

Investigação da função biológica dos eventos de *arm switching* de microRNAs e sua potencial associação com a metilação de RNA m<sup>6</sup>A utilizando *zebrafish* como organismo modelo

**Arthur Casulli de Oliveira**

UNIVERSIDADE ESTADUAL PAULISTA  
"Júlio de Mesquita Filho"  
INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

Investigação da função biológica dos eventos de *arm switching* de microRNAs e sua potencial associação com a metilação de RNA m<sup>6</sup>A utilizando *zebrafish* como organismo modelo

**Candidato:** Arthur Casulli de Oliveira

**Orientador:** Danillo Pinhal

**Co-orientador:** Luiz Augusto Bovolenta

Tese apresentada ao Instituto de Biociências, Campus de Botucatu, UNESP, para obtenção do título de Doutor no Programa de Pós- Graduação em Ciências Biológicas (Genética).

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP

BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Oliveira, Arthur Casulli de.

Investigação da função biológica dos eventos de *arm switching* de microRNAs e sua potencial associação com a metilação de RNA m<sup>6</sup>A utilizando *zebrafish* como organismo modelo / Arthur Casulli de Oliveira. - Botucatu, 2021

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Danillo Pinhal

Coorientador: Luiz Augusto Bovolenta

Capes: 20204000

1. Biologia computacional. 2. Bioinformática.  
3. MicroRNAs. 4. Sequenciamento de nucleotídeo. 5. *Zebrafish*.  
6. miRNAs.

Palavras-chave: Biogênese de miRNAs; Bioinformática;  
Modificação de RNAs; RNA-seq; RNAs não-codificadores.

Dedico este trabalho à minha mãe e aos meus avós, que me forneceram todo carinho e suporte necessário no decorrer do meu doutorado e à minha noiva, que sempre esteve ao meu lado, me apoiando, me incentivando e me fazendo crescer.

## Agradeço

Ao meu orientador, Prof. Dr. Danillo Pinhal, pelo suporte, orientação científica e amizade durante todos estes anos em que fui seu aluno e por sempre me incentivar a encarar novos desafios nesta jornada acadêmica.

À minha mãe, Fernanda Bressanelli Casulli, por todo amor e carinho e por sempre estar ao meu lado me apoiando e educando, nos momentos felizes e tristes da minha vida.

Aos meus avós, Ida Maria e Pilade, por todo amor, carinho e suporte que sempre me deram durante todos os anos da minha vida e, principalmente, durante anos de graduação e pós-graduação.

Ao meu irmão, Victor Casulli de Oliveira, por todos estes anos de amizade, brincadeiras e brigas e por sempre estar ao meu lado, nas risadas e nas tristezas.

À minha noiva e amor da minha vida, Karina Gabriele Alves Dias (Nuros), por ser essa pessoa incrível e mostrar todo o amor e carinho que você tem comigo, por estar sempre ao meu lado em todos os momentos da minha vida, me apoiando, me incentivando, me ajudando e me mostrando como crescer e ser sempre alguém melhor, e por querer compartilhar sua vida e seu futuro comigo. Te amo para sempre.

Aos meus sogros, Régis e Dona Melita e minha cunhada Melissa, por todo o carinho que sempre me deram e por terem me acolhido na família, me fazendo sempre me sentir em casa.

A todos os colegas do Laboratório Genômica e Evolução Molecular pelas conversas e discussões sobre os mais diversos temas relacionados à pesquisa e à ciência.

Aos amigos de trabalho Luiz (Chokito), Lucas Di Pietro e Lucas Alves, por todas as brincadeiras, risadas e colaborações realizadas durante todo o período do meu doutorado e por sempre estarem dispostos a colaborar durante o desenvolvimento deste projeto.

A todos os amigos feitos em meu novo hobby jogos de tabuleiro, que me abriu novas portas de conhecimento e crescimento pessoal e profissional.

Aos funcionários da seção de Pós-Graduação por todo trabalho realizado e dúvidas esclarecidas.

Ao Laboratório Genômica e Evolução Molecular, ao Departamento de Ciências Químicas e Biológicas, à Pós-graduação em Ciências Biológicas (Genética), ao Instituto de Biociências de Botucatu e à Universidade Estadual Paulista pela estrutura cedida para a realização deste trabalho.

A todas as pessoas que de alguma forma, direta ou indiretamente, auxiliaram na realização e finalização deste trabalho.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de financiamento 001 e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) processo nº 2017/17510-2, cujo o apoio financeiro foi essencial para realização dessa pesquisa.

*“Se eu vi mais longe, foi por estar sobre ombros de gigantes.”*

*- Isaac Newton*

## Resumo

Os microRNAs (miRNAs) são atualmente conhecidos como uma ampla classe de moléculas de RNAs reguladores da expressão gênica. Durante a biogênese, uma fita ("braço") do miRNA *duplex* é preferencialmente selecionada como um miRNA funcionalmente maduro, enquanto a outra fita é usualmente degradada. Eventos de *arm switching* ocorrem quando há mudança na preferência do braço selecionado. Isso se reflete na alteração do braço identificado como transcrito funcional, e varia de forma bastante dinâmica entre estágios do desenvolvimento, tipos de tecidos e mesmo entre espécies. Devido às particularidades de seu mecanismo de ação, os *arm switching* concedem aos miRNAs um ampliado potencial regulatório ao modificar o conjunto de alvos e, comumente, os processos biológicos por eles regulados. Entretanto, os mecanismos moleculares que promovem episódios de *arm switching* ainda necessitam ser melhor caracterizados. A identificação de metilações m<sup>6</sup>A em pri-miRNAs de linhagens celulares humanas reportada na literatura forneceu indícios de que este tipo de metilação poderia estar envolvida no processo de *arm switching*, principalmente devido à alteração que m<sup>6</sup>A promove na estrutura secundária de RNAs. Para testar esta hipótese, utilizamos o *zebrafish*, um vertebrado modelo, para avaliar se a metilação m<sup>6</sup>A poderia estar influenciando na ocorrência de *arm switching* nesta espécie. Para isso, confrontamos o perfil global de expressão dos braços 5p e 3p dos miRNAs pela análise de uma ampla coleção de dados de RNA-seq de períodos iniciais do desenvolvimento embrionário e diversos tecidos adultos de *zebrafish* obtidos da literatura. Adicionalmente, aplicamos a técnica de sequenciamento de metilações m<sup>6</sup>A (m<sup>6</sup>A-seq) no transcriptoma para identificar potencial correlação entre os picos de metilação presentes nos miRNAs e os eventos de *arm switching* identificados. Este conjunto de análises nos conduziu à inédita identificação de 14 eventos de *arm switching* em *zebrafish*. Pelo estudo detalhado destes casos e sua comparação aos eventos descritos em outras espécies verificamos que eventos de *arm switching* estão majoritariamente associados ao desenvolvimento ontogenético dos vertebrados. Adicionalmente, nossas análises sugerem que a expressão diferencial de isoformas de miRNAs (isomiRs) pode ser um importante modulador na prevalência dos braços. Também demonstramos, via análises de predição de alvos e enriquecimento funcional, que os eventos de *arm switching* de fato potencializam a regulação exercida pelos miRNAs, ao aumentarem o número de genes alvos, redes regulatórias e funções biológicas aos quais estão associados. No entanto, surpreendentemente, a análise criteriosa dos dados de m<sup>6</sup>A-seq nos forneceu indícios de que os pri-miRNAs de *zebrafish* não são efetivamente metilados, ainda que contenham sítios de metilação m<sup>6</sup>A em suas sequências. A

discrepância entre os resultados obtidos neste trabalho e aqueles reportados em humanos e, mais recentemente, em *Arabidopsis thaliana*, sugerem que pressões seletivas atuantes na biogênese dos miRNAs poderiam estar ocorrendo em diferentes estágios nestas espécies. Ainda, considerando dados recentes da literatura sobre a interação entre a proteína efetora de metilações m<sup>6</sup>A e moléculas de RNA, conjecturamos que a estrutura secundária dos pri-miRNAs pode ser um fator permissivo desta modulação. Nesse contexto, concluímos que o *arm switching* é um fenômeno conservado e que está majoritariamente associado ao desenvolvimento ontogenético dos organismos. Também não encontramos evidências de que a metilação m<sup>6</sup>A interfira na prevalência dos braços de miRNAs em *zebrafish*, dada sua ausência aparente nos transcritos de pri-miRNAs desta espécie. Os resultados obtidos neste estudo trazem novas contribuições para a compreensão de mecanismos regulatórios inerentes à atividade funcional de miRNAs. Adicionalmente, fornecem subsídios para pesquisas futuras interessadas em explorar os mecanismos moleculares subjacentes às diferentes taxas de metilação e sua interferência na função e dinâmica evolutiva dos miRNAs nos metazoários.

**Palavras-chave:** RNAs não-codificadores, biogênese de miRNAs, modificação de RNAs, bioinformática, RNA-seq, m<sup>6</sup>A-seq.

## Abstract

MicroRNAs (miRNAs) are currently known as a large class of regulatory RNAs. During biogenesis, one strand ("arm") of the duplex miRNA is preferably selected as a functionally mature miRNA, while the other strand is usually degraded. Arm switching events occur when there is a change in the preference of the selected arm. This is reflected in the alteration of the arm identified as the most expressed and functional transcript, and has a dynamically variation between stages of development, types of tissues and even between species. Due to the particularities of their mechanism of action, arm switching grants miRNAs an increased regulatory potential by modifying the set of targets and, commonly, the biological processes they regulate. However, the molecular mechanisms that promote arm switching episodes still need to be better characterized. The identification of m<sup>6</sup>A methylations in pri-miRNAs from human cell lines reported in the literature provided evidence that this type of methylation could be involved in the arm switching process, mainly due to the change that m<sup>6</sup>A promotes in the secondary structure of RNAs. To test this hypothesis, we used zebrafish, a model vertebrate, to assess whether m<sup>6</sup>A methylation could be influencing the occurrence of arm switching in this species. This way, we compared the global expression profile of the miRNAs 5p and 3p arms by analyzing a wide collection of RNA-seq data from early periods of embryonic development and various adult zebrafish tissues obtained from the literature. Additionally, we applied the m<sup>6</sup>A methylation sequencing technique (m<sup>6</sup>A-seq) in the transcriptome to identify potential correlation between the methylation peaks present in the miRNAs and the identified arm switching events. This set of analyzes led us to the unprecedented identification of 14 arm switching events in zebrafish. Through the detailed study of these cases and their comparison with the events described in other species, we verified that arm switching events are mostly associated with the ontogenetic development of vertebrates. Additionally, our data suggest that the differential expression of miRNA isoforms (isomiRs) may be an important modulator in the prevalence of arms. We have also demonstrated, via target prediction and functional enrichment analyzes, that arm switching events in fact enhance the regulation exercised by miRNAs, by increasing the number of target genes, regulatory networks and biological functions with which they are associated. However, surprisingly, careful analysis of the m<sup>6</sup>A-seq data provided us with evidence that zebrafish pri-miRNAs are not effectively methylated, even though they contain m<sup>6</sup>A methylation sites in their sequences. The discrepancy between the results obtained in this work and those reported in humans and, more recently, in *Arabidopsis thaliana*, suggest

that selective pressures acting on the biogenesis of miRNAs may be occurring at different stages in these species. Still, considering recent data from the literature on the interaction between the m<sup>6</sup>A methylating reader protein and RNA molecules, we conjecture that the secondary structure of the pri-miRNAs may be a permissive factor in this modulation. In this context, we conclude that arm switching is a conserved phenomenon and that it is mostly associated with the ontogenetic development of organisms. We also found no evidence that m<sup>6</sup>A methylation interferes with the prevalence of miRNA arms in zebrafish, given their apparent absence in pri-miRNA transcripts of this species. The results obtained in this study bring new contributions to the understanding of regulatory mechanisms inherent to the functional activity of miRNAs. Additionally, they provide subsidies for future research interested in exploring the molecular mechanisms underlying the different m<sup>6</sup>A methylation rates and their interference in the miRNAs function and evolutionary dynamics in metazoans.

**Key-words:** non-coding RNAs, miRNA biogenesis, RNA modifications, bioinformatics, RNA-seq, m<sup>6</sup>A-seq.

## Lista de Figuras

|   |    |
|---|----|
| <b>Figura 1.</b> Via canônica de biogênese dos miRNAs .....   | 21 |
| <b>Figura 2.</b> Resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 1 ....   | 28 |
| <b>Figura 3.</b> Amostras utilizadas para identificação de miRNAs e detecção de padrões de arm switching.....                                       | 30 |
| <b>Figura 4.</b> Esquema de alinhamento para identificação dos braços 5p e 3p dos miRNAs .....  | 31 |
| <b>Figura 5.</b> Estratégia de obtenção dos valores de energia livre do duplex miRNA-5p/miRNA-3p .....  | 34 |
| <b>Figura 6.</b> Panorama geral dos dados de RNA-seq .....  | 35 |
| <b>Figura 7.</b> Eventos de arm switching identificados durante fases do desenvolvimento e em tecidos adultos de zebrafish .....                    | 37 |
| <b>Figura 8.</b> Distribuição de eventos de arm switching entre zebrafish, tilápia do Nilo, galinha e camundongo.....                               | 39 |
| <b>Figura 9.</b> Razão da expressão dos braços 5p e 3p de eventos de arm switching ocorrendo durante os períodos de desenvolvimento estudados ..... | 41 |
| <b>Figura 10.</b> Eventos de arm switching entre cópias parálogas dos miRNAs dre-mir-92a and dre-mir-153 .....                                      | 44 |
| <b>Figura 11.</b> Enriquecimento funcional dos alvos do dre-miR-27b expressos no embrião 24hpf.....   | 47 |
| <b>Figura 12.</b> Padrões de isomiR do dre-miR-153a-2 expressos no cérebro .....  | 50 |
| <b>Figura 13.</b> Análise das propriedades do duplex dos miRNAs .....   | 53 |
| <b>Figura 14.</b> Resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 2 ..  | 55 |
| <b>Figura 15.</b> Verificação do sucesso do enriquecimento da fração nuclear nas amostras de RNA.....   | 58 |
| <b>Figura 16.</b> Identificação de motifs m6A em pri-miRNAs .....   | 63 |
| <b>Figura 17.</b> Panorama geral dos dados de m6A-seq .....   | 65 |
| <b>Figura 18.</b> Funções biológicas da metilação m6A ao longo do mRNA .....  | 66 |

|  |    |
|--|----|
| <b>Figura 19.</b> Genes diferencialmente expressos e diferencialmente metilados entre o cérebro e embrião 24hpf..... | 67 |
| <b>Figura 20.</b> Lista de miRNAs com picos de metilação m6A detectados .....  | 68 |
| <b>Figura 21.</b> Picos de metilação associados à pri-miRNAs.....  | 70 |
| <b>Figura 22.</b> Picos de metilação falso positivos associados à pri-miRNAs.....                                    | 70 |
| <b>Figura 23.</b> Níveis de metilação m6A em diferentes estruturas secundárias .....                                 | 72 |

### Lista de Tabelas

|  |    |
|--|----|
| <b>Tabela 1.</b> Número de alvos preditos para os braços 5p e 3p dos miRNAs dre-mir-27b-1/-2, dre-mir-135b, dre-miR-137b-1/-2 e dre-mir-31 no cérebro e no embrião 24 hpf..... | 46 |
| <b>Tabela 2.</b> Reads representativas dos braços 5p e 3p nos tecidos embrião 24hpf e cérebro.....   | 50 |
| <b>Tabela 3.</b> Concentração e pureza das amostras de RNA nuclear .....   | 59 |
| <b>Tabela 4.</b> Qualidade das amostras de RNA nuclear .....   | 60 |

### Lista de abreviaturas

|   |  |
|---|--|
| 3'UTR – 3' <i>untranslated region</i>                                       | <i>hnRNPC</i> – <i>heterogeneous Nuclear Ribonucleoprotein C</i> |
| 5' UTR – 5' <i>untranslated region</i>                                      |  |
| Ago – Proteínas Argonautas  | <i>Hsc70</i> – <i>heat shock cognate 71 kDa protein</i>          |
| BAM – <i>Binary Alignment Map</i>   | <i>Hhp90</i> – <i>heat shock protein 90</i>                      |
| BSA - <i>Bovine Serum Albumin</i>   | Hpf – Horas pós-fecundação                                       |
| cDNA – DNA complementar   | <i>Ints12</i> – <i>Integrator Complex Subunit 12</i>             |
| CDS – <i>Coding Sequence Region</i>   | IP – imunoprecipitado  |
| <i>Coil</i> – <i>Coilin</i>   | Kb – kilobase  |
| Covid-19 – <i>Corona Virus Disease 2019</i>                                 | KEGG – <i>Kyoto Encyclopedia of Genes and Genomes</i>            |
| <i>Dgcr8</i> - <i>DGCR8 microprocessor complex subunit</i>                  | ncRNAs – <i>Non-coding RNAs</i>                                  |
| EAD – Ensino a distância  | <i>Nop10</i> – <i>NOP10 Ribonucleoprotein</i>                    |
| EDTA – Acido etilenodiamino tetra-acético                                   | <i>Notch</i> – <i>Notch signaling pathway</i>                    |
| <i>Eftud2</i> – <i>Elongation Factor Tu GTP Binding Domain Containing 2</i> | m <sup>5</sup> C – <i>5-methylcytosine</i>                       |
| <i>eIF3</i> – <i>eukaryotic initiation factor 3</i>                         | m <sup>6</sup> A – <i>N6-methyladenosine</i>                     |
| <i>eIF4F</i> – <i>eukaryotic initiation factor 4F</i>                       | miRNA – microRNA   |
| EMT - <i>epithelial-to-mesenchymal transition</i>                           | miRISC – <i>miRNA Induced Silence Complex</i>                    |
| FC – <i>fold-change</i>   | MREs – <i>microRNA Recognition Elements</i>                      |
| FPKM – <i>Fragments Per Kilebase Million</i>                                | mRNA – RNA mensageiro  |
| GO – <i>Gene Ontology</i>   | <i>Mttl3</i> - <i>Methyltransferase Like 3</i>                   |
| <i>hnRNPA2B1</i> – <i>heterogeneous nuclear ribonucleoproteins A2/B1</i>    | <i>Mttl14</i> – <i>Methyltransferase Like 14</i>                 |
|   | pré-miRNA – microRNA precursor                                   |
|   | pri-miRNA – microRNA primário                                    |
|   | <i>Prpf4</i> – <i>Pre-MRNA Processing Factor 4</i>               |

RLC – *RISC Loading Complex*

RPM – Rotações por minuto

rRNA – RNA ribossomal

RVC - *ribonucleoside vanadyl complexes*

*Snrpa1 - Small Nuclear Ribonucleoprotein*

*Polypeptide A*

*Snrpb2 – Small Nuclear Ribonucleoprotein*

*Polypeptide B2*

*SNRNP70 - Small Nuclear  
Ribonucleoprotein U1 Subunit 70*

*TGF-β – Transforming growth factor beta*

TMMP – *Target mediated miRNA  
protection*

TS – Estabilidade termodinâmica

*TUT4 - terminal uridylyltransferase 4*

*TUT7 - Terminal uridylyltransferase 7*

*Wnt – Wnt signaling pathway*

*Ythdc1 – YTH domain-containing protein 1*

## Sumário

|   |    |
|---|----|
| 1. Introdução.....  | 17 |
| 1.1. <i>MicroRNAs e seu papel na regulação gênica</i> .....   | 17 |
| 1.2. <i>A biogênese dos microRNAs e a origem dos eventos de arm switching</i> .....   | 19 |
| 1.3. <i>A metilação m<sup>6</sup>A e seu impacto na estrutura e biogênese de microRNAs</i> .....                                    | 22 |
| 1.4. <i>Hipótese do trabalho</i> .....  | 24 |
| 1.5. <i>O zebrafish como organismo modelo</i> .....   | 24 |
| 2. Objetivos.....   | 26 |
| 2.1. <i>Objetivo Geral</i> .....  | 26 |
| 2.2. <i>Objetivos Específicos</i> .....   | 26 |
| 3. Capítulo I: <i>Arm switching</i> e seu impacto no potencial regulatório dos microRNAs.....                                       | 27 |
| 3.1. <i>Material e métodos</i> .....  | 28 |
| 3.1.1. <i>Resumo do workflow</i> .....  | 28 |
| 3.1.2. <i>Obtenção dos dados de RNA-seq</i> .....   | 29 |
| 3.1.3. <i>Tratamento das amostras, identificação de microRNAs e caracterização dos perfis de expressão dos braços 5p e 3p</i> ..... | 30 |
| 3.1.4. <i>Normalização dos dados e identificação dos eventos de arm switching</i> .....   | 32 |
| 3.1.5. <i>Análise de predição de alvos e enriquecimento funcional</i> .....   | 32 |
| 3.1.6. <i>Identificação e caracterização dos padrões de isomiRs dos microRNAs sob eventos de arm switching</i> .....                | 33 |
| 3.1.7. <i>Cálculo de prevalência e incorporação diferencial dos braços 5p e 3p pela Ago</i> ....                                    | 33 |
| 3.2. <i>Resultados e discussão</i> .....  | 35 |
| 3.2.1. <i>Visão geral dos dados de RNA-seq</i> .....  | 35 |
| 3.2.2. <i>Eventos de arm switching estão majoritariamente associados ao desenvolvimento ontogenético dos vertebrados</i> .....      | 36 |
| 3.2.3. <i>Arm switching é um fenômeno conservado, porém seus casos são pontuais e espécie-específicos</i> .....                     | 42 |

|  |    |
|--|----|
| 3.2.4. <i>Arm switching amplia potencial regulatórios dos microRNAs em zebrafish, alterando os alvos e funções biológicas por eles controladas entre tecidos distintos</i> ..... | 45 |
| 3.2.5. <i>Expressão diferencial de isomiRs é capaz de alterar a isoforma representativa dos microRNAs</i> .....  | 49 |
| 4. Capítulo II: Metilação m <sup>6</sup> A e seu papel no mecanismo de atuação dos microRNAs .....   | 54 |
| 4.1. <i>Material e Métodos</i> .....   | 55 |
| 4.1.1. <i>Resumo do workflow</i> .....   | 55 |
| 4.1.2. <i>Cuidados com zebrafish e coleta de amostras</i> .....  | 55 |
| 4.1.3. <i>Enriquecimento de RNA nuclear</i> .....  | 56 |
| 4.1.4. <i>Imunoprecipitação m<sup>6</sup>A, preparação das bibliotecas de cDNA e m<sup>6</sup>A-seq</i> .....  | 59 |
| 4.1.5. <i>Tratamento dos dados de m<sup>6</sup>A-seq e identificação de picos de metilação</i> .....   | 61 |
| 4.2. <i>Resultados e discussão</i> .....   | 62 |
| 4.2.1. <i>Sequências motif de metilação m<sup>6</sup>A estão presentes nos pri-miRNAs de zebrafish</i> ..  | 62 |
| 4.2.2. <i>Visão geral dos dados de m<sup>6</sup>A-seq</i> .....  | 64 |
| 4.2.3. <i>Não há indícios da ocorrência efetiva de metilação m<sup>6</sup>A em pri-miRNAs de zebrafish, apesar do enriquecimento de sequências motif</i> .....                   | 68 |
| 4.2.4. <i>Hipóteses alternativas sobre a ocorrência de arm switching em zebrafish e possibilidades de pesquisas futuras</i> .....  | 74 |
| 5. <i>Conclusões e considerações finais</i> .....  | 77 |
| 5.1. <i>Conclusões</i> .....   | 78 |
| 5.2. <i>Considerações finais</i> .....   | 79 |
| 6. <i>Referências bibliográficas</i> .....   | 80 |
| 7. <i>Apêndice</i> .....   | 93 |
| 7.1. <i>Material suplementar</i> .....   | 93 |
| 7.2. <i>Atividades acadêmicas realizadas durante o período de doutoramento</i> .....   | 93 |

## 1. Introdução

### 1.1. MicroRNAs e seu papel na regulação gênica

MicroRNAs (miRNAs) são pequenos RNAs não codificadores (ncRNAs) de aproximadamente 22 nucleotídeos que regulam a expressão gênica pós-transcricionalmente. Primeiramente descritos como atuantes durante o desenvolvimento em *Caenorhabditis elegans* (Lee et al., 1993), possuem atualmente a reconhecida importância de participar do controle de virtualmente todos os processos celulares já caracterizados de animais, plantas e, inclusive, alguns vírus (Lee et al., 1993; Zhang et al., 2006; Jia et al., 2008). Estes processos perfazem desde o desenvolvimento ontogenético, proliferação, diferenciação e homeostasia celular, até respostas a estímulos ambientais e doenças tais como os diversos tipos de câncer (Qiu et al., 2012; Shenoy e Blalock, 2014; Reddy, 2015; Gebert e MacRae, 2018).

A via canônica de mecanismo de ação dos miRNAs ocorre pelo seu acoplamento com proteínas da família argonauta (*Ago*), formando o complexo miRISC (*miRNA Induced Silence Complex*), que por sua vez interage com transcritos de RNA mensageiro (mRNA) alvos, promovendo o silenciamento da expressão gênica. O pareamento dos miRNAs com seus alvos ocorre preferencialmente na região 3'UTR (*3' Untranslated region*) (Dexheimer e Cochella, 2020), embora também haja evidências de interação em éxons (Reckzo et al., 2012; Hausser et al., 2013) e na região 5'UTR (Devlin et al., 2010; Zhou e Rigoutsos, 2014).

Molecularmente, a interação miRNA-alvo ocorre pelo pareamento simples entre as bases nitrogenadas destas moléculas, geralmente apresentando complementaridade total nas plantas e parcial nos animais (embora haja exceções). Em animais, a complementaridade dos miRNAs com seus alvos se dá preferencialmente por uma sequência de sete nucleotídeos (nucleotídeos 2 a 8 na porção 5' do miRNA), chamada de sequência *seed*, embora interações menos comuns, não baseadas na *seed* também tenham sido reportadas (Chi et al., 2012; Clark et al., 2014) assim como foi demonstrado que a porção 3' de alguns miRNAs pode ser tão relevante quanto a região *seed* na detecção de seus alvos em animais (Broughton et al., 2016). Uma vez estabelecida a interação miRNA-alvo, a regulação da expressão gênica ocorre principalmente por meio da repressão da tradução – mediante à inibição do reconhecimento do 5'-Cap do mRNA alvo e desacoplamento prematuro do ribossomo – e pela degradação prematura do mRNA – via mecanismos de deadenilação (He e Hannon, 2004; Petersen et al., 2006; Mathonnet et al., 2007; Fabian et al., 2010).

Durante a biogênese dos miRNAs, uma importante etapa consiste no processamento da molécula precursora em um *duplex* de RNA que é reconhecido pelas proteínas *Ago* para formar o complexo miRISC (Lee et al., 2003). Neste reconhecimento, um dos braços do *duplex* é selecionado para incorporar o complexo miRISC, se tornando o miRNA funcional, enquanto o outro braço é liberado e rapidamente degradado (Griffiths-Jones et al., 2011). Para alguns miRNAs, um dos braços tende a acumular na célula, sendo preferencialmente selecionado pela *Ago*, enquanto o outro braço possui baixos níveis de acúmulo celular (Hutvagner, 2005; Griffiths-Jones et al., 2011). Já para outros miRNAs, verifica-se que as sequências precursoras podem ser processadas de maneira a produzir uma quantidade significativa de transcritos funcionais oriundos de ambos os braços (Hutvagner, 2005; Okamura et al., 2008; Griffiths-Jones et al., 2011). Entretanto, em alguns casos, o braço selecionado pela *Ago* para ser incorporado ao complexo miRISC e tornar-se o miRNA funcional é amplamente variável entre diferentes contextos biológicos, como por exemplo entre diferentes períodos do desenvolvimento, tecidos e condições fisiológicas. O resultado deste processo, que conduz à seleção diferencial de braço, é conhecido como *arm switching* (Ro et al., 2007; Wit et al., 2009; Griffiths-Jones et al., 2011; Ludwig et al., 2016; Kern et al., 2020).

Uma vez que o reconhecimento dos alvos pelos miRNAs ocorre devido à complementaridade entre suas sequências (especialmente entre a região *seed* dos miRNAs e os chamados *microRNA recognition elements* (MREs) ou sítios de interação, presentes no mRNA), a ocorrência de eventos de *arm switching* permite aos miRNAs modular um conjunto potencialmente diferente de alvos, a depender do braço preferencialmente expresso. De fato, análises de predição de alvos de miRNAs realizados em *Drosophila* e humano demonstram que os braços distintos de um mesmo miRNA tendem a regular transcritos diferentes de mRNAs. Adicionalmente, não só estes transcritos diferem entre si, como o papel biológico por eles desempenhados também tende a ser distinto (Marco et al., 2012). Desta maneira, os eventos de *arm switching* decorrentes da complexa biogênese dos miRNAs emergem como um poderoso mecanismo de regulação gênica, capaz de modular vias regulatórias envolvidas em múltiplas funções biológicas relevantes em diversos contextos normais e patológicos.

A ocorrência de *arm switching* já foi detectada em diversas espécies de vertebrados, tais como em humano (*Homo sapiens*), camundongo (*Mus musculus*), coelho (*Oryctolagus cuniculus*) e galinha (*Gallus gallus*) (Ro et al., 2007; Glazov et al., 2008; Li et al., 2011; Guo et al., 2015); em espécies de invertebrados, como a mosca doméstica (*Drosophila melanogaster*) e o besouro marrom (*Tribolium castaneum*) (Griffiths Jones et al., 2011); entre diferentes

espécies de nematodas (Wit et al., 2009); e ainda ocorrentes durante o desenvolvimento ontogenético de abelhas (*Apis mellifera*) (Freitas et al., 2017). Recentemente, nosso grupo de pesquisa descreveu *arm switching* em peixes a partir da investigação da ampla caracterização do miRNoma da tilápia do Nilo (*Oreochromis niloticus*) (Pinhal et al., 2018). Estes dados sugerem que a ocorrência de *arm switching* é um evento comum nos metazoários.

## 1.2. A biogênese dos microRNAs e a origem dos eventos de *arm switching*

Apesar do *arm switching* ser um fenômeno conhecido, os mecanismos moleculares que promovem a seleção diferencial dos braços 5p e 3p pela *Ago* ainda não foram muito bem elucidados. Atualmente, a instabilidade termodinâmica das extremidades do *duplex* de miRNA, bem como a identidade do nucleotídeo 5' de cada braço são reconhecidos como os principais fatores responsáveis por guiar a seleção do braço a ser incorporado na *Ago* (Khvorova et al., 2003; Tsai et al., 2016; Kim et al., 2020). Entretanto, ainda é necessário um maior conhecimento sobre os mecanismos capazes de promover alterações nas propriedades destes fatores, permitindo a ocorrência da seleção alternativa dos braços de miRNAs.

