

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
CURSO DE GRADUAÇÃO EM BIOTECNOLOGIA**

THAIS GONZAGA GONTIJO DE SOUSA

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DAS PRINCIPAIS PROTEÍNAS ENVOLVIDAS
NA VIA DE PROCESSAMENTO DE miRNAs EM 8 ESPÉCIES DE ABELHAS**

PATOS DE MINAS - MG

JANEIRO DE 2023

THAIS GONZAGA GONTIJO DE SOUSA

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DAS PRINCIPAIS PROTEÍNAS ENVOLVIDAS
NA VIA DE PROCESSAMENTO DE miRNAs EM 8 ESPÉCIES DE ABELHAS**

Projeto de pesquisa apresentado ao Instituto de Biotecnologia da Universidade Federal de Uberlândia como requisito final para a obtenção do título de Bacharel em Biotecnologia.

PATOS DE MINAS - MG

JANEIRO DE 2023

THAIS GONZAGA GONTIJO DE SOUSA

Identificação e caracterização das principais proteínas envolvidas na via de processamento de miRNAs em 8 espécies de abelhas

Projeto de pesquisa apresentado ao Instituto de Biotecnologia da Universidade Federal de Uberlândia como requisito final para a obtenção do título de Bacharel em Biotecnologia.

Banca examinadora:

Dr. Matheus de Souza Gomes - Universidade Federal de Uberlândia (IBTEC)
Presidente

Dr. Raoni Siqueira - Universidade Federal de Uberlândia (IBTEC)
Membro

Dra. Lívia Carneiro Fidélis Silva - Universidade Federal de Uberlândia (IBTEC)
Membro

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa que se encontra no Sistema Eletrônico de Informações (SEI) da Universidade Federal de Uberlândia.

Patos de minas, 18 de janeiro de 2023.

“Pesquisa Básica gera conhecimento novo. Novos produtos e novos processos não aparecem do nada. São sustentados por novos princípios e novas ideias, que por sua vez são desenvolvidas a duras penas pela pesquisa nas áreas mais puras da ciência”

Science – The endless Frontier

RESUMO

As abelhas representam o maior grupo de polinizadores na agricultura e polinizam mais de 90% dos 107 principais cultivos agrícolas estudados no mundo. Apesar de serem elementos críticos para manutenção de ecossistemas naturais e agrícolas, as evidências apontam para uma redução brusca na população de várias espécies de abelhas em todo o mundo. Com os genomas elucidados recentemente, as espécies *Apis mellifera*, *Bombus impatiens*, *Megachile rotundata*, *Dufourea novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa*, *Lasioglossum albipes* e *Melipona quadrifasciata* apresentam uma vasta coleção de dados ainda não estudados, que possuem a capacidade de tornar possível a identificação, prevenção e mitigação dos impactos que levam a redução de populações de abelhas. Dentre os possíveis alvos de estudos, os miRNA, apresentam uma importante possibilidade no estudo da regulação da expressão gênica. Na família *Apidae*, foram citados como reguladores de comportamentos, relacionados a mudança de funções desempenhadas pelas operárias ao longo do tempo, por exemplo. O início da cadeia de eventos que leva as condições efectoras desses pequenos RNAs não codificadores de proteínas é a sua biogênese. Ela ocorre em várias etapas e exige um número considerável de proteínas para poder exercer sua função, entre elas: Drosha, Pasha, Dicer, Argonauta, RANBP21, Pasha, Tudor, FMR1, VIG2 e Loquacious. Nesse estudo, o objetivo é identificar e caracterizar, por métodos *in silico*, as principais proteínas – Drosha, Dicer e AGO - envolvidas na via de biogênese de miRNAs, através do proteoma das 8 espécies. O trabalho analisou os domínios conservados, sítios ativos e realiza filogenia das prováveis proteínas em questão. Foram identificadas e analisadas duas famílias de proteínas, RNAsesIII – 8 Drosha e 16 Dicer - e 17 Argonautas. Os resultados obtidos corroboram com caracterização das proteínas da espécie modelo *D. melanogaster* descritas na literatura, em que as proteínas caracterizadas apresentam a disposição de domínios e motifs do sítio catalítico semelhantes ou até iguais. A identificação e caracterização de proteínas desempenha um papel introdutório, mas extremamente relevante, na cadeia de entendimento sobre uma espécie. Caracterizar uma enzima é, por consequência, conhecer um pouco, também, de seu produto.

Palavras-chave: Abelhas. Argonauta. Dicer. Drosha. microRNAs.

ABSTRACT

*Bees represent the largest group of pollinators in agriculture and pollinate more than 90% of the 107 main agricultural crops studied in the world. Despite being critical elements for the maintenance of natural and agricultural ecosystems, evidence points to a critical reduction in the population of several species of bees worldwide. With recently elucidated genomes, the species *Apis mellifera*, *Bombus impatiens*, *Megachile rotundata*, *Dufourea novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa*, *Lasioglossum albipes* and *Melipona quadrifasciata* present a vast collection of data not yet studied, which have the capacity to make possible the identification, prevention, and mitigation of impacts that lead to the reduction of bees' populations. Among all the possible study targets, the miRNAs present a relevant opportunity to the study of the regulation of gene expression. In the Apidae family, they were cited as behavior regulators, related to the change of functions performed by workers over time, for example. The beginning of the chain of events that leads to the effector conditions of these small non coding protein RNAs is their biogenesis. It occurs in several stages and requires a considerable number of proteins to perform its function, including: Droscha, Pasha, Dicer, Argonaute, RANBP21, Pasha, Tudor, FMRI, VIG2 and Loquacious. In this study, the objective is to identify and characterize, by in silico methods, the main proteins - Droscha, Dicer and AGO - involved in the biogenesis pathway of miRNAs, through the proteome of the 8 species. The work analyzed the conserved domains, active sites and performs phylogeny of the probable proteins in question. Two families of proteins, RNAsesIII - 8 Droscha and 16 Dicer - and 17 Argonauts were identified and analyzed. The results obtained corroborate the characterization of the proteins of the model species *D. melanogaster* described in the literature, in which the characterized proteins present similar or even the same disposition of domains and motifs of the catalytic site. The identification and characterization of proteins plays an introductory role, but extremely relevant, in the chain of understanding about a species. Characterizing an enzyme is, therefore, also knowing a little about its product.*

Keywords: Argonaut. Bees. Dicer. Droscha. microRNAs.

LISTA DE ABREVIATURAS E SIGLAS

CDD: *Conserved Domains Database*

DGCR8: *DiGeorge Syndrome Critical Region 8*

DSRM: motif de ligação de RNA de fita dupla

dsRNA: RNA de fita dupla

mRNA: RNA mensageiro

NCBI: *National Center for Biotechnology Information*

PFAM: Protein families database

piRNA: RNA de interação com Piwi

pré-miRNA: precursor de miRNA

pri-miRNA: miRNA primário

RISC: *RNA-induced silence complex*

siRNA: pequenos RNAs interferentes

SUMÁRIO

1 INTRODUÇÃO	15
OBJETIVO	18
OBJETIVOS ESPECÍFICOS	18
2 MATERIAL E MÉTODOS	19
2.1 A BUSCA DE SEQUÊNCIAS HOMÓLOGAS DE ORGANISMOS MODELO	19
2.2 PREDIÇÃO DAS PROTEÍNAS ENVOLVIDAS NA BIOGÊNESE DE MIRNA EM ABELHAS	19
2.3 ANÁLISE DE DOMÍNIOS CONSERVADOS, DE SÍTIO ATIVO, ALINHAMENTO MÚLTIPLO E ANÁLISE FILOGENÉTICA	20
3 RESULTADOS E DISCUSSÃO	20
3.1 DROSHA	22
3.2 DICER	24
3.3 ARGONAUTA	27
4 CONCLUSÃO	32
REFERÊNCIAS	34
APÊNDICE	39

1 INTRODUÇÃO

A polinização, definição dada no ‘1º Relatório Temático sobre polinização, polinizadores e produção de alimentos no Brasil’ (2019), é a transferência de grãos de pólen entre órgãos masculinos para femininos nas flores. Este representa um processo essencial para a reprodução das plantas, principalmente angiospermas, resultando na formação de sementes e frutos.

O serviço ecológico de polinização é extremamente importante para as comunidades de plantas nativas e ecossistemas agrícolas. Para além da manutenção da biodiversidade, as plantas que são polinizadas são capazes de produzir frutos em maior quantidade e qualidade (IMPERATRIZ-FONSECA; NUNES-SILVA, 2010). Calcula-se que a polinização relacionada à produção agrícola tem um valor anual de US\$ 12 bilhões no Brasil (GIANNINI *et al.*, 2015). O processo pode ser feito através do vento, da água ou de animais, como: aves, morcegos, moscas, mariposas e abelhas. Entretanto, são os animais os protagonistas, em áreas tropicais 94% das plantas são polinizadas por eles (OLLERTON *et al.*, 2011), e são em sua grande maioria insetos. Dentro desse grupo, as abelhas representam o maior grupo de polinizadores na agricultura. Apesar de serem famosas por produzirem vários produtos apícolas como mel e própolis, sua contribuição mais relevante é o serviço de polinização, polinizando mais de 90% dos 107 principais cultivos agrícolas estudados no mundo, sendo capazes de aumentar consideravelmente a frutificação nas áreas em que atuam. (KLEIN *et al.*, 2007). Por serem aproximadamente 20mil espécies, os indivíduos pertencentes à família *Apidae*, estão presentes em vários tipos de ecossistemas, e por consequência estão adaptadas a climas e culturas distintas (DOMINGOS *et al.*, 2016).

Apesar de serem elementos críticos para manutenção de ecossistemas naturais e agrícolas, as evidências apontam para uma redução brusca na população de várias espécies de abelhas em todo o mundo (LOZIER; ZAYED, 2017). Um estudo realizado em Curitiba – PR apontou um declínio de 37% da abundância, 44% da riqueza de espécies de abelhas na área desde 1980 (CARDOSO; GONÇALVES, 2018). Segundo o ‘1º Relatório Temático sobre polinização, polinizadores e produção de alimentos no Brasil’, o motivo dessa redução é multifatorial, perpassando questões ambientais e biológicas. As mudanças no uso da terra, como o desmatamento, para a criação de pastos e a plantação intensiva e de larga escala, é a principal causa. A manutenção de um habitat rico em biodiversidade mantém as abelhas supridas de recursos alimentares – como: Pólen, néctar, óleos e tecidos florais –, abrigo e resinas durante a

todo o ano. A falta de paisagens heterogêneas, com a presença de espécies variadas de plantas entremeadas as plantações, pode levar a perda completa de grupos polinizadores, considerando que uma única espécie de planta raramente consegue suprir todas as necessidades desses polinizadores, e mesmo quando consegue, a espécie florescerá apenas de forma sazonal. O uso indiscriminado de agrotóxicos e mudanças climáticas globais também são questões relevantes a serem apontadas. Entre os fatores biológicos, é possível citar patógenos, pragas e competição entre espécies invasoras.

