UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUACÃO EM GENÉTICA

JOÃO PACIFICO BEZERRA NETO

CARACTERIZAÇÃO GENÉTICA E ANÁLISE EVOLUTIVA *IN SILICO* DE AQUAPORINAS EM FEIJÃO-CAUPI (*Vigna unguiculata* (L.) WALP.)

RECIFE - PE

2012

JOÃO PACIFICO BEZERRA NETO

CARACTERIZAÇÃO E ANÁLISE EVOLUTIVA IN SILICO DE AQUAPORINAS NO TRANSCRIPTOMA DO FEIJÃO-CAUPI (*Vigna unguiculata* (L.) WALP.)

Dissertação apresentada ao Programa de Pós-Graduação *Stricto sensu* do Centro de Ciências Biológicas da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Mestre em Genética.

Orientadora: Prof.ª Dr.ª Ana Maria Benko Iseppon

Co-orientadora: Dr.ª Valesca Pandolfi

Recife, PE

2012

Bezerra Neto, João Pacífico

Caracterização genética e análise evolutiva *In silico* de aquaporinas em feijão-caupi (*Vigna unguiculata* (L.) Walp.)/ João Pacífico Bezerra Neto– Recife: O Autor, 2012.

133 folhas : il., fig., tab.
Orientadora: Ana Maria Benko Iseppon
Coorientadora: Valesca Pandolfi
Dissertação (mestrado) – Universidade Federal de Pernambuco, Centro de Ciências Biológicas, Ciências Biológicas, 2012.
Inclui bibliografia

1. Feijão- caupi 2. Bioinformática 3. Proteínas I. Iseppon, Ana Maria Benko (orientadora) II. Pandolfi, Valesca (coorientadora) III. Título

583.74 CDD (22.ed.) UFPE/CCB- 2012- 194

JOÃO PACIFICO BEZERRA NETO

CARACTERIZAÇÃO E ANÁLISE EVOLUTIVA IN SILICO DE AQUAPORINAS NO TRANSCRIPTOMA DO FEIJÃO-CAUPI (*Vigna unguiculata* (L.) WALP.)

APROVADO EM ___/___/_____

BANCA EXAMINADORA:

_____	_____
1º Examinador	**1º Suplente**
Prof. Dr. Mauro Guida dos Santos	**Prof. Dr. Ana Christina Brasileiro Vidal**
UFPE	**UFPE**
_____	_____
2º Examinador	**2º Suplente**
Prof. Dr. Tercílio Calsa Jr.	**Prof. Dr. Paulo Paes de Andrade**
UFPE	**UFPE**
_____	_____
3º Examinador/Orientadora	**Co-orientadora**
Prof.ª Dr.ª Ana Maria Benko Iseppon	**Drª. Valesca Pandolfi**
UFPE	**UFPE**

Recife, PE

2011

À minha família, pedra fundamental da construção do que hoje sou e aos amigos que permaneceram ao meu lado.

Dedico.

AGRADECIMENTOS

Agradeço em primeiro lugar a **Deus**, que me faz ver além do agora, me faz confiar no que não conheço, confiar no que já está planejado para minha vida. Sem minha Fé, nada teria sentido.

Agradeço à minha mãe, **Ivonete Pacifico de Souza**, por ser para mim um exemplo de determinação e perseverança, por ter me ensinado que pelos filhos todo sacrifício é válido, por me apoiar em cada decisão difícil, sendo motivação para seguir em frente.

Agradeço também ao meu pai, **Adalberto Vanderlei de Souza Filho**, por ter feito o possível e impossível para permitir minha formação acadêmica e pessoal, por me mostrar que podemos ir além dos limites impostos.

Ao meu irmão, **Ivanberto Pacifico**, por ter me apoiado e incentivado nesta jornada, de um jeito estranho que ainda hoje não compreendo, mas valorizo.

Aos meus familiares, em especial à minha tia **Ivone Pacifico**, minha segunda mãe, por todo o carinho e confiança em mim depositados.

Agradeço enormemente à Profª Drª **Ana Maria Benko Iseppon**, por ter me orientado, por ter confiado em mim e no meu trabalho, além de ser um exemplo de pessoa e profissional.

Agradeço a Drª **Valesca Pandolfi** (Vale), por sua co-orientação, por ter surgido na hora certa em minha vida acadêmica, pela ajuda nos momentos mais complicados e principalmente por me fazer rir nos momentos certos.

Estendo meus agradecimentos à **Nina da Mota** e ao **Luis Belarmino**, meus grandes co-orientadores, desde minha entrada no laboratório até os dias de hoje, por terem me ajudado a desvendar os segredos da bioinformática.

Aos meus colegas do **LGBV**, que hoje são muitos, mas cada um com seu espaço na minha vida, especialmente o grupo da Bioinfo e seus agregados, **Hévila Mendes, Lidiane Amorim, Marx Lima, Pedro Lima, Rômulo Santos, Silvany Araújo, Ana Rafaela** que compartilham comigo grandes momentos, por transformarem meu dia-a-dia no laboratório mais animado e divertido.

Agradeço aos que trabalharam comigo, **Diego Guerra, Gabriela Rodrigues, Rafaela Cavalcanti**, vocês marcaram minha vida para sempre, aprendi com vocês que os vínculos não precisam ser só profissionais.

Agradeço a **Vanessa Souza** pelo auxílio na dissolução dos inúmeros problemas que teimavam em surgir, por sua amizade e orientações essenciais.

Aos meus mestres, em especial Dr. **Tercílio Calsa,** Drª **Neide Santos**, Dr. **Rodrigo Torres**, Dr. **Diego Ástua**, Dr. **Mauro Guida**, e Drª **Ana Christina**, por terem contribuído significativamente na minha formação.

Agradeço a **Suane Maria da Silva**, por me ceder um pouco de sua sabedoria, por sua amizade e por ser para mim um exemplo de pessoa e profissional.

Aos meus eternos companheiros de jornada do Curso de Ciências Biológicas, agradeço a todos, principalmente **Carolina Liberal, Morgana Xavier, Raquel Barbosa, Rodrigo César, Sarah Nigro**, por sua amizade durante e depois da graduação, pois caminhamos juntos, mesmo que não lado a lado. Agradeço à **Cassia Maria Rodrigues**, minha companhia nos cinemas e eventos da vida, alguém que merece minha confiança e gratidão.

Agradeço à **Waléria Kelly**, a minha louca predileta, a quem eu não escolhi, mas aceitei de todo meu coração.

Aos meus amigos irmãos do grupo **FORJ,** por terem entendido minha ausência, por me apoiar e incentivar e por serem "os melhores amigos que eu posso ter". Agradeço em especial**, Diego Henrique** (Animal)**, Wanessa Karla, Jéssica Correia, Willames César, Raquel Ester, Ilan Ramos, Renata Vieira, Ênia, Rogério Henrique,** os Moura **(July, João e Lidi), Gabriel Jacob, Rayanne Brito, Walter Pinto, Rosely e Roseany, Alice e Paulo,** por se tornarem muito mais que coadjuvantes em minha vida.

Agradeço o suporte financeiro oferecidos pelas agências de fomento à pesquisa, **CAPES** (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior), **CNPq** (Conselho Nacional de Desenvolvimento Científico e Tecnológico), **BNB** (Banco do Nordeste do Brasil), **FINEP** (Financiadora de Estudos e Projetos) e **FACEPE** (Fundação de Amparo à Ciência e Tecnologia de Pernambuco)**.**

E a todos que de alguma forma contribuíram para a realização e conclusão deste trabalho.

Muito obrigado.

“Aceitar-se como um ser humano cheio de limites e fraquezas é acima de tudo, sinal de equilíbrio, paz consigo mesmo e felicidade.”

(Pe. Fábio de Melo)

RESUMO

O feijão-caupi é uma das mais antigas leguminosas cultivadas, com destacada importância em áreas tropicais e subtropicais. No Brasil, é uma das principais fontes de nutrientes para a dieta humana e animal, representando 80% da produção de grãos nas regiões norte e nordeste. Entretanto, fatores como seca e salinidade (as duas principais adversidades ambientais em regiões semiáridas) têm prejudicado tanto a quantidade quanto a qualidade da produção do feijão-caupi. Neste sentido, o estudo de proteínas que agem como facilitadoras da passagem de água entre as membranas biológicas e na prevenção da passagem de íons e outros solutos na célula é fundamental para o entendimento de como algumas plantas respondem a tais adversidades. Entre estas proteínas as aquaporinas merecem atenção, tratando-se de uma família de pequenas proteínas transmembranas (24-30 kDa), dividida em quatro subfamílias, incluindo as Proteínas de Membrana Plasmática (PIP), Proteínas de Membrana de Tonoplasto (TIP), Proteínas de Membrana de Nódulos (NIP) e Pequenas Proteínas Básicas Intrínsecas de Membrana (SIP). Apesar da grande quantidade de dados sobre aquaporinas nos vegetais, nada se sabe sobre estes transportadores em algumas culturas de importância econômica, como o feijão-caupi. Utilizando dados de EST (*Expressed Sequence Tags*), 37 transcritos candidatos a aquaporinas foram identificados em feijão-caupi por meio de buscas com ferramentas de bioinformática, onde 27 apresentavam o domínio íntegro e 10 encontravam-se incompletas. Análises fenéticas das sequências de aminoácidos dividiram esta família nas quatro subfamílias já conhecidas para outras espécies vegetais, sendo 11 PIPs (cinco PIP1 e seis PIP2), 10 TIPs (quatro TIP3 e seis TIP), quatro NIPs e duas SIPs. Os alinhamentos múltiplos gerados a partir das sequências com domínio MIP completo, indicaram que as regiões ar/R das aquaporinas do feijão-caupi, em sua maioria, são similares às presentes em *Arabidopsis*, milho e arroz, apontando para afinidades com o mesmo tipo de soluto transportado. A expressão dos transcritos foi observada em todos os tecidos vegetais, estando associadas ao transporte de diferentes tipos de solutos nos compartimentos celulares. As aquaporinas identificadas neste estudo representam um importante recurso genético, fornecendo alvos em potencial para modificar as propriedades do uso de água em feijão-caupi ou em outras leguminosas relacionadas.

Palavras-chave: bioinformática; leguminosa; NordEST; Expressed Sequence Tag; Modelagem.

ABSTRACT

Cowpea is one of most ancient legumes cultivated by man, with outstanding importance in tropical and subtropical areas. In Brazil this legume is one of the main sources of nutrients for human and animal diet, representing 80% of grain production in the north and northeastern regions. However, factors like drought and salinity (the two main environmental adversities in semiarid regions) have affected both quantity and quality of cowpea production. In this sense, the analysis of proteins that act as water flux facilitators across biological membranes preventing the penetration of ions and other solutes in cells is critical to understand how some plants respond to such adversities. Among these proteins, the aquaporins deserve special attention, regarding a family of small transmembrane proteins (24-30 kDa), divided into four subfamilies, including the Plasma Intrinsic Proteins (PIP), Tonoplast Intrinsic Proteins (TIP), Nodulin26-like Intrinsic Proteins (NIP) and Small and Basic Intrinsic Proteins (SIP). Despite the large amount of data over plant aquaporins, no evaluation was carried out about these transporters in some economically important crops, such as cowpea. Using EST data (Expressed Sequence Tags), 37 aquaporin candidates have been identified in cowpea by searching with bioinformatic tools, where 27 sequences presented complete domains and 10 were incomplete. Phenetic analysis of amino acid sequences has grouped this family into four subfamilies known for other plant species, 11 PIPs (five PIP1 and six PIP2), 10 TIPs (four TIP3 and six other TIPs), four NIPs and two SIPs. The multiple alignments from complete sequences, indicated that most of the ar/R regions on cowpea aquaporins were similar to those found in *Arabidopsis*, maize and rice, indicating similarities concerning the type of solute transport. The transcript expression was observed in all plant tissues, associated with solutes transport in different cellular compartments. The aquaporins identified in this study represent important genetic resources, providing potential targets to modify the properties of water usage in cowpea or other related legumes.

Keywords: bioinformatics, legume, NordEST; Expressed Sequence Tag; modelling.

LISTA DE ILUSTRAÇÕES

Página

Figura 1. Esquema geral da resposta vegetal ao estresse abiótico, em especial seca e salinidade. As células vegetais recebem os vários sinais através de uma variedade de sensores (muitos ainda desconhecidos), os sinais são traduzidos para as diferentes vias por meio de hormônios vegetais, transdutores de sinal e reguladores de transcrição. Genes estresse induzidos são regulados por diversos sinais, e muitos regulados por fatores de transcrição (TFs) induzidos por estresse (uma cascata transcricional). Muitos destes genes codificam para proteínas diretamente envolvidas na tolerância ao estresse (Fonte: Benko-Iseppon et al., 2012). **19**

Figura 2. Via genérica da resposta vegetal frente ao estresse hídrico, salino e baixas temperaturas, evidenciando a comunicação cruzada entre as diferentes vias de sinalização. (Adaptado de Mahajan e Tuteja, 2005). **21**

Figura 3. Modelo do transporte de moléculas de água através das aquaporinas nas membranas celulares. (Adaptada de Zhao et al., 2008). **23**

Figura 4. Representação da estrutura molecular da aquaporina de espinafre (*Spinacia oleracea*, Chenopodiaceae). A. Estrutura da conformação aberta da SoPIP2;1 [Protein Data Bank ID: 36623], arranjo tetramérico típico; B e C. Estrutura do monômero, conformação lateral e superior, em amarelo evidenciados os dois motivos NPA (Asn-Pro-Ala). **28**

Figura 5. Modelo da estrutura de uma aquaporina, evidenciando as principais características estruturais. Alfa-hélices representadas pelos retângulos azuis. Nela estão presentes os seis domínios transmembrana (H1-H6), conectados pelos cinco loops (LA-LE). Dois domínios alfa-hélices nos loops B e E, contendo os dois motivos NPA, formadores do poro (esquema do autor). **29**

Figura 6. Diagrama esquemático da localização das diferentes famílias de aquaporinas nas organelas vegetais. PIPs: localizadas na membrana **31**

plasmática; TIPs: localizadas no vacúolo; NIPs: localizadas no retículo endoplasmático e membrana plasmática; SIPs: Localizada no retículo endoplasmático. Adicionalmente ao transporte de água, transportam: PIP (CO2 e ureia), TIP (ureia, amônia, glicerol e H2O2), NIP (amônia, ureia, glicerol, boro e lactato). Adaptado de Maeshima e Ishikawa (2008).

Figura 7. Funções integradas das aquaporinas vegetais. Fatores endógenos e ambientais atuando no organismo, ativando a funcionalidade das aquaporinas nos diferentes órgãos e tecidos. Adaptado de Maurel et al. (2008). **32**

Figura 8. Representação esquemática da construção da biblioteca de SAGE. 1. Preparação do mRNA; 2. Síntese de cDNA; 3. Clivagem do cDNA "biotinilado"; 4. Ligação do cDNA "biotinilado" às esferas magnéticas; 5. Ligação dos adaptadores ao cDNA (to bound cDNA); 6. Liberação das tags de cDNA utilizando a enzima "Tagging"; 7. Liberação das extremidades cegas das tags (blunt ending release); 8. Ligação de tags para formar ditags; 9. Amplificação das ditags via PCR; 10. Isolamento das ditags; 11. Purificação das ditags; 12. Ligação das ditags para formação do concatâmero; 13. Clonagem e sequenciamento dos concatâmeros. Adaptado de Velculescu et al. (1999). **41**

Figura 9. Fluxograma da análise de dados de SuperSAGE, visando à descoberta de genes-candidatos a partir de tags de 26 pb. Tags diferencialmente expressas são selecionadas e seu perfil de expressão é analisado. Um subgrupo de Tags, as quais apresentavam um perfil de expressão diferencial significativo (reposta rápida e/ou expressão tecido-específica) foi utilizado na descoberta de genes de interesse por metodologias específicas (ex. RACE-PCR), sendo sua expressão confirmada por RT-PCR. **42**

Figura 10. A bioinformática como uma ferramenta integradora da informação gerada por diversas ciências. Fonte: Malone et al. (2006). **45**

Figura 11. Diagrama do processo geral de prospecção, caracterização e modelagem de aquaporinas no genoma expresso de feijão-caupi via ferramentas de bioinformática. **58**

Figura 12. Análise fenética de membros da família de aquaporinas de *Vigna unguiculata* pelo método de Neighbor Joining, com valores de bootstrap ao nível de 1000 replicações.O dendrograma evidencia as 27 sequências de domínio completo e seu agrupamento em quatro subfamílias com relação à similaridade de suas sequências. Escala de 0,1 refere-se a 10% de divergência entre as sequências. Legenda: PIP (Proteínas de Membrana Plasmática), TIP (Proteínas de Membrana de Tonoplasto), NIP (Proteínas de Membrana de Nódulos) e SIP (Pequenas Proteínas Básicas Intrínsecas de Membrana). **70**

Figura 13. Análise fenética para as VuPIPs de Vigna e PIPs de *Arabidopsis*. Divisão em dois grupos distintos. Escala de 0,02 refere-se a 2% de divergência entre as sequências. Legenda: PIP (Proteínas de Membrana Plasmática). **71**

Figura 14. Análise fenética para as VuTIPs de Vigna e TIPs de Arabidopsis. Divisão em cinco grupos distintos, onde a escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: TIP (Proteínas de Membrana de Tonoplasto). **72**

Figura 15. Análise fenética para as VuNIPs de Vigna e NIPs de *Arabidopsis*. Subfamília dividida em sete grupos segundo classificação de *Arabidopsis*. Escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: NIP (Proteínas de Membrana de Nódulos). **73**

Figura 16. Análise fenética para as VuSIPs de Vigna e SIPs de *Arabidopsis*. Divisão em dois grupos distintos: VuSIP1 e VuSIP2, onde a escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: SIP (Pequenas Proteínas Básicas Intrínsecas de Membrana). **74**

Figura 17. Relações fenéticas entre aquaporinas vegetais. Árvore construída pelo método de Neighbour Joining, mostrando a comparação entre as 27 aquaporinas completas de *V. unguiculata* em conjunto com aquaporinas de milho (*Zea mays*), *Arabidopsis thaliana*, *Ricinus communis* e uva (*Vitis vinifera*). As quatro subfamílias encontradas e feijão-caupi estão indicadas pelas diferentes cores. A barra indica distância de 0,05 indica 5% de alteração na sequência de aminoácidos. **75**

Figura 18. Perfil de expressão obtido após clusterização hierárquica realizada pelo programa CLUSTER 3.0 dos transcritos de aquaporinas de feijão-caupi. Os quadrados em laranja indicam alta expressão, em laranja mais claro indicam baixa expressão e em amarelo, ausência de expressão. Abreviações: Mi1: mix de tecidos; Lfm: folha e meristema apical; Dsd: sementes em desenvolvimento; Rt3: raiz; Lf3: folha; Bud: botão floral; RtS: raízes sob salinidade. **78**

Figura 19. Modelos tridimensionais das aquaporinas de feijão-caupi, obtidas via modelagem comparativa, a partir do modelo SoPIP2-1 disponível no PDB. **82-88**

Figura 20. Gráficos representativos do Z-score dos modelos gerados, comparativamente aos modelos de cristalografia de raios-x (em azul claro) e ressonância magnética (em azul escuro) depositados no PDB. Os pontos em preto referem-se aos modelos gerados para as subfamílias PIP (A), TIP (B), NIP (C) e SIP (D); os pontos em vermelho referem-se ao Z-score do modelo de aquaporina de espinafre (SoPIP2-1). **89**

Figura 21. Visualização da similaridade dos fragmentos isolados do cDNA de *Vigna unguiculata* e seus respectivos primers contra os transcritos caracterizados como aquaporinas no banco de dados NordEST utilizando o resultado do BLASTn no programa Circoletto. Regiões coloridas representam os níveis de similaridade baseando-se nos scores: Azul – score menor que 100; Verde – score entre 100 e 170; Laranja – score entre 170 e 240; Vermelho – score superior a 240. **91**

LISTA DE TABELAS

Páginas

Tabela 1: Conjunto de primers selecionados para a validação das aquaporinas de feijão-caupi. Fw: foward; Rv: reverse. **60**

Tabela 2. Categorias de aquaporinas identificadas no transcriptoma do feijão-caupi. **65**

Tabela 3. Sumário das principais regiões e sítios conservados de aquaporinas em *Vigna*. **67**

Tabela 4. Padrões de Filtro de seletividade ar/R de feijão-caupi. **68**

Tabela 5. Distribuição das *reads* nas diferentes bibliotecas de feijão-caupi. Mi1: mix de tecidos; Lfm: folha e meristema apical; Dsd: sementes em desenvolvimento; Rt3: raiz; Lf3: folha; Bud: botão floral; RtS: raízes sob salinidade. **77**

Tabela 6. Valores de similaridade entre o modelo de espinafre (SoPIP2-1) e os diferentes membros de aquaporinas de feijão-caupi. **79**

Tabela 7: Valores de qualidade e energia para os modelos estruturais de aquaporinas de feijão-caupi obitidos via plataforma SWISS-MODEL e ferramenta Procheck. **81**

Tabela 8: Valores de similaridade dos fragmentos amplificados do cDNA de feijão-caupi contra os bancos NordEST e NCBI, seus respectivos primers e tamanhos da sequência. **90**

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

Sigla	Definição
aa	Aminoácidos
AFLP	*Amplified Fragment Length Polymorphism;* Polimorfismo de restrição por amplificação do fragmento
ar/R	Região arginina/aromática
BLAST	*Basic Local Alignment Search Tool* Ferramenta de Busca por Alinhamento Local
CD	*Conserved Domain* Domínio Conservado
cDNA	DNA complementar
cDNA-AFLP	Polimorfismo de restrição por amplificação do fragmento baseado em cDNA
CGKB	*Cowpea Genomics Knowledge Base* Base de Conhecimentos Genômicos de Feijão-Caupi
DD	*Differential display* Display diferencial
DDBJ	*DNA Database of Japan* Banco de Dados de DNA do Japão
DNA	Ácido desoxirribonucleico
EMBL	*European Molecular Biology Laboratory* Laboratório Europeu de Biologia Molecular
EST	*Expressed Sequence Tag* Sequência de etiquetas expressa
FT	Fator de Transcrição
GenBank	Banco de Genes - banco de dados norte-americano
GIP	Homóloga aos canais de glicerol bacteriano

GSS	*Cowpea Genespace Sequences* Sequências de Feijão-Caupi no Genespace
HARVEST	*HARVEST Database* Banco de dados HARVEST
HIP	Proteínas híbridas intrínsecas de membrana
INSD	*International Nucleotide Sequence Database* Banco de Dados Internacional de Sequências de Nucleotídeos
KEGG	*Kyoto Encyclopedia of Genes and Genomes* Enciclopédia de Genes e Genomas de Kyoto
LongSAGE	Variante da SAGE
MEGA	*Molecular Evolutionary Genetic Analysis* Análises Genéticas de Evolução Molecular
MIP	*Major Intrinsic Protein* Proteínas Intrínsecas de Membrana
MPSS	*Massively Parallel Signature Sequencing* Sequenciamento massivo em paralelo de assinaturas
mRNA	RNA mensageiro
NCBI	*National Center for Biotechnology Information* Centro Nacional para Informação Biotecnológica
NIP	*Nodulin26-like Intrinsic Proteins* Proteínas de membrana de nódulos
NJ	*Neighbor-Joining* Agrupamento por Vizinhança
NordEST	Banco de dados de EST para o feijão-caupi da Rede Nordeste de Biotecnologia
ORF	*Open Reading Frame* Quadro Aberto de Leitura
ORF-Finder	*Open Reading Frame Finder* Identificador de Quadros Abertos de Leitura
Pb	Pares de bases em nucleotídeos

PCR	*Polymerase Chain Reaction* Reação em cadeia da polimerase
PDB	*Protein Data Bank* Banco de Dados de Proteínas
Pfam	*Protein families database of alignments and HMMs* Banco de dados de famílias proteicas de alinhamentos e HMMs
PI	Ponto isoelétrico
PIP	*Plasma Membrane Intrinsic Protein* Proteínas de membrana plasmática
PIR	*Protein Information Resource* Recursos de Informações Proteicas
PRINTS	Banco de Dados de Assinaturas, Motivos e Domínios Proteicos
PROSITE	Banco de Dados de Domínios Proteicos
PubMed	Banco de Dados da Biblioteca Nacional de Medicina
ReNorBio	Rede Nordeste de Biotecnologia
RNA	Ácido ribonucleico
RT-qPCR	*Real Time quantitative PCR* PCR quantitativo em tempo real
SAGE	*Serial Analysis of Gene Expression* Análise serial de genesn expressos
SIP	*Small Basic Intrinsic Proteins* Pequenas proteínas básicas intrínsecas de membranas
SSH	*Supression Subtractive Hybridization* Biblioteca subtrativa suprimida
SuperSAGE	Variante da SAGE
Tag	Etiquetas de sequências
TIP	*Tonoplast Intrinsic Proteins* Proteínas de membrana de tonoplasto

TMH	*Transmembrane hélices* Hélices transmembrana
UPGMA	*Unweighted Pair Group Method with Arithmetic Means* Método não polarizado de Agrupamentos aos Pares com Médias Aritméticas
XIP	*Uncharacterized X Intrinsic Protein* Proteína intrínseca desconhecida X

SUMÁRIO

1. INTRODUÇÃO 12

2. REVISÃO DA LITERATURA 15

2.1 A IMPORTÂNCIA DA ÁGUA NAS PLANTAS 15

2.2 PLANTAS SOB ESTRESSE 17

2.3 ESTRESSES ABIÓTICOS E VIAS DE SINALIZAÇÃO EM PLANTAS 18

2.4 PLANTAS SOB DÉFICIT HÍDRICO E SALINIDADE 19

2.4.1 Déficit hídrico 19

2.4.2 Estresse salino 21

2.5 AQUAPORINAS: DIVERSIDADE, CLASSIFICAÇÃO E ATIVIDADE 22

2.5.1 Proteínas Intrínsecas de Membrana Plasmática – PIP 25

2.5.2 Proteínas Intrínsecas de Tonoplasto – TIP 26

2.5.3 Proteínas Intrínsecas de Nódulos – NIP 27

2.5.4 Pequenas Proteínas Básicas Intrínsecas de Membrana – SIP 27

2.6 ESTRUTURA E LOCALIZAÇÃO DAS AQUAPORINAS 28

2.7 FUNÇÃO FISIOLÓGICA E EXPRESSÃO 31

2.7.1 Aquaporinas e respostas a estresses 35

2.8 IMPORTÂNCIA DAS "ÔMICAS" NA PROSPECÇÃO E ANÁLISE DE GENES 36

2.8.1 TRANSCRIPTÔMICA 37

2.8.1.1 Análise serial da expressão gênica (SAGE) e seus derivados 39

2.9 BIOINFORMÁTICA: BANCOS DE DADOS E FERRAMENTAS 44

2.10 FEIJÃO-CAUPI 49

2.10.1 CLASSIFICAÇÃO, ORIGEM E DISTRIBUIÇÃO 49

2.10.2 IMPORTÂNCIA ECONÔMICA 50

2.10.3 MELHORAMENTO GENÉTICO DO FEIJÃO-CAUPI 51

3. OBJETIVOS 54

3.1 Objetivo Geral 54

3.2 Objetivos Específicos 54

4. MATERIAIS E MÉTODOS 55

4.1 Identificação e Análise das Aquaporina de Feijão-caupi 55

4.2 Modelagem Comparativa e Estrutura Tridimensional 57

4.3 Validação dos Transcritos por Clonagem e Sequenciamento 58

4.3.1 Desenho de Primers 58

4.3.2 Obtenção do cDNA e clonagem das sequências relacionadas a aquaporinas 59

4.3.2.1 Obtenção do material vegetal 59

4.3.2.2 Extração do RNA total e síntese do cDNA 61

4.3.2.3 Obtenção dos produtos da PCR 61

4.3.2.4. Clonagem e Sequenciamento 61

4.3.3 Análise das sequências 63

5. RESULTADOS 64

5.1 Avaliação dos Resíduos Conservados 66

5.2 Análise Comparativa e Filogenia 69

5.3 Expressão das Aquaporinas de Feijão-caupi 75

5.4 Modelagem Comparativa 78

5.5. Validação dos transcritos 88

6. DISCUSSÃO 92

6.1 Avaliação dos Resíduos Conservados 93

6.2 Análise Comparativa e Filogenia 95

6.3 Expressão e Localização Subcelular 96

6.4 Modelagem Comparativa 99

6.5 Validação dos transcritos 100

7. CONCLUSÕES 102

8. REFERÊNCIAS BIBLIOGRÁFICAS 104

1. INTRODUÇÃO

O feijão-caupi é uma das leguminosas mais antigas utilizadas pelo homem, sendo cultivado desde regiões tropicais, subtropicais de baixa altitude até altitudes superiores a 1.600 m (EHLERS; HALL, 1997). Esta cultura apresenta importância vital na Ásia, África e América (MUCHERO et al., 2009), com mais de 8,5 milhões de ha cultivados e produção anual superior a três milhões de toneladas, sendo considerada grande fonte de proteínas, carboidratos, cálcio, ferro e vitaminas na alimentação humana (PHILLIPS et al., 2003). Em regiões semi-áridas, o feijão-caupi representa uma importante forrageira e alimentícia, destacando-se como um produto valioso para a produção e o comércio de grãos (TIMKO et al., 2007).

Entretanto, nos países em desenvolvimento, o feijão-caupi tem recebido pouca atenção, principalmente com relação a sua importância econômica e social. Porém, avanços no melhoramento desta cultura vêm surgindo com o desenvolvimento das várias ômicas, juntamente com o conhecimento adquirido nestas áreas com espécies modelos de leguminosas (SINGH, 2005; TIMKO et al., 2007).

Dentre os estresses ambientais que afetam drasticamente a produtividade do feijão-caupi, o déficit hídrico destaca-se como um dos mais importantes, principalmente por afetar todos os aspectos relacionados ao desenvolvimento e crescimento vegetal, incluindo mudanças bioquímicas e fisiológicas. A água, elemento indispensável para as células vivas, compreende o solvente universal presente em processos como a respiração celular e fotossíntese, fundamental para sobrevivência e desenvolvimento vegetal (CHAUMONT et al., 2005). Para facilitar o transporte de água e de outras pequenas moléculas polares através das membranas hidrofóbicas, os organismos vivos dispõem de uma variedade de canais proteicos integrantes de membranas. As aquaporinas ou *Major Intrinsic Proteins* (MIPs – Proteínas Intrínsecas de Membrana) formam uma grande e evolutivamente conservada superfamília de canais de membrana, encontradas em todos os tipos de organismos, incluindo eubactéria, archebactérias, fungos, animais e plantas (AGRE; KOZONO, 2003; HEYMANN; ENGEL, 1999).

Estas proteínas apresentam funções chave na permeabilidade à água através das membranas em uma variedade de processos fisiológicos, como a captação de

água do solo, transporte célula-célula e raiz-caule, homeostase celular e controle do gradiente osmótico (TYEMAN et al., 2002; MAUREL et al., 2008; UEHLEIN; KALDENHOFF, 2008; HEINEN et al., 2009). Estudos recentes também relatam membros de aquaporinas com funções importantes em outros processos, como a absorção e translocação de nutrientes, remobilização de nitrogênio e transdução de sinais (KALDENHOFF; FISCHER, 2006; BIENERT et al., 2008; GOMES et al., 2009; SOTO et al., 2010).

As aquaporinas vegetais formam uma grande e divergente superfamília de proteínas com mais isoformas do que em outros organismos. Até o momento, foram identificados mais de 35 membros da superfamília MIP (*Major Intrinsic Proteins*) em *Arabidopsis thaliana* (JOHANSON et al., 2001), 31 em *Zea mays* (CHAUMONT et al., 2001), 33 em *Oryza sativa* (SAKURAI et al., 2005). Mais recentemente, 23 isoformas foram descritas na briófita *Physcomitrella patens* (DANIELSON; JOHANSON, 2008), 55 na arbórea *Populus tricocharpa* (GUPTA; SANKARARAMAKRISHNAN, 2009) e 71 em algodão (PARK et al., 2010). O número de diferentes isoformas de aquaporinas nos organismos vegetais implica na importância que estas possuem em eventos como o rápido crescimento celular e adaptação à seca por meio da regulação da regulação do transporte de água (ALEXANDERSSON et al., 2005).

Baseando-se na similaridade de suas sequências, esta superfamília é dividida em cinco subfamílias evolucionariamente distintas: (i) PIP, proteínas de membrana plasmática (*Plasma Membrane Intrinsic Proteins*); (ii) TIP, Proteínas de membrana de tonoplasto (*Tonoplast Intrinsic Proteins*); (iii) NIP, Proteínas de membrana de nódulos (*Nodulin26-like Intrinsic Proteins*); (iv) SIP, Pequenas proteínas básicas intrínsecas de membranas (*Small Basic Intrinsic Proteins*) e a recentemente descoberta, XIP, uma proteína intrínseca de membrana não caracterizada, denominada de X (CHAUMONT et al., 2001; JOHANSON et al., 2001; DANIELSON;JOHANSON, 2008; BIENERT et al., 2011).

Estruturalmente esses peptídeos apresentam uma estrutura quaternária tetramérica, onde cada monômero consiste em seis hélices transmembranas (H1-H6) conectados por cinco "*loops"* (A-E) e duas hemi-hélices transmembranas nos *loops* B e E, cada uma contendo os motivos conservados NPA (asparagina-prolina-alanina), responsáveis pela formação da constrição central do poro (CHAUMONT et al., 2005; KRUSE et al., 2006). A segunda constrição de seletividade das

aquaporinas, a região arginina/aromática (ar/R) é formada por quatro resíduos de aminoácidos (R1-R4) que contribuem para o tamanho da barreira de exclusão e pontes de hidrogênio necessárias para o transporte efetivo de substrato (MURATA et al., 2000; LUDEWIG; DYNOWSKI, 2009). Tais resíduos são localizados na hélice transmembrana 2 (H2), hélice transmembrana 5 (H5) e *loop* E (LE1 e LE2) (BIENERT et al., 2011), sendo determinantes para a especificidade de transporte para diferentes moléculas como amônia (LIU et al., 2003; JAHN et al., 2004; LOQUE et al., 2005), ácido lático (CHOI; ROBERTS, 2007), B (boro) (TAKANO et al., 2006), silício (MA et al., 2006), CO_2 (F UEHLEIN et al., 2003; LEXAS et al., 2006), entre outros.

Apesar da sua importância no transporte de água tanto em nível celular quanto fisiológico, o entendimento por completo das funções fisiológicas e mecanismos regulatórios das aquaporinas é de difícil elucidação (HACHEZ et al., 2006), dada sua vasta diversidade na especificidade de substrato, localização, regulação transcricional (em resposta aos estresses) e pós-traducional (MAUREL, 2007). Além disso, algumas aquaporinas são expressas constitutivamente, enquanto outras são controladas por determinadas condições ambientais e possuem especificidade tecidual (ALEXANDERSSON et al., 2005; BROUSIAC et al., 2005).

Para o completo entendimento das relações hídricas e sua associação ao transporte de pequenas moléculas polares em nível molecular faz-se necessário não somente identificar todo o conjunto de aquaporinas, mas também entender como funcionam seus aspectos de especificidade, bem como traçar o perfil de expressão dessas proteínas (DANIELSON; JOHANSON, 2008). Neste sentido, a utilização de métodos *in silico,* utilizando bancos de dados de genoma completo ou bancos de sequências expressas (ESTs - *Expressed Sequence Tags*), tem sido uma estratégia bastante eficiente tanto na identificação de genes codificantes para aquaporinas, bem como na determinação do conjunto total de aquaporinas de um dado organismo (CHAUMONT et al., 2001; PARK et al., 2010). Assim, neste estudo objetivou-se à identificação e à caracterização de genes codificadores para aquaporinas no transcriptoma do feijão-caupi, com vistas à catalogação e caracterização de elementos úteis para programas de melhoramento contra seca e salinidade desta importante cultura.

2. REVISÃO DA LITERATURA

2.1 A IMPORTÂNCIA DA ÁGUA NAS PLANTAS

A água é uma das mais importantes substâncias do nosso planeta, tratando-se do meio onde a vida evoluiu primordialmente. As plantas adaptaram-se às mais diversas condições terrestres, desenvolvendo mecanismos morfofisiológicos que possibilitaram essa adaptação, até mesmo em ambientes com baixas condições hídricas (NEPOMUCENO et al., 2001). Essas adaptações incluíram raízes, que permitiam a absorção de água e nutrientes a partir de grandes volumes de solo; um sistema vascular, que permitia um rápido transporte da água e de produtos da fotossíntese; e uma cutícula bem desenvolvida, com estômatos, que permitia a entrada de dióxido de carbono, mas capaz de controlar a perda de água dos tecidos (KRAMER; BOYER, 1995).

A disponibilidade de água impõe forte pressão seletiva aos organismos vegetais, uma vez que a mesma é fundamental em todos os aspectos fisiológicos (BOHNERT, et al., 1995; BRAY, 1997). É componente de muitos processos fisiológicos, incluindo a fotossíntese, a hidrólise do amido, além do seu papel na turgescência celular e, consequentemente, no crescimento vegetal (SILVA; FREITAS, 1998). A distribuição e abundância das plantas em sistemas naturais e agrícolas são incisivamente determinadas pela disponibilidade da água (MCKAY, et al., 2003). Estudar as relações entre a água, o solo, as plantas e a atmosfera, torna-se importante devido à diversidade de funções bioquímicas, fisiológicas e ecológicas que a mesma exerce sobre o meio e sobre as plantas.

Para extrair os nutrientes do solo necessários ao seu desenvolvimento, as plantas precisam de certa quantidade de água, a qual serve como veículo de transporte destes e de fotoassimilados. A entrada de água na planta ocorre por meio da absorção em toda a sua superfície, sendo que a maior parte do suprimento vem do solo (PAIVA; OLIVEIRA, 2006). A quantidade de água absorvida pelo sistema radicular depende da quantidade de água do solo disponível para a planta, do arejamento, da temperatura do solo, da concentração da solução e da taxa de transpiração (KERBAUY, 2004). A água disponível refere-se à quantidade de água

que as plantas podem reter ou armazenar entre a capacidade de campo e o ponto de murcha (BERNARDO et al., 2006). A capacidade de campo corresponde ao conteúdo de água em um solo depois de ter sido saturado com água, cujo excesso foi drenado (TAIZ; ZEIGER, 2004). Já o ponto de murcha corresponde ao momento em que a planta não consegue mais retirar água do solo (BERNARDO et al., 2006).

Entre os recursos que a planta necessita para crescer e se desenvolver, a água é o mais abundante e também o mais limitante (KERBAUY, 2004). Embora amplamente dispersa por toda a planta, a maior parte da água é perdida para a atmosfera na forma de vapor, via transpiração. Através da transpiração, a água atua tanto na refrigeração das folhas, como também participa ativamente do metabolismo vegetal (KLAR, 1991).

A entrada da água na planta através da superfície radicular ocorre normalmente através de processos osmóticos (não envolvendo, portanto, gasto de energia), ou seja, é conduzida do meio menos concentrado para o meio mais concentrado em solutos, sendo importante para o balanço hídrico – a diferença entre a água absorvida e a água perdida num determinado espaço de tempo (RAVEN et al., 2007). Entre os principais processos envolvidos no balanço hídrico de uma planta estão: a absorção, a condução e a perda de água. Assim, para que o balanço hídrico seja mantido a níveis razoáveis, é necessário que as taxas relativas a estes três processos se ajustem (KERBAUY, 2004; TAIZ; ZEIGER, 2004).

Na tentativa de evitar o desbalanço hídrico, algumas plantas otimizam a absorção de água; outras reduzem a perda de água aumentando a resistência à difusão, reduzindo a taxa de transpiração (o que garante maior armazenamento de água nos tecidos), sendo que muitas regulam o potencial osmótico por meio de canais água-específicos localizados nas membranas celulares. Todos estes aspectos podem alterar a morfologia das plantas na tentativa de evitar tanto a dessecação, bem como a indução da expressão de genes associados à tolerância à dessecação (NEPOMUCENO et al., 2001; PAIVA; OLIVEIRA, 2006).

2.2 PLANTAS SOB ESTRESSE

As plantas são frequentemente expostas a diferentes condições de estresses bióticos ou abióticos, tais como: temperaturas extremas, seca, salinidade, estresse oxidativo, ataque de insetos, de herbívoros e de patógenos. Estes fatores afetam o crescimento da planta, impedindo que alcance todo seu potencial genético e - numa situação extrema e irreversível - resultem na morte do vegetal (MAHAJAN; TUTEJA, 2005).

Como forma de adaptação, as plantas desenvolveram, ao longo da evolução, mecanismos para lidar e se adaptar a estresses impostos por ambientes, sem as quais sua ocorrência e sobrevivência seria ameaçada nestes novos ambientes. Juntos, os estresses abióticos são responsáveis por desencadear uma série de respostas das plantas, que podem ser percebidas através das modificações morfológicas, fisiológicas, moleculares e metabólicas, a fim de tolerar estes estresses, embora o custo metabólico e energético destes mecanismos possa, em alguns casos, resultar em baixa produtividade (WANG et al., 2001; NOGUEIRA et al., 2005). Quando tais estresses abióticos ocorrem, várias respostas bioquímicas e fisiológicas são induzidas nas plantas, de forma a propiciar a tolerância ou aumentar as chances de sobrevivência às condições adversas (BENKO-ISEPPON et al., 2012; *in press*).

Estudos apontam estresses abióticos como os fatores mais prejudiciais com relação ao crescimento vegetal e a produtividade agrícola em nível mundial (BROWN et al., 2003; MITTLER, 2006). Perdas devido a estes estresses levam a queda do rendimento médio das colheitas em mais de 50%, representando assim um grande desafio diante de uma população mundial crescente (BRAY, et al., 2000; MAHAJAN; TUTEJA, 2005). Dentre os estresses abióticos críticos para a agricultura, o déficit hídrico e a salinidade destacam-se como os principais problemas ambientais que limitam a sobrevivência, crescimento e produtividade das plantas.

2.3 ESTRESSES ABIÓTICOS E VIAS DE SINALIZAÇÃO EM PLANTAS

Os diferentes estresses abióticos quase sempre induzem a ativação de genes de vias de sinalização celular (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2000; KNIGHT; KNIGHT, 2001; ZHU, 2001, 2002) e respostas celulares semelhantes como a produção de proteínas de estresse e a superexpressão de substâncias antioxidantes (CUSHMAN; BOHNERT, 2000). De uma forma geral, os estresses abióticos iniciam com a percepção do estresse, desencadeando uma cascata de eventos moleculares, levando à transmissão e o processamento de sinais para geração de respostas mediadas por fitormônios, mensageiros secundários, fatores de transcrição (FTs), genes e ou proteínas sinalizadoras (Figura 1) (BRAY, 1993; PANDEY, 2008).

A percepção e a transdução de sinais são passos importantes na determinação da sobrevivência e reprodução de plantas frente a diferentes ambientes (CHINNUSAMY et al., 2004). Dentre os mecanismos de proteção estão o controle redox, a remoção de espécies de oxigênio reativo, a acumulação de metabólitos e a homeostase iônica (BOHNERT, et al., 2001).

A regulação do potencial osmótico e a compartimentalização de íons ocorre à custa do gradiente eletroquímico de H^+ e do controle integrado de diferentes ATPases e de outros transportadores associados com membranas celulares (BRAY, 1993). Alguns desses transportadores são membros das aquaporinas, canais específicos para água, íons e outros pequenos solutos. À medida que as proteínas canal acumulam-se no tonoplasto (membrana do vacúolo) durante o estresse, o movimento da água e dos solutos do vacúolo para o citoplasma é promovido alterando tanto o teor de água quanto o potencial osmótico do citoplasma (NEPOMUCENO et al., 2001).

