

Universidade Federal do Rio Grande do Sul
Instituto de Biociências
Departamento de Genética
Programa de Pós-graduação em Genética e Biologia Molecular

Validação de genes normalizadores em *Echinococcus granulosus* s.s. (G1) e *Echinococcus ortleppi* para estudos de expressão gênica através de PCR em tempo real

Sergio Martin Espínola

Dissertação submetida ao Programa de Pós-graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para obtenção do grau de Mestre

Orientador: Dr. Arnaldo Zaha

Porto Alegre, Março de 2014.

Agradecimentos

Ao professor Arnaldo Zaha, pela orientação, confiança e constante acompanhamento durante o trabalho, sendo parte fundamental nesta etapa da minha carreira.

Ao professor Henrique B. Ferreira, pela participação na minha formação.

Ao professor Rogerio Margis, por suas sugestões na análise de dados.

Aos membros da banca pela sua disponibilidade e comprometimento na correção deste trabalho.

Aos colegas dos laboratórios 204, 206 e 210 pela amizade, companheirismo, e sempre estarem disponíveis a ajudar, em especial a Tatiana, Martin, Guilherme, Edileuza, João, Karina L. e Carol.

Ao PPGBM e todos os professores das disciplinas que contribuíram na minha formação durante estes dois anos de mestrado.

Ao CNPq, CAPES e FAPERGS pelo apoio financeiro.

A toda minha família, tios, primos, sobrinhos e afilhada.

A minha mãe, meu pai, e meus irmãos Ariel e Gabriela, para quem dedico esta e todas minhas conquistas.

A todos aqueles que contribuíram de alguma forma para o sucesso deste trabalho.

Resumo

Recentemente, uma quantidade significativa de dados de sequência (tanto genômicas como transcritômicas) para *Echinococcus* spp. foi publicado, facilitando a análise de genes expressos durante um estágio específico ou envolvidos no desenvolvimento do parasito. Para realizar uma análise adequada de quantificação da expressão gênica, o uso de genes de referência validados é fortemente recomendado. Portanto, o objetivo deste trabalho foi identificar e validar genes de referência para permitir a normalização confiável da expressão gênica de determinados genes de interesse em *Echinococcus granulosus* sensu stricto (s.s.) (G1) e *Echinococcus ortleppi* durante o desenvolvimento inicial da forma pré-adulta do parasito. Protoescólices não tratados (PS) e protoescólices tratados com pepsina (PSP) do metacestódeo de *E. granulosus* s.s. (G1) e *E. ortleppi* foram usados para extração de RNA total e análise da expressão gênica. A estabilidade na expressão gênica de onze genes candidatos (*TUB*, *NDUFV2*, *RPL13*, *TBP*, *CYP*, *RNApol sub2*, *EF-1 α* , *β ACT*, *GAPDH*, *ETIF4* e *ERK*) foi avaliada usando os programas geNorm, NormFinder e RefFinder. Nossos dados de RT-qPCR mostraram uma alta correlação com dados recentemente publicados de RNA-seq. Em relação à estabilidade na expressão, *EF-1 α* e *TBP* foram os genes mais estáveis para ambas as espécies. Entretanto, *β ACT* (gene comumente utilizado como normalizador), *GAPDH* e *ETIF4* (previamente identificados como genes *housekeeping*) não tiveram um comportamento estável em nossas condições de ensaio. Desta maneira, propomos o uso de *EF-1 α* como gene de referência para estudos que envolvam análises de expressão gênica no início do desenvolvimento da forma pré-adulta de *E. granulosus* s.s. (G1) e *E. ortleppi*. Para demonstrar sua aplicabilidade, *EF-1 α* foi usado como gene normalizador na quantificação relativa de transcritos para genes que codificam as proteínas ribossômicas L10, L14, e S15, e para as subunidades do antígeno B de *E. granulosus*. O mesmo gene de referência *EF-1 α* poderia ser usado em estudos com outras espécies de *Echinococcus* sensu lato. Este é o primeiro relato que valida genes de referência adequados para a classe Cestoda, phylum Platyhelminthes, portanto fornecendo uma base para futuras validações em espécies de cestódeos epidemiologicamente importantes, como aquelas do gênero *Taenia*.

ÍNDICE

CAPÍTULO 1. INTRODUÇÃO	5
Equinococose.....	5
<i>Echinococcus</i> spp.	7
Antígeno B de <i>E. granulosus</i>	10
Expressão gênica e Genes de Referência	11
Justificativa.....	17
Objetivos	18
CAPÍTULO 2. MANUSCRITO: VALIDATION OF SUITABLE REFERENCE GENES FOR EXPRESSION NORMALIZATION IN ECHINOCOCCUS SPP. LARVAL STAGES[†]	19
Abstract	20
Introduction.....	21
Materials and Methods.....	23
<i>Sample collection, treatments and genotyping</i>	23
<i>Total RNA extraction and cDNA synthesis</i>	23
<i>qPCR analysis</i>	24
<i>Gene expression stability analysis</i>	25
<i>Normalization of selected target genes</i>	25
Results	26
<i>qPCR</i>	26
<i>Gene expression stability analysis</i>	27
<i>Relative quantification of the selected target genes</i>	28
Discussion	28
Acknowledgments	31
Supporting information	31
References	32
Figure legends	34
Tables.....	34
Figures	39
CAPÍTULO 3. DISCUSSÃO GERAL E CONCLUSÕES	43
Perspectivas.....	45
Referências bibliográficas	46
Anexos	54

Capítulo 1. Introdução

Equinococose

A equinococose é reconhecida como uma das 17 doenças zoonóticas negligenciadas pela Organização Mundial da Saúde (World Health Organization) (WHO, 2011a). Ela ocorre em humanos apresentando-se principalmente de duas formas: a equinococose cística (EC), causada pela infecção por *Echinococcus granulosus*, e a equinococose alveolar (EA), causada pela infecção por *Echinococcus multilocularis*. *E. granulosus* possui uma distribuição mundial (Figura 1), com áreas altamente endêmicas na parte oriental da região Mediterrânea, norte da África, sul e leste da Europa, no extremo sul da América do Sul, na Ásia Central, Sibéria e China ocidental. No Brasil, a doença ocorre no estado de Rio Grande do Sul, principalmente nas fronteiras com Uruguai e Argentina. Por outro lado, a EA esta confinada ao hemisfério norte, particularmente em regiões da China, Rússia, alguns países da Europa e América do Norte. Ambas as doenças são consideradas subnotificadas, no entanto, os dados indicam que a equinococose está re-emergindo como um importante problema na saúde pública. Há mais de 1 milhão de pessoas em todo o mundo afetadas com estas doenças em qualquer momento (Budke *et al.*, 2006; Torgerson *et al.*, 2010). De acordo ao impacto da doença, o custo anual da EC está estimado em US\$ 3 bilhões, o que inclui custos para o tratamento dos casos em humanos e as perdas registradas na indústria pecuária (Budke *et al.*, 2006). Em regiões onde a EC é endêmica, a incidência em humanos pode exceder a 50/100000 pessoas-ano (Craig *et al.*, 2007). Por outro lado, na pecuária, a taxa de EC encontrada em matadouros nas áreas hiperendêmicas de Latino-América variam desde 20% até 95% dos animais abatidos (Pan American Health Organization, 2001). As perdas na produção pecuária atribuíveis a EC dependem de qual espécie esta infectada, mas incluem o fígado como não apto para o consumo, redução no peso da carcaça, diminuição no valor da pele do animal, diminuição na produção do leite e redução da fecundidade (Budke *et al.*, 2006).

Varias estratégias foram e estão sendo desenvolvidas para o controle da equinococose. Por exemplo, para o ano 2015, projetos piloto para a validação de estratégias eficazes para o controle da equinococose serão desenvolvidos em três países selecionados (Quirguistão, Peru, e Tunísia). Estas estratégias e pacotes de controle integrado para as principais zoonoses relacionadas com cães (raiva e equinococose) estarão disponíveis em 2018. Este projeto será desenvolvido até o ano 2020

com um custo aproximado de US\$ 10 bilhões (WHO, 2011b). O grupo de trabalho sobre equinococose da WHO desenvolveu um consenso sobre tratamento da EC onde existem quatro opções: (i) os tratamentos percutâneos, incluindo punção, aspiração, injeção, e re-aspiração (conhecido como PAIR, siglas em inglês), (ii) a cirurgia, (iii) tratamento com medicamentos, ou (iv) observar e esperar. A escolha deve principalmente ser feita com base nos resultados das imagens de ultrassom do cisto hidático, mas também depende da infraestrutura médica e dos recursos humanos disponíveis.

A EC é evitável porque envolve espécies de animais domesticados como hospedeiros definitivos e intermediários. Com o tratamento periódico dos cães, garantindo medidas de controle no abate de animais eliminando as vísceras contaminadas, em conjunto com a educação pública, foi verificada uma taxa de transmissão da doença muito mais baixa (WHO, 2014). Um programa que combine a vacinação de ovelhas (hospedeiro intermediário do parasito), o tratamento dos cães (hospedeiro definitivo do parasito), e descarte dos ovinos mais velhos, pode levar à eliminação da doença em seres humanos em menos de 10 anos (Conchedda *et al.*, 2012).

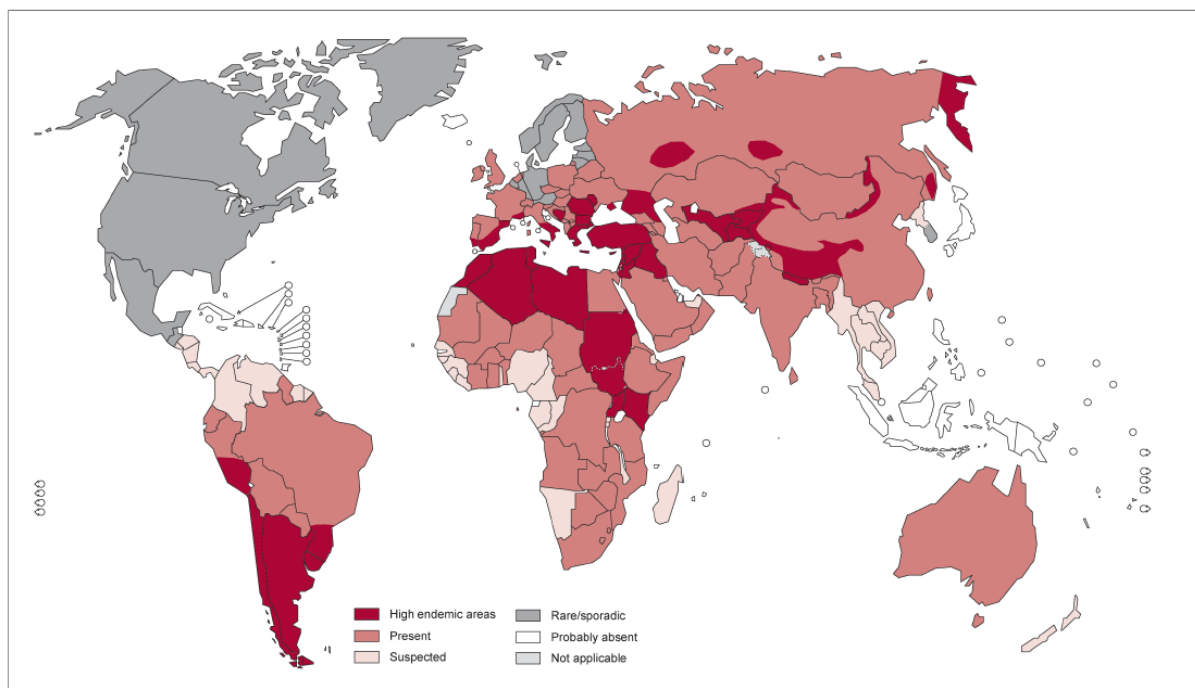


Figura 1. Distribuição mundial de *E. granulosus* e equinococose cística. Fonte: World Health Organization Map Production: Control of Neglected Tropical Diseases (NTD). World Health Organization (2011).

***Echinococcus* spp.**

Para completar e manter seu ciclo de vida, *Echinococcus* spp. precisa de dois hospedeiros, o hospedeiro intermediário (HI) (ungulados), e o hospedeiro definitivo (HD) (canídeos). O homem pode infectar-se acidentalmente ingerindo os ovos do parasito (Figura 2). Nos órgãos internos dos HI (principalmente nos pulmões e fígado) ocorre a formação do cisto hidático, dando início ao estágio larval ou metacestódeo de *Echinococcus* spp. (McManus and Smyth, 1986; Thompson and McManus, 2001).

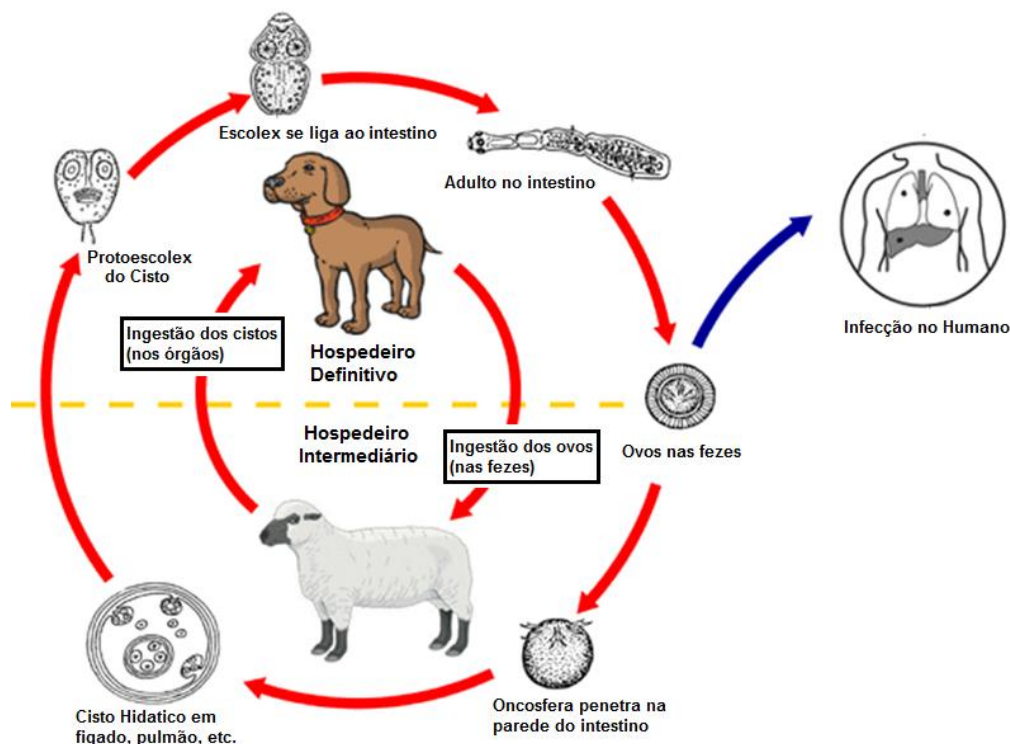


Figura 2. Ciclo de vida de *E. granulosus*. O verme adulto desenvolve-se no trato gastrointestinal do HD (canídeos). Quando os ovos do verme adulto são eliminados pelas fezes do HD e são ingeridos pelos ungulados domésticos (HI), pode ocorrer o desenvolvimento da fase larval ou metacestódeo. Nesta fase ocorre a formação do cisto hidático, onde a forma pré-adulta (protoescólices) do parasito começa a se desenvolver. Se as vísceras infectadas do HI são ingeridas pelo HD, começa o desenvolvimento da forma adulta, completando desta maneira o ciclo. O homem pode acidentalmente se infectar ao entrar em contato com os ovos do parasito eliminados pelas fezes do HD. Do mesmo modo, o cisto hidático é formado, fechando assim com o ciclo de vida. Modificado do Centro para o Controle e Prevenção de Doenças Infecciosas (www.dpd.cdc.gov).

Morfologicamente o cisto hidático é unilocular e cheio de líquido, rodeado por uma camada germinativa interna e coberta externamente por outra camada laminar não celular e rica em carboidratos. Por último e mais externamente, um tecido fibroso chamado de camada adventícia é produzido pelo hospedeiro como resposta à infecção do parasito (Thompson, 1995) (Figura 3, A). A

camada germinativa é a responsável pela formação da camada laminar e das cápsulas prolíferas, as quais se originam como pequenas massas nucleadas que proliferam em direção à cavidade do cisto. Destas cápsulas surgem por proliferação assexuada os protoescólices (PS), a forma pré-adulta do parasito. Os PS desenvolvem-se em vermes adultos quando são ingeridos pelo HD, mas também podem se re-diferenciar em cistos novos e completamente desenvolvidos (cistos secundários ou microcistos) quando são liberados na cavidade do corpo do HI, após a ruptura do cisto (Figura 4). O verme adulto (Figura 3, B) possui um corpo ou estróbilo caracteristicamente alongado e segmentado em unidades reprodutivas (proglótides), cada uma em vários estágios de desenvolvimento. Na sua porção anterior, apresenta um órgão especializado para fixação no hospedeiro, chamado de escólex.