Durante a biogênese dos miRNAs uma longa molécula conhecida como miRNA primário (pri-miRNA) é transcrita no núcleo pela *RNA polimerase II* como transcritos individuais, em forma de clusters de miRNAs (transcritos policistrônicos) ou a partir de íntrons de mRNAs (Lee et al. 2004; Treiber et al., 2018). A *RNase* do tipo III denominada *Drosha* reconhece a estrutura em formato de *hairpin* do pri-miRNA (~ 33 nucleotídeos) e cliva esta região 11 nucleotídeos acima de sua junção basal e 22 nucleotídeos abaixo de sua região apical (Alarcón et al., 2015a), dando origem ao miRNA precursor (pré-miRNA) de aproximadamente 60 nucleotídeos, contendo dois nucleotídeos sobressalentes na posição 3' característicos de clivagens por *RNases* do tipo III (Lee et al., 2003; Ha e Kin, 2014). Desta maneira, as junções basal e apical do pri-miRNA servem como principais referências para determinar o sítio de clivagem da *Drosha* (Zeng et al., 2008; Ma et al., 2013).

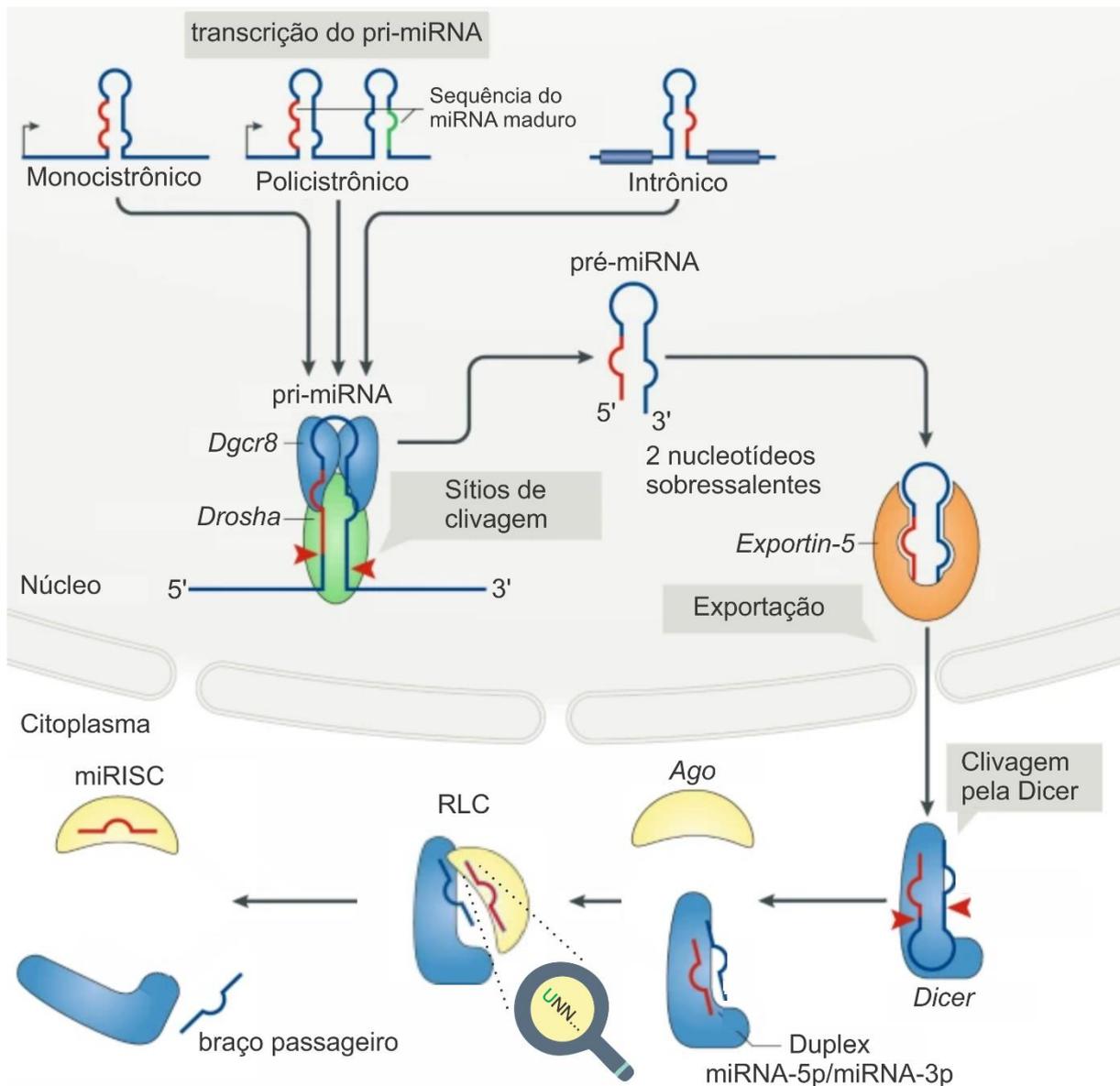
Após este processamento, a molécula de pré-miRNA é exportada do núcleo para o citoplasma pela proteína *Exportin-5* e, uma vez no citoplasma, a *RNase* do tipo III *Dicer* reconhece os dois nucleotídeos sobressalentes (Okada et al., 2009) e cliva o pré-miRNA em sua porção apical, aproximadamente 22 nucleotídeos acima de sua porção basal, previamente delimitada pela *Drosha*, produzindo a molécula de RNA *duplex* miRNA-5p/miRNA-3p,

contendo dois nucleotídeos sobressalentes em cada lado nas extremidades 3' (Ketting et al., 2001). Assim, o miRNA-5p é o braço proveniente da porção 5' enquanto o miRNA 3p é o braço proveniente da porção 3' do pré-miRNA.

Após o corte realizado pela *Dicer* as proteínas *Ago* se associam ao complexo riboproteico formado pela por ela e pelo *duplex*, originando uma estrutura estável conhecida como *Risc Loading Complex* (RLC) (Treiber et al., 2018). Durante a formação do RLC, a *Ago* ocupa a posição pertencente à região *loop* do pré-miRNA no período anterior à clivagem. Desta maneira, o domínio PAZ da proteína *Ago* fica localizado próximo aos nucleotídeos sobressalentes da porção 3' do miRNA-5p, recém-formados pelo corte da *Dicer*, enquanto o domínio PAZ da proteína *Dicer* está ancorada aos nucleotídeos sobressalentes da porção 3' do miRNA-3p (Treiber et al., 2018). Uma vez que este complexo está estável, ocorre a transferência do duplex para a *Ago*, graças às modificações estruturais promovidas pela chaperona *Hsc70-Hsp90*, formando o complexo pré-RISC (Kwak e Tomari, 2012).

As modificações estruturais da *Ago* durante esta etapa promovem uma tensão contorcional no duplex, ocasionando a quebra do pareamento de sua porção terminal 5' mais instável (Kawak e Tomari, 2012). Uma vez que este pareamento é quebrado, a porção liberada tende a se ligar ao domínio MID da *Ago*. Simultaneamente, a modificação estrutural da *Ago* força seu domínio N-terminal contra sua outra extremidade, promovendo a quebra do pareamento de diversos pares de base desta região (Kawak e Tomari, 2012). De maneira interessante, o domínio MID apresenta afinidades diferentes, a depender da identidade do nucleotídeo inicial do miRNA, possuindo maior afinidade por Uracilas, seguido por Adeninas, Guaninas e, por último, Citosinas (Krol et al., 2004; Suzuki et al., 2015). Assim, a associação destes dois fatores (instabilidade relativa das porções 5' e identidade do nucleotídeo inicial do miRNA), em paralelo, são fundamentais para determinar a porção 5' de qual braço, 5p ou 3p, será ligado ao domínio MID da *Ago* (Suzuki et al., 2015). Devido aos *mismatches* comumente presentes na porção interior do duplex, as quebras no pareamento de ambas as porções terminais são suficientes para desestabilizar toda a estrutura, separando os braços 5p e 3p. O braço cuja porção 5' está associada ao domínio MID permanece no complexo, dando origem à miRISC funcional e se tornando o braço canônico, enquanto o outro braço, o braço passageiro, é liberado e rapidamente degradado. O miRNA proveniente do braço 5p é chamado de miRNA-#-5p (e.g., miR-1-5p), enquanto o miRNA proveniente do braço 3p é chamado de miRNA-#-3p (e.g., miR-499-3p).

A Figura 1 apresenta um resumo do processo canônico de biogênese de miRNAs.



**Figura 1. Via canônica de biogênese dos miRNAs.** O pri-miRNA é transcrito a partir de genes individuais, de regiões policistrônicas ou a partir de íntrons de mRNAs. O pri-miRNA então é processado pelo complexo *Drosha-Dgcr8*, dando origem a molécula de pré-miRNA. A proteína *Exportin-5* exporta o pré-miRNA para o citoplasma e a enzima *Dicer* cliva o pré-miRNA dando origem ao duplex miRNA-5p/miRNA-3p. As proteínas *Ago* se ligam ao complexo Riboproteico formando dando origem ao RLC e alterações estruturais promovem a separação dos braços dos miRNAs. As propriedades das extremidades do duplex (instabilidade e nucleotídeo inicial) determinam qual braço será incorporado pela *Ago*. O braço associado às proteínas *Ago* se torna o miRNA funcional e forma o complexo miRISC enquanto o outro braço é degradado. Figura modificada de Treiber et al. (2018).

Desta maneira, o corte do pri-miRNA promovido pela enzima *Drosha*, determina a instabilidade termodinâmica de uma extremidade, bem como a identidade de seu nucleotídeo 5', enquanto guia a posição do corte realizado pela *Dicer*, que determinará as características da outra extremidade do duplex, sendo, portanto, a etapa mais crítica da biogênese de um miRNA (Wu et al., 2009). Alterações na estrutura do pri-miRNA poderiam ocasionar modificações nas posições dos cortes promovidos pela *Drosha* e pela *Dicer*, acarretando no processamento de isoformas variantes de miRNAs maduros, conhecidos como isomiRs. IsomiRs são moléculas variantes de miRNAs que diferem do miRNA representativo devido à modificações em suas sequências por alterações na clivagem da região 5' ou 3', como inserções e deleções, e inclusive substituições como adições não-modelo na porção 3' (Morin et al., 2008; Neilsin et al., 2012; Guo et al., 2014).

A ocorrência de isomiRs permite a formação de *duplexes* com estruturas e propriedades distintas que, em última instância promoveriam eventos de *arm switching*. De fato, Ma et al. (2013) demonstram que sítios de clivagem alternativos de pri-miRNAs realizados pela *Drosha* podem ser manipulados, possibilitando a indução da clivagem de sítios específicos por meio da alteração da estrutura do pri-miRNA mediante mutações de inserção e deleção que promovem a aproximação ou distanciamento das junções basal e apical do pri-miRNA. Adicionalmente, Wu et al. (2009) demonstram que este processamento alternativo é capaz de produzir isomiRs distintos que possibilitam a seleção diferencial dos braços 5p e 3p pela *Ago* e a ocorrência de *arm switching*. Estes dados indicam que o estudo de modificações de RNA ocorrentes em pri-miRNAs, capazes de alterar sua estrutura secundária poderiam fornecer subsídios para um melhor entendimento dos processos moleculares potencialmente envolvidos na gênese de eventos de *arm switching*.

### 1.3. A metilação m<sup>6</sup>A e seu impacto na estrutura e biogênese de microRNAs

Uma ampla gama de modificações de bases nitrogenadas de RNA foi caracterizada nos últimos anos, como as metilações m<sup>6</sup>A e m<sup>5</sup>C, inosina e pseudouridina (Liu e Pan, 2016). Dentre os diversos tipos de modificações de RNA, a metilação m<sup>6</sup>A é a mais frequente entre os eucariotos, atuando em uma ampla variedade de RNAs (Bokar, 2005). Esta metilação ocorre no grupo N6-metil de adeninas (A), devido à atuação das proteínas *Mttl3* e *Mttl14* em sequências RRACH (R = A/G e H = A/C/U) e é capaz de alterar a estrutura secundária dos RNAs (Liu e Pan, 2016).

Uma metilação m<sup>6</sup>A em uma adenina pareada é capaz de promover a desestabilização do pareamento, dependendo da localização relativa de seu sítio com o final do duplex (Liu et al., 2017). De maneira oposta, caso esta metilação ocorra em uma adenina que está adjacente a uma bolha 5', ela pode intensificar a estabilidade de seu pareamento (Liu et al., 2018). Isto ocorre porque esta bolha promove a flexibilização necessária para permitir o encaixe ótimo entre o grupo metil e a porção 3' do nucleotídeo adjacente (Liu et al., 2018). Com isso, a metilação m<sup>6</sup>A pode proporcionar diversas modificações à estrutura secundária final do RNA, promovendo alterações nas interações RNA-RNA, RNA-DNA e RNA-proteína.

Alarcón et al. (2015a) demonstra *in vitro* em células humanas que durante a biogênese dos miRNAs, o pri-miRNA sofre metilações m<sup>6</sup>A decorrentes da atividade da proteína *Mtll3* em sítios GGAC. Notadamente, estas ocorrem em regiões próximas à junção basal, principal região de referência para determinar o sítio de clivagem da *Drosha* (Ha e Kin, 2014). Estas metilações são reconhecidas pela proteína *reader hnRNPA2B1*, e servem como ponte entre o pri-miRNA metilado e o complexo microprocessador composto pelas proteínas *Dgcr8* e *Drosha* (Alarcón et al., 2015b).

Alarcón et al. (2015a) também detectaram que os pri-miRNAs de *H. sapiens* podem apresentar mais de um sítio de metilação e que nem todos estes sítios apresentam o mesmo enriquecimento m<sup>6</sup>A. Outros autores, no caso investigando alguns mRNAs, detectaram que a metilação m<sup>6</sup>A por vezes apresenta perfis tecido- e célula-específicos (Meyer et al., 2012; Chen et al., 2015), e que transcritos apresentando tais metilações atuam em processos biológicos chave para o correto funcionamento daquele tipo celular, tais como manutenção de células tronco e diferenciação neuronal (Chen et al., 2015).

Adicionalmente à metilação m<sup>6</sup>A, outras modificações de RNA como uridilação e e modificação *A-to-I* também foram identificadas ocorrentes durante a biogênese dos miRNAs (Paul et al., 2017; Kim et al., 2020). De fato, eventos de *arm switching* associados a estas modificações foram previamente descritos na literatura (Iizasa et al., 2010; Li et al., 2018; Kim et al., 2020), demonstrando o potencial das modificações de RNA em alterar a razão do acúmulo dos braços 5p e 3p e promover eventos de *arm switching*.

#### 1.4. Hipótese do trabalho

De acordo com estes dados, a hipótese avaliada neste trabalho é que a metilação m<sup>6</sup>A dos pri-miRNAs, além de servir como guia para o acoplamento da proteína *Drosha*, também auxilia no modelamento de sua estrutura secundária em forma de *hairpin*, ditando à *Drosha* onde o corte do pri-miRNA será realizado. Essa metilação, portanto, seria capaz de promover alterações na instabilidade das porções terminais 5' e 3' do duplex miRNA-5p/miRNA-3p, bem como na assinatura de seus nucleotídeos 5'. Adicionalmente, é provável que picos diferenciais de metilação m<sup>6</sup>A tecido- e célula-específicos também ocorram nos pri-miRNAs, de maneira a direcionar a maturação do miRNA para o braço atuante no controle de genes responsáveis pelo desenvolvimento e manutenção do tecido em questão, promovendo os eventos de *arm switching*.

#### 1.5. O *zebrafish* como organismo modelo

Nas últimas duas décadas, o *zebrafish* (*Danio rerio*) se consolidou como um excelente modelo biológico para estudos genéticos e relacionados ao desenvolvimento. Este organismo consiste num dos modelos biológicos com melhor qualidade de anotações no genoma completo sequenciado (Ensembl – vGRCZ11), igualando-se ao genoma de modelos como o camundongo e humano, fornecendo assim, valiosas informações para análises estruturais e funcionais do genoma desta espécie e estudos de genômica comparativa.

Um dos principais atributos que levaram à popularização do *zebrafish* como organismo modelo é a possibilidade de sua utilização no estudo de diversas doenças humanas, como doenças neurológicas (Babin et al., 2014; Saleem e Kannan, 2018), distrofia muscular (Guyon et al., 2007; Li et al., 2017), regeneração do tecido cardíaco (Beffagna, 2019; Chavéz et al., 2020), insuficiência renal (Outtandy et al., 2019), doenças imunológicas (Novoa e Figueiras, 2012), câncer (Amatruda et al., 2002; Hason e Bartunek, 2019) e até odontologia (Huyseune et al., 2004; Wise e Stock, 2010). De fato, 70% dos genes humanos associados a doenças possuem similaridades funcionais aos ortólogos de *zebrafish* (Santoriello et al., 2012).

Outros atributos favoráveis ao uso do *zebrafish* como modelo biológico incluem seu pequeno tamanho corporal, relativo baixo custo de manutenção, grande capacidade reprodutiva e rápido desenvolvimento, permitindo a análise de indivíduos em larga escala (Kimmel et al.,

1989). Adicionalmente, a fecundação externa e o corpo transparente do embrião oferecem grandes vantagens ao estudo *in vivo* quando comparado a outros modelos de vertebrados, permitindo a observação direta de diversos padrões morfológicos, e a validação experimental de interações entre produtos gênicos (ex., RNA ou proteínas) via ensaios de fluorescência (Bakkers, 2011).

O *zebrafish* também tem sido amplamente utilizado no estudo de diversas classes de ncRNAs, como miRNAs (Zayed et al., 2019; Ji et al., 2020; Ahkin Chin Tai et al., 2020), long non-coding RNAs (lncRNAs) (Hu et al., 2018; Mathew et al., 2019; Valenzuela-Muñoz et al., 2019), piwi-RNAs (piRNAs) (Zhou et al., 2010; Kaaij et al., 2013) e RNAs circulares (circRNAs) (Shen et al., 2017; Sharma et al., 2019). Essas investigações trouxeram inúmeras contribuições para a compreensão da dinâmica evolutiva e funcionalidade dos miRNAs nos vertebrados.

A ocorrência de *arm switching* já foi documentada nesta espécie, em um estudo avaliando o envolvimento dos miRNAs durante o desenvolvimento gonadal, com um evento pontual da troca da prevalência dos braços 5p e 3p do miR-202 ocorrendo entre 12 e 24 spf (semanas pós-fertilização) durante o desenvolvimento dos testículos (Pressauler et al., 2017). Adicionalmente, outros estudos também validaram a ocorrência de metilações m<sup>6</sup>A e seus papéis biológicos no transcriptoma de *zebrafish*, embora estes trabalhos não tenham avaliado sua ocorrência em miRNAs (Zhao et al., 2017; Xia et al., 2018; Kontur et al., 2020). Desta maneira, o *zebrafish* se apresenta como um organismo modelo sólido para o estudo funcional e comparativo destes elementos regulatórios.

## **2. Objetivos**

### **2.1. Objetivo Geral**

Avaliar a potencial associação entre a metilação de RNA m<sup>6</sup>A e a biogênese e o processo de *arm switching* de microRNAs em *zebrafish*

### **2.2. Objetivos Específicos**

- Identificar os eventos de *arm switching* presentes no desenvolvimento e tecidos adultos de *zebrafish*
- Comparar os eventos identificados em *zebrafish* com os detectados em outras espécies modelo de vertebrados
- Avaliar as funções biológicas associadas aos eventos de *arm switching* identificados e seu impacto na regulação gênica
- Avaliar a ocorrência de isormiRs e seu impacto no *arm switching*
- Identificar os picos de metilação m<sup>6</sup>A presentes nos miRNAs de *zebrafish*
- Correlacionar os perfis de metilação m<sup>6</sup>A e os eventos de *arm switching* identificados

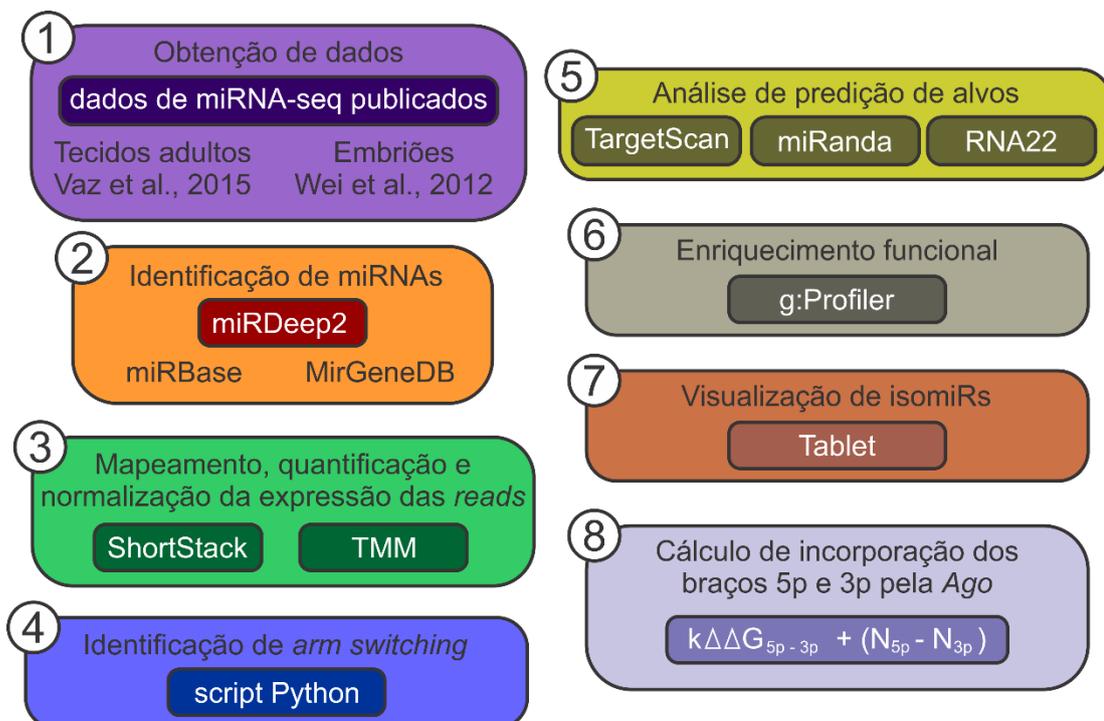
### **3. Capítulo I:**

*Arm switching e seu impacto no potencial regulatório dos microRNAs*

### 3.1. Material e métodos

#### 3.1.1. Resumo do workflow

A Figura 2 abaixo, apresenta um breve resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 1. Cada etapa será detalhadamente descrita nos tópicos a seguir.



**Figura 2. Resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 1.**

1) Dados de RNA-seq de diversos tecidos e períodos-chave do desenvolvimento inicial do *zebrafish* foram obtidos da literatura. 2) Estes dados foram tratados e os miRNAs identificados utilizando-se a ferramenta *MirDeep2*. 3) As amostras foram Mapeadas e quantificadas utilizando o módulo *Unique* da ferramenta *Shortstack* normalizadas e a normalização da expressão dos braços 5p e 3p foi feita pela ferramenta *TMM*. 4) Os eventos de *arm switching* foram identificados utilizando um *script* em Python por nós desenvolvido. 5) Foi realizada a análise de predição de alvos, nos tecidos de interesse, dos miRNAs com eventos de *arm switching* identificados. 6) A análise de enriquecimento funcional dos alvos foi realizada para avaliar as principais funções biológicas controladas pelos braços 5p e 3p dos miRNAs. 7) Os isomiRs expressos em cada braço nos diferentes tecidos foram visualizados utilizando a ferramenta *Tablet*, para avaliar suas expressões relativas. 8) O cálculo de incorporação dos braços 5p e 3p pela *Ago* foi realizado utilizando-se as isoformas representativas de cada tecido para avaliar como os diferentes isomiRs poderiam interferir no *arm switching*.

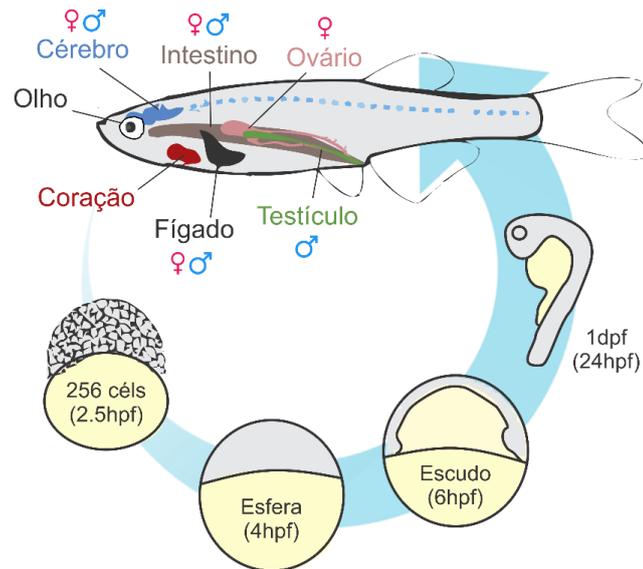
### 3.1.2. Obtenção dos dados de RNA-seq

Para a identificação dos eventos de *arm switching* foram utilizados dados brutos de RNA-seq disponíveis na literatura de diversos tecidos adultos e períodos de desenvolvimento de *zebrafish* (Figura 3). Desta maneira, foram refeitas no presente estudo todas as análises necessárias para normalização dos dados dos diferentes estudos prévios, incluindo o reprocessamento dos transcriptomas de miRNAs, e a identificação e o cálculo de expressão diferencial dos braços 5p e 3p em todos os contextos biológicos avaliados comparativamente.

Os dados referentes às amostras de tecidos adultos foram obtidos do banco de dados SRA da plataforma NCBI, estudo SRP041544 (Vaz et al., 2015), que realizou uma ampla análise exploratória de miRNAs, sem contudo avaliar a prevalência de braços 5p e 3p. Esses dados, gerados na plataforma *Illumina HiSeq 2000*, englobaram RNA-seq de machos e fêmeas referentes às amostras de cérebro, intestino e fígado, além de *pools* de olho e coração, e amostras de ovário e testículo de cada sexo (n = 3).

Os dados referentes às amostras de períodos embrionários foram obtidos do do banco de dados SRA, estudo SRP028895 (Wei et al., 2012), que contempla dados referentes a quatro períodos-chave do desenvolvimento ontogenético dos vertebrados, incluindo 256 células (2.5 horas pós-fecundação – hpf; n = 1), esfera (4 hpf, n = 2), escudo (6 hpf, n = 1) e 24 hpf (n = 1), gerados na plataforma *Illumina Genome Analyzer II*. O estágio de 256 células é caracterizado por possuir majoritariamente RNAs de origem materna, os quais são, no estágio de esfera, degradados e gradativamente substituídos por RNAs transcritos pelo próprio embrião. Durante o estágio de escudo, ocorre a formação das três camadas germinativas, ao passo que às 24 hpf, em *zebrafish*, a maioria dos órgãos encontram-se em estágio final de formação (Wei et al., 2012).

Desta maneira, a utilização dos dados provenientes destes trabalhos possibilitam a execução de uma análise detalhada sobre o perfil de expressão e eventos de *arm switching* de miRNAs, ao longo de pontos chaves do desenvolvimento inicial, englobando desde as fases iniciais de diferenciação tecidual até a completa formação dos órgãos.



**Figura 3. Amostras utilizadas para identificação de miRNAs e detecção de padrões de *arm switching*.** Dados de tecidos adultos obtidos de Vaz et al. (2015) e dados de períodos embrionários obtidos de Wei et al. (2012).

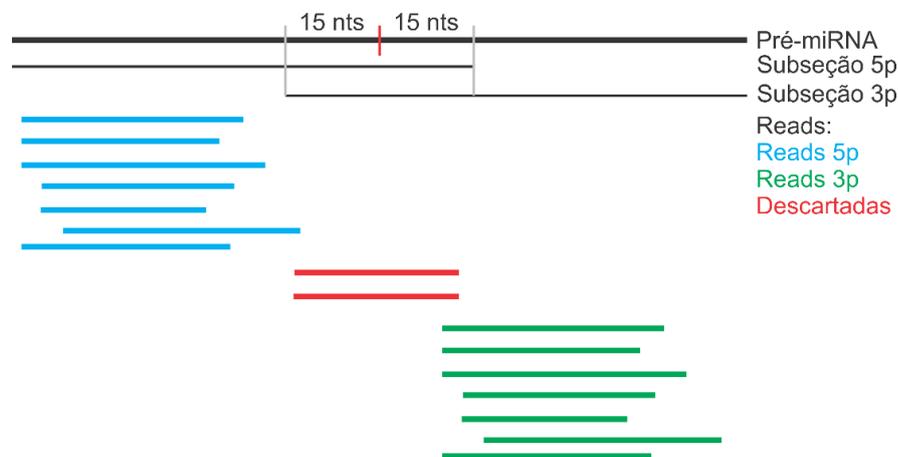
### 3.1.3. Tratamento das amostras, identificação de microRNAs e caracterização dos perfis de expressão dos braços 5p e 3p

Uma vez que os dados de RNA-seq foram obtidos, as amostras foram avaliadas quanto à qualidade pela ferramenta *FastQC* (v0.11.5; Wingett e Andrews, 2018) para assegurar a integridade dos dados utilizados. Posteriormente, as *reads* contendo sequências de adaptadores foram removidas utilizando-se a ferramenta *cutadapt* (v1.14; Martin, 2011) e as sequências remanescentes foram convertidas do formato fastq para o formato fasta utilizando-se a ferramenta *FASTX-Toolkit* (v2.8.1). Finalmente, as *reads* foram filtradas por comprimento (16 - 26 nts) e complexidade ( $\geq 3$  nucleotídeos distintos) e as *reads* alinhadas com outros tipos de ncRNAs como rRNAs e tRNAs foram removidas, utilizando-se o módulo de filtragem da ferramenta *UEA small RNA Workbench* (v3.2, release 19; Stocks et al., 2012).

