

UNIVERSIDADE FEDERAL DE VIÇOSA

Assinaturas transcritômicas e proteômicas que sustentam a adaptação ninfal e a produção da espuma na cigarrinha-das-pastagens *Mahanarva spectabilis*

Monique da Silva Bonjour
Magister Scientiae

**VIÇOSA - MINAS GERAIS
2026**

MONIQUE DA SILVA BONJOUR

Assinaturas transcritômicas e proteômicas que sustentam a adaptação ninfal e a produção da espuma na cigarrinha-das-pastagens *Mahanarva spectabilis*

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica e Biotecnologia, para obtenção do título de *Magister Scientiae*.

Orientador: Humberto J. de O. Ramos

Coorientadora: Maria G. de A. Oliveira

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

B715a
2026 Bonjour, Monique da Silva, 2000-
Assinaturas transcritômicas e proteômicas que sustentam a
adaptação ninfal e a produção da espuma na
cigarrinha-das-pastagens *Mahanarva spectabilis* / Monique da
Silva Bonjour. – Viçosa, MG, 2026.

1 dissertação eletrônica (71 f.): il. (algumas color.).

Inclui apêndice.

Orientador: Humberto Josué de Oliveira Ramos.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Bioquímica e Biologia Molecular, 2026.

Referências bibliográficas: f. 62-68.

DOI: <https://doi.org/10.47328/ufvbbt.2026.483>

Modo de acesso: World Wide Web.

1. Cigarrinha (Inseto) - Genética. 2. Regulação da expressão
gênica. 3. Proteínas. 4. Enzimas. 5. Pastagens - Doenças e
pragas. 6. Relação inseto-planta. I. Ramos, Humberto Josué de
Oliveira, 1969-. II. Universidade Federal de Viçosa.
Departamento de Bioquímica e Biologia Molecular. Programa de
Pós-Graduação em Bioquímica e Biotecnologia. III. Título.

CDD 22. ed. 595.75413

MONIQUE DA SILVA BONJOUR

Assinaturas transcritômicas e proteômicas que sustentam a adaptação ninfal e a produção da espuma na cigarrinha-das-pastagens *Mahanarva spectabilis*

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica e Biotecnologia, para obtenção do título de *Magister Scientiae*.

APROVADA: 3 de agosto de 2026.

Assentimento:

Monique da Silva Bonjour
Autora

Humberto Josue de Oliveira Ramos
Orientador

Essa dissertação foi assinada digitalmente pela autora em 18/08/2026 às 20:28:18 e pelo orientador em 18/08/2026 às 20:52:41. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **F8FU.HIB3.TINN** e clique no botão 'Validar documento'.

Aos meus pais, Isabel e José Milton, por todo o amor, apoio incondicional e pelos inúmeros sacrifícios realizados para que eu pudesse estudar e alcançar meus objetivos. Mesmo sem terem tido as mesmas oportunidades, sempre valorizaram o conhecimento e foram os meus maiores incentivadores. A vocês, que são minha base e minha maior inspiração, dedico.

AGRADECIMENTOS

Ao meu orientador, Prof.Dr. Humberto Josué de Oliveira Ramos, pela orientação segura, disponibilidade e valiosas contribuições ao longo de todo o desenvolvimento deste trabalho. Sua postura profissional e comprometimento com a pesquisa foram fundamentais para a minha formação acadêmica.

Aos meus coorientadores, Dr. Angelo José Rinaldi, Dr. Eulálio Gutemberg Dias dos Santos e Profa.Dra Maria Goreti de Almeida Oliveira, agradeço pelas valiosas contribuições científicas, pelas discussões enriquecedoras e pela disponibilidade em compartilhar conhecimentos ao longo do desenvolvimento deste trabalho. Suas sugestões, apoio técnico e olhar crítico foram fundamentais para o aprimoramento das análises e para a qualidade desta dissertação.

Aos colegas do Laboratório de Enzimologia e Bioquímica de Proteínas e Peptídeos, pelo apoio técnico, pelas trocas de conhecimento e pelo ambiente de trabalho colaborativo, essenciais para a execução das atividades experimentais, Em especial, à Gabriely Teixeira Bhering Faria, pela dedicação, comprometimento e genuíno entusiasmo pela ciência, que muito contribuiu para o desenvolvimento deste trabalho.

Ao Programa de Pós-Graduação em Bioquímica e Biotecnologia da Universidade Federal de Viçosa, pela excelência na formação oferecida, pela infraestrutura disponibilizada e pelo suporte institucional necessário à realização desta pesquisa. À Embrapa Gado de Leite e ao Núcleo de Análises de Biomoléculas (NUBIOMOL), pelo suporte dado à realização dos experimentos.

À FINEP, SisNANO/MCTI e Instituto Nacional de Ciências e Tecnologia em Interações Planta-Praga (INCTIPP) pelo apoio financeiro que viabilizou a realização deste estudo.

A todos que contribuíram, direta ou indiretamente, para a concretização desta dissertação, registro meus sinceros agradecimentos.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

RESUMO

BONJOUR, Monique da Silva, M.Sc., Universidade Federal de Viçosa, agosto de 2026. **Assinaturas transcritômicas e proteômicas que sustentam a adaptação ninfal e a produção da espuma na cigarrinha-das-pastagens *Mahanarva spectabilis***. Orientador: Humberto Josue de Oliveira Ramos. Coorientadora: Maria Goreti de Almeida Oliveira.

A cigarrinha-das-pastagens *Mahanarva spectabilis* (Distant, 1909) (Hemiptera: Cercopidae) é uma importante praga de gramíneas forrageiras na América do Sul, onde suas ninfas causam danos às pastagens ao se alimentarem da seiva do xilema e produzirem uma espuma característica que as protege contra estressores ambientais. Para investigar a base molecular dessa adaptação, integramos a análise de RNA-seq de ninfas com a proteômica por LC-MS/MS da glândula de Batelli, principal fonte da secreção da espuma. A montagem de novo de 100.666 unigenes revelou ampla diversidade funcional, com forte representação de enzimas de desintoxicação (CYP450s, GSTs, UGTs e carboxilesterases), transportadores e bombas iônicas, proteínas cuticulares e genes relacionados ao estresse e à imunidade. Aproximadamente 16% dos loci apresentaram splicing alternativo, especialmente em famílias gênicas de desintoxicação, quimiorrecepção e osmorregulação, evidenciando variabilidade transcriptômica. As predições de peptídeo sinal e de proteínas secretadas identificaram 168 proteínas secretadas candidatas de alta confiança, incluindo enzimas de desintoxicação, proteases, proteínas estruturais e fatores relacionados à imunidade, várias das quais são consistentes com funções antimicrobianas e surfactantes atribuídas às espumas das cigarrinhas. O perfil proteômico da glândula de Batelli confirmou 500 proteínas, enriquecidas em chaperonas, enzimas metabólicas, vias de desintoxicação e componentes de osmorregulação, sendo as proteínas mais abundantes correspondentes a chaperonas Hsp70, ATP sintases, proteínas cuticulares e anidrases carbônicas. Em conjunto, esses resultados fornecem uma visão integrativa transcriptômica e proteômica das ninfas de *M. spectabilis*, destacando genes e proteínas associados à alimentação no xilema, à produção de espuma e a respostas potencialmente relacionadas à tolerância ao estresse ambiental. Esse conjunto abrangente de dados não apenas amplia a compreensão da biologia das cigarrinhas, mas também identifica alvos moleculares candidatos que podem subsidiar estratégias inovadoras para o controle dos estágios ninfais e a mitigação dos danos causados por essas pragas em sistemas forrageiros.

Palavras-chave: Defesa de plantas; Forrageiras; Secretoma; Interação

inseto-planta; Enzimas de detoxificação

ABSTRACT

BONJOUR, Monique da Silva, M.Sc., Universidade Federal de Viçosa, August, 2026. **Transcriptomic and proteomic signatures underlying nymphal adaptation and foam production in the forage pest *Mahanarva spectabilis***. Adviser: Humberto Josue de Oliveira Ramos. Co-adviser: Maria Goreti de Almeida Oliveira.

The spittlebug *Mahanarva spectabilis* (Distant, 1909) (Hemiptera: Cercopidae) is an important pest of forage grasses in South America, where its nymphs cause pasture damage by feeding on xylem sap and producing a characteristic foam that protects them against environmental stressors. To investigate the molecular basis of this adaptation, we integrated RNA-seq analysis of nymphs with LC-MS/MS proteomics of the Batelli gland, the primary source of foam secretion. De novo assembly of 100,666 unigenes revealed broad functional diversity, with strong representation of detoxification enzymes (CYP450s, GSTs, UGTs, carboxylesterases), transporters and ion pumps, cuticle proteins, and stress- and immunity-related genes. Nearly 16% of loci exhibited alternative splicing, particularly within detoxification, chemosensory, and osmoregulatory gene families, highlighting evidence of transcriptomic variability. Signal peptide and secreted protein predictions identified 168 high-confidence candidate secreted proteins, including detoxification enzymes, proteases, structural proteins, and immune-related factors, several of which are consistent with antimicrobial and surfactant-related functions for spittlebug foams. Proteomic profiling of the Batelli gland confirmed 500 proteins, enriched in chaperones, metabolic enzymes, detoxification pathways, and osmoregulatory components, with the most abundant proteins corresponding to Hsp70 chaperones, ATP synthases, cuticle proteins, and carbonic anhydrases. Together, these results provide an integrative transcriptomic and proteomic overview for *M. spectabilis* nymphs, highlighting genes and proteins associated with xylem feeding, foam production and responses potentially related to environmental stress tolerance. This comprehensive dataset not only advances the understanding of spittlebug biology but also identifies candidate molecular targets that may inform innovative strategies for controlling nymphal stages and mitigating spittlebug damage in forage systems.

Keywords: Plant defense; Forage crops; Secretome; Insect-plant interaction; Detoxification enzymes

Lista de ilustrações

Figura 1- Anotação funcional de unigenes de ninfas de <i>Mahanarva spectabilis</i> em múltiplos bancos de dados	33
Figura 2 - Anotação funcional dos unigenes do transcriptoma de ninfas de <i>Mahanarva spectabilis</i>	36
Figura 3 - Genes com e sem splicing alternativo nas diferentes categorias funcionais do transcriptoma de ninfas de <i>Mahanarva spectabilis</i> . ..	42
Figura 4 - Principais anotações funcionais do secretoma com base nos bancos Pfam (verde), KEGG (azul) e Gene Ontology (GO; laranja).	44
Figura 5 - Classificação funcional das proteínas identificadas na glândula de Batelli de ninfas de <i>Mahanarva spectabilis</i>	50
Figura 6 - Distribuição das proteínas por categoria funcional a partir da análise proteômica.	54

Lista de Tabelas

Tabela 1 - Resumo das categorias de transcritos identificadas no transcriptoma ninfal de <i>Mahanarva spectabilis</i> relacionadas à interação com a planta hospedeira e à adaptação ao ambiente.....	38
Tabela 2 - Principais categorias funcionais do secretoma previsto a partir da análise de RNA-seq de <i>Mahanarva spectabilis</i>	44
Tabela 3 - Proteínas secretadas por <i>Mahanarva spectabilis</i> previstas da glândula de Batelli, identificadas pela presença de peptídeos sinal N-terminais (SignalP/Phobius).....	46
Tabela 4 - 15 proteínas mais abundantes da glândula de Batelli a partir de análises proteômicas.....	52

Lista de abreviaturas e siglas

AS	Alternative splicing
BCA	Bicinchoninic acid
BLAST	Basic Local Alignment Search Tool
BP	Biological process
cDNA	Complementary DNA
COG	Clusters of Orthologous Groups
CSP	Chemosensory protein
CPR	Cuticular protein
CYP450	Cytochrome P450 monooxygenase
DDA	Data-Dependent Acquisition
DEG	Diferentially expressed gene
FDR	False discovery rate
FPKM	Fragments per kilobase of transcript per million mapped reads
GO	Gene Ontology
GST	Glutathione S-transferase
HCD	High-Energy Collision Dissociation
HSP	Heat shock protein
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC–MS/MS	Liquid chromatography coupled to tandem mass spectrometry
MF	Molecular function
MOFA	Multi-Omics Factor Analysis
mRNA	Messenger RNA
MS	Mass spectrometry
Na⁺/K⁺-ATPase	Sodium/potassium ATPase
NCE	Scaled normalized collision energy
NGS	Next-generation sequencing
OBP	Odorant-binding protein

ORF	Open reading frame
PBP	Pheromone-binding protein
PIB	Produto Interno Bruto
PCA	Principal component analysis
PDI	Protein disulfide isomerase
PGRP	Peptidoglycan recognition proteins
RIN	RNA integrity number
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
SLC	Solute carrier
SP	Signal peptide
SP3	Single-Pot Solid-Phase-enhanced Sample Preparation
SRA	Sequence Read Archive
SSR	Simple sequence repetition
TIC	Total ion chromatogram
TM	Transmembrane helices
TPM	Transcripts per million
v-ATPase	Vacuolar-type H ⁺ -ATPase

SUMÁRIO

1.	INTRODUÇÃO.....	14
2.	REVISÃO DE LITERATURA.....	16
2.1	Impactos da cigarrinha-das-pastagens na agropecuária.....	16
2.2	Aspectos biológicos de <i>Mahanarva spectabilis</i>	17
2.3	Glândula de Batelli e produção da espuma.....	18
2.4	O uso de tecnologias ômicas para a compreensão de sistemas biológicos na agricultura	19
3.	OBJETIVOS.....	21
3.1	Objetivo Geral	21
3.2	Objetivos Específicos	21
4.	METODOLOGIA.....	22
4.1	Insetos e extração de RNA.....	22
4.2	Construção de bibliotecas e sequenciamento ilumina.....	22
4.3	Processamento dos dados e controle de qualidade	23
4.4	Montagem <i>de novo</i> do transcriptoma	23
4.5	Análise e quantificação da expressão	24
4.6	Predição de peptídeo sinal e definição de candidatos ao secretoma	24
4.7	Análise de <i>splicing</i> alternativo	26
4.8	Análise proteômica da glândula de Batelli por LC/MS.....	27
4.8.1	Extração de proteínas e digestão.....	27
4.8.2	Identificação de proteínas por LC-MS/MS.....	28
5.	RESULTADOS	30
5.1	Qualidade do sequenciamento e estatística da montagem dos dados do RNA-seq	30
5.2	Estatísticas de montagem e mapeamento <i>de novo</i> do transcriptoma de ninfas de <i>Mahanarva spectabilis</i>	31

5.3	Anotação funcional dos unigenes.....	32
5.4	Categorização funcional dos unigenes.....	33
5.5	Assinaturas transcriptômicas envolvidas na interação planta- inseto	38
5.6	Predição de <i>splicing</i> alternativo no transcriptoma da ninfa.....	40
5.7	Predição do secretoma (SignalP e Phobius) através dos transcritos do RNA-seq.....	42
5.8	Anotação funcional do secretoma obtido por RNA-seq.....	42
5.9	Proteínas secretadas ligadas à interação planta-hospedeiro e à sobrevivência ambiental	46
5.10	Identificação e perfil quantitativo do proteoma da glândula de Batelli por LC-MS/MS	47
5.11	Anotação funcional e interpretação biológica do proteoma da glândula de Batelli	48
5.12	Proteínas candidatas potencialmente relevantes para a interação planta-hospedeiro e sobrevivência ambiental, identificadas por análise proteômica.....	53
6.	DISCUSSÃO.....	55
7.	CONCLUSÃO	61
	Referências Bibliográficas	62
	Apêndice	69

1. INTRODUÇÃO

A cigarrinha-das-pastagens *Mahanarva spectabilis* (Distant, 1909) (Hemiptera: Cercopidae) é um inseto praga de gramíneas forrageiras. Infestações desse inseto reduzem o teor de clorofila da planta e aumentam a porcentagem de matéria seca, levando a expressivas perdas produtivas e econômicas (Resende *et al.*, 2013). O controle efetivo das ninfas é crucial, uma vez que reduz diretamente a emergência de adultos e a disseminação das infestações nos ecossistemas de pastagens (Alvarenga *et al.*, 2019).

Uma característica distintiva das ninfas de cigarrinhas é a produção de uma secreção espumosa que as envolve. Essa espuma, produzida principalmente pela glândula de Batelli em associação com secreções dos túbulos de Malpighi, desempenha múltiplas funções protetoras (Farina *et al.*, 2022; Guilbeau, 1908; Marshall, 1965). Ela atua como uma barreira física e bioquímica contra a dessecação, variações de temperatura, predadores e colonização microbiana, além de criar um microambiente tamponado que favorece a sobrevivência ninfal sob condições de campo (Aquad *et al.*, 2012; Alzani *et al.*, 2023; Hoch *et al.*, 2024; Sahayaraj *et al.*, 2021; Tonelli *et al.*, 2018; Tonelli *et al.*, 2020). Estudos em espécies relacionadas, como *Mahanarva fimbriolata* e *Poophilus costalis*, demonstraram que as espumas das cigarrinhas contêm proteínas, carboidratos e lipídios com propriedades isolantes e antimicrobianas (Sahayaraj *et al.*, 2021; Tonelli *et al.*, 2018). No entanto, os mecanismos moleculares subjacentes à produção da espuma e às suas funções ainda permanecem pouco caracterizados.

Avanços recentes em sequenciamento de RNA de alto rendimento (RNA-seq) e em proteômica têm possibilitado a caracterização funcional abrangente de transcriptomas e secretomas de insetos. Estudos transcriptômicos em hemípteros, como *Empoasca vitis* (Shao *et al.*, 2021), *Myzus persicae* (Sharma *et al.*, 2024.) e *Lipaphis erysimi* (Sharma *et al.*, 2024), revelaram expansões em famílias gênicas associadas à desintoxicação, osmorregulação e quimiorrecepção, destacando seu papel na adaptação às plantas hospedeiras. Abordagens proteômicas combinadas à transcriptômica também têm sido aplicadas às secreções salivares de pulgões (*Pseudoregma bambucicola*) (Zhang *et al.*, 2023), carrapatos (*Ornithodoros erraticus*) (Pérez-sánchez *et al.*, 2022) e até mesmo a estruturas especializadas de cigarrinhas (*Nephotettix*

cincticeps) (Wu *et al.*, 2023), permitindo a identificação de enzimas, efetores e proteínas estruturais cruciais para as interações hospedeiro-parasita. Em conjunto, esses estudos evidenciam como abordagens ômicas integrativas podem revelar as estratégias moleculares utilizadas pelos insetos para se adaptarem aos seus nichos ecológicos.

Apesar da relevância agrônômica de *M. spectabilis*, poucos estudos moleculares têm abordado o estágio ninfal ou a glândula de Batelli, responsável pela produção da espuma protetora. Neste trabalho, combinamos análises de RNA-seq de ninfas de *M. spectabilis* com o perfil proteômico da glândula de Batelli obtido por LC–MS/MS (cromatografia líquida acoplada à espectrometria de massas em tandem), fornecendo a primeira caracterização molecular integrada da produção de espuma em cigarrinhas-das-pastagens. Os genes e proteínas identificados foram anotados e categorizados em classes funcionais, incluindo desintoxicação, resposta ao estresse, quimiorrecepção, osmorregulação, imunidade e remodelamento da cutícula. Ao integrar assinaturas transcriptômicas com evidências proteômicas de proteínas secretadas, este estudo não apenas amplia a compreensão da biologia das cigarrinhas, como também identifica moléculas candidatas com potencial para o desenvolvimento de estratégias inovadoras de controle. Considerando que o controle das ninfas reduz as populações subsequentes de adultos, os *insights* obtidos aqui podem contribuir para o desenvolvimento de abordagens sustentáveis para o manejo de infestações de cigarrinhas em sistemas de pastagens.

2. REVISÃO DE LITERATURA

2.1 Impactos da cigarrinha-das-pastagens na agropecuária

Nas décadas de 1960 e 1970, a produção agropecuária brasileira era insuficiente para atender à demanda interna, tornando o país dependente da importação de alimentos. A partir de investimentos em pesquisa científica e no desenvolvimento de tecnologias voltadas para a agricultura tropical, o agronegócio passou por intensa expansão, consolidando-se como um dos principais setores da economia nacional (Torres, 2023). Em 2025, o setor respondeu por aproximadamente 25,13% do Produto Interno Bruto (PIB) brasileiro (Cepea; CNA, 2026). Nesse contexto, as pastagens ocupam a maior parte das áreas destinadas à agropecuária no país, abrangendo cerca de 165 milhões de hectares (MapBiomas, 2025).

