

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-graduação em Bioinformática

Karla Pollyanna Vieira de Oliveira

**ANÁLISE COMPARATIVA DA EVOLUÇÃO DE MIRNAS UTILIZANDO INSETOS
COMO MODELO**

Belo Horizonte
2016

Karla Pollyanna Vieira de Oliveira

**ANÁLISE COMPARATIVA DA EVOLUÇÃO DE MIRNAS UTILIZANDO INSETOS
COMO MODELO**

Dissertação de Mestrado apresentada ao Curso de Pós-graduação em Bioinformática da Universidade Federal de Minas Gerais como requisito para a obtenção do grau de Mestre em Bioinformática.

Orientador: Prof. Dr. João Trindade Marques

Coorientador: Dr. Eric Roberto Guimarães Rocha Aguiar

Belo Horizonte
2016

043

Oliveira, Karla Pollyanna Vieira de.

Análise comparativa da evolução de miRNAs utilizando insetos como modelo [manuscrito] / Karla Pollyanna Vieira de Oliveira. – 2016.

115 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. João Trindade Marques. Co-orientador: Dr. Eric Roberto Guimarães Rocha Aguiar.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Insetos. 3. MicroRNAs. 4. Análise de Sequência. I. Marques, João Trindade. II. Aguiar, Eric Roberto Guimarães Rocha. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004



DECLARAÇÃO

Declaro, para os devidos fins, que **Karla Pollyanna Vieira de Oliveira** cumpriu todos os requisitos regulamentares necessários à obtenção e atribuição do grau de Mestre em Bioinformática, tendo defendido a dissertação intitulada "**Análise comparativa da evolução de miRNAs utilizando insetos como modelo**" em 29 de janeiro de 2016, no Programa Interunidades de Pós-Graduação em Bioinformática do Instituto de Ciências Biológicas da UFMG.

Belo Horizonte, 29 de janeiro de 2016.

Dr. Vasco Ariston de C. Azevedo
Prof. Titular e Coordenador do Programa de
Pós-Graduação em Bioinformática
ICB/UFMG

Prof. Dr. Vasco Ariston de Carvalho Azevedo
Coordenador do Programa Interunidades de Pós-Graduação em Bioinformática
ICB - UFMG

AGRADECIMENTOS

Tudo começou com a “tia Fátima”, minha professora de Ciências da 4ª série do Ensino Fundamental. Foi amor à primeira vista... No Ensino Médio, veio a Jackie Chan. Juro que não me lembro de seu nome verdadeiro, apenas de sua fisionomia e de sua influência na minha vida acadêmica... Pouco tempo depois vieram os professores do Curso Técnico em Biotecnologia: Solange, Lúcia, Murilo, Michelyne... E se já na 7ª série eu dizia que seria cientista, no curso técnico isso só se confirmou.

Comecei a graduação em Ciências Biológicas na UNA, pois tinha medo das greves da UFMG. Continuei cursando o ensino técnico e por insistência do Prof. Murilo, prestei um concurso para minha área (Biotecnologia) sem saber ao menos o que isso significava. Para minha surpresa, passei. Tive que correr para entregar a monografia, terminar estágio, enfim, conseguir me formar...

Com 18 anos estava empregada. Não conseguia conciliar trabalho e faculdade, de modo que tive de prestar novo vestibular. Dessa vez passei na UFMG. Foram 7 anos de história na Funed. Quando decidi abandonar meu emprego no serviço público, eu sentia que não seria fácil. Foram muitos anos trabalhando e tendo certa estabilidade na carreira. Foram 7 anos também sonhando em ser pesquisadora (fora todos os anos anteriores que eu não vou mencionar, pois a idade de uma mulher é algo sigiloso).

Durante o período em que lá trabalhei, fiquei sabendo que um novo professor chegara à UFMG tendo como linha de pesquisa os RNAs de Interferência. Eu tinha acabado de me apaixonar pela área, a partir das aulas da Prof. Glória Franco. Mandeí email com um currículo. Três horas depois a surpresa da resposta.

Comecei a Iniciação Científica durante as férias da Funed, em janeiro de 2011. Queria saber se possuía aptidão para a pesquisa. Devo ter, pois o João me aceitou como IC durante o período que eu tinha disponível. No ano seguinte tive a oportunidade de estagiar no Lab. do Prof. Richard Carthew na Northwestern University (NU). Ele fora orientador do João durante pós-doutorado, o que aumentava de forma considerável a responsabilidade que eu tinha em não decepcionar ambos. Lembro da Xuan (que conheci na NU) falando que o *paper* dependia dos experimentos que eu iria realizar. Passei um mês ininterrupto sonhando com *Drosophila* e com o João me cobrando resultados. Voltei ao Brasil, 5 meses depois.

Já no fim da graduação ouvi da Prof. Denise algo que me fez chorar. Ela nos

desejou que jamais precisássemos trabalhar. O “precisar” implica algo doloroso. Invés disso, ela nos desejou que sentíssemos prazer em exercer nossas profissões. Ao ouvir tais palavras, imediatamente me lembrei do meu pai. Ele sempre orientou a mim e ao meu irmão que fizéssemos o que nos torna felizes, independente da profissão ou cargo que escolhêssemos. Esse foi o melhor conselho que alguém já me deu. Senti que todas as minhas escolhas sempre foram pautadas no que me faz feliz. Eu sempre soube o que quero. Apesar de tudo isso, sinto que não me preparei o suficiente para os enfrentamentos que viriam com as minhas decisões. Na verdade, acho que somente o caminho nos prepara. E por falar em caminho, sei que falta um longo percurso para que eu chegue onde quero. Todavia completo hoje mais uma etapa de minha trajetória.

Essa dissertação é fruto de uma história e nada mais justo que agradecer aos que dela participaram, de forma direta ou indireta. Por isso, quero agradecer imensamente aos meus pais e ao meu irmão pelo suporte, por me apoiarem quando estava dando tudo errado, quando chorava e duvidava que conseguiria. Essa dissertação é dedicada a vocês.

Agradeço também aos meus amigos. Obrigada por dividirem seus momentos comigo e permitir que eu dividisse os meus. Independente da distância ou da frequência com que nos vejamos. Independente da universidade, cidade, estado ou país em que vocês se encontrem agora. Embora seja uma frase bastante clichê, não menosprezem a veracidade dela: o caminho é mais prazeroso pois os tenho em minha vida.

Quero agradecer aos professores que tive durante toda minha vida acadêmica. Obrigada por me mostrarem um mundo fascinante. Posso dizer com propriedade que sei a importância da influência de vocês nas escolhas de um aluno, de um cidadão. Muito obrigada também àqueles com quem tive o prazer de trabalhar até agora. Seja na Funed, onde me formei profissionalmente; no Laboratório de RNA de Interferência, que apesar de pequeno e de vez ou outra nós termos dificuldade em encontrar um espaço desocupado, é um ótimo lugar pra se trabalhar e possui pessoas incríveis, dispostas a ajudar em experimentos, em discussões científicas, discussões sobre o sentido da vida ou apenas aproveitar o momento bebendo uma cerveja e comendo o churrasco preparado pelo Marcílio e pelo Vicente; seja ainda no Laboratório do Rich, o oposto em termos de espaço e um lugar onde aprendi muito também e que mostrou sou capaz de ir além.

Chego agora aos agradecimentos aos meus orientadores. Primeiro ao meu co-orientador, Eric Aguiar. Eric, muitíssimo obrigada por ter aceitado o convite, por todo o apoio, pelas conversas, pelas orientações tanto quanto à execução dos comandos, quanto às discussões sobre outros tantos aspectos. Quero agradecer de forma especial ao meu orientador: Prof. João Marques. João, obrigada por ter respondido àquele primeiro email, ter aceitado que eu ingressasse no laboratório e por ter conferido a mim a confiança por meio de cada oportunidade.

Por fim, apesar de um pouco cética e de sempre me questionar a respeito de sua existência, agradeço à Deus, ou à força motriz do universo, qualquer que seja seu nome, pela existência de cada um desses na minha vida.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”.

(Martin Luther King)

“Do. Or do not. Try, there is not.”
Mestre Yoda

RESUMO

MicroRNAs (miRNAs) são uma classe de pequenos RNAs não-codificantes descobertos em 1993. Estas sequências se originam de um RNA precursor endógeno, que é processado em um pequeno RNA cujo tamanho varia de 20 a 24 nucleotídeos. Os miRNAs são conservados em eucariotos e possuem importante função na regulação gênica. Apesar da sua importância e da conservação das sequências de miRNAs entre espécies, pouco se sabe se outras características que definem os miRNAs são também conservadas, uma vez que não há estudos de tais aspectos. Para responder a essa questão, foram geradas bibliotecas de pequenos RNAs derivadas de três insetos (*Drosophila melanogaster*, *Aedes aegypti* e *Lutzomyia longipalpis*), todos da ordem Diptera, e cuja distância evolutiva é de aproximadamente 250 milhões de anos. As bibliotecas foram sequenciadas e analisadas sendo os miRNAs de cada espécie avaliados não só quanto à conservação de sequência, mas também com relação ao perfil de tamanho, homogeneidade da região 5', origem genômica, origem no precursor, e preferência de base. De modo geral, houve grande conservação de miRNAs entre as três espécies não só em termo de suas sequências, mas também de todas as demais características avaliadas. Todavia, analisando os miRNAs espécie-específicos, foi observada maior divergência dos padrões característicos dos miRNAs. Os miRNAs conservados surgiram há mais tempo e portanto possuem maior tempo de evolução e seleção natural. Em contraste, os miRNA espécie-específicos seriam os mais recentes e por isso podem ainda estar passando por processos seletivos. Assim, esses resultados sugerem que a pressão evolutiva sofrida pelos miRNAs é refletida não somente na sua sequência, mas também em outras características intrínsecas provavelmente ligadas à sua estrutura geral que determina como será sua biogênese.

Palavras-chave: miRNAs, sequenciamento, insetos.

ABSTRACT

MicroRNAs (miRNAs) are a small non-coding RNA class discovered in 1993. These sequences originate from the endogenous precursors, which are processed into small RNAs ranging size from 20 to 24 nt. MiRNAs are conserved in eukaryotes and they are important gene regulators. Despite their importance and the sequence conservation among species, other miRNAs' features conservation remains unclear since there are no studies of such aspects. To answer this question, we generated small RNAs libraries – derived from three dipteran insects whose evolutionary distance is approximately 250 million years ago. The libraries were sequenced, and the miRNAs of each organism were analyzed according to species conservation, sequence, size profile, 5' homogeneity, genomic origin, precursor origin, and preferably base. Overall, it was observed high conservation of miRNAs between the three dipteran insects not only in terms of their sequences, but also in terms of all the other features. However, when we analyzed miRNAs that were found in only one specie, it was observed a higher divergence of miRNAs patterns than observed in conserved miRNAs. These observations can be explained by the time of miRNAs evolution. Conserved miRNAs are evolutionarily older than non-conserved miRNAs and they have been naturally selected. In contrast, the miRNA-specific species were the most recent and therefore have not been selected. Thus, these results suggest that the evolutionary pressure suffered by miRNAs is reflected not only in their sequence, but also in other intrinsic characteristics probably linked to its overall structure that determines how it is their biogenesis.

Keywords: miRNAs, sequencing, insects

LISTA DE FIGURAS

Figura 1: Via canônica de miRNAs em insetos.	21
Figura 2: Origens genômicas de miRNAs.	22
Figura 3: Representação dos tipos de sítios-alvo de miRNAs.	28
Figura 4: Esquema de um miRNA primário (pri-miRNA) de transcrição independente.	31
Figura 5: Evolução simplificada das vias de miRNA.	32
Figura 6: Conservação da família de miRNAs let-7 em diferentes clados.	33
Figura 7: Árvore filogenética das famílias de insetos da ordem Diptera.	36
Figura 8: <i>Pipeline</i> de identificação e análise dos miRNAs em insetos da ordem Diptera.	45
Figura 9: Precursores únicos conservados entre os diferentes insetos analisados.	64
Figura 10: Origem genômica dos miRNAs. A origem genômica dos miRNAs foi dividida em três categorias (exônica, intergênica e intrônica) e sua contribuição percentual foi verificada para cada organismo estudado.	69
Figura 11: Média da expressão normalizada (RPM) dos miRNAs conservados entre as três espécies estudadas.	71
Figura 12: <i>Boxplot</i> da comparação de expressão (RPM) dos miRNAs classificados como de “expressão similar” e “expressão discrepante” entre os insetos estudados.	74
Figura 13: Comparação par a par dos níveis de expressão (RPM) de miRNAs compartilhados entre os três insetos.	75
Figura 14: Comparação entre os percentuais de 10 dos miRNAs mais expressos em cada organismo.	78
Figura 15: Média da expressão normalizada (RPM) dos miRNAs não-conservados entre as três espécies estudadas.	79
Figura 16: Relação entre a expressão dos braços 5p e 3p (<i>ratio</i>) dos miRNAs conservados entre os diferentes insetos.	82
Figura 17: Enriquecimento de base na região 5' de todas as sequências mapeadas em miRNAs anotados para cada espécie.	85
Figura 18: Enriquecimento de base na região 5' dos miRNAs anotados para cada espécie.	86
Figura 19: <i>Heatmap</i> do perfil de tamanho das sequências mapeadas nos miRNAs conservados de cada espécie.	87
Figura 20: Perfil de tamanho dos miRNAs conservados entre os insetos estudados através da normalização por Z-score	88
Figura 21: <i>Heatmap</i> do perfil de tamanho das sequências mapeadas nos miRNAs não-conservados de cada espécie	92
Figura 22: Perfil de tamanho dos miRNAs não-conservados entre os insetos estudados através da normalização por Z-score	93

LISTA DE TABELAS

Tabela 1: Principais características dos componentes das vias de RNAi no organismo modelo <i>Drosophila melanogaster</i>	18
Tabela 2: Descrição das bibliotecas de pequenos RNAs utilizadas.	44
Tabela 3: Número de miRNAs e precursores do subfilo Hexapoda disponíveis no miRBase (versão 21) e utilizados para predição em <i>Lutzomyia longipalpis</i>	48
Tabela 4: Scores para avaliação das características dos miRNAs conservados	56
Tabela 5: Número de sequências após pré-processamento e seu respectivo mapeamento no genoma.	58
Tabela 6: Análise da presença de SNPs em miRNAs.	61
Tabela 7: Resumo de precursores e miRNAs maduros anotados em cada organismo.	63
Tabela 8: pre-miRNAs conservados entre três organismos da ordem Diptera.	65
Tabela 9: miRNAs conservados entre os insetos analisados.	67
Tabela 10: Padrão de expressão de miRNAs dos insetos avaliados segundo similaridade.	73
Tabela 11: 20 miRNAs mais expressos em cada organismo.	76
Tabela 12: miRNAs com todas as características conservadas (anotação pela mesma metodologia, tamanho e expressão conservados) entre os diferentes insetos estudados.	89
Tabela 13: miRNAs escolhidos para avaliação de perfil de tamanho dos precursores.	90

LISTA DE ABREVIATURAS

A	Adenina
Aae	<i>Aedes aegypti</i>
Ago	Argonauta
Ago1	Argonauta-1
Ago2	Argonauta-2
ATP	Adenosina trifosfato
C	Citosina
Dcr1	Dicer-1
Dcr2	Dicer-2
DENV	Dengue virus
Dme	<i>Drosophila melanogaster</i>
DNA	Ácido desoxirribonucleico
dpa	Dias após alimentação
Exp5	Exportina 5
G	Guanina
G0	Estágio quiescente do ciclo celular
GFF3	Formato das características gerais – versão 3. (<i>General Feature Format</i>)
GO	Ontologia do gene (<i>Gene ontology</i>)
GTP	Guanina trifosfato GW182 Proteína Gawky 182
H3K4me1	Metilação da lisina K4 da histona H3
ICB	Instituto de Ciências Biológicas
IGBMC	<i>Institut de Génétique e Biologie Moléculaire et Cellulaire</i>
kb	Kilobase
Llo	<i>Lutzomyia longipalpis</i>
Loqs	Loquacious
m ⁷ G	7-metilguanossina
MFE	Energia Mínima Livre (<i>Minimum free energy</i>)
miRNA	MicroRNA – Sequência madura
mirna	microrna – Precursor
miRors	miRNAs espelho
miRtron	miRNAs intrônicos
mRNA	RNA mensageiro (<i>messenger RNA</i>)
N/C	Não consta
ncRNA	RNA não codificante (<i>non-coding RNA</i>) nt Nucleotídeo
Pol	Polimerase
R ²	Coeficiente de determinação
Ran	Proteína nuclear relativa a RAs (<i>RAs-related nuclear protein</i>)
RISC	Complexo de Silenciamento Induzido por RNA
RNA	Ácido ribonucleico
RNAi	RNA de interferência
RPM	Sequências por milhão (<i>Reads per million</i>)
rRNA	RNA ribossômico
siRNA	Pequenos RNAs de interferência (<i>small interfering RNA</i>)
SNP	Polimorfismo único de nucleotídeo
stRNA	Pequenos RNAs temporais (<i>small temporal RNAs</i>)
tRNA	RNA Transportador
TSS	Sítio de início de transcrição (<i>transcription start site</i>)

U	Uracila
UFMG	Universidade Federal de Minas Gerais
UTR	Região não traduzida (<i>Untranslated region</i>)
VSV	<i>Vesicular Stomatitis virus</i>
σ	Desvio-padrão

SUMÁRIO

1. INTRODUÇÃO	18
1.1. Aspectos históricos	18
1.2 Biogênese dos miRNAs	20
1.2.1 Transcrição.....	20
1.2.2 Maturação	23
1.2.3 Carregamento	25
1.3 Mecanismos de regulação do alvo	26
1.4 Nomenclatura	30
1.5 Evolução da via de miRNAs	31
1.6 miRNAs e sua conservação em insetos da ordem Diptera	34
2. JUSTIFICATIVA	39
3. OBJETIVOS	41
3.1 Objetivo geral	41
3.2 Objetivos específicos	41
4. MATERIAIS E MÉTODOS	43
4.1 Origem das amostras	43
4.2 Pre-processamento das sequências	46
4.3 Identificação de miRNAs baseada no genoma disponível	46
4.3.1 <i>Lutzomyia longipalpis</i>	47
4.3.2 <i>Aedes aegypti</i>	49
4.4 Identificação de miRNAs baseada em comparação com organismos próximos	49
4.5 Análise da presença de SNPs nos miRNAs identificados por comparação com organismos próximos	50
4.6 Curadoria manual	50
4.7 Anotação de miRNAs	52
4.8 Análises comparativas	52
4.8.1 miRNAs compartilhados entre os três organismos	52
4.8.2 Origem genômica dos miRNAs	53
4.8.3 Avaliação da expressão	53
4.8.4 Expressão dos miRNAs segundo origem no precursor de cada braço no grampo do pre-miRNA	54
4.8.5 Perfil de tamanho dos miRNAs	55
4.8.6 Enriquecimento de base.....	55
4.8.7 Comparação de miRNAs conservados e com perfis de tamanho distintos entre os organismos.....	55
4.8.8 Avaliação das diferenças de perfis de miRNAs segundo expressão e tamanho	56
5. RESULTADOS E DISCUSSÃO	58

5.1	Bibliotecas de pequenos RNAs	58
5.2	Identificação e anotação de miRNAs e precursores	59
5.3	Conservação de miRNAs e precursores entre os insetos estudados...	64
5.4	Origem genômica	69
5.5	Expressão dos miRNAs	70
5.5.1	Expressão geral	70
5.5.2	Expressão dos miRNAs segundo origem no precursor	80
5.6	Padrões Moleculares: perfil de tamanho e preferência de base.....	84
5.6.1	Preferência de Base.....	84
5.6.2	Perfil de tamanho das sequências mapeadas nos miRNAs.....	86
5.7	Considerações finais.....	94
6.	<i>CONCLUSÕES</i>	96
7.	<i>PERSPECTIVAS</i>.....	98
8.	<i>REFERÊNCIAS BIBLIOGRÁFICAS</i>.....	100

1. INTRODUÇÃO

1.1. Aspectos históricos

Os miRNAs fazem parte de uma família de pequenos RNAs, chamados de RNAs de interferência (RNAi) em que se destacam 3 grupos (**TABELA 1**): siRNAs, miRNAs e piwiRNAs (piRNAs). Destas, as duas primeiras possuem em comum a regulação transcricional por pequenos RNAs, enquanto a família dos piRNAs está relacionada principalmente com a regulação negativa de elementos transponíveis em células germinativas (Revisado por Chu e Rana, 2007; Revisado por Kim *et al.*, 2009; Revisado por Ha e Kim, 2014). Estas vias possuem em comum seu mecanismo associado às proteínas da família Argonauta (Ago), que serão discutidas mais detalhadamente à frente, com foco principal na atuação dos miRNAs, objetivo deste trabalho.

Tabela 1: Principais características dos componentes das vias de RNAi no organismo modelo *Drosophila melanogaster*.

Característica	miRNA	siRNA	piRNA
Tamanho	~ 22nt	~21nt	24~29nt
Origem do pequeno RNA	Fita simples de RNA	Fita dupla de RNA	Fita simples de RNA
Enzimas de processamento	Drosha e Dicer-1	Dicer-2	Zucchini
Proteínas Argonauta	Argonauta 1	Argonauta 2	Aubergine e Piwi
Mecanismo de Ação	Inibição da tradução; Clivagem do mRNA alvo	Clivagem do RNA alvo	Clivagem do RNA alvo
Função	Regulação de genes codificadores de proteínas	Regulação de genes codificadores de proteínas; Defesa antiviral	Regulação de elementos transponíveis; Defesa antiviral

Adaptado de Ha e Kim (2014)

MicroRNAs (miRNAs) são pequenos RNAs endógenos, fita simples, não codificantes (ncRNAs), de aproximadamente 20 a 24 nucleotídeos (nt), cuja função é relacionada à regulação pós-transcricional de RNAs mensageiros (mRNAs) mediante complementaridade de sequência do tipo Watson-Crick (Lee *et al.*, 1993; Wightman *et al.*, 1993; Reinhart *et al.*, 2000; Lai, 2002; Ambros *et al.*, 2003; Bartel, 2004a; Kim, 2005; Kim e Nam, 2006; Bartel, 2009; Lucas e Raikhel, 2013; Ha e Kim, 2014).

Os dois primeiros miRNAs descobertos foram descritos no verme *Caenorhabditis elegans*, e foram chamados lin-4 e let-7 (Lee *et al.*, 1993; Reinhart *et al.*, 2000). Sua descoberta deu-se através de estudos do controle temporal do desenvolvimento das larvas desse verme e acreditava-se que estariam envolvidos somente nesse processo. Foram então chamados de pequenos RNAs temporais (stRNAs), não sendo creditado a eles reconhecimento como extensa classe de moléculas reguladoras. No entanto, trabalhos posteriores mostraram sua abundância em número, conservação entre espécies e, pela descoberta de novas sequências com aproximadamente 22 nt, passaram a ser chamados miRNAs (Pasquinelli *et al.*, 2000; Lagos-Quintana *et al.*, 2001b; Lau *et al.*, 2001; Lee e Ambros, 2001; Lagos-Quintana, 2003). Atualmente sabe-se inclusive, que alguns vírus podem codificar miRNAs (Who; Skalsky e Cullen, 2010; Grundhoff e Sullivan, 2011).

Com o avanço dos estudos na área, os miRNAs tiveram sua importância reconhecida nos mais diversos processos biológicos, tais como: proliferação celular (Brennecke *et al.*, 2003); desenvolvimento do padrão neuronal (Johnston e Hobert, 2003); metabolismo (Barrio *et al.*, 2014); imunidade (Liu *et al.*, 2015), dentre outros. Como exemplos, podemos citar o primeiro miRNA descrito no organismo modelo *Drosophila melanogaster*, miR-14 – que apresenta função regulatória de um ativador de morte celular e que quando suprimido ocasiona redução na sobrevivência dessas moscas – que deu início a inúmeras publicações que contribuíram enormemente para o entendimento desta via neste e em diversos outros organismos (Xu *et al.*, 2003). Foi mostrado posteriormente em *C. elegans* que alterações na disponibilidade de alguns miRNAs possuem impacto também na sobrevivência durante estágios específicos do desenvolvimento (Biemar *et al.*, 2005; Zhang *et al.*, 2011).

Os miRNAs possuem como principal função o silenciamento gênico, uma vez que são capazes de levar à diminuição da expressão de um gene ou conjunto de genes. No entanto, também foi demonstrada a existência de miRNAs capazes de regular positivamente a expressão de seus alvos (Place *et al.*, 2008; Valinezhad Orang

et al., 2014). Além disso, alterações nos níveis de alguns miRNAs foram associadas a patologias, evidenciando a importância do papel por eles desempenhado (Croce, 2009; Schonrock *et al.*, 2010; Alvarez-Erviti *et al.*, 2013).

1.2 Biogênese dos miRNAs

1.2.1 Transcrição

O processo de biogênese dos miRNAs (**FIGURA 1**) inicia-se com sua transcrição pela RNA polimerase II (RNA Pol II) e consequente formação da sequência primária (pri-miRNA), cujo tamanho pode ser superior a 1 kb (Lee *et al.*, 2004). Esta sequência sofre adição de um grupo 7-metilguanossina (m^7G) à extremidade 5' e a adição da cauda de poli-adeninas na extremidade 3', de maneira semelhante ao processamento dos mRNAs celulares (Cai *et al.*, 2004).

Os pri-miRNAs podem ser codificados a partir de regiões intergênicas, exônicas ou mesmo de a partir de regiões intrônicas (Kim e Kim, 2007; Hinske *et al.*, 2010; Isik *et al.*, 2010; Slezak-Prochazka *et al.*, 2013; Ramalingam *et al.*, 2014). Atualmente são descritas pelo menos 5 origens genômicas de um miRNA (**FIGURA 2**), a saber: a) transcritos independentes; b) dispostos em grupo; c) intrônicos; d) miRtrons convencionais e e) miRNAs com prolongamentos na extremidade 3'.

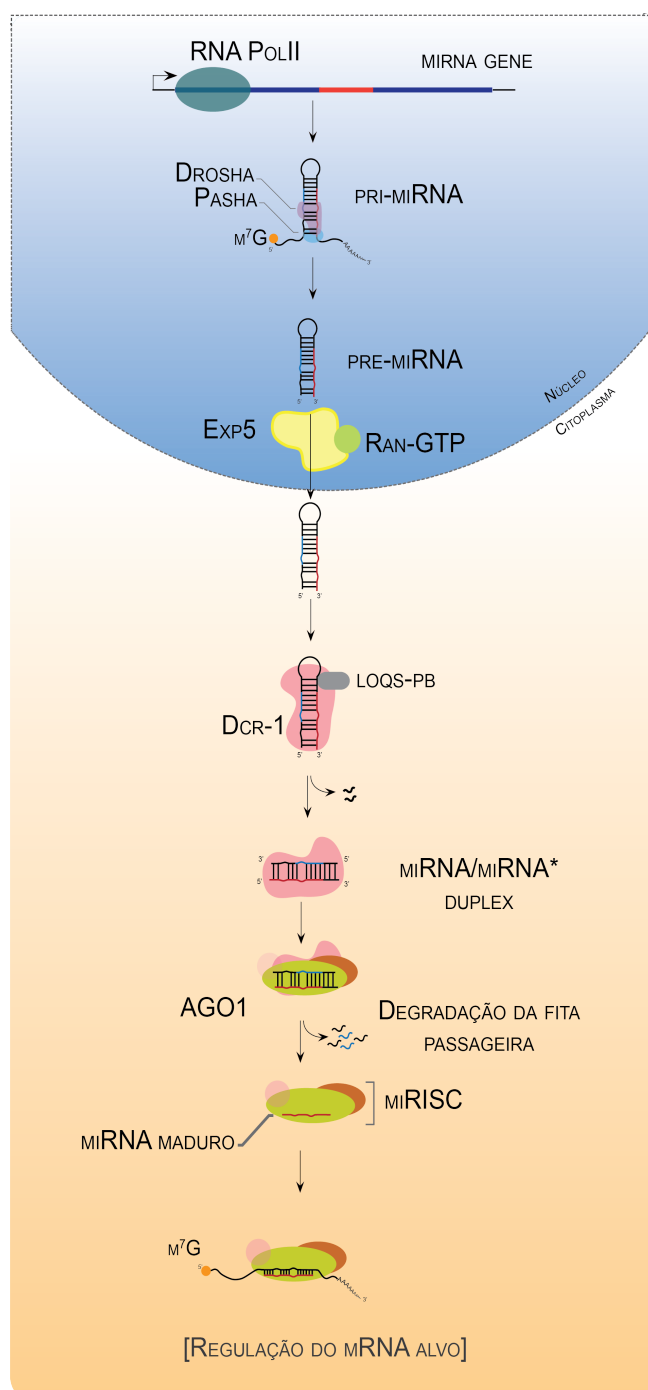


Figura 1: Via canônica de miRNAs em insetos. A via clássica de biogênese do miRNA começa com a transcrição pela RNA Pol II e clivagem do pri-miRNA, ainda no núcleo, pelo complexo microprocessador (constituído pelas enzimas Drosha e Pasha) gerando o grampo conhecido como pre-miRNA que será então exportado ao citosol através da proteína exportina 5. Uma vez no citoplasma, ocorre clivagem do pre-miRNA pelo complexo Dicer-1/Loqs-PB a fim de gerar o *duplex* que será carregado na proteína Ago1. A seguir, apenas uma das fitas permanece associada a Ago1, formando o complexo miRISC maduro, o qual atuará na regulação do mRNA alvo. **RNA Pol II:** RNA Polimerase II; **Exp5:** Exportina 5; **Ran:** Proteína nuclear relativa a RAS; **Dcr-1:** Dicer-1; **Loqs-PB:** Loquacious – isoforma PB; **miRISC:** Complexo de Silenciamento Induzido por miRNA; **miRNA*:** miRNA estrela ou fita passageira.

