



**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
MESTRADO EM AGRONOMIA**

DILIANE HARUMI YAGUINUMA

**ANÁLISE *IN SILICO* E EXPRESSÃO GÊNICA DE CANDIDATOS DA FAMÍLIA
DAS AQUAPORINAS EM *Coffea canephora* SUBMETIDOS AO ESTRESSE
HÍDRICO**

DILIANE HARUMI YAGUINUMA

**ANÁLISE *IN SILICO* E EXPRESSÃO GÊNICA DE CANDIDATOS DA FAMÍLIA
DAS AQUAPORINAS EM *Coffea canephora* SUBMETIDOS AO ESTRESSE
HÍDRICO**

Dissertação apresentada Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade do Oeste Paulista, como parte dos requisitos para obtenção do título de Mestre em Agronomia – Área de concentração: Produção Vegetal.

Orientadora:
Prof.^a Dra. Alessandra Ferreira Ribas
Coorientador:
Dr. Tiago Benedito dos Santos

633.73
Y12a

Yaguinuma, Diliane Harumi.

Análise *in silico* e expressão gênica de candidatos da família das aquaporinas em *Coffea canephora* submetidos ao estresse hídrico / Diliane Harumi Yaguinuma. – Presidente Prudente, 2019.
95 f.: il.

Dissertação (Mestrado em Agronomia) -
Universidade do Oeste Paulista – Unoeste, Presidente Prudente, SP, 2019.

Bibliografia.

Orientadora: Alessandra Ferreira Ribas.

Coorientador: Tiago Benedito dos Santos.

1. Café. 2. Aquaporina. 3. Classificação. 4. Estresse.
I. Título.

CERTIFICADO DE APROVAÇÃO

TÍTULO: "ANÁLISE *IN SILICO* E EXPRESSÃO GÊNICA DE CANDIDATOS DA FAMÍLIA DAS AQUAPORINAS EM *Coffea canephora* SUBMETIDO AO ESTRESSE HÍDRICO"

AUTOR(A): DILIANE HARUMI YAGUINUMA

ORIENTADOR(A): ALESSANDRA FERREIRA RIBAS

Aprovado(a) como parte das exigências para obtenção do título de MESTRE em AGRONOMIA

Área de Concentração PRODUÇÃO VEGETAL, pela Comissão Examinadora:

Alessandra Ferreira Ribas

Profa. Dra. Alessandra Ferreira Ribas

UNOESTE - Universidade do Oeste Paulista / Presidente Prudente (SP)

Leonardo Cardoso Alves

Dr. Leonardo Cardoso Alves

SuperBAC BioTechnology Solutions / São Paulo (SP)

Nelson Barbosa Machado Neto

Prof. Dr. Nelson Barbosa Machado Neto

UNOESTE - Universidade do Oeste Paulista / Presidente Prudente (SP)

Data da realização: Presidente Prudente, 25 de novembro de 2019.

DEDICATÓRIA

À minha avó de coração Cergina de Jesus Mendes (in memorian) e meu pai Yoshinori Yaguinuma (in memorian) que foram exemplos de caráter, dedicação, paciência e amor incondicional.

AGRADECIMENTOS

Em primeiro lugar, não posso deixar de agradecer a minha orientadora, Professora Dra. Alessandra Ferreira Ribas, por toda a paciência, amizade, empenho e sentido prático com que sempre me orientou neste trabalho e em todas as outras apresentações e participações em simpósios, encontros e seminários. Ela foi essencial em todo processo e também a pessoa mais importante na conclusão do mesmo, sempre foi meu apoio e incentivo. Minha eterna admiração e gratidão.

Agradeço meu Coorientador Professor Dr. Tiago Benedito dos Santos, a quem me apresentou a Bioinformática e me mostrou um leque de novas possibilidades de pesquisa. Muito obrigada pela ajuda, suporte e paciência.

O meu muito obrigada ao Professor Dr. Edson Assunção Mareco pela ajuda e participação na Banca de Qualificação.

À Banca de Defesa de Mestrado composta pelo Dr. Leonardo Cardoso Alves e especialmente ao Professor Dr. Nelson Barbosa Machado Neto, que contribuiu e acompanhou todas as fases desde o início deste trabalho. Gratidão.

Agradeço de coração todos os meus colegas do Mestrado em Agronomia em especial minhas “migans” Giovana e a Bruna que são meus presentes preciosos da Pós em Agronomia. Sem o apoio e o carinho delas não teria ido tão longe. Amo vocês! À Vivi Cacefo, minha companheira querida de viagem, sempre prestativa e atenciosa, obrigada pela ajuda, amizade e risadas soltas. À Flaive, Lorryne, Jéssica, Camila, Jardel, Wilton obrigada pela companhia, conversas aleatórias, cujo apoio e amizade estiveram presentes em todos os momentos.

Às funcionárias dos Laboratórios de Genética Molecular (Maiara) e Cultura de Tecidos Vegetais (Cris) e a secretária da Pós (Keid) que foram sempre prestativas e me ajudaram muito na execução das análises e documentos solicitados

À todos os funcionários da Universidade, que sempre me receberam com sorriso no rosto, um “bom dia” e “até mais”. Desde a senhora da limpeza, o porteiro, o guarda, em especial o Ivanildo que sempre foi muito prestativo e simpático e também a querida Elis do Centro de Cópias.

Gratidão ao meu namorado Eder Garbin, pelo apoio, incentivo e paciência. Sem ele não teria iniciado o mestrado consecutivamente ao término da graduação. Obrigada meu amor por sempre me proporcionar o melhor que pôde.

À Universidade do Oeste Paulista e ao Curso de Pós-graduação em Agronomia.

Aos professores queridos pelos ensinamentos e dedicação.

Por último e não menos importante, quero agradecer à minha família e amigos da Biologia (Larissa, Takata, Dévlin, Luiz, Renan, Caio, Camila, Marcelo e Pedro) pelo apoio incondicional que me deram, pelos momentos de descontração e alegria! Vocês são DEMAIS!

“O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – (Brasil) CAPES – Código de Financiamento 001”.

“A evolução por si mesma não tem sentido se não formos capazes de fazer coisas grandes com o que somos agora”.

Jane Goodall

RESUMO

Análise *in silico* e expressão gênica de candidatos da família das aquaporinas em *Coffea canephora* submetidos ao estresse hídrico

As aquaporinas são proteínas pertencentes a família das MIPs (Major Intrinsic Proteins) e que estão inseridas na membrana plasmática formando canais que atuam diretamente no transporte de água e outros solutos essenciais para o desenvolvimento e crescimento de plantas. Genes dessa família são chaves para o estudo da resposta das plantas aos estresses hídricos. O objetivo desse trabalho foi a caracterização *in silico* da família das aquaporinas em *Coffea canephora* e avaliar a expressão relativa de membros CcAQPs em dois genótipos contrastantes de *C. canephora* submetidas ao déficit hídrico, sendo, o clone 14 tolerante e o clone 109A suscetível à seca. A plataforma Coffee Genome Hub foi utilizada para a busca das sequências das aquaporinas em *C. canephora*. As sequências foram analisadas quanto sua estrutura, localização, similaridade, presença de domínios e motivos conservados e relacionamento filogenético com outras espécies. Foram identificadas 33 aquaporinas em *C. canephora*, subdivididas em cinco famílias sendo: PIPs, TIPs, NIPs, SIPs e XIPs. O tamanho das proteínas variou entre 118 a 327 aminoácidos. O peso molecular variou entre 12364.41 a 35109.04 Mw e o ponto isoelétrico entre 5.12 a 9.66 pl. As aquaporinas de *C. canephora* estão localizadas na membrana celular e no vacúolo ou simultaneamente nos dois. A predição do filtro de seletividade aromático / arginina (ar / R), as posições e especificidade de Froger mostraram uma diferença notável na especificidade do substrato entre subfamílias. A expressão *in silico* variou entre os membros de cada subfamília e conforme o tecido analisado. Em geral, membros das subfamílias PIP e TIP apresentaram níveis de expressão mais elevados. Um experimento para imposição de déficit hídrico foi conduzido em casa de vegetação. As plantas foram submetidas a quatro tratamentos: sem estresse (ψ – 1,34 MPa), estresse moderado (ψ – 2,39 MPa), estresse severo (ψ – 4,5 MPa) e re-irrigado até o potencial similar ao sem estresse. O RNA foi extraído das folhas e convertidos em cDNA. Foram desenhados iniciadores para os membros CcPIPs1;2, 2;3, 2,4 e CcTIPs1;2, 2;1. A quantificação da expressão relativa para cada gene foi realizada pela RTqPCR calculada pelo método ΔCt utilizando o gene UBIQ10 como normalizador. O déficit hídrico afetou a expressão relativa das isoformas CcPIPs e CcTIPs analisadas em folhas de *C. canephora*. A expressão relativa de todas as isoformas reduziu no clone 14 enquanto aumentou no clone 109A quando as plantas foram submetidas ao estresse severo. Este estudo foi o primeiro a caracterizar *in silico* genes da família das aquaporinas e analisar a expressão relativa de membros das subfamílias PIP e TIP em *C. canephora*.

Palavras-chave: Café, Transporte de água, MIPs.

ABSTRACT

Genome wide and gene expression of candidates from the aquaporine family in *Coffea canephora* submitted to water stress

Aquaporins are proteins that belongs to a family of MIPs (Major Intrinsic Proteins) that are inserted into the plasmatic membrane forming channels that act directly in the transport of water and other essential solutes for the development and growth of plants. Genes of this family are the key to the study of the plants response of water stress. The objective of this work was to characterize in silico of the aquaporin family in *Coffea Canephora* and evaluate the relative expression of CcAQPs members in two contrasting of *C. canephora* genotypes submitted to water deficit, therefore, clone 14 is tolerant and clone 109A is susceptible to drought. The Coffee Genome Hub platform was used to search for the sequences of aquaporins in *C. canephora*. The sequences were analyzed regarding its structure, location, similarity, presence of domains and conserved motifs and phylogenetic relationship with other species. Were identified 33 aquaporins in *C. canephora*, subdivided into five families: PIPs, TIPs, NIPs, SIPs and XIPs. The size of the proteins varied between 118 to 327 amino acids. The molecular weight varied between 12364.41 to 35109.04 Mw and the isoelectric point varied between 5.12 to 9.66 pl. The aquaporins of *C. canephora* are located on the cell membrane and in the vacuole or simultaneously in both. The prediction of the aromatic / arginine selectivity filter (ar / R) Froger's positions and specificity showed a noticeable difference in substrate specificity between subfamilies. The expression in silico varied between the members of each subfamily according to the tissue analyzed. In general, members of the PIP and TIP subfamilies showed higher levels of expression. The plants were subjected to four treatments: no stress (ψ - 1.34 MPa), moderate stress (ψ - 2.39 MPa), severe stress (ψ - 4.5 MPa) and re-irrigated treatment until the potential similar to the no stress treatment. The RNA was extracted and converted from the leaves to cDNA. Primers were designed to the members CcPIPs1;2, 2;3, 2;4 e CcTIPs1;2, 2;1. The quantification of the relative expression for each gene was performed by RTqPCR calculated by the ΔC_t method using the UBIQ10 gene as a normalizer. The water deficit affected the relative expression of the CcPIPs and CcTIPs isoforms analyzed in *C. canephora* leaves. The relative expression of all isoforms decreased in the clone 14 while increased in the clone 109A when the plants were subjected to severe stress. This study was the first to characterize in silico genes of the aquaporin family and to analyze the relative expression of members of the subfamilies PIP and TIP in *C. canephora*.

Keywords: Coffee, Water transporters, MIPs.

LISTA DE SIGLAS

AQP: Aquaporina

ar/R: Aromático / arginina

BLAST - Ferramenta básica de pesquisa de alinhamento local

bp: Pares de base

CDS: Região codificadora

Chr: Cromossomo

cDNA: DNA complementar.

GIP: Proteína intrínseca do glicerol

HIP: Proteína intrínseca híbrida

H2: Hélice 2

H5: Hélice 5

LB: Alça B

LE: Alça E

MIP: Proteína intrínseca principal

MW: Peso molecular

NIP: Proteína intrínseca semelhante a Nod26

NPA: Asparafina-Prolina-Alanina

ORF: Moldura de leitura aberta

PCR – Reação em cadeia de polimerase

P1-P5: Resíduos nas posições P1 a P5

pI: Ponto isoelétrico

PIP: Proteína intrínseca da membrana plasmática

RNA: Ácido ribonucléico

RNA-seq: Sequenciamento de RNA

RPKM: *Reads* por Kilobase por Milhão de *Reads* Mapeados

SIP: Proteína intrínseca básica pequena

TIP: Proteína intrínseca de tonoplasto

TMH: Hélice transmembranar

XIP: Proteína intrínseca não categorizada

LISTA DE FIGURAS

CAPÍTULO I

Figura 1-	Disposição e origem geográfica dos principais grupos genéticos de <i>C. canephora</i> no continente Africano. I: local de origem do grupo Guineano. II: local de origem do grupo genético Congolês. Subgrupo SG1: considerado tolerante à seca. Subgrupo SG2: considerado sensível ao déficit hídrico; Áreas com hachuras: local do último período glacial (20.000 anos atrás)	24
Figura 2-	Ilustração do monômero de aquaporina. (A): A estrutura primária de MIPs compreende 6 domínios transmembrana (1-6) conectados por cinco alças (A-E), as alças B e E possuem subunidades helicoidais (HB e HE) compostas de um motivo NPA altamente conservado. (B): As subunidades helicoidais em <i>Loop</i> B e E dobram-se internamente, justapondo-se para formar um poro de água.....	28
Figura 3-	Segmentos conservados de aquaporinas responsáveis pela seletividade: o motivo de assinatura NPA e o filtro de seletividade (SF) no lado extracelular do canal. Esses dois elementos guiam as moléculas de água através de um poro estreito, onde a seletividade da água é conferida por fatores eletrostáticos e estéricos.....	28
Figura 4-	Distribuição dos genes <i>CcAQP</i> nos dez cromossomos de <i>C.canephora</i> . O programa utilizado para localizar e mapear os genes nos cromossomos foi o MapChart.....	38
Figura 5-	Perfil de expressão dos genes <i>CcAQP</i> em diferentes tecidos e órgãos. O <i>heatmap</i> foi clusterizado com base no valor <i>RPKM</i> dos <i>CcAQPs</i> . A escala representa a intensidade relativa do sinal dos valores do <i>RPKM</i> , onde a cor vermelha representa maior expressão, enquanto a cor verde representa a menor expressão.....	40
Figura 6-	Árvore filogenética de AQP de <i>C. canephora</i> . Os números nos nós internos mostram os resultados da análise de bootstrap (n = 1000). Ao lado a estrutura éxon-íntron de genes de aquaporinas em canéfora. Traços amarelos são regiões de éxon, as linhas são representadas pelos íntrons, regiões na cor azul são sequências downstream e upstream....	42
Figura 7-	Árvore filogenética de <i>C. canephora</i> e <i>A. thaliana</i> usando o parâmetro padrão de alinhamento <i>ClustalW</i> . O dendrograma filogenético foi gerado pelo MEGA X usando o método de Maximum Likelihood com 1.000 réplicas de bootstrap.....	43
Figura 8-	Alinhamento parcial das prováveis sequências XIP em três espécies de café (<i>C. arabica</i> (Ca), <i>C. eugenioides</i> (Ce) e <i>C. canephora</i> (Cc)) utilizando a ferramenta <i>ClustalW</i> através da plataforma MEGA X. Destacado em vermelho está o motivo NPARC.....	44

Figura 9-	Árvore filogenética das MIPs de café canéfora, arábica, eugenioides, arabidopsis, tomate, batata e arroz. Foi criada a partir do método Maximum Likelihood derivada do alinhamento <i>ClustalW</i> . Os números nos nós internos mostram os resultados da análise de bootstrap (n = 1000)	45
Figura 10-	Distribuição de motivos (motifs) conservados entre AQPs de canephora, onde diferentes motivos são representados por diferentes blocos de cores, conforme indicado na parte inferior da figura, e o mesmo bloco de cores em diferentes proteínas indica um determinado motivo.....	49
CAPÍTULO II		
Figura 1-	Amplificação de fragmentos dos genes <i>CcAQPs</i> utilizando o cDNA bulk pela reação em cadeia da polimerase e corrida em gel de agarose 1,5%. O teste foi realizado em duplicata. M – marcador ladder 100 pb, B branco, 1-2: <i>CcPIP</i> 1;2, 3 e 4 <i>CcPIP</i> 2;3, 5 e 6 <i>CcTIP</i> 2;1, 7 e 8 <i>CcTIP</i> 1;2, 9 e 10 <i>CcSIP</i> 1;1, 11 e 12 <i>CcNIP</i> 7;1, 13 e 14 <i>CcNIP</i> 5;2, 15 e 16 <i>CcPIP</i> 2;4.....	67
Figura 2-	Amplificação de fragmentos dos genes normalizadores com o bulk de cDNAs pela reação em cadeia da polimerase e corrida em gel de agarose 1,5%. M – marcador ladder 100 pb, 1 – branco, 2 – <i>CcUBQ10</i> , 3 – <i>CcGAPDH</i> , 4 - <i>CcEF1α</i>	67
Figura 3-	Curvas de dissociação produzidas pelo produtos de amplificação por RTqPCR para os <i>CcAPQs</i> utilizados neste estudo (cada curva inclui três replicatas técnicas do conjunto de cDNA de todas as amostras)	68
Figura 4-	Expressão relativa de gene AQPs em folhas de <i>C. Canephora</i> submetidas a diferentes condições de irrigação. Os resultados são expressos em níveis de expressão relativa dos genes alvo em relação a expressão do gene da ubiquitina como referência interna (método ΔCt). Os valores representam a média ($\pm$ DP) calculada a partir de três repetições biológicas em cada tratamento. A análise estatística entre os diferentes tratamentos foi realizada por meio de ANOVA unidirecional seguida do teste de Tukey ($p < 0,05$). Letras maiúsculas e minúsculas indicam comparações em diferentes condições de irrigação para os clones 109A e 14, respectivamente a comparação entre os clones em cada tratamento foi realizada pelo teste t (* $p < 0,05$)	69
Figura S1–	Árvore filogenética de <i>C. arabica</i> e <i>A. thaliana</i> usando o parâmetro padrão de alinhamento <i>ClustalW</i> . O dendrograma filogenético foi gerado pelo MEGA X usando o método de <i>Maximum Likelihood</i> com 1.000 réplicas de bootstrap.....	75
Figura S2–	Árvore filogenética de <i>C. eugenioides</i> e <i>A. thaliana</i> usando o parâmetro padrão de alinhamento <i>ClustalW</i> . O dendrograma filogenético foi gerado pelo MEGA X usando o método de <i>Maximum Likelihood</i> com 1.000 réplicas de bootstrap.....	76
Figura S3–	Predição dos domínios transmembrana em sequências deduzidas de aminoácidos de <i>CcAQP</i> através do site	

	<i>TMHMM</i> Server v. 2.0.....	77
Figura S4–	Alinhamento múltiplos de sequências de aminoácidos deduzidos de CcAQP utilizando a ferramenta <i>ClustalW</i> 2.1 disponível no programa MEGA X. Motivos NPA, negrito; Filtros de seletividade ar / R (H2, H5, LE1 e LE2) são destacados em vermelho; As 5 posições de Froger (P1 - P5) destacadas em verde. Posições das hélices transmembranas sublinhadas.....	81

LISTA DE TABELAS

CAPÍTULO I

Tabela 1 -	Genes <i>AQPs</i> identificados em <i>C. canephora</i>	36
Tabela 2 -	Motivos conservados de NPA, filtro de seletividade ar / R (H2, H5, LE1 e LE2), e posições de Froger (P1-P5) de CcAQPs.....	48

CAPÍTULO II

Tabela 1 -	Iniciadores desenhados para análise da expressão relativa por RTqPCR em CcAQPs.....	66
------------	---	----

SUMÁRIO

CAPÍTULO I- ANÁLISE GENÔMICA *IN SILICO* DA FAMÍLIA DAS AQUAPORINAS EM *C. canephora*

1	INTRODUÇÃO	20
2	OBJETIVOS.....	22
2.1	Objetivo geral	22
2.2	Objetivos específicos.....	22
3	REVISÃO DE LITERATURA	23
3.1	<i>Coffea canephora</i>	23
3.2	Importância econômica	25
3.3	Estresse hídrico	26
3.4	Aquaporinas	26
3.5	Genoma de <i>Coffea canephora</i>	31
4	MATERIAL E MÉTODOS	32
4.1	Busca e Identificação das sequências relacionadas aos genes de AQP	32
4.2	Alinhamento das sequências e análise filogenética	32
5	Caracterização das AQP em <i>C. canephora</i>	34
5.1	Análise de estrutura dos genes	34
5.2	Análise de localização subcelular e perfil transcricional de genes <i>CcAQP</i> em <i>C. canephora</i>	34
5.3	Análise dos motivos das proteínas <i>CcAQP</i> em <i>C. canephora</i>	35
6	RESULTADOS.....	36
6.1	Identificação dos genes <i>CcAQP</i> em <i>Coffea canephora</i>	36
6.2	Mapeamento de genes nos cromossomos de <i>C. canephora</i>	38
6.3	Perfil transcricional de genes <i>CcAQP</i> em <i>C. canephora</i>	39
6.4	Análise filogenética de proteínas <i>CcAQP</i> e estrutura dos genes	41
6.5	Predição dos domínios transmembranas e análise das dos motivos conservados	46
6.6	Análise dos motivos proteicos de <i>CcAQP</i> em <i>C. canephora</i>	49
7	DISCUSSÃO	50
7.1	Caracterização de AQPs em <i>C. canephora</i>	50
7.2	Estrutura Gênica	51
7.3	Expressão <i>in silico</i>	52
7.4	Seletividade das proteínas codificadas pelos genes <i>CcAQP</i>	52
8	CONCLUSÃO	55

**CAPÍTULO II- ANÁLISE DA EXPRESSÃO GÊNICA DE MEMBROS DA FAMÍLIA
CcAQP EM DOIS GENÓTIPOS CONTRASTANTES de *C.
canephora* SUBMETIDO AO DÉFICIT HÍDRICO**

1	INTRODUÇÃO	58
2	OBJETIVO	60
3	MATERIAL E MÉTODOS	61
3.1	Experimento de estresse hídrico	61
3.2	Extração de RNA total	62
3.3	Síntese de cDNA.....	62
3.4	Desenho de <i>Primers</i>.....	63
3.5	Eficiência de amplificação dos iniciadores.....	63
3.6	Análise da expressão gênica	64
4	RESULTADOS.....	65
5	DISCUSSÃO	72
6	CONCLUSÃO	74
7	MATERIAL SUPLEMENTAR.....	75
	REFERÊNCIAS.....	86

CAPÍTULO I

ANÁLISE *IN SILICO* DOS GENES DA FAMÍLIA DAS AQUAPORINAS EM *Coffea canephora*

RESUMO

Análise *in silico* dos genes da família das aquaporinas em *Coffea canephora*

As aquaporinas pertencem a família das MIPs (Major Intrinsic Proteins) e desempenham um papel fundamental no transporte de água e outros solutos através das membranas celulares. Genes dessa família são chaves para o estudo da resposta das plantas aos estresses hídricos. O objetivo desse trabalho foi a caracterização *in silico* da família aquaporina em *Coffea canephora*. Essa espécie, também conhecida como Robusta, representa cerca de 30% da produção brasileira de café. O sequenciamento do genoma de *C. canephora* disponível na plataforma Coffee Genome Hub foi utilizado para a busca de sequências. As aquaporinas identificadas foram analisadas quanto sua estrutura, localização, similaridade, presença de domínios e motivos conservados e relacionamento filogenético com outras espécies. Foram identificadas 33 aquaporinas em *C. canephora* subdivididas em cinco famílias sendo sete proteínas intrínsecas da membrana plasmática (PIPs), nove proteínas intrínsecas do Tonoplasto (TIPs), onze proteínas intrínsecas tipo nodulina NOD26 (NIPs), três pequenas proteínas intrínsecas básicas (SIPs) e três proteínas intrínsecas não especificadas (XIPs). O tamanho das proteínas variou entre 118 a 327 aminoácidos. O peso molecular ficou entre 12364.41 a 35109.04 Mw e o ponto isoelétrico teve variação entre 5.12 a 9.66 pI. As aquaporinas de *C. canephora* estão localizadas na membrana celular e no vacúolo ou simultaneamente nos dois. A predição do filtro de seletividade aromático / arginina (ar / R), as posições e especificidade de Froger mostraram uma diferença notável na especificidade do substrato entre subfamílias. A expressão *in silico* de cada membro variou conforme o tecido analisado. Em geral membros das subfamílias PIP e TIP foram os que apresentaram maiores níveis de expressão. Este estudo é o primeiro a caracterizar *in silico* genes da família das aquaporinas em *C. canephora* permitindo a seleção genes candidatos para a validação funcional em resposta aos estresses abióticos nessa espécie.

