

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
BACHARELADO EM BIOTECNOLOGIA**

Gustavo Fernando Ferreira Gonçalves

**ANÁLISE DE EXPRESSÃO DIFERENCIAL ENTRE ESTÁGIOS DE
DESENVOLVIMENTO DE *Mahanarva spectabilis* (Hemiptera) PARA
PROSPECÇÃO DE GENES ALVOS NO CONTROLE MEDIADO POR RNAi**

São Carlos - SP

2025

**UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
BACHARELADO EM BIOTECNOLOGIA**

Gustavo Fernando Ferreira Gonçalves

**ANÁLISE DE EXPRESSÃO DIFERENCIAL ENTRE ESTÁGIOS DE
DESENVOLVIMENTO DE *Mahanarva spectabilis* (Hemiptera) PARA
PROSPECÇÃO DE GENES ALVOS NO CONTROLE MEDIADO POR RNAi**

**Trabalho de Conclusão de Curso
apresentado à Universidade Federal de
São Carlos como requisito para a
obtenção do título de Bacharel em
Biotecnologia.**

**Orientadora: Dra. Bianca Baccili
Zanotto Vigna**

**Coorientadora: Profa. Dra. Andrea
Soares da Costa Fuentes**

São Carlos - SP

2025

AGRADECIMENTOS

À minha mãe, Suzana Aparecida Ferreira, e ao meu pai, Ademilson Gonçalves, por todo o amor, apoio e confiança. São os seus ensinamentos, conselhos, cuidados e sonhos que me guiam e me permitem trilhar na vida os mais belos caminhos.

Aos amigos Dario, Victor, Thomas, Sabrina e Matheus, por todos os momentos que pudemos compartilhar nestes anos, e que tenhamos ainda muitos pela frente.

À minha companheira Eduarda Valentim, pelo amor, pela ajuda, e por estar sempre ao meu lado me trazendo a paz e a leveza necessária.

Ao Dr. Marcos Gusmão por me receber, me motivar, e por acreditar em minhas capacidades.

Ao Dr. Wilson Malagó por toda a tutoria e companhia nas atividades em bancada.

À Ma. Isabela Begnami pela inspiração e pelo auxílio com as atividades em bioinformática.

À Dra. Bianca Vigna pela orientação, pelo direcionamento e pela confiança depositada em mim.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelas bolsas de iniciação científica que possibilitaram a realização deste trabalho.

*"A coisa mais bonita sobre o
aprendizado é que ninguém pode
tirá-lo de você."*

(B. B. King)

RESUMO

A cigarrinha-das-pastagens *Mahanarva spectabilis* (Distant, 1909) é considerada uma praga-chave das pastagens, pois ao sugar a seiva do xilema pode levar a perdas de até 20% na produção de biomassa em gramíneas. A resistência a estes ataques é uma característica desejável em qualquer cultivar de forrageira a ser lançada no Brasil, além de outras gramíneas de grande porte (como a cana de açúcar) que também são afetadas. A busca por caminhos ecologicamente sustentáveis para o manejo de insetos pragas envolve o uso de ferramentas moleculares como a interferência por RNA (RNAi). Este mecanismo é utilizado como um método de indução de resistência que traz como vantagens maior especificidade de alvos, maior eficiência de controle e menor impacto sobre o ecossistema. Para o desenvolvimento de biodefensivos utilizando a tecnologia de RNAi, algumas etapas são importantes, como a busca de genes alvos candidatos a serem silenciados de forma que o dsRNA (double-strand RNA, ou RNA dupla fita) desenvolvido não encontre alvos nas plantas cultivadas ou em organismos não-alvos ("off targets"). Além disso, os genes alvos precisam estar envolvidos em processos fisiológicos críticos para a sobrevivência do inseto, de modo que, ao serem silenciados, gerem um fenótipo preferencialmente letal. Neste estudo foi realizada uma montagem *de novo* do transcriptoma de *M. spectabilis* a partir de amostras de ovos, ninfas e adultos, a investigação da existência do mecanismo de RNAi no organismo e a prospecção de genes alvo para serem silenciados. O transcriptoma obtido foi considerado adequado em termos de tamanho e representação dos RNAs fornecidos, com 197.003 *contigs*, 25.492 deles (12,94%) com anotação positiva em Blast, tamanho médio de 364 pb por *contig*, N50 de 853 pb e 34,48% de conteúdo GC. Os resultados revelam o alto número de transcritos referentes a componentes essenciais na resposta de RNAi encontrados no transcriptoma (como Dicer 1-2, Argonauta 1-2-3, Piwi, Spindle, Gawky, Armi, dentre outros), apontando a viabilidade do uso desta ferramenta para controle do inseto. A análise de expressão gênica categorizou 10.928 *contigs* (5,54%) presentes no transcriptoma como DEGs, sendo as comparações de ovos com adultos a mais distinta e ninfas pequenas com ninfas grandes a mais semelhante. Além disso, a partir da análise de função e de expressão, cinco genes envolvidos em diferentes etapas fundamentais da sobrevivência do inseto e com alta expressão (V-ATPase, trealase, calmodulina, apolipoproteína D e hexoquinase tipo 2) foram selecionados como

possíveis alvos de um dsRNA que direcionaria um silenciamento eficiente e especializado.

Palavras-chave: cigarrinha-das-pastagens, transcriptoma, silenciamento gênico.

LISTA DE ILUSTRAÇÕES

Figura 1 – Fases do desenvolvimento da cigarrinha-das-pastagens.	9
Figura 2 – Diferentes abordagens de extração de RNA total.	14
Figura 3 – Imagens das eletroforeses dos RNAs totais das 24 amostras de <i>M. spectabilis</i>	22
Figura 4 – Representação do valor médio de qualidade em cada posição de base do sequenciamento de RNA mensageiro de <i>M. spectabilis</i>	23
Figura 5 – Representação da distribuição da quantidade de transcritos diferencialmente expressos entre as seis comparações de amostras.	30
Figura 6 – Representação das mudanças de logFC entre comparações de amostras de <i>M. spectabilis</i> sequenciadas..	31
Figura 7 – Comparação entre contagem média por biblioteca de cDNA dos cinco genes candidatos a alvo no controle de <i>M. spectabilis</i>	36
Figura A1 – Representação da comparação entre as expressões relativas de genes em RNA-Seq e RT-qPCR.	43
Figura A2 – Representação da dispersão entre as expressões em RNA-Seq e RT-qPCR.	44

LISTA DE TABELAS

Tabela 1 – Delineamento experimental das doze bibliotecas de cDNA construídas.	15
Tabela 2 – Determinação dos genes utilizados nas reações de RT-qPCR para validação.	20
Tabela 3 – Quantidade média de reads do sequenciamento de <i>M. spectabilis</i> mantidas durante as etapas de filtragem com Trimmomatic e SortMeRNA.	24
Tabela 4 – Parâmetros do transcriptoma de <i>M. spectabilis</i> montado com o Trinity.	25
Tabela 5 – Parâmetros do transcriptoma de <i>M. spectabilis</i> após filtragem com CD-HIT.	25
Tabela 6 – Resultados de análise com BUSCO v3.	26
Tabela 7 – Resultados de análise com Bowtie2.	26
Tabela 8 – Genes da maquinaria de RNAi identificados no transcriptoma montado de <i>M. spectabilis</i> , categorizados por grupo de função.	28
Tabela 9 – Pares de primers utilizados na validação de expressão por RT-qPCR em <i>M. spectabilis</i>	33
Tabela 10 – Genes selecionados como alvo para silenciamento por RNA a partir da literatura e do transcriptoma de <i>M. spectabilis</i> , com anotação e perfil de expressão destes.	34

SUMÁRIO

1	INTRODUÇÃO.....	9
1.1	Cigarrinhas-das-pastagens	9
1.2	Interferência por RNA.....	10
1.3	Transcriptoma e prospecção de genes	11
2	OBJETIVOS.....	12
2.1	Objetivo geral	12
2.2	Objetivos específicos	13
3	MATERIAL E MÉTODOS	13
3.1	Coleta de amostras	13
3.2	Extração de RNA total.....	14
3.3	Sequenciamento	15
3.4	Montagem <i>de novo</i> e anotação de transcriptoma	16
3.5	Identificação de genes da maquinaria de RNAi.....	17
3.6	Análise de expressão diferencial	18
3.7	Seleção de candidatos à alvo	18
3.8	Validação da expressão gênica	19
4	RESULTADOS E DISCUSSÃO.....	21
4.1	Qualidade de extrações e bibliotecas	21
4.2	Sequenciamento, montagem e anotação	23
4.3	Maquinaria de RNAi	28
4.4	Transcritos diferencialmente expressos.....	29
4.5	Validação por RT-qPCR	32
4.6	Potenciais alvos	34
5	PERSPECTIVAS.....	37
6	REFERÊNCIAS	38
	APÊNDICE A – PROCEDIMENTO DE VALIDAÇÃO DA EXPRESSÃO GÊNICA	43

1 INTRODUÇÃO

1.1 Cigarrinhas-das-pastagens

As cigarrinhas-das-pastagens são insetos sugadores pertencentes à ordem *Hemiptera*, subordem *Auchenorrhyncha* e família *Cercopidae*, embora uma classificação ainda em uso as mantenha na ordem Homoptera (VALÉRIO, 2009). Estão associados principalmente às gramíneas e seu desenvolvimento ocorre por hemimetabolia, passando pelas fases de ovo, ninfa e estágio adulto, como exemplificado na Figura 1. As fêmeas depositam ovos no solo ou em restos vegetais, com uma média de cem ovos por fêmea, podendo ocorrer um período de incubação prolongado (conhecido como diapausa) de até duzentos dias em condições adversas (VALÉRIO, 2009).

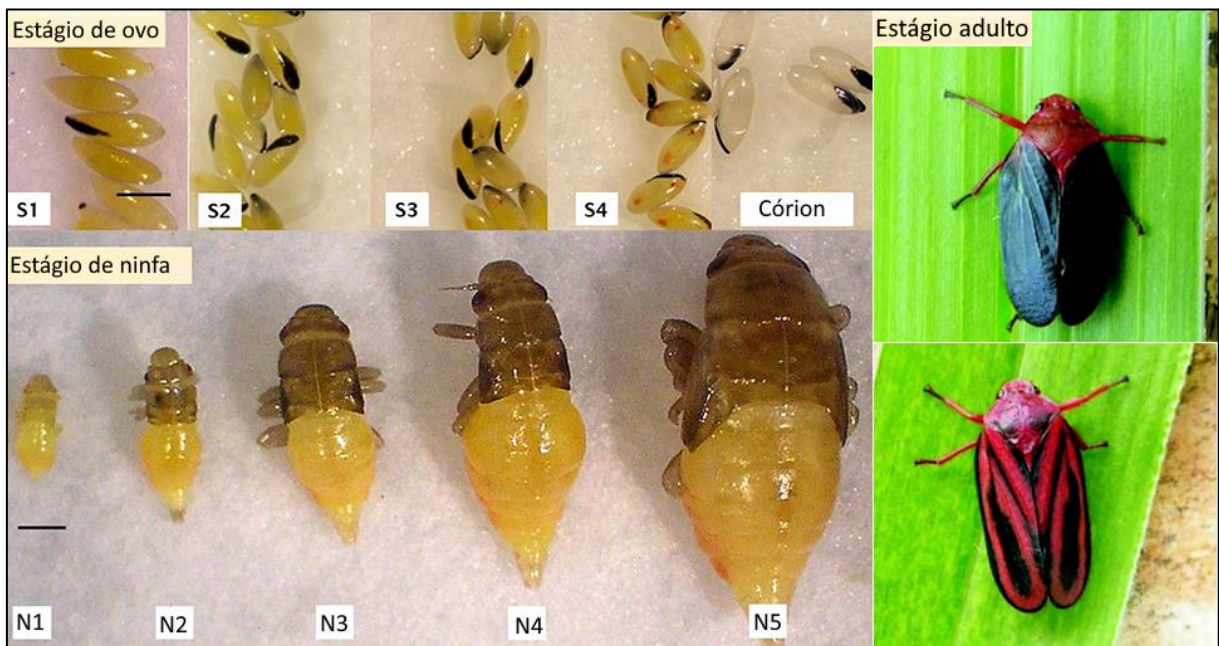


Figura 1 – Fases do desenvolvimento de cigarrinhas da família *Cercopidae*. Ovos (*Aeneolamia varia*), ninfas (*Prosapia simulans*) e adultos (*Mahanarva spectabilis*). Fonte: adaptado de ESPITIA et al., 2022 e PAULA-MORAES et al., 2006.

