



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE BIOLOGIA
BACHARELADO EM CIÊNCIAS BIOLÓGICAS

ITALO GOIS XAVIER

**DIVERSIDADE DAS AQUAPORINAS NO GENOMA DE *Culex quinquefasciatus*:
UM ESTUDO *in silico***

Recife
2023

ITALO GOIS XAVIER

**DIVERSIDADE DAS AQUAPORINAS NO GENOMA DE *Culex quinquefasciatus*:
UM ESTUDO *in silico***

Monografia apresentada à Coordenação do curso de Bacharelado em Ciências Biológicas da Universidade Federal Rural de Pernambuco como parte dos requisitos para conclusão de curso.

Orientadora: Bióloga Dra. Iêda Ferreira de Oliveira

Recife
2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

X3d Xavier, Italo Gois Xavier
 DIVERSIDADE DAS AQUAPORINAS NO GENOMA DE *Culex quinquefasciatus*: UM ESTUDO in silício
 / Italo Gois Xavier Xavier. - 2023.
 58 f. : il.

 Orientadora: Ieda Ferreira de Oliveira.
 Inclui referências e apêndice(s).

 Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco,
 Bacharelado em Ciências Biológicas, Recife, 2023.

 1. Aquagliceroporinas. 2. Aquaporinas. 3. Culicidae. 4. Entomogliceroporinas. 5. Mosquitos. I. Oliveira,
 Ieda Ferreira de, orient. II. Título

CDD 574

ITALO GOIS XAVIER

**DIVERSIDADE DAS AQUAPORINAS NO GENOMA DE *Culex quinquefasciatus*:
UM ESTUDO *in silico***

Monografia apresentada à Coordenação do curso de Bacharelado em Ciências Biológicas da Universidade Federal Rural de Pernambuco como parte dos requisitos para conclusão de curso.

Aprovado em: ___ / ___ / _____

Banca Examinadora:

Biol. Dra. Iêda Ferreira de Oliveira (Orientadora e Presidente)
Universidade Federal Rural de Pernambuco

Profa. Dra. Maria de Mascena Diniz Maia (Titular)
Universidade Federal Rural de Pernambuco

Profa. Dra. Jéssica Barbara Vieira Viana (Titular)
Universidade Federal Rural de Pernambuco

MSc. Viviane Maria da Silva Santos (Suplente)
Universidade Federal Rural de Pernambuco

Ao meu pai, minha mãe e todos que sempre estiveram ao meu lado, não importa o quão árduo fosse o caminho.

RESUMO

Culex quinquefasciatus é uma espécie cosmopolita transmissora do nematoide *Wuchereria bancrofti*, agente etiológico da filariose linfática. Ao longo do seu ciclo de vida, esse mosquito necessita sobreviver a condições ambientais adversas e a diversos estressores biológicos. Nesse processo, mecanismos moleculares são essenciais para o restabelecimento da homeostase do organismo e as aquaporinas são um grupo de biomoléculas potencialmente importantes nessas respostas. As aquaporinas são uma superfamília de proteínas formadoras de canais transmembranares, altamente hidrofóbicas, que estão presentes em todos os organismos vivos, responsáveis primariamente pelo transporte facilitado de água. O objetivo deste trabalho foi identificar, classificar e caracterizar as aquaporinas presentes no genoma do mosquito *Culex quinquefasciatus* por ferramentas *in silico*. Para isso, foram feitas análises comparativas estruturais e funcionais das sequências gênicas e protéicas das aquaporinas identificadas no banco de dados NCBI, usando informações genômicas das espécies *Drosophila melanogaster*, *Anopheles gambiae* e *Aedes aegypti*. Seis prováveis genes e 15 isoformas foram identificados no genoma de *Culex quinquefasciatus*, todos localizados no cromossomo 2, e classificados nas subfamílias DRIP, BIB, PRIP, Eglp1 e Eglp3. Todas as sequências proteicas foram confirmadas para a topologia, arquitetura, localização subcelular e outros parâmetros físico-químicos característicos das aquaporinas. Os motivos e estrutura gênica das aquaporinas de *Culex quinquefasciatus* mostraram-se conservados dentro de uma mesma família, mas sutilmente diferentes entre os membros de diferentes famílias, refletindo suas diferentes funções biológicas. Os resultados provenientes destas análises poderão ajudar a conhecer melhor os aspectos moleculares dessas proteínas de transporte transmembrana, que são de grande importância para a biologia e evolução dessa espécie de importância médico-veterinária.

Palavras-chave: Aquagliceroporinas; Aquaporinas; Culicidae; Entomogliceroporinas; Mosquitos.

ABSTRACT

Culex quinquefasciatus is a cosmopolitan species that transmits the nematode *Wuchereria bancrofti*, the etiologic agent of lymphatic filariasis. Throughout its life cycle, this mosquito needs to survive adverse environmental conditions and various biological stressors. In this process, molecular mechanisms are essential for reestablishing the body's homeostasis and aquaporins are a group of potentially important biomolecules in these responses. Aquaporins are a superfamily of highly hydrophobic, transmembrane channel-forming proteins that are present in all living organisms, primarily responsible for the facilitated transport of water. The objective of this work was to identify, classify and characterize the aquaporins present in the genome of the mosquito *Culex quinquefasciatus* using *in silico* tools. For this, comparative structural and functional analysis of the gene and protein sequences of the aquaporins identified in the NCBI database were carried out, using genomic information from the species *Drosophila melanogaster*, *Anopheles gambiae* and *Aedes aegypti*. Six probable genes and 15 isoforms have been identified in the *Culex quinquefasciatus* genome, all located on chromosome 2, and classified into the DRIP, BIB, PRIP, Eglp1 and Eglp3 subfamilies. All protein sequences were confirmed for topology, architecture, subcellular localization and other physicochemical configurations characteristic of aquaporins. The motifs and gene structure of aquaporins from *Culex quinquefasciatus* were conserved within the same family, but subtly different among members of different families, reflecting their different biological functions. The results of these analyzes may help to better understand the molecular aspects of these transmembrane transport proteins, which are of great importance for the biology and evolution of this species of medical and veterinary importance.

Keywords: Aquaglyceroporins; Aquaporins; Culicidae; Entomoglyceroporins; Mosquitoes.

LISTA DE ILUSTRAÇÕES

Figura 1.	Fêmea do mosquito <i>Culex quinquefasciatus</i> antes do repasto sanguíneo.	15
Figura 2.	Distribuição geográfica do mosquito <i>Culex quinquefasciatus</i> .	16
Figura 3.	Ciclo de vida do mosquito <i>Culex quinquefasciatus</i> , mostrando as fases de ovo, larva, pupa e adulto.	17
Figura 4.	Estrutura das hélices e alças em uma aquaporina.	22
Figura 5.	Moléculas de água passando através do canal de uma aquaporina.	23
Figura 6.	Filogenia molecular das aquaporinas ortodoxas de artrópodes.	27
Figura 7.	Análise fenética refletindo os padrões de distribuição das aquaporinas no genoma de <i>Culex quinquefasciatus</i> .	31
Figura 8.	Predição dos segmentos transmembranares e topologia da sequência XP_038109078 de <i>Culex quinquefasciatus</i> .	32
Figura 9.	Motivos conservados das aquaporinas identificadas no genoma de <i>Culex quinquefasciatus</i> . Cada barra representa um motivo detectado, nas quais barras de mesma cor representam motivos em comum.	34
Figura 10.	Alinhamento múltiplo de sequências de aquaporinas identificadas no genoma de <i>Culex quinquefasciatus</i> .	36
Figura 11.	Distribuição cromossômica das aquaporinas no genoma de <i>Culex quinquefasciatus</i> . A escala está em milhões de bases. Todas estão localizadas no cromossomo dois (Chr 2).	38
Figura 12.	Estrutura e organização íntron-éxon de genes aquaporina em <i>Culex quinquefasciatus</i> .	39

LISTA DE TABELAS

- | | | |
|------------------|---|----|
| Tabela 1. | Dados da caracterização estrutural e funcional das aquaporinas de <i>Culex quinquefasciatus</i> . | 33 |
| Tabela 2. | Dados quanto ao nome do gene, ID do transcrito, ID da proteína, número de éxons, localização na fita do DNA, localização cromossômica e início e fim dos genes aquaporinas no genoma de <i>Culex quinquefasciatus</i> . | 40 |

LISTA DE ABREVIATURAS E SIGLAS

AQP	Proteína aquaporina
AQP	Gene aquaporina
BIB	<i>Drosophila</i> Big Brain
BLAST	Basic Local Alignment Search Tool
CDS	Coding Sequence (Sequência Codificante)
DNA	Ácido Desoxirribonucleico
DNAc	Ácido Desoxirribonucleico Complementar
DRIP	<i>Drosophila</i> Integral Protein
Eglp	Entomogliceroporina
ET	Elemento de transposição
GLP	Aquagliceroporina
HTM	Hélice Transmembrana
JH	Johannesburgo
JTT	Jones-Taylor-Thornton
kDa	Kilodalton
NCBI	National Center for Biotechnology Information
NPA	Motivo Asparagina-Prolina-Alanina
PRIP	Pyrocoelia Rufa Integral Protein
RNA _m	Ácido Ribonucleico Mensageiro
UTR	Untranslated Region (Região Não-Traduzida)
VESL	Vírus da encefalite de Saint Louis
VNO	Vírus do Nilo Ocidental

SUMÁRIO

1. INTRODUÇÃO	13
2. OBJETIVOS	14
2.1. Objetivo geral	14
2.2. Objetivos específicos	14
3. REFERENCIAL TEÓRICO	15
3.1. <i>Culex quinquefasciatus</i> Say, 1823	15
3.1.1. Distribuição geográfica	16
3.1.2. Biologia	16
3.1.3. Importância médico-veterinária	19
3.1.4. Genômica	19
3.2. Aquaporinas	21
3.2.1. Descoberta da primeira aquaporina	21
3.2.2. Estrutura proteica e função	22
3.2.3. Diversidade molecular das aquaporinas	24
3.2.4. Aquaporinas em insetos	25
4. METODOLOGIA	28

4.1. Identificação genômica de aquaporinas em <i>Culex quinquefasciatus</i>	28
4.2. Identificação de domínios e motivos NPA	28
4.3. Caracterização das aquaporinas	28
4.4. Classificação	29
4.5. Localização cromossômica e estrutura gênica	29
5. RESULTADOS	30
5.1. Identificação e classificação das aquaporinas em <i>Culex quinquefasciatus</i>	30
5.2. Caracterização molecular das aquaporinas de <i>Culex quinquefasciatus</i>	32
5.3. Estrutura gênica e localização cromossômica das aquaporinas de <i>Culex quinquefasciatus</i>	37
6. DISCUSSÃO	41
7. CONCLUSÕES	44
8. REFERÊNCIAS BIBLIOGRÁFICAS	45
9. APÊNDICES	52

1. INTRODUÇÃO

Culex quinquefasciatus Say, 1823 é um membro da ordem diptera, família Culicidae. São popularmente conhecidos como muriçocas, pernilongos e carapanãs. É uma espécie cosmopolita e associada a aglomerados humanos urbanos e rurais. É uma espécie de grande importância para a saúde pública no Brasil, pois está envolvida na transmissão do nematoide *Wuchereria bancrofti*, agente etiológico da filariose linfática (LOVEZOI, 2011).

Em 2010, a publicação do sequenciamento genômico de *Culex quinquefasciatus* trouxe à comunidade científica novas informações sobre a genética desta importante espécie vetora (ARENSBURGER *et al.*, 2010). *Culex quinquefasciatus* possui características únicas, o que justifica a importância de conhecer melhor seus aspectos genéticos e moleculares, pois refletem diretamente na sua fisiologia, interação com o ambiente, relação hospedeiro-parasita e resistência a inseticidas (REDDY; LABBÉ; CORBEL, 2012).

As aquaporinas são uma superfamília de proteínas de canais transmembranares, altamente hidrofóbicas, que estão presentes em todos os organismos vivos (JUNG *et al.*, 1994). São moléculas de extrema importância para o sistema de transporte de água e alguns solutos.

Existe atualmente uma grande necessidade de desenvolvimento de inseticidas novos e eficazes para combater insetos vetores, uma vez que os inseticidas em uso atualmente são baseados apenas em um número limitado de compostos ativos. Devido à sua importância vital nas larvas dos insetos e na homeostase dos adultos, as aquaporinas podem se tornar alvos para o desenvolvimento de novos inseticidas. A análise mais aprofundada dessas moléculas e da sua regulação possuem o potencial de contribuir para o desenvolvimento futuro de novas estratégias anti-vetores (REDDY; LABBÉ; CORBEL, 2012).

Desse modo, o presente trabalho propõe identificar, classificar e caracterizar as aquaporinas presentes no genoma do mosquito *Culex quinquefasciatus* por ferramentas *in silico*. Os resultados provenientes dessas análises ajudarão a conhecer melhor os aspectos moleculares dessas proteínas de transporte transmembrana que são importantes para a biologia e evolução dessa espécie de importância médico-veterinária.

2. OBJETIVOS

2.1. Objetivo geral

Analisar a diversidade estrutural e funcional dos genes que codificam a família das aquaporinas no mosquito *Culex quinquefasciatus* por ferramentas *in silico*.

2.2. Objetivos específicos

- Identificar e classificar prováveis membros da superfamília das aquaporinas no genoma de *Culex quinquefasciatus*.
- Verificar a presença de domínios conservados nas sequências proteicas das aquaporinas candidatas.
- Caracterizar as aquaporinas quanto ao peso molecular, ponto isoelétrico e localização subcelular.
- Comparar os motivos conservados entre os membros de cada família de proteínas.
- Demonstrar a estrutura gênica das aquaporinas de *Culex quinquefasciatus*.
- Mapear os genes que codificam aquaporinas nos cromossomos de *Culex quinquefasciatus*.

3. REFERENCIAL TEÓRICO

3.1 *Culex quinquefasciatus* Say, 1823

Culex quinquefasciatus (Figura 1) é um membro da ordem diptera, família Culicidae. Esta família está dividida nas subfamílias Anophelinae, Culicinae e Toxorhynchitinae. *Anopheles* é o gênero mais importante da subfamília Anophelinae, enquanto *Aedes* e *Culex* são os gêneros de maior importância da subfamília Culicinae (CLEMENTS, 1992). São popularmente conhecidos como muriçocas, pernilongos e carapanãs. São vetores de diversos agentes patogênicos e causam grande incômodo à população nos locais de sua ocorrência (PERUZIN, 2009).

Figura 1. Fêmea do mosquito *Culex quinquefasciatus* antes do repasto sanguíneo.



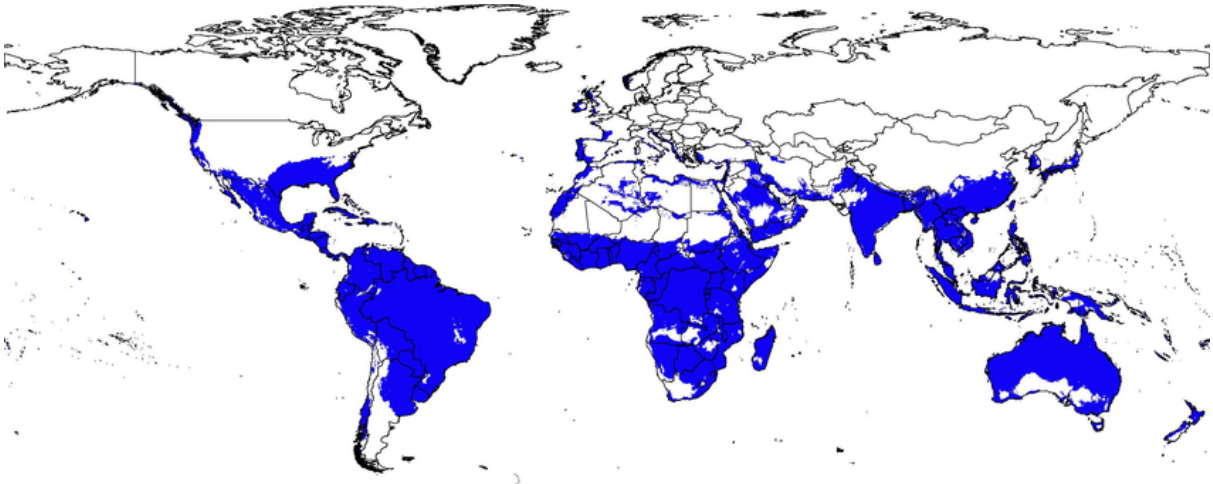
Fonte: Centers for Disease Control and Prevention (2023). Disponível em: https://www.cdc.gov/mosquitoes/gallery/culex/images/LVV7_Culex_quinquefasciatus_Adult_Resting_2022_005.jpg?_=06316.

3.1.1. Distribuição geográfica

Culex quinquefasciatus é uma espécie cosmopolita. Encontra-se distribuído em regiões tropicais e subtropicais, desde o nível do mar até altitudes moderadas (HARBACH, 2012). Amplamente adaptado ao ambiente urbano, é facilmente encontrado em habitações humanas e animais (REIS *et al.* 2023).

Ocorre em parte da Ásia, África, Austrália e nas Américas, desde o sul dos Estados Unidos à Argentina (Figura 2). No Brasil, possui ampla ocorrência na maioria dos estados (LOZOVEI, 2011).

Figura 2. Distribuição geográfica do mosquito *Culex quinquefasciatus*.



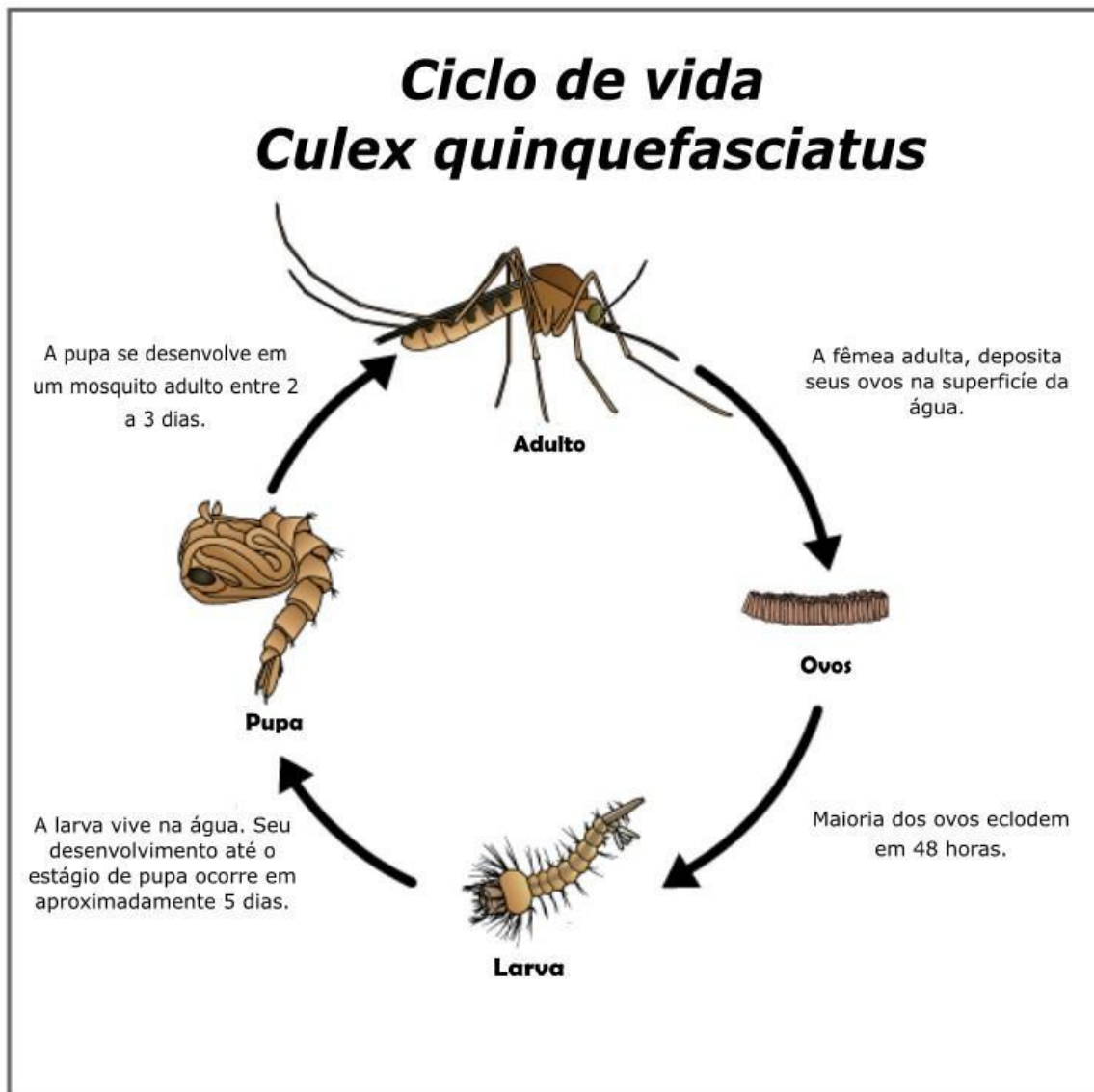
Fonte: SAMY *et al.* (2016).

3.1.2. Biologia

Culex quinquefasciatus é um mosquito que apresenta ciclo de vida holometábolo, ou seja, sofre metamorfose completa. Durante seu desenvolvimento, passa pelos estágios de ovo, larva (com os estádios L1, L2, L3 e L4), pupa e adulto (Figura 3) (BORROR; DeLONG, 1988).

Quando fecundadas, as fêmeas depositam, na superfície da água, seus ovos unidos entre si. Estes formam uma estrutura flutuante, semelhante a uma jangada, que pode conter aproximadamente 90 ovos. Eles ainda apresentam em sua extremidade superior feromônios, que atraem outras fêmeas para o sítio de oviposição (BARBOSA *et al.* 2007).

Figura 3. Ciclo de vida do mosquito *Culex quinquefasciatus*, mostrando as fases de ovo, larva, pupa e adulto.



Fonte: Centers for Disease Control and Prevention (2023). Disponível em: https://www.cdc.gov/mosquitoes/es/about/life-cycles/images/Culex_mosquito_lifecycle_web_ES.jpg?_=22890 (modificado).

As larvas de *Culex quinquefasciatus* possuem o corpo dividido em: cabeça, tórax e abdômen. Na cabeça são encontrados os principais órgãos sensoriais e peças bucais. No abdômen, observam-se oito segmentos com numerosas cerdas. Na extremidade do abdômen, apresentam um sifão respiratório, por onde se abrem os espiráculos para respiração aérea. Estas formas imaturas alimentam-se por filtração

dos detritos orgânicos e microrganismos do ambiente onde vivem (TIMMERMANN; BRIEGEL, 1999). Podem ser encontradas em qualquer tipo de hábitat que contenha água que varie de doce e clara a salobra, turva e poluída com matéria orgânica em decomposição, tais como: lixo e dejetos humanos acumulados em piscinas subterrâneas, valas, ralos, esgotos, depósitos de lixo, latrinas, fossas sépticas, tanques e em vários tipos de recipientes artificiais (SIRIVANAKARN, 1976).

Por sua vez, as pupas não se alimentam. Elas possuem a forma de vírgula, com o corpo dividido em cefalotórax e abdômen. Em seu cefalotórax, existem duas estruturas tubulares, chamadas de trombas ou trombetas respiratórias, por onde se abrem seus únicos espiráculos. Estas permitem a respiração do ar atmosférico, enquanto estão mergulhadas na água (CONSOLI; OLIVEIRA, 1994).

As formas adultas, machos e fêmeas, possuem tegumento marrom claro ou escuro, recoberto por escamas amarelo-douradas, estreitas, alongadas e curvas (LOVEZOI, 2011). Os adultos possuem aproximadamente meio centímetro de comprimento, e os machos costumam ser um pouco menores do que as fêmeas. O corpo dos adultos é delgado, com pernas longas, asas estreitas e alongadas. Os adultos apresentam dimorfismo sexual: os machos apresentam antenas plumosas e as fêmeas possuem antenas pilosas. Além disso, as fêmeas são dotadas de palpos curtos, enquanto os machos possuem palpos longos (MELO, 2020).

Machos e fêmeas alimentam-se de seiva de plantas ou outras fontes ricas em carboidratos, os quais fornecem energia para o vôo e realização de outras atividades. Porém, apenas as fêmeas realizam a hematofagia. Tal comportamento é necessário para o desenvolvimento de seus ovários e promoção da vitelogênese, já que o sangue é rica fonte de aminoácidos e outros nutrientes essenciais (ROSEN *et al.*, 1983).

As peças bucais das fêmeas são do tipo picador-sugador, com apêndices bucais estiliformes e alongados. São formadas pelo labro, mandíbulas, maxilas e hipofaringe, que são protegidas pelos lábios. Na extremidade distal, possuem dois lobos chamados labelos. O canal alimentar é formado pelo labro e pelo canal salivar. Suas maxilas e mandíbulas possuem borda serrilhada, com função cortante, auxiliando a penetração das peças bucais na derme dos hospedeiros vertebrados durante o repasto sanguíneo. Após a travessia do epitélio, as peças bucais alcançam o capilar sanguíneo, onde o inseto retira o sangue diretamente da luz do vaso em um processo denominado solenofagia. Nos machos, a probóscide e a mandíbula são mais curtas do que nas fêmeas (MELO, 2020). Eles não possuem a capacidade de perfurar

tecidos, nem animais nem vegetais. Por essa razão, não conseguem realizar hematofagia, alimentando-se apenas de líquidos açucarados de plantas (FORATTINI, 2002).

Os adultos são abundantes durante todo o ano. São comumente encontrados descansando durante o dia em ambientes escuros, como no interior de guarda-roupas e banheiros, debaixo de mesas ou em locais semelhantes dentro das casas. As fêmeas picam em ambientes fechados ou ao ar livre. Alimentam-se principalmente de sangue humano desde o pôr do sol até o amanhecer. Elas também atacam frequentemente aves e, em menor grau, outros animais domésticos, incluindo cães, gatos e porcos. A gama de hospedeiros para a refeição de sangue é ampla, mas diferentes populações variam a esse comportamento (BECKER, 2020).

3.1.3. Importância médico-veterinária

Culex quinquefasciatus é uma espécie de relevância para a saúde pública no Brasil, pois está envolvida na transmissão do nematoide *Wuchereria bancrofti*, agente etiológico da filariose linfática. Também é importante vetor de *Brugia malayi* (filariose malaia), *Dirofilaria immitis* (dirofilariose canina) e alguns plasmódios que parasitam aves (LOZOVEI, 2011).

É um vetor do vírus da encefalite de Saint Louis (VESL) e do vírus do Nilo Ocidental (VNO) (SIRIVANAKARN, 1976). Já foi relatado como vetor dos arbovírus causadores da chikungunya e encefalite japonesa. Um estudo realizado por GUEDES *et al.* (2017) sugere que *Culex quinquefasciatus* também pode estar envolvido na transmissão do Zika Vírus no Nordeste do Brasil. Todavia, não há evidências conclusivas de que desempenhe papel na transmissão natural dessas doenças (RIOS *et al.*, 2022).

3.1.4. Genômica

Em 2010, a publicação do sequenciamento genômico de *Culex quinquefasciatus* trouxe à comunidade científica novas informações sobre a genética de uma importante espécie vetora de doenças (ARENSBURGER *et al.*, 2010). Até então, apenas os genomas de *Anopheles gambiae* (HOLT *et al.*, 2002) e de *Aedes aegypti* (NENE *et al.*, 2007) estavam disponíveis.

O genoma de referência de *Culex quinquefasciatus* está depositado na base de dados do National Center for Biotechnology Information (NCBI), sob o número de acesso GCF_015732765.1. As amostras utilizadas estão descritas como *Culex quinquefasciatus* var. Johannesburgo (JHB), provenientes da África do Sul. Uma fração significativa (29%) do genoma de *Culex quinquefasciatus* apresentou elementos de transposição (ETs). Tais valores foram um pouco menores que em *Aedes aegypti* (42%-47%), mas maiores que em *Anopheles gambiae* (11%-16%). No total, 9% dos genes (1.768 genes) foram mapeados nos seus três cromossomos. Destes, 803 mostraram-se ortólogos a *Anopheles gambiae* e 641 em *Drosophila melanogaster*.

O repertório gênico de *Culex quinquefasciatus*, com 18.883 genes codificadores de proteínas, apresentou-se 22% maior que o de *Aedes aegypti* (15.419) e 52% maior que o de *Anopheles gambiae* (12.457). O aumento relativo no número de genes em *Culex quinquefasciatus* é explicado em parte pela presença de famílias gênicas significativamente expandidas, incluindo receptores olfativos e gustativos, genes relacionados ao sistema imunológico, bem como genes com possíveis funções de desintoxicação xenobiótica (ARENSBURGER *et al.*, 2010).

A curadoria especializada de famílias de genes revelou expansão no grupo das glutathione S-transferases (GSTs) citosólicas e na superfamília da citocromo P450 (CYP450). Especula-se que o grande repertório de CYP450 pode refletir adaptação a habitats larvais poluídos. Alternativamente, essa característica pode ter contribuído para a evasão de *Culex quinquefasciatus* quanto aos programas de controle vetorial baseados no uso de inseticidas químicos, visto que CYP450s estão associadas à resistência a inseticidas. A consequente diversidade em muitos genes diferentes pode ser um importante fator que levou à ampla distribuição geográfica de *Culex quinquefasciatus* (ARENSBURGER *et al.*, 2010).

Portanto, a genômica comparativa entre *Culex quinquefasciatus* e outras espécies de mosquitos mostrou que este vetor possui características únicas que justificam a importância de conhecer melhor seus aspectos genéticos e moleculares, pois refletem diretamente na sua biologia, interação hospedeiro-parasita, resistência a inseticidas e aspectos evolucionários. Dessa forma, *Culex quinquefasciatus* requer perguntas específicas e, talvez, as respostas a essas questões tragam luz à solução de problemas referentes a outras espécies (REDDY; LABBÉ; CORBEL, 2012).

3.2. Aquaporinas

3.2.1. Descoberta da primeira aquaporina

Em 1988, Agre e sua equipe de pesquisadores, durante estudos com o fator Rh presente nas membranas de eritrócitos humanos, encontraram uma banda contaminante em um dos géis analisados, com aproximadamente 28 kiloDaltons (kDa). Logo, esta proteína despertou grande interesse da comunidade científica (DENKER *et al.*, 1988).

Mais tarde, Brad Denker e Barbara Smith isolaram a proteína de 28 kDa por SDS-PAGE, uma técnica simples baseada na solubilidade em detergente. Porém, somente após coloração com prata foi possível revelar uma banda discreta de 28 kDa a partir dos extratos insolúveis em detergente, já que não foi possível pelo uso dos corantes convencionais de proteínas, como o azul de Coomassie. A partir daí, a proteína foi purificada em grandes quantidades a partir de membranas de eritrócitos humanos e melhor estudada (DENKER *et al.*, 1988; SMITH; AGRE, 1991).

Trabalhos posteriores mostraram que o isolado possuía características que sugeriam se tratar de uma proteína tetramérica que atravessava a membrana plasmática. Assim, surgiu a hipótese de que seria uma molécula de canal. Por este motivo, foi nomeada inicialmente como CHIP28 (channel-forming integral protein). Utilizando anticorpos específicos, foram analisados outros tecidos e foi observado que a proteína também era abundante no rim humano (PRESTON *et al.*, 1992). Estas informações aumentaram a suspeita de que a molécula de 28 kDa seria uma proteína envolvida no transporte de água.

Então, Peter Agre em colaboração com Bill Guggino, em 1988, testaram a possível função de transporte de água pela proteína, utilizando ovócitos da rã *Xenopus laevis*. Este é um sistema particularmente útil, visto que os oócitos de rã possuem baixa permeabilidade à água e resistem mesmo em tanques de água doce. Para isso, oócitos controle foram injetados apenas com água e oócitos teste foram injetados com DNA complementar (DNAc) codificante para a proteína. Após três dias de síntese proteica, os oócitos pareciam essencialmente idênticos. Em seguida, eles foram submetidos a estresse, sendo colocados em água destilada, e uma diferença surpreendente foi imediatamente observada: possuindo permeabilidade à água extremamente baixa, os oócitos do grupo controle não

alteraram seu tamanho; porém, os oócitos teste, altamente permeáveis à água, sofreram lise celular (PRESTON *et al.*, 1992). Conseqüentemente, a proteína foi nomeada “aquaporina” ou “AQP1”, a primeira proteína de canal de água funcionalmente definida (AGRE *et al.*, 1993).

3.2.2. Estrutura proteica e função

As aquaporinas são proteínas transmembranares altamente hidrofóbicas (Figura 4). Elas consistem em seis hélices (HTM1-6) que atravessam as membranas biológicas e estão separadas por cinco loops (A-E). Os loops B e E são altamente simétricos e contêm os motivos característicos Asparagina-Prolina-Alanina (NPA). Ambos os terminais amino (N) e carboxil (C) da proteína são intracelulares.

Figura 4. Estrutura das hélices e alças em uma aquaporina.

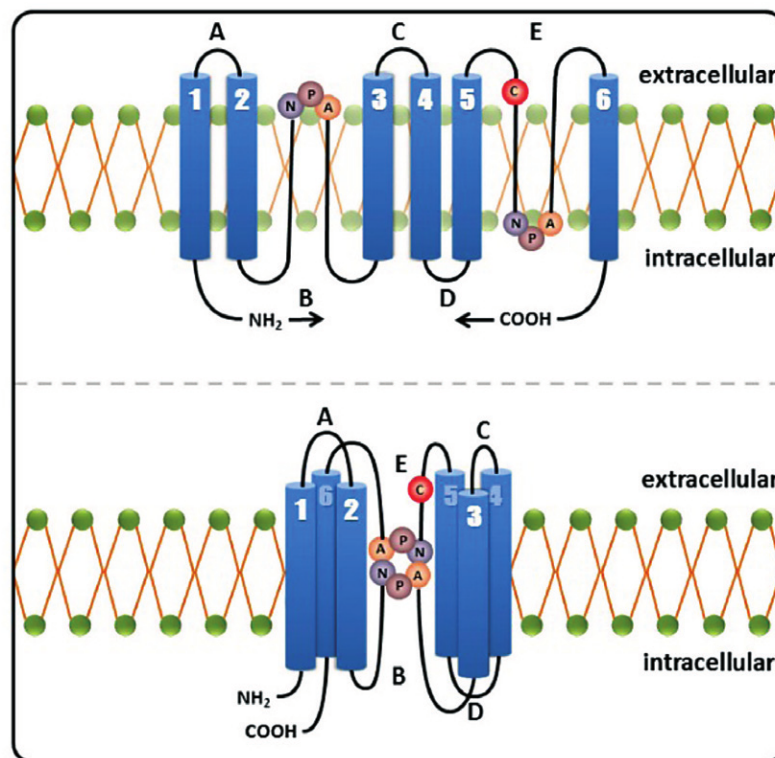


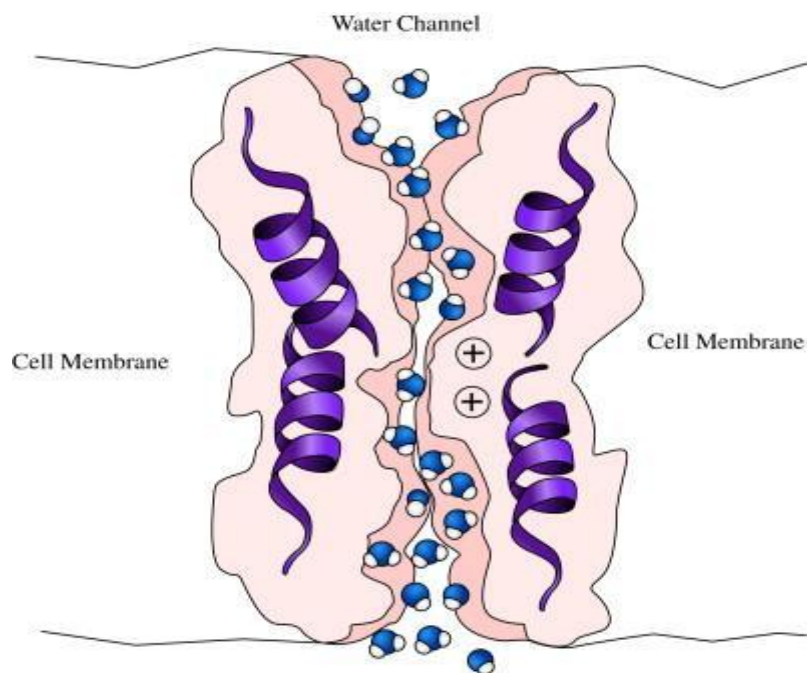
Imagem superior: Diagrama mostrando a arquitetura e topologia da aquaporina.

Imagem inferior: Conformação funcional da aquaporina segundo o modelo ampulheta.

Fonte: GEADKAEW (2016).

Sua estrutura tridimensional é denominada de modelo “ampulheta” (Figura 5) (JUNG *et al.*, 1994). As alças B e E dobram-se de lados opostos na região do poro, alinhando suas respectivas regiões NPA. As prolínas interagem de modo que as duas asparaginas carregadas positivamente formam a parede do poro. Esta região é uma das duas principais constrições e que conferem às aquaporinas sua impermeabilidade iônica. Comumente, os prótons são três vezes mais móveis em solução aquosa do que as próprias moléculas de água. Portanto, qualquer poro aquoso deveria ser um canal altamente condutor de prótons. Mas este não é o caso das aquaporinas (CHAKRABARTI *et al.*, 2004; FU; LU, 2007; TAJKHORSHID *et al.*, 2002).

Figura 5. Moléculas de água passando através do canal de uma aquaporina. É possível observar a redução do tamanho do poro na região central, o que confere a aparência de uma ampulheta.



Fonte: STILLWELL (2013).

A segunda maior constrição e o filtro primário de seletividade estão oito angstroms (8 Å) mais próximos da superfície extracelular do que o filtro de prótons. Esta é a região aromática/arginina (ar/R), que é formada por um resíduo da HTM2 e HTM5 e dois resíduos da alça E. Esses resíduos são conservados tanto nas

aquaporinas quanto nas aquagliceroporinas, e permitem a passagem de água e de solutos de baixo peso molecular, como o glicerol. Este filtro tem aproximadamente 3,0 Å de diâmetro ou um pouco menos (até 2,8 Å) em aquaporinas, um pouco mais em aquagliceroporinas (SPRING *et al.*, 2008). As interações hidrofóbicas são essenciais para a seletividade (CHEN, H. *et al.*, 2006).

A maioria das aquaporinas são inibidas reversivelmente por mercúrio (Hg^{2+}), sendo a inibição reduzida pelo agente redutor β -mercaptoetanol. Os íons mercúrio interagem principalmente com o grupamento sulfidril do resíduo de cisteína localizado próximo ao motivo NPA no loop E e, secundariamente, com a alanina do loop B. Isso bloqueia efetivamente a entrada de água no poro. Essas informações são essenciais para estudos funcionais (AGRE *et al.* 1993).

3.2.3. Diversidade molecular das aquaporinas

As aquaporinas são moléculas de importância vital para o sistema de transporte celular dos organismos vivos. Já foram identificadas em todos os domínios e reinos biológicos, a exemplo de Archaea, Eubacteria, Fungi, Plantae e Animalia.

Em geral, são codificadas por famílias gênicas e o número de membros pode variar de acordo com a espécie analisada. *Homo sapiens* codifica 13 subfamílias de aquaporinas AQP0 – AQP12 (AZAD *et al.*, 2021). *Arabidopsis thaliana* codifica 35 AQPs, que se agrupam em quatro subfamílias: proteínas intrínsecas da membrana plasmática (PIPs), proteínas intrínsecas do tonoplasto (TIPs), pequenas proteínas intrínsecas básicas (SIPs) e proteínas intrínsecas semelhantes à nodulina 26 (Nod26) (REINHARDT *et al.*, 2016).

Quanto à organização gênica, o número e a extensão de éxons e íntrons pode variar. Podem sofrer splicing alternativo, levando à formação de transcritos alternativos. Diferentes isoformas podem estar presentes e, comumente, a proteína apresenta um tamanho entre 250 a 290 aminoácidos em média (BENGA, 2009).

Quanto à classificação funcional das aquaporinas, os membros com permeabilidade exclusiva à água são denominados aquaporinas (AQPs), enquanto os membros permeáveis à água, glicerol e outros solutos são referidos como aquagliceroporinas (GLPs) (KRUZE *et al.*, 2006).

3.2.4. Aquaporinas em insetos

Os insetos são o grupo de organismos mais bem-sucedido na história da vida, com estimativas atuais representando mais da metade da biodiversidade eucariótica do mundo (FOOTITT; ADLER, 2009). Dentre suas estratégias de vida, muitas espécies precisam sobreviver a condições ambientais adversas (exposição a xenobióticos, frio, calor, umidade, seca, por exemplo) ou até mesmo a fatores estressores biológicos, como infecções por patógenos e parasitas. Nesse processo, mecanismos moleculares são essenciais para o restabelecimento da homeostase do organismo e as aquaporinas são um grupo de biomoléculas potencialmente importantes nessas respostas.

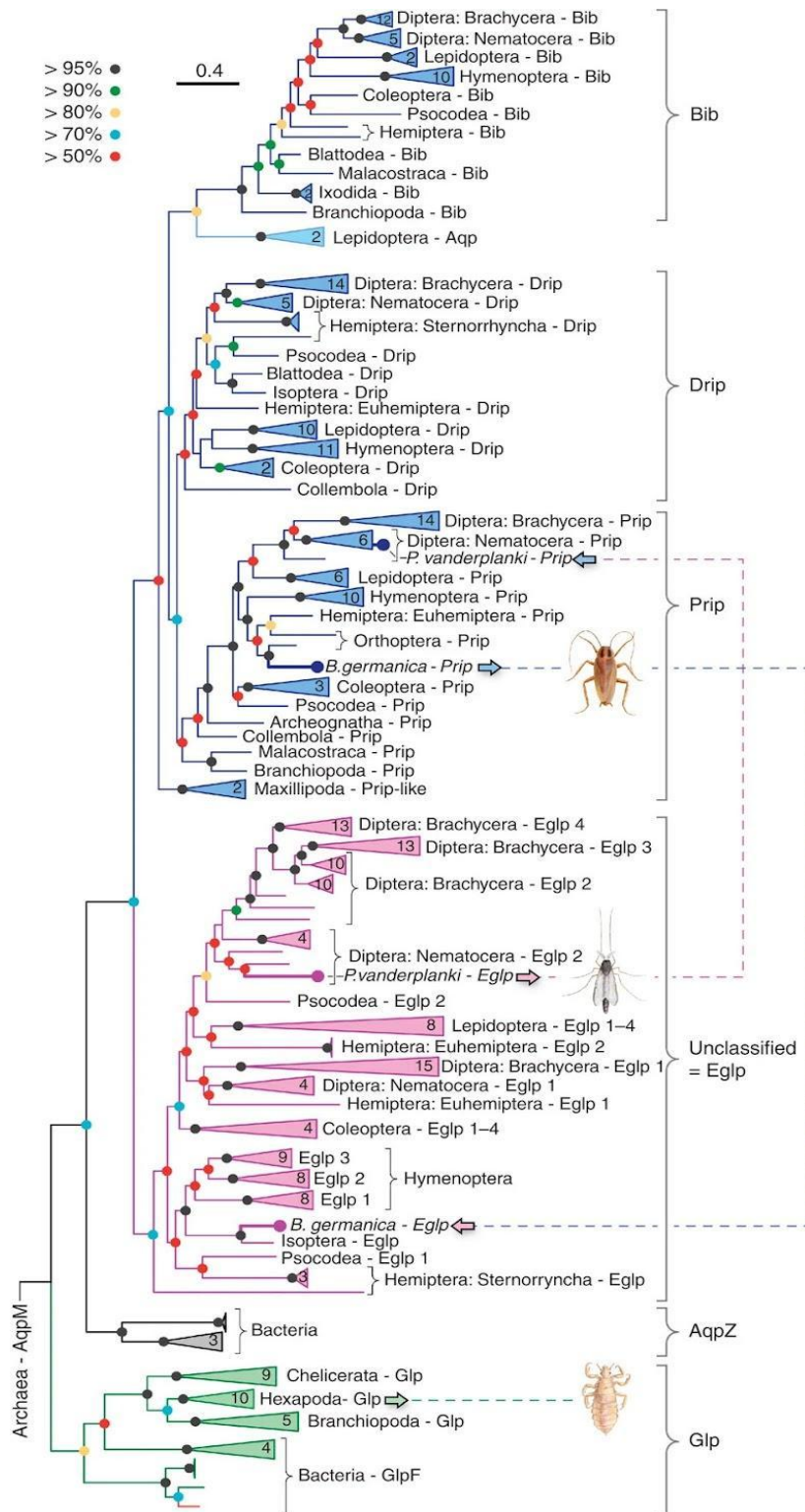
Um estudo com *Aedes aegypti* demonstrou que cada uma das seis aquaporinas identificadas apresentaram padrão específico de expressão gênica em distintas partes do corpo ou órgãos. AaAQP1 homóloga da *Drosophila* DRIP (*Drosophila* Integral Protein), é um canal específico para a água altamente seletivo na mosca das frutas. O gene *AaAQP1* foi expresso em todos os órgãos e partes do corpo, com níveis de expressão mais elevados nos túbulos de Malpighi. *AaAQP2* mostrou altos níveis de expressão em todos os órgãos e partes do corpo, exceto no tecido adiposo. Destaque para os túbulos de Malpighi, intestino médio e ovários. *AaAQP3* é homóloga da BIB de *Drosophila*. BIB não funciona como canal de água na mosca das frutas, mas está envolvida na regulação da adesão celular. *AaAQP3* foi fracamente expresso nos túbulos de Malpighi, intestino médio e ovários. *AaAQP4* foi altamente expresso em túbulos de Malpighi e mostrou-se significativamente aumentado três horas após uma refeição de sangue. Foi regulado negativamente no intestino médio durante o início da vitelogênese. *AaAQP5* foi expresso em todos os órgãos, exceto nos ovários. Apresentou expressão aumentada na maioria dos órgãos após repasto sanguíneo. Por fim, *AaAQP6* foi predominantemente expresso no tórax. Os níveis mais altos de expressão de RNAm foram encontrados no intestino anterior (DRAKE *et al.*, 2010).

Nos domínios Prokaryota e Eukaryota, é conhecido o transporte transmembrana de glicerol facilitado por aquagliceroporinas (TONG *et al.*, 2019). No entanto, as primeiras informações genômicas disponíveis sugeriram a perda desses genes em insetos holometábolos durante o curso evolutivo (ABASCAL *et al.*, 2014; DRAKE *et al.*, 2015). Por outro lado, outros grupos de cientistas deparavam-se com

uma contradição, visto com as evidências de seus estudos experimentais demonstravam que algumas aquaporinas de diferentes espécies de insetos também podiam transportar glicerol (DRAKE *et al.*, 2015; KATAOKA; MIYAKE; AZUMA, 2009).

Finn *et al.* (2015) investigaram mais profundamente as relações filogenéticas de aquaporinas de insetos e transportadores de glicerol a partir de repertórios genômicos e também realizaram estudos de mutagênese direcionada para identificar experimentalmente a origem e evolução funcional desses transportadores. Seus resultados revelaram que, nas linhagens basais de insetos hexápodos, mutações sítio-específicas nos canais de água sofreram conversão para a função de transporte de glicerol (neofuncionalização). Estas novas formas, denominadas de entomogliceropoinas (EgIp), por serem mais eficientes no transporte de glicerol foram positivamente selecionadas nos insetos holometábolos, resultando na substituição das aquagliceroporinas ortodoxas como principais veículos de condução de polióis. Na subordem Brachycera, a qual a *Drosophila melanogaster* é um membro, ocorrem quatro entomogliceroporinas (EgIp1-4). A subordem Nematocera, a qual pertencem os mosquitos, apresentam apenas duas entomogliceropoinas (EgIp1 e EgIp2) (Figura 6).

Figura 6. Filogenia molecular das aquaporinas ortodoxas de artrópodes.



Fonte: FINN et al. (2015).

4. METODOLOGIA

4.1. Identificação genômica de aquaporinas em *Culex quinquefasciatus*

Sequências genômicas de *Culex quinquefasciatus* foram retiradas do banco de dados públicos do National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>). Sequências de aquaporinas devidamente anotadas e curadas no banco de dados UniProt (<https://www.uniprot.org/>), que foram utilizadas para a busca por sondas adicionais de aquaporinas de dípteros correspondentes às espécies *Drosophila melanogaster*, *Anopheles gambiae* e *Aedes aegypti*. Foram usados como *query* para executar pesquisas via BLASTp contra o proteoma de *Culex quinquefasciatus* no programa BioEdit (version 7.0.9.0). Um *e-value* igual ou inferior a e^{-05} foi aplicado para identificar correspondências significativas.

4.2. Identificação de domínios e motivos NPA

A ferramenta Batch Web CD-Search Tool do NCBI (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) foi empregada para examinar a presença de domínios conservados nas proteínas candidatas identificadas. Os motivos foram visualizados através da ferramenta MEME Suite 5.5.4 (BAILEY *et al.*, 2015; <http://memesuite.org/tools/meme>).

4.3. Caracterização das aquaporinas

Os parâmetros físico-químicos das aquaporinas candidatas como o ponto isoelétrico e o peso molecular foram determinados usando o programa JVirGel 2.0 (HILLER *et al.*, 2006). A localização subcelular foi determinada pelo programa Cello 2.5 (YU *et al.*, 2006; <http://cello.life.nctu.edu.tw/>). Por sua vez, a quantidade de hélices transmembranas foram previstas através da ferramenta DeepTMHMM (HALLGREN *et al.*, 2022; <https://dtu.biolib.com/DeepTMHMM>).

4.4. Classificação

A análise fenética das sequências candidatas foi realizada de modo a observar seu agrupamento baseado na similaridade. Inicialmente, um alinhamento múltiplo de sequências foi realizado com todas as sequências candidatas usando a ferramenta ClustalW (THOMPSON *et al.*, 1994) presente no programa MEGA 11 (KUMAR *et al.*, 2016).

Para gerar a árvore filogenética pelo método de Neighbor-Joining, foi utilizado o modelo de substituição de Jones-Taylor-Thornton (JTT) com *bootstrap* de 1000 repetições. Foi realizada utilizando as sequências ortólogas dos dípteros *Aedes aegypti*, *Anopheles gambiae* e *Drosophila melanogaster*.

4.5. Localização cromossômica e estrutura gênica

Para determinar a localização cromossômica dos genes que codificam aquaporinas em *Culex quinquefasciatus* e a visualização da sua estrutura gênica, foram utilizados os recursos disponíveis no programa TBtools (CHEN, C. *et al.*, 2020).

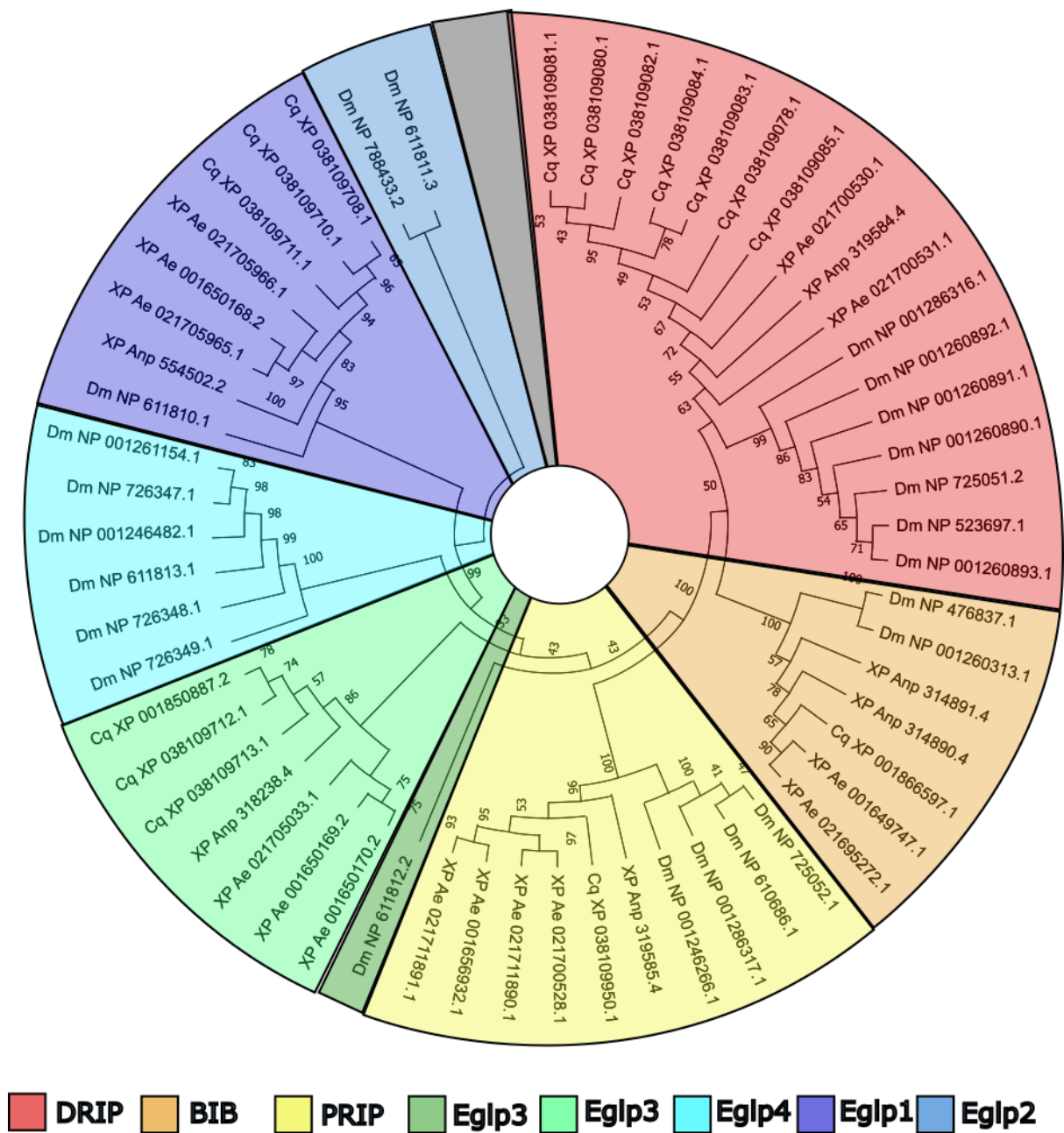
5. RESULTADOS

5.1. Identificação e classificação das aquaporinas em *Culex quinquefasciatus*

As pesquisas conduzidas com as sequências-sonda (*query*) das aquaporinas (AQPs) de *Drosophila melanogaster*, *Anopheles gambiae* e *Aedes aegypti* permitiram a identificação de 15 sequências codificando prováveis AQPs no genoma de *Culex quinquefasciatus*, incluindo as isoformas. Os IDs das proteínas estão descritos na Tabela 1.

A classificação das aquaporinas presentes no genoma de *Culex quinquefasciatus* procedeu conforme sua semelhança com as AQPs de *Drosophila melanogaster*. Desse modo, o dendrograma construído a partir de 15 sequências de *Culex quinquefasciatus*, 23 sequências de *Drosophila melanogaster*, seis sequências de *Anopheles gambiae* e 14 sequências de *Aedes aegypti* revelou importantes associações entre as aquaporinas (Figura 7). Sete sequências de *Culex quinquefasciatus* formaram um clado junto a dez sequências de outras espécies, incluindo as aquaporinas DRIP de *Drosophila*. A sequência XP_001866597.1 de *Culex quinquefasciatus* foi a única presente no clado contendo BIB de *Drosophila*, junto a outras duas sequências de *Anopheles gambiae* e duas de *Aedes aegypti*. No clado PRIP de *Drosophila*, verificou-se a presença de um único ortólogo de *Culex quinquefasciatus*, junto a uma sequência de *Anopheles gambiae* e quatro de *Aedes aegypti*. Por fim, observou-se a presença de sequências de *Culex quinquefasciatus* nos grupos correspondentes às entomogliceroporinas 1 (Eg1p1) e 3 (Eg1p3) de *Drosophila*.

Figura 7. Análise fenética refletindo os padrões de distribuição das aquaporinas no genoma de *Culex quinquefasciatus*.

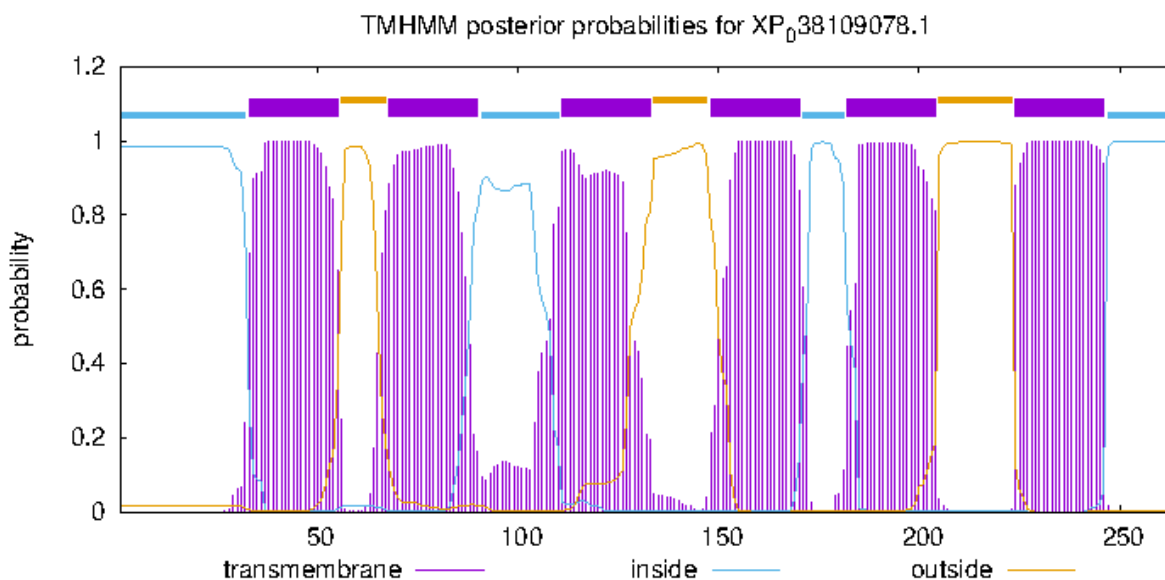


5.2. Caracterização molecular das aquaporinas de *Culex quinquefasciatus*

De acordo com os dados na Tabela 1, o peso molecular (PM) variou de 26,72 kDa (XP_038109713.1 e XP_038109712.1) a 69,50 kDa (XP_001866597.1). O ponto isoeletrico (PI) variou de 5,32 (XP_038109711.1) a 9,30 kDa (XP_001866597.1). O tamanho das aquaporinas variou de 250 (XP_038109713.1 e XP_038109712.1) a 638 aminoácidos (XP_001866597.1).

Com relação à análise de domínios, as 15 sequências candidatas (isoformas) apresentaram o domínio MIP (Major Intrinsic Proteins) (Apêndice 1). Todas as aquaporinas previstas apresentaram o mesmo número de hélices transmembranas (HTM 1-6), com as regiões N e C terminais localizadas no citoplasma e alças (A-E) características que ligam as seis hélices transmembranas. Destas, três alças foram externas (A, C e E) e duas alças foram internas (B e D) (Figura 8). Quanto à localização subcelular, todas as isoformas foram previstas para a membrana plasmática, com exceção de XP_001866597.1, que também foi associada ao núcleo (Tabela 1).

Figura 8. Predição dos segmentos transmembranares e topologia da sequência XP_038109078 de *Culex quinquefasciatus*.



Fonte: O autor (2023).

Tabela 1. Dados da caracterização estrutural e funcional das aquaporinas de *Culex quinquefasciatus*.

ID da proteína	Peso molecular (kDa)	Ponto isoelétrico (PI)	Tamanho (aa)	Localização subcelular	Nº Hélices TM
XP_038109084.1	26,78	7,88	256	Membrana citoplasmática	6
XP_038109083.1	26,78	7,88	256	Membrana citoplasmática	6
XP_038109082.1	26,78	7,88	256	Membrana citoplasmática	6
XP_038109081.1	26,78	7,88	256	Membrana citoplasmática	6
XP_038109080.1	26,78	7,88	256	Membrana citoplasmática	6
XP_038109085.1	26,73	8,33	253	Membrana citoplasmática	6
XP_038109078.1	27,77	8,32	263	Membrana citoplasmática	6
XP_038109950.1	27,69	6,48	263	Membrana citoplasmática	6
XP_038109713.1	26,72	7,06	250	Membrana citoplasmática	6
XP_038109712.1	26,72	7,06	250	Membrana citoplasmática	6
XP_001850887.2	28,01	7,53	261	Membrana citoplasmática	6
XP_038109708.1	31,8	5,9	296	Membrana citoplasmática	6
XP_038109711.1	28,86	5,32	267	Membrana citoplasmática	6
XP_038109710.1	31,39	6,04	294	Membrana citoplasmática	6
XP_001866597.1	69,5	9,3	638	Núcleo/ Membrana citoplasmática	6

Legenda de cores: Vermelho: DRIP; Amarelo: PRIP; Verde:

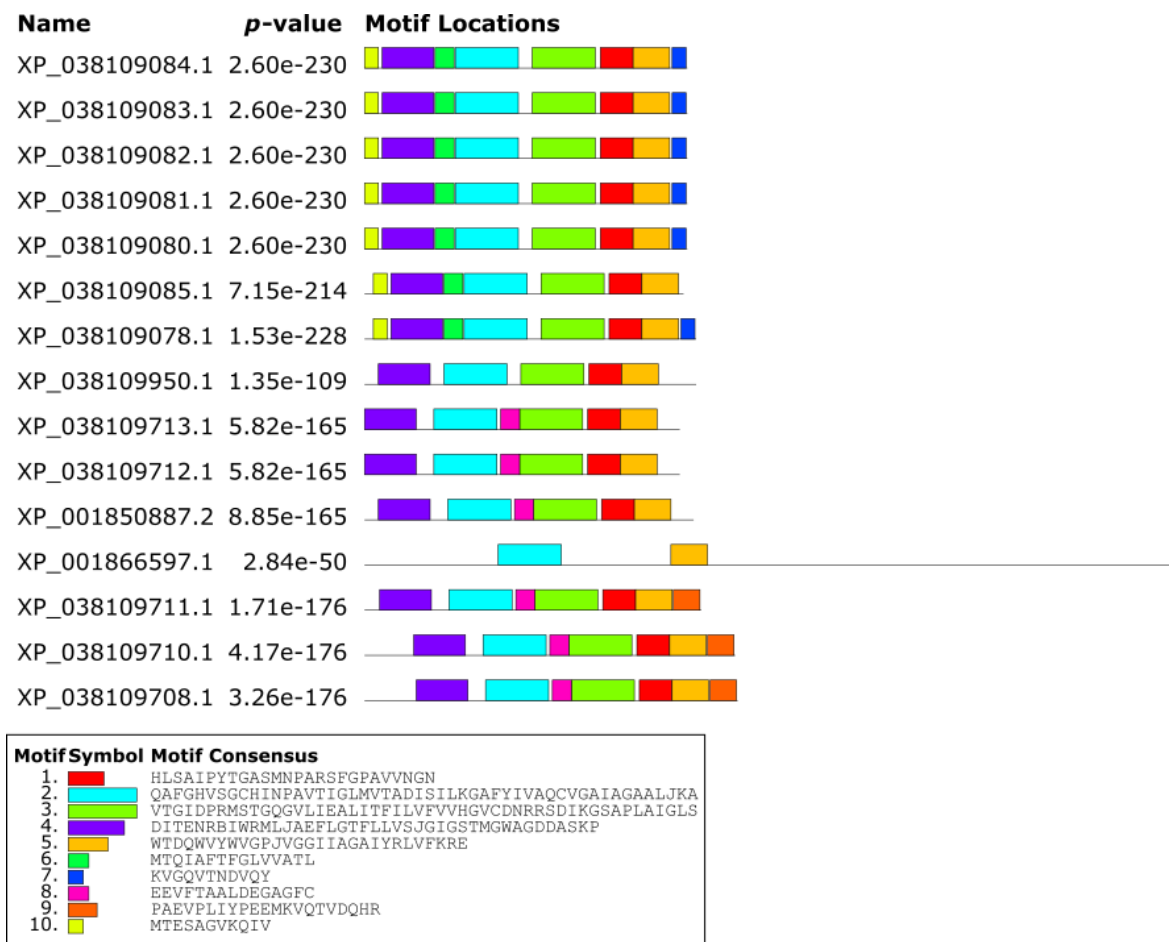
Entomogliceroporinas;

Laranja: BIB.

Fonte: O autor (2023).

Adicionalmente, foram analisados os motivos conservados dentro dos domínios das aquaporinas candidatas. Por meio da busca através do software MEME, foram identificados dez motivos, variando de dois a oito por sequência. Uma análise mais detalhada revelou que foi obtido um padrão associado com cada uma das famílias (DRIP, BIB, PRIP, Eglp1 e Eglp3) (Figura 9). A maioria das proteínas da mesma subfamília apresentam motivos semelhantes. Os motivos 2 e 5 estão presentes em todas as subfamílias (Figura 9). As sequências compartilham alta similaridade. De 10 motivos, a metade destes (1, 2, 3, 4 e 5) foram encontrados em todos os membros, exceto a XP_001866597.1 ortóloga da BIB, que apresentou apenas os motivos 2 e 5.

Figura 9. Motivos conservados das aquaporinas identificadas no genoma de *Culex quinquefasciatus*. Cada barra representa um motivo detectado, nas quais barras de mesma cor representam motivos em comum.

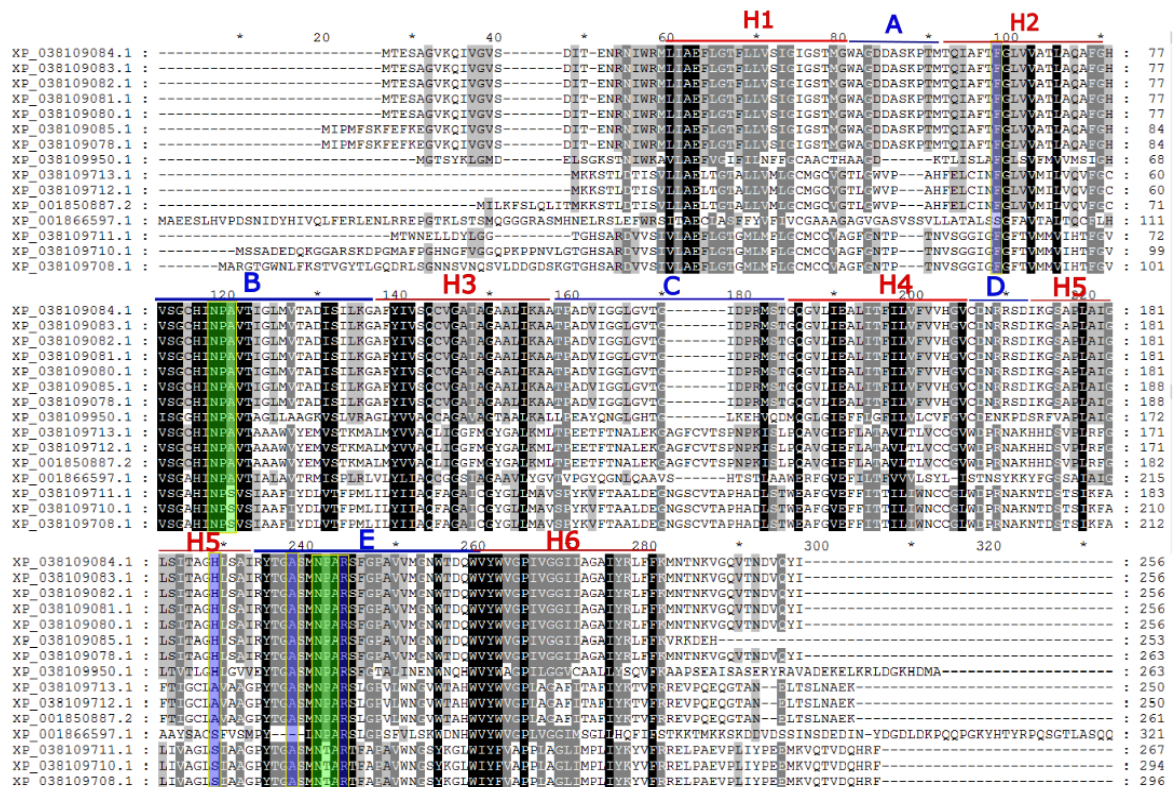


Fonte: O autor (2023).

Os motivos também foram analisados por meio do alinhamento múltiplo de sequências (Figura 10). Das 15 sequências de *Culex quinquefasciatus* analisadas, todas apresentaram duas caixas NPA conservadas, com exceção de três proteínas pertencentes ao grupo da entomogliceroprina 1 (Eglp1). Em sua primeira caixa, o aminoácido alanina foi substituído por serina (N-P-S). Na segunda caixa, houve a substituição de prolina por tirosina (N-T-A).

Os filtros de seletividade ar/R foram observados e marcados no alinhamento (Figura 10). Um deles está localizado em HTM2, outro em HTM5 e dois na alça E. O primeiro filtro, localizado em HTM2, corresponde a fenilalanina (F). Há uma modificação neste filtro apenas no grupo BIB, na qual a fenilalanina é substituída por uma serina (F/S). Este mantém-se conservado em todas as outras subfamílias. O segundo filtro, localizado em HTM5, apresenta histidina (H) nos grupos DRIP e PRIP, alanina (A) no grupo Eglp3; serina nos grupos Eglp1 e BIB. No terceiro filtro, localizado na alça E, todos os grupos possuem alanina (A) conservada. O último filtro, localizado após a segunda caixa NPA, na alça E, possui arginina (R) conservada em todos os grupos analisados.

Figura 10. Alinhamento múltiplo de sequências de aquaporinas identificadas no genoma de *Culex quinquefasciatus*.



Legenda: Caixas em verde: motivos NPA; Caixas em azul: Filtros de seletividade ar/R; H1-H6: hélices transmembranas; A-E: alças.

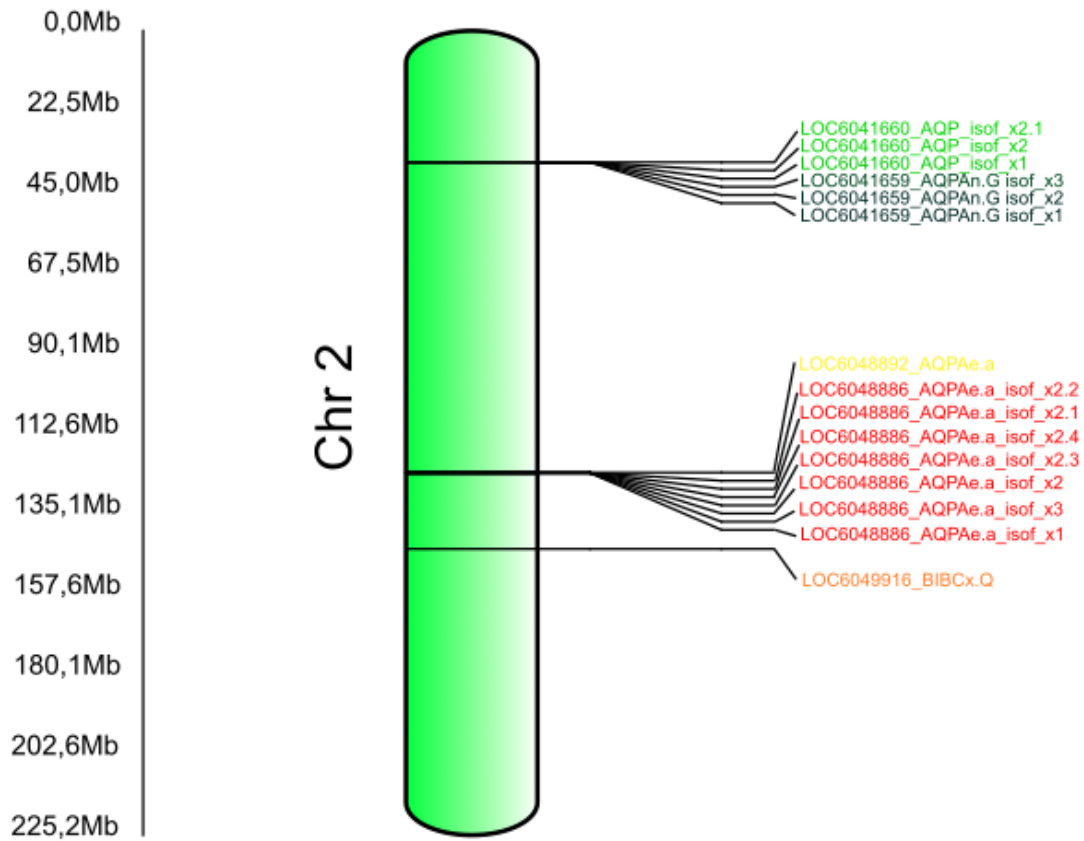
Fonte: O autor (2023).

5.3. Estrutura gênica e localização cromossômica das aquaporinas de *Culex quinquefasciatus*

Quanto à localização cromossômica, todas as sequências identificadas ancoraram no cromossomo de número dois (Chr 2) (Figura 11). Observa-se que as isoformas que compõem os grupos DRIP e PRIP estão fisicamente localizadas próximas entre si, na posição aproximada de 125 Mb, do mesmo modo que as entomogliceroporinas estão agrupadas na posição aproximada de 30 Mb. Já a sequência correspondente à BIB encontra-se mais próxima do primeiro grupo, na posição próxima a 145 Mb.

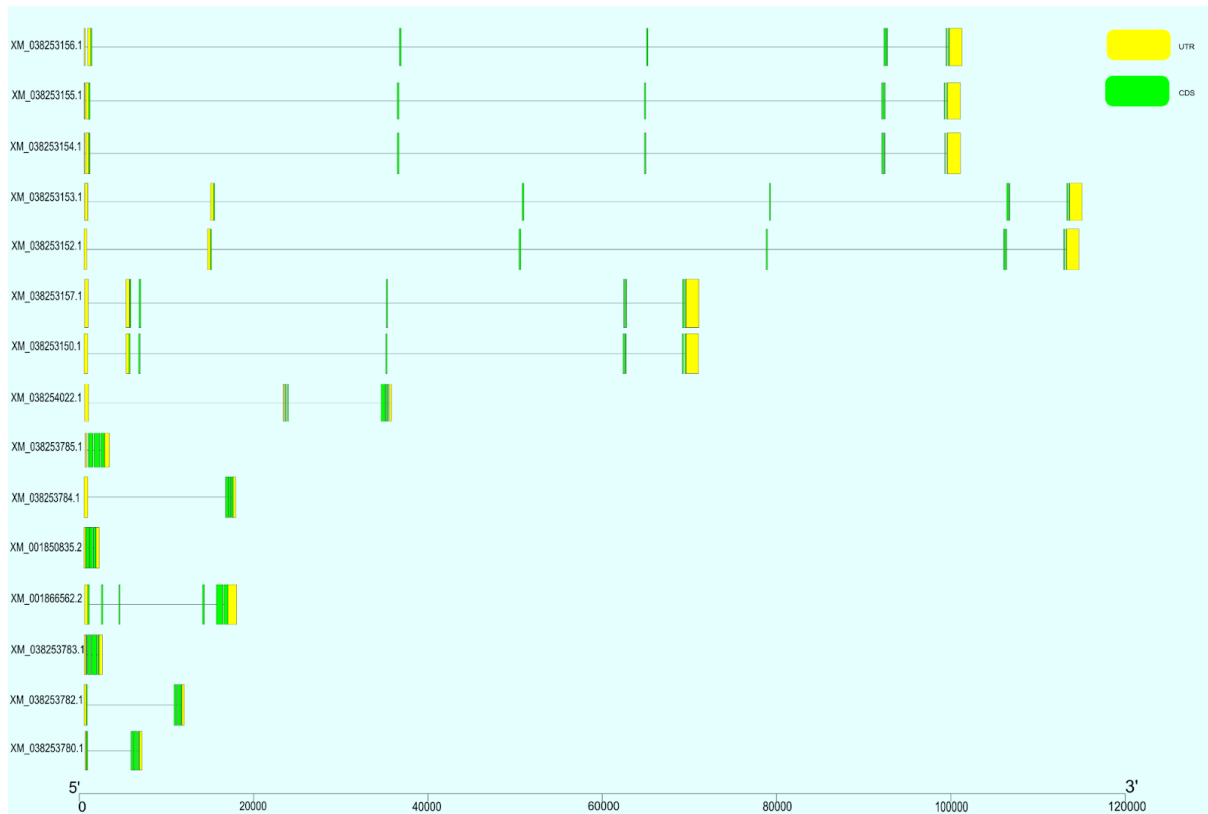
Grandes variações na estrutura e organização íntron-éxon no genoma de *Culex quinquefasciatus* foram observadas (Figura 12; Tabela 2; Apêndice 2). Os transcritos variaram de 1,3 kb (XP_038109713.1) a 115,1 kb (XP_038109081.1) em tamanho. Com relação à estrutura gênica, o número de éxons variou de três (XP_001850887.2) a oito (CqPRIP). Observa-se que a maioria dos íntrons foram significativamente maiores que os éxons, o que explica a grande diferença de tamanho observada nesses genes, mas menos significativas com relação ao tamanho da proteína. Em geral, o número de éxons foi conservado dentro das subfamílias. O grupo DRIP mostrou oito éxons, enquanto BIB e PRIP seis éxons. As Eglps mostraram predominantemente quatro éxons.

Figura 11. Distribuição cromossômica das aquaporinas no genoma de *Culex quinquefasciatus*. A escala está em milhões de bases. Todas estão localizadas no cromossomo dois (Chr 2).



Fonte: O autor (2023).

Figura 12. Estrutura e organização íntron-éxon de genes aquaporina em *Culex quinquefasciatus*.



Legenda: CDS: Coding Sequence (Sequência Codificante); UTR: Untranslated Region (Região Não-Traduzida).

Fonte: O autor (2023).

Tabela 2. Dados quanto ao nome do gene, ID do transcrito, ID da proteína, número de éxons, localização na fita do DNA, localização cromossômica e início e fim dos genes aquaporinas no genoma de *Culex quinquefasciatus*.

Nome do gene	ID do transcrito	ID da proteína	Exon nº	Fita	Cromossomo	Início	Fim
LOC6048886	XM_038253156.1	XP_038109084.1	8	+	Chr 2	123737820	123838759
LOC6048886	XM_038253155.1	XP_038109083.1	8	+	Chr 2	123738069	123838759
LOC6048886	XM_038253154.1	XP_038109082.1	8	+	Chr 2	123738058	123838759
LOC6048886	XM_038253153.1	XP_038109081.1	8	+	Chr 2	123723657	123838759
LOC6048886	XM_038253152.1	XP_038109080.1	8	+	Chr 2	123724019	123838759
LOC6048886	XM_038253157.1	XP_038109085.1	8	+	Chr 2	123767908	123838759
LOC6048886	XM_038253150.1	XP_038109078.1	8	+	Chr 2	123767908	123838759
LOC6048892	XM_038254022.1	XP_038109950.1	6	-	Chr 2	123651007	123685710
LOC6041660	XM_038253785.1	XP_038109713.1	4	-	Chr 2	36797269	36798581
LOC6041660	XM_038253784.1	XP_038109712.1	4	-	Chr 2	36797269	36814309
LOC6041660	XM_001850835.2	XP_001850887.2	3	-	Chr 2	36797269	36798448
LOC6049916	XM_001866562.2	XP_001866597.1	6	-	Chr 2	145040559	145058108
LOC6041659	XM_038253783.1	XP_038109711.1	4	-	Chr 2	36825660	36827060
LOC6041659	XM_038253782.1	XP_038109710.1	4	-	Chr 2	36825660	36836785
LOC6041659	XM_038253780.1	XP_038109708.1	4	-	Chr 2	36825660	36831517

Fonte: O autor (2023)

6. DISCUSSÃO

No presente estudo, foi realizada a identificação genômica e a caracterização molecular *in silico* das aquaporinas presentes em *Culex quinquefasciatus* para tentar compreender melhor o papel biológico dessas proteínas na espécie. O primeiro trabalho que sistematicamente identificou esses homólogos em espécies de mosquitos foi realizado por Drake *et al.* (2010). Eles utilizaram as aquaporinas de seres humanos e de *Drosophila melanogaster*, até então as principais conhecidas, como modelos para comparação com as demais. Neste estudo, foram identificadas seis sequências genômicas correspondentes a aquaporinas em *Aedes aegypti*. Para reconstruir a história evolutiva dessa família de proteínas, foram utilizadas 12 sequências do genoma de *Homo sapiens*, oito sequências da espécie *Drosophila melanogaster*, sete sequências de *Anopheles gambiae*, além de sequências para outras cinco espécies disponíveis (*Pediculus humanus*, *Tribolium castaneum*, *Leishmania major*, *Plasmodium sp.* e *Pichia pastoris*). Visto que em *Culex quinquefasciatus* há provavelmente a ocorrência de seis genes e 15 isoformas, este valor está coerente com o esperado. Ressaltamos que naquele momento não havia informações provenientes de transcriptomas como atualmente.

Todos os genes que codificam aquaporina em *Culex quinquefasciatus* foram direcionados para o cromossomo 2. Para outras espécies de mosquitos, não existem trabalhos com esta informação, constando apenas nas bases de dados. Todavia, nem sempre tal informação está clara. Uma análise no FlyBase (<https://flybase.org/search/#/page/1>) mostrou que em *Drosophila melanogaster* os genes aquaporinas estão, pelo menos, mapeados nos cromossomos 2R e 2L.

Com relação à classificação, as principais subfamílias de aquaporinas reportadas para *Drosophila melanogaster* também foram observadas nas espécies de mosquitos analisadas: DRIP e PRIP (canais específicos para o transporte altamente seletivo de água), BIB (envolvida na regulação da adesão celular) e o grupo das entomogliceroporinas. Há alguns anos, estudos filogenéticos anteriores não foram capazes de identificar em insetos holometábolos um grupo especial de aquaporinas adaptado para o transporte de outras pequenas moléculas, como solutos e glicerol (ABASCAL *et al.*, 2014; DRAKE *et al.*, 2010), presente na grande maioria dos seres vivos (TONG *et al.*, 2019). Todavia, trabalhos experimentais apontavam para sua existência (DRAKE *et al.*, 2010; 2015; KATAOKA; MIYAKE;

AZUMA, 2009).

Somente há alguns anos, esse paradoxo foi resolvido com a descoberta das entomogliceroporinas, uma subfamília exclusiva de aquaporinas em insetos que evoluiu a partir de uma aquaporina ancestral, cuja função primordial era o transporte de água. Ao sofrer mutação, essa variante passou a desempenhar função no transporte de glicerol e outras moléculas. Com o passar do tempo, mostrou vantagem adaptativa em relação à aquagliceroporina ortodoxa, levando à perda desta na linhagem nos holometábolos e o surgimento daquela novo grupo, onde em Brachycera encontramos as Eglps 1-4 e em Nematocera apenas as Eglps 1 e 2. Este processo de evolução gênica, é denominado neofuncionalização (FINN *et al.*, 2015).

De modo geral, as aquaporinas identificadas em *Culex quinquefasciatus* mostraram peso molecular, tamanho da cadeia de aminoácidos, presença de domínios e motivos característicos para essa proteína (BENGA, 2009). Particularmente, os motivos NPA mostraram sua assinatura característica, onde todas as aquaporinas ortodoxas mantiveram sua sequência canônica conservada. Embora diferentes nas entomogliceroporinas, essas caixas NPA revelaram-se conservadas dentro dos mesmo grupos de subfamílias (NPA-NPA/NPS-NTA).

As alterações observadas nas regiões correspondentes aos filtros de seletividade ar/R apontam para mudanças de aminoácidos relacionadas à sua função na estrutura tridimensional da proteína canal e, conseqüentemente, ao tipo de soluto transportado via canal. Infelizmente, este trabalho não analisou a estrutura proteica das aquaporinas de *Culex quinquefasciatus*, não sendo possível confirmar essa hipótese. Experimentos futuros de modelagem de aquaporinas podem trazer à luz novas informações relevantes sobre esses sítios modificados (DRAKE *et al.*, 2015; FINN *et al.*, 2015).

A estrutura gênica das aquaporinas observada em *Culex quinquefasciatus* refletiu sua distribuição filogenética entre as subfamílias, mostrando-se conservada dentro delas, mas diferentes entre elas. Do mesmo modo, as propriedades bioquímicas e físicas também estiveram de acordo com as informações filogenéticas. Em plantas, vários genomas apontam para essa organização (BEZERRA-NETO *et al.*, 2019).

Este foi o primeiro trabalho que analisou a diversidade estrutural e funcional dos genes que codificam a família das aquaporinas no mosquito *Culex*

quinquefasciatus. Pelos resultados obtidos, pode-se afirmar que ferramentas de análise in silico são poderosos recursos para acessar os genomas das espécies e contribuir para o maior conhecimento de sua biologia. Futuros estudos, seja quanto ao padrão específico de expressão gênica dessas moléculas ou ao efeito das mutações em sua estrutura e função, por exemplo, ajudarão a ampliar nossa visão sobre esses canais transmembrana. Desse modo, entenderemos melhor sobre seu papel funcional, sobretudo, no restabelecimento da homeostase do vetor, quando desafiado por agentes estressores ambientais e endógenos.

7. CONCLUSÕES

- No genoma de *Culex quinquefasciatus* estão presentes seis genes que provavelmente codificam aquaporinas. Foram encontradas 15 possíveis candidatas, incluindo isoformas, que foram classificadas nas subfamílias DRIP, BIB, PRIP, Eglp1 e Eglp3.
- A análise *in silico* das sequências proteicas das aquaporinas candidatas de *Culex quinquefasciatus* confirmou a presença dos domínios MIP (Major Intrinsic Proteins) conservados, assim como a topologia prevista, composta por seis hélices transmembranas (HTM1-6) ligadas por cinco alças (A-E).
- A maioria das aquaporinas identificadas no genoma de *Culex quinquefasciatus* apresentou peso molecular em torno de 28 kDa, assim como localização subcelular e outros parâmetros físico-químicos característicos.
- As aquaporinas de *Culex quinquefasciatus* mostraram motivos e estrutura gênica conservados dentro de uma mesma família, mas sutilmente diferentes entre os membros de diferentes famílias, o que reflete suas diferentes funções.
- Todos os genes que codificam aquaporinas em *Culex quinquefasciatus* estão localizados no cromossomo 2 (Chr 2).

8. REFERÊNCIAS BIBLIOGRÁFICAS

AGRE, P. *et al.* Aquaporin CHIP: the archetypal molecular water channel. *American journal of physiology. Renal physiology.* v. 265, n. 4, p. F463–F476, 1993. Disponível em: <<http://dx.doi.org/10.1152/ajprenal.1993.265.4.f463>>.

ARENSBURGER, P. *et al.* Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science (New York, N.Y.).* v. 330, n. 6000, p. 86–88, 2010. Disponível em: <<http://dx.doi.org/10.1126/science.1191864>>.

AZAD, A. K. *et al.* Human aquaporins: Functional diversity and potential roles in infectious and non-infectious diseases. *Frontiers in genetics,* v. 12, 2021.

BAILEY, T. L. *et al.* The MEME suite. *Nucleic acids research,* v. 43, n. W1, p. W39–W49, 2015.

BARBOSA, R. M. R. *et al.* Laboratory and field evaluation of an oviposition trap for *Culex quinquefasciatus* (Diptera: Culicidae). *Memórias do Instituto Oswaldo Cruz.* v. 102, n. 4, p. 523–529, 2007. Disponível em: <<http://dx.doi.org/10.1590/s0074-02762007005000058>>.

BECKER, N. *et al.* *Mosquitoes: Identification, ecology and control.* 3. ed. Cham, Switzerland: Springer Nature, 2020. 279 p.

BENGA, G. Water channel proteins (later called aquaporins) and relatives: Past, present, and future. *IUBMB life.* v. 61, n. 2, p. 112–133, 2009. Disponível em: <<http://dx.doi.org/10.1002/iub.156>>.

BENOIT, J. B. *et al.* Emerging roles of aquaporins in relation to the physiology of blood-feeding arthropods. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology.* v. 184, n. 7, p. 811–825, 2014. Disponível em: <<http://dx.doi.org/10.1007/s00360-014-0836-x>>.

BEZERRA-NETO, J. P. *et al.* Plant Aquaporins: Diversity, Evolution and Biotechnological Applications. *Current Protein Peptide Science.*v. 20, n. 4, p.

368-395, 2019. Disponível em:
<<http://dx.doi.org/10.2174/1389203720666181102095910>>.

BORROR, D.J.; DeLONG, D.M. Introdução ao estudo dos insetos. São Paulo. Ed. Blücher, 1988. 653 p.

CHAKRABARTI, N. *et al.* Molecular basis of proton blockage in aquaporins. *Structure* (London, England: 1993). v. 12, n. 1, p. 65–74, 2004. Disponível em:
<<http://dx.doi.org/10.1016/j.str.2003.11.017>>.

CHEN, C. *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Molecular plant*. v. 13, n. 8, p. 1194–1202, 2020. Disponível em:
<<http://dx.doi.org/10.1016/j.molp.2020.06.009>>.

CHEN, H.; WU, Y.; VOTH, G. A. Origins of proton transport behavior from selectivity domain mutations of the aquaporin-1 channel. *Biophysical journal*. v. 90, n. 10, p. L73–L75, 2006. Disponível em: <<http://dx.doi.org/10.1529/biophysj.106.084061>>.

CLARK, M. S.; WORLAND, M. R. How insects survive the cold: molecular mechanisms. *Journal of comparative physiology. B, Biochemical, systemic, and environmental physiology*. v. 178, n. 8, p. 917–933, 2008. Disponível em:
<<http://dx.doi.org/10.1007/s00360-008-0286-4>>.

CLEMENTS, A. *Biology of mosquitoes : Development nutrition and reproduction*. London, England: Chapman and Hall, 1992. 509 p.

CONSOLI, R. A. G. B.; OLIVEIRA, R. L. *Principais mosquitos de importância sanitária no Brasil*. Rio de Janeiro: Editora Fiocruz, 1994. 228p.

DENKER, B. M. *et al.* Identification, purification, and partial characterization of a novel 28,000 integral membrane protein from erythrocytes and renal tubules. *The journal of biological chemistry*. v. 263, n. 30, p. 15634–15642, 1988. Disponível em:
<[http://dx.doi.org/10.1016/s0021-9258\(19\)37635-5](http://dx.doi.org/10.1016/s0021-9258(19)37635-5)>.

DRAKE, L. L. *et al.* The aquaporin gene family of the yellow fever mosquito, *Aedes aegypti*. *PloS one*. v. 5, n. 12, p. e15578, 2010. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0015578>>.

DRAKE, L. L.; RODRIGUEZ, S. D.; HANSEN, I. A. Functional characterization of aquaporins and aquaglyceroporins of the yellow fever mosquito, *Aedes aegypti*. *Scientific reports*. v. 5, n. 1, 2015. Disponível em: <<http://dx.doi.org/10.1038/srep07795>>.

DUMAN, J. G. Antifreeze and ice nucleator proteins in terrestrial arthropods. *Annual review of physiology*. v. 63, n. 1, p. 327–357, 2001. Disponível em: <<http://dx.doi.org/10.1146/annurev.physiol.63.1.327>>.

FINN, R. N. *et al.* The lineage-specific evolution of aquaporin gene clusters facilitated tetrapod terrestrial adaptation. *PloS one*. v. 9, n. 11, p. e113686, 2014. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0113686>>.

FLYBASE. Disponível em: <https://flybase.org/search/#/page/1>.

FOOTTIT, R. G.; ADLER, P. H. *Insect biodiversity: Science and society*. Chichester, England: Wiley-Blackwell, 2009. 632p.

FORATTINI, O.P. *Culicidologia médica*. São Paulo: Editora da USP, 2002. v.2, 860 p.

FU, D.; LU, M. The structural basis of water permeation and proton exclusion in aquaporins (Review). *Molecular membrane biology*. v. 24, n. 5–6, p. 366–374, 2007. Disponível em: <<http://dx.doi.org/10.1080/09687680701446965>>.

GEADKAEW, A. Human parasitic aquaporins: New anti-parasitic drug targets. Disponível em: <<http://dx.doi.org/10.13140/RG.2.1.2123.6085>>.

GUEDES, D. R. D. *et al.* Zika virus replication in the mosquito *Culex quinquefasciatus* in Brazil. *Emerging microbes & infections*, 2017. v. 6, n. 1, p. 1–11. Disponível em: <<http://dx.doi.org/10.1038/emi.2017.59>>.

HALLGREN, J. *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. 2022. Disponível em: <<http://dx.doi.org/10.1101/2022.04.08.487609>>.

HARBACH, R. E. *Culex Pipiens: Species Versus Species Complex - Taxonomic History and Perspective.* J. Am. Mosq. Control. Assoc. v. 28, p. 10–23, 2012.

