UNIVERSIDADE ESTADUAL DE LONDRINA

THAYNÁ BISSON FERRAZ LOPES

INFLUÊNCIA DA VARIAÇÃO DE TEMPERATURA NA EXPRESSÃO GÊNICA DE *Bombyx mori* LINNAEUS, 1758 (LEPIDOPTERA: BOMBYCIDAE)

Londrina
2023

THAYNÁ BISSON FERRAZ LOPES

INFLUÊNCIA DA VARIAÇÃO DE TEMPERATURA NA EXPRESSÃO GÊNICA DE *Bombyx mori* LINNAEUS, 1758 (LEPIDOPTERA: BOMBYCIDAE)

Londrina
2023

THAYNÁ BISSON FERRAZ LOPES

INFLUÊNCIA DA VARIAÇÃO DE TEMPERATURA NA EXPRESSÃO GÊNICA DE *Bombyx mori* LINNAEUS, 1758 (LEPIDOPTERA: BOMBYCIDAE)

Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito final para a obtenção do título de Doutor.

Orientador: Prof. Dra. Renata da Rosa.
Coorientador: Prof. Dr. Rogério Fernandes de Souza.

Londrina
2023

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Lopes, Thayná Bisson Ferraz .
Influência da variação de temperatura na expressão gênica de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) / Thayná Bisson Ferraz Lopes. - Londrina, 2023.
88 f. : il.

Orientador: Renata da Rosa.
Coorientador: Rogério Fernandes de Souza.
Tese (Doutorado em Genética e Biologia Molecular) - Universidade Estadual de Londrina, Centro de Ciências Biológicas, Programa de Pós-Graduação em Genética e Biologia Molecular, 2023.
Inclui bibliografia.

1. Bicho-da-seda - Tese. 2. Transcriptoma - Tese. 3. Expressão diferencial - Tese. 4. Elementos transponíveis - Tese. I. Rosa, Renata da. II. Souza, Rogério Fernandes de. III. Universidade Estadual de Londrina. Centro de Ciências Biológicas. Programa de Pós-Graduação em Genética e Biologia Molecular. IV. Título.

CDU 575.1

THAYNÁ BISSON FERRAZ LOPES

INFLUÊNCIA DA VARIAÇÃO DE TEMPERATURA NA EXPRESSÃO GÊNICA DE *Bombyx mori* LINNAEUS, 1758 (LEPIDOPTERA: BOMBYCIDAE)

Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito final para a obtenção do título de Doutor.

BANCA EXAMINADORA

Orientador: Profa. Dra. Renata da Rosa
Universidade Estadual de Londrina – UEL

Prof. Dr. Alexandre Rossi Paschoal
Universidade Tecnológica do Paraná – UTFPR

Prof. Dr. Roberto Ferreira Artoni
Universidade Estadual de Ponta Grossa – UEPG

Profa. Dra. Silvia Helena Sofia
Universidade Estadual de Londrina – UEL

Prof. Dr. Laurival Antônio Vilas-Bôas
Universidade Estadual de Londrina – UEL

Londrina, 15 de fevereiro de 2023.

A todos os cientistas deste País. Em especial, aos envolvidos nas pesquisas de vacinas durante a pandemia que nos assolou nos últimos anos. Vocês nos ensinaram sobre persistência e salvam vidas todos os dias. Muitos de nós só chegamos aonde estamos hoje, vivos, graças a vocês.

Ser cientista é um ato de resistência.

AGRADECIMENTOS

Aos meus pais, Marilene e Marcos, e a minha irmã, Ana Beatriz, por todo o incentivo ao estudo desde sempre, por todo apoio, amor e carinho.

Ao meu companheiro de vida, Fernando. Obrigada por sempre fazer o possível e, muitas vezes, até o impossível, por mim e por minhas conquistas. Obrigada por todo incentivo, carinho e dedicação durante todos esses anos.

A minha orientadora, Prof. Dra. Renata da Rosa, pela oportunidade concedida novamente, pela confiança, por todos os ensinamentos e pela parceria de trabalho há nove anos. Fazer parte de projetos sob sua orientação como este, que buscam melhorar a qualidade de trabalho da comunidade externa, que valorizam o trabalho no campo e o trabalho feminino, é a maior realização que eu poderia ter como cientista.

Ao meu coorientador, Prof. Dr. Rogério Fernandes de Souza. Agradeço imensamente por tudo que me ensinou, desde quando foi meu professor de Evolução na graduação em Ciências Biológicas, até esses últimos quatro anos onde trabalhamos juntos durante o meu Doutorado. Obrigada por todo o conhecimento e prontidão. A sua contribuição foi essencial no presente trabalho e em minha formação.

Aos membros do Laboratório de Citogenética e Entomologia Molecular, que me auxiliaram e me ensinaram muito ao longo de todos os nove anos em que faço parte desse time. É um prazer imenso fazer parte de um grupo de pesquisa tão querido e divertido.

A BRATAC, pela parceria de trabalho e fomento das nossas pesquisas, e pela oportunidade de conhecer a fundo toda a cadeia da sericicultura.

Ao Prof. Dr. Mário Sérgio Mantovani, a Profa. Dra. Sandra Regina Lepri, a Profa. Dra. Simone Cristine Semprebon, e todos os membros Laboratório de Genética Toxicológica pelas contribuições nas análises moleculares.

Aos meus amigos. Os de longa data, os que a Universidade me proporcionou, os que moram longe e os que moram perto. Vocês tornam a minha vida mais divertida até mesmo nos momentos mais difíceis. Em especial, gostaria de agradecer a Aline e Erik, meus queridos amigos há tão pouco tempo, mas que foram essenciais, especialmente nessa reta final. Obrigada.

Agradeço aos amigos que o Programa de Doutorado Sanduíche no Exterior (PDSE) me proporcionou. Monyse e Rafael, que já faziam parte da minha turma de pós-graduação, mas que graças ao PDSE nos tornamos próximos. Agradeço também a Gabriela, que conheci

durante esse processo. Vocês três foram essenciais em um momento repleto de tantas mudanças e desafios. Passar por tudo isso com a amizade de vocês tornou tudo mais fácil e mais divertido.

Ao Prof. Dr. Michal Žurovec, por ter me recebido de braços abertos no Laboratório de Genética Molecular no Instituto de Entomologia (Biology Centre, Czech Academy of Sciences) em České Budějovice, República Tcheca. Desde meu primeiro contato sempre esteve de prontidão para me auxiliar. Durante esses seis meses sob sua orientação aprendi muito e nunca serei capaz de agradecer o suficiente por todos os seus ensinamentos, por sua preocupação comigo e com o meu trabalho, e por me apresentar lugares incríveis na República Tcheca. Agradeço também a todos os membros do laboratório por terem me recebido tão bem e por tudo que me ensinaram.

Ao meu querido amigo Peter Kmeť. Quando embarquei na jornada do Doutorado Sanduíche, jamais imaginei que pudesse ganhar um amigo tão especial no laboratório de destino. Obrigada por tudo. Sinto sua falta todos os dias e espero que em breve possamos nos reunir novamente.

A banca examinadora, pela disposição em avaliar e contribuir com este trabalho.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular, pela estrutura para a realização deste trabalho, e por todos os docentes e discentes que contribuíram e formaram meu conhecimento acadêmico.

A Universidade Estadual de Londrina. No dia 17 de janeiro de 2013 fui aprovada para o ingresso nesta Universidade. Neste ano de 2023, completei dez anos que faço parte de uma das melhores Universidades deste País. Agradeço por toda a estrutura física para a realização da minha graduação, Mestrado e Doutorado, por todo o conhecimento de docentes excepcionais que fizeram parte da minha formação e por me proporcionar grande parte do que sou hoje, não somente como pesquisadora, mas como pessoa. Foi nesta Universidade que vivi momentos muito difíceis, mas ao mesmo tempo vivi dez anos incríveis.

A CAPES, órgão de fomento à pesquisa científica, pelo apoio financeiro.

"A ciência é muito simples. Quando se torna complicada, em geral é porque o mundo é complicado ou porque nós é que somos complicados. Quando nos afastamos assustados da ciência, porque ela parece difícil demais (ou porque não fomos bem ensinados), abrimos mão da capacidade de cuidar de nosso futuro. Ficamos privados dos direitos civis. A nossa autoconfiança se deteriora.

Mas quando ultrapassamos essa barreira, quando as descobertas e os métodos da ciência se tornam claros para nós, quando compreendemos e empregamos esse conhecimento, sentimos uma profunda satisfação. (...) Sei pessoalmente, tanto por terem me explicado a ciência como pelas minhas tentativas de explicá-la aos outros, o quanto é gratificante quando a compreendemos, quando os termos obscuros de repente adquirem sentido, quando entendemos afinal do que se trata, quando maravilhas profundas nos são reveladas."

(Carl Sagan, 1996)

LOPES, Thayná Bisson Ferraz. **Influência da variação de temperatura na expressão gênica de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae).** 2023. 67 f. Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito final para a obtenção do título de Doutor.

RESUMO

O bicho-da-seda (*Bombyx mori*) é um inseto domesticado economicamente importante para a produção de seda e seus derivados. No entanto, variações de temperatura influem no desenvolvimento desse lepidóptero e impactam diretamente na produção de casulos de seda, ocasionando grandes perdas econômicas. Neste trabalho utilizamos o sequenciamento de RNA para a obtenção das bibliotecas e posterior montagem dos transcriptomas, utilizando um genoma de referência, a fim de entender a expressão gênica em duas condições de temperatura, 26°C e 34°C. O estudo comparativo dos transcritos revelou a expressão diferencial de vários genes relacionados a imunidade, fatores de transcrição, proteínas de choque térmico e genes envolvidos na produção de seda. Os dados de qRT-PCR validaram as regulações positivas e negativas de seis desses genes. Além disso, buscamos explorar a influência das duas temperaturas no desenvolvimento e produção de casulos de seda. As lagartas mantidas na temperatura mais elevada apresentaram média menor de peso corporal. O peso dos casulos produzidos e a taxa de conversão de seda também foram diretamente impactados pela temperatura de 34°C em comparação com as lagartas mantidas a 26°C, e representaram diferenças estatisticamente significativas. Enquanto isso, as taxas de mortalidade foram semelhantes às obtidas em campos de criação e consideradas normais. Esses dados evidenciam que, apesar da temperatura não ter ocasionado mortalidade anormal, o desenvolvimento e a produção de casulos foram prejudicados. Outro aspecto abordado neste estudo foi a identificação e a caracterização de elementos transponíveis (TEs) no transcriptoma de *B. mori* utilizando seis bancos de dados diferentes. Aproximadamente 40% dos transcritos totais apresentaram similaridade com sequências de TEs em pelo menos um dos bancos de dados. A análise de expressão diferencial entre sequências com semelhança a TEs também foi realizada. Dentre as sequências diferencialmente expressas com similaridades a TEs, os retrotransposons representaram a maior fração. Enquanto isso, a quantidade de sequências reguladas negativamente foi 2,5 vezes maior que as sequências reguladas positivamente. Nossos dados demonstram que a temperatura pode estar relacionada com a modulação de TEs. Também foi possível observar a possibilidade da ocorrência de inserções de TEs em genes específicos, devido ao elevado número de sequências com identificação com TEs e a anotação das sequências que codificam diferentes funções em *B. mori*. Os resultados obtidos nesse estudo fornecem informações importantes para o entendimento da influência da temperatura no desenvolvimento e expressão de genes em *B. mori*, bem como a identificação e expressão de TEs nesse inseto. Esses dados podem direcionar futuros estudos relacionados a temperatura em *B. mori* da linhagem comercial brasileira e auxiliar em futuras estratégias de climatização de barracões de produção de casulos de seda no Brasil e no mundo.

Palavras-chave: Transcriptoma; Estresse térmico; Expressão diferencial; Elementos transponíveis.

LOPES, Thayná Bisson Ferraz. **Temperature variation influence in *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) gene expression.** 2023. 67 p. Tese apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular, da Universidade Estadual de Londrina, como requisito final para a obtenção do título de Doutor.

ABSTRACT

Silkworms (*Bombyx mori*) are an economic important domesticated insects to producing silk and its byproducts. However, fluctuations in temperature affect the lepidopteran development and impact silk cocoons production, which cause great monetary losses. In this present study, RNA sequencing is used to obtain libraries and assemble the transcriptomes, using a reference genome in order to comprehend the gene expression in different temperatures – 26 ºC and 34 ºC. The comparative study of the transcripts reveals the differential expression of several genes related to immunity, transcription factors, heat shock proteins and silk production. Both positive and negative regulations in six of such genes were validated by qRT-PCR. Furthermore, we aimed to explore both temperatures' influence in the development and production of silk cocoons. Silkworms kept in the higher temperature displayed lower average body weight. Produced cocoons' weight and silk conversion ratio were also impacted by the 34 ºC temperature in comparison to the silkworms kept at 26 ºC, presenting significant statistical differences. Meanwhile, mortality rates were similar to the ones obtained in sericulture farms and considered normal. The data suggests that even though mortality rates were not abnormal, growth and production of cocoons were impaired. Other aspect approached by this study were the identification and characterization of transposable elements (TEs) in *B. mori* transcriptome consulting six distinct databases. Approximately 40% of the transcripts presented similarities to TEs sequences in at least one of the consulted databases. Differential expression analysis among the similar TEs sequences were also performed. Retrotransposons represented the largest portion amongst the differentially expressed TEs sequences. Concurrently the number of sequences negatively regulated were 2.5 times larger than those positively regulated. The data suggests that temperature might be related to the TEs modulation. TEs insertions in specific genes might be possible due to the elevated number of sequences similar to TEs and the annotation of sequences that code different functions in *B. mori*. Results obtained in the present study provided valuable information on the influence of temperature in *B. mori* development and gene expression, as well as TEs identification and their expression in silkworms. Such data might direct further studies related to the temperature in the Brazilian commercial lineage of *B. mori* and aid future climatization strategies for sericulture worldwide.

Key-words: Transcriptome; Heat stress; Differential expression; Transposable elements.

LISTA DE FIGURAS

Figura 1 - Bicho-da-seda *(Bombyx mori)* durante o quinto instar do estágio larval e o casulo de seda (Emery, 2019)..... 05

Figura 2 - Morfologia de um casulo de *Bombyx mori* no plano da parede do casulo e em sua espessura. Superior: morfologia da superfície do casulo e sua estrutura, onde cada fibra consiste em fios fundidos exsudados do par de glândulas de seda. Inferior: morfologia multicamadas do casulo (Chen et al., 2012)..... 05

Figura 3 - Glândula sericígena de *B. mori*. (1) Porção anterior, (2) porção média e (3) porção posterior. Elaborado pelo autor (2022)..... 06

Figura 4 - Corte transversal das fibras de seda de *B. mori*, evidenciando a morfologia microestrutural. A fibra de seda é composta por dois filamentos centrais de fibroína envoltos externamente pela sericina (Guan et al., 2017)..... 07

Figura 5 - Sistema de classificação para os elementos transponíveis segundo Wicker et al. (2007). A classificação hierárquica divide os elementos em duas principais classes (retrotransposons e DNA transposons) com base na presença ou ausência de RNA como intermediário, respectivamente. Em seguida, são divididos em subclasses, ordens superfamílias 13

CAPÍTULO I: "Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae)"

Figura 1 - *Pipeline* aplicada na análise da expressão gênica e da produção de casulos de *Bombyx mori*. Abreviação: DE = Diferencialmente expresso 26

Figura 2 - Termos mais frequentes do Gene Ontology para cada uma das três categorias principais, nos transcritos diferencialmente expressos de *B. mori*: **a.** *up* regulados, **b.** *down* regulados (em lagartas criadas a 34 ºC em relação às criadas a 26 ºC)..... 34

Figura 3 - Domínios Pfam mais frequentes para transcritos diferencialmente expressos de *B. mori*. **a.** *Up* regulados, **b.** *down* regulados (em lagartas criadas a 34 ºC em relação às criadas a 26 ºC) 36

Figura 4 - *Heatmap* com o resumo dos principais genes diferencialmente expressos no transcriptomas de *Bombyx mori*. Os níveis de mRNA são representados pelas contagens estimadas normalizadas em DESeq2 (rlog). As diferentes tonalidades de vermelho indicam genes *up* (em tons mais escuros) e *down* regulados (destacados em tons mais claros) em cada biblioteca. Os nomes dos genes estão à direita com as respectivas identificações dos transcritos e os grupos de genes envolvidos na resposta à temperatura identificados à esquerda 37

Figura 5 - Análise de RT-qPCR (log2foldchange, método $2^{-\Delta\Delta Ct}$) de genes diferencialmente expressos na análise de RNA-seq. Os dados foram normalizados com os genes de α-tubulina e RpL3 e estão apresentados com o erro padrão médio (SEM). L-fib, H-fib, Ser1 e Hsp70 foram regulados negativamente, enquanto Ser2 e Hsp19.5 foram regulados positivamente 38

CAPÍTULO II: "Variação na temperatura de criação de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) e sua influência na expressão de elementos transponíveis"

Figura 1 - Domínios proteicos Pfam característicos de elementos transponíveis nas sequências do transcriptoma montado de *B. mori* e o número de transcritos relacionados a cada domínio. O domínio RVT_1 foi o mais numeroso 55

Figura 2 - Distribuição de sequências relacionadas a TEs identificados no transcriptoma de *B. mori* em duas classes, elementos de classe I (retrotransposons) e elementos de classe II (DNA transposons), e em suas respectivas ordens. Os elementos de Classe I foram predominantes 56

Figura 3 - Distribuição de sequências relacionadas a TEs identificados no transcriptoma de *B. mori*. No eixo vertical, o número de sequências correspondentes a cada elemento. No eixo horizontal, a divisão entre

duas classes (I e II) e classificação dos elementos em superfamílias/clado..... 57

Figura 4 - Representação de elementos transponíveis diferencialmente expressos no transcriptoma de *B. mori*. Os níveis de RNAm são representados pelas contagens estimadas normalizadas em DESeq2 (rlog). As diferentes tonalidades de vermelho indicam genes *up* (em tons mais escuros) e *down* regulados (destacados em tons mais claros) em cada biblioteca. Os nomes putativos dos genes estão à direita com as identificações dos transcritos correspondentes 59

Figura 5 - Representação esquemática da organização dos domínios de codificação de proteínas na sequência STRG 9105.1 (*fibrohexamerin*) em *B. mori*. A sequência foi representada em número de aminoácidos. Abreviações: RT = transcriptase reversa; ORF = *Open Reading Frame* 59

LISTA DE TABELAS

CAPÍTULO I: "Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae)"

Tabela 1 - *Primers* utilizados na qRT-PCR 30

Tabela 2 - Resumo das estatísticas de sequenciamento das seis bibliotecas *paired-end* de *B. mori* 31

Tabela 3 - Avaliação da qualidade da montagem do transcriptoma de *B. mori* em comparação com bancos de dados (*_odb10) de Lepidoptera e Insecta por meio da ferramenta BUSCO 32

Tabela 4 - Peso das lagartas (g) de *B. mori* nos dias 1, 3 e 6 do quinto instar 39

Tabela 5 - Medição dos casulos de *B. mori* produzidos 39

TABELAS SUPLEMENTARES

Acesso:

https://drive.google.com/drive/folders/1BBgI09UdNoi1-AomLdDLMd7rihGWaZYa?usp=share_link

CAPÍTULO I: "Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae)"

Tabela S1 - Identificação, tamanho das sequências e contagem de *reads* dos transcritos de *B. mori* montados pelo StringTie, utilizando um genoma de referência ..... Link externo

Tabela S2 - Transcritos diferencialmente expressos de *B. mori*. Log2FoldChange positivo (vermelho): sequências *up* reguladas em lagartas criadas a 34 ºC em relação às criadas a 26 ºC; log2FoldChange negativo (azul): sequências *down* reguladas em lagartas criadas a 34 ºC em relação às criadas a 26 ºC ..... Link externo

Tabela S3a - Distribuição dos transcritos diferencialmente expressos *up* regulados em termos do Gene Ontology em três categorias principais: processo biológico (em vermelho), componente celular (em azul) e função molecular (em amarelo) ..... Link externo

Tabela S3b - Distribuição dos transcritos diferencialmente expressos *down* regulados em termos do Gene Ontology em três categorias principais: processo biológico (em vermelho), componente celular (em azul) e função molecular (em amarelo)..... Link externo

Tabela S4a - Mapeamento das vias metabólicas relacionadas aos transcritos diferencialmente expressos *up* regulados..... Link externo

Tabela S4b - Mapeamento das vias metabólicas relacionadas aos transcritos diferencialmente expressos *down* regulados ..... Link externo

Tabela S5 - Domínios Pfam mais frequentes para transcritos diferencialmente expressos de *B. mori* em lagartas submetidas a 34 ºC em relação às submetidas a 26 ºC ..... Link externo

Tabela S6 - Famílias de RNA identificadas nos transcritos diferencialmente expressos de *B. mori* pelo Rfam..... Link externo

CAPÍTULO II: "Variação na temperatura de criação de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) e sua influência na expressão de elementos transponíveis"

Tabela S1 - Transcritos relacionados a TEs em comparação com o banco de dados Repbase (Taxon: Hexapoda). Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente..... Link externo

Tabela S2 - Transcritos relacionados a TEs em comparação com o banco de dados ArTEDB (Taxon: Insecta). Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente ..... Link externo

Tabela S3 - Identificação de famílias de TEs do transcriptoma montado no banco de dados das famílias curadas do Dfam. Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente ..... Link externo

Tabela S4 - Transcritos relacionados a TEs das ordens LTR e TIR no banco de dados do Gypsy Database (GyDB). Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente ..... Link externo

Tabela S5 - Identificação de domínios proteicos relacionados a TEs no banco de dados Pfam. Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente. ..... Link externo

Tabela S6 - Identificação de TEs no transcriptoma montado de *B. mori* a partir da identificação de TEs no genoma de referência, utilizando o RepeatModeler2. Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos,

respectivamente ..... Link externo

Tabela S7 - Classificação dos elementos transponíveis no transcriptoma de *B. mori* ..... Link externo

Tabela S8 - Anotação de elementos transponíveis diferencialmente expressos em *B. mori*. Os transcritos destacados mostraram-se diferencialmente expressos em lagartas mantidas a 34 ºC em relação a lagartas de *B. mori* mantidas a 26 ºC durante o quinto instar. As cores vermelho e azul representam a *up* e *down* regulação destes transcritos, respectivamente ..Link externo

LISTA DE ABREVIATURAS E SIGLAS

AP	Proteinase aspártica (*Aspartic proteinase*)
cDNA	DNA complementar
DE	Diferencialmente expresso
DNA	Ácido desoxirribonucleico
DIRS	Sequência de Repetição Intermediária de *Dictyostelium* (*Dictyostelium Intermediate Repeat Sequence*)
GO	*Gene Ontology*
INT	Integrase
kDa	Quilodalton
LINE	Elementos nucleares longos intercalados (*Long Interpersed Nuclear Element*)
LTR	Sequências terminais repetidas (*Long Terminal Repeat*)
ORF	Fase de leitura aberta (*Open reading frame*)
PLE	*Penelope-like*
qRT-PCR	*Real-Time Quantitative Reverse Transcription PCR*
RH	Ribonuclease H
RNA	Ácido ribonucleico
RNAm	RNA mensageiro
RNAr	RNA ribossômico
RNAt	RNA transportador
RT	Transcriptase reversa (*Reverse transcriptase*)
SINE	Elementos nucleares curtos intercalados (*Short Interspersed Nuclear Element*)
TE	Elementos genéticos móveis (*Transposable elements*)
TIR	Repetição terminal invertida (*Terminal Inverted Repeat*)
UTR	Região não traduzida (*Untranslated Region*)

SUMÁRIO

1 INTRODUÇÃO 01

2 REVISÃO BIBLIOGRÁFICA 02
2.1 Sericicultura: histórico e economia 02
2.2 *Bombyx mori* Linnaeus, 1758 04
2.3 Influência da variação térmica na criação das larvas de *Bombyx mori* 08
2.4 Estudo de transcriptomas e análise de expressão diferencial 09
2.5 Elementos genéticos móveis 11

3 OBJETIVOS 16
3.1 Objetivo geral 16
3.2 Objetivos específicos 16

REFERÊNCIAS 17

5 CAPÍTULO I 23
Resumo 24
5.1 Introdução 25
5.2 Material e métodos 26
5.2.1 Criação das lagartas 26
5.2.2 Extração de RNA e sequenciamento 27
5.2.3 Montagem do transcriptoma e análise de expressão diferencial 28
5.2.4 Validação dos genes diferencialmente expressos (RT-qPCR) 29
5.2.5 Avaliação morfológica das lagartas e medição dos casulos produzidos 30
5.3 Resultados 31
5.3.1 Sequenciamento e montagem do transcriptoma 31
5.3.2 Análise de expressão diferencial e anotação funcional 32
5.3.3 RT-qPCR 38
5.3.4 Parâmetros morfológicos das lagartas e casulos 38
5.4 Discussão 40
Agradecimentos 43

Referências 44

6 CAPÍTULO II 48

Resumo 49

6.1 Introdução 50

6.2 Material e métodos 51

6.2.1 Criação das lagartas 51

6.2.2 Extração de RNA e sequenciamento 51

6.2.3 Montagem do transcriptoma e análise de expressão diferencial 52

6.2.4 Anotação de TEs 53

6.3 Resultados 54

6.3.1 Identificação e caracterização de elementos transponíveis no transcriptoma de *Bombyx mori* 54

6.3.2 Expressão diferencial 58

6.4 Discussão 60

Agradecimentos 63

Referências 64

7 CONCLUSÕES FINAIS 67

1. INTRODUÇÃO

A seda é considerada um dos tecidos mais antigos e nobres comercializados no mundo. É uma fibra ecológica de alta resistência, sendo utilizada além da confecção de tecidos, como na elaboração de materiais eletrônicos, aeronáuticos e médicos (Liu et al., 2010; Babu, 2015). A sericicultura é atividade agrícola que compreende a criação da lagarta *Bombyx mori* (Linnaeus, 1758), popularmente conhecida como bicho-da-seda, a qual durante a fase de pupa/crisálida produz os casulos de onde são retirados os fios de seda (Roy et al., 2012). A atividade compreende desde o plantio de amoreiras para a alimentação das lagartas, sua criação e produção dos casulos, até o processo industrial de fiação.

