

Victor Fernandes de Oliveira

**Identificação e caracterização de microRNAs de origem
intrônica em *Schistosoma mansoni***

Ouro Preto - MG, março de 2013



UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
DEPARTAMENTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

**Identificação e caracterização de microRNAs de origem
intrônica em *Schistosoma mansoni***



AUTOR: VICTOR FERNANDES DE OLIVEIRA

ORIENTADORA: PROF^a. DR^a. RENATA GUERRA DE SÁ COTA

Ouro Preto - MG, março de 2013

Este trabalho foi desenvolvido no Laboratório de Bioquímica e Biologia Molecular NUPEB/ICEB/UFOP, com apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Dedico esta dissertação aos meus pais, Alvimar e Zenaide, aos meus irmãos, Igor e Isis e a todos meus familiares. Agradeço pelo apoio e compreensão em todos os momentos dessa etapa na minha vida.

AGRADECIMENTOS

Primeiramente a Deus;

Aos meus pais, Alvimar e Zenaide, pelo apoio, carinho e amor. Obrigado por tudo. Amo vocês!;

Aos meus irmãos, Igor e Isis, pelo amor, apoio, amizade, e companheirismo em todos os momentos dessa etapa;

A todos os outros familiares, pela preocupação e carinho;

À Prof.^a Dr.^a Renata Guerra de Sá Cota, que me deu a oportunidade de conhecer e trabalhar com a biologia molecular desde a graduação até o mestrado. Obrigado pelos ensinamentos práticos e teóricos, pelo carinho, paciência e, sobretudo pela orientação;

Aos meus grandes amigos quase que irmãos, Wagner e Fred, que conviveram comigo durante todo o período em Ouro Preto;

Ao meu grande amigo Bruno, pela amizade construída e sobretudo, pelo apoio nos momentos mais difíceis;

A todos os amigos que estão presentes ou que passaram pelo laboratório (LBBM): Leandro, Roberta V., Karina, Sávio, Natália, Matheus, Raquel, Isabel, Carrol, Nayara M., Soraya, Kelvin, Ester, Érica, Cíntia, Walmir, Fabiano, Suianne, Andréia, Roberta D., Tiago... Obrigado! Essa trajetória foi mais agradável com a ajuda de vocês;

Aos demais amigos da graduação, e do mestrado;

Aos professores doutores do laboratório Elísio, Leandro e Karen pela amizade, aprendizado e convívio agradável;

Aos técnicos Ezequiel e Eduardo pelo auxílio;

Ao professor William pela amizade e ensinamentos no trabalho;

Aos demais professores da UFOP que participaram da minha formação acadêmica;

Aos amigos do Museu de Ciência e Técnica Escola de Minas/UFOP;

E por todos aqueles que participaram de forma direta ou indireta na realizaço deste trabalho.

“A verdadeira viagem de descobrimento não consiste em procurar novas paisagens, e sim em ter novos olhos.”

Marcel Proust

Sumário

LISTA DE FIGURAS	I
LISTA DE TABELAS	V
ABREVIATURAS E SIGLAS	VI
RESUMO	VII
ABSTRACT	IX
1 - INTRODUÇÃO	1
1.1 - HISTÓRICO E EPIDEMIOLOGIA	2
1.2 - CICLO DE VIDA DO <i>S. MANSONI</i> E PATOLOGIA DA DOENÇA	4
1.3 - GENOMA E TRANSCRISSOMA DO <i>S. MANSONI</i>	7
1.4 - REGULAÇÃO DA EXPRESSÃO GÊNICA EM <i>S. MANSONI</i>	8
1.5 - MICRORNAS	11
1.6 - MIRNAS NO GÊNERO SCHISTOSOMA.....	14
2 - OBJETIVOS	16
2.1 - OBJETIVO GERAL	17
2.2 - OBJETIVOS ESPECÍFICOS	17
3 - MATERIAIS E MÉTODOS	18
3.1 - ESTABELECIMENTO DO PIPELINE PARA A IDENTIFICAÇÃO DE MIRNA INTRÔNICO EM <i>S. MANSONI</i>	19
3.1.1 - <i>Extração das sequências do genoma de S. mansoni</i>	19
3.1.2 - <i>Predição de miRNAs maduros intrônicos e seus precursores (pré-miRNAs)</i>	19
3.1.3 - <i>Busca por homologia no mirBASE</i>	20
3.1.4 - <i>Predição dos alvos de miRNAs no genoma de S. mansoni</i>	20
3.2 - VALIDAÇÕES EXPERIMENTAIS.....	21
3.2.1 - <i>Obtenção dos Parasitos</i>	21
3.2.2 - <i>Análise da expressão gênica por qRT-PCR</i>	24
4 - RESULTADOS E DISCUSSÕES	36
4.1 - IDENTIFICAÇÃO DE MIRNAS NO GENOMA DE <i>S. MANSONI</i>	37
4.2 - EXPRESSÃO DOS MIRNAS INTRÔNICOS EM <i>S. MANSONI</i>	47
4.3 - ANÁLISE DA EXPRESSÃO DOS GENES HOSPEDEIROS.	52
4.5 - ANÁLISE FUNCIONAL <i>IN SILICO</i> - INTERAÇÃO DO MIRNA COM O GENE ALVO	58
5 - CONCLUSÃO	62
6 - PERSPECTIVAS	64
7 - ANEXOS	66

ANEXO A: ESTRUTURAS SECUNDÁRIAS DOS 19 CANDIDATOS A PRÉ-MIRNA DE <i>SCHISTOSOMA MANSONI</i>	67
ANEXO B: SCRIPTS DESENHADOS NA LINGUAGEM DE PROGRAMAÇÃO PERL.	72
8 - REFERÊNCIAS BIBLIOGRÁFICAS	78
.....	

LISTA DE FIGURAS

Figura 1: Esquistossomose, países ou áreas em risco, 2011. Distribuição da esquistossomose no mundo (Adaptado de WHO, 2012).

Figura 2: Áreas endêmicas e focais da esquistossomose mansônica. Porcentagem de prevalência da doença por município (adaptado de COURA et al., 2004).

Figura 3: Ciclo de vida do parasito *S. mansoni*. No hospedeiro vertebrado, os esquistossômulos (A) originam-se em vermes adultos (B) os quais se acasalam e produzem ovos (C) liberando-os para o ambiente aquático. Os ovos desenvolvem e eclodem em formas larvais chamadas de miracídios (D), que penetram nos caramujos originando por reprodução assexuada os esporocistos (E). Eles finalmente geram as cercárias (F) que deixam o caramujo e são liberadas na água, infectando o hospedeiro vertebrado fechando o ciclo (Adaptado de VERJOVSKI-ALMEIDA et al., 2003).

Figura 4: Genes relacionados ao desenvolvimento do *S. mansoni*. Maior expressão dos genes em esporocisto (18/31) quando comparada com esquistossômulo de três dias (11/31) e cercária (8/31). Os estágios estão representados em E (Esporcocisto), C (Cercária) e em D3 (Esquistossômulos com três dias de desenvolvimento) (Adaptado de PARKER-MANUEL et al., 2011).

Figura 5: Expressão gênica em *S. mansoni*. Expressão gênica diferenciada entre os estágios baseado nos projetos transcrissoma do parasito (Adaptado de HAN et al., 2009).

Figura 6: Localização dos miRNAs no genoma. Os miRNAs podem ser classificados em relação a sua localização genômica sendo miRNAs intragênicos (quando localizado dentro dos genes) e intergênicos (quando localizados entre genes). Os intragênicos podem ainda ser subdivididos em relação à posição no gene (presente no íntron ou no éxon).

Figura 7: Biogênese dos miRNAs. Os miRNAs podem ser codificados de uma região intergênica ou de um íntron proveniente de um gene hospedeiro sendo transcritos como miRNA primário (pri-miRNA) ou como um transcrito primário. O pri-miRNA é processado pela Ribonuclease 3 (Drosha) e seu cofator DGCR8, liberando uma estrutura em grampo (pré-miRNA), que também pode resultar a partir do processamento de um transcrito primário via *splicing*. Os pré-miRNAs são exportados para o citoplasma e processados pela Dicer em cerca de 22 nucleotídeos (miRNAs maduros). O complexo RISC interage com o miRNA que se liga em vários genes alvos, mediando a repressão da tradução ou a degradação do mRNA. Embora a expressão dos genes-alvos individuais seja modestamente reprimida, a regulação simultânea de múltiplos genes que operam em uma rede de regulamentação (representado por genes alvos A-D) podem ter efeitos sinérgicos sobre as funções biológicas (Adaptado de MIYAKI & ASAHARA, 2012).

Figura 8: miRNAs em parasitos. Grupos de parasitos incluindo as espécies em que os miRNAs foram relatados até o momento (MANZANO-ROMÁN, 2012).

Figura 9: Esquistossômulos cultivados *in vitro*: Os esquistossômulos foram cultivados por 24 horas em estufa de CO₂ 5% a 37°C em meio 169 (DE SOUZA GOMES, 2008).

Figura 10: RNA total obtido a partir de estágios evolutivos do *S. mansoni*. O RNA total obtido a partir de cercárias (1), esquistossômulos com 3,5 horas (2); esquistossômulos com 24 horas (3); vermes adultos (4); miracídios (5); ovos (6).

Figura 11: Curva padrão referente ao gene *eIF4E*. No eixo X estão representados os valores de Log da concentração de cDNA e no eixo Y os valores de Cq correspondes a cada diluição realizada. O coeficiente de linearidade e de *slope* estão representados na figura. Para a realização da curva de eficiência, foram utilizadas amostras de cDNA de cercária em uma diluição seriada de 4 vezes.

Figura 12: Gráfico de amplificação referente à curva de eficiência do gene *eIF4E*. No eixo X está representado o valor dos ciclos da qRT-PCR e no eixo Y os valor de Delta Rn. Amostras de cDNA de cercaria foram utilizadas em uma diluição seriada de 4 vezes.

Figura 13: Curva de dissociação do *amplicon* referente ao gene *Smp-170800*. No eixo Y está representado a derivada do valor emitido pela fluorescência e no eixo X a temperatura de dissociação do produto gerado pela qRT-PCR.

Figura 14: Síntese do cDNA a partir do miRNA. Inicialmente ocorre uma reação de poliadenilação nas moléculas de RNA. Um primer oligo-dT fornecido pelo kit é utilizado para sintetizar a fita de cDNA. Esse cDNA é submetido ao primeiro ciclo da reação de PCR no qual o primer específico do miRNA de interesse é utilizado. Os ciclos seguintes ocorrem a partir do primer específico e de um primer universal fornecido pelo kit (Adaptado de miScript II RT Kit - QIAGEN).

Figura 15: Curva de dissociação do miRNA 18-3p. Representado na figura somente um pico de amplificação da curva de dissociação do *sma-miR-new 18-3p*.

Figura 16: Pré-miRNAs em *S. mansoni*. *Weblogo* representando a frequência de cada nucleotídeo presente nos precursores preditos de *S. mansoni* indicados pelos logos U (uracila), A (adenina), C (Citosina), G (Guanina). Figura (A) 15 nucleotídeos iniciais da extremidade 5', (B) 15 nucleotídeos finais da extremidade 3', (C) 50 nucleotídeos finais da extremidade 3'.

Figura 17: Estrutura precursora do *sma-miR-new_i12 3p*. Representado na figura a estrutura secundária de origem intrônica do gene *Smp_048030.1*, precursor dos miRNAs maduros *sma-miR-new_i12 5p* e *sma-miR-new_i12 3p*. Resultados da expressão relativa demonstram que somente a sequência madura *sma-miR-new_i12 3p* foi expresso.

Figura 18: Gel de agarose dos miRNAs. O gel de agarose 1,2% exemplifica nas canaletas de 1-3 três candidatos a miRNAs que não apresentaram produtos de PCR esperados (respectivamente *sma-new-mir_i1 5p*, *sma-new-mir_i3 5p*, *sma-new-mir_i8 5p*). As canaletas 4-8 representam miRNAs em que apresentaram somente um pico de dissociação e um produto esperado (respectivamente *sma-new-mir_i11 3p*, *sma-new-mir_i5 5p*, *sma-new-mir_i6 5p*, *sma-new-mir_i9 5p*, *sma-new-mir_i15 5p*).

Figura 19: miRNAs com menor expressão relativa em *S. mansoni*. A expressão dos miRNAs foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios). Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. Os resultados foram multiplicados por 10^2 . Estatística não representada na figura.

Figura 20: miRNAs com maior expressão relativa no *S. mansoni*. A expressão dos miRNAs foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios). Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. Os resultados foram multiplicados por 10^2 . Estatística não representada na figura.

Figura 21: miRNA 17-5p é super expresso no *S. mansoni*. A expressão do miRNA 17-5p foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios). Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * diferente de cercária, ** diferente de EMT-3,5 horas, *** diferente de EMT-24 horas, # diferente de verme adulto, δ diferente de ovos. Os resultados foram multiplicados por 10^2 .

Figura 22: Expressão relativa dos genes hospedeiros (5, 6, 7, 9) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

Figura 23: Expressão relativa dos genes hospedeiros (11, 12, 13, 15) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

Figura 24: Expressão relativa dos genes hospedeiros (16, 17, 18, 19) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi

realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

Figura 25: Correlação de Pearson entre miRNAs e os seus genes hospedeiros. O nível de expressão relativa dos pares miRNAs - Genes hospedeiros foram analisados pelo coeficiente de Pearson (r). Os miRNAs *sma-new-miR_i15-5p* e *sma-new-miR_i18-3p* apresentaram uma correlação positiva em relação aos seus genes hospedeiros.

Figura 26: Rede de interação ente o miRNA e seus alvos. A rede foi criada a partir do programa Cytoscape 3.0 entre os 18 miRNAs (símbolos de cor preta) e os seus 993 alvos mRNAs (símbolos de cor azul). Quanto mais próximo o alvo ficou do miRNA, menor foi o valor de ΔG indicando uma maior possibilidade de ligação dele com seus alvos.

LISTA DE TABELAS

Tabela 1: Meio 169 com seus componentes e concentração.

Tabela 2: Sequências dos iniciadores dos genes hospedeiros dos miRNAs.

Tabela 3: Sequências do oligonucleotídeos iniciadores referentes aos miRNAs.

Tabela 4: Pré-miRNAs identificados e respectivos genes hospedeiros em *S. mansoni*

Tabela 5: Características estruturais e termodinâmicas dos pré-miRNAs de origem intrônica de *S. mansoni*.

Tabela 6: Características dos miRNAs maduros preditos que foram expressos.

Tabela 7: Número de alvos dos miRNAs preditos e validados.

ABREVIATURAS E SIGLAS

AMFE: Energia mínima livre ajustada

DNA: Ácido desoxirribonucléico

DNase: Desoxirribonuclease

dNTP: Desoxirribonucleosídeos trifosfatados

EMT: Esquistossômulos mecanicamente transformados

ESTs: do inglês, *Expressed Sequence Tags*

FAPESP: Fundação de Amparo à Pesquisa do Estado de São Paulo

Kb: Kilobase

Mb: Mega bases

MFE: Energia livre mínima

MFEE: Energia mínima livre do conjunto termodinâmico

MFEI: Índice de energia mínima livre

miRNA: microRNA

mRNA: RNA mensageiro

nt: Nucleotídeos

pb: Pares de bases

PCR: do inglês, Polymerase Chain Reaction

PERL: do inglês, Practical Extraction and Report Language (Linguagem Prática de Extração e Geração de Relatório)

pri-miRNA: miRNA primário

qRT-PCR: do inglês, *Quantitative Reverse Transcription PCR*

RISC: do inglês, *RNA-induced silencing complex*

RNase: Ribonuclease

TIGR: Institute for Genomic Research (Instituto de Pesquisas Genômicas)

UTR: Untranslated Region (Região não traduzida)

WHO: do inglês, World Health Organization

RESUMO

microRNAs (miRNAs) são uma classe de reguladores pós-transcricionais com aproximadamente 22 nucleotídeos. Estas moléculas regulam a expressão gênica por se ligarem a sequências complementares existentes na região 3'UTR de mRNAs específicos, induzindo sua degradação ou silenciamento. Utilizando abordagens computacionais, nosso grupo de pesquisa identificou 42 novos miRNAs precursores em *Schistosoma mansoni*, sendo que 5 destes, eram de origem intrônica. Considerando que cerca de metade dos miRNAs humanos conhecidos estão localizados em íntrons de genes codificantes de proteínas e que muitos deles são expressos somente quando o gene é transcrito, levantamos a hipótese de que parte do repertório de miRNAs de *S. mansoni* poderia ser de origem intrônica. Para investigar esta hipótese foi utilizada a versão 5.1 do genoma deste parasito, bem como analisada o perfil de expressão utilizando a metodologia de qRT-PCR nos seguintes estágios evolutivos do parasito: cercárias, esquistossômulos jovens (3,5 e 24h de cultivo *in vitro*), vermes adultos, ovos e miracídios. Após a recuperação do arquivo contendo as sequências de íntrons presentes nos genes de *S. mansoni*, foram recuperadas somente as que continham entre 50 a 120 nucleotídeos e que apresentavam o mínimo de energia livre de $\Delta G < -25$ Kcal/mol, estimado pelo software RNAfold. Essas análises mostraram um conjunto de 38 candidatos a moléculas precursoras de miRNAs. Posteriormente utilizando a ferramenta *Mature Bayes* foram preditos os miRNAs maduros, e a seguir, utilizados para busca de homologia no o banco de dados miRBase. Os resultados mostraram que os miRNAs preditos não apresentam homólogos no mirBase, levantando a hipótese de que este conjunto seja específico da espécie *S. mansoni*. A análise do perfil de expressão mostrou que dos 38 miRNAs preditos, 19 apresentaram expressão em pelo menos um estágio evolutivo analisado. Não foram identificados miRNAs estágios específicos, apesar de todos serem diferencialmente expressos. Com relação ao perfil de expressão dos genes hospedeiros, também foi observado uma expressão diferencial entre os estágios analisados. Finalmente para a predição dos alvos, foi utilizado o programa miRanda, evidenciando um conjunto de 993 alvos potenciais, sugerindo que 10% dos genes preditos em *S. mansoni* poderiam ser regulados por miRNAs. Em conjunto os resultados sugerem que parte dos miRNAs de *S. mansoni* estão localizados em genes que codificam proteínas. Esta observação levanta uma série de questões interessantes que serão futuramente investigadas.

Palavras-chave: miRNAs, gene hospedeiro, alvo mRNA, *Schistosoma mansoni*, expressão gênica, co-expressão, regulação pós-transcricional.

ABSTRACT

microRNAs (miRNAs) are a class of post-transcriptional regulators with approximately 22 nucleotides. These molecules regulate gene expression by interaction to complementary sequences in the 3'UTR region of specific mRNA and inducing its degradation or silencing. Using computational approaches, our research group has identified 42 new miRNAs precursor in *Schistosoma mansoni*, and 5 of these were of intron origin. Whereas about half of the known human miRNAs are located in introns of protein coding genes and many of them are express only when the gene is transcribed, hypothesized was that part of the miRNAs repertoire in *S. mansoni* could be intron origin. To investigate this hypothesis we used version 5.1 of this parasite genome and analyzed the expression profile using qRT-PCR methodology in the following evolutionary stages of the parasite: cercariae, schistosomula young (3.5 and 24 hours of in vitro culture), adult worms, eggs and miracidium. After recovery of the file containing the sequence of introns present in the genes of *S. mansoni*, the sequences that contained between 50 and 120 nucleotides and who had at least $\Delta G < -25$ Kcal/mol free energy, estimated by RNAfold software were recovered. Together, these analyzes revealed a set of 38 miRNA precursor candidates. Subsequently, the software *Mature Bayes* was used to miRNAs mature candidates and searched for homology using the miRBase database. The results showed that predicted miRNAs have no counter parts in mirBase, suggesting that this miRNA set is *S. mansoni* species-specific. The expression profile analysis showed that the 38 predicted miRNAs, 19 were expression in at least one developmental stage analyzed. miRNAs were not identified as stage specific expression, despite all being differentially expressed. Regarding the profile of expression of host genes was also observed a differential expression between stages analyzed. Finally for the prediction of targets, we used the software miRanda, revealing a set of 993 potential targets, suggesting that 10% of the predicted genes in *S. mansoni* could be regulated by miRNAs. Together these results suggest that parts of *S. mansoni* miRNAs are located in genes that encode proteins. This observation raises a number of interesting questions that will be further investigated.

Keywords: miRNAs, host gene, mRNA target, *Schistosoma mansoni*, gene expression, co-expression, post-transcriptional regulation.

1 - Introdução

1.1 - Histórico e epidemiologia

Estima-se que a esquistossomose humana exista há pelo menos 4.000 anos, uma vez que foram encontrados ovos do gênero *Schistosoma* em múmias chinesas e egípcias, demonstrando que a infecção esteve presente em ambas às civilizações. Ela foi descrita inicialmente por Theodor Maximilian Bilharz em 1851, que identificou pela primeira vez, vermes adultos da espécie *Schistosoma haematobium* durante autópsias feitas em camponeses do Egito. Em 1902, o médico escocês Patrick Manson (MANSON, 1902) verificou diferenças morfológicas nos ovos de *Schistosomas* africanos, indicando a presença de duas espécies distintas, a já identificada *S. haematobium* e outra desconhecida. Posteriormente em 1907 essa espécie foi finalmente descrita o por Louis Westenra Sambon (SAMBON, 1907), e nomeado em homenagem a Manson de *Schistosoma mansoni*.

A esquistossomose é a segunda doença tropical de maior impacto socioeconômico no mundo, superada somente pela malária e afetando principalmente comunidades pobres, sem água potável e saneamento básico adequado (WHO, 2008). Ela é causada por parasitos do gênero *Schistosoma*, sendo as principais espécies conhecidas que podem infectar o homem: *S. mansoni*, *S. japonicum*, *S. hematobium*, *S. intercalatum*, *S. mekongi*, *S. guineensis*. Atinge cerca de 240 milhões de pessoas em todo o mundo, e mais de 700 milhões estão vivendo em áreas endêmicas (WHO, 2012). Sua distribuição é prevalente em áreas tropicais e subtropicais, incluindo a África, América Central, Sudeste da Ásia, China e Brasil (**Figura 1**).

Dentre as espécies citadas acima, somente o *S. mansoni* ocorre no Brasil causando a esquistossomose mansônica humana, que também é conhecida popularmente como “xistose” ou “barriga d’água”. Acredita-se que a introdução desse parasito no país, tenha ocorrido no período colonial com a chegada de escravos africanos que foram obrigados a trabalhar na lavoura canavieira do nordeste brasileiro. Devido às características da região do país, como condições ambientais propícias, falta de saneamento, possível presença de moluscos do gênero *Biomphalaria*, hospedeiros intermediários e hospedeiros definitivos como o homem, possibilitou assim a instalação da esquistossomose no Nordeste brasileiro (BARRETO, 1982).

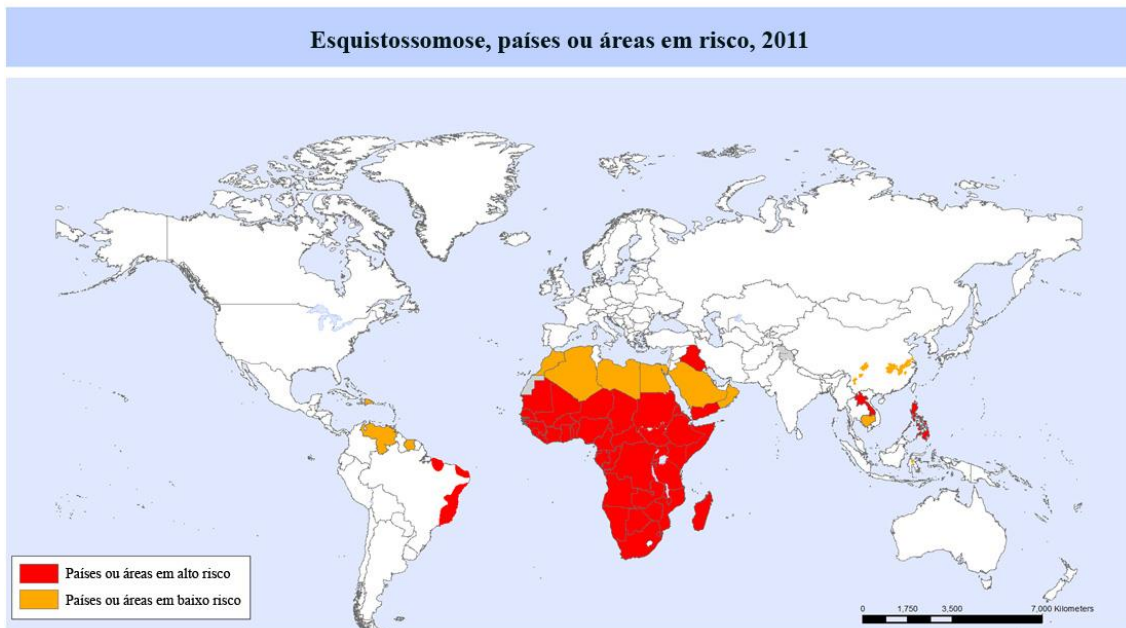


Figura 1: Esquistossomose, países ou áreas em risco, 2011. Distribuição da esquistossomose no mundo (Adaptado de WHO, 2012).

Em 1908, o médico brasileiro Pirajá da Silva, registrou o primeiro caso de infecção no Brasil, mas foi somente em 1950 que Pellon & Teixeira evidenciaram através de um inquérito nacional, a existência da esquistossomose mansônica em 612 das 877 localidades pesquisadas na região Nordeste, no Estado de Minas Gerais e posteriormente em mais seis estados. (FREITAS, 1972).

Os últimos dados demonstram que a doença é encontrada em 18 estados e no Distrito Federal (**Figura 2**), dentre os quais destacam-se: Alagoas, Bahia, Pernambuco, Rio Grande do Norte (faixa litorânea), Paraíba, Sergipe, Espírito Santo e Minas Gerais. No Pará, Maranhão, Piauí, Ceará, Rio de Janeiro, São Paulo, Santa Catarina, Paraná, Rio Grande do Sul, Goiás e no Distrito Federal, a transmissão é predominantemente focal, não atingindo grandes áreas (BRASIL, 2009). Os dados mais recentes do Ministério da Saúde do Brasil sobre a doença indicam que cerca de seis milhões de indivíduos estão infectados e 25 milhões expostos ao risco de contaminação (BRASIL, 2009).

infectante denominada miracídio (NEVES, 2005). Essas larvas ciliadas conseguem nadar ativamente no ambiente aquático até serem atraídas para o seu hospedeiro intermediário invertebrado (BRASIO, 1985). Elas também possuem glândulas adesivas e um conjunto de espículas que são importantes na penetração na homocela do caramujo. Uma vez dentro do tegumento do molusco, os miracídios sofrem alterações perdendo parte de suas estruturas transformando-se em um aglomerado de células germinativas chamadas de esporocisto. Essa é a fase denominada assexuada em que essas células (esporocistos primários) passam por processos de multiplicação, dobrando o seu tamanho. Dentro do esporocisto primário, numerosos esporocistos secundários, desenvolvem-se por multiplicação assexuada dando origem à terceira geração de larvas, denominadas cercárias (NEVES, 2005).