A boa notícia é que nem todas as espécies estão em perigo, algumas permanecem bem-sucedidas em seus habitats. Isso poderia indicar que pode haver traços genéticos que aumentam a resistência a fatores que até então afetariam a manutenção da vida da população (LOZIER; ZAYED, 2017). Entretanto, atualmente não há métodos ou protocolos para identificar, de acordo com o contexto e a espécie, os motivos que levam o declínio de uma população em específico (GROZINGER; ZAYED, 2020). Uma possível solução é aplicar a genômica para melhorar a capacidade da ciência atual para identificar, prever e mitigar os impactos que levam a redução de populações de abelhas.

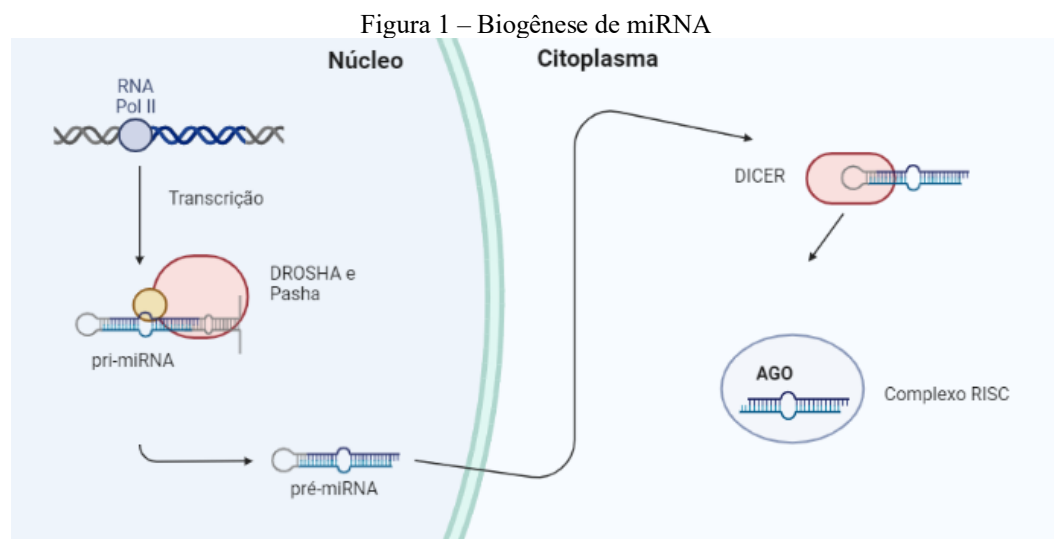
Com os genomas elucidados recentemente, as espécies *Apis mellifera*, *Bombus impatiens*, *Megachile rotundata*, *Dufourea novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa*, *Lasioglossum albipes* e *Melipona quadrifasciata* apresentam uma vasta coleção de dados ainda não estudados (SADD *et al.*, 2015). Ao se caracterizar o genoma de uma espécie, abre-se meios de entender mecanismos evolutivos importantes, que auxiliarão no estudo geral do organismo e espécies próximas filogeneticamente. Uma via importante é a de regulação da expressão gênica, que está relacionado diretamente à multicelularidade e à evolução complexa da vida vegetal e animal (PEREIRA, 2015). Essa regulação, ocorre por diferentes vias, uma delas é com o auxílio de pequenos RNAs, entre eles os que são melhor caracterizados são os microRNAs (SOUZA, 2009).

Os miRNAs, também chamados de microRNAs, pertencem a classe de RNA não codificadores de proteínas. São denominados micro por possuírem aproximadamente 22 nucleotídeos. Eles estão envolvidos no controle da expressão gênica no nível pós-transcricional (LIU *et al.*, 2019), podendo desempenhar a regulação com inibição da tradução, pelo pareamento da sequência do miRNA com o mRNA-alvo, causando impedimento estérico ou clivagem do mRNA-alvo (LUCAS; RAIKHEL, 2012).

Os miRNAs exercem um controle refinado e altamente específico dentro dos processos biológicos em que atuam, o desenvolvimento embrionário e a diferenciação tecidual, são alguns exemplos (MORAN *et al.*, 2017). Em insetos, há evidências que indicam que estão envolvidos

na regulação da metamorfose e no desenvolvimento das asas (LOZANO *et al.*, 2015; BELLES *et al.*, 2017). Especificamente em abelhas, eles já foram citados como reguladores de comportamentos, relacionados a mudança de funções desempenhadas pelas operárias ao longo do tempo (ASHBY *et al.*, 2016).

A biogênese é uma via de várias etapas, que abrange os eventos que ocorrem desde sua transcrição à incorporação do miRNA a um complexo proteico chamado RISC (KIM *et al.*, 2009). Ela exige um número considerável de proteínas para poder exercer sua função, entre elas: Drosha, Pasha (DGCR8 em mamíferos e nematoides), Dicer, Argonata, RANBP21, Pasha, Tudor, FMR1, VIG2 e Loquacious (LUCAS; RAIKHEL, 2012). A Figura 1 traz as etapas dessa via de forma simplificada, destacando as principais proteínas.



Fonte: Autor

A via se inicia com a ativação de uma RNA polimerase, de forma majoritária a Pol II, que transcreve uma cadeia longa de RNA, o pri-miRNA. Esse é caracterizado pelo grampo (hairpin), por possuírem uma 7-metilguanossina (CAP) no terminal 5' e pela cauda poliadenilada no 3' (AXTELL *et al.*, 2011). Ainda no núcleo, uma outra proteína, RNase III DROSHA juntamente com Pasha, cliva o pri-miRNA, gerando o pré-miRNA, reduzindo consideravelmente o número de nucleotídeos componentes (LEE *et al.*, 2003). Em seguida, esse precursor é exportado do núcleo para o citosol, como a utilização da proteína Exportina-5, com o cofator RAN-GTP.

Agora no citoplasma, a RNA III Dicer é a responsável pela segunda clivagem, tendo como produto o miRNA, fita dupla com um tamanho que varia de 21- 25 nucleotídeos (ZHANG *et al.*, 2004). Esse produto é incorporado ao complexo RISC, que possui a proteína Argonata

como principal componente, mas possui outras proteínas como FMR1, TSN e VIG (KETTING *et al.*, 2001) É importante ressaltar que apenas uma das fitas do duplex permanece na associação ao RISC, é exatamente esse componente que se liga ao mRNA para a regulação. Caso o pareamento entre as duas moléculas seja perfeito, há a clivagem do mRNA, senão o complexo continua associado ao mRNA, impedindo a sua tradução (LEE *et al.*, 2003; LEE *et al.*, 2004).

A *Drosophila melanogaster*, um organismo modelo muito utilizado, possui cerca de 30% de todos os genes controlados por miRNAs (RUBY *et al.*, 2007). Segundo Asgari (2015), a porcentagem de miRNA conhecidos em insetos atualmente é consideravelmente baixa, considerando que o sequenciamento da classe Insecta é também pequena.

Diante da crescente importância desses RNAs não codificantes na regulação da expressão gênica, é importante estudarmos, inicialmente, os genes envolvidos na via de miRNA, para, em consequência, melhorar a compreensão das características e função do miRNA em abelhas.

Sendo assim, o presente projeto se justifica pela necessidade de dar sentido aos dados já gerados pelo sequenciamento, com possíveis aplicações econômicas no futuro, considerando o valor ambiental e econômico das abelhas. Para tal, o trabalho faz uso de ferramentas de bioinformática com o intuito de identificar e caracterizar as proteínas de via de miRNAs, por métodos de análises de domínios conservados, sítios ativos e filogenia.

OBJETIVO

Identificar e caracterizar, por métodos *in silico*, as principais proteínas (Drosha, Dicer e Argonata) envolvidas na via de biogênese de miRNAs, através do proteoma de 8 espécies de abelhas: *Apis mellifera*, *Bombus impatiens*, *Megachile rotundata*, *Dufourea novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa*, *Lasioglossum albipes* e *Melipona quadrifasciata*.

OBJETIVOS ESPECÍFICOS

1. Buscar, no NCBI, as proteínas envolvidas na via de miRNAs em espécies modelo tais como *D. melanogaster*;

2. Buscar, nos respectivos proteomas, disponíveis no BeeBase, as proteínas envolvidas na via de miRNAs nas espécies de interesse, baseado na homologia das proteínas da espécie *D. melanogaster*;
3. Analisar os domínios conservados das proteínas preditas envolvidas na via de miRNAs nas espécies de interesse;
4. Analisar os aminoácidos dos sítios ativos nas proteínas preditas envolvidas na via de miRNAs nas espécies de interesse;
5. Analisar filogeneticamente as proteínas preditas envolvidas na via de miRNAs nas espécies de interesse;

2 MATERIAL E MÉTODOS

Neste projeto a utilização de ferramentas de bioinformática são propostas a fim de proporcionar um maior entendimento sobre as principais proteínas envolvidas na via de biogênese de miRNAs. Tais análises foram realizadas no Laboratório Bioinformática Análises Moleculares (LBAM) da Universidade Federal de Uberlândia, Campus de Patos de Minas.

2.1 A busca de sequências homólogas de organismos modelo

Para obter um conjunto de proteínas de referência, das espécies modelos, o NCBI (<http://www.ncbi.nlm.nih.gov/>) foi utilizado como banco de dados. Dentro da plataforma a ferramenta *BLASTp* foi a escolhida, e o *Database* empregado foi o *RefSeq Select proteins*, conjunto que apresenta apenas sequências obtidas através de dados experimentais.

2.2 Predição das proteínas envolvidas na biogênese de miRNA em abelhas

A busca das sequências de interesse foi realizada utilizando o proteoma predito das espécies de abelha, depositados no banco de dados BeeBase (<http://hymenopteragenome.org/beebase/>) - com as seguintes versões: *Apis mellifera* – 4.5, *Bombus impatiens* – 1.0, *Megachile rotundata* – 1.1, *Dufourea novaeangliae* – 1.1, *Eufriesea Mexicana* – 1.1, *Habropoda laboriosa* – 1.2, *Lasioglossum albipes* – 5.42 e *Melipona*

quadrifasciata – 1.1 -. Para obtenção das prováveis proteínas foi utilizado o *BLASTp*, com seqüências referência do organismo modelo *D. melanogaster*, obtidas previamente.

2.3 Análise de domínios conservados, de sítio ativo, alinhamento múltiplo e análise filogenética

As análises das famílias de proteínas e domínios foram realizadas com o auxílio do PFAM (v 31.0) (<http://pfam.xfam.org/>). Para a localização dos resíduos de aminoácidos dos sítios ativos, a ferramenta utilizada foi o CDD (<http://www.ncbi.nlm.nih.gov/cdd/>). Todas as seqüências obtidas foram, também, submetidas ao alinhamento múltiplo de seqüência utilizando o ClustalX 2.0 (v 2.1) (<http://www.clustal.org/>) com configurações padrão (disponível em <http://www.clustal.org/clustal2/>; LARKIN *et al.*, 2007). O ClustalX foi responsável por indicar possíveis erros na montagem do proteoma. Assim como, alinhar todos os sítios ativos para comparação global. Na mesma análise de sítio ativo, os logotipos de seqüência foram gerados usando o WebLogo 2.8.2 (<http://weblogo.berkeley.edu/logo.cgi>).

As árvores filogenéticas foram realizadas utilizando o método *Neighbor-joining* e as divergências de seqüências foram estimadas usando o modelo Jones – Taylor – Thornton (JTT). A confiabilidade estatística das ramificações internas foi avaliada usando 2.000 réplicas de *bootstrap*. A análise filogenética foi conduzida no programa MEGA5 (<https://www.megasoftware.net/>) (TAMURA *et al.*, 2011).