Embora as pastagens ocupem grande parte da área agropecuária brasileira, a produtividade dos sistemas pecuários permanece abaixo de seu potencial, evidenciando oportunidades para intensificação sustentável da produção (Arantes *et al.*, 2018). Diversos fatores contribuem para essa baixa produtividade, destacando-se a reduzida fertilidade dos solos, o manejo inadequado das pastagens e os estresses abióticos. Somam-se a esses fatores os danos provocados por plantas daninhas, doenças e insetos-praga, que comprometem o desenvolvimento das forrageiras (Carrascosa *et al.*, 2024). Como as gramíneas forrageiras constituem a principal fonte alimentar dos rebanhos bovinos criados em sistemas de pastejo, reduções na produtividade e na qualidade das pastagens comprometem diretamente o desempenho animal e, conseqüentemente, a produção de carne e leite (Congio *et al.*, 2020).

Os insetos-praga constituem um dos principais entraves à produção agrícola mundial, comprometendo a produtividade das culturas e gerando expressivos prejuízos econômicos, sociais e ambientais (Tiwari, 2024). Entre os principais insetos associados às pastagens brasileiras destacam-se as cigarrinhas-das-pastagens, especialmente espécies dos gêneros *Aeneolamia*, *Deois*, *Mahanarva*, *Zulia*, *Notozulia*, e *Prosapia*. Esses insetos provocam alterações significativas na estrutura e no desenvolvimento das forrageiras (Aquad e Silva, 2019, p. 371). Em capim-marandu (*Urochloa brizantha*), Congio (2010) demonstrou que a infestação por *Mahanarva* resulta em plantas de menor altura, maior perfilhamento, aumento da proporção de folhas e material morto,

redução da proporção de hastes e alterações nas características morfológicas e estruturais da planta. Posteriormente, Congio *et al* (2020) verificou que infestações por espécies de *Mahanarva* reduzem em até 74% a produtividade de novilhas Nelore mantidas em pastagens de capim-marandu.

Diante da importância econômica das cigarrinhas-das-pastagens, seu manejo requer estratégias integradas que conciliem eficiência no controle, viabilidade econômica e sustentabilidade ambiental (Oerke, 2006). Esses insetos atacam uma ampla variedade de gramíneas forrageiras, incluindo espécies dos gêneros *Urochloa*, *Cenchrus*, *Megathyrsus* e *Cynodon*. Entretanto, a intensidade dos danos varia de acordo com a espécie de cigarrinha, o genótipo da planta hospedeira e as condições ambientais (Alvarenga *et al.*, 2017).

Os impactos desse grupo de insetos também se estendem a outras culturas de importância econômica. Em cana-de-açúcar, por exemplo, ataques de cigarrinhas podem reduzir em até 50% o rendimento da cultura, além de comprometer a eficiência da moagem, dificultar a recuperação de sacarose e inibir o processo fermentativo, causando prejuízos expressivos para a indústria sucroenergética (Dinardo-Miranda *et al.*, 2016; Madaleno *et al.*, 2008).

2.2 Aspectos biológicos de *Mahanarva spectabilis*

A cigarrinha-das-pastagens *Mahanarva spectabilis* (Distant, 1909) é um inseto da ordem Hemiptera, subordem Auchenorrhyncha e família Cercopidae, considerado uma das principais pragas de gramíneas forrageiras no Brasil. Os adultos medem aproximadamente 12 mm de comprimento e são maiores que a maioria das demais espécies de cigarrinhas associadas às pastagens. Assim como os demais cercopídeos, possuem aparelho bucal do tipo sugador, adaptado à alimentação por meio da sucção da seiva do xilema das plantas hospedeiras (Santos & Silva, 2023).

Mahanarva spectabilis apresenta metamorfose incompleta (hemimetabolia), passando pelas fases de ovo, ninfa e adulto. Os ovos são alongados e podem entrar em diapausa durante períodos desfavoráveis, especialmente no outono, contribuindo para a sobrevivência da espécie entre as estações. Após a eclosão, as ninfas migram para a base das gramíneas, onde se alimentam continuamente da seiva do xilema e secretam uma espuma característica que envolve completamente seus corpos. Essa espuma desempenha funções essenciais para a sobrevivência do inseto, protegendo-o

contra a dessecação, oscilações de temperatura, inimigos naturais e outros fatores ambientais. Em razão da dependência de condições de elevada umidade, as ninfas são mais abundantes durante o período chuvoso, especialmente no verão (Valério *et al.*, 2001; Auad *et al.*, 2020).

O desenvolvimento ninfal compreende cinco ínstaes, culminando na emergência dos adultos, que passam a habitar a parte aérea das forrageiras e deixam de produzir espuma (Souza *et al.*, 2007). Nessa fase, os insetos alimentam-se das folhas por meio da sucção da seiva do xilema e injetam saliva diretamente nos tecidos vegetais. Estudos metabolômicos baseados em cromatografia gasosa e líquida acopladas à espectrometria de massas (GC-MS e LC-MS) demonstraram que as glândulas salivares de *M. spectabilis* contêm elevadas concentrações de ácidos graxos de cadeia longa, como o ácido octadecanoico, além de diversos fitormônios, incluindo ácido salicílico, ácido indol-3-acético e metiljasmonato (Saraiva *et al.*, 2021; Monteiro *et al.*, 2021). Esses compostos podem atuar na modulação das respostas fisiológicas e de defesa da planta hospedeira, sendo proposto que sejam sintetizados ou sequestrados pelo inseto, ou ainda produzidos por microrganismos simbiotes associados, para posteriormente serem introduzidos nos tecidos vegetais durante a alimentação. Como consequência dessa interação, observam-se alterações fisiológicas que incluem lesões necróticas, redução do teor de clorofila e diminuição da capacidade fotossintética, resultando no aumento de matéria seca, na perda da qualidade nutricional da forragem e, consequentemente, no comprometimento do desempenho animal (Resende *et al.*, 2013).

Embora os danos provocados pelos adultos sejam mais facilmente observados, estudos demonstram que, em espécies do gênero *Mahanarva*, as ninfas também contribuem significativamente para os prejuízos econômicos, podendo ocasionar danos de magnitude semelhante aos dos adultos em decorrência do seu maior porte em comparação com ninfas de outros gêneros de cigarrinha-das-pastagens, elevada taxa de alimentação e permanência prolongada sobre as plantas hospedeiras (Alvarenga *et al.*, 2019).

2.3 Glândula de Batelli e produção da espuma

As glândulas de Batelli são glândulas epidérmicas localizadas nos sétimo e oitavo segmentos abdominais de cigarrinhas. Estas estão ativas em ninfas de cigarrinhas e cessam a produção de secreção quando o inseto atinge a idade adulta. Análises histológicas revelaram poros igualmente distribuídos pela superfície, por onde saem um líquido transparente que se mistura com outros fluidos do intestino (Guilbeau, 1908; Marshall, 1966). Bolhas de ar são adicionadas à secreção por movimentos de contração e expansão do abdômem, dando origem à espuma que é secretada pelo ânus. Esta espuma funciona como um micro-habitat, que confere proteção às ninfas. Diversas funções têm sido atribuídas à esta espuma, como isolamento térmico, proteção contra radiação solar e atividades antimicrobianas (Chen *et al*, 2018; Sahayaraj, 2021; Tonelli *et al*, 2018).

Muitas classes de moléculas já foram identificadas em amostras de espuma de cigarrinha-das-pastagens, como proteínas, lipídeos e carboidratos (Aquad *et al.*, 2012; Tonelli *et al.*, 2018). Rinaldi *et al* (2026a) identificou 196 proteínas que constituem essa espuma através de abordagem proteômica *shotgun* LC-MS/MS. Modelagens em 3D usando AlphaFold demonstraram que as proteínas estruturais da espuma compartilham moldes ricos em alfa-hélices estendidas e domínios de adesão (como repetições de anquirina e domínios WSC). Essas características permitem que as proteínas formem dímeros e construam matrizes supramoleculares altamente coesas e adesivas, garantindo a integridade mecânica e protetora do microambiente da ninfa (Rinaldi *et al*, 2026b).

Ao avaliar a espuma de ninfas alimentadas em diferentes gramíneas, constatou-se que o perfil proteico da secreção muda de acordo com a resistência do hospedeiro. Incluindo uma fração expressiva de proteínas não anotadas e altamente abundantes, sugerindo funções especializadas associadas à formação e manutenção da espuma (Rinaldi *et al*, 2026a).

2.4 O uso de tecnologias ômicas para a compreensão de sistemas biológicos na agricultura

A evolução das ciências agrárias nas últimas décadas foi marcada pela transição de uma abordagem focada em genes individuais para uma perspectiva sistêmica, impulsionada pelo advento das tecnologias de sequenciamento de nova geração (NGS) e métodos de alto rendimento. Esse avanço transformou a

pesquisa agrícola em uma disciplina intensiva em dados, permitindo o surgimento de estudos multiômicos que integram genômica, transcriptômica, proteômica e metabolômica (Sarfraz, 2025; Wang, 2022). Essas abordagens são fundamentais para elucidar a complexidade das interações biológicas, oferecendo uma visão holística dos mecanismos moleculares que regem a adaptação dos organismos a estresses bióticos e abióticos (Sarfraz, 2025).

Estudos transcriptômicos em hemípteros têm ampliado significativamente o conhecimento sobre os mecanismos moleculares envolvidos na adaptação às plantas hospedeiras. Em espécies como *Empoasca vitis* (Shao *et al.*, 2021), *Myzus persicae* e *Lipaphis erysimi* (Sharma *et al.*, 2024), análises de RNA-Seq revelaram a expansão de famílias gênicas relacionadas à desintoxicação, osmorregulação e quimiorrecepção, evidenciando sua importância na adaptação a diferentes hospedeiros e às condições ambientais. De forma complementar, a integração entre transcriptômica e proteômica tem sido empregada na caracterização de secreções salivares e estruturas especializadas de artrópodes, permitindo a identificação de enzimas, proteínas estruturais e efetores envolvidos na alimentação, na modulação das respostas de defesa do hospedeiro e na interação hospedeiro-parasita. Essa abordagem já foi aplicada em pulgões (*Pseudoregma bambucicola*) (Zhang *et al.*, 2023), carrapatos (*Ornithodoros erraticus*) (Pérez-Sánchez *et al.*, 2022) e cigarrinhas (*Nephotettix cincticeps*) (Wu *et al.*, 2023).

Apesar do potencial, a integração desses dados apresenta desafios, como a necessidade de lidar com a heterogeneidade dos conjuntos de dados e a dimensionalidade das variáveis. Ferramentas computacionais e estatísticas, como a Análise de Fatores Multiômicos (MOFA) e métodos baseados em redes (ex: MixOmics, DIABLO), têm sido desenvolvidas para identificar *clusters* de elementos significativos, permitindo a identificação de biomarcadores e a compreensão de fenótipos complexos (Wang, 2022).

Portanto, a aplicação de tecnologias ômicas integradas transcende a simples identificação de moléculas isoladas, proporcionando uma compreensão profunda da fisiologia do organismo em seu nicho ecológico. Essa abordagem sistêmica é indispensável para decifrar as estratégias moleculares de adaptação de pragas agrícolas, fornecendo subsídios para o desenvolvimento de estratégias de manejo mais precisas e sustentáveis frente às mudanças climáticas e aos desafios da agricultura moderna.

3. OBJETIVOS

3.1 Objetivo Geral

Caracterizar e integrar os perfis transcriptômico e proteômico de ninfas e glândula de Batelli de *Mahanarva spectabilis*, visando identificar genes e proteínas potencialmente envolvidos nos mecanismos moleculares associados à produção da espuma protetora.

3.2 Objetivos Específicos

3.2.1 Obter dados de RNAseq de alta qualidade das ninfas de quinto ínstar de *Mahanarva spectabilis*.

3.2.2 Caracterizar funcionalmente os unigenes obtidos por RNAseq por múltiplos bancos de dados.

3.2.3 Identificar e caracterizar assinaturas transcriptômicas potencialmente envolvidas na interação planta–inseto–ambiente.

3.2.4 Predizer e selecionar proteínas candidatas ao secretoma com base em análises *in silico*, realizando sua caracterização funcional e discutindo seu possível papel na interação planta–inseto–ambiente.

3.2.5 Integrar dados transcriptômicos e proteômicos, correlacionando os transcritos expressos com as proteínas secretadas identificadas na glândula de Batelli.

3.2.6 Caracterizar funcionalmente o proteoma da glândula de Batelli e discutir sua contribuição para a biologia da espuma e para os mecanismos moleculares envolvidos na interação planta–inseto.

4. METODOLOGIA

4.1 Insetos e extração de RNA

As ninfas de quinto instar foram identificadas com base no tamanho do corpo e, principalmente, no desenvolvimento avançado das tecas alares, que se estendiam além do segundo segmento abdominal, de acordo com os critérios morfológicos estabelecidos para Cercopidae (Marucci *et al.*, 1999).

Ninfas de *Mahanarva spectabilis* foram coletadas em campos de pastagem naturalmente infestados, localizados na estação experimental da Embrapa Gado de Leite (Juiz de Fora, MG, Brasil). As ninfas foram encontradas na base das plantas hospedeiras (*Cenchrus purpureus* (sin. *Pennisetum purpureum*) cultivar BRS Capiáçu), envoltas em suas massas de espuma características, e se alimentavam da seiva do xilema de gramíneas forrageiras no momento da coleta. Após a coleta, os insetos foram congelados instantaneamente em nitrogênio líquido e armazenados a -80°C até a extração de RNA e proteína. Não houve alimentação artificial ou criação em laboratório antes do preparo das amostras.

Para a extração do RNA, 50mg de tecido de inseto foram homogeneizadas em TRIzol (Invitrogen, Carlsbad, CA, EUA), e o RNA foi isolado de acordo com as instruções do fabricante, incluindo separação de fases, precipitação do RNA com isopropanol, lavagem com etanol a 75% e ressuspensão em água livre de nucleases. A concentração e a pureza do RNA foram avaliadas utilizando um espectrofotômetro NanoDrop 2000 (Thermo Scientific), e a integridade foi verificada por meio de um Bioanalyzer 2100 da Agilent (Agilent Technologies, Santa Clara, CA, EUA). Apenas amostras de RNA de alta qualidade, com número de integridade do RNA (RIN) $> 7,0$ e razão OD260/280 entre 1,9 e 2,1, foram utilizadas para a construção das bibliotecas.

4.2 Construção de bibliotecas e sequenciamento ilumina

Um único conjunto de dados de RNA-seq foi gerado a partir de uma amostra agrupada composta por oito ninfas inteiras de *M. spectabilis* para gerar um transcriptoma de referência abrangente.

O mRNA poli(A)+ foi enriquecido a partir do RNA total utilizando esferas magnéticas acopladas a oligo(dT), fragmentado em pequenos fragmentos (200–300 pb) e utilizado como molde para a síntese de cDNA (DNA complementar). A síntese da primeira fita de cDNA foi realizada utilizando primers hexâmeros aleatórios, e a RNase H foi empregada durante a síntese da segunda fita de cDNA para remover os moldes de RNA após a síntese da primeira fita, de acordo com os protocolos padrão de preparação de bibliotecas da Illumina. A síntese da segunda fita de cDNA foi concluída utilizando a DNA polimerase I. Após a purificação, os fragmentos de cDNA passaram por etapas de reparo de extremidades, adição de adenina na extremidade 3' (A-tailing) e ligação de adaptadores. Os produtos ligados foram selecionados por tamanho, amplificados por PCR e purificados para a geração das bibliotecas de cDNA. A qualidade das bibliotecas e a distribuição do tamanho dos insertos foram verificadas em um Qsep400 (BiOptic Inc.), quantificadas por qPCR (PCR quantitativo) e sequenciadas na plataforma Illumina HiSeq 4000 (*paired-end 150 bp reads*, PE150) (Illumina Inc., San Diego, CA, EUA).

4.3 Processamento dos dados e controle de qualidade

As leituras brutas (*raw reads*) foram processadas para a remoção de sequências de baixa qualidade, contaminação por adaptadores e leituras contendo mais de 10% de bases ambíguas, utilizando o software Trimmomatic v0.39 (Bolger *et al.*, 2014) e o FastQC v0.11.9 (Andrews, 2010). A filtragem de qualidade foi realizada sob a codificação Phred33, com base nos parâmetros Q20, Q30 e conteúdo GC.

4.4 Montagem *de novo* do transcriptoma

As leituras de alta qualidade foram montadas *de novo* utilizando o Trinity v2.8.5 (Grabherr *et al.*, 2011), que emprega uma estratégia baseada em grafos de Bruijn. De forma resumida, o Trinity é composto por três módulos sequenciais: (i) *Inchworm*, responsável por montar os dados de RNA-seq em sequências únicas de transcritos; (ii) *Chrysalis*, que agrupa os contigs e constrói grafos de Bruijn para cada agrupamento; e (iii) *Butterfly*, que processa os grafos para gerar transcritos completos e isoformas. A isoforma mais longa de cada agrupamento

foi designada como unigene. A qualidade da montagem foi avaliada por meio de métricas como N50, comprimento médio dos transcritos e número total de transcritos montados.

A qualidade da montagem *de novo* do transcriptoma de *Mahanarva spectabilis* foi avaliada usando o BUSCO v6.0.0 (Benchmarking Universal Single-Copy Orthologs). A análise foi realizada no modo transcriptoma, utilizando o conjunto de dados insecta_odb12, que compreende 3.114 ortólogos conservados representativos de Insecta. O BUSCO foi executado com os parâmetros padrão e com *multithreading* habilitado para avaliar a proporção de ortólogos completos (cópia única e duplicados), fragmentados e ausentes, fornecendo uma estimativa da qualidade geral e da completude do transcriptoma montado.

4.5 Análise e quantificação da expressão

As sequências brutas foram inicialmente submetidas a um filtro de qualidade para gerar sequências limpas. Esse processo incluiu a remoção de sequências contaminadas por adaptadores, sequências contendo mais de 10% de nucleotídeos ambíguos (N) e sequências em que mais de 50% das bases apresentavam um escore de qualidade Phred ≤ 10 . As sequências limpas de alta qualidade resultantes foram retidas para análises subsequentes.

As sequências limpas foram alinhadas de volta ao conjunto de unigenes montado usando o Bowtie2 v2.3.4.1 (Langmead & Salzberg, 2012) com parâmetros padrão. Os níveis de expressão gênica foram posteriormente estimados usando o RSEM v1.3.1 (Li & Dewey, 2011), que calcula a abundância de transcritos com base no número de fragmentos de leitura mapeados. Os valores de expressão foram normalizados como FPKM (*Fragments per kilobase of transcript per million mapped reads*), uma métrica que corrige tanto o comprimento do transcrito quanto a profundidade de sequenciamento.

A distribuição dos níveis de expressão dos unigenes foi avaliada usando gráficos de densidade e diagramas de caixa para avaliar os padrões gerais de expressão e a consistência dos dados em todo o transcriptoma.

4.6 Predição de peptídeo sinal e definição de candidatos ao secretoma

As sequências proteicas preditas utilizadas nesta análise foram derivadas de quadros de leitura abertos (*open reading frames*, ORFs) identificados no conjunto de unigenes obtido a partir da montagem do transcriptoma, especificamente de transcritos expressos na glândula de Batelli. Essas proteínas preditas foram geradas por tradução *in silico* das sequências codificadoras derivadas dos unigenes. As sequências proteicas resultantes foram utilizadas como entrada em formato FASTA para a predição de peptídeos sinal e vias de secreção, conforme descrito a seguir.

As sequências proteicas preditas do conjunto de dados da glândula de Batelli foram utilizadas como entrada em formato FASTA. Antes da análise, as sequências passaram por controle de qualidade, incluindo a remoção de caracteres de parada terminal ("*"), a conversão de aminoácidos não padronizados para símbolos canônicos quando possível e a exclusão de fragmentos com menos de 18 aminoácidos. Quando múltiplos genes/proteínas compartilhavam a mesma anotação funcional (nova montagem do genoma), apenas um representante foi mantido, priorizando-se (i) o maior comprimento proteico e (ii) o maior *bit score* do BLASTP contra o UniProt, de modo a favorecer a isoforma mais completa.