As cercárias abandonam os caramujos devido a estímulos externos, como temperatura e luminosidade sendo esta a que exerce um maior papel na manutenção do ritmo de liberação (LUTZ, 1919). Elas nadam ativamente por movimentação da cauda e permanecem agrupadas em águas rasas, onde podem penetrar em vários vertebrados, mas desenvolvendo-se somente no seu hospedeiro definitivo correto. Em contato com a pele humana, geralmente se fixam entre os folículos pilosos com auxílio de duas ventosas e de uma substância muco protéica secretada por suas glândulas acetabulares. Com ajuda da ação mecânica (vibração da cauda), promovem a penetração do corpo cercariano e perdem a cauda (McKERROW & SALTER, 2002).

Algumas horas após a infecção as cercárias começam a adaptar as condições fisiológicas do hospedeiro transformando-se em esquistossômulos. Por volta de 24 a 72 horas, elas migram para o tecido subcutâneo podendo atingir a circulação venosa sendo então levadas passivamente ao pulmão. Nesse trajeto os esquistossômulos sofrem um alongamento do seu corpo, mas não ocorrendo divisão celular (CLEGG, 1965). Após alguns dias, os vermes migram para o sistema porta-hepático, onde se alimentam e desenvolvem atingindo a fase adulta de macho/fêmea por volta de 25-28 dias após a penetração. Ao se acasalarem, a fêmea começa o seu processo de ovoposição no qual libera de 300 a 1000 ovos por dia (NEVES, 2005). Muitos desses ovos permanecem na circulação atingindo órgãos vitais, sendo os demais seguindo no intestino até serem eliminados com as fezes fechando o ciclo (WILSON, 1980) (**Figura 3**).

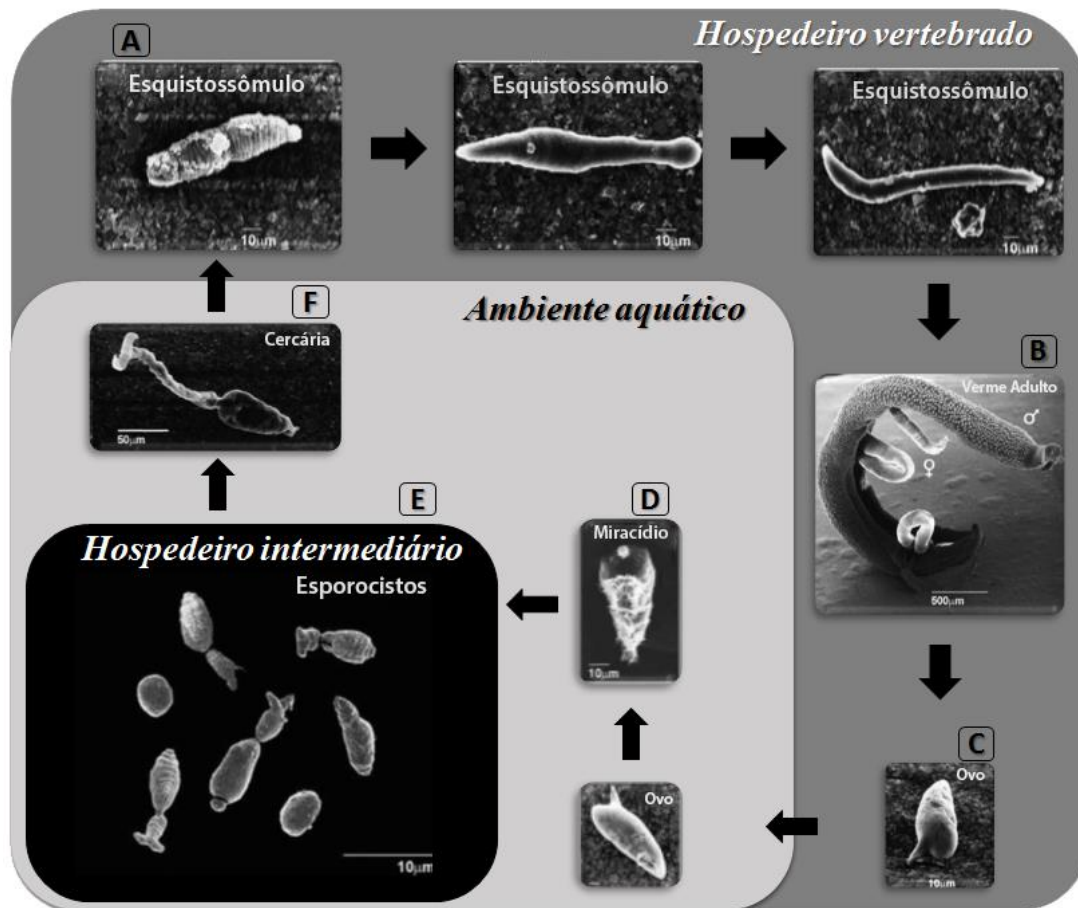


Figura 3: Ciclo de vida do parasito *S. mansoni*. No hospedeiro vertebrado, os esquistossômulos (A) originam-se em vermes adultos (B) os quais se acasalam e produzem ovos (C) liberando-os para o ambiente aquático. Os ovos desenvolvem e eclodem em formas larvais chamadas de miracídios (D), que penetram nos caramujos originando por reprodução assexuada os esporocistos (E). Eles finalmente geram as cercárias (F) que deixam o caramujo e são liberadas na água, infectando o hospedeiro vertebrado fechando o ciclo (Adaptado de VERJOVSKI-ALMEIDA et al., 2003).

Vários fatores estão relacionados à patogenicidade da esquistossomose, dentre eles a carga parasitária, idade, estado nutricional e a resposta imunitária do indivíduo infectado. Dos estágios que interagem com o hospedeiro definitivo, o principal responsável pela patogenicidade são os ovos. Quando estes permanecem no indivíduo, eles podem ficar retidos nas paredes de órgãos como o fígado e baço, formando granulomas que originam a fibrose. Esses granulomas são formados inicialmente em resposta aos antígenos sintetizados pelo miracídio em desenvolvimento dentro do ovo. O excesso desses antígenos leva posteriormente à migração de neutrófilos, macrófagos e eosinófilos para a região afetada (BOROS, 1989; MELO & COELHO, 2005). Em casos mais graves, ocorre a formação de ascite e hepatoesplenomegalia, podendo levar o indivíduo ao óbito.

1.3 - Genoma e Transcrissoma do *S. mansoni*

O estudo genômico do *S. mansoni* iniciou em 1982 a partir de experimentos feitos por Simpson, no qual estimou que este organismo possuísse um genoma com aproximadamente 270 mega bases (Mb) (SIMPSON et al., 1982). Ele ainda foi agrupado em oito pares de cromossomos, sendo divididos em sete autossômicos e um sexual. O parasito fêmea nesse caso foi identificado como heterogamético (ZW) e o macho homogamético (ZZ) (SHORT & MENZEL, 1960; SHORT et al., 1979).

Devido a iniciativas realizadas pelo Programa de Pesquisas em Doenças Tropicais da Organização Mundial de Saúde (OMS), foi estabelecido a criação do Projeto Genoma do *S. mansoni*. Ele teve início em 1992 no qual foram identificados e caracterizados novos genes do parasito. Até o ano de 2002, aproximadamente 12.000 *ESTs* (*Expressed Sequence Tags*) foram geradas pelas bibliotecas de cDNA construídas a partir de estágios distintos do *S. mansoni*. (FRANCO et al., 1995; FRANCO et al., 1997).

No ano de 2001 foi iniciado um projeto em larga escala chamado “*Schistosoma mansoni EST Genome Project*” no qual participaram vários laboratórios de pesquisa em São Paulo (Rede ONSA). Este projeto foi financiado pela FAPESP, e seus dados foram publicados no final do ano de 2003 (VERJOVSKI-ALMEIDA et al, 2003). Ele gerou um total de 180.000 *ESTs* obtidas de cada fase do ciclo de vida do parasita, incluindo cercárias, esquistossômulos, vermes adultos, ovos, miracídios e esporocistos. Eles foram ainda agrupados em 30.000 *contigs* (conjuntos não redundantes de sequencias expressas), estimando a existência de 14.000 genes no genoma.

Finalmente em 2009, *The Institute for Genomic Research (TIGR)*, e o *Wellcome Trust Sanger Institute* em parceria com outros laboratórios no mundo, finalizaram a montagem do genoma, disponibilizando-o no banco de dados GeneDB (<http://www.genedb.org>) (BERRIMAN et al., 2009).

Apesar disso, o genoma do *S. mansoni* atualmente ainda não é considerado fechado. Isto se deve a característica desse genoma constituir de 40% de elementos repetitivos sendo fragmentado em *scaffolds* (BERRIMAN et al., 2009). Na sua última anotação, ele apresentou 380Mb sendo dividido em 885 *scaffolds* (PROTASIO et al., 2012).

O número total de genes preditos nesse genoma vem sendo extensivamente modificados ao longo dos anos, sendo que esse número passou de 11.807 para 10.852 genes (PROTASIO et al., 2012). Dentre os motivos para essas mudanças, verificaram-se eventos de trans-splicing que foram identificados em pelo menos 11% dos genes preditos e possíveis erros de sequenciamento e montagem dessas sequências.

O genoma do *S. mansoni* apresenta ainda aspectos muito peculiares como a presença de regiões intrônicas grandes, o que não é característico de parasitos eucariotas primitivos. Esses íntrons apresentam em média 1.692 pb, sendo que o maior já encontrado no genoma foi de 33.8 Kb (BERRIMAN et al., 2009). Outras estruturas que não são características desse grupo, mas que estão presentes no *S. mansoni* são as MEGs (Micro Exons Genes), que são genes que apresentam pequenos exons (6-36pb) que aumentam a variabilidade dos seus transcritos por *splicing* alternativo (BERRIMAN et al., 2009).

Esses conjuntos de dados do genoma abriram assim, novas perspectivas para a utilização de técnicas moleculares visando o entendimento da biologia do parasito. Os genes descritos podem ainda ser utilizados para identificação de alvos para novas drogas contra a doença (VERJOVSKI-ALMEIDA et al, 2003; OLIVEIRA et al., 2008).

1.4 - Regulação da expressão gênica em *S. mansoni*

Como os demais eucariotos multicelulares, o conjunto de genes expressos em *S. mansoni* é influenciado por vários fatores, tais como: os nutricionais e ambientais. Particularmente, no caso do *S. mansoni*, fatores do hospedeiro (tanto vertebrado como invertebrado) também estão envolvidos na regulação da expressão gênica. Desta forma, o transcissoma do parasito é regulado em diferentes níveis; transcricional, pós-transcricional, ao nível da tradução e pós-traducionais. Dentre esses processos, a regulação pós-transcricional vem adquirindo uma atenção maior devido a crescentes estudos na área demonstrando a sua importância na regulação gênica (WEAVER, 2007). Essa regulação envolve principalmente interações entre RNAs e proteínas nas quais levam ao silenciamento gênico ou degradação dos transcritos. Dentre as modificações pós-transcricionais, podemos citar a adição de cauda poli-A, adição de 5`CAP, *splicing*, *trans-splicing*, silenciamento gênico por pequenos RNAs (como microRNAs) entre outros (HELM, 2005; KEENE, 2007; ANDERSON et al., 2009).

Acredita-se que durante o seu ciclo de vida, o *S. mansoni* expressa um conjunto de genes essenciais para viabilizar a transição entre as diferentes fases evolutivas o que lhe permite uma adaptação rápida aos diferentes ambientes e hospedeiros. Trabalhos recentes que utilizam técnicas para avaliação em larga escala da expressão gênica, como microarranjos e sequenciamento de RNA, evidenciaram um padrão de expressão diferencial em diversos genes durante as fases de cercária, esquistossômulos e esporocistos (PARKER-MANUEL et al., 2011 e HAN et al., 2009). Foi demonstrado, por exemplo, que dos 31 genes envolvidos com o desenvolvimento do parasito, 18 foram mais expressos em esporocistos do que em cercárias e esquistossômulos com três dias de desenvolvimento *in vitro* (PARKER-MANUEL et al., 2011- Figura 4).

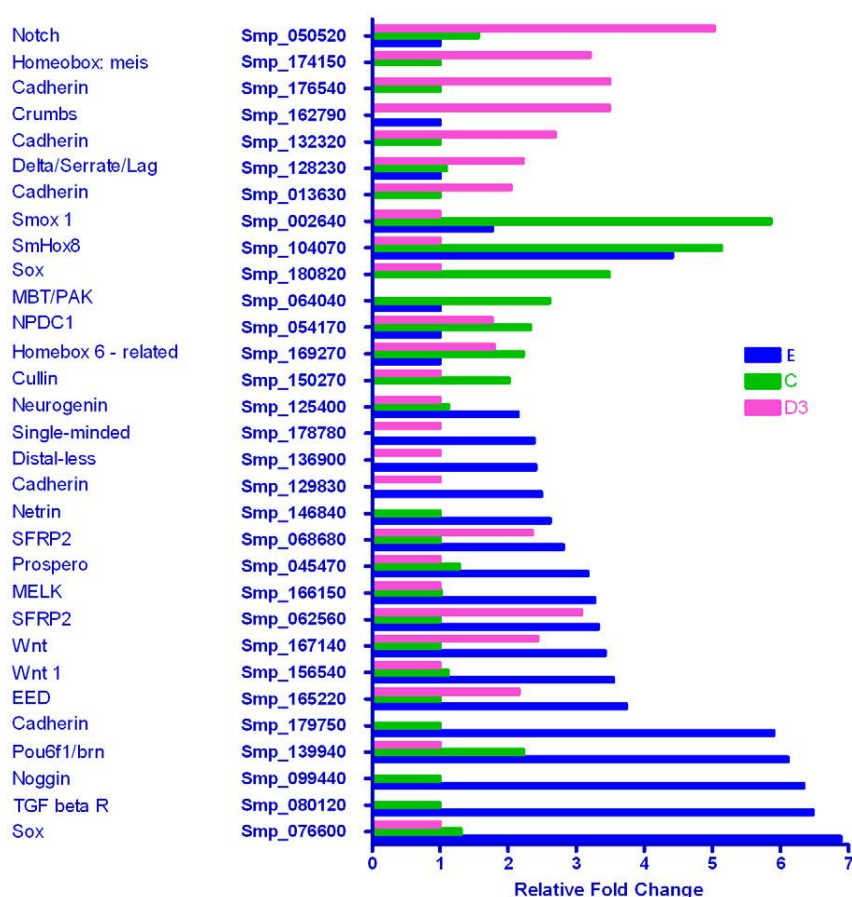


Figura 4: Genes relacionados ao desenvolvimento do *S. mansoni*. Maior expressão dos genes em esporocisto (18/31) quando comparada com esquistossômulo de três dias (11/31) e cercária (8/31). Os estágios estão representados em E (Esporcisto), C (Cercária) e em D3 (Esquistossômulos com três dias de desenvolvimento) (Adaptado de PARKER-MANUEL et al., 2011).

Esses dados enfatizam que a diferenciação e o desenvolvimento entre os estágios seja dependente da regulação da expressão gênica na qual está diretamente relacionada à composição de suas proteínas sintetizadas (HAN et al., 2009) (**Figura 5**).

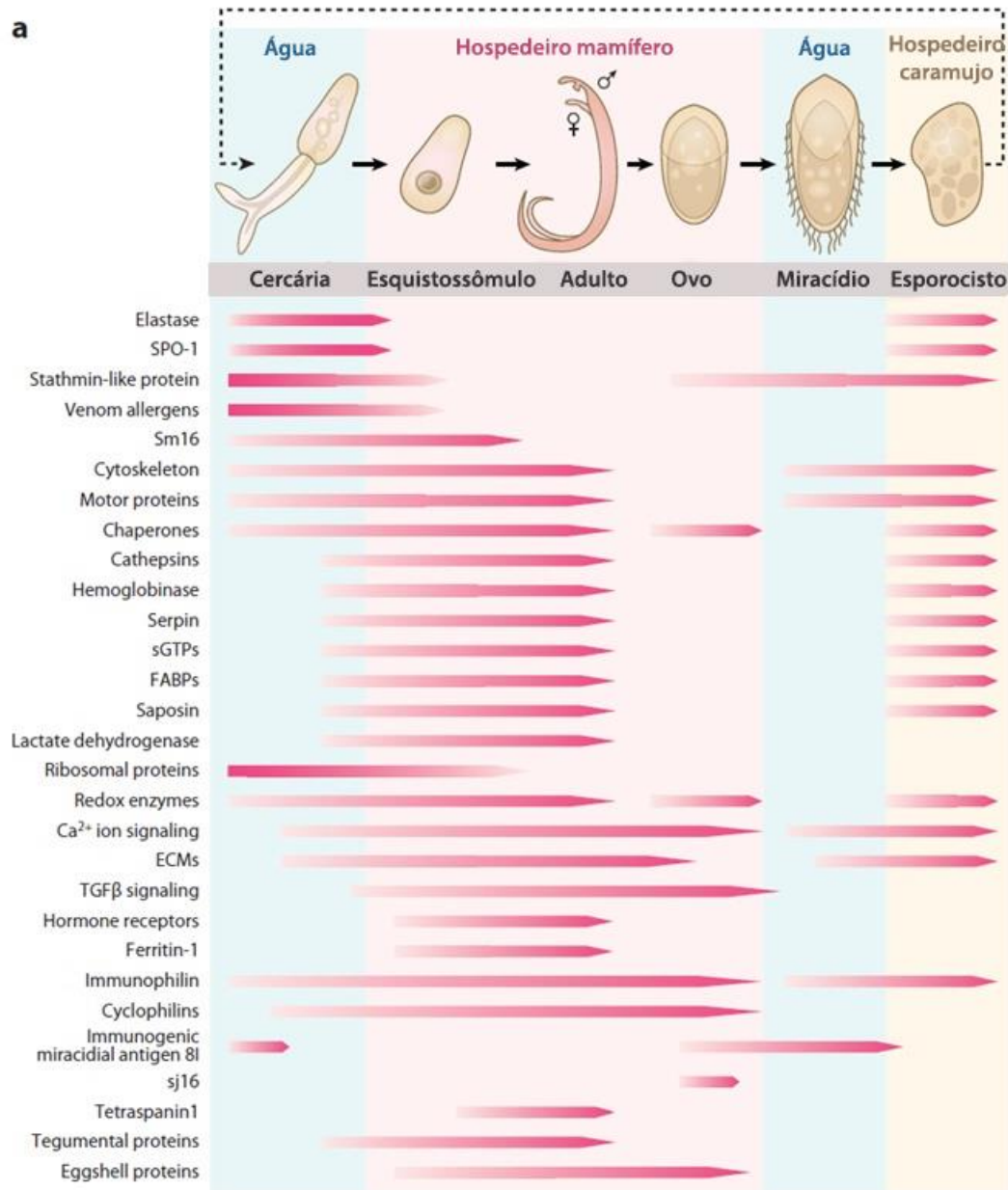


Figura 5: Expressão gênica em *S. mansoni*. Expressão gênica diferenciada entre os estágios baseado nos projetos transcrissoma do parasito (Adaptado de HAN et al., 2009).

Essa integração da informação genômica, juntamente com o transcrissoma e proteoma, proporcionou uma visão ampla sobre a biologia e patogenia do parasito, além de elucidar muitas das interações parasita-hospedeiro. Sendo assim a manipulação desse

genoma estabelece uma base para o desenvolvimento de novas vacinas, alvos de drogas, e marcadores de diagnóstico para o tratamento e controle da esquistossomose.

1.5 - MicroRNAs

Eles foram descobertos inicialmente em 1993 por um estudo realizado por Victor Ambros e colaboradores, no qual identificaram dois pequenos transcritos de RNA (22 e 61 nt) em *Caenorhabditis elegans*. Esses transcritos eram provenientes de um gene *lin-4* relacionado ao desenvolvimento do *C. elegans*, e apresentavam uma região complementar a sequência 3'UTR do gene *lin-14*. Essa observação sugeriu que esses pequenos RNAs pudessem regular a tradução do gene *lin-14* através de uma interação RNA-RNA caracterizando assim como um novo mecanismo de regulação da expressão gênica (LEE et al., 1993). Anos depois, Pasquinelli e colaboradores identificaram um pequeno RNA não-codificante, denominado de *let-7* em várias espécies incluindo o *Homo sapiens*, sugerindo que essa molécula também poderia fazer parte dessa classe de pequenos RNAs (PASQUINELLI et al., 2000). Posteriormente essas moléculas foram então denominadas de microRNAs (miRNAs) (LAGOS-QUINTANA et al., 2001).

Os miRNAs são definidos como moléculas endógenas de RNA com aproximadamente 22 nucleotídeos capazes de atuar como reguladores pós-transcricionais. Eles apresentam características em comum por serem processados a partir de um precursor em forma de grampo. Além disso, conseguem se ligar a sequências complementares na região 3'UTR de um determinado mRNA resultando no seu silenciamento ou na sua degradação. Os miRNAs apresentam diferenças em relação a sua localização e biogênese, sendo classificados em miRNAs intergênicos e intragênicos (LEE & AMBROS, 2001). Os intragênicos podem ainda ser subdivididos em relação à posição no gene estando presentes nos íntrons ou éxons, como representado na **Figura 6**.

Estudos demonstram que os miRNAs possuem importantes atividades biológicas regulando aproximadamente 20 a 30% de todos os genes humanos, sendo que um único miRNA pode ser capaz regular vários genes em um organismo (LOA et al., 2007, AMBROS, 2004; PLASTERK, 2006).

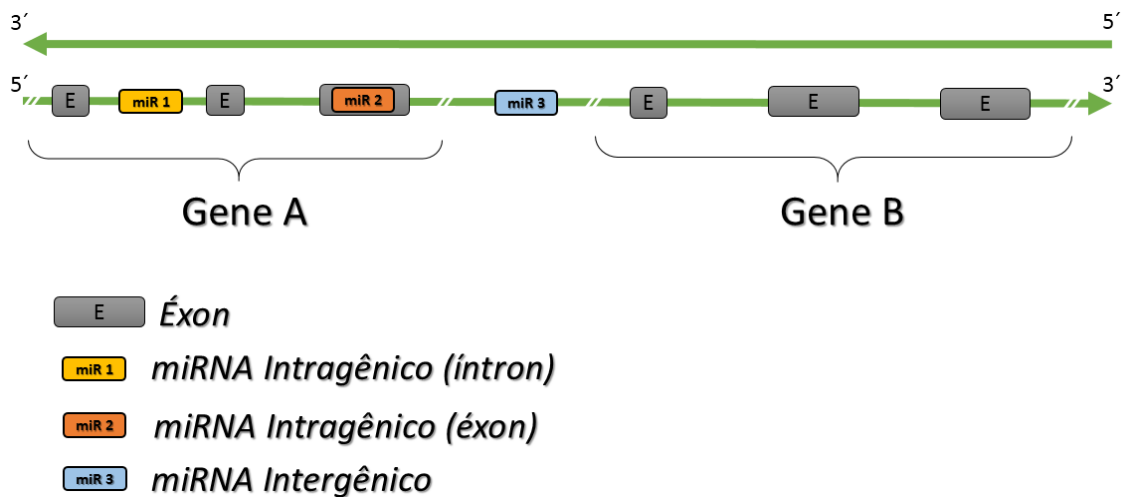


Figura 6: Localização dos miRNAs no genoma. Os miRNAs podem ser classificados em relação a sua localização genômica sendo miRNAs intragênicos (quando localizado dentro dos genes) e intergênicos (quando localizados entre genes). Os intragênicos podem ainda ser subdivididos em relação à posição no gene (presente no íntron ou no éxon).

A biogênese dos miRNAs (**Figura 7**) atualmente é amplamente conhecida, sendo classificada por muitos autores como via canônica e a via não-canônica (intrônica). Inicialmente, os miRNAs são transcritos pela RNA polimerase II ou III no núcleo, originando miRNAs primários (pri-miRNAs). Estes são posteriormente clivados pela enzima RNase III endonuclease também chamada de Drosha e seu cofator DGCR8 formando uma estrutura precursora de aproximadamente 60-120 nucleotídeos nomeada pré-miRNA. (LUND et al., 2004; LEE et al., 2003).

Esse pré-miRNA gerado é exportado para o citoplasma de forma ativa pelo complexo Ran-GTP/Exportina-5, sendo processado novamente por outra enzima da classe RNase III, Dicer, resultando em um miRNA duplex, contendo aproximadamente 22 nucleotídeos. Na maioria dos casos, somente uma das fitas desse duplex 5'-3' e 3'-5' é ativa sendo a outra degradada. Após essa clivagem, o miRNA maduro interage com o complexo de indução do silenciamento de RNA (RISC) formando o complexo miRNA-RISC no qual desempenha a função de regulação do RNA mensageiro (mRNA alvo) (BARTEL, 2004). Quando essa regulação ocorre em animais, ela geralmente atua reprimindo a tradução dos genes por apresentar um pareamento incompleto na região 3'UTR (LEWIS et al., 2005).

