



PARTICIPAÇÃO DE HP1 NAS MODIFICAÇÕES PÓS-TRADUCIONAIS EM *Schistosoma mansoni*: UMA ANÁLISE *IN SÍLICO*

Programa de iniciação científica voluntária PIBIC

Vigência: : 01/09/2022 - 31/08/2023

FAPESP: 2021/14982-6

FABIANA DOS SANTOS DA SILVA¹, NATÁLIA DA SILVA TRINDADE¹, FERNANDA JANKU CABRAL¹

IB-UNICAMP, Departamento de Biologia Animal, Laboratório de Parasitologia, Instituto de Biologia, Universidade Estadual de Campinas, Campinas-SP

Palavras-Chave: HP1, *Schistosoma mansoni*, Silenciamento gênico, Proteína, Modificações pós-traducionais.

INTRODUÇÃO

O *Schistosoma mansoni* é o parasita trematódeo responsável por causar a doença conhecida por esquistossomose. Seu ciclo de vida passa por dois hospedeiros, o molusco *Biomphalaria glabrata* (hospedeiro intermediário) e por pequenos mamíferos e o ser humano (hospedeiros definitivos), além de ser um ciclo bastante complexo devido às mudanças genéticas e morfológicas, que auxiliam no mecanismo de escape do parasito em relação ao sistema imunológico do hospedeiro [1, 2]. Essas mudanças envolvem metilação de histonas que são essenciais para regulação do estado da cromatina [2]. Neste caso, a Heterochromatin Protein 1 (HP1) é um importante co-regulador, inicialmente descoberta em estudos envolvendo *Drosophila sp.*, e que sugeriu/inferiu que a proteína desempenha diferentes funções [3]. No entanto, a principal função desempenhada pela proteína é como fator de transcrição, sendo responsável pela regulação da expressão gênica, e pelo estado repressivo da cromatina [4]. A proteína HP1 faz parte da família de proteínas chromobox, porém além de possuir o domínio chromo, tem um segundo domínio na sua cauda C-terminal, chromo_shadow, que direciona a HP1 para a heterocromatina [5]. Pesquisas recentes de interactoma em cercárias, indicam que a HP1 tem um papel significativo para o desenvolvimento do ciclo de vida do *Schistosoma mansoni*. Estes estudos também apontam a presença de regiões de conservação de domínio entre diferentes espécies [6].

OBJETIVO

Investigar e compreender o papel da proteína HP1 como modulador epigenético em *Schistosoma mansoni* e criar um banco de dados contendo informações sobre as proteínas HP1 encontradas em diferentes organismos.

METODOLOGIA

Através do código de entrada da proteína HP1 de *Schistosoma mansoni* (Smp_179650), foi obtido o formato FASTA da proteína através da interface do WormBase ([WormBase Parasite](#)), para encontrar similaridade entre a proteína HP1 do parasito com os proteínas que seriam obtidas no processo de triagem. Com isto, foi realizado uma triagem por meio do recurso NCBI ([National Center for Biotechnology Information](#)) para selecionar possíveis proteínas HP1 de organismos diferentes no database “Protein”. Os dados obtidos para pesquisa por “Heterochromatin protein 1”, foi selecionado apenas as proteínas que o código de entrada inicia com XP ou NP, que passaram pelo programa BLAST projetado no NCBI para análise de similaridade de sequências dessas proteínas com a proteína de *Schistosoma mansoni*. As proteínas que o programa encontrou similaridades entre as sequências foram coletadas e organizadas em uma tabela no excel.

Feito isso, com o auxílio da interface do InterPro ([Classification of protein families](#)), analisamos o formato FASTA de cada proteína selecionada com o objetivo de confirmar se essas proteínas continham os domínios chromobox e chromo_shadow. Com as informações obtidas pela pesquisa, foi organizada uma nova tabela com a localização dos domínios na sequência FASTA, juntamente com o código de entrada no InterPro.

Concluído a análise no InterPro, utilizamos o banco de dados UniProt ([Base de Conhecimento de Proteína Universal](#)) para pesquisar os códigos de acesso de cada proteína que foi selecionada na triagem, para coleta de dados funcionais. As proteínas que tinham uma função descrita no banco de dados, foram organizadas em uma nova tabela, enquanto as que não possuem uma função descrita, foram verificadas as publicações relacionadas a elas com o objetivo de sugerir qual a possível função que a proteína possui.

RESULTADO

Utilizando o formato FASTA da proteína HP1 do parasito *Schistosoma mansoni* na interface do WormBase, realizamos uma análise de similaridade de sequenciamento, comparando com as proteínas registradas no NCBI, utilizando a ferramenta projetada pelo banco de dados, o BLAST. Desta análise foi obtido possíveis proteínas HP1 de oito organismos: *Cephus cinctus*, *Dermacentor andersoni*, *Echinococcus granulosus*, *Fopius arisanus*, *Opisthorchis viverrini*, *Oryzias melastigma*, *Rousettus aegyptiacus*, *Schistosoma haematobium* e *Trichogramma pretiosum*.