A presença de peptídeo sinal (SP) e o sítio putativo de clivagem Sec/SPI (peptídeo sinal dependente da via Sec, peptidase sinal I) foram preditos utilizando o SignalP 6.0 com o modelo para eucariotos e configurações padrão (Teufel *et al.*, 2022).

Em paralelo, a topologia e sequências sinal adicionais na região N-terminal foram inferidas com o Phobius v1.01 (saída longa), que modela conjuntamente peptídeos sinal e hélices transmembrana. Para cada proteína, foram registrados o resultado do SignalP ("SP" vs. "OTHER") e a posição de clivagem mais provável, juntamente com as chamadas do Phobius para SP (S/N) e o número de hélices transmembrana (TM) preditas.

Em seguida, foi aplicado um consenso simples e transparente para classificar a probabilidade de exportação. As proteínas foram rotuladas como "alta confiança" quando eram positivas no SignalP (Sec/SPI) e, adicionalmente, (a) o Phobius também previa um SP ou (b) o Phobius previa ≤ 1 hélice TM no total (permitindo uma única TM atribuível ao peptídeo sinal). As proteínas foram rotuladas como "confiança média" quando pelo menos uma das ferramentas (SignalP ou Phobius) indicava a presença de SP, mas os critérios de Alta

confiança não eram atendidos. Todas as demais proteínas foram rotuladas como “baixa confiança”. Os sítios de clivagem preditos pelo SignalP foram mantidos como os cortes operacionais do peptídeo sinal para as análises subsequentes.

Para construir o conjunto de trabalho do secretoma, foram retidas todas as proteínas classificadas como positivas no SignalP e/ou com consenso de Confiança média ou Alta, e removidas as proteínas claramente *multipass* de membrana (≥ 2 hélices TM além da região do peptídeo sinal). O conjunto final não redundante foi ranqueado de acordo com: (1) nível de consenso (alto > médio), (2) presença de um sítio de clivagem bem definido, (3) ausência de segmentos TM adicionais, (4) comprimento proteico e (5) *bit score* do BLASTP.

4.7 Análise de *splicing* alternativo

Para avaliar a complexidade do transcriptoma decorrente do *splicing* alternativo (AS), utilizamos a estrutura de isoformas gerada pela montagem *de novo* do transcriptoma realizada com o Trinity v2.15.1 (Grabherr *et al.*, 2011). Na nomenclatura do Trinity, transcritos que compartilham o mesmo identificador de “gene” (por exemplo, TRINITY_DNxxxxx_cY_gZ), mas apresentam sufixos de isoforma distintos (i1, i2, etc.), são considerados isoformas de um único *cluster* de unigene. O arquivo FASTA de transcritos (Demo.Transcripts.fa) foi processado utilizando scripts personalizados em Python 3.10 para contabilizar o número de isoformas por gene do Trinity. Para cada gene, foram calculados o número de isoformas, geradas distribuições de frequência e estimadas estatísticas descritivas, incluindo o número total de genes, o número máximo e médio de isoformas por gene e a proporção de *loci* com múltiplas isoformas. Genes com ≥ 2 isoformas foram definidos como *loci* multi-isoforma e considerados potenciais candidatos a *splicing* alternativo.

Para integrar as informações de AS com a anotação funcional, o conjunto de isoformas foi combinado com os resultados de Pfam, KEGG e de expressão gênica. Os domínios Pfam foram identificados utilizando o HMMER v3.3.2 contra o banco de dados Pfam-A (Mistry *et al.*, 2021). Os ortólogos e as vias do KEGG foram atribuídos por meio do KAAS (Moriya *et al.*, 2007), enquanto os valores de expressão dos unigenes foram estimados como FPKM a partir do alinhamento das leituras de RNA-seq utilizando o Bowtie2 v2.4.5 (Langmead & Salzberg, 2012) e o RSEM v1.3.3 (Li & Dewey, 2011).

Os genes foram classificados em categorias funcionais relevantes para interações hospedeiro–ambiente (Desintoxicação, Quimiorrecepção, Osmorregulação/Transporte, Cutícula/Barreira, Resposta ao estresse, Defesa imune e Digestão/interação com plantas) por meio da aplicação de filtros baseados em expressões regulares curadas sobre os campos de anotação Pfam e KEGG. Para cada categoria, foi calculado o número e a porcentagem de genes que apresentaram *splicing* alternativo em relação ao total de genes anotados.

4.8 Análise proteômica da glândula de Batelli por LC/MS

A análise proteômica foi realizada especificamente na glândula de Batelli, um órgão secretor especializado que desempenha papel central na produção da espuma e é essencial para a proteção e sobrevivência das ninfas de *M. spectabilis*.

4.8.1 Extração de proteínas e digestão

As glândulas de Batelli foram cuidadosamente dissecadas de ninfas do quinto instar de *Mahanarva spectabilis*. Brevemente, o tórax foi separado do abdômen, e a cutícula abdominal foi aberta longitudinalmente utilizando um bisturi estéril sob um estereomicroscópio (Olympus HZS10) com objetiva de 0,75×. As glândulas de Batelli foram então removidas manualmente da cavidade abdominal com o auxílio de pinças finas e imediatamente transferidas para água acidificada (pH 3,0), com o objetivo de preservar a integridade glandular e minimizar a degradação proteica. Aproximadamente 30 segmentos glandulares, obtidos a partir de 10 ninfas de quinto instar, foram reunidos e submetidos à lise celular por sonicação em tampão de lise SP3 (100 mM bicarbonato de trietilamônio [TEAB], 1% dodecil sulfato de sódio [SDS], 10 mM tris(2-carboxietil)fosfina [TCEP] e 40 mM cloroacetamida [CAA]), seguindo o protocolo Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) (Hughes *et al*, 2019). Após a lise, os detritos insolúveis foram removidos por centrifugação, e o sobrenadante clarificado foi coletado para quantificação proteica utilizando o ensaio do ácido bicinconínico (BCA). As proteínas foram posteriormente digeridas com tripsina utilizando esferas magnéticas, seguindo o fluxo de

trabalho do SP3, para a geração de peptídeos destinados à análise por espectrometria de massas.

4.8.2 Identificação de proteínas por LC-MS/MS

Os peptídeos foram ressuspensos em 50 μL de ácido fórmico 0,1% e analisados em um sistema nano-UHPLC (UltiMate R 3000, Dionex, San Jose, EUA) equipado com uma coluna Nano-Trap Acclaim PepMap100 C18 (100 μm d.i. $\times$ 20 mm, 5 μm , 100 Å; Thermo Scientific, Waltham, MA, EUA) e uma coluna analítica Acclaim PepMap100 C18 RSLC (75 μm d.i. $\times$ 150 mm, 2 μm , 100 Å; Thermo Scientific), operando em conjunto a um fluxo constante de 0,3 $\mu\text{L min}^{-1}$. Os solventes utilizados foram: (A) ácido fórmico 0,1% e (B) 80% acetonitrila contendo 0,1% de ácido fórmico. O gradiente multietapas foi aplicado da seguinte forma: etapa de condicionamento com 3,8% de B por 3 min, seguida de um gradiente de 3,8 a 30% de B ao longo de 120 min; posteriormente, aumento de 30 a 55% de B até 150 min; rampa final para 99% de B até 162 min, seguida de uma etapa de recondicionamento com 3,8% de B até 180 min.

Os dados espectrais foram adquiridos em um espectrômetro de massas Q-Exactive (Thermo Scientific, Bremen, Alemanha), operando no modo *full scan*/MS2. A fonte de ionização Nanospray Flex foi ajustada para 3,8 kV em modo positivo, com temperatura do capilar de 250 °C e nível do S-lens ajustado para 55. O método de Aquisição Dependente de Dados (DDA) foi realizado selecionando os 12 íons mais intensos, com estados de carga entre +2 e +4, em uma janela de 1,2 m/z. As varreduras MS1 foram configuradas com resolução de 70.000 e faixa de massa de 300 a 2.000 m/z. Para MS/MS, os íons selecionados foram fragmentados por Dissociação por Colisão de Alta Energia (HCD), com energia de colisão normalizada escalonada (NCE) de 28–30. Os espectros de fragmentação foram adquiridos com resolução de 17.500, tempo máximo de injeção de 60 ms, exclusão dinâmica de 40 s e valor alvo de AGC de 5×10^5 (AGC mínimo de $6,25 \times 10^3$).

Os dados brutos foram processados usando o PEAKS Studio v12.5, que realizou a identificação de proteínas por meio de uma abordagem combinada de sequenciamento *de novo* e busca em banco de dados. As buscas no banco de dados foram realizadas no banco de dados de unigenes obtido a partir da análise do transcriptoma de ninfas de *M. spectabilis* (Seções 4.4 e 4.5).

Os parâmetros usados no programa foram: unigenes de *M. spectabilis* como banco de dados de referência; oxidação de metionina como modificação variável; e carbamidometilação de cisteína como modificação fixa; uma clivagem perdida, estados de carga de 2+, 3+ e 4+; enzima de clivagem tipo tripsina e um erro de massa de 20 ppm. Para uma identificação confiável, considerou-se uma taxa de falsos positivos (FDR) < 0,5% para peptídeos e proteínas identificados usando FDR < 1%, mostrando pelo menos 2 peptídeos únicos.

5. RESULTADOS

5.1 Qualidade do sequenciamento e estatística da montagem dos dados do RNA-seq

O sequenciamento Illumina do transcriptoma de ninfas de *M. spectabilis* gerou um total de 63,0 milhões de *reads* limpos, correspondendo a 18,86 Gb de bases de alta qualidade. As *reads* brutas foram processadas para a remoção de adaptadores e sequências de baixa qualidade, gerando arquivos FASTQ limpos, os quais foram depositados no NCBI Sequence Read Archive (SRA) sob o BioProject PRJNA1404840. O conteúdo global de GC foi de 39,06%, e 94,09% das bases apresentaram valor de Q30 ou superior, indicando que o conjunto de dados é de excelente qualidade, com taxas muito baixas de erro na chamada de bases. De fato, a avaliação da qualidade das bases mostrou que Q20, Q30 e Q40 corresponderam a acurácias de 99%, 99,9% e 99,99%, respectivamente, confirmando a confiabilidade da corrida de sequenciamento.

A taxa média de erro ao longo das *reads* foi extremamente baixa ($<0,0003$), com um leve aumento próximo ao final tanto da *Read1* quanto da *Read2*, o que é consistente com os perfis típicos de sequenciamento Illumina (Figura S1). É importante destacar que não foi observada deterioração abrupta da qualidade, indicando uma chamada de bases confiável ao longo de todo o comprimento das *reads*. A análise da composição de bases mostrou uma distribuição quase uniforme de nucleotídeos após os ciclos iniciais, nos quais pequenas flutuações são esperadas devido a efeitos de iniciadores randômicos (Figura S2). Após aproximadamente a posição 20, as proporções se estabilizaram, com adenina (A) e timina (T) representando cerca de 30% cada, e guanina (G) e citosina (C) correspondendo a aproximadamente 20% cada. Essa representação equilibrada reflete a ausência de viés sistemático de nucleotídeos e sugere que a biblioteca de sequenciamento apresentou alta complexidade. A proporção de bases indeterminadas (N) foi negligenciável ao longo das *reads*. Esses parâmetros indicam que o sequenciamento da amostra de ninfas de *M. spectabilis* produziu *reads* de alta qualidade, adequadas para a montagem do transcriptoma e análises de expressão gênica subsequentes.

Para avaliar a qualidade e a representatividade da biblioteca de sequenciamento, foram realizadas diversas análises diagnósticas (Figura S4). A

avaliação da aleatoriedade da distribuição das *reads* ao longo dos transcritos demonstrou cobertura uniforme do extremo 5' ao 3' dos genes, com apenas um leve enriquecimento nas regiões centrais (Figura S4A). Esse padrão indica que a construção da biblioteca foi isenta de vieses e que as *reads* foram distribuídas de maneira homogênea ao longo dos transcritos. A análise da distribuição do tamanho dos insertos revelou que a maioria dos fragmentos *paired-end* variou entre 300 e 500 pb, com um pico bem definido em aproximadamente 400 pb (Figura S4B). Esse tamanho de inserto é consistente com o esperado para a preparação da biblioteca.

A análise de saturação gênica, utilizando fragmentos por quilobase de transcrito por milhão de *reads* mapeadas (FPKM $\geq 0,1$), mostrou que o número de genes detectados atingiu um platô após cerca de 10 milhões de *reads*, alcançando aproximadamente 100.000 transcritos (Figura S4C). Esse resultado indica que a profundidade de sequenciamento foi suficiente para capturar a maior parte dos genes expressos no estágio ninfal de *M. spectabilis*.

5.2 Estatísticas de montagem e mapeamento *de novo* do transcriptoma de ninfas de *Mahanarva spectabilis*

Após a filtragem de qualidade, 63,5% das *reads* (40.015.826) foram mapeadas com sucesso contra os unigenes montados, das quais aproximadamente metade (49,1%) apresentou mapeamento único, enquanto os 50,9% restantes corresponderam a mapeamentos múltiplos.

A montagem *de novo* do transcriptoma produziu um conjunto abrangente de unigenes, com transcritos variando desde fragmentos curtos até sequências de comprimento completo. No total, foram montados 100.666 unigenes. A distribuição de comprimento mostrou que a maioria dos unigenes se concentrou na faixa de 300–1.000 nt (64,4%), enquanto 13,9% apresentaram comprimento entre 1.000 e 2.000 nt, e uma fração significativa (8,6%) excedeu 2.000 nt (Figura S3). Esses unigenes mais longos provavelmente representam transcritos completos ou quase completos, sendo particularmente valiosos para a anotação funcional e análises subsequentes. Uma fração substancial de unigenes ainda mais longos (>3.000 nucleotídeos) também foi recuperada, refletindo a presença de transcritos completos ou quase completos.

A montagem *de novo* resultou em 99.669 unigenes, com N50 de unigenes de 1.259 pb (N50 de transcritos = 1.598 pb) e comprimento médio de unigene de 878 pb, indicando boa continuidade e representatividade da montagem.

A qualidade da montagem do transcriptoma de *M. spectabilis* foi avaliada utilizando o BUSCO com o conjunto de dados de linhagem insecta_odb12. A análise recuperou 89,2% de BUSCOs completos, dos quais 88,4% estavam em cópia única e 0,8% estavam duplicados. Adicionalmente, 2,9% dos BUSCOs foram classificados como fragmentados, enquanto 7,9% estavam ausentes. Esses resultados indicam um alto nível de completude e baixa redundância, sustentando a robustez da montagem *de novo*.

5.3 Anotação funcional dos unigenes

A anotação funcional dos unigenes foi realizada com base na homologia de sequência e na identificação de domínios proteicos. As sequências dos unigenes foram comparadas com os bancos de dados NR, Swiss-Prot, COG, KOG, eggNOG (v4.5) e KEGG utilizando o DIAMOND. As atribuições de ortologia KEGG foram obtidas utilizando o KOBAS. Os termos da Gene Ontology (GO) foram atribuídos utilizando o InterProScan com base em domínios proteicos conservados e assinaturas funcionais integradas. Nenhuma análise de enriquecimento da Gene Ontology foi realizada neste estudo; os termos GO foram utilizados exclusivamente para a classificação funcional e análises descritivas do transcriptoma anotado.

Para obter informações sobre as funções biológicas dos unigenes montados, diversas bases de dados funcionais foram consultadas, incluindo COG, GO, KEGG, KOG, Pfam, eggNOG, NR e SwissProt. No total, uma fração substancial dos unigenes pôde ser anotada com sucesso, destacando a robustez do conjunto de dados para análises posteriores. Entre os recursos de anotação, o GO (Gene Ontology) forneceu informações funcionais para 13.282 unigenes, enquanto o KEGG (Kyoto Encyclopedia of Genes and Genomes) atribuiu 10.768 unigenes a vias metabólicas e de sinalização. As buscas por domínios no Pfam resultaram em 12.945 correspondências, e a classificação KOG (Eukaryotic Orthologous Groups) identificou 8.404 unigenes. Além disso, a análise COG (Clusters of Orthologous Groups) anotou 4.958 unigenes, refletindo funções ortólogas conservadas entre diferentes espécies.

A representação integrativa em “gráfico em flor” (Figura 1) ilustra a sobreposição entre as diferentes bases de dados de anotação. Um conjunto central de 2.638 unigenes foi consistentemente anotado em múltiplos bancos de dados, ressaltando sua forte conservação funcional. Além do conjunto central, anotações únicas adicionais foram obtidas de recursos individuais, como 2.901 unigenes anotados exclusivamente no eggNOG, 2.065 no NR e 467 no SwissProt, entre outros. Esses resultados indicam que uma grande proporção do transcriptoma da ninfa de *M. spectabilis* pode ser caracterizada funcionalmente, fornecendo uma base valiosa para a exploração da função gênica e a análise de vias metabólicas nesse estágio de desenvolvimento. Dada a ausência de um genoma de referência para *M. spectabilis*, a identificação de 2.638 unigenes consistentemente anotados em múltiplos bancos de dados representa uma fração substancial e biologicamente significativa do transcriptoma, refletindo um núcleo robusto de genes conservados e funcionalmente suportados.