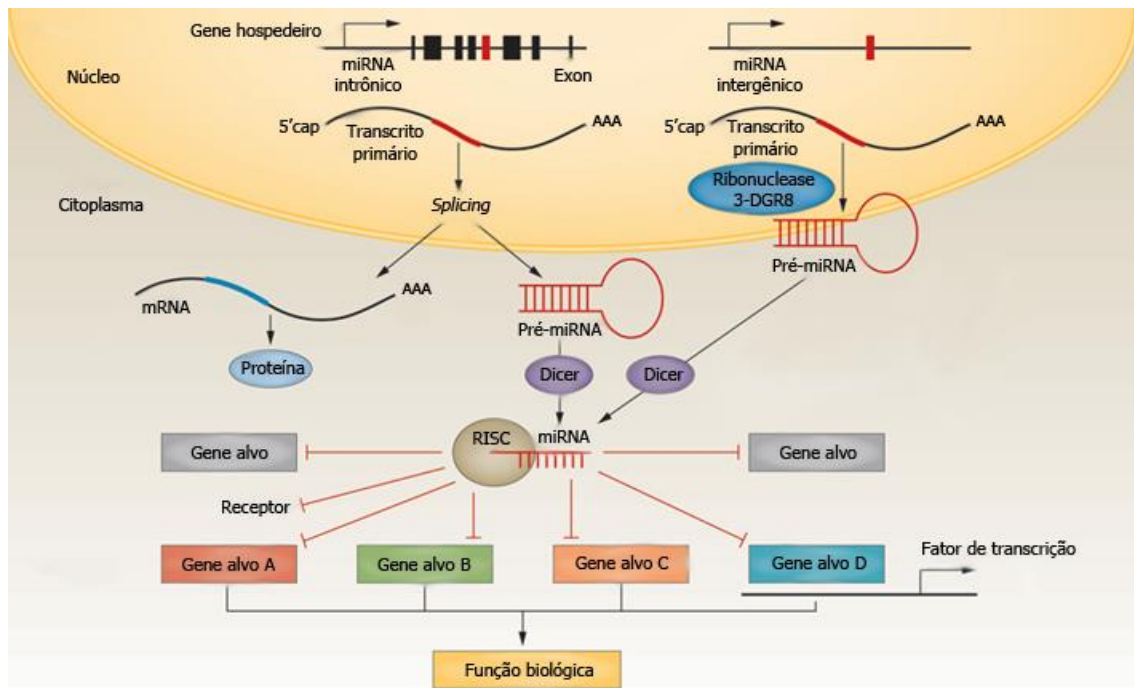


Figura 7: Biogênese dos miRNAs. Os miRNAs podem ser codificados de uma região intergênica ou de um íntron proveniente de um gene hospedeiro sendo transcritos como miRNA primário (pri-miRNA) ou como um transcrito primário. O pri-miRNA é processado pela Ribonuclease 3 (Drosha) e seu cofator DGCR8, liberando uma estrutura em grampo (pré-miRNA), que também pode resultar a partir do processamento de um transcrito primário via *splicing*. Os pré-miRNAs são exportados para o citoplasma e processados pela Dicer em cerca de 22 nucleotídeos (miRNAs maduros). O complexo RISC interage com o miRNA que se liga em vários genes alvos, mediando a repressão da tradução ou a degradação do mRNA. Embora a expressão dos genes-alvos individuais seja modestamente reprimida, a regulação simultânea de múltiplos genes que operam em uma rede de regulamentação (representado por genes alvos A-D) podem ter efeitos sinérgicos sobre as funções biológicas (Adaptado de MIYAKI e ASAHARA, 2012).

Na biogênese dos miRNAs de origem intrônica ou “mirtrons” (também representado na **Figura 7**), o processamento inicia-se pelo mecanismo de *splicing* no íntron de um transcrito primário. Estudos demonstraram que esses miRNAs intrônicos parecem ignorar a enzima Drosha na qual é essencial para a biogênese dos miRNAs (RUBY et al., 2007a). Uma alternativa a esse mecanismo, é que esse íntron seja removido pela maquinaria de *splicing* sendo posteriormente processado por uma enzima desramificadora formando um pré-miRNA. Esse pré-miRNA a partir desse ponto, segue os mesmos caminhos da biogênese na via canônica (WESTHOLM & LAI, 2001; RUBY et al., 2007a; OKAMURA et al., 2007).

1.6 - miRNAs no gênero *Schistosoma*

Os primeiros trabalhos de identificação de miRNAs no gênero *Schistosoma* iniciaram no *S. japonicum* e posteriormente em *S. mansoni* (XUE et al., 2008; COPELAND et al., 2009). Nesses estudos, foram identificados os primeiros miRNAs conservados e não conservados do gênero. Dentre eles podemos citar os que foram altamente conservados em *S. japonicum* como, por exemplo, o let-7, miR-71, bantam e miR-125 (XUE et al., 2008).

Para a validação e identificação desses miRNAs, foram associadas estratégias computacionais e experimentais. Essas estratégias seguiram a partir da predição dos miRNAs no genoma utilizando análises computacionais, e posteriormente a sua validação experimental. As análises *in silico*, aumentam ainda mais a eficácia da predição, uma vez que baseiam-se em características específicas dessas moléculas utilizando diversos programas de predição, filtros e homologia abordando identidade e similaridade entre as espécies (LI et al., 2010a). Para as análises experimentais, as principais técnicas para validação dessas sequências foram utilizadas podendo citar; sequenciamento, Northern blot e qRT-PCR.

Apesar de até o momento somente 20 miRNAs de *S. mansoni* estarem depositados no miRBASE, alguns estudos publicados identificaram muitos outros candidatos, como por exemplo; 43 miRNAs incluindo conservados e não conservados (DE SOUZA GOMES et al., 2011), 211 miRNAs (SIMOES et al., 2011). Esses trabalhos ainda demonstraram a importância desses novos dados, facilitando a compreensão da biologia desses parasitos, principalmente no desenvolvimento ao longo do ciclo e na instalação definitiva no hospedeiro pelos vermes adultos.

O número de trabalhos tendo miRNAs como foco em parasitos vem crescendo a cada dia, principalmente em experimentos nos quais utilizam a tecnologia de sequenciamento em larga escala. Novas sequências estão sendo constantemente adicionadas ao principal banco de dados de miRNA, o miRBASE (www.mirbase.org). A **Figura 8** representa os grupos de parasitos nos quais já foram identificados miRNAs até a data de agosto de 2012.

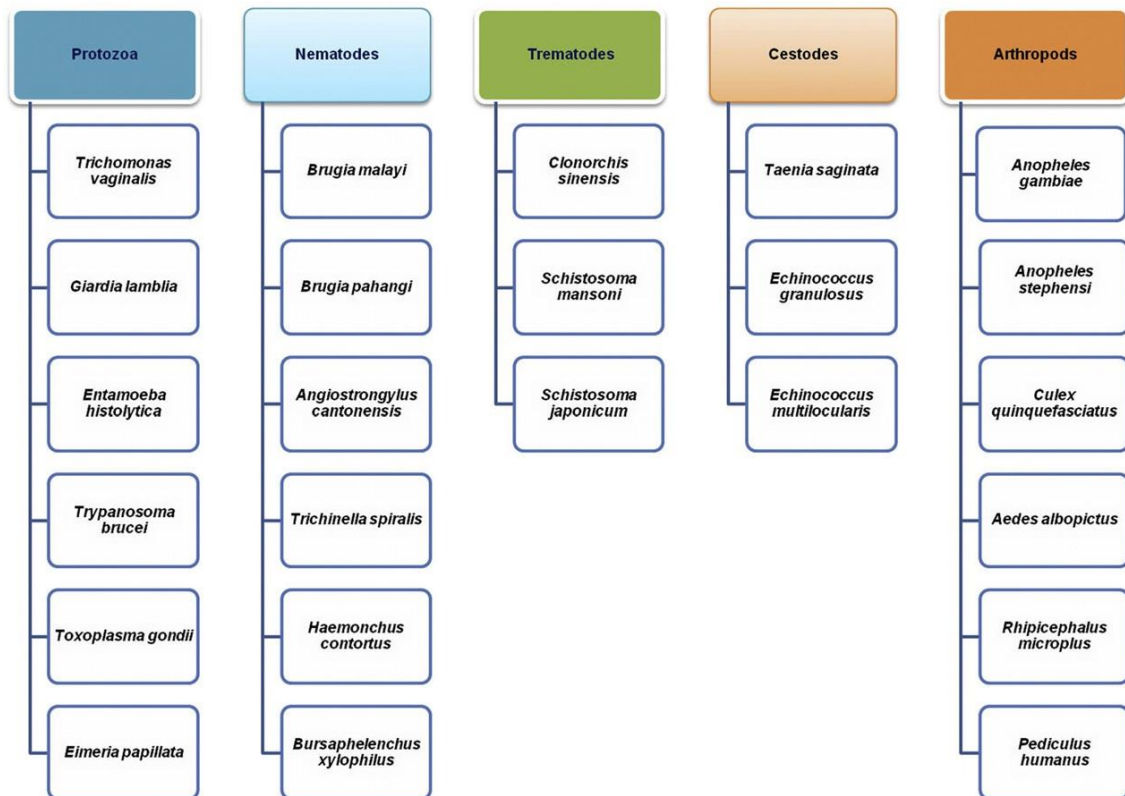


Figura 8: miRNAs em parasitos. Grupos de parasitos incluindo as espécies em que os miRNAs foram relatados até o momento (MANZANO-ROMÁN, 2012).

Durante o desenvolvimento da tese “miRNAs em *Schistosoma mansoni*: biogênese, predição, validação e alvos potenciais” (DE SOUZA GOMES et al., 2011), foi observado que vários miRNAs de *S. mansoni* apresentavam origem intrônica, incluindo um bem conservado, o *miR-190*. Continuando esta linha de investigação, nesse trabalho buscou-se identificar miRNAs de origem intrônica no *S. mansoni*, utilizando abordagens computacionais e experimentais com a finalidade de serem utilizados em futuros trabalhos como potenciais biomarcadores para o monitoramento do desenvolvimento do parasito e da evolução da patologia da esquistossomose mansoniana.

2 - Objetivos

2.1 - Objetivo Geral

Identificar e validar miRNAs de origem intrônica no desenvolvimento do parasito *S. mansoni* utilizando abordagens computacionais e experimentais.

2.2 - Objetivos Específicos

- Estabelecer um *pipeline* para busca e caracterização de miRNAs intrônicos e seus precursores utilizando dados de genoma e transcrissoma de *S. mansoni*;
Avaliar o perfil de expressão dos miRNAs identificados em diferentes fases de desenvolvimento de *S. mansoni*;
- Identificar e analisar o perfil de expressão dos genes hospedeiros dos miRNAs intrônicos validados de *S. mansoni*;
- Predizer os principais alvos de miRNAs no genoma de *S. mansoni*;

3 - Materiais e Métodos

3.1 - Estabelecimento do pipeline para a identificação de miRNA intrônico em *S. mansoni*

3.1.1 - Extração das sequências do genoma de *S. mansoni*

A partir da versão 5.1 do genoma de *S. mansoni* (Genome Assembly v5. 1), foram recuperados os arquivos *v4.19.05.11.chado.filtered.gff* e *sma_V5.1.chr.fasta* presentes no banco de dados GeneDB (<http://www.genedb.org>) (BERRIMAN e cols., 2009). Eles foram posteriormente utilizados na criação de um *pipeline* e para análise de dados.

3.1.2 - Predição de miRNAs maduros intrônicos e seus precursores (pré-miRNAs)

Para busca de novos candidatos a miRNAs de origem intrônica e seus precursores no genoma do *S. mansoni*, foi desenvolvido um *pipeline* no qual foi otimizado com diferentes filtros, utilizando diversos *scripts* na linguagem de programação Perl sendo descrito abaixo.

Inicialmente os arquivos GFF e FASTA obtidos no banco de dados, foram formatados para que gerassem um maior acesso de leitura. Para busca de todos os íntrons no genoma, foi executado o *script* “*recintron-new.pl*” gerando o arquivo de saída *recintron.out.new*, que continha somente as sequências entre exons positivos e negativos. Ambas foram separadas pelos *scripts* “*selecionar-minus.pl*” e “*selecionar-plus.pl*” e salvas em arquivos separados, *recintron_new.tab.plus.fasta* e *recintron_new.tab.minus.fasta*, contendo apenas sequências positivas e negativas, respectivamente. A partir disso foi feita a conversão do formato FASTA para o formato TABULAR, sendo isso possível através da plataforma Galaxy (<http://main.g2.bx.psu.edu/>). Essa etapa foi realizada para aperfeiçoar as análises posteriores, uma vez que o formato tabular gera uma ou mais colunas de textos separados por tabulações. Somente com as sequências negativas foi necessário fazer o complemento reverso da fita, rodando a linha de comando “*seqret-srevrecintron_new.tab.minus.fasta_recintron_new.tab.minus_reverso.fasta.plus*”. Feito isso, os dois arquivos, *recintron_new.tab.plus.fasta* e *recintron_new.tab.minus_reverso.fasta.plus*, foram novamente salvos em um único arquivo fasta *recintron_new_final*.

O primeiro filtro aplicado foi selecionar somente as sequências que possuíam entre 50 a 120 nucleotídeos (LUND et al., 2004) utilizando o script “filtra_tamanho_50_120.pl”, no qual gerou o arquivo de saída *seq_filtradas50_120.out*. A partir dele executamos o “fasta_semql.pl”, para obter um arquivo fasta sem quebra de linha.

A segunda etapa foi analisar a estrutura secundária dos pré-miRNAs e eliminar as sequências que não apresentavam potencial para ser um precursor de miRNA intrônico. A principal característica analisada foi o valor da energia mínima livre (MFE) das moléculas, obtido pelo programa RNAfold utilizando os seguintes parâmetros: “-p -d2 -noLP” (HOFACKER, 2009). Por meio do script “filtra_estrut_miRNA.pl” foi selecionado somente as sequências com $MFE = \Delta G < -25$ Kcal/mol que é considerada uma energia baixa, para que o precursor possa ser estável no sistema biológico, gerando assim miRNAs maduros (XUE et. al., 2008).

Posteriormente foi utilizado a ferramenta *MatureBayes* (<http://mirna.imbb.forth.gr/MatureBayes.html>) para a busca de sequências de miRNAs maduros nos pré-miRNAs preditos (GKIRTZOU et al., 2010).

3.1.3 - Busca por homologia no mirBASE

As sequências precursoras e maduras preditas foram comparadas com todas as sequências presentes no banco de dados do mirBASE utilizando a ferramenta Blastn. A última versão do bando de dados (Release 19,) apresentou 21264 precursores expressando 25141 miRNA maduros de 193 espécies (<ftp://mirbase.org/pub/mirbase/>).

3.3.4 - Predição dos alvos de miRNAs no genoma de *S. mansoni*

A predição computacional dos genes alvos dos miRNAs validados foram realizadas utilizando as sequências dos miRNAs maduros, e as regiões 3`UTRs (Regiões da extremidade 3` da fita de RNA mensageiro não traduzidas em proteína) do genoma do *S. mansoni*.

A partir dos arquivos GFF e GFF3 depositados no banco de dados GeneDB (www.genedb.org) foram recuperadas as sequências 3`UTRs dos mRNAs preditos de *S. mansoni*, utilizando as informações posicionais do códon de parada do mRNA e da

região não traduzida da molécula de RNA. As sequências 3'UTRs encontradas na fita antisense do DNA foram invertidas, a fim de se obter a fita correspondente ao pareamento correto com o miRNA maduro. Esta etapa foi realizada utilizando a ferramenta do pacote EMBOSS “seqret”. Todos os arquivos dessas sequências foram editados para o formato FASTA e analisados posteriormente.

Finalmente para a predição dos alvos, o programa miRanda foi utilizado com os parâmetros ajustados baseados em identificações de alvos de miRNAs em diversas espécies (JOHN et al., 2004; CHEN et al., 2005). Os valores desses parâmetros foram: penalidade de abertura de espaço (“*gap opening penalty*”) -8, penalidade sobre expansão de espaço (“*gap extension penalty*”) -2, pontuação mínima aceitável entre o par miRNA maduro e 3'UTR 120, energia livre máxima entre o par miRNA maduro e 3'UTR -15 kcal/mol, escala utilizada para medida da pontuação de complementaridade entre o par miRNA e 3'UTR. Todos estes parâmetros foram aplicados em linha de comando na plataforma Linux, e os resultados positivos foram recuperados do arquivo final utilizando um script escrito em Perl.

3.2 - Validações Experimentais

3.2.1 - Obtenção dos Parasitos

Foram utilizados nesse estudo parasitos da espécie *S. mansoni* linhagem LE, em diferentes fases do ciclo. Os estágios de verme adulto e cercária, foram cedidos pelo Moluscário do CPqRR/FIOCRUZ (Centro de Pesquisa René Rachou, Fundação Oswaldo Cruz) em Belo Horizonte/MG e pelo Laboratório de Enzimologia e Proteômica (LEP) da UFOP. Os demais estágios de esquistossômulo, ovo e miracício foram obtidos somente no ciclo recentemente mantido na UFOP. Os experimentos foram realizados com a permissão da Comissão de Ética no Uso de Animais CEUA/UFOP sob protocolo nº 2011/55.

As cercárias foram obtidas após 27 a 31 dias de infecção por miracídios no hospedeiro intermediário *Biomphalaria glabrata*. Cada molusco foi colocado em contacto com cerca de 10 miracídios. O sobrenadante e as impurezas presentes no meio foram retirados, colocando as cercárias em tubos de 1,5 ml sendo em seguida

congeladas em nitrogênio líquido e armazenadas no freezer a -80°C para posterior utilização.

A obtenção dos vermes adultos ocorreu a partir da perfusão do sistema porta-hepático de camundongos da linhagem *Swiss Webster* infectados com aproximadamente 100 cercárias pela via subcutânea após 50 dias (SMITHERS & TERRY, 1965; BASCH, 1981). Após a coleta, os vermes adultos foram armazenados em tubo de 1,5 ml a -80°C para posterior utilização.

Os esquistossômulos foram obtidos seguindo a metodologia de Harrop e Wilson (HARROP & WILSON, 1993). Inicialmente, o recipiente com as cercárias foram colocados no gelo para sedimentação, retirando as impurezas e sendo em seguida transferidas para tubos *falcons* de 15 mL. Foram ressuspendidas em 10 mL de meio RPMI 1640 (Invitrogen) filtrado, e deixado no gelo por 10 minutos. Esse processo de lavagem foi realizado por três vezes visando retirar o máximo de impurezas. Finalmente, cada tubo foi separado com aproximadamente 200.000 cercárias e cinco mL de RPMI 1640. Esses tubos foram vigorosamente agitados em vórtex, sob velocidade máxima, durante 90 segundos para que ocorresse a separação mecânica da cauda do corpo cercariano. Após a ruptura, todo volume foi adicionado a um novo recipiente de cultura adicionando RPMI 1640 suplementado com 1% de penicilina cristalina G (1000 UI/mL) e estreptomicina (1000 $\mu\text{g}/\text{mL}$), e incubado em estufa de CO_2 5% a 37°C por três horas e meia. Posteriormente foram realizadas em um fluxo laminar, várias lavagens com um intervalo de quarto minutos em cada, para sedimentação dos esquistossômulos e remoção das caudas presentes no sobrenadante. Esse processo foi acompanhado em um microscópio de luz invertida. O precipitado com apenas o corpo cercariano sem a presença de caudas, foi considerado como os esquistossômulos com 24 horas de cultivo *in vitro* (**Figura 9**). Estes foram separados e utilizados para cultivo de um dia ou armazenados em tubos de 1,5 mL congelados em nitrogênio líquido e mantidos a -80°C até o momento do uso.



Figura 9: Esquistossômulos cultivados *in vitro*: Os esquistossômulos foram cultivados por 24 horas em estufa de CO₂ 5% a 37°C em meio 169 (DE SOUZA GOMES, 2008).

Os esquistossômulos mecanicamente transformados foram cultivados durante 3,5 horas e 24 horas (EMT-3,5 horas, EMT-24 horas). O cultivo EMT-24 horas foi realizado em placas de seis poços contendo 8 mL de meio 169 (**Tabela 1**) suplementado com 10% de soro fetal bovino (Gibco) e 1% de penicilina cristalina G (1000 UI/mL)/estreptomicina (1000 µg/mL) por poço. (BASCH, 1981)

Tabela 1: Meio 169 com seus componentes e concentração.

Composição	Concentração
Hidrolisado de lactoalbumina	0,1%
Hidrocortisona	1×10^{-6} M
Hipoxantina	5×10^{-7} M
Triiodotironina (T3)	2×10^{-7} M
Serotonina	1×10^{-6} M
Glicose	0,1%
Meio mínimo vitamina	0,5%
Meio Schneider	5,0%
HEPES	20 mM
RPMI 1640 (Invitrogen)	q.s.p. 500 mL

Para obtenção dos ovos, cerca de 10 fígados de camundongos (linhagem *Swiss Webster* com 50 dias de infecção) foram triturados e homogeneizados com auxílio de um liquidificador doméstico em solução tampão de 200 mL (0,06 M de Na₂HPO₄,

0,0033 M de KH₂PO₄, pH 8,3), contendo 20 mg de tripsina (Invitrogen, Carlsbad, CA, EUA), e mantidos por 3h a temperatura de 37°C no banho-maria. Em seguida foram filtrados em peneiras de malhas de 300 e 180 µm em solução salina isotônica sucessivamente (CASTRO-BORGES, 2007). Os ovos foram lavados com essa solução salina e decantados por 5 minutos. Separou-se o sobrenadante e repetiu-se o passo acima por várias vezes até não se retirar mais ovos do material. Após a coleta, os ovos foram transferidos para tubos de 1,5 ml sendo em seguida congelados em nitrogênio líquido e armazenados no freezer a -80°C para posterior utilização.

A eclosão dos miracídios foi realizada a partir dos ovos coletados depositando-os com água em uma placa de Petri exposta à luz. A coleta do sobrenadante com os miracídios foi realizada com auxílio de uma pipeta tipo Pasteur e transferindo para um tubo tipo falcon, mantido em banho de gelo. Após a sedimentação, os miracídios foram transferidos para tubos tipo eppendorf sendo em seguida congelados em nitrogênio líquido e armazenados em freezer a -80°C até o momento do uso.

3.2.2 - Análise da expressão gênica por qRT-PCR

A metodologia para a expressão dos genes hospedeiros e dos miRNAs utilizada neste trabalho foi realizada seguindo as etapas já padronizadas no Laboratório de Bioquímica e Biologia Molecular (LBBM) e brevemente descritas abaixo.

3.2.2.1 - Determinação do perfil de expressão de mRNAs (Genes hospedeiros)

3.2.2.1.1 - Extração do RNA total

Para extração de RNA total foram utilizadas aproximadamente 50 a 100 mg de vermes adultos, cercárias, esquistossômulos de 3,5 horas e de 24 horas utilizando o kit de RNA total (*SV total RNA Isolation System - PromegaTM*) seguindo o protocolo do fabricante.

Todas as amostras biológicas foram homogeneizadas em 1 mL de TRIzol[®] Reagent (InvitrogenTM) com auxílio de um homogeneizador tipo *politron*. Posteriormente, os homogeneizados foram transferidos para tubos *eppendorf* de 1,5 mL e incubados por 5 minutos à temperatura ambiente para permitir a completa dissociação dos complexos de nucleoproteínas. A seguir foram adicionados 0,2 mL de clorofórmio

(Sigma - St. Louis, MO, USA) para cada 1,0mL de TRIzol. A mistura foi homogeneizada vigorosamente com auxílio de um vórtex. Posteriormente, as amostras foram centrifugadas por 12 minutos a 12.000g a temperatura ambiente. A fase aquosa foi transferida para um novo tubo seguido da adição de volume equivalente de etanol 95% v/v (preparado com água livre de RNAses) e homogeneizado suavemente por inversão o tubo, por três vezes, para a precipitação do RNA. A seguir o RNA total foi purificado com o kit SVRNA System conforme instrução do boletim técnico. A qualidade das preparações foi avaliada em gel de agarose/formaldeído, como mostra a **Figura 10**. A pureza e quantificação dos RNAs foram determinadas utilizando o aparelho Nano Vue Spectrophotometer (GEHealthcare) avaliando as relações entre os comprimentos de onda 260nm/280nm e 260nm/230nm indicativos de pureza do amostra.

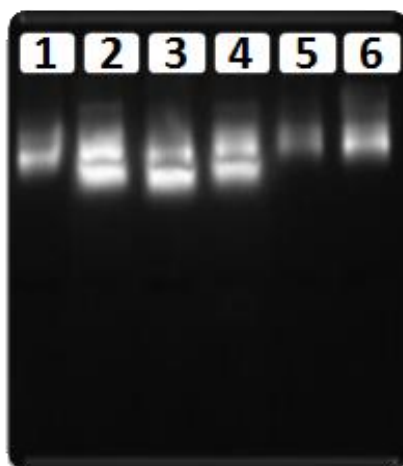


Figura 10: RNA total obtido a partir de estágios evolutivos do *S. mansoni*. O RNA total obtido a partir de cercárias (1), esquistossômulos com 3,5 horas (2); esquistossômulos com 24 horas (3); vermes adultos (4); miracídios (5); ovos (6).

3.2.2.1.2 - Oligonucleotídeos iniciadores (*Primers*)

Os genes hospedeiros dos miRNAs utilizados nesse estudo foram amplificados a partir de oligonucleotídeos iniciadores idealizados com auxílio do programa *Gene Runner (Version 3.05)*, conforme a **Tabela 2**. Esses iniciadores foram baseados nas sequências de cDNA depositadas no bancos de dados GeneDB.

Tabela 2: Sequências dos iniciadores dos genes hospedeiros dos miRNAs.

