



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO  
CENTRO DE SAÚDE E TECNOLOGIA RURAL  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS  
FLORESTAIS**



**GISELA FORMIGA QUEIROZ NÓBREGA**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE AQUAPORINAS EM  
*Cnidioscolus quercifolius* Pohl**

**Patos-PB**

**2017**

**GISELA FORMIGA QUEIROZ NÓBREGA**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE AQUAPORINAS EM**  
*Cnidoscolus quercifolius* Pohl

Dissertação apresentada ao Programa de Pós Graduação em Ciências Florestais da Universidade Federal de Campina Grande como parte dos requisitos exigidos para obtenção do título de Mestre em Ciências Florestais.

**Orientador:** Prof. Dr. Carlos Eduardo Alves Soares

**Patos-PB**

**2017**

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO CSRT DA UFCG

754i  
N

Nóbrega, Gisela Formiga Queiroz

Identificação e caracterização de genes de aquaporinas em *Cnidocolus quercifolius* Pohl / Gisela Formiga Queiroz. – Patos, 2017.  
104 f. : il. color.

Dissertação (Mestrado em Ciências Florestais) – Universidade Federal de Campina Grande, Centro de Saúde e Tecnologia Rural, 2017.

"Orientação: Prof. Dr. Carlos Eduardo Alves Soares".

Referências.

1. Bioinformática. 2. Euphobiaceas. 3. Transportadores de membrana. I. Título.

630\*2

CDU

**GISELA FORMIGA QUEIROZ NÓBREGA**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE AQUAPORINAS EM  
*Cnidocolus quercifolius* Pohl**

Dissertação apresentada ao Programa de Pós Graduação em Ciências Florestais da Universidade Federal de Campina Grande como parte dos requisitos exigidos para obtenção do título de Mestre em Ciências Florestais.

**Aprovada em: 28 de março de 2017.**

---

**Prof. Dr. Carlos Eduardo Alves Soares**  
**Universidade Federal Rural do Semiárido (CCBS/UFERSA)**  
**(Orientador)**

---

**Dr. João Pacífico Bezerra Neto**  
**Universidade Federal do Pernambuco (CB/ Dep. Genética/UFPE)**  
**(1º Examinador)**

---

**Prof<sup>a</sup>. Dr<sup>a</sup>. Ivonete Alves Bakke**  
**Universidade Federal de Campina Grande (CSTR/PPGCF/UFCG)**  
**(2º Examinador)**

## **AGRADECIMENTOS**

A Deus, pelo dom da vida e por todas as bênçãos a mim concedidas, por ser o meu Senhor, guiando-me e fazendo-me sempre forte e perseverante.

A minha mãe, Maria Betânia, minha avó, Terezinha Maciel, e ao meu irmão Reginaldo Júnior, pelo carinho, cuidados e apoio durante toda minha vida. Pelo amor!

Ao meu pai, Marcos Antônio, por todo carinho, apoio e incentivo, sempre me estimulando a crescer profissionalmente e como pessoa, por ser meu exemplo de personalidade.

Ao professor Dr. Carlos Eduardo Alves Soares, pela orientação e confiança, pela amizade, pelos bons conselhos e oportunidades oferecidas. Agradecer por sempre me estimular ao crescimento acadêmico.

À professora Dr<sup>a</sup>. Ana Maria Benko-Iseppon, por ter me acolhido em seu laboratório, pela atenção e confiança, por ter me proporcionado experiências únicas que me fizeram valorizar ainda mais meus estudos e despertar novos interesses.

À Dr<sup>a</sup>. Valesca Pandolfi, por todos os ensinamentos, pela disponibilidade, por ter acreditado em mim, por ter me feito perseverar em momentos difíceis, pela amizade e carinho. Serei sempre grata por tudo que fez por mim.

Aos Drs. José Ribamar Costa Ferreira Neto e João Pacífico Bezerra Neto, e ao Me. Marx Oliveira Lima, pelo conhecimento repassado, pela disponibilidade e, sobretudo, pela amizade e atenção.

A Tia Fátima, pela acolhida e apoio, pelos cuidados, carinho e amizade, por ser uma pessoa de muita fé e exemplo de determinação e força de vontade.

Ao meu namorado, Ewerton Medeiros, por todo amor, dedicação, incentivo e apoio, por estar ao meu lado em todos os momentos, sempre me ajudando nas horas mais difíceis e por ser exemplo de caráter e determinação.

Aos amigos que fiz durante o estágio no LGBV, Marislane, Jéssica, Flávia, Artemisa, Bruna, Caroline, Maria, Stephany, Gabriella, Laís, Diogo, Rodrigo e Manasses, em geral, a todos os membros do LGBV e LGM que contribuíram para a realização da parte experimental do meu trabalho, não apenas com o conhecimento, mas com a amizade, a alegria de trabalhar junto, a paciência, a compreensão, a disponibilidade. Serei sempre grata a todos!

À professora Dr<sup>a</sup> Ivonete Bakke, pelas valiosas contribuições, conselhos, ensinamentos e amizade.

Ao professor Dr. Olaf Bakke, pelos ensinamentos e pelas sementes cedidas, que foram muito importantes para a condução do experimento.

A todos os professores do PPGCF, por todo conhecimento e contribuições para minha formação acadêmica e pessoal.

Ao secretário do PPGCF, Paulo César, pela competência, disponibilidade e amizade.

À “Família Medeiros”, Edvania, Mylena, Jamilysson e Bean por todo carinho, apoio e amizade. Pela torcida!

Aos meus afilhados Hian Osmar e Fabiana Evilyn, por me ajudarem nas horas de estresse, por me trazerem tanta alegria! À minha prima Islanna, pela ajuda durante o meu afastamento e pelo carinho.

Aos amigos Pablo Forlan e Ana Cláudia, pela ajuda nos momentos de coleta e pela amizade sincera.

A todos a minha gratidão, pelas contribuições indispensáveis para o sucesso desse trabalho, da minha vida profissional e pessoal.

“E tudo quanto fizerdes, fazei-o de todo o coração, como ao Senhor, e não aos homens”.

(Apóstolo Paulo)

## LISTA DE FIGURAS

- Figura 1.** *Cnidoscopus quercifolius*. (A) Espécime; (B) Folhas e tricomas urticantes; (C) Flores; (D) Fruto. (Fonte: próprio autor) .....23
- Figura 2.** Representação esquemática das interações planta - estresses abióticos, percepção do estresse e ativação de cascatas metabólicas para gerar de respostas fisiológicas. Fonte: adaptado de Bezerra Neto (2016), com modificações. ....26
- Figura 3.** Estrutura do transporte de moléculas de água através dos canais de água (aquaporinas). Adaptada de Zhao et al. (2008) .....26
- Figura 4.** Modelo estrutural de aquaporina, demonstrando suas principais características. Alfa-hélices (retângulos azuis), nelas estão presentes domínios transmembrana (H1-H6), conectados pelos cinco loops (LA-LE). Dois domínios alfa-hélices (loops B e E), contendo os motivos NPA, responsáveis pela formação do poro. Adaptado de Bezerra Neto (2012). ....31
- Figura 5.** Estrutura e formação do poro das aquaporinas. Os retângulos em azul representam as alfa-hélices, onde estão presentes domínios transmembrana (H1-H6). Os cinco *loops* de conexão entre as alfa-hélices estão representados por LA-LE. No *loops* B e E, estão os motivos NPA, responsáveis pela formação do poro. Adaptado de Cheidde e Schor (1999) ...32
- Figura 6.** Diagrama esquemático da localização dos membros das subfamílias de aquaporinas de planta. Adaptado de Maeshima e Ishikawa (2008). ....33
- CAPÍTULO I - CARACTERIZAÇÃO DE AQUAPORINAS *IN SÍLICO* DE *Manihot esculenta* CRANTZ e *Ricinus communis* L. PARA O DESENHO DE INICIADORES E ISOLAMENTO DE SEQUÊNCIAS EM *Cnidoscopus quercifolius* POHL**
- Figura 1.** Análise fenética de membros das famílias de aquaporinas de *M. esculenta* (Fig. 7A) e de *R. communis* (Fig.7B), pelo método de *Neighbor Joining* (*bootstrap* ao nível de 1.000 réplicas). ....61
- Figura 2.** Análise fenética das aquaporinas preditas em *M. esculenta* (MeMIPs) e de *A. thaliana* (AtMIPs), por subfamília, separadamente. (A) MePIPs e AtPIPs. (B) MeTIPs e AtTIPs. (C) MeNIPs e AtNIPs. (D) MeSIPs e AtSIPs. (E) MeXIPs e LjXIPs. A escala de 0,05 refere-se a 5%, 0,02 a 2% e 0,1 a 10% de divergência entre as sequências. ....62
- Figura 3.** Análise fenética das aquaporinas preditas em *R. communis* (RcMIPs) e de *A. thaliana* (AtMIPs), por subfamília, separadamente. (A) RcPIPs e AtPIPs. (B) RcTIPs e AtTIPs. (C) RcNIPs e AtNIPs. (D) RcSIPs e AtSIPs. (E) RcXIPs e LjXIPs. ....63
- Figura 4.** Alinhamento múltiplo entre MePIPs e AtPIPs gerados com o auxílio da ferramenta *online* Clustal Omega. Destaque para os motivos conservados AEF e NPA. ....68
- Figura 5.** Relações fenéticas entre aquaporinas vegetais. Árvore construída pelo método de *Neighbour Joining*, mostrando a comparação entre as sequências de supostas aquaporinas presentes no genoma de *M. esculenta* e *R. communis* juntamente com aquaporinas de *A.*



*thaliana* e *L. japonicus*. As divisões entre as subfamílias encontram-se indicadas pelas diferentes cores. .... 71

**Figura 6.** Visualização da disposição dos genes de MeMIPs (*M. esculenta*) entre os 18 cromossomos de *M. esculenta*, de acordo com dados do Phytozome. Gráfico gerado pelo programa Circos. .... 72

**Figura 7.** Modelos tridimensionais das (A) MePIPs e (B) RcPIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína, com destaque para os motivos NPA. .... 74

**Figura 8.** Modelos tridimensionais das (A) MeTIPs e (B) RcTIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína, com destaque para os motivos NPA. .... 75

**Figura 9.** Modelos tridimensionais das (A) MeNIPs e (B) RcNIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína, com destaque para os motivos NPA. .... 76

**Figura 10.** Modelos tridimensionais das (A) MeSIPs e (B) RcSIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína, com destaque para os motivos NPA. .... 77

**Figura 11.** Modelos tridimensionais das (A) MeXIPs e (B) RcXIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína, com destaque para os motivos NPA. .... 78

**Figura 12.** Gráficos representativos do Z-score dos modelos gerados para MePIP e RcPIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB. .... 79

**Figura 13.** Gráficos representativos do Z-score dos modelos gerados para MeTIP e RcTIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB. .... 80

**Figura 14.** Gráficos representativos do Z-score dos modelos gerados para MeNIP e RcNIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB. .... 80

**Figura 15.** Gráficos representativos do Z-score dos modelos gerados para MeSIP e RcSIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB. .... 81

**Figura 16.** Gráficos representativos do Z-score dos modelos gerados para MeXIP e RcXIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB. .... 81

**Figura 17.** Análise de confiabilidade do modelo gerado para MePIP via Ramachandran Plot. .... 82

<b>Figura 18.</b> Análise de confiabilidade do modelo gerado para MeTIP via Ramachandran Plot. .....	82
<b>Figura 19.</b> Análise de confiabilidade do modelo gerado para MeNIP via Ramachandran Plot. .....	83
<b>Figura 20.</b> Análise de confiabilidade do modelo gerado para MeSIP via Ramachandran Plot. .....	83
<b>Figura 21.</b> Análise de confiabilidade do modelo gerado para MeXIP via Ramachandran Plot. .....	84
<b>Figura 22.</b> Análise de confiabilidade do modelo gerado para RcPIP via Ramachandran Plot. .....	84
<b>Figura 23.</b> Análise de confiabilidade do modelo gerado para RcTIP via Ramachandran Plot. .....	85
<b>Figura 24.</b> Análise de confiabilidade do modelo gerado para RcNIP via Ramachandran Plot. .....	85
<b>Figura 25.</b> Análise de confiabilidade do modelo gerado para RcSIP via Ramachandran Plot. .....	86
<b>Figura 26.</b> Análise de confiabilidade do modelo gerado para RcXIP via Ramachandran Plot .....	86
<b>Figura 27.</b> Resultado de testes de amplificação em DNA de <i>C. quercifolius</i> . (A) Utilizando iniciadores para aquaporinas desenhados com base em sequências genômicas de <i>M. esculenta</i> (Me) e <i>R. communis</i> (Rc), (B) Utilizando iniciadores para aquaporinas desenhados com base em sequências expressas de <i>V. unguiculata</i> . Usando como padrão de massa molecular DNA Ladder de 100 pb .....	88
<b>Figura 28.</b> Dendrograma de membros das famílias de aquaporinas de <i>C. quercifolius</i> juntamente com <i>A. thaliana</i> pelo método de <i>Neighbor Joining</i> ( <i>bootstrap</i> ao nível de 1.000 replicas). .....	90

## LISTA DE TABELAS

<b>Tabela 1.</b> Caracterização de genes de AQPs em <i>M. esculenta</i> (43 sequências) e de <i>R. communis</i> (36 sequências).....	65
<b>Tabela 2.</b> Sumário das principais regiões e sítios conservados de aquaporinas em <i>Manihot esculenta</i> .....	68
<b>Tabela 3.</b> Sumário das principais regiões e sítios conservados de aquaporinas em <i>Ricinus communis</i> (RcAQP).....	70
<b>Tabela 4.</b> Valores de similaridade entre o modelos utilizados para modelagem e os diferentes membros de aquaporinas de <i>M. esculenta</i> e <i>R. communis</i> .....	73
<b>Tabela 5.</b> Valores de qualidade e energia para os modelos estruturais de aquaporinas de <i>M. esculenta</i> e <i>R. communis</i> obtidos via plataforma SWISS-MODEL. (Valores mais próximos a um representam modelos com maior qualidade).....	79
<b>Tabela 6.</b> Lista de iniciadores que apresentaram resultados de amplificação no processo de transferabilidade entre <i>M. esculenta</i> , <i>R. communis</i> e <i>V. unguiculata</i> em <i>C. quercifolius</i> . .....	87
<b>Tabela 7.</b> Resultado de alinhamento simples realizado contra bando de dados não redundantes utilizando como <i>query</i> sequências de <i>C. quercifolius</i> . .....	89
<b>Tabela 8.</b> Identificação e descrição de domínio conservado em aquaporinas de <i>C. quercifolius</i> . .....	91

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

<b>Sigla</b>	<b>Definição</b>
aa	Aminoácidos
ar/R	Região arginina/aromática
BLAST	<i>Basic Local Alignment Search Tool</i> Ferramenta de Busca por Alinhamento Local
CD	<i>Conserved Domain</i> Domínio Conservado
cDNA	DNA complementar
DDBJ	<i>DNA Database of Japan</i> Banco de Dados de DNA do Japão
DNA	Ácido desoxirribonucleico
EMBL	<i>European Molecular Biology Laboratory</i> Laboratório Europeu de Biologia Molecular
<i>GenBank</i>	Banco de Genes - banco de dados norte-americano
GIP	Homóloga aos canais bacterianos de glicerol
INSDB	<i>International Nucleotide Sequence Database</i> Banco de Dados Internacional de Sequências de Nucleotídeos
MEGA	<i>Molecular Evolutionary Genetic Analysis</i> Análises Genéticas de Evolução Molecular
MIP	<i>Major Intrinsic Protein</i> Proteínas Intrínsecas de Membrana
MM	Massa molecular
NCBI	<i>National Center for Biotechnology Information</i> Centro Nacional para Informação Biotecnológica
NIP	<i>Nodulin26-like Intrinsic Proteins</i> Proteínas intrínsecas do nódulo
NJ	<i>Neighbor-Joining</i> Agrupamento por Vizinhança
<i>ORF-Finder</i>	<i>Open Reading Frame Finder</i> Identificador de Pautas Abertas de Leitura

Pb	Pares de bases em nucleotídeos
PCR	<i>Polymerase Chain Reaction</i> Reação em cadeia da polimerase
PDB	<i>Protein Data Bank</i> Banco de Dados de Proteínas
PEG	Polietilenoglicol
pI	Ponto isoelétrico
PIP	<i>Plasma Membrane Intrinsic Protein</i> Proteínas intrínsecas da membrana plasmática
RNA	Ácido ribonucleico
SIP	<i>Small Basic Intrinsic Proteins</i> Pequenas proteínas básicas intrínsecas de membranas
TIP	<i>Tonoplast Intrinsic Proteins</i> Proteínas de membrana de tonoplasto
UPGMA	<i>Unweighted Pair Group Method with Arithmetic Means</i> Método não polarizado de Agrupamentos aos Pares com Médias Aritméticas
XIP	<i>Uncharacterized X Intrinsic Protein</i> Proteína intrínseca desconhecida

## SUMÁRIO

1 INTRODUÇÃO.....	17
2. OBJETIVOS.....	19
2.1 OBJETIVO GERAL.....	19
2.2 OBJETIVOS ESPECÍFICOS.....	19
3 REFERENCIAL TEÓRICO.....	20
3.1 O bioma Caatinga como fonte de recursos genéticos vegetais.....	20
3.1.1 Importância de membros da família Euphorbiaceae.....	21
3.2. Relações hídricas e seus efeitos sobre as plantas.....	24
3.3 Aquaporinas vegetais: função, diversidade e classificação.....	26
3.3.1 Subfamílias de aquaporinas em plantas superiores.....	27
3.3.2 Estrutura e localização celular das aquaporinas.....	31
3.4 Avanço das “ômicas” e ferramentas computacionais na prospecção e análise de genes...34	
3.4.1 Bancos de dados biológicos.....	34
3.4.2 Ferramentas computacionais para caracterização e anotação de sequências.....	36
4 REFERÊNCIAS.....	38
CAPÍTULO 1	
CARACTERIZAÇÃO DE AQUAPORINAS <i>in silico</i> DE <i>Manihot esculenta</i> CRANTZ E <i>Ricinus communis</i> L. PARA O DESENHO DE INICIADORES E ISOLAMENTO DE SEQUÊNCIAS EM <i>Cnidoscolus quercifolius</i> POHL	
1 INTRODUÇÃO.....	53
2. MATERIAIS E MÉTODOS.....	55
2.1 Identificação e caracterização de MeAQPs e RcAQPs.....	55
2.1.1 Seleção de sondas e busca por aquaporinas homólogas em <i>M. esculenta</i> e <i>R. communis</i> .....	55
2.1.2 Caracterização de sequências.....	55
2.1.3 Alinhamento de sequências e análise fenética.....	56
2.1.4 Distribuição cromossômica e localização subcelular.....	56
2.1.5 Modelagem comparativa.....	56
2.2 Obtenção e caracterização de aquaporinas em <i>C. quercifolius</i> – CqAQPs.....	57
2.2.1 Desenho de iniciadores.....	57
2.2.2 Material vegetal.....	58
2.2.3 Extração de DNA.....	58

2.2.4 Extração de RNA e síntese de cDNA .....	58
2.2.5 Reação em Cadeia da Polimerase – PCR .....	59
2.2.6 Confirmação das <i>CqAQP</i> s via sequenciamento genético .....	59
2.2.7 Obtenção e análise de bioinformática das sequências <i>CqAQP</i> .....	59
3 RESULTADOS .....	60
3.1 Busca e classificação das MeMIPs e RcMIPs .....	60
3.1.1 Caracterização e localização subcelular .....	65
3.1.2 Motivos e domínios conservados .....	67
3.1.3 Análise fenética e distribuição cromossômica.....	71
3.1.4 Modelagem comparativa .....	73
3.2 Transferabilidade de iniciadores em <i>C. quercifolius</i> .....	87
4 DISCUSSÃO .....	91
4.1 Identificação e classificação de MeAQPs e RcAQPs.....	91
4.1.1 Análise comparativa e fenética das subfamílias de MeAQPs e RcAQPs.....	93
4.1.2 Caracterização de sequências de MeAQPs e RcAQps .....	94
4.1.3 Motivos e domínios conservados .....	94
4.1.4 Análise fenética e distribuição cromossômica.....	95
4.1.5 Modelagem comparativa .....	96
4.2 Transferabilidade de iniciadores em <i>C. quercifolius</i> .....	96
5 CONSIDERAÇÕES FINAIS .....	97
6 CONCLUSÃO.....	98
7 REFERÊNCIAS .....	99

Nóbrega, Gisela Formiga Queiroz. **IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE AQUAPORINAS EM *Cnidoscolus quercifolius* Pohl**. Dissertação de Mestrado em Ciências Florestais. CSTR/UFCG, Patos - PB. 2017. 104p.

## RESUMO

A família botânica Euphorbiaceae Juss. é uma das mais diversificadas e complexas, apresentando cerca de 50 gêneros e 240 espécies no Nordeste brasileiro, grande parte distribuídos em áreas de Caatinga. Dentre essas, destacam-se *Manihot esculenta* Cartz, *Ricinus communis* L. e *Cnidoscolus quercifolius* Pohl, que, devido à tolerância ao déficit hídrico, apresentam-se como espécies de importância econômica para regiões tropicais, especialmente o Brasil. *C. quercifolius* é uma espécie endêmica da Caatinga, sendo bastante explorada devido às várias possibilidades de utilização de suas partes (produção de forragem, óleo comestível e farinha, utilizada na regeneração de áreas degradadas e para fins medicinais). Entre os mecanismos de adaptação às condições ambientais limitantes existentes em plantas superiores, ressalta-se a importância do reajuste dos padrões de expressão gênica, que atua ativando e desativando uma variedade de genes, considerados genes estresse-responsivos, que são responsáveis pelas alterações do metabolismo celular que permite à planta a geração de respostas fisiológicas frente à exposição ao estresse. Um exemplo são as aquaporinas vegetais, que são proteínas de membrana responsáveis pelo transporte de água, gases e outros pequenos solutos e que desempenham um papel fundamental na adaptação das plantas mediante estresses como salinidade e déficit hídrico. A identificação de genes estresse-responsivos só foi possível devido às novas tecnologias de sequenciamento e às ferramentas de bioinformática, que possibilitaram, não apenas a obtenção de genomas completos, como também a manipulação e análise desses dados, facilitando a identificação de novas moléculas envolvidas em processos fisiológicos essenciais às plantas. Até o momento, não se encontra disponível em bancos de dados biológicos qualquer tipo de informação genética de *C. quercifolius*. A identificação de aquaporinas em *C. quercifolius* foi realizada pelo processo de transferibilidade de iniciadores entre espécies próximas, tendo em vista a elevada conservação entre as sequências desta família gênica. Foram identificadas sete sequências parciais de CqAQPs, sendo este o primeiro relato de aquaporinas nesta espécie, podendo, assim, ser usado para o desenho de iniciadores específicos para a espécie, favorecendo a identificação de potenciais alvos biotecnológicos.

**PALAVRAS-CHAVE:** Bioinformática, Euforbiáceas, Transportadores de membrana.



Nóbrega, Gisela Formiga Queiroz. **IDENTIFICATION AND CHARACTERIZATION OF AQUAPORINES GENES IN *Cnidoscolus quercifolius* Pohl.** – Master's thesis in Forest Science. CSTR/UFCG, Patos - PB. 2017. 104p.

### ABSTRACT

Euphorbiaceae Juss. is one of the most diversified and complex species of the botanical family, comprising approximately 50 genera and 240 kinds in the Brazilian Northeast, widely distributed in Caatinga areas. Among them are *Manihot esculenta* Cartz, *Ricinus communis* L. and *Cnidoscolus quercifolius* Pohl, which, due to its tolerance to water shortage, are species of great economic importance for tropical regions, including Brazil. *C. quercifolius* is an endemic species of the Caatinga, and it has been extensively exploited for various uses (forage production, edible oil and flour, use in the regeneration of degraded areas and for medicinal purposes). Among the mechanisms of adaptation to the limiting environmental conditions existing in higher plants, we emphasize the importance of the readjustment of genetic gender patterns, which activates and deactivates a variety of genes that are considered genes-responsive genes responsible for cellular metabolism, which allows a generation of physiological responses to a plant from stress. An example of that are the aquaporins plants, which are membrane proteins responsible for transporting of water, gases and other small solutes, in addition to playing a key role in the adaptation of plants to adverse conditions such as salinity and water shortage. The identification of stress-responsive genes was only possible from new sequencing technologies and bioinformatics tools, which enabled not only the attainment of a complete genome, but also the manipulation and analysis of such data, facilitating the identification of new molecules involved in essential physiological processes to plants. To date, it is not available in biological databases any genetic information on *C. quercifolius*. The identification of aquaporins in *C. quercifolius* was conducted by the process of transferability of initiators between close species, considering the high conservation among the sequences of this family. Seven partial sequences of CQAQPs were identified, and this is the first report of aquaporins in that species; therefore, enabling its use for designing of specific initiators for the species, aiding the identification of potential biotechnological targets.

**KEYWORDS:** Bioinformatics, Euphorbiaceae, Membrane transporters.

## 1 INTRODUÇÃO

As plantas são frequentemente submetidas à situação de estresse por fatores bióticos e/ou abióticos, presentes no meio em que estão inseridas. Em geral, elas respondem a esses estresses, e essas respostas provocam danos ao crescimento e desenvolvimento vegetal. A capacidade que a planta possui de reagir a determinadas condições depende de um conjunto de fatores, entre os quais, é possível ressaltar a importância do patrimônio genético da espécie e o estágio de desenvolvimento no qual ela se encontra (WANG et al., 2001; BENKO-ISEPPON et al., 2011).

Dentre os principais fatores que limitam o crescimento vegetal, destacam-se o déficit hídrico, a salinidade e a escassez de nutrientes no solo, inviabilizando a realização de atividades importantes, comprometendo o desenvolvimento e produtividade da planta (NEPOMUCENO et al., 2001). O déficit hídrico, além de impedir a realização de processos indispensáveis como a fotossíntese, interfere na disponibilidade de nutrientes e aumento da salinidade do solo, favorecendo a ocorrência de outros estresses simultaneamente (MAHAJAN; TUTEJA, 2005).

A vegetação da Caatinga destaca-se como uma importante candidata para estudos genéticos e prospecção de genes estresse-responsivos, tendo em vista a sua evidente adaptação morfológica e fisiológica, determinantes na sobrevivência à baixa precipitação pluviométrica e demais características climáticas dessa região. Além disso, muitos dos mecanismos adaptativos presentes nesses vegetais ainda permanecem desconhecidos, cuja escassez de informações é ainda maior em espécies endêmicas do bioma (SAMPAIO, 2010).

Neste contexto, torna-se de grande importância compreender quais mecanismos biológicos estão envolvidos no desenvolvimento da capacidade adaptativa dessas plantas às condições ambientais, em especial, aqueles que participam dos processos de captação e aproveitamento da água pela planta, já que o déficit hídrico é uma constante na Caatinga.

Entre os membros de plantas representantes da Caatinga, destaca-se o gênero *Cnidoscolus* Pohl. Das aproximadas 70 espécies descritas para o referido gênero, cinco espécies: *C. quercifolius*, *C. obtusifolius*, *C. vitifolius*, *C. bahianus* e *C. urnigerus* foram registradas apenas na Caatinga (MELO; SALES, 2008). A espécie *C. quercifolius* Pohl é vista como uma importante candidata a ser explorada em análises *in silico* e moleculares, por ser uma planta nativa da Caatinga e extremamente tolerante ao déficit hídrico.

Através de ferramentas de bioinformática, juntamente com as novas tecnologias de sequenciamento atualmente disponíveis, é possível a obtenção de genomas completos, além

de uma grande variedade de sequências expressas, as quais nos permitem prever padrões de sequências biológicas, facilitando a identificação de novas moléculas envolvidas em processos fisiológicos essenciais às plantas, favorecendo a identificação de potenciais alvos biotecnológicos.

As aquaporinas vegetais, proteínas de membrana, apresentam papel fundamental na adaptação das plantas frente a estresses como salinidade e déficit hídrico, uma vez que facilitam o transporte de água e outros solutos importantes para o metabolismo celular (WALLACE et al., 2006; HEINEN; CHAUMONT, 2009; WUDICK et al., 2009;). Considerando que até o momento não existe qualquer tipo de informação de sequências de *C. quercifolius* disponível em banco de dados biológicos, este estudo foi conduzido com base em trabalhos desenvolvidos com outras espécies, tendo em vista a elevada conservação entre as sequências de aquaporinas, como descrito anteriormente por Chaumont et al. (2001), Deshmukh et al. (2016) e Zou et al. (2016).

Neste contexto, o estudo aqui desenvolvido envolveu a identificação e caracterização de genes de aquaporinas em euforbiáceas, com vistas ao isolamento e caracterização de sequências homólogas em *Cnidoscolus quercifolius*, uma planta nativa do Bioma Caatinga, extremamente tolerante a seca. As informações obtidas nesta pesquisa poderão ser utilizadas no melhoramento de plantas, utilizando as informações genéticas de aquaporinas de espécie endêmica da Caatinga.

## 2. OBJETIVOS

### 2.1 OBJETIVO GERAL

Identificar e caracterizar genes de aquaporinas em *Cnidoscolus quercifolius* Pohl.

### 2.2 OBJETIVOS ESPECÍFICOS

- Realizar a prospecção *in silico* de genes codificantes para aquaporinas vegetais em bancos de dados públicos, usando sequências “sondas” em representantes mais próximos da espécie em estudo;
- Localizar representantes de aquaporinas na espécie estudada;
- Realizar análises de bioinformática das sequências obtidas, reconhecendo seus domínios conservados;
- Efetuar alinhamentos múltiplos, inferindo sobre a estrutura e as relações do citado gene com sequências de nucleotídeos e de proteínas;
- Realizar modelagem comparativa de sequências para os diferentes membros identificados, visando à obtenção de modelos tridimensionais, viáveis para estudos relacionados à especificidade e eficiência de transporte.
- Com base nas sequências obtidas de aquaporinas, realizar caracterização *in silico* para desenho de iniciadores para isolamento de genes homólogos de aquaporinas em *Cnidoscolus quercifolius*.

### 3 REFERENCIAL TEÓRICO

#### 3.1 O bioma Caatinga como fonte de recursos genéticos vegetais

O domínio Caatinga é exclusivamente brasileiro, eleito como um dos mais importantes do Brasil, com extensão territorial de aproximadamente 844.453 Km<sup>2</sup>, ocupa cerca de 50 % do Nordeste Brasileiro, o que equivale aproximadamente a 10% do território nacional (IBGE, 2014). Primeiramente, a Caatinga foi caracterizada como um bioma pobre e de baixa biodiversidade, cuja potencialidade permaneceu explorada de forma desordenada. Assim, cerca de 80 % dos ecossistemas naturais foram alterados, principalmente pela prática de desmatamentos e queimadas (MMA, 2010). Hoje se sabe que além de uma vasta biodiversidade, a Caatinga também possui alto índice de endemismo (GIULIETTI; CONCEIÇÃO; QUEIROZ, 2006; MACIEL, 2010).

A Caatinga está submetida a intensa luminosidade e temperaturas elevadas, com pouca variação, mesmo entre os meses mais frios e mais quentes (médias anuais entre 25 - 30° C). O baixo índice pluviométrico (entre 250 e 700 mm/ano) (SAMPAIO, 2003; 2010) e a predominância de solos rasos e pedregosos, com baixa capacidade de armazenamento de água e escassez de nutrientes (LEAL et al, 2003), constituem os principais fatores limitantes para o crescimento vegetal, caracterizando a semiaridez climática predominante na Caatinga (relação precipitação/evapotranspiração potencial < 0,65) (SAMPAIO, 2010).

A estrutura da vegetação é o reflexo das variações das condições ambientais associadas ao antropismo. Em regiões onde há presença de corpos de água, ou locais alagados, por exemplo, as vegetações não apresentam características de adaptação à aridez (FRANÇA et al., 2003). Embora estas áreas sejam relativamente pequenas (considerando em escala regional), apresentam grande importância, principalmente para o aumento da diversidade de vegetação (GIULIETTI; CONCEIÇÃO; QUEIROZ, 2006).

A Caatinga é representada, principalmente, por árvores e arbustos de pequeno porte, já que o desenvolvimento de maior parte da sua vegetação é limitado pelas condições climáticas menos favoráveis. A maior parte da vegetação apresenta adaptações como microfilia, caducifolia e controle da atividade estomática, que permite um melhor aproveitamento da água, favorecendo sobrevivência do vegetal (SANTOS; RIBEIRO; SAMPAIO, 1992). As espécies predominantes são representadas por indivíduos das famílias Fabaceae, Cactaceae e Euforbiaceae (GOLFARI; CASER, 1977).

Diante do exposto, a cobertura vegetal da Caatinga representa uma importante fonte de genes com alto potencial biotecnológico, principalmente aqueles envolvidos direta e/ou indiretamente nos processos de captação e armazenamento de água, que permitem a adaptação das plantas às condições de restrição hídrica.

### 3.1.1 Importância de membros da família Euphorbiaceae

A família Euphorbiaceae Juss. é uma das mais diversificadas e complexas entre as Angiospermas (WEBSTER, 1994), reunindo mais de 8.000 espécies, distribuídas em 317 gêneros (ENDRESS et al., 2013; BIJEKAR; GAYATRI, 2014). No Brasil, estima-se a ocorrência de 1.100 espécies e 72 gêneros (SOUZA; LORENZI, 2008). Apresenta variedade de hábitos, o que confere a esta família maior capacidade adaptativa e melhor eficiência do uso da água (OLIVEIRA, 2013), favorecendo amplamente sua distribuição em regiões tropicais e subtropicais, principalmente nos continentes americano e africano (LUCENA; AMORIN; ALVES, 2009).

O Nordeste Brasileiro possui alta diversidade de euporbiaceas (CORDEIRO; CARNEIRO-TORRES, 2006), com cerca de 240 espécies e 50 gêneros (LUCENA; ALVES, 2010), sendo a maioria distribuída em áreas de Caatinga, dentre as quais encontra-se um grande número de espécies endêmicas (taxa de endemismo de Euphorbiaceae: cerca de 60 %) (SÁTIRO; ROQUE, 2008; FORZZA et al., 2010).

As Euphorbiaceae estão entre as famílias de maior importância econômica entre as Eudicotiledôneas, principalmente nos setores industrial, madeireiro, farmacológico, ornamental e na produção alimentícia (ALVES, 1998). Entre as espécies utilizadas na produção de alimentação humana, principalmente na região Nordeste do Brasil, destaca-se a *Manihot esculenta* Crantz (*M. esculenta*) (CORRÊA et al., 2002). No âmbito industrial ressalta-se o potencial econômico da *R. communis* (*Ricinus communis* L.) (OGUNNIYI, 2006). Já, entre as espécies endêmicas utilizadas para fins medicinais e produção de forragem animal, está a *Cnidoscolus quercifolius* Pohl (*C. quercifolius*) (SOUZA, 2012; PEIXOTO SOBRINHO et al., 2012).

A *M. esculenta* (*M. esculenta* Crantz) representa a terceira espécie botânica de maior fonte de calorias (depois do arroz e do milho) para mais de um bilhão de pessoas em todo o mundo (CEBALLOS et al., 2010). Apresenta alta capacidade de desenvolvimento sob condições ambientais desfavoráveis, como regime irregular de chuvas e baixa fertilidade do solo (JARVIS et al., 2012), sendo cultivada principalmente nas regiões tropicais e

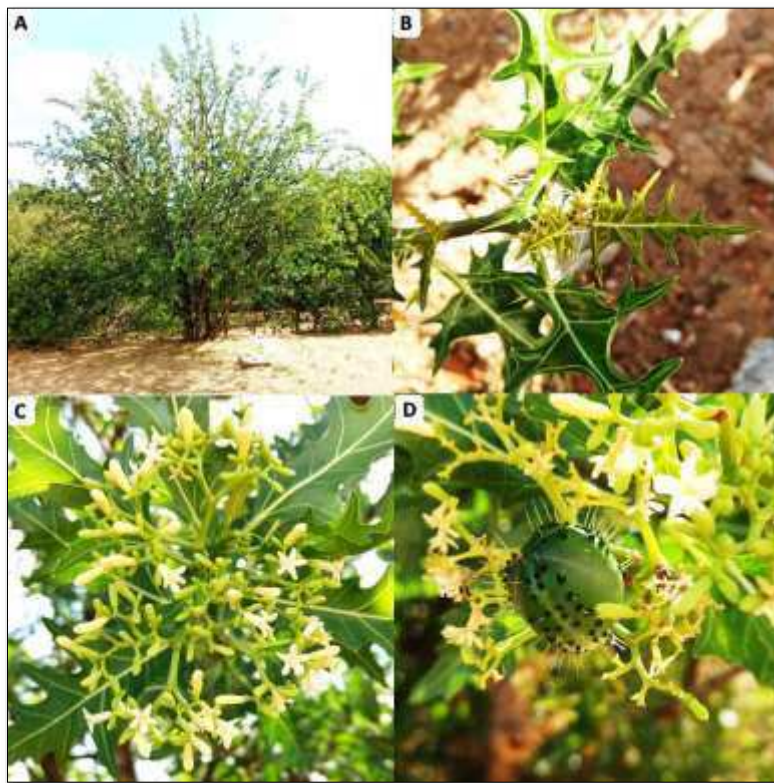
subtropicais da África, Ásia e América Latina, onde é considerada uma cultura de grande importância socioeconômica (CEBALLOS et al., 2010).