Através da sua alimentação tanto em fase de ninfa quanto adulta, o inseto é capaz de causar grandes perdas na disponibilidade e na qualidade de gramíneas forrageiras (VALÉRIO, 2009). O controle químico da cigarrinha é possível, mas não eficaz pois acaba sendo antieconômico e antiecológico, especialmente devido ao tamanho das plantações (AUAD et al., 2012). Fungos entomopatogênicos do gênero *Metarhizium* são destacados como casos de sucesso no controle biológico de

cigarrinhas (IWANICKI et al., 2019), porém sua eficiência a longo prazo e usabilidade ainda permanecem em aberto. Em comparação, o manejo mais econômico concentra-se no controle por meio de plantas resistentes ao ataque desse inseto, resistência esta advinda de antibiose (a planta tem um efeito negativo na aptidão do inseto) e/ou tolerância (a capacidade da planta de suportar os danos da praga reparando lesões, com crescimento ou com aumento da reprodução (CARDONA et al., 2004)).

1.2 Interferência por RNA

A interferência por RNA (RNAi) é um mecanismo regulatório presente na maioria das células eucarióticas que utiliza pequenas moléculas de RNA dupla fita (dsRNA) como “gatilhos” que direcionam, através de complementariedade, o silenciamento de genes (KIM et al., 2008). A presença de dsRNA livre desencadeia a clivagem dessas moléculas em fragmentos menores, os siRNAs (small interfering RNAs), que serão depois atraídos para um complexo proteico denominado RISC (RNA-induced silencing complex), que contém uma nuclease e uma helicase de RNA (ALBERTS et al., 2017, p. 429). Os siRNA's direcionam o complexo enzimático para outras moléculas de RNA que possuem sequências de nucleotídeos complementares e as degradam (ALBERTS et al., 2017, p. 431).

Cada vez que cliva um novo RNA, o RISC é “regenerado” com uma pequena molécula de RNA, de maneira que uma única molécula de dsRNA original pode atuar cataliticamente para destruir muitos RNAs complementares. Além disso, os próprios pequenos produtos de clivagem do dsRNA podem ser replicados por enzimas celulares adicionais, possibilitando uma amplificação da atividade de RNAi (ALBERTS et al., 2017, p. 431).

No contexto de controle de pragas agrícolas o mecanismo de RNAi representa uma ferramenta importante que, ao silenciar um ou mais genes específicos do inseto, pode ocasionar a morte do mesmo e conseqüentemente evitar prejuízos advindos da infestação (CAGLIARI et al., 2020).

Na literatura estão descritos diferentes genes testados como alvos de silenciamento, com resultados igualmente diversos em taxas de mortalidade atingidas (MEZZETTI et al, 2020). Estas diferenças são atribuídas a características bioquímicas e fisiológicas do inseto, como a presença de nucleases em suas glândulas salivares,

intestino médio ou hemolinfa, que acabam degradando as moléculas de dsRNA antes de serem clivadas (PARISE et al, 2024).

Coleópteros, como o besouro da batata do Colorado (*Leptinotarsa decemlineata*) são particularmente sensíveis ao controle com RNAi, porém no caso de hemípteros há uma maior variação nas respostas quando foram administrados dsRNAs (PARISE et al, 2024). Apesar disso, diversos trabalhos obtiveram resultados promissores ao se utilizarem desta ferramenta em diferentes espécies de Hemiptera como cicadélídeos, psilídeos e fulgoromorpha (PARISE et al, 2024).

A entrega deste dsRNA direcionador ao inseto praga pode ser feita através de um produto de uso tópico, onde este dsRNA é previamente sintetizado e estabilizado em solução para que possa ser pulverizado em baixa ou alta pressão, permanecendo na superfície da folha (DALAKOURAS et al, 2024). Outra alternativa seriam plantas transplastômicas que expressam dsRNAs no cloroplasto, uma vez que os cloroplastos não possuem maquinaria de RNAi e consequentemente, os dsRNAs expressos permanecem não-processados e intactos para absorção por insetos (DALAKOURAS et al, 2024).

1.3 Transcriptoma e prospecção de genes

Para que a interferência por RNA atinja níveis desejados de eficiência, é primordial que seja feita uma boa escolha de genes alvo do silenciamento. Um gene ideal é aquele vital para a sobrevivência e/ou desenvolvimento do inseto, que esteja conservado dentro da população e que apresente altos níveis de expressão (MAMTA et al., 2017). Além disso, o gene alvo do silenciamento deve apresentar menor impacto possível sobre o ecossistema, ou seja, não pode desencadear malefícios sobre as plantas cultivadas nem sobre os demais organismos que convivem naquele ambiente, tendo como alvo exclusivo o inseto praga (MAMTA et al., 2017).

Visando identificar os alvos de RNAi mais estratégicos podemos utilizar ferramentas moleculares e de bioinformática para o levantamento e a análise de expressão gênica diferencial em cigarrinhas-das-pastagens, considerando as diferentes fases de desenvolvimento do inseto. O transcriptoma refere-se ao conjunto total de transcritos observados em determinada amostra biológica, sendo que normalmente estas informações estão atreladas a condições específicas e distintas entre si (RAGHAVAN et al., 2022). Neste caso, a avaliação de amostras coletadas em

diferentes estágios de desenvolvimento permite identificar variações na expressão gênica, relacionadas às distintas fases da hemimetabolia e em vias metabólicas importantes para o inseto, além de permitir verificar a existência e expressão de genes da maquinaria de RNAi.

Considerando-se a incerteza acerca da conservação do mecanismo de RNAi na classe *Insecta*, uma etapa adicional comumente realizada na bibliografia corrente (PARISE et al., 2024; SCHVARTZMAN et al., 2022; CAMARGO et al., 2015) tem sido a verificação da presença ou ausência de transcritos relacionados a genes referentes à maquinaria de RNAi. A família de proteínas Dicer, Argonautas e proteínas envolvidas em transporte, ligação e absorção de dsRNA são exemplos dos genes da maquinaria de RNAi buscados no transcriptoma (CAMARGO et al., 2015).

Diante da importância da seleção de genes alvo específicos e eficientes e do papel dos genes relacionados à maquinaria de RNAi em *M. spectabilis*, algumas questões relevantes a serem investigadas são: Quais seriam os genes expressos mais diferentes ou semelhantes entre os diferentes estágios de desenvolvimento da hemimetabolia? Quão completo está o repertório de proteínas envolvidas no mecanismo de RNAi nesta espécie? Quais seriam bons genes alvo cujo silenciamento acarretaria no aumento significativo na mortalidade do inseto?

Nossas hipóteses são (I) os genes ideais para o silenciamento são aqueles envolvidos em funções essenciais para sobrevivência e maturação do inseto e que estes podem ser identificados de maneira precisa e segura através de uma análise do transcriptoma do inseto; (II) dentre as fases de ovo, ninfa e adulto, as principais diferenças estariam entre os adultos e as demais pois são os mais morfologicamente e fisiologicamente diferenciados; e (III) a maquinaria de RNAi está completa em *M. spectabilis*, considerando que este é o caso para a maior parte da classe *Insecta*, apesar de algumas exceções descritas na literatura.

2 OBJETIVOS

2.1 Objetivo geral

O presente trabalho teve como objetivo confirmar a presença das proteínas envolvidas no mecanismo de RNAi em *Mahanarva spectabilis* (Distant, 1909) e identificar genes alvos candidatos para serem silenciados por RNAi, visando suas aplicações em gramíneas de importância agrônômica.

2.2 Objetivos específicos

- I. Extrair e purificar RNA total de amostras de *M. spectabilis* em fase de ovo, ninfa e adulto;
- II. Realizar montagem *de novo* do transcriptoma das diferentes fases de desenvolvimento através das sequências obtidas no sequenciamento;
- III. Verificar a expressão de genes relacionados à maquinaria de RNAi;
- IV. Verificar os genes diferencialmente expressos entre as diferentes fases de desenvolvimento;
- V. Estabelecer candidatos a genes alvos (junto de seus ortólogos) para serem silenciados no inseto;
- VI. Validar a expressão observada *in silico* através de PCR por transcrição reversa quantitativo de genes diferencialmente expressos e genes de referência.

3 MATERIAL E MÉTODOS

3.1 Coleta de amostras

Durante a safra de 2020 e 2021 foram coletados indivíduos adultos de *M. spectabilis* em pasto de *Urochloa brizantha* cultivar Marandu, na Embrapa Pecuária Sudeste, utilizando-se puçá entomológico. Através destes espécimes, foram obtidos em laboratório ovos do inseto que foram mantidos em B.O.D a 25°C até atingirem o último estágio de desenvolvimento, sendo assim coletados em microtubo e congelados em freezer -80°C. Em março de 2021 foram coletadas amostras de ninfas nos diferentes estágios de desenvolvimento (N1 a N5) e adultos machos e fêmeas no mesmo campo. Os indivíduos foram colocados em microtubos e imediatamente imersos em nitrogênio líquido, para posterior armazenamento em freezer -80°C.

Todas as coletas e preparos de amostras foram realizadas por demais colaboradores do trabalho, de modo que o autor foi incluído ao grupo quando todas as amostras de ovos, ninfas e adultos já estavam armazenadas em freezer e prontas para a extração.

3.2 Extração de RNA total

Devido a variedade de tamanhos dos insetos e de conteúdos de quitina nas amostras, duas abordagens diferentes de extração de RNA total foram utilizadas visando otimizar a quantidade e qualidade de material obtido.

Para amostras de ovos e ninfas de estágio 1, 2 e 3 (pequenas quantidades de tecido), foi utilizado o kit Quick-RNA™ Tissue/Insect Microprep (Zymo Research) com protocolo fornecido pela fabricante (Figura 2-A). Já no caso de amostras de ninfas nos estágios 4 e 5 e de adultos, onde a quantidade de tecido era maior, assim como a quantidade de quitina, a abordagem utilizada foi o TRI Reagent da Sigma-Aldrich com protocolo adaptado (Figura 2-B). Para os ovos, foram utilizados 100 ovos para duas replicatas cada e uma terceira replicata com 87 ovos. Para as ninfas e adultos, cada amostra corresponde a um único animal.

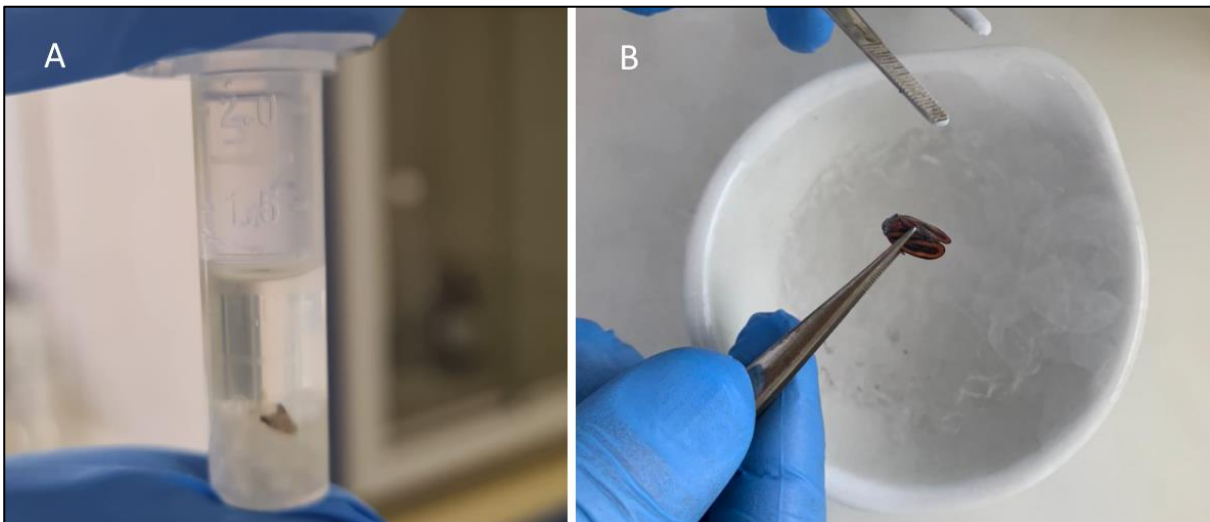


Figura 2 – Diferentes abordagens de extração de RNA total. (A) Microtubo contendo microesferas e ninfa de terceiro instar após agitação e homogeneização. (B) Inseto adulto no momento de pulverização com almofariz e pistilo. Fonte: elaborado por Isabela S. Begnami.

Após as extrações de RNA total, obtivemos 24 amostras categorizadas como: ovos (OVO), ninfas de 1° instar (N1), ninfas de 2° instar (N2), ninfas de 3° instar (N3), ninfas de 4° instar (N4), ninfas de 5° instar (N5), adulto macho (AM) e adulto fêmea (AF), cada categoria em triplicata.

Visando estimar os parâmetros de qualidade e quantidade dos RNAs extraídos, todas as amostras foram submetidas a análise com o espectrofotômetro NanoDrop®. Foram avaliadas as absorvâncias nos comprimentos de onda de 230, 260 e 280 nm

e as razões 260/230 e 260/280, parâmetros característicos de avaliações de ácido ribonucleico. A primeira razão indica a pureza do produto de extração em relação à presença de compostos orgânicos ou sais caotrópicos e a segunda indica possível contaminação por proteínas remanescentes da purificação (BRIAN, 2015).