As imagens realizadas no projeto foram construídas utilizando Softwares livres: o Bio render para a imagem de Biogênese de miRNA(<https://biorender.com/>) e o InkScape (<https://inkscape.org/pt-br/>) para todas as relacionadas com os resultados da pesquisa.

3 RESULTADOS E DISCUSSÃO

Os proteomas preditos permitiram a identificação de nove prováveis famílias de proteínas envolvidas na via de miRNA- Drosha, Dicer, Argonauta, RANBP21, Pasha, Tudor, FMR1, VIG2 e Loquacious – (Tabela Suplementar I - Apêndice). Dentre elas, duas famílias de proteínas foram escolhidas para análise posterior, devido ao seu papel primordial no processamento e maturação dos miRNAs: RNase III (DICER e DROSHA) e Argonauta. A

análise foi realizada relacionando o tamanho da proteína, a conservação e distribuição de domínios, de acordo com a homologia com *D. melanogaster*. Ao final, 41 prováveis proteínas foram identificadas: 24 RNases III (8 Drosha e 16 Dicers) e 17 Argonautas. Cada espécie apresentou três proteínas RNase III, 1 Drosha e 2 Dicers, e duas Argonautas, apenas a espécie *Megachile rotundata* exibiu três proteínas Argonauta.

Os miRNAs são evolutivamente muito bem conservados, a análise de estruturas secundárias dos pré-miRNAs é um passo essencial na identificação de sequências desconhecidas e recentemente sequenciadas, em diferentes espécies (MISHRA & LOBIYAL, 2008). Segundo Sassan (2013) os miRNAs em insetos, dentre todos os outros grupos de animais, são os mais conservados. Chen *et al.* (2010), identificou 81 miRNAs através de homologia com outras espécies de insetos, indicando que além de serem miRNAs autênticos, também poderiam apresentar funções similares. Isso indica que há conservação, também, das proteínas da biogênese. Não há, até o momento, estudos sistemáticos dentro da Classe Insecta que indiquem uma variação significativa na via canônica de miRNA.

Em insetos, as pesquisas sobre miRNAs é caracterizada de forma aprofundada na espécie *D. melanogaster* (CHAWLA & SOKOL, 2011); foi nessa espécie que a molécula foi originalmente isolada por clonagem (BEHURA, 2007). Muitos dos avanços científicos referentes a biogênese de miRNAs em insetos ainda se restringem a *D. melanogaster* (PERREIRA, 2015). Há algumas especificidades da espécie modelo em relação a outros animais. A primeira é o auxílio da proteína Pasha à Drosha, na transformação de pri-miRNA para pré-miRNA. Em mamíferos e nematoides a proteína utilizada é a DGCR8. A segunda é a subdivisão de Dicer 1 e 2. A maioria das espécies animais contém apenas uma Dicer, como mamíferos e nematoides; no entanto, outras classes apresentam homólogos com diferentes funções. Na *D. melanogaster* a Dicer -1 é responsável pela formação dos duplex de miRNA e a Dicer -2 está envolvida na síntese de outros pequenos RNAs (LEE *et al.*, 2004). Além disso, as proteínas R2D2, Loquacious e FMR1, são as proteínas auxiliares a Dicer. Apesar de não participarem na clivagem do miRNA, elas participam da formação e ação do complexo proteico, aumentando a afinidade da Dicer com o pequeno RNA e na transferência do miRNA para proteína Argonauta (CHAWLA & SOKOL, 2011). Ademais, diferente de mamíferos, em que todas Argonautas (1 - 4) participam na via do miRNA, mas apenas AGO2 está presente na via do siRNA, no organismo modelo, apenas a AGO-1 está relacionada com miRNA, enquanto AGO2 auxilia na via do RNAi (CHAWLA & SOKOL, 2011).

Os resultados deste trabalho corroboram com a manutenção da conservação em relação ao organismo modelo utilizado, em que as proteínas caracterizadas apresentam a disposição de

domínios e motifs do sítio catalítico semelhantes ou até iguais. Além da identificação com e-value próximo de zero para as proteínas auxiliares encontradas na *D. melanogaster* (Pasha, FMR1 e Loquacious) e na identificação de 2 Dicercs por espécie.

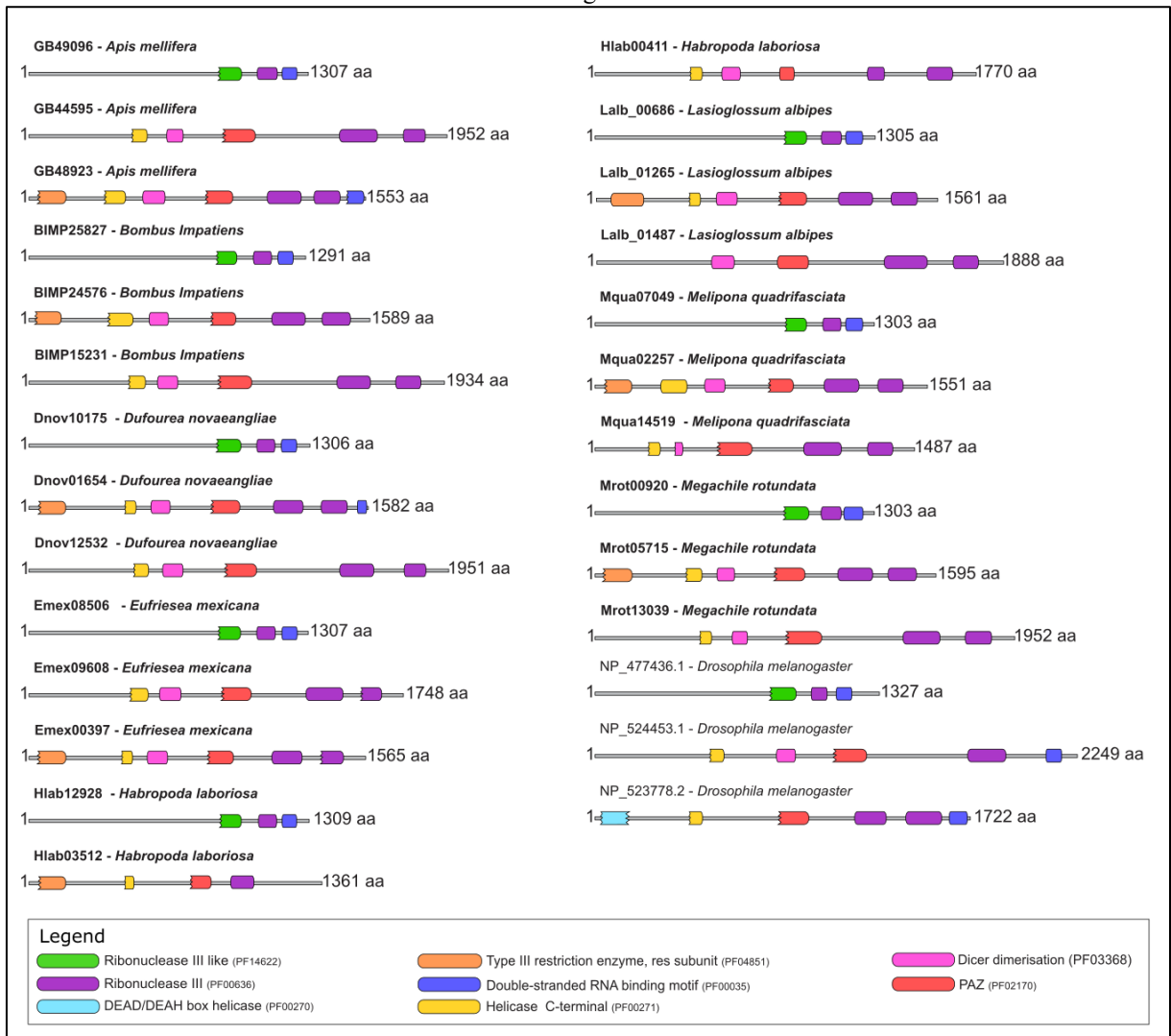
Apesar dos miRNAs serem muito importantes para a regulação da expressão gênica em abelhas, ainda há muito a ser entendido sobre os miRNAs em geral, inclusive sobre sua biogênese e suas especificidades. Assim, caracterizar a via de miRNA e os miRNAs em si, em outros insetos além da mosca da fruta, é uma tendência emergente na pesquisa de insetos (SHAFI *et al.*, 2022).

3.1 DROSHA

A Drosha é uma nuclease que possui um domínio de ligação dsRNA e dois domínios RNase III. Cada domínio RNase III está envolvido na clivagem de um dos braços pri-miRNA (LEE *et al.*, 2003). A clivagem realizada pela enzima precisa ser exata, porque define a porção terminal do microRNA (HAN *et al.*, 2006).

Na análise, foram identificadas oito possíveis proteínas Drosha - GB49096, BIMP25827, Dnov10175, Emex08506, Hlab12928, Lalb_00686, Mqua07049 e Mrot00920 -. Todas elas apresentaram os três domínios necessários: Ribonuclease III-like, Ribonuclease III e Double-stranded RNA binding motif (DSRM). As posições dos domínios são apresentadas na Tabela 1 e sua distribuição na Figura 2.

Figura 2- Domínios conservados encontrados nas proteínas da família RNase III (Drosha e Dicer) das espécies estudadas e de organismos modelo



Fonte: Autor

Os resíduos catalíticos do domínio Riboc I apresentaram o motif catalítico do sítio ativo (EDNE): ácido glutâmico (E), aspartato (D) e asparagina (N). O segundo Riboc apresentou o motivo catalítico do sítio ativo (EDDE): ácido glutâmico (E), aspartato (D) (Figura 3). Os dois domínios Ribonuclease III identificados nas possíveis proteínas mostraram um sítio ativo altamente conservado comparado aos ortólogos. As posições para cada aminoácido estão na Tabela 1.

A análise filogenética das prováveis proteínas Drosha mostrou a relação evolutiva com seus ortólogos e parálogos. Todas as prováveis proteínas Drosha analisadas se agruparam junto a outras proteínas da mesma subfamília, mais especificamente, dentro do clado dos artrópodes, como esperado (Figura 4).