Apesar de serem semelhantes às espécies do gênero *Taenia*, os vermes adultos do gênero *Echinococcus* possuem apenas alguns milímetros de comprimento e raramente apresentam mais de cinco proglótides (Thompson, 1995).

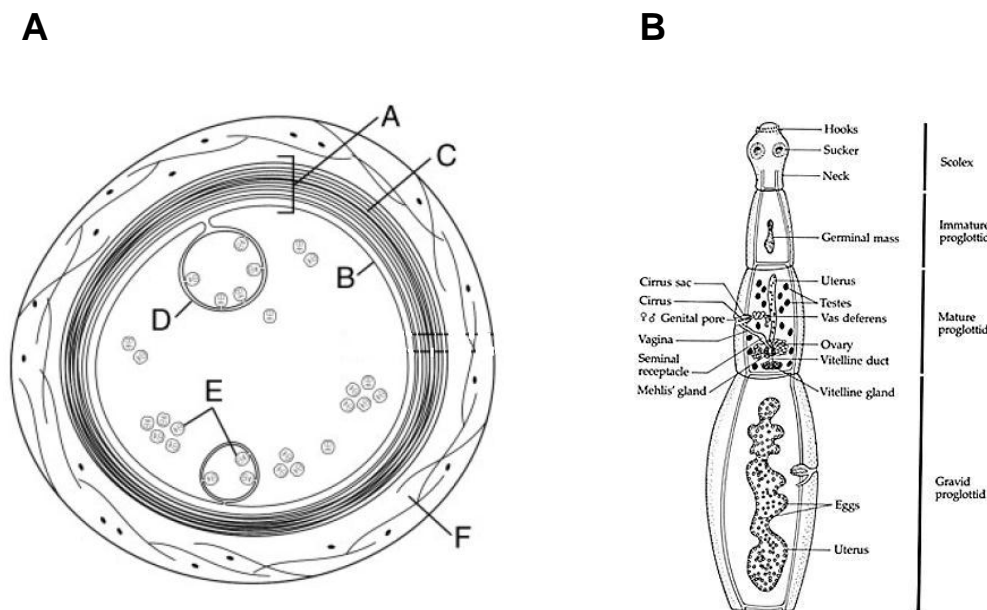


Figura 3. Representação esquemática do cisto hidático (esquerda) e do verme adulto (direita) de *E. granulosus*. A- Cisto hidático. A- parede do cisto hidático, B- camada germinativa, C- camada laminar, D- cápsula prolífera, E- protoescólices, F- camada adventícia. B. Verme adulto. As barras da direita mostram as divisões do corpo. Fontes: Monteiro *et al.* (2010); www.msudenver.edu/rowdy/.

A taxonomia de *Echinococcus* spp. tem sido constantemente revisada, já que novos dados moleculares, morfológicos, e epidemiológicos vem fornecendo evidências para a designação de

espécies que anteriormente eram reconhecidas apenas como genótipos ou linhagens. Segundo a última revisão (Nakao *et al.*, 2013) *Echinococcus granulosus* sensu lato (s.l.) inclui cinco espécies independentes: *E. granulosus* sensu stricto (s.s.) (G1–G3), responsável pela maioria dos casos humanos de EC, e *Echinococcus equinus* (G4), *Echinococcus ortleppi* (G5), *Echinococcus canadensis* (G6–G10), e *Echinococcus felidis*, responsáveis de um menor número de casos de EC em humanos ou sem casos ainda descritos (McManus, 2013).



Figura 4. Desenvolvimento bidirecional de *Echinococcus* spp. A re-diferenciação (reproduzida em condições *in vitro*, lado esquerdo) ocorre pela ruptura do cisto hidático, liberando os PS do cisto e dando origem à formação de microcistos. A diferenciação ao verme adulto (lado direito) ocorre no tracto gastrointestinal do HD. Fonte: Elissondo *et al.* (2011); www.phsource.us/index.html

Diferentes estudos já relataram a ocorrência das espécies *E. granulosus* s.s. (G1), *E. ortleppi* (G5) e *E. canadensis* (G7) em bovinos, ovinos, ou suínos no Rio Grande do Sul (Haag *et al.*, 1999; de la Rue *et al.*, 2006; Badaraco *et al.*, 2008; Monteiro *et al.*, 2014). As análises das frequências genotípicas das espécies *E. granulosus* s.s. (G1) e *E. ortleppi* (descritos como haplótipos G1 e G5 respectivamente, Balbinoti *et al.* (2012)) realizadas no sul do Brasil demonstram que 56,6% de 638 cistos férteis pertenciam à espécie *E. granulosus* s.s. (G1), enquanto 43,4% pertenciam à espécie *E. ortleppi*. Além disso, constataram a predileção de cada espécie para um determinado órgão, com *E. ortleppi* sendo comum nos pulmões e *E. granulosus* s.s. (G1) sendo comum no fígado. Por outro lado, também existem ocorrências das espécies neotropicais *E. vogeli* e *E. oligartha*, descritas principalmente na região norte do Brasil (Acre, Amapá, Amazônia, Pará e Rondônia) (de Siqueira *et al.*, 2013; Santos *et al.*, 2012; Soares *et al.*, 2004; Rodrigues-Silva *et al.*, 2002). *E. oligartha* tem sido relatada em várias espécies de felídeos neotropicais desde Argentina até o norte do México (D'Alessandro e Rausch, 2008). As infecções humanas são raras, apenas quatro casos foram relatados até hoje na América do Sul (Soares *et al.*, 2013.). A localização basal de *E. vogeli* tem sido demonstrado em várias filogenias de *Echinococcus* spp. (Bowles *et al.*, 1995; Nakao *et al.*, 2007), assim como uma relação de espécies irmãs com *E. oligartha* num estudo baseado em genes

nucleares (Knapp *et al.*, 2011). Uma análise genética populacional recente baseado em genes nucleares e mitocondriais revelou uma forte estrutura geográfica de *E. vogeli* na Amazônia brasileira, diferenciando metapopulações do oeste e leste (Santos *et al.*, 2012). Os resultados sugerem que *E. vogeli*, ao contrário de muitas outras espécies de *Echinococcus* spp., não foi afetada pela recente rápida dispersão antrópica ou natural de animais hospedeiros, e o fluxo gênico entre as regiões ocidental e oriental da Amazônia é restrito. Assim, o conhecimento da presença de determinadas espécies ou genótipos numa determinada região geográfica é muito importante em estudos epidemiológicos e no controle da hidatidose cística devido à existência de diferenças entre as espécies ou genótipos em relação ao desenvolvimento, bioquímica, e resistência a drogas anti-helmínticas do parasito (Thompson *et al.*, 1984; McManus, 2009; Stamatakos *et al.*, 2009).

Antígeno B de *E. granulosus*

O Antígeno B (AgB) é a principal proteína secretada por *E. granulosus* no estágio larval e está envolvida nas interações chaves parasito-hospedeiro durante a infecção. Desde sua primeira descrição em 1971 (Oriol *et al.*, 1971), tem sido a proteína mais estudada devido a seu papel na biologia do parasito e a seu potencial para a aplicação como ferramenta de controle. O AgB foi descrito como envolvido em vários mecanismos que promovem o estabelecimento e sobrevivência do parasito no HI, como a inibição de proteases (Shepherd *et al.*, 1991), ligação a lipídios (Chemale *et al.*, 2005) e imunomodulação (Rigano *et al.*, 2001; 2007), além do mais, o AgB é altamente imunogênico nas infecções humanas, apresentando um alto valor de diagnóstico para a hidatidose cística (Virginio *et al.*, 2003; Lorenzo *et al.*, 2005). O AgB tem uma alta homologia e aparentemente apresenta similaridade estrutural com um grupo de proteínas que ligam compostos hidrofóbicos ricos em α -hélices (HLBPs) (Shaghir *et al.*, 2001). Algumas destas proteínas parecem estar envolvidas na detoxificação, transporte, e metabolismo de lipídios com sua propriedade de ligação a ácidos graxos (Barrett *et al.*, 1997; Shaghir *et al.*, 2001).

Estruturalmente, o AgB de *E. granulosus* (EgAgB) é uma lipoproteína oligomérica de 120-160 kDa composta por subunidades de 8 kDa (EgAgB8), as quais são codificadas por uma família multigênica que inclui ao menos cinco membros (*EgAgB8/1-5*) altamente polimórficos (Kamenetzky *et al.*, 2005). Até o momento foram identificados os cinco genes que codificam as subunidades de 8

kDa do EgAgB, *EgAgB8/1* (Frosch *et al.*, 1994), *EgAgB8/2* (Fernández *et al.*, 1996), *EgAgB8/3* (Chemale *et al.*, 2001), *EgAgB8/4* (Arend *et al.*, 2002), *EgAgB8/5* (Haag *et al.*, 2004). No entanto, análises de Southern blot mostram a presença de várias bandas na hibridação com DNA genômico de *E. granulosus*, identificando ao menos sete genes que codificariam diferentes subunidades do EgAgB (Haag *et al.*, 2004). Resultados similares foram encontrados em *E. multilocularis* (Olson *et al.*, 2012) onde os sete loci de AgB parecem formar um *cluster* com alta homologia a *E. granulosus* em relação à estrutura gênica. Este *cluster* estaria composto por uma cópia de *AgB8/1*, *AgB8/2*, *AgB8/4* e *AgB8/5*, e três cópias ligeiramente diferentes de *AgB8/3* (nomeadas *AgB8/3-1*, *AgB8/3-2* e *AgB8/3-3*). Os dados genômicos recentemente publicados para *E. granulosus* (Zheng *et al.*, 2013) são consistentes com os sete genes descritos para as EgAgB1-5. Contrariamente a estes achados, na caracterização dos genes de AgB a partir de isolados obtidos em diferentes regiões geográficas, Zhang *et al.* (2010) identificaram 10 genes únicos, que apresentam uma alta homologia entre os isolados e uma baixa quantidade de polimorfismos (Haag *et al.*, 2004; Olson *et al.*, 2012).

Análises de transcrição reversa seguida de PCR quantitativa (RT-qPCR) revelaram que os níveis de transcrição de alguns genes do AgB em *E. multilocularis* (*EmAgB*) são regulados positiva ou negativamente dependendo do estágio de desenvolvimento do parasito (Mamuti *et al.*, 2006). Por outro lado, Zhang *et al.* (2010) também detectaram variações na expressão gênica das subunidades do AgB nos diferentes estágios do desenvolvimento de *E. granulosus*. Devido a estes escassos estudos de expressão gênica realizados para uma das proteínas mais importantes na interação parasito-hospedeiro, e ao uso de genes de referência não validados para quantificar estas variações, a identificação de genes normalizadores (ou genes estavelmente expressos) permitiria a quantificação relativa confiável e precisa de alvos específicos de grande relevância nas espécies do gênero *Echinococcus* spp.

Expressão gênica e Genes de Referência

Recentemente, uma enorme quantidade de dados de sequenciamento (tanto genômicos como transcritômicos) foi publicada para *E. granulosus* e espécies relacionadas (Tsai *et al.*, 2013; Zheng *et al.*, 2013). Como analisado por Zheng *et al.* (2013), existem centenas de genes que são específicos de um estágio em particular e outros genes aparentemente exclusivos e/ou envolvidos no

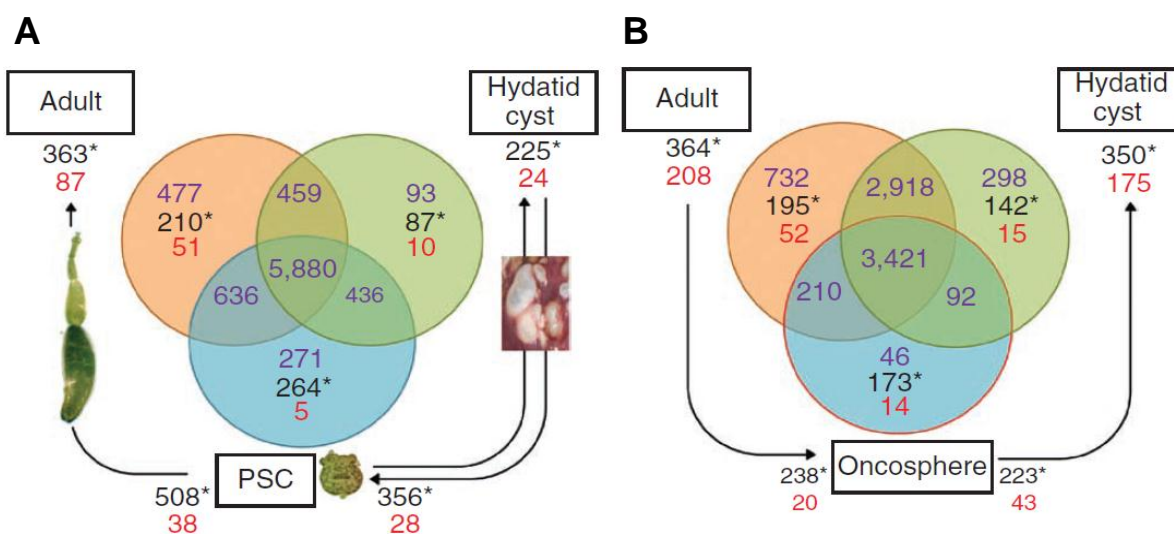


Figura 5. Regulação dos genes no adulto, oncosfera, PS, e cisto hidático. **A.** Distribuição de transcritos em PS, adulto e cisto hidático provavelmente envolvidos no desenvolvimento bidirecional dos PS em direção ao verme adulto ou a cistos secundários. **B.** Genes transcritos nos estágios de adulto, oncosfera e cisto hidático. Os números em azul indicam o número de genes transcritos em cada estágio. O número com asterisco indica o número de genes regulados positivamente de forma significativa em cada estágio comparado com cada um dos estágios imediatamente ligados no ciclo de vida ($P < 0.00001$). O número em vermelho representa os genes regulados positivamente expressos em cada estágio individual comparado com cada um dos estágios imediatamente ligados. Fonte: Zheng *et al.* (2013).

desenvolvimento bidirecional do parasito (Figura 5). No adulto, oncosfera, cisto hidático e PS foram identificados um total de 1.156 genes regulados positiva ou negativamente de forma significativa num estágio com relação aos outros três. Por exemplo, durante o desenvolvimento de PS ao verme adulto, 212 genes (2,5% do total de genes transcritos) foram descritos como regulados positivamente quando comparados estes transcritos em PS ou oncosfera. Além dos produtos gênicos envolvidos no processo metabólico normal (34 genes), os produtos gênicos regulados positivamente foram associados com a reprodução sexual (11 genes envolvidos no desenvolvimento do espermatozóide, a fertilização, a diferenciação sexual e morfogênese embrionária), sinalização (9 genes), transporte mediado por vesículas (10 genes), desenvolvimento do sistema nervoso (9 genes), proteólise (8 genes), processo de oxidação-redução (7 genes), processo metabólico da auxina (6 genes) e resposta a estímulos (21 genes). Genes transportadores e receptores de ácidos biliares provavelmente envolvidos no estímulo ao desenvolvimento dos PS ao verme adulto também foram encontrados. Moléculas quimiotáticas expressas em PS e adultos poderiam estar envolvidas na busca e aquisição dos alimentos assim como na escolha (dos PS) do local para se desenvolver ao longo do intestino do HD. Várias sequências codificadas para policistinas, envolvidas na via de

sinalização de neuropeptídios, foram também achadas em PS e adultos, sugerindo que nas fases iniciais do desenvolvimento ao adulto, *E. granulosus* produz estruturas sensoriais para detectar o movimento de fluidos no intestino cão. Estes e outros alvos específicos identificados durante o desenvolvimento do PS ao adulto (e também nos outros estágios do ciclo de vida) são obtidos de dados recentemente descritos, abrindo as portas para inúmeros estudos no campo da genômica funcional. Desta maneira, de ampla relevância são os estudos de expressão gênica para a detecção de um mRNA determinado.