As *reads* filtradas foram então agrupadas em *reads* únicas utilizando-se um *script* em linguagem *Python* por nós desenvolvido e foram mapeadas no genoma de *zebrafish* (vGRCZ10) utilizando o *script perl, Mapper* (v2.0.0.7; An et al., 2013) e os miRNAs foram identificados utilizando a ferramenta *miRDeep2* (v2.0.0.7; An et al., 2013). Para isso, sequências de pré-miRNAs e miRNAs maduros de *zebrafish* e miRNAs maduros de outras espécies foram obtidas dos bancos de dados *miRBase* (v22; Kozomara e Griffiths-Jones, 2011) e *MirGeneDB* (v2.0;

Fromm et al., 2019) e utilizadas como referência, permitindo-se o limite máximo de 1 *mismatch* fora da região *seed*. *Reads* que alinham com um dos miRNAs de referência foram descritos como miRNAs conhecidos, caso contrário foram preditos como miRNAs novos. Os miRNAs preditos como novos foram ainda filtrados pelo seu *score* (*miRDeep2 score*  $\geq 5$ ) e estrutura secundária em formato de hairpin de sua molécula precursora predita (*randfold p-value*  $< 0.05$ ).

Uma vez que os miRNAs foram identificados, os braços 5p e 3p de cada miRNA foram caracterizados utilizando-se um *script* em *Python* por nós desenvolvido. Para isso, as sequências de pré-miRNAs obtidas foram divididas em duas subseções com 30 nucleotídeos de sobreposição. As *reads* filtradas foram então mapeadas nestas subseções utilizando-se a ferramenta *Shortstack* (v1.1.2; Axtell, 2013) no modo *Unique* e classificadas como 5p e 3p. As *reads* que alinham exclusivamente na região de sobreposição foram descartadas por não corresponderem a sequências de miRNAs maduros funcionais (Figura 4).



**Figura 4. Esquema de alinhamento para identificação dos braços 5p e 3p dos miRNAs.** Traços azuis representam as *reads* alinhadas na subseção 5p, identificadas como miRNAs 5p. Traços verdes representam as *reads* alinhadas na subseção 3p, identificadas como miRNAs 3p. Traços vermelhos representam as *reads* que alinham de maneira exclusiva na região de intersecção das subseções 5p e 3p e foram descartadas.

O durante o mapeamento das *reads*, o módulo *Unique* da ferramenta *Shortstack* alinha primeiramente as *reads* mapeadas em apenas uma subseção. Posteriormente, as *reads* que podem ser mapeadas em mais de uma subseção são distribuídas entre elas de acordo com a proporção das *reads* únicas mapeadas, resolvendo parcialmente o problema de pareamento em múltiplos loci, permitindo uma quantificação mais precisa de miRNAs com cópias parálogas e a possibilitando a verificação de eventos de *arm switching* entre elas.

### 3.1.4. Normalização dos dados e identificação dos eventos de *arm switching*

A expressão das *reads* dos braços 5p e 3p de cada amostra foi normalizada com o pacote *edgeR* (R Bioconductor) utilizando o método TMM (*trimmed mean of M values*; Robinson e Oshlack, 2010). Com os dados normalizados, a razão das expressões dos braços 5p e 3p dos miRNAs identificados foi obtida em cada tecido para a identificação dos eventos de *arm switching*. Para isso, foram considerados os casos nos quais as razões 5p/3p possuíam um valor de *fold-change*  $\geq 2$  e  $\leq -2$  em pelo menos dois tecidos nos quais ambos os braços possuíam pelo menos 10 *raw reads* expressas. Ao final, o teste estatístico de Fisher ( $p\text{-value} \leq 0.05$ ) foi aplicado para filtrar os eventos de alta confiabilidade e garantir a reprodutibilidade dos dados. Tais análises foram realizadas seguindo o mesmo protocolo de identificação de eventos de *arm switching* por nós aplicado na tilápia do Nilo (Pinhal et al., 2018).

### 3.1.4. Análise de predição de alvos e enriquecimento funcional

A predição de alvos foi realizada utilizando a união dos resultados das ferramentas *miRanda* (Enright et al., 2003; [www.microrna.org](http://www.microrna.org)), *RNA22* (v2, Miranda et al., 2006; <https://cm.jefferson.edu/rna22/>) e *TargetsScan* (Agarwal et al., 2015; [www.targetscan.org](http://www.targetscan.org)). Esta metodologia foi executada por retornar a melhor performance no balanço entre especificidade e sensibilidade da predição, conforme demonstrado por nosso grupo de pesquisa (Oliveira et al., 2017). Os resultados foram posteriormente filtrados pela presença de expressão dos alvos identificados nos tecidos de interesse. Para isso dados de RNA-seq da literatura foram obtidos e o FPKM (*Fragments Per Kilobase Million*) dos mRNAs alvo foi calculado, tendo sido descartados alvos com valores de FPKM  $< 5$  (Soh et al., 2015; Staton et al., 2017; Pinhal et al., 2018).

Os alvos preditos remanescentes das filtragens foram submetidos à análise de enriquecimento funcional para avaliação do perfil regulatório dos miRNAs 5p e 3p nos tecidos avaliados. Para isso, foi utilizada a ferramenta *g:Profiler* (Reimand et al., 2016), utilizando-se a função *multiquery* e o método de significância *g:SCS threshold* ( $p\text{-value} \leq 0.05$ ), buscando pelos termos associados a Processos Biológicos, Componentes Celulares, Funções Moleculares (Gene Ontology; Harris et al., 2004) e Vias Biológicas (KEGG; Kanehisa et al., 2000).

### 3.1.6. Identificação e caracterização dos padrões de isomiRs dos microRNAs sob eventos de *arm switching*

Para analisar a ocorrência e os padrões de isomiRs presentes nos miRNAs que exibiram *arm switching*, os arquivos BAM, obtidos durante a execução da ferramenta *Shortstack*, foram carregados na ferramenta *Tablet* (Milne et al., 2013) permitindo a visualização das *reads* presentes nos tecidos de interesse. Para isso, o arquivo contendo as subseções dos pré-miRNAs foi utilizado como referência.

Uma vez que os arquivos foram carregados, os padrões de isomiRs e as *reads* canônicas de cada tecido foram obtidas, permitindo identificar as principais modificações pós-transcricionais e traçar os principais cortes realizados pelas enzimas *Drosha* e *Dicer*, nos pri-miRNA e pré-miRNA respectivamente, durante o processamento e maturação dos miRNAs, bem como suas diferenças entre os tecidos analisados.

### 3.1.7. Cálculo de prevalência e incorporação diferencial dos braços 5p e 3p pela *Ago*

Foi realizada a montagem conceitual dos duplex miRNA-5p/miRNA-3p, os quais são incorporados pela *Ago* para maturação final do miRNA e formação do complexo RISC. Para isso, as *reads* 5p e 3p mais expressas de cada tecido foram usadas como referência para a montagem do duplex conceitual, contendo 2 nucleotídeos 3' sobressalentes em cada extremidade (Krol et al., 2004; Suzuki et al., 2015).

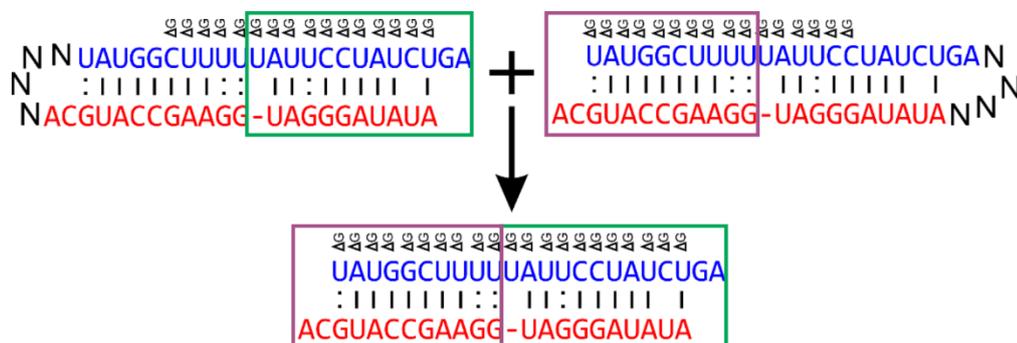
Uma vez que o braço do duplex a ser incorporado ao complexo RISC é majoritariamente selecionado de acordo com a instabilidade relativa das porções 5' – favorecendo o braço com maior instabilidade nesta região – e com a identidade do nucleotídeo inicial da porção 5' – havendo uma preferência por Uracilas, seguido por Adenina, Guanina e Citosina, respectivamente (Krol et al., 2004; Suzuki et al., 2015; Kim et al., 2020) – o cálculo da prevalência de seleção dos braços pelas proteínas Argonautas foi realizado baseado nestes fatores para verificar se as propriedades dos duplex gerados nos diferentes tecidos reproduzem as prevalências detectadas nos dados de RNA-seq. Para isso, foi utilizada a seguinte fórmula, descrita por Suzuki et al., 2015:

$$k\Delta\Delta G_{5p-3p} + (N_{5p} - N_{3p})$$

no qual  $k$  é a constante da estabilidade termodinâmica (TS), no valor de  $k = 0,66005$ .  $\Delta\Delta G_{5p-3p}$  é o valor de TS relativo entre os braços ( $\Delta\Delta G_{5p-3p} = \Delta G_{5p} - \Delta G_{3p}$ ).  $N_{5p(3p)}$  é o valor dado ao primeiro nucleotídeo do respectivo braço, sendo  $N_{(U)} = 2,63963$ ;  $N_{(A)} = 1,05674$ ;  $N_{(G)} = 0,37825$  e  $N_{(C)} = 0$ . Para o cálculo do  $\Delta\Delta G_{5p-3p}$  foi considerada a energia livre do dinucleotídeo inicial de cada braço uma vez que esse é o intervalo que melhor representa a TS relativa entre os braços (Krol et al., 2004; Suzuki et al., 2015).

Para o cálculo da energia livre a ferramenta utilizada foi a ferramenta *mfold* (Zuker, 2003). Primeiramente, esta ferramenta calcula a estrutura secundária do pré-miRNA. Assim que a estrutura mais estável é identificada, a energia livre de cada par de base formado é calculada de acordo com o método *nearest neighbor*, levando em consideração todos os elementos de desestabilização, como loops internos e bolhas.

Como o programa é capaz de calcular a estabilidade apenas de estruturas secundárias, não sendo capaz de calcular tais valores para duplex de RNAs, duas pré-estruturas de *hairpins* foram artificialmente sintetizadas e posteriormente fundidas, para a obtenção dos valores de energia livre do duplex (Figura 5; Krol et al., 2004). Com as pré-estruturas fundidas e o duplex conceitual formado, a fórmula descrita acima é aplicada e o valor obtido demonstra a razão de incorporação no complexo miRISC e maturação dos braços 5p e 3p. Valores positivos indicam prevalência do braço 5p enquanto valores negativos indicam prevalência dos braços 3p.



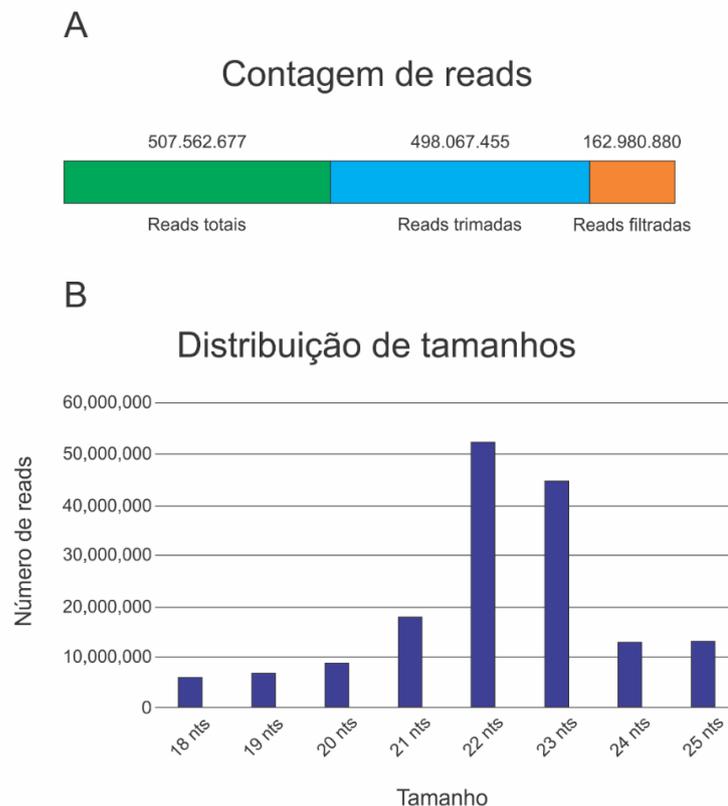
**Figura 5. Estratégia de obtenção dos valores de energia livre do duplex miRNA-5p/miRNA-3p.**

Dois hairpins sintéticos foram gerados adicionando nucleotídeos artificialmente em ambas as extremidades do duplex miRNA-5p/miRNA-3p para o cálculo de sua energia livre pelo programa *mfold*. Uma vez calculado, os valores de energia livre de ambas as extremidades foram combinados, para se verificar as diferenças de energia livre nas extremidades do duplex. Letras azuis representam o braço 5p. Sequências vermelhas o braço 3p. “N”, representa os nucleotídeos incorporados artificialmente.  $\Delta G$  representa o valor da energia livre de cada pareamento. Traços contínuos representam pareamento A-U/C-G. “:” representa o pareamento oscilante G-U.

## 3.2. Resultados e discussão

### 3.2.1. Visão geral dos dados de RNA-seq

O reprocessamento dos dados brutos de RNA-seq dos estudos SRP041544 (Vaz et al., 2015) e SRP028895 (Wei et al., 2012) resultou num compilado de aproximadamente 508 milhões de *reads* brutas. Das aproximadamente 498 milhões de *reads* remanescentes após a avaliação de qualidade e remoção dos adaptadores, aproximadamente 163 milhões passaram por todas as etapas de filtragem (Figura 6a).



**Figura 6. Panorama geral dos dados de RNA-seq.** A) Contagem das *reads* ao longo das etapas de processamento. B) Distribuição das *reads* identificadas por comprimento.

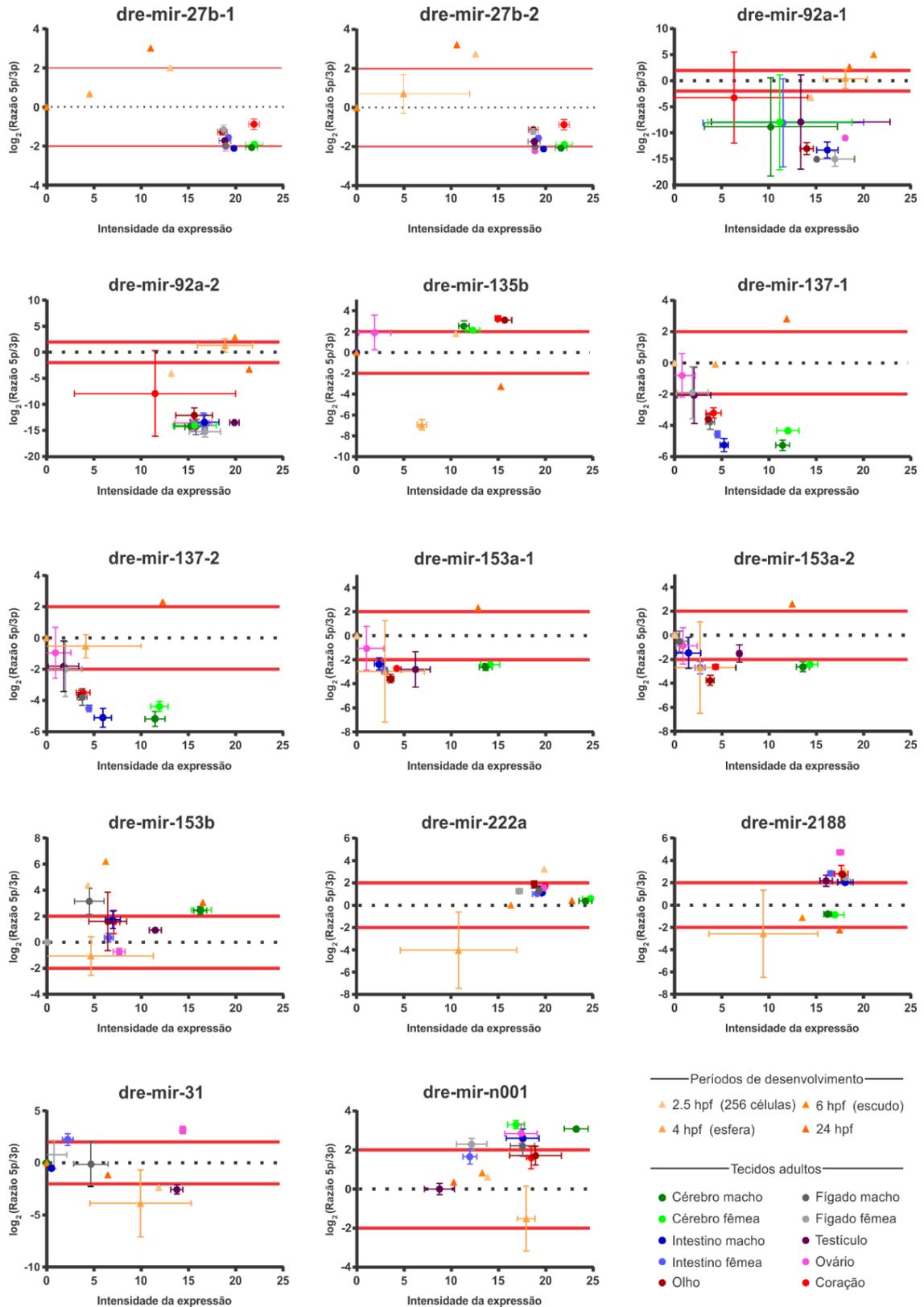
Aproximadamente 70% das *reads* apresentaram comprimentos entre 21 a 23 nucleotídeos (Figura 6b). O comprimento prevalente das *reads* identificadas foi de 22 nucleotídeos na maioria dos tecidos, com exceção dos cérebros de machos e fêmeas que apresentaram comprimento prevalente de 23 nucleotídeos. As amostras de tecido gonadal ainda apresentaram

um pico adicional em 25 nucleotídeos, característicos de transcritos referentes a piwi-RNAs (piRNAs) (Tabela Suplementar S1). Os piRNAs são abundantes nos tecidos gonadais, atuando no silenciamento de transposons e retrotransposons (Lakshimi e Agrawal, 2008) e contribuindo para a proteção e manutenção da integridade e estabilidade do genoma de células gaméticas (Sharp, 2009).

### **3.2.2. Eventos de *arm switching* estão majoritariamente associados ao desenvolvimento ontogenético dos vertebrados**

Nossas análises identificaram 1318 miRNAs maduros distintos (674 miRNA-5p e 674 miRNA-3p), oriundos de 674 loci de miRNAs. Destes, 578 miRNAs já haviam sido previamente caracterizados, enquanto 66 foram considerados miRNAs novos (Tabelas Suplementares S2 e S3). Wei et al. (2012) identificaram 198 miRNAs conhecidos e 8 miRNAs preditos como novos na época de sua publicação utilizando os dados do estudo SRP028895. Segundo estes autores, para a maioria dos miRNAs identificados apenas foram detectadas as cópias de um dos braços. Já Vaz et al. (2015; estudo SRP041544) identificaram aproximadamente 225 miRNAs maduros e 459 pré-miRNAs preditos como novos. A diferença entre os resultados obtidos no presente trabalho e nos estudos cujos dados de transcriptoma originalmente foram publicados mostram que os dados de RNA-seq de alta qualidade disponíveis na literatura estão longe de se tornarem saturados. Com o aprimoramento das ferramentas de bioinformática e novas técnicas de análises de dados de sequenciamento, diversos experimentos podem se beneficiar de dados previamente publicados, tanto para uma reanálise de resultados anteriores utilizando uma metodologia mais robusta, quanto para utilização dos dados em abordagens diferentes, que é o foco deste projeto.

Por meio de uma cuidadosa avaliação dos perfis de expressão dos miRNAs nos diversos tecidos avaliados, fomos capazes de identificar 14 eventos de *arm switching*, dos quais 13 ocorreram em miRNAs conhecidos (*dre-mir-27b-1*, *dre-mir-27b-2*, *dre-mir-92a-1*, *dre-mir-92a-2*, *dre-mir-135b*, *dre-mir-137-1*, *miR-137-2*, *dre-mir-153a-1*, *dre-mir-153a-2*, *dre-mir-153b*, *dre-mir-222a*, *dre-mir-2188* e *dre-mir-31*) e 1 ocorreu em um miRNA novo (*dre-mir-n001*) (Figura 7; Tabela Suplementar S4).



**Figura 7.** Eventos de *arm switching* identificados durante fases do desenvolvimento e em tecidos adultos de *zebrafish*. As linhas vermelhas representam o limiar de 2-fold para caracterização do evento de *arm switching* (log<sub>2</sub>(Razão 5p/3p)). Triângulos representam períodos do desenvolvimento e círculos os tecidos adultos. Tecidos adultos (n=3), 4 hpf (n=2), 2.5, 6 e 24 hpf (n=1).

Ao avaliarmos os padrões de ocorrência dos 14 eventos de *arm switching* identificados, pudemos separá-los em quatro cenários biológicos comparativos: (i) períodos do desenvolvimento embrionário versus tecidos adultos; (ii) entre os diferentes períodos de desenvolvimento embrionário; (iii) gônadas de machos versus gônadas de fêmeas e; (iv) entre cópias parálogas de miRNAs.

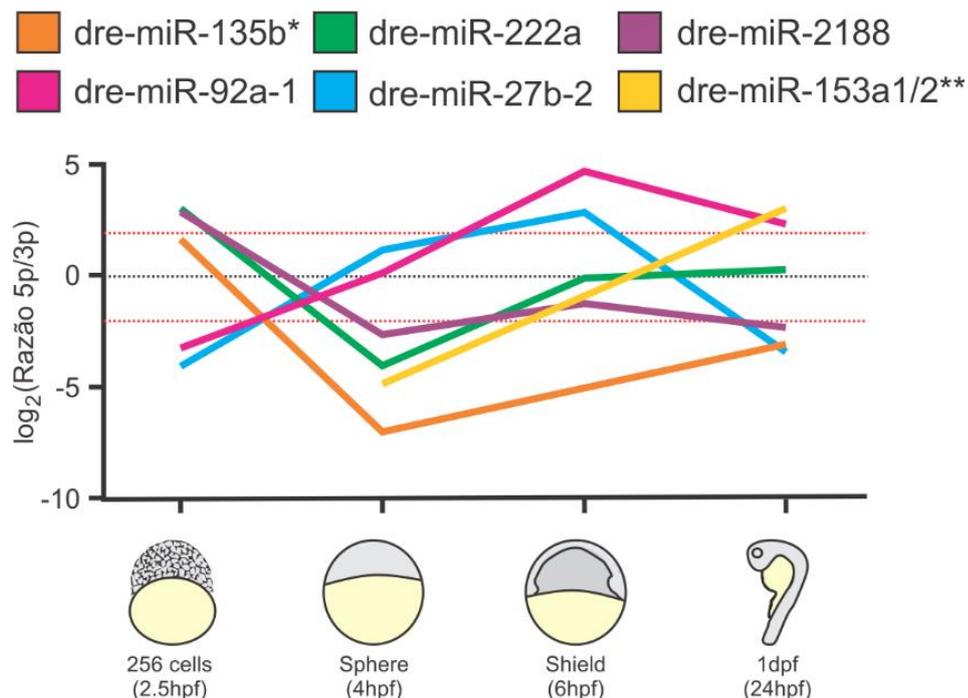
O primeiro cenário, que contempla a comparação entre períodos do desenvolvimento e tecidos adultos, congregou todos os 14 eventos de *arm switching* identificados neste estudo (Figura 7). A clara predominância deste cenário biológico sugere que a ocorrência de *arm switching* pode estar ligada majoritariamente ao controle de processos vitais do desenvolvimento dos organismos, no qual um braço atua regulando genes expressos durante a embriogênese dos organismos enquanto, no indivíduo adulto, o outro braço passa a atuar sobre genes associados ao controle da homeostasia dos tecidos. Eventualmente a desregulação desse processo por alteração da expressão dos braços 5p e 3p poderia trazer efeitos deletérios ao organismo.

De encontro com este raciocínio, vários trabalhos identificaram a ocorrência de eventos de *arm switching* em diversos tipos de câncer (Kuo et al., 2016; Tsai et al., 2016; Chen et al., 2018; Lin et al., 2018). De fato, reprogramações epigenéticas ou mutações em miRNAs podem levar à desregulação de uma ampla gama de genes envolvidos em circuitos regulatórios atuantes durante o desenvolvimento, tais como as vias *Notch*, *Wnt* e *Headgehog*, que também estão presentes em diversos estágios da progressão tumoral (Aiello et al., 2016). Além disso, diversas modificações morfológicas ocorridas durante a embriogênese foram detectadas em alguns tumores. Por exemplo, durante a transição do epitélio para mesênquima (epithelial-to-mesenchymal transition [EMT]), as células epiteliais perdem suas características, como a polaridade basal-apical e a adesão celular, aumentando assim sua mobilidade, sendo este evento extremamente importante durante a gastrulação. O mesmo mecanismo ocorre em células tumorais, aumentando sua mobilidade e facilitando a propagação do câncer (Aiello et al., 2016).

Tais semelhanças moleculares e morfofuncionais ocorrentes entre a embriogênese e doenças tumorais corroboram com a hipótese de que, em ambos os casos, a ativação ou desativação de genes e vias biológicas atuantes durante o desenvolvimento embrionário acarretam na alteração do padrão de expressão dos braços 5p e 3p de alguns miRNAs, levando à ocorrência de eventos de *arm switching*. Futuros trabalhos investigando a expressão global de

miRNAs em linhagens celulares embrionárias e linhagens tumorais poderiam auxiliar na confirmação da ocorrência de *arm switching* do mesmo miRNA em ambos os casos.

Adicionalmente aos eventos de *arm switching* recorrentes entre períodos de desenvolvimento e os tecidos adultos, o segundo cenário identificado em nossas análises detectou sete eventos ocorrendo entre diferentes fases do desenvolvimento: *dre-miR-27b*, *dre-miR-135b*, *dre-miR-192a-1*, *dre-miR-153a-1*, *dre-miR-153a-2*, *dre-miR-222a* e *dre-miR-2188* (Figura 8).



**Figura 8. Razão da expressão dos braços 5p e 3p de eventos de *arm switching* ocorrendo durante os períodos de desenvolvimento estudados.** As linhas vermelhas representam o limiar de 2-fold para caracterização do evento de *arm switching*. \*A expressão do *dre-miR-135b* não foi detectada no período de 6hpf. \*\*A expressão dos *dre-mir-153a-1/-2* não foram detectadas nos períodos de 2,5 e 6hpf.

Com exceção dos miRNAs *dre-mir-153a-1/-2*, dos quais não foram detectados níveis de expressão à 2.5 hpf, a razão 5p/3p à 2.5 hpf destes miRNAs sempre acompanhava razão dos tecidos adultos. Até o período de 2.5 hpf (256 células) o conjunto de miRNAs maduros presente no embrião é totalmente de origem materna (Wei et al., 2012). Assim, é esperado que a proporção 5p/3p encontrada neste período acompanhe a verificada no ovário. De maneira oposta, o *pmi-mir-31* é uma exceção a este caso, no qual a razão 5p/3p a 2.5 hpf é inversa à razão observada no ovário (Figura 7).

No período de 4 hpf (esfera) em *zebrafish*, a maquinaria embrionária começa a produzir miRNAs zigóticos, enquanto o RNA materno passa a ser degradado (Wei et al., 2012). Seguindo este raciocínio, no período de 4 hpf foram identificados eventos de *arm switching* em relação ao período anterior (2.5 hpf), possivelmente ocasionados pelo início da transcrição endógena embrionária (Figura 8). A partir deste período, diversos padrões puderam ser observados.

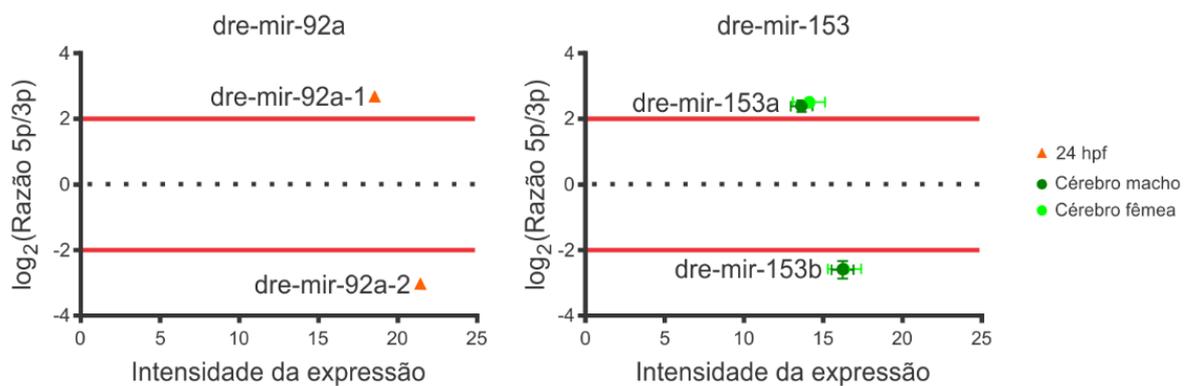
Para o *dre-miR-27b*, por exemplo, há uma predominância do *dre-miR-27b-3p* no período de 2.5 hpf, com uma inversão desta razão ocorrendo no período de 4 hpf. Entretanto, no período de 24 hpf um novo evento de *arm switching* ocorre, no qual o *dre-miR-27b-3p* volta a ser o braço predominante. Para o *dre-miR-2188* e para os *dre-miR-135b*, após a ocorrência do *arm switching* entre os períodos de 2.5 e 4 hpf, a razão entre os braços 5p e 3p se estabilizam, com a prevalência do braço 3p nos períodos subsequentes. Já para o *dre-miR-222a*, o *dre-miR-222a-5p* encontra-se prevalente no período de 2.5 hpf, enquanto que no período de 4 hpf, a prevalência dos braços se inverte, com predominância do braço *dre-miR-222a-3p*. Entretanto, após esse período ambos os braços passam a ser igualmente expressos. No caso do *dre-miR-92a*, o evento completo de *arm switching* leva um pouco mais de tempo para ocorrer. No período de 2.5 hpf, é possível verificar a predominância do *dre-miR-92a-3p*, enquanto que no período de 4 hpf ambos os braços possuem níveis de expressão similares. A inversão da razão verificada no período de 2.5 hpf é finalmente observada no período de 6 hpf, no qual o *dre-miR-92a-5p* é o braço dominante. Destaca-se que estes miRNAs exemplificam o caráter súbito do *arm switching*, dado que inversões completas podem ocorrer em períodos curtos de tempo, como entre os períodos de 2.5 e 4 hpf, possibilitando uma resposta rápida de modulação durante períodos de alta variação na expressão gênica como os períodos iniciais do desenvolvimento ontogenético.