Palavras-chave: Café, Transporte de água, MIPs.

ABSTRACT

Genome wide identification of aquaporin genes in *Coffea canephora*

Aquaporins belong to the *Major Intrinsic Proteins* (MIPs) family and play a key role in transport of water and other solutes across cell membranes. These genes are key to the study of plant response to water stress. *Coffea canephora*, also known as Robusta, represents about 30% of Brazilian coffee production. The whole genome sequencing from this species is available on the Coffee Genome Hub platform make it possible the in silico analyses. The objective of this work was characterizes in silico the aquaporin family in *C. canephora*. Thirty-three *C. canephora* aquaporins were identified and subdivided into five families: seven intrinsic plasma membrane proteins (PIPs), nine intrinsic Tonoplast proteins (TIPs), eleven intrinsic nodulin-like proteins named NOD26 (NIPs), three small intrinsic proteins (SIPs) and three unspecified intrinsic proteins (XIPs). Protein size ranged from 118 to 327 amino acids. Molecular weight ranged from 12364.41 to 35109.04 Mw and the isoelectric point ranged from 5.12 to 9.66 pI. They are located in the cell membrane and in the vacuole or in both simultaneously. The prediction of the aromatic / arginine selectivity filter (ar/R), Froger's positions and specificity showed a noticeable difference in substrate specificity between subfamilies. In silico expression varied according to the tissue analyzed. In general, members of the PIP and TIP subfamilies had the highest expression levels. This study is the first to characterize gene to aquaporin family in silico *C. canephora* allowing selection of candidate genes for functional validation in response to abiotic stresses in this situation.

Keywords: Coffee, Water Transporters, MIPs.

1 INTRODUÇÃO

Aquaporinas (AQPs), pertencem à família das principais proteínas intrínsecas (MIPs, *Major Intrinsic Proteins*), as quais desempenham um papel importante para facilitar a difusão de água e outras moléculas pequenas tais como glicerol, ureia, amônia (NH₃), dióxido de carbono (CO₂), peróxido de hidrogênio (H₂O₂), boro e silício através das membranas biológicas (GOMES *et al.*, 2009; MAUREL *et al.*, 2015; POMMERENIG *et al.*, 2015). AQPs existem em todos os organismos vivos: bactérias, fungos, animais e plantas. Na literatura existem muitos estudos sobre as AQPs envolvidos em diferentes processos fisiológicos, por exemplo em resposta a estresses abióticos (KHAN *et al.*, 2015 revisado por KAPILAN; VAZIRI; ZWIAZEK, 2018), fotossíntese e condutância estomática e mesofílica (PEREZ-MARTIN *et al.*, 2014), longevidade e viabilidade das sementes (MAO; SUN, 2015), entre outras.

Nas plantas, a primeira MIP foi identificada em feijoeiro e representou aproximadamente 2% do total das proteínas extraídas do cotilédone (JOHNSON *et al.*, 1989), sendo demonstrada a sua função de canal de água primeiramente por Maurel *et al.* (1993) em *Arabidopsis thaliana*. As aquaporinas em plantas apresentam alta multiplicidade de isoformas, as quais teriam ocorrido pela diferenciação das regiões carboxi e amino-terminal (C e N-terminal, respectivamente) da proteína durante o processo evolutivo (REIZER *et al.*, 1993). Nas plantas, as AQPs apresentam estruturas altamente conservadas com peso molecular médio entre 23-31 kDa, composto por seis hélices transmembranas e cinco alças de conexão sendo as caudas C e N-terminal inseridas no citosol responsáveis pela formação dos canais seletivos de água (CHAUMONT *et al.*, 2001; TORNROTH-HORSEFIELD *et al.*, 2006; GOMES *et al.*, 2009), além de contar com motivos conservados asparagina-prolina-alanina (Asp-Pro-Ala / NPA) (revisado por REGON *et al.*, 2014).

O número de AQPs é bastante variável entre as espécies. A partir do sequenciamento completo de vários genomas, estabeleceram-se o número de genes homólogos de muitas espécies eudicotiledoneas e monocotiledôneas sendo 35 em *A. thaliana* (JOHANSON *et al.*, 2001), 23 em *Vitis vinifera* (SHELDEN *et al.*, 2009), 47 em *Solanum lycopersicum* (REUSCHER *et al.*, 2013), 34 em *Citrus sinensis* (MARTINS *et al.*, 2015), 44 em *Hevea brasiliensis* Muell. Arg. (ZOU *et al.*,

2015a), 37 em *Ricinus communis* L. (ZOU *et al.*, 2015b), 41 em *Sorghum bicolor* L. (REDDY *et al.*, 2015), 26 em *Phyllostachys edulis* (SUN *et al.*, 2016), 28 em *Beta vulgaris* (KONG *et al.*, 2017) 42 em *Manihot esculenta* (ZOU *et al.*, 2019), 41 em *Zea mays* L. (BARI *et al.*, 2018), 35 em *Eutrema salsugineum* (QIAN *et al.*, 2019).

As aquaporinas podem ser classificadas em cinco subfamílias baseadas em homologia entre as sequências: proteínas intrínsecas de membrana plasmática (PIPs), proteínas intrínsecas do tonoplasto (TIPs), proteínas intrínsecas semelhantes a nodulina (NIPs), pequenas proteínas intrínsecas básicas (SIPs) e proteínas intrínsecas não classificadas (XIPs), sendo estas últimas descobertas recentemente e com função ainda desconhecida (MAUREL *et al.*, 2015). Além disso, são documentadas duas subfamílias adicionais, as proteínas intrínsecas do glicerol (*GlypF-like*) (GIPs) e as proteínas intrínsecas híbridas (HIPs), as quais encontram-se exclusivamente em musgo, e não em plantas vasculares (ANDERBERG *et al.*, 2011; 2012).

Como relatado acima existe grande variabilidade de aquaporinas em plantas, e estas são relacionadas a seletividade do substrato, localização celular e tecidual, compartimentalização das células vegetais e funções celulares responsáveis pelo balanço de água (WUDICK *et al.*, 2009).

Essa família de proteínas é de fundamental importância devido ao seu papel no transporte de água nas células vegetais implicando em todos os processos de desenvolvimento e crescimento das plantas. No entanto, existem poucas informações sobre a superfamília dos genes da AQP no gênero *Coffea* (SANTOS; MAZZAFERA, 2013; MINIUSSI *et al.*, 2015). O café é uma das mais importantes *commodities* do comércio internacional, a espécie *Coffea canephora* Pierre ex A. Froehner ($2n = 2x = 22$) é diploide, perene (LASHERMES *et al.*, 1999) e representa aproximadamente 40% da produção mundial segundo a OIC (ORGANIZAÇÃO INTERNACIONAL DO CAFÉ, 2019). O genoma de *C. canephora* foi sequenciado em 2014 e a disponibilização desses dados possibilitou a identificação de novos genes com potencial biotecnológico (DENOEUDE *et al.*, 2014). Neste contexto, o objetivo desse estudo foi identificar e classificar as aquaporinas no genoma de *C. canephora*.

As análises de bioinformática compreendem comparações de domínios conservados, avaliação da estrutura e classificação de genes, mapeamento em cromossomos e inferência filogenética. Além disso, também foi avaliado o perfil de expressão dos genes de algumas classes de AQPs em resposta ao déficit hídrico.

2 OBJETIVOS

2.1 Objetivo geral

O objetivo do presente estudo foi identificar e caracterizar os membros da superfamília das AQP no genoma de *C. canephora*.

2.2 Objetivos específicos

- Analisar *in silico* a superfamília das aquaporinas no genoma de *C. canephora*.
- Analisar a relação filogenética das aquaporinas de *C. canephora* com outras espécies.
- Comparar os domínios conservados entre os membros de cada família.
- Mapear os membros nos cromossomo de *C. canephora*.
- Analisar a expressão relativa de genes da AQP em *C. canephora* submetidas ao déficit hídrico.

3 REVISÃO DE LITERATURA

3.1 *Coffea canephora*

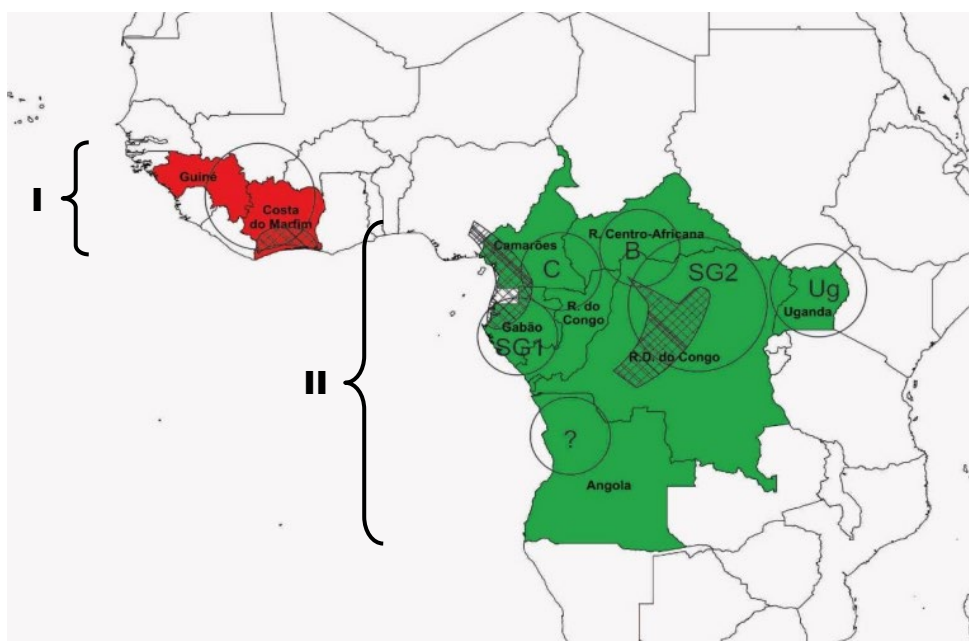
C. canephora é uma espécie diplóide ($2n = 2x = 22$) alógama e auto incompatível do tipo gametofítica, caracterizado pela fecundação cruzada como único meio de reprodução sexuada. Essa espécie apresenta uma melhor adaptação às planícies quentes e úmidas equatoriais, além de ser considerada mais resistente a doenças e pragas (SOUZA *et al.*, 2004).

A espécie *C. canephora* pertence à família Rubiaceae (DAVIS *et al.*, 2011), são arbustos multicaules que apresentam desenvolvimento inicial mais lento do que a *Coffea arabica*, mas ao atingirem a maturidade suas copas são mais desenvolvidas e de porte mais elevado (MENDES; GUIMARÃES, 1996; MATIELLO, 1998). A variedade botânica Conilon é caracterizada pelo crescimento de arbustos, folhas alongadas, maturação precoce, maior tolerância à pragas, doenças e deficiência hídrica. A variedade botânica Robusta tem como característica o crescimento ereto, folhas maiores, grãos maiores que a variedade Conillon, maturação tardia, maior resistência a doenças e pragas e menor tolerância à seca. (MONTAGNON *et al.*, 2012; SANTOS *et al.*, 2017).

O café é uma das culturas mais importantes do mundo, cerca de 2,5 bilhões de xícaras da bebida são consumidas diariamente (DENOEUDE *et al.*, 2014). A espécie *C. canephora* é constituída por populações polimórficas que possuem características distintas em relação a sua forma natural de reprodução, formato e tamanho dos grãos com maturação uniforme dos mesmos, tem origem e distribuição natural em uma ampla área do continente Africano (CHARRIER; BERTHAUD, 1985; BERTHAUD, 1986; LASHERMES *et al.*, 1999). Na Figura 1 em “I” pode-se observar a disposição e origem dos cinco grupos genéticos no continente Africano: Guineense representado por indivíduos da zona da Costa do Marfim e Guiné; em “II” têm-se o Grupo “C”, representado pelo café conhecido como Nana coletado nas fronteiras de Camarões, África Central e Congo; SG1 (subgrupo 1) representado pelo café Luki e Niaouli originário do sul do Gabão e os Conilons; SG2 (subgrupo 2) é o maior, composto pelo café cultivado na Bacia do Congo e café do sul da República Centro-

Africana; Grupo "Ug" ou "Ugandan" correspondente às populações selvagens originárias do Uganda.

Figura 1- Disposição e origem geográfica dos principais grupos genéticos de *C. canephora* no continente Africano. I: local de origem do grupo Guineano. II: local de origem do grupo genético Congolês. Subgrupo SG1: considerado tolerante à seca. Subgrupo SG2: considerado sensível ao déficit hídrico; Áreas com hachuras: local do último período glacial (20.000 anos atrás).



Fonte: (MONTAGNON *et al.*, 2012 e revisto por SOUZA *et al.*, 2015).

O Subgrupo SG1 possui grande proximidade genética ao café Conilon cultivado no Brasil (MONTAGNON *et al.*, 2012). O cultivo do café Conilon e Robusta no Brasil, se expandiu para áreas mais secas (DAMATTA *et al.*, 2006). Segundo a Companhia Nacional de Abastecimento CONAB (2017), os estados que mais produzem o café canéfora são: Espírito Santo, Rondônia e Bahia perfazendo um total de 381,58 mil hectares que estão em produção e 45,41 mil hectares em formação.

3.2 Importância econômica

O Brasil é o maior produtor e exportador de café em escala mundial, exportou 35,15 milhões de sacas de 60 kg com média mensal de 2,92 milhões de sacas em 2018. A produção mundial de café no ano cafeeiro de 2017/18 ficou 5,7%, acima de 164,81 milhões de sacas, com a produção de *C. arabica* aumentando em 2,2%, para 101,82, e o *C. canephora* (robusta), em 11,7%, para 62,99 milhões de sacas de acordo com a OIC (ORGANIZAÇÃO INTERNACIONAL DO CAFÉ, 2019).

Segundo a CONAB, a produção total dos Cafés do Brasil para a safra de 2019 está estimada em 50,92 milhões de sacas de 60 kg, onde 36,98 milhões são de *C. arabica* e 13,94 milhões de café *C. canephora* (conilon). Esses valores podem ser convertidos em uma redução de 17,4% na safra de 2019 em comparação com a de 2018. Desse volume físico, 47,48 milhões de sacas foram de café arábica e 14,17 milhões de *C. canephora* (conilon). Comparando a safra de 2018 à de 2019, a produção de *C. arabica* teve queda de 22,1% na produção acompanhada de *C. canephora* (conilon) que teve redução de apenas 1,7% de acordo com dados obtidos do 2º Levantamento da Safra de Café de 2019.

O Brasil é destaque na produção e exportação de café em nível mundial, as duas espécies mais comercializadas de café são: *C. arabica* L. e *C. canephora* (LEROY *et al.*, 2014). A espécie *C. arabica* é requisito em *blends* de alta qualidade enquanto que *C. canephora* é utilizado na indústria de café solúvel (SOUZA *et al.*, 2004). Segundo a EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária), a indústria do café solúvel utiliza cerca de quatro milhões de sacas da variedade conilon enquanto que a indústria de café torrado e moído utiliza oito milhões de sacas do produto.

Fatores que causem estresses abióticos, como altas temperaturas, períodos de seca e mudanças climáticas abruptas, podem afetar diretamente a produção de café. Os estresses abióticos ameaçam a qualidade e o fornecimento do produto, podendo afetar o comércio nacional e internacional.

3.3 Estresse hídrico

As plantas estão susceptíveis a estresses como déficit hídrico no decorrer do seu desenvolvimento, podendo variar a frequência, intensidade e também estar associado a outros estresses ambientais como salinidade e calor (SHAAR-MOSHE *et al.*, 2017). Esse tipo de estresse nas plantas gera respostas complexas e interligadas, indicando a percepção do sinal de estresse, resultando em respostas adaptativas a níveis morfológicos, fisiológicos e moleculares (HU; XIONG, 2014).

O déficit hídrico tem sido considerado o principal fator limitante para o crescimento e desenvolvimento das plantas em ambientes naturais e agrícolas. Fatores abióticos como temperaturas elevadas são uma das causas da mudança no potencial de membrana celular, perda de turgor da célula e desintegração dos processos metabólicos (MANAVALAN; NGUYEN, 2017).

As AQPs são peça-chave neste cenário, visto que as plantas desenvolveram uma ampla variedade de respostas à escassez de água em níveis fisiológico, bioquímico e molecular durante o percurso da evolução. Estudos dos padrões de expressão em diversas AQPs referentes a resposta ao estresse hídrico tiveram resultados altamente complexos que mudaram dependendo da isoforma investigada, estes nem sempre conseguiram definir uma correlação direta entre o nível de expressão e a abundância da proteína (ALEXANDERSSON *et al.*, 2005). A resposta das AQPs devido ao estresse hídrico pode direcionar a uma regulação ascendente ou descendente da expressão gênica ou até nenhuma alteração dependendo da duração e intensidade do estresse (GALMES *et al.*, 2007).

O estudo em tabaco sugeriu que as AQPs podem exercer um papel importante na tolerância à seca em plantas. A utilização de uma versão antisense do gene *NtAQP1* nas plantas de tabaco produziu uma redução na condutividade hidráulica das raízes e, portanto, uma maior susceptibilidade ao estresse pela seca (SIEFRITZ *et al.*, 2002).

3.4 Aquaporinas

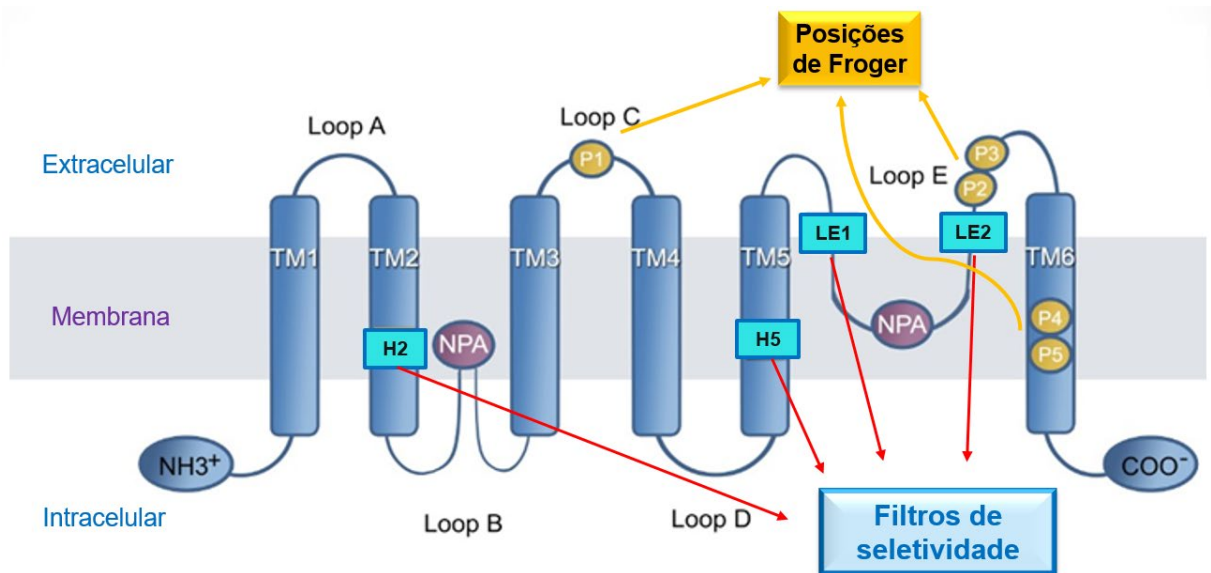
Aquaporinas (AQPs) são canais proteicos localizados na membrana celular, constituídos de proteínas integrais que facilitam o transporte de água e / ou

pequenos solutos neutros e gases através de membranas, diversificando sua distribuição entre os diferentes tipos de organismos (MAUREL *et al.*, 2008, VENKATESH *et al.*, 2013).

As proteínas de membrana helicoidal, aquelas que possuem α -hélices como estrutura secundária, são descritas de acordo com sua topologia, ou seja, a maneira como a proteína se dobra, em função do número e orientações das alças transmembranas assim como a localização dos terminais N e C (DE MAROTHY; ELOFSSON, 2015). As proteínas membranares em geral seguem um padrão de organização topológica contendo hélices transmembranas anti-paralelas as quais atravessam a membrana através de alças hidrofílicas para dentro e para fora do citoplasma (DE MAROTHY; ELOFSSON, 2015).

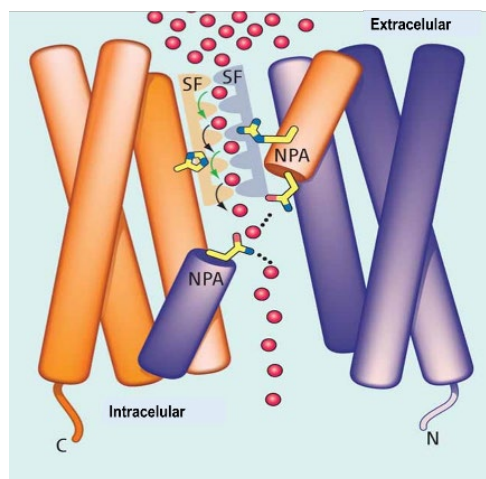
As AQPs são compostas por seis α -hélices transmembrana conectadas por cinco alças A – E (*loops*) nas quais ambos os N e C terminal estão voltadas para o interior citoplasmático. As alças B e D são citoplasmáticas enquanto que A, C e E são extracitoplasmáticas. Eles formam estruturas tetraméricas complexas nas quais cada subunidade funciona como um canal de água. Os poros dos canais são caracterizados por duas regiões de constrição que promovem o transporte seletivo de água e outros solventes neutros. A primeira constrição é formada no centro do poro por uma repetição interna de resíduos de asparagina-prolina-alanina altamente conservada que formam o motivo NPA localizados nas alças B e E. Essa região é extremamente hidrofóbica e de extrema importância para o funcionamento das AQPs (revisado por KAPILAN; VAZIRIN; ZWIAZEK, 2018). Uma representação esquemática das APQ pode ser visualizada na Figura 2. A segunda constrição é formada um resíduo de arginina altamente conservado no anel aromático, chamado de filtro de seletividade (ar / R) localizado na borda extracelular do poro (Figura 3). O filtro de seletividade (SF) é uma tétrade formada por quatro resíduos localizados nas hélices (H2 e H5) e na alça E (LE1 e LE2) (WALACE; ROBERTS, 2004).