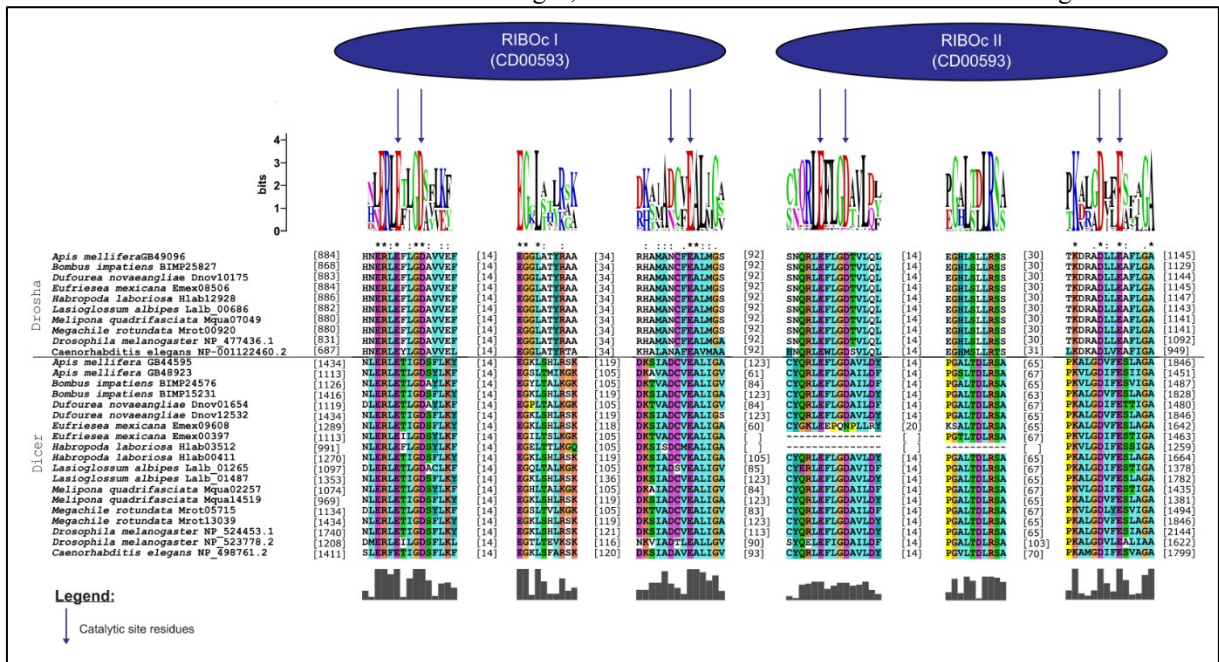
Tabela 1- Domínios conservados e posições de resíduos catalíticos de proteínas Drosha preditas

Family	Conserved domains	<i>Apis mellifera</i>				<i>Bombus impatiens</i>				<i>Dufourea novaeangliae</i>			
		GB49096				BIMP25827				Dnov10175			
		Alignment				Alignment				Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
PF14622	Ribonuclease III like	883	990	867	974	882	989						
PF00636	Ribonuclease III	1065	1151	1049	1135	1064	1150						
PF00035	DSRM	1179	1250	1163	1234	1178	1249						
Active Site: cd00593 - RIBOc	Riboc I	889	893	962	965	873	877	946	949	888	892	961	964
		E	D	N	E	E	D	N	E	E	D	N	E
	Riboc II	1068	1072	1137	1140	1052	1056	1121	1124	1067	1071	1136	1139
		E	D	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Eufriesea mexicana</i>				<i>Habropoda laboriosa</i>				<i>Lasioglossum albipes</i>			
		Emex08506				Hlab12928				Lalb_00686			
		Alignment				Alignment				Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
PF14622	Ribonuclease III like	883	990	885	992	881	988						
PF00636	Ribonuclease III	1065	1151	1067	1153	1063	1149						
PF00035	DSRM	1179	1250	1181	1252	1177	1248						
Active Site: cd00593 - RIBOc	Riboc I	889	893	962	965	E	D	N	E	887	896	960	963
		E	D	N	E	891	895	964	967	E	D	N	E
	Riboc II	1068	1072	1137	1140	1070	1074	1139	1142	1066	1070	1135	1138
		E	D	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Melipona quadrifasciata</i>				<i>Megachile rotundata</i>							
		Mqua07049				Mrot00920							
		Alignment				Alignment							
		Start	End	Start	End	Start	End	Start	End				
PF14622	Ribonuclease III like	879	986	879	986								
PF00636	Ribonuclease III	1061	1147	1061	1147								
PF00035	DSRM	1175	1246	1175	1246								
Active Site: cd00593 - RIBOc	Riboc I	885	889	958	961	885	889	958	961				
		E	D	N	E	E	D	N	E				
	Riboc II	1064	1068	1133	1136	1064	1068	1133	1136				
		E	D	D	E	E	D	D	E				

Fonte: Autor

Figura 4

Figura 3 - Análise dos domínios conservados da RNase III (Riboc I e II) das proteínas DICER e DROSHA encontrados nas abelhas e seus ortólogos, utilizando as ferramentas ClustalX e WebLogo



Fonte: Autor

3.2 DICER

A enzima Dicer participa do processamento do precursor do RNA (pré-RNA) no citoplasma, clivando-o perto da porção terminal da alça, liberando o miRNA duplex (Chawla & Sokol, 2011).

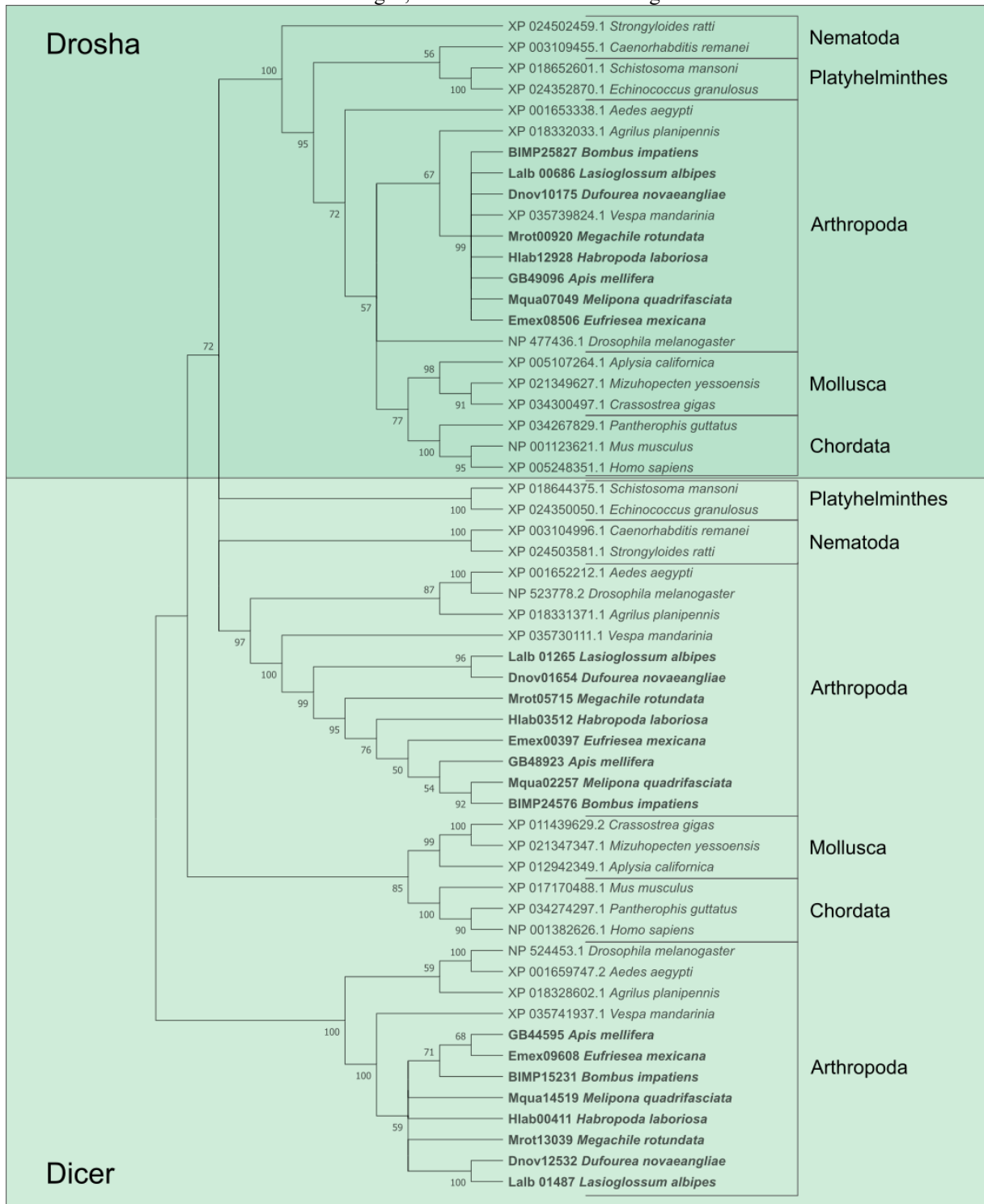
Na análise, dezesseis proteínas foram analisadas e identificadas como Dicer - GB48923, GB44595, BIMP24576, BIMP15231, Dnov01654, Dnov12532, Emex00397, Emex09608, Hlab00411, Hlab03512, Lalb_01265, Lalb_01487 Mqua02257, Mqua14519, Mrot05715 e Mrot13039. A Tabela 2 mostra a posição dos sete domínios conservados – Ribonuclease III, Ribonuclease III, PAZ, Dicer dimerization, Type III restriction enzyme subunit, Helicase conserved C-terminal e DSRM –. A Figura 2 apresenta como os domínios identificados estão distribuídos.

Tabela 1- Domínios conservados e posições de resíduos catalíticos de proteínas Drosha preditas

Family	Conserved domains	<i>Apis mellifera</i>				<i>Bombus impatiens</i>				<i>Dufourea novaeangliae</i>			
		GB49096				BIMP25827				Dnov10175			
		Alignment				Alignment				Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
PF14622	Ribonuclease III like	883	990	867	974	882	989						
PF00636	Ribonuclease III	1065	1151	1049	1135	1064	1150						
PF00035	DSRM	1179	1250	1163	1234	1178	1249						
Active Site: cd00593 - RIBOc	Riboc I	889	893	962	965	873	877	946	949	888	892	961	964
		E	D	N	E	E	D	N	E	E	D	N	E
	Riboc II	1068	1072	1137	1140	1052	1056	1121	1124	1067	1071	1136	1139
		E	D	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Eufriesea mexicana</i>				<i>Habropoda laboriosa</i>				<i>Lasioglossum albipes</i>			
		Emex08506				Hlab12928				Lalb_00686			
		Alignment				Alignment				Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
PF14622	Ribonuclease III like	883	990	885	992	881	988						
PF00636	Ribonuclease III	1065	1151	1067	1153	1063	1149						
PF00035	DSRM	1179	1250	1181	1252	1177	1248						
Active Site: cd00593 - RIBOc	Riboc I	889	893	962	965	E	D	N	E	887	896	960	963
		E	D	N	E	891	895	964	967	E	D	N	E
	Riboc II	1068	1072	1137	1140	1070	1074	1139	1142	1066	1070	1135	1138
		E	D	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Melipona quadrifasciata</i>				<i>Megachile rotundata</i>							
		Mqua07049				Mrot00920							
		Alignment				Alignment							
		Start	End	Start	End	Start	End	Start	End				
PF14622	Ribonuclease III like	879	986	879	986								
PF00636	Ribonuclease III	1061	1147	1061	1147								
PF00035	DSRM	1175	1246	1175	1246								
Active Site: cd00593 - RIBOc	Riboc I	885	889	958	961	885	889	958	961				
		E	D	N	E	E	D	N	E				
	Riboc II	1064	1068	1133	1136	1064	1068	1133	1136				
		E	D	D	E	E	D	D	E				

Fonte: Autor

Figura 4- Análise filogenética das proteínas Drosha e Dicer encontradas nas espécies de interesse e seus ortólogos, utilizando a ferramenta Mega



Uma das proteínas apresentou um domínio inesperado, Lalb_01265. Esta proteína putativa apresentou domínio de autoantígeno de câncer de cólon centrossomal, o que poderia inferir a um erro de montagem do proteoma. Provavelmente o domínio apresentado pertence a

uma proteína diferente. Após alinhá-la com sua proteína homóloga, a NP_524453.1, seu códon inicial parece ser 704Gln. Tornando o aminoácido 18Arg o início do primeiro domínio, Type III restriction enzyme subunit, o que corrobora com os resultados de outras proteínas, como GB48923, BIMP24576, Dnov01654, Emex00397, Hlab03512, Mqua14519 e Mrot05715.

