

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

LAYSA GOMES PORTILHO

**ANÁLISE *IN SILICO* DO SISTEMA UBIQUITINA-
PROTEASSOMA EM *Biomphalaria glabrata***

PATOS DE MINAS – MG
FEVEREIRO DE 2019

LAYSA GOMES PORTILHO

**ANÁLISE *IN SILICO* DO SISTEMA UBIQUITINA-
PROTEASSOMA EM *Biomphalaria glabrata***

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Biotecnologia, Mestrado Acadêmico em Biotecnologia, com vistas à obtenção do título de Mestre em Biotecnologia.

Orientação: Dr. Matheus de Souza Gomes

PATOS DE MINAS – MG

FEVEREIRO DE 2019

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

P852a Portilho, Laysa Gomes, 1994-
2019 Análise in silico do sistema Ubiquitina-proteassoma em
 Biomphalaria glabrata [recurso eletrônico] / Laysa Gomes Portilho. -
 2019.

 Orientador: Matheus de Souza Gomes.
 Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Biotecnologia.
 Modo de acesso: Internet.
 Disponível em: <http://dx.doi.org/10.14393/ufu.di.2019.330>
 Inclui bibliografia.
 Inclui ilustrações.

 1. Biotecnologia. 2. Ubiquitinas. 3. Esquistossomose. 4. . I. Gomes,
Matheus de Souza, 1981- (Orient.) II. Universidade Federal de
Uberlândia. Programa de Pós-Graduação em Biotecnologia. III. Título.

CDU: 60

Maria Salete de Freitas Pinheiro - CRB6/1262

LAYSA GOMES PORTILHO

**ANÁLISE *IN SILICO* DO SISTEMA UBIQUITINA-
PROTEASSOMA EM *Biomphalaria glabrata***

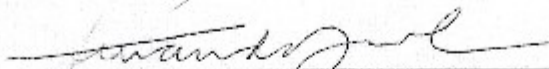
Dissertação de mestrado apresentada ao Programa
de Pós-Graduação em Biotecnologia, Mestrado
Acadêmico em Biotecnologia, com vistas à
obtenção do título de Mestre em Biotecnologia.

Aprovado em: 21 / 02 / 19

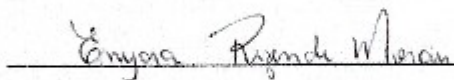
BANCA EXAMINADORA



Professor Dr. Mathheus de Souza Gomes



Professora Dra. Fernanda Janku Cabral



Professora Dra. Enyara Rezende Moraes

PATOS DE MINAS – MG

2019

Dedico esta dissertação aos meus pais, alicerces e exemplos da minha vida, Jaci e Enio, ao meu irmão, Lucas e a todos meus familiares, especialmente aos meus primos Sóter, Marlene, Emanuel, Sader e Débora, por me incentivarem incondicionalmente. Dedico também a todos que estiveram comigo e agradeço pelo apoio imensurável.

AGRADECIMENTOS

A Deus, pela graça da vida, pelas oportunidades que me são dadas todos os dias e pelo dom do altruísmo, fundamental para um cientista;

Aos meus pais, Jaci e Enio, minha base, meus exemplos e maiores incentivadores, por não medirem esforços para verem meus sonhos realizados, por me ensinarem todos os valores pelos quais luto e compartilho todos os dias. Por serem os meus maiores amores, meus espelhos de fé, humildade, perseverança e empatia. Espero dar-lhes orgulho sempre;

Ao meu irmão, Lucas, pelo companheirismo, união e amor incondicionais, pelo entendimento das ausências, pelo ombro amigo e as piadas nos dias difíceis. Obrigada por ser meu exemplo de sabedoria, pessoa, líder, profissional e amigo;

A todos os meus familiares, especialmente meus primos-irmãos Sóter, Marlene, Emanuel, Sader e Débora, por acreditarem na minha capacidade, pelo carinho e preocupação, pelo incentivo incondicional, pelas orações, pelo amor e amizade que construímos. À Aline e à Bárbara, por serem minha inspiração para ser sempre o meu melhor;

À Nina, minha fiel escudeira, maior e melhor fonte de renovação de energias e alegria que eu possa ter. Meu anjo de quatro patas;

A todos os meus amigos, sem exceção, por serem tão presentes, mesmo que alguns estejam distantes fisicamente, demonstrando tanto afeto, cuidado e amor, por me ensinarem muito sobre ser humano. Em especial agradeço as minhas amigas Laís, Larissa, Mali, Tacy, Fran, Gabi e Anninha que me auxiliaram muito nessa caminhada, pela credibilidade que por vezes nem eu mesma conseguia encontrar, pelo entendimento das faltas, pelo tempo investido para jogarmos conversas fora, por manterem minha sanidade mental em dia e por me compreenderem tão bem. Gratidão pelos anjos da guarda que são em minha vida;

Aos colegas de mestrado por compartilharem comigo das sensações difíceis de angústias, medos, frustrações, mas também das vitórias e conquistas. Por vezes foram as únicas pessoas que conseguiram me entender. Ao Diego, Paulinha e Cíntia meu agradecimento por estarem presentes em minha vida, por sempre possuírem um abraço amigo e um sorriso convidador para oferecer, por fazerem a diferença e por serem pessoas tão maravilhosas e iluminadas, um caminho brilhante espera por vocês. Obrigada por tudo;

À Débora e Hugo, que sempre encontram um espaço na casa e no coração para me acolherem em Belo Horizonte, pela amizade, pelos conselhos, por compartilharem comigo mais que uma república. Obrigada por todo carinho, cuidado e atenção de sempre. Minha família de BH. Vocês são muito especiais.

Ao Prof. Dr. Matheus de Souza Gomes, pelos seis anos de convívio, de aprendizado, de confiança e credibilidade, de crescimento profissional e pessoal. Obrigada pelo exemplo de ser humano, professor e pesquisador dedicado que tive o prazer de compartilhar conhecimentos e me espelhar;

Às professoras Dr.^a Enyara Rezende Moraes, Dr.^a Roberta Lima Caldeira, Dr.^a Fernanda Janku Cabral e ao professor Dr. Aulus Barbosa por todo o acolhimento, pela credibilidade, ensinamentos, empolgação e pelas conversas instigantes e esclarecedoras de sempre;

Ao pessoal do LBAM pela parceria, convivência, pelas reuniões, projetos e risadas. Deixo meu grande agradecimento à Bruna que foi fundamental para o desenvolvimento deste trabalho, pela presteza, amizade e companheirismo. O sucesso é intrínseco seu, minha menina;

Aos pesquisadores do Instituto de Pesquisas René Rachou por me receberem sempre tão bem, pela disponibilidade de todos os recursos para o meu crescimento como cientista, por compartilharem comigo de conhecimentos valiosos, desde frustrações a conquistas. Agradeço imensamente ao Ms. Fábio Ribeiro Queiroz, ao Dr. Wander Jesus Jeremias e ao Dr. Elio Ideo Babá pela paciência, pelo enorme aprendizado teórico e prático ganho nestes últimos dois anos. Vocês foram de fundamental importância para meu amadurecimento pessoal e profissional.

Ao pessoal dos laboratórios de helmintologia e malacologia médica e de estudos em esquistossomose, gratidão pela ciência ensinada. Às meninas Cíntia, Sandra, Fernanda, Izabela, Amanda, Jéssica, Carol, Renatinha e Malu, obrigada pela ajuda, pelas discussões sempre muito precisas e esclarecedoras, pelos ensinamentos, pela busca por soluções, pelo ombro amigo, por tudo. À Silvia pelo laço criado, pelo apoio e amor demonstrados, pelos auxílios em todos os momentos dentro e fora do laboratório, não tenho palavras que demonstrem minha gratidão;

Ao moluscário Lobato Paraense pela disponibilização dos moluscos, em especial à professora Liana Janotti, à Delza e Dilce, pela alegria e satisfação diárias;

Ao Aparecido, Carla e Milton, técnicos extremamente prestativos, que foram essenciais para a conclusão deste trabalho e desta fase da minha vida. Obrigada pelos conselhos, amizade, presteza e amor pela profissão. Vocês são exemplos.

Ao CNPQ e CAPES pelo financiamento desta pesquisa com o auxílio fornecido para o projeto vinculado sob o número 479890/2013-7 e a bolsa de estudos concedida.

E a todos aqueles que, direta ou indiretamente, participaram da realização deste trabalho e que torceram pelo meu sucesso.

*“Estamos aqui para fazer alguma diferença
no universo, se não, por que estar aqui?”*

Steve Jobs, 1955-2011

RESUMO

O sistema ubiquitina-proteassoma (SUP) é um dos principais reguladores gênicos, pois é capaz de realizar uma regulação pós-traducional através da marcação de substratos-alvo pela ubiquitina que, posteriormente, são direcionados para sinalizações celulares e/ou são degradados pelo proteassoma 26S. O SUP está relacionado a diversos mecanismos, incluindo respostas imunes e pode apresentar uma vasta atividade em moluscos do gênero *Biomphalaria*, principalmente relacionada a resistência/suscetibilidade do caramujo. O *B. glabrata*, altamente suscetível e principal responsável pela transmissão do *S. mansoni*. Muito da biologia do molusco já é conhecida, mas estudos envolvendo sua regulação gênica ainda são escassos. Dessa maneira, o objetivo deste trabalho foi prever, caracterizar e avaliar a expressão de genes envolvidos na via de ubiquitinação utilizando ferramentas de bioinformática em dados de genoma e transcrito de *B. glabrata*. Os supostos genes e proteínas do SUP em *B. glabrata* foram preditos utilizando o BLASTp usando proteínas de organismos modelos como *queries*. As prováveis proteínas foram caracterizadas usando o software PFAM e CDD descrevendo os domínios conservados e os sítios ativos. A análise filogenética foi realizada usando ClustalX2 e MEGA5.2. A avaliação da expressão foi realizada a partir de 12 tecidos de caramujo utilizando RPKM (*Reads Per Kilobase of transcript, per Million mapped reads*). Foram identificadas 86 sequências: uma de ubiquitina, 6 enzimas ativadoras de ubiquitina (E1), 22 enzimas conjugadoras de ubiquitina (E2), 39 enzimas ubiquitina ligases (E3) e 18 proteínas acessórias e adaptadoras. Os principais domínios conservados das proteínas identificadas foram: *Ubiquitin Family*, UQ_con, HECT, U-box e domínios homólogos a *RING finger*. Os sítios ativos encontrados foram os resíduos de lisina responsáveis pelas cadeias de poliubiquitinas na sequência de ubiquitina e resíduos de cisteína das enzimas da via de ubiquitinação que se ligam à ubiquitina, entre outros. A filogenia demonstrou uma distribuição evolutiva muito bem definida entre todos os grupos de sequências, corroborando com a árvore da vida. A sequência de ubiquitina foi a única com perfil de alta expressão encontrado para todas as bibliotecas, inferindo sua ampla gama de desempenho. Esses resultados evidenciam a presença, conservação e a expressão de SUP em *B. glabrata*, fornecendo base para novos conhecimentos envolvendo esta via no molusco em estudo, assim como oferece novos caminhos para o entendimento da regulação de genes que conferem a característica de suscetibilidade/resistência do caramujo, possibilitando maiores perspectivas para o controle da esquistossomose.

Palavras-chave: Modificação pós-traducional, caramujo, doença negligenciada, via de sinalização, análises *in silico*.

ABSTRACT

Schistosomiasis is a parasitic disease caused by helminths of the genus *Schistosoma* that affects more than 200 million people worldwide. Brazil is the South American country with the highest incidence of the disease, due, among other factors, to the vast territorial distribution of mollusks of the genus *Biomphalaria*. *B. glabrata*, highly susceptible and main responsible for *S. mansoni* transmission. Much of the biology of the mollusk is known, but studies involving its gene regulation are still scarce. The ubiquitin-proteasome system (USP) is one of the major gene regulators since it is capable of performing post-translational regulation by labeling target substrates by ubiquitin, which are subsequently targeted for cellular signaling and/or degraded by the 26S proteasome. This system is related to several mechanisms, including immune responses. Thus, the objective of this work was to predict, characterize and evaluate the expression of genes involved in the ubiquitination pathway using bioinformatics tools in genome and transcriptome data from *B. glabrata*. The alleged USP genes and proteins in *B. glabrata* were predicted using BLASTp using proteins from model organisms as queries. We characterize these putative proteins using the PFAM and CDD software describing the conserved domains and the active sites. Phylogenetic analysis was performed using ClustalX2 and MEGA5.2. Expression evaluation was performed from 12 snail tissues using RPKM. A total of 86 sequences were identified: one ubiquitin sequence, 6 ubiquitin activating enzymes (E1), 22 ubiquitin conjugating enzymes (E2), 39 ubiquitin ligases (E3) enzymes and 18 accessory and adapter proteins. The main conserved domains of the proteins identified were Ubiquitin Family, UQ_con, HECT, U-box and domains homologous to RING finger. The active sites found were the lysine residues responsible for the polyubiquitin chains, cysteine residues that bind to ubiquitin, among others. Phylogeny showed a very well defined evolutionary distribution between all groups of sequences, corroborating with the tree of life. The ubiquitin sequence was the only with high expression profile found for all libraries, inferring its wide range of performance. These results evidenced the presence, conservation and expression of USP data of *B. glabrata*, providing a basis for new knowledge involving this pathway in the study mollusk, as well as offering new ways of understanding the regulation of genes that confer the susceptibility/resistance of the snail, allowing greater perspectives for the control of schistosomiasis.

Key words: Post-translational modification, snail, neglected disease, signaling pathway, in silico analyzes.

SUMÁRIO

CAPÍTULO I.....	12
1 INTRODUÇÃO.....	12
2 REFERENCIAL TEÓRICO.....	15
2.1 Esquistossomose mansônica	15
2.2 Hospedeiro intermediário	18
2.3 Ubiquitina-proteassoma	21
2.4 Utilização das análises <i>in silico</i>	24
CAPÍTULO II.....	27
ARTIGO	27
CONCLUSÃO.....	53
REFERÊNCIAS	54
ANEXO	60
Anexo 1- Instruções para autores da revista “Memórias do Instituto Oswaldo Cruz”.....	60

CAPÍTULO I

1 INTRODUÇÃO

De todas as espécies conhecidas de *Schistosoma*, apenas o *S. mansoni* foi capaz de desenvolver seu ciclo evolutivo no Brasil, devido à presença do hospedeiro intermediário *Biomphalaria glabrata* presentes em praticamente todo o território nacional. Dessa forma, esses moluscos apresentam uma grande importância epidemiológica, pois auxiliam no processo de disseminação da esquistossomose.

A esquistossomose é uma parasitose ocasionada por helmintos do gênero *Schistosoma* capaz de atingir uma enorme população em países espalhados por todo o mundo, principalmente no continente africano. O Brasil é um dos países que apresenta grande incidência da doença, especificamente nas regiões nordeste e sudeste uma vez que se trata de áreas endêmicas para o caramujo hospedeiro intermediário principal da doença. O principal tratamento realizado é feito por meio da quimioterapia utilizando o praziquantel, um fármaco de baixo custo, o qual não é capaz de evitar reinfecções pelo parasito e nem atua sobre vermes jovens.

Ao ser infectado pelos miracídios de *S. mansoni*, o sistema interno de defesa dos moluscos desencadeia alterações no perfil de expressão proteica e consequentemente, no padrão de defesa inata dos mesmos, na busca de combater o parasito. O controle e/ou combate da infecção pelo caramujo pode estar relacionada à regulação e degradação de determinadas proteínas. Dentre os níveis de regulação proteica existentes uma das modificações pós-traducionais mais conhecidas e relevantes envolve o sistema ubiquitina-proteassoma que é composto pela via de ubiquitinação e o proteassoma 26S. A via de ubiquitinação marca substratos proteicos alvos que, posteriormente, participarão de outros processos celulares ou serão degradados pela molécula proteassomal. Assim, o sistema ubiquitina-proteassoma pode estar estritamente ligado à expressão da característica de suscetibilidade/resistência existente no caramujo.