De acordo com a Organização das Nações Unidas para a Agricultura e Alimentação (FAO), a produção mundial de *M. esculenta*, nos últimos cinco anos, apresentou um crescimento médio de 13,9%, atingindo 275 milhões de toneladas no ano de 2013 (SEAB-DERL, 2016). Em 2016, o Brasil atingiu uma produção de cerca de 23 milhões de toneladas, ocupando o 2º lugar no *ranking* mundial, ficando atrás da Nigéria (38,57 milhões de toneladas) (IBGE, 2016). Das 76 espécies do gênero *Manihot* existentes no Brasil, 66 são endêmicas (CORDEIRO et al., 2015), correspondendo a mais de 70% da diversidade conhecida no mundo (ORLANDIN; LIMA, 2014), ou seja, grande parte da diversidade genética atualmente utilizada em todo o mundo vem do Brasil (OLIVEIRA et al., 2015).

Outra euforbiácea de relevante importância socioeconômica é a *Ricinus communis* L., uma espécie oleaginosa singular, de considerável valor energético e de ampla utilização industrial (OGUNNIYI, 2006; ZENG, 2010). Extraído das sementes de *R. communis*, o “óleo de rícino” é uma importante matéria-prima utilizada para fins industriais, medicinais e cosméticos (OGUNNIYI, 2006). Embora originária da África, a *R. communis* é agora cultivada em muitas regiões temperadas tropicais, subtropicais ao redor do mundo, especialmente sob condições de solos pobres, pouca disponibilidade de água e salinidade (RIVAROLA, 2011). De acordo com dados da FAO, a Índia é o maior produtor mundial de *R. communis*, atingindo no ano de 2013, mais de 1,6 milhão de toneladas (CONAB, 2016). Atualmente, o Brasil ocupa o 5º lugar entre os maiores produtores da *R. communis*, perdendo somente para Índia, China, Paquistão e Tailândia. Em termos de óleo de *R. communis*, o Brasil, juntamente com a Índia e a China, constituem os três maiores produtores mundiais. No Brasil, a produção de *R. communis* concentra-se na região Nordeste, principalmente no Semiárido (FAOSTAT, 2013; EMBRAPA, 2012).

Popularmente conhecida como “faveleira, faveleiro ou favela”, a espécie *C. quercifolius* é classificada como arbusto, arvoreta ou árvore, atingindo de 2 a 12 m de altura, em função do local e das condições ambientais (MELO; SALES, 2008). Apresenta-se distribuída por todo o semiárido da região da Caatinga, o que inclui os Estados do Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Piauí, e Bahia (OLIVEIRA, et al., 2008). A maioria dos indivíduos apresentam tricomas urticantes aciculiformes, recobrendo estruturas como ramos, pecíolo, lâmina foliar, perianto e frutos, sendo esta uma característica marcante da espécie (MELO; SALES, 2008; CAVALCANTI et al., 2011) (Figura 1).

Figura 1. *Cnidoscolus quercifolius*. (A) Espécime; (B) Folhas e tricomas urticantes; (C) Flores; (D) Fruto. (Fonte: próprio autor)



A *C. quercifolius* é uma espécie bastante explorada devido às várias possibilidades de utilização de suas partes, contribuindo para o sustento da população e sobrevivência dos animais. Destaca-se como uma das forrageiras de grande potencial da região, sendo considerada uma opção sustentável de suplementação alimentar para os animais devido aos altos valores de proteína bruta (SOUZA, 2012). Suas sementes são empregadas na produção de óleo comestível e farinha, sendo utilizadas na alimentação de animais e humanos (CAMPOS, 2010; ARRIEL et al., 2006). A casca do caule da *C. quercifolius* é popularmente usada para fins medicinais, no tratamento de processos inflamatórios da próstata e ovários (AGRE et al., 2006). O potencial antibacteriano contra *Staphylococcus* torna a *C. quercifolius* uma excelente candidata na identificação de novas drogas com atividade antimicrobiana (PEIXOTO SOBRINHO et al., 2012).

Uma das grandes propriedades da *C. quercifolius* é a capacidade de adaptação a ambientes de clima semiárido, suportando longos períodos de estiagem (ALOUFA; MEDEIROS, 2016). Devido a esta característica, ela é considerada uma planta fundamental para o equilíbrio da Caatinga, além de ser utilizada, também, no processo de recuperação de áreas degradadas (ARRIEL et al., 2006; MEDEIROS; ALOUFA, 2015). A presença de características xerófitas sugere a existência de estratégias de captação e aproveitamento de



água por parte da planta. Isso ressalta a importância do estudo de moléculas envolvidas na captura e transporte seletivo de água e íons através das membranas (GASPAR, 2011; HUANG et al., 2012).

### **3.2. Relações hídricas e seus efeitos sobre as plantas**

Fatores de estresses abióticos atuam sobre a planta provocando danos, muitas vezes irreversíveis, inviabilizando a realização de atividades fisiológicas indispensáveis para que o vegetal alcance seu potencial genético (MAHAJAN; TUTEJA, 2005). No ambiente, estes fatores podem atuar de forma isolada ou simultaneamente, expondo as plantas a uma combinação de estresses, que agem de forma diferente sobre os organismos vegetais (BRAY, 2004). Entre os fatores abióticos, o déficit hídrico, altas temperaturas, elevada incidência luminosa, a salinidade do solo e a escassez de nutrientes estão entre os principais estresses, aos quais a vegetação da Caatinga é constantemente exposta (SAMPAIO, 2010).

Devido à importância que exerce na sobrevivência dos seres vivos, a água destaca-se como recurso fundamental na realização da maioria dos processos fisiológicos da planta (KERBAUY, 2004), já que o estresse osmótico diminui a atividade fotossintética, interferindo no crescimento e produtividade do vegetal. A falta de água interfere também na disponibilidade de nutrientes e aumento da salinidade do solo, provocando a ocorrência de outros estresses simultaneamente (ZHU, 2002; MAHAJAN; TUTEJA, 2005).

A quantidade de água disponível para a realização das reações metabólicas nos vegetais é o resultado do balanço hídrico entre a absorção, condução e a perda de água (VINNAKOTA et al., 2015). Normalmente, a redução da perda de água ocorre pelo controle da transpiração, processo pelo qual a maior parte da água presente na planta é perdida para o ambiente na forma de vapor. A transpiração é necessária para manter o resfriamento das folhas, por isto algumas plantas de clima seco perdem as folhas em períodos de seca, evitando, assim, a desidratação (NOGUEIRA et al., 2005; VINNAKOTA et al., 2015).

A conservação e o manejo da água em plantas, especialmente durante períodos de seca, são críticos para a sobrevivência do vegetal. O déficit hídrico compromete o metabolismo minimizando (ou até mesmo bloqueando) processos como a fotossíntese e a hidrólise de amido. A redução destas atividades inibe o desenvolvimento de todo o organismo, impõe forte pressão seletiva e força a planta a desenvolver adaptações que possibilitem suportar as mais diversas condições ambientais (NEPOMUCENO et al., 2001).

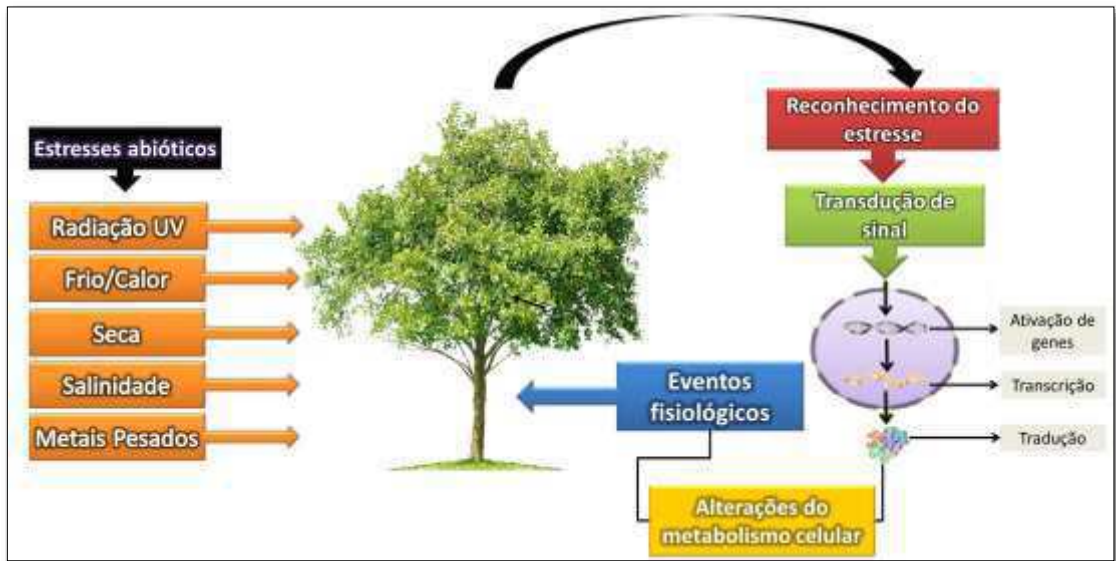
A absorção de água pela planta ocorre principalmente pelo sistema radicular. Desta forma, a captação depende, tanto da disponibilidade de água do ambiente, temperatura e concentração de solutos no solo, como também da eficiência da raiz em realizar tal processo. Neste contexto, destaca-se a habilidade de transportadores de água presentes nas membranas celulares (as proteínas intrínsecas das membranas - MIPs), já que a entrada de água, bem como a condução da mesma entre os tecidos vegetais, ocorre pelo processo de difusão (RAVEN; EVERT; EICHHORN, 2007; VINNAKOTA et al., 2015).

Os efeitos causados pelo déficit hídrico em plantas dependem diretamente do estágio de desenvolvimento no qual se encontra o vegetal, bem como da intensidade e velocidade de aplicação do estresse sobre o mesmo (BENKO-ISEPPON et al., 2011). Como forma de adaptação, as plantas desenvolveram mecanismos específicos que lhes permitem tolerar tais condições de estresse, por meio de modificações bioquímicas, fisiológicas, morfológicas, genéticas e metabólicas (WANG et al., 2001), na tentativa de minimizar danos e conservar recursos indispensáveis ao seu crescimento e reprodução (MITTLER; BLUMWALD, 2010; BENKO-ISEPPON et al., 2011).

Entre os principais mecanismos de adaptação às condições ambientais limitantes existentes em planta superiores, destaca-se o reajuste dos padrões de expressão gênica (BARTELS; SUNKAR, 2005). Os vários tipos de estresses aos quais as plantas são constantemente submetidas induzem uma variedade de genes (BENKO-ISEPPON et al., 2011). A partir do reconhecimento das condições ambientais desfavoráveis, produtos de genes induzidos por estresse ativam vias de transdução de sinais, que transmitem as informações levando a alterações de processos fisiológicos e metabólicos, em resposta ao estresse (NAKASHIMA; YAMAGUCHI-SHINOZAKI, 2009), visando minimizar danos e conservar recursos valiosos para seu crescimento e reprodução (MITTLER; BLUMWALD, 2010) (Figura 2).

As respostas das plantas ao estresse abiótico são complexas e envolvem vários mecanismos regulados pela “*crosstalk*” entre genes relacionados à sinalização hormonal, fotossíntese, respiração e regulamentações transcricionais (CHEN et al., 2016). Estudos de genômica funcional associados aos vários projetos de sequenciamento e perfis de expressão gênica têm possibilitado a prospecção de genes candidatos responsivos a estresses (SEKI et al., 2002). Entre os vários selecionados, encontram-se os genes envolvidos na captura e transporte seletivo de água e íons como as aquaporinas (BENKO-ISEPPON et al., 2011; GASPAR, 2011; HUANG et al., 2012).

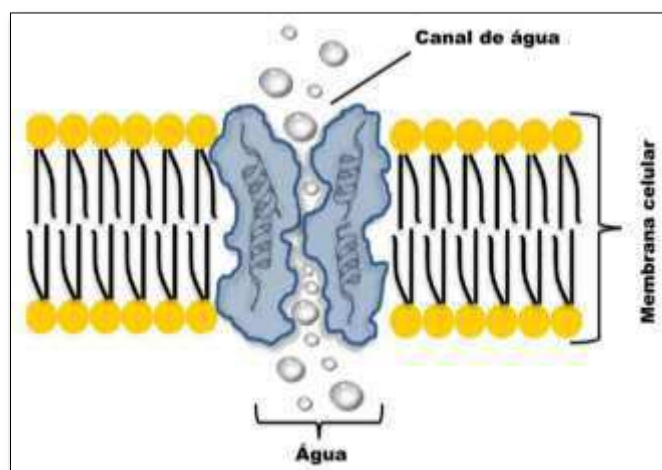
Figura 2. Representação esquemática das interações planta - estresses abióticos, percepção do estresse e ativação de cascatas metabólicas para gerar de respostas fisiológicas. Fonte: adaptado de Bezerra Neto (2016), com modificações.



### 3.3 Aquaporinas vegetais: função, diversidade e classificação

A atividade de distribuição e regulação do volume de água nos tecidos vegetais depende diretamente da capacidade de captação e transporte que as células apresentam. Estes processos são responsáveis pela manutenção do metabolismo vegetal e, conseqüentemente, interferem no desenvolvimento do organismo (VINNAKOTA et al., 2015). O transporte executado pelas membranas celulares é um tipo de mecanismo biofísico que depende da presença de estruturas especializadas e de proteínas de membrana a exemplo de aquaporinas – proteínas conhecidas como “canais de água” (MAUREL et al., 2008) (Figura 3).

Figura 3. Estrutura do transporte de moléculas de água através dos canais de água (aquaporinas). Adaptada de Zhao et al. (2008)



As aquaporinas (AQPs) pertencem a uma superfamília altamente conservada de Proteínas Intrínsecas de Membrana (MIP - *Major Intrinsic Protein*) (AGRE, 2004), presentes em todos os organismos (GOMES et al., 2009; ABASCAL; IRISARRI; ZARDOYA, 2014). Embora originalmente descritas como canal de água, estudos têm reforçado o importante papel das aquaporinas como facilitadoras no transporte de gases, íons e outros pequenos solutos (MAUREL et al., 2008), entre os quais estão glicerol, ureia, amônia (NH<sub>3</sub>), dióxido de carbono (CO<sub>2</sub>), peróxido de hidrogênio (H<sub>2</sub>O<sub>2</sub>), e metalóides, tais como boro e silício (MAUREL et al., 2008; GOMES et al., 2009; POMMERREINIG; DIEHN; BIENERT, 2015).

Em plantas, a quantidade de isoformas de aquaporinas é consideravelmente superior a animais: por exemplo, em *Populus trichocarpa*, foram identificados 58 genes de aquaporinas (GUPTA; SANKARARAMAKRISHNAN, 2009); em *Hevea brasiliensis* Muell. Arg., 51 (ZOU et al., 2015a); *Brassica oleracea*, 67 genes (DESHMUKH et al., 2015) e, em *Glycine max*, 72 (DESHMUKH et al., 2013), em comparação aos 13 genes identificados em humanos (MAGNI et al., 2006).

O elevado número de isoformas em plantas pode estar associado à natureza sésil das plantas e à ausência de um sistema circulatório (FORREST; BHAVE, 2007), podendo, também, estar relacionada à seletividade aos diferentes substratos, à localização celular e tecidual e às diversas funções celulares que promovem o equilíbrio hídrico e osmótico entre os meios extra e intracelular (WUDICK et al., 2009). Além disso, sugere-se que representam vantagens adaptativas, em decorrência do processo de seleção, em função das diversidades ambientais (CHAUMONT; MOSHELION; DANIELS, 2005; GASPAR, 2011).

As aquaporinas estão presentes em todos os tecidos vegetais, sendo reguladas espacial e temporalmente, a depender do estágio de desenvolvimento e condições ambientais, estando envolvidas não apenas na captação e transporte de água (SIEFRITZ et al., 2002; JAVOT et al., 2003), como também na reprodução (BOTS et al., 2005), no alongamento das células (HUKIN et al., 2002), na fotossíntese e germinação das sementes (VINNAKOTA et al., 2015), além de importantes componentes de adaptações a estresses salinos e hídricos (WALLACE et al., 2006; WUDICK et al., 2009; HEINEN; CHAUMONT, 2009).

### 3.3.1 Subfamílias de aquaporinas em plantas superiores

Com base na homologia de sequências, as aquaporinas vegetais foram classificadas em sete subfamílias: PIP (*Plasma Membrane Intrinsic Protein*); TIP (*Tonoplast Intrinsic Proteins*) (MAUREL et al., 2002, WALLACE et al., 2006); NIP (*Nodulin26-like Intrinsic*

*Proteins*) (WALLACE et al., 2006), SIP (*Small and Basic Intrinsic Proteins*) (GOMES et al., 2009), XIP (*Membrane Intrinsic X*), HIPs (*Hybrid Intrinsic Proteins*) e GIP (*GlpF-like Intrinsic Proteins*) (DANIELSON; JOHANSON, 2008).

A diversificação precoce das MIPs resultou em um grande número de subfamílias ainda em plantas terrestres primitivas, sendo que, durante a evolução, algumas plantas superiores foram perdendo duas destas subfamílias (HIP e GIP), enquanto as subfamílias restantes se expandiram e, em alguns casos, diversificara-se, resultando na formação de grupos mais especializados dentro destas subfamílias (DANIELSON; JOHANSON, 2008). No caso das monocotiledôneas, a subfamília XIP é ausente em todo o grupo, assim como nas *Brassicaceae* (GUPTA; SANKARARAMAKRISHNAN, 2009; DIEHN et al., 2015; DESHMUKH et al., 2015; 2016).

- Proteínas Intrínsecas de Membrana Plasmática – PIP

As Proteínas Intrínsecas de Membrana Plasmática (PIPs) constituem a maior subfamília entre as aquaporinas vegetais. As proteínas classificadas como PIPs encontram-se divididas em dois grupos menores as PIP1 e PIP2, que diferem entre si pelo tamanho de suas extremidades N- e C- terminal, como também em relação à permeabilidade à água, sendo que as PIP2 são mais eficientes no transporte de água (5 a 20 vezes mais eficazes), quando comparadas com as PIP1 (KALDENHOFF; FISCHER, 2006; DANIELSON; JOHANSON, 2010; BEZERRA-NETO, 2012).

Estudos realizados com *Physcomitrella patens* (briófita) sugere que esses grupos devem ter se formado no início da evolução das plantas terrestres e são de fundamental importância na fisiologia da planta (DANIELSON; JOHANSON, 2008). Membros desta subfamília são localizados principalmente em tecidos que apresentam alto fluxo de água, como no caso de tecidos vasculares e células guarda de estômatos, apresentando-se altamente expressivas em raízes, elementos condutores do floema e de folhas, podendo ter importante papel na comunicação célula-a-célula (DYNOWSKI et al., 2008).

As PIPs parecem estar envolvidas na tolerância a déficit hídrico, presente em alguns vegetais. Em estudo realizado com AQPs de *Oryza sativa*, durante exposição a diferentes tipos de estresse abiótico, foi demonstrado um aumento nos seus níveis de expressão das OsPIPs, sobretudo em situações de déficit hídrico (GUO et al., 2006). Foi também relatado que, no tratamento com polietilenoglicol (PEG), os genes das AQPs: *OsPIP1.2*, *OsPIP1.3*, *OsPIP2.1* e *OsPIP2.5* em raízes, assim como *OsPIP1.2* e *OsPIP1.3* em folhas, foram

superexpressos, sugerindo o importante papel destas aquaporinas durante o estresse hídrico (LIAN et al., 2006).

- Proteínas Intrínsecas de Tonoplasto – TIP

A subfamília TIP apresenta grande número de subgrupos entre as MIPs, apresentando cinco grandes grupos (TIP1, TIP2, TIP3, TIP4 e TIP5). Os membros dos grupos TIP1, TIP2 e TIP4 foram localizados nas membranas de vacúolos centrais, enquanto que os membros do grupo TIP3 estão localizados na membrana dos corpos proteicos (MAESHIMA et al., 1994; MAESHIMA, 2001).

A presença das TIPs permite o rápido ajuste osmótico, atuando na manutenção da osmolaridade e do turgor (GOMES et al., 2009). As TIPs são encontradas na membrana do vacúolo celular, atuando nos processos de regulação, sinalização e degradação celular. Várias destas proteínas, além de realizar o transporte de pequenos solutos e gases, transportam metabólitos importantes envolvidos no ciclo da ureia e de síntese de aminoácidos (KALDENHOFF; FISCHER, 2006).

Estudo realizado em *A. thaliana*, mostra que aquaporinas classificadas como *AtTIP1;1*, *AtTIP1;2* e *AtTIP2;3* foram caracterizadas como capazes de transportar também peróxido de hidrogênio (H<sub>2</sub>O<sub>2</sub>) (BIENERT et al., 2007). Além disso, elas podem estar envolvidas na desintoxicação do citoplasma por um mecanismo de armadilha ácida, que facilita o transporte de amônia (LOQUE et al., 2005).

- Proteínas Intrínsecas de Nódulos – NIP

As NIPs são um tipo de proteínas transportadoras normalmente presentes em membranas de raízes, caules, sementes e na membrana de simbiossomo. Em vários experimentos, estas proteínas apresentaram capacidade de transporte tanto para água quanto para glicerol, porém exibem uma baixa permeabilidade para a água (KALDENHOFF; FISCHER, 2006; GOMES et al., 2009), sendo responsáveis também pelo transporte de ureia e metaloides (WALLACE; ROBERTS, 2004; BIENERT; SCHUSSLER; JAHN, 2008)

As proteínas desta subfamília apresentam expressão inferior à de PIP e TIP, já que a sua distribuição é restrita a determinadas regiões, a exemplo das zonas de alongação radicular e por estarem associada a especializações celulares. Devido à sua localização e alta

especificidade, a maioria dos processos nos quais as NIPs atuam são ainda desconhecidos (GOMES et al., 2009).

- Pequenas Proteínas Básicas Intrínsecas de Membrana - SIP

A subfamília SIPs é a mais divergente entre as cinco (CHAUMONT et al., 2001; DANIELSON; JOHANSON, 2010; BEZERRA NETO, 2012). As proteínas agrupadas nesta subfamília representam o menor grupo dentre as MIPs (KALDENHOFF; FISCHER, 2006).

Apresentam ponto isoelétrico básico, o que difere das outras MIPs. Possuem uma extremidade N-terminal curta e uma modificação no motivo NPA, localizado no *loop* B, o que confere uma maior abertura no poro, gerando uma alta especificidade por substratos diferentes (GOMES et al., 2009).

Ainda são desconhecidas as reais funções desempenhadas pelas SIPs vegetais, porém estudos *in vitro* indicam que estas podem estar localizadas na membrana do retículo endoplasmático, atuando na regulação do volume e na concentração de íons no interior do lúmen, sendo responsáveis também pela manutenção da morfologia da organela (FORREST; BHAVE, 2007).

- Proteínas Intrínsecas Desconhecidas - XIP

A subfamília XIP encontra-se presente apenas em dicotiledôneas (DESHMUKH et al., 2015; 2016). O gênero *Populus* possui o maior número de isoformas XIP, com nove sequências PtXIPs, distribuídas dentro de três grupos. Análise realizada com sequências XIPs encontradas em bancos de dados revelou que os membros desta subfamília se distribuem em, pelo menos, cinco subgrupos (o que é suportado pela semelhança entre as sequências e as modificações existentes no filtro arginina aromático (ar/R) (LOPEZ et al., 2012).

Como estas proteínas apresentam variações de aminoácidos nas regiões NPA e ar/R (MITANI-UENO. et al, 2011), foi levantada a hipótese de que XIPs não são funcionais como canais de água, mas, em vez disso, realizam o transporte de solutos hidrofóbicos (DANIELSON; JOHANSON, 2008; GUPTA; SANKARARAMAKRISHNAN, 2009). Esta hipótese foi recentemente validada por Bienert et al. (2011). As modificações presentes na região ar/R permitem a formação de um poro largo, altamente hidrofóbico, que se assemelha ao grupo NIP3, o que se mostrou facilitar a difusão de ureia, glicerol, e metaloides, mas é

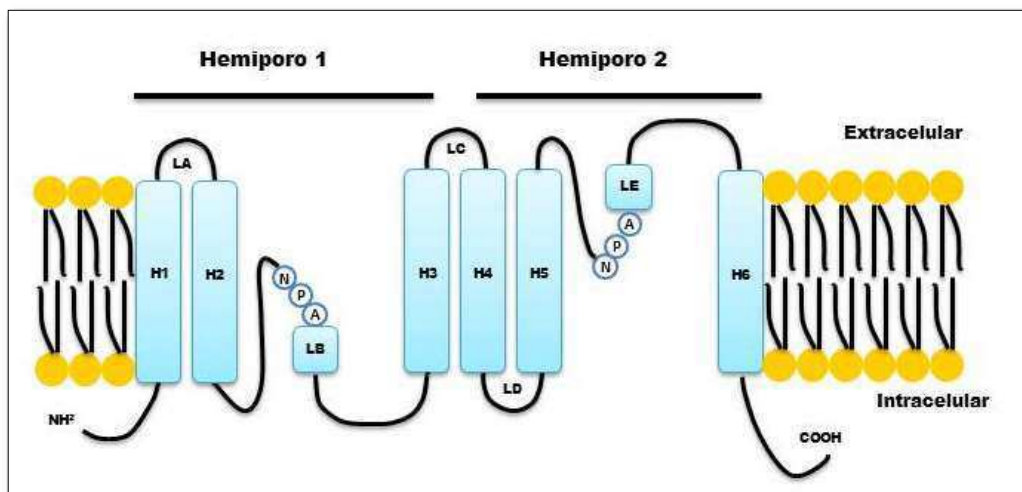
fracamente permeável à água (BIENERT; SCHUSSLER; JAHN., 2008; DYNOWSKI et al., 2008).

### 3.3.2 Estrutura e localização celular das aquaporinas

Todas as aquaporinas até hoje caracterizadas agrupam-se na membrana como tetrâmeros, e a composição desses tetrâmeros pode determinar o seu direcionamento para a membrana plasmática (ZELAZNY et al., 2007). As aquaporinas apresentam uma estrutura composta por seis alfa-hélices hidrofóbicas que formam os domínios transmembrana (H1-H6), as quais estão arranjadas em duas metades semelhantes (hemiporo 1=H1-H3 e hemiporo 2=H4-H6), provavelmente originadas por duplicação gênica durante a evolução (GOMES et al., 2009).

A conexão entre as alfa-hélices ocorre por meio de cinco alças (*loops* A-E), onde os *loops* A, C e E estão localizados no lado exoplasmático e os *loops* B e D, onde se encontram os motivos NPA (Asn-Pro-Ala), localizados no lado citoplasmático, juntamente com as regiões N e C-terminais (Figura 4).

Figura 4. Modelo estrutural de aquaporina, demonstrando suas principais características. Alfas-hélices (retângulos azuis), nelas estão presentes domínios transmembrana (H1-H6), conectados pelos cinco loops (LA-LE). Dois domínios alfa-hélices (loops B e E), contendo os motivos NPA, responsáveis pela formação do poro. Adaptado de Bezerra Neto (2012).

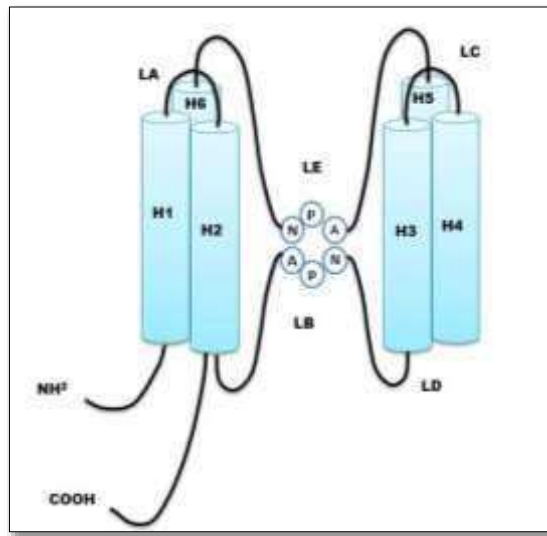


Na conformação tridimensional, os hemi-poros estão posicionados um de frente para o outro, em orientação oposta à membrana, formando o chamado “modelo ampulheta” (JUNG et al. 1994) (Figura 5). Os motivos NPA (regiões altamente conservadas que são parte da assinatura da superfamília MIP) (PRESTON et al., 1994) e a região arginina/aromática (ar/R),



conhecida por filtro seletivo (formada pela interação entre quatro resíduos conservados localizados nas hélices H2 e H5 e na alça E (LE1 e LE2) (AGRE; KOZONO, 2003), exercem papel fundamental sobre a especificidade de substratos transportados através das aquaporinas (WALLACE; ROBERTS, 2004; FORREST; BHAVE, 2007).

Figura 5. Estrutura e formação do poro das aquaporinas. Os retângulos em azul representam as alfas-hélices, onde estão presentes domínios transmembrana (H1-H6). Os cinco *loops* de conexão entre as alfas-hélices estão representados por LA-LE. No *loops* B e E, estão os motivos NPA, responsáveis pela formação do poro. Adaptado de Cheidde e Schor (1999).



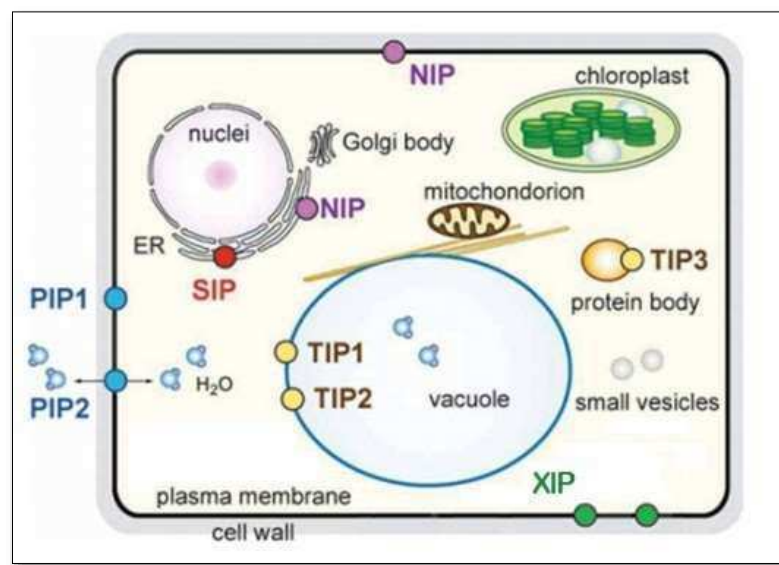
Na região NPA, as asparaginas presentes no centro do poro formam ligações de hidrogênio com as moléculas transportadas e podem desempenhar um papel na exclusão de prótons (FORREST; BHAVE, 2007), já o filtro seletivo (ar/R) estabelece pontes de hidrogênio e ligação de *Van der Waals* com os solutos transportados e / ou moléculas de água (SUI et al., 2001).

Além dos motivos NPA e do filtro seletivo, as aquaporinas PIPs, TIPs e NIPs apresentam outra região conservada, denominada de motivo AEF (Ala-Glu-Phe), localizada no domínio N-terminal, a qual, até o momento, não tem função conhecida. Os motivos e resíduos são regiões encontradas conservados dentro dos grandes grupos, porém, entre eles, é possível identificar variações, as quais são responsáveis pela variedade de moléculas transportadas pelas aquaporinas (ZARDOYA; VILLALBA, 2001).

As aquaporinas vegetais podem ser encontradas em todos os compartimentos subcelulares, sendo extremamente importantes para o controle do transporte de água e pequenos solutos, tanto na membrana celular externa, como através das membranas intracelulares (MAUREL et al., 2008). Os membros das subfamílias PIP e XIP encontram-se

localizados na membrana plasmática (LOPEZ et al., 2012). Em geral, as TIPs estão localizadas na membrana do vacúolo, enquanto as SIPs encontram-se no retículo endoplasmático (FORREST; BHAVE, 2007). Já os membros das subfamílias NIP possuem localização subcelular incerta (CHAUMONT; MOSHELION; DANIELS, 2005), embora tenha sido evidenciada presença de membros desta subfamília em membranas do retículo endoplasmático e na membrana plasmática de células do simbiossomo, nos nódulos radiculares (MAUREL, 2007; MAESHIMA; ISHIKAWA, 2008) (Figura 6).

Figura 6. Diagrama esquemático da localização dos membros das subfamílias de aquaporinas de planta. Adaptado de Maeshima e Ishikawa (2008).



A estrutura principal, o substrato de transporte, a regulação funcional, o perfil de expressão gênica, bem como a abundância e localização intracelular das aquaporinas são altamente diversificados (MAESHIMA; ISHIKAWA, 2008). Desta forma, não é possível traçar um perfil de expressão geral para as MIPs, já que as diversas isoformas existentes podem ser superexpressas ou reprimidas, dependendo do tecido estudado e dos estímulos ambientais recebidos.

A ausência de um padrão de expressão permite a adaptação das plantas a condições desfavoráveis. Em situações de estresses abióticos, como seca e salinidade, alguns membros são superexpressos para garantir o transporte eficiente de água e alguns outros solutos (KAWASAKI et al., 2001), enquanto outros são reprimidos (BOURSIAC et al., 2005), sugerindo que as diferentes isoformas respondem de formas distintas aos diferentes estresses (AROCA; PORCEL; RUIZ-LOZANO, 2011; BEZERRA NETO, 2012).

### 3.4 Avanço das “ômicas” e ferramentas computacionais na prospecção e análise de genes

A genômica é definida como a parte da Biologia Molecular responsável pelo estudo dos genes dos organismos vivos, ou seja, todo conjunto de DNA que ele carrega em suas células (DELOUKAS et al., 1998; ECHEVERRIGARAY, 2007). Os avanços recentes das tecnologias de sequenciamento revolucionaram o campo das ômicas, apresentando inovações sem precedentes, além de implementarem aplicações viáveis como o sequenciamento de genomas completos em um curto intervalos de tempo, ressequenciamento de genoma inteiro para análise de variação, sequenciamento de RNA (RNA seq) para análise de transcriptoma (BURGESS, 2016; CHEN et al., 2016), detecção quantitativa de dinâmica epigenômica e análise de Chip-seq para interações DNA-proteína (LISTER et al., 2009).

Além dessas, outras abordagens foram desenvolvidas, incluindo a análise interativa para redes formadas por interações proteína-proteína (Arabidopsis Interactome Mapping Consortium, 2011), entre outras (KOJIMA et al., 2009; SAITO; MATSUDA, 2010). As informações geradas para espécies de plantas-modelo, como *A. thaliana*, *Oryza sativa* (arroz) e *Populus trichocarpa*, deram início à era das “ômicas” em plantas. Esta nova era na genética molecular compreende o estudo de transcritos (transcriptômica), de proteínas (proteômica), dos metabólitos celulares (metabolômica) e várias outras (BHALLA; NARASIMHAN; SWARUP, 2005).

Porém, o sequenciamento das moléculas de DNA e RNA, embora fosse indiscutivelmente relevante, eram limitado e não conseguia explicar os processos biológicos, em especial, o que ocorre com o próprio material genético após a duplicação, a transcrição e a tradução do material genético (PIMENTA, 2003). O desenvolvimento das ferramentas de Bioinformática foi crucial para o sucesso das ômicas, não apenas por possibilitar a organização dos dados, mas por permitir identificar genes e suas funções, proporcionando a compreensão dos processos biológicos, tais como: a diferenciação celular, a morfogênese, a determinação fenotípica, a resistência a patógenos e a adaptabilidade, além de disponibilizar estas informações em banco de dados (SOUZA; RHODEN; PAMPHILE, 2014).

#### 3.4.1 Bancos de dados biológicos

Com a grande demanda de sequências biológicas geradas pelos sequenciamentos em larga escala, foi necessário o desenvolvimento e aprimoramento de ferramentas

computacionais que permitam o gerenciamento destas informações, facilitando o armazenamento e a interpretação adequada das mesmas. A partir de então, a bioinformática passou a ocupar uma posição significativa na interpretação dos dados provenientes de diversas tecnologias (PROSDOCINI et al., 2002; AMARAL; REIS; SILVA, 2007).

O desenvolvimento de programas e algoritmos desta natureza permitiu a manipulação de uma grande variedade de dados biológicos, marcando um momento de grande desenvolvimento para a genética e as demais ciências que, de alguma forma, utilizam-se dessas informações. Através destas ferramentas, é possível armazenar, processar, avaliar, descrever estruturas, delinear relações entre moléculas e interpretar grande quantidade de sequências (BORÉM; SANTOS, 2001).