Produto do gene	Sequência do primer	Nº de acesso
	F: 5'-TCCA CTGTGCCTCTTGATG-3'	
Cysteinyl tRNA synthetase, cytoplasmic	R: 5'-GGCTGCGTCCAAGGTATC-3'	Smp_170800
	F: 5'-GCCTTTGACGGGTACACTAG-3'	
Protocadherin 11	R: 5'-TGGGTCATCAGCATCGTC-3'	Smp_141740
	F: 5'-GGTCGTGGTTCCTTGTGG-3'	
Par 3 partitioningdefective 3	R: 5'-GATCCTGTGGTCCGAGCTTC-3'	Smp_157550
	F: 5'-GTGGTGTGCCGATGAACTG-3'	
Cyclindependent kinase 14	R: 5'-ATCCTGGAGGGTTGTGAACC-3'	Smp_155330
	F: 5'-ATGGTTCGAGGCGAAGATC-3'	
1 Phosphatidylinositol 4,5 bisphosphate	R: 5'-CACTTGCCACGAATGCTG-3'	Smp_146330
	F: 5'-CCGATATGCGAAGGTGAG-3'	
Zinc finger protein 362	R: 5'-GCAGCAGCCTCTTTATGACG-3'	Smp_169260
	F: 5'-ATGTGGGTTTCGAGAGTCC-3'	
Skiinter acting protein	R: 5'-CCTCCTGATTGCGTGGTAC-3'	Smp_105860
	F: 5'-AAACACCTGCGGCAACTG-3'	
Hypothetical protein	R: 5'-CGGATGCTGGTGTGATTTG-3'	Smp_130330
	F: 5'-ATTGACAACCTGCGACAGC-3'	
Acetyl-CoAacetyltransferase, cytosolic (ACAT2)	R: 5'-CATTCTCGACGGAGATAGAC-3'	Smp_129330
	F: 5'-TGCTGGGCTTGATCTGTC-3'	
Chromatin assembly factor 1 subunit b	R: 5'-GCACGTTTCATCTGGAGAC-3'	Smp_153610
	F: 5'-TGGAAGCACTGGACGACC-3'	
Ryanodine receptor	R: 5'-GAACGCTCAACCTGCAGATC-3'	Smp_163570
	F: 5'-AATGGCGTTATGGGTGTG-3'	
Hypothetical protein	R: 5'-TGGAATGCTGAGTGAGTG-3'	Smp_048030
	F: 5'-TTCGATTGGACCCACAG-3'	
Venom allergen (val) protein	R: 5'-GTCCTCTTCGCCACACAG-3'	Smp_141560
	F: 5'-CTGTGCCTGTTTCTCTCC-3'	
Hypothetical protein	R: 5'-GGCTCATGCTGTCCATTG-3'	Smp_121240
	F: 5'-GTTGGGTGGGCATTGGTTAG-3'	
RNA polymerase II associatedfactor 1	R: 5'-ATGTACCGTGTCCCTCCTG-3'	Smp_045940
	F: 5'-ACCTTCCTCTGCTTGCTG-3'	
Mindbomb	R: 5'-GAACCACAGGACCTCTCTC-3'	Smp_171440
	F: 5'-CTGGTGTCAAACTTGG-3'	
Acid sphingomyelinase phosphodiesterase 3b	R: 5'-GGAAGAGAAAGAAGGACG-3'	Smp_133330
	F: 5'-GCAACGTGCCATCCTGTG-3'	
Helicase	R: 5'-CGCGAAATCAGCAACTACGG-3'	Smp_059250
	F: 5'-GTCTATCCTGCTGATTGTGC-3'	
Cercarial elastase (S01 family)	R: 5'-TACACAGGATGGCACGTTG-3'	Smp_006520
	F: 5'-TGTTCCAACCACGGTCTCG-3'	
Eukaryotic translation initiation factor 4E (EIF4E)	R: 5'-TCGCCTTCCAATGCTTAGG-3'	Smp_001500

3.2.2.1.3 - RT-PCR e obtenção dos cDNAs

A primeira fita do cDNA das amostras de RNA total foram sintetizadas utilizando 1 µg de RNA total extraído e o Kit *High Capacity RT-PCR System* (Applied Biosystems), seguindo as recomendações dadas pelo fabricante.

Para cada 1µg de RNA total extraído, foram utilizados: 2 µL de tampão da reação, 2µL de *primers* randômicos, 10 mM de dNTPs, 1,0 µL de Transcriptase reversa Multiscribe, 1 µL de inibidor de RNase e água livre de nuclease para um volume final de 10 µL. Homogeneizou-se o tubo pipetando para cima e para baixo 2 vezes para completa mistura do RNA total na mistura.

Posteriormente, o tubo contendo a mistura de reagentes e o RNA total foi incubado em termociclador (ThermoHybaid Px2) seguindo o seguinte programa: 10 minutos a 25° C, 120 minutos a 37° C para produção da primeira fita de DNA (cDNA), 85°C por 5 minutos para inativação da enzima e por fim 4°C. A amostra de cDNA foi estocada a -80°C até o momento do uso.

3.2.2.1.4 - PCR quantitativa em Tempo Real (qRT-PCR)

A análise da expressão dos genes em estudo foi realizada a partir da técnica de PCR quantitativa em tempo real. As reações foram realizadas utilizando o kit *Platinum SYBR Green qPCR Super Mix-UDG ROX* (Applied Biosystems), utilizando 2µL dos iniciadores na concentração de 300 nM, 5µL de SYBR Green, 3µL de cDNA diluído 5x com água livre da enzima DNase, somando um volume final de 10 µL de reação. Todos os ensaios foram realizados em triplicata técnica e biológica para todos os genes hospedeiros dos miRNAs intrônicos em *S. mansoni*, com o normalizador *eIF4E* presente na mesma placa que os genes avaliados. Os valores do *threshold* foram fixados em 0,2 e os do *baseline* ajustados para 3-15 ciclos.

As análises foram feitas pelo método de quantificação relativa da expressão gênica (ΔCq), que permite quantificar diferenças no nível de expressão de um alvo específico entre as diferentes amostras. Os níveis dos genes alvos foram normalizados pelos níveis do controle endógeno (Gene normalizador). Os resultados foram alcançados pela fórmula aritmética $2^{-\Delta Cq} = 2^{-(Cq_{\text{Gene}} - Cq_{\text{Gene Normalizador } eIF4E})}$. A reação de

qRT-PCR foi conduzida conforme programação contida no aparelho ABI 7300 Applied Biosystems.

3.2.2.1.5 - Curva de eficiência dos oligonucleotídeos iniciadores

A eficiência da amplificação dos iniciadores utilizados foi analisada através da realização de 5 diluições seriadas (1:4) de uma amostra de cDNA de cercária. Este ensaio foi realizado em triplicatas com concentração inicial de *primer* de 300 nM. Os dados obtidos geraram um gráfico de regressão linear onde o eixo X representa o log das concentrações de cDNA e o eixo Y o valor de Cq.

A eficiência da amplificação é determinada pelo *slope* da curva representada pela fórmula: $\text{Eficiência} = [10^{(-1/\text{slope})} - 1] \times 100$ (**Figura 11**). Essa eficiência foi considerada apropriada para avaliação da expressão gênica, quando obtidos valores da reação acima de 95% e abaixo de 105%. Os valores do *threshold* foram fixados em 0,2 e os do *baseline* ajustados para 3-15 ciclos como representados na **Figura 12**.

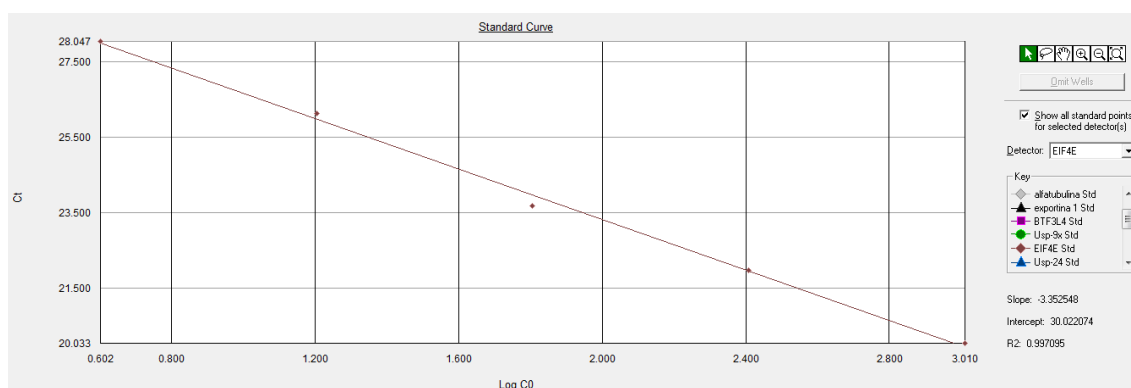


Figura 11: Curva padrão referente ao gene *EIF4E*. No eixo X estão representados os valores de Log da concentração de cDNA e no eixo Y os valores de Cq correspondes a cada diluição realizada. O coeficiente de linearidade e de *slope* estão representados na figura. Para a realização da curva de eficiência, foram utilizadas amostras de cDNA de cercária em uma diluição seriada de 4 vezes.

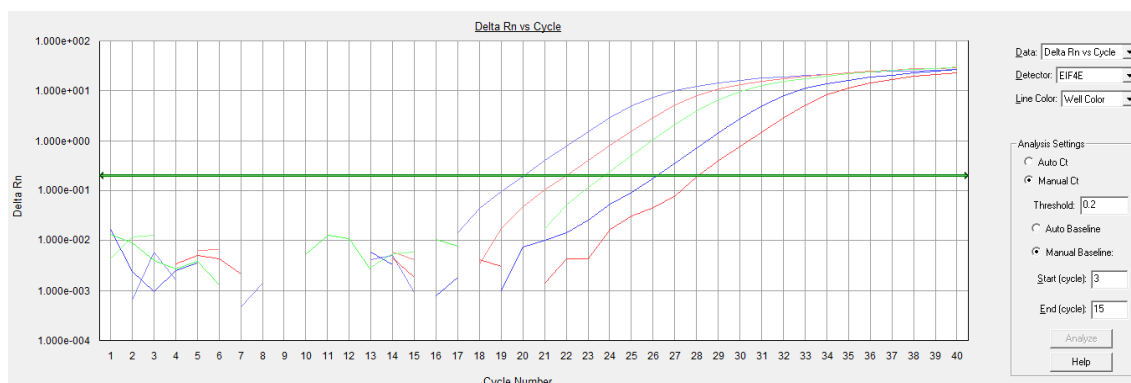


Figura 12: Gráfico de amplificação referente à curva de eficiência do gene *eIF4E*. No eixo X está representado o valor dos ciclos da qRT-PCR e no eixo Y os valor de Delta Rn. Amostras de cDNA de cercaria foram utilizadas em uma diluição seriada de 4 vezes.

3.2.2.1.6 - Qualidade dos produtos amplificados

Os produtos amplificados na qRT-PCR juntamente com o padrão de peso molecular de *100pb DNA Ladder* (Invitrogen) foram analisados em gel de agarose a 1,2% (90 volts por aproximadamente 80 minutos). A finalidade dessa metodologia foi confirmar a presença de uma única banda com peso molecular esperado para cada um dos *amplicons* obtidos.

3.2.2.1.7 - Curva de dissociação dos *amplicons*

A realização da curva de dissociação dos *amplicons* teve como o objetivo observar a presença de possíveis amplificações inespecíficas. Ao final dos 40 ciclos da programação da qRT-PCR, a temperatura é elevada gradualmente de 60°C á 95°C, mantendo-se por 15s em cada temperatura, durante o qual é feito a leitura da emissão de fluorescência. O sinal fluorescente emitido pelo SYBR™ Green é reduzido quando ocorre a desnaturação dos *amplicons*. Essa desnaturação ocorre em temperaturas específicas sendo dependente do tamanho e do número de bases GCAT.

O gráfico resultante permite verificar se há um ou mais produtos de PCR presentes em cada reação devido a diferenças de temperatura de dissociação da sequência (*Melting temperature* - T_m). Neste caso a temperatura desejável foi de 75°C como exemplificado na **Figura 13** no gene de entrada *Smp_170800*.

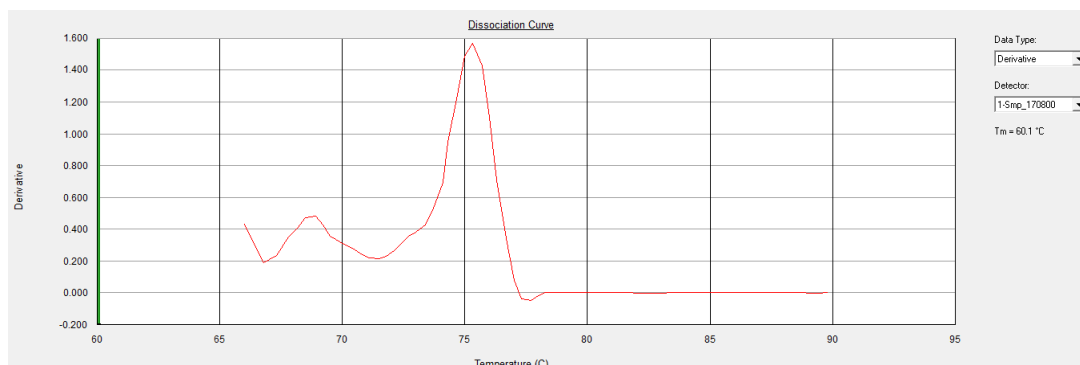


Figura 13: Curva de dissociação do amplicon referente ao gene *Smp-170800*. No eixo Y está representado a derivada do valor emitido pela fluorescência e no eixo X a temperatura de dissociação do produto gerado pela qRT-PCR.

3.2.2.1.8 Análises estatísticas

A expressão relativa dos mRNAs nos estágios do parasito foi comparada utilizando a análise da variância por ONE-WAY (Teste de Tukey ou *T-Student*). Foi considerado estatisticamente significativo $p < 0.05$. A análise estatística foi feita utilizando o programa GraphPad Prism 5.

3.2.2.2 - Determinação do perfil de expressão de microRNAs por qRT-PCR

3.2.2.2.1 - Extração do RNA total

Aproximadamente 100 mg de vermes adultos, cercárias e esquistossômulos de 3,5 e 24 horas foram utilizados para extração de RNA total utilizando o kit miRNeasy (Qiagen) seguindo o protocolo do fabricante.

As amostras biológicas foram homogeneizadas em 700 μ l de *QIAzol Lysis Reagent* em tubos *ependorf* de 1,5 ml com auxílio de um homogeneizador do tipo politron. O homogeneizado foi incubado por cinco minutos à temperatura ambiente permitindo assim a completa dissociação dos complexos de nucleoproteínas. Posteriormente foram adicionados 140 μ l de clorofórmio (Sigma - St. Louis, MO, USA) em cada amostra. A mistura foi homogeneizada vigorosamente com auxílio de um vórtex. Em seguida, as amostras foram centrifugadas por 12 minutos a 12.000g a 4°C. A fase aquosa foi transferida para um novo tubo seguido da adição de 1,5 vezes o volume equivalente de etanol 100% e homogeneizado suavemente por inversão o tubo, por três vezes, para a precipitação do RNA. A seguir o RNA total foi purificado conforme instrução do boletim técnico do fabricante.

A padronização da metodologia foi avaliada em gel de agarose/formaldeído. A pureza e quantificação dos RNAs foram determinadas utilizando o aparelho Nano Vue Spectrophotometer (GE Health care) avaliando as relações entre os comprimentos de onda 260nm/280nm e 260nm/230nm indicativos de pureza do amostra.

3.2.2.2.2 - Oligonucleotídeos iniciadores (*Primers*)

Os oligonucleotídeos iniciadores específicos para os miRNAs foram idealizados com base nas sequências maduras dos miRNAs identificados neste trabalho (**Tabela 3**). Estes *primers* foram utilizados para posterior qRT-PCR.

Tabela 3: Sequências do oligonucleotídeos iniciadores referentes aos miRNAs.

microRNA	Tamanho	Fita	Sequência do <i>primer</i>
sma-mir-new_i1 5p	22	5'	CAACTGTTGCGTATTTTTGGAG
sma-mir-new_i1 3p	22	3'	TTTTGGAGTTGCTGATCACACT
sma-mir-new_i2 5p	22	5'	GAATTTAATGGAGAGTGGAAGG
sma-mir-new_i2 3p	22	3'	CAAGAAGCTAGATTTTTTGTGC
sma-mir-new_i3 5p	22	5'	TAACATTGTTATTGATGATGAT
sma-mir-new_i3 3p	22	3'	ATAAGTCATCAATGACATGTTA
sma-mir-new_i4 5p	22	5'	CAGCGGCGGCAGTCGCGGCTGC
sma-mir-new_i4 3p	22	3'	CTGCTGCCGCTGCAGTTACCCT
sma-mir-new_i5 5p	22	5'	GTGCCGGTGGTGGTAGTAGTGG
sma-mir-new_i5 3p	22	3'	GTAGTGGTGGTCTTGCCGGTAC
sma-mir-new_i6 5p	22	5'	ATTCTTCTGCTGCTGCTGCCGC
sma-mir-new_i6 3p	22	3'	CTGCGGCGGCCGCTGTCGCTGC
sma-mir-new_i7 5p	22	5'	ACATAAATACTTCCAAGAAATG
sma-mir-new_i7 3p	22	3'	TGGAACTCCATTCGATGTGTTT
sma-mir-new_i8 5p	22	5'	CTACTACTACTACTACTACTAC
sma-mir-new_i8 3p	22	3'	TAGTAGTGGTAATAACATTA
sma-mir-new_i9 5p	22	5'	ACAGCTGCTGGCTCAGTTGGTG
sma-mir-new_i9 3p	22	3'	TCAGGCGGAGCGGATATAGTGC
sma-mir-new_i10 5p	22	5'	CATCTGCTCTGCCAGCTGAAGA
sma-mir-new_i10 3p	22	3'	AGTGTTGGAGAAGTGGGTCCAG
sma-mir-new_i11 5p	22	5'	CACCAGGACATAGTGCAACATC
sma-mir-new_i11 3p	22	3'	TGGGTGCACTAGTAACGGAAAC
sma-mir-new_i12 5p	22	5'	TACAACCAAAGTTCTTCAGCGG
sma-mir-new_i12 3p	22	3'	GTCGGGGTGTAGATTGTTTCTA

sma-mir-new_i13 5p	22	5'	AGCCTAGTGCTTCGCACTCTAA
sma-mir-new_i13 3p	22	3'	TTCGCACTCTAATCTCAACGCG
sma-mir-new_i14 5p	22	5'	TATTACGTTTTCTTTTGCATTG
sma-mir-new_i14 3p	22	3'	CAAACAAGGTTGCGTAAGTGGA
sma-mir-new_i15 5p	22	5'	CAGTGTCTGGCTGAGGAAAAC
sma-mir-new_i15 3p	22	3'	AAAGGTGGAGTCCAGGTGAGTG
sma-mir-new_i16 5p	22	5'	GCAGGTCTAACTTGGGCTCAAA
sma-mir-new_i16 3p	22	3'	TGCTGGTCCAATCCTTTTACCT
sma-mir-new_i17 5p	22	5'	TGGAACTAATCCTGGCGGCGGT
sma-mir-new_i17 3p	22	3'	AAATGGTGCATTACCTATTTCT
sma-mir-new_i18 5p	22	5'	CATGTGGCTCAGTATATTCATG
sma-mir-new_i18 3p	22	3'	TCATGATTTCTTAGGTACGCGG
sma-mir-new_i19 5p	22	5'	ACACTGAGACACATTGTTGTTC
sma-mir-new_i19 3p	22	3'	ATTCGCAGTGTGCGCAGTGGGT

3.2.2.2.3 - Obtenção da primeira fita de DNA (cDNA)

A primeira fita do cDNA foi sintetizada utilizando 1µg de RNA total extraído dos diferentes estágios do *S. mansoni*, utilizando o Kit miScript Reverse Transcription (miScript II RT Kit - QIAGEN), seguindo as recomendações dadas pelo fabricante. A reação dessa transcrição reversa ocorre inicialmente com uma poliadenilação das moléculas de RNA. A etapa posterior é a hibridação de um *primer* oligo-dT na cauda poli-A para a síntese da fita de cDNA que será utilizada na reação de qRT-PCR (**Figura 14**).

Para cada 1µg de RNA total enriquecido com pequenos RNAs, foram utilizados 4 µL miScript RT Buffer 5x, 1µL de transcriptase reversa (miScript Reverse Transcriptase Mix) e água livre da enzima RNase para um volume final de 20 µL. A mistura foi incubada a 35°C por 60 minutos, seguidos de 5 minutos a 95°C, com auxílio de um termociclador (Biocycler, version 3.2). Todas as amostras de cDNA foram estocadas a -80°C e utilizadas em no máximo duas semanas.

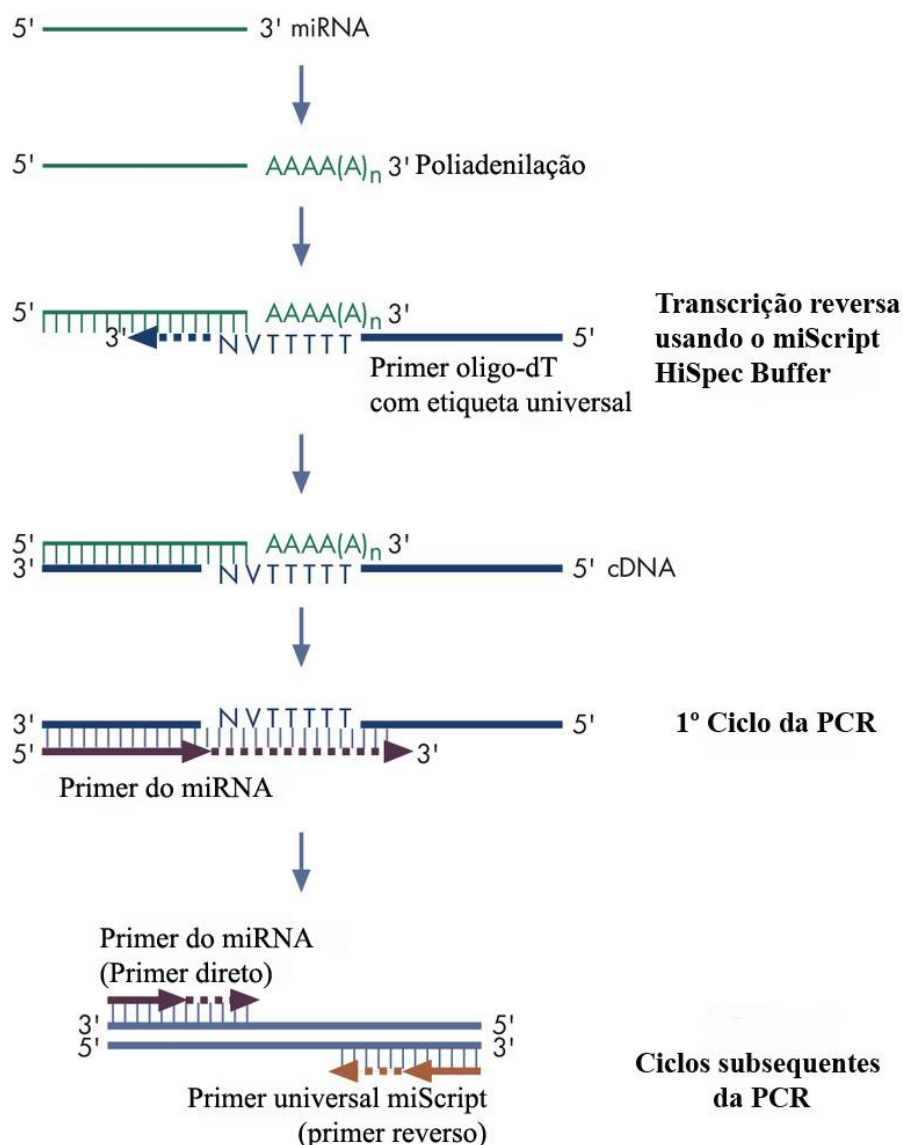


Figura 14: Síntese do cDNA a partir do miRNA. Inicialmente ocorre uma reação de poliadenilação nas moléculas de RNA. Um primer oligo-dT fornecido pelo kit é utilizado para sintetizar a fita de cDNA. Esse cDNA é submetido ao primeiro ciclo da reação de PCR no qual o primer específico do miRNA de interesse é utilizado. Os ciclos seguintes ocorrem a partir do primer específico e de um primer universal fornecido pelo kit (Adaptado de miScript II RT Kit - QIAGEN).

3.2.2.2.4 - Expressão dos miRNAs detectada por qRT-PCR

A análise da expressão dos miRNAs preditos foi realizada utilizando a técnica de PCR em tempo real. As reações foram realizadas pelo kit *SYBR Green detection* (miScript Primer Assay – Qiagen) em placas de 96 poços (MicroAmp® Optical 96 Well Reaction Plate – Applied Biosystems) e seladas com adesivo óptico (MicroAmp® Optical Adhesive Film – Applied Biosystems) ao final do procedimento.

Inicialmente, foram pipetados 2 μL dos oligonucleotídeos iniciadores (na concentração de 2,5 μM) em triplicata, 2 μL de cDNA diluído 10 vezes, 1 μL de 10x miScript Universal Primer e 5 μL de 2x QuantiTect SYBR Green PCR Master Mix, totalizando um volume de reação de 10 μL em cada poço. Todos os ensaios foram realizados em triplicata biológica para os miRNAs avaliados, com o controle endógeno presente em todas as placas, sendo nesse caso o pequeno RNA nuclear U6. Os valores de *baseline* (ciclos iniciais em que há pequenas alterações na fluorescência) foram ajustados para 3-15 ciclos. O *threshold*, desvio padrão médio do repórter normalizado (Rn) para os ciclos iniciais da PCR multiplicado por um fator ajustável, foi ajustado à região associada ao crescimento exponencial do produto da PCR e, portanto, fixado em 0,02 para todas as amostras, uma vez que se comparou o mesmo miRNA nos diferentes estágios. O Rn refere-se à relação entre a intensidade de fluorescência emitida pelo corante repórter pela intensidade de fluorescência emitida pelo corante da referência passiva (ROX).

As análises foram feitas pelo método de quantificação relativa da expressão gênica (ΔCT), que permite quantificar diferenças no nível de expressão de um alvo específico entre as diferentes amostras. Os níveis dos miRNAs alvos foram normalizados pelos níveis do controle endógeno. Os resultados foram alcançados por uma fórmula aritmética que considera a quantidade do alvo, normalizado ($2^{-\Delta\text{CT}}$). A reação de qRT-PCR foi conduzida seguindo a programação: 15 min a 95 °C, seguido por 40 ciclos de 94 °C por 15 s, 55 °C por 30 s e 70 °C por 30 s.

3.2.2.2.5 - Verificação dos amplicons

Os produtos amplificados na qRT-PCR juntamente com o padrão de peso molecular de 100pb DNA Ladder (Invitrogen) foram analisados em gel de agarose a 1,2% (90 volts por aproximadamente 80 minutos). A finalidade dessa metodologia foi confirmar a presença de uma única banda próxima a 100pb referente aos miRNAs.

3.2.2.2.6 - Curva de dissociação dos amplicons

A realização da curva de dissociação dos amplicons teve como o objetivo observar a presença de possíveis amplificações inespecíficas. Ao final dos 40 ciclos da programação da qRT-PCR para miRNAs, a temperatura foi elevada gradualmente de

60°C á 95°C, mantendo-se por 15s em cada temperatura, durante o qual é feito a leitura da emissão de fluorescência.

O gráfico resultante permite verificar se há um ou mais produtos de PCR presentes em cada reação devido a diferenças de temperatura de dissociação da sequência (*Melting temperature* - T_m). Neste caso a temperatura desejável foi entre 75°C e 80°C como exemplificado na **Figura 15**.