Outro evento que merece destaque foi o *arm switching* do *dre-miR-31* entre os tecidos gonadais, no qual detectamos uma predominância do *dre-miR-31-5p* no ovário, enquanto o *dre-miR-31-3p* é mais expresso no testículo (Figura 7). Este caso é o terceiro cenário biológico identificado, sendo o único evento detectado ocorrendo entre tecidos adultos.

Estudos anteriores demonstraram a importância do *miR-31* na diferenciação gonadal durante o desenvolvimento de galinhas (Cutting et al., 2012). No início do dimorfismo sexual, o *gga-miR-31* encontra-se altamente expresso nos machos, enquanto que em estágios mais avançados do desenvolvimento, quando a diferenciação sexual já foi concluída os níveis de

expressão do *gga-miR-31* se assemelham em indivíduos de ambos os sexos (Cutting et al., 2012). De maneira similar, observamos altos níveis de expressão do *dre-miR-31* apenas durante os estágios iniciais de seu desenvolvimento (2.5 e 4 hpf) e nos tecidos testículo e ovário. Nos estágios de 2.5 hpf, 4 hpf e testículo observa-se uma predominância do *dre-miR-31-3p*, enquanto que no ovário, observamos uma predominância do *dre-miR-31-5p* (Figura 7). Apesar da diferenciação sexual de *zebrafish* ser um evento poligênico, análises de reversão sexual e experimentos de mapeamento de sítios de DNA sugerem a presença de um sistema ZZ/ZW semelhante ao presente nas aves (Sharma et al., 1998; Wilson et al., 2014; Kossack e Draper et al., 2019). Análises computacionais de predição de alvos indicam que nas aves o *gga-miR-31* regula componentes da via de sinalização *TGF- $\beta$*  (Cutting et al., 2012), que possui um papel fundamental no desenvolvimento gonadal (Drummond, 2005; Fan et al., 2011). Desta maneira, é plausível que o *dre-miR-31* atue de maneira semelhante durante o desenvolvimento das estruturas reprodutivas de *zebrafish*.

O último cenário biológico identificado em nossas análises corresponde a *arm switching* entre cópias parálogas de miRNAs. Para o *dre-miR-92a*, verificamos uma variação na predominância dos braços 5p e 3p entre as cópias *dre-mir-92a-1* e *dre-mir-92a-2* no período de 24 hpf. Já para o *dre-miR-153*, tal variação foi observada para as cópias *dre-miR-153a* e *dre-miR-153b*, nas amostras de cérebro (Figura 9).



**Figure 9. Eventos de *arm switching* entre cópias parálogas dos miRNAs *dre-mir-92a* and *dre-mir-153*.** As linhas vermelhas representam o limiar de 2-fold para caracterização do evento de *arm switching*. Amostras de cérebro (n=3) e 24 hpf (n=1).

A regulação da expressão gênica proporcionada pelos miRNAs é altamente conservada entre os organismos, com seu desbalanço podendo levar a graves consequências (Cao e Zen, 2018; Long et al., 2018; Russell et al., 2018). Desta maneira, o surgimento de cópias parálogas de um miRNA qualquer poderia atenuar a pressão seletiva exercida sobre a manutenção da expressão em uma das cópias, facilitando a ocorrência de eventos de *arm switching*.

Segundo Griffiths-Jones et al. (2011), a ocorrência de duplicações gênicas podem dar origem a eventos de *arm switching* de duas formas. A primeira consiste nos casos em que ambos os braços são expressos igualmente. Nestes casos, o surgimento de uma cópia paróloga pode levar à sub-funcionalização de cada cópia, na qual cada uma se especializaria na regulação proporcionada por um dos braços, levando ao aumento da expressão de braços opostos em cada cópia paróloga. A segunda maneira consiste nos casos em que a expressão de um dos braços é predominante sobre a outra. Nestes casos, a duplicação gênica seguida de um evento de *arm switching* pode levar à neofuncionalização desta nova cópia (Ruby et al., 2007). Em ambos os casos as pressões seletivas impostas em cada cópia podem resultar em futuros genes de miRNAs distintos, com baixa identidade de sequência, apesar da ancestralidade comum (Ruby et al., 2007).

A análise minuciosa dos cenários biológicos nos quais os eventos de *arm switching* foram identificados demonstra uma alta correlação entre a ocorrência de eventos de *arm switching* e o desenvolvimento ontogenético inicial de *zebrafish*, no qual todos os eventos detectados ocorreram entre pelo menos um período do desenvolvimento avaliado e uma amostra de tecido adulto. Adicionalmente, para metade destes miRNAs, pudemos observar eventos ocorrendo também entre períodos de desenvolvimento distintos. Cabe destacar que apenas um caso de *arm switching* foi identificado entre tecidos adultos. Entretanto, análises realizadas em outras espécies sugerem que esta ocorrência pode estar associada ao desenvolvimento gonadal de *zebrafish*.

### **3.2.3. *Arm switching* é um fenômeno conservado, porém seus casos são pontuais e espécie-específicos**

Eventos de *arm switching* já foram descritos em um amplo número de espécies (Landgraf et al., 2007; Ro et al., 2007; Glazov et al., 2008; Wit et al., 2009; Chiang et al., 2010; Griffiths-Jones et al., 2011; Pinhal et al., 2018). Entretanto, em cada estudo, um conjunto diferente de miRNAs exibiu variações na prevalência de seus braços, e apenas um número

limitado de eventos foram identificados para cada uma delas. Nosso grupo de pesquisa realizou uma análise em larga escala de miRNAs na tilápia do Nilo utilizando os mesmos critérios descritos neste trabalho e detectamos 9 eventos de *arm switching* (de um total de 368 miRNAs identificados) (Pinhal et al., 2018), ocorrendo majoritariamente entre o período de 48 hpf e tecidos adultos. Entretanto, não foi encontrado nenhum evento em comum entre o *zebrafish* e a tilápia do Nilo. Adicionalmente, um estudo em larga escala realizado em camundongos identificou *arm switching* em 21 miRNAs (de 506 miRNAs analisados) ao comparar diversos períodos do desenvolvimento (ES, e7.5, e9.5, e12.5 e recém-nascido) aos tecidos adultos (testículo, ovário e cérebro) (Chiang et al, 2010). Novamente, nenhum evento identificado foi similar aos verificados em *zebrafish* e apenas um caso (e.g., miR-30e-1) foi encontrado tanto em camundongo quanto na tilápia do Nilo.

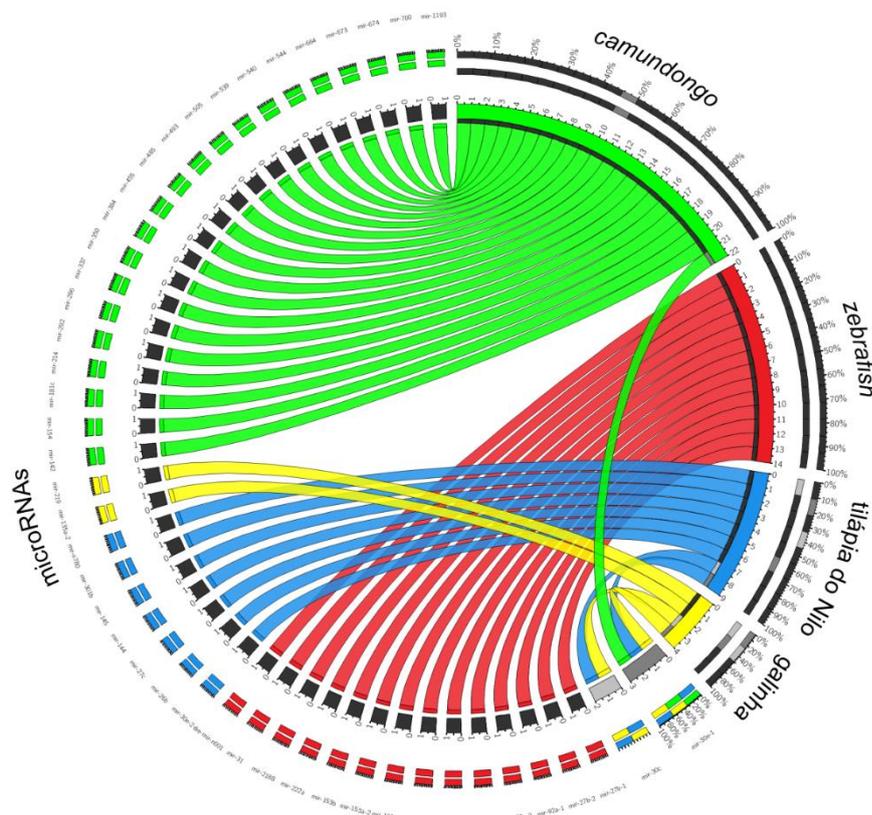
Adicionalmente a estes experimentos de larga escala, outros eventos de *arm switching* foram descritos em contextos específicos. Landgraf et al. (2007) demonstraram que em mamíferos, os miRNAs mir-100 e miR-125 possuíam o braço 5p mais expresso, mas em alguns tecidos o braço 3p foi detectado como o braço dominante. Glazov et al. (2008) identificaram quatro eventos de *arm switching* durante o desenvolvimento em galinha (miR-135a-2, miR-30e, miR-219 e miR-30c). Um evento de *arm switching* no miR-30e também foi detectado entre estômago e baço em camundongos (Ro et al, 2007). Adicionalmente, *arm switching* em oito miRNAs foram detectados entre quatro espécies de nematodas comparativamente avaliados (*Caenorhabditis elegans*, *C. briggsae*, *C. remanei* e *Pristionchus pacificus*; Wit et al., 2009) e dois eventos foram descritos ocorrendo entre duas espécies de insetos (*Drosophila melanogaster* and *Tribolium castaneum*; Griffiths-Jones et al., 2011).

Em uma análise de maior escala, Wit et al. (2009) também avaliaram a prevalência global dos braços 5p e 3p em 21 espécies com pelo menos 50 miRNAs descritos no miRBase e identificaram que enquanto em vertebrados há uma predominância dos braços 5p, para invertebrados como *Drosophila melanogaster*, *D. pseudoobscura*, e as espécies de nematodas analisadas, há uma prevalência da expressão do braço 3p, sugerindo que houve uma grande troca na prevalência dos braços de miRNAs durante algum ponto da evolução dos metazoários.

Os resultados obtidos em nosso trabalho, aliado aos achados descritos na literatura demonstram que os eventos de *arm switching* compreendem apenas uma fração muito pequena dos miRNAs expressos no genoma. Estes dados indicam que apesar de o processo de *arm switching* ser um fenômeno conservado, eles possuem uma ocorrência rara e pontual ao nível

de espécie. Ao analisarmos eventos específicos, verificamos que os casos observados em uma espécie geralmente não ocorrem em outra (Figura 10), demonstrando que os eventos de *arm switching* são em sua maioria espécie-específicos e podem estar relacionados às particularidades do ambiente celular no qual o miRNA está expresso. Uma interessante exceção é a família do miR-30, no qual eventos de *arm switching* para o miR-30c foram identificados para a tilápia do Nilo e galinha, enquanto eventos para o miR-30e-1 foram identificados para tilápia do Nilo, galinha e camundongo.

Adicionalmente, de maneira similar aos dados obtidos em nosso trabalho, a ocorrência de eventos de *arm switching* associados ao desenvolvimento também foi amplamente observada para as espécies de camundongo, galinha e tilápia do Nilo. Estes achados sugerem que eventos de *arm switching* são quase uma constante durante o desenvolvimento ontogenético nos vertebrados.



**Figura 10. Distribuição de eventos de *arm switching* entre zebrafish, tilápia do Nilo, galinha e camundongo.** Segmentos verdes, vermelhos, azuis e amarelos representam eventos de *arm switching* em camundongos, zebrafish, tilápia do Nilo e galinha respectivamente. Quadrados pretos representam eventos de *arm switching* identificados apenas em uma espécie. O quadrado cinza claro (miR-30c) representa o evento de *arm switching* detectado em tilápia do Nilo e galinha. O quadrado cinza escuro (miR-30e-1) representa o evento de *arm switching* detectado em tilápia do Nilo e galinha e camundongo.

### 3.2.4. *Arm switching* amplia potencial regulatórios dos microRNAs em *zebrafish*, alterando os alvos e funções biológicas por eles controladas entre tecidos distintos

Uma vez que metade dos casos de *arm switching* identificados em *zebrafish* ocorreram entre as amostras de 24 hpf e cérebro, as análises subsequentes de predição de alvos, enriquecimento funcional, padrões de isomiR e identificação de picos de metilação m<sup>6</sup>A foram realizadas exclusivamente para estes miRNAs, nestes tecidos (*dre-mir-27b-1/-2*, *dre-mir-135b*, *dre-mir-137-1/-2* e *dre-mir-153a-1/-2*). Isso nos permitiu a realização de uma análise mais coesa e direcional, sem a presença de ruídos externos causados por variações de múltiplos tecidos.

A análise de predição de alvos retornou um total de 16.593 e 13.681 potenciais alvos para os miRNAs analisados, nas amostras de cérebro e embrião 24 hpf, respectivamente. Uma vez que a sequência dos braços 5p e 3p dos miRNAs diferem uma da outra, alterando, principalmente a composição da sequência *seed*, principal região responsável pelo pareamento do miRNA com seu mRNA alvo (Bartel, 2009), existem duas possibilidades pelas quais os braços de um miRNA poderiam interagir com seus alvos: (i) cada braço regularia um conjunto diferente de genes ou (ii) os braços interagiriam com genes semelhantes porém em *MREs* distintas.

De maneira similar aos achados em outros trabalhos (Griffiths-Jones et al., 2011; Marco et al., 2012), nossos dados demonstraram uma baixa sobreposição de genes entre os alvos preditos para cada braço de um mesmo miRNA. Uma exceção a isso foi o *dre-miR-153a-1/-2*, que retornou um maior número de alvos preditos para o braço 5p que a média, acarretando em um maior número de alvos em comum com o braço 3p (Tabela 1).

Griffiths-Jones et al. (2011) destaca que para os miRNAs avaliados em seu trabalho (*miR-10*, *miR-993*, *miR-100* e *miR-125*) o número de alvos preditos para o braço canônico é menor do que para o braço passageiro e argumenta que talvez haja uma pressão seletiva contra a presença de sítios de pareamento da cópia mais expressa. Porém, nossos dados demonstram que nem sempre isto ocorre. Os braços *dre-miR-27b-3p* (cérebro), *dre-miR-135b-3p* (24 hpf), *dre-miR-137b-5p* (24 hpf) e *dre-miR153a-5p* (24 hpf) possuem mais alvos preditos que seu braço irmão apesar de serem os braços mais expressos nestes tecidos. De maneira oposta, os braços *dre-miR-27b-5p* (24 hpf), *dre-miR-135b-5p* (cérebro), *dre-miR-137b-3p* (cérebro) e o *dre-miR-153a-3p* (cérebro) são os braços mais expressos e possuem menos alvos que seus braços irmãos (Tabela 1).

**Tabela 1.** Número de alvos preditos para os braços 5p e 3p dos miRNAs *dre-mir-27b-1/-2*, *dre-mir-135b*, *dre-miR-137b-1/-2* e *dre-mir-31* no cérebro e no embrião 24 hpf. Valores em negrito representam os alvos preditos no o braço mais expresso. A lista completa pode ser encontrada na Tabela Suplementar S5.

| miRNA                    | Tecido  | Alvos braço 5p | Alvos braço 3p | Alvos em comum |
|--------------------------|---------|----------------|----------------|----------------|
| <i>dre-mir-27b-1/-2</i>  | Cérebro | 2356           | <b>2465</b>    | 1057           |
|                          | 24 hpf  | <b>1882</b>    | 2035           | 838            |
| <i>dre-mir-135b</i>      | Cérebro | <b>1065</b>    | 1461           | 378            |
|                          | 24 hpf  | 1302           | <b>1749</b>    | 469            |
| <i>dre-mir-137b-1/-2</i> | Cérebro | 1656           | <b>1358</b>    | 448            |
|                          | 24 hpf  | <b>1298</b>    | 1165           | 340            |
| <i>dre-miR-153a-1/-2</i> | Cérebro | 3974           | <b>1733</b>    | 1147           |
|                          | 24 hpf  | <b>3343</b>    | 1432           | 947            |

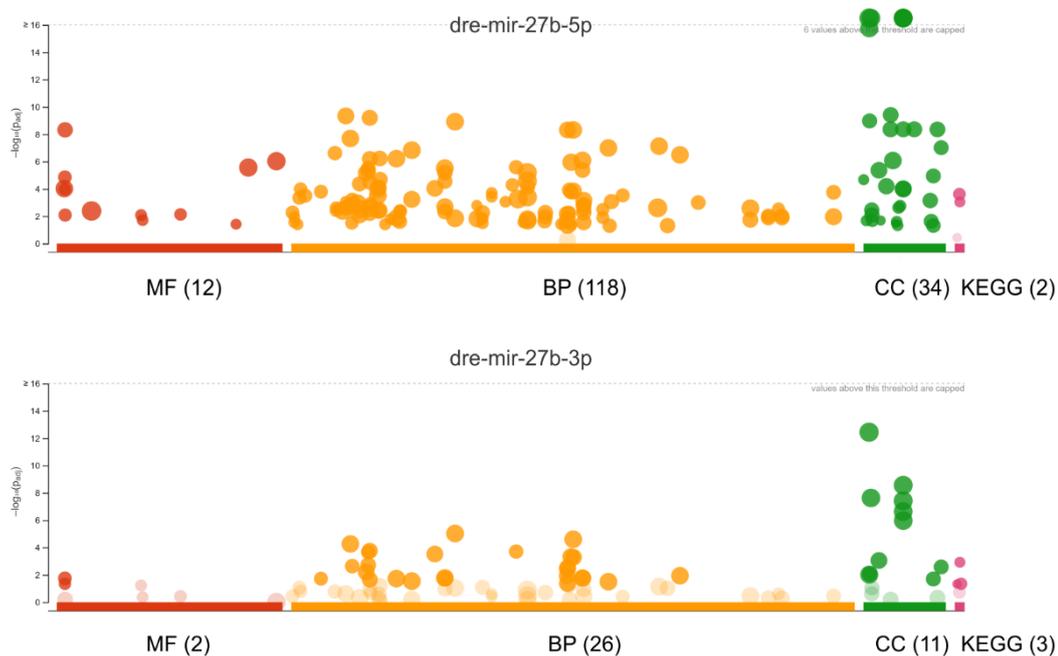
As divergências nos resultados obtidos por nós e os dados identificados por Griffiths-Jones et al. (2011) podem estar associadas à metodologia empregada na predição de alvos. Em seu trabalho, Griffiths-Jones et al. (2011) considerou a predição global destes miRNAs, sem levar em consideração sua presença ou ausência em tecidos específicos, enquanto que nossos dados levaram em consideração a expressão dos genes-alvo nos tecidos avaliados. Desta maneira, verificamos que a relação entre o número de alvos e a expressão do miRNA é dependente do contexto celular em que se insere, e não necessariamente ocorre uma pressão seletiva contra sítios de interação do braço mais expresso em todos os cenários.

Notadamente, as análises de enriquecimento funcional mostraram que não só há discrepância no conjunto de alvos regulados por cada braço, como também há variação no papel funcional que estes alvos desempenham (Figura 11; Tabela Suplementar S6).

As divergências mais acentuadas ocorrem nas categorias funções moleculares e vias biológicas, enquanto que para a categoria processos biológicos, apesar de haver um expressivo número de processos controlados exclusivamente por um dos dois braços, ainda é possível verificar a ocorrência de termos em comum. Estes resultados demonstram uma separação na função regulatória exercida pelos braços 5p e 3p, uma vez que, embora haja semelhanças entre

alguns dos processos biológicos regulados por ambos os braços, as funções moleculares atuantes nestes processos, bem como as diversas vias biológicas neles inseridas tendem a ser reguladas por braços distintos.

É importante destacar que a ocorrência de termos altamente enriquecidos em ambos os braços (e.g. *developmental process* (GO:0032502) e *anatomical structure development* (GO:0048856)) pode derivar do viés metodológico recorrente da técnica de enriquecimento funcional. No consórcio *Gene Ontology*, cada gene pode apresentar diversos termos associados, nos quais cada um possui um ou mais termos parentais associados a ele. Desta maneira, o enriquecimento de termos específicos distintos pode levar a ocorrência de enriquecimento de termos parentais em comum em ambos os braços.



**Figura 11. Enriquecimento funcional dos alvos do *dre-miR-27b* expressos no embrião 24 hpf.**

Pontos vermelhos = Função Molecular. Pontos amarelos = Processos Biológicos. Pontos verdes = Componentes Celulares. Pontos rosas = Vias Biológicas. A posição dos pontos no eixo Y indicam a significância estatística do termo ( $p \leq 0.05$ ). Pontos escuros indicam termos estatisticamente significativos e pontos claros indicam termos encontrados sem significância estatística. A posição dos pontos no eixo X representam os termos distintos. O diâmetro dos pontos indica o número de genes associados com o termo. Os resultados detalhados de todos os miRNAs analisados podem ser encontrados na Tabela Suplementar S4.

Outro padrão interessante detectado é que, apesar das diferenças verificadas nas funções moleculares e processos biológicos dos alvos, os componentes celulares em que os alvos de ambos os braços atuam tendem a possuir um maior grau de semelhança (Figura 11; Tabela Suplementar S4). Tal característica vai de encontro ao processo de biogênese de um miRNA, no qual ambos os braços acumulados são oriundos da mesma molécula precursora de pré-miRNA. Entretanto, é válido salientar que a quantidade de termos descritos para “Componentes Celulares” no Consórcio *Gene Ontology* é inferior aos termos associados à “Processos Biológicos”, “Funções Moleculares” e “Vias Biológicas”, podendo este ser um fator impactante na maior semelhança encontrada para estes termos.

Para o *dre-miR-27b* expresso no embrião 24hpf, apenas uma função molecular (*transcription coregulator activity*; GO:0003712) e uma via KEGG (*Notch signaling pathway*; KEGG:04330) foram enriquecidos tanto no *dre-miR-27b-5p* quanto no *dre-miR-27b-3p*. Por outro lado, 11 dos 34 componentes celulares e 22 dos 119 dos processos biológicos enriquecidos foram observados em ambos os braços. Chama-se a atenção que os termos enriquecidos para apenas um braço foram observados apenas para o *dre-miR-27b-5p*, o braço mais expresso no embrião 24 hpf, para os termos função molecular, componente celular e processo biológico, enquanto 2 termos de vias KEGG (*Insulin signaling pathway*; KEGG:04910 e *Terpenoid backbone biosynthesis*; KEGG:00900) foram enriquecidos apenas no *dre-miR-27b-3p* (Figura 5; Tabela Suplementar S4).

Ao avaliar os termos enriquecidos dos alvos do *dre-miR-27b* expressos no cérebro, cujo *dre-miR-27b-3p* é o braço mais expresso, observamos 21 funções moleculares e uma via KEGG (*Endocytosis*; KEGG:04144) enriquecida em ambos os braços. Para os termos referentes a componentes celulares e processos biológicos, identificamos 17 dos 46 e 37 dos 115 termos enriquecidos em ambos os braços, respectivamente. Diferentemente dos dados de enriquecimento no embrião 24 hpf, identificamos um maior número de termos enriquecidos exclusivamente para os braços 5p ou 3p (Tabela Suplementar S4). Os resultados detalhados da referentes ao enriquecimento funcional de todos os miRNAs analisados podem ser encontrados na Tabela Suplementar S4.

As diferenças observadas, tanto no conjunto de alvos quanto em seus papéis biológicos, destacam o amplo potencial regulatório mediado por miRNAs, capaz de produzir duas moléculas maduras com funções distintas a partir de uma única molécula precursora. Tal modelo potencializa o impacto dos eventos de *arm switching* na regulação gênica, permitindo

à célula alternar entre a expressão dos braços 5p e 3p dos miRNAs e, conseqüentemente, modular conjuntos de genes distintos de acordo com o contexto funcional demandado.

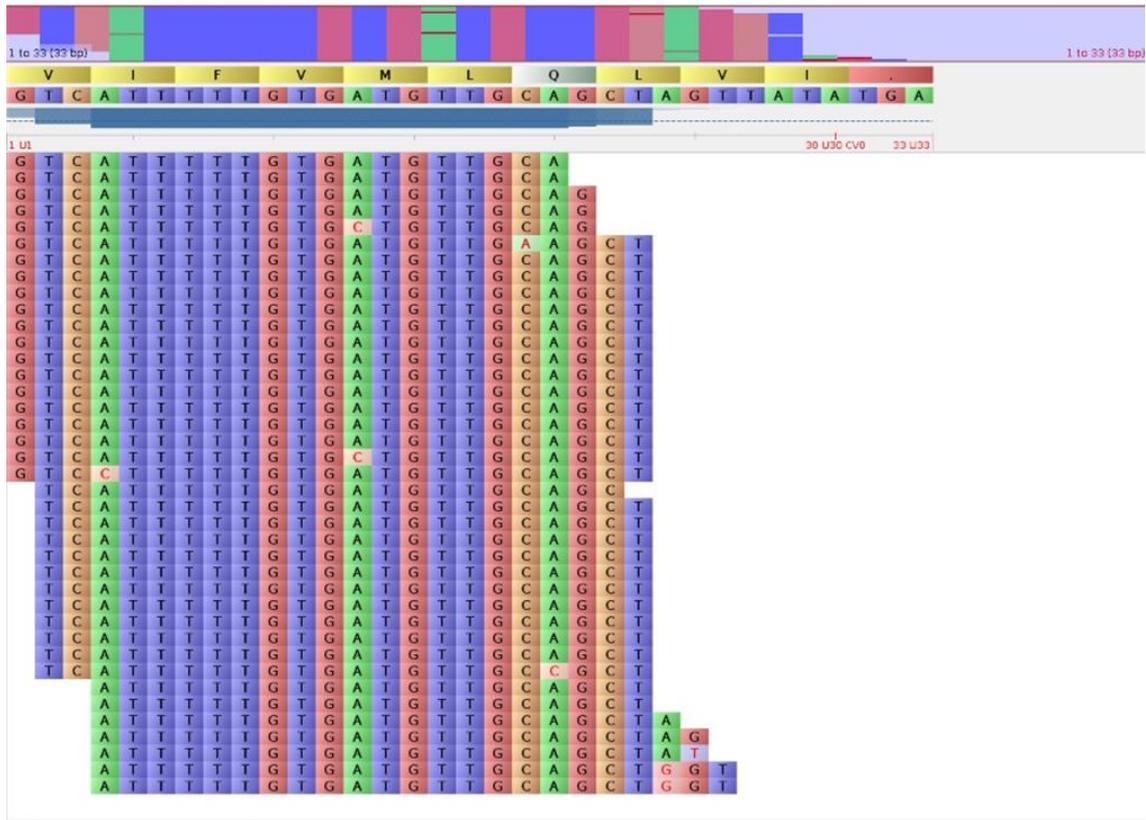
### 3.2.5. Expressão diferencial de isomiRs é capaz de alterar a isoforma representativa dos microRNAs

Uma vez que a ocorrência de isomiRs pode promover alterações tanto na instabilidade relativa das porções terminais do duplex miRA-5p/miRNA-3p, quanto alterar a identidade dos nucleotídeos iniciais de cada braço, nós buscamos analisar os padrões de isomiR presentes nos eventos de *arm switching* identificados para avaliar como eles impactam a seleção dos braços durante a incorporação nas proteínas Ago.

Nossas análises identificaram diversas modificações nas sequências das *reads* expressas, ambas relativas a modificações modelo e modificações não-modelo, caracterizando uma ocorrência típica de isomiRs (Figura 12). Entre os isomiRs detectados, a grande maioria advinha de prováveis alterações nos cortes promovidos pelas enzimas *Dicer* e *Drosha*, durante a biogênese destes miRNAs, enquanto apenas uma pequena porção (~0.72%) correspondia a modificações não modelo, como substituições ou adições não modelo a região 3'.

Interessantemente, os mesmos tipos de isomiRs foram detectados em ambos os tecidos cérebro e embrião 24 hpf, enquanto isomiRs exclusivos de um tecido apresentavam rara ocorrência. Entretanto, a expressão relativa destas isoformas compartilhadas variou bastante entre os tecidos analisados. Para cinco dos sete miRNAs avaliados, tais variações na expressão dos isomiRs foram suficientes para alterar a isoforma representativa (i.e. a isoforma mais expressa) entre os tecidos de pelo menos um dos braços dos miRNAs (Tabela 2). Por exemplo o *dre-miR-135b-5p*, é menos expresso que o *dre-miR-135b-3p*, no embrião 24 hpf e sua isoforma representativa possui 23 nucleotídeos. Entretanto, na amostra de cérebro há um evento de *arm switching*, na qual o *dre-miR-135b-5p* torna-se o braço mais expresso. Este evento é seguido por uma troca na isoforma representativa deste miRNA, que possui uma sequência menor que sua versão no embrião 24 hpf, com uma adenina a menos na região 3' (Tabela 2).

Estes dados sugerem que a troca da isoforma representativa poderia promover a ocorrência de *arm switchings*, uma vez que esta troca possui o potencial de alterar as propriedades do duplex e, portanto, interferir na escolha de qual miRNA será incorporado pela Ago e exercerá seu papel na regulação gênica.



**Figura 12. Padrões de IsomiR do *dre-miR-153a-2* expressos no cérebro.** Diversos padrões de isomiR foram identificados em ambos os braços, como *reads* com substituições, inserções, deleções, adições não-modelo na porção 3' e modificações na clivagem da região 5' ou 3'.