Com o avanço da bioinformática e suas ferramentas, várias áreas vêm emergindo nos últimos anos, com destaque para a mineração de texto e biologia de sistemas (KEMPER et al., 2010). Além de design racional de proteínas, análise de imagem, evolução dirigida e computação em grade (BIASINI et al., 2014), bioinformática relacionada ao microRNA (LIU; ZHOU; WHITE, 2014).

Além das várias funções citadas anteriormente, existia também a necessidade de compartilhamento de dados, possibilitando que estes fossem acessados por pesquisadores de diversas localidades do mundo, contribuindo para a troca de informações, como também possibilitar a comparação entre as sequências, facilitando sua identificação, o que foi solucionado com a criação de bancos de dados públicos e privados (MORAIS, 2003).

Banco de dados (*DB, data bank*) é definido como o espaço adequado para o armazenamento de sequências biológicas, sejam elas de nucleotídeos ou aminoácidos. Alguns desses bancos podem ser de acesso restrito, outros classificados como públicos de acesso ilimitado, sem a necessidade de algum tipo de autorização prévia. A inserção de sequências biológicas em bancos de dados possibilita a realização de comparações usando as sequências já caracterizadas como referências para a identificação de novas sequências (AMARAL; REIS; SILVA, 2007).

Centenas de bancos de dados vêm disponibilizando informações de uma grande variedade de espécies ou grupos de organismos, incluindo sequências nucleotídicas e proteicas (FERNANDEZ-SUAREZ; GALPERIN, 2012). Entre os vários bancos de dados públicos existentes, destacam-se o DDBJ (*DNA Databank of Japan*), o EMBL (*European Molecular Biology Laboratory*) e o GenBank (*National Center for Biotechnology Information*), que são considerados os três maiores bancos mundiais. Estes bancos trabalham em cooperação para estabelecer formatos de dados e protocolos que facilitem a submissão

segura dos dados, além de proporcionar o intercâmbio contínuo de dados em todo o mundo, formando o INSD (*International Nucleotide Sequence Database*; Banco de Dados Internacional de Sequências de Nucleotídeos) (TATENO et al., 2002; NAKAMURA; COCHRANE; KARSCH-MIZRACHI, 2012).

O *GenBank* (Banco de Genes), sediado nos EUA, é um dos mais conhecidos e mais usados bancos de dados. Nele são armazenadas e disponibilizadas sequências para consulta pública, abrangendo, desde sequências pequenas, como DNA, RNA e proteínas, até genomas completos. Além de sequências primárias, o *GenBank* dispõe de informações sobre estruturas proteicas, taxonomia, mapas gênicos, entre outros. É importante destacar também o seu abrangente sistema de busca bibliográfica - o *PubMed*, um serviço da Biblioteca Nacional de Medicina. O *GenBank* está acessível no NCBI (*National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica) (BENSON; KARSCH-MIZRACHI; LIPMAN, 2000).

No âmbito da biologia vegetal, o Phytozome é um banco de dados que fornece acesso a informações de 52 genomas de plantas. Neste banco, é possível encontrar informações sobre famílias de genes, genes individuais, dados de diversidade e expressão (GOODSTEIN et al., 2012).

Juntamente com os bancos de dados, foram criadas, pela bioinformática, diversas ferramentas que ensejam o acesso e análise da grande quantidade de dados disponíveis nos bancos. Estas ferramentas permitem, por exemplo, a avaliação *in silico* de sequências desconhecidas através de análises de alinhamentos comparativos (KENT et al., 2002). A mais comum destas ferramentas é o BLAST (*Basic Local Alignment Search Tool*; Ferramenta de Busca por Alinhamento Local), um algoritmo que possibilita a comparação entre sequências disponíveis em domínio público (AMARAL; REIS; SILVA, 2007)

### **3.4.2 Ferramentas computacionais para caracterização e anotação de sequências**

O BLAST (*Basic Local Alignment Search Tool*) é uma ferramenta básica de pesquisa por alinhamento local disponível no NCBI, que utiliza matriz de substituição para avaliar a similaridade entre sequências de aminoácidos ou de nucleotídeos (PEARSON; LIPMAN, 1988; ALTSCHUL et al., 1990), sendo uma das mais utilizadas para busca e alinhamento de sequências. Os alinhamentos de sequência frequentemente proporcionam a primeira ligação entre DNA ou proteína recentemente sequenciada e sequências já categorizadas. O BLAST

também fornece estatísticas que estimam a probabilidade de uma correspondência ocorrer por acaso (BORATYN et al., 2013).

Esta ferramenta realiza a comparação entre uma sequência de DNA ou proteína (denominada “*Query*”) e todas as sequências genômicas disponíveis em domínio público, resultando na seleção de sequências disponíveis na base de dados (“*Subject*”) que apresentam maior homologia com a “*Query*”. O BLAST pode traduzir sequências de nucleótidos conforme necessário, permitindo, assim, a consulta de nucleotídeos contra uma base de dados de proteínas ou uma consulta de proteínas contra uma base de dados de nucleotídeos (BORATYN et al., 2013).

O BLAST possui cinco subdivisões, que desempenham diferentes modalidades de alinhamento, entre as quais pode-se escolher aquela que, de acordo com o tipo de molécula de interesse e da resposta que se deseja obter através da comparação, possa apresentar o melhor resultado (AMARAL; REIS; SILVA, 2007). A modalidade BLASTn (*nucleotide blast*) busca por homologia entre sequências de nucleotídeos, a modalidade BLASTp (*protein blast*) busca por homologia entre sequências de proteínas e, na modalidade BLASTx uma sequência de nucleotídeos é usada como “*Query*” para realizar comparações contra banco de proteínas. No caso do tBLASTx, tanto a “*Query*” como a base de dados “*Subject*” são sequências de nucleotídeos. Já no tBLASTn, a sequência de aminoácidos é comparada em um banco de dados de nucleotídeos (GIBAS; JAMBECK, 2001).

Outras ferramentas disponíveis no NCBI são *ORF-Finder* (*Open Reading Frame Finder*) e *CD-search* (*Conserved Domain Search*). O *ORF-Finder* é um identificador que gera uma análise gráfica indicando todas as pautas abertas de leitura de uma sequência nucleotídica. O programa retorna o intervalo de cada ORF, juntamente com a sua tradução em proteínas (NCBI, 2011), enquanto o *CD-search* é uma ferramenta que permite detectar domínios estruturais e funcionais em sequências de proteínas, usando a heurística do BLAST (MARCHLER-BAUER; BRYANT, 2004).

A análise comparativa de sequências permite a reconstrução da história evolutiva das espécies e das famílias multigênicas, estimando as taxas da evolução molecular (KUMAR et al., 2004). Para a análise comparativa global de sequências nucleotídicas ou proteicas, o ClustalW (THOMPSON et al., 1997) é um dos softwares mais utilizados para efetuar estes alinhamentos múltiplos biologicamente informativos. O alinhamento realizado pelo ClustalW é executado em três etapas: alinhamento *pairwise*, geração de árvore-guia e alinhamento progressivo (KOU-BIN, 2003).

O programa MEGA (*Molecular Evolutionary Genetics Analysis*) analisa caracteres evolutivamente informativos, presentes em sequências nucleotídicas, proteicas ou marcadores moleculares (TAMURA et al., 2007; 2013), permitindo a análise da matriz de dados através de métodos utilizados para calcular as distâncias genéticas, como distância P (NEI, 1991), distância de Jukes-Cantor (JUKES-CANTOR, 1969), distância de Tajima-Nei (TAJIMA; NEI, 1984), distância de Kimura-2- parâmetros (KIMURA, 1980), distância de Tamura (TAMURA, 1992; TAMURA et al., 2013).

O MEGA também disponibiliza algoritmos como UPGMA (*Unweighted Pair Group Method with Arithmetic Means*) (SNEATH; SOKAL, 1973), NJ (*Neighbor-Joining*) (SAITOU; NEI, 1987) e máxima parcimônia (ECK; DAYHOFF, 1966), permitindo a realização de inferências filogenéticas e fenéticas através da construção de dendrogramas (KUMAR et al., 2004), podendo ser visualizados, por exemplo, no programa *TreeView* (PAGE, 1996; THOMPSON et al., 1997). O *TreeView* possibilita a visualização de dendrogramas, sendo este, capaz de ler diferentes formatos de arquivos, como *Nexus*, *Phylip*, *Nona*, *Mega* e *ClustalW/X*. (PAGE, 1996; MORAIS, 2003).

#### 4 REFERÊNCIAS

ABASCAL, F., I. IRISARRI, R. ZARDOYA. Diversity and evolution of membrane intrinsic proteins. **Biochim. Biophys. Acta**, v.1840, p. 1468–1481, 2014.

AGRE, P. "**The aquaporin water channels.**"Proceedings of the American Thoracic Society. v. 3.1 p. 5-13, 2006.

AGRE, P. **Aquaporin water channels** (Nobel lecture). *Angewandte Chemie International Edition*, v. 43, n. 33, p. 4278-4290, 2004.

AGRE. P.; KOZONO, D. Aquaporin water channels: molecular mechanisms for human diseases. **FEBS Letters**, v. 555, n. 1, p. 72-78, 2003.

ALOUFA M. A. I., MEDEIROS J.A., Valorização e preservação da *C. quercifolius* (*C. quercifolius*) para o desenvolvimento sustentável do semiárido brasileiro. **Revista OKARA: Geografia em debate**, v. 10, n. 3, p. 453-476, 2016.

ALTSCHUL, S.F. et al. Basic local alignmentsearch tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ALVES, M.V. Checklist das espécies de Euphorbiaceae Juss. ocorrentes no semi-árido pernambucano, Brasil. **Acta Bot. Bras.**, v. 12(3), p. 485-495, 1998.

AMARAL, A.M.; REIS, M.S.; SILVA, F.S. O programa BLAST: guia prático de utilização. **Embrapa recursos genéticos e biotecnologia**, Brasília, 1676 -1340; 224, p.24. 2007

Arabidopsis Interactome Mapping Consortium. Evidence for Network Evolution in an *Arabidopsis* Interactome Map. **Science** (New York, NY), v. 333, n. 6042, p. 601-607, 2011.

AROCA, R.; PORCEL, R.; RUIZ-LOZANO, J.M. Regulation of root water uptake under abiotic stress conditions. **Journal of Experimental Botany**.v. 13, 2011.

ARRIEL, E. F. et al. Divergência genética entre progênes de *Cnidocolus phyllacanthus* submetidas a três regimes hídricos. **Científica**. Vol. 34, n.2, p. 229- 237, 2006.

BARTELS, D., SUNKAR R. Drought and salt tolerance in plants. **Crit. Rev. in Plant Sci**, 241-36, 2005.

BENKO-ISEPPON, A.M., et al., Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. **Revista Brasileira de Geografia Física**, v. 06, p. 1112-1134, 2011.

BENSON, D.A.; KARSCH-MIZRACHI, I.; LIPMAN, D.J. GenBank. **Nucleic Acids Research**, v. 28, p. 15-18, 2000.

BEZERRA-NETO, J.P. **Cracterização e análise evolutiva in silico de quaporinas no transcriptoma do feijão-caupi (*Vigna unguiculata* (L.) Walp.)** Dissertação de Mestrado. UFPE. 2012.

BEZERRA-NETO, J.P. **Respostas transcricionais e estresse-induzidas em genótipos contrastantes de *Glycine max* (soja) e *Vigna unguiculata* (feijão-caupi)**. Tese de Doutorado. UFPE, 2016.

BHALLA, R.; NARASIMHAN, K.; SWARUP, S. Metabolomics and its role in understanding cellular responses in plants. **Plant Cell Reports**, Berlin, v. 24, n. 10, p. 562-71, 2005.

BIASINI M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Res**, v.42 (W1), p. W252-W258, 2014.

BIENERT, G.P. et al. Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. **The Plant Journal**, 2011.

BIENERT, G.P. et al. Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. **Journal of Biological Chemistry**, v. 282, p. 1183-1192, 2007.

BIENERT, G.P.; SCHÜSSLER, M.D.; JAHN, T.P. Metalloids: essential, beneficial or toxic? Major intrinsic proteins sort it out. **Trends in Biochemical Sciences**, v. 33, p. 20–26, 2008.

BIJEKAR, S; GAYATRIM.C. Ethanomedicinal properties of Euphorviaceae family-A comprehensive review. **International Journal of Phytomedicine**. v. 6,p. 144-156, 2014.



BORATYN G.M. et al. BLAST: a more efficient report with usability improvements. **Nucleic Acids Res**, v.41 (W1), p.W29-W33, 2013.

BORÉM, A.; SANTOS, F.R. **Biotecnologia simplificada**. Editora Suprema. Viçosa, MG, 2001.

BOTS, M. et al. PIP1 and PIP2 aquaporins are differentially expressed during tobacco anther and stigma development. **Journal of Experimental Botany**, v. 56, p. 113-121, 2005.

BOURSIAC, Y. et al. Early effects of salinity on water transport in *Arabidopsis* roots. Molecular and cellular features of aquaporin expression. **Plant Physiology**, v. 139, n. 2, p. 790-805, 2005.

BRAY, E. Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. **Journal of Experimental Botany**, Oxford, v. 55, n. 407, 2004.

BURGESS D. J. A space for transcriptomics. **Nature Reviews Genetics** .2016.  
<http://www.nature.com/nrg/journal/v17/n8/full/nrg.2016.94.html>

AMPOS, G. N. F. **Clonagem de *Cnidocolus Phyllacanthus* (Mart.) Pax et K. Hoffm. (*C. quercifolius*) por alporquia**. Dissertação, UFCG, Patos, 45p.2010.

CAVACANTI, M. T. **Utilização das sementes da faveleira (*Cnidocolus phyllacanthus* (Mart.) Pax et K. Hoffm.) em produtos alimentícios**. 2011. 330f. Tese (Doutorado em Engenharia de processos). Campina Grande, PB: Universidade Federal de Campina Grande, Centro de Ciência e Tecnologia, 2011.

CEBALLOS, H. et al. Cassava. Cap. 2, p. 53–96. In: BRADSHAW, J. E. (ed.). **Root and tuber crops**, New York: Springer, vol. 7, 2010.

CHAUMONT, F.; BARRIEU, F.; WOJCIK, E.; CHRISPEELS, M.J.; JUNG, R. Aquaporins constitute a large and highly divergent protein family in maize. **Plant Physiology**, v. 125, n. 1206-1215, 2001.

CHAUMONT, F.; MOSHELION, M.; DANIELS, M.J. Regulation of plant aquaporin activity. **Biology of the Cell**, v. 97, p. 749-764, 2005.

CHEN, W. et al. Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. **Front. Plant Sci.** v. 7, p. 1044, 2016.

CHOU, K.C. AND SHEN, H.B. Cell-PLoc: A package of web-servers for predicting subcellular localization of proteins in various organisms. **Nature Protocols**, v. 3, p. 153-162, 2008.

CONAB – Companhia Nacional de Abastecimento. **Conjuntura mensal: Mandioca**, 2016.

CORDEIRO, I.; SECCO, R.; SILVA, M. J. DA; SODRÉ, R. C.; MARTINS, M. L. L. Manihot in Lista de Espécies da Flora do Brasil. **Jardim Botânico do Rio de Janeiro**. 2015

- CORDEIRO, I.;CARNEIRO-TORRES, D.S. **Euphorbiaceae**. In: M.R.V. Barbosa, C. Sothers, S. Mayo, C.F.L. Gamarra & A.C. Mesquita. (orgs.). Checklist das plantas do Nordeste brasileiro v.1: Angiospermas e Gymnospermas. Ministério da Ciência e Tecnologia, Brasília, p.71-74, 2006.
- CORRÊA, A. D. et al. Farinha de folhas de *M. esculenta*: I. efeito da secagem das folhas sobre a atividade da linamarase. **Ciências e Agrotecnologia**, Lavras, v. 26, n. 2, p. 368-374, 2002.
- DANIELSON, J. A H.; JOHANSON, U. Phylogeny of major intrinsic proteins. **Advances in experimental medicine and biology**, v. 679, p. 19-31, 2010.
- DANIELSON, J.A. et al. The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. **Plant Physiology**, v. 126, n. 1358–1369, 2001.
- DANIELSON, J.A.; JOHANSON, U. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. **BMC Plant Biology**, v. 8, n. 45, 2008.
- DELOUKAS, P. et al. A physical map of 30.000 human genes. **Science**, v. 282, p. 744-746, 1998.
- DESHMUKH R.K. et al. Identification and functional characterization of sílicon transporters in soybean using comparative genomics of major intrinsic proteins in *Arabidopsis* and rice. **Plant Mol Biol.**, v. 83(4–5), p.303–315, 2013.
- DESHMUKH RK, SONAH H AND BÉLANGER RR. Plant Aquaporins: Genome-Wide Identification, Transcriptomics, Proteomics, and Advanced Analytical Tools. **Front. Plant Sci.** v.7, p.1896, 2016.
- DESHMUKH, R. K. et al. A precise spacing between the NPA domains of aquaporins is essential for sílicon permeability in plants. **Plant J.** v. 83, p. 489–500, 2015.
- DIEHN, T. A. et al. Genome-wide identification of aquaporin encoding genes in *Brassica oleracea* and their phylogenetic sequence comparison to *Brassica* crops and *Arabidopsis*. **Front. Plant Sci.** v. 6, p. 166, 2015.
- DYNOWSKI, M. et al. Plant plasma membrane water channels conduct the signalling molecule H<sub>2</sub>O<sub>2</sub>. **Biochemical Journal**, v. 414, p. 53-61, 2008.
- ECHEVERRIGARAY S. **Genômica: decifrando a linguagem da vida**. Educ. v. 1(9), p. 36-42, 2007.
- ECK, R.V.; DAYHOFF, M.O. Atlas of protein sequence and structure. **Silver Spring**, p. 1-68, 1966.
- EMBRAPA. Empresa Brasileira de Pesquisa Agropecuária. **Árvore do conhecimento – R. communis**, 2012. Disponível em: <http://www.agencia.cnptia.embrapa.br/gestor/R.communis/arvore/CONT000h4pitb4s02wx7ha0awymtyiscijnl.html>

ENDRESS, P. K., DAVIS, C. C., MATTHEWS, M. L. Advances in the floral structural characterization of the major subclades of Malpighiales, one of the largest orders of flowering plants. **Ann. Bot.** v. 111, p. 969–985, 2013.

FAO - FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATION. **Save and grow: Cassava**. A guide to sustainable production intensification. Rome: Food and Agriculture Organization of the United Nation, 2013.

FAOSTAT - Food and Agriculture Organization of the United Nations Statistics Division. **Production quantities by country: Average 1993 – 2013**.

FAOSTAT. **Food and Agriculture Organization of the United Nations**, 2013. Disponível: <http://www.agencia.cnptia.embrapa.br/gestor/R.communis/arvore/CONT000h4pitb4s02wx7ha0awymtyiscijnl.html>

FERNÁNDEZ-SUÁREZ X.M, GALPERIN M.Y. The 2013 *Nucleic Acids Research* Database Issue and the online Molecular Biology Database Collection. **Nucleic Acids Res**, v. 41 (D1), p. 1-7, 2013.

FORREST, K.L.; BHAVE, M. Major intrinsic proteins (MIPs) in plants: a complex gene family with major impacts on plant phenotype. **Functional & Integrative Genomics**, v. 7, p. 263-289, 2007.

FORZZA, R. C., org., et al. Instituto de Pesquisas Jardim Botânico do Rio De Janeiro. **Catálogo de plantas e fungos do Brasil [online]**. Rio de Janeiro: Andrea Jakobsson Estúdio: Instituto de Pesquisa Jardim Botânico do Rio de Janeiro, 2010. Introdução: as angiospermas do Brasil, v. 1, p. 78-89, 2010.

FRANÇA, F. et al. Flora vascular de açudes de uma região do semi-árido da Bahia, Brasil. **Acta Botanica Brasilica**, v. 17, p. 549-559, 2003.

GASPAR, M. Aquaporinas: de canais de água a transportadores multifuncionais em plantas. **Revista Brasil. Bot.**, v.34, n.4, p.481-49. 2011.

GIBAS, C.; JAMBECK, P. **Desenvolvendo Bioinformática-Ferramentas de software para aplicações em Biologia**. Rio de Janeiro. Editora campus, p. 440, 2001.

GIULIETTI, A. M.; CONCEIÇÃO, A.; QUEIROZ, L.P. **Nordeste Semiárido: Caracterização Geral e Lista das Fanerógamas**. In: GIULIETTI, A. M.; CONCEIÇÃO, A.; QUEIROZ, L.P. (Eds.). *Diversidade e caracterização das fanerógamas do semi-árido brasileiro*, Recife: Associação Plantas do Nordeste, p. 15-39, 2006.

GOLFARI, L., CASER, R. L. **Zoneamento ecológico da região nordeste para experimentação florestal**. Série técnica. PRODEPEF, Brasília (10): 1 - 116, 1977.

GOMES, D. et al. Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. **Biochimica et Biophysica Acta**, v. 1788, p. 1213-1228, 2009.

GOODSTEIN D. M., et al. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Res**, v.40 (D1), p. D1178-D1186, 2012.

GUO, L. et al. Expression and functional analysis of the rice plasma-membrane intrinsic protein gene family. **Cell Res**, v. 16, p. 277–286, 2006.

GUPTA, A.B.; SANKARARAMAKRISHNAN, R. Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: characterization of XIP subfamily of aquaporins from evolutionary perspective. **BMC Plant Biology**, v. 9, p. 134, 2009.

HEINEN, R.B.; YE, Q.; CHAUMONT, F. Role of aquaporins in leaf physiology. **J. Exp. Bot.**, v. 60, n. 2971-2985, 2009.

HIANG, L. et al. The protein storage vacuole: a unique compound organelle. **J Biol Chem** v. 155, p. 991–1002, 2001.

HUANG, G.T. et al. Signal transduction during cold, salt, and drought stresses in plants. **Molecular Biology Reports**, v. 39 p. 969-987, 2012.

HUKIN, D. et al. Sensitivity of cell hydraulic conductivity to mercury is coincident with symplasmic isolation and expression of plasmalemma aquaporin genes in growing maize roots. **Planta**, v.215, p.1047–1056, 2002.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Levantamento Sistemático da Produção Agrícola (LSPA)**, 2016. Disponível em:  
<http://www.sidra.ibge.gov.br/bda/prevsaf/default.asp?t=2&z>

IBGE. **Instituto Brasileiro de Geografia e Estatística**. Levantamento Sistemático da Produção Agrícola, 2016. Disponível:  
<http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/default.shtm>.

IBGE. **Instituto Brasileiro de Geografia e Estatística**. Mapa de Biomas e de Vegetação, 2014. Disponível:  
<http://www.ibge.gov.br/home/presidencia/noticias/21052004biomashtml.shtm>

JARVIS A., RAMIREZ-VILLEGAS J., CAMPO B.V.H, NAVARRO-RACINES C., Is Cassava the Answer to African Climate Change Adaptation?. **Tropical Plant Biol.** v. 5, p.9-29, 2012.

JAVOT H, MAUREL C. The role of aquaporins in root water uptake. **Ann Bot.** v. 90, p.301–13, 2003.

JUKES TH & CANTOR CR. **Evolution of protein molecules**. In Munro HN, Mammalian Protein Metabolism, Academic Press, New York, p. 21-132, 1969.

JUNG, J. S. ET AL. Molecular characterization of an aquaporin cDNA from brain: candidate osmoreceptor and regulator of water balance. **Proc. Natl. Acad. Sci. USA.** v. 91, p. 13052–13056, 1994.

KALDENHOFF, R.; FISCHER, M. Functional aquaporin diversity in plants. **Biochimica et Biophysica Acta**, n. 1758, p. 1134–1141, 2006.

KAWASAKI, S. et al. Gene expression profiles during the initial phase of salt stress in rice. **Plant Cell**, v. 13, p. 889–905, 2001.

KEMPER B., et al. PathText: a text mining integrator for biological pathway visualizations. **Bioinformatics**. v. 26, p. 374–381, 2010.

KENT J. W. BLAT—The BLAST-Like Alignment Tool. **Genome Res.** v. 12(4), p.656–664, 2002.

KERBAUY, G.B. **Fisiologia Vegetal**. Editora Guanabara Koogan S.A., Rio de Janeiro, 452 p, 2004.

KIMURA M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. **Journal of Molecular Evolution.** v.16, p. 111-120, 1980.

KOJIMA M. et al. Highly sensitive and high-throughput analysis of plant hormones using MS-probe modification and liquid chromatography–tandem mass spectrometry: an application for hormone profiling in *Oryza sativa*, **Plant Cell Physiol.**, v.50, p.1201-1214. 2009.

KUMAR S,TAMURA K.,NEI M. MEGA3: an integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. **Brief Bioinform**,v. 5, p 150-163, 2004.

KUO-BIN LI. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. **Bioinformatics**, v. 19 n. 12 , p. 1585–1586, 2003.

LEAL, I. R, TABARELLI, M., SILVA, J. M. C. **Ecologia e Conservação da Caatinga**. 1<sup>a</sup> ed. Editora Universitária, UFPE, Recife, 2003.

LIAN H.L. et al. Upland rice and lowland rice exhibited different PIP expression under water deficit and ABA treatment. **Cell Res.** v.16, p. 651–60, 2006.

LISTER R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. **Nature**, v. 462(7271), p.315-22, 2009.

LIU Y, ZHOU J., WHITE K.P. RNA-seq differential expression studies: more sequence or more replication?. **Bioinformatics**, v. 30 (3), p. 301-304, 2014.

LOPEZ, G. et al. Insights into Populus XIP aquaporins: evolutionary expansion, protein functionality, and environmental regulation. **Journal of Experimental Botany**, v. 63, n. 5, p. 2217–2230, 2012.

LOQUE, D. et al. Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH<sub>3</sub> transport into the vacuole. **Plant Physiology**, v. 137, n. 2, p. 671-680, 2005.

LUCENA, M. F. A.; ALVES, M. Notas taxonômicas para Euphorbiaceae s.l. do Nordeste do Brasil. **Hoehnea**, v.37(1), p. 71-85, 2010.

- LUCENA, M. F. A.; AMORIM, B. S.; ALVES, M. Sinopse das espécies de Euphorbiaceae s. l. do Parque Nacional Serra de Itabaiana, Sergipe, Brasil. **Revista Caatinga**, v. 22(4), p. 214-224, 2009.
- SOUZA L.L, RHODEN S.A., PAMPHILE J.A.. A importância das Ômicas como ferramentas para o estudo da prospecção de microrganismos: perspectivas e desafios. **Revista Uningá Review**. v.18,n.2,p.16-21, 2014.
- MACIEL, B. A. **Uso sustentável e conservação dos recursos florestais da Caatinga**. Misistério do Meio Ambiente. Brasília-DF: Serviço Florestal Brasileiro. p 76-81, 2010.
- MAESHIMA M. et al. Accumulation of vacuolar H<sup>+</sup>-pyrophosphatase and H<sup>+</sup>-ATPase during reformation of the central vacuole in germinating pumpkin seeds. **Plant Physiol**, v.106, p. 61–69, 1994.
- MAESHIMA M. Tonoplast transporters: organization and function. Annu. Ver. **Plant Mol.** v. 52: p. 469-497, 2001.
- MAESHIMA, M.; ISHIKAWA, F. ER membrane aquaporins in plants. **European journal of physiology**, v. 456, n. 4, p. 709-16, 2008.
- MAGNI, F. et al. Proteomic knowledge of human aquaporins. **Proteomics**. v. 6, p. 5637-5649, 2006.
- MAHAJAN, S.; TUTEJA, N. Cold salinity and drought stresses: an overview. **Archives of Biochemistry and Biophysics**, v. 444, n. 139-158, 2005.
- MARCHLER-BAUER A., BRYANT S. H. Cd-Search: Protein Domain Annotations on the fly. **Nucleic Acids Res.**327-331, 2004.
- MAUREL C. et al. Molecular Physiology Of Aquaporins In Plants. **International Review Of Cytology**. v. 215, p. 105–148, 2002.
- MAUREL, C. Plant aquaporins: novel functions and regulation properties. **FEBS Letters**. v. 581, p. 2227–2236, 2007.
- MAUREL, C.; VERDOUCQ, L.; LUU, D.T.; SANTONI, M. Plant aquaporins: membrane channels with multiple integrated functions. **Annual Review of Plant Biology**, v. 59, p. 595-624, 2008.
- MEDEIROS J. A., ALOUFA M.A.I. Revegetação de área em processo de desertificação com a *C. quercifolius* (*Cnidoscolus quercifolius* Pohl) no município de São José do Seridó/RN. **Revista Brasileira de Geografia Física** v. 08, n. 04, p. 1158-1175, 2015.
- MELO, A. L.,SALES, M. F. O gênero *Cnidoscolus* Pohl (Crotonoideae-Euphorbiaceae) no Estado do Pernambuco, Brasil. **Acta Botanica Brasilica**. 22(3):806-827. 2008.
- MITANI-UENO N. et al .The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for sílicon, boron, and arsenic. **Journal of Experimental Botany** 62, 4391–4398, 2011.

MITTLER R, BLUMWALD E. Genetic engineering for modern agriculture: challenges and perspectives. **Annual Review of Plant Biology**, v. 61, p. 443–462, 2010.

MMA- **Ministério do Meio Ambiente**, 2010. Disponível em:  
[http://aprender.ead.unb.br/pluginfile.php/28053/mod\\_resource/content/1/Drummond\\_etal\\_2010\\_UC\\_legislacao\\_historico.pdf](http://aprender.ead.unb.br/pluginfile.php/28053/mod_resource/content/1/Drummond_etal_2010_UC_legislacao_historico.pdf)

MORAIS, D.A.L. **Análise bioinformática de genes de resistência a patógenos no genoma expresso da cana-de-açúcar**. Dissertação de Mestrado, UFPE, Recife, 2003.

NAKAMURA, Y.; COCHRANE, G.; KARSCH-MIZRACHI, I. The International Nucleotide Sequence Database Collaboration. **Nucleic Acids Research**, v. 41, n. D1, p. D21–D24, 2012.

NAKASHIMA, K.; ITO, Y.; YAMAGUCHI-SHINOZAKI, K. Transcriptional regulatory networks in response to abiotic stresses in *Arabidopsis* and grasses. **Plant physiology**, v. 149, n. 1, p.88-95, 2009.

NCBI- **National Center For Biotechnology Information**, ORF-Finder Open Reading Frame Finder, 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/orffinder>>

NEPOMUCENO, A.L. et al. Tolerância à seca em plantas: Mecanismos fisiológicos e moleculares. **Biociência e Desenvolvimento**, v. 23, n. 12-18, 2001

NEPOMUCENO, A.L.; NEUMAIER, N.; FARIAS, J.R.B.; OYA, T. Tolerância à seca em plantas: Mecanismos fisiológicos e moleculares. **Biociência e Desenvolvimento**, v. 23, n. 12-18, 2001.

NOGUEIRA, R.M.C. et al. **Estresses ambientais: danos e benefícios em plantas**. Imprensa Universitária, Recife, 2005.

OGUNNIYI DS. Castor oil: a vital industrial raw material. **Bioresour Technol**, v. 97(9), p. 1086–1091, 2006.

OLIVEIRA D. G. A família Euphorbiaceae Juss. em um fragmento de Caatinga em Sergipe, **Scientia Plena**, v. 9, n. 4, 2013

OLIVEIRA, D. M. et al. Intoxicação por *Cnidioscolus phyllacanthus* (Euphorbiaceae) em caprinos. **Pesquisa Veterinária Brasileira**. Vol. 28, n. 1, p. 36-42, 2008.

OLIVEIRA, E. J.; SANTANA, F. A.; OLIVEIRA, L. A.; SANTOS, V. S. Genotypic variation of traits related to quality of cassava roots using affinity propagation algorithm. **Scientia Agricola**, v. 72, n. 1, p. 53-61, 2015.

ORLANDIN, P.; LIMA, L. R. Sinopse do gênero *Manihot* Mill. (Euphorbiaceae) no Estado de São Paulo, Brasil. **Hoehnea**, v.41, p.51-60, 2014.

PAGE, R.D.M. TreeView: An application to display phylogenetic trees on personal computers. **Computer applications in the biosciences**, v. 12, p. 357-358, 1996.

PANDEY, A.; MANN, M. Proteomics to study genes and genomes. **Nature**, v. 405, p. 837-846, 2000.

PEARSON WR, LIPMAN DJ: Ferramentas melhoradas para comparação de seqüências biológicas. **Proc Natl Acad Sci.USA**, v. 85, p. 2444-2448, 1988.

PEIXOTO SOBRINHO et al, Phytochemical screening and antibacterial activity of four *Cnidocolus* species (Euphorbiaceae) against standard strains and clinical isolates. **Journal of Medicinal Plants Research**, v. 6(21), p. 3742-3748, 2012.

PIMENTA, A.M.C. **Os desafios do proteoma**. Ciência Hoje, Rio de Janeiro, v. 32, n. 192, p. 16- 22, 2003.

POMMERENIG B., DIEHN T. A., BIENERT G. P. Metalloido-porins: essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. (2015). **Plant Sci**. v. 238, p. 212–227, 2015.

PRESTON, G. M. et al. Mutations in aquaporin-1 in Phenotypically Normal Humans Without Functional CHIP Water Channels. **Science**, v. 265, n. 9, p. 1585-1587, 1994.

PROSDOCINI, F. et al. **Bioinformática: manual do usuário**. Biotecnologia Ciência e Desenvolvimento, v. 29, n. 12-25, 2002.

RAVEN, P.H.; EVERT, R.F.; EICHHORN, S.E.; **Biologia Vegetal**. 7 ed. Rio de Janeiro: Guanabara Koogan. 830p., 2007.

RIVAROLA M. et al. Castor bean organelle genome sequencing and worldwide genetic diversity analysis. **PLoS One**, v. 6(7), e21743, 2011.

SAITO K., MATSUDA F. Metacolumics for functional genomics, systems biology, and biotechnology. **Annual review of plant biology**, v. 61, p. 463-89, 2010.

SAITOU N, NEI M. The Neighbor-Joining method—a new method for reconstructing phylogenetic trees, **Mol Biol Evol**, vol. 4, 1987.

SAMPAIO, E. V. S. B. et al. **Desertificação no Brasil: conceitos, núcleos e tecnologias de recuperação e convivência**. Recife: Ed. Universitária da UFPE, 2003.

SAMPAIO, E. V. S.B. **Uso sustentável e conservação dos recursos florestais da Caatinga. Misistério do Meio Ambiente**. Brasília-DF: Serviço Florestal Brasileiro. p. 29-48. 2010.

SANTOS, M.F.A.V.; RIBEIRO, M.R.; SAMPAIO, E.V.S.B. Semelhanças vegetacionais em sete solos da caatinga. **Pesquisa Agropecuária Brasileira** 27: 305-314, 1992.

SÁTIRO L.N., ROQUE N., A família Euphorbiaceae nas caatingas arenosas do médio rio São Francisco, BA, Brasil. **Acta Bot. Bras**. v. 22(1), p.99-118. 2008.

SCHAFFNER, A.R.; BOUCHEZ, D. Role of a single aquaporin isoform in root water uptake. **Plant Cell**, v. 15, 509–522, 2003.



SEAB – Secretaria de Estado da Agricultura e do Abastecimento DERAL - Departamento de Economia Rural. **Análise da conjuntura agropecuária *M. esculenta* - safra 2015/16.** Disponível em: <[http://www.agricultura.pr.gov.br/arquivos/File/deral/Prognosticos/2016/M.esculenta\\_2015\\_16.pdf](http://www.agricultura.pr.gov.br/arquivos/File/deral/Prognosticos/2016/M.esculenta_2015_16.pdf)>

SEKI, M. et al. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold, and high-salinity stresses using a full-length microarray. **The Plant Journal**, v.31, p. 279-292, 2002.

SIEFRITZ, F. et al. PIP1 plasma membrane aquaporins in tobacco: from cellular effects to function in plants. **Plant Cell**, v. 14, p. 869-876, 2002.

SILVA, J. G. I.; SANTOS, M. R.; SOUSA, R. M.; PEREIRA, N. B. 11089 – Protocolo para propagação rápida de mandioca nas condições de Uruçuí-PI. **Cadernos de Agroecologia**, v. 6, n. 2, 2012.

SNEATH, P.H.A.; SOKAL, R.R. **Numerical taxonomy — the principles and practice of numerical classification.** (W. H. Freeman: San Francisco.), 1973.

SOUZA, B.B. et al., Utilização da favela (*Cnidocolus phyllacanthus*) como fonte de suplementação alimentar para caprinos e ovinos no semiárido brasileiro. **Agropecuária Científica no Semiárido**, v.8, p. 23-12, 2012.

SOUZA, V.C., LORENZI, H. **Botânica Sistemática - Guia ilustrado para a identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APG II.** Nova Odessa, Plantaru, 2008.

SUI, H. et al. Structural basis of waterspecific transport through the AQP1 water channel. **Nature**, v. 414, p. 872–878, 2001.

TAJIMA F & NEI M. Estimation of evolutionary distance between nucleotide sequences. **Molecular Biology and Evolution**, v.1, p. 269-285, 1984.