Dessa maneira, visto a importância de *B. glabrata* como hospedeiro intermediário no ciclo da esquistossomose e o desafio em tratar e controlar a disseminação da doença, o problema enfatizado neste trabalho se refere à esquistossomose que atinge um grande número de pessoas, mas que é uma doença negligenciada. Além de demonstrar a importância do

caramujo como vetor da parasitose e do sistema ubiquitina-proteassoma como regulador gênico.

Considerando que até o momento esse sistema não foi descrito no molusco em estudo, a principal hipótese existente no presente trabalho é a de que tais moléculas, envolvidas na via de ubiquitinação, estejam presentes em dados de genoma e transcrito de *B. glabrata*. Além disso, outras hipóteses fomentadas são as de que elas possam ser caracterizadas quanto a seus domínios conservados, sítios ativos e filogenia frente a sequências homólogas de organismos referência utilizando ferramentas de bioinformática. A última, mas não menos importante, hipótese fomentada é a de que os prováveis genes pertencentes à via de ubiquitinação estejam sendo expressos em bibliotecas de transcrito de 12 tecidos diferentes que descrevem o organismo em condições normais de desenvolvimento (não infectados) de caramujos adultos.

Avaliando essas informações, o objetivo deste trabalho foi identificar e caracterizar proteínas envolvidas na via de ubiquitinação, assim como avaliar o papel e a expressão dos principais genes envolvidos na via de ubiquitinação em *B. glabrata*. Especificamente objetiva-se prever a existência das proteínas e enzimas da via de ubiquitinação através de análises *in silico*, além de analisar o perfil de expressão dos genes codificadores dessas moléculas em dados de transcrito de 12 tecidos diferentes de moluscos adultos, estando entre eles o sistema nervoso central, rim e ovotestis.

Nos últimos anos a interação entre *B. glabrata* e *S. mansoni* tem sido observada por alguns pesquisadores, recebendo importância e destaque para seu melhor entendimento. No entanto, essa relação precisa ser melhor explorada e elucidada, uma vez que ainda existem perguntas a serem respondidas, como por exemplo, o fato de algumas espécies do molusco apresentarem graus distintos de suscetibilidade ao miracídio. Assim, a interação entre parasito e hospedeiro se torna um intrigante objeto de estudo, podendo ser base para buscas de novas alternativas para controlar a disseminação da esquistossomose.

Dessa maneira, os conhecimentos sobre o sistema ubiquitina-proteassoma podem estar relacionados à interação existente entre parasito e hospedeiro intermediário, uma vez que estudos prévios já demonstraram um perfil de expressão diferente para as moléculas envolvidas nesse sistema presentes nos hemócitos de caramujos infectados e não infectados. Devido ao papel que o sistema ubiquitina-proteassoma desempenha nos organismos, sendo capaz de realizar uma regulação pós-traducional específica, pode estar relacionado com o padrão de expressão gênica de proteínas no caramujo, principalmente àquelas envolvidas com o seu sistema interno de defesa, responsável por combater o invasor.

As pesquisas realizadas até o momento não demonstraram a existência e funcionamento da maquinaria da via de ubiquitinação na sua totalidade no principal organismo para estudos de interação da esquistossomose mansoni, *B. glabrata*. Pelo papel que o sistema ubiquitina-proteassoma desempenha, sabe-se que este está diretamente relacionado a diversos processos de desenvolvimento animal e acredita-se que ele possa influenciar na característica de resistência/suscetibilidade do molusco. Sendo assim, a descrição dessas moléculas neste caramujo, assim como a avaliação de sua expressão em diferentes condições de infecção, pode fornecer respostas às indagações geradas referentes ao processo de envolvimento da via em estudo no controle da expressão de genes responsáveis por desenvolver a resistência nos moluscos no momento da infecção do hospedeiro intermediário por *S. mansoni*, abrindo perspectivas para novas alternativas de controle da doença.

2 REFERENCIAL TEÓRICO

Diversos estudos já foram realizados buscando o melhor entendimento da relação existente entre o hospedeiro intermediário e o parasito causador da esquistossomose. No entanto, este é o primeiro trabalho correlacionando a via de ubiquitinação, pertencente ao sistema ubiquitina-proteassoma, ao principal caramujo hospedeiro invertebrado da esquistossomose, *B. glabrata*.

2.1 Esquistossomose mansônica

Segundo a Organização Mundial da Saúde (OMS), a esquistossomose humana é uma doença causada por helmintos do gênero *Schistosoma* que atinge atualmente mais de 200 milhões de pessoas pelo mundo e, aproximadamente, 700 milhões vivem em áreas endêmicas e áreas de risco de transmissão da doença. Dentre os 78 países relatados com a ocorrência desta patologia estão àqueles pertencentes às regiões asiática, africana e americana, em que 52 deles são considerados regiões endêmicas para a parasitose (WHO, 2019). O Brasil é o país, pertencente à América do Sul, mais afetado pela esquistossomose. Em 2016, de acordo com o Ministério da Saúde, foram contabilizados cerca de 20 mil casos da doença, sendo encontrados, principalmente, nas regiões nordeste e sudeste do país, embora existam relatos de casos distribuídos por todo o território brasileiro (DA MARTINS et al., 2015; MINISTÉRIO DA SAÚDE, 2019). Além dos casos expostos, a OMS estima que 1,5 milhão de pessoas necessitem de um tratamento preventivo no país (MINISTÉRIO DA SAÚDE, 2019).

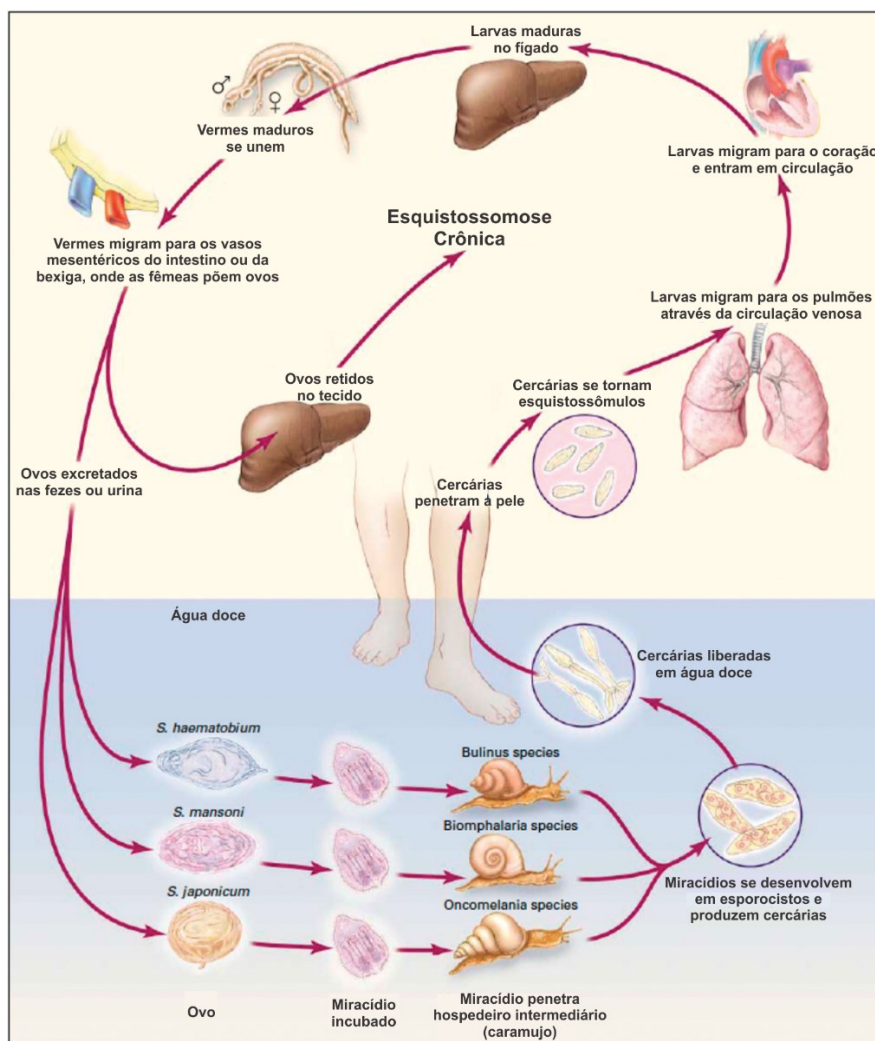
Seis espécies do gênero *Schistosoma* são descritas possuindo a capacidade de infectar seres humanos, sendo elas: *S. mansoni*, *S. japonicum*, *S. haematobium*, *S. makongi*, *S. intercalatum* e *S. malyensis*. Apenas *S. haematobium* atinge o sistema urogenital humano, já as demais espécies estão relacionadas com a esquistossomose intestinal. Acredita-se que essas espécies de helmintos tenham sido introduzidas no país através da imigração africana e asiática. Porém, apenas *S. mansoni* foi capaz de perpetuar a espécie e continuar seu ciclo de vida, uma vez que este encontrou moluscos aquáticos hospedeiros intermediários suscetíveis à infecção (COURA; AMARAL, 2004). Dessa maneira, a única espécie do gênero *Schistosoma* existente no Brasil e, conseqüentemente, responsável pela patologia conhecida como xistose ou barriga d'água é o *S. mansoni* (GRYSEELS et al., 2006; WHO, 2019).

A esquistossomose, apesar de ter sido descoberta por Theodore Bilharz em 1851, só foi descrita no Brasil em 1908 por Manuel Augusto Pirajá da Silva, identificando quatro casos de uma nova espécie de *Schistosoma* encontrada por Sambon (1907) e nomeada de *S. mansoni* (SILVA, 1908). O ciclo de vida desse helminto ocorre quando ovos de *S. mansoni* são excretados juntamente com as fezes de mamíferos infectados, próximos a locais com água doce disponível. Posteriormente à deposição de ovos, ocorre a eclosão de larvas natantes, chamadas miracídeos, capazes de infectar caramujos de água doce do gênero *Biomphalaria*. Esses moluscos são hospedeiros intermediários desse parasito e cruciais na disseminação da doença, uma vez que, no seu tecido, esses miracídeos passam por reprodução assexuada e sofrem alterações morfofisiológicas se tornando esporocistos. Através de estímulos luminosos e de temperatura esses hospedeiros intermediários liberam larvas capazes de infectar os seus hospedeiros definitivos, ou seja, mamíferos, especialmente os seres humanos (Figura 1) (GRYSEELS et al., 2006; CARVALHO et al., 2008; EL-ZAHAR e EL-KARIM, 2012; WHO, 2019).

Ao penetrarem pela pele do hospedeiro vertebrado, as cercárias novamente sofrem alterações morfológicas se tornando esquistossômulos, caracterizados pela ausência da cauda. Após a passagem pela epiderme, os esquistossômulos migram para as veias da derme, e ao encontrar a artéria pulmonar estes parasitos são levados até os pulmões. Dos pulmões, os esquistossômulos migram para a circulação venosa e seguem para o sistema porta-hepático, no qual ocorre à maturação sexual e o acasalamento dos parasitos machos e fêmeas completando desse modo o ciclo de vida do parasito. A maioria dos ovos depositados pela fêmea adulta fica retida no tecido, porém, os ovos que são excretados nas fezes do hospedeiro definitivo marcam o reinício do ciclo. Dessa maneira, a esquistossomose é considerada uma doença crônico-debilitante e um grave problema de saúde pública (GRYSEELS et al., 2006; CARVALHO et al., 2008; EL-ZAHAR e EL-KARIM, 2012; WHO, 2019).

Pessoas infectadas com *S. mansoni* apresentam sintomas que estão diretamente relacionados à deposição de ovos pelas fêmeas, sendo que cerca de 300 ovos/dia/fêmea são produzidos. Juntamente com a deposição destes, alguns antígenos solúveis são liberados no tecido hepático e ocasionam uma grande reação inflamatória, levando a um processo de fibrose. Além dessa característica, outros sinais e sintomas estão envolvidos com a patologia como diarreia, dores abdominais e sangue nas fezes. Entretanto, esplenomegalia e ascite surgem como sintomas na forma crônica da esquistossomose (CALDAS et al., 2008; WHO, 2019).

Figura 1 – Ciclo evolutivo da esquistossomose.



Fonte: Adaptado de EL-ZAHAR e EL-KARIM, 2012.

Para que a doença seja diagnosticada, a OMS indica que seja utilizado o método Kato-Katz em que é realizada uma busca de ovos depositados nas fezes do hospedeiro vertebrado (KATZ et al., 1972; WHO, 2019), porém, essa é uma ferramenta não muito efetiva em indivíduos com baixa carga parasitária. Dessa forma, técnicas com maior sensibilidade são necessárias para diagnosticar a esquistossomose. Uma delas é a Reação em Cadeia da Polimerase (PCR) em tempo real, que se apresenta como uma excelente alternativa devido a sua grande precisão. Contudo, é uma técnica que possui certa dificuldade ao ser aplicada em campo, além do alto custo relacionado e a necessidade de um profissional capacitado para interpretar os resultados (ESPÍRITO-SANTO et al., 2014).

O tratamento realizado atualmente para a esquistossomose é por meio da quimioterapia utilizando o praziquantel, um fármaco ou droga com baixo custo. No entanto, sua eficácia é colocada à prova, uma vez que não atua sob vermes jovens e é incapaz de evitar

reinfeções pelo parasito pelo fato de não ser uma vacina. Além disso, ele tem sido largamente utilizado a mais de 40 anos, podendo ocasionar uma seleção do helminto ao fármaco (BERGQUIST et al., 2017; WHO, 2019).

O fornecimento de saneamento básico e tratamento de água adequado à população são medidas necessárias para evitar a disseminação da doença. Já o controle dos caramujos é realizado a partir do uso de moluscidas que também não se mostram completamente eficazes e são utilizados como último recurso (SILVA et al., 2008; WHO, 2019).

2.2 Hospedeiro intermediário

Moluscos do gênero *Biomphalaria* apresentam uma grande importância epidemiológica, uma vez que a sua presença compõe uma condição crucial e indispensável para a esquistossomose da esquistossomose mansônica, causada pelo *S. mansoni*. Assim, o hospedeiro intermediário é infectado pelo parasito na sua forma de miracídio e suas fases assexuadas se desenvolvem nos tecidos do molusco (BRASIL, 2008).

As espécies pertencentes a esse gênero pertencem ao filo Mollusca, classe Gastropoda, subclasse Pulmonata, ordem Basommatophora e família Planorbidae possuindo características peculiares como, por exemplo, a concha espiral plana com um aprofundamento no giro central e lados aproximadamente paralelos. Essa particularidade também indica o nome, em latim, dado ao gênero *Biomphalaria*, em que *bis* indica a expressão “duas vezes”, referindo-se aos dois lados praticamente paralelos da concha e *omphalos* que significa umbigo, mencionando o aprofundamento na concha (PARAENSE, 1975a).

Pertencentes a esse gênero, 27 espécies e subespécies já foram descritas, existindo a possibilidade de algumas delas estarem em sinonímia. Dentre as localidades em que foram encontradas estão a América do Sul, América Central, sul dos Estados Unidos da América, África e Caribe (SILVA, 2012). Entretanto, de todas as espécies descritas, apenas três apresentaram a característica de serem infectadas naturalmente por *S. mansoni*, *Biomphalaria glabrata*, *tenagophila* e *straminea* (PARAENSE; CORRÊA, 1978; SOUZA, 1986).

B. glabrata (Figura 2) é o mais relevante molusco hospedeiro intermediário desse parasito nas Américas, tanto pela ampla distribuição geográfica em que se encontra como pelo alto grau de compatibilidade parasito-hospedeiro que apresenta, ou seja, uma grande suscetibilidade à infecção. Assim, essa espécie contribui para que um maior número de cercárias sejam liberadas na água doce e que, conseqüentemente, possam atingir um maior

número de hospedeiros definitivos ocasionando uma grande disseminação da doença e uma continuidade do ciclo evolutivo do parasito (DE SOUZA et al., 1995; PARAENSE, 2008).