O mercado da sericicultura é capaz de ajudar a manutenção de empregos no campo, evitando o êxodo rural e garantindo remuneração, requerendo esta atividade, pequenos investimentos que proporcionam matéria-prima para as indústrias têxteis (CONAB, 2016). Além disso, é relevante no aspecto social pois é considerada uma atividade de baixo impacto ambiental, contribuindo, em quase totalidade, nos 17 Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas (ONU). A produção de seda promove a participação da mulher no mercado de trabalho e igualdade de gênero, não competindo com outras produções agrícolas, colaborando com a conservação do solo e permitindo o estabelecimento de parcerias globais entre diversos países desenvolvidos e países em desenvolvimento (IDR Paraná, 2022).

A produção brasileira de seda, embora ainda pareça pouco expressiva em termos quantitativos (1% da produção mundial), já é suficiente para colocar o Brasil em 6º lugar no ranking mundial, sendo o único exportador em escala comercial do Ocidente, exportando 90% de sua produção, principalmente para o Japão, Coréia do Sul, França e Itália (DERAL, 2018; Inserco, 2023). O Estado do Paraná, em especial, contribui com 83% da produção nacional total de casulos de seda, distribuídos em 165 municípios do Estado e contando com a participação de quase 1.900 produtores (SEAB, 2020). Na safra 2019/2020, 4.744,76 hectares de área produziram 2.213,42 toneladas de casulos de *B. mori*, com valor bruto de produção de aproximadamente R$ 44 milhões (SEAB, 2020). Além disso, o fio de seda brasileiro tem ganhado notoriedade internacional pela excelente qualidade, sendo considerado atualmente o melhor fio de seda do mundo (IBGE, 2018).

A fibra de seda natural é um filamento contínuo de proteína, confeccionada em forma de casulos, que proporcionam uma proteção para elevar as chances de sobrevivência

do bicho-da-seda no ambiente, evitando temperaturas extremas (Roy et al., 2012). A temperatura considerada adequada para o seu desenvolvimento se situa entre 20 e 30ºC e, a umidade adequada fica em torno de 80%. Fora desses parâmetros, a lagarta diminui sua alimentação, chegando a interrompê-la, ocasionando sua morte (Hussain et al., 2011).

A composição e organização do casulo são essenciais para garantir a sobrevivência das lagartas de *B. mori* frente às oscilações ambientais. O desenvolvimento ótimo das lagartas e a produção de casulos de alta qualidade são os grandes objetivos da sericicultura, uma vez que estes constituem a matéria-prima base da produção de seda. Sendo assim, compreender como o fator temperatura atua no desenvolvimento e mortalidade das lagartas, bem como na confecção dos casulos é fundamental para a elaboração de estratégias de seleção de linhagens adaptadas às oscilações deste parâmetro no ambiente. Neste sentido, a obtenção e análise do transcriptoma de lagartas do *B. mori* submetidos a diferentes condições de temperatura pode fornecer importantes parâmetros e respostas, permitindo que novas estratégias possam ser desenvolvidas para o aumento da produção de casulos, aumento da produtividade industrial e valorização da sericicultura.

2. REVISÃO BIBLIOGRÁFICA

2.1 Sericicultura: histórico e economia

A sericicultura é uma atividade originária da China e, ao longo dos anos, foi introduzida no Japão, na Ásia Central e Grécia. Em 1740, *B. mori* passou a ser criado em outros países da Europa, como a Espanha, França, Itália e Áustria. No território brasileiro a introdução da sericicultura ocorreu no estado do Rio de Janeiro (1848) e posteriormente na cidade de Campinas/SP (1922). A partir da década de 1930, a atividade sericícola tornou-se importante para a agroindústria brasileira e, atualmente, o estado do Paraná é o maior produtor nacional de casulos de bicho-da-seda (CONAB, 2016; DERAL, SEAB, 2020; IDR Paraná, 2022).

A atividade sericícola se inicia com o cultivo da amoreira (*Morus* sp.), alimento exclusivo da dieta das lagartas de *B. mori*, e passa pelo recebimento dos ovos - fornecidos pela indústria - e segue com a criação destas lagartas em galpões/barracões denominados sirgarias, se completando confecção dos casulos, processamento industrial e venda da seda processada. No Brasil, o fornecimento de ovos e acompanhamento do processo produtivo, assim como o processamento de casulos produzidos e exportação é realizado pela BRATAC

– Fiação de seda, localizada em Londrina/PR, e é a única em atividade no País. A qualidade de seus produtos alcança reconhecimento nos melhores mercados consumidores de fios de seda, tais como Europa, Japão e EUA (BRATAC, 2021).

A criação de *B. mori* e a produção de fios de altíssimo valor comercial ocasiona a integração entre a produção rural e a indústria (Júnior et al., 2018). A atividade serve como fonte de incremento da renda para a população rural e permanência dos sistemas de produção familiares, aumentando a atratividade para a sucessão familiar, favorecendo a permanência dos jovens no campo com qualidade de vida para as famílias; além de contribuir para a preservação do ambiente rural, principalmente pela não utilização de agroquímicos na sericicultura (Júnior et al., 2018; IDR Paraná, 2022). Além disso, a sericicultura tem um papel muito importante na valorização do trabalho feminino, tanto no campo quanto na indústria (Inserco, 2022).

Durante o auge da sericicultura no estado do Paraná, mais de 14 mil produtores estavam envolvidos na atividade. No entanto, a redução da mão de obra disponível no meio rural reduziu as áreas de cultivo de amoreiras, a produção de casulos e o número de sericicultores na atividade e nos últimos anos. A atividade sericícola que anteriormente demandava grandes áreas cultivadas e apresentava baixos índices de produtividade, visa aumentar os níveis de produtividade por hectare, o que tem atraído novos agricultores e jovens para atividade atualmente. Além disso, o processo de modernização da sericicultura é responsável por novas oportunidades de geração de renda e torna-se ainda mais atrativa devido a necessidade de um baixo capital de giro e por toda a produção ser voltada para o mercado internacional e cultura correta, orgânica e viavelmente econômica (Júnior et al., 2018).

No entanto, a atividade sericícola é afetada por diferentes fatores, alguns em comum com a produção agropecuária de modo geral: condições climáticas, a oscilação econômica, a cotação do dólar, entre outros (CONAB, 2016; Cirio, 2018). A variação climática, em especial, ocasiona uma pausa na atividade dos sericicultores, especialmente durante o período de entressafra (inverno nos meses de junho, julho e agosto), pois os insetos têm sua taxa metabólica reduzida em temperaturas muito baixas, ocasionando a produção de uma fibra de seda de baixa qualidade (Li et al. 2012). Além disso, os sericicultores ainda relatam uma elevada taxa de mortalidade de insetos em meses de verão intenso, como nos meses de novembro, dezembro e janeiro, e mesmo os poucos insetos sobreviventes não produzem um casulo de qualidade: sua taxa metabólica é elevada e a produção de seda

desenfreada ocasiona casulos com estrutura e qualidade inferior (Inf. Pessoal – Encontros de Sericultura; Li et al. 2012; IDR Paraná, 2022).

Em um acompanhamento das safras de 2014/2015, 2015/2016, 2016/2017 e 2017/2018, envolvendo 21 sericicultores distribuídos por 16 municípios de diferentes regiões produtoras do Estado do Paraná, Júnior et al. (2018) descreve uma significante diferença de resultados em unidades onde os produtores, com ajustes e cuidados na criação (automatização dos processos, maior atenção na qualidade e quantidade das folhas fornecidas aos insetos e até mesmo a estrutura dos barracões), atingiram produtividades superiores a 1.500 kg de casulos/ha de amoreira, enquanto a média de produtividade paranaense no mesmo período foi de 635 kg casulos/ha de amoreira. Isso evidencia a importância da pesquisa e da implementação de inovações na cadeia sericícola. Além disso, o autor ainda sugere como estratégia potencial a ser pesquisada e implementada no sistema produtivo, o controle de temperatura nos barracões de criação, que poderia elevar a produção a um número muito maior – e estender o número de meses de produtivos.

2.2 *Bombyx mori* Linnaeus, 1758

Bombyx mori (Lepidoptera: Bombycidae) é um inseto domesticado pelo homem há pelo menos cinco mil anos. É considerado um sistema modelo para estudos biológicos devido a sua fácil criação em laboratório e seu tamanho corporal grande em comparação com outras espécies de lepidópteros, o que facilita a obtenção de material biológico, além de sua importância econômica na sericicultura (Goldsmith et al., 2005; Takasu et al., 2016).

São animais holometábolos, com quatro estágios de vida bem definidos: ovo, larva (com cinco ínstares), pupa/crisálida (casulo) e adulto (mariposa), com duração de aproximadamente cinquenta dias, sendo o estágio larval o mais longo (aproximadamente 28 dias). O processo de síntese do casulo, que é a matéria prima da cadeia produtiva de seda, pode levar de três a quatro dias. Os níveis de hormônio juvenil circulantes na hemolinfa diminuem no final do estágio larval, o que ocasiona o início da confecção do casulo pelas lagartas em torno de si mesmas (Hanada & Watanabe, 1986). Cada casulo é constituído por um único fio de seda que, quando desenrolado, pode alcançar 1000 a 1500 metros de comprimento (Chen et al., 2012; Babu, 2015). (Figura 2).

Figura 1. Bicho-da-seda *(Bombyx mori)* durante o quinto instar do estágio larval e o casulo de seda (Emery, 2019).

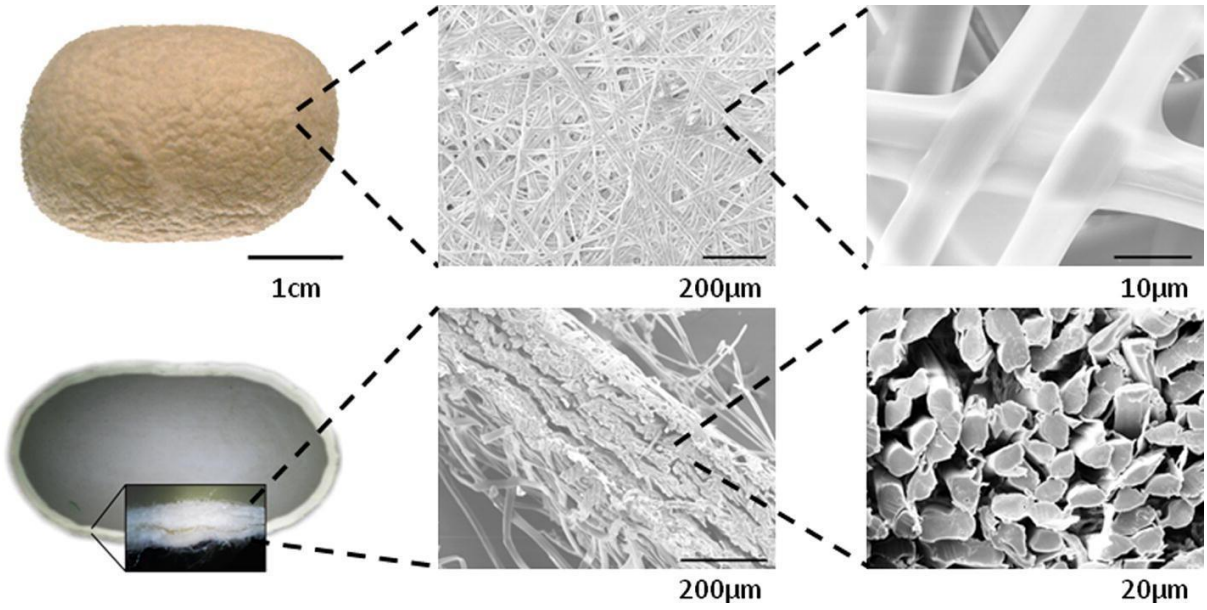

Figura 2. Morfologia de um casulo de *Bombyx mori* no plano da parede do casulo e em sua espessura. Superior: morfologia da superfície do casulo e sua estrutura, onde cada fibra consiste em fios fundidos exsudados do par de glândulas de seda. Inferior: morfologia multicamadas do casulo (Chen et al., 2012).

As fibras de seda que compõe os casulos são compostas por duas proteínas principais: a fibroína (75-83%) e a sericina (17-25%), além de gorduras e carboidratos, que representam menos de 2% da composição (Babu, 2015). Ambas as proteínas são produzidas nas glândulas sericígenas do inseto, as quais são divididas em três porções: posterior, média e anterior (Figura 2). A fibroína, produzida na porção posterior e armazenada na porção mediana, representa a proteína estrutural do fio da seda. A sericina é sintetizada na porção mediana e desempenha o papel de "goma proteica", unindo os filamentos de fibroína e mantendo a estrutura protetiva do casulo (Hanada & Watanabe, 1986; Nicodemo et al., 2014; Guan et al., 2017). Estruturalmente, os fios de seda são compostos por dois filamentos centrais de fibroína envoltos por uma camada externa de sericina (Nicodemo et al., 2014; Guan et al., 2017) (Figura 3). Dentre os aminoácidos mais abundantes na composição do fio de seda estão a glicina (~40%), alanina (~28%) e serina (~10%), seguidos de tirosina (~5%) e valina (~2%) (Babu, 2015).

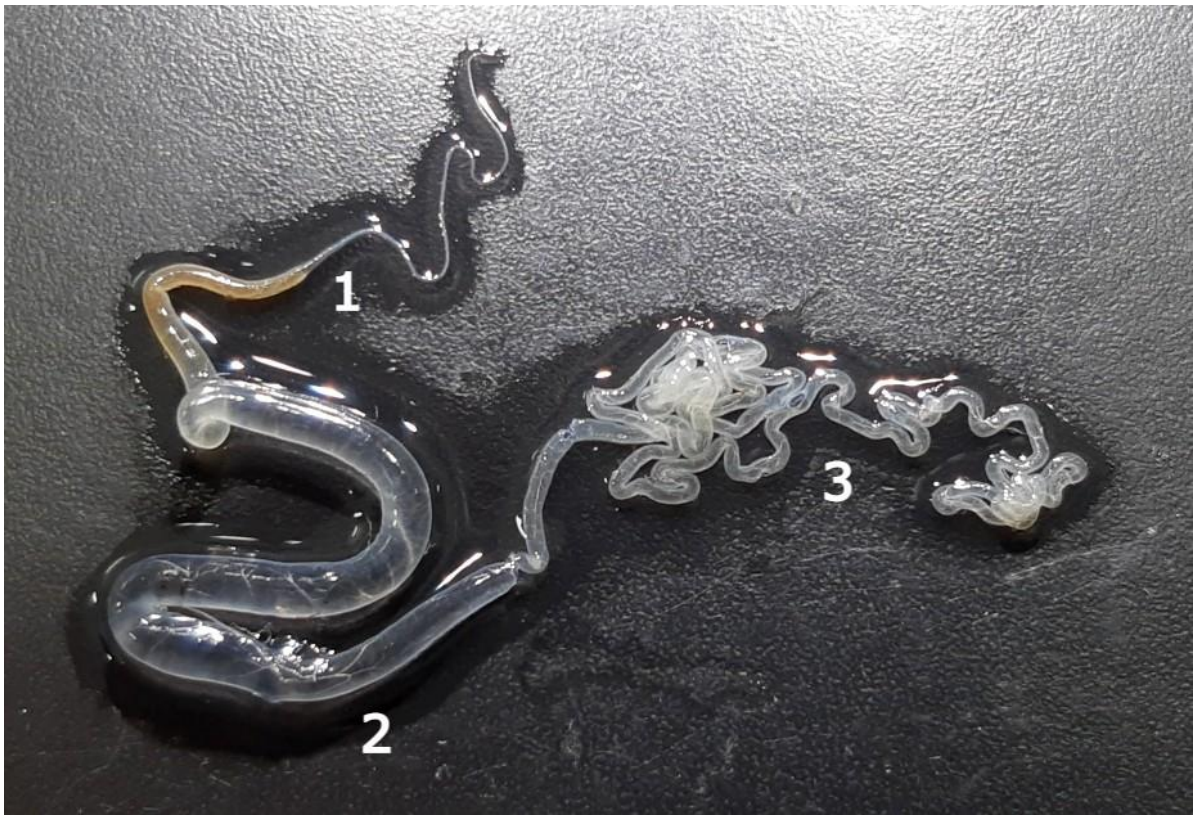

Figura 3. Glândula sericígena de *B. mori*. (1) Porção anterior, (2) porção média e (3) porção posterior. Elaborado pelo autor (2022).

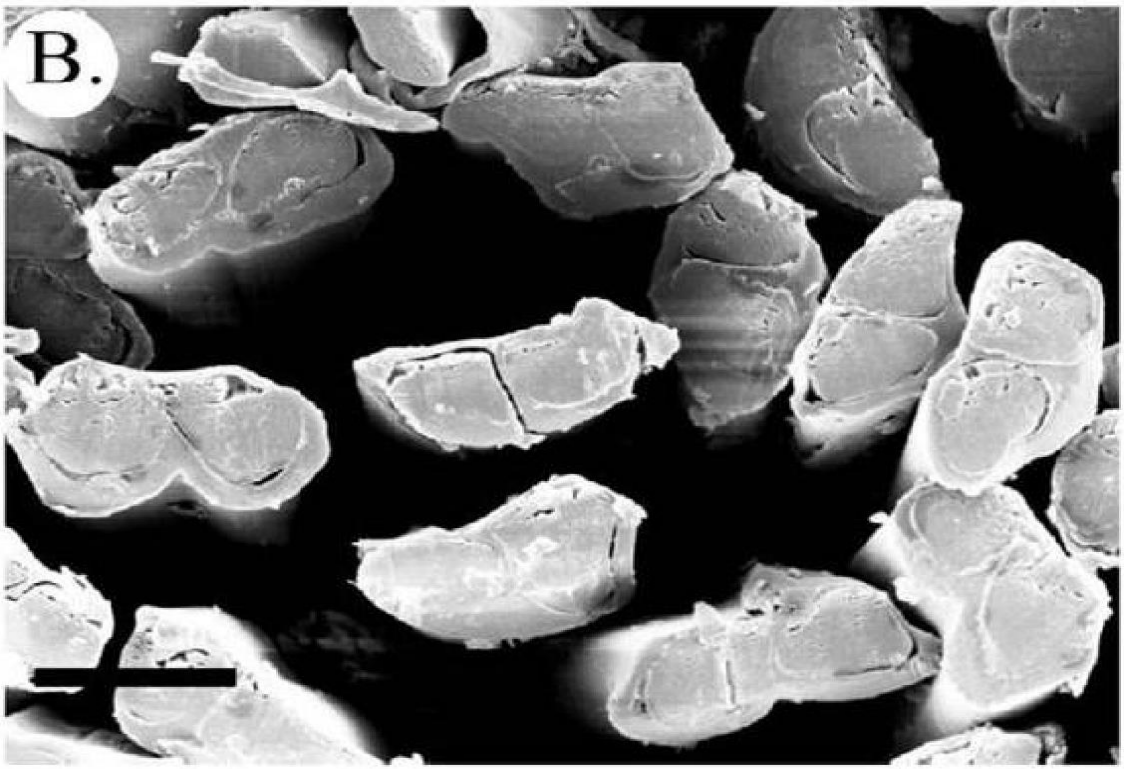

Figura 4. Corte transversal das fibras de seda de *B. mori*, evidenciando a morfologia microestrutural. A fibra de seda é composta por dois filamentos centrais de fibroína envoltos externamente pela sericina (Guan et al., 2017).

O elevado teor proteico do fio de seda, assim como sua qualidade, se dá principalmente por dois fatores: alimentação (disponibilidade e qualidade do alimento) e suas condições de criação (temperatura, umidade e manejo). Durante todo o período larval as lagartas alimentam-se exclusivamente com folhas de amoreira (*Morus alba* L.) até multiplicarem o seu tamanho (em média doze mil vezes em relação ao peso inicial) e atingirem o estágio de pupa (casulo). Todos os nutrientes e água necessários ao desenvolvimento das lagartas de *B. mori* são provenientes das folhas da amoreira, principalmente proteínas, carboidratos, lipídios, sais inorgânicos, água e vitaminas (Hanada & Watanabe, 1986; Dong et al., 2017). Na fase adulta, as mariposas cessam a alimentação. A qualidade e a disponibilidade das folhas para as lagartas durante seu crescimento são vitais para a produção de casulos de ótima qualidade.

Os caracteres biológicos e relacionados ao casulo são influenciados pela temperatura ambiente, épocas de cultivo, qualidade da folha da amoreira e constituição genética das linhagens de *B. mori*. As diferenças sazonais nos componentes ambientais afetam

consideravelmente a expressão genotípica na forma de produção fenotípica da cultura do bicho-da-seda, como o peso do casulo, peso da casca e proporção da casca do casulo (Rahmathulla, 2012). Em comparação com outras espécies selvagens produtoras de seda, como *Antheraea pernyi* e *Antheraea mylitta*, *B. mori* domesticado produz casulos menores, fios menos compactos e capacidade de regulação térmica inferior (Zhang et al., 2013). Essas características constituem um fator negativo para as incubadoras de criação e manutenção do inseto, uma vez que as lagartas domesticadas não toleram súbitas oscilações térmicas, resultando em um elevado índice de mortalidade. O intervalo entre 20 e 30ºC, é considerado como o rol de temperaturas adequadas para o desenvolvimento, com umidade ideal em torno de 80%. Mudanças nesses parâmetros ocasionam a morte do inseto e consequentemente queda da produção de seda (Hanada & Watanabe, 1986; Hussain et al., 2011).

2.3 Influência da variação térmica na criação das larvas de *Bombyx mori*

A contínua domesticação de *B. mori* por séculos tornou a espécie suscetível ao ambiente onde determinadas alterações, por menores que sejam, podem ser prejudiciais ou até mesmo letais, devido ao "estresse" ou choque para esses insetos sensíveis (Howrelia et al., 2011). Temperatura, umidade, gases, luz e outros fatores externos, têm efeito significativo sobre a fisiologia do inseto, dependendo da combinação de fatores e estágios de desenvolvimento que afetam o crescimento, desenvolvimento, produtividade e qualidade de seda (Rahmathulla, 2012).

A temperatura externa é fator regulador do desenvolvimento e do comportamento dos insetos, influindo na sua alimentação e crescimento. Para *B. mori*, a temperatura ótima para o desenvolvimento ideal é de 25 ºC (Guo et al., 2018). Devido à grande influência da temperatura durante o desenvolvimento deste inseto e o alto índice de mortalidade relacionado às variações térmicas, especialmente em épocas do ano onde as condições de temperatura são adversas, vários artifícios técnicos são utilizados para controlar ou minimizar seus efeitos na criação de lagartas de *B. mori* (Porto, 2012; Guo et al., 2018). Além da escolha do local e época de criação, do material utilizado para construção dos barracões de criação, e até mesmo o tipo e posicionamento das instalações, a estratégia de se criar linhagens ou híbridos de *B. mori* adaptados às diferentes condições ambientais é bastante difundida nos países produtores, tendo em vista a relação direta entre a ocorrência das linhagens com o ambiente de origem (Kumari et al., 2011).

De acordo com a diferença de temperatura apresentada em quatro dos principais países produtores mundiais de seda (China, Japão, Brasil e Índia), o inseto pode apresentar diferença no número de ciclos anuais, sendo monovoltino (um ciclo anual) que usualmente ocorre nas regiões frias, bivoltino (dois ciclos anuais) observado nas regiões subtropicais, e finalmente, polivoltino (vários ciclos anuais) detectado nas regiões tropicais (Matsuoka & Sakamoto, 2018). Os sericicultores brasileiros utilizam híbridos de origem japonesa e chinesa que apresentam ciclo univoltino e bivoltino (Inf. Pessoal - BRATAC). A existência de diversas linhagens em diferentes regiões geográficas e geneticamente melhoradas faz com que *B. mori* apresente não somente diferenças em relação às características mendelianas, mas também nas características quantitativas, tais como o tamanho de corpo, os dias de alimentação, a tolerância térmica e a resistência a doenças (Fujii, 1998).