TAMURA K. et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. **Molecular Biology and Evolution**, v.24, p.1596-1599, 2007.

TAMURA K. et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Mol Biol Evol.**, v. 30(12), p. 2725–2729, 2013.

TAMURA N, et al. synthesis and glutamate-agonistic activity of (s)-2-amino-3-(2,5-dihydro-5-oxo-3-isoxazolyl)-propanoic acid derivatives. **Chem Pharm Bull.**v. 40(2), p. 381-6, 1992.

TATENO, Y. et al. DNA databank of Japan (DDBJ) for genome scale research in life science. **Nucleic Acids Research**, v. 25, p. 4876- 4882, 2002.

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, J. et al. The Clustal\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, v. 25, p. 4876-4882, 1997.

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 Res.**, v. 22, p. 4673–4680, 1994.

VINNAKOTA R. et al. A comparison of aquaporin function in mediating stomatal aperture gating among drought-tolerant and sensitive varieties of rice (*Oryza sativa* L.). **Protoplasma**, v. 253(6), p. 1593-1597, 2015.

WALLACE, I. S.; CHOI, W.-GYU; ROBERTS, D. M. The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins. **Society**, v. 1758, p. 1165 – 1175, 2006.

WALLACE, I. S.; ROBERTS, D. M. Homology Modeling of Representative Subfamilies of *Arabidopsis* Major Intrinsic Proteins. Classification Based on the Aromatic / Arginine Selectivity Filter 1 [ w ]. **Proteins**, v. 135, p. 1059-1068, 2004.

WANG, W.X. et al. Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. **Acta Horticulturae**, v. 560, p. 285-292, 2001.

WEBSTER, G.L. **Systematics of the Euphorbiaceae**. Annals of the Missouri Botanical Garden, v. 81(1), p. 1-144 (1994).

WUDICK, M.M.; LUU, D.T.; MAUREL, C. A look inside: localization patterns and functions of intracellular plant aquaporins. **New Phytology**, v. 184, p. 289-302, 2009.

ZARDOYA, R. VILLALBA, S. A Phylogenetic Framework for the Aquaporin Family in Eukaryotes. **J Mol Evol**, v. 52, p. 391–404, 2001.

ZELAZNY, E. et al. FRET imaging in living maize cells reveals that plasma membrane aquaporins interact to regulate their subcellular localization. **Proc. Natl. Acad. Sci. USA**, v. 104, p. 12359–12364, 2007.

ZENG XL. Research progress for resistance and tolerance mechanism of castor. Chinese Agricultural Science Bulletin. v. 26(4), p. 123–125, 2010.

ZHAO, C.X.; SHAO, H.B.; CHU, L.Y. Aquaporin structure – function relationships: Water flow through plant living cells. **Colloids and Surfaces**, v. 62, p. 163-172, 2008.

ZOU, Z. et al. Gene structures, evolution, classification and expression profiles of the aquaporin gene family in castor bean (*Ricinus communis* L.). **PLoS ONE** 10:e0141022, 2015b.

ZOU, Z. et al. Genomewide identification of *Jatropha curcas* aquaporin genes and the comparative analysis provides insights into the gene family expansion and evolution in *Hevea brasiliensis*. **Front. Plant Sci.** v. 7, p.395, 2016.

ZOU, Z. et al. Genome-wide identification of rubber tree (*Hevea brasiliensis* Muell. Arg.) aquaporin genes and their response to ethephon stimulation in the laticifer, a rubber-producing tissue. **BMC Genomics**, v. 16,p.1001, 2015a.

## CAPÍTULO 1

**CARACTERIZAÇÃO DE AQUAPORINAS *in silico* DE *Manihot esculenta* Crantz E *Ricinus communis* L. PARA O DESENHO DE INICIADORES E ISOLAMENTO DE SEQUÊNCIAS EM *Cnidoscolus quercifolius* Pohl**

Nóbrega, Gisela Formiga Queiroz. **CARACTERIZAÇÃO DE AQUAPORINAS *in silico* DE *Manihot esculenta* CRANTZ e *Ricinus communis* L. PARA O DESENHO DE INICIADORES E ISOLAMENTO DE SEQUÊNCIAS EM *Cnidocolus quercifolius* POHL.** Dissertação de Mestrado Em Ciências Florestais. CSTR/UFCG, Patos - PB. 2017. 104p.

## RESUMO

Euphorbiaceae Juss. destaca-se como uma das famílias botânicas de maior importância econômica, com expressiva diversidade de espécies e ampla variedade de hábitos, o que confere uma maior capacidade adaptativa e melhor eficiência no uso da água. Espécies de euforbiáceas representam importantes fontes de genes com alto potencial biotecnológico, principalmente aqueles envolvidos nos processos de captação de água, que permitem a adaptação das plantas às condições de restrição hídrica. Entre os genes responsivos ao estresse estão as aquaporinas membros da superfamília MIP (*Major Intrinsic Protein*), proteínas canais responsáveis pelo transporte de água, gases, íons e outros pequenos solutos. Através dos bancos de dados genômicos de *M. esculenta* (*Manihot esculenta*) e *R. communis* (*Ricinus communis*), foram identificadas 79 aquaporinas, sendo 43 em *M. esculenta* e 36 em *R. communis*. A partir de análises fenéticas, as sequências de aminoácidos foram classificadas nas cinco subfamílias já descritas em vegetais superiores (24 PIPs, 22 TIPs, 16 NIPs, 7 SIPs e 9 XIPs). As sequências foram caracterizadas quanto à presença de domínios e resíduos conservados, massa molecular, ponto isoelétrico, localização subcelular e distribuição cromossômica. Foram selecionados membros de cada uma das cinco subfamílias de cada espécie para realização da modelagem comparativa. De posse das sequências devidamente classificadas e caracterizadas, foi possível a obtenção de 16 pares de iniciadores, além de outros 15 pares desenhados com base em sequências expressas de feijão-caupi (*V. unguiculata*), os quais foram utilizados na amplificação de sequências homólogas de aquaporinas em *C. quercifolius* (*Cnidocolus quercifolius*), uma espécie nativa da Caatinga, extremamente tolerante ao déficit hídrico. Dos 31 pares de iniciadores testados, 14 amplificaram 17 sequências. Destas, sete foram confirmadas (via sequenciamento e Blastn) com similaridade (média de 89%) às aquaporinas, as quais foram classificadas com base na homologia de sequências e nos dendogramas gerados nas análises fenéticas. Foram detectadas subfamílias PIP (3 membros) e TIP (4 membros). Tendo em vista que a *C. quercifolius* não dispõe de informações biológicas em bancos de dados e o fato da alta conservação entre as aquaporinas vegetais, neste estudo, foi possível, não somente a obtenção de aquaporinas entre 3 espécies diferentes de euforbiáceas (*M. esculenta*, *R. communis* e *C. quercifolius*), mas entre dois membros de duas diferentes famílias: o feijão-caupi (Fabaceae) e a *C. quercifolius* (Euphorbiaceae), refletindo o potencial de transferibilidade de iniciadores entre essas importantes espécies vegetais. Embora não tenha sido possível a obtenção de sequências completas de aquaporinas em *C. quercifolius* (apenas parcialmente), com base nos resultados obtidos, é possível confirmar que, além de presente no seu genoma, as aquaporinas de *C. quercifolius* estão sendo expressas, uma vez que as amplificações foram observadas no seu genoma expresso (cDNA).

**PALAVRAS-CHAVE:** Bioinformática, Euphorbiaceae, Transportadores de membrana

Nóbrega, Gisela Formiga Queiroz. **CHARACTERIZATION OF AQUAPORINS in silico of *Manihot esculenta* CRANTZ and *Ricinus communis* L. FOR THE DESIGN OF INITIATORS AND THE SEQUENCE ISOLATION IN *Cnidoscolus quercifolius* POHL.** Master's degree Thesis in Forest Sciences. CSTR/UFCG, Patos - PB. 2017. 104p.

### ABSTRACT

Euphorbiaceae Juss stands out as one of the most economically important botanical species, with an expressive diversity of types and a wide variety of habits, which allows a greater capacity of adaptation as well as better techniques for water use. Euphorbiaceous species represent important sources of genes with high biotechnological potential, especially those involved in the processes of water abstraction that allows adaptation of plants to water restriction conditions. Among stress responsive genes, there are the aquaporins - members of the MIP (*Major Intrinsic Protein*) superfamily, which are channel proteins responsible for transportation of water, gases, ions and other small solutes. Through the genomic databases of cassava (*Manihot esculenta*) and castor bean (*Ricinus communis*), 79 aquaporins were identified, where 43 were in the cassava and 36 in the castor bean. From phenotypic analyzes, amino acid sequences were classified in the five subfamilies already described in higher plants (24 PIPs, 22 TIPs, 16 NIPs, 7 SIPs and 9 XIPs). Sequences were characterized according to the presence of domain and conserved residues, molecular weight, isoelectric point, subcellular location and chromosomal distribution. Members of each of the five subfamilies of each species were selected for comparative modeling. In possession of the properly classified and characterized sequences, it was possible to obtain 16 primers of indicators, as well as 15 other designed pairs based on the expressed sequences of cowpea (*V. unguiculata*) that were used in the amplification of homologous sequences of the aquaporins in *Cnidoscolus quercifolius*, a native species of the Caatinga, extremely tolerant to water shortage. Out of 31 pairs of primers tested, 14 amplified 17 sequences. From these, seven were confirmed (through the sequencing method and the Blastn online tool) with an average similarity of 89% to aquaporins, which were classified based on homology sequence and on the dendrograms generated in phenotypic analyzes. It was also detected PIP (3-member) and TIP (4-member) subfamilies. Considering that the *C. quercifolius* has no biological data in databases and the high conservation among aquaporins, in this study, it was possible not only to obtain aquaporins from 3 different species of euphorbiaceae (*M. esculenta*, *R. communis* e *C. quercifolius*), but also from two members of two different families: the cowpea (Fabaceae) and the *C. quercifolius* (Euphorbiaceae), thus demonstrating the potential of transferability between those important plant species. Although it was not possible to obtain complete sequences of aquaporins in *C. quercifolius* (only partially), based on the achieved results, it was possible to confirm that, in addition to being present in their genome, the aquaporins from *C. quercifolius* are being expressed, as the amplifications were observed in their genome-express (cDNA).

**Keywords:** Bioinformatics, Euphorbiaceae, Membrane Transporters.

## 1 INTRODUÇÃO

Os fatores de estresses abióticos atuam sobre as plantas causando diversos danos ao seu desenvolvimento e produtividade, muitas vezes, comprometendo a sobrevivência, afetando interações entre espécies e alterando processos ambientais (HELLMANN et al., 2008). O déficit hídrico destaca-se como um dos principais fatores que limitam o crescimento vegetal, já que este estresse inviabiliza a realização de atividades importantes, como a fotossíntese e a hidrólise de amido (NEPOMUCENO et al., 2001), interferindo, também, na disponibilidade de nutrientes, aumento da salinidade do solo, favorecendo a ocorrência de outros estresses simultaneamente (ZHU, 2002; MAHAJAN; TUTEJA, 2005).

Para conseguir sobreviver em situações de pouca disponibilidade de água, as plantas desenvolveram mecanismos que lhes permitem ajustar o metabolismo de acordo com as necessidades. Estes ajustes podem ocorrer por meio de alterações fisiológicas, bioquímicas, morfológicas ou moleculares (MCDOWELL et al., 2008; CLAEYS; INZE, 2013). Tais mecanismos refletem a importância do patrimônio genético da espécie, assim como o estágio de desenvolvimento no qual a planta se encontra frente ao estresse (WANG et al., 2001; WANG et al., 2003; BENKO-ISEPPON et al., 2011).

De forma geral, estresses abióticos induzem a ativação de genes de vias de sinalização celular (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000; ZHU, 2002), muitos destes genes codificam para proteínas diretamente envolvidas na tolerância ao estresse (BENKO-ISEPPON et al., 2011). Entre os produtos de genes responsivos ao estresse abiótico em plantas, destacam-se a família das aquaporinas (SHINOZAKI et al., 2003; PARK; CAMPBELL, 2015; AFZAL et al., 2016). Igualmente conhecidas como proteínas canais, as aquaporinas pertencem a uma superfamília altamente conservada de Proteínas Intrínsecas de Membrana – MIPs (*Major Intrinsic Protein*) (AGRE, 2004). Estão presentes em todos os tecidos vegetais, sendo reguladas espaço-temporal, dependendo do estágio de desenvolvimento e das condições ambientais, atuando na adaptação das plantas frente aos estresses como salinidade e déficit hídrico (WALLACE et al., 2006; WUDICK; LUU; MAUREL, 2009; HEINEN; CHAUMONT, 2009).

A vegetação da Caatinga é bastante variável, de constituição arbustiva-arbórea, com predominância de leguminosas, cactáceas e euforbiáceas (GOLFARI; CASER, 1977; DRUMMOND et al., 2000). A família Euphorbiaceae Juss. possui 8.000 espécies descritas, distribuídas em 317 gêneros (ENDRESS et al., 2013; BIJEKAR; GAYATRIM, 2014). No Brasil, estima-se a ocorrência de 1.100 espécies e 72 gêneros (BARROSO et al., 2002;

SOUZA; LORENZI, 2008), dos quais, 240 são encontradas na região Nordeste (LUCENA; ALVES, 2010) e, em grande parte, na Caatinga, além de muitas espécies endêmicas (SÁTIRO; ROQUE, 2008; FORZZA et al., 2010).

Muitas espécies de euforbiáceas apresentam importância socioeconômica para o mundo, entre as quais se destacam a *Manihot esculenta* Crantz (CORRÊA et al., 2002) e a *Ricinus communis* L. (OGUNNIYI, 2006). Para a região Nordeste do Brasil, destaca-se a *Cnidoscolus quercifolius* Pohl, espécie endêmica da Caatinga (MELO; SALES, 2008; PEIXOTO SOBRINHO et al., 2012). Membros de euforbiáceas apresentam alta capacidade adaptativa à seca (MELO; SALES, 2008; QIU, 2010; RIVAROLA, 2011; JARVIS et al., 2012), embora muitos mecanismos adaptativos permaneçam ainda pouco elucidados. A escassez de informações é ainda maior quando relacionada a espécies endêmicas (SAMPAIO, 2010).

Desta forma, as espécies da Caatinga compõem uma importante fonte de genes com alto potencial biotecnológico, principalmente aqueles que participam direta e/ou indiretamente dos mecanismos de adaptação das plantas frente às condições de restrição hídrica. Neste contexto, estudos envolvendo o uso de tecnologias moleculares em larga escala, aliados ao grande avanço nas plataformas de sequenciamento de alto rendimento, têm permitido a descoberta de um grande número de genes em um curto espaço de tempo (SEKI et al., 2002). Por outro lado, estas informações, muitas vezes, são pouco exploradas, ou, ainda, limitadas a organismo de interesse. Tão importante quanto à geração de dados biológicos é a extrapolação destas informações.

Tendo em vista a correlação existente entre as aquaporinas vegetais com os mecanismos de tolerância à seca, este trabalho teve como objetivo isolar genes homólogos de aquaporinas em *Cnidoscolus quercifolius*, constituindo o primeiro relato dessa importante família gênica nessa espécie.

## 2. MATERIAIS E MÉTODOS

### 2.1 Identificação e caracterização de MeAQPs e RcaAQPs

#### 2.1.1 Seleção de sondas e busca por aquaporinas homólogas em *M. esculenta* e *R. communis*

A seleção de sequências sondas para aquaporinas foi realizada por meio da busca por palavras-chave específicas nos genomas de organismos de referência: *Arabidopsis thaliana*, *Zea mays*, *Populus trichocarpa*, *Gycine max* e *Lotus japonicus*, sendo obtidas no GeneBank/NCBI (*National Center for Biotechnology Information*; [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). Essas sequências foram utilizadas como sequências sondas (*seed sequences*) para a busca por seus homólogos nos genomas de *Manihot esculenta* e *Ricinus communis*, disponíveis no Phytozome v12.0 (<https://phytozome.jgi.doe.gov/pz/portal.html>), utilizando a ferramenta tBLASTn para o alinhamento local, utilizando o valor de E-value  $< 10^{-5}$  como ponto de corte.

#### 2.1.2 Caracterização de sequências

As sequências candidatas, obtidas nos genomas das espécies de interesse como aquaporinas, foram traduzidas para identificação da melhor pauta aberta de leitura - ORF (*Open Reading Frame*) com a ferramenta *ORF-Finder* (<http://www.ncbi.nlm.nih.gov/orffinder/>). Em seguida, através da ferramenta Batch CD-Search (<http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>), as sequências de aminoácidos foram avaliadas de acordo com a presença de assinaturas, domínios, motivos e resíduos conservados, bem como a integridade dessas regiões.

A partir dos candidatos traduzidos em aminoácidos, foram realizados alinhamentos múltiplos com as sequências de *M. esculenta* e de *R. communis*, utilizando a ferramenta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>), com visualização gerada por meio do programa JalView (WATERHOUSE et al., 2009), a fim de se avaliar o grau de conservação e identificar as posições dos resíduos conservados entre as sequências. As sequências de aminoácidos também foram submetidas a uma eletroforese bidimensional (*in silico*), via ferramenta web JVirGel (<http://www.jvirgel.de/>), para determinação das massas moleculares e pontos isoelétricos das sequências protéicas candidatas (HILLER et al., 2006).



### 2.1.3 Alinhamento de sequências e análise fenética

O grau de similaridade entre as sequências, bem como a determinação da nomenclatura para as supostas aquaporinas, nas diferentes subfamílias, foram determinados tomando por base os dendrogramas construídos por meio do método de *Neighbor-Joining*, com valores de *bootstrap* ao nível de 1.000 replicações, utilizando o programa MEGA6 (TAMURA, 2013) para a geração dos dendrogramas, com visualização destes por meio do software TreeView (PAGE, 1996).

A análise fenética foi realizada a partir de sequências de aquaporinas de monocotiledôneas e dicotiledôneas, juntamente com as aquaporinas identificadas em *M. esculenta* e, posteriormente, com as de *R. communis*. As sequências já identificadas de *Arabidopsis thaliana*, *Zea mays*, *Glycine max*, *Populus trichocarpa* e *Lotus japonicus* foram utilizadas na construção dos dendrogramas, a fim de verificar a relação entre as diferentes subfamílias de aquaporinas vegetais. A nomenclatura sugerida para as aquaporinas encontradas no genoma da *M. esculenta* e da *R. communis* foi baseada na homologia de sequências e nos dendrogramas gerados nas análises fenéticas, seguindo a regra de nomenclatura já utilizada em outros organismos vegetais (JOHANSON et al., 2001).

### 2.1.4 Distribuição cromossômica e localização subcelular

Utilizando a plataforma online do genome browser Phytozome, foram anotadas as informações acerca da posição dos diferentes locus gênicos codificantes para aquaporinas apenas dentro do genoma da *M. esculenta*, uma vez que, para o genoma da *R. communis*, ainda não existe representação em cromossomos no Phytozome. Adicionalmente foi realizada a predição da localização subcelular das sequências proteicas de aquaporinas, identificadas, em ambas as espécies deste trabalho, por meio da ferramenta Cell-Ploc (CHOU; SHEN, 2008; <http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>).

### 2.1.5 Modelagem comparativa

As sequências de aminoácidos preditas para aquaporinas em *M. esculenta* e *R. communis* foram submetidas a uma modelagem comparativa, visando à determinação de sua estrutura terciária, representando a primeira modelagem para os membros desta família de transportadores nestas espécies. Como modelo para a modelagem tridimensional, foram

utilizados os *templates*, 1h6i.1.A, 1j4n.1.A, 1rc2.2.A, 2b5f.1.C, 3cn5.1.A, 3ne2.1.A e 5i32.1.A de aquaporina depositados no PDB (*Protein Data Bank*) e obtidos por métodos experimentais, tais como a difração de raios-X. A modelagem foi realizada na plataforma SWISS-MODEL Workspace (<http://swissmodel.expasy.org/>) (LASKOWSKI et al., 1993), com as informações acerca da qualidade do modelo determinadas via ProSA (<https://prosa.services.came.sbg.ac.at/prosa.php>), uma ferramenta *online* que compara os padrões de energia e estrutura do modelo com outras proteínas depositadas no banco de dados do PDB (SIPPL, 1993; WIEDERSTEIN; SIPPL, 2007). A avaliação de qualidade foi determinada através da análise do gráfico de Ramachandran e QMEAN (BENKERT; SCHWEDE; TOSATTO, 2009), que determinam a qualidade estereoquímica do modelo.

## 2.2 Obtenção e caracterização de aquaporinas em *C. quercifolius* – CqAQPs

### 2.2.1 Desenho de iniciadores

As sequências de aquaporinas identificadas em *M. esculenta* e *R. communis* foram utilizadas para o desenho de iniciadores através da ferramenta Primer3 (<http://frodo.wi.mit.edu/primer3/>), seguindo os seguintes parâmetros: tamanho (18 - 22 pb), temperatura de anelamento (57° C -63°C) e conteúdo GC (40 % - 60 %) (ROZEN; SKALETSKY, 2000). Os pares de iniciadores foram desenhados com base em sequências de PIP1, PIP2, TIP1, TIP2, TIP3, NIP1, NIP2 e XIP1, abrangendo as cinco subfamílias MIPs presentes nestas espécies. Adicionalmente, iniciadores para cDNA de aquaporinas, desenhados com base em sequências de *Vigna unguiculata*, foram disponibilizados pela equipe do Laboratório de Genética e Biotecnologia Vegetal, da Universidade Federal de Pernambuco – UFPE, sendo incluídos nesta análise.

Para as ampliações das aquaporinas em *C. quinidoscolus*, foram utilizados 31 pares de iniciadores (16 para DNA genômico, desenhados com base nas sequências de *R. communis* e *M. esculenta*, e 15 para cDNA, desenhados com base em sequências de *V. unguiculata*), representantes das subfamílias PIP (10 pares), TIP (11 pares), NIP (6 pares), SIP (2 pares) e XIP (2 pares).

### 2.2.2 Material vegetal

Folhas jovens de plantas adultas (média de 3 metros de altura) da espécie *C. quercifolius* com espinho (Reg. 5938/ CSTR-UFCG) foram coletadas de um conjunto de plantas do Centro de Saúde e Tecnologia Rural da UFCG, Patos-PB, Brasil (07.03'29.3"S, 037.16'30.4" W e 275m de altitude) em junho de 2016. A coleta do material vegetal foi realizada com base na integridade física da planta, evitando aquelas com desgastes causados pelo ataque de animais e/ou patógenos em geral. O material foi armazenado em solução saturada de NaCl e CTAB (*cetyl trimethyl ammonium bromide*) (20 g NaCl, 350 g e CTAB, 20g, por litro de solução) (ROGSTAD, 1992) e mantidas a 4°C até a extração do DNA.

### 2.2.3 Extração de DNA

A extração do DNA de folhas de *C. quercifolius* foi realizada utilizando o método descrito por Weising et al. (2005), com algumas modificações. Após o processo de extração, o DNA foi tratado para eliminação de polissacarídeos segundo protocolo de Michaels et al. (1994). As reações de PCR foram conduzidas em uma solução contendo: 20 ng de DNA, 2,5 µL de solução tampão 1X, 1,5 µL de MgCl<sub>2</sub> (25 mM) e 0,2 µL de Taq DNA polimerase (0,5 U/reação) 0,3 µL de dNTPs (10 mM), 0,75 µL de cada iniciador (F e R, 5 mM) num volume final de 25 uL com água ultrapura estéril.

### 2.2.4 Extração de RNA e síntese de cDNA

Para obtenção de RNA total, utilizaram-se 150mg de tecido foliar macerado na presença de nitrogênio líquido. O processo de extração foi realizado mediante a utilização do kit comercial *SV Total RNA Isolation System* (Promega), seguindo as recomendações do fabricante. O nível de concentração das amostras foi verificado por meio da técnica de fluorimetria (Qubit-Invitrogen), e a qualidade das mesmas foi avaliada através da técnica de eletroforese em gel de agarose a 2,0%.

Para realização da síntese de cDNA, foram utilizadas amostras com 1µg do RNA total. O processo foi realizado com o kit GoScript™ Reverse Transcriptase (Promega), de acordo com o protocolo do fabricante, sendo as amostras conduzidas em um termociclador "TGradiente" (Biometra), com a seguinte programação: 5 minutos a 70°C, 5 minutos a 4°C, 5 minutos a 25°C, uma hora a 42°C e 15 minutos a 70°C.

### 2.2.5 Reação em Cadeia da Polimerase – PCR

Todas as reações de PCR foram realizadas em termociclador “TGradiente” (Biometra). O programa foi configurado para 4 min a 95°C (desnaturação inicial), seguido de 40 ciclos de: 1 min a 95°C (desnaturação), 1 min à temperatura de anelamento para cada par de iniciador e 1 min a 72°C (extensão), com uma extensão final de 72°C por 5 min. Os produtos das ampliações foram visualizados em eletroforese 2 %, em tampão TBE 1X (10 mM Tris-borato, 1 mM EDTA, pH 8.0) a 120 V, usando como padrão de massa molecular *DNA Ladder* de 100 pb (Ludwig Biotec). Em seguida, os produtos (amplicons) foram purificados através do kit “GFX PCR DNA e Gel Band Purification Kit” (GE Healthcare, USA), de acordo com protocolo do fabricante. A quantificação foi realizada em fluorômetro Qubit (Invitrogen).

### 2.2.6 Confirmação das *CqAQP*s via sequenciamento genético

A reação de sequenciamento foi realizada com 20 ng de DNA ou 1 ng cDNA, 0,7 µL de Big-dye, versão 3.1, 3,2 pmoles de iniciador, 150 µL de tampão de sequenciamento e completado para 10 µL com água Milli-Q estéril. A reação de PCR de sequenciamento foi realizada em equipamento ABI9700 (GeneAmp PCR SYSTEM Applied Biosystems). Ao final desta etapa, a reação foi precipitada com 90 µL, uma solução composta de 3µL de acetado de potássio a 3M, 62,5 µL etanol absoluto, 24,5 µL de H<sub>2</sub>O Milli-Q estéril em cada tubo. Após agitação suave por 15s, as placas foram incubadas em gelo por 20 min, sendo em seguida centrifugadas por 30 min / 3.000g / 4°C. Descartado o sobrenadante, os precipitados foram lavados em 100 µL de etanol 70 % gelado e centrifugados por mais 10 min, nas mesmas condições. Após esse processo, a placa foi seca no escuro por 1 h. Em seguida, foram adicionados 10 µL de formamida Hi-Di (Applied Biosystems) em cada amostra, sendo a placa submetida a eletroforese capilar no sequenciador automático ABI 3100 (Applied Biosystems).

### 2.2.7 Obtenção e análise de bioinformática das sequências *CqAQP*

As sequências de nucleotídeos geradas no sequenciamento foram avaliadas quanto à qualidade (Phred/Phrap  $\geq 30$ ), via ferramenta BioEdit 7.0 (HALL, 1999). As sequências consenso foram obtidas no seu formato FASTA, para serem confrontadas com o banco de dados (não redundante, *nr*) do NCBI (*National Center for Biotechnology Information*;

www.ncbi.nlm.nih.gov/Blast), utilizando a ferramenta BLASTn (ALTSCHUL et al., 1990), a fim de se realizar a anotação funcional das mesmas.

Após a anotação, as sequências nucleotídicas foram traduzidas em todos os seis quadros abertos de leitura (ORFs), através da ferramenta *ORF-Finder* (<http://www.ncbi.nlm.nih.gov/orffinder/>). A presença de domínio conservado foi avaliada via ferramenta Batch CD-Search (<http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>). As sequências de aminoácidos deduzidas das *CqAQP* foram alinhadas com AQPs de *Arabidopsis thaliana*, usando o programa ClustalW (THOMPSON et al. 1994), gerando uma árvore fenética, construída pelo programa MEGA 6.0 (TAMURA, 2013), utilizando-se para isto, o método de *Neighbor-Joining* (SAITOU, 1987), com *bootstrap* ao nível de 1.000 replicações, sendo os dados visualizados no TreeView (PAGE, 1996). A nomenclatura sugerida para as aquaporinas de *C. quercifolius* foi baseada na homologia de sequências, seguindo a regra de nomenclatura já utilizada em outros organismos vegetais (JOHANSON et al., 2001).

### 3 RESULTADOS

#### 3.1 Busca e classificação das MeMIPs e RcMIPs

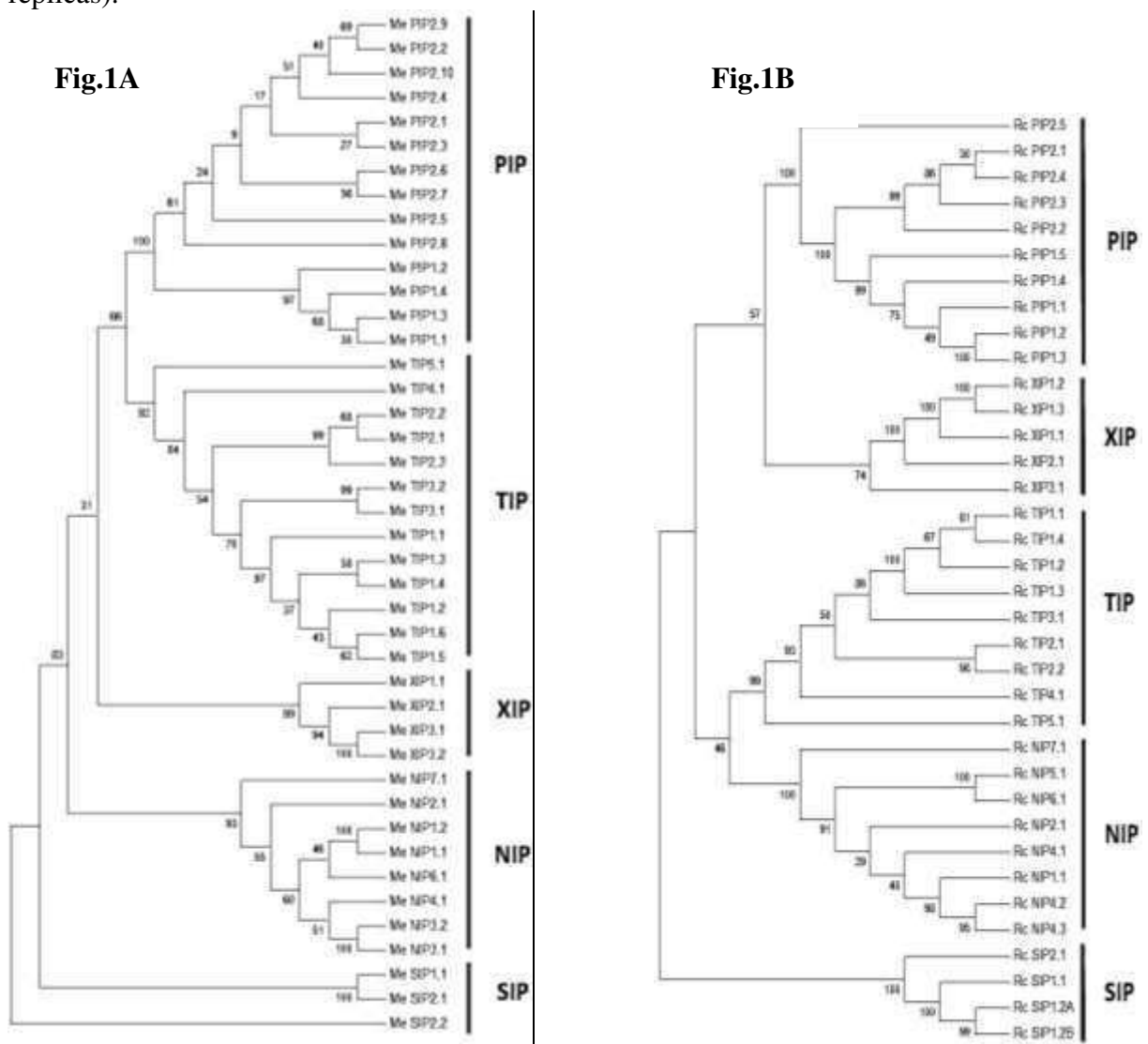
Através da busca por meio de palavras-chave, foram identificadas 69 sequências de aquaporinas (entre monocotiledôneas e dicotiledôneas), sendo: 35 em *Arabidopsis thaliana*, 16 em *Zea mays*, 12 em *Gycine max*, 5 em *Populus trichocarpa* e 1 em *Lotus japonicus*, abrangendo todas as subfamílias e suas subdivisões, as quais foram usadas como sondas para busca das sequências de interesse, bem como para construção de fenogramas, utilizados para identificar e classificar as sequências analisadas.

No Phytozome v12.0, após a realização do tBlastn, utilizando as sequências sondas contra o genoma de *M. esculenta* e *R. communis*, foi obtido um total de 82 sequências candidatas a genes codificantes para aquaporinas, em *M. esculenta* e *R. communis*, respectivamente. Destas, somente 4 sequências foram descartadas (não apresentando os critérios de qualidade pré-definidos, como e-value  $<10^{-5}$ ), resultando em 78 sequências, sendo 42 de *M. esculenta* e 36 de *R. communis*.

As relações fenéticas entre as aquaporinas de referência e as sequências encontradas nos genomas das duas espécies de interesse, observadas através da árvore gerada pelo método de *Neighbor-Joining*, permitiram agrupar as sequências de *M. esculenta* e *R. communis* nas cinco subfamílias de aquaporinas (PIP, TIP, NIP e SIP, considerados grupos ortólogos, e XIP,

um grupo de proteína intrínseca de membrana ainda não caracterizada), já descritas para outras espécies de vegetais superiores. Com base nesta análise e na homologia entre as sequências, foi sugerida a nomenclatura para as sequências de *M. esculenta* e *R. communis*, de acordo com a nomenclatura padrão para MIPs (Figura 1) (CHAUMONT et al., 2001; JOHANSON et al., 2001; DANIELSON; JOHANSON, 2008; BIENERT et al., 2011; LOPEZ et al., 2012).

Figura 1. Análise fenética de membros das famílias de aquaporinas de *M. esculenta* (Fig. 7A) e de *R. communis* (Fig.7B), pelo método de *Neighbor Joining* (*bootstrap* ao nível de 1.000 réplicas).



Com base nos dendogramas, foi possível classificar as 42 candidatas existentes no genoma da *M. esculenta* em 14 PIP (dez PIP1 e quatro PIP2); 13 TIPs (seis TIP1, três TIP2 e quatro outras TIPs); oito NIPs (duas NIP1, duas NIP3 e quatro outras NIPS); três SIPs e

quatro XIPs (Fig. 1A). As 36 seqüências de *R. communis* foram classificadas em dez PIPs (cinco PIP1 e cinco PIP2); nove TIPs (quatro TIP1, duas TIP2 e três outras TIPs); oito NIPs (três NIP4 e cinco outras NIPs); quatro SIPs e cinco XIPs (Fig. 1B), semelhante ao descrito por Zou et al. (2015).

As MIPs de *Manihot* (MeMIPs) e *Ricinus* (RcMIPs) foram analisadas separadamente por subfamília e de forma comparativa às MIPs de *Arabidopsis thaliana* (AtMIPs), para confirmação da divisão entre as cinco subfamílias e seus subgrupos (Figuras 2 e 3).

Figura 2. Análise fenética das aquaporinas preditas em *M. esculenta* (MeMIPs) e de *A. thaliana* (AtMIPs), por subfamília, separadamente. (A) MePIP e AtPIP. (B) MeTIP e AtTIP. (C) MeNIP e AtNIP. (D) MeSIP e AtSIP. (E) MeXIP e LjXIP. A escala de 0,05 refere-se a 5%, 0,02 a 2% e 0,1 a 10% de divergência entre as seqüências.