Figura 2- Ilustração do monômero de aquaporina. (A): A estrutura primária de MIPs compreende 6 domínios transmembrana (TM1-6) conectados por cinco alças (LA-LE), as alças LB e LE possuem subunidades helicoidais (H2 e H5) compostas de um motivo NPA altamente conservado. (B): As subunidades helicoidais em LB e LE dobram-se internamente, justapondo-se para formar um poro de água.



Fonte: A Autora (baseada em HOVE; BHAVE, 2011)

Figura 03- As aquaporinas contêm dois segmentos conservados responsáveis pela seletividade: o motivo de assinatura NPA e o filtro de seletividade (SF) no lado extracelular do canal. Esses dois elementos guiam as moléculas de água através de um poro estreito, onde a seletividade da água é conferida por fatores eletrostáticos e estéricos.



Fonte: (ABRAMSON; VARTANIAN, 2013).

As AQPs são pertencentes à conservada superfamília das principais proteínas intrínsecas (MIP) e identificados em todos os grupos de organismos vivos (AGRE *et al.*, 1993). As AQPs são eficazes no controle do balanço hídrico da planta, desde o crescimento até sob condições de estresse, evidenciando a importância de sua presença (AROCA *et al.*, 2012; LUU; MAUREL, 2005; MAUREL *et al.*, 2008). Podem estar relacionadas também a alguns fatores como o alto grau de compartimentação em células; aperfeiçoamento funcional de cada isoforma AQP; seleção de substratos; tipos de tecidos órgãos e células de localização específica. (MAUREL *et al.*, 2008, WUDICK *et al.*, 2009).

Em plantas, as AQPs foram organizadas em cinco subfamílias: proteínas intrínsecas da membrana plasmática (PIPs), proteínas intrínsecas do Tonoplasto (TIPs), proteínas intrínsecas tipo nodulina (NIPs), pequenas proteínas intrínsecas básicas (SIPs) e proteínas intrínsecas não especificadas (XIPs) (CHAUMONT *et al.*, 2001, DANIELSON; JOHANSON, 2008; JOHANSON *et al.*, 2001; SAKURAI *et al.*, 2005). Sendo que a subfamília PIP foi dividida em dois grupos: PIP1 e PIP2 (SCHAFFNER, 1998).

Algumas isoformas PIP são altamente expressas em raiz, evidenciando a função das AQPs na obtenção do balanço hídrico em plantas e contribuindo na absorção de água disponível no solo (JAVOT *et al.*, 2003; KALDENHOFF *et al.*, 2008; POSTAIRE *et al.*, 2010). Em *A. thaliana* e *Nicotiana tabacum* foram realizados estudos em que PIP1 e PIP2 foram silenciados para demonstrar a importância dessas proteínas na realização do transporte da água pelas raízes (MARTRE *et al.*, 2002; SIEFRITZ *et al.*, 2002).

As TIPs podem conduzir tanto água (MAUREL *et al.*, 1993) como amônia (JAHN *et al.*, 2004; DYNOWSKI *et al.*, 2004), são encontrados no tonoplasto e constituem até 40% da membrana vacuolar (MAESHIMA 2001) e podem ampliar a eficiência de absorção de nitrogênio; outra função é a desintoxicação pelo acúmulo de ácido de íons amônio em vacúolos (LOQUÉ *et al.*, 2005). As TIPs estão presentes em todas as plantas terrestres e os cinco subgrupos especializados (TIP1–5) estão em plantas superiores e evoluídas, apenas o TIP 6 está presente em plantas primitivas como os musgos (ANDERBERG *et al.*; 2012).

As NIPs são encontradas em abundância em plantas primitivas, sabe-se que o subgrupo filogenético (NIP3) é derivado de um ancestral comum de musgos e plantas superiores (DANIELSON; JOHANSON, 2008). As NIPs transportam uma

variedade de pequenas moléculas neutras como, ureia, amônia, ácido bórico e silício. O grupo NIP2 contém os NIPs melhor caracterizados quanto à sua função fisiológica, captando e distribuindo silício dentro da planta (MA *et al.*, 2006).

Em SIP o transporte de água foi descrito apenas para o subgrupo SIP1 (ISHIKAWA *et al.*, 2005). Possivelmente essa diferença na especificidade de transporte seja devido a grandes diferenças no filtro ar / R de MIPs de plantas, como foi proposto para MIPs em *A. thaliana*, *Z. mays* e *Oriza sativa* (BANSAL; SANKARARAMAKRISHNAN, 2007)

Até o momento as XIPs foram encontradas apenas em plantas dicotiledôneas e em plantas superiores. Se agrupam em XIP1s e XIP2s, embora estes últimos tenham sido identificados apenas em *R. communis* e *Populus trichocarpa* (ZHOU *et al.*, 2015; GUPTA; SANKARARAMAKRISHNAN, 2009), essas proteínas detêm variações de aminoácidos em regiões correspondentes ao filtro primário de seleção assim como um ponto de verificação para a permeabilidade do soluto [no motivo Asn-Pro-Ala (NPA) e região aromática / arginina (ar / R)] (BEITZ *et al.* 2006; TORNROTH-HORSEFIELD *et al.* 2006; MITANI-UENO *et al.*, 2011). Pensava-se que os XIPs eram proteínas não-funcionais como canais de água, mas são utilizados no transporte de solutos hidrofóbicos (DANIELSON; JOHANSON, 2008; GUPTA; SANKARARAMAKRISHNAN, 2009). Bienert *et al.* (2011) validou tais predições em XIPs de *Nicotiana tabacum*, *S. lycopersicum* e *Solanum tuberosum*.

Além do transporte de água estudos com genes em AQP envolvem a regulação do desenvolvimento e crescimento das plantas. Grande parte dos estudos realizados foram orientados usando plantas modelo assim como em *A. thaliana* (JOHANSON *et al.*, 2001), *Z. mays* (BARI *et al.*, 2018), *Physcomitrella patens* (DANIELSON; JOHANSON, 2008) e *O. sativa* (SAKURAI *et al.*, 2005). Atualmente algumas plantas superiores como *M. esculenta* (ZOU; YANG, 2019), *C. sinensis* (WEI *et al.*, 2019), *E. salicigineum* (QIAN *et al.*, 2019) tiveram seus genes AQPs caracterizados. Ainda pouco se sabe sobre aquaporinas em *C. canephora*.

A análise da estrutura das aquaporinas tem um papel fundamental para a compreensão do seu funcionamento e seletividade. A conservação de resíduos em determinadas posições na sequências das proteínas aquaporinas são consideradas como um padrão de assinatura para essas proteínas em relação a estrutura, função e a seletividade do substrato para o transporte de água e pequenos solutos neutros tal como glicerol. Cinco posições-chave na sequência proteica (P1 – P5) nas quais

os resíduos estão associados a função e as propriedades físico-químicas específicas para cada sub-grupos foram denominadas como posições de Froger (FROGER *et al.*, 1998). Neste sentido a análise dessas cinco posições nas sequências proteicas pode dar indícios sobre sua especificidade e funcionalidade.

3.5 Genoma de *Coffea canephora*

O sequenciamento de *Coffea sp* possibilita o estudo da estrutura e a manipulação de genes de interesse à partir de buscas nos bancos de dados. Um exemplo é o estudo *in silico* da superfamília das aquaporinas (diretamente relacionadas ao transporte da água) que possibilitaria a escolha de determinados genes candidatos para tolerância ao estresse hídrico, os quais poderão ser utilizados em programas futuros de melhoramento, a fim de aumentar a resistência de determinada cultivar.

O sequenciamento completo do genoma *C. canephora* resultou na geração de um genoma de referência, que permite a genotipagem em escala genômica, possibilitando um grande avanço para o melhoramento genético no desenvolvimento de diversas características de interesse. A espécie *C. canephora* pertence à família das angiospermas é diploide e heterozigota exibindo uma ordem de genes cromossômicos conservados (LASHERMES *et al.*, 1999 revisado por DENOEUDE *et al.*, 2014). O sequenciamento completo foi realizado a partir da colaboração de várias equipes e em particular do *International Coffee Genome Sequencing Consortium*. O sequenciamento resultou em um mapa genético de alta densidade abrangendo 349 *scaffolds* e compreendendo ~ 64% do conjunto (364 Mb) e 86% dos genes anotados foram ancorados aos 11 cromossomos de *C. canephora*. Toda informação gerou um banco de dados que está disponível em *Coffee Genome Hub* (<http://coffee-genome.org/>) para pesquisas (DENOEUDE *et al.*, 2014).



4 MATERIAL E MÉTODOS

4.1 Busca e identificação das sequências relacionadas aos genes de AQP

Foram realizadas buscas para identificação dos membros da família AQP usando o banco de dados genoma café (<http://coffee-genome.org/>; DENOEUDE *et al.*, 2014) através da palavra-chave *Aquaporin*. Um catálogo de todas as sequências identificadas de aminoácidos e CDS completos para AQPs de café foi organizado em uma tabela.

Com a finalidade de realizar análises comparativas, foram também realizadas buscas de sequências AQPs em outras espécies como arábida (*A. thaliana*), tomate (*S. lycopersicum*), arroz (*O. sativa*) e batata (*S. tuberosum*). Para as espécies: *A. thaliana*, *S. lycopersicum*, *O. sativa* e *S. tuberosum* as sequências de aminoácidos de AQP foram recuperadas a partir dos genomas das referidas espécies disponíveis na plataforma *Phytozome v12.0* (<http://www.phytozome.net>; GOOSTEIN *et al.*, 2011).

Para identificação dos membros de AQP em outras espécies de café (*C. arabica* e *C. eugenioides*) utilizou-se o Centro Nacional de Informações sobre Biotecnologia (NCBI) onde encontra-se um grande conjunto de recursos *on-line* para informações e dados biológicos, incluindo o *GenBank*[®], banco de dados de sequência de ácido nucléico (<https://www.ncbi.nlm.nih.gov/>). Para busca de membros MIPs utilizou-se a palavra-chave *Aquaporin*, resultando em uma lista com sequências de aminoácidos e CDS completos.

4.2 Alinhamento das sequências e análise filogenética

Para realizar o alinhamento das sequências, foi utilizado a plataforma *ClustalW* (<http://www.genome.jp/tools-bin/clustalw>; THOMPSON; HIGGINS; GIBSON, 1994). Posteriormente, foi realizado um alinhamento múltiplo de sequências proteicas com espécies de plantas superiores, incluindo *A. thaliana*, *S. lycopersicum*, *S. tuberosum*, *O. sativa*, *C. arabica* e *C. eugenioides* na intenção de obter uma melhor classificação dos membros das CcAQPs. As árvores filogenéticas

foram construídas utilizando o *software* MEGA X (KUMAR *et al.*, 2018) com método de Máxima verossimilhança (*Maximum Likelihood - ML*) usando a confiabilidade e foi realizada pelo método de *bootstrap* com 1.000 réplicas. A classificação das subfamílias de *C. canephora*, *C. arabica* e *C. eugenioides*, foi elaborada à partir do relacionamento filogenético com as MIPS de *A. thaliana*.

5 CARACTERIZAÇÃO DAS AQP EM *C. canephora*

5.1 Análise de estrutura dos genes

A análise da estrutura dos genes de *CcAQPs* foi predita pelo programa *Gene Structure Display Server* (GSDS; <http://gsds.cbi.pku.edu.cn/> HU *et al.*, 2015). A informação da posição cromossômica foi detectada para cada um dos genes *CcAQPs* identificados pelo banco de dados do genoma *C. canephora* (DENOEUDE *et al.*, 2014). Além disso, o mapeamento de genes foi realizado utilizando o Mapchart 2.2 (VOORRIPS, 2002) onde localização física dos genes está representada graficamente.

5.2 Análise de localização subcelular e perfil transcricional de genes *CcAQP* em *C. canephora*

Foram determinadas características físico-químicas, como pesos moleculares (Mw) e pontos isoelétricos (pI) das *CcAQPs* utilizando o programa ProtParam (<https://web.expasy.org/protparam/> GASTEIGER *et al.*, 2005). A predição da localização subcelular de prováveis *CcAQP* foi realizada utilizando o programa *Plant-mPLoc* com parâmetro padrão (CHOU; SHEN, 2010). A predição dos domínios transmembranas de *CcAQPs* foi representada por meio de diagramas utilizando a plataforma Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/>; KELLEY; STERNBERG, 2009).

Para a análise de expressão das AQPS dos diferentes órgãos e tecidos de *C. canephora* os valores de *RPKM* (*Reads Per Kilobase Million*) foram recuperados no *JBrowse* disponível na base de dados *Coffee Genome Hub* (<http://coffee-genome.org/> / DENOEUDE *et al.*, 2014; DEREEPER *et al.*, 2015). Os dados foram compilados em uma planilha que posteriormente foi utilizada como dados de entrada para o perfil da expressão *in silico* utilizando o algoritmo do programa *Heatmapper* (<http://www1.heatmapper.ca/>; BABICKI *et al.*, 2016).

A escala de cores representa a intensidade relativa do sinal dos valores de *RPKM*. Os níveis de expressão dos transcritos são representados em um código de

cores com verde e vermelho, que mostra os níveis mais baixos e mais altos de expressão de transcritos, respectivamente. 

5.3 Análise dos motivos das proteínas *CcAQP* em *C. canephora*

As sequências deduzidas das proteínas de *CcAQP* foram analisadas por meio da ferramenta MEME versão 5.1 (Multiple EM for Motif Elicitation <http://meme-suite.org/index.html>; BAILEY *et al.*, 2006) para identificação dos motivos conservados. Para confirmação dos domínios transmembrana foi utilizado o programa *TMHMM Server* v.2.0 (<http://www.cbs.dtu.dk/serviços/TMHMM>; KROGH; RAPACKI, 2016). Por meio da plataforma *ClustalW* (<http://www.genome.jp/tools-bin/clustalw>; THOMPSON; HIGGINS; GIBSON, 1994) foi realizado o alinhamento múltiplo das sequências de *CcAQP*. A identificação dos motivos conservados de NPA, filtro de seletividade ar / R (H2, H5, LE1 e LE2) e análise das posições de Froger (P1-P5) foi realizada por inspeção visual dos alinhamentos em comparação com os de *AtAQP*. Visto que a subfamília XIP é inexistente em *A. thaliana* os motivos conservados desta subfamília em *CcAQP* foram identificados por comparação com outras espécies que também possuem essa subfamília tais como *R. communis* (ZHOU *et al.*, 2015) e *Gossypium hirsutum* (PARK *et al.*, 2010).

6 RESULTADOS

6.1 Identificação dos genes *CcAQP* em *Coffea canephora*

Estudos para identificação de genes da família das aquaporinas em plantas têm sido realizados por meio de metodologia *in silico*. Em *CcAQP* as sequências foram obtidas através de buscas na plataforma *Coffee Genome Hub* (<http://coffee-genome.org> / DENOEUDE *et al.*, 2014) e listadas na Tabela 1. Foram encontradas um total de 33 prováveis membros da família AQP em *C. canephora*.

O tamanho das proteínas variou entre 118 a 327 aminoácidos. O peso molecular variou entre 12364.41 a 35232.98 Mw e o ponto isoelétrico (pI) teve variação entre 5.12 a 9.66. A maioria das *CcAQPs* foram preditas com seis TMHs, enquanto que *CcNIP4;2* e *CcTIP4;1* obtiveram apenas duas hélices transmembranares. Quanto a localização subcelular observa-se que os *CcAQP* estão localizados na membrana celular e no vacúolo ou simultaneamente nos dois (Tabela 1). Todas as NIPs, PIPs, SIPs e XIPs estão localizadas na membrana celular. Além disso as *CcNIP1;3* e *3;1* as *CcPIP1;1* e *2;1* e a *CcSIP1;1* estão também presentes no vacúolo. Por outro lado, todas as TIPs estão localizadas no vacúolo com exceção da *CcTIP5;1* a qual está presente também na membrana celular (Tabela 1).

Tabela 1- Genes AQPs identificados em *C. canephora*.

Café MIPs*	Identificação do gene canephora	Posição do cromossomo	Proteína (aa)	Peso molecular (Mw)	TM hélices	pI	Localização
							<i>Plant-mPLOC</i>
CcNIP1;1	Cc02_g15840	chr2:14277549..14280032	265	27850.20	5	6.58	membrana celular
CcNIP1;2	Cc02_g15820	chr2:14270763..14273962	284	29914.55	6	8.96	membrana celular
CcNIP1;3	Cc02_g15890	chr2:14390802..14393029	235	25081.39	6	8.37	membrana celular/vacúolo
CcNIP2;1	Cc07_g03390	chr7:2365243..2367853	295	31021.71	6	7.70	membrana celular
CcNIP3;1	Cc10_g07150	chr10:6073710..6074611	178	19325.66	4	7.67	membrana celular/vacúolo
CcNIP4;1	Cc07_g08140	chr7:5983226..5984796	281	30353.32	7	5.73	membrana celular
CcNIP4;2	Cc11_g08690	chr11:26075747..26077737	154	17025.70	2	5.18	membrana celular
CcNIP5;1	Cc04_g15640	chr4:25353055..25357132	299	30565.42	5	9.10	membrana celular
CcNIP5;2	Cc08_g17120	chr8:31492021..31498389	297	30611.58	5	8.88	membrana celular
CcNIP6;1	Cc04_g04990	chr4:3713193..3715430	303	30958.93	6	7.75	membrana celular
CcNIP7;1	Cc03_g03200	chr3:2417270..2418614	254	26390.03	6	8.72	membrana celular
CcPIP1;1	Cc02_g06820	chr2:5399077..5400193	213	22439.32	5	9.66	membrana celular/vacúolo
CcPIP1;2	Cc02_g38120	chr2:52874143..52875950	250	22526.22	5	9.57	membrana celular/vacúolo
CcPIP1;3	Cc08_g14850	chr8:29720914..29723324	190	20382.88	5	9.51	membrana celular
CcPIP2;1	Cc01_g11110	chr1:29854916..29857498	288	30754.63	6	7.73	membrana celular
CcPIP2;2	Cc09_g09880	chr9:20778382..20779828	277	29621.57	6	9.13	membrana celular

Tabela 1 continuação

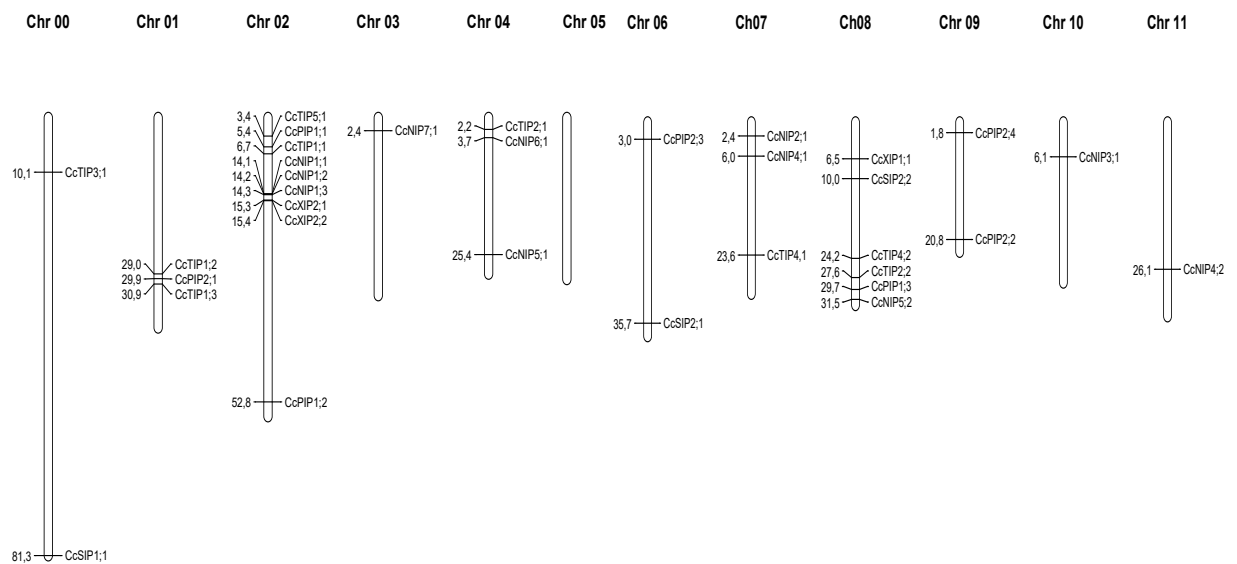
CcPIP2;3	Cc06_g03900	chr6:3002393..3003704	287	30566.41	6	8.20	membrana celular
CcPIP2;4	Cc09_g02170	chr9:1811779..1813234	199	21084.54	5	9.51	membrana celular
CcSIP1;1	Cc00_g09500	chr0:81327025..81329499	240	25412.00	5	9.48	membrana celular/vacúolo
CcSIP2;1	Cc06_g22730	chr6:35700588..35705369	256	28146.31	6	9.34	membrana celular
CcSIP2;2	Cc08_g05530	chr8:10009003..10009419	138	15831.13	4	9.46	membrana celular
CcTIP1;1	Cc02_g08360	chr2:6693536..6694585	252	25647.68	7	5.78	vacúolo
CcTIP1;2	Cc01_g10300	chr1:28971849..28973231	257	26534.99	6	6.28	vacúolo
CcTIP1;3	Cc01_g12260	chr1:30940570..30941579	251	25924.17	6	5.35	vacúolo
CcTIP2;1	Cc04_g02840	chr4:2193273..2194443	248	24951.23	7	6.48	vacúolo
CcTIP2;2	Cc08_g12280	chr8:27602314..27603735	251	25640.88	7	5.12	vacúolo
CcTIP3;1	Cc00_g01530	chr0:10099611..10100645	261	27818.23	6	6.22	vacúolo
CcTIP4;1	Cc07_g20570	chr7:23586574..23586930	118	12364.41	2	7.00	vacúolo
CcTIP4;2	Cc08_g09400	chr8:24159544..24161331	247	25777.15	6	6.02	vacúolo
CcTIP5;1	Cc02_g04340	chr2:3431090..3432414	249	25307.45	6	7.93	membrana celular/vacúolo
CcXIP1;1	Cc08_g04450	chr8:6507004..6508354	327	35232.98	6	6.54	membrana celular
CcXIP2;1	Cc02_g16640	chr2:15342830..15344090	326	35109.04	7	6.30	membrana celular
CcXIP2;2	Cc02_g16660	chr2:15356931..15358283	326	35026.98	7	5.63	membrana celular

*Nova classificação sugerida neste trabalho.