Os resíduos catalíticos dos domínios Riboc I e II apresentaram o motif catalítico do sítio ativo (EDDE): ácido glutâmico (E), aspartato (D). As posições de cada aminoácido estão na Tabela 2 e a análise dos domínios conservados Riboc I e II estão na Figura 3. Apenas um deles, Hlab03512, não apresentou todos os motivos para Riboc II, o que pode significar que a proteína está truncada.

A análise filogenética foi feita simultaneamente com a proteína Droscha. Todas as prováveis proteínas Dicer analisadas se agruparam junto a outras proteínas da mesma subfamília, e, também, dentro do clado dos artrópodes (Figura 4).

3.3 ARGONAUTA

A família de proteínas Argonauta pode ser dividida em duas subfamílias, AGO e PIWI. Essas subfamílias desempenham papéis importantes nas vias de RNA não codificantes. As proteínas AGO constituem a principal proteína do complexo RISC para microRNA e para siRNA. Por sua vez, as proteínas PIWI são responsáveis pelo complexo RISC em piRNA (HUTVAGNER *et al.*, 2001; CARMELL *et al.*, 2002).

A subfamília Argonauta apresenta dois domínios altamente conservados, PIWI e PAZ. O domínio PAZ reconhece a extremidade 3' do miRNA a ser anexado no complexo RISC, enquanto o domínio PIWI possui um sítio catalítico responsável pela ligação e posterior clivagem do RNA mensageiro (mRNA), alvo do miRNA (WANG *et al.*, 2008). O sítio ativo é composto por três resíduos de aminoácidos: DDH (Aspartato, Aspartato e Histidina) ou DDD (três Aspartatos) (KAWAMATA *et al.*, 2009; KIRIAKIDO *et al.*, 2007).

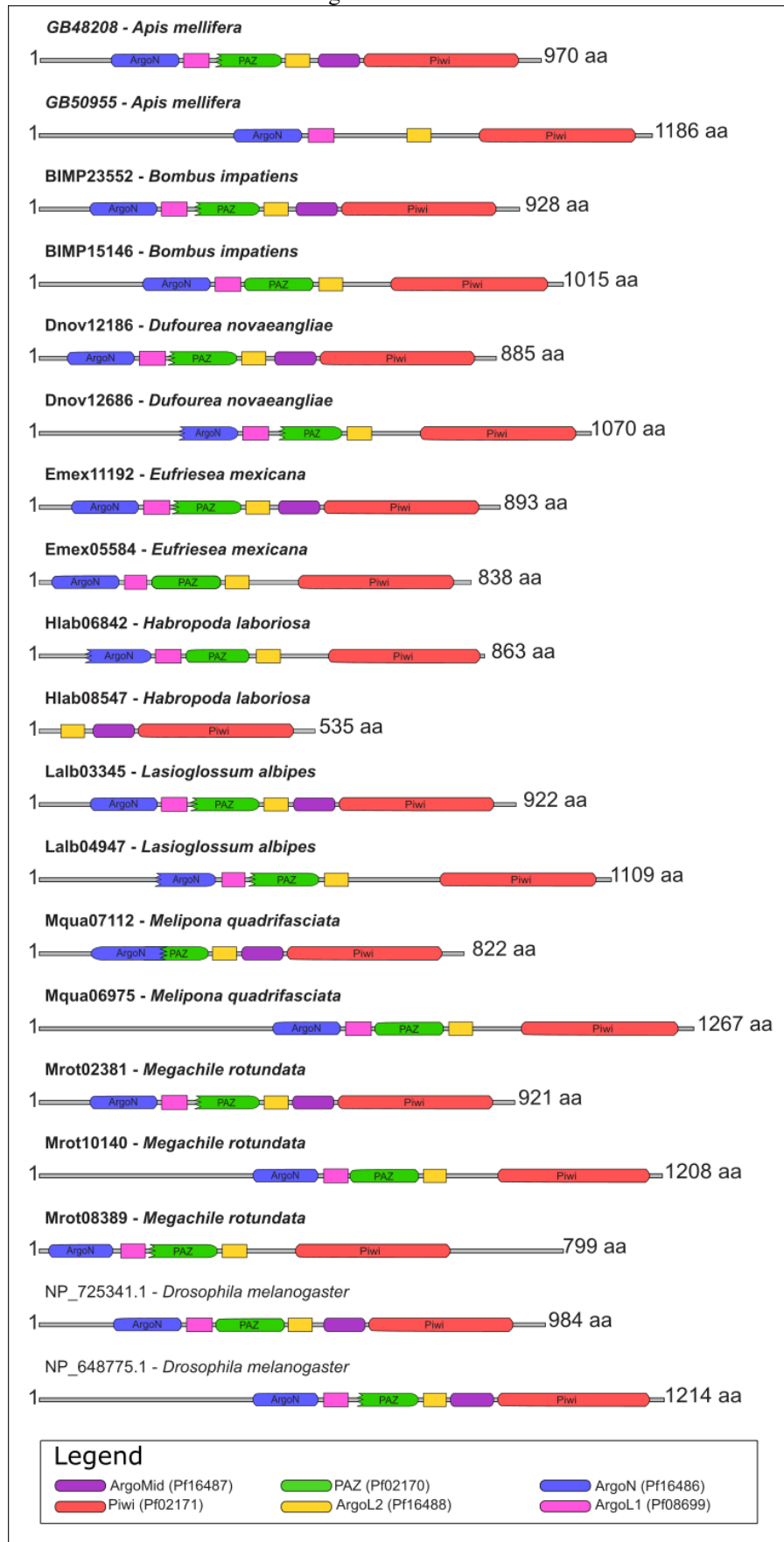
Dezessete proteínas foram identificadas como Argonauta - GB48208, GB50955, BIMP23552, BIMP15146, Dnov12168, Dnov12686, Emex11192, Emex05584, Hlab06842, Hlab08547, Lalb_03345, Lalb_04947, Mqua07112, Mqua06975, Mrot02381, Mrot10140 e Mrot08389 -. A Tabela 3 mostra os seis domínios conservados – Piwi, Mid domain of argonaute, N-terminal domain of argonaute, PAZ, Argonaute linker 1 and Argonaute linker 2 -. A Figura 5 mostra a disposição dos domínios conservados. Apesar de todas apresentarem o domínio PIWI, três delas não apresentaram o domínio PAZ: GB50955, Hlab08547, Mrot10140. Podendo significar que as proteínas estão truncadas.

Tabela 2- Domínios conservados e posições de resíduos catalíticos de proteínas Dicer preditas

Family	Conserved domains	<i>Apis mellifera</i>								<i>Bombus impatiens</i>							
		GB 48923				GB 44595				BIMP24576				BIMP15231			
		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End				
PF00636	Ribonuclease III	1115	1272	1436	1611	1128	1289	1418	1593								
PF00636	Ribonuclease III	1334	1456	1731	1852	1370	1492	1713	1834								
PF02170	PAZ	820	953	906	1063	844	967	880	1039								
PF03368	Dicer dimerisation	534	627	623	714	564	653	599	690								
PF04851	Type III restriction enzyme, res subunit	21	177	-	-	17	175	-	-								
PF00271	Helicase conserved C-terminal	331	454	488	548	368	485	463	524								
PF00035	DSRM	1487	1549	-	-	-	-	-	-								
Active Site: cd00593 - RIBOc	Riboc I	1118	1122	1262	1265	1439	1443	1597	1600	1131	1135	1275	1278	1421	1425	1579	1582
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E
	Riboc II	1337	1341	1443	1446	1734	1738	1838	1841	1373	1377	1479	1482	1716	1720	1820	1823
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Dufourea novaeangliae</i>								<i>Eufriesea mexicana</i>							
		Dnov01654				Dnov12532				Emex00397				Emex09608			
		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End				
PF00636	Ribonuclease III	1121	1281	1436	1611	1115	1276	1291	1465								
PF00636	Ribonuclease III	1363	1486	1731	1852	1361	1468	1547	1648								
PF02170	PAZ	849	985	908	1064	818	950	898	1042								
PF03368	Dicer dimerisation	566	660	625	716	536	623	615	706								
PF04851	Type III restriction enzyme, res subunit	21	177	-	-	21	176	-	-								
PF00271	Helicase conserved C-terminal	429	486	491	550	428	481	469	527								
PF00035	DSRM	1531	1578	-	-	-	-	-	-								
Active Site: cd00593 - RIBOc	Riboc I	1124	1128	1268	1271	1439	1443	1597	1600	1118	1122	1262	1265	1294	1298	1451	1454
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E
	Riboc II	1366	1370	1472	1475	1734	1738	1838	1841	1356	1359	1455	1458	1525	1529	1634	1637
		E	D	D	E	E	D	D	E	E	D	D	E	E	N	D	E
Family	Conserved domains	<i>Habropoda laboriosa</i>								<i>Lasioglossum albipes</i>							
		Hlab00411				Hlab03512				Lalb_01265				Lalb_01487			
		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End				
PF00636	Ribonuclease III	1272	1447	993	1152	1801	1962	1355	1547								
PF00636	Ribonuclease III	1549	1670	-	-	2044	2167	1667	1788								
PF02170	PAZ	864	1029	727	853	1519	1654	829	995								
PF03368	Dicer dimerisation	589	680	-	-	1233	1326	554	646								
PF04851	Type III restriction enzyme, res subunit	-	-	22	176	720	886	-	-								
PF00271	Helicase conserved C-terminal	441	501	431	483	1101	1154	-	-								
PF00035	DSRM	-	-	-	-	-	-	-	-								
Active Site: cd00593 - RIBOc	Riboc I	1275	1279	1433	1436	996	1000	1140	1143	1803	1807	1948	1951	1670	1674	1774	1777
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E
	Riboc II	1552	1556	1656	1659	-	-	1251	1254	2047	2051	2153	2156	1358	1362	1533	1536
		E	D	D	E	-	-	D	E	E	D	D	E	E	D	D	E
Family	Conserved domains	<i>Melipona quadrifasciata</i>								<i>Megachile rotundata</i>							
		Mqua02257				Mqua14519				Mrot05715				Mrot13039			
		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment		Alignment			
		Start	End	Start	End	Start	End	Start	End	Start	End	Start	End				
PF00636	Ribonuclease III	1076	1235	971	1146	1136	1293	1436	1611								
PF00636	Ribonuclease III	1318	1440	1266	1387	1377	1500	1731	1852								
PF02170	PAZ	807	928	576	733	835	978	895	1059								
PF03368	Dicer dimerisation	512	607	371	409	566	650	619	710								
PF04851	Type III restriction enzyme, res subunit	22	17	-	-	23	175	-	-								
PF00271	Helicase conserved C-terminal	309	430	248	307	429	487	485	544								
PF00035	DSRM	-	-	-	-	-	-	-	-								
Active Site: cd00593 - RIBOc	Riboc I	1079	1083	1223	1226	974	978	1132	1135	1139	1143	1283	1286	1439	1443	1597	1600
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E
	Riboc II	1321	1325	1427	1430	1269	1273	1373	1376	1380	1384	1486	1489	1734	1738	1838	1841
		E	D	D	E	E	D	D	E	E	D	D	E	E	D	D	E