Figura 2 – *Biomphalaria glabrata*, caramujo hospedeiro intermediário de *S. mansoni*.



Fonte: Keila Maia, CPqRR.

Por ser a espécie com maior grau de suscetibilidade frente ao *S. mansoni*, o *B. glabrata* se tornou a principal espécie vetora para estudos que envolvem a interação trematódeo-molusco. Devido à grande importância que esse organismo demonstra no seguimento da doença, trabalhos envolvendo a biologia molecular do molusco foram financiados e realizados buscando evidenciar a relação existente entre genoma, transcrissoma e suas características fenotípicas, como a resistência/suscetibilidade. Além de oferecer diferentes perspectivas para o entendimento da regulação gênica do molusco (ADEMA et al., 2006; LOCKYER et al., 2008; ADEMA et al., 2017; QUEIROZ et al., 2017).

A interação entre molusco e parasito é complexa e foi demonstrada em um trabalho realizado por Richards e Shade (1987) que demonstraram a grande diversidade de genes que indicavam a suscetibilidade dos hospedeiros intermediários e a infectividade dos parasitas, sugerindo a geração de variabilidades genéticas adaptativas em ambos os organismos. Além destas variabilidades, demonstradas em linhagens de caramujos completamente suscetíveis e outras parcialmente resistentes à infecção pelo trematódeo, os hemócitos, células do sistema interno de defesa do caramujo, também apresentaram grande influência para que a infecção ocorresse efetivamente ou para que ela seja evitada (LOKER e BAYNE, 2001).

O sucesso da infecção envolve muitos fatores sendo que por um caminho, o sistema interno de defesa dos caramujos é o responsável por reconhecer e destruir o parasito, mas por outro lado, o verme apresenta a característica de evasão e autodefesa contra o ataque do sistema de defesa do caramujo envolvendo habilidades como mimetizar moléculas do hospedeiro, fazendo com que o sistema interno de defesa do molusco não seja capaz de reconhecê-lo (PLOWS et al., 2005; LEHR et al., 2008; BAYNE, 2009; NEGRAO-CORREA et al., 2012). Ainda assim, várias reações e alterações no perfil de expressão gênica do caramujo são desencadeadas quando o trematódeo é reconhecido, principalmente envolvendo os hemócitos, principais células em suspensão que atuam no sistema de defesa desses animais contra organismos invasores (QUEIROZ, 2015)

Além dos hemócitos, o sistema interno de defesa do caramujo compartilha de componentes humorais da hemolinfa, responsáveis pelo desencadeamento da resposta imunológica nos moluscos no momento em que a infecção acontece. Assim, provoca alterações no perfil de expressão de algumas proteínas e consequentemente, no padrão de defesa que o molusco apresenta. Algumas destas proteínas estão relacionadas a espécies reativas de oxigênio e nitrogênio (LOCKYER et al., 2008; LOKER e BAYNE, 2001; NEGRAO-CORREA et al., 2012; LARSON et al., 2014). A transferência da hemolinfa, retirada de caramujos resistentes e implantada em caramujos suscetíveis, é capaz de conferir um aumento da resistência naqueles que são mais suscetíveis à infecção por *S. mansoni* (BAYNE et al., 1980a; b; GRANATH; YOSHINO, 1984; VAN DER KNAAP; LOKER, 1990; BAYNE, 2009).

Além dessas moléculas, alguns trabalhos avaliaram a expressão gênica de hemócitos em caramujos infectados e não infectados, suscetíveis e resistentes, e foram capazes de mostrar um perfil diferencial de outros componentes, como sinalizadores celulares e genes envolvidos na via de ubiquitinação (LOCKYER et al., 2008; ZAHOOR et al., 2014). Indicando que existe uma modificação do padrão de regulação gênica quando o caramujo é infectado, assim como quando são comparados moluscos com maior e menor grau de resistência. Queiroz e colaboradores (2017) demonstram uma via envolvida na regulação gênica do molusco, uma vez que foram capazes de descrever a via de biossíntese de pequenos RNAs e avaliar a expressão das principais enzimas e proteínas relacionadas a essa via. Os pequenos RNAs são responsáveis, juntamente com as proteínas argonauta, por realizar uma regulação de genes bastante específica, devido a sua capacidade de se ligar complementarmente à fita de RNA mensageiro. Outros sistemas envolvidos na regulação gênica são bastante conhecidos e bem descritos. Um dos mais importantes é o sistema

ubiquitina-proteassoma, responsável por executar uma regulação pós-traducional e largamente conservado em organismos eucariotos (TANAKA; CHIBA, 1998; GLICKMAN; CIECHANOVER, 2002;). Sendo assim, é importante avaliar os sistemas reguladores da expressão gênica em organismos vetores, principalmente no molusco usado como modelo em estudos para a esquistossomose envolvendo a interação parasito-hospedeiro, uma vez que a alteração no perfil proteico de hemócitos pode estar relacionado aos sistemas de regulação gênica.

2.3 Ubiquitina-proteassoma

O sistema ubiquitina-proteassoma (SUP) é um complexo ATP-dependente que envolve a marcação de substratos proteicos alvos, ou seja, realiza uma regulação pós-traducional através da marcação dos mesmos por moléculas de ubiquitina que podem sofrer uma posterior degradação intracelular pelo proteassoma (HERSHKO; CIECHANOVER, 1998; TANAKA; CHIBA, 1998; GLICKMAN; CIECHANOVER, 2002). Este sistema está relacionado a diversas vias e mecanismos celulares distintos, uma vez que é capaz de degradar e inativar proteínas mutadas e defeituosas que estejam ligadas a processos celulares importantes, como a regulação do ciclo celular, a resposta ao estresse e moduladores extracelulares, reparo do DNA e a regulação de respostas imunes e inflamatórias quando envolve o denominado imunoproteassoma, por exemplo (CIECHANOVER, 1998; RAMOS, 2002; LI et al., 2016; EBNER et al., 2017).

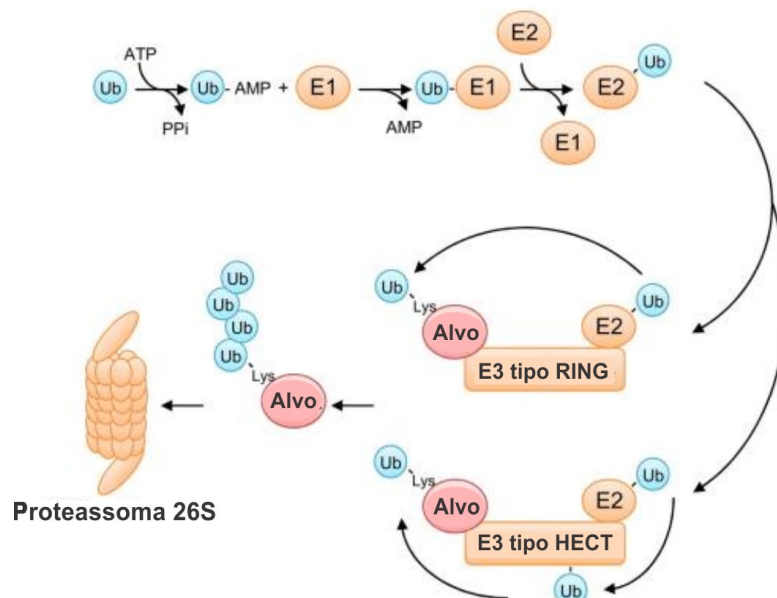
Dois elementos principais compõem o sistema, 1) a Via de Ubiquitinação e 2) o Proteassoma proteolítico 26S. Primeiramente, a partir da via de ubiquitinação, as proteínas alvo são marcadas com uma ou mais moléculas de ubiquitina e depois são direcionadas para outras vias de sinalização ou degradadas pela ação do proteassoma 26S, composto por duas estruturas: o núcleo proteolítico 20S e regulatório 19S, parte de reconhecimento dos substratos-alvo (HERSHKO; CIECHANOVER, 1998; GLICKMAN; CIECHANOVER, 2002).

A ubiquitina é um polipeptídeo com 76 aminoácidos, altamente conservado entre os organismos eucariotos e que possui a capacidade de se ligar, através da sua porção C-terminal, a resíduos de lisina ou à extremidade N-terminal de substratos proteicos, se tornando o componente principal da via de ubiquitinação (PICKART; EDDINS, 2004; LI; YE, 2008; ZUIN et al., 2014). Para que via de ubiquitinação ocorra de forma correta é necessária a

atuação de três classes enzimáticas diferentes e, além disso, a presença de moléculas de ATP (HERSHKO; CIECHANOVER, 1998; GLICKMAN; CIECHANOVER, 2002).

A primeira das três classes enzimáticas envolvidas na via de ubiquitinação são as enzimas ativadoras de ubiquitina, também conhecidas como E1, que de forma ATP-dependente ativam a molécula de ubiquitina e geram um composto intermediário de alta energia, o E1-tioestér de ubiquitina. Subsequente à ação das E1, as enzimas conjugadoras de ubiquitina E2 realizam a função de catalisar a reação de ligação covalente da ubiquitina marcada ao substrato proteico alvo. Ao final dessa cascata, as enzimas ubiquitina ligases E3 reconhecem a proteína específica a ser degradada e auxiliam na transferência da ubiquitina de E2 a tal substrato, podendo se ligar tanto a E2 quanto à molécula de ubiquitina marcada provinda dessa enzima conjugadora (SCHEFFNER et al., 1995; GLICKMAN; CIECHANOVER, 2002; ZHENG; SHABEK, 2017). A via de ubiquitinação é demonstrada na Figura 3.

Figura 3 – Esquema da via de ubiquitinação. As siglas utilizadas demonstram as moléculas envolvidas na via em que Ub: ubiquitina, E1: enzimas ativadoras de ubiquitina, E2: enzimas conjadoras de ubiquitina e E3: enzimas ubiquitina ligases.



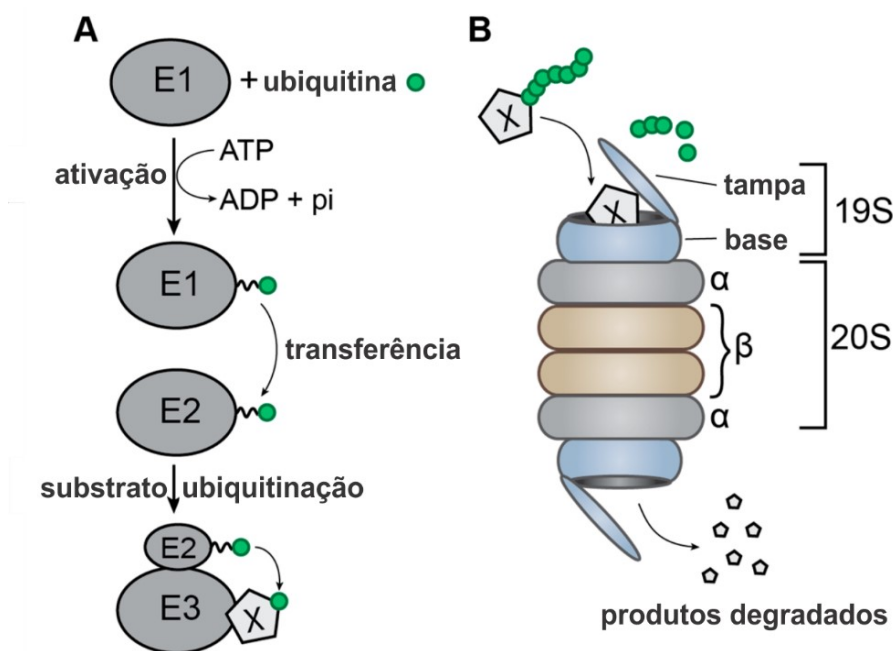
Fonte: Adaptado de MAGORI; CITOVSKY, 2011.

As E3 são consideradas peças chave na via de ubiquitinação, uma vez que apresentam a função de reconhecer o substrato alvo específico e catalisar a transferência da molécula de ubiquitina a ele. Dentro desta classe existem subdivisões que estão relacionadas, principalmente, ao sítio catalítico da enzima. Quatro dessas classes já foram descritas, sendo

elas a HECT Ub ligase, E3 do tipo U-box, RING *finger* de ligação simples e RING *finger* de múltiplas subunidades, sendo a última considerada a maior delas (LI; YE, 2008; ZHENG; SHABEK, 2017). Na classe RING *finger*, existem proteínas adaptadoras e acessórias, que auxiliam no processo de reconhecimento da E3 à proteína-alvo. A presença das proteínas Cullin são cruciais para o desempenho da função da E3, uma vez que apresentam a capacidade de posicionar o substrato proteico alvo próximo à enzima E2 facilitando a transferência da molécula de ubiquitina ativada (SARIKAS et al., 2011).

Posterior à marcação das proteínas alvo e a formação da cauda de poliubiquitina realizada através da cascata enzimática, esses substratos podem ser encaminhados para outras vias de sinalização celular, assim como podem ser degradados pelo complexo proteolítico do Proteassoma 26S que a reconhece. Esta protease é formada a partir de uma associação entre uma partícula regulatória 19S, também conhecida como PA700 e um componente central denominado 20S. A partícula 19S é dividida em dois subcomplexos, chamados de base e tampa e estão ligados de maneira reversível e ATP-dependente ao 20S (Figura 4) (ARENDT; HOCHSTRASSER, 1997; TANAKA; CHIBA, 1998; GLICKMAN; CIECHANOVER, 2002).

Figura 4 – Sistema ubiquitina-proteassoma. A) Esquema da via de ubiquitinação e atuação das enzimas envolvidas. B) Constituição molecular e demonstração da atuação do proteassoma 26S.



Fonte: Adaptado de SEISSLER et al., 2017.

A partícula regulatória que compõe o proteassoma 26S apresenta uma função que envolve reconhecer, desdobrar e translocar o substrato até a partícula central (BUDENHOLZER et al., 2017). Já o componente central 20S possui sítios catalíticos que desempenham o papel de degradar a proteína alvo, pois apresentam funções semelhantes à caspase, quimotripsina e tripsina (SALOMONS et al., 2010). O núcleo 20S é composto por, aproximadamente, 70 subunidades, sendo divididas entre α e β . As subunidades do tipo α possuem uma função voltada para a estrutura e regulação do núcleo proteassomal, enquanto as subunidades do tipo β desempenham o papel proteolítico propriamente dito, com atividade catalítica de degradação (JUNG et al., 2009; PAPAEOGENIOU e CHONDROGIANNI, 2014). Em *Caenorhabditis elegans*, um dos principais organismos eucariotos considerado modelo, 14 sequências correspondentes a essas subunidades foram identificadas em um trabalho realizado por Papaevgeniou e Chondrogianni (2014).

Ao final do SUP, as proteínas alvo que foram inicialmente marcadas pela ubiquitina e que não foram encaminhadas para outros processos celulares são dirigidas ao proteassoma 26S e então degradadas em peptídeos menores. Já as moléculas de ubiquitina presentes na cauda formada na ligação com o substrato alvo são liberadas, podendo ser recicladas pela célula em um novo processo (GLICKMAN; CIECHANOVER, 2002).

2.4 Utilização das análises *in silico*

A análise *in silico* e utilização da bioinformática tem se tornado ferramentas importantes e largamente utilizadas na descrição de moléculas e vias de sinalização em diversos organismos, sejam eles eucariotos ou procariotos. O uso dessa análise possibilita o maior entendimento, clareza e funcionamento de vias metabólicas, assim como a identificação de genes ainda não descritos em dados de genomas e transcrito (FÖRSTER et al., 2002). Adema e colaboradores (2017) publicaram o genoma de *B. glabrata* tendo em vista a importância que o molusco apresenta como vetor na transmissão da esquistossomose, a partir desses dados, foi possível descrever e identificar algumas sequências encontradas nesses dados, como sequências de miRNAs maduros e precursores de miRNAs, enzimas e proteínas envolvidas na via de pequenos RNAs, entre outros componentes importantes para o ciclo de vida do molusco (ADEMA et al., 2017).