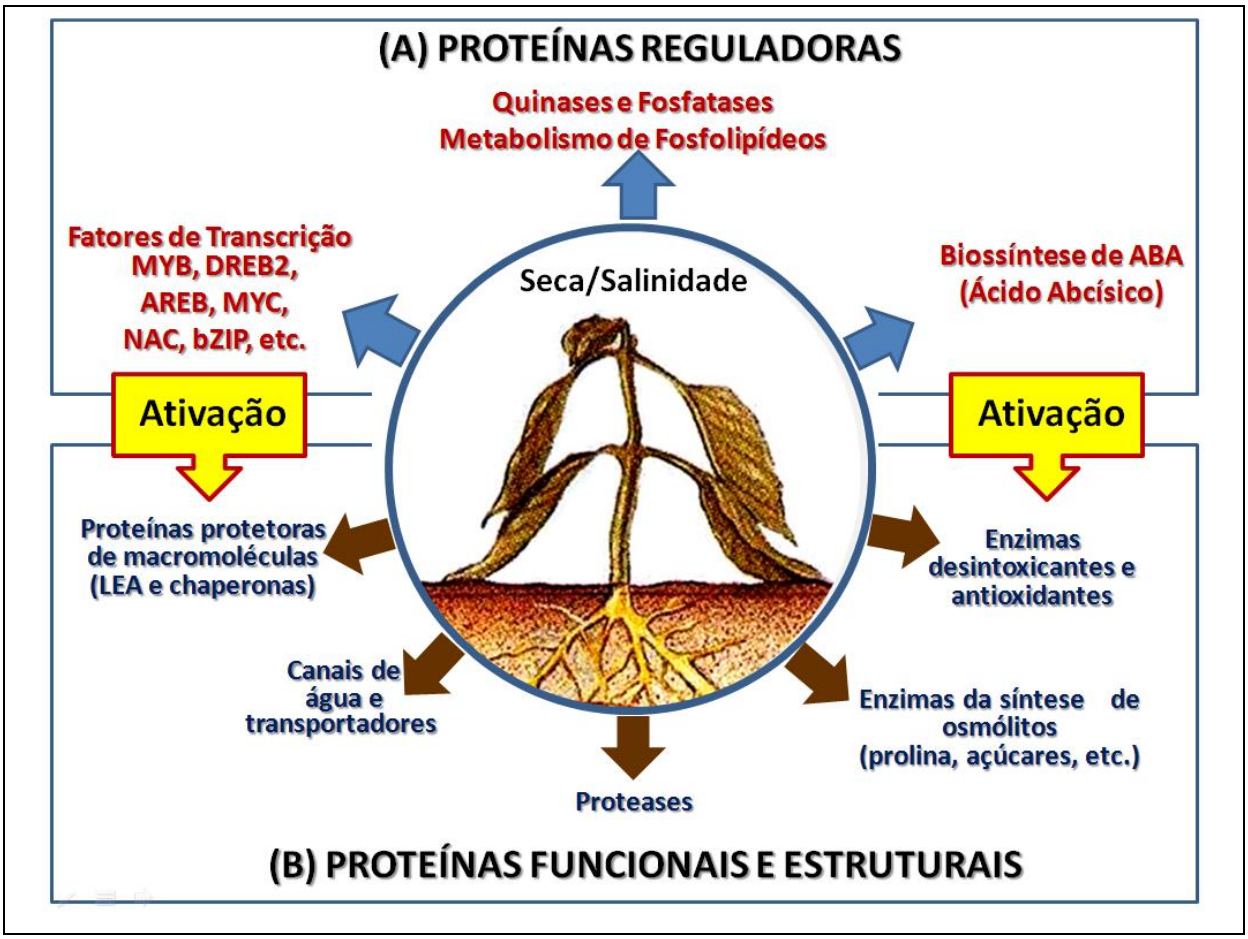


Figura 1. Esquema geral da resposta vegetal ao estresse abiótico, em especial seca e salinidade. As células vegetais recebem os vários sinais através de uma variedade de sensores (muitos ainda desconhecidos), os sinais são traduzidos para as diferentes vias por meio de hormônios vegetais, transdutores de sinal e reguladores de transcrição. Genes estresse induzidos são regulados por diversos sinais, e muitos regulados por fatores de transcrição (TFs) induzidos por estresse (uma cascata transcricional). Muitos destes genes codificam para proteínas diretamente envolvidas na tolerância ao estresse (Fonte: Benko-Iseppon et al., 2012).

2.4 PLANTAS SOB DÉFICIT HÍDRICO E SALINIDADE

2.4.1. Déficit hídrico

O estresse causado pela deficiência hídrica em plantas inicia um complexo de respostas, começando com a percepção do estresse, o qual desencadeia uma cascata de sinais moleculares induzindo vários níveis de respostas bioquímicas, fisiológicas e de desenvolvimento (BRAY, 1993). De uma forma geral, tais modificações visam manter o crescimento e a reprodução da planta em ambientes onde há pouca disponibilidade de água.

A maioria dessas mudanças envolve um aumento na resistência difusiva ao vapor de água, mediante o fechamento dos estômatos, reduzindo a transpiração e o suprimento de dióxido de carbono para o processo fotossintético, diminuição do crescimento celular e em alguns casos, aumento da fotorrespiração (SHINOZAKI; YAMAGUCHI-SHINOZAKI, 2007). A sua sobrevivência depende, principalmente, da velocidade de resposta aos estímulos externos (VIDAL et al., 2005). Os mecanismos fisiológicos de tolerância à seca também estão relacionados ao uso moderado da água pela planta, pelo controle da perda de água pelas folhas e pela habilidade das raízes em explorar camadas mais profundas do solo (NGUYEN et al., 1997). Tais respostas envolvem maior relação entre raiz e parte aérea, diminuição do volume das células, aumento da concentração do protoplasto, diminuição do tamanho das folhas, dissipação de energia da folha, ajuste osmótico mais eficiente, regulação estomática, acúmulo de metabólitos e, consequentemente, uma maior resistência à desidratação (KERBAUY, 2004; TAIZ; ZEIGER, 2004).

Ao longo dos anos, muitos estudos têm se voltado para a elucidação dos principais efeitos da seca no crescimento das plantas; entretanto, ainda existe carência de conhecimento sobre os efeitos primários do déficit hídrico em níveis bioquímicos e moleculares (ZHU, 2002; CHAVES et al., 2003; YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2005). Neste sentido, é necessário entender como são ativadas e como ocorrem essas respostas adaptativas, sendo este o ponto principal para o desenvolvimento de novas cultivares comerciais que sejam tolerantes à seca (ASSAD, 2002). Adicionalmente, parece existir uma comunicação cruzada entre a indução dos mesmos genes após diferentes tipos de estresses, como seca, salinidade e baixa temperatura. Tal mecanismo tem sido referido como "*cross-talk*" (reação cruzada), onde diferentes vias de sinalização compartilham de um ou mais intermediários (CHINNUSAMY et al., 2004). Um exemplo de *cross-talk* apresenta-se ilustrado na Figura 2 (YAMAGUCHI-SHINOZAKI; SHINOZAKI, 2005).

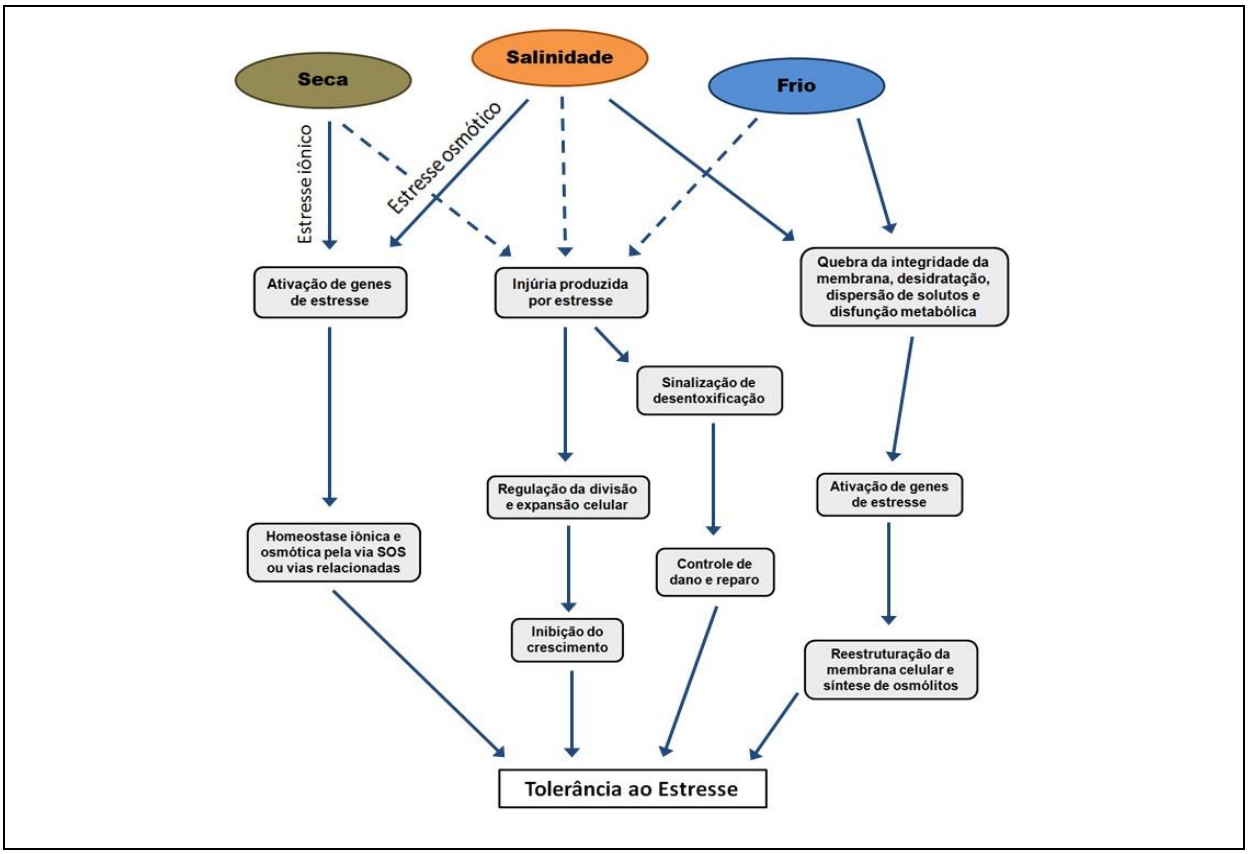


Figura 2. Via genérica da resposta vegetal frente ao estresse hídrico, salino e baixas temperaturas, evidenciando a comunicação cruzada entre as diferentes vias de sinalização. (Adaptado de Mahajan e Tuteja, 2005).

2.4.2 Estresse salino

De acordo com Paz et al. (2000) o termo salinidade se refere à presença de sais solúveis no solo a ponto de prejudicar o rendimento econômico das culturas. Além dos fatores naturais, a ação antrópica pode também induzir tal processo, principalmente com a utilização de água salina associada a uma irrigação inadequada, ou com uma drenagem insuficiente em solos com baixa condutividade hidráulica (ORCUTT; NILSEN, 2000). O manejo inadequado da água de irrigação aliado ao uso intensivo de fertilizantes têm contribuído para o aumento de áreas agrícolas com problemas de salinidade. Tais fatores são particularmente importantes nas regiões áridas e semi-áridas, principalmente devido à escassez da precipitação pluvial e à alta demanda evaporativa, fatores que dificultam a lixiviação dos sais localizados na camada arável do solo. Estima-se que no Brasil existam aproximadamente nove milhões de hectares com problemas de salinidade,

localizando-se a maior parte dessa área nos perímetros irrigados do Nordeste (CARNEIRO et al., 2002).

Apesar da existência de variabilidade genética para tolerância à salinidade (SHANNON; GRIEVE, 1998), os mecanismos bioquímicos e fisiológicos que contribuem para tal fator ainda são pouco conhecidos (MANSOUR et al., 2003). Além disso, o limite de tolerância depende da concentração do sal, do tempo de exposição, bem como do estádio de desenvolvimento das plantas (AYRES; WESTCOT, 1999).

O efeito da salinidade sobre as plantas pode ser restrito a dois componentes distintos: (I) o componente osmótico – resultante da elevada concentração de solutos do solo, provocando um déficit de água pela redução do potencial osmótico e (II) o componente iônico – decorrente dos elevados teores de Na+ e Cl-, e da alterada relação K+/Na+ (WILLADINO; CAMARA, 2010). Diversos trabalhos têm evidenciado os efeitos danosos dos íons que contribuem para a salinidade do solo sobre processos fisiológicos importantes para o crescimento das plantas, tais como de redução da taxa de absorção de água e osmólitos, ao enrijecimento da parede ou à queda no turgor celular (BETHKE; DREW, 1992; COSGROVE, 1993; YAHYA, 1998).

Um mecanismo comumente citado para tolerância à salinidade tem sido a capacidade das plantas em acumular íons no vacúolo e/ou solutos orgânicos de baixo peso molecular no citoplasma - um processo denominado ajustamento osmótico, permitindo a manutenção da absorção de água e da turgescência celular (HOPKINS, 1999). O acúmulo excessivo de sais pode ser evitado pela exclusão, suculência e redistribuição de sais. Por exemplo, algumas plantas retém íons em órgãos como raízes, parte superior do caule, pedúnculo da flor e pecíolo da folha, reduzindo a quantidade destes que chegam às folhas e aos frutos jovens (LACERDA et al., 2001).

2.5 AQUAPORINAS: DIVERSIDADE, CLASSIFICAÇÃO E ATIVIDADE

Sabe-se que o crescimento e o desenvolvimento vegetal dependem da regulação entre a captação e o transporte de água através das membranas celulares

e tecidos. Também não é recente o conhecimento dos mecanismos biofísicos que participam do transporte de água através de membranas biológicas. Entretanto, a descoberta de proteínas de membranas capazes de facilitar o transporte de água e/ou pequenos solutos e gases, tem popularizado o conceito de “canais de água de membrana” para membros de uma família de proteínas, chamadas aquaporinas (Agre, 2004).

Primeiramente descoberta na década de 90, a aquaporina nomeada de *CHIP28* foi isolada em eritrócitos humanos (PRESTON; AGRE, 1991). Em plantas o primeiro transportador identificado em *A. thaliana*, foi caracterizado por Maurel et al. (1993). Vários estudos vêm indicando que estas proteínas representam uma importante via seletiva para a movimentação da água através das membranas (Figura 3), permitindo o maior entendimento de como as plantas regulam o fluxo de água em diferentes condições fisiológicas (JAVOT; MAUREL, 2002).

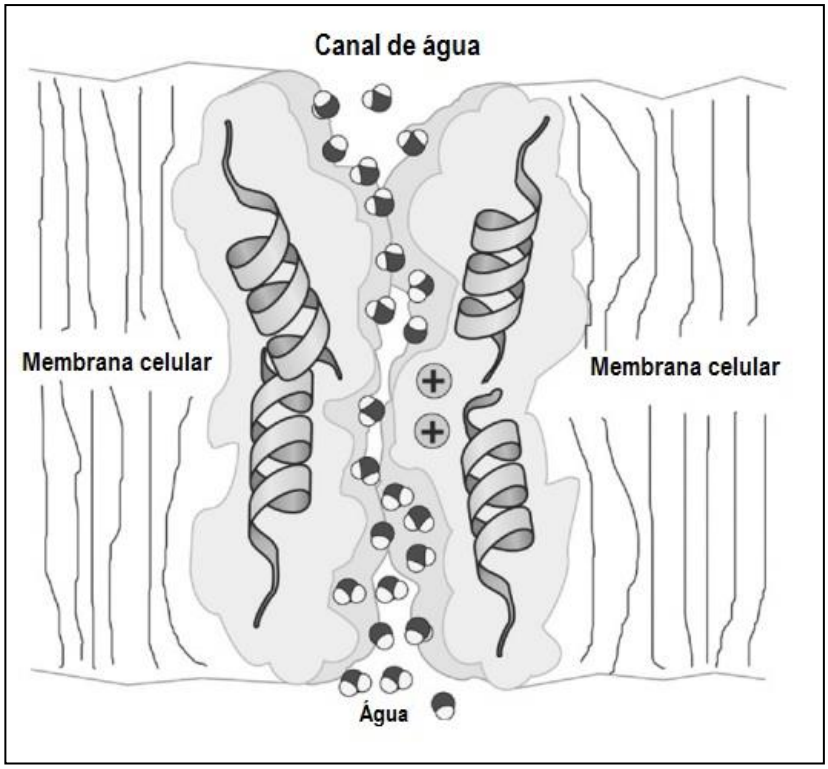


Figura 3. Modelo do transporte de moléculas de água através das aquaporinas nas membranas celulares. (Adaptada de Zhao et al., 2008).

As aquaporinas definem uma classe funcional de proteínas transportadoras de água que pertencem às MIPs (*Major Instrinsic Protein*; Proteínas Intrínsecas de Membrana), com 26-30 kDa, presentes em todos os reinos (FORREST; BHAVE,

2007). Estes transportadores especializaram-se em termos de localização, função (envolvidos no transporte de uma diversidade de substratos), além de apresentarem diversos efeitos na fisiologia vegetal (JOHANSON et al., 2001; FORREST; BHAVE, 2007; GOMES et al., 2009).

Ao contrário do que ocorre em animais, as plantas possuem um grande número de isoformas para aquaporinas. Segundo CHAUMONT et al. (2005), esta variedade de isoformas em plantas sugere que as duplicações proporcionam vantagens adaptativas para o crescimento sob as mais variadas condições, possivelmente como resultado das diferentes seletividades e mecanismos regulatórios. Até o momento 35 aquaporinas foram identificadas na planta-modelo *A. thaliana* (JOHANSON et al., 2001; QUIGLEY et al., 2001), 36 em milho (*Zea mays*) (CHAUMONT et al., 2001), 33 em arroz (*Oryza sativa*) (SAKURAI et al., 2005), 55 em *Populus tricocharpa* (GUPTA; SANKARARAMAKRISHNAN, 2009) e 28 isoformas para aquaporinas em uva (*Vitis vinifera*) (FOUQUET et al., 2008). Estudos baseados na similaridade entre sequências permitiram a classificação das aquaporinas vegetais em quatro subfamílias: proteínas intrínsecas de membrana plasmática (PIP – *Plasma Membrane Intrinsic Protein*); proteínas intrínsecas de tonoplasto (TIP – *Tonoplast Intrinsic Proteins*); proteínas intrínsecas de nódulos (NIP – *Nodulin26-like Intrinsic Proteins*) e pequenas proteínas básicas intrínsecas de membrana (SIP – *Small and Basic Intrinsic Proteins*) (CHAUMONT et al., 2001; ZARDOYA, 2005; MAUREL, 2007; ZHAO et al., 2008; GOMES et al., 2009).

A diversificação desta família proteica em quatro subfamílias foi estabelecida na emergência das primeiras plantas terrestres. Estas quatro subfamílias foram identificadas na briófita *Physcomitrella patens*, que possui 23 isoformas destes transportadores de membrana (BORSTLAP, 2002; DANIELSON; JOHANSON, 2008). Recentemente, três subfamílias adicionais foram reportadas para a primitiva *P. patens*, sendo: GIP (Homóloga aos canais de glicerol bacteriano) (GUSTAVSSON et al., 2005) encontrada em bactérias Gram-positivas; HIP (Proteínas híbridas intrínsecas de membrana) e XIP (Proteínas intrínseca desconhecida X) (DANIELSON; JOHANSON, 2008). Por se tratarem de novas subfamílias, poucas informações têm sido disponibilizadas. A única informação até o momento refere-se à subfamília XIP (BIENERT et al., 2011) a qual apresenta função similar à subfamília

NIP, com afinidade por H_2O_2, boro e glicerol, mas com pouca eficiência no transporte de água.

Várias proteínas de membrana da família MIP transportam outras substâncias que não água, podendo atuar no transporte adicional de NH_3, CO_2, H_2O_2 e outros pequenos solutos (TYERMAN et al., 2002; LUDEWIG; DYNOWSKI, 2009). Membros da subfamília PIP atuam no transporte de água, enquanto que os membros da subfamília TIP transportam água, amônia, ureia e glicerol. Os membros da subfamília NIP estão associados ao transporte de água, glicerol ou boro (raízes), enquanto SIP apresenta afinidade de transporte por pequenos solutos básicos (MAESHIMA; ISHIKAWA, 2008).

2.5.1 Proteínas Intrínsecas de Membrana Plasmática – PIP

As proteínas intrínsecas de membrana plasmática (PIP – *Plasma Membrane Intrinsic Protein*) apresentam peso molecular de aproximadamente 30 kDa e um ponto isoelétrico (pI) de 9,0 em decorrência de diversos aminoácidos básicos na extremidade C-terminal (GOMES et al., 2009). De um modo geral, essa subfamília localiza-se em tecidos com alto fluxo de água, como tecidos vasculares, células guarda de estômatos e flores. Membros da subfamília PIP são altamente expressos em raízes, elementos de floema e folhas, podendo ter importante papel na comunicação célula-a-célula (DYNOWSKI et al., 2008).

Constituem a maior subfamília de aquaporinas vegetais, primordialmente localizadas em membranas plasmáticas. Filogeneticamente este grupo pode ser dividido em PIP1 e PIP2, que diferem entre si com relação ao tamanho das extremidades N- e C- terminal, assim como na permeabilidade à água (KALDENHOFF; FISCHER, 2006). Estudo de expressão heteróloga em oócitos de *Xenopus* revelou maior eficiência (5-20 vezes maior) no transporte de água dos membros pertencentes a PIP2, quando comparados com representantes do grupo PIP1 (DANIELS et al., 1994).

Os representantes de PIP1 possuem uma longa extremidade N-terminal, mas uma cauda C-terminal curta (GOMES et al., 2009). Já os membros de PIP2 apresentam uma cauda amino-terminal curta e uma extremidade carboxi-terminal longa, além de 4-10 aminoácidos adicionais no primeiro *loop* extracitosólico. Os membros de PIP2 apresentam funções fisiológicas associadas ao transporte de água em raízes, folhas, órgãos reprodutivos e à germinação de sementes (LOPEZ et al., 2003; BOTS et al., 2005).

Todas as PIPs apresentam na região arginina/aromática (ar/R) os resíduos Phe-His-Thr-Arg (F..H..T..R). Em contraste, uma variedade de combinações nesses resíduos é observada em TIP, NIP e SIP, incluindo diferenças dentro dos subgrupos nas diferentes subfamílias, onde cada modificação está provavelmente associada a especializações funcionais. Por exemplo, o resíduo Arg na segunda posição do *loop* E mostra-se importante no transporte de água (FORREST; BHAVE, 2007).

2.5.2 Proteínas Intrínsecas de Tonoplasto – TIP

A subfamília TIP de aquaporinas apresenta um peso molecular entre 25 a 28 kDa, com ponto isoelétrico em torno de 6,0. Tratam-se de canais de transporte para água, mas que também podem ser permeáveis para ureia, glicerol e amônia. Em *A. thaliana*, as aquaporinas *AtTIP1;1*, *AtTIP1;2* e *AtTIP2;3*, foram caracterizadas como capazes de transportar também peróxido de hidrogênio (H_2O_2) (BIENERT et al., 2007).

Membros de TIP são localizados na membrana do vacúolo celular, sendo responsáveis pelo fluxo de água e pequenos solutos através da membrana, atuando também nos processos de regulação do turgor, sinalização e degradação celular. Diversas TIPs estão relacionadas com a osmorregulação, além do transporte de pequenos solutos e gases, juntamente com o transporte de metabólitos para vias importantes como o ciclo da ureia ou de síntese de aminoácidos (KALDENHOFF; FISCHER, 2006).

A expressão de membros da subfamília TIP ocorre predominantemente na membrana de tonoplasto, sendo essencial para a movimentação intracelular de água e outras pequenas moléculas. Tratam-se das proteínas mais abundantes no

tonoplasto, permitindo o rápido ajustamento osmótico, atuando na manutenção da osmolaridade e do turgor (GOMES et al., 2009).

2.5.3 Proteínas Intrínsecas de Nódulos – NIP

Em diversos experimentos, membros das NIPs mostraram-se funcionais para o transporte de água e glicerol, mas exibem primariamente baixa permeabilidade para a água. Seus transcritos são identificados em sementes, caules e raízes, além das membranas do simbiossomo, quando estas existem (KALDENHOFF; FISCHER, 2006; GOMES et al., 2009). Esta subfamília é considerada como a detentora das aquagliceroporinas vegetais.

Estruturalmente as NIPs podem ser divididas em dois grupos com relação ao filtro ar/R, o primeiro grupo relacionado ao transporte de água e de glicerol, e o segundo associado ao transporte de solutos maiores, como a ureia. A expressão dos membros desta subfamília é inferior à de PIP e TIP, estando associada a especializações celulares, tais como zonas de elongação radicular. Devido à sua localização e expressão particulares, os processos em que este grupo atua, permanecem em grande parte desconhecidos (GOMES et al., 2009).

2.5.4 Pequenas Proteínas Básicas Intrínsecas de Membrana - SIP

A subfamília de pequenas proteínas básicas intrínsecas de membrana representa o menor grupo dentre as MIPs, sendo identificado primariamente por meio de mineração de dados nos bancos de sequências, associado à análise filogenética. As proteínas desta família são pequenas como as integrantes da subfamília TIP, mas altamente básicas (KALDENHOFF; FISCHER, 2006).

Esta família proteica possui um ponto isoelétrico básico e alta divergência com relação às outras MIPs, incluindo uma modificação no motivo NPA localizado no *loop* B e uma extremidade N-terminal curta. Essas modificações nos aminoácidos conferem uma maior abertura no poro, havendo uma potencial especificidade por substratos diferentes (GOMES et al., 2009). A função das SIPs vegetais ainda encontra-se indefinida, entretanto, estudos *in vitro* indicam que estas atuam no

transporte de água através da membrana do retículo endoplasmático, regulando o volume e a concentração iônica do lúmen, assim como a morfologia da organela (FORREST; BHAVE, 2007).

2.6 ESTRUTURA E LOCALIZAÇÃO DAS AQUAPORINAS

As aquaporinas são geralmente descritas como sendo tetraméricas, tanto *in vitro* quanto *in vivo*, onde cada tetrâmero é formado por quatro canais independentes (Figura 4) (ENGEL et al., 2000). A tetramerização parece ser resultado da interação entre monômeros vizinhos (MURATA et al., 2000). Embora vários estudos relatem a formação de homotetrâmeros, várias isoformas vegetais apresentam estruturas heteromórficas, as quais influenciam a afinidade de soluto e eficiência de transporte (CHAUMONT et al., 2005).

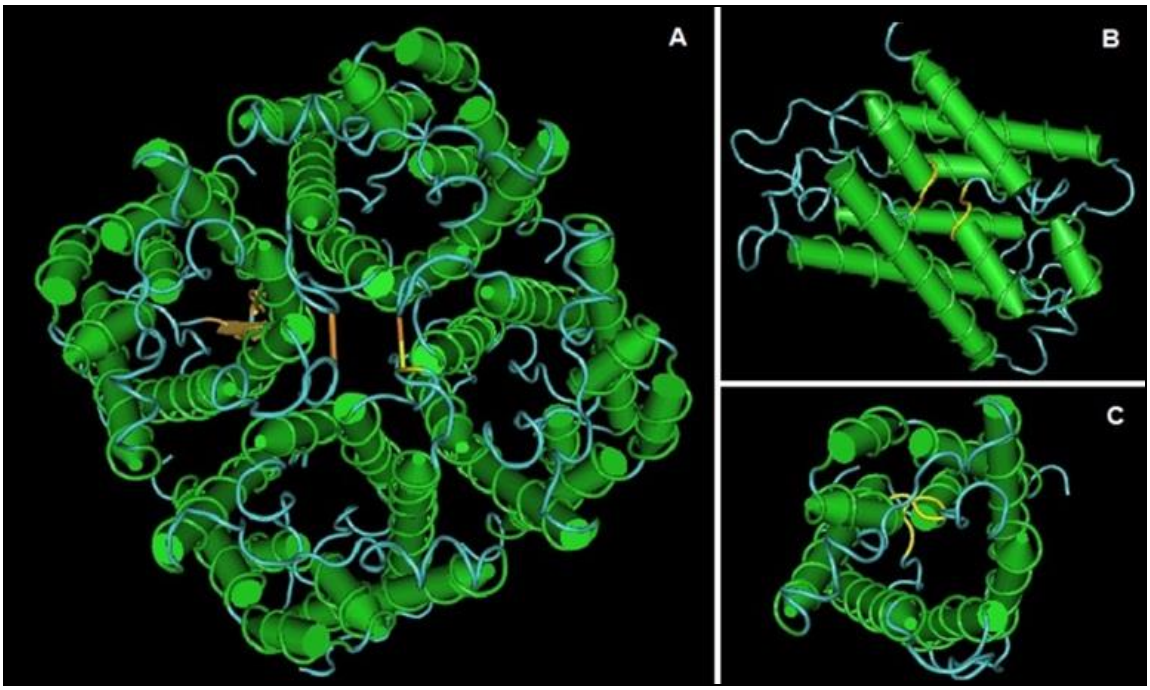


Figura 4. Representação da estrutura molecular da aquaporina de espinafre (*Spinacia oleracea*, Chenopodiaceae). A. Estrutura da conformação aberta da SoPIP2;1 [Protein Data Bank ID: 36623], arranjo tetramérico típico; B e C. Estrutura do monômero, conformação lateral e superior, em amarelo evidenciados os dois motivos NPA (Asn-Pro-Ala).

As estruturas das MIPs (Figura 5) são altamente conservadas, apresentando caracteristicamente: seis alfa-hélices hidrofóbicas que formam os domínios transmembrana (H1-H6); cinco *loops* entre as alfa-hélices (A-E, onde os *loops* B e E

são hidrofóbicos, parcialmente inseridos nas membranas e os demais são hidrofílicos), havendo dois motivos NPA (asparagina-prolina-alanina) altamente conservados nos loops B e E (FORREST; BHAVE 2007; GOMES et al., 2009).

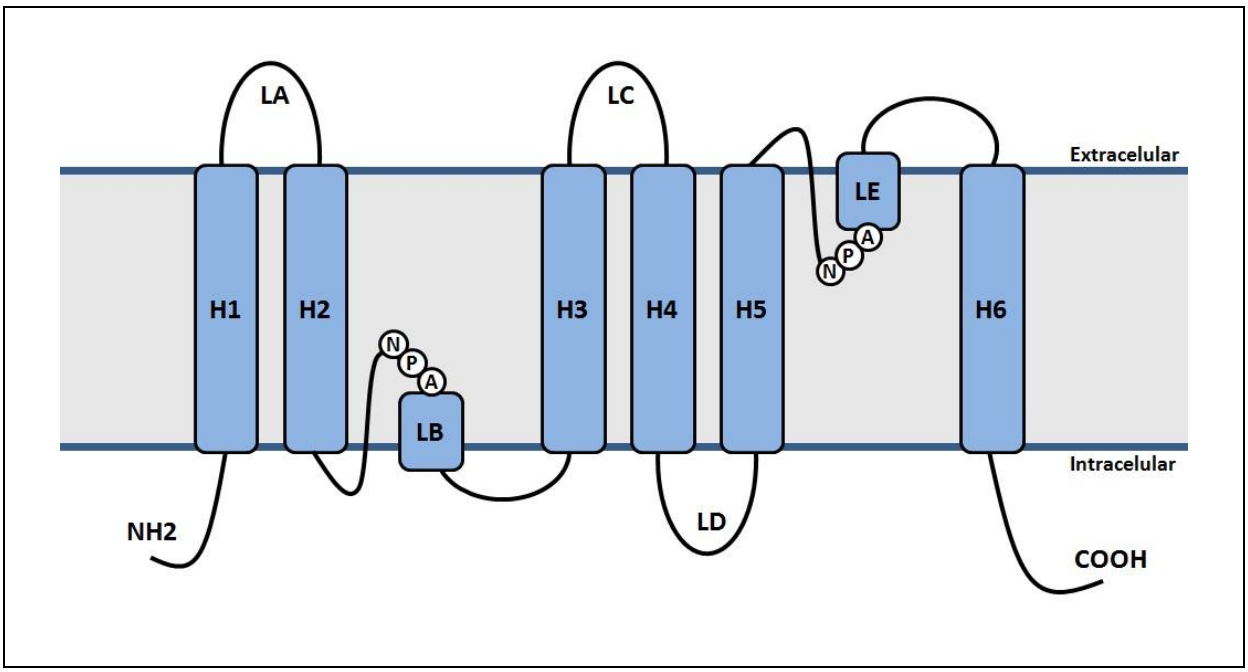


Figura 5. Modelo da estrutura de uma aquaporina, evidenciando as principais características estruturais. Alfa-hélices representadas pelos retângulos azuis. Nela estão presentes os seis domínios transmembrana (H1-H6), conectados pelos cinco loops (LA-LE). Dois domínios alfa-hélices nos loops B e E, contendo os dois motivos NPA, formadores do poro (esquema do autor).

Com base nesta estrutura, as aquaporinas podem ser divididas em duas metades, cada uma com três hélices transmembranas H1-H3 ou H4-H6 (TMH, *transmembrane helices*) e um *loop* hidrofóbico contendo o motivo NPA, constituindo o modelo *hourglass* (FORREST; BHAVE, 2007).

A base molecular para a especificidade de substrato tem sido amplamente estudada e vem demonstrando que apesar da estrutura geral ser conservada, existem pontos específicos nas MIPs que determinam diferenças nos canais de seletividade. Os motivos NPA são amplamente reconhecidos como importantes na atividade transportadora devido à sua conservação em diversas MIPs de animais, plantas, protozoários, fungos, eubactéria e Archeae (ZARDOYA, 2005). Estes motivos localizam-se no centro das duas metades, nos *loops* B e E (LB e LE, Figura 5), formando o poro seletivo, que contribui para a função transportadora do canal (FORREST; BHAVE, 2007).

A segunda estrutura associada à seletividade de substrato nas diferentes MIPs para água ou outras moléculas polares são os resíduos de aminoácidos conservados, incluindo a região arginina/aromática (ar/R). Esta região compreende uma tétrade formada por dois resíduos conservados localizados nas alfa-hélices H2 e H5, bem como duas no *loop* E (LE1 e LE2), localizada no centro do poro, formando uma segunda constrição, separada do motivo NPA, atuando na afinidade por substratos (FORREST; BHAVE, 2007). O resíduo de arginina altamente conservado torna-se importante por prover pontes de hidrogênio para o transporte de moléculas de água ou glicerol, ou ainda repelir cátions do poro (SUI et al., 2001). A propriedade desses resíduos em controlar a especificidade de substrato, pode ser um eficiente modelo de predição funcional para essas proteínas, definindo a afinidade de transporte das mesmas (THOMAS et al., 2002).

As aquaporinas vegetais localizam-se em todos os compartimentos subcelulares. Este padrão de localização reflete a alta compartimentalização celular nas plantas, sendo necessário para o controle do transporte de água e pequenos solutos não apenas na membrana celular externa, mas também, através das membranas intracelulares (MAUREL et al., 2008).

De um modo geral, as aquaporinas das subfamílias PIP e TIP estão localizadas nas membranas plasmáticas e vacúolos, respectivamente. Já os membros das subfamílias NIP e SIP possuem localização subcelular incerta (CHAUMONT et al., 2005), embora tenha sido evidenciada a localização desses membros em membranas do retículo endoplasmático e na membrana plasmática de células do simbiossomo, nos nódulos radiculares (Figura 6) (MAUREL, 2007; MAESHIMA; ISHIKAWA, 2008).

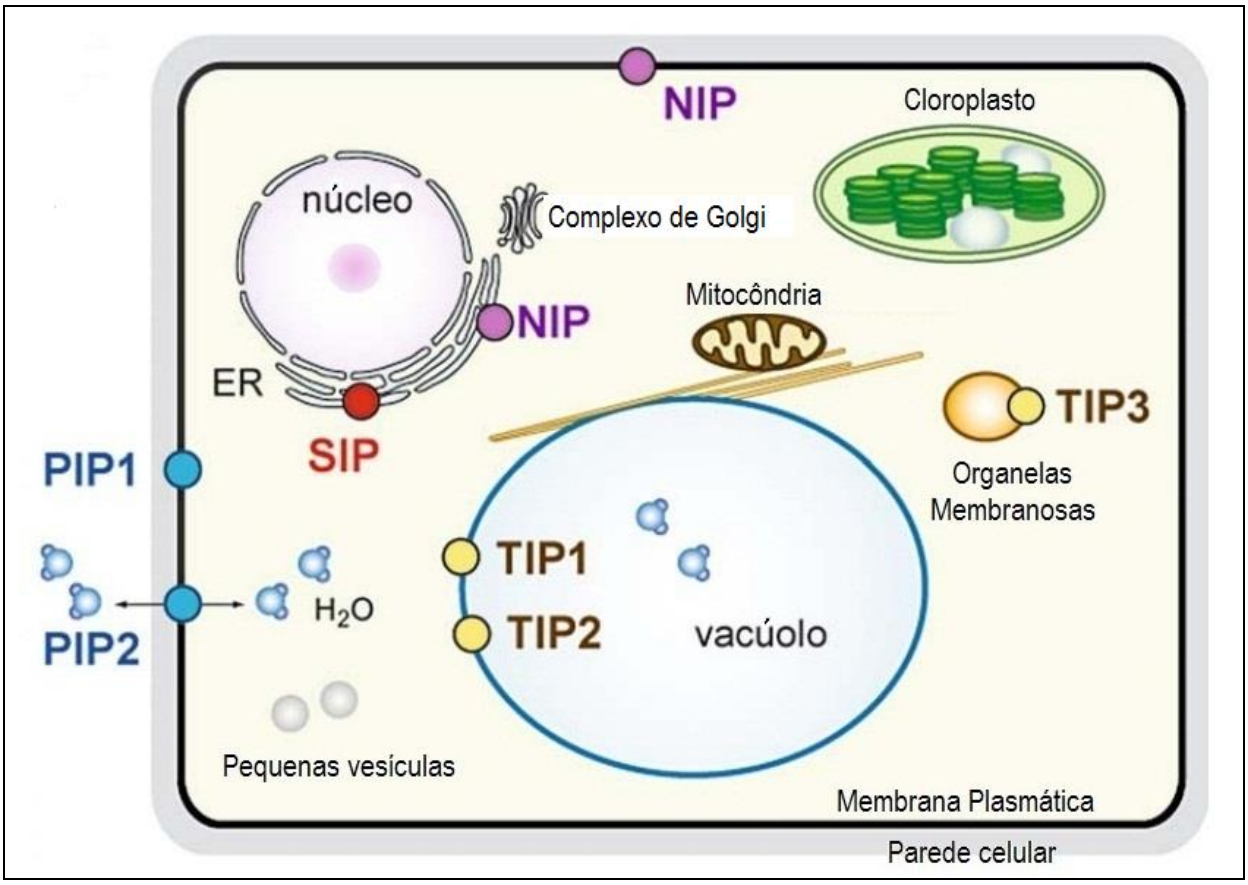


Figura 6. Diagrama esquemático da localização das diferentes famílias de aquaporinas nas organelas vegetais. PIPs: localizadas na membrana plasmática; TIPs: localizadas no vacúolo; NIPs: localizadas no retículo endoplasmático e membrana plasmática; SIPs: Localizada no retículo endoplasmático. Adicionalmente ao transporte de água, transportam: PIP (CO2 e ureia), TIP (ureia, amônia, glicerol e H2O2), NIP (amônia, ureia, glicerol, boro e lactato). Adaptado de Maeshima e Ishikawa (2008).

2.7. FUNÇÃO FISIOLÓGICA E EXPRESSÃO

Monitorando o perfil de expressão gênica para aquaporinas em diversas espécies, tecidos, tipos celulares ou em resposta a fitormônios e fatores ambientais, verificou-se que as MIPs controlam o transporte extensivo de água das raízes para as folhas durante o processo de transpiração, bem como no transporte de assimilados nos elementos de floema, apresentando atividade na abertura e fechamento estomático, movimentos foliares e no controle da homeostase (Figura 7) (TYERMAN et al., 2002; SANTONI, 2006; UEHLEIN; KALDENHOFF, 2008).

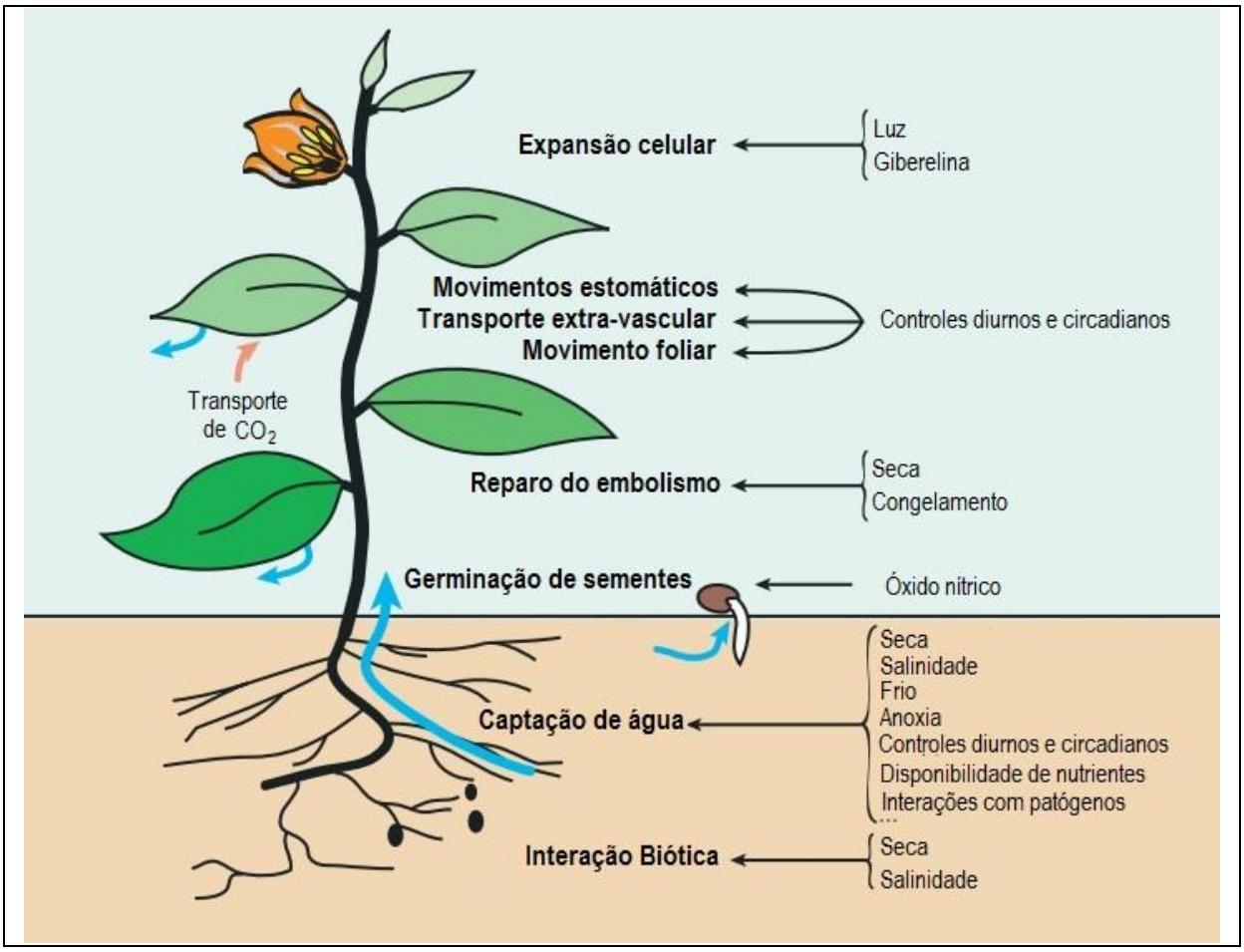


Figura 7. Funções integradas das aquaporinas vegetais. Fatores endógenos e ambientais atuando no organismo, ativando a funcionalidade das aquaporinas nos diferentes órgãos e tecidos. Adaptado de Maurel et al. (2008).

A regulação de diferentes processos por parte das aquaporinas representa o mecanismo básico de modulação da permeabilidade hídrica frente às diferentes condições ambientais e à disponibilidade hídrica. Em plantas, por exemplo, a atividade das MIPs pode ser influenciada por estresse salino e pelo estresse osmótico, os quais têm efeito no seu estado de fosforilação, em seus níveis de transcrição e/ou nos níveis de degradação dessas proteínas. Além desses mecanismos, a realocação dessas proteínas em resposta ao sal tem sido demonstrada em diversos membros das subfamílias TIP e PIP, onde transportadores inicialmente localizados nas membranas do tonoplasto são realocadas em invaginações intravacuolares ou outras membranas vesiculares (GOMES et al., 2009).

De um modo geral, a expressão de MIPs é regulada durante o desenvolvimento vegetal através de sinalização hormonal e condições ambientais

adversas (ex. seca, infecção por nematoides, baixas temperaturas, salinidade), sendo que esta regulação da expressão (super expressão / repressão) pode variar conforme os estímulos e/ou órgãos analisados (HACHEZ et al., 2006). Já, quando analisada a expressão de diversos membros das subfamílias PIP e TIP nos diferentes órgãos vegetais sob condições normais (sem estresse), observa-se uma maior prevalência destes transportadores em raízes do que em folhas (JA VOT; MAUREL, 2002; HACHEZ et al., 2008; ZHANG et al., 2008), apesar de alguns estudos apontarem membros das MIPs como sendo exclusivos em folhas quando comparadas a raízes (HEINEN et al., 2009).