Fonte: Autor

Figura 5 - Domínios conservados encontrados em proteínas Argonauta das espécies estudadas e em organismos modelo



Fonte: Autor

Tabela 3- Domínios conservados e posições de resíduos catalíticos de proteínas Argonauta preditas

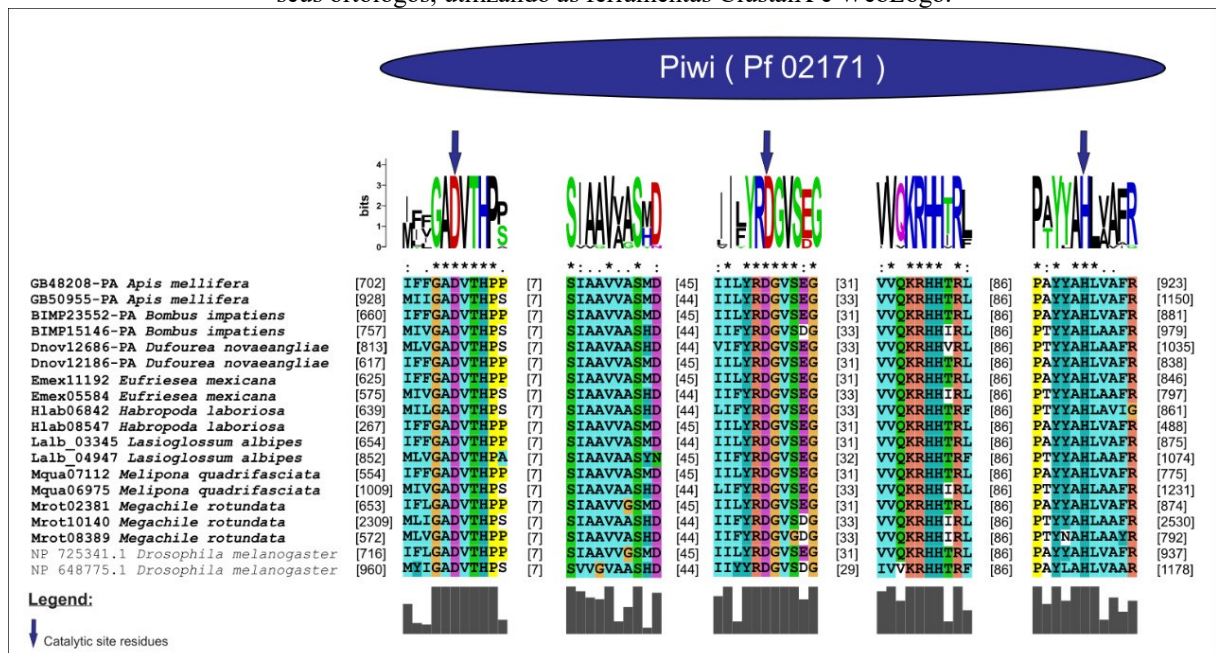
		<i>Apis mellifera</i>				<i>Bombus impatiens</i>										
Family	Conserved domains	GB48208		GB50955		BIMP23552		BIMP15146								
		Alignment		Alignment		Alignment		Alignment								
		Start	End	Start	End	Start	End	Start	End							
PF02171	Piwi	628	928	852	1155	586	886	681	984							
PF16487	Mid domain of argonaute	540	620	-	-	498	578	-	-							
PF16486	N-terminal domain of argonaute	140	269	378	509	98	227	200	331							
PF02170	PAZ	343	468	-	-	301	426	397	531							
PF08699	Argonaute linker 1	279	329	520	569	237	287	342	391							
PF16488	Argonaute linker 2	477	523	712	758	435	481	541	587							
Active Site: cd04657	Piwi_ago-like	707	780	918	933	1005	1145	665	738	876	762	834	974			
		D	D	H	D	D	H	D	D	H	D	D	H			
		<i>Dufourea novaeangliae</i>				<i>Eufriesea mexicana</i>										
Family	Conserved domains	Dnov12168		Dnov12686		Emex11192		Emex05584								
		Alignment		Alignment		Alignment		Alignment								
		Start	End	Start	End	Start	End	Start	End							
PF02171	Piwi	543	843	739	1040	551	851	501	802							
PF16487	Mid domain of argonaute	455	534	-	-	463	543	-	-							
PF16486	N-terminal domain of argonaute	55	184	271	384	63	192	24	155							
PF02170	PAZ	258	383	472	586	266	391	217	351							
PF08699	Argonaute linker 1	194	244	395	443	202	252	166	207							
PF16488	Argonaute linker 2	392	438	596	642	400	446	361	407							
Active Site: cd04657	Piwi_ago-like	622	695	833	818	890	1030	630	703	841	580	652	792			
		D	D	H	D	D	H	D	D	H	D	D	H			
		<i>Habropoda laboriosa</i>				<i>Lasioglossum albipes</i>										
Family	Conserved domains	Hlab06842		Hlab08547		Lalb_03345		Lalb_04947								
		Alignment		Alignment		Alignment		Alignment								
		Start	End	Start	End	Start	End	Start	End							
PF02171	Piwi	560	858	193	493	580	880	778	1079							
PF16487	Mid domain of argonaute	-	-	105	184	492	571	-	-							
PF16486	N-terminal domain of argonaute	121	216	-	-	99	228	246	341							
PF02170	PAZ	284	405	-	-	302	427	424	542							
PF08699	Argonaute linker 1	226	275	-	-	238	288	352	399							
PF16488	Argonaute linker 2	420	466	42	88	436	482	552	598							
Active Site: cd04657	Piwi_ago-like	644	716	856	272	345	483	659	732	870	857	930	1069			
		D	D	H	D	D	H	D	D	H	D	D	H			
		<i>Melipona quadrifasciata</i>				<i>Megachile rotundata</i>										
Family	Conserved domains	Mqua07112		Mqua06975		Mrot02381		Mrot10140		Mrot08389						
		Alignment		Alignment		Alignment		Alignment		Alignment						
		Start	End	Start	End	Start	End	Start	End	Start	End					
PF02171	Piwi	480	780	933	1235	579	879	2233	2535	496	796					
PF16487	Mid domain of argonaute	392	472	-	-	490	571	-	-	-	-					
PF16486	N-terminal domain of argonaute	100	226	453	584	98	227	1754	1883	17	143					
PF02170	PAZ	202	327	663	783	301	426	-	-	230	345					
PF08699	Argonaute linker 1	-	-	594	642	237	287	1894	1944	158	205					
PF16488	Argonaute linker 2	336	382	793	839	435	481	2093	2139	354	401					
Active Site: cd04657	Piwi_ago-like	1014	1086	1226	1014	1086	1226	658	731	869	2314	2386	2525	577	649	787
		D	D	H	D	D	H	D	D	H	D	D	H	D	D	H

Fonte: Autor

Duas delas mostraram resultados de tamanho inesperados, Mrot10140 e Emex05584. Ambas as prováveis proteínas apresentaram um tamanho muito superior em relação às demais, o que poderia inferir a um erro de montagem do genoma. Após alinhá-los com suas proteínas homólogas, o início de Mrot10140 e o final de Emex05584 foram alterados. Agora, o primeiro códon de Mrot10140 é 1359Phe, e o último códon de Emex05584 é 838Pro.

Os resíduos catalíticos do domínio PIWI apresentaram o motivo catalítico do sítio ativo (DDH): aspartato (D) e Histidina (H). As posições para cada aminoácido estão na Tabela 3 e a análise da conservação do motivo catalítico do domínio Piwi está na Figura 6.

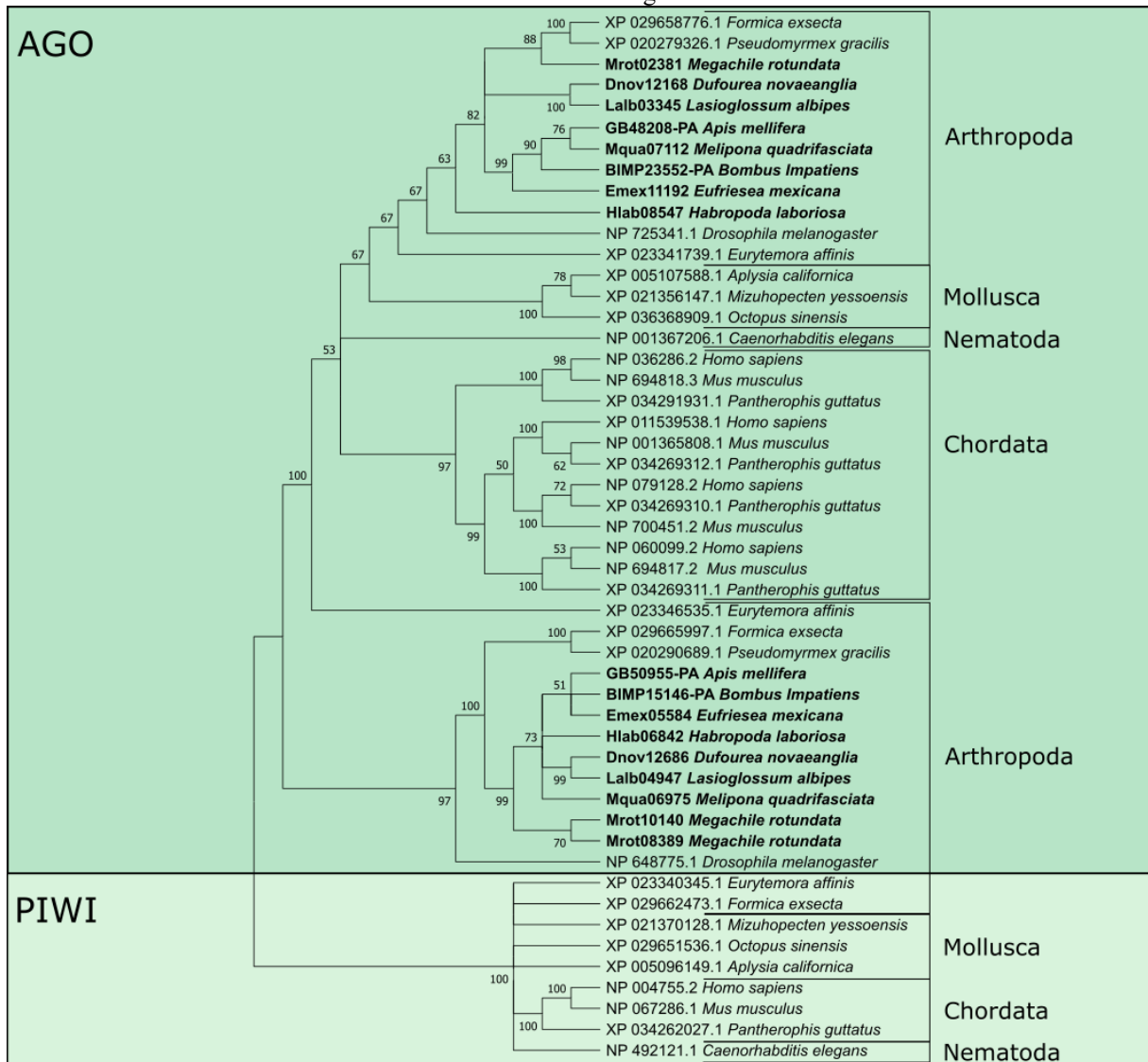
Figura 6 - Análise de domínios conservados PIWI em proteínas Argonauta encontradas nas espécies de abelhas e seus ortólogos, utilizando as ferramentas ClustalX e WebLogo.