O proteassoma e as moléculas envolvidas na via de ubiquitinação foram descritas em alguns trabalhos envolvendo as análises *in silico* e desenhos de estruturas tridimensionais. No trabalho relacionado à molécula proteassomal, são comparados dados encontrados em

bactérias como *Thermoplasma acidophilum* e *Escherichia coli* frente aos dados obtidos a partir de organismos eucarióticos como as leveduras, sendo possível verificar a estrutura tridimensional do complexo e como é realizada a sua montagem em cada um desses clados (BUDENHOLZER et al., 2017). Para as enzimas e proteínas relacionadas à via de ubiquitinação, suas estruturas são descritas e esquematizadas, possibilitando a inferência das funções desempenhadas por elas e como ocorre sua regulação, assim como a identificação de novos sítios de ubiquitinação proteica (RADIVOJAC et al., 2010; ZHENG; SHABEK, 2017). Proctor e colaboradores (2007) foram capazes de desenvolver um modelo para o SUP utilizando bioinformática confrontando com dados experimentais, além de permitir que modelos relacionados possam ser incrementados ao longo de novas descobertas. Isso permite uma maior compreensão da homeostase proteica e de que maneira o sistema ubiquitina-proteassoma pode estar envolvido. Estes trabalhos fornecem dados importantes para a identificação e caracterização de prováveis sequências proteicas e enzimáticas a serem descobertas e descritas. Algumas delas abrangem características estruturais conservadas das sequências funcionais pertencentes e envolvidas no SUP, como a posição de resíduos de aminoácidos, a conservação destes aminoácidos, a presença de domínios conservados, sítios ativos, motivos estruturais, entre outros. No entanto, até o momento nenhum estudo foi realizado demonstrando, caracterizando e analisando filogeneticamente a existência desse sistema tão importante no caramujo *B. glabrata*.

Além disso, no principal parasito de *B. glabrata*, *S. mansoni*, algumas moléculas envolvidas no SUP foram identificadas, analisadas, descritas e caracterizadas a partir de ferramentas de bioinformática e análises experimentais, como as deubiquitinases e as proteases ubiquitina-específicas, pertencentes à via de ubiquitinação (PEREIRA et al., 2015a, b). Envolvendo a avaliação do SUP em diferentes estágios de vida do trematódeo, Mathieson e colaboradores (2011) estudaram os estágios de evolução do ovo do helminto demonstrando que o sistema é capaz de mudar o seu perfil de expressão com o seu desenvolvimento. Já Castro-Borges e colaboradores (2007) trabalharam visando as mudanças ocasionadas quando as cercárias se transformam em esquistossômulos e observaram que ocorrem modificações na estrutura do proteassoma no momento que estão diretamente ligadas a remodelação morfofisiológicas do parasito. Por outro lado, Nabhan e colaboradores (2007), através de análises de bioinformática e o desenvolvimento de RNAs de interferência, buscaram entender com maior clareza a relevância biológica do proteassoma no esquistossomo e perceberam a presença de uma subunidade essencial para o desenvolvimento e sobrevivência do trematódeo.

Muitos estudos abrangendo a relação existente entre o parasito e o sistema ubiquitina-proteassoma, como demonstrado, vêm sendo realizados. No entanto, quando se relaciona o SUP ao hospedeiro invertebrado de *S. mansoni*, este é o primeiro trabalho realizado a fim de identificar e caracterizar esse sistema como um todo no caramujo, assim como relacioná-lo com o desenvolvimento do organismo, oferecendo conhecimentos que levem a novos entendimentos, assim como possibilitar o entendimento da maquinaria envolvida na regulação gênica pós-traducional no molusco, podendo controlar características fenotípicas do mesmo, como a resistência/suscetibilidade. Além disso, ainda pode proporcionar base para novas buscas quanto à relação parasito/hospedeiro intermediário, uma vez que o SUP está diretamente ligado à regulação gênica pós-traducional.

Considerando a regulação gênica como um caminho a ser seguido no melhor entendimento da relação tremaódeo/molusco e na característica de resistência do caramujo, um trabalho envolvendo o sequenciamento do genoma de *B. glabrata* foi publicado, abrindo possibilidade para novas pesquisas que poderão utilizar os dados obtidos por essa tecnologia (ADEMA et al., 2017).

Dessa maneira, a junção de técnicas experimentais de biologia molecular e ferramentas de bioinformática oferecem grandes vantagens e novas perspectivas às pesquisas relacionadas à descrição e caracterização de vias metabólicas e de sinalização nos organismos vivos. Esse aglomerado de conhecimentos pode oferecer maiores evidências da presença, ativação, funcionalidade e interação do sistema ubiquitina-proteassoma no molusco em estudo. Além do mais, pode fornecer uma base de novas descobertas e dados que podem ser usados a fim de entender melhor o modo com que essa via pode estar envolvida na regulação de genes que conferem a característica fenotípica de resistência/suscetibilidade do caramujo frente à infecção pelos miracídeos de seu principal parasito, *S. mansoni*.

CAPÍTULO II

ARTIGO

Genome-wide identification, characterization and expression profiling of the ubiquitin-proteasome genes in *Biomphalaria glabrata*

Laysa Gomes Portilho¹, Bruna Custódio Dias Duarte¹, Fábio Ribeiro Queiroz², Thales Henrique Cherubino Ribeiro³, Wander Jesus Jeremias⁴, Elio Hideo Babá², Paulo Marcos Zech Coelho², Enyara Rezende Moraes⁵, Roberta Lima Caldeira⁶, Matheus de Souza Gomes¹

1- Laboratório de Bioinformática e Análises Moleculares, Universidade Federal de Uberlândia Campus Patos de Minas, Patos de Minas, Minas Gerais, Brasil.

2- Grupo de Pesquisa em Biologia do *Schistosoma mansoni* e sua Interação com o Hospedeiro, Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Minas Gerais, Brasil.

3- Departamento de Biologia, Seção de Fisiologia de Plantas, Laboratório de Fisiologia Molecular de Plantas (LFMP), Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil.

4- UniBH - Centro Universitário de Belo Horizonte, Belo Horizonte, Minas Gerais, Brasil.

5- Laboratório de Bioquímica e Biologia Molecular, Universidade Federal de Uberlândia Campus Patos de Minas, Patos de Minas, Minas Gerais, Brasil.

6- Grupo de Pesquisa em Helminologia e Malacologia Médica, Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Minas Gerais, Brasil.

* Corresponding author

Corresponding Author: Matheus de Souza Gomes, Federal University of Uberlandia, Institute of Genetics and Biochemistry, Laboratory of Bioinformatics and Molecular Analysis, Getúlio

Vargas Avenue, 230, Patos de Minas, MG, Brazil. Phone: +55 34 38233714 –
matheusgomes@ufu.br.

Abstract

BACKGROUND: *Biomphalaria glabrata* is the major species for study of schistosomiasis-related parasite-host relationship and its gene regulation may aid in this understanding. The ubiquitin-proteasome system (UPS) performs a post-translational relationship and is related to several mechanisms, including immune responses. **OBJECTIVE:** The aims of this work were to identify and characterize the putative genes involved in UPS using bioinformatic tools and also their expression on different tissues of *B. glabrata*. **METHODS:** The putative genes and proteins of UPS in *B. glabrata* were predicted using BLASTp and as queries reference proteins from model organism. We characterized these putative proteins using the PFAM and CDD software describing the conserved domains and active sites. The phylogenetic analysis was performed using ClustalX2 and MEGA5.2. Expression evaluation was performed from 12 snail tissue using RPKM. **FINDINGS:** 119 putative genes involved in the UPS in *B. glabrata* were identified, which 86 have been related to the ubiquitination pathway and 33 to proteasome. In addition, the conserved domains found were Ubiquitin Family, UQ_con, HECT, U-box and Proteasome. Evidence of site formation, catalytic cleft and structural motifs infer the enzymatic function of genes. The active sites found were the lysine residues responsible for polyubiquitin chains, cysteine residues that bind to ubiquitin, among others. The phylogeny shows a distribution that corroborates with that demonstrated in the tree of life of the animals for all groups of sequences analyzed. The ubiquitin sequence was the only one with high expression profile found for all libraries inferring its wide range of performance. **MAIN CONCLUSIONS:** Our results evidenced the presence, conservation and expression profile of the UPS in the mollusk, providing basis and new knowledge to other studies

involving this system. Due to the importance of UPS and *B. glabrata*, this work may influence the search for new methodologies for the control of schistosomiasis.

Key words: snail – UPS – signaling pathway – bioinformatics – schistosomiasis.

Sponsorships

This work was supported by CNPq, National Research Council, through project number 479890/2013-7.

Introduction

Biomphalaria glabrata is a specie of snails that presents large epidemiological importance, since their presence composes a crucial condition for the dissemination of schistosomiasis, reaching more than 200 million people around the world. These mollusk are considered intermediate hosts of *Schistosoma mansoni* and present a high degree of susceptibility to helminth infection. In addition, they have a wide geographical distribution in the Americas⁽¹⁻²⁾. Due to the importance of this organism, many studies seek new knowledge about its biology and it is already known that the interaction between mollusk and trematode is complex, involving genes expressing host susceptibility/resistance and parasite infectivity⁽³⁾. Thus, one of the ways to better understand this relationship is to deepen the studies related to the gene regulation of *B. glabrata*, because the infection carried out by *S. mansoni* leads to changes in the expression profile of some proteins and, consequently, in the defense pattern of the mollusk⁽⁴⁻⁵⁾. Thus, it is necessary to evaluate the regulatory systems of gene expression in these organisms, which one of the main components is the ubiquitin-proteasome system that performs a post-translational regulation.

Given the epidemiological importance of *B. glabrata*, Adema and contributors (2017), including the present working group, published a complete analysis of the snail genome. Genes involved in communication with the aquatic environment, stress, innate immunity and regulation of biological processes were identified and described, as well as several small RNAs related to gene regulation. In addition, transcripts involved in the ubiquitin-proteasome system were found, from genes encoding enzymes to accessory proteins and the subunits that form the proteolytic 26S proteasome⁽⁶⁾.

The ubiquitin-proteasome system (UPS) is composed of two main elements: the ubiquitination pathway and the proteolytic macromolecule 26S proteasome. First, the ubiquitination pathway labels proteins with more than one ubiquitin molecule that are then degraded by the 26S proteasome⁽⁷⁾. This system is able to degrade mutated and defective proteins that are involved with important cellular processes, such as cell cycle regulation, stress response and extracellular modulators, DNA repair, and the regulation of immune and inflammatory responses⁽⁸⁻⁹⁾.

The first class acting on the ubiquitination pathway are ubiquitin-activating enzymes (E1) that activate the ubiquitin molecule in an ATP-dependent manner and generate an ubiquitin E1-thioester. Subsequently, ubiquitin conjugating enzymes (E2) catalyze the covalent binding reaction of activated ubiquitin with target protein substrate. At the end of this cascade, the enzymes ubiquitin ligases (E3) recognize the specific protein to be degraded and assist in the transfer of the ubiquitin present in E2 to the substrate, being able to bind both E2-ubiquitin and target substrate concomitantly or at different times^(7, 10-11). At the end of the first step, the labeled substrates are degraded by the proteolytic complex 26S proteasome. This protease is formed from an association between a 19S (PA700) regulatory particle divided between lid and base which are reversibly connected and ATP-dependent to the central component 20S^(7, 12-13). The regulatory particle recognizes, unfolds and translocates the

substrate to the central particle⁽¹⁴⁾. However, the central component 20S has proteolytic sites that play the role of degrading the target protein, since they have similar functions to caspase-like, chymotrypsin-like and trypsin-like⁽¹⁵⁾. Soon after, the target proteins are degraded into smaller peptides and the ubiquitin molecules present in the tail are released and recycled⁽⁷⁾.

The peptides generated from the digestion performed by the proteasome are related to immunity. In humans, these peptides are recognized as epitopes by MHC class I⁽¹⁶⁾ and, thus, the proteasome plays an essential role in this recognition⁽¹⁷⁾. However, the immunoproteasome is more efficient for the generation of antigenic peptides due to the presence of three different subunits than those in the 20S nucleus of the standard proteasome. These specific subunits modify peptidase activity and increase the generation of these epitopes⁽¹⁸⁾. In addition, the thymoproteasoma is another isoform specifically expressed in the thymus and both are involved with cell-mediated immunity⁽¹⁹⁻²⁰⁾. Thus, the relationship involving the catalytic proteasome and immunity leads to believe that it is also related to the resistance/susceptibility that some organisms present when they are infected by pathogens.

Genes involved in the ubiquitination pathway and in the proteasome formation have already been reported in works involving the cells of the internal defense system, the hemocytes. These genes demonstrated a different expression profile when analyzed in infected and uninfected snails⁽²¹⁻²³⁾. Thus, the hypothesis of this work was that UPS has been conserved in the genome of *B. glabrata* and expressed at transcriptional level in all parts of the snail body. Thus, the aims of this work were to identify and characterize UPS in the genome and transcriptome data of the snail using *in silico* analyzes, as well as to characterize the classes involved in this system and to evaluate the expression profile of the sequences identified in different tissues of the adult snail.

Materials and Methods

Identification of UPS genes in *B. glabrata*

The genome and transcriptome data used for analysis *in silico* were retrieved from the last version of VectorBase database, Bglab1 do genome de *B. glabrata* (<https://www.vectorbase.org/organisms/biomphalaria-glabrata>). Initially, the prediction of genes involved in ubiquitin-proteasome pathway was performed through bibliographic searches and the KEGG (<https://www.genome.jp/kegg/pathway.html>) in model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*. The sequences belonging to the model organisms were obtained from the NCBI database (<https://www.ncbi.nlm.nih.gov/>) and later used as a queries against the genome and transcriptome of the snail seeking to find UPS genes. Subsequently, the sequences identified in *B. glabrata* data were grouped according to the role they play and also analyzed for their conserved domains, active sites and gene expression.

Sequences analysis

Sequences were fed into the protein family database, PFAM (<https://pfam.xfam.org/>), to identify the major conserved characteristic domains belonging to each previously organized sequence group. The CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) was used to search for amino acid residues involved in active site formations, structural motifs and catalytic cleft.

Phylogenetic analysis

To determine the evolutionary organization and distribution of the sequences found, as well as to provide further evidence of the presence of probable proteins found in the snail, phylogenetic analysis was performed using MEGA5.2. For the analysis, amino acid sequences

of deuterostomes and protostomes organisms such as *Homo sapiens*, *Mus musculus*, *D. melanogaster*, *C. elegans* and molluscs such as *Aplysia californica* and *Lottia gigantea*, obtained from NCBI were used. ClustalX 2.1 was used to perform multiple sequence alignment and their output fed in the MEGA 5.2 software, using the neighbor-joining method and JTT model for calculations of evolutionary distance. The consensus tree was inferred from a bootstrap of 1000 replicates to represent the evolutionary history of the study⁽²⁴⁾.

Expression analysis

For expression analysis, all transcripts identified in the *B. glabrata* genome data were individually submitted and analyzed against a RNASeq data set for a library of 12 different snail tissues (Table S1)⁽⁶⁾. Quality control and adapter removal were conducted using Trimmomatic (v 0.36). Single-end reads were aligned against pre-selected sequences using bowtie2 (v 2.3.0) with the “--very-sensitive-local parameter”. Alignment sam files were sorted and converted to bam files with samtools (v 1.6). Expression values were extracted from the alignment results using *express* (v 1.5.1) and RPKM (Reads Per Kilobase of transcript, per Million mapped reads) were calculated after libraries size normalization with the Bioconductor package *edgeR* in the R statistical environment. The expression profile was plotted so that the more intense the red color, the less expressed the gene was presented to the corresponding library and how much the more intense the white color, the greater its expression found in the data used for the analysis.