Além disso, para visualização da integridade e fragmentação dos RNAs, os mesmos foram submetidos à eletroforese em gel de agarose 1% corado com brometo de etídeo 0,5 µg/mL. Para tal, foram aplicadas alíquotas das amostras com tampão de corrida TBE (Tris-Borato-EDTA, uma mistura de base tamponante Tris, ácido bórico e ácido etilenodiamino tetra-acético) 1X no gel, o qual foi submetido a uma corrente de 100V por 1 hora. Na sequência o gel foi transportado para o equipamento Bio-Rad Gel Doc™ XR Molecular Imager® no qual pudemos visualizar o resultado.

3.3 Sequenciamento

A partir das 24 amostras extraídas foram construídas 12 bibliotecas de DNA complementar (cDNA) para sequenciamento (Tabela 1). Para isso, os materiais de ninfa 1, 2 e 3 foram agrupados (a partir de informações sobre concentração, para garantir quantidades iguais de RNA em cada biblioteca) e nomeados como “ninfa pequena, NP”, idem para ninfas 4 e 5 que receberam o nome “ninfa grande, NG”. Adultos macho e fêmea também foram mesclados em apenas uma amostra categorizada como “adulto, AD”.

Tabela 1 – Delineamento experimental das doze bibliotecas de cDNA construídas. (NG): Ninfas grandes; (NP): Ninfas pequenas; (AD): Adultos; Fonte: elaborado pelo autor.

Bibliotecas construídas		
Ovos	Ninfas	Adultos
OVO1	NP1 (N1+N2+N3)	AD1 (AM+AF)
OVO2	NP2 (N1+N2+N3)	AD2 (AM+AF)
OVO3	NP3 (N1+N2+N3)	AD3 (AM+AF)
	NG1 (N4+N5)	
	NG2 (N4+N5)	
	NG3 (N4+N5)	

As bibliotecas de cDNA foram construídas no Laboratório de Análise Genética e Molecular (LAGM) da Unicamp utilizando o kit TruSeq Stranded mRNA (Illumina), de acordo com instruções do fabricante. A qualidade das bibliotecas obtidas foi avaliada

através de eletroforese automática no equipamento 4200 Tape Station System (Agilent). Estas etapas foram realizadas pela estudante de doutorado Isabela dos Santos Begnami associada a este trabalho.

O Sequenciamento de Nova Geração (NGS) foi realizado utilizando a plataforma Illumina NextSeq 550 2x75pb no Laboratório de Sequenciamento em Larga Escala (SELA) da Faculdade de Medicina da Universidade de São Paulo como prestação de serviço.

A qualidade dos dados obtidos foi avaliada com FastQC (ANDREWS, 2010) e os resultados agregados em um relatório único através do MultiQC (EWELS et al., 2016).

3.4 Montagem *de novo* e anotação de transcriptoma

A partir dos dados brutos recebidos do sequenciamento em forma de sequências individuais detectadas (*reads*), uma filtragem com o software Trimmomatic (BOLGER et al., 2014) foi realizada visando recortar os adaptadores Illumina (por meio do módulo ILLUMINACLIP e uma biblioteca específica de sequências de adaptadores), as bases e *reads* de baixa qualidade (com filtros LEADING:3 e TRAILING:3) e aquelas sequências cujo tamanho fosse menor que 35 pb (MINLEN:35).

Uma filtragem adicional com o software SortMeRNA (KOPYLOVA et al., 2012) foi realizada para descartar *reads* que fossem referentes à RNA de função ribossomal residual indesejados para análise da expressão gênica, utilizando-se do banco de dados `smr_v4.3_sensitive_db.fasta` e parâmetro `-num_alignments` de 1 para considerar apenas o melhor alinhamento por leitura.

Devido à ausência de genoma referencial para *M. spectabilis* foi adotada a abordagem *de novo* na montagem do transcriptoma. O software Trinity (GRABHERR et al., 2011), especializado neste modelo de montagem e referência presente na bibliografia atual (KNUTSEN, 2023; MUHAMMAD et al., 2021; HAAS et al, 2013), realizou o processo a partir dos doze arquivos filtrados e seguintes os parâmetros específicos: `--normalize_max_read_cov 30`, `--run_as_paired` e `-NO_SEQTK`. Considerando a alta quantidade de *reads* fornecidos para a tarefa, utilizamos a função de normalização *in silico* para atenuar os possíveis efeitos negativos da superabundância de leituras. Além disso, após a montagem utilizamos o *script* de

filtragem *longest* (`get_longest_isoform_seq_per_trinity_gene.pl`) para manter apenas a isoforma mais longa de cada *cluster*. A principal unidade de informação que compõe o transcriptoma são as chamadas *contigs*, formadas a partir de sobreposições estatisticamente ajustadas entre *reads* concatenadas em sequências contíguas mais longas.

O programa CD-HIT (FU et al., 2012) com parâmetro de `-c 0.8` (80% de identidade) e `-n 5` foi utilizado para eliminar a redundância do transcriptoma montado e a qualidade do mesmo foi averiguada através do BUSCO v3 (SIMÃO et al., 2015) em modo *default* (sem parâmetros específicos), onde suas sequências foram comparadas com um conjunto de genes ortólogos conservados em *Hemiptera*. Por fim, pareamos as *reads paired-end* com o transcriptoma final através do Bowtie2 com até 20 alinhamentos por leitura (`-k 20`) (LANGMEAD; SALZBERG, 2012).

Para inferir e atribuir informações relativas a propriedades biológicas funcionais e evolutivas das sequências levantadas, realizamos a anotação funcional do transcriptoma montado utilizando métodos *in silico* (RAGHAVAN et al., 2022). Esta anotação foi executada via o software DIAMOND versão 2.1.8.162 (BUCKFINK et al., 2021), onde as sequências foram alinhadas com BLASTPX e BLASTP contra o banco de dados da SwissProt (UniProt - <https://www.uniprot.org/>), com um critério de significância $\leq 1e-5$ e maior sensibilidade de busca ativada (`--sensitive`).

3.5 Identificação de genes da maquinaria de RNAi

A busca por genes referentes a maquinaria de interferência por RNA é motivada principalmente pelo fato de que fora constatado em alguns hemípteros a ausência de proteínas fundamentais para o RNAi (JAIN et al., 2020), tornando-se, portanto, muito pouco eficiente o uso deste no controle destes insetos. Assim, a fim de garantir que a entrega do dsRNA resulte em um silenciamento efetivo do gene alvo e aumento da mortalidade de *M. spectabilis*, almejamos diagnosticar a espécie acerca de sua maquinaria necessária.

Para isso, foi realizado um levantamento dos principais genes associados a proteínas envolvidas no mecanismo de resposta RNAi na bibliografia vigente referente à controle de insetos (SINGH et al., 2019; CAGLIARI et al., 2020; CAMARGO et al., 2015; GUPTA et al., 2021). Estes foram catalogados e categorizados por função,

estabelecendo-se assim uma lista dos componentes essenciais para o funcionamento e propagação do silenciamento via RNAi.

Utilizando a base de dados UniProt, selecionamos o “*UniProt name*” dos genes procurados e fizemos a busca no transcriptoma anotado para investigar a presença ou ausência destes em *M. spectabilis*.

3.6 Análise de expressão diferencial

As informações de quantidade de expressão dos transcritos no transcriptoma (expressa em termos de “contagens”, ou *counts*) foram inferidas através de pseudoalinhamento realizado com o software Salmon (PATRO et al., 2017), com um tamanho de k-mer de 31 nucleotídeos (-k 31). Estas informações individuais foram importadas e sumarizadas para o software RStudio (RStudio Team, 2010) através do pacote tximport (SONESON et al., 2015). Assim, obteve-se matrizes resumidas que continham estimativas de contagens para cada transcrito, em cada biblioteca.

Para a análise empírica de expressão diferencial entre as bibliotecas foi utilizado o pacote edgeR 4.4 (CHEN et al., 2024). Nele recursos estatísticos como distribuição binomial negativa e modelos lineares generalizados são empregados na modelagem de dados de *counts* considerando réplicas de *reads*, já que o edgeR implementa estimadores de Bayes empíricos para permitir inferências confiáveis quando o número de réplicas é pequeno (CHEN et al., 2024). Sendo assim, a partir das matrizes de contagens brutas, ou *raw counts*, conseguimos acessar informações sobre *contigs* que apresentem perfil de expressão de um “gene diferencialmente expresso”, ou DEG (differentially expressed gene), sendo estes aqueles que apresentam diferenças relevantes no número de *counts* inferidos de cada biblioteca.

Para isso, estas matrizes brutas foram primeiramente normalizadas mantendo-se apenas *contigs* com contagem relevante em todas as amostras de ao menos um grupo, além de no mínimo 20 contagens por milhão (CPM, medida utilizada para simplificação). Em cada teste de expressão diferencial foi utilizado um limite de taxa de falsas descobertas das detecções (FDR, *False Discovery Rate*) menor ou igual a 0,01 e um logFC (*log fold change*, medida de razão que expressa a diferença de expressão entre amostras) mínimo de 2.

3.7 Seleção de candidatos à alvo

A seleção de genes para silenciamento via RNAi se iniciou por uma consulta na literatura (SINGH et al., 2019; CAGLIARI et al., 2020; SCHVARTZMAN et al., 2022; WANG et al., 2018; HUNTER et al., 2021; ZHANG et al., 2024; YANG et al., 2022) relacionada a RNAi em *Insecta/Hemiptera*, registro e categorização dos candidatos. Em seguida estabelecemos quais os perfis de expressão destes nos nossos dados através de consulta às análises de expressão, verificando se apresentavam ou não comportamento de DEG.

Assim, concatenando informações da literatura acerca de funções biológicas com os padrões de expressão inferidos pelas análises realizadas neste trabalho, selecionou-se um total de cinco genes candidatos para serem alvos no silenciamento.

3.8 Validação da expressão gênica

Como forma de validação das diferenças de expressão apontadas *in silico* realizou-se uma bateria de reações de PCR quantitativo (RT-qPCR). Para isso, com base nos dados de expressão selecionamos transcritos com perfil de desejado, ou seja, que fossem diferencialmente expressos, que apresentassem um nível mínimo de contagens para garantir que sejam satisfatoriamente detectados na PCR, e que se possível tivessem sido acrescidos de informação pela anotação.

Três grupos de genes foram delineados: sete selecionados dentre aqueles caracterizados como DEGs; cinco referentes a diferentes componentes da maquinaria de RNAi (além de também serem DEGs); e quatro normalizadores, dois estabelecidos a partir da literatura (ZHANG et al., 2021; SINGH et al., 2019) e dois escolhidos no transcriptoma (Tabela 2).

Tabela 2 – Determinação dos genes utilizados nas reações de RT-qPCR para validação. *O sétimo gene dentre os DEGs não possuía informações de anotação, mas foi escolhido pelo perfil de expressão. Fonte: elaborado pelo autor.

Genes utilizados para validação por RT-qPCR			
Tipo	Anotação	Descrição da proteína	Abreviação
Perfil de DEG	CU16_MANSE	Proteína de cutícula larval	CU
	SSP1_BOMMO	Proteína de armazenamento sexo-específica 1	SS
	SP63_STRPU	Proteína de membrana flagelar de esperma de 63 kDa	SP
	HP8_BOMMO	Serina protease HP8 contendo domínio CLIP	HPI
	BETA_ECO57	Colina desidrogenase dependente de oxigênio	BE
	PGSB1_DROME	Proteína de reconhecimento de peptidoglicano SB1	PG
	Não anotado*	Não anotado*	NA
Maquinaria de RNAi	AGO2_RABIT	Proteína argonauta 2	AG
	SIWI_BOMMO	Proteína Siwi semelhante a Piwi	PI
	SIDT2_HUMAN	Membro 2 da família transmembrana SID1	SI
	SCRB1_BOVIN	Membro 1 da classe B do receptor Scavenger	SC
	GAWKY_DROME	Proteína Gawky	GA
Normalizador	TSR3_MOUSE	RNA ribossômico 18S	18
	NSUN5_DROME	RNA ribossômico 28S	28
	FUT10_CANLF	Alfa(1,3)fucosiltransferase 10	FU
	APT_RAT	Adenina fosforribosiltransferase	API

Utilizando as sequências destas *contigs* selecionadas, desenhamos *primers* (pequenos fragmentos iniciadores de síntese) que seriam necessários nas reações. Com o uso das ferramentas Primer3Plus (UNTERGASSER et al., 2007) e NetPrimer (<http://www.premierbiosoft.com/netprimer>), estabelecemos as sequências mais adequadas para garantir a maior taxa de alinhamento, analisando características como comprimento do *primer* (entre 18 e 25 nucleotídeos), temperatura de *melting* (entre 55°C e 65°C), conteúdo de GC (mantido entre 40% e 60%), formação de estruturas secundárias (energia de formação superior a -2 kcal/mol para evitar *hairpins* e dímeros) e assim averiguamos as performances esperadas.