PCR em tempo real ou PCR quantitativa (qPCR) fornece uma medida simultânea da expressão gênica em várias amostras para um número limitado de genes, e é especialmente adequado quando se tem disponível um pequeno número de células. Esta metodologia tem a vantagem de velocidade, rendimento e alto potencial de automatização quando comparado aos métodos de quantificação convencional, como os análises de *northern-blot*, ensaio de proteção da ribonuclease, ou RT-PCR competitiva. No entanto, esta técnica requer o mesmo tipo de normalização como os métodos tradicionais de quantificação do mRNA (Vandesompele *et al.*, 2002).

Certas variáveis precisam ser controladas na análise da expressão gênica, como a quantidade de material inicial, eficiência da transcrição reversa (RT), componentes inibitórios, diferenças entre tecidos ou células na atividade transcricional global, erros na pipetagem, entre outros (Bustin *et al.*, 2005). Varias estratégias tem sido aplicadas para normalizar essas variações (Vandesompele *et al.*, 2002). Sob condições controladas de extração de RNA total de boa qualidade e reproduzível, o número de transcritos pode ser idealmente padronizado ao número de células, porém, a contagem de células é muitas vezes impedido quando, por exemplo, o material de partida é um tecido sólido. Outra escala utilizada para normalização é a quantidade de massa, especialmente em análises do tipo *northern blot*; porém, existem vários argumentos em contra do uso da quantidade de massa. A qualidade do RNA e a eficiência das reações enzimáticas não são levadas em consideração nesta abordagem. Além disso, o RNA total consiste predominantemente de moléculas de rRNA, e nem sempre é representativo da fração de mRNA. Outras desvantagens do uso das moléculas de 18S ou 28S como *standards* se devem à ausência em amostras de mRNA purificado, e à abundância quando comparado a transcritos de mRNA alvos. Até hoje, os genes de controle endógenos são os mais frequentemente usados para normalizar os parâmetros acima descritos. A

expressão destes controles internos (frequentemente chamados de genes *housekeeping*) não deveria variar nas células ou tecidos em estudo, ou em resposta aos tratamentos experimentais. No entanto, vários estudos utilizam estes genes expressos de forma constitutiva sem uma adequada validação da sua presumida estabilidade na expressão. Além disso, a literatura mostra que a expressão dos genes *housekeeping* (embora ocasionalmente constante num determinado tipo celular ou tecido ou condição experimental) pode variar consideravelmente (Vandesompele *et al.*, 2002). Com o incremento da sensibilidade, reprodutibilidade e amplo alcance dinâmico dos métodos de RT-PCR em tempo real, os requerimentos de um gene de controle endógeno tornaram-se rigorosamente estritos. Além disso, sem uma previa validação, a acurácia dos dados de expressão gênica obtidos por metodologias de alta sensibilidade, como a PCR em tempo real, estaria diretamente afetada. Frente a esta questão, diferentes algoritmos estatísticos foram desenvolvidos para escolher os genes de referência mais apropriados para a normalização de dados de PCR em tempo real a partir de um determinado painel de genes candidatos. Através dos programas geNorm (Vandesompele *et al.*, 2002), BestKeeper (Pfaffl *et al.*, 2004), NormFinder (Andersen *et al.*, 2004), é possível classificar os genes escolhidos como candidatos em mais ou menos estáveis.

O geNorm calcula a medida M de estabilidade na expressão de um gene de referência como a variação média par a par V desse gene contra todos os outros genes candidatos testados. A exclusão passo a passo do gene com o valor M mais alto (menos estável) permite o *ranking* dos genes testados de acordo a sua estabilidade na expressão. geNorm também determina o número ideal de genes de referência necessários para normalizar a expressão de certos alvos de acordo com variação par a par entre dois fatores de normalização subsequentes $V_{n/n+1}$, onde “n” indica o número de genes envolvidos na comparação par a par. Este fator de normalização da expressão gênica é calculado para cada amostra do tecido com base na média geométrica de um número definido de genes de referência candidatos. Uma ampla variação implica que a adição do gene tem um efeito significativo e deve preferivelmente ser incluído para o cálculo de um fator de normalização confiável. Baseado no trabalho que descreve o algoritmo utilizado pelo geNorm, os autores propõem o valor de 0,15 como *cut-off*, abaixo do qual não é necessária a inclusão de um gene de referência adicional. Por exemplo, se o valor da $V_{3/4}$ é de 0,22, o fator de normalização deveria conter preferivelmente no mínimo os 4 melhores genes de referência. Subsequentemente, se o valor da $V_{4/5}$ for de 0,14, não haveria necessidade real de incluir um quinto gene no fator de normalização. O número ideal de

genes para a normalização neste exemplo seria, portanto, quatro. Ao contrário de geNorm que se baseia em comparações par a par, NormFinder baseia-se no modelo matemático da análise da variância (ANOVA) para estimar a variações intra- e inter-grupo, combinando assim estas estimativas para fornecer uma medida direta da variação na expressão de cada gene. Tanto no geNorm como NormFinder, os dados de entrada (*input*) devem ser quantidades relativas, diferente do Bestkeeper, onde os dados brutos (valores de ΔCt) devem ser inseridos. Este último programa gera um índice usando a média geométrica dos valores de Cq dos genes de referência candidatos. Ele também calcula o desvio padrão (DP) e o coeficiente de variação (CV) baseado nos valores de Cq de todos os genes candidatos. Os genes de referência identificados como os mais estáveis serão aqueles que apresentarem o menor coeficiente de variação e desvio padrão ($CV \pm DP$). Adicionalmente, existe uma ferramenta chamada de RefFinder (<http://www.leonxie.com/referencegene.php>) que integra cada um dos algoritmos explicados anteriormente (geNorm, NormFinder e Bestkeeper), além do método comparativo ΔCt . Esta ferramenta compara e classifica os genes de referência candidatos baseado nos rankings de genes mais ou menos estáveis obtidos por cada um dos programas, ele atribui um peso adequado a um gene individual e calcula a media geométrica dos seus pesos para a classificação geral final.

Atualmente não existe para o gênero *Echinococcus*, tampouco para outros organismos pertencentes à classe Cestoda uma validação adequada de genes normalizadores para ser utilizada em diferentes estudos de expressão gênica. Na literatura observa-se que diversos autores utilizam ou propõem diferentes genes como normalizadores, como a actina (Siles-Lucas *et al.*, 2001; Muzulin *et al.*, 2007; Cabrera *et al.*, 2008; Pan *et al.*, 2009; Graichen *et al.*, 2007; Matsumoto *et al.*, 2006; Zhang *et al.*, 2010), ciclofilina (Colebrook *et al.*, 2002), fator da iniciação da tradução eucariótica (Zhang *et al.*, 2010), e kinase reguladora da sinal extracelular (Li *et al.*, 2011). Embora o gene da actina seja um dos mais utilizados, Zhang *et al.* (2010) verificaram que o mesmo é significativamente regulado de forma positiva em adultos maduros de *E. granulosus*, variando em outros estágios. No presente estudo, serão selecionados, a partir de um painel de potenciais genes de referência para *E. granulosus* s.s. (G1) e *E. ortleppi*, aqueles que se encontram mais estavelmente expressos durante o estágio larval e em diferentes condições de indução do desenvolvimento do parasito. Após a obtenção do/s gene/s normalizador/es, se feita a avaliação da expressão gênica das subunidades do

antígeno B de *E. granulosus* s.s. e *E. ortleppi*, tendo como perspectiva a avaliação de outros genes alvo relevantes na interação parasito-hospedeiro.

Justificativa

Existem vários estudos em *Echinococcus* spp. que utilizam determinados genes como “controles” endógenos sem uma adequada validação de sua presumida estabilidade na expressão. A avaliação do comportamento de genes candidatos a normalizar a expressão de certos genes alvo é necessária e adequada para ter resultados precisos e confiáveis. Com a obtenção de genes estavelmente expressos, é possível analisar produtos gênicos de grande relevância na interação parasito-hospedeiro como as subunidades do antígeno B.

Objetivos

Geral

Validar genes normalizadores em *E. granulosus* sensu stricto (G1) e *E. ortleppi* em diferentes estágios do desenvolvimento do parasito empregando a técnica de PCR em tempo real.

Específicos

Avaliar a expressão de diferentes genes candidatos a normalizadores para análises de expressão gênica por RT-qPCR em diferentes estágios de desenvolvimento dos parasitos *E. granulosus* s.s. (G1) e *E. ortleppi*.

Determinar os genes cuja expressão seja mais estável nas diferentes condições experimentais a serem avaliadas e sugerir a utilização destes genes em posteriores estudos de expressão gênica.

Utilizando o/s gene/s normalizador/es validado/s neste estudo, analisar a expressão dos genes que codificam as subunidades do AgB em *E. granulosus* sensu stricto (G1) e *E. ortleppi*.

1 **Capítulo 2. Manuscrito: Validation of suitable reference genes for expression**
2 **normalization in Echinococcus spp. larval stages †**

3

4 **Sergio Martin Espínola¹, Henrique Bunselmeyer Ferreira² and Arnaldo Zaha^{1,2*}**

5

6 ¹ Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio
7 Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

8 ² Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências; and Centro de
9 Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

10

11 * Corresponding author

12 Corresponding author's email address: zaha@cbiot.ufrgs.br

13 † *Artigo submetido (24 de fevereiro de 2014), revisado (30 de maio de 2014) e aceito para publicação*
14 *(16 de junho de 2014) na revista PLoS ONE.*

15

16

17 **Abstract**

18 In recent years, a significant amount of sequence data (both genomic and transcriptomic) for
19 *Echinococcus* spp. has been published, thereby facilitating the analysis of genes expressed during a
20 specific stage or involved in parasite development. To perform a suitable gene expression
21 quantification analysis, the use of validated reference genes is strongly recommended. Thus, the aim
22 of this work was to identify suitable reference genes to allow reliable expression normalization for
23 genes of interest in *Echinococcus granulosus* sensu stricto (s.s.) (G1) and *Echinococcus ortleppi*
24 upon induction of the early pre-adult development. Untreated protoscoleces (PS) and pepsin-treated
25 protoscoleces (PSP) from *E. granulosus* s.s. (G1) and *E. ortleppi* metacestode were used. The gene
26 expression stability of eleven candidate reference genes (β TUB, NDUFV2, RPL13, TBP, CYP-1, RPII,
27 EF-1 α , β ACT-1, GAPDH, ETIF4A-III and MAPK3) was assessed using geNorm, Normfinder, and
28 RefFinder. Our qPCR data showed a good correlation with the recently published RNA-seq data.
29 Regarding expression stability, EF-1 α and TBP were the most stable genes for both species.
30 Interestingly, β ACT-1 (the most commonly used reference gene), and GAPDH and ETIF4A-III
31 (previously identified as housekeeping genes) did not behave stably in our assay conditions. We
32 propose the use of EF-1 α as a reference gene for studies involving gene expression analysis in both
33 PS and PSP experimental conditions for *E. granulosus* s.s. and *E. ortleppi*. To demonstrate its
34 applicability, EF-1 α was used as a normalizer gene in the relative quantification of transcripts from
35 genes coding for antigen B subunits. The same EF-1 α reference gene may be used in studies with
36 other *Echinococcus* sensu lato species. This report validates suitable reference genes for species of
37 class Cestoda, phylum Platyhelminthes, thus providing a foundation for further validation in other
38 epidemiologically important cestode species, such as those from the *Taenia* genus.

39

40 Introduction

41 Echinococcosis, one of the 17 neglected tropical diseases prioritized by WHO [1], is caused
42 by species from the genus *Echinococcus*. According to the latest revision [2], this genus includes nine
43 species, two of which have medical and public health relevance in humans: *Echinococcus granulosus*
44 sensu lato (s.l.), which is the causative agent of cystic echinococcosis (CE), and *Echinococcus*
45 *multilocularis*, which is responsible for alveolar echinococcosis (AE). Previously, the complex *E.*
46 *granulosus* was represented by several genotypes or strains (G1-G10) of the same species [3,4].
47 Currently, *E. granulosus* s.l. includes five independent species as follows: *E. granulosus* sensu stricto
48 (s.s.) (G1–G3), which is responsible for the majority of human CE, and *E. equinus* (G4), *E. ortleppi*
49 (G5), *E. canadensis* (G6–G10), and *E. felidis*, which are responsible for fewer human CE cases [2,5].
50 CE results in a loss of 1–3 million disability-adjusted life years (DALYs) per annum, and it shows a
51 cosmopolite distribution, having a hyperendemic pattern in several regions of East Africa, Central
52 Asia, China, and South America (including Argentina, Uruguay, Chile, Peru, and Southern Brazil). In
53 these areas, a CE prevalence of 20-90% is observed in slaughtered animals, leading to annual losses
54 of approximately US\$ 3 billion [1].

55 The life cycle of *Echinococcus* species encompasses two different hosts as follows: the
56 intermediate host (generally domestic ungulates) where the infective larval stage (metacestode)
57 occurs, resulting in protoscoleces, which is the pre-adult parasite form, and the definitive host (usually
58 canids) where the differentiation of protoscoleces to the adult form occurs. Through sexual
59 reproduction, the adult form generates eggs that, once eliminated by the definitive host feces, are
60 ingested by the intermediate host, thereby completing the cycle. In *E. granulosus* s.l. species, the
61 metacestode stage is represented by the formation of a unilocular hydatid cyst, filled with hydatid
62 fluid, and more internally by a germinal layer, which gives rise to protoscoleces [6,7]. A particular
63 developmental characteristic of protoscoleces is the possibility to re-differentiate into secondary cysts
64 or microcysts [6], which naturally occurs by hydatid cyst rupture and release of its content in the
65 intermediate host, with each protoscolex being able to generate an individual secondary cyst. The
66 ability to infect and survive in a wide range of intermediate hosts and the peculiar bidirectional
67 development features make the *Echinococcus* species an interesting model to study host-parasite
68 relationships and parasite development. The high impact of this disease on livestock and public health

69 also highlights the necessity of basic studies on *Echinococcus* spp. to identify molecular targets and
70 develop new strategies for CE control and eradication.

71 In recent years, a significant amount of *Echinococcus* spp. sequence data (both genomic and
72 transcriptomic) have been published [8-10], facilitating the analysis of genes differentially expressed
73 in specific life cycle stages or potentially involved developmental events of the parasite. Several
74 upregulated or downregulated genes have been recently described in a transcriptomic survey of *E.*
75 *granulosus* [8].

76 The microarrays and the RNA-seq are some of the several methodologies to detect different
77 gene expression levels, and commonly are used to have a global vision of the gene expression.
78 Moreover, when the focus is a limited number of target genes, the quantitative PCR (qPCR) is the
79 gold standard method to detect the gene expression variations of a specific mRNA [11]. Comparing to
80 conventional methods of quantification, such as the RNase protection assay or northern blot, the
81 qPCR have the advantage of high sensitivity, specificity, reproducibility, and broad dynamic range,
82 making it one of the most widespread techniques in many areas of research [12,13].

83 To perform a suitable and reliable gene expression quantification analysis, reference genes
84 need to be validated. Reference (or normalizer) genes are defined as those with a stable expression
85 under previously defined conditions, thus appropriate to quantify gene expression levels of specific
86 targets. To date, there are no reports on the validation of reference genes for members of the class
87 Cestoda, phylum Platyhelminthes. There have been several reports on the identification of
88 housekeeping genes for the larval infective stage of *E. granulosus* [14-18]. Furthermore, because the
89 widely used β ACT-1 gene is upregulated in immature adults of *E. granulosus* and is variably
90 expressed in the other developmental stages [19], the eukaryotic translation initiation factor (*Eg-eif*)
91 has been proposed as an alternative housekeeping gene. Without previous validation of the reference
92 gene, the accuracy of gene expression data obtained by high sensitivity methodologies, such as
93 qPCR, will be directly affected.

94 In this study, we describe the identification of a suitable and reliable reference gene for
95 normalizing the expression of specific target genes in *E. granulosus* s.s. and *E. ortleppi* protoscoleces
96 upon induction of early pre-adult development by pepsin treatment. Using the validated reference
97 gene, we quantified the relative mRNA expression of differential and constitutive expressed genes

98 between both PS and PSP conditions, as previously described [10,20]. Finally, we analyze the
99 expression of the *E. granulosus* genes coding for antigen B (EgAgB) subunits and compare with
100 previous results of RNA-seq and qPCR data [8,19,21].