HILLER, K. *et al.* JVirGel 2.0: computational prediction of proteomes separated via two-dimensional gel electrophoresis under consideration of membrane and secreted proteins. *Bioinformatics (Oxford, England).* v. 22, n. 19, p. 2441–2443, 2006. Disponível em: <<http://dx.doi.org/10.1093/bioinformatics/btl409>>.

HOLT, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science (New York, N.Y.).* v. 298, n. 5591, p. 129–149, 2002. Disponível em: <<http://dx.doi.org/10.1126/science.1076181>>.

JUNG, J. S. *et al.* Molecular structure of the water channel through aquaporin CHIP. The hourglass model. *The journal of biological chemistry.* v. 269, n. 20, p. 14648–14654, 1994 Disponível em: <[http://dx.doi.org/10.1016/s0021-9258\(17\)36674-7](http://dx.doi.org/10.1016/s0021-9258(17)36674-7)>.

KATAOKA, N.; MIYAKE, S.; AZUMA, M. Aquaporin and aquaglyceroporin in silkworms, differently expressed in the hindgut and midgut of *Bombyx mori*. *Insect molecular biology.* v. 18, n. 3, p. 303–314, 2009. Disponível em: <<http://dx.doi.org/10.1111/j.1365-2583.2009.00871.x>>.

KRUSE, E.; UEHLEIN, N.; KALDENHOFF, R. The aquaporins. *Genome Biology.* v. 7, n. 2, p. 206, 2006. Disponível em: <<http://dx.doi.org/10.1186/gb-2006-7-2-206>>.

LOZOVEI, A. L. *Culicidae (Mosquitos).* In: MARCONDES, C. B. *et al.* *Entomologia Médica e Veterinária.* 2. ed. São Paulo: Atheneu, 2011. Cap. 8, p. 107-144.

MELO, L. L. *Adaptação à hematofagia: como fêmeas e machos de *Culex quinquefasciatus* (diptera: culicidae) enfrentam o estresse oxidativo e o sistema complemento humano após repasto sanguíneo.* 63 f Dissertação (Mestrado) - Universidade Federal de Minas Gerais, Belo Horizonte, 2020.

NENE, V. *et al.* Genome Sequence of *Aedes aegypti*, a Major Arbovirus Vector. *Science* (New York, N.Y.). v. 316, n. 5832, p. 1718–1723, 2007. Disponível em: <<http://dx.doi.org/10.1126/science.1138878>>.

PERUZIN, M.C.J. Análises comparativas populacionais de *Culex quinquefasciatus* em dois locais do Estado de São Paulo. 2009. 97 f. Dissertação (Mestrado) - Universidade de São Paulo, São Paulo, 2009.

PRESTON, G. M. *et al.* Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* (New York, N.Y.). v. 256, n. 5055, p. 385–387, 1992. Disponível em: <<http://dx.doi.org/10.1126/science.256.5055.385>>.

REDDY, B. P. N.; LABBÉ, P.; CORBEL, V. *Culex* genome is not just another genome for comparative genomics. *Parasites & vectors*. v. 5, n. 1, 2012. Disponível em: <<http://dx.doi.org/10.1186/1756-3305-5-63>>.

REINHARDT, H. *et al.* Tonoplast aquaporins facilitate lateral root emergence. *Plant physiology*, v. 170, n. 3, p. 1640–1654, 2016.

REIS, L. A. M. *et al.* Vector Competence of *Culex quinquefasciatus* from Brazil for West Nile Virus. *Tropical medicine and infectious disease*. v. 8, n. 4, p. 217, 2023. Disponível em: <<http://dx.doi.org/10.3390/tropicalmed8040217>>.

RIOS, F. G. F. *et al.* *Culex quinquefasciatus* predominance during integrated mosquito surveillance in an urban area of the Brazilian Amazon. *Brazilian Journal of Biology*. v. 82, 2022. Disponível em: <<http://dx.doi.org/10.1590/1519-6984.266219>>.

ROSEN, L. *et al.* Transovarial Transmission of Dengue Viruses by Mosquitoes: *Aedes albopictus* and *Aedes aegypti*. *The American journal of tropical medicine and hygiene*. v. 32, n. 5, p. 1108–111, 1983. Disponível em: <<http://dx.doi.org/10.4269/ajtmh.1983.32.1108>>.

SAMY, A. M. *et al.* Climate change influences on the global potential distribution of the mosquito *Culex quinquefasciatus*, vector of west Nile virus and lymphatic

filariasis. PloS one. v. 11, n. 10, p. e0163863, 2016. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0163863>>.

SIRIVANAKARN, S. Medical entomology studies. A revision of the subgenus Culex in the Oriental region (Diptera: Culicidae). Contributions of the American Entomological Institute. v. 12, n. 2, p. 1-272, 1976.

SMITH, B. L.; AGRE, P. Erythrocyte Mr 28,000 transmembrane protein exists as a multisubunit oligomer similar to channel proteins. The journal of biological chemistry. v. 266, n. 10, p. 6407–6415, 1991. Disponível em: <[http://dx.doi.org/10.1016/s0021-9258\(18\)38133-x](http://dx.doi.org/10.1016/s0021-9258(18)38133-x)>.

SPRING, J. H.; ROBICHAUX, S. R.; HAMLIN, J. A. The role of aquaporins in excretion in insects. The journal of experimental biology. v. 212, n. 3, p. 358–362, 2009. Disponível em: <<http://dx.doi.org/10.1242/jeb.024794>>.

STILLWELL, W. An introduction to biological membranes: From bilayers to rafts. London, England: Elsevier Science, 2013. 367 p.

TAJKHORSHID, E. *et al.* Control of the selectivity of the aquaporin water channel family by global orientational tuning. Science (New York, N.Y.). v. 296, n. 5567, p. 525–530, 2002. Disponível em: <<http://dx.doi.org/10.1126/science.1067778>>.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research. v. 22, n. 22, p. 4673–4680, 1994. Disponível em: <<http://dx.doi.org/10.1093/nar/22.22.4673>>.

TIMMERMANN, S. E.; BRIEGEL, H. Larval growth and biosynthesis of reserves in mosquitoes. Journal of insect physiology, 1999. v. 45, n. 5, p. 461–470. Disponível em: <[http://dx.doi.org/10.1016/s0022-1910\(98\)00147-4](http://dx.doi.org/10.1016/s0022-1910(98)00147-4)>.

TONG, H.; HU, Q.; ZHU, L. DONG, X. Prokaryotic Aquaporins. Cells. v. 24; n. 8, p.1316. 2019.

WALLACE, I. S. *et al.* Acyrthosiphon pisum AQP2: A multifunctional insect aquaglyceroporin. *Biochimica et biophysica acta. Biomembranes*. v. 1818, n. 3, p. 627–635, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.bbamem.2011.11.032>>.

9. APÊNDICES

Apêndice 1. Resultado da análise de domínio via NCBI Batch CD-search tool

Query	PSSM-ID	From	To	E-Value	Bitscore	Accession	Short name	Incomplete	Superfamily
Q#1 - >XP_0381090 84.1	444743	23	235	2,03E-69	225 659	cl00200	MIP superfamily	-	-
Q#2 - >XP_0381090 83.1	444743	23	235	2,03E-69	225 659	cl00200	MIP superfamily	-	-
Q#3 - >XP_0381090 82.1	444743	23	235	2,03E-69	225 659	cl00200	MIP superfamily	-	-
Q#4 - >XP_0381090 81.1	444743	23	235	2,03E-69	225 659	cl00200	MIP superfamily	-	-
Q#5 - >XP_0381090 80.1	444743	23	235	2,03E-69	225 659	cl00200	MIP superfamily	-	-
Q#6 - >XP_0381090 85.1	444743	30	242	4,61E-70	227.2	cl00200	MIP superfamily	-	-
Q#7 - >XP_0381090 78.1	444743	30	242	5,65E-70	227 585	cl00200	MIP superfamily	-	-
Q#8 - >XP_0381099 50.1	444743	15	226	3,84E-63	210 251	cl00200	MIP superfamily	-	-
Q#9 - >XP_0381097 13.1	444743	13	228	1,67E-44	162 038	cl00200	MIP superfamily	-	-
Q#10 - >XP_0381097 12.1	444743	13	228	1,67E-44	162 038	cl00200	MIP superfamily	-	-
Q#11 - >XP_0018508 87.2	444743	24	239	2,19E-44	162 038	cl00200	MIP superfamily	-	-
Q#12 - >XP_0018665 97.1	395174	50	265	2,64E-75	253 394	pfam00230	MIP	-	-
Q#13 - >XP_0381097 11.1	444743	25	240	5,75E-36	140 852	cl00200	MIP superfamily	-	-
Q#14 - >XP_0381097 10.1	444743	52	267	7,26E-37	143 933	cl00200	MIP superfamily	-	-
Q#15 - >XP_0381097 08.1	444743	54	269	3,30E-37	144 704	cl00200	MIP superfamily	-	-

Apêndice 2. Planilha contendo ID das proteínas; Número de exons e introns; Início e fim da proteína, introns e exons; Fita; Tamanho total em kilobases.

ID Proteína Nº Exon e introns	Início	Fim	Fita	Total kb
XP_038109084.1	123737820	123838759	+	100,939
Exon 1	123737820	123737938		
Intron 1	123738582	123738597		
Exon 2	123738244	123738597		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		
XP_038109083.1	123738069	123838759	+	100,69
Exon 1	123738069	123738123		
Intron 1	123738582	123738597		
Exon 2	123738244	123738597		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		

XP_038109082.1	123738058	123838759	+	100,701
Exon 1	123738058	123738127		
Intron 1	123738582	123738597		
Exon 2	123738244	123738597		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		
XP_038109081.1	123723657	123838759	+	115,102
Exon 1	123723657	123723781		
Intron 1	123738582	123738597		
Exon 2	123738244	123738597		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		
XP_038109080.1	123724019	123838759	+	114,74
Exon 1	123724019	123724325		

Intron 1	123738582	123738597		
Exon 2	123738244	123738597		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		
XP_038109078.1	123767908	123838759	+	70,851
Exon 1	123767908	123768314		
Intron 1	123773117	123773153		
Exon 2	123772750	123773153		
Intron 2	123774179	123774381		
Exon 3	123774179	123774381		
Intron 3	123802696	123802839		
Exon 4	123802696	123802839		
Intron 4	123830064	123830273		
Exon 5	123830064	123830273		
Intron 5	123830343	123830423		
Exon 6	123830343	123830423		
Intron 6	123837030	123837095		
Exon 7	123837030	123837095		
Intron 7	123837299	123837346		
Exon 8	123837299	123838759		
XP_038109950.1	123651007	123685710	-	34,703
Exon 1	123685275	123685710		
Intron 1	123662914	123662998		
Exon 2	123662914	123663017		
Intron 2	123662733	123662839		

Exon 3	123662733	123662839		
Intron 3	123651712	123652146		
Exon 4	123651712	123652146		
Intron 4	123651523	123651651		
Exon 5	123651523	123651651		
Intron 5	123651245	123651277		
Exon 6	123651007	123651277		
XP_038109713.1	36797269	36798581	-	1,312
Exon 1	36798485	36798581		
Intron 1	36798152	36798400		
Exon 2	36798152	36798409		
Intron 2	36797800	36798093		
Exon 3	36797800	36798093		
Intron 3	36797533	36797739		
Exon 4	36797269	36797739		
XP_038109712.1	36797269	36814309	-	17,04
Exon 1	36813891	36814309		
Intron 1	36798152	36798400		
Exon 2	36798152	36798409		
Intron 2	36797800	36798093		
Exon 3	36797800	36798093		
Intron 3	36797533	36797739		
Exon 4	36797269	36797739		
XP_001850887.2	36797269	36798448	-	1,179
Exon 1	36798152	36798448		
Intron 1	36798152	36798433		
Exon 2	36797800	36798093		
Intron 2	36797800	36798093		
Exon 3	36797269	36797739		
Intron 3	36797533	36797739		
XP_001866597.1	145040559	145058108	-	17,549
Exon 1	145057580	145058108		
Intron 1	145057580	145057752		

Exon 2	145056008	145056174		
Intron 2	145056008	145056174		
Exon 3	145054032	145054158		
Intron 3	145054032	145054158		
Exon 4	145044301	145044508		
Intron 4	145044301	145044508		
Exon 5	145042136	145042917		
Intron 5	145042136	145042917		
Exon 6	145040559	145042017		
Intron 6	145041561	145042017		
XP_038109711.1	36825660	36827060	-	1,4
Exon 1	36826907	36827060		
Intron 1	36826907	36826940		
Exon 2	36826554	36826804		
Intron 2	36826554	36826804		
Exon 3	36826120	36826494		
Intron 3	36826120	36826494		
Exon 4	36825660	36826054		
Intron 4	36825914	36826054		
XP_038109710.1	36825660	36836785	-	11,125
Exon 1	36836435	36836785		
Intron 1	36836435	36836549		
Exon 2	36826554	36826804		
Intron 2	36826554	36826804		
Exon 3	36826120	36826494		
Intron 3	36826120	36826494		
Exon 4	36825660	36826054		
Intron 4	36825914	36826054		
XP_038109708.1	36825660	36831517	-	5,857
Exon 1	36831359	36831517		
Intron 1	36831359	36831479		
Exon 2	36826554	36826804		
Intron 2	36826554	36826804		
Exon 3	36826120	36826494		

Intron 3	36826120	36826494		
Exon 4	36825660	36826054		
Intron 4	36825914	36826054		