As etapas seguintes da validação (síntese de cDNA, montagem de reações RT-qPCR e análise de resultados) foram realizadas pela doutoranda Isabela dos Santos Begnami, com acompanhamento do autor. Montagem, execução e resultados estão descritos no Apêndice A.

4 RESULTADOS E DISCUSSÃO

4.1 Qualidade de extrações e bibliotecas

Os resultados das leituras de RNA total em espectrofotômetro foram registrados em planilhas, onde averiguamos que as extrações foram satisfatórias em termos de quantidade já que as concentrações ng/ μ L registradas apresentaram valor médio de 165,66 para os ovos, 572,92 para as ninfas pequenas, 1071,28 para as ninfas grandes e 950,01 para os adultos, e que a quantidade de entrada requisitada para o kit de construção de bibliotecas é de 0,1 – 1 μ g.

Na razão A260/A280 obtivemos média de 2,04 para os ovos, 2,10 para as ninfas pequenas, 2,09 para as ninfas grandes e 2,08 para adultos. Já a razão A260/A230 apresenta 1,79 para os ovos, 2,10 para as ninfas pequenas, 2,05 para ninfas grandes e 2,02 para adultos. De acordo com BRIAN (2015), uma razão ~2,0 é geralmente aceita como “pura” para RNA, portanto averiguamos que as extrações foram satisfatórias também em termos de pureza.

Três corridas de eletroforese em gel de agarose foram realizadas e os resultados foram registrados em imagens (Figura 3).

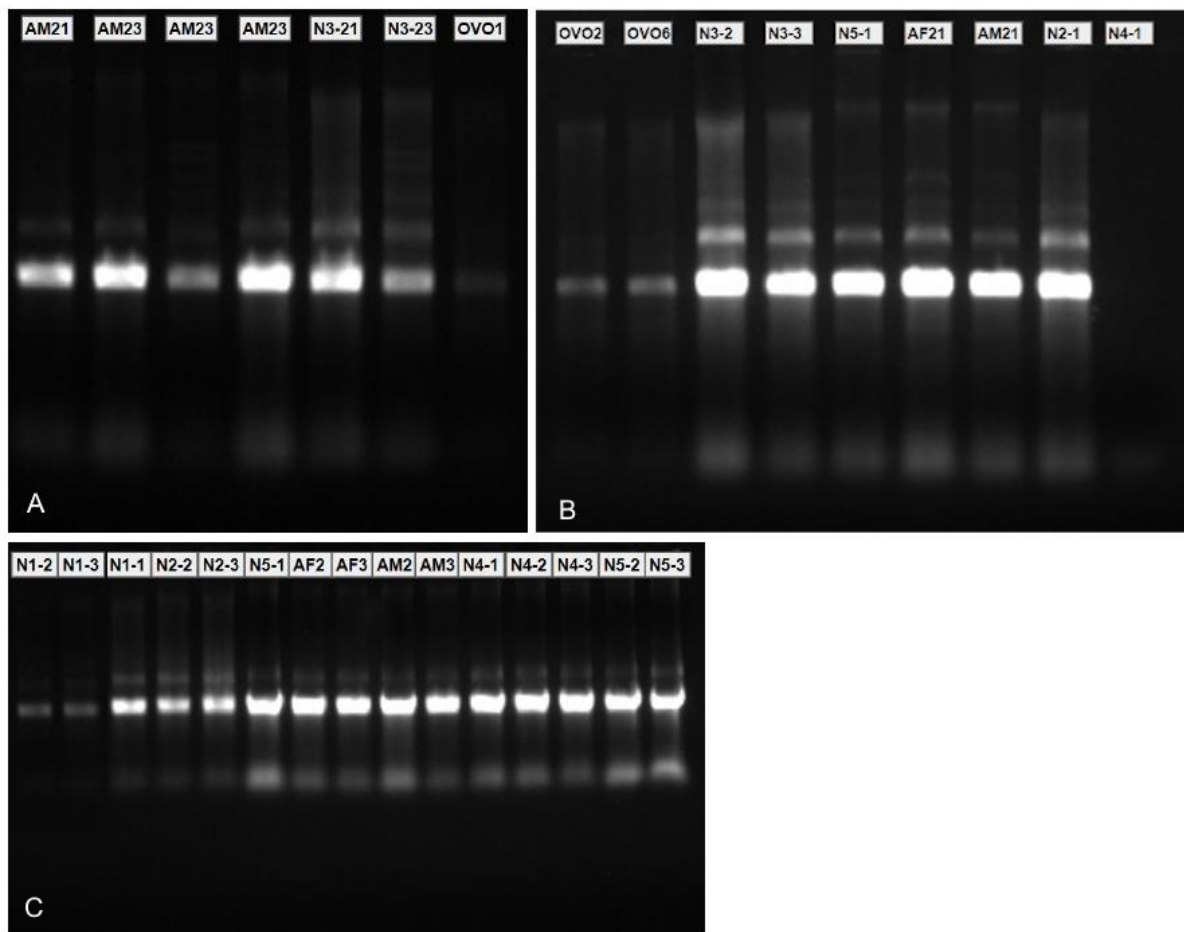


Figura 3 – Imagens das eletroforeses dos RNAs totais das 24 amostras de *M. spectabilis*. (A) Gel com sete amostras, (B) Gel com nove amostras e (C) Gel com 15 amostras, cronologicamente, de diferentes eletroforeses em gel de agarose 1% corados com brometo de etídeo 0,5 µg/mL em tampão de corrida TBE 1X, por 1 hora a 100V, e visualizados em sistema de fotodocumentação Bio-Rad Gel Doc™ XR Molecular Imager®. Os nomes das amostras estão indicados nas imagens, de acordo com a seguinte codificação: N refere-se a ninfa; AM refere-se a adulto macho e AF refere-se a adulto fêmea; Fonte: elaborado pelo autor.

Para a maioria das amostras houve a formação de uma banda intensa com posição central, indicando a presença e integridade esperada do RNA total extraído. No caso do acesso N4-1 (ninfa em quarto instar, primeira repetição) que originalmente não apresentou banda alguma devido à extração com kit (Figura 3B), na segunda tentativa com TRI Reagent (Figura 3C) passou a exibir o padrão esperado.

Já a eficiência e integridade das bibliotecas de cDNA construídas foi atestada com o sistema TapeStation 4200, onde para a maior parte das bibliotecas as análises apontaram tamanhos entre 200 e 400 pb (medidas consideradas satisfatórias).

4.2 Sequenciamento, montagem e anotação

O sequenciamento resultou em uma série de arquivos no formato FASTQ contendo as *reads* e os parâmetros de qualidade para as mesmas, que quando concatenados representam o rendimento bruto por amostra. A quantidade total de *reads* geradas foi de 1.172.433.138, com uma média de 48.851.381 *reads* por amostra, ou 97.702.762 se considerarmos *paired-end*.

Parâmetros gerais como pontuação de qualidade por sequência, conteúdo de guanina-citosina por sequência, conteúdo de nitrogênio por sequência e distribuição de comprimento de sequência foram avaliados via FastQC, com os resultados individuais de cada amostra sendo agrupados via MultiQC (Figura 4).

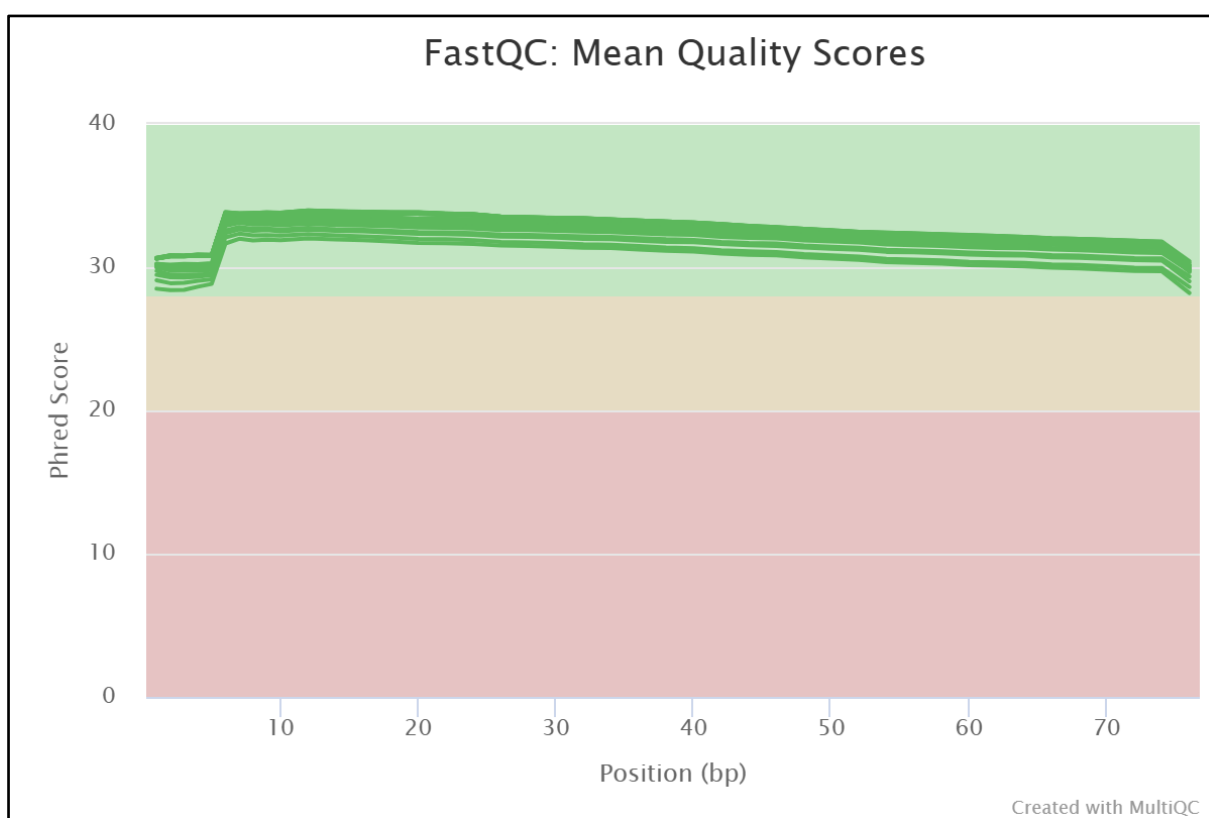


Figura 4 – Representação do valor médio de qualidade em cada posição de base do sequenciamento de RNA mensageiro de *M. spectabilis*. Valores de Phred Score foram obtidos pelo FastQC e agregado no programa MultiQC. Linhas verdes representam diferentes sequências de diferentes amostras. Regiões de fundo verde, amarelo e vermelho representam Phred Score bom, médio e ruim, respectivamente. Fonte: elaborado pelo autor.

De maneira geral a qualidade reportada pelo FastQC foi considerada boa na grande maioria dos parâmetros, para a maioria das amostras, baseado em referências do software (ANDREWS, 2010).

Na filtragem com os programas Trimmomatic e SortMeRNA, diferentes quantidades de reads foram retiradas de cada amostra, resultando em distintas porcentagens restantes ao final (Tabela 3).

Tabela 3 - Quantidade média de *reads* do sequenciamento de *M. spectabilis* mantidas durante as etapas de filtragem com Trimmomatic e SortMeRNA. Fonte: elaborado pelo autor.

Grupo de amostras	Pré filtro	Pós Trimmomatic	Pós SortMeRNA	% Mantida
OVO	43653377	38788294	37544170	85,79
NP	48429580	43346873	42633463	88,12
NG	55172794	49866236	49326517	89,44
AD	48170772	43409286	42944490	89,16
Total	48856631	43852672	43112160	88,13

Após filtragem com Trimmomatic os arquivos apresentavam uma média de 43.852.672 *reads* por amostra (média de 10,31% de seqüências descartadas) e quantidade total de 1.052.464.134 *reads* considerando R1+R2. Destas, 17.772.294 (cerca de 1,68%) foram descartadas com SortMeRNA.

Um total de 1.034.691.840 *reads* foi obtido ao final (média de 43.112.160 por amostra), significando uma desqualificação de 11,76% do total obtido por sequenciamento. Para garantir a qualidade do transcriptoma final na abordagem *de novo*, para faz-se necessário que haja um número suficiente de *reads* “livres de erro” que cubram todo o genoma, e neste contexto a quantidade obtida foi considerada satisfatória.

O resultado do transcriptoma obtido pelo Trinity foi avaliado através do arquivo ‘TrinityStats.pl’ (Tabela 4), onde temos representado o total de “genes”, de transcritos, a porcentagem de bases guanina-citosina, e as características das *contigs*.