Algumas doenças comuns e prejudiciais no processo de criação de *B. mori* e produção de seda se desenvolvem em determinadas condições ambientais, tais como temperatura e umidade, as quais coincidem com as utilizadas para a criação das lagartas, o que aumenta a probabilidade de contágio e ocasionando a reprodução e disseminação destes microrganismos (Jiang & Xia, 2014; Guo et al., 2018). Algumas doenças fúngicas (sendo as mais comuns pelos fungos *Beauveria bassiana* e *Metarhizium anisopliae*, causadores das doenças calcinose branca e calcinose preta) são adquiridas pelo contato das lagartas com folhas contaminadas da amoreira. Doenças virais também são comuns, em especial a causada por nucleopolyhedrovirus de *B. mori* (BmNPV), sendo considerada a ameaça mais prevalente para a sericultura em quase todas países (Gomi et al., 1999; Rahman & Gopinathan, 2004; Jiang & Xia, 2014).

2.4 Estudo de transcriptomas e análise de expressão diferencial

Os lepidópteros incluem as pragas agrícolas que mais causam prejuízos, onde centenas de espécies de lagartas ocasionam perdas econômicas generalizadas em plantas de cultivo de alimentos e fibras, florestas e grãos armazenados. No entanto, eles também são importantes indicadores de diversidade e saúde do ecossistema, em especial como polinizadores (Mita et al., 2003). Sendo assim, as informações dos genomas de lepidópteros têm um forte impacto na ciência dos insetos e em indústrias, dentre elas as de inseticidas, controle de pragas e produção de seda.

Bombyx mori, possui também uma importância muito grande na pesquisa para a produção de biomateriais (Altman et al., 2003). Além disso, novos dados genéticos são relevantes para a genética comparativa e a genômica, podendo ajudar a elucidar os caminhos evolutivos conservados e sua diversificação além de identificar novos genes e sistemas de genes como alvos para a transgênese e outras tecnologias para melhoramento (Heckel, 2003), aliados a características produtivas da espécie, como peso do casulo, ao teor da seda e a taxa de mortalidade dos insetos (De Toledo et al., 2002; Porto, 2004), que são informações cruciais para elaboração de estratégias de aumento da produção sericícola.

O transcriptoma é um conjunto completo de transcritos de RNA de uma determinada linhagem celular, tecidos, órgãos ou organismos, e o seu estudo é muito importante para compreender os elementos funcionais do genoma, suas vias biológicas e mecanismos moleculares. A obtenção do transcriptoma se dá pelo sequenciamento do DNA complementar (cDNA) e os fragmentos gerados são denominados *reads* que, em seguida, são alinhados a um genoma ou transcriptoma de referência ou, na ausência de uma referência, é utilizada a estratégia de montagem *de novo* (Riesgo et al., 2012).

A expressão gênica é um processo pelo qual a informação de um gene é processada em um produto gênico, como o RNA e as proteínas. Mensurando a quantidade de RNA é possível distinguir quais genes estão sendo expressos, sendo que essa quantidade varia de a depender dos genes, células, tecidos ou das variações das condições experimentais (Chen et al., 2005). O intuito de uma análise de expressão diferencial é testar se os dados conseguem gerar indícios suficientes para concluir que essa variação de expressão é realmente significativa. Sendo assim, para cada transcrito, um pacote em linguagem R, como o DESeq2, realiza um teste de hipótese para determinar se as evidências são relevantes contra a hipótese nula (H0). Porém, isso é apenas uma probabilidade, o que pode causar falsos negativos sendo descartados pela aleatoriedade dos resultados (Love et al., 2014). Assim, o objetivo final de uma análise de expressão diferencial é destacar apenas os genes que realmente variaram significativamente em abundância nas condições experimentais e podendo também realizar a atribuição biológica desses genes (Li et al., 2010). Para a quantificação e validação dos resultados de uma análise de expressão diferencial é realizada por reação em cadeia da polimerase com transcrição reversa (RT-PCR) (Nonis et al., 2014).

Sendo assim, identificar os transcritos diferencialmente expressos entre *B. mori* expostos a diferentes condições de temperatura é essencial para o entendimento dos

mecanismos de regulação térmica da linhagem comercial brasileira. Uma melhor compreensão da influência da temperatura no processo de produção de seda é fundamental para o aprimoramento da produção de seda e consequente aumento da produção industrial e valorização da atividade rural. Nosso estudo poderá fornecer dados importantes para a elaboração de novas estratégias de criação do inseto e aprimoramento da qualidade de seus casulos na atividade sericícola, fortalecendo ainda mais este setor no Brasil e no mundo.

2.5 Elementos genéticos móveis

Os elementos genéticos móveis ou elementos transponíveis (TEs, do inglês *Transposable elements*) são sequências de DNA que se transportam para outros locais dentro do genoma, criando cópias durante o processo de mobilização (Kazazian, 2004; Drongitis et al., 2019). Esses elementos foram inicialmente descritos nos anos 1940 por Barbara McClintock ao realizar experimentos com milho, e inicialmente denominados por ela como "elementos de controle", em função da hipótese de que estes regulariam a expressão gênica. Esses elementos constituem parte significativa em genomas eucariotos: em *Drosophila melanogaster* (Diptera: Drosophilidae) correspondem a 20% do genoma total (McCullers & Steiniger, 2017), em humanos representam a quase metade do genoma (Fueyo et al., 2022), em *B. mori* representam aproximadamente 46% do genoma total (Kawamoto et al., 2019), podendo ainda corresponder a mais de 80% do genoma de algumas espécies de plantas, a exemplo do milho (Stitzer, 2021).

Os elementos genéticos móveis são classificados de acordo com sua estrutura e mecanismo de transposição em duas classes principais, proposto por Finnegan (1989): (1) os elementos de classe I, ou retroelementos, são TEs cuja fonte de mobilidade é a capacidade de fazer cópias de DNA de seus transcritos de RNA, que são então integrados em novos locais no genoma (mecanismo de transposição "copiar e colar") (2) elementos de classe II, ou elementos do tipo DNA, aqueles que são capazes de manipular diretamente o DNA para se propagar dentro o genoma (mecanismo de transposição "recortar e colar"); em resumo: essa divisão se dá pela presença (1) ou ausência (2) de um intermediário de RNA. A partir de estudos mais aprofundados sobre os TEs e a necessidade de um sistema de classificação ainda mais abrangente, um novo sistema foi proposto por Wicker et al. (2007), onde os TEs passaram a ser classificados em subclasses, ordens, superfamílias, famílias e subfamílias.

Ademais, os TEs também são classificados como elementos autônomos e não autônomos, conforme a sua capacidade de codificar todos os domínios necessários para a transposição.

Partindo desse sistema de classificação, os retrotransposons (classe I) são subdivididos em cinco ordens: LTR, DIRS, PLE, LINE e SINE. Os elementos LTRs (do inglês, *Long Terminal Repeats*) são a ordem predominante em plantas e menos abundantes em animais (Wicker et al., 2007). Evolutivamente, os retrotransposons LRTs e os retrovírus estão intimamente ligados (Frankel & Young, 1998; Seelamgari et al., 2004; Wicker et al., 2007). Esses TEs são flanqueados em ambas as extremidades por repetições terminais longas e apresentam uma fase de leitura aberta (ORF; do inglês, *Open Reading Frame*) com o gene gag, que codifica proteínas estruturais que formam partículas semelhantes às de vírus, e o gene *pol*, que codifica a proteinase aspártica (AP), transcriptase reversa (RT), RNase H (RH) e a integrase (INT) (Wicker et al., 2007). A ordem LTR compreende as superfamílias *Copia*, *Gypsy*, *Bel-Pao*, *Retrovirus* e *ERV* (Figura 4).

Enquanto isso, os retrotransposons da ordem DIRS (do inglês, *Dictyostelium Intermediate Repeat Sequence*) são filogeneticamente relacionados aos *Gypsy*. No entanto, esses elementos codificam uma recombinase de tirosina no lugar da integrasse apresentando regiões terminais incomuns, com repetições diretas divididas. Sendo assim, os elementos DIRS se inserem em um novo local por recombinação e não integração (Cappello et al., 1985; Goodwin & Poulter, 2001; Wicker et al., 2007). Enquanto isso, os elementos da ordem PLE (do inglês, *Penelope-like*), encontrados pela primeira vez em *Drosophila virilis*, codificam uma transcriptase reversa que é mais intimamente relacionada à telomerase do que da transcriptase reversa dos próprios elementos retrotransposons. Ademais, estes elementos codificam uma endonuclease relacionada à proteína de reparo de DNA bacteriano UvrC (Evgen'ev et al., 1997; Wicker et al., 2007). Por sua vez, a ordem LINE (do inglês, *Long Interspersed Nuclear Element*) é composta por elementos que apresentam uma região 5’ UTR (do inglês, *Untranslated Region*) com atividade promotora, uma ou duas ORFs que codificam a transcriptase reversa e uma endonuclease e uma região 3’ UTR com cauda poli-A (Burke et al., 2002; Ostertag & Kazazian, 2005; Wicker et al., 2007).

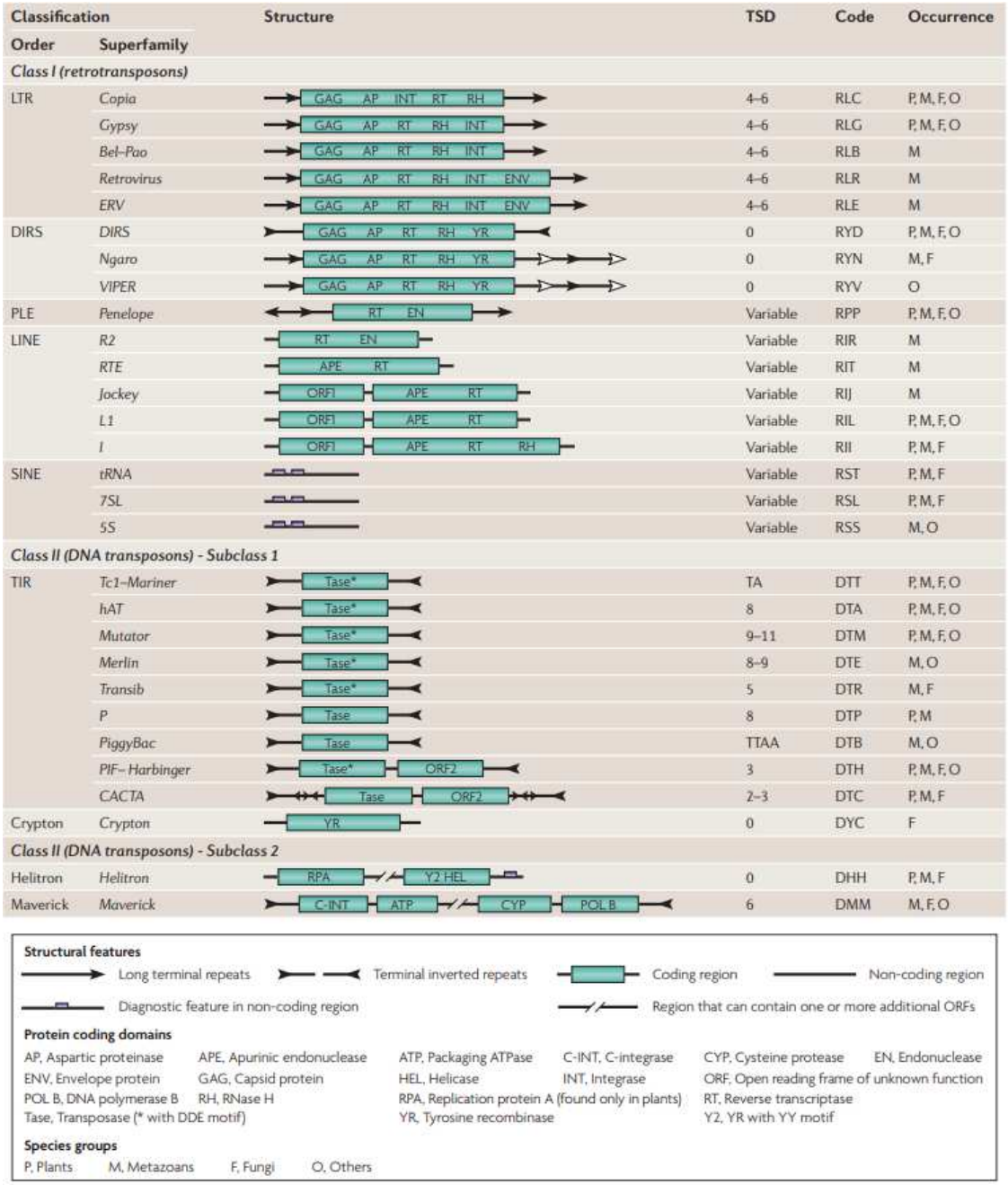

Figura 5. Sistema de classificação para os elementos transponíveis segundo Wicker et al. (2007). A classificação hierárquica divide os elementos em duas principais classes (retrotransposons e DNA transposons) com base na presença ou ausência de RNA como intermediário, respectivamente. Em seguida, são divididos em subclasses, ordens superfamílias.

Os elementos SINEs (do inglês, *Short Interspersed Nuclear Elements*) são considerados elementos não autônomos, pois dependem da maquinaria enzimática de elementos LINEs autônomos para sua transposição. Esses elementos apresentam de 80 a

500 pb, possuindo um promotor de RNA polimerase III interno que permite a sua expressão e regiões 3’ terminais ricas em adenina. A região que abriga o promotor da polimerase III define as superfamílias, sendo elas RNAt, 7SL e 5S (Kramerov & Vassetsky, 2005, Wicker et al., 2007).

Ainda segundo Wicker et al. (2007), os DNA transposons (elementos de classe II) são agrupados em duas subclasses, baseado no número de fitas de DNA que são cortadas durante o processo de mobilização. A subclasse I consiste em TEs que utilizam o mecanismo de “recortar e colar”, com transposição mediada pela transposase (Wicker et al., 2007). Essa subclasse é dividida em duas ordens, sendo a primeira composta por elementos TIRs, caracterizados por suas repetições terminais invertidas de comprimentos variados, e a ordem *Crypton*, formada por elementos encontrados apenas em fungos (Goodwin et al., 2003). A ordem de DNA transposons TIR é subdividida em nove superfamílias: *Tc1-Mariner*, *hAT*, *Mutator*, *Merlin*, *Transib*, *P*, *PiggyBac*, *PIF-Harbinger* e *CACTA*, sendo a superfamília *Tc1-Mariner* amplamente distribuída em eucariotos e compreendendo a maioria dos transposons descritos, apresentando uma estrutura simples com repetições terminais invertidas e uma ORF que codifica a transposase (Shao & Tu, 2001; Wicker et al., 2007). Por sua vez, a subclasse II é representada por elementos em que a transposição envolve a replicação sem clivagem da dupla fita, com o deslocamento de apenas uma delas. Dentro da subclasse II os elementos são classificados nas ordens (1) Helitrons, que se replicam usando um mecanismo círculo-rolante, codificando uma recombinase de tirosina, com um domínio helicase e atividade iniciadora da replicação (Kapitonov & Jurka, 2006), e (2) Polintons (Mavericks), que são transposons de tamanho grande e que apresentam repetições invertidas longas, replicando-se usando uma polimerase de DNA auto-codificada (Pritham et al., 2007).

Com o avanço dos estudos, diversas famílias de TEs têm sido descritas e caracterizadas em diversas espécies devido a sua importância na diversidade genética e em processos evolutivos, além de resposta a estresses bióticos, como competição, predação e parasitismo (Hof et al., 2016; Hou et al., 2019; Bogdan et al., 2020) e abióticos, como choque térmico, danos ao DNA, radiação UV, compostos químicos e condições climáticas (Bradshaw & McEntee, 1989; Capy et al., 2000; Ikeda et al., 2001; El-Sawy et al., 2005; Kale et al., 2005; Stribinkis & Ramos, 2006; Stoycheva et al., 2010; Tanaka et al., 2012; Jardim et al., 2015; De Oliveira et al., 2021). A sua influência na arquitetura e evolução dos genomas (Romano & Fanti, 2022) se dá pelo fato de, além de serem transferidos verticalmente, de pai para filho, estes podem ser transferidos horizontalmente, de uma espécie para outra, podendo contribuir

para a diversidade genética e potenciais efeitos deletérios na nova espécie hospedeira (Wallau et al., 2012). Por exemplo, os TEs que se movem utilizando o mecanismo de "recortar e colar" podem ocasionar mutações de deslocamento do quadro de leitura por inserção em sequências codificantes ou causando pequenas deleções (excisão imprecisa). Ademais, TEs podem ainda induzir inversões cromossômicas ou grandes deleções e duplicações, após recombinação entre diferentes cópias do mesmo elemento (Romano & Fanti, 2022). A inserção de TEs em uma sequência reguladora podem até mesmo ocasionar a superexpressão do gene, ou perturbar essa região, inativando-o (Drongitis et al., 2019).

Em espécies animais, estudos demonstraram a resposta ao estresse ambiental moduladas por TEs. Em *Drosophila*, alterações ambientais podem impulsionar a variabilidade genética, especificamente, junto da complexa regulação de *heat-shock proteins* (HSPs), induzindo um relaxamento na repressão de TEs (Romano & Fanti, 2022). A transcrição do gene Hsp70, que é ativada por fatores de choque térmico (HSFs) em situações de estresse, torna-se rapidamente atenuada quando as condições de estresse diminuem, para mitigar a influência que Hsp70 tem no crescimento e divisão celular (Feder et al., 1992; Morimoto, 1998). Uma vez produzida, a Hsp70 interage com a Hsp90 "distraindo-a" de sua função regulatória de *piwi-RNAs* e permitindo a expressão de TEs (Cappucci et al., 2019). Sendo assim, em *Drosophila*, as condições do estresse, que consequentemente refletem no nível de expressão de Hsp70, modulam a consequente ativação dos TEs. Em humanos, TEs são estudados principalmente devido a sua relação com doenças. Um exemplo são os elementos ERV-LRT, que enriquecem vários sítios de ligação de DNA de p53 (proteína supressora de tumor, responsiva a danos de DNA e estresse celular), afetando a expressão de genes alvo de p53 (Romano & Fanti, 2022).

Apesar de alguns trabalhos já caracterizarem a composição de TEs em diferentes linhagens de *B. mori* (Osanai-Futahashi et al., 2008; The International Silkworm Genome Consortium, 2008; Kawamoto et al., 2019), as informações sobre a relação entre elementos genéticos móveis e a sua atuação na resposta a diferentes condições de temperatura ainda são escassas. Sendo assim, a primeira identificação e caracterização de TEs na linhagem comercial brasileira de *B. mori* e, a análise de expressão diferencial de TEs entre grupos experimentais mantidos em temperaturas distintas podem auxiliar no entendimento do mecanismo de regulação térmica do inseto.

3. OBJETIVOS

3.1 Objetivo geral

O presente trabalho teve por objetivo geral a análise integrada da influência de duas temperaturas de criação na expressão gênica e na produção de casulos de *Bombyx mori*.

3.2 Objetivos específicos

- Obtenção do transcriptoma de lagartas de *B. mori* submetidas a duas temperaturas distintas, 26 ºC e 34 ºC, durante o quinto instar;
- Identificação de genes diferencialmente expressos a fim de associá-los ao metabolismo de regulação térmica e relacioná-los com os dados obtidos em literatura;
- Avaliação dos parâmetros morfológicos das lagartas durante o quinto ínstar e a produção de casulos a 26 ºC e 34 ºC;
- Identificação e caracterização de TEs dentre os transcritos de *B. mori* e sua possível relação com o estresse térmico;
- Contribuição com novas informações que possam auxiliar a sericicultura, principalmente com dados que possam trazer inovações, proporcionando desenvolvimento tecnológico do Brasil, do Estado do Paraná, Londrina e região, como referenciais de produção de seda de alta qualidade.

REFERÊNCIAS

Altman, G. H., Diaz, F., Jakuba, C., Calabro, T., Horan, R. L., Chen, J., Lu, H., Richmond, J. & Kaplan, D. L. (2003). Silk-based biomaterials. *Biomaterials*, *24*(3), 401-416.

De Toledo, J. O. A., Takahashi, R., & Braga Malheiros, E. (2002). Biological development after 5th instar among pure breed and hybrids of silkworm (*Bombyx mori* L.). *Acta Sci. Anim. Sci.*, 869-874.

Babu, K. M. (2015). Natural textile fibres: Animal and silk fibres. In *Textiles and Fashion – Materials, Design and Technology*, 57-78.

Bogdan, L., Barreiro, L., & Bourque, G. (2020). Transposable elements have contributed human regulatory regions that are activated upon bacterial infection. *Philosophical Transactions of the Royal Society B*, *375*(1795), 20190332.

Bradshaw, V. A., & McEntee, K. (1989). DNA damage activates transcription and transposition of yeast Ty retrotransposons. *Molecular and General Genetics MGG*, *218*(3), 465-474.

BRATAC. (2021). Disponível em: http://www.bratac.com.br/bratac/pt/index.php

Burke, W. D., Malik, H. S., Rich, S. M., & Eickbush, T. H. (2002). Ancient lineages of non-LTR retrotransposons in the primitive eukaryote, *Giardia lamblia*. *Molecular biology and evolution*, *19*(5), 619-630.

Cappello, J., Handelsman, K., & Lodish, H. F. (1985). Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. *Cell*, *43*(1), 105-115.

Cappucci, U., Noro, F., Casale, A. M., Fanti, L., Berloco, M., Alagia, A. A., Grassi, L., Pera, L. L., & Pimpinelli, S. (2019). The Hsp70 chaperone is a major player in stress-induced transposable element activation. *Proceedings of the National Academy of Sciences*, *116*(36), 17943-17950.

Capy, P., Gasperi, G., Biémont, C., & Bazin, C. (2000). Stress and transposable elements: co-evolution or useful parasites?. *Heredity*, *85*(2), 101-106.

Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., Barbisin, M., Xu, N.L., Mahuvakar, V.R, Lao, K. Q., Livak, K. J. & Guegler, K. J. (2005). Real-time quantification of microRNAs by stem–loop RT–PCR. *Nucleic acids research*, *33*(20), e179-e179.

Chen, F., Porter, D., & Vollrath, F. (2012). Silk cocoon (*Bombyx mori*): multi-layer structure and mechanical properties. *Acta Biomaterialia*, *8*(7), 2620-2627.

Cirio, G. M. (2018). Evolução e condição atual da sericicultura no Paraná. *Inovações na sericicultura do Paraná: tecnologias, manejo e rentabilidade, 1ª ed. Londrina ABRASEDA: IAPAR*, 21-32.

CONAB - Companhia Nacional de Abastecimento. Indicadores da Agropecuária. Observatório Agrícola. (2016). ISSN: 2317-7535.

DERAL - Departamento de Economia Rural do Paraná, SEAB - Secretaria da Agricultura e do Abastecimento. (2018). Relatório TAKII. Sericicultura no Estado do Paraná. Safra 2017/2018. Disponível em: https://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/documento/2019-09/sericicultura_2019_v1.pdf

De Oliveira, D. S., Rosa, M. T., Vieira, C., & Loreto, E. L. (2021). Oxidative and radiation stress induces transposable element transcription in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, *34*(4), 628-638.

Dong, H. L., Zhang, S. X., Tao, H., Chen, Z. H., Li, X., Qiu, J. F., Cui, W. Z., Sima, Y. S., Cui, W. Z. & Xu, S. Q. (2017). Metabolomics differences between silkworms (*Bombyx mori*) reared on fresh mulberry (*Morus*) leaves or artificial diets. *Scientific reports*, *7*(1), 1-16.

Drongitis, D., Aniello, F., Fucci, L., & Donizetti, A. (2019). Roles of transposable elements in the different layers of gene expression regulation. *International Journal of Molecular Sciences*, *20*(22), 5755.

El-Sawy, M., Kale, S. P., Dugan, C., Nguyen, T. Q., Belancio, V., Bruch, H., Roy-Engel, A. M., & Deininger, P. L. (2005). Nickel stimulates L1 retrotransposition by a post-transcriptional mechanism. *Journal of molecular biology*, *354*(2), 246-257.

Emery, C. (2019). 10 Interesting Silkworm Facts. Everything Silkworms. Disponível em: https://everythingsilkworms.com.au/10-interesting-silkworm-facts-will-blow-you-away/

Evgen'ev, M. B., Zelentsova, H., Shostak, N., Kozitsina, M., Barskyi, V., Lankenau, D. H., & Corces, V. G. (1997). Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proceedings of the National Academy of Sciences*, *94*(1), 196-201.

Finnegan, D. J. (1989). Eukaryotic transposable elements and genome evolution. *Trends in genetics*, *5*, 103-107.

Feder, J. H., Rossi, J. M., Solomon, J., Solomon, N., & Lindquist, S. (1992). The consequences of expressing hsp70 in *Drosophila* cells at normal temperatures. *Genes & development*, *6*(8), 1402-1413.

Frankel, A. D., & Young, J. A. (1998). HIV-1: fifteen proteins and an RNA. *Annual review of biochemistry*, *67*, 1.

Fueyo, R., Judd, J., Feschotte, C., & Wysocka, J. (2022). Roles of transposable elements in the regulation of mammalian transcription. *Nature Reviews Molecular Cell Biology*, 1-17.

Fujii, H. (1998). Genetical stocks and mutations of *Bombyx mori*: *Important genetic resources*. Institute of Genetic Resources, Faculty of Agriculture, Kyushu University, Fukuoka, Japan.

Goldsmith, M. R., Shimada, T., & Abe, H. (2005). The genetics and genomics of the silkworm, *Bombyx mori. Annual Review of Entomology*, *50*, 71-100.