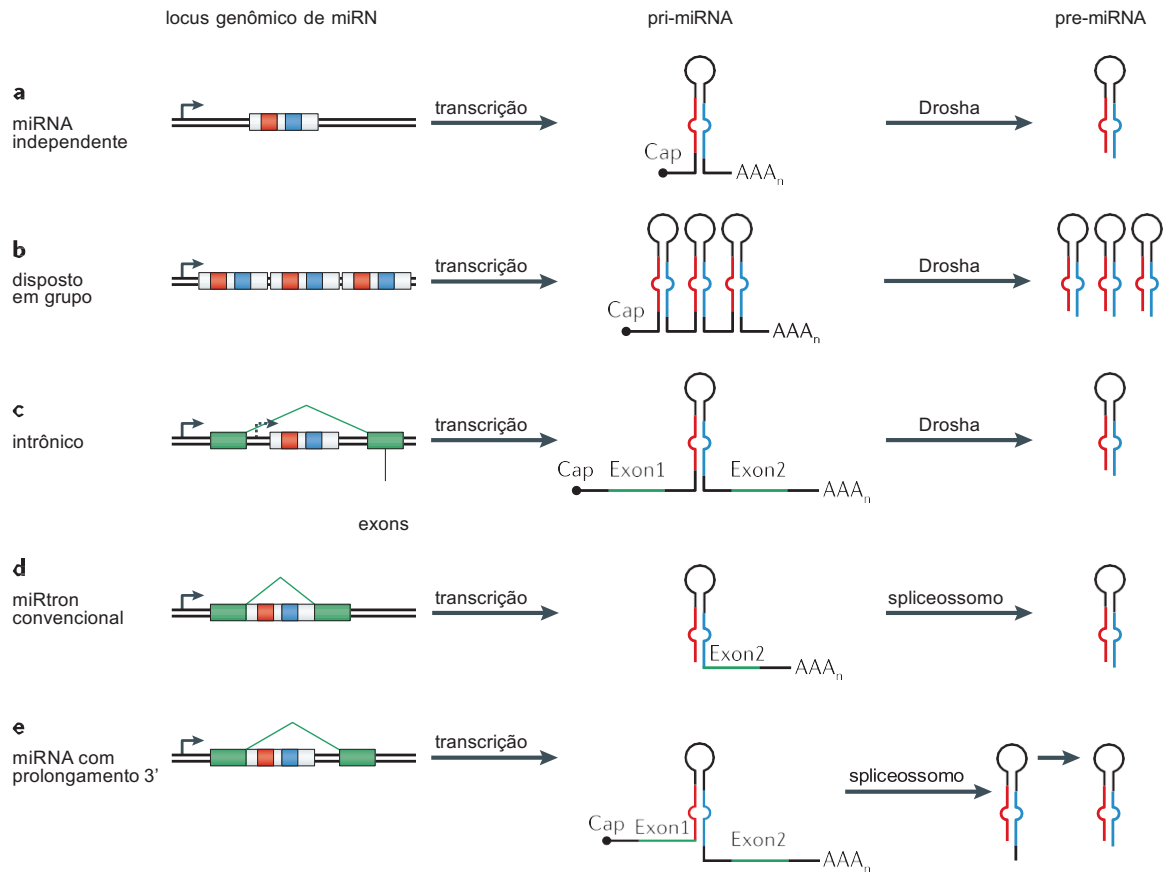


Figura 2: Origens genômicas de miRNAs. A figura indica 5 origens genômicas de miRNAs. Os pri-miRNAs sofrem adição de 7-metilguanossina (*cap*) na extremidade 5' e de poli-adeninas na extremidade 3'. O pre-miRNA é gerado a partir da clivagem pela enzima Droscha ou pela maquinaria de spliceossomo celular. **a)** miRNA de origem independente no genoma, o qual apresenta seqüências de início de transcrição próprias; **b)** miRNAs dispostos em grupo são transcritos em conjunto pela RNA polimerase II; **c)** miRNAs intrônicos são clivados pela Droscha; **d)** miRtrons convencionais apresentam suas extremidades 5' e 3' dispostas no sítio de clivagem pela maquinaria de spliceossomo celular; **e)** miRNAs que apresentam prolongamentos da extremidade 3' são clivados e têm sua extremidade 3' novamente processada para a geração do pre-miRNA. Adaptado de Berezikov (2011).

1.2.2 Maturação

A principal característica dos pri-miRNAs baseia-se na estrutura secundária gerada pelo transcrito de RNA, que forma alças através do pareamento interno de suas bases (**FIGURA 1**). Uma vez transcrito e enovelado, ainda no núcleo, o pri-miRNA é reconhecido pela enzima Drosha, uma endonuclease do tipo III, através de características de sua estrutura secundária como: tamanho do laço terminal do grampo (*loop*); tamanho da haste do grampo; e tamanho da fita dupla (*duplex*) (Conesa *et al.*, 2005; Auyeung *et al.*, 2013). Após o reconhecimento, a Drosha associa-se à proteína Pasha, também no núcleo, e cliva ambas as fitas do pri-miRNA próximo à base de seu grampo (Tomari e Zamore, 2005).

A sequência do grampo liberado contém aproximadamente 70 nt de extensão, com um grupo fosfato livre em sua extremidade 5' e um grupo hidroxila livre na extremidade 3'. Além disso, sua extremidade 3' geralmente apresenta 2 nt não pareados, comumente chamados nucleotídeos 3' *overhang* (Lee *et al.*, 2003; Tomari e Zamore, 2005; Zeng *et al.*, 2005). Cabe ressaltar que esta é uma característica inerente ao processamento por endonucleases (Nicholson, 1999).

Ao grampo de miRNA, produto da clivagem pelo complexo proteico formado por Drosha e Pasha, dá-se o nome de miRNA precursor (pre-miRNA) (**FIGURA 1**). Este é então exportado ativamente ao citosol pela proteína Exportina 5 (Exp5), que requer a hidrólise de guanosina trifosfato (GTP) e possui a proteína Ran como cofator (Yi *et al.*, 2003; Bohnsack, 2004; Lund *et al.*, 2004). A Exp5 reconhece os nucleotídeos não pareados da extremidade 3' do grampo e parece atuar também protegendo o pre-miRNA contra degradação (Lund *et al.*, 2004; Zeng e Cullen, 2004).

Uma exceção à geração canônica dos miRNAs é vista na classe dos miRtrons (**FIGURA 2E**), que não passam pela etapa de processamento da Drosha (Okamura *et al.*, 2007; Ruby *et al.*, 2007). Os miRtrons localizam-se em íntrons de regiões codificadoras de proteínas, mas diferem de miRNAs canônicos de origem intrônica, uma vez que suas porções terminais coincidem com as junções de *splicing* clivadas pela maquinaria do spliceossomo celular, não sendo necessária ação da Drosha. No entanto, de modo semelhante ao descrito para miRNAs canônicos, os pre-miRNA derivados dos miRtrons são exportados ao citosol pela Exp5.

Uma vez no citoplasma, o pre-miRNA é clivado por outra RNase do tipo III, conhecida como Dicer-1 (Dcr-1) (Grishok *et al.*, 2001; Hutvagner *et al.*, 2001). Essa

enzima foi caracterizada primeiramente no organismo modelo *Drosophila melanogaster* (Bernstein *et al.*, 2001). De modo análogo à clivagem pelo complexo Drosha/Pasha, a clivagem pela Dcr-1 também resulta em extremidade 5' fosfato livre e extremidade 3' hidroxila livre dotada de 2 nt não pareados. A Dcr-1 possui dois domínios helicase e um domínio PAZ, que reconhecem o tamanho do laço terminal e sua distância até a extremidade 3' da haste do pre-miRNA – características determinantes para clivar a dupla fita através dos domínios RNase que estão a distâncias fixas da extremidade 3' não pareada (Tsutsumi *et al.*, 2011). Além disso, essa enzima parece discriminar as sequências do pre-miRNA a fim de selecionar o sítio de clivagem, evitando por exemplo, o corte em posições que produzam miRNAs com o nucleotídeo guanina (G) na extremidade 5' (Starega-Roslan *et al.*, 2015a).

Ao final da etapa de clivagem pela Dcr-1, o grampo deixa de apresentar o laço, tornando-se apenas uma dupla fita (*duplex*) transiente de RNA (**FIGURA 1**). Este *duplex* é então carregado na proteína Argonauta-1 (Ago1) e após a seleção de uma das fitas, com consequente descarte da fita complementar, promove a regulação dos seus genes-alvo (Okamura *et al.*, 2004). O *duplex* do miRNA é composto pelo miRNA maduro, carregado e selecionado pela proteína Ago1, e pela sua fita complementar, que é descartada pela Ago1. Originalmente, a fita complementar era conhecida por miRNA* (lê-se miRNA “estrela”), no entanto esta nomenclatura tem sido substituída, uma vez que ambas as fitas de um mesmo *duplex* podem ser capazes de gerar miRNAs maduros através da variação da escolha pela Ago1 que pode variar segundo tecido ou estágio celular (Lagos- Quintana *et al.*, 2001a; Lee *et al.*, 2002; Biemar *et al.*, 2005).

A capacidade de geração de miRNAs maduros pela Dcr-1 e seu carregamento na Ago1 em *D. melanogaster* é afetada pela ausência de uma proteína acessória chamada Loquacious (Loqs) cuja isoforma envolvida nesse processo é conhecida como PB (Forstemann *et al.*, 2005; Jiang *et al.*, 2005; Saito *et al.*, 2005). Esta descoberta deu-se a partir da observação de que na via de siRNAs há também a necessidade de uma proteína acessória à Dcr-2, chamada R2D2, para o carregamento do pequeno RNA na proteína Ago2 (Liu *et al.*, 2006). Como as vias são semelhantes em funcionamento e mecanismo, Saito e colaboradores, bem como Jiang e colaboradores, investigaram e observaram a associação da proteína Loqs-PB com Dcr-1 em *Drosophila*.

1.2.3 Carregamento

O número de proteínas da família Argonauta varia entre as espécies (Hutvagner e Simard, 2008). Em *D. melanogaster*, por exemplo, foram identificadas 5 proteínas tipo Argonauta, as quais são divididas em dois subclados: Ago e Piwi. No subclado Ago, encontram-se Ago1 e Ago2, as quais foram originalmente descritas na participação das vias de miRNA e siRNA, respectivamente. Em mamíferos e outros organismos, a biogênese de miRNAs ocorre de forma semelhante ao observado em *D. melanogaster*, uma vez que a via de miRNAs é altamente conservada em eucariotos (Berezikov, 2011). No entanto, existem algumas particularidades referentes às proteínas envolvidas neste processo. Uma dessas particularidades refere-se ao fato que, de modo distinto ao encontrado em insetos, em mamíferos há 4 proteínas Ago e parece não haver preferência pelo direcionamento dos miRNAs especificamente para cada uma delas, o que sugere haver redundância funcional dessas proteínas (Su *et al.*, 2009).

Ainda com relação à redundância funcional das proteínas Argonautas, em insetos, estudos demonstraram que em casos específicos a escolha do carregamento de alguns *duplexes* de siRNA ou de miRNA na Ago1 ou Ago2 podem ser intercambiáveis e são influenciados por suas características estruturais (Doerks *et al.*, 2002; Okamura *et al.*, 2004; Forstemann *et al.*, 2007; Tomari *et al.*, 2007; Okamura *et al.*, 2009). Somado a isso, após a escolha do miRNA maduro pela Ago1, a fita complementar descartada pode não ser rapidamente degradada e associar-se a Ago2, modulando mRNAs complementares (Forstemann *et al.*, 2007; Czech *et al.*, 2009; Okamura *et al.*, 2009; Ghildiyal *et al.*, 2010). No entanto, vale ressaltar que estes casos constituem exceções à regra e parece ocorrer discriminação entre as vias de miRNA e siRNA nesses organismos.

Como regra geral, a existência de nucleotídeos não pareados no *duplex* de RNA ocasiona a preferência da Ago1 por esses substratos (Pratt, A. J. e Macrae,

I. J., 2009). De fato, a extremidade 5' do pre-miRNA escolhido pela Ago1 apresenta maior instabilidade, decorrente dos nucleotídeos não pareados. Essa instabilidade é visualizada através dos maiores valores de energia livre (MFE) (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003; Kawamata *et al.*, 2009). É importante mencionar que a Ago1 também apresenta preferência pelo nucleotídeo uracila na primeira posição da sequência a ser carregada (Hu *et al.*, 2009; Okamura *et al.*, 2009; Ghildiyal *et al.*, 2010; Seitz *et al.*, 2011; Wang, 2013). Estudos reportam que pre-miRNAs, em geral, possuem menores valores de MFE que sequências aleatórias, o que lhes é conferido pelo tipo de suas ligações nucleotídicas (Bonnet *et al.*, 2004). De modo simples, os valores de MFE são inversamente proporcionais à estabilidade da molécula. Desse modo, a maior instabilidade observada na extremidade 5' da fita escolhida como miRNA maduro apresenta valores de MFE ligeiramente maiores que outras regiões do *duplex*. Isso ocorre tanto em insetos, quanto em vermes e mamíferos (Khvorova *et al.*, 2003). Em suma, cabe à proteína Argonauta a escolha da fita madura, sendo importante destacar que dado à estabilidade, nem todos os precursores geram miRNAs maduros de ambas as fitas do *duplex* do miRNA.

O complexo de silenciamento induzido por RNA (RISC) constitui o efetor da regulação gênica promovida pelos pequenos RNAs (**FIGURA 1**). A associação do pequeno RNA à uma proteína da família Argonauta é suficiente para formar o complexo RISC e promover regulação dos alvos, embora possa haver a participação de outras proteínas (Rivas *et al.*, 2005; Hutvagner e Simard, 2008; Pratt, A. J. e Macrae, I. J., 2009). Como existem proteínas tipo Argonauta específicas para cada uma das vias de RNAi em insetos, foram nomeados então complexos RISC específicos: o miRISC, possui como principal componente a Ago1 (via de miRNA); o siRISC, cujo principal componente é a Ago2 (via de siRNA); e piRISC, composto pela proteína Aubergine ou pela proteína Piwi (via de piRNA).

1.3 Mecanismos de regulação do alvo

Os miRNAs atuam através do pareamento do tipo Watson-Crick com RNAs-alvo. Como apresentado na **FIGURA 1**, essa complementariedade não precisa ser total, podendo apresentar nucleotídeos não pareados, também chamados de *mismatches* ou “bolhas”. A regulação do alvo, no entanto, depende do reconhecimento e da

complementariedade entre os nucleotídeos 2 a 7 da extremidade 5' do miRNA, os quais são conhecidos como semente ou *seed* (Wightman *et al.*, 1991; Wightman *et al.*, 1993; Nicholson, 1999; Lai, 2002; Brennecke *et al.*, 2005). Os miRNAs que apresentam a região da *seed* com sequência conservada e derivam de um ancestral comum são agrupados em famílias, podendo haver variação dos demais nucleotídeos de sua sequência (Griffiths-Jones, 2003; Kozomara e Griffiths-Jones, 2011; Tehler *et al.*, 2011; Marco, Hooks, *et al.*, 2012). Além disso, miRNAs de uma mesma família tendem a regular RNAs alvos em comum, discutido mais detalhadamente à frente.

A complementariedade à *seed* é crucial para regulação dos alvos (**FIGURA 3**), e mutações nessa região ocasionam perda de função do miRNA (Okamura *et al.*, 2009). Todavia, existem ainda outros nucleotídeos no miRNA que podem potencializar sua atuação como através do pareamento de um maior número de bases da sua extremidade 3' ao RNA alvo (**FIGURA 3**) (Brennecke *et al.*, 2005; Grimson *et al.*, 2007; Nielsen *et al.*, 2007; Bartel, 2009). Além disso, utilizando a estratégia de pareamento de nucleotídeos da região 3' do miRNA, sítios-alvo podem compensar possíveis nucleotídeos não pareados na *seed* (**FIGURA 3G**), possibilitando a ação do miRNA (Grimson *et al.*, 2007; Revisado por: Bartel, 2009).

nesse processo foi proposto em células de mamíferos, e apesar de não ter sido completamente elucidado, parece ser dependente da ligação da Agonauta ao *cap* de metilguanossina do mRNA (m^7G), impedindo sua tradução (Kiriakidou *et al.*, 2007).

Além do mecanismo de silenciamento pela repressão da tradução, o complexo miRISC pode ocasionar o recrutamento de deadenilases e consequente degradação do mRNA pelos mecanismos celulares convencionais de reciclagem de mRNAs, num processo dependente de ATP e da proteína GW182 em mamíferos (Eulalio *et al.*, 2008; Iwasaki *et al.*, 2009; Pratt, A. J. e Macrae, I. J., 2009). Em casos onde há extensa complementariedade da sequência do miRNA com a sequência do mRNA alvo é possível que haja clivagem direta do mRNA pela Ago1, também levando o mRNA à degradação por mecanismos celulares convencionais (Valencia-Sanchez *et al.*, 2006).

Embora haja inúmeros estudos relativos ao silenciamento de alvos por miRNAs, algumas dessas moléculas parecem aumentar a transcrição de genes alvos, como o *Ccnb1* – cuja expressão é aumentada na presença de miR-744 e miR-1186 em células de camundongos (Huang *et al.*, 2012). Nesse modelo, ambos os miRNAs associados à Ago1 parecem se ligar a regiões próximas ao sítio de início de transcrição (TSS), no qual a Ago1 se liga e promove aumento da transcrição pela RNA Pol II e modificações na cromatina, como a metilação da lisina K4 da histona H3 (H3K4me1). Outro trabalho interessante demonstra a capacidade de miRNAs específicos ativarem ou reprimirem alvos dependendo do estágio do ciclo celular em que a célula se encontra (Vasudevan *et al.*, 2007). Os autores propõem que, em seu modelo de estudos, o mecanismo de repressão por miRNAs é uma propriedade de células que estão se proliferando, e os mesmos miRNAs inibem a tradução em células no estágio quiescente (G0) do ciclo celular. Estes aspectos da regulação de alvos pela via de miRNA demonstra sua importância, complexidade, e revela seu potencial como um vasto campo de estudos nas mais diversas áreas da biologia.

1.4 Nomenclatura

A nomenclatura dos miRNAs segue a sistematização definida por Ambros (2003), atualmente mantida pelo banco de dados miRBase (Ambros, 2003; Griffiths-Jones, 2003; 2004; Griffiths-Jones *et al.*, 2006; Griffiths-Jones *et al.*, 2008; Kozomara e Griffiths-Jones, 2011; 2014). Esta nomenclatura é dada segundo o organismo e origem do miRNA no precursor. Como esta sistematização ocorreu após a identificação de alguns miRNAs, tais como let-7 e lin-4 – descobertos em 1993 e 2000 respectivamente em *C. elegans*, isto justifica a existência de miRNAs derivados de animais que não seguem estas regras. Além disso, casos específicos não foram considerados, como a existência de miRNAs-espelho (miRors) que derivam da transcrição bidirecional do seu *locus* gênico (Tyler *et al.*, 2008; Scott *et al.*, 2012). Por isso, recentemente alguns pesquisadores propuseram sua revisão (Desvignes *et al.*, 2015). Entretanto, até o momento não houve manifestação da comunidade científica para uma nova normatização da nomenclatura dos miRNAs.

Os miRNAs recebem um acrônimo de três letras referente ao organismo do qual derivam (por exemplo: dme, para sequências derivadas de *D. melanogaster*; hsa, para sequências de *Homo sapiens* e mmu, para sequências derivadas de *Mus musculus*), seguido de hífen (-) e as informações mir (para precursor) ou miR (para sequência madura) seguido de novo hífen e a numeração da sequência na ordem que foram depositados no miRBase, ou de acordo com ortólogos em outros organismos. Tratando-se de miRNAs maduros, os sufixos -5p ou -3p são atribuídos de acordo com a origem do miRNA a partir de seu precursor (por exemplo: dme-miR-1-5p). Além disso, neste trabalho foi utilizada a nomenclatura “miRNA” para designar a sequência madura, e “mirna” para designar seu precursor. Cabe ressaltar que a adição dos sufixos -5p e -3p só foi incorporada ao nome dos miRNAs a partir da versão 17 do miRBase.

A **FIGURA 4** representa esquematicamente a estrutura de um pri-miRNA, sendo explicitados os sítios de clivagem das enzimas Drosha e Dicer, e em cores a representação dos braços que originam os miRNAs. As sequências derivadas da extremidade 5' do pri-miRNA recebem o sufixo -5p, e as sequências derivadas da extremidade 3' recebem o sufixo -3p.

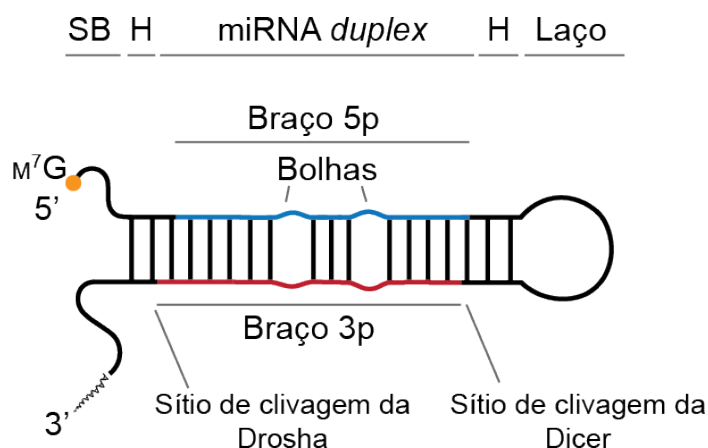


Figura 4: Esquema de um miRNA primário (pri-miRNA) de transcrição independente. São representadas esquematicamente a estrutura de um pri-miRNA, os sítios de clivagem das enzimas Drosha e Dicer; a haste, o laço e as bolhas internas do grampo que permanecem após o corte feito pela Drosha e conferem maior instabilidade ao miRNA nas posições específicas em que se encontram. **Em azul:** miRNA derivado da extremidade 5' do precursor; **Em vermelho:** miRNA derivado da extremidade 3' do precursor; **m⁷G:** 7-metilguanosa ou *cap*; **SB:** Segmento Basal; **H:** Haste.

Cópias idênticas de um mesmo miRNA que apresente origem de diferentes *loci* gênicos num mesmo organismo recebem as numerações -1, -2 sucessivamente, antes dos sufixos 5p ou 3p (por exemplo: dme-miR-6-1-3p). Cópias que apresentam 1 ou 2 nucleotídeos não pareados fora da região da *seed*, recebem as letras do alfabeto latino a, b, c sucessivamente, logo após o número da sequência (por exemplo: dme-miR-92a-3p). Nomenclatura semelhante é adotada para precursores, não havendo apenas os sufixos 5p e 3p mencionados.

1.5 Evolução da via de miRNAs

Um aspecto interessante acerca da via de miRNAs reside na sua conservação entre diferentes espécies, a qual começou a ser verificada através dos estudos do let-7, ainda nos anos 2000 (Pasquinelli *et al.*, 2000). Pasquinelli e colaboradores apresentaram a conservação de sequências deste miRNA em *H. sapiens*, *D. melanogaster*, *Danio rerio* (zebrafish) e *C. elegans*. A partir desse trabalho, foi observada a existência de inúmeros miRNAs ortólogos como exemplos observados entre vermes e humanos, os quais distam em aproximadamente 600 milhões de anos (Lim *et al.*, 2003). De fato, há indícios da presença de miRNAs até nos ramos mais basais da evolução dos metazoários (Berezikov *et al.*, 2011). O miR-100 é conservado evolutivamente em eumetazoários e em bilaterais e parece ter havido grande expansão das sequências de miRNAs (Hertel e Stadler, 2006; Grimson *et al.*, 2008 {Hertel, 2006 #60}). Em vertebrados, parece ter ocorrido nova expansão dos

miRNAs, e a perda destas sequências parece ser incomum nas linhagens descendentes (Berezikov, 2011). Uma visão simplificada da evolução do número de miRNAs é mostrada na **FIGURA 5**.

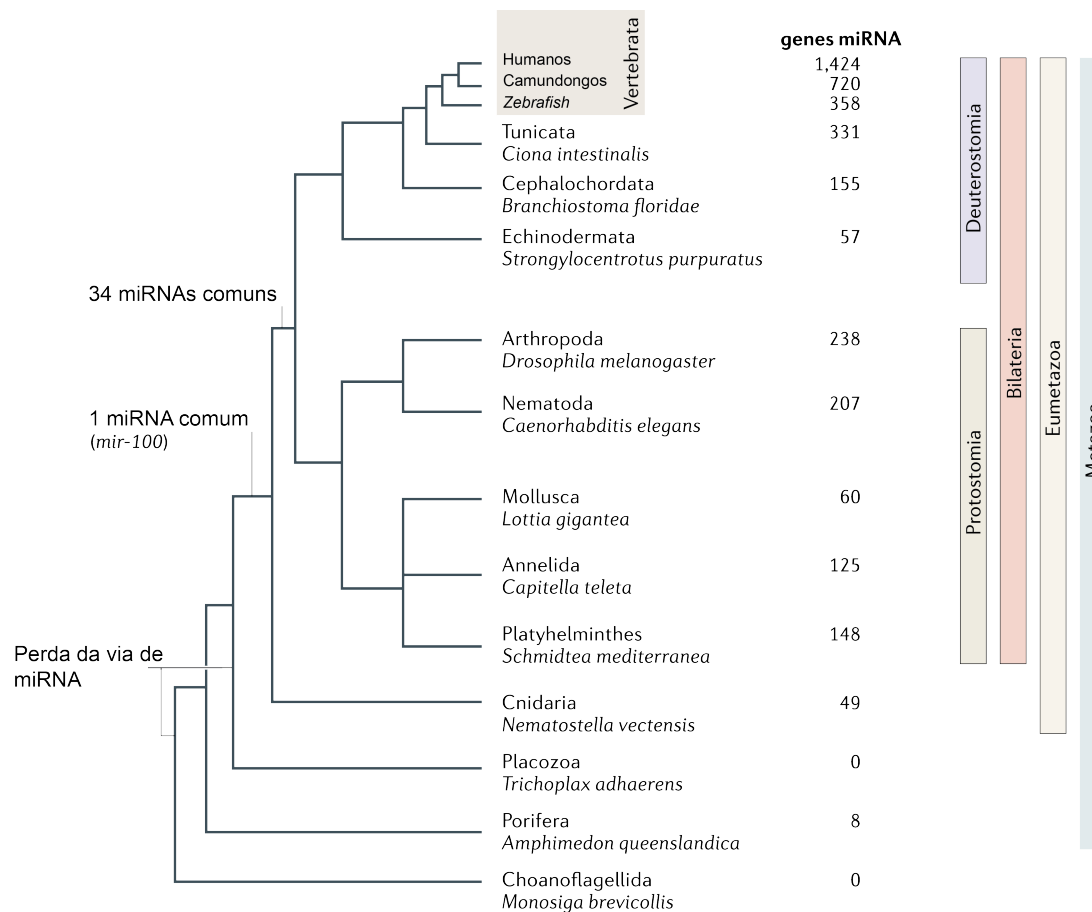


Figura 5: Evolução simplificada das vias de miRNA. O número de miRNAs reflete a complexidade do organismo. A via de miRNAs existe desde ramos mais basais de metazoários e sofreu grande expansão em vertebrados. Adaptado de Berezikov (2011).