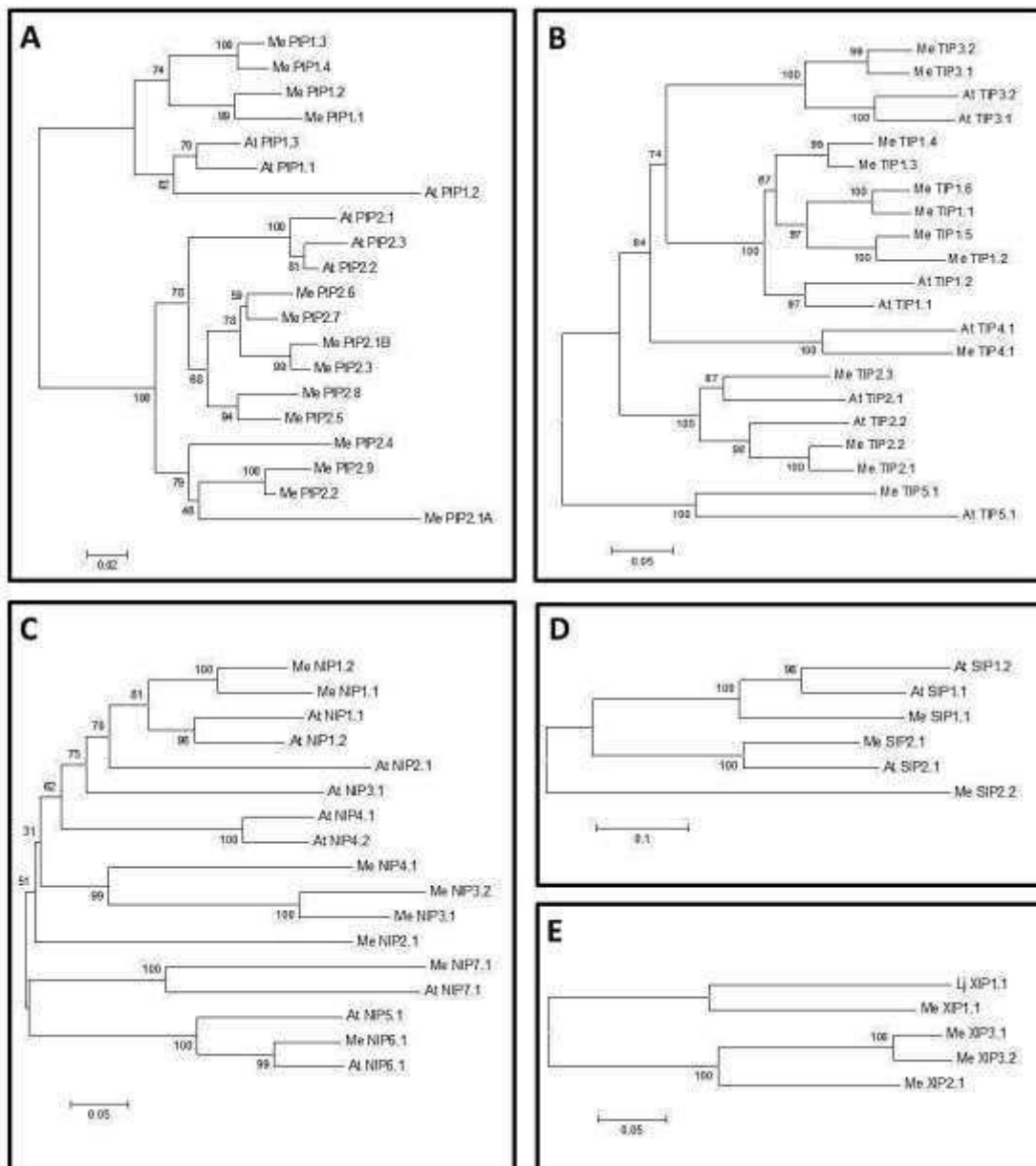
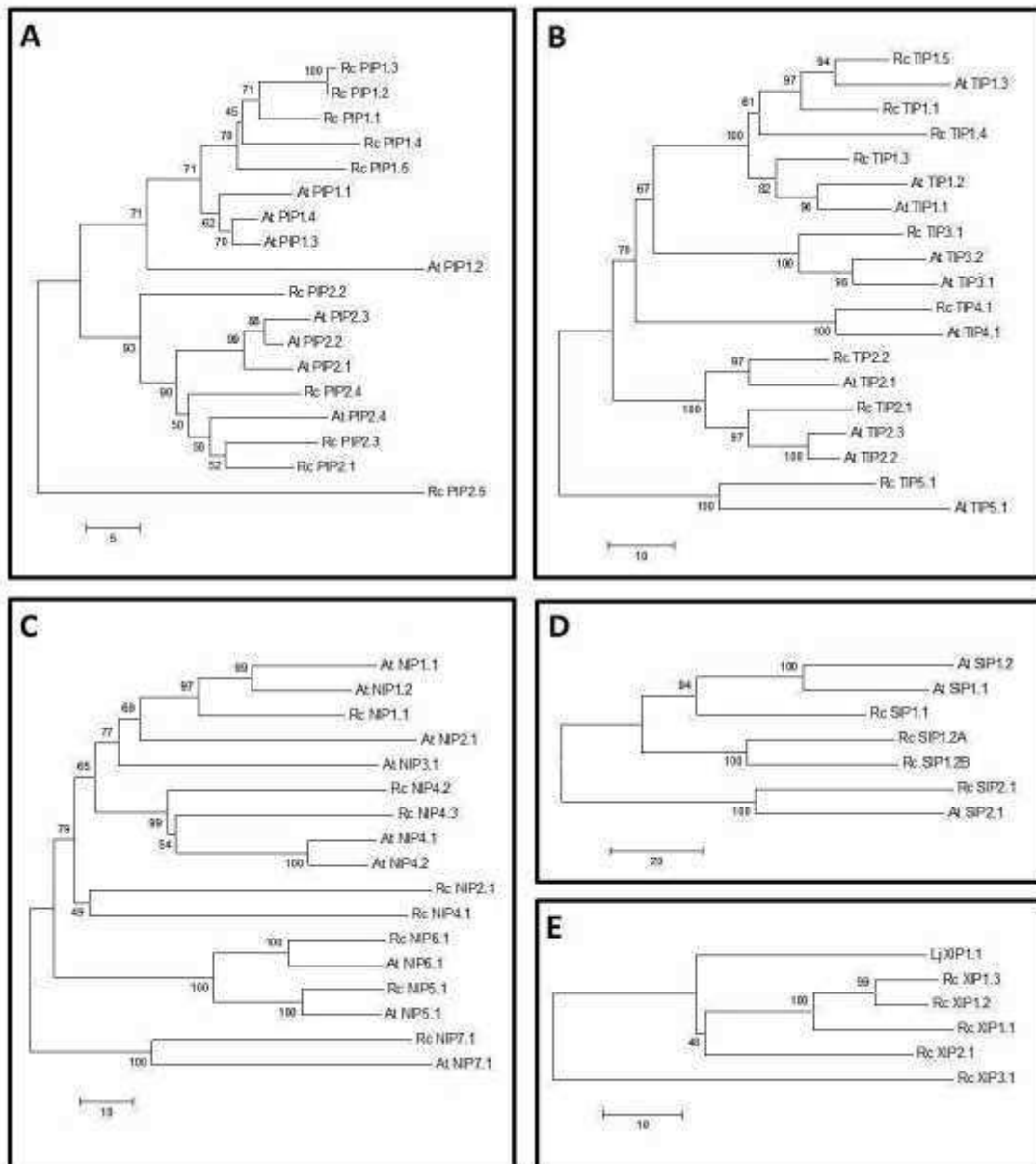


Figura 3. Análise fenética das aquaporinas previstas em *R. communis* (RcMIPs) e de *A. thaliana* (AtMIPs), por subfamília, separadamente. (A) RcPIPs e AtPIPs. (B) RcTIPs e AtTIPs. (C) RcNIPs e AtNIPs. (D) RcSIPs e AtSIPs. (E) RcXIPs e LjXIPs.



Os dados de MePIPs produziram um alinhamento de 311 posições, onde 150 foram variáveis, 116 parcimônio-informativos e 153 sítios foram conservados. O fenograma mostra a divisão de MePIPs em dois grupos distintos: MePIP1 e MePIP2, com quatro e dez membros, respectivamente (Figura 2A).

De forma semelhante, as RcPIPs produziram um alinhamento de 311 posições, onde 157 foram variáveis, 108 parcimônio-informativos e 150 sítios foram conservados entre os diferentes membros de PIP, apresentando cinco RcPIP1 e cinco RcPIP2 (Figura 3A).



As sequências MeTIPs produziram um alinhamento de 273 posições, com 213 sítios variáveis, 185 parcimônio-informativos e 58 conservados para todas as sequências (Figura 2B). Os grupos encontrados foram TIP1 (seis membros), TIP2 (três membros), TIP3 (dois membros), TIP4 (um membro) e TIP5 (um membro). Já as RcTIPs produziram um alinhamento de 268 posições, com 209 sítios variáveis, 184 parcimônio-informativos e 59 conservados para todas as sequências (Figura 3B). Os grupos encontrados foram TIP1 (quatro membros), TIP2 (dois membros), TIP3, TIP4 e TIP5 (com um membro cada).

Os dados de MeNIP resultaram em um alinhamento de 363 posições, onde 304 representavam sítios variáveis, 241 sítios parcimônio-informativos e 32 sítios conservados. As sequências pertencentes à MeNIPs foram agrupadas em seis subgrupos, distribuindo-se em NIP1 (dois membros), NIP2 (um membro), NIP3 (dois membros), NIP4 (um membro), NIP6 (um membro) e NIP7 (um membro) (Figura 2C).

Já as sequências de RcNIP resultaram em um alinhamento de 426 posições, onde 325 representavam sítios variáveis, 236 sítios parcimônio-informativos e 36 sítios conservados nas sequências analisadas, sendo estas agrupadas também em seis subgrupos, distribuindo-se em NIP1 (um membro), NIP2 (um membro), NIP4 (três membros), NIP5 (um membro), NIP6 (um membro) e NIP7 (um membro) (Figura 3C).

O alinhamento com MeSIPs apresentou 246 posições, com 218 sítios variáveis, 27 conservados e 63 parcimônio-informativos. As sequências pertencentes à MeSIPs foram agrupadas em SIP1 (um membro) e SIP2 (dois membros) (Figura 2D). O alinhamento com RcSIPs apresentou 252 posições, com 213 sítios variáveis, 34 conservados e 140 parcimônio-informativos. As sequências foram agrupadas em SIP1 (três membros) e SIP2 (um membro) (Figura 3D).

As sequências MeXIPs produziram um alinhamento de 322 posições, com 229 sítios variáveis, 79 parcimônio-informativos e 77 conservados para todas as sequências analisadas. Conforme a Figura 2E, as sequências pertencentes à MeXIPs foram agrupadas em XIP1 (um membro), XIP2 (um membro) e XIP3 (dois membros), enquanto que as sequências RcXIPs produziram um alinhamento de 346 posições, com 231 sítios variáveis, 52 parcimônio-informativos e 80 conservados, permanecendo, estas, agrupadas em XIP1 (três membros) e XIP2 (um membro) (Figura 3E) (CHAUMONT et al., 2001; DANIELSON; JOHANSON, 2010).

A classificação realizada com base no dendograma confirma o agrupamento das aquaporinas em PIP (com dois subgrupos); TIP (com cinco subgrupos); NIP (com sete

subgrupos); SIP (com dois subgrupos) e XIP (com três subgrupos), corroborando a classificação proposta por Zou et al. (2015 a, b).

### 3.1.1 Caracterização e localização subcelular

As MeAQPs preditas mostraram um comprimento de 95 a 306 (média de 260) aminoácidos, massa molecular de 11,69 a 32,52 (média de 28,11kDa), com valor de pI variando de 4,55 a 10,18 (média de 7,50) (Tabela 1). A maioria (93%) das sequências apresentou domínio MIP completo, com exceção das sequências MeNIP1.1, MeSIP1.1 e MeSIP2.1, que se apresentam incompletas na região N-terminal (Tabela 1).

Da mesma forma, para RcAQPs, foi observado um comprimento variável entre 198 a 367 (média de 268) aminoácidos, com massa molecular de 21,20 a 39,50 (média de 28,42 kDa) e com um pI básico (4,29 a 11, 88) (média de 8,00) (Tabela 1). A maioria (86%) das sequências apresentou domínio MIP completo, com exceção das sequências RcPIP2.5, RcSIP1.1, RcSIP1.2A e RcSIP2.1, que se apresentam incompletas na região N-terminal e RcXIPs 2.1 na região C-terminal.

A análise realizada com a ferramenta *online* Cell-Ploc mostrou que a maioria das MeAQPs e RcAQPs apresentam-se localizadas na membrana plasmática (MP). As exceções foram 4 MeAQPs (MePIP2.3, MeTIP1.3, MeTIP1.5, MeTIP1.6), de localização vacuolar (VC), e uma RcAQP (RcNIP2.1), localizada no cloroplasto (CP) (Tabela 1).

Tabela 1. Caracterização de genes de AQPs em *M. esculenta* (43 sequências) e de *R. communis* (36 sequências).

Nome da Proteína	(Phytozome)	(JVirGel)	(ORF)	(CD-search)	(CellPloc)			
	Locus ID	E	MM (KDa)	pI	(pblaa)	Domínio Incomp.	LSC	
MePIP1.1	Manes.01G059600.1	4	30.59	8.06	864 287	MIP sf	-	MP
MePIP1.2	Manes.02G020100.1	4	25.97	5.50	864 287	MIP sf	-	MP
MePIP1.3	Manes.07G126600.1	4	29.93	6.58	864 287	MIP sf	-	MP
MePIP1.4	Manes.10G016600.1	4	30.55	7.18	669 222	MIP sf	-	MP
MePIP1.5	Manes.02G109200.1	4	29.52	9.18	852 283	MIP sf	-	MP
MePIP2.1	Manes.11G145300.1	4	25.14	6.19	723 240	MIP sf	-	MP
MePIP2.2	Manes.04G076500.1	4	32.03	8.37	843 280	MIP sf	-	MP
MePIP2.3	Manes.04G021200.1	4	32.52	8.09	861 286	MIP sf	-	VC
MePIP2.4	Manes.05G144100.1	4	26.09	8.10	846 281	MIP sf	-	MP
MePIP2.5	Manes.07G100500.1	4	25.65	4.74	858 285	MIP sf	-	MP
MePIP2.6	Manes.08G006800.1	4	30.79	8.46	867 288	MIP sf	-	MP
MePIP2.7	Manes.09G068800.1	4	30.42	9.13	867 288	MIP sf	-	MP

Continuação da Tabela 1

Nome da Proteína	(Phytozome)	(JVirGel)	(ORF)	(CD-search)	(CellPloc)			
	Locus ID	E	MM (KDa)	pI	(pblaa)	Domínio Incomp.	LSC	
MePIP2.8	Manes.10G046200.1	4	25.83	4.94	861 286	MIP sf	-	MP
MePIP2.9	Manes.11G096600.1	4	30.70	7.18	843 280	MIP sf	-	MP
MeTIP1.1	Manes.04G030400.1	3	25.96	6.15	759 252	MIP sf	-	MP
MeTIP1.2	Manes.07G111500.1	3	30.73	8.46	759 252	MIP sf	-	MP
MeTIP1.3	Manes.08G012800.1	2	30.39	8.87	759 252	MIP sf	-	VC
MeTIP1.4	Manes.09G062300.1	2	25.84	5.02	759 252	MIP sf	-	MP
MeTIP1.5	Manes.10G035000.1	2	30.04	8.60	759 252	MIP sf	-	VC
MeTIP1.6	Manes.11G134600.1	3	25.26	5.63	759 252	MIP sf	-	VC
MeTIP2.1	Manes.01G081600.1	3	30.74	8.99	753 250	MIP sf	-	MP
MeTIP2.2	Manes.02G040800.1	3	27.31	6.88	753 250	MIP sf	-	MP
MeTIP2.3	Manes.11G036500.1	3	30.57	8.33	747 248	MIP sf	-	MP
MeTIP3.1	Manes.01G160000.1	3	29.76	8.95	777 258	MIP sf	-	MP
MeTIP3.2	Manes.02G118300.1	3	25.68	4.55	777 258	MIP sf	-	MP
MeTIP4.1	Manes.03G062300.1	3	31.82	6.31	744 247	MIP sf	-	MP
MeTIP5.1	Manes.14G036400.1	3	30.62	8.75	759 252	MIP sf	-	MP
MeNIP1.1	Manes.16G044000.1	4	25.41	4.86	684 227	MIP sf	N	MP
MeNIP1.2	Manes.17G061100.1	5	27.45	7.51	858 285	MIP sf	-	MP
MeNIP2.1	Manes.01G091200.1	5	25.78	6.14	867 288	MIP sf	-	MP
MeNIP3.1	Manes.12G133900.1	5	25.92	5.98	822 273	MIP sf	-	MP
MeNIP3.2	Manes.13G093600.1	5	31.03	8.38	825 274	MIP sf	-	MP
MeNIP4.1	Manes.02G152300.1	5	30.52	8.70	852 283	MIP sf	-	MP
MeNIP5.1	Manes.01G001400.1	4	30.45	8.45	897 298	MIP sf	-	MP
MeNIP6.1	Manes.04G104100.1	5	30.39	9.60	921 306	MIP sf	-	MP
MeNIP7.1	Manes.03G183400.1	5	24.26	9.66	900 299	MIP sf	-	MP
MeSIP1.1	Manes.09G144400.1	3	31.58	7.80	720 239	MIP sf	N	MP
MeSIP2.1	Manes.09G074000.1	2	29.31	6.58	495 164	MIP sf	N	MP
MeSPI2.2	Manes.09G074100.1	2	31.86	8.10	318 105	MIP sf	-	MP
MeXIP1.1	Manes.04G078900.1	2	28.87	4.73	885 294	MIP sf	-	MP
MeXIP2.1	Manes.11G089300.1	3	26.07	9.30	915 304	MIP sf	-	MP
MeXIP3.1	Manes.11G089200.1	3	18.11	10.18	921 306	MIP sf	-	MP
MeXIP3.2	Manes.11G089100.1	1	11.69	9.71	288 95	MIP sf	-	MP
RcPIP1.1	30190.m011229	4	30.66	8.17	864 287	MIP sf	-	MP
RcPIP1.2	29669.m000809	4	30.77	7.72	867 288	MIP sf	-	MP
RcPIP1.3	29669.m000808	4	30.68	7.79	864 287	MIP sf	-	MP
RcPIP1.4	29969.m000266	4	30.77	7.22	867 288	MIP sf	-	MP
RcPIP1.5	30174.m008614	4	30.71	8.72	861 286	MIP sf	-	MP
RcPIP2.1	28076.m000411	4	30.59	7.70	864 287	MIP sf	-	MP
RcPIP2.2	29869.m001194	4	29.71	8.96	843 280	MIP sf	-	MP
RcPIP2.3	30078.m002337	4	30.65	8.07	867 288	MIP sf	-	MP

Continuação da Tabela 1

Nome da Proteína	(Phytozome)	(JVirGel)	(ORF)	(CD-search)	(CellPloc)			
	Locus ID	E	MM (KDa)	pI	(pblaa)	Domínio Incomp.	LSC	
RcPIP2.4	27516.m000174	4	30.30	8.16	852 283	MIP sf	-	MP
RcPIP2.5	28962.m000437	3	21.20	8.85	597 198	MIP sf	N	MP
RcTIP1.1	29788.m000338	3	30.68	7.79	759 252	MIP sf	-	MP
RcTIP1.3	30078.m002254	2	30.77	7.22	756 251	MIP sf	-	MP
RcTIP1.4	29589.m001261		30.71	8.72	759 252	MIP sf	-	MP
RcTIP1.5	28180.m000392	3	30.59	7.70	759 252	MIP sf	-	MP
RcTIP2.1	30101.m000372	3	29.71	8.96	612 203	MIP sf	-	MP
RcTIP2.2	30146.m003542	3	30.65	8.07	747 248	MIP sf	-	MP
RcTIP3.1	29681.m001366	3	30.30	8.16	768 255	MIP sf	-	MP
RcTIP4.1	29794.m003419	3	21.20	8.85	744 247	MIP sf	-	MP
RcTIP5.1	30147.m014231	3	26.04	11.88	759 252	MIP sf	-	MP
RcNIP1.1	30026.m001488	5	28.86	9.42	816 271	MIP sf	-	MP
RcNIP2.1	27860.m000040	5	31.49	9.00	894 297	MIP sf	-	CP
RcNIP4.1	29908.m006033	5	30.64	6.22	849 282	MIP sf	-	MP
RcNIP4.2	28827.m000171	5	26.39	8.30	759 252	MIP sf	-	MP
RcNIP4.3	29816.m000676	8	39.50	7.13	1104 367	MIP sf	-	MP
RCNIP5.1	30068.m002640	4	30.94	8.64	897 298	MIP sf	-	MP
RcNIP6.1	29588.m000860	5	31.84	8.37	927 308	MIP sf	-	MP
RcNIP7.1	29844.m003330	5	31.57	8.27	897 298	MIP sf	-	MP
RcSIP1.1	29950.m001177	3	25.92	9.91	720 239	MIP sf	N	MP
RcSIP1.2A	30010.m000659	1	25.62	10.02	720 239	MIP sf	N	MP
RcSIP1.2B	30010.m000658	1	25.24	10.31	705 234	MIP sf	-	MP
RcSIP 2.1	30045.m000487	3	26.04	11.88	723 240	MIP sf	N	MP
RcXIP1.1	28846.m000049	1	30.80	7.71	858 285	MIP sf	-	MP
RcXIP1.2	28846.m000048	2	32.55	6.99	912 303	MIP sf	-	MP
RcXIP1.3	28929.m000055	2	33.22	7.00	930 309	MIP sf	-	MP
RcXIP2.1	28929.m000054	2	22.49	7.63	627 208	MIP sf	C	MP
RcXIP3.1	28747.m000131	5	25.67	5.71	729 242	MIP sf	-	MP

sf= superfamily; N= domínio incompleto na região N-terminal; C= domínio incompleto na região C-terminal; MP= membrana plasmática; VC= vacúolo; CP= cloroplasto; E= éxons; MM= massa molecular; pI= ponto isoelétrico; pb= nucleotídeos; aa= aminoácidos e LSC= localização subcelular .

### 3.1.2 Motivos e domínios conservados

As aquaporinas apresentam uma estrutura conservada dentro dos diferentes reinos (AZAD et al., 2012), a seletividade do poro está diretamente relacionada à presença de motivos e resíduos conservados como os motivos NPA e os resíduos da região

arginina/aromática (ar/R), que têm se mostrado importantes na determinação da afinidade de substrato e passagem destes através das membranas biológicas (MAUREL et al., 2008).

Por meio de alinhamentos múltiplos gerados com o auxílio da ferramenta *online* Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) e visualizados no Jalview (WATERHOUSE et al., 2009), foi possível observar a presença e integridade dos motivos AEF e NPA (Figura 4), a integridade e posição dos resíduos de aminoácidos responsáveis pela formação do filtro seletivo (ar/R) das aquaporinas e as posições de Forger's, responsáveis pela distinção entre aquaporinas e aquagliceroporinas (Tabelas 2 e 3).

Figura 4. Alinhamento múltiplo entre MePIPs e AtPIPs gerados com o auxílio da ferramenta *online* Clustal Omega. Destaque para os motivos conservados AEF e NPA.

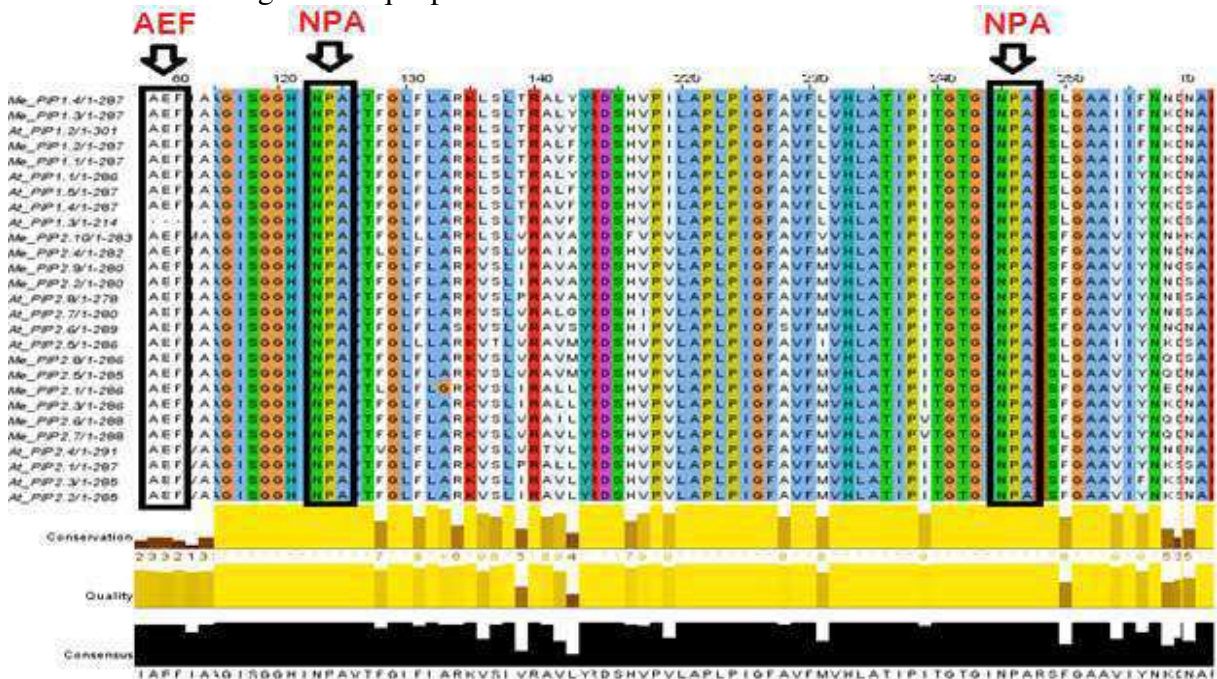


Tabela 2. Sumário das principais regiões e sítios conservados de aquaporinas em *Manihot esculenta*.

Classificação	Motivo	Ar/R						Forger's Position				
		NPA (LB)	NPA (LE)	H2	H5	LE1	LE2	P1	P2	P3	P4	P5
<b>PIP</b>	MePIP1.1	NPA	NPA	F	H	T	R	G	S	A	F	W
	MePIP1.2	NPA	NPA	F	H	T	R	G	S	A	F	W
	MePIP1.3	NPA	NPA	F	H	T	R	G	S	A	F	W
	MePIP1.4	NPA	NPA	F	H	T	R	G	S	A	F	W
	MePIP2.1	NPA	NPA	F	H	T	R	G	S	A	F	W
	MePIP2.2	NPA	NPA	F	H	T	R	A	S	A	F	W
	MePIP2.3	NPA	NPA	F	H	T	R	A	S	A	F	W
	MePIP2.4	NPA	NPA	F	H	T	R	A	S	A	F	W

Continuação da Tabela 2

Classificação	Motivo		Ar/R				Forges's Position				
	NPA (LB)	NPA (LE)	H2	H5	LE1	LE2	P1	P2	P3	P4	P5
MePIP2.5	NPA	NPA	F	H	T	R	G	S	A	F	W
MePIP2.6	NPA	NPA	F	H	T	R	G	S	A	F	W
MePIP2.7	NPA	NPA	F	H	T	R	A	S	A	F	W
MePIP2.8	NPA	NPA	F	H	T	R	A	S	A	F	W
MePIP2.9	NPA	NPA	F	H	T	R	A	S	A	F	W
MePIP2.10	NPA	NPA	F	H	T	R	A	S	A	F	W
<b>TIP</b>											
MeTIP1.1	NPA	NPA	H	I	A	V	F	S	A	Y	W
MeTIP1.2	NPA	NPA	H	I	A	V	F	S	A	Y	W
MeTIP1.3	NPA	NPA	H	I	A	V	F	S	A	Y	W
MeTIP1.4	NPA	NPA	H	I	A	V	L	S	A	Y	W
MeTIP1.5	NPA	NPA	H	I	A	V	F	S	A	Y	W
MeTIP1.6	NPA	NPA	H	I	A	V	F	S	A	Y	W
MeTIP2.1	NPA	NPA	H	I	G	R	F	S	A	Y	W
MeTIP2.2	NPA	NPA	H	I	G	R	L	S	A	Y	W
MeTIP2.3	NPA	NPA	H	I	G	R	A	S	A	Y	W
MeTIP3.1	NPA	NPA	H	I	A	R	L	A	A	Y	W
MeTIP3.2	NPA	NPA	H	I	A	R	L	A	A	Y	W
MeTIP4.1	NPA	NPA	H	I	A	R	Y	A	A	Y	W
MeTIP5.1	NPA	NPA	N	V	G	C	V	A	A	Y	W
<b>NIP</b>											
MeNIP1.1	NPA	NPA	W	V	A	R	L	S	A	Y	I
MeNIP1.2	NPA	NPA	W	V	A	R	L	S	A	Y	I
MeNIP2.1	NPA	NPA	G	S	G	R	V	T	A	Y	L
MeNIP3.1	NPA	NPA	W	I	A	R	V	S	A	Y	V
MeNIP3.2	NPA	NPA	W	M	A	R	V	S	A	Y	I
MeNIP4.1	NPA	NPA	W	A	A	R	V	S	A	Y	I
MeNIP6.1	NPA	NPV	T	I	A	R	G	T	A	Y	L
MeNIP7.1	NPA	NPA	A	V	G	R	C	S	A	Y	I
<b>SIP</b>											
MeSIP1.1	NPT	NPA	V	V	P	N	E	A	A	Y	W
MeSIP2.1	*	NPA	F	Q	G	S	E	V	A	Y	W
MeSIP2.2	NPL	NPA	F	Q	G	S	L	V	A	Y	W
<b>XIP</b>											
MeXIP1.1	SPV	NPA	V	I	V	R	S	C	A	F	W
MeXIP2.1	NPV	NPA	I	T	V	R	A	C	A	F	W
MeXIP3.1	NPL	NPA	V	T	A	R	A	C	A	F	W
MeXIP3.2	NPL	NPA	V	T	A	R	A	C	A	F	W

LB= loop B; LE= loop E; H2= alfa-hélice 2; H5= alfa-hélice 5; P=posição.

Tabela 3. Sumário das principais regiões e sítios conservados de aquaporinas em *Ricinus communis* (RcAQP).

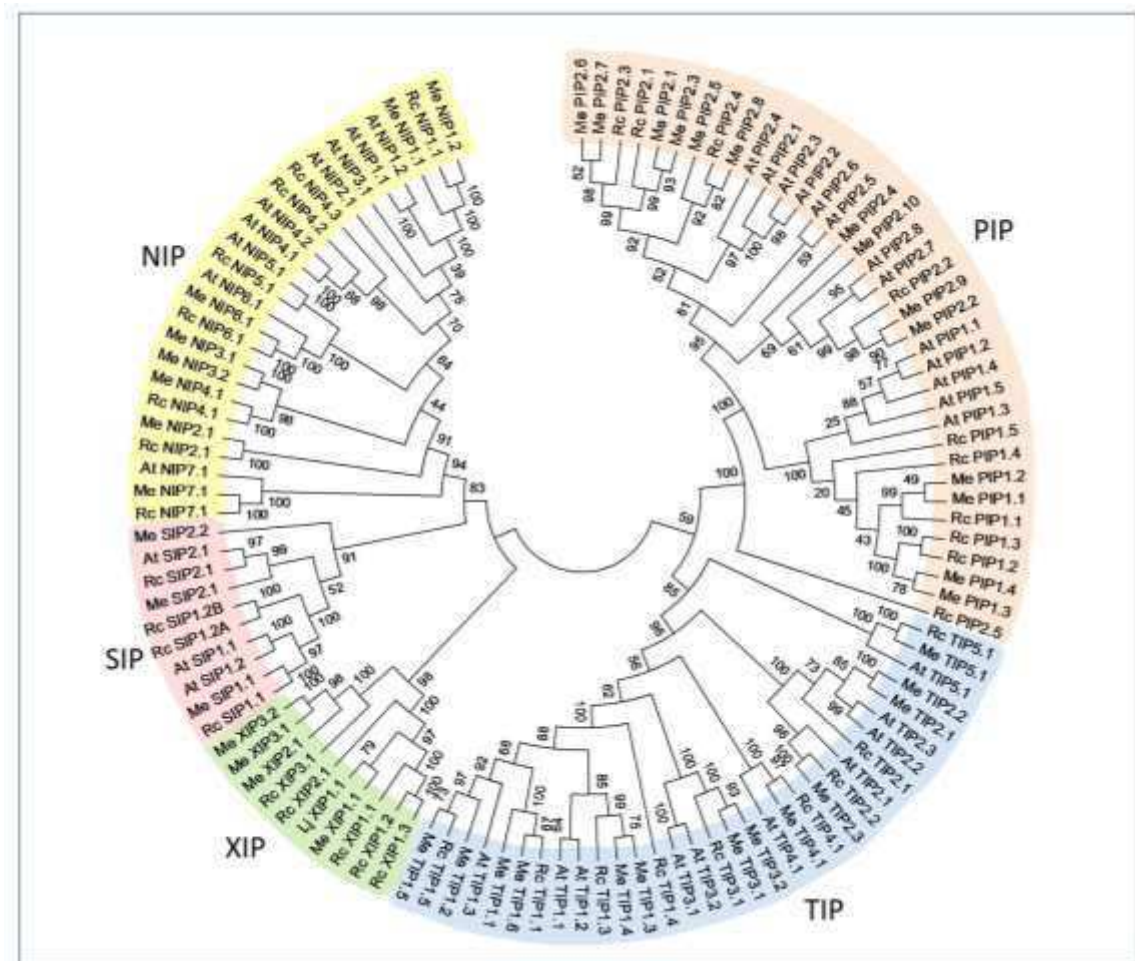
Classificação	Motivo		Ar/R				Forges's Position					
	NPA (LB)	NPA (LE)	H2	H5	LE1	LE2	P1	P2	P3	P4	P5	
<b>PIP</b>	RcPIP1.1	NPA	NPA	F	H	T	R	E	S	A	F	W
	RcPIP1.2	NPA	NPA	F	H	T	R	E	S	A	F	W
	RcPIP1.3	NPA	NPA	F	H	T	R	E	S	A	F	W
	RcPIP1.4	NPA	NPA	F	H	T	R	E	S	A	F	W
	RcPIP1.5	NPA	NPA	F	H	T	R	Q	S	A	F	W
	RcPIP2.1	NPA	NPA	F	H	T	R	Q	S	A	F	W
	RcPIP2.2	NPA	NPA	F	H	T	R	M	S	A	F	W
	RcPIP2.3	NPA	NPA	F	H	T	R	Q	S	A	F	W
	RcPIP2.4	NPA	NPA	F	H	T	R	Q	S	A	F	W
	RcPIP2.5	NPA	NPA	F	H	T	R	M	S	A	P	F
<b>TIP</b>	RcTIP1.1	NPA	NPA	H	I	A	V	T	S	A	Y	W
	RcTIP1.3	NPA	NPA	H	I	A	V	T	S	A	Y	W
	RcTIP1.4	NPA	NPA	H	I	A	V	T	S	A	Y	W
	RcTIP1.5	NPA	NPA	H	I	A	V	T	S	A	Y	W
	RcTIP2.1	NPA	NPA	H	I	G	R	T	S	A	Y	W
	RcTIP2.2	NPA	NPA	H	I	G	R	T	S	A	Y	W
	RcTIP3.1	NPA	NPA	H	I	A	R	T	A	A	Y	W
	RcTIP4.1	NPA	NPA	H	I	A	R	T	S	A	Y	W
	RcTIP5.1	NPA	NPA	N	V	G	C	A	A	A	Y	W
	<b>NIP</b>	RcNIP1.1	NPA	NPA	W	V	A	R	F	S	A	Y
RcNIP2.1		NPA	NPA	G	S	G	R	L	T	A	Y	I
RcNIP4.1		NPA	NPA	W	A	A	R	F	S	A	F	I
RcNIP4.2		NPA	NPA	W	V	A	R	V	S	A	Y	I
RcNIP4.3		NPA	NPA	W	V	A	R	F	T	A	Y	M
RcNIP5.1		NPS	NPV	A	I	G	R	F	T	A	Y	L
RcNIP6.1		NPS	NPV	T	I	A	R	F	T	A	Y	L
RcNIP7.1		NPA	NPA	A	V	G	R	Y	S	A	Y	I
<b>SIP</b>	RcSIP1.1	NPT	NPA	V	V	P	N	E	A	A	Y	W
	RcSIP1.2A	NPA	NPA	A	A	P	N	Q	V	A	Y	W
	RcSIP1.2B	NPN	NPA	A	A	T	N	Q	A	A	Y	W
	RcSIP 2.1	NPL	NPA	S	H	G	S	Q	V	A	Y	W
<b>XIP</b>	RcXIP1.1	SPV	NPA	V	V	V	R	M	C	V	F	W
	RcXIP1.2	SPT	NPA	V	V	V	R	M	C	A	F	W
	RcXIP1.3	SPT	NPA	V	V	V	R	M	C	A	F	W
	RcXIP2.1	SPV	NPA	V	V	V	R	M	C	A	F	W
	RcXIP3.1	NPV	NPA	I	L	A	R	M	C	A	F	W

LB= loop B; LE= loop E; H2= alfa-hélice 2; H5= alfa-hélice 5; P=posição.

### 3.1.3 Análise fenética e distribuição cromossômica

De posse das sequências renomeadas, as mesmas foram novamente alinhadas juntamente com as sequências de *Arabidopsis thaliana* e *Lotus japonicus*, para a construção do dendograma, a fim de verificar a relação fenética entre as diferentes subfamílias de aquaporinas vegetais (Figura 5). A árvore criada pelo método *Neighbor-Joining* confirma a classificação das aquaporinas presentes em *M. esculenta* e *R. communis* entre as cinco subfamílias já descritas para vegetais superiores, apresentando candidatas da subfamília XIP, presentes apenas em algumas dicotiledôneas.