Vários genes de aquaporinas têm sido caracterizados em diferentes espécies, como sendo importantes desde a germinação até a reprodução vegetal. A germinação de sementes representa uma etapa crucial do desenvolvimento vegetal, na qual a presença de aquaporinas pode regular o transporte de água requerido durante a embebição, o crescimento embrionário e a protrusão da radícula. Dois genes de aquaporinas isolados de sementes de canola (*Brassica napus*), *BnPIP1* e *BnTIP2*, foram relacionados a etapas distintas do processo germinativo. O gene *BnPIP1* atuaria preferencialmente no transporte de água necessário para a ativação das enzimas e mobilização das reservas nos estágios iniciais da germinação, enquanto a expressão de *BnTIP2* estaria relacionada ao alongamento celular associado à protrusão da radícula (GAO et al., 1999).

Estudos realizados em sementes de ervilha (*Pisum sativum*) evidenciam o envolvimento de aquaporinas no fluxo de água e solutos através da membrana das células parenquimáticas do tegumento, responsáveis pelo fornecimento de nutrientes, tanto para o desenvolvimento dos tecidos, quanto para a síntese de compostos de reserva. Além da função de transporte, uma aquaporina do tipo NIP, expressa no tegumento da semente, apresentou alta capacidade de transporte de glicerol quando testada em sistema heterólogo (SCHUURMANS et al., 2003). Uma análise mais ampla do perfil de expressão de diversas aquaporinas de *Arabidopsis* pela técnica de microarranjo demonstrou que dois genes *TIP3* e um *TIP5* são altamente expressos na semente quiescente, mas durante a germinação estes tem sua expressão reduzida, sendo expressos preferencialmente genes das subfamílias *TIP1* e *TIP2*. Membros do tipo *PIP* em sementes de

Arabidopsis apresentaram expressão discreta, elevando seus níveis de expressão apenas na fase final de crescimento embrionário (WILLIGEN et al., 2006).

Diversas aquaporinas se mostram importantes no processo reprodutivo, estando relacionadas com o desenvolvimento das estruturas de reprodução, ou com a própria viabilidade do grão de pólen. Em tabaco, genes de aquaporinas de membrana plasmática (*PIP1* e *PIP2*) foram identificados durante o desenvolvimento da antera, além disso, plantas de tabaco transgênicas com silenciamento do gene *PIP2* apresentaram desidratação mais lenta e atraso na deiscência das anteras (BOTS et al. 2005). Em *Arabidopsis*, foram identificados dois genes de aquaporinas de tonoplasto, denominados *TIP5;1* e *TIP1;3*, com alta expressão nos grãos de pólen maduros e capazes de transportar água e uréia. Mutantes deficientes desses genes possuem tubos polínicos mais curtos quando germinados em meio com baixas concentrações de nitrogênio, sugerindo que nestas condições estes genes são importantes para o alongamento do tubo polínico (SOTO et al., 2008). Ainda em *Arabidopsis*, oito genes *PIP* foram detectados nos botões florais, sugerindo o envolvimento das aquaporinas no desenvolvimento floral (QUIGLEY et al., 2001).

Nas plantas, o movimento de água e solutos é altamente regulado, sugerindo a presença de aquaporinas em quase todos os processos fisiológicos relacionados ao crescimento e desenvolvimento vegetal. Estudos com *Arabidopsis* e tabaco para os genes *PIP1* e *PIP2* demonstraram que estas subfamílias estão diretamente relacionadas ao transporte de água em raízes, sobretudo em situação de déficit hídrico (KALDENHOFF et al.,1998; SIEFRITZ et al., 2002). Já em arroz (*Oryza sativa*), a super-expressão de *PIP2;1* levou a um aumento de aproximadamente 150% na biomassa de raízes e na condutividade hidráulica. No entanto, o cultivo em 100 mM de NaCl levou a uma diminuição do crescimento após duas semanas, sugerindo uma maior sensibilidade das plantas transgênicas ao estresse salino (KATSUHARA et al., 2003).

Assim como em raízes, o fluxo radial de água em folhas, que é controlado pela evapotranspiração, implica no transporte de água através das membranas e, consequentemente, na presença de proteínas de canais de água. Também em folhas, a expressão de diferentes aquaporinas obedece a um padrão de regulação temporal e espacial. Enquanto algumas isoformas são expressas em folhas jovens em expansão, outras são expressas preferencialmente em folhas completamente

expandidas, sendo que as aquaporinas expressas nas folhas maduras estão provavelmente envolvidas em processos fisiológicos como carregamento do floema e do xilema, abertura e fechamento estomático, transporte de CO_2 na fotossíntese e movimento foliar (HEINEN et al., 2009).

Sarda et al. (1997) identificaram dois genes de aquaporina expressos em folhas de girassol, mais especificamente nas células-guarda dos estômatos, denominados *SunTIP7* e *SunTIP20*. A variação no nível de transcritos de *SunTIP7* acompanhava as alterações de condutância estomática ao longo do dia e aumentava em plantas submetidas ao déficit hídrico, sugerindo o papel desta aquaporina na modulação da abertura e fechamento estomático (SARDA et al. 1997). Em cevada, uma aquaporina funcional, a *HvPIP1;6*, é expressa especificamente na epiderme foliar, com maior expressão na zona de alongamento, sugerindo a atuação dessa proteína no transporte de água associado ao crescimento celular (WEI et al., 2007).

2.7.1 Aquaporinas e resposta a estresses

Embora as aquaporinas tenham um papel chave na regulação do transporte de água, resultados contrastantes têm sido observados em situações de estresse. Como relatado acima, a condutividade hidráulica de raízes é regulada parcialmente por aquaporinas, no entanto, não é possível encontrar uma resposta comum dessa classe de proteínas em raízes submetidas a uma situação de déficit hídrico. A expressão de genes *PIP* foi analisada em condição de seca, mostrou que, de 37 genes *PIP* estudados, 15 tiveram sua expressão diminuída, 13 tiveram aumento de expressão e nove mantiveram seus níveis de transcritos inalterados, indicando que cada gene *PIP* deve ter uma função específica na resposta ao estresse (AROCA et al., 2011).

Da mesma forma, respostas diferenciais à seca foram observadas em plantas transgênicas com inibição ou indução de determinadas isoformas de PIP. Plantas de *Arabidopsis* com silenciamento de genes *PIP* apresentaram queda na capacidade de recuperação após imposição de déficit hídrico (MARTRE et al., 2002), enquanto plantas de arroz com aumento de expressão de aquaporinas *PIP2*, se mostraram hipersensíveis aos estresses hídrico e salino (KATSUHARA et al. 2003). Com

relação às aquaporinas de tonoplasto, comportamentos diferenciais das plantas transformadas também foram observados. Sade et al. (2009) demonstraram que a expressão constitutiva de *TIP2;2* em plantas de tomate foi capaz de modular a transpiração em diferentes situações ambientais e manter as funções fisiológicas, crescimento e produtividade em níveis normais, mesmo em condições de estresse hídrico e salino.

Contrariamente ao observado em tomate, plantas de *Arabidopsis* super-expressando uma *TIP* de soja se mostraram mais suscetíveis à desidratação (WANG et al., 2011). Segundo Hussain et al. (2011), existem atualmente interpretações divergentes dos resultados obtidos com plantas transgênicas. Para alguns autores, o aumento nos níveis de aquaporina melhora a capacidade das plantas em lidar com o estresse hídrico, enquanto outros acreditam que as plantas evitam perdas excessivas de água ao "desligar" as aquaporinas durante a desidratação.

Além da seca, diversas isoformas de aquaporinas são moduladas por estresse salino e osmótico, por frio e por fitormônios, com grande variedade de respostas, como observado para o déficit hídrico (SAKURAI et al., 2005; PENG et al. 2007). No entanto, é importante ressaltar a complexidade dos mecanismos de resposta à seca e aos demais estresses, assim, em conjunto com as aquaporinas, outros fatores são responsáveis pela resposta de tolerância (MAUREL et al., 2008).

2.8 IMPORTÂNCIA DAS “OMICAS” NA PROSPECÇÃO E ANÁLISE DE GENES

A genômica pode ser 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). A partir da década de 90, com o sequenciamento completo do primeiro organismo de vida livre em 1995 (FLEISCHMANN et al., 1995) e, logo em seguida o sequenciamento do genoma de seres complexos, incluindo plantas, animais e seres humanos, acreditava-se que a genômica seria a solução para todos os problemas. Entretanto, a constatação de que os dados trazidos pelo sequenciamento das moléculas de DNA genômico, embora indiscutivelmente relevantes, eram limitados, não sendo possível elucidar muitos dos processos biológicos, em especial, o que ocorre após a duplicação, a

transcrição e a tradução do material genético (PANDEY; MANN, 2000; PIMENTA, 2003).

É preciso ter conhecimento, por exemplo, sobre quais proteínas estão realmente sendo expressas, quando estão sendo expressas e em quais níveis esta expressão ocorre, assim como as possíveis modificações pós-transcricionais. Tanto dados de genomas, como a identificação dos genes e suas funções num determinado organismo são fundamentais para a compreensão dos processos biológicos tais como: a diferenciação celular, a morfogênese, a determinação fenotípica, a resistência a patógenos e a adaptabilidade. Neste sentido, a grande quantidade de informações geradas para espécies de plantas-modelo como *A. thaliana*, arroz e *Populus* deram início à era das "ômicas" em plantas. Esta nova era na genética molecular compreende o estudo de transcritos (transcriptômica), de proteínas (proteômica) (CHEN; HARMON, 2006), dos metabólitos celulares (metabolômica) (BHALLA et al., 2005), com o objetivo maior de traçar perfis funcionais da célula e/ou tecido, fornecendo uma visão holística da interação genes-fenótipo.

2.8.1 TRANSCRIPTÔMICA

A transcriptômica (genômica funcional) refere-se ao conjunto de transcritos de uma célula ou tecido, expressos em um dado momento, representando o primeiro passo na caracterização funcional dos genes e identificação de grupos de genes funcionalmente associados. O transcriptoma reflete os genes ativos e inativos em determinado tempo e condição ambiental, tornando-se importante no entendimento de diversos processos biológicos, tanto na identificação, como na compreensão das complexas redes metabólicas a que estão sujeitos (ALBA et al., 2004; SUBRAMANIAN et al., 2005). Assim, com base na identificação dos genes, é possível estabelecer os padrões de expressão e suas funções.

Apesar de serem bem definidos, sensíveis e robustos, os métodos tradicionais de análise da expressão gênica (*Northern Blot*, *Hibridização in situ*, *Polymerase Chain Reaction, etc.)* são restritos à análise de um número pequeno de genes. Assim, o estabelecimento de perfis transcricionais tem proporcionado um enorme avanço, pois estes permitem uma análise eficiente em alta escala (*high-*

throughput) de um grande número de genes ao mesmo tempo (DONSON et al., 2002; MEYERS et al., 2004).

As metodologias utilizadas para estabelecer perfis transcricionais baseiam-se na geração de populações de cDNA a partir da população de RNAm, as quais podem ser divididas em métodos de análise abertos (globais) e métodos fechados. Os métodos globais permitem o acesso a potencialmente todos os transcritos expressos em um determinado momento porque não requerem informações prévias sobre o número de transcritos para a espécie. Já os métodos fechados são limitados pela quantidade de informação disponível para a espécie, uma vez que utilizam sondas e/ou primers específicos (ALBA et al., 2004).

Tais metodologias são amplamente utilizadas em genomas com grandes regiões não codificadoras, a exemplo dos eucariotos, onde são sequenciadas apenas as regiões expressas. Uma destas metodologias baseia-se na geração de bibliotecas de Etiquetas de Sequências Expressas (*Expressed Sequence Tags* - ESTs), cuja base consiste no fato de que o nível de mRNA tecido-específico presente em uma espécie é refletido pela frequência da ocorrência de seu EST correspondente em uma biblioteca de cDNA (ADAMS et al.,1991). Os ESTs, no entanto, não representam qualitativamente o conjunto de RNAm expressos já que transcritos com baixo número de cópias não são representados nas bibliotecas (BHALERAO et al., 2003; SUN et al., 2004). Por outro lado, o sequenciamento de ESTs e o acúmulo de informações genômicas são fundamentais em outras abordagens como, por exemplo, para a construção de plataformas de análise baseadas em arranjos de DNA (ex. *Microarrays*).

Entre as metodologias disponíveis de análise de transcriptomas a metodologia de *microarrays* (compreendendo conjuntos de cDNA ou de oligonucleotídeos), também conhecidos como plataformas de chips de nucleotídeos, tem sido a mais amplamente utilizada em análise de transcritos em plantas (MOREAU et al., 2005; ANDERSSON-GUNNERAS et al., 2006). No caso dos microarrays de cDNA, a plataforma pode ser preparada diretamente a partir de sequências de cDNA disponíveis nos bancos de dados públicos (DRMANAC; DRMANAC, 1999), além poder utilizar a mesma plataforma na hibridização de duas ou mais amostras diferentes, usando florescências de cores diferentes para produção de perfis comparativos (AHARONI; VORST, 2001). De outra forma, os

microarrays de oligonucleotídeos possibilitam a análise de um número muito maior de genes com hibridização específica de um único gene com cada oligonucleotídeo, além de apresentarem-se bastante uniforme em termos de temperatura de anelamento e concentração, sendo a variação experimental reduzida, aumentando assim a precisão e confiabilidade estatística (ALBA et al., 2004).

Além de bibliotecas de ESTS, outras metodologias estão atualmente disponíveis, permitindo uma análise global do transcriptoma, tais como: a tecnologia SAGE (Análise serial da expressão gênica; *Serial Analysis of Gene Expression* - SAGE) (VELCULESCU et al., 1995), e seus derivados, a Expressão gênica diferencial (*Differential display* – DD) (LIANG; PARDEE, 1992), Biblioteca subtrativa suprimida (*Supression Subtractive Hybridization* - SSH) (DIATCHENKO et al., 1996), *Amplified Fragment Length Polymorphism* baseado em cDNA (cDNA-AFLP) (BACHEM et al., 1996), bem como o Sequenciamento Massivo de Assinatura Paralela (*Massively Parallel Signature Sequencing* - MPSS) (BRENNER et al., 2000).

2.8.1.1 Análise Serial da Expressão Gênica (SAGE) e seus derivados

A técnica de SAGE (*Serial Analysis of Gene Expression*), desenvolvida por Velculescu et al. (1995) permite identificar quais genes são expressos em um dado momento, possibilitando uma análise quantitativa acurada da expressão de um grande número de transcritos simultaneamente, inclusive daqueles ainda não conhecidos, com a exclusividade de permitir que o perfil dos transcritos seja apresentado na forma de dados digitais (MATSUMURA et al., 1999).

Nesta metodologia, um fragmento de DNA de 9-11 pares de base (*tag*) adjacente ao sítio de restrição *Nla* III, localizado próximo à extremidade 3’ de cada cDNA, é extraído por uma série de ligações com adaptadores e digestão com enzimas de restrição. As *tags* geradas são amplificadas por PCR, concatenadas, clonadas em um vetor plasmidial e por último sequenciadas (Figura 8). Devido à concatenação, o sequenciamento de uma biblioteca SAGE proporciona maior economia quando comparado ao sequenciamento de bibliotecas de ESTs, que se baseiam na clonagem de sequências expressas que representam apenas um único transcrito por clone. Dessa forma, a partir de um mesmo número de clones

sequenciados, na técnica SAGE é possível identificar um número muito maior de transcritos (SUN et al., 2004).

O ponto crítico do método é a associação de cada *tag* ao transcrito de origem que deve ser realizada através de um banco de dados de sequências, processo denominado de *tag mapping*. Essa associação é feita através de um banco de dados de sequências, onde somente devem ser considerados os resultados nos quais a *tag* encontra-se posicionada imediatamente após o último sítio de reconhecimento da enzima utilizada, antes da cauda poli(A), na extremidade 3´. As *tags* devem, preferencialmente, ser mapeadas em bancos de ESTs 3´ ou em sequências completas de cDNA, o que é muito difícil no caso de plantas e principalmente para espécies arbóreas, uma vez que os investimentos concentram-se principalmente nos estudos envolvendo plantas herbáceas e humanos (LASH et al., 2000).

Dessa forma, algumas alternativas têm sido utilizadas na tentativa de se realizar o *tag mapping*. Uma delas é o SAGEmap disponível no site do NCBI (http://www.ncbi.nlm.nih.gov/sage), que extrai *tags* "virtuais" a partir de sequências expressas (ESTs e cDNAs) depositadas que contenham a terminação poli(A) ou o sinal de poliadenilação e as compara com o conjunto de *tags* obtidas experimentalmente (PLEASANCE et al., 2003). Uma alternativa adicional envolve o uso de bancos genômicos, onde as possíveis *tags* "virtuais" são extraídas do genoma de uma espécie e comparadas com um conjunto das *tags* obtidas experimentalmente, permitindo não só uma maior porcentagem de identificação dos genes, como também evidenciando a regulação por *splicing* (FIZAMES et al., 2004; ROBINSON et al., 2004).

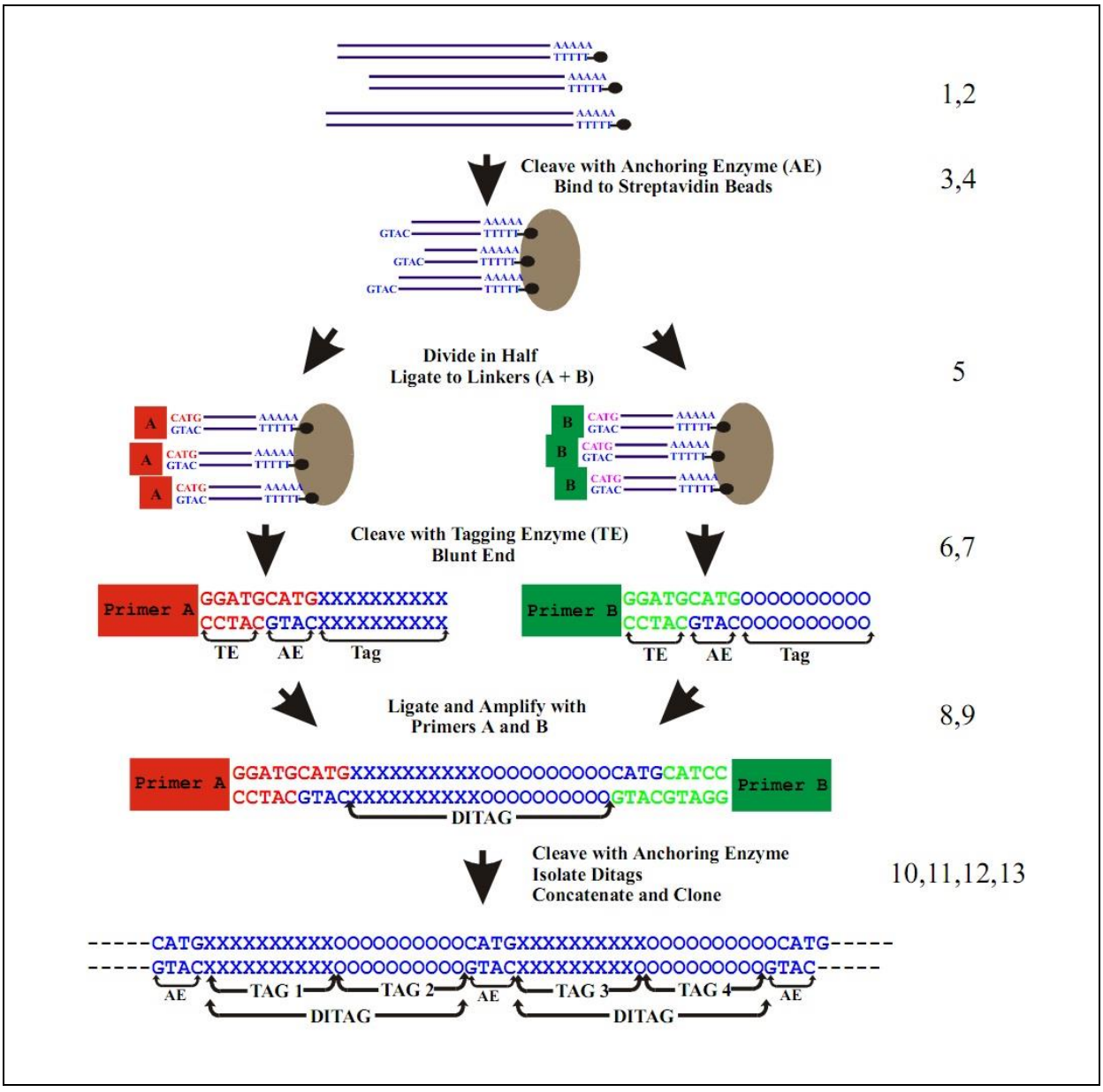


Figura 8. Representação esquemática da construção da biblioteca de SAGE. 1. Preparação do mRNA; 2. Síntese de cDNA; 3. Clivagem do cDNA "biotinilado"; 4. Ligação do cDNA "biotinilado" às esferas magnéticas; 5. Ligação dos adaptadores ao cDNA (to bound cDNA); 6. Liberação das tags de cDNA utilizando a enzima "Tagging"; 7. Liberação das extremidades cegas das tags (blunt ending release); 8. Ligação de tags para formar ditags; 9. Amplificação das ditags via PCR; 10. Isolamento das ditags; 11. Purificação das ditags; 12. Ligação das ditags para formação do concatâmero; 13. Clonagem e sequenciamento dos concatâmeros. Adaptado de Velculescu et al. (1999).

Adicionalmente, o uso do SAGE propicia uma maior sensibilidade à análise, permitindo o acesso a transcritos de baixa abundância, os quais muito provavelmente devem representar reguladores essenciais da expressão gênica, atuantes tanto na transcrição quanto na tradução, os quais são facilmente perdidos em outras abordagens metodológicas (SUN et al., 2004). Além disso, as *tags* de SAGE podem ser utilizadas como oligonucleotídeos (*primers*) ou como “sondas” na

identificação de genes desconhecidos em genomas (Figura 9). (COEMANS et al., 2005). Por outro lado, é importante salientar que a sensibilidade, precisão e potencial de identificação da SAGE são limitados pela quantidade de clones / *tags* sequenciados e pelo processo de *tag mapping*, ou seja, pela disponibilidade de sequências para a espécie em questão ou uma espécie próxima. Assim, embora o método não necessite do conhecimento prévio de sequências para geração das *tags*, outras metodologias devem ser agregadas à técnica para possibilitar a identificação de genes para os quais ainda não há informação disponível.

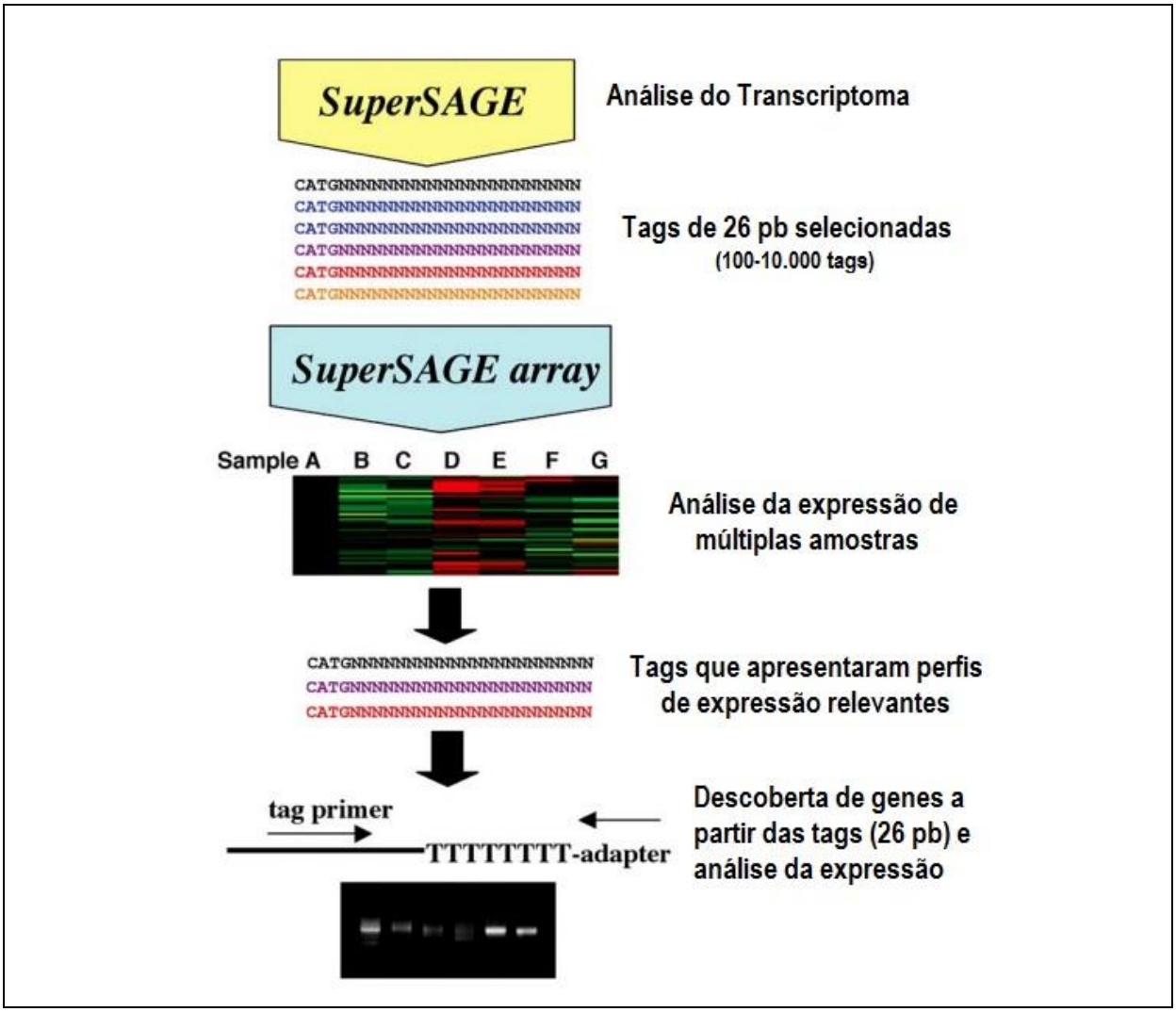


Figura 9. Fluxograma da análise de dados de SuperSAGE, visando à descoberta de genes-candidatos a partir de tags de 26 pb. Tags diferencialmente expressas são selecionadas e seu perfil de expressão é analisado. Um subgrupo de Tags, as quais apresentavam um perfil de expressão diferencial significativo (reposta rápida e/ou expressão tecido-específica) foi utilizado na descoberta de genes de interesse por metodologias específicas (ex. RACE-PCR), sendo sua expressão confirmada por RT-PCR.

Apesar de todas as vantagens e das aplicações que a técnica oferece, alguns autores ainda questionam o potencial de uso das *tags* geradas pela SAGE na identificação de genes únicos, principalmente para genes que pertencem a famílias gênicas, uma vez que nesse caso, uma única *tag* pode identificar mais de um gene (GE et al., 2006). Para aumentar a fidelidade do *tag mapping*, diversas modificações vêm sendo propostas e incorporadas à técnica, buscando aumentar a eficiência da clonagem e tamanho dos insertos. Saha et al. (2002) melhoraram o SAGE convencional utilizando a enzima de restrição *Mme*I, obtendo uma *tag* com 21pb, nomeando esta técnica de LongSAGE. Por sua vez, o SuperSAGE é baseado no uso de uma enzima de restrição tipo III para gerar *tags* maiores que 25 pb (MATSUMURA et al., 2003), produzidas tanto a partir da região 5', quanto da região 3' dos transcritos, aumentando assim a especificidade de identificação dos genes (WEI et al., 2004).

Sem dúvida, a aplicação mais importante da SAGE é a identificação da expressão diferencial de genes, tratando-se de um método efetivo para o reconhecimento de genes que são expressos em tecidos específicos em resposta a estresses ou doenças e, portanto, bastante indicado para estudos evolvendo vegetais superiores (MATSUMURA et al.,1999).

O primeiro trabalho empregando SAGE no estudo da expressão gênica em plantas foi realizado por MATSUMURA et al. (1999), em arroz (*O. sativa*) usando a técnica de SuperSAGE. Com base na eficiência desta metodologia, outros trabalhos com a mesma técnica foram conduzidos na identificação de genes envolvidos na resistência a doenças (MATSUMURA et al., 2003), bem como na caracterização do padrão de expressão gênica em diferentes tecidos tais como folhas, sementes (GIBBINGS et al., 2003) e plântulas (MATSUMURA et al., 1999) de arroz. Em leguminosas, análises de SuperSAGE para a identificação de genes relacionados a seca em raízes de grão-de-bico (*Cicer arietinum*), identificaram uma variedade de genes diferencialmente expressos, dentre eles diversas aquaporinas, que apresentavam membros superexpressos e reprimidos, relacionados a resistência/tolerância (MOLINA et al., 2008). Em soja (*Glycine max*) mais de 5.000 sequências únicas relacionadas a seca foram superexpressas (up-regulated) quando comparadas ao controle, estando primordialmente relacionadas como genes de resposta a hormônios, reposta aos estresses hídricos, salinos, e estresse oxidativo,

atuando na resistência/tolerância a seca nesta cultura (FERREIRA-NETO et al., 2012; *in press*).

Experimentos de SuperSAGE para o entendimento da interação planta-patógeno em soja, permitiram a identificação de mais de 100.000 tags únicas, sendo estas sequências caracterizadas como genes relacionados a regulação da transcrição, fotossíntese e resposta à estresse abiótico, envolvidos principalmente com mecanismos regulatórios e de sinalização ligados a resistência vegetal a doença, tratando-se de elementos importantes para programas de melhoramento (KIDO et al., 2012; *in press*).

2.9 BIOINFORMÁTICA: BANCO DE DADOS E FERRAMENTAS

Um dos aspectos decisivos para o grande avanço da genética é o uso intensivo de técnicas computacionais. A bioinformática surgiu na segunda metade da década de 80, quando ocorreu um expressivo aumento na quantidade de sequências geradas a partir dos primeiros projetos de sequenciamento em larga escala, exigindo recursos computacionais cada vez mais eficientes para o armazenamento e interpretação dos resultados obtidos (ROUZÉ, et al., 1999; PROSDOCINI et al., 2002). Esta ciência surgiu como resultado da união entre diversas linhas de conhecimento (Figura 10), entre elas, a engenharia de *softwares*, matemática, estatística, ciência da computação e a biologia molecular, com objetivo de gerar novos conhecimentos a partir do eficaz acesso e manuseio de grande quantidade de dados (SABBATINI, 1999; CARRARO; KITAJIMA, 2002).

A bioinformática, de um modo geral, surgiu como um novo ramo da ciência que tem por objetivo responder a questões biológicas através de análises com ferramentas computacionais. O potencial desta abordagem tem mudado a maneira de fazer ciência básica e aplicada, auxiliando pesquisas realizadas em laboratório (BAXEVANIS, 2001). Por meio das ferramentas de bioinformática é possível manipular uma grande diversidade de dados biológicos; sendo os programas e algoritmos desenvolvidos capazes de armazenar, processar, analisar, decifrar estruturas, traçar relações entre moléculas e vias e interpretar grande quantidade de informações (BORÉM; SANTOS, 2001).

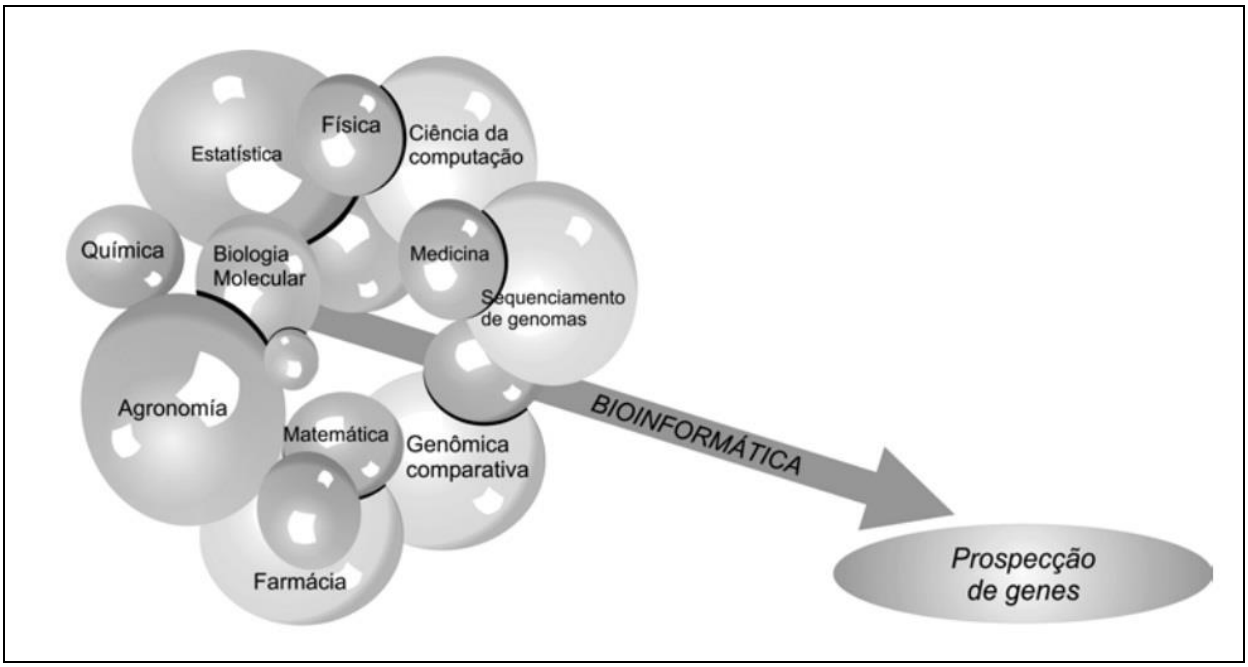


Figura 10. A bioinformática como uma ferramenta integradora da informação gerada por diversas ciências. Fonte: Malone et al. (2006).

Esta vasta quantidade de dados biológicos gerados levou ao desenvolvimento e evolução de uma diversidade de ferramentas computacionais que vêm facilitando a interpretação e análise de dados genômicos e pós-genômicos, permitindo a integração das informações relacionadas à transcriptômica, proteômica, metabolômica e outras "ômicas". Assim, a bioinformática tem ocupado papel significante na integração de vários dados gerados por estas diferentes tecnologias (EDWARDS; BATLEY, 2004).

Dada a grande quantidade de informações (sequências de nucleotídeos, aminoácidos e/ou proteínas) produzidas ao longo das últimas décadas, bem como a necessidade de acesso a tais dados, grandes bancos de dados públicos e privados, e de redes de acesso, têm sido disponibilizados à comunidade científica, permitindo interação entre diferentes grupos, bem como acesso e depósito contínuo de dados e ferramentas úteis para sua manipulação (MORAIS, 2003).

Um dos principais bancos de dados é o *GenBank* (Banco de Genes – sediado nos EUA) o qual armazena e disponibiliza todas as sequências publicamente disponíveis de DNA (de sequências pequenas a genomas inteiros), RNA e proteínas, estando acessível no NCBI (*National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica). Atualmente, este banco concentra

98.868.465 sequências, contabilizando mais de 99.116.431.942 bases, incluindo sequências de mais de 435.810 organismos (NCBI, 2011). Além de sequências primárias de genes e proteínas, o *GenBank* oferece informações sobre a taxonomia, genomas completos, mapas gênicos, informações sobre estruturas proteicas, etc., além de um amplo sistema de busca bibliográfica - o *PubMed,* um serviço da Biblioteca Nacional de Medicina (BENSON, et al., 2000).

Junto com o *Genbank*, fazem parte do INSD (*International Nucleotide Sequence Database*; Banco de Dados Internacional de Sequências de Nucleotídeos) o DDBJ (*DNA Database of Japan*; Banco de Dados de DNA do Japão) (http://www.ddbj.nig.ac.jp/) e o EMBL (*European Molecular Biology Laboratory*; Laboratório Europeu de Biologia Molecular; http://www.ebi.ac.uk/embl/) que também compartilham informações, além de manter um importante acordo de cooperação e de intercâmbio de dados com o NCBI (TATENO et al., 2002).

Entre os bancos que buscam informações acerca de funções proteicas está o InterPro (http://www.ebi.ac.uk/interpro), onde é feita uma combinação das assinaturas proteicas mais comuns encontradas nos bancos específicos, as quais são reunidas em um mesmo banco, possibilitando a caracterização única e não-redundante de famílias proteicas, domínios ou sítios funcionais (MULDER et al., 2002). Também estão integradas ao InterPro bases de dados específicas, como o Pfam (*Protein families database of alignments and HMMs*; Banco de dados de famílias proteicas de alinhamentos e HMMs), um banco de famílias de domínios (STUDHOLME et al., 2003); o PROSITE, um banco de domínios e famílias proteicas mais sensível na detecção de organismos divergentes (FALQUET et al., 2002), bem como o PRINTS, um banco que abriga uma coleção de assinaturas, motivos e domínios de famílias proteicas (ATTWOOD ET AL., 2003). Além desses, destacam-se o PDB (*Protein Data Bank*; Banco de Dados de Proteínas), o PIR (*Protein Information Resource*; Recursos de Informações Proteicas) e o KEGG (*Kyoto Encyclopedia of Genes and Genomes*; Enciclopédia de Genes e Genomas de Kyoto) que também mantêm um constante intercâmbio de dados com os demais bancos (PROSDOCINI et al., 2002; TATENO et al., 2002).

Paralelamente à criação dos bancos de dados, várias ferramentas que permitem o acesso e análise deste grande número de dados foram desenvolvidas pela bioinformática. Tais ferramentas permitem, por exemplo, uma análise de

predição *in silico* de sequências desconhecidas e novos genes, por meio de análises de alinhamentos comparativos (KENT et al., 2002).

A ferramenta mais comum de comparação de sequências de DNA com os bancos de dados genômicos é o BLAST (*Basic Local Alignment Search Tool;* Ferramenta de Busca por Alinhamento Local) (ALTSCHUL et al., 1990). Este algoritmo permite comparar uma sequência (DNA ou proteína) qualquer (denominada "*Query"*) com todas sequências genômicas de domínio público. O resultado desta busca (cuja pesquisa pode ser limitada a um dado organismo, por exemplo) retorna aquelas sequências (DNA ou proteínas) depositadas na base de dados ("*Subject*") com maior homologia.

O BLAST é subdividido em diferentes modalidades, dependendo do tipo de molécula de interesse e da resposta que se deseja nesta comparação. Na modalidade tBLASTx, por exemplo, tanto a "Query" como a base de dados "Subject" são sequências de nucleotídeos. Neste caso, são feitas as seis traduções possíveis de cada sequência de nucleotídeos. O resultado é o alinhamento do par proteína (Query) - proteína (Subject), uma vez que proteínas de dois organismos são mais parecidas entre si que os nucleotídeos que as codificam. No caso de tBLASTn, sequência de aminoácidos em um banco de dados de nucleotídeos, também traduzido nos seis quadros de leitura. Outras modalidades buscam homologia entre sequências de nucleotídeos (BLASTn), sequências de proteínas (BLASTp) ou entre sequências de nucleotídeos e proteínas (BLASTx) (ALTSCHUL et al., 1990; GIBAS; JAMBECK, 2001).

Outra ferramenta disponível no NCBI é o *ORF-Finder* (*Open Reading Frame Finder*; Identificador de Quadros Abertos de Leitura), uma ferramenta de análise gráfica, responsável pela tradução de sequências nucleotídicas (DNA/RNA) em todos os seis quadros abertos de leitura (NCBI, 2011).

Enquanto o BLAST está envolvido em análises locais de similaridade, o programa CLUSTAL é responsável por executar alinhamentos múltiplos biologicamente informativos, onde as sequências envolvidas, tanto de nucleotídeos como de aminoácidos, são analisadas em níveis globais de similaridade. Além disso, o programa permite a construção de cladogramas e fenogramas, para inferência filogenética e fenética, podendo ser visualizados, por exemplo, no programa *TreeView* (PAGE, 1996; THOMPSON et al., 1997), visto que – apesar de possuir o

algoritmo para gerá-los – o CLUSTALX não possui uma interface para sua visualização. O *TreeView* é um programa simples para visualização de dendrogramas, sendo capaz de ler diferentes formatos de arquivos, como *Nexus*, *Phylip*, *Nona*, *Mega* e *ClustalW/X*. O programa gera opções para visualizar as árvores de diferentes formas, como radialmente, em forquilha, em paralelo, etc. (PAGE, 1996; MORAIS, 2003).

A análise comparativa de sequências tornou-se essencial não só pela possibilidade de avaliar as similaridades entre sequências de vários organismos, como por proporcionar a reconstrução da história evolutiva das espécies e das famílias multigênicas, estimando as taxas da evolução molecular e inferindo sobre a natureza e a extensão das forças seletivas que atuaram no curso da evolução dos genes e dos genomas (KUMAR et al., 2004). Neste sentido, foi desenvolvido o programa MEGA *(Molecular Evolutionary Genetics Analysis*; Análises Genéticas de Evolução Molecular), um programa para análise de caracteres evolutivamente informativos, podendo ser usado para análises com sequências nucleotídicas, proteicas e marcadores moleculares (Tamura et al., 2007). Além disso, ele permite calcular matrizes de distância genética e analisar a composição de sequências nucleotídicas e proteicas. O programa também disponibiliza algoritmos como UPGMA (*Unweighted Pair Group Method with Arithmetic Means*; Método não polarizado de Agrupamentos aos Pares com Médias Aritméticas) (SNEATH; SOKAL, 1973), NJ (*Neighbor-Joinning*; Agrupamento por Vizinhança) (SAITOU; NEI, 1987) e máxima parcimônia (ECK; DAYHOFF, 1966; FITCH, 1971), permitindo a realização de inferências filogenéticas e fenéticas através da construção de dendrogramas (KUMAR et al., 2004).

O CLUSTER, amplamente utilizado na avaliação do padrão de expressão de genes, transformou-se em uma ferramenta importante para investigar como e onde o organismo responde às condições adversas do ambiente. Através desta ferramenta, sequências genômicas e dados gerados por experimentos de Microarray, SAGE, EST, entre outros, podem ser avaliados, incluindo a geração de mapas, agrupamento de médias K (*K-Means Clustering*) e clusterizações hierárquicas (EISEN et al., 1998). Permite também a visualização dos padrões de expressão relativa dos genes estudados nos diferentes tecidos e condições testadas.

2.10 FEIJÃO-CAUPI

2.10.1 CLASSIFICAÇÃO, ORIGEM E DISTRIBUIÇÃO

O feijão-caupi [*Vigna unguiculata* (L.) Walp.] é uma dicotiledônea (Magnoliopsida) que pertence à ordem Fabales, família Fabaceae (leguminosas), subfamília Faboideae, tribo Phaseoleae, subtribo Phaseolinea e gênero *Vigna* (APG III, 2009). Trata-se de uma cultura autógama, possuidora de um dos menores genomas deste grupo (450-500 Mb), com nível diploide de 2n=22 cromossomos (BENKO-ISEPPON, 2001; TEÓFILO et al., 2001).

O feijão-caupi é uma planta herbácea anual, propagada por sementes, de crescimento morfológico diversificado, apresentando variação de porte, desde ereto até prostrado e hábito de crescimento determinado a indeterminado (OLIVEIRA; CARVALHO, 1998). As flores, completas, têm os órgãos masculinos e femininos bem protegidos pelas pétalas, em número de cinco, de coloração branca, amarela ou violeta (TEÓFILO et al., 2001).

O gênero *Vigna* compreende 170 espécies, das quais 120 ocorrem na África (66 espécies endêmicas), 22 na Índia e sudeste da Ásia (16 espécies endêmicas), havendo poucas espécies descritas para as Américas (FARIS, 1965; GHAFFOR et al., 2001). Análises deste gênero quanto à variação morfológica, citogenética, molecular e fitogeográfica situam o feijão-caupi entre as espécies mais primitivas ocorrentes na África (STEELE E MEHRA, 1980; NG; MARÉCHAL, 1985). Contudo, através de análises moleculares Simon et al. (2007) sugerem que o gênero tenha se originado no sudoeste da Ásia (reconhecido berço das subtribos Phaseolinae, Galegeae, Loteae, Trifolieae e Vicieae), enquanto o sul da África representaria sua região de diversificação.