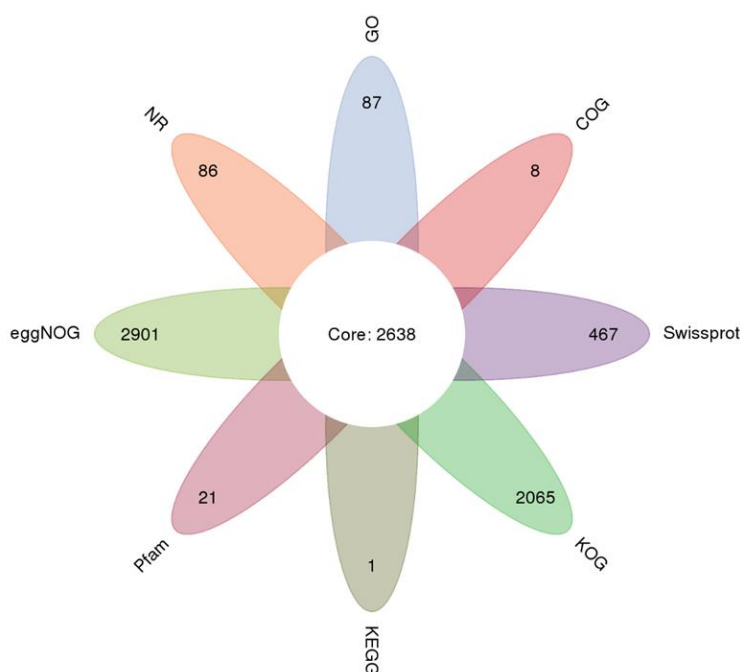
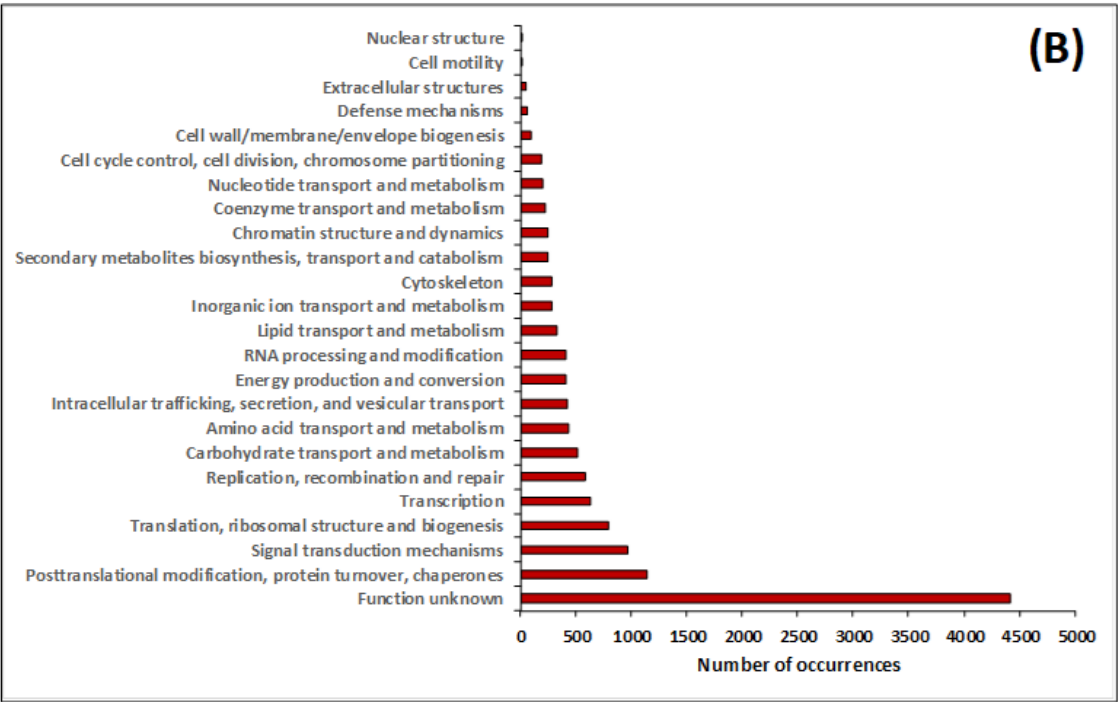
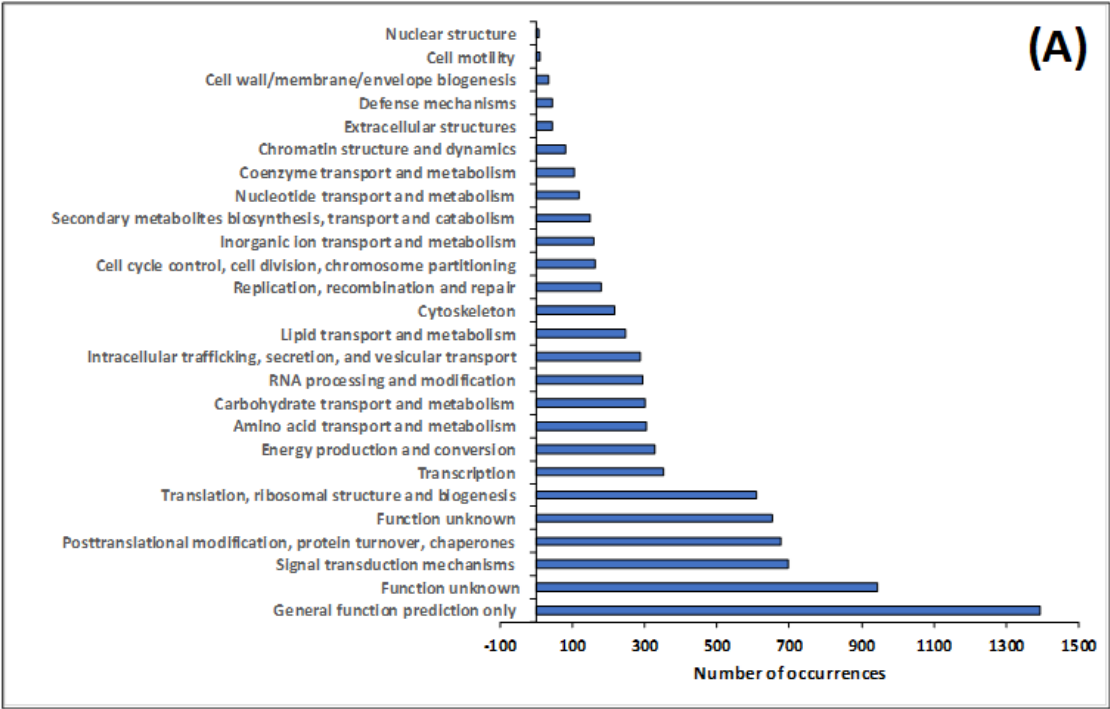


Figura 1- Anotação funcional de unigenes de ninfas de *Mahanarva spectabilis* em múltiplos bancos de dados. O gráfico de flores ilustra anotações compartilhadas e únicas, com um conjunto central de 2.638 unigenes caracterizados funcionalmente.

5.4 Categorização funcional dos unigenes

Para obter uma visão funcional abrangente do transcriptoma da ninfa de *M. spectabilis*, os unigenes montados foram anotados em relação a múltiplos bancos de dados de referência, incluindo eggNOG/COG/KOG, Pfam, KEGG, GO, SwissProt e NR (Figura 2 e Tabela S1). Cada banco de dados forneceu camadas complementares de informação, permitindo uma caracterização funcional ampla e detalhada.

Por exemplo, o Pfam capturou assinaturas em nível de domínio não detectadas em KOG/eggNOG, enquanto eggNOG/KOG forneceu previsões funcionais filogeneticamente conservadas que contextualizam adaptações específicas de insetos. Essas anotações confirmam que o transcriptoma da ninfa de *M. spectabilis* codifica um amplo repertório de genes envolvidos em processos biológicos essenciais, vias de sinalização e funções metabólicas, além de revelar um grande conjunto de transcritos com funções desconhecidas que podem representar novos genes ou inovações específicas da linhagem.



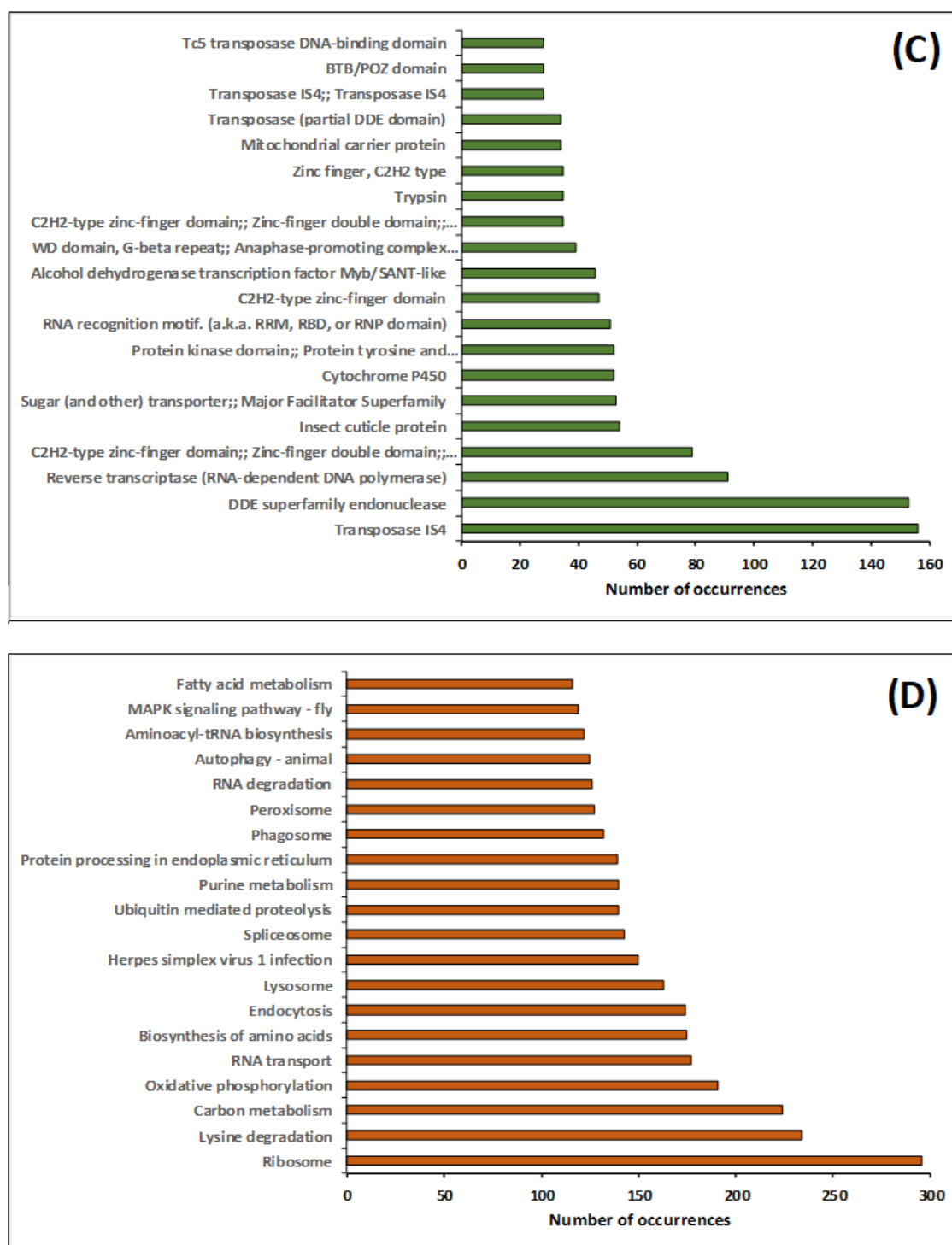


Figura 2 - Anotação funcional dos unigenes do transcriptoma de ninfas de *Mahanarva spectabilis*. (A) Classificação funcional KOG dos unigenes. As maiores categorias foram predição de função geral apenas, seguidas por mecanismos de transdução de sinais e modificação pós-traducional, turnover proteico e chaperonas. (B) Classificação funcional eggNOG dos unigenes. A categoria predominante foi função desconhecida, enquanto um número substancial de transcritos esteve envolvido em modificação pós-traducional, transdução de sinais, tradução e estrutura ribossomal, e transcrição. (C) Os 20 principais domínios Pfam identificados entre os unigenes montados. Os domínios proteicos mais frequentes incluíram transposase IS4, endonuclease da superfamília DDE, transcriptase reversa (DNA polimerase

dependente de RNA), domínios zinc-finger do tipo C2H2 e proteínas da cutícula de insetos, refletindo a diversidade de motivos estruturais e enzimáticos presentes no transcriptoma. (D) Os 20 principais caminhos metabólicos KEGG representados no transcriptoma de ninfas de *M. spectabilis*. A tabela lista os caminhos mais enriquecidos, incluindo o identificador KEGG e o número de unigenes atribuídos a cada via. As categorias mais representadas incluíram biogênese de ribossomos, metabolismo de aminoácidos, fosforilação oxidativa, transporte de RNA e metabolismo central do carbono, destacando a forte representação de processos celulares e metabólicos fundamentais no estágio ninfal.

A base de dados Pfam identificou milhares de domínios proteicos conservados, incluindo famílias de peptidases, transportadores, fatores de transcrição e proteínas ligantes de RNA, destacando a diversidade de motivos estruturais e catalíticos codificados no transcriptoma. Os domínios mais abundantes incluíram transposase IS4, endonucleases da superfamília DDE, transcriptase reversa (polimerase de DNA dependente de RNA), domínios do tipo dedo de zinco C2H2 e proteínas da cutícula de insetos, refletindo tanto genes de manutenção celular quanto repertórios gênicos específicos de insetos (Figura 2C).

A classificação por meio das bases de dados eggNOG e KOG atribuiu os unigenes a grandes categorias funcionais relacionadas ao metabolismo, aos processos celulares e ao armazenamento e processamento de informação. No KOG, os grupos mais abundantes corresponderam a “predição de função geral”, mecanismos de transdução de sinais e modificação pós-traducional, *turnover* proteico e chaperonas (Figuras 2A e 2B). No eggNOG, a categoria predominante foi “função desconhecida”, seguida por transcritos envolvidos em modificação pós-traducional, transdução de sinais, tradução e estrutura ribossomal, e transcrição (Figura 2C). A elevada proporção de unigenes classificados como de função desconhecida ou geral é uma característica comum de transcriptomas de insetos não modelo e reflete a presença de genes conservados cujas funções biológicas específicas ainda são pouco caracterizadas, juntamente com uma ampla representação de processos celulares e regulatórios essenciais.

A integração entre os diferentes recursos de anotação revelou um conjunto central de transcritos anotados por múltiplos bancos de dados, aumentando a confiança na atribuição funcional. Ao mesmo tempo, cada recurso individual contribuiu com informações únicas. Por exemplo, o Pfam capturou assinaturas em nível de domínio que não foram detectadas pelo KOG/eggNOG,

enquanto o eggNOG/KOG forneceu predições funcionais filogeneticamente conservadas que auxiliam na contextualização de adaptações específicas de insetos.

A anotação KEGG revelou que 8.836 unigenes foram mapeados em vias metabólicas e de sinalização, fornecendo uma visão geral ampla do panorama funcional do transcriptoma de ninfas de *M. spectabilis* (Figura 2D). As categorias mais representadas estiveram associadas ao metabolismo primário, particularmente ao metabolismo energético e de carboidratos. Entre essas, glicólise/gliconeogênese (103 unigenes), ciclo do citrato (ciclo do TCA; 77 unigenes) e via das pentoses fosfato (58 unigenes) foram as mais enriquecidas, ressaltando a forte representação do metabolismo central do carbono.

Além do metabolismo de carboidratos, diversas vias relacionadas à biossíntese e degradação de aminoácidos, metabolismo de lipídios e biossíntese de cofatores e vitaminas também estiveram altamente representadas. Notavelmente, unigenes também foram atribuídos a vias envolvidas na transdução de sinais e no processamento de informações ambientais, incluindo genes relacionados às cascatas de sinalização MAPK e PI3K-Akt, frequentemente associadas a respostas ao estresse celular e à adaptação de insetos às plantas hospedeiras.

5.5 Assinaturas transcriptômicas envolvidas na interação planta-inseto

Nossa análise do transcriptoma revelou um amplo repertório de genes classificados por anotação funcional como potencialmente envolvidos nas interações planta–inseto e ambiente–ninfa (Tabelas 1, S2 e S3). Essas inferências baseiam-se em termos conservados de Gene Ontology, domínios Pfam e vias do KEGG, e não implicam validação experimental de suas funções em *M. spectabilis*. Em conjunto, essas categorias correspondem a aproximadamente 6–7% do conjunto total de unigenes e representam um subconjunto biologicamente direcionado, selecionado por sua relevância para a sobrevivência das ninfas, proteção mediada pela espuma e interações com a planta e o ambiente.

Tabela 1 - Resumo das categorias de transcritos identificadas no transcriptoma ninfal de *Mahanarva spectabilis* relacionadas à interação com a planta hospedeira e à adaptação ao ambiente. Listas detalhadas de genes são apresentadas nas Tabela Suplementar S2.

Category	PFAM hits	KEGG hits	Overlap
Detoxification	317	217	96
Chemosensory	11	18	0
Osmoregulation_transport	42	0	0
Cuticle_barrier	65	0	0
Stress_response	218	100	20
Immune_defense	43	0	0
Plant_interaction_digestion	65	33	4

Genes envolvidos na detoxificação de metabólitos secundários de plantas e xenobióticos estiveram fortemente representados. Um total de 317 transcritos apresentou domínios Pfam típicos, incluindo monooxigenases do citocromo P450 (CYP450), glutathione S-transferases (GSTs), UDP-glicosiltransferases (UGTs), carboxilesterases e redutases de cadeia curta/aldo-ceto (SDR/AKR). Em paralelo, 217 anotações do KEGG estiveram associadas a funções relacionadas à detoxificação, com 96 transcritos apresentando sobreposição tanto com Pfam quanto com KEGG, indicando forte suporte para a capacidade de detoxificação. As vias do KEGG enriquecidas incluíram Metabolismo da glutathione (ko00480; 80 genes), Metabolismo de xenobióticos pelo citocromo P450 (ko00980; 39 genes), Metabolismo de fármacos via P450 (ko00982; 36 genes) e Metabolismo de fármacos por outras enzimas (ko00983; 100 genes). Entre os candidatos com maior nível de expressão destacaram-se transportadores ABC, GSTs, carboxilesterases e monooxigenases CYP450.

Assinaturas transcriptômicas relacionadas à quimiossensação, essencial para o reconhecimento da planta hospedeira e seleção do habitat, também foram detectadas. As análises por Pfam e KEGG recuperaram 11 e 1 transcritos relacionados à quimiossensação, respectivamente, incluindo proteínas ligantes de odor (OBPs), proteínas ligantes de feromônios (PBP), proteínas quimiossensoriais (CSPs) e receptores das famílias OR/GR/IR. O KEGG também identificou dois genes associados à via de transdução olfatória (ko04740). Notavelmente, os candidatos quimiossensoriais mais altamente expressos pertenciam à família PBP/GOBP.

Genes associados à osmorregulação e ao transporte iônico, críticos para insetos alimentadores de xilema, foram altamente abundantes. Identificamos 42 anotações Pfam e 100 anotações KEGG correspondentes a v-ATPases, Na⁺/K⁺-ATPases, transportadores da família solute carrier (SLC) e anidrases carbônicas, com 20 transcritos sustentados por ambos os sistemas de anotação. Diversas subunidades de v-ATPase, ATPases E1–E2 e anidrases carbônicas figuraram entre os transcritos mais expressos, refletindo as demandas fisiológicas impostas por uma dieta diluída e rica em água.

A adaptação ambiental também foi evidenciada por transcritos envolvidos na formação e remodelação da cutícula. Um total de 158 transcritos codificou proteínas cuticulares (famílias CPR RR-1, RR-2, CPF e CPLCG) e enzimas do metabolismo da quitina, incluindo quitinases, quitina desacetilases e lacases (Tabelas 1, S2 e S3). Essas famílias gênicas são essenciais para a formação de barreiras protetoras e provavelmente contribuem para a resistência à dessecação e à invasão microbiana no ambiente da espuma.

Por fim, transcritos relacionados a respostas ao estresse e imunidade inata foram abundantes. A anotação Pfam identificou 218 transcritos codificando proteínas de choque térmico (HSPs) e enzimas antioxidantes, como superóxido dismutases, catalases e peroxirredoxinas. Componentes de vias de imunidade inata também estiveram representados, com 43 anotações Pfam e 33 KEGG englobando serino proteases/serpinas, lisozimas e receptores de reconhecimento de patógenos, incluindo elementos das vias PGRP e Toll. Em conjunto, esses conjuntos gênicos destacam mecanismos moleculares que sustentam a sobrevivência, a tolerância ao estresse e a defesa contra microrganismos associados às plantas.

5.6 Predição de *splicing* alternativo no transcriptoma da ninfa

Os padrões de *splicing* alternativo relatados aqui são baseados em previsões computacionais de todo o transcriptoma, derivadas da montagem *de novo*, e representam estimativas da diversidade de isoformas, em vez de eventos de *splicing* validados experimentalmente.

A montagem *de novo* do transcriptoma de ninfas de *M. spectabilis* revelou um grau substancial de diversidade de isoformas. Com base no agrupamento Trinity, um total de 99.669 agrupamentos de unigenes foram identificados, dos

quais 16.374 (16,43%) produziram pelo menos duas isoformas, indicando splicing alternativo (SA). O número médio de isoformas por unigene foi de 1,40, com um máximo de 75 isoformas observadas em um único locus. A distribuição da contagem de isoformas apresentou forte assimetria à direita (Tabelas S4–S6), com a maioria dos genes apresentando uma única isoforma e um número progressivamente menor de loci carregando múltiplas isoformas. O status da isoforma indicou que ~84% dos agrupamentos de unigenes eram de isoforma única, enquanto ~16% eram de múltiplas isoformas.

Para avaliar se o *splicing* alternativo afetava preferencialmente famílias de genes funcionais específicas, os loci com múltiplas isoformas foram comparados com as anotações do Pfam e do KEGG e agrupados em categorias relevantes para as interações hospedeiro-planta e hospedeiro-ambiente (Figura 3, Tabela S6). A análise revelou que diversas categorias funcionalmente importantes abrigavam uma proporção notável de genes com splicing alternativo. Transcritos relacionados à desintoxicação (citocromos P450, GSTs, UGTs, carboxilesterases e transportadores ABC) exibiram extensa diversidade de isoformas, consistente com seu papel na plasticidade metabólica e no processamento de xenobióticos. Da mesma forma, genes quimiossensoriais (receptores de odor/gustativos/ionotrópicos, OBPs e CSPs) e transportadores osmorregulatórios (v-ATPases, aquaporinas e trocadores iônicos) também apresentaram múltiplas isoformas, sugerindo modulação em nível de isoforma na detecção do hospedeiro, no equilíbrio hídrico e na homeostase. Proteínas cuticulares com múltiplas isoformas, fatores de resposta ao estresse (proteínas de choque térmico, antioxidantes) e genes relacionados à imunidade (lisozimas, PGRPs, serina proteases, defensinas) também foram abundantes, refletindo a contribuição do splicing alternativo para a adaptação fisiológica e os mecanismos de defesa. Finalmente, enzimas de interação com plantas (celulases, pectinases, glicosidases, amilases) também incluíram vários loci com múltiplas isoformas, apontando para uma potencial especialização em nível de isoforma na digestão da parede celular vegetal.