Goodwin, T. J., & Poulter, R. T. (2001). The DIRS1 group of retrotransposons. *Molecular biology and evolution*, *18*(11), 2067-2082.

Goodwin, T. J., Butler, M. I., & Poulter, R. T. (2003). Cryptons: a group of tyrosine-recombinase-encoding DNA transposons from pathogenic fungi. *Microbiology*, *149*(11), 3099-3109.

Guan, J., Zhu, W., Liu, B., Yang, K., Vollrath, F., & Xu, J. (2017). Comparing the microstructure and mechanical properties of *Bombyx mori* and *Antheraea pernyi* cocoon composites. *Acta biomaterialia*, *47*, 60-70.

Gomi, S., Majima, K., Maeda, S., & Efstathiou, S. (1999). Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. *Journal of general virology*, *80*(1), 1323-1337.

Guo, H., Huang, C., Jiang, L., Cheng, T., Feng, T., & Xia, Q. (2018). Transcriptome analysis of the response of silkworm to drastic changes in ambient temperature. *Applied microbiology and biotechnology*, *102*(23), 10161-10170.

Hanada, Y., & Watanabe, J. K. (1986). *Manual de criação do bicho da seda*. Associação de Crédito e Assistência Rural do Paraná, Paraná (Brasil). 224p.

Heckel, D. G. (2003). Genomics in pure and applied entomology. *Annual Review of Entomology*, *48*(1), 235-260.

Hof, A. E., Campagne, P., Rigden, D. J., Yung, C. J., Lingley, J., Quail, M. A., Hall, N., Darby, A. C., & Saccheri, I. J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature*, *534*(7605), 102-105.

Hou, J., Lu, D., Mason, A. S., Li, B., Xiao, M., An, S., & Fu, D. (2019). Non-coding RNAs and transposable elements in plant genomes: emergence, regulatory mechanisms and roles in plant development and stress responses. *Planta*, *250*(1), 23-40.

Howrelia, J. H., Patnaik, B. B., Selvanayagam, M., & Rajakumar, S. (2011). Impact of temperature on heat shock protein expression of *Bombyx mori* crossbreed and effect on commercial traits. *Journal of Environmental Biology*, *32*(1), 99-103.

Hussain, M., Naeem, M., Khan, S. A., Bhatti, M. F., & Munawar, M. (2011). Studies on the influence of temperature and humidity on biological traits of silkworm (*Bombyx mori* L.; Bombycidae). *African Journal of Biotechnology*, *10*(57), 12368-12375.

IBGE – Instituto Brasileiro de Geografia e Estatística. (2018). Disponível em https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-denoticias/noticias/20901-a-arte-da-sericicultura-ganha-espaco-no-pais.html

Paraná – Instituto de Desenvolvimento Rural do Paraná. (2022). Disponível em: http://www.idrparana.pr.gov.br/Pagina/Sericicultura

Ikeda, K., Nakayashiki, H., Takagi, M., Tosa, Y., & Mayama, S. (2001). Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. *Molecular Genetics and Genomics*, *266*(2), 318-325.

Inserco. (2023). Statistics. Disponível em: https://inserco.org/en/statistics

International Silkworm Genome Consortium. (2008). The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect biochemistry and molecular biology*, *38*(12), 1036-1045.

Jardim, S. S., Schuch, A. P., Pereira, C. M., & Loreto, E. L. S. (2015). Effects of heat and UV radiation on the mobilization of transposon mariner-Mos1. *Cell Stress and Chaperones*, *20*(5), 843-851.

Jiang, L., & Xia, Q. (2014). The progress and future of enhancing antiviral capacity by transgenic technology in the silkworm *Bombyx mori*. *Insect biochemistry and molecular biology*, *48*, 1-7.

Júnior, D. S., Almeida, E. L. D., & Pádua, O. S. (2018). Inovações na sericicultura do Paraná: tecnologias, manejo e rentabilidade. ABRASEDA, IAPAR, Londrina, Brasil.

Kale, S. P., Moore, L., Deininger, P. L., & Roy-Engel, A. M. (2005). Heavy metals stimulate human LINE-1 retrotransposition. *International journal of environmental research and public health*, *2*(1), 14-23.

Kapitonov, V. V., & Jurka, J. (2006). Self-synthesizing DNA transposons in eukaryotes. *Proceedings of the National Academy of Sciences*, *103*(12), 4540-4545.

Kawamoto, M., Jouraku, A., Toyoda, A., Yokoi, K., Minakuchi, Y., Katsuma, S., Fujiyama, A., Kiuchi, T., Yamamoto, K., & Shimada, T. (2019). High-quality genome assembly of the silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, *107*, 53-62.

Kazazian Jr, H. H. (2004). Mobile elements: drivers of genome evolution. *Science*, 303(5664), 1626-32.

Kramerov, D. A., & Vassetzky, N. S. (2005). Short retroposons in eukaryotic genomes. *International review of cytology*, *247*, 165-221.

Kumari, S. S., Subbarao, S. V., Misra, S., & Murtyd, U. S. (2011). Screening strains of the mulberry silkworm, Bombyx mori, for thermotolerance. *Journal of Insect Science*, *11*(1).

Li, H. & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, *26*(5), 589-595.

Li, J., Ye, L., Lan, T., Yu, M., Liang, J., & Zhong, B. (2012). Comparative proteomic and phosphoproteomic analysis of the silkworm (*Bombyx mori*) posterior silk gland under high temperature treatment. *Molecular biology reports*, *39*(8), 8447-8456.

Liu, T. L., Miao, J. C., Sheng, W. H., Xie, Y. F., Huang, Q., Shan, Y. B., & Yang, J. C. (2010). Cytocompatibility of regenerated silk fibroin film: a medical biomaterial applicable to wound healing. *Journal of Zhejiang University SCIENCE B*, *11*, 10-16.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 1-21.

Matsuoka, D., & Sakamoto, K. (2018). Effects of mild and low temperature incubation on heat tolerance in *Bombyx mori* embryos. *American Journal of Entomology*, 2, 6-9.

McCullers, T. J., & Steiniger, M. (2017). Transposable elements in *Drosophila*. *Mobile genetic elements*, *7*(3), 1-18.

Mita, K., Morimyo, M., Okano, K., Koike, Y., Nohata, J., Kawasaki, H., Okuda, K. K., Yamamoto, K., Suzuki, M. G., Shimada, T., Goldsmith, M. R. & Maeda, S. (2003). The construction of an EST database for *Bombyx mori* and its application. *Proceedings of the National Academy of Sciences*, *100*(24), 14121-14126.

Morimoto, R. I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes & development*, *12*(24), 3788-3796.

Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C. N., Richardson, A. O., Okumoto, Y., Tanisaka, T., & Wessler, S. R. (2009). Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature*, *461*(7267), 1130-1134.

Nicodemo, D., Oliveira, J. E., Sedano, A. A., Marconcini, J. M., & Tonoli, G. H. (2014). Impact of different silkworm dietary supplements on its silk performance. *Journal of materials science*, *49*(18), 6302-6310.

Nonis, A., De Nardi, B., & Nonis, A. (2014). Choosing between RT-qPCR and RNA-seq: a back-of-the-envelope estimate towards the definition of the break-even-point. *Analytical and Bioanalytical Chemistry*, 406, 3533–3536.

Osanai-Futahashi, M., Suetsugu, Y., Mita, K., & Fujiwara, H. (2008). Genome-wide screening and characterization of transposable elements and their distribution analysis in the silkworm, Bombyx mori. *Insect biochemistry and molecular biology*, *38*(12), 1046-1057.

Ostertag, E. M., & Kazazian, H. H. (2005). LINEs in mind. *Nature*, *435*(7044), 890-891.

Porto, A. J. (2004). Avaliação da heterose em cruzamentos do bicho-da-seda (*Bombyx mori* L.). *Boletim de Indústria Animal*, *61*(2), 135-140.

Porto, A. J. (2012). Desempenho de raças de *Bombyx mori* L. sob condições limites de temperatura. *Semina: Ciências Agrárias*, *33*(4), 1263-1271.

Pritham, E. J., Putliwala, T., & Feschotte, C. (2007). Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene*, *390*(1-2), 3-17.

Rahman, M. M., & Gopinathan, K. P. (2004). Systemic and in vitro infection process of *Bombyx mori* nucleopolyhedrovirus. *Virus research*, *101*(2), 109-118.

Rahmathulla, V. K. (2012). Management of climatic factors for successful silkworm (*Bombyx mori* L.) crop and higher silk production: a review. *Psyche*, *2012*.

Riesgo, A., Andrade, S., Sharma, P. P., Novo, M., Pérez-Porro, A. R., Vahtera, V., González, V. L., Kawauchi, G. Y & Giribet, G. (2012). Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. *Frontiers in zoology*, 9(1), 1-24.

Romano, N. C., & Fanti, L. (2022). Transposable Elements: Major Players in Shaping Genomic and Evolutionary Patterns. *Cells*, *11*(6), 1048.

Roy, M., Meena, S. K., Kusurkar, T. S., Singh, S. K., Sethy, N. K., Bhargava, K., Sarkar, S. & Das, M. (2012). Carbondioxide gating in silk cocoon. *Biointerphases*, 7(1), 45.

SEAB - Secretaria da Agricultura e do Abastecimento. (2020). Valor bruto da produção. Disponível em: https://www.agricultura.pr.gov.br/vbp

Secco, D., Wang, C., Shou, H., Schultz, M. D., Chiarenza, S., Nussaume, L., Ecker, J. R., Whelan, J., & Lister, R. (2015). Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *elife*, *4*, e09343.

Shao, H., & Tu, Z. (2001). Expanding the diversity of the IS630-Tc1-mariner superfamily: discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons. *Genetics*, *159*(3), 1103-1115.

Seelamgari, A., Maddukuri, A., Berro, R., de la Fuente, C., Kehn, K., Deng, L., Dadgar, S., Bottazzi, M. E., Ghedin, E., Pumfery, A., & Kashanchi, F. (2004). Role of viral regulatory and accessory proteins in HIV-1 replication. *Frontiers in Bioscience-Landmark*, *9*(4), 2388-2413.

Stitzer, M. C., Anderson, S. N., Springer, N. M., & Ross-Ibarra, J. (2021). The genomic ecosystem of transposable elements in maize. *PLoS genetics*, *17*(10), e1009768.

Stoycheva, T., Pesheva, M., & Venkov, P. (2010). The role of reactive oxygen species in the induction of Ty1 retrotransposition in *Saccharomyces cerevisiae*. *Yeast*, *27*(5), 259-267.

Stribinskis, V., & Ramos, K. S. (2006). Activation of human long interspersed nuclear element 1 retrotransposition by benzo (a) pyrene, an ubiquitous environmental carcinogen. *Cancer research*, *66*(5), 2616-2620.

Takasu, Y., Tamura, T., Goldsmith, M., & Zurovec, M. (2016). Targeted mutagenesis in *Bombyx mori* using TALENs. In *TALENs* (pp. 127-142). *Humana Press*, New York, NY.

Tanaka, A., Nakatani, Y., Hamada, N., Jinno-Oue, A., Shimizu, N., Wada, S., Funayama, T., Mori, T., Islam, S., Hoque, S. A., Shinagawa, M., Ohtsuki, T., Kobayashi, Y., & Hoshino, H. (2012). Ionising irradiation alters the dynamics of human long interspersed nuclear elements 1 (LINE1) retrotransposon. *Mutagenesis*, *27*(5), 599-607.

Wallau, G. L., Ortiz, M. F., & Loreto, E. L. S. (2012). Horizontal transposon transfer in eukarya: detection, bias, and perspectives. *Genome Biology and Evolution*, *4*(8), 801-811.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, *8*(12), 973-982.

Zhang, J., Rajkhowa, R., Li, J. L., Liu, X. Y., & Wang, X. G. (2013). Silkworm cocoon as natural material and structure for thermal insulation. *Materials & Design*, *49*, 842-849.

CAPÍTULO I

Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae)

O manuscrito será submetido ao periódico *Gene*.

Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae)

Thayná Bisson Ferraz Lopes[1], Rachel Colauto Milanezi Aguiar[1], Rogério Fernandes de Souza[1], Cristianne Cordeiro Nascimento[2], Jaqueline Fernanda Dionísio[1], Mario Sergio Mantovani[1], Simone Cristine Semprebon[1], Renata da Rosa[1]

[1] Departamento de Biologia Geral, Universidade Estadual de Londrina, Londrina, Paraná, Brasil

[2] Departamento de Design Gráfico, Centro de Educação, Comunicação e Artes, Universidade Estadual de Londrina, Londrina, Paraná, Brasil

Resumo

O bicho-da-seda (*Bombyx mori*) é um lepidóptero economicamente importante para a produção de seda em todo o mundo. No entanto, grandes perdas econômicas são ocasionadas por fatores que afetam diretamente o rendimento e a qualidade da produção dos casulos de seda, como doenças e variações de temperatura. Compreender seus mecanismos de resposta às temperaturas de criação durante o período mais crítico do seu ciclo de vida, o quinto instar, pode fornecer informações sobre sua adaptação ao ambiente e melhoramento da cadeia sericícola. Neste estudo, realizamos a análise transcricional de dois grupos de *B. mori* criados em temperaturas de 26 ºC e 34 ºC durante todo o quinto instar, e medição do peso das lagartas e dos casulos produzidos nas duas temperaturas. Das 29.157 sequências obtidas na montagem do transcriptoma, 3.115 transcritos foram diferencialmente expressos (DE). Destes, 1.696 foram *down* regulados e 1.419 *up* regulados nas lagartas criadas a 34 ºC em relação às criadas a 26 ºC. Dentre os transcritos DE em resposta às condições temperatura do presente estudo destacamos os genes envolvidos na resposta imune, fatores de transcrição, genes do mecanismo de produção de seda e proteínas de choque térmico. As lagartas que foram mantidas a 34 ºC apresentaram média menor de peso corporal, afetando diretamente o peso dos casulos produzidos e a taxa de conversão de seda. Essas diferenças em comparação com as lagartas mantidas a 26 °C foram estatisticamente significativas. As taxas de mortalidade (6 e 9%, a 26 °C e 34 ºC, respectivamente) foram semelhantes às obtidas em campos de criação. Os resultados fornecem informações sobre os processos biológicos envolvidos no mecanismo de resposta a temperatura das lagartas de *B. mori* e nos trazem informações que podem futuramente auxiliar nos processos de climatização durante a criação das larvas em campo.

Palavras-chave: Expressão diferencial; RNA-Seq; estresse térmico, transcriptoma.

1. Introdução

Os insetos são animais que têm seu crescimento, desenvolvimento, sobrevivência, comportamento e reprodução diretamente afetados pela temperatura externa (Xiao et al., 2017). Considerado um inseto modelo, o bicho-da-seda domesticado *Bombyx mori* Linnaeus, 1758 é um lepidóptero que produz um casulo durante o estágio final do desenvolvimento larval como um mecanismo de sobrevivência a fatores externos, entre eles a temperatura. O casulo é formado pelos fios de seda, compostos por duas proteínas principais, fibroína e sericina, sendo amplamente utilizados na indústria para a confecção de tecidos, e recentemente também na produção de biomateriais (Sahi et al., 2023; Yang et al., 2023).

Alguns fatores afetam diretamente o rendimento e a qualidade da produção dos casulos de seda, como doenças e variações de temperatura, que são intimamente relacionados e ocasionam grandes perdas econômicas (Jiang & Xia, 2014; Guo et al, 2018). A temperatura elevada pode afetar principalmente o desenvolvimento do inseto, e influir na disseminação e na incubação de vírus e fungos na atividade sericícola. “Flacherie” é um termo em francês comumente utilizado para descrever essas duas situações, se referindo às lagartas de *B. mori* que apresentam um aspecto flácido devido à exposição a temperaturas excessivamente altas, e que tem como consequência a morte desses insetos em grande escala (James & Li, 2012).

O Brasil, atualmente o 6º produtor de casulos de seda no mundo (Inserco, 2023), é um país tropical, com temperaturas elevadas especialmente durante os meses de verão intenso. Essas condições dificultam o processo produtivo de seda de *B. mori* não somente da linhagem comercial brasileira, mas também mundial. Diversos países produtores de casulos de seda estão localizados em regiões tropicais e são afetados pelas condições extremas. Sendo assim, um melhor entendimento da resposta do inseto à temperatura é essencial para compreender os mecanismos envolvidos no processo e obter novas informações sobre a expressão e regulação gênica de *B. mori* utilizado na produção de seda.

Diante disso, neste trabalho foram avaliados o transcriptoma e a expressão gênica diferencial entre lagartas mantidas em duas temperaturas diferentes. Adicionalmente, dados sobre a morfologia das lagartas e confecção de casulos também foram avaliados. Os resultados podem fornecer dados para elaboração de estratégias de climatização de locais de produção, amenizando os efeitos da temperatura e elevando a produção de seda no Brasil e no mundo.

2. Material e Métodos

2.1 Criação das lagartas

O presente estudo está cadastrado no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SisGen) sob registro A368283. Os métodos e ferramentas usadas no presente trabalho foram resumidas na Figura 1.

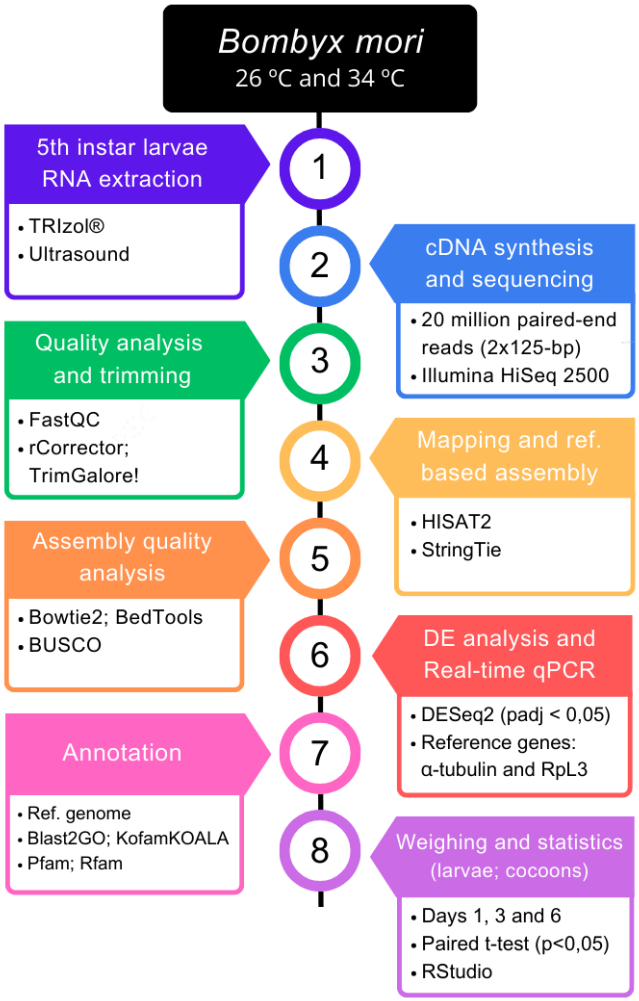

Figura 1. *Pipeline* aplicada na análise da expressão gênica e da produção de casulos de *Bombyx mori*. Abreviação: DE = Expressão diferencial.

O cruzamento das mariposas, a incubação dos ovos e o nascimento das larvas de *B. mori* (linhagem comercial) foram promovidos e monitorados pela Fiação de Seda Bratac, localizada em Londrina-PR. Foram fornecidas 1.000 larvas da linhagem comercial (híbrido: ♂ japonês x ♀ chinesa), de mesma idade (4º instar), peso e origem parental. As larvas de cada

grupo experimental (N=500 / grupo) foram mantidas em quadruplicadas para cada grupo em bandejas plásticas (60cm x 40cm x 15cm) e em sala de criação com temperatura, umidade e fotoperíodo controlados sob as seguintes condições: 26 ºC, ±75% UR e 14h claro/10h escuro, sendo alimentadas com folhas frescas de amoreira (*Morus alba* L.) cinco vezes ao dia com quantidades recomendadas por Takahashi et al. (2009). A partir do primeiro dia do quinto estágio de desenvolvimento larval, duas temperaturas de criação foram testadas, uma para cada grupo experimental: (1) 26 ºC, e (2) 34 ºC. O quinto instar é o período mais crítico da vida das lagartas, onde elas mais se alimentam e se preparam para a confecção do casulo, e tem duração de aproximadamente sete dias. Ao final do quinto instar, três lagartas de cada grupo experimental foram selecionadas aleatoriamente e imediatamente congeladas e mantidas em freezer -80°C para posterior extração de RNA. As lagartas remanescentes continuaram sendo mantidas nas condições previamente descritas até a finalização da confecção dos casulos para análise morfológica.

2.2 Extração de RNA e sequenciamento

Três lagartas de cada grupo foram selecionadas para extração de RNA e sequenciamento. Os indivíduos criados a 26 ºC foram identificados como BM26A, BM26B e BM26C, e, os criados a 34 ºC, como BM34A, BM34C e BM34C. A extração de RNA foi realizada utilizando 1mL do reagente TRIzol® (Invitrogen) por amostra, e posteriormente submetidas a 12 minutos em gelo no aparelho de ultrassom (50-60Hz – Cristófoli Biossegurança, Brasil). Foi adicionado 0,2mL de clorofórmio e após incubação de três minutos, 400 µL de sobrenadante foi retirado e 400 µL de etanol 70% foi adicionado. A purificação seguiu o protocolo descrito pelo fabricante, utilizando PureLink RNA mini kit (Invitrogen). A integridade do RNA extraído foi avaliada através de eletroforese com gel de agarose 1% e a qualidade das amostras foi estimada através do espectrofotômetro BiodropTM (Biochrom).

As amostras de RNA total (três réplicas biológicas por tratamento) foram submetidas para preparação de bibliotecas de cDNA usando protocolo *dUTP-stranded* e um mínimo de 20 milhões de *paired-end reads* (2x125-bp) por amostra foram sequenciados usando a tecnologia Illumina HiSeq 2500. O processamento dos dados brutos foi feito utilizando o Illumina CASAVA pipeline, para obtenção dos dados de sequencias no formato FASTQ.

2.3 Montagem do transcriptoma e análise de expressão diferencial

Os parâmetros de qualidade das leituras geradas após o sequenciamento para cada biblioteca foram avaliados com o auxílio da ferramenta FastQC v0.11.3 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). A correção dos *reads* raros foi realizada no programa rCorrector v1.0.4 (https://github.com/mourisl/Rcorrector). Os *reads* não corrigidos foram retirados com o *script* FilterUncorrectabledPEfastq.py (https://github.com/harvardinformatics/TranscriptomeAssemblyTools/blob/master/FilterUncorrectabledPEfastq.py). A limpeza das bibliotecas para a remoção de pares de bases com baixa qualidade, sequências adaptadoras adicionadas durante o sequenciamento e fragmentos curtos foi realizada com o programa TrimGalore! v0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).

Os transcriptomas das seis bibliotecas sequenciadas foram alinhadas (alinhamento *splice-aware*) ao genoma de referência disponível no NCBI (Ref Seq GCF_014905235_1) com programa HISAT2 v2.1.0 (Kim et al., 2019) e a montagem guiada realizada com o programa StringTie v2.1.1 (http://ccb.jhu.edu/software/stringtie/). O mapeamento foi realizado a fim de verificar quais *reads* foram usados na montagem utilizando o software Bowtie2 v2.3.5 (http://bowtiebio.sourceforge.net/bowtie2/index.shtml) (Langmead e Salzberg, 2012) e BedTools v2.30.0 (https://bedtools.readthedocs.io/). A qualidade da montagem também foi avaliada por meio do banco de dados BUSCO v.5.3.2 (https://busco-data.ezlab.org/v4/data/lineages/), comparando com os conjuntos de dados insecta_odb10 e lepidoptera_odb10.

A análise de expressão gênica diferencial entre os grupos experimentais foi avaliada com o pacote DESeq2 v1.26.0 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) (Love et al., 2014), implementados no programa RStudio v2021.09.0 Build 351 (https://rstudio.com/) em linguagem R. Os transcritos diferencialmente expressos, que apresentaram uma probabilidade ajustada de p menor que o limite alfa de 5% (padj < 0,05) foram isolados com a ferramenta Samtools v1.11 e convertidos em sequências peptídicas com o programa CodAn v1.2. A ferramenta *blastx* foi utilizada para a identificação das proteínas a partir dos transcritos anotados do genoma, o mesmo utilizado na montagem, também disponíveis no NCBI (RefSeq GCF_014905235_1). As mesmas sequências nucleotídicas também foram submetidas a anotação automática no software Blast2GO v6.0.3 (https://www.blast2go.com/) (Conesa et

al., 2005), com um valor de corte de 10^{-3} por meio da busca dos termos do Gene Ontology (GO) associados aos *hits* encontrados para cada sequência após o BLAST. As sequências peptídicas foram submetidas ao KofamKOALA (https://www.genome.jp/tools/kofamkoala/) para identificação das vias metabólicas associadas disponibilizadas pelo KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa e Goto, 2000; Kanehisa et al., 2016; Aramaki et al., 2019). Os transcritos tiveram a sua classificação em famílias de proteínas, bem como a predição de domínios a partir de bancos do InterPro, dentre eles o Pfam (https://www.ebi.ac.uk/interpro/) e também pesquisadas em banco de dados de famílias de RNA, Rfam (https://rfam.org/).