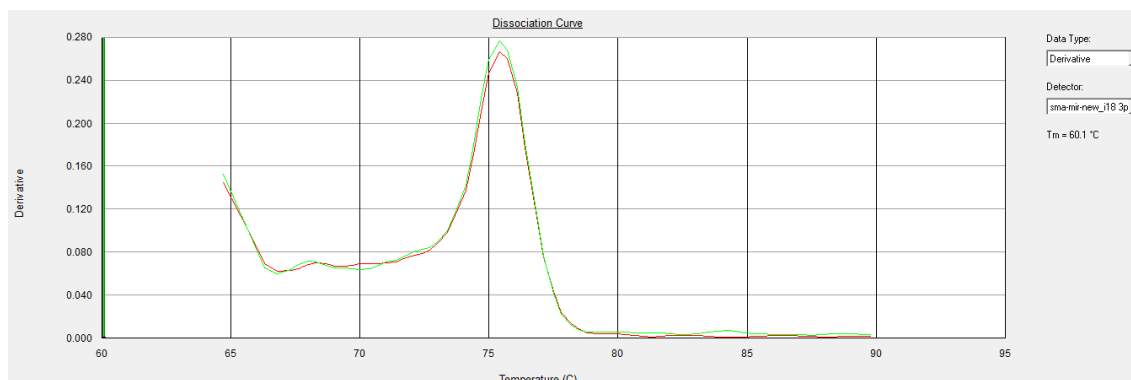


Figura 15: Curva de dissociação do miRNA 18-3p. Representado na figura somente um pico de amplificação da curva de dissociação do *sma-miR-new 18-3p*.

3.2.2.2.7 Análises estatísticas

A expressão relativa dos miRNAs nos estágios do parasito foi comparada utilizando a análise da variância por ONE-WAY (Teste de Tukey ou *T-Student*). Foi considerado estatisticamente significativo $p < 0.05$. A análise estatística foi feita utilizando o programa GraphPad Prism 5. A correlação de Pearson foi utilizada para avaliar o coeficiente entre a expressão dos miRNAs intrônicos e genes hospedeiros considerando como significativo $p < 0,05$.

4 - Resultados e Discussões

4.1 - Identificação de miRNAs no genoma de *S. mansoni*

Os arquivos do genoma e do transcrito utilizados neste trabalho foram obtidos do banco de dados GeneDB. Eles foram editados para o formato fasta para que gerassem um maior acesso de leitura nas análises posteriores. A última versão do genoma 5.1 foi utilizada neste estudo, na qual apresenta 380Mb, distribuídas entre os 8 cromossomos e nos seus 885 *scaffolds*. As sequências de ESTs relativas ao transcrito, foram compostas de aproximadamente 30.000 contigs provenientes de bibliotecas de cDNA de todas as fases do parasito.

A identificação dos miRNAs maduros e precursores de origem intrônica foi realizada utilizando o *pipeline* descrito na sessão materiais e métodos. Inicialmente, todas as sequências intrônicas foram extraídas do genoma gerando um arquivo fasta contendo 57.528 íntrons. Esse arquivo foi posteriormente submetido a vários filtros e parâmetros com a finalidade de aumentar a acurácia da predição. A próxima etapa filtrou somente os íntrons de tamanho entre 50 e 120 nucleotídeos gerando um conjunto de 1.901 sequências, organizadas no mesmo formato. Esse arquivo foi então submetido ao programa RNAfold, sendo analisados os seguintes parâmetros: características das moléculas precursoras dos miRNAs como o tamanho, conteúdo de GC, energia mínima livre (MFE), energia mínima ajustada (AMFE), índice de energia livre (MFEI), energia mínima livre do conjunto termodinâmico (MFEE), diversidade do conjunto termodinâmico e frequência da estrutura de energia mínima livre no conjunto termodinâmico. A seguir, foram selecionadas apenas as estruturas que tiveram o MFE de $\Delta G < -25$ Kcal/mol, sendo esse valor de energia de estrutura secundária considerado muito estável (XUE et al., 2008; REHMSMEIER et al., 2004). A escolha dessas características para a predição foi feita baseando-se em vários estudos (BONNET et al., 2004, ZHANG et al., 2009 ZHANG et al., 2005) evitando a ocorrência de precursores falsos positivos no genoma do *S. mansoni*.

A **Tabela 4** mostra as 19 sequências preditas como moléculas precursoras de miRNAs de origem intrônica, bem como o número de acesso dos respectivos genes hospedeiros, e na **Tabela 5** as suas características estruturais e termodinâmicas.

A seguir, as sequências foram analisadas no banco de dados mirBASE, utilizando a ferramenta Blastn para a identificação de sequências homologas.

Entretanto, não foi verificada homologia entre as sequências depositadas no banco, sugerindo que os candidatos a miRNAs identificados sejam espécie-específicos.

Smp_163570.1	GUGGUUCACAAGGUUCACCAGGACAUAGUGCAACAUCAGCUGACGAUGAUCCUGGGUGCACUAGUAACGGAAACGGU AACA AUGGUAGUGGUAGUUCUGGUGGAGCUGAU AACGUAG
Smp_048030.1	GUUACCUGAGGACUUGUACAACCAAAGUUCUUCAGCGGUGCUGUUUACAAACGCUUCGGAUUGAAGACCAAUCGGUA UGAAAAGCGUCGGGGUGUAGAUUGUUUCUACUAUAUUUAG
Smp_141560.1	GUGAGUUUGAGAUGAUGUGGAAGCCUAGUGCUUCGCACUCUAAUCUCAACGCGUUUCAUCACAUUGCUC AUGAAUGC ACGAUU AUUAGUUUGCUAUGGUUUUU AUGCUUGUAUCAG
Smp_121240.1	GUAACA UUUU AUUACGUUUUCUUUUGCAUUGCAAGAUUUUUCUAGCAAACAAGGUUGCGUAAGUGGAAAGCUGUGUG CAU AUGUAAUGAAA UGCACAUGUUUAUCAG
Smp_045940.1	GUCAGUGUCCUGGCUGAGGAAAACAGAAUUAUUAUCUACUGAGCUUACAAUCGUUGGAACAAAUCUGAAAAGGUGG AGUCCAGGUGAGUGUUUCACUUGUCUUUACUUCAAAG
Smp_171440.1	GUGGGUCGAUACUUCUAGGACCAACAGGAUCAGCCAUUGC UUGCAGGUCU AACUUGGGCUCAAACAACGAAAAAACG UCUGCUGGUCCAUCUUUUACCUACACAAG
Smp_133330.1	GUAUGAAUGGAACUAAUCCUGGCGGCGGUGGUUUUGGUGGUCGUGGUGGUGGUGGUGGAAAUGGUGCAUUAACCUAUU UCUUCAUUUCUUUAUGGAUCUCGUUCA AUGAUUAAAG
Smp_059250.1	GUAAGUCAUGUGGCUCAGUAUAUUCAUGAUUUCUAGGUACGCGGUUGC UAAAAGUCUGAACCAAACCAGCGCGUU AUUUACACGACGCCUAUCAAGGCCUGAGCAAUCAGAAG
Smp_006520.1	GUAAGAGAACACACUGAGACACAUUGUUGUCAAUGUAAUCGAUUA AAAUUCGCAGUGUGCGCAGUGGGUUUGCAUAG AUCUAAUUUGACAUACUUUGUCGUUUUCAG

Tabela 5: Características estruturais e termodinâmicas dos pré-miRNAs de origem intrônica de *S. mansoni*.

miRNA ID	Gene Hospedeiro	Intron N°	Conteúdo GC (%)	Tamanho	MFE	AMFE	MFEI	MFEI	MFEI	FREQUÊNCIA (%)	DIVERSIDADE (%)	Cromossomo/Scaffold	Início	Fim	Fita
sma-mir-new_i1	Smp_170800.1	5	50.45	111	-28,5	-25.68	-0.51	-24.90	5.98	22.66	Schisto_mansoni.Chr_1	26307363	26307475	-	
sma-mir-new_i2	Smp_141740.1	3	36.89	103	-26,3	-25.53	-0.69	-27.00	7.09	11.24	Schisto_mansoni.Chr_1	35275458	35275562	-	
sma-mir-new_i3	Smp_157550.1	6	30.63	111	-32,6	-29.37	-0.96	-23.99	2.09	17.49	Schisto_mansoni.Chr_1	57853672	57853784	-	
sma-mir-new_i4	Smp_155330.1	2	64.67	75	-42,8	-57.07	-0.86	-42.10	24.10	5.81	Schisto_mansoni.Chr_1	1164947	1165023	-	
sma-mir-new_i5	Smp_146330.1	2	55.91	93	-25	-26.88	-0.48	-24.05	1.24	17.40	Schisto_mansoni.Chr_2	9417066	9417160	-	
sma-mir-new_i6	Smp_169260.1	3	59.14	93	-34,1	-36.67	-0.62	-33.80	14.92	10.36	Schisto_mansoni.Chr_3	6225576	6225670	-	
sma-mir-new_i7	Smp_105860.1	4	36.44	118	-25,5	-21.62	-0.59	-24.60	1.33	8.68	Schisto_mansoni.Chr_3	11571155	11571274	-	
sma-mir-new_i8	Smp_130330.1	1	29.73	111	-40,7	-36.67	-1.23	-33.54	2.77	37.61	Schisto_mansoni.Chr_3	27060590	27060702	-	
sma-mir-new_i9	Smp_129330.1	16	52.94	102	-27,8	-27.25	-0.51	-14.15	2.05	29.15	Schisto_mansoni.Chr_5	367233	367336	-	
sma-mir-new_i10	Smp_153610.1	5	50	114	-25,3	-22.19	-0.44	-24.30	0.50	18.83	Schisto_mansoni.SC_0013	1343629	1343744	-	
sma-mir-new_i11	Smp_163570.1	43	48.72	117	-37,4	-31.97	-0.66	-36.40	3.84	11.50	Schisto_mansoni.SC_0013	3164407	3164525	-	
sma-mir-new_i12	Smp_048030.1	2	42.74	117	-28,1	-24.02	-0.56	-26.40	1.48	23.96	Schisto_mansoni.Chr_2	31231102	31231220	+	
sma-mir-new_i13	Smp_141560.1	1	40.52	116	-26,5	-22.84	-0.56	-17.40	0.32	30.96	Schisto_mansoni.Chr_6	13017146	13017263	+	
sma-mir-new_i14	Smp_121240.1	3	33.33	108	-25,2	-23.35	-0.70	-22.22	2.01	19.33	Schisto_mansoni.Chr_7	8209436	8209545	+	
sma-mir-new_i15	Smp_045940.1	4	39.66	116	-26,7	-23.02	-0.58	-19.10	0.47	40.83	Schisto_mansoni.Chr_W	36999127	36999244	+	
sma-mir-new_i16	Smp_171440.1	4	47.22	108	-27,2	-25.19	-0.53	-16.48	0.83	28.10	Schisto_mansoni.Chr_W	42455993	42456102	+	
sma-mir-new_i17	Smp_133330.1	6	43.86	114	-25,9	-22.72	-0.52	-11.11	0.46	30.85	Schisto_mansoni.SC_0080	454020	454135	+	
sma-mir-new_i18	Smp_059250.1	2	44.83	116	-26,6	-22.93	-0.51	-16.66	1.75	23.76	Schisto_mansoni.SC_0129	841470	841587	+	
sma-mir-new_i19	Smp_006520.1	1	37.96	108	-25,3	-23.43	-0.62	-20.89	6.83	24.36	Schisto_mansoni.SC_0243	100564	100673	+	

Muitos precursores de miRNAs intrônicos apresentam na sua estrutura sítios receptores de *splicing*, sugerindo que a sua remoção do transcrito primário siga a regra clássica desse processamento de íntrons (WESTHOLM & LAI, 2011). Os sítios são representados pelos nucleotídeos GU presentes na extremidade 5', e os nucleotídeos AG, na extremidade 3' (BLACK, 2003; CLANCY, 2008). Outra característica, é que eles apresentam muitas pirimidinas na sua extremidade 3' antes do sítio AG, e uma adenina chamada de ponto de ramificação, situada entre 20 e 50 nucleotídeos antes da extremidade 3'. Essa adenina, entretanto pode não apresentar conservação entre grupos como mamíferos (CLANCY, 2008).

Para verificar a presença desses sítios, foi realizado um *weblogo* nos 15 nucleotídeos finais e iniciais nos candidatos a pré-miRNAs intrônicos. Foi analisado também os últimos 50 nucleotídeos na posição 3' demonstrando se as sequências apresentariam um número elevado de pirimidinas e uma adenina (A) com potencial para a participação no mecanismo de remoção de íntrons. Observou-se que todas as 19 sequências intrônicas precursoras apresentaram os sítios conservados GU-AG nas extremidades da estrutura. Em relação aos 50 nucleotídeos próximos a extremidade 3', a maioria da sequências apresentaram muitas pirimidinas e possíveis adeninas de ramificação. Entretanto a localização dessas bases não foi compartilhada na maioria das estruturas analisadas como representado na **Figura 16**.

Esses resultados caracterizam essas estruturas como boas candidatas a pré-miRNAs intrônicos, como descritos em trabalhos anteriores (WESTHOLM & LAI, 2011; CLANCY, 2008).

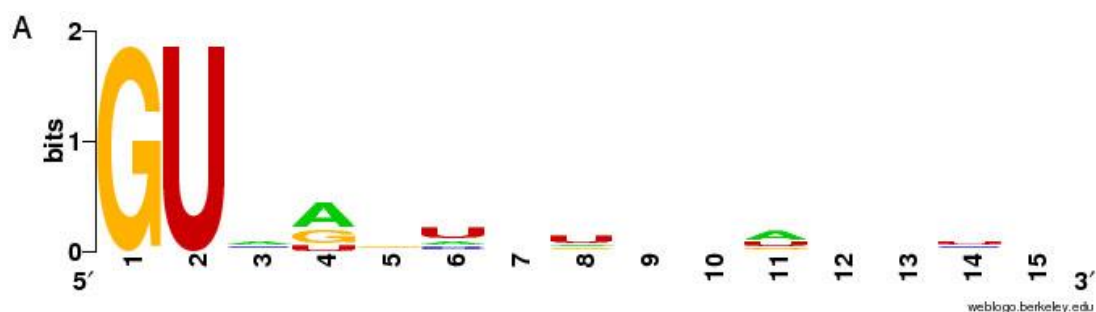


Tabela 6: Características dos miRNAs maduros preditos que foram expressos.

microRNA	Fita	Tamanho	Sequências maduras	Gene Hospedeiro	Produto do Gene
<i>sma-miR-new_i5 5p</i>	5'	22	GUGCCGGUGGUGGUAGU AGUGG	Smp_146330.1	Phospholipase C beta, putative
<i>sma-miR-new_i6 5p</i>	5'	22	AUUCUUCUGCUGCUGCU GCCGC	Smp_169260.1	Zinc finger protein, putative
<i>sma-miR-new_i6 3p</i>	3'	22	CUGC GGCCGCUGUC GCUGC	Smp_169260.1	Zinc finger protein, putative
<i>sma-miR-new_i7 5p</i>	5'	22	ACAUAAA UACUCCAAG AAAUG	Smp_105860.1	Nuclear protein SkiP, putative
<i>sma-miR-new_i7 3p</i>	3'	22	UGGAACUCCA UUCGAUG UGUUU	Smp_105860.1	Nuclear protein SkiP, putative
<i>sma-miR-new_i9 5p</i>	5'	22	ACAGCUGCUGGCUCAGU UGGUG	Smp_129330.1	Acetyl-CoA acetyltransferase, cytosolic (ACAT2)
<i>sma-miR-new_i11 5p</i>	5'	22	CACCAGGACA UAGUGCA ACAUC	Smp_163570.1	Ryanodine receptor related
<i>sma-miR-new_i11 3p</i>	3'	22	UGGGUGCACUAGUAACG GAAAC	Smp_163570.1	Ryanodine receptor related
<i>sma-miR-new_i12 3p</i>	3'	22	GUCGGGGUGUAGAUUGU UUCUA	Smp_048030.1	Hypothetical protein
<i>sma-miR-new_i13 3p</i>	3'	22	UUCGCACUCUAAUCUCA ACGCG	Smp_141560.1	Venom allergen-like (VAL) 25 protein
<i>sma-miR-new_i15 5p</i>	5'	22	CAGUGUCCUGGCUGAGG AAAAC	Smp_045940.1	Procollagen-lysine,2-oxoglutarate 5- dioxygenase, putative
<i>sma-miR-new_i15 3p</i>	3'	22	AAAGGUGGAGUCCAGGU GAGUG	Smp_045940.1	Procollagen-lysine,2-oxoglutarate 5- dioxygenase, putative
<i>sma-miR-new_i16 5p</i>	5'	22	GCAGGUCUAA CUUGGGC UCAAA	Smp_171440.1	Mindbomb, putative

<i>sma-miR-new_i17 5p</i>	5'	22	UGGAACUAAUCCUGGCG GCGGU	Smp_133330.1	Acid Sphingomyelinase Phosphodiesterase 3b
<i>sma-miR-new_i17 3p</i>	3'	22	AAAUGGUGCAUUACCUA UUUCU	Smp_133330.1	Acid Sphingomyelinase Phosphodiesterase 3b
<i>sma-miR-new_i18 5p</i>	5'	22	CAUGUGGCUCAGUAUUAU UCAUG	Smp_059250.1	Helicase, putative
<i>sma-miR-new_i18 3p</i>	3'	22	UCAUGAUUUCUUAGGUA CGCGG	Smp_059250.1	Helicase, putative
<i>sma-miR-new_i19 5p</i>	5'	22	ACACUGAGACACAUUGU UGUUC	Smp_006520.1	Cercarialelastase (S01 family)

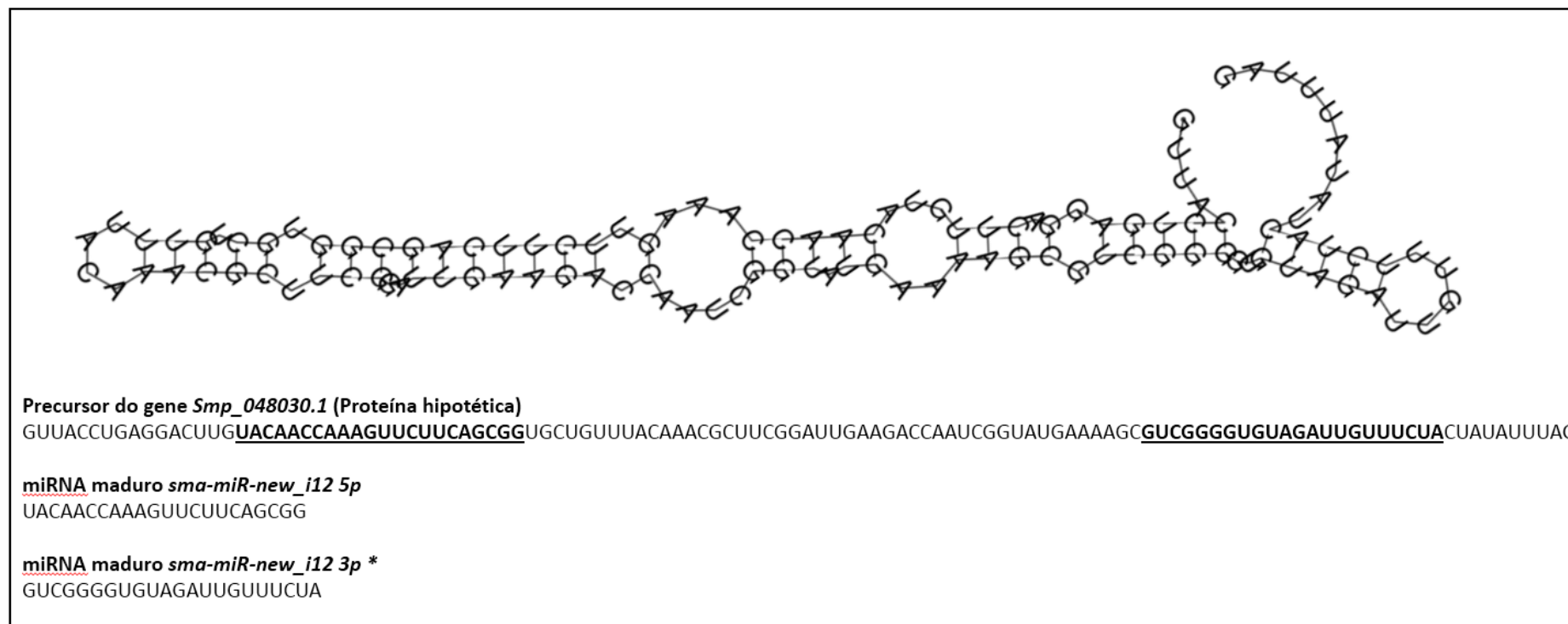


Figura 17: Estrutura precursora do *sma-miR-new_i12 3p*. Representado na figura a estrutura secundária de origem intrônica do gene *Smp_048030.1*, precursor dos miRNAs maduros *sma-miR-new_i12 5p* e *sma-miR-new_i12 3p*. Resultados da expressão relativa demonstram que somente a sequência madura *sma-miR-new_i12 3p* foi expresso.

4.2 - Expressão dos miRNAs intrônicos em *S. mansoni*.

Várias metodologias para a análise da expressão gênica de miRNAs tem sido utilizadas no gênero *Schistosoma* (WANG et al., 2010; CAI et al., 2011; SIMOES et al. 2011), sendo as principais técnicas: qRT-PCR, *Northern Blot*, e sequenciamento de nova geração. A técnica escolhida para a validação dos miRNAs maduros preditos neste trabalho foi a de qRT-PCR por apresentar uma alta sensibilidade utilizando baixas quantidades de RNA.

Foram testados todos os 38 miRNAs maduros preditos ao final do *pipeline* descrito na metodologia. Eles foram avaliados durante os seguintes estágios do parasito: cercárias, esquistossômulos mecanicamente transformados de 3,5 horas (EMT-3,5) e 24 horas (EMT-24), vermes adultos, ovos e miracídios. Observou-se que, dos 38 miRNAs preditos, somente 18 apresentaram uma expressão considerada real. Para essa validação da expressão dos miRNAs foram analisados: a curva de dissociação, os valores de Cq e os *amplicons* obtidos na qRT-PCR, visualizados no gel de agarose.

Para análise da curva de dissociação, foram descartados todos os candidatos a miRNAs que não apresentaram picos definidos plotados no gráfico, ou que apresentaram dois ou mais picos, sendo caracterizados como produtos inespecíficos (BENGESTRATE et al., 2011).

A análise no gel de agarose foi corroborada com a curva de dissociação. Quando não apresentava picos no gráfico, não era observado produto no gel, e quando apresentava dois ou mais picos em temperaturas diferentes eram encontrados produtos de tamanhos muito diferentes ao esperado para um miRNA (**Figura 18**).

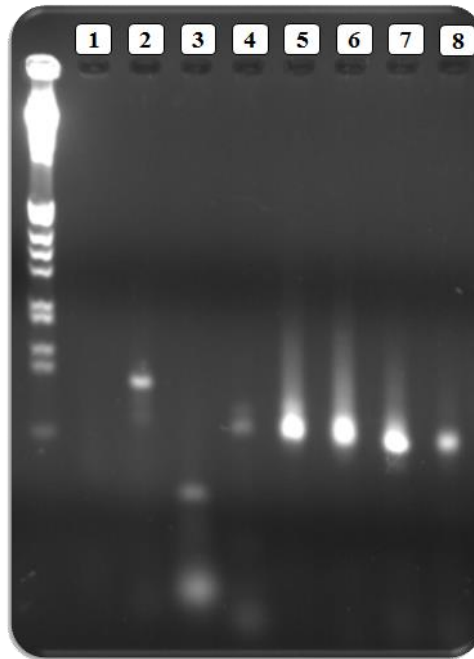


Figura 18: Gel de agarose dos miRNAs. O gel de agarose 1,2% exemplifica nas canaletas de 1-3 três candidatos a miRNAs que não apresentaram produtos de PCR esperados (respectivamente *sma-new-mir_i1 5p*, *sma-new-mir_i3 5p*, *sma-new-mir_i8 5p*). As canaletas 4-8 representam miRNAs em que apresentaram somente um pico de dissociação e um produto esperado (respectivamente *sma-new-mir_i11 3p*, *sma-new-mir_i5 5p*, *sma-new-mir_i6 5p*, *sma-new-mir_i9 5p*, *sma-new-mir_i15 5p*).

Em relação aos valores de Cq, foram considerados somente os miRNAs que apresentaram valores menores que 35. Alguns autores consideram que os valores de Cq próximos de 20 são ideais para miRNAs maduros, valores entre 25 a 35 são típicos (MESTDAGH et al., 2009) e valores de Cq > 35, são muitas vezes considerados *background* ou ruídos da técnica de qRT-PCR (MESTDAGH et al., 2008; BECKER et al., 2010).

Sendo assim, os miRNAs obtidos que seguiram os critérios acima foram categorizados em três níveis de expressão e agrupados pelos valores de Cq. A **Figura 19** mostra o perfil de expressão de 11 miRNAs com faixa de amplificação entre 30 a 35. Podemos observar que a maioria dos miRNAs apresentaram uma maior expressão em verme adulto, sendo praticamente não detectado nos estágios larvais. Também foi possível detectar que dentre esse conjunto de miRNAs, o *sma-new-miR_i11 5p*, *sma-new-miR_i15 3p* e *sma-new-miR_i12 3p*, apresentaram uma expressão maior em cercária e valores tendendo a zero no estágio de EMT-24h.