Tabela 4 - Parâmetros do transcriptoma de *M. spectabilis* montado com o Trinity. Fonte: elaborado pelo autor.

Estatísticas total Trinity	
Total de 'genes' Trinity	309002
Total de transcritos	434018
Percentual de GC	35,67
Contigs	
Contig N10	4405
Contig N20	2834
Contig N30	2004
Contig N40	1418
Contig N50	980
Contigs Médias	
Comprimento médio de contigs	338
Média de contigs	634
Total de bases montadas	275290569

Com um total de 309.002 *contigs*, este transcriptoma foi submetido ao CD-HIT para eliminar quaisquer níveis de redundância presentes (Tabela 5).

Tabela 5 - Parâmetros do transcriptoma de *M. spectabilis* após filtragem com CD-HIT. Fonte: elaborado pelo autor.

Estatísticas cd-hit longest	
Total de genes	197003
Total de transcritos	197003
Percentual de GC	34,48
Contigs	
Contig N10	3800
Contig N20	2468
Contig N30	1721
Contig N40	1206
Contig N50	853
Contigs Médias	
Comprimento médio de contigs	364
Média de contigs	622
Total de bases montadas	122563784

Um dos principais parâmetros para avaliação da qualidade do transcriptoma montado é o N50. Esta é uma métrica simples que descreve o comprimento da sequência em que metade dos nucleotídeos na montagem do transcriptoma estão em sequências iguais ou maiores que N. No caso, 853 pb é considerado um valor

aceitável para montagens *de novo*, já que para uma montagem eucariótica típica a bibliografia sugere que com valores N50 da ordem de 50 Kpb produzem-se conjuntos de genes bastante completos (SIMÃO et al., 2015). Obteve-se um total de 197.003 *contigs*, quantidade satisfatória para os fins de análise de expressão e prospecção de genes.

Embora a quantidade exata de genes presentes em um inseto varie muito conforme a espécie, estima-se uma média em torno de 10 mil a 20 mil genes para qualquer genoma realista (MEI et al., 2022). Sendo assim, 197 mil *contigs* aparenta estar muito fora do limite plausível de genes esperados para um genoma de *M. spectabilis*. Porém esta discrepância pode ser explicada por artefatos inerentes ao processo de montagem *de novo* com Trinity, como ambiguidades na leitura das *reads*, erros de sobreposição, isoformas de *splicing*, erros na montagem de regiões muito repetitivas e transcritos codificados por *loci* adjacentes que se sobrepõem e formam as chamadas transcrições quiméricas (GRABHERR et al., 2011).

Outras formas utilizadas para assegurar a qualidade da montagem do transcriptoma foram através do BUSCO v3 (Tabela 6) e do Bowtie2 (Tabela 7).

Tabela 6 - Resultados de análise com BUSCO v3. a) Sequências comparadas com conjunto de genes ortólogos conservados em Hemiptera; b) Sequências comparadas com conjunto de genes ortólogos conservados em Eucariotos. Fonte: elaborado pelo autor.

a) Estatísticas BUSCO (hemiptera_odb10)		b) Estatísticas BUSCO (eukaryota_odb10)	
BUSCOs completos	2423	BUSCOs completos	250
Completos e cópia única	2291	Completos e cópia única	219
Completos e duplicado	132	Completos e duplicado	31
Fragmentado	27	Fragmentado	2
Faltando	60	Faltando	3
Total de grupos pesquisados	2510	Total de grupos pesquisados	255

Tabela 7 – Resultados de análise com Bowtie2. Fonte: elaborado pelo autor.

Estatísticas Bowtie2
517345920 reads; sendo destas:
517345920 (100.00%) pareadas; sendo destas:
85701514 (16.57%) alinhadas concordantemente 0 vezes
325533933 (62.92%) alinhado concordantemente exatamente 1 vez
106110473 (20.51%) alinhado concordantemente >1 vez
85701514 pares alinhados concordantemente 0 vezes; sendo destes:

11571637 (13.50%) alinhados discordantemente 1 vez
74129877 pares alinhados 0 vezes concordante ou discordantemente; sendo destes:
148259754 associados formam os pares; destes:
97345316 (65.66%) alinhados 0 vezes
27535249 (18.57%) alinhados exatamente 1 vez
23379189 (15.77%) alinhados >1 vez
Taxa de alinhamento geral de 90,59%

A partir das estatísticas do BUSCO concluímos, acerca dos transcritos presentes no transcriptoma, que: 96,53% destes foram encontrados em ortólogos de Hemiptera, sendo em sua maioria completos e em cópia única; 98% destes foram encontrados em ortólogos de Eukaryota, sendo em sua maioria completos e em cópia única. Isso indica que, em termos de conteúdo, o transcriptoma obtido é significativamente representativo do domínio e ordem ao qual a cigarrinha pertence, atestando a confiabilidade e qualidade do mesmo.

As estatísticas do Bowtie2 acrescentam que, quando confrontadas com o transcriptoma final, o total de *reads* filtradas apresentou uma taxa de alinhamento geral de 90,59%. Consideramos, portanto, que a montagem foi eficiente em incorporar uma parcela suficiente dos transcritos fornecidos, atestando a cobertura e amplitude do transcriptoma em relação ao genoma.

O transcriptoma filtrado para redundância foi anotado a partir do banco de dados UniProt e obteve-se metadados acerca de 25.492 (12,94%) *contigs* com BLASTX (em termos de transcrito) e 18.302 (9,29%) *contigs* com BLASTP (em termos de proteína). Informações como ontologia gênica, proteína a que se referem e identificação foram obtidas para estes, sendo esta última uma identificação própria da SwissProt.

Através deste processo de anotação dados relevantes foram adicionados e permitiram a investigação, em termos de proteína e função, dos diferentes perfis de expressão em cada fase, diferenciando-se genes cuja atuação esteja restrita à um intervalo, ou aqueles que estão sendo acionados de forma constante. Esta porcentagem foi consideravelmente menor do que as obtidas em transcriptomas de *Piezodorus guildinii* (SCHVARTZMAN et al., 2022) e *Euschistus heros* (CAGLIARI et al., 2020), o que pode ser explicado pelo menor número de *contigs* investigadas

nestes casos e pelo fato da anotação ter sido realizada a partir de uma única base de dados.

4.3 Maquinaria de RNAi

Com o levantamento em bibliografia organizamos cinco principais categorias de genes que codificam: proteínas envolvidas na clivagem de dsRNA (família Dicer); proteínas com atividade de endonuclease dirigida contra a fita de mRNA (Argonautas); proteínas de ligação de dsRNA; proteínas de transporte e captação de dsRNA e proteínas auxiliares associadas ao RISC.

Dentro dessas categorias, selecionamos 30 genes para serem procurados, e muitos deles apresentaram correspondência no transcriptoma anotado. Os padrões encontrados variam entre expresso de maneira uniforme em todas as amostras e diferencialmente expresso entre amostras.

De maneira geral, para cada uma das categorias pelo menos três principais genes foram identificados no transcriptoma anotado, e parte destas sequências foi selecionada para exibição (Tabela 8).

Tabela 8 - Genes da maquinaria de RNAi identificados no transcriptoma montado de *M. spectabilis*, categorizados por grupo de função. Fonte: elaborado pelo autor.

Categoria	Gene	Anotação	Descrição da proteína	Média de contagens de reads			
				Ovo	Ninfa pequena	Ninfa grande	Adulto
Clivagem de dsRNA	Dicer-1	DCR1_DROME	Endoribonuclease Dcr-1	37,03	37,04	197,61	150,91
	Dicer-2	DCR2_DROME	Endoribonuclease Dcr-2	965,13	1070,85	2013,66	1910,43
	Drosha	RNC_MOUSE	Ribonuclease 3	2450,08	1083,24	1472,26	1405,42
Endonucleases	Ago1	AGO1_ARATH	Protein argonaute 1	0,00	1,59	1,33	1,33
	Ago2	AGO2_RABIT	Protein argonaute 2	412,38	709,93	2935,77	2763,16
	Ago3	AGO3_BOMMO	Protein argonaute 3	367,72	278,89	768,85	612,69
	Piwi	SIWI_BOMMO	Piwi-like protein Siwi	12853,89	3889,71	5635,65	2521,39
Ligação de dsRNA	Spindle	SPNE_AEDAE	ATP-dependent RNA helicase spindle-E	2535,59	1182,45	1346,11	1391,02
	Loqs	LOQS_DROME	Protein Loquacious	750,00	767,82	1286,99	720,54
	Pasha	DGCR8_MOUSE	Microprocessor complex subunit DGCR8	506,67	318,37	432,67	286,99
Transporte e captação de dsRNA	SID	SIDT2_HUMAN	SID1 transmembrane family member 2	191,88	207,11	646,45	808,42
	Chc	CLH_DROME	Clathrin heavy chain	5354,55	8629,08	15026,96	12067,93
	Innexin	INX2_SCHAM	Innexin inx2	1229,87	1610,88	7112,21	4060,07
	Epsin	EPN2_RAT	Epsin-2	72,05	229,02	324,26	371,51

	Gawky	GAWKY_DROME	Protein Gawky	606,92	758,81	2607,08	2538,32
Fatores auxiliares associados ao RISC	Armi	ARMI_DROME	Probable RNA helicase armi	1683,45	1132,11	1614,36	1430,49
	Staufen	STAU2_DANRE	Double-stranded RNA-binding protein Staufen homolog 2	4719,21	2230,90	3133,68	2039,89

Identificamos a anotação de mais de 100 *contigs* correspondentes a 30 diferentes genes envolvidos nas principais funções do mecanismo. A expressão da maioria destes varia entre média e alta, indicando que o sistema está conservado no genoma de *M. spectabilis* e sendo transcrito em mRNAs. Esta evidência aponta um sinal positivo de que, em resposta a dsRNAs livres, o organismo da *M. spectabilis* desencadearia as reações necessárias que resultem no silenciamento de um gene de interesse. A conservação deste mecanismo na ordem Hemiptera tem sido atestada (CAGLIARI et al., 2020; SINGH et al., 2020, SCHVARTZMAN et al., 2022), e casos de sucesso no uso de RNAi para aumento na taxa de mortalidade foram reportados em alguns membros (YANG et al., 2022; SCHVARTZMAN et al., 2022), sugerindo que esta pode ser uma estratégia viável em *M. spectabilis*.

4.4 Transcritos diferencialmente expressos

Com as análises de expressão diferencial, 10.928 *contigs* presentes no transcriptoma foram caracterizados como DEGs. Estes se distribuem de forma distinta nas seis comparações entre as amostras (Figura 5), sendo estas ovo contra ninfa pequena (OVvsNP), ovo contra ninfa grande (OVvsNG), ovo contra adulto (OVvsAD), ninfa pequena contra ninfa grande (NPvsNG), ninfa pequena contra adulto (NPvsAD) e ninfa grande contra adulto (NGvsAD).

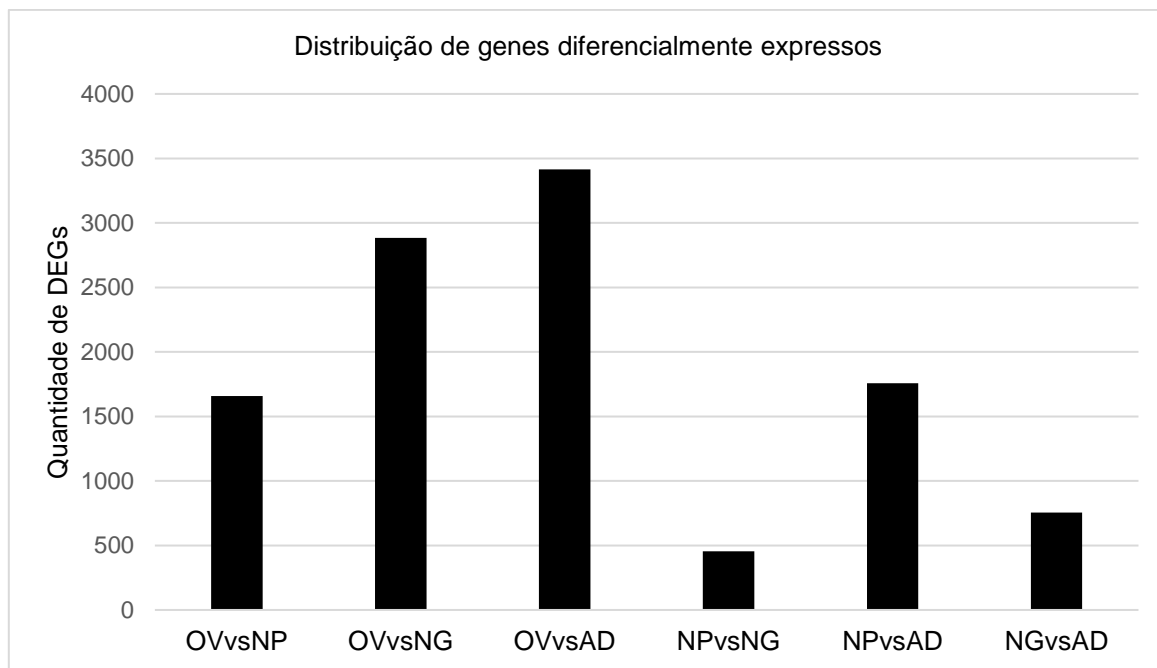


Figura 5 – Representação da distribuição da quantidade de transcritos diferencialmente expressos do transcriptoma de *M. spectabilis* entre as seis comparações de amostras. OVvsNP: comparação entre ovo e ninfa pequena; OVvsNG: comparação entre ovo e ninfa grande; OVvsAD: comparação entre ovo e adulto; NPvsNG: comparação entre ninfa pequena e ninfa grande; NPvsAD: comparação entre ninfa pequena e adulto; NGvsAD: comparação entre ninfa grande e adulto. Fonte: elaborado pelo autor.