6.2 Mapeamento de genes nos cromossomos de *C. canephora*

Trinta e três genes foram mapeados desigualmente em dez cromossomos (Figura 4). Nove genes de *CcAQP* foram mapeados no cromossomo Chr02, seis no Chr08, três nos cromossomos Chr01, Chr04 e Chr07, dois em Chr00, Chr06 e Chr09 e apenas um gene de *CcAQP* nos Chr03 e Chr10. Nenhum AQP putativo foi encontrado em Chr05.

Figura 4- Distribuição dos genes *CcAQP* nos dez cromossomos de *C.canephora*, de acordo com a ferramenta *MapChart*.



6.3 Perfil transcricional de genes *CcAQP* em *C. canephora*

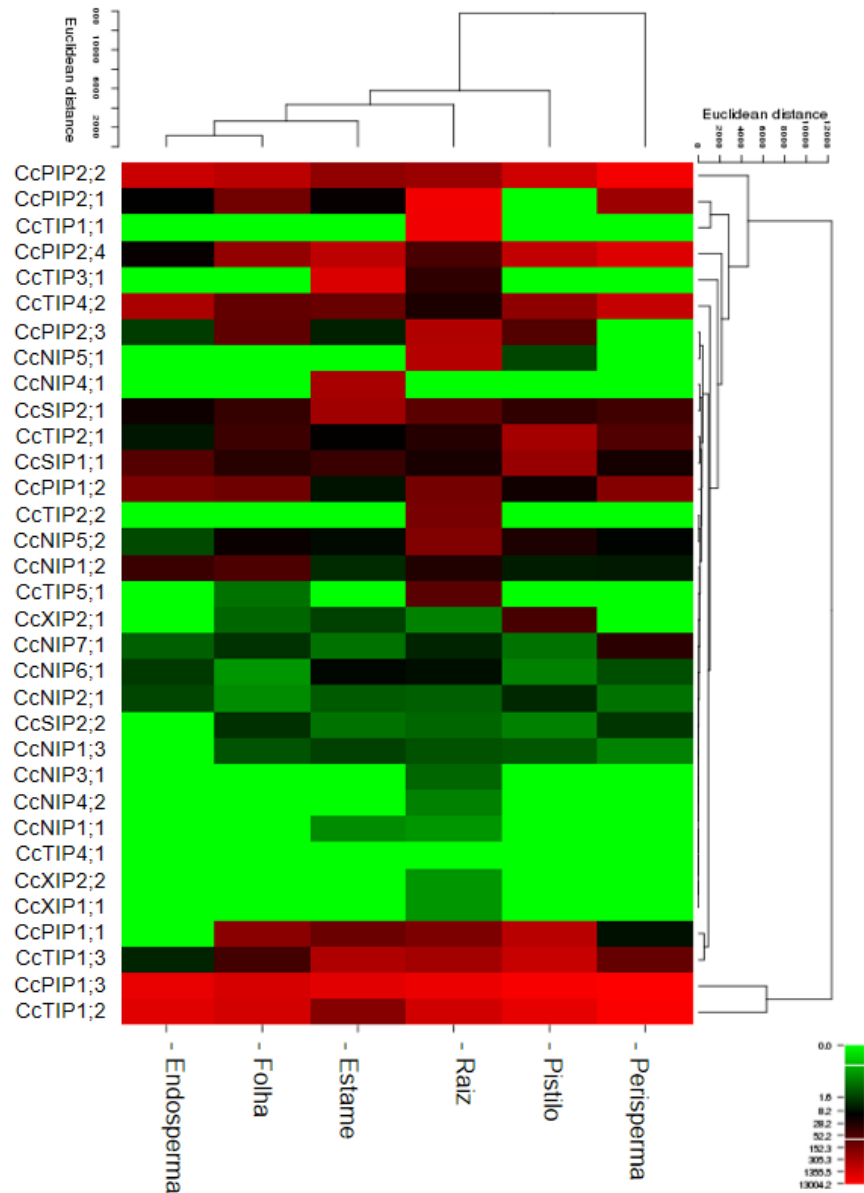
Para identificar os genes potenciais em resposta ao estresse abiótico, foram analisados o perfil dos padrões de expressão em todas as 33 AQPs de café e os valores estão demonstrados na Figura 5.

Em geral os genes *CcNIPs* foram os que apresentaram os mais baixos perfis de expressão. A mais alta expressão foi detectada em *CcTIP1;2*, *CcPIP1;3* e *CcPIP2;2* em todos os tecidos e órgãos. De modo geral membros das subfamílias PIP e TIP apresentaram maiores perfis de expressão, indicando a importância dos dois grupos em relação a respostas ao estresse. Em TIP apenas 4 membros mostraram expressão, sendo que em NIP apenas *CcNIP1;2* teve uma expressão relativa enquanto os genes obtiveram baixos níveis ou nenhuma expressão.

As *CcPIP1;3*, *CcPIP2;2*, *CcTIP1;2* e *CcTIP4;2* apresentaram alto perfil de expressão em todos os tecidos, indicando uma expressão constitutiva. Em *CcTIP1;1*, *CcTIP2;2*, *CcTIP5;1*, *CcPIP2;3* e *CcNIP5;1* obtiveram uma alta expressão específica em raiz.

Todas as XIPs tiveram uma baixa expressão, indicando pouca participação em resposta ao estresse, com exceção de *CcXIP2;1* em pistilo.

Figura 5- Perfil de expressão dos genes *CcAQP* em diferentes tecidos e órgãos. O *heatmap* foi clusterizado com base no valor *RPKM* dos *CcAQPs*. A escala representa a intensidade relativa do sinal dos valores do *RPKM*, onde a cor vermelha representa maior expressão, enquanto a cor verde representa a menor expressão.



6.4 Análise filogenética de proteínas *CcAQP* e estrutura dos genes

A análise das estruturas dos genes de *CcAQPs* foi gerada pelo programa *Gene Structure Display Server*. A subfamília PIP foi caracterizada por uma estrutura conservada de 4 éxons em todos os membros. A maioria dos membros da subfamília NIP foi caracterizada por 4 e 5 éxons, com exceção ao *CcNIP4;2* que apresenta apenas 3 éxons. A maioria dos membros da subfamília TIP possui 3 éxons, enquanto *CcTIP1;2* possui 2 éxons e *CcTIP4;1* possui apenas 1 éxon com 300pb sem íntrons, o que demonstra variação. Os genes da pequena subfamília SIPs contêm 3 éxons em dois membros e apenas 1 éxon em *CcSIP2;2*. A subfamília XIP que também é pequena foi caracterizada por uma estrutura conservada de 3 éxons (Figura 6).

Com base na homologia entre as sequências dos 33 prováveis genes aquaporinas identificados a partir do genoma de *C. canephora*, foi gerada uma árvore filogenética a qual apresenta quatro clusters distintos. Os genes de *CcAQP*, *CaAQP* e *CeAQP* foram nomeados com base na nomenclatura utilizada em outras plantas e guiados por similaridade de sequência e análise filogenética com *A. thaliana* como descrita por JOHANSON *et al.* (2001) com exceção da subfamília XIP (Figura 7, S1 e S2).

Figura 6- Árvore filogenética de AQP de *C. canephora* construída a partir do método *Maximum Likelihood* derivada do alinhamento *ClustalW*. Os números nos nós internos mostram os resultados da análise de *bootstrap* (n = 1000). Ao lado a estrutura éxon-íntron de genes de aquaporinas em canéfora. Traços amarelos são regiões de éxon, as linhas são representadas pelos íntrons, regiões na cor azul são sequências *downstream* e *upstream*.

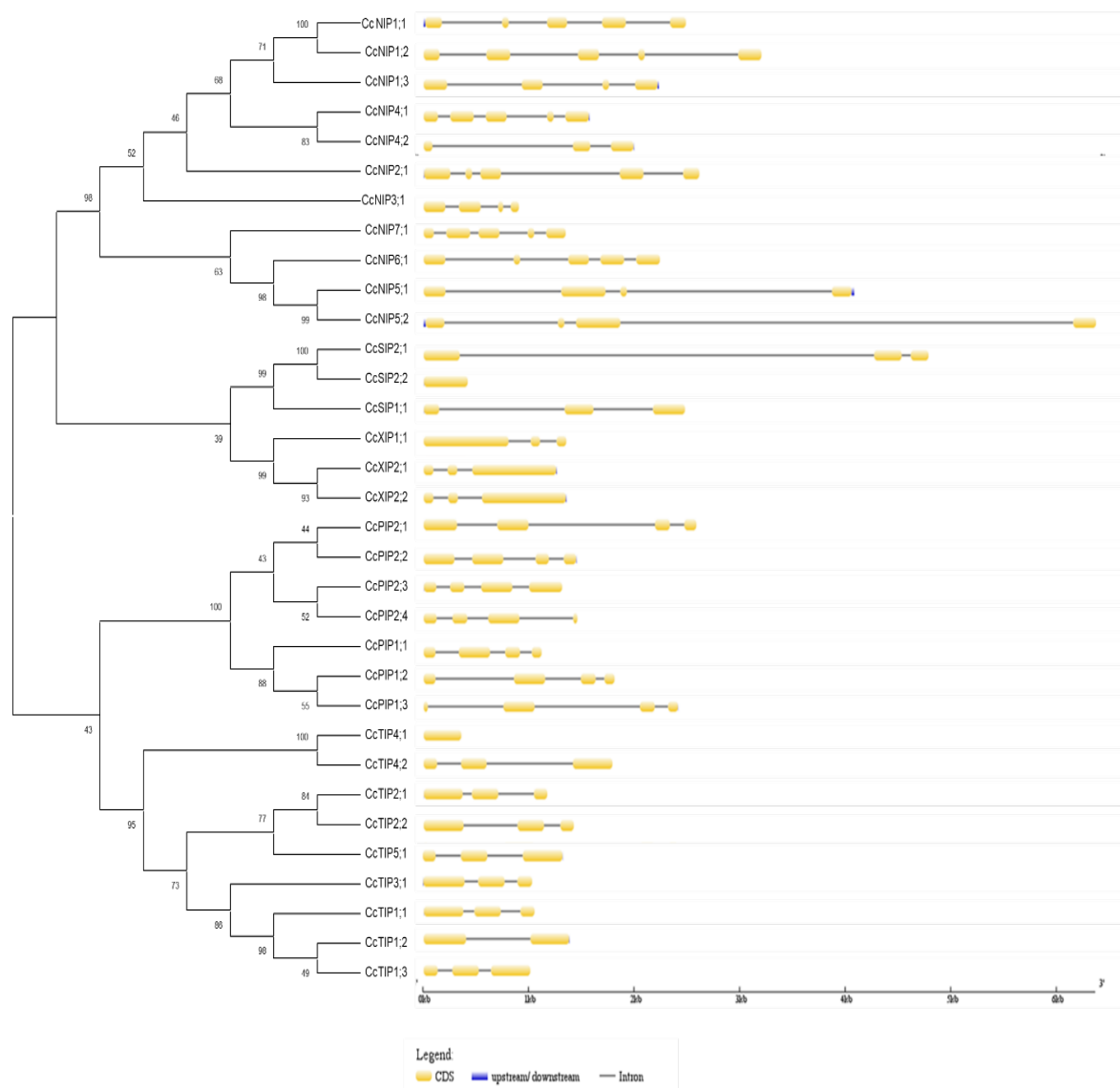
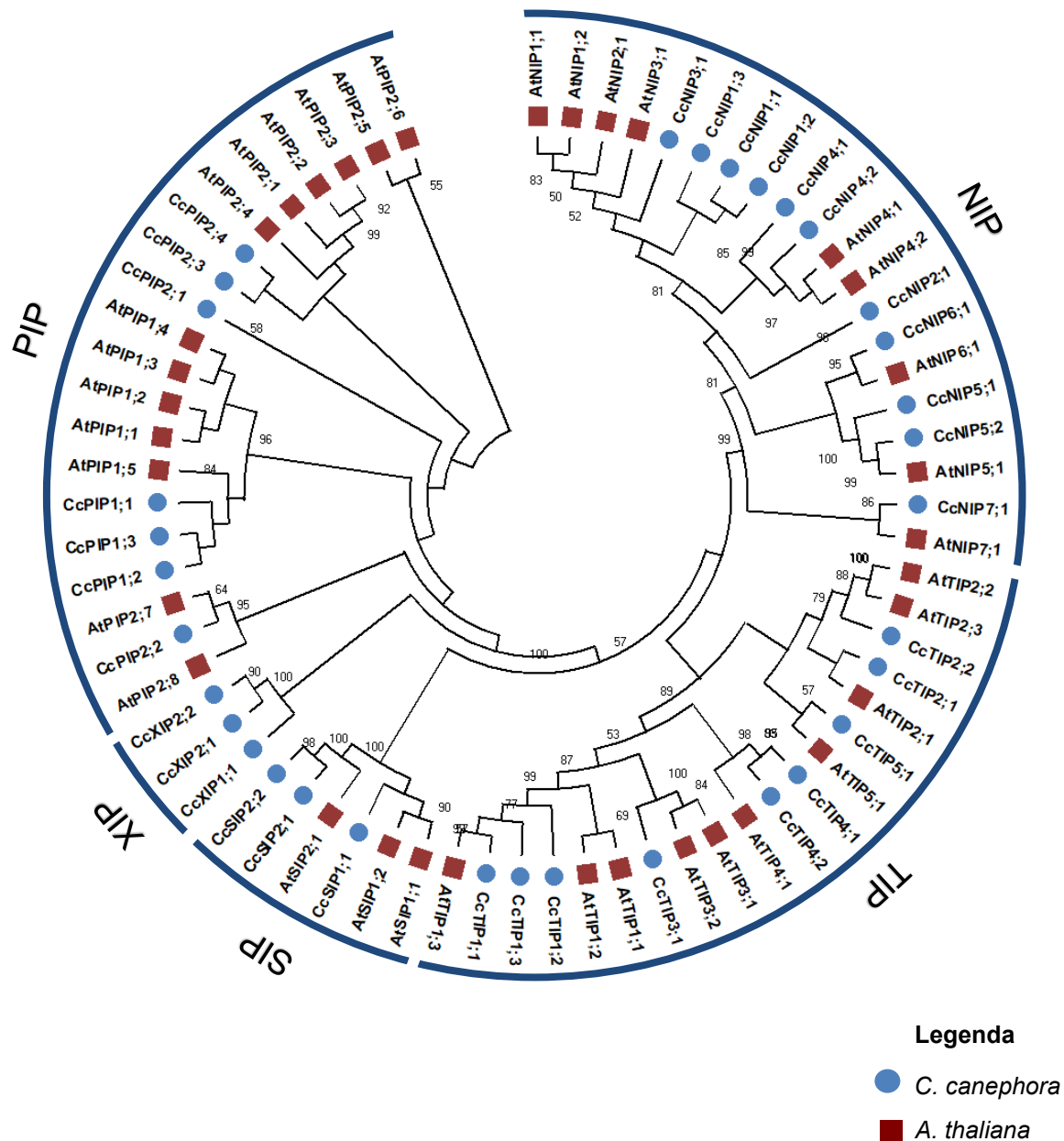


Figura 7- Árvore filogenética de *C. canephora* e *A. thaliana* usando o parâmetro padrão de alinhamento *ClustalW*. O dendrograma filogenético foi gerado pelo MEGA X usando o método de *Maximum Likelihood* com 1.000 réplicas de *bootstrap*.

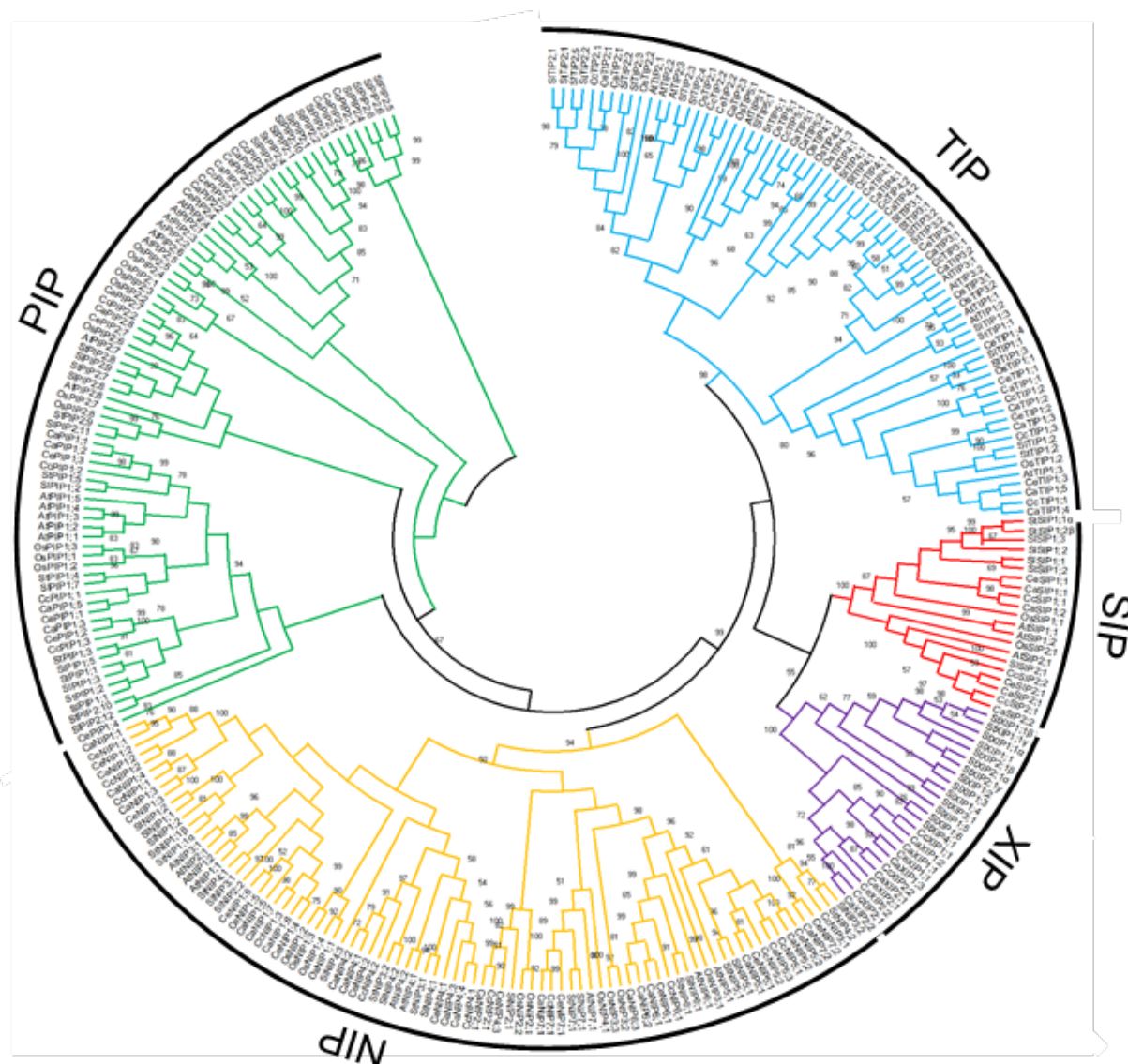


A relação filogenética entre e os genes identificados em *C. canephora* com outras espécies incluindo *C. arabica*, *C. eugenioides*, *A. thaliana*, *S. lycopersicum*, *S. tuberosum* e *O. sativa* revelaram que os membros de cada subfamília foram corretamente agrupados formando clados claramente distintos (Figura 9). As CcAQPs foram classificadas em 5 subfamílias divididas em 7 PIPs, 9 TIPs, 11 NIPs, 3 SIPs e 3 XIPs (Figura 6). A subfamília XIP, não está presente em *A. thaliana*, para classifica-las em café, utilizou-se as espécies *R. communis* (ZHOU *et al.*, 2015) e *G. hirsutum* (PARK *et al.*, 2010) como referência. A subfamília XIP possui uma particularidade para sua classificação, sendo necessário conter o segundo motivo NPA amplificado para NPARC, o qual está presente apenas nessa subfamília (DANIELSON; JOHANSON; 2008). Para verificar a presença do motivo estendido (NPARC) nas sequências de XIPs de *C. canephora*, *C. arabica* e de *C. eugenioides*, foi realizado o alinhamento e inspeção. Resultando na presença do motivo (NPARC) presente em todas as sequências de XIP (Figura 8).

Figura 8- Alinhamento parcial das prováveis sequências XIP em três espécies de café (*C. arabica* (Ca), *C. eugenioides* (Ce) e *C. canephora* (Cc)) utilizando a ferramenta *ClustalW* através da plataforma MEGA X. Destacado está o motivo NPARC.

Protein Sequences																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Species/Abbrev	*	*	*		*		*		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Figura 9- A árvore filogenética das MIPs de *C. canephora*, *C. arabica*, *C. eugenioides*, *A. thaliana*, *L. lycopersicum*, *L. tuberosum* e *O. sativa*. Foi criada a partir do método *Maximum Likelihood* derivada do alinhamento *ClustalW*. Os números nos nós internos mostram os resultados da análise de *bootstrap* (n = 1000).



6.5 Predição dos domínios transmembranas e análise dos motivos conservados

As sequências de aminoácidos dos CcAQPs foram alinhadas e os domínios transmembranas (TMHs), motivos NPA, filtro de seletividade ar / R e posição de Froger foram analisados visualmente para compreender o papel fisiológico e especificidade do substrato de cada aquaporina (Tabela 2).

Dentre as 33 CcAQPs analisadas, 16 sequências foram preditas contendo seis TMHs sendo três *CcPIPs* (2;1, 2;2 e 2;3), cinco *CcTIPs* (1;2, 1;3, 3;1, 5;1 e 4;2), seis *CcNIPs* (1;1, 1;2, 1;3, 2;1, 6;1 e 7;1) uma *CcSIP2;1* e uma *CcXIP1;1*. Sete sequências apresentaram cinco TMHs entre elas: *CcPIP1;1*, 1;2, 1;3, 2;4, *CcNIP5;1* e 5;2 e *CcSIP1;1*. Seis sequências apresentaram sete TMHs: *CcTIP1;1*, 2;1 e 2;2, *CcNIP4;1* e *CcXIP2;1* e 2;2. As sequências *CcNIP3;1* e *CcSIP2;2* apresentaram 4 TMHs e as sequências *CcTIP4;1* e *CcNIP4;2* apresentaram a predição de somente dois domínios transmembranas em cada (Figura S3).

A análise dos resíduos nas sequências de aminoácidos revelou que todas as PIPs e TIPs e a maioria das *CcNIPs* mostraram primeiro motivo NPA conservado na primeira alça *do Loop B* (LB), com exceção da *CcNIP4;2* que não apresentou motivo e *CcNIP5;2* que apresentou o (S) na terceira posição do motivo (Tabela 2). Nas três *CcXIP* o resíduo A (arginina) da terceira posição foi substituída pelo resíduo isoleucina (I). Já nas *CcSIP* o resíduo A (arginina) da terceira posição variou entre treonina (T) para *CcSIP1;1* e ausentes em *CcSIP2;1* e 2;2.