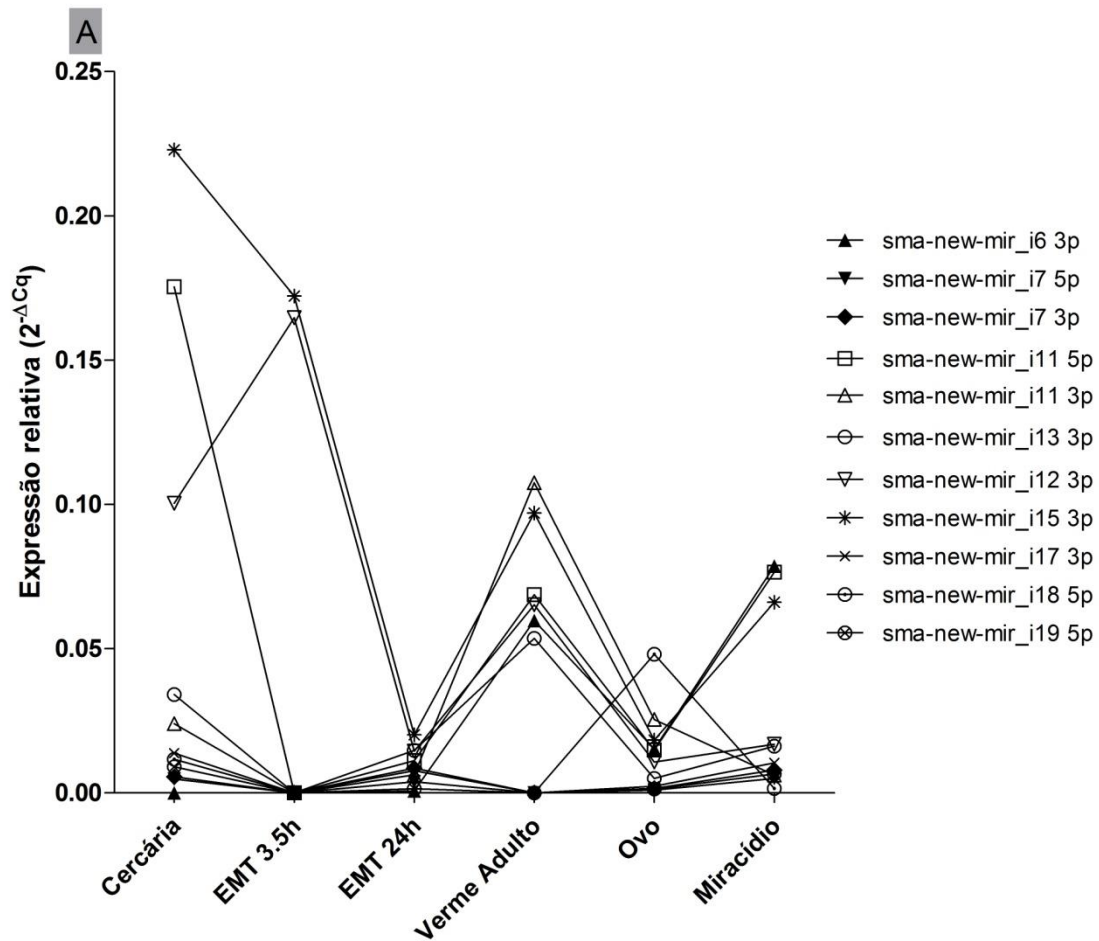


Figura 19: miRNAs com menor expressão relativa em *S. mansoni*. A expressão dos miRNAs foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. Os resultados foram multiplicados por 10^2 . Estatística não representada na figura.

Os seis miRNAs que apresentaram níveis intermediários de expressão relativa (valores de Cq entre 25 a 30) estão apresentados na **Figura 20**. Podemos observar um padrão de expressão diferencial entre os estágios do ciclo de vida do *S. mansoni*. Os miRNAs: *sma-new-miR_i6 5p*, *sma-new-miR_i18 3p* e *sma-new-miR_i16 5p* também apresentaram uma queda gradativa da expressão quando comparamos o estágio de cercária e EMT-24h. Esses resultados seguem o padrão de muitos miRNAs

identificados no trabalho realizado com *S. japonicum*, no qual foi verificada a expressão alta em cercária e quase nula em esquistossômulos (XUE et al., 2008). Nos estágios de verme adulto, a maioria dos miRNAs apresentaram um leve aumento da expressão relativa. Entretanto, esse aumento não foi significativo quando comparado entre eles, mas em comparação aos 11 miRNAs menos expressos apresentados pela **Figura 19**, esse aumento foi muito significativo. O miRNA *sma-new-miR_i6 5p* apresentou ainda, um aumento da expressão relativa em miracídio.

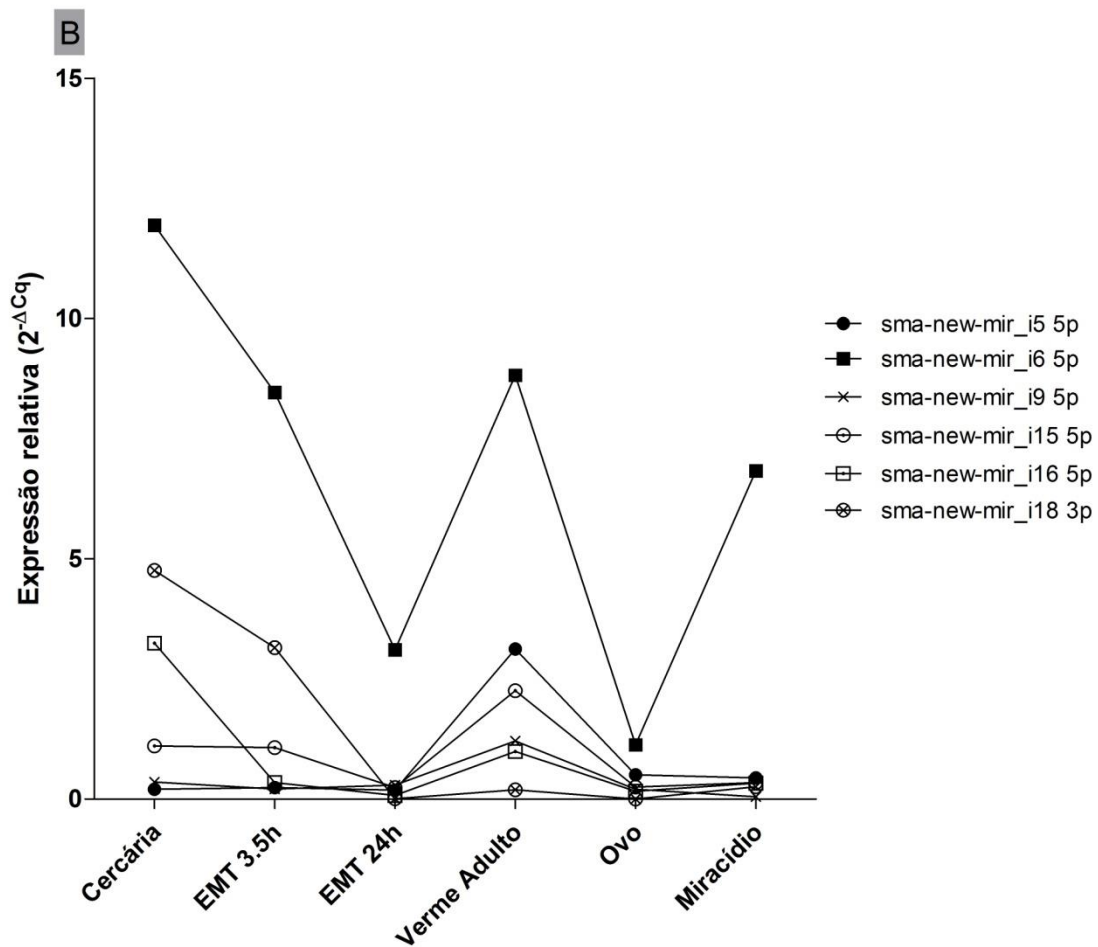


Figura 20: miRNAs com maior expressão relativa no *S. mansoni*. A expressão dos miRNAs foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios). Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. Os resultados foram multiplicados por 10^2 . Estatística não representada na figura.

Do conjunto de miRNAs cuja expressão foi avaliada por qRT-PCR, o miRNA *sma-new-miR_i17 5p* apresentou um Cq que variou de 19 a 21, dependendo do estágio evolutivo. Como podemos observar na **Figura 21**, o *sma-new-miR_i17 5p* foi cerca de 10 vezes mais expresso em verme adulto do que em cercária. Também foi observado baixos níveis de expressão em EMT24h e ovos. Em relação aos demais miRNAs avaliados neste trabalho, o *sma-new-miR_i17 5p* é pelo menos 100 vezes mais abundante em vermes adultos.

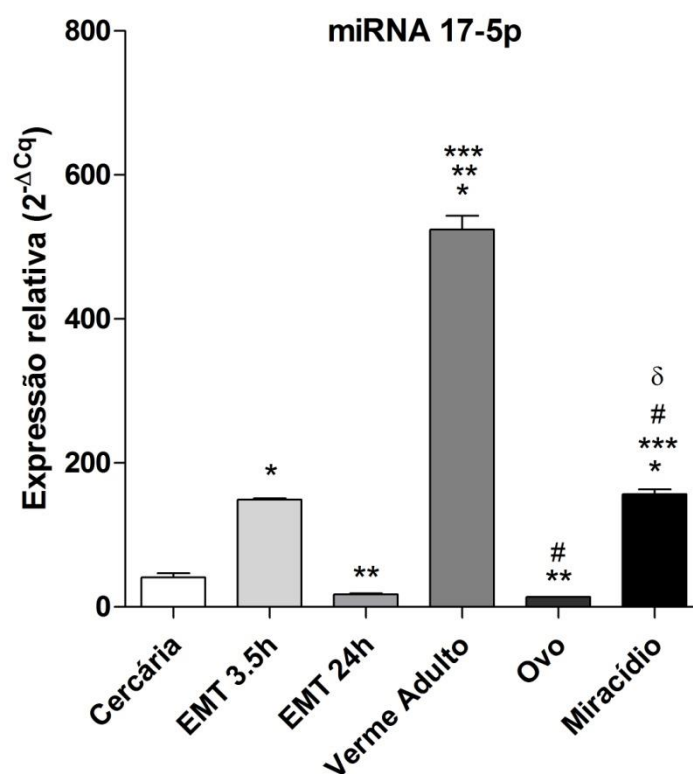


Figura 21: miRNA 17-5p é super expresso no *S. mansoni*. A expressão do miRNA 17-5p foi avaliada nos estágios de cercárias, esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas, vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *snRNAU6* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * diferente de cercária, ** diferente de EMT-3,5 horas, *** diferente de EMT-24 horas, # diferente de verme adulto, δ diferente de ovos. Os resultados foram multiplicados por 10^2 .

4.3 - Análise da expressão dos genes hospedeiros.

Considerando que os miRNAs intragênicos sejam derivados do mesmo transcrito primário e sendo assim, co-expressos com os seus genes do hospedeiros (BASKERVILLE & BARTEL 2005), foi investigada a expressão de cada um dos genes hospedeiros com a finalidade de verificar e comparar com a expressão dos miRNAs de origem intrônica.

A técnica de qRT-PCR foi utilizada para avaliar a expressão dos genes hospedeiros dos 18 miRNAs de *S. mansoni* validados anteriormente. Foram analisados durante as seguintes fases do parasito: cercárias, EMT-3,5h, EMT-24h, vermes adultos, ovos e miracídios, sendo os dados normalizados com o gene *eIF4E* para todos eles. O padrão de expressão desses genes foi apresentado nas figuras 22, 23 e 24.

Analisando a **Figura 22**, verifica-se que o transcrito da *Phospholipase C beta* (*Smp_146330*) foi expresso em todos os estágios com uma “down regulation” em ovos sendo em média três vezes mais baixo que quando comparados a miracídio por exemplo.

O transcrito da proteína *Zinc Finger* (*Smp_169260*) apresentou no geral uma expressão quase nula, sendo o estágio mais expresso o de verme adulto. Proteínas dessa família já foram descritas por participar em processos de regulação, além de serem identificadas como importante classe de fatores de transcrição no parasito na fase de verme adulto (BITAR et al., 2013; DRUMMOND et al., 2008).

O gene da proteína *Nuclear Skip* (*Smp_105860*) não apresentou muitas diferenças no seu perfil de expressão, sendo menos expresso em verme adulto e ovos. Já o gene *Smp_129330* apresentou uma expressão mais alta nos estágios iniciais, com o maior aumento em EMT-3,5h sendo cerca de quatro vezes maior que verme adulto, ovo e miracídio.

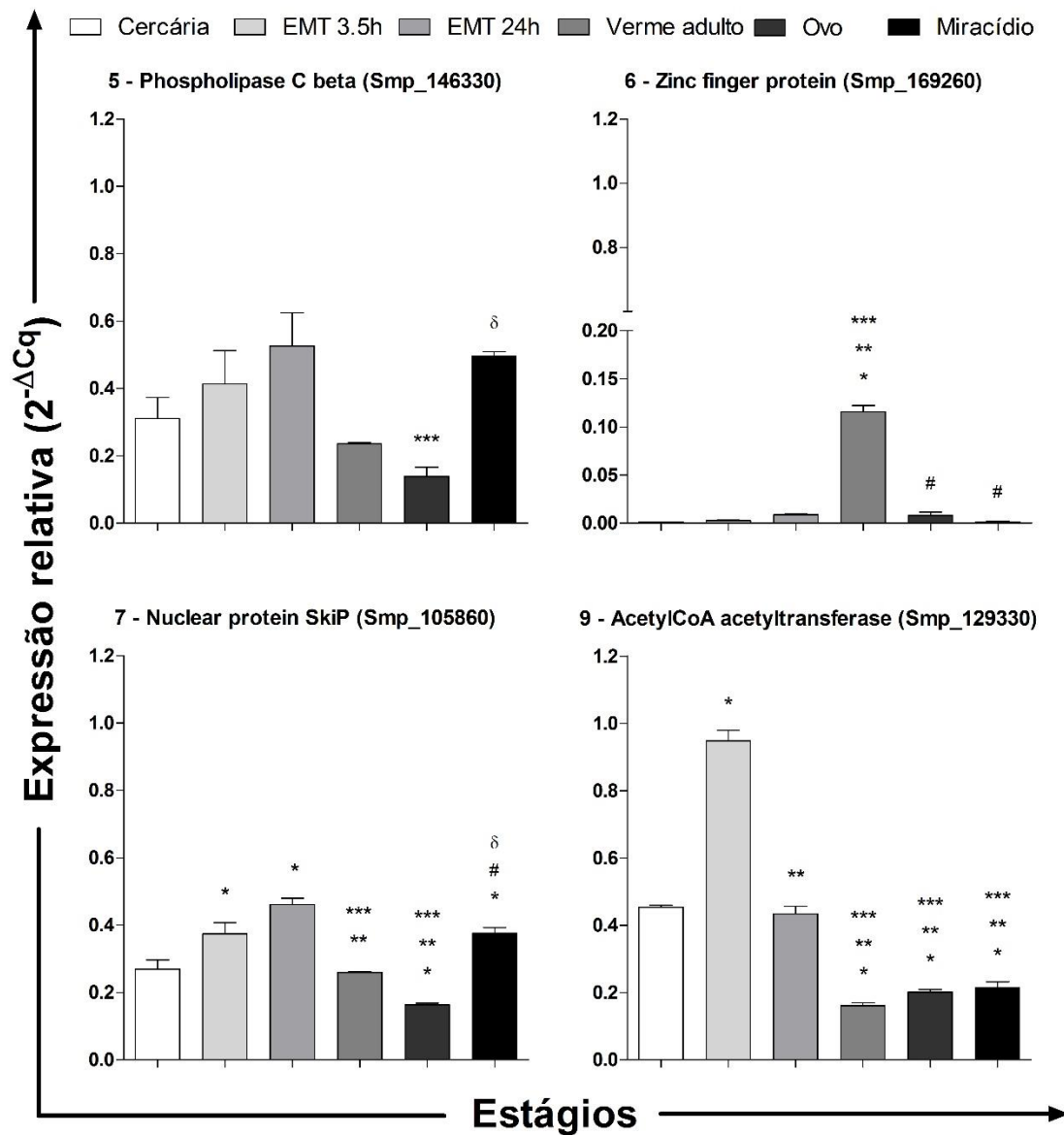


Figura 22: Expressão relativa dos genes hospedeiros (5, 6, 7, 9) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

Na **Figura 23**, todos os transcritos *Ryanodine receptor* (*Smp_163579*), *Hypothetical protein* (*Smp_048030*) e *Venom allergen-like - VAL* (*Smp_141560*) apresentaram um perfil de expressão muito baixo durante os estágios. Trabalhos anteriores utilizando 28 genes de proteínas VAL (PARKER-MANUEL et al., 2011, CHALMERS et al., 2007) corroboram com o resultado relativo a expressão do gene *Smp_141560* da proteína VAL 25 demonstrando uma baixa expressão nos diferentes estágios. O gene *Procollagen-lysine* (*Smp_045940*) apresentou um perfil diferente dos demais, com uma baixa gradativa da expressão passando pelos estágios de cercária até miracídio.

Finalmente na **Figura 24**, os quatro genes apresentam praticamente o mesmo perfil de expressão no qual foi mais elevado nos estágios de cercária e EMT-24h, e mantiveram basicamente constantes durante os outros estágios. Entre esses genes, destaca-se o *Mindbomb* (*Smp_171440*) no qual é uma ubiquitina-ligase, que é essencial para a ativação eficiente de sinalização da via *Notch*. Essa via já foi descrita em *Schistosoma mansoni* (MAGALHÃES, 2009), apresentando uma importância crítica na diferenciação celular do parasito.

Para a maioria dos genes hospedeiros analisados, ainda não se conhece na literatura o perfil de expressão relativa de seus transcritos durante os estágios do parasito estudados.

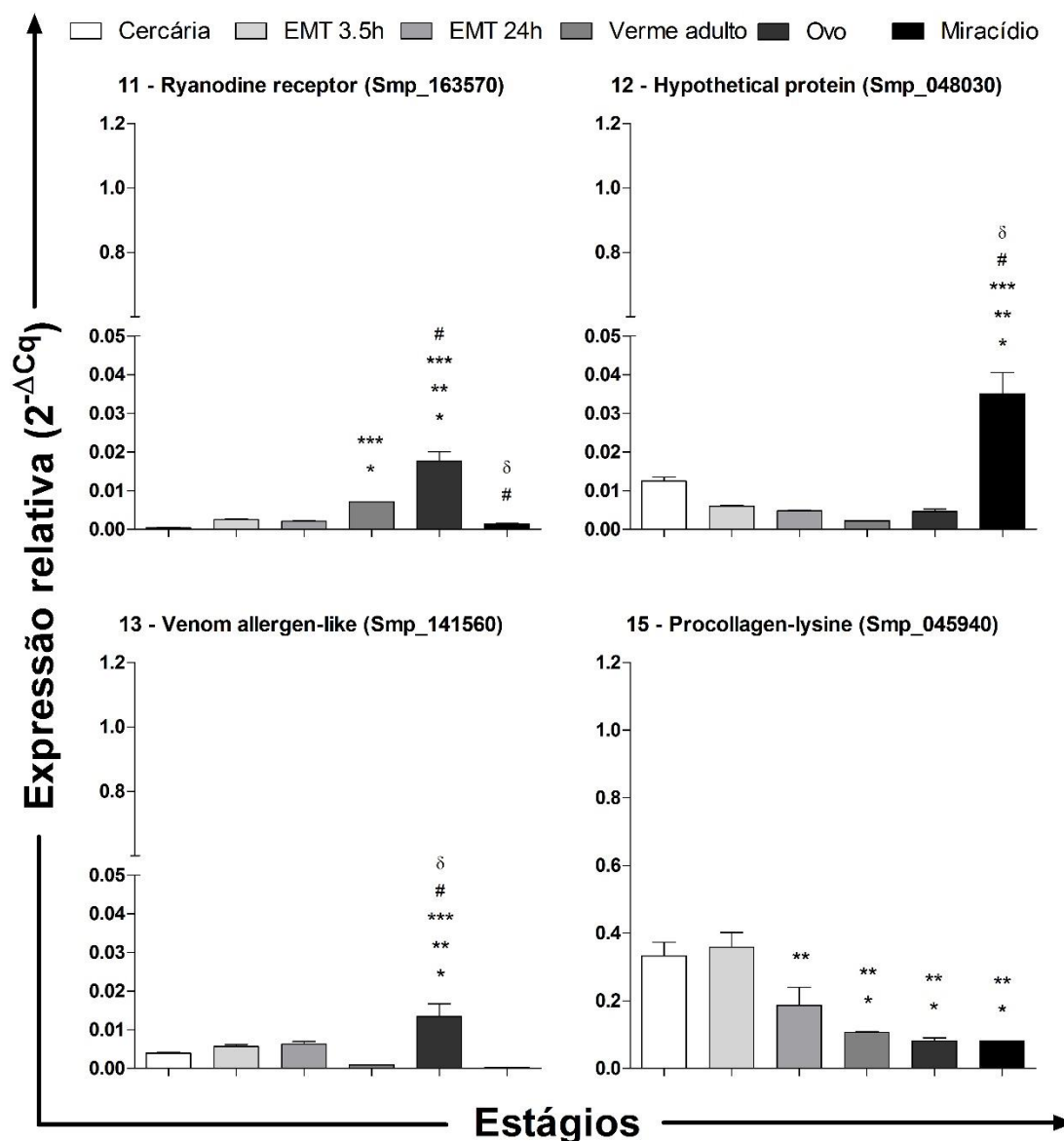


Figura 23: Expressão relativa dos genes hospedeiros (11, 12, 13, 15) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

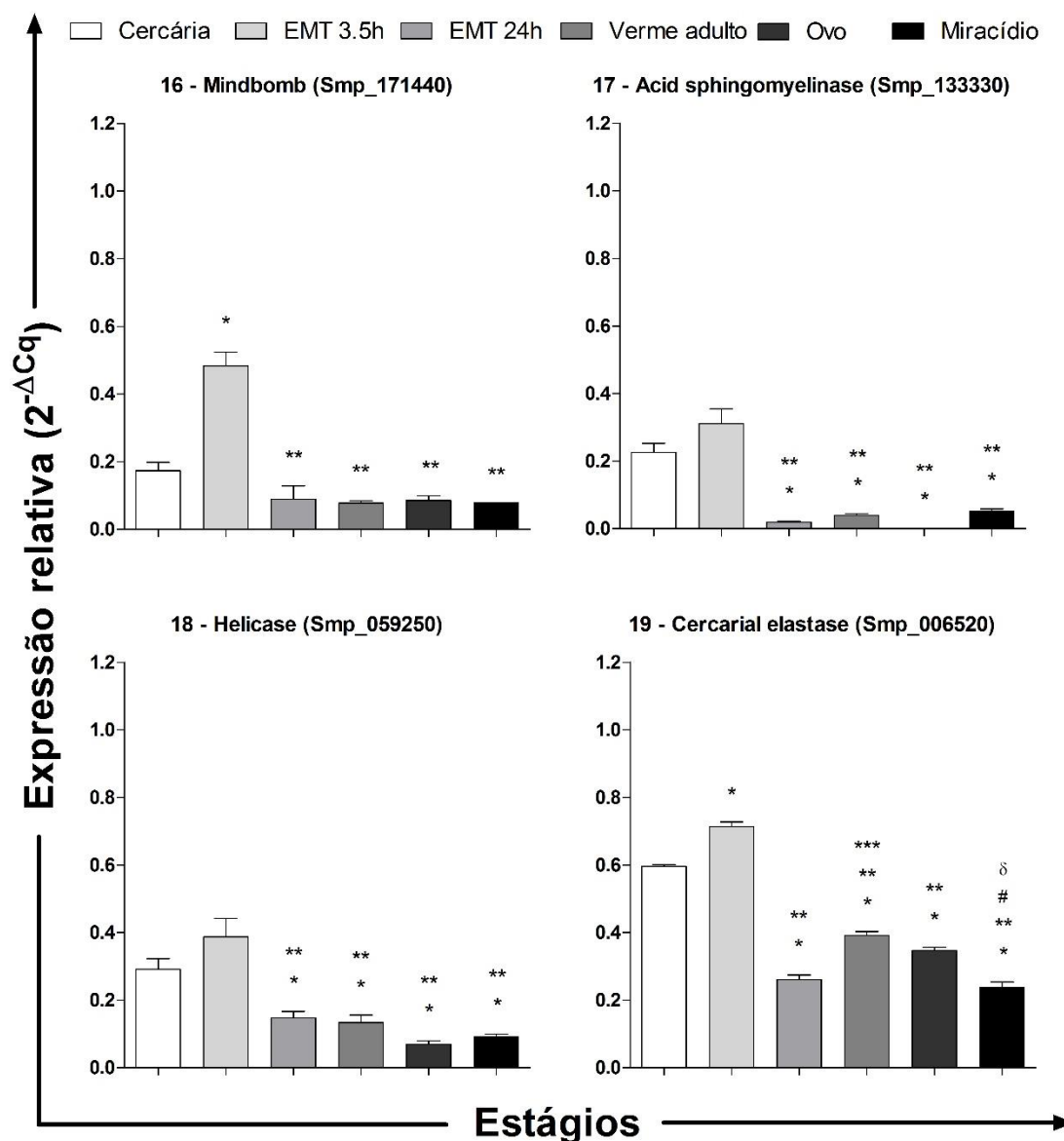


Figura 24: Expressão relativa dos genes hospedeiros (16, 17, 18, 19) dos miRNAs. Os níveis de transcritos dos genes foram avaliados nos estágios de cercárias, EMT: esquistossômulos mecanicamente transformados (EMT-3,5 horas, EMT-24 horas), vermes adultos, ovos e miracídios. Como gene constitutivo foi utilizado o *eIF4E* e a expressão gênica relativa, determinada pelo método do $2^{-\Delta Cq}$. A análise estatística foi realizada utilizando o teste de ANOVA *oneway* (Tukey) com $p < 0,05$ no programa GraphPad Prism 5. As diferenças estatísticas entre os estágios foram determinadas pelos sinais: * comparado com cercária, ** comparado com EMT-3,5 horas, *** comparado com EMT-24 horas, # comparado com verme adulto, δ comparado com ovos.

4.4 - Co-regulação entre miRNA intrônicos e seus genes hospedeiros

Pouco se sabe sobre a co-regulação dos miRNAs intrônicos e seus genes hospedeiros em uma célula. Acredita-se que quando transcritos, esses miRNAs sigam a mesma orientação do seu pré-mRNA, podendo assim estar sobre o controle do promotor do gene hospedeiro (BASKERVILLE & BARTEL, 2005). Entretanto, estudos recentes têm demonstrado que os miRNAs intrônicos podem ser transcritos de forma independente do seu gene hospedeiro (HE, et al., 2012; LEE et al., 2008; OBERNOSTERER, 2006). A alternativa possível para isso acontecer é que o miRNA intrônico seja transcrito por meio de elementos de regulação próprios, principalmente quando apresentam tamanhos muito grandes (XIE et al., 2005).

Para investigar se a expressão dos miRNAs foram dependentes dos seus genes hospedeiros, foi analisado a correlação entre os seus perfis de expressão utilizando o coeficiente de Pearson. Dos 18 miRNAs analisados somente os miRNAs *sma-new-miR_i15-5p* e *sma-new-miR_i18-3p* apresentaram uma correlação positiva entre os seis estágios de desenvolvimento do parasito como mostra a **Figura 25**. Nessa correlação, o miRNA acompanhou a expressão do seu gene hospedeiro de maneira diretamente proporcional.