Na América latina, o feijão-caupi foi provavelmente trazido da Europa e do Oeste da África pelos colonizadores europeus e pelos escravos africanos durante os séculos 16 e 17. O processo ocorreu primeiramente nas colônias espanholas e logo após no Brasil, possivelmente no estado da Bahia e posteriormente para outras regiões do nordeste brasileiro (FREIRE-FILHO et al., 1988; 1999). Por outro lado, Simon et al. (2007) sugerem que mais de eventos de introdução de diferentes

regiões da África devam ter ocorrido, especialmente considerando que houve migração de escravos africanos de diferentes regiões para nosso país.

2.10.2 IMPORTÂNCIA ECONÔMICA

O feijão-caupi (também conhecido como feijão macassar, fradinho, ou de corda) constitui uma das principais leguminosas cultivadas no Brasil e no mundo. Estima-se que ocupe uma área mundial de, aproximadamente, 14 milhões de ha, sendo 9.3 milhões na África e o restante distribuído na América do Sul, América Central e Ásia, com pequenas áreas espalhadas pelo sudoeste da Europa, sudoeste dos Estados Unidos e Oceania. Os principais produtores mundiais são Nigéria, Niger e o Brasil (LANGYINTUO et al., 2003; SINGH et al., 2003). No Brasil, é cultivado predominantemente nas regiões Norte e Nordeste, onde é utilizado para fins alimentares (subsistência ou comercial), representando 95 a 100% do total das áreas ocupadas nessas regiões (LIMA et al., 2007).

Esta leguminosa possui um elevado valor nutricional, sendo fonte de proteínas (média de 23-25%), apresentando todos os aminoácidos essenciais, carboidratos (média de 62%), vitaminas, minerais, além de possuir quantidade significativa de fibras dietéticas, baixa quantidade de gordura (teor de óleo em torno de 2%). Por não apresentar colesterol, o feijão-caupi pode ser utilizado na substituição ou complementação dos alimentos derivados dos animais, tornando-a extremamente importante para populações de baixa renda, onde muitas pessoas não têm acesso a outras fontes de proteínas (BRESSANI; ELIAS, 1980; PRINYAWIWATKUL et al., 1996; MAIA et al., 2000; FROTA et al., 2009).

O sucesso de seu plantio se deve à sua considerável tolerância à seca e ao calor, à alta salinidade nos solos, bem como a temperaturas elevadas, além de proporcionar uma excelente cobertura vegetal, inibindo o crescimento de ervas daninhas, ajudando no combate a erosão e contribuindo para a melhoria da fertilidade do solo, aumentando a disponibilidade de fósforo e contribuindo para a fixação de nitrogênio, através da associação com bactérias fixadoras de N_2 atmosférico (SILVEIRA et al., 2001; VALENZUELA; SMITH, 2002). Assim, por sua capacidade de fertilizar naturalmente o solo, suportar condições ambientais

desfavoráveis e impedir a infecção e reprodução de organismos oportunistas, o feijão-caupi destaca-se como uma cultura de excelência para a agricultura (FERY, 1990; EHLERS; HALL, 1997).

Em estimativas realizadas, a produção média anual do feijão-caupi no Brasil (terceiro produtor mundial desta leguminosa) foi superior a 420.000 ton., com produtividade de 366 Kg/ha (SILVA, 2009), perfazendo 32% da área plantada com algum tipo de feijão em nível nacional (FREIRE-FILHO et al., 1999; 2005). Trata-se do único feijão capaz de sobreviver com sucesso na região norte (alta umidade, muita chuva e solo argiloso) e no Nordeste (seca, solo arenoso, por vezes salino e muito sol) (BARRETO, 1999). Com relação ao aspecto socioeconômico, a cultura do feijão-caupi representa a geração de quase 1.500.000 empregos/ano no Brasil, com o valor de produção estimado em US$ 250.000.000/ano (IBGE, 2010).

Apesar disso, a cultura do feijão-caupi está muito aquém do seu potencial produtivo, principalmente em função de fatores de estresses. Dentre os fatores que limitam o desenvolvimento do feijão-caupi pode-se citar a deficiência hídrica, causada pela escassez e irregularidade das chuvas; a salinização dos solos; plantios predominantemente de subsistência; utilização de cultivares com potencial genético reduzido, além da grande incidência de doenças e pragas (CASTRO, 2000).

2.10.3 MELHORAMENTO GENÉTICO DO FEIJÃO-CAUPI

Durante muito tempo os programas de melhoramento do feijão-caupi baseavam-se apenas em métodos tradicionais de cruzamento, com seleção de genótipos adaptados às condições específicas de cada região (XAVIER et al., 2005). Neste sentido, esforços foram direcionados para a geração de cruzamentos visando à obtenção de caracteres como aumento do tamanho dos grãos, aumento da produtividade média, porte ereto das plantas, floração precoce, bem como para a identificação de linhagens resistentes à salinidade e às doenças que mais prejudicam a produção do feijão-caupi (BARRETO, 1999; FREIRE-FILHO et al., 1999). De fato, a resistência múltipla a doenças e pragas, melhoria do porte, da tolerância ao calor, ao frio, à salinidade e à seca não se limita a um problema local,

ou regional, mas sim, objetos de estudo de programas de melhoramento em nível mundial (TIMKO; SINGH, 2008).

No Brasil, os primeiros trabalhos visando ao melhoramento do feijão-caupi foram iniciados na década de 1960 e tinham como objetivo o aumento da produtividade da cultura (PAIVA et al., 1970). Nos últimos anos cresceram as perspectivas de colheita e melhoria da qualidade do feijão-caupi através de programas de melhoramento que visam o desenvolvimento de cultivares com alta qualidade de grãos, alto potencial produtivo, adaptado aos sistemas de cultivo e com elevada qualidade nutricional (EHLERS; HALL, 1997).

O uso de ferramentas biotecnológicas modernas pode proporcionar não só condições de competitividade e características que atendam às necessidades comerciais internacionais do feijão-caupi (TIMKO, 2002), mas também informações sobre a estrutura e a composição do seu genoma e proteoma, auxiliando na interpretação de sua evolução e, definitivamente, contribuindo substancialmente para o melhoramento dessa cultura (SIMON et al., 2007; TIMKO et al., 2008).

Nos últimos anos o sequenciamento do genoma e do transcriptoma do feijão-caupi foi desenvolvido por alguns grupos e disponibilizado em bancos de dados públicos, como o CGKB (*Cowpea Genomics Knowledge Base*; Base de Conhecimentos Genômicos de Feijão-Caupi), um banco de dados de sequências genômicas baseado em informações derivadas de 298.848 sequências ricas em genes (GSS - *Cowpea Genespace Sequences*), geradas através da filtragem de DNA genômico metilado (CHEN et al., 2007). Outro grande banco de dados de feijão-caupi é o HarvEST, um banco de dados internacional, com mais de 180.000 ESTs geradas a partir de 17 bibliotecas oriundas de diversos tecidos, estando disponível há mais de cinco anos para acesso público (HARVEST, 2010).

Outro banco de dados de feijão-caupi teve início em 2004, através da Rede Nordestina de Biotecnologia (ReNorBio) e sob coordenação da Universidade Federal de Pernambuco. Nesta rede, foi realizada a análise genômica funcional, estrutural e comparativa em *V. unguiculata*, o projeto NordEST, o qual teve como objetivo identificar genes candidatos potencialmente úteis para fins de melhoramento desta cultura. O projeto integra 10 instituições e 11 Laboratórios, onde foram obtidas 27.453 sequências de EST e mais de 21 milhões de sequências de SuperSAGE relacionadas à expressão de genes de resistência a fatores bióticos e abióticos

importantes para o melhoramento do feijão-caupi (BENKO-ISEPPON et al., 2010). O banco NordEST representa um dos mais elaborados para a espécie, com 12 bibliotecas de EST sob condições de estresse biótico (vírus do mosaico severo e potyvirus) e abiótico (salinidade), capacitando o acesso à informações transcricionais de uma diversidade de genes em vários momentos de resposta vegetal frente as situações contrastantes aplicadas experimentalmente (BENKO-ISEPPON, 2009).

3. OBJETIVOS

3.1 Objetivo Geral

Identificar, quantificar e caracterizar as estruturas das aquaporinas e de seus genes codificantes, bem como analisar o padrão de expressão desses genes no feijão-caupi (*Vigna unguiculata*), concentrando-se em elementos com potencial filogenético e biotecnológico.

3.2 Objetivos Específicos

- Identificar e caracterizar genes relacionados ao transporte de água no banco de dados NordEST, reconhecer seus domínios conservados, comparando-os àqueles de genes anteriormente caracterizados;
- Efetuar alinhamentos múltiplos amplos e parciais, inferindo sobre a estrutura e as relações dos citados genes ao nível de nucleotídeos e de proteínas;
- Avaliar regiões filogeneticamente informativas com programas de análise filogenética;
- Estabelecer um perfil da expressão *in silico* dos genes analisados, a partir da análise de sua presença/ausência nos diferentes tecidos e condições de isolamento efetuados na montagem do banco de ESTs da cultura;
- Realizar modelagem comparativa para os diferentes membros identificados, visando à obtenção de modelos tridimensionais viáveis para estudos relacionados a especificidade e eficiência de transporte;

4. MATERIAIS E MÉTODOS

4.1 Identificação e Análise das Aquaporina de Feijão-caupi

A triagem de genes codificantes para aquaporinas em feijão-caupi foi realizada, inicialmente, no banco NordEST (http://bioinfo03.ibi.unicamp.br/vigna/) - um banco de dados do transcriptoma do feijão-caupi - utilizando modelos probabilísticos baseados no Modelo Oculto de Markov (EDDY, 1996), através do pacote de ferramentas HMMER3, inferência que permite maior eficiência na identificação de homólogos pela geração de modelos a partir de alinhamentos múltiplos de proteínas (EDDY, 2009). Para o desenho do modelo foram utilizadas sequências proteicas completas das 35 aquaporinas descritas em *A. thaliana* (JOHANSON et al.,2001). O alinhamento foi realizado para as sequências candidatas, através da ferramenta hmmalign, incluída no HMMER. Com base neste alinhamento foi realizada a construção do modelo, com auxílio da ferramenta hmmbuild, também incluída no pacote, seguindo-se pela busca de sequências ortólogas via hmmsearch, em um banco de dados local, traduzido nos seis quadros (*frames*) de leitura, correspondente a todos os transcritos de *Vigna unguiculata* disponíveis, utilizando um ponto de corte "*cut-off"* menor ou igual a e^{-05}.

As sequências anotadas foram então confrontadas com o banco de proteínas não redundantes do NCBI (*National Center for Biotechnology Information* – www.ncbi.nlm.nih.gov) através de alinhamentos recíprocos com a ferramenta BLASTx (ALTSCHUL et al., 1990). Os candidatos obtidos no banco de dados NordEST foram traduzidos no quadro de leitura indicado pelo BLASTx, por meio do programa ORF-finder (disponível em http://www.ncbi.nlm.nih.gov/projects/gorf/). Tais polipeptídios preditos foram, então, submetidos à ferramenta BLASTp, para a identificação de proteínas homólogas com função conhecida e, em paralelo, foi avaliada a integridade de seus domínios, via ferramenta CD-search, vinculada ao algoritmo rpsBLAST.

Levando em consideração integridade do domínio conservado MIP, um alinhamento múltiplo para as sequências que apresentassem domínios completamente íntegros foi conduzido por meio do ClustalW (THOMPSON et al.,

1994). Nesta análise foi possível observar a integridade dos motivos funcionais NPA, a integridade e posição dos resíduos de aminoácidos responsáveis pela formação do filtro seletivo (ar/R) das aquaporinas, e as posições de Froger's, responsáveis pela distinção entre aquaporinas e aquagliceroporinas.

Para determinar a nomenclatura das aquaporinas de feijão-caupi e analisar a relação das diferentes aquaporinas nas subfamílias propostas para o grupo, foram construídos dendrogramas, utilizando o método de *Neighbor-Joining*, com valores de *bootstrap* ao nível de 1000 replicações, com a utilização do programa MEGA4 (TAMURA et al., 2007) e visualização via TreeView (PAGE, 1996). Como regra, a nomenclatura das aquaporinas identificadas em feijão-caupi foi baseada na homologia de sequências e análises filogenéticas, de acordo com a nomenclatura já utilizada em outros organismos vegetais (JOHANSON et al., 2001). Além de sequências proteicas de feijão-caupi, sequências de aquaporinas completas descritas e caracterizadas em *Arabidopsis*, *Z. mays*, *Ricinus communis* e *Vitis vinifera* foram utilizadas na construção dos dendrogramas, a fim de verificar a relação filogenética entre as diferentes subfamílias de aquaporinas vegetais.

Através da contagem direta das *reads (*aquaporinas em feijão-caupi), foi traçado um perfil de expressão das sequências obtidas na plataforma NordEST. Tais *reads* foram agrupadas segundo a biblioteca/tecido de expressão, após a normalização dos dados, ou seja, considerando-se o número de *reads* em cada biblioteca e sua frequência relativa nas diferentes bibliotecas. A análise do perfil de expressão das aquaporinas em *V. unguiculata* foi realizada através do programa CLUSTER 3.0 (EISEN et al., 1998), sendo a clusterização hierárquica gerada visualizada no programa TreeView (PAGE, 1996). Em tal visualização gráfica, a cor amarela indica a ausência de expressão e os níveis de laranja, a intensidade de expressão na biblioteca em questão, objetivando a identificação dos principais sítios de expressão das aquaporinas no feijão-caupi.

Para elucidar a estrutura e função destes genes, as sequências de aminoácidos das aquaporinas foram submetidas à ferramenta WoLF PSORT (http://wolfpsort.org/) para a determinação de sua localização subcelular, além de determinar o número de segmentos transmembrana para cada membro de aquaporina em feijão-caupi, através do programa Philius (http://www.yeastrc.org/philius; REYNOLDS et al., 2008)

4.2 Modelagem Comparativa e Estrutura Tridimensional

As sequências de aminoácidos preditas para aquaporinas em feijão-caupi foram submetidas a uma modelagem comparativa, visando à determinação de sua estrutura terciária, representando a primeira modelagem para os membros desta família de transportadores nesta espécie.

Como modelo (*template)* para a modelagem tridimensional foi utilizada a aquaporina SoPIP2-1 [1z98] de espinafre (*S. oleracea*) depositada no PDB (*Protein Data Bank*) e obtida por difração de raios-X (TÖRNROTH-HORSEFIELD et al., 2006), uma vez que se trata da única aquaporina vegetal disponível no banco, além de apresentar a maior similaridade com as aquaporinas preditas para feijão-caupi.

A modelagem foi executada no programa MODELLER 9v9 (ESWAR et al., 2008), através da rotina *model*, que utiliza para o cálculo, as coordenadas atômicas do modelo, o alinhamento gerado entre o modelo e as sequências submetidas à comparação, em conjunto com as coordenadas atômicas do molde. As estruturas geradas foram visualizadas e os diversos modos de visualização gerados no programa NOC 3.01 (CHEN et al., 2007).

A avaliação de qualidade foi determinada pela ferramenta *online Protein Structure & Model Assessment,* da plataforma SWISS-MODEL Workspace (http://swissmodel.expasy.org/), onde foram obtidos os dados para a avaliação via Procheck (LASKOWSKI et al., 1993), que determina a qualidade estereoquímica do modelo através da análise do gráfico de Ramachandran e QMEAN (BENKERT et al., 2009), o qual compara o modelo com estruturas depositadas de mesmo tamanho. Informações acerca da qualidade do modelo também foram 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; WIEDRSTEIN; SIPPL, 2007). Por fim, as sequências de aminoácidos também foram submetidas a uma eletroforese bidimensional *in silico* via ferramenta *web* JVirGel (http://www.jvirgel.de/), que determina os prováveis pesos moleculares e pontos isoelétricos das sequências proteicas analisadas (HILLER et al., 2006). A hieraquia das análises computacionais encontra-se descrita abaixo na figura 11.

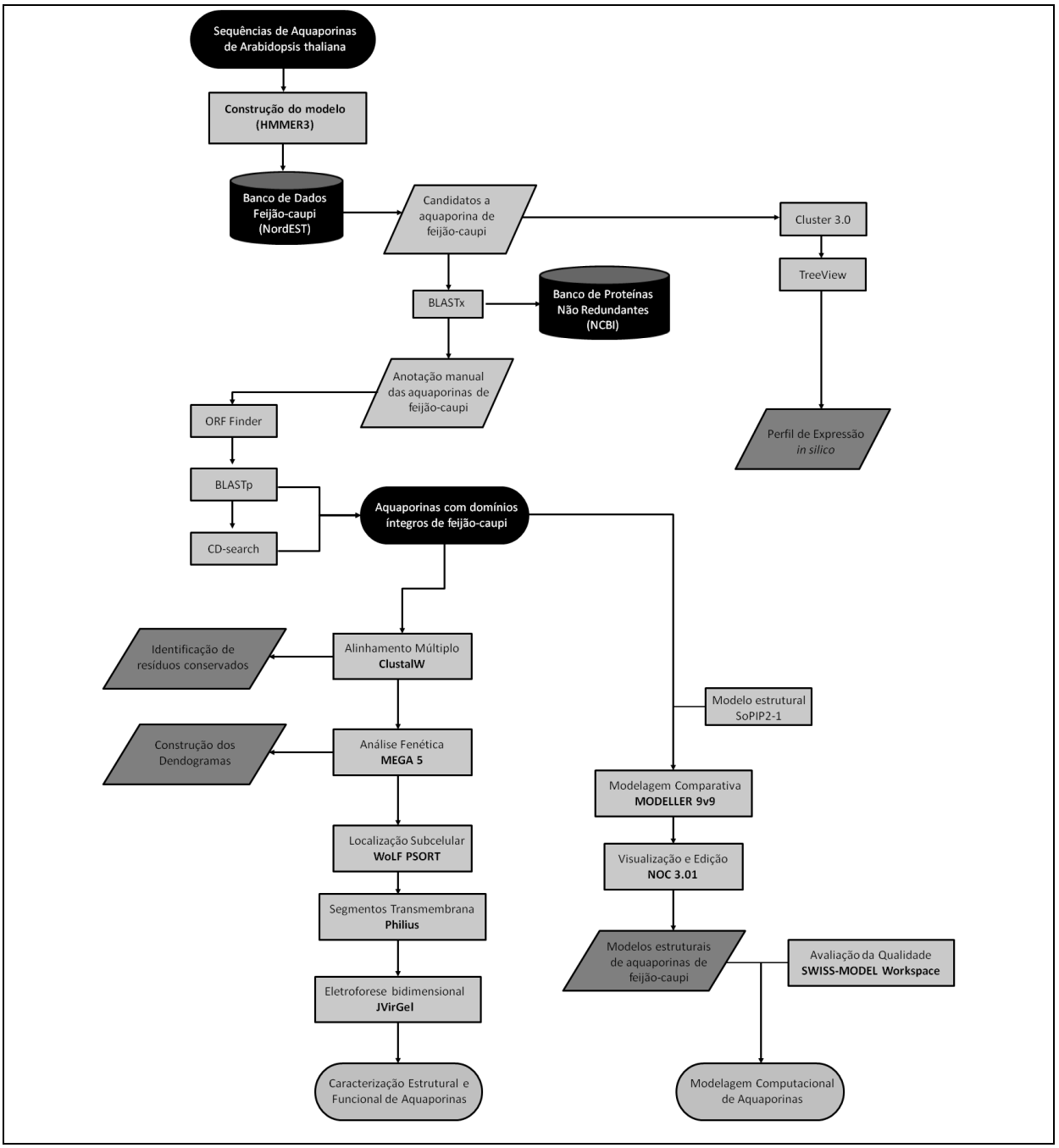


Figura 11. Diagrama do processo geral de prospecção, caracterização e modelagem de aquaporinas no genoma expresso de feijão-caupi via ferramentas de bioinformática.

4.3 Validação dos Transcritos por Clonagem e Sequenciamento

4.3.1 Desenho de *Primers*

Realizada a caracterização *in silico* das sequências candidatas, as mesmas foram sequenciadas para confirmação como sendo aquaporinas. Para tal, pares de

primers foram desenhados com base em 15 sequências (ESTs) de prováveis aquaporinas. As sequências selecionadas para desenho de *primers* foram aquelas com os valores de expressão mais significativos e que contemplassem membros de todas as subfamílias (Tabela 1). Os *primers* foram desenhados através da ferramenta web Primer3 (http://frodo.wi.mit.edu/primer3/; ROZEN; SKALETSKY, 2000) levando-se em consideração os seguintes critérios: tamanho do oligonucleotídeo (mínimo: 18 bases; ótimo: 20 bases; máximo: 22 bases); temperatura de anelamento (mínimo: 57º C; ótimo: 60º C; máximo 63º C); e conteúdo em GC (mínimo: 40%; ótimo: 50%; máximo: 60%).

4.3.2 Obtenção do cDNA e clonagem das sequências relacionadas a aquaporinas

4.3.2.1 Obtenção do material vegetal

Sementes de genótipos contrastantes de feijão-caupi, com características de tolerância (Pitiúba) e suscetibilidade (Canapu Amarelo) ao estresse salino e osmótico, foram semeadas em vasos contendo areia lavada (cinco sementes por vaso), utilizando-se três vasos por tratamento (Controle, 2 horas e 8 horas após estresse salino). A rega inicial foi constituída de água (primeira semana), sendo que após o início da germinação foram realizadas regas com solução nutritiva de Hoagland três vezes por semana com ½ concentração.

Cerca de um mês após a semeadura foi aplicado tratamento com 200 mL de NaCl 100mM Os vasos controles foram regados com o mesmo volume de solução nutritiva desprovida de NaCl. Os tempos de coleta (2 e 8 horas) foram escolhidos considerando-se dados da literatura, especialmente estudos realizados com *A. thaliana* que demonstram um pico de ativação entre duas a três horas da introdução do estresse salino, bem como outro pico de ativação (de um segundo grupo de genes) oito horas após o estresse (SEKI et al., 2002). Nos tempos estabelecidos as raízes das plantas tratadas foram separadas da parte aérea, rapidamente lavadas para a retirada da areia e imediatamente colocadas em nitrogênio líquido. Ao final da

coleta as amostras foram armazenadas em ultra-freezer (-80°C) até serem submetidas à extração do RNA.

Tabela 1: **Conjunto de *primers* selecionados para a validação das aquaporinas de feijão-caupi. Fw: foward; Rv: reverse.**

Gene	Sentido	Primers	Tamanho Primer	Tamanho Fragmento
VuPIP11	Fw	5'-TGTTCCACAGTTGGTGTTCAAG-3'	22	143
	Rv	5'-AGAGATAACTTGCGTGCCAAGA-3'	22	
VuPIP13	Fw	5'-CCCTCAGAGCTTACTTCATGGT-3'	22	151
	Rv	5'-AAGCGATTCCTTGAATACCAAC-3'	22	
VuPIP14	Fw	5'-CGCTAGAGACTCACACGTTCCT-3'	22	151
	Rv	5'-GCCAAGGTCTTTGTTGAAGATG-3'	22	
VuPIP21	Fw	5'-AAGGACTACCATGACCCTCCTC-3'	22	150
	Rv	5'-CTGGCTCTTGTAACCAATGACA-3'	22	
VuPIP23	Fw	5'-ACCGAGGTACGAACTGAGTTGA-3'	22	148
	Rv	5'-TGGTGACGTAGAGGAAGAGGAG-3'	22	
VuPIP24	Fw	5'-ATAAGGGCGTTGTTGTACATGG-3'	22	160
	Rv	5'-TAATCTCAGCACCCAAAGCAGT-3'	22	
VuTIP11	Fw	5'-GTTAGCTGGACCTGGGACAAC-3'	21	130
	Rv	5'-AGTAGTCAGTGGTTGGGAGCTG-3'	22	
VuTIP23	Fw	5'-CCCTCCTAACCGGTTTCTTCTA-3'	22	149
	Rv	5'-AATCTCAAACACCACACCTTGG-3'	22	
VuTIP21	Fw	5'-TTTAGCAGCAGGACCATTCTCT-3'	22	151
	Rv	5'-CACATTGCCATAGATGAGACCA-3'	22	
VuTIP22	Fw	5'-TACATTGCCGAGTTCATCTCAA-3'	22	152
	Rv	5'-ACAGCAACGAAGAGAGCAAAAC-3'	22	
VuNIP11	Fw	5'-TCCGACTCCTTTGTCTCTGTTC-3'	22	149
	Rv	5'-CAAACAATTGCTATCCCAGGAA-3'	22	
VuNIP12	Fw	5'-GCAGTTGGATCTACGGTACTGC-3'	22	147
	Rv	5'-TCCAAGAGTTGTTGACACCAGA-3'	22	
VuNIP71	Fw	5'-CTCGTGATGCTCTTTCAGGACT-3'	22	154
	Rv	5'-GCAAGCTTTGACACTCTGCTCT-3'	22	
VuSIP11	Fw	5'-CTCATCATGATCAAGGGTCCTC-3'	22	156
	Rv	5'-GTCGTGCCATTTCTCAACGTAT-3'	22	
VuSIP21	Fw	5'-CGCTTCAACAGAGGTAGCACTT-3'	22	155
	Rv	5'-AAACTGTGGTTGAAGGGTTGAA-3'	22	

4.3.2.2 Extração do RNA total e síntese do cDNA

Aproximadamente 200 mg de material vegetal (raízes de feijão-caupi de ambos os genótipos, e em todos os tratamentos) foram usados para a extração de RNA total através do kit “SV Total RNA Isolation System” (Promega) de acordo com recomendações do fabricante. Após tratamento com DNase, 1 µg de RNA foi utilizado para a síntese de cDNA utilizando o kit “QuantiTect® Reverse Transcription Kit” (Qiagen).

4.3.2.3 Obtenção dos produtos da PCR

As amostras de cDNAs foram utilizadas para a amplificação dos segmentos relacionados a aquaporinas em feijão-caupi, utilizando os *primers* desenhados com base nas sequências de EST descritas previamente (Tabela 1). As condições da PCR foram as seguintes: 50 ng de cDNA, 3,5 µL de MgCl2 (25 mM), 1,6 µL de dNTPs (10 mM), 1 µL de cada primer (10 mM), 1.0 U Taq polimerase (Fermentas), 2,5 µL de solução tampão 10x (500 mM KCl, 100 mM Tris-HCl, pH8.3 e 1,4 mM MgCl2) e água Milli-Q num volume final de 25 µL. A reação foi programada para: 10 min a 95°C (desnaturação inicial), seguido de 38 ciclos de 1 min a 95°C (desnaturação), 1 min a 54°C (anelamento) e 1 min a 72°C (extensão) e finalizando com uma extensão final de 10 min a 72°C. As reações foram realizadas em termociclador “Techne TC-412” (Barloworld Scientific). Os produtos da PCR foram submetidos à eletroforese em gel de agarose a 1,8 % a 4 V/cm por 3 horas, usando como padrão de peso molecular 100 pb DNA Ladder (Promega). Com auxílio de um bisturi, fragmentos entre 300 e 500 pb foram excisados do gel, purificados de acordo com o kit ”GFX PCR DNA e Gel Band Purification Kit” (GE Healthcare, USA).Os fragmentos foram quantificados através de gel de agarose 1,2 % comparando-se com o padrão de peso molecular DNA de fago *lambda* (λ-DNA) em diferentes concentrações (20, 50 e 100 ηg/µL).

4.3.2.4. Clonagem e Sequenciamento

A clonagem dos fragmentos foi realizada em pGEM, através do Kit “pGEM-T Easy vector System I” (Promega). Aproximadamente 150-300 ng de cada fragmento

purificado foi utilizado para ligação em 50 ng de vetor pGEM-T-easy. Uma alíquotas de 4 μL da mistura de cada reação de ligação (fragmento + vetor) foi utilizada para transformação em 100 µL de células quimiocompetentes “MAX Efficiency DH10B™ Competent Cells (Invitrogen) por choque térmico (incubação da reação no gelo, por 30 min, choque térmico a 42°C por 45 s, seguido por incubação em gelo novamente por um período de 2-5 min). Realizada a transformação, foram adicionados 500 µL de meio de cultura S.O.C (triptona 2 %; extrato de levedura 0,5 %; NaCl 5M; MgCl2 1 M; KCl 1 M e MgSO4 1 M) à suspensão de células, que foram então incubadas a 37ºC, sob agitação (240 rpm), durante 1 h. Uma alíquota de 150 µL da suspensão de células transformantes foi plaqueada em placas de Petri contendo meio LB sólido, suplementado com ampicilina 1,2 % (100 mg/L). Aproximadamente 30 min antes do plaqueamento, foram adicionados ao meio de cultura sólido 100 µL de X-Gal (20 mg/ mL) e 40 µL de IPTG (100 mM). As culturas foram incubadas a 37°C durante 16 h, sendo posteriormente transferidas para 4°C para melhor identificação das colônias recombinantes. A seleção das colônias recombinantes foi realizada a partir da diferenciação de sua cor na presença do IPTG (isopropil tiogalactosídeo) e X-GAL (5-bromo-4-cloro-3-indol-beta-D-galactopiranosídeo

Os clones positivos (recombinantes) foram multiplicados em microplacas de cultivo de 96 poços, contendo 1 mL de meio de cultura LB suplementado com ampicilina 1,2 % (100 mg/ L), e incubadas a 37°C, 320 rpm, por 22 h. Alíquotas de 100 µL de cada cultura foram transferidas para uma nova microplaca contendo 100 µL de glicerol 50 % e armazenadas a -80°C. O restante da cultura foi utilizado para a extração do DNA plasmidial (Miniprep) através do método de lise alcalina para microplacas, conforme descrito em: http://sucest.lad.ic.unicamp.br/public. Após esse processo, os fragmentos de DNA obtidos foram quantificados em comparação com o marcador molecular pGEM (100 ng).

A reação de sequenciamento foi realizada com 150 ng de DNA plasmidial, 0,7 µL de Big-dye versão 3.1, 3,2 pmoles de primer T7, 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 H2O Milli-Q estéril em cada tubo. Após agitação suave por 15

s, as placas foram incubadas em gelo por 20 min, sendo em seguida centrifugadas por 30 min / 3.000g / 4ºC. Descartado o sobrenadante, os *pellets* 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).

4.3.3 Análise das sequências

As sequências obtidas foram inicialmente analisadas via BioEdit 7 (HALL, 1999), para eliminação dos segmentos de baixa qualidade e obtidas as sequências de nucleotídeos a serem submetidas ao VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen) para retirada dos trechos de vetor. De posse das sequências tratadas, foi conduzido um alinhamento e montagem de “Contigs” ainda no BioEdit, obtendo-se assim os candidatos a validadores das aquaporinas obtidas via bioinformática no genoma expresso de feijão-caupi.

Após essa etapa, as sequências de nucleotídeos foram confrontadas com as sequências obtidas no Banco de Dados NordEST utilizando a ferramenta BLASTn do NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sendo o resultado submetido à ferramenta web Circoletto (DARZENTAS, 2010) para visualização de similaridades entre sequências.

5. RESULTADOS

O presente estudo apresenta a primeira descrição das “Proteínas Intrínsecas de Membrana” (MIPs - *Major Intrinsic Proteins*), também conhecidas de aquaporinas, no transcriptoma do feijão-caupi. Para este trabalho, após mineração no banco de dados NordEST, através de métodos de buscas em bioinformática (HMMER), foram identificados 37 candidatos a genes codificantes para aquaporinas (Tabela 2) . Das 37 sequências candidatas, somente dez (quatro PIPs, três NIPs, duas TIPs e uma SIP) deduzidas de aminoácidos apresentaram-se incompletas. As demais 27 candidatas apresentaram sequência completa de aminoácidos para aquaporinas de feijão-caupi, sendo filogeneticamente analisadas juntamente com os membros descritos para *A. thaliana.* Esta análise foi realizada tanto pelo método de *Neighbor-Joining* como pelo método de máxima parcimônia e de máxima verossimilhança, cujos resultados foram similares (dados não mostrados). Assim, por comparação das sequências de aminoácidos, estas aquaporinas foram classificadas em 11 PIPs (cinco PIP1s e seis PIP2s); 10 TIPs (quatro TIP3s e seis outras TIPs); quatro NIPs e duas SIPs, não sendo encontrado nenhum membro da subfamília XIP.

As sequências preditas para as subfamílias apresentaram um tamanho entre 190 e 302 aminoácidos, mas com variações internas a cada subtipo de transportador. Os 11 membros da subfamília PIP foram classificados em dois grupos. O *open reading frame* (ORF) das sequências completas dos transcritos de PIP1 codificavam para proteínas de 214-290 aminoácidos de comprimento, com 82-94% de identidade entre as sequências. O tamanho das ORFs de PIP2 variou de 279 a 288 aminoácidos e identidade variando entre 75 e 93%. Os agrupamentos PIP1 e PIP2 quando comparados entre si, mostraram uma identidade variando entre 19 e 78% e tamanhos estimados das ORFs predominantemente similares.

No caso da subfamília TIP, as sequências protéicas com domínios completos variaram entre 218 e 257 aminoácidos, mostrando 49 a 98 % de identidade. Nesta subfamília, o grupo TIP3 apresentou maior identidade entre seus membros, com valores superiores a 98 % (Tabela 2). A subfamília NIP apresentou quatro sequências entre 190 e 302 aminoácidos, com identidade mínima de 45 % e máxima de 79%. Já para a subfamília SIP, foi observado que os polipeptídeos eram

relativamente curtos, com 248 a 245 aminoácidos e com baixa identidade (45%) entre suas sequências.

Tabela 2. Categorias de aquaporinas identificadas no transcriptoma do feijão-caupi.

	Nome	No. acesso (NordEST)	Tamanho		Maior Similaridade (%)		Domínio Conservado	Estado*
			(nt)	(aa)	Caupi	NCBI		
PIPs	VuPIP1-1	Contig15975	858	213	VuPIP1-2 (85.98)	ACJ84965 (92.34)	MIP superfamily	CP
	VuPIP1-2	Contig640	600	213	VuPIP1-1 (85.98)	DAA33860 (90.46)	MIP superfamily	CP
	VuPIP1-3	Contig14532	656	289	VuPIP1-4 (94.48)	AAB72149 (95.86)	MIP superfamily	CP
	VuPIP1-4	Contig16565	492	289	VuPIP1-3 (94.48)	ACU20400 (97.58)	MIP superfamily	CP
	VuPIP1-5	Contig7208	858	287	VuPIP1-3 (85.81)	AAY22204 (98.26)	MIP superfamily	CP
	VuPIP2-1	Contig12096	750	285	VuPIP2-2 (93.21)	ACU20229 (96.84)	MIP superfamily	CP
	VuPIP2-2	Contig2737	864	287	VuPIP2-1 (93.21)	ABU94631 (97.21)	MIP superfamily	CP
	VuPIP2-3	Contig2789	180	284	VuPIP2-2 (81.60)	ABD32807 (90.37)	MIP superfamily	CP
	VuPIP2-4	Contig1057	771	285	VuPIP2-5 (89.24)	ACU20043 (97.20)	MIP superfamily	CP
	VuPIP2-5	Contig13160	747	287	VuPIP2-4 (89.24)	NP_001105639 (96.52)	MIP superfamily	CP
	VuPIP2-6	Contig2256	870	278	VuPIP2-4 (79.93)	BAD90701 (90.390	MIP superfamily	CP
	VuPIPa	Contig10897	906	200	VuPIP2-2 (100.00)	ACU23034 (94.77)	MIP superfamily	IC
	VuPIPb	Contig10949	735	164	VuPIP1-3 (90.65)	CAA06745 (90.57)	MIP superfamily	IC
	VuPIPc	Contig16439	642	126	VuPIP2-1 (100.00)	ACU20229 (96.45)	MIP superfamily	IC
	VuPIPd	Contig736	378	83	VuPIP1-3 (97.26)	AAB72149 (92.41)	MIP superfamily	IC
TIPs	VuTIP1-1	Contig70	870	250	VuNIP6-1 (80.00)	AAF82790 (94.67)	MIP superfamily	CP
	VuTIP1-2	Contig8535	634	252	VuTIP1-1 (75.10)	AAA02947 (94.07)	MIP superfamily	CP
	VuTIP2-1	Contig7680	819	248	VuTIP2-2 (87.20)	ACU14104 (95.97)	MIP superfamily	CP
	VuTIP2-2	Contig3059	192	248	VuTIP2-1 (87.20)	ACU23209 (85.19)	MIP superfamily	CP
	VuTIP2-3	Contig12252	837	249	VuNIP6-1 (80.00)	AAZ78659 (94.63)	MIP superfamily	CP
	VuTIP3-1	Contig109	864	219	VuTIP3-3 (98.60)	ACU18941 (84.46)	MIP superfamily	CP
	VuTIP3-2a	Contig140	855	256	VuTIP3-2b (98.83)	P23958 (94.92)	MIP superfamily	CP
	VuTIP3-2b	Contig7183	747	256	VuTIP3-2a (98.83)	P23958 (93.75)	MIP superfamily	CP
	VuTIP3-3	Contig348	744	254	VuTIP3-1 (98.60)	ACU18941 (84.40)	MIP superfamily	CP
	VuTIP5-1	Contig15230	765	244	VuNIP6-1 (80.00)	ACX37451 (62.30)	MIP superfamily	CP
	VuTIPa	Contig16579	344	208	VuNIP6-1 (80.00)	ACU18235 (91.96)	MIP superfamily	IC
	VuTIPb	Contig3907	735	114	VuTIP1-1 (92.73)	ADQ08684 (79.63)	MIP superfamily	IC
NIPs	VuNIP1-1	Contig16718	534	272	VuNIP1-2 (79.35)	ABS72446 (96.34)	MIP superfamily	CP
	VuNIP1-2	Contig7093	642	189	VuNIP1-1 (79.35)	ACU23597 (92.02)	MIP superfamily	CP
	VuNIP6-1	Contig15045	788	301	VuTIP1-1 (80.00)	AAT35231 (82.95)	MIP superfamily	CP
	VuNIP7-1	Contig6501	753	263	VuTIP5-1 (66.67)	ACU23509 (83.52)	MIP superfamily	CP
	VuNIPa	Contig137	570	60	VuNIP6-1 (100.00)	AAT35231 (80.88)	MIP superfamily	IC
	VuNIPb	Contig16969	771	64	VuNIP1-2 (81.36)	CAB45652 (80.77)	MIP superfamily	IC

	VuNIPc	Contig5471	864	178	VuTIP1-1 (80.00)	ACJ84850 (76.10)	MIP superfamily	IC
SIPs	VuSIP1-1	Contig3455	249	247	VuPIP2-4 (80.00)	ACU22985 (79.35)	MIP superfamily	CP
	VuSIP2-1	Contig4744	747	244	VuPIP2-4 (66.67)	ACU20408 (78.24)	MIP superfamily	CP
	VuSIPa	Contig14057	759	249	VuPIPa (80.00)	ACU23100 (90.73)	MIP superfamily	IC

*CP: Domínio completo; IC: Domínio incompleto

5.1 Avaliação de resíduos conservados

Informações a respeito das sequências estão sumarizadas na Tabela 2, onde são evidenciados tais motivos “NPAs”, resíduos conservados no filtro “ar/R” e nos sítios “*Froger’s*, juntamente com a predição dos domínios transmembrana e localização subcelular das aquaporinas de feijão-caupi.

Os *loops* B e E apresentam os motivos NPA (Asn-Pro-Ala), reconhecidos como formadores de uma das duas constrições do poro. Todos os membros da subfamília PIP apresentaram os dois motivos NPA conservados (não havendo substituição em nenhuma das posições). Para os membros da subfamília TIP, com exceção da sequência VuTIP3-1 [que no segundo motivo LE apresenta uma lisina (K) no lugar da asparagina (N)], todas as sequências apresentaram motivo íntegro nas duas posições. Como já estabelecido (CHAUMONT et al., 2001; GOMES et al., 2009), modificações nestes motivos são mais acentuadas para as subfamílias NIP e SIP. Para VuNIP1-1 e VuNIP6-1, o resíduo de alanina (A) foi substituído por valina (V), enquanto que para VuSIPs ocorreu uma modificação de alanina (A) por Treonina (T) no primeiro NPA de VuSIP1-1 e por serina (S) em VuSIP2-1 (Tabela 3)

A segunda constrição, também conhecida como constrição arginina/aromática (ar/R), tem sido relatada como sendo o maior delimitador de permeabilidade ao soluto. Alguns resíduos desta região são responsáveis pelo fornecimento de pontes de hidrogênio para as moléculas de água a serem transportadas (WALLACE; ROBERTS, 2004). Os filtros de seletividade identificados nos representantes VuPIPs, VuTIPs, VuNIPs e VuSIPs foram muito similares ao que foi descrito em outras espécies (LUDEWIG; DYNOWSKI, 2009). Análises de homologia mostraram que apenas oito membros das MIPs de feijão-caupi apresentaram filtros de seletividade ar/R não descritos previamente, mas que mesmo assim, compartilham

de homologia em algumas posições e resíduos com outros filtros identificados para as diferentes subfamílias (Tabela 4).

Tabela 3. Sumário das principais regiões e sítios conservados de aquaporinas em *Vigna*.

Name	Motivo NPA (LB)	Motivo NPA (LE)	Ar/R selectivity filter				Froger's Position					TM*
			H2	H5	LE1	LE2	P1	P2	P3	P4	P5	
VuPIP1-1	NPA	NPA	F	H	T	R	Q	S	T	F	W	5
VuPIP1-2	NPA	NPA	F	H	T	R	R	S	A	F	W	4
VuPIP1-3	NPA	NPA	F	H	T	R	E	S	A	F	W	5
VuPIP1-4	NPA	NPA	F	H	T	R	E	S	A	F	W	6
VuPIP1-5	NPA	NPA	F	H	T	R	E	S	A	F	W	6
VuPIP2-1	NPA	NPA	F	H	T	R	Q	S	A	F	W	6
VuPIP2-2	NPA	NPA	F	H	T	R	Q	S	A	F	W	6
VuPIP2-3	NPA	NPA	F	H	T	R	Q	S	A	F	W	6
VuPIP2-4	NPA	NPA	F	H	T	R	Q	S	A	F	W	6
VuPIP2-5	NPA	NPA	F	H	T	R	Q	S	A	Y	W	6
VuPIP2-6	NPA	NPA	F	H	T	R	M	S	A	F	W	6
VuTIP1-1	NPA	NPA	H	I	A	V	T	T	A	Y	W	7
VuTIP1-2	NPA	NPA	H	I	A	V	T	S	A	Y	W	6
VuTIP2-1	NPA	NPA	H	I	G	R	T	S	A	Y	W	7
VuTIP2-2	NPA	NPA	H	I	G	R	T	S	A	Y	W	6
VuTIP2-3	NPA	NPA	H	I	G	R	T	S	A	Y	W	6
VuTIP3-1	NPA	KPA	H	I	G	R	T	G	A	S	-	5
VuTIP3-2a	NPA	NPA	H	I	A	L	T	A	S	F	W	6
VuTIP3-2b	NPA	NPA	H	I	A	L	T	A	S	F	W	6
VuTIP3-3	NPA	NPA	H	I	G	R	T	A	A	F	W	6
VuTIP5-1	NPA	NPA	N	T	A	A	S	S	A	Y	W	6
VuNIP1-1	NPA	NPV	W	V	A	R	F	S	A	Y	I	5
VuNIP1-2	NPA	NPA	-	V	A	R	F	S	A	Y	L	4
VuNIP6-1	NPA	NPV	T	I	G	R	Y	T	A	Y	L	6
VuNIP7-1	NPA	NPA	A	S	G	R	F	T	A	Y	F	5
VuSIP1-1	NPT	NPA	I	T	P	F	M	A	A	Y	M	6
VuSIP2-1	NPS	NPA	N	T	P	N	M	A	A	Y	M	6

(*) TM: Número de hélices transmembrana preditos via Philius.

