmente fígado e pulmões. Nesses órgãos se desenvolve em um cisto que aumenta gradualmente, produzindo protoescólecies e hidátides filhas que preenchem o interior do cisto. O hospedeiro definitivo se infecta por ingestão de cistos existentes nos órgãos dos hospedeiros intermediários infectados. Após a ingestão, os protoescólecies evaginam, fixam-se à mucosa intestinal e transformam-se em vermes adultos. Humanos infectam-se pela ingestão de ovos, com resultante liberação de oncosferas no intestino delgado e desenvolvimento de cistos em diversos órgãos. Modificado a partir de figura da biblioteca de imagens do *Center for Disease Control & Prevention* (<http://www.dpd.cdc.gov>).

com sua região apical (ventosas, rostelo e ganchos) invaginados na região basal do tegumento, de forma a proteger o escólex até que seja estimulado a evaginar.

O hospedeiro definitivo é infectado ao ingerir vísceras com cistos hidáticos contendo protoescóleces viáveis. As alterações ambientais (temperatura, pH, pressão osmótica) estimulam a evaginação dos protoescóleces, que se tornam muito ativos e se estabelecem no intestino delgado do hospedeiro, onde se desenvolvem em vermes adultos. A parasitose é assintomática nos hospedeiros definitivos.

A doença causada pela forma larval do parasita é denominada hidatidose e sua gravidade provém do crescimento do cisto hidático, o qual comprime os órgãos adjacentes a sua localização. O rompimento de um cisto hidático no hospedeiro intermediário pode ocasionar a formação de múltiplos cistos secundários ou provocar um choque anafilático. Seres humanos se contaminam ao ingerirem os ovos liberados nas fezes dos cães, pelo contato íntimo com cães parasitados ou ingestão dos ovos presentes em verduras e água contaminadas. O tratamento de pessoas infectadas pode ser cirúrgico, mas a cirurgia é de alto risco. Medicamentos como o albendazol e praziquantel também têm sido utilizados com relativo sucesso (Anadol *et al.*, 2001; Saimot, 2001).

Na literatura é possível encontrar sinônimos para a hidatidose causada por *E. granulosus*: equinococose cística, equinococose devida à *E. granulosus* e doença hidática. Às vezes, o termo equinococose é atribuído a infecções causadas tanto pelo verme adulto no hospedeiro definitivo, quanto pelo metacestódeo no hospedeiro intermediário e o termo hidatidose fica reservado à infecção pelo metacestódeo apenas (Lightowers & Gottstein, 1995). Kassai (2001) propõe a utilização do termo equinococose granulosus em contrapartida à recomendação do *International Nomenclature of Diseases*, de se utilizar o termo equinococose devido a *E. granulosus*. Neste texto, a fim de facilitar a identificação do estágio infectante (adulto ou larva) e do hospedeiro (definitivo ou intermediário), o termo equinococose será adotado para se referir à infecção causada pelo verme adulto no hospedeiro definitivo e hidatidose para se referir à infecção causada pelo estágio larval (metacestódeo) no hospedeiro intermediário.

O gênero *Echinococcus* é composto por quatro espécies: *E. granulosus*, *E. multilocularis*, *E. vogeli* e *E. oligarthrus*, que diferem em relação à distribuição geográfica, hospedeiros definitivos e intermediários, além de morfologia e localização preferencial do metacestódeo (Thompson, 1995).

1.2. CARACTERIZAÇÃO DE LINHAGENS EM *E. granulosus*

Sete linhagens são descritas para *E. granulosus*: a dos cervídeos, da ovelha, da ovelha da Tasmânia, do cavalo, do boi, do porco e do camelo. Duas outras possíveis linhagens, do búfalo e do leão, ainda não foram geneticamente caracterizadas e são separadas com base em características morfológicas, biológicas e epidemiológicas (Thompson, 1995). Deve-se notar que o nome da linhagem está relacionado ao hospedeiro intermediário no qual uma maior proporção de cistos com determinada variante genética foi identificada, não significando que uma linhagem infecte somente o hospedeiro intermediário que lhe deu o nome. Por exemplo, a linhagem da ovelha infecta, além de ovelhas, boi, cabras, porcos, cavalos e é a principal causa de infecções em humanos (Schantz *et al.*, 1995).

Bowles *et al.* (1995) propõem a seguinte definição para linhagens de *Echinococcus*: linhagem é um grupo de indivíduos que difere significativamente de outros grupos da mesma espécie em frequências gênicas ou seqüências de ácidos nucleicos e em uma ou mais características de significância para o controle e epidemiologia da hidatidose.

As sete linhagens descritas para *E. granulosus* se enquadram na definição acima, visto que apresentam diferenças na distribuição geográfica, no grau de infectividade para diferentes hospedeiros intermediários (Schantz *et al.*, 1995), nos sistemas metabólicos (McManus & Bryant, 1995) e nas seqüências dos genes mitocondriais COI (citocromo *c* oxidase I) e ND1 (NADH desidrogenase 1) (Bowles *et al.*, 1992, 1995).

A taxonomia de *Echinococcus* é um assunto controvertido há décadas. Thompson & McManus (2002) argumentam que as diferenças na morfologia, desenvolvimento e seqüências de DNA entre as linhagens da ovelha, do cavalo e do boi de *E. granulosus* suportam o reconhecimento do *status* de espécie para estas linhagens, as quais deveriam ser designadas por *E. granulosus*, *E. equinus* e *E. ortleppi*, respectivamente, conforme sugerido anteriormente por outros autores. Além disso, argumentam que possivelmente as linhagens do camelo, do porco e do cervídeo também possam ser elevadas à categoria de espécie, porém ainda é necessária uma melhor caracterização molecular de isolados dessas linhagens procedentes de diferentes áreas geográficas antes da alteração de sua nomenclatura.

Dois modelos contrastantes, mas não mutuamente exclusivos, foram propostos para explicar os processos que geram variação dentro do gênero *Echinococcus*. Smyth & Smyth (1964) sugeriram que as mutações ocorridas são sempre detectáveis porque são transmitidas

tanto na forma homozigota quanto heterozigota por auto-fertilização obrigatória em adultos; se as mutações não forem selecionadas negativamente no hospedeiro intermediário, podem ser amplificadas por reprodução assexuada para gerar uma nova linhagem.

Rausch (1985) propôs que os adultos se reproduzem normalmente por reprodução cruzada e que as linhagens se originam por diferenciação genética gradual das populações isoladas através de suas associações com diferentes espécies de hospedeiros domésticos.

1.3. EPIDEMIOLOGIA DE *E. granulosus*

E. granulosus é uma espécie cosmopolita, estando presente em todos os continentes, excluindo-se a Antártida (Fig. 1.2). A mais alta prevalência do parasita é encontrada em partes da Eurásia, África, Austrália e América do Sul. Nas zonas endêmicas a prevalência do parasita varia de esporádica a alta, mas somente poucos países podem ser considerados livres de *E. granulosus* (Eckert *et al.*, 2000).

Os principais fatores que influenciam a persistência, re-emergência e o espalhamento da infecção por *E. granulosus*, de acordo com Eckert *et al.* (2000), são os seguintes: 1) presença de grande número de cachorros com alta prevalência de equinococose; 2) fácil acesso dos cachorros aos órgãos de animais de fazenda infectados com cistos de *E. granulosus*, especialmente no interior dos países; 3) tratamento anti-helmíntico insuficiente dos cães; 4) número restrito ou falta de pequenos canis municipais; 5) abate e inspeção ineficientes da carne; 6) ineficiência ou ausência de oportunidade de destruição de vísceras infectadas; 7) abate ilegal ou caseiro não-inspecionados de animais de fazenda e 8) ausência de educação sanitária adequada. Além disso, a conscientização pública do problema e a disponibilidade de recursos econômicos para programas de controle e prevenção efetivos também influenciam a persistência e emergência/re-emergência da hidatidose.

Os programas de controle da hidatidose são baseados predominantemente no controle da população de cachorros, tratamento regular dos cães de estimação com praziquantel para eliminação do parasita e educação sanitária (Schantz *et al.*, 1995; Eckert *et al.*, 2000).

Na América do Sul mais de 2000 pessoas sofrem cirurgia para hidatidose anualmente, incluindo 464 na Argentina (1,42/100.000 habitantes), 367 no Uruguai (12,4/100.000), 573 no Chile (3,4/100.000) e 244 no Peru (2,4/100.000) (Eckert *et al.*, 2000).

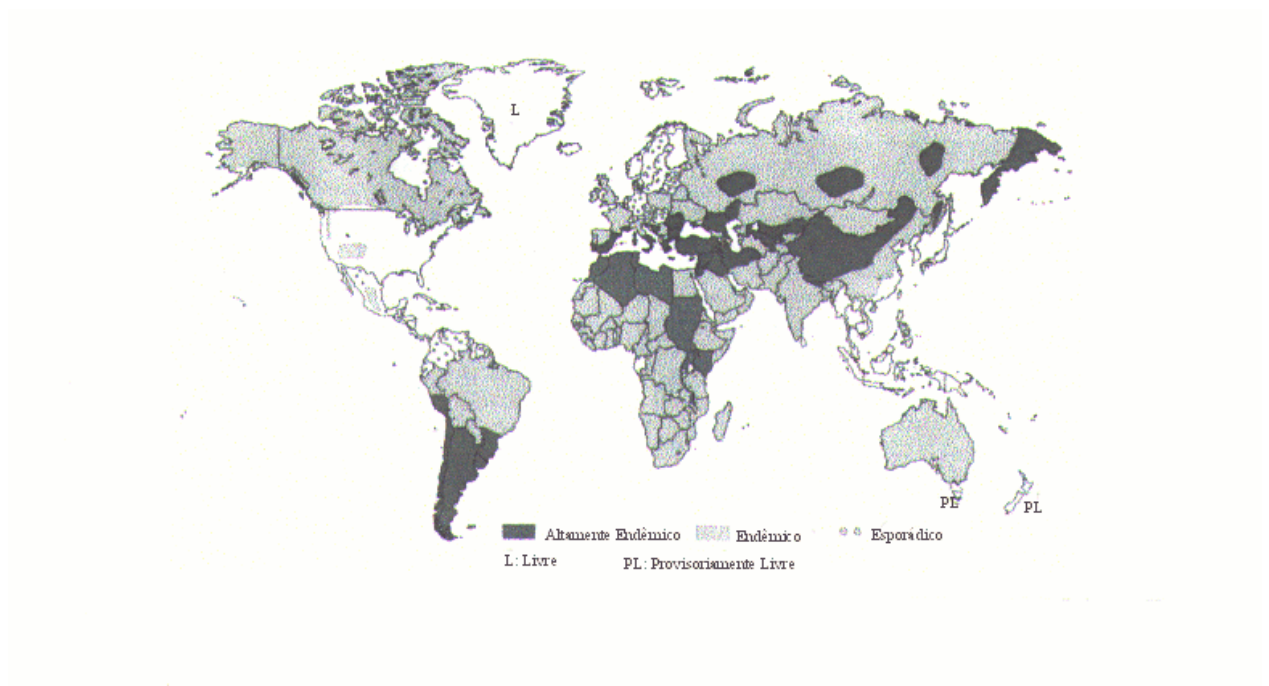


Figura 1.2. Distribuição geográfica aproximada de *Echinococcus granulosus*. Modificado de Eckert *et al.*, 2000.

No Brasil a hidatidose é um problema sério de saúde pública no estado do Rio Grande do Sul, onde a doença é considerada endêmica entre animais domésticos e também, seres humanos. A maior incidência de casos ocorre nas fronteiras com o Uruguai e a Argentina, locais de intensa criação de ovelhas. Estudos realizados em 1990 pela Divisão de Zoonoses e Vetores da Secretaria de Saúde e Ambiente do Rio Grande do Sul identificaram a hidatidose em 31,25% dos bovinos e 23,15% dos ovinos abatidos na década anterior. Dados do Serviço de Inspeção Federal do Ministério da Agricultura obtidos em abatedouros do Rio Grande do Sul entre 1990 e 1993 mostraram a ocorrência de hidatidose em 25% dos bovinos abatidos, sendo que nas regiões de fronteira com Uruguai e Argentina a ocorrência variou de 30 a 40%. No mesmo período a ocorrência de hidatidose em ovinos foi 6%. Em humanos a incidência cirúrgica anual da doença entre 1982 e 1987 variou de 2,4 a 4,9/100.000 habitantes (Schantz *et al.*, 1995).

1.4. O QUE SÃO MICROSSATÉLITES?

Microssatélite é o termo atualmente mais utilizado para descrever repetições em tandem de seqüências curtas de DNA, com no máximo 6 bases, que têm sido encontradas em todos os organismos estudados até o momento (Hancock, 1999).

A principal característica dos loci de microssatélites é a variabilidade encontrada no número de repetições da seqüência-motivo. Os primeiros trabalhos a detectar esta variabilidade datam de 1989 (Litt & Luty, 1989; Weber & May, 1989; Tautz, 1989).

Os microssatélites tornaram-se populares na década passada porque se mostraram ferramentas moleculares extremamente versáteis. As aplicações relacionadas com os microssatélites variam desde sua utilização como marcadores moleculares muito acessíveis para segmentos cromossômicos, passando pela identificação de indivíduos, até o rastreamento da história biológica de populações (Chambers & MacAvoy, 2000).

À primeira vista, a popularidade dos microssatélites pode parecer surpreendente, visto que não são seqüências fáceis de serem isoladas e trabalhadas em laboratório. Além disso, a tecnologia de análise de um microssatélite em particular nem sempre é transferível de uma espécie para outra. O fator chave que leva a sua ampla utilização reside no poder que os microssatélites apresentam para resolver problemas biológicos (Chambers & MacAvoy, 2000).

Na literatura encontram-se diferentes termos usados como sinônimos de microssatélites, tais como STR (Simple Tandem Repeats), SSR (Simple Sequence Repeat), SSLP (Simple Sequence Length Polymorphism) e VNTR (Variable Number of Tandem Repeats). O termo microssatélite foi primeiramente utilizado por seus descobridores (Litt & Luty, 1989; Weber & May, 1989; Tautz, 1989), em analogia ao termo minissatélite, empregado para definir repetições em tandem de seqüências com mais de 10 pb, sendo atualmente a nomenclatura mais utilizada.

Existem controvérsias a respeito da definição de um microssatélite, em relação ao número de nucleotídeos da seqüência repetida, por exemplo, 1 a 6 pb (Goldstein & Pollock, 1997), 1 a 5 pb (Schlötterer, 1999), 2 a 8 pb (Armour *et al.*, 1999). Chambers & MacAvoy (2000) sugerem uma definição de 2 a 6 pb para o tamanho da unidade repetitiva, seguindo a descrição dos autores originais, além de um tamanho mínimo de 8 pb para que um arranjo seja considerado um microssatélite. Assim, um microssatélite com repetição de dinucleotídeos precisa de no mínimo 4 repetições para ser considerado como tal. Arranjos de repetições com menos de 8 nucleotídeos devem ser considerados como seqüências simples crípticas.

Ainda segundo estes autores, o intervalo de 4 nucleotídeos entre as definições de tamanho para mini e microssatélites deve ser mantido, visto que estes dois tipos de seqüências repetitivas evoluem por mecanismos diferentes e pouco se sabe a respeito da evolução e comportamento genético de elementos repetitivos de DNA com unidades de repetição entre 7 e 10 nucleotídeos.

As seguintes definições devem ser utilizadas para as diferentes classes de seqüências repetitivas em tandem, de acordo com Chambers & MacAvoy (2000):

Satélites: segmentos altamente repetitivos de 100 nucleotídeos ou mais formando arranjos relativamente uniformes com 10^3 - 10^7 nucleotídeos de comprimento.

Minissatélites: segmentos moderadamente repetitivos de 10 a 100 nucleotídeos formando arranjos relativamente uniformes de 10^2 - 10^5 pb de comprimento.

Microssatélites: segmentos curtos de 2 a 6 nucleotídeos repetidos em arranjos relativamente uniformes de até $\sim 10^2$ pb de comprimento.

Arranjos de mononucleotídeos: trechos uniformes de um único nucleotídeo de qualquer tamanho.

Os microssatélites são classificados quanto à presença ou ausência de interrupção na unidade repetitiva e ainda, quanto à presença de mais de um tipo de unidade repetitiva (Goldstein & Schlötterer, 1999).

Microssatélites perfeitos consistem de um único motivo de repetição, não sendo interrompidos ao longo da seqüência repetida por nenhuma base que altere o padrão da repetição (por exemplo, CACACACACACA).

Microssatélites imperfeitos apresentam uma ou mais repetições que contém uma base que altera o padrão do motivo repetitivo (por exemplo, GAGAGAGACAGAGA).

Microssatélites interrompidos apresentam a inserção de um pequeno número de bases que não se encaixam na estrutura da repetição (por exemplo, CACACACACAGGGCACACA).

Microssatélites compostos consistem de dois ou mais microssatélites adjacentes com motivos repetitivos diferentes (por exemplo, CACACACAGATGATGATGATGAT).

Em teoria, existem 16, 64, 256, 1024 e 4096 motivos repetitivos possíveis para microssatélites de di, tri, tetra, penta e hexanucleotídeos, respectivamente. Entretanto, como os microssatélites são repetidos em tandem, alguns motivos equivalem a outros. Duas regras podem ser utilizadas para identificar a equivalência de motivos. Um motivo A é considerado equivalente ao motivo B se: 1) o motivo A é inversamente complementar ao motivo B; 2) o motivo A difere do motivo B ou da seqüência inversamente complementar do motivo B na matriz de leitura. Por exemplo, $(GAAA)_n$ é equivalente a $(AGAA)_n$ ou $(AAGA)_n$, a $(AAAG)_n$ ou $(TTTC)_n$, a $(TTCT)_n$ ou $(TCTT)_n$ ou a $(CTTT)_n$. Assim, devido à equivalência entre motivos, é possível considerar a existência de 4, 10, 33, 102 e 350 possíveis motivos de repetição para microssatélites de di, tri, tetra, penta e hexanucleotídeos, respectivamente (Jin *et al.*, 1994).

1.4.1. BIOLOGIA DOS MICROSSATÉLITES

Os microssatélites apresentam um “ciclo de vida”, no qual eles nascem, crescem e morrem (Amos, 1999). Messier *et al.* (1996) relatam o nascimento de um microssatélite a partir da comparação de seqüências ortólogas em 13 espécies de primatas (Fig. 1.3). Uma única substituição nucleotídica nos hominóides, de G para A, ocorreu na linha do ancestral comum entre os macacos africanos grandes (gorila, bonobo e chimpanzé) e os humanos. Esta

substituição gerou uma repetição de tetranucleotídeos, (ATGT)₂, a qual se expande para (ATGT)₄ nos macacos africanos e para (ATGT)₅ nos humanos. Outra substituição nucleotídica ocorreu nos macacos do novo mundo na linha levando ao macaco-coruja, criando uma repetição de dinucleotídeos (GT)₆ ou (TG)₇, que se expandiu em uma unidade nesta espécie.



Figura 1.3. Pequena porção de uma alinhamento de seqüências do pseudogene da η-globina de primatas, mostrando o nascimento de microssatélites com repetições de tetranucleotídeos nos hominóides e de dinucleotídeos no macaco-coruja, a partir de mutações de ponto. Modificado de Messier *et al.*(1996).

A instabilidade dos locos de microssatélites torna-os extremamente úteis para estudos evolutivos e genéticos. Esta instabilidade é predominantemente devida à expansão ou diminuição do número de cópias da repetição do microssatélite.

Dois modelos foram propostos para explicar o mecanismo de instabilidade dos microssatélites. Um dos modelos sugere que a instabilidade é causada por uma taxa elevada de *crossing-over* desigual entre cromossomos homólogos imperfeitamente alinhados ao longo das repetições de microssatélites. O modelo alternativo (SSM – *Slip Strand Mismatching*) propõe que a instabilidade é causada por uma taxa elevada de deslizamento das fitas de DNA

ao longo das repetições dos microssatélites durante a replicação (Eisen, 1999). O segundo modelo tem sido mais aceito e amplamente citado na literatura, embora os dois modelos não sejam mutuamente exclusivos.

As evidências que dão suporte ao modelo SSM e contradizem o modelo de *crossing-over* desigual, segundo Eisen (1999), são as seguintes:

1) A estabilidade dos microssatélites não é afetada por defeitos em genes com papéis principais na recombinação, como *recA* em *Escherichia coli* e *rad52* em *Saccharomyces cerevisiae*.

2) Em humanos, alterações no número de repetições em microssatélites podem ser geradas sem a troca de marcadores genéticos flanqueadores e assim, provavelmente sem recombinação.

3) Em *S. cerevisiae*, a estabilidade do microssatélite é similar em células na meiose e na mitose. Como a recombinação ocorre mais frequentemente na meiose do que na mitose, os microssatélites deveriam ser mais instáveis durante a meiose.

4) A estabilidade dos microssatélites é reduzida por defeitos em genes envolvidos nas vias de correção de erro durante a replicação. Esta observação é consistente com o modelo SSM, o qual requer a replicação do DNA para ocorrer. Além disso, experimentos bioquímicos e genéticos mostram que essas vias de correção de erro podem reconhecer e reparar os tipos de alças que seriam criadas por SSM.

5) A orientação do microssatélite em relação às fitas de síntese contínua e descontínua influencia a estabilidade. Esta observação também é consistente com o modelo SSM, já que as fitas de síntese contínua e descontínua têm mecanismos de replicação um pouco diferenciados.

O deslizamento na replicação refere-se ao alinhamento fora de fase das duas fitas de DNA após a dissociação, no momento em que a DNA polimerase atinge a região repetitiva. Se a unidade repetitiva mais a 3' da fita nascente renatura com uma unidade repetitiva complementar a 3' da repetição correta na fita molde, será formada uma alça na fita nascente e a nova fita apresentará um tamanho maior que a fita molde após o término do alongamento. Inversamente, se o alinhamento incorreto ocorre à 5' na fita molde, a nova fita será menor que a seqüência molde (Fig. 1.4) (Ellegren, 2000a).

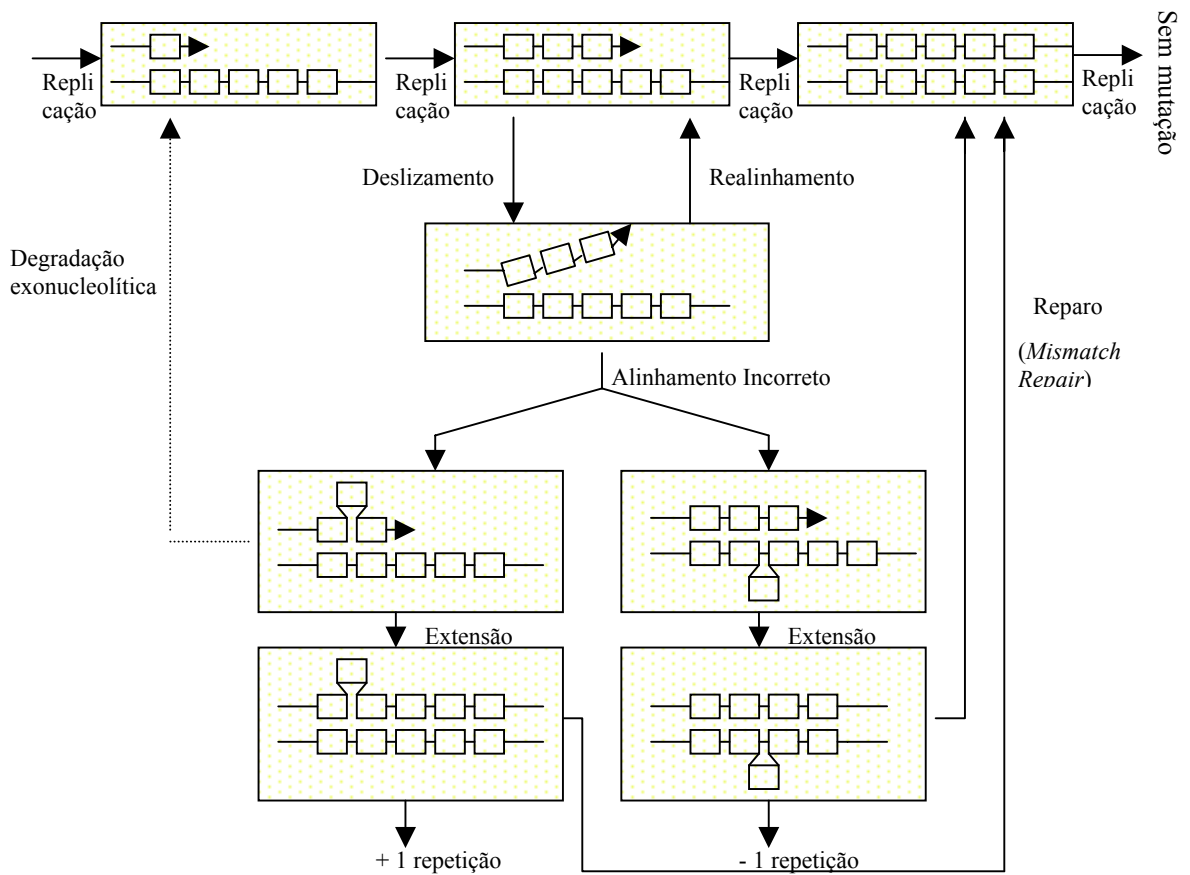


Figura 1.4. Modelo de mutação dos microssatélites a partir de deslizamento na replicação (SSM). Representações do DNA dupla fita contendo um repetição de microssatélites são mostrados em diferentes estágios da replicação e do processo mutacional. As fitas de DNA são representadas por linhas finas, as repetições de microssatélites por pequenos retângulos e a replicação em andamento por pequenas setas. Setas apontando para baixo indicam os passos que levam à mutação e setas apontando para cima indicam passos que impedem a ocorrência das mutações. Modificado de Eisen (1999).

Nem todos os erros causados por deslizamento durante a replicação tornam-se mutações, pois alguns são corrigidos por mecanismos de reparo de DNA. Duas vias de correção de erro foram demonstradas como sendo importantes no reparo destas seqüências, a “prova de leitura” exonucleolítica feita pela DNA polimerase e o reparo pós-replicação (*mismatch repair*) (Eisen, 1999).

A instabilidade dos microssatélites aumenta em função do aumento no número de repetições. Microssatélites de dinucleotídeos CA e de trinucleotídeos ricos em GC são os mais instáveis e formam quatro grupos (Ramel, 1997):

- 1) Alelos com pequeno número de repetições: microssatélite estável.
- 2) Alelos com repetições de tamanho médio: microssatélite polimórfico, mas estável entre gerações.
- 3) Alelos longos: a instabilidade é aumentada em várias ordens de magnitude, constituindo alelos pré-mutacionais, os quais não são estáveis entre gerações.
- 4) Alelos longos e extremamente instáveis, os quais também exibem instabilidade mitótica.

Embora o mecanismo de deslizamento durante a replicação permita tanto o aumento, quanto o decréscimo no número de repetições, vários estudos encontraram evidências para uma tendência ao aumento no número de unidades repetitivas (Amos *et al.*, 1996; Primmer *et al.*, 1996; Ellegren, 2000a).

A observação de que microssatélites raramente apresentam comprimentos superiores a poucas dezenas de unidades repetitivas levantou uma discussão a respeito de quais mecanismos impediriam o crescimento infinito destas seqüências, sendo que alguns estudos indicam que a perda de repetições pode ser mais comum ou envolver deleções maiores entre alelos longos, do que entre alelos curtos (Ellegren, 2000b).

Além disso, as mutações de ponto em microssatélites perfeitos contribuem para impedir o crescimento infinito, pois geram alelos que contêm dois trechos menores de repetições separados por um nucleotídeo, ou seja, geram um microssatélite imperfeito. Locus de microssatélites imperfeitos são menos variáveis que os perfeitos, devido à uma redução na taxa de mutação por deslizamento nos alelos imperfeitos ou interrompidos (Jin *et al.*, 1996).

As distribuições dos tamanhos das repetições de um locus de microssatélite no equilíbrio podem ser o resultado de um balanço entre eventos de deslizamento e mutação de ponto (Kruglyak *et al.*, 1998; Kruglyak *et al.*, 2000; Santibáñez-Koref *et al.*, 2001). Neste

caso, duas forças mutacionais opostas operam nas seqüências de microssatélites: o deslizamento durante a replicação, o qual favorece o crescimento, e as mutações de ponto, as quais dividem um arranjo longo de repetições em duas ou mais seqüências menores (Ellegren, 2002).

De acordo com esta idéia, a evolução do microssatélite é um processo no qual o deslizamento na replicação arbitrariamente dirige alguns alelos para um crescimento em tamanho até que uma mutação de ponto corte o arranjo repetitivo perfeito em dois trechos de repetição muito menores. Neste ponto, a velocidade de expansão seria muito diminuída. Pode-se dizer que este é o primeiro passo para a “morte” do microssatélite, já que mutações adicionais podem eventualmente converter o arranjo repetitivo em uma mistura de seqüências únicas e trechos repetidos curtos (Taylor *et al.*, 1999; Ellegren, 2000b).

Entretanto, embora existam estudos mostrando evidências para a morte dos microssatélites, também já foram encontradas evidências para a “purificação” de microssatélites interrompidos por mutações pontuais, provavelmente por deslizamento durante a replicação, de tal forma que a fita molde é deslocada para trás antes que a interrupção seja sintetizada na nova fita, que assim perde a interrupção (Harr *et al.*, 2000).

Paradoxalmente, seqüências simples crípticas podem se originar a partir da acumulação de mutações de ponto em microssatélites, mas os microssatélites também podem se originar a partir de mutações de ponto em seqüências simples crípticas. Esses dois processos não são necessariamente mutuamente exclusivos, sendo que a morte de um microssatélite devido à uma mutação de ponto poderia predispor ao nascimento de um novo microssatélite (Hancock, 1999).

Vários modelos teóricos de mutação foram considerados para os microssatélites (revisados em Estoup & Cornuet, 1999). O Modelo da Mutação em Passos (*Stepwise Mutation Model*) descreve a mutação de alelos de microssatélites por meio de ganho ou perda de uma única unidade de repetição em tandem e conseqüentemente, alelos podem sofrer mutações para estados alélicos já presentes na população.

O Modelo do Alelo Infinito (*Infinite Allele Model*), em contraste, descreve que uma mutação envolve qualquer número de repetições em tandem e sempre resulta em um estado alélico previamente não encontrado na população.

No Modelo de Duas Fases (*Two Phase Model*), as mutações introduzem um ganho ou perda de x repetições. Com probabilidade p , x é igual a 1 (o que corresponde ao modelo da mutação em passos) e com probabilidade $1-p$, x segue uma distribuição geométrica.

Os modelos acima assumem que as diferenças alélicas são inteiramente devidas a alterações no número de unidades repetitivas básicas. A questão de qual modelo teórico deve ser aplicado a um marcador do tipo microsatélite é essencial porque diversas estimativas de parâmetros populacionais dependem do modelo assumido para o marcador.

1.4.2. FUNÇÃO DOS MICROSSATÉLITES NO GENOMA

Muitas evidências a respeito de possíveis funções dos microsatélites como elementos regulatórios ou codificantes têm se acumulado a partir da metade da década passada.

Como seqüências regulatórias, os microsatélites são encontrados ubiquamente em regiões promotoras a 5' de seqüências codificantes e em alguns casos são conservados em localizações correspondentes no mesmo gene em diferentes espécies (Xu *et al.*, 1997). Regiões promotoras contendo microsatélites servem como elementos *enhancer* em construções para expressão e as construções com deleção do microsatélite tem atividade de *enhancer* reduzida (Akai *et al.*, 1999). Também já foi demonstrado que microsatélites ligam-se a proteínas, o que é típico de seqüências ativadoras a 5' (Kashi & Soller, 1999).

Repetições do tetranucleotídeo TGGA, *in vitro*, apresentam baixa afinidade por octâmeros de histonas e impedem a formação de nucleossomos, podendo estar relacionadas à alta atividade transcricional (Cao *et al.*, 1998).

A regulação da expressão gênica por microsatélites também pode ocorrer por atenuação, conforme demonstrado para duas repetições de trinucleotídeos presentes no exon terminal do gene para o receptor de imunoglobulina polimérica de ratos (Aoki *et al.*, 1997).

Trechos de poliglutaminas ou poliprolinas foram encontrados em pelo menos 67 fatores de transcrição, incluindo a proteína humana que se liga à seqüência TATA e o receptor de androgênio de várias espécies. A variação no número de repetições afeta a função dos trechos de poliglutamina, em geral diminuindo a ativação transcricional com o aumento do número de repetições CAG (Kashi & Soller, 1999).

Várias doenças genéticas humanas resultam da expansão exagerada de microsatélites de trinucleotídeos nas extremidades não-codificantes de genes (como a repetição CGG na

síndrome do X Frágil e a repetição CTG na distrofia miotônica) ou dentro das regiões codificantes, como a expansão da repetição CAG na doença de Huntington e na ataxia espinocerebelar tipo 1, codificando poliglutamina (Ramel, 1997).

Uma importante função dos microssatélites na evolução tem sido mostrada em bactérias patogênicas como *Neisseria gonorrhoeae* e *Haemophilus influenzae*. Por exemplo, *N. gonorrhoeae* possui uma família de 12 proteínas na membrana externa codificadas pelos genes *Opa*. Estas proteínas permitem a adesão da bactéria a células epiteliais e a conseqüente invasão das mesmas. Cada gene *Opa* contém um microssatélite de pentanucleotídeos. Alterações no número de repetições podem ocorrer durante a replicação do DNA que precede a divisão da bactéria. A deleção de unidades repetitivas nos genes *Opa* leva à produção de proteínas incapazes de aderir à célula hospedeira e invadi-la. Um deslizamento subsequente pode adicionar a repetição perdida, permitindo que seja produzida novamente uma proteína funcional. Este mecanismo é importante para a sobrevivência da bactéria, visto que em certos momentos é útil que a bactéria entre nas células hospedeiras, mas em outras circunstâncias é estrategicamente vantajoso que a bactéria não interaja com células hospedeiras, principalmente com fagócitos que irão destruí-la (Stern *et al.*, 1986).

Um levantamento realizado a partir de seqüências de proteínas depositadas no SWISS-PROT (70.822 proteínas no total) revelou que 14% das proteínas apresentam uma ou mais duplicações internas estatisticamente significantes, sendo que as proteínas de eucariotos apresentam três vezes mais repetições internas do que as de procariotos (Marcotte *et al.*, 1998). Os autores sugerem que as repetições de proteínas eucarióticas apresentam funções únicas a este super-reino, já que as classes de proteínas mais propensas a apresentar repetições são predominantemente únicas de eucariotos, como por exemplo, proteínas do tecido conectivo, proteínas citoesqueléticas, proteínas musculares, proteínas sinápticas e cerebrais e proteínas de adesão celular.

Diante das possíveis funções expostas acima, a classificação dos microssatélites como DNA “parasita”, “egoísta” ou “lixo”, anteriormente empregada na literatura, torna-se inadequada para estas seqüências com tantas peculiaridades.

1.4.3. DISTRIBUIÇÃO E FREQUÊNCIA DE MICROSSATÉLITES NOS GENOMAS

Os microssatélites são abundantes nos genomas de praticamente todos os organismos eucarióticos estudados até o momento. Estudos realizados buscando-se por microssatélites em genomas completamente seqüenciados ou com cromossomos inteiros seqüenciados mostram que os microssatélites estão preferencialmente localizados em regiões não-codificantes (Dokholyan *et al.*, 2000; Katti *et al.*, 2001).

Os microssatélites presentes em regiões codificadoras limitam-se a repetições de trinucleotídeos ou repetições com múltiplos de três nucleotídeos, as quais não alteram a matriz de leitura (Katti *et al.*, 2001).

Os microssatélites com repetições CA são os mais abundantes em vários genomas de animais, seguidos por repetições AT e GA (Beckmann & Weber, 1992; Stallings, 1992; Katti *et al.*, 2001). Por outro lado, foi observada uma ausência de repetições GC em todos os organismos estudados (Dokholyan *et al.*, 2000; Katti *et al.*, 2001).

Nos genomas de plantas, a repetição AT é a mais freqüente seguida por GA (Lagercrantz *et al.*, 1993; Katti *et al.*, 2001).

Quanto à freqüência de microssatélites de trinucleotídeos, as repetições CAG são as mais freqüentes no genoma humano (Stallings, 1994) e de *Drosophila melanogaster* (Katti *et al.*, 2001), enquanto no genoma de *Arabidopsis thaliana* e de *Caenorhabditis elegans* as freqüências de repetições de trinucleotídeos AAG são comparativamente mais altas (Katti *et al.*, 2001).

A seqüência mais comum repetida em tandem no genoma humano é o arranjo de mononucleotídeo poli-A (Beckmann & Weber, 1992; Katti *et al.*, 2001). Foi sugerido que a abundância dessa seqüência no genoma humano é um reflexo de seu espalhamento por retrotransposons, sendo que a seqüência poli-A corresponderia à cauda de poli-A do mRNA transcrito reversamente (Beckmann & Weber, 1992; Hancock, 1999).

1.5. MICROSATÉLITES EM PARASITAS

Em diversos parasitas, o estudo de microssatélites tem fornecido importantes informações. Em um estudo feito em três espécies de *Leishmania*, utilizando-se dois loci de microssatélites do tipo CA, foi possível a identificação de polimorfismos entre diferentes linhagens do parasita, além da identificação de alguns heterozigotos, sugerindo que *Leishmania* pode ser diplóide, com reprodução predominantemente por auto-fertilização (Rossi *et al.*, 1994).

Oito loci de microssatélites CA foram utilizados para a tipagem genética de linhagens de *Trypanosoma cruzi*, agente causador da doença de Chagas. Os resultados indicaram que o parasita é diplóide, com reprodução sexual rara ou ausente, com estrutura populacional clonal, alto grau de polimorfismo e que clones se associam para formar linhagens multiclonais (Oliveira *et al.*, 1996). A análise de mais linhagens com os mesmos loci revelou que isolados de pacientes na fase crônica da doença de Chagas também podem ser multiclonais, embora em menor extensão do que linhagens isoladas de insetos vetores, mamíferos e humanos com a doença na fase aguda (Oliveira *et al.*, 1999).

No nematódeo parasita de cabras e ovelhas, *Haemonchus contortus*, a análise de 10 loci de microssatélites CA imperfeitos revelou extensa variabilidade genética entre e dentro de populações, mostrando que uma população originária da África divergiu em grande extensão de duas populações europeias e uma asiática devido ao isolamento geográfico. Além disso, 8 dos 13 loci isolados no estudo apresentaram ligação com um elemento repetitivo direto de 128 pb (Hoekstra *et al.*, 1997).

Outros 59 loci de microssatélites CA imperfeitos foram isolados em *H. contortus*. Entre estes, 26 também estavam associados a repetição direta de 128pb, permitindo o desenho de um *primer* genérico para a amplificação de alguns destes loci. Quatro populações foram estudadas pela análise de 30 loci, revelando que 8 loci podem discriminar completamente entre duas ou mais populações. Duas populações apresentaram pouco polimorfismo intrapopulacional, uma foi levemente mais variável e a última apresentou significativamente mais alelos que as outras três. A análise também permitiu verificar a ocorrência de contaminação genética em uma linhagem propagada experimentalmente em ovelhas, além de permitir a identificação da geração onde se deu a contaminação e das possíveis fontes de contaminação (Otsen *et al.*, 2000).

Dois locos de microsatélites de *Brugia malayi* (agente causador da filariose), o primeiro localizado no cromossomo Y (consequentemente, específico de machos) e o outro de localização desconhecida, mas segregando independentemente do primeiro, mostraram polimorfismo entre isolados da Malásia e da Indonésia, em um estudo preliminar (Underwood *et al.*, 2000).

O estudo de microsatélites também tem sido importante no mapeamento genético de *Plasmodium falciparum*, agente causador da malária, para o qual desenvolveu-se um mapa de ligação de alta resolução, consistindo de centenas de locos de microsatélites (Ferdig & Su, 2000).

Somente três espécies de platelmintos tiveram microsatélites isolados de seus genomas: *Dugesia polychroa* (Ramachandran *et al.*, 1997), *Echinococcus multilocularis* (Bretagne *et al.*, 1996) e *Schistocephalus solidus* (Binz *et al.*, 2000), sendo os dois últimos pertencentes à classe Cestoda.

O microsatélite descrito para *E. multilocularis* é uma repetição de pentanucleotídeos presente no gene que codifica o snRNA U1, o qual está repetido em tandem mais de 50 vezes. A análise de 41 isolados permitiu a identificação de 3 perfis, um encontrado nos isolados provenientes da Europa, outro nos isolados do Alasca e outro nos isolados do Japão (Bretagne *et al.*, 1996).

1.6. POR QUE ESTUDAR MICROSSATÉLITES EM *E. granulosus*?

Os microsatélites têm se mostrado marcadores moleculares altamente polimórficos nos mais diferentes organismos estudados e extremamente úteis para responder a questões biológicas. Entretanto, até o momento não foram isolados microsatélites do genoma de *E. granulosus*.

O isolamento de microsatélites em *E. granulosus* e posterior utilização como marcador molecular em estudos de estrutura de populações pode permitir um melhor entendimento da biologia deste parasita.

A partir da análise de alelos presentes em diferentes locos de microsatélites pode-se verificar se há algum padrão genotípico específico para as diferentes linhagens, o que permitiria identificá-las utilizando-se estas seqüências.

Variação genética entre diferentes linhagens de *E. granulosus* tem sido relatada, com a utilização de diferentes métodos moleculares, com mais frequência do que variação dentro de linhagens (Bowles *et al.*, 1992, 1995; McManus & Bryant, 1995; Gasser & Chilton, 1995; Siles-Lucas *et al.*, 1996; Haag *et al.*, 1999). Por serem muito polimórficos, os microssatélites tendem a apresentar vários alelos em cada locus, podendo facilitar a detecção de variação genética dentro de linhagens.

Por serem marcadores moleculares com padrão de herança mendeliana co-dominante, a análise dos microssatélites também pode permitir que se identifique um padrão predominante de cruzamento, se auto-fertilização ou fertilização cruzada.

Devido à importância da hidatidose como zoonose no Rio Grande do Sul, o isolamento de marcadores moleculares altamente polimórficos pode facilitar análises populacionais e colaborar para o controle epidemiológico da hidatidose, pela correlação dos dados moleculares com as diferenças quanto à infectividade, transmissão, aspectos bioquímicos e do desenvolvimento das linhagens de *E. granulosus*.

1.7. OBJETIVOS

1.7.1. OBJETIVO GERAL

Este estudo teve como objetivo o isolamento e a caracterização de locus de microssatélites do genoma de *E. granulosus*.

1.7.2. OBJETIVOS ESPECÍFICOS

- Verificação da presença e abundância de seqüências repetitivas do tipo microssatélites no genoma de *E. granulosus*.
- Isolamento de microssatélites do genoma de *E. granulosus*, clonagem e caracterização das seqüências repetitivas.
- Seleção de seqüências contendo microssatélites candidatas à utilização como marcadores moleculares.
- Avaliação da estabilidade de um locus de microssatélite em protoescoléces isolados provenientes de um único cisto.
- Teste de locus de microssatélites como marcadores moleculares no estudo de populações do parasita.
- Isolamento do microssatélite contido no gene do snRNA U1, a partir de seqüência ortóloga descrita para *E. multilocularis* (Bretagne *et al.*, 1996) e verificação de sua utilidade como marcador molecular.

2. MATERIAL E MÉTODOS

2.1. CISTOS HIDÁTICOS

Os cistos hidáticos foram extraídos de bovinos provenientes de regiões endêmicas do estado do Rio Grande do Sul, durante o abate nos frigoríficos Líder (Viamão, RS) e Rost (atual Cooperleo, São Leopoldo, RS). No laboratório, o líquido hidático era aspirado utilizando-se uma seringa estéril com agulha de 50 x 20. Os protoescócecs, quando presentes no líquido, eram sedimentados por gravidade, separados do sobrenadante e armazenados a -80°C até o momento do uso.

2.2. LINHAGEM BACTERIANA, VETOR DE CLONAGEM, ENZIMAS, REAGENTES, OLIGONUCLEOTÍDEOS, ANTIBIÓTICOS, MEIOS DE CULTURA E SOLUÇÕES UTILIZADOS

A linhagem XL1Blue de *Escherichia coli* foi utilizada nas transformações bacterianas (Sambrook & Russel, 2001). O plasmídeo pGEM-T Easy Vector (Promega) foi utilizado como vetor de clonagem.

As enzimas de restrição e de modificação utilizadas, assim como os respectivos tampões de reação, foram provenientes de diferentes fabricantes, como Cenbiot, Gibco, Amersham-Pharmacia, Promega, USB e Biolabs.

Os dNTPs utilizados nas reações em cadeia da polimerase foram adquiridos da Amersham-Pharmacia. Os oligonucleotídeos utilizados foram sintetizados por diferentes fabricantes, entre eles BioSynthesis, Operon Technologies, The Midland Certified Reagent Company, Only DNA e Gibco.

Foram utilizados os meios de cultura LB (líquido e sólido) e SOC, preparados de acordo com Sambrook & Russel (2001).

Os antibióticos utilizados foram ampicilina e tetraciclina (Amersham-Pharmacia), os quais foram preparados e utilizados de acordo com Sambrook & Russel (2001).

As soluções e meios foram preparados de acordo com Sambrook & Russel (2001), utilizando-se reagentes de diferentes procedências (Gibco, Amersham-Pharmacia, Merck, Reagen, Synth).

2.3. POPULAÇÕES DE PARASITAS

Na análise populacional foram utilizadas 100 amostras de DNA de uma genoteca pertencente ao Laboratório de Biologia Molecular de Cestódeos, obtidas a partir de protoescoléces ou camada germinativa de metacestóides provenientes de diferentes hospedeiros intermediários (humanos, bovinos e ovinos) e localidades (região sul do Brasil, Argentina e África).

2.4. EXTRAÇÃO DE DNA GENÔMICO DE ISOLADOS DE *Echinococcus granulosus*

A extração de DNA de protoescoléces a partir de cistos únicos foi feita conforme descrito anteriormente (McManus *et al.*, 1985), com pequenas modificações. Aproximadamente 100 µL de protoescoléces foram incubados em solução de lise (Tris-HCl 100mM, EDTA (ácido etilenodiaminotetraacético) 100mM, NaCl 250mM) contendo 800 µg de proteinase K e SDS (dodecilsulfato de sódio) 2%, a 56°C durante 90 minutos. Em seguida, foram realizadas três extrações com fenol-clorofórmio (1:1) e uma extração com clorofórmio. O DNA foi precipitado com dois volumes de etanol e 0,01 volume de NaCl 5M, centrifugado a 13000 rpm por 15 minutos e o precipitado lavado com etanol 70%. Depois de seco o DNA foi ressuscitado em TE (Tris-HCl 10mM, EDTA 1mM) e tratado com RNaseI. A solução de DNA foi então dialisada contra TE durante 12 horas, havendo três trocas de tampão durante este período.

Uma pequena alíquota do DNA assim obtido (1 a 2 µL) foi misturada ao tampão de amostra (25% Ficoll 400, 0,5% azul de bromofenol, 0,5% xilenocianol), aplicada em gel de agarose 0,6% e submetida à eletroforese em tampão TEB 1x para verificação do peso molecular. A concentração do DNA na amostra foi determinada através de um espectrofotômetro.

2.5. SOUTHERN BLOT

2.5.1. CLIVAGEM DO DNA COM ENDONUCLEASES DE RESTRIÇÃO E TRANSFERÊNCIA PARA MEMBRANA DE NÁILON

Cinco μg de DNA genômico de *E. granulosus* foram clivados com as seguintes endonucleases de restrição: *AluI* (Gibco BRL), *TaqI* (Cenbiot), *RsaI* (Gibco BRL), *EcoRI* (Gibco BRL) ou *HindIII* (Gibco BRL), utilizando os tampões de reação fornecidos pelos fabricantes.

O DNA clivado foi aplicado em gel de agarose 0,8% e submetido à eletroforese em tampão TEB 1x a 1V/cm, para separação dos fragmentos.

Após a separação por eletroforese, os fragmentos de DNA foram transferidos para membranas de náilon (Hybond N⁺, Amersham-Pharmacia) por capilaridade em tampão 10x SSC (Sambrook & Russel, 2001) e fixados à membrana através de aquecimento a 80°C durante duas horas. As membranas foram utilizadas na hibridização com sondas de oligonucleotídeos contendo repetições e com uma sonda contendo parte do gene do snRNA U1.

2.5.2. MARCAÇÃO RADIOATIVA DAS SONDAS DE OLIGONUCLEOTÍDEOS E HIBRIDIZAÇÃO

Dez pmol de oligonucleotídeos (Tabela 1) foram marcados na extremidade 5' com 50 μCi de $\gamma^{32}\text{P}$ -ATP (5000 Ci/mmol, Amersham-Pharmacia), utilizando 3U da enzima T4 polinucleotídeo-quinase (USB), conforme o protocolo do fabricante da enzima.

As membrana de náilon contendo os fragmentos clivados foram pré-hibridizadas durante 16 horas a 65°C em 5x SSC, 5x solução de Denhardt (0,1% Ficoll 400, 0,1% polivinilpirrolidona, 0,1% albumina sérica bovina), 0,1% SDS e 4% leite em pó desnatado. A hibridização foi feita em uma solução contendo 5x SSC, 5x solução de Denhardt, 0,1% SDS e 10 pmol de uma das sondas marcadas, durante 4 horas. A temperatura de hibridização e as condições de lavagem da membrana variaram de acordo com a T_m dos oligonucleotídeos marcados radioativamente (Tabela 1). As membranas foram expostas a filme de raios-X, a -80°C.

Tabela 1. Seqüências dos oligonucleotídeos utilizados como sondas e condições utilizadas nos experimentos de hibridização.

Seqüência do Oligo	Tm do Oligo (°C)	Temperatura de Hibridização	Condições de Lavagem da Membrana
(GT) ₁₅	90	65	2xSSC/0,1% SDS, 50°C, 10 minutos (2x) 1xSSC/0,1% SDS, 50°C, 10 minutos (1x)
(CT) ₁₅	90	65	2xSSC/0,1% SDS, 50°C, 10 minutos (2x) 1xSSC/0,1% SDS, 50°C, 10 minutos (1x)
(AT) ₁₅	60	40	5xSSC/0,1% SDS, room temperature, 5 minutos (2x)
(CG) ₁₅	Não determinada	65	2xSSC/0,1% SDS, 50°C, 10 minutos (2x) 1xSSC/0,1% SDS, 50°C, 10 minutos (2x)
(CAT) ₁₀	80	60	2xSSC/0,1% SDS, 50°C, 10 minutos (3x)
(CAA) ₁₀	80	60	2xSSC/0,1% SDS, 50°C, 10 minutos (3x)
(CGG) ₁₀	Não determinada	65	2xSSC/0,1% SDS, 50°C, 10 minutos (2x) 1xSSC/0,1% SDS, 50°C, 10 minutos (1x)
(CATA) ₁₀	Não determinada	65	2xSSC/0,1% SDS, 50°C, 10 minutos (3x)

2.5.3. MARCAÇÃO RADIOATIVA DA SONDA snRNA U1 E HIBRIDIZAÇÃO

Cem ng do produto de PCR do gene snRNA U1 (descrito no item 2.6.2) foram marcados por *random priming* com 20 µCi de $\alpha^{32}\text{P}$ -dATP (3000 Ci/mmol, Amersham Pharmacia), segundo Sambrook & Russel (2001).

A membrana de náilon contendo os fragmentos de DNA foi pré-hibridizada a 60°C por uma hora em uma solução contendo 5x SSC, 5x solução de Denhardt, 0,1% SDS e 4% leite em pó desnatado. A hibridização foi feita a 60°C, durante 16 horas, em uma solução contendo 5x SSC, 5x solução de Denhardt, 0,1% SDS e a sonda radioativa (aproximadamente 10^7 cpm) previamente desnaturada. Posteriormente as membranas foram lavadas por 20 min, sucessivamente, em tampão 5x SSC/0,1% SDS; 2x SSC /0,1% SDS; 1x SSC/0,1% SDS e

finalmente em 0,2x SSC/0,1% SDS. A membrana foi exposta a filme de raios-X, durante 24 horas, a -80°C .

2.6. ISOLAMENTO DOS MICROSSATÉLITES A PARTIR DE DNA GENÔMICO DE *E. granulosus*

2.6.1. ISOLAMENTO DE SEQUÊNCIAS CONTENDO REPETIÇÕES DO TIPO CA, GA, AAC E TA

O isolamento dos microssatélites contendo repetições $(\text{CA})_n$, $(\text{GA})_n$, $(\text{AAC})_n$ e $(\text{TA})_n$ foi realizado conforme previamente descrito (Refseth *et al.*, 1997; Oliveira *et al.*, 1998), com algumas modificações. DNA genômico foi digerido com *TaqI* (Gibco BRL) e 500 ng de fragmentos foram ligados a 50 pmol de um adaptador a 16°C durante 16 horas usando 1 unidade de T4 DNA ligase (Gibco BRL).

O adaptador para *TaqI* foi formado pelo anelamento dos oligonucleotídeos *TaqI*-20 Mer (5' ATGAAGCCTTGGTACTGGAT 3') e *TaqI*-22 Mer (5' CGATCCAGTACCAAGGCTTCAT 3') (MCRC). Cinquenta pmol de cada oligo foram misturados, aquecidos a 70°C por cinco minutos e resfriados até 10°C , diminuindo-se 1°C por minuto, em um termociclador.

A captura dos fragmentos contendo repetições foi feita baseando-se na propriedade das moléculas de biotina de se ligarem à estreptavidina (figura 2.1). Assim, oligonucleotídeos biotinizados contendo repetições $(\text{GATGATCCGACGCAT}(\text{CA})_{12}$, $\text{GATGATCCGACGCAT}(\text{GA})_{12}$, $\text{GATGATCCGACGCAT}(\text{AAC})_{10}$ ou $\text{GATGATCCGACGCAT}(\text{TA})_{12}$), foram anelados ao DNA genômico clivado de *E. granulosus* e capturados utilizando-se microesferas magnéticas recobertas com estreptavidina. Cem ng de fragmentos de DNA ligados ao adaptador foram misturados a 20 pmol de cada um dos oligonucleotídeos biotinizados acima, separadamente, em 50 μL de tampão TE contendo 2 μM do oligonucleotídeo *TaqI*-20 Mer para evitar a formação de grampos. A mistura foi aquecida a 95°C durante 10 minutos e anelada a 60°C (ou 40°C para o oligo AT) durante 1 minuto. Cem μg de esferas magnéticas recobertas com estreptavidina (Dynabeads M280, Dynal) foram adicionadas à mistura e ressuspensas em 500 μL de TE/1 M NaCl. Após 30

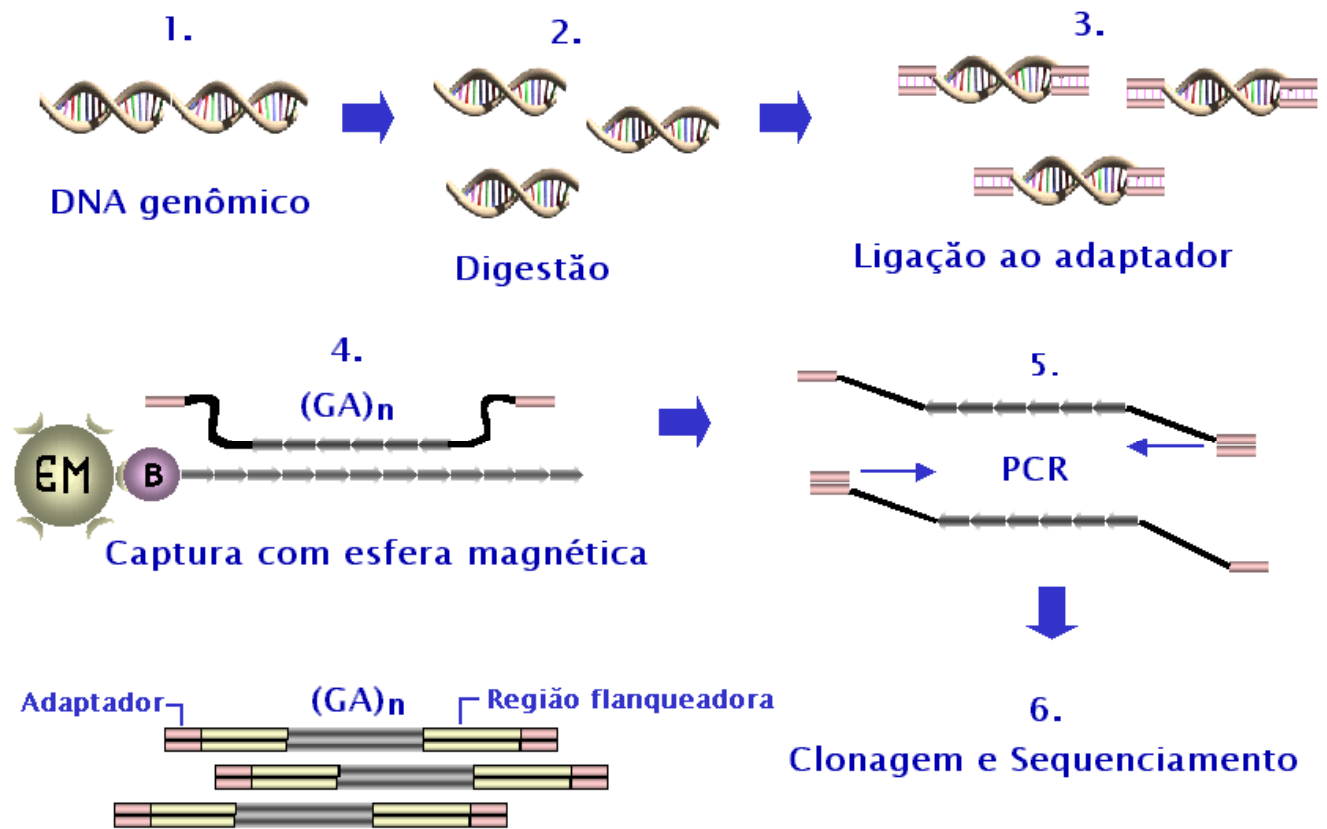


Figura 2.1. Captura dos fragmentos contendo microssatélites. Após clivagem do DNA genômico (1) de *E. granulosus* com endonuclease de restrição (2), os fragmentos obtidos eram ligados a oligonucleotídeos adaptadores (3), desnaturados e hibridizados com um oligonucleotídeo repetitivo biotilado na extremidade 5'. Os fragmentos contendo seqüências complementares ao oligonucleotídeo biotilado eram capturados com microesferas magnéticas recobertas com estreptavidina e separados com o auxílio de um imã (4). Em seguida, os fragmentos capturados eram submetidos à PCR (5), clonados em plasmídeos e seqüenciados (6). B = biotina; EM = esfera magnética.

minutos de incubação a temperatura ambiente com agitação suave, as esferas foram coletadas com o auxílio de um separador magnético (Gibco BRL), lavadas duas vezes com 2x SSC/0,5% SDS a 50°C durante cinco minutos, uma vez com TE/2 M NaCl e uma vez com TE/1 M NaCl a temperatura ambiente. Para a captura de seqüências repetidas AT, somente as duas últimas lavagens foram realizadas.

Os fragmentos de DNA fita simples imobilizados foram eluídos das esferas magnéticas em 50 µL de H₂O destilada a temperatura ambiente (no caso dos oligos biotinilados AT e AAC) ou a 90°C (no caso dos oligos biotinilados CA e GA).

A obtenção de fragmentos fita-dupla foi feita através da amplificação do DNA capturado por PCR. A reação, preparada em um volume total de 50 µl, continha 10 µl de fragmentos capturados, 10 pmol do oligonucleotídeo *TaqI*-20 mer, 2.5 unidades de *Taq* DNA polimerase (Cenbiot), 100 µm de cada dNTP e tampão para *Taq* DNA polimerase (10mM tris-HCl pH8,5; 50 mM KCl; 4mM MgCl₂).

As amostras foram incubadas a 94°C durante 3 minutos e submetidas a 23 ciclos consistindo de 1 minuto a 94°C para desnaturação, 3 minutos a 72°C para anelamento e extensão e então incubadas para uma extensão final durante 5 minutos a 72°C. Os produtos de PCR foram visualizados sob luz ultravioleta após eletroforese em gel de agarose 0,8% corado com brometo de etídeo (0,15µg/mL) e purificados em colunas GFX (Amersham-Pharmacia).

2.6.2. ISOLAMENTO DO MICROSSATÉLITE COM REPETIÇÕES DE PENTANUCLEOTÍDEOS CONTIDO NO GENE *snRNA U1*

O isolamento de uma repetição de pentanucleotídeos contida no gene do *snRNAU1* de *E. granulosus* foi feito através da amplificação de um segmento gênico de 258 pb contendo o microssatélite, utilizando-se *primers* heterólogos, projetados a partir da seqüência do gene do *snRNAU1* descrita para *E. multilocularis* (Bretagne *et al.*, 1996). A PCR continha 20 ng de DNA genômico, 20 pmol de cada *primer* (*snRNA U1-F*: 5'ATTGTCGTTGCCATCTCTCC3' e *snRNA U1-R*: 5'CGTCTCCATCACACACATC3'), 100µM de cada dNTP, 1 U *Taq* DNA polimerase e tampão para *Taq* DNA polimerase (10 mM tris-HCl, 50mM KCl) com 1,5 mM MgCl₂. A amostra foi submetida a 20 ciclos de 1 minuto a 94°C para desnaturação, 1 minuto a 50°C para anelamento com diminuição de 1°C a cada ciclo e 1 minuto a 72°C para extensão, seguidos por mais 20 ciclos com temperatura de anelamento de 40°C e de uma

extensão final a 72°C durante 5 minutos. O produto de PCR foi visualizado sob luz ultravioleta após eletroforese em gel de agarose 1,2% corado com brometo de etídeo (0,15µg/ml) e purificado em coluna Microspin S400 (Pharmacia).

O produto de PCR obtido foi utilizado para clonagem e também, para um *Southern* genômico, descrito no item 2.5.3.

2.7. CLONAGEM DOS FRAGMENTOS CONTENDO MICROSSATÉLITES

Os fragmentos amplificados por PCR, contendo microssatélites com repetições CA, GA e AAC ou a repetição de pentanucleotídeos do gene snRNA U1 foram clonados no vetor pGEM T-Easy (Promega), de acordo com as instruções do fabricante.

2.8. TRANSFORMAÇÃO

Bactérias *Escherichia coli* XL1Blue eletrocompetentes foram misturadas com a solução contendo plasmídeo e inserto, transferidas para cubetas de eletroporação de 0,2 cm de largura e transformadas com o auxílio de um eletroporador (Genepulser, Biorad), ajustado para 2,5 kV de voltagem, 200 Ω de resistência e 25 µF de capacitância. Após o pulso elétrico, as bactérias foram incubadas em meio SOC a 37°C durante 1 hora, semeadas em placas contendo meio LB sólido, ampicilina (100 µg/mL), IPTG (80 µM) e X-Gal (32 µg/mL) e incubadas a 37°C durante 16 horas para o crescimento de colônias isoladas.

2.9. SELEÇÃO DE COLÔNIAS RECOMBINANTES

As colônias brancas, possivelmente provenientes de transformações com plasmídeos contendo insertos de microssatélites CA ou AAC, foram transferidas com o auxílio de um palito estéril para uma membrana de náilon quadriculada e numerada, colocada sobre uma placa de petri contendo meio LB sólido, ampicilina, IPTG e X-Gal (nas concentrações descritas no item anterior) e repicadas em uma placa réplica, também quadriculada e numerada. As placas foram incubadas por 16 horas a 37°C. Em seguida, a placa réplica foi armazenada a 4 °C .

A membrana contendo colônias foi tratada por 10 minutos com cada uma das seguintes soluções: SDS 10%, solução desnaturante (0,5M NaOH/1,5M NaCl), solução neutralizante (1,5M NaCl) e 20x SSC. Após secagem ao ar, a membrana foi fixada a 80°C durante 2 horas.

A membrana foi então hibridizada com sonda (CA)_n ou (AAC)_n conjugadas à fosfatase alcalina (LIFECODES), de acordo com as instruções do fabricante e expostas a filme de raios-X durante 1 hora. Após a revelação e fixação do filme, as colônias contendo microssatélites foram identificadas e transferidas da placa réplica para tubos de ensaio contendo 3 mL de meio LB líquido e ampicilina (50 µg/mL) e incubadas por 15 horas a 37°C com agitação horizontal a 200 rpm.

Para as colônias contendo plasmídeos com insertos repetitivos dos tipos (GA)_n, (TA)_n ou o microssatélite do gene snRNA U1, para os quais não se dispunha de sondas, foi feita a transferência direta das colônias de coloração branca para tubos de ensaio contendo meio LB e ampicilina.

As culturas foram submetidas ao processo de lise alcalina (Sambrook & Russel, 2001) para extração de DNA plasmidial. Os plasmídeos extraídos foram clivados com *EcoRI* e submetidos à eletroforese em gel de agarose 1,5% corado com brometo de etídeo para visualização sob luz UV dos fragmentos gerados pela clivagem e confirmação da presença de um inserto.

2.10. SEQÜENCIAMENTO DOS PLASMÍDEOS RECOMBINANTES

Os plasmídeos recombinantes foram submetidos ao seqüenciamento manual, utilizando-se o “Thermo Sequenase Radiolabeled Terminator Cycle Sequencing kit” (USB) ou ao seqüenciamento automático, utilizando-se o “DYEnamic Dye Terminator Cycle Sequencing kit” para MegaBACE (Amersham-Pharmacia). Ambas as fitas dos insertos foram seqüenciadas, utilizando-se os *primers* universais T7 e SP6.

2.11. ELABORAÇÃO DE PRIMERS PARA AS REGIÕES FLANQUEADORAS DOS MICROSSATÉLITES CA E GA

Após o seqüenciamento dos plasmídeos recombinantes foram projetados *primers* posicionados nas regiões flanqueadoras dos microssatélites com o auxílio do programa OLIGO (Rychlik & Rhoads, 1989).

2.12. ESTABELECIMENTO DE CONDIÇÕES DE PCR PARA A AMPLIFICAÇÃO DE UM LÓCUS DE MICROSSATÉLITE CA E UM LÓCUS GA

As condições de PCR testadas para a amplificação dos microssatélites compreenderam variações nas temperaturas de anelamento dos *primers* e nas concentrações de sal ($MgCl_2$) e *primers*, mantendo-se constantes as concentrações de tampão para *Taq* DNA polimerase (10 mM tris-HCl, 50mM KCl), dNTPs (100 μ M cada um) e *Taq* DNA polimerase (2,5 U) por reação (volume total 50 μ L). Em cada reação foram utilizados 10 a 20 ng de DNA genômico de *E. granulosus* como molde.

Foram testadas as concentrações de 10 e 20 pmol de cada *primer* projetado (as seqüências estão descritas no item Resultados) por reação e as concentrações de 0,5 mM, 1mM, 1,5 mM, 2 mM, 2,5 mM e 3 mM $MgCl_2$.

Foram testados 4 diferentes programas (TD 60/50, TD 65/55, TD 70/60, TD 75/65) para PCR em termociclador, listados na tabela 2.

Como controle positivo de amplificação foram utilizados os plasmídeos contendo os microssatélites inseridos.

Reações contendo DNA de hospedeiros intermediários de *E. granulosus* (bovino, ovino e humano) e reações sem DNA foram utilizadas como controles negativos de amplificação.

Tabela 2. Descrição dos programas utilizados.

Passo	Temperatura °C (TD 60/50)	Temperatura °C (TD 65/55)	Temperatura °C (TD 70/60)	Temperatura °C (TD 75/65)	Tempo (minutos)
1	94	94	94	94	5
2	60 -1°C/ciclo	65 -1°C/ciclo	70 -1°C/ciclo	75 -1°C/ciclo	1
3	72	72	72	72	1
4	94	94	94	94	1
5	60 -1°C/ciclo	65 -1°C/ciclo	70 -1°C/ciclo	75 -1°C/ciclo	1
6	72	72	72	72	1
7	94	94	94	94	1
8	10 x para o passo 2	10 x para o passo 2	10 x para o passo 2	10 x para o passo 2	
9	50	55	60	65	1
10	72	72	72	72	1
11	94	94	94	94	1
12	20 x para o passo 9	20 x para o passo 9	20 x para o passo 9	20 x para o passo 9	
13	50	55	60	65	2
14	72	72	72	72	10
15	4	4	4	4	infinito

2.13. ELETROFORESE DOS FRAGMENTOS DE DNA EM GEL DE POLIACRILAMIDA

Após a PCR, 5 μL da reação com os fragmentos amplificados contendo microssatélites foram adicionados a 2 μL de tampão de amostra (25% Ficoll 400, 0,5% azul de bromofenol, 0,5% xilenocianol) e submetidos à eletroforese em gel longo (40cm) de poliacrilamida 8%, não-desnaturante em tampão TEB 1x, durante 6 horas a 15W.

2.14. COLORAÇÃO DO GEL DE POLIACRILAMIDA

Após o término da eletroforese, os géis foram imediatamente removidos para uma solução de HNO_3 1%, na qual permaneceram por 4 minutos. Em seguida, os géis foram lavados em H_2O destilada duas vezes e corados com uma solução de nitrato de prata 0,2% durante 20 minutos, novamente lavados em H_2O destilada e então revelados com uma solução de carbonato de sódio 3% e formaldeído 0,05% até o aparecimento das bandas de DNA. A coloração foi interrompida com ácido acético 10% e os géis foram secos dentro de celofane embebido em solução de metanol 30% e glicerol 1%.

2.15. TESTE DA ESTABILIDADE DE UM LÓCUS DE MICROSSATÉLITE ATRAVÉS DE PCR DE PROTOESCÓLECES INDIVIDUAIS

Protoescóleces retirados de seis cistos de hospedeiros bovinos provenientes do estado do RS foram diluídos em água destilada. Cem μL de solução foram transferidos para poços de uma placa de ELISA e visualizados sob microscópio óptico. Os poços que apresentavam um único protoescólece tiveram seu conteúdo transferido para tubos de PCR.

Após breve centrifugação dos tubos (1 minuto a 13.000 rpm), foram retirados 75 μL de água de cada tubo, permanecendo portanto, 25 μL de água mais o protoescólece em cada um. A cada tubo foram adicionados 2 μg de proteinase K e em seguida, os tubos foram incubados a 50°C durante 1 hora. Adicionou-se óleo mineral aos tubos que foram então incubados a 95°C durante 20 minutos.

O material lisado foi utilizado em uma PCR, que continha 10 pmol de cada *primer* (Egmsca1-F e Egmsca1-R, projetados conforme descrito no item 2.11), 50 μM de cada

dNTP, 2,5 U *Taq* polimerase, tampão para amplificação 1x (10 mM tris-HCl, 50 mM KCl) e 1,5 mM MgCl₂. No total, 150 amostras de um cisto e 20 amostras dos outros 5 cistos foram submetidas à amplificação, utilizando-se o programa TD 70/60 (listado na Tabela 2).

As amostras amplificadas foram submetidas à eletroforese em gel de poliacrilamida 8 %, conforme descrito nos itens 2.13 e 2.14.

2.16. AMPLIFICAÇÃO DO LÓCUS *Egmsca1* EM ISOLADOS DE DIFERENTES PROCEDÊNCIAS

Após o estabelecimento das condições ideais de amplificação para o lócus *Egmsca1*, DNAs genômicos (10 a 20 ng) de 100 isolados (item 2.3) foram amplificados em reações contendo 20 pmol de cada *primer* (*Egmsca1*-F e *Egmsca1*-R), 100 µM cada dNTP, 2,5 U *Taq* polimerase, tampão de amplificação 1x (10 mM tris-HCl, 50mM KCl) e 1,5 mM MgCl₂.

As amostras foram submetidas à eletroforese em gel de poliacrilamida, conforme descrito nos itens 2.13 e 2.14.

2.17. ANÁLISE ESTATÍSTICA

A análise dos dados de polimorfismo do lócus *Egmsca1* foi feita com o auxílio do programa GENEPOP (Raymond & Rousset, 1995a).

Somente as populações da linhagem da ovelha provenientes do Brasil e da Argentina (com $n > 10$) foram testadas em relação ao equilíbrio de Hardy-Weinberg e quanto à diferenciação gênica e genotípica.

As proporções de Hardy-Weinberg foram testadas através de um teste exato em tabelas de contingência (Weir, 1990). Neste teste o valor P corresponde à soma das probabilidades de todas as tabelas (com a mesma contagem alélica) com a mesma ou menor probabilidade que a probabilidade das amostras observadas, a qual é utilizada para definir a zona de rejeição.

Os testes para diferenciação gênica e genotípica (relacionada com a distribuição de alelos e genótipos nas várias populações) foram realizados em tabelas de contingência por um teste exato de Fisher (Raymond & Rousset, 1995b) e por teste G (Goudet *et al.*, 1996). As estimativas de F_{is} e F_{st} foram calculadas seguindo Weir & Cockerham (1984).

2.18. COMPARAÇÃO DAS SEQÜÊNCIAS CONTENDO MICROSSATÉLITES ISOLADAS COM SEQÜÊNCIAS DEPOSITADAS EM BANCOS DE DADOS.

As seqüências contendo microssatélites foram comparadas a seqüências depositadas no Genbank utilizando-se o programa BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>).

3. RESULTADOS

3.1. VERIFICAÇÃO DA ABUNDÂNCIA DE SEQUÊNCIAS REPETITIVAS NO GENOMA DE *E. granulosus* UTILIZANDO SONDAS DE OLIGONUCLEOTÍDEOS MARCADAS RADIOATIVAMENTE

Dentre as oito sondas utilizadas para a hibridização com o DNA genômico de *E. granulosus*, seis sondas [(GT)₁₅, (CT)₁₅, (CAT)₁₀, (CAA)₁₀, (CGG)₁₀, (CATA)₁₀] geraram padrões de bandas únicas ou múltiplas variando de aproximadamente 0,3 a 10 kb (Fig. 3.1).

As sondas (GT)₁₅, (CAA)₁₀, (CATA)₁₀ e (CT)₁₅ geraram padrões de bandas múltiplas, independentemente da enzima utilizada para a clivagem do DNA. Para as três primeiras sondas, a quantidade de bandas geradas foi tão grande e uniformemente distribuída que não foi possível contá-las.

A sonda (CAT)₁₀ gerou um padrão de bandas únicas para cada enzima utilizada para clivar o DNA e a sonda (CGG)₁₀ gerou de uma a três bandas de hibridização, variando de acordo com a enzima utilizada.

Duas sondas, (GC)₁₅ e (AT)₁₅, não apresentaram bandas de hibridização, indicando uma possível ausência dessas repetições no genoma de *E. granulosus*.

3.2. VERIFICAÇÃO DA FREQUÊNCIA DO GENE snRNA U1 NO GENOMA DE *E. granulosus*

A hibridização da sonda snRNA U1 com o DNA genômico clivado de *E. granulosus* (Fig. 3.2) gerou uma única banda de aproximadamente 10kb no DNA digerido com a enzima *Hind*III. Além de uma banda de aproximadamente 10 kb, a hibridização da sonda com o DNA digerido com *Eco*RI revelou algumas bandas adicionais, correspondendo a fragmentos de menor tamanho, muito fracas.

Nove bandas, variando de 0,5 a 10 kb, foram visualizadas no DNA digerido com *Rsa*I. Duas bandas fortes, com aproximadamente 1 e 1,7 kb foram identificadas no DNA digerido com *Alu*I e com *Taq*I, além de bandas fracas.

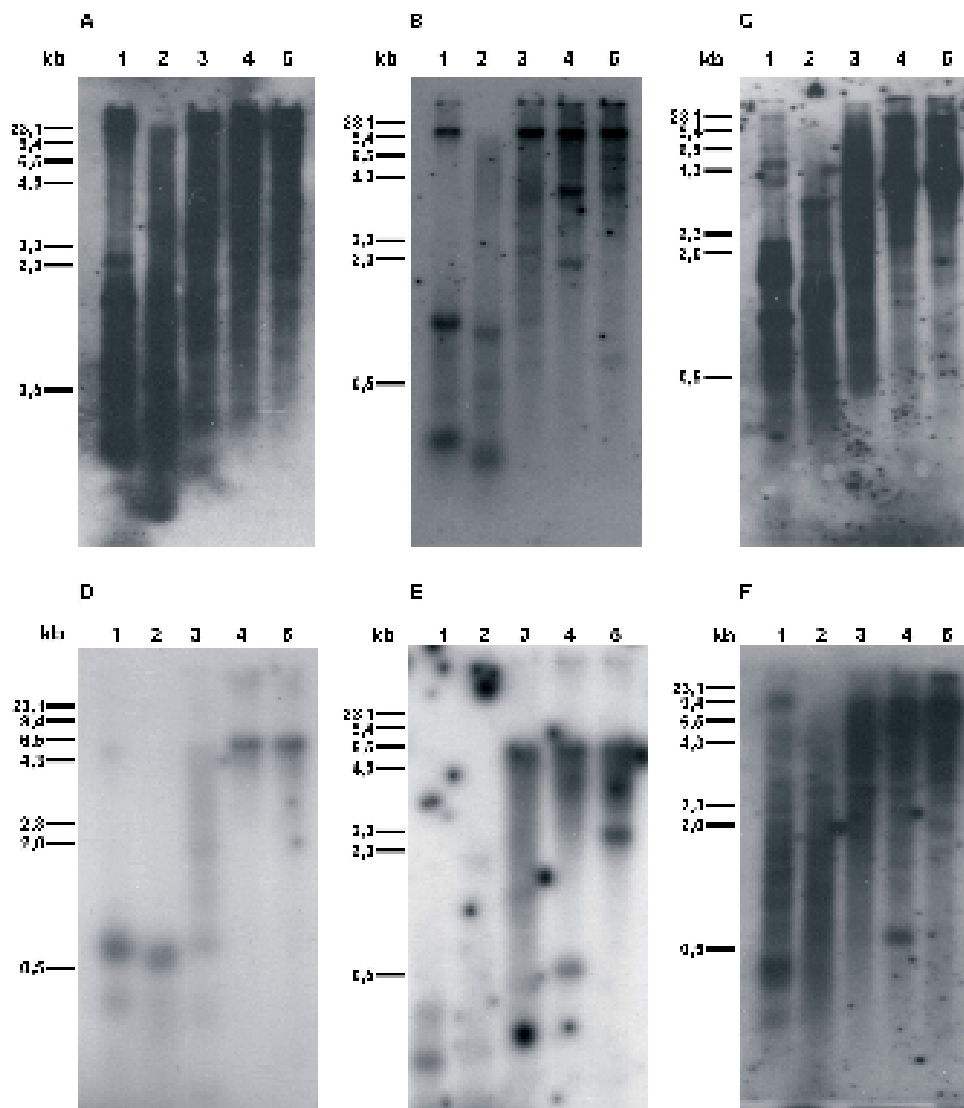


Figura. 3.1. Presença e abundância de microssatélites em *E. granulosus*. O DNA total obtido a partir de protoescoléces de um único cisto hidático foi digerido com *AluI* (1), *TaqI* (2), *RsaI* (3), *EcoRI* (4) ou *HindIII* (5), submetido à eletroforese em gel de agarose 0,8%, transferido para membranas de náilon e hibridizado com as seguintes sondas de oligonucleotídeos marcadas com ^{32}P : (A) $(\text{GT})_{15}$; (B) $(\text{CT})_{15}$; (C) $(\text{CAA})_{10}$; (D) $(\text{CAT})_{10}$; (E) $(\text{CGG})_{10}$; (F) $(\text{CATA})_{10}$.

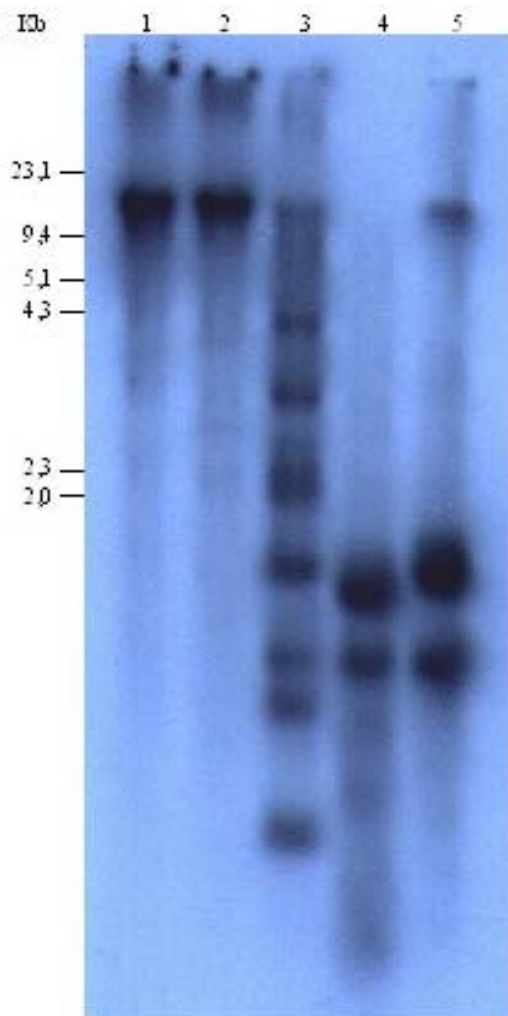


Figura 3.2. Presença do gene snRNA U1 em *E. granulosus*. O DNA total obtido a partir de protoescoléces de um único cisto hidático foi digerido com *Hind*III (1), *Eco*RI (2), *Rsa*I (3), *Taq*I (4) ou *Alu*I (5), submetido à eletroforese em gel de agarose, transferido para membranas de náilon e hibridizado com a sonda snRNA U1.

3.3. ISOLAMENTO DO MICROSSATÉLITE CONTIDO NO GENE snRNA U1

Foram isoladas seis seqüências em *E. granulosus* que apresentaram similaridade com parte da seqüência do gene snRNA U1 descrita para *E. multilocularis* por Bretagne *et al.* (1996), através de comparações feitas utilizando-se o programa BLAST.

As seqüências apresentaram variação no motivo e/ou número de unidades repetitivas do microsatélite de pentanucleotídeos e na seqüência flanqueadora das repetições (Fig 3.3).

Três seqüências apresentaram microsatélites imperfeitos, com o motivo principal GACGA, duas seqüências apresentaram microsatélites compostos (GACGA)(GCGAG) e uma seqüência apresentou um microsatélite composto (GACGA)(GGCGA).

(A) snRNA U1-1

attgtcgttgccatctctccacctcggctgtggcagtggg**GACGAGACGAGACG** : 54
AGACGAGACGAGACGAGGCGAgttagggggaggtactactgaagtggagaagta :108
ggtgggagtagagtgcaggtgggtagtgagatggtggagaggagagggaggtg :162
gtgatgagtggcgtgtgaggagtagggtgcgtgggcgtaggaaagcagaagaga :216
cgtggtggtgtgagggcgatgtgtggtgatggagacg :252

(B) snRNA U1-2

attgtcgttgccatctctccatctctcaccctcggctgtggcggtggg**GACGAG** : 54
ACGAGACGAGACGAGGCAGGCAGGGCAGGGCAGGGCAGGGCAGGGCAGGGCAGGGCAG :108
GCGAGttaggggaggtactactgaagtggagaagtaggtgggagtagagtgcag :162
gtgggtagtgagatggtggagaggagagggaggtggtgatgagtggcgtgtga :216
ggagtagggtgcgtgggcgtaggaaagcagaagagacgtggtggtgtgagggcga :270
tgtgtggtgatggagacg :288

(C) snRNA U1-3

attgtcgttgccatctctccacctcggctgtggcagtggg**GACGAGACGAGACG** : 54
AGACGAGACGAGACGAGACGAGACGAGGCGAgttagggggaggtactactgaag :108
tggagaagtaggtgggagtagagtgcaggtgggtagtgagatggtggagagga :162
gagggaggtggtgatgagtggcgtgtgaggagtagggtgcgtgggcgtaggaaa :216
gcagaagagacgtggtggtgtgagggcgatgtgtggtgatggagacg :262

(D) snRNA U1-4

attgtcgttgccatctctccgcggaccacttcccaccctcgtctgtggcggtggg : 54
GACGAGACGAGACGAGACGAGGCGAGGCGAGGCGAgttagggggaggtactact :108
gaagtggagaagtaggtgggagtagagtgcaggtgggtagtgagatggtggag :162
aggagagggaggtggtgatgagtggcgtgtgaggagtagggtgcgtgggcgtag :216
gaaagcagaagagacgtggtggtgtgagggcgatgtgtggtgatggagacg :266

(E) snRNA U1-5

attgtcgttgccatctctccacccggcgggcggggt**GACGAGACGAGACGAGAC** : 54
GAGACGAGCGAGGCGAGGCGAGGCGAGGCGAGtagggagacacgaggagaagta :108
ggtgggagtagagtgcaggtgggtagtgagatggtggagaggagagggaggtg :162
gtgatgagtggcgtgtgaggagtagggtgcgtgggcgtaggaaagcagaagaga :216
cgtggtggtgtgagggcgatgtgtggtgatggagacg : 252

(F) snRNA U1-6

attgtcgttgccatctctccgccaccccaccggctgtggcggtggg**GACGAGA** : 54
CGAGACGAGACGAGCGAGGCGAGGCGAGGCGAGtcggggagtagtagtactgactga :108
gtggagaagtaggtgggagtagagtgcaggtgggtagtgagatggtggagagg :162
agagggaggtggtgatgagtggcgtgtgaggagtagggtgcgtgggcgtaggaa :216
agcagaagagacgtggtggtgtgagggcgatgtgtggtgatggagacg :263

Figura. 3.3. Sequências de microssatélites com repetições de pentanucleotídeos e de suas regiões flanqueadoras, contidos nos genes do complexo snRNA U1. (A) Sequência com um microssatélite imperfeito (GACGA)₆(GGCGA); (B) Sequência com um microssatélite imperfeito (GACGA)₄(GGCAG)(GCAGG)₂(GCGAG)₃(ACGAG)(GCGAG)₂; (C) Sequência com um microssatélite imperfeito (GACGA)₈(GGCGA); (D) Sequência com um microssatélite composto (GACGA)₄(GGCGA)₃; (E) Sequência com um microssatélite composto (GACGA)₅(GCGAG)₅; (F) Sequência com um microssatélite composto (GACGA)₄(GCGAG)₄. As sequências correspondentes aos microssatélites estão em letras maiúsculas e em negrito e as sequências correspondentes aos primers utilizados na obtenção dos fragmentos clonados estão sublinhadas.

3.4. AMPLIFICAÇÃO DO MICROSSATÉLITE PRESENTE NO GENE snRNA U1

A amplificação do microsatélite presente no gene snRNA U1 gerou vários fragmentos de diferentes tamanhos nos isolados analisados (Fig. 3.4), tornando-o inadequado para estudos populacionais.

3.5. ISOLAMENTO DE SEQÜÊNCIAS CONTENDO REPETIÇÕES CA, GA, AAC E TA

Foram isoladas seis seqüências contendo um microsatélite CA (Fig 3.5). Destas, cinco apresentaram microsatélites perfeitos, com número de unidades repetitivas variando de 9 (locus Egmsca2) a 13 (Egmsca6).

O locus Egmsca4 apresentou um microsatélite imperfeito com cinco unidades repetitivas e uma transversão guanina → citosina na quarta unidade.

Dois locus com microsatélites CA (Egmsca5 e Egmsca6) apresentaram uma das extremidades flanqueadoras com a seqüência inicial do oligo nucleotídeo biotilado utilizado no isolamento das seqüências repetitivas (Fig 3.5.E e F). O locus Egmsca2 apresentou uma das extremidades curta devido à presença de um sítio de clivagem para a enzima *TaqI* próximo ao microsatélite (Fig. 3.5B). Para estes locus não é possível a projeção de *primers*.

Foram projetados *primers* para as seqüências flanqueadoras do locus Egmsca1, os quais têm suas seqüências representadas na figura 3.5.A.

Quatro seqüências contendo repetições GA foram isoladas, três delas (Egmsga2, Egmsga3, Egmsga4) com a seqüência inicial do oligonucleotídeo biotilado utilizado no isolamento dos microsatélites em uma das extremidades flanqueadoras (Fig 3.6.B, C e D). O locus Egmsga2 apresentou o maior número de unidades repetitivas (21) entre todas as seqüências isoladas neste estudo.

O locus Egmsga1, com 13 repetições GA, apresentou ambas as seqüências flanqueadoras longas, para as quais foram projetados *primers* (Fig 3.6.A).

Foram isoladas oito seqüências contendo repetições do tipo AAC/GTT (Fig. 3.7). Destas, sete apresentaram microsatélites perfeitos, com nove (EgmsAAC9), dez

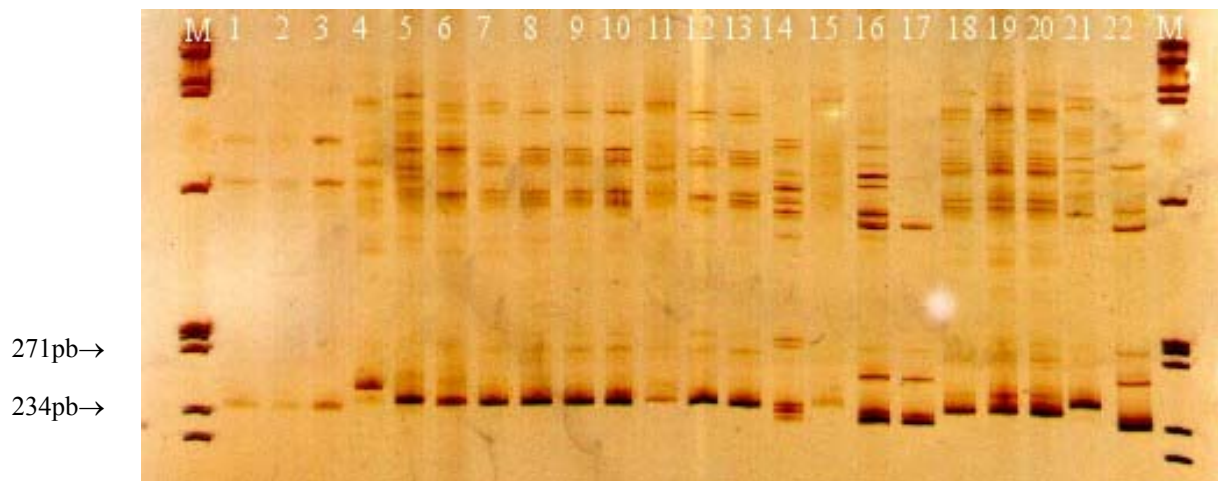


Figura 3.4. Gel de poliacrilamida com produtos de amplificação do microsatélite de pentanucleotídeos presente no complexo gênico snRNA U1 em diferentes isolados de *E. granulosus*. M = Marcador de Peso Molecular (ϕ X 174/*Hae*III); 1 a 22 = amostras de isolados.

(EgmsAAC7), 11 (EgmsAAC2, EgmsAAC13 e EgmsAAC15) e 12 repetições (EgmsAAC4 e EgmsAAC14). Uma seqüência (EgmsAAC5) apresentou um microssatélite interrompido, contendo 11 repetições AAC, uma interrupção de 23 nucleotídeos e mais seis repetições AAC.

Todas as seqüências contendo repetições AAC isoladas apresentaram uma extremidade flanqueadora contendo a seqüência do oligo biotilado utilizado na etapa de captura dos microssatélites e não puderam ser utilizadas para a projeção de *primers*.

Nenhum locus isolado apresentou um microssatélite composto.

Utilizando-se o programa BLAST não foi possível constatar similaridade entre as seqüências de microssatélites CA, GA e AAC isoladas e seqüências depositadas no Genbank.

(A) Egmsca1:
cgaaagtgatgacaaaccaattaataaaggggtacaaa**CACACACACACACAC** : 54
ACACA^{Egmsca1-F}aaacgcacgctggaacagcaaataaacgacctcatctccatcaagcccaat :108
^{Egmsca1-R}
 ttatacagacaactggcaaaaacgataaacttcacaaacactaagtgcctgtga :162
 aattacttacactgaggccaattctgcaaaatatttgaaggaaacactgaaaag :216
 tttegctaaaacgacagttgcaataactaataatatacatattccacctgcctcc :270
 cat : 273

(B) Egmsca2:
 cgaacaccatcacccaacttatcaacgattgcctgagtaacgctcacgacacca : 54
 cctcacaactgccactgctgcctcaattcaaa**CACACACACACACACACA**tcc :108
 act : 111

(C) Egmsca3
 cgatcccgggtcgcggcatgcgtcagtatgtaccgcatatgccacctaccttta : 54
 tgcacgtcactgcgtgcatg**CACACACACACACACACACACA**cggacaagaagc :108
 gaggtgtgcagcaagacgagcgcagtcattgccctgatggagcacaggagtgtg :162
 ccacacagactgctgccacagacagcaagggtgagcatt : 201

(D) Egmsca4
 cgaacgagtgacagacactctgatggttgatagacgtgagtcacaaggtcaccca : 54
 ttcactgtaagtacgcgtgaccagtacggcggagcgacc**CACACAGACA**agctc :108
 cggcaaggaaagcaggaaacgtgcacggcgcattgcttacgaactgtgggtatgt :162
 ctagagaataattataagccttgcttccgcccagcaaaaataaaatatttcatttt :216
 atagacctttcccatatttttctaccgaggcactccatgtggtagtgggcgtcgg :270
 ctttaagggtgtaaaatttacgtctacttgctctactgcgactaccattacgtg :324
 gtaatttggtatgtattttactctctaaaagtttaattcaatttaactttcatt :378
 tacaagtgttaatgaaagatat : 400

(E) Egmsca5
 gatgatccgacgcat**CACACACACACACACACACACA**ccccactcgatcca : 54
 gtaccaaggcttcataatcactagtgaattcgcggccgcctgcaggt : 101

(F) Egmsca6
 gatgatccgacgcat**CACACACACACACACACACACA**tccact : 47

Figura 3.5. Sequências de seis microssatélites CA e de suas regiões flanqueadoras isoladas de *E. granulosus*. (A) Sequência contendo um microssatélite perfeito com dez repetições CA; (B) Sequência contendo um microssatélite perfeito com nove repetições CA; (C) Sequência contendo um microssatélite perfeito com 11 repetições CA; (D) Sequência contendo um microssatélite imperfeito com cinco unidades repetitivas e uma mutação C → G; (E) Sequência contendo um microssatélite perfeito com 12 repetições CA; (F) Sequência contendo um microssatélite perfeito com 13 repetições CA. As sequências correspondentes aos microssatélites estão em letras maiúsculas e em negrito.

(A) Egmsga1

cgactgcttccccgcctcactacacctggtcacacgcaaattgcatgctcat : 54

gctgcgctccatgggtgtgtgtgcacggcttaggaatgggaagggtgacggcgatg :108

atgagatagtgttggagatggcgaggtggtg**GAGAGAGAGAGAGAGAGAGAGAG** :162

AGAggagggcgaggcttgcggatggcaatg**cagtgtagcgtgtggcaaggcagc** :216

cttcgtggtcgcaatcggtttagcgcggttcggctgtg**Egmsga1-R**taactgaaagggttgggtggt :269

(B) Egmsga2:

gatgatccgacgcat**AGA** : 54

GAGcgtgtgactgtttgtgcggt : 76

(C) Egmsga3:

gatgatccgacgcat**AGAGAGAGAGAGAGAGAGAGAGAGAGAGAG**cggaagacagagaa : 54

tactaatgctggtagtaatagtgagaacaataaatggatttgagttgaggaggc :108

gataaggggttgacattggcaattaaagttatgaagtgataaatcggctctccg :162

tttacacgcaaaggaaaaaaagcattaagatgctcaaattaggcactcaatgt :216

gaagtgagtgtgttgaaattctttggtgtgacgggaaacgtgaacttaaaaaac :270

ctctttcctgaccttattagattaagaaatggattacctgagaaaagactgta :324

agccaacttcacaaagtgtaaattttattcattttgatgccact : 367

(D) Egmsga4:

gatgatccgacgcat**AGAGAGAGAGAGAGAGAGAGAGAGAGAGAG**gagaggggagagggaa : 54

agagaggtagagggagaatgggaattgctgagggtggtgatggtttaaaacga :108

atggagactaagctgtggggaagactttcagtcttgccaaatgtaacgagagtg :162

atgagacaattgatgtatgcctaaatcaagcaagtggggaaaggggaagtcgtgga :216

agt : 219

Figura. 3.6. Sequências de quatro microssatélites GA e de suas regiões flanqueadoras isoladas de *E. granulosus*. (A) Sequência contendo um microssatélite perfeito com 13 repetições GA; (B) Sequência contendo um microssatélite perfeito com 21 repetições GA; (C) Sequência contendo um microssatélite perfeito com 12 repetições GA; (D) Sequência contendo um microssatélite perfeito com 12 repetições GA. As sequências correspondentes aos microssatélites estão em letras maiúsculas e em negrito.

3.6. ESTABELECIMENTO DAS CONDIÇÕES DE AMPLIFICAÇÃO PARA OS LÓCUS *Egmsca1* E *Egmsga1*

Entre as condições de amplificação testadas para o locus *Egmsca1*, a que melhor resultado apresentou (menor número de bandas inespecíficas amplificado), foi aquela na qual se utilizou o programa TD70/60 e a seguinte condição: 20 pmoles de cada *primer* (*Egmsca1-F* e *Egmsca1-R*), 100 µM cada dNTP, 2,5 U *Taq* DNA Polimerase, tampão de amplificação 1x (10 mM Tris-HCl, 50mM KCl) e 1,5 mM MgCl₂. Essa condição foi utilizada para o estudo populacional.

Em todas as condições de amplificação testadas para o locus *Egmsga1* foram geradas várias bandas de amplificação (“bandas espelho”), provavelmente em decorrência de erros introduzidos pela *Taq* DNA Polimerase (adição ou deleção de unidades repetitivas) ou ainda, porque o locus *Egmsga1* pode estar situado dentro de uma região repetitiva maior, o que levaria à amplificação de diversos fragmentos parálogos. Como não foi possível definir uma condição adequada para a amplificação deste locus e posterior interpretação dos resultados, o locus *Egmsga1* não foi utilizado como marcador molecular em estudos populacionais de *E. granulosus*.

3.7. ANÁLISE DE POLIMORFISMO NO LÓCUS *Egmsca1*

Não foi observado polimorfismo no locus *Egmsca1* ao se analisar 150 protoescoléces provenientes de um único cisto hidático e 20 protoescoléces de cada um de 5 cistos adicionais. Todos os indivíduos analisados apresentaram genótipo homozigoto CA₇/CA₇ (um cisto), CA₈/CA₈ (um cisto) e CA₁₀/CA₁₀ (quatro cistos).

Entretanto, observou-se polimorfismo quando protoescoléces provenientes de diferentes isolados foram analisados. Na tabela 3.1 podemos observar os genótipos encontrados para cada isolado, bem como a origem geográfica, linhagem a qual pertence o parasita e hospedeiro intermediário do qual foi retirado cada cisto. Os dados estão sintetizados na tabela 3.2.

Tabela 3.1. Procedência, hospedeiro intermediário, linhagem e genótipos do locus Egmsca1 de cada isolado analisado.

Isolado	Procedência	Hospedeiro intermediário	Linhagem	Genótipo
HB4	Argentina	humano	ovelha	CA10/CA10
HN7	Argentina	humano	ovelha	CA10/CA10
HN8	Argentina	humano	ovelha	CA10/CA10
HN12	Argentina	humano	ovelha	CA10/CA10
HN16	Argentina	humano	ovelha	CA10/CA10
HR2	Argentina	humano	ovelha	CA10/CA10
HR6	Argentina	humano	ovelha	CA10/CA10
HR7	Argentina	humano	ovelha	CA8/CA8
HR8	Argentina	humano	ovelha	CA10/CA10
HR22	Argentina	humano	ovelha	CA10/CA10
HT8	Argentina	humano	ovelha	CA10/CA10
HT10	Argentina	humano	ovelha	CA10/CA10
HSC4	Argentina	humano	ovelha	CA10/CA10
HSC5	Argentina	humano	ovelha	CA10/CA10
OSF1	Argentina	Ovino	ovelha	CA10/CA10
OCH0	Argentina	Ovino	ovelha	CA10/CA10
OCH1	Argentina	ovino	ovelha	CA10/CA10
HT4	Argentina	humano	Ovelha Tasm.	CA10/CA10
HT5	Argentina	humano	Ovelha Tasm.	CA8/CA10
HT6	Argentina	humano	Ovelha Tasm.	CA10/CA10
HT7	Argentina	humano	Ovelha Tasm.	CA10/CA10
HT9P	Argentina	humano	Ovelha Tasm.	CA8/CA10
HN4	Argentina	humano	camelo	CA7/CA7
HN15	Argentina	humano	camelo	CA7/CA7
HN17	Argentina	humano	camelo	CA7/CA7

Tabela 3.1. Continuação

MHN23	Argentina	humano	camelo	CA7/CA7
MHN24	Argentina	humano	camelo	CA7/CA7
HR3	Argentina	humano	camelo	CA7/CA7
HR4	Argentina	humano	camelo	CA7/CA7
7	Brasil	bovino	Ovelha	CA10/CA10
8b	Brasil	bovino	Ovelha	CA10/CA10
3	Brasil	bovino	Ovelha	CA10/CA10
4c	Brasil	bovino	Ovelha	CA10/CA10
6	Brasil	bovino	Ovelha	CA10/CA11
17	Brasil	bovino	Ovelha	CA10/CA10
20	Brasil	bovino	Ovelha	CA10/CA10
16a	Brasil	bovino	Ovelha	CA11/CA11
27d	Brasil	bovino	Ovelha	CA10/CA10
36	Brasil	ovino	Ovelha	CA10/CA10
19a	Brasil	ovino	Ovelha	CA10/CA10
9b	Brasil	ovino	Ovelha	CA10/CA10
10a	Brasil	ovino	Ovelha	CA10/CA10
11b	Brasil	ovino	Ovelha	CA10/CA10
12a	Brasil	ovino	Ovelha	CA10/CA10
13	Brasil	ovino	Ovelha	CA10/CA10
34	Brasil	ovino	Ovelha	CA8/CA10
35	Brasil	ovino	Ovelha	CA8/CA10
37	Brasil	ovino	Ovelha	CA8/CA10
38MG	Brasil	ovino	Ovelha	CA10/CA10
39MG	Brasil	ovino	Ovelha	CA10/CA10
40MG	Brasil	ovino	Ovelha	CA10/CA10
41MG	Brasil	ovino	Ovelha	CA10/CA10
42MG	Brasil	ovino	Ovelha	CA10/CA10
44MG	Brasil	ovino	Ovelha	CA10/CA10

Tabela 3.1. Continuação

45MG	Brasil	ovino	Ovelha	CA10/CA10
46MG	Brasil	ovino	Ovelha	CA10/CA10
47MG	Brasil	ovino	Ovelha	CA10/CA10
48MG	Brasil	ovino	Ovelha	CA10/CA10
49MG	Brasil	ovino	Ovelha	CA10/CA10
50MG	Brasil	ovino	Ovelha	CA10/CA10
51MG	Brasil	ovino	Ovelha	CA10/CA10
52MG	Brasil	ovino	Ovelha	CA10/CA10
53MG	Brasil	ovino	Ovelha	CA10/CA10
54MG	Brasil	ovino	Ovelha	CA10/CA10
55MG	Brasil	ovino	Ovelha	CA10/CA10
56MG	Brasil	ovino	Ovelha	CA10/CA10
57MG	Brasil	ovino	Ovelha	CA10/CA10
1	Brasil	bovino	Boi	CA7/CA7
5b	Brasil	bovino	Boi	CA7/CA7
21b	Brasil	bovino	Boi	CA7/CA7
25	Brasil	bovino	Boi	CA7/CA7
14a	Brasil	bovino	Boi	CA7/CA7
15	Brasil	bovino	Boi	CA7/CA7
M5/1	Etiópia	bovino	ND*	CA7/CA7
M14/1	Etiópia	bovino	ND	CA7/CA7
M16/1	Etiópia	bovino	ND	CA7/CA7
M33/1	Etiópia	bovino	ND	CA7/CA7
M37/1	Etiópia	bovino	ND	CA7/CA7
M46/1	Etiópia	bovino	ND	CA7/CA7
M58/1	Etiópia	bovino	ND	CA7/CA7
M63/1	Etiópia	bovino	ND	CA7/CA7
M83/1	Etiópia	bovino	ND	CA7/CA7
M32/2	Etiópia	bovino	ND	CA10/CA10

Tabela 3.1. Continuação

M39/7	Etiópia	bovino	ND	CA10/CA10
M66/1	Etiópia	bovino	ND	CA10/CA10
M68/1	Etiópia	bovino	ND	CA10/CA10
M80/1	Etiópia	bovino	ND	CA10/CA10
A52/1	Etiópia	bovino	ND	CA10/CA10
A74/1	Etiópia	bovino	ND	CA10/CA10
A86/1	Etiópia	bovino	ND	CA10/CA10
A83/3	Etiópia	bovino	ND	CA10/CA10
A97/7	Etiópia	bovino	ND	CA10/CA10
M49/1	Etiópia	bovino	ND	CA8/CA8
M50/1	Etiópia	bovino	ND	CA9/CA10
A75/7	Etiópia	bovino	ND	CA8/CA8
S21/1	Etiópia	ovino	ND	CA8/CA10
S24/3	Etiópia	ovino	ND	CA10/CA10
S29/2	Etiópia	ovino	ND	CA10/CA10
S34/1	Etiópia	ovino	ND	CA10/CA10
S35/1	Etiópia	ovino	ND	CA8/CA8

*ND = não determinada

Tabela 3.2. Número de isolados analisados de acordo com a origem, linhagem e hospedeiro intermediário do qual o cisto foi removido.

Origem dos isolados	Linhagem	Hospedeiro Intermediário	Número
Brasil	Ovelha	Bovino, Ovino	38
Brasil	Boi	Bovino	6
Argentina	Ovelha	Humanos, Ovino	17
Argentina	Ovelha da Tasmânia	Humanos	5
Argentina	Camelo	Humanos	7
Etiópia	Desconhecida	Bovino, Ovino	27

Cinco alelos foram encontrados no locus Egmsca1 entre os 100 isolados analisados: (CA)₇, (CA)₈, (CA)₉, (CA)₁₀ e (CA)₁₁ (Figura 3.8). As frequências alélicas para cada população analisada estão apresentadas na tabela 3.3.

Tabela 3.3. Frequência alélica no locus Egmsca1 para cada população.

População - linhagem	Alelos				
	(CA) ₇	(CA) ₈	(CA) ₉	(CA) ₁₀	(CA) ₁₁
Brasil – ovelha	0,000	0,039	0,000	0,922	0,039
Brasil – boi	1,000	0,000	0,000	0,000	0,000
Argentina – ovelha	0,000	0,059	0,000	0,941	0,000
Argentina – ovelha da Tasmânia	0,000	0,200	0,000	0,800	0,000
Argentina – camelo	1,000	0,000	0,000	0,000	0,000
Etiópia – desconhecida	0,333	0,130	0,019	0,519	0,000

O alelo (CA)₇ foi compartilhado por isolados das linhagens do boi e do camelo. As linhagens da ovelha e da ovelha da Tasmânia compartilharam os alelos (CA)₈ e (CA)₁₀. A linhagem da ovelha também apresentou o alelo (CA)₁₁, o qual ocorreu em uma frequência muito baixa, somente em isolados provenientes do Brasil.

O alelo (CA)₉ foi encontrado em heterozigose com o alelo (CA)₁₀ em um único isolado proveniente da África e de linhagem desconhecida.

As frequências genotípicas no locus Egmsca1 apresentaram desvios em relação ao equilíbrio de Hardy-Weinberg na população argentina da linhagem da ovelha, a qual apresentou deficiência de heterozigotos ($P < 0,05$). A população brasileira da linhagem da ovelha estava de acordo com as expectativas de Hardy-Weinberg e não apresentou deficiência significativa de heterozigotos ($P > 0,05$). As populações da linhagem do camelo e da ovelha da Tasmânia, ambas provenientes da Argentina, não foram testadas em relação ao equilíbrio devido ao pequeno número de isolados analisados ($n < 10$). Os isolados provenientes da África também não foram testados quanto ao equilíbrio em função das linhagens dessa população ainda não terem sido identificadas.

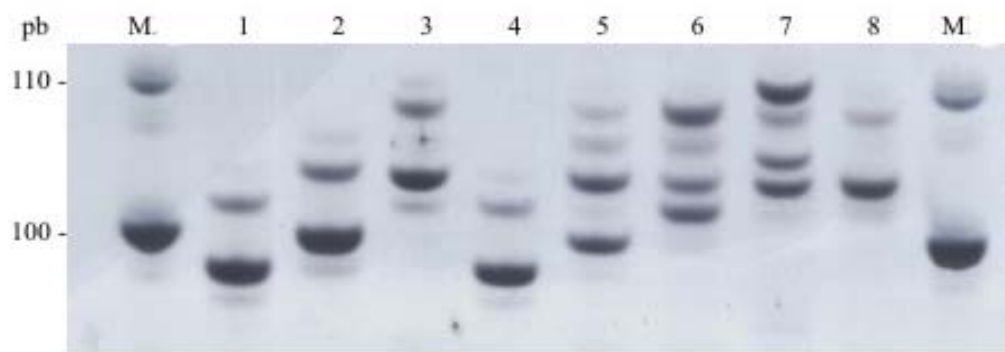


Figura 3.8. Diferentes genótipos encontrados para o locus Egmsca1. M = marcador de peso molecular; 1 = (CA)₇/(CA)₇; 2 = (CA)₈/(CA)₈; 3 = (CA)₁₀/(CA)₁₀; 4 = (CA)₇/(CA)₇; 5 = (CA)₈/(CA)₁₀; 6 = (CA)₉/(CA)₁₀; 7 = (CA)₁₀/(CA)₁₁; 8 = (CA)₁₀/(CA)₁₀.

A tabela 3.4 apresenta as frequências genóticas e a heterozigosidade observadas para cada população.

Não se observou diferenciação gênica e genotípica entre as populações brasileira e argentina da linhagem da ovelha ($P > 0,05$, para ambos os testes). Para estas populações, os valores encontrados de F_{is} e F_{st} foram 0,48 e $-0,02$, respectivamente.

Tabela 3.4. Frequências genótípicas e heterozigosidades para o locus Egmsca1 em cada população analisada.

População*	Frequências genótípicas							Hobs**	Hesp***
	(CA) ₇ /(CA) ₇	(CA) ₈ /(CA) ₈	(CA) ₁₀ /(CA) ₁₀	(CA) ₁₁ /(CA) ₁₁	(CA) ₈ /(CA) ₁₀	(CA) ₉ /(CA) ₁₀	(CA) ₁₀ /(CA) ₁₁		
LO – B	0	0	0.868	0.026	0.080	0	0.026	0.11	0.15
LB – B	1.000	0	0	0	0	0	0	0	-
LO – A	0	0.059	0.941	0	0	0	0	0	0.11
LOT – A	0	0	0.600	0	0.400	0	0	0.40	0.36
LC – A	1.000	0	0	0	0	0	0	0	-
LD – E				0			0	-	-

* LO – B = linhagem da ovelha do Brasil; LB – B = linhagem do boi do Brasil; LO – A = linhagem da ovelha da Argentina; LOT – A = linhagem da ovelha da Tasmânia da Argentina; LC – A = linhagem do camelo da Argentina; LD – E = linhagem desconhecida da Etiópia.

** Hobs = heterozigosidade observada.

*** Hesp = heterozigosidade esperada (calculada para as populações não-monomórficas de linhagem conhecida)

4. DISCUSSÃO

O isolamento de microssatélites e suas seqüências flanqueadoras a partir de genomas é freqüentemente uma tarefa cara e trabalhosa, com exceção dos organismos cujos genomas foram completamente seqüenciados ou que possuam um grande número de seqüências depositadas em bancos de dados. Nestes casos, programas computacionais, como o *Tandem Repeats Finder* (Benson, 1999), podem ser utilizados para a localização dos microssatélites.

Várias técnicas têm sido empregadas para o isolamento ou para a construção de bibliotecas enriquecidas em microssatélites, entre elas, hibridização de oligonucleotídeos repetitivos marcados radioativamente com bancos genômicos (Naom *et al.*, 1995), fixação de oligonucleotídeos repetitivos em membrana de náilon e hibridização com fragmentos de DNA genômico gerados por clivagem com endonuclease de restrição e ligados a oligonucleotídeo adaptador, seguido de amplificação por PCR dos fragmentos hibridizados (Edwards *et al.*, 1996), hibridização de oligonucleotídeos repetitivos biotinizados com fragmentos de DNA genômico ligado a adaptador e posterior captura das seqüências contendo microssatélites utilizando-se esferas magnéticas recobertas por estreptavidina seguida de amplificação por PCR dos fragmentos capturados (Refseth *et al.*, 1997). Entretanto, o isolamento dos microssatélites em geral é feito sem um conhecimento prévio da presença e da abundância da repetição a ser isolada.

A análise da abundância de microssatélites por intermédio de experimentos de *Southern blot* utilizando oligonucleotídeos repetitivos como sondas pode fornecer informações importantes antes do isolamento da repetição.

Neste estudo foram utilizadas oito sondas diferentes de oligonucleotídeos repetitivos marcados radioativamente com ^{32}P para hibridização com DNA de *E. granulosus*. As sondas (GT)₁₅, (CAA)₁₀, (CATA)₁₀ e (CT)₁₅ apresentaram o maior número de bandas hibridizadas em *E. granulosus*. As repetições GT (ou CA) têm sido relatadas como sendo os motivos de microssatélites mais freqüentes em vários genomas animais (Tautz, 1989; Weber & May, 1989; Stallings *et al.*, 1991; Katti *et al.*, 2001). Os microssatélites AAC estão entre as repetições de trinucleotídeos mais freqüentes nos genomas humano, de *Drosophila* e de levedura, e os microssatélites CATA são as repetições de tetranucleotídeos mais freqüentes em *Drosophila* (Katti *et al.*, 2001).

Possivelmente as repetições AT e GC não ocorrem no genoma de *E. granulosus*, ou ocorrem em frequências muito baixas, pois não foram encontrados sinais de hibridização para estas sondas. Analisando-se seqüências depositadas no Genbank, a ausência ou raridade de repetições GC foi relatada nos genomas humano, de *C. elegans*, *Drosophila melanogaster*, *A. thaliana* e de levedura (Dokholyan *et al.*, 2000; Katti *et al.*, 2001). Frequências baixas de dinucleotídeos CpG em genomas de vertebrados têm sido atribuídas à metilação da citosina, a qual aumentaria a chance de mutação para timina por desaminação (Schorderet & Garder, 1992). Entretanto, a supressão de CpG por este mecanismo não poderia explicar a raridade de microssatélites (GC)_n em leveduras, *C. elegans* e *Drosophila*, visto que não apresentam metilação da citosina (Katti *et al.*, 2001). Outra explicação para a falta de repetições GC poderia ser a sua habilidade de formar Z-DNA nas regiões supertorcidas negativamente durante a transcrição (Dokholyan *et al.*, 2000).

Por outro lado, as repetições AT, que não foram encontradas em *E. granulosus*, são freqüentes nos genomas de diversos organismos, como em humanos, ratos e leveduras (Katti *et al.*, 2001). Em plantas as repetições AT são as mais comuns (Lagercrantz *et al.*, 1993).

A presença de microssatélites CA e GA no genoma de *E. granulosus* foi confirmada pela clonagem e seqüenciamento dessas repetições. Todas as tentativas de isolar seqüências repetitivas AT falharam, confirmando o resultado negativo obtido nos experimentos de *Southern blot*. Deve-se ressaltar que as tentativas de isolamento de microssatélites AT foram realizadas, a despeito da ausência de sinal de hibridização com a sonda (AT)₁₅, em virtude da baixa T_m deste oligonucleotídeo, o que poderia acarretar em perda do sinal de hibridização durante o processo de lavagem da membrana. Entretanto, a falha no isolamento de seqüências AT corrobora o resultado obtido no experimento de *Southern blot* e justifica o seu emprego antes de se partir para o isolamento de microssatélites.

A análise de microssatélites por meio de *Southern blot* genômico mostrou-se extremamente útil para estudar a distribuição e abundância destes marcadores moleculares em *E. granulosus*. As informações obtidas permitem que os esforços sejam direcionados para o isolamento de microssatélites freqüentes no genoma, aumentando as chances de que lócus polimórficos sejam isolados.

Neste estudo, foram isolados seis lócus com repetições CA, quatro com repetições GA e oito com repetições AAC do genoma de *E. granulosus*. Entretanto, dois lócus CA, três lócus GA e todos os oito lócus AAC apresentaram apenas uma extremidade na qual pode ser

identificado o oligonucleotídeo adaptador. A outra extremidade desses lócus apresentou a seqüência correspondente aos primeiros nucleotídeos do oligonucleotídeo biotilado utilizado no processo de captura dos microssatélites. A ocorrência deste artefato da técnica também foi relatada por Refseth *et al.* (1997). Provavelmente, o oligonucleotídeo biotilado age como um *primer* interno durante a PCR, quando liberado das esferas magnéticas recobertas por estreptavidina durante a captura. O conhecimento de apenas uma das extremidades flanqueadoras não permite a projeção de *primers* para os lócus em questão, os quais não poderão ser utilizados como marcadores moleculares em *E. granulosus*.

Além disso, um lócus com repetições CA (Egmsca2) apresentou uma de suas extremidades curta para a projeção de um *primer*, em função da existência de um sítio para a endonuclease *TaqI* (utilizada na clivagem do DNA genômico para posterior captura dos microssatélites) muito próximo ao microssatélite. Desta forma, apenas quatro lócus (Egmsca1, Egmsca3, Egmsca4 e Egmsga1) apresentaram extremidades flanqueadoras apropriadas para a projeção de *primers*.

O lócus Egmsca4 apresentou um microssatélite CA imperfeito com cinco unidades repetitivas e com uma mutação C → G na quarta unidade. Possivelmente, este lócus não é polimórfico, tanto devido ao pequeno número de unidades repetitivas, quanto devido ao fato de ser imperfeito. O polimorfismo de marcadores CA perfeitos aumenta com o aumento do número de unidades repetitivas (Weber, 1992; Beckmann & Weber, 1992), presumivelmente porque uma configuração estável das fitas do DNA alinhadas de forma errada durante a replicação é maior para segmentos repetitivos mais longos. Além disso, microssatélites imperfeitos são mais estáveis que os perfeitos. Bases que interrompem um microssatélite podem estabilizar o segmento repetitivo, reduzindo a probabilidade de alinhamento errôneo e resultando em níveis menores de polimorfismo quando comparados a microssatélites perfeitos com um número similar de repetições (Estoup & Cornuet, 1999).

Com exceção do lócus Egmsca4, o qual apresenta um microssatélite imperfeito e do lócus EgmsAAC5, o qual apresenta uma interrupção de 23 nucleotídeos, todos os outros lócus isolados no presente estudo apresentaram microssatélites perfeitos. Interessantemente, 72 lócus de microssatélites CA isolados do genoma do parasita nematódeo *H. contortus* apresentaram microssatélites imperfeitos (Hoekstra *et al.*, 1997, Otsen *et al.*, 2000).

O número mínimo de unidades repetitivas encontrado nas seqüências clonadas foi cinco para o lócus Egmsca4 contendo um microssatélite imperfeito e nove para os lócus Egmsca2 e

EgmsAAC9, os quais contêm microssatélites perfeitos. O número máximo de repetições encontrado foi 21 no locus Egmsga2.

Amos (1999) fez duas considerações a respeito da distribuição do tamanho de locus de microssatélites entre diferentes *taxa*. Primeiramente, ele considera que seria de se esperar que populações grandes carreguem níveis maiores de variabilidade neutra do que populações equivalentes pequenas. Assim, a história da “vida” de um microssatélite iniciaria mais cedo e ocorreria mais rapidamente em populações grandes, implicando em um tempo médio de “vida” menor e menos tempo gasto em comprimentos maiores. Como a taxa de origem de novos microssatélites é independente do tamanho da população, uma longevidade menor dos microssatélites em populações maiores prediz que populações grandes carregarão menos microssatélites por quilobase e que o comprimento médio destes microssatélites será menor.

A segunda consideração diz respeito à temperatura corporal. Os microssatélites provavelmente sofrem mutação por deslizamento molecular, um processo que poderia aumentar em taxa com o aumento da temperatura corporal e assim diminuir o comprimento do ciclo de vida dos microssatélites. Espécies com temperatura corporal mais alta poderiam apresentar um limiar menor de estabilidade e comprimento. Combinando os efeitos de tamanho de população e temperatura corporal poderia ser predito que aves e mamíferos deveriam ter menos e menores microssatélites que peixes, e que invertebrados, como os insetos, os quais têm tamanhos populacionais grandes, apresentariam poucos microssatélites e muito curtos.

Ainda segundo Amos (1999), estudos publicados parecem confirmar essas previsões. Assim, pode-se dizer que os locus de microssatélites são mais frequentes e maiores em vertebrados do que em invertebrados e que entre os vertebrados, espécies de sangue frio como tartarugas e peixes apresentam arranjos repetitivos maiores. Entretanto, o autor adverte que embora estas observações sejam interessantes, mais estudos são necessários, pois muitos dados disponíveis podem estar sujeitos a tendências de publicação. Primeiramente, resultados negativos são difíceis de serem publicados e assim, espécies com microssatélites variáveis, longos e abundantes ficariam super-representadas. Segundo, uma preferência geral por microssatélites longos tenderá a comprimir a amplitude dos comprimentos relatados. Terceiro, falhas em encontrar repetições de dinucleotídeos apropriadas, podem levar ao isolamento de repetições de tri ou tetra nucleotídeos, as quais em geral têm comprimentos menores. Assim, em estudos comparativos entre diferentes espécies, considerando-se microssatélites diferentes,

pode-se chegar a conclusões que favoreçam um maior comprimento dos microssatélites nas espécies onde a proporção de repetições de dinucleotídeos estudada é maior.

A partir das considerações feitas por Amos (1999), esperaria-se encontrar microssatélites curtos em *E. granulosus*. De fato, no presente estudo o número máximo de repetições encontrado foi 21, mas não é possível atribuir o pequeno tamanho do locus ao fato de nosso organismo de estudo ser um invertebrado e nem ao seu tamanho populacional, pois poucos locus foram isolados. Além disso, a técnica utilizada para o isolamento dos microssatélites pode ter favorecido o isolamento de seqüências curtas.

O isolamento dos microssatélites foi feito clivando-se o DNA genômico de *E. granulosus* com endonuclease de restrição e ligando-se os fragmentos a oligonucleotídeos adaptadores. Os fragmentos eram desnaturados e misturados a um oligonucleotídeo repetitivo biotilado, permitindo a hibridização dos fragmentos contendo repetições complementares ao oligonucleotídeo, os quais eram coletados com esferas magnéticas recobertas por estreptavidina, proteína que apresenta alta afinidade pela molécula de biotina. Os fragmentos assim capturados eram amplificados por PCR, para posterior clonagem. É possível que, na etapa de amplificação, os fragmentos menores fossem preferencialmente amplificados e dessa forma, microssatélites longos, os quais necessariamente estariam em fragmentos de maior tamanho, não seriam detectados.

Dokholyan *et al.* (2000) verificaram (analisando seqüências depositadas no Genbank) que o número máximo de repetições de dinucleotídeos nos genomas humano e de camundongo é 40, enquanto que em *C. elegans* pode chegar a 100, inversamente à predição de Amos (1999). Katti *et al.* (2001) encontraram locus de microssatélites de di, tri e tetranucleotídeos com mais de 40 unidades de repetição nos genomas humano, de levedura, de *Arabidopsis*, de *Drosophila* e de *C. elegans*; entretanto constataram que a freqüência de microssatélites decresce exponencialmente com o aumento do comprimento da repetição.

Nenhum dos locus com repetições de dinucleotídeos isolados neste estudo apresentou um microssatélite composto, provavelmente em função do pequeno número de locus isolados. Microssatélites compostos, como $(CA)_n(GA)_n$ são relativamente comuns (Weber, 1990) e apresentam uma característica interessante, a homoplasia de alelos.

Originalmente, homoplasia foi um conceito utilizado por evolucionistas para caracterizar o fato de que um certo caráter presente em duas espécies não é derivado do mesmo caráter em uma espécie ancestral comum, mas que a similaridade é devida a fatores

como convergência, paralelismo ou conversão. Transpondo para o nível gênico, dois alelos são homoplásicos quando são idênticos em estado (têm o mesmo tamanho), embora não idênticos por descendência (não têm a mesma estrutura e história genética comum) (Estoup & Cornuet, 1999).

Como a variação entre alelos de microssatélites é revelada por eletroforese de produtos de PCR, microssatélites compostos podem ter alelos com diferentes números de cada uma das repetições, mas com o mesmo número de unidades repetitivas no total e dessa forma os diferentes alelos eletromorfos não são detectados.