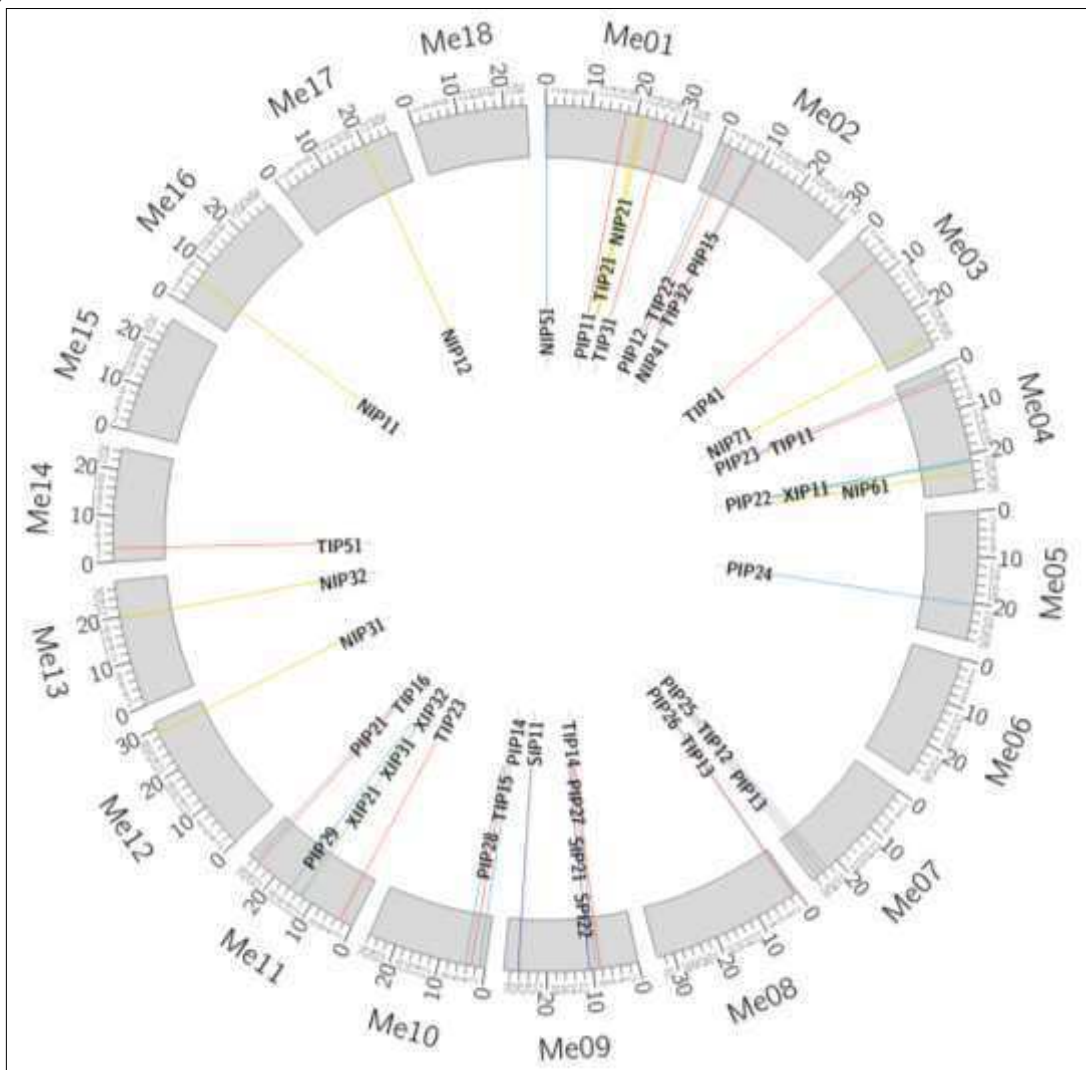
Figura 5. Relações fenéticas entre aquaporinas vegetais. Árvore construída pelo método de *Neighbour Joining*, mostrando a comparação entre as sequências de supostas aquaporinas presentes no genoma de *M. esculenta* e *R. communis* juntamente com aquaporinas de *A. thaliana* e *L. japonicus*. As divisões entre as subfamílias encontram-se indicadas pelas diferentes cores.





O genoma de *Manihot esculenta*, disponível no JBrowse do Phytozome, foi utilizado para verificar a distribuição dos genes de aquaporinas ao longo dos 18 cromossomos presentes na espécie (Figura 6). Foram identificados 43 locus, distribuídos em 15 cromossomos. O cromossomo 11 apresenta o maior número de aquaporinas, onde estão a maioria das MeXIPs (com exceção apenas da MeXIP1.1). Nos demais cromossomos, as AQPs se distribuem aleatoriamente: cromossomos 1,2,4 e 9 com 5 locus (Me09-onde estão todos os genes SIP), o 7 e o 10 com 3 locus, o 3 e o 8 com 2. A maioria dos cromossomos (5,12,13,14,16,17) apresentam apenas um locus. Com base na representação esquemática das estruturas gênicas de cada sequência, gerada a partir do JBrowse/Phytozome, foi possível estimar o número de éxons, os quais variavam entre 1 e 5.

Figura 6. Visualização da disposição dos genes de MeMIPs (*M. esculenta*) entre os 18 cromossomos de *M. esculenta*, de acordo com dados do Phytozome. Gráfico gerado pelo programa Circos.



### 3.1.4 Modelagem comparativa

A modelagem foi realizada através do alinhamento das sequências de MeMIPS e RcMIPS, de maneira comparativa a modelos conhecidos (estruturas determinadas experimentalmente), o que permitiu visualizar a arquitetura geral destes transportadores de membrana. De acordo com o nível de similaridade e semelhanças estruturais, estes modelos podem ser considerados tão eficientes quanto os métodos de Difração de Raios-X e Ressonância Magnética Nuclear.

Neste estudo, os modelos foram construídos a partir de modelos “*templates*” de aquaporinas, disponíveis no Banco de Dados de Proteínas – PDB (*Protein Data Bank*), que apresenta a melhor similaridade com as aquaporinas de *M. esculenta* e *R. communis*, além de representar a única aquaporina vegetal disponível. Entre as MeMIPs e RcMIPs, foi selecionado um representante de cada subfamília para realização da modelagem computacional, as quais apresentaram variações na similaridade após o alinhamento contra o modelo (Tabela 4).

Tabela 4. Valores de similaridade entre os modelos utilizados para modelagem e os diferentes membros de aquaporinas de *M. esculenta* e *R. communis*.

MeMIPs	Similaridade (%)	RcMIPs	Similaridade (%)
MePIP	73,95%	RcPIP	74,62%
MeTIP	38,29%	RcTIP	60,48%
MeNIP	34,30%	RcNIP	26,67%
MeSIP	21,52%	RcSIP	20,98%
MeXIP	27,23%	RcXIP	30,99%

Os modelos propostos para as aquaporinas de *M. esculenta* e *R. communis*, gerados via SWISS MODEL (Figura 7 - 11), apresentaram, para a maioria de seus membros, todas as características principais de aquaporinas, sendo: seis domínios transmembrana (H1-H6) interligados por cinco loops (LA-LE), juntamente com os dois domínios alfa-hélice onde se encontram os motivos conservados NPA, que interagem internamente, formando o poro seletivo (GOMES et al., 2009).

Figura 7. Modelos tridimensionais das (A) MePIPs e (B) RcPIPs, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína com destaque para os motivos NPA.

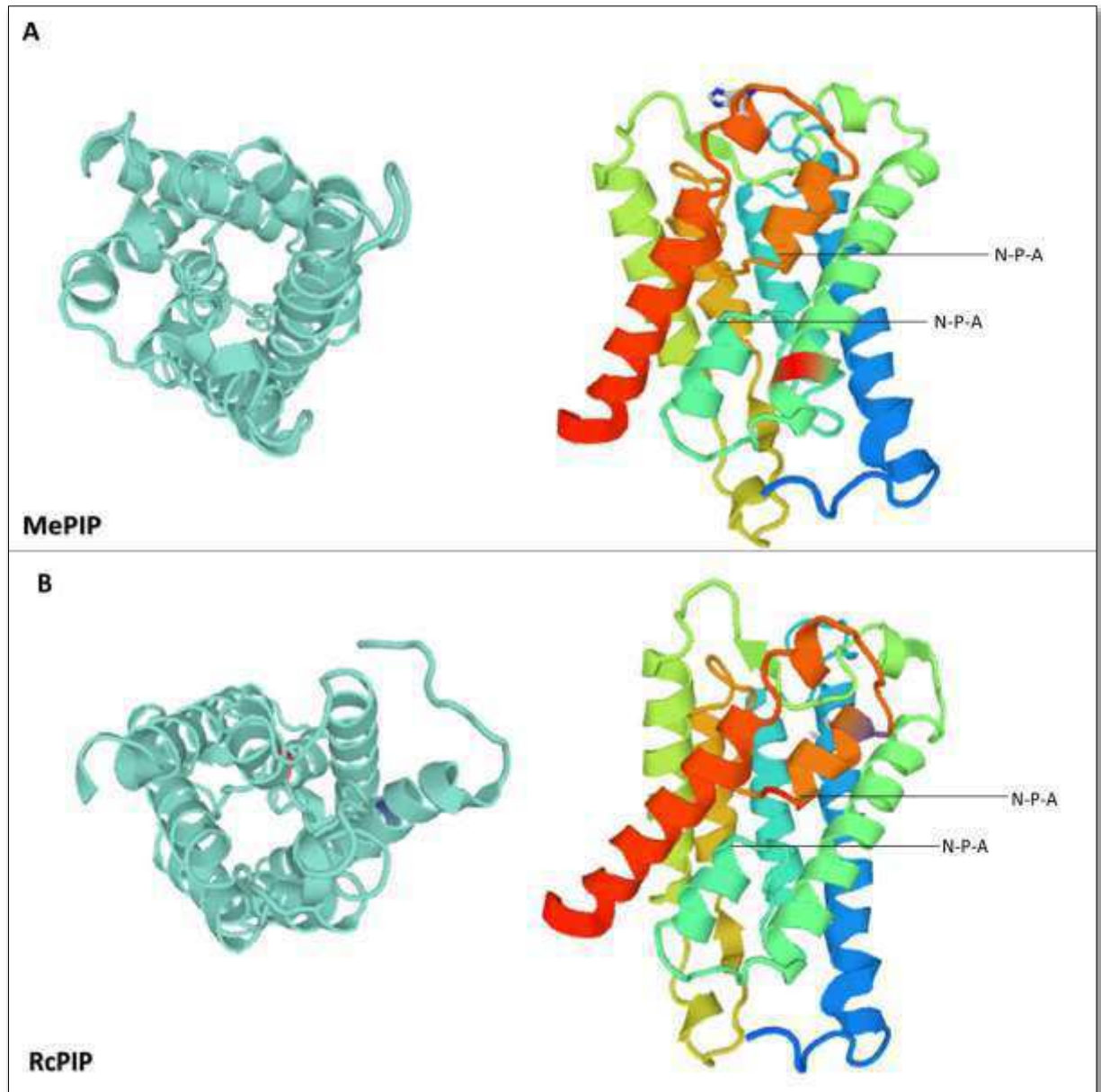


Figura 8. Modelos tridimensionais das (A) MeTIPs e (B) RcTIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína com destaque para os motivos NPA.

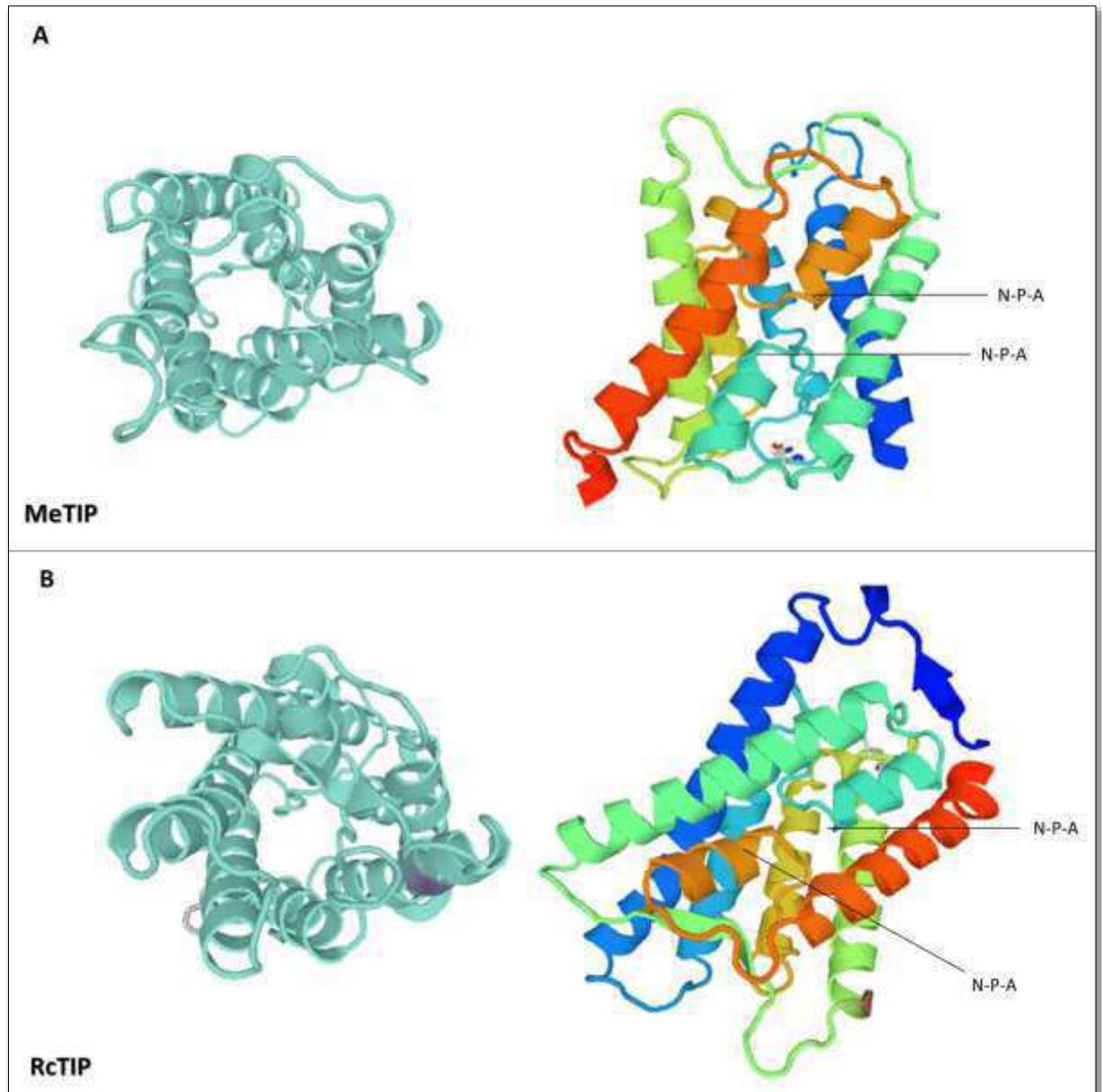


Figura 9. Modelos tridimensionais das (A) MeNIPs e (B) RcNIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína com destaque para os motivos NPA.

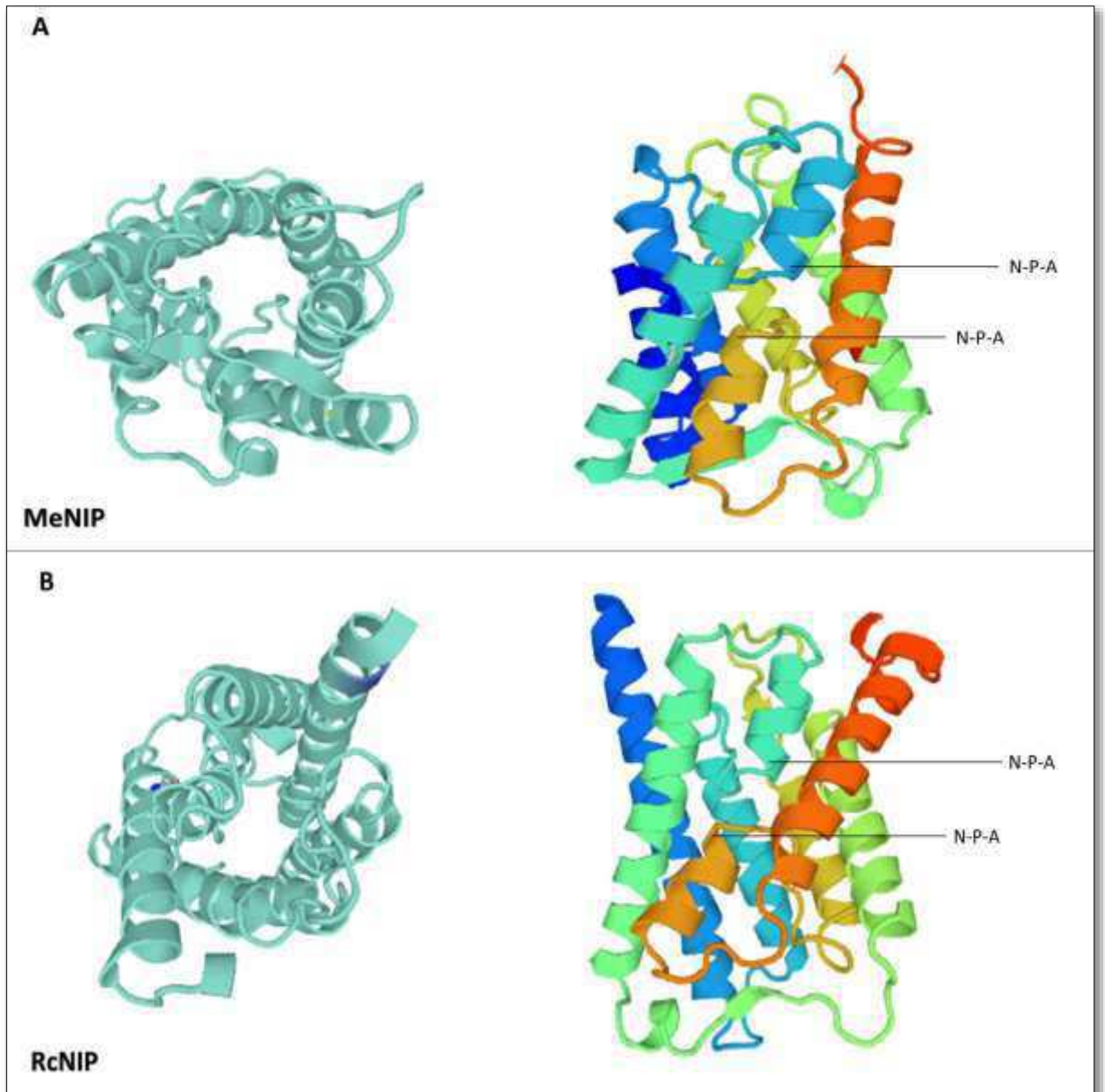


Figura 10. Modelos tridimensionais das (A) MeSIPs e (B) RcSIPs, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína com destaque para os motivos NPA.

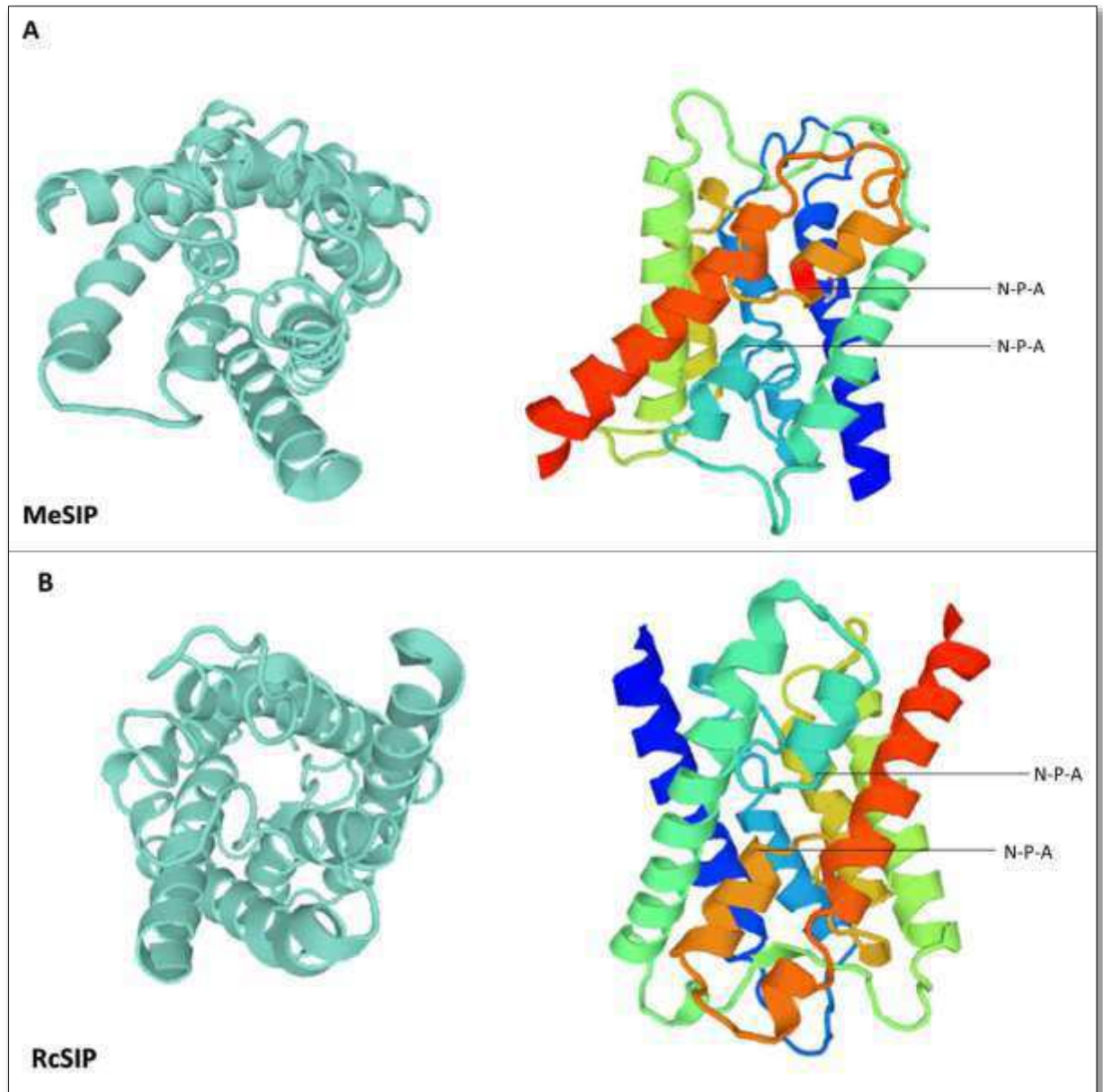
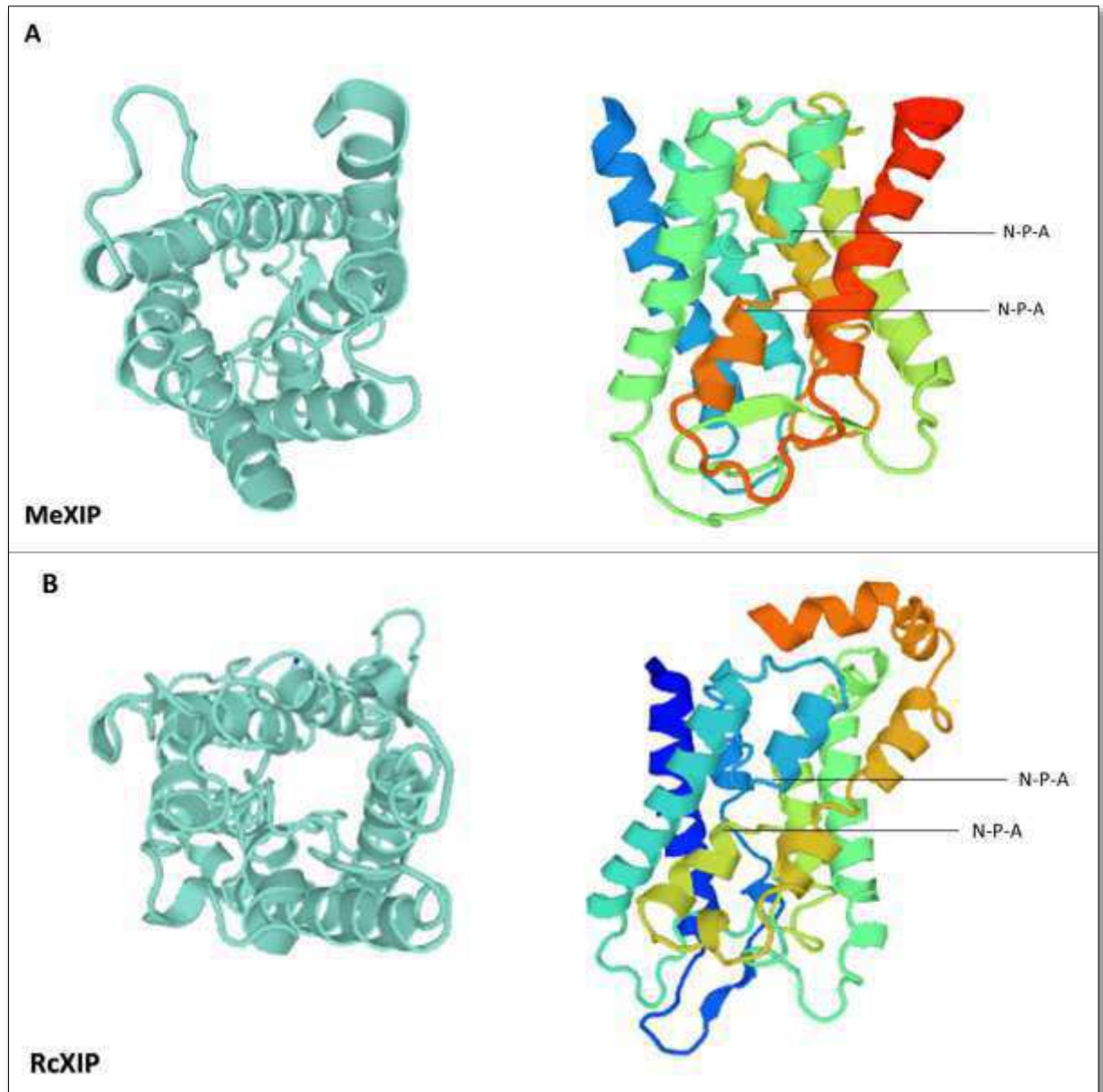


Figura 11. Modelos tridimensionais das (A) MeXIPs e (B) RcXIPS, obtidas via modelagem comparativa. À esquerda, imagem frontal do poro. À direita, imagem lateral da proteína com destaque para os motivos NPA.



A análise de confiabilidade dos modelos gerados para as aquaporinas da *M. esculenta* e *R. communis* foi realizada via ProSA (SIPPL, 1993; WIEDRSTEIN; SIPPL, 2007) (Figuras 12 -16) e Ramachandran Plot (BENKERT et al., 2009), que apresentaram valores dentro dos padrões de confiabilidade (Tabela 5) (Figuras 17 -26).

Para confirmação da qualidade do modelo, a plotagem no gráfico do *Z-score* da estrutura protéica depositada no PDB deve estar situada dentro da área azul (clara ou escura), a qual reflete o método de identificação. Todas as aquaporinas avaliadas situam-se dentro da área referida a padrões de qualidade desejada.

Tabela 5. Valores de qualidade e energia para os modelos estruturais de aquaporinas de *M. esculenta* e *R. communis* obtidos via plataforma SWISS-MODEL. (Valores mais próximos a um representam modelos com maior qualidade).

MIPs	SWISS-MODEL	ProSA	Ramachandran			
	QMEAN	Z-score	RMF	RPA	RGP	RNP
MePIP	-5,39	-5,1	86,80%	8,10%	5,10%	0,00%
MeTIP	-7,29	-2,95	84,50%	11,00%	4,60%	0,00%
MeNIP	-6,13	-2,88	91,30%	5,30%	3,40%	0,00%
MeSIP	-6,92	-2,71	85,90%	10,10%	4,00%	0,00%
MeXIP	-7,13	-3,71	89,70%	7,30%	3,00%	0,00%
RcPIP	-5,39	-5,41	87,70%	7,70%	4,70%	0,00%
RcTIP	-2,28	-4,56	98,75%	0,90%	0,40%	0,00%
RcNIP	-3,79	-3,07	94,10%	4,50%	1,40%	0,00%
RcSIP	-6,13	-4,04	93,00%	5,70%	1,30%	0,00%
RcXIP	-5,79	-3,13	91,70%	5,60%	2,80%	0,00%

RMF: Região Mais Favorável; RPA: Região Permitida Adicional; RGP: Região Generosamente Permitida; RNP: Região Não Permitida.

Figura 12. Gráficos representativos do Z-score dos modelos gerados para MePIP e RcPIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB.

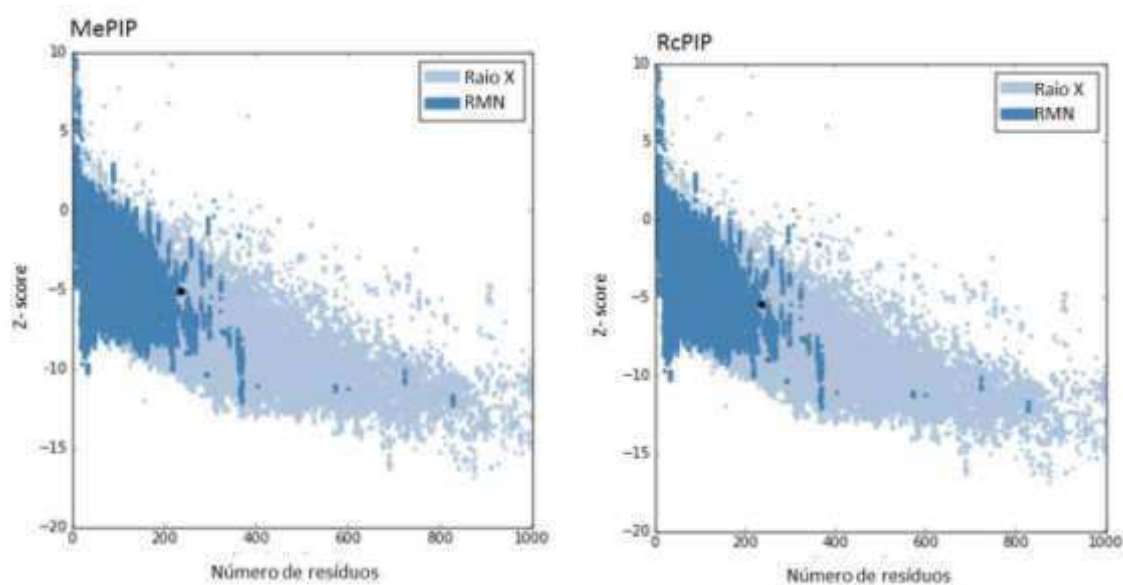




Figura 13. Gráficos representativos do Z-score dos modelos gerados para MeTIP e RcTIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB.

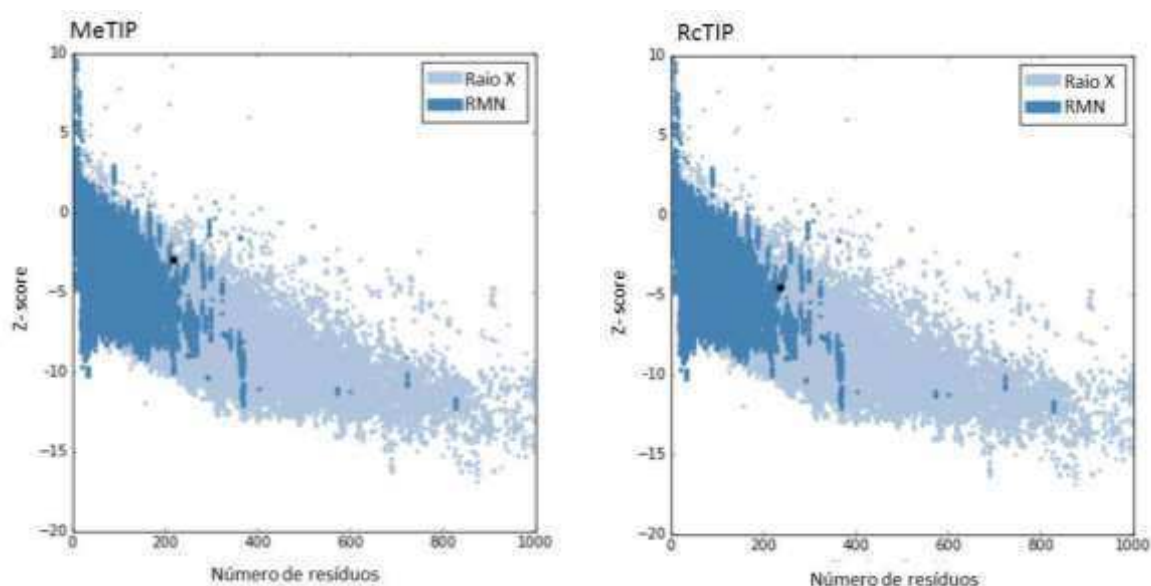


Figura 14. Gráficos representativos do Z-score dos modelos gerados para MeNIP e RcNIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB.

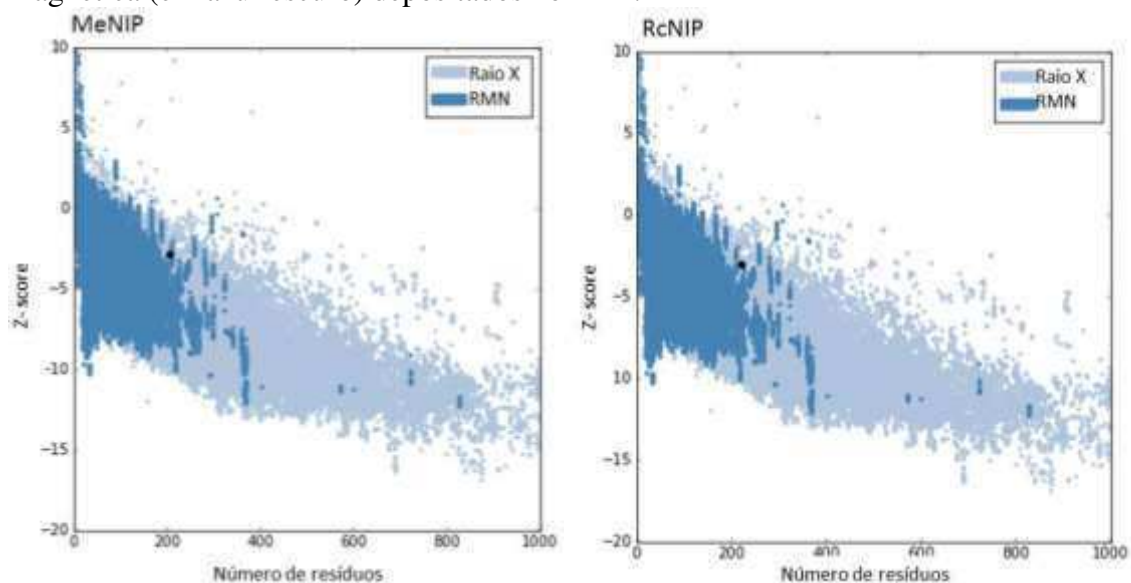


Figura 15. Gráficos representativos do Z-score dos modelos gerados para MeSIP e RcSIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB.

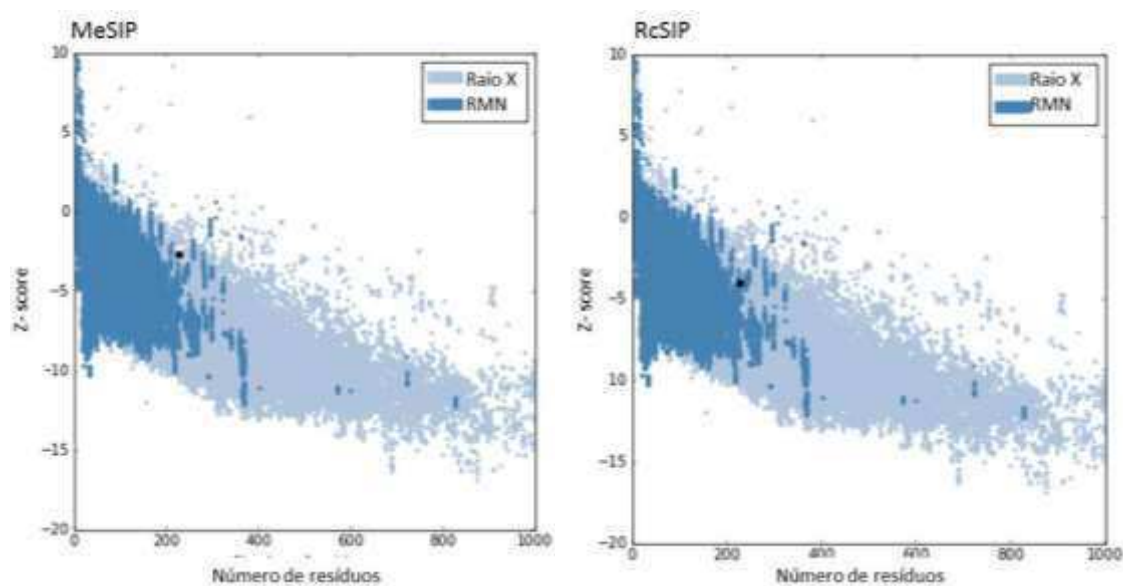


Figura 16. Gráficos representativos do Z-score dos modelos gerados para MeXIP e RcXIP, comparativamente aos modelos de cristalografia de raios-x (em azul-claro) e ressonância magnética (em azul-escuro) depositados no PDB.