2.4 Validação dos genes diferencialmente expressos (RT-qPCR)

A síntese de cDNA foi realizada em termociclador T100TM (Bio-Rad), utilizando 12,4µl de RNA total (1000 ng / mL), 1µl de oligo dT (20 pmol / mL) e 2µl de dNTPs (2,5 mM) para cada amostra. Em seguida, as amostras foram submetidas ao termociclador por 10 minutos a 65 °C. Imediatamente após, as amostras foram resfriadas no gelo e, em seguida, foi adicionado 4µl de MLV 5x *First Strand Buffer* (M-MLV - Invitrogen, Life Technologies) e incubadas por 2 minutos no gelo. Após incubação, foi adicionado 0,1µl de RNA*out* (Invitrogen, Life Technologies) e 0,5µl da enzima transcriptase reversa (M-MLV-RT, Invitrogen, Life Technologies). Na sequência, as amostras foram submetidas ao termociclador por 50 minutos a 37° C e 15 minutos a 70ºC. O cDNA foi armazenado em -20ºC.

A reação de qRT-PCR foi realizada em termociclador LightCycler®Nano (Roche) utilizando 5µl de Sybr *Green Super Mix* (Invitrogen), 0,5µl de cada par de oligonucleotídeos e 5 µl de cDNA de cada amostra, nas seguintes condições: 95 ° C por 2 minutos e 40 ciclos [95°C / 15s, 60°C / 15s, 72°C / 60s]. A análise da curva de *melting* foi usada para analisar a especificidade dos produtos da qRT-PCR. As reações foram realizadas em duplicatas para cada uma das amostras experimentais.

A sequência dos *primers* utilizados na qRT-PCR está descrita na (Tabela 1). Os *primers* para amplificação da sericina-1 (Ser1; RefSeq: NM_001044041) e da 19.5 kDa *heat shock protein* mRNA (Hsp19.5) (GenBank: GQ884144.1) foram desenhados utilizando o *software* PrimerPlus (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi), a partir de sequências previamente depositadas no NCBI e GenBank, Em adição, os dados da qRT-PCR foram normalizados pelos valores da expressão dos genes de referência α-tubulina

e RpL3 (Fang et al. (2015) e analisados no software Rest (Pfaffl et al., 2002), com nível de significância fixado em $p<0,05$.

Tabela 1. *Primers* utilizados na qRT-PCR.

Gene	Produto (pb)	Referência	Sequência (5'-3')
α-tubulin	146	Fang et al. (2015)	ACATGGCTTGCTGTATGCT GGGTGGCTGGTAGTTGATA
RpL3	161	Fang et al. (2015)	CGGTGTTGTTGGATACATTGAG GCTCATCCTGCCATTTCTTACT
FibL	155	Hu et al. (2018)	CCGGAGGTGGAAGAATCTAT GGTTATGTAGGCAGCGATGT
FibH	150	Naik et al. (2022)	CAGGGGATACGGACAAGGTG ACACAAGGCAGTGCTCTGAA
Ser1	192	Desenhado pelo autor no *software* PrimerPlus	CTACCGTACAGTCATCCAC GTCTGTTCACGCCTGAATG
Ser2	520	Takasu et al. (2010)	GCTCTAGACACTTTTCGGGGGCTTAGTT CGCAAGCTTTTCGTATTCGGAACCTTTGC
Hsp19.5	200	Desenhado pelo autor no *software* PrimerPlus	ACAGTCATGGCATCGAAGGG GAAGTGACTCCTCGTCCAGC
Hsp70	400	Velu et al. (2008)	GAACACACTCGCTGCACATC GAGGAGTGCCCAAGATCGAC

2.5 Avaliação morfológica das lagartas e medição dos casulos produzidos

Um total de 100 lagartas foram selecionadas aleatoriamente e pesadas individualmente a cada dois dias durante o quinto instar (três medições foram realizadas: dia 1, dia 3 e dia 6) para avaliação da influência da temperatura no desenvolvimento. Após a confecção dos casulos pelas lagartas remanescentes, um total de 100 casulos foram selecionados aleatoriamente de cada grupo experimental para a medição do peso dos casulos e das cascas. Para a medição do peso das cascas, foram descartadas as crisálidas no interior dos casulos. A taxa de conversão de seda foi estabelecida a partir do cálculo: (peso da casca / peso do casulo total) x 100. Os dados foram expressos em média e desvio padrão. Os dados de peso larval, de peso dos casulos e peso da casca foram avaliados por meio do teste T pareado. Foram considerados como nível de significância os valores de $p<0,05$. Todos os dados morfológicos foram analisados no RStudio v2021.09.0 Build 351 (https://rstudio.com/) em linguagem R.

3. Resultados

3.1 Sequenciamento e montagem do transcriptoma

O sequenciamento de bibliotecas de cDNA gerou 208.626.156 *reads paired-end* de aproximadamente 125 pb, em média. Cerca de 0,8% das leituras obtidas foram descartadas após a remoção das sequências e adaptadores de baixa qualidade (Tabela 2).

Tabela 2. Resumo das estatísticas de sequenciamento das seis bibliotecas *paired-end* de *B. mori*.

Bibliotecas	***Paired-end***	***Reads* totais**	***Reads* trimados**
BM26A	R1	22.544.321	22.385.786
	R2	22.544.321	22.385.786
BM26B	R1	22.778.091	22.614.954
	R2	22.778.091	22.614.954
BM26C	R1	12.727.035	12.616.251
	R2	12.727.035	12.616.251
BM34A	R1	13.198.100	13.112.247
	R2	13.198.100	13.112.247
BM34B	R1	15.633.957	15.459.660
	R2	15.633.957	15.459.660
BM34C	R1	17.431.574	17.311.651
	R2	17.431.574	17.311.651
Total		208.626.156	207.001.098
***Reads* removidos**			1.625.058 (0,8%)

A montagem do transcriptoma guiada por um genoma referência foi realizada com os 207.001.098 *reads* restantes, resultando em 29.157 transcritos. A taxa de alinhamento geral dos *reads* de todas as bibliotecas de RNA-Seq utilizadas para a montagem do transcriptoma foi de 90,70%. A comparação do transcriptoma com o banco de dados de Lepidoptera usando a ferramenta BUSCO identificou 82,9% de sequências completas, sendo 50,14% cópia única e 32,8% duplicadas. Além disso, 2,3% das sequências estavam fragmentadas e 14,8% eram faltantes. A comparação com o banco de dados de Insecta disponíveis também foi realizada (Tabela 3).

Tabela 3. Avaliação da qualidade da montagem do transcriptoma de *B. mori* em comparação com bancos de dados (*_odb10) de Lepidoptera e Insecta por meio da ferramenta BUSCO.

BUSCOs	lepidoptera_odb10	insecta_odb10
Completos	4382 (82,9%)	1269 (92,8%)
Completos e de cópia única	2647 (50,1%)	774 (56,6%)
Completos e duplicados	1735 (32,8%)	495 (36,2%)
Fragmentados	119 (2,3%)	36 (2,6%)
Faltantes	785 (14,8%)	62 (4,6%)
Total	5286 (100%)	1367 (100%)

3.2 Análise de expressão diferencial e anotação funcional

Após a eliminação dos transcritos redundantes, histonas e RNAr, o transcriptoma foi reduzido a 28.683 transcritos (Tabela Suplementar 1), submetidos a análise de expressão diferencial e resultando em 3.115 sequências diferencialmente expressas (padj < 0,05; Log2FoldChange ≥ 2,0). Destas, 1.696 foram *down* reguladas e 1.419 *up* reguladas nas lagartas criadas a 34 ºC em relação às criadas a 26 ºC (Tabela Suplementar 2).

Bombyx mori Linnaeus, 1758 e *Bombyx mandarina (*Moore*,* 1872) apresentaram-se no topo dos *hits* de similaridade com os dados analisados. Destes, 2.753 (88,4%) apresentaram identificação com sequências no banco de dados do NCBI, sendo que 2.720 (87,3%) sequências obtiveram *hits,* 1.872 (60,1%) foram mapeadas pelo Gene Ontology (GO) e 1.405 (45,1%) tiveram anotação automática de GOs.

A anotação funcional contra o banco de dados GO dos transcritos diferencialmente expressos (DE) foi realizada com base em sua semelhança com as sequências com funções previamente conhecidas. Os termos GO foram resumidos nas três categorias principais: processo biológico, componente celular e função molecular. Entre estes, a função molecular apresentou o maior número de termos GO anotados, com 2.186 termos, seguido por processo biológico, com 1.644 termos, e componente celular que apresentou

1.286 termos anotados. A descrição de cada sequência e cada anotação funcional foram listadas na Tabelas Suplementar S2. A distribuição das sequências em cada termo GO estão listadas nas Tabelas Suplementares 3a e 3b.

Dentre os transcritos DE *up* regulados (2.494 anotações), os principais termos GO para a função molecular (1.112 anotações) foram ligação de ATP (55 anotações), atividade de endopeptidase do tipo serina (53 anotações) e ligação íon zinco (40 anotações). Dentro da categoria de processo biológico (817 anotações), proteólise (87 anotações), transporte transmembrana (47 anotações) e resposta imune inata (31 anotações) foram os termos mais anotados. Na categoria componente celular (565 anotações), os termos principais foram membrana (240 anotações), região extracelular (47 anotações) e membrana plasmática (36 anotações) (Figura 2a, Tabela Suplementar 3a).

Entre os transcritos DE *down* regulados (2.628 anotações), os principais termos GO para a categoria de função molecular (1.074 anotações) foram de atividade de transportador transmembrana (70 anotações), ligação íon metálica (53 anotações) e atividade de endopeptidase do tipo serina (52 anotações). Na categoria de processo biológico (827 anotações), os principais termos foram transporte transmembrana (104 anotações), regulação negativa da atividade endopeptidase (58 anotações) e proteólise (44 anotações). Dentro da função componente celular (727 anotações), os principais termos GO foram membrana (355 anotações), região extracelular (133 anotações) e membrana plasmática (51 anotações) (Figura 2b, Tabela Suplementar 3b).

Além disso, um total de 696 sequências DEs foram atribuídas a Enciclopédia de Genes e Genomas de Kyoto (KEGG), envolvidas em vias metabólicas. Para as sequências *up* reguladas foram identificadas 130 vias metabólicas, e destacam-se as vias relacionadas ao metabolismo da glicina, serina e treonina e biossíntese de metabólitos secundários. Para as sequências *down* reguladas foram identificadas 215 vias metabólicas, dentre elas vias de digestão e absorção de proteínas, metabolismo das purinas, pirimidinas e dos lisossomos. Os caminhos identificados podem ser observados nas Tabelas Suplementares 4a e 4b.

Um total de 775 transcritos foram associados a pelo menos um código de enzima. As enzimas identificadas mais abundantes foram as hidrolases (364 sequências), incluindo principalmente as peptidases (135 sequências), e as transferases (212 sequências), sendo as principais as glicosiltransferases (45 sequências) (Tabela Suplementar 2).

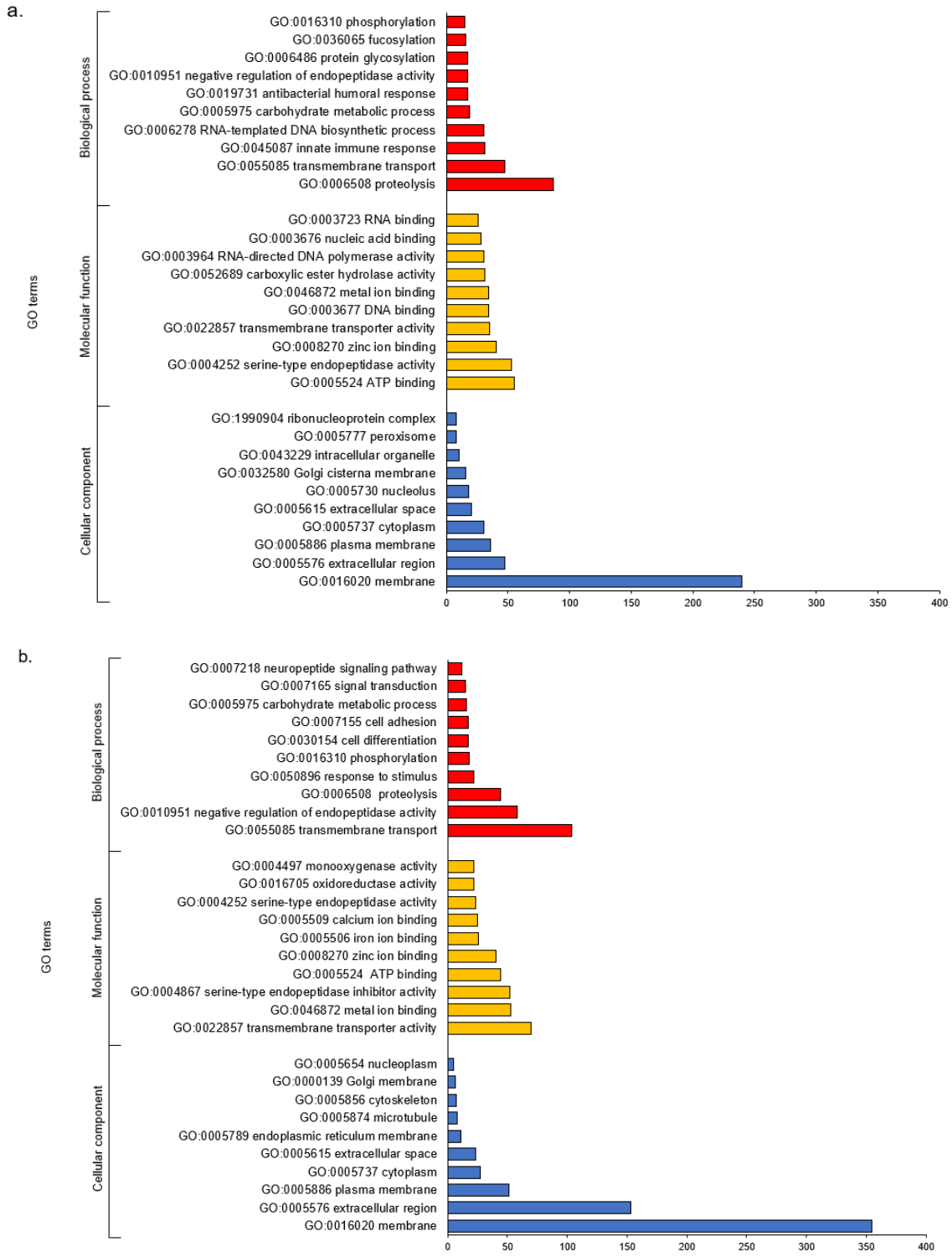

Figura 2. Termos mais frequentes do Gene Ontology para cada uma das três categorias principais, nos transcritos diferencialmente expressos de *B. mori*: **a.** *up* regulados, **b.** *down* regulados (em lagartas criadas a 34 ºC em relação às criadas a 26 ºC).

A análise Pfam permitiu a anotação de 2.531 sequências em domínios Pfam. Entre as sequências *up* reguladas (Figura 3a, Tabela Suplementar 5), o domínio das lipoproteínas de baixo peso molecular de lepidópteros compreendeu o maior número de sequências enquanto o domínio das tripsinas (serina proteases) compreenderam o maior número de sequências entre as *down* reguladas (Figura 3b, Tabela Suplementar 5). A pesquisa de famílias de RNA contra o banco de dados Rfam identificou 25 sequências distribuídas em 9 famílias de RNAs, sendo 7 delas de miRNAs (Tabela Suplementar 6).

Transcritos DEs envolvidos em diferentes processos metabólicos foram identificados, incluindo genes relacionados ao metabolismo de detoxificação, como glutationa S-transferase (maioria *down* regulados), carboxilesterase (todos *up* regulados) e citocromo P450 (maioria *down* regulados). Entre genes envolvidos em respostas imunológicas de *B. mori* foram identificados inibidores de serina protease (SERPINs; maioria *down* regulados), proteína Spatzle-like (*down* regulada) e receptor Toll-like (*down* regulado). Outros genes importantes relacionados ao sistema imune e inibição de apoptose foram os da família das lipoproteínas de baixa massa molecular 30 kDa (30KPs, todos *down* regulados), Leb-3 (atividade antibiótica e antimicrobial, *up* regulado), moricina B3 (resposta de defesa a bactérias, *down* regulada), cecropinas (todos *up* regulados) e gloverinas (todos *up* regulados) (Figura 4, Tabela Suplementar 1).

Diversos fatores de transcrição foram identificados, em especial proteínas com motivos do tipo dedo de zinco (10 sequências *up* reguladas e 11 *down* reguladas). Genes de choque térmico (Hsps) foram identificados, dentre eles sHsp19.5 (*up* regulado) e (Hsp70 *down* regulado). Dentre os genes envolvidos na produção de seda, sericinas (Ser 1, sendo todos *down* regulados; e Ser2, todos *up* regulados) e fibroínas (FibL, cadeia leve; FibH, cadeia pesada; e P25, fibrohexamerina) foram identificados. Todos as sequências de FibL e FibH foram *down* reguladas. Cinco sequências de fibrohexamerina P25 identificadas também foram *down* reguladas, enquanto duas sequências mostraram-se *up* reguladas. (Figura 4, Tabela Suplementar 2).

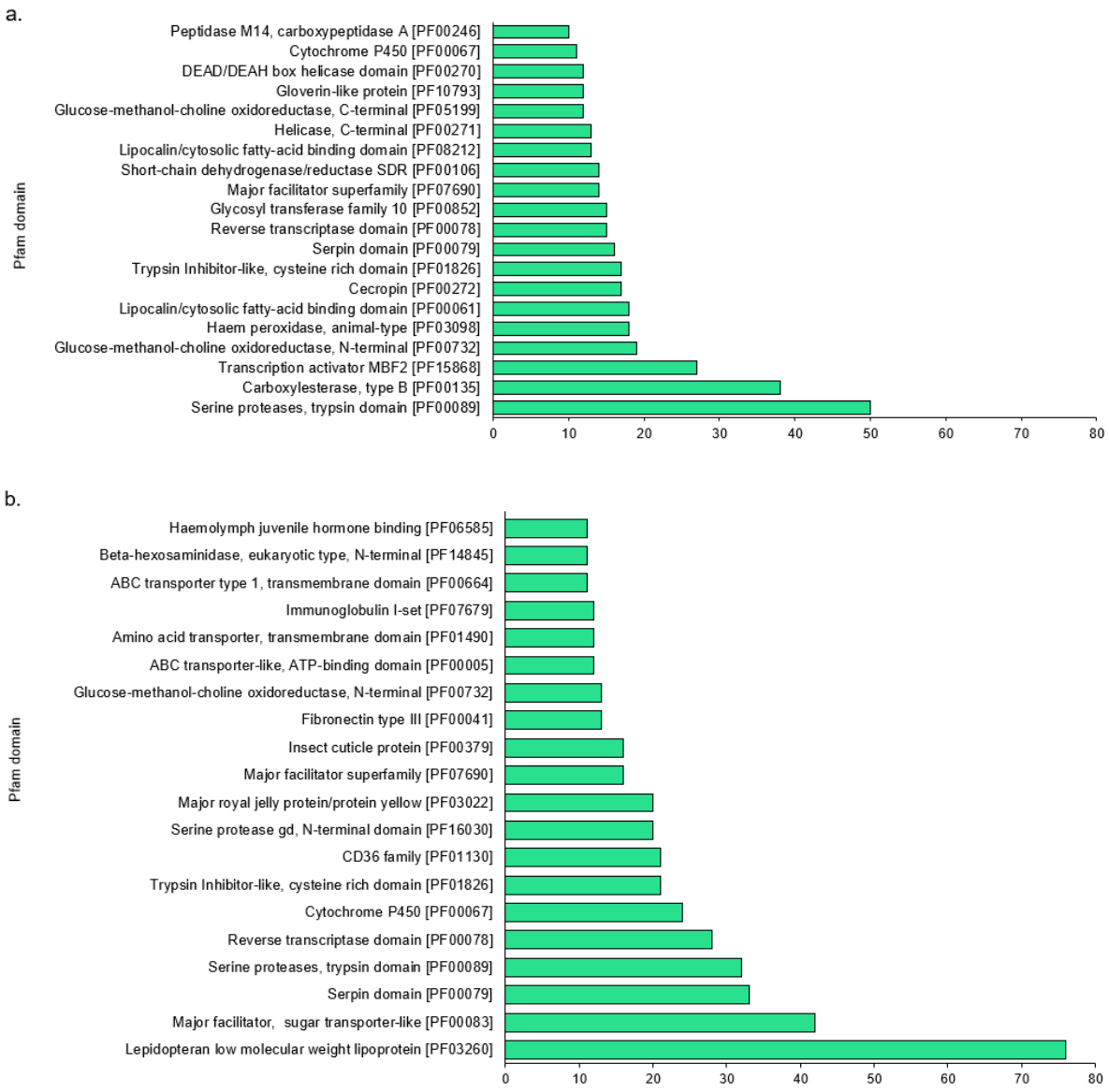

Figura 3. Domínios Pfam mais frequentes para transcritos diferencialmente expressos de *B. mori*. **a.** *Up* regulados, **b.** *down* regulados (em lagartas criadas a 34 ºC em relação às criadas a 26 ºC).

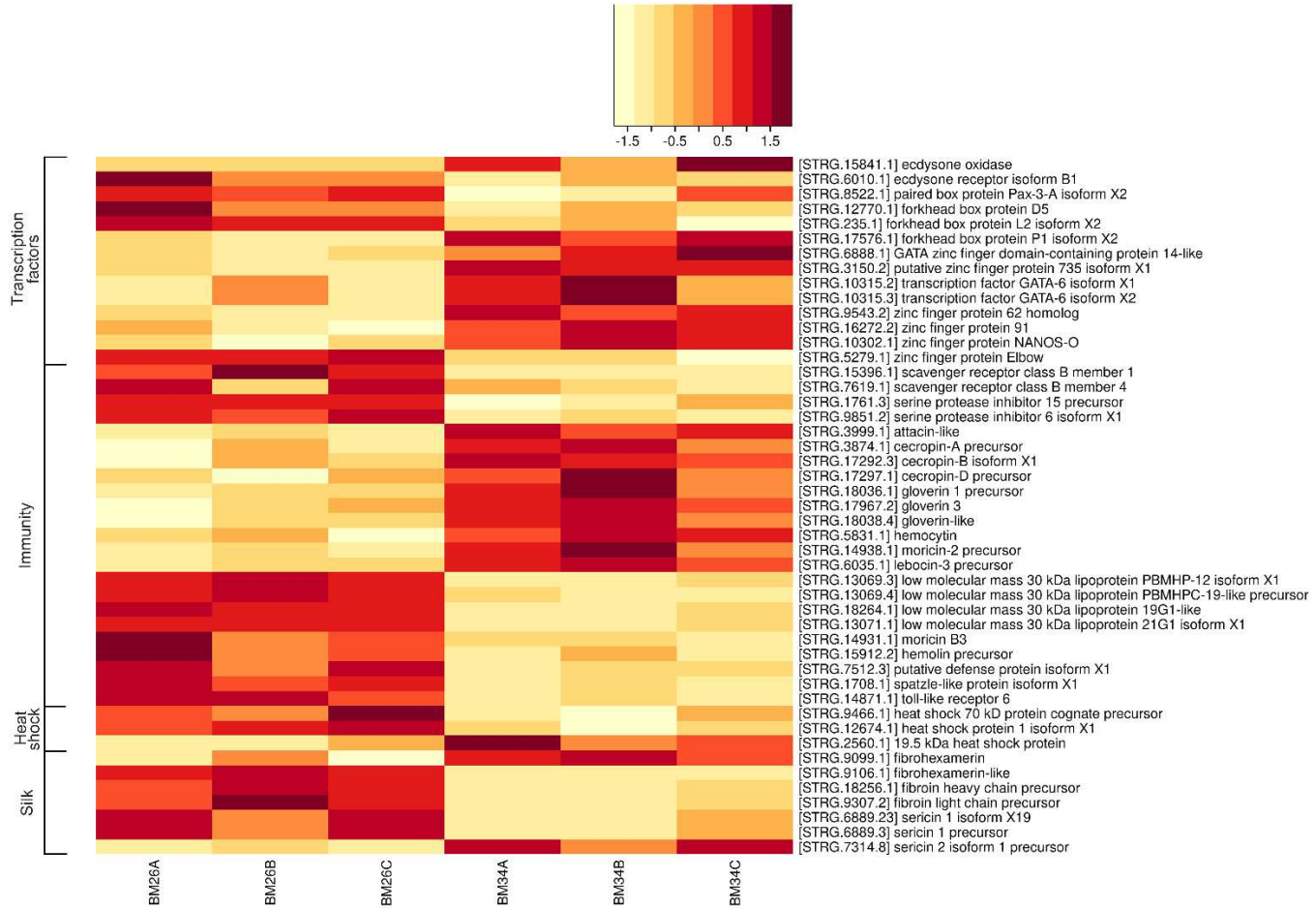

Figura 4. *Heatmap* com o resumo dos principais genes diferencialmente expressos no transcriptomas de *Bombyx mori*. Os níveis de mRNA são representados pelas contagens estimadas normalizadas em DESeq2 (rlog). As diferentes tonalidades de vermelho indicam genes *up* (em tons mais escuros) e *down* regulados (destacados em tons mais claros) em cada biblioteca. Os nomes dos genes estão à direita com as respectivas identificações dos transcritos e os grupos de genes envolvidos na resposta à temperatura identificados à esquerda.