O segundo motivo de NPA encontra-se conservado nas subfamílias PIP, SIP, XIP e TIP com exceção de *CcTIP4;1* e *CcSIP2;2* que encontram-se ausentes. Em NIP foi observada maior diversidade. Nas sequências *CcNIP5;1*, *CcNIP5;2*, *CcNIP6;1* a terceira posição (Arginina) foi substituída por uma valina (V) enquanto que em *CcNIP3;1* a segunda e terceira posições do motivo NPA estão ausentes. Em *CcNIP1;1* a primeira letra do motivo de asparagina (N) foi substituído por serina (S) (Tabela 2).

O filtro de seletividade para as *CcNIPs* na H2 apresentou o resíduo G na 2;1 e T para 6;1 e A para as 5;1, 5;2 e 7;1 as demais apresentaram o resíduo W e sendo ausente na 4;2. Na H5 variou entre V, S ou I. No LE1 A ou G, W e R para LE2 com exceção de 7;1 com C da 1;1 com G e ausente em 3;1. As posições de Frogers na posição P1 variaram em F, L, e Y para 4;2 e 7;1. Em P2 S, T e R e ausente em 3;1 a

posição P3 e P4 altamente conservada apresentando os resíduos A e Y e ausente em 3;1 e C na P4. Em P5 variou entre I, L ou V e em 3;1 mostrou-se um N. Para as *CcPIPs* os filtros de seletividade foram altamente conservados apresentando os resíduos F em H2, resíduo H em H5, T e R em LE1 e LE2. As posições de Froger também foram altamente conservadas em *CcPIPs* apresentando os resíduos Q, S, A, F e W em P1 a P5. A única exceção foi a *CcPIP1;3* e *CcPIP2;1* as quais apresentaram os resíduos E e M na posição P1.

Para as *CcTIPs* o filtro de seletividade em H2 apresentou o resíduo H exceto para as *CcTIP1;2* e *5;1* que contém os resíduos Y e N. Na H5 apresentaram o resíduo I ausente apenas em *4;1*. Em LE1 A ou G e em LE2 V ou R com exceção da *5;1* que apresentou Y. As posições de Froger para as sequências *CcTIPs* foram também altamente conservadas apresentando os resíduos T, S, A, Y e W para as posições de P1 a P5. A sequência *CcTIP4;1* foi a única que apresentou resíduos L na P1 e ausente em todos os demais. Entre as sequências de *CcSIP* a *2;2* apresentou somente resíduos no filtro de seletividade H2 (K) e nas posições P4 (L) e P5 (F) estando ausente nas demais.

Entre as sequências *CcSIP1;1* e *2;1* os filtros de seletividade apresentaram-se variáveis TVPN e KHGS, respectivamente. Na posição de Froger apresentaram variação nas posições P1 (M/F) e P2 e A/V e conservada nas posições P3 a P5 (AYW). Em *CcXIP* os resíduos do filtro de seletividade apresentaram os resíduos LFAR. Somente em *CcXIP1;1* houve a troca da leucina (L) por valina (V) na posição H2. Todos os resíduos das posições de Froger, apresentaram-se altamente conservados entre as sequências (CCAFW) (Tabela 2).

As posições de Froger P3, P4 e P5 mostraram-se altamente conservadas. A posição P3 apresentou o resíduo hidrofóbico, alanina (A) em todas as subfamílias com exceção em *CcSIP2;2* que teve a posição ausente. Na posição P4 as subfamílias NIP, SIP e TIP apresentaram o resíduo Y enquanto que em PIP e XIP o resíduo F ambos aromáticos. Já na posição P5 as NIPs apresentaram os resíduos alifáticos, ou seja de cadeia aberta I, L ou V e as para as demais subfamílias o resíduo aromático W. as únicas exceções em P4 e P5 foram em *CcSIP2;2*, *CcTIP4;1* (Tabela 2).

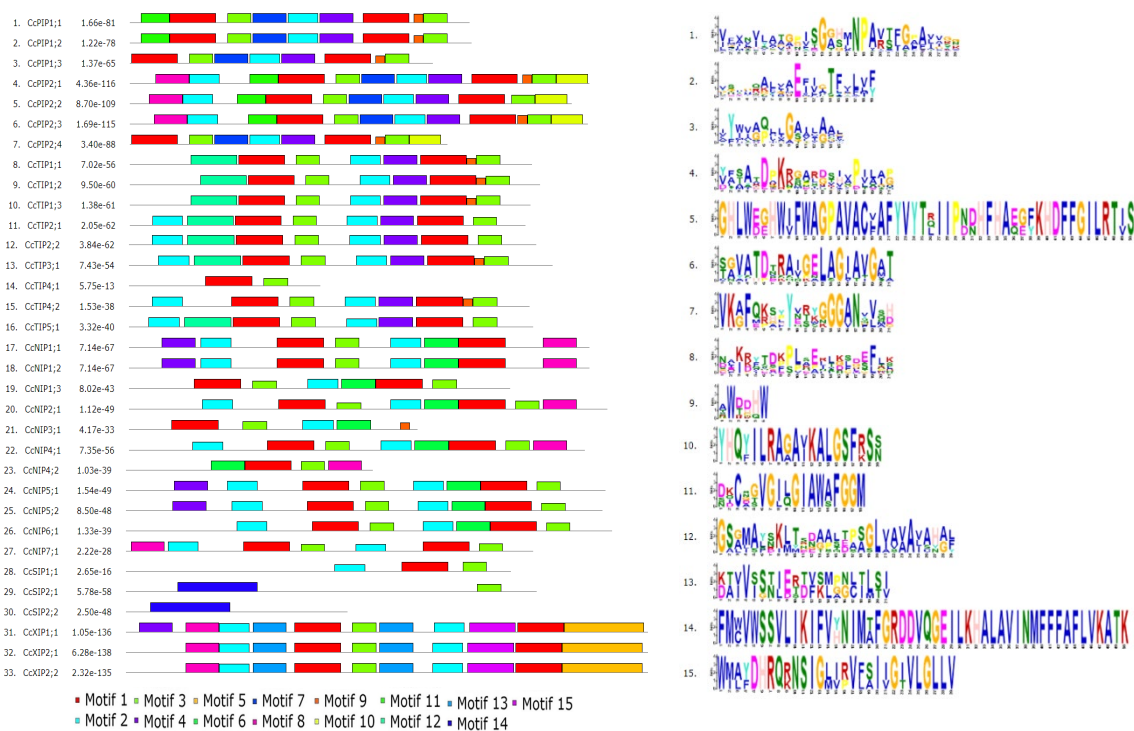
Tabela 2- Motivos conservados de NPA, filtro de seletividade ar / R (H2, H5, LE1 e LE2), e posições de Froger (P1-P5) de CcAQPs.

Subfamília	Gene	NPA		Filtro seletivo Ar / R				Posição de Froger				
		LB	LE	H2	H5	LE1	LE2	P1	P2	P3	P4	P5
NIP	<i>CcNIP1;1</i>	NPA	SPA	W	V	G	G	F	R	A	Y	I
	<i>CcNIP1;2</i>	NPA	NPA	W	V	A	R	F	S	A	Y	I
	<i>CcNIP1;3</i>	NPA	NPA	W	S	A	R	F	S	A	Y	I
	<i>CcNIP2;1</i>	NPA	NPA	G	S	G	R	L	S	A	Y	V
	<i>CcNIP3;1</i>	NPA	N--	W	S	W	*	F	*	*	C	N
	<i>CcNIP4;1</i>	NPA	NPA	W	V	A	R	F	S	A	Y	I
	<i>CcNIP4;2</i>	*	NPA	*	V	A	R	Y	S	A	Y	I
	<i>CcNIP5;1</i>	NPA	NPV	A	I	G	R	F	T	A	Y	L
	<i>CcNIP5;2</i>	NPS	NPV	A	I	G	R	F	T	A	Y	L
	<i>CcNIP6;1</i>	NPA	NPV	T	I	A	R	F	T	A	Y	L
	<i>CcNIP7;1</i>	NPA	NPA	A	V	A	C	Y	S	A	Y	L
PIP	<i>CcPIP1;1</i>	NPA	NPA	F	H	T	R	Q	S	A	F	W
	<i>CcPIP1;2</i>	NPA	NPA	F	H	T	R	Q	S	A	F	W
	<i>CcPIP1;3</i>	NPA	NPA	F	H	T	R	E	S	A	F	W
	<i>CcPIP2;1</i>	NPA	NPA	F	H	T	R	Q	S	A	F	W
	<i>CcPIP2;2</i>	NPA	NPA	F	H	T	R	M	S	A	F	W
	<i>CcPIP2;3</i>	NPA	NPA	F	H	T	R	Q	S	A	F	W
	<i>CcPIP2;4</i>	NPA	NPA	F	H	T	R	Q	S	A	F	W
SIP	<i>CcSIP1;1</i>	NPT	NPA	T	V	P	N	M	A	A	Y	W
	<i>CcSIP2;1</i>	NP-	NPA	K	H	G	S	F	V	A	Y	W
	<i>CcSIP2;2</i>	NP-	*	K	*	*	*	*	*	*	L	F
TIP	<i>CcTIP1;1</i>	NPA	NPA	H	I	A	V	T	S	A	Y	W
	<i>CcTIP1;2</i>	NPA	NPA	Y	I	A	V	T	S	A	Y	W
	<i>CcTIP1;3</i>	NPA	NPA	H	I	A	V	T	S	A	Y	W
	<i>CcTIP2;1</i>	NPA	NPA	H	I	G	R	T	S	A	Y	W
	<i>CcTIP2;2</i>	NPA	NPA	H	I	G	R	T	S	A	Y	W
	<i>CcTIP3;1</i>	NPA	NPA	H	I	A	R	T	S	A	Y	W
	<i>CcTIP4;1</i>	NPA	*	H	*	*	*	L	*	*	*	*
	<i>CcTIP4;2</i>	NPA	NPA	H	I	A	R	T	S	A	Y	W
	<i>CcTIP5;1</i>	NPA	NPA	N	I	G	Y	T	S	A	Y	W
XIP	<i>CcXIP1;1</i>	NPI	NPA	V	F	A	R	C	C	A	F	W
	<i>CcXIP2;1</i>	NPI	NPA	L	F	A	R	C	C	A	F	W
	<i>CcXIP2;2</i>	NPI	NPA	L	F	A	R	C	C	A	F	W

6.6 Análise dos motivos proteicos de CcAQP em *C. canephora*

Os motivos conservados foram preditos usando MEME versão 5.1 (*Multiple EM for Motif Elicitation* <http://meme-suite.org/index.html>; BAILEY *et al.*, 2006). Entre os 15 motivos identificados, os motivos 1, 2 e 3 são amplamente encontrados nas cinco subfamílias; O motivo 4 está presente nas subfamílias PIP, TIP, XIP e em quatro membros NIP; O motivo 5 encontra-se exclusivamente na subfamília XIP; O motivo 6 está presente na apenas na subfamília NIP; O motivo 7 está presente na subfamílias PIP; O motivo 8 está amplamente presente nas subfamílias PIP, XIP e NIP; O motivo 9 está amplamente presente nas subfamílias PIP e TIP; O motivo 10 está limitado a subfamília PIP (Figura 10).

Figura 10- Distribuição de motivos (*motifs*) conservados entre AQPs de *C. canephora*, onde diferentes motivos são representados por diferentes blocos de cores, conforme indicado na parte inferior da figura, e o mesmo bloco de cores em diferentes proteínas indica um determinado motivo.



7 DISCUSSÃO

7.1 Caracterização de AQPs em *C. canephora*

O sequenciamento completo do genoma de *C. canephora* (DENOEUDE *et al.*, 2014) e a disponibilização desses dados na plataforma *Coffee Genome Hub* (<http://coffee-genome.org>) oferece a oportunidade para estudo de várias famílias de gênicas, visando uma melhor entendimento da estrutura, função e evolução.

As aquaporinas (AQPs) são proteínas do canal que facilitam o rápido transporte de água e pequenas moléculas através das membranas celulares das plantas (MAUREL *et al.*, 2015).

A análise filogenética mostrou que os CcAQPs foram agrupados e distribuídas em cinco subfamílias, como observado em plantas superiores, exceto em monocotiledôneas e em membros da família das Brassicas que não possuem a subfamília XIP (BARI *et al.*, 2018; SUN *et al.*, 2018; DIEHN *et al.*, 2015). Nesse trabalho, 33 CcAQPs foram identificadas no genoma de *C. canephora* e divididas em 5 subfamílias, distribuídas em 7 PIPs, 9 TIPs, 11 NIPs, 3 SIPs e 3 XIPs. Número similar de AQPs é encontrado em *A. thaliana* (35), foram divididas em apenas 4 subfamílias contendo: 13 PIPs, 10 TIPs, 9 NIPs e 3 SIPs (JOHANSON *et al.*, 2001) em *O. sativa* (33) sendo 10 NIPs, 10 TIPs, 11 PIPs e 2 SIPs em (SAKURAI *et al.*, 2005). Outras espécies como em *Lotus japonica* (DESHMUKH; BÉLANGER, 2016) e *Cicer arietinum* (DEOKAR; TAR'AN, 2016) a subfamília XIP também está ausente.

Em espécies como as solanáceas, tomate e batata há um número maior de AQPs que foi identificado e também subdivididas em 5 famílias pois estas também apresentam a subfamília XIP, Em *S. lycopersicum* foram encontrados 47 AQPs, destas são 14 PIPs, 11 TIPs, 12 NIPs, 4 SIPs e 6 XIPs (REUSCHER, S. *et al.*, 2013). Em *S. tuberosum* foram identificadas 15 PIPs, 11 TIPs, 10 NIPs, 3 SIPs e 8 XIPS, perfazendo 47 membros (VENKATESH, *et al.*, 2013). Outras espécies de café o número de AQPs também foram identificadas nesse trabalho e apresentaram um total de 52 em *C. arabica* e 37 em *C. eugenioides* ambas divididas nas cinco subfamílias. O número maior de AQPs encontrados em *C. arabica* pode ser devido a sua ploidia visto que é uma espécie tetraploide originada pela hibridação natural

entre os genomas de *C. canephora* e *C. eugenoides*, ambas diploides (LASHERMES *et al.*, 1999).

As sequências pertencentes a subfamília XIP de *C. canephora*, *C. arábica* e *C. eugenoides*, agruparam-se com as XIPs de *S. lycopersicum* e *S. tuberosum*, utilizadas para gerar a árvore filogenética. As sequências foram analisadas para a presença do motivo estendido NPARC (DANIELSON; JOHANSON; 2008), a qual caracteriza a subfamília XIP. Nas três espécies de café analisadas o referido motivo foi encontrado nessas sequências e ausentes nas demais subfamílias, confirmando assim serem pertencentes a subfamília XIP. Portanto algumas sequências nesse trabalho foram renomeadas e encontram-se diferentes dos nomes depositados no *NCBI*. Nenhuma das sequências estava identificada como XIP no *NCBI*, isso pode ser devido ao fato das sequências terem sido nomeadas usando a homologia com *A. thaliana*, que não possui a família XIP.

7.2 Estrutura Gênica

A comparação dos arranjos de íntrons e éxons (número e posição) entre os 33 *CcAPQ* revelou grande variação na estrutura gênica assim como observado em outras plantas tal como *Linum usitatissimum* (SHIVARAJ *et al.*, 2017); *R. cominus* (ZHOU *et al.*, 2015); *B. vulgaris* (KONG *et al.*, 2017).

Para as *CcAQPs* a massa molecular variou entre 12364.41 a 35232.98 Mw. As primeiras aquaporinas caracterizadas em planta possuíam propriedades estruturais típicas como peso molecular de 25-30 kDa caudas terminal N e C voltadas para o interior da célula, definindo assim seis domínios putativos que transpõe a membrana (MURATA *et al.*, 2000). Com o crescente número de espécies sequenciadas e um maior número de aquaporinas analisadas foi observada uma maior variação na massa molecular. As sequências *CcNIP5;1*, *CcTIP3;1* e as *CcXIP 1;1* e *1;2* apresentaram massa moleculares menores que 20 KDa inferiores aos observados em outras espécies como em *RcAQPs* 22.49–33.22 kDa (ZHOU *et al.*, 2015) e 25.09 a 35.08 em *BvAQPs* (KONG *et al.*, 2017).

Em relação ao números de transmembranas preditas pelo site *TMHMM Server v. 2.0* (<http://www.cbs.dtu.dk/services/TMHMM/>; KROGH; RAPACKI, 2016), variaram entre 2 e 7. Porém quando as sequências foram inspecionadas

visualmente e comparadas a outras espécies com *A. thaliana* (JOHANSON *et al.*, 2001) e *R. communis* (ZOU *et al.*, 2015) e *G. hirsutum* (PARK *et al.*, 2010), 6 domínios transmembranas (hélices) foram encontrados (Figura S4). Assim o programa *TMHMM Server v. 2.0* parece não ser eficiente em detectar todos os domínios TMHs.

7.3 Expressão *in silico*

Em geral, os genes *NIP* mostram um baixo nível de expressão e uma especificidade para um tipo de tecido e /ou célula com expressão padronizada (LIU *et al.*, 2009; ALEXANDERSSON *et al.*, 2009; SAKURAI *et al.*, 2005), o mesmo foi observado em *CcNIPS* em diferentes tecidos. Apesar da diversidade de *CcNIPs* a expressão em folhas é irrelevante, Takano *et al.* (2006) explica que *NIP5;1* é um transcrito específico de raiz, dado que *NIP6;1* se expressa especificamente dentro de tecido vascular do floema de nódulos jovens em desenvolvimento (TANAKA *et al.*, 2009), enquanto que *NIP7;1* tem expressão em tecidos florais, porém pode ser pouco expresso em siliqua e raiz (LI *et al.*, 2011).

O transporte de água realizado por XIP é muito baixo (BIENERT *et al.* 2011). Em *CcXIP* a expressão foi constitutiva e baixa com exceção de *CcXIP1;2* em pistilo. PIPs e TIPs apresentaram os maiores níveis de expressão *in silico* em *CcAQP*, outros trabalhos também reportaram resultados semelhantes em *B. vulgaris*, *Brassica rapa ssp. pekinensis*, *Arachis duranensis* e *Arachis ipaensis* (KONG *et al.*, 2017; TAO *et al.*, 2014; SHIVARAJ *et al.*, 2019)

7.4 Seletividade das proteínas codificadas pelos genes *CcAQP*

Os motivos duplos NPA (Asn-Pro-Ala), filtro de seletividade ar / R e posições de Froger, estão intimamente relacionadas à especificidade do substrato das proteínas AQP, devido a resolução atômica e o estímulo molecular. Os AQPs em plantas tem função de permeabilidade água, além de facilitar o transporte de algumas pequenas moléculas neutras, como glicerol, ácido bórico, ácido silícico, uréia, NH₃, CO₂ e H₂O₂ (LI *et al.*, 2014; AFZAL *et al.*, 2016). As sequências

preditas em *CcAQPs* foram verificadas visualmente em relação a presença de domínios e motivos característicos e funcionalmente importantes da família MIP. Para *C. canephora* verificou-se que o primeiro motivo NPA é marcadamente conservado em NIP, TIP e PIP, enquanto que o segundo em NIP é moderadamente variante, podendo ser substituído por NVI ou NPV.

Os motivos NPA e filtros de seletividade ar / R são semelhantes em PIP e são altamente conservados, resultado de uma forte pressão de seletividade realizada nos canais de água dos genes *PIP* e cada discrepância de aminoácidos pode ser considerado uma desvantagem seletiva para a planta, resultante de uma manutenção ativa (DIEHN *et al.*, 2015).

Em *C. canephora* há diversificação e abundância no subgrupo NIP, quando comparada a outras espécies superiores. Segundo Liu e Zhu (2010), essa variação na subfamília NIP pode ser consequência à uma reação adaptativa e evolutiva em relação a um ambiente mutável, onde a planta possa ter melhor resposta ao estresse e as doenças. Os NIPs em plantas são responsáveis pelo transporte, absorção, translocação de diversos metaloides não carregados (BIENERT; BIERNET, 2017). As AQPs de plantas apresentam-se divididas em cinco grupos principais, juntamente com a família NIP, que é composta por três subgrupos NIP I, NIP II e NIP III estes diferem sua seletividade para substratos (HOVE; BHAVE, 2011).

Em CcNIP3;1 mostrou ausência de algumas proteínas no filtro seletivo em LE1, nota-se também que no *loop E1* o motivo não é conservado. Em CcNIP4;2 há ausência em H2 e na posição P1 de Froger. Os domínios de proteína ausentes podem resultar em efeito negativo dominante nas funções da proteína, ainda que isso precise ser investigado. Porém em CcNIP3;1 mostrou ausência de algumas proteínas no filtro seletivo em LE1, nota-se também que no *loop E1* o motivo não é conservado em CcNIP4;2.

As estruturas de *CcAQPs* da subfamília PIP em *C. canephora* manteve-se quase que uniforme, com aminoácidos idênticos. Sabe-se que PIP2s são excelentes canais de água, seus filtros de seletividade e a composição de seus aminoácidos permanecem invariáveis. (HOOIJMAIJERS *et al.*, 2012).

A repetição do motivo (NPA) Asn-Pro-Ala na proteína é considerada fato consolidado para MIPs, mas exceções a esse regra já foram identificadas em NIPs e SIPs de plantas, (FORREST; BHAVE, 2007; DANIELSON; JOHANSON, 2008) o mesmo acontece em *C. canephora* onde verificou-se a substituição na terceira

posição do primeiro motivo NPA por (T/L) em todos os membros CcSIP, enquanto que em CcNIP apenas alguns membros sofreram mudanças no segundo motivo NPA, mostraram (V/ I).

Todos os membros da subfamília CcXIP mostraram (NPI) no primeiro motivo Asn-Pro-Ala, enquanto o segundo motivo foi ampliado para um motivo com resíduo de cisteína (NPARC), conservado em membros da subfamília XIP de outras plantas (REUSCHER, 2013; DANIELSON; JOHANSON, 2008). Algumas sequências de XIP em café arábica e eugeníodes estavam classificadas incorretamente ou não possuíam classificação no banco de dados, isso demonstra que possivelmente foram classificadas baseadas em *A. thaliana* na qual essa subfamília é ausente. Apesar dessas variações, os resíduos Asn e Pro permanecem perfeitamente conservados, mantendo a função apresentada para estes resíduos, tanto para estabilização do poro aquoso quanto na permeação de água (MURATA *et al.*, 2000).

As XIPs de plantas e fungos apresentam variação significativa de aminoácidos no primeiro motivo NPA e no filtro ar / R. Essa variabilidade também é observada em NIPs, possivelmente os XIPs se tornaram diversificados em termos de especificidade de substrato, enquanto conservam outras características de sequência específicas de de MIPs. Baseado nos quatro aminoácidos que definem os filtros ar / R, os XIPs de eudicotiledôneas podem ser divididos em quatro subclasses (GUPTA; SANKARARAMAKRISHNAN, 2009) dois dos quais têm uma assinatura de Ar / R semelhante em alguns NIPs de plantas, ao mesmo tempo os outros dois são ainda mais hidrofóbicos. Para mostrar a função de XIPs de plantas, foi realizada uma caracterização molecular e funcional completa de XIPs de *N. tabacum*, *S. tuberosum* e *S. lycopersicum* (BIENERT *et al.*, 2011).