101

102 **Materials and Methods**

103 *Sample collection, treatments and genotyping*

104 Bovine hydatid cysts were obtained from the Cooperleo Abattoir (São Leopoldo, Rio Grande
105 do Sul, Brazil). The slaughtered animals came from different regions of the Rio Grande do Sul, mostly
106 from farms located in the south and southwest of the state. The protoscoleces were collected by
107 hydatid cyst fluid aspiration and washed at least five times with 1X phosphate buffered saline (PBS).
108 Viability was assessed through optical microscope observation and trypan blue staining. Only
109 protoscoleces with viability greater than 90% were used for further analysis. For species
110 determination, a high-resolution melting (HRM) genotyping method was performed using part of the
111 cytochrome c oxidase subunit I (*cox1*) gene, which has distinct melting curves that allow
112 discrimination between *E. granulosus* s.s. (G1) and *E. ortleppi* [22].

113 Focusing on the early development of the pre-adult form, we used two different conditions: 1)
114 protoscoleces were directly extracted from hydatid cyst and washed with 1X PBS (PS) and 2)
115 protoscoleces were treated with pepsin (PSP). For the PSP group, after washing with 1X PBS, the
116 samples were treated with pepsin (2 mg/mL) for 15-20 min at pH 2 to mimic the contact with the
117 digestive enzymes of the definitive host, thus achieving an “activated” or development-induced state
118 [6]. After pepsin treatment, the evagination of protoscoleces and their flame cell movements were
119 clearly evident by microscope observation.

120

121 *Total RNA extraction and cDNA synthesis*

122 PS or PSP samples (50-100 μ L, containing approximately 5000-10000 individuals) were
123 mixed with 1 mL of TRIzol reagent and immediately frozen in liquid nitrogen until the total RNA
124 extraction. Total RNA was isolated using TRIzol reagent in conjunction with the PureLink RNA Mini Kit
125 according to the manufacturer’s protocol. After treatment with RNase-free DNase I (Thermo

126 SCIENTIFIC) for 30 min at 37°C to remove all genomic DNA, total RNA concentration was
127 determined using a Nanodrop ND2000 spectrophotometer (Thermo SCIENTIFIC). Quality and
128 integrity were assessed by 1.5% agarose gel electrophoresis and by an Agilent 2100 Bioanalyzer
129 using an RNA 6000 Pico Chip Kit. The first strand of cDNA was synthesized from 100 ng of total RNA
130 using RevertAid reverse transcriptase (Thermo SCIENTIFIC) and Oligo (dT)₁₈ (0.5 µg/µL) as the
131 anchor primer. The reaction mixture was incubated at 42°C for 1 h followed by 70°C for 10 min to
132 terminate the reaction and brought to a final volume of 20 µL. The final cDNA product was diluted 50-
133 fold with nuclease-free water prior to use in qPCR analysis.

134

135 *qPCR analysis*

136 Eleven genes were selected for expression studies, namely beta tubulin (*βTUB*), NADH
137 dehydrogenase ubiquinone flavoprotein 2 (*NDUFV2*), L13 ribosomal protein (*RPL13*), TATA-box
138 binding protein (*TBP*), cyclophilin 1 (*CYP-1*), RNA polymerase II subunit RPB2 (*RPII*), elongation
139 factor 1 alpha (*EF-1α*), beta actin 1 (*βACT-1*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*),
140 eukaryotic translation initiation factor 4A-III (*ETIF4A-III*) and mitogen activated protein kinase 3
141 (*MAPK3*) genes. Gene sequences obtained from different databases (LophDB, GeneDB, and
142 GenBank) were used to design specific primers using Vector NTI software, except for the *βACT-1*
143 gene, for which the primer sequences were obtained from a previous study in *E. multilocularis* [23]
144 and already used in *E. granulosus* as reference gene [19]. For primer design, the following
145 characteristics were considered: an amplification product between 100–200 bp, annealing
146 temperature of 60±1°C, and location of the amplified sequence close to the 3' end. The details of
147 each selected gene and the characteristics of each primer are shown in Table 1 and Table 2,
148 respectively. The qPCR reactions were performed using an ABI Real-Time 7500 PCR system
149 (Applied Biosystems) with the following reaction mixture: 10 µL of diluted cDNA as template, 0.1X
150 SYBR Green I (Invitrogen), 0.1 µM of each primer, 1X PCR Buffer (20 mM Tris-HCl pH 8.4, 50 mM
151 KCl), 3 mM MgCl₂, 25 µM dNTPs, 0.25 U Platinum Taq DNA polymerase, and MilliQ water in a final
152 reaction volume of 20 µL (for each qPCR reagent, the final concentration is showed). A reverse
153 transcription negative control (without reverse transcriptase) for each synthesized cDNA and a non-
154 template negative control for each gene run were included to confirm the absence of genomic DNA

155 and contamination of PCR reactions, respectively. The qPCR conditions were as follows: initial
156 activation at 94°C for 5 min, followed by 40 cycles with denaturation at 94°C for 15 s, annealing at
157 60°C for 10 s and extension at 72°C for 35 s. A dissociation step from 94°C to 50°C with ramping
158 increments of 0.1°C/s was added to assess the amplification specificity for each gene through melting
159 curve analyses in SDS software (provided in ABI Real-Time 7500 system). The specificity of the
160 amplified products was also analyzed by electrophoresis in 2% agarose gels. To determine the PCR
161 amplification efficiency for each candidate to reference gene, standard cDNA dilutions were prepared
162 using seven 2-fold serial dilutions. To calculate the PCR efficiency, the LinRegPCR software was
163 used [24]. All qPCR reactions for each sample and each gene were performed in triplicate.

164

165 *Gene expression stability analysis*

166 geNorm [25] and NormFinder [26] are the two gene normalization algorithms generally used
167 to analyze the expression stability of candidate reference genes. geNorm calculates the gene
168 expression stability measure (M value) for a reference gene as the average pairwise variation (V) for
169 that gene with all other tested reference genes. Moreover, geNorm determines an optimal number of
170 reference genes for reliable normalization. NormFinder algorithm is based on the analysis of variance
171 (ANOVA) mathematical model and allows the estimation of intra- and intergroup variation as well as
172 the calculation of reference gene stability values. In addition, we used the RefFinder tool
173 (<http://www.leonxie.com/referencegene.php>) for the assessment and screening of reference genes.
174 The RefFinder tool integrates the currently available major computational programs (geNorm,
175 Normfinder, BestKeeper, and the comparative $\Delta\Delta\text{Ct}$ method) to compare and rank the tested
176 candidate reference genes.

177

178 *Normalization of selected target genes*

179 Once the most stably expressed genes were detected, we used the $\Delta\Delta\text{Ct}$ method to quantify
180 the expression of three selected target genes: two ribosomal proteins, L14 and s15, previously
181 described as differentially expressed between *E. granulosus* PS and PSP conditions [10]; and the
182 ezrin-radixin-moesin (ERM)-like protein (*ELP*) gene, described as constitutively expressed in PS and
183 PSP conditions for *E. multilocularis* [20]. Also, we quantified the relative expression of the five genes

184 that encode the widely studied *EgAgB1-5* genes. We used the available RNA-seq data [8] and
185 previous qPCR analysis [19,21] to discuss the gene expression of the different EgAgB subunits. The
186 characteristics of the specific primers for these eight target genes (designed with Vector NTI software)
187 are summarized in Table 2. Using the $2^{-\Delta\Delta CT}$ values, we compared the PS and PSP experimental
188 groups through the paired samples *t* test. Furthermore, the differences between the relative quantities
189 of each *EgAgB1-5* gene were assessed by ANOVA. Statistical analyses were performed using SPSS
190 software. In this assay, three biological replicates and two technical replicates were used.

191

192 **Results**

193 A total of 10 cysts were collected, with 3 belonging to *E. granulosus* s.s. (G1) and 7 belonging
194 to *E. ortleppi*. We used paired PS and PSP samples for all *E. granulosus* s.s. (G1) and for five *E.*
195 *ortleppi* cysts. For the other two *E. ortleppi* cysts, only a PS or PSP sample (unpaired samples) was
196 analyzed due to the low amount of parasite material. Neither genomic DNA nor RNA degradation was
197 observed for any of the total RNA samples analyzed (Figure S1). As expected for several
198 Platyhelminthes species [27], a single band of total RNA was observed on the agarose gel and in the
199 Bioanalyzer analysis in all samples (Figure S1).

200

201 *qPCR*

202 For all PCR products, we did not detect amplification of nonspecific products, formation of
203 primer dimers or any PCR contaminants (Figure S2). The amplification efficiency for each candidate
204 to reference gene and for each target gene is shown in Table 2. For the *EgAgB5* gene the
205 amplification efficiency value were not calculated due to its very low expression level, avoiding the
206 formation of the *plateau*, and thus, the amplification efficiency calculation by the LinRegPCR software.
207 The amplification curves used to calculate the amplification efficiency of the candidate reference
208 genes are shown in Figure S3. Once all the cycle quantification (Cq) values of the qPCR were
209 obtained, we assessed the transcript abundance of each gene in *E. granulosus* s.s. (G1) and *E.*
210 *ortleppi* (Figure 1). In both species, the transcript distribution was almost the same, which was
211 expected for closely related species and apparent housekeeping genes. Moreover, we found that
212 *βTUB*, *RPL13*, *RPII* and *βACT-1* contain the most dispersal Cq values in *E. ortleppi*, while in *E.*

213 *granulosus* s.s. (G1) it is observed for βTUB and *RPII*. Furthermore, because of the greater number of
214 biological replicates for *E. ortleppi*, we based on the mean of both PS and PSP values of each gene
215 to compare our qPCR data versus the RNA-seq data for *E. granulosus* spp. protoscolex (expressed
216 as Reads Per Kilobase per Million (RPKM) [8] and Fragments Per Kilobase per Million (FPKM) [9]). A
217 linear correlation (R^2) of 0.67 and 0.56 in *E. granulosus* s.s. (G1) and 0.82 and 0.73 in *E. ortleppi* was
218 observed for RPKM and FPKM, respectively (Figure 1, inset).

219

220 *Gene expression stability analysis*

221 The most common algorithms available to assess gene expression stability were employed,
222 and the results were compared [25,26]. These algorithms required the transformation of Cq values to
223 relative quantities. Thus, the mean Cq values from triplicate runs were converted into relative
224 quantities by the $\Delta\Delta Ct$ method and used as input data for both the geNorm and NormFinder
225 algorithms. Figure 2 shows the two output charts from the geNorm program for each species.
226 Regarding the average expression stability values, *EF-1 α* and *TBP* were the most stable genes in *E.*
227 *ortleppi*, and *EF-1 α* and *RPL13* were the most stable genes in *E. granulosus* s.s. (G1). In contrast,
228 βTUB , *ETIF4A-III*, $\beta ACT-1$ and *NDUFV2* were the least stable genes for both species. The other
229 geNorm chart showed the optimal number of reference genes that would be necessary for suitable
230 gene expression normalization. In both species, the V2/3 values were less than 0.15 (the cut-off value
231 recommended by geNorm authors), thereby indicating that the optimal number of reference genes for
232 normalization is 2. The other algorithm that allows the identification of the most reliable reference
233 genes from a set of candidate genes is Normfinder. Here, the ranking and stability values obtained for
234 each gene (and also in geNorm) were exactly the same as those obtained by RefFinder. For this
235 reason, we plotted only the RefFinder output data in Table 3. The comprehensive ranking for *E.*
236 *ortleppi* was similar to that of *E. granulosus* s.s. (G1). However, the ordinal order given by each gene
237 expression stability algorithm was more variable in *E. granulosus* s.s. (G1) than in *E. ortleppi*.
238 Performing random samplings of the Cq values for 3 paired samples of *E. ortleppi* and by placing
239 these data as input in RefFinder, we corroborated that the differences in the ordinal order obtained
240 was clearly a consequence of the number of biological replicates used (Table S1A, S1B and S1C).
241 For *E. ortleppi*, *TBP*, *EF-1 α* and *CYP-1* were the most stable genes, and *ETIF4A-III* and βTUB were
242 the least stable genes (Table 3). For *E. granulosus* s.s. (G1), *TBP*, *EF-1 α* and *GAPDH* were the most

243 stables genes, and *RPII* and *βTUB* were the least stable genes (Table 3). Interestingly, for both
244 species, none of the genes described as housekeeping genes in the *Echinococcus* literature (*βACT-1*,
245 *ETIF4A-III*, *CYP-1* and *MAPK3*) demonstrated a good expression stability value (Figure 2).

246

247 *Relative quantification of the selected target genes*

248 Firstly, with the most stable gene identified, we performed a relative quantification analysis of
249 three selected target genes, the *RPL14*, the *RPs15* and the *ELP* genes. Significant statistical
250 differences was solely observed for the *RPL14* gene ($p < 0.05$) (Figure 3A). The *RPs15* gene showed a
251 clear tendency to increase in the PSP experimental condition, with $p = 0.06$, in contrast to the *ELP*
252 gene, which exhibit no significant difference for both *E. granulosus* s.s. (G1) and *E. ortleppi* (p values
253 of 0.14 and 0.17, respectively) (Figure 3A). Based on these consistent results with the previously
254 described [10,20], we normalized the *EgAgB1-5* gene expression using *EF-1α* as the reference gene.
255 The significant differences between each *EgAgB1-5* gene are shown in Figure 3B (inset). As
256 described previously [19,21], *EgAgB3* was the most abundant, and *EgAgB5* was expressed at a low
257 level. *EgAgB1* showed a moderate expression level, followed by *EgAgB2*, *EgAgB4* and *EgAgB5*.
258 Among the PS and PSP experimental groups, *EgAgB1*, *EgAgB2*, and *EgAgB4* showed considerable
259 variation as indicated by the bar graph (Figure 3) and qPCR curves (ΔCq variation of 0.6-3.35 for
260 *EgAgB1*, 1.66-3.57 for *EgAgB2* and 0.86-4.25 for *EgAgB4*), but no significant differences were found
261 for all *EgAgB1-5* genes.

262

263 **Discussion**

264 Faced with a vast amount of gene expression data, it is important to recognize and
265 understand which genes are upregulated or downregulated as well as which genes are specific to a
266 particular life cycle stage. The elucidation of the dynamic behavior of gene expression is also
267 expected to provide new insights into parasite development and host-parasite relationships. Recently,
268 a large amount of sequencing (both genomic and transcriptomic) data for *Echinococcus* genus and
269 other related parasites was published and made available in public databases [8,9]. However, there
270 are very few publications that validate the transcriptomic data. A more comprehensive overview
271 considering the available protein expression data for *E. granulosus* [28,29] could help to obtain and

272 understand the correlations between transcription and translation pathways. In our analysis, we
273 focused on the identification of genes that are stably expressed (normalizer or reference genes) in PS
274 and PSP experimental conditions of *E. granulosus* s.s. (G1) and *E. ortleppi*. In addition to confirming
275 the transcriptomic data of several selected genes, we showed a suitable and reliable validation of
276 gene expression stability for *Echinococcus* spp. in the initial stage of the pre-adult development.

277 We selected eleven candidates for reference genes, seven of which were usually employed in
278 previous studies (*βTUB*, *NDUFV2*, *RPL13*, *TBP*, *CYP-1*, *RPII*, and *EF-1α*) and four of which were
279 previously identified as housekeeping genes in *Echinococcus* spp. (*βACT-1*, *GAPDH*, *ETIF4A-III*, and
280 *MAPK3*). Another criterion for selecting candidate reference genes was that the genes were involved
281 in different functions or pathways (metabolism, structural, translation, and signal transduction) to
282 avoid a possible co-regulation between selected genes. Despite the lack of validated reference
283 genes, several publications involving gene expression analysis on *Echinococcus* spp. have been
284 published [16-19,30-33]. The *βACT-1* gene is widely used as a housekeeping gene, but this gene has
285 been shown to be significantly upregulated in immature adult worms and to be variable in the other
286 stages. Thus, the translation initiation factor of *E. granulosus* (*Eg-eif*) was proposed as an alternative
287 housekeeping gene [19]. Importantly, this previous study showed that one of the most commonly
288 used genes is not a good normalizer for gene expression analysis. However, the gene proposed by
289 the authors was not validated, thereby creating a new uncertainty regarding reliable gene expression
290 quantification. Other genes identified with apparently constitutive expression in *E. granulosus* were
291 cyclophilin in protoscoleces [14] and the extracellular signal-regulated kinase in the cyst wall and
292 protoscoleces [15]. Although the selection of the candidate reference genes was not based on
293 recently published RNA-seq quantification data for *Echinococcus* spp., we identified stably expressed
294 genes to be used as reference genes. It is worth noting that our qPCR data for several selected
295 genes confirmed and validated those obtained from transcriptome analysis of *E. granulosus* spp. [8,9]
296 (Figure 1) and from other studies where *CYP-1*, *GAPDH* and *βACT-1* generally showed high
297 transcript abundance [34-36].