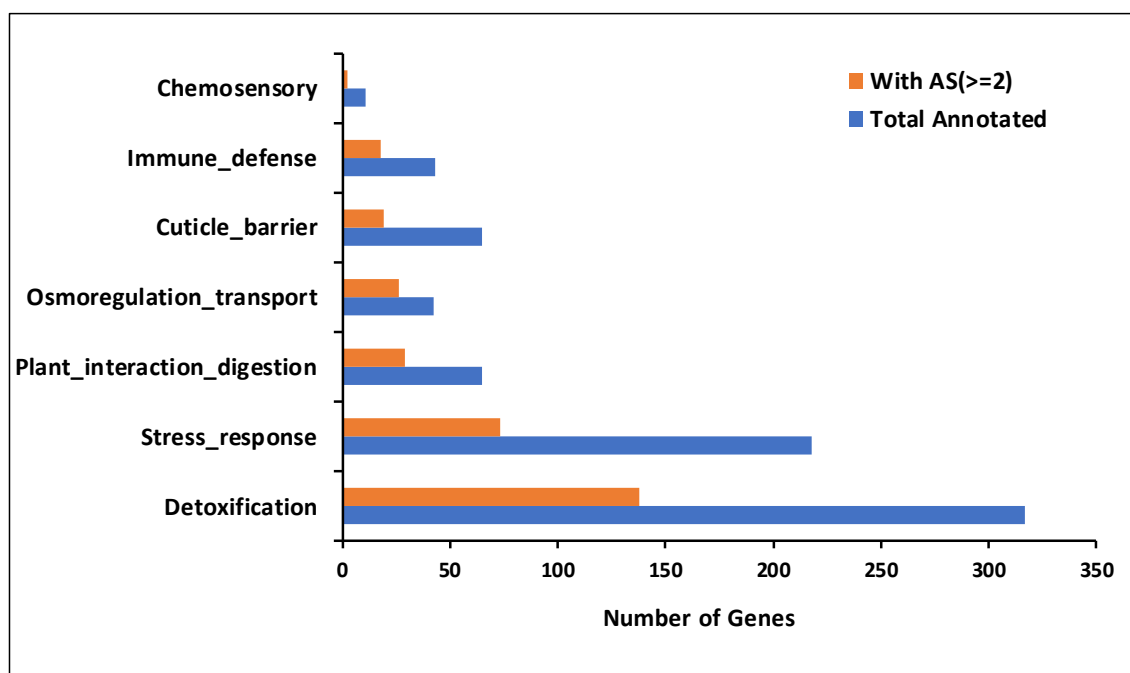


Figura 1 - Genes com e sem splicing alternativo nas diferentes categorias funcionais do transcriptoma de ninfas de *Mahanarva spectabilis*. Gráfico de barras mostrando o número de genes anotados em cada categoria funcional, classificados como de isoforma única (azul, “Sem AS”) ou múltiplas isoformas (laranja, “Com AS ≥ 2 ”). As categorias incluem *desintoxicação*, *quimiossensoriais*, *osmorregulação/transporte*, *cutícula/barreira*, *resposta ao estresse*, *defesa imune* e *interação com plantas/digestão*.

5.7 Predição do secretoma (SignalP e Phobius) através dos transcritos do RNA-seq

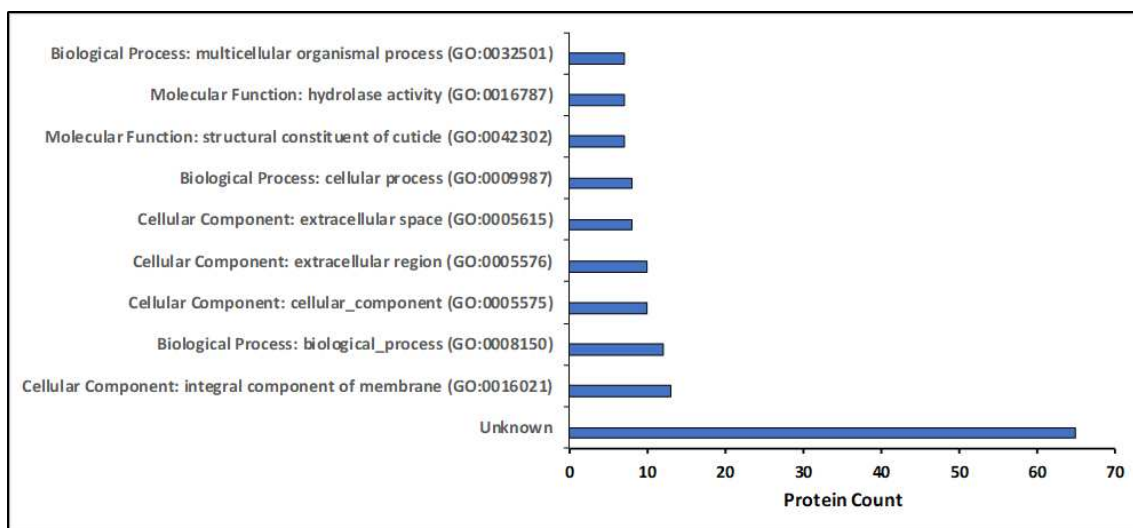
A predição do peptídeo sinal foi realizada no conjunto completo de proteínas previstas a partir da montagem do transcriptoma. O SignalP identificou 168 proteínas como candidatas à secreção, e o Phobius confirmou independentemente os sinais de secreção em 147 dessas candidatas. A forte concordância entre os preditores corrobora sua classificação como proteínas putativas secretadas.

Esse conjunto de proteínas secretadas inclui enzimas (por exemplo, oxidoredutases e isomerases de dissulfeto proteico), chaperonas (como a BiP do retículo endoplasmático), glicoproteínas estruturais (como proteínas relacionadas à cutícula) e proteínas de armazenamento (como hexamerinas), todas potencialmente envolvidas na composição bioquímica da espuma protetora secretada pela glândula.

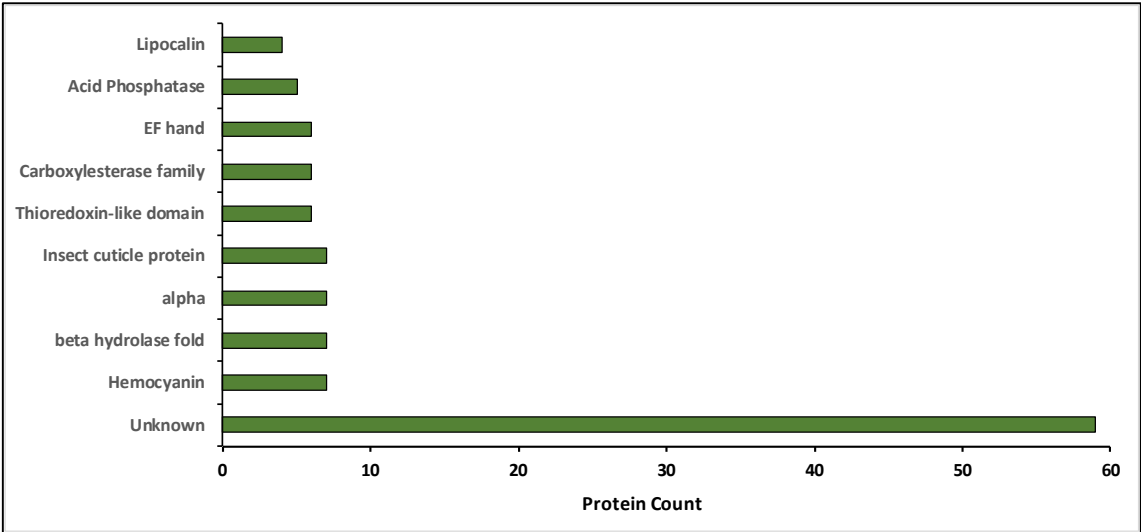
5.8 Anotação funcional do secretoma obtido por RNA-seq

A classificação funcional das 168 proteínas secretadas, baseada em GO, KEGG e PFAM, revelou enriquecimento em categorias relacionadas à localização extracelular, funções estruturais e atividade enzimática (Figura 4 e Tabela S8). A anotação GO destacou termos relacionados à membrana e à localização extracelular, com 'componente integral da membrana' (13 proteínas) e 'região extracelular' (10 proteínas) entre as categorias mais representadas (Tabela 2). As atribuições funcionais também incluíram proteínas associadas a processos metabólicos, ligação a ATP e atividades catalíticas gerais, corroborando o papel do secretoma na atividade enzimática extracelular e nas interações moleculares. A análise KEGG recuperou múltiplas famílias de enzimas, incluindo K17619 (6 proteínas), K13809 (6 proteínas) e K03927 (6 proteínas), que estão ligadas ao metabolismo redox e às vias de processamento de proteínas (Tabela 2). Anotações adicionais foram atribuídas a enzimas envolvidas na desintoxicação e no transporte, o que está de acordo com a necessidade da glândula de gerenciar compostos derivados de plantas e contribuir para a composição da matriz da espuma.

A)



B)



C)

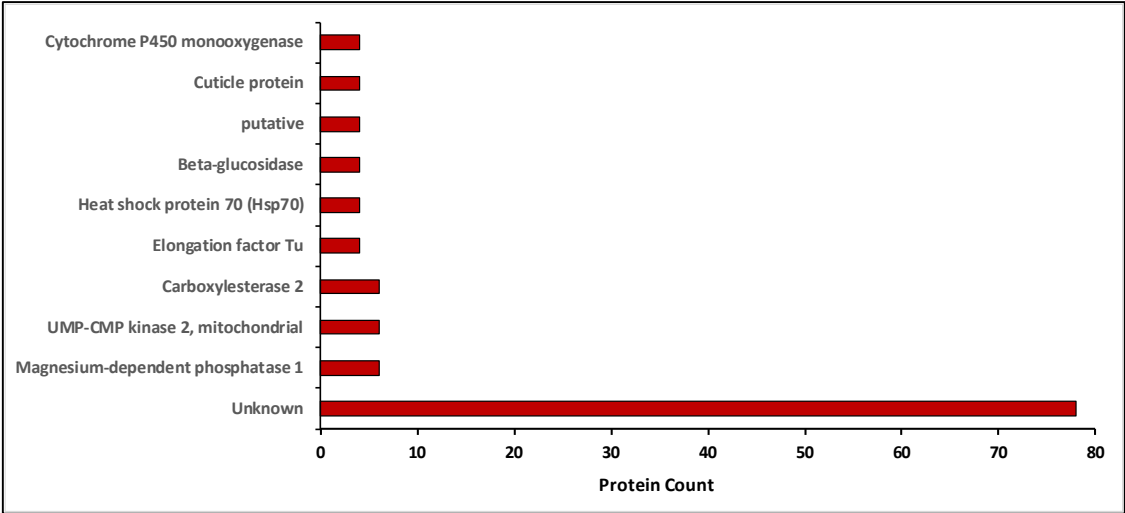


Figura 2 - Principais anotações funcionais do secretoma com base nos bancos Pfam (verde), KEGG (azul) e Gene Ontology (GO; laranja). O gráfico de barras apresenta as dez categorias mais representadas recuperadas de cada base de dados, com os identificadores KEGG substituídos por nomes funcionais descritivos.

Tabela 2 - Principais categorias funcionais do secretoma previsto a partir da análise de RNA-seq de *Mahanarva spectabilis*. A tabela resume os termos mais representados de Gene Ontology (GO), KEGG e Pfam. O número de proteínas associadas a cada termo de anotação é mostrado para comparação entre os bancos de dados.

Termo	Contagem	Fonte
Desconhecidas	65	GO
Componente celular: proteínas integrais de membrana (GO:0016021)	13	GO
Processos biológicos: processos biológicos (GO:0008150)	12	GO
Componente celular: componente celular (GO:0005575)	10	GO
Componente celular: região extracelular (GO:0005576)	10	GO

Componente celular: espaço extracelular (GO:0005615)	8	GO
Processos biológicos: processos celulares (GO:0009987)	8	GO
Função molecular: constituinte estrutural da cutícula (GO:0042302)	7	GO
Função molecular: atividade de hidrolase (GO:0016787)	7	GO
Processos biológicos: processos de organismos multicelulares (GO:0032501)	7	GO
Desconhecidas	78	KEGG
Fosfatase magnésio dependente 1	6	KEGG
UMP-CMP quinase 2, mitocondrial	6	KEGG
Carboxilesterase 2	6	KEGG
Fator de elongação Tu	4	KEGG
Proteína do choque térmico70 (Hsp70)	4	KEGG
Beta-glucosidase	4	KEGG
Putativa	4	KEGG
Proteína da cutícula	4	KEGG
Citocromo P450 monooxygenase	4	KEGG
Desconhecidas	59	PFAM
Hemocianina	7	PFAM
Dobra beta hidrolase	7	PFAM
Alfa	7	PFAM
Proteína da cutícula de inseto	7	PFAM
Domínio Tiorredoxina -like	6	PFAM
Família Carboxilesterase	6	PFAM
Motivo EF hand	6	PFAM
Acido fosfatase	5	PFAM
Lipocalina	4	PFAM

A análise de domínio Pfam revelou a predominância de motivos estruturais e enzimáticos (Tabelas 2 e S8). Entre eles, domínios semelhantes à hemocianina (7 proteínas), dobras beta-hidrolase (7 proteínas) e domínios de proteínas da cutícula de insetos (7 proteínas) foram especialmente abundantes, reforçando os papéis das glicoproteínas ligadoras de oxigênio, enzimas hidrolíticas e proteínas estruturais na composição do secretoma.

A análise KEGG (Tabelas 2 e S8) recuperou múltiplas famílias de enzimas, incluindo a fosfatase 1 dependente de magnésio (K17619), a UMP-CMP quinase 2 mitocondrial (K13809) e a carboxilesterase 2 (K03927), cada uma representada por diversas proteínas. Essas enzimas estão ligadas ao metabolismo redox, ao metabolismo de nucleotídeos, à desintoxicação e às vias de processamento de proteínas. Outros termos KEGG mapeados incluem aqueles envolvidos no transporte e na secreção extracelular.

5.9 Proteínas secretadas ligadas à interação planta-hospedeiro e à sobrevivência ambiental

O conjunto de 168 proteínas secretadas de alta confiança incluiu múltiplas classes funcionais potencialmente associadas à interação planta–inseto e à sobrevivência ambiental (Tabela 3). Entre elas, enzimas relacionadas à desintoxicação e ao estado redox foram representadas por nove candidatas, incluindo carboxilesterases, oxidorreduções associadas à glutatona e proteínas relacionadas ao citocromo P450, sugerindo um papel no processamento extracelular de compostos derivados de plantas e xenobióticos.

Tabela 3 - Proteínas secretadas por *Mahanarva spectabilis* previstas da glândula de Batelli, identificadas pela presença de peptídeos sinal N-terminais (SignalP/Phobius). As proteínas estão agrupadas em categorias funcionais relevantes para a interação planta-hospedeiro e sobrevivência ambiental. A tabela lista o número de proteínas secretadas previstas em cada categoria e os respectivos IDs de proteína.

Functional Group	Secreted Proteins	ProteinID
Chaperones / Protein folding	4	Gene_25378, Gene_35197, Gene_50668, Gene_99264
Chemosensory	1	Gene_126242
Cuticle / Structural	10	Gene_41268, Gene_38689, Gene_53331, Gene_29966, Gene_31904, Gene_24102, Gene_34920, Gene_126386, Gene_56704, Gene_81905
Detoxification / Redox	9	Gene_47720, Gene_114391, Gene_79666, Gene_24447, Gene_23337, Gene_125559, Gene_100906, Gene_123132, Gene_78940
Immune / Antimicrobial	3	Gene_32042, Gene_89309, Gene_28231
Osmoregulation / Enzymes	8	Gene_28076, Gene_18758, Gene_117898, Gene_47660, Gene_60582, Gene_64326, Gene_33482, Gene_6183
Proteases / Protease inhibitors	11	Gene_52229, Gene_17926, Gene_23192, Gene_177, Gene_125745, Gene_60582, Gene_42899, Gene_52512, Gene_25589, Gene_95455, Gene_46772
Storage / Lipocalins	5	Gene_32042, Gene_89309, Gene_27730, Gene_99181, Gene_76286

Proteases e inibidores de proteases também foram proeminentes, com onze proteínas contendo peptídeo sinal, incluindo proteases de serina e serpinas. Essas proteínas podem contribuir para a remodelação da matriz extracelular dentro da espuma e para processos relacionados à defesa. Além disso, dez proteínas secretadas foram classificadas como componentes

cuticulares ou estruturais, indicando uma possível contribuição para a formação ou o reforço de barreiras na interface espuma–inseto.

Proteínas associadas a funções imunes e antimicrobianas foram representadas por três candidatas com assinaturas semelhantes a lisozima ou fenoloxidase, consistentes com papéis protetores no microambiente da espuma. Além disso, oito proteínas secretadas foram relacionadas à osmorregulação e ao transporte de solutos, incluindo anidrases carbônicas e transportadores associados à membrana, refletindo adaptações fisiológicas ao ambiente aquoso e ionicamente diluído característico de insetos sugadores de seiva.

O secretoma também incluiu quatro chaperonas da via secretória, como isomerases de dissulfeto de proteína e proteínas de choque térmico, que podem auxiliar no dobramento e na estabilidade de proteínas exportadas no meio extracelular. Por fim, cinco candidatas secretadas foram anotadas como proteínas carreadoras ou relacionadas ao armazenamento, potencialmente envolvidas no tamponamento da composição química da espuma ou na interação com moléculas derivadas de plantas. Uma única proteína de ligação a odorantes/quimiossensorial contendo peptídeo sinal foi detectada, sugerindo um possível papel na ligação a compostos voláteis ou semioquímicos na interface planta–espuma.