**Tabela 2. Reads representativas dos braços 5p e 3p nos tecidos embrião 24hpf e cérebro.** Letras em vermelho representam nucleotídeos presentes na *read* representativa de apenas um tecido. Reads em negrito representam as *reads* do braço mais expresso daquele tecido.

| miRNA                    | Embrião 24 hpf                  | Cérebro                        |
|--------------------------|---------------------------------|--------------------------------|
| <i>dre-miR-135b-5p</i>   | TATGGCTTTTTATTCCCTATCTGA        | <b>TATGGCTTTTTATTCCCTATCTG</b> |
| <i>dre-miR-135b-3p</i>   | <b>ATATAGGGATGGAAGCCATGCA</b>   | TATAGGGATGGAAGCCATGCA          |
| <i>dre-miR-137-5p-1</i>  | <b>ACGGGTATTCTTGGGTGGATAATA</b> | ACGGGTATTCTTGGGTGGATA          |
| <i>dre-miR-137-3p-1</i>  | TTATTGCTTAAGAATACGCGT           | <b>TTATTGCTTAAGAATACGCGT</b>   |
| <i>dre-miR-137-5p-2</i>  | <b>ACGGGTATTCTTGGGTGGATAATA</b> | ACGGGTATTCTTGGGTGGATAA         |
| <i>dre-miR-137-3p-2</i>  | TTATTGCTTAAGAATACGCGT           | <b>TTATTGCTTAAGAATACGCGT</b>   |
| <i>dre-miR-27b-5p-1</i>  | AGAGCTTAGCTGATTGGTGAACA         | <b>AGAGCTTAGCTGATTGGTGAAC</b>  |
| <i>dre-miR-27b-3p-1</i>  | <b>TTCACAGTGGCTAAGTTCTG</b>     | TTCACAGTGGCTAAGTTCTG           |
| <i>dre-miR-27b-5p-2</i>  | AGAGCTTAGCTGATTGGTGAACA         | <b>AGAGCTTAGCTGATTGGTGAAC</b>  |
| <i>dre-miR-27b-3p-2</i>  | <b>TTCACAGTGGCTAAGTTCTG</b>     | TTCACAGTGGCTAAGTTCTG           |
| <i>dre-miR-153a-5p-1</i> | <b>GTCATTTTGTGATGTTGCAGCT</b>   | GTCATTTTGTGATGTTGCAGCT         |
| <i>dre-miR-153a-3p-1</i> | TTGCATAGTCACAAAAGTGATC          | <b>TTGCATAGTCACAAAAGTGATC</b>  |
| <i>dre-miR-153a-5p-2</i> | <b>GTCATTTTGTGATGTTGCAGCT</b>   | GTCATTTTGTGATGTTGCAGCT         |
| <i>dre-miR-153a-3p-2</i> | TTGCATAGTCACAAAAGTGATC          | <b>TTGCATAGTCACAAAAGTGATC</b>  |

Os miRNAs *dre-miR-153-1/-2* foram uma exceção a esta regra, possuindo a mesma *read* representativa dos braços 5p e 3p em ambos os tecidos, apesar da troca na prevalência da expressão destes braços. Neste caso, verificamos que a expressão relativa entre os isomiRs de um mesmo braço não se alteraram de um tecido para o outro, apesar da ocorrência dos eventos de *arm switching*.

Nossas análises também revelaram padrões interessantes referentes a ocorrência de isomiRs, particulares de cada tecido. Primeiramente, as *reads* representativas expressas do embrião 24h apresentaram comprimento maior que as *reads* detectadas no cérebro nestes miRNAs (Tabela 2). Interessantemente, estes dados diferem dos dados referentes às *reads* totais expressas nestes tecidos, dos quais o embrião apresenta uma prevalência de *reads* com 22 nucleotídeos, enquanto no cérebro verifica-se um pico no comprimento de 23 nucleotídeos.

Adicionalmente, nós detectamos uma maior ocorrência de isomiRs modelo-independente no cérebro (média de 1,14%) em relação ao embrião (média de 0,29%), nos miRNAs analisados ( $p < 0.0001$ ). Diversos trabalhos publicados recentemente demonstram ocorrências de isomiRs tecido específicos (Telonis et al., 2017; Woldemariam et al., 2019; Wang et al., 2019), podendo inclusive apresentar funções tecido específicas (Colaiacovo et al., 2012; Sablok et al., 2015). Com isso, a alta taxa de isomiRs modelo-independente detectada no cérebro pode possuir um papel importante na correta regulação gênica deste tecido (Cloonan et al., 2011). Análises futuras direcionadas ao estudo dos impactos dos isomiRs na regulação gênica exercida pelos miRNAs no tecido nervoso podem fornecer maiores informações sobre suas funções biológicas.

Outro caso interessante detectado foi o do *dre-miR-135b*, no qual a *read* representativa do *dre-miR-135b-3p* no embrião 24 hpf, possui uma adenina adicional na porção 5' em relação à *read* representativa identificada no cérebro. IsomiRs decorrentes de alterações na porção 5' do miRNA são extremamente raros, e promovem eventos chamados de *seed shifting* ou troca da *seed*, no qual a região *seed*, principal região guia do pareamento do miRNA com seu alvo, possui sua sequência alterada. Estes isomiRs são capazes de regular um conjunto amplamente distinto de alvos em relação as isoformas que possuam sequências *seed* diferentes, ampliando o potencial regulatório do miRNA em questão (Guo e Chen et al., 2014; Haseeb et al., 2017).

A somatória destes eventos, ou seja, a ocorrência de eventos de *arm switching* associadas a prevalência de diferentes isomiRs demonstram a grande potencial modulatório que

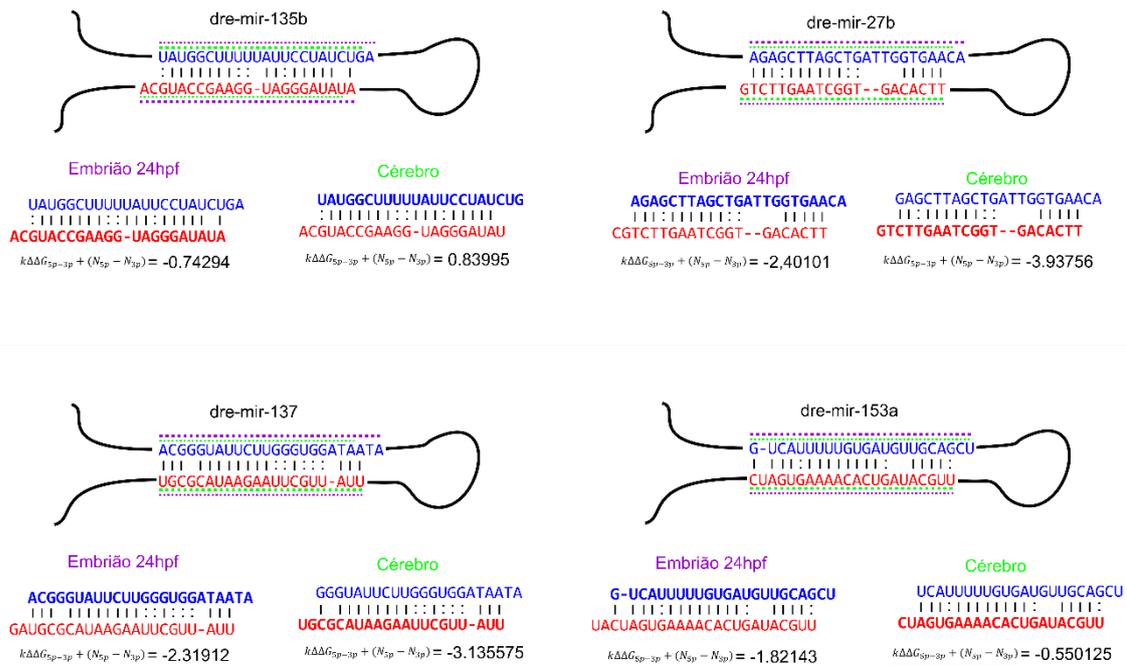
um único gene de miRNA é capaz de promover em diversas condições específicas demandadas para uma ampla e eficiente regulação gênica.

Para avaliar como os isomiRs poderiam afetar as propriedades do duplex e a seleção diferencial dos braços 5p e 3p, utilizamos a sequência representativa do braço mais expresso de cada tecido para a construção de duplex conceituais contendo 2 nucleotídeos sobressalentes nas porções 3' e calculamos a razão de incorporação dos braços 5p e 3p destes duplex no complexo RISC seguindo a fórmula descrita por Suzuki et al. (2015) (Material e Métodos item 3.1.6; Figura 4).

De todos os casos analisados, apenas o *dre-miR-135b* apresentou uma correta correlação entre o resultado obtido na fórmula proposta por Suzuki et al. (2015) e as proporções observadas experimentalmente: valores negativos para o duplex representativo do embrião, cujo braço mais expresso é o *dre-miR-135b-3p* e valores positivos para o duplex representativo do cérebro, cujo braço mais expresso é o braço 5p (Figura 13). Nos demais casos, ambos os duplex apresentaram valores negativos, apesar de em um deles o *dre-miR-135b-5p* ser mais expresso naquele tecido.

Curiosamente, o *dre-miR-135b* também foi o único entre os miRNAs estudados cujas *reads* representativas dos braços 5p e 3p formavam de maneira natural o duplex conceitual contendo dois nucleotídeos sobressalentes nas porções 3' em ambos os tecidos (Figura 12). Com isso, é possível que os duplex dos outros miRNA, construídos a partir de modelos, não representem fielmente a realidade dos duplex expressos de maneira endógena, uma vez que diversas etapas de processamento que não são captadas pelo sequenciamento podem ocorrer. Apesar de estar bem estabelecido na literatura a ocorrência de um corte comprimento e forma fixa promovido pela *Dicer*, devido à distância e características moleculares entre os domínios *PAZ* e *RNAse III* (Lingel et al., 2004; Macrae et al., 2006), não se pode descartar a ocorrência de cortes imprecisos realizados por esta enzima (Cloonan et al., 2011), que promoveriam duplex com propriedades distintas (e.g. um ou nenhum nucleotídeo sobressalente nas regiões 3').

Uma vez que a técnica de RNA-seq não é capaz de identificar as moléculas pareadas no estágio de duplex, tais propriedades não são possíveis de serem observadas. Desta maneira, estes dados não são suficientes para avaliar como os isomiRs estariam afetando o reconhecimento molecular do duplex pela *Ago*. Estudos futuros de co-precipitação e análises funcionais de modificações *in vitro* destas estruturas poderiam fornecer novos conhecimentos sobre o assunto.



**Figura 13. Análise das propriedades do duplex dos miRNAs.** Na parte superior de cada miRNA, há uma representação de seu pri-miRNA, contendo as sequências representativas dos braços 5p (em azul) e 3p (em vermelho) dos tecidos em que estes braços são mais expressos. Linhas roxas representam comprimento da molécula madura expressa no embrião 24 hpf, enquanto linhas verdes representam comprimento da molécula expressa no cérebro. Abaixo, há a estrutura conceitual dos duplex formados por ambos os braços, bem como a razão de incorporação destes duplex no complexo RISC. Valores positivos indicam prevalência de incorporação do braço 5p e valores negativos prevalência de incorporação do braço 3p. Sequências em negrito indicam o braço mais expresso naquele tecido.

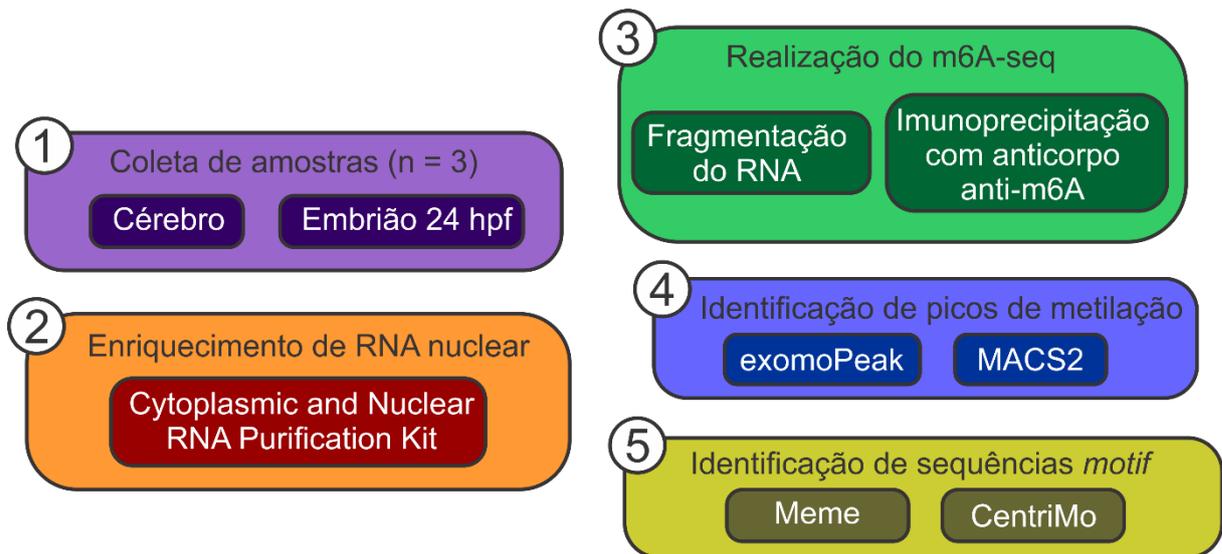
## **4. Capítulo II:**

*Metilação m<sup>6</sup>A e seu papel no mecanismo de atuação dos microRNAs*

## 4.1. Material e Métodos

### 4.1.1. Resumo do *workflow*

A Figura 14 abaixo, apresenta um breve resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 2. Cada etapa será detalhadamente descrita nos tópicos a seguir.



**Figura 14. Resumo das atividades realizadas referentes aos resultados discutidos no Capítulo 2.** 1) Amostras de cérebro e embrião 24hpf (n = 3) foram coletadas, para a extração de RNA. 2) A fração de RNA nuclear foi enriquecida e purificada, utilizando-se o kit *Cytoplasmic and Nuclear RNA purification kit*, para enriquecer as amostras com transcritos de pri-miRNAs. 3) As amostras foram fragmentadas e a imunoprecipitação com anticorpo anti-m6A recuperou os fragmentos metilados. Os fragmentos metilados foram então sequenciados. 4) A identificação dos picos de metilação foi realizada utilizando-se as ferramentas exomePeak, para a identificação de metilações em regiões anotadas do genoma e o programa MAC2S, para a identificação de picos de metilação em regiões flanqueadoras de pré-miRNAs anotados, caracterizando-as como metilações nos pri-miRNAs não anotados. 6) As sequências *motif* foram identificadas utilizando a ferramenta Meme e a ferramenta CentriMo calculou a distância dos picos de metilação das sequências *motif* identificadas.

### 4.1.2. Cuidados com *zebrafish* e coleta de amostras

Espécimes de *zebrafish* da linhagem TAB foram mantidos no biotério do Laboratório Genômica e Evolução Animal (LGEM) do Departamento de Ciências Químicas e Biológicas

(DCQB), IBB, UNESP sob condições ótimas controladas de temperatura (27-28°C), pH (7.0), condutividade (900  $\mu$ S) e ciclo claro/escuro (14/10 horas). Os peixes foram alimentados quatro vezes ao dia com artemia e ração industrial até saciedade aparente.

Para reprodução e obtenção de embriões, machos e fêmeas foram alocados em aquários próprios para este fim e os embriões foram produzidos por fecundação natural. Os embriões obtidos foram mantidos em incubadoras a 28°C para a coleta de RNA no período de tempo 24 hpf. Adicionalmente, sessões de reprodução para a manutenção da população do sistema foram realizadas periodicamente. Nestas ocasiões, os embriões produzidos foram mantidos durante cinco dias nas incubadoras a 28°C antes de serem transferidos para o sistema de aquários do biotério. Uma vez nos aquários, embriões foram alimentados com *paramécia* quatro vezes ao dia até saciedade aparente e até atingirem tamanho adequado para serem alimentados com artêmias e ração industrializada.

Os animais foram sacrificados por overdose do anestésico MS-222 (50 mg/L tricaine-methanensulfonate; Sigma-Aldrich, USA). Amostras de cérebro de indivíduos do sexo masculino e amostras de embriões 24 hpf (N = 3) foram coletadas para a realização do m<sup>6</sup>A-seq. Para cada replicata, a quantidade de amostras utilizadas foi o suficiente para totalizar 20  $\mu$ g de RNA nuclear. Imediatamente após a coleta, as amostras foram submersas em nitrogênio líquido para estabilização do RNA e armazenadas em *biofreezer* -80°C até sua utilização.

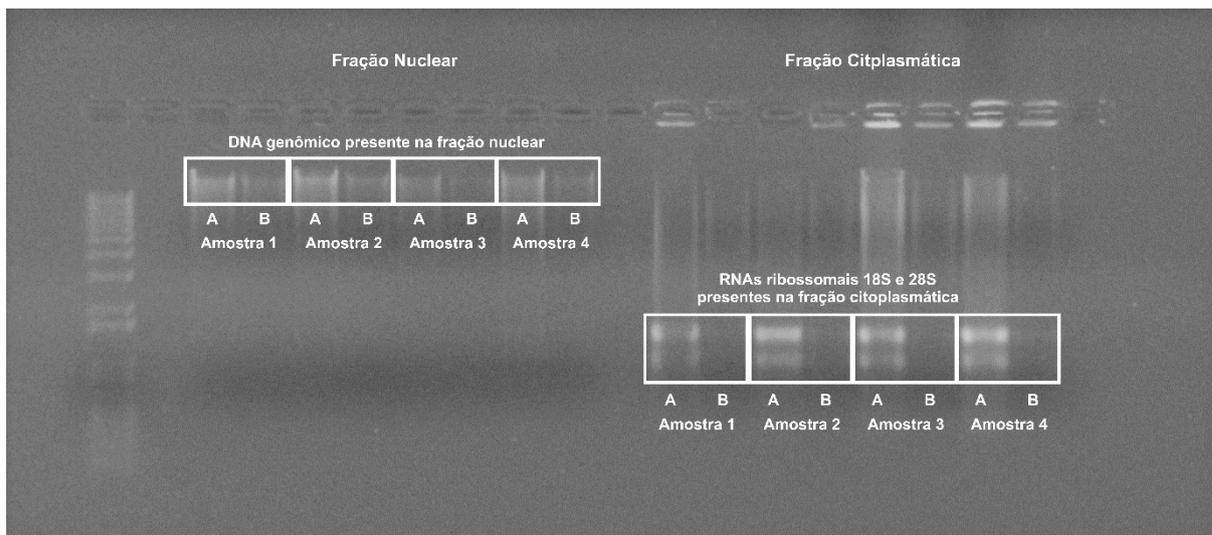
Todos os procedimentos realizados foram aprovados pelo comitê de ética (CEUA protocolo n° 1008 - 2017).

#### 4.1.3. Enriquecimento de RNA nuclear

A extração de RNA nuclear foi realizada utilizando-se o protocolo publicado por Dominissini et al. (2013), com modificações para adaptá-lo ao estudo de miRNAs. Uma etapa importante deste processo de adaptação é enriquecer as amostras com pri-miRNAs, devido à alta concentração de sítios de metilação m<sup>6</sup>A em RNAs ribossomais (rRNAs), que podem gerar um grande ruído e promover baixa resolução e difícil interpretação dos dados (Dominissini et al., 2013). Uma vez que os pri-miRNAs são exclusivamente acumulados no núcleo, enquanto que os rRNAs após expressos se concentram no citoplasma, a primeira etapa deste procedimento consiste em segregar as frações nuclear e citoplasmática para purificar amostras de RNA de maneira isolada.

Para isso, foi utilizado o kit *Cytoplasmic and Nuclear RNA Purification Kit* (Norgen, Canadá). Diversas etapas de padronização e modificações ao protocolo fornecido pelo kit foram realizadas devido à baixa eficiência obtida em testes iniciais. Após todos os testes, o protocolo padronizado foi definido como descrito abaixo:

Primeiramente, foram adicionados 250  $\mu$ l de *buffer J* contendo 10  $\mu$ l de  $\beta$ -*mercaptoethanol* por ml de *buffer* em até 35 mg de amostras de cérebro ou até 100 embriões 24 hpf coletados. O tecido foi mantido em solução por 10 min no gelo e então homogeneizado utilizando o equipamento *Beadblaster<sup>TM</sup>* (Merck, Alemanha) com *beads* de metal. A solução foi transferida para um tubo de microcentrífuga e centrifugada à 14.000 RPM por 10 min à temperatura ambiente. Após a centrifugação, o sobrenadante, contendo a porção citoplasmática, foi transferido para um novo tubo, enquanto o *pellet*, contendo a porção nuclear foi mantido. Adicionou-se 200 e 400  $\mu$ l de *buffer SK* contendo 10  $\mu$ l de  $\beta$ -*mercaptoethanol* por ml de *buffer* às porções citoplasmática e nuclear, respectivamente. Posteriormente, foram adicionados 200  $\mu$ l de etanol 100% em ambos os tubos e as misturas foram transferidas às colunas de centrifugação. Nesta etapa, a fração nuclear foi passada por uma seringa de 25G 10 vezes, para fragmentar o DNA genômico, evitando o entupimento da coluna e melhorando a eficiência da extração. As colunas foram então centrifugadas a 6.000 RPM e o *flowthrough* descartado. As amostras foram lavadas três vezes com 400  $\mu$ l de *Wash buffer*, seguidos de centrifugação a 14.000 RPM por 1 min a temperatura ambiente. Após a última centrifugação, uma nova centrifugação a 14.000 RPM por 2 min a temperatura ambiente foi realizada para garantir máxima secagem da coluna. As colunas foram então transferidas para novos tubos de microcentrífuga e adicionou-se 50  $\mu$ l de *Elution buffer E*. As colunas foram centrifugadas a 2.000 RPM por 2 min, seguindo de uma centrifugação de 1,5 min a 14.000 RPM, ambas em temperatura ambiente. Uma segunda rodada de eluição foi realizada passando a solução uma segunda vez pela coluna para maximizar a recuperação das amostras de RNA. A integridade das amostras e o sucesso do enriquecimento de RNA nuclear foram verificados por eletroforese em gel de agarose a 1% (Figura 15). Finalmente, as amostras foram armazenadas em Biofreezer -80°C até sua utilização. As frações nucleares foram utilizadas para a realização do sequenciamento m<sup>6</sup>A-seq e as frações citoplasmáticas armazenadas para eventual utilização futura.



**Figura 15. Verificação do sucesso do enriquecimento da fração nuclear nas amostras de RNA.** Resultado da purificação das frações nucleares e citoplasmáticas de quatro amostras de RNA de cérebro. Os 8 primeiros poços consistem nas frações nucleares. Os 8 poços seguintes consistem nas frações citoplasmáticas. Amostras “A” (banda mais espessa) representam a primeira eluição ao final da purificação. Amostras “B” (banda mais branda) representam uma segunda rodada de eluição para maximização da recuperação de RNA das colunas.

As amostras de RNA nuclear foram tratadas com o kit *Turbo DNase* (Ambion, USA) para remoção de DNA genômico seguindo as instruções do fabricante. Para isso, 5  $\mu$ l de *10X TURBO™ DNase Buffer*, 1  $\mu$ l *TURBO™ DNase* e 4  $\mu$ l de *RNAse-free water* foram adicionados a 40  $\mu$ l de cada amostra, completando 50  $\mu$ l de reação, e as amostras foram incubadas por 30 min a 37°C. Foram então adicionadas 5  $\mu$ l de *DNase Inactivation Reagent* e as amostras incubadas por 5 min à temperatura ambiente. Finalmente, as amostras foram centrifugadas 10.000 RPM por 1,5 min a 4°C e o sobrenadante contendo as amostras de RNA purificadas armazenadas em *biofreezer* a -80°C até sua utilização.

A concentração e a pureza das amostras obtidas foram medidas utilizando-se o equipamento *NanoDrop 1000* (Thermo Scientific ©, USA) para assegurar a quantidade mínima de 20  $\mu$ g por replicata demandada para enviar as amostras para a realização do sequenciamento m<sup>6</sup>A-se. A integridade das amostras, após o tratamento com DNase, foi verificada em gel de agarose a 1% (Tabela 3).

**Tabela 3. Concentração e pureza das amostras de RNA nuclear.** Br = amostras de cérebro. En = amostras de embrião 24 hpf. N = fração nuclear. A última letra de cada amostra indica sua replica. A amostra BrNA foi degradada durante a etapa de tratamento por DNase e substituída pela amostra BrND.

| Amostras    | 260/280 | 260/230 | [ ng/ $\mu$ l ] | Total (ng) |
|-------------|---------|---------|-----------------|------------|
| <b>BrNB</b> | 1,90    | 1,91    | 225,1           | 27.012     |
| <b>BrNC</b> | 1,89    | 1,96    | 213,8           | 25.625     |
| <b>BrND</b> | 1,88    | 1,90    | 203,1           | 24.372     |
| <b>EnNA</b> | 1,90    | 1,69    | 133,3           | 21.328     |
| <b>EnNB</b> | 1,91    | 1,26    | 108,4           | 21.680     |
| <b>EnNC</b> | 1,87    | 1,51    | 111,9           | 22.380     |

As amostras foram enviadas à empresa LC Sciences (Texas, EUA) para execução do serviço contratado m<sup>6</sup>A-seq, que englobou a realização das técnicas de imunoprecipitação, preparação das bibliotecas de cDNA e sequenciamento. Para isso, as amostras foram preparadas adicionando-se 1/10 de volume de acetato de sódio 3 M (pH 5.2), 1 $\mu$ l glicogênio (100  $\mu$ g ml<sup>-1</sup>) e 2.5 volumes de etanol 100% e armazenando-as em gelo seco.

#### 4.1.4. Imunoprecipitação m<sup>6</sup>A, preparação das bibliotecas de cDNA e m<sup>6</sup>A-seq

Assim que as amostras foram recebidas pela empresa, elas foram purificadas por precipitação por etanol 75% e sua concentração e integridade foram avaliadas pelo equipamento *Agilent 2100 Bioanalyzer*, com o kit *RNA 6000 Nano LabChip Kit* (Agilent Technologies, USA). Com a qualidade das amostras assegurada, o m<sup>6</sup>A-seq foi realizado no conjunto total de seis amostras (Tabela 4), sendo três as triplicatas biológicas analisadas por grupo experimental.

**Tabela 4. Qualidade das amostras de RNA nuclear.** Br = amostras de cérebro. En = amostras de embrião 24 hpf. N = fração nuclear. A última letra de cada amostra indica sua replica. Amostras avaliadas pela empresa LC Sciences após sua purificação por etanol 75%.

| Amostras    | 260/280 | 260/230 | [ ng/µl ] | Total (ng) | RIN  |
|-------------|---------|---------|-----------|------------|------|
| <b>BrNB</b> | 1,80    | 2,54    | 0,195     | 12.690     | 8.80 |
| <b>BrNC</b> | 1,82    | 2,53    | 0,175     | 11.350     | 6.70 |
| <b>BrND</b> | 1,77    | 2,56    | 0,182     | 11.800     | 8.80 |
| <b>EnNA</b> | 1,80    | 2,42    | 0,185     | 12.050     | 9.00 |
| <b>EnNB</b> | 1,85    | 2,40    | 0,172     | 11.200     | 9.40 |
| <b>EnNC</b> | 1,81    | 2,41    | 0,188     | 12.190     | 9.70 |

Primeiramente, as amostras de RNA foram fragmentadas em oligonucleotídeos de ~100 nucleotídeos. Para isso, 2 µl de *fragmentation buffer 10X* (800 µl de *molecular biology-grade, RNase-free water* contendo 100 µl de Tris-HCl 1M (pH 7.0) e 100 µl de ZnCl<sub>2</sub> 1M) foram adicionados a tubos de coleta contendo 18 µl de amostra, completando 20 µl de reação. As soluções foram incubadas a 94°C por 5 min e posteriormente foram adicionados 2 µl de EDTA 0.5 M e levadas ao gelo. As soluções pertencentes a uma mesma amostra foram então reagrupadas e foram adicionadas 1/10 de volume de acetato de sódio 3M (pH 5.2), 1 µl de glicogênio 100 µg ml<sup>-1</sup> e 2.5 volumes de etanol 100% e foram incubadas em *biofreezer -80°C overnight*. Posteriormente, os fragmentos de RNA foram purificados por precipitação por etanol 75%.

Das amostras fragmentadas, uma fração foi separada e armazenada em *biofreezer -80°C* para ser utilizada como controle, enquanto o restante foi utilizado para a reação de imunoprecipitação. Para isso, o volume das amostras de RNA foram ajustadas para 755 µl com *RNase-free water* e foram adicionadas 10 µl de *RNAasin* (40U µl<sup>-1</sup>), 10 µl RVC (200 mM), 200 µl de em solução *IP buffer* (50 mM Tris-HCl, 750 mM NaCl e 0.5% Igepal CA-630) e 25 µl de anticorpos específicos anti-m<sup>6</sup>A (0.5 mg ml<sup>-1</sup>; No. 202003, Synaptic Systems, Germany) e a solução foi incubada por 2h a 4°C. Uma reação paralela a esta foi realizada, contendo a mesma quantidade de RNA fragmentado porém sem a presença do anticorpo específico anti-m<sup>6</sup>A para avaliar os níveis de ruído e eficiência da técnica de imunoprecipitação.

Durante o período de incubação as *protein-A beads* foram lavadas com 1ml de *IP buffer 1x*, ressuspensas em *IP buffer 1x* suplementado com BSA (0.5 mg ml<sup>-1</sup>) e divididas em 2

tubos de microcentrifugação. Após a incubação, as amostras de ambos os tratamentos foram transferidas para as soluções contendo as *protein-A beads* e foram novamente incubadas por 2h em rotação a 4°C. As *beads* foram então precipitadas e o sobrenadante removido e armazenado para verificação de qualidade das amostras. As *beads* foram então lavadas três vezes com solução *IP buffer 1x* e foram adicionados 100 µl de *elution buffer* (IP buffer 1x e 6.7mM de m<sup>6</sup>A livres) em cada amostra. As soluções foram incubadas por 1h em movimentação contínua a 4°C, as *beads* foram centrifugadas e o sobrenadante contendo os fragmentos de RNA foram recuperados. Foram adicionados 1/10 de volume de acetato de sódio 3 M (pH 5.2) e 2.5 volumes de etanol 100% e as amostras foram incubadas a -80°C *overnight*. Finalmente, o RNA foi purificado por precipitação por etanol 75%.

As amostras de RNA imunoprecipitados (IP) com e sem anticorpo anti-m<sup>6</sup>A foram quantificadas utilizando-se o kit *Quant-iT RNA assay kit* (Invitrogen, EUA), na qual confirmou-se o sucesso das imunoprecipitações e a não detecção de RNA presente no tratamento sem os anticorpos anti-m<sup>6</sup>A. As amostras de RNA IP e controle foram então convertidas em cDNA para a preparação das bibliotecas, com fragmentos de ~ 100±50 pares de base. Ao final, o sequenciamento paired-end (2×150bp) foi realizado utilizando-se a plataforma *Illumina Novaseq™ 6000*.