As proteínas obtidas apenas da busca de similaridade de sequência de *Schistosoma mansoni*, foram organizadas no excel com o nome científico do organismo, descrição da proteína, *E value*, porcentagem de identidade e o código de entrada. Foram selecionadas proteínas com o percentual de identidade acima de 45%, esse valor indica a similaridade da sequência de aminoácidos com a sequência do organismo modelo. Este conjunto de informações contribuem para a confiabilidade dos dados e diminui as dúvidas relacionadas às proteínas [7]. Neste contexto, o *E value* também desempenha um importante papel para a confiabilidade dos dados, quando o valor é mais próximo de zero (0), maior a confiabilidade do alinhamento [8].

Diante disso, foi necessário realizar um novo refinamento devido ao resultado insuficiente de proteínas obtido da análise de similaridade da proteína do organismo modelo, para obter mais proteínas para as futuras comparações. Através do NCBI e a ferramenta BLAST, foi realizada a pesquisa por novos organismos com possíveis proteínas HP1. Dessa forma, obtivemos um total de 108.036 resultados do banco de dados, no entanto somente 171 foram analisados, permanecendo as proteínas de organismos que não eram repetidos e que possuíam similaridade com a proteína HP1 de *Schistosoma mansoni*, levando em consideração os códigos de entrada NP/XP, por serem códigos relacionados a uma proteínas.

Diante disso, foi necessário realizar um novo refinamento devido ao resultado insuficiente de proteínas obtido da análise de similaridade da proteína do organismo modelo, para obter mais proteínas para as futuras comparações. Através do NCBI e a ferramenta BLAST, foi realizada a pesquisa por novos organismos com possíveis proteínas HP1. Dessa forma, obtivemos um total de 108.036 resultados do banco de dados, no entanto somente 171 foram analisados, permanecendo as proteínas de organismos que não eram repetidos e que possuíam similaridade com a proteína HP1 de *Schistosoma mansoni*, levando em consideração os códigos de entrada NP/XP, por serem códigos relacionados a uma proteínas. Deste refinamento, selecionamos 47 proteínas ([Tabela 01](#)) de organismos diferentes, sendo duas (2) espécies de *Lucilia sp.* e *Rhagoletis sp.*, quatro (4) espécies de *Plasmodium sp.*, dezenove (19) espécies de *Drosophila sp.*, e *Agrilus pomonella*, *Culex quinquefasciatus*, *Folsomia candida*, *Glossina fuscipes*, *Loa loa*, *Melanaphis sacchari*, *Musca domestica*, *Neurospora crassa*, *Plutella xylostella*, *Rhagoletis sp.* *Tetranychus urticae*. No entanto, ao analisarmos estas proteínas na interface do InterPro, para identificar os domínios chromo e chromo_shadow na sequência FASTA, dois domínios que conferem a confiabilidade de uma proteína HP1 [5], onze (11) proteínas foram descartadas por não conter os dois domínios ([Tabela 02](#)). Esses resultados são extremamente relevantes para realização das comparações entre as proteínas selecionadas. As proteínas que continham ambos os domínios, tiveram cada domínio respectivamente identificado no formato FASTA e registrado em um arquivo docs. No caso dos organismos que possuíam mais de uma espécie selecionada (*Drosophila*, *Plasmodium*, *Rhagoletis* e *Lucila*), a sequência FASTA eram predominantemente iguais entre as espécies, com uma mínima diferença na localização dos domínios chromo e chromo_shadow.

Contudo, para ser melhor avaliado é necessário entender a função que cada uma dessas proteínas possuem, e para isso utilizamos o UniProt para verificar os dados funcionais das proteínas. Consequentemente, a grande maioria das proteínas não foram encontradas no banco de dados com seu código de acesso, resultando no descarte de trinta e três (33) proteínas, uma vez que não seria possível utilizá-las. Como forma de contornar essa dificuldade e ter pelo menos cinco (5) proteínas para comparar, foi feita uma revisão bibliográfica na literatura disponível para encontrar mais possíveis proteínas HP1 de outros organismos que foram estudados. Desta revisão bibliográfica foram obtidas quatro (4) proteínas de organismos diferentes: *Drosophila melanogaster*, *Homo sapiens*, *Mus musculus* e *Plasmodium falciparum*. Com estes resultados foi verificado os dados funcionais no UniProt, porém apenas a proteína de *Homo sapiens* e *Mus musculus* tinham função descrita. Para não perder as outras proteínas selecionadas uma nova revisão bibliográfica foi feita, desta vez utilizando as publicações relacionadas a proteína dos organismos: *Schistosoma mansoni*, *Neurospora crassa*, *Plasmodium falciparum*, *Drosophila melanogaster* e *Drosophila mojavensis* ([Tabela 03](#)). Por meio desta revisão, foi possível sugerir uma função para estas proteínas ([Tabela 04](#)). É importante ressaltar que, mesmo a proteína HP1 desempenhe diferentes funções, devido a sua alta conservação dos domínios chromo e chromo_shadow entre organismos, o esperado é que a função seja a mesma ou parecida [4, 6].