298 Several different algorithms are available to identify the relative stability of genes from a
299 given set of candidate reference genes. Generally, these algorithms show a stability value and an
300 ordinal ranking that allow selection of the best reference gene for further analysis of gene expression
301 quantification. geNorm was the first program to be published [25], and it is currently the most used to

302 identify normalizer genes. In addition to the stability value, geNorm gives the number of reference
303 genes that would be necessary for suitable gene expression normalization. For *E. granulosus* s.s.
304 (G1), *EF-1 α* and *RPL13* were the recommended reference genes. However, when we compared the
305 geNorm ranking with those generated by other programs (Table 3), we found that *TBP* and *EF-1 α*
306 were the most stable genes and that *RPL13* was not included within the most stable genes. Similar
307 results were obtained for *E. ortleppi*, where the distribution of the most and least stable genes for
308 each method was correlated (Table 3), which may have been due to the number of samples used in
309 this species. The difference in the number of biological samples was due to differences in the species
310 frequency in the region sampled, as discussed previously [37]. For *E. ortleppi*, *TBP* and *EF-1 α* were
311 the most stable genes, which was the same as the RefFinder ranking observed for *E. granulosus* s.s.
312 (G1). An interesting observation in the *E. granulosus* s.s. (G1) RefFinder ranking was that *TBP* was
313 the most stable gene regarding the comparative $\Delta\Delta\text{Ct}$ and NormFinder methods but not for the
314 geNorm program, suggesting that the comparative analysis performed by RefFinder is an efficient
315 strategy to determine a suitable ranking from a set of candidate reference genes, mainly when the
316 number of replicates is small. Based on this comparative analysis between the two related species,
317 we propose the use of *EF-1 α* and *TBP* genes as reference genes for studies that involve gene
318 expression analysis of *E. granulosus* s.s. (G1) and *E. ortleppi* PS and PSP experimental conditions. In
319 accordance with the similar results obtained for each species, we suggest that *EF-1 α* and *TBP* could
320 be used to normalize the gene expression in studies involving other *E. granulosus* s.l. species. It is
321 important to note that we focused on the early pre-adult development of *Echinococcus* spp., excluding
322 the germinal layer tissue and other parasite life cycle stages, such as the onchosphere or the adult
323 form. Thus, for gene expression quantification studies including these stages, new reference genes
324 should be validated.

325 Finally, we performed a relative quantification analysis of different selected target genes. At
326 first, studying constitutive and differentially expressed genes between the PS and PSP conditions in,
327 we showed that consistent data are obtained when the most stable *EF-1 α* gene is used as normalizer.
328 Based in these results, we measure the expression levels of the *EgAgB1-5* genes using *EF-1 α* as a
329 reference gene in both PS and PSP experimental groups for *E. granulosus*. The results for *EgAgB1*,
330 *EgAgB3* and *EgAgB5* were consistent with RNA-seq data [8]. *EgAgB3* was the most abundant,
331 followed by *EgAgB1* with a moderate gene expression level and *EgAgB5* displaying a low gene

332 expression level in PS samples but a slightly increased expression level in PSP samples. Previous
333 proteomic analyses have detected the presence of the EgAgB1 subunit in protoscolecemes [28] as well
334 as in the hydatid fluid and germinal layer [29]. Our results were also consistent with the *EgAgB1-5*
335 expression data where *CYP-1* was used as a reference gene in *E. multilocularis* [21]. Although *CYP-1*
336 was not the most stable gene in our analysis, its expression stability remained among the most stable,
337 thus leading to analogous results. Based on these findings, we show that *EF-1 α* can be reliably used
338 as a reference gene in expression studies involving protoscolecemes and/or pepsin-treated
339 protoscolecemes.

340 In this work, we focused on the early pre-adult development of *Echinococcus* spp. that occurs
341 in the most crucial stage of the parasite life cycle, the metacestode stage, which can accidentally
342 occur in humans causing CE or AE. We identified that *EF-1 α* is a suitable and reliable reference gene
343 for gene expression normalization, both in protoscolecemes and their pepsin “activated” stage obtained
344 experimentally *in vitro*. This report validates suitable reference genes for gene expression studies in
345 two species of the class Cestoda, phylum Platyhelminthes, and provides a basis for further analysis in
346 other species, such as those that cause Taeniasis, another important neglected tropical disease as
347 indicated by WHO.

348

349 **Acknowledgments**

350 We would like to thank Dr. Rogerio Margis and Dr. Martin Cancela (CBiot, UFRGS, Porto Alegre,
351 Brazil) for qPCR experimental advice as well as Dra. Laura Kamenetzky (IMPAM, UBA-CONICET,
352 Buenos Aires, Argentina) for their comments and suggestions that helped improve the manuscript.

353

354 **Supporting information**

355 **Table S1.** RefFinder output tables for *E. ortleppi*. Based on three random samplings of 3 paired
356 samples of *E. ortleppi*, the results obtained (A, B and C) were similar to those in the *E. granulosus* s.s.
357 (G1) RefFinder output in Table 3.

358

359 **Figure S1.** Total RNA isolation. In all samples a single band of total RNA was observed on the 1.5%
360 agarose gel (left) and in the Bioanalyzer analysis (right). Total RNA extraction also displays absence
361 of genomic DNA and RNA degradation. Eg refers to *E. granulosus* and Eo to *E. ortleppi*.

362

363 **Figure S2.** Amplification specificity of the primers. A single band was observed for each amplicon in
364 the 2% agarose gel (top). The melting curves obtained for each gene (below the agarose gel) also
365 show a specific curve without any contaminants.

366

367 **Figure S3.** Amplification curves used to calculate the amplification efficiency for each selected gene.
368 Numbers from 1 to 7 (in red color) indicate the dilutions that were used for the amplification efficiency
369 calculation of the candidate reference genes.

370

371 **References**

- 372 1. World Health Organization website. Available: <http://www.who.int/echinococcosis/en/>. Accessed
373 2014 February 23.
- 374 2. Nakao M, Lavikainen A, Yanagida T, Ito A (2013) Phylogenetic systematics of the genus
375 *Echinococcus* (Cestoda: Taeniidae). *Int J Parasitol* 43: 1017-1029.
- 376 3. Thompson RC, McManus DP (2002) Towards a taxonomic revision of the genus *Echinococcus*.
377 *Trends Parasitol* 18: 452-457.
- 378 4. Thompson RC (2008) The taxonomy, phylogeny and transmission of *Echinococcus*. *Exp Parasitol*
379 119: 439-446.
- 380 5. McManus DP (2013) Current status of the genetics and molecular taxonomy of *Echinococcus*
381 species. *Parasitology* 140: 1617-1623.
- 382 6. RCA T (1995) Biology and systematics of *Echinococcus*. In: Thompson RCA, Lymbery A, editors.
383 *Echinococcus and hydatid disease*. Wallingford: CAB International. pp. 1–50. .
- 384 7. Cameron TW, Webster GA (1969) The histogenesis of the hydatid cyst (*Echinococcus* spp.). I. Liver
385 cysts in large mammals. *Can J Zool* 47: 1405-1410.
- 386 8. Zheng H, Zhang W, Zhang L, Zhang Z, Li J, et al. (2013) The genome of the hydatid tapeworm
387 *Echinococcus granulosus*. *Nat Genet* 45: 1168-1175.
- 388 9. Tsai IJ, Zarowiecki M, Holroyd N, Garcarrubio A, Sanchez-Flores A, et al. (2013) The genomes of
389 four tapeworm species reveal adaptations to parasitism. *Nature* 496: 57-63.
- 390 10. Parkinson J, Wasmuth JD, Salinas G, Bizarro CV, Sanford C, et al. (2012) A transcriptomic analysis
391 of *Echinococcus granulosus* larval stages: implications for parasite biology and host
392 adaptation. *PLoS Negl Trop Dis* 6: e1897.
- 393 11. Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription
394 polymerase chain reaction assays. *J Mol Endocrinol* 25: 169-193.
- 395 12. Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR):
396 trends and problems. *J Mol Endocrinol* 29: 23-39.
- 397 13. Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR--a perspective. *J*
398 *Mol Endocrinol* 34: 597-601.
- 399 14. Colebrook AL, Jenkins DJ, Jones MK, Tatarczuch L, Lightowlers MW (2004) Effect of cyclosporin A
400 on the survival and ultrastructure of *Echinococcus granulosus* protoscoleces in vitro.
401 *Parasitology* 129: 497-504.
- 402 15. Li J, Zhang CS, Lü GD, Wang JH, Wen H, et al. (2011) Molecular characterization of a signal-
403 regulated kinase homolog from *Echinococcus granulosus*. *Chin Med J (Engl)* 124: 2838-2844.
- 404 16. Siles-Lucas M, Nunes CP, Zaha A (2001) Comparative analysis of the 14-3-3 gene and its
405 expression in *Echinococcus granulosus* and *Echinococcus multilocularis* metacestodes.
406 *Parasitology* 122: 281-287.
- 407 17. Muzulin PM, Kamenetzky L, Gutierrez AM, Guarnera EA, Rosenzvit MC (2008) *Echinococcus*
408 *granulosus* antigen B gene family: further studies of strain polymorphism at the genomic and
409 transcriptional levels. *Exp Parasitol* 118: 156-164.

- 410 18. Cabrera G, Cabrejos ME, Morassutti AL, Cabezón C, Orellana J, et al. (2008) DNA damage, RAD9
411 and fertility/infertility of *Echinococcus granulosus* hydatid cysts. *J Cell Physiol* 216: 498-506.
- 412 19. Zhang W, Li J, Jones MK, Zhang Z, Zhao L, et al. (2010) The *Echinococcus granulosus* antigen B
413 gene family comprises at least 10 unique genes in five subclasses which are differentially
414 expressed. *PLoS Negl Trop Dis* 4: e784.
- 415 20. Koziol U, Rauschendorfer T, Zanon Rodríguez L, Krohne G, Brehm K (2014) The unique stem cell
416 system of the immortal larva of the human parasite *Echinococcus multilocularis*. *Evodevo* 5:
417 10.
- 418 21. Mamuti W, Sako Y, Xiao N, Nakaya K, Nakao M, et al. (2006) *Echinococcus multilocularis*:
419 developmental stage-specific expression of Antigen B 8-kDa-subunits. *Exp Parasitol* 113: 75-
420 82.
- 421 22. Santos GB, Espínola SM, Ferreira HB, Margis R, Zaha A (2013) Rapid detection of *Echinococcus*
422 species by a high-resolution melting (HRM) approach. *Parasit Vectors* 6: 327.
- 423 23. Matsumoto J, Müller N, Hemphill A, Oku Y, Kamiya M, et al. (2006) 14-3-3- and II/3-10-gene
424 expression as molecular markers to address viability and growth activity of *Echinococcus*
425 *multilocularis* metacestodes. *Parasitology* 132: 83-94.
- 426 24. Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, et al. (2009) Amplification efficiency:
427 linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37: e45.
- 428 25. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization
429 of real-time quantitative RT-PCR data by geometric averaging of multiple internal control
430 genes. *Genome Biol* 3: RESEARCH0034.
- 431 26. Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse
432 transcription-PCR data: a model-based variance estimation approach to identify genes suited
433 for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64: 5245-5250.
- 434 27. van Keulen H, Loverde PT, Bobek LA, Rekosh DM (1985) Organization of the ribosomal RNA genes
435 in *Schistosoma mansoni*. *Mol Biochem Parasitol* 15: 215-230.
- 436 28. Cui SJ, Xu LL, Zhang T, Xu M, Yao J, et al. (2013) Proteomic characterization of larval and adult
437 developmental stages in *Echinococcus granulosus* reveals novel insight into host-parasite
438 interactions. *J Proteomics* 84: 158-175.
- 439 29. Monteiro KM, de Carvalho MO, Zaha A, Ferreira HB (2010) Proteomic analysis of the
440 *Echinococcus granulosus* metacestode during infection of its intermediate host. *Proteomics*
441 10: 1985-1999.
- 442 30. Pan D, Bera AK, Bandyopadhyay S, Das SK, Bhattacharyya S, et al. (2010) Relative expression of
443 antigen B coding gene of bubaline isolates of *Echinococcus granulosus* in fertile and sterile
444 cysts. *J Helminthol* 84: 241-244.
- 445 31. Pan D, Bera AK, De S, Bandyopadhyay S, Das SK, et al. (2010) Relative expression of the 14-3-3
446 gene in different morphotypes of cysts of *Echinococcus granulosus* isolated from the Indian
447 buffalo. *J Helminthol* 84: 394-397.
- 448 32. Zhang W, Li J, You H, Zhang Z, Turson G, et al. (2003) Short report: *Echinococcus granulosus* from
449 Xinjiang, PR China: cDNAs encoding the EG95 vaccine antigen are expressed in different life
450 cycle stages and are conserved in the oncosphere. *Am J Trop Med Hyg* 68: 40-43.
- 451 33. Graichen DA, Gottstein B, Matsumoto J, Müller N, Zanotto PM, et al. (2007) Expression and
452 diversity of *Echinococcus multilocularis* AgB genes in secondarily infected mice: evaluating
453 the influence of T-cell immune selection on antigenic variation. *Gene* 392: 98-105.
- 454 34. Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM, et al. (2012) Reference genes for
455 accurate transcript normalization in citrus genotypes under different experimental
456 conditions. *PLoS One* 7: e31263.
- 457 35. Liu S, Cai P, Hou N, Piao X, Wang H, et al. (2012) Genome-wide identification and characterization
458 of a panel of house-keeping genes in *Schistosoma japonicum*. *Mol Biochem Parasitol* 182:
459 75-82.

- 460 36. Swijssen A, Nelissen K, Janssen D, Rigo JM, Hoogland G (2012) Validation of reference genes for
461 quantitative real-time PCR studies in the dentate gyrus after experimental febrile seizures.
462 BMC Res Notes 5: 685.
- 463 37. Balbinotti H, Santos GB, Badaraco J, Arend AC, Graichen D, et al. (2012) Echinococcus ortleppi
464 (G5) and Echinococcus granulosus sensu stricto (G1) loads in cattle from Southern Brazil. Vet
465 Parasitol 188: 255-260.

466

467 Figure legends

468 **Figure 1. Assessment of transcript abundance and comparison with RNA-seq data.** Note the
469 similarity in the boxplot distribution for each gene and the high dispersal values for several genes,
470 such as *βTUB* and *RPII*, in both species. The mean Cq value for each gene (including both PS and
471 PSP experimental groups) in *E. ortleppi* was compared with the published RNA-seq data for *E.*
472 *granulosus* spp. protoscolex (FPKM from the Cuffdiff program [9]; RPKM from the ERANGE program
473 [8]) as shown in the top right. Due to a high standard deviation, the *βTUB* gene was excluded in the
474 qPCR vs. RNA-seq comparison.

475

476 **Figure 2. Gene expression stability (left) and determination of the optimal number of reference**
477 **genes (right) by geNorm.** The top charts correspond to *E. granulosus* s.s. (G1), and the bottom
478 charts correspond to *E. ortleppi*. The dotted line represents the cut-off of 0.15 proposed by geNorm
479 authors to determine the optimal number of reference genes to be used.

480

481 **Figure 3. Relative mRNA expression of the selected target genes using *EF-1α* as reference**
482 **gene.** The constitutive (*ELP*) and differential (*RPs15* and *RPL14*) gene expression between both PS
483 and PSP conditions are shown in (A). Here, the PS treatment of the *EgELP* gene was used as
484 calibrator to obtain the $2^{-\Delta\Delta CT}$ values and, subsequently, the bar graph bar. *E. ortleppi* was used only
485 for the *ELP* gene (*EoELP*) expression analysis, and *E. granulosus* was used to analyze the *ELP* gene
486 (*EgELP*) and all other target genes. Results for the *EgAgB1-5* genes (B) are consistent with previous
487 works in *Echinococcus* spp. [8,21], where *EgAgB1* and *EgAgB3* are the most abundant in
488 protoscoleces, and *EgAgB5* exhibit a very low gene expression. The increase of *EgAgB1* and the
489 decline of *EgAgB2*, *EgAgB3* and *EgAgB4* genes in the PSP group is comparable to previously
490 described for the immature adult form of *E. multilocularis* [21]. The statistical significance between
491 *EgAgB* subunits in both PS and PSP groups (*) as well as only for PSP (¤) or only for PS (§) are
492 shown in the inset. Here, the PS treatment of the *EgAgB3* subunit was used as a calibrator. A
493 significance level of $\alpha=0.05$ was considered.