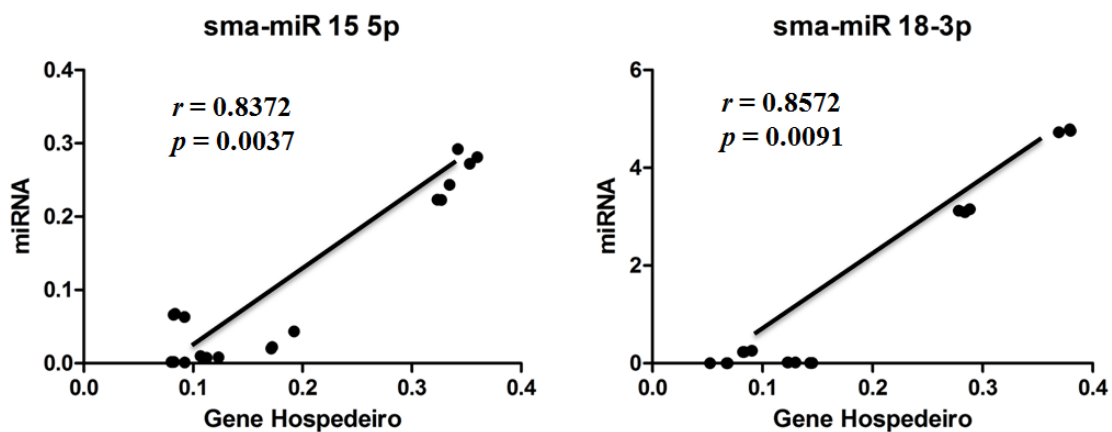


Figura 25: Correlação de Pearson entre miRNAs e os seus genes hospedeiros. O nível de expressão relativa dos pares miRNAs - Genes hospedeiros foram analisados pelo coeficiente de Pearson (r). Os miRNAs *sma-new-miR_i15-5p* e *sma-new-miR_i18-3p* apresentaram uma correlação positiva em relação aos seus genes hospedeiros.

Os demais miRNAs não apresentaram nenhuma correlação (negativa ou positiva) quando comparados ao seus genes hospedeiros. Dentre as justificativas que explicariam esse fato, foi que eles poderiam apresentar promotores próprios sendo que alguns desses miRNAs estão presentes em transcritos grandes apresentando regiões intrônicas muito extensas.

Um estudo recente demonstrou que de um conjunto de miRNAs conservados e não conservados analisados, os considerados miRNAs novos (não conservados) apresentaram uma taxa de co-expressão com o seu gene hospedeiro significativamente menor, quando comparado com os miRNAs conservados (HE et al., 2012). Um dos motivos que ocasionaria esse fato pode ser a ocorrência de regulação ao nível da cromatina. Eles sugeriram assim, que a co-expressão de miRNA intragênicos com seus genes hospedeiros possa estar fortemente relacionada com a sua evolução e grau de conservação entre as espécies.

Outra possibilidade é que os genes hospedeiros possam ser regulados por outros miRNAs ou até mesmo ser auto-regulados. Mesmo que o gene hospedeiro e seu miRNA intrônico sejam expressos a partir do mesmo promotor, eles podem apresentar níveis de expressão diferenciados entre si, devido a essa regulação pós-transcricional (RADFAR et al., 2011; OBERNOSTERER et al., 2006; THOMSON et al., 2006). A auto regulação, não foi identificada entre os alvos e miRNAs analisados, mas a alternativa de serem regulados por outros miRNAs ainda não descritos é possível.

4.5 - Análise funcional *in silico* - interação do miRNA com o gene alvo

A predição dos alvos dos miRNAs validados foi realizada utilizando o arquivo GFF disponível no banco de dados do genoma do *S. mansoni*. Atualmente somente 6261 genes dos 10.852 apresentam as regiões 3'UTRs mapeadas nesse arquivo. O programa de predição utilizado foi o miRanda que baseia-se no alinhamento entre miRNAs - alvos candidatos e na energia de hibridização entre eles (LINDOW, 2011). Assim os parâmetros foram ajustados com a finalidade de aumentar a estringência da predição sendo considerados apenas os valores de *score* mínimo de 130 e delta G mínimo de -15 Kcal/mol (YOICHI et al., 2008). A partir desse algoritmo, foi possível fazer a predição de 993 alvos dentre os genes mapeados, sendo os miRNAs que

apresentaram mais alvos os *sma-miR-new-i15-3p*, *sma-miR-new-i5-5p*, *sma-miR-new-i19-5p* e *sma-miR-new-i11-5p*. (**Tabela 7**).

Esse número representa aproximadamente 10% dos genes de todo o genoma do *S. mansoni*, sendo alvos potenciais nos quais estão envolvidos muitos processos importantes no desenvolvimento do parasito. Da lista de 18 miRNAs, 12 apresentaram alvos em potencial da via ubiquitina-proteassoma que é muito importante para o desenvolvimento do parasito. Dentre esses miRNAs, destacam-se os que apresentaram mais alvos para essa via sendo o *sma-miR-new 9 5p*, *sma-miR-new 11 5p*, *sma-miR-new 15 3p*, *sma-miR-new 5 5p*. Foram cerca de 4 alvos distintos para cada miRNA. Esses resultados demonstram grande importância para biologia do parasito, pois como já elucidado em alguns estudos, a inibição de genes dessa via foi capaz de reduzir o número de esquistossômulos, a carga de vermes adultos e consequentemente, diminuir a produção de ovos em modelos murinos experimentais (GUERRA-SÁ et al., 2005; MATHIESON et al., 2011). Foi verificado também mRNA alvos para genes que apresentam uma analogia com a via de ubiquitinação. Nessa via estão presentes moléculas com homologia com a ubiquitina, como a NEED8, que também apresentou em um estudo, importante ação regulatória celular durante o desenvolvimento de estágios larvais como cercárias e esquistossômulos até o estágio de verme adulto (PEREIRA et al., 2013).

É importante ressaltar que dentre todos os alvos preditos, 323 codificam proteínas hipotéticas, ou seja, que não apresentam similaridade com proteínas descritas em outros organismos. Esse fato é interessante, pois dentre esses alvos podemos encontrar proteínas específicas do parasito que consequentemente apresentariam funções essenciais no desenvolvimento do mesmo.

Para representar a relação dos 18 miRNAs e seus respectivos mRNA alvos, foi realizada uma rede de interação antes eles a partir do aplicativo Cytoscape 3.0 (**Figura 26**). Nela foi possível visualizar que muitos alvos são regulados por dois ou mais miRNAs ao mesmo tempo, como já demonstrado em diversos estudos (LIU et al., 2010; SATOH & TABUNOKI, 2010; GENNARINO et al., 2012; LU et al., 2011; PETER, 2001). A força de interação entre o miRNA - Alvo foi calculada a partir dos valores de ΔG que variaram entre -15 a -33,8 kcal/mol. Quanto menor o valor, mais provável foi a predição da ligação do miRNA com os seus alvos.

Tabela 7: Número de alvos dos miRNAs preditos e validados.

miRNA	Nº de alvos	Alvo mais provável no miRanda
sma-miR-new_i5_5p	100	Far upstream (fuse) binding protein, putative (Smp_044550.2)
sma-miR-new_i6_3p	8	Far upstream (fuse) binding protein, putative (Smp_044550.1)
sma-miR-new_i6_5p	65	Hypothetical protein (Smp_103340.1)
sma-miR-new_i7_3p	59	RNA lariat debranching enzyme, putative (Smp_007190)
sma-miR-new_i7_5p	29	Groucho related (Smp_165280)
sma-miR-new_i9_5p	46	Hypothetical protein (Smp_002920.1)
sma-miR-new_i11_3p	57	Hypothetical protein (Smp_038700.1)
sma-miR-new_i11_5p	86	Hypotheticalprotein (Smp_121440)
sma-miR-new_i12_3p	41	Protein C10orf118 (CTCL_tumor_antigen_HD-CL-_01_L14-2), putative (Smp_034940.3)
sma-miR-new_i13_3p	25	Threonine dehydratasedeaminase, putative (Smp_137520)
sma-miR-new_i15_3p	144	Hypothetical protein (Smp_016240)
sma-miR-new_i15_5p	57	Hypothetical protein (Smp_178030)
sma-miR-new_i16_5p	25	CD36-like class B scavenger receptor (Smp_011680.1)
sma-miR-new_i17_3p	29	Hypothetical protein (Smp_184780)
sma-miR-new_i17_5p	48	Hypothetical protein (Smp_123330)
sma-miR-new_i18_3p	80	Aspartyl-tRNA synthetase, putative (Smp_053510.2)
sma-miR-new_i18_5p	47	F-box_only_protein, putative (Smp_150070.1)
sma-miR-new_i19_5p	96	Hypothetical protein (Smp_130440)

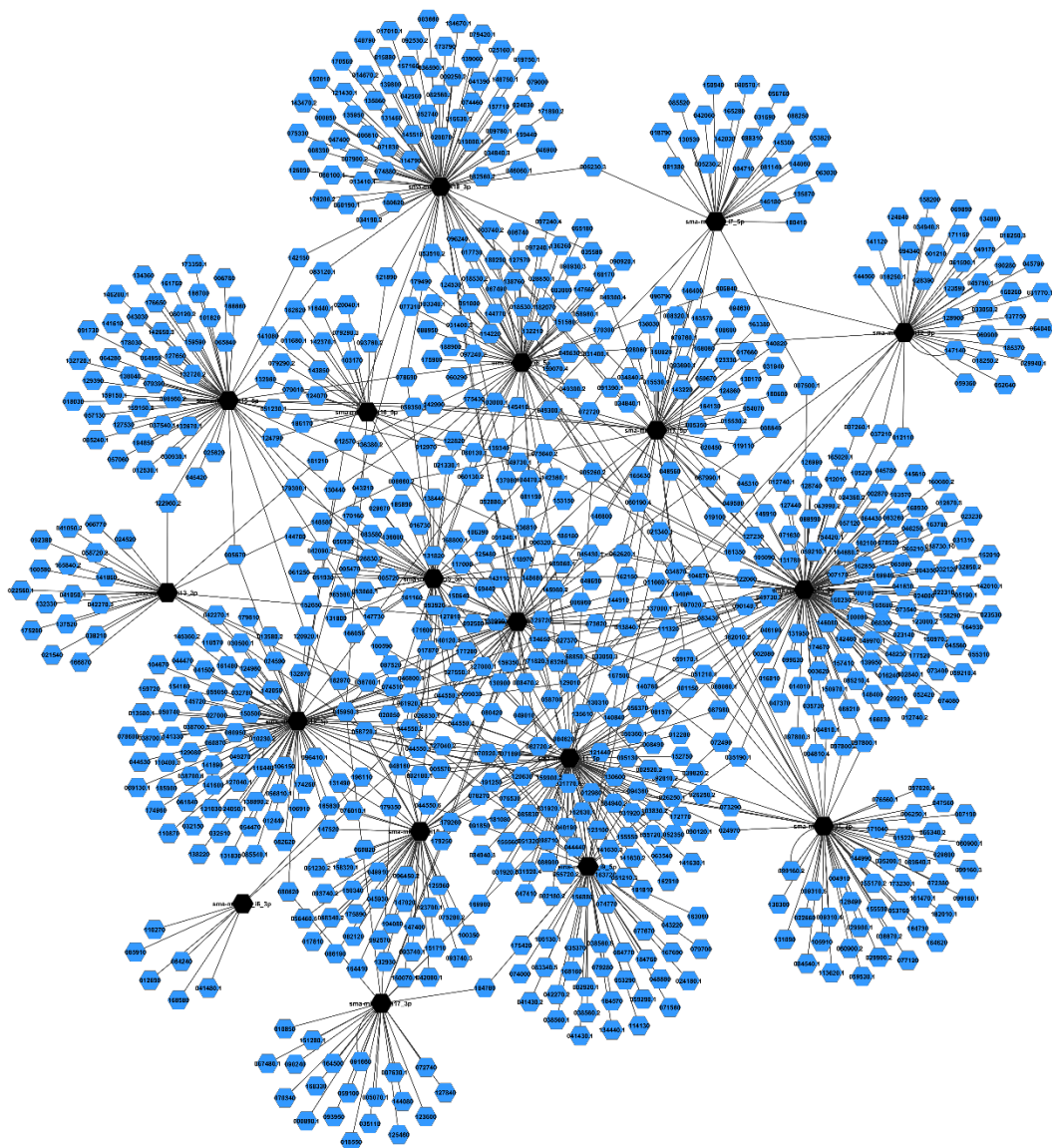


Figura 26: Rede de interação ente o miRNA e seus alvos. A rede foi criada a partir do programa Cytoscape 3.0 entre os 18 miRNAs (símbolos de cor preta) e os seus 993 alvos mRNAs (símbolos de cor azul). Quanto mais próximo o alvo ficou do miRNA, menor foi o valor de ΔG indicando uma maior possibilidade de ligação dele com seus alvos.

5 - Conclusão

Dentre os principais resultados desse trabalho, destacam-se:

- 18 candidatos a miRNAs não conservados identificados em *S. mansoni*;
- 993 alvos preditos para os miRNAs intrônicos representando quase 10% do genoma do parasito, sendo 323 constituídos de proteínas hipotéticas;
- Um miRNA *sma-miR-new-i17-5p* apresenta uma super expressão em *S. mansoni*;
- Somente 2 miRNAs (*sma-new-miR_i15-5p* e *sma-new-miR_i18-3p*) apresentaram co-expressão em relação aos seus genes hospedeiros;

Sendo assim, esse conjunto de resultados sugerem que o *S. mansoni* apresente também como alternativa para a biogênese de miRNAs, a via de biogênese intrônica. Essa via aumenta ainda mais a possibilidade de regulação ao nível pós-transcricional por miRNAs nos diferentes estágios do parasito, contribuindo para a sua adaptação aos diferentes ambientes.

6 - Perspectivas

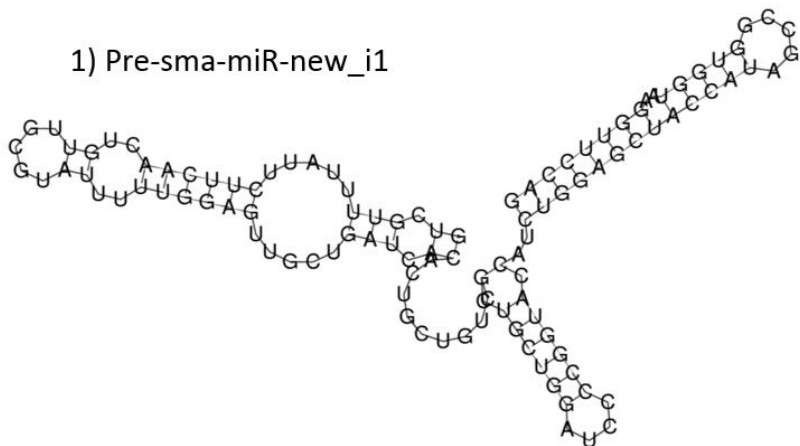
Como perspectivas desse estudo, futuros experimentos serão realizados para:

- 1) Validar as estruturas precursoras dos miRNAs identificados nesse trabalho;
- 2) Utilizar novos softwares de predição para descartar a hipótese que não tenha isoformas desses miRNAs;
- 3) Validar experimentalmente os principais alvos preditos dos miRNAs e correlacionar as suas expressões relativas (miRNA-Alvo).

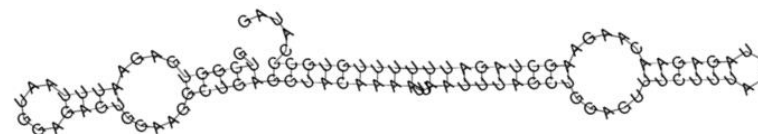
7 - Anexos

Anexo A: Estruturas secundárias dos 19 candidatos a Pré-miRNA de *Schistosoma mansoni*.

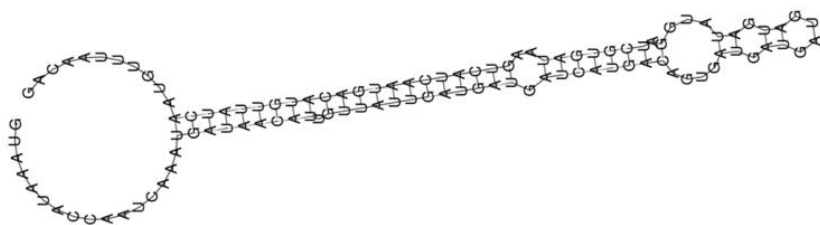
1) Pre-sma-miR-new_i1



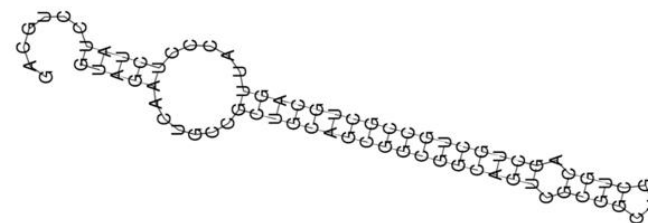
2) Pre-sma-miR-new_i2



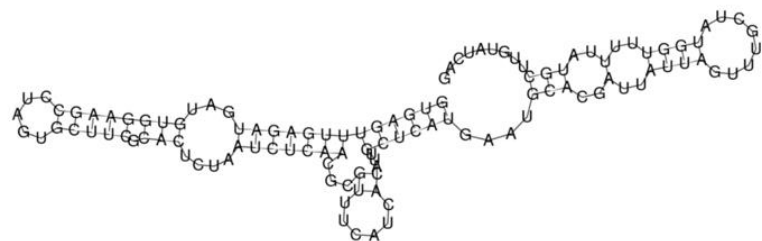
3) Pre-sma-miR-new_i3



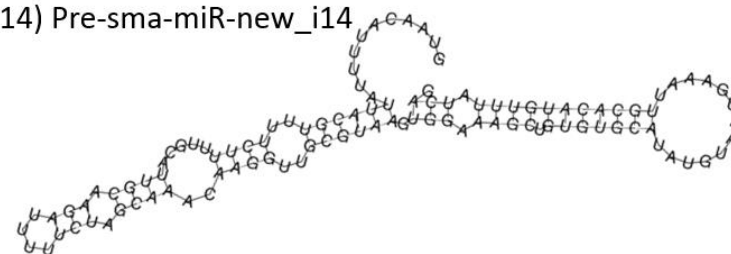
4) Pre-sma-miR-new_i4



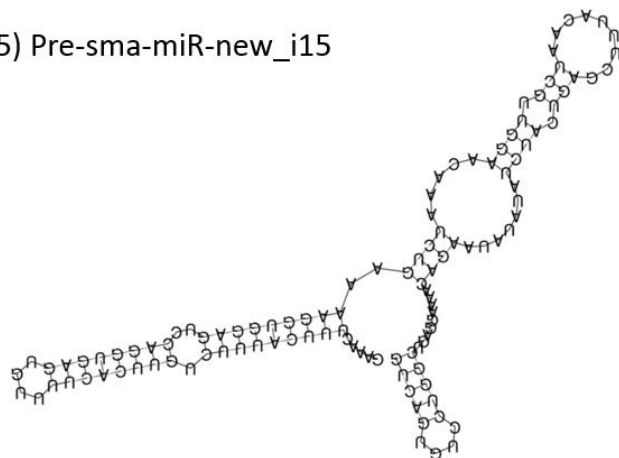
13) Pre-sma-miR-new_i13



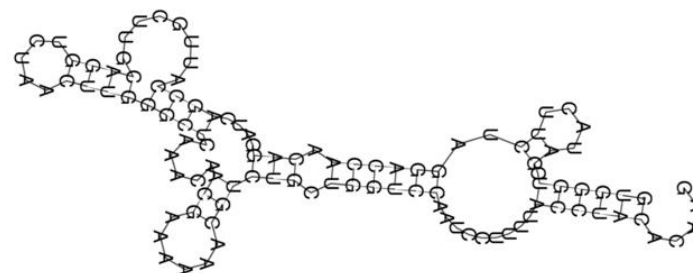
14) Pre-sma-miR-new_i14



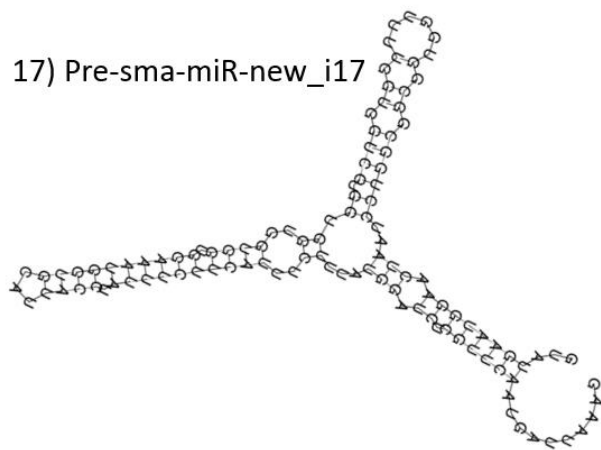
15) Pre-sma-miR-new_i15



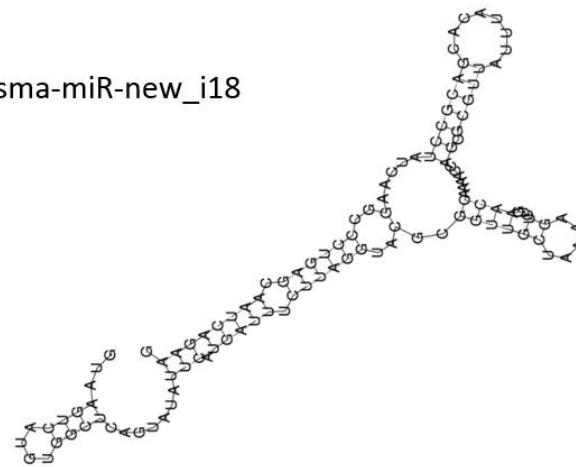
16) Pre-sma-miR-new_i16



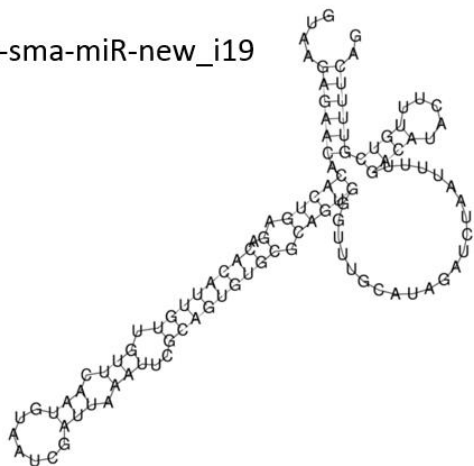
17) Pre-sma-miR-new_i17



18) Pre-sma-miR-new_i18



19) Pre-sma-miR-new_i19



Anexo B: Scripts desenhados na linguagem de programação Perl.