Diante desses resultados, foi feita uma busca para obter a estrutura tridimensional dessas proteínas no banco de dados PDB ([RCSB PDB: Homepage](#)), que irão ser comparados com a estrutura do organismo modelo. Com esta comparação será possível analisar como a proteína HP1 interage com as marcas epigenéticas H3K9me3 e H2K27me3. Estudos corroboram que a metilação da histona 3 nas lisinas 9 e 27 é uma marca epigenética que está relacionada ao silenciamento gênico, em contrapartida, a acetilação da lisina 9 é uma marca da eucromatina que facilita o processo de transcrição [2]

DISCUSSÃO

Os resultados possibilitaram aprimorar o refinamento da HP1 em diferentes organismos e através dos dados obtidos no NCBI e na revisão bibliográfica, o *e-value* nos permitiu observar a confiabilidade dos dados do alinhamento [7]. Com os domínios chromo e chromo_shadow devidamente identificados na sequência, corroborando com a confiabilidade da proteína devido a HP1 ser uma proteína da família cromobox e ter o domínio chromo, além de na cauda C-terminal ter o domínio chromo_shadow [5]. Através desse refinamento, conseguimos sugerir a função das proteínas que não tem uma função descrita no PDB com base nas publicações relacionadas a elas. Estes resultados vão possibilitar a compreensão da função dessas proteínas, como as modificações pós-traducionais afetam o organismo modelo *Schistosoma mansoni*, além de conseguirmos estudar as interações entre HP1 e as marcas epigenéticas H3K9me3 e K2K27me3.

CONCLUSÃO.

Não é possível concluir alguma hipótese acerca das interações envolvendo a HP1 como modulador epigenético em *Schistosoma mansoni*, além que a HP1 parece desempenhar funções parecidas em diferentes organismos, estando relacionada a regulação gênica e ao estado repressivo da cromatina. Após concluirmos a etapa de comparação das estruturas tridimensionais das proteínas HP1, seja possível inferir uma hipótese como essas interações ocorrem e quais as suas implicações no organismo modelo *Schistosoma mansoni*.

ANEXOS.

- Tabela 01: Proteínas selecionadas ( Proteínas selecionadas)
- Tabela 02: Possíveis proteínas HP1 ( Possíveis HP1)
- Tabela 03: Refinamento final. ( Refinamento final)
- Tabela 04: Funções ( Funções Proteínas)

REFERÊNCIAS.

- [1] COSSEAU, Céline et al. (Epi)genetic Inheritance in *Schistosoma mansoni*: A Systems Approach. **Trends in Parasitology**, [S. l.], v. 33, n. 4, p. 285–294, 2017. Disponível em: <http://dx.doi.org/10.1016/j.pt.2016.12.002>
- [2] ROQUIS, David et al. The epigenome of *Schistosoma mansoni* provides insight about how cercariae poise transcription until infection. **PLoS Neglected Tropical Diseases**, [S. l.], v. 9, n. 8, p. 1–22, 2015.
- [3] Vermaak, D., & Malik, H. S. (2009). Multiple Roles for Heterochromatin Protein 1 Genes in *Drosophila*. <https://doi.org/10.1146/Annurev-Genet-102108-134802>, 43, 467–492.
- [4] GEYER, Kathrin K. et al. Methyl-CpG-binding (SmMBD2/3) and chromobox (SmCBX) proteins are required for neoblast proliferation and oviposition in the parasitic blood fluke *Schistosoma mansoni*. **PLoS Pathogens**, [S. l.], 2018.
- [5] Lomberk, G., Wallrath, L. L., & Urrutia, R. (2006). The Heterochromatin Protein 1 family. **Genome Biology**, 7(7), 228. <https://doi.org/10.1186/GB-2006-7-7-228>

- [6] DA TRINDADE, Natália Silva et al. Schistosoma mansoni Heterochromatin Protein 1 (HP1) nuclear interactome in cercariae. **Journal of Proteomics**, [S. l.], v. 239, n. July 2020, p. 104170, 2021. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1874391921000695>
- [7] Hennell, J. R., D'Agostino, P. M., Lee, S., Khoo, C. S., & Sucher, N. J. (2012). Using genbank® for genomic authentication: A tutorial. **Methods in Molecular Biology**, **862**, 181–200. https://doi.org/10.1007/978-1-61779-609-8_15
- [8] Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. **BMC Bioinformatics**, **10**, 421. <https://doi.org/10.1186/1471-2105-10-421>