8 CONCLUSÃO

Este estudo é o primeiro a apresentar a identificação e análises *in silico* dos genes da família das aquaporinas em *C. canephora*. A caracterização dessa família de genes permite a selecionar genes candidatos para a validação funcional em resposta aos estresses abióticos nessa espécie com grande importância econômica.

CAPÍTULO II

ANÁLISE DA EXPRESSÃO GÊNICA DE MEMBROS DA FAMÍLIA CcAQP EM DOIS GENÓTIPOS CONTRASTANTES de *C. canephora* SUBMETIDO AO DÉFICIT HÍDRICO

RESUMO

Análise da expressão gênica de membros da família CCAQP em dois genótipos contrastantes de *C. canephora* submetido ao déficit hídrico

O Brasil é o maior produtor e exportador mundial de café. A espécie *Coffea canephora*, conhecida como Robusta é utilizada para a indústria de café solúvel. O déficit hídrico é um dos maiores limitantes para a produção agrícola. A resposta ao estresse hídrico em nível molecular ocorre pela alteração da expressão de vários genes, dentre eles os da família das aquaporinas. Essa família de genes desempenham papel-chave na regulação da condutância hidráulica regulando o fluxo de água para dentro e fora das células. O objetivo desse trabalho foi avaliar a expressão de membros da família CcAQPs por RT- qPCR em dois genótipos contrastantes de *C. canephora* submetidas ao déficit hídrico. Sendo, o clone 14 tolerante e o clone 109A suscetível à seca; As plantas foram submetidas a quatro tratamentos: (1) sem estresse (ψ – 1,34 MPa), (2) estresse moderado (ψ – 2,39 MPa), (3) estresse severo (ψ – 4,5 MPa) e (4) re-irrigado até o potencial similar ao sem estresse. O RNA foi extraído das folhas e convertidos em cDNA. Foram desenhados iniciadores para os membros CcPIPs1;2, 2;3; 2;4, duas CcTIPs1;2, 2;1. A quantificação da expressão relativa para cada gene foi calculada pelo método ΔC_t onde nível de expressão dos genes alvo (CcAQPs) menos do gene normalizador *UBI10*. O déficit hídrico afeta a expressão relativa das isoformas CcPIPs e CcTIPs em folhas de *C. canephora*. O clone 14 mostrou redução na expressão enquanto que o clones 109A apresentou aumento no nível na expressão relativa quando as plantas foram submetidas ao estresse severo.

Palavras chave: Seca; Aquaporina; Café

ABSTRACT

Analysis of gene expression of *CCAQP* family members in two *C. canephora* genotypes contracting submitted to water deficit

Brazil is the world's largest producer and exporter of coffee. The species *Coffea canephora*, known as Robusta is used for the instant coffee industry. Water deficit is one of the major constraints on agricultural production. The response to water stress at the molecular level occurs by altering the expression of several genes, including those in the aquaporin family. This gene family plays a key role in regulating hydraulic conductance by regulating the flow of water into and out of cells. The aim of this study was to evaluate the expression of *CcAQPs* family members by RT-qPCR in two *C. canephora* genotypes contracting submitted to water deficit. Therefore, clone 14 tolerant and clone 109A susceptible to drought; The plants were submitted to four treatments: (1) no stress ($\psi - 1.34$ MPa), (2) moderate stress ($\psi - 2.39$ MPa), (3) severe stress ($\psi - 4.5$ MPa) and (4) re-irrigated even potential similar to stress free. RNA was extracted from leaves and converted to cDNA. Primers were designed for the *CcPIPs*1, 2, 2,3; 2,4, two *CcTIPs*1, 2, 2, 1 Quantification of relative expression for each gene was calculated by the ΔCt method where expression level of the target genes (*CcAQPs*) less than the normalizing gene *UBI10*. Water deficit affects the relative expression of *CcPIPs* and *CcTIPs* isoforms in *C. canephora* leaves. Clone 14 showed reduction in expression while clones 109A showed increase in relative expression level when plants were subjected to severe stress.

Keywords: Drought; Aquaporine; Coffee.

1 INTRODUÇÃO

O café é uma das mais importantes commodities do comércio internacional, a espécie diploide perene *C. canephora* ($2n = 2x = 22$) representa aproximadamente 40% da produção mundial de acordo com a OIC (ORGANIZAÇÃO INTERNACIONAL DO CAFÉ, 2019). Os estresses abióticos, principalmente o déficit hídrico é um dos fatores mais limitantes para a produção de café na maioria dos países produtores (DA MATTA *et al.*, 2006). Os genótipos podem ser caracterizados como tolerante ou suscetível ao estresse hídrico em função dos parâmetros fisiológicos e respostas oxidativas que apresentam. Em *C. canephora* var. Conillon o clone 14 foi caracterizado como tolerante, enquanto que o clone 109A como suscetível à seca (MARRACINI *et al.*, 2011; LIMA *et al.*, 2002; 2015).

Em nível molecular, a resposta ao estresse hídrico ocorre pela alteração da expressão de vários genes, dentre eles os da família das aquaporinas. Em plantas, as aquaporinas ocorrem em múltiplas isoformas e desempenham papéis-chave na regulação da condutância hidráulica em raízes e folhas, regulando o fluxo de água para dentro e fora das células, e a transferência de água através das membranas (SHEKOOFA; SINCLAIR, 2018). Essa família de proteínas é de fundamental importância devido ao seu papel no transporte de água nas células vegetais, pois afeta todos os processos de desenvolvimento e crescimento das plantas

No entanto, existem poucas informações sobre a expressão de genes da superfamília da AQP no gênero *Coffea* em relação ao déficit hídrico. Os primeiros estudos da família das AQPs no gênero *Coffea* (SANTOS; MAZZAFERA, 2013), foram realizados utilizando bibliotecas EST (*Expressed Sequence Tags*) de *C. arabica*, *C. canephora* e *Coffea racemosa* (VIEIRA *et al.*, 2006) sendo que apenas seis unigenes completos foram identificados e quatro genes (3 PIPs e 1 TIP) foram analisados revelando que as isoformas *PIP2;1* e *PIP2;2* foram reguladas positivamente nas folhas e raízes pelo déficit hídrico, sugerindo o possível envolvimento desses genes no controle do estado hídrico das plantas e na recuperação pós-estresse (SANTOS; MAZZAFERA, 2013).

O sequenciamento completo do genoma de *C. canephora* (DENOEUDE *et al.*, 2014) e a disponibilização dos dados na plataforma *Coffee Genome Hub* (coffee-genome.org/) abriu novas possibilidades para uma investigação mais completa sobre a família das aquaporinas em *C. canephora*. No capítulo 1 desse trabalho foi

realizada a caracterização *in silico* de 33 sequências de aquaporinas encontradas no genoma dessa espécie e permitindo a seleção de genes candidatos para experimentos de validação funcional.

2 OBJETIVO

Avaliar a expressão de membros candidatos da família *CcAQPs* por RT-qPCR em dois genótipos contrastantes de *C. canephora* submetidas ao déficit hídrico.

3 MATERIAL E MÉTODOS

3.1 Experimento de estresse hídrico

O experimento foi realizado em parceria com o Instituto Agronômico do Paraná (IAPAR), localizado em Londrina – PR em condições de casa de vegetação climatizada. Ao todo, foram utilizadas dez plantas de dois genótipos de *C. canephora* var. Conillon caracterizadas como sendo tolerante à deficiência hídrica (clone 14: tolerante à seca - D^T e clone 109A: suscetível à seca - D^S) com a mesma idade e tamanho similar (~1 ano). As plantas foram transplantadas em vasos com aproximadamente 10 Kg de substrato Bioterra®. Após o estabelecimento das plantas. A irrigação foi suspensa até que as plantas atingiram os níveis de estresse descritos a seguir: sem estresse (ψ – 1,34 MPa), estresse moderado (ψ – 2,39 MPa), estresse severo (ψ – 4,5 MPa) e re-irrigado até as plantas atingirem o potencial hídrico estabelecido como sem estresse. As coletas dos materiais para análise foram realizadas conforme as leituras do potencial hídrico. Para cada coleta de folhas foi realizado um *pool* de dois pares de folhas de cada vaso na mesma condição. As folhas coletadas foram imersas imediatamente em N líquido e armazenadas em freezer – 80°C até o momento de análise.

O potencial de água na folha foi monitorado através de psicrômetros de termopar acoplados a um datalogger. A leitura foi feita ao amanhecer e outra ao meio dia. Foram coletadas amostras de aproximadamente 2 cm² da parte mediana da folha do ramo plagiotrópico posicionado no meio da planta (folha totalmente madura e sem sinais danos ou deficiências) e colocada nos psicrômetros. O datalogger foi programado para que as leituras fossem efetuadas a cada 10 min até o equilíbrio da pressão de vapor na câmara. A microvoltagem fornecida pelo sistema foi convertida em potencial da água (MPa) em função de prévia calibração dos sensores com soluções de cloreto de sódio. Após a obtenção do potencial total da água (ψ_t), os sensores foram imersos durante 4 min em nitrogênio líquido e as leituras retomadas para obtenção do potencial osmótico (ψ_s). O cálculo do potencial de pressão foi realizado através de $\psi_t - \psi_s$. A coleta das plantas para a realização dos *pool/s* foi baseada no potencial hídrico em função dos níveis de estresse

estabelecidos, realizados através do monitoramento do potencial hídrico de cada planta por psicrometria.

3.2 Extração de RNA total

Foram realizadas as extrações de RNA total de folhas de *C. canephora* para os experimentos de déficit hídrico descritos anteriormente. As extrações foram conduzidas usando o tampão de extração utilizando o kit PureLink RNA Reagent (Invitrogen, Carlsbad, CA, USA). Todo material de manipulação foi esterilizado a seco overnight (180°C), para evitar a ação de RNases e possível degradação dos ácidos nucleicos e proteínas. Foram usados 500 mg do material macerado em nitrogênio líquido, segundo as recomendações do fabricante. O sobrenadante foi descartado e o pelete ressuspensionado em H₂O tratada com DEPC (dietil pirocarbonato).

Após a extração foi verificada a integridade do RNA das amostras, através de eletroforese em gel de agarose 1% (p/v), com tampão TAE (2M Tris – acetato, 0,05 M EDTA), contendo brometo de etídio (0,5 mg mL⁻¹) e visualizados sob luz UV em transluminador.

3.3 Síntese de cDNA

A concentração do RNA e a pureza foram determinadas usando espectrometria a 260 e 280 nm e as razões 260/280 e 260/230 com um Biophotometer®. Foram utilizadas as amostras que apresentaram razão acima de 1,8. O DNA complementar (cDNA) foi sintetizado por um Transcriptase Reversa SuperScript IV® (Invitrogen), seguindo as instruções do fabricante, em volume final de 20 µl utilizando 5 µg de RNA total. Para cada 1 µg do RNA total extraído do tecido do cafeeiro e misturado com 500 ng de Oligo (dT)₁₂₋₁₈ e 1 mM de dNTPmix. A remoção do RNA original foi feita por meio de incubação dessa reação a 37°C por 20 minutos com 2 unidades de RNase H. Os produtos finais de cDNA foram diluídos 10 vezes antes de serem utilizados em RT-qPCR.

3.4 Desenho de *Primers*

As sequências CDS dos genes de *CcAQP* foram buscadas no banco de dados do *Coffee Genome Hub* (<http://coffee-genome.org/>) e alinhadas através da plataforma *ClustalW* (<https://www.genome.jp/tools-bin/clustalw>; THOMPSON; HIGGINS; GIBSON, 1994) para identificar regiões conservadas e desenhar os iniciadores específicos fora dessa região para cada isoforma. Para o desenho foi utilizado o programa Primer Express® v3.0,1 (Applied Biosystems). Para escolha dos iniciadores foram utilizados os seguintes parâmetros: temperatura de anelamento de 58 a 60°C, amplicons de 80 a 150 pares de bases, comprimento de 9 a 40 pares de bases e conteúdo GC de 30 a 80%. A confirmação da especificidade foi realizada através da ferramenta *Blast* disponível em *Coffee Genome Hub* (<http://coffee-genome.org/blast>). Os iniciadores foram avaliados também através da plataforma *Multiple Primer Analyzer* (ThermoFisher) (<https://www.thermofisher.com/multiple-primer-analyzer.html>) para verificar a formação de estruturas secundárias (dímeros e grampos – *hairpins*). Foram utilizadas para a tentativa de desenho dos iniciadores as isoformas de cada subfamília (PIPs, NIPs, TIPs, SIPs E XIPs) que apresentaram altos níveis de expressão *in silico* para folhas (Figura 5, Capítulo I). Também foram testados os genes *CcEF1α* (SANTOS *et al.*, 2015), *CcUBQ10* (MARRACINI *et al.*, 2012) e *CcGAPDH* (desenhados para esse trabalho) como normalizadores para os cálculos da expressão relativa.

A amplificação dos iniciadores desenhados foram testados utilizando amostras em *bulk* (agrupadas) em PCR convencional. O produto das amplificações foram observadas por eletroforese em gel de agarose 1,5% (p/v), em tampão TAE (2M Tris – acetato, 0,05 M EDTA), contendo brometo de etídio (0,5 mg mL⁻¹) e visualizados sob luz UV em transluminador.

3.5 Eficiência de amplificação dos iniciadores

Primeiramente foram realizadas amplificações para avaliar a eficiência dos iniciadores. As reações de RT-qPCR foram realizadas em placas utilizando 5 µl de SYBR Green, 0,4 µl do iniciador senso (5 µM), 0,4 µl do iniciador antisenso (5 µM), 1

μL do *pool* de cDNA, completando o volume final para 10 μL com água Milli-Q autoclavada. O *pool* de cDNA foi feito com todas amostras para testar a eficiência dos iniciadores. Foi incluso também uma amostra de controle negativo (sem cDNA). Cada reação foi montada em triplicata. As placas foram centrifugadas brevemente (*spin*) e analisadas pelo equipamento StepOnePlus™ Real-Time PCR System (Applied Biosystems), seguindo as seguintes condições: desnaturação a 95°C por 2 min, anelamento a 95°C por 30 s e extensão a 60°C por 30 s, ciclo este repetido 40 vezes. Ao término da reação, a curva de dissociação foi determinada para cada par de iniciadores visando verificar a presença de somente um pico específico para cada gene e ausência de dímeros. Os dados foram exportados para o Excel e a eficiência de amplificação individual de cada par de iniciadores, foi calculada utilizando o programa *LinRegPCR* (RAMAKERS *et al.*, 2003).

3.6 Análise da expressão gênica

Os perfis transcricionais dos genes selecionados foram detectados a partir do equipamento *StepOnePlus™* Real-Time PCR System (Applied Biosystems) utilizando 5 μL de SYBR Green, 0,4 μL do iniciador senso (5 μM), 0,4 μL do iniciador antisenso (5 μM), 1 μL do cDNA 1:10 (v/v), completando o volume final para 10 μL com água Milli-Q autoclavada. Para cada amostra, as reações foram montadas em triplicata para cada uma das três repetições biológicas. A amplificação foi conduzida seguindo as seguintes condições: 40 ciclos de desnaturação a 95°C por 2 min, anelamento a 95°C por 30 s e extensão a 60°C por 30 s. A quantificação da expressão relativa para cada gene foi calculada pelo método ΔCt onde nível de expressão dos genes alvo (*CcAQPs*) menos do gene normalizador. Em todas as análises, os níveis de transcrição dos genes alvo foram normalizados contra o perfil transcricional do gene *CcUBQ10* (MARRACINI *et al.*, 2012).

Análises estatísticas entre os diferentes tratamentos para cada clone testado foi realizado usando one-way ANOVA seguido pelo teste de Tukey teste ($p > 0,05$). A comparação entre os dois clones em cada tratamento foi realizada pelo teste t para comparação de duas médias (* $p < 0,05$).

4 RESULTADOS

Após a realização do experimento de imposição do déficit hídrico nos clones 14 e 109A de *C. canephora* foi realizada a extração do RNA das folhas de todos os tratamentos e em seguida foi realizada a síntese do cDNA.

Inicialmente todas as isoformas foram utilizadas para a tentativa de desenho de iniciadores para os experimentos de RTqPCR. Os iniciadores desenhados foram checados por *Blast* para a especificidade da isoforma, presença de estruturas secundárias (dímeros e grampos). Foram descartadas as sequências para as quais não foi possível desenhar iniciadores específicos e/ou apresentaram a formação de estruturas secundárias restando então oito genes: três *CcPIPs*1;2, 2,3; 2,4, duas *CcTIPs*1;2, 2;1, duas *NIPs*7;1, 5,2 e uma *CcSIP*1;1 (Tabela 1). Não foi possível obter um par de iniciadores adequados para nenhuma das *CcXIPs*. A especificidade de amplificação dos oligonucleotídeos iniciadores desenhados para diferentes isoformas foi testada por PCR convencional utilizando um *bulk* das amostras (Figura 1) e para os normalizadores (Figura 2). Foi observada a amplificação de uma única banda tanto para os *CcAQPs* quanto para os normalizadores com exceção da *CcNIP* 7;1 (Figuras 1 e 2 respectivamente).

A eficiência de ampliação dos iniciadores foi testada pela PCR em tempo real utilizando o mesmo *bulk* de amostras. A curva de dissociação para cada iniciador foi obtida confirmando a presença de amplificação única pela obtenção de um único pico para a maioria dos genes *CcAQPs* com exceção da *CcSIP*1;1 a qual apresentou picos múltiplos e para a *CcNIP*7;1 que apresentou um dois picos (Figura 3). Os cálculos de eficiência de amplificação foram realizados de acordo com os valores médios da obtidos pela análise individual de cada curva de amplificação para cada gene. As sequências dos iniciadores, tamanho do amplicon e eficiência são apresentados na Tabela 1. A eficiência de amplificação para cada um genes variou entre 95,75 e 103,4% sendo considerados adequadas para uso em estudos de RTqPCR (Tabela 1). Visto que os genes *CcSIP*1;1 e *CcNIP*7;1 apresentaram picos múltiplos, a eficiência de amplificação não foi calculada para esses genes.

Tabela 1- Iniciadores desenhados para análise da expressão relativa por RTqPCR em *CcAQPs*.

Nome do gene	Descrição	ID	Locus <i>C. canephora</i>	Iniciador senso (5'-3')	Iniciador antisenso (5'-3')	Amplicon (pb)	Eficiência (%)*
<i>CcPIP1;2</i>	Proteína intrínseca da membrana plasmática	Cc02_g38120	chr2:52874143..52875950	GGTTTCAGCCATCGCCATAT	TTTCAGCACCAAGCCCATCT	90	95,75
<i>CcPIP2;3</i>	Proteína intrínseca da membrana plasmática	Cc06_g03900	chr6:3002393..3003704	TTCATCGCCACCCTCTTGTT	TTGGCCTTGGTGTGAGATTG	80	102,8
<i>CcPIP2;4</i>	Proteína intrínseca da membrana plasmática	Cc09_g02170	chr9:1811779..1813234	CATGGTGGCTCAGTGTGAGGT	TCCTCCGCCGTATGTGACAT	88	103,4
<i>CcTIP1;2</i>	Proteína intrínseca do tonoplasto	Cc01_g10300	chr1:28971849..28973231	GGCGCATGGTGTATCTGTGT	GGCTGTGGCATAGACGGTGTA	85	99,0
<i>CcTIP2;1</i>	Proteína intrínseca do tonoplasto	Cc04_g02840	chr4:2193273..2194443	GCCATGGCCTACAGCAAACCT	GAGCAAAACCATGGCAAACCTG	85	99,3
<i>CcNIP5;2</i>	Proteína intrínseca semelhante a nodulina	Cc08_g17120	chr8:31492021..31498389	GCAGAATCGGATGGTGGAA	CGAACCGACGAAAACAATGG	80	99,3
<i>CcNIP7;1</i>	Proteína intrínseca semelhante a nodulina	Cc03_g03200	chr3:2417270..2418614	TCCAGGTCTCCGGAAGGTTT	TGCTCGCAATGATTCCACAT	80	----
<i>CcSIP1;1</i>	Pequena proteína intrínseca básica	Cc00_g09500	chr0:81327025..81329499	TCATCACCACGGTGCTTGTT	AGACCGACGGCATGAAAGG	115	----
<i>UBQ10</i>	Ubiquitina 10	GW488515		AAGACAGCTTCAACAGAGTACAGCAT	GGCAGGACCTTGGCTGACTATA	104	Marracini <i>et al.</i> , 2012
<i>CcEF1α</i>	Fator de alongamento 1α	Cc01_g19830	chr1:36565137..36570059	GCATCTTTAGTTTCTCCTCGCTTC	GCGGTGGACGCTTCATTT	82	-----
<i>GAPDH</i>	Gliceraldeído-3- fosfato	GW445811	—	TTGAAGGGCGGTGCAAA	AACATGGGTGCATCCTTGCT	59	Marracini <i>et al.</i> , 2012

*Os valores de eficiência foram obtidos a partir do programa LinRegPCR (RAMAKERS *et al.*, 2003).

Figura- 1 Amplificação de fragmentos dos genes *CcAQPs* utilizando o cDNA *bulk* pela reação em cadeia da polimerase e corrida em gel de agarose 1,5%. O teste foi realizado em duplicata. M – marcador ladder 100 pb; B branco, 1-2: *CcPIP1*;2, 3 e 4 *CcPIP2*;3, 5 e 6 *CcTIP1*;2, 7 e 8 *CcTIP1*;2, 9 e 10 *CcSIP1*;1, 11 e 12 *CcNIP7*;1, 13 e 14 *CcNIP5*;2, 15 e 16 *CcPIP2*;4.

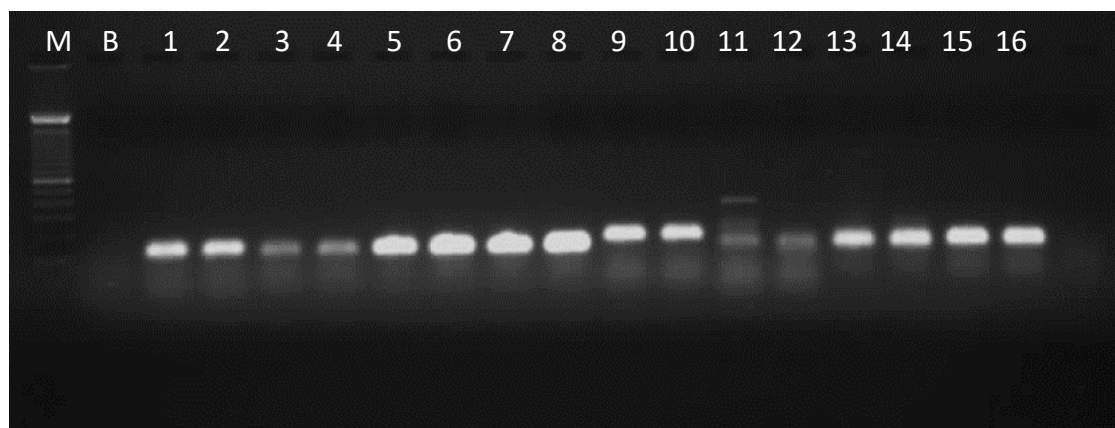


Figura- 2 Amplificação de fragmentos dos genes normalizadores com o *bulk* de cDNAs pela reação em cadeia da polimerase e corrida em gel de agarose 1,5%. M – marcador ladder 100 pb, 1 – branco, 2 – *CcUBQ10*, 3 – *CcGAPDH*, 4 - *CcEF1 α* .