Fonte: Autor

A análise filogenética foi realizada simultaneamente com a subfamília PIWI. Todas as dezessete proteínas analisadas se agruparam junto a subfamília Argonauta, no clado de artrópodes (Figura 7).

Figura 7- Análise filogenética das proteínas Argonauta encontradas nas espécies de abelhas e seus ortólogos, utilizando a Mega.



Fonte: Autor

4 CONCLUSÃO

Corroborando com os dados na Literatura, as 8 espécies de abelhas apresentaram proteínas com características similares as do ortólogo modelo, *D. melanogaster*. Entende-se, portanto, que há fortes indícios de manutenção na conversação estrutural e funcional das proteínas caracterizadas.

Apesar da identificação e caracterização de proteínas desempenhar um papel introdutório, esse trabalho é extremamente relevante na cadeia de entendimento sobre uma

espécie. Caracterizar uma enzima é, por consequência, conhecer um pouco, também, de seu produto.

Os resultados deste trabalho, portanto, é um prelúdio a novas possibilidades. Um novo caminho para o estudo de miRNA na biologia das abelhas ajudando a compreender, a partir da análise comparativa de elementos moleculares, semelhanças e diferenças que justifiquem modificações morfológicas, fisiológicas e comportamentais está sendo aberto. Assim como, para a origem e manutenção da sociabilidade, interações gênicas que levam ao desenvolvimento de castas, base genética da infertilidade em operárias, alvos moleculares relacionados à regulação do comportamento polinizador, ação equilibrada de hormônios, entre outros.

Portanto, estudos sobre a via de miRNA em abelhas podem ajudar a compreender melhor como essas moléculas regulam a expressão gênica e, futuramente, como elas podem ser usadas em aplicações médicas e biotecnológicas. Além disso, esse tipo de conhecimento pode ajudar na compreensão de como essas moléculas atuam na saúde desses indivíduos, o que é importante para a conservação desses importantes polinizadores.

REFERÊNCIAS

- ASGARI, S. Regulatory role of cellular and viral microRNAs in insect–virus interactions. **Elsevier**, v. 8, p. 104, 2015.
- ASHBY R *et al.*. MicroRNAs in honeybee caste determination. **Scientific Reports**, v. 6, n.18, p. 794, 2016.
- AXTELL, M.; WESTHOLM, J.; ERIC C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. **Genome Biology**, v. 12, n. 4, p. 221, 2011.
- BELLES X. MicroRNAs and the evolution of insect metamorphosis. **Annu Rev Entomol**, v. 62, p. 111, 2017.
- BEHURA, S. K. Insect microRNAs: Structure, function, and evolution. **Insect Biochemistry and Molecular Biology**, v. 37, p. 3, 2007.
- CARDOSO, M. C.; GONÇALVES, R. B. Reduction by half: the impact on bees of 34 years of urbanization. **Urban Ecosystems**, v. 21, n. 5, p. 943, 2018.
- CARMELL, M. A. et al.. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. **Genes Dev**, v. 16, n.2, p. 733, 2002.
- CHEN, X. et al.. Next-generation small RNA sequencing for microRNAs profiling in the honey bee *Apis mellifera*. **Insect Molecular Biology**, v. 10, n. 6, p. 799, 2010.
- DOMINGOS, A.T.S.; NÓBREGA, M.M; SILVA, R.A. Biologia das abelhas *Apis mellifera*: Uma revisão bibliográfica. **ACTA Apicola Brasilica**, v. 04, n.2, p.08-12, 2016.
- CHAWLA, G.; SOKOL, N. S. MicroRNAs in *Drosophila* development. **Int Rev Cell Mol Biol**, n. 65, v. 1, p. 286, 2011;
- GIANNINI, C. et al.. The Dependence of Crops for Pollinators and the Economic Value of Pollination in Brazil. **Journal of Economic Entomology**, v. 108, n. 3, p. 849-857, 2015.
- GKIRTZOU, K. et al.. MatureBayes: A Probabilistic Algorithm for Identifying the Mature miRNA within Novel Precursors. **PLoS ONE**, v. 5, n. 8, p. 11843, 2010.

GOMES, M. et al.. Preliminary analysis of miRNA pathway in *Schistosoma mansoni*. **Parasitology International**, v. 58, n. 1, p. 61, 2009.

GOMES, M. S. et al.. Preliminary analysis of miRNA pathway in *Schistosoma mansoni*. **Parasitology International** , v. 58, p. 61, 2009.

GROZINGER, C. M.; ZAYED, A. Improving bee health through genomics. **Nature Reviews**, v. 21, p. 277, 2020.

IMPERATRIZ-FONSECA, L; NUNES-SILVA, P. V. As abelhas, os serviços ecossistêmicos e o Código Florestal Brasileiro. **In Biota Neotrop**, v. 10, n. 4, 2010.

HAN, J. et al.. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. **Cell**, v. 125, n. 5, p. 887, 2006.

HUTVAGNER G, et al.. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. **Science**, v. 293, p. 834, 2001.

JIANG, P. et al.. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. **Nucleic Acids Research**, v. 35, n. Web Server, p. 339, 2007.

KAWAMATA T. et al.. Structural determinants of miRNAs for RISC loading and slicer **independent unwinding**. , v. 16, n. 9, p. 953 , 2009.

KETTING R, et al.. Dicer functions in RNA interferen and in synthesis of small RNA involved in developmental timing in *C. elegans*. **Ele Genes Dev**, v. 15, n.20, p. 2654, 2001.

KIM, V.; NARRY, H.; JINJUSIOMI, M. Biogenesis of small RNAs in animals. Nature Reviews **Molecular Cell Biology**, v. 10, n. 2, p. 126, 2009.

KIRIAKIDOU, M. et al.. An mRNA m7G cap binding-like motif within human Ago2 represses **translation**. **Cell**, v. 129, p. 1141, 2007.

KLEIN, A. et al.. Importance of pollinators in changing landscapes for world crops. **Biological Sciences**, v. 274, n. 1608, p. 303, 2006.

- LARKIN M.A. et al.. Clustal W and Clustal X version 2.0. *Bioinformatics*, v. 23, p. 2947, 2007.
- LEE, Y. et al.. MicroRNA genes are transcribed by RNA polymerase II. **The EMBO Journal**, v. 23, n. 20, p. 4051, 2004.
- LEE, Y. et al.. The nuclear RNase III Droscha initiates microRNA processing. **Nature**, v. 425, n. 6956, p. 415, 2003.
- LIU H. et al.. Analysis of miRNAs in the Heads of Different Castes of the Bumblebee *Bombus lantschouensis* (Hymenoptera: Apidae). **Insects**, v.10, n.10, p. 349, 2019.
- LOZANO J, et al.. MiR-2 family regulates insect metamorphosis by controlling the juvenile hormone signaling pathway. **Proc Natl Acad Sci USA**, v. 112, n;12, p. 3740, 2015.
- LOZIER, J. D.; ZAYED, A. Bee conservation in the age of genomics. **Conservation Genetics**, v. 18, n. 3, p. 713, 2017.
- LUCAS, K.; RAIKHEL A. S. Insect MicroRNAs: biogenesis, expression profiling and biological functions. **Insect Biochem Mol Biol**, v. 43, n.1, p.24, 2012.
- MISHRA ,A.K.; LOBIYAL, D.K. Exploring dominating features from *Apis Mellifera* pre-miRNA. **International Conference on Advanced Computer Theory and Engineering**, p 363, 2008.
- MORAN Y. et al.. The evolutionary origin of plant and animal microRNAs. **Nat Ecol Evol**. v.1, n. 3, p.27, 2017.
- NÓBREGA, M.; SILVA, R. Biologia das abelhas *Apis Mellifera*: Uma revisão bibliográfica. Pombal: **ACTA Apicola Brasilica**, v. 04, n.2, p.08 – 12, 2016.
- OLLERTON, J.; WINFREE, R.; TARRANT, S. How many flowering plants are pollinated by animals?. **Oikos**, v. 120, n. 3, p. 321-326, 2011.
- PALMIER, K.; SHEFFIELD, C. First records of the Common Eastern Bumble Bee, *Bombus impatiens* Cresson (Hymenoptera: Apidae, Apinae, Bombini) from the Prairies Ecozone in Canada. **Biodiversity Data Journal**, v. 7, 2019.
- PEREIRA, T., C. et al.. Introdução ao Mundo dos MicroRNAs. 1. ed. Riberão Preto: Sociedade Brasileira de Genética, 2015.

PETERSEN, J.; REINERS, S.; BRIAN, A. Pollination Services Provided by Bees in Pumpkin Fields Supplemented with Either *Apis mellifera* or *Bombus impatiens* or Not Supplemented.

PLoS ONE, v. 8, n. 7, p. e69819, 2013.

RUBY, J. et al.. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. **Genome Research**, v. 17, n. 12, p. 1850-1864, 2007.

SADD, B et al.. The genomes of two key bumblebee species with primitive eusocial organization. **Genome Biology**, v. 16, n. 1, 2015.

SAITOU, N. et al.. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, n. 4, p. 406–425, 1987.

SASSAN, A. MicroRNA functions in insects. **Insect Biochemistry and Molecular Biology**, v. 43, p. 388, 2013.

SHAFI, I. *et al.*. The potential role of long non-coding RNAs and micro RNAs in insects: From junk to luxury. Eurasian **Journal of Molecular and Biochemical Sciences**, v. 1, p. 28, 2022.

TAMURA, K. et al.. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. **Molecular Biology and Evolution**, v. 24, n. 8, p. 1596-1599, 2007.

THOMPSON, D.; GIBSON, T.; HIGGINS, G. Multiple Sequence Alignment Using ClustalW and ClustalX. **Current Protocols in Bioinformatics**, v. 04, n. 1, 2002.

TRAPNELL, C. et al.. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, v. 7, n. 3, p. 562-578, 2012.

VELTHUIS, H.; VAN-DOORN, A. A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. **Apidologie**, v. 37, n. 4, p. 421-451, 2006.

WANG, Y. et al.. Structure of the guide-strand-containing argonaute silencing complex. **Nature**, v. 456, p. 209, 2008.

WHIDDEN, L. The fidelity of commercially reared colonies of *Bombus impatiens cresson* (hymenoptera: apidae) to lowbush blueberry in southern new brunswick. *The Canadian Entomologist*, v. 128, n. 5, p. 957-958, 1996.

WILLIAMS, H. Phylogenetic relationships among bumble bees (*Bombus* Latr.): a reappraisal of morphological evidence. *Systematic Entomology*, v. 19, n. 4, p. 327-344, 1994.

WOLOWSKI, M et al.. 1º Relatório temático sobre Polinização, Polinizadores e Produção de Alimentos no Brasil. **BPBES/REBIPP**, v. 1, p. 148-256, 2019.

ZHANG, H. et al.. Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell*, v. 118, n. 1, p. 57-68, 2004.