5.10 Identificação e perfil quantitativo do proteoma da glândula de Batelli por LC-MS/MS

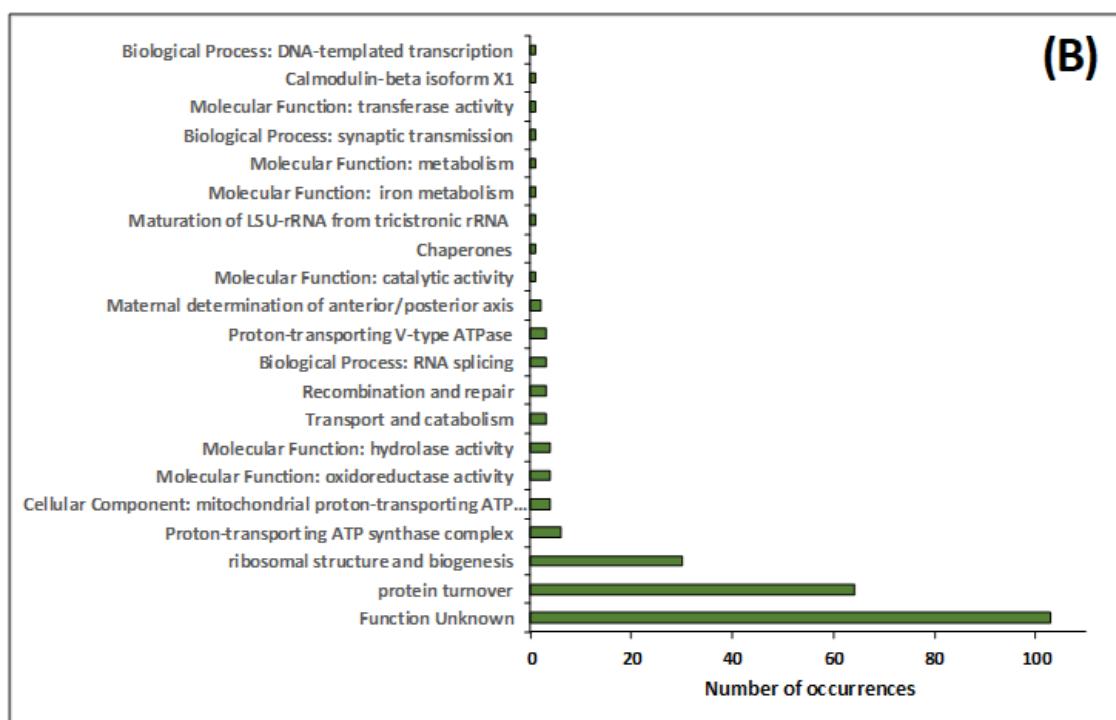
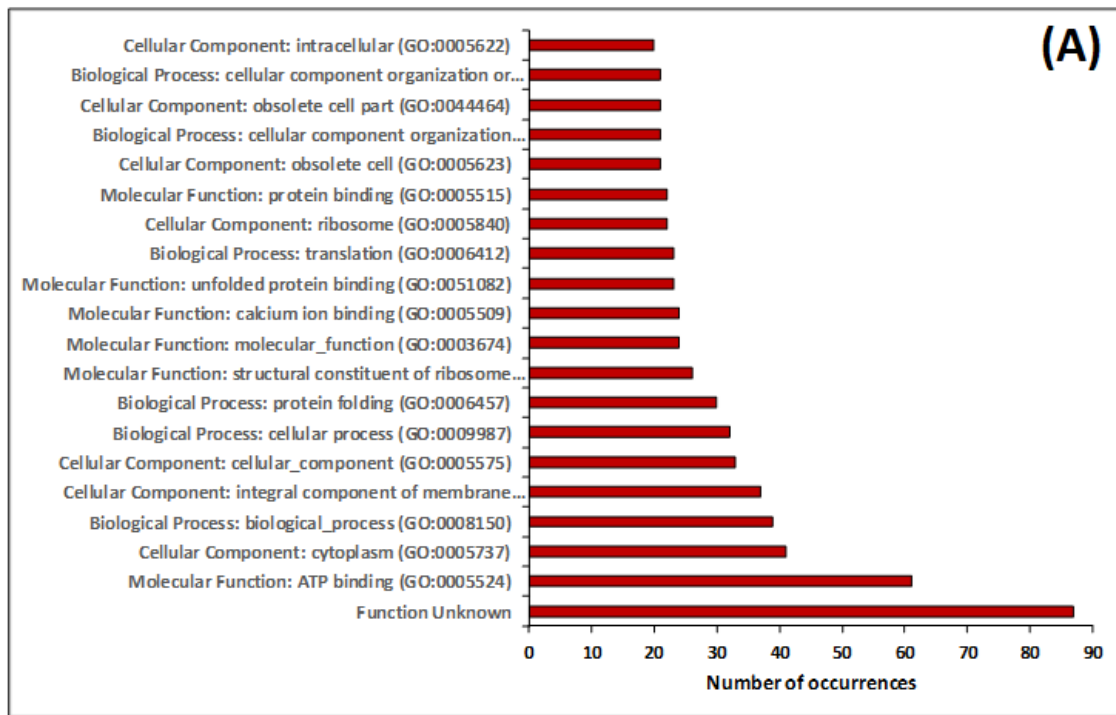
A análise shotgun por LC–MS/MS da glândula de Batelli de ninfas de *M. spectabilis* gerou 30.270 varreduras MS e 14.444 varreduras MS/MS (Orbitrap MS com LIT MS/MS; nano-ESI, CID de alta energia). A busca em banco de dados no PEAKS (modo homologia) e os filtros aplicados (peptídeo –10lgP $\geq$ 17; proteína –10lgP $\geq$ 20; $\geq$ 2 peptídeos únicos; de novo ALC $\geq$ 50%) permitiram a identificação de 8.094 PSMs, 3.806 peptídeos distintos e 500 proteínas agrupadas em 493 grupos proteicos, com FDRs de 0,5% (PSM), 0,8% (peptídeo) e 0,0% (proteína). A maioria das proteínas (371) foi suportada por mais de dois peptídeos únicos (129 com exatamente dois). Os controles de qualidade indicaram predominância de peptídeos totalmente trípticos (3.681 sem clivagens perdidas; 125 com uma clivagem perdida), e 3.650 espectros foram classificados

como “*de novo only*”. As distribuições de erro de massa do precursor e de escores são apresentadas na saída do controle de qualidade (CQ).

5.11 Anotação funcional e interpretação biológica do proteoma da glândula de Batelli

Com base nas proteínas identificadas experimentalmente por LC-MS/MS na Seção 5.10, examinamos a seguir sua anotação funcional e significado biológico usando análises de Pfam, KEGG e Gene Ontology.

A análise proteômica do secretoma da glândula de Batelli identificou um conjunto diverso de 500 proteínas abrangendo categorias funcionais associadas à homeostase celular, adaptação ao estresse e interações com o ambiente hospedeiro (Figura 5 e Tabela S9). A anotação funcional do proteoma da glândula de Batelli utilizando Pfam, KEGG e Gene Ontology (GO) revelou ampla representação de domínios e vias relacionadas à tolerância ao estresse, atividade metabólica e secreção extracelular.



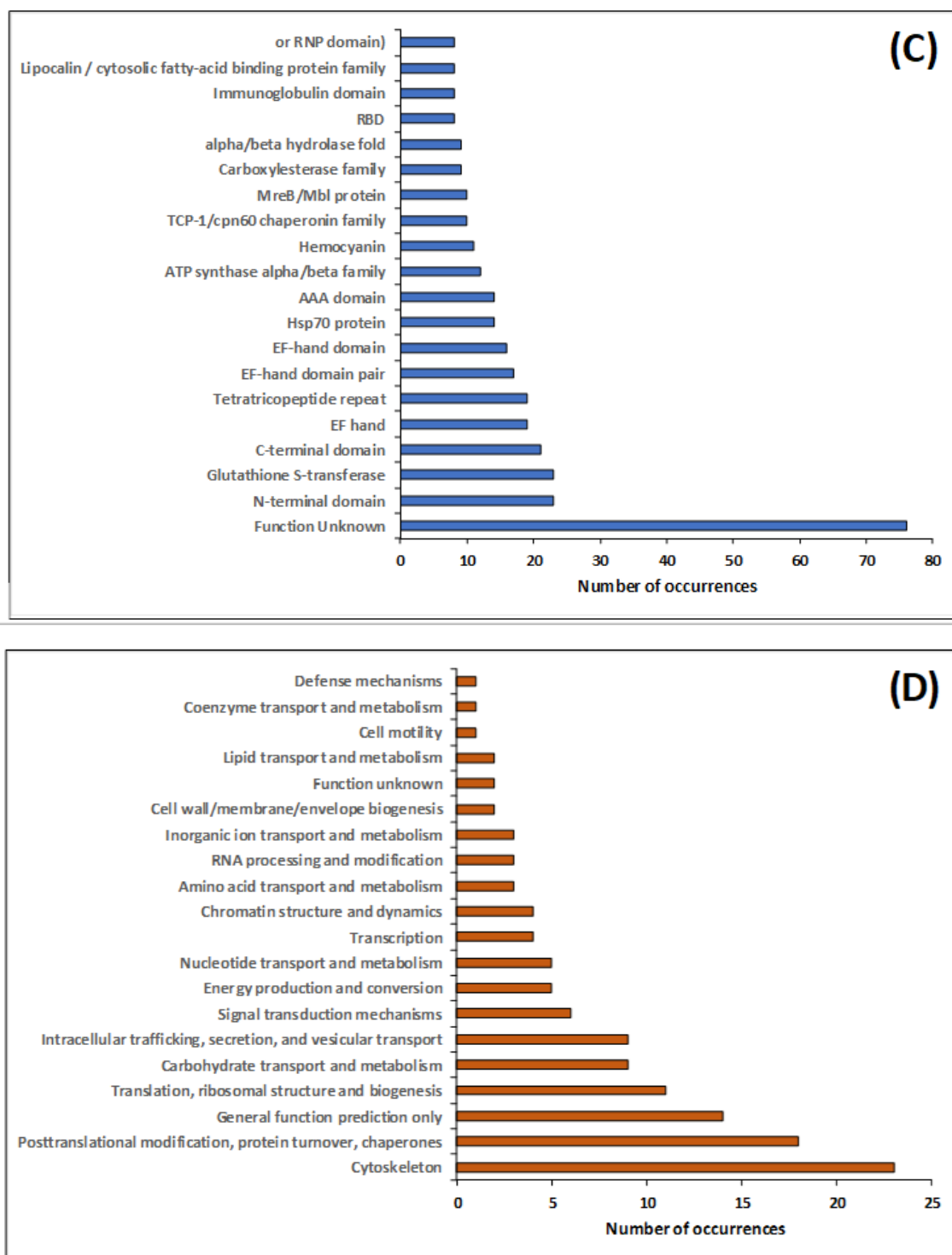


Figura 3 - Classificação funcional das proteínas identificadas na glândula de Batelli de ninfas de *Mahanarva spectabilis*. (A) Distribuição dos 20 domínios Pfam mais frequentes, destacando motivos estruturais conservados e módulos enzimáticos prevalentes no conjunto de dados. (B) Anotação Gene Ontology (GO) dos 20 termos mais enriquecidos, incluindo categorias relacionadas à função molecular, processo biológico e componente celular, ilustrando a diversidade funcional do proteoma glandular. (C) As 20 principais vias KEGG representadas entre as proteínas identificadas, revelando rotas metabólicas e de sinalização potencialmente associadas à secreção, estabilização da espuma e interações inseto-planta. (D) As 20 principais categorias funcionais KOG, organizadas por processos celulares e metabólicos de ordem superior, refletindo o contexto evolutivo e fisiológico das funções proteicas na glândula de Batelli.

A análise de domínios Pfam destacou famílias proteicas abundantes, como glutatona S-transferases (23 ocorrências), proteínas ligantes de cálcio do

tipo EF-hand (19 ocorrências) e domínios relacionados a chaperonas, consistentes com funções em desintoxicação, sinalização e dobramento proteico. Uma proporção substancial das proteínas apresentou domínios gerais N- e C-terminais, sugerindo motivos estruturais conservados, enquanto outras estiveram associadas a funções enzimáticas, incluindo oxidoredutases e transferases.

A classificação pelo KEGG recuperou 103 proteínas sem anotação definida; entretanto, entre as entradas funcionalmente caracterizadas, as categorias enriquecidas incluíram proteínas mitocondriais (29 ocorrências), enzimas de vias de desintoxicação e funções relacionadas ao transporte. Diversas proteínas foram mapeadas em rotas metabólicas relevantes para a adaptação ao hospedeiro, incluindo metabolismo de xenobióticos e produção de energia.

A anotação por GO corroborou esses achados. Os termos mais representados incluíram ligação a ATP (61 proteínas), refletindo elevada demanda energética; localização citoplasmática (41 proteínas); e funções associadas à membrana (37 proteínas), muitas das quais correspondem a transportadores iônicos e enzimas. Os termos de processos biológicos destacaram respostas ao estresse e processos metabólicos proteicos, alinhando-se aos papéis previstos da glândula na manutenção da estabilidade proteica e na contribuição para a formação da espuma. Notavelmente, 35 proteínas ($\approx 7\%$ do conjunto de dados) permaneceram sem anotação funcional, evidenciando a presença de componentes potencialmente novos ou pouco caracterizados do proteoma da glândula.

A identificação de proteínas realizada utilizando o algoritmo PEAKS a partir de dados de LC-MS/MS permitiu inferir a abundância relativa de proteínas nas amostras com base na abordagem de contagem espectral. Assim, proteínas com maior número de peptídeos únicos e/ou maior cobertura de sequência foram consideradas relativamente mais abundantes, visto que esses parâmetros são comumente associados a maior abundância de proteínas em análises proteômicas sem marcação. Esses resultados fornecem informações sobre os principais constituintes proteicos de cada banda associada à espuma e servem como base para a subsequente anotação funcional e estrutural.

Quando a abundância proteica foi inferida a partir do número de peptídeos detectados, o conjunto de dados foi dominado por chaperonas altamente

expressas, reguladores pós-traducionais e enzimas metabólicas (Tabela 4). As proteínas mais abundantes incluíram múltiplos homólogos de HSP70 (por exemplo, proteínas do tipo DnaK), com mais de 80 peptídeos, enfatizando seu papel crítico na proteostase. Outras chaperonas abundantes incluíram chaperoninas de 60 kDa, cada uma representada por cerca de 70 peptídeos, evidenciando o forte investimento da glândula em dobramento e estabilização proteica.

Tabela 4 - 15 proteínas mais abundantes da glândula de Batelli a partir de análises proteômicas

ID da proteína	Contagem de peptídeos	Categoria funcional	Anotação/Termo correspondente
Gene_2 8070	81	Chaperona	Proteína chaperona DnaK
Gene_1 09130	79	Modificação pós traducional	Proteína de renovação (turnover)
Gene_2 8076	79	Parede celular/membrana/ biogênese do envelope	Osmoregulação e transporte
Gene_2 6924	72	Produção de energia	ATPase transportadora de H ⁺ do tipo F subunit d
Gene_2 9120	69	Chaperona	60 kDa chaperonina
Gene_8 6297	69	Transporte de carboidratos e metabolismo	Triosefosfato isomerase
Gene_2 9122	69	Chaperona	60 kDa chaperonina
Gene_2 5378	67	Proteína de renovação (turnover)	Chaperona
Gene_2 6709	64	Proteína de renovação (turnover)	Superfamília RNI-like
Gene_7 3327	63	Constituinte estrutural do ribossomo	Proteína ribossômica 60S acídica P2
Gene_2 8072	60	Modificação pós traducional	Proteína hipotética B566
Gene_3 8689	60	Constituinte estrutural da cutícula	Proteína da barreira cuticular
Gene_2 5201	58	Desconhecida	Desconhecida
Gene_2 8090	58	ATP sintase subunidade beta	ATP sintase subunidade beta
Gene_3 153	58	Estrutura do ribossomo e biogênese	Proteína ribossomal P1

Proteínas associadas à modificação pós-traducional e ao *turnover* proteico também foram altamente representadas (>70 peptídeos), ressaltando a intensa regulação do controle de qualidade proteica. Candidatos ligados à osmorregulação e ao transporte de membrana foram detectados em níveis semelhantes, sugerindo homeostase iônica ativa e balanço hídrico dentro da glândula. A resiliência metabólica foi refletida na detecção da subunidade β da

ATP sintase, com quase 60 peptídeos, indicando elevadas demandas energéticas sustentadas. A adaptação estrutural foi ainda sugerida por proteínas anotadas como envolvidas na remodelação da barreira cuticular, enquanto diversas proteínas abundantes, porém não caracterizadas (60–70 peptídeos), podem representar novos contribuintes para a fisiologia da glândula. Esse perfil proteômico ressalta a ênfase da glândula de Batelli na estabilidade proteica, atividade metabólica, osmorregulação e funções de barreira extracelular. Essas características são consistentes com seu papel especializado na secreção de espuma protetora enriquecida em proteínas estruturais e defensivas.

5.12 Proteínas candidatas potencialmente relevantes para a interação planta-hospedeiro e sobrevivência ambiental, identificadas por análise proteômica

A análise proteômica da glândula de Batelli identificou diversos grupos de proteínas potencialmente associados à interação com a planta hospedeira e à adaptação ambiental (Figura 6 e Tabela S10). Entre as proteínas anotadas, 33 candidatas foram atribuídas a funções relacionadas à desintoxicação, incluindo monooxigenases do citocromo P450, glutathione S-transferases (GSTs) e outras oxidoredutases. Essas proteínas são comumente associadas ao processamento de compostos derivados de plantas e xenobióticos e estiveram bem representadas no proteoma da glândula.

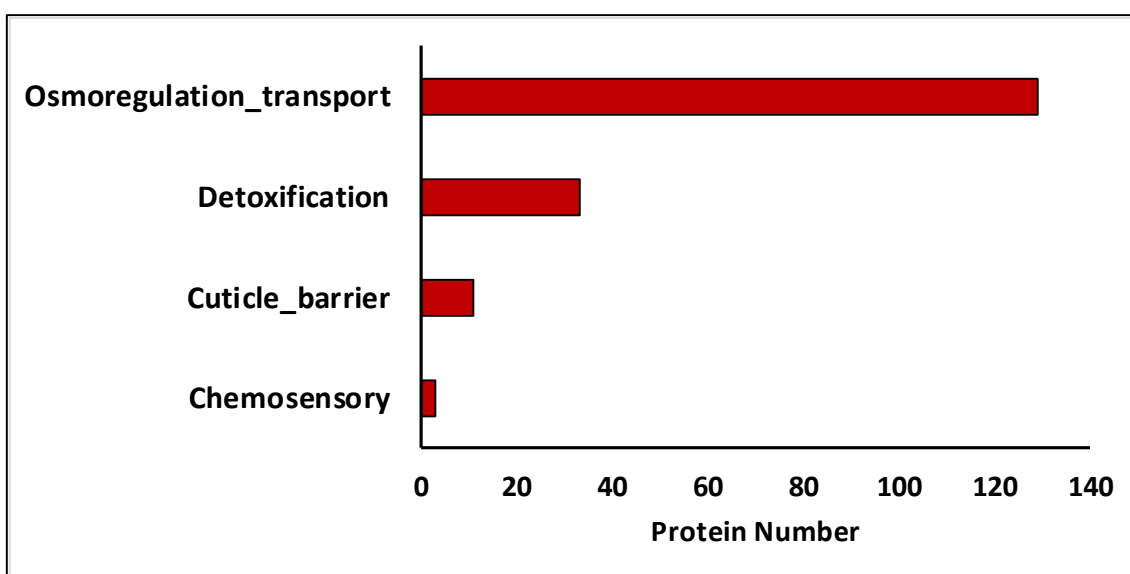


Figura 4 - Distribuição das proteínas por categoria funcional a partir da análise proteômica. As barras indicam o número de proteínas anotadas por categoria (Desintoxicação, Quimiossensoriais, Cutícula/Barreira, Osmorregulação/Transporte).

Um conjunto menor, composto por três proteínas, foi classificado como relacionado à quimiorrecepção, incluindo proteínas de ligação a odorantes (OBPs/PBPs) e proteínas quimiossensoriais (CSPs). A presença dessas proteínas no proteoma da glândula de Batelli sugere uma possível associação com a percepção ou o processamento de sinais químicos na interface entre o inseto, a espuma e a planta hospedeira.

Proteínas envolvidas na osmorregulação e no transporte de íons constituíram uma fração proeminente do conjunto de dados, com 129 proteínas anotadas, incluindo subunidades de v-ATPase, Na^+/K^+ -ATPases, transportadores da família solute carrier e anidrases carbônicas. Essas proteínas estão tipicamente associadas à homeostase iônica e osmótica e podem refletir exigências fisiológicas ligadas à alimentação com seiva do xilema. Além disso, 11 proteínas foram atribuídas à remodelação da cutícula e de barreiras, incluindo proteínas estruturais cuticulares e enzimas relacionadas ao metabolismo de quitina.

6. DISCUSSÃO

A cigarrinha-das-pastagens é uma praga de gramíneas forrageiras, uma vez que tanto adultos quanto ninfas se alimentam da seiva do xilema, causando reduções substanciais no rebrote, na qualidade e na produtividade das plantas. As ninfas são particularmente danosas porque permanecem fixadas aos tecidos basais das plantas por longos períodos, produzem uma espuma protetora e criam condições que predispõem os hospedeiros a infecções secundárias, acelerando ainda mais o declínio das pastagens. Diante disso, estratégias para o controle sustentável de surtos de cigarrinhas vêm sendo amplamente investigadas. Entre elas estão a indução de resistência das plantas aos ataques da cigarrinha (Alvarenga *et al.*, 2017; Alvarenga *et al.*, 2019; Leite *et al.*, 2014), a diversificação de pastagens (Alvarenga *et al.*, 2017), o uso de compostos vegetais como agentes de controle (Dias *et al.*, 2019; Nascimento *et al.*, 2021), a adubação do solo (Aguilar *et al.*, 2014; Alvarenga *et al.*, 2019) e o uso de fungos entomopatogênicos (Campagnani *et al.*, 2017; Campagnani *et al.*, 2024). Somadas a essas estratégias, as adaptações moleculares que sustentam sua sobrevivência na seiva do xilema pobre em nutrientes e sua capacidade de resistir às defesas das plantas foram propostas nesta pesquisa para a espécie *Mahanarva spectabilis*.