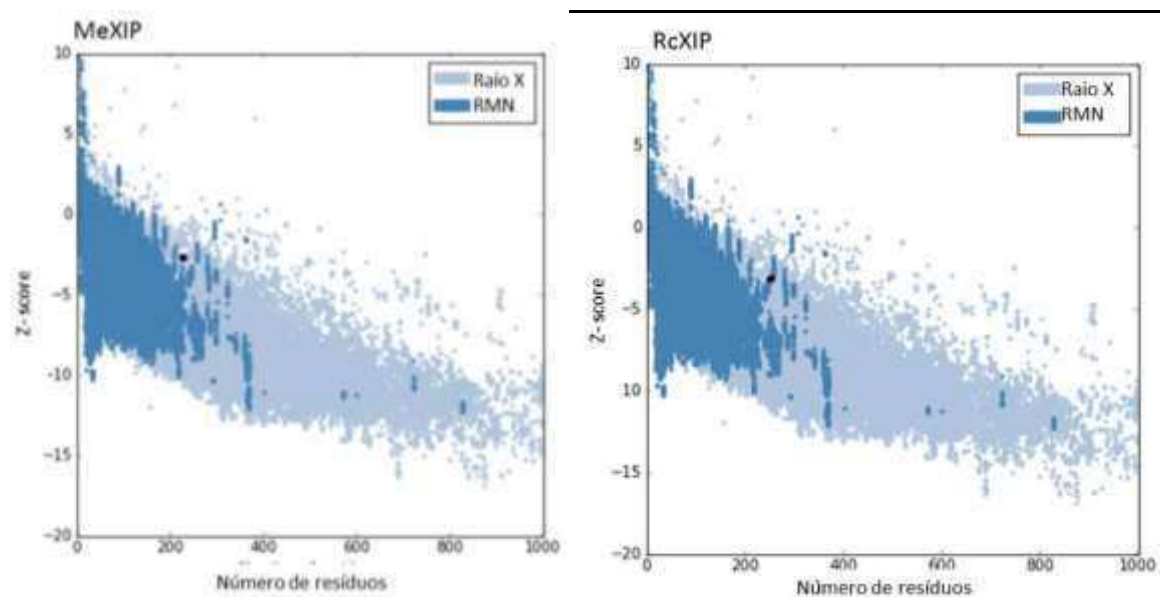


Figura 17. Análise de confiabilidade do modelo gerado para MePIP via Ramachandran Plot.

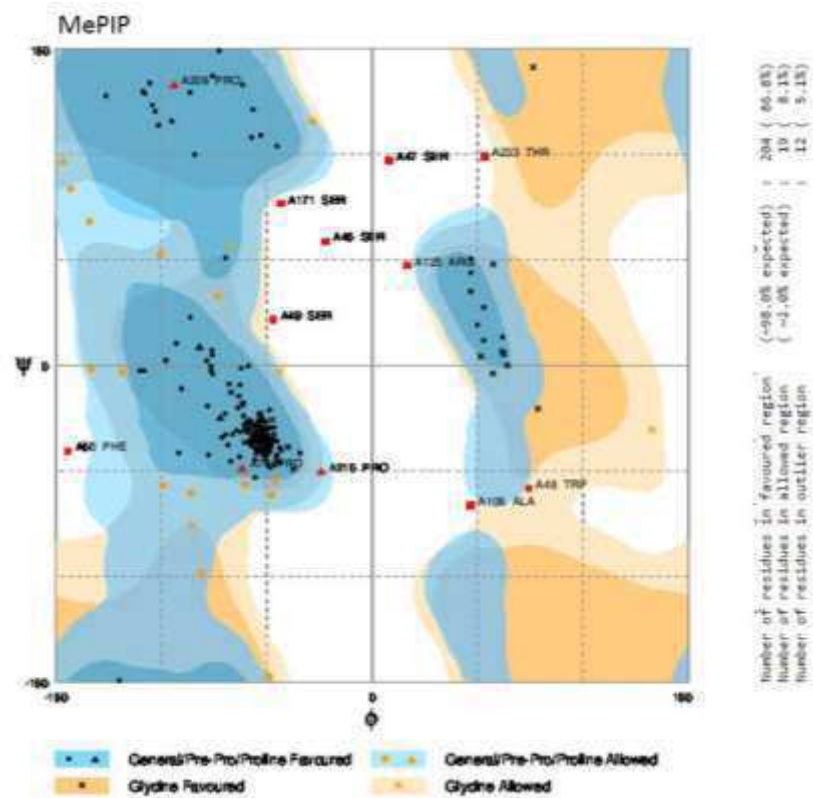


Figura 18. Análise de confiabilidade do modelo gerado para MeTIP via Ramachandran Plot.

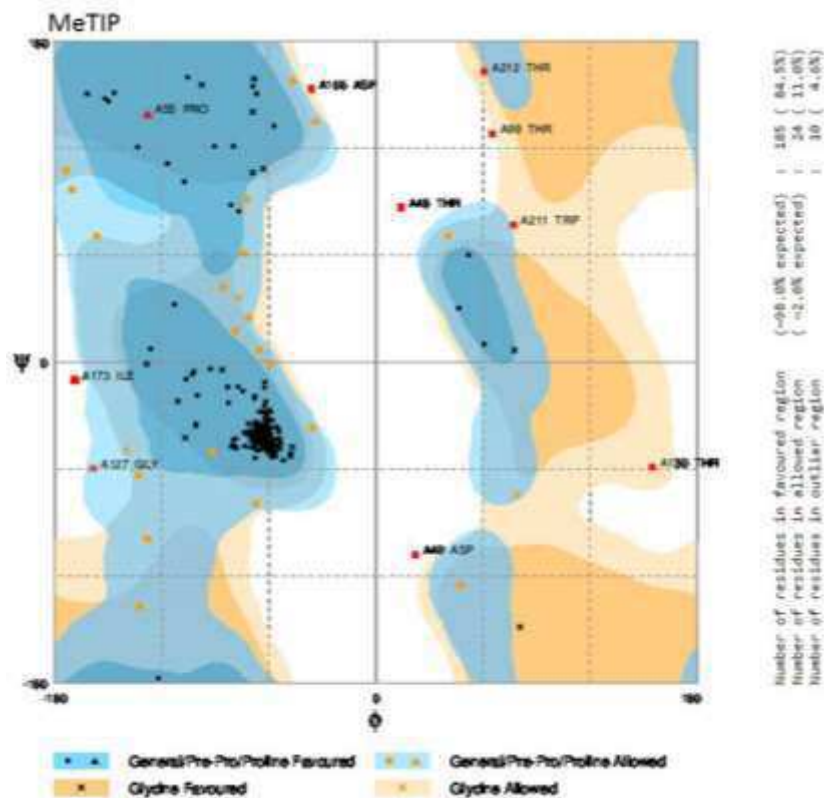




Figura 21. Análise de confiabilidade do modelo gerado para MeXIP via Ramachandran Plot.

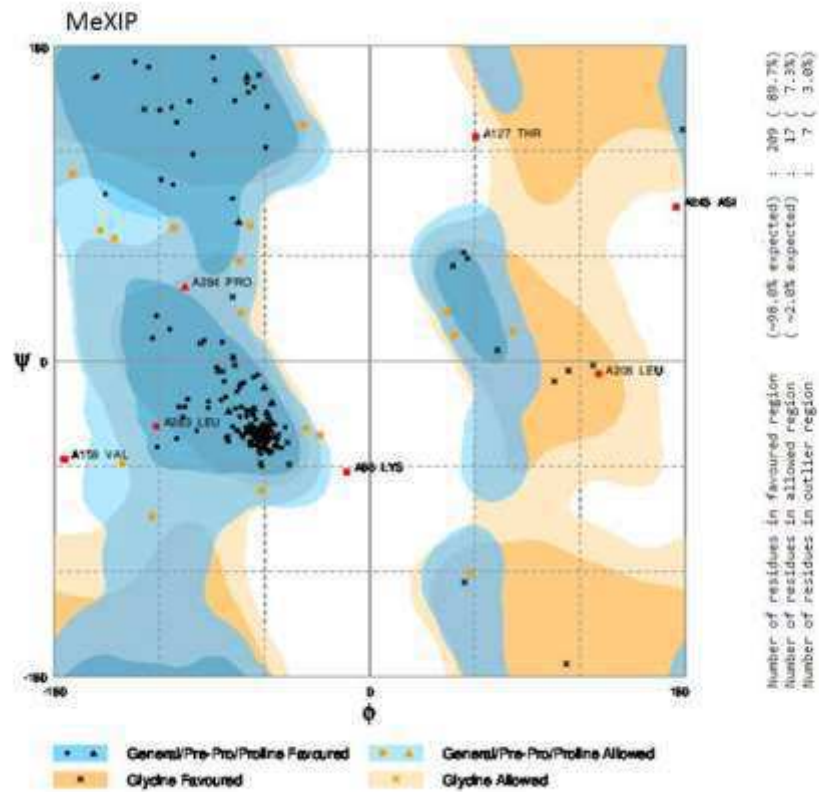


Figura 22. Análise de confiabilidade do modelo gerado para RcPIP via Ramachandran Plot.

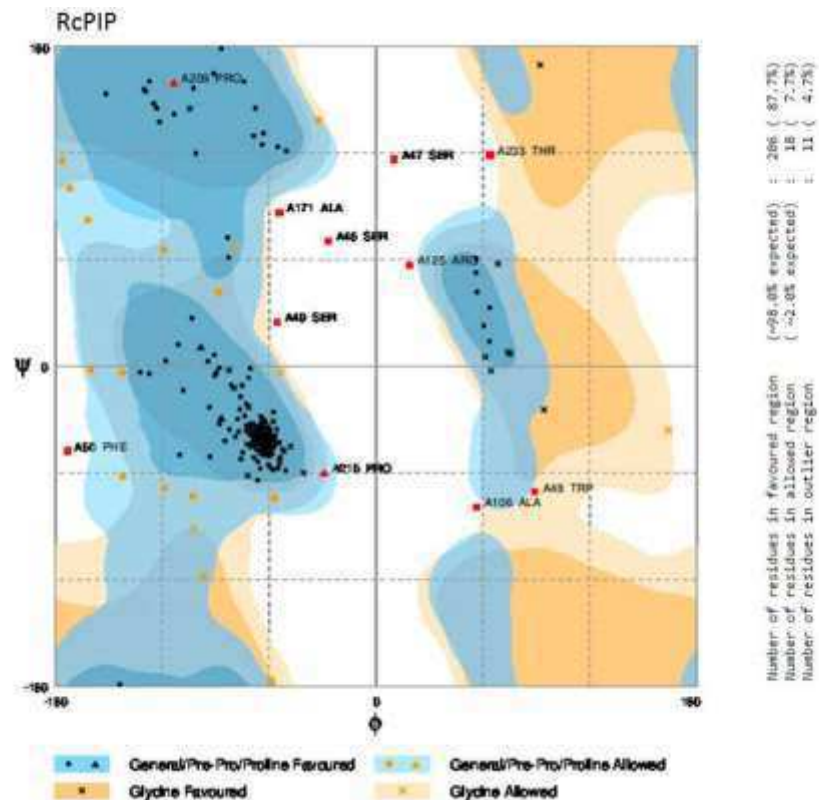


Figura 23. Análise de confiabilidade do modelo gerado para RcTIP via Ramachandran Plot.

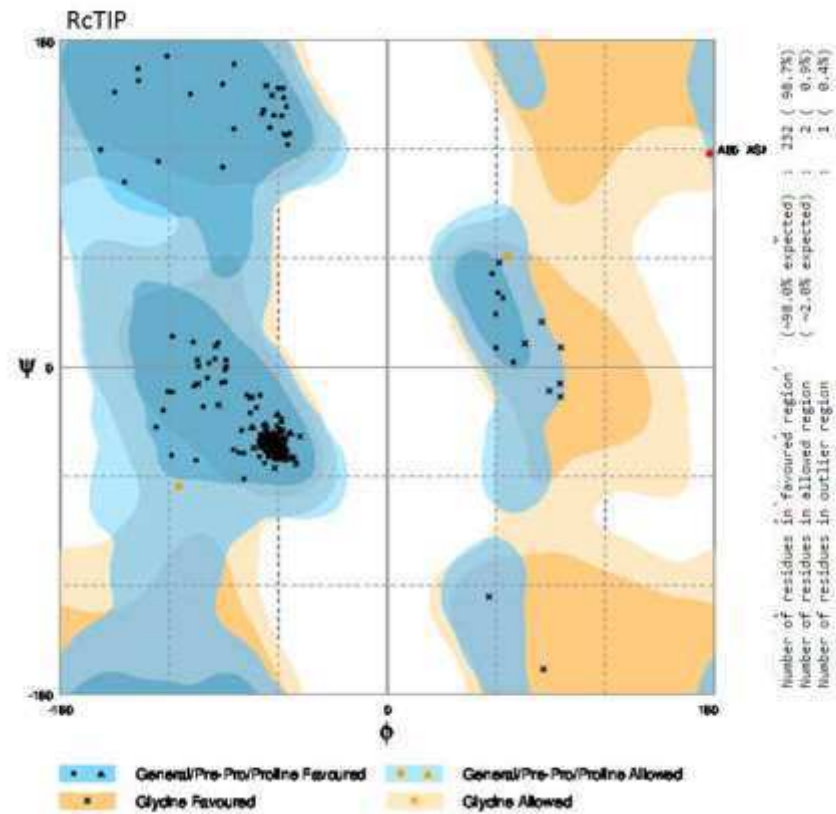


Figura 247. Análise de confiabilidade do modelo gerado para RcNIP via Ramachandran Plot.

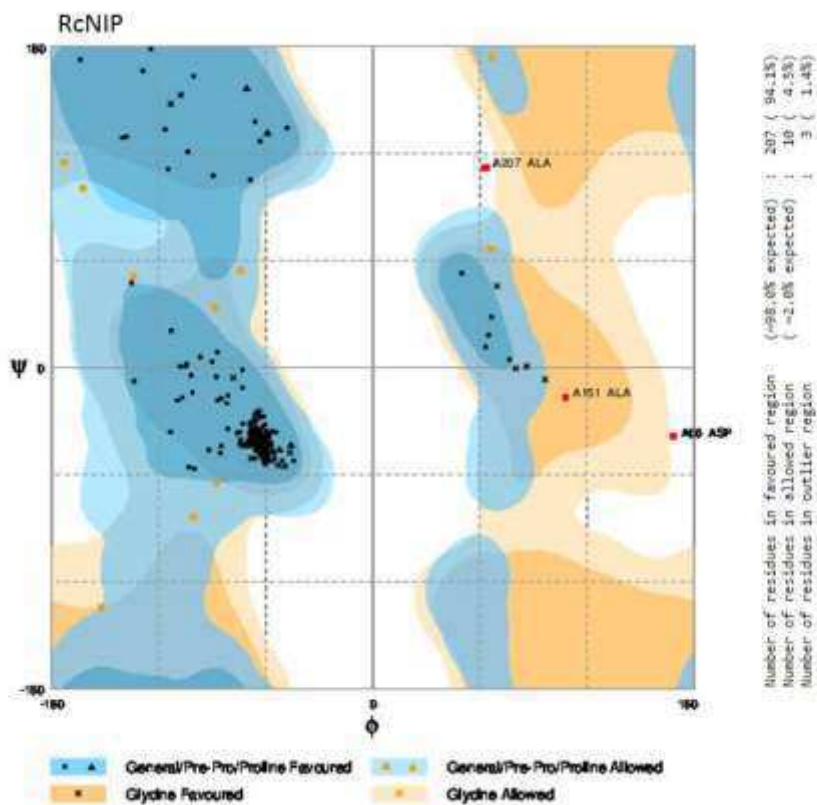


Figura 25. Análise de confiabilidade do modelo gerado para RcPIP via Ramachandran Plot.

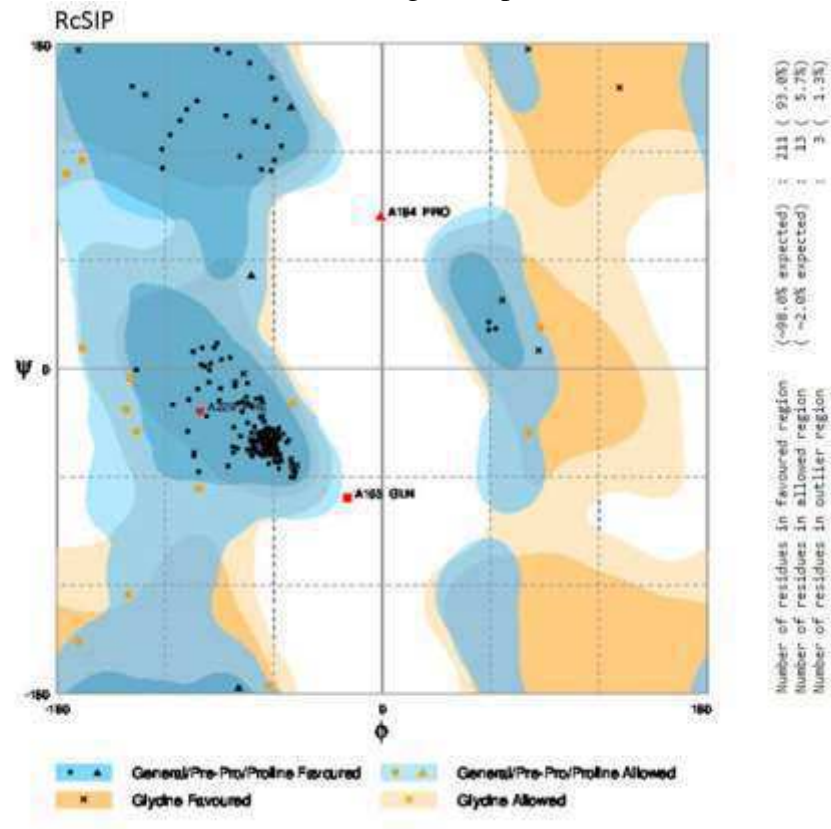
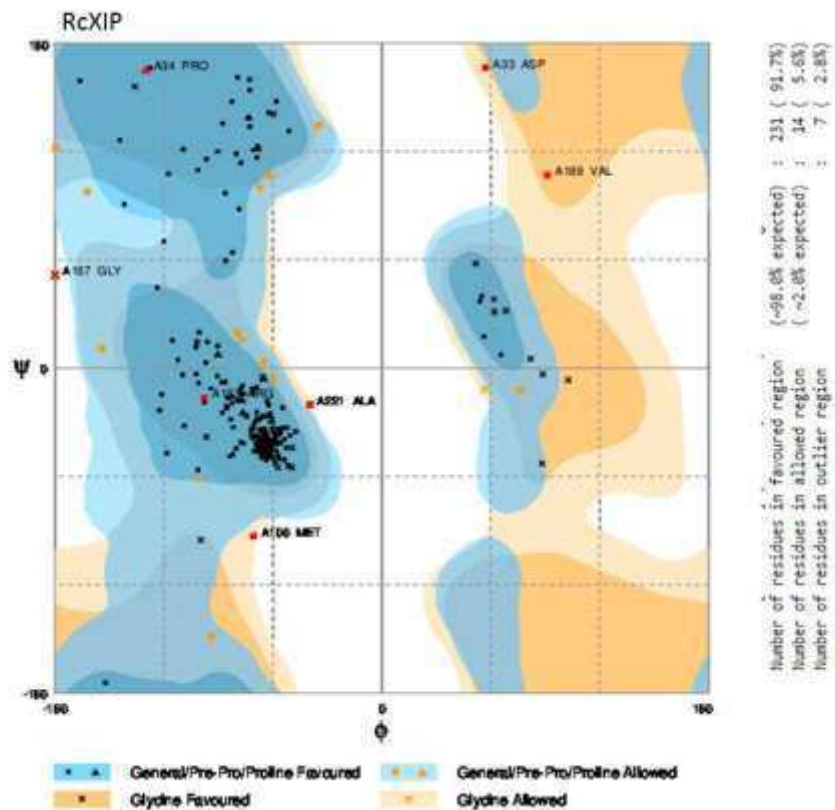


Figura 26. Análise de confiabilidade do modelo gerado para RcXIP via Ramachandran Plot.



### 3.2 Transferabilidade de iniciadores em *C. quercifolius*

Dos 31 pares de iniciadores testados em *C. quercifolius*, 14 apresentaram resultados de amplificação próximos aos tamanhos esperados, abrangendo quatro das cinco subfamílias de aquaporinas descritas para organismos vegetais. Os iniciadores apresentaram temperaturas de anelamento entre 54° a 58°C. Os iniciadores que apresentaram resultado de amplificação em *C. quinidoscolus* são listados na Tabela 6.

Tabela 6. Lista de iniciadores que apresentaram resultados de amplificação no processo de transferabilidade entre *M. esculente*, *R. communis* e *V. unguiculata* em *C. quercifolius*.

Amostra	Iniciadores		Tipo	TA*	TAO**	
1	MePIP2.2	F	5'-TCACCCATCAATCACTCGCT-3'	DNA	55°	1500 bp
		R	5'-TTCTTGTTGGCGGGATTTC-3'			
2	MeXIP1.1	F	5'-TTCAGGCTCATGGATTTGTTCG-3'	DNA	56°	450 bp
		R	5'-CTTCCCATCCTATCCCTTGGT-3'			
3	MeTIP2.2	F	5'-GCAAGAATGGCGAGGATAGC-3'	DNA	56°	400 bp
		R	5'-TGAGACGCAAACCAAGCAA-3'			
4	MeNIP1.1	F	5'-TGAGTGATCGATGTGCAGTTG-3'	DNA	58°	1600 bp
		R	5'-GACGTGCTCTGCTCTTGATG-3'			
5	RcTIP1.1	F	5'-ATCCTTCACCTACGCTCTCC-3'	DNA	58°	1100 bp
		R	5'-AGATAAATCCAGCCAGCCCA-3'			
6	VuPIP1.3	F	5'-CCCTCAGAGCTTACTTCATGGT-3'	cDNA	54°	200 bp
		R	5'-AAGCGATTTCCTTGAATACCAAC-3'			
7	VuPIP1.4	F	5'-CGCTAGAGACTCACACGTTCTCCT-3'	cDNA	54°	180 bp
		R	5'-GCCAAGGTCTTTGTTGAAGATG-3'			
8	VuPIP2.1	F	5'-AAGGACTACCATGACCCCTCCTC-3'	cDNA	54°	200 bp
		R	5'-CTGGCTCTTGTAACCAATGACA-3'			
9	VuPIP2.4	F	5'-ATAAGGGCGTTGTTGTACATGG-3'	cDNA	54°	200 bp
		R	5'-TAATCTCAGCACCCAAAGCAGT-3'			
10	VuTIP1.1	F	5'-GTTAGCTGGACCTGGGACAAC-3'	cDNA	54°	180 bp
		R	5'-AGTAGTCAGTGGTTGGGAGCTG-3'			
11	VuTIP2.1	F	5'-TTTAGCAGCAGGACCATTCTCT-3'	cDNA	58°	200 bp
		R	5'-CACATTGCCATAGATGAGACCA-3'			
12	VuNIP1.1	F	5'-TCCGACTCCTTTGTCTCTGTTC-3'	cDNA	54°	100 bp
		R	5'-CAAACAATTGCTATCCCAGGAA-3'			
13	VuNIP7.1	F	5'-CTCGTGATGCTCTTTCAGGACT-3'	cDNA	54°	100 bp
		R	5'-GCAAGCTTTGACACTCTGCTCT-3'			
14	VuSIP2.1	F	5'-CGCTTCAACAGAGGTAGCACTT-3'	cDNA	54°	100 bp
		R	5'-AAACTGTGGTTGAAGGGTTGAA-3'			

TA= temperatura de anelamento; TAO=Tamanho do amplicom obtido

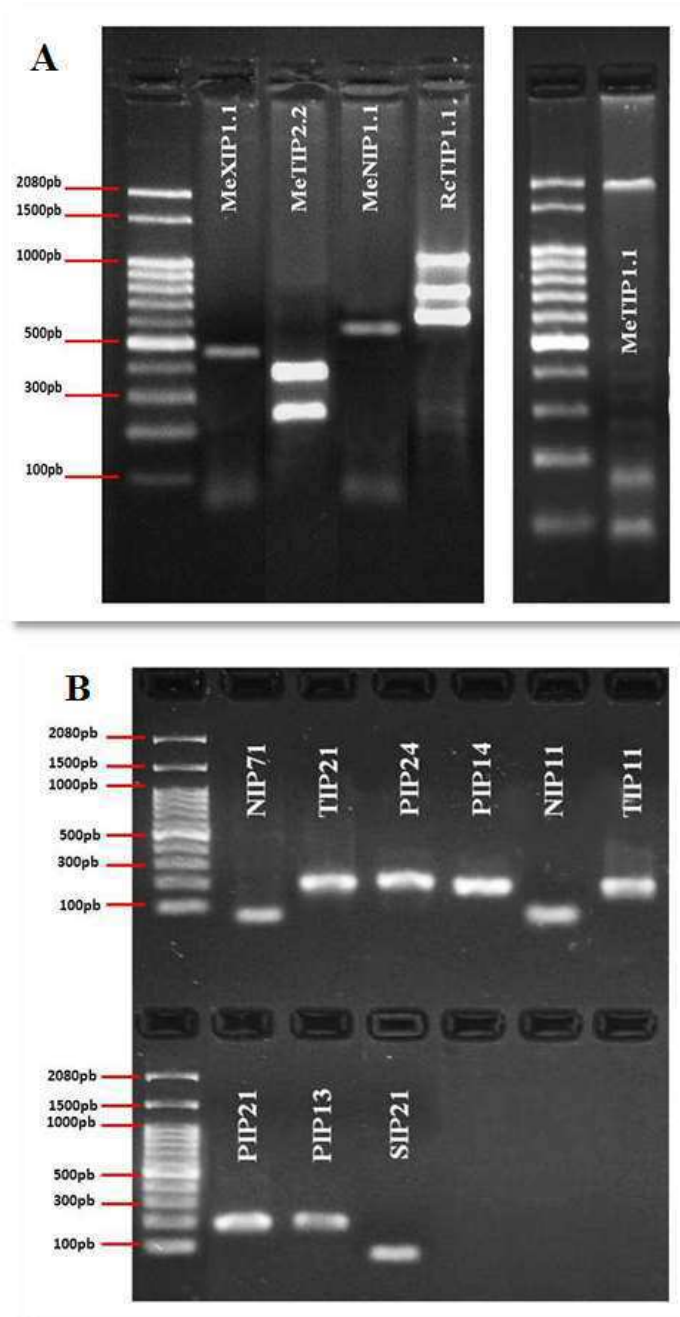
Este resultado mostra que iniciadores para aquaporinas podem ser transferíveis entre membros de diferentes espécies, mesmo quando se trata de espécies filogeneticamente distantes. A taxa média de transferência de iniciadores para aquaporinas (razão entre o número de iniciadores amplificáveis e o número total de iniciadores testados) foi de 45,16%, a



taxa de transferência (relação entre o número de isoformas identificadas e o número total de iniciadores) foi de 54,82%.

A maioria dos iniciadores obteve como resultado de amplificação um único fragmento, porém alguns deles apresentaram bandas múltiplas, as quais foram recortadas separadamente para realização da purificação (Figura 27).

Figura 27. Resultado de testes de amplificação em DNA de *C. quercifolius*. (A) Utilizando iniciadores para aquaporinas desenhados com base em sequências genômicas de *M. esculenta* (Me) e *R. communis* (Rc), (B) Utilizando iniciadores para aquaporinas desenhados com base em sequências expressas de *V. unguiculata*. Usando como padrão de massa molecular DNA Ladder de 100 pb.



Das 17 sequências candidatas obtidas através do processo de transferabilidade de iniciadores e submetidas ao sequenciamento, sete foram confirmadas como aquaporinas, por meio de alinhamentos simples realizados com o auxílio da ferramenta Blastn. Todas as sequências confirmadas como aquaporinas foram provenientes das reações realizadas com iniciadores desenhados com base em sequências expressas de *V. unguiculata*.

Apesar de se tratar de uma espécie de outra família, os iniciadores de *V. unguiculata* apresentaram melhor resultado de amplificação em relação aos iniciadores desenhados com base em sequências de *R. communis* e *M. esculenta* (espécies pertencentes à família euforbiacia, a mesma que *C. quercifolius*). O resultado do alinhamento mostra que as aquaporinas de *C. quercifolius* apresentam alta similaridade com espécies de outras famílias, confirmando o elevado nível de conservação existente entre estas proteínas (Tabela 7).

Tabela 7. Resultado de alinhamento simples realizado contra bando de dados não redundantes utilizando como *query* sequências de *C. quercifolius*.

Iniciador	T	BLASTn					
		Organismo	Família	Descrição	Classif	E-value	Ident
VuPIP1.4 R	188	<i>Vigna angularis</i>	Fabaceae	Provável aquaporina	PIP	9,00E-23	84%
		<i>Phaseolus vulgaris</i>	Fabaceae	Proteína intrínseca da membrana plasmática	PIP1.3	4,00E-21	83%
		<i>Glycine max</i>	Fabaceae	Provável aquaporina	PIP1.2	3,00E-17	79%
VuTIP1.1 F	104	<i>Phaseolus vulgaris</i>	Fabaceae	Proteína intrínseca do tonoplasto	TIP	7,00E-19	87%
		<i>Vigna angularis</i>	Fabaceae	Aquaporina Tipo - TIP	TIP	3,00E-17	86%
		<i>Vigna radiata</i>	Fabaceae	Proteína intrínseca do tonoplasto	TIP1.1	2,00E-15	85%
VuTIP1.1 R	110	<i>Musa acuminata</i> subsp	Musaceae	Provável aquaporina	TIP1.2	1,00E-06	100%
		<i>Hevea brasiliensis</i>	Euphorbiaceae	Proteína intrínseca do tonoplasto	TIP1	1,00E-06	97%
		<i>Solanum tuberosum</i>	Solanaceae	Aquaporina	TIP1.3	7,00E-04	92%
VuPIP1.3 F	122	<i>Hevea brasiliensis</i>	Euphorbiaceae	aquaporina	PIP1.1	9,00E-33	92%
		<i>Populus tremula</i> <i>Populus alba</i>	Salicaceae	Proteína intrínseca da membrana plasmática	PIP1.3	4,00E-21	90%
		<i>Populus trichocarpa</i>	Salicaceae	Proteína intrínseca da membrana plasmática	PIP	4,00E-21	90%
VuTIP2.1 F	119	<i>Hevea brasiliensis</i>	Euphorbiaceae	Proteína intrínseca do tonoplasto	TIP3	7,00E-29	91%
		<i>Malus prunifolia</i>	Rosaceae	Aquaporina	TIP2.1	6,00E-25	89%
		<i>Vernicia fordii</i>	Euphorbiaceae	Aquaporina	*	9,00E-23	93%
VuTIP2.1 R	122	<i>Hevea brasiliensis</i>	Euphorbiaceae	Proteína intrínseca do tonoplasto	TIP3	2,00E-30	91%
		<i>Populus euphratica</i>	Salicaceae	Aquaporina	TIP2.1	2,00E-30	91%
		<i>Ricinus communis</i>	Euphorbiaceae	Aquaporina	TIP2.1	6,00E-30	90%

Continuação da Tabela 7

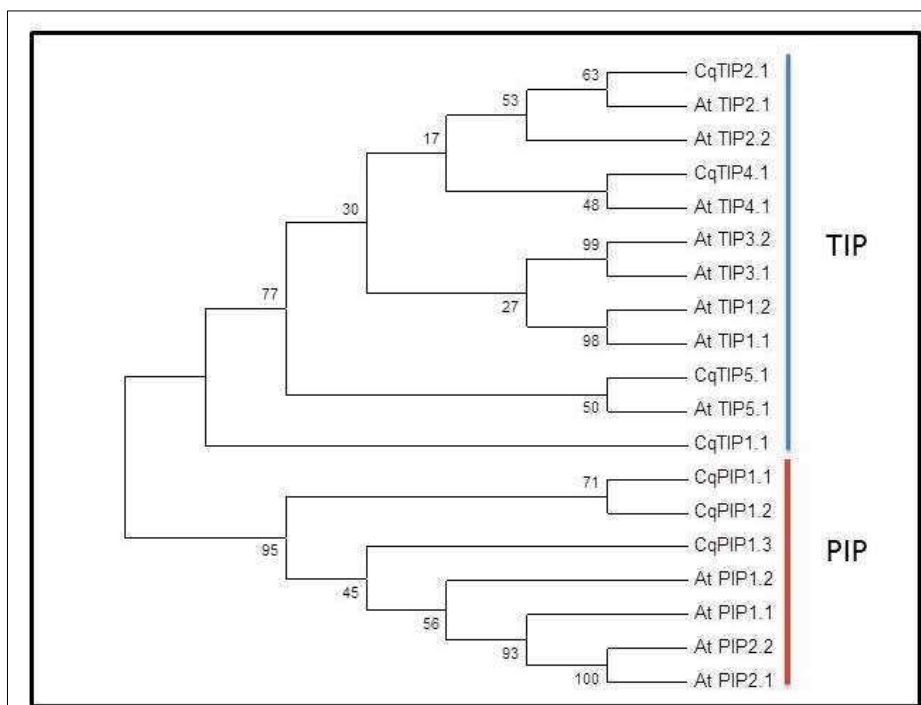
VuTIP2.1 F-R	153	<i>Hevea brasiliensis</i>	Euphorbiaceae	Proteína intrínseca do tonoplasto	TIP3	2,00E-45	91%
		<i>Ricinus communis</i>	Euphorbiaceae	Aquaporina	TIP2.1	3,00E-38	88%
		<i>Populus euphratica</i>	Salicaceae	Aquaporina	TIP2.1	2,00E-39	87%

T=tamanho do produto

Após a confirmação das sequências como aquaporinas, as MIPs de *C. quercifolius* (CqMIPs) foram alinhadas juntamente com sequências bem caracterizadas de *Arabidopsis thaliana*, para classificação das mesmas entre as cinco subfamílias descritas para organismos vegetais. As sete sequências de CqMIPs se distribuíram apenas entre as subfamílias PIP e TIP, confirmando novamente o resultado obtido no Blastn.

O alinhamento realizado com o auxílio da ferramenta ClustalW, bem como o fenograma construído pelo método de *Neighbor-Joining*, permitiram avaliar o grau de similaridade entre as sequências de aquaporinas de *A. thaliana* e *C. quercifolius*. O resultado obtido através da análise frenética possibilitou inferir nomenclatura para as aquaporinas de *C. quercifolius*, tendo como base a homologia entre as sequências. As sete CqMIPs se distribuem entre as subfamílias PIP e TIP da seguinte forma: PIP1 (três membros), TIP1 (um membro), TIP2 (um membro), TIP4 (um membro) e TIP5 (um membro) (Figura 28).

Figura 28. Dendrograma de membros das famílias de aquaporinas de *C. quercifolius* juntamente com *A. thaliana* pelo método de *Neighbor Joining* (*bootstrap* ao nível de 1.000 réplicas).



As sequências proteicas obtidas via *ORF-Finder* foram analisadas através da ferramenta Batch CD-Search, para avaliar a presença de domínio conservado. Apesar de se tratar de sequências curtas, todas apresentaram regiões parciais de domínio MIP (incompletos em uma das extremidades), confirmando o resultado obtido na busca por ortólogos realizada via Blastn (Tabela 8).

Tabela 8. Identificação e descrição de domínio conservado em aquaporinas de *C. quercifolius*.

ORF		CD-Search Blastx			
Sequência	bp  aa	Domínio	DN	Descrição	E-value
CqPIP1.1	186 61	C	MIP sf	Membros da superfamília MIP	6e-12
CqTIP1.1	102 33	N	MIP sf	Membros da superfamília MIP	2.50e-13
CqPIP1.2	120 39	N	MIP sf	Membros da superfamília MIP	6e-15
CqTIP5.1	117 38	N	MIP sf	Membros da superfamília MIP	5.09e-12
CqPIP1.3	120 39	C	MIP sf	Membros da superfamília MIP	6e-17
CqTIP2.1	150 49	N	MIP sf	Membros da superfamília MIP	44e-24
CqTIP4.1	153 40	N	MIP sf	Membros da superfamília MIP	2.79e-19

DN= Domain name; N= domínio incompleto na extremidade N-terminal; C= domínio incompleto na extremidade C-terminal; sf= superfamília; pb=pares e bases; aa= aminoácidos.