Estima-se que, em média, cerca de 0,5 a 1% dos genes de um organismo codificam para miRNAs (Bartel, 2004a). As famílias de miRNAs apresentam elevada similaridade entre si, *seed* idêntica e um ancestral comum (Griffiths-Jones, 2003; Berezikov, 2011). Estas famílias possuem altos níveis de conservação entre diferentes organismos e entre os miRNAs membros da mesma família. Além disso, algumas delas são encontradas ubiquamente com altas taxas de identidade na totalidade de suas sequências, como exemplificado pela família do miRNA let-7 (**FIGURA 6**) – encontrada desde artrópodes até vertebrados (Lee et al., 2007). É possível observar também, que miRNAs como o miR-98 surgiram recentemente na evolução, uma vez que são encontrados apenas na classe Mammalia. A ausência desse miRNA em

organismos mais basais na escala evolutiva pode ser explicada de duas formas: a) houve duplicação gênica seguida de diferenciação do miRNA; ou b) o miRNA ainda não foi encontrado nos demais organismos por limitações técnicas. Essas explicações baseiam-se na premissa de que a perda de um miRNA que existiu no ancestral de uma espécie é um evento raro, embora existam trabalhos que critiquem essa ideia, sugerindo que a evolução da via de miRNAs é mais complexa que os métodos de parcimônia possam explicar (Thomson et al., 2014).

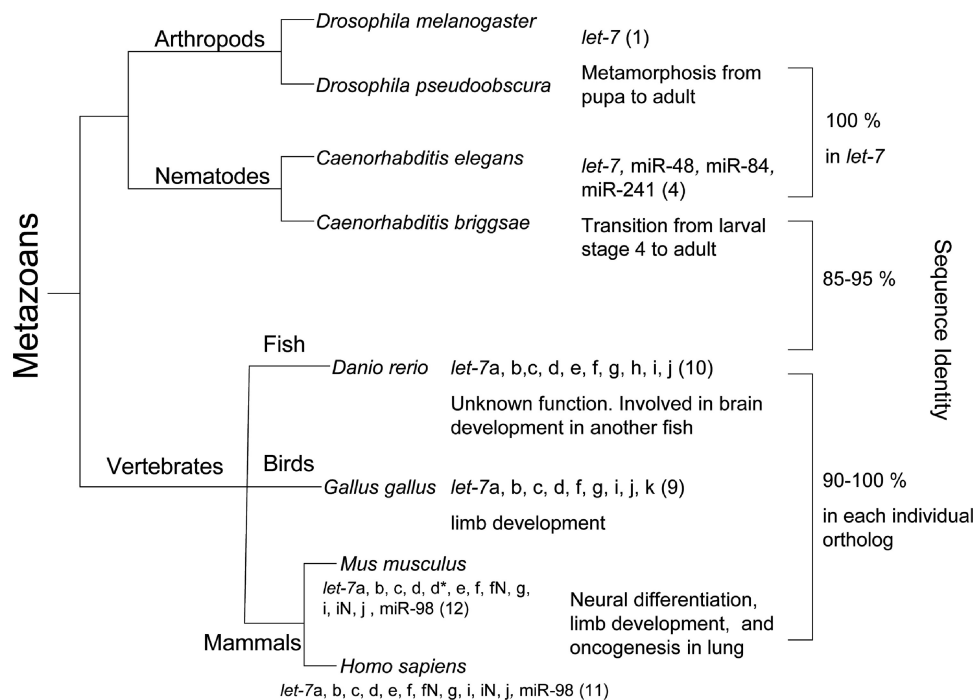


Figura 6: Conservação da família de miRNAs *let-7* em diferentes clados. O número de sequências encontrado em cada espécie, até 2007 (quando o trabalho foi publicado), varia de 1 a 11 (números entre parênteses). A família *let-7* é encontrada em diferentes metazoários – desde artrópodes até vertebrados e o grau de conservação é alto entre todos os membros desta família. Adaptado de: Lee et al. (2007).

Recentemente, Thomson e colaboradores (2014) avaliaram os métodos de parcimônia empregados nas análises evolutivas de miRNAs, os quais se baseiam apenas na presença/ausência das famílias de miRNAs nos organismos, e sugeriram o emprego de análise bayesiana. Esses pesquisadores encontraram diferenças significativas nas árvores filogenéticas propostas em outros estudos. Além disso, como mencionado por eles, a ausência de um miRNA por si não indica sua inexistência na espécie, mas pode estar relacionada a erros de amostragem. Por isso, o desenvolvimento de metodologias que permitam suplantar possíveis fontes de erros (como o estado de montagem e anotação do genoma) faz-se necessário.

Quanto às proteínas da via, estas apresentam alto nível de conservação de sequências entre eucariotos (Shabalina e Koonin, 2008). A fusão entre proteínas que possuem domínio helicase e domínio RNase – observada nas proteínas Drosha e Dicer –, é uma sinapomorfia de eucariotos. De fato, a maquinaria da via de miRNA destes organismos parece ter sido originada a partir de proteínas de arqueas, bactérias e fagos, permitindo a transição entre a dependência de processos biológicos mediados por DNA para processos biológicos mediados por RNA, o que ocasionou aumento substancial da complexidade evolutiva evidenciada no número de miRNAs observados (Shabalina e Koonin, 2008).

Atualmente, uma ferramenta chamada **miRviewer** (Kiezun *et al.*, 2012) foi desenvolvida para apresentar uma a visão global da homologia dos miRNAs em diversos organismos onde foram identificados e encontra-se disponível *online* em: <http://people.csail.mit.edu/akiezun/microRNAviewer/index.html>.

1.6 miRNAs e sua conservação em insetos da ordem Diptera

Desde o princípio dos estudos da via de miRNAs, a presença das proteínas e de diversas sequências de miRNAs têm sido descritas em insetos da ordem Diptera. Nestes organismos são observados miRNAs espécie-específicos, porém também são observados miRNAs com alta taxa de conservação. Todavia, para entender melhor tais níveis de conservação e as origens das divergências entre estes organismos, faz-se necessário compreender alguns aspectos da sua evolução filogenética.

As “moscas verdadeiras” – como são conhecidos alguns insetos da ordem Diptera – compreendem uma das quatro superradiações de insetos e correspondem a aproximadamente 10% de todas as espécies descritas atualmente. As infraordens

Culicinomorpha (em que são classificados os mosquitos *A. aegypti*) e Psycodomorpha (como os mosquitos-palha – *L. longipalpis*) são alguns dos representantes de dípteras basais (Wiegmann *et al.*, 2011). Moscas da fruta (*D. melanogaster*), por sua vez, pertencem à subordem Brachycera, seção Schizophora, cuja radiação se deu de forma rápida e recente (Wiegmann *et al.*, 2011) (**FIGURA 7**).

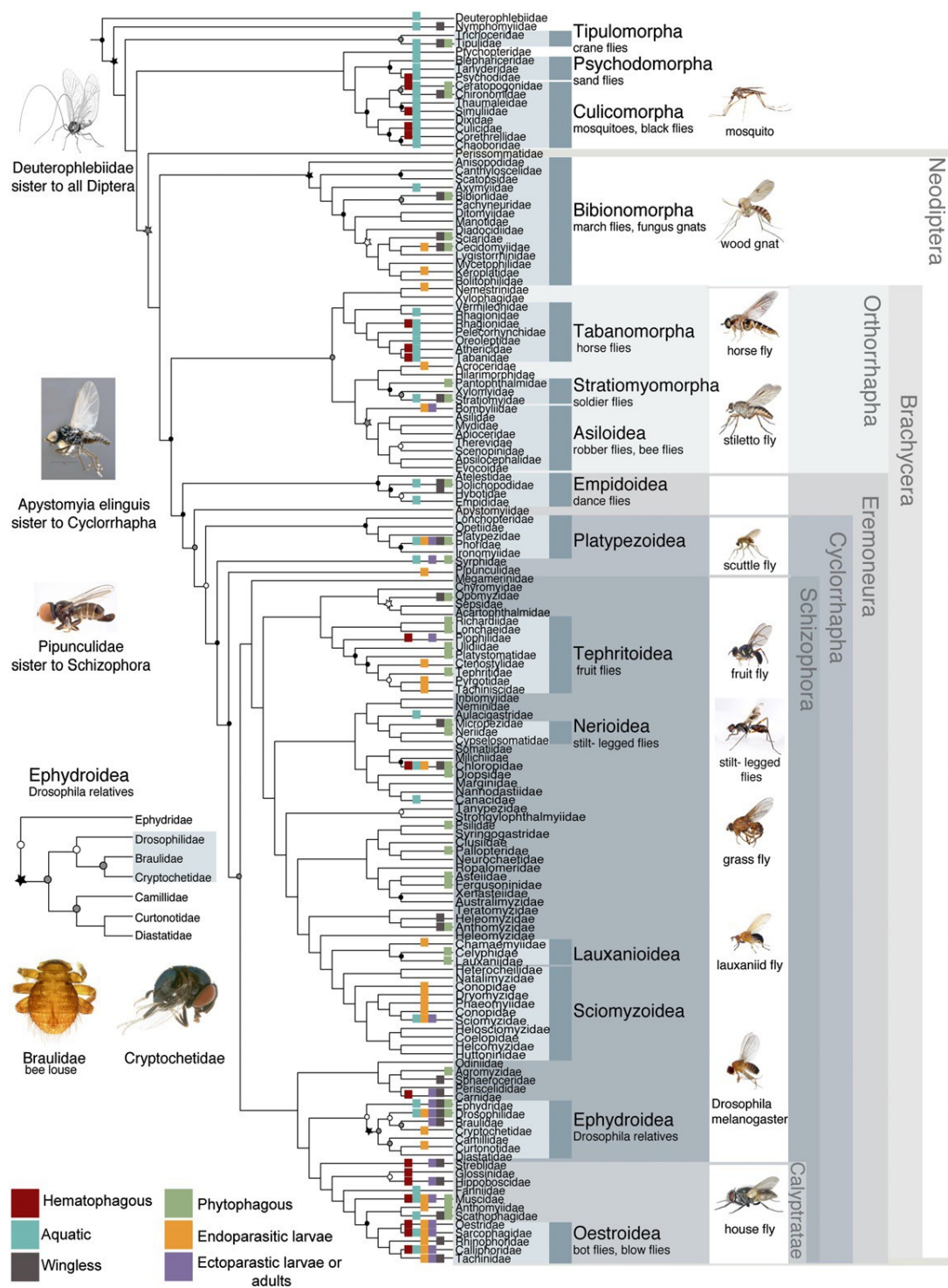


Figura 7: Árvore filogenética das famílias de insetos da ordem Diptera. Em Culicinomorpha encontram-se mosquitos do gênero *Aedes*; em Psychodomorpha encontram-se os insetos do gênero *Lutzomyia* e na seção Schizophora encontram-se o gênero *Drosophila*. O último ancestral comum entre esses três insetos data de aproximadamente 250 milhões de anos. Adaptado de Wiegmann *et al.* (2011).

O organismo modelo *D. melanogaster* é vastamente empregado em estudos de imunologia, biologia molecular e genética (Ashburner e Bergman, 2005). Somado a isso, diversas descobertas nas vias de RNAi foram feitas a partir deste inseto e mostraram-se conservadas em outros Diptera. Como exemplo, tem-se a existência de miRNAs conservados entre *D. melanogaster* e *A. aegypti* (Li et al., 2009); entre *D. melanogaster* e o bicho-da-seda *Bombyx mori* (Cao et al., 2008); e entre *D. melanogaster* e abelhas *Apis mellifera* (Zondag et al., 2012). Tais aspectos reforçam que, apesar da distância evolutiva entre os organismos desta ordem (**FIGURA 7**), a via de miRNAs é conservada e mantém sua função nestes insetos. Os mesmos trabalhos mencionados anteriormente também reportam a existência de miRNAs espécie-específicos. Contudo, esses miRNAs parecem estar sofrendo pressão em sua biogênese, uma vez que as proteínas da via de miRNAs parecem apresentar alta conservação entre os organismos, como previamente discutido, e não há aparente redundância com outras proteínas das demais vias de RNA de interferência. Portanto, essas características tornam os organismos da ordem Diptera bons modelos para estudos de evolução da via de miRNAs.

Justificativa

2. JUSTIFICATIVA

Os miRNAs compreendem uma das classes de pequenos RNAs não codificantes cuja associação com proteínas Argonauta (Ago) forma o complexo RISC – responsável por mediar a regulação pós-transcricional da expressão de inúmeros genes. Uma grande quantidade de proteínas Ago e de miRNAs é encontrada em diferentes eucariotos e, a despeito dessas informações, a maioria dos trabalhos baseia-se em comparações simples das sequências de miRNAs entre espécies ou de seus alvos, sem considerar a biogênese dos mesmos.

Em Diptera, as proteínas Argonauta apresentam função relativamente independente, sendo possível compreender vias específicas de pequenos RNAs, (como a de miRNAs) sem a redundância característica observada em outros organismos como mamíferos. Desse modo, as análises de evolução na biogênese de miRNAs podem ser mais bem estudadas em modelos dessa ordem.

Para estudar esse processo e avaliar a conservação de características intrínsecas aos miRNAs, foram escolhidos três insetos da ordem Diptera que apresentaram alta conservação das proteínas da via de miRNAs em trabalhos anteriores realizados no Laboratório de RNA de Interferência (ICB/UFMG). Dentre esses organismos, dois apresentam importância vetorial de doenças humanas ou de caráter zoonótico (*L. longipalpis* e *A. aegypti*) e o outro trata-se de um modelo animal bem caracterizado (*D. melanogaster*).

Deste modo, este trabalho propõe-se a avaliar características específicas dos miRNAs que permitam compreender aspectos de sua biogênese em diferentes espécies através de ferramentas de bioinformática.

Objetivos

3. OBJETIVOS

3.1 Objetivo geral

Avaliar a evolução dos miRNAs e suas características intrínsecas, utilizando insetos da ordem Diptera como modelo.

3.2 Objetivos específicos

- a) Realizar a predição e anotação dos miRNAs de *Lutzomyia longipalpis*;
- b) Melhorar a anotação dos miRNAs de *Aedes aegypti*;
- c) Definir grupos de miRNAs conservados e não conservados entre *Drosophila melanogaster*, *Lutzomyia longipalpis* e *Aedes aegypti*;
- d) Comparar as características intrínsecas dos miRNAs conservados e não conservados em cada inseto avaliado.

Materiais e métodos

4. MATERIAIS E MÉTODOS

4.1 Origem das amostras

As bibliotecas de pequenos RNAs provenientes de *L. longipalpis* foram obtidas através da coleta desses organismos na cidade de Teresina (região centro-norte do Estado do Piauí), ou através de cultivo celular. Essas amostras foram infectadas com o vírus da estomatite vesicular (VSV). As bibliotecas derivadas de *A. aegypti* foram obtidas de colônias estabelecidas em laboratório após coleta de ovos desses mosquitos em diferentes regiões do Rio de Janeiro e também passaram por infecção, com o vírus da dengue (DENV-4).

Colônias de moscas de fruta selvagens, provenientes de laboratório e infectadas com vírus distintos aos supracitados, foram processadas para geração das bibliotecas de pequenos RNAs oriundas de *D. melanogaster*.

As bibliotecas preparadas fazem parte de trabalhos de outros alunos do Laboratório de RNA de Interferência (ICB/UFMG) ou de colaboradores do grupo (Jean-Luc Imler's Lab – Estrasburgo/França) e, portanto, o processo detalhado de preparação das amostras não faz parte do escopo dessa dissertação.

Todavia, cabe ressaltar que as amostras foram igualmente processadas. Para realização da extração de RNAs foi utilizado o reagente **Trizol®** e metodologia recomendada pelo fabricante. Para sequenciamento das amostras, foi utilizada a plataforma **Illumina HiSeq™ 2500**, situada na *Plateforme Biopuces & Séquençage, Institut de Génétique e Biologie Moléculaire et Cellulaire* (IGBMC), em Estrasburgo, na França. A **TABELA 2** descreve as bibliotecas utilizadas neste trabalho.

Tabela 2: Descrição das bibliotecas de pequenos RNAs utilizadas.

Biblioteca número	Organismo	Descrição
1	Llo*	Mock pool 2dpa
2	Llo*	VSV-GFP pool 2dpa
3	Llo*	Mock pool 2dpa
4	Llo*	Mock pool 4dpa
5	Llo*	Mock pool 6dpa
6	Llo	VSV-GFP pool 2dpa
7	Llo	VSV-GFP pool 4dpa
8	Llo	VSV-GFP pool 6dpa
9	Llo*	LL5 – cultivo celular
10	Llo*	Lulo – cultivo celular
11	Aae	DENV-4 pool 4dpa
12	Aae	DENV-4 pool 8dpa
13	Aae	DENV-4 pool 14dpa
14	Dme	pool VSV injetado
15	Dme	pool DCV injetado
16	Dme	pool SINV injetado

Llo: *Lutzomyia longipalpis*; **Aae:** *Aedes aegypti*; **Dme:** *Drosophila melanogaster*; **dpa:** dias pós alimentação; **dpi:** dias pós injeção; **VSV:** Vesicular Stomatitis Virus; **GFP:** proteína verde fluorescente (tradução livre); **DENV-4:** Dengue virus sorotipo 4; **DCV:** Drosophila C virus; **SINV:** Sindbis virus; *: Biblioteca não-utilizada nas análises comparativas.

Neste trabalho, foi utilizada a penúltima versão do genoma de *L. longipalpis* disponível no banco de dados de “Vetores Invertebrados de Patógenos Humanos” – tradução livre de VectorBase – sob identificação LlonJ1.1 (disponibilizada *online* a partir de 25/02/2014 através do endereço eletrônico: <https://www.vectorbase.org/organisms/lutzomyia-longipalpis/jacobina/LlonJ1.1>). Esta versão apresenta 11.532 *scaffolds* e nenhum cromossomo completo, o que reflete a necessidade de melhoria.

Quanto a *A. aegypti*, foi utilizado genoma disponível no VectorBase sob identificação AaegL3.3, e disponibilizada *online* a partir de 21/10/2014 através do endereço eletrônico <http://www.vectorbase.org/organisms/aedes-aegypti/liverpool/AaegL3.3>.

Após as etapas iniciais de preparo, sequenciamento e pré-processamento das amostras, o *pipeline* desenvolvido (**FIGURA 8**) foi dividido em duas partes (**FIGURA 8B** e **FIGURA 8C**). Na primeira estão representadas as etapas de identificação e comparação das características dos miRNAs identificados através do programa **miRDeep2** (Friedlander *et al.*, 2012). A segunda, por sua vez, apresenta a metodologia

desenvolvida para identificação e análise comparativa das características dos miRNAs identificados através de comparação com organismos próximos. A descrição detalhada de cada etapa encontra-se nos tópicos a seguir.

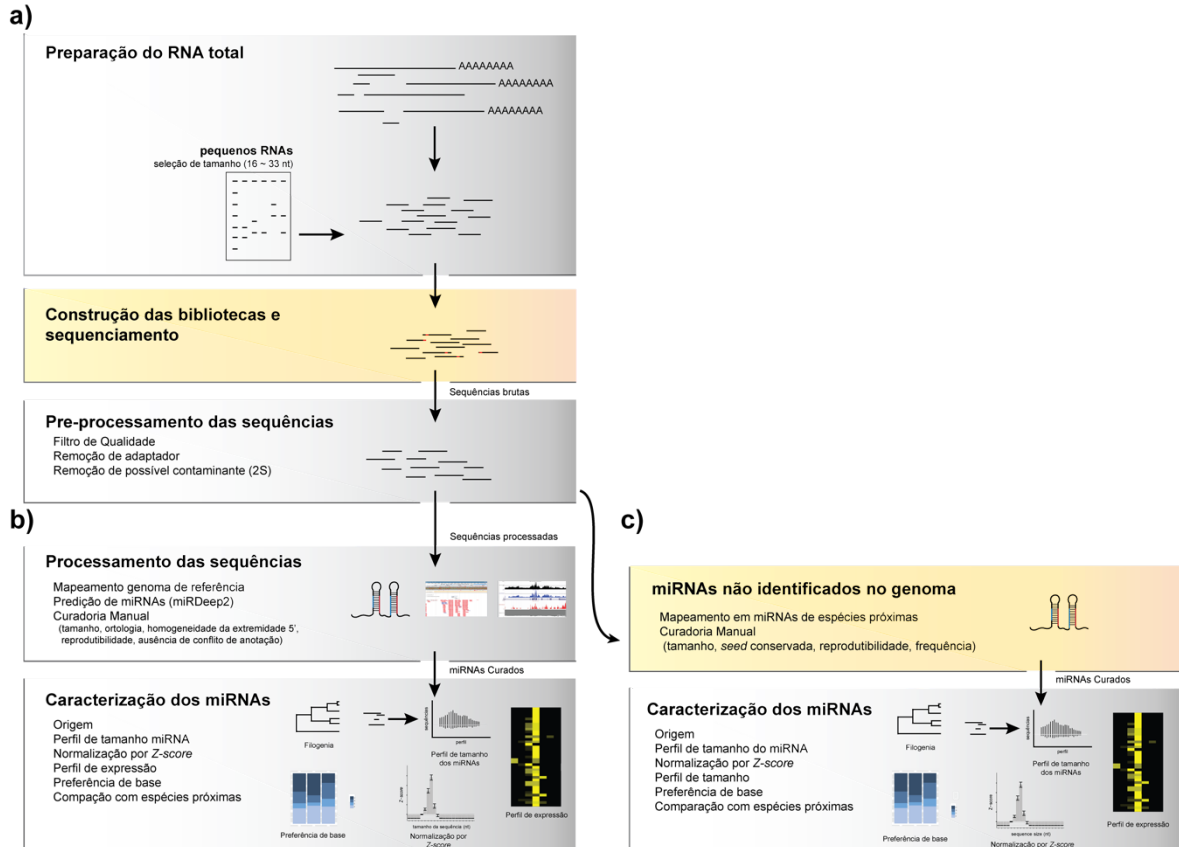


Figura 8: Pipeline de identificação e análise dos miRNAs em insetos da ordem Diptera. Após preparo de amostras, sequenciamento e pré-processamento de sequências (**a**), o *pipeline* foi dividido em duas partes, de acordo com a estratégia utilizada. **b)** Análises desenvolvidas através do uso do programa miRDeep2. **c)** Análises desenvolvidas para identificação de miRNAs não preditos pelo programa miRDeep2, possivelmente devido a falhas de montagem do genoma.

4.2 Pre-processamento das sequências

As sequências derivadas das bibliotecas de pequenos RNAs foram submetidas a filtro de qualidade e remoção dos adaptadores através dos programas **fastx_quality_filter** e **fastq_clipper**, respectivamente. Ambos os programas fazem parte do pacote **FastX-Toolkit** (versão 0.0.14) disponível *online* no endereço: http://hannonlab.cshl.edu/fastx_toolkit/index.html. Sequências com qualidade *phred* menor que 33 e com tamanho inferior a 15 nt, após retirada dos adaptadores, foram descartadas.

A seguir, as sequências derivadas das bibliotecas de *L. longipalpis* foram mapeadas utilizando-se o programa **Bowtie** versão 1.1.1 – (Langmead *et al.*, 2009) contra possível contaminante derivado do RNA ribossomal 2S

(TCTCGGAAACTGCATGGACCACGTATGGTTGAGTGTTCGTAAAATAGCAAT)

(Pavlakakis *et al.*, 1979), uma vez que não foi feita depleção de rRNA durante o preparo das bibliotecas. Somente as sequências não mapeadas nesse processo foram utilizadas para predição de miRNAs.

4.3 Identificação de miRNAs baseada no genoma disponível

Para predição dos miRNAs de *L. longipalpis* e *A. aegypti*, foi utilizado o programa miRDeep2. Este programa permite a identificação de miRNAs em bibliotecas de pequenos RNAs derivados de sequenciamento em larga escala.

Foram utilizados dois *scripts* do pacote miRDeep2: **mapper.pl** e **miRDeep2.pl**. O primeiro tem por finalidade mapear as sequências no genoma de referência indexado. O segundo, por sua vez, consiste no *core* do programa, o qual estende regiões mapeadas no genoma, através do primeiro *script*, e avalia a estrutura secundária do precursor, bem como similaridade com miRNAs já anotados (quando utilizados como *input* ou modelo).

Para o primeiro *script*, foram utilizados os parâmetros: `-d -e -h -i -j -m -v -p`. O parâmetro `-d` permite utilização de múltiplas bibliotecas cuja identificação é dada em um arquivo de texto chamado `config.txt`. O parâmetro `-e` indica que as bibliotecas estão em formato `fastq`; `-h` refere-se à transformação desse formato em

fasta; *-i* indica que as sequências são de RNA e devem ser transformadas no alfabeto de sequências DNA. Além disso, quaisquer outras posições que contenham caracteres diferentes de *a, c, g, t, u, n, A, C, G, T, U, N* são excluídas, mediante o parâmetro *-j*.

Para o segundo *script*, foi utilizado o parâmetro referente ao nome do arquivo que continha as sequências de miRNAs maduras, os demais parâmetros utilizados foram *default*.

Por se tratar de um organismo bem caracterizado, as bibliotecas derivadas de *D. melanogaster* não foram submetidas a processo de predição de miRNAs. Nesse caso, foram utilizadas as sequências anotadas disponíveis no banco de dados miRBase (versão 21).

4.3.1 *Lutzomyia longipalpis*

Por se tratar de um genoma cuja montagem e anotação não estão finalizados, e por esse motivo não existir miRNAs anotados, foram utilizados os miRNAs do subfilo Hexapoda (**TABELA 3**), disponíveis no banco de dados miRBase (v.21) como modelo para que o miRDeep2 realizasse a predição dos miRNAs de *L. longipalpis*.