494

495 Tables

496

Table 1. Descriptions of candidate reference genes

Gene symbol	Gene name	Function	^a Databases & Acc. number	^b GeneDB Acc. Number
<i>βTUB</i>	Beta tubulin 2C chain	Cytoskeletal structural protein (microtubules)	LophDB EGC04893	EgrG_002026000
<i>βACT-1</i>	Beta actin-1	Cytoskeletal structural protein (microfilaments)	GenBank L07773	EgrG_000406900
<i>GAPDH</i>	Glyceraldehyde-3-phosphate dehydrogenase	Glycolytic enzyme	LophDB EGC00305	EgrG_000254600
<i>NDUFV2</i>	NADH dehydrogenase ubiquinone flavoprotein 2	Oxidoreductase activity	GeneDB	EgrG_001114700
<i>RPL13</i>	L13 ribosomal protein	Structural component of the large 60S ribosomal subunit	LophDB EGC01259	EgrG_000517800
<i>ETIF4A-III</i>	Eukaryotic translation initiation factor 4A-III	Translation	LophDB EGC00363	EgrG_001193600
<i>TBP</i>	TATA-Box binding protein	Transcription	GeneDB	EgrG_000972300
<i>CYP-1</i>	Cyclophilin	Protein folding and protein interactions	GenBank AF430707	EgrG_000920600
<i>RPII</i>	DNA directed RNA polymerase II subunit RPB2	Polymerization	GenBank FN566850	EgrG_000604200
<i>EF-1α</i>	Elongation Factor 1 alpha	Protein synthesis	GenBank AB306934.1	EgrG_000982200
<i>MAPK3</i>	Mitogen activated protein kinase 3	Signal transduction	GenBank HQ585923	EgrG_000803700

497 ^a Databases and accession numbers used in this work.498 ^b Accession number obtained from GeneDB database after the *E. granulosus* genome annotation.

499

Table 2. Details of each primer designed for the candidate reference genes and target genes.

Gene symbol	Primer sequence (5'-3') forward/reverse	Amplicon length (bp)	Tm	Amplification efficiency (%)
<i>βTUB</i>	CGTTCAGGCTACCGCCGGTT/GGAGCCCGTTGGGTCTACTCCGT	146	85.9	77.0
<i>βACT-1</i>	CGCGATCTCACCGACTGG/CTCCAGAGAGGAGCTAGTG	161	87.5	78.0
<i>GAPDH</i>	ACTCCGTCAATGTTGTCGCTGTCA/TAACCAACTTGCCGCCATCAACCT	128	84.0	92.1
<i>NDUFV2</i>	GACACCGCCATCAATAACAGGGAT/CCATTCTGCCGTTGTGCAATGT	146	85.3	85.2
<i>RPL13</i>	GAAGTGGCAATTCATGGTAAGGACG/CACAAGCAGGTTTGGGAGCGA	110	87.9	84.0
<i>ETIF4A-III</i>	AGTTCTCATTCTGTGCGCTACACGC/GACATTAGTGCCGCCATAGCAGG	115	85.4	82.4
<i>TBP</i>	TTCCAGCGCTCAGGCACACA/CGTGCGCTTTGAGCTATCCGTCT	165	86.5	87.5
<i>CYP-1</i>	CGACATCTCCATTGGCGGTAAGC/TTGTATCCGAAACCCTTCTCACCG	120	86.8	90.0
<i>RPII</i>	CATCTGCCGCCCGCTTGTTA/TCATGGCTGTCTCCTCCTCCAAA	163	85.6	90.2

<i>EF-1α</i>	TTTGAGAAAGAGGCGGCTGAGATG/TAATAAAGTCACGATGACCGGGCG	174	87.9	92.0
<i>MAPK3</i>	AAAGTACAGCAGTTGAGTCGCGAGC/GCTTCAAATCTCGGTGCAAAACGTT	106	83.3	87.0
<i>RPL14</i>	TCCTTATCGATGGACCTTGTTCCGG/TTCTGCCTGCGCCAATTATCCT	136	85.0	88.0 ^a
<i>RPs15</i>	AATACGTCGAGGTCTCGGAACCAA/CAGGTGAGTTTTAACAACCTGCCGGT	109	84.3	91.0 ^a
<i>ELP</i>	GACACGCGATCAGTCGAAAATGC/TGTTGCCCTTGCGAATGTTGC	107	84.7	90.4 ^a
<i>EgAgB1</i>	AAATGTTTGGCGAAGTGAAGT/ACCTGAGTGCCATGCGTAGCTTCT	126	84.5	90.0 ^{a, b}
<i>EgAgB2</i>	AAAGCACACATGGGGCAAGTG/GTGTCCCGACGCATGACTTA	218	85.4	86.4 ^{a, b}
<i>EgAgB3</i>	GAAGGGTGTGATGAAGGCCAT/ATACTCCTTCAGTGCCATGCGTGCC	145	85.4	90.0 ^{a, b}
<i>EgAgB4</i>	CGAGAGATGCAAGTGCCTCAT/GTGTCCCGACGCATGACTTA	219	86.0	86.0 ^{a, b}
<i>EgAgB5</i>	GAAGATGACATCGATTGAAA/GATCGAGCTTTTGTCTCTGCC	155	83.4	N/C ^{b, c}

501 ^a Values correspond to the amplification efficiency average of the samples (both PS and PSP groups) used in the gene
502 expression experiments and obtained with the LinRegPCR software.

503 ^b Amplification products confirmed by sequencing.

504 ^c N/C = not calculated

505

506
507
508
509
510
511

Table 3. Gene expression stability for *E. granulosus* s.s. (G1) (top) and *E. ortleppi* (bottom) as assessed by RefFinder

	Ranking Order (Better→Good→Average)											512
Method	1	2	3	4	5	6	7	8	9	10	11	513
Delta CT	TBP	EF-1α	GAPDH	MAPK3	CYP-1	RPII	βACT-1	NDUFV2	RPL13	ETIF4A-III	βTUB	514
BestKeeper	ETIF4A-III	βACT-1	GAPDH	TBP	NDUFV2	CYP-1	EF-1α	MAPK3	RPL13	RPII	βTUB	
Normfinder	TBP	GAPDH	βACT-1	NDUFV2	EF-1α	MAPK3	CYP-1	RPII	RPL13	ETIF4A-III	βTUB	
geNorm	RPL13 EF-1α		MAPK3	CYP-1	RPII	GAPDH	TBP	NDUFV2	βACT-1	ETIF4A-III	βTUB	
Recommended comprehensive ranking	TBP	EF-1α	GAPDH	βACT-1	MAPK3	RPL13	CYP-1	ETIF4A-III	NDUFV2	RPII	βTUB	
	Ranking Order (Better→Good→Average)											
Method	1	2	3	4	5	6	7	8	9	10	11	
<u>Delta CT</u>	TBP	EF-1α	CYP-1	RPII	NDUFV2	MAPK3	GAPDH	βACT-1	RPL13	ETIF4A-III	βTUB	
<u>BestKeeper</u>	TBP	EF-1α	CYP-1	GAPDH	MAPK3	NDUFV2	RPII	βACT-1	RPL13	βTUB	ETIF4A-III	
<u>Normfinder</u>	TBP	EF-1α	CYP-1	NDUFV2	RPII	MAPK3	βACT-1	GAPDH	RPL13	ETIF4A-III	βTUB	
geNorm	TBP EF-1α		CYP-1	RPII	GAPDH	RPL13	MAPK3	NDUFV2	βACT-1	ETIF4A-III	βTUB	
Recommended comprehensive ranking	TBP	EF-1α	CYP-1	RPII	NDUFV2	GAPDH	MAPK3	βACT-1	RPL13	ETIF4A-III	βTUB	

515 **Table S1.** RefFinder outputs tables for *E. ortleppi*. Based on three random sampling of 3 paired samples of *E. ortleppi* is possible to obtain a similar result (A,
 516 B and C) when compared with the *E. granulosus* s.s. (G1) RefFinder output on Table 3.

517

518 **A**

Method	Ranking Order (Better→Good→Average)										
	1	2	3	4	5	6	7	8	9	10	11
Delta CT	RPL13	CYP-1	NDUFV2	βTUB	EF-1α	ETIF4A-III	TBP	GAPDH	MAPK3	βACT-1	RPII
BestKeeper	GAPDH	MAPK3	EF-1α	βTUB	NDUFV2	TBP	CYP-1	ETIF4A-III	RPL13	βACT-1	RPII
Normfinder	RPL13	CYP-1	NDUFV2	βTUB	EF-1α	ETIF4A-III	TBP	GAPDH	MAPK3	βACT-1	RPII
Genorm	NDUFV2 RPL13		CYP-1	βTUB	EF-1α	ETIF4A-III	TBP	GAPDH	MAPK3	βACT-1	RPII
Recommended comprehensive ranking	RPL13	NDUFV2	CYP-1	βTUB	EF-1α	GAPDH	MAPK3	ETIF4A-III	TBP	βACT-1	RPII

519

520 **B**

Method	Ranking Order (Better→Good→Average)											
	1	2	3	4	5	6	7	8	9	10	11	
Delta CT	TBP	CYP-1	RPII	EF-1α	NDUFV2	MAPK3	βACT-1	GAPDH	RPL13	ETIF4A-III	βTUB	
BestKeeper	CYP-1	βACT-1	EF-1α	TBP	NDUFV2	MAPK3	RPII	GAPDH	RPL13	ETIF4A-III	βTUB	
Normfinder	TBP	RPII	NDUFV2	CYP-1	EF-1α	βACT-1	MAPK3	GAPDH	RPL13	ETIF4A-III	βTUB	
Genorm	CYP-1 EF-1α		TBP	RPII	MAPK3	NDUFV2	GAPDH	RPL13	βACT-1	ETIF4A-III	βTUB	
Recommended comprehensive ranking	CYP-1	TBP	EF-1α	RPII	NDUFV2	βACT-1	MAPK3	GAPDH	RPL13	ETIF4A-III	βTUB	

521

522 **C**

Method	Ranking Order (Better→Good→Average)										
	1	2	3	4	5	6	7	8	9	10	11
Delta CT	EF-1α	NDUFV2	GAPDH	RPII	TBP	RPL13	CYP-1	MAPK3	βACT-1	βTUB	ETIF4A-III
BestKeeper	NDUFV2	EF-1α	GAPDH	TBP	MAPK3	RPL13	CYP-1	RPII	TUB	ETIF4	βACT-1
Normfinder	EF-1α	NDUFV2	RPII	TBP	GAPDH	MAPK3	RPL13	CYP-1	βACT-1	βTUB	ETIF4A-III
Genorm	CYP-1 RPL13		GAPDH	TBP	RPII	EF-1α	NDUFV2	MAPK3	βACT-1	βTUB	ETIF4A-III
Recommended comprehensive ranking	EF-1α	NDUFV2	GAPDH	RPL13	TBP	CYP-1	RPII	MAPK3	βACT-1	βTUB	ETIF4A-III