A subfamília PIP em feijão-caupi destacou-se como a mais conservada, apresentando o filtro ar/R conservado, com os resíduos Phe (F), His (H), Thr (T), Arg (R) presentes nas posições H2, H5, LE1 e LE2 em todas as sequências encontradas (Tabela 3). Para VuTIPs, o filtro não foi conservado em todas as posições para as diferentes sequências, apresentando algumas substituições. Os membros de TIP1

apresentam His (H), Ile (I), Ala (A) e Val (V), enquanto que os membros de TIP2 apresentam His (H), Ile (I), Gly (G) e Arg (R) como padrão de resíduos. Entretanto, os membros de TIP3 mesmo com His (H), Ile (I) como resíduos conservados nas posições H2 e H5, apresentaram variação nos resíduos das posições LE1 e LE2, que podem ser resíduos de Gly (G) ou Ala (A) em LE1 e Arg (R) ou Lis (L) em LE2. Já TIP5, com seu único membro, apresentou os resíduos mais variáveis do filtro ar/R [Asn (N), Thr (T), Ala (A) e Ala (A)]. O filtro de seletividade para VuNIP, apresentou-se mais conservado para os dois membros de NIP1 que apresentam três resíduos conservados nas posições H5, LE1 e LE2, com Val (V) Ala (A) e Arg (R). O sítio H2 apresentou um Trp (T) em VuNIP1-1, sendo que para VuNIP1-2, não foi possível identificar um resíduo nesta posição. A subfamília SIP, tida como a mais divergente quando comparada às demais, apresentou dois padrões de resíduos no filtro ar/R, ambos não observados em outras espécies vegetais (Tabela 4).

Tabela 4: Padrões de Filtro de seletividade ar/R de feijão-caupi.

Membros MIP	H2	H5	LE1	LE2
PIPs				
VuPIP1-1; VuPIP1-2; VuPIP1-3; VuPIP1-4; VuPIP1-5; VuPIP2-1; VuPIP2-2; VuPIP2-3; VuPIP2-4; VuPIP2-5; VuPIP2-6	F	H	T	R
TIPs				
VuTIP1-1; VuTIP1-2	H	I	A	V
VuTIP3-2a; VuTIP3-2b	H	I	A	*L*
Vu TIP2-1; VuTIP2-2; VuTIP2-3; VuTIP3-1; VuTIP3-3	H	I	G	R
VuTIP5-1	N	T	A	*A*
NIPs				
VuNIP1-1; VuNIP1-2	W	V	A	R
VuNIP6-1	*T*	I	G	R
VuNIP7-1	A	S	G	R
SIPs				
VuSIP1-1	*I*	*T*	P	*F*
VuSIP2-1	*N*	*T*	P	*N*

* Em itálico as sequências e seus correspondentes resíduos não identificados em *Arabidopsis*, milho e arroz.

As posições conservadas de Froger's, reconhecidas como resíduos conservados que separam aquaporinas de aquagliceroporinas, provendo especificidades funcionais, também foram analisadas nas aquaporinas de feijão-

caupi. A subfamília PIP apresentou dois dos cinco sítios (P1-P5) conservados em todas as sequências, possuindo uma Ser (S) no sítio P2 e um Trp (W) no sítio P5. O sítio P1 foi o menos conservado para as VuPIPs, com os aminoácidos Glu (E), Met (M) e Gln (Q) aparecendo nesta posição, enquanto que P3 apresentou uma Thr (T) ao invés de Ala (A) na sequência VuPIP1-1 e um Tyr (Y) substituía o resíduo de Phe (F) na aquaporina VuPIP2-5 (Tabela 3).

Os cinco sítios Froger's apareceram conservados em todas as posições preditas para os membros de TIP2, como Thr (T), Ser (S), Ala (A), Tyr (Y), e Trp (W), mas relativamente variáveis nos demais membros VuTIPs, cuja conservação de aminoácidos ocorreu apenas nos sítios P1 e P5, com uma Thr (T) e um Trp (W), respectivamente. Adicionalmente, a subfamília NIP também mostrou variação em termos de resíduos de aminoácidos nos sítios Froger's. Nesta subfamília, houve conservação nos sítios P3, com uma Ala (A) e P4 com um Tyr (Y), os demais sítios mostraram-se variáveis, com Phe (F) ou Tyr (Y) na posição P1, Ser (S) ou Thr (T) no sítio P2, e a posição P5, onde pode ser encontrado Ile (I), Leu (L) ou Phe (F). As duas sequências pertencentes a VuSIP apresentaram similaridade com relação aos resíduos de aminoácidos, com Met (M), Ala (A), Ala (A), Tyr (Y) e Met (M) nas posições de P1 à P5.

5.2 Análise comparativa e Filogenia

A análise filogenética das aquaporinas de domínio completo em feijão-caupi confirmou a classificação deste grupo de transportadores em quatro grupos ortólogos (Figura 12). As diferentes subfamílias foram analisadas independentemente, e de forma comparativa às MIPs de *Arabidopsis* (AtMIPs), agrupando-se em *clusters* menores de acordo com a similaridade de sequência com as AtMIPs inseridas no alinhamento. Como em outros organismos vegetais, a subfamília de aquaporinas PIP apresentou divisão em dois subgrupos, entretanto, a divisão em cinco subgrupos para a subfamília TIP e sete para NIP, previamente descritos para *Arabidopsis*, não foi observada em *Vigna*, pois os membros encontrados foram passíveis de classificação em apenas quatro grupos distintos em TIP e três em NIP. Os membros da subfamília PIP foram divididos em PIP1 e PIP2, TIP foi agrupado em TIP1, TIP2, TIP3 e TIP5, enquanto que três e dois subgrupos apareceram para NIP (NIP1, NIP6 e NIP7) e SIP (SIP1 e SIP2). Nos nossos dados

não foram identificados membros de alguns grupos já descritos para outros vegetais, estes subgrupos incluem TIP4, NIP2, NIP3, NIP4 e NIP5.

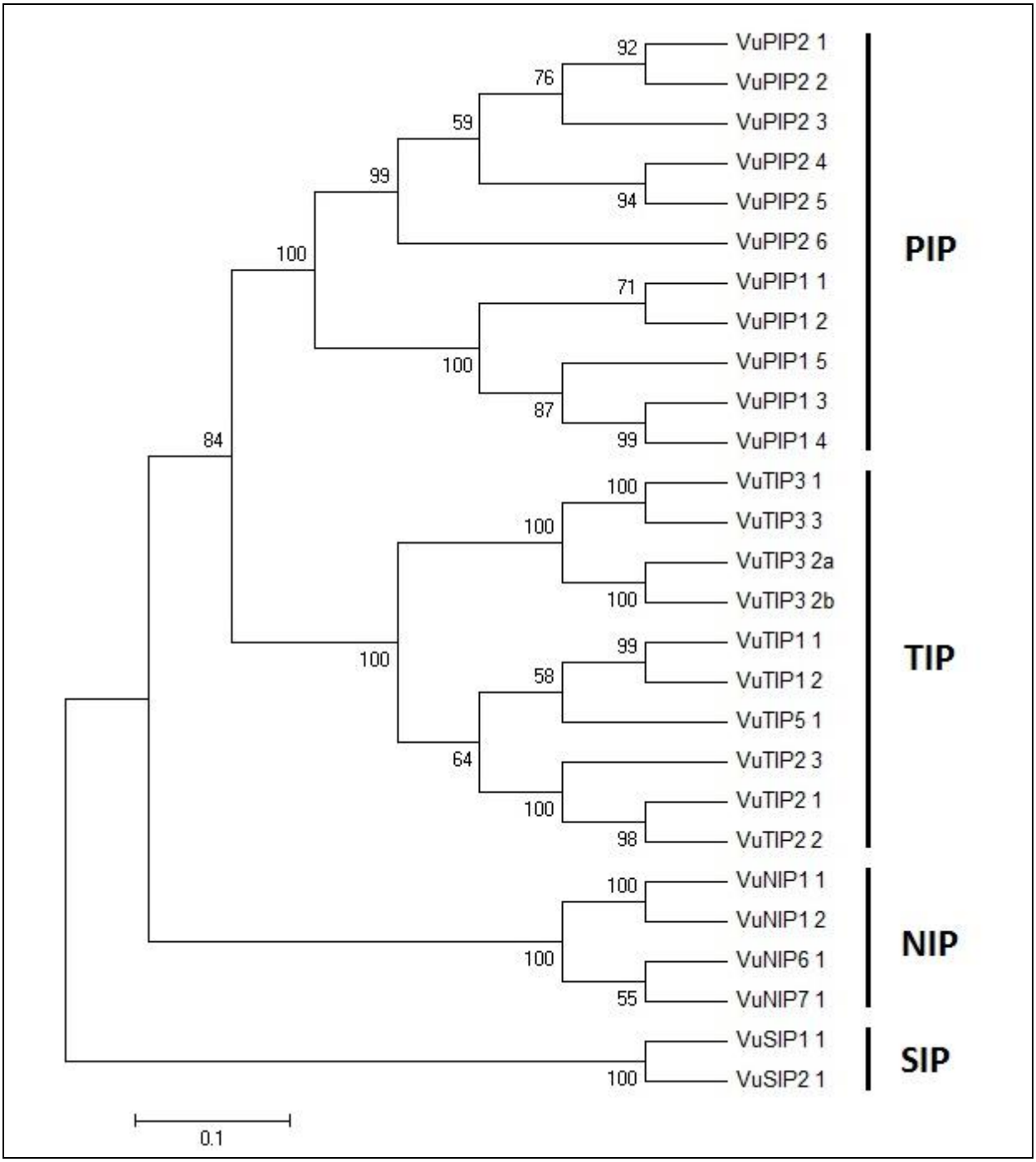


Figura 12. Análise fenética de membros da família de aquaporinas de *Vigna unguiculata* pelo método de *Neighbor Joining*, com valores de *bootstrap* ao nível de 1000 replicações.O dendrograma evidencia as 27 sequências de domínio completo e seu agrupamento em quatro subfamílias com relação à similaridade de suas sequências. Escala de 0,1 refere-se a 10% de divergência entre as sequências. Legenda: PIP (Proteínas de Membrana Plasmática), TIP (Proteínas de Membrana de Tonoplasto), NIP (Proteínas de Membrana de Nódulos) e SIP (Pequenas Proteínas Básicas Intrínsecas de Membrana).

Os dados para VuPIP produziram um alinhamento de 305 posições, onde 147 foram variáveis, 106 parcimônio-informativos e 158 sítios foram conservados entre os diferentes membros de PIP, indicando que PIP representa um grupo bastante homogêneo, comparativamente com baixos níveis de divergência. A análise filogenética suportam a separação de VuPIPs em dois grupos distintos: VuPIP1 e VuPIP2 (Figura 13), com cinco e seis membros respectivamente (DANIELSON; JOHANSON, 2010).

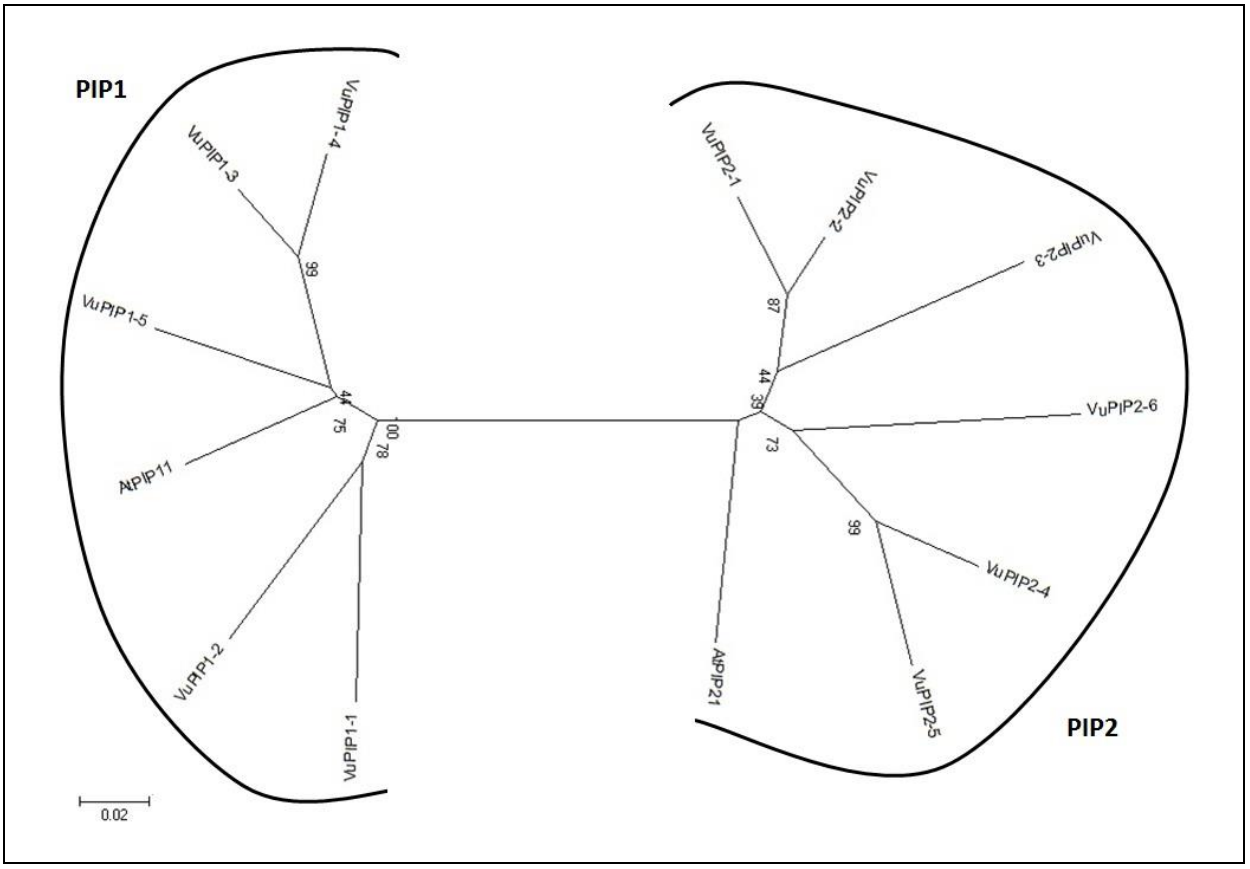


Figura 13. Análise fenética para as VuPIPs de *Vigna* e PIPs de *Arabidopsis*. Divisão em dois grupos distintos. Escala de 0,02 refere-se a 2% de divergência entre as sequências. Legenda: PIP (Proteínas de Membrana Plasmática).

Em feijão-caupi, as sequências VuTIPs produziram um alinhamento de 270 posições, com 210 sítios variáveis, 167 parcimônio-informativos e 48 conservados para todas as sequências (Figura 14). Esta análise confirma tanto a divisão proposta para TIPs em quatro subgrupos (ZARDOYA et al., 2005), como para cinco grupos aceitos para a classificação desta subfamília (JOHANSON et al., 2001;

DANEIELSON; JOHANSON, 2010). Os grupos encontrados foram TIP1, TIP2, TIP3 e TIP5, com dois, três, quatro e um membro respectivamente.

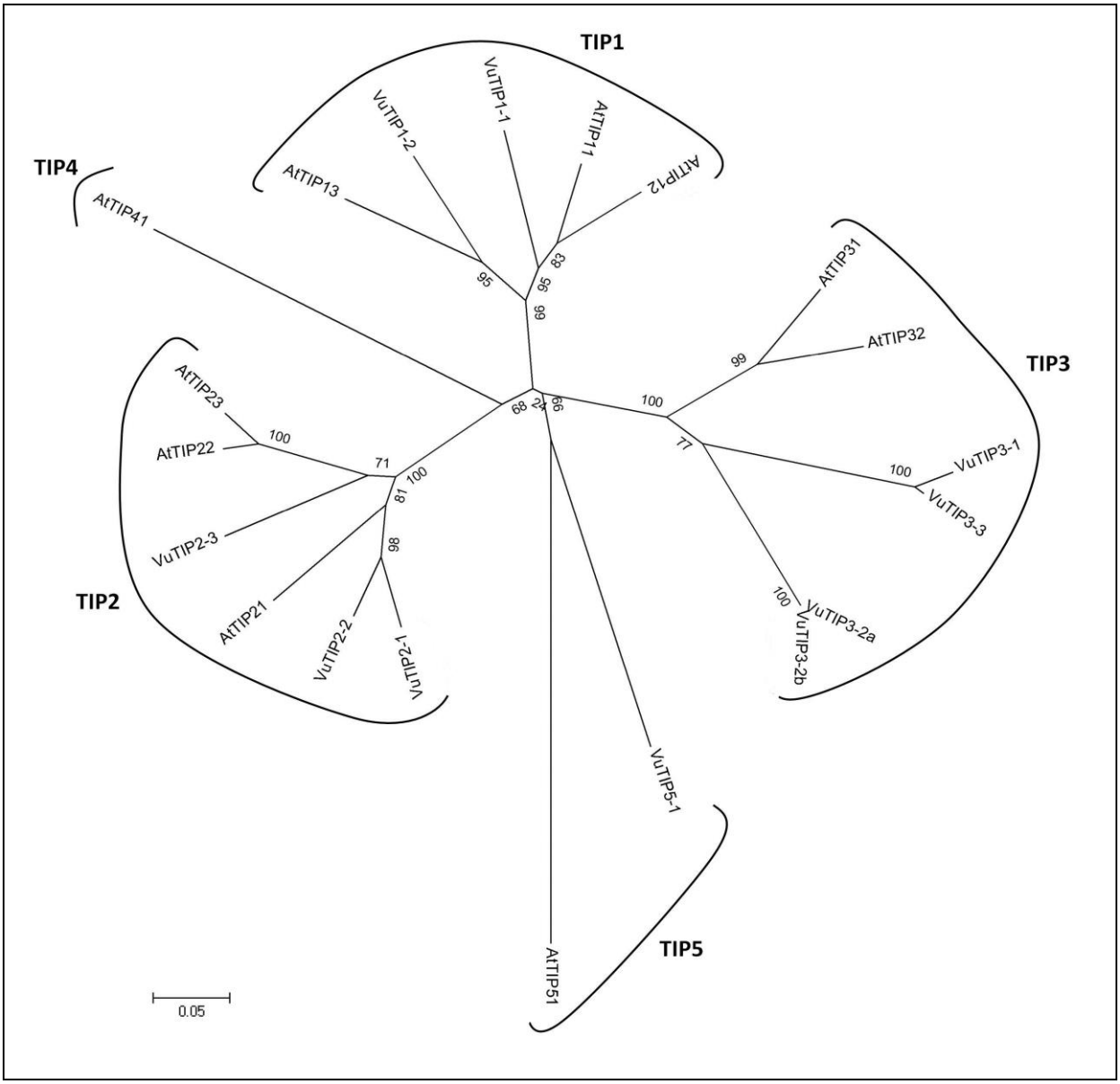


Figura 14. Análise fenética para as VuTIPs de *Vigna* e TIPs de *Arabidopsis*. Divisão em cinco grupos distintos, onde a escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: TIP (Proteínas de Membrana de Tonoplasto).

Os dados para VuNIP resultaram em um alinhamento de 335 posições, onde 247 representavam sítios variáveis, 114 sítios parcimônio-informativos e 67 sítios conservados nas sequências analisadas. As sequências pertencentes à VuNIPs foram agrupadas em NIP1 (dois membros), NIP6 (um membro) e NIP7 (um membro) (Figura 15), classificação esta oriunda da dicotiledônea *Arabidopsis*, que comporta

sete subgrupos em sua classificação (QUIGLEY et al., 2001; JOHANSON et al., 2001), mas alguns autores observaram que em milho e algumas outras espécies vegetais são apenas identificados 3 grupos na subfamília NIP (CHAUMONT et al., 2001; DANIELSON; JOHANSON, 2010).

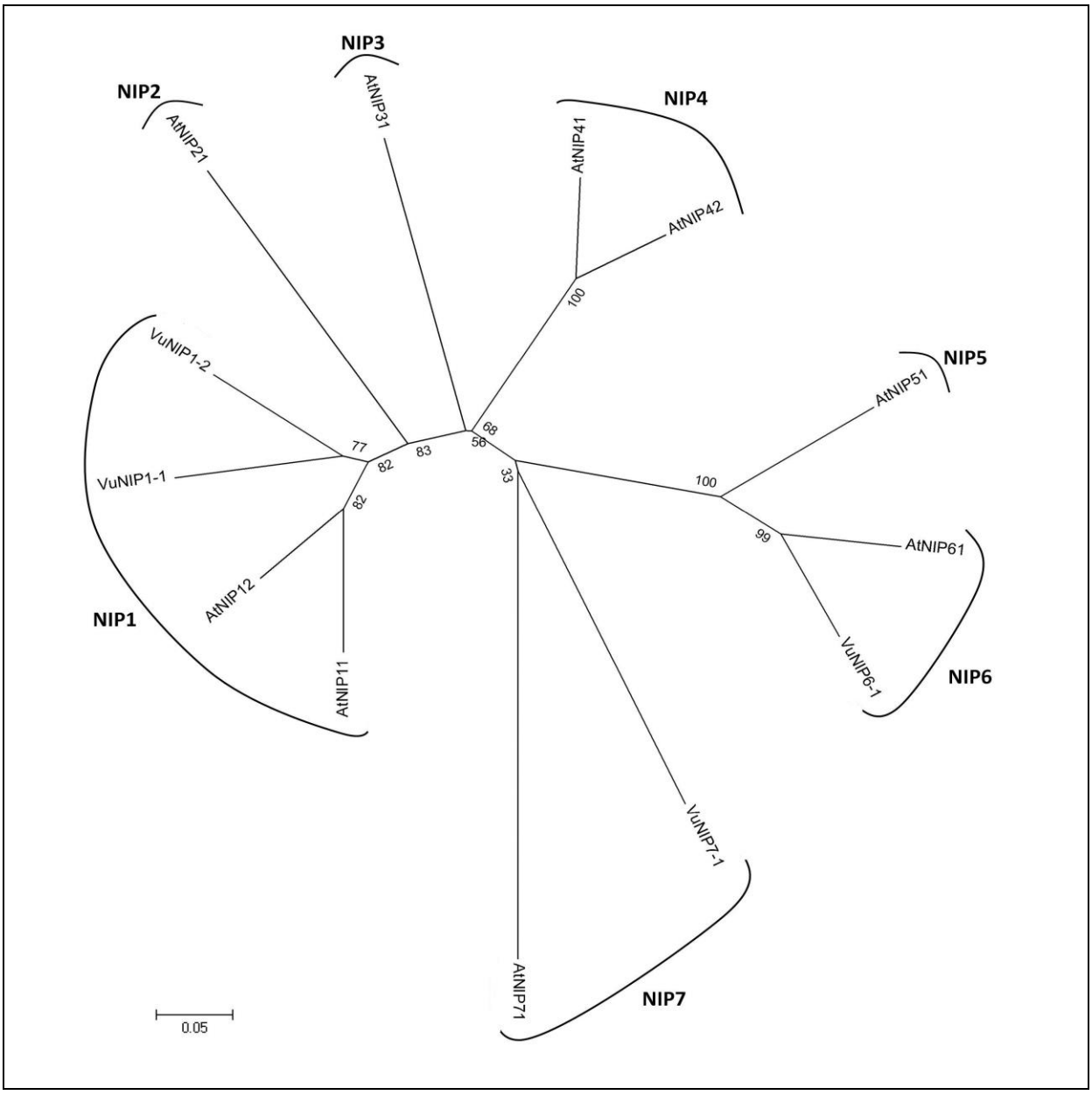


Figura 15. Análise fenética para as VuNIPs de *Vigna* e NIPs de *Arabidopsis*. Subfamília dividida em sete grupos segundo classificação de *Arabidopsis*. Escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: NIP (Proteínas de Membrana de Nódulos).

A análise fenética com VuSIPs apontam esta subfamília como sendo a mais divergente. Tal alinhamento apresentou 253 posições, com 161 sítios variáveis, 83

conservados, mas nenhum parcimônio-informativo, agrupando-se nas duas divisões descritas para monocotiledôneas e dicotiledôneas (Figura 16).

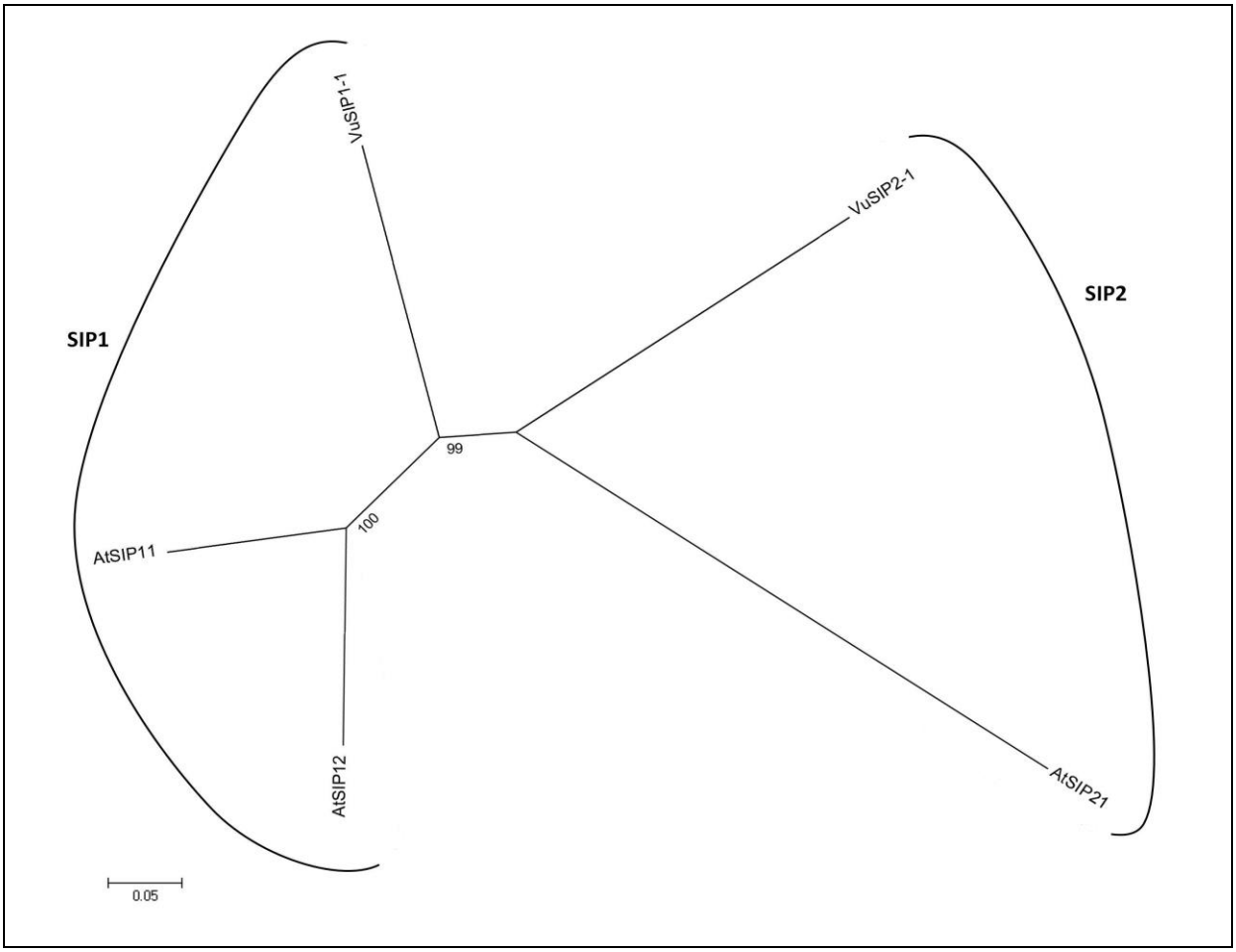


Figura 16. Análise fenética para as VuSIPs de *Vigna* e SIPs de *Arabidopsis*. Divisão em dois grupos distintos: VuSIP1 e VuSIP2, onde a escala de 0,05 refere-se a 5% de divergência entre as sequências. Legenda: SIP (Pequenas Proteínas Básicas Intrínsecas de Membrana).

Para determinar relações filogenéticas entre as aquaporinas de *V. unguiculata* com as aquaporinas de outras espécies vegetais, sequências completas de *A. thaliana*, *Z. mays*, *R. communis* e *V. vinifera* foram também analisadas (Figura 17). Como já visto em outras espécies (FOUQUET et al., 2008; GUPTA; SANKARARAMAKRISHNAN, 2009). Pela árvore gerada, a partir dos alinhamentos pelo método de *Neighbor-Joining*, pode-se observar uma distribuição das MIPs de *V. unguiculata* (VuMIPs) nas quatro subfamílias descritas nos vegetais superiores, excluindo-se a subfamília XIP, não encontrada em nossas análises.

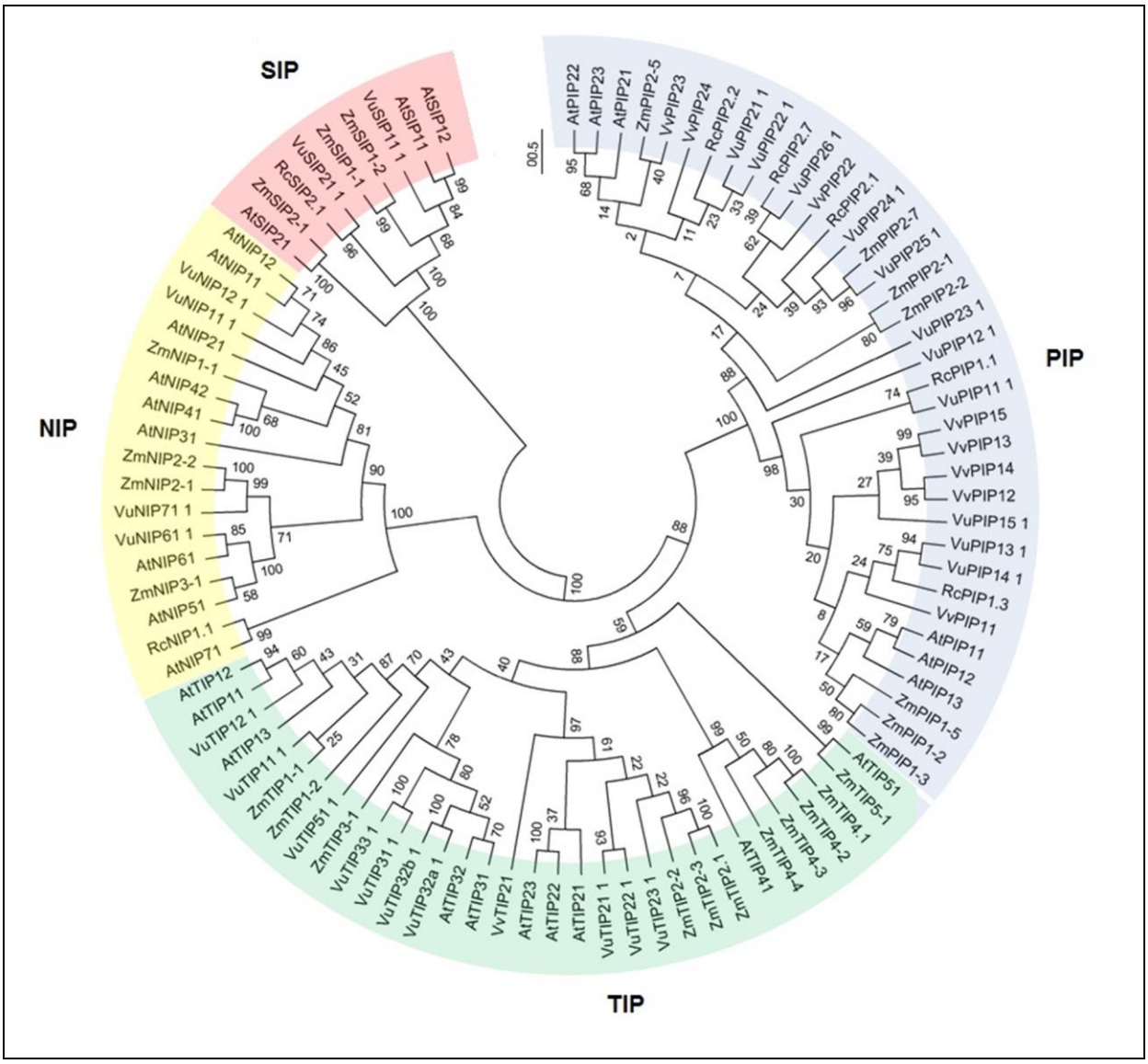


Figura 17. Relações fenéticas entre aquaporinas vegetais. Árvore construída pelo método de *Neighbour Joining*, mostrando a comparação entre as 27 aquaporinas completas de *V. unguiculata* em conjunto com aquaporinas de milho (*Zea mays*), *Arabidopsis thaliana*, *Ricinus communis* e uva (*Vitis vinifera*). As quatro subfamílias encontradas e feijão-caupi estão indicadas pelas diferentes cores. A barra indica distância de 0,05 indica 5% de alteração na sequência de aminoácidos.

5.3 Expressão das aquaporinas de feijão-caupi

Como em nossas análises foram considerados dados oriundos de diversos bancos de dados (públicos ou de acesso restrito); diferentes tempos de desenvolvimentos e condições, o número de vezes em que um transcrito específico

é identificado nas bibliotecas não representa toda população, mas apenas uma estimativa da abundância deste na população de mRNA.

Os transcritos de aquaporinas em feijão-caupi correspondem a 2.225 reads distribuídas em oito bibliotecas. Destas, as sequências mais abundantes nas diferentes bibliotecas foram VuTIP1-1 e a maioria dos membros de VuPIP, tanto VuPIP1 e VuPIP2 (Tabela 5). Para 11 transcritos foram encontrados menos de 10membros nas bibliotecas, fazendo com que algumas famílias fossem fracamente representadas. Para a subfamília NIP, o transcrito VuNIP1-1 representou o único membro com contagem superior a 10 vezes e VuSIP1-1 e VuSIP2-1 apareceram seis e quatro vezes, respectivamente.

Utilizando os dados de expressão obtidos para os oito órgãos, um perfil de expressão por clusterização hierárquica foi construído baseado na abundância de transcritos (Figura 18a). Em geral as VuMIPs apresentaram expressão nos diversos órgãos vegetais, como pode ser observado na biblioteca Mi1, onde apenas os membros VuTIP3-1, VuTIP3-2a, VuTIP3-2b, VuTIP3-3, VuNIP1-1, VuNIP7-1 e VuPIP1-1 não apresentaram níveis de expressão significativos após a normalização dos dados. Tais transcritos apresentaram expressão diferencial em tecidos/órgãos específicos, sugerindo uma expressão sítio-específica. Os membros VuTIPs, ausentes constitutivamente, apresentaram uma expressão preferencial em sementes nos estágios iniciais de desenvolvimento, já VuNIP1-1 e VuNIP7-1 foram mais expressos em raiz e folha, respectivamente. O membro VuPIP1-1, diferentemente dos demais membros desta subfamília, apresentou expressão apenas em tecidos radiculares, com expressão suprimida (*down-regulated*) quando submetida ao estresse salino, em comparação ao controle (raízes sem estresse com NaCl).

Com exceção da biblioteca Mi1 (mix de tecidos), entre as bibliotecas analisadas, a que apresentou um maior números de transcritos relacionados a aquaporinas foi a biblioteca de raiz, que corresponde a um dos principais órgão vegetais relacionados à absorção de água, seguido da biblioteca de folha, órgão responsável pelo controle da perda de água pelo processo de transpiração vegetal.

As raízes apresentam membros de todas as quatro subfamílias sendo expressos, em especial VuNIPs que são característicos no processo de nodulação em Leguminosas (WALLACE et al., 2006), e VuPIPs, família com alta eficiência no transporte de água (CHAUMONT et al., 2000).

Tabela 5: Distribuição não normalizada das reads nas diferentes bibliotecas de feijão-caupi. Mi1: mix de tecidos; Lfm: folha e meristema apical; Dsd: sementes em desenvolvimento; Rt3: raíz; Lf3: folha; Bud: botão floral; RtS: raiz sob salinidade.

Trancrito	Mi1	Lfm	Dsd	Rt3	Lf3	Bud	RtS	Total
VuPIPs	**1061**	**27**	**3**	**11**	**5**	**0**	**62**	**1169**
VuPIP1-1	0	0	0	2	0	0	2	4
VuPIP1-2	5	1	0	0	0	0	0	6
VuPIP1-3	387	1	1	0	0	0	17	406
VuPIP1-4	25	1	0	0	0	0	12	38
VuPIP1-5	122	6	0	0	4	0	2	134
VuPIP2-1	27	6	0	4	0	0	5	42
VuPIP2-2	123	4	0	2	0	0	2	131
VuPIP2-3	2	0	0	0	0	0	3	5
VuPIP2-4	76	7	0	0	1	0	9	93
VuPIP2-5	159	1	2	3	0	0	9	174
VuPIP2-6	135	0	0	0	0	0	1	136
VuTIPs	**866**	**8**	**76**	**14**	**0**	**2**	**42**	**1008**
VuTIP1-1	656	1	19	0	0	2	26	704
VuTIP1-2	6	0	0	5	0	0	0	11
VuTIP2-1	98	2	0	4	0	0	1	105
VuTIP2-2	25	0	0	1	0	0	1	27
VuTIP2-3	75	1	0	4	0	0	14	94
VuTIP3-1	0	0	2	0	0	0	0	2
VuTIP3-2ª	0	2	6	0	0	0	0	8
VuTIP3-2b	0	2	36	0	0	0	0	38
VuTIP3-3	1	0	13	0	0	0	0	14
VuTIP5-1	5	0	0	0	0	0	0	5
VuNIPs	**5**	**6**	**4**	**2**	**1**	**0**	**20**	**38**
VuNIP1-1	1	0	0	0	0	0	18	19
VuNIP1-2	1	2	0	0	0	0	1	4
VuNIP6-1	3	0	0	0	0	0	0	3
VuNIP7-1	0	4	0	0	0	0	1	5
VuSIPs	**8**	**2**	**0**	**0**	**0**	**0**	**0**	**10**
VuSIP1-1	6	0	0	0	0	0	0	6
VuSIP2-1	2	2	0	0	0	0	0	4

Para as VuPIPs os principais órgãos de expressão também foram as raízes (em especial para os transcritos VuPIP1-1, VuPIP1-4 e VuPIP2-3) e folhas (sequências VuPIP1-2, VuPIP1-5 e VuPIP2-1). A expressão de membros desta família não foi observada em botões florais. Entretanto, em sementes foi verificada a expressão da aquaporina VuPIP2-5. Os membros da subfamília VuTIPs clusterizaram de forma semelhante a análise filogenética, ficando a expressão de TIP3 situada em sementes e folhas (membros VuTIP3-2a e VuTIP3-2b). Para as

demais VuTIPs, a expressão foi situada tanto em tecidos radiculares (VuTIP1-2 e VuTIP2-3) como em botões florais (VuTIP1-1). Os menores níveis de expressão foram os da subfamília VuSIP, que se restringiram ao tecido foliar para VuSIP2-1, não sendo possível a determinação de outros sítios de expressão órgão-específica (Figura 18b).

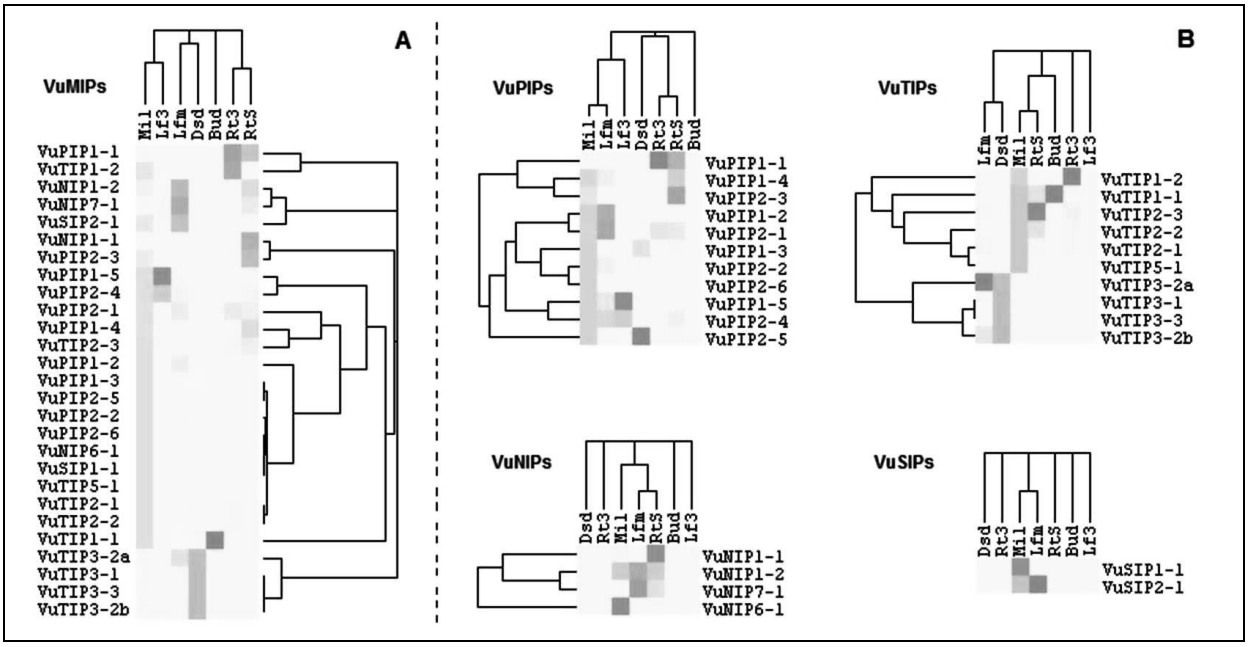


Figura 18. Perfil de expressão obtido após clusterização hierárquica realizada pelo programa CLUSTER 3.0 dos transcritos de aquaporinas de feijão-caupi. Os quadrados em laranja indicam alta expressão, em laranja mais claro indicam baixa expressão e em amarelo, ausência de expressão. Abreviações: Mi1: mix de tecidos; Lfm: folha e meristema apical; Dsd: sementes em desenvolvimento; Rt3: raiz; Lf3: folha; Bud: botão floral; RtS: raízes sob salinidade.

5.4 Modelagem comparativa

A modelagem comparativa foi baseada na utilização de um modelo conhecido (com estrutura determinada experimentalmente), alinhado com as sequências para as quais se desejava determinar a estrutura terciária. Esta análise permitiu uma visualização da arquitetura geral destes transportadores de membrana. A estrutura terciária de uma proteína se dá pelo dobramento de suas estruturas secundárias e domínios. Dada a similaridade ao nível secundário, semelhanças estruturais podem ser consideradas tão eficientes quanto os métodos experimentais (Difração de

Raios-X e Ressonância Magnética Nuclear). Neste estudo, os modelos foram construídos a partir do modelo "*template*" 1Z98 de espinafre (*S. oleracea*, SoPIP2;1), com 2.10 Å de resolução, disponível no Banco de Dados de Proteínas - PDB(*Protein Data Bank*), que apresenta a melhor similaridade com as aquaporinas de feijão-caupi, além de representar a única aquaporina vegetal disponível. Todas as aquaporinas com domínio completo foram submetidas à modelagem computacional, apresentando variações na similaridade após o alinhamento contra o modelo, cujos valores de identidade foram de 84 % (VuPIP2-6) e 25 % em VuSIP1-1 (Tabela 6).

Tabela 6. Valores de similaridade entre o modelo de espinafre (SoPIP2-1) e os diferentes membros de aquaporinas de feijão-caupi.

VuPIPs	Similaridade (%)	VuTIPs	Similaridade (%)	VuNIPs	Similaridade (%)	VuSIPs	Similaridade (%)
VuPIP2-6	84.34	VuTIP2-1	42.53	VuNIP1-1	31.28	VuSIP2-1	26.67
VuPIP2-1	79.34	VuTIP2-2	41.85	VuNIP6-1	30.53	VuSIP1-1	25.85
VuPIP2-4	77.82	VuTIP1-1	40.43	VuNIP1-2	30.34		
VuPIP2-2	77.19	VuTIP3-3	39.74	VuNIP7-1	27.78		
VuPIP2-5	76.22	VuTIP3-2a	39.55				
VuPIP2-3	75.18	VuTIP3-2b	39.55				
VuPIP1-2	74.88	VuTIP2-3	39.37				
VuPIP1-5	74.81	VuTIP3-1	38.58				
VuPIP1-1	71.70	VuTIP1-2	37.61				
VuPIP1-3	70.74	VuTIP5-1	34.82				
VuPIP1-4	69.63						

Os modelos propostos para as aquaporinas de *V. unguiculata* (gerados via MODELLER; Figura 19) apresentaram, para a maioria de seus membros, todas as características principais: seis domínios transmembrana (H1-H6) conectados por cinco loops (LA-LE), juntamente com os dois domínios alfa-hélice portadores dos motivos conservados NPA, que interagem internamente, formando o poro seletivo (GOMES et al., 2009). Como estabelecido para as aquaporinas, os modelos apresentam extremidades carboxi- e amino-terminal, juntamente como os loops A, C e E, voltados para a região citoplasmática, além de exibirem a conformação "*hourglass*", evidenciando a simetria dos segmentos transmembranas e alfa-hélices portadoras dos motivos NPA. As estruturas protéicas modeladas apresentaram peso molecular predito variando entre 31, 08 kDa e 20,16 kDa e pontos isoelétricos entre 5,36 e 9,96.