#### 4.1.5. Tratamento dos dados de m<sup>6</sup>A-seq e identificação de picos de metilação

Uma vez que os dados de m<sup>6</sup>A-seq foram obtidos, as amostras foram avaliadas quanto à qualidade pela ferramenta *FastQC* (v0.11.5; Wingett e Andrews, 2018). Posteriormente, as sequências de adaptadores foram removidas utilizando-se a ferramenta *cutadapt* (v1.14; Martin, 2011), as *reads* foram filtradas por um *script* em Perl pela qualidade de suas bases e pela presença de bases indeterminadas e as sequências foram convertidas de Fastq para Fasta utilizando-se a ferramenta *FASTX-Toolkit*. As *reads* filtradas foram mapeadas no genoma de *zebrafish* (vGRCZ11) utilizando a ferramenta *Bowtie* (Langmead e Salzberg, 2012).

Para a identificação dos picos de metilação e a ocorrência de metilações diferenciais entre as amostras (p-value ≤ 0.05 e FDR < 5%), foi utilizada a ferramenta *exomePeak* (Meng et al., 2014) do pacote *edgeR* (R Bioconductor). A anotação dos picos identificados foi realizada pela ferramenta *ChIPseeker* (Yu et al., 2015). A ferramenta *StringTie* (Pertea et al., 2015) foi utilizada para calcular os perfis de expressão de mRNAs, utilizando as bibliotecas controle

como referência e os mRNAs diferencialmente expressos entre os tecidos foram identificados ( $\log_2 \text{fold change} >1$  ou  $< -1$  e  $p \text{ value} \geq 0.05$ ).

A identificação das sequências *motif* de metilação m<sup>6</sup>A presente nos dados oriundos do m<sup>6</sup>A-seq foi realizada utilizando-se a ferramenta *Meme* (Bailey et al., 2009). Para isso, foram utilizadas regiões flanqueadoras de  $\pm 50$ nts a partir do pico de metilação dos 1.000 picos com melhor *score* identificados e a ferramenta *Meme* calculou as 3 principais sequências consenso enriquecidas nas amostras. A presença da sequência *motif* RRAC entre as sequências enriquecidas foi então verificada manualmente. Finalmente, a ferramenta *CentriMo* (Bailey et al., 2012) foi utilizada para determinar a distância entre as sequências *motif* identificada e os picos de metilação.

Para a identificação de picos de metilação associados a miRNAs uma etapa adicional foi realizada. Uma vez que sítios de metilação de pri-miRNAs muitas vezes encontram-se fora das regiões anotadas de pré-miRNAs (Alarcón et al., 2015a), um *script* em *Python* por nós desenvolvido foi utilizado para identificar picos de metilação em regiões flanqueadoras de miRNAs anotados. Para isso, regiões flanqueadoras de  $\pm 500$ ,  $\pm 1.000$  e  $\pm 10.000$  nts a partir do nucleotídeo central dos miRNAs foram consideradas e os resultados obtidos em cada grupo comparados. A identificação de picos nestas regiões foi realizada utilizando-se a ferramenta *MAC2S* (Zhang et al., 2008) com a função *callpeak* ( $p\text{-value} \leq 0.05$  e  $\text{FDR} < 5\%$ ). Os resultados obtidos foram somados aos dados obtidos pela ferramenta *exomePeak* para uma identificação completa de eventos de metilação em miRNAs. Para a identificação de sequências *motif* presentes nas sequências de pri-miRNAs, a ferramenta *Meme* foi utilizada levando-se em consideração as regiões flanqueadoras de  $\pm 1.000$  nts de todos os miRNAs anotados.

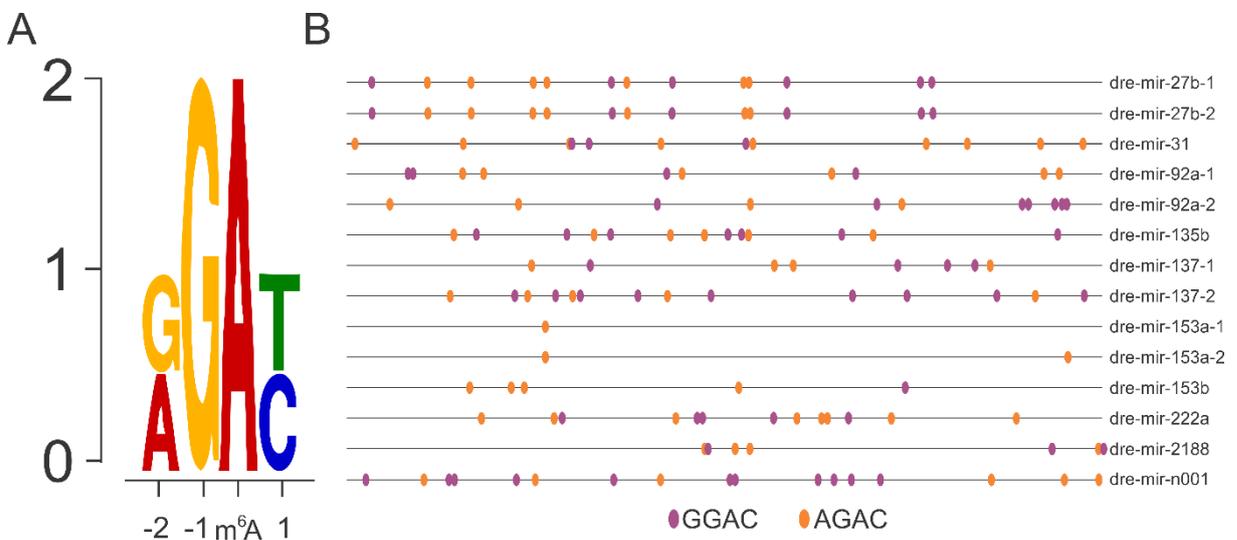
## 4.2. Resultados e discussão

### 4.2.1. Sequências *motif* de metilação m<sup>6</sup>A estão presentes nos pri-miRNAs de *zebrafish*

Para verificar se a metilação m<sup>6</sup>A poderia atuar na ocorrência dos eventos de *arm switching*, primeiramente realizamos uma busca por sequências *motif* desta metilação nos pri-miRNAs de *zebrafish*. Devido à dificuldade de diferenciar pri-miRNAs de outros tipos de RNA, são poucos os trabalhos que avaliaram o comprimento médio de pri-miRNAs, sendo casos pontuais e específicos. Lee et al. (2002) identificaram pri-miRNAs com comprimento aproximado de 403 nts para o cluster *mir-23/27/24-2* e 791 nts para o cluster *mir-*

17/18/19a/20/19b-1. Adicionalmente, estes autores também verificaram o comprimento aproximado de 600 nts para o *miR-30a*. Outros autores detectaram ainda um comprimento de aproximadamente 3.5kb do *pri-miR-21* (Cai et al., 2004). Com isso, consideramos como pri-miRNAs sequências de 2kb, com  $\pm 1000$  nts flanqueadores a partir do nucleotídeo central do pré-miRNA anotado. (Lee et al., 2002; Cai et al., 2004).

Esta análise identificou uma sequência consenso RGAY (R = A/G; Y = C/U), compatível com a sequência *motif* RRAC descrita na literatura (Figura 16a) para a metilação m<sup>6</sup>A, enriquecida nos pri-RNAs de *zebrafish*. De fato, as sequências GGAC e AGAC, principais sequências de reconhecimento do complexo N6-metiltransferase (Meiser et al., 2020), foram detectadas em todos os pri-miRNAs dos miRNAs com eventos de *arm switching* identificados (Figura 16b).



**Figura 16. Identificação de motifs m<sup>6</sup>A em pri-miRNAs.** A) Sequência consenso RGAY identificada nos pri-miRNAs. B) Sequências GGAC (em roxo) e AGAC (em laranja) identificadas nos 14 pri-miRNAs com eventos de *arm switching* identificados.

Uma vez que foram observados potenciais sítios de metilação m<sup>6</sup>A nos pri-miRNAs, o m<sup>6</sup>A-seq foi realizado nas amostras de cérebro e embrião 24 hpf para identificarmos os padrões de enriquecimento destas metilações nestes tecidos e avaliarmos se elas poderiam estar influenciando a ocorrência de *arm switching* nos miRNAs.

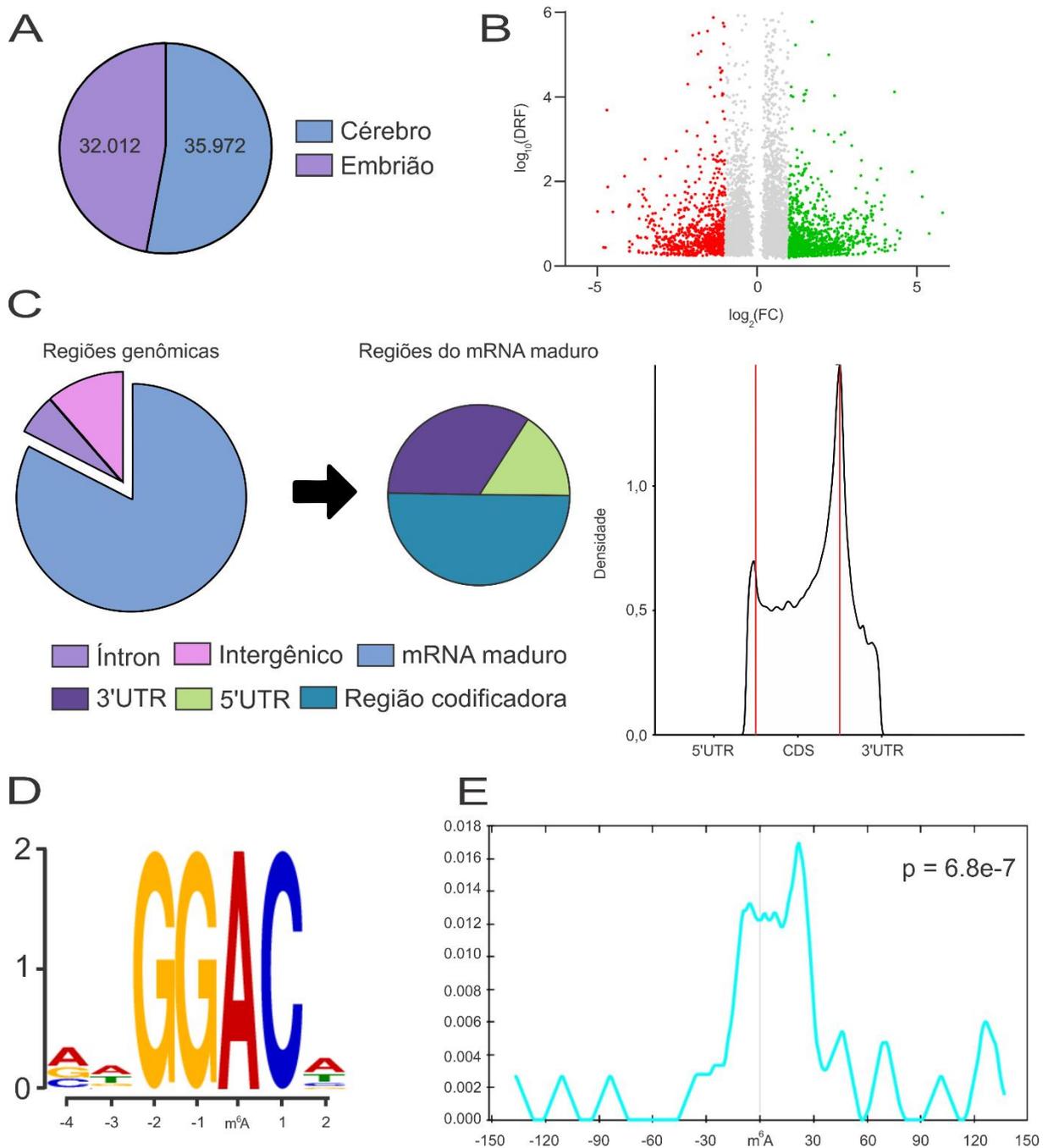
#### 4.2.2. Visão geral dos dados de m<sup>6</sup>A-seq

Com o sequenciamento das amostras, foram compilados aproximadamente 78,5 bilhões de *raw reads*, que foram avaliadas quanto à qualidade e tiveram seus adaptadores removidos, obtendo-se um total de aproximadamente 76,6 bilhões de *reads* válidas e mapeadas no genoma.

Foram identificados um total de 35.972 picos de metilação m<sup>6</sup>A enriquecidos no cérebro e 32.012 no embrião, caracterizando um número similar de picos totais identificados em ambos os tecidos (Figura 17a; Tabelas Suplementares S7 e S8). Adicionalmente, foram detectadas, 5.155 regiões metiladas em ambos dos tecidos, das quais, 1.057 encontravam-se hipermetiladas no cérebro em relação ao embrião ( $\log_2(fc) \geq 1$ ), 862 encontravam-se hipermetiladas no embrião 24hpf em relação ao cérebro ( $\log_2(fc) \leq -1$ ) e 3.236 encontravam-se igualmente metilados em ambos os tecidos ( $\log_2(fc) \leq 1$  e  $\geq -1$ ; Figura 17b; Tabela Suplementar S9). Estes dados mostram que apesar do número de picos metilados terem sido similares nos diferentes tecidos estudados, há uma alta variabilidade quanto aos padrões e níveis de metilação, sendo certo que apenas 9% do número total de regiões metiladas encontraram-se igualmente enriquecidas em ambos os tecidos.

Ao analisar a distribuição dos picos de metilação sobre diferentes regiões do genoma, detectamos um alto enriquecimento em regiões referentes a mRNAs maduros se comparadas ao observado em regiões intrônicas e intergênicas para ambos os tecidos. Para o mRNA maduro, identificamos aproximadamente 16% dos picos na região 5'UTR, 35% na região 3'UTR e 49% na região codificadora, com um acúmulo próximo ao *stop codon* (Figura 17c). Tais dados corroboram com os achados descritos outros trabalhos sobre a distribuição dos picos de metilação m<sup>6</sup>A em embriões de *zebrafish* (Zhang et al., 2017) e em linhagens de células humanas e de camundongo (Dominissini et al., 2012; Meyer et al., 2012; Mao et al., 2019).

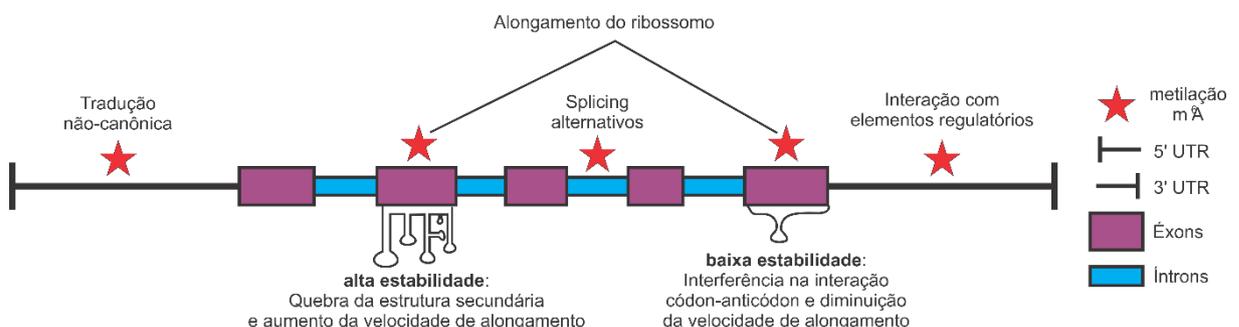
Metilações m<sup>6</sup>A localizadas em regiões intrônicas possuem um papel importante no processamento e no *splicing* de mRNAs (Ru et al., 2020). O nocaute da enzima *Mtll3*, é capaz alterar a ocorrência de *splicings* (Ping et al. 2014), enquanto que diversas proteínas *readers* desta metilação como *Ythdc1*, *hnRNPC* e *hnRNPA2B1* podem promover *splicings* alternativos nas moléculas de mRNA (Ping et al., 2014; Xiao et al., 2016; Liu et al., 2015; Liu et al., 2017).



**Figura 17. Panorama geral dos dados de m<sup>6</sup>A-seq.** A) Identificação de picos de metilação nas amostras de cérebro e embrião 24 hpf. B) Vulcano plot demonstrando os picos diferencialmente metilados entre as amostras de embrião 24 hpf (em vermelho) e cérebro (em verde). Picos detectados em ambas as amostras, porém sem níveis de metilação diferenciais estão representados em cinza. C) Distribuição dos picos de metilação ao longo do genoma. Linhas vermelhas indicam as regiões do mRNA maduro com maior densidade de picos de metilação. D) Sequência consenso GGAC identificada, correspondendo com a sequência *motif* de metilação m<sup>6</sup>A descrita na literatura. E) Posição relativa das sequência consenso em relação aos picos de metilação identificados. Os picos de metilação identificados concentram-se sobre a sequência consenso GGAC identificada.

A distribuição diferencial dos picos de metilação em regiões distintas do mRNA maduro é um indicador de uma ocorrência de funções divergentes em cada uma delas (Figura 18; Mao et al., 2019). Estudos mostram que metilações m<sup>6</sup>A presentes na região 5'UTR podem promover eventos de tradução não-canônicos, de maneira Cap-independente, com a participação dos genes *eIF3* e *eIF4F* (Meyer et al., 2015; Coots et al., 2017). Já a região 3'UTR é composta por um grande número de sítios de interação com ncRNAs, que promovem regulações pós-transcricionais no mRNA, como miRNAs, por exemplo. Chen et al. (2015) demonstram que um elevado número de sequências *motifs* de metilação m<sup>6</sup>A RRACH possuem alta complementaridade com a sequência *seed* de diversos miRNAs e que, de fato, não só estes miRNAs são capazes de parear com as regiões *motifs*, como seu pareamento pode promover a metilação destas regiões. Devido à grande capacidade destas metilações em alterar a estrutura secundária das regiões que estão inseridas, seja por romper o pareamento da adenina metilada, seja por aumentar sua estabilidade (Liu et al., 2017; 2018), as metilações m<sup>6</sup>A formadas podem ainda ser capazes de alterar a acessibilidade destes e de outros sítios de interações nas vizinhanças, promovendo uma complexa modulação na interação entre os miRNAs e seus mRNA alvos.

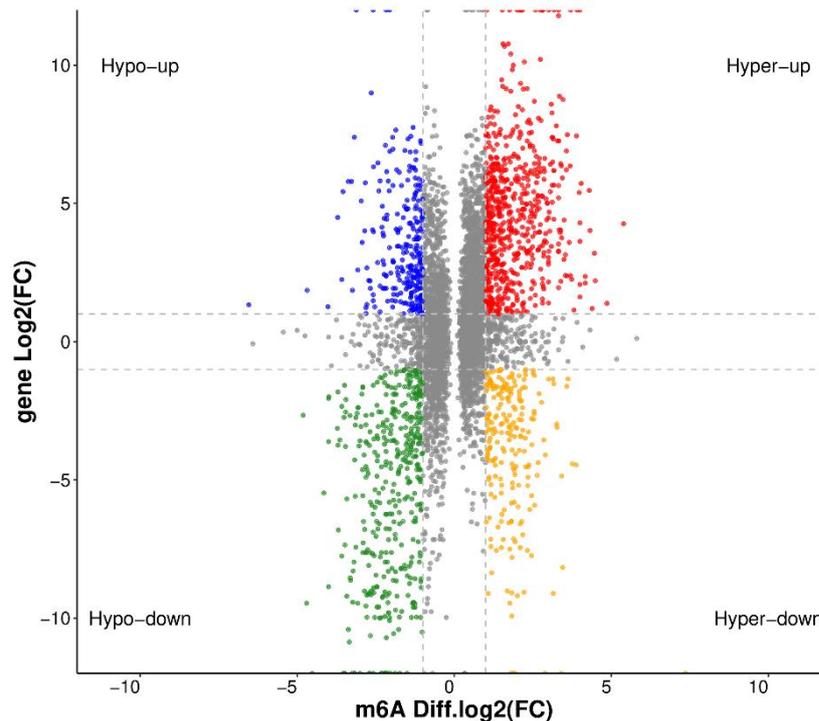
No caso dos dados de *zebrafish*, praticamente metade dos picos identificados estão presentes na região CDS. A presença de metilações m<sup>6</sup>A nesta região interfere diretamente na eficiência da tradução dos mRNAs, impactando na velocidade de alongamento dos ribossomos (Mao et al., 2019). Enquanto a metilação m<sup>6</sup>A, por si só prejudica a eficiência do alongamento por interferir no pareamento códon-anticódon, sua presença promove a desestabilização de estruturas secundárias, facilitando a passagem do ribossomo (Mao et al., 2019). Desta maneira, a metilação m<sup>6</sup>A no CDS pode apresentar funções antagônicas, a depender da região inserida, servindo como um resistor em regiões com estruturas secundárias instáveis, ou como um condutor em regiões com estruturas secundárias estáveis.



**Figura 18. Funções biológicas da metilação m<sup>6</sup>A ao longo do mRNA.**

Ao buscarmos pelas sequências consenso presentes nos picos de metilação, a sequência GGAC identificada corresponde à descrita na literatura como sequência *motif* de reconhecimento da enzima *Mtll3* para a ocorrência da modificação m<sup>6</sup>A (Figura 17d; Dominissini et al., 2012). Adicionalmente, a posição destas sequências nas *reads* mapeadas batem com os picos de metilação identificados (Figura 17e), comprovando o enriquecimento das metilações nestas sequências.

Adicionalmente, a análise comparativa dos picos de metilação associados à expressão dos genes metilados identificou diversos genes hipo- e hiper-metilados diferencialmente expressos entre os tecidos (Figura 19; Tabela Suplementar S10).

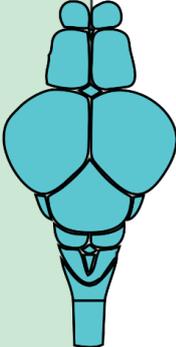


**Figura 19. Genes diferencialmente expressos e diferencialmente metilados entre o cérebro e embrião 24hpf.** Pontos azuis representam genes hipo-metilados e *up-regulated* no cérebro em relação ao embrião. Pontos vermelhos representam genes hiper-metilados e *up-regulated* no cérebro em relação ao embrião. Pontos verdes representam genes hipo-metilados e *down-regulated* no cérebro em relação ao embrião. Pontos amarelos representam genes hiper-metilados e *down-regulated* no cérebro em relação ao embrião.

Estes dados de correlação entre os níveis de metilação e expressão diferenciais permitirão a execução de análises futuras sobre os impactos biológicos associados à ocorrência da metilação m<sup>6</sup>A nas condições biológicas avaliadas. Tais análises preveem avaliar em detalhes o perfil funcional dos transcritos expressos para compreender os papéis biológicos que a metilação m<sup>6</sup>A estaria desempenhando nestes tecidos. Adicionalmente, os dados gerados permitem ainda avaliarmos como a ocorrência de metilações m<sup>6</sup>A em regiões próximas a MREs interferem no reconhecimento dos alvos de miRNAs.

#### 4.2.3. Não há indícios da ocorrência efetiva de metilação m<sup>6</sup>A em pri-miRNAs de zebrafish, apesar do enriquecimento de sequências *motif*

Foram detectados picos de metilação em 14 miRNAs nas amostras de cérebro e 14 miRNAs nas amostras de embrião (Figura 20). Destes, 3 miRNAs apresentaram picos em ambos os tecidos (*dre-mir-219-1*, *dre-miR-454a* e *dre-mir-2193*). Inesperadamente, não foram identificados picos de metilação em nenhum dos sete miRNAs nos quais foram detectados eventos de *arm switching* ocorrendo entre os tecidos de cérebro e embrião 24 hpf.

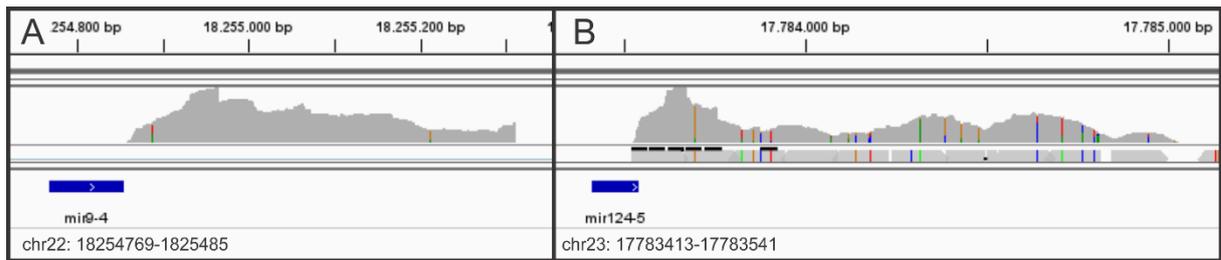
|   |  |   |
|---|--|---|
|  | dre-mir-9-2<br>dre-mir-9-4<br>dre-mir-24-1<br>dre-mir-30c<br>dre-mir-101b<br>dre-mir-124-5<br>dre-mir-124-6<br>dre-mir-125c<br><b>dre-mir-219-1*</b><br>dre-mir-375-2<br><b>dre-mir-454a</b><br>dre-mir-735<br>dre-mir-736<br><b>dre-mir-2193</b>  |  |
|   | dre-mir-9-1<br>dre-mir-10b<br>dre-mir-15a-1<br>dre-mir-15a-2<br>dre-mir-16b<br>dre-mir-23b<br>dre-mir-124-4<br>dre-mir-129-4<br><b>dre-mir-219-1*</b><br><b>dre-mir-454a</b><br>dre-mir-728<br>dre-mir-2188<br>dre-mir-2191<br><b>dre-mir-2193</b> |   |

**Figura 20.** Lista de miRNAs com picos de metilação m<sup>6</sup>A detectados. Na esquerda, em amarelo, miRNAs metilados no embrião. Na direita, em azul, miRNAs metilados no cérebro. Nos miRNAs destacados e em negrito foram encontrados picos de metilação em ambos os estágios de vida analisados. Para o *\*dre-mir-219-1* foram detectados picos de metilação a ~950 bases do nucleotídeo central do miRNA.

Nossas análises detectaram a presença de picos de metilação em apenas 3% dos miRNAs descritos em *zebrafish*. Estes resultados divergem dos dados encontrados para humano, no qual 70% dos miRNAs apresentaram variações em seus níveis de expressão após o silenciamento da enzima *Mtll3* (Alarcón et al., 2015a). Adicionalmente, um estudo recentemente publicado, contemplando uma ampla biblioteca de modificações m<sup>6</sup>A no transcriptoma de sete espécies, incluindo humano, camundongo, rato, *zebrafish*, *Drosophila*, *Arabidopsis* e levedura, destaca que há uma baixa conservação de sítios de metilação m<sup>6</sup>A entre os vertebrados (apenas 1,641% de sítios identificados em comum entre humanos e *zebrafish*; Tang et al., 2021). Tais dados sugerem uma possível divergência na ocorrência e na frequência de metilações m<sup>6</sup>A em miRNAs, nos diferentes clados de vertebrados.

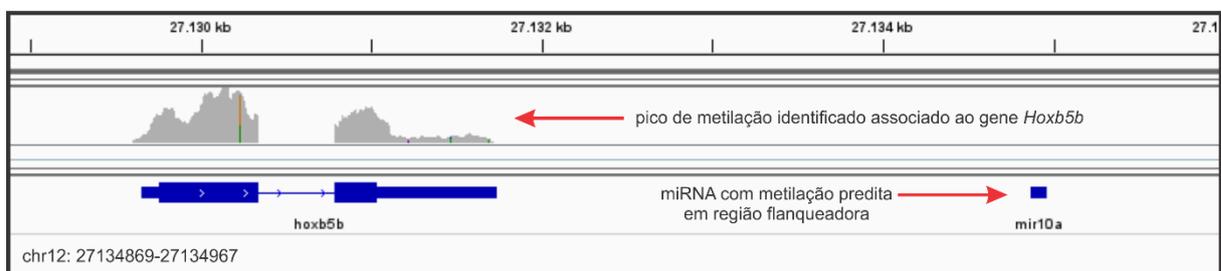
Uma vez que detectamos um baixo número de eventos de metilação nos miRNAs, buscamos identificar as prováveis razões que poderiam estar promovendo tal ocorrência. Primeiramente, questionamos se o comprimento inicialmente selecionado de 2kb para a sequências de pri-miRNAs não poderia estar restringindo o número de picos detectados, por estar excluindo picos de metilação genuínos que acabaram por ficar fora da região delimitada. Desta maneira, realizamos uma análise de identificação de picos de metilação em regiões flanqueadoras de miRNAs assumindo a priori dois outros comprimentos distintos de pri-miRNAs e comparamos estes resultados aos anteriores. Um comprimento avaliado constituiu-se de regiões flanqueadoras de  $\pm 500$ b a partir do nucleotídeo central do pré-miRNA anotado e o outro fez um extensivo segmento de  $\pm 10$ kb, a partir do nucleotídeo central do miRNA, constituindo-se em pri-miRNAs de comprimentos hipotéticos de 1kb e 20kb respectivamente. Estes foram avaliados em adição aos pri-miRNAs inicialmente utilizados de 2kb.

Os pri-miRNAs de 1kb e 2kb retornaram um número similar de picos de metilação identificados, sendo que apenas o *dre-miR-219-1* foi identificado exclusivamente nos pri-miRNAs de 2kb, com o pico identificado a -936 nt de seu nucleotídeo central no embrião 24 hpf e -969 nt distante no cérebro, (Figura 20). Ao checarmos visualmente a localização destes picos de metilação em relação a região anotada do miRNA identificado, confirmamos que estas metilações estavam de fato contidas em regiões não anotadas de seus pri-miRNAs (Figura 21). Entretanto, quando realizamos o teste *T de student*, verificamos que a frequência observada de picos de metilação em miRNAs corresponde à frequência esperada para uma distribuição aleatória de picos de metilação no genoma, em ambos os casos ( $p = 0.8536$  para os pri-miRNAs de 1kb e  $p = 0.4861$  para os pri-miRNAs de 2kb).



**Figura 21. Picos de metilação associados à pri-miRNAs.** Visualização das posições relativas dos picos de metilação com a região anotada dos miRNAs associados nos pri-miRNAs de 1kb e 2kb. Os picos de metilação identificados encontra-se exclusivamente adjacentes à anotação do miRNA. A) Pico de metilação identificado no dre-miR-9-4 na amostra de cérebro. B) Pico de metilação identificado no dre-miR-124-5 na amostra de embrião 24 hpf.