## 4 DISCUSSÃO

### 4.1 Identificação e classificação de MeAQPs e RcAQPs

O número elevado de genes descrito nesse estudo é semelhante ao observado em outras espécies vegetais, tal como *Oryza sativa*, com 34 membros (SAKURAI et al., 2005), 36 de *Zea mays* (CHAUMONT et al., 2001), *Phaseolus vulgaris*, com 41 membros (ARIANI; GEPTS, 2015), e outras espécies da família Euphobiaceae, como *Jatropha curcas*, com 32 membros (ZOU et al., 2016) e *Hevea brasiliensis*, com 51 membros (ZOU et al., 2015). Esta diversidade de isoformas existente em organismos vegetais está relacionada à variedade de funções que estes transportadores desempenham, bem como às diferentes seletividades e afinidades por solutos (LIU et al., 2009).

O número de AQPs de *R. communis* identificado nesse estudo foi semelhante ao descrito por Zou et al. (2015). Em comparação às AQPs em euforbiáceas, *R. communis* apresentou número de genes aproximado ao encontrado em *Jatropha curcas* (32 AQPs) (ZOU et al., 2016), além de outros genomas de referência, a exemplo de *Arabidopsis thaliana* (35 AQPs) (JOHANSON et al., 2001), *Zea mays* (36 AQPs) (CHAUMONT et al., 2001) e *Oryza*

*sativa* (33 AQPs) (SAKURAI et al., 2005). Os resultados são ainda mais interessantes quando comparamos as 42 AQPs de *M. esculenta* aqui identificadas, cujo número foi próximo-superior às 41 AQPs descritas em *Symphytum tuberosum* (VENKATESH et al., 2013) e *Phaseolus vulgaris* (ARIANI; GEPTS, 2015), e as 40 AQPs em *Cicer arietinum* (DEOKAR; TAR'AN, 2016).

Ambas as espécies apresentaram número de genes inferior a *Hevea brasiliensis*, com 51 AQPs (ZOU et al. 2015), bem como a espécies de outras famílias, como *Musa acuminata* (51) (DESHMUKH et al., 2015), *Populus trichocarpa* (55) (GUPTA; SANKARARAMAKRISHNAN, 2009), *Brassica oleracea* (67) (DESHMUKH et al., 2015) e *Glycine max* (72) (DESHMUKH et al., 2013).

O número de genes preditos em *Hevea brasiliensis* é duas vezes maior do que em *R. communis* (CHAN et al., 2010; RAHMAN et al., 2013; WU et al., 2015), uma vez que a *R. communis* não foi submetida a nenhum evento recente de duplicação de genoma completo (WGD), apresentando número de genes mais próximo da *Jatropha curcas*, o qual também não foi submetido a nenhum processo de WGD (CHAN et al., 2010; WU et al., 2015). Os 51 genes encontrados na seringueira (ZOU et al., 2015a) são comparáveis aos 55 em álamo (GUPTA; SANKARARAMAKRISHNAN, 2009), uma espécie de árvore da família Salicaceae que sofreu recente duplicação de genoma (TUSKAN et al., 2006).

A poliploidia resultante da duplicação do genoma é muito comum em plantas, refletindo sua extensa redundância genética. No caso da *Glycine max*, 75% de seus genes têm mais de uma cópia como resultado de dois eventos de duplicação do genoma (13 e 59 milhões de anos atrás) (ROULIN et al., 2013). Em *Populus trichocarpa*, apenas seis dos 55 genes das aquaporinas encontrados no genoma estão presentes como uma única cópia (COHEN et al., 2013).

A realização da identificação de aquaporinas baseada em homologia de sequências em espécies que apresentam genoma sequenciado é viável devido ao alto nível de conservação destas proteínas entre espécies de plantas, apresentando elevada similaridade entre sequências pertencentes a outras espécies, o que já foi demonstrado em trabalhos anteriores como CHAUMONT et al. (2001), DESHMUKH et al. (2016) e ZOU et al. (2016).

A diversificação precoce das MIPs resultou em um grande número de subfamílias ainda em plantas terrestres primitivas, sendo que, durante a evolução, algumas plantas superiores foram perdendo duas destas subfamílias (HIP e GIP), enquanto as subfamílias restantes se expandiram e, em alguns casos, diversificaram-se, resultando na formação de grupos mais especializados dentro destas subfamílias (DANIELSON; JOHANSON, 2008).

No caso das monocotiledôneas, a subfamília XIP também se encontra ausente em todo o grupo, assim como nas *Brassicaceae* (DESHMUKH et al., 2015; DESHMUKH et al., 2016). Com base nisso, DESHMUKH et al. (2016) sugerem que, para a realização da análise de busca por homologia, é importante ampliar o número de sequências sondas para o maior número de espécies possíveis (mono e dicotiledôneas), favorecendo, assim, a chance de obter o maior número de subfamílias e subgrupos de AQPs.

#### 4.1.1 Análise comparativa e fenética das subfamílias de MeAQPs e RcAQPs

De acordo com os fenogramas apresentados (Figuras 2A e 3A), as sequências MePIPs foram divididas em dois grupos distintos: MePIP1 e MePIP2, semelhante à classificação realizada com sequências de AQPs de outras espécies, a exemplo do feijão-caupi, que apresenta quatro PIP1 e nove PIP2, confirmando as PIPs como a subfamília com maior número de membros entre as AQPs (JOHANSON et al., 2001; BEZERRA-NETO, 2012)

Os resultados de ambos os alinhamentos demonstram baixos níveis de divergência entre as sequências de PIPs. Estudos realizados com o musgo *Physcomitrella patens* (briófita) sugerem que esses grupos devem ter se formado no início da evolução das plantas terrestres e são de fundamental importância na fisiologia da planta (DANIELSON; JOHANSON, 2008).

A análise fenética (Figuras 2B e 3B), realizada com as sequências de MeTIPS e RcTIPS, confirma a divisão proposta para subfamília em cinco subgrupos. Acredita-se que as TIPS em plantas superiores evoluíram com um maior número de subgrupos em relação às PIPs devido às exigências funcionais especiais de diferentes vacúolos (JOHANSON et al., 2001; DANIELSON; JOHANSON, 2008, 2010).

Para a subfamília NIP, ambas as espécies apresentam classificação diferente da proposta para *Arabidopsis*, que comporta sete subgrupos, enquanto *M. esculenta* e *R. communis* apresentam apenas seis subgrupos. *M. esculenta* não apresenta membros do subgrupo NIP5 e *R. communis* do NIP3 (QUIGLEY et al., 2001; JOHANSON et al., 2001).

O resultado obtido nos alinhamentos realizados com as sequências da subfamília SIPs sugerem que esta é a mais divergente entre as cinco subfamílias descritas em vegetais superiores, corroborando resultados obtidos em estudos anteriores (CHAUMONT et al., 2001; DANIELSON; JOHANSON, 2010; BEZERRA NETO, 2012).

AS XIPs presentes no genoma de *Manihot* e *Ricinus* apresentaram-se distribuídas em três grupos, semelhante à classificação proposta para as XIPs de *Populus* (que é suportado pela semelhança entre as sequências e as modificações existentes no filtro de ar / R) e

diferentemente da maioria das espécies que possuem apenas XIP1 e XIP2 (LOPEZ et al., 2012)

Os dados gerados a partir das tecnologias de sequenciamento tornam possível não apenas a identificação e classificação de sequências de fundamental importância para o desenvolvimento dos organismos (BURGESS, 2016; CHEN et al, 2016), como possibilita, também, integrar informações por meio de caracterização *in silico* para uma melhor compreensão das funções gênicas (SONAH et al., 2016; SONG et al., 2016).

#### 4.1.2 Caracterização de sequências de MeAQPs e RcAQPs

A caracterização molecular de AQPs permite destacar vários aspectos importantes para compreensão da estrutura e funcionalidade destas proteínas, ressaltando informações como distribuição, organização genética, evolução e características conservadas (DESHMUKH et al., 2016)

Os valores médios de massa molecular e ponto isoelétrico encontrados para MeMIPs e RcMIPs são semelhantes aos valores propostos para AQPs de outras Euphorbiaceae, como *Jatropha curcas* (ZOU et al., 2016 a) e *Hevea brasiliensis* (ZOU et al., 2016 b), como do próprio *Ricinus communis*, proposto por Zou et al. (2015), como também para *Sorghum bicolor* (REDDY et al., 2015) e *Musa acuminata* (HEMMATI, 2016).

As AQPs das plantas mostraram ter propriedades bioquímicas diferentes associadas às características da função (JOHANSON; GUSTAVSSON, 2002; HOVE; BHAVE, 2011).

#### 4.1.3 Motivos e domínios conservados

O motivo AEF é uma região conservada presente apenas em PIP, TIP, NIP. Localizado na extremidade N-terminal, aproximadamente entre 50 e 60 aminoácidos antes do motivo NPA presente no loop B. A subfamília PIP apresenta a menor variação de aminoácidos nessa região. Apesar de se tratar de um motivo conservado, até o momento, não foi identificada a função desempenhada por este (ZARDOYA; VILLALBA, 2001).

Esses peptídeos apresentam dois motivos NPA (asparagina-prolina-alanina) localizados nos loops B e E, os quais são responsáveis pela formação da constrição central do poro (CHAUMONT et al., 2005; KRUSE; UEHLEIN; KALDENHOFF, 2006). Outra região responsável pela seletividade das aquaporinas é a região arginina/aromática (ar/R), formada por quatro resíduos de aminoácidos (R1-R4), que estão diretamente relacionados à

seletividade do substrato transportado (BIENERT et al., 2008; LUDEWIG; DYNOWSKI, 2009).

Os motivos NPA e os resíduos ar / R são bastante conservados em PIPs, sugerindo que todas estas, independente de subgrupo, apresentam especificidade semelhante para substratos (DANIELSON; JOHANSON, 2008). As PIPs possuem, na posição H5 da região ar/R, um resíduo de histidina (H), que proporciona uma alta especificidade no transporte de água, entretanto o subgrupo PIP2 apresenta maior eficiência de transporte em relação a PIP1 (CHAUMONT et al., 2000; WALLACE; ROBERTS, 2004; FORREST; BHAVE, 2007; MAUREL et al., 2008). Essa diferença na eficiência de transporte levanta a hipótese de que, mesmo sendo a região ar/R a principal determinante da afinidade de transporte, devem existir outras regiões relacionadas ao controle do transporte nestas aquaporinas.

As TIPs apresentam motivos NPAs bem conservados, com alterações apenas nos resíduos da região ar/R presentes no LE1 e LE2, que possivelmente estão relacionados aos substratos os quais essas proteínas têm a capacidade de transportar. As sequências de SIPs apresentam extremidade N-terminal curta e modificações no motivo NPA, localizado no *loop* B, o que confere uma maior abertura no poro, gerando uma alta especificidade por substratos diferentes (GOMES et al., 2009).

As sequências de XIPs apresentam variações nos aminoácidos nas regiões correspondentes ao primeiro motivo NPA e na região ar/R (MITANI-UENO *et al.*, 2011). Devido a essas alterações, foi levantada a hipótese de que XIPs não são funcionais como canais de água, sendo estas responsáveis pelo transporte de solutos hidrofóbicos (GUPTA; SANKARARAMAKRISHNAN, 2009). Estudos realizados por Bienert et al. (2011) com as espécies *Nicotiana tabacum*, *Solanum Lycopersico* e *Solanum tuberosum* confirmaram a funcionalidade das XIPs como transportadores de solutos hidrofóbicos.

#### 4.1.4 Análise fenética e distribuição cromossômica

A análise confirmou a ausência de membros da subfamília XIP em *A. thaliana*, em comparação com a *M. esculenta* e *R. communis*, que apresentaram quatro e cinco membros, respectivamente. A subfamília SIP apresenta em *M. esculenta*, três membros, semelhantemente ao proposto para *A. thaliana*, diferentemente de *R. communis*, que apresenta quatro membros.

Com a exceção da subfamília XIP, a classificação das subfamílias de MeMIP e RcMIP em subgrupos é semelhante à *A. thaliana*, ou seja, 2 subgrupos PIP, 5 subgrupos TIP, 7



subgrupos NIP e 2 subgrupos SIP. Uma vez que nenhum homólogo de XIP foi encontrado no genoma de *A. thaliana*, a nomenclatura proposta por Lopez et al. (2012) foi adotada para classificar as MeXIPs e RcXIPs em três subgrupos (ZOU et al., 2015a, b).

Foi constatada a presença de MeAQPs em 15 dos 18 cromossomos existentes na espécie. Nenhum loci foi detectado nos cromossomos 6, 15 e 18 (Figura 6), sugerindo que as AQPs são genes abundantes que participam do metabolismo celular (REDDY et al., 2015).

A distribuição dos genes MeAQPs entre os cromossomos, aparentemente, não obedece a nenhum padrão, como já observado em outras espécies, a exemplo de *Sorghum bicolor* (REDDY et al., 2015) e *Musa acuminata* (HEMMATI, 2016), apenas os membros da subfamília SIP se encontram todos no mesmo cromossomo.

#### 4.1.5 Modelagem comparativa

Na análise do desvio total de energia via Z-score, foi possível visualizar que os modelos MeMIPs e RcMIPS gerados encontravam-se nas posições referentes aos modelos obtidos via raios-X e ressonância magnética (Figura 9), indicando que os modelos estão intimamente relacionados às estruturas terciárias de proteínas obtidas experimentalmente.

A estrutura estereoquímica dos modelos propostos para aquaporinas de *M. esculenta* e *R. communis* (via Ramachandaran, programa Procheck) apresentou valor médio de 90,3% de quantidade de resíduos distribuídos nas regiões energeticamente mais favoráveis para um modelo de qualidade, em todos os modelos propostos para as MeMIPs e Rc MIPs. O valores de Zscores variaram entre -2,71 e -5,41 (média de -3,76) (quanto mais negativos, maior nível de confiabilidade para o modelo), confirmando a qualidade dos modelos gerados.

#### 4.2 Transferabilidade de iniciadores em *C. quercifolius*

Das 17 sequências candidatas obtidas através do processo de transferabilidade de iniciadores e subsequentemente submetidas ao sequenciamento, sete foram confirmadas como aquaporinas, por meio de alinhamentos simples realizados com o auxílio da ferramenta Blastn. Todas as sequências confirmadas como aquaporinas foram provenientes das reações realizadas com iniciadores desenhados com base em sequências expressas de *V. unguiculata*.

Apesar de se tratar de uma espécie de outra família, os iniciadores de *V. unguiculata* apresentaram melhor resultado de amplificação em relação aos iniciadores desenhados com base em sequências de *R. communis* e *M. esculenta* (espécies pertencentes à mesma família

que a *C. quercifolius*). O resultado do alinhamento mostra que as aquaporinas de *C. quercifolius* apresentam alta similaridade com espécies de outras famílias, confirmando o elevado nível de conservação existente entre estas proteínas (Tabela 7) (AGRE, 2004).

Por se tratar de um processo de transferabilidade de iniciadores entre espécies distintas, é esperado que, com os pares de iniciadores de PIP e TIP, fossem observados resultados de amplificação. Isso foi possível por estas subfamílias serem consideradas mais conservadas entre as cinco existentes em plantas superiores (CHAUMONT et al., 2001; DESHMUKH; SONAH; BÉLANGER, 2016).

A diferenciação da expressão de aquaporinas está relacionada aos órgãos específicos, estágio de desenvolvimento do vegetal e condições fisiológicas do mesmo (HEINEN et al., 2009), o que pode ter favorecido a amplificação de sequências de CqPIPs e CqTIPs, já que o material genético utilizado para reações é proveniente unicamente de folhas e trata-se de cDNA. As subfamílias NIP, SIP e XIP são mais específicas no transporte de pequenos solutos, sendo sua expressão menos abundante nos diferentes tecidos, quando comparadas às subfamílias PIP e TIP (PARK et al., 2010).

A presença de aquaporinas da subfamília PIP na folha da *C. quercifolius* confere mais um suporte à adaptação a regiões de clima semiárido, já que, além do transporte de água, as PIPs também participam da regulação da atividade fotossintética (AROCA et al., 2006). Algumas plantas apresentam menores números de isoformas de AQPs em relação a outras, o que pode ser o caso desta espécie, atuando como mecanismo de defesa da planta para evitar a perda excessiva água (HUSSAIN et al., 2011). Para alguns autores, as plantas impedem a desidratação ao “desligar” as aquaporinas durante períodos de seca (GASPAR, 2011).

## 5 CONSIDERAÇÕES FINAIS

As novas tecnologias de sequenciamento têm permitido a geração de um número expressivo de dados, favorecendo a prospecção de genes candidatos responsivos a estresses, incluindo aqueles envolvidos na captura e transporte seletivo de água e íons, como as aquaporinas. A identificação baseada em homologia de AQPs em espécies sequenciadas é viável devido ao alto nível de conservação de sequências destas proteínas entre espécies de plantas.

A caracterização genômica de AQPs destacou vários aspectos importantes, como distribuição, organização genética, conservação, estrutura e evolução, possibilitando uma melhor compreensão do papel desempenhado pelas mesmas. Com as preocupações atuais

sobre os recursos hídricos, muitos estudos têm sido desenvolvidos na perspectiva de compreender o papel das aquaporinas no sistema de transporte de água e pequenos solutos, favorecendo o desenvolvimento do vegetal sobre condições ambientais estressantes, fazendo destas proteínas alvos biotecnológicos importantes para o desenvolvimento de cultivares mais tolerantes, bem como possibilitar o manejo de plantas melhor adaptadas às condições ambientais.

A maior parte dos trabalhos com aquaporinas são realizados com plantas modelo e espécies de interesse comercial. Estudos realizados com espécies típicas de regiões semiáridas, que apresentam importantes adaptações morfológicas, fisiológicas e bioquímicas a condições de déficit hídrico, têm sido negligenciados. Os resultados obtidos neste estudo permitem concluir que iniciadores de aquaporinas podem ser transferíveis entre espécies de plantas, até mesmo entre espécies de diferentes famílias, possibilitando a busca de sequências de AQPs em plantas que não apresentam sequências disponíveis em banco de dados, como é o caso da maioria das espécies das regiões tropicais.

Foram identificados, no genoma da *C. quercifolius*, apenas representantes das subfamílias PIP e TIP. O material genético utilizado nesta pesquisa foi extraído apenas de tecido foliar, e os iniciadores utilizados foram desenhados com base em sequências de *M. esculenta*, *R. communis* e *V. unguiculata*. Assim, para uma melhor caracterização e identificação de genes da família MIP em *C. quercifolius*, faz-se necessário o desenho de iniciadores específicos e a extração de DNA das várias partes da planta (em diferentes estágios do desenvolvimento).

## 6 CONCLUSÃO

Foram obtidas sete sequências de CqAQPs por meio de sequenciamento. As mesmas foram classificadas e caracterizadas *in silico*, confirmando se tratar de sequências de aquaporinas. Os produtos de sequenciamento consistem apenas em isolamento parcial de CqMIPs, não podem, assim, serem caracterizadas adequadamente, porém este é o primeiro relato de aquaporinas em *C. quercifolius*, podendo então serem usadas para o desenho de iniciadores específicos para a espécie.

## 7 REFERÊNCIAS

- AFZAL, Z., et al. The Roles of Aquaporins in Plant Stress Responses. **J. Dev. Biol.**, v. 4, p. 9, 2016.
- AGRE, P. Aquaporin water channels (Nobel lecture). **Angewandte Chemie International Edition**, v. 43, n. 33, p. 4278-4290, 2004.
- ALTSCHUL, S.F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.
- ARIANI, A., GEPTS, P. Genome-wide identification and characterization of aquaporin gene family in common bean (*Phaseolus vulgaris* L.). **Mol. Genet. Genomics**, v. 290, p. 1771–178, 2015
- AROCA, R. et al. Drought, abscisic acid and transpiration rate effects on the regulation of PIP aquaporin gene expression and abundance in *Phaseolus vulgaris* plants. **Annals of botany**, v. 98, n. 6, p. 1301-1310, 2006.
- AZAD, A. K et al. Substitution of a single amino acid residue in the aromatic/arginine selectivity filter alters the transport profiles of tonoplast aquaporin homologs. **Biochimica et Biophysica Acta**, v. 1818, n. 1, p. 1-11, 2012 (*in press*).
- BARROSO, G. M.; COSTA, C. G.; GUIMARÃES, E. F.; ICHASO, C. L.; PEIXOTO, A. L. **Sistemática de angiospermas do Brasil**. 2. ed. Viçosa, MG: UFV, v. 1, p. 304. 2002.
- BENKERT, P.; SCHWEDE, T.; TOSATTO, S.C.E. QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. **BMC Structural Biology**, n. 20, p. 9-35, 2009.
- BENKO-ISEPPON, A.M. et al. Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. **Revista Brasileira de Geografia Física**, v. 06, p. 1112-1134, 2011.
- BEZERRA-NETO, J.P. **Caracterização e análise evolutiva in silico de aquaporinas no transcriptoma do feijão-caupi (*Vigna unguiculata* (L.) Walp.** Dissertação de Mestrado. UFPE. 2012.
- BIENERT, G.P. et al. Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. **The Plant Journal**, 2011.
- BIENERT, G.P.; SCHÜSSLER, M.D.; JAHN, T.P. Metalloids: essential, beneficial or toxic? Major intrinsic proteins sort it out. **Trends in Biochemical Sciences**, v. 33, p. 20–26, 2008.
- BIJEKAR, S; GAYATRIM.C. Ethanomedicinal properties of Euphorviaceae family-A comprehensive review. **International Journal of Phytomedicine**, v. 6, , p. 144-156, 2014.
- BURGESS, D. J. Gene expression: a space for transcriptomics. **Nat. Rev. Genet.** v. 17, p. 436–437, 2016.

CHAN A.P. et al. Draft genome sequence of the oilseed species *Ricinus communis*. **Nat Biotechnol**, v. 28(9), p. 951–956, 2010.

CHAUMONT, F. et al. Aquaporins constitute a large and highly divergent protein family in maize. **Plant Physiology**, v. 125, p. 1206–1215, 2001.

CHAUMONT, F.; MOSHELION, M.; DANIELS, M.J. Regulation of plant aquaporin activity. **Biology of the Cell**, v. 97, p. 749–764, 2005.

CHEN, W. et al. Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. **Front. Plant Sci.** v. 7, p. 1044. 2016.

CHOU, K.C. AND SHEN, H.B. Cell-PLoc: A package of web-servers for predicting subcellular localization of proteins in various organisms. **Nature Protocols**, v. 3, p. 153–162, 2008.

CLAEYS H, INZE D. The agony of choice: how plants balance growth and survival under water-limiting conditions. **Plant Physiology**, v. 162, p. 1768–1779, 2013.

COHEN D., et al. "Developmental and environmental regulation of Aquaporin gene expression across *Populus* species: divergence or redundancy?". **PLoS One**, v. 8.2, e55506, 2013.

CORRÊA, A. D. et al. Farinha de folhas de *M. esculenta*: I. efeito da secagem das folhas sobre a atividade da linamarase. Tese de Doutorado, **Ciências e Agrotecnologia**, Lavras, v. 26, n. 2, p. 368–374, 2002.

DANIELSON, J. A H.; JOHANSON, U. Phylogeny of major intrinsic proteins. **Advances in experimental medicine and biology**, v. 679, p. 19–31, 2010.

DANIELSON, J.A.; JOHANSON, U. Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. **BMC Plant Biology**, v. 8, n. 45, 2008.

DEOKAR A.A., TAR'AN B. Genome-Wide Analysis of the Aquaporin Gene Family in Chickpea (*Cicer arietinum* L.). **Front. Plant Sci.**, v.7, p. 1802, 2016.

DESHMUKH R.K., SONAH H., BÉLANGER R.R. Plant Aquaporins: Genome-Wide Identification, Transcriptomics, Proteomics, and Advanced Analytical Tools. **Front. Plant Sci.**, v. 7, p. 1896. 2016

DESHMUKH R.K. et al. Identification and functional characterization of silicon transporters in soybean using comparative genomics of major intrinsic proteins in *Arabidopsis* and rice. **Plant Mol Biol**. v. 83(4–5), p. 303–315. 2013.

DESHMUKH, R. K. et al. A precise spacing between the NPA domains of aquaporins is essential for silicon permeability in plants. **Plant J.**, v. 83, p. 489–500, 2015.

DRUMOND et al, **Estratégias para o Uso Sustentável da Biodiversidade da Caatinga**. Petrolina. 2000.

- ENDRESS, P. K., DAVIS, C. C., MATTHEWS, M. L. Advances in the floral structural characterization of the major subclades of Malpighiales, one of the largest orders of flowering plants. **Ann. Bot.** 111, 969–985. 2013
- FORREST, K.L.; BHAVE, M. Major intrinsic proteins (MIPs) in plants: a complex gene family with major impacts on plant phenotype. **Functional & Integrative Genomics**, v. 7, p. 263-289, 2007.
- FORZZA, R. C. et al. INSTITUTO DE PESQUISAS JARDIM BOTÂNICO DO RIO DE JANEIRO. *Catálogo de plantas e fungos do Brasil* [online]. Rio de Janeiro: Andrea Jakobsson Estúdio: Instituto de Pesquisa Jardim Botânico do Rio de Janeiro,. Introdução: as angiospermas do Brasil, p. 78-89. Vol. 1. ISBN 978-85-8874-242-0, 2010.
- GASPAR, M. Aquaporinas: de canais de água a transportadores multifuncionais em plantas. **Revista Brasil. Bot.**, v.34, n.4, p.481-49. 2011.
- GOLFARI, L., CASER, R. L. **Zoneamento ecológico da região nordeste para experimentação florestal**. Série técnica. PRODEPEF, Brasília (10): 1 - 116, 1977
- GOMES, D. et al. Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. **Biochimica et Biophysica Acta**, v. 1788, p. 1213-1228, 2009.
- GUPTA, A.B.; SANKARARAMAKRISHNAN, R. Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: characterization of XIP subfamily of aquaporins from evolutionary perspective. **BMC Plant Biology**, v. 9, p. 134, 2009.
- HALL, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95-98, 1999.
- HEINEN, R.B.; YE, Q.; CHAUMONT, F. Role of aquaporins in leaf physiology. **J. Exp. Bot.**, v. 60, n. 2971-2985, 2009.
- HELLMANN, J. J. et al. Five potential consequences of climate change for invasive species. **Conservation Biology**, v. 22, p. 534-543, 2008.
- HEMMATI, S., Predicting the functionality of major intrinsic proteins: An *in silico* analysis in *Musa* . **Trends in Pharmaceutical Sciences**, v. 2(2), p. 139-150, 2016.
- 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. 2441-3, 2006.
- HOVE, R. M., AND BHAVE, M. Plant aquaporins with non-aqua functions: deciphering the signature sequences. **Plant Mol. Biol.** v. 75, p. 413–430, 2011.
- HUSSAIN, S.S.; IQBAL, M.T.; ARIF, M.A.; AMJAD, M. Beyond osmolytes and transcription factors: drought tolerance in plants via protective proteins and aquaporins. **Biologia Plantarum**, v. 55, p. 401-413, 2011.

JARVIS A. et al. Is Cassava the Answer to African Climate Change Adaptation?. **Tropical Plant Biol**, v. 5, p. 9-29, 2012.

JOHANSON, U. et al. The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. **Plant Physiology**, v. 126, p. 1358–1369, 2001.

JOHANSON U, GUSTAVSSON S. A new subfamily of major intrinsic proteins in plants. **Mol Biol Evol**, v. 19, p. 456–461, 2002.

KRUSE, E.; UEHLEIN, N.; KALDENHOFF, R. The aquaporins. **Genome Biology**, v. 7, p. 1-6, 2006.

KRZYWINSKI M. et al. Circos: an information aesthetic for comparative genomics. **Genome Res.** (2009) 19: 1639-1645

LASKOWSKI, R.A. et al. A program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283-291, 1993.

LIU, Q.; WANG, H.; ZHANG, Z.; et al. Divergence in function and expression of the NOD26-like intrinsic proteins in plants. **BMC genomics**, v. 10, p. 313, 2009.

LOPEZ, G. et al. Insights into *Populus* XIP aquaporins: evolutionary expansion, protein functionality, and environmental regulation. **Journal of Experimental Botany**, v. 63, n. 5, p. 2217–2230, 2012

LUCENA, M. F. A.; ALVES, M. **Notas taxonômicas para Euphorbiaceae s.l. do Nordeste do Brasil**. Hoehnea, v. 37(1), p. 71-85, 2010.

LUDEWIG, U.; DYNOWSKI, M. Plant aquaporin selectivity: where transport assays, computer simulations and physiology meet. **Cellular and molecular life sciences: CMLS**, v. 66, n. 19, p. 3161-75, 2009.

MAHAJAN, S.; TUTEJA, N. Cold salinity and drought stresses: an overview. **Archives of Biochemistry and Biophysics**, v. 444, p. 139-158, 2005.

MAUREL, C. et al. Plant aquaporins: membrane channels with multiple integrated functions. **Annual Review of Plant Biology**, v. 59, p. 595-624, 2008.

MCDOWELL N, POCKMAN WT, ALLEN CD, *et al*. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought?. **New Phytologist**, v.178, p. 719–739, 2008.

MELO, A. L.;SALES, M. F. O gênero *Cnidocolus* Pohl (Crotonoideae-Euphorbiaceae) no Estado do Pernambuco, Brasil. **Acta Botanica Brasilica**. v. 22(3), p. 806-827, 2008.

MICHAELS, S. D.; JOHN, M. C.; AMASINO, R. M. Removal of polysaccharides from plant DNA by ethanol precipitation. **Biotechniques**, v. 17, p. 274-276p, 1994.

MITANI-UENO N, YAMAJI N, ZHAO FJ, MA JF. The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for silicon, boron, and arsenic. **Journal of Experimental Botany**. v. 62, p.4391–4398, 2011.

NEPOMUCENO, R.C. et al. Inclusão da quirera de arroz em rações de suínos na fase de creche. *Rev. Bras. Saúde Prod. Anim.*, v.12, p.718-728, 2011

OGUNNIYI DS. Castor oil: a vital industrial raw material. **Bioresour Technol**, v. 97(9), p. 1086–1091, 2006.

MELO, A. L.; SALES, M. F. O gênero *Cnidocolus* Pohl (Crotonoideae-Euphorbiaceae) no Estado do Pernambuco, Brasil. **Acta Botanica Brasilica**. 22(3):806-827. 2008.

PAGE, R.D.M. TreeView: An application to display phylogenetic trees on personal computers. **Computer applications in the biosciences**, v. 12, p. 357-358, 1996.

PARK, W. J, CAMPBELL, B. T. Aquaporins as targets for stress tolerance in plants: genomic complexity and perspectives. **Turk J Bot.**, v. 39, p. 879-886, 2015.

PARK, W. et al. Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). **BMC Plant Biology**, v. 10, p. 142, 2010.

PEIXOTO SOBRINHO et al, Phytochemical screening and antibacterial activity of four *Cnidocolus* species (Euphorbiaceae) against standard strains and clinical isolates. **Journal of Medicinal Plants Research**. v. 6(21), p. 3742-3748, 2012

QIU L. et al. Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). **BMC Plant Biol.**, v. 10, p. 278, 2010.

QUIGLEY, F. et al. From genome to function: the *Arabidopsis* aquaporins. **Genome biology**, v. 3, n. 1, p. 1-17, 2001.

RAHMAN, A. Y. et al. Draft genome sequence of the rubber tree *Hevea brasiliensis*. **BMC Genomics**, v. 14, p. 75, 2013.

REDDY P.S. et al., Genome-wide identification and characterization of the aquaporin gene family in *Sorghum bicolor* (L.). **Plant Gene**, v. 1, p. 18–28, 2015.

RIVAROLA M. et al. Castor bean organelle genome sequencing and worldwide genetic diversity analysis. **PLoS One**, v. 6(7), e21743, 2011.

ROGSTAD, S.H. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. **Taxon**. v. 41, p. 701–708, 1992.

ROULIN A. et al. The fate of duplicated genes in a polyploid plant genome. *Plant J.*, v.73, p. 143–153, 2013.

ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. In: KRAWETZ, S.; MISENER, S. **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. Humana Press, p. 365- 386, 2000.



SAITOU, N.; NEI, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p. 406-425, 1987.

SAKURAI, J. et al. Identification of 33 rice aquaporin genes and analysis of their expression and function. **Plant and Cell Physiology**, v. 46, n. 1568-1577, 2005.

SAMPAIO, E. V. S.B. **Uso sustentável e conservação dos recursos florestais da Caatinga. Misistério do Meio Ambiente**. Brasília-DF: Serviço Florestal Brasileiro. p. 29-48. 2010.

SÁTIRO L.N., ROQUE N., A família Euphorbiaceae nas caatingas arenosas do médio rio São Francisco, BA, Brasil. **Acta bot. bras.**, v.22(1), p. 99-118. 2008.

SEKI, M. et al. Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold, and high-salinity stresses using a full-length microarray. **The Plant Journal**, v. 31, p. 279-292, 2002.

SHINOZAKI K, YAMAGUCHI-SHINOZAKI K, SEKI M. Regulatory network of gene expression in the drought and cold stress responses. **Curr Opin Plant Biol** 6:410–417, 2003.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. **Current Opinion of Plant Biology**, v. 3, n. 217-223, 2000.

SIPPL, M.J. Recognition of Errors in Three-Dimensional Structures of Proteins. **Proteins**, v. 17, p. 355-362, 1993.

SONAH, H. et al. Genome-wide identification and characterization of Xylanase Inhibitor Protein (XIP) genes in cereals. **Indian J. Genet. Plant Breed**, v. 76, p. 159–166, 2016.

SONG, L. et al. Soybean TIP gene family analysis and characterization of GmTIP1; 5 and GmTIP2; 5 in water transport. **Front. Plant Sci**, v. 7, p. 1564, 2016.

SOUZA, V.C., LORENZI, H. **Botânica Sistemática - Guia ilustrado para a identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APG II**. Nova Odessa, Plantaru, 2008.

TAMURA K. et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Mol Biol Evol.**, v. 30(12), p. 2725–2729, 2013.

THOMPSON JD, HIGGINS DG, GIBSON TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res.**, v. 22(22), p. 4673–4680, 1994

TUSKAN G.A. et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). **Science.**, v. 313(5793), p. 1596–1604, 2006.

VENKATESH, J.; YU, J. W.; PARK, S. W. Genome-wide analysis and expression profiling of the *Solanum tuberosum* aquaporins. **Plant Physiol. Biochem.** 73, 392–404, 2013.

WALLACE, I. S.; CHOI, W.-GYU; ROBERTS, D. M. The structure, function and regulation of the nodulin 26-like intrinsic protein family of plant aquaglyceroporins. **Society**, v. 1758, p. 1165 - 1175, 2006.

WALLACE, I. S.; ROBERTS, D. M. Homology Modeling of Representative Subfamilies of *Arabidopsis* Major Intrinsic Proteins. Classification Based on the Aromatic / Arginine Selectivity Filter 1 [ w ]. **Biochimica et Biophysica Acta (BBA) - Biomembranes**, v. 135, p. 1059-1068, 2004.

WANG, W.; VINOCUR, B.; ALTMAN, A. Plant responses to drought salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta**, v. 218, n. 1-14, 2003.

WANG, W.X. et al. Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. **Acta Horticulturae**, v. 560, p. 285-292, 2001.

WATERHOUSE A.M. et al. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, v. 25 (9), p. 1189-1191, 2009.

WEISING, K. et al. **DNA Fingerprinting in Plants. Principles, Methods and Applications**. 2 Edition. Taylor & Francis Group, p. 444, 2005.

WIEDERSTEIN, M.; SIPPL, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, v. 35, W407-W410, 2007

WU, P. et al. Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. **Plant J.**, v. 81, p. 810–821, 2015

WUDICK, M.M.; LUU, D.T.; MAUREL, C. A look inside: localization patterns and functions of intracellular plant aquaporins. **New Phytology**, v. 184, p. 289-302, 2009.

ZARDOYA, R. VILLALBA, S. A Phylogenetic Framework for the Aquaporin Family in Eukaryotes. **J Mol Evol.**, v. 52: p. 391–404, 2001.

ZHU, J.K. Salt and drought stress signal transduction in plants. **Annual Review of Plant Biology**, v. 53, p. 247-273, 2002.

ZOU, Z. et al. Genome-wide identification of rubber tree (*Hevea brasiliensis* Muell. Arg.) aquaporin genes and their response to ethephon stimulation in the laticifer, a rubber-producing tissue. **BMC Genomics**, v. 16, p.1001, 2015b.

ZOU, Z. et al. Gene structures, evolution, classification and expression profiles of the aquaporin gene family in castor bean (*Ricinus communis* L.). **PLoS One**, v.10, e0141022, 2015a.

ZOU, Z. et al. Genomewide identification of *Jatropha curcas* aquaporin genes and the comparative analysis provides insights into the gene family expansion and evolution in *Hevea brasiliensis*. **Front. Plant Sci.**, v. 7, p. 395, 2016.