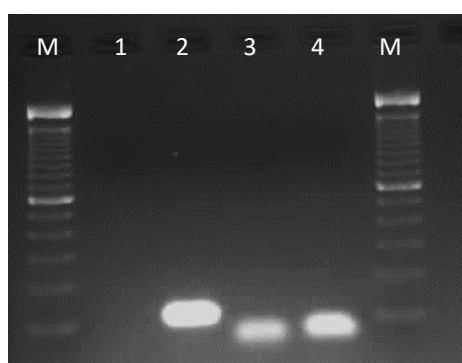
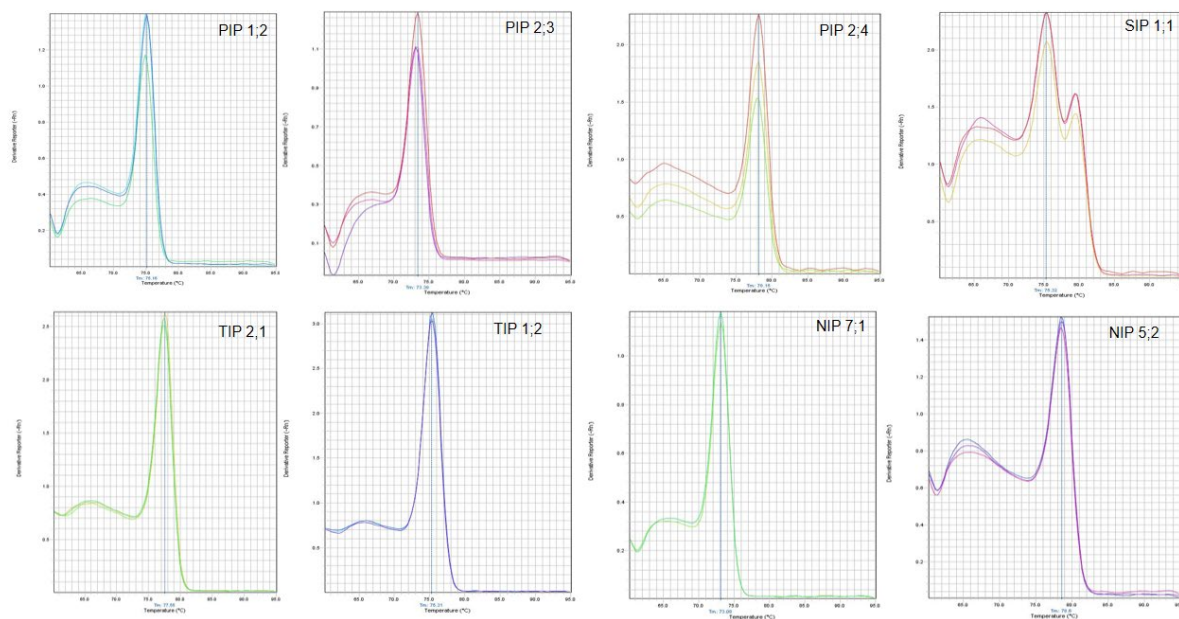


Figura- 3 Curvas de dissociação produzidas pelo produtos de amplificação por RTqPCR para os *CcAPQs* utilizados neste estudo (cada curva inclui três replicatas técnicas do conjunto de cDNA de todas as amostras).



Para a análise da expressão relativa foram descartados os genes *CcSIP1;1* e *CcNIP7;1* por apresentarem amplificações inespecíficas. O gene *CcNIP5;2* também foi descartado por apresentarem baixa expressão *in silico* em folhas (Figura 4, Capítulo 1). Portanto, cinco genes sendo três *CcPIP1;2*, *2;3*, *2;4* e duas *CcTIP2;1*, *2;2* foram selecionados para as análises da expressão relativa por apresentarem especificidade de amplificação e alta expressão *in silico* em folhas (Figura 4, Capítulo 1).

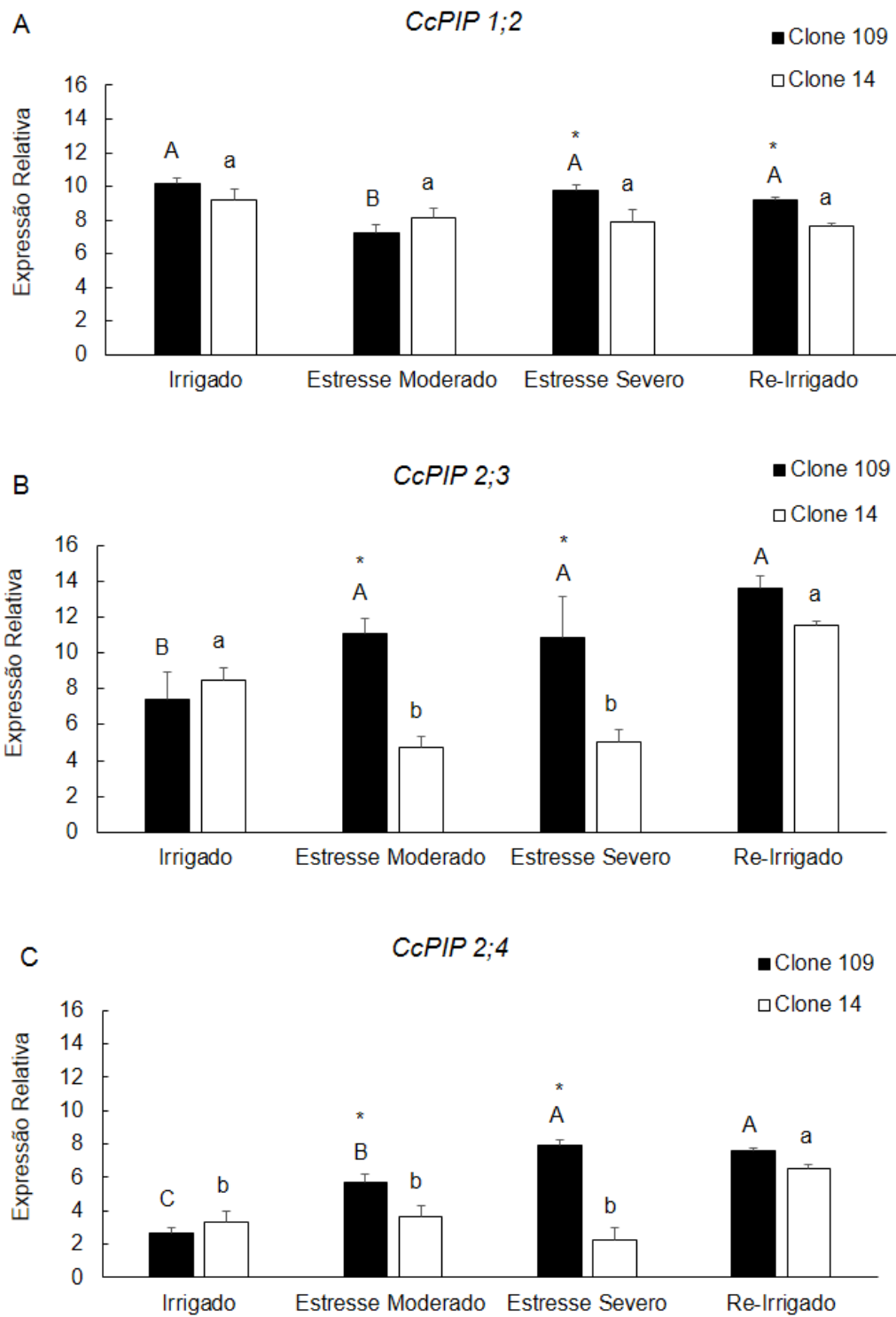
A expressão relativa em folhas de *C. canephora* foi calculada para cada uma das isoforma de *CcAQP* selecionada contra a expressão do gene normalizador *CcUB10* (ubiquitina) (MARRACINI *et al.*, 2012). Esse normalizador foi selecionado por apresentar a maior expressão (Figura 2). Para a isoforma *CcPIP1;2* a expressão do clone 109A foi significativamente mais elevada no estresse severo e na re-irrigação em relação ao clone 14, porém não diferiram dos valores expressos no tratamento irrigado para cada clone avaliado. (Figura 3). A expressão relativa das isoformas *CcPIP2;3* e *CcPIP2;4* foi significativamente mais elevada nos estresses

moderado e severo no clone 109A quando comparados ao clone 14. A comparação dentro de cada clone revelou que para o clone 109A o estresse severo e a re-irrigação diferiram do estresse moderado e este por sua vez diferiu do irrigado. Já para o clone 14 apenas o re-irrigado diferiu dos demais tratamentos. Em relação as três isoformas de *CcPIPs* o nível de expressão em *CcPIP2,4* foi mais baixo do que o observado para as outras isoformas analisadas nesse trabalho.

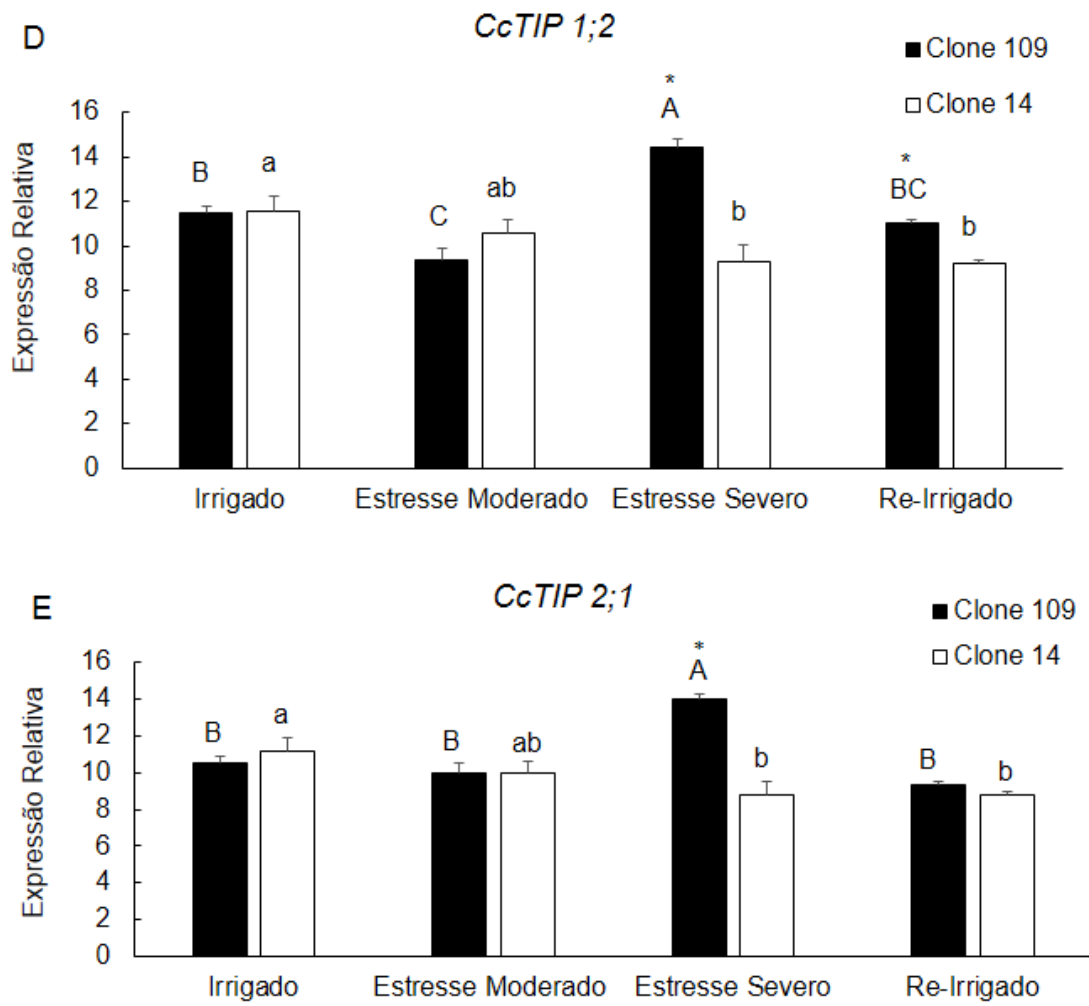
Para as duas isoformas de *CcTIPs* (1;2 e 2;1) analisadas o clone 109^a demonstrou aumento significativo da expressão relativa quando as plantas foram submetidos ao estresse severo e também no tratamento re-irrigado para a *CcTIP1;2*. Para o clone 14 em ambas as isoformas de *CcTIP* houve redução na expressão relativa nos tratamentos estresse severo e re-irrigado quando comparados ao tratamento irrigado.

Em geral, considerando as cinco isoformas (três *CcPIPs* e duas *CcTIPs*) analisadas nesse trabalho, o clone 109A apresentou aumento no nível na expressão relativa nas quando submetido as plantas e café foram submetidas ao estresse severo.

Figura 4. Expressão relativa de gene AQPs em folhas de *C. Canephora* submetidas a diferentes condições de irrigação: sem estresse (ψ – 1,34 MPa), estresse moderado (ψ – 2,39 MPa), estresse severo (ψ – 4,5 MPa) e re-irrigado até as plantas atingirem o potencial hídrico estabelecido como sem estresse. Os resultados são expressos em níveis de expressão relativa dos genes alvo em relação a expressão do gene da ubiquitina como referência interna (método ΔCt). Os valores representam a média ($\pm$ DP) calculada a partir de três repetições biológicas em cada tratamento. A análise estatística entre os diferentes tratamentos foi realizada por meio de ANOVA unidirecional seguida do teste de Tukey ($p < 0,05$). Letras maiúsculas e minúsculas indicam comparações em diferentes condições de irrigação para os clones 109A e 14, respectivamente a comparação entre os clones em cada tratamento foi realizada pelo teste t (* $p < 0,05$).



Continua ...



5 DISCUSSÃO

Neste trabalho foi realizada a caracterização de cinco isoformas de genes aquaporinas (3 *CcPIPs* e 2 *CcTIPs*) em resposta à seca, em dois cultivares contrastantes *C. canephora*. O clone 14 e o 109A são caracterizados como tolerante e suscetível ao déficit hídrico, respectivamente (MARRACINI *et al.*, 2011; LIMA *et al.*, 2002; 2015). A expressão foi medida em folhas, e entre as *CcPIPs* as isoformas *CcPIP2;3* e *CcPIP2;4* apresentaram redução na expressão para o clone 14 enquanto as isoformas *CcPIP1;2* e *CcPIP2;3* tiveram aumento da expressão durante o estresse e a reidratação para o clone 109A. Já entre as *CcTIPs* ambas apresentaram perfis transicionais semelhantes, no qual o aumento no nível de expressão foi observado apenas para o genótipo suscetível (109A) quando submetido ao estresse severo. Em um experimento com genótipos contrastantes de feijão comum em relação a tolerância a seca revelou que o déficit hídrico induziu uma redução da expressão gênica em *PvPIP2;7* e *PvTIP1;1* sendo revertido na reidratação. Enquanto que a expressão dos *PvPIP2;7* *PvTIP4;1* foi reduzida na cultivar mais tolerante à seca (ZUPIN *et al.*, 2017). Esse fato também foi observado em *C. canephora*, em geral a expressão durante o estresse severo aumentou no genótipo suscetível (clone 109A) e reduziu no tolerante (clone 14).

Tem sido sugerido que as AQPs em plantas têm um papel importante na indução na regulação da transpiração em solos secos e em elevado déficit de pressão de vapor os quais permitem que as plantas conservem a água do solo para sustentar as condições fisiológicas das plantas (SHEKOOFA; ZUPIN, 2018). Uma análise geral dos 35 AQPs de aquaporina em *A. thaliana* submetidas ao déficit hídrico revelou que a maioria dos genes PIPs e TIPs foram regulados negativamente, ou seja, tiveram redução na expressão podendo ser uma maneira para a planta minimizar o fluxo de água através membranas celulares e manter o turgor das folhas, por outro lado, pode refletir uma regulação negativa geral da transcrição/tradução de genes não cruciais para a planta durante o estresse hídrico (ALEXANDERSSON *et al.*, 2005).

A identificação de isoformas diferencialmente reguladas em resposta ao déficit hídrico pode permitir a seleção de genes candidatos para a produção de cultivares tolerantes a seca utilizando ferramentas biotecnológicas. Nesse estudo, as

isoformas *CcPIPs*2;3 e 2;4 isoladas a partir do clone 14 seriam genes candidatos potenciais para novos estudos.

A avaliação de genes aquaporinas em genótipos contrastantes de *Phaseolus vulgaris* (tolerantes e suscetíveis ao estresse hídrico) sugerem que a capacidade da planta conservar água durante o déficit hídrico em cultivares tolerantes a seca envolvem a redução da expressão de isoformas específicas aquaporinas (ZUPIN *et al.*, 2017).

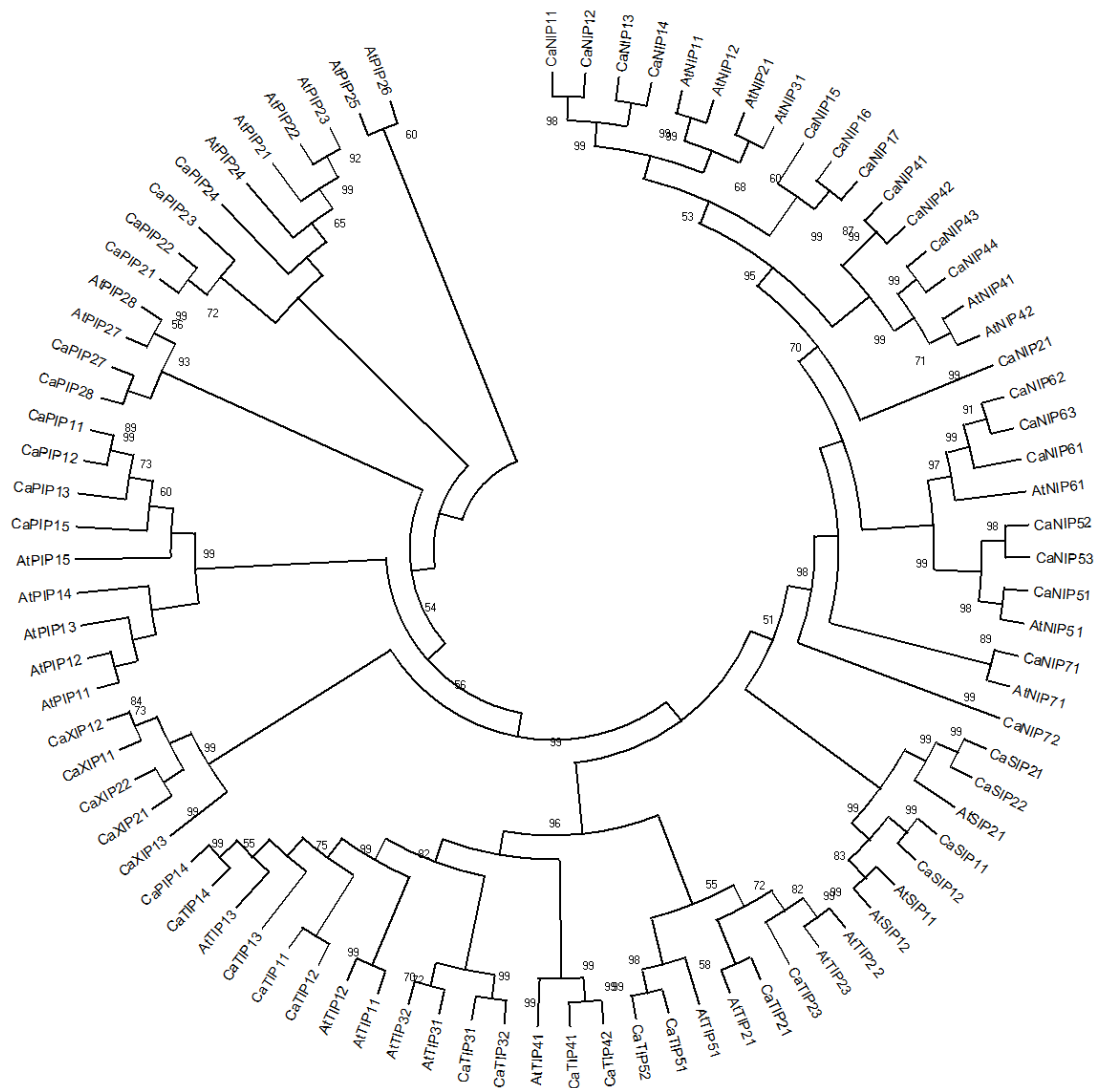
Uma análise global da expressão dos genes AQPs em *Citrus sinensis* submetidos ao déficit hídrico revelou que todas as *CsTIPs* e *CsXIPs* duas *CsPIPs* (1;1 e 2;4) uma *CsNIP1;1* e uma *CsSIP1;21* tiveram aumento na expressão em folhas (MARTINS *et al.*, 2015) indicado que cada isoforma responde de forma específica a falta de água.

6 CONCLUSÃO

O déficit hídrico afeta a expressão relativa das isoformas CcPIPs e CcTIPs em folhas de *Coffea canephora* avaliadas nesse trabalho. A diferença nos níveis de expressão das aquaporinas entre os clones 109A e 14 pode ser utilizado como ferramenta para auxiliar a caracterização dos genótipos em suscetível ou tolerante ao déficit hídrico e para selecionar genes candidatos para a tolerância à seca.

7 MATERIAL SUPLEMENTAR

Figura- S1. Árvore filogenética de *C. arabica* e *A. thaliana* usando o parâmetro padrão de alinhamento *ClustalW*. O dendrograma filogenético foi gerado pelo MEGA X usando o método de *Maximum Likelihood* com 1.000 réplicas de *bootstrap*.