As fases mais distintas (com maior número de DEGs) são, respectivamente, ovos e adultos, ovos e ninfa grandes e ovos e ninfas pequenas. Pode-se deduzir que esta diferença observada está vinculada ao alto número de diferentes genes necessários para o desenvolvimento pré-eclosão (destaca-se que 72,8% dos DEGs estão presentes nas comparações envolvendo ovos e as demais bibliotecas). Já as bibliotecas mais semelhantes (com menor número de DEGs) seriam de ninfa pequena e ninfa grande, e ninfa grande e adultos, indicando que são poucas as diferenças em termos de transcritos entre os últimos ínstares da hemimetabolia. Estas constatações *in silico* correspondem com as observações *in vivo* considerando que as cigarrinhas-das-pastagens se desenvolvem por hemimetabolia, ou seja, uma metamorfose incompleta (VALÉRIO 2009). Sendo assim, durante os estágios de ninfa estas passam por mudanças graduais em que a expressão gênica tende a mudar de forma mais linear e progressiva, quando comparado ao desenvolvimento de holometábolos.

Um gráfico de escala multidimensional das amostras foi produzido para representar as principais mudanças de logFC (consequentemente as diferenças em conteúdo de expressão) entre as bibliotecas (Figura 6).

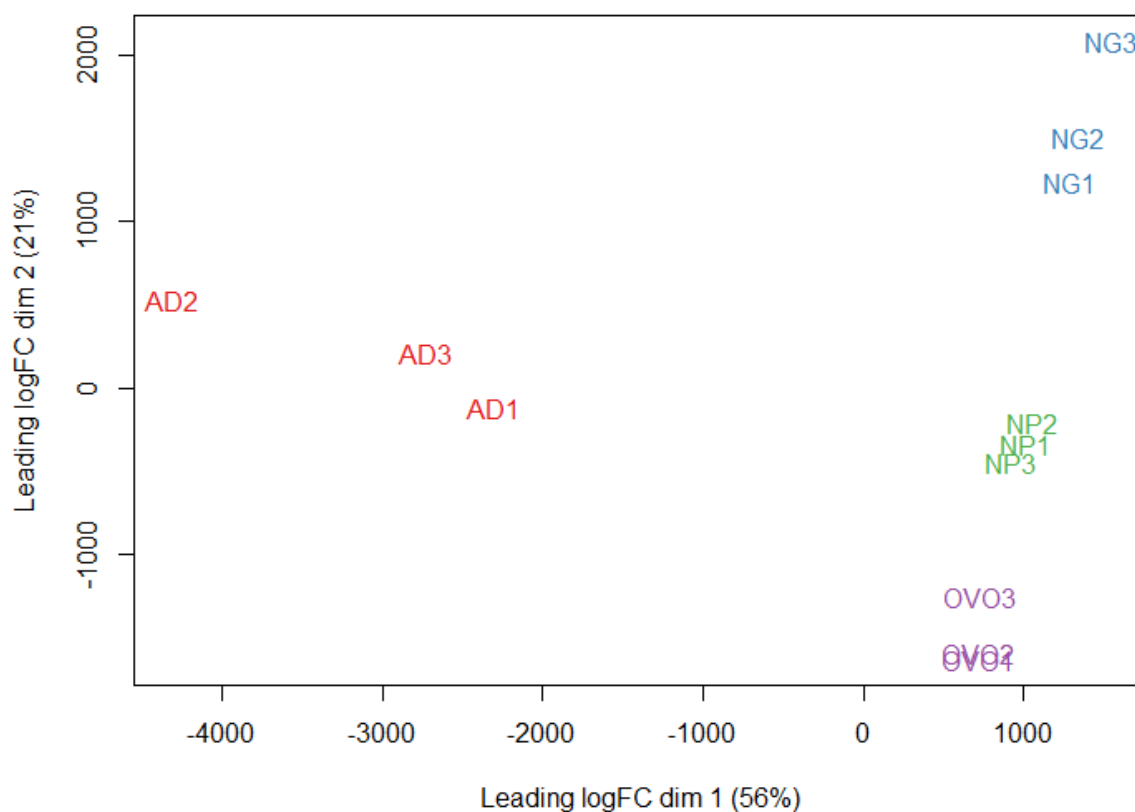


Figura 6 – Representação das mudanças de logFC entre comparações de amostras de *M. spectabilis* sequenciadas. Cada amostra é representada pela sua sigla e agrupada com suas réplicas pela cor da fonte. As distâncias entre cada sigla correspondem às principais mudanças de logFC, ou às relativas similaridades entre cada amostra. AD1, AD2 e AD3 referem-se a adultos; NP1, NP2 e NP3 referem-se a ninfas pequenas; NG1, NG2 e NG3 referem-se a ninfas grandes; OVO1, OVO2 e OVO3 referem-se a ovos. Fonte: elaborado pelo autor.

As proximidades observadas entre as triplicatas de cada grupo corroboram a esperada semelhança das mesmas em termos de conteúdo de *contigs*. Além disso, o fato de AD1, AD2 e AD3 estarem consideravelmente distantes das demais aponta como a fase adulta apresenta transcritos altamente diferentemente expressos das anteriores, principalmente de ovos (3.416 DEGs nesta comparação). Sendo assim temos a confirmação da hipótese de que os insetos em fase adulta seriam os mais diferenciados em termos de expressão. Esta dedução advém do fato de que, após o quinto e último instar ninfal, surgem mudanças morfológicas (como a disposição de asas completamente desenvolvidas, a deposição de pigmentos de coloração),

fisiológicas (como a alimentação na parte aérea da planta, um metabolismo adaptado ao voo) e comportamentais (adultos são altamente móveis e se dispersam, acasalam e ovipositam)(VALÉRIO 2009) que afetam enormemente toda a dinâmica de proteínas necessárias nas células.

4.5 Validação por RT-qPCR

Ao todo desenhamos pares de *primers* para 16 genes (Tabela 9), com destaque para API e HPI que propositalmente foram desenhados na ponta de um éxon com regiões de *introns* (não-codificante, portanto, não presente em RNA total) no interior do *amplicon*, visando-se avaliar possíveis contaminações de DNA nas amostras.

Tabela 9 – Genes utilizados na validação de expressão por RT-qPCR em *M. spectabilis*. *Primers forward* e *reverse* referem-se ao sentido de referência em que os *primers* serão anelados ao molde. *Melting* refere-se a etapa na qual metade dos fragmentos de molde estão na forma desnaturada (temperatura de desnaturação). Fonte: elaborado pelo autor.

Genes utilizados no RT-qPCR						
Abreviação do gene	Eficiência	Fita forward (5'-3')	Temperatura de <i>melting</i>	Fita reverse (5'-3')	Temperatura de <i>melting</i>	Tamanho do Amplicon (pb)
CU	95,30%	ATAACCTCGCTCCATTCTGTG	54,6 °C	TACCCAGGAGATAACGGTTGA	55 °C	133
SS	104,20%	CGAGTGTAGCAAGCTGATTTG	54,1 °C	CACTACAAATGGAGCAACTGGT	55,3 °C	91
SP	104%	ACAGGAAGAGGGAAATGTGATG	54,7 °C	GTTGAATACAGTGCGGAAAGAC	54,1 °C	192
HPI	102,50%	GCTGGATGGGGGATTACTA	53,5 °C	CGCAGTATGATTCCACACAG	53,5 °C	87
BE	-	AAAGGGCATTGGAGGAAGTA	-	AATGTGGGAGCAAGTCTTCA	-	125
PG	0,953	CCCGCATTATGAGATTCTG	51,1 °C	TGCATCATTGGCAACTACAC	53,5 °C	77
NA	1,01	GGCTTGCTTTGTTGCTTCAT	54,6 °C	TCCAGCGACCAGTGAGTATT	55,7 °C	142
AG	0,962	CCTTCAGTTGATGTGGCAGT	55,3 °C	TGAAAGTCTTCCACCTGCTC	54,7 °C	87
PI	0,973	CGGGTGGTTGAATACTGAATCT	54,7 °C	TACGCTCTAATCCCATGTTTCA	54,4 °C	144
SI	-	AGTTGCACCAGTTCGAGAAA	-	TCCAGCTCCAACAACAACA	-	83
SC	-	AAAGAACTGTAGCGGAGAACG	-	TCAAGTCAGTCGGTTCTGTTC	-	73
GA	0,95	AGTGGCAGTCAGTCTCAAGG	56,7 °C	GTACCTGTAGCTGGCGTGTT	57,2 °C	131
18	0,968	GTTGATTGTTCTTGGGCTTG	52,3 °C	GCCTCAACACATGACAACCTG	54,2 °C	137
28	0,973	CAGCACCTGGTATGAAGACAAG	55,6 °C	CTGGACACATATCAGGACCTATC	54,2 °C	183
FU	-	CCTCTGTCCAAAGAATACCGAA	-	GTTGGTTGAACGGGATTAACCTG	-	73
API	-	GATGTAAAAGCAGATAAAAATTCAA	-	AAGAAGCCTCGTGCATCTAA	-	206

Baseado nos resultados descritos pelo Apêndice A, concluímos que a expressão diferencial estimada *in silico* no transcriptoma verdadeiramente corresponde àquela presente nas amostras *in vivo*, sendo considerada válida para o estudo e análises das diferenças entre os estágios de desenvolvimento do inseto.

4.6 Potenciais alvos

Foram selecionados, a partir de aprofundamento em bibliografia e consulta ao transcriptoma, cinco genes candidatos para silenciamento com perfis de expressão adequados (Tabela 10).

Tabela 10 – Genes selecionados como alvo para silenciamento por RNA a partir da literatura e do transcriptoma de *M. spectabilis*, com anotação e perfil de expressão destes. Fonte: elaborado pelo autor.

Genes candidatos a alvo para silenciamento em <i>M. spectabilis</i>						
Gene	Referência	Anotação	Média de contagens de reads			
			Ovo	Ninfa pequena	Ninfa grande	Adulto
V-ATPase A	SCHVARTZMAN et al., 2022	VATA_AEDAE	4274,32	21617,69	19089,06	16020,64
Trealase	HUNTER et al., 2021	TREA_APIME	740,94	4064,91	5578,41	11312,62
Calmodulina	WANG et al., 2018	CALM_STRIE	23301,63	23174,75	17722,19	14004,54
Apolipoproteína D	ZHANG et al., 2024	APOD_HUMAN	2799,88	3013,05	4788,95	11856,16
Hexoquinase	YANG et al., 2022	HXK2_DROME	2827,08	13963,01	12812,40	26155,12

4.6.1 V-ATPase

A ATPase de bombeamento de prótons do tipo vacuolar (V-ATPase) é uma enzima muito conservada evolutivamente, responsável pelo transporte de prótons (H⁺) através das membranas e pela acidificação dos compartimentos celulares (FUTAI et al., 2019). A subunidade A desta proteína é um alvo conhecido de silenciamento em insetos, onde demonstrou causar mortalidade em diferentes ordens (SCHVARTZMAN et al., 2022).

4.6.2 Trealase

A trealose é o principal açúcar circulante no sangue dos insetos, e a enzima trealase está envolvida no metabolismo energético e regulação dos níveis de trealose

nas células (HUNTER et al., 2021). A trealase desempenha papéis fundamentais no desenvolvimento e sobrevivência dos hemípteros, e com a supressão da mesma a síntese de trealose é reduzida gerando um aumento na mortalidade de psílídeos, fulgoromorpha e outros invertebrados (HUNTER et al., 2021).

4.6.3 Calmodulina

Calmodulina é uma proteína ligante de cálcio importante em diversas funções celulares, pois o complexo Ca^{2+} /calmodulina interage e regula várias enzimas e proteínas-alvo, conhecidas como proteínas ligantes de calmodulina (WANG et al., 2018). Esta proteína é responsável por mudanças celulares de Ca^{2+} e regula eventos de sinalização subsequentes, estando envolvida em uma variedade de atividades fisiológicas (WANG et al., 2018).