Tabela 3: Número de miRNAs e precursores do subfilo Hexapoda disponíveis no miRBase (versão 21) e utilizados para predição em *Lutzomyia longipalpis*.

miRBase – versão 21				
Acrônimo	Organismo	mirnas	miRNAs	Versão do genoma
aae	<i>Aedes aegypti</i>	101	124	AaegL1
aga	<i>Anopheles gambiae</i>	66	65	AgamP3
ame	<i>Apis mellifera</i>	254	259	AMEL4.5
api	<i>Acyrtosiphon pisum</i>	123	103	Acyr_2.0
bmo	<i>Bombyx mori</i>	487	563	SILKDB2.0
cqu	<i>Culex quinquefasciatus</i>	76	75	CpipJ1
dan	<i>Drosophila ananassae</i>	76	75	dana_r1.3_FB2010_03
der	<i>Drosophila erecta</i>	81	78	dere_r1.3_FB2010_02
dgr	<i>Drosophila grimshawi</i>	82	72	dgri_r1.3_FB2010_02
dme	<i>Drosophila melanogaster</i>	256	466	BDGP5.0
dmo	<i>Drosophila mojavensis</i>	71	71	dmoj_r1.3_FB2010_02
dpe	<i>Drosophila persimilis</i>	75	69	dper_r1.3_FB2010_02
dps	<i>Drosophila pseudoobscura</i>	210	273	dpse_r2.26_FB2012_01
dse	<i>Drosophila sechellia</i>	78	76	dsec_r1.3_FB2010_02
dsi	<i>Drosophila simulans</i>	135	178	dsim_caf1
dvi	<i>Drosophila virilis</i>	134	237	FB2011_07
dwi	<i>Drosophila willistoni</i>	77	72	dwil_r1.3_FB2010_02
dya	<i>Drosophila yakuba</i>	76	75	dyak_caf1
hme	<i>Heliconius melpomene</i>	92	97	HelMel1.1
lmi	<i>Locusta migratória</i>	7	14	N/C
mse	<i>Maduca sexta</i>	98	93	Msex1.0
ngi	<i>Nasonia giraulti</i>	32	32	N/C
nlo	<i>Nasonia longicornis</i>	28	28	N/C
nvi	<i>Nasonia vitripennis</i>	53	53	Nvit_2.0
tca	<i>Tribolium castaneum</i>	220	430	Tcas3.0

mirnas: Precursores; **miRNAs:** Sequências maduras; **N/C:** Não consta

Para aumentar a cobertura da predição, foram incluídas todas as bibliotecas de *L. longipalpis* sequenciadas (TABELA 4 – tópico 5.1). Todavia, somente as replicatas de uma mesma origem (isto é, *pool* de *L. longipalpis* infectados e coletados em diferentes tempos) foram utilizadas para as análises de comparação das sequências de miRNAs maduros.

4.3.2 *Aedes aegypti*

Para melhorar a anotação dos miRNAs de *A. aegypti*, as bibliotecas de pequenos RNAs derivadas desse organismo foram submetidas à predição de miRNAs pelo mesmo programa, tendo o arquivo “modelo” substituído por outro no qual constam os 101 miRNAs da espécie, disponíveis no banco de dados miRBase (v.21).

4.4 Identificação de miRNAs baseada em comparação com organismos próximos

Todos os miRNAs de *D. melanogaster* disponíveis no miRBase (v.21) foram utilizados como referência para mapeamento das sequências derivadas das bibliotecas de *L. longipalpis* e de *A. aegypti*. Além disso, os miRNAs previamente preditos e curados de cada uma dessas espécies também foram utilizados como referência para mapear as bibliotecas derivadas dos diferentes organismos estudados.

Essas análises foram realizadas através do programa **Bowtie** (versão 1.1.1), parâmetros $-n\ 0\ -l\ 9\ -v\ 2$. O parâmetro $-n$ refere-se ao número de *mismatches* permitidos na região da *seed*, e cujo tamanho é definido pelo parâmetro $-l$. O parâmetro $-v$, por sua vez, indica o número de *mismatches* aceitos. Como se pode perceber, a análise foi estridente ao definir a *seed* ($-l\ 9$) como uma sequência de tamanho superior à definição real (nucleotídeos nas posições 2 a 7 da extremidade 5' do miRNA). Além disso, não foram aceitos *mismatches* nessa região ($-n\ 0$).

Os miRNAs cujas sequências mapeadas apresentaram: perfil de tamanho entre 19 e 24 nt; expressão em mais de uma biblioteca da triplicata avaliada, e que estavam ausentes na predição anterior (utilizando somente o genoma de referência do organismo do qual as bibliotecas foram derivadas), passaram a ser referidos como miRNAs derivados de regiões do genoma não presentes na versão da montagem utilizada. Além disso, todas as bibliotecas foram mapeadas contra os vírus com os quais os organismos foram infectados a fim de avaliar possíveis contaminações entre as amostras durante o preparo das bibliotecas. As proporções dos miRNAs avaliados entre as bibliotecas também foram consideradas a fim de garantir que os resultados encontrados fossem acurados.

4.5 Análise da presença de SNPs nos miRNAs identificados por comparação com organismos próximos

A fim de garantir que os miRNAs identificados através do mapeamento direto das bibliotecas nos miRNAs de outros organismos não fossem oriundos de artefato ou de contaminação entre as bibliotecas, foi feita análise da presença de polimorfismos únicos de nucleotídeos (SNPs) que poderiam ser reflexo da diferença entre o miRNA nestes organismos. Deste modo, era esperado que os miRNAs compartilhados e cuja sequência não fora identificada previamente no organismo do qual as bibliotecas foram preparadas, apresentassem alguns nucleotídeos diferentes dos seus respectivos ortólogos. Esta diferença, no entanto, não deveria ser observada na região da *seed*.

Para isto, os resultados do mapeamento das bibliotecas contra os miRNAs oriundos dos demais organismos foram avaliados utilizando o programa **Tablet** (versão 1.15.09.01) (Milne *et al.*, 2013). Sequências com percentual de *mismatches* igual ou superior a 0,6% (percentual escolhido com o objetivo de evitar alterações por erro de sequenciamento, invés de SNPs) tiveram as posições de seus nucleotídeos verificadas a fim de identificar o correto nucleotídeo daquela posição. Frequências similares de nucleotídeos diferentes na mesma posição foram consideradas formas alternativas e ambos nucleotídeos foram reportados como corretos.

4.6 Curadoria manual

Os miRNAs preditos através do programa miRDeep2 foram submetidos à curadoria manual. O não atendimento a 3 ou mais critérios, ou o não atendimento de pelo menos duas características importantes – assinaladas por “*” – foram utilizados como critérios de exclusão dos miRNAs preditos das análises subsequentes. Além disso, em caso de dúvidas, utilizou-se a frequência mínima de 100 mapeamentos totais no miRNA como critério de exclusão.

Nesta etapa, foram avaliadas as seguintes características:

- Estrutura do grampo (*): a estrutura (presença de bolhas entre 8-20 nt, presença de 2-3 nt 3' *overhang*) e o tamanho (60-80 nt) foram verificados, uma vez que são importantes características para correto processamento do miRNA pelas enzimas Droscha e Dicer (Lee *et al.*, 2003; Forstemann *et al.*, 2005; Jiang *et al.*, 2005; Saito

et al., 2005; Tomari e Zamore, 2005; Zeng *et al.*, 2005; Macrae *et al.*, 2007; Ruby *et al.*, 2007; Tsutsumi *et al.*, 2011; Okamura *et al.*, 2013; Starega-Roslan *et al.*, 2015b);

- Distribuição de tamanho das sequências mapeadas (*): A maioria dos miRNAs descritos para insetos apresenta tamanho entre 20 e 24 nt (Revisado por Bartel, 2004b). Assim, foram avaliados se os perfis de tamanho do miRNAs preditos se enquadravam majoritariamente nessa faixa;
- Ausência de anotação de outros genes/RNAs disponíveis no banco de dados Vector Base (*) (Giraldo-Calderon *et al.*, 2015);
- **JBrowse**: presença de *phasing* e origem em ambas as fitas não deveriam ser encontrados (*) (Skinner *et al.*, 2009; Chiang *et al.*, 2010);
- Ausência de conflito com regiões repetitivas obtidas através de análise pelo programa **Repeat Masker** (Smit, 1996; Chiang *et al.*, 2010);
- Reprodutibilidade: o miRNA deve estar presente em mais de uma biblioteca analisada;
- Homogeneidade: $\geq 70\%$ de todas as sequências mapeadas no miRNA devem iniciar-se na mesma posição 5', uma vez que as variações na posição dos nucleotídeos iniciais ocasionam alteração de *seed* e, deste modo, dos respectivos alvos (Chiang *et al.*, 2010; Marco *et al.*, 2010);
- Presença de sequência passageira, para garantir que os miRNAs derivaram de um mesmo precursor e não representam mapeamento inespecífico (Chiang *et al.*, 2010);

4.7 Anotação de miRNAs

Após a predição baseada na versão do genoma disponível, ou identificação através de comparação com organismos próximos, e curadorias, os miRNAs de *L. longipalpis* foram anotados segundo ortologia (isto é, receberam a mesma numeração do miRNA com o qual apresentaram maior similaridade) e armazenados em formato GFF3.

Arquivos GFF3 são arquivos de anotação formados por 9 colunas (separadas por espaçamento tabular), cuja sequência anotada recebe uma identificação unívoca e suas posições inicial e final do genoma são indicadas.

Além disso, são incluídas a orientação no genoma (fita positiva ou negativa) e outros atributos que auxiliem no seu rastreamento. Esses arquivos podem ser tipo específico (exemplo: GFF3 de anotação de miRNAs) ou podem conter toda a anotação referente a determinado genoma (como regiões codificadoras, tRNA, íntrons, entre outros – em um mesmo arquivo).

A mesma metodologia foi aplicada para melhoria da anotação dos miRNAs de *A. aegypti*, uma vez que parte deles se encontra na base de dados miRBase e parte está disponível em dados da literatura (quando identificados e ainda não anotados) (Akbari *et al.*, 2013; Hu *et al.*, 2015).

4.8 Análises comparativas

4.8.1 miRNAs compartilhados entre os três organismos

Os mirnas únicos, isto é, apenas uma cópia de cada precursor (quando existirem mais de uma) em um mesmo organismo, foram analisados com o objetivo de verificar sua presença em cada espécie. Para isso, foi gerado um diagrama de Venn através da ferramenta disponível *online* no endereço: <http://bioinformatics.psb.ugent.be/webtools/Venn/>.

Os miRNAs maduros gerados por cada um dos precursores foram avaliados a fim de verificar sua presença em cada organismo. Aqueles cuja existência foi compartilhada, foram considerados conservados e aqueles encontrados em apenas uma das espécies foram classificados como não- conservados. Ambos os grupos

foram tratados separadamente para as análises de expressão e perfil de tamanho, as quais serão descritas em tópicos posteriores.

4.8.2 Origem genômica dos miRNAs

Os arquivos GFF3 referentes aos miRNAs (identificados nos genomas disponíveis) anotados para cada um dos três organismos estudados foram comparados com o GFF3 de anotação genômica total, através do programa **IntersectBed** (v2.17.0) (Quinlan e Hall, 2010). Os resultados das análises foram plotados no programa **R** (Team, 2011).

Nessa etapa não foram utilizados os miRNAs identificados através da comparação com organismos próximos.

4.8.3 Avaliação da expressão

Para análise de expressão dos miRNAs curados manualmente, as bibliotecas de pequenos RNAs de cada organismo foram mapeadas contra seus respectivos miRNAs através do programa Bowtie, permitindo-se 1 *mismatch* (no caso dos miRNAs anotados após predição com miRDeep2) ou 2 *mismatches* fora da *seed* (no caso dos miRNAs anotados através da análise de comparação com organismos próximos).

Além disso, as sequências mapeadas foram filtradas por tamanho, entre 19 e 24 nt e a soma destes foi utilizada para normalização de “sequências por milhão” (RPM), cuja média das triplicatas foi comparada entre as espécies. A normalização por RPM refere-se à contagem de determinada sequência, dividida pela somatória das sequências mapeadas e multiplicada por 1 milhão. Os miRNAs foram ordenados em ordem decrescente segundo a expressão (em RPM) de *D. melanogaster* e sua correlação foi avaliada através do teste correlação de *Pearson* utilizando o programa **R** por meio da função embutida **cor.test**.

Os 20 miRNAs mais expressos em cada organismo foram comparados segundo sua ordem de expressão. A partir da contagem normalizada (RPM), seguiu-se a verificação do percentual correspondente, dividindo-se o RPM de cada miRNA pela soma dos RPMs de todos os miRNAs mapeados nas bibliotecas do organismo.

4.8.4 Expressão dos miRNAs segundo origem no precursor de cada braço no grampo do pre-miRNA

O miRNA foi nomeado de forma diferente, de acordo com a origem no precursor, isto é, recebeu os sufixos -5p e -3p caso derivasse das porções 5' ou 3' do grampo, respectivamente, as quais são também conhecidas como “braços”. Essa nomenclatura seguiu o padrão adotado pelo miRBase.

Desse modo, a expressão de cada “braço” é diferente e pode-se calcular a razão (*ratio*) dessa expressão. Essa análise já foi realizada em outros trabalhos sob o nome de *Relative Arm Usage* – uso relativo do braço, em tradução livre (Marco *et al.*, 2010).

Para calcular a expressão relativa dos miRNAs em cada braço, a contagem normalizada de cada miRNA derivada do braço 5p foi dividida pela contagem normalizada do miRNA derivado do braço 3p do mesmo precursor. Os valores foram plotados em escala logarítmica. Valores positivos, referem-se à maior expressão do braço 5p e os valores negativos, por sua vez, referem-se à maior expressão do braço 3p. O valor 0 indica expressão idêntica de ambos os braços. A correlação entre a expressão dos miRNAs de espécies diferentes foi avaliada através do teste de correlação de *Pearson* no programa **R**, por meio da função **cor.test**.

Para análise da diferença de expressão dos miRNAs conservados em cada organismo, foi calculado o *ratio* de cada miRNA conservado entre os insetos, dividindo-se seu valor RPM pelo respectivo valor correspondente em outro inseto. A seguir, foi calculado o desvio-padrão (σ) da média de todas as razões através da função **desvpad.a** do programa **Microsoft Excel** (Pacote Office 2016 – sistema operacional Windows). Os valores *ratio* diferentes em pelo menos $1,5\sigma$ foram considerados “discrepantes” e os demais valores “similares”. Onze miRNAs de cada grupo foram, então, analisados estatisticamente a fim de verificar padrões de expressão que permitissem entender as diferenças observadas.

4.8.5 Perfil de tamanho dos miRNAs

As sequências mapeadas em cada miRNA foram separadas segundo seu tamanho através de *scripts* em Perl (desenvolvido por membros do Laboratório de RNA de Interferência – ICB/UFMG) e plotadas em **R**, por meio da função **heatmap.2** do pacote **ggplots**.

Essas análises também foram feitas utilizando-se normalização por *Z-score* através de *script* desenvolvido em Perl, tal como mencionado anteriormente. De forma resumida, a média das sequências mapeadas de tamanho entre 19 e 25 nt foi subtraída da contagem de mapeamento por tamanho, dividindo-se o valor encontrado pelo desvio-padrão da média. Os resultados foram plotados usando a linguagem **R**.

4.8.6 Enriquecimento de base

Para avaliar os padrões moleculares associados aos miRNAs, foi verificada a preferência de base da posição 5' dos pequenos RNAs mapeados nos miRNAs descritos.

A partir da contagem total de nucleotídeos por posição, seguiu-se a verificação do percentual correspondente, dividindo-se o valor obtido para cada nucleotídeo pelo número total de nucleotídeos na posição referida.

4.8.7 Comparação de miRNAs conservados e com perfis de tamanho distintos entre os organismos

Os perfis de tamanho dos miRNAs conservados foram categorizados conforme similaridade entre os três insetos em: a) idêntico – quando o pico no *heatmap* em todos os organismos apresentou dominância de um dos tamanhos;

b) similar – quando o pico apresentou dominância por um dos tamanhos e ligeira variação entre os demais; c) discrepante – quando não atendeu às exigências anteriores. Foram escolhidos 10 miRNAs para avaliação estatística de tamanho e estabilidade dos precursores preditos e comparados a um grupo “controle”.

Para isto, foi feita verificação do tamanho do grampo, bem como o grau de estabilidade – conferido pelo MFE – através do programa **RNAfold** (versão 2.1.9) (Hofacker, 2003). Os valores obtidos foram comparados com os resultados de outros

miRNAs conservados cujo perfil de tamanho não variou entre os insetos. A tabela dos miRNAs usados nesta análise (**TABELA 12**) encontra-se no tópico 5.5.

4.8.8 Avaliação das diferenças de perfis de miRNAs segundo expressão e tamanho

A fim de entender as diferenças observadas nos perfis de alguns miRNAs, foi criado um sistema para atribuir valor a cada característica conservada. Neste, são atribuídos pesos às características analisadas. A contribuição dos *scores* parciais pode ser interpretada segundo **TABELA 4**.

Tabela 4: Scores para avaliação das características dos miRNAs conservados

Score	Característica	Significado
2	Expressão	Expressão conservada
1	Expressão	Expressão discrepante
2	Perfil de tamanho	Perfil idêntico nas 3 espécies
1	Perfil de tamanho	Perfil similar nas 3 espécies
0	Perfil de tamanho	Perfil discrepante entre as espécies

Como é possível verificar, os *scores* parciais variaram entre 0 e 2, segundo característica avaliada. O perfil de tamanho apresenta três atribuições possíveis:

a) perfil idêntico – quando *heatmap* apresentou prevalência de tamanho idêntica em todos os organismos; b) perfil similar – quando o *heatmap* apresentou dominância por um tamanho, mas houve variação em um ou mais *heats*; c) perfil discrepante – quando não é possível identificar as similaridades mencionadas anteriormente.

Resultados e Discussão

5. RESULTADOS E DISCUSSÃO

5.1 Bibliotecas de pequenos RNAs

Na **TABELA 5** estão indicadas as bibliotecas utilizadas, o número de sequências inicial, o número de sequências após pré-processamento e o número de sequências mapeadas no respectivo genoma de referência.

Tabela 5: Número de sequências após pré-processamento e seu respectivo mapeamento no genoma.

Biblioteca número	Organismo	Número de sequências inicial	Após pré-processamento (%)	Mapeadas no genoma (%)
1	Llo*	12.546.269	11.050.675 (88,08)	8.466.872 (76,61)
2	Llo*	16.630.933	16.378.540 (98,48)	12.389.829 (75,65)
3	Llo*	10.679.360	3.579.913 (33,52)	2.719.107 (75,95)
4	Llo*	9.463.241	3.357.798 (35,48)	2.353.024 (70,08)
5	Llo*	8.766.757	4.128.224 (47,09)	2.615.516 (63,36)
6	Llo	8.500.448	3.902.405 (45,91)	2.893.103 (74,14)
7	Llo	8.109.613	2.713.458 (33,46)	1.778.859 (65,56)
8	Llo	8.587.131	4.345.701 (50,61)	3.034.281 (69,82)
9	Llo*	-	9.916.482	7.982.228 (80,49)
10	Llo*	-	6.841.508	4.840.232 (70,75)
11	Aae	11.101.469	10.561.397 (95,14)	9.427.128 (89,26)
12	Aae	16.702.736	15.867.172 (95,00)	14.113.442 (88,95)
13	Aae	17.333.444	15.212.488 (87,76)	12.859.115 (84,53)
14	Dme	24.474.261	21.695.377 (88,65)	14.098.022 (64,98)
15	Dme	15.786.440	15.779.078 (99,95)	11.448.335 (72,55)
16	Dme	4.234.079	4.231.439 (99,94)	3.188.759 (72,36)

Llo: *Lutzomyia longipalpis*; **Aae:** *Aedes aegypti*; **Dme:** *Drosophila melanogaster*; *: Biblioteca não-utilizada nas análises comparativas.

5.2 Identificação e anotação de miRNAs e precursores

Através do *pipeline* inicial (**FIGURA 8B**) foram identificados 94 precursores de miRNAs em *L. longipalpis* e 130 em *A. aegypti*. Comparados aos 256 precursores de miRNAs existentes em *D. melanogaster*, tal resultado sugere que provavelmente ainda não foram identificados todos os miRNAs dos demais insetos. Além disso, menos de 50 precursores foram identificados como conservados entre as três espécies, dos quais estavam ausentes alguns dos miRNAs mais expressos em *D. melanogaster* (como miR-1 e miR-8), cuja função está relacionada, entre outras coisas, à regulação da diferenciação do músculo cardíaco e ao padrão de pigmentação e habilidade de eclosão, respectivamente.

Para verificar se era possível expandir o número de miRNAs identificados, as bibliotecas de miRNAs de *A. aegypti* foram mapeadas contra os miRNAs de *L. longipalpis* e de *D. melanogaster* (o mesmo foi feito para *L. longipalpis*). Esta etapa permitiu encontrar novas sequências por comparação com organismos próximos e, portanto, foi adicionada ao *pipeline* inicial (**FIGURA 8C**).

A estratégia global (**FIGURA 8**) permitiu a identificação de 124 precursores em *L. longipalpis*, dos quais derivam pelo menos 206 miRNAs maduros. É importante ressaltar que o número de novos precursores anotados em *L. longipalpis* através da comparação com organismos próximos (30) corresponde a aproximadamente 1/3 dos 94 inicialmente anotados. A não identificação destas sequências pelo programa miRDeep2 pode ser explicada pelo fato de o mesmo utilizar o genoma disponível como referência, o qual ainda está em fase de montagem e anotação.

Quando aplicada a *A. aegypti*, a mesma metodologia permitiu ampliar a anotação de precursores e de miRNAs depositados no miRBase (v.21). Houve o acréscimo de 48 precursores e 84 miRNAs maduros, aos 101 precursores e 124 miRNAs já anotados, o que corresponde a um aumento de 47,5% e 67,7% precursores e miRNAs maduros, respectivamente. Além disso, as origens dos miRNAs segundo o braço do precursor (isto é, derivadas da porção 5' e 3' – sufixos -5p e -3p, respectivamente) também foram acrescentadas a todos os miRNAs já anotados que não dispunham de tal informação.

Os miRNAs não identificados pelo programa miRDeep2 foram então analisados segundo a presença em mais de uma biblioteca, tamanho conservado entre 19 e 24 nt e quanto à presença de SNPs, com o objetivo de encontrar diferenças nos miRNAs

maduros que refletissem a sequência do genoma do qual se originaram. Como discutido anteriormente, as famílias de miRNAs apresentam alta conservação da *seed* entre seus membros, porém a identidade total de uma sequência pode ser de aproximadamente 85% a 95% (como observado na **FIGURA 6**, para a família do let-7). Portanto, é factível sugerir que o percentual restante da variação seja devido à pressão evolutiva imposta pelos alvos, quanto pela evolução natural do genoma no qual o miRNA deriva.

O programa Tablet identifica o percentual de *mismatches* encontrados nas sequências utilizadas e sinaliza os nucleotídeos que variaram em determinada posição (SNPs), sendo possível identificar casos em que dois ou mais nucleotídeos podem ser igualmente encontrados. É possível verificar que os percentuais da presença de SNPs em cada miRNA varia de forma diferente entre os organismos comparados (**TABELA 6**). Em *D. melanogaster*, por exemplo, miRNA miR-87-3p apresenta apenas 0,2% de *mismatches*, contrastando com os 9,6% apresentados em *L. longipalpis*.

Os miRNAs cujos percentuais de mismatches são mais similares entre esses dois organismos foram encontrados em miR-281-1-5p e miR-13b-3p. Ainda assim, as sequências observadas variam entre eles. Em *A. aegypti*, miR- 956-5p apresenta o maior percentual de SNPs (9,2%), percentual este que varia apenas 0,2% em *D. melanogaster*.

Tais resultados, portanto, sugerem que os miRNAs identificados através da comparação com organismos próximos não derivam de artefatos ou contaminação cruzada entre as bibliotecas. Provavelmente os *loci* correspondentes a estes miRNAs não estão presentes nas versões montadas dos genomas de *A. aegypti* e *L. longipalpis* utilizados como referência para o mapeamento das bibliotecas. Logo, estes miRNAs podem ser considerados conservados entre os organismos, embora isto ainda deva ser confirmado por sequenciamento do genoma ou análise de *contigs* não incorporados na montagem final.

Tabela 6: Análise da presença de SNPs em miRNAs

miR	Mismatch (%) em Llo	Mismatch (%) em Dme	Dme - sequência após análise de SNPs	Aae - sequência após análise de SNPs
llo-miR-87-3p	9,6	0,2	T[T GAGCAA]AATTT C AGGTGTG	A [TGAGCAA]ATTT T CAGGTGTG
llo-miR-305-3p	9,6	0,2	C[GGCACAT]GTTGAAGTACTCA	C[GGCACAT]GTT G AGTACTCA
llo-miR-980-3p	9,3	0,2	T[AGCTGCC]TTGTGAAGGGCTTA	T[AGCTGCC]TT A CGAAGGGCTTA
llo-miR-318-3p	9,2	0,1	T[CACTGGG]CTTTGTTTATCTCA	T[CACTGGG]CTT A GTTT A GCTCA
llo-miR-282-3p	9,1	0,2	A[CATAGCC]TATAAGAGGTTAGG	A[CATAGCC]T G AAGAGGTTAGG
llo-miR-1-5p	9,1	0,1	C[CATGCTT]CCTTGCATTCAATA	C[CATGCTT]CCTTGCATT C GGTA
llo-miR-306-5p	5,7	0,2	T[CAGGTAC]TTAGTGACTCTCAA	T[CAGGTAC]T G AGTGACTCTCAA(G)
llo-miR-989-3p	5,3	0,2	T[GTGATGT]GACGTAGTGGAA	T[GTGATGT]GACGTAGTGG T A
llo-miR-282-5p	5	0,6	T[AGCCTCT]ACTAGGCTTTGTCT(A)GT	T[AGCCTCT] C CTAGGCTTTGTCTGT
llo-miR-281-1-5p	5	4,3	A[AAGAGAG]CTG(A)TCCGTCGACAGT	A[AAGAGAG]CT A TCCGTCGACAGT
llo-miR-283-5p	4,9	0,1	T[GACTAGA]TTTCATGCTCGTCT	T[GACTAGA]TT A CATGCTCGTCT
llo-miR-2a-1-5p	3,2	0,3	C[TCTCAA]GTGGCTGTGAAAC	C[TCTCAA]GTGGCTGTGAA A T
llo-miR-281-3p	2,9	0,2	T[GTCATGG]AATTGCTCTCTTTGT	T[GTCATGG]AATTGCTCTCTTT A T
llo-miR-2a-1-3p	2,4	3,8	T[ATCACAG]CCAGCTTTG A (G)GAGC	T[ATCACAG]CCAGCTTTG A (T)GA
llo-miR-92a-3p	2,1	0,2	T[ATTGCAC]TTGTCCCGGCCTAT	T[ATTGCAC]TTGTCCCGGCCT G (A)T
llo-miR-2b-3p	1,1	1	T[ATCACAG]CCAGCTTTG A (T)GAGC	T[ATCACAG]CCATTTT G (C (T))GAGCT
llo-miR-13b-3p	0,7	0,4	T[ATCACAG]CCATTTT G ACGAGTT	T[ATCACA(C) G]CCATTTT G ACGAGTT
llo-miR-263a-5p	0,6	0,2	A[ATGGCAC]TGGAGAATTCACGGG	A[ATGGCA(C) C]TGGAGAATTCACGGG

Tabela 6: Análise da presença de SNPs em miRNAs (continuação).

miR	Mismatch (%) em Llo	Mismatch (%) em Dme	Dme - sequência após análise de SNPs	Aae - sequência após análise de SNPs
aae-miR-956-5p	9,2	0,2	G[TGTTTGG]AATGGTCTCGTTAGCT	G[CGTTTGA]AATGGTCTCGTTAACT
aae-miR-2a-5p	4	0,3	C[TCTCAA]GTGGTTGTGAAATG	A[CTCTCAA]AGTGGC(T)TGTGAAA
aae-miR-956-3p	3,6	0,1	T[TTGAGA]CCACTCTAATCCATT	T[TTGAGA]CCACTC(G)T(C)AAT(A)CCATT
aae-miR-2b-3p	3	1	T[ATCACAG]CCAGCTTTGAG(T)GAG C	T[ATCACAG]CCAGCTTTGAA(T)GAGCA
aae-miR-304-5p	1,4	0,1	T[AATCTCA]ATTTGTAAATGTGAG	T[AATCTCA]AT(A)TTGTTAAAT(A)GTGAG
aae-miR-2a-3p	1,3	3,8	T[ATCACAG]CCAGCTTTGAAG	T[ATCACAG]CCAGCTTTGAA(T)G
aae-miR-92b-3p	1	0,3	A[AT(G A)TGCAC]TAGTCCCGG CCT(A G)GC	A(T)[ATTGCAC]TTGTCCCGGCCTGC(T)
aae-miR-92a-3p	0,9	0,2	T[ATTGCAC]TTGTCCCGGCCTAT	T(A)[ATTGCAC]TTGTCCCGGCCTAT
aae-miR-281-3p	0,6	0,2	T[GTCATGG]AATTGCTCTCTTA	T[GTCATGG]AATTGCTCTCTTTA(T)
aae-miR-276-5p	0,6	0,2	A[GCGAGGT]ATAGAGTTCCTAC	A[GCGAGGT]ATAGAGTTCCTAC(T)

miR: miRNA; **Llo:** *L. longipalpis*; **Dme:** *D. melanogaster*; **Aae:** *A. aegypti*; **SNPs:** Polimorfismo único de sequências. As diferenças encontradas foram ressaltadas em vermelho e entre parênteses estão as alternativas aos nucleotídeos imediatamente à esquerda da posição em que se encontram. O símbolo “ | ” indica que ambos os nucleotídeos apareceram em frequências similares sendo, portanto, considerados igualmente possíveis naquela posição. Os colchetes “ [] ” indicam a seed, compreendendo dos nucleotídeos 2 a 8.

A comparação entre o número de precursores e de miRNAs anotados frente aos dados atualmente disponíveis no miRBase para *A. aegypti* e *D. melanogaster* é apresentada (TABELA 7), sendo possível concluir que a metodologia desenvolvida é robusta e capaz de superar problemas decorrentes do estado de montagem e anotação do genoma do organismo estudado.

Todavia, um aspecto que restringe o número de miRNAs anotados para *L. longipalpis* e *A. aegypti* refere-se ao número de espécies de insetos estudadas. Em outras palavras, a aplicação do *pipeline* a outros organismos da ordem Diptera pode permitir anotação de novas sequências ausentes até o momento, ressaltando similaridades ou diferenças entre esses organismos.

Tabela 7: Resumo de precursores e miRNAs maduros anotados em cada organismo.

Sequências encontradas nos genomas disponíveis			
	Llo	Dme	Aae
Precursor	94	256	130
miRNA maduro	163	466	198
Precursos disponíveis no miRBase (v.21)	0	256	101
miRNAs maduros disponíveis no miRBase	0	466	124
Sequências únicas encontradas através de comparação com organismos próximos			
	Llo	Dme	Aae
Precursor	30	0	9
miRNA maduro	43	0	10
Mapeado no precursor de <i>D. melanogaster</i>	27	0	9
Mapeado no precursor de <i>A. aegypti</i>	3	0	0
Mapeado no precursor de <i>L. longipalpis</i>	0	0	0
Sequências Totais (incluindo cópias)			
	Llo	Dme	Aae
Precursor	124	256	149
miRNA maduro	206	466	208

Llo: *L. longipalpis*; **Dme:** *D. melanogaster*; **Aae:** *A. aegypti*.

5.3 Conservação de miRNAs e precursores entre os insetos estudados

Foram encontrados 75 precursores compartilhados (TABELA 8) e 70 precursores de cópia única conservados entre os três insetos estudados (FIGURA 9) o que corresponde a 61,4% e 82,3% de todos os precursores únicos de *A. aegypti* e *L. longipalpis*, respectivamente.

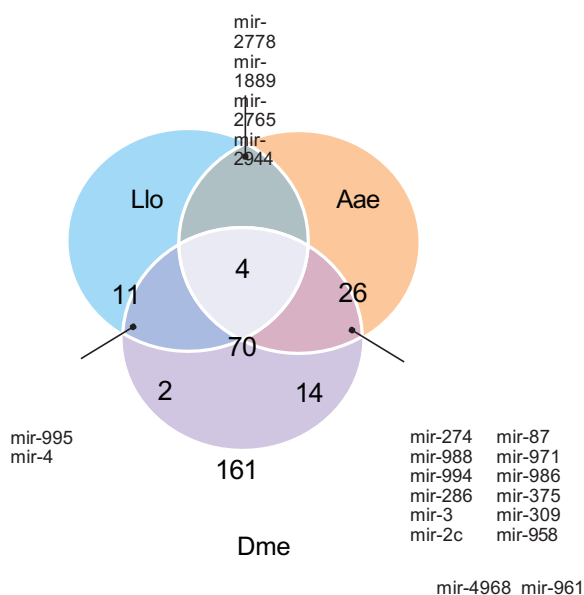


Figura 9: Precursosres únicos conservados entre os diferentes insetos analisados. Dos mirnas anotados e/ou disponíveis no banco de dados miRBase (v.21), 11 precursores únicos foram encontrados apenas em Llo. Destes, a grande maioria (8 mirnas) foram preditos e curados neste trabalho e não apresentam ortologia com outros organismos. Em Aae foram encontrados 26 precursores únicos, dos quais 15 são espécie-específicos. Em Dme, o número de precursores únicos encontrados foi de 157. Quanto aos mirnas compartilhados entre as três espécies, foram encontrados 70 precursores únicos. **Llo:** *L. longipalpis*; **Aae:** *A. aegypti*; **Dme:** *D. melanogaster*.

Tabela 8: pre-miRNAs conservados entre três organismos da ordem Diptera.

pre-miRNAs conservados entre <i>L. longipalpis</i> , <i>D. melanogaster</i> e <i>A. aegypti</i>				
bantam	mir-190	mir-2a	mir-4	mir-970
let-7	mir-193	mir-2b	mir-4968	mir-980
mir-1	mir-210	mir-304	mir-5	mir-981
mir-10	mir-219	mir-305	mir-7	mir-989
mir-100	mir-252	mir-306	mir-79	mir-993
mir-1000	mir-263	mir-307a	mir-8	mir-996
mir-11	mir-275	mir-308	mir-87	mir-998
mir-12	mir-276	mir-314	mir-927	mir-999
mir-124	mir-277	mir-315	mir-929	mir-9a
mir-125	mir-278	mir-316	mir-92a	mir-9b
mir-133	mir-279	mir-317	mir-92b	mir-9c
mir-137	mir-281	mir-318	mir-932	mir-iab-4
mir-13b	mir-282	mir-31a	mir-956	mir-iab-8
mir-14	mir-283	mir-33	mir-957	
mir-184	mir-285	mir-34	mir-965	

Entre os precursores compartilhados por *A. aegypti* e *L. longipalpis*, são encontrados mirnas de *Bombyx mori* – bicho da seda – (mir-2778 e mir-2765); *Anopheles gambiae* (mir-1889); *Culex quinquefasciatus* (mir-1889); *Apis mellifera* – abelha – (mir-2765 e mir-2944); *Tribolium castaneum* (mir-2944); *Heliconius melpomene*, *Acyrtosiphon pisum* e *Manduca sexta* (mir-2765).

Dos miRNAs derivados de precursores compartilhados apenas entre *D. melanogaster* e *L. longipalpis*, mir-995 parece estar presente no ancestral comum dos três insetos avaliados no presente trabalho, como proposto por Marco (2014). Deste modo, sua ausência em mosquitos sugere que estes o perderam. O precursor mir-4, por sua vez, faz parte do grupo ou *cluster* mir-309~6, o qual codifica um transcrito policistrônico que gera 8 miRNAs conservados e que foi inicialmente proposto como específico do gênero *Drosophila*. No entanto, esse miRNA é um homólogo de mir-9 e foi reportado em *A. aegypti* (Ninova *et al.*, 2014).

A existência de miRNAs compartilhados apenas entre *D. melanogaster* e *A. aegypti*, também pode indicar a perda destas sequências em *L. longipalpis*, bem como limitações técnicas para sua descoberta.

Quanto aos precursores não compartilhados entre os três insetos, foram encontradas 8 novas sequências em *L. longipalpis*, as quais não apresentam similaridade com outras espécies. Estas sequências foram nomeadas como llo-mir-new1 a llo-mir-new8 e apenas uma delas gera miRNAs maduros de ambos os braços (llo-mir-new2). Em *A. aegypti*, dos 26 precursores únicos não- conservados, 15 não possuem ortólogos descritos. Grande parte desses miRNAs já foi reportada em estudos anteriores (Akbari *et al.*, 2013; Hu *et al.*, 2015). Neste trabalho, sua nomenclatura foi sistematizada de acordo com as regras do miRBase apenas nesse trabalho.

Quanto às sequências maduras, foram encontrados 109 miRNAs compartilhados pelos três organismos que corresponde a aproximadamente 50% de todos os miRNAs de *A. aegypti* e de *L. longipalpis* e cerca de 23% em *D. melanogaster*.

A **TABELA 9** apresenta os miRNAs maduros que são compartilhados entre as espécies estudadas. Nela figura um dos primeiros miRNAs descritos (let-7) – cuja descoberta se deu no verme *C. elegans* (Pasquinelli *et al.*, 2000). Let-7 apresenta função ligada ao controle temporal da diferenciação celular durante a mudança do estágio larval para o estágio adulto do desenvolvimento desse verme e sua conservação é descrita em eucariotos (Pasquinelli *et al.*, 2000). Além de let-7, são observados miRNAs envolvidos em processos de: regulação hormonal (miR-8) (Lee *et al.*, 2015); diferenciação de tecidos mesodérmicos e desenvolvimento muscular durante desenvolvimento larval (miR-1) (Sokol e Ambros, 2005); crescimento e controle de proliferação celular e do gene pro- apoptótico *hid* (bantam) (Brennecke *et al.*, 2003; Parrish *et al.*, 2009), entre outros. Estas funções são descritas na literatura a partir de diversos estudos realizados em *D. melanogaster*.

Tabela 9: miRNAs conservados entre os insetos analisados.

Precursor (mir)	miRNA maduro	Precursor (mir)	miRNA maduro
bantam	bantam-3p bantam-5p	mir-304	miR-304-5p
let-7	let-7-3p let-7-5p	mir-305	miR-305-3p miR-305-5p
mir-1	miR-1-3p miR-1-5p	mir-306	miR-306-5p
mir-10	miR-10-3p miR-10-5p	mir-307a	miR-307a-3p miR-307a-5p
mir-100	miR-100-3p miR-100-5p	mir-308	miR-308-5p
mir-1000	miR-1000-5p	mir-315	miR-315-5p
mir-11	miR-11-3p miR-11-5p	mir-316	miR-316-5p
mir-12	miR-12-5p	mir-317	miR-317-3p
mir-124	miR-124-3p	mir-318	miR-318-3p
mir-125	miR-125-3p miR-125-5p	mir-31a	miR-31a-5p
mir-133	miR-133-3p miR-133-5p	mir-33	miR-33-3p miR-33-5p
mir-137	miR-137-3p	mir-34	miR-34-3p miR-34-5p
mir-13b	miR-13b-3p miR-13b-5p	mir-5	miR-5-5p
mir-14	miR-14-3p miR-14-5p	mir-7	miR-7-3p miR-7-5p
mir-184	miR-184-3p miR-184-5p	mir-79	miR-79-3p miR-79-5p
mir-190	miR-190-3p miR-190-5p	mir-8	miR-8-3p miR-8-5p
mir-210	miR-210-3p miR-210-5p	mir-87	miR-87-3p
mir-219	miR-219-5p	mir-927	miR-927-3p miR-927-5p
mir-252	miR-252-3p miR-252-5p	mir-929	miR-929-3p miR-929-5p
		mir-92a	miR-92a-3p miR-92a-5p
		mir-92b	miR-92b-3p
		mir-932	miR-932-5p
		mir-956	miR-956-3p

Tabela 9: miRNAs conservados entre os insetos analisados (continuação).

mir-263a	miR-263a-3p	mir-956	miR-956-5p
	miR-263a-5p	mir-957	miR-957-3p
mir-263b	miR-263b-5p	mir-965	miR-965-3p
mir-275	miR-275-3p	mir-970	miR-970-3p
	miR-275-5p		miR-970-5p
mir-276a	miR-276a-3p	mir-980	miR-980-3p
	miR-276a-5p	mir-981	miR-981-3p
mir-277	miR-277-3p	mir-989	miR-989-3p
	miR-277-5p		
mir-278	miR-278-3p	mir-993	miR-993-3p
mir-279	miR-279-3p	mir-996	miR-996-3p
	miR-279-5p	mir-998	miR-998-3p
mir-281	miR-281-1-5p	mir-999	miR-999-3p
	miR-281-3p		
mir-282	miR-282-3p	mir-9a	miR-9a-3p
	miR-282-5p		miR-9a-5p
mir-283	miR-283-5p	mir-9b	miR-9b-3p
			miR-9b-5p
mir-285	miR-285-3p	mir-9c	miR-9c-3p
	miR-285-5p		miR-9c-5p
mir-2a-1	miR-2a-1-5p	mir-iab-4	miR-iab-4-3p
	miR-2a-3p		miR-iab-4-5p
mir-2b	miR-2b-3p		

Embora tenha sido observada a presença do precursor mir-iab-8 em *A. aegypti* (descrito nesse trabalho) e sua anotação seja compartilhada entre os insetos avaliados, as replicatas utilizadas para mosquitos-palha (**TABELA 3**) não apresentaram expressão de miRNAs maduros oriundos desse precursor e por isso não foram indicados na **TABELA 8**.

Um aspecto interessante sobre as sequências maduras, refere-se às alterações de origem do braço dos seguintes miRNAs: llo-miR-275a-5p^{inv-3p} (o qual apresenta *seed* conservada com dme-miR-275-3p – derivado de *D. melanogaster*) e llo-miR-2778-5p^{inv-3p} (o qual apresenta *seed* conservada com bmo-miR-2778-3p – derivado de *B. mori*). Como indicado, a nomenclatura foi sugerida para evitar-se interpretações errôneas, uma vez que a ortologia de um miRNA é utilizada para propor sua nomenclatura. Tais alterações não foram encontradas entre os miRNAs de *A. aegypti*.

Diferente do que fora reportado em trabalhos anteriores (Marco *et al.*, 2010; Griffiths-Jones *et al.*, 2011), essa mudança não se trata da variação na origem do braço mais processado (5p ou 3p entre diferentes organismos e

conhecida como *switch* ou *shift*), tampouco miRNA espelho. Trata-se sim, de um miRNA de mesma *seed* que deriva de braços diferentes em organismos distintos. Entretanto, é factível sugerir que esses miRNAs possuam alvos diferentes, visto que já foram reportadas contribuições de nucleotídeos fora da *seed* no reconhecimento do mRNA alvo (Revisado por Bartel, 2009).

5.4 Origem genômica

No início dos estudos de miRNAs, eram creditadas às regiões intergênicas a responsabilidade por codificá-los (Kim, 2005). Contudo, trabalhos posteriores em diversos organismos têm demonstrado que regiões codificadoras de proteínas (ou exônicas) podem codificá-los (Slezak-Prochazka *et al.*, 2013). O mesmo se aplica aos íntrons, os quais também são capazes de codificar miRNAs funcionais (Kim e Kim, 2007; Ruby *et al.*, 2007; Ramalingam *et al.*, 2014). Além disso, alguns trabalhos sugerem que os alvos variam segundo a origem genômica do miRNA (Berillo *et al.*, 2013). Nesse caso miRNAs intrônicos tendem a regular os genes do qual derivam (Hinske *et al.*, 2010), por exemplo.

Diante do exposto, torna-se importante saber a origem do miRNA segundo sua posição no genoma. Para isto, os miRNAs foram categorizados em: exônicos, intrônicos ou intergênicos. Observa-se que em *L. longipalpis* a proporção de miRNAs derivados de éxons e de íntrons é bastante similar (**FIGURA 10**).

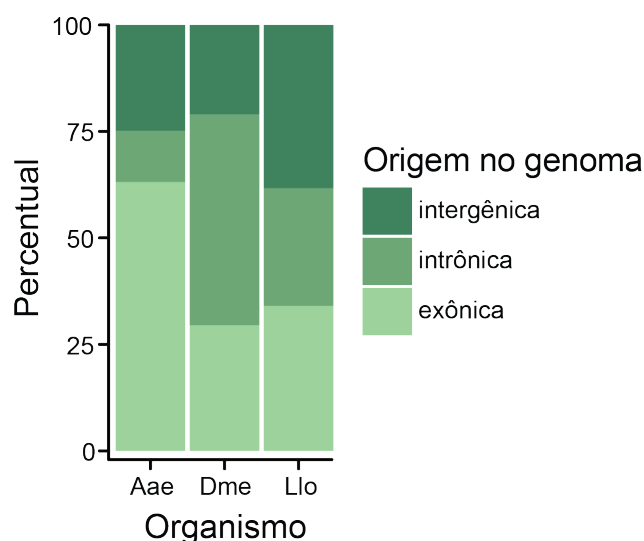


Figura 10: Origem genômica dos miRNAs. A origem genômica dos miRNAs foi dividida em três categorias (exônica, intergênica e intrônica) e sua contribuição percentual foi verificada para cada organismo estudado. **Llo:** *L. longipalpis*; **Aae:** *A. aegypti*; **Dme:** *D. melanogaster*.

O percentual de miRNAs exônicos encontrados nesse organismo é próximo ao encontrado em *D. melanogaster*. Em contraste, em *A. aegypti* esse valor é superior a 50% e corresponde à soma dos percentuais de miRNAs derivados de íntrons e éxons em *L. longipalpis*.

Por outro lado, o número de miRNAs de origem intrônica em *Aedes aegypti* corresponde a aproximadamente 12% de todos os miRNAs encontrados no genoma disponível desse inseto. Além disso, esse percentual aproxima-se do encontrado em *D. melanogaster* para mesma categoria.

Tais análises, entretanto, podem apresentar viés decorrente dos estados de montagem e anotação do genoma de cada organismo, sendo dos genomas analisados o da mosca da fruta mais bem caracterizado. Nesse caso, a caracterização da montagem do genoma pode melhorar os aspectos de sua anotação, afetando, por exemplo, a informação de que uma determinada posição se refere à um éxon ao invés de um íntron. Em suma, os resultados encontrados parecem refletir as diferenças da montagem dos genomas e não apresentam significado biológico.

5.5 Expressão dos miRNAs

5.5.1 Expressão geral

É possível observar a similaridade dos níveis de expressão dos miRNAs presentes no terço superior dos gráficos e maior variação no terço inferior. No terço superior, são encontrados miRNAs muito expressos. O oposto ocorre no terço inferior (**FIGURA 11**).

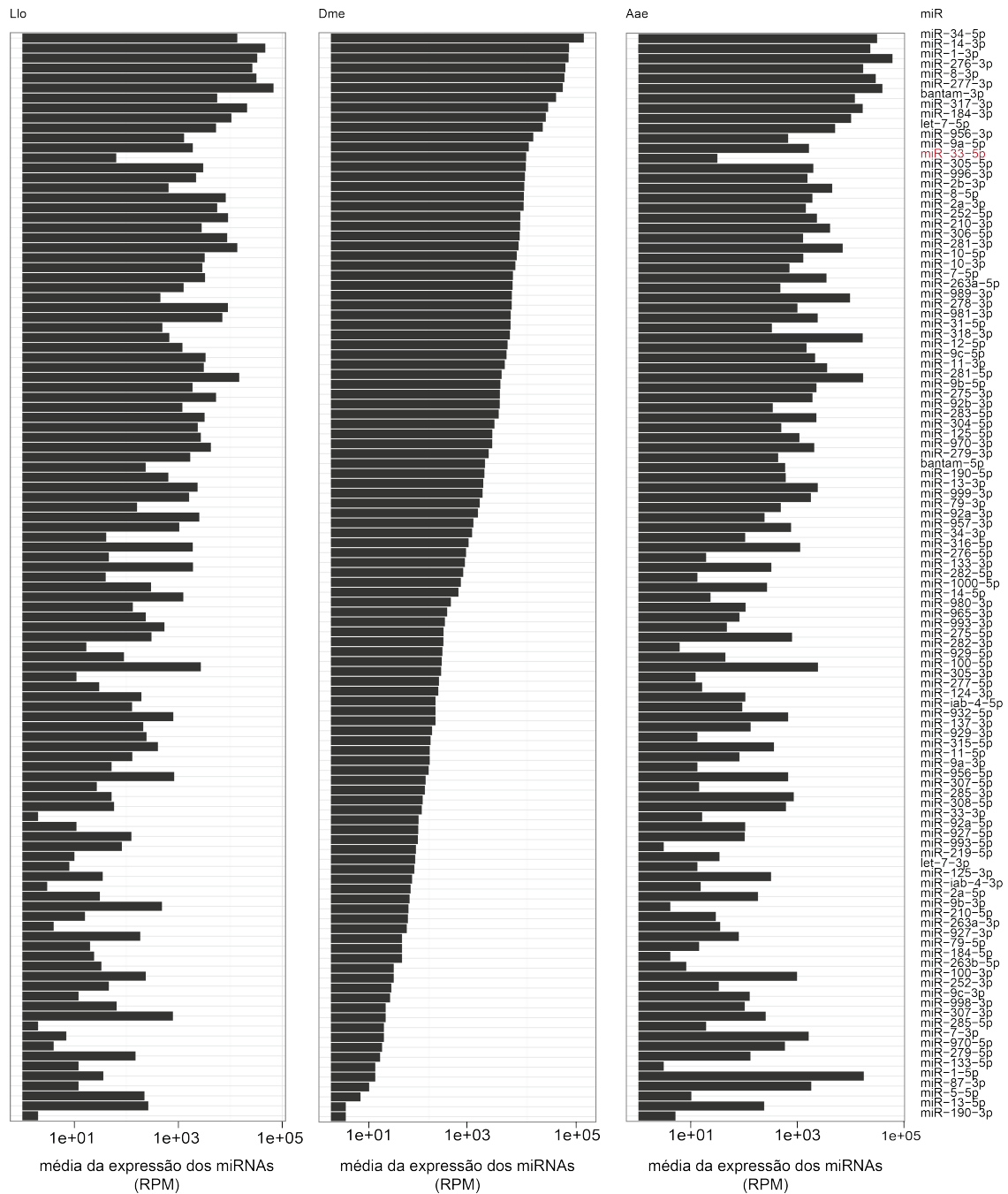


Figura 11: Média da expressão normalizada (RPM) dos miRNAs conservados entre as três espécies estudadas. As médias consistem na expressão do miRNA na triplicata de bibliotecas analisadas para cada organismo. Os miRNAs foram ordenados segundo a expressão em Dme. **Em vermelho:** miRNA discutido no texto. **Lio:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **miR:** miRNA; **RPM:** sequências por milhão.

Uma exceção à similaridade encontrada nos miRNAs mais expressos é vista em miR-33-5p. A mudança de braço de miR-33 já fora reportada em trabalhos anteriores onde *Tribolium castaneum* apresenta dominância do braço 3p, de modo semelhante ao observado em *A. aegypti* e em *L. longipalpis* (Marco *et al.*, 2010). Esse resultado indica que essa mudança ocorreu após o último ancestral comum dos organismos avaliados nesse trabalho, em *D. melanogaster*.

Para verificar a existência de padrões de expressão entre os miRNAs conservados, todos os 109 miRNAs foram avaliados segundo a similaridade de expressão entre os insetos. Assim, miRNAs cuja expressão variou pelo menos $1,5\sigma$ da razão entre dois organismos foram considerados “discrepantes” e os miRNAs com variação inferior a $1,5\sigma$, por sua vez, foram classificados como “similares”. Foram utilizados 11 miRNAs de cada grupo para análises de padrão de expressão entre os insetos e observou-se que a expressão de cada braço de um mesmo precursor independe do outro (**TABELA 10**).

Deste modo, ocorre fina regulação entre eles, podendo existir miRNAs com perfis distintos: miR-1-3p, por exemplo, apresentou maior similaridade entre *L. longipalpis*, *D. melanogaster* e *A. aegypti* que miR-1-5p – o qual, segundo classificação adotada, foi incluído no grupo “discrepante”. O mesmo ocorreu com miR-7, o qual gerou miR-7-5p – de “expressão similar” entre os organismos, e miR-7-3p – cuja expressão foi classificada como “discrepante”. Além disso, a expressão de ambos os grupos (similar e discrepante) apresenta diferenças estatisticamente significativas (**FIGURA 12**), sendo o grupo classificado como “similar”, o que apresenta maior expressão ($p = 0,002006 \sim 95\%$).

Tabela 10: Padrão de expressão de miRNAs dos insetos avaliados segundo similaridade.

miRNA	Expressão (RPM) <i>L. longipalpis</i>	Expressão (RPM) <i>D.</i> <i>melanogaster</i>	Expressão (RPM) <i>A. aegypti</i>	Tipo de expressão
miR-33-5p	64	9501	31	Discrepante
miR-14-5p	1247	397	23	Discrepante
miR-33-3p	2	71	16	Discrepante
miR-993-5p	82	59	3	Discrepante
miR-307a-3p	788	13	255	Discrepante
miR-7-3p	7	12	1655	Discrepante
miR-970-5p	4	11	588	Discrepante
miR-1-5p	36	8	18238	Discrepante
miR-87-3p	12	6	1852	Discrepante
miR-5-5p	223	4	10	Discrepante
miR-13b-5p	264	2	238	Discrepante
miR-34-5p	13617	142885	32587	Similar
miR-14-3p	47059	71362	24422	Similar
miR-1-3p	32925	69713	63749	Similar
let-7-5p	5318	20765	5230	Similar
miR-9a-5p	1896	10752	1674	Similar
miR-305-5p	3010	9349	2024	Similar
miR-996-3p	2193	8976	1573	Similar
miR-2b-3p	648	8724	4609	Similar
miR-8-5p	8152	8565	1947	Similar
miR-7-5p	3262	5117	3596	Similar

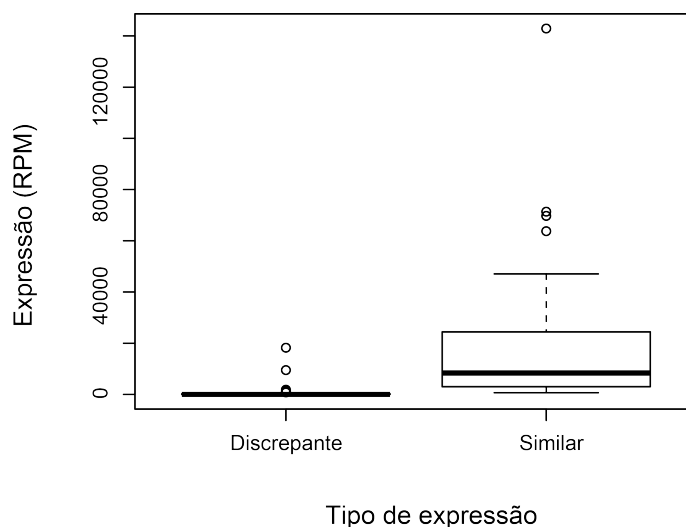


Figura 12: Boxplot da comparação de expressão (RPM) dos miRNAs classificados como de “expressão similar” e “expressão discrepante” entre os insetos estudados. A análise refere-se aos RPMs de expressão de cada miRNA em cada organismo. **Similar:** miRNAs cuja expressão variou menos que $1,5\sigma$ do *ratio* de expressão RPM em dois organismos; **Discrepante:** miRNAs cuja expressão variou pelo menos $1,5\sigma$ do *ratio* de expressão RPM de dois organismos; **RPM:** sequências por milhão.

De forma resumida, parece haver dois padrões de expressão: os miRNAs muito expressos, que tendem a manter níveis semelhantes entre as espécies, e os miRNAs com expressão classificada como discrepante, que tendem a possuir baixa expressão. Outro dado interessante refere-se ao fato de não haver padrão que caracterize as diferenças estatísticas globais dos grupos de miRNAs discrepantes derivados de *D. melanogaster* quando comparados aos mesmos miRNAs das demais espécies. Essas análises permitem sugerir que os miRNAs cuja expressão fora maior (e similar) entre os organismos desempenhem funções importantes nos três insetos. Por outro lado, a inexistência de diferenças estatísticas globais entre os miRNAs de padrão discrepante entre as espécies (isto é, quando se avalia o grupo, invés de um ou outro miRNA isoladamente) sugere que estes estejam sofrendo pressões evolutivas similares.