Para a análise de confiabilidade dos modelos propostos para as aquaporinas de feijão-caupi, diversos programas de avaliação de qualidade foram aplicados, visando à validação dos desenhos computacionais. Os modelos analisados via Procheck (LASKOWSKI et al., 1993), QMEAN (BENKERT et al., 2009) e ProSA (SIPPL, 1993; WIEDRSTEIN; SIPPL, 2007) apresentaram para os modelos propostos, valores dentro dos padrões de confiabilidade (Tabela 7), com *score* QMEAN variáveis dentro do espectro 0,62 - 0,30 (valores mais próximos a um representam modelos com maior qualidade).

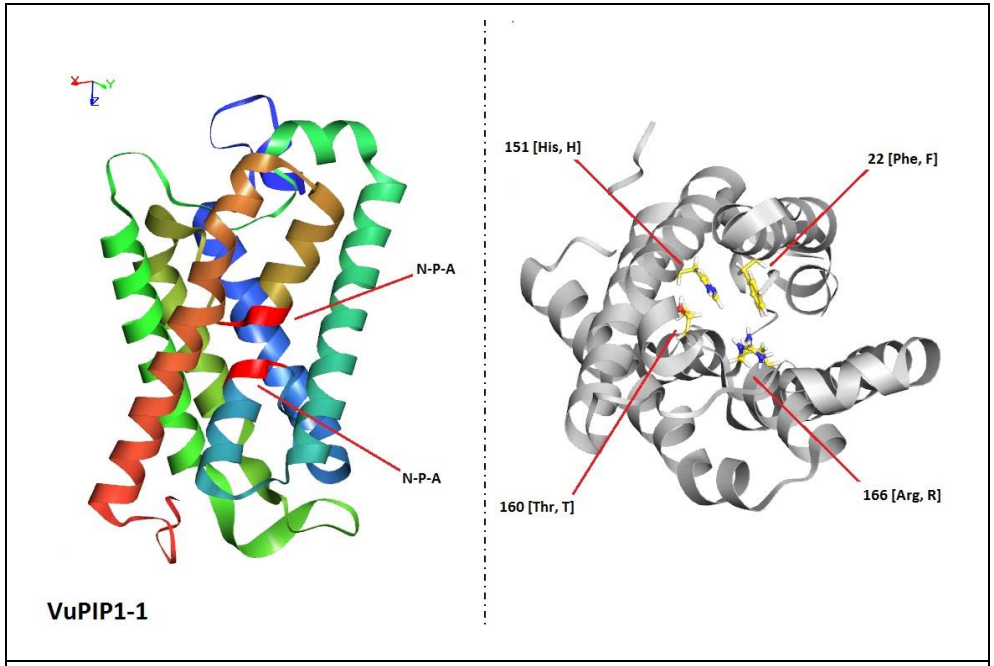
A estrutura estereoquímica dos modelos (via gráfico de Ramachandaran, programa Procheck) apresentou valores superiores a 83 % na quantidade de resíduos localizados nas regiões energeticamente mais favoráveis para um modelo de qualidade, em todos os modelos propostos para as VuMIPs. As inferências com relação à qualidade da geometria dos modelos revelaram que estes apresentaram z-scores entre -1,97 e -4,77 (quanto mais negativos, maior nível de confiabilidade para o modelo).

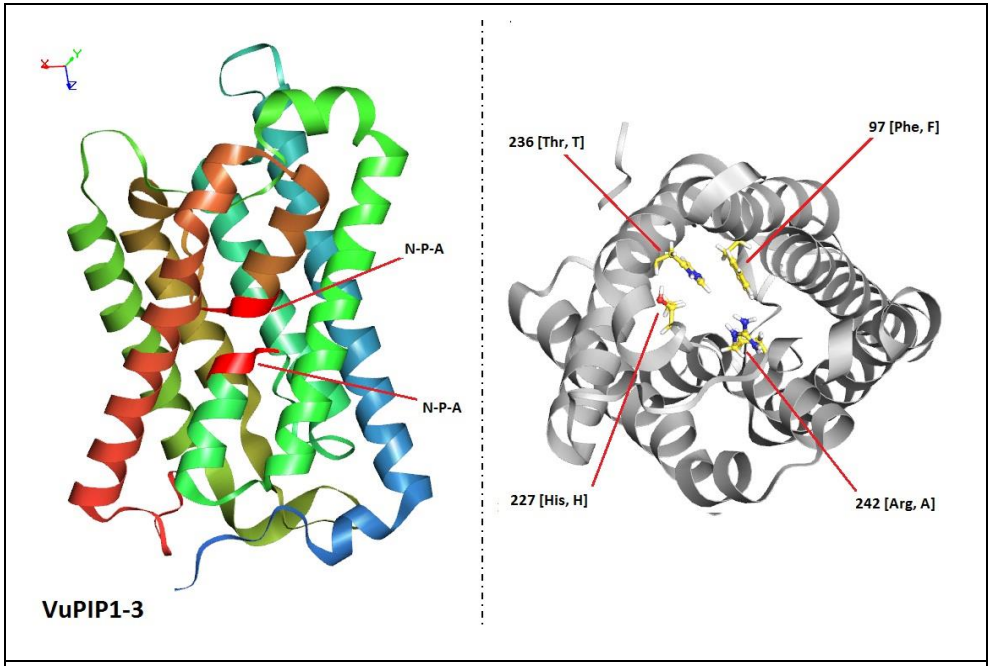
Embora *z-score* negativos sejam a indicação de um bom modelo, também deve ser considerada a plotagem do modelo no gráfico dos *scores* de todas as estruturas protéicas depositadas no PDB, a qual deve estar situada dentro da área azul (clara ou escura), a qual reflete o método de identificação, que pode ser cristalografia de raios-X ou ressonância magnética nuclear RMN). Com base nesta análise, as diferentes subfamílias de aquaporinas modeladas situam-se dentro da área referida a padrões de qualidade desejada, muitas delas próximas ao modelo estrutural utilizado na modelagem comparativa (Tabela 7 e Figura 20A-D).

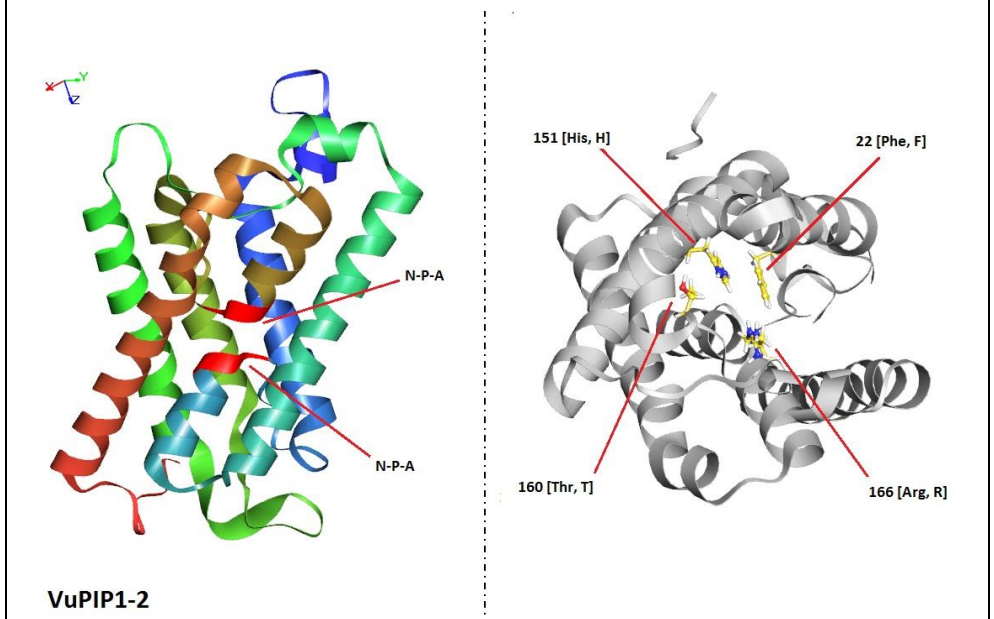
Tabela 7: Valores de qualidade e energia para os modelos estruturais de aquaporinas de feijão-caupi obitidos via plataforma SWISS-MODEL e ferramenta Procheck.

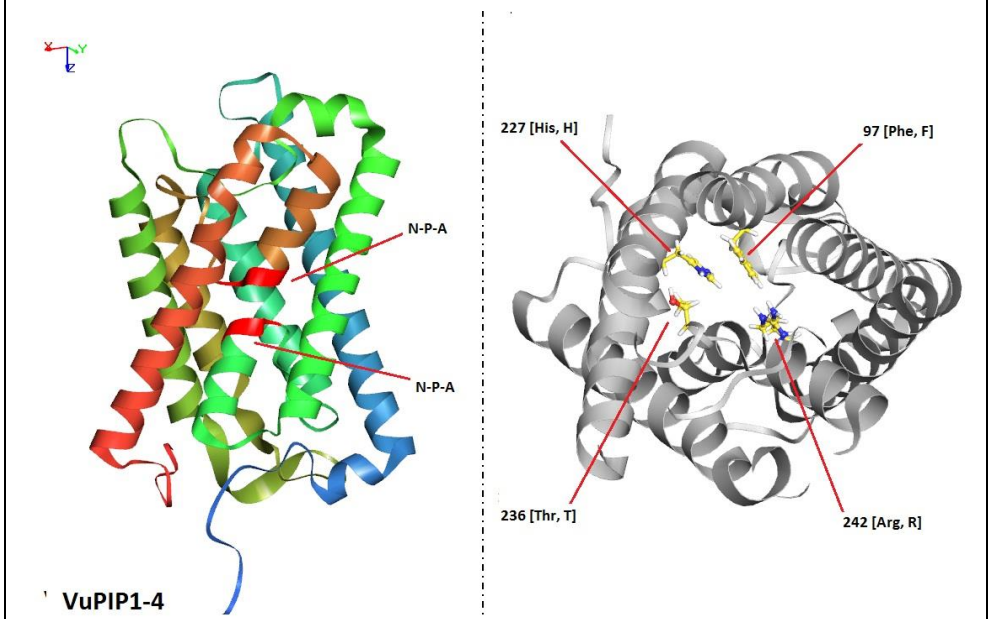
Name	Mass (kDa)	pI	Length (aa)	Z-score	QMEAN	Ramachandran*			
						RMF	RPA	RGP	RNP
SoPIP21	*29,90*	*9.03*	*251*	*-4,36*	*0,52*	*90,0%*	*9,3%*	*0,7%*	*0,0%*
VuPIP11	22,57	9.69	213	-3,69	0,41	91,4%	7,4%	0,6%	0,6%
VuPIP12	22,52	9.65	213	-3,62	0,43	92,6%	6,2%	0,6%	0,6%
VuPIP13	30,82	8.79	289	-4,64	0,54	93,3%	5,9%	0,8%	0,0%
VuPIP14	30,80	7.79	289	-4,37	0,55	92,4%	5,9%	1,7%	0,0%
VuPIP15	30,87	8.98	287	-4,61	0,57	92,9%	5,4%	0,8%	0,8%
VuPIP21	30,40	8.27	285	-4,77	0,56	92,9%	6,7%	0,4%	0,0%
VuPIP22	30,72	8.78	287	-4,41	0,59	91,7%	7,1%	1,2%	0,0%
VuPIP23	30,93	8.77	284	-3,99	0,57	92,0%	6,7%	1,3%	0,0%
VuPIP24	30,61	6.5	285	-4,69	0,57	92,9%	5,4%	0,8%	0,8%
VuPIP25	30,78	7.66	287	-4,77	0,62	92,2%	7,0%	0,9%	0,0%
VuPIP26	29,57	9.31	278	-4,36	0,56	92,8%	5,5%	1,3%	0,4%
VuTIP11	25,48	5.6	250	-2,00	0,51	87,2%	11,3%	0,5%	1,0%
VuTIP12	25,94	5.7	252	-3,34	0,53	89,2%	8,0%	2,4%	0,5%
VuTIP21	25,11	5.75	248	-2,77	0,51	87,9%	11,1%	0,5%	0,5%
VuTIP22	25,19	6.1	248	-2,90	0,53	88,4%	10,1%	1,5%	0,0%
VuTIP23	25,49	5.36	249	-2,95	0,51	87,9%	10,7%	1,0%	0,5%
VuTIP31	22,70	8.6	219	-3,36	0,50	87,4%	10,4%	1,1%	1,1%
VuTIP32a	27,13	6.58	256	-3,54	0,50	88,3%	8,5%	2,3%	0,9%
VuTIP32b	27,05	6.78	256	-3,65	0,50	85,9%	10,8%	2,8%	0,5%
VuTIP33	26,74	6.14	254	-4,49	0,55	87,6%	9,0%	1,0%	2,4%
VuTIP51	25,75	9.19	244	-3,48	0,49	89,3%	9,2%	1,0%	0,5%
VuNIP11	28,87	9.16	272	-1,97	0,37	85,1%	11,0%	2,6%	1,3%
VuNIP12	20,16	9.74	189	-2,29	0,30	83,2%	12,4%	3,1%	1,2%
VuNIP61	31,08	8.25	301	-4,18	0,41	86,6%	10,5%	2,0%	0,8%
VuNIP71	27,26	6.17	263	-3,43	0,48	85,8%	11,4%	1,8%	0,9%
VuSIP11	26,67	7.1	247	-2,76	0,39	86,8%	10,8%	1,9%	0,5%
VuSIP21	26,15	9.96	244	-3,62	0,44	86,6%	8,6%	2,9%	1,9%

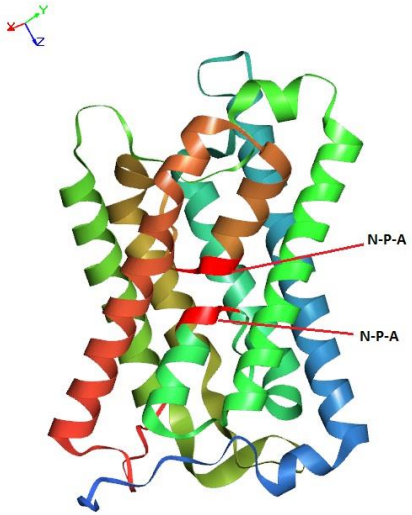
(*) RMF: Região Mais Favorável; RPA: Região Permitida Adicional; RGP: Região Generosamente Permitida; RNP: Região Não Permitida.

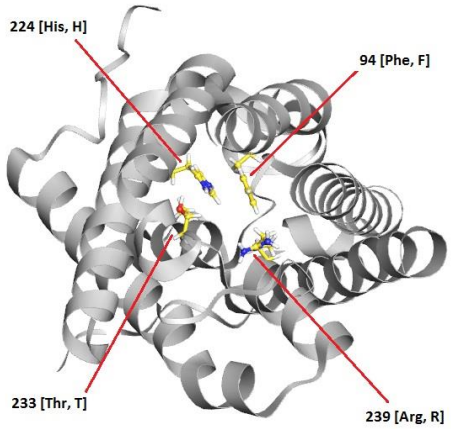

N-P-A
N-P-A
151 [His, H]
22 [Phe, F]
160 [Thr, T]
166 [Arg, R]
VuPIP1-1


N-P-A
N-P-A
236 [Thr, T]
97 [Phe, F]
227 [His, H]
242 [Arg, A]
VuPIP1-3

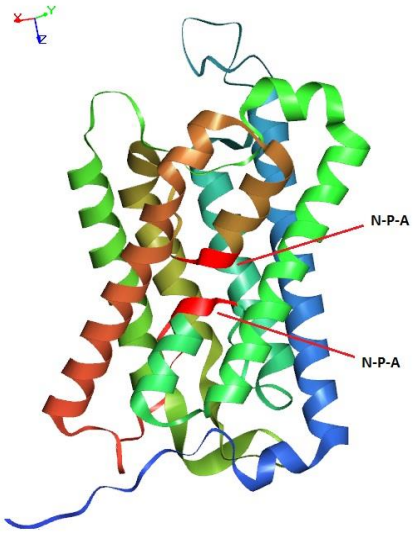

N-P-A
N-P-A
151 [His, H]
22 [Phe, F]
160 [Thr, T]
166 [Arg, R]
VuPIP1-2

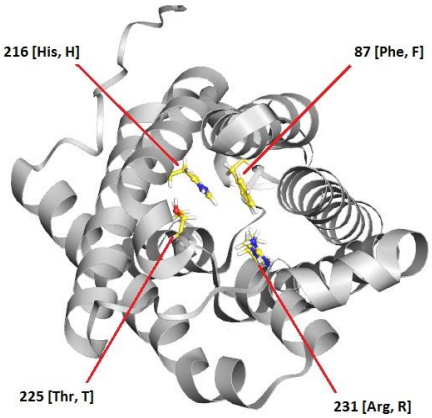

N-P-A
N-P-A
227 [His, H]
97 [Phe, F]
236 [Thr, T]
242 [Arg, R]
VuPIP1-4



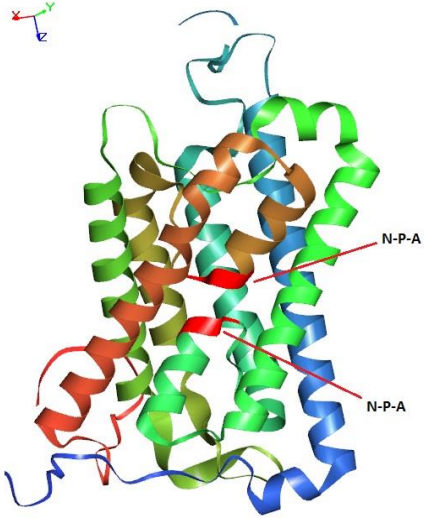


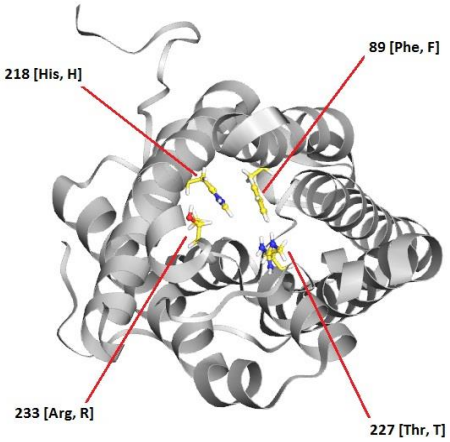
VuPIP1-5



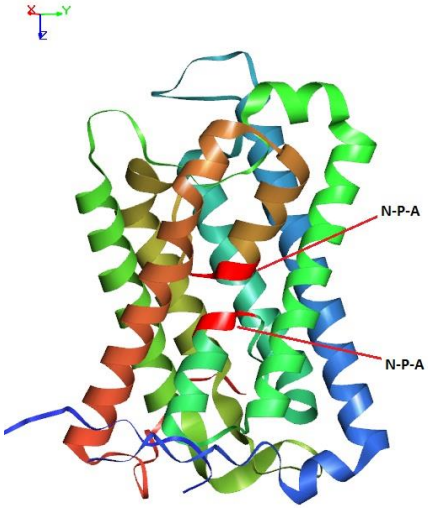


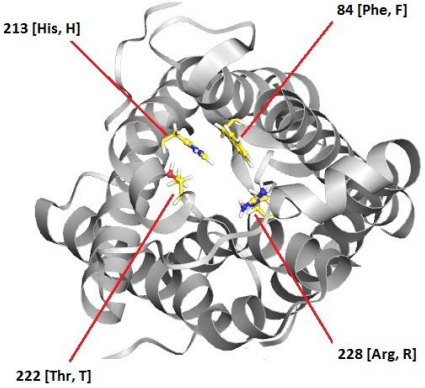
VuPIP2-2



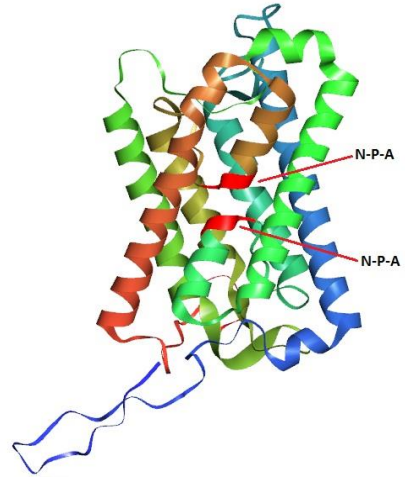


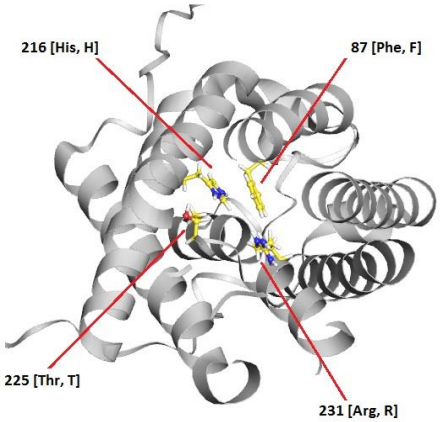
VuPIP2-1



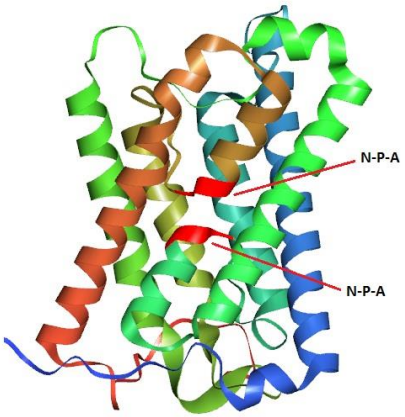


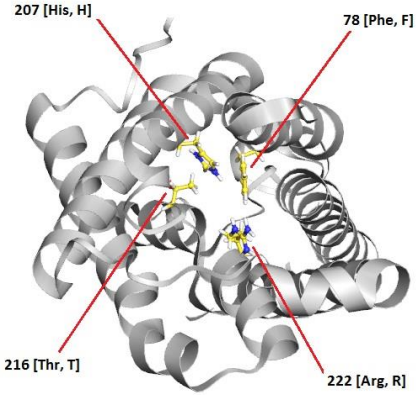
VuPIP2-3



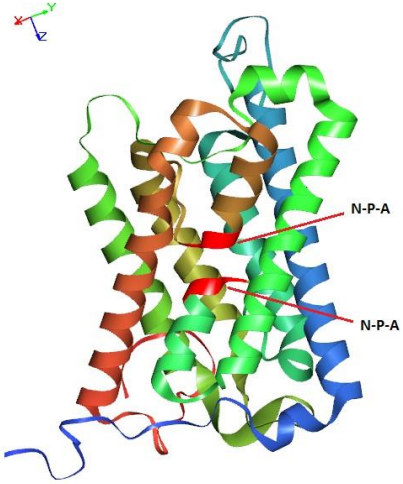


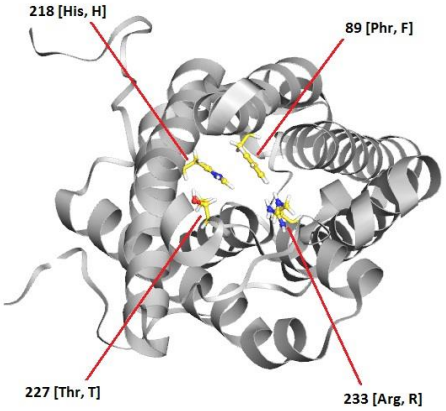
VuPIP2-4



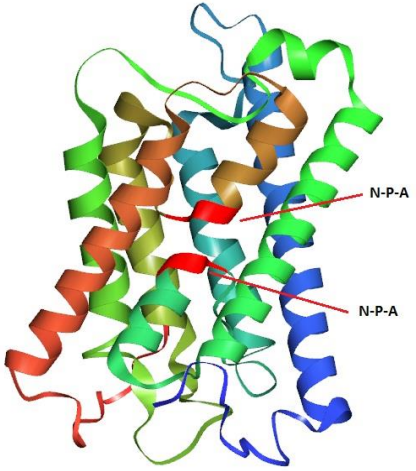


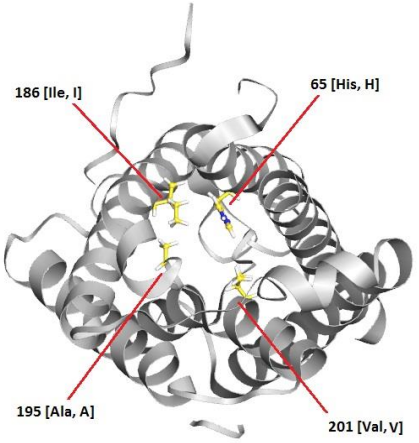
VuPIP2-6



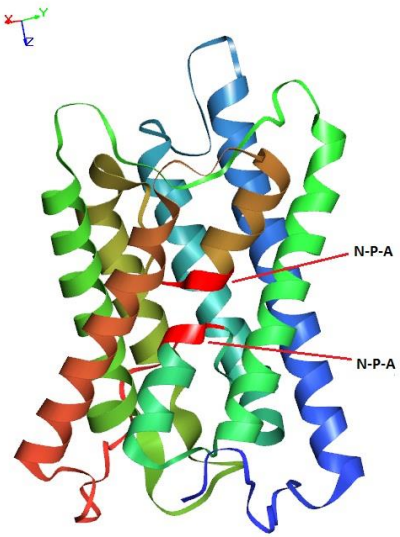


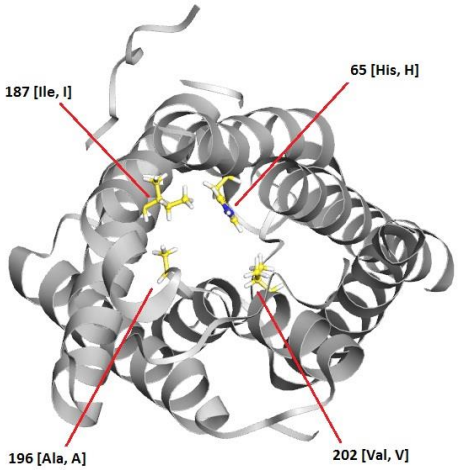
VuPIP2-5



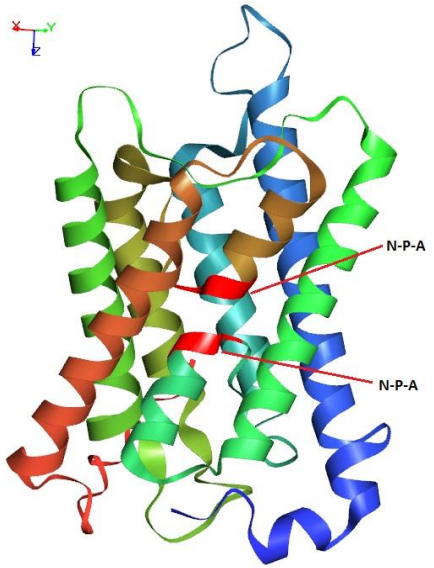


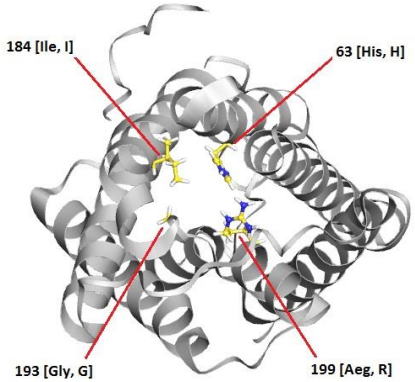
VuTIP1-1



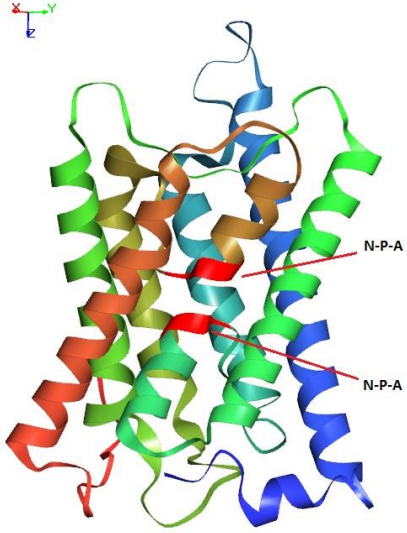


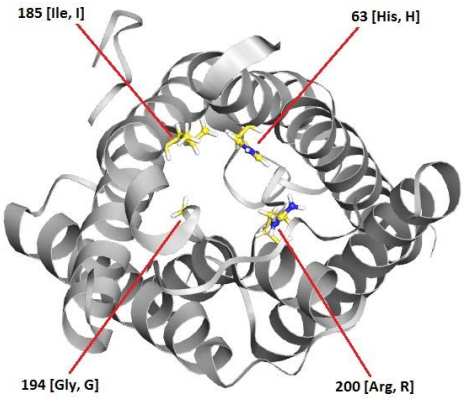
VuTIP1-2



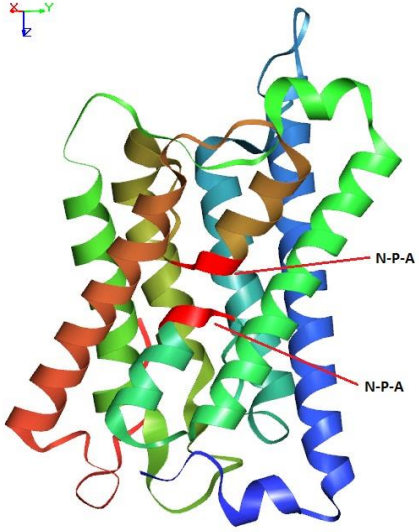


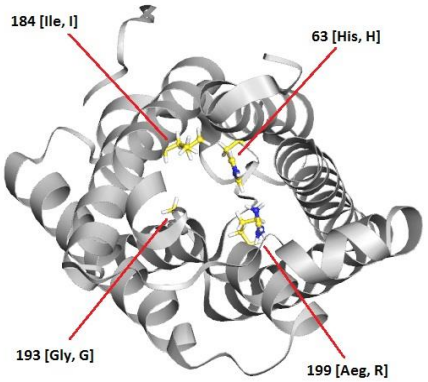
VuTIP2-2



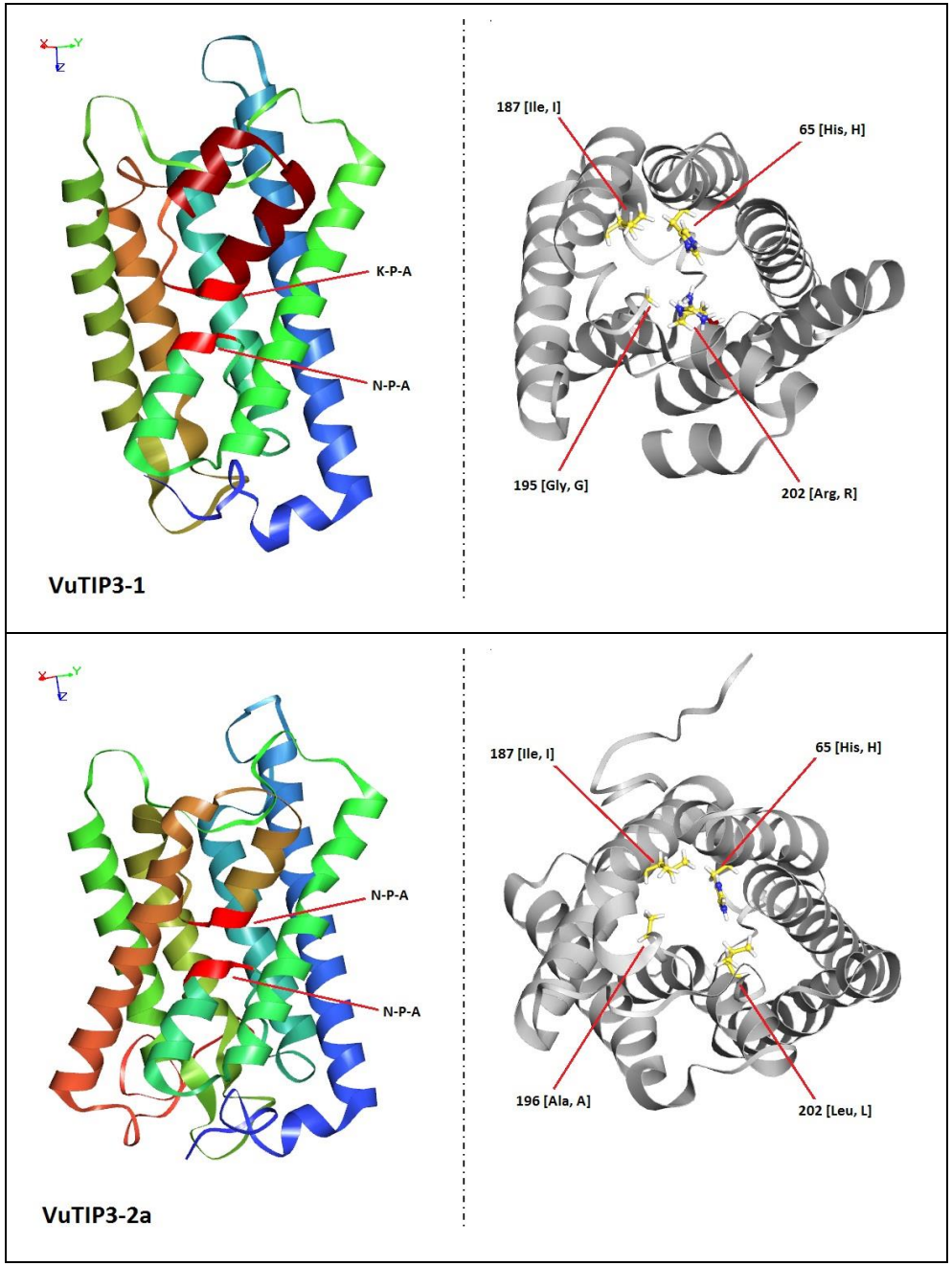


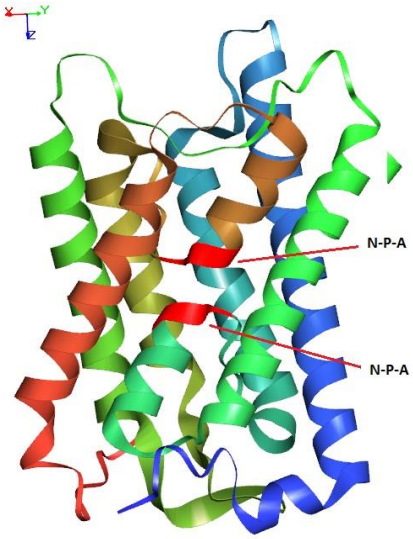
VuTIP2-1

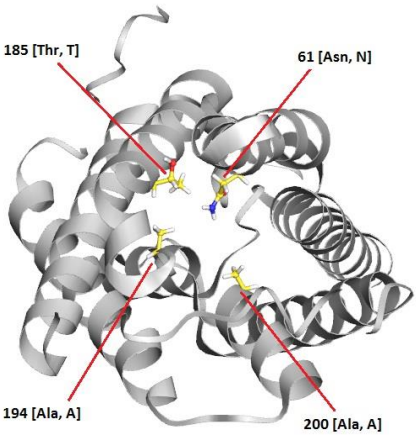




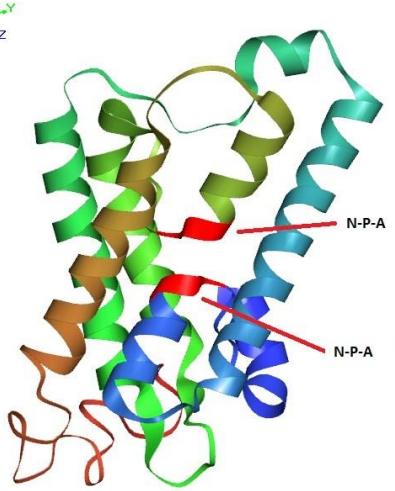
VuTIP2-3

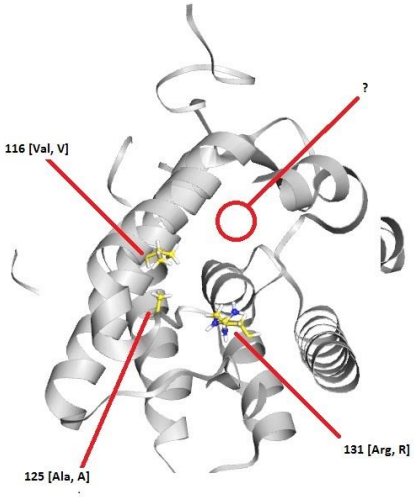

K-P-A
N-P-A
187 [Ile, I]
65 [His, H]
195 [Gly, G]
202 [Arg, R]
VuTIP3-1
N-P-A
N-P-A
187 [Ile, I]
65 [His, H]
196 [Ala, A]
202 [Leu, L]
VuTIP3-2b
N-P-A
N-P-A
187 [Ile, I]
65 [His, H]
196 [Ala, A]
202 [Leu, L]
VuTIP3-2a
N-P-A
N-P-A
187 [Ile, I]
65 [His, H]
195 [Gly, G]
202 [Arg, R]
VuTIP3-3



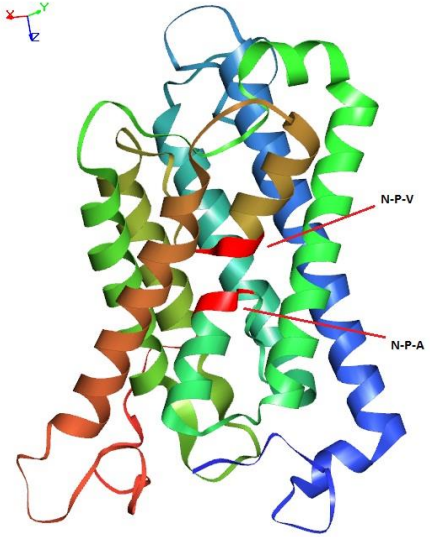


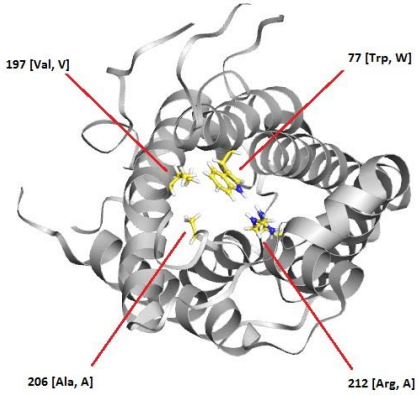
VuTIP5-1



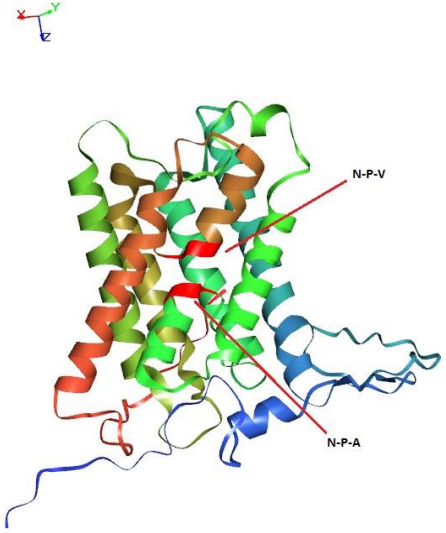


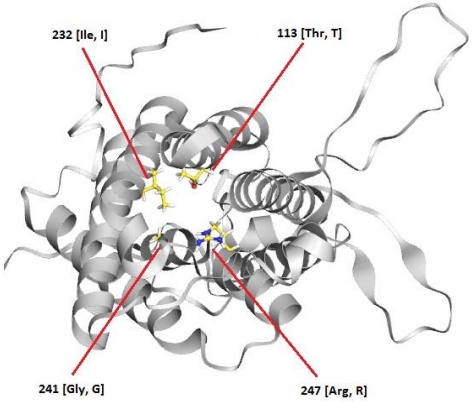
VuNIP1-2





VuNIP1-1





VuNIP6-1

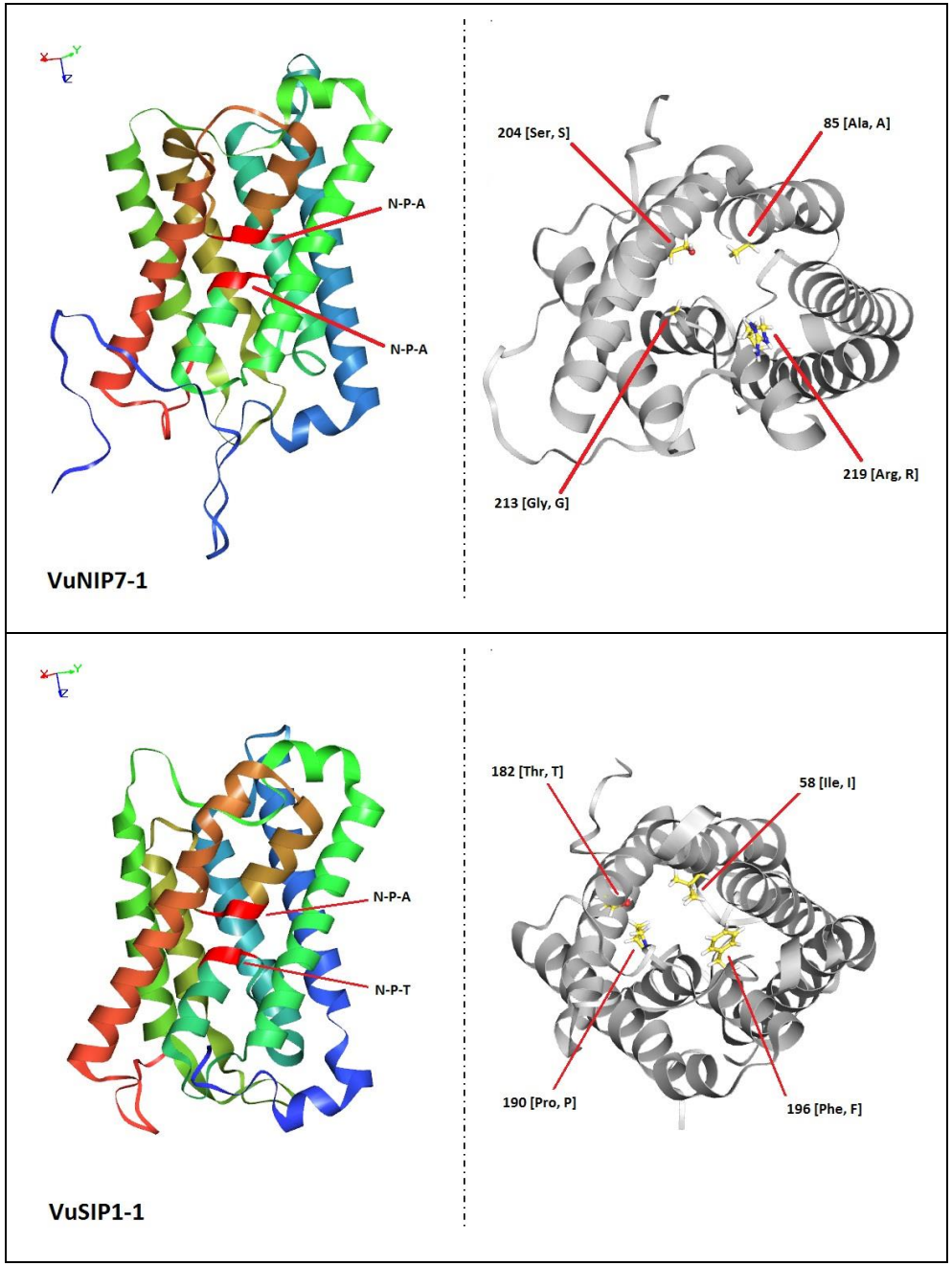


Figura 19. Modelos tridimensionais das aquaporinas de feijão-caupi, obtidas via modelagem comparativa, a partir do modelo SoPIP2-1 disponível no PDB.