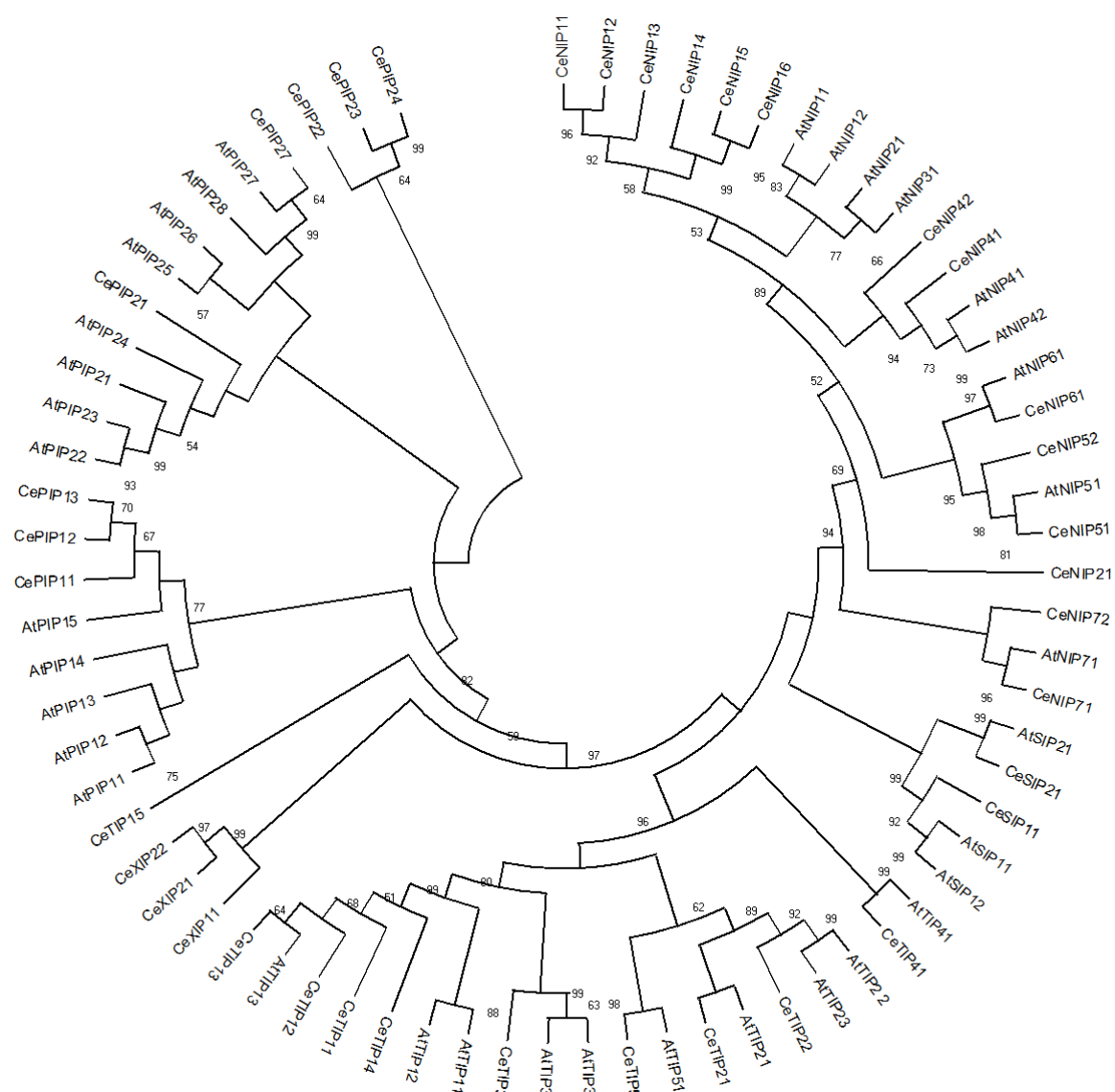
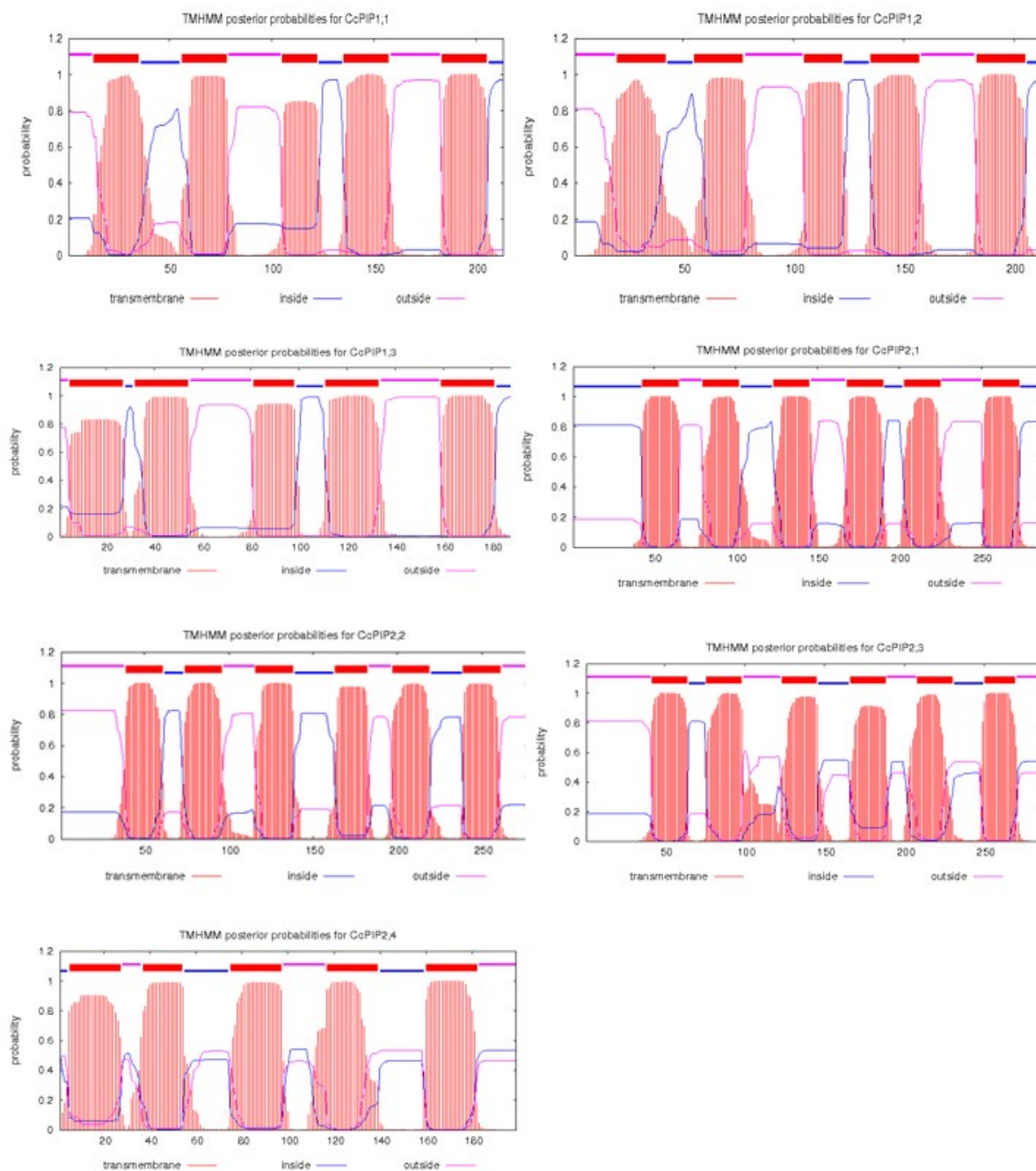


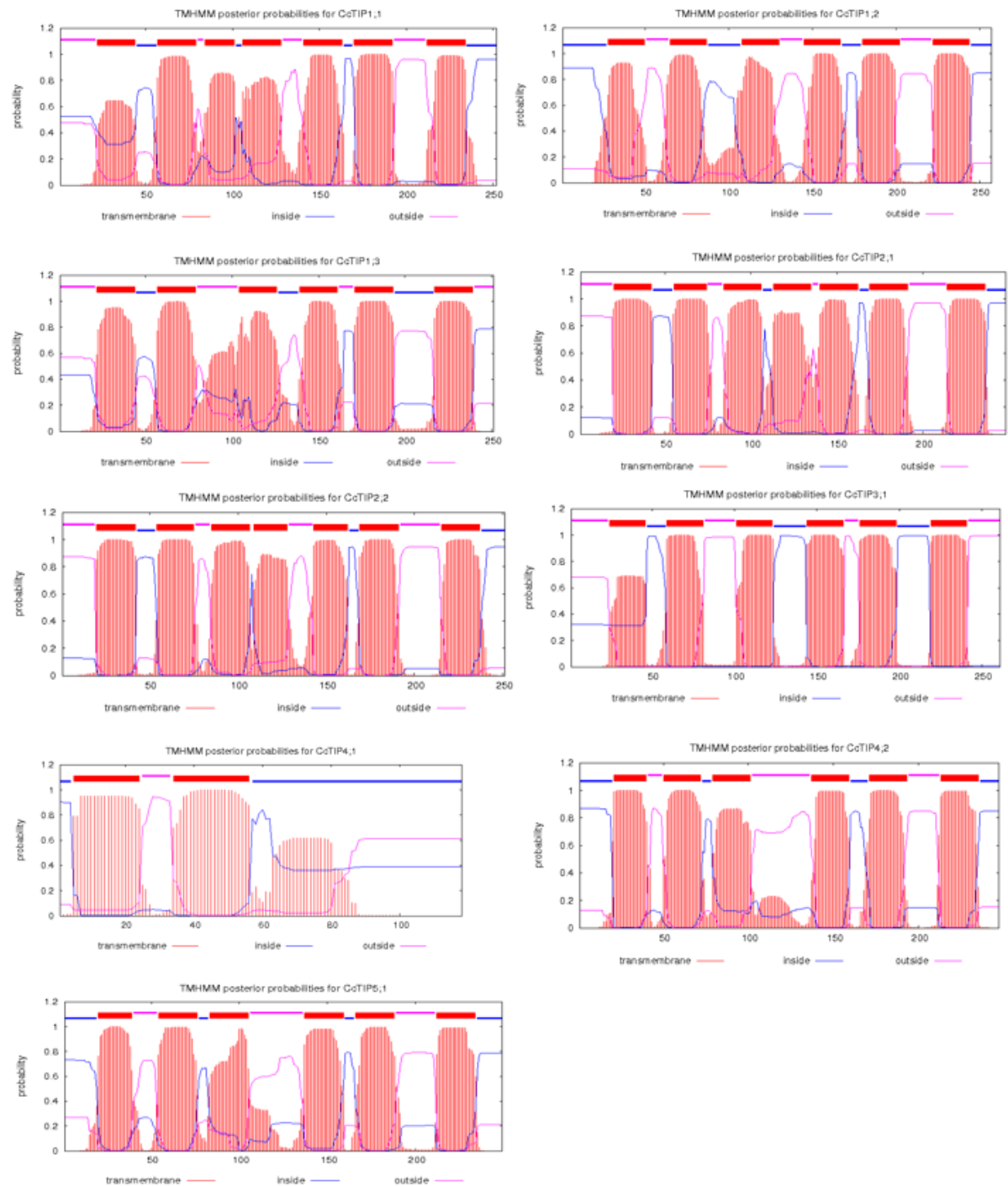
Figura- S3. Predição dos domínios transmembrana em sequências deduzidas de aminoácidos de CcAQP através do site *TMHMM Server v. 2.0*.

Subfamília PIP



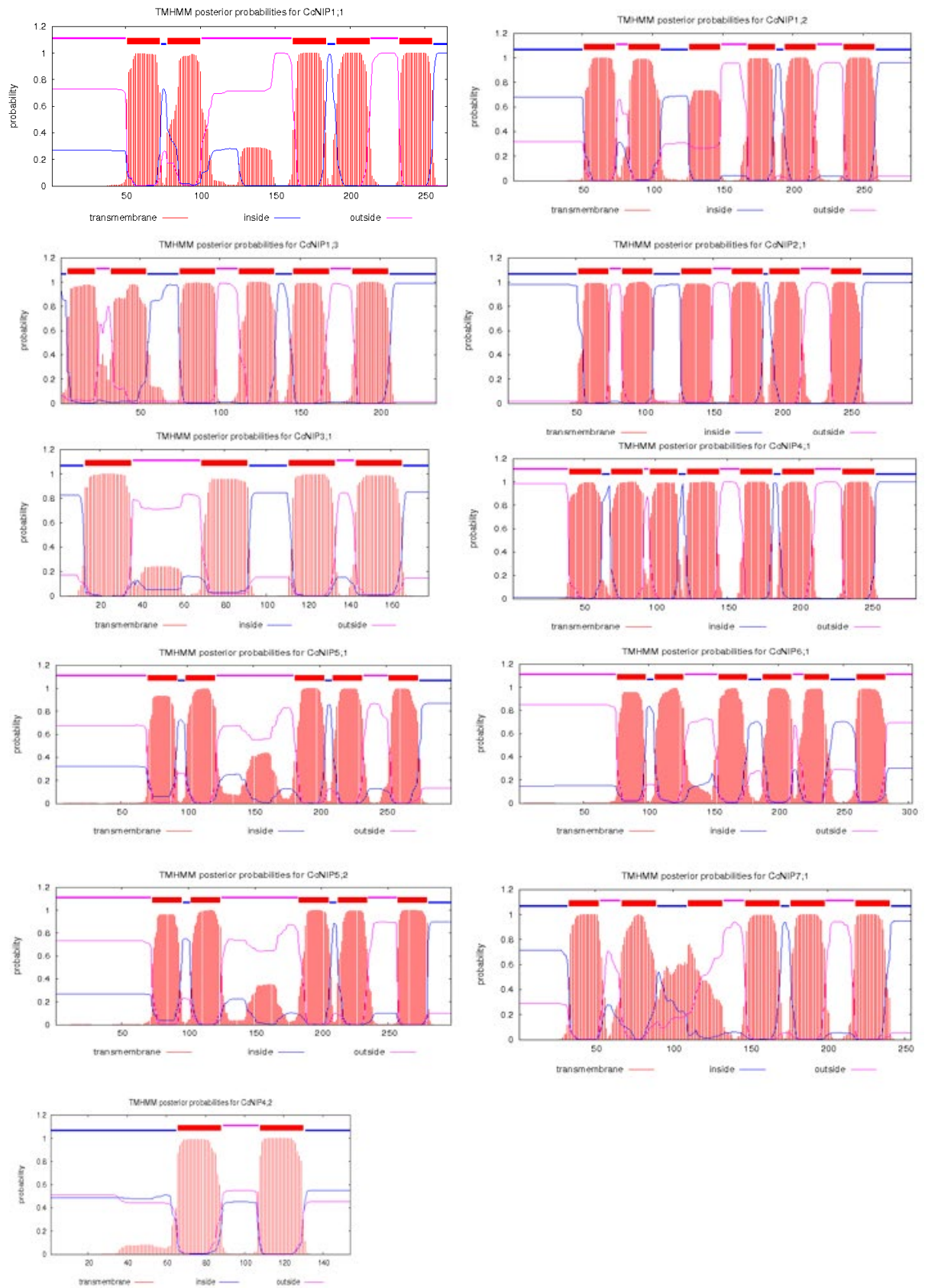
Continua...

Subfamilia TIP



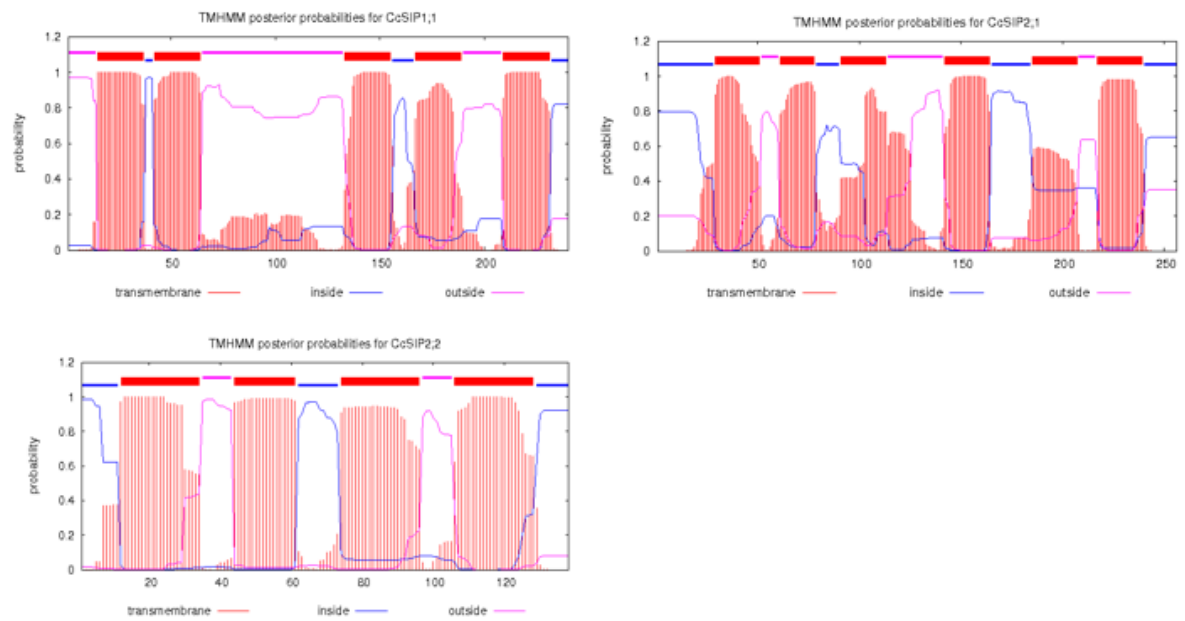
Continua...

Subfamilia NIP

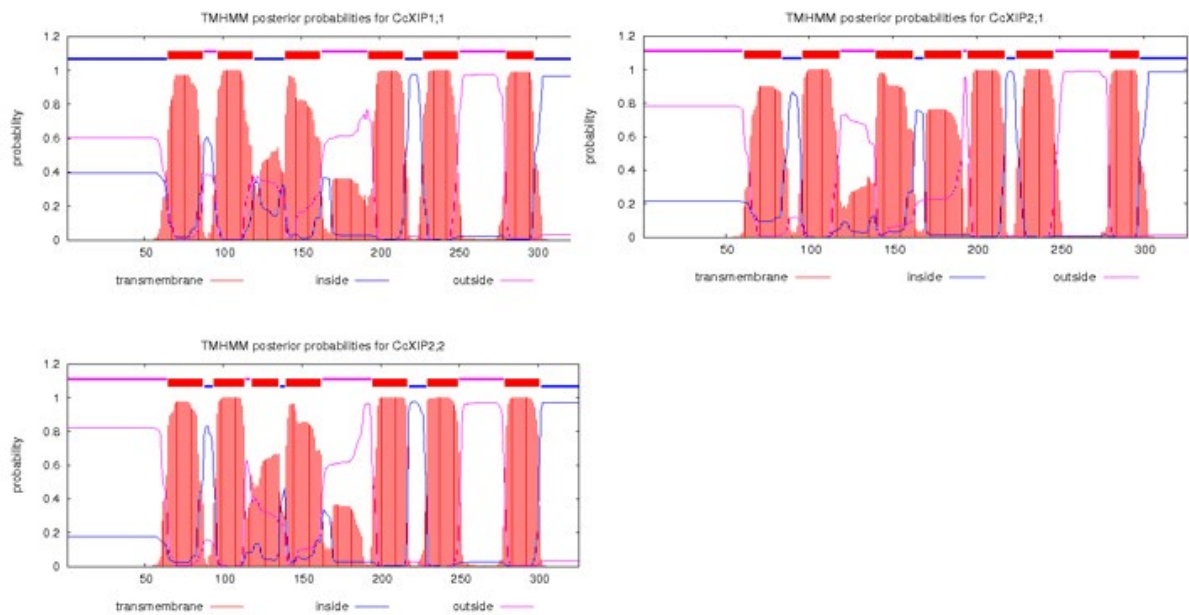


Continua...

Subfamilia SIP



Subfamilia XIP



[illegible]

Continua...

Subfamilia NIP

```

CcNIP1_1  -----MAEASGANGNHGVALDVKDEDLSHKDVKNAES-----KT
CcNIP1_2  -----MAEASEANGNHGVALDVKDEDLSYRDGKNSAS-----KT
CcNIP1_3  -----
CcNIP4_1  -----MPAGIDRVEEEEISRMEEGVESNKREC-----DY
CcNIP4_2  -----MSEKEIAALEEGNASG-----
CcNIP2_1  -----MENGQAKVNASSRKEEHGSVSTERQQISQMPLS-----SN
CcNIP5_1  MAQSEPGSPRGTA STPGSTPTPLFTAVRIDSMADDRKMP-ACKCFPVLAG-----PQ
CcNIP5_2  MAESDGGTPRASAPPTPG-TPTPLFSSVRVDSMSGDRKSMP-RCKCLPVEVPTWGA--PH
CcNIP6_1  -MSEDVPSAPSTPVTPTPGAPLFGGIKPERSGNRRSLSSCKCFTVEAWALEEGTLP
CcNIP7_1  -----MGTDAPKISGEPLQENQS-----
CcNIP3_1  -----

```

```

CcNIP1_1  ISDTSGFLFTVPFMQKLVAEIIAGTYYLIFAGVGSVVINAARDNIITFPGTSVVWGLAAMV
CcNIP1_2  VSDTSGCLLTVPFMQKLIAEIIAGTYYLIFAGCGAVVINAVRNNITITAPGISIVWGLAVMV
CcNIP1_3  -----MAELIGTYLLMFTGFTSIVVNLGMDNIIVTFPGVAIIWGLDVMV
CcNIP4_1  HKGFFSSKDVIIILQKIVAEEVGTGYFVIFAGCGSVVVKNIYG-NVTFPGVCVTWGLIVMV
CcNIP4_2  -----
CcNIP2_1  LGRLYGECCPPGFIRKVTAEIIATYMLVFVTCGSAALSASDEHKVSKLGASVAGGLIVTV
CcNIP5_1  SCITNFPAPDISLPRKVGAEFLGTFILIFGATAGPIVNQKYDGAETLIGNAACAGLAVLI
CcNIP5_2  TCLTDFPAPDVSLTRKLGAEFVGTFILIFGAAAGPIVNQKYNGAETLIGNAACAGLAVMI
CcNIP6_1  SVSCALPPPPVSLARKVGAEFIGTLILIFSGTATAIVNQKTQGSSETLIGLSASTGLAVMI
CcNIP7_1  --FRGFPRAMCPGLRKVLAEAVGTFILVFCICGIIASMELMKGRIGLMEYATTAAALTVVV
CcNIP3_1  -----MCGGSALVDRIITPLTIVGIGLIWGLVIMA

```

hélice 1 hélice 2

```

CcNIP1_1  VVYSVGHISGAHLNPAITIAFATCKRFPWKQVP-----AYISAQVLGSTLASGTLRL
CcNIP1_2  VVYSVGHISGAHLNPAVTIAFATCKRFPWKQVP-----AYISAQVLGSILASGTLRL
CcNIP1_3  MVYAVGHVSGAHLNPAVTIAFATCKKFAWKRP-----AYICAQLFASTLASGTVWL
CcNIP4_1  MVYTVGHVSGAHLNPAVTVAFAIFGGFPWREVP-----LYILAQLMGSTLASCTLAL
CcNIP4_2  -----F
CcNIP2_1  MIYAVGHISGAHMLNPAVTLFAA VRHFPWKQAP-----FYAGAQLTG AISAGFTLHV
CcNIP5_1  VILATGHISGAHLNPAVTIAFAALRHFPWAQVP-----AYIAAQVSGSICASFALKA
CcNIP5_2  VILATGHISGAHLNPSLTIAFAALRHFPWVQVP-----AYIAAQVSGSICASFALKG
CcNIP6_1  VILATGHISGAHLNPAVTIAFAALKHFPWKHVP-----AYIAAQVLASLAAAFILKG
CcNIP7_1  VVYAIGTISGAHLNPAITIAFATVGFPFPWSTVP-----LYLSAQLGGSVLATYAGKF
CcNIP3_1  LIYALGHVSGAