

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCIÊNCIAS

SEQÜENCIAMENTO DE ESTs E ANÁLISE DE SEQÜÊNCIAS RELACIONADAS
AO ANTÍGENO B DE *Echinococcus granulosus*

Manuel Rodrigues Loncan

Trabalho apresentado como um dos requisitos
para obtenção do grau de Bacharel no Curso
de Ciências Biológicas Ênfase Molecular,
Celular e Funcional.

Orientador: Dr. Henrique B. Ferreira

Porto Alegre
Dezembro, 2004



Este trabalho foi inteiramente realizado no Centro de Biotecnologia do Estado do Rio Grande do Sul, sendo financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS).

Agradecimentos

Ao Dr. Henrique B. Ferreira, pela oportunidade que me foi dada de trabalhar no grupo LBMC, pela orientação durante as etapas deste trabalho e pela amizade.

Ao Dr. Arnaldo Zaha, por sua co-orientação e amizade.

À Dra. Cecília Fernández, por ter disponibilizado as bibliotecas de cDNA, matéria-prima para todo meu trabalho e ao Dr. John Parkinson, pelo processamento das seqüências.

Ao Cris, por todos os ensinamentos, pelas discussões biológicas, científicas e filosóficas, pela grande amizade e por ter sido um verdadeiro guia para mim neste laboratório e no mundo da biologia molecular.

A todos os colegas dos grupos LBMC, LEPA e Genoma pela amizade e por terem tornado a vivência nos laboratórios uma experiência tão agradável e frutífera. Em especial, àqueles que colaboram de alguma forma com meu trabalho: Anelise, pela introdução ao laboratório; Cris, Christhyane, Roberta, Adri, Antônio, Leandro, Fernandinha e Cláudio pelo trabalho em conjunto na bancada; Gustavo pela ajuda com o Antígeno B e ao bioinformata Marcos pela ajuda na “parte de bioinformática”.

Aos funcionários do Centro de Biotecnologia.

A todos os professores do curso de ciências biológicas.

À minha família, pelo carinho, incentivo e apoio incondicionais e por todos os cafés, coca-colas e outras fontes de cafeína em horários inusitados.

Índice

Lista de abreviaturas, símbolos e unidades	6
Lista de figuras	9
Lista de tabelas	10
1. Introdução	11
1.1 A biologia de <i>Echinococcus granulosus</i> e a hidatidose cística	12
1.1.2 O antígeno B de <i>Echinococcus granulosus</i>	17
1.2 Seqüenciamento de ESTs	19
1.3 Objetivos gerais	22
1.3.1 Objetivos específicos	22
2. Materiais e Métodos	24
2.1 Bibliotecas de cDNA	24
2.2 Triagem das colônias transformantes	24
2.3 Estocagem dos clones	25
2.4 Multiplicação dos clones bacterianos em meio de cultura líquido	25
2.5 Extração de DNA plasmidial em pequena escala	26
2.6 Eletroforese em gel de agarose	26
2.7 Seqüenciamento de DNA	27
2.7.1 Reação de seqüenciamento	27
2.7.2 "Primers"	28
2.7.3 Seqüenciamento automático	28
2.8 Processamento, anotação e submissão das seqüências	29
2.9 Identificação e análise de seqüências relacionadas às diferentes subunidades do Antígeno B	30
3. Resultados e Discussão	31
3.1 Triagem das colônias transformantes e estocagem de clones	31
3.2 Extração de DNA plasmidial em pequena escala	31
3.3 Seqüenciamento, processamento e submissão de ESTs de <i>Echinococcus granulosus</i>	32
3.4 Identificação e análise de seqüências relacionadas às diferentes subunidades do antígeno B de <i>Echinococcus granulosus</i>	35

4. Perspectivas	43
5. Referências bibliográficas	44

Lista de abreviaturas, símbolos e unidades

'	minuto
"	segundo
°C	graus Celsius
µL	microlitro
µm	micrômetro
A	alanina
A	adenina
aa	aminoácido
AgB1	subunidade 1 de 8 kDa do antígeno B
AgB3	subunidade 3 de 8 kDa do antígeno B
AgB4	subunidade 4 de 8 kDa do antígeno B
BAC	"bacterial artificial cromossome"
C	cisteína
C	citosina
cDNA	DNA complementar
Cf	concentração final
cm	centímetro
CW	"cyst wall"
D	ácido aspártico
DNA	ácido desoxirribonucléico
E	ácido glutâmico
EDTA	ácido etilenodiaminotetracético
EgAgB8/1	subunidade 1 de 8 kDa do antígeno B
EgAgB8/2	subunidade 2 de 8 kDa do antígeno B
EgAgB8/3	subunidade 3 de 8 kDa do antígeno B
EgAgB8/4	subunidade 4 de 8 kDa do antígeno B
EgAgB8/5	subunidade 5 de 8 kDa do antígeno B
ESTs	"expressed sequence tags"
ET	"energy transfer"
EtBr	brometo de etídeo

F	fenilalanina
fmol	fentomol = 10^{-15} mol
G	glicina
G	guanina
g	grama
gDNA	DNA genômico
GR	GeneRacer
GSS	"genome sequence survey"
H	histidina
I	isoleucina
IPTG	1-isopropil- β -D-1-tiogalactopiranosídeo
K	lisina
kb	quilobase
kDa	quilodalton
kV	quilovolt
L	leucina
L	litro
M	metionina
mL	mililitro
mm	milímetro
mRNA	RNA mensageiro
mtDNA	DNA mitocondrial
N	asparagina
ng	nanograma
nm	nanometro
nt	nucleotídeo
P	prolina
pb	par de bases
PCR	reação em cadeia da polimerase
pH	potencial hidrogeniônico
PS	protoescólicas
PS/P	protoescólicas tratados com pepsina

Q	glutamina
R	arginina
rAgB8/1	AgB8/1 recombinante
rAgB8/2	AgB8/2 recombinante
RNA	ácido ribonucléico
rpm	rotações por minuto
rRNA	RNA ribossômico
RT-PCR	PCR por transcrição reversa
S	serina
SDS	duodecilsulfato de sódio
SL	“spliced leader”
snRNA	“small nuclear RNA”
T	treonina
T	timina
TAE	Tris-acetato-EDTA
V	valina
V	volt
Y	tirosina
YAC	“yeast artificial cromossome”
pmol	picomol = 10^{-12} mol

Lista de figuras

Figura 1. Ciclo de vida de <i>Echinococcus granulosus</i>	15
Figura 2. Eletroforese em gel de agarose dos produtos de extração de DNA plasmidial em pequena escala.....	32
Figura 3. Análise feita pelo programa MegaBACE Sequencing Scorecard.....	33
Figura 4. Melhores alinhamentos gerados pelo algoritmo TBLASTN para as sequencias EgAgB8/3 (CHEMALE <i>et al.</i> , 2001), EgAgB8/4 (AREND <i>et al.</i> , 2002) e EgAgB8/1 (ROTT <i>et al.</i> , 2000).....	36
Figura 5. Alinhamento múltiplo entre as seqüências AgB3 (CHEMALE <i>et al.</i> , 2001), AgB3 (R5) (HAAG <i>et al.</i> , 2004), AgB4 (AREND <i>et al.</i> , 2002) e AgB1 (ROTT <i>et al.</i> , 2000) e as seqüências relacionadas a elas encontradas nas ESTs seqüenciadas.....	37

Lista de tabelas

Tabela 1. Resumo dos resultados gerados no LBMC.....	34
Tabela 2. Panorama atual das seqüências publicamente disponíveis para <i>E. granulosus</i>	34
Tabela 3. Listagem dos clones relacionados ao antígeno B de <i>E. granulosus</i> encontrados nas ESTs.....	39

UFRGS - BIBLIOTECA
INST. BIOCÊNCIAS

1. Introdução

O filo Platyhelminthes é composto por vermes com uma posição filogenética bastante interessante para estudos de biologia do desenvolvimento e de evolução comparada, além de incluir espécies parasitas de grande importância médica e veterinária.

Os platelmintos foram considerados por muito tempo como sendo os primeiros animais triploblásticos e com simetria bilateral que surgiram na Terra (BARNES, 1980). No entanto, novos estudos filogenéticos têm apontado que o processo de segmentação do plano corporal dos organismos bilatérios é, na verdade, uma condição derivada ao invés de ancestral (ADOUTTE *et al.*, 2000; MALLATT & WINCHELL, 2002; PATEL, 2003). Com base nesta hipótese, uma nova árvore filogenética foi proposta para os organismos metazoários, colocando os platelmintos e demais acelomados no ramo Lophotrochozoa, um super-clado que inclui todos organismos que possuem uma estrutura alimentar lofófora ou que apresentem durante seu desenvolvimento larvas trocóforas.

Os platelmintos da classe Cestoda possuem ciclos vitais heteroxenos, envolvendo mais de uma forma larval e diferentes hospedeiros que influenciam diretamente os processos básicos de crescimento e diferenciação do parasito. Dessa forma, o conhecimento dos processos de interação parasito-hospedeiro envolvidos na diferenciação celular e morfológica destes organismos, encerra, em si, elementos fundamentais para o estabelecimento da história evolutiva dos mecanismos envolvidos no desenvolvimento dos metazoários bilatérios em geral.

Em todo o mundo, infecções com parasitos platelmintos da classe Cestoda (cestodíases) causam uma morbidade substancial em populações humanas e de animais domésticos. Os cestódeos mais freqüentemente encontrados parasitando o homem e animais domésticos são os pertencentes à ordem Cyclophyllidea. A família Taeniidae é a mais importante da ordem Cyclophyllidea, incluindo, por exemplo, vermes dos gêneros *Echinococcus* e *Taenia* (HOBERG, 2002).

1.1 A biologia de *Echinococcus granulosus* e a hidatidose cística

A taxonomia de *Echinococcus* tem sido extensivamente revisada, uma vez que diversas espécies e subespécies foram descritas ainda na primeira metade do século XX (THOMPSON & MCMANUS, 2002). A variação intraespecífica encontrada em *Echinococcus granulosus* é de grande significância epidemiológica e o conceito de linhagens foi introduzido para refletir as variações fenotípicas encontradas em diferentes isolados do parasito (THOMPSON & LYMBERY, 1988). Estudos moleculares, utilizando principalmente seqüências de mtDNA (genes da NADH-desidrogenase 1, da citocromo c-oxidase 1) identificaram nove genótipos diferentes (G1-G9) em *E. granulosus* (THOMPSON & MCMANUS, 2002; MCMANUS *et al.*, 2003). Além, de *E. granulosus*, outras três espécies são consideradas taxonomicamente válidas dentro do gênero: *E. multilocularis*, *E. vogeli* e *E. oligarthrus*. As duas últimas estão de uma forma geral restritas a ciclos parasitários selvagens, em animais silvestres. No entanto, *E. granulosus* e *E. multilocularis*, constituem um grave problema em saúde pública, uma vez que, acidentalmente, podem utilizar o homem como hospedeiro intermediário no seu desenvolvimento. *E. granulosus* causa ainda, perdas econômicas significativas, pois é capaz de utilizar ungulados domésticos (bovinos, ovinos, suínos e eqüinos) como hospedeiros intermediários em seu ciclo de vida natural.

O verme adulto de *E. granulosus*, é uma pequena tênia, com 2 a 7 mm de comprimento (raramente chegando a 11 mm) (THOMPSON & LYMBERY, 1995; ECKERT *et al.*, 2001). O corpo do verme é composto de três regiões que são claramente distinguidas: o escólex, o colo e o estróbilo. O escólex possui quatro ventosas e duas fileiras de acúleos, formando o rostelo. É através desta estrutura que o parasito fixa-se ao intestino do hospedeiro definitivo. O colo não apresenta órgãos diferenciados, mas possui células parenquimáticas em intensa atividade mitótica, sendo responsável pelo alongamento da porção posterior, o estróbilo. Essa última região é composta de quatro (podendo chegar a 6) segmentos ou proglótides. A penúltima proglótide é sexualmente madura possuindo as genitálias masculina e feminina. Na última proglótide ocorre a ovulação e fecundação dos ovos.

Os vermes adultos vivem no intestino delgado do hospedeiro definitivo, onde fixam-se entre as vilosidades, alguns invadindo as criptas de Lieberkühn. A maturidade sexual é alcançada entre 4 e 5 semanas e a reprodução pode ocorrer por autofecundação ou por fecundação cruzada (HAAG *et al.* 1998). Diferentemente do que ocorre nas teníases, a fase parasitária no hospedeiro definitivo de *Echinococcus* é caracteristicamente assintomática. Cada cão infectado contém em média de 200 a 400 vermes no intestino (GEMMEL *et al.*, 2001), mas este número pode chegar até 50.000 (KAMMERER & SCHANTZ, 1993).

Após a fecundação, ocorre a produção dos ovos, cujo número pode chegar a 1500 na última proglótide (THOMPSON & LYMBERY, 1995). Os ovos são esféricos e medem de 30 a 40 μm de diâmetro. Dentro do ovo está o embrião hexacanto ou oncosfera, circundado por várias camadas. O embrióforo, a camada mais proeminente, é queratinizada e dá sustentação e resistência ao ovo. As proglótides grávidas ou os ovos são liberados no ambiente juntamente com as fezes do hospedeiro definitivo. Cada cão infectado libera em média 8470 ovos por dia (GEMMEL *et al.*, 2001), podendo chegar a 1 milhão (KAMMERER & SCHANTZ, 1993).

Os hospedeiros intermediários contaminam-se pela ingestão dos ovos do parasito. No estômago do hospedeiro, o ovo sofre a ação de enzimas digestivas liberando a oncosfera do embrióforo. A oncosfera é então ativada e penetra na parede intestinal, onde tem acesso a vasos sanguíneos ou linfáticos. Através deles, ela é transportada passivamente até o fígado, pulmão ou outro órgão, onde se fixa e desenvolve-se na forma larval do cisto hidático. O cisto hidático ou metacestódeo é a forma patogênica do *E. granulosus*, causadora da chamada hidatidose cística. A hidatidose cística é caracterizada por sintomas variáveis, que dependem do local de desenvolvimento do cisto, do seu tamanho e dos tecidos e órgãos comprimidos pela sua massa em crescimento.

O cisto hidático de *E. granulosus* é tipicamente unilocular e, apresenta uma estrutura subesférica, com sua cavidade interna sendo preenchida pelo líquido hidático (CAMERON & WEBSTER, 1969). O cisto consiste de uma camada

interna, a camada germinativa, que é sustentada externamente por uma camada elástica acelular, a camada laminar (THOMPSON & LYMBERY, 1995). A camada laminar, por sua vez é circundada por uma camada fibrosa de tecido conjuntivo produzida pelo hospedeiro, chamada de camada adventícia.

A camada laminar é constituída principalmente de mucopolissacarídeos e de uma pequena fração protéica, que são secretados pela camada germinativa (THOMPSON & LYMBERY, 1995). Ela confere uma proteção importante ao parasito uma vez que é capaz de inibir a ativação do sistema complemento do hospedeiro (DÍAZ *et al.*, 1999; FERREIRA *et al.*, 2000).

Na camada germinativa encontram-se vários tipos celulares, como células armazenadoras de glicogênio, musculares, tegumentares e células indiferenciadas de grande capacidade proliferativa (THOMPSON & LYMBERY, 1995). A partir destas células, formam-se as cápsulas prolíferas. Estas se originam como pequenas massas nucleares, proliferam e vacuolizam, podendo ficar presas por um pedúnculo curto à camada germinativa ou soltar-se no interior da cavidade cística.

As cápsulas prolíferas, por sua vez, diferenciam-se para dar origem aos protoescólices (THOMPSON, 1976; ROGAN & RICHARDS, 1987). Os protoescólices permanecem num estado de dormência no interior do cisto hidático até que a integridade deste seja afetada (THOMPSON & LYMBERY, 1990). Quando ocorre a ruptura do cisto, os protoescólices podem se vesicular e formar novos cistos (cistos secundários). Quando ingeridos pelo hospedeiro definitivo, os protoescólices diferenciam-se em vermes adultos, fechando o ciclo do parasito (THOMPSON & LYMBERY, 1995). Este papel é geralmente exercido pelos cães, que adquirem os protoescólices através da ingestão de vísceras de ungulados infectados (hospedeiros intermediários). As cápsulas prolíferas são liberadas do cisto hidático pela ação mecânica do aparelho digestivo do cão. As cápsulas prolíferas por sua vez, sofrem ação proteásica por parte da pepsina encontrada no estômago do cão, liberando então os protoescólices. O ciclo de vida de *E. granulosus* está representado na Figura 1.

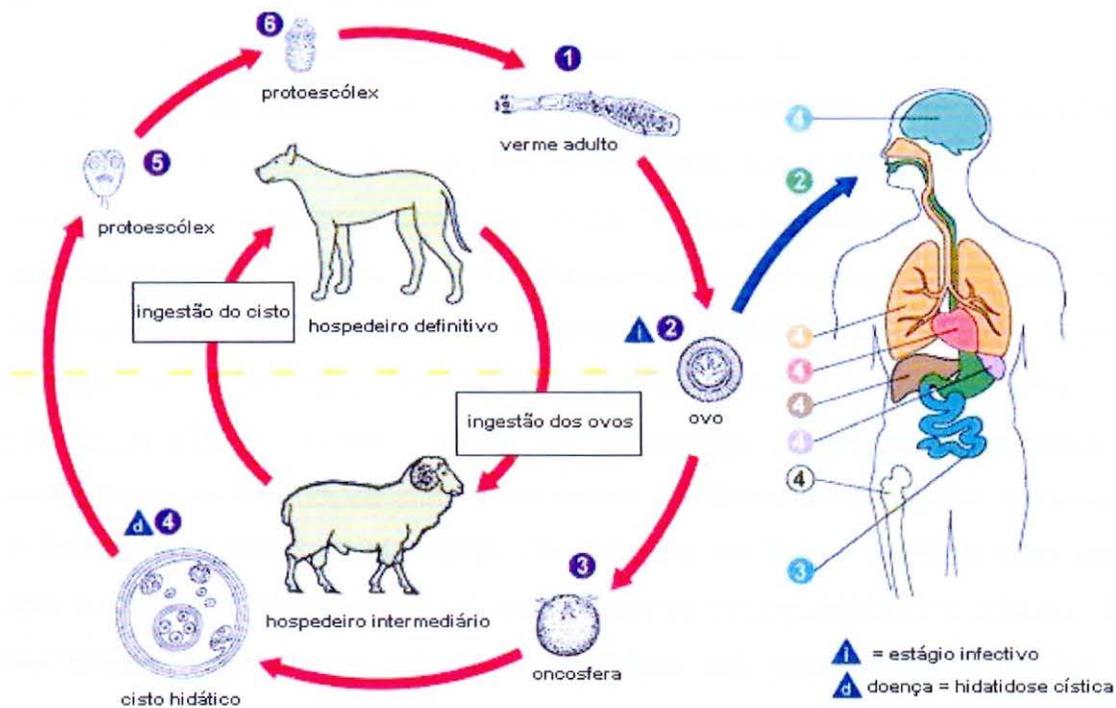


Figura 1. Ciclo de vida de *Echinococcus granulosus*. O verme adulto (1) vive no intestino delgado do hospedeiro definitivo (canídeos). As proglótides contendo os ovos (2) são liberadas juntamente com as fezes do hospedeiro no ambiente. Ao serem ingeridos pelo hospedeiro intermediário (ovinos, bovinos, suínos, cavalos, camelos, humanos, etc.) os ovos liberam a oncosfera (3), que penetra na parede intestinal e migra pelo sistema circulatório ou linfático até atingir um órgão, especialmente o fígado e pulmões. Nestes órgãos, a oncosfera diferencia-se num cisto (4) que cresce gradualmente, produzindo protoescolices (5). O hospedeiro definitivo infecta-se através da ingestão de vísceras do hospedeiro intermediário contaminadas com cistos do parasito. Após a ingestão, os protoescolices evaginam (6) e prendem-se à mucosa intestinal do hospedeiro definitivo. Após 4-5 semanas os protoescolices tornam-se adultos (1), fechando o ciclo de vida do parasito. Modificado do Centro para o controle e Prevenção de Doenças infecciosas (www.dpd.cdc.gov, CDC – USA).

O líquido hidático que preenche o cisto hidático consiste de uma mistura de uma série de componentes bioquímicos, orgânicos e inorgânicos, que vão desde íons até proteínas (FRAYHA & HADDAD, 1980; SANCHEZ & SANCHEZ, 1971). Estas últimas podem ser de origem do hospedeiro, ou, então, produtos de

excreção e secreção dos protoescólices e da camada germinativa, que interagem com o sistema imune do hospedeiro.

O *E. granulosus* apresenta uma ampla distribuição geográfica e a prevalência da hidatidose cística humana está altamente relacionada à prevalência de criação de gado (ROMIG, 2003). Ele é encontrado em todos os continentes, com maiores prevalências na Eurásia (especialmente nos países do Mediterrâneo), na Rússia, estados independentes adjacentes e China, norte e leste da África, América do Sul e Austrália (MCMANUS *et al.*, 2003). O parasita também é encontrado no subcontinente Indiano e no Oriente Médio (TORGERSON & BUDKE, 2003). Há evidências claras de emergência e re-emergência da hidatidose cística em partes da China, Ásia central, leste europeu e Israel (MCMANUS *et al.*, 2003). Na América do Sul as regiões mais afetadas são a Argentina, Uruguai e os Andes peruanos (TORGERSON & BUDKE, 2003). No Brasil, a hidatidose cística é endêmica em populações humanas e em rebanhos de ungulados domésticos no Estado do Rio Grande do Sul, especialmente em municípios próximos à fronteira com o Uruguai ou a Argentina (KLOETZEL & PEREIRA, 1992; DE LA RUE *et al.*, 1994; ARAMBULO, 1997). Entre os anos de 1981 e 1999 foram registrados 716 casos de infecções em humanos no Rio Grande do Sul (referido em FARIAS *et al.*, 2004).

Talvez o efeito mais importante da hidatidose cística seja o potencial impacto econômico causado pela infecção nas criações de animais (TORGERSON & BUDKE, 2003). Estes efeitos foram analisados e é possível que em alguns países como Jordânia, Uruguai e Reino Unido, os prejuízos causados pela infecção representem a principal causa de perdas econômicas afetando a criação de ovelhas, na ordem de milhões de dólares nas áreas endêmicas (TORGERSON *et al.*, 2000a; TORGERSON *et al.*, 2000b; TORGERSON *et al.*, 2001).

1.1.2 O antígeno B de *Echinococcus granulosus*

O antígeno B (AgB) do parasito *Echinococcus granulosus* é um dos principais componentes antigênicos presente no líquido hidático (ORIOLE *et al.*, 1971). É uma lipoproteína polimérica, que em condições não-desnaturantes apresenta uma massa molecular de 120-160 kDa. Sob condições desnaturantes este antígeno apresenta subunidades de 8, 16 e 24 kDa (ORIOLE *et al.*, 1971; LIGHTOWLERS *et al.*, 1989). O AgB é altamente imunogênico em infecções humanas e é considerado um antígeno com alta sensibilidade e especificidade em testes imunodiagnósticos (LIGHTOWLERS *et al.* 1989; LIGHTOWLERS *et al.*, 1995; BARBIERI *et al.*, 1998; ROTT *et al.*, 2000; GONZÁLEZ-SAPIENZA *et al.*, 2000).

Experimentos *in vitro* mostraram que o AgB é capaz de inibir elastase e também o recrutamento de neutrófilos (SHEPHERD *et al.*, 1991). Além disso, este antígeno pode estar envolvido na evasão da resposta imunológica do hospedeiro, inibindo a migração de células polimorfonucleadas e provocando uma resposta celular não-protetora de linfócitos do tipo Th2 (RIGANO *et al.*, 2001). A função do AgB na biologia do parasito ou na interação hospedeiro-parasito permanece desconhecida, mas estes resultados sugerem que o AgB pode estar envolvido em diferentes mecanismos de evasão da resposta imune do hospedeiro.

O AgB é homólogo à proteínas pertencentes a um novo grupo de proteínas ricas em α -hélice que ligam-se a lipídeos (*helix rich hydrophobic ligand binding proteins*, HLBP) (CHEMALE, 2002), específicas de helmintos (SAGHIR *et al.*, 2000). Possíveis membros desta família foram descritos em *Moniezia expansa* (Barret *et al.*, 1997), *Taenia solium* (SAKO *et al.*, 2000), *Taenia crassiceps* (ZARLENGA *et al.*, 1994) e *Hymenolepis diminuta* (SAGHIR *et al.*, 2000). Como o AgB, estas proteínas oligoméricas são formadas por pequenas subunidades (7 a 11 kDa) e são abundantes em alguns estágios destes parasitos. Muitas delas são altamente imunogênicas (SAGHIR *et al.*, 2000; SAKO *et al.*, 2000; ZARLENGA *et al.*, 1994) e antígenos potenciais para o imunodiagnóstico das respectivas parasitoses. Como experimentalmente demonstrado, algumas destas proteínas ligam-se a ácidos graxos, o que sugere uma possível função destas proteínas no

metabolismo de lipídeos (BARRET *et al.*, 1997; SAGHIR *et al.*, 2000; CHEMALE, 2002).

Já foram identificados pelo menos quatro diferentes genes que codificam as subunidades de 8 kDa do AgB: EgAgB8/1 (FROSCH *et al.*, 1994), EgAgB8/2 (FERNÁNDEZ *et al.*, 1996), EgAgB8/3 (CHEMALE *et al.*, 2001) e EgAgB8/4 (AREND *et al.*, 2002), mostrando que o antígeno B é codificado por uma família de genes. HAAG *et al.* (2004) identificou uma seqüência que potencialmente codifica uma isoforma adicional do AgB (EgAgB8/5). A expressão de pelo menos alguns destes genes no metacestódeo foi confirmada pelo isolamento dos respectivos cDNAs. GONZÁLEZ *et al.* (1996) demonstrou que pelo menos duas subunidades diferentes, EgAgB8/1 e EgAgB8/2, fazem parte da composição dos oligômeros de 16, 24 e 32 kDa do AgB.

As subunidades EgAgB8/1 e EgAgB8/2 foram expressas como proteínas recombinantes em nosso laboratório e tiveram o seu potencial diagnóstico para a hidatidose cística avaliado (ROTT *et al.*, 2000; VIRGINIO *et al.*, 2003). A subunidade recombinante rAgB8/2 mostrou ser de alto valor diagnóstico, apresentando sensibilidade e especificidade superiores às da rAgB8/1, do AgB nativo e de outras proteínas antigênicas recombinantes. Nenhuma reação cruzada foi observada com infecções por outras espécies de *Echinococcus* utilizando-se rAgB8/2 e, devido a isto, foi proposto que a subunidade EgAgB8/2 é espécie-específica (ROTT *et al.*, 2000).

Com o conhecimento atual, não é possível ainda determinar quais genes são expressos e quais são as proporções relativas das diferentes subunidades na proteína nativa polimérica. É possível, ainda, que existam outras subunidades que ainda não tenham sido identificadas. HAAG *et al.*, (2004), através de experimentos de "Southern blot", demonstrou que estão presentes pelo menos sete cópias de genes do AgB no genoma de *E. granulosus*.

As proteínas codificadas pelos diferentes genes do AgB apresentam similaridade de 44-81% em suas seqüências de aminoácidos (HAAG *et al.*, 2004). Os estudos têm mostrado ainda que os diferentes loci gênicos do AgB podem apresentar diferentes graus de polimorfismo (FROSCH *et al.*, 1994; CHEMALE *et*

al., 2001; AREND *et al.*, 2002; AREND *et al.*, 2004; HAAG *et al.*, 2004). Acredita-se que a variabilidade antigênica apresentada por diferentes isoformas do AgB em um metacestódeo tenha a função de modular a resposta imune do hospedeiro (HAAG *et al.*, 2004). Além das importantes implicações biológicas para a função do AgB na evasão da resposta imune, esta variabilidade tem implicações para o diagnóstico da hidatidose cística, uma vez que diferenças no padrão de expressão das subunidades entre diferentes cistos produzem preparações antigênicas de diferente valor diagnóstico.

1.2 Seqüenciamento de ESTs

O advento do seqüenciamento de alta performance e a relativa diminuição dos custos permitiu que fossem seqüenciados por inteiro os genomas de um grande número de organismos, incluindo 202 espécies procarióticas e 31 eucarióticas (<http://www.genomesonline.org/>). Para espécies eucarióticas esses dados de seqüências são geralmente obtidos na forma de marcadores ("tags") de seqüências expressas (ESTs) (PARKINSON *et al.*, 2004a). ESTs são leituras feitas uma única vez a partir de fragmentos de cDNAs escolhidos aleatoriamente. Assumindo que as ESTs refletem os níveis e a diversidade reais de mRNAs, a estratégia de seqüenciamento de ESTs fornece informações acerca do perfil de expressão gênica em nível de mRNAs e da atividade metabólica, salientando processos importantes que ocorrem em um determinado transcrito (STÜRZENBAUM *et al.*, 2003). Criando-se bibliotecas com base em um tecido ou estágio de desenvolvimento específico, é possível, então, estabelecer associações entre mudanças no panorama transcricional e a biologia de um organismo.

As ESTs têm se mostrado uma ferramenta importante para projetos de seqüenciamento de genomas completos, porque facilitam a identificação de genes. Em organismos para os quais o seqüenciamento completo do genoma ainda é uma meta distante, a análise de ESTs torna-se uma estratégia com uma relação custo-benefício favorável para a descoberta de genes (PARKINSON *et*

al., 2002). A utilidade desta abordagem é ilustrada pela diversidade filogenética de organismos representados no dbEST (<http://www.ncbi.nlm.nih.gov/dbEST/>), o banco de dados de ESTs do NCBI.

Contudo, as ESTs são geralmente mais curtas do que os mRNAs completos dos quais elas são derivadas (até 700 pb) e são sujeitas a erros de seqüenciamento (PARKINSON *et al.*, 2004b). Outros problemas comuns incluem uma redundância significativa nos conjuntos de dados (alguns genes podem ser seqüenciados múltiplas vezes) e a falta de anotação consistente entre os diferentes projetos de seqüenciamento de ESTs. Em geral, estes dados não são bem organizados, o que torna difícil interpretá-los em um contexto genômico. Uma estratégia para otimizar a análise de ESTs consiste em processar e organizar estes dados em um contexto genômico que facilite análises em grande escala. Uma maneira efetiva de superar estes problemas é agrupar as ESTs em conjuntos que representam genes ("clusters"), e providenciar a anotação para estes conjuntos. Uma vez que ESTs representam apenas uma fração dos genes disponíveis para um determinado organismo, estes conjuntos de dados analisados são referidos como "genomas parciais" (PARKINSON *et al.*, 2004a).

A estratégia de seqüenciamento e análise de ESTs em grande escala tem sido aplicada com grande sucesso para diversos organismos parasitas, especialmente na forma de consórcios multinacionais. Estes consórcios envolvem centros de pesquisa tanto de países desenvolvidos como de países em desenvolvimento e visam o seqüenciamento de genomas de tamanho médio. Desta maneira foram iniciados projetos para o seqüenciamento dos genomas de parasitos como: *Plasmodium falciparum*, *Schistosoma mansoni*, *Trypanosoma cruzi*, *Trypanosoma brucei*, *Leishmania major* e *Brugia malayi* entre outros (DEGRAVE *et al.*, 2001). Numa primeira etapa, quando são formadas as redes, é feito o mapeamento do genoma e são gerados os recursos (construção de bibliotecas ESTs, GSS, "phage", cosmídeos, BACs e YACs). O seqüenciamento é feito inicialmente com o objetivo de descobrir novos genes, e depois passa a ser efetuado em uma escala cromossômica e genômica.

Um consórcio envolvendo os grupos Nematode Genomics (Edinburgo, Reino Unido), o Wellcome Trust Sanger Institute (Cambridge, Reino Unido) e o Genomic Sequencing Center (Saint Louis, Estados Unidos) foi estabelecido para seqüenciar ESTs de diversos nematódeos parasitas do homem, animais e plantas (PARKINSON *et al.*, 2004b). Apesar da existência de um repositório central para armazenamento de ESTs, o dbEST, foi criada uma base de dados (NEMBASE; <http://www.nematodes.org/nematodeESTs/nembase.html>) com o objetivo de fornecer aos usuários o acesso aos dados para futura análise e correlação com o contexto biológico. O NEMBASE atualmente abriga em torno de 100.000 ESTs de 10 diferentes espécies de nematódeos e permite acesso a todos os metadados associados.

Em um esforço realizado em diversos centros de pesquisa do Brasil (rede ONSA), foram geradas mais de 160.000 seqüências ESTs de *Schistosoma mansoni*, o agente causador da esquistossomíase (VERJOVSKI-ALMEIDA *et al.*, 2003), representando a primeira iniciativa em grande escala para o seqüenciamento de um organismo bilatério acelomado. O processamento destas seqüências resultou em um total de 31.000 ESTs agrupadas, das quais 77%, representam novos parálogos, ou ortólogos ou genes de função desconhecida. Estima-se que este número represente uma cobertura de 92% do transcrito de *S. mansoni*, cujo genoma possui aparentemente 14.000 genes.

Projetos-piloto de seqüenciamento de ESTs dos parasitos *Fasciola hepatica* e *Schistosoma haematobium* estão em andamento no Wellcome Trust Sanger Center (Cambridge, Reino Unido), com o objetivo de seqüenciar 15.000 ESTs de cada um deles (MCMANUS *et al.*, 2004).

É esperado que o conhecimento gerado com estas iniciativas resulte em novas ferramentas para o controle de doenças (GUTIERREZ *et al.*, 2000; DEGRAVE *et al.*, 2001; ALMEIDA *et al.*, 2002; MCMANUS *et al.*, 2004). Isto inclui a identificação de novos genes com funções essenciais para as células, que sirvam potencialmente de alvos para novas drogas, e a identificação de novos antígenos, que possam ser úteis no diagnóstico e no desenvolvimento de novas vacinas, ou ainda, que levem a compreensão de fenômenos de modulação da

resposta imune no hospedeiro. Esta imensa quantidade de dados que esta sendo gerada a partir de uma ampla variedade de organismos, permite ainda comparações em larga escala entre diferentes taxa ao nível molecular. Estas análises podem revelar relações evolutivas e os mecanismos moleculares subjacentes à diversidade biológica e a especialização entre os organismos.

Recentemente, foi lançado um projeto para a descoberta de genes em *Echinococcus granulosus* (FERNÁNDEZ *et al.*, 2002). Neste trabalho, foram construídas seis bibliotecas de cDNA enriquecidas em transcritos completos derivados de populações de mRNAs de protoescólices e da parede cística, permitindo estudos de expressão diferencial de genes no metacéstódeo. Na construção destas bibliotecas, foi aplicada uma metodologia que restringe as espécies de mRNA a somente aquelas codificadas pelo núcleo, também excluindo contaminações por rRNAs e gDNA (derivado do hospedeiro ou parasito). Para armazenar os dados gerados por este e outros projetos foi criada uma base de dados para seqüências de Lofotrocozoários (LophDB; <http://nema.cap.ed.ac.uk/Lopho/LophDB.php>).

1.3 Objetivos gerais

Este projeto tem como meta estabelecer o perfil preliminar de genes expressos em diferentes estágios de desenvolvimento de *E. granulosus*.

1.3.1 Objetivos específicos

(1) em colaboração com o Wellcome Trust Sanger Institute, gerar 2000 seqüências a partir de cada uma das bibliotecas de cDNA de *E. granulosus* e disponibiliza-las na forma de ESTs no dbEST e paralelamente no LophDB, na forma de “clusters” que representam genes.

(2) analisar comparativamente os dados obtidos de cada uma das bibliotecas entre si e com outros bancos de dados de domínio público de

nucleotídeos, proteínas e domínios funcionais, permitindo a caracterização estrutural e funcional do transcrito da fase larval (metacestódeo) de *E. granulosus*.

(3) identificar e analisar qualitativa e quantitativamente as seqüências relacionadas às diferentes subunidades do AgB presentes nas seqüências ESTs .

UFRGS - BIBLIOTECA
INST. BIOCÊNCIAS

2. Materiais e métodos

2.1 Bibliotecas de cDNA

O Laboratório de Biologia Molecular de Cestódeos (LBMC) recebeu alíquotas das bibliotecas de cDNA PS-SL, PS-GR, PS/P-SL, PS/P-GR e CW-SL, construídas pela Dra. Cecília Fernández (Universidad de la República, Uruguai), a partir de diferentes tecidos que foram obtidos de cistos hidáticos de bovinos naturalmente infectados (FERNÁNDEZ *et al.*, 2002). Os nomes das bibliotecas fazem referência ao material parasitário e ao tipo de cDNA que contêm. Os cDNAs foram sintetizados a partir de mRNAs da parede cística (CW), protoescólices (PS) e de protoescólices tratados com pepsina (PS/P) anteriormente à extração de RNA. A pepsina aumenta a atividade metabólica dos protoescólices, sendo um sinal naturalmente encontrado por eles quando ingeridos pelo hospedeiro definitivo. O kit GeneRacer (GR) (Invitrogen) foi utilizado com a finalidade de favorecer o isolamento de cDNAs completos. Similarmente, foi utilizado um “primer” que anela ao “spliced leader” (SL), que corresponde a um fragmento de snRNA que é adicionado por “trans-splicing” à extremidade 5' de cerca de 25% dos transcritos de *E. granulosus* (BREHM *et al.*, 2000). Os cDNAs foram clonados direcionalmente no vetor pSPORT1 (Invitrogen) e transformados por eletroporação em *E. coli* EletroMAX DH-10B (Invitrogen).

2.2 Triagem das colônias transformantes

Foi realizada uma triagem nas bibliotecas PS/P-SL, CW-SL e PS-GR para exclusão de clones contendo vetores religados através do plaqueamento em meio L.B sólido na presença do indutor IPTG (Invitrogen), do agente cromogênico X-Gal (Invitrogen) e do antibiótico ampicilina (Calbiochem) conforme descrito em SAMBROOK *et al.* (1989). As placas foram incubadas *overnight* em estufa a 37°C e a identificação das colônias portadoras de clones recombinantes foi feita de

forma visual: colônias bacterianas que foram transformadas com os plasmídios recombinantes apresentam-se brancas, enquanto que aquelas transformadas apenas com o vetor religado apresentam coloração azul.

2.3 Estocagem dos clones

As colônias brancas foram repicadas das placas de meio sólido com o auxílio de palitos de dente e foram inoculadas em microplacas de 96 poços para cultura de células (3799 - Costar Corning Inc.) contendo 100 μL /poço de meio Circle Grow (Bio 101 Systems) (40 g/L) e ampicilina (Calbiochem) (100 $\mu\text{g}/\text{mL}$) que então foram incubadas 20-24 horas em estufa a 37°C. Após o crescimento era então adicionado um volume de 52 μL /poço de glicerol 87% ($C_f = 30\%$). As placas eram então seladas com filme adesivo (Axygen), identificadas e embaladas em filme plástico e papel alumínio sendo então estocadas em um freezer a uma temperatura de -20°C .

2.4 Multiplicação de clones bacterianos em meio de cultura líquido

A partir de 10 μL de cultura de estoque (item 2.3) eram feitos inóculos para posterior extração de DNA plasmidial em pequena escala. As placas de estoque a serem inoculadas eram retiradas do freezer e colocadas em recipiente contendo gelo para evitar choque térmico. Os clones isolados eram crescidos em placas de 96 poços fundos (RK-SW Riplate) contendo 1 mL/poço de meio Circle Grow (40 g/L) e ampicilina (100 $\mu\text{g}/\text{mL}$). As placas eram seladas com filme adesivo no qual eram feitos dois pequenos orifícios sobre cada poço, para permitir uma melhor aeração durante a incubação. As bactérias eram então colocadas em uma incubadora de bancada por um período de 18-22 horas a 37°C com agitação circular de 320 rpm.

2.5 Extração de DNA plasmidial em pequena escala

1 mL de cultivo bacteriano (item 2.4) era utilizado para a extração de DNA plasmidial em pequena escala pelo método da lise alcalina (BIRNBOIM & DOLY, 1979). O método consiste essencialmente, na lise das células por um detergente (SDS) que expõe o DNA cromossômico a um pH alcalino (NaOH) causando a sua desnaturação seletiva e na neutralização do lisado pela adição de um ácido (Acetato de Potássio) que renatura a massa de DNA cromossômico na forma de uma rede insolúvel. Simultaneamente esta mudança de pH induz a precipitação de RNAs de alto peso molecular e proteínas complexadas ao SDS. Desta forma, as três principais moléculas contaminantes são co-precipitadas e podem ser removidas por centrifugação. O DNA plasmidial pode então ser recuperado a partir do sobrenadante por precipitação com etanol. Foi utilizado um protocolo semelhante ao descrito por BIRNBOIM & DOLY (1979), com adaptações ao uso de microplacas (Costar Corning Inc.). Foi incluída uma etapa de passagem por uma placa com filtro (MAGVN22, Millipore). O DNA plasmidial recuperado ao final do procedimento era ressuspensão em um volume de 30 μ L de H₂O MilliQ e armazenado no congelador a -4°C . O rendimento total das extrações variou de 1,5 a 6,0 μg (400-1600 fmol) de DNA plasmidial.

2.6 Eletroforese em gel de agarose

Anteriormente à reação de seqüenciamento, o produto da extração era submetido a uma análise por eletroforese em gel de agarose para quantificação do DNA. Para tanto, foi utilizado um sistema de eletroforese horizontal (Sunrise 96, Life Technologies – Gibco BRL) em gel submerso. A concentração de agarose dos géis utilizada foi de 0,8%, tendo estes sido preparados com 100 mL de TAE 1X. A concentração utilizada de EtBr, responsável pela coloração das duplas fitas de DNA, era de 0,5 $\mu\text{g}/\text{mL}$. Os géis mediam 12 cm x 24 cm x 3,5 mm e continham 104 canaletas dispostas em 4 fileiras de 26 cada. Eram verificadas 96 amostras

de duas placas diferentes (item 2.5) em cada gel. Em cada canaleta, era aplicado um volume final de 9 μL correspondente a 1 μL de DNA; 1,5 μL de tampão de amostra 6X ($C_f = 1X$) e 6,5 μL de H_2O deionizada. Utilizou-se como padrão de peso molecular massas de 100, 200 ou 400 ng do vetor pUC18 íntegro (2,6 kb). As voltagens aplicadas nas eletroforeses foram de 80 V ou 100 V em tempos que variaram de 60 a 90 minutos. Após a corrida os géis eram visualizados e fotografados sobre um transluminador de luz ultravioleta.

2.7 Seqüenciamento de DNA

2.7.1 Reação de seqüenciamento

A reação de seqüenciamento requer a mistura de um “pré-mix” de reação (8 μL), das amostras de DNA (11 μL) e de um “primer forward” ou “reverse” (1 μL) e envolve etapas cíclicas de desnaturação das amostras, anelamento dos “primers” e extensão/terminação da cadeia que são feitas de forma automática em um termociclador, utilizando-se um programa similar aos empregados em reações de PCR. No entanto, como é utilizado um único “primer” em cada reação, a amplificação do produto é linear em relação ao número de ciclos empregados. A preparação das reações do seqüenciamento automático seguiu os critérios do sistema DYEnamic Dye Terminator Cycle Sequencing Kit for MegaBACE (Amersham Biosciences). As reações foram realizadas em um termociclador automático (Eppendorf Mastercycler Gradient) utilizando-se o seguinte programa:

95°C por 20”

50°C por 15”

60°C por 1’

Repete 29X

4°C

Os produtos da reação eram precipitados com acetato de amônio (1 µL), etanol absoluto (27,5 µL) e etanol 70% (150 µL) para remoção de “terminadores” marcados não incorporados. O material precipitado era então finalmente dissolvido em um volume de 10 µL da solução de “carregamento” para separação e detecção no seqüenciador automático. Estes procedimentos também foram realizados segundo os critérios do sistema DYEnamic Dye Terminator Cycle Sequencing Kit for MegaBACE (Amersham Biosciences).

2.7.2 “Primers”

Os “primers” utilizados nas reações de seqüenciamento nos sentidos “forward” e “reverse” foram respectivamente, os oligonucleotídeos M13F e M13R (Invitrogen) que anelam às sequências promotoras de fago M13 que flanqueiam o sítio de policlonagem do vetor pSPORT1 (Invitrogen). A concentração dos “primers” utilizada foi de 3 µmol/µL. As seqüências destes oligonucleotídeos são as seguintes:

M13F – GTTTTCCCAGTCACGACGTTGTA

M13R – TCACACAGGAAACAGCTATGAC

2.7.3 Seqüenciamento automático

O seqüenciamento automático de DNA foi feito em uma máquina MegaBACE 1000 (Amersham Biosciences) tendo sido utilizado o sistema DYEnamic ET Dye Terminator (Amersham Biosciences). O sistema DYEnamic ET Dye Terminator é um método de seqüenciamento bastante sensível e robusto que está baseado em uma modificação do método tradicional de didesoxirribonucleotídeos descrito por SANGER *et al.* (1977). No sistema DYEnamic ET Dye Terminator, cada um dos quatro didesoxirribonucleotídeos (ddG, ddA, ddT e ddC) está ligado a duas moléculas marcadoras, fluoresceína e

uma de quatro diferentes rodaminas (rodamina 110, rodamina-6-G, tetrametil rodamina e rodamina X). A fluoresceína possui um grande coeficiente de extinção no comprimento de onda do laser de íon argônio (488 nm) utilizado pelo seqüenciador. Atuando como uma molécula marcada doadora, a fluoresceína absorve energia do laser e a transfere para a rodamina que é a molécula marcada acceptora na mesma molécula de didesoxirribonucleotídeo. Ao absorver a energia da fluoresceína, cada uma das rodaminas emite luz em um comprimento de onda que é característico, tornando possível a detecção e identificação de qual didesoxirribonucleotídeo terminou a extensão do fragmento de DNA. Os parâmetros utilizados durante as corridas foram:

- Voltagem de injeção das amostras: 2 kV
- Voltagem durante a corrida: 9 kV
- Tempo de injeção das amostras: 100 segundos
- Tempo de corrida: 120 minutos

2.8 Processamento, anotação e submissão das seqüências

Os eletroferogramas foram processados, anotados e submetidos automaticamente ao dbEST utilizando-se o software PartiGene (PARKINSON *et al.*, 2004a), um pacote de ferramentas integradas para análise de seqüências ESTs. O processamento inclui a verificação da qualidade dos cromatogramas e eliminação de seqüências de baixa qualidade, a remoção de seqüências do vetor, "primers", adaptadores, cauda de poli-adenilação, seqüência líder (SL) e do oligonuclotídeo GeneRacer (GR). O programa Phred atribui um valor de qualidade a cada base lida pelo seqüenciador que está relacionado à probabilidade de erro na identificação de uma base (EWING & GREEN, 1998). Um valor de qualidade de 20, por exemplo, significa que há chance de 1/100 de uma base ter sido identificada erroneamente. A anotação primária foi feita com base em seqüências homólogas de maior valor de similaridade presentes no banco de dados não redundante de seqüências de proteínas do NCBI, o nr (GenBank, PDB, SwissProt, PIR e PRF), identificadas através do algoritmo BLASTX (ALTSCHUL

et al., 1997). As seqüências contendo mais de 150 bases de alta qualidade (valor de Phred acima de 20) foram automaticamente submetidas ao dbEST. Paralelamente as seqüências foram disponibilizadas no LophDB, onde elas podem ser acessadas na forma de ESTs, “contigs” e “clusters” de seqüências.

2.9 Identificação e análise de seqüências relacionadas às diferentes subunidades do antígeno B

Para identificar as seqüências correspondentes às diferentes subunidades do AgB de *Echinococcus granulosus*, foi feito localmente, um experimento de TBLASTN (ver. 2.2.9) (ALTSCHUL *et al.*, 1997), que busca em um banco de dados, seqüências de nucleotídeos cuja seqüência predita de aminoácidos seja semelhante à seqüência de aminoácidos de uma proteína especificada pelo usuário. Foram utilizadas como banco de dados (“subject”), todas as seqüências ESTs das 6 bibliotecas enriquecidas com transcritos completos presentes no dbEST. As seqüências de aminoácido utilizadas como “sonda” (“query”) para identificação de seqüências relacionadas ao AgB e seus respectivos números de acesso no GenBank são: EgAgB8/1 (AF143813) (ROTT, *et al.*, 2000); EgAgB8/2 (L48620) (FERNÁNDEZ *et al.*, 1996); EgAgB8/3 (AF362442) (CHEMALE *et al.*, 2001); e EgAgB8/4 (AY357110) (AREND *et al.*, 2002) e (AY569365) (KAMENETZKY *et al.*, submetido a publicação). Foram utilizadas ainda as seqüências de aminoácidos correspondentes a dois clones genômicos R5 (EgAgB8/3) e Q7 (EgAgB8/5) isolados por HAAG *et al.* (2004).

3. Resultados e discussão

3.1 Triagem das colônias transformantes e estocagem de clones

Após plaqueamento em meio de cultura sólido (item 2.2) as bibliotecas PS/P-SL, CW-SL e PS-GR apresentaram respectivamente 82,2%; 81,9% e 71,5% de colônias brancas que foram repicadas individualmente e inoculadas em meio de cultura líquido em microplacas de cultura (item 2.3). Somando-se a estas as colônias das bibliotecas PS/P-GR e PS-SL (que não foram selecionadas previamente), um total de 7488 clones foi distribuído em 78 microplacas de estoque.

3.2 Extração de DNA plasmidial em pequena escala

O rendimento total das extrações variou de 1,5 a 6,0 μg (400-1600 fmol) de DNA plasmidial conforme a quantificação visual efetuada no gel de agarose (item 2.6). A Figura 2 ilustra o resultado deste procedimento. Este rendimento é suficiente para o seqüenciamento no sistema MegaBACE 1000, uma vez que a quantidade de DNA plasmidial recomendada pelo fabricante é de 80 a 800 fmol.

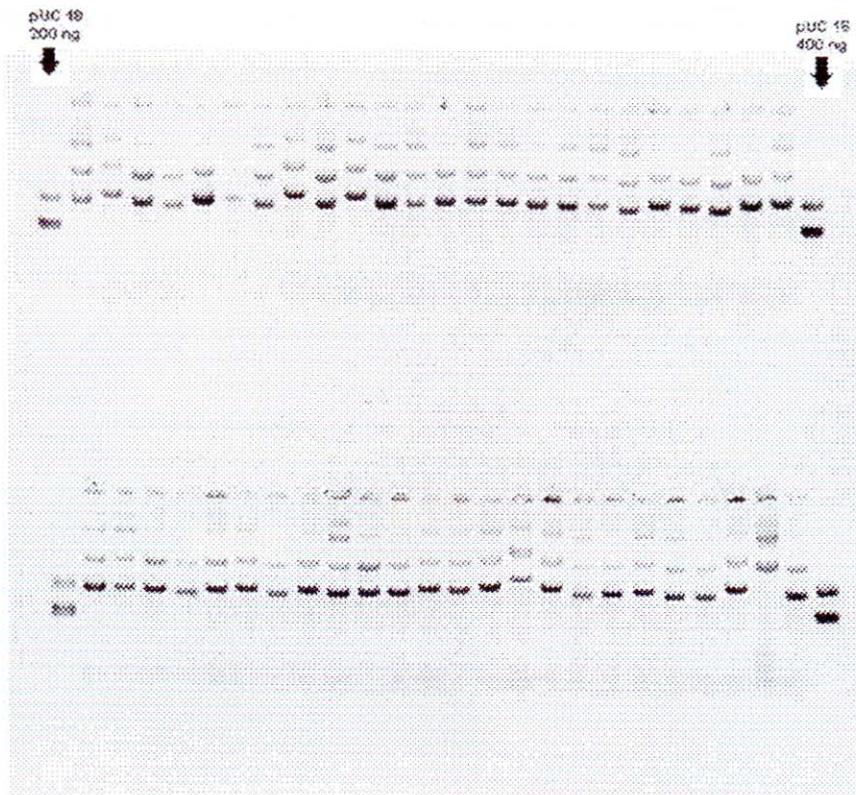


Figura 2. Eletroforese em gel de agarose dos produtos de extração de DNA plasmidial em pequena escala. Em um gel de 104 canaletas eram verificadas 96 amostras de duas placas diferentes. Em cada canaleta, era aplicado um volume de 1 μ L de DNA e 8 μ L de tampão de amostra 1X. Utilizou-se como padrão de peso molecular massas de 100, 200 ou 400 ng do vetor pUC18 íntegro (2,6 kb). A foto representa a metade de um gel.

3.3 Sequenciamento, processamento e submissão de ESTs de *Echinococcus granulosus*

Foi seqüenciado um total de 54 placas, que deram origem a 4869 seqüências ESTs. O tamanho das seqüências geradas foi satisfatório, a julgar pela análise individual dos cromatogramas gerados. O tamanho dos insertos clonados nas bibliotecas de protoescolícos variam de 350 a 1800 pb e nas bibliotecas de parede cística de 400 a 2800 pb (FERNÁNDEZ *et al.*, 2002). As seqüências geradas variam bastante em tamanho, ficando a maior proporção entre 500 e 800 pb, o que pode ser considerado um bom tamanho para

seqüências ESTs, que em geral não são maiores do que 700 pb (PARKINSON *et al.*, 2004b). A Figura 3 mostra, na tela do computador, a análise feita pelo programa MegaBACE Sequencing Scorecard, que avalia a qualidade obtida no seqüenciamento de uma placa com base no tamanho das seqüências lidas.

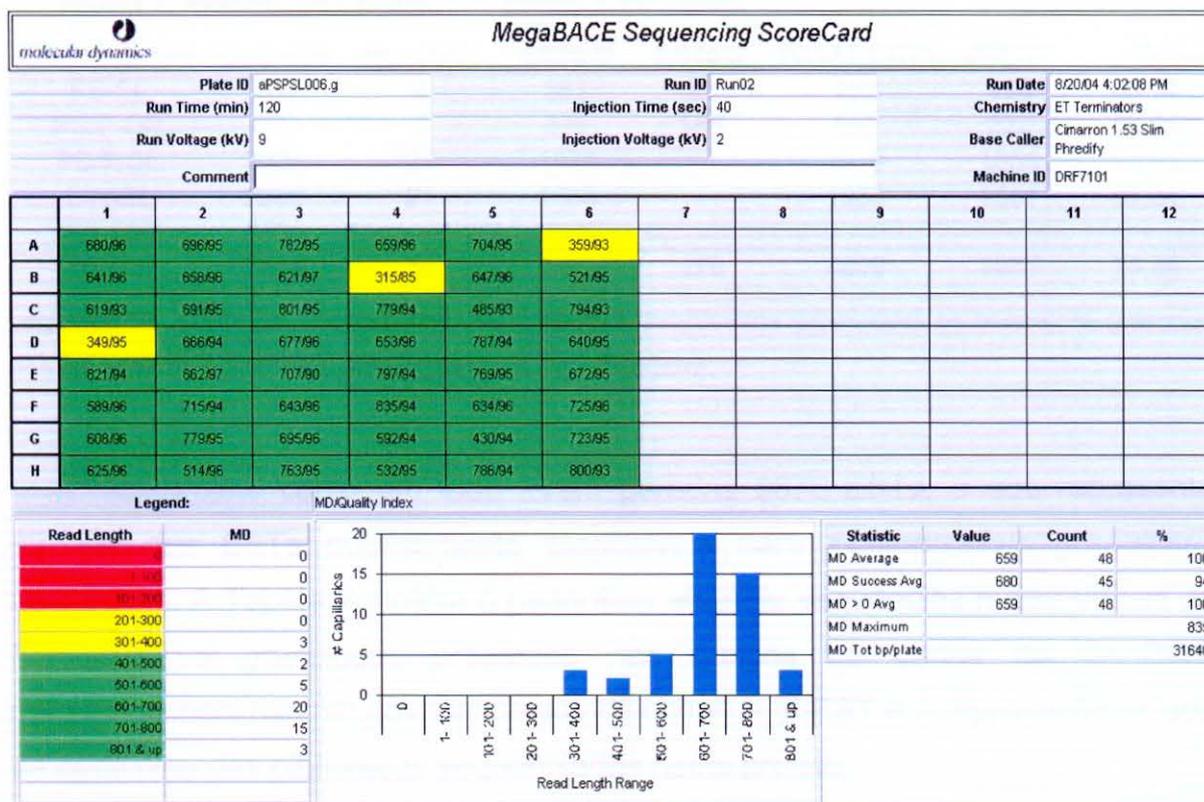


Figura 3. Análise feita pelo programa MegaBACE Sequencing Scorecard. O programa avalia a qualidade obtida no seqüenciamento de uma placa com base no intervalo de tamanho em que se encontram as seqüências geradas em cada reação individual de seqüenciamento.

Após o processamento, um total de 3651 ESTs (75%) apresentaram mais de 150 bases de alta qualidade (valor de Phred acima de 20) e foram depositadas no dbEST (<http://www.ncbi.nlm.nih.gov/dbEST/>). Este número representa uma parcela de 37% do total de seqüências ESTs de *Echinococcus granulosus* (9703) disponível hoje (08/12/2004) nesta base de dados. A Tabela 1 resume os resultados gerados pelo LBMC neste projeto. Da esquerda para a direita: as bibliotecas de cDNA, total de clones em estoque, total de placas seqüenciadas,

número de seqüências geradas com o “primer reverse”, número de seqüências geradas com o “primer forward”, total de seqüências geradas, total de seqüências submetidas ao dbEST e aproveitamento das seqüências (seqüências submetidas/seqüências geradas) em porcentagem.

Tabela 1. Resumo dos resultados gerados no LBMC.

BIBLIOTECAS	CLONES	PL. SEQ.	ESTs 5'	ESTs 3'	TOTAL ESTs	dbEST	%
PS-SL	768	4	361	91	452	195	43,14
PS/P-GR	672	5	370	185	555	272	49,01
PS/P-SL	2304	24	1979	-	1979	1653	83,53
CW-SL	2208	21	1883	-	1883	1531	81,31
PS-GR	1536	-	-	-	-	-	-
TOTAL	7488	54	4593	276	4869	3651	74,98

Legenda: PL. SEQ. (placas seqüenciadas); dbEST (seqüências submetidas ao banco); % (parcela das seqüências que foram submetidas, em porcentagem).

No projeto como um todo foram geradas 9678 ESTs, o que representa 99,74% das ESTs publicamente disponíveis para *Echinococcus granulosus* atualmente. A Tabela 2 mostra o panorama atual de seqüências nucleotídicas de *Echinococcus granulosus* presentes nos bancos de dados do GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) e dbEST e a representatividade de cada uma das bibliotecas seqüenciadas neste projeto.

Tabela 2. Panorama atual das seqüências publicamente disponíveis para *E. granulosus*.

PS/P-GR	1640
PS-GR	1482
PS/P-SL	1886
PS-SL	1385
CW-GR	1210
CW-SL	1836
PS	239
Outras ESTs	25
GenBank	526
TOTAL	10229

3.4 Identificação e análise de seqüências relacionadas às diferentes subunidades do antígeno B de *Echinococcus granulosus*

As ESTs foram anotadas automaticamente com base em homologia com seqüências de proteínas presentes em bases de dados públicas (item 2.8). Dessa forma, diversas seqüências de genes de interesse para o grupo foram identificadas (ex.: membros da família das actinas, proteínas de ligação ao cálcio, proteínas imunogênicas).

O AgB é uma molécula altamente antigênica presente no líquido hidático e que pode estar envolvido em mecanismos de evasão da resposta imune do hospedeiro. O potencial diagnóstico do AgB para a hidatidose cística tem sido explorado em nosso laboratório (ROTT *et al.*, 2000; VIRGINIO *et al.*, 2003). Devido a sua grande importância o AgB foi selecionado como primeiro alvo a ser analisado.

As ESTs relacionadas ao AgB encontradas com o algoritmo TBLASTN (ALTSCHUL *et al.*, 1997) foram classificadas manualmente, com base na homologia apresentada com as seqüências utilizadas como "sonda". A Figura 4 demonstra respectivamente os melhores alinhamentos gerados com este programa para as seqüências EgAgB8/3 (CHEMALE *et al.*, 2001), EgAgB8/4 (AREND *et al.*, 2002) e EgAgB8/1 (ROTT *et al.*, 2000). O alinhamento múltiplo das seqüências ESTs e as seqüências do antígeno B às quais são relacionadas está representado na Figura 5.

```

Query: EgAgB8/3 (Bligeno43)
Subject: Eg_PSGRS_11001_SP6
Length = 402
Score = 145 bits (367), Expect = 5e-037
Identities = 74/75 (98%), Positives = 74/75 (98%)
Frame = -1

Query: 1  FVVAHADDDEVTKTKKGVMKAISEIKHFFQSDPLGKKLVEVMKDVASVCEMVRKKAR 60
          FVVA  ADDDDDEVTKTKKGVMKAISEIKHFFQSDPLGKKLVEVMKDVASVCEMVRKKAR
Sbjct: 294 FVVARADDDEVTKTKKGVMKAISEIKHFFQSDPLGKKLVEVMKDVASVCEMVRKKAR 115

Query: 61  MALKEYVRKLVKEDE 75
          MALKEYVRKLVKEDE
Sbjct: 114 MALKEYVRKLVKEDE 70

Query: EgAgB8/4 (Arend)
Subject: Eg_CWGRS_03F02_SP6
Length = 312
Score = 161 bits (407), Expect = 1e-041
Identities = 81/81 (100%), Positives = 81/81 (100%)
Frame = -3

Query: 1  ALVAFVAVVQAKAEPERCKCLIMRKLGEIRDFFRSDPLGQKLAALGRDLTAICQKLOLKV 60
          ALVAFVAVVQAKAEPERCKCLIMRKLGEIRDFFRSDPLGQKLAALGRDLTAICQKLOLKV
Sbjct: 298 ALVAFVAVVQAKAEPERCKCLIMRKLGEIRDFFRSDPLGQKLAALGRDLTAICQKLOLKV 119

Query: 61  HEVLKKYVKDLLEEEDEDDLK 81
          HEVLKKYVKDLLEEEDEDDLK
Sbjct: 113 HEVLKKYVKDLLEEEDEDDLK 56

Query: EgAgB8/1
Subject: Eg_CWGRS_15E04_SP6
Length = 360
Score = 137 bits (346), Expect = 1e-034
Identities = 69/81 (85%), Positives = 78/81 (96%)
Frame = -1

Query: 1  MLLALALVSFVWVTQADDGLTSTSRSMVMKMFGEVYFFERDPLGQKLVDLLKEEEVFQL 60
          +LLALALVSFVWVTQADDGLTSTSRSM+MKM GE++YFFERDPLGQK+VDLLKEEEVFQ+
Sbjct: 312 LLLALALVSFVWVTQADDGLTSTSRSMKMLGEMRYFFERDPLGQKLVDLLKEEEVFQM 133

Query: 61  LRKKLRMALRSHLRGLIAEGE 81
          LRKKLR AL+SHLR L+AEG+
Sbjct: 132 LRKKLRTALKSHLRELVAEGK 70

```

Figura 4. Melhores alinhamentos gerados pelo algoritmo TBLASTN para as seqüências EgAgB8/3 (CHEMALE *et al.*, 2001), EgAgB8/4 (ARENDE *et al.*, 2002) e EgAgB8/1 (ROTT *et al.*, 2000). A seqüência "Query" é aquela que está sendo buscada no banco de dados. "Subject" é uma seqüência presente no banco que possui homologia com a seqüência buscada. "Length" é o tamanho (nt) da seqüência encontrada no banco. O valor de "Expect" é um parâmetro que descreve o número de seqüências com similaridade que podem ser encontradas ao acaso numa busca em um banco de dados de tamanho determinado. O "Score" é calculado em função dos parâmetros Karlin-Altschul. A expressão do escore em termos de bits é independente da matriz utilizada. "Identities" corresponde ao número de aminoácidos idênticos presentes no alinhamento das seqüências e "Positives" o número de aminoácidos com propriedades físico-químicas equivalentes. O valor do "Frame" é equivalente a posição da seqüência encontrada no alinhamento em relação às diferentes fases de leitura possíveis para a seqüência buscada.

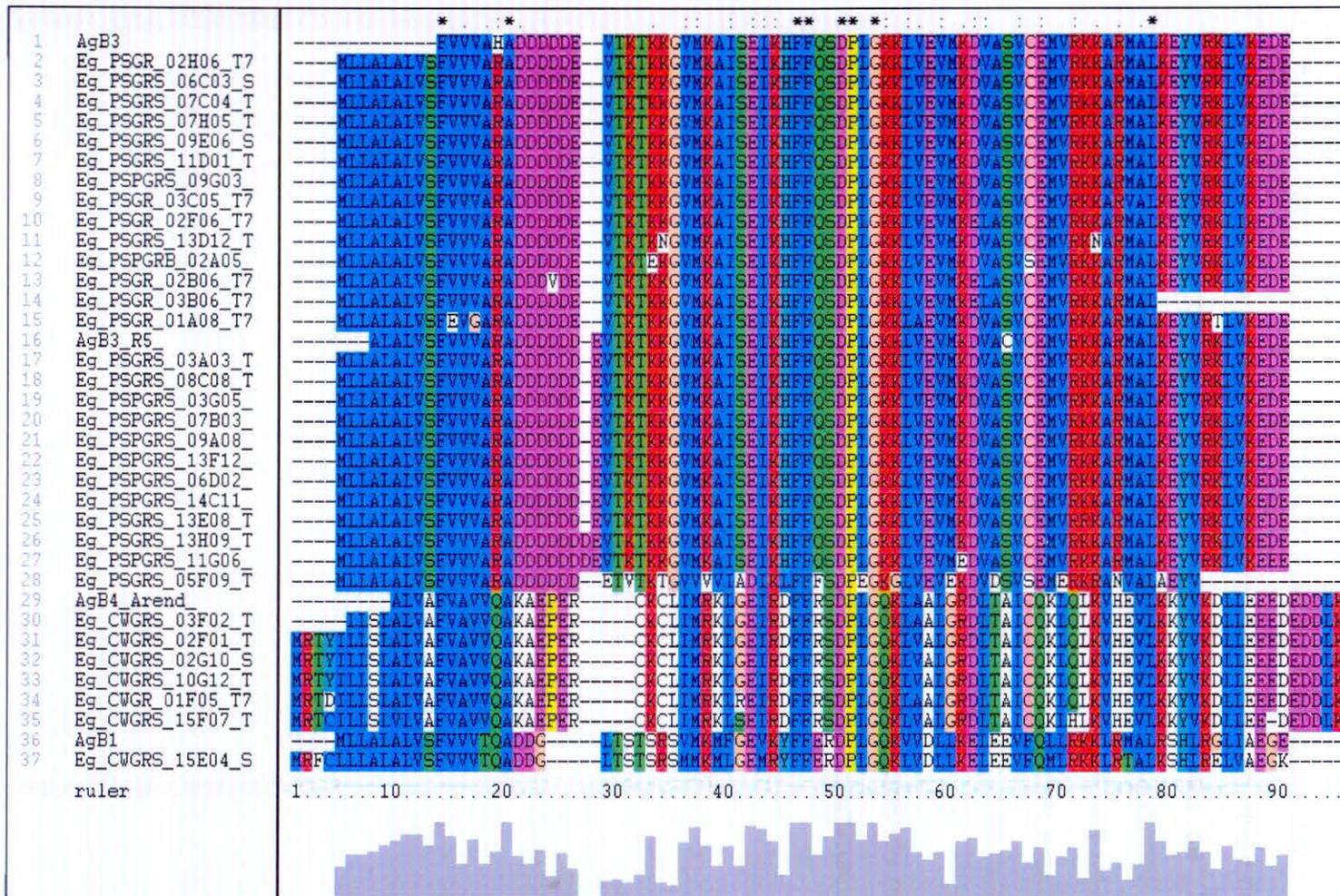


Figura 5. Alinhamento múltiplo entre as seqüências AgB3 (CHEMALE *et al.*, 2001), AgB3 (R5) (HAAG *et al.*, 2004), AgB4 (AREND *et al.*, 2002) e AgB1 (ROTT *et al.*, 2000) e as seqüências relacionadas a elas encontradas nas ESTs seqüenciadas. Alinhamento gerado através do programa Clustal X. O histograma ao pé da figura representa o grau de conservação encontrado para cada posição dos resíduos.

O resultado da análise comparativa das seqüências por homologia com seqüências do AgB indicam que estão sendo expressas seqüências relacionadas ao antígeno B nas três bibliotecas de cDNA que foram amplificadas a partir do oligonucleotídeo GeneRacer (PS/P-GR, PS-GR e CW-GR). Porém, nenhuma seqüência com homologia significativa foi encontrada nas bibliotecas cujos cDNAs foram amplificados a partir da seqüência "spliced-leader" (SL). Este resultado era esperado, uma vez que, até o momento não foram identificadas seqüências relacionadas ao antígeno B associadas ao processo de "trans-splicing" em *Echinococcus granulosus*. A Tabela 3 apresenta os resultados mais significativos para cada uma das buscas. Foram encontrados 33 clones relacionados ao AgB: 26 relacionados a AgB3, sendo que 14 deles apresentam maior homologia com a seqüência descrita por CHEMALE *et al.*, 2001 (AgB3) e as outras 12 seqüências sendo mais semelhantes a AgB3 (R5) (HAAG *et al.*, 2004); 6 seqüências relacionadas a AgB4 (AREND *et al.*, 2002) e uma seqüência relacionada a AgB1 (ROTT *et al.*, 2000). A tabela informa os valores dos parâmetros calculados pelo algoritmo TBLASTN.

Tabela 3. Listagem dos clones relacionados ao antígeno B de *E. granulosus* encontrados nas ESTs.

Clones	Tam. (nt)	Esc. (bits)	Valor E	Ident.	Posit.	Abert.	Fase	Sub.
Eg_PSGR_02H06_T7	379	145	5,00E-37	74/75	74/75	0	+2	AgB3
Eg_PSGRS_06C03_SP6	377	145	5,00E-37	74/75	74/75	0	-3	AgB3
Eg_PSGRS_07C04_T7	375	145	5,00E-37	74/75	74/75	0	+1	AgB3
Eg_PSGRS_07H05_T7	375	145	5,00E-37	74/75	74/75	0	+1	AgB3
Eg_PSGRS_09E06_SP6	375	145	5,00E-37	74/75	74/75	0	-1	AgB3
Eg_PSGRS_11D01_T7	375	145	5,00E-37	74/75	74/75	0	+1	AgB3
Eg_PSPGRS_09G03_T7	375	145	5,00E-37	74/75	74/75	0	+1	AgB3
Eg_PSGR_03C05_T7	375	144	1,00E-36	73/75	74/75	0	+1	AgB3
Eg_PSGR_02F06_T7	375	142	4,00E-36	71/75	74/75	0	+1	AgB3
Eg_PSGRS_13D12_T7	404	142	7,00E-36	72/75	72/75	0	+3	AgB3
Eg_PSPGRB_02A05_M13R	373	140	2,00E-35	72/75	73/75	0	+2	AgB3
Eg_PSGR_02B06_T7	373	139	4,00E-35	71/75	73/75	0	+1	AgB3
Eg_PSGR_03B06_T7	373	139	5,00E-35	70/75	73/75	0	+1	AgB3
Eg_PSGR_01A08_T7	367	137	2,00E-34	70/75	70/75	0	+2	AgB3
Eg_PSGRS_03A03_T7	381	157	2,00E-40	81/82	81/82	0	+3	AgB3(R5)
Eg_PSGRS_08C08_T7	380	157	2,00E-40	81/82	81/82	0	+3	AgB3(R5)
Eg_PSPGRS_03G05_SP6	378	157	2,00E-40	81/82	81/82	0	-1	AgB3(R5)
Eg_PSPGRS_07B03_T7	382	157	2,00E-40	81/82	81/82	0	+2	AgB3(R5)
Eg_PSPGRS_09A08_T7	381	157	2,00E-40	81/82	81/82	0	+3	AgB3(R5)
Eg_PSPGRS_13F12_T7	381	157	2,00E-40	81/82	81/82	0	+3	AgB3(R5)
Eg_PSPGRS_06D02_T7	372	157	2,00E-40	81/82	81/82	0	+1	AgB3(R5)
Eg_PSPGRS_14C11_T7	383	157	2,00E-40	81/82	81/82	0	+3	AgB3(R5)
Eg_PSGRS_13E08_T7	380	156	4,00E-40	80/82	81/82	0	+3	AgB3(R5)
Eg_PSGRS_13H09_T7	385	152	4,00E-39	81/83	81/83	1/83	+3	AgB3(R5)
Eg_PSPGRS_11G06_T7	383	149	3,00E-38	79/83	81/83	1/83	3	AgB3(R5)
Eg_PSGRS_05F09_T7	383	90.5	3,00E-20	54/78	59/78	0	+3	AgB3(R5)
Eg_CWGRS_03F02_T7	312	161	1,00E-41	81/81	81/81	0	+3	AgB4(Arend)
Eg_CWGRS_02F01_T7	387	159	3,00E-41	80/81	80/81	0	+1	AgB4(Arend)
Eg_CWGRS_02G10_SP6	387	15	3,00E-41	80/81	80/81	0	-3	AgB4(Arend)
Eg_CWGRS_10G12_T7	387	159	3,00E-41	80/81	80/81	0	+3	AgB4(Arend)
Eg_CWGR_01F05_T7	390	158	1,00E-40	80/81	80/81	0	+2	AgB4(Arend)
Eg_CWGRS_15F07_T7	366	147	2,00E-37	76/80	76/80	0	+2	AgB4(Arend)
Eg_CWGRS_15E04_SP6	360	137	1,00E-34	69/81	78/81	0	-1	AgB1

Aparentemente, estão sendo transcritas predominantemente pela camada germinativa da parede cística moléculas relacionadas à subunidade EgAgB8/4, uma vez que, à exceção de uma seqüência relacionada à subunidade EgAgB8/1, as demais ESTs encontradas nas bibliotecas CW apresentam seqüências preditas de aminoácidos que assemelham-se bastante à seqüência descrita por AREND *et al.*, (2002).

As seqüências ESTs relacionadas a subunidade EgAgB8/4 confirmam a seqüência de aminoácidos predita para a extremidade amino-terminal do primeiro éxon da seqüência descrita por KAMENETZKY *et al.* (submetido a publicação).

Estas seqüências apresentam variações, como inserções/deleções e substituições de aminoácidos principalmente no segundo éxon.

Em protoescólices por sua vez, pôde-se confirmar apenas a transcrição de moléculas relacionadas à subunidade EgAgB8/3. Foi identificado um número considerável de transcritos cujas traduções das seqüências apresentam valores altos de similaridade com as seqüências de aminoácido das seqüências AgB3 (CHEMALE *et al.*, 2001) e AgB3 (R5) (HAAG *et al.*, 2004). As seqüências ESTs relacionadas a subunidade EgAgB8/3 confirmam a seqüência de aminoácidos predita para a extremidade amino-terminal do primeiro éxon da seqüência descrita por CHEMALE (comunicação pessoal). As seqüências que formam este grupo também apresentam variações, como inserções/deleções e substituições de aminoácidos na região correspondente ao segundo éxon. Algumas ESTs 5' e 3' seqüenciadas a partir de um mesmo clone que resultam em seqüências idênticas de aminoácidos suportam a hipótese de que esteja ocorrendo variabilidade em nível de seqüência de aminoácidos. A presença de polimorfismos do mesmo locus gênico em diferentes tipos celulares de um mesmo cisto, ou entre, cistos diferentes já foi observada para seqüências do AgB (HAAG *et al.*, 2004; AREND *et al.*, 2004). No entanto, uma outra hipótese, a de que esta variabilidade pode ser devida a artefatos introduzidos durante a clonagem e/ou o seqüenciamento, não pode ser descartada. Erros no seqüenciamento poderão ser verificados pela análise individual dos cromatogramas que contêm a informação de qualidade de leitura de cada base presente na seqüência.

Foi identificado apenas um clone cuja seqüência predita de aminoácidos, é relacionada à subunidade EgAgB8/1. Esta seqüência difere da seqüência de aminoácidos utilizada como "sonda" para esta subunidade principalmente nas suas extremidades amino e carboxi-terminais, apresentando, inclusive, quatro resíduos a mais na extremidade amino-terminal. Não foram encontradas, porém, quaisquer seqüências relacionadas à subunidade EgAgB8/2 ou EgAgB8/5. A aparente ausência de transcrição de seqüências relacionadas à subunidade EgAgB8/2 nesta "população" de RNAs mensageiros é um achado intrigante, uma vez que já foi demonstrado que esta subunidade, juntamente com EgAgB8/1 é um

dos componentes das subunidades maiores (16 e 24 kDa) e do AgB nativo (GONZÁLEZ *et al.*, 1996).

Os dados encontrados podem ser indicativos de um padrão de expressão de subunidades do AgB, que do ponto de vista qualitativo é bastante informativo, no entanto, do ponto de vista quantitativo os dados devem ser interpretados com uma maior cautela. O perfil transcricional apresentado pelas diferentes bibliotecas pode ser afetado por uma certa tendenciosidade, uma vez que a chance que diferentes seqüências possuem de serem amplificadas e/ou clonadas não são as mesmas. Desta forma, mesmo que uma seqüência não esteja representada nas bibliotecas de cDNA, a sua ausência não pode ser extrapolada para o nível biológico.

A presença de seqüências que apresentam um alto grau de similaridade com as seqüências de aminoácido das subunidades EgAgB8/3, EgAgB8/4 e EgAgB8/1 indica que moléculas relacionadas a estas subunidades estão sendo transcritas no metacestódeo. A transcrição de EgAgB8/1, EgAgB8/3 e EgAgB8/4 em protoescólices já havia sido demonstrada anteriormente (FERNÁNDEZ *et al.*, 1996; CHEMALE *et al.*, 2001; AREND *et al.*, 2004). No entanto, neste trabalho foram também identificadas seqüências que foram isoladas a partir da camada germinativa de cistos hidáticos (FERNÁNDEZ *et al.*, 2002), indicando possivelmente expressão diferencial de subunidades do AgB em tecidos diferentes do metacestódeo. Serão necessárias novas abordagens para confirmar estes dados.

4. Perspectivas

Ficam como perspectivas de continuidade desse trabalho as análises globais das seqüências geradas. Serão comparadas as seqüências das bibliotecas GR e SL para determinar a real parcela de transcritos que utilizam a seqüência “spliced leader” no genoma de *E. granulosus*. A comparação entre a biblioteca de protoescólices tratados com pepsina com as bibliotecas de protoescólices que não sofreram o tratamento informará sobre genes que possivelmente têm sua transcrição ativada no momento ruptura das cápsulas prolíferas e que tem papel importante na fixação ao intestino delgado e na diferenciação dos protoescólices em vermes adultos no hospedeiro definitivo. Ainda, a comparação entre as bibliotecas de protoescólices e de parede cística, fornecerá dados sobre genes que possivelmente sofrem expressão diferencial nos tecidos diferenciados que compõem o metacestódeo. Serão feitas ainda análises comparativas do transcrito de *E. granulosus* com o de outros platelmintos como *Schistosoma mansoni* e *Mesocestoides corti* (cestódeo utilizado como modelo em desenvolvimento e biologia molecular em nosso laboratório) e de outros organismos cujo genoma já está bem caracterizado (*Drosophila*, *Saccharomyces*). Outras análises envolverão a pesquisa em diversos bancos de dados de domínio público de nucleotídeos, proteínas e domínios funcionais visando a caracterização estrutural e funcional de genes de *E. granulosus*. Em relação às seqüências relacionadas ao antígeno B poderão ser feitas análises qualitativas e quantitativas de transcritos individuais, utilizando-se as seqüências identificadas neste trabalho para projetar “primers” para serem utilizados em experimentos de PCR em tempo real. Poderá-se ainda, utilizar estas seqüências como sondas em experimentos de “Southern blot”, ou para a projeção de “primers” para o isolamento de seqüências genômicas.

5. Referências bibliográficas

1. ADOUTTE, A.; BALAVOINE, G.; LARTILLOT, N.; LESPINET, O.; PRUD'HOMME, B. & DE ROSA, R. The new animal phylogeny: reliability and implications. *Proc. Natl. Acad. Sci. U. S. A.*, 97:4453-4456, 2000.
2. ALMEIDA, R.; NORRISH, A.; LEVICK, M.; VETRIE, D.; FREEMAN, T.; VILO, J.; IVENS, A.; LANGE, U.; STOBER, C.; MCCANN, S. & BLACKWELL, J. M. From genomes to vaccines: *Leishmania* as a model. *Phil. Trans. R. Soc. Lond. B.*, 357:5-11, 2002.
3. ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W. & LIPMAN D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 25(17):3389-3402, 1997.
4. ARAMBULO, P. Public health importance of cystic echinococcosis in Latin America. *Acta Trop.*, 67: 113-124, 1997.
5. AREND, A.C.; ALVES-JUNIOR, L.; CHEMALE, G.; ZAHA, A.; HAAG, K.L. (2002) Cloning and characterization of sequences encoding an *Echinococcus granulosus* antigen (AgB) of diagnostic value. *Am. J. Trop. Med. Hyg.*, 67:299-300, 2002 (Supplement).
6. AREND, A. C.; ZAHA, A.; AYALA, F. J. & HAAG, K. L. The *Echinococcus granulosus* antigen B shows a high degree of genetic variability. *Exp. Parasitol.* 108: 76-80, 2004.
7. BARBIERI, M.; FERNÁNDEZ, V.; GONZÁLEZ-SAPIENZA, G.; LUACES, V. M. & NIETO, A. Diagnostic evaluation of a synthetic peptide derived from a novel antigen B subunit as related to other available peptides and native antigens used for serology of cystic hydatidosis. *Parasite Immunol.*, 20:51-61, 1998.
8. BARNES, R. D. *Invertebrate zoology*. 4a ed. Saunders, Philadelphia, 1980.
9. BARRET, J.; SAGUIR, N.; TIMANOVA, A.; CLARKE, K.; BROPHY, P. M. Characterization and properties of an intracellular lipid-binding protein from the tapeworm *Moniezia expansa*. *Eur. J. Biochem.*, 250:269-275, 1997.
10. BIRNBOIM, H. C. & DOLY, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, 7 (6):1513-1523, 1979.
11. BLAXTER, M.; DAUB, J.; GUILIANO, D.; PARKINSON, J.; WHITTON, C. & THE FILARIAL GENOME PROJECT. The *Brugia malayi* genome project: expressed sequence tags and gene discovery. *Trans. R. Soc. Trop. Med. Hyg.* 96(1):7-17, 2002.

12. BREHM, K.; JENSEN, K. & FROSCH, M. mRNA trans-splicing in the human parasitic cestode *Echinococcus multilocularis*. *J. Biol. Chem.*, 275: 38311-38318, 2000.
13. CAMERON, T. W. M. & WEBSTER, G. A. The histogenesis of the hydatid cyst (*Echinococcus spp.*). Part 1. Liver cysts in large mammals. *Can. J. Zool.*, 47:1405-10, 1969.
14. CHEMALE, G. Caracterização do antígeno B e análise proteômica de *Echinococcus granulosus*. Tese de doutorado. Programa de Pós-graduação em Biologia Molecular e Celular. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.
15. CHEMALE, G.; HAAG, K. L.; FERREIRA, H. B.; ZAHA, A. *Echinococcus granulosus* antigen B is encoded by a gene family. *Mol. Biochem. Parasitol.* 116: 233-237, 2001.
16. DE LA RUE, M. L.; OSAIDA, I.; SILVEIRA, F. T. & MOREIRA, W. S. Human hydatidosis in the population of endemic regions of the Rio Grande do Sul state, Brazil. *Rev. Latinoam. Microbiol.*, 36: 205-206, 1994.
17. DEGRAVE, W. M.; MELVILLE, S.; IVENSC, A. & ASLETT, M. Parasite genome initiatives. *Int. J. Parasitol.*, 31:532-536, 2001.
18. DÍAZ, A.; IRIGOÍN, F.; FERREIRA, A. M. & SIM, R. B. Control of host complement activation by the *Echinococcus granulosus* hydatid cyst. *Immunopharmacology*, 42: 91-98, 1999.
19. ECKERT, J.; GEMMEL, M. A.; MESLIN, F. X.; PAWLOWSKI, Z. S. *Manual on Echinococcosis in humans and animals: a public health problem of global concern*. Paris, WHO/OIE, 265pp., 2001.
20. EWING, B. & GREEN, P. Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Res.*, 8:186-194, 1998.
21. FARIAS, L. N.; MALGOR, R.; CASSARAVILLA, C.; BRAGANÇA, C. & DE LA RUE, M. L. Echinococcosis in southern Brazil: efforts toward implementation of a control program in Santana do Livramento, Rio Grande do Sul. *Rev. Inst. Med. trop. S. Paulo*, 46(3):153-156, 2004.
22. FERNÁNDEZ, C.; GREGORY, W. F.; LOKE, P.; MAIZELS, R. M. Full-length-enriched cDNA libraries from *Echinococcus granulosus* contain separate populations of oligo-capped and trans-spliced transcripts and a high level of predicted signal peptide sequences. *Mol. Biochem. Parasitol.*, 122(2):171-180, 2002.
23. FERNANDÉZ, V.; FERREIRA, H. B.; FERNÁNDEZ, C.; ZAHA, A. & NIETO A. Molecular characterization of a novel 8 kDa subunit of *Echinococcus granulosus* antigen B. *Mol. Biochem. Parasitol.*, 77:247-250, 1996.

24. FERREIRA, A. M.; IRIGOÍN, F.; BREIJO, M.; SIM, R. B. & DÍAZ, A. How *Echinococcus granulosus* deals with complement. *Parasitol. Today*, 16(4): 168-172, 2000.
25. FRAYHA, G. J. & HADDAD, R. Comparative chemical composition of protoscoleces and hydatid cyst fluid of *Echinococcus granulosus* (Cestoda). *Int. J. Parasitol.*, 10:359-364, 1980.
26. FROSCH, P.; HARTMANN, M.; MÜHLSCHLEGEL, F. & FROSCH, M. Sequence heterogeneity of the echinococcal antigen B. *Mol. Biochem. Parasitol.*, 64:171-175, 1994.
27. GUTIERREZ, J. A. Genomics: from novel genes to new therapeutics in parasitology. *Int. J. Parasitol.*, 30:247-252, 2000.
28. GEMMEL, M. A.; ROBERTS, M. G.; BEARD, T. C.; LAWSON, J. R. Quantitative epidemiology and transmission dynamics with special reference to *Echinococcus granulosus*. In: *Manual on Echinococcosis in humans and animals: a public health problem of global concern*. ECKERT, J.; GEMMEL, M. A.; MESLIN, F. X.; PAWLOWSKI, Z. S. Paris, WHO/OIE, pp.143-156, 2001.
29. GONZÁLEZ, G.; NIETO, A.; FERNÁNDEZ, C.; ÖRN, A.; WERNSTEDT, C. & HELLMAN, U. Two different 8 kDa monomers are involved in the oligomeric organization of the native *Echinococcus granulosus* antigen B. *Parasite Immunol.*, 18:587-596, 1996.
30. GONZÁLEZ-SAPIENZA, G.; LORENZO, C. & NIETO, A. Improved immunodiagnosis of cystic hydatid disease by using a synthetic peptide with higher diagnostic value than that of its parent protein, *Echinococcus granulosus* antigen. *B. J. Clin. Microbiol.*, 38(11):3979-83, 2000.
31. HAAG, L. L.; ALVES-JUNIOR, L.; ZAHA, A. & AYALA, F. J. Contingent, non-neutral evolution in a multicellular parasite: natural selection and gene conversion in the *Echinococcus granulosus* antigen B gene family. *Gene*, 333:157-167, 2004.
32. 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(5):695-702, 1998.
33. HOBERG, E. P. *Taenia* tapeworms: their biology, evolution and socioeconomic significance. *Microbes Infect.*, 4:859-866, 2002.
34. KAMMERER, W. S.; SCHANTZ, P. M. Echinococcal disease. *Infect. Dis. Clin. North. Am.*, 7:605, 1993.
35. KLOETZEL, K. & PEREIRA, J. A. Human hydatidosis in Rio Grande do Sul (Brazil): an estimate of its significance for the public health of the country. *Rev. Inst. Med. Trop. Sao Paulo*, 34:549-555, 1992.

36. LIGHTOWLERS, M. W. & GOTTSTEIN, B. Echinococcosis/hydatidosis: antigens, immunological and molecular diagnosis. In: Thompson, R. C. A. & Lymbery, A. J. *Echinococcus* and hydatid disease. Wallingford, CAB International, pp. 335-410, 1995.
37. LIGHTOWLERS, M. W.; LIU, D.; HARALAMBOUS, A. & RICKARD, M. D. Subunit composition and specificity of the major cyst fluid antigens of *Echinococcus granulosus*. *Mol. Biochem. Parasitol.*, 37: 171-182, 1989.
38. MCMANUS, D. P.; LEA, T. L. & BLAIR, D. Genomics of parasitic flatworms. *Int. J. Parasitol.*, 34:153-158, 2004.
39. MCMANUS, D. P.; ZHANG, W.; LI, J. & BARTLEY, P.B. Echinococcosis. *Lancet Infect. Dis.*, 362:1295-304, 2003.
40. MALLATT, J. & WINCHELL, C. J. Testing the new animal phylogeny: first use of combined large-subunit and small-subunit rRNA gene sequences to classify the protostomes. *Mol. Biol. Evol.*, 19:289-301, 2002.
41. ORIOL, C. & ORIOL, R. Physicochemical properties of a lipoprotein antigen of *Echinococcus granulosus*. *Am. J. Trop. Med. Hyg.*, 24(1):96-100, 1975.
42. ORIOL, R.; WILLIAMS, J. F.; ESANDI, M. V. P & ORIOL, C. Purification of lipoprotein antigens of *Echinococcus granulosus* from sheep hydatid fluid. *Am. J. Trop. Med. Hyg.*, 20(4): 569-575, 1971.
43. ORTONA, E.; RIGANO, R.; BUTTARI, B.; DELUNARDO, F.; IOPPOLO, S.; MARGUTTI, P.; PROFUMO, E.; TEGGI, A.; VACCARI, S. & SIRACUSANO, A. An update on immunodiagnosis of cystic echinococcosis. *Acta Trop.*, 85:165-171, 2003.
44. PARKINSON, J.; ANTHONY, A.; WASMUTH, J.; SCHIMID, R.; HEDLEY, A. & BLAXTER, M. PartiGene - constructing partial genomes. *Bioinformatics*, 20 (9):1398-1404, 2004.
45. PARKINSON, J.; GUILIANO, D. B. & Blaxter, M. Making sense of EST sequences by CLOBBing them. *BMC Bioinformatics*, 3(31), 2002.
46. PARKINSON, J.; WHITTON, C.; SCHIMID, R.; THOMSON, M. & BLAXTER, M. NEMBASE: a resource for parasitic nematode ESTs. *Nucleic Acids Res.*, 32: 427-430, 2004.
47. PATEL, N. H. The ancestry of segmentation. *Dev. Cell.*, 5(1):2-4, 2003.
48. RIGANO, R.; PROFUMO, E.; BRUSCHI, F.; CARULLI, G.; AZZARA, A.; IOPPOLO, S.; BUTTARI, B.; ORTONA, E.; MARGUTTI, P.; TEGGI, A.; SIRACUSANO, A. Modulation of human immune response by *Echinococcus*

- granulosus* antigen B and its possible role in evading host defenses. *Infect. Immun.*, 69:288-96, 2001.
49. ROGAN, M. T. & RICHARS, K. S. *Echinococcus granulosus* changes in the surface ultrastructure during protoscolex formation. *Parasitology*, 94:359-367, 1987.
50. ROMIG, T. Epidemiology of echinococcosis. *Langenbecks Arch. Surg.*, 388:209-217, 2003.
51. ROTT, M. B.; FERNÁNDEZ, V.; FARIAS, S.; CENI, J.; FERREIRA, H. B.; HAAG, K. L. & ZAHA, A. Comparative analysis of two different subunits of antigen B from *Echinococcus granulosus*: gene sequences, expression in *Escherichia coli* and serological evaluation. *Acta Trop.* 75:331-340, 2000.
52. SAMBROOK, J.; RUSSELL, D. W.; SAMBROOK, J. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. 2001. 999 pg.
53. SAGHIR, N.; CONDE, P. J.; BROPHY, P. M.; BARRET, J. A new diagnostic tool for neurocysticercosis is a member of a cestode specific hydrophobic ligand binding protein family. *FEBS Lett.*, 487:181-184, 2000.
54. SAKO, Y.; NAKAO, M.; IKEJIMA, T.; PIAO, X. Z.; NAKAYA, K.; ITO, A. Molecular characterization and diagnostic value of *Taenia solium* low-molecular weight antigen genes. *J. Clin. Microbiol.*, 38(12):4439-4444, 2000.
55. SANCHEZ, F. A. & SANCHEZ, A. C. Estudio de algunas propiedades físicas y componentes químicos del líquido y pared germinativa de quistes hidáticos de diversas especies y de diferentes localización. *Rev. Ib. Parasitol.*, 31:347-366, 1971.
56. SHEPERD, J. C.; AITKEN, A. & MCMANUS, D. P. A protein secreted in vivo by *Echinococcus granulosus* inhibits elastase activity and neutrophil chemotaxis. *Mol. Biochem. Parasitol.*, 44: 81-90, 1991.
57. STÜRZENBAUM, S. R.; PARKINSON, J.; BLAXTER, M. ; MORGAN, J.; KILLE, P. & GEORGIEV, O. The earthworm Expressed Sequence Tag project. *Pedobiologia*, 47, 2003.
58. THOMPSON R. C. A. The development of brood capsules and protoscolexes in secondary hydatid cyst of *Echinococcus granulosus*. A histological study. *Z. Parasitenk.*, 51:31-26, 1976.
59. THOMPSON, R. C. A. & LYMBERY, A. J. *Echinococcus* and hydatid disease. Wallingford, CAB International, 477pp., 1995.
60. THOMPSON, R. C. A. & LYMBERY, A. J. *Echinococcus*: biology and strain variation. *Int. J. Parasitol.*, 20:457-470, 1990.

61. THOMPSON, R. C. A. & LYMBERY, A. J. The nature, extent and significance of variation within the genus *Echinococcus*. *Adv. Parasitol.*, 27:209–258, 1988.
62. TORGERSON, P. R. & BUDKE, C. M. Echinococcosis – an international public health challenge. *Res. Vet. Sci.*, 74: 191–202, 2003.
63. TORGERSON, P. R.; CARMONA C. & BONIFACINO, R. Estimating the economic effects of cystic echinococcosis: Uruguay, a developing country with upper-middle income. *Ann. Trop. Med. Parasitol.*, 94(7):703-13, 2000.
64. TORGERSON, P. R. & DOWLING, P. M. Estimating the economic effects of cystic echinococcosis. Part 2: an endemic region in the United Kingdom, a wealthy, industrialized economy. *Ann. Trop. Med. Parasitol.*, 95(2):177-85, 2001.
65. TORGERSON, P. R., DOWLING, P. M. & ABO-SHEHADA, M. N. Estimating the economic effects of cystic echinococcosis. Part 3: Jordan, a developing country with lower-middle income. *Ann. Trop. Med. Parasitol.*, 95(6):595-603, 2000.
66. VERJOVSKI-ALMEIDA, S.; DEMARCO, R.; MARTINS, E. A.; GUIMARÃES, P. E.; OJOPI, E. P.; PAQUOLA, A. C.; PIAZZA,.; NISHIYAMA, M. Y. JR; KITAJIMA, J. P.; ADAMSON, R. E.; ASHTON, P. D.; BONALDO, M. F.; COULSON, P. S.; DILLON, G. P.; FARIAS, L. P.; GREGORIO, S. P.; HO, P. L.; LEITE, R. A.; MALAQUIAS, L. C.; MARQUES, R. C.; MIYASATO, P. A.; NASCIMENTO, A. L.; OHLWEILER, F. P.; REIS, E. M.; RIBEIRO, M. A.; SA, R. G.; STUKART, G. C.; SOARES, M. B.; GARGIONI, C.; KAWANO, T.; RODRIGUES, V.; MADEIRA, A. M.; WILSON, R. A.; MENCK, C. F.; SETUBAL, J. C.; LEITE, L. C. & DIAS-NETO, E. Transcriptome analysis of the acelomate human parasite *Schistosoma mansoni*. *Nat. Genet.* 35:148-157, 2003.
67. VIRGINIO, V. G.; HERNÁNDEZ, A.; ROTT, M. B.; MONTEIRO, K. M.; ZANDONAI, A. F.; NIETO, A.; ZAHA, A & FERREIRA, H. B. A set of recombinant antigens from *Echinococcus granulosus* with potential for use in the immunodiagnosis of human cystic hydatid disease. *Clin. Exp. Immunol.*, 132: 309–315, 2003.
68. ZARLENGA, D. S.; RHOADS, M. L.; AL-YAMAN, F. M. A *Taenia crassiceps* cDNA sequence encoding a putative immunodiagnostic antigen for bovine cysticercosis. *Mol. Biochem. Parasitol.*, 67:215-223, 1994.