3.3 RT-qPCR

Seis transcritos diferencialmente expressos identificados foram selecionados para análise de expressão gênica por RT-qPCR (Tabela 1, Figura 5). Os dados da análise corroboraram os padrões de expressão dos genes dos exemplares de *B. mori* submetidos a 34 ºC em relação aos submetidos a 26 ºC.

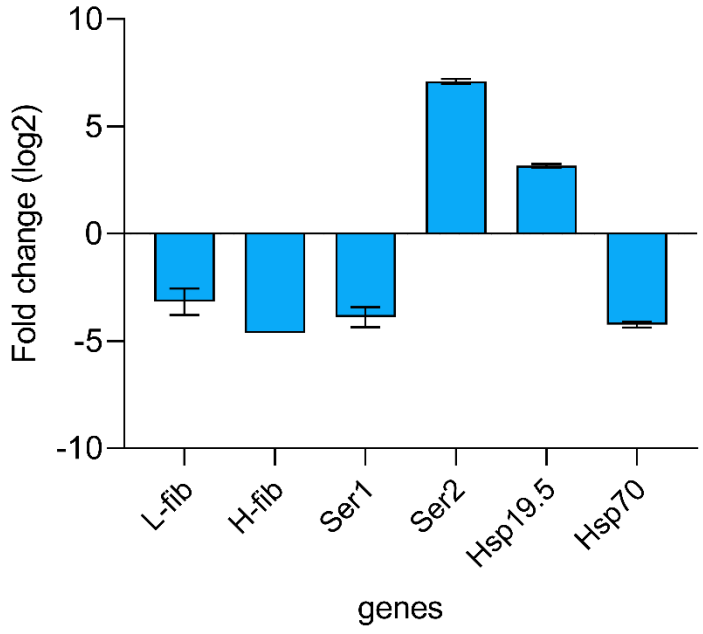

Figura 5. Análise de RT-qPCR (log2foldchange, método $2^{-\Delta\Delta Ct}$) de genes diferencialmente expressos na análise de RNA-seq. Os dados foram normalizados com os genes de α-tubulina e RpL3 e estão apresentados com o erro padrão médio (SEM). L-fib, H-fib, Ser1 e Hsp70 foram regulados negativamente, enquanto Ser2 e Hsp19.5 foram regulados positivamente.

3.4 Parâmetros morfológicos das lagartas e casulos

Todos os grupos iniciaram o quinto instar com peso corporal similar e apresentaram aumento progressivo atingindo um peso aproximadamente cinco vezes maior no dia 6 em relação ao peso inicial para as lagartas criadas a 26 ºC; e apenas 3,8 vezes em para as lagartas criadas a 34 ºC (Tabela 4). As lagartas criadas a 34 ºC tiveram uma diferença significativa ($p<0,05$) de peso em relação ao grupo criado a 26 ºC tanto no dia 3 quanto no dia 6 do quinto instar.

Tabela 4. Peso das lagartas (g) de *B. mori* nos dias 1, 3 e 6 do quinto instar.

	26 ºC	**34 ºC**
Dia 1	0,743 (0,087)	0,745 (0,116)
Dia 3	1,801 (0,236)	1,701 (0,182)*
Dia 6	3,765 (0,261)	2,821 (0,161)*

*Diferença significativa ($p<0,05$) das criadas lagartas a 34 ºC em relação às lagartas criadas a 26 ºC.

O grupo experimental criado a 26 ºC apresentou taxa de mortalidade de 6% enquanto o grupo experimental mantido a 34 ºC apresentou taxa de mortalidade de 9%. No 7º dia do quinto instar, após a última pesagem, as lagartas deram início ao processo de encasulamento. A pesagem total dos casulos produzidos e das cascas foi realizada (Tabela 5).

Tabela 5. Medição dos casulos de *B. mori* produzidos.

	26 ºC	**34 ºC**
Peso dos casulos (g)	1,889 (0,064)	1,711 (0,008)*
Peso da casca (g)	0,356 (0,036)	0,288 (0,014)*
Taxa de seda (%)	18,85	16,83*

*Diferença significativa ($p<0,05$) das criadas lagartas a 34 ºC em relação às lagartas criadas a 26 ºC.

Os casulos totais e o peso das cascas produzidos pelas lagartas criadas a 34 ºC tiveram uma diferença significativa ($p<0,05$) de peso em relação ao grupo criado a 26 ºC. Houve também redução significativa na taxa de conversão de seda.

4. Discussão

A tolerância de insetos à temperatura é geralmente considerada um importante fator para adaptação às mudanças ambientais. Estudos sobre os mecanismos de tolerância dos insetos a temperaturas elevadas são de grande importância para a criação de insetos de interesse econômico de forma eficiente, especialmente em países tropicais. Diversos estudos já foram realizados para o entendimento do fator temperatura em *B. mori*, incluindo a análise de seus transcritos em determinadas condições de temperatura (Xiao et al., 2017; Guo et al., 2018). No entanto, avaliar a resposta desse inseto a duas condições de temperatura ao longo de todo o quinto instar, considerado um período crítico da vida do inseto – e um dos períodos mais sensíveis a alta temperatura, e sua preparação para a confecção do casulo, matéria prima da cadeia produtora de seda, pode fornecer respostas mais profundas sobre a atuação da temperatura no desenvolvimento das lagartas e na seleção de melhores estratégias de criação mais rentáveis de *B. mori*.

Dentre os principais impactos negativos observados da 34 ºC em *B. mori* em relação a criação a 26 ºC incluem: (a) A inibição do crescimento e desenvolvimento do inseto; (b) redução da síntese de seda; (c) redução da resistência a patógenos.

Inibição do crescimento e desenvolvimento do inseto

As lagartas que foram mantidas na temperatura mais elevada durante o quinto instar apresentaram média menor de peso (0,944g em relação ao outro grupo experimental) (Tabela 4). É bem entendido que a maioria das características genéticas economicamente importantes de *B. mori* são de natureza qualitativa e que a expressão fenotípica é grandemente influenciada por fatores ambientais, como umidade relativa, luz, nutrição e, em especial, a temperatura (Chatterjee et al., 1993; Wu & Hou, 1993; Ramesha et al., 2009; Rahmathulla, 2012). Em relação aos parâmetros econômicos das lagartas, os grupos experimentais apresentaram taxas de mortalidade similares (6 e 9%) às obtidas em campos de criação (Takahashi et al., 2009), sugerindo que a diferença de temperaturas durante o quinto instar não afetou a taxa de mortalidade. No entanto, é possível observar a influência direta da temperatura elevada principalmente no crescimento e desenvolvimento das lagartas e na quantidade de seda produzida. As lagartas do grupo mantido a 34 ºC produziram, em média, uma quantidade de seda menor (Tabela 5). Isso acontece pois em durante processos de infecção ou estresse, como a temperatura elevada, a tendência é que as lagartas de *B. mori* se

alimentem menos e consequentemente ocasione menores taxas de crescimento e produção de seda (Aneesha & Kumar, 2022). Enquanto isso, proteínas de dedo de zinco são importantes fatores de transcrição, que podem ser ligados linearmente em conjunto para reconhecer as sequências de ácido nucleico de comprimentos variados (Klug, 2010; Gillinder et al., 2017). Os níveis de expressão aumentados nas lagartas criadas a 34 ºC em relação às criadas em 26 ºC implicam que *B. mori* possivelmente regula positivamente a transcrição desses genes essenciais para a sobrevivência sob alta temperatura. No entanto, a contínua transcrição elevada de mRNA e síntese de proteínas de genes importantes para a sobrevivência consumiriam a maior parte da energia e recursos biológicos da lagarta, o que afetaria a transcrição de outros genes, incluindo síntese reduzida de seda (Li et al., 2012; Guo et al., 2018).

Redução da síntese de seda

A redução da síntese de seda pode ser observada no presente estudo, já que os genes de fibroína (todos os de FibL e FibH, e a grande maioria de P25), proteína mais abundante na composição de casulos de *B. mori* e de função estrutural, tiveram seus níveis de expressão reduzidos durante a criação em temperatura mais elevada (Figura 4). Alguns outros estudos já demonstraram que quando a seda é confeccionada em temperaturas elevadas, isso ocasiona também alterações na estrutura da sericina (Rahmathulla, 2012), a segunda proteína mais abundante na composição dos fios de seda, e afinamento e espessamento abrupto do filamento produzido (Ramachandra et al., 2001), resultando em casulos mais leves e de baixa qualidade. Um outro importante grupo de proteínas também envolvidas no processo de produção de seda são as seroínas, que são expressas nas porções posterior e média das glândulas de seda. A expressão do gene que codifica essas proteínas flutua durante todo o desenvolvimento do inseto: o mRNA é elevado nas glândulas de seda das lagartas durante a alimentação, diminui durante a ecdise, atinge um máximo durante a fiação do casulo e, posteriormente, cai rapidamente para um nível indetectável (Žurovec et al., 1998). No presente estudo, os níveis de seroína em *B. mori* tiveram seus níveis de expressão reduzidos em lagartas criadas a 34 ºC em relação às criadas em 26 ºC durante o quinto instar, período em que as lagartas se preparam para a confecção do casulo.

Redução da resistência a patógenos

O sistema imune inato dos insetos consiste em respostas celulares e humorais (Wu & Yi, 2008). A resposta imune celular refere-se à fagocitose e encapsulamento, mediadas por hemócitos circulantes (Hillyer, 2016), enquanto a resposta imune humoral inclui a síntese de peptídeos antimicrobianos (AMPs), ativação do sistema profenoloxidase (proPO) e geração de oxigênio reativo espécies (ROS) no corpo gorduroso, hemócitos ou outros tecidos (Hoffmann, 1995). Estudos anteriores relataram a expressão das Hsps (do inglês, *Heat shock proteins*), e seu importante papel no processo de resistência ao calor em insetos. São uma família de proteínas que ajudam os organismos a se adaptarem e se protegerem contra danos celulares induzidos pelo ambiente, sendo também essenciais na regulação da apoptose e ativação do sistema imunológico inato (Lindquist & Craig, 1998).

Em relação ao presente estudo, uma das proteínas Hsp identificada, Hsp70, mostrou-se *down* regulada (Figuras 4 e 5, Tabela Suplementar 1) Segundo Wu & Yi (2008), a regulação positiva de Hsp70 pode ativar as vias de sinalização relacionadas ao sistema imunológico e consequentemente aumentar as respostas imunes celulares, a exemplo encapsulamento e fagocitose, auxiliando os organismos a lidar com os riscos ambientais. Sendo assim, as lagartas criadas a 34 ºC estão possivelmente tendo seus processos de respostas imunes afetadas, assim como foi observado por Li et al. (2013) em lagartas da linhagem Nistari mantidas a 32ºC durante o quinto instar. Enquanto isso, a sHsp19.15 mostrou-se *up* regulada (Figuras 4 e 5, Tabela Suplementar 1). A sHsp19.5 faz parte do grupo das sHsps (do inglês, *Small heat shock proteins*), e são extremamente importantes devido a sua função de chaperonas moleculares, sendo responsáveis por protegerem outras proteínas da desnaturação durante processos de estresse (Haslbeck et al. 2005; Li et al., 2009; Basha et al. 2012). Essas sHsps também estão envolvidas em processos celulares como apoptose e autofagia, actina e dinâmica de filamentos intermediários, organização do citoesqueleto e fluidez de membrana (Haslbeck 2002; Quinlan, 2002; Tsvetkova et al. 2002; Sun & MacRae, 2005). Apesar da importância semelhante das Hsps, o diferente padrão de expressão de genes Hsp no presente estudo já foi observado em outros trabalhos, e indica que pode haver diferenças na expressão e função desses genes em resposta à alta temperatura (Xiao et al., 2017).

Um outro grupo importante de proteínas envolvidas no mecanismo de resistência a patógenos são as lipoproteínas mais abundantes na hemolinfa durante o quinto instar até estágio de pupa em *B. mori*, que tem baixa massa molecular de 30 kDa (30KPs). 30KPs

possuem função de inibir a morte celular programada (Rhee et al. 2002) prolongando a sobrevivência celular (Shi et al., 2015). 30KPs podem ser conjugadas a glicose, dextrano, maltose e glicoproteínas, podendo agir como um mecanismo de defesa (Ujita et al., 2005, Ueno et al., 2006; Shi et al., 2015). Um mecanismo do sistema imunológico inato de *B. mori* contra fungos envolve a interação de 30KPs de ligação a b-glicanas fúngicas, que ativa a profenoloxidase (proPO), interferindo assim no crescimento de hifas de modo a proteger *B. mori* da infestação fúngica (Ujita et al. 2005; Shi et al., 2015). No presente estudo, todos os genes de 30KPs se demonstraram *down* regulados, evidenciando que a temperatura de 34 ºC pode estar afetando diretamente a resposta imunológica a patógenos, assim como pode reduzir a sobrevivência celular, prejudicando o desenvolvimento das lagartas e sua produção de seda.

Em conclusão, os nossos dados indicam que a criação de lagartas de *B. mori* a 34 ºC em comparação a lagartas mantidas a 26 ºC durante o quinto instar afetam diretamente o crescimento, desenvolvimento e o sistema imunológico desses insetos, tendo como consequência um efeito prejudicial na produção de seda. Entender a resposta a determinadas temperaturas pode ser uma estratégia interessante para o aprimoramento da produção de seda, visto que é um dos fatores que mais impactam a cadeia produtiva. Novas abordagens experimentais para ampliar o conhecimento sobre a temperatura são necessárias, a exemplo, outras temperaturas de criação, e até mesmo a combinação de diferentes temperaturas em diferentes estágios de desenvolvimento das lagartas de *B. mori* para estabelecer a melhor temperatura de criação e rentabilidade.

Agradecimentos

Os autores agradecem a Unidade Gestora do Fundo Paraná (UGF – TC 55/2018); ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Finance code 001) pelo financiamento; à Fiação de seda BRATAC (Londrina, Paraná) pela parceria e fornecimento das larvas; aos pesquisadores Dr. Ruy Yamaoka e Dr. Dimas Soares Júnior do Instituto Agronômico do Paraná (Instituto de Desenvolvimento Rural do Paraná – IDR-PR) pela permissão ao acesso do Campo de amoreiras; a Financiadora de Estudos e Projetos (FINEP) e a Fundação Araucária pelo apoio a infraestrutura dos laboratórios.

Referências

Aneesha, U., & Kumar, C. S. (2022). Foliar supplementation of ascorbic acid modulates biochemical performance of silkworm, *Bombyx mori* under thermal stress. *Journal of Thermal Biology*, 103184. https://doi.org/10.1016/j.jtherbio.2021.103184

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., & Ogata, H. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, *36*(7), 2251-2252. https://doi.org/10.1093/bioinformatics/btz859

Basha, E., O'Neill, H., & Vierling, E. (2012). Small heat shock proteins and α-crystallins: dynamic proteins with flexible functions. *Trends in biochemical sciences*, *37*(3), 106-117. https://doi.org/10.1016/j.tibs.2011.11.005

Chatterjee, S. N., Rao, C. G. P., Chatterjee, G. K., Ashwath, S. K., & Patnaik, A. K. (1993). Correlation between yield and biochemical parameters in the mulberry silkworm, *Bombyx mori* L. *Theoretical and Applied Genetics*, *87*(3), 385-391. https://doi.org/10.1007/BF01184928

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610

Fang, S. M., Hu, B. L., Zhou, Q. Z., Yu, Q. Y., & Zhang, Z. (2015). Comparative analysis of the silk gland transcriptomes between the domestic and wild silkworms. *BMC genomics*, *16*(1), 1-12. https://doi.org/10.1186/s12864-015-1287-9

Gillinder, K. R., Ilsley, M. D., Nébor, D., Sachidanandam, R., Lajoie, M., Magor, G. W., Tallack, M. R., Bailey, T., Landsberg, M. J., Mackay, J. P., Parker, M. W., Miles, L. A, Graber, J. H., Peters, L. L., Bieker, J. J. & Perkins, A. C. (2017). Promiscuous DNA-binding of a mutant zinc finger protein corrupts the transcriptome and diminishes cell viability. *Nucleic acids research*, *45*(3), 1130-1143. https://doi.org/10.1093/nar/gkw1014

Guo, H., Huang, C., Jiang, L., Cheng, T., Feng, T., & Xia, Q. (2018). Transcriptome analysis of the response of silkworm to drastic changes in ambient temperature. *Applied microbiology and biotechnology*, *102*(23), 10161-10170. https://doi.org/10.1007/s00253-018-9387-5

Haslbeck, M. (2002). sHsps and their role in the chaperone network. *Cellular and Molecular Life Sciences CMLS*, *59*(10), 1649-1657. https://doi.org/10.1007/PL00012492

Haslbeck, M., Franzmann, T., Weinfurtner, D., & Buchner, J. (2005). Some like it hot: the structure and function of small heat-shock proteins. *Nature structural & molecular biology*, *12*(10), 842-846. https://doi.org/10.1038/nsmb993

Hillyer, J. F. (2016). Insect immunology and hematopoiesis. *Developmental & Comparative Immunology*, *58*, 102-118. https://doi.org/10.1016/j.dci.2015.12.006

Hoffmann, J. A. (1995). Innate immunity of insects. *Current opinion in immunology*, *7*(1), 4-10. https://doi.org/10.1016/0952-7915(95)80022-0

Hu, J. H., Cheng, X. Y., Li, J. X., Xue, B., Tian, J. H., Hu, J. S., & Li, B. (2018). Apoptosis of posterior silk gland of Bombyx mori during spinning period and the role of PI3K/Akt pathway. *Archives of insect biochemistry and physiology*, *98*(1), e21450. https://doi.org/10.1002/arch.21450

Inserco. (2023). Statistics. Disponível em: https://inserco.org/en/statistics

James, R. R.; Li, Z. (2012). Chapter 12 - From Silkworms to Bees: Diseases of Beneficial Insects. Editors: Vega, F. E.; Kaya, H. K. Insect Pathology (Second Edition), *Academic Press*, 425-459. ISBN 9780123849847 https://doi.org/10.1016/B978-0-12-384984-7.00012-9.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, *28*(1), 27-30. https://doi.org/10.1093/nar/28.1.27

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of molecular biology*, *428*(4), 726-731. https://doi.org/10.1016/j.jmb.2015.11.006

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology*, *37*(8), 907-915. https://doi.org/10.1038/s41587-019-0201-4

Klug, A. (2010). The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual review of biochemistry*, 79, 213-231. https://doi.org/10.1146/annurev-biochem-010909-095056

Langmead, B., Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359. https://doi.org/10.1038/nmeth.1923

Li, J., Ye, L., Lan, T., Yu, M., Liang, J., & Zhong, B. (2012). Comparative proteomic and phosphoproteomic analysis of the silkworm (*Bombyx mori*) posterior silk gland under high temperature treatment. *Molecular Biology Reports*, *39*(8), 8447–8456. https://doi.org/10.1007/s11033-012-1698-5

Li, J., Moghaddam, S. H. H., Du, X., Zhong, B. X., & Chen, Y. Y. (2012). Comparative analysis on the expression of inducible HSPs in the silkworm, *Bombyx mori*. *Molecular biology reports*, 39, 3915-3923. https://doi.org/10.1007/s11033-011-1170-y

Li, Z. W., Li, X., Yu, Q. Y., Xiang, Z. H., Kishino, H., & Zhang, Z. (2009). The small heat shock protein (sHSP) genes in the silkworm, *Bombyx mori*, and comparative analysis with other insect sHSP genes. *BMC Evolutionary Biology*, *9*(1), 1-14. https://doi.org/10.1186/1471-2148-9-215

Lindquist, S., & Craig, E. A. (1988). The heat-shock proteins. *Annual review of genetics*, 22(1), 631-677. https://doi.org/10.1146/annurev.ge.22.120188.003215

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), 1-21. https://doi.org/10.1101/002832

Naik, K. S., Ismail, S., Pradeep, A. R., & Mishra, R. K. (2022). Molecular Characterization of the Functional Genes Associated with Silk Assembly, Transport, and Protection in the Silk Glands of Popular Multivoltine Breeds of Silkworm Bombyx mori. L. *Applied Biochemistry and Biotechnology*, 1-24. https://doi.org/10.1007/s12010-022-04158-2

Pfaffl, M. W., Horgan, G. W., & Dempfle, L. (2002). Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic acids research*, *30*(9), e36-e36. https://doi.org/10.1093/nar/30.9.e36

Quinlan, R. (2002). Cytoskeletal competence requires protein chaperones. *Small stress proteins*, 219-233. https://doi.org/10.1007/978-3-642-56348-5_12

Rahmathulla, V. K. (2012). Management of climatic factors for successful silkworm (*Bombyx mori* L.) crop and higher silk production: a review. *Psyche*. https://doi.org/10.1155/2012/121234

Ramachandra, Y. L., Bali, G., & Rai, S. P. (2001). Effect of temperature and relative humidity on spinning behaviour of silkworm (*Bombyx mori* L.). PMDI: 11349535

Ramesha, C., Seshagiri, S. V., & Rao, C. G. P. (2009). Evaluation and identification of superior polyvoltine crossbreeds of mulberry silkworm, *Bombyx mori* L. *Journal of Entomology*, *6*(4), 179-188. https://doi.org/10.3923/je.2009.179.188

Rhee, W. J., Kim, E. J., & Park, T. H. (2002). Silkworm hemolymph as a potent inhibitor of apoptosis in Sf9 cells. *Biochemical and Biophysical Research Communications*, *295*(4), 779-783. https://doi.org/10.1016/S0006-291X(02)00746-5

Sahi, A. K., Gundu, S., Kumari, P., Klepka, T., & Sionkowska, A. (2023). Silk-Based Biomaterials for Designing Bioinspired Microarchitecture for Various Biomedical Applications. *Biomimetics*, *8*(1), 55. https://doi.org/10.3390/biomimetics8010055

Shi, X. F., Li, Y. N., Yi, Y. Z., Xiao, X. G., & Zhang, Z. F. (2015). Identification and characterization of 30 K protein genes found in *Bombyx mori* (Lepidoptera: Bombycidae) transcriptome. *Journal of Insect Science*, *15*(1). https://doi.org/10.1093/jisesa/iev057

Sun, Y., & MacRae, T. H. (2005). Small heat shock proteins: molecular structure and chaperone function. *Cellular and Molecular Life Sciences CMLS*, *62*(21), 2460-2476. https://doi.org/10.1007/s00018-005-5190-4

Takahashi, R., Takahashi, K. M., & Takahashi, L. S. (2001). *Sericicultura: uma promissora exploração agropecuária*. Funep.

Takasu, Y., Hata, T., Uchino, K., & Zhang, Q. (2010). Identification of Ser2 proteins as major sericin components in the non-cocoon silk of *Bombyx mori*. *Insect biochemistry and molecular biology*, *40*(4), 339-344. https://doi.org/10.1016/j.ibmb.2010.02.010

Tsvetkova, N. M., Horváth, I., Török, Z., Wolkers, W. F., Balogi, Z., Shigapova, N., Crowe, L. M., Vierling, E., Crowe, J. H., & Vígh, L. (2002). Small heat-shock proteins regulate membrane lipid polymorphism. *Proceedings of the National Academy of Sciences*, *99*(21), 13504-13509. https://doi.org/10.1073/pnas.192468399

Ueno, Y., He, N., Ujita, M., Yamamoto, K., Banno, Y., Fujii, H., & Aso, Y. (2006). Silkworm midgut proteins interacting with a hemolymph protease inhibitor, CI-8. *Bioscience, biotechnology, and biochemistry*, *70*(7), 1557-1563. https://doi.org/10.1271/bbb.50548

Ujita, M., Katsuno, Y., Kawachi, I., Ueno, Y., Banno, Y., Fujii, H., & Hara, A. (2005). Glucan-binding activity of silkworm 30-kDa apolipoprotein and its involvement in defense against fungal infection. *Bioscience, biotechnology, and biochemistry*, *69*(6), 1178-1185. https://doi.org/10.1271/bbb.69.1178

Velu, D., Ponnuvel, K. M., & Qadri, S. M. H. (2008). Expression of the heat shock protein genes in response to thermal stress in the silkworm *Bombyx mori*. *International Journal of Industrial Entomology*, *16*(1), 21-27. https://koreascience.kr/article/JAKO200833338747379.page

Wu, D. J., & Hou, R. F. (1993). Relationship between thermotolerance and heat-stable esterase in the silkworm, *Bombyx mori* (Lepidoptera: Bombycidae). *Applied Entomology and Zoology*, *28*(3), 371-377. https://doi.org/10.1303/aez.28.371

Wu, G., & Yi, Y. (2018). Transcriptome analysis of differentially expressed genes involved in innate immunity following *Bacillus thuringiensis* challenge in *Bombyx mori* larvae. *Molecular immunology*, *103*, 220-228. https://doi.org/10.1016/j.molimm.2018.10.006

Xiao, W., Chen, P., Xiao, J., Wang, L., Liu, T., Wu, Y., Dong, F., Jiang, Y., Pan, M., Zhang, Y. & Lu, C. (2017). Comparative transcriptome profiling of a thermal resistant vs. sensitive silkworm strain in response to high temperature under stressful humidity condition. *PLoS One*, *12*(5), e0177641. https://doi.org/10.1371/journal.pone.0177641

Yang, C., Yao, L., & Zhang, L. (2023). Silk sericin-based biomaterials shine in food and pharmaceutical industries. *Smart Materials in Medicine*. https://doi.org/10.1016/j.smaim.2023.01.003

Žurovec, M., Yang, C., Kodrık, D., & Sehnal, F. (1998). Identification of a novel type of silk protein and regulation of its expression. *Journal of Biological Chemistry*, *273*(25), 15423-15428. https://doi.org/10.1074/jbc.273.25.15423

CAPÍTULO II

Variação na temperatura de criação de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) e sua influência na expressão de elementos transponíveis

O manuscrito será submetido ao periódico *Insect Molecular Biology*.