Na comparação par a par entre os valores de RPM de todos os 109 miRNAs conservados observam-se Coeficientes de Determinação (R^2) acima de 0,6 para todos os pares testados – isto é, os níveis de expressão dos 109 miRNAs conservados entre as espécies são estatisticamente comparáveis (**FIGURA 13**). Além disso o teste de correlação de *Pearson* apresentou significância estatística em todas as análises ($p < 2,2e-16$, 95% de confiança) e é possível identificar maior correlação dos níveis de miRNAs conservados entre *L. longipalpis* e *A. aegypti* ($R^2 = 0,802$), seguidos da correlação entre *A. aegypti* e *D. melanogaster* ($R^2 = 0,790$) e entre *L. longipalpis* e *D. melanogaster* ($R^2 = 0,692$). Esses resultados são condizentes com a filogenia entre os insetos estudados (**FIGURA 7**).

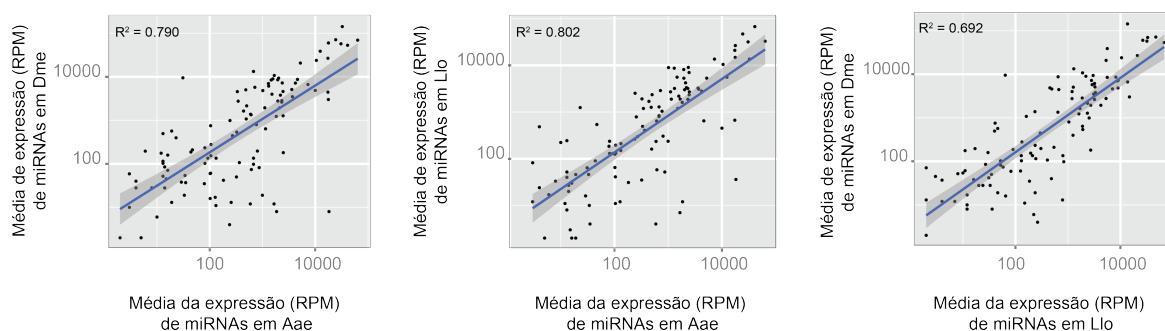


Figura 13: Comparação par a par dos níveis de expressão (RPM) de miRNAs compartilhados entre os três insetos. Foram utilizados todos os miRNAs compartilhados entre as espécies estudadas. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **RPM:** sequências por milhão.

Foram avaliados ainda, os 20 miRNAs mais expressos em cada inseto analisado (**TABELA 11**). Os resultados encontrados estão de acordo com dados da literatura, nos quais: bantam; miR-184; miR-8; miR-34 e let-7 são alguns dos miRNAs mais expressos em culturas de células e em tecidos (embrionário, ovário, testículos e cabeça) de *D. melanogaster* (Wen *et al.*, 2014).

Tabela 11: 20 miRNAs mais expressos em cada organismo.

<i>Lutzomyia longipalpis</i>			
Ordem	miR	RPM Total	% do Total
1	llo-miR-277-3p	67641,35	16,24
2	llo-miR-14-3p	47059,48	11,30
3	llo-miR-1-3p	32925,33	7,91
4	llo-miR-8-3p	31837,02	7,65
5	llo-miR-276a-3p	26573,75	6,38
6	llo-miR-317-3p	20976,05	5,04
7	llo-miR-281-1-5p	14802,76	3,55
8	llo-miR-281-3p	13624,59	3,27
9	llo-miR-34-5p	13616,84	3,27
10	llo-miR-184-3p	10474,13	2,52
11	llo-miR-252-5p	9037,57	2,17
12	llo-miR-278-3p	8965,75	2,15
13	llo-miR-306-5p	8704,21	2,09
14	llo-miR-8-5p	8151,78	1,96
15	llo-miR-981-3p	7033,37	1,69
16	llo-bantam-3p	5593,67	1,34
17	llo-miR-2a-1-3p	5589,41	1,34
18	llo-let-7-5p	5318,42	1,28
19	llo-miR-275a-5p	5311,10	1,28
20	llo-miR-970-3p	4226,23	1,01
<i>Drosophila melanogaster</i>			
Ordem	miR	RPM Total	% do Total
1	dme-miR-34-5p	142885,48	18,68
2	dme-miR-14-3p	71362,07	9,33
3	dme-miR-1-3p	69713,15	9,11
4	dme-miR-276a-3p	60131,82	7,86
5	dme-miR-8-3p	57469,13	7,51
6	dme-miR-277-3p	53221,31	6,96
7	dme-bantam-3p	38795,42	5,07
8	dme-miR-317-3p	26738,55	3,50

Tabela 11: 20 miRNAs mais expressos em cada organismo (continuação).

9	dme-miR-184-3p	23914,68	3,13
10	dme-let-7-5p	20764,86	2,71
11	dme-miR-956-3p	13308,91	1,74
12	dme-miR-9a-5p	10751,68	1,41
13	dme-miR-33-5p	9501,22	1,24
14	dme-miR-305-5p	9348,85	1,22
15	dme-miR-996-3p	8976,06	1,17
16	dme-miR-2b-3p	8723,59	1,14
17	dme-miR-8-5p	8565,46	1,12
18	dme-miR-2a-3p	8476,52	1,11
19	dme-miR-252-5p	7241,38	0,95
20	dme-miR-210-3p	7191,17	0,94
<i>Aedes aegypti</i>			
	miR	RPM Total	%
1	aae-miR-1-3p	63748,63	15,94
2	aae-miR-277-3p	41031,54	10,26
3	aae-miR-34-5p	32587,41	8,15
4	aae-miR-8-3p	30742,43	7,69
5	aae-miR-14-3p	24421,66	6,11
6	aae-miR-1-5p	18237,67	4,56
7	aae-miR-276-3p	17787,27	4,45
8	aae-miR-281-5p	17721,45	4,43
9	aae-miR-317-3p	17393,98	4,35
10	aae-miR-318-3p	17393,98	4,35
11	aae-bantam-3p	12410,02	3,10
16	aae-miR-184-3p	10521,15	2,63
17	aae-miR-989-3p	10016,32	2,50
18	aae-miR-281-3p	7271,58	1,82
19	aae-let-7-5p	5230,40	1,31
20	aae-miR-2b-3p	4608,66	1,15

Dos 20 miRNAs mais expressos em cada organismo, 10 destes são conservados e compartilham também o *ranking* dos mais expressos em cada organismo (**FIGURA 14**). Além disso, todos foram classificados como “similarmente expressos” nas análises anteriores. Dos 10 miRNAs, 8 derivam da região 3’ do precursor e apenas os miRNAs let-7-5p e miR- 277-5p são oriundos da extremidade oposta.

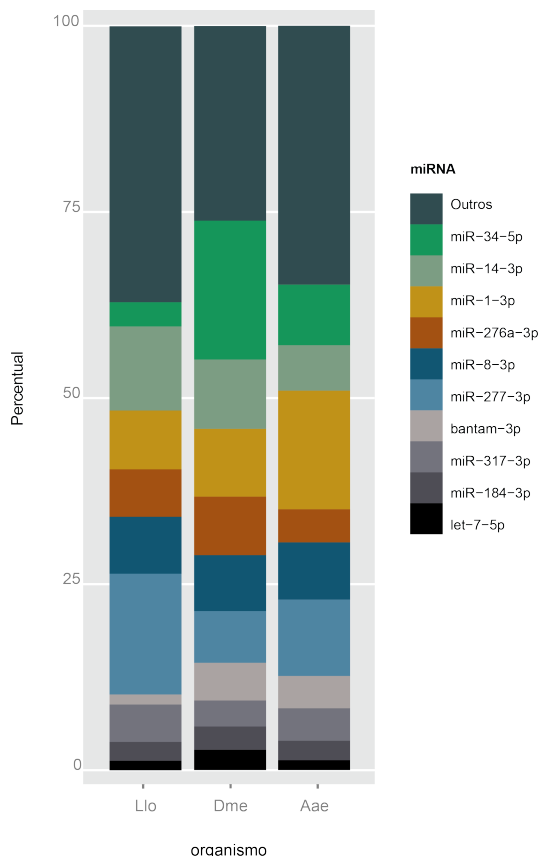


Figura 14: Comparação entre os percentuais de 10 dos miRNAs mais expressos em cada organismo. Dos 20 miRNAs mais expressos em cada inseto, 10 são compartilhados entre os três organismos. Destes, 8 derivam da região 3' de seu respectivo precursor. MiR-8-3p e miR-184-5p contribuem com percentual quase idêntico para o total de miRNAs em cada inseto. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*.

Outro dado interessante refere-se à expressão de miR-8-3p e miR-184-3p, os quais contribuem com percentuais muito similares em cada espécie – sendo miR-8-3p o mais expresso entre esses dois miRNAs. É importante mencionar que a identificação de miR-8-3p se deu por comparação com espécies próximas, demonstrando o impacto da qualidade dos genomas de referências em diversos estudos e reforçando a robustez do *pipeline* desenvolvido no presente trabalho.

Além dos miRNAs conservados entre os três diferentes insetos, o perfil de expressão dos miRNAs não-conservados também foi verificado (**FIGURA 15**).

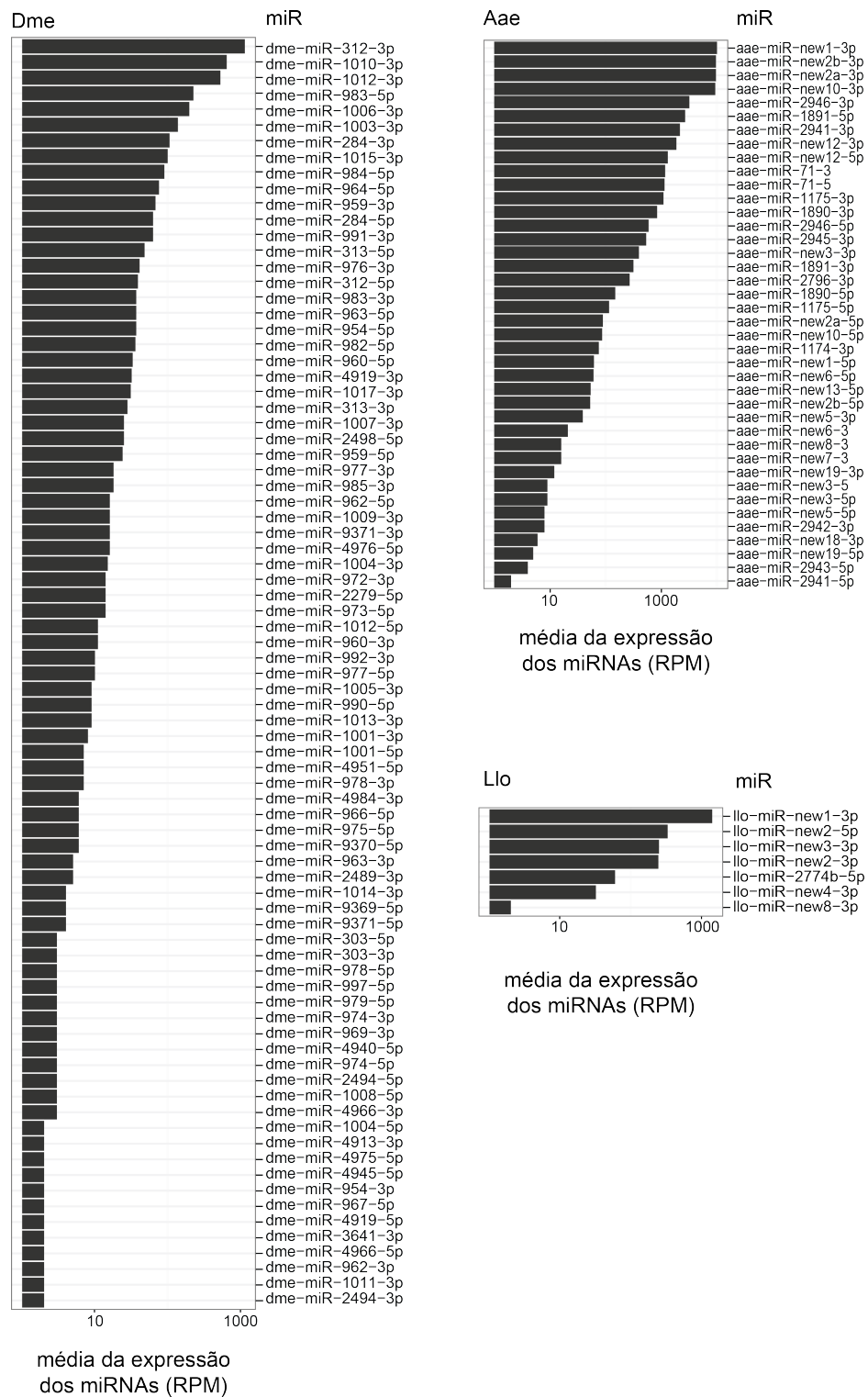


Figura 15: Média da expressão normalizada (RPM) dos miRNAs não-conservados entre as três espécies estudadas. As médias consistem na expressão do miRNA na triplicata de bibliotecas analisadas para cada organismo. Cada espécie contém números diferentes de miRNAs não-conservados com os demais insetos estudados e alguns desses miRNAs foram anotados no presente trabalho. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **miR:** miRNA; **RPM:** sequências por milhão.

Como pode-se perceber, o número de miRNAs não-conservados varia entre as espécies. Além disso, os miRNAs não-conservados derivados de *D. melanogaster* tendem a apresentar expressão menor que o observado nos miRNAs conservados desta espécie (Kruskal-Wallis teste: $p < 2,2e-16$). Para os demais organismos, não há diferenças estatísticas significativas entre ambos os grupos (em *L. longipalpis* – Kruskal-Wallis teste: $p = 0,2063$; em *A. aegypti* – $p = 0,01355$).

5.5.2 Expressão dos miRNAs segundo origem no precursor

A especificidade dos alvos é dada pela escolha da fita de RNA a ser inserida no complexo RISC, o que constitui etapa fundamental do processamento do pre-miRNA. Essa escolha é feita pela enzima Ago1 através das características termodinâmicas do *duplex* nela carregado, onde a extremidade 5' do *duplex* que apresenta menor estabilidade é escolhida como miRNA maduro, uma vez que esta região será acomodada no domínio MID da proteína supracitada (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003; Okamura *et al.*, 2009).

Além dessas características, o não pareamento dos quatro primeiros nucleotídeos da extremidade 5' (Schwarz *et al.*, 2003) e *mismatches* entre o *duplex* na região da *seed* e na região 3'-mid (nucleotídeos 12-15) (Kawamata *et al.*, 2009) são características de escolha da fita que irá compor o complexo RISC maduro (miRISC) e a alteração de um único nucleotídeo é suficiente para a mudança no carregamento do miRNA maduro, como apresentado por Schwarz e colaboradores (2003).

É importante salientar que a proteína Ago2 também pode ser carregada com uma das fitas do *duplex* de miRNA, mediando a clivagem de alvos complementares a ela (Doerks *et al.*, 2002; Forstemann *et al.*, 2007; Czech *et al.*, 2009; Ghildiyal *et al.*, 2010). Todavia, os padrões das sequências envolvidos nesse processo são diferentes dos envolvidos no carregamento pela Ago1.

Outro ponto que merece destaque refere-se à preferência da expressão de um ou outro braço do pre-miRNA em diferentes espécies, o que reflete processos evolutivos distintos sofridos pelas espécies (Marco *et al.*, 2010; Griffiths-Jones *et al.*, 2011). Diante dessas informações, estudar a conservação dos miRNAs

segundo a origem do braço do precursor pode fornecer pistas evolutivas significativas sobre os processos de seleção envolvidos.

Para isso, a expressão em escala logarítmica do RPM de cada braço dos grampos que apresentaram anotação conservada entre *A. aegypti*, *D. melanogaster* e *L. longipalpis* foi comparada (**FIGURA 16**). Valores positivos indicam maior expressão do braço 5p. Valores negativos indicam, por sua vez, indicam maior expressão do braço 3p. Valores acima e abaixo de $1e+05$ e $1e-04$ indicam expressão apenas do braço 5p e 3p, respectivamente.

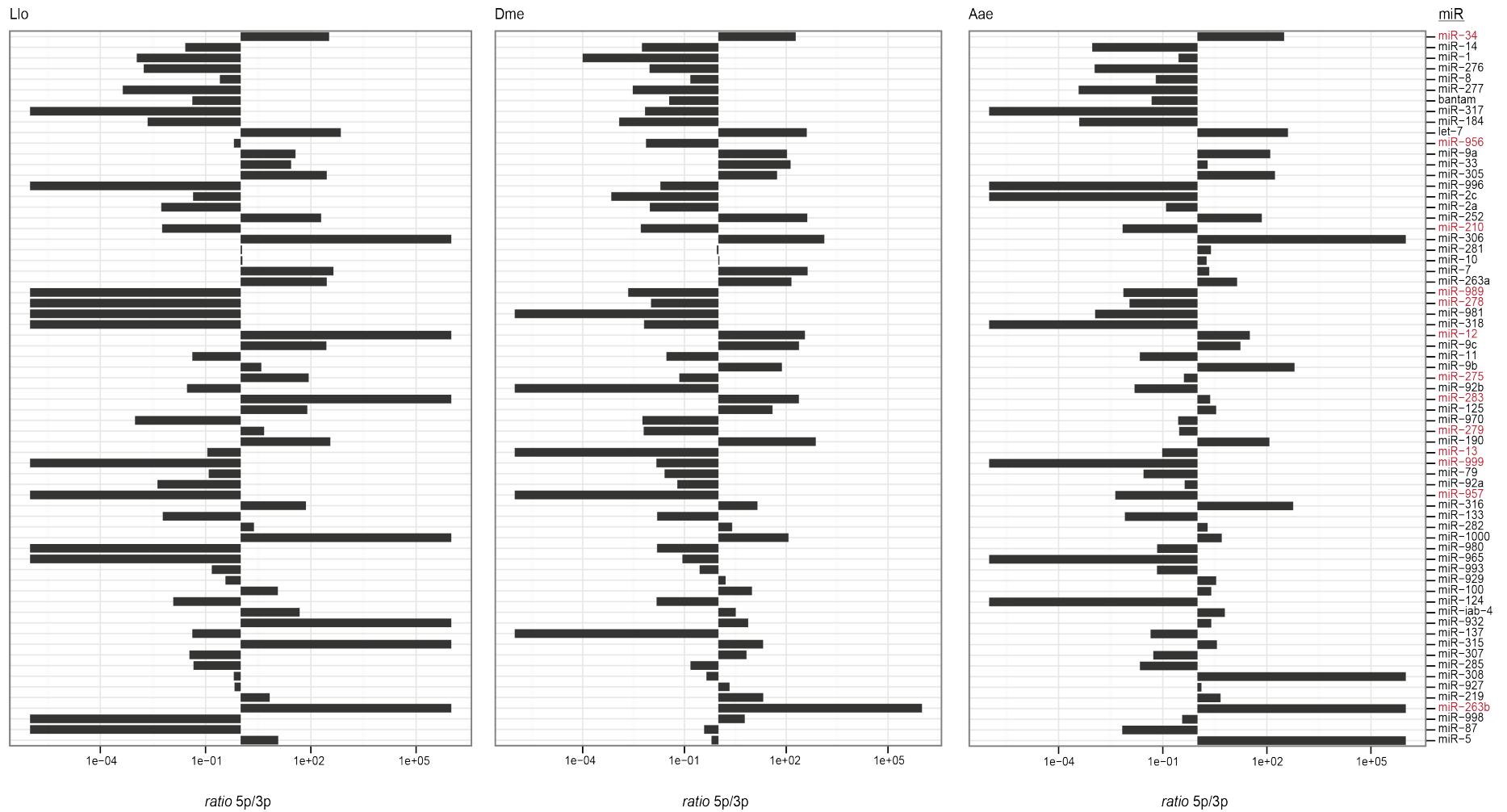


Figura 16: Relação entre a expressão dos braços 5p e 3p (*ratio*) dos miRNAs conservados entre os diferentes insetos. Llo: *Lutzomyia longipalpis*; Dme: *Drosophila melanogaster*; Aae: *Aedes aegypti*; miR: miRNA.; Em vermelho: miRNAs discutidos no texto.

Os perfis encontrados são bastante similares entre os organismos e parece haver preferência de expressão do braço 3p em todos eles. Além disso, observa-se distribuição mais uniforme (e próxima do eixo central) na mosca da fruta que nos demais organismos. Em outras palavras, *D. melanogaster* parece gerar miRNAs de ambos os braços de um mesmo precursor em proporções mais próximas que *L. longipalpis* e *A. aegypti*.

A partir da **FIGURA 16** podem ser destacados alguns pontos (miRNAs marcados em vermelho):

- a) Os *ratios* de miR-34 – o qual é herdado matematicamente (Soni *et al.*, 2013); de let-7 de e bantam são bem similares entre os insetos e corrobora dados de expressão discutidos no item 5.4.1;
- b) O precursor mir-956 gera miRNAs em ambos os braços com expressão quase idêntica em *A. aegypti*;
- c) Diferentemente dos demais organismos, mir-989 e mir-278 geram apenas o braço 3p em *L. longipalpis*. Trabalhos anteriores reportam processos de regulação celular e migração como alvos de miR-989 (Kugler *et al.*, 2013) e controle da homeostase energética como principal processo controlado por miR-278 (Doerks *et al.*, 2002);
- d) Apenas o braço 5p é gerado pelos precursores mir-12 e mir-283 em *L. longipalpis*;
- e) Ocorre mudança de braço para os mir-275 e mir-279 em *L. longipalpis*. Em *A. aegypti* e em *D. melanogaster*, diferentemente do organismo anterior, 3p constitui o braço mais expresso e dessa forma, espera-se que os alvos desses miRNAs não sejam conservados (Marco, Macpherson, *et al.*, 2012);
- f) O precursor mir-13 gera somente o braço 3p em *D. melanogaster*;
- g) Apenas *D. melanogaster* apresenta o braço 5p de mir-999;
- h) Somente *A. aegypti* apresenta o braço 5p de mir-957;
- i) Em todos os organismos, mir-263b gera apenas o braço 5p.

Além disso, foram feitos testes de correlação de *Pearson*, tendo sido encontrados os índices de: 0,325 (entre *A. aegypti* e *L. longipalpis*, $p < 0,007141$, 95% de confiança); 0,361 (entre *L. longipalpis* e *D. melanogaster*, $p < 0,002688$, 95% de confiança); 0,489 (entre *A. aegypti* e *D. melanogaster*, $p < 2,675e-05$, 95% de

confiança). Todavia, as correlações encontradas parecem indicar que este não é um bom modelo para esse tipo de análise, uma vez que o perfil apresentado (maior correlação entre *A. aegypti* e *D. melanogaster*, seguido da correlação entre *D. melanogaster* e *L. longipalpis*) não reflete a topologia da árvore filogenética das espécies estudadas (**FIGURA 7**) e difere do encontrado nas análises de correlação dos perfis de tamanho dos miRNAs conservados. De forma geral, observa-se que a expressão dos miRNAs é uma característica mais conservada que a razão entre os braços dos miRNAs derivados de um mesmo precursor.

5.6 Padrões Moleculares: perfil de tamanho e preferência de base

5.6.1 Preferência de Base

Estudos anteriores demonstraram a existência de preferência do nucleotídeo uracila (U) na extremidade 5' de miRNAs maduros (Hu *et al.*, 2009; Okamura *et al.*, 2009; Ghildiyal *et al.*, 2010; Seitz *et al.*, 2011; Wang, 2013). Isso pode ser explicado pela interação dessa base com o domínio MID da proteína Ago1, e pela relativa facilidade na quebra de duas pontes de hidrogênio formadas entre adenina e uracila (A-U) quando comparado às três pontes formadas por guanina e citosina (G-C). Tais aspectos acabam por favorecer a escolha da fita de RNA que apresenta essa característica como miRNA maduro.

Assim, foi realizada a verificação do nucleotídeo inicial de todas as sequências que mapearam nos miRNAs anotados (**FIGURA 17**).

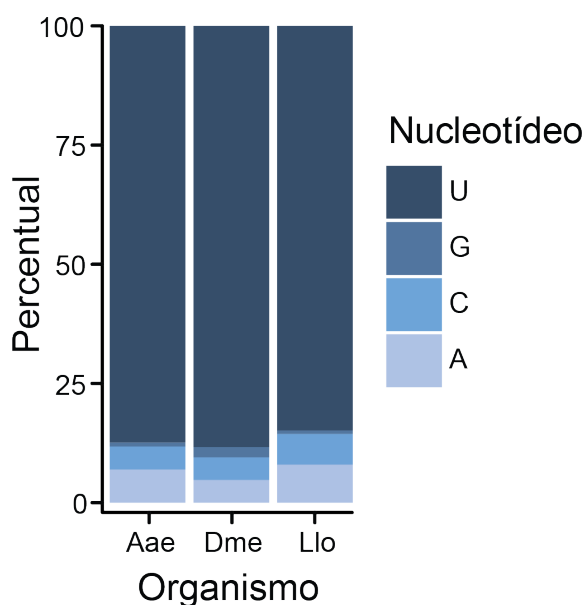


Figura 17: Enriquecimento de base na região 5' de todas as sequências mapeadas em miRNAs anotados para cada espécie. A análise mostra um enriquecimento de U na região 5' de Llo, Dme e Aae, seguido de A e C. G, por sua vez teve a menor representatividade na posição inicial em todos os organismos que compõem o estudo. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **U:** uracila; **G:** guanina; **C:** citosina; **A:** adenosina.

Os resultados encontrados corroboram dados da literatura, uma vez que é observado enriquecimento majoritário de U na posição 5' nos miRNAs anotados para cada inseto. Além disso, como descrito em trabalhos prévios, o nucleotídeo adenina (A) é a segundo mais encontrado nesta posição nos miRNAs avaliados. Outros aspectos importantes referem-se ao alto grau de conservação dos percentuais de representatividade de cada base entre os insetos estudados e ao fato de a preferência por U na primeira posição da região 5' (5',1) ser conservada ao longo da evolução em diversos organismos como apresentado no trabalho de Wang (2013).

A preferência de U como primeiro nucleotídeo é observada também quando são comparados miRNAs conservados e não conservados entre de cada espécie (**FIGURA 18**). No entanto, em *L. longipalpis*, observam-se que aproximadamente 8% dos miRNAs não-conservados apresentam o nucleotídeo guanina (G) na primeira posição, contrastando com menos de 1% do observado em miRNAs conservados (Fisher: $p < 0,01$).

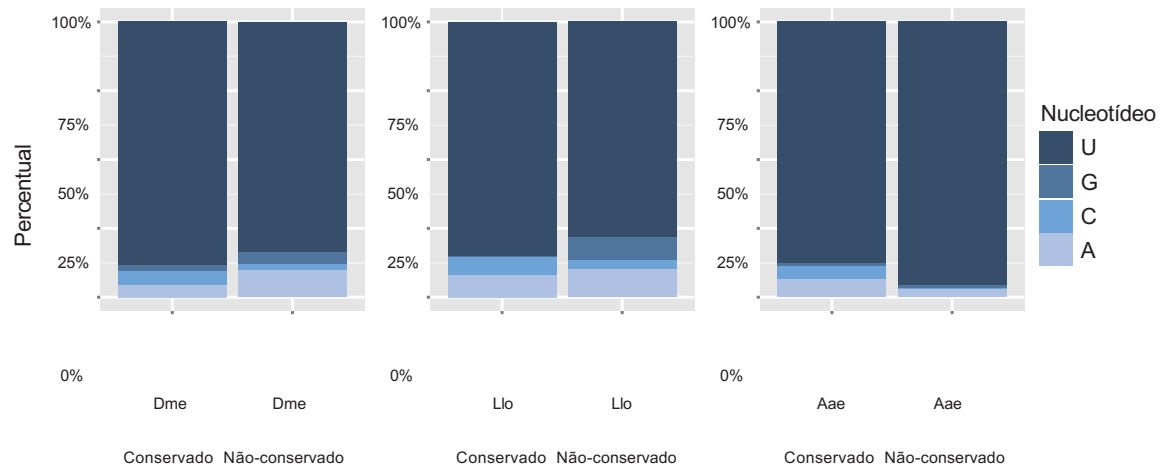


Figura 18: Enriquecimento de base na região 5' dos miRNAs anotados para cada espécie. A análise mostra um enriquecimento de U na região 5' de Llo, Dme e Aae. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **U:** uracila; **G:** guanina; **C:** citosina; **A:** adenosina.