Results

Identification and characterization of putative proteins of the UPS complex: alignment and phylogenetic analysis

119 putative sequences were found in genome, transcriptome and predicted proteome of *B. glabrata* involved in the UPS. Among them 86 were found to be involved with the ubiquitination pathway and 33 with the 26S proteasome formation. Among the 86 genes of the ubiquitination pathway, 1 ubiquitin gene, 6 E1, 22 E2 and 39 E3, divided into the HECT, RING finger and U-box groups were identified. In addition, 19 genes were found to be related to the formation of 26S proteasome and involved with the regulatory portion, while 14 involved in form the nucleus of the molecule (Table 1). The genes were grouped according to their characteristics and functions based on the analysis of conserved domains, active sites and phylogenetic analysis.

Ubiquitination pathway

The analysis performed by PFAM and CDD showed that all the genes found for the ubiquitination pathway in *B. glabrata* have conserved domains and active sites/catalytic clefts/structural motifs. These conserved domains identified for each set of sequences correspond to their previously known structures. For Ubiquitin (BGLB020284-PA) gene, the Ubiquitin Family domain was identified, as well as found in the organisms *D. melanogaster* and *C. elegans* (Fig. S1 A). This domain is approximately 72 amino acids in size and shows some important residues for the molecule as those involved in the interaction of the molecule with E1 and E2. In addition, lysine residues at specific positions enable the formation of polyubiquitin chains (Fig. S1 B).

In the genes identified as E1 some domains were found distributed and overlapped (Table 2). These sequences showed the ThiF domain as the main representative showing a fairly characteristic size and location between the sequences of *B. glabrata* (Fig. 1). Besides that, it was possible to observe the presence of a cysteine residue as active site of these enzymes and other amino acid residues relevant to the performance of the catalytic activity

that it possess (Fig. 2). On the other hand, E2 are molecules capable of interacting with the other two classes of enzymes of the ubiquitination pathway because they are in the middle of this cascade. Genes found in the *B. glabrata* shared the UBQ_con domain and have a conserved cysteine residue as the active site, responsible for the role of conjugation of the ubiquitin molecule with the E3 (Fig. S2 A and B).

The ubiquitin ligases 3 were divided into HECT, RING finger and U-box and grouped based on conserved domains HECT domain, U-box and domains homologous to RING finger, demonstrating the conservation of the catalytic activity of each group of enzymes (Fig. S3 A, S4 A and S5 A). The analyzes performed by the CDD show that the genes of the HECT group had a catalytic cleft formed from the presence and distribution of some amino acid residues in specific and conserved positions (Fig. S3 B). The genes classified in the RING finger and U-box groups evidenced the formation of a structural motif for each group. For the RING finger gene group, the cysteine residue was conserved in all *B. glabrata* genes (Fig. S4 B). On the other hand, the motif found in the U-box gene group was determined by the distribution of various residues located along the sequences, but there is a highly conserved aspartic acid residue (Fig. S5 B).

Phylogenetic analysis was performed to evaluate the evolutionary relationship of the genes found in *B. glabrata* against the genes of orthologous organisms and to demonstrate the degree of conservation. For all classes of enzymes, this result showed that there was a very organized division between the clades of the enzymes and that they were distributed evolutionarily as presented by the evolutionary tree of life (Fig. S6, S7, S8 and S9). This infers that the sequences found in the snail were likely enzymes involved in the ubiquitination pathway. In addition, the distribution found among the organisms used for this analysis was demonstrated in figure 3, showing the analysis done for E1.

227 **26S proteasome formation**

228 The genes involved with the proteasome formation identified in *B. glabrata* were
 229 divided between regulatory particles and core particles. Groups called PAS and PBS were
 230 related to the nucleus of the molecule, while the RPT and RPN groups form the recognition
 231 portions. All genes identified as PAS showed the conserved domain named Proteasome (Fig.
 232 4) and an active site consisting of amino acid residues distributed at five specific positions
 233 throughout the seven sequences (Fig. 5). In the same sense, the PBS group was also shown to
 234 share the proteasome conserved domain in their sequences and, in addition, evidenced the
 235 presence of an active site consisting of amino acid residues in eight distinct positions (Fig.
 236 S10 A and B). This feature may indicate a direct relationship formed between site-position
 237 and not between site and amino acid conservation.

238 The other identified genes were divided into the RPT and RPN groups that formed the
 239 regulatory and recognition portion of 26S proteasome. The gene group called RPT showed no
 240 amino acid residue on active site formation, however, all six sequences presented the
 241 conserved domain AAA in the C-terminal portion (Fig. S11). Unlike the RPT group, the
 242 sequences of the RPN group did not share the common domain (Fig. S12 A) and showed a
 243 poorly conserved active site, suggesting that it is more involved with the target substrate (Fig.
 244 S12 B).

245 The phylogenetic analysis performed for the sequences involved with the formation of
 246 26S proteasome showed an evolutionary distribution that corroborates with that known
 247 through the tree of life of the animals. In addition the genes identified in *B. glabrata* and their
 248 orthologous from other organisms were organized in deuterostomes and protostomes as seen
 249 for the results found in the phylogeny concerning the ubiquitination pathway (Fig. S13, S14
 250 and S15). Figure 6 showed the phylogenetic distribution of the PAS group using those
 251 identified in the snail and the the orthologous genes.

Expression profile of the genes involved in the UPS identified in B. glabrata

The heatmap demonstrated the expression profile of all 119 sequences found in the *B. glabrata* in 12 libraries of tissues: albumen gland (AG), buccal mass (BUC), central nervous system (CNS), digestive gland/hepatopancreas (DG/HP), muscular part of the headfoot (FOOT), heart including amebocyte producing organ (HAPO), kidney (KID), mantle edge (MAN), ovotestis (OVO), salivary gland (SAL), stomach (STO) and terminal genitalia (TRG). The vast majority of the UPS related sequences presented low to medium expression for all libraries used (Fig. 7).

The ubiquitin (ubq-1) gene, identified as BGLB020284-RA was the only one presented highly expressed in all libraries. Other genes such as rpt-2 (BGLB012308-RB), wwp-1-a (BGLB008139-RC) and wwp-1-b (BGLB008139-RB) also demonstrated satisfactory, but not high, expression in all libraries. However, most other genes were not able to evidence a high expression profile for any of the libraries. Only a few transcripts such as pbs-3 (BGLB010370-RB), uba1-a (BGLB013827-RB) and rpn-10 (BGLB003999-RB) were able to display good expression in specific libraries such as terminal genitalia (TRG), ovotestis (OVO) and central nervous system (CNS), respectively (Fig. 7).

The genes with lower expression indicated in the results of the heatmap were related to the class of ubiquitin ligases (E3). All isoforms of the chn-1 genes (BGLB001813-PI, BGLB001813-PH, BGLB001813-PG, BGLB001813-PF, BGLB001813-PE, BGLB001813-PD, BGLB001813-PC and BGLB001813-PB) and traf-6 (BGLB026988-PC, BGLB026988-PB and BGLB026988-PA) demonstrated a very low RPKM-based expression profile. In the same sense, a very low expression was perceived for the isoforms of the fizzy genes (BGLB002564-PD, BGLB002564-PC, BGLB002564-PB, BGLB005726-PB and

BGLB005726-PC) and skpt-1 (BGLB032982-PA) that were involved in the formation of E3 complexes as target recognizing subunit (Fig 7).

Discussion

The ubiquitin-proteasome system (UPS) is one of the most conserved components among eukaryotic organisms and has been described in model organisms such as *D. melanogaster*, *C. elegans* and *Homo sapiens*. Most sequences identified in *B. glabrata* have an amount of amino acids close to or even equal to the number of amino acids found in the proteins of the model organism used, *C. elegans* and/or *D. melanogaster*. The ubiquitin is a polypeptide having 76 amino acids in its structure capable of labeling target substrates⁽²⁵⁾. However, the gene identified in *B. glabrata* encodes a gene product of 229 amino acids due to a polygene responsible for translating this molecule. In addition, in *C. elegans* and *D. melanogaster* the same situation occurs (Fig. S1 A). The difference found in the size of these sequences can be explained by the annotation performed for the mollusk sequence, suggesting that a new annotation process be performed. This situation refers to the existence of polygenes that have tandem repeats of the coding region of ubiquitin in their sequences. Thus, most organisms present sequences of polygenes for ubiquitin and from that, they have become one of the best models for the study of evolution from clusters⁽²⁶⁾. All three organisms used for this analysis have the conserved Ubiquitin Family domain (PF00240) in the same positions, showing that the presence and location of the domain in the sequence may be of great importance for the role that this polypeptide plays.

The Ubiquitin Family domain in *B. glabrata* genes shared very important and highly conserved amino acids against the orthologous organisms. Some of these residues have the ability to form a site of interaction with the ubiquitin conjugating enzymes (E2) and to

catalyze the transfer (Fig. S1 B). Another three lysine (K) residues play a crucial role in labeling the target substrate. These residues at positions 29, 48 and 63 of the domain are considered binding sites for novel ubiquitin molecules capable of forming polyubiquitin chains. The chains formed at positions 29 and 48 induce degradation by the proteasome, while the chains originating at the lysine residue at position 63 are related to DNA damage response signaling, for example⁽²⁷⁾. On the other hand, monoubiquitin proteins are involved in other cellular activities such as gene transcription⁽²⁸⁾.

It was not possible to trace the evolutionary history of ubiquitin through phylogenetic analyzes because it is an extremely conserved sequence (Fig. S1 B). There is a high degree of conservation involving the ubiquitin encoding gene, suggesting that it is able to undergo few combined evolutionary events⁽²⁹⁾. The reason why this conservation is so high is still not fully understood, but the properties presented by the ubiquitin molecule in the cells were selected and fixed in a eukaryotic ancestor in the early stages of evolution⁽³⁰⁾.

E1 are the first enzymes recruited into the ubiquitination pathway and play the role of activating the ubiquitin molecule to be transferred to the substrate. Activation occurs by the ubiquitin adenylation from an ATP-dependent thioester bond between the E1 cysteine residue and the glycine of the C-terminal portion of ubiquitin. The activated ubiquitin molecule is transferred to an E2⁽³¹⁻³²⁾. An existing hypothesis is that the E1 found in eukaryotes have evolved from the bacterial enzymes MoeB and ThiF. In this way, the ThiF domains present a great participation in the activation reaction of the ubiquitin molecule⁽³³⁾. All sequences identified as E1 in *B. glabrata* showed the ThiF domain with a similar amount of amino acids, showing great conservation, except in the sequence translated from the *rfl-1* gene which possibly has incomplete annotation. However, other conserved domains were also found in these sequences (Table 2) capable of assisting in the process of ubiquitin activation. A highly conserved cysteine (C) residue was observed as an active site among all sequences identified

as E1 in *B. glabrata*, the acceptor portion of ubiquitin in the enzyme. In addition, other amino acid residues in close positions are capable of aiding in the formation of the catalytic site and consequently in the activity that the enzyme plays (Fig. 2). Thus, these residues are probably involved with ATP binding and the formation of the thioester intermediate, activating the ubiquitin molecule for a subsequent transfer to E2⁽³⁰⁾.

E2 are enzymes from the center of the ubiquitination pathway and are related to the other two enzymatic classes in the cascade. In addition, they play the role of catalyzing the binding reaction of the ubiquitin molecule with the target substrate⁽³⁰⁾. The UQ_con domain was identified in all sequences identified as E2 present in the *B. glabrata* data, sharing close and highly conserved sizes. On the other hand, the UBA domain was found only in the sequence encoded from the *ubc-20* gene (BGLB036058-PA) and has a role of limiting the formation of the polyubiquitins chain (Fig. S2 A). In addition, these enzymes have a cysteine residue as active site capable of binding the ubiquitin molecule to E2 and mediates the interaction with E3⁽³⁰⁾. Other amino acid residues were important in the sequences because they were related to the interaction E2 makes with E3, however, they do not show great conservation (Fig. S2 B).

E3 perform the process of transferring the ubiquitin to the N-terminal portion of the lysine residue (K) belonging to the target substrate, are more related to the target substrate it recognizes than to the ubiquitin molecule and they are encoded in hundreds in eukaryotic cells, allowing the labeling of different proteins in a specific way and are organized into different classes (Table 1)⁽³⁴⁾. Thus, each of these classes shares specific domains that differ by recognition of the E2-ubiquitin complex⁽³⁵⁾. All E3-HECT sequences identified in *B. glabrata* showed a fairly conserved HECT domain in their C-terminal portion, maintaining a position pattern at the end of the protein chain. The members of this family are divided into three classes: NEDD4, HERC and other HECTs. The NEDD4 enzyme class shares the

presence of the WW domain in several copies and the C2 domain in its N-terminal portions⁽³⁶⁾, corroborating with the results found in this work for the sequences of the smurf-1-a, smurf-1-b, wwp-a, wwp-b and Y92H12A.2 (Fig. S3A). The absence of the C2 domain in the sequences encoded by the smurf-1 gene can be suggested by failure in the protein annotation process. On the other hand, sequences designated as herc-2, herc-4-a, herc-4-b, herc-4-c belong to the HERC family and share the RCC1 and HECT domains⁽³⁶⁾ (Fig. S3 A). In the HECT domain it was possible to find amino acid residues involved in the formation of a catalytic cleft and a cysteine as active site (Fig. S3B) where a thiol linkage between ubiquitin and ligase is performed. Although this residue was present in the C-terminal portion of the domain, its N-terminal portion is responsible for interacting with E2 and with the substrate that is bound in distinct regions along the enzyme sequence and outside the domain region⁽³⁶⁾. However this binding does not happen directly, because a recruitment of the E2-ubiquitin complex is first made to the active site of the ligase. Subsequently, ubiquitin is transferred to the lysine residue belonging to the target substrate through a transesterification reaction^(27, 30, 36).

The E3-RING finger class has a direct transfer of ubiquitin, since a concomitant recruitment between the E2-ubiquitin complex and target substrate is performed. This mediates and facilitates the formation of the binding between ubiquitin and the target protein^(27, 36-37). The conserved domains zf-C3HC4_3 (PF13920), zf-MIZ (PF02891), zf-RING_2 (PF13639), Sina (PF03145), Prok-RING_4 (PF13923) e zf-C3HC4_2 (PF14447) are homologous to the RING finger domain (Fig. S4 A) and were found distributed among all the sequences identified in the mollusk, presenting a profile consistent with the paper performance they have⁽³⁸⁻⁴¹⁾. A highly conserved distribution of cysteine, cysteine, histidine and cysteine was found in the results of this work, providing evidence that there was a structural motif formation (Fig. S4 B) and that it is related to the bond made with zinc⁽⁴²⁻⁴³⁾.

Like E3-RING finger, the E3-U-box class also recruits the E2-ubiquitin complex and target substrate concomitantly. In addition, they share a common organization and architecture, suggesting that U-box domains are modified RING domains because they do not bind to zinc and are formed from hydrogen bonds⁽⁴⁴⁻⁴⁵⁾. Thirteen of the fifteen E3-U-box sequences identified in *B. glabrata* demonstrated the U-box domain (CL0229) and the other two inferred other correlated domains (Fig. S5 A). The *ufd-2* (BGLB026024-PA) sequence showed only the Ufd2P_core domain (PF10408) which escapes ubiquitinated proteins to the proteasome and was directly linked to the U-box domain in the C-terminal portion of the sequence⁽⁴⁶⁾. The other gene was designated *cyn-4* (BGLB022189-PA), which was found the Rtf2 domain (PF04641), similar to the RING finger domain, but has a three-dimensional ring-like structure and only one site for binding with zinc⁽⁴⁷⁾. The absence of U-box domain may be related to the annotation process of these sequences. The results obtained showed that in the U-box domain there exists the formation of a structural motif in which not all amino acid residues were highly conserved, but there are some that were present in all sequences used in the analysis (Fig. S5 B). In this way, it was possible to infer that the residues of amino acids involved in the formation of structural motifs were related to the performance of the function of these sequences, since enzymes have a direct relation between activity and three-dimensional structure.