As ninfas de *M. spectabilis* também se destacam por produzirem uma espuma característica que envolve seus corpos. Essa espuma, produzida pelas secreções da glândula de Batelli em combinação com os fluidos excretórios associados aos túbulos de Malpighi, foi descrita em cigarrinhas como uma barreira multifuncional contra dessecação, flutuações de temperatura, predação e crescimento microbiano (Rakitov, 2002; Szterk *et al.*, 2024; Tonelli *et al.*, 2020).

Estudos recentes com outras espécies de cigarrinhas, como *M. fimbriolata* e *Poophilus costalis*, confirmam que essas espumas contêm misturas complexas de proteínas, carboidratos e lipídios que garantem estabilidade térmica e proteção antimicrobiana (Sahayaraj *et al.*, 2021; Tonelli *et al.*, 2018). Ao focar no transcriptoma das ninfas e no proteoma da glândula de Batelli, nosso trabalho oferece uma visão abrangente dos componentes moleculares subjacentes à produção da espuma e de suas potenciais funções nas interações planta–inseto–ambiente.

As estatísticas de montagem confirmaram a confiabilidade do conjunto de dados, com um comprimento N50 dentro da faixa esperada para transcriptomas de insetos e representação suficiente em diferentes categorias de comprimento. A presença de unigenes curtos e longos indica uma montagem eficiente em genes altamente e moderadamente expressos, fornecendo um recurso valioso para a descoberta de genes e análise de expressão.

O RNA-seq das ninfas forneceu um recurso genômico em larga escala de genes expressos, anotados nas categorias Pfam, KEGG e GO, enquanto a proteômica por LC-MS/MS revelou centenas de proteínas secretadas pela glândula de Batelli. A integração desses conjuntos de dados permitiu a identificação de genes e proteínas com funções previstas em desintoxicação, percepção quimiossensorial, osmorregulação, respostas ao estresse, imunidade e remodelação da cutícula. Abordagens integrativas semelhantes foram aplicadas com sucesso em outros hemípteros, como a cigarrinha *Nephotettix cincticeps* (Wu *et al.*, 2023) e o pulgão *Metopolophium dirhodum* (Gebrekidan *et al.*, 2024), confirmando que o acoplamento transcriptoma-proteoma é uma estratégia poderosa para elucidar a base molecular de estruturas secretadas em insetos sugadores de seiva.

Detectamos um amplo repertório de genes de desintoxicação, incluindo CYPs, GSTs, UGTs, CCEs e AKRs. Essas famílias gênicas apresentaram forte enriquecimento em vias KEGG, como metabolismo da glutathione e metabolismo de fármacos/xenobióticos. Enzimas de desintoxicação são amplamente reconhecidas como mediadoras centrais da adaptação ao hospedeiro em hemípteros, como demonstrado em ninfas de *Empoasca vitis*, nas quais muitos genes de desintoxicação são regulados positivamente durante a alimentação (Shao *et al.*, 2021). De forma semelhante, expansões de CCEs, GSTs e P450s foram relatadas em *Picromerus lewisi* (Li *et al.*, 2022), destacando sua importância evolutiva nas interações inseto-planta. Notavelmente, o metabolismo da glutathione também foi relatado como uma via chave associada à resistência em *Paspalum regnellii* sob ataque de *M. spectabilis*, destacando seu papel central na modulação das respostas de defesa da planta e dos processos de desintoxicação do inseto dentro dessa interação (Begnami *et al.*, 2025). O robusto repertório de desintoxicação em *M. spectabilis* provavelmente reflete a adaptação a aleloquímicos e metabólitos secundários presentes nas gramíneas forrageiras.

Os transcritos quimiossensoriais identificados incluem proteínas ligantes de odor (OBPs), proteínas ligantes de feromônios (PBP), proteínas quimiossensoriais (CSPs) e receptores gustativos e olfatórios. Essas proteínas desempenham papéis essenciais na detecção de voláteis derivados do hospedeiro e são cruciais para orientação, alimentação e oviposição. Em pulgões como *Myzus persicae*, análises de RNA-seq mostraram que a expressão de genes quimiossensoriais varia entre ninfas e adultos, refletindo necessidades ecológicas e de desenvolvimento (Sharma *et al.*, 2024). A alta expressão de membros da família PBP/GOBP em *M. spectabilis* pode estar igualmente relacionada à localização do hospedeiro e ao comportamento gregário das ninfas dentro da espuma.

A alimentação com seiva do xilema impõe desafios osmóticos singulares devido ao alto teor de água e baixo valor nutricional. Nossos dados revelaram quase 4.000 transcritos relacionados ao transporte, incluindo múltiplas subunidades de v-ATPases, Na⁺/K⁺-ATPases, aquaporinas, transportadores da família solute carrier e transportadores iônicos. Anidrases carbônicas e ATPases bombeadoras de prótons estiveram entre os transcritos mais abundantes. Esses achados estão em concordância com observações em cigarras e outros insetos sugadores de xilema, nos quais o aparato de osmorregulação é altamente expandido (Beckett *et al.*, 2019). Tais adaptações permitem que os insetos mantenham o equilíbrio iônico e osmótico enquanto processam grandes volumes de seiva diluída.

Identificamos transcritos para proteínas cuticulares (CPRs) e enzimas do metabolismo da quitina, sugerindo remodelação do exoesqueleto. A presença dessas proteínas entre os transcritos associados às glândulas de Batelli sugere uma possível contribuição indireta para o ambiente da espuma, possivelmente reforçando a interface entre o corpo do inseto e a espuma ou fornecendo propriedades estruturais ou antimicrobianas. Isso é consistente com relatos em *Aphrophora alni*, nos quais proteínas e componentes associados à quitina foram detectados na espuma (Szterk *et al.*, 2024). Contudo, o papel dessas proteínas na estabilidade da espuma deve ser considerado uma hipótese funcional, e não um mecanismo comprovado. A cutícula e a espuma parecem, portanto, formar uma barreira integrada que protege as ninfas contra danos mecânicos e invasão microbiana.

A abundância de transcritos para proteínas de choque térmico e enzimas antioxidantes confirma a importância da tolerância ao estresse. Além disso, transcritos relacionados à imunidade, como serino proteases, serpinas, lisozimas e proteínas de reconhecimento de peptidoglicano (PGRPs), refletem a necessidade de defesa dentro do habitat espumoso. A atividade antimicrobiana já foi demonstrada em espumas de cigarrinhas (Tonelli *et al.*, 2018), e proteínas do tipo lisozima também foram previstas em nosso secretoma, fortalecendo a relação entre proteínas secretadas e defesa microbiana.

Aproximadamente 16% dos *loci* expressos foram classificados como potencialmente sujeitos a *splicing* alternativo, incluindo enzimas de desintoxicação, receptores e genes relacionados à imunidade. Essa plasticidade transcriptômica é consistente com achados em pulgões e cigarrinhas, nos quais mudanças no hospedeiro vegetal induzem ajustes rápidos na expressão gênica (Rivera-Vega *et al.*, 2017; Chongtham *et al.*, 2024). O *splicing* alternativo pode, portanto, ser um mecanismo crucial que permite a *M. spectabilis* adaptar-se a diferentes hospedeiros de pastagem e estressores ambientais.

As análises por SignalP e Phobius previram 168 proteínas secretadas de alta confiança em *M. spectabilis*, incluindo múltiplas enzimas de desintoxicação, proteínas estruturais, diversas proteases, fatores relacionados à imunidade, chaperonas e proteínas de armazenamento/transporte. A predição funcional foi sustentada por anotações nos bancos Pfam, KEGG e GO, e a identidade proteica foi validada por evidências de LC/MS. Abordagens integrativas semelhantes em hemípteros sugadores de seiva, como pulgões e cigarrinhas, revelaram conjuntos de proteínas secretadas implicadas na interação planta–inseto, desintoxicação e modulação da defesa (Wu *et al.*, 2023; Zhang *et al.*, 2023). Esses estudos demonstram que análises multi-ômicas podem identificar e caracterizar sistematicamente proteínas que contribuem para o sucesso da alimentação em seiva e para a adaptação a diferentes plantas hospedeiras.

Detectamos altos níveis de enzimas de desintoxicação e transportadores no proteoma da glândula de Batelli, indicando um papel secretor no processamento de xenobióticos derivados do hospedeiro. Proteases e inibidores de proteases foram fortemente representados, sugerindo tanto a remodelação da matriz da espuma quanto a ativação de cascatas imunes, como observado em outros proteomas salivares de hemípteros (Da Silva Vaz Junior *et al.*, 2024). Proteínas estruturais, como proteínas cuticulares, e enzimas do metabolismo da

quitina também estavam presentes, sustentando sua contribuição para a estabilidade da espuma, característica demonstrada em análises de bioespuma de cigarrinhas (Sahayaraj *et al.*, 2021). Proteínas imunes, incluindo lisozimas e fenoloxidases, bem como chaperonas moleculares e fatores de dobramento como BiP e isomerases de dissulfeto proteico (PDIs), reforçam o caráter multifuncional da matriz da espuma (Hendershot *et al.*, 2024). Notavelmente, a identificação de proteínas de armazenamento, incluindo hexamerinas e lipocalinas, enfatiza funções adicionais de tamponamento e transporte das proteínas secretadas.

O proteoma da glândula de Batelli demonstra a multifuncionalidade e a importância ecológica da espuma das cigarrinhas. Ela atua como uma barreira bioquímica e estrutural, assegura a regulação osmótica, medeia a desintoxicação de aleloquímicos vegetais e fornece defesa antimicrobiana. Estudos ecológicos confirmaram que a espuma mantém um microclima favorável e protege contra dessecação e predadores (Beckett *et al.*, 2019; Tonelli *et al.*, 2018). A grande fração de proteínas secretadas não anotadas sugere a presença de componentes novos, possivelmente específicos de cigarrinhas, que podem representar inovações evolutivas únicas contribuindo para o sucesso adaptativo de *M. spectabilis*.

Quimicamente, a espuma das cigarrinhas é um material complexo, anfipático e viscoelástico, contendo proteínas (incluindo glicopeptídeos e proteoglicanos), carboidratos e, dependendo da espécie, ácidos graxos como os ácidos palmítico e esteárico. Essa composição é consistente com uma matriz estrutural capaz de capturar ar e resistir ao colapso. Funcionalmente, a espuma amortece variações de temperatura e umidade, protege contra a luz e dificulta a ação de inimigos naturais, além de apresentar propriedades antimicrobianas, incluindo atividades antibacteriana e antifúngica. Comunidades microbianas já foram descritas em espumas de espécies relacionadas de *Mahanarva*, sugerindo interações ecológicas adicionais moduladas pela química da espuma (Sahayaraj *et al.*, 2021).

Em consonância com esses relatos, as proteínas derivadas da glândula de Batelli, positivas para SignalP e caracterizadas neste estudo, provavelmente constituem o núcleo proteico da espuma e sustentam múltiplas funções não excludentes. Proteínas secretadas, glicosiladas e de baixa complexidade, como aquelas com domínios do tipo mucina e regiões ricas em Pro/Thr/Ser, juntamente

com módulos associados a proteoglicanos, podem contribuir para a viscosidade, reduzir a tensão superficial e aumentar a persistência das bolhas, em acordo com descrições históricas de glicopeptídeos e proteoglicanos na espuma. Paralelamente, hidrolases e enzimas redox secretadas, incluindo possíveis quitinases e proteínas do tipo oxidoredutase, podem sustentar as atividades antifúngica e antibacteriana relatadas para a espuma, contribuindo para o controle da colonização microbiana no microhabitat ninfal. Outras proteínas com propriedades semelhantes a surfactantes, juntamente com proteínas ligantes ou transportadoras de lipídios, são compatíveis com os ácidos graxos documentados na espuma e podem manter um filme de bolhas estável que regula temperatura, exposição à luz e propriedades reológicas. Mecanicamente, a espuma é gerada por bombeamento abdominal periódico que injeta ar nas secreções das glândulas de Batelli e dos túbulos de Malpighi. A secreção da glândula de Batelli, em particular, facilita a formação das bolhas, tornando o fluido anal espumoso, conforme descrito em relatos (Guilbeau, 1908; Marshal, 1965). Assim, a interpretação mais coerente é que a glândula de Batelli é a principal fonte dos componentes proteicos e lipídicos que, juntamente com polissacarídeos, constroem a matriz extracelular da espuma protetora, e que as proteínas positivas para SignalP identificadas aqui representam candidatos imediatos secretados para formar e manter essa bioespuma.

De modo geral, o secretoma da glândula de Batelli está alinhado com a visão estabelecida de que essas secreções glandulares são necessárias e suficientes para a produção da espuma das cigarrinhas. Ao fornecer um arcabouço proteico enriquecido em glicoproteínas, surfactantes e enzimas antimicrobianas, em combinação com lipídios e carboidratos, a glândula de Batelli assegura a estabilidade das bolhas, a regulação do microclima e uma primeira barreira defensiva durante o estágio ninfal de *M. spectabilis*, caracterizado pela alimentação em seiva do xilema.

7. CONCLUSÃO

Ao integrar dados transcriptômicos e proteômicos, este estudo apresenta a primeira caracterização molecular aprofundada do estágio ninfal de *M. spectabilis*. A identificação de múltiplas categorias funcionais — incluindo enzimas de desintoxicação (por exemplo, citocromos P450, GSTs), receptores quimiossensoriais (OBPs, CSPs, PBPs), proteínas envolvidas na osmorregulação (aquaporinas, ATPases), proteínas cuticulares e enzimas do metabolismo da quitina, fatores de resposta ao estresse (proteínas de choque térmico e enzimas antioxidantes) e diversas proteínas secretadas — destaca as adaptações multifatoriais que permitem às ninfas de cigarrinhas-das-pastagens sobreviver e se desenvolver no ambiente quimicamente e fisicamente estressante das gramíneas forrageiras.

Esses resultados corroboram e ampliam achados de análises transcriptômicas e proteômicas recentes em outros hemípteros, avançando a compreensão das estratégias moleculares e ecológicas que sustentam a adaptação de insetos sugadores de seiva. De forma relevante, esse arcabouço molecular pode servir como base para o desenvolvimento de estratégias inovadoras e direcionadas de manejo de cigarrinhas-praga em ecossistemas de pastagens, incluindo abordagens baseadas na interrupção de famílias proteicas chave ou de processos fisiológicos relacionados à formação da espuma.

Referências Bibliográficas

- AGUIAR, D.M., AUAD, A.M., FONSECA, M.G. & LEITE, M.V. ***Brachiaria ruzizensis* responses to different fertilization doses and to the attack of *Mahanarva spectabilis* (Hemiptera: Cercopidae) nymphs and adults.** *The Scientific World Journal*, 1–8, 2014.
- ALTSCHUL, S. F. *et al.* **Basic local alignment search tool.** *Journal of Molecular Biology*, v. 215, n. 3, p. 403–410, 1990.
- ALVARENGA, R. *et al.* **Spittlebugs (Hemiptera: Cercopidae) and their host plants: a strategy for pasture diversification.** *Applied Entomology and Zoology*, v. 52, p. 653–660, 2017.
- ALVARENGA, R. *et al.* **Tolerance to nymphs and adults of *Mahanarva spectabilis* (Hemiptera: Cercopidae) by forage plants in fertilized soils.** *Pest Management Science*, v. 75, n. 8, p. 2242–2250, 2019.
- ALZANI, P. *et al.* **Spittlebug invisibility cloak: experimental tests on the antipredatory effect of the froth of *Philaenus spumarius*.** 2023.
- ANDREWS, S. **FastQC: a quality control tool for high throughput sequence data.** 2010.
- ARANTES, A. E.; COUTO, V. R. M.; SANO, E. E.; FERREIRA, L. G. **Livestock intensification potential in Brazil based on agricultural census and satellite data analysis.** *Pesquisa Agropecuária Brasileira*, Brasília, v. 53, n. 9, p. 1053–1060, 2018. DOI: 10.1590/S0100-204X2018000900009.
- AUAD, A. M. *et al.* **Seasonal dynamics of egg diapause in *Mahanarva spectabilis*.** *Arquivos do Instituto Biológico*, v. 78, n. 2, p. 325–330, 2020.
- AUAD, A. M. *et al.* **Spittle protein profile of *Mahanarva spectabilis* fed various elephant grass genotypes.** *Genetics and Molecular Research*, v. 11, n. 4, p. 3601–3606, 2012.
- AUAD, A. M.; SILVA, S. E. B. da. Pasture. In: SOUZA, B.; VÁZQUEZ, L. L.; MARUCCI, R. C. (org.). **Natural Enemies of Insect Pests in Neotropical Agroecosystems.** Cham: Springer, 2019. p. 369–382.
- BECKETT, K.I.S., ROBERTSON, A.B. & MATTHEWS, P.G.D. **Studies on gas exchange in the meadow spittlebug, *Philaenus spumarius*: the metabolic cost of feeding on, and living in, xylem sap.** *Journal of Experimental Biology*, 222(3), 2019.
- BEGNAMI, I.D.S., AONO, A.H., GRACIANO, D.D.S., CARMELLO-GUERREIRO, S.M., ULBRICHT FERREIRA, R.C., MALAGÓ JR. W. *et al.* **Elucidating molecular responses to spittlebug attack in *Paspalum regnellii*.** *Plant Molecular Biology Reporter*, 43, 307–323. 2025. Available from: <https://doi.org/10.1007/s11105-024-01487-w>

BERGMAN, E. A. *The biomechanics and energetics of xylem feeding in the meadow spittlebug (Philaenus spumarius)*. 2021. Tese (Doutorado) – University of British Columbia.

BOLGER, A. M.; LOHSE, M.; USADEL, B. **Trimmomatic: a flexible trimmer for Illumina sequence data**. *Bioinformatics*, v. 30, n. 15, p. 2114–2120, 2014.

BURMESTER, T. **Evolution and function of the insect hexamerins**. *European Journal of Entomology*, v. 96, n. 3, p. 213–225, 1999.

CAMPAGNANI, M.O., AUAD, A.M., MAURÍCIO, R.M., MADUREIRA, A.P., CANGUSSÚ, M., ROSA, L.H. et al. **Endophytic capacity of entomopathogenic fungi in a pasture grass and their potential to control the spittlebug *Mahanarva spectabilis* (Hemiptera: Cercopidae)**. *Agronomy*, 14, 943, 2024. Available from: <https://doi.org/10.3390/agronomy14050943>

CAMPAGNANI, M.O., CAMPOS, W.G., AMORIM, S.S., ROSA, L.H., AUAD, A.M., CANGUSSÚ, M.A. et al. **Prospection and fungal virulence associated with *Mahanarva spectabilis* (Hemiptera: Cercopidae) in an Amazon silvopastoral system**. *Florida Entomologist*, 100(2), 426–432, 2017. Available from: <https://doi.org/10.1653/024.100.0204>

CARRASCOSA, A.; MORENO, G.; RODRIGO, S.; ROLO, V. **Unravelling the contribution of soil, climate and management to the productivity of ecologically intensified Mediterranean wood pastures**. *Science of The Total Environment*, Volume 957, 177575, 2024. <https://doi.org/10.1016/j.scitotenv.2024.177575>.

CENTRO DE ESTUDOS AVANÇADOS EM ECONOMIA APLICADA (CEPEA). PIB do Agronegócio Brasileiro. Piracicaba: Cepea/Esalq-USP; Confederação da Agricultura e Pecuária do Brasil (CNA), 2026. Disponível em: <https://www.cepea.org.br/upload/kceditor/files/PIB-4tri2025-27.ABRIL.2026.pdf>. Acesso em: 6 jul. 2026.

CHEN, X. *et al.* **The role of biofoam in shielding spittlebug nymphs (Insecta, Hemiptera, Cercopidae) against bright light**. *Ecological Entomology*, vol. 43, 2018, <https://doi.org/10.1111/een.12496>.

CHONGTHAM, R.; SHARMA, M.; SHUKLA, R.N.; JOSHI, G.; KUMAR, A.; GOEL, S.; AGARWAL, M.; JAGANNATH, A. **De Novo Transcriptome Profiling of Mustard Aphid (*Lipaphis erysimi*) and Differential Expression of Transcripts Associated with Feeding and Non-Feeding Conditions and Developmental Stages**. *Insects*, 2024, 15, 682. <https://doi.org/10.3390/insects15090682>

CONGIO, G. F. S. *Caracterização e quantificação dos danos causados por cigarrinhas adultas do gênero *Mahanarva* sp. em *Brachiaria brizantha* cv. Marandu*. 2010. Dissertação (Mestrado) – ESALQ/USP, Piracicaba, 2010.