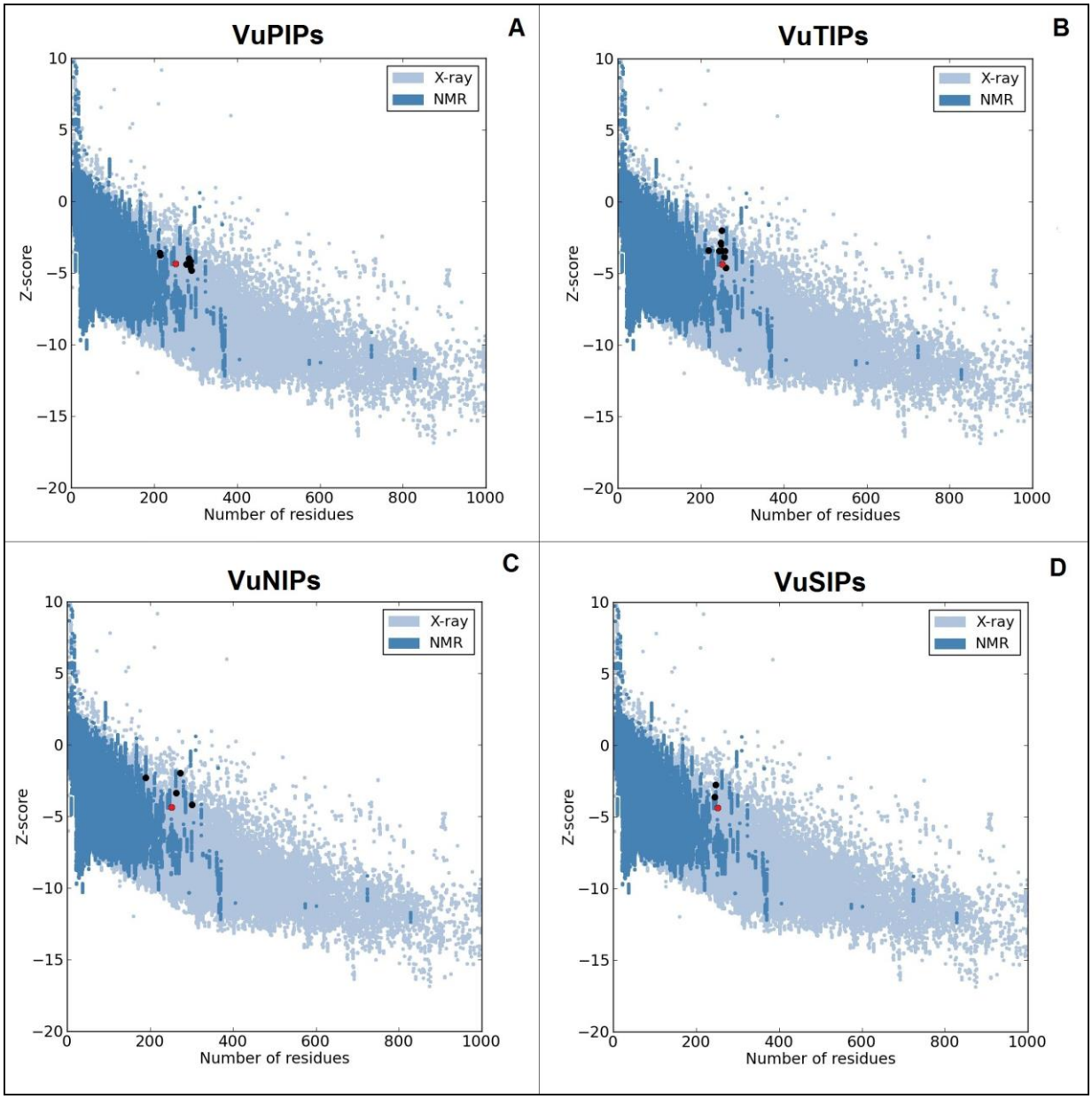


Figura 20. Gráficos representativos do Z-score dos modelos gerados, comparativamente aos modelos de cristalografia de raios-x (em azul claro) e ressonância magnética (em azul escuro) depositados no PDB. Os pontos em preto referem-se aos modelos gerados para as subfamílias PIP (A), TIP (B), NIP (C) e SIP (D); os pontos em vermelho referem-se ao Z-score do modelo de aquaporina de espinafre (SoPIP2-1).

5.5. Validação dos transcritos

Através do uso de 15 pares primers desenhados para aquaporinas de *Vigna unguiculata*, foi possível isolar 13 fragmentos a partir do cDNA da variedade Canapú amarelo, não sendo obtidos amplicons apenas para os primers (VuSIP21F, VuSIP21R, VuTIP21F e VuTIP21R). Os 13 fragmentos obtidos após sequenciamento e montagem dos contigs possuem de 49 a 160 nucleotídeos de comprimento, apresentando similaridades com os transcritos de feijão-caupi ao nível de nucleotídeos superior a 91% de identidade no BLASTn (Tabela 8). Todos os fragmentos gerados a partir do cDNA correspondem a sua sequência (transcrito) de origem (Figura 21), confirmando (validando) a existência destes genes inicialmente preditos apenas por ferramentas de bioinformática, bem como confirmando a expressão destes no tecido radicular, visto que o cDNA foi construído a partir do conjunto total de mRNA deste tecido.

Tabela 8: Valores de similaridade dos fragmentos amplificados do cDNA de feijão-caupi contra os bancos NordEST e NCBI, seus respectivos primers e tamanhos da sequência.

Contig	Primers para amplificação	Tamanho (bp)	Similaridades (%)	
			Transcritos de feijão-caupi (BLASTn)	NCBI (BLASTx)
ContigPIP11	VuPIP11F / VuPIP11R	119	VuPIP11 (91.00)	ACF74346 (85.00)
ContigPIP13	VuPIP13F / VuPIP13R	151	VuPIP13 (100.00)	ACF74346 (92.00)
ContigPIP14	VuPIP14F / VuPIP14R	145	VuPIP14 (97.00)	XP_003552478 (96.00)
ContigPIP21	VuPIP21F / VuPIP21R	63	VuPIP21 (100.00)	AAK83979 (100)
ContigPIP23	VuPIP23F / VuPIP23R	103	VuPIP23 (97.00)	AAY22203 (74.00)
ContigPIP24	VuPIP24F / VuPIP24R	160	VuPIP24 (99.00)	ADK25057 (96.00)
ContigTIP11	VuTIP11F / VuTIP11R	49	VuTIP11 (96.00)	CAD33928 (92.00)
ContigTIP22	VuTIP22F / VuTIP22R	151	VuTIP22 (94.00)	EFH59247 (67.00)
ContigTIP23	VuTIP23F / VuTIP23R	151	VuTIP23 (97.00)	AAZ78659 (89.00)
ContigNIP11	VuNIP11F / VuNIP11R	121	VuNIP11 (94.00)	ABS72446 (92.00)
ContigNIP12	VuNIP12F / VuNIP12R	65	VuNIP12 (100.00)	AES64037 (95.00)
ContigNIP71	VuNIP71F / VuNIP71R	156	VuNIP71 (100.00)	NP_001240190 (83.00)
ContigSIP11	VuSIP11F / VuSIP11R	127	VuSIP11 (91.00)	XP_003548052 (79.00)

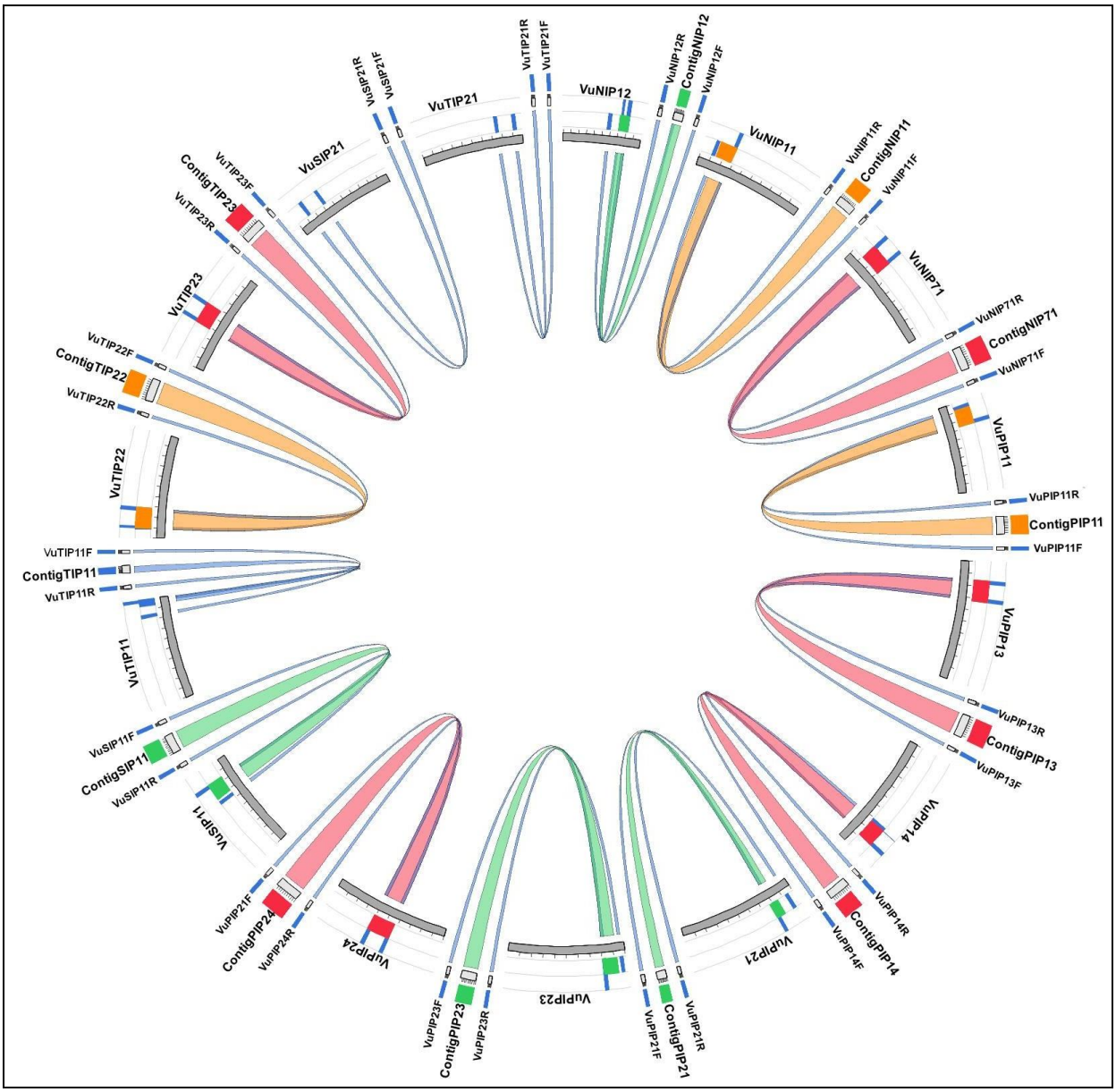


Figura 21: Visualização da similaridade dos fragmentos isolados do cDNA de V*igna unguiculata* e seus respectivos *primers* contra os transcritos caracterizados como aquaporinas no banco de dados NordEST utilizando o resultado do BLASTn no programa Circoletto. Regiões coloridas representam os níveis de similaridade baseando-se nos *scores*: Azul – score menor que 100; Verde – score entre 100 e 170; Laranja – score entre 170 e 240; Vermelho – score superior a 240.

6. DISCUSSÃO

A importância das aquaporinas como uma grande família multigência está ficando cada vez mais evidente em estudos que visam à otimização do uso da água e nutrientes nos organismos vegetais. Por se mostrar bastante diversa nas plantas, acredita-se que nela estejam elementos funcionalmente diversificados e especializados, atuantes no crescimento vegetal nas mais variadas condições ambientais. Então, o entendimento das relações hídricas vegetais a nível molecular, é dependente da identificação do conjunto total de MIPs, juntamente como suas especificidades de substratos e perfis de expressão.

O número de aquaporinas completas descrito nesse estudo é inferior ao observado em outras espécies vegetais, tal como *Arabidopsis* (35 membros) e arroz (33 membros), se devendo este fato provavelmente à indisponibilidade de sequências genômicas, ou mesmo pelas limitações do banco de EST, não abrangente a toda diversidade de transcritos existente em feijão-caupi, mas ainda assim superior ao número de isoformas identificadas em outros organismos. Este aumento do número de isoformas em organismos vegetais aponta para a amplitude de funções que estes transportadores exercem, bem como as diferentes seletividades e afinidades por solutos (LIU et al., 2009).

Apesar de possivelmente existirem outros genes codificantes para aquaporinas ainda não identificados em feijão-caupi, a proporção observada entre os membros das diferentes subfamílias (PIP, TIP, NIP e SIP) é similar ao que é encontrado em outras plantas, onde a ausência da recém descoberta subfamília XIP se deve provavelmente ao baixo nível de expressão dos membros e provavelmente a um baixo número de isoformas para este grupo, que ainda permanece desconhecido em termos de localização e especificidade de soluto (BIENERT et al., 2011).

A diversidade das aquaporinas nos organismos multicelulares reflete a necessidade destes transportadores na osmorregulação e movimento de água através das membranas nos diferentes tecidos, órgãos e estágios de desenvolvimento. A localização subcelular de todos os membros de aquaporinas em feijão-caupi foi predita como sendo de membrana plasmática, onde o compartimento subcelular não foi determinado. A abundância destes canais é importante para a

movimentação de água e de outras pequenas moléculas, facilitando as atividades regulatórias relacionadas ao transporte de água e pequenos solutos (KRUSE et al., 2006).

6.1 Avaliação de resíduos conservados

As aquaporinas possuem uma estrutura conservada dentro dos diferentes reinos (Azada et al., 2011), apresentando motivos funcionais e resíduos conservados com importante papel na atividade e seletividade do poro. A base molecular para a seletividade das aquaporinas é decorrente dos dois filtros formadores do poro, onde o primeiro corresponde ao conservado motivo NPA e o segundo é formado pela região ar/R, sendo este o principal determinante da seletividade por diferentes substratos.

Os resíduos conservados dos motivos NPA encontram-se, de um modo geral, altamente conservados nas subfamílias PIP e TIP, com substituições de resíduos encontradas apenas em NIP e SIP. Os motivos NPA para a subfamília NIP parecem apresentar pouco efeito na seletividade do substrato, observando-se que a substituição da alanina (A) por valina (V) no motivo presente no *loop* E é compatível com o transporte de formamida, uréia e glicerol, apresentando baixos níveis de transporte de água, indicando que esta substituição de resíduos afeta mais a capacidade de transporte de água, não sendo o determinante da seletividade pelo soluto (FORREST; BHAVE, 2007).

Por sua vez, as VuSIPs são divergentes quando ao motivo NPA, mostrando um assinatura diferente no *loop* B onde Asn-Pro-Ala (NPA) foi substituída por Asn-Pro-Thr em VuSIP1-1 ou Asn-Pro-Ser em VuSIP1-2, indicando que as VuSIPs podem apresentar especificidade, seletividade e eficiência de transporte diferente das demais aquaporinas (MAESHIMA; ISHIKAWA, 2008).

Os resíduos da região arginina/aromática altamente conservados têm se mostrado funcionalmente importantes na determinação da afinidade de substrato e passagem destes através das membranas biológicas (MAUREL et al., 2008). Assim, as regiões ar/R e sítios Froger's possibilitam o entendimento do espectro da atividade das aquaporinas em feijão-caupi. A região ar/R é o ponto crítico para

seletividade de transporte, as plantas geralmente possuem mais de uma composição para este filtro, indicando grande diversidade estrutural e propriedades funcionais para cada uma dessas proteínas (WALLACE; ROBERTS, 2004).

Baseando-se na homologia de sequências, a subfamília PIP corresponde ao grupo mais conservado dentro das aquaporinas, possuindo da região ar/R um resíduo de histidina (H) na posição H5, que provê uma alta especificidade pelo transporte de água graças às pontes de hidrogênio geradas, entretanto a eficiência de transporte é maior para os membros de PIP2, que formam canais de água com alta condutância (CHAUMONT et al., 2000; WALLACE; ROBERTS, 2004; FORREST; BHAVE, 2007; MAUREL et al., 2008). O fato de PIP1 e PIP2 apresentarem a mesma configuração ar/R com eficiências díspares levanta uma questão importante, que mesmo sendo a região arginina/aromática o principal determinante da afinidade de transporte, existem outras regiões que devem contribuir para o controle do transporte de água e outros pequenos solutos.

Em contraste, uma variedade de assinaturas para filtros seletivos ar/R são encontrados nas subfamílias TIP, NIP e SIP, divergindo da estrutura clássica de aquaporinas, incluindo diferenças nos diferentes grupos de uma mesma subfamília, sugerindo especializações funcionais das diferentes isoformas, onde - por exemplo - a substituição de arginina (R) por valina pode modificar a função de transporte de água (BANSAL; SANKARARAMAKRISHNAN, 2007; FORREST; BHAVE, 2007).

Na maioria das MIPs a posição H2 corresponde a um resíduo aromático hidrofóbico como triptofano (W) ou fenilalanina (F), mas para a subfamília TIP um resíduo hidrofílico ocupa esta posição, que na grande maioria dos membros de *V. unguiculata* corresponde a uma histidina (H). A histidina e a isoleucina conservadas nas posições H2 e H5 constituem a configuração de especificidade de transporte por água nesta subfamília de transportadores (WALLACE; ROBERTS, 2004). Conservações nos filtros ar/R e nos sítios Froger's nos trazem evidências que as aquaporinas pertencentes a TIP2, provavelmente compartilham similaridades funcionais, tanto pelo produto transportado, quanto pela eficiência de transporte, diferindo muito provavelmente no tecido em que se expressam.

Em *Arabidopsis* foi observada a presença de dois tipos de filtros ar/R para a subfamília NIP, também observada em feijão-caupi. A região caracterizada como Trp (W), Val (V) ou Ile (I), Ala (A) e Arg (R), identificada para as NIP1 é sabidamente

transportadora de água, glicerol, ácido lático. A segunda conformação, apresentando Thr (T) ou Ala (A), Ala (A)/Ile (I) ou Val (V), Gly (G) ou Ala (A) e Arg (R) forma poros de maior diâmetro, permeáveis a solutos maiores, assim como o silício (LIU et al., 2009).

O filtro ar/R para os membros VuNIP1-1 e VuNIP1-2 indica que estes transportadores estão relacionados ao transporte primário de glicerol, onde a arginina na posição LE2 fornece as pontes de hidrogênio para a interação com o glicerol e os demais resíduos fornecem a hidrofobicidade necessária, além de aumentar o diâmetro do poro, permitindo a passagem de solutos maiores (WALLACE et al., 2002). Segundo experimentos de Wallace et al. (2002), quando se substitui por mutação o triptofano por histidina na posição H2 o transporte passa a ser de água e não mais de glicerol.

6.2 Análises comparativas e filogenia

Para feijão-caupi, as aquaporinas agrupam-se em quatro grupos em termos de similaridade. A separação das aquaporinas vegetais em subfamílias distintas é reflexo da diversificação deste grupo em termos de especialização de função, expressão e localização. As subfamílias PIP, TIP, NIP e SIP, identificadas em feijão-caupi, são conhecidas, tendo tais categorias sido identificadas desde a briófita *Physcomitrella patens* até as plantas superiores. Diferenças nas diversas subclasses e suas subdivisões são resultado de eventos de duplicação recentes durante a evolução, evidentes em estudos filogenéticos (DANIELSON; JOHANSON, 2008).

As análises fenéticas realizadas indicam que as VuMIPs apresentam baixa divergência entre si (Figura 12), bem como entre membros deste grupo nas outras espécies analisadas (Figura 17), tanto com monocotiledôneas e dicotiledôneas, indicando que a diversificação deste grupo de transportadores é anterior a esta separação, com indicações que é anterior mesmo à separação das plantas vasculares, sendo resultado de eventos de duplicação gênica (ZARDOYA, 2005). Neste contexto, a subfamília PIP apresenta a menor divergência em termos de sequência quando comparada às demais, esta elevada similaridade mesmo nos dois

grupos obtidos (PIP1 e PIP2) se deve provavelmente à diversificação mais recente deste grupo, que possivelmente ocorreu pouco antes da separação entre monocotiledôneas e dicotiledôneas, refletindo antes modificações em sítios de nucleotídeos que diferenças na sequência protéica, visto que esta é bastante conservada dentro do grupo (JOHANSON et al., 2001). Em contraste, os grupos encontrados para a subfamília TIP descritos paras as plantas superiores, divergiram mais tardiamente, diversificando-se nas plantas vasculares, em vista de sua baixa diversidade de isoformas nos vegetais inferiores (DANIELSON; JOHANSON, 2010).

Os primeiros estudos apontavam a subfamília NIP como sendo presente apenas em plantas formadoras de nódulos radiculares. Porém, atualmente sabe-se que este grupo está presente em plantas leguminosas e não-leguminosas, apresentando diversas isoformas em todos os organismos vegetais estudados até o momento, observação que indica que a função desta subfamília não é restrita apenas aos nódulos radiculares. A presença de membros deste grupo na briófita *P. patens*, sugere que este grupo surgiu na passagem das plantas do ambiente aquático para o terrestre, com alguns autores sugerindo transferência vertical a partir das aquagliceroporinas bacterianas (WALLACE et al., 2006). Análises filogenéticas sugerem a existência de diferentes subgrupos para NIP. Em *Arabidopsis* esta subfamília apresenta sete subgrupos, em milho apresenta três, enquanto em arroz cinco grupos podem ser identificados com relação à similaridade das sequências. As vias fisiológicas em que estes grupos atuam, permanecem obscuras, mas estudos recentes (DANIELSON; JOHANSON 2011) indicam diferenças nas propriedades funcionais para esta separação.

6.3 Expressão e Localização Subcelular

Estudos anteriores sugerem que as aquaporinas são presentes em todos os tecidos vegetais, sendo reguladas espacial e temporalmente a depender do estágio de desenvolvimento e condições ambientais, sendo importantes componentes de adaptações a estresses salinos e hídricos (WALLACE et al., 2006; WUDICK et al., 2009; HEINEN; CHAUMONT, 2009). Embora a expressão das aquaporinas ocorra na maioria dos tecidos (Tabela 5 e Figura 18), a distribuição de cada isoforma de

aquaporina está controlada pelas condições fisiológicas e desenvolvimento vegetal, sugerindo papéis específicos para cada uma dessas proteínas. Por exemplo, várias aquaporinas da subfamília TIP têm padrões associados com alongamento e diferenciação das células (MAUREL, 1997). A variedade dessas isoformas e padrões de expressão pode muito bem refletir a necessidade de diferentes tipos celulares e tecidos para acomodar diferentes necessidades hídricas em determinadas condições fisiológicas.

Um perfil de expressão geral para as MIPs não pode ser efetuado, pois as diferentes isoformas podem ser super-expressas ou reprimidas a depender dos estímulos ambientais ou do órgão estudado. Esta ausência de um padrão geral de expressão é o que acaba por permitir que as plantas estejam aptas a se adaptar a condições desfavoráveis. No caso da seca e da salinidade, enquanto alguns membros são expressos para garantir o transporte de água e alguns outros solutos (KAWASAKI et al., 2001) outros são reprimidos em sua expressão (BOURSIAC et al., 2005), esta diferença sugere que as diferentes isoformas possuem vias distintas de resposta aos diferentes estresses. Além de fatores ambientais, diversos hormônios, como o ácido abscísico e giberelinas podem atuar no controle da expressão das aquaporinas nos vegetais (GOMES et al., 2009; PARENT et al., 2009).

O perfil de expressão de aquaporinas para feijão-caupi nos diferentes órgãos e estágios trazem consigo informações a respeito das vias fisiológicas que os diferentes membros atuam. Muitos transcritos mostram claramente especificidade por alguns tecidos, alguns até por certos estágios de desenvolvimento (Figura 17), essa especificidade reflete as diferentes atividades fisiológicas que as VuMIPs apresentam (SAKURAI et al., 2005). A coexistência dos membros PIP subfamília TIP levanta as questões das suas funções específicas e sua possível colaboração em diferentes condições fisiológicas (FLEURAT-LESSARD et al., 2005). PIPs e TIPs são geralmente encontradas nas membranas plasmática e do tonoplasto respectivamente, estando envolvidas diretamente em vias de diferenciação fisiológica (ZARDOYA, 2005). Apresentam expressão em folhas e raízes, sendo sua expressão em folhas menor que seu nível de expressão em raízes,variação esta relacionada diretamente com a fase de desenvolvimento vegetal (HEINEN et al., 2009).

Em *Phaseolus vulgaris,* foi observado um aumento na expressão de aquaporinas PIP no tecido foliar quando estas são submetidas à condição de seca, favorecendo a diminuição na taxa de transpiração vegetal, por controlarem a abertura e fechamento estomático (AROCA et al., 2006). Diversos estudos afirmam que níveis elevados de expressão para aquaporinas, principalmente PIPs e TIPs, estão associados com expansão e diferenciação celular, células com intensa mudança de volume (ex. células guarda) ou com intenso fluxo de água, além de células com grandes vacúolos (TYERMAN et al., 2002).

As subfamílias NIP e SIP estão mais relacionadas no que diz respeito à estrutura e função, estando ambas relacionas ao tranporte de pequenos solutos. Sendo sua expressão menos abundante nos diferentes tecidos, quando comparada às subfamílias PIP e TIP (PARK et al., 2010).

Comparando-se as aquaporinas mais expressas PIP e TIP, a subfamília NIP é geralmente expressa em baixos níveis (ALEXANDERSSON et al., 2005) e em células/órgãos especializados, como nódulos fixadores de nitrogênio nas raízes, indicando que este grupo apresenta expressão tecido-específica restrito a alguns tipos celulares vegetais (WALLACE et al., 2006). Além do transporte de água verificado nas membranas do simbiossomo, para a subfamília NIP é conhecido o transporte de glicerol, amônia, de forma semelhante às aquagliceroporinas (NIEMIETZ et al., 2000).

Os membros da subfamília SIP são membros menos abundante nos tecidos, restrindindo-se de um modo geral aos tecidos foliares (MAUREL et al., 2008). A função da subfamília SIP ainda é incerta, mas supõe-se que esta estja mais intimamente relacionada ao transporte de pequenos solutos em vias de sinalização celular, estando situada nas membranas do retículo endoplasmático (ZARDOYA, 2005). Os memebros da subfamília NIP possuem funções distintas, em leguminosas atuam no processo de formação do simbiossomo, mas pode atuar em outros processos, como a elongação e diferenciação celular.

Aquaporinas são diferencialmente expressas em membranas de tecidos específicos, possuindo nos organismos vegetais um considerável número de ortólogos, sendo altamente controladas ao nível transcricional, modificando sua expressão sob uma variedade de sinais e estresses (MAATHUIS et al., 2003).

Apesar dos métodos de predição não elucidarem a localização subcelular nas diferentes membranas internas, sabe-se que a localização das aquaporinas reflete o alto grau de compartimentalização das células vegetais, levando a uma maior necessidade de controlar o fluxo de água e pequenos solutos não somente através das membranas biológicas, mas também nas membranas internas. Em outros organismos vegetais tem-se proposto que membros da subfamília SIP encontram-se predominantemente localizados em membranas do retículo endoplasmático (MAUREL et al., 2008). Outro ponto levantado recentemente refere-se ao fato de muitos membros estarem situados em sítios diferentes dos tidos como padrão, sendo necessários estudos histoquímicos ou de fluorescências para elucidar os sítios de localização dos diversos membros das aquaporinas em feijão-caupi e outras espécies.

6.4 Modelagem comparativa

A análise da similaridade de sequências entre o modelo (1z98) e as aquaporinas de *V. unguiculata* variou de 25% à 84%. A modelagem comparativa é frequentemente restrita a sequências com similaridade acima de 30%, onde valores de identidade inferiores a este nível acarretam problemas no alinhamento do modelo com a sequência alvo, resultando em erros mais significantes na modelagem (KHALED et al., 2011). Nos nossos dados as identidades inferiores a 30% resultaram em modelos dentro dos principais parâmetros de validação, tais como Procheck, QMEAN e Z-score, indicando que por ser uma família extremamente conservada em termos de sequências, as aquaporinas – mesmo com baixa similaridade em relação ao modelo – têm na modelagem comparativa uma ferramenta eficiente na determinação da estrutura tridimensional.

Outro ponto a ser levando em consideração é o fato de que pequenas alterações ao nível de sequência secundária resultam em poucas alterações na estrutura terciária, sendo as estruturas tridimensionais de famílias protéicas mais conservadas entre si que a estrutura secundária propriamente dita. De um modo geral, alinhamentos como valores entre 30% e 50% tendem a gerar alteração nos

ângulos e posições em no máximo 3,5 Å da posição correta, enquanto que identidades acima de 50% são comparáveis a um modelo estrutural obtido por raios-X de 3 Å de resolução, podendo ser aplicados em estudos onde um modelo estrutural é requerido (ZHANG; SKOLNICK, 2005; ZHANG, 2008).

As análises via gráfico de Ramachandran mostraram que os modelos encontravam-se como pelo menos 83% dos resíduos de aminoácidos localizados nas regiões energeticamente mais favoráveis, implicando assim, que os modelos de aquaporinas gerados para feijão-caupi apresentam uma boa qualidade estereoquímica. Na análise do desvio total de energia via Z-score foi possível visualizar que os modelos VuMIPs gerados encontravam-se nas posições referentes aos modelos obtidos via raios-X e ressonância magnética (Figura 19), indicando que os modelos estão intimamente relacionados as estruturas terciárias de proteínas obtidas experimentalmente.

6.5 Validação dos transcritos

Os fragmentos obtidos após a amplificação, clonagem e sequenciamento apresentaram elevados níveis de similaridade com aquaporinas disponíveis no banco de dados do NCBI (Tabela 8), o que sugere que as regiões amplificadas correspondem a este tipo de transportador de membrana. A similaridade foi ainda maior quando o alinhamento ocorreu contra os transcritos caracterizados como aquaporinas de feijão-caupi (Figura 21), confirmando a existência destes transcritos inicialmente caracterizados apenas por ferramentas computacionais, descartando assim, a possibilidade dos transcritos identificados serem resultado de erros no processo de clusterização do banco NordEST.

Todas as subfamílias identificadas via métodos *in silico*, foram também identificadas no cDNA construído a partir do tecido radicular de indivíduos de feijão-caupi submetidos a salinidade. Segundo Vandeleur et al. (2009) a expressão de aquaporinas permite um aumento na eficiência e capacidade de transporte de água nos diferentes tecidos vegetais. No caso do estresse salino, a exposição ao sal

resulta em uma regulação diferencial da expressão. Em linhagens tolerantes, muitas aquaporinas apresentam expressão aumentada (KAWASAKI et al. 2001). Pelo uso de diferentes amostras de cDNA obtidas em diferentes tempos de desenvolvimento e sob diferentes condições de salinidade encontram-se expressos membros de todas as famílias de aquaporinas, relacionadas o transporte água e outros pequenos solutos.

Aquaporinas vegetais são modelos únicos para o estudo do transporte específico de água, possibilitando o incremento do conhecimento a respeito destes mecanismos. Os estudos *in silico* permitem a identificação de um grande número de isoformas destes transportadores em organismos vegetais, a partir de dados genômicos e de ESTs. Apesar da existência de diversos estudos baseados apenas em metodologias de bioinformática, existe a necessidade de validação destes dados por meio de técnicas moleculares, tanto no que se refere à sequência, quanto ao nível de expressão e sítios específicos de cada isoforma identificada. Estas informações contribuem para o delineamento do mecanismo de ação destes transportadores frente a estresses bióticos e abióticos, possibilitando a otimização deste processo em futuros melhoramentos.

7. CONCLUSÕES

- A cultura do feijão-caupi apresenta em seu transcriptoma representantes das quatro subfamílias de aquaporinas caracterizadas em diversos organismos vegetais, estando ausentes apenas os representantes na subfamília recém descoberta XIP, apontando para a existência de mecanismos de transporte de água e pequenos solutos similares aos processos amplamente conhecidos em plantas modelo, como *A. thaliana*.

- Considerando cada subfamília de transportadores de membrana, a maioria das proteínas apresenta alto grau de conservação, especialmente na região dos motivos e domínios, confirmando a ancestralidade destes transportadores nos principais grupos de organismos.

- No feijão-caupi a maior expressão de aquaporinas se dá em tecidos radiculares, tecido este que primeiro percebe e responde às mudanças nas condições de solo, seguido de folha, órgão responsável pelas trocas gasosas e, dependente do fluxo de água e pequenos solutos.

- O número de aquaporinas encontradas no NordEST foi inferior ao encontrado em outras espécies vegetais. Apesar disso, pôde-se perceber que o estudo baseado em no banco de ESTs disponível é satisfatório no que se refere à identificação de aquaporinas completas de todas as classes procuradas, identificando-se os principais transportadores de membranas.

- A topologia dos dendrogramas gerados com as diferentes isoformas de aquaporinas indica que o elevado número de isoformas, bem como a distribuição heterogênea destes nos organismos vegetais, é resultado de processos seletivos e adaptativos em resposta às pressões ambientais com relação à necessidade de eficiência do transporte de água e de outras moléculas. Os dados confirmam proposições de que a maioria das isoformas já existia antes da divergência evolutiva de mono e dicotiledôneas.

- Os modelos tridimensionais gerados mostraram-se satisfatórios com relação aos principais valores estatísticos, apontando para a possibilidade da utilização

destes em outras análises, tal como modelos gerados por cristalografia de raios-x ou ressonância magnética. Indicam também que as aquaporinas identificadas tem todas as estruturas necessárias para permitir sua funcionalidade, apresentando potencial para inferências biotecnológicas.

8. REFERÊNCIAS

ADAMS, M.D.; KELLEY, J.M.; GOCAYNE, J.D.; DUBNICK, M.; POLYMEROPOULOS, M.H.; XIAO, H.; MERRIL, C.R.; WU, A.; OLDE, B.; MORENO, R.F.; KERLAVAGE, A.K.; McCOMBIE, W.R.; VENTER, J.C. Complementary DNA sequencing: expressed sequence tags and human genome project. **Science**, Washington, v. 252, p. 1651–1656, 1991.

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.

AHARONI, A.; VORST, O. DNA microarrays for functional plant genomics. **Plant Molecular Biology**, New York, v. 48, p. 99–118, 2001.

ALBA, R.; FEI, Z.; PAYTON, P.; LIU, Y.; MOORE, S.L.; DEBBIE, P.; COHN, J.; D'ASCENZO, M.; GORDON, J.S.; ROSE, J.K.C.; MARTIN, G.; TANKSLEY, S.D.; BOUZAYEN, M.; JAHN, M.M.; GIOVANNONI, G. ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. **The Plant Journal**, Oxford, v. 39, n. 5, p. 697-714, 2004.

ALEXANDERSSON, E.; FRAYSSE, L.; SJOVALL-LARSEN, S.; GUSTAVSSON, S.; FELLERT, M.; KARLSSON, M.; JOHANSON, U.; KJELLBOM, P. Whole gene family expression and drought stress regulation of aquaporins. **Plant Molecular Biology**, v. 59, n. 3, p. 469-484, 2005.

ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ANDERSSON-GUNNERAS, S.; MELLEROWICZ, E.J.; LOVE, J.; SEGERMAN, B.; OHMIYA, Y.; COUTINHO, P.M.; NILSSON, P.; HENRISSAT, B.; MORITZ, T.; SUNDBERG, B. Biosynthesis of cellulose-enriched tension wood in Populus: global analysis of transcripts and metabolites identifies biochemical and

developmental regulators in secondary wall biosynthesis. **The Plant Journal**, Oxford, v. 45, p. 144-165, 2006.

APG, Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. **Botanical Journal of the Linnean Society**, v. 161, p. 105-121, 2009.

AROCA, R.; FERRANTE, A.; VERNIERI, P.; CHRISPEELS, M. J. 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. 1301-10, 2006.

AROCA, R.; PORCEL, R.; RUIZ-LOZANO, J.M. Regulation of root water uptake under abiotic stress conditions. **Journal of Experimental Botany**. First published online September 13, 2011. (acesso em 19/01/2011).

ASSAD, E.D.; PINTO, H.S.; ZULLO-JUNIOR, J.; BRUNINI, O. Mudanças climáticas: o aquecimento global e a agricultura. **Reportagens - Com Ciência**, 2002.

ATTWOOD, T.K.; BRADLEY, P.; FLOWER, D.R.; GAULTON, A.; MAUDLING, N.; MITCHELL, A.L.; MOULTON, G.; NORDLE, A.; PAINE, K.; TAYLOR, P.; UDDIN, A.; ZYGOURI, C. PRINTS and its automatic supplement, prePRINTS. **Nucleic Acids Research**, v. 31, n. 1, p. 400-402, 2003.

AYRES, R.S.; WESTCOT, D.W. **A qualidade da água na agricultura**. Tradução de GHEYI, H.R.; MEDEIROS, J.F.; DAMASCENO, F.A.V. Campina Grande, UFPB, 153p, 1999.

AZAD, A. K.; YOSHIKAWA, N.; ISHIKAWA, T.; SAWA, Y.; SHIBATA, H. 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*).

BACHEM, C.; VAN DER HOEVEN, R.; DE BRUIJN, S.; VREUGDENHIL, D.; ZABEAU, M.; VISSER, R. Visualisation of differential gene expression using a novel method of RNA finger- printing based on AFLP: analysis of gene

expression during potato tuber development. **The Plant Journal**, Oxford, v. 9, p. 745–753, 1996.

BANSAL, A.; SANKARARAMAKRISHNAN, R. Homology modeling of major intrinsic proteins in rice, maize and Arabidopsis: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. **BMC Structural Biology**, v. 7, p. 27, 2007.

BARRETO, P.D. Recursos genéticos e programa de melhoramento de feijão-de-corda no Ceará: Avanços e perspectivas. In: QUEIRÓS, M.A.; DE GOEDERT, C.O.; RAMOS, S.R.R. (eds) Recursos genéticos e melhoramento de plantas para o Nordeste **EMBRAPA**, CPATSA, Petrolina, 1999.

BAXEVANIS, A.D. Bioinformatics and the internet. **Methods Biochem Anal**, v.43, p. 1-17. 2001.

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. Estudos moleculares no caupi e em espécies relacionadas: Avanços e perspectivas. **Embrapa Documentos**, v. 56, p. 327-332, 2001.

BENKO-ISEPPON, A.M. Identificação e Validação de Genes Importantes para a Biotecnologia e Produtividade a partir do Transcriptoma do Feijão-Caupi (*Vigna unguiculata*). Projeto Submetido ao Edital MCT/CNPq/CT-AGRO/CT-BIOTEC Nº 42/2009. Programa Genoprot - Rede Integrada de Estudos Genômicos e Proteômicos, p.1-34, 2009.

BENKO-ISEPPON, A.M.; KIDO, E.A.; PANDOLFI, V. et al. Brazilian cowpea transcriptome project: over 20 million expressed sequence tags to analyze and breed salinity and virus resistance. **Resumos do III Congresso Brasileiro de Biotecnologia**, Fortaleza, CE, p. 97-98, 2010.

BENKO-ISEPPON, A.M.; SOARES-CAVALCANTI, N.M.; BELARMINO, L.C; BEZERRA-NETO J.P.; AMORIM, L.L.B.; FERREIRA-NETO, J.R.C.;

PANDOLFI, V.; AZEVEDO, H.M.A.; SILVA, R.L.O.; SANTOS, M.G.; ALVES, M.V.S.; KIDO, E.A. Prospecção de genes de resistência à seca e à salinidade em plantas nativas e cultivadas. **Revista Brasileira de Geografia Física**, 2012 (*in press*). ISSN/ISBN: 19842295

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

BERNARDO, S.; SOARES, A. A.; MANTOVANI, E. C. Manual de irrigação, **Editora UFV**, 8a edição, Viçosa, 2006, 625 p.

BETHKE, P.C.; DREW, M.C. Stomatal and non-stomatal components to inhibition of photosynthesis in leaves of *Capsicum annuum* during progressive exposure to NaCl salinity. **Plant Physiology**, Rockville, v.99, p.219–226, 1992.

BHALERAO, R.; KESKITALO, J.; STERKY, F.; ERLANDSSON, R.; BJORKBACKA, H.; BIRVE, S.J.; KARLSSON, J.; GARDESTROM, P.; GUSTAFSSON, P.; LUNDEBERG, J.; JANSSON, S. Gene expression in autumn leaves. **Plant Physiology**, v.131, n. 2, p. 430-442, 2003.

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.

BIENERT, G.P.; BINERT, M.D.; JAHN, T.P.; BOUTRY, M.; CHAUMONT, F. Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. **The Plant Journal**, 2011.

BIENERT, G.P.; MOLLER, A.L.B.; KRISTIANSEN, K.A.; SCHULZ, A.; MOLLER, I.M.; SCHJOERRING, J.K.; JAHN, T.P. 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.

BOHNERT, H.J.; AYOUBI, P.; BORCHERT, C. et al. A genomic approach towards salt stress tolerance. **Plant Physiol and Biochemistry**, v. 39, p. 295-311, 2001.

BOHNERT, H.J.; NELSON, D.E.; JENSEN, R.G. Adaptations to environmental stresses. **Plant Cell**, Rockville, v.7, n.6, p.1099-1111, 1995.

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

BORSTLAP, A.C. Early diversification of plant aquaporins. **Trends in Plant Science**, v. 7, p. 529–530, 2002.

BOTS, M.; FERON, R.; UEHLEIN, N.; WETERINGS, K.; KALDENHOFF, R.; MARIANI, T. PIP1 and PIP2 aquaporins are differentially expressed during tobacco anther and stigma development. **Journal of Experimental Botany**, v. 56, p. 113-121, 2005.

BOTS, M.; VERGELDT, F.; WOLTERS-ARTS, M.; WETERINGS, K.; VAN AS, H.; MARIANI, C. Aquaporins of the PIP2 class are required for efficient anther dehiscence in tobacco. **Plant Physiology**, v. 137, p. 1049-1056, 2005.

BOURSIAC, Y.; CHEN, S.; LUU, D.T.; SORIEUL, M.; VAN DEN DRIES, N.; MAUREL, C. 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.A. Molecular responses to water deficit. **Plant Physiology**, v. 103, p. 1035-1040, 1993.

BRAY, E.A. Plant responses to water deficit. **Trends in Plant Science**, v. 2, p.48-54, 1997.

BRAY, E.A.; BAILEY-SERRES, J.; WERETILNYK, E. Responses to abiotic stresses. In: Gruissem W Buchannan B Jones R (Eds) Biochemistry and molecular biology of plants. **American Society of Plant Physiologists**, Rockville, MD, p. 1158-1249, 2000.

BRENNER, S.; JOHNSON, M.; BRIDGHAM, J.; GOLDA, G.; LLOYD, D.H.; JOHNSON, D.; LUO, S.; MCCURDY, S.; FOY, M.; EWAN, M.; ROTH, R.; GEORGE, D.; ELETR, S.; ALBRECHT, G.; VERMAAS, E.; WILLIAMS, S.R.; MOON, K.; BURCHAM, T.; PALLAS, M.; DUBRIDGE, R.B.; KIRCHNER, J.; FEARON, K.; MAO, J.; CORCORAN, K. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. **Nature Biotechnology**, New York, v.18, n. 6, p. 630-4, 2000.

BRESSANI, R.; ELIAS L.G. Nutritional value of legume crops for humans and animals. In: SUMMERFIELD, R.J.; BUNTING, A.H. **Advances in Legume Science**, England: Royal Botanic Gardens, n. 1, p. 135-155, 1980.

BROWN, R.L.; KAZAN, K.; MCGRATH, K.C. et al. A role for the GCC-box in jasmonate-mediated activation of The PDF1.2 gene of Arabidopsis. **Plant Physiology**, v. 132, p. 1020-1032, 2003.

CARNEIRO, P.T.; FERNANDES, P.D.; GHEYI, H.R.; SOARES, F.A.L. Germination and initial growth of precocious dwarf cashew genotypes under saline conditions. **Revista Brasileira de Engenharia Agrícola**, Campina Grande, v.6, p.199-206, 2002.

CARRARO, D.M.; KITAJIMA, J.P. Sequenciamento e bioinformática de genomas bacterianos. **Biotecnologia Ciência e Desenvolvimento**, v. 28, p. 16-20, 2002.

CASTRO, N.R. Caracterização fisiológica de *Cercospora cruenta* Sacc. e controle genético de cercosporiose em caupi. (Tese de Mestrado). 48 f. Recife. **Universidade Federal Rural de Pernambuco**. 2000.

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.

CHAVES, M.M.; MAROCO, J.P.; PEREIRA, J.S. Understanding plant responses to drought - from genes to the whole plant. **Functional Plant Biology,** v. 30, p. 239-264, 2003.

CHEN, S.; HARMON, A.C. Advances in plant proteomics. **Proteomics**, Weinheim, v. 6, n. 20, p. 5504-16, 2006.

CHEN, X.; LAUDEMAN, T.W.; RUSHTON, P.J.; SPRAGGINS, T.A.; TIMKO, M.P. Database CGKB: an annotation knowledge base for cowpea (*Vigna unguiculata* L.) methylation filtered genomic genespace sequences. **BMC Bioinformatics**, v. 8, p. 129, 2007.