The proteasome is a proteolytic complex formed from an association between a 19S regulatory particle (PA700) and a central component 20S, composed of approximately 70 subunits. The 19S particle is divided into base and lid and are reversibly bound to 20S. The regulatory portion recognizes the labeled substrate while the 20S nucleus is responsible for degrading the protein in minor peptides^(7, 17-18).

19 genes from the RPN and RPT groups were identified in our analyzes (Table 1), encompassing sequences of major importance such as RPN-11, RPN-15 and RPN-3, involved

with target substrate recognition and removal of the ubiquitin chain formed^(19, 48-50). In the same sense, some RPT subunits, such as RPT-2 and RPT-5 were also found in the data of *B. glabrata*, however, they are related to activation of the 20S nucleus⁽⁵¹⁾. The results of this work showed that RPT shares of the conserved AAA domain, but not of an active site, structural motif or catalytic cleft, inferring that the function of these molecules was linked to the sequence that the conserved domain possesses (Fig. S11). On the other hand, sequences from the RPN group did not have a common domain, but have been shown to have an active site which seems to be more related to the distribution and position of amino acid residues than their conservation (Fig. S12 A e B). A suggested explanation for this feature was that these subunits may be more involved with the target substrate than with the catalytic portion of the 26S proteasome and therefore do not share highly conserved sequences.

The nucleus of the proteasome is composed of α and β subunits. Subunits of the α -type have a structural and regulatory role whereas β possesses proteolytic and degradation activity. 14 sequences involved with the formation of 20S proteasome between α and β subunits were identified in *C. elegans*⁽⁵¹⁾, corroborating the number of sequences found in *B. glabrata* data and identified as particles involved in the formation of the proteasome nucleus (Table 1). All sequences belonging to the PAS (α -subunit) and PBS (β -subunit) groups have the conserved domain Proteasome in their N-terminal portions (Fig. 4 and Fig. S10 A) demonstrating that the presence of this domain in this region is important for the activity they play. An active site composed of five relatively conserved amino acid residues was identified in the results obtained for the PAS group genes among all sequences (Fig. 5). On the other hand, for the PBS group a less conserved active site was found but possibly involved with the proteolysis function that this subunit possesses (Fig. S10 B), be it with activities chymotrypsin-like, trypsin-like or caspase-like^(20, 52).

Through the phylogenetic analyzes it was possible to describe the evolutionary history of the sequences under study and from these results to infer their identity by the location in the tree generated, the proximity to the orthologous organisms and model organisms. Our results show a clear and specific distribution among all data obtained from the analysis performed using MEGA5.2. The evolutionary organization of each set of sequences, both the sequences involved in the ubiquitination pathway and those related to the formation of the 26S proteasome (Fig. 3 and Fig. 6), corresponds to what was already known in the tree of life of the animals. There was a well-defined division between the clade of deuterostomia and protostomia organisms, separated between phyla, and grouped according to the location of the species in the evolutionary process⁽⁵³⁾. In the same sense, each sequence identified in *B. glabrata* presents well-defined ramifications involving its orthologous organisms, providing strong evidence that the putative sequences found in the mollusk transcriptome are real sequences involved in the ubiquitin-proteasome complex.

Our results demonstrated that the gene responsible for translating the ubiquitin (ubq-1) molecule was the only one to present a high expression profile in the RNAseq data of all 12 libraries used, that is, in all analyzed tissues of the adult snail. Ubiquitin is a key molecule in UPS, however, it is also involved in other cellular signaling processes such as plasma membrane transport, DNA repair, among others⁽²⁷⁾. This only happens because of the role it plays in labeling target substrates with one or more ubiquitins, forming a polyubiquitin chain. In addition to these functions, ubiquitin is also involved in autophagy^(27, 54-55).

The high expression that the ubq-1 transcript has demonstrated can be explained by the diversity of cellular processes in which ubiquitin is involved, being related to virtually all cellular processes, directly or indirectly⁽⁵⁵⁾. This also explains the high expression in all organs of *B. glabrata* used for analysis. In addition, the formation of polyubiquitin chains

influences that a greater amount of this gene was transcribed and may be translated, since more than one molecule of ubiquitin is required to label only a target substrate.

Most of the genes involved in UPS and identified in the *B. glabrata* genome presented a median expression for all organs of uninfected adult snails evaluated. This expression profile infers that both the transcripts involved in the ubiquitination pathway and genes participating in the formation of the 26S proteasome do not show a high expression under normal conditions of mollusc survival. Thus, it can be inferred that these transcripts share a behavior related to the expression very similar for all organs analyzed. Our results indicate that this was a pathway capable of performing its role in the organism without large amounts of circulating molecules in its basal state. However, when referring to infected snails, the expected results for these transcripts may be different, demonstrating greater expression, since UPS plays a role related to immunity.

Some transcripts appear to have a direct relationship with specific tissues because they have a very low expression profile for 11 libraries and relatively high for only one, suggesting that the function they play may be linked to their location in the organism. For example, *rpn-10* (BGLB003999-RB) demonstrates better expression in the central nervous system (CNS) than in the other tissues (Fig. 7). The deletion of *rpn-10* in *D. melanogaster* leads to lethality and demonstrates cause abnormalities in the mitotic cycle such as aneuploidies and absence of chromosomal segregation in CNS larval cells, besides accumulating multiubiquitin proteins in the tissue, since this is a gene participant of the 26S proteasome regulatory portion⁽⁵⁶⁾. In this way, it is interesting to investigate in depth the relationship between transcripts and tissues demonstrated in this work, its functions and importance in organisms.

Four families of genes showed a low expression for all libraries used. All were involved with ubiquitin ligases (E3), two of which refer to transcripts of E3 enzymes and two were genes encoding molecules that form part of complexes. The enzymes ubiquitin ligases

are transcribed from the genome in the hundreds, because they have the characteristic of being related directly to the substrate and not with the molecule of ubiquitin^(7, 10-11). In the same sense, because of the specificity they present for the labeled substrate, they do not need to be transcribed in large quantities into the mollusk tissues in non-stress situations. Traf-6 is a family of genes that presented one of the smallest expressions according to the results obtained, however, several works relate this transcript to the cellular signaling involving the immune system⁽⁵⁷⁻⁵⁹⁾. Thus, the low expression can be explained by the fact that the data from the analyzed libraries refer to uninfected adult snail transcripts. Traf-6 becomes a possible gene to be investigated in future research using data from molluscs under infection situation, suggesting that this expression may be increased in front of the parasite. In addition to traf-6, other genes involved in the ubiquitination pathway found in our results have also been described involved with immunity, such as diap1 (BGLB013412-RB) and ubc-13 (BGLB033527-RA and BGLB033527-RB). All presented a medium expression profile, but also suggest that genes may have a differential expression in cases of infected snails.

Our results allow the beginning of the search for new knowledge and perspectives involving *B. glabrata*, increasing the amount of information on the regulation of the mollusk and the relation between parasite/host. This work evidences the presence of UPS in the genome and transcriptome of the snail, besides offering a delineation of the expression profile of the 119 sequences identified for mollusk in their normal development. Our work also provides subsidies for new research related to the expression profile in different phases of the life of the organism, as well as the evaluation of the behavior of this pathway when the intermediate host is infected by *S. mansoni*, suggesting a super expression of genes participating in UPS and involved in the internal defense system. Therefore, these results may offer new ways of controlling infection and, consequently, schistosomiasis.

Competing interests

We declare that there is no conflict of interest, whether financial or non-financial origin, involved in the publication of this article.

Author's contributions

LGP and BCDD accomplished the in silico analyzes and drafted the manuscript. FRQ, THCR, ERM, MSG and RLC have participated in the experimental design and writing of the manuscript. WJJ, PMZC and EHB participated in the study design. MSG also coordinated the project.

Acknowledgements

We appreciate to the funding agency CNPq (National Research Council), the researchers responsible for moluscario Lobato Paraense and group of research involved in the sequencing of the genome and transcriptome of *Biomphalaria glabrata*.

References

1. Souza CP de, Jannotti-Passos LK, Freitas JR de 1995. Degree of host-parasite compatibility between *Schistosoma mansoni* and their intermediate molluscan hosts in Brazil. *Memórias do Inst. Oswaldo Cruz* 90: 5–10.
2. Carvalho O dos S, Coelho PMZ, Lenzi HL 2008. *Schistosoma Mansoni* e Esquistossomose: Uma Visão Multidisciplinar. *Schistosoma mansoni e esquistossomose uma visão Multidiscip.*: 960.
3. Richards CS, Shade PC 1987. The genetic variation of compatibility in *Biomphalaria glabrata* and *Schistosoma mansoni*. *J. Parasitol.* 73: 1146–1151.
4. Negrão-Corrêa D, Mattos ACA, Pereira CAJ, Martins-Souza RL, Coelho PMZ 2012. Interaction of *Schistosoma mansoni* sporocysts and hemocytes of *Biomphalaria*. *J. Parasitol. Res.* 2012.

- 530 5. Larson MK, Bender RC, Bayne CJ 2014. Resistance of *Biomphalaria glabrata* 13-16-R1
531 snails to *Schistosoma mansoni* PR1 is a function of haemocyte abundance and
532 constitutive levels of specific transcripts in haemocytes. *Int. J. Parasitol.* 44: 343–353.
533
- 534 6. Adema CM, Hillier LW, Jones CS, Loker ES, Knight M, Minx P, Oliveira G, Raghavan
535 N, Shedlock A, Amaral LR do, Arican-Goktas HD, Assis JG, Baba EH, Baron OL,
536 Bayne CJ, Bickham-Wright U, Biggar KK, Blouin M, Bonning BC, Botka C, Bridger
537 JM, Buckley KM, Buddenborg SK, Lima Caldeira R, Carleton J, Carvalho OS, Castillo
538 MG, Chalmers IW, Christenssens M, Clifton S, Cosseau C, Coustau C, Cripps RM,
539 Cuesta-Astroz Y, Cummins SF, Stephano L di, Dinguirard N, Duval D, Emrich S,
540 Feschotte C, Feyereisen R, FitzGerald P, Fronick C, Fulton L, Galinier R, Gava SG,
541 Geusz M, Geyer KK, Giraldo-Calderón GI, Souza Gomes M de, Gordy MA, Gourbal B,
542 Grunau C, Hanington PC, Hoffmann KF, Hughes D, Humphries J, Jackson DJ, Jannotti-
543 Passos LK, Jesus Jeremias W de, Jobling S, Kamel B, Kapusta A, Kaur S, Koene JM,
544 Kohn AB, Lawson D, Lawton SP, Liang D, Limpanont Y, Liu S, Lockyer AE, Lovato
545 TL, Ludolf F, Magrini V, McManus DP, Medina M, Misra M, Mitta G, Mkoji GM,
546 Montague MJ, Montelongo C, Moroz LL, Munoz-Torres MC, Niazi U, Noble LR,
547 Oliveira FS, Pais FS, Papenfuss AT, Peace R, Pena JJ, Pila EA, Quelais T, Raney BJ,
548 Rast JP, Rollinson D, Rosse IC, Rotgans B, Routledge EJ, Ryan KM, Scholte LLS,
549 Storey KB, Swain M, Tennesen JA, Tomlinson C, Trujillo DL, Volpi E V., Walker AJ,
550 Wang T, Wannaporn I, Warren WC, Wu X-J, Yoshino TP, Yusuf M, Zhang S-M, Zhao
551 M, Wilson RK 2017. Whole genome analysis of a schistosomiasis-transmitting
552 freshwater snail. *Nat. Commun.* 8: 15451.
553
- 554 7. Glickman MH, Ciechanover A 2002. The Ubiquitin-Proteasome Proteolytic Pathway:
555 Destruction for the Sake of Construction. *Physiol. Rev.* 82: 373–428.
556
- 557 8. Ciechanover A 1998. The ubiquitin-proteasome pathway: On protein death and cell life.
558 *EMBO J.* 17: 7151–7160.
559
- 560 9. Ebner P, Versteeg GA, Ikeda F 2017. Ubiquitin enzymes in the regulation of immune
561 responses. *Crit. Rev. Biochem. Mol. Biol.* 52: 425–460.
562
- 563 10. Scheffner M, Nuber U, Huibregtse JM 1995. Protein ubiquitination involving an E1–E2–
564 E3 enzyme ubiquitin thioester cascade. *Nature* 373: 81–83.
565
- 566 11. Zheng N, Shabek N 2017. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu.*
567 *Rev. Biochem.* 86: 129–157.
568
- 569 12. Arendt C, Hochstrasser M 1997. Identification of the yeast 20S proteasome catalytic
570 centers and subunit interactions required for active-site formation. *Proc. Natl.* 94: 7156–
571 7161.
572

- 573 13. Tanaka K, Chiba T 1998. The proteasome: A protein-destroying machine. *Genes to Cells*
574 3: 499–510.
575
- 576 14. Budenholzer L, Cheng CL, Li Y, Hochstrasser M 2017. Proteasome Structure and
577 Assembly. *J. Mol. Biol.* 429: 3500–3524.
578
- 579 15. Jung T, Catalgol B, Grune T 2009. The proteasomal system. *Mol. Aspects Med.* 30: 191–
580 296.
581
- 582 16. Morel S, Lévy F, Burlet-Schiltz O, Brasseur F, Probst-Kepper M, Peitrequin AL,
583 Monsarrat B, Velthoven R Van, Cerottini JC, Boon T, Gairin JE, Eynde BJ Van den
584 2000. Processing of some antigens by the standard proteasome but not by the
585 immunoproteasome results in poor presentation by dendritic cells. *Immunity* 12: 107–117.
586
- 587 17. Rock KL, Gramm C, Rothstein L, Clark K, Stein R, Dick L, Hwang D, Goldberg AL
588 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the
589 generation of peptides presented on MHC class I molecules. *Cell* 78: 761–771.
590
- 591 18. Groettrup M, Khan S, Schwarz K, Schmidtke G 2001. Interferon- γ inducible exchanges
592 of 20S proteasome active site subunits: Why? In: *Biochimie*, pp. 367–372.
593
- 594 19. Murata S, Sasaki K, Kishimoto T, Niwa SI, Hayashi H, Takahama Y, Tanaka K 2007.
595 Regulation of CD8+T cell development by thymus-specific proteasomes. *Science* (80-.).
596 316: 1349–1353.
597
- 598 20. Murata S, Takahama Y, Kasahara M, Tanaka K 2018. The immunoproteasome and
599 thymoproteasome: functions, evolution and human disease. *Nat. Immunol.* 19: 923–931.
600
- 601 21. Lockyer AE, Spinks J, Kane RA, Hoffmann KF, Fitzpatrick JM, Rollinson D, Noble LR,
602 Jones CS 2008. *Biomphalaria glabrata* transcriptome: cDNA microarray profiling
603 identifies resistant- and susceptible-specific gene expression in haemocytes from snail
604 strains exposed to *Schistosoma mansoni*. *BMC Genomics* 9.
605
- 606 22. Zahoor Z, Lockyer AE, Davies AJ, Kirk RS, Emery AM, Rollinson D, Jones CS, Noble
607 LR, Walker AJ 2014. Differences in the gene expression profiles of haemocytes from
608 schistosoma-susceptible and -resistant *Biomphalaria glabrata* exposed to *Schistosoma*
609 *mansoni* excretory-secretory products. *PLoS One* 9.
610
- 611 23. Dinguirard N, Cavalcanti MGS, Wu XJ, Bickham-Wright U, Sabat G, Yoshino TP 2018.
612 Proteomic Analysis of *Biomphalaria glabrata* Hemocytes During *in vitro* Encapsulation
613 of *Schistosoma mansoni* Sporocysts. *Front Immunol.* 9: 2773.
614
- 615 24. Queiroz FR, Silva LM, Jesus Jeremias W De, Hideo Babá É, Caldeira RL, Coelho PMZ,
616 Souza Gomes M De 2017. Differential expression of small RNA pathway genes

associated with the *Biomphalaria glabrata*/*Schistosoma mansoni* interaction. *PLoS One* 12.