De forma geral, a preferência por U na primeira base sugere que a Ago1 manteve a seleção da extremidade 5' do *duplex* que apresenta a maior instabilidade como miRNA maduro ao longo da evolução.

5.6.2 Perfil de tamanho das sequências mapeadas nos miRNAs

O perfil de tamanho das sequências mapeadas nos miRNAs maduros (**FIGURA 19**) mostra que 63 das 109 sequências possuem tamanho bastante similar (em torno de 22 nt) entre os organismos estudados. Esse número corresponde a aproximadamente 58% de todos os miRNAs avaliados. Como exemplo, tem-se os o perfil encontrado em: miR-34-5p, miR-14-3p, miR-133-3p, miR-306-5p, miR-1-5p e miR-999-3p.

Entretanto, foram encontradas sequências mapeadas em miRNAs conservados cujo perfil de tamanho entre as espécies varia consideravelmente. Nesse grupo figuram, por exemplo, miR-210-3p, miR-989-3p, miR-31a-5p, miR-318-3p, miR-12-5p, miR-279-3p, bantam-5p, miR-275-5p e miR-79-5p.

A variação desse perfil não parece estar relacionada à expressão do miRNA, uma vez que dos 46 miRNAs que apresentaram mapeamento diferente entre pelo menos duas espécies, apenas 5 variaram também a expressão.

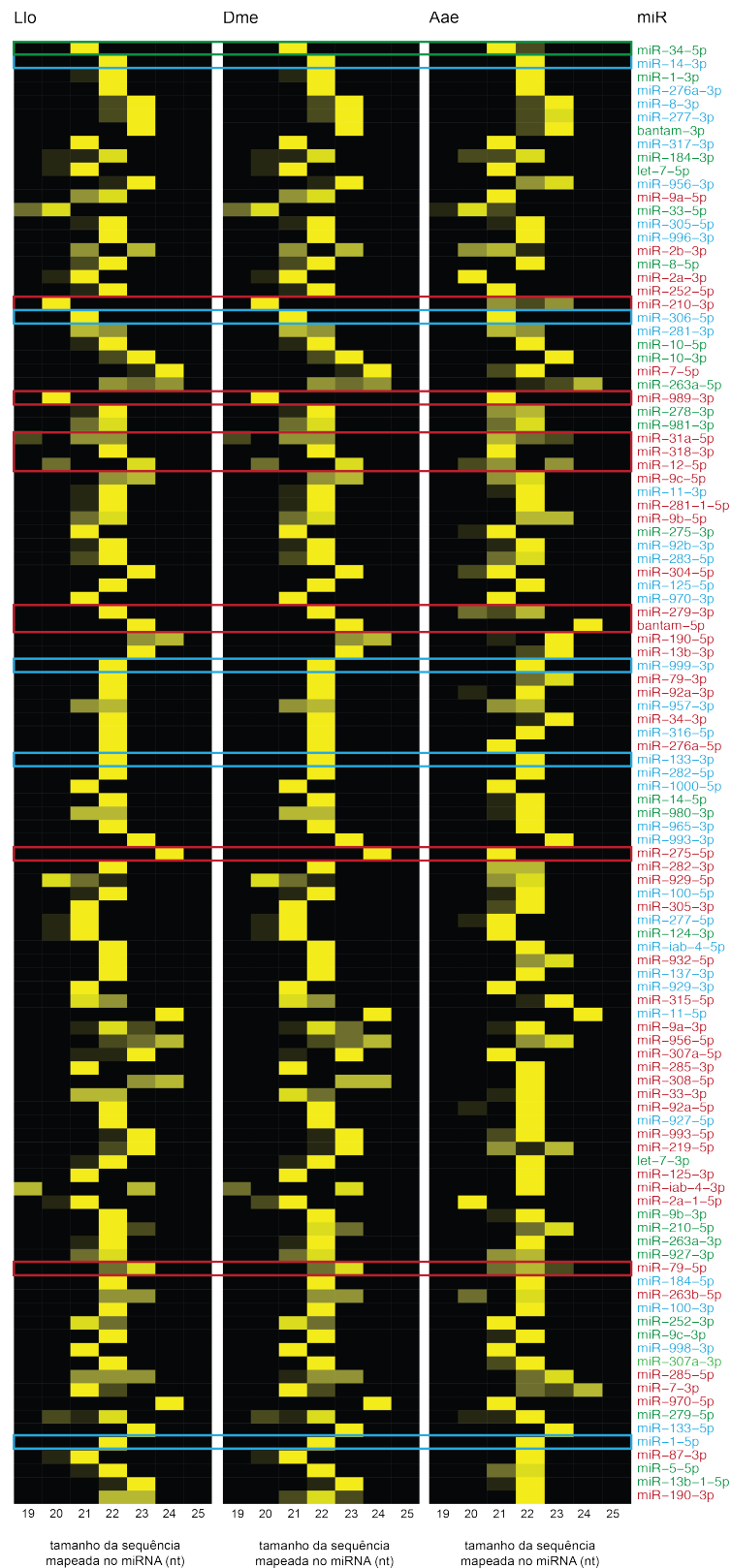


Figura 19: Heatmap do perfil de tamanho das seqüências mapeadas nos miRNAs conservados de cada espécie. Os miRNAs são apresentados em ordem decrescente de acordo com a expressão em Dme. **Legendas em vermelho:** Perfil discrepante; **Em azul:** Perfil idêntico; **Em verde:** Perfil similar; **Caixas no Heatmap:** miRNAs discutidos no texto; **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*.

Os miRNAs com perfil de tamanho variável não ocorreram devido à comparação com organismos próximos uma vez que alguns dos miRNAs anotados por essa estratégia apresentaram perfil de tamanho idêntico entre as três espécies (miR-989-3p, miR-276a-3p, miR-8-3p, miR-305, miR-996-3p, miR-282-5p, miR-929-3p, miR-1-5p entre outros).

A fim de avaliar quantitativamente a contribuição de cada tamanho das sequências mapeadas em todos os miRNAs conservados (por espécie), foi feita a normalização por *Z-score* e observou-se que miRNAs conservados entre insetos da ordem Diptera estudados tendem a ter tamanho preferencial de 22 nt (**FIGURA 20**), conforme esperado.

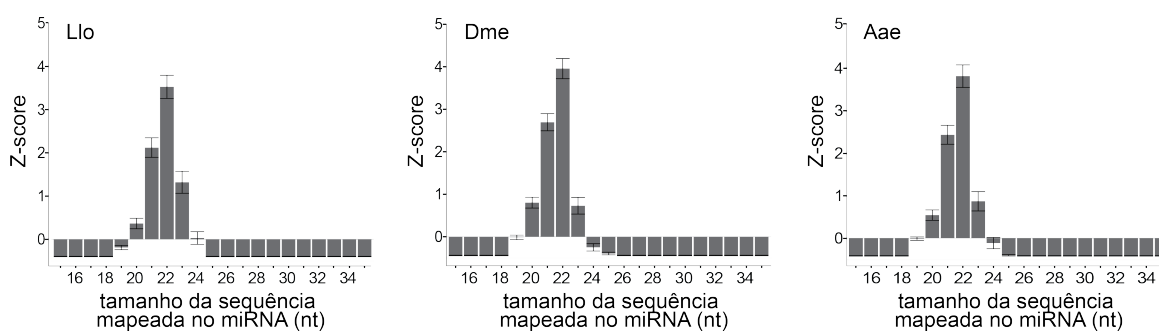


Figura 20: Perfil de tamanho dos miRNAs conservados entre os insetos estudados através da normalização por *Z-score*. Llo: *Lutzomyia longipalpis*; Dme: *Drosophila melanogaster*; Aae: *Aedes aegypti*; nt: nucleotídeo.

O teste de *Pearson* apresenta pequena diferença na correlação dos organismos do que fora apresentado em 5.4.2 (Expressão dos miRNAs segundo origem do precursor). Nestas análises, é observada maior correlação entre *L. longipalpis* e *A. aegypti* ($p < 0,786$, 95% de confiança), seguido de *A. aegypti* e *D. melanogaster* ($p < 0,776$, 95% de confiança) e *D. melanogaster* e *L. longipalpis* ($p < 0,751$, 95% de confiança). Em suma, observa-se alta correlação entre nos perfis de tamanho dos miRNAs das espécies estudadas.

Pela análise de diferenças entre os miRNAs, foram encontrados 22 miRNAs com *score* igual a 4 (**TABELA 12**), o que significa conservação de todas as características avaliadas entre os insetos estudados, incluindo identificação através do programa miRDeep2 e curadoria manual.

Tabela 12: miRNAs com todas as características conservadas (anotação pela mesma metodologia, tamanho e expressão conservados) entre os diferentes insetos estudados.

miR	Score Tamanho	Score Expressão	Score total	Interpretação
miR-14-3p	2	2	4	Controle
miR-277-3p	2	2	4	Controle
miR-317-3p	2	2	4	Controle
miR-956-3p	2	2	4	Controle
miR-11-3p	2	2	4	Controle
miR-92b-3p	2	2	4	Controle
miR-125-5p	2	2	4	Controle
miR-970-3p	2	2	4	Controle
miR-999-3p	2	2	4	Controle
miR-316-5p	2	2	4	Controle
miR-133-3p	2	2	4	Controle
miR-965-3p	2	2	4	Controle
miR-993-3p	2	2	4	Controle
miR-100-5p	2	2	4	Controle
miR-277-5p	2	2	4	Controle
miR-iab-4-5p	2	2	4	Controle
miR-137-3p	2	2	4	Controle
miR-11-5p	2	2	4	Controle
miR-927-5p	2	2	4	Controle
miR-184-5p	2	2	4	Controle
miR-100-3p	2	2	4	Controle
miR-133-5p	2	2	4	Controle

Com o objetivo de entender a divergência dos perfis de tamanho de alguns miRNAs, sem a influência de quaisquer outras variáveis, 14 miRNAs do grupo “Variação de tamanho” foram escolhidos para análises mais detalhadas de estrutura do precursor cujas observações feitas foram comparadas às de 14 miRNAs do grupo “Controle”. Esses miRNAs são apresentados na tabela (**TABELA 13**).

Tabela 13: miRNAs escolhidos para avaliação de perfil de tamanho dos precursores.

miR	Grupo
miR-14-3p	Controle
miR-277-3p	Controle
miR-317-3p	Controle
miR-11-3p	Controle
miR-92b-3p	Controle
miR-125-5p	Controle
miR-970-3p	Controle
miR-999-3p	Controle
miR-316-5p	Controle
miR-133-3p	Controle
miR-965-3p	Controle
miR-993-3p	Controle
miR-100-5p	Controle
miR-iab-4-5p	Controle
miR-9a-5p	Varição de Tamanho
miR-2b-3p	Varição de Tamanho
miR-252-5p	Varição de Tamanho
miR-210-3p	Varição de Tamanho
miR-7-5p	Varição de Tamanho
miR-31a-5p	Varição de Tamanho
miR-9c-5p	Varição de Tamanho
miR-9b-5p	Varição de Tamanho
miR-279-3p	Varição de Tamanho
bantam-5p	Varição de Tamanho
miR-190-3p	Varição de Tamanho
miR-13b-3p	Varição de Tamanho
miR-92a-3p	Varição de Tamanho
miR-275-5p	Varição de Tamanho

Os resultados de tamanho do precursor e MFE – também do precursor – não foram estatisticamente diferentes nas comparações intra-espécies dos grupos “Varição de tamanho” e “Controle”. Todavia, foi observado maior tamanho de precursores e menores valores de MFE de sequências derivadas de *Drosophila melanogaster* quando comparadas aos demais organismos. Até o momento, estes achados não tiveram sua justificativa elucidada, embora especula-se que a expressão tecido-específica ou os alvos para cada um deles possam ter influência sobre o perfil de tamanho do miRNA.

A diferença do tamanho dos miRNAs parece não se correlacionar com as alterações do padrão de expressão do miRNA entre as espécies, uma vez que de 46 miRNAs que apresentaram variação no perfil de tamanho, apenas miR-33-3p, miR-

993-5p, miR-970-5p e miR-7-3p pertencem ao grupo em que foram observadas alterações de ambas as características (perfil de tamanho e alteração de expressão).

Todavia, parece haver relação entre a razão de expressão dos braços 5p/3p e o perfil de tamanho dos miRNAs. Essa relação se observa, por exemplo nos miRNAs miR-33-3p e miR-7-3p, cuja expressão dos braços 5p/3p é quase idêntica em *A. aegypti*, mas varia até 400 vezes nos demais organismos. Entretanto, a metodologia para avaliação dessas diferenças está em processo de desenvolvimento, bem como a implementação de *scores* que considerem a conservação da sequência do miRNA entre os organismos.

Perfis diferentes de mapeamento das sequências são encontrados quando se comparam miRNAs não conservados entre as três espécies (**FIGURA 21** e **FIGURA 22**).

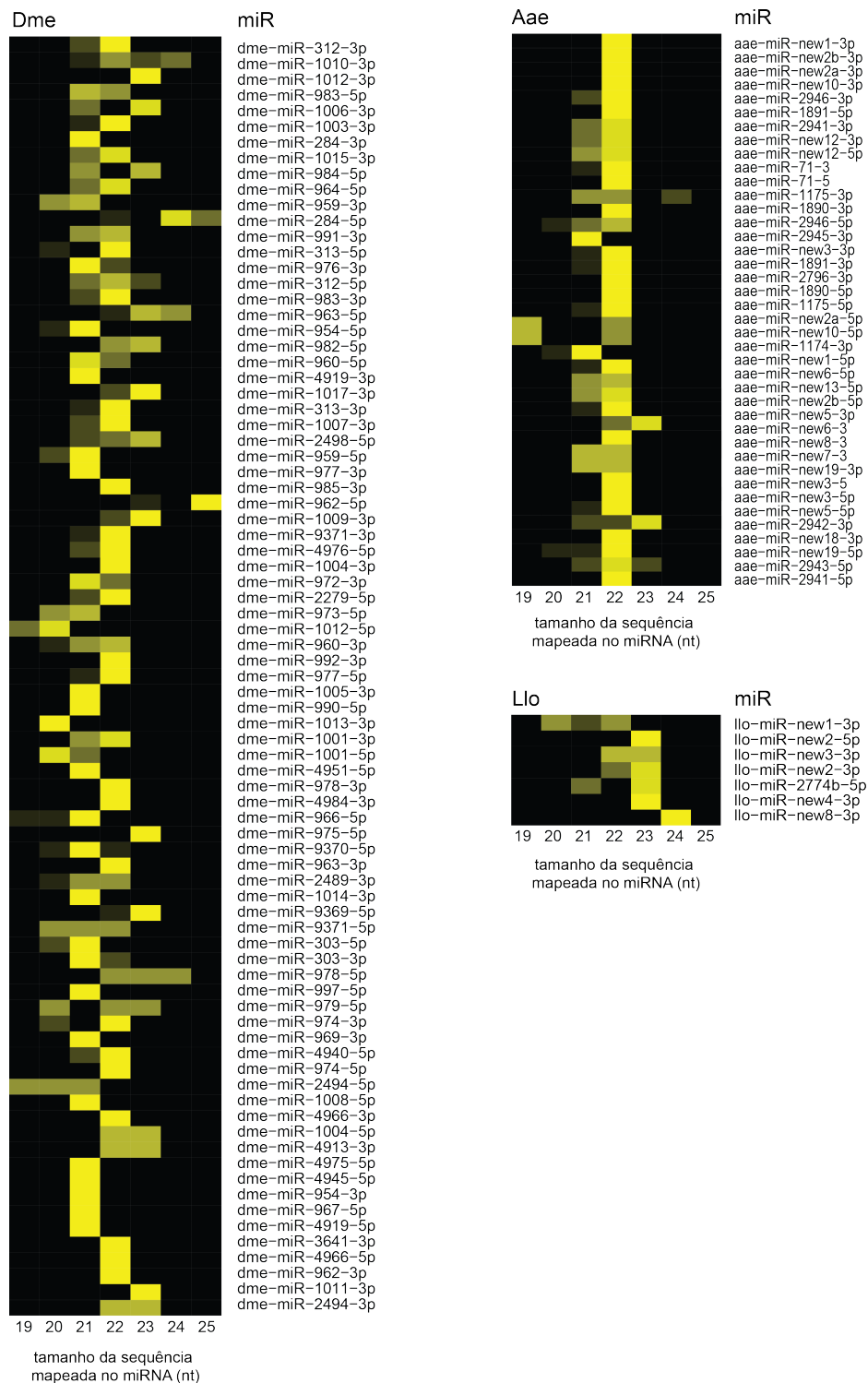


Figura 21: Heatmap do perfil de tamanho das sequências mapeadas nos miRNAs não-conservados de cada espécie. São apresentados somente miRNAs que tiveram expressão nas bibliotecas avaliadas. **Llo:** *Lutzomyia longipalpis*; **Dme:** *Drosophila melanogaster*; **Aae:** *Aedes aegypti*; **miR:** microRNA; **nt:** nucleotídeo.

A princípio, podem-se identificar diferenças entre o tamanho preferencial de miRNAs por eles gerados. O *heatmap* sugere que *L. longipalpis* possui um viés para

o tamanho de 23 nt. Em *A. aegypti* esse viés é apresentado no tamanho de 22 nt, embora alguns miRNAs sejam encontrados em tamanho bem distintos, como os miRNAs aae-new-10-3p e aae-new-3-3p. Esse resultado pode indicar uma preferência pela produção de tamanhos diferentes em tecidos distintos, como já foi reportado em trabalhos anteriores (Babak *et al.*, 2004). *Drosophila melanogaster*, por outro lado, apresenta o maior espectro de tamanho, sendo possível inclusive, encontrar miRNA de 25 nt (dme-miR-962-5p).

Quando verificada a contribuição de cada tamanho no número total de sequências mapeadas (**FIGURA 22**), a normalização por Z-score confirma os resultados do *heatmap*, isto é, há preferência de tamanho dos miRNAs espécie-específica.

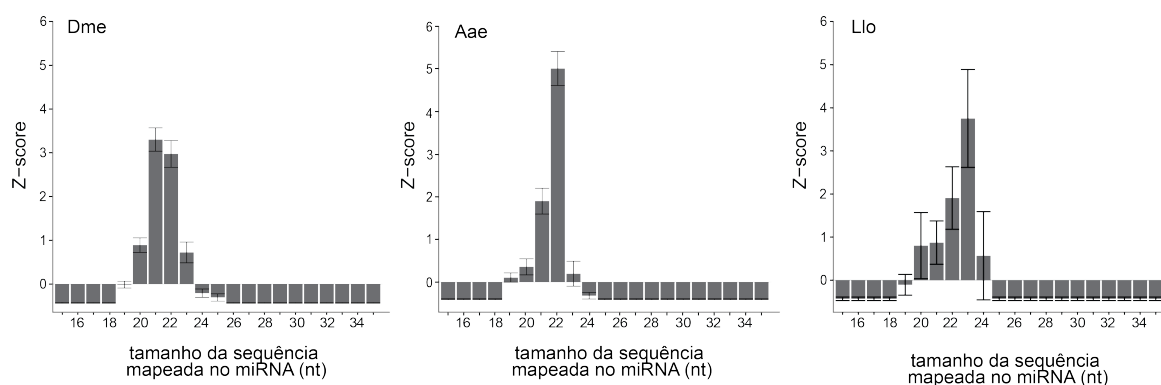


Figura 22: Perfil de tamanho dos miRNAs não-conservados entre os insetos estudados através da normalização por Z-score. Llo: *Lutzomyia longipalpis*; Dme: *Drosophila melanogaster*; Aae: *Aedes aegypti*; nt: nucleotídeos; nt: nucleotídeo.

Como mencionado, *L. longipalpis* apresentou Z-score maior para o tamanho de 23 nt e desvio-padrão elevado, devido ao baixo número de sequências analisadas. Em *A. aegypti* esse tamanho foi de 22 nt. Por outro lado, a contribuição dos tamanhos de 21 e 22 nt em *D. melanogaster* foi muito similar.

Tais resultados sugerem a existência de pressão evolutiva para a conservação do tamanho do miRNA, além de sua sequência. Assim, há pelo menos duas explicações para a diferença no tamanho de miRNAs não conservados entre as espécies estudadas: a) esses miRNAs estão em processo de seleção purificadora; b) esses miRNAs têm como alvos genes não-essenciais ao inseto, porém foram selecionados por conferir alguma vantagem evolutiva.

Como argumento a favor da hipótese definida em “a”, os miRNAs conservados surgiram há mais tempo que os miRNAs não-conservados, e deste modo sofreram o

processo de seleção natural. Em contraste, os miRNAs não conservados pode tratar-se de sinapomorfias desses organismos, não tendo sofrido as mesmas pressões desde o ancestral comum a todos.

Como argumento a favor da hipótese apresentada em “b”, a pressão evolutiva sofrida pelos miRNAs seria menor, justificando a existência do tamanho diferente ao observado em miRNAs conservados.

Contudo, para responder a essas questões, faz-se necessário avaliar os alvos desses miRNAs e o estabelecer dos processos biológicos em que eles participam.

5.7 Considerações finais

Neste trabalho mostrou-se que a via de miRNAs dos insetos da ordem Diptera estudados apresenta alta conservação do ponto de vista de sua biogênese, não se limitando à conservação proteica ou de sequência dos miRNAs como demonstrado em trabalhos anteriores. Foi observado que essa conservação se mantém em características tais como tamanho do miRNA e níveis de expressão, incluindo-se o braço mais expresso e análises estatísticas realizadas, além dos padrões moleculares que corroboram esses achados. Tais aspectos refletem a topologia da árvore filogenética dos organismos estudados, onde *L. longipalpis* e *A. aegypti* são mais próximos filogeneticamente entre si que entre *D. melanogaster*, o que demonstra que esse tipo de análise pode ser utilizado em estudos de filogenia.

A existência do tamanho de miRNAs espécie-específico poderia estar relacionada ao tempo de evolução da via de miRNAs. Neste caso, o fato de *D. melanogaster* ter surgido mais recentemente na escala evolutiva que *A. aegypti* e que *L. longipalpis*, poderia justificar a existência do espectro de tamanho de miRNAs observado neste inseto, isto é, não houve tempo evolutivo suficiente para selecionar o tamanho desses miRNAs.

Conclusões

6. CONCLUSÕES

A partir desse trabalho conclui-se:

1. Foi desenvolvida uma metodologia robusta, que permite a identificação e caracterização dos miRNAs independente do estado e disponibilidade da anotação e do genoma de referência utilizado;
2. Foi realizada a predição e anotação de 124 precursores de miRNAs, dos quais derivam 206 miRNAs maduros em *Lutzomyia longipalpis*. Além disso, este é um trabalho seminal, uma vez que não existem outros dados na literatura para miRNAs deste organismo;
3. A anotação de miRNAs em *Aedes aegypti*, foi melhorada. O número de sequências maduras passou de 124 (miRNAs disponíveis no banco de dados miRBase – versão 21), para 208 – a partir deste trabalho. Esse acréscimo corresponde a 67,7% de todos os miRNAs anotados para esta espécie. Além disso, foram incluídas as origens no precursor, isto é, braço 5p ou 3p, para todos os miRNAs que não dispunham de tal anotação;
4. A partir da metodologia proposta, foi possível identificar 9 novos precursores de miRNAs em *A. aegypti* e 30 novos precursores em *L. longipalpis*;
5. Foram encontrados 70 precursores únicos e 109 miRNAs conservados entre os três insetos estudados;
6. Existe alteração da origem de braço de dois miRNAs de *L. longipalpis* (llo-miR-275a-5p^{inv-3p} e llo-miR-2778-5p^{inv-3p}). Isto significa que ambos os miRNAs derivam de braços distintos daqueles observados em seus respectivos ortólogos. Observação semelhante ainda não foi descrita na literatura;
7. Existe alta conservação das características biológicas dos miRNAs conservados, tais como: preferência de base na região 5'; expressão; e perfil de tamanho;
8. A razão da expressão entre os braços 5p/3p é a característica biológica menos conservada entre as três espécies estudadas;
9. Quanto aos miRNAs não conservados, isto é, encontrados em apenas uma das espécies estudadas, esse número corresponde a 11 sequências em *L. longipalpis*, 26 em *A. aegypti* e 161 em *D. melanogaster*;
10. Os miRNAs não conservados apresentaram perfis de tamanho divergentes entre as espécies quando comparados aos miRNAs conservados

Perspectivas

7. PERSPECTIVAS

1. Desenvolver o *score* de conservação da sequência dos miRNAs entre os organismos;
2. Estabelecer a relação entre a variação dos perfis de tamanho e razão da expressão dos braços 5p e 3p dos miRNAs derivados de um mesmo precursor;
3. Avaliar a composição nucleotídica dos miRNAs conservados com perfil de tamanho similar e com perfil de tamanho diferente entre os três insetos a fim de verificar a ocorrência de preferência de bases que justifique a o tamanho espécie-específico observado entre os miRNAs não conservados;
4. Avaliar os alvos dos 109 miRNAs conservados entre os três insetos, estabelecer as vias biológicas envolvidas e verificar a conservação destas;
5. Avaliar os alvos dos miRNAs não conservados entre os três insetos, estabelecer as vias biológicas envolvidas e verificar a divergência destas;
6. Validar experimentalmente alvos compartilhados entre os três insetos.

Referências bibliográficas

8. REFERÊNCIAS BIBLIOGRÁFICAS

Websites utilizados:

FastX-Tool.

Disponível em: < http://hannonlab.cshl.edu/fastx_toolkit/index.html >

Diagrama de Venn.

Disponível em: < <http://bioinformatics.psb.ugent.be/webtools/Venn/>

Artigos:

AKBARI, O. S. et al. The developmental transcriptome of the mosquito *Aedes aegypti*, an invasive species and major arbovirus vector. **Genes, Genomics, Genetics**, v. 3, n. 9, p. 1493-509, Sep 2013. ISSN 2160-1836 (Electronic) 2160-1836 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23833213> >.

ALVAREZ-ERVITI, L. et al. Influence of microRNA deregulation on chaperone-mediated autophagy and alpha-synuclein pathology in Parkinson's disease. **Cell Death Dis**, v. 4, p. e545, 2013. ISSN 2041-4889 (Electronic). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23492776> >.

AMBROS, V. A uniform system for microRNA annotation. **Rna**, v. 9, n. 3, p. 277- 279, 2003. ISSN 13558382.

AMBROS, V. et al. MicroRNAs and Other Tiny Endogenous RNAs in *C. elegans*. **Current Biology**, v. 13, n. 10, p. 807-818, 2003. ISSN 09609822.

ASHBURNER, M.; BERGMAN, C. M. *Drosophila melanogaster*: a case study of a model genomic sequence and its consequences. **Genome Res**, v. 15, n. 12, p. 1661-7, Dec 2005. ISSN 1088-9051 (Print) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16339363> >.

AUYEUNG, V. C. et al. Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. **Cell**, v. 152, n. 4, p. 844- 58, Feb 14 2013. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23415231> >.

BABAK, T. et al. Probing microRNAs with microarrays: tissue specificity and functional inference. **RNA**, v. 10, n. 11, p. 1813-9, Nov 2004. ISSN 1355-8382 (Print) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15496526> >.

BARRIO, L.; DEKANTY, A.; MILAN, M. MicroRNA-mediated regulation of Dp53 in the *Drosophila* fat body contributes to metabolic adaptation to nutrient deprivation. **Cell Rep**, v. 8, n. 2, p. 528-41, Jul 24 2014. ISSN 2211-1247 (Electronic). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25017064> >.