CONGIO, G. F. S. *et al.* **Spittlebug damage on tropical grass and its impact in pasture-based beef production systems**. *Scientific Reports*, v. 10, p. 10758, 2020.

DA SILVA VAZ JUNIOR, I., LU, S., PINTO, A.F.M., DIEDRICH, J.K., YATES, J.R., III, MULENGA, A. *et al*. **Changes in saliva protein profile throughout *Rhipicephalus microplus* blood feeding.** *Parasites & Vectors*, 17, 36. 2024

DIAS, M.; AUAD, A. M.; RESENDE, T. T. **Insecticidal effects of thymol on *Mahanarva spectabilis*.** *Journal of Agricultural and Crop Research*, v. 7, p. 181–185, 2019.

DINARDO-MIRANDA, L. L. *et al*. **Resistance mechanisms of sugarcane cultivars to spittlebug *Mahanarva fimbriolata*.** *Scientia Agricola*, v. 73, p. 115–124, 2016.

FARINA, P.; BEDINI, S.; CONTI, B. **Multiple functions of Malpighian tubules in insects: a review.** *Insects*, v. 13, 2022.

GEBREKIDAN, A.G., ZHANG, Y. & CHEN, J. **A comprehensive transcriptomic and proteomics analysis of candidate secretory proteins in rose grain aphid, *Metopolophium dirhodum*.** *Current Issues in Molecular Biology*, 46, 13383–13404, 2024.

GRABHERR, M. G. *et al*. **Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data.** *Nature Biotechnology*, v. 29, n. 7, p. 644–652, 2011.

GUILBEAU, B. H. **The origin and formation of the froth in spittle-insects.** *The American Naturalist*, v. 42, p. 783–798, 1908.

HENDERSHOT, L. M.; BUCK, T. M.; BRODSKY, J. L. **Molecular chaperones and folding enzymes in ER homeostasis.** *Journal of Molecular Biology*, v. 436, n. 14, p. 168418, 2024.

HOCH, H. *et al*. **Adhesive properties of Aphrophoridae spittlebug foam.** *Journal of the Royal Society Interface*, v. 21, 2024.

HUGHES, C.S., MOGGRIDGE, S., MÜLLER, T. *et al*. **Single-pot, solid-phase-enhanced sample preparation for proteomics experiments.** *Nature Protocols* 14, 68–85, 2019. <https://doi.org/10.1038/s41596-018-0082-x>

LANGMEAD, B.; SALZBERG, S. L. **Fast gapped-read alignment with Bowtie 2.** *Nature Methods*, v. 9, n. 4, p. 357–359, 2012.

LEITE, M.V., AUAD, A.M., RESENDE, T.T., FRIAS, M.P., FONSECA, M.G. & CASTRO, R.J.C. **Do salicylic acid, nitric oxide and feeding by *Mahanarva spectabilis* nymphs induce a resistance response in elephant grass?** *Experimental Agriculture*, 50, 498–504, 2014. Available from: <https://doi.org/10.1017/S0014479713000689>

LI, B.; DEWEY, C. N. **RSEM: accurate transcript quantification from RNA-Seq data.** *BMC Bioinformatics*, v. 12, p. 323, 2011.

LI, W., WANG, X., JIANG, P., YANG, M., LI, Z., HUANG, C. *et al*. **A full-length transcriptome and gene expression analysis of three detoxification gene**

families in a predatory stink bug, *Picromerus lewisi*. *Frontiers in Physiology*, 10(13), 1016582. 2022.

MADALENO, L. L. *et al.* Influence of *Mahanarva fimbriolata* injury on cane juice quality. *Neotropical Entomology*, v. 37, p. 68–73, 2008.

MAPBIOMAS. **Nota técnica: Uso do fogo em pastagens plantadas no Brasil**. São Paulo: MapBiomass, 2025. Disponível em: https://brasil.mapbiomas.org/wp-content/uploads/sites/4/2025/01/Nota_Tecnica_Fogo-e-Pastagem_2024-1.pdf. Acesso em: 6 jul. 2026

MARSHALL, A. T. **Batelli glands of cercopoid nymphs**. *Nature*, v. 205, p. 925, 1965.

MARUCCI, R. C.; CAVICHIOLI, R. R.; ZUCCHI, R. A. **A key to five sharpshooter leafhoppers (Hemiptera: Cicadellidae: Cicadellinae) vectors of the citrus variegated chlorosis (CVC)**. *Anais da Sociedade Entomológica do Brasil*, v. 28, n. 3, p. 439–446, 1999. DOI: 10.1590/S0301-80591999000300008.

MISTRY, J. *et al.* **Pfam: the protein families database**. *Nucleic Acids Research*, v. 49, p. D412–D419, 2021.

MONTEIRO, L. P. *et al.* **Protein and phytohormone profiles of *Mahanarva spectabilis* salivary glands**. *Archives of Insect Biochemistry and Physiology*, v. 106, n. 3, e21773, 2021.

MORIYA, Y. *et al.* **KAAS: an automatic genome annotation server**. *Nucleic Acids Research*, v. 35, p. W182–W185, 2007.

NASCIMENTO, V. F.; AUAD, A. M.; RESENDE, T. T. **Olfactory response of *Mahanarva spectabilis* to plant extracts**. *Agronomy*, v. 11, p. 856, 2021.

OERKE, E. C. **Crop losses to pests**. *Journal of Agricultural Science*, v. 144, p. 31–43, 2006.

PÉREZ-SÁNCHEZ, R. *et al.* **Proteomics informed by transcriptomics in tick saliva**. *Parasites & Vectors*, v. 15, p. 1, 2022.

RAKITOV, R.A. **Structure and function of the Malpighian tubules, and related behaviors in juvenile cicadas: evidence of homology with spittlebugs (Hemiptera: Cicadoidea & Cercopoidea)**. *Zoologischer Anzeiger – A Journal of Comparative Zoology*, 241(2), 117–130, 2002. Available from: <https://doi.org/10.1078/0044-5231-00025>

RAMOS, H.J.O. **Transcriptomic and Proteomic Signatures Underlying Nymphal Adaptation and Foam Production in the Forage Pest *Mahanarva spectabilis***. 2026a. Massive MSV000101533. <https://doi.org/10.25345/C59Z90S0P>.

RAMOS, H.J.O. **Transcriptomic data from nymphs of *Mahanarva spectabilis* associated with foam production and adaptation**. NCBI BioProject,

PRJNA1404840. 2026b. Available at:
<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1404840>

RESENDE, T. *et al.* **Damage capacity of *Mahanarva spectabilis* adults on *Brachiaria ruziziensis*.** *The Scientific World Journal*, v. 2013, p. 1–5, 2013.

RINALDI, A. J. *et al.* **The adaptive nature of the foam proteome produced by *Mahanarva spectabilis*.** *Scientific Reports*, 2026a.

RINALDI, A. J.; BONJOUR, M. da S.; PINTO, I. de P. A.; FARIA, G. T. B.; LIMA, L. L.; PEREIRA, M. C.; AUAD, A. M.; PEREIRA, J. F.; OLIVEIRA, M. G. A.; RAMOS, H. J. de O. **Molecular and Structural Characterization of Foam Proteins from *Mahanarva spectabilis* (Distant, 1909) (Hemiptera: Cercopidae) Nymphs Reveals Adaptive Features and Potential Targets for Pest Control.** *Archives of Insect Biochemistry and Physiology*, v. 17, 2026b.

RIVERA-VEGA, L.J., GALBRAITH, D.A., GROZINGER, C.M. & FELTON, G.W. **Host plant driven transcriptome plasticity.** *PLoS One*, 12(8), e0182636, 2017.

SAHAYARAJ, K.; SARANYA, B.; SAYED, S.; ESTELLE, L. Y. L.; MADASAMY, K. **Biofoam of Spittlebug, *Poophilus costalis* (Walker): Preferential Sites, Temperature Regulation, Chemical Composition and Antimicrobial Activity.** *Insects*, Basel, v. 12, n. 4, p. 340, 2021. DOI: 10.3390/insects12040340.

SANTOS, R. S.; SILVA, S. A. **Registro de *Mahanarva spectabilis* em *Urochloa brizantha* cv. Piatã no Acre.** *Entomology Beginners*, 2023.

SARAIVA, N. B. *et al.* **Molecular profiling of *Mahanarva spectabilis* salivary glands.** *International Journal of Tropical Insect Science*, v. 41, n. 1, p. 849–860, 2021.

SARFRAZ, Z. *et al.* **Plant biochemistry in the era of omics: integrated omics approaches to unravel the genetic basis of plant stress tolerance.** *Plant Breeding*, 0:1–23, 2025. <https://doi.org/10.1111/pbr.13277>

SHARMA, M.; ORAON, P. K.; SRIVASTAVA, R.; CHONGTHAM, R.; GOEL, S.; AGARWAL, M. *et al.* **Comparative transcriptomics of aphids.** *Frontiers in Plant Science*, Lausanne, v. 15, p. 1415628, 2024. DOI: 10.3389/fpls.2024.1415628.

SHAO, E., SONG, Y., WANG, Y., LIAO, Y., LUO, Y., LIU, S. *et al.* **Transcriptomic and proteomic analysis of putative digestive proteases in the salivary gland and gut of *Empoasca* (*Matsumurasca*) *onukii* Matsuda.** *BMC Genomics*, 27, 271, 2017. Available from: <https://doi.org/10.1186/s12864-021-07578-2>

SHAO, E., SONG, Y., WANG, Y. *et al.* **Transcriptomic and proteomic analysis of putative digestive proteases in the salivary gland and gut of *Empoasca* (*Matsumurasca*) *onukii* Matsuda.** *BMC Genomics* 22, 271, 2021. <https://doi.org/10.1186/s12864-021-07578-2>

SOUZA, J. C. *et al.* **Cigarrinhas-das-pastagens: histórico, bioecologia, prejuízos, monitoramento e medidas de controle.** EPAMIG, 2007.

SUN, Y. L. *et al.* **Chemosensory protein 1 binding properties in *Helicoverpa armigera***. *Journal of Agricultural and Food Chemistry*, v. 73, n. 11, p. 6622–6631, 2025.

SZTERK, A. *et al.* **Chemical composition of spittlebug foam**. *Journal of Asia-Pacific Entomology*, v. 27, n. 1, p. 102185, 2024.

TEUFEL, F. *et al.* **SignalP 6.0 predicts signal peptides**. *Nature Biotechnology*, v. 40, n. 7, p. 1023–1025, 2022.

THIEL, T. *et al.* **Gene-derived SSR markers in barley**. *Theoretical and Applied Genetics*, v. 106, n. 3, p. 411–422, 2003.

TIWARI, A. K. **Insect pests in agriculture Identifying and Overcoming Challenges through IPM**. *Archives of Current Research International*, v. 24, n. 3, p. 124–130, 2024.

TONELLI, M. *et al.* **Spittlebugs produce foam as a thermoregulatory adaptation**. *Scientific Reports*, v. 8, 2018.

TONELLI, M. *et al.* **Bacterial community in the foam produced by *Mahanarva fimbriolata***. *Brazilian Journal of Microbiology*, v. 51, p. 1151–1157, 2020.

TORRES, D. A. P. **Investimento em pesquisa e produtividade agropecuária**. *Revista de Política Agrícola*, v. 32, n. 2, p. 106–122, 2023.

VALÉRIO, J. R. *et al.* **Spittlebugs: bioecology, host plant resistance and advances in IPM**. 2001.

WANG, J. *et al.* **Bioinformatics for plant and agricultural discoveries in the age of multiomics: A review and case study of maize nodal root growth under water déficit**. *Physiologia Plantarum*, 2022.

WESSING, A. *et al.* **Carbonic anhydrase supports electrolyte transport in *Drosophila***. *Journal of Insect Physiology*, v. 43, n. 1, p. 17–28, 1997.

Wu, W.; Ye, Z.; Mao, Q.; Shan, H.-W.; Li, J.-M.; Chen, J.-P. **Combined Transcriptome and Proteome Analysis of the Protein Composition of the Brochosomes of the Leafhopper *Nephotettix cincticeps***. *Insects* 2023, 14, 784. <https://doi.org/10.3390/insects14100784>

XINGTAO, Q. *et al.* **Chemosensory protein 18 in *Plutella xylostella***. *Journal of Insect Science*, v. 24, n. 1, p. 3, 2024.

YAN, S.; CUI, F.; QIAO, C. **Carboxylesterases and insecticide resistance**. *Protein & Peptide Letters*, v. 16, n. 10, p. 1181–1188, 2009.

ZHANG, H. *et al.* **Odorant-binding and chemosensory proteins in *Diaphorina citri***. *Pest Management Science*, v. 76, n. 8, p. 2734–2744, 2020.

ZHANG, H., LIN, R., LIU, Q., LU, J., QIAO, G., HUANG, X. **Transcriptomic and proteomic analyses provide insights into host adaptation of a bamboo-feeding aphid.** *Frontiers in Plant Science*, v. 13, p. 1098751, 2023.

ZHANG, Y. *et al.* **Host plant volatile responses in *Bradysia odoriphaga*.** *ACS Omega*, v. 4, n. 2, p. 3800–3811, 2019.

Apêndice

Figura S1. Taxa média de erro ao longo das leituras pareadas Illumina do RNA-seq da ninfa de *Mahanarva spectabilis*. A taxa de erro permaneceu extremamente baixa ($<0,0003$) entre as leituras 1 e 2, com apenas um ligeiro aumento em direção ao final das leituras, consistente com os perfis típicos de sequenciamento Illumina.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0001-FigureS1.docx>

Figura S2. Composição de bases ao longo das leituras pareadas Illumina do RNA-seq da ninfa de *Mahanarva spectabilis*. Após os primeiros ~20 ciclos, as proporções de nucleotídeos estabilizaram em ~30% para A e T e ~20% para G e C, indicando uma distribuição equilibrada de bases. A proporção de bases indeterminadas (N) foi insignificante em ambas as leituras.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0002-FigureS2.docx>

Figura S3. Distribuição do comprimento dos unigenes montados a partir do transcriptoma da ninfa de *Mahanarva spectabilis*. A maioria dos unigenes variou entre 300 e 1000 nucleotídeos, com 8,6% excedendo 2000 nt, representando possíveis transcritos de comprimento total.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0003-FigureS3.docx>

Figura S4: (A) Distribuição aleatória de leituras ao longo dos transcritos na biblioteca de RNA-seq da ninfa de *Mahanarva spectabilis*. (B) As leituras foram distribuídas uniformemente pelas regiões 5'–3' do gene, com apenas um pequeno enriquecimento central. (C) Distribuição do tamanho do inserto das leituras pareadas, mostrando um pico em ~400 pb. Curva de saturação gênica 4D (FPKM $\geq 0,1$) indicando que a profundidade de sequenciamento foi suficiente para recuperar quase todos os transcritos expressos.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0004-FigureS4.docx>

Tabela S1. Anotação funcional abrangente de transcritos expressos em ninfas de *Mahanarva spectabilis* com base em dados de RNA-Seq. Cada entrada corresponde a um transcrito codificador de proteína previsto, com seu respectivo ID de equivalência gênica e anotações funcionais derivadas de múltiplos bancos de dados, incluindo Pfam (domínios proteicos), Gene Ontology (função molecular e processo biológico), KEGG (vias metabólicas) e outras classificações funcionais. Este conjunto de dados fornece a estrutura de referência utilizada para identificar proteínas putativas secretadas e funcionalmente relevantes expressas na glândula de Batelli.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0005-TableS1.xls>

Tabela S2. Anotação funcional de transcritos de ninfas de *Mahanarva spectabilis* relacionados a interações planta-inseto e ambientais. A tabela lista os genes anotados agrupados por categorias biológicas principais, incluindo transdução de sinal, respostas ao estresse e à defesa, digestão, metabolismo e adaptação ambiental. A classificação funcional foi baseada em termos da Gene Ontology (GO), vias KEGG e domínios Pfam inferidos a partir de anotações de transcritos de RNA-Seq.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0006-TableS2.xlsx>

Tabela S3. Resumo da contagem de genes por categoria funcional identificada no transcriptoma da ninfa de *Mahanarva spectabilis*. As categorias representam os principais grupos funcionais inferidos a partir das anotações de RNA-Seq, incluindo desintoxicação, percepção quimiossensorial, formação de cutícula e barreira, e processos de osmorregulação e transporte. O número de genes únicos atribuídos a cada categoria é indicado.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0007-TableS3.docx>

Tabela S4. Distribuição de isoformas de transcritos por gene no transcriptoma da ninfa de *Mahanarva spectabilis*. A tabela mostra o número de genes que produzem um determinado número de isoformas de transcritos, variando de genes com uma única isoforma a loci altamente alterados por splicing alternativo, gerando mais de 40 isoformas. Essa distribuição destaca a prevalência de genes com uma ou duas isoformas e o número relativamente pequeno de loci que exibem splicing alternativo extenso.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0008-TableS4.docx>

Tabela S5. Resumo dos genes com splicing alternativo no transcriptoma da ninfa de *Mahanarva spectabilis*. A tabela lista os genes que apresentam múltiplas isoformas de transcritos (≥ 2) agrupados por categorias funcionais principais, incluindo metabolismo, transdução de sinal, resposta ao estresse e componentes estruturais. Esses dados fornecem uma visão geral da distribuição funcional dos loci com splicing alternativo identificados pela análise de RNA-Seq.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0009-TableS5.xlsx>

Tabela S6. Classificação funcional de genes com splicing alternativo associados a interações planta-ambiente no transcriptoma da ninfa de *Mahanarva spectabilis*. A tabela lista genes que apresentam duas ou mais isoformas de transcritos, agrupados por categorias funcionais relacionadas a respostas da planta, adaptação ambiental e sinalização de estresse, conforme inferido a partir da análise de splicing baseada em RNA-Seq.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0010-TableS6.xlsx>

Tabela S7. Predição de proteínas secretadas (secretoma) no transcriptoma da ninfa de *Mahanarva spectabilis*. A tabela lista transcritos codificadores de proteínas que, segundo previsões, contêm peptídeos sinalizadores ou características associadas à secreção, identificadas por meio dos programas SignalP e Phobius. Esses resultados foram derivados de dados de RNA-Seq e representam proteínas secretadas putativas, potencialmente envolvidas em funções extracelulares ou glandulares.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0011-TableS7.xlsx>

Tabela S8. Anotação funcional das proteínas secretadas previstas (secretoma) do transcriptoma da ninfa de *Mahanarva spectabilis*. A tabela lista as proteínas secretadas putativas identificadas por meio das previsões do SignalP e do Phobius, juntamente com seus termos de Gene Ontology (GO), vias KEGG e domínios Pfam associados. Esta análise funcional fornece informações sobre os potenciais papéis biológicos das proteínas secretadas expressas na glândula de Batelli.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0012-TableS8.xlsx>

Tabela S9. Anotação funcional das proteínas identificadas na glândula de Batelli de ninfas de *Mahanarva spectabilis* por análise proteômica LC-MS/MS. A tabela lista as proteínas detectadas experimentalmente juntamente com suas classificações funcionais, incluindo domínios Pfam, termos da Gene Ontology (GO), vias KEGG e categorias relacionadas ao metabolismo, resposta ao estresse e secreção. Essas anotações fornecem informações sobre a composição molecular e os potenciais papéis biológicos do secretoma da glândula de Batelli.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0013-TableS9.xls>

Tabela S10. Anotação funcional de proteínas da glândula de Batelli de ninfas de *Mahanarva spectabilis* relacionadas a interações planta-hospedeiro e ambientais. As proteínas foram identificadas por LC-MS/MS e anotadas de acordo com os termos da Gene Ontology (GO), vias KEGG e domínios Pfam. O conjunto de dados destaca proteínas potencialmente envolvidas no reconhecimento do hospedeiro, modulação da defesa, adaptação ao estresse e outros processos relevantes para a planta.

<https://resjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fimb.70050&file=imb70050-sup-0014-TableS10.xls>