ANÁLISE *IN SÍLICO* DA AMINOPEPTIDASE AAEL003012 DE *Aedes aegypti*

Autores: Eduarda Rodrigues Araujo Alves, Laís Marins Menezes, Eduardo Mendes de Oliveira, Bruno Regis Lyrio Ferraz, Tatiane Sanches Soares

Universidade Federal do Espírito Santo – UFES, Campus Alegre, Alto Universitário, S/N, Guararema – 29500-000 – Alegre - ES, Brasil, eduardarodriguesaraujoalves5@gmail.com, laismarins28@gmail.com, eduardo.oliveira.75@edu.ufes.br, bruno.ferraz@ufes.br, tatiane.s.soares@ufes.br.

Resumo
O *Aedes aegypti* é o principal vetor de arboviroses de importância epidemiológica no Brasil, tornando seu controle um importante desafio de saúde pública. Nesse contexto, a identificação de novos alvos moleculares é fundamental para o desenvolvimento de estratégias mais eficazes de controle vetorial. Este estudo realizou a caracterização *in sílica* da proteína codificada pelo gene AAEL003012 de *Aedes aegypti*, integrante da família M1 de aminopeptidases dependentes de zinco. A análise pelo InterProScan identificou domínios catalíticos característicos dessa família. O alinhamento múltiplo com sequências homólogas de *Aedes albopictus*, *Armigeres subalbatus* e *Toxorhynchites rutilus septentrionalis* revelou a conservação do motivo catalítico HExxH na posição 331, com similaridade entre 41% e 70%. A predição estrutural pelo AlphaFold apresentou 80,3% dos resíduos com alta confiabilidade, enquanto a análise no PyMOL evidenciou a disposição do domínio catalítico e do ERAP1-C. Esses resultados indicam propriedades compatíveis com atividade proteolítica, destacando o AAEL003012 como alvo relevante para estudos funcionais e controle de *Aedes aegypti*.
Palavras-chave: *Aedes aegypti*. Aminopeptidase M1. Bioinformática estrutural. Modelagem molecular. Controle vetorial.

Área do Conhecimento: Ciências Biológicas.
Subárea do conhecimento: Bioquímica.

Introdução

O *Aedes aegypti* é reconhecido como o principal vetor urbano de arboviroses de importância mundial, como dengue, Zika, chikungunya e febre amarela, configurando um dos maiores desafios atuais para a saúde pública (BHATT et al., 2013; KRAEMER et al., 2015). As abordagens tradicionais de controle vetorial vêm mostrando limitações crescentes, como a resistência a inseticidas, o que impulsiona a busca por alternativas moleculares mais direcionadas e eficazes (BENELLI & MEHLHORN, 2016). Nesse cenário, as metaloproteinases constituem uma classe de enzimas proteolíticas que utilizam o zinco (Zn^{2+}) no sítio ativo para promover a hidrólise de ligações peptídicas, atuando em processos como digestão, remodelação de tecidos, sinalização celular e resposta imune (RAWLINGS & BARRETT, 1993; NAGASE & WOESSNER, 1999). Em insetos holometábolos, as metaloproteinases participam do remodelamento e da degradação de tecidos durante a metamorfose.

Em *Aedes aegypti*, diferentes metaloproteinases apresentam expressão associada à metamorfose e à digestão do repasto sanguíneo (PAGE-MCCAW et al., 2003; KANTOR et al., 2017). Apesar de sua importância biológica, o repertório de metaloproteinases de *Aedes aegypti* permanece pouco caracterizado, com escassos membros estudados funcionalmente (WATERHOUSE et al., 2007). No conjunto das metaloproteinases ainda não caracterizadas, destaca-se o gene AAEL003012, registrado no banco de dados VectorBase e identificado em análises transcriptômicas como ativo durante diferentes fases do desenvolvimento do inseto (GIRALDO-CALDERÓN et al., 2015). Análises *in silico* indicam que a proteína codificada por esse gene integra a família M1 de aminopeptidases dependentes de zinco, com possível atuação em digestão proteica e processamento de peptídeos regulatórios (ISOE et al., 2009; RAWLINGS et al., 2014). Apesar de sua provável importância biológica, a função da

proteína codificada pelo gene AAEL003012 permanece sem validação experimental, representando uma lacuna relevante no conhecimento bioquímico desse vetor.

Nesse contexto, o presente trabalho tem como objetivo realizar a caracterização bioinformática da proteína codificada pelo gene AAEL003012 de *Aedes aegypti*, incluindo a identificação de domínios conservados, análises de similaridade interespecífica e predição estrutural tridimensional, de modo a ampliar o entendimento sobre sua função e importância biológica.

Metodologia

A sequência gênica da metaloproteinase foi coletada nos bancos VectorBase (GIRALDOCALDERÓN et al., 2015) e NCBI (2024), a partir do identificador gênico AAEL003012, sendo a sequência proteica obtida em formato FASTA para as análises estruturais e funcionais. A caracterização funcional, incluindo identificação de família proteica, domínios conservados e sítio catalítico, foi realizada por meio do InterProScan (JONES et al., 2014).

A sequência proteica foi analisada por meio do BLASTp (ALTSCHUL et al., 1997), contra o banco de dados não redundante (nr) do NCBI, com o objetivo de identificar homólogos e selecionar sequências representativas de *Aedes aegypti*, *Aedes albopictus*, *Armigeres subalbatus* e *Toxorhynchites rutilus septentrionalis*, com base em critérios de similaridade, cobertura e qualidade de anotação. As sequências selecionadas foram alinhadas por meio do BLAST, permitindo a visualização de regiões conservadas e a identificação do motivo catalítico ao longo das espécies analisadas.

O alinhamento múltiplo foi realizado no MEGA versão 12 utilizando o método ClustalW, produzindo uma matriz de distâncias par a par utilizada para o cálculo de similaridade entre as sequências, obtida pela fórmula:

Similaridade = (1 - distância) x 100%. Os valores obtidos foram expressos em porcentagem.

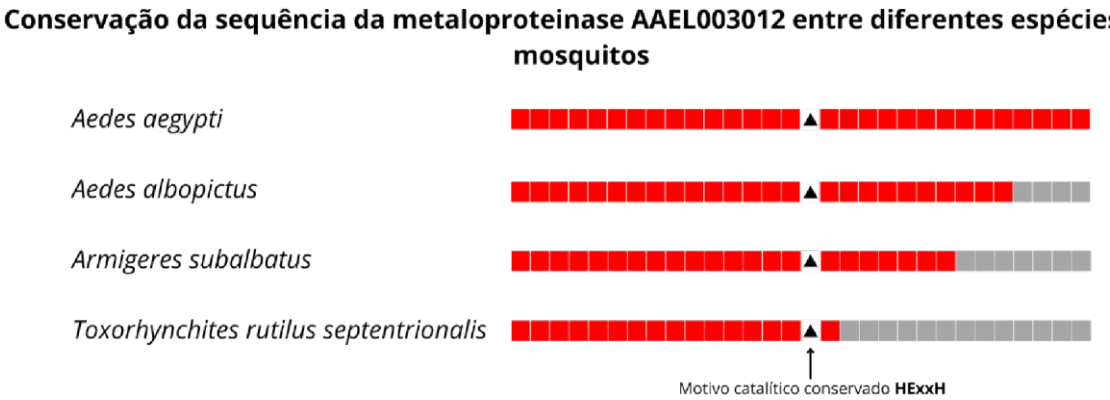
A estrutura tridimensional da proteína foi predita utilizando o AlphaFold (JUMPER et al., 2021) e analisada no PyMOL (SCHRÖDINGER, LLC, 2020), onde foram destacadas regiões estruturais relevantes, incluindo domínios funcionais e o sítio catalítico putativo.

Resultados

A análise do gene AAEL003012 por meio do InterProScan identificou domínios característicos de aminopeptidases dependentes de íons metálicos na proteína de 863 aminoácidos, incluindo Metalloproteases (zincins), catalytic domain (234–489 aa), Neutral Protease Domain 2 (229–482 aa), Tricorn interacting factor F3 domain (15–226 aa), Leukotriene A4 hydrolase N-terminal (21–224 aa) e o domínio C-terminal da ERAP1 (551–674 aa).

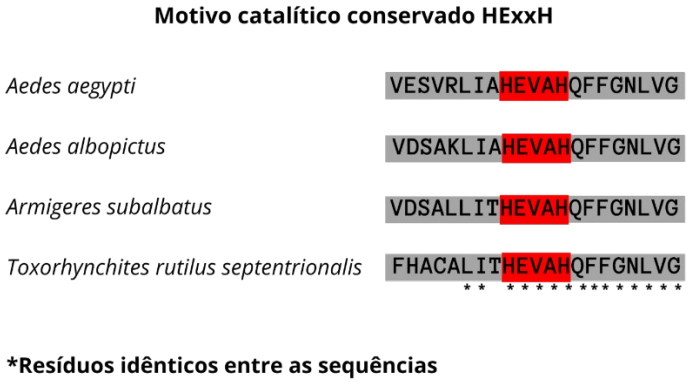
O alinhamento múltiplo da sequência proteica do AAEL003012 de *Aedes aegypti* com homólogos de *Aedes albopictus*, *Armigeres subalbatus* e *Toxorhynchites rutilus septentrionalis*, realizado no BLAST e no MEGA, revelou regiões conservadas ao longo da sequência, com destaque para o motivo catalítico HExxH, idêntico entre todas as espécies analisadas (Figuras 1 e 2).

Figura 1 - Representação esquemática da conservação da sequência da metaloproteinase AAEL003012 entre diferentes espécies de mosquitos. Em vermelho, regiões conservadas; em cinza, regiões de maior variabilidade. O triângulo indica a posição do motivo catalítico HExxH.



Fonte: Autoria própria

Figura 2 - Alinhamento do motivo catalítico conservado HExxH da metaloproteinase AAEL003012 em *Aedes aegypti*, *Aedes albopictus*, *Armigeres subalbatus* e *Toxorhynchites rutilus septentrionalis*. Asteriscos indicam resíduos idênticos entre as sequências.



Fonte: Autoria própria

Os valores de similaridade entre as sequências analisadas variaram entre aproximadamente 41% e 70%, conforme retratado na Tabela 1.

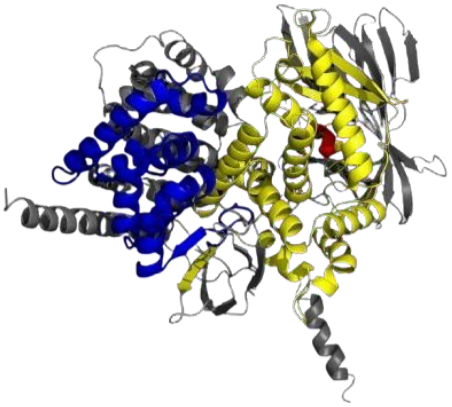
Tabela 1 — Porcentagem de similaridade da sequência proteica codificada pelo gene AAEL003012 de *Aedes aegypti* com espécies homólogas.

Espécies	Similaridade com *Aedes aegypti*
Aedes albopictus	70%
Armigeres subalbatus	60%
Toxorhynchites rutilus septentrionalis	41%

Fonte: Autoria própria

A estrutura tridimensional da proteína codificada pelo gene AAEL003012 foi predita pelo AlphaFold e visualizada no PyMOL. O modelo apresentou 80,3% dos resíduos com confiança alta ou muito alta e 19,7% com confiança baixa ou muito baixa, conforme indicado pelo índice pLDDT. A visualização revelou o domínio catalítico contendo o motivo conservado e o domínio ERAP1-C localizado na face oposta ao domínio catalítico (Figura 3).

Figura 3 - Estrutura tridimensional predita da proteína AAEL003012 visualizada no PyMOL, com destaque para o domínio catalítico em amarelo, motivo conservado em vermelho e o domínio ERAP1C em azul.



Fonte: Autoria própria

Discussão

A análise *in silico* da proteína codificada pelo AAEL003012 identificou domínios característicos de aminopeptidases dependentes de zinco, evidenciando o motivo catalítico HExxH, em que resíduos de histidina coordenam o íon Zn^{2+} e o glutamato participa como base catalítica na hidrólise de ligações peptídicas (STOCKER; BODE, 1995). Esse motivo constitui a assinatura molecular típica das metaloproteinases da família M1 e sua identificação na sequência do AAEL003012 confirma a função enzimática proposta. A presença dos demais domínios identificados pelo InterProScan, como o Neutral Protease Domain 2, o Tricorn interacting factor F3 domain e o Leukotriene A4 hydrolase N-terminal, sustenta a classificação da proteína na família M1 de aminopeptidases, visto que esses domínios são característicos e relacionados à atividade catalítica e ao reconhecimento de substratos (RAWLINGS et al., 2018). A identificação desses domínios via abordagem in silico representa um passo importante na caracterização funcional do AAEL003012, considerando que o repertório de metaloproteinases de *Aedes aegypti* permanece pouco estudado e que a maior parte dos membros dessa família ainda não teve sua função confirmada experimentalmente (WATERHOUSE et al., 2007).

A conservação do motivo catalítico HExxH entre *Aedes aegypti*, *Aedes albopictus*, *Armigeres subalbatus* e *Toxorhynchites rutilus septentrionalis* reforça a relevância funcional dessa região, indicando que pressões seletivas foram determinantes para preservar o sítio catalítico ao longo da evolução dessas espécies. Esse padrão é particularmente relevante considerando que as espécies analisadas apresentam diferentes hábitos alimentares e ciclos de vida. Diferentemente de *Aedes aegypti* e *Aedes albopictus,* que atuam como hematófagos e vetores de arboviroses, *Toxorhynchites rutilus septentrionalis* apresenta comportamento predatório sobre larvas de mosquito e não realiza repasto sanguíneo, indicando que a função do AAEL003012 pode ser relevante mesmo sem associação ao hábito hematófago. Os índices de similaridade variam aproximadamente entre 41% e 60%, com maior conservação observada entre espécies mais próximas filogeneticamente, padrão compatível com proteínas que preservam função catalítica essencial, embora atuem em diferentes contextos biológicos (RAWLINGS et al., 2014).

A predição estrutural realizada pelo AlphaFold produziu um modelo no qual 80,3% dos resíduos apresentaram níveis de confiança elevados, indicando uma estrutura tridimensional bem definida e adequada para análises funcionais. As regiões de baixa confiança, correspondentes a 19,7% dos resíduos, tendem a corresponder a regiões intrinsecamente desordenadas ou a alças expostas ao solvente, propriedades típicas de proteases que demandam flexibilidade conformacional para interagir com substratos diversos (JUMPER et al., 2021). A análise estrutural no PyMOL mostrou que o domínio ERAP1-C se encontra na face oposta ao domínio catalítico, disposição compatível com o observado em outras aminopeptidases da família M1, nas quais os domínios C-terminais regulam a especificidade de substrato e a abertura do sítio ativo (McCLUSKEY et al., 2010). Em conjunto, a análise in silico

demonstra que o AAEL003012 possui características estruturais compatíveis com atividade proteolítica, representando um candidato relevante para estudos funcionais experimentais voltados à biologia digestiva de *Aedes aegypti* e ao desenvolvimento de estratégias de controle vetorial.

Conclusão

O presente trabalho de caracterização *in silico* do gene AAEL003012 de *Aedes aegypti* indicou que a proteína codificada pertence à família das aminopeptidases M1, contendo domínios típicos vinculados à função exopeptidásica e ao metabolismo de peptídeos. A presença do motivo conservado HEVAH na posição 331, detectado em todas as espécies avaliadas, reforça a preservação do mecanismo catalítico dependente de Zn^{2+} ao longo da evolução. As análises de similaridade demonstraram conservação parcial (40–70%) entre as espécies, com maior correspondência entre aquelas mais próximas filogeneticamente, sugerindo que pressões seletivas atuaram para manter a integridade funcional da proteína. A predição estrutural tridimensional gerou um modelo de alta confiabilidade (80,3%), possibilitando a caracterização dos domínios funcionais e revelando a disposição espacial entre o domínio catalítico e o domínio ERAP1-C, localizado na face contrária. Esses resultados reforçam que o AAEL003012 representa um alvo relevante para análises funcionais experimentais direcionadas à compreensão da biologia digestiva de *Aedes aegypti* e ao delineamento de estratégias de controle vetorial.

Referências

ALTSCHUL, S. F. et al**.** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997. DOI: https://doi.org/10.1093/nar/25.17.3389.

BENELLI, G.; MEHLHORN, H. Declining malaria, rising dengue and Zika virus: insights for mosquito vector control. **Parasitology Research**, v. 115, n. 5, p. 1747–1754, 2016. DOI: 10.1007/s004360164971-z.

BHATT, S. et al. The global distribution and burden of dengue. **Nature**, v. 496, n. 7446, p. 504–507, 2013. DOI: 10.1038/nature12060.

GIRALDO-CALDERÓN, G. I. et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. **Nucleic Acids Research**, v. 43, Database issue, p. D707–D713, 2015 (publicado online em dez. 2014). DOI: 10.1093/nar/gku1117.

GIRALDO-CALDERÓN, G. I. et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. **Nucleic Acids Research**, v. 43, Database issue, p. D707–D713, 2015 (publicado online em dez. 2014). DOI: 10.1093/nar/gku1117.

ISOE, J.; ZAMORA, J.; MIESFELD, R. L. Molecular analysis of the *Aedes aegypti* carboxypeptidase gene family. **Insect Biochemistry and Molecular Biology**, v. 39, n. 1, p. 68–73, 2009. DOI: 10.1016/j.ibmb.2008.09.006.

ISOE, J.; ZAMORA, J.; MIESFELD, R. L. Molecular analysis of the *Aedes aegypti* carboxypeptidase gene family. **Insect Biochemistry and Molecular Biology**, v. 39, n. 1, p. 68–73, 2009. DOI: 10.1016/j.ibmb.2008.09.006.

JONES, P. et al**.** InterProScan 5: genome-scale protein function classification. **Bioinformatics**, 2014.

JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, p. 583–589, 2021. DOI: https://doi.org/10.1038/s41586-021-03819-2.

KANTOR, A. M. et al. Identification and initial characterization of matrix metalloproteinases in the yellow fever mosquito, *Aedes aegypti*. **Insect Molecular Biology**, v. 26, n. 1, p. 113–126, 2017. DOI: 10.1111/imb.12275**.**

KRAEMER, M. U. G. et al. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. **eLife**, v. 4, e08347, 2015. DOI: 10.7554/eLife.08347.

NAGASE, H.; WOESSNER, J. F. Matrix metalloproteinases. **Journal of Biological Chemistry**, v. 274, n. 31, p. 21491–21494, 1999.

NAGASE, H.; WOESSNER, J. F. Matrix metalloproteinases. **Journal of Biological Chemistry**, v. 274, n. 31, p. 21491–21494, 1999.

NCBI – NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **National Center for Biotechnology Information**. Bethesda: National Library of Medicine, 2024.
PAGE-MCCAW, A.; SERANO, J.; SANTÉ, J. M.; RUBIN, G. M. Drosophila matrix metalloproteinases are required for tissue remodeling, but not embryonic development. **Developmental Cell**, v. 4, n. 1, p. 95–106, 2003. DOI: 10.1016/S1534-5807(02)00400-8**.**

PAGE-MCCAW, A.; SERANO, J.; SANTÉ, J. M.; RUBIN, G. M. Drosophila matrix metalloproteinases are required for tissue remodeling, but not embryonic development. **Developmental Cell**, v. 4, n. 1, p. 95–106, 2003. DOI: 10.1016/S1534-5807(02)00400-8**.**

PYMOL. Schrödinger, LLC. *PyMOL* ***Molecular Graphics System*****.** 2020.

RAWLINGS, N. D.; BARRETT, A. J. Evolutionary families of peptidases. **Biochemical Journal**, v. 290, n. 1, p. 205–218, 1993. DOI: 10.1042/bj2900205.

RAWLINGS, N. D.; WALLER, M.; BARRETT, A. J.; BATEMAN, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. **Nucleic Acids Research**, v. 42, Database issue, p. D503–D509, 2014.

RIDDIFORD, L. M. Hormones and *Drosophila* development. In: BATE, M.; MARTINEZ-ARIAS, A. (eds.). **The Development of Drosophila melanogaster.** Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1993. p. 899–939.

STOCKER, W.; BODE, W. Structural features of a superfamily of zinc peptidases: the metzincins. ***Current Opinion in Structural Biology*****,** 1995.

VECTORBASE. **VectorBase: a genomic resource for invertebrate vectors of human pathogens**. VectorBase, 2015.

WATERHOUSE, R. M. et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. **Science**, v. 316, n. 5832, p. 1738–1743, 2007. DOI: 10.1126/science.1139862.