A caracterização dos eletromorfos por seqüenciamento pode produzir resultados muito interessantes. Por exemplo, um estudo com microssatélites em uma espécie de leão-marinho revelou dois locos compostos. Um dos locos apresentou os alelos $(CT)_9 + (CA)_{18}$, $(CT)_{11} + (CA)_{16}$ e $(CT)_8 + (CA)_{19}$, os três com arranjos de 27 repetições. O outro locus apresentou os alelos $(CT)_{10} + (CA)_{18}$ e $(CT)_9 + (CA)_{19}$, representando 28 unidades repetitivas (Ashton, 1999 *apud* Chambers & MacAvoy, 2000).

Em outro estudo, foi encontrado um microssatélite composto $(CA)_n(GA)_n$ localizado na região HLA (antígenos leucocitários humanos) de classe II, aproximadamente 12 kb centromérico ao gene *HLA-DQA1* e 1-2 kb telomérico ao gene *HLA-DQB1* (Macaubas *et al.*, 1997). Os alelos associados ao subtipo não-*DQI* apresentam o tetrâmero GCAA entre os dois motivos repetitivos, mas os alelos associados ao subtipo *DQI* apresentam a repetição GA imediatamente após a CA. À medida que o número de repetições CA aumenta nos alelos associados a *DQI*, o número de unidades GA diminui, mantendo constante o número total de unidades repetitivas. Os números de unidades repetitivas encontrados nos alelos associados a *DQI* foram: $(CA)_8 + (GA)_4$, $(CA)_9 + (GA)_3$, $(CA)_{10} + (GA)_2$, totalizando 12 repetições e $(CA)_{11} + (CA)_3$ e $(CA)_{12} + (GA)_2$, totalizando 14 repetições. Nos alelos associados ao subtipo não-*DQI*, interrompidos por quatro nucleotídeos que separam os dois motivos repetitivos, a variabilidade ocorre somente no motivo CA, sendo o motivo GA monomórfico com cinco repetições. Também foram observadas substituições nucleotídicas nas seqüências flanqueadoras dos microssatélites.

Bell & Jurka (1997) sugerem que até um terço das repetições longas GA e um décimo das repetições longas AT podem estar associadas a repetições longas AC. Analisando muitos locos de microssatélites AC em humanos e roedores, os autores observaram que há um único dímero alternativo nos microssatélites compostos e que esse dímero difere de AC por um

único par de bases. Este fato sugere que o processo que inicia um microssatélite composto pode ser uma substituição de base ou possivelmente a existência do dímero alternativo em uma extremidade da repetição AC, seguida por múltiplos deslizamentos das fitas do DNA e correção do heteroduplex. Também foi observado que, curiosamente, o dímero alternativo em geral difere de AC por uma transversão.

Os microssatélites são seqüências que apresentam alta freqüência de mutação, devido ao deslizamento de uma das fitas do DNA durante a replicação, favorecido pela natureza repetitiva da seqüência, sendo que as repetições de dinucleotídeos são as mais sujeitas a alterações no número de repetições ocasionadas por este processo (Eisen, 1999). Assim, seria possível a ocorrência de mutações em protoescóceles de *E. granulosus* de um único cisto hidático gerados por reprodução assexuada.

Durante o desenvolvimento da fase larval (ou metacestóide) de *E. granulosus* no hospedeiro intermediário, há a formação do cisto hidático, o qual apresenta três camadas: a camada adventícia, mais externa, secretada pelo hospedeiro; a camada laminar acelular e a camada germinativa, mais interna. Esta última prolifera assexuadamente, formando as cápsulas prolíferas, dentro das quais o processo de reprodução assexuada se repete, levando à produção de inúmeros protoescóceles (Thompson *et al.*, 1995). Um único cisto hidático, considerado como um isolado, pode conter centenas a milhares de protoescóceles, que ao serem ingeridos pelo hospedeiro definitivo, se desenvolvem nos vermes adultos.

A estabilidade do locus Egmsca1, contendo um microssatélite com repetição de dinucleotídeos CA, foi avaliada em 150 protoescóceles provenientes de um único cisto retirado de hospedeiro bovino, além de 20 protoescóceles de 5 cistos adicionais, por meio de amplificação por PCR e eletroforese em gel de poliacrilamida. Todos os indivíduos analisados de um mesmo cisto apresentaram o mesmo padrão de migração eletroforética, correspondente genótipos homozigotos.

Apesar da geração de inúmeros protoescóceles por reprodução assexuada aumentarem as chances de se detectar alguma mutação, não foi encontrada instabilidade no locus Egmsca1, ao se analisar amostras de 150 indivíduos (um cisto) ou 20 indivíduos (5 cistos) tomados ao acaso. Estes resultados validam a utilização de protoescóceles de um mesmo cisto, agrupados como um isolado, em estudos populacionais do parasita, utilizando microssatélites como marcadores moleculares.

O locus Egmsca1 mostrou-se adequado para uso como marcador molecular em *E. granulosus*, pois além de estável em protoescoléces gerados por reprodução assexuada, confirmando um padrão de herança mendeliana para o microsatélite, este locus apresentou-se polimórfico quando 100 isolados diferentes foram analisados.

Os isolados foram agrupados de acordo com a linhagem a qual pertencem (caracterizada em estudos prévios, por outros autores) e com a origem geográfica de cada um, com exceção de isolados africanos provenientes da Etiópia, os quais ainda não tiveram suas linhagens caracterizadas. Desta forma, foram estabelecidas seis populações: 1) isolados da linhagem da ovelha, provenientes do Brasil (RS); 2) isolados da linhagem do boi, provenientes do Brasil (RS); 3) isolados da linhagem da ovelha, provenientes da Argentina; 4) isolados da linhagem da ovelha da Tasmânia, provenientes da Argentina; 5) isolados da linhagem do camelo, provenientes da Argentina; 6) isolados de linhagens indeterminadas, provenientes da Etiópia.

Atualmente, o gênero *Echinococcus* apresenta quatro espécies taxonomicamente válidas, *E. granulosus*, *E. multilocularis*, *E. vogeli* e *E. oligarthrus*. Entretanto, um total de 16 espécies já foi descrito com base em características morfológicas. As variações intra-específicas ainda apresentam categoria taxonômica incerta (Kumaratilake & Thompson, 1982; Thompson, 1995).

As linhagens de *E. granulosus* foram descritas com base em características morfológicas, bioquímicas e epidemiológicas. Bowles *et al.* (1992) identificaram sete genótipos mitocondriais, denominados de G1 a G7, correspondentes às linhagens da ovelha (G1), da ovelha da Tasmânia (G2), do búfalo (G3), do cavalo (G4), do boi (G5), do camelo (G6) e do porco (G7). Foi sugerido que o genótipo G3 corresponde a uma variação do genótipo G1 da linhagem da ovelha ou a um grupo geneticamente distinto, mas intimamente relacionado. Posteriormente foram descritos os genótipos G8, correspondente à linhagem dos cervídeos (Bowles *et al.*, 1995) e G9, encontrado em protoescoléces retirados de cistos de pacientes humanos da Polônia (Scott *et al.*, 1997). Este último pode ser uma variante da linhagem do porco (G7). Assim, com exceção dos genótipos G3 e G9, os outros sete têm sido utilizados para a caracterização de linhagens de *E. granulosus*.

A técnica de SSCP também tem sido utilizada com sucesso na caracterização de algumas linhagens de *E. granulosus*, por meio da análise de seqüências codificantes e não-codificantes (Haag *et al.*, 1999). Os isolados analisados no presente estudo foram caracterizados previamente utilizando-se uma das duas metodologias expostas acima.

A análise do locus *Egmsca1* em isolados de diferentes linhagens permitiu que resultados interessantes fossem observados.

Foram encontrados cinco alelos no locus *Egmsca1*, diferindo entre si por uma única unidade repetitiva: CA_7 , CA_8 , CA_9 , CA_{10} e CA_{11} . O alelo $(CA)_9$ foi encontrado somente em um isolado da Etiópia (linhagem desconhecida), em heterozigose com o alelo $(CA)_{10}$.

Este padrão mutacional se encaixa em um modelo de pequenas alterações no número de unidades repetitivas via um mecanismo de deslizamento das fitas do DNA durante a replicação, como o modelo de mutação em passos (Kimura & Otah, 1978).

Vários estudos apontam para um modelo de mutação no qual uma unidade repetitiva é adicionada ou perdida em cada evento mutacional (Macaúbas *et al.*, 1997; Bell & Jurka, 1997; Estoup & Cornuet, 1999). Entretanto, em um trabalho realizado com uma população humana da Sardenha, que supostamente evoluiu em considerável isolamento, utilizando dez locus *CA*, Di Rienzo *et al.* (1994) concluíram que para oito dos dez locus analisados o modelo de mutação em um passo pode ser rejeitado em favor do modelo de duas fases, no qual mutações introduzem um ganho ou perda de x repetições (quando $x = 1$, tem-se o modelo de mutação em um passo).

Muitos modelos têm sido propostos para explicar a evolução dos locus de microssatélites, e possivelmente, cada modelo pode explicar os eventos mutacionais em alguns locus, mas não em todos, pois parece não haver um único modo pelo qual as repetições de microssatélites podem aumentar ou diminuir em tamanho.

Uma característica da amplificação de microssatélites de dinucleotídeos é a representação de cada alelo por múltiplas bandas que presumivelmente se originam a partir de mecanismos de deslizamento das fitas durante a PCR e da atividade de transferase terminal da *Taq* polimerase (Mellersh & Sampson, 1993). A maioria das bandas espúrias pode ser eliminada com uma otimização cuidadosa das condições de amplificação. Apesar disto, algumas bandas “espelho” podem ocorrer. Para o locus *Egmsca1* não foi possível eliminar totalmente as bandas “espelho”, entretanto, a análise não foi prejudica pelo aparecimento destas. Ao contrário, as bandas “espelho” funcionaram como uma espécie de marcador de peso molecular, auxiliando na identificação dos alelos, visto que para cada alelo identificado pode ser atribuída uma banda “espelho” de peso constante.

A análise do locus *Egmsca1* também foi útil para ajudar a responder a questão relativa ao sistema de reprodução predominante em *E. granulosus*.

Durante vários anos o modelo proposto por Smyth & Smyth (1964) a respeito do modo de reprodução de *E. granulosus*, que seria amplamente ou exclusivamente por autofertilização, foi aceito, no qual as mutações ocorridas seriam sempre expressas porque são transmitidas tanto na forma homozigota quanto heterozigota por autofertilização obrigatória em adultos; se as mutações não forem selecionadas negativamente no hospedeiro intermediário, podem ser amplificadas por reprodução assexuada, gerando uma nova linhagem.

Rausch (1985) propôs um modelo alternativo, não mutuamente exclusivo ao anterior, no qual as linhagens se originariam por adaptação aos diferentes hospedeiros e por isolamento geográfico e os adultos se reproduziriam por fertilização cruzada.

Várias análises eletroforéticas de isozimas e de DNA deram suporte à proposição de que *E. granulosus* é um complexo de linhagens geneticamente distintas, com cada linhagem sendo completamente homozigota e monomórfica, como resultado da autofertilização (Lymbery & Thompson, 1988).

Entretanto, com a análise de um número maior de locos isoenzimáticos e com o surgimento de técnicas moleculares que permitem acessar uma maior quantidade de variabilidade genética, tornou-se claro que *E. granulosus* se reproduz também por fecundação cruzada (Lymbery *et al.*, 1990; Lymbery *et al.*, 1997; Haag *et al.*, 1999). Hoje em dia, a questão não mais é se ocorre ou não fecundação cruzada em *E. granulosus*, e sim, qual é o sistema predominante de reprodução.

Os nossos resultados corroboram a observação da ocorrência de fecundação cruzada em *E. granulosus*, pois a população composta de isolados da linhagem ovina provenientes do Brasil mostrou-se em equilíbrio de Hardy-Weinberg, não apresentando deficiência significativa de heterozigotos. Estes resultados só são possíveis se fecundação cruzada estiver ocorrendo naquela população.

Por outro lado, a população composta de isolados da linhagem ovina da Argentina não se mostrou em equilíbrio e uma deficiência de heterozigotos foi detectada. Esses dados confirmam que tanto autofertilização quanto fertilização cruzada ocorrem em *E. granulosus*, embora ocorra endocruzamento significativo, conforme revelado pelo índice $F_{is} = 0,48$.

Não deve ser descartada a possibilidade de ausência de heterozigotos na população argentina da linhagem ovina em função de fecundação cruzada entre indivíduos

geneticamente homogêneos derivados de um mesmo cisto (geitonogamia), o que leva a resultados semelhantes aos de autofecundação (LyMBERY *et al.*, 1997).

A deficiência de heterozigotos na população argentina também pode ser em decorrência do efeito Wahlund, que ocorre quando populações subdivididas são agrupadas como uma só. Este efeito seria possível devido à ausência de informação sobre a origem exata de isolados compreendendo cada área amostrada.

Possivelmente, um balanço entre fertilização cruzada e autofertilização foi a melhor solução encontrada pelo parasita para manter sua evolução. Parece que as taxas de recombinação não podem ser tão altas a ponto de romper complexos gênicos co-adaptados nem tão baixas a ponto de impedir alterações adaptativas (Haag *et al.*, 1998).

O coeficiente F_{st} estima a porção de variação que é atribuível a divergência entre populações. O valor negativo encontrado para as populações argentina e brasileira da linhagem ovina indica que estas duas populações não estão estruturadas, resultado confirmado pela ausência de diferenciação gênica ou genotípica entre as duas populações.

Embora um pequeno número de indivíduos das linhagens do boi e do camelo tenha sido analisado, parece que estas linhagens, as quais apresentaram somente um alelo no locus Egmsca1 (monomórficas) são menos variáveis neste locus que as linhagens da ovelha e da ovelha da Tasmânia.

O alelo (CA)₇ foi compartilhado pelas linhagens do boi e do camelo. Em três árvores parcimoniosas construídas a partir de conjuntos de dados mitocondriais (Bowles *et al.*, 1995), as linhagens do camelo e do boi são mais relacionadas entre si do que ambas com as linhagens da ovelha ou da ovelha da Tasmânia. Por outro lado, estas análises mostraram que as linhagens da ovelha e da ovelha da Tasmânia são proximamente relacionadas, uma possível explicação para o compartilhamento dos alelos (CA)₈ e (CA)₁₀ por estas duas linhagens.

Analisando quatro seqüências nucleotídicas de *E. granulosus*, duas delas seqüências parciais de íntrons de um gene de actina (EgactII) e de um gene homeobox (Hbx2) e as outras duas seqüências parciais de regiões codificantes do gene de uma das subunidades do antígeno B (AgB8/1) e de um gene mitocondrial (ND1), Haag *et al.* (1998) verificaram a ausência de variação dentro das linhagens do boi, do cavalo, da ovelha e do porco, embora houvesse variação entre as linhagens. As linhagens do boi e do porco apresentaram alta similaridade quando as seqüências foram utilizadas para o cálculo de distâncias genéticas.

Embora apenas o locus Egmsca1 tenha sido utilizado para um estudo populacional, os resultados acerca da utilização de microssatélites como marcadores moleculares em *E. granulosus* são extremamente promissores, pois foi possível observar variabilidade dentro de linhagens, como nos casos das linhagens da ovelha e da ovelha da Tasmânia. Em estudos prévios, a maioria dos loci analisados apresentou monomorfismo dentro de linhagens ou grande deficiência de heterozigotos (Lymbery & Thompson, 1988; Lymbery *et al.*, 1990; Lymbery *et al.*, 1995; Haag *et al.*, 1998), indicando que embora fecundação cruzada ocorra, não deve ser o sistema de reprodução predominante.

Conforme citado anteriormente, o fato da população de isolados do Brasil pertencentes à linhagem da ovelha não ser monomórfica, apresentar seus alelos em equilíbrio de Hardy-Weinberg e não mostrar deficiência de heterozigotos só é possível se nesta população estiver ocorrendo fecundação cruzada.

A análise de vários loci de microssatélites em *E. granulosus* poderá ajudar a responder à questão de qual sistema de reprodução é predominante neste parasita.

No presente estudo também foram isoladas seqüências de microssatélites dentro do complexo gênico snRNA U1. Os snRNAs (*small nuclear RNAs*) são pequenas moléculas que se associam a proteínas formando as snRNPs (ribonucleoproteínas pequenas nucleares) e que participam no processamento do RNA mensageiro.

Bretagne *et al.* (1991) isolaram a seqüência completa do gene snRNA U1 de *E. multilocularis* e definiram a estrutura do gene, o qual está repetido no genoma do parasita. O número de repetições foi estimado entre 50 e 100 cópias. Os genes estão localizados em um *cluster* e apresentam uma organização similar a dos genes rRNA. A unidade de repetição tem 1300pb e consiste de espaçadores e uma região transcrita de 156 pb. As regiões espaçadoras apresentam diferenças consistindo de um número variável de microssatélites de tri, tetra e pentanucleotídeos.

Embora haja a estimativa de 50 a 100 cópias do gene snRNA U1 nos genomas de *E. multilocularis* e de *E. granulosus* (Bretagne *et al.*, 1996), o número máximo de bandas de hibridização encontrado no presente estudo em um experimento de *Southern blot* foi nove. Possivelmente cada fragmento hibridizado com a sonda possui mais de uma unidade de repetição deste gene.

Utilizando um *primer* senso descrito por Bretagne *et al.* (1996) para *E. multilocularis* e um *primer* antisenso projetado para amplificar um fragmento de 256 pb contendo um

microssatélite de pentanucleotídeos, foram obtidos fragmentos do tamanho esperado em *E. granulosus*. A clonagem e seqüenciamento deste fragmento revelaram seis clones com diferenças nas seqüências e número de unidades repetitivas dos microssatélites. As seqüências encontradas apresentaram similaridade com a seqüência do snRNA U1 de *E. multilocularis* (Bretagne *et al.*, 1996) quando comparadas utilizando-se o programa BLAST.

Das seis seqüências obtidas, três apresentaram microssatélites imperfeitos e três microssatélites compostos. Um microssatélite composto apresentou a seqüência (GACGA)₄(GGCGA)₃, a qual pode ter se originado a partir de uma transição A → G e posterior expansão do segmento com a mutação. Dois microssatélites compostos apresentaram a seqüência (GACGA)_n(GCGAG)_n. A origem dessa seqüência composta pode ter ocorrido pela deleção de uma base A em um microssatélite perfeito, desde que a última repetição fosse seguida por uma base G.

O teste de amplificação com DNA de vários isolados e com os *primers* utilizados para o isolamento da seqüência do snRNA U1 em *E. granulosus* mostrou que esse microssatélite não é ideal para uso como marcador molecular. Muitas bandas de amplificação puderam ser observadas. Estas bandas provavelmente refletem a amplificação de segmentos dentro de diversas unidades repetitivas do complexo gênico, ou seja, segmentos homólogos parálogos provavelmente estão sendo amplificados e estes segmentos diferem entre si pelas seqüências dos microssatélites. Seqüências parálogas são aquelas resultantes de eventos de duplicação gênica, enquanto que seqüências ortólogas são decorrentes de eventos de especiação.

Bretagne *et al.* (1996) projetaram *primers* para a amplificação de um segmento de 420 pb contendo o microssatélite de pentanucleotídeos em *E. multilocularis*. A análise, realizada em um seqüenciador automático, de 41 isolados revelou três perfis eletroforéticos. O primeiro perfil, chamado A, foi encontrado em isolados europeus e foi caracterizado pela presença de um pico em 414 pb e pela relação entre a área do pico em 414 pb e a área do pico em 429 pb sempre maior ou igual a 2,5.

O perfil B, encontrado em isolados do Alasca e do Japão, caracterizou-se pela ausência do pico em 414 pb e pela relação entre a área do pico em 421 pb e a área do pico em 429 pb menor ou igual a 0,5. O perfil C, encontrado em isolados do Alasca e de Montana, foi caracterizado por produtos de PCR mais longos, entre 411 e 439 pb. Embora estes resultados sejam interessantes, o ideal é a utilização de lócus de microssatélites localizados fora de regiões repetitivas, para não haver a amplificação de seqüências parálogas.

A impossibilidade de detectar os fragmentos ortólogos para comparação entre os diferentes isolados dificulta a utilização do segmento amplificado como marcador molecular em *E. granulosus*. Além disso, o fato dos microssatélites encontrados no complexo gênico snRNA U1 serem imperfeitos ou compostos também dificulta sua utilização como marcador molecular. Locus de microssatélites imperfeitos tendem a ser menos polimórficos e o polimorfismo detectado pode realmente ser devido à comparação de segmentos parálogos. Microssatélites compostos, conforme citado anteriormente, podem apresentar eletromorfos com diferentes números de cada unidade repetitiva, o que só é possível detectar seqüenciando os produtos amplificados.

5. CONCLUSÕES

1. Os microssatélites estão amplamente distribuídos no genoma de *E. granulosus*, sendo que as seqüências repetitivas CA, AAC, CATA e GA são as mais freqüentes entre as que foram analisadas.

2. Os experimentos de *Southern blot*, utilizando oligonucleotídeos repetitivos como sondas, são extremamente úteis quando realizados antes da etapa de isolamento dos microssatélites, permitindo que esforços sejam direcionados para o isolamento de microssatélites abundantes no genoma.