- 1) Extração das posições das sequências após a metodologia utilizada na plataforma galaxy:

```
#!/usr/bin/perl
#encontra a linha do arquivo de saida inv e a linha dos numeros

open(MYINPUTFILE, "<depois-galaxy-new.gff");
open(MYOUTPUTFILE, ">depois-galaxy-new.gff.out");

#print MYOUTPUTFILE "Gene\tRefSeq\tOrganism\tSize\tSequence\n" ;

while(<MYINPUTFILE>)
{
my($line) = $_;
chomp($line);
if ( $line =~ m/exon exon/ig)
{
print MYOUTPUTFILE "$line\n" ;
}
}

close(MYINPUTFILE);
close(MYOUTPUTFILE);

#$act8 = `perl -p -i -e s/'\t+\t/' '\tplus\t'/g
v4.19.05.11.chado.filtered.gff.out`;

#$act9 = `perl -p -i -e s/'\t-\t/' '\tminus\t'/g
v4.19.05.11.chado.filtered.gff.out`;
```

- 2) Encontra a linha do arquivo de saída e a linha dos números:

```
#!/usr/bin/perl
#encontra a linha do arquivo de saida inv e a linha dos numeros

open(MYINPUTFILE, "<seqs-entre-60-120-new.gff");
open(MYOUTPUTFILE, ">seqs-entre-60-120-new.gff.pl");

#print MYOUTPUTFILE "Gene\tRefSeq\tOrganism\tSize\tSequence\n" ;
#$act8 = `perl -p -i -e s/'\t+\t/' '\tplus\t'/g
v4.19.05.11.chado.filtered.gff.out`;
#$act9 = `perl -p -i -e s/'\t-\t/' '\tminus\t'/g
v4.19.05.11.chado.filtered.gff.out`;

while(<MYINPUTFILE>)
{
```

```

my($line) = $_;
chomp($line);

#ID=Smp_186980.1:exon:1;Name=Smp_186980.1:exon:1;Parent=Smp_1869
80:mRNA;RP=1      11220 12411+      exon exon schisto_mansoni.Chr_1
      1191.0
#ID=Smp_173620.1:exon:7;Name=Smp_173620.1:exon:7;Parent=Smp_1736
20:mRNA;RP=1644 30803377 30803435 +      exon exon
      Schisto_mansoni.Chr_1 58.0

#Chromossome 1 plus

if ( $line =~
m/ID=(Smp_.+?exon:.) .+?\t(\d{1,})\t(\d{1,})\t\+\t.+(Schisto_man
soni.Chr_1)\t(\d{1,})/ig)
{
if ($2<30000000)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.1.fasta -reg ",
"$start..$end", " stdout -separate >$1_$4_size_$5_plus.txt`";
"\n" ;
}
if ($2>30000000)
{
$somastart=$2-30000000;
$somaend=$3-30000000;
$start=$somastart+1;
$end=$somaend-1;

print MYOUTPUTFILE "\$act \= \`extractseq $4.2.fasta -reg ",
"$start..$end", " stdout -separate >$1_$4_size_$5_plus.txt`";
"\n" ;
}
}

#Chromossome 1 minus

if ( $line =~ m/ID=(Smp_.+?exon:.) .+?\t(\d{1,})\t(\d{1,})\t\+
\t.+(Schisto_mansoni.Chr_1)\t(\d{1,})/ig)
{
if ($2<30000000)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.1.fasta -reg ",
"$start..$end", " stdout -separate >$1_$4_size_$5_minus.txt`";
"\n" ;
}
if ($2>30000000)
{
$somastart=$2-30000000;
$somaend=$3-30000000;
$start=$somastart+1;
$end=$somaend-1;

```

```

print MYOUTPUTFILE "\$act \= \`extractseq $4.2.fasta -reg ",
"$start..$end'", " stdout -separate >$1_$4_size_$5_minus.txt`";",
"\n" ;
}
}

#Chromossome W plus

if ( $line =~
m/ID=(Smp_.+?exon:.)+?\t(\d{1,})\t(\d{1,})\t+\t.+(Schisto_manson
i.Chr_W)\t(\d{1,})/ig)
{
if ($2<30000000)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.1.fasta -reg ",
"$start..$end'", " stdout -separate >$1_$4_size_$5_plus.txt`";",
"\n" ;
}
if ($2>30000000)
{
$somastart=$2-30000000;
$somaend=$3-30000000;
$start=$somastart+1;
$end=$somaend-1;

print MYOUTPUTFILE "\$act \= \`extractseq $4.2.fasta -reg ",
"$start..$end'", " stdout -separate >$1_$4_size_$5_plus.txt`";",
"\n" ;
}
}

#Chromossome W minus

if ( $line =~ m/ID=(Smp_.+?exon:.)+?\t(\d{1,})\t(\d{1,})\t-\t.+(Schisto_manson
i.Chr_W)\t(\d{1,})/ig)
{
if ($2<30000000)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.1.fasta -reg ",
"$start..$end'", " stdout -separate >$1_$4_size_$5_minus.txt`";",
"\n" ;
}
if ($2>30000000)
{
$somastart=$2-30000000;
$somaend=$3-30000000;
$start=$somastart+1;
$end=$somaend-1;

```

```

print MYOUTPUTFILE "\$act \= \`extractseq $4.2.fasta -reg ",
"$start..$end"," stdout -separate >$1_$4_size_$5_minus.txt`";",
"\n" ;
}
}

#Remaining sequences

if ( $line =~
m/ID=(Smp_.+?exon:.) .+?\t(\d{1,})\t(\d{1,})\t+\t.+?(Schisto.+?)
\t(\d{1,})/ig)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.fasta -reg ",
"$start..$end"," stdout -separate >$1_$4_size_$5_plus.txt`";",
"\n" ;
}
elsif ( $line =~ m/ID=(Smp_.+?exon:.) .+?\t(\d{1,})\t(\d{1,})\t\
\t.+?(Schisto.+?)\t(\d{1,})/ig)
{
$start=$2+1;
$end=$3-1;
print MYOUTPUTFILE "\$act \= \`extractseq $4.fasta -reg ",
"$start..$end"," stdout -separate >$1_$4_size_$5_minus.txt`";",
"\n" ;
}
}

close(MYINPUTFILE);
close(MYOUTPUTFILE);

$sact = `perl -p -i -e s/'Chr'/'chr'/g seqs-entre-55-120-
new.gff.pl`;
$sact = `perl -p -i -e s/'Chr_W'/'chr_w'/g seqs-entre-55-120-
new.gff.pl`;
$sact = `perl -p -i -e s/'_W'/'_w'/g seqs-entre-55-120-
new.gff.pl`;
$sact = `perl -p -i -e s/'SC_'/'sc_'/g seqs-entre-55-120-
new.gff.pl`;
$sact = `perl -p -i -e s/'Schisto_mansoni'/'schisto_mansoni'/g
seqs-entre-55-120-new.gff.pl`;
$sact = `perl -p -i -e s/' :exon:'/'_exon_'/g seqs-entre-55-120-
new.gff.pl`;

#$sact = `perl seqs-entre-60-120-new.gff.pl`;

open(MYINPUTFILE, "<seqs-entre-60-120-new.gff.pl");
open(MYOUTPUTFILE, ">seqs-entre-60-120-new.gff-1.pl");

#print MYOUTPUTFILE "Gene\tRefSeq\tOrganism\tSize\tSequence\n" ;

```

```

#Sact8 = `perl -p -i -e s/'\t+\t'/'\tplus\t'/g
v4.19.05.11.chado.filtered.gff.out`;
#Sact9 = `perl -p -i -e s/'\t-\t'/'\tminus\t'/g
v4.19.05.11.chado.filtered.gff.out`;

while(<MYINPUTFILE>)
{
my($line) = $_;
chomp($line);

#ID=Smp_186980.1:exon:1;Name=Smp_186980.1:exon:1;Parent=Smp_1869
80:mRNA;RP=1 11220 12411+ exon exon schisto_mansoni.Chr_1
1191.0
#ID=Smp_173620.1:exon:7;Name=Smp_173620.1:exon:7;Parent=Smp_1736
20:mRNA;RP=1644 30803377 30803435 + exon exon
Schisto_mansoni.Chr_1 58.0

if ( $line !~ m/schisto_mansoni.chr_(1|w).fasta -reg/ig)
{
print MYOUTPUTFILE "$line", "\n" ;
}

}

close(MYINPUTFILE);
close(MYOUTPUTFILE);

```

3) RNAfold para fasta:

```

#!/usr/bin/perl -w

open(ENTRADA, "seqfitr_estrut_miRNA");
open(SAIDA, '>seqfitr_estrut_miRNA_fasta2');

while($linha = <ENTRADA>){
    if($linha =~ /^>/){
        print SAIDA $linha;
    }elseif($linha =~ /[AUCG]+/){
        print SAIDA $linha;
    }
}

```

4) Filtra tamanho da sequências de 60 a 120:

```

#!/usr/bin/perl -w

use Bio::SeqIO;

$seqin = Bio::SeqIO->new('-file' => 'dmel-all-intron-
r5.41.fasta', '-format' => 'fasta');
open(SAIDA, '>seq_filtradas_drosophila.out');

```

```
while($seqfasta = $seqin->next_seq){
    $seq = $seqfasta->seq;
    $nome = $seqfasta->display_id.' '.$seqfasta->desc;
    if(length($seq) >= 50 && length($seq) <= 120){
        print SAIDA ">", $nome, "\n", $seq, "\n";
    }
}
```

8 - Referências Bibliográficas

- AMBROS, V. The functions of animal microRNAs. *Nature*, v.431, p.350-355, 2004.
- ANDERSON, P. & KEDERSHA, N. *Nature reviews. Molecular cell biology* n.10, p.430-6, 2009.
- BARRETO, M. L. Esquistossomose Mansônica. Distribuição da Doença e Organização Social do Espaço. Tese de Mestrado, Salvador: Departamento de Medicina Preventiva, Universidade Federal da Bahia, 1982.
- BARTEL, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, v.116, p.281-97, 2004.
- BASCH, P.F. Cultivation of *Schistosoma mansoni* in vitro. I. Establishment of cultures from cercariae and development until pairing. *J Parasitol* v.67, p.179-85, 1981.
- BASKERVILLE, S. AND D.P. BARTEL, Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, v.11, p. 241-7, 2005.
- BECKER, C., HAMMERLE-FICKINGER, A., RIEDMAIER, I. & PFAFFL, M. mRNA and microRNA quality control for RT-qPCR analysis. *Methods (San Diego, Calif.)* v.50, 237-43, 2010.
- BENGESTRATE, L., VIRTUE, S., CAMPBELL, M., VIDAL-PUIG, A., HADASCHIK, D., HAHN, P. & BIELKE, W. Genome-wide profiling of microRNAs in adipose mesenchymal stem cell differentiation and mouse models of obesity. *PloS one* 6, e21305, 2011.
- BITAR, MAINÁ ; DRUMMOND, MARCELA GONÇALVES ; COSTA, MAURICIO GARCIA SOUZA ; LOBO, FRANCISCO PEREIRA ; CALZAVARA-SILVA, CARLOS EDUARDO ; BISCH, PAULO MASCARELLO ; MACHADO, CARLOS RENATO ; MACEDO, ANDRÉA MARA ; PIERCE, RAYMOND J. ; FRANCO, GLÓRIA REGINA. Modeling the zing finger protein SmZF1 from *Schistosoma mansoni*: Insights into DNA binding and gene regulation. *Journal of Molecular Graphics & Modelling*, v. 39, p. 29-38, 2013.

- BLACK, DOUGLAS L. Mechanisms of alternative pre-messenger RNA splicing". *Annual Reviews of Biochemistry* v.72, p.291-336, 2003.
- BONNET, E., WUYTS, J., ROUZÉ, P. & VAN DE PEER, Y. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics (Oxford, England)*, v.20, p.2911-7, 2004.
- BOROS, D.L. Immunopathology of *Schistosoma mansoni* infection. *Clinical Microbiology Reviews* v.2 n.3: p.250-69, 1989.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. *Guia de Vigilância Epidemiológica*, 7.ed. Brasília, DF; 2009.
- BRASIO, B. C. et al. Atração de miracídios de *Schistosoma mansoni* por hospedeiros invertebrados: comportamento de miracídios frente a girinos de *Hylafuscovaria*. *Saúde Pública*, São Paulo, v. 19, n. 1, 1985.
- CAI, P., HOU, N., PIAO, X., LIU, S., LIU, H., YANG, F., WANG, J., JIN, Q., WANG, H. AND CHEN, Q. Profiles of small non-coding RNAs in *Schistosoma japonicum* during development. *PLoS Negl Trop Dis* n.5 p. e1256, 2011.
- CASTRO-BORGES, W., CARTWRIGHT, J., ASHTON, P., BRASCHI, S., GUERRA SA, R., RODRIGUES, V., WILSON, R. & CURWEN, R. The 20S proteasome of *Schistosoma mansoni*: a proteomic analysis. *Proteomics*, n.7, 1065-75, 2007.
- CHALMERS, I., MCARDLE, A., COULSON, R., WAGNER, M., SCHMID, R., HIRAI, H. & HOFFMANN, K. Developmentally regulated expression, alternative splicing and distinct sub-groupings in members of the *Schistosoma mansoni* venom allergen-like (SmVAL) gene family. *BMC genomics* n.9, v.89, 2007.
- CHEN, P.Y., MANNINGA, H., SLANCHEV, K., CHIEN, M., RUSSO, J.J., JU, J., SHERIDAN, R., JOHN, B., MARKS, D.S., GAIDATZIS, D., SANDER, C., ZAVOLAN, M. AND TUSCHL, T. The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev* 19 p.1288-93, 2005.
- CLANCY, SUZANNE. "RNA Splicing: Introns, Exons and Spliceosome". *Nature Education* n.1 v.1, 2008.

- CLEGG, J.A. In Vitro Cultivation of *Schistosoma mansoni*. *Exp Parasitol* v16, p. 133-47, 1965.
- COPELAND, C., MARZ, M., ROSE, D., HERTEL, J., BRINDLEY, P., SANTANA, C., KEHR, S., ATTOLINI, C. & STADLER, P. Homology-based annotation of non-coding RNAs in the genomes of *Schistosoma mansoni* and *Schistosoma japonicum*. *BMC genomics*, v.10, p.464, 2009.
- COURA, J.R.A., R. S. Epidemiological and control aspects of schistosomiasis in Brazilian endemic areas. *Mem Inst Oswaldo Cruz* v.99, p. 13-9, 2004.
- DE SOUZA GOMES, M. Caracterização inicial dos constituintes da maquinaria de silenciamento gênico mediada por microRNAs em *Schistosoma mansoni*. 122f. Dissertação (Mestrado em Ciências Biológicas) - NUPEB, Universidade Federal de Minas Gerais, Ouro Preto, 2008.
- DE SOUZA GOMES, M. miRNAs em *Schistosoma mansoni*: biogênese, predição, validação e alvos potenciais. Tese (Doutorado em Ciências Biológicas) - NUPEB, Universidade Federal de Minas Gerais, Ouro Preto, 2011.
- DE SOUZA GOMES, M., MUNIYAPPA MK, CARVALHO SG, GUERRA-SÁ R, SPILLANE C. Genome-wide identification of novel microRNAs and their target genes in the human parasite *Schistosoma mansoni*. *Genomics*, v.98, n.2, p.96-111, 2011.
- DRUMMOND, M., CALZAVARA-SILVA, C., D'ASTOLFO, D., CARDOSO, F., RAJÃO, M., MOURÃO, M., GAVA, E., OLIVEIRA, S., MACEDO, A., MACHADO, C., PENA, S., KITTEN, G. & FRANCO, G. Molecular characterization of the *Schistosoma mansoni* zinc finger protein SmZF1 as a transcription factor. *PLoS neglected tropical diseases* v.3, e547, 2008.
- FRANCO, G.R., ADAMS, M.D., SOARES, M.B., SIMPSON, A.J., VENTER, J.C. AND PENA, S.D. Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library. *Gene* v.152, p.141-7, 1995.
- FRANCO, G.R., RABELO, E.M., AZEVEDO, V., PENA, H.B., ORTEGA, J.M., SANTOS, T.M., MEIRA, W.S., RODRIGUES, N.A., DIAS, C.M., HARROP, R., WILSON, A., SABER, M., ABDEL-HAMID, H., FARIA, M.S., MARGUTTI,

- M.E., PARRA, J.C. AND PENA, S.D. Evaluation of cDNA libraries from different developmental stages of *Schistosoma mansoni* for production of expressed sequence tags (ESTs). *DNA Res* v.4, p.231-40, 1997.
- FREITAS, C. A. Situação atual da esquistossomose no Brasil. *Revista Brasileira de Malariologia e Doenças Tropicais*, v.24, p.03-63, 1972.
- GENNARINO, V.A., D'ANGELO, G., DHARMALINGAM, G., FERNANDEZ, S., RUSSOLILLO, G., SANGES, R., MUTARELLI, M., BELCASTRO, V., BALLABIO, A., VERDE, P., SARDIELLO, M. & BANFI, S. Identification of microRNA-regulated gene networks by expression analysis of target genes. *Genome Research* v.22, 2012.
- GKIRTZOU, K., TSAMARDINOS, I., TSAKALIDES, P. AND POIRAZI, P. MatureBayes: a probabilistic algorithm for identifying the mature miRNA within novel precursors. *PLoS One* 5, p. e11843, 2010.
- GUERRA-SÁ, R., CASTRO-BORGES, W., EVANGELISTA, E., KETTELHUT, I. & RODRIGUES, V. *Schistosoma mansoni*: functional proteasomes are required for development in the vertebrate host. *Experimental parasitology*, v.109, p.228-36, 2005.
- HAN, Z.G., BRINDLEY, P.J., WANG, S.Y. AND CHEN, Z. *Schistosoma* genomics: new perspectives on schistosome biology and host-parasite interaction. *Annu Rev Genomics Hum Genet*, v.10 p.211-40, 2009.
- HARROP, R.; WILSON, R.A. Protein synthesis and release by cultured schistosomula of *Schistosoma mansoni*. *Parasitology*, v.107, n.3, 265-274, 1993.
- HE, C., LI, Z., CHEN, P., HUANG, H., HURST, L. & CHEN, J. Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA-host gene coevolution. *Nucleic acids research*, v.40, p.4002-12, 2012.
- HELM, M. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic acids research*, v.34, p.721-33, 2005.

- HOFACKER, I. RNA secondary structure analysis using the Vienna RNA package. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis*, Chapter 12, Unit12.2, 2009.
- JOHN, B., ENRIGHT, A.J., ARAVIN, A., TUSCHL, T., SANDER, C. AND MARKS, D.S. HUMAN. MicroRNA targets. *PLoS Biol* n.2, p.e363, 2004.
- KEENE, J. RNA regulons: coordination of post-transcriptional events. *Nature reviews. Genetics* v.8, 533-43, 2007.
- LAGOS-QUINTANA, M., et al., Identification of novel genes coding for small expressed RNAs. *Science*, v.294(5543) p. 853-8, 2001.
- LEE, E., BAEK, M., GUSEV, Y., BRACKETT, D., NUOVO, G. & SCHMITTGEN, T. Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors. *RNA (New York, N.Y.)*, v.14, p.35-42, 2008.
- LEE, R.C., FEINBAUM, R.L. AND AMBROS, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* v.75 p. 843-54, 1993.
- LEE, RC, AMBROS V. An extensive class of small RNAs in *Caenorhabditis elegans*. **Science** **294**(5543):862-4, 2001.
- LEE, Y., AHN, C., HAN, J., CHOI, H., KIM, J., YIM, J., LEE, J., PROVOST, P., RADMARK, O., KIM, S. AND KIM, V.N. The nuclear RNase III Droscha initiates microRNA processing. *Nature*, v.425 p.415-9, 2003.
- LEWIS, B.P., BURGE, C.B. AND BARTEL, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, v.120 p. 15-20, 2005.
- LI, L., XU, J., YANG, D., TAN, X. AND WANG, H. Computational approaches for microRNA studies: a review. *Mamm Genome*, v.21, p.1-12, 2010a.
- LINDOW, M. Prediction of targets for microRNAs. *Methods Mol Biol*, v.703, p.311-7, 2011.

- LIU, C., YU, J., YU, S., LAVKER, R., CAI, L., LIU, W., YANG, K., HE, X. & CHEN, S. MicroRNA-21 acts as an oncomir through multiple targets in human hepatocellular carcinoma. *Journal of hepatology* v.53, p.98-107, 2010.
- LOA, K. et al. Real Time PCR profiling of 330 human micro-RNAs. *J. Biotechnol.*; n.2, p.1-3, 2007.
- LU, Y., ZHOU, Y., QU, W., DENG, M. & ZHANG, C. A Lasso regression model for the construction of microRNA-target regulatory networks. *Bioinformatics (Oxford, England)*, v.27, p.2406-13, 2011.
- LUND, E., GUTTINGER, S., CALADO, A., DAHLBERG J. E. E KUTAY, U. Nuclear export of microRNA precursors. *Science*, v.303, p.95-98, 2004.
- LUTZ, A.O. Schistosomum mansoni e a schistosomose segundo observações feitas no Brasil. *MemInst Oswaldo Cruz* n.11, p.121-144, 1919.
- MAGALHÃES, L.G. Via *Notch* em *Schistosoma mansoni*: Aspectos moleculares e funcionais, 195f. Tese de Doutorado - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2009.
- MANSON, P. Report of a case of bilharzia from West Indies. *Brit Med J* n.2, p.1894-1995, 1902.
- MANZANO-ROMÁN, R. & SILES-LUCAS, M. MicroRNAs in parasitic diseases: potential for diagnosis and targeting. *Molecular and biochemical parasitology* 186, 81, 2012.
- MATHIESON, W., CASTRO-BORGES, W. & WILSON, R. The proteasome-ubiquitin pathway in the *Schistosoma mansoni* egg has development- and morphology-specific characteristics. *Molecular and biochemical parasitology* v.175, p.118-25, 2011.
- MCKERROW, J. H.; SALTER, J. Invasion of skin by *Schistosoma cercariae*. *Trends in parasitology*, v.18, p.193-195, 2002.
- MELO, A. L.; COELHO, P. M. Z. *Schistosoma mansoni* e a Doença. In. NEVES, D.P. *Parasitologia Humana*. 11 ed., São Paulo: Atheneu, p.193-212, 2005.

- MESTDAGH, P., FEYS, T., BERNARD, N., GUENTHER, S., CHEN, C., SPELEMAN, F. & VANDESOMPELE, J. High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. *Nucleic acids research*, v.36, e143, 2008.
- MESTDAGH, P., VAN VLIERBERGHE, P., DE WEER, A., MUTH, D., WESTERMANN, F., SPELEMAN, F. & VANDESOMPELE, J. A novel and universal method for microRNA RT-qPCR data normalization. *Genome biology*, v.10, R64, 2009.
- MIYAKI, S. & ASAHARA, H. Macro view of microRNA function in osteoarthritis. *Nat. Rev. Rheumatol*, 2012.
- NEVES, D. P.; MELO, A. L.; GENARO, O.; LINARDI, P. M. *Parasitologia Humana*, 11 ed, São Paulo, Atheneu, 2005.
- OBERNOSTERER, G., et al., Post-transcriptional regulation of microRNA expression. *RNA*, V.12, p.1161-7, 2006.
- OKAMURA, K., HAGEN, J.W., DUAN, H., TYLER, D.M., e LAI, E.C. The mirtron pathway generates microRNAs-class regulatory RNAs in Drosophila. *Cell* v.130, 89-100, 2007.
- OLIVEIRA, G., FRANCO, G.; VERJOVSKI-ALMEIDA, S. The Brazilian contribution to the study of the *Schistosoma mansoni* transcriptome. *Acta Trop*, v.108, p.179-82, 2008.
- PARKER-MANUEL SJ, IVENS AC, DILLON GP, WILSON RA. Gene Expression Patterns in Larval *Schistosoma mansoni* Associated with Infection of the Mammalian Host. *PLoS Negl Trop Dis*, e1274, 2011.
- PASQUINELLI, A.E., et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, v.408, p. 86-9, 2000.
- PEREIRA, R., DE SOUZA GOMES, M., OLMO, R., SOUZA, D., JANNOTTI-PASSOS, L., BABA, E., CASTRO-BORGES, W. & GUERRA-SÁ, R. NEDD8 conjugation in *Schistosoma mansoni*: Genome analysis and expression profiles. *Parasitology international* v.62, p.199-207, 2013.

- PETER, M. Targeting of mRNAs by multiple miRNAs: the next step. *Oncogene*, v.29, p.2161-4, 2010.
- PLASTERK, R.H. Micro RNAs in animal development. *Cell*, v.124, p.877-881, 2006.
- PROTASIO, A., TSAI, I., BABBAGE, A., NICHOL, S., HUNT, M., ASLETT, M., DE SILVA, N., VELARDE, G., ANDERSON, T., CLARK, R., DAVIDSON, C., DILLON, G., HOLROYD, N., LOVERDE, P., LLOYD, C., MCQUILLAN, J., OLIVEIRA, G., OTTO, T., PARKER-MANUEL, S., QUAIL, M., WILSON, R., ZERLOTINI, A., DUNNE, D. & BERRIMAN, M. A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. *PLoS neglected tropical diseases*, n.6, v.1455, 2012.
- RADFAR M.H., WONG W., MORRIS Q. Computational Prediction of Intronic microRNA Targets using Host Gene Expression Reveals Novel Regulatory Mechanisms. *PLoS ONE*, n.6, e.19312, 2011.
- REHMSMEIER, M., STEFFEN, P., HOCHSMANN, M. & GIEGERICH, R. Fast and effective prediction of microRNA/target duplexes. *RNA (New York, N.Y.)* 10, p.1507-17, 2004.
- REY, L. Parasitologia: parasitos e doenças parasitárias do homem nas Américas e na África. *Guanabara* n1, p.856, 2001.
- RUBY, J. G., JAN, C.H., e BARTEL, D.P. Intronic microRNA precursors that bypass Drosha processing. *Nature*, v.448, p. 83-86, 2007a.
- SAMBON, L.W. Remarks on *Schistosoma mansoni*. *Am J Trop Med Hyg*, v.10, p. 303-304, 1907.
- SATOH, J.-I. & TABUNOKI, H. Comprehensive analysis of human microRNA target networks. *BioData mining* v.4, p.17, 2010.
- SIMÕES, M.C., LEE, J., DJIKENG, A., CERQUEIRA, G.C., ZERLOTINI, A., DA SILVA-PEREIRA, R.A., DALBY, A.R., LOVERDE, P., EL-SAYED, N.M. AND OLIVEIRA, G. Identification of *Schistosoma mansoni* microRNAs. *BMC Genomics*, v.12 , p. 47, 2011.

- SIMPSON, A.J., SHER, A. AND MCCUTCHAN, T.F. The genome of *Schistosoma mansoni*: isolation of DNA, its size, bases and repetitive sequences. *Mol Biochem Parasitol* 6, p.125-37, 1982.
- SMITHERS, S.R. & TERRY, R.J. The infection of laboratory hosts with cercariae of *Schistosoma mansoni* and the recovery of the adult worms. *Parasitology*, v.55, p.695-700, 1965.
- THOMSON, J.M., ET AL., Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev*, p.2202-7, 2006.
- VERJOVSKI-ALMEIDA, S., DEMARCO, R., MARTINS, E.A., GUIMARAES, P.E., OJOPI, E.P., PAQUOLA, A.C., PIAZZA, J.P., NISHIYAMA, M.Y., JR., KITAJIMA, J.P., ADAMSON, R.E., ASHTON, P.D., BONALDO, M.F., COULSON, P.S., DILLON, G.P., FARIAS, L.P., GREGORIO, S.P., HO, P.L., LEITE, R.A., MALAQUIAS, L.C., MARQUES, R.C., MIYASATO, P.A., NASCIMENTO, A.L., OHLWEILER, F.P., REIS, E.M., RIBEIRO, M.A., SA, R.G., STUKART, G.C., SOARES, M.B., GARGIONI, C., KAWANO, T., RODRIGUES, V., MADEIRA, A.M., WILSON, R.A., MENCK, C.F., SETUBAL, J.C., LEITE, L.C. and DIAS-NETO, E. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. *Nat Genet* v.35, p.148-57, 2003.
- WANG, Z., XUE, X., SUN, J., LUO, R., XU, X., JIANG, Y., ZHANG, Q. AND PAN, W. An "in-depth" description of the small non-coding RNA population of *Schistosoma japonicum schistosomulum*. *PLoS Negl Trop Dis*, v.4, p.e596. 2010.
- WEAVER, R. J. Part V: Post-transcriptional events. *Molecular Biology*. Boston: McGraw Hill Higher Education, 2007.
- WESTHOLM, J. & LAI, E. Mirtrons: microRNA biogenesis via splicing. *Biochimie* v.93, p.1897-904, 2011.
- WHO. International travel and health. Geneva: World Health Organization, 2012.
- WHO. Prevention and control of schistosomiasis and soiltransmitted helminthiasis. *Technical report series*. Geneva: World Health Organization, 2008.

- WILSON, R. A. Introdução à Parasitologia. Tradução: Cláudio Santos Ferreira e Annete Silva Foronda. São Paulo; EPU; Ed. da Universidade de São Paulo, 1980.
- XIE, X., et al., Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, v.434 p. 338-45, 2005.
- XUE, X., SUN, J., ZHANG, Q., WANG, Z., HUANG, Y. & PAN, W. Identification and characterization of novel microRNAs from *Schistosoma japonicum*. *PLoS one*, v.3, e4034, 2008.
- YOICHI, R., FUGI AND NITIN, K.SKSENA. Viral Infection-Related MicroRNAs in Viral and Host Genomic Evolution. RNA and the Regulation of the Expression. Cap. 7, 2008.
- ZHANG, B., PAN, X., COX, S., COBB, G. & ANDERSON, T. Evidence that miRNAs are different from other RNAs. *Cellular and molecular life sciences : CMLS* v.63, p.246-54, 2005.
- ZHANG, Y.Q., CHEN, D.L., TIAN, H.F., ZHANG, B.H. AND WEN, J.F. Genome-wide computational identification of microRNAs and their targets in the deep-branching eukaryote *Giardia lamblia*. *Comput Biol Chem*, v.33, p.391-6, 2009c.