Variação na temperatura de criação de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae) e sua influência na expressão de elementos transponíveis

Thayná Bisson Ferraz Lopes[1], Rogério Fernandes de Souza[1], Cristianne Cordeiro Nascimento[2], Renata da Rosa[1]

[1] Departamento de Biologia Geral, Universidade Estadual de Londrina, Londrina, Paraná, Brasil

[2] Departamento de Design Gráfico, Centro de Educação, Comunicação e Artes, Universidade Estadual de Londrina, Londrina, Paraná, Brasil

Resumo

Os elementos transponíveis (TEs) são sequências de DNA com capacidade de se mover de um local genômico para outro, constituindo uma fração significativa dos genomas de grande parte dos eucariotos. Esses elementos são importantes por sua influência na diversidade genética e evolução genômica, podendo também impactar na expressão gênica. No presente estudo identificamos e caracterizamos os TEs em transcritos de *B. mori*, lepidóptero de importância comercial na produção de seda no Brasil e no mundo, submetidos a duas temperaturas de criação, 26°C e 34°C. A identificação e comparação foi realizada a partir de buscas baseadas em bancos de dados utilizando a ferramenta blastn, além de buscas baseadas em *Hidden Markov Model*. Foram identificados 11.697 transcritos com semelhança significativas a TEs e/ou domínios característicos. Os TEs identificados foram classificados em ordens, superfamílias e outros clados. Os retrotransposons representaram 79% de transcritos identificados relacionados a TEs, em especial os da ordem LINE, enquanto os transposons de DNA representaram cerca de 21% dos transcritos identificados, com a maioria das sequências pertencentes a ordem TIR. Os transcritos sem classificação representaram aproximadamente 6%. Por sua vez, a análise de expressão diferencial de TEs em lagartas criadas a 26 ºC e 34 ºC indicou a presença de 784 sequências semelhantes a TEs, sendo 227 sequências *up* reguladas e 557 *down* reguladas. Dessa forma, os dados contribuem para um melhor entendimento da distribuição de elementos genéticos móveis na espécie e demonstram que a temperatura pode estar relacionada com a modulação de TEs.

Palavras-chave: Estresse térmico; retrotransposons; transposons de DNA; transcriptoma.

1. Introdução

Os elementos genéticos móveis ou elementos transponíveis são sequências de DNA que possuem a capacidade de mudar sua posição dentro de um genoma, podendo criar cópias durante o processo de mobilização (Kazazian, 2004; Bourque et al., 2018; Drongitis et al., 2019). Esses elementos compreendem uma parte considerável do genoma em diversas espécies. Em *Bombyx mori* (Lepidoptera: Bombycidae), considerado um inseto modelo usado em vários estudos, e economicamente relevante na cadeia produtiva de seda, assim como nas indústrias de biotecnologia, os TEs representam em torno de 46% do genoma total (Kawamoto et al., 2019).

Os TEs normalmente são classificados em classes com base em seu mecanismo de transposição, de acordo com a presença ou ausência de um intermediário de RNA (retrotransposons ou classe I, e transposons de DNA ou classe II, respectivamente). Podem também ser classificados em subclasses, ordens, superfamílias, famílias e subfamílias, segundo o sistema de classificação proposto por Wicker et al. (2007), e ainda como autônomos e não autônomos (dependendo da sua capacidade de codificar os domínios necessários para sua transposição).

Os TEs de classe I utilizam um mecanismo de "copiar e colar" (Wicker et al., 2007; Bourque et al., 2018; Yokoi et al., 2022) e são divididos em cinco ordens, sendo elas: *Long Terminal Repeats* (LTRs), *Dictyostelium Intermediate Repeat Sequence* (DIRS), Penelope-like (PLE), *Long Interspersed Nuclear Element* (LINEs), e *Short Interspersed Nuclear Elements* (SINEs). Essas ordens ainda podem ser divididas em superfamílias. Os LTRs, por exemplo, são divididos em cinco superfamílias, sendo elas: *Copia*, Gypsy, *Bel-Pao*, *Retrovirus* e *ERV*. Enquanto isso, os TEs de classe II movem-se usando um mecanismo de "recortar e colar" (Wicker et al., 2007; Bourque et al., 2018; Yokoi et al., 2022). Os TEs classe II também são divididos em ordens, dentre elas *Terminal Inverted Repeat* (TIRs) e Crypton (subclasse I), e Helitron e Marverick (subclasse 2). Cada ordem é também classificada em superfamílias, das quais algumas são divididas em várias famílias.

Diversos efeitos de TEs nas espécies hospedeiras já foram relatados. A mobilização, expressão e inserção de TEs podem ser deletérias se não forem controladas, com efeitos negativos em seus hospedeiros (Zamudio & Bourc'His, 2010; Drongitis et al., 2019). No entanto, eles também envolvidos na formação da estrutura genômica, servem como fonte de mutações e estão envolvidos também no mecanismo de regulação da expressão gênica

(Slotkin & Martienssen, 2007) e de diversos processos biológicos, como transcrição de genes, *splicing* alternativo e tradução do RNA (DeSouza et al., 2013; Drongitis et al., 2019). Dessa forma, é possível dizer que, por um lado, os TEs podem induzir mutações deletérias, causando disfunções, doenças e até letalidade nos indivíduos. Por outro lado, os TEs podem aumentar a variabilidade genética, proporcionando às populações uma melhor resposta adaptativa a determinadas mudanças ambientais. Sendo assim, os TEs podem ser considerados ao mesmo tempo "parasitas" e "simbiontes" do genoma, de acordo com a modulação de sua atuação.

Diversos estudos têm sugerido um papel adaptativo de TEs na resposta ao estresse (Capy et al., 2000; Fablet & Vieira, 2011; Chénais et al., 2012; Casacuberta & González, 2013; Rey et al., 2016). Entretanto, informações sobre a relação entre elementos transponíveis e sua atuação na resposta a diferentes condições de temperatura em *B. mori* ainda não foi elucidada. Dessa forma, no presente estudo realizamos a identificação e caracterização de TEs nos transcritos de *Bombyx mori* e a análise de expressão diferencial de TEs em lagartas mantidas a 26 ºC e 34 ºC durante o quinto instar, a fim de entender a resposta desses elementos a situação de estresse térmico durante o período mais crítico da vida desse inseto.

2. Material e Métodos

2.1 Criação das lagartas

A criação das lagartas para obtenção do transcriptoma foi realizada segundo Lopes et al., (2023) (em fase de pré-publicação). Duas temperaturas de criação foram testadas, uma para cada grupo experimental: (1) 26 ºC, e (2) 34 ºC durante o quinto instar do estágio larval. Ao final desse estágio, três lagartas de cada grupo experimental foram selecionadas aleatoriamente e imediatamente mantidas em freezer -80°C para posterior extração de RNA.

2.2 Extração de RNA e sequenciamento

Três lagartas de cada grupo foram selecionadas para extração de RNA e sequenciamento. Os indivíduos criados a 26 ºC foram identificados como BM26A, BM26B e BM26C, e, os criados a 34 ºC, como BM34A, BM34C e BM34C. A extração, purificação e avaliação da integridade do RNA seguiram as metodologias descritas por Lopes et al., (2023) (em fase de pré-publicação).

As amostras de RNA total (três réplicas biológicas por tratamento) foram submetidas para preparação de bibliotecas de cDNA usando protocolo *dUTP-stranded* e um mínimo de 20 milhões de *paired-end reads* (2x125-bp) por amostra foram sequenciados usando a tecnologia Illumina HiSeq 2500. O processamento dos dados brutos foi feito utilizando o Illumina CASAVA pipeline, para obtenção dos dados de sequencias no formato FASTQ.

2.3 Montagem do transcriptoma e análise de expressão diferencial

Os parâmetros de qualidade das leituras geradas após o sequenciamento para cada biblioteca foram avaliados com o auxílio da ferramenta FastQC v0.11.3 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), seguidos da correção dos *reads* raros no programa rCorrector v1.0.4 (https://github.com/mourisl/Rcorrector). Os *reads* não corrigidos foram retirados com o *script* FilterUncorrectabledPEfastq.py (https://github.com/harvardinformatics/TranscriptomeAssemblyTools/blob/master/FilterUncorrectabledPEfastq.py). A remoção de pares de bases de baixa qualidade, fragmentos curtos e sequências adaptadoras do processo de sequenciamento foram realizadas com o programa TrimGalore! v0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).

Os transcriptomas das seis bibliotecas sequenciadas foram alinhados ao genoma de referência disponível no NCBI, de número de referência Ref Seq GCF_014905235_1 utilizando o programa HISAT2 v2.1.0 (Kim et al., 2019). A montagem guiada do transcriptoma realizada com o programa StringTie v2.1.1 (http://ccb.jhu.edu/software/stringtie/). A qualidade da montagem obtida foi descrita por Lopes et al., (2023) (em fase de pré-publicação).

A análise de expressão gênica diferencial entre os grupos experimentais de lagartas mantidas a 26°C e 34°C foi avaliada com o pacote DESeq2 v1.26.0 (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) (Love et al., 2014), implementado no programa RStudio v2021.09.0 Build 351 (https://rstudio.com/). Somente os transcritos diferencialmente expressos com probabilidade ajustada de p menor que o limite alfa de 5% (padj $< 0,05$) e log2foldchange $\geq 2,0$ foram considerados.

2.4 Anotação de TEs

Para a identificação e caracterização de sequências de TEs em *B. mori*, os transcritos foram anotados de acordo com a sua similaridade a elementos previamente caracterizados disponíveis em diferentes bancos de dados.

A identificação e caracterização de TEs no transcriptoma montado foram realizadas com os bancos de dados Repbase (Taxon: Hexapoda) (Bao et al., 2015) e Arthropod Transposable Elements Database (ArTEdb) (Taxon: Insecta) (http://artedb.net/), utilizando a ferramenta blastn e considerando apenas sequências com identidade mínima de 60%. utilizando a ferramenta blastn.

A identificação de famílias proteicas de TEs no transcriptoma montado foi realizada utilizando o Dfam v 3.7 (https://www.dfam.org/), considerando apenas as famílias curadas na busca, e Gypsy Database (LTRs e TIRs Seq. DB) (Llorens et al., 2011), utilizando a ferramenta blastn e considerando apenas sequências com identidade mínima de 60%. Para a localização de domínios proteicos característicos de TEs, foi utilizado o Pfam (https://www.ebi.ac.uk/interpro/).

A identificação de famílias de TEs no genoma de referência para a montagem do transcriptoma foi utilizada como banco de dados para posterior comparação com o transcriptoma montado. A identificação dos elementos no genoma foi realizada no programa RepeatModeler v.2.0.3 e, posteriormente, o CD-HIT (http://cd-hit.org), para agrupar sequências biológicas com objetivo de reduzir a redundância de sequências e melhorar o desempenho das análises, utilizando um *threshold* de identidade ≥ 80%. A comparação entre os TEs identificados no genoma e o transcriptoma montado foi realizado pela ferramenta blastn.

3. Resultados

3.1 Identificação e caracterização de elementos transponíveis no transcriptoma de *Bombyx mori*

A anotação usando o banco de dados Repbase previu a presença de 7.402 transcritos semelhantes à TEs, distribuídos em 1.007 identificações diferentes e agrupados em mais de 20 superfamílias/clados (Tabela Suplementar 1), com porcentagem de similaridade variando de 64 a 100% e comprimento de alinhamento até 5.957 pb. O elemento BM1 foi o mais abundante, com 1.448 anotações, seguidos de Zenon_BM (650 anotações) e elementos *Gypsy* (450 anotações).

Enquanto isso, o alinhamento blastn do transcriptoma montado ao banco de dados ArTEDB permitiu a identificação de 11.235 sequências de TEs, distribuídas em 1.256 diferentes identificações. As sequências de maior identificação se deram com sequências de *Bombyx mori* (Lepidoptera: Bombycidae) (9.772 sequências), seguidos de *Melitaea cinxia* Lepidoptera: Nymphalidae) (410 sequências) e *Locusta migratoria* (Orthoptera: Acrididae) (287 sequências), variando de 63 a 100% de identificação (Tabela Suplementar 2). A sequência STRG.14683.1 apresentou o maior comprimento de alinhamento (8.264 pb), com 99,7% de identificação e caracterizado como um elemento da ordem LTR, superfamília *Bel-Pao* BEL-13 BomMor LTR/Pao bm 1226 (e-value 0.0).

A busca de domínios proteicos na base de dados do Dfam resultou em um total de 3.117 transcritos com ao menos um domínio relacionado a transposição de TEs. A frequência das principais famílias e domínios de proteína estão identificados na Tabela Suplementar 3. Os elementos identificados mais abundantes tiveram maior similaridade com sequências de borboletas do gênero *Heliconius* (Lepidoptera: Nymphalidae), sendo as mais numerosas CR1-Zenon-2_Diul e CR1-Zenon-4_Diul, e com insetos da espécie *Sitophilus oryzae* (Coleoptera: Curculionidae), sendo a sequência a TE0535_SO2_FAM0152 mais abundante e pertencente a ordem dos retrotransposons e da ordem LINE, superfamília *I*.

A busca no Gypsy Database (LTRs e TIRs Seq. DB) permitiu a identificação de 240 sequências de retrotransposons da ordem LTR e de 8 transposons de DNA da ordem TIR, com elementos variando de 73% a 100% de identificação (Tabela Suplementar 4). Os TEs LTRs mais abundantes foram LTR3'_Mdg1, LTR3'_Del e LTR3'_Skipper, enquanto dentro da ordem TIR foram identificados somente os TEs BmGINGER2a e HmGINGER1-3.

A busca por domínios proteicos característicos de TEs no Pfam, baseada em *Hidden Markov Model*, resultou em um total de 699 transcritos agrupados em 17 domínios relacionados a transposição de TEs (Tabela Suplementar 5). Os domínios mais frequentes foram: *Reverse transcriptase (RNA-dependent DNA polymerase)* (RVT_1), *Endonuclease-reverse transcriptase* (Exo_endo_phos_2) e *DDE superfamily endonuclease* (DDE_Tnp_4) (Figura 1).

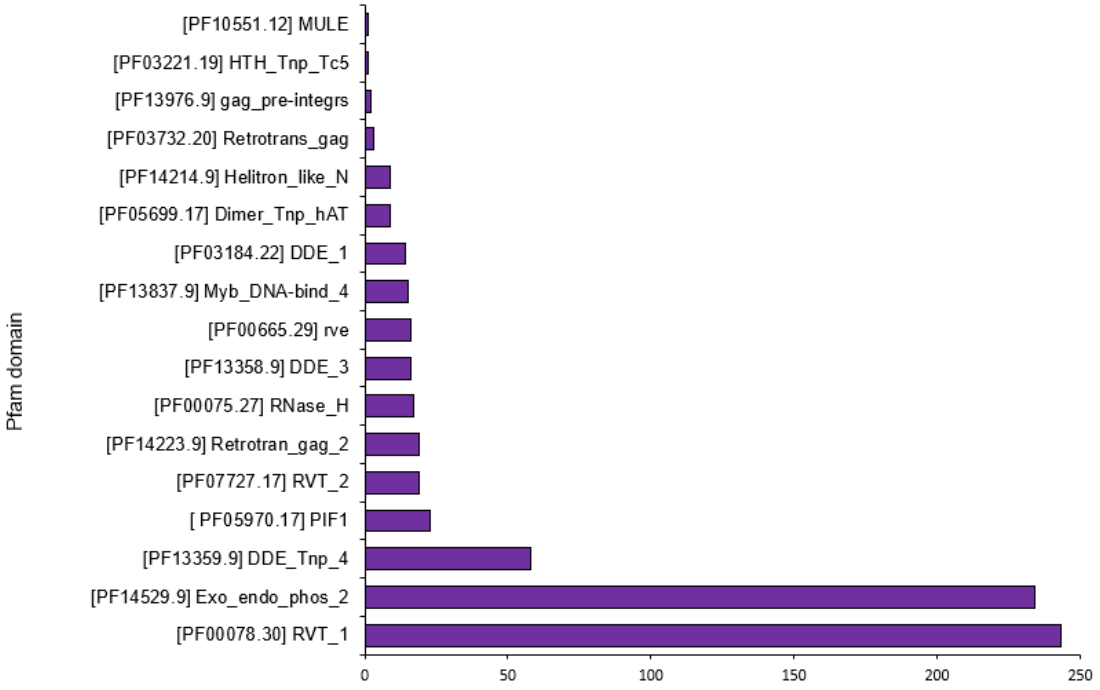

Figura 1. Domínios proteicos Pfam característicos de elementos transponíveis nas sequências do transcriptoma montado de *B. mori* e o número de transcritos relacionados a cada domínio. O domínio RVT_1 foi o mais numeroso.

A anotação de TEs usando o blastn mostrou que 7.910 sequências têm correspondência com os elementos do genoma de referência identificado pelo RepeatModeler2, sendo 3.614 delas distribuídos em 13 superfamílias conhecidas (Tabela Suplementar S6). A ordem mais abundante de TEs foi a ordem dos retrotransposons LINE (sendo os TEs mais abundantes *CR1-Zenon, Jockey* e *RTE-BovB*), seguida de transposons de DNA da ordem TIR (sendo *P, TcMar-Mariner* e *TcMar-Tc1* os mais abundantes) (Tabela Suplementar 6).

A reunião de todas as informações obtidas em diferentes bancos de dados (Repbase, ArTEDB, Gypsy, Dfam, Pfam e RepeatModeler2) permitiu a identificação de

11.697 sequências semelhantes a TEs no transcriptoma de *B. mori*, que apresentaram identificação com pelo menos um banco de dados. A maioria das sequências relacionadas a TEs puderam ser classificadas em diferentes categorias de elementos genéticos móveis (classes, ordens e superfamílias/clados) (Figuras 2 e 3, Tabela Suplementar 7).

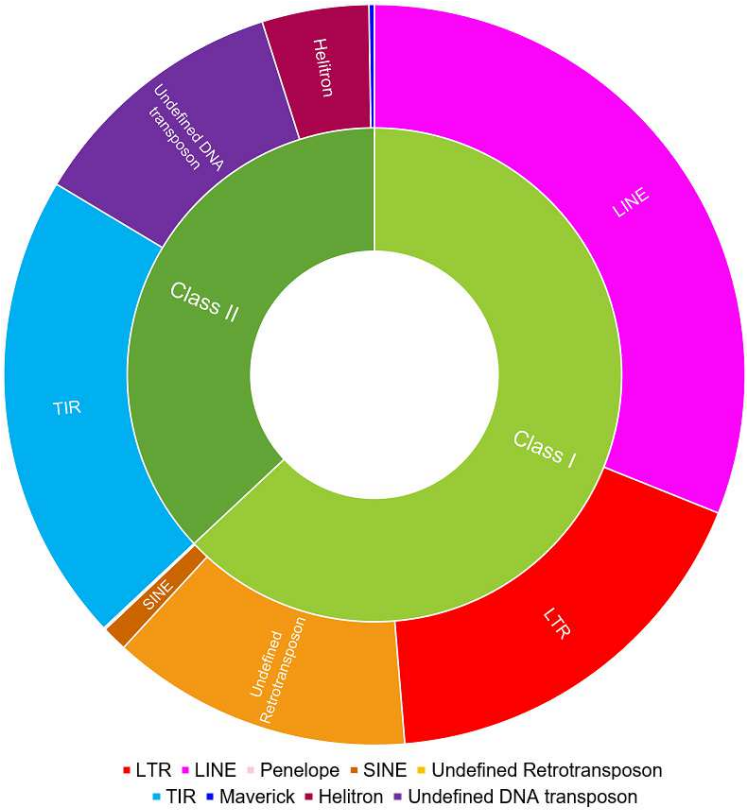

Figura 2. Distribuição de sequências relacionadas a TEs identificados no transcriptoma de *B. mori* em duas classes, elementos de classe I (retrotransposons) e elementos de classe II (DNA transposons), e em suas respectivas ordens. Os elementos de Classe I foram predominantes.

Os retrotransposons representaram aproximadamente 79% dos transcritos totais que foram classificados em diferentes clados (8.339 transcritos), com destaque para as ordens LINE e LTR, que representaram 50 e 28%, respectivamente. Dentre os retrotransposons, a ordem LINE além de mais numerosa também foi a mais diversa, abrangendo 12 superfamílias/clados (Figura 3). Enquanto isso, na classe dos DNA transposons, que representaram em torno de 21% dos transcritos identificados, a ordem a ordem TIR compreendeu 56% dos DNA transposons e sendo a ordem mais diversa (17 superfamílias/clados), seguida da ordem Helitron, que representou em torno de 13% dos DNA transposons.

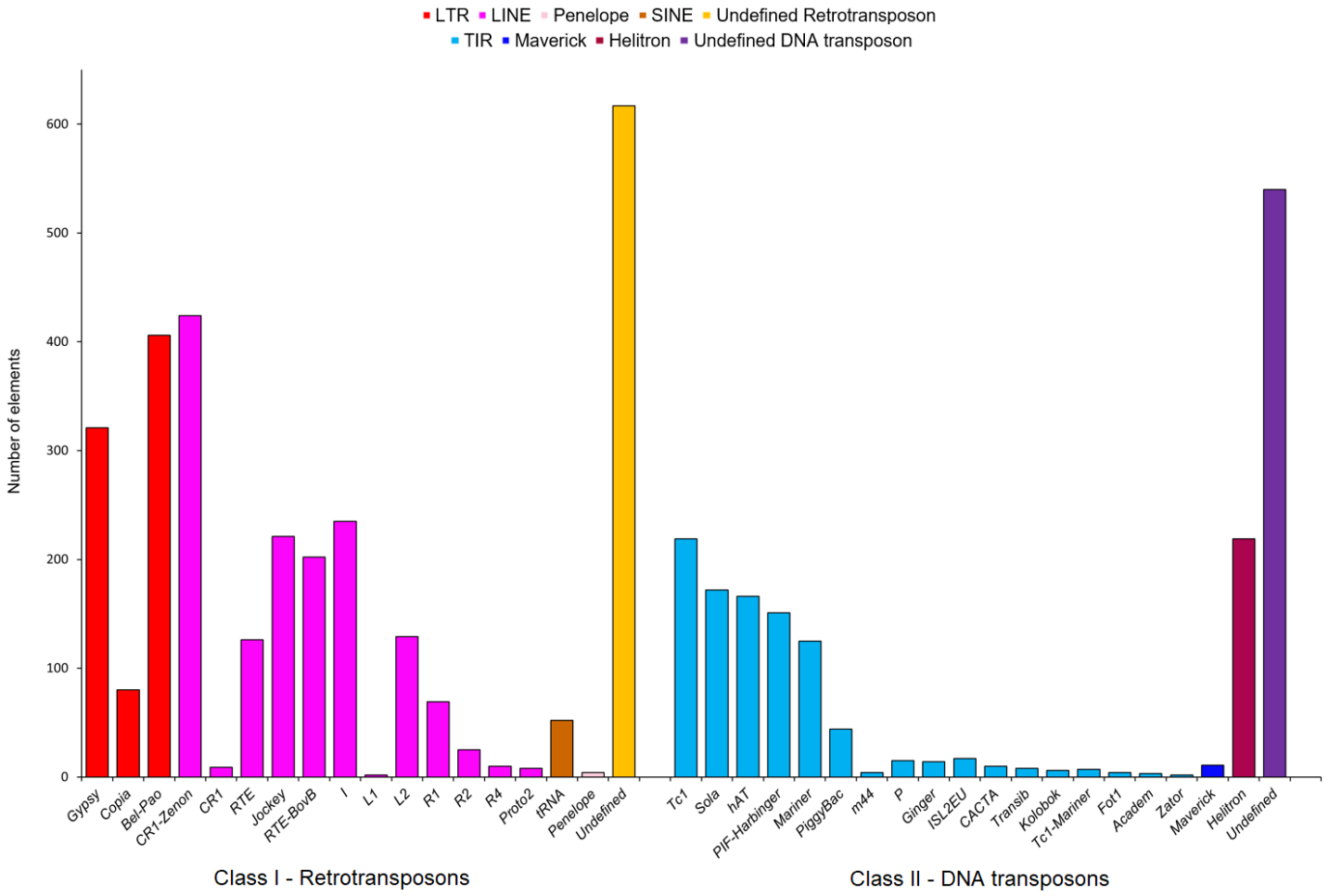

Figura 3. Distribuição de sequências relacionadas a TEs identificados no transcriptoma de *B. mori*. No eixo vertical, o número de sequências correspondentes a cada elemento. No eixo horizontal, a divisão entre duas classes (I e II) e classificação dos elementos em superfamílias/clados.