25. Hershko A, Ciechanover A 1992. The Ubiquitin System for Protein Degradation. *Annu. Rev. Biochem.* 61: 761–807.

26. Müller WEG 1995. Molecular phylogeny of metazoa (animals): Monophyletic origin. *Naturwissenschaften* 82: 321–329.

27. Ramos PC 2002. A via proteolítica dependente de ubiquitina/proteassoma. *Sociedade Portuguesa de Química.* 57-63.

28. Schnell JD, Hicke L 2003. Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *J. Biol. Chem.* 278: 35857–35860.

29. Sharp PM, Li WH 1987. Ubiquitin genes as a paradigm of concerted evolution of tandem repeats. *J. Mol. Evol.* 25: 58–64.

30. Zuin A, Isasa M, Crosas B 2014. Ubiquitin Signaling: Extreme Conservation as a Source of Diversity. *Cells* 3: 690–701.

31. Hershko A, Heller H, Elias S, Ciechanover A 1983. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J. Biol. Chem.* 258: 8206–8214.

32. Dou T, Gu S, Liu J, Chen F, Zeng L, Guo L, Xie Y, Mao Y 2005. Isolation and characterization of ubiquitin-activating enzyme E1-domain containing 1, UBE1DC1. *Mol. Biol. Rep.* 32: 265–271.

33. Lois LM, Lima CD 2005. Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *EMBO J.* 24: 439–451.

34. Rodrigo-Brenni MC, Morgan DO 2007. Sequential E2s Drive Polyubiquitin Chain Assembly on APC Targets. *Cell* 130: 127–139.

35. Amerik AY, Hochstrasser M 2004. Mechanism and function of deubiquitinating enzymes. *Biochim. Biophys. Acta - Mol. Cell Res.* 1695: 189–207.

36. Rotin D, Kumar S 2009. Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* 10: 398–409.

37. Pickart CM, Eddins M.J 2004. UBIQUITIN: Structures, functions, mechanisms. *Biochim. Biophys. Acta* 1695: 55–72.

38. Barlow PN, Luisi B, Milner A, Elliott M, Everett R 1994. Structure of the C3HC4 domain by ¹H-nuclear magnetic resonance spectroscopy: A new structural class of zinc-finger. *J. Mol. Biol.* 237: 201–211.
39. Wu L, Wu H, Sangiorgi F, Wu N, Bell JR, Lyons GE, Maxson R 1997. Miz1, a novel zinc finger transcription factor that interacts with Msx2 and enhances its affinity for DNA. *Mech. Dev.* 65: 3–17.
40. Hu G, Chung YL, Glover T, Valentine V, Look AT, Fearon ER 1997. Characterization of human homologs of the Drosophila seven in absentia (sina) gene. *Genomics* 46: 103–111.
41. Burroughs AM, Iyer LM, Aravind L 2011. Functional diversification of the RING finger and other binuclear treble clef domains in prokaryotes and the early evolution of the ubiquitin system. *Mol. Biosyst.* 7: 2261–2277.
42. Joazeiro CAP, Weissman AM 2000. RING finger proteins: Mediators of ubiquitin ligase activity. *Cell* 102: 549–552.
43. Metzger MB, Hristova VA, Weissman AM 2012. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J. Cell Sci.* 125: 531–537.
44. Aravind L, Koonin E V. 2000. The U box is a modified RING finger - A common domain in ubiquitination [1]. *Curr. Biol.* 10.
45. Morreale FE, Walden H 2016. *SnapShot: Types of Ubiquitin Ligases.* *Cel.* 165.
46. Tu D, Li W, Ye Y, Brunger AT 2007. Structure and function of the yeast U-box-containing ubiquitin ligase Ufd2p. *Proc. Natl. Acad. Sci.* 104: 15599–15606.
47. Inagawa T, Yamada-Inagawa T, Eydmann T, Mian IS, Wang TS, Dalgaard JZ 2009. Schizosaccharomyces pombe Rtf2 mediates site-specific replication termination by inhibiting replication restart. *Proc. Natl. Acad. Sci.* 106: 7927–7932.
48. Kish-Trier E, Hill CP 2013. Structural Biology of the Proteasome. *Annu. Rev. Biophys.* 42: 29–49.
49. Collins GA, Goldberg AL 2017. The Logic of the 26S Proteasome. *Cell* 169: 792–806.
50. Kato K, Satoh T 2018. Structural insights on the dynamics of proteasome formation. *Biophys. Rev.* 10: 597–604.

- 702 51. Smith DM, Chang SC, Park S, Finley D, Cheng Y, Goldberg AL 2007. Docking of the
 703 Proteasomal ATPases' Carboxyl Termini in the 20S Proteasome's α Ring Opens the Gate
 704 for Substrate Entry. *Mol. Cell* 27: 731–744.
 705
- 706 52. Rivett AJ, Hearn AR 2004. Proteasome function in antigen presentation:
 707 immunoproteasome complexes, Peptide production, and interactions with viral proteins.
 708 *Curr Protein Pept Sci* 5: 153–161.
 709
- 710 53. Donoghue MJ, Cracraft J 2004. Charting the tree of life. In: *Assem. Tree Life*, pp. 1–4.
 711
- 712 54. Papaevgeniou N, Chondrogianni N 2014. The ubiquitin proteasome system in
 713 *Caenorhabditis elegans* and its regulation. *Redox Biol.* 2: 333–347.
 714
- 715 55. Li W, Ye Y 2008. Polyubiquitin chains: Functions, structures, and mechanisms. *Cell.*
 716 *Mol. Life Sci.* 65: 2397–2406.
 717
- 718 56. Szlanka T 2003. Deletion of proteasomal subunit S5a/Rpn10/p54 causes lethality,
 719 multiple mitotic defects and overexpression of proteasomal genes in *Drosophila*
 720 *melanogaster*. *J. Cell Sci.* 116: 1023–1033.
 721
- 722 57. Hu H, Sun SC 2016. Ubiquitin signaling in immune responses. *Cell Res.* 26: 457–483.
 723
- 724 58. Zinngrebe J, Montinaro A, Peltzer N, Walczak H 2014. Ubiquitin in the immune system.
 725 *EMBO Rep.* 15: 28–45.
 726
- 727 59. Li J, Chai QY, Liu CH 2016. The ubiquitin system: A critical regulator of innate
 728 immunity and pathogen-host interactions. *Cell. Mol. Immunol.* 13: 560–576.

CONCLUSÃO

O estudo do sistema ubiquitina-proteassoma é amplo, envolvendo diversos organismos e hipóteses, demonstrando ser um ótimo e promissor complexo a ser avaliado para responder alguns questionamentos que ainda intrigam a ciência. Dentre eles, estão principalmente aqueles que envolvem a regulação gênica e a expressão de determinadas características importantes e cruciais para o desenvolvimento dos organismos. Nosso grupo de trabalho acredita que o SUP esteja ligado ao processo de interação que acontece entre o parasito e o hospedeiro invertebrado na esquistossomose mansoni, uma vez que o SUP é responsável por exercer uma regulação pós-traducional e algumas proteínas do molusco mudam seu padrão de expressão quando o mesmo é infectado.

Os resultados expostos demonstram que o SUP está presente em dados de transcritoma de *B. glabrata*, incluindo sequências correspondentes à proteínas e enzimas envolvidas na via de ubiquitinação, assim como sequências de genes participantes da formação estrutural e funcional do proteassoma 26S. A partir da presença de domínios conservados, sítios ativos e análise filogenética é possível inferir que as sequências encontradas nos dados do molusco em estudo sejam prováveis sequências de moléculas envolvidas com o SUP. No mesmo sentido, quando os transcritos desse sistema foram confrontados com as informações de organismos modelo como *D. melanogaster* e *C. elegans* houve uma grande compatibilidade, elevando as suposições de que há uma transcrição ativa desses genes em *B. glabrata*.

Os dados obtidos para a análise da expressão indicam a presença dos transcritos nos dados analisados, porém, não garante que o processo traducional destes aconteçam, sendo necessários mais estudos para comprovar a presença física destas moléculas no caramujo. Além disso, a ubiquitina, molécula chave no SUP, foi a única sequência a apresentar uma alta expressão, corroborando com o esperado diante da vasta gama de funções que desempenha nos organismos. No entanto, é preciso novas buscas experimentais para que a relação entre SUP e as características de resistência/suscetibilidade do caramujo seja confirmada, assim como responder a indagação sobre a existência da variação do perfil de expressão desse sistema diante da infecção por *S. mansoni*. Ainda assim, esse trabalho é capaz de oferecer base para novos entendimentos a cerca desse sistema no caramujo vetor da esquistossomose e da sua participação no desenvolvimento do molusco.

REFERÊNCIAS

- ADEMA, C. M. et al. A bacterial artificial chromosome library for *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni*. *Memorias do Instituto Oswaldo Cruz. Anais...* 2006. DOI:10.1590/S0074-02762006000900027.
- ADEMA, C. M. et al. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. *Nature Communications*, v. 8, p. 15451, 2017. DOI: 10.1038/ncomms15451.
- ARENDT, C. S.; HOCHSTRASSER, M. Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation. *Proceedings of the National Academy of Sciences of the United States of America*, v. 94, n. 14, p. 7156–7161, 1997. DOI: 10.1073/pnas.94.14.7156.
- BAYNE, C. J. Successful parasitism of vector snail *Biomphalaria glabrata* by the human blood fluke (trematode) *Schistosoma mansoni*: A 2009 assessment. *Molecular and Biochemical Parasitology*, 2009. DOI: 10.1016/j.molbiopara.2009.01.005.
- BAYNE, C. J.; BUCKLEY, P. M.; DEWAN, P. C. Macrophagelike hemocytes of resistant *Biomphalaria glabrata* are cytotoxic for sporocysts of *Schistosoma mansoni* in vitro. *The Journal of parasitology*, v. 66, n. 3, p. 413–9, 1980a. DOI: 10.2307/3280740.
- BAYNE, C. J.; BUCKLEY, P. M.; DEWAN, P. C. *Schistosoma mansoni*: Cytotoxicity of hemocytes from susceptible snail hosts for sporocysts in plasma from resistant *Biomphalaria glabrata*. *Experimental Parasitology*, v. 50, n. 3, p. 409–416, 1980b. DOI: 10.1016/0014-4894(80)90043-0.
- BUDENHOLZER, L. et al. Proteasome Structure and Assembly. *Journal of Molecular Biology*, 2017. DOI: 10.1016/j.jmb.2017.05.027.
- BRASIL. Vigilância e controle de moluscos de importância epidemiológica: diretrizes técnicas: Programa de Vigilância e Controle da Esquistossomose (PCE). *Editora do Ministério da Saúde*, n. 2ª edição, p. 177, 2008.
- CALDAS, I. R. et al. Human schistosomiasis mansoni: Immune responses during acute and chronic phases of the infection. *Acta Tropica*, v. 108, n. 2–3, p. 109–117, 2008. DOI: 10.1016/j.actatropica.2008.05.027.

CARVALHO, O. DOS S.; COELHO, P. M. Z.; LENZI, H. L. *Schistosoma Mansoni* e Esquistossomose: Uma Visão Multidisciplinar. ***Schistosoma mansonii* e esquistossomose: uma visão multidisciplinar**, p. 960, 2008. DOI: 10.7476/9788575413708.

CASTRO-BORGES, W. et al. The 20S proteasome of *Schistosoma mansonii*: A proteomic analysis. **Proteomics**, v. 7, n. 7, p. 1065–1075, 2007. DOI: 10.1002/pmic.200600166.

CIECHANOVER, A. The ubiquitin-proteasome pathway: On protein death and cell life. **EMBO Journal**, 1998. DOI: 10.1093/emboj/17.24.7151.

COURA, J. R.; AMARAL, R. S. **Epidemiological and control aspects of schistosomiasis in Brazilian endemic areas**. Memórias do Instituto Oswaldo Cruz. **Anais...**2004. DOI: 10.1590/S0074-02762004000900003.

DA MARTINS, D. S. et al. Schistosomiasis in Southern Brazil 17 years after the confirmation of the first autochthonous case. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 48, n. 3, p. 354–357, 2015. DOI: 10.1590/0037-8682-0240-2014.

DE SOUZA, C. P.; JANNOTTI-PASSOS, L. K.; DE FREITAS, J. R. Degree of host-parasite compatibility between *Schistosoma mansonii* and their intermediate molluscan hosts in Brazil. **Memórias do Instituto Oswaldo Cruz**, v. 90, n. 1, p. 5–10, 1995. DOI: 10.1590/S0074-02761995000100003.

EBNER, P.; VERSTEEG, G. A.; IKEDA, F. Ubiquitin enzymes in the regulation of immune responses. **Critical Reviews in Biochemistry and Molecular Biology**, v. 9238, n. June, p. 1–36, 2017. DOI: 10.1080/10409238.2017.1325829.

EL-ZAHAR, M.; EL-KARIM, S. Schistosomiasis (chemoprophylaxis and treatments). **Egyptian Pharmaceutical Journal**, v. 11, p. 1-15, 2012. DOI: 10.7123/01.EPJ.0000415037.38143.d4.

ESPÍRITO-SANTO, M. et al. Evaluation of real-time PCR assay to detect *Schistosoma mansonii* infections in a low endemic setting. **BMC infectious diseases**, v. 14, n. 1, p. 558. doi:10.1186/s12879-014-0558-4, 2014.

FÖRSTER, J.; GOMBERT, A. K.; NIELSEN, J. A functional genomics approach using metabolomics and in silico pathway analysis. **Biotechnology and Bioengineering**, v. 79, n. 7, p. 703–712, 2002. DOI: 10.1002/bit.10378.

GLICKMAN, M. H.; CIECHANOVER, A. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. **Physiological Reviews**, v. 82, n. 2, p. 373–428, 2002. DOI: 10.1152/physrev.00027.2001.

GRANATH, W. O.; YOSHINO, T. P. *Schistosoma mansoni*: Passive transfer of resistance by serum in the vector snail, *Biomphalaria glabrata*. **Experimental Parasitology**, v. 58, n. 2, p. 188–193, 1984. DOI: 10.1016/0014-4894(84)90034-1.

GRYSEELS, B. et al. Human schistosomiasis. **Lancet**, 2006. DOI: 10.1016/S0140-6736(13)61949-2.

HERSHKO, A; CIECHANOVER, A. The ubiquitin system for protein degradation. **Annu. Rev. Biochem.** v.61, p.761-807, 1992. DOI: 10.1146/annurev.bi.61.070192.003553.

JUNG, T.; CATALGOL, B.; GRUNE, T. The proteasomal system. **Molecular Aspects of Medicine**, 2009. DOI: 10.1016/j.mam.2009.04.001.

KATZ, N.; CHAVES, A.; PELLEGRINO, J. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. **Revista do Instituto de Medicina Tropical de São Paulo**, v. 14, n. 6, p. 397–400, 1972. ISSN: 0036-4665.

LARSON, M. K.; BENDER, R. C.; BAYNE, C. J. Resistance of *Biomphalaria glabrata* 13-16-R1 snails to *Schistosoma mansoni* PR1 is a function of haemocyte abundance and constitutive levels of specific transcripts in haemocytes. **Int J Parasitol**, v. 44, n. 6, p. 343-53, 2014. DOI: 10.1016/j.ijpara.2013.11.004.

LEHR, T. et al. Localization of carbohydrate determinants common to *Biomphalaria glabrata* as well as to sporocysts and miracidia of *Schistosoma mansoni*. **Parasitology**, v. 135, n. 8, p. 931–942, 2008. DOI: 10.1017/S0031182008004514.

LI, J.; CHAI, Q. Y.; LIU, C. H. The ubiquitin system: A critical regulator of innate immunity and pathogen-host interactions. **Cellular and Molecular Immunology**, 2016. DOI: 10.1038/cmi.2016.40.

LI, W.; YE, Y. Polyubiquitin chains: Functions, structures, and mechanisms. **Cellular and Molecular Life Sciences**, 2008. DOI: 10.1007/s00018-008-8090-6.

LOCKYER, A. E. et al. *Biomphalaria glabrata* transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to *Schistosoma mansoni*. **BMC Genomics**, v. 9, 2008. DOI: 10.1186/1471-2164-9-634.

LOKER, E. S.; BAYNE, C. J. Molecular studies of the molluscan response to digenean infection. **Adv Exp Med Biol**, v. 484, p. 209-22, 2001. ISSN 0065-2598 (Print) 0065-2598 (Linking).

MAGORI, S.; CITOVSKY, V. Hijacking of the Host SCF Ubiquitin Ligase Machinery by Plant Pathogens. **Frontiers in Plant Science**, v. 2, 2011. DOI: 10.3389/fpls.2011.00087.

MATHIESON, W.; CASTRO-BORGES, W.; WILSON, R. A. The proteasome-ubiquitin pathway in the *Schistosoma mansoni* egg has development- and morphology-specific characteristics. **Molecular and Biochemical Parasitology**, v. 175, n. 2, p. 118–125, 2011.

MINISTÉRIO DA SAÚDE. **PCE-Programa de Controle da Esquistossomose-Brasil**, 2017. Disponível em: < <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinan/pce/cnv/pcebr.def> >. Acesso em: 02 de jan. 2019.

NABHAN, J. F. et al. The 26S proteasome in *Schistosoma mansoni*: Bioinformatics analysis, developmental expression, and RNA interference (RNAi) studies. **Experimental Parasitology**, v. 117, n. 3, p. 337–347, 2007. DOI: 10.1016/j.exppara.2007.08.002.

NEGRAO-CORREA, D. et al. Interaction of *Schistosoma mansoni* Sporocysts and Hemocytes of *Biomphalaria*. **J Parasitol Res**, v. 2012, p. 743920, 2012. DOI: 10.1155/2012/743920.

PAPAEVGENIOU, N.; CHONDROGIANNI, N. The ubiquitin proteasome system in *Caenorhabditis elegans* and its regulation. **Redox Biology**, 2014. DOI: 10.1016/j.redox.2014.01.007.

PARAENSE, W.L. Estado atual da sistemática dos planorbídeos brasileiros. Arquivo Do Museu Nacional Do Rio de Janeiro. 1975a; 55: 105-28.

PARAENSE, W. L. e CORRÊA, L. R. Differential susceptibility of *Biomphalaria tenagophila* populations to infection with a strain of *Schistosoma mansoni*. **Journal of Parasitology**, v. 64, n. 4, p. 822-826, 1978.

PARAENSE WL. Histórico do Gênero *Biomphalaria*, Morfologia e Sistemática Morfológica. In: Carvalho OS, Coelho PMZ, Lenzi HL, editores. *Schistosoma mansoni* e esquistossomose: uma visão multidisciplinar. Rio de Janeiro: Fiocruz; 2008. P. 285-308. DOI: 10.7476/9788575413708.

PEREIRA, R. V. et al. MJD and OTU deubiquitinating enzymes in *Schistosoma mansoni*. **Parasitology Research**, v. 114, n. 8, p. 2835–2843, 2015a. DOI: 10.1007/s00436-015-4484-1.

PEREIRA, R. V. et al. Ubiquitin-specific proteases are differentially expressed throughout the *Schistosoma mansoni* life cycle. **Parasites & Vectors**, v. 8, n. 1, p. 349, 2015b. DOI: 10.1186/s13071-015-0957-4.

PICKART, C.M.; EDDINS, M.J. UBIQUITIN: Structures, functions, mechanisms. *Biochim. Biophys. Acta*, v. 1695, p. 55–72. 2, 2004. DOI: 10.1016/j.bbamcr.2004.09.019.

PLOWS, L. D. et al. Carbohydrates that mimic schistosome surface coat components affect ERK and PKC signalling in *Lymnaea stagnalis* haemocytes. **International Journal for Parasitology**, v. 35, n. 3, p. 293–302, 2005. DOI: 10.1016/j.ijpara.2004.11.012.

PROCTOR, C. J.; TSIRIGOTIS, M.; GRAY, D. A. An in silico model of the ubiquitin-proteasome system that incorporates normal homeostasis and age-related decline. **BMC Systems Biology**, v. 1, 2007. DOI: 10.1186/1752-0509-1-17.

QUEIROZ, F. R. Caracterização da maquinaria de processamento de miRNAs e piRNAs em *Biomphalaria glabrata* e o efeito da infecção por *Schistosoma mansoni* no processo. 2015. 118 f. Dissertação (Mestrado em Ciências da Saúde) – Centro de Pesquisas René Rachou, Instituto Oswaldo Cruz, Belo Horizonte. 2015. ISSN: 1471-0080.

QUEIROZ, F. R. et al. Differential expression of small RNA pathway genes associated with the *Biomphalaria glabrata*/*Schistosoma mansoni* interaction. **PLoS ONE**, v. 12, n. 7, 2017. DOI: 10.1371/journal.pone.0181483.

RADIOVOJAC, P. et al. Identification, analysis, and prediction of protein ubiquitination sites. **Proteins**, v. 78, n. 2, p. 365–80, 2010. DOI: 10.1002/prot.22555.

RAMOS, P. C., A via proteolítica dependente de ubiquitina/proteassoma. *Campus de Gambelas. Sociedade Portuguesa de Química*, p. 57-63, 2002.

RICHARDS, C. S.; SHADE, P. C. The genetic variation of compatibility in *Biomphalaria glabrata* and *Schistosoma mansoni*. **The Journal of parasitology**, v. 73, n. 6, p. 1146–51, 1987.

SALOMONS, F. A.; ÁCS, K.; DANTUMA, N. P. Illuminating the ubiquitin/proteasome system. **Experimental Cell Research**, 2010. DOI: 10.1016/j.yexcr.2010.02.003.

SAMBON, L.W. Remarks on *Schistosoma mansoni*. **Am J Trop Med Hyg**, v. 10, p. 303-304, 1907.

SARIKAS, A.; HARTMANN, T.; PAN, Z.-Q. The cullin protein family. **Genome biology**, v. 12, n. 4, p. 220, 2011. DOI: 10.1186/gb-2011-12-4-220.

SCHEFFNER, M.; NUBER, U.; HUIBREGTSE, J. M. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. **Nature**, 1995. DOI: 10.1038/373081a0.

SEISSLER, T.; MARQUET, R.; PAILLART, J. C. Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins. **Viruses**, v. 9(11), 2017. DOI: 10.3390/v9110322.

SILVA, C. A. Avaliação e retificação da identificação específica de moluscos do gênero *Biomphalaria* Preston, 1910 do acervo da Coleção de Malacologia Médica (Fiocruz-CMM). 2012. 104 f. Dissertação (Mestrado em Ciências da Saúde) – Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, Belo Horizonte. 2012.

SILVA, N. F. S. et al. Bioensaio de atividade moluscicida adaptado para avaliação de extratos de plantas medicinais. **Arq. Ciências Vet Zool**, v. 11, p. 179-181, 2008. DOI: 10.25110/arqvet.v11i2.2008.2575.

SILVA, M. A. P. Contribuição para o estudo da schistosomíase na Bahia. **Brazil-méd**, 22, p. 281-283, 1908. ISBN 978-85-7541-370-8.

SOUZA, C. P. Estudo de moluscos do gênero *Biomphalaria* de Minas Gerais, com relação a adaptação parasito hospedeiro e importância na epidemiologia da esquistossomose. **Revista do Instituto de Medicina Tropical**, v. 28, n.5, p. 287-929, 1986. DOI: 10.1590/S0036-46651986000500002.

TANAKA, K.; CHIBA, T. The proteasome: A protein-destroying machine. **Genes to Cells**, 1998. DOI: 10.1046/j.1365-2443.1998.00207.x.

VAN DER KNAAP, W. P. W.; LOKER, E. S. Immune mechanisms in trematode-snail interactions. **Parasitology Today**, 1990. DOI: 10.1016/0169-4758(90)90349-9.

WHO. Schistosomiasis. **Fact sheet N° 115**, 2017. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs115/en/>>. Acesso em: 02 de jan. 2019.

ZAHOOR, Z. et al. Differences in the Gene Expression Profiles of Haemocytes from Schistosome-Susceptible and -Resistant *Biomphalaria glabrata* Exposed to *Schistosoma mansoni* Excretory-Secretory Products. **PLoS ONE**, v. 9, n. 3, 2014. DOI: 10.1074/mcp.M800538-MCP200.

ZHENG, N.; SHABEK, N. Ubiquitin ligases: Structure, Function, and Regulation. **Annual Review of Biochemistry**, v. 86, p. 129-157, 2017. DOI: 10.1146/annurev-biochem-060815-014922.

ZUIN, A.; ISASA, M.; CROSAS, B. Ubiquitin Signaling: Extreme Conservation as a Source of Diversity. **Cells**, v. 3, n. 3, p. 690–701, 2014. DOI: 10.3390/cells3030690.

ANEXO

Anexo 1 – Instruções para autores da revista “Memórias do Instituto Oswaldo Cruz”.

INSTRUCTIONS TO AUTHORS

The Memórias' content is freely accessible to readers and no publication fees are charged to authors. The Memórias do Instituto Oswaldo Cruz has decided to simplify the requirements regarding the format of submitted manuscripts. From now on, all manuscripts may be submitted in any text format as long as the common subdivision of scientific articles are followed, e.g. introduction, materials and methods, results, discussion and references. For Reviews, Perspectives and similar articles, authors may use the sections that best suit the structure and content of the proposed manuscript. All manuscripts should contain, besides the title and abstract, full details of authors and institutions, acknowledgements of any technical or financial assistance as well as state any conflicts of interest. This flexible text format will be used for the initial analysis and peer review. If the manuscript is accepted, authors will be requested to edit the text in accordance with the publication style of the Memorias."

Upon acceptance, the manuscript should be arranged in the following format:

The manuscript should be prepared using standard word processing software and should be printed (font size 12) double-spaced throughout the text, figure captions, and references (must be up to 30 references), with margins of at least 3 cm. The figures should come in the extension tiff, with a minimum resolution of 300 dpi. Tables and legends to figures must be submitted all together in a single file. Figures, must be uploaded separately as supplementary file.

Running title: not required

Title: with up to 250 characters

Author's names: without titles or graduations

Institutional affiliations: full address of the corresponding author only

Abstracts: Provide an abstract of between 250- 300 words (100 words in case of short communications, technical notes, genome announcements or reviews). Abstracts of original articles should be structured into 5 sections as follows: BACKGROUND, OBJECTIVES, METHODS, FINDINGS and MAIN CONCLUSIONS, each section addressing respectively

the problem, the aim of the study, the main methodological approach, the most important findings and the conclusions of the study.

Key words: 3-6 items must be provided. Terms from the Medical Subject Headings (Mesh) list of Index Medicus should be used.

Sponsorships: indicating the sources of financial support and change of address.

Introduction: should set the purpose of the study, give a brief summary (not a review) of previous relevant works, and state what new advance has been made in the investigation. It should not include data or conclusions from the work being reported.

Materials and Methods: should briefly give clear and sufficient information to permit the study to be repeated by others. Standard techniques need only be referenced.

Ethics: when reporting experiments on human subjects, indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 1983. When reporting experiments on animals, indicate whether the institution's or a national research council's guide for, or any national law on the care and use of laboratory animals was followed.

Results: should be a concise account of the new information discovered, with the least personal judgement. Do not repeat in text all the data in the tables and illustrations.

In case of describing New Species, should follow:

Name of the new species, authors (when it is the case), sp. nov., (Figs x-y)
[Ex: *An. (Nyssorhynchus) atacamensis* González and Sallum, sp. nov. (Figs 1-4)]

Previous reference to the new species (when it is the case)
[Ex: *An. pictipennis* of Rueda et al. (2008): 448.]

Diagnosis (or Description; all stages are described);

Type host (when it is the case);

Site of Infection (when it is the case);

Type-locality;

Type data and depository;

Other material examined (when it is the case);

Distribution;

Host-parasite data (such prevalence and other important data, when it is the same case);

Bionomics;

Etymology;

Taxonomic discussion (or simply DISCUSSION as internal title).

Discussion: should be limited to the significance of the new information and relate the new findings to existing knowledge. Only unavoidable citations should be included.

Acknowledgements: should be short and concise, and restricted to those absolutely necessary.

Author's contribution: state each author's contribution to the work.

REFERENCES

Must be accurate. Only citations that appear in the text should be referenced. Unpublished papers, unless accepted for publication, should not be cited. Work accepted for publication should be referred to as "in press" and a letter of acceptance of the journal must be provided. Unpublished data should only be cited in the text as "unpublished observations", and a letter of permission from the author must be provided. The references at the end of the paper should be listed in numerical order, and in the same order in which they are cited in text. CLICK HERE [\[+\]](#)

FIGURES AND TABLES MUST BE UNDERSTANDABLE WITHOUT REFERENCE TO THE TEXT

Figures: presented in tiff format with a minimum of 300 dpi and photographs must be sharply focused, well contrasted, and if mounted onto a plate, the figures should be numbered consecutively with Arabic numbers. Magnification must be indicated by a line or bar in the figure, and referenced, if necessary in the caption (e.g., bar = 1 mm). Plates and line figures should either fit one column (8 cm) or the full width (16.5 cm) of the page and should be shorter than the page length to allow inclusion of the legend. Letters and numbers on figures should be of a legible size upon reduction or printing. A colour photograph illustrates the cover of each issue of the Journal and authors are invited to submit illustrations with legends from their manuscript for consideration for the cover.

Tables: should supplement, not duplicate, the text and should be numbered with Roman numerals. A short descriptive title should appear above each table, with any explanations or footnotes (identified with a, b, c, etc.) below.

Review: Papers in “review” format are accepted only by means of invitations made by the editor or associated editors.

Technical Notes: Technical Notes should communicate rapidly single novel techniques or original technical advances. The entire note should occupy no more than three printed pages including figures and/or tables (it means around 10 double-spaced typed Word file maximum). The text must not be divided into sections. Therefore, the state of art must be very briefly presented; results must be rapidly presented and discussed at a time. Complementary tables and figures may be published as supplementary data. References must be limited to few essential ones and cited at the end of the note, using the same format as in full papers. A brief summary and three key words must be provided.

Short communications: should communicate rapidly single results or techniques. They should occupy no more than three printed pages including figures and/or tables. They should not contain excessive references. References should be cited at the end of the paper using the same format as in full papers. A brief summary and three key words must be provided.

Genome Announcement and Highlights: this section is dedicated to publish new genome information from eukaryote parasites, virus, bacteria and their respective vectors. Authors who wants a fast peer review and publication cycle for their research results covering new genome sequences, re-sequencing and comparative genome analysis as well as the expression pattern of genomes are invited to submitted papers under the short communication format.

Alternative format: manuscripts may be submitted following the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" produced by the International Committee of Medical Journal Editors also known as the Vancouver Style. In this case, authors should follow the guidelines in the fifth edition (Annals of Internal Medicine 1997; 126: 36-47, or at the website <http://www.acponline.org/journals/resource/unifreq/htm>) and will be responsible for modifying the manuscript where it differs from the instructions given here, if the manuscript is accepted for publication.

Authors should also follow the Uniform Requirements for any guidelines that are omitted in these Instructions.

In case of clinical trials it's mandatory to inform the registration number of the REBEC platform.

A statement that the data/results of the manuscript are not plagiarism and have not been published elsewhere.

ONCE A PAPER IS ACCEPTED FOR PUBLICATION, THE AUTHORS MUST PROVIDE:

Page charges: there will be no page charges.

Proofs: one set of page proofs will be supplied for the author to check for typesetting accuracy, to be returned by the stipulated date. No changes to the original manuscript will be allowed at this stage.