ZUKER, M.; STIEGLER, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, v. 9, n. 1, p. 133-148, 1981.

APÊNDICE

Tabela Suplementar 1 – Resultado do alinhamento através da ferramenta BLASTP entre proteínas de interesse e proteínas do organismo modelo *D. melanogaster*

ID protein	Putative name <i>Bombus Impatiens</i>	Length (aa)	ID protein <i>D. melanogaster</i>	E-value	Length (aa)
GB49096	AMEL_Drosha	1307	-	-	-
GB44595	AMEL_Dicer	1952	NP_524453.1	0.0	2249
			NP_523778.2	1,00E-67	1722
GB48923	AMEL_Dicer	1553	NP_524453.1	1,00E-91	2249
			NP_523778.2	0.0	1722
GB48208	AMEL_Argonaute	970	NP_725341.1	0.0	984
			NP_648775.1	3,00E-102	1214
			NP_001036627.2	7,00E-39	867
GB50955	AMEL_Argonaute	1186	NP_725341.1	1,00E-146	984
			NP_648775.1	5,00E-148	1214
			NP_001036627.2	2,00E-36	867
GB47231	AMEL_Ranbp21	1217	NP_001285453.1	0.0	1241
GB40545	AMEL_Pasha	638	NP_651879.1	0.0	642
GB53973	AMEL_Tudor	2693	NP_476773.1	1,00E-54	2515
GB50259	AMEL_Fmr1	621	NP_611645.1	1,00E-165	684
GB46268	AMEL_Vig2	414	NP_733084.1	1,00E-33	412
GB47214	AMEL_Loquacious	383	NP_723813.1	1,00E-105	419
GB54480	AMEL_Loquacious	337	NP_723813.1	9,00E-28	419
BIMP25827	BIM_Drosha	1291	NP_477436.1	0.0	1327
BIMP15231	BIM_Dicer	1934	NP_524453.1	0.0	2249
			NP_523778.2	9,00E-75	1722
BIMP24576	BIM_Dicer	1589	NP_524453.1	8,00E-92	2249
			NP_523778.2	9,00E-75	1722
BIMP23552	BIM_Argonaute	928	NP_725341.1	0.0	984
			NP_648775.1	5,00E-103	1214
			NP_001036627.2	1,00E-38	867

*Apis mellifera**Bombus impatiens*

			NP_725341.1	4,00E-152	984
BIMP15146	BIM_Argonaute	1015	NP_648775.1	6,00E-161	1214
			NP_001036627.2	2,00E-38	867
BIMP17218	BIM_Ranbp21	1217	NP_001285453.1	0.0	1241
BIMP21618	BIM_Pasha	639	NP_651879.1	0.0	642
BIMP25242	BIM_Tudor	2462	NP_476773.1	2E-48	2515
BIMP14884	BIM_Fmr1	381	NP_611645.1	2,00E-167	684
BIMP23765	BIM_Vig2	623	NP_733084.1	1,00E-32	412
BIMP17206	BIM_Loquacious	381	NP_723813.1	3,00E-106	419
BIMP21027	BIM_Loquacious	332	NP_723813.1	1,00E-29	419
Dnov10175	DNOV_Drosha	1306	-	-	-
			NP_524453.1	5E-82	2249
Dnov01654	DNOV_Dicer	1582	NP_523778.2	1,00E-119	1722
			NP_524453.1	0.0	2249
Dnov12532	DNOV_Dicer	1951	NP_523778.2	1,00E-67	1722
			NP_725341.1	0.0	984
Dnov12168	DNOV_Argonaute	885	NP_648775.1	5,00E-103	1214
			NP_001036627.2	2,00E-38	867
			NP_725341.1	1,00E-132	984
Dnov12686	DNOV_Argonaute	1070	NP_648775.1	1,00E-167	1214
			NP_001036627.2	2,00E-32	867
Dnov05799	DNOV_Ranbp21	1176	NP_001285453.1	0.0	1241
Dnov08200	DNOV_Pasha	606	NP_651879.1	0.0	642
Dnov11113	DNOV_Tudor	2582	NP_476773.1	6E-53	2515
Dnov04084	DNOV_Fmr1	643	NP_611645.1	1,00E-158	684
Dnov06058	DNOV_Vig2	420	NP_733084.1	3,00E-31	412
Dnov04997	DNOV_Loquacious	376	NP_723813.1	2,00E-99	419
Dnov00381	DNOV_Loquacious	678	NP_723813.1	2,00E-27	419
Emex08506	EMEX_Drosha	1307	NP_477436.1	0.0	1327
			NP_524453.1	0.0	2249
Emex00397	EMEX_Dicer	1565	NP_523778.2	3,00E-64	1722

*Dufourea novaeangliae**Eufriesea
mexicana*

Habropoda laboriosa

Emex09608	EMEX_Dicer	1748	NP_524453.1	2,00E-82	2249
			NP_523778.2	0.0	1722
Emex11192	EMEX_Argonaute	893	NP_725341.1	0.0	984
			NP_648775.1	3,00E-103	1214
			NP_001036627.2	2,00E-39	867
Emex05584	EMEX_Argonaute	1640	NP_725341.1	1,00E-132	984
			NP_648775.1	7,00E-137	1214
			NP_001036627.2	2,00E-35	867
Emex06719	EMEX_Ranbp21	1217	NP_001285453.1	0.0	1241
Emex12383	EMEX_Pasha	639	NP_651879.1	0.0	642
Emex08003	EMEX_Tudor	2545	NP_476773.1	1E-44	2515
Emex12114	EMEX_Fmr1	597	NP_611645.1	2,00E-160	684
Emex09481	EMEX_Vig2	415	NP_733084.1	8,00E-33	412
Emex10064	EMEX_Loquacious	292	NP_723813.1	1,00E-68	419
Emex08666	EMEX_Loquacious	332	NP_723813.1	5,00E-30	419
Hlab12928	HLAB_Drosha	1309	NP_477436.1	0.0	1327
Hlab03512	HLAB_Dicer	1361	NP_524453.1	0.0	2249
			NP_523778.2	5,00E-68	1722
Hlab00411	HLAB_Dicer	1770	NP_524453.1	3,00E-39	2249
			NP_523778.2	4,00E-67	1722
			NP_725341.1	0.0	984
Hlab08547	HLAB_Argonaute	535	NP_648775.1	6,00E-77	1214
			NP_001036627.2	1,00E-39	867
			NP_725341.1	6,00E-124	984
Hlab06842	HLAB_Argonaute	863	NP_648775.1	3,00E-131	1214
			NP_001036627.2	1,00E-37	867
			NP_001285453.1	0.0	1241
Hlab01347	HLAB_Ranbp21	1219	NP_001285453.1	0.0	1241
Hlab04648	HLAB_Pasha	626	NP_651879.1	0.0	642
Hlab11079	HLAB_Tudor	2549	NP_476773.1	2E-54	2515
Hlab01311	HLAB_Fmr1	636	NP_611645.1	3,00E-173	684
Hlab09905	HLAB_Vig2	382	NP_733084.1	1,00E-30	412
Hlab07815	HLAB_Loquacious	362	NP_723813.1	2,00E-101	419

	Hlab05217	HLAB_Loquacious	329	NP_723813.1	5,00E-24	419
<i>Lasioglossum albipes</i>	Lalb_00686	LALB_Drosha	1305	NP_477436.1	0.0	1327
	Lalb_01487	LALB_Dicer	1888	NP_524453.1	0.0	2249
				NP_523778.2	2,00E-55	1722
	Lalb_01265	LALB_Dicer	2263	NP_524453.1	3,00E-94	2249
				NP_523778.2	0.0	1722
	Lalb_03345	LALB_Argonaute	922	NP_725341.1	0.0	984
				NP_648775.1	5,00E-104	1214
				NP_001036627.2	2,00E-37	867
	Lalb_04947	LALB_Argonaute	1109	NP_725341.1	2,00E-75	984
				NP_648775.1	2,00E-80	1214
				NP_001036627.2	2,00E-29	867
	Lalb_12680	LALB_Ranbp21	1202	NP_001285453.1	0.0	1241
	Lalb_07036	LALB_Pasha	644	NP_651879.1	0.0	642
	Lalb_03756	LALB_Tudor	2553	NP_476773.1	1,00E-57	2515
	Lalb_03119	LALB_Fmr1	615	NP_611645.1	5,00E-144	684
Lalb_02581	LALB_Vig2	423	NP_733084.1	2,00E-34	412	
Lalb_13234	LALB_Loquacious	381	NP_723813.1	5,00E-105	419	
Lalb_11838	LALB_Loquacious	338	NP_723813.1	3,00E-25	419	
<i>Melipona quadrifasciata</i>	Mqua07049	MQUA_Drosha	1303	NP_477436.1	0.0	1327
	Mqua14519	MQUA_Dicer	1487	NP_524453.1	0.0	2249
				NP_523778.2	3,00E-56	1722
	Mqua02257	MQUA_Dicer	1551	NP_524453.1	8,00E-90	2249
				NP_523778.2	0.0	1722
	Mqua07112	MQUA_Argonaute	822	NP_725341.1	0.0	984
				NP_648775.1	2,00E-87	1214
				NP_001036627.2	4,00E-38	867
	Mqua06975	MQUA_Argonaute	1267	NP_725341.1	1,00E-151	984
				NP_648775.1	4,00E-149	1214
				NP_001036627.2	7,00E-32	867
	Mqua07008	MQUA_Ranbp21	1204	NP_001285453.1	0.0	1241
	Mqua04933	MQUA_Pasha	623	NP_651879.1	0.0	642

Megachile rotundata

Mqua05689	MQUA_Tudor	2484	NP_476773.1	1,00E-48	2515
Mqua11574	MQUA_Fmr1	637	NP_611645.1	7,00E-172	684
Mqua05168	MQUA_Vig2	432	NP_733084.1	2,00E-29	412
Mqua09094	MQUA_Loquacious	400	NP_723813.1	2,00E-95	419
Mqua01890	MQUA_Loquacious	333	NP_723813.1	1,00E-27	419
Mrot00920	MROT_Droscha	1303	NP_477436.1	0.0	1327
Mrot05715	MROT_Dicer	1595	NP_524453.1	6E-91	2249
			NP_523778.2	0.0	1722
Mrot13039	MROT_Dicer	1952	NP_524453.1	0.0	2249
			NP_523778.2	4,00E-71	1722
Mrot02381	MROT_Argonaute	921	NP_725341.1	0.0	984
			NP_648775.1	3,00E-109	1214
			NP_001036627.2	5,00E-37	867
Mrot10140	MROT_Argonaute	2566	NP_725341.1	6,00E-145	984
			NP_648775.1	1,00E-142	1214
			NP_001036627.2	7,00E-34	867
Mrot04147	MROT_Ranbp21	1193	NP_001285453.1	0.0	1241
Mrot10599	MROT_Pasha	635	NP_651879.1	0.0	642
Mrot08863	MROT_Tudor	2568	NP_476773.1	2E-42	2515
Mrot07844	MROT_Fmr1	722	NP_611645.1	1,00E-165	684
Mrot05578	MROT_Vig2	415	NP_733084.1	3,00E-25	412
Mrot12322	MROT_Loquacious	378	NP_723813.1	2,00E-107	419
Mrot11283	MROT_Loquacious	268	NP_723813.1	6,00E-27	419