3.2 Expressão diferencial

A análise de expressão diferencial foi realizada com um total de 28.683 transcritos do transcriptoma montado pelo StringTie. Como resultado, 3.115 sequências foram diferencialmente expressas (padj < 0,05; Log2FoldChange ≥ 2,0). Destes, 1.696 foram *down* reguladas e 1.419 *up* reguladas nas lagartas criadas a 34 ºC em relação às criadas a 26 ºC. Dentre as sequências diferencialmente expressas, 784 sequências apresentaram pelo menos uma identificação com sequências TEs nos bancos de dados utilizados, sendo 227 sequências de TEs *up* reguladas e 557 sequências *down* reguladas (Tabela Suplementar 8, Figuras 4 e 5). Os retrotransposons também representaram a maioria das sequências diferencialmente expressas identificadas, aproximadamente 57%.

Entre os transcritos *up* regulados, a superfamília de retrotransposon LTR *Bel-Pao* foi a mais numerosa, compreendendo 31 sequências. Entre os transcritos *down* regulados, a superfamílias *Bel-Pao* e *Gypsy* da ordem LTR foram as mais numerosas, com 42 e 34 sequências, respectivamente. Enquanto isso, a superfamília *Copia*, pertencente a mesma ordem, apresentou somente uma sequência.

Enquanto isso, somente 29 sequências *up* reguladas foram identificadas como DNA transposons. A ordem TIR foi a mais abundante, e o clado Tc1 o mais numeroso da ordem, com 6 sequências. Entre os transcritos *down* regulados, a quantidade de transcritos DNA transposons foi quase quatro vezes maior (114 sequências), sendo o clado *Sola* o mais numeroso (27 sequências), seguido das superfamílias *hAT* e *Helitron* (12 transcritos cada).

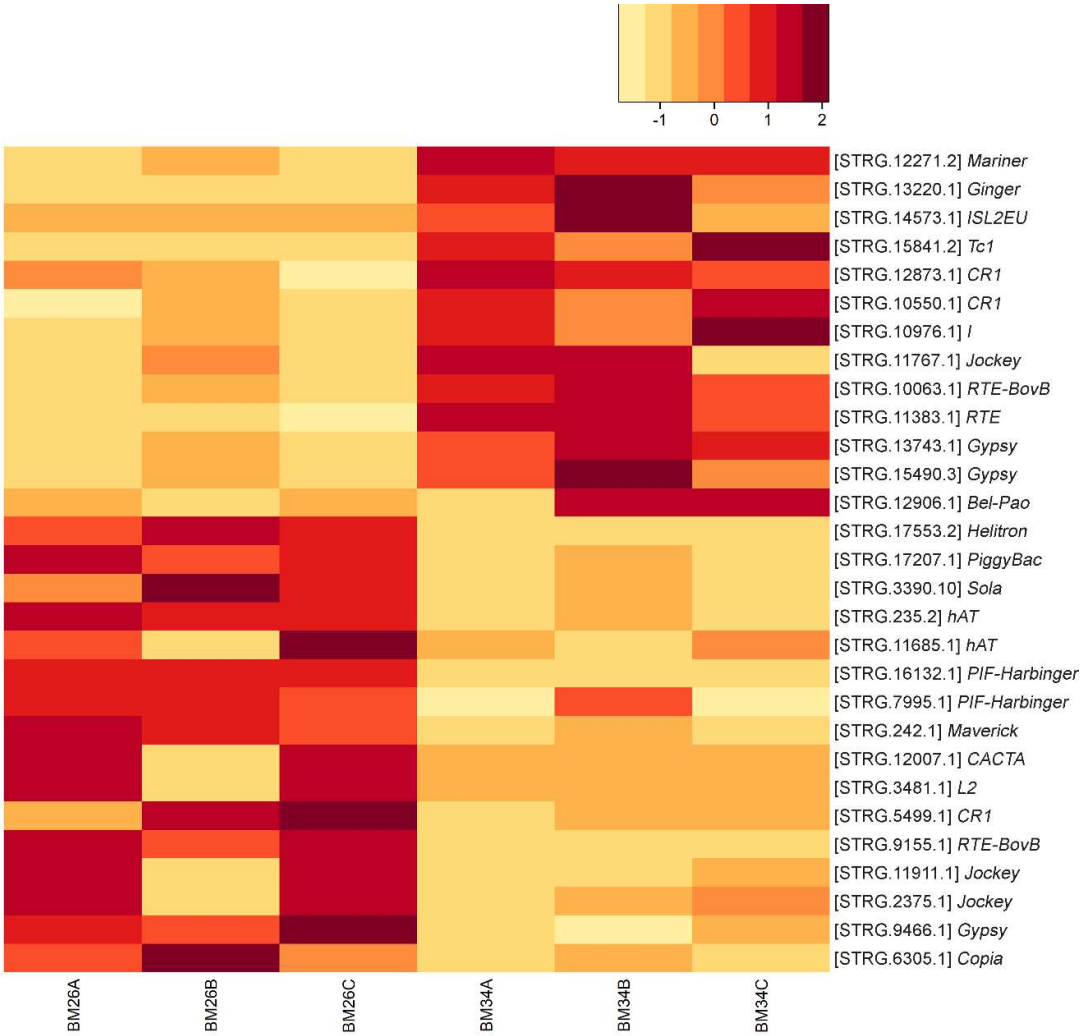

Figura 4. Representação de elementos transponíveis diferencialmente expressos no transcriptoma de *B. mori*. Os níveis de RNAm são representados pelas contagens estimadas normalizadas em DESeq2 (rlog). As diferentes tonalidades de vermelho indicam genes *up* (em tons mais escuros) e *down* regulados (destacados em tons mais claros) em cada biblioteca. Os nomes putativos dos genes estão à direita com as identificações dos transcritos correspondentes.

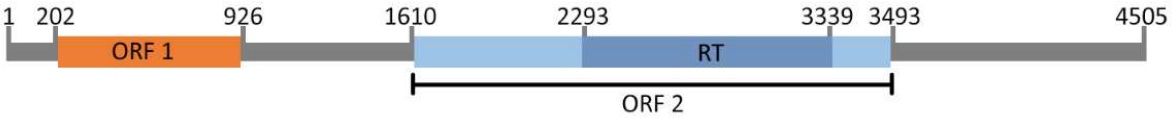

Figura 5. Representação esquemática da organização dos domínios de codificação de proteínas na sequência STRG 9105.1 (*fibrohexamerin*) em *B. mori*. A sequência foi representada em número de aminoácidos. Abreviações: RT = transcriptase reversa; ORF = *Open Reading Frame.*

4. Discussão

No presente estudo, 11.697 transcritos apresentaram semelhança com pelo menos um dos seis bancos de dados utilizados na pesquisa desses elementos, representando 40% dos transcritos totais montados. Esses números elevados de transcritos com semelhanças a TEs se dão pois em muitos casos, as sequências de TEs são encontradas de forma parcial nos genomas e transcriptomas (Osanai-Futahashi et al., 2008). A busca de TEs por homologia utilizando diferentes bancos de dados permite que a quantidade e a qualidade das informações sejam maiores, principalmente quando se estuda elementos com estrutura e sequências tão variáveis. Ademais, a utilização de diferentes bancos de dados também se justifica pela quantidade reduzida de informações sobre TEs em espécies de invertebrados em comparação com espécies de vertebrados, especialmente mamíferos. No presente estudo foi possível observar que muitos transcritos não puderam ser caracterizados pois suas informações eram desconhecidas (Figuras 2 e 3, Tabela Suplementar 7).

Os elementos genéticos móveis representam uma parte significativa no genoma de eucariotos. Em *B. mori*, por exemplo, esses elementos representam cerca de 46% da composição do genoma (Kawamoto et al., 2019), e muitas vezes podem ser transcricionalmente ativos. Com o avanço das tecnologias e os sequenciamentos de nova geração (NGS, do inglês *Next Generation Sequencing*), os transcriptomas obtidos por RNA-Seq tem permitido um melhor entendimento da expressão de TEs com uma variedade de agentes de estresse (Mombach et al., 2022) e proporcionando uma melhor identificação e caracterização desses elementos. Diversos estudos sugerem que situações de estresse, biótico ou abiótico, podem estimular a expressão ou mobilização desses elementos em diversos organismos. Em *Drosophila*, por exemplo, um inseto com uma das melhores qualidades de genomas anotados e um dos conjuntos de sequências TE mais bem caracterizados, já é sugerido que o estresse não leva à ativação geral de TEs, mas diferentes situações de estresse podem ativar ou reprimir conjuntos de TEs (Mombach et al., 2022). Dessa maneira, fazer uma investigação dos transcritos de *B. mori*, uma espécie amplamente estudada como inseto modelo e protagonista da cadeira produtora de seda no Brasil e no mundo se torna algo muito relevante.

A proporção de retrotransposons e DNA transposons (79 e 21% respectivamente) das sequências classificadas é similar às encontradas em outro estudo de *B. mori* (Osanai-Futahashi et al., 2008). Os retrotransposons normalmente constituem a maior porção repetitiva nos genomas eucariotos devido ao seu mecanismo de mobilização, onde, a cada

ciclo de replicação uma nova cópia do elemento é produzida (Kazazian, 2004; Finnegan, 2012).

Os retrotransposons, ou elementos de classe I são subdivididos em cinco ordens, sendo elas LTR, DIRS, PLE, LINE e SINE. Transcritos da ordem DIRS não foram identificados no transcriptoma de *B. mori*, A ausência da identificação desses elementos também foi observada em outros estudos (Osanai-Futahashi et al., 2008; Kawamoto et al., 2019). Os elementos LTRs são caracterizados por suas repetições terminais longas flanqueando diretamente uma região de codificação interna e seu mecanismo de transposição é compartilhado com os retrovírus (Wicker et al., 2007; Finnegan, 2012). Das cinco superfamílias de retrotransposons, somente três foram observadas no presente estudo (*Bel-Pao*, *Copia* e *Gypsy*). Os elementos *Gypsy* foram os mais abundantes (551 transcritos), seguidos de *Bel-Pao* (407 transcritos) e *Copia* (80 transcritos). Sendo assim, os LTRs foram o segundo grupo mais numeroso de retrotransposons em *B. mori.* Esses resultados diferem dos estudos do genoma de *B. mori* de Osanai-Futahashi et al. (2008) e Kawamoto et al. (2019), onde a ordem LTR foi a mais e menos que outras ordens, respectivamente.

Os elementos LINE, segundo a classificação proposta por Wicker et al. (2007), eram classificados somente em cinco grandes superfamílias: *R2, L1*, *RTE*, *I* e *Jockey*. No entanto, devido à ausência de um consenso entre as classificações, além do avanço nos estudos dos TEs, novas classificações foram propostas (Kapitonov et al., 2009; Piégu et al., 2015), e isso permitiu que pudéssemos classificar as sequências relacionadas a TEs no presente estudo em 12 diferentes clados: *CR1-Zenon*, *CR1*, *RTE*, *RTE-BovB*, *I*, *L1*, *L2*, *R1*, *R2*, *R4* e *Proto2*, visto que alguns bancos de dados já possuem essas classificações mais abrangentes. Os elementos da ordem LINE foram os mais abundantes no transcriptoma de *B. mori* no presente estudo. Esse resultado é similar ao estudo do genoma de *B. mori* de Osanai-Futahashi et al. (2008). No entanto, em comparação com outras espécies de insetos, como *Anopheles gambiae* (Diptera: Culicidae) e *Drosophila melanogaster* (Diptera: Drosophilidae), esses resultados diferem consideravelmente. Em *A. gambiae*, LTRs e SINEs compreendem a maior porção de sequências de TEs, enquanto em *D. melanogaster*, a maior porção de sequências de TEs pertencem a ordem dos LTRs, Diferentemente da composição do genoma de *A. gambiae* e do estudo de Kawamoto et al. (2019) em *B. mori*, os elementos SINEs em *B. mori* foram os menos abundantes dentre os retrotransposons no presente estudo, seguidos dos elementos PLE.

A classe de transposons de DNA, também chamada de classe II, é subdividida em duas subclasses, sendo a Subclasse 1 constituída pelas ordens TIR e Crypton, que compreendem os TEs que utilizam um mecanismo de "cortar e colar" e são caracterizados por repetições terminais invertidas (Wicker et al., 2007). No transcriptoma de *B. mori*, foram identificadas as superfamílias e outros clados da ordem TIR: *Tc1*, *Sola*, *hAT*, *PIF-Harbinger, Mariner*, *PiggyBac*, *P*, *Ginger*, *m44*, *ISL2EU*, *Fot1*, *CACTA*, *Transib*, *Kolobok*, *Tc1- Mariner, Academ* e *Zator*. Os clado *Tc1*, pertencente a superfamília Tc1-Mariner, e a superfamília *Sola* representaram os grupos de TEs mais abundantes, com 221 e 173 sequências anotadas, respectivamente. O resultado condiz com o estudo do genoma de *B. mori* Osanai-Futahashi et al. (2008), onde os elementos da superfamília *Tc1-Mariner* foram os mais abundantes, e condiz com outros estudos anteriores da espécie (Robertson & Asplund, 1996; Tomita et al., 1997; Robertson e Walden, 2003). Enquanto isso, a subclasse II dos transposons de DNA compreende as ordens Helitron e Maverick (Polintons), que utilizam um mecanismo de transposição que envolve a replicação sem clivagem da dupla fita. No presente estudo, foram identificadas 198 sequências com semelhança a TEs Helitrons, enquanto apenas 11 sequências apresentaram similaridade com elementos Maverick.

No presente estudo, foi possível observar a possibilidade da ocorrência de inserções de TEs em genes específicos. O elevado número de sequências com identificação com TEs e a anotação das sequências que codificam diferentes funções em *B. mori* foi observado, em especial nos genes diferencialmente expressos (*craniofacial development protein 2-like, ecdysone 20-hydroxylase, cytocrome P450, fibrohexamerin,* entre outros).

Em relação a análise de expressão diferencial, também foi possível observar que a quantidade de transcritos com semelhança a TEs regulados positivamente foi 2,5 vezes menor que a quantidade de transcritos *down* regulados (Tabela Suplementar 8). Transcritos *zinc finger BED domain-containing protein 4-like down* regulados foram identificados e classificados como DNA transposons da família *hAT* a partir dos dados dos seis bancos de dados utilizados. A relação entre fatores de transcrição *zinc fingers* e resistência a temperatura já foi estabelecida em alguns estudos de *B. mori* (Li et al., 2012; Guo et al., 2018). A identificação das sequências desses transcritos diferencialmente expressos com sequências de DNA transposons no presente estudo se dá possivelmente devido a possibilidade das proteínas do domínio ZBED terem derivado da superfamília de DNA transposons *hAT* (Hayward et al., 2013).

Enquanto isso, em busca de transcritos envolvidos com a resistência a temperatura, o transcrito *down* regulado STRG.9466.1 (*heat shock 70 kDa protein*) foi identificado. Já é conhecido que a transcrição do gene Hsp70, que é ativada por fatores de choque térmico (HSFs) em situações de estresse, torna-se rapidamente atenuada quando as condições de estresse diminuem, para mitigar a influência que Hsp70 tem no crescimento e divisão celular (Feder et al., 1992; Morimoto, 1998). Uma vez produzida, a Hsp70 interage com a Hsp90 "distraindo-a" de sua função regulatória de piwi-RNAs, que tem papel conhecido de reprimir e silenciar transposons por mecanismos transcricionais ou pós-transcricionais e mantêm a integridade do genoma, e permitindo a expressão de TEs (Cappucci et al., 2019). Sendo assim, em *Drosophila*, já é conhecido que as condições do estresse, que consequentemente refletem no nível de expressão de Hsp70, modulam a consequente ativação de TEs.

Neste estudo apresentamos a identificação e a caracterização de TEs no transcriptoma de *Bombyx mori* da linhagem comercial brasileira utilizando seis diferentes bancos de dados com sequências de elementos genéticos móveis. Nossos resultados revelaram uma grande diversidade de TEs na espécie. Além disso, avaliamos a expressão relativa de elementos genéticos móveis em lagartas criadas a 26º e 34 ºC durante todo o quinto instar. Nossos dados contribuem para o conhecimento da composição desses elementos dentro do transcriptoma de *B. mori* e evidenciam um elevado número de sequências com identificação com TEs que representam diferentes funções em *B. mori*. Ademais, a temperatura pode estar relacionada com a modulação de sequências semelhantes a TEs.

Agradecimentos

Agradecemos à Fiação de Seda Bratac pelo fornecimento das larvas e apoio logístico; à Unidade Gestora do Fundo Paraná (UGF), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento; à Universidade Estadual de Londrina, à Fundação Araucária e à Financiadora de Estudos e Projetos (FINEP) pelo suporte.

Referências

Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile Dna*, *6*, 1-6. https://doi.org/10.1186/s13100-015-0041-9

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin., H. L., Macfarlan, T. S., Mager, D. L., & Feschotte, C. (2018). Ten things you should know about transposable elements. *Genome biology*, *19*(1), 1-12. https://doi.org/10.1186%2Fs13059-018-1577-z

Capy, P., Gasperi, G., Biémont, C., & Bazin, C. (2000). Stress and transposable elements: co-evolution or useful parasites?. *Heredity*, *85*(2), 101-106. https://doi.org/10.1046/j.1365-2540.2000.00751.x

Cappucci, U., Noro, F., Casale, A. M., Fanti, L., Berloco, M., Alagia, A. A., Grassi, L., Pera, L. L., & Pimpinelli, S. (2019). The Hsp70 chaperone is a major player in stress-induced transposable element activation. *Proceedings of the National Academy of Sciences*, *116*(36), 17943-17950. https://doi.org/10.1073/pnas.1903936116

Casacuberta, E., & González, J. (2013). The impact of transposable elements in environmental adaptation. *Molecular ecology*, *22*(6), 1503-1517. https://doi.org/10.1111/mec.12170

Chénais, B., Caruso, A., Hiard, S., & Casse, N. (2012). The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. *Gene*, *509*(1), 7-15. https://doi.org/10.1016/j.gene.2012.07.042

De Souza, F. S., Franchini, L. F., & Rubinstein, M. (2013). Exaptation of transposable elements into novel cis-regulatory elements: is the evidence always strong?. *Molecular biology and evolution*, *30*(6), 1239-1251. https://doi.org/10.1093/molbev/mst045

Drongitis, D., Aniello, F., Fucci, L., & Donizetti, A. (2019). Roles of transposable elements in the different layers of gene expression regulation. *International Journal of Molecular Sciences*, *20*(22), 5755. https://doi.org/10.3390%2Fijms20225755

Fablet, M., & Vieira, C. (2011). Evolvability, epigenetics and transposable elements. *Biomol Concepts*. https://doi.org/10.1515/BMC.2011.035

Feder, J. H., Rossi, J. M., Solomon, J., Solomon, N., & Lindquist, S. (1992). The consequences of expressing hsp70 in *Drosophila* cells at normal temperatures. *Genes & development*, *6*(8), 1402-1413. https://doi.org/10.1101/gad.6.8.1402

Finnegan, D. J. (2012). Retrotransposons. *Current Biology*, *22*(11), R432-R437. https://doi.org/10.1016/j.cub.2012.04.025

Guo, H., Huang, C., Jiang, L., Cheng, T., Feng, T., & Xia, Q. (2018). Transcriptome analysis of the response of silkworm to drastic changes in ambient temperature. *Applied microbiology and biotechnology*, *102*(23), 10161-10170. https://doi.org/10.1007/s00253-018-9387-5

Hayward, A., Ghazal, A., Andersson, G., Andersson, L., & Jern, P. (2013). ZBED evolution: repeated utilization of DNA transposons as regulators of diverse host functions. *PloS one*, *8*(3), e59940. https://doi.org/10.1371/journal.pone.0059940

Kapitonov, V. V., Tempel, S., & Jurka, J. (2009). Simple and fast classification of non-LTR retrotransposons based on phylogeny of their RT domain protein sequences. *Gene*, *448*(2), 207-213. https://doi.org/10.1016/j.gene.2009.07.019

Kawamoto, M., Jouraku, A., Toyoda, A., Yokoi, K., Minakuchi, Y., Katsuma, S., Fujiyama, A., Kiuchi, T., Yamamoto, K., & Shimada, T. (2019). High-quality genome assembly of the silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, *107*, 53-62. https://doi.org/10.1016/j.ibmb.2019.02.002

Kazazian Jr, H. H. (2004). Mobile elements: drivers of genome evolution. *Science*, 303(5664), 1626-32. https://doi.org/10.1126/science.1089670

Li, J., Moghaddam, S. H. H., Du, X., Zhong, B. X., & Chen, Y. Y. (2012). Comparative analysis on the expression of inducible HSPs in the silkworm, *Bombyx mori*. *Molecular biology reports*, 39, 3915-3923. https://doi.org/10.1007/s11033-011-1170-y

Llorens, C., Futami, R., Covelli, L., Domínguez-Escribá, L., Viu, J. M., Tamarit, D., Aguilar-Rodríguez, J., Vicente-Ripolles, M., Fuster, G., Bernet, G.P., Maumus, F., Munoz-Pomer, A., Sempere, J. M., Latorre, A., & Moya, A. (2010). The Gypsy Database (GyDB) of mobile genetic elements: release 2.0. *Nucleic acids research*, *39*(suppl_1), D70-D74. https://doi.org/10.1093/nar/gkq1061

Lopes, T. B. F., Aguiar, R. C. M., De Souza, R. F., Nascimento, C. C., Dionísio, J. F., Mantovani, M. S., Semprebon, S. C., Zurovec, M., & Da Rosa, R. (2023). Influência da variação de temperatura na expressão gênica e produção de casulos em *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae). (Em fase de pré-publicação).

Mombach, D. M., da Fontoura Gomes, T. M. F., & Loreto, E. L. S. (2022). Stress does not induce a general transcription of transposable elements in *Drosophila*. *Molecular Biology Reports*, 1-8. https://doi.org/10.1007/s11033-022-07839-7

Morimoto, R. I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes & development*, *12*(24), 3788-3796. https://doi.org/10.1101/gad.12.24.3788

Osanai-Futahashi, M., Suetsugu, Y., Mita, K., & Fujiwara, H. (2008). Genome-wide screening and characterization of transposable elements and their distribution analysis in the silkworm, *Bombyx mori*. *Insect biochemistry and molecular biology*, *38*(12), 1046-1057. https://doi.org/10.1016/j.ibmb.2008.05.012

Piegu, B., Bire, S., Arensburger, P., & Bigot, Y. (2015). A survey of transposable element classification systems–a call for a fundamental update to meet the challenge of their diversity and complexity. *Molecular phylogenetics and evolution*, *86*, 90-109. https://doi.org/10.1016/j.ympev.2015.03.009

Rey, O., Danchin, E., Mirouze, M., Loot, C., & Blanchet, S. (2016). Adaptation to global change: a transposable element–epigenetics perspective. *Trends in ecology & evolution*, *31*(7), 514-526. https://doi.org/10.1016/j.tree.2016.03.013

Robertson, H. M., & Asplund, M. L. (1996). Bmmarl: a basal lineage of the mariner family of transposable elements in the silkworm moth, *Bombyx mori*. *Insect biochemistry and molecular biology*, *26*(8-9), 945-954.

Robertson, H. M., & Walden, K. K. O. (2003). Bmmar6, a second mori subfamily mariner transposon from the silkworm moth *Bombyx mori*. *Insect molecular biology*, *12*(2), 167-171.

Slotkin, R. K., & Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature reviews genetics*, *8*(4), 272-285. https://doi.org/10.1038/nrg2072

Tomita, S., Sohn, B. H., & Tamura, T. (1997). Cloning and characterization of a mariner-like element in the silkworm, *Bombyx mori*. *Genes & genetic systems*, *72*(4), 219-228. https://doi.org/10.1266/ggs.72.219

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, *8*(12), 973-982. https://doi.org/10.1038/nrg2165

Yokoi, K., Kimura, K., & Bono, H. (2022). Revealing Landscapes of Transposable Elements in *Apis* Species by Meta-Analysis. *Insects*, *13*(8), 698. https://doi.org/10.3390/insects13080698

Zamudio, N., & Bourc'His, D. (2010). Transposable elements in the mammalian germline: a comfortable niche or a deadly trap?. *Heredity*, *105*(1), 92-104. https://doi.org/10.1038/hdy.2010.53

7. CONCLUSÕES FINAIS

Os dados apresentados neste estudo são inovadores para o estudo de *B. mori* da linhagem comercial brasileira, evidenciando vias metabólicas importantes no processo de resposta a temperatura. Essas informações iniciais são relevantes para futuras estratégias que permitam o melhor desempenho da produção de seda no Brasil e no Estado do Paraná. Entre as principais contribuições, podemos citar:

- A criação de lagartas de *B. mori* a 34 ºC em comparação a lagartas mantidas a 26 ºC durante o quinto instar afetam diretamente o crescimento, desenvolvimento e o sistema imunológico desses insetos, tendo como consequência um efeito prejudicial na produção de casulos de seda;
- A caracterização do transcriptoma de *B. mori* utilizando diferentes bancos de dados de elementos genéticos móveis revelou que grande parte dos transcritos da espécie possuem identificação com sequências de TEs, e que a maior porção desses transcritos são retrotransposons da ordem LINE;
- A pesquisa de elementos transponíveis no transcriptoma de *B. mori* revelou uma diversidade considerável desses elementos em sequências transcritas dessa linhagem e contribuem significativamente para o conhecimento da composição desses elementos dentro do transcriptoma de *B. mori*.