A calmodulina apresenta particular relevância durante certas fases de desenvolvimento de insetos, e o seu silenciamento afeta o crescimento, a motilidade, e a vitelogênese (WANG et al., 2018).

4.6.4 Apolipoproteína D

A espuma produzida pelas cigarrinhas contém vários carboidratos e lipídios que são essenciais para manter a estabilidade, e o transporte de lipídios depende da apolipoproteína D (ZHANG et al., 2024). Sendo assim, é esperado que o silenciamento de proteínas que se ligam à lipídios e os transportam no organismo ocasione aumento da mortalidade de cigarrinhas ainda na fase de ninfa, onde a espuma produzida é de suma importância para proteção.

4.6.5 Hexoquinase tipo 2

Hexoquinase é uma enzima chave na rota de biossíntese de quitina em insetos e desempenha um papel importante no desenvolvimento e regulação de energia, além de influenciar na síntese de glicose-6-fosfato (YANG et al., 2022). Aparentemente o silenciamento da hexoquinase pode inibir a formação de quitina durante as fases iniciais de crescimento e conseqüentemente levar a anormalidades nos insetos, crescimento atrofiado e pupação tardia (YANG et al., 2022).

Com base no aprofundamento na literatura foram encontrados onze genes envolvidos em funções essenciais e comumente utilizados como alvo (proteína de classificação vacuolar SNF7, V-ATPase A, proteína de reconhecimento de peptidoglicano, lacase 1, trealase, calmodulina, apolipoproteína D, lisozima 2, hexoquinase 2, syntaxina 1A e fator de splicing Cactin) e, após investigação do perfil de expressão em *M. spectabilis*, foram escolhidos V-ATPase A, trealase, calmodulina, apolipoproteína D e hexoquinase tipo 2 (Figura 7).

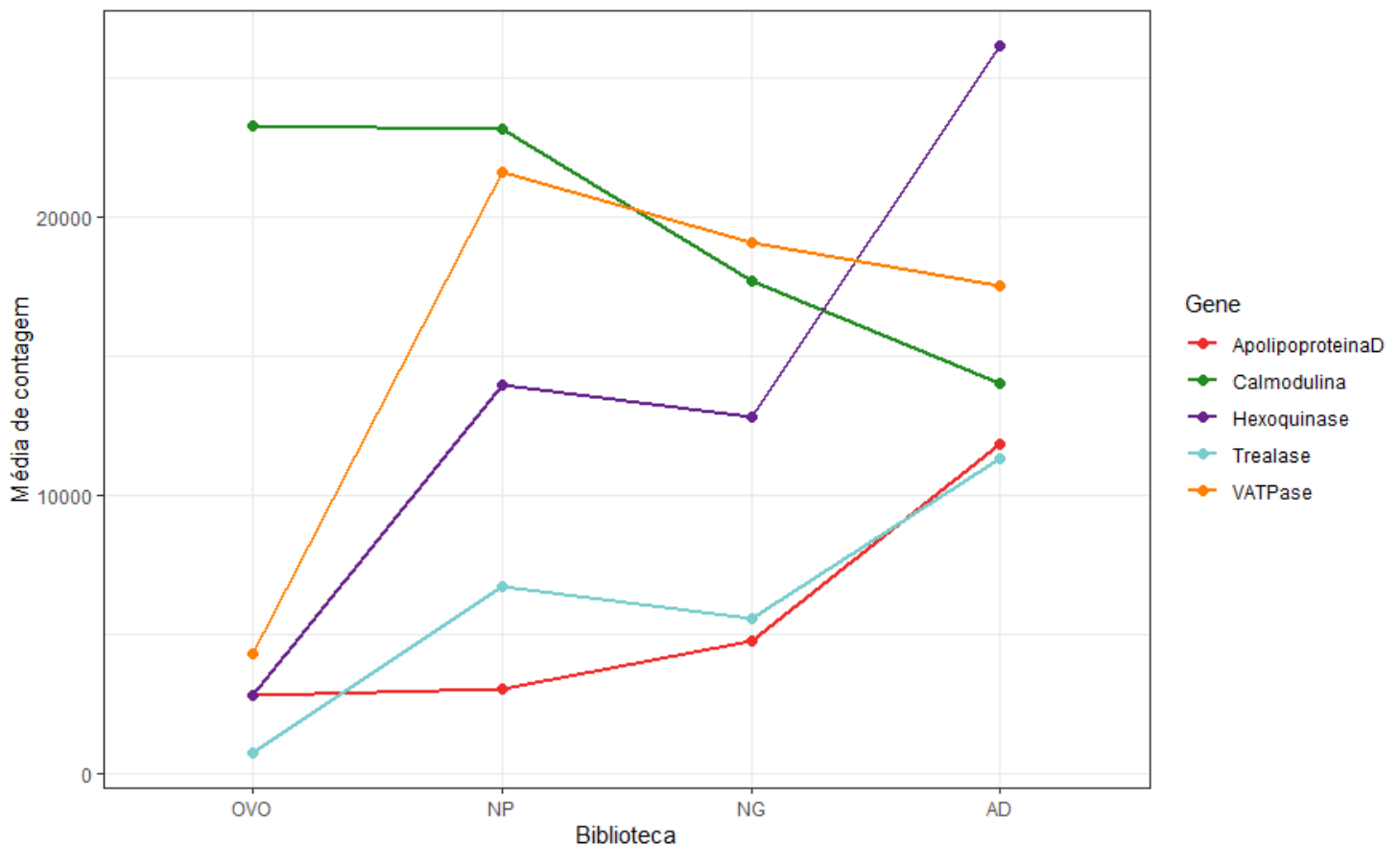


Figura 7 – Comparação entre contagem média por biblioteca de cDNA dos cinco genes candidatos a alvo no controle de *M. spectabilis*. OVO: Ovos; NP: Ninfas pequenas; NG: Ninfas grandes; AD: Adultos. Fonte: elaborado pelo autor.

A escolha optou, tanto quanto possível, por transcritos relacionados a diferentes espécies de cigarrinhas-das-pastagens, aos hemípteros e em última instância aos insetos. Além disso, uma alta expressão também foi fator determinante, garantindo que os escolhidos fossem relevantes o suficiente para que, quando suprimidos, afetem de maneira satisfatória a sobrevivência. Dos cinco escolhidos quatro apresentam expressão crescente conforme desenvolvimento, enquanto apenas

calmodulina decresce entre ovos e adultos (apesar disso ainda se mantem altamente expresso após quinto instar ninfal).

Em muitos casos, a V-ATPase tem sido o alvo majoritário para testes de eficiência de RNAi em Hemiptera (BASNET et al., 2018; SCHVARTZMAN et al., 2022), e outros insetos (CHANG et al., 2021; ZENG et al., 2021), porém o aprofundamento nos genes diferencialmente expressos em fases distintas de desenvolvimento é uma estratégia capaz de prover alvos mais específicos, ecologicamente seguros e possivelmente mais eficazes (NIU et al., 2024). Recentemente um novo inseticida baseado na tecnologia de RNAi se tornou destaque por ser o primeiro a receber aprovação da Agência de Proteção Ambiental dos Estados Unidos (RODRIGUES et al., 2021), após uma etapa de revisão e certificação de quatro anos. Neste caso, estudos prévios já elucidavam a conservação e o funcionamento da maquinaria de RNAi na espécie, além de terem sido testados diversos genes essenciais que demonstravam o potencial do silenciamento (RODRIGUES et al., 2021). O alvo ideal que propiciou o escalonamento da fase experimental para o desenvolvimento e comercialização de um produto foi a subunidade beta 5 do proteassoma (dsPSMB5), que quando silenciado via dsRNA aplicado de forma tópica por pulverização foi eficaz em causar mortalidade no besouro da batata do Colorado, *Leptinotarsa decemlineata*.

Assim, o presente trabalho foi bem-sucedido em estabelecer um transcriptoma de diferentes estágios de desenvolvimento de *M. spectabilis* e, através dele, atestar a presença dos genes da maquinaria de RNAi na espécie, prospectar cinco possíveis genes vitais (V-ATPase A, trealase, calmodulina, apolipoproteína D e hexoquinase tipo 2) que possam ser alvo de silenciamento e determinar o perfil de expressão destes considerando as diferentes fases de desenvolvimento.

5 PERSPECTIVAS

Como perspectivas para a continuidade do que foi feito, planeja-se: síntese de dsRNAs e uso destes em ensaios de resistência para verificação da eficiência dos alvos na mortalidade de insetos; elaboração de análises de redes com *network hubs* para estabelecer novos genes alvo ainda não explorados pela literatura; uso de ferramentas metabolômicas para investigação dos compostos específicos da espuma produzida pelas ninfas, sendo estes possíveis alvos quando identificados no transcriptoma; análises *in silico* de predição de possíveis alvos não desejados, para

garantir a segurança ecológica dos dsRNAs caso fossem utilizados em um produto final.

6 REFERÊNCIAS

ALBERTS, Bruce et al. **Biologia molecular da célula**. 6. ed. Porto Alegre: Artmed, 2017.

ANDREWS, Simon. **FastQC: a quality control tool for high throughput sequence data**. 2010. Disponível em: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

AUAD, Alexandre M. et al. **Spittle protein profile of *Mahanarva spectabilis* (Hemiptera: Cercopidae) fed various elephant grass genotypes**. 2012.

BASNET, Sanjay et al. **RNAi-mediated knockdown of vATPase subunits affects survival and reproduction of bed bugs (Hemiptera: Cimicidae)**. Journal of medical entomology, v. 55, n. 3, p. 540-546, 2018.

BAUM, James A. et al. **Control of coleopteran insect pests through RNA interference**. Nature biotechnology, v. 25, n. 11, p. 1322-1326, 2007.

BOLGER, Anthony M. et al. **Trimmomatic: a flexible trimmer for Illumina sequence data**. Bioinformatics, v. 30, n. 15, p. 2114-2120, 2014.

BRIAN, Matlock. **Assessment of Nucleic Acid Purity**. Wilmington, MA, USA: Thermo Fisher Scientific, 2015.

BUCHFINK, Benjamin et al. **Sensitive protein alignments at tree-of-life scale using DIAMOND**. Nature methods, v. 18, n. 4, p. 366-368, 2021.

CAGLIARI, Deise et al. **First transcriptome of the Neotropical pest *Euschistus heros* (Hemiptera: Pentatomidae) with dissection of its siRNA machinery**. Scientific reports, v. 10, n. 1, p. 4856, 2020.

CAMARGO, Roberto de A. et al. **De novo transcriptome assembly and analysis to identify potential gene targets for RNAi-mediated control of the tomato leafminer (*Tuta absoluta*)**. BMC genomics, v. 16, p. 1-17, 2015.

CAMARGO, Roberto de A. et al. **RNA interference as a gene silencing tool to control *Tuta absoluta* in tomato (*Solanum lycopersicum*)**. PeerJ. 2016.

CARDONA, Cesar et al. **Antibiosis and tolerance to five species of spittlebug (Homoptera: Cercopidae) in *Brachiaria spp.*: implications for breeding for resistance**. Journal of economic entomology, v. 97, n. 2, p. 635-645, 2004.

CHANG, Ya-Wen et al. **RNA interference of genes encoding the vacuolar-ATPase in *Liriomyza trifolii***. Insects, v. 12, n. 1, p. 41, 2021.

CHEN, Yunshun et al. **edgeR 4.0: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets**. bioRxiv, p. 2024.01. 21.576131, 2024.

DALAKOURAS, Athanasios et al. **DsRNA-based pesticides: Considerations for efficiency and risk assessment**. Chemosphere, p. 141530, 2024.

ESPITIA BUITRAGO, Paula A. et al. **Spittlebugs (Hemiptera: Cercopidae): Integrated pest management on gramineous crops in the Neotropical ecozone**. Frontiers in Sustainable Food Systems, v. 6, p. 891417, 2022.

EWELS, Philip et al. **MultiQC: summarize analysis results for multiple tools and samples in a single report**. Bioinformatics 32(19):3047–3048, 2016.

FU, Limin et al. **CD-HIT: accelerated for clustering the next-generation sequencing data**. Bioinformatics, v. 28, n. 23, p. 3150-3152, 2012.

FUTAI, Masamitsu et al. **Vacuolar-type ATPase: A proton pump to lysosomal trafficking**. Proceedings of the Japan Academy, Series B, v. 95, n. 6, p. 261-277, 2019.

GRABHERR, Manfred G. et al. **Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data**. Nature biotechnology, v. 29, n. 7, p. 644, 2011.