CHINNUSAMY, V.; SCHUMAKER, K.; ZHU, J.K. Molecular genetic perspectives on cross-talk and specificity in abiotic stress sinalling in plants. **Journal of Experimental Botany,** v. 55, n. 395, p. 225-236, 2004.

CHOI, W.G.; ROBERTS, D.M. Arabidopsis NIP2;1: A major intrinsic protein transporter of lactic acid induced by anoxic stress. **Journal of Biological Chemistry**, v. 282, n. 33, p. 24209-24218, 2007.

COEMANS, B.; MATSUMURA, H.; TERAUCHI, R.; REMY, S.; SWENNEN, R.; SAGI, L. SuperSAGE combined with PCR walking allows global gene expression profiling of banana (*Musa acuminata*), a non-model organism. **Theoretical and Applied Genetics**, Berlin, v. 111, n. 6, p. 1118-26, 2005.

COSGROVE, D.J. Water uptake by growing cells: an assessment of the controlling roles of wall relaxation, solute uptake and hydraulic conductance. **International Journal of Plant Science**, v. 154, p. 10-21, 1993.

CUSHMAN, J.C.; BOHNERT, H.J. Genomic approaches to plant stress tolerance. **Current Opinion in plant Biology**, v. 3, p. 117-124, 2000.

DANIELS, M.J.; MIRKOV, T.E.; CHRISPEELS, M.J. The plasma membrane of *Arabidopsis thaliana* contains a mercury-insensitive aquaporin that is homolog of the tonoplast water channel protein TIP. **Plant Physiology**, v. 106, p. 1325-1333, 1994.

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.

DARZENTAS, N. Circoletto: visualizing sequence similarity with Circos. **Bioinformatics (Oxford, England)**, v. 26, n. 20, p. 2620-2621, 2010.

DELOUKAS, P. et al. A physical map of 30.000 human genes. **Science**, v. 282, p. 744-746, 1998.

DIATCHENKO, L.; LAU, Y.F.; CAMPBELL, A.P.; CHENCHIK, A.; MOQADAM, F.; HUANG, B.; LUKYANOV, S.; LUKYANOV, K.; GURSKAYA, N.; SVERDLOV, E.D.; SIEBERT, P.D. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, n. 12, p. 6025-30, 1996.

DONSON, J.; FANG, Y.; ESPIRITU-SANTO, G.; XING, W.; SALAZAR, A.; MIYAMOTO, S.; ARMENDAREZ, V.; VOLKMUTH, W. Comprehensive gene expression analysis by transcript profiling. **Plant Molecular Biology**, Dordrecht, v. 48, p. 75-97, 2002.

DRMANAC, R.; DRMANAC, S. cDNA screening by array hybridization. **Methods in Enzymology**, New York, v. 303, p. 165–178, 1999.

DYNOWSKI, M.; SCHAAF, G.; LOQUE, D.; MORAN, O.; LUDEWIG, U. Plant plasma membrane water channels conduct the signalling molecule H2O2. **Biochemical Journal**, v. 414, p. 53-61, 2008.

ECK, R.V.; DAYHOFF, M.O. Atlas of protein sequence and structure. **Silver Spring**, p. 1-68, 1966.

EDDY, S. R. A new generation of homology search tools based on probabilistic inference. **Genome informatics. International Conference on Genome Informatics**, v. 23, n. 1, p. 205-11, 2009.

EDDY, S. R. Hidden Markov models. **Current Opinion in Structural Biology**, v. 6, p. 361–365, 1996.

EDWARDS, D.; BATLEY, J. Plant bioinformatics: from genome to Phenome. **Trends in Biotechnology**, v. 22, n. 5, p. 232-237, 2004.

EHLERS, J.D.; HALL, A.E. Cowpea (*Vigna unguiculata* Walp L.). **Field Crops Research**, v. 53, p. 187-204, 1997.

EISEN, M.B.; SPELLMAN, P.T.; BROWN, P.O.; BOTSTEIN, D. Cluster analysis and display of genomic-wide expression pattern. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 14863-14868, 1998.

ENGEL, A.; FUJIYOSHI, Y.; AGRE, P. The importance of aquaporina water channel protein structures. **The EMBO Journal**, v. 19, n. 5, p. 800-806, 2000.

ESWAR, N.; ERAMIAN, D.; WEBB, B.; SHEN, M.Y.; SALI, A. Protein Structure Modeling with Modeller. **Methods in Molecular Biology**, v. 426, p. 145-159, 2008.

FALQUET, L.; PAGNI, M.; BUCHER, P.; HULO, N.; SIGRIST, C.J.; HOFMANN, K.; BAIROCH, A. The PROSITE database, its status in 2002. **Nucleic Acids Research**, v. 30, n. 1, p. 235-238, 2002.

FARIS, D.G. The origin and evolution of the cultivated forms of *Vigna sinensis*. **Canadian Journal of Genetis and Cytology**, Ottawa, n. 7, p. 433-452, 1965.

FERREIRA-NETO, J.R.C.; BARBOSA, P.K.A.; PANDOLFI, V.; SILVA, R.L.O.; ARAÚJO. F.; ANDRADE, P.P.; NASCIMENTO, L.C.; CARAZZOLLE, M.F.; PEREIRA, G.G.; BENKO-ISEPPON, A.M.; KIDO, E.A. SuperSAGE analysis of soybean drought stress-responsive transcriptome. **Genetics and Molecular Biology**, 2012 (*in press*).

FERY, R.L. The cowpea: production, utilization and research in the United States. **Horticultural Reviews**, v. 12, p. 197-222, 1990.

FITCH, W.M. Toward defining the course of evolution: minimum change for a specific tree topology. **Systematic Zoology**, v. 20, n. 406-416, 1971.

FIZAMES, C.; MUÑOS, S.; CAZETTES, C.; NACRY, P.; BOUCHEREZ, B.; GAYMARD, F.; PIQUEMAL, D.; DELORME, V.; COMMES, T.; DOUMAS, P.; COOKE, R.; MARTI, J.; SENTENAC, H.; GOJON, A. The Arabidopsis root transcriptome by Serial Analysis of Gene Expression. Gene identification using the genome sequence. **Plant Physiology**, London, v. 134, p. 67-80, 2004.

FLEISCHMANN, R.D.; ADAMS, M.D.; WHITE,O.; CLAYTON, R.A.; KIRKNESS, E.F.; KERLAVAGE, A.R.; BULT, C.J.; TOMB, J.F.; DOUGHERTY, B.A.; MERRICK, J.M.; ET AL. Whole-genome random sequencing and assembly of *Haemophilus influenza.* **Rd Science**, Washington, v. 269, n. 5223, p. 496-512, 1995.

FLEURAT-LESSARD, P.; MICHONNEAU, P.; MAESHIMA, M.; DREVON, J.J.; SERRAJ, R. The distribution of aquaporin subtypes (PIP1, PIP2 and $_{\gamma}$-TIP) is tissue dependent in soybean (*Glycine max*) root nodules. **Annals of Botany**, v. 96, p. 457-460, 2005.

FLEXAS, J.; RIBAS-CARBO, M.; HANSON, D.T.; BOTA, J.; OTTO, B.; CIFRE, J.; MCDOWELL, N.; MEDRANO, H.; KALDENHOFF, R. Tobacco aquaporin NtAQP1 is involved in mesophyll conductance to CO2 in vivo. **Plant Journal**, v. 48, n.3, p. 427-439, 2006.

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.

FOUQUET, R.; LÉON, C.; OLLAT, N.; BARRIEU, F. Identification of grapevine aquaporins and expression analysis in developing berries. **Plant Cell Reports**, v.27, p. 1541-1550, 2008.

FREIRE-FILHO, F.R.. Cowpea taxonomy and introduction to Brazil. In: WATT, E.E.; ARAÚJO, J.P.P. (eds) Cowpea research in Brazil IITA, **EMBRAPA**, Brasília, p. 3-10, 1988.

FREIRE-FILHO, F.R.; RIBEIRO, V.Q.; BARRETO, P.D.; SANTOS, C.A.F. Melhoramento genético do caupi (Vigna unguiculata (L.) Walp.) na região nordeste. In: QUEIRÓS, M.A.; GOEDERT, C.O.; RAMOS, S.R.R. (eds). Recursos genéticos e melhoramento de plantas para o Nordeste Brasileiro, **EMBRAPA**, 1999.

FREIRE-FILHO, F.R.; ROCHA, M.M.; RIBEIRO, V.Q.; LOPES, A.C.A. Adaptabilidade e estabilidade produtiva de feijão-caupi. **Ciência Rural**, v. 35, n. 24-30, 2005.

FROTA, K.M.G.; MORGANO, M.A.; SILVA, M.G.; ARAÚJO, M.A.M.; MOREIRA-ARAÚJO, R.S.R. Utilização da farinha de feijão-caupi (*Vigna unguiculata* L. Walp) na elaboração de produtos de panificação. **Ciência e Tecnologia de Alimentos**, v. 30, p. 44-50, 2009.

GAO, Y-P.; YOUNG, L; BONHAM-SMITH, P; GUSTA L.V. Characterization and expression of plasma and tonoplast membrane aquaporins in primed seed of Brassica napus during germination under stress conditions. **Plant Molecular Biology**, v. 40, p. 635-644, 1999.

GE, X.; WU, Q.; JUNG, Y.C.; CHEN, J.; WANG, S.M. A large quantity of novel human antisense transcripts detected by LongSAGE. **Bioinformatics**, Oxford, v. 22, n. 20, p. 2475-2479, 2006.

GHAFOOR, A.; SHARIF, A.; AHMAD, Z.; ZAHID, M. A.; RABBANI, M. A. Genetic diversity in blackgram (*Vigna mungo* L. Hepper). **Field Crops Research**, n.69, p.183- 190, 2001.

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

GIBBINGS, J.G.; COOK, B.P.; DUFAULT, M.R.; MADDEN, S.L.; JHURI, S.; TURNBULL, C.J.; DUNWELL, J.M. Global transcript of rice leaf and seed using

SAGE technology. **Plant Biotechnology Journal**, Oxford, v. 1, p. 271-285, 2003.

GOMES, D.; AGASSE, A.; THIEBAUD, P.; DELROT, S.; GEROS, H.; CHAUMONT, F. 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.

GUSTAVSSON, S.; LEBRUN, A.S.; NORDEN, K.; CHAUMONT, F.; JOHANSON, U. A novel plant major intrinsic protein in Physcomitrella patens most similar to bacterial glycerol channels. **Plant Physiology**, v. 139, p. 287-295, 2005.

HACHEZ, C.; HEINEN, R.B.; DRAYE, X.; CHAUMONT, F. The expression pattern of plasma membrane aquaporins in maize leaf highlights their role in hydraulic regulation. **Plant Molecular Biology**, v. 68, p. 337–353, 2008.

HACHEZ, C.; MOSHELION, M.; ZELAZNY, E.; CAVEZ, D.; CHAUMONT, F. Localization and quantification of plasma membrane aquaporin expression in maize primary root: a clue to understanding their role as cellular plumbers. **Plant Molecular Biology**, v. 62, p. 305–323, 2006.

HACHEZ, C.; ZELAZNY, E.; CHAUMONT, F. Modulating the expression of aquaporin genes in planta: a key to understand their physiological functions? **Biochimica et Biophysica Acta,** v. 1758, p. 1142–1156, 2006.

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.

HARVEST (Harvest Database), Disponível em: <http://www.harvest.ucr.edu/>. Acessado em: 15 mar 2010.

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

HEYMANN, J.B.; ENGEL, A. Aquaporins: Phylogeny, Structure, and Physiology of Water Channels. **News Physiology Science**, v. 14, p. 187-193, 1999.

HILLER, K.; GROTE, A.; MANECK, M.; MÜNCH, R.; JAHN, D. 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, out 2006.

HIRAYAMA, T.; SHINOZAKI, K. Research on plant abiotic stress responses in the post- genome era: past, present and future. **Plant Journal**, v. 61, p. 1041–1052, 2010.

HOPKINS, G.W. Introduction to plant physiology, **New York: John Wiley & Sons**, 512 p, 1999.

HORTON, P.; PARK, K.-J.; OBAYASHI, T.; et al. WoLF PSORT: protein localization predictor. **Nucleic acids research**, v. 35, p. 585-587, jul 2007.

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.

IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em: <http://www.ibge.com.br>. Acessado em: 15 mar 2010.

JAHN, T.P.; MOLLER, A.L.; ZEUTHEN, T.; HOLM, L.M.; KLAERKE, D.A.; MOHSIN, B.; KUHLBRANDT, W.; SCHJOERRING, J.K. Aquaporin homologues in plants and mammals transport ammonia. **FEBS Letters**, v. 574, n. 1–3, p. 31-36, 2004.

JAVOT, H.; MAUREL, C. The role of aquaporins in rootwater uptake, **Annals of Botany**, v. 90, n. 301–313, 2002.

JOHANSON, U.; KARLSSON, M.; JOHANSSON, I.; GUSTAVSSON, S.; SJOVALL, S.; FRAYSSE, L.; WEIG, A.R.; KJELLBOM, P. 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.

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

KALDENHOFF, R.; GROTE, K.; ZHU, J.-J.; ZIMMERMANN, U. Significance of plasmalemma aquaporins for water-transport in *Arabidopsis thaliana*. **Plant Journal**, v. 14, p. 121-128, 1998

KATSUHARA, M.; KOSHIO, K.; SHIBASAKA, M.; HAYASHI, Y.; HAYAKAWA, T.; KASAMO, K. Over-expression of a barley aquaporin increased the shoot/root ratio and raised salt sensitivity in transgenic rice plants. **Plant Cell Physiology**, v. 44, p. 1378-1383, 2003.

KAWASAKI, S.; BORCHERT, C.; DEYHOLOS, M.; WANG, H.; BRAZILLE, S.; KAWAI, K.; GALBRAITH, D.; BOHNERT, H.J. Gene expression profiles during the initial phase of salt stress in rice. **Plant Cell**, v. 13, p. 889–905, 2001.

KIDO, EA.; PANDOLFI, V.; FERREIRA-NETO, J.R.C.; BARBOSA, P.K.A.; SILVA, R.L.O.; NASCIMENTO, L.C.; CARAZZOLLE, M.F.; PEREIRA, G.G.; BENKO-ISEPPON, A.M. Global SuperSAGE transcriptional profile of Phakopsora pachyrhizi-infected soybean. **Genetics and Molecular Biology**, 2012 (*in press*).

KENT, W.J.; SUGNET, C.W.; FUREY, T.S.; ROSKIN, K.M. et al. The human genome browser at UCSC. **Genome Research**, v. 12, p. 996-1006, 2002.

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

KHALED, K.; KARIM, K.; MOURAD, O.; ABDEKARIM, K.; NOUREDINE, R. PIP Aquaporin Protein of in Durum Wheat, Homology Modelling and 3D Structure Analysis. **BIOMIRROR**, v. 2, p. 1-7, 2011.

KLAR, A.E. Irrigação: Frequência e quantidade de aplicação, **Editora Nobel**, São Paulo, 1991, 156 p.

KNIGHT, H.; KNIGHT, M,R. Abiotic stress signalling pathways: specificity and cross-talk. **Trends in Plant Science**, v. 6, n. 262–267, 2001.

KRAMER, P.J.; BOYER, J.S. **Water relations of plants and soils.** San Diego: Academic, 1995. 495p

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

KUMAR, S.; TAMURA, K.; NEI, M. MEGA 3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. **Briefings in Bioinformatics**, v. 5, p. 150-163, 2004.

LACERDA, C.F.; CAMBRAIA, J.; CANO, M.A.O.; RUIZ, H.A. Plant growth and solute accumulation and distribution in two sorghum genotypes, under NaCl stress. **Revista Brasileira de Fisiologia Vegetal**, Lavras, v.13, p. 270-284, 2001.

LANGYINTUO, A.S.; LOWENBERG-DEBOER, J.; FAYE, M.; LAMBERT, D.; IBRO, G.; MOUSSA, B.; KERGNA, A.; KUSHWAHA, S.; MUSA, S; NTOUKAM, G. Cowpea supply and demand in west and central Africa. **Field Crops Research**, v. 82, n. 2-3, p. 215- 231, 2003.

LASH, A.E.; TOLSTOSHEV, C.M.; WAGNER, L.; SCHULER, G.D.; STRAUSBERG, R.L.; RIGGINS, G.J.; ALTSCHUL, S.F. SAGEmap: A public gene expression resource. **Genome Research**, v. 10, p. 1051-1060, 2000.

LASKOWSKI, R.A.; MACARTHUR, M.W.; MOSS, D.S.; THORNTON, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283-291, 1993.

LIANG, P.; PARDEE, A. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. **Science**, Washington, v. 257, p. 967–971, 1992.

LIMA, C.J.G.S.; OLIVEIRA, F.A; MEDEIROS, J.F.; OLIVEIRA, M.K.T.; JÚNIOR, A.B.A. Resposta do feijão caupi a salinidade da água de irrigação. **Revista Verde**, v. 2, n. 2, p. 79–86, 2007.

LIU, L.H.; LUDEWIG, U.; GASSERT, B.; FROMMER, W.B.; VON WIREN, N. Urea transport by nitrogen-regulated tonoplast intrinsic proteins in Arabidopsis. **Plant Physiology**, v. 133, n. 3, p. 1220-1228, 2003.

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, F.; BOUSSER, A.; SISSOËFF, I.; GASPAR, M.; LACHAISE, B.; HOARAU, J.; MAHÉ, A. Diurnal regulation of water transport and aquaporin gene expression in maize roots: contribution of PIP2 proteins. **Plant Cell Physiology** v. 44, p. 1384–1395, 2003.

LOQUE, D.; LUDEWIG, U.; YUAN, L.; VON WIREN, N. Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH3 transport into the vacuole. **Plant Physiology**, v. 137, n. 2, p. 671-680, 2005.

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.

MA, J.F; TAMAI. K.; YAMAJI, N.; MITANI, N.; KONISHI, S.; KATSUHARA, M.; ISHIGURO, M.; MURATA, Y.; YANO, M. A silicon transporter in rice. **Nature**, v. 440, n. 7084, p. 688-691, 2006.

MAATHUIS, F.J.M.; FILATOV, V.; HERZYK, P.; KRIJGER, G.C.; AXELSEN, K.B.; CHEN, S.; GREEN, B.J.; LI, Y.; MADAGAN, K.L.; SANCHEZ-FERNANDEZ, R.; FORDE, B.G.; PALMGREN, M.G.; REA, P.A.; WILLIAMS, L.E.; SANDERS, D.; AMTMANN, A. Transcriptome analysis of root transporters reveals participation of multiple gene families in the response to cation stress. **Plant Journal**, v. 35, p. 675–692, 2003.

MAESHIMA, M.; ISHIKAWA, F. ER membrane aquaporins in plants. **European journal of physiology**, v. 456, n. 4, p. 709-16, 2008.

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

MAIA, F.M.M.; OLIVEIRA, J.T.A.; MATOS, M.R.T.; MOREIRA, R.A.; VASCONCELOS, I.M. Proximate Composition, Amino Acid Content and Haemagglutinating and Trypsin Inhibiting Activities of Some Brazilian *Vigna unguiculata* (L.) Walp. cultivars. **Journal of Science and Food Agriculture**, n. 80, p. 453-458, 2000.

MALONE, G.; ZIMMER, P.D.; MENEGHELLO, G.E.; BINNECK, E.; PESKE, S.T. Prospecção de Genes em Bibliotecas de cDNA. **Revista Brasileira de. Agrociência**, Pelotas, v. 12, n.1, p. 7-13, 2006.

MANSOUR, M.M.F.; SALAMA, K.H.A.; Al-MUTANA, M.M. Transport protein and salt tolerance in plants. **Plant Science**, Limerik, v. 146, p. 891-900, 2003.

MARTRE, P.; MORILLON, R.; BARRIEU, F.; NORTH, G.B.; NOBEL, P.S; CHRISPEELS, M.J. Plasma membrane aquaporins play a significant role during recovery from water deficit. **Plant Physiology**, n. 130, p. 2101-2110, 2002.

MATSUMURA, H.; NIRASAWA, S.; TERAUCHI, R. Transcript profiling in rice (*Oryza sativa* L.) seedlings using the serial analysis of gene expression (SAGE). **The Plant Journal**, Oxford, v. 20, n. 6, p. 719-726, 1999.

MATSUMURA, H.; REICH, S.; ITO, A.; SAITOH, H.; KAMOUN,S.; WINTER, P.; KALS, G.; REUTER, M.; KRÜGER, D.H., TERAUCHI, R. Gene expression analysis of plant host-pathogen interactions by SuperSAGE. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 100, n. 26, p. 15718-15723, 2003.

MAUREL, C. Aquaporins and Water Permeability of Plant Membranes. **Annual review of plant physiology and plant molecular biology**, v. 48, p. 399-429, 1997.

MAUREL, C. Plant aquaporins: novel functions and regulation properties. **FEBS Letters** v. 581, p. 2227–2236, 2007.

MAUREL, C.; REIZER, J.; SCHROEDER, J.I.; CHRISPEELS, M.J. The vacuolar membrane protein γ-TIP creates water specific channels in *Xenopus* oocytes. **EMBO Journal**, v. 12, p. 2241–2247, 1993.

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.

MCKAY, J.K.; RICHARDS, J.H.; MITCHELL-OLDS, T. Genetics of drought adaptation in *Arabidopsis thaliana*: I pleiotropy contributes to genetic correlations among ecological traits. **Molecular Ecology**, v. 12, p. 1137-1151, 2003.

MEYERS, B.C.; GALBRAITH, D.W.; NELSON, T.; AGRAWAL, V. Methods for Transcriptional Profiling in Plants. Be Fruitful and Replicate. **Plant Physiology**, Lancaster, v. 135, p. 637-652, 2004.

MITTLER, R. Abiotic stress, the field environment and stress combination. **Trends In Plant Science**, v. 11, p. 15–19, 2006.

MOLINA, C.; ROTTER, B.; HORRES, R.; UDUPA, S.M.; BESSER, B.; BELLARMINO, L.; BAUM, M.; MATSUMURA, H.; TERAUCHI, R.; KAHL, G.; WINTER, P. SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. **BMC Genomics**, v. 9, p. 553, 2008.

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

MOREAU, C.; AKSENOV, N.; LORENZO, M. G.; SEGERMAN, B.; FUNK, C.; NILSSON, P.; JANSSON, S.; TUOMINEN, H. A genomic approach to investigate developmental cell death in woody tissues of Populus trees. **Genome Biology**, London, v. 6, n. 4, p. 1-14, 2005.

MUCHERO, W.; DIOP, N.N.; BHAT, P.R. et al. A consensus genetic map of cowpea [*Vigna unguiculata* (L.) Walp.] and synteny based on EST-derived SNPs.

Proceedings of the National Academy of Sciences of the United States of America, v. 106, n. 43, p. 18159–18164, 2009.

MULDER, N.J.; APWEILER, R.; ATTWOOD, T.K.; BAIROCH, A.; BATEMAN, A.; BINNS, D.; BISWAS, M.; BRADLEY, P.; BORK, P.; BUCHER, P.; COPLEY, R.; COURCELLE, E.; DURBIN, R.; FALQUET, L.; FLEISCHMANN, W.; GOUZY, J.; GRIFFITH-JONES, S.; HAFT, D.; HERMJAKOB, H.; HULO, N.; KAHN, D.; KANAPIN, A.; KRESTYANINOVA, M.; LOPEZ, R.; LETUNIC, I.; ORCHARD, S.; PAGNI, M.; PEYRUC, D.; PONTING, C.P.; SERVANT, F.; SIGRIST, C.J. InterPro: an integrated documentation resource for protein families, domains and functional sites. **Briefings in Bioinformatics**, v. 3, n. 3, p. 225-235, 2002.

MURATA, K.; MITSUOKA, K.; HIRAI, T.; WALZ, T.; AGRE, P.; HEYMANN, J.B.; ENGEL, A.; FUJIYOSHI, Y. Structural determinants of water permeation through aquaporin-1. **Nature**, v. 407, n. 599–605, 2000.

NCBI, *National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica. Disponível em <http://www.ncbi.nlm.nih.gov/protein>. Acesso em: 3 nov 2011.

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

NG, N.Q.; MARÉCHAL, R. Cowpea taxonomy, origin germplasm. In: SINCH, S.R.; RACHIE, K.O. eds. **Cowpea research, production and utilization.** Cheichecter, John Wiley. p. 11-21, 1985.

NGUYEN, H.T.; BABU, R.C.; BLUM, A. Breeding for drought resistance in rice: physiology and molecular genetics considerations. **Crop Science**, v. 37, p. 1426- 1437, 1997.

NIEMIETZ, C.M.; TYERMAN, S.D. Channel-mediated permeation of ammonia gas through the peribacteroid membrane of soybean nodules. **FEBS Letters**, v. 465, p. 110–114, 2000.

NOGUEIRA, R.M.C.; ARAÚJO, E.L.; WILLADINO, L.G.; CAVALCANTE, U.M.T. Estresses ambientais: danos e benefícios em plantas. **Imprensa Universitária**, Recife, 2005

OLIVEIRA, I.P.; CARVALHO, A.M. A cultura do caupi nas condições de clima e solo dos trópicos úmidos de semi-árido do Brasil. In: ARAUJO, J.P.P.; WATT. E.E. (Org.) O caupi no Brasil. Brasília, DF: **Embrapa-CNPAF**; Ibadan: IITA, 1998, p.63-96.

ORCUTT, D.M.; NILSEN, E.T. The physiology of plants under stress-soil and biotic factors. **New York: John Wiley and Sons**, 2000.

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.

PAIVA, J.B.; CARMO, C.M.; TAVORA, F.J.A.; ALMEIDA, F.G. et al. Melhoramento, experimentação e fitossanidade com feijão (*Vigna sinensis*), realizadas no estado do Ceará (1967/68). **Pesquisa Agropecuária do Nordeste**, v. 2, n. 2, p. 99-113, 1970.

PAIVA, R.; OLIVEIRA, L.M. Fisiologia e Produção Vegetal. Lavras. **Ed. UFLA**, 2006. 104p.

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

PANDEY, G.K. Emergence of a novel calcium signaling pathway in plants: CBL-CIPK signaling network. **Physiology and Molecular Biology of Plants**, v. 14, p. 51-68, 2008.

PARENT, B.; HACHEZ, C.; REDONDO, E.; SIMONNEAU, T. Drought and Abscisic Acid Effects on Aquaporin Content Translate into Changes in Hydraulic Conductivity and Leaf Growth Rate: A Trans-Scale Approach 1 [W][OA]. **Growth (Lakeland)**, v. 149, p. 2000-2012, 2009.

PARK, W.; SCHEFflER, B.E.; BAUER, P.J.; CAMPBELL, B.T. Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). **BMC Plant Biology**, v. 10, p. 142, 2010.

PENG, Y.H.; LIN, W.L.; CAI, W.M.; ARORA, R. Overexpression of a *Panax ginseng* tonoplast aquaporin alters sat tolerance, drought tolerance and cold acclimation ability in transgenic Arabidopsis plants. **Planta**, v. 226, p. 729-740, 2007.

PHILLIPS, R.D.; MCWATTERS, K.H.; CHINNAN, M.S.; HUNG, Y.B.; LARRY, R.; SEFA-DEDEH, S.; SAKYI-DAWSON, E.; NGODDY, P.; NNANYELUGO, D.; ENWERE, J.; KOMEY, N.S.; LIU, K.; MENSA-WILMOT, Y.; NNANNA, I.A.; OKEKE, C.; PRINYAWIWATKUL, W.; SAALIA, F.K. Utilization of cowpeas for human food. **Field Crops Research**, v. 82, n. 2-3, p. 193-213, 2003.

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

PLEASANCE, E.D.; MARRA, M.A.; JONES, S.J.M. Assessment of SAGE in transcript identification. **Genome Research**, Cold Spring Harbor, v. 13, p. 1203-1215, 2003.

PRESTON, G.M.; AGRE, P. Isolation of the cDNA for erythrocyte integral membrane protein of 28 kilodaltons: member of an ancient channel family. **Proceedings of the National Academy of Sciences of the United States of America**, v. 88, p. 11110–11114, 1991.

PRINYAWIWATKUL, W.; MCWATTERS, K.H.; BEUCHAT, L.R.; PHILLIPS R.D. Cowpea flour: A potential ingredient in food products. **Critical Reviews in Food Sciences and Nutrition**, n. 36, p. 413-436, 1996.

PROSDOCINI, F.; CERQUEIRA, G.C.; BINNECK, E.; SILVA, A.F. et al. Bioinformática: manual do usuário. **Biotecnolia Ciência e Desenvolvimento**, v. 29, n. 12-25, 2002.

QUIGLEY, F.; ROSENBERG, J. M.; SHACHAR-HILL, Y.; BOHNERT, H. J. From genome to function: the Arabidopsis aquaporins. **Genome biology**, v. 3, n. 1, p. 1-17, 2001.

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

REYNOLDS, S.M.; KÄLL, L.; RIFFLE, M.E.; BILMES, J.A.; NOBLE, W.S. Transmembrane Topology and Signal Peptide Prediction Using Dynamic Bayesian Networks. **PLoS Computational Biology**, v. 4, n. 11, p. e1000213, 2008.

ROBINSON, S.J.; CRAM, D.J.; LEWIS, C.T.; PARKIN, I.A.P. Maximizing the efficacy of SAGE analysis identifies novel transcripts in Arabidopsis. **Bioinformatics**, Oxford, v. 136, p. 3223-3233, 2004.

ROUZÉ, P.; PAVY, N.; ROMBAUTS, S. Genome annotation: witch tools do we have for it? **Current Opinion of Plant Biology**, v. 2, n. 90-95, 1999.

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.

SABBATINI, R.M.E. A bioinformática chegou. Rev Info Médica, v. 2, p. 4-6, 1999.

SADE, N.; VINOCUR, B.J.; DIBER, A.; SHATIL, A.; NISSAN, H.; WALLACH, R.; KARCHI, H; MOSHELION, M. Improving plant stress tolerance and yield production: is the tonoplast aquaporin SITIP2;2 a key to isohydric to anisohydric conversion. **New Phytologist**, v. 181, p. 651-661, 2009.

SAHA, S.; SPARKS, A.B.; RAGO, C.; AKMAEV, V.; WANG, C.J.; VOGELSTEIN, B.; KINZLER, K.W.; VELCULESCU, V.E. Using the transcriptome to annotate the genome. **Nature Biotechnology**, v.20, p.508-512. 2002.

SAHNOUN, I.; DÉHAIS, P.; MONTAGU, M.V.; ROSSIGNOL, M.; ROUZÉ, P. PPMdb: a plant plasma membrane database. **Journal of Biotechnology**, v. 78, p. 235-246, 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.; ISHIKAWA, F.; YAMAGUCHI, T.; UEMURA, M.; MAESHIMA, M. Identification of 33 rice aquaporin genes and analysis of their expression and function. **Plant and Cell Physiology**, v. 46, n. 1568-1577, 2005.

SANTONI, V.; VERDOUCQ, L.; SOMMERER, N.; et al. Methylation of aquaporins in plant plasma membrane. **The Biochemical journal**, v. 400, n. 1, p. 189-97, 2006.

SARDA, X.; TOUSCH, D.; FERRARE, K.; LEGRAND, E.; DUPUIS, J.M.; CASSE-DELBART, F; LAMAZE, T. Two TIP-like genes encoding aquaporins are expressed in sunflower guard cells. **Plant Journal**, v. 12, p. 1103-1111, 1997.

SCHUURMANS, J.A.; VAN DONGEN, J.T.; RUTJENS, B.P.; BOONMAN, A.; PIETERSE, C.M; BORSTLAP, A.C. Members of the aquaporin family in the developing pea seed coat include representatives of the PIP, TIP, and NIP subfamilies. **Plant Molecular Biology**, v. 53, p. 633-645, 2003.

SEKI, M.; NARUSAKA, M.; ISHIDA, J.; NANJO, T.; FUJITA, M.; OONO, Y.; KAMIYA, A.; NAKAJIMA, M.; ENJU, A.; SAKURAI, T.; SATOU, M.; AKIYAMA, K.; TAJI, T.; YAMAGUCHI-SHINOZAKI, K.; CARNINCI, P.; KAWAI, J.; HAYASHIZAKI, Y.; SHINOZAKI, K. Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold, and high-salinity stresses using a full-length cDNA microarray. **Plant Journal**, v. 31, p. 279-292, 2002.

SHANNON, M.C.; GRIEVE, C.M. Tolerance of vegetable crops to salinity. **Scientia Horticulturae**, Amsterdam, v. 78, p. 5- 38, 1998.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene networks involved in drought stress response and tolerance. **Journal of Experimental Botany**, v. 58, p. 221-227, 2007.

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.

SIEFRITZ, F.; TYREE, M.T.; LOVISOLO, C.; SCHUBERT, A; KALDENHOFF, R. PIP1 plasma membrane aquaporins in tobacco: from cellular effects to function in plants. **Plant Cell**, v. 14, p.869-876, 2002.

SILVA, K.J.D. **Estatística da produção de feijão-caupi**. 2009. Disponível em: <http://www.portaldoagronegocio.com.br/conteudo.php?id=34241>. Acesso em: 03 mar. 2010.

SILVA, L.B.; FREITAS, H.M.B. **Texto Acadêmico - Os Vegetais e a Água. UFBA / Projeto Qualibio, Salvador**, 1998. Disponível em: <http://www.qualibio.ufba.br/012.html>. Acesso em: 04/08/2011.

SILVEIRA, J.A.G.; MELO, A.R.B.; VIÉGAS, R.A.; OLIVEIRA, J.T.A. Salinity-induced effects on nitrogen assimilation related to growth in cowpea plants. **Environmental and Experimental Botany**, v. 46, n. 171-179, 2001.

SIMON, M.V.; BENKO-ISEPPON, A.M.; RESENDE, L.V.; WINTER, P. et al. Genetic diversity and phylogenetic relationships in *Vigna* Savi germplasm revealed by DNA amplification fingerprinting. **Genome**, v. 50, n. 538-547, 2007.

SINGH, B.B. Cowpea [*Vigna unguiculata* (L.) Walp. In: SINGH, R.J.; JAUHAR, P.P. editor. Genetic Resources, Chromosome Engineering and Crop Improvement. v. 1. Boca Raton, **FL: CRC Press**, 2005, pp. 117–162.

SINGH, B.B.; AJEIGBE, H.A.; TARAWALI, S.A.; FERNANDEZ-RIVERA, S. and ABUBAKAR, M. Improving the production and utilization of cowpea as food and fodder. **Field Crops Research**, v. 84, n. 1–2, p.169–177, 2003.

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

SNEATH, P.H.; SOKAL, R.R. Numerical taxonomy: The principles and practice of numerical classification. San Francisco: **W.H. Freeman**, 573 p., 1973.

SOTO, G.; ALLEVA, K.; MAZZELLA, M.A.; AMODEO, G.; MUSCHIETTI, J.P. AtTIP1;3 and AtTIP5;1, the only highly expressed Arabidopsis pollen-specific

aquaporins, transport water and urea. **FEBS Letters**, v. 582, p. 4077-4082, 2008.

SOTO, G.; FOX, R.; AYUB, N.; ALLEVA, K.; GUAIMAS, F.; ERIJMAN, E.J.; MAZZELLA, A.; AMODEO, G.; MUSCHIETTI, J. TIP5;1 is an aquaporin specifically targeted to pollen mitochondria and is probably involved in nitrogen remobilization in *Arabidopsis thaliana*. **Plant Journal**, v. 64, p. 1038–1047, 2010.

STEELE, W.M.; MEHRA, K.L. Structure, evolution and adaptation to farming system and environment in *Vigna*. In: Summerfield DR, Bunting AH eds. Advances in legume science. **Royal Botanical Gardens.**, p. 459-468, 1980.

STUDHOLME, D.J.; RAWLINGS, N.D.; BARRETT, A.J.; BATEMAN, A. A comparison of Pfam and MEROPS: two databases, one comprehensive, and one specialised. **BMC Bioinformatics**, v. 4, n. 1, p. 17, 2003.

SUBRAMANIAN, A. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, p.15545-15550. 2005.

SUI, H.; HAN, B.G.; LEE, J.K.; WALIAN, P.; JAP, B.K. Structural basis of water-specific transport through the AQP1 water channel. **Nature**, v. 414, n. 872–878, 2001.

SUN, M.; ZHOU, G.; LEE, S.; CHEN, J.; SHI, R.Z.; WANG, S.M. SAGE is far more sensitive than EST for detecting low-abundance transcripts. **BMC Genomics**, London, v. 5, n. 1, p.1-4, 2004.

TAIZ, L.; ZEIGUER, E. Fisiologia Vegetal, Tradução de: Elaine R. Santarém, [et al.], 3ª ed., Porto Alegre, **Artimed**, 719 p., 2004.

TAKANO, J.; WADA, M.; LUDEWIG, U.; SCHAAF, G.; VON WIREN, N.; FUJIWARA, T. The Arabidopsis major intrinsic protein NIP5;1 is essential for efficient boron

uptake and plant development under boron limitation**. Plant Cell**, v. 18, n. 6, p. 1498-1509, 2006.

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

TATENO, Y.; IMANISHI, T.; MIYAZAKI, S.; FUKAMI-KOBAYASHI, K. et al. DNA databank of Japan (DDBJ) for genome scale research in life science. **Nucleic Acids Research**, v. 25, p. 4876- 4882, 2002.

TEÓFILO, E.M.; PAIVA, J.B.; MEDEIROS-FILHO, S. Polinização artificial em feijão caupi. (*Vigna unguiculata* (L.) Walp.). **Ciência e Agrotecnologia**, Lavras, v. 25, n. 1, p. 220-223, 2001.

THOMAS, D.; BRON, P.; RANCHY, G.; DUCHESNE, L.; CAVALIER, A.; ROLLAND, J.P.; RAGUENES-NICOL, C.; HUBERT, J.F.; HAASE, W.; DELAMARCHE, C. Aquaglyceroporins, one channel for two molecules. **Biochimica et Biophysica Acta**, v. 1555, p. 181-186, 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.

TIMKO, M.P. Molecular cloning in cowpea: perspectives on the status of genome characterization and gene isolation for crop improvement. In: FATOKUN, C.A.; TARAWALI, S.A.; SINGH, B.B.; KORMAWA, P.M.; TAMÒ, M. (Eds.) Challenges and opportunities for enhancing sustainable cowpea production. **Proceedings of the World Cowpea Conference III held at the International Institute of Tropical Agriculture (IITA) Nigeria**, p. 197-212, 2002.

TIMKO, M.P.; EHLERS, J.D.; ROBERTS, P.A. COWPEA. In: KOLE, C. editor. Genome Mapping and Molecular Breeding in Plants Pulses, Sugar and Tuber Crops. v. 3, **Berlin Heidelberg: Springer-Verlag**, 2007, pp. 49–68.

TIMKO, M.P.; RUSHTON, P.J.; LAUDEMAN, T.W.; BOKOWIEC, M.T. et al. Sequencing and analysis of the gene-rich space of cowpea. **BMC Genomics**, v. 9, p.103, 2008.

TIMKO, M.P.; SINGH, B.B. Cowpea, a Multifunctional Legume. In: MOORE, P.H.; MING, R. (Eds.) **Genomics of Tropical Plants (Plant Genetics and Genomics: Crops and Models)**. Springer, v. 1, p. 227-258, 2008.

TÖRNROTH-HORSEFIELD, S.; WANG, Y.; HEDFALK, K.; JOHANSON, U.; KARLSSON, M.; TAJKHORSHID, E.; NEUTZE, R.; KJELLBOM, P. Structural mechanism of plant aquaporin gating, **Nature**, v. 439, p. 688–694, 2006.

TYERMAN, S.D.; NIEMIETZ, C.M.; BRAMLEY, H. Plant aquaporins: multifunctional water and solute channels with expanding roles. **Plant Cell Environ**, v. 25, p. 173–194, 2002.

UEHLEIN, N.; KALDENHOFF, R. Aquaporins and plant leaf movements. **Annals of botany**, v. 101, n. 1, p. 1-4, 2008.

UEHLEIN, N.; LOVISOLO, C.; SIEFRITZ, F.; KALDENHOFF, R. The tobacco aquaporin NtAQP1 is a membrane CO2 pore with physiological functions. **Nature**, v. 425, n. 6959, p. 734-737, 2003.

VALENZUELA, H.; SMITH, J. Cowpea. In: Sustainable Agriculture Green Manure Crops. **Cooperative Extension Service, College of Tropical Agriculture & Human resources**, University of Hawai, v. 6, p.1-3, 2002.

VELCULESCU, V.E.; ZHANG, L.; VOGELSTEIN, B.; KINZLER, K.W. Serial analysis of gene expression. **Science**, Washington, v. 270, p. 484-487, 1995.

VIDAL, M.S.; CARVALHO, J.M.F.C.; MENESES, C.H.S.G. Déficit Hídrico: Aspectos Morfofisiológicos, Documentos 142, **Embrapa**, Campina Grande, 2005.

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.

WALLACE, I.S.; WILLS, D.M.; GUENTHER, J.F.; ROBERTS, D.M. Functional selectivity for glycerol of the nodulin 26 subfamily of plant membrane intrinsic proteins. **FEBS Letters**, v. 523, p. 109–112, 2002.

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

WANG, W.X.; VINOCUR, B.; SHOSEYOV, O.; ALTMAN, A. Biotechnology of plant osmotic stress tolerance: physiological and molecular considerations. **Acta Horticulturae**, v. 560, p. 285-292, 2001.

WANG, X.; LI, Y.; JI, W.; BAI, X.; CAI, H.; ZHU, D.; SUN, X.-L.; CHEN, L-J; ZHU, Y-M. A novel Glycine soja tonoplast intrinsic protein gene responds to abiotic stress and depresses salt and dehydration tolerance in transgenic *Arabidopsis thaliana*. **Journal of Plant Physiology**, v. 161, p. 1241-1248, 2011.

WEI, C.L.; PATRICK, N.G.; CHIU, K.P.; WONG, C.H.; ANG, C.C.; LIPOVICH, L.; LIU, E.T.; RUAN, Y. 5' Long serial analysis of gene expression (LongSAGE) and 3' LongSAGE for transcriptome characterization and genome annotation. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, n. 32, p. 11701-11706, 2004.

WEI, W.; ALEXANDERSSON, E.; GOLLDACK, D.; MILLER, A.J.; KJELLBOM, P.O; FRICKE, W. HvPIP1;6, a barley (*Hordeum vulgare* L.) plasma membrane water channel particularly expressed in growing compared with non-growing leaf tissues. **Plant and Cell Physiology**, v. 48, p. 1132-1147, 2007.

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.

WILLADINO, L.; CAMARA, T. R. Tolerância das plantas à salinidade: aspectos fisiológicos e bioquímicos. **Enciclopédia Biosfera**, v. 6, n. 11, p. 1-23, 2010.

WILLIGEN, C.V.; POSTAIRE, O.; TOURNAIRE-ROUX, C.; BOURSIAC, Y.; MAUREL, C. Expression and inhibition of aquaporins in germinating Arabidopsis Seeds. **Plant Cell Physiology**, v. 47, p. 1241-1250, 2006.

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.

XAVIER, G.R.; MARTINS, L.M.V.; RUMJANEK, N.G.; FREIRE-FILHO, F.R. Variabilidade genética em acessos de caupi analisada por meio de marcadores RAPD. **Pesquisa Agropecuária Brasileira**, v. 40, p. 353-359, 2005.

YAHYA, A. Salinity effects on growth and on uptake and distribution of sodium and some essential mineral nutrients in sesame. **Journal of Plant Nutrition**, New York, v. 21, p. 1439-1451, 1998.

YAMAGUCHI-SHINOZAKI, K; SHINOZAKI, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. **Trends in Plant Science,** v. 10, p. 88-94, 2005.

ZARDOYA, R. Phylogeny and evolution of the major intrinsic protein family. **Biology of Cell**, v. 97, p. 397-414, 2005.

ZHANG, Y. Progress and challenges in protein structure prediction. **Current opinion in structural biology**, v. 18, n. 3, p. 342-8, 2008.

ZHANG, Y.; SKOLNICK, J. The protein structure prediction problem could be solved using the current PDB library. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 4, p. 1029-34, 2005.

ZHANG, Y.; WANG, Z.; CHAI, T.; WEN, Z.; ZHANG, H. Indian mustard aquaporin improves drought and heavy-metal resistance in tobacco. **Molecular Biotechnology**, v. 40, p. 280–292, 2008.

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.

ZHU, J.K. Cell signaling under salt water and cold stresses. **Current Opinion of Plant Biology**, v. 4, p. 401-406, 2001.

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