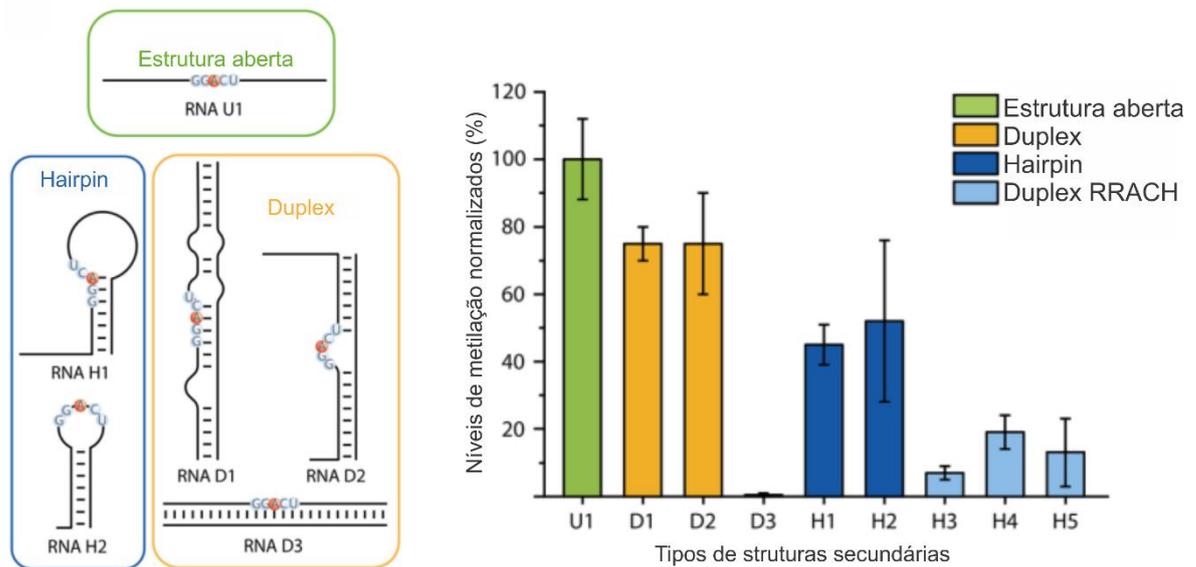
Já para os pri-miRNAs de 20kb foram detectados aproximadamente 55 casos de metilação m<sup>6</sup>A. Apesar do elevado comprimento da região avaliada, o número de casos detectados ainda foi baixo, contemplando 13% dos miRNAs descritos. Destaca-se que o teste *T de student* mostrou ainda que tal valor é abaixo do esperado para uma distribuição aleatória ( $p = 0.0114$ ), demonstrando que mesmo um aumento expressivo no comprimento da região avaliada não foi suficiente para recuperar um aumento significativo no número de miRNAs metilados. Além disso, ao avaliarmos visualmente a localização picos adicionais obtidos neste cenário, verificamos que, na verdade, estes picos correspondiam a regiões de mRNAs metilados, caracterizando-os como resultados falso-positivo (Figura 22).



**Figura 22. Picos de metilação falso positivos associados à pri-miRNAs.** Visualização das posições relativas dos picos de metilação com a região anotada dos miRNAs associados nos pri-miRNAs de 20kb. Os picos de metilação identificados encontra-se em regiões anotadas de outros genes.

Adicionalmente, questionamos se esta aparente ausência de um número significativo de metilações em pri-miRNAs não estaria associada a um possível insucesso no enriquecimento da fração nuclear durante a etapa de purificação de RNA das amostras. Embora seu sucesso tivesse sido visualmente confirmado via gel de agarose (Figura 15), uma possível contaminação de fração citoplasmática, imperceptível durante a visualização da corrida das amostras em gel de agarose, poderia estar impossibilitando a identificação de metilações presentes na fração nuclear. Felizmente, ao analisarmos os resultados de picos de metilação identificados, detectamos a presença de picos enriquecidos em diversos RNAs expressos exclusivamente no núcleo em ambas as amostras, como moléculas de RNAs codificadores de proteínas como *Ints12*, *Eftud2*, *Coil*, *Prpf4*, *Nop10*, entre outras e moléculas de snRNAs como *Snrpa1*, *Snrpb2* e *Snrnp70* (Tabela Suplementar S7 e S8). Desta maneira, a presença da anotação destes RNAs nos dados referentes aos picos de metilação identificados confirmam o sucesso de nossas análises em detectar metilações em moléculas presentes exclusivamente no núcleo, como é o caso dos pri-miRNAs, minimizando a possibilidade de a aparente ausência de metilações nos pri-miRNAs de zebrafish estar associada a uma falha metodológica.

Com isso, apesar de sítios *motif* de metilação m<sup>6</sup>A estarem presentes em pri-miRNAs, nossos resultados fornecem indícios de que estas moléculas não são efetivamente metiladas em *zebrafish*. Desta maneira, é possível que algum mecanismo esteja impossibilitando a interação da enzima *Mttl3* com os pri-miRNAs. Um artigo recentemente publicado analisou em detalhes a estrutura tridimensional do complexo *Mttl3-Mttl14* e como ocorre sua interação com as moléculas de RNA *in vitro* para promover a reação catalítica de metilação da adenina (Meiser et al., 2020). Estes autores verificaram que em adição à sequência *motif*, a estrutura secundária da região do RNA a ser metilada é um fator determinante para a ocorrência da metilação. A presença de sítios de metilação em estruturas secundárias “frágeis” (i.e. com alta energia livre), já são suficientes para diminuir os níveis de metilação da região em ~25%; sua presença em estruturas *hairpin* podem afetar a eficiência da reação em ~50% enquanto que a presença de sítios em estruturas secundárias “fortes” (i.e. com baixa energia livre) bloqueiam totalmente a ocorrência de metilação (Figura 23; Modificada de Meiser et al. (2020). Desta maneira é plausível que os sítios presentes nos pri-miRNAs de *zebrafish* não estejam localizados em regiões favoráveis para que o complexo da enzima *Mttl3* possa promover sua metilação. Isso justificaria nossos achados sobre a ocorrência de picos de metilações em miRNAs em frequência similares às esperadas para eventos aleatórios.



**Figura 23. Níveis de metilação m<sup>6</sup>A em diferentes estruturas secundárias.** Verde = níveis de metilação de sequências GGACU em estruturas abertas de RNA. Laranja = níveis de metilação de sequências GGACU em estruturas de duplex de RNA. Azul escuro = níveis de metilação de sequências GGACU em estruturas em *hairpin*. Azul claro = níveis de metilação de outras sequências RRACH em estruturas de *hairpin*. Figura modificada e resultados obtidos por Meiser et al. (2020).

De maneira oposta aos resultados obtidos em nossas análises, a ocorrência efetiva de metilações m<sup>6</sup>A em pri-miRNAs foi identificada por outros autores em duas espécies. Adicionalmente, estas modificações foram capazes de aumentar a eficiência do processamento dos pri-miRNAs pela *Drosha* (Alarcón et al., 2015a; Han et al., 2019; Bhat et al., 2020). Alarcón et al. (2015a; 2015b) demonstram que, em linhagens celulares humanas, os pri-miRNAs são metilados durante sua transcrição pela enzima *Mtll3* e que esta modificação de RNA auxilia no reconhecimento da molécula pelo complexo da enzima *Drosha-Dgcr8*, pela interação com a enzima hnRNPA2B1, que promove a ponte entre o pri-miRNA metilado e o complexo *Drosha-Dgcr8* e aumentando a eficiência de seu processamento. Estes autores mostram ainda que o nocaute da enzima *Mtll3*, nesta espécie, foi capaz de diminuir a expressão global de 70% dos miRNAs, caracterizando uma clara atuação da metilação m<sup>6</sup>A na biogênese de miRNAs em humanos. Tal interação foi ainda confirmada por outro trabalho em linhagens de células tumorais (Han et al., 2019). Mais recentemente, a ocorrência de metilações m<sup>6</sup>A em pri-miRNAs e seu impacto no processamento destas moléculas também foram descritos em *Arabidopsis thaliana* (Bhat et al., 2020). No entanto, estes autores verificaram uma atuação mais branda nesta espécie, na qual apenas 25% dos miRNAs avaliados (56 de 230) foram impactados

pelo nocaute da enzima *Mta* (enzima homóloga da *Mttl3*). Destes, uma relação inversa entre acúmulo de pri-miRNAs não processados e redução de miRNAs maduros foram observadas em 20 dos 56 miRNAs. Estes dados sugerem a ocorrência de uma ampla variação na taxa de metilação m<sup>6</sup>A em pri-miRNAs de diferentes espécies.

Dados observados para miRNAs, demonstram que miRNAs “jovens”, em termos evolutivos, tendem a possuir níveis de expressão baixos e pontuais e um número baixo de interações com mRNAs-alvo, enquanto que miRNAs altamente conservados geralmente possuem uma alta e ampla expressão no genoma, bem como um número maior de alvos (Cuperus et al., 2011; Qin et al., 2014; Patel e Capra, 2017). O baixo impacto biológico de miRNAs recentemente formados pode derivar da falta de elementos regulatórios que promovam uma expressão diferencial, bem como a ocorrência de processamentos imprecisos durante sua biogênese, devido ao baixo reconhecimento de suas estruturas secundárias (Cuperus et al., 2011). À medida que uma função é adquirida em um miRNA nascente, processos seletivos tendem a fixá-la e a refinar seu potencial biológico, aumentando a especificidade de sua estrutura, sua expressão e número de alvos ao longo do tempo. Tal processo de ganho de função, fixação e refinamento, pode também estar acontecendo para a ocorrência de metilação m<sup>6</sup>A em pri-miRNAs de diferentes espécies, na qual um aumento de estruturas secundárias favoráveis à metilação de suas sequências *motif* poderia estar ocorrendo.

Uma vez que não há anotações de sequências de pri-miRNAs completos na literatura, não é possível verificarmos se a ocorrência e distribuição de estruturas secundárias nos sítios de metilação dos pri-miRNAs de *zebrafish* poderia estar, de fato, impedindo sua metilação. Isso ocorre pois a utilização da sequência de comprimento correto é imprescindível para a identificação precisa de suas estruturas secundárias, principalmente em sequências longas como o pri-miRNA. O aperfeiçoamento de técnicas de isolamento de pri-miRNAs poderia fornecer os subsídios necessários para a execução destas análises. De maneira similar, experimentos de ganho de função, introduzindo pri-miRNAs metilados sinteticamente em células de *zebrafish*, poderiam averiguar se a metilação m<sup>6</sup>A, caso presente, aumentaria a eficiência do processamento destas moléculas, de maneira similar aos dados observados para humano.

Adicionalmente, trabalhos que avaliam a ocorrência de metilações m<sup>6</sup>A em pri-miRNAs ainda são muito escassos na literatura e estudos envolvendo um número maior de espécies poderiam auxiliar na compreensão de como este mecanismo evoluiu e impacta na eficiência de processamento dos pri-miRNAs.

Uma vez que nossos resultados fornecem indícios de que os pri-miRNAs de *zebrafish* não sofrem modificações do tipo metilação m<sup>6</sup>A, a hipótese inicialmente proposta neste projeto, de que a metilação m<sup>6</sup>A interfere na ocorrência dos eventos de *arm switching* em *zebrafish*, não foi corroborada. Entretanto, não podemos afirmar que metilações m<sup>6</sup>A não impactam eventos de *arm switching* em espécies nas quais esta interação foi verificada. A realização de estudos semelhantes aos realizados neste trabalho com linhagens de células humanas podem auxiliar numa compreensão mais ampla deste fenômeno.

#### **4.2.4. Hipóteses alternativas sobre a ocorrência de *arm switching* em *zebrafish* e possibilidades de pesquisas futuras**

Visto que os eventos de *arm switching* em *zebrafish* não são suficientemente justificados pela ocorrência de metilações m<sup>6</sup>A, buscamos propor outras hipóteses que podem ser futuramente testadas para explicar este efeito.

Nossos resultados demonstram uma possível correlação entre os eventos de *arm switching* identificados e a ocorrência de isomiRs (Capítulo I; Tópico 2.3.5). Desta maneira, outras modificações de RNA em miRNAs poderiam ser um facilitador da ocorrência dos eventos de *arm switching*, por serem capaz de modificar sua estrutura secundária ou seu reconhecimento por proteínas envolvidas em sua biogênese, promovendo cortes distintos pelos complexos *Drosha* e *Dicer*. De fato, um trabalho recentemente publicado demonstrou que a uridilação do pré-miR-324 pelas enzimas *TUT4* e *TUT7* é capaz de reposicionar a *Dicer* no pré-miRNA e alterar seu sítio de clivagem, promovendo o evento de *arm switching* neste miRNA em camundongo (Kim et al., 2020). Outra modificação de RNA, a modificação *A-to-I*, é identificada em diversas regiões dos miRNAs, como ocorrendo em pri-miRNAs, pré-miRNAs e miRNAs maduros (Yang et al., 2006; Kawahara et al. 2007a; Kawahara et al., 2007b; Wulf e Nishikura, 2015). Apesar de um número elevado destas modificações ser capaz de inibir o corte promovido pelas enzimas *Drosha* e *Dicer*, interrompendo a biogênese do miRNA (Yang et al., 2006; Kawahara et al. 2007a), elas também são detectadas em miRNAs maduros (Kawahara et al., 2007b). De fato, eventos de *arm switching* derivados de modificações *A-to-I* já foram observados em linhagens celulares humanas e de camundongos (Iizasa et al., 2010; Li et al., 2018).

Adicionalmente às modificações de RNA, outras moléculas associadas à biogênese dos miRNAs também podem ter um papel importante na ocorrência dos eventos de *arm switching*. O nocaute do TRBP, co-fator da *Ago*, em linhagens celulares humanas foi capaz de alterar o braço preferencialmente incorporado do miR-30a (Kim et al., 2014), no qual o braço 5p é preferencialmente selecionado em linhagens selvagens, ao passo que o braço 3p passou a ser acumulado após o nocaute. Isso ocorre pois durante o processamento do pré-miRNA, há uma associação do TRBP com a Dicer, estabilizando seu processamento (Chendrimada et al. 2005; Haase et al. 2005; Lee e Doudna et al., 2012). Desta maneira, a ausência do TRBP leva a ocorrência de cortes alternativos pela Dicer, promovendo duplexes de miRNAs com instabilidades distintas, promovendo a ocorrência de *arm switching* (Kim et al., 2014).

Paralelamente à ocorrência de eventos de *arm switching* associadas a seleção diferencial dos braços 5p e 3p pela *Ago* durante a biogênese dos miRNA, estudos vem sugerindo que a razão da expressão relativa dos braços miR-5p/miR-3p pode ser amplamente influenciada pela disponibilidade de seus alvos na célula.

Neste contexto, diversos trabalhos demonstraram que os mRNAs-alvo podem promover o acúmulo diferencial dos braços de miRNAs maduros, de maneira independente às propriedades estruturais do duplex em situações específicas, em um processo chamado de *target-mediated miRNA protection* (TMMP) (Chatterjee et al., 2011; Tsai et al., 2016; Zhang et al., 2019), no qual o pareamento dos miRNAs com seus alvos os protegem de degradações por exoribonucleases. Chatterjee et al. (2011) foram os primeiros a descrever este processo e demonstraram que ao super-expressar alvos do braço passageiro do *cel-miR-241* houve um acúmulo deste braço na célula, sem haver modificações no nível de expressão de seu pré-miRNA. De maneira similar, a super-expressão de alvos do braço canônica *cel-miR-241* também potencializa seu acúmulo. Chatterjee et al. (2011) observaram que o braço passageiro do *cel-miR-241* possui uma instabilidade maior em sua porção 5' que o braço canônico do *cel-miR-241*, portanto, deveria ser preferencialmente acumulado, mas que os alvos do braço canônico *cel-miR-241* são mais abundantes. Entretanto, estes autores mostram que o acúmulo diferencial dos braços pelo TMMP ocorre apenas em casos que a instabilidade das pontas do duplex sejam relativamente semelhantes. Ao repetir estas análises para o *cel-let-7*, no qual há uma grande divergência entre as instabilidades do braço canônico e do braço passageiro, este fenômeno não foi observado.

De maneira oposta ao TMMP, outros trabalhos demonstram que os alvos dos miRNAs também são capazes de promover sua degradação em um processo chamado *target-direct miRNA degradation* (TDMD) (Ameres et al., 2010; de la Mata et al., 2015; Fuchs Wightman et al., 2018). Primeiramente descrita por Ameres et al. (2010), a ocorrência de TDMD está associada a um pareamento extensivo da região 3' dos miRNAs em animais. Análises em *Drosophila* e linhagens celulares de mamíferos demonstram que o pareamento com alta complementariedade de sua região 3' pode promover eventos de alongamento 3', clivagem 3' ou degradação do miRNA maduro (Fuchs Wightman et al., 2018). Isso ocorre pois o pareamento complementar da região 3' promove uma má acomodação do miRNA na Ago, expondo sua região 3' para enzimas que promovam sua degradação (Sheu-Gruttadauria et al., 2019). De maneira similar ao TMMP, eventos de arm switching associados a ocorrência de TDMD já foram observados. Zhang et al., (2019) demonstraram que durante a progressão de câncer gástrico, há uma inversão da razão no acúmulo dos braços 5p e 3p pelo TDMD. Adicionalmente, estes autores verificaram que este arm switching da alta complementariedade 3' com seus alvos é um fator determinante para agravar a progressão tumoral.

Desta maneira, a ocorrência dinâmica de ambos TMMP e TDMD ampliam a complexidade da rede regulatória existente entre os miRNAs e seus alvos, aumentando a profundidade da regulação gênica proporcionada pelos miRNAs. Estes fenômenos também poderiam justificar o porquê dos cálculos de incorporação da dos braços na Ago, descritos no Capítulo I, (Tópico 3.2.5), não condizerem com os resultados de expressão observados.

Os achados descritos na literatura, aliados aos resultados obtidos em nossas análises, abrem possibilidades para futuras investigações acerca dos eventos de *arm switching*, e como estes eventos se comportam em diferentes espécies. Análises exploratórias de diversas modificações de RNA, por exemplo, podem permitir a identificação de novos mecanismos que promovam a seleção diferencial dos braços de miRNAs pela Ago, de maneira semelhante ao trabalho realizado por Kim et al. (2020). Adicionalmente, análises de expressão em diferentes tecidos e condições biológicas de *zebrafish*, podem auxiliar na compreensão de como os alvos dos miRNAs atuam na modulação da expressão de seus braços 5p e 3p. Finalmente, estudos comparativos sobre as similaridades e particularidades destes eventos de *arm switching* em diferentes espécies podem fornecer novas perspectivas sobre como eles evoluem e impactam a regulação gênica de seus organismos.

## **5. Conclusões e considerações finais**

## 5.1. Conclusões

Os resultados obtidos neste estudo trazem importantes contribuições para a compreensão de mecanismos regulatórios atuantes nos miRNAs. As análises de investigação de eventos de *arm switching* demonstram que, apesar de seus casos particulares tenderem a ocorrer de maneira espécie-específica, este é um mecanismo majoritariamente associado ao desenvolvimento ontogenético dos vertebrados. A rápida velocidade na modulação da razão da expressão dos braços 5p e 3p e seu potencial de regular conjuntos distintos de funções biológicas, a depender do braço mais expresso, potencializam o controle e manutenção da célula promovido pelos miRNAs durante a intensa variação na expressão gênica ocorrente durante o desenvolvimento dos organismos.

Apesar de não identificarmos os fatores que expliquem o impacto dos isomiRs na seleção dos braços do duplex pela *Ago*, nossos dados sugerem que a expressão diferencial dos isomiRs e a alteração da *read* canônica em diferentes tecidos possa ser um fator importante na ocorrência destes eventos. Adicionalmente, esta expressão diferencial dos isomiRs se apresenta como uma potencializadora do papel regulatório dos eventos de *arm switching*, ampliando o leque de processos e vias biológicas controlados pelos miRNAs.

Surpreendentemente, ao investigarmos o papel da metilação m<sup>6</sup>A na ocorrência dos eventos de *arm switching* detectamos que, diferentemente dos dados identificados para as espécies de humano e *Arabidopsis*, não foram detectados indícios de que pri-miRNAs de *zebrafish* sejam metilados. Além disso, os dados da literatura também mostram altas taxas de variação nos níveis de metilação de pri-miRNAs nas espécies estudadas. Apesar de não termos identificado os fatores moleculares responsáveis por estas diferenças, é possível que a estrutura secundária dos pri-miRNAs seja um importante modulador deste evento. O aprimoramento de técnicas de purificação de pri-miRNAs e experimentos de ganho de função podem auxiliar na compreensão deste fenômeno.

Os dados obtidos neste trabalho ressaltam que apesar de muitos mecanismos regulatórios apresentarem funções conservadas entre os vertebrados, como o já estabelecido papel da regulação de mRNAs promovido pela metilação m<sup>6</sup>A, ou a importância dos eventos de *arm switching* durante o desenvolvimento ontogenético aqui descrita, suas particularidades e casos específicos podem apresentar grandes variações entre diferentes organismos. Com isso, o estudo destes processos em um número cada vez maior de espécies poderá fornecer subsídios

para compreendermos como estes eventos emergem e evoluem e como suas particularidades impactam a regulação gênica em diferentes organismos.

## **5.2. Considerações finais**

Adicionalmente aos resultados obtidos, o desenvolvimento desta pesquisa e a realização das análises aqui descritas foram fundamentais para minha formação. Os desafios superados e discussões realizadas, bem como o aprendizado de diversas técnicas de bancada e computacionais ao longo de todo o período de doutoramento, proporcionaram a aquisição uma de rica experiência na área de biologia e evolução molecular. De maneira similar, as experiências obtidas durante este período me permitiram contribuir de maneira efetiva em diversas outras pesquisas, buscando sempre que possível transmitir os conhecimentos adquiridos aos colegas de laboratório.

A participação em atividades paralelas ao desenvolvimento da pesquisa deste projeto, como a realização de estágios docência, participação de eventos, ministração de palestras e aulas em disciplinas, co-orientação de alunos de graduação e mestrado, entre outras atividades, também foram fundamentais para minha formação acadêmica e profissional.

As recentes restrições oriundas da pandemia Covid-19, que iniciaram em abril de 2020 e perduram até os dias de hoje, prejudicaram a execução de algumas análises funcionais até o período de defesa do doutoramento. Entretanto, o tempo inicialmente destinado a estes experimentos pode ser convertido na execução de análises computacionais inicialmente não previstas, que geraram resultados inéditos e dados que ainda poderão ser explorados futuramente. Adicionalmente, a realização de estágio docência em meio a paralisação das aulas presenciais, no primeiro semestre de 2020, me possibilitou presenciar os desafios e mudanças didáticas requeridas no EAD, complementando a minha formação na área de educação.

## 6. Referências Bibliográficas

- Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife*. 12;4.
- Ahkin Chin Tai JK, Freeman JL (2020) Zebrafish as an integrative vertebrate model to identify miRNA mechanisms regulating toxicity. *Toxicol Rep*. 7:559-570.
- Aiello NM, Bajor DL, Norgard RJ, Sahmoud A, Bhagwat N, Pham MN, Cornish TC, Iacobuzio-Donahue CA, Vonderheide RH, Stanger BZ (2016) Metastatic progression is associated with dynamic changes in the local microenvironment. *Nat Commun*. 7:12819.
- Alarcón CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF (2015a) N6-methyladenosine marks primary microRNAs for processing. *Nature*. 519(7544):482-5.
- Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF (2015b) HNRNPA2B1 Is a Mediator of m(6)A- Dependent Nuclear RNA Processing Events. *Cell*. 162(6):1299-308.
- An J, Lai J, Lehman ML, Nelson CC (2013) miRDeep\*: an integrated application tool for miRNA identification from RNA sequencing data. *Nucleic Acids Res*. 41(2): 727–737.
- Amatruda JF, Shepard JL, Stern HM, Zon LI (2002) Zebrafish as a cancer model system. *Cancer Cell*. 1(3):229-31.
- Ameres SL, Horwich MD, Hung JH, Xu J, Ghildiyal M, Weng Z, Zamore PD (2010) Target RNA-directed trimming and tailing of small silencing RNAs. *Science*. 328(5985):1534-9.
- Axtell MJ (2013) ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA*. 19(6): 740–751.
- Babin PJ, Goizet C, Raldúa D (2014) Zebrafish models of human motor neuron diseases: advantages and limitations. *Prog Neurobiol*. 118:36-58.
- Bakkers J (2011) Zebrafish as model to study cardiac development and human cardiac disease. *Cardiovasc Res*. 91(2):279-88.
- Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell*. 136(2):215-33.
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 37(Web Server issue):W202-8.
- Bailey TL, Machanick P (2012) Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res*. 40(17):e128.
- Beffagna G (2019) Zebrafish as a Smart Model to Understand Regeneration After Heart Injury: How Fish Could Help Humans. *Front Cardiovasc Med*. 6:107.
- Bhat SS, Bielewicz D, Gulanicz T, Bodi Z, Yu X, Anderson SJ, Szewc L, Bajczyk M, Dolata J, Grzelak N, et al. (2020) mRNA adenosine methylase (MTA) deposits m6A on pri-

- miRNAs to modulate miRNA biogenesis in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A*. 117(35):21785-21795.
- Bokar JA (2005) The biosynthesis and functional roles of methylated nucleosides in eukaryotic mRNA. *Topics in Current Genetics*. 12:141–177.
- Broughton JP, Lovci MT, Huang JL, Yeo GW, Pasquinelli AE (2016) Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Mol Cell*. 64(2):320-333.
- Cai X, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*. 10(12):1957-66.
- Cao T, Zhen XC (2018) Dysregulation of miRNA and its potential therapeutic application in schizophrenia. *CNS Neurosci Ther*. 24(7):586-597.
- Chatterjee S, Fasler M, Büssing I, Grosshans H (2011) Target-mediated protection of endogenous microRNAs in *C. elegans*. *Dev Cell*. 20(3):388-96
- Chávez MN, Morales RA, López-Crisosto C, Roa JC, Allende ML, Lavandero S (2020) Autophagy Activation in Zebrafish Heart Regeneration. *Sci Rep*. 10(1):2191.
- Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*. 436: 740–744
- Chen T, Hao YJ, Zhang Y, Li MM, Wang M, Han W, Wu Y, Lv Y, Hao J, Wang L, Li A, Yang Y, Jin KX, Zhao X, Li Y, Ping XL, Lai WY, Wu LG, Jiang G, Wang HL, Sang L, Wang XJ, Yang YG, Zhou Q (2015) m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell*. 16(3):289-301.
- Chen L, Sun H, Wang C, Yang Y, Zhang M, Wong G (2018) miRNA arm switching identifies novel tumour biomarkers. *EBioMedicine*. 38:37-46.
- Chi SW, Hannon GJ, Darnell RB (2012) An alternative mode of microRNA target recognition. *Nat Struct Mol Biol* 19: 321–32.
- Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, Blelloch R, Schroth GP, Nusbaum C, Bartel DP (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev*. 24(10):992-1009.
- Clark PM, Loher P, Quann K, Brody J, Londin ER, Rigoutsos I. 2014. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. *Sci Rep* 4: 5947.
- Cloonan N, Wani S, Xu Q, Gu J, Lea K, Heater S, Barbacioru C, Steptoe AL, Martin HC, Nourbakhsh E, et al. (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol*. 12(12):R126

- Colaiacovo M, Bernardo L, Centomani I, Crosatti C, Giusti L, Orrù L, Tacconi G, Lamontanara A, Cattivelli L, Faccioli P (2012) A Survey of MicroRNA Length Variants Contributing to miRNome Complexity in Peach (*Prunus Persica* L.). *Front Plant Sci.*3:165.
- Coots RA, Liu XM, Mao Y, Dong L, Zhou J, Wan J, Zhang X, Qian SB (2017) m6A Facilitates eIF4F-Independent mRNA Translation. *Mol Cell.* 68(3):504-514.e7.
- Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of MIRNA genes. *Plant Cell.* 23(2):431-42.
- Cutting AD, Bannister SC, Doran TJ, Sinclair AH, Tizard MV, Smith CA (2012) The potential role of microRNAs in regulating gonadal sex differentiation in the chicken embryo. *Chromosome Res.* 20(1):201-13.
- de la Mata M, Gaidatzis D, Vitanescu M, Stadler MB, Wentzel C, Scheiffele P, Filipowicz W, Großhans H (2015) Potent degradation of neuronal miRNAs induced by highly complementary targets. *EMBO Rep.* 16(4):500-11.
- Devlin AH, Thompson P, Robson T, McKeown SR (2010) Cytochrome P450 1B1 mRNA untranslated regions interact to inhibit protein translation. *Mol Carcinog* 49: 190–199.
- Dexheimer PJ, Cochella L (2020) MicroRNAs: From Mechanism to Organism. *Front Cell Dev Biol.* 8:409.
- Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, Sorek R, Rechavi G (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature.* 485(7397):201-6.
- Dominissini D, Moshitch-Moshkovitz S, Salmon-Divon M, Amariglio N, Rechavi G (2013) Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. *Nat Protoc.* 8(1):176-89. doi: 10.1038/nprot.2012.148.
- Drummond AE (2005) TGFbeta signalling in the development of ovarian function. *Cell Tissue Res.* 322(1):107-15.
- Enright AJ1, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in *Drosophila*. *Genome Biol.* 5(1):R1.
- Fabian MR, Sonenberg N, Filipowicz W (2010) Regulation of mRNA translation and stability by microRNAs. *Ann. Rev. Biochem.* 79: 351-379.
- Fan YS, Hu YJ, Yang WX (2011) TGF- $\beta$  superfamily: how does it regulate testis development. *Mol Biol Rep.* 39(4):4727-41.
- Fromm B, Domanska D, Høye E, Ovchinnikov V, Kang W, Aparicio-Puerta E, Johansen M, Flatmark K, Mathelier A, Hovig E, Hackenberg M, Friedländer MR, Peterson KJ. (2019) MirGeneDB 2.0: the metazoan microRNA complement. *Nucleic Acids Res.* 48(D1):D132-D141.