Figures

Figure 1

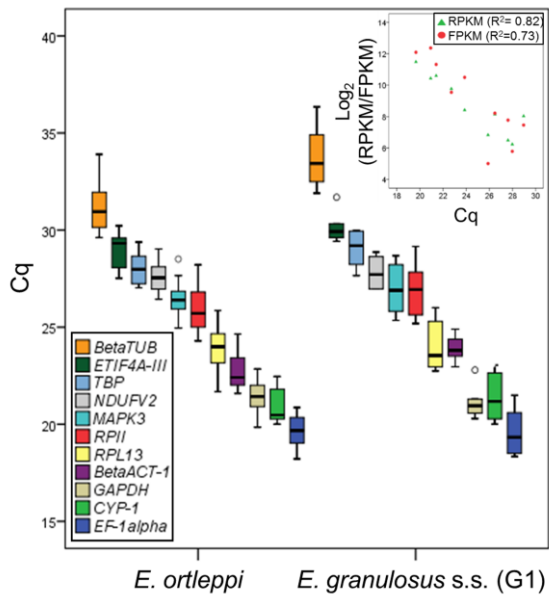


Figure 2

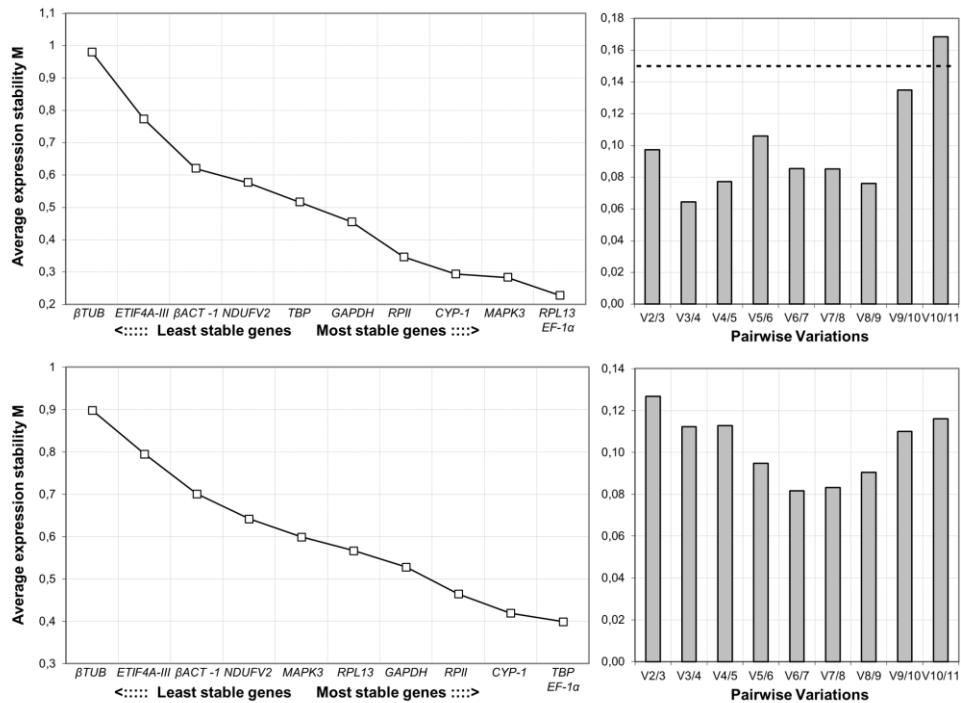


Figure 3

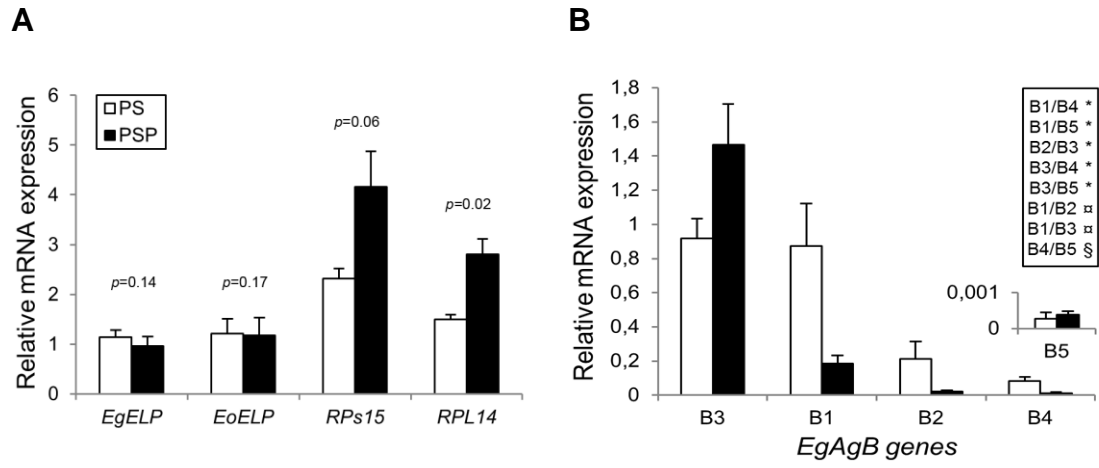


Figure S1

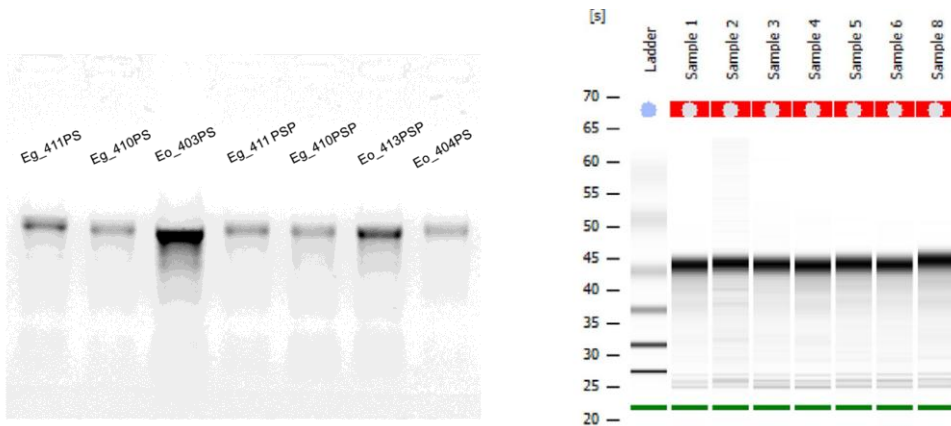


Figure S2

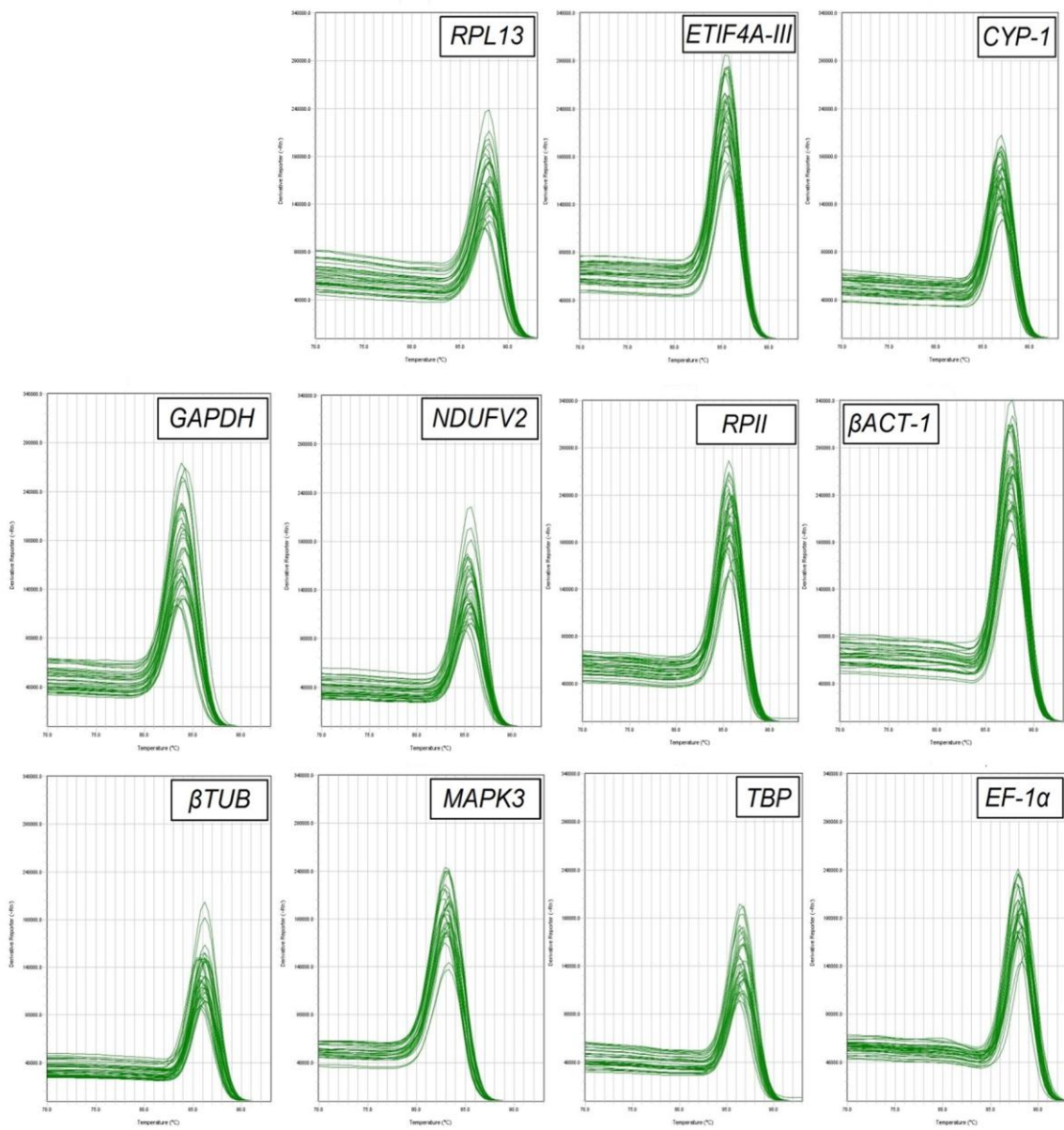
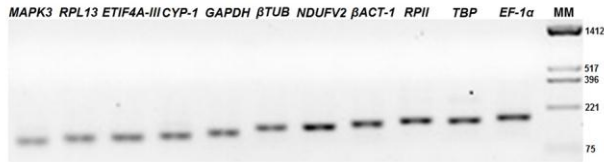
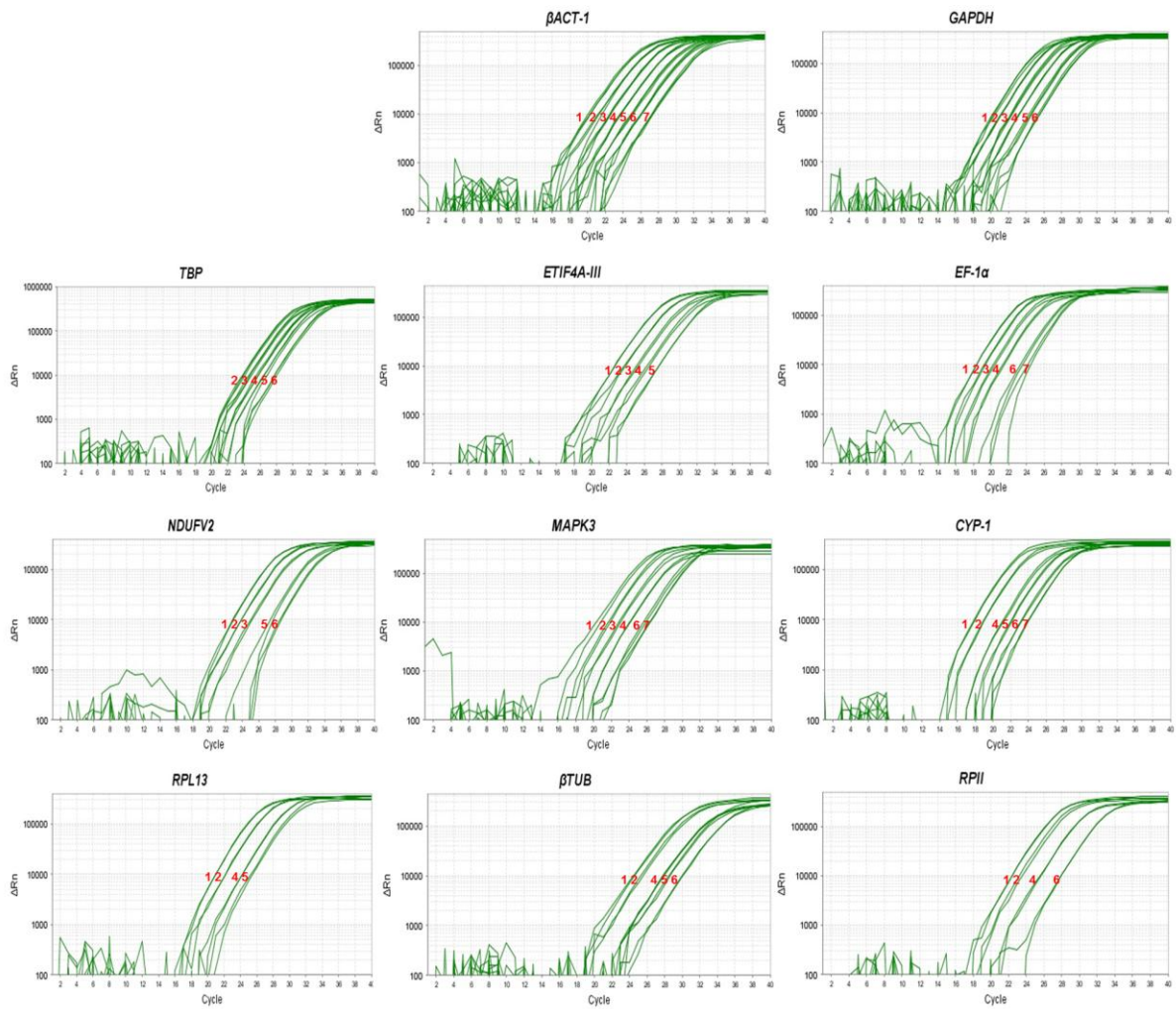


Figure S3



Capítulo 3. Discussão geral e Conclusões

Atualmente, estão disponíveis varias técnicas para quantificar os níveis de expressão gênica global, como os micro arranjos ou mais recentemente o sequenciamento do RNA (RNA-seq), porém, a PCR em tempo real continua sendo a técnica padrão tanto para validar esses dados de sequenciamento massivo como para estudos de expressão de genes alvo previamente selecionados. A sensibilidade, especificidade, reprodutibilidade e amplo alcance dinâmico (Bustin *et al.*, 2004, Kubista *et al.*, 2006) são as principais vantagens que fazem desta técnica sua ampla utilização entre as mais diversas áreas de pesquisa. Na relação parasito-hospedeiro uma série de interações permite a sobrevivência, crescimento e permanência do parasito no hospedeiro. O estudo dos níveis de expressão gênica ao longo do ciclo de vida do parasito através de qPCR permite identificar os componentes chaves que estão sendo expressos e que promovem o estabelecimento da doença.

Para analisar e quantificar genes específicos, genes de referência ou normalizadores são utilizados para mensurar as mudanças na expressão gênica. Estes genes devem ser previamente validados para confirmar que seu comportamento em certas condições experimentais é realmente estável. Neste trabalho, uma vez detectado os genes mais estáveis (ver Capítulo 2), foram avaliados genes previamente descritos como de expressão constitutiva (Koziol *et al.*, 2014) e diferencialmente expressos entre PS e PSP (Parkinson *et al.*, 2012). Assim, a expressão do gene ezrin-radixin-moesin (ERM)-like protein (*ELP*) e dos genes que codificam as proteínas ribossômicas L14 e S15 foi analisada. Como resultado (ver Capítulo 2), para o gene *ELP* não foram observadas diferenças significativas entre os tratamentos analisados tanto para *E. ortleppi* como para *E. granulosus* s.s. (G1). Em relação aos outros alvos analisados, somente foram observadas diferenças significativas para um dos genes (valores de p igual a 0,02 para L14; e 0,06 para S15, com $\alpha=0,05$), embora seja possível observar uma leve tendência ao aumento nos níveis de expressão em PSP para cada um dos alvos analisados, consistente com os dados de *expressed sequence tags* (ESTs) obtidos por Parkinson *et al.* (2012) (PS=11 e PSP=21 para RPL14, PS=2 e PSP=24 para RPLS15). O tratamento com pepsina, enzima que simularia o contato do parasito com os sucos gástricos do HD, tem o principal papel de ativar e/ou “preparar” ao parasito para estabelecer-se no seu novo microambiente,

o trato gastrointestinal, e se desenvolver na forma adulta. Desta maneira, *Echinococcus* spp. abandona seu estado de metabolismo basal, começando assim a síntese de numerosas proteínas, o que poderia explicar o aumento no nível de expressão destas proteínas ribossômicas em PSP para cada um dos genes analisados. Uma diferença ainda maior entre PS e PSP poderia se esperar com um maior tempo de exposição da enzima a uma concentração um pouco menor, como realizada por Parkinson *et al.* (2012); porém, estas condições não seriam as que na realidade se supõe que acontecem dentro do HD, o que provavelmente provocaria desvios na análise dos resultados. Por outro lado, aumentando o número de réplicas biológicas, o desvio padrão provavelmente diminuiria, obtendo desta maneira diferenças significativas entre os dois tratamentos analisados. Numerosas são as variações provocadas pelo material de partida (PS), como o local onde é coletada a amostra, o HI (no nosso caso, somente bovinos) e a idade do animal, o órgão em que se encontra o cisto, o tempo/momento em que ocorreu a infecção, entre outros (ver Anexo, Tabela 1). Assim, devido à consistência dos nossos dados das proteínas ribossomais com os publicados previamente, estudamos a expressão gênica das subunidades do antígeno B de *E. granulosus* s.s. (G1) (ver Capítulo 2).

Em conclusão, no presente trabalho identificamos genes estavelmente expressos em PS e PSP de *E. granulosus* s.s. (G1) e *E. ortleppi* para serem utilizados como genes de referência em posteriores estudos de expressão gênica. Devido à similaridade entre os resultados obtidos nas duas espécies, estes mesmos genes poderiam ser utilizados para outras espécies de *E. granulosus* s.l. Nossos dados de RT-qPCR apresentaram uma alta correlação com dados de RNA-seq previamente publicados. Por último, a expressão gênica de alvos selecionados (genes que codificam proteínas ribossômicas e subunidades do AgB) permitiu avaliar a aplicabilidade do gene normalizador e comparar com dados previamente obtidos do transcrito e de RT-qPCR. Este primeiro relato de genes de referência dentro da Classe Cestoda, Phylum Platyhelminthes, abre as portas para a realização de novos estudos de expressão utilizando alvos específicos, além de proporcionar uma base para estudos futuros de validações dentro do grupo.

Perspectivas

Devido ao complexo ciclo de vida de *Echinococcus* spp. envolvendo dois hospedeiros e diversos estágios durante o desenvolvimento, um estudo mais abrangente de genes de referência (partindo dos dados de RNA-seq disponíveis) permitiria uma análise mais completa de genes expressos durante a vida do parasito. Uma das limitações para estes casos é a difícil obtenção do material em estágios que não sejam o larval. A cultura *in vitro* de PS para o desenvolvimento na forma adulta (ver Anexo, Figura 1) já está sendo padronizada no nosso laboratório, porém, uma grande diminuição e perda de material é observada durante o desenvolvimento, indicando que novos procedimentos precisarão ser padronizados para otimizar a realização de extrações de RNA total.

Outra perspectiva a destacar é a realização do sequenciamento (e quantificação) massivo de transcritos (RNA-seq) em diferentes estágios do desenvolvimento de *E. ortleppi* para fazer comparações com os dados de RNA-seq de *E. granulosus*; identificar alvos específicos envolvidos no desenvolvimento do parasito; analisar posteriormente estes alvos através de diversas técnicas ou ferramentas de genômica funcional como PCR em tempo real, RNA de interferência; entre outros.

Referências bibliográficas

- Andersen, C.L.; Jensen J.L.; Orntoft T.F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research* 64(15):5245-5250, 2004.
- Arend, A.C.; Alves-Junior, L.; Chemale, G.; Zaha, A.; Haag, K.L. Cloning and characterization of sequences encoding an *Echinococcus granulosus* antigen (AgB) of diagnostic value. *The Am. J. Trop. Med. Hyg.* 67, 299– 300, 2002.
- Badaraco, J.L.; Ayala, F.J.; Bart, J.M.; Gottstein, B.; Haag, K.L. Using mitochondrial and nuclear markers to evaluate the degree of genetic cohesion among *Echinococcus* populations. *Exp. Parasitol.* 119, 453–459, 2008.
- Balbinotti H.; Santos G.B.; Badaraco J.; Arend A.C.; Graichen D.; *et al.* *Echinococcus ortleppi* (G5) and *Echinococcus granulosus* sensu stricto (G1) loads in cattle from Southern Brazil. *Vet Parasitol* 188: 255-260, 2012.
- Barrett, J.; Saghir, 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.
- Bowles, J.; Blair, D.; McManus, D.P. Genetic variants within the genus *Echinococcus* identified by mitochondrial sequencing. *Mol. Biochem. Parasitol.* 54, 165–174, 1992.
- Budke, C.; Deplazes, P.; Torgerson, P. Global socioeconomic impact of cystic echinococcosis. *Emerging Infections Diseases*, 12:296–302, 2006.
- Bustin, S.A.; Nolan, T. Pitfalls of quantitative real-time reverse transcription polymerase chain reaction. *J Biomol Tech* 15:155-166, 2004.