BARTEL, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. **Cell**, v. 116, p. 281-297, 2004a.

BARTEL, D. P. MicroRNAs: target recognition and regulatory functions. **Cell**, v. 136, n. 2, p. 215-33, Jan 23 2009. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19167326> >.

BEREZIKOV, E. Evolution of microRNA diversity and regulation in animals. **Nat Rev Genet**, v. 12, n. 12, p. 846-60, Dec 2011. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22094948> >.

BEREZIKOV, E. et al. Deep annotation of *Drosophila melanogaster* microRNAs yields insights into their processing, modification, and emergence. **Genome Res**, v. 21, n. 2, p. 203-15, Feb 2011. ISSN 1549-5469 (Electronic) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21177969> >.

BERILLO, O. A. et al. Characteristics of binding sites of intergenic, intronic and exonic miRNA with mRNA oncogenes coding intronic miRNAs. **African Jour. of Biotech.**, 2013.

BERNSTEIN, E. et al. Role for a bidentate ribonuclease in the initiation step of RNA interference. **Nature**, 2001.

BIEMAR, F. et al. Spatial regulation of microRNA gene expression in the *Drosophila* embryo. **Proc Natl Acad Sci U S A**, v. 102, n. 44, p. 15907-11, Nov 1 2005. ISSN 0027-8424 (Print) 0027-8424 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16249329> >.

BOHNSACK, M. T. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. **Rna**, v. 10, n. 2, p. 185-191, 2004. ISSN 1355-8382.

BONNET, E. et al. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. **Bioinformatics**, v. 20, n. 17, p. 2911-7, Nov 22 2004. ISSN 1367-4803 (Print) 1367-4803 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15217813> >.

BRENNECKE, J. et al. bantam Encodes a Developmentally Regulated microRNA that Controls Cell Proliferation and Regulates the Proapoptotic Gene hid in *Drosophila*. **Cell**, 2003.

BRENNECKE, J. et al. Principles of microRNA-target recognition. **PLoS Biol**, v. 3, n. 3, p. e85, Mar 2005. ISSN 1545-7885 (Electronic) 1544-9173 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15723116> >.

CAI, X.; HAGEDORN, C. H.; CULLEN, B. R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. **RNA**, v. 10, n. 12, p. 1957-66, Dec 2004. ISSN 1355-8382 (Print) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15525708> >.

CAO, J. et al. Identification of conserved microRNAs in *Bombyx mori* (silkworm) and regulation of fibroin L chain production by microRNAs in heterologous system. **Insect Biochem Mol Biol**, v. 38, n. 12, p. 1066-71, Dec 2008. ISSN 1879-0240 (Electronic) 0965-1748 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18977439> >.

CHIANG, H. R. et al. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. **Genes Dev**, v. 24, n. 10, p. 992-1009, May 15 2010. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20413612> >.

CHU, C. Y.; RANA, T. M. Small RNAs: regulators and guardians of the genome. **J Cell Physiol**, v. 213, n. 2, p. 412-9, Nov 2007. ISSN 0021-9541 (Print) 0021-9541 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17674365> >.

CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics (Oxford, England)**, v. 21, p. 3674-6, 2005. Disponível em: < <http://bioinformatics.oxfordjournals.org/content/21/18/3674.full.pdf> >.

CROCE, C. M. Causes and consequences of microRNA dysregulation in cancer. **Nat Rev Genet**, v. 10, n. 10, p. 704-14, Oct 2009. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19763153> >.

CZECH, B. et al. Hierarchical rules for Argonaute loading in Drosophila. **Mol Cell**, v. 36, n. 3, p. 445-56, Nov 13 2009. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19917252> >.

DESVIGNES, T. et al. miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants. **Trends Genet**, v. 31, n. 11, p. 613-26, Nov 2015. ISSN 0168-9525 (Print) 0168-9525 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/26453491> >.

DOERKS, T. et al. Systematic identification of novel protein domain families associated with nuclear functions. **Genome Res**, v. 12, n. 1, p. 47-56, Jan 2002. ISSN 1088-9051 (Print) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11779830> >.

EULALIO, A.; HUNTZINGER, E.; IZAURRALDE, E. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. **Nat Struct Mol Biol**, v. 15, n. 4, p. 346-53, Apr 2008. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18345015> >.

FORSTEMANN, K. et al. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. **Cell**, v. 130, n. 2, p. 287-97, Jul 27 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17662943> >.

FORSTEMANN, K. et al. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. **PLoS Biol**, v. 3, n. 7, p. e236, Jul 2005. ISSN 1545-7885 (Electronic) 1544-9173 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15918770> >.

FRIEDLANDER, M. R. et al. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. **Nucleic Acids Res**, v. 40, n. 1, p. 37-52, Jan 2012. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em:

< <http://www.ncbi.nlm.nih.gov/pubmed/21911355> >.

GHILDIYAL, M. et al. Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. **RNA**, v. 16, n. 1, p. 43-56, Jan 2010. ISSN 1469-9001 (Electronic) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19917635> >.

GIRALDO-CALDERON, G. I. et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. **Nucleic Acids Res**, v. 43, n. Database issue, p. D707-13, Jan 2015. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25510499> >.

GRIFFITHS-JONES, S. Rfam: an RNA family database. **Nucleic Acids Research**, v. 31, n. 1, p. 439-441, 2003. ISSN 13624962.

GRIFFITHS-JONES, S. The microRNA Registry. **Nucleic Acids Res**, v. 32, n. Database issue, p. D109-11, Jan 1 2004. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14681370> >.

GRIFFITHS-JONES, S. et al. miRBase: microRNA sequences, targets and gene nomenclature. **Nucleic Acids Res**, v. 34, n. Database issue, p. D140-4, Jan 1 2006. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16381832> >.

GRIFFITHS-JONES, S. et al. MicroRNA evolution by arm switching. **EMBO Rep**, v. 12, n. 2, p. 172-7, Feb 2011. ISSN 1469-3178 (Electronic) 1469-221X (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21212805> >.

GRIFFITHS-JONES, S. et al. miRBase: tools for microRNA genomics. **Nucleic Acids Res**, v. 36, n. Database issue, p. D154-8, Jan 2008. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17991681> >.

GRIMSON, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. **Mol Cell**, v. 27, n. 1, p. 91-105, Jul 6 2007. ISSN 1097-2765 (Print) 1097-2765 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17612493> >.

GRIMSON, A. et al. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. **Nature**, v. 455, n. 7217, p. 1193-7, Oct 30 2008. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18830242> >.

GRISHOK, A. et al. Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control *C. elegans* Developmental Timing. **Cell**, 2001.

GRUNDHOFF, A.; SULLIVAN, C. S. Virus-encoded microRNAs. **Virology**, v. 411, n. 2, p. 325-43, Mar 15 2011. ISSN 1096-0341 (Electronic) 0042-6822 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21277611> >.

HA, M.; KIM, V. N. Regulation of microRNA biogenesis. **Nat Rev Mol Cell Biol**, v. 15, n. 8, p. 509-24, Aug 2014. ISSN 1471-0080 (Electronic) 1471-0072 (Linking).

Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25027649> >.

HERTEL, J.; STADLER, P. F. Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. **Bioinformatics**, v. 22, n. 14, p. e197- 202, Jul 15 2006. ISSN 1367-4811 (Electronic) 1367-4803 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16873472> >.

HINSKE, L. C. et al. A potential role for intragenic miRNAs on their hosts' interactome. **BMC Genomics**, v. 11, p. 533, 2010. ISSN 1471-2164 (Electronic) 1471-2164 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20920310> >.

HOFACKER, I. L. Vienna RNA secondary structure server. **Nucleic Acids Research**, v. 31, n. 13, p. 3429-3431, 2003. ISSN 1362-4962.

HU, H. Y. et al. Sequence features associated with microRNA strand selection in humans and flies. **BMC Genomics**, v. 10, p. 413, 2009. ISSN 1471-2164 (Electronic) 1471-2164 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19732433> >.

HU, W. et al. MicroRNAs of two medically important mosquito species: *Aedes aegypti* and *Anopheles stephensi*. **Insect Mol Biol**, v. 24, n. 2, p. 240-52, Apr 2015. ISSN 1365-2583 (Electronic) 0962-1075 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25420875> >.

HUANG, V. et al. Upregulation of Cyclin B1 by miRNA and its implications in cancer. **Nucleic Acids Res**, v. 40, n. 4, p. 1695-707, Feb 2012. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22053081> >.

HUTVÁGNER, G. et al. A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA. **Science**, 2001.

HUTVAGNER, G.; SIMARD, M. J. Argonaute proteins: key players in RNA silencing. **Nat Rev Mol Cell Biol**, v. 9, n. 1, p. 22-32, Jan 2008. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18073770> >.

ISIK, M.; KORSWAGEN, H. C.; BEREZIKOV, E. Expression patterns of intronic microRNAs in *Caenorhabditis elegans*. **Silence**, 2010.

IWASAKI, S.; KAWAMATA, T.; TOMARI, Y. *Drosophila* argonaute1 and argonaute2 employ distinct mechanisms for translational repression. **Mol Cell**, v. 34, n. 1, p. 58-67, Apr 10 2009. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19268617> >.

JIANG, F. et al. Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. **Genes Dev**, v. 19, n. 14, p. 1674-9, Jul 15 2005. ISSN 0890-9369 (Print) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15985611> >.

JOHNSTON, R. J.; HOBERT, O. A microRNA controlling left/right neuronal asymmetry

in *Caenorhabditis elegans*. **Nature**, Dec 18 2003. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14685239> >.

KAWAMATA, T.; SEITZ, H.; TOMARI, Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. **Nat Struct Mol Biol**, v. 16, n. 9, p. 953-60, Sep 2009. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19684602> >.

KHVOROVA, A.; REYNOLDS, A.; JAYASENA, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. **Cell**, 2003.

KIEZUN, A. et al. miRviewer: a multispecies microRNA homologous viewer. **BMC Res. Notes**, 2012.

KIM, V. N. MicroRNA biogenesis: coordinated cropping and dicing. **Nat Rev Mol Cell Biol**, v. 6, n. 5, p. 376-85, May 2005. ISSN 1471-0072 (Print) 1471-0072 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15852042> >.

KIM, V. N.; HAN, J.; SIOMI, M. C. Biogenesis of small RNAs in animals. **Nat Rev Mol Cell Biol**, v. 10, n. 2, p. 126-39, Feb 2009. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19165215> >.

KIM, V. N.; NAM, J. W. Genomics of microRNA. **Trends Genet**, v. 22, n. 3, p. 165- 73, Mar 2006. ISSN 0168-9525 (Print) 0168-9525 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16446010> >.

KIM, Y.-K.; KIM, V. N. Processing of intronic microRNA. **EMBO**, 2007.

KIRIAKIDOU, M. et al. An mRNA m7G cap binding-like motif within human Ago2 represses translation. **Cell**, v. 129, n. 6, p. 1141-51, Jun 15 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17524464> >.

KOZOMARA, A.; GRIFFITHS-JONES, S. miRBase: integrating microRNA annotation and deep-sequencing data. **Nucleic Acids Res**, v. 39, n. Database issue, p. D152-7, Jan 2011. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21037258> >.

KOZOMARA, A.; GRIFFITHS-JONES, S. miRBase: annotating high confidence microRNAs using deep sequencing data. **Nucleic Acids Res**, v. 42, n. Database issue, p. D68-73, Jan 2014. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24275495> >.

KUGLER, J. M. et al. miR-989 is required for border cell migration in the *Drosophila* ovary. **PLoS One**, v. 8, n. 7, p. e67075, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23843983> >.

LAGOS-QUINTANA, M. New microRNAs from mouse and human. **Rna**, v. 9, n. 2,

p. 175-179, 2003. ISSN 13558382.

LAGOS-QUINTANA, M. et al. Identification of tissue-specific microRNAs from mouse. **Current Biology**, 2001a.

LAGOS-QUINTANA, M. et al. Identification of Tissue-Specific MicroRNAs from Mouse. **Curr Biol**, 2001b.

LAI, E. C. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. **Nat Genet**, v. 30, n. 4, p. 363-4, Apr 2002.

ISSN 1061-4036 (Print) 1061-4036 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11896390> >.

LANGMEAD, B. et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biol**, v. 10, n. 3, p. R25, 2009. ISSN 1474-760X (Electronic) 1474-7596 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19261174> >.

LAU, N. C. et al. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in *Caenorhabditis elegans*. **Sicie**, 2001.

LEE, C. T.; RISOM, T.; STRAUSS, W. M. Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. **DNA Cell Biol**, v. 26, n. 4, p. 209-18, Apr 2007. ISSN 1044-5498 (Print) 1044-5498 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17465887> >.

LEE, G. J.; JUN, J. W.; HYUN, S. MicroRNA miR-8 regulates multiple growth factor hormones produced from *Drosophila* fat cells. **Insect Mol Biol**, v. 24, n. 3, p. 311-8, Jun 2015. ISSN 1365-2583 (Electronic) 0962-1075 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25492518> >.

LEE, R. C.; AMBROS, V. An extensive class of small RNAs in *Caenorhabditis elegans*. **Science**, v. 294, n. 5543, p. 862-4, Oct 26 2001. ISSN 0036-8075 (Print) 0036-8075 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11679672> >.

LEE, R. C.; FEINBAUM, R. L.; V., A. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. **Cell**, 1993.

LEE, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. **Nature**, 2003.

LEE, Y. et al. MicroRNA maturation: stepwise processing and subcellular localization. **EMBO** 2002.

LEE, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. **EMBO**, 2004.

LI, S. et al. Direct sequencing and expression analysis of a large number of miRNAs in *Aedes aegypti* and a multi-species survey of novel mosquito miRNAs. **BMC Genomics**, v. 10, p. 581, 2009. ISSN 1471-2164 (Electronic)

1471-2164 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19961592> >.

LIM, L. P. et al. The microRNAs of *Caenorhabditis elegans*. **Genes Dev**, v. 17, n. 8, p. 991-1008, Apr 15 2003. ISSN 0890-9369 (Print)

0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/12672692> >.

LIU, X. et al. Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. **RNA**, v. 12, n. 8, p. 1514-20, Aug 2006. ISSN 1355-8382 (Print)

1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16775303> >.

LIU, Y. et al. The expression profile of *Aedes albopictus* miRNAs is altered by dengue virus serotype-2 infection. **Cell Biosci**, v. 5, p. 16, 2015. ISSN 2045-3701 (Electronic)

2045-3701 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25922662> >.

LUCAS, K.; RAIKHEL, A. S. Insect microRNAs: biogenesis, expression profiling and biological functions. **Insect Biochem Mol Biol**, v. 43, n. 1, p. 24-38, Jan 2013. ISSN 1879-0240 (Electronic) 0965-1748 (Linking). Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/23165178> >.

LUND, E. et al. Nuclear export of microRNA precursors. **Science**, v. 303, n. 5654, p. 95-8, Jan 2 2004. ISSN 1095-9203 (Electronic)

0036-8075 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14631048> >.

MACRAE, I. J.; ZHOU, K.; DOUDNA, J. A. Structural determinants of RNA recognition and cleavage by Dicer. **Nat Struct Mol Biol**, v. 14, n. 10, p. 934-40, Oct 2007. ISSN 1545-9993 (Print) 1545-9985 (Linking). Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/17873886> >.

MARCO, A. Sex-biased expression of microRNAs in *Drosophila melanogaster*.

Open Biol, v. 4, p. 140024, 2014. ISSN 2046-2441 (Electronic)

2046-2441 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24694940> >.

MARCO, A.; HOOKS, K.; GRIFFITHS-JONES, S. Evolution and function of the extended miR-2 microRNA family. **RNA Biol**, v. 9, n. 3, p. 242-8, Mar 2012. ISSN 1555-8584 (Electronic) 1547-6286 (Linking). Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/22336713> >.

MARCO, A. et al. Functional shifts in insect microRNA evolution. **Genome Biol Evol**, v. 2, p. 686-96, 2010. ISSN 1759-6653 (Electronic)

1759-6653 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20817720> >.

MARCO, A. et al. MicroRNAs from the same precursor have different targeting properties. **Silence**, 2012.

MILNE, I. et al. Using Tablet for visual exploration of second-generation sequencing data. **Brief Bioinform**, v. 14, n. 2, p. 193-202, Mar 2013. ISSN 1477-4054 (Electronic)

1467-5463 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22445902> >.

NICHOLSON, A. W. Function, mechanism and regulation of bacterial ribonucleases. **FEMS Microbiology Review**, 1999.

NIELSEN, C. B. et al. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. **RNA**, v. 13, n. 11, p. 1894-910, Nov 2007. ISSN 1355-8382 (Print) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17872505> >.

NINOVA, M.; RONSHAUGEN, M.; GRIFFITHS-JONES, S. Fast-evolving microRNAs are highly expressed in the early embryo of *Drosophila virilis*. **RNA**, v. 20, n. 3, p. 360-72, Mar 2014. ISSN 1469-9001 (Electronic) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24448446> >.

OKAMURA, K. et al. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. **Cell**, v. 130, n. 1, p. 89-100, Jul 13 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17599402> >.

OKAMURA, K. et al. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. **Genes Dev**, v. 18, n. 14, p. 1655-66, Jul 15 2004. ISSN 0890-9369 (Print) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15231716> >.

OKAMURA, K. et al. Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals. **Genes Dev**, v. 27, n. 7, p. 778-92, Apr 1 2013. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23535236> >.

OKAMURA, K.; LIU, N.; LAI, E. C. Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. **Mol Cell**, v. 36, n. 3, p. 431-44, Nov 13 2009. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19917251> >.

OROM, U. A.; NIELSEN, F. C.; LUND, A. H. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. **Mol Cell**, v. 30, n. 4, p. 460-71, May 23 2008. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18498749> >.

PARRISH, J. Z. et al. The microRNA bantam functions in epithelial cells to regulate scaling growth of dendrite arbors in *Drosophila* sensory neurons. **Neuron**, v. 63, n. 6, p. 788-802, Sep 24 2009. ISSN 1097-4199 (Electronic) 0896-6273 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19778508> >.

PASQUINELLI, A. E. et al. Conservation across animal phylogeny of the sequence

and temporal regulation of the 21 nucleotide let-7 heterochronic regulatory RNA. **Nature**, 2000.

PAVLAKIS, G. N. et al. Sequence and secondary structure of *Drosophila melanogaster* 5.8S and 2S RNAs and of the processing site between them. **Nucleic Acids Res**, 1979.

PLACE, R. F. et al. MicroRNA-373 induces expression of genes with complementary promoter sequences. **Proc Natl Acad Sci U S A**, v. 105, n. 5, p. 1608-13, Feb 5 2008. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18227514> >.

PRATT, A. J.; MACRAE, I. J. The RNA-induced silencing complex: a versatile gene-silencing machine. **J Biol Chem**, v. 284, n. 27, p. 17897-901, Jul 3 2009. ISSN 1083-351X (Electronic) 0021-9258 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19342379> >.

PRATT, A. J.; MACRAE, I. J. The RNA-induced Silencing Complex: A Versatile Gene-silencing Machine. **J Biol Chem.**, 2009.

QUINLAN, A. R.; HALL, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. **Bioinformatics**, v. 26, n. 6, p. 841-2, Mar 15 2010. ISSN 1367-4811 (Electronic) 1367-4803 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20110278> >.

RAMALINGAM, P. et al. Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. **RNA**, v. 20, n. 1, p. 76-87, Jan 2014. ISSN 1469-9001 (Electronic) 1355-8382 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24226766> >.

REINHART, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. **Nature**, 2000.

RIVAS, F. V. et al. Purified Argonaute2 and an siRNA form recombinant human RISC. **Nat Struct Mol Biol**, v. 12, n. 4, p. 340-9, Apr 2005. ISSN 1545-9993 (Print) 1545-9985 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15800637> >.

RUBY, J. G.; JAN, C. H.; BARTEL, D. P. Intronic microRNA precursors that bypass Drosha processing. **Nature**, v. 448, n. 7149, p. 83-6, Jul 5 2007. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17589500> >.

SAITO, K. et al. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. **PLoS Biol**, v. 3, n. 7, p. e235, Jul 2005. ISSN 1545-7885 (Electronic) 1544-9173 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15918769> >.

SCHONROCK, N. et al. Neuronal microRNA deregulation in response to Alzheimer's disease amyloid-beta. **PLoS One**, v. 5, n. 6, p. e11070, 2010. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). Disponível em: <

<http://www.ncbi.nlm.nih.gov/pubmed/20552018> >.

SCHWARZ, D. S. et al. Asymmetry in the Assembly of the RNAi Enzyme Complex. **Cell**, 2003.

SCOTT, H. et al. MiR-3120 is a mirror microRNA that targets heat shock cognate protein 70 and auxilin messenger RNAs and regulates clathrin vesicle uncoating. **J Biol Chem**, v. 287, n. 18, p. 14726-33, Apr 27 2012. ISSN 1083-351X (Electronic) 0021-9258 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22393045> >.

SEITZ, H.; TUSHIR, J. S.; ZAMORE, P. D. A 5'-uridine amplifies miRNA/miRNA* asymmetry in Drosophila by promoting RNA-induced silencing complex formation. **Silence**, 2011.

SHABALINA, S. A.; KOONIN, E. V. Origins and evolution of eukaryotic RNA interference. **Trends Ecol Evol**, v. 23, n. 10, p. 578-87, Oct 2008. ISSN 0169-5347 (Print) 0169-5347 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18715673> >.

SKALSKY, R. L.; CULLEN, B. R. Viruses, microRNAs, and host interactions. **Annu Rev Microbiol**, v. 64, p. 123-41, 2010. ISSN 1545-3251 (Electronic) 0066-4227 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/20477536> >.

SKINNER, M. E. et al. JBrowse: a next-generation genome browser. **Genome Res**, v. 19, n. 9, p. 1630-8, Sep 2009. ISSN 1549-5469 (Electronic) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19570905> >.

SLEZAK-PROCHAZKA, I. et al. Cellular localization and processing of primary transcripts of exonic microRNAs. **PLoS One**, v. 8, n. 9, p. e76647, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24073292> >.

SMIT, A. F. The origin of interspersed repeats in the human genome. **Curr Opin Genet Dev**, v. 6, n. 6, p. 743-8, Dec 1996. ISSN 0959-437X (Print) 0959-437X (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/8994846> >.

SOKOL, N. S.; AMBROS, V. Mesodermally expressed Drosophila microRNA-1 is regulated by Twist and is required in muscles during larval growth. **Genes Dev**, v. 19, n. 19, p. 2343-54, Oct 1 2005. ISSN 0890-9369 (Print) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16166373> >.

SONI, K. et al. miR-34 is maternally inherited in Drosophila melanogaster and Danio rerio. **Nucleic Acids Res**, v. 41, n. 8, p. 4470-80, Apr 2013. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23470996> >.

STAREGA-ROSLAN, J.; GALKA-MARCINIAK, P.; KRZYZOSIAK, W. J. Nucleotide

sequence of miRNA precursor contributes to cleavage site selection by Dicer. **Nucleic Acids Res**, Sep 30 2015a. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/26424848> >.

SU, H. et al. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. **Genes Dev**, v. 23, n. 3, p. 304-17, Feb 1 2009. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19174539> >.

TEAM, R. D. C. R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**, 2011.

TEHLER, D.; HOYLAND-KROGHSBO, N. M.; LUND, A. H. The miR-10 microRNA precursor family. **RNA Biol**, v. 8, n. 5, p. 728-34, Sep-Oct 2011. ISSN 1555-8584 (Electronic) 1547-6286 (Linking). Disponível em:< <http://www.ncbi.nlm.nih.gov/pubmed/21881411> >.

THOMSON, R. C. et al. A critical appraisal of the use of microRNA data in phylogenetics. **Proc Natl Acad Sci U S A**, v. 111, n. 35, p. E3659-68, Sep 2 2014. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25071211> >.

TOMARI, Y.; DU, T.; ZAMORE, P. D. Sorting of Drosophila small silencing RNAs. **Cell**, v. 130, n. 2, p. 299-308, Jul 27 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/17662944> >.

TOMARI, Y.; ZAMORE, P. D. MicroRNA biogenesis: drosha can't cut it without a partner. **Curr Biol**, v. 15, n. 2, p. R61-4, Jan 26 2005. ISSN 0960-9822 (Print) 0960-9822 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15668159> >.

TSAI, N. P.; LIN, Y. L.; WEI, L. N. MicroRNA mir-346 targets the 5'-untranslated region of receptor-interacting protein 140 (RIP140) mRNA and up-regulates its protein expression. **Biochem J**, v. 424, n. 3, p. 411-8, Dec 15 2009. ISSN 1470-8728 (Electronic) 0264-6021 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/19780716> >.

TSUTSUMI, A. et al. Recognition of the pre-miRNA structure by Drosophila Dicer-1. **Nat Struct Mol Biol**, v. 18, n. 10, p. 1153-8, Oct 2011. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). Disponível em:< <http://www.ncbi.nlm.nih.gov/pubmed/21926993> >.

TYLER, D. M. et al. Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. **Genes Dev**, v. 22, n. 1, p. 26-36, Jan 1 2008. ISSN 0890-9369 (Print) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/18172163> >.

VALENCIA-SANCHEZ, M. A. et al. Control of translation and mRNA degradation by miRNAs and siRNAs. **Genes Dev**, 2006. ISSN 1088-9051 (Print) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/11779830> >.

VALINEZHAD ORANG, A.; SAFARALIZADEH, R.; KAZEMZADEH-BAVILI, M.

Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. **Int J Genomics**, v. 2014, p. 970607, 2014. ISSN 2314-436X (Print) 2314-436X (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/25180174> >.

VASUDEVAN, S.; TONG, Y.; STEITZ, J. A. Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation **Science**, 2007.

WANG, B. Base Composition Characteristics of Mammalian miRNAs. **J Nucleic Acids**, v. 2013, p. 951570, 2013. ISSN 2090-0201 (Print) 2090-0201 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23710337> >.

WEN, J. et al. Diversity of miRNAs, siRNAs, and piRNAs across 25 Drosophila cell lines. **Genome Res**, v. 24, n. 7, p. 1236-50, Jul 2014. ISSN 1549-5469 (Electronic) 1088-9051 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24985917> >.

WHO, W. H. O. Global strategy for Dengue prevention and control. ISSN 978 92 4 150403 4.

WIEGMANN, B. M. et al. Episodic radiations in the fly tree of life. **Proc Natl Acad Sci U S A**, v. 108, n. 14, p. 5690-5, Apr 5 2011. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21402926> >.

WIGHTMAN, B. et al. Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development **Genes Dev**, 1991.

WIGHTMAN, B.; HA, I.; RUVKUN, G. Posttranscriptional Regulation of the Heterochronic Gene lin-14 by lin-4 Mediates Temporal Pattern Formation in C. elegans. **Cell**, v. 75, 1993.

XU, P. et al. The Drosophila MicroRNA Mir-14 Suppresses Cell Death and Is Required for Normal Fat Metabolism. **Current Biology**, v. 13, n. 9, p. 790-795, 2003. ISSN 09609822.

YI, R. et al. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. **Genes Dev**, v. 17, n. 24, p. 3011-6, Dec 15 2003. ISSN 0890-9369 (Print) 0890-9369 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/14681208> >.

ZENG, Y.; CULLEN, B. R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. **Nucleic Acids Res**, v. 32, n. 16, p. 4776-85, 2004. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/15356295> >.

ZENG, Y.; YI, R.; CULLEN, B. R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. **EMBO**, 2005.

ZHANG, X. et al. microRNAs play critical roles in the survival and recovery of Caenorhabditis elegans from starvation-induced L1 diapause. **PNAS**, 2011.

ZONDAG, L.; DEARDEN, P. K.; WILSON, M. J. Deep sequencing and expression of

microRNAs from early honeybee (*Apis mellifera*) embryos reveals a role in regulating early embryonic patterning. **BMC Evol Biol**, v. 12, p. 211, 2012. ISSN 1471-2148 (Electronic) 1471-2148 (Linking). Disponível em: <
<http://www.ncbi.nlm.nih.gov/pubmed/23121997> >.