- Fuchs Wightman F, Giono LE, Fededa JP, de la Mata M (2018) Target RNAs Strike Back on MicroRNAs. *Front Genet.* 9:435.
- Gebert LFR and MacRae IJ (2018) Regulation of microRNA function in animals. *Nat Rev Mol Cell Biol.* 20(1):21–37.
- Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML (2008) A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res.* 18(6):957-64.
- Guo L, Chen F (2014) A challenge for miRNA: multiple isomiRs in miRNAomics. *Gene.* 2014 Jul 1;544(1):1-7.
- Guo L, Yu J, Yu H, Zhao Y, Chen S, Xu C, Chen F (2015) Evolutionary and expression analysis of miR-#-5p and miR-#-3p at the miRNAs/isomiRs levels. *Biomed Res Int.* 168358.
- Guyon JR, Steffen LS, Howell MH, Pusack TJ, Lawrence C, Kunkel LM (2007) Modeling human muscle disease in zebrafish. *Biochim Biophys Acta.* 1772(2):205-15.
- Griffiths-Jones S, Hui JH, Marco A, Ronshaugen M (2011) MicroRNA evolution by arm switching. *EMBO Rep.* 12(2):172-7.
- Ha M, Kim VN (2014) Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol.* 15(8):509-24.
- Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, Filipowicz W (2005) TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* 6: 961–967
- Han J, Wang JZ, Yang X, Yu H, Zhou R, Lu HC, Yuan WB, Lu JC, Zhou ZJ, Lu Q, Wei JF, Yang H (2019) METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. *Mol Cancer.* 18(1):110.
- Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C et al (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32(Database issue):D258-61.
- Hason M, Bartunek P (2019) Zebrafish Models of Cancer-New Insights on Modeling Human Cancer in a Non-Mammalian Vertebrate. *Genes (Basel).* 10(11):935.
- Haseeb A, Makki MS, Khan NM, Ahmad I, Haqqi TM (2017) Deep sequencing and analyses of miRNAs, isomiRs and miRNA induced silencing complex (miRISC)-associated miRNome in primary human chondrocytes. *Sci Rep.* 2017 Nov 9;7(1):15178.
- Hausser J, Syed AP, Bilen B, Zavolan M (2013) Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res* 23: 604–615.
- He L e Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5: 522-531.

- Hu X, Chen W, Li J, Huang S, Xu X, Zhang X, Xiang S, Liu C (2018) ZFLNC: a comprehensive and well-annotated database for zebrafish lncRNA. Database (Oxford). 2018:bay114.
- Hutvagner G (2005) Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. *FEBS Lett.* 579: 5850–5857.
- Huyseune A, Thesleff I (2004) Continuous tooth replacement: the possible involvement of epithelial stem cells. *Bioessays.* 26(6):665-71.
- Iizasa H, Wulff BE, Alla NR, Maragkakis M, Megraw M, Hatzigeorgiou A, Iwakiri D, Takada K, Wiedmer A, Showe L, Lieberman P, Nishikura K (2010) Editing of Epstein-Barr virus-encoded BART6 microRNAs controls their dicer targeting and consequently affects viral latency. *J Biol Chem.* 285(43):33358-70
- Ji X, Jiang P, Luo J, Li M, Bai Y, Zhang J, Han B (2020) Identification and characterization of miRNAs involved in cold acclimation of zebrafish ZF4 cells. *PLoS One.* 15(1):e0226905.
- Jia W, Li Z, Lun Z (2008) Discoveries and functions of virus-encoded MicroRNAs. *Chinese Science Bulletin.* 53:169–177.
- Kaaij LJ, Hoogstrate SW, Berezikov E, Ketting RF (2013) piRNA dynamics in divergent zebrafish strains reveal long-lasting maternal influence on zygotic piRNA profiles. *RNA.* 19(3):345-56.
- Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28(1):27-30.
- Kawahara Y, Zinshteyn B, Chendrimada TP, Shiekhattar R, Nishikura K (2007a) RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO reports.* 8:763–769.
- Kawahara Y, Zinshteyn B, Sethupathy P, Iizasa H, Hatzigeorgiou AG, Nishikura K (2007b) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science.* 315(5815):1137-40.
- Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15: 2654–2659.
- Kern F, Amand J, Senatorov I, Isakova A, Backes C, Meese E, Keller A, Fehlmann T (2020) miRSwitch: detecting microRNA arm shift and switch events. *Nucleic Acids Res.* 48(W1):W268-W274.
- Kim H, Kim J, Yu S, Lee YY, Park J, Choi RJ, Yoon SJ, Kang SG, Kim VN. (2020) A Mechanism for microRNA Arm Switching Regulated by Uridylation. *Mol Cell.* 78(6):1224-1236.e5.
- Kim Y, Yeo J, Lee JH, Cho J, Seo D, Kim JS, Kim VN. Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP. *Cell Rep.* 9(3):1061-74.

- Kimmel CB (1989) Genetics and early development of zebrafish. *Trends Genet.* 5, 283–288.
- Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell.* 115(2):209-16.
- Kontur C, Jeong M, Cifuentes D, Giraldez AJ (2020) Ythdf m6A Readers Function Redundantly during Zebrafish Development. *Cell Rep.* 33(13):108598.
- Kossack ME, Draper BW (2019) Genetic regulation of sex determination and maintenance in zebrafish (*Danio rerio*). *Curr Top Dev Biol.* 134:119-149.
- Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* D152-7.
- Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ (2004) Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem.* 279(40):42230-9.
- Kuo WT, Su MW, Lee YL, Chen CH, Wu CW, Fang WL, Huang KH, Lin WC (2016) Bioinformatic Interrogation of 5p-arm and 3p-arm Specific miRNA Expression Using TCGA Datasets. *J Clin Med.* 4(9):1798-814.
- Kwak PB, Tomari Y (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol.* 19(2):145-51.
- Lakshmi SS, Agrawal S (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic acids res.* 36:D173–D177.
- Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell.* 129(7):1401-14.
- Lee HY, Doudna JA (2012) TRBP alters human precursor microRNA processing in vitro. *RNA.* 18(11):2012-9.
- Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell.* 75:843–54.
- Lee Y, Jeon K, Lee JT, Kim S, Kim VN (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* 21(17):4663-70.
- Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature.* 425(6956):415-9.
- Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23(20):4051-60.
- Li M, Hromowyk KJ, Amacher SL, Currie PD (2017) Muscular dystrophy modeling in zebrafish. *Methods Cell Biol.* 138:347-380.

- Li SC, Liao YL, Chan WC, Ho MR, Tsai KW, Hu LY, Lai CH, Hsu CN, Lin WC (2011) Interrogation of rabbit miRNAs and their isomiRs. *Genomics*. 98(6):453-9.
- Li L, Song Y, Shi X, Liu J, Xiong S, Chen W, Fu Q, Huang Z, Gu N, Zhang R (2018) The landscape of miRNA editing in animals and its impact on miRNA biogenesis and targeting. *Genome Res*. 28(1):132-143.
- Lin MH, Chen YZ, Lee MY, Weng KP, Chang HT, Yu SY, Dong BJ, Kuo FR, Hung LT, Liu LF, Chen WS, Tsai KW (2018) Comprehensive identification of microRNA arm selection preference in lung cancer: miR-324-5p and -3p serve oncogenic functions in lung cancer. *Oncol Lett*. 15(6):9818-9826.
- Lingel A, Simon B, Izaurralde E, Sattler M (2004) Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol*. 11(6):576-7.
- Liu B, Merriman DK, Choi SH, Schumacher MA, Plangger R, Kreutz C, Horner SM, Meyer KD, Al-Hashimi HM (2018) A potentially abundant junctional RNA motif stabilized by m6A and Mg<sup>2+</sup>. *Nat Commun*. 9(1):2761.
- Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*. 518(7540):560-4.
- Liu N, Pan T (2016) N6-methyladenosine-encoded epitranscriptomics. *Nat Struct Mol Biol*. 23(2):98-102.
- Liu N, Zhou KI, Parisien M, Dai Q, Diatchenko L, Pan T (2017) N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res*. 45(10):6051-6063.
- Long H, Wang X, Chen Y, Wang L, Zhao M, Lu Q (2018) Dysregulation of microRNAs in autoimmune diseases: Pathogenesis, biomarkers and potential therapeutic targets. *Cancer Lett*. 428:90-103.
- Ludwig N, Leidinger P, Becker K, Backes C, Fehlmann T, Pallasch C, Rheinheimer S, Meder B, Stähler C, Meese E. et al. (2016) Distribution of miRNA expression across human tissues. *Nucleic Acids Res*. 44:3865–3877.
- Ma H, Wu Y, Choi JG, Wu H (2013) Lower and upper stem-single-stranded RNA junctions together determine the Drosha cleavage site. *Proc. Natl Acad. Sci*. 110: 20687–20692.
- Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by Dicer. *Science*. 311(5758):195-8.
- Mao Y, Dong L, Liu XM, Guo J, Ma H, Shen B, Qian SB (2019) m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun*. 10(1):5332.
- Mathew S, Sivadas A, Sehgal P, Kaushik K, Vellarikkal SK, Scaria V, Sivasubbu S (2019) Methods to Study Long Noncoding RNA Expression and Dynamics in Zebrafish Using RNA Sequencing. *Methods Mol Biol*. 1912:77-110.

- Marco A, Macpherson JI, Ronshaugen M, Griffiths-Jones S (2012) MicroRNAs from the same precursor have different targeting properties. *Silence*. 27;3(1):8.
- Martin M (2011) Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnet.journal*. 17:1.
- Mathonnet G, Fabian MR, Svitkin YV, Parsyan A, Huck L, Murata T, Biffo S, Merrick WC, Darzynkiewicz E, Pillai RS, Filipowicz W, Duchaine TF, Sonenberg N (2007) MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science*. 317(5845):1764-7.
- Meiser N, Mench N, Hengesbach M (2020) RNA secondary structure dependence in METTL3-METTL14 mRNA methylation is modulated by the N-terminal domain of METTL3. *Biol Chem*. Online ahead of Print 2020 Oct 19
- Meng J, Lu Z, Liu H, Zhang L, Zhang S, Chen Y, Rao MK, Huang Y. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods*. 2014 Oct 1;69(3):274-81.
- Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR (2012) Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*. 22;149(7):1635-46.
- Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D (2013) Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform*. 14(2):193-202.
- Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*. 126(6):1203-17.
- Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, et al (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res*. 18:610–21.
- Neilsen CT, Goodall GJ, Bracken CP (2012) IsomiRs--the overlooked repertoire in the dynamic microRNAome. *Trends Genet*. 28(11):544-9.
- Novoa B, Figueras A (2012) Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases. *Adv Exp Med Biol*. 946:253-75.
- Okada C, Yamashita E, Lee SJ, Shibata S, Katahira J, Nakagawa A, Yoneda Y, Tsukihara T (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. *Science*. 326(5957):1275-9.
- Okamura K, Liu N, Lai EC (2009) Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. *Mol Cell*. 36: 431–444.

- Oliveira AC, Bovolenta LA, Nachtigall PG, Herkenhoff ME, Lemke N, Pinhal D (2017) Combining Results from Distinct MicroRNA Target Prediction Tools Enhances the Performance of Analyses. *Front Genet.* 8:59.
- Outtandy P, Russell C, Kleta R, Bockenbauer D (2019) Zebrafish as a model for kidney function and disease. *Pediatr Nephrol.* 34(5):751-762.
- Patel VD, Capra JA (2017) Ancient human miRNAs are more likely to have broad functions and disease associations than young miRNAs. *BMC Genomics.* 18(1):672.
- Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 2015 Mar;33(3):290-5.
- Petersen CP, Bordeleau ME, Pelletier J, Sharp PA (2006) Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell.* 21: 533-542.
- Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, Adhikari S, Shi Y, Lv Y, Chen YS, et al. (2014) Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24(2):177-89.
- Pinhal D, Bovolenta LA, Moxon S, Oliveira AC, Nachtigall PG, Acencio, ML, Patton JG, Hilsdorf AWS, Lemke N, Martins C (2018) Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. *Sci. Rep.* 8(1):8248.
- Presslauer C, Tilahun Bizuayehu T, Kopp M, Fernandes JM, Babiak I (2017) Dynamics of miRNA transcriptome during gonadal development of zebrafish. *Sci Rep.* 7:43850.
- Qin Z, Li C, Mao L, Wu L (2014) Novel insights from non-conserved microRNAs in plants. *Front Plant Sci.* 5:586.
- Qiu C, Chen G, Cui Q (2012) Towards the understanding of microRNA and environmental factor interactions and their relationships to human diseases. *Sci Rep.* 2:318.
- Reczko M, Maragkakis M, Alexiou P, Grosse I, Hatzigeorgiou AG (2012) Functional microRNA targets in protein coding sequences. *Bioinformatics.* 28(6):771-6.
- Reddy KB (2015). MicroRNA (miRNA) in cancer. *Cancer. Cell. Int.* 15:38.
- Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H2, Vilo J (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* 44(W1):W83-9.
- Ro S, Park C, Young D, Sanders KM, Yan W (2007) Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res.* 35: 5944–5953.
- Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11(3):R25.

- Ru W, Zhang X, Yue B, Qi A, Shen X, Huang Y, Lan X, Lei C, Chen H (2020) Insight into m6A methylation from occurrence to functions. *Open Biol.* 10(9):200091.
- Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res.* 17: 1850–186.
- Russell AP, Ghobrial L, Ngo S, Yerbury J, Zacharewicz E, Chung R, Lamon S (2018) Dysregulation of microRNA biogenesis machinery and microRNA/RNA ratio in skeletal muscle of amyotrophic lateral sclerosis mice. *Muscle Nerve.* 57(5):838-847.
- Sablok G, Srivastva AK, Suprasanna P, Baev V, Ralph PJ (2015) isomiRs: Increasing Evidences of isomiRs Complexity in Plant Stress Functional Biology. *Front Plant Sci.* 6:949.
- Saleem S, Kannan RR (2018) Zebrafish: an emerging real-time model system to study Alzheimer's disease and neurospecific drug discovery. *Cell Death Discov,* 4:45.
- Santoriello C, Zon LI (2012) Hooked! Modeling human disease in zebrafish. *J Clin Invest.* 122(7):2337-43.
- Sharma K, Sharma O, Tripathi M (1998) Female heterogamety in *Danio rerio* (Cypriniformes: Cyprinidae) *Proc Ind Natl Sci Acad.* 68:123–6
- Sharma, D., Sehgal, P., Mathew, S. et al (2019) A genome-wide map of circular RNAs in adult zebrafish. *Sci Rep.* 9:3432.
- Sharp PA (2009) The centrality of RNA. *Cell.* 136:577–580
- Shen Y, Guo X, Wang W (2017) Identification and characterization of circular RNAs in zebrafish. *FEBS Lett.* 591(1):213-220.
- Shenoy A, Blelloch RH (2014) Regulation of microRNA function in somatic stem cell proliferation and differentiation. *Nat Rev Mol Cell Biol.* 15(9):565-576.
- Sheu-Gruttadauria J, Pawlica P, Klum SM, Wang S, Yario TA, Schirle Oakdale NT, Steitz JA, MacRae IJ (2019) Structural Basis for Target-Directed MicroRNA Degradation. *Mol Cell.* 75(6):1243-1255.e7.
- Soh YQ1, Junker JP2, Gill ME1, Mueller JL3, van Oudenaarden A2, Page DC4 (2015) A Gene Regulatory Program for Meiotic Prophase in the Fetal Ovary. *PLoS Genet.* 17;11(9):e1005531.
- Staton KA, Edger PP, Puzey JR, Kinser T, Cheng P, Vernon DM, Forsthoefel NR, Cooley AM (2017) A Whole-Transcriptome Approach to Evaluating Reference Genes for Quantitative Gene Expression Studies: A Case Study in *Mimulus*. *G3.* 7(4): 1085–1095.
- Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T, Moulton V (2012) The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. *Bioinformatics.* 28(15):2059-61.

- Suzuki HI, Katsura A, Yasuda T, Ueno T, Mano H, Sugimoto K, Miyazono K (2015) Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nat Struct Mol Biol.* 2015 Jul;22(7):512-21.
- Tang Y, Chen K, Song B, Ma J, Wu X, Xu Q, Wei Z, Su J, Liu G, Rong R, Lu Z, de Magalhães JP, Rigden DJ, Meng J (2021) m6A-Atlas: a comprehensive knowledgebase for unraveling the N6-methyladenosine (m6A) epitranscriptome. *Nucleic Acids Res.* 49(D1):D134-D143.
- Telonis AG, Magee R, Loher P, Chervoneva I, Londin E, Rigoutsos I (2017) Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* 45(6):2973-2985.
- Treiber T, Treiber N, Meister G (2018). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nat Rev Mol Cell Biol.* 20(1):5-20.
- Tsai KW, Leung CM, Lo YH, Chen TW, Chan WC, Yu SY, Tu YT, Lam HC, Li SC, Ger LP, Liu WS, Chang HT (2016) Arm Selection Preference of MicroRNA-193a Varies in Breast Cancer. *Sci Rep.* 6:28176.
- Valenzuela-Muñoz V, Pereiro P, Álvarez-Rodríguez M et al (2019) Comparative modulation of lncRNAs in wild-type and rag1-heterozygous mutant zebrafish exposed to immune challenge with spring viraemia of carp virus (SVCV). *Sci Rep.* 9, 14174.
- Vaz C, Wee CW, Lee GPS, Ingham PW, Tanavde V, Mathavan S (2015) Deep sequencing of small RNA facilitates tissue and sex associated microRNA discovery in zebrafish. *BMC Genomics.* 16:950.
- Wang S, Zheng Z, Chen P, Wu M (2019) Tumor classification and biomarker discovery based on the 5'isomiR expression level. *BMC Cancer.* 19(1):127.
- Wei C, Salichos L, Wittgrove CM, Rokas A, Patton JG (2012) Transcriptome-wide analysis of small RNA expression in early zebrafish development. *RNA.* 18(5): 915–929.
- Wilson CA, High SK, McCluskey BM, Amores A, Yan YL, Titus TA, Anderson JL, Batzel P, Carvan MJ 3rd, Schartl M, Postlethwait JH (2014) Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics.* 198(3):1291-308.
- Wingett SW, Andrews S (2018) FastQ Screen: A tool for multi-genome mapping and quality control. *eCollection.* 7:1338.
- Wise SB, Stock DW (2010) bmp2b and bmp4 are dispensable for zebrafish tooth development. *Dev Dyn.* 239(10):2534-46.
- Wit E, Linsen S, Cuppen E, Berezikov E (2009) Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Res.* 19: 2064–2074.
- Woldemariam NT, Agafonov O, Høyheim B, Houston RD, Taggart JB, Andreassen R (2019) Expanding the miRNA Repertoire in Atlantic Salmon; Discovery of IsomiRs and miRNAs Highly Expressed in Different Tissues and Developmental Stages. *Cells.* 8(1). pii: E42.

- Wu H, Ye C, Ramirez D, Manjunath N (2009) Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS One*. 4(10):e7566.
- Wulff BE, Nishikura K (2015) Modulation of microRNA expression and function by ADARs. *Curr Top Microbiol Immunol*. 353: 91–109
- Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, Sun HY, Li A, Ping XL, Lai WY, et al. (2016) Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell*. 61(4):507-519.
- Xia H, Zhong C, Wu X, Chen J, Tao B, Xia X, Shi M, Zhu Z, Trudeau VL, Hu W (2018) *Mettl3* Mutation Disrupts Gamete Maturation and Reduces Fertility in Zebrafish. *Genetics*. 208(2):729-743
- Yang W, Chendrimada TP, Wang Q, Higuchi M, Seeburg PH, Shiekhattar R, Nishikura K (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol*. 13(1):13-21.
- Yu G, Wang LG, He QY (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 31(14):2382-3.
- Zayed Y, Qi X, Peng C (2019) Identification of Novel MicroRNAs and Characterization of MicroRNA Expression Profiles in Zebrafish Ovarian Follicular Cells. *Front Endocrinol (Lausanne)*. 10:518.
- Zhang C, Chen Y, Sun B, Wang L, Yang Y, Ma D, Lv J, Heng J, Ding Y, Xue Y, Lu X, Xiao W, Yang YG, Liu F (2017) m6A modulates haematopoietic stem and progenitor cell specification. *Nature*. 549(7671):273-276.
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*. 9(9):R137.
- Zhang B, Pan X, Cobb GP, Anderson TA (2006) Plant microRNA: a small regulatory molecule with big impact. *Dev Biol*. 289(1):3-16.
- Zhang Z, Pi J, Zou D, Wang X, Xu J, Yu S, Zhang T, Li F, Zhang X, Zhao H, Wang F, Wang D, Ma Y, Yu J (2019) microRNA arm-imbalance in part from complementary targets mediated decay promotes gastric cancer progression. *Nat Commun*. 10(1):4397.
- Zhao BS, Wang X, Beadell A, Beadell AV, Lu Z, Shi H, Kuuspalu A, Ho RK, He C (2017) m6A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* 542, 475–478.
- Zeng Y, Yi R, Cullen BR (2008) Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J*. 24, 138–148.
- Zhou H, Rigoutsos I (2014) MiR-103a-3p targets the 5' UTR of GPRC5A in pancreatic cells. *RNA*. 2014 Sep;20(9):1431-9.

Zhou X, Zuo Z, Zhou F, Zhao W, Sakaguchi Y, Suzuki T, Suzuki T, Cheng H, Zhou R (2010) Profiling sex-specific piRNAs in zebrafish. *Genetics*. 186(4):1175-85.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 31(13):3406-15.

## 7. Apêndice

### 7.1. Material suplementar

O material suplementar deste trabalho está disponível pelo link <https://drive.google.com/file/d/1HqwzkX3yTAwSGl1UeL2mbL29gm3OZONM/view>.

### 7.2. Atividades acadêmicas realizadas durante o período de doutoramento

#### 7.2.1. Artigos publicados como primeiro autor

- **Oliveira, AC**; Bovolenta, LA; Nachtigall, PG; Herkenhoff, ME; Lemke, N; Pinhal, D (2017) Combining Results from Distinct MicroRNA Target Prediction Tools Enhances the Performance of Analyses. *Frontiers in Genetics*, v. 8, p. 1-10.
- **Oliveira, AC**; Bovolenta, LA; Alves, L; Figueiredo, L; Ribeiro, AO; Campos, VF; Lemke, N; Pinhal, D (2019) Understanding the Modus Operandi of MicroRNA Regulatory Clusters. *Cells*, v. 8, p. 1103.

#### 7.2.2. Artigos publicados como coautor

- Herkenhoff, ME; **Oliveira, AC**; Nachtigall, PG; Costa, JM; Campos, VF; Hilsdorf, AWS; Pinhal, D (2018) Fishing Into the MicroRNA Transcriptome. *Frontiers in Genetics*, v. 9, p. 88.
- Pinhal, D; Bovolenta, LA; Moxon, S; **Oliveira, AC**; Nachtigall, PG; Acenio, ML; Patton, JG; Hilsdorf, AWS; Lemke, N; Martins, C (2018) Genome-wide microRNA screening in Nile tilapia reveals pervasive isomiRs' transcription, sex-biased arm switching and increasing complexity of expression throughout development. *Scientific Reports*, v. 8, p. 1-18, 2018.
- Bovolenta, LA; Pinhal, D; Acenio, ML; **Oliveira, AC**; Moxon, S; Martins, C; Lemke, N (2020) miRTil: An Extensive Repository for Nile Tilapia microRNA Next Generation Sequencing Data. *Cells*, v. 9, p. 1752.
- Herkenhoff, ME; Ribeiro, AO; Costa, JM; **Oliveira, AC**; Dias, MAD; Reis Neto, RV; Hilsdorf, AWS. ; Pinhal, D (2020) Expression profiles of growth-related genes in two Nile tilapia strains and their crossbred provide insights into introgressive breeding effects. *Animal Genetics*, v. 51, p. 611-616.

### 7.2.3. Artigos submetidos em revisão

- **Oliveira AC**, Bovolenta LA, Figueiredo L, Pinhal D. A switch in the development: microRNA arm usage screening in zebrafish suggests an important role of arm switching events on ontogenesis. *Genomics*.
- Herkenhoff ME, Bovolenta LA, Brödel O, Dos Santos LD, **Oliveira AC**, Chuffa LGA, Ribeiro AO, Lupi Jr LA, Dias MAD, Hilsdorf AWS, Frohme M, Pinhal D. Variant expression signatures of microRNAs and protein related to growth in a crossbreed between two strains of Nile tilapia (*Oreochromis niloticus*). *Frontiers in Genetics*.

### 7.2.4. Premiações

- Melhor apresentação de projeto de doutorado (Investigação do papel da metilação de RNA m6A na biogênese e processo de arm switching de microRNAs utilizando zebrafish como modelo biológico). VI Simpósio do Programa de Pós-Graduação em Ciências Biológicas – Genética. 2017.
- Melhor Painel (Variable microRNA regulatory intensity over target genes provides distinctive biological functional patterning). XVII Workshop de Genética. 2017.

### 7.2.5. Co-orientações

- Lucas Di Pietro Neves Figueiredo Pinto. Mapeamento in silico de interações microRNA-alvo em vias biológicas associadas às cardiomiopatias hipertrófica e dilatada. 2017. Trabalho de Conclusão de Curso (Ciências Biomédicas) - Universidade Estadual Paulista Júlio de Mesquita Filho, CNPq.
- Lucas Alves da Costa Silva. Mapeamento de vias biológicas reguladas por microRNAs com base na modulação diferencial do nível de expressão de genes alvos. 2019. Trabalho de Conclusão de Curso (Ciências Biomédicas) - Universidade Estadual Paulista Júlio de Mesquita Filho, FAPESP.

**7.2.6. Cursos ministrados**

- **Oliveira, AC;** Nachtigall, PG; Herkenhoff, ME (2017) Evolução humana: passado, presente e futuro. XVII Workshop de Genética. Botucatu, Brasil; 07 a 09 de abril de 2017.
- **Oliveira, AC.;** Figueiredo, L; Pereira, BJA (2018) Do unicelular ao indivíduo: como RNAs não codificantes atuam na evolução e desenvolvimento. XVIII Workshop de Genética. Botucatu, Brasil; 06 a 08 de abril de 2018.

**7.2.7. Palestra ministrada**

- **Oliveira, AC;** Pinhal, D. Unraveling microRNAs function in zebrafish: strategies and advances. XXXIV Reunião Anual da Federação de Sociedades de Biologia Experimental – FeSBE. Campos do Jordão, Brasil; 09 a 13 de setembro de 2019.

**7.2.8. Parecer de Trabalho de Conclusão de Curso**

- Felipe Garcia da Silva Sucupira. Buscando evidências da ocorrência de *Gonipterus* spp. em fezes de morcegos empregando marcadores moleculares. 2019. Trabalho de Conclusão de Curso (Ciências Biológicas - Bacharelado) - Universidade Estadual Paulista Júlio de Mesquita Filho.

**7.2.9. Revisor de periódicos**

- BMC Genomics. 2018. Detecting TF-miRNA-gene network based modules for 5hmC and 5mC brain samples: an intra- and inter-species case-study between human and rhesus.
- Gene. 2017. MicroRNA-203a regulates fast muscle differentiation by targeting *dmrt2a* in zebrafish embryos.

#### **7.2.10. Organização de eventos**

- IV Simpósio Zebrafish Como Modelo Animal de Pesquisa. Botucatu, Brasil; 26 a 28 de outubro de 2017.
- Darwin Day UNESP. Botucatu, Brasil; 30 de março de 2017.

#### **7.2.11. Apresentações de trabalhos e resumos publicados em anais de eventos**

- **Oliveira, AC**; Pinhal, D. Investigação do papel da metilação de RNA m6A na biogênese e processo de arm switching de microRNAs utilizando zebrafish como modelo biológico. VI Simpósio do Programa de Pós-Graduação em Ciências Biológicas – Genética. Botucatu, Brasil; 09 e 10 de novembro de 2017.
- **Oliveira, AC**; Bovolenta, LA; Herkenhoff, ME; Nachtigall, PG ; Lemke, N ; Pinhal, D. Variable microrna regulatory intensity over target genes provides distinctive biological functional patterning. XVII Workshop de Genética. Botucatu, Brasil; 07 a 09 de abril de 2017.
- **Oliveira, AC**; Bovolenta, LA; Figueiredo, L; Alves, L; Pinhal, D. Investigation of miRNA arm switching during zebrafish development. EMBO | EMBL Symposium: The Non-Coding Genome. Heidelberg, Alemanha; 16 a 19 de outubro de 2019.

#### **7.2.12. Resumos de trabalhos publicados em anais de eventos como coautor**

- Pinto, LDPNF; Bovolenta, LA; Nachtigall, PG ; **Oliveira, AC**; PINHAL, D. Validation of the *tbx2b*/microRNA-21 interaction during the atrio-ventricular delay in zebrafish heart. IV Simpósio Zebrafish Como Modelo Animal de Pesquisa. Botucatu, Brasil; 26 a 28 de outubro de 2017.
- Costa, JM ; Fernandes, C ; Zambuzzi, W ; **Oliveira, AC**; Pinhal, D. Effect of dietary phosphorus on the activity of alkaline phosphatase in zebrafish (*Danio rerio*). Simpósio Zebrafish Como Modelo Animal de Pesquisa. Botucatu, Brasil; 26 a 28 de outubro de 2017.
- Costa, JM; Zambuzzi, W; Fernandes, C; **Oliveira, AC**; Pinhal, D. Apparent digestibility of zebrafish (*Danio rerio*) bichlate phosphate. IV Simpósio Zebrafish Como Modelo Animal de Pesquisa. Botucatu, Brasil; 26 a 28 de outubro de 2017.

- Herkenhoff, ME; **Oliveira, AC**; Bovolenta, LA; Santos, LD; Chuffa, LGA; Ribeiro, AO; Hilsdorf, AWS; Pinhal, D. Perfis de expressão de microRNAs e proteínas em híbridos de tilápia do Nilo (*Oreochromis niloticus*). XVIII Simpósio de Citogenética e Genética de Peixes – SCGP. Cascavel, Brasil; 04 a 07 de novembro de 2018.
- Figueiredo, L; **Oliveira, AC**; Nachtigall, P; Bovolenta, LA; Pinhal, D. In silico and functional analyses of the regulatory activity of mir-21 over the *tbx2b* gene in zebrafish. Congresso Brasileiro de Genética. Águas de Lindóia, Brasil; 17 a 20 de setembro de 2019.
- Alves, LAC; **Oliveira, AC**; Bovolenta, LA; Figueiredo, L; Ribeiro, AO; Campos, VF; Lemke, N; Pinhal, D. Understanding the modus operandi of microRNA regulatory clusters. Congresso Brasileiro de Genética. Águas de Lindóia, Brasil; 17 a 20 de setembro de 2019.
- **Oliveira, AC**; Bovolenta, LA; Alves, L; Figueiredo, L; Ribeiro, AO; Campos, VF; Lemke, N; Pinhal, D. Understanding the modus operandi of microRNA regulatory clusters. Evolution 2019. Rhode Island, EUA; 21 a 25 de junho de 2019.