- Bustin, S.A.; Benes, V.; Nolan, T.; Pfaffl, M.W. Quantitative real-time RT-PCR-a perspective. *J Mol Endocrinol* 34:597-601, 2005.
- Cabrera, G.; Cabrejos, M.E.; Morassutti, A.L.; Cabezon, C.; Orellana, J.; *et al.* DNA damage, RAD9 and fertility/infertility of *Echinococcus granulosus* hydatid cysts. *J Cell Physiol* 216: 498–506, 2008.
- 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.
- Chemale, G.; Ferreira, H.B.; Barrett, J.; Brophy, P.M.; Zaha, A. *Echinococcus granulosus* antigen B hydrophobic ligand binding properties. *Biochim Biophys Acta* 1747: 189–194, 2005.
- Colebrook, A.L.; Jenkins, D.D.; Lightowers, M.W. Anti-parasitic effect of cyclosporin A on *Echinococcus granulosus* and characterization of the associated cyclophilin protein. *Parasitology* 125(Pt 5):485-93, 2002.
- Conchedda, M.; Seu, V.; Capra, S.; Caredda, A.; Pani, S.P; *et al.* Cystic echinococcosis in sheep in Sardinia. Changing pattern and present status. *Acta Tropica*, 122:52–58, 2012.
- Craig, P.; McManus D.P.; Lightowers, M.W.; Chabalgoity, J.A.; Garcia, H.H.; *et al.* Prevention and control of cystic echinococcosis. *Lancet Infectious Diseases*, 7:385–394, 2007.
- D'Alessandro, A.; Rausch, R.L. New aspects of neotropical polycystic (*Echinococcus vogeli*) and unicystic (*Echinococcus oligarthrus*) echinococcosis. *Clin. Microbiol. Rev.* 21, 380–401, 2008.
- de la Rue, M.L.; Dinkel, A.; Mackenstedt, U.; Romig, T. New data on *Echinococcus spp.* in southern Brazil. *Rev. Inst. Med. Trop. S. Paulo* 48, 103–104, 2006.
- de Siqueira, N.G.; de Siqueira, C.M.; Rodrigues-Silva, R.; Soares, C., Póvoa M.M. Polycystic echinococcosis in the state of Acre, Brazil: contribution to patient diagnosis, treatment and prognosis. *Mem Inst Oswaldo Cruz.* Aug;108(5):533-40, 2013.

- Elissondo, M.C.; Pensel, P.E.; Denegri, G. Optimización del cultivo in vitro en dirección vesicular de *Echinococcus granulosus*. XXXIV Jornadas Internacionales de Hidatidología, San Salvador de Jujuy, Jujuy, Argentina, 2011.
- Fernández, 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.
- Frosch, P.; Hartmann, M.; Mühlischlegel, F.; Frosch, M. Sequence heterogeneity of the echinococcal antigen B. *Mol Biochem Parasitol* 64, 171–175, 1994.
- Graichen, D.A.; Gottstein, B.; Matsumoto, J.; Muller, N.; Zanotto, P.M.; *et al.* Expression and diversity of *Echinococcus multilocularis* AgB genes in secondarily infected mice: evaluating the influence of T-cell immune selection on antigenic variation. *Gene* 392: 98–105, 2007.
- Haag, K.L.; Araujo, 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.; 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.
- Knapp, J.; Nakao, M.; Yanagida, T.; Okamoto, M.; Saarma, U.; *et al.* Phylogenetic relationships within *Echinococcus* and *Taenia* tapeworms (Cestoda: Taeniidae): an inference from nuclear protein-coding genes. *Mol. Phylogenet. Evol.* 61, 628–638, 2011.
- Kamenetzky, L.; Muzulin, P.M.; Gutierrez, A.M.; Angel, S.O.; Zaha, A.; *et al.* High polymorphism in genes encoding antigen B from human infecting strains of *Echinococcus granulosus*. *Parasitology* 131(Pt 6):805-15, 2005.
- Koziol, U.; Rauschendorfer, T.; Zanon Rodríguez, L.; Krohne, G.; Brehm, K. The unique stem cell system of the immortal larva of the human parasite *Echinococcus multilocularis*. *Evodevo* 5: 10. 2014.

- Kubista, M.; Andrade, J.M.; Bengtsson, M.; Forootan, A.; Jonak, J.; *et al.* The real-time polymerase chain reaction. *Mol Aspects Med* 27:95-125, 2006.
- Li, J.; Zhang, C.; Lü G.; Wang, J.; Wen, H.; *et al.* Molecular characterization of a signal-regulated kinase homolog from *Echinococcus granulosus*. *Chin. Med. J.* 124(18):2838-2844, 2011.
- Lorenzo, C.; Ferreira, H.B.; Monteiro, K.M.; Rosenzvit, M.; Kamenetzky L.; *et al.* Comparative analysis of the diagnostic performance of six major *Echinococcus granulosus* antigens assessed in a double-blind, randomized multicenter study. *J. Clin. Microbiol.* 43: 2764–2770, 2005.
- Mamuti, W.; Sako, Y.; Xiao, N.; Nakaya, K.; Nakao, M.; *et al.* *Echinococcus multilocularis*: developmental stage-specific expression of antigen B 8-kDa-subunits, *Exp. Parasitol.* 113, 75–82, 2006.
- Matsumoto, J.; Muller, N.; Hemphill, A.; Oku, Y.; Kamiya, M.; *et al.* 14-3-3- and II/3-10-gene expression as molecular markers to address viability and growth activity of *Echinococcus multilocularis* metacestodes. *Parasitology* 132: 83–94, 2006.
- McManus, D.P. e Smyth, J.D. Hidatidoses: changing concepts in epidemiology and speciation. *Parasitol. Today* 2: 163-167, 1986.
- McManus, D.P. Reflections on the biochemistry of *Echinococcus*: past, present and future. *Parasitology* 136, 1643–1652, 2009.
- McManus, D.P. Current status of the genetics and molecular taxonomy of *Echinococcus* species. *Parasitology* 140: 1617-1623, 2013.
- Monteiro, D.U.; Botton, S.A.; Tonin A.A.; Azevedo, M.I.; Graichen, D.A.S.; *et al.* *Echinococcus canadensis* (G7) and *Echinococcus granulosus* sensu stricto (G1) in swine of southern Brazil. *Veterinary Parasitology.* 2014.
- Monteiro, K.M.; de Carvalho M.O.; Zaha, A.; Ferreira, H.B. Proteomic analysis of the *Echinococcus granulosus* metacestode during infection of its intermediate host. *Proteomics* 10: 1985-1999. 2010.

- Muzulin, P.M.; Kamenetzky, L.; Gutierrez, A.M.; Guarnera, E.A.; Rosenzvit, M.C. *Echinococcus granulosus* antigen B gene family: Further studies of strain polymorphism at the genomic and transcriptional levels. *Exp Parasitol* 118: 156–164, 2007.
- Nakao, M.; McManus, D.P.; Schantz, P.M.; Craig, P.S.; Ito, A. A molecular phylogeny of the genus *Echinococcus* inferred from complete mitochondrial genomes. *Parasitology* 134, 713–722, 2007.
- Nakao, M.; Lavikainen A.; Yanagida T.; Ito, A. Phylogenetic systematics of the genus *Echinococcus* (Cestoda: Taeniidae). *Int J Parasitol* 43: 1017-1029, 2013.
- Olson, P.D.; Zarowiecki, M.; Kiss, F.; Brehm, K. Cestode genomics - progress and prospects for advancing basic and applied aspects of flatworm biology. *Parasite Immunol.* 34(2-3):130-50, 2012.
- Oriol, R.; Williams, J.F.; Pérez Esandi, M.V.; Oriol, C. Purification of lipoprotein antigens of *Echinococcus granulosus* from sheep hydatid fluid. *Am J Trop Med Hyg* 20: 569–574, 1971.
- Pan American Health Organization. Zoonoses and communicable diseases common to man and animals, vol. III, 3rd ed. Washington, DC, (Scientific and Technical Publication No. 580), 2001.
- Pan, D.; Bera, A.K.; Bandyopadhyay, S.; Das, S.K.; Bhattacharyya, S.; *et al.* Relative expression of antigen B coding gene of bubaline isolates of *Echinococcus granulosus* in fertile and sterile cysts. *J Helminthol.* 2009.
- Parkinson, J.; Wasmuth, J.D.; Salinas, G.; Bizarro, C.V.; Sanford, C.; *et al.* A transcriptomic analysis of *Echinococcus granulosus* larval stages: implications for parasite biology and host adaptation. *PLoS Negl Trop Dis* 6: e1897, 2012.
- Pfaffl, M.W.; Tichopad, A.; Prgomet, C.; Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnology Letters* 26(6):509-515, 2004.

- Rigano, R.; Profumo, E.; Bruschi, F.; Carulli G.; Azzara, A.; *et al.* Modulation of human immune response by *Echinococcus granulosus* antigen B and its possible role in evading host defenses. *Infect Immun* 69: 288–296, 2001.
- Rigano, R.; Buttari, B.; Profumo, E.; Ortona, E.; Delunardo, F.; *et al.* *Echinococcus granulosus* antigen B impairs human dendritic cell differentiation and polarizes immature dendritic cell maturation towards a Th2 cell response. *Infect Immun* 75: 1667–1678, 2007.
- Rodrigues-Silva, R.; Peixoto, J.R.; de Oliveira, R.M.; Magalhães Pinto, R.; Gomes, D.C.; *et al.* An autochthonous case of *Echinococcus vogeli* Rausch & Bernstein, 1972 polycystic echinococcosis in the state of Rondônia, Brazil. *Mem Inst Oswaldo Cruz*. Jan;97(1):123-6, 2002.
- Santos, G.B.; Badaraco, J.; Arend, A.C.; Graichen, D.A.; *et al.* *Echinococcus ortleppi* (G5) and *Echinococcus granulosus* sensu stricto (G1) loads in cattle from Southern Brazil. *Vet Parasitol* 188(3-4):255-60, 2012.
- Shaghir, N.; Conde P.J.; Brophy P.M.; Barrett J. Biochemical characterization of a hydrophobic ligand binding protein from the tapeworm *Hymenolepis diminuta*. *Int. J. Parasitol.* 31, 653–660, 2001.
- Shepherd, 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.
- Siles-Lucas, M.; Nunes, C.P.; Zaha, A. Comparative analysis of the 14-3-3 gene and its expression in *Echinococcus granulosus* and *Echinococcus multilocularis* metacestodes. *Parasitology*. 122:281–7, 2001.
- Soares, C.; Moreira-Silva, C.A.; Alves, M.M.; Nunes, H.M.; Amaral, I.A.; *et al.* Polycystic echinococcosis in the Eastern Brazilian Amazon: an update. *Rev Soc Bras Med Trop*. 37 Suppl 2:75-83, 2004.
- Soares, M.C.P.; dos Santos Rodrigues, A.L.; Moreira Silva, C.A.; de Figueiredo Brito, E.M.; Gomes-Gouvêa, M.S.; *et al.* Anatomico-clinical and molecular description of liver neotropical

- echinococcosis caused by *Echinococcus oligarthrus* in human host. *Acta Trop.* 125,110–114, 2013.
- Stamatakis, M.; Sargedi, C.; Stefanaki, Ch.; Safioleas, C.; Matthaopoulou, I.; Safioleas, M. Anthelmintic treatment: an adjuvant strategy against *Echinococcus granulosus*. *Parasitol. Int.* 58, 115–120, 2009.
- Thompson, R.C.A.; Kumaratilake, L.M.; Eckert, J. Observations on *Echinococcus granulosus* of the cattle origin in Switzerland. *Int. J. Parasitol.* 14, 283–291, 1984.
- Thompson, R.C.A. Biology and systematics of *Echinococcus*. In Thompson R.C.A. *Echinococcus* and hydatid disease. CAB International, Wallingford, 1-50, 1995.
- Thompson, R.C.A. e McManus D.P. Aetiology: parasites and life cycles. In: WHO/OIE manual on echinococcosis in humans and animals: a public health problem of global concern. OIE/WHO. Paris, cap. 1, 2001.
- Torgerson, P.R.; Keller, K.; Magnotta, M.; Ragland, N. The global burden of alveolar echinococcosis. *PLoS Neglected Tropical Diseases*, 4:e722, 2010.
- Tsai, I.J.; Zarowiecki, M.; Holroyd, N.; Garcarrubio, A.; Sanchez-Flores, A.; *et al.* The genomes of four tapeworm species reveal adaptations to parasitism. *Nature* 496: 57-63, 2013.
- Vandesompele, J.; De Preter, K.; Pattyn, F.; Poppe, B.; Van Roy, N.; *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3:research0034, 2002.
- Virginio, V.G.; Hernández, A.; Rott, M.B.; Monteiro, K.M.; Zandonai, A.F.; *et al.* 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.
- World Health Organization (WHO), <http://www.who.int/en/> (Março, 2014)

World Health Organization. The control of neglected zoonotic diseases: community-based interventions for prevention and control. Geneva, 2011a (WHO/HTM/NTD/NZD/2011.1).

World Health Organization. The interagency meeting on planning the prevention and control of neglected zoonotic diseases (NZDs). Geneva, 2011b (WHO/HTM/NTD/NZD/2011.3).

Zhang, W.; Li, J.; Jones, M.K.; Zhang, Z.; Zhao, L.; *et al.* The *Echinococcus granulosus* antigen B gene family comprises at least 10 unique genes in five subclasses which are differentially expressed. *Trop Dis.* 4:e784, 2010.

Zheng, H.; Zhang, W.; Zhang, L.; Zhang, Z.; Li, J.; *et al.* The genome of the hydatid tapeworm *Echinococcus granulosus*. *Nat Genet* 45: 1168-1175, 2013.

Anexos

Tabela 1. Descrição das amostras coletadas neste trabalho.

Amostra	Cidade	Órgão	Espécie	Tratamentos		
				No induzido (NI)	Induzido (I)	Indução Des. (ID)
382	São Gabriel	Fígado	<i>E. granulosus</i> s.s. (G1)	X	X	Não
386	Dom Pedrito	Pulmão	<i>E. ortleppi</i>	X	X	Não
388	Bagé	Pulmão	<i>E. ortleppi</i>	X	X	Não
389	Bagé	Pulmão	<i>E. ortleppi</i>	Não	X	Não
390	Itaqui	Pulmão	<i>E. ortleppi</i>	X	X	Não
391	Sem Reporte	Pulmão	<i>E. ortleppi</i>	X	X	Não
403	Herval	Fígado	<i>E. ortleppi</i>	X	X	X (4 dias)
404	Herval	Pulmão	<i>E. ortleppi</i>	X	Não	Não
410	Itaqui	Pulmão	<i>E. granulosus</i> s.s. (G1)	X	X	Não
411	Itaqui	Pulmão	<i>E. granulosus</i> s.s. (G1)	X	X	Não

Observações: Na coluna dos “Tratamentos”, o sinalado com “X” indica as amostras utilizadas no trabalho, porém, o tratamento de Indução ao Desenvolvimento (ID) não foi analisado.

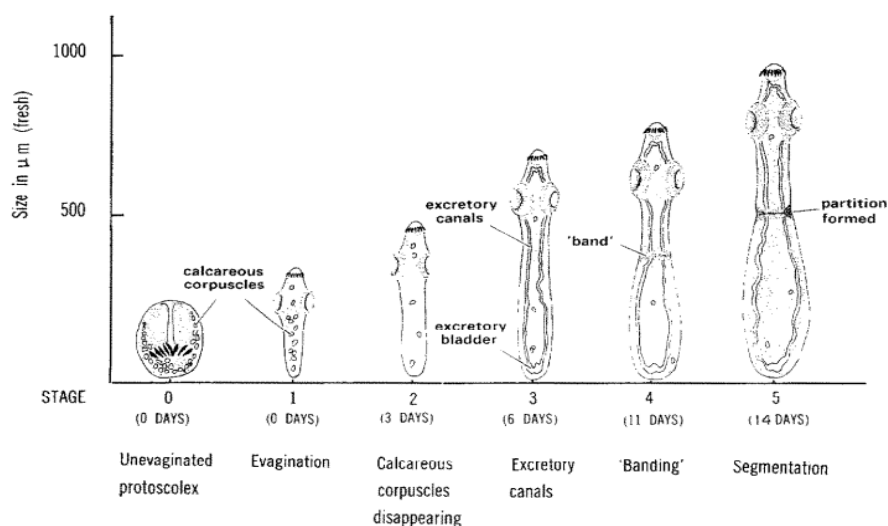


Figura 1. Desenvolvimento inicial da forma pré-adulta de *E. granulosus*. No início, os PS encontram-se invaginados, com seu metabolismo basal. Logo depois da indução ao desenvolvimento, eles evaginam e começam a se desenvolver até o verme adulto segmentado. O desaparecimento dos corpúsculos calcáreos, e conseqüentemente, a formação dos canais excretórios, são os sinais morfológicos visíveis do desenvolvimento. Fonte: Smyth and Davies (1974).