3. O locus Egmsca1 é um bom marcador molecular em *E. granulosus*, pois apresentou estabilidade entre protoescóleces originados por reprodução assexuada de um único cisto hidático e polimorfismo entre isolados de linhagens diferentes e de uma mesma linhagem.

4. Os resultados obtidos confirmam a ocorrência de fertilização cruzada em *E. granulosus*, além de autofertilização.

5. O microssatélite com repetições de pentanucleotídeos contido no complexo gênico do snRNA U1 não é um marcador molecular ideal em *E. granulosus*, pois apresentou repetições imperfeitas ou compostas, além da amplificação de muitos fragmentos parálogos.

6. PERSPECTIVAS

O isolamento de mais loci de microssatélites e a utilização desses loci, juntamente com o locus Egmsca1 e o locus Egmsca3 (isolado neste estudo, mas ainda não testado) possivelmente permitirão a caracterização das linhagens de *E. granulosus* e a verificação de variabilidade dentro de linhagens, a qual tem sido pouco relatada na literatura com a utilização de outros métodos bioquímicos e moleculares, mas que pode ser constatada para o locus Egmsca1. Também auxiliarão a responder a questão sobre qual é o sistema predominante de reprodução neste parasita.

O estudo de vários loci de microssatélites em *E. granulosus* poderá aumentar os conhecimentos relativos à estrutura e dinâmica populacional do parasita, as quais estão relacionadas aos aspectos epidemiológicos da hidatidose. Os microssatélites poderão ser utilizados para detectar as fontes de transmissão aos seres humanos em um determinado foco endêmico, ou ainda, para determinar se condições clínico-patológicas da hidatidose (incluindo respostas a drogas quimioterápicas) ou diferenças quanto à prevalência da doença podem se correlacionar a alelos de microssatélites.

A amplificação dos loci isolados também poderá ser testada nas outras três espécies de *Echinococcus* e em outros tenídeos. Em caso de sucesso na amplificação, permitirá a realização de análises filogenéticas e também, o estudo da evolução das seqüências dos microssatélites nestas espécies.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- AKAI, J.; KIMURA, A. & HATA, R.-I. Transcriptional regulation of the human type I collagen $\alpha 2$ (*COL1A2*) gene by the combination of two dinucleotide repeats. **Gene** **239**: 65-73, 1999.
- AMOS, W. A comparative approach to the study of microsatellite evolution. In: GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.
- AMOS, W.; SAWCER, S. J.; FEAKES, R. W. & RUBINSZTEIN, D. C. Microsatellites show mutational bias and heterozygote instability. **Nat. Genet.** **13**: 390-391, 1996.
- ANADOL, D.; OZÇELIK, U.; KIPER, N. & GÖÇMEN, A. Treatment of hydatid disease. **Paediat. Drugs** **3**: 123-135, 2001.
- AOKI, T.; KOCH, K. & LEFFERT, H. L. Attenuation of gene expression by a trinucleotide repeat-rich tract from the terminal exon of the rat hepatic polymeric immunoglobulin receptor gene. **J. Mol. Biol.** **267**: 229-236, 1997.
- ARMOUR, J. A. L.; ALEGRE, S. A.; MILES, S.; WILLIAMS, L. J. & BADGE, R. M. Minisatellites and mutation processes in tandemly repetitive DNA. In: GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.
- ASHTON, R. M. Genetic diversity and population structure of the New Zealand sea lion (*Phocarctos hookeri*). Dissertação de Mestrado, Universidade de Auckland. 1999. *Apud* CHAMBERS, G. K. & MACAVOY, E. S. Microsatellites: consensus and controversy. **Comp. Biochem. Phys. Part B** **126**: 455-476, 2000.
- BECKMANN, J. S. & WEBER, J. L. Survey of human and rat microsatellites. **Genomics** **12**: 627-631, 1992.

- BELL, G. I. & JURKA, J. The length distribution of perfect dimer repetitive DNA is consistent with its evolution by an unbiased single-step mutation process. **J. Mol. Evol.** **44**:414-421, 1997.
- BENSON, G. Tandem repeats finder: a program to analyze DNA sequences. **Nucleic Acids Res.** **27**: 573-580, 1999.
- BINZ, T., REUSCH, T. B. H., WEDEKING, C., SCHÄRER, L., STERNBERG, J. M. & MILINSKI, M. Isolation and characterization of microsatellite loci from the tapeworm *Schistocephalus solidus*. **Mol. Ecol.** **9**: 1926-1927, 2000.
- BOWLES, J., BLAIR, D. & McMANUS, D. P. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. **Mol. Biochem. Parasitol.** **54**: 165-174, 1992.
- BOWLES, J., BLAIR, D. & McMANUS, D. P. A molecular phylogeny of the genus *Echinococcus*. **Parasitology** **110**: 317-328, 1995.
- BRETAGNE, S., ASSOULINE, B., VIDAUD, D., HOUIN, R. & VIDAUD, M. *Echinococcus multilocularis*: microsatellite polymorphism in U1 snRNA genes. **Exp. Parasitol.** **82**: 324-328, 1996.
- BRETAGNE, S.; ROBERT, B.; VIDAUD, D.; GOOSSENS, M. & HOUIN, R. Structure of the *Echinococcus multilocularis* U1 snRNA gene repeat. **Mol. Biochem. Parasitol.** **46**: 285-292, 1991.
- CAO, H.; WIDLUND, H. R.; SIMONSSON, T. & KUBISTA, M. TGGA repeats impair nucleosome formation. **J. Mol. Biol.** **281**: 253-260, 1998.
- CHAMBERS, G. K. & MACAVOY, E. S. Microsatellites: consensus and controversy. **Comp. Biochem. Phys. Part B** **126**: 455-476, 2000.

- DI RIENZO, A.; PETERSON, A. C.; GARZA, J. C.; VALDES, A. M.; SLATKIN, M. & FREIMER, N. B. Mutational processes of simple-sequence repeat loci in human populations. **Proc. Natl. Acad. Sci. USA** **91**: 3166-3170, 1994.
- DOKHOLYAN, N. V., BULDYREV, S. V., HAVLIN, S. & STANLEY, H. E. Distributions of dimeric tandem repeats in non-coding and coding DNA sequences. **J. Theor. Biol.** **202**: 273-282, 2000.
- ECKERT, J., CONRATHS, F. J. & TACKMANN, K. Echinococcosis: an emerging or re-emerging zoonosis? **Int. J. Parasitol.** **30**:1283-1294, 2000.
- EDWARDS, K. J., BARKER, J. H. A., DALY, A., JONES, C. & KARP, A. Microsatellite libraries enriched for several microsatellite sequences in plants. **Biotechniques** **20**:758-760, 1996.
- EISEN, J.A. (1999). Mechanistic basis for microsatellite instability. In **Microsatellites: evolution and applications** (ed. Goldstein, D. B. & Schlotterer, C.), pp. 34-48. Oxford University Press, Oxford.
- ELLEGREN, H. Microsatellite mutations in the germline: implications for evolutionary inference. **Trends Genet.** **16**: 551-558, 2000a.
- ELLEGREN, H. Heterogeneous mutation processes in human microsatellite DNA sequences. **Nat. Genet.** **24**: 400-402, 2000b.
- ELLEGREN, H. Microsatellite evolution: a battle between replication slippage and point mutation. **Trends Genet.** **18**: 70, 2002.
- ESTOUP, A. & CORNUET, J.-M. Microsatellite evolution: inferences from population data. In: GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.

- FERDIG, M. T. & SU, X. -Z. Microsatellite markers and genetic mapping in *Plasmodium falciparum*. **Parasitol. Today** **16**: 307-312, 2000.
- FISHER, M. C. & VINEY, M. E. Microsatellites of the parasitic nematode *Strongyloides ratti*. **Mol. Biochem. Parasitol.** **80**: 221-224, 1996.
- GASSER, R. B. & CHILTON, N. B. Characterization of taeniid cestode species by PCR-RFLP of ITS2 ribosomal DNA. **Acta Trop.** **59**: 31-40, 1995.
- GOLDSTEIN, D. B. & POLLOCK, D. D. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. **J. Hered.** **88**: 335-342, 1997.
- GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.
- GOUDET, J., RAYMOND, M., DE MEEÛS, T. & ROUSSET, F. Testing differentiation in diploid populations. **Genetics** **144**, 1933-1940, 1996.
- HAAG, K.L., ARAÚJO, A. M., GOTTSTEIN, B., SILES-LUCAS, M., THOMPSON, R. C. A. & ZAHA, A. Breeding systems in *Echinococcus granulosus* (Cestoda; Taeniidae): selfing or outcrossing? **Parasitology** **118**: 63-71, 1999.
- HAAG, K. L.; ARAÚJO, A. M.; GOTTSTEIN, B. & ZAHA, A. Selection, recombination and history in a parasitic flatworm (*Echinococcus*) inferred from nucleotide sequences. **Mem. Inst. Oswaldo Cruz** **93**: 695-702, 1998.
- HANCOCK, J. M. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.

- HARR, B.; ZANGERL, B. & SCHLÖTTERER, C. Removal of microsatellite interruptions by DNA replication slippage: phylogenetic evidence from *Drosophila*. **Mol. Biol. Evol.** **17**: 1001-1009, 2000.
- HOEKSTRA, R., CRIADO-FORNELIO, A., FAKKELDIJ, J., BERGMAN, J. & ROOS, M. H. Microsatellites of the parasitic nematode *Haemonchus contortus*: polymorphism and linkage with a direct repeat. **Mol. Biochem. Parasitol.** **89**: 97-107, 1997.
- JIN, L.; ZHONG, Y. & CHAKRABORTY, R. The exact number of possible microsatellite motifs. **Am. J. Hum. Genet.** **55**: 582-583, 1994.
- JIN, L.; MACAUBAS, C.; HALLMAYER, J.; KIMURA, A. & MIGNOT, E. Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. **Proc. Natl. Acad. Sci. USA** **93**: 15285-15288, 1996.
- KASHI, Y. & SOLLER, M. Functional roles of microsatellites and minisatellites. In: GOLDSTEIN, D. B. & SCHÖLTTERER, C. (ed.) **Microsatellites: evolution and applications**. Oxford University Press, Oxford, 1999.
- KASSAI, T. Disease nomenclature. **Trends Parasitol.** **17**: 217-218, 2001.
- KATTI, M. V., RANJEKAR, P. K. & GUPTA, V. S. Differential distribution of simple sequence repeats in eukaryotic genome sequences. **Mol. Biol. Evol.** **18**: 1161-1167, 2001.
- KIMURA, M. & OHTA, T. Stepwise mutation model and distribution of allelic frequencies in a finite population. **Proc. Natl. Acad. Sci. USA** **75**, 2868-2872, 1978.
- KRUGLYAK, S.; DURRETT, R.; SCHUG, M. D. & AQUADRO, C. F. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. **Proc. Natl. Acad. Sci. USA** **95**: 10774-10778, 1998.

- KRUGLYAK, S.; DURRETT, R.; SCHUG, M. D. & AQUADRO, C. F. Distribution and abundance of microsatellites in the yeast genome can be explained by a balance between slippage events and point mutations. **Mol. Biol. Evol.** **17**: 1210-1219, 2000.
- KUMARATILAKE, L. M. & THOMPSON, R. C. A. A review of the taxonomy and speciation of the genus *Echinococcus* Rudolphi 1801. **Z. Parasitenkd.** **68**: 121-146, 1982.
- LAGERCRANTZ, U.; ELLEGREN, H. & ANDERSSON, L. The abundance of various polymorphic motifs differs between plants and vertebrates. **Nucleic Acids Res.** **21**: 1111-1115, 1993.
- LIGHTOWLERS, M. W. & GOTTSTEIN, B. Echinococcosis/Hydatidosis: antigens, immunological and molecular diagnosis. In: THOMPSON, R. C. A. & LIMBERY, A. J. (ed.). **Echinococcus and hydatid disease**. CAB INTERNATIONAL, Wallingford, 1995.
- LYMBERY, A. J.; CONSTANTINE, C. C. & THOMPSON, R. C. A. Self-fertilization without genomic or population structuring in a parasitic tapeworm. **Evolution** **51**: 289-294, 1997.
- LYMBERY, A. J. & THOMPSON, R. C. A. Electrophoretic analysis of genetic variation in *Echinococcus granulosus* from domestic hosts in Australia. **Int. J. Parasitol.** **18**: 803-811, 1988.
- 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. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. **Am. J. Hum. Genet.** **44**, 397-401, 1989.

- MACAUBAS, C.; JIN, L.; HALLMAYER, J.; KIMURA, A. & MIGNOT, E. The complex mutation pattern of a microsatellite. **Gen. Res.** **7**: 635-641, 1997.
- MARCOTTE, E. M.; PELLEGRINI, M.; YEATES, T. O. & EISENBERG, D. A census of protein repeats. **J. Mol. Biol.** **293**: 151-160, 1998.
- MCMANUS, D. P. & BRYANT, C. Biochemistry, physiology and molecular biology of *Echinococcus*. In: THOMPSON, R. C. A. & LIMBERY, A. J. (ed.). ***Echinococcus and hydatid disease***. CAB INTERNATIONAL, Wallingford, 1995.
- MCMANUS, D. P.; KNIGHT, M. & SIMPSON, A. G. Isolation and characterization of nucleic acids from the hydatid organisms, *Echinococcus* spp. (Cestoda). **Mol. Biochem. Parasitol.** **16**: 251-266, 1985.
- MELLERSH, C. & SAMPSON, J. Simplifying detection of microsatellite length polymorphisms. **Biotechniques** **15**: 138-139, 1993.
- MESSIER, W.; LI, S.-H. & STEWART, C.-B. The birth of microsatellites. **Nature** **381**: 483, 1996.
- NAOM, I.S., MATHEW, C. G., TOWN, M. -M. Genetic mapping with microsatellites. In: GLOVER, D. M. & HAMES, B. D. (ed.). **DNA cloning 3: a practical approach**, pp. 195-217. Oxford University Press, Oxford, 1995.
- OLIVEIRA, R. P., BROUDE, N. E., MACEDO, A. M., CANTOR, C. R., SMITH, C. L. & PENA, S. D. J. Probing the genetic structure of *Trypanosoma cruzi* with polymorphic microsatellites. **Proc. Natl Acad. Sci. USA** **95**: 3776-3780, 1998.
- OLIVEIRA, R. P.; MELO, A. I. R.; MACEDO, A. M.; CHIARI, E. & PENA, S. D. J. The population structure of *Trypanosoma cruzi*: expanded analysis of 54 strains using eight polymorphic CA-repeat microsatellites. **Mem. Inst. Oswaldo Cruz** **94**: 65-70, 1999.

- OTSEN, M., PLAS, M. E., LENSTRA, J. A., ROOS, M. H. & HOEKSTRA, R. Microsatellite diversity of isolates of the parasitic nematode *Haemonchus contortus*. **Mol. Biochem. Parasitol.** **110**, 69-77, 2000.
- PRIMMER, C. R., ELLEGREN, H.; SAINO, N. & MOLLER, A. P. Directional evolution in germline microsatellite mutations. **Nat. Genet.** **13**: 391-393, 1996.
- RAMACHANDRAN, S.; BEUKEBOOM, L. W.; GERACE, N.; PAVLOVIC, N.; CARRANZA, S. & MICHIELS, N. K. Isolation and characterization of microsatellite loci from the planarian *Dugesia polychroa* (Schmidt) (Platyhelminthes: Tricladida). **Mol. Ecol.** **6**: 389-391, 1997.
- RAMEL, C. Mini- and microsatellites. **Environ. Health Persp.** **105**: 781-789, 1997.
- RAUSCH, R. L. Parasitology: retrospect and prospect. **J. Parasitol.** **71**: 139-151, 1985.
- RAYMOND, M. & ROUSSET, F. Genepop (version 1.2): population genetics software for exact tests and ecumenecism. **J. Hered.** **86**, 248-249, 1995a.
- RAYMOND, J. & ROUSSET, F. An exact test for population differentiation. **Evolution** **49**, 1280-1283, 1995b.
- REFSETH, U. H., FANGAN, B. M. & JAKOBSEN, K. S. Hybridization capture of microsatellites directly from genomic DNA. **Electrophoresis** **18**, 1519-1523, 1997.
- ROSSI, V., WINCKER, P., RAVEL, C., BLAINEAU, C., PAGÉS, M. & BASTIEN, P. Structural organisation of microsatellite families in the *Leishmania* genome and polymorphisms at two (CA)_n loci. **Mol. Biochem. Parasitol.** **65**, 271-282, 1994.
- RYCHLIK, W. & RHOADS, R. E. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and *in vitro* amplification of DNA. **Nucleic Acids Res.** **17**: 8543-8551, 1989.

- SAIMOT, A. G. Medical treatment of liver hydatidosis. **World J. Surg.** **25**: 15-20, 2001.
- SAMBROOK, J. & RUSSELL, D.W. **Molecular cloning: a laboratory manual**. 3^a ed. Cold Spring Harbor Laboratory Press, New York, 2001.
- SANTIBÁÑEZ-KOREF, M. F.; GANGESWARAN, R. & HANCOCK, J. M. A relationship between lengths of microsatellites and nearby substitution rates in mammalian genomes. **Mol. Biol. Evol.** **18**: 2119-2123, 2001.
- SCHANTZ, P. M.; CHAI, J.; CRAIG, P. S.; ECKERT, J.; JENKINS, D. J.; MACPHERSON, C. N. L. & THAKUR, A. Epidemiology and control of hydatid disease. In: THOMPSON, R. C. A. & LIMBERY, A. J. (ed.). **Echinococcus and hydatid disease**. CAB INTERNATIONAL, Wallingford, 1995.
- SCHLÖTTERER, C. MICROSATELLITES. IN: HOELZEL, A. R. (ED.). **Molecular Genetics analysis of populations: a practical approach**. IRL Press, Oxford, pp. 237-261, 1999.
- SCHORDERET, D. E. & GARTLER, S. M. Analysis of CpG suppression in methylated and nonmethylated species. **Proc. Natl. Acad. Sci. USA** **89**: 957-961, 1992.
- SCOTT, J. C.; STEFANIAK, J.; PAWLOWSKI, Z. S. & MCMANUS, D. P. Molecular genetic analysis of human cystic hydatid cases from Poland: identification of a new genotype group (G9) of *Echinococcus granulosus*. **Parasitology** **114**: 37-43, 1997.
- SILES-LUCAS, M.; BENITO, M. C. & CUESTA-BANDERA, C. *Echinococcus granulosus*: genomic and isoenzymatic study of Spanish strains isolated from different intermediate hosts. **Vet. Parasitol.** **63**: 273-282, 1996.
- SMYTH, J. D. & SMYTH, M. M. 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, 1964.

- STALLINGS, R. L. Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequence: implications for human genetic diseases. **Genomics** **21**: 116-121, 1994.
- STALLINGS, R. L., FORD, A. F., NELSON, D., TORNEY, D. C., HILDEBRAND, C. E. & MOYZIZ, R. K. Evolution and distribution of (GT)_n repetitive sequences in mammalian genomes. **Genomics** **10**: 807-815, 1991.
- STERN, A. BROWN, M; NICKEL, P. & MEYER, T. F. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. **Cell** **47**: 61-71, 1986.
- TAUTZ, D. Hypervariability of simple sequences as a general source of polymorphic DNA markers. **Nucleic Acids Res.** **17**: 6463-6471, 1989.
- TAYLOR, J. S.; DURKIN, J. M. H. & BREDEN, F. The death of a microsatellite: a phylogenetic perspective on microsatellite interruptions. **Mol. Biol. Evol.** **16**: 567-572, 1999.
- THOMPSON, R. C. A. Biology and systematics of Echinococcus. In: THOMPSON, R. C. A. & LIMBERY, A. J. (ed.). **Echinococcus and hydatid disease**. CAB INTERNATIONAL, Wallingford, 1995.
- THOMPSON, R. C. A. & MCMANUS, D. P. Towards a taxonomic revision of the genus Echinococcus. **Trends Parasitol.** **18**: 452-457, 2002.
- UNDERWOOD, A. P., SUPALI T., WU, Y. & BIANCO, A. E. Two microsatellite loci from *Brugia malayi* show polymorphisms among isolates from Indonesia and Malaysia. **Mol. Biochem. Parasitol.** **106**: 299-302, 2000.
- WEBER, J. L. Informativeness of human (dC-dA)_n-(dG-dT)_n polymorphisms. **Genomics** **7**: 524-530, 1990.

WEBER, J. L. & MAY, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. **Am. J. Hum. Genet.** **44**, 388-396, 1989.

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

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

XU, W.; LIU, L.; EMSON, P. C.; HARRINGTON, C. R. & CHARLES, I. G. Evolution of a homopurine-homopyrimidine pentanucleotide repeat sequence upstream of the human inducible nitric oxide synthase gene. **Gene** **204**: 165-170, 1997.

8. ANEXO – ARTIGO CIENTÍFICO

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

Aceito para publicação pela revista *Parasitology* em 02/01/2003.