GUPTA, Mridula et al. **Transcriptome analysis unravels RNAi pathways genes and putative expansion of CYP450 gene family in cotton leafhopper *Amrasca biguttula* (Ishida)**. Molecular Biology Reports, v. 48, n. 5, p. 4383-4396, 2021.

HAAS, Brian J. et al. **De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis.** Nature protocols, v. 8, n. 8, p. 1494-1512, 2013.

HUNTER, Wayne B. et al. **Optimizing efficient RNAi-mediated control of hemipteran pests (psyllids, leafhoppers, whitefly): modified pyrimidines in dsRNA triggers.** Plants, v. 10, n. 9, p. 1782, 2021.

IWANICKI, Natasha S. A. et al. **Monitoring of the field application of *Metarhizium anisopliae* in Brazil revealed high molecular diversity of *Metarhizium* spp in insects, soil and sugarcane roots.** Scientific reports, v. 9, n. 1, p. 4443, 2019.

JAIN, Ritesh G. et al. **RNAi-based functional genomics in Hemiptera.** Insects, v. 11, n. 9, p. 557, 2020.

KIM, Daniel H. et al. **RNAi mechanisms and applications.** Biotechniques, v. 44, n. 5, p. 613-616, 2008.

KNUTSEN, Morgan C. **RNAi-Mediated Gene Silencing in the Exotic Redbay Ambrosia Beetle, *Xyleborus glabratus*, and Insect-Fungal Interactions within the Laurel Wilt Complex.** 2023.

KOPYLOVA, Evguenia et al. **SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.** Bioinformatics, v. 28, n. 24, p. 3211-3217, 2012.

LANGMEAD, Ben; SALZBERG, Steven L. **Fast gapped-read alignment with Bowtie 2.** Nature methods, v. 9, n. 4, p. 357-359, 2012.

MAMTA, B. et al. **RNAi technology: a new platform for crop pest control.** Physiol Mol Biol Plants 23, 487–501, 2017.

MEI, Yang et al. **InsectBase 2.0: a comprehensive gene resource for insects.** Nucleic acids research, v. 50, n. D1, p. D1040-D1045, 2022.

MEZZETTI, Bruno et al. **RNAi: What is its position in agriculture?** Journal of Pest Science, v. 93, n. 4, p. 1125-1130, 2020.

MUHAMMAD, Nor Azlan Nor et al. **Data on RNA-seq analysis of the cocoa pod borer pest *Conopomorpha cramerella* (Snellen) (Lepidoptera: Gracillariidae).** Data in Brief, v. 34, p. 106638, 2021.

NIU, Jinzhi et al. **RNA interference in insects: the link between antiviral defense and pest control.** Insect science, v. 31, n. 1, p. 2-12, 2024.

PAULA-MORAES, S. V. de et al. **Variabilidade genética de *Mahanarva spectabilis* (Distant, 1909) coletada em *Brachiaria brizantha* cv. Marandu no estado de Tocantins, com base em marcadores RAPD.** Planaltina, DF: Embrapa Cerrados, 2006.

PARISE, Cecilia et al. **RNA interference protocols for gene silencing in the spittlebug *Philaenus spumarius*, vector of *Xylella fastidiosa*.** Scientific Reports, v. 14, n. 1, p. 25812, 2024.

PATRO, Rob et al. **Salmon provides fast and bias-aware quantification of transcript expression.** Nature methods, v. 14, n. 4, p. 417-419, 2017.

RAGHAVAN, Venket et al. **A simple guide to *de novo* transcriptome assembly and annotation.** Briefings in bioinformatics, v. 23, n. 2, p. bbab563, 2022.

RODRIGUES, Thais B. et al. **First sprayable double-stranded RNA-based biopesticide product targets proteasome subunit beta type-5 in Colorado potato beetle (*Leptinotarsa decemlineata*).** Frontiers in Plant Science, v. 12, p. 728652, 2021.

RStudio Team (2020). **RStudio: Integrated Development for R.** RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

SCHVARTZMAN, Claudia et al. **RNAi in *Piezodorus guildinii* (Hemiptera: Pentatomidae): transcriptome assembly for the development of pest control strategies.** Frontiers in Plant Science, v. 13, p. 804839, 2022.

SIMÃO, Felipe A. et al. **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** Bioinformatics, v. 31, n. 19, p. 3210-3212, 2015.

SINGH, Satnam et al. **Using de novo transcriptome assembly and analysis to study RNAi in *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae)**. Scientific reports, v. 9, n. 1, p. 13710, 2019.

SINGH, Satnam et al. **Reference gene selection in *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae) and their normalization impact on gene expression in RNAi studies**. Journal of economic entomology, v. 112, n. 1, p. 371-381, 2019.

SONESON, Charlotte et al. **Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences**. F1000Research, v. 4, 2015.

UNTERGASSER, Andreas et al. **Primer3Plus, an enhanced web interface to Primer3**. Nucleic acids research, v. 35, n. suppl_2, p. W71-W74, 2007.

VALÉRIO, José R. **Cigarrinhas-das-pastagens / José Raul Valério**. — Dados eletrônicos. — Campo Grande, MS : Embrapa Gado de Corte, 2009.

WANG, Weixia et al. **Double-stranded RNA targeting calmodulin reveals a potential target for pest management of *Nilaparvata lugens***. Pest Management Science, v. 74, n. 7, p. 1711-1719, 2018.

YANG, Shan et al. **Knockdown of hexokinase in *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) by RNAi inhibits chitin synthesis and leads to abnormal phenotypes**. Pest Management Science, v. 78, n. 10, p. 4303-4313, 2022.

ZENG, Jie et al. **Evaluation of three vacuolar ATPase genes as potential RNAi target in *Henosepilachna vigintioctopunctata***. Journal of Asia-Pacific Entomology, v. 24, n. 2, p. 55-63, 2021.

ZHANG, Yue et al. **Stability evaluation of candidate reference genes for RT-qPCR normalization in *Lasioderma serricorne* (F.)**. Journal of Stored Products Research, v. 94, p. 101877, 2021.

ZHANG, Xiao et al. **Genomic and Transcriptomic Insights into the Genetic Basis of Foam Secretion in Rice Spittlebug *Callitettix versicolor***. International Journal of Molecular Sciences, v. 25, n. 4, p. 2012, 2024.

APÊNDICE A – PROCEDIMENTO DE VALIDAÇÃO DA EXPRESSÃO GÊNICA

Doze transcritos foram selecionados para validar os resultados de expressão diferencial *in silico*, incluindo aqueles anotados como maquinaria de RNAi com diferença considerável de expressão. Para isso, os cDNAs foram sintetizados com o kit GoScript™ Reverse Transcription System A5000 (Promega) e as reações de PCR quantitativo foram realizadas utilizando o mix comercial Taq DNA Polymerase Master Mix Red (Ampliqon).

Três dos genes selecionados foram descartados devido à baixa eficiência na amplificação: colina desidrogenase dependente de oxigênio (BE), membro 1 da classe B do receptor Scavenger (SC) e membro 2 da família transmembrana SID1 (SI). Além disso, dois dos quatro genes de referência testados foram selecionados como controles normalizadores com base na menor variação no ciclo de quantificação (Cq), sendo eles os RNA's ribossômicos 18 e 28s.

Os resultados das reações foram analisados através do programa 7500 Software v2.3 e os valores de logFC calculados. Para visualização e interpretação, as informações foram representadas de maneira gráfica (Figura A1 e Figura A2).

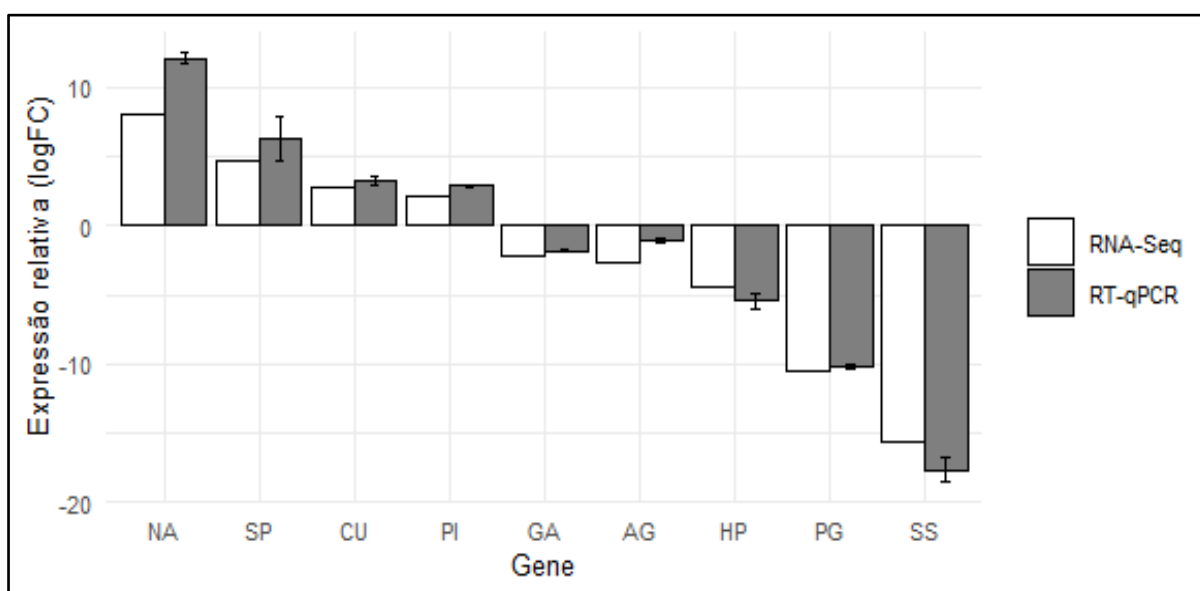


Figura A1 – Representação da comparação entre as expressões relativas de genes em RNA-Seq e RT-qPCR. Colunas brancas representam dimensões dos valores de logFC obtidos via RNA Seq para cada gene, colunas cinzas igualmente porém para RT qPCR. SP: Proteína de membrana flagelar de espermatozoides de 63 Kda; CU: Proteína de cutícula larval; PI: Proteína Siwi semelhante a Piwi; GA: Proteína

Gawky; AG: Proteína argonauta 2; HP: Serina protease HP8 contendo domínio CLIP; PG: Proteína de reconhecimento de peptidoglicano SB1; SS: Proteína de armazenamento sexo-específica 1; NA: Proteína não anotada.

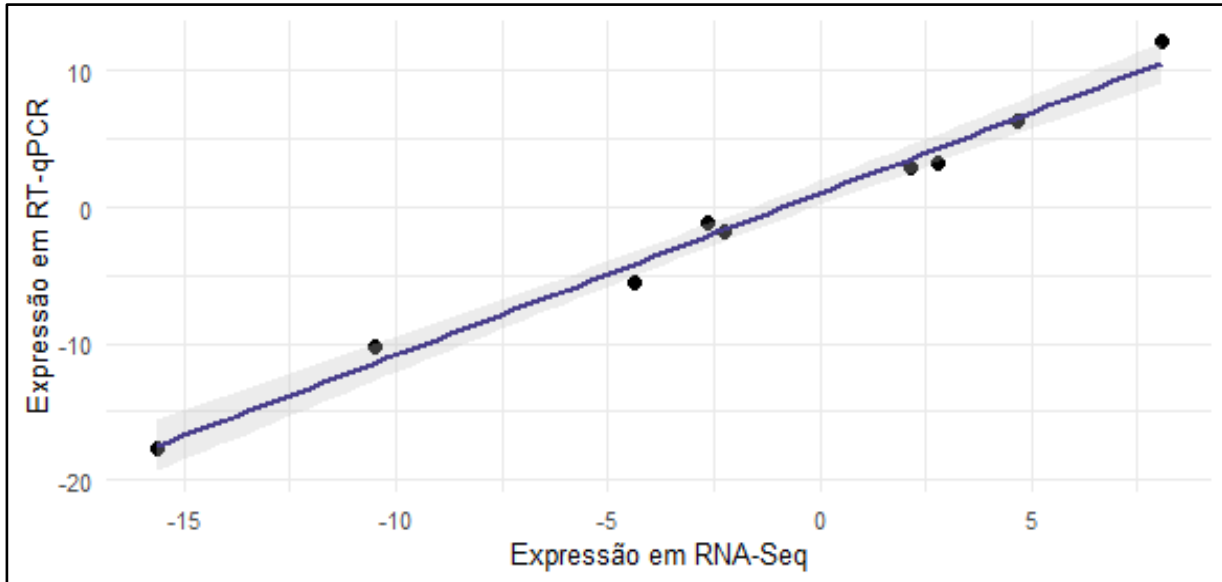


Figura A2 – Representação da dispersão entre as expressões em RNA-Seq e RT-qPCR.

Considerando que os padrões de expressão observados na análise RT-qPCR foram consistentes com os dados de RNA-Seq do transcriptoma, confirmou-se a confiabilidade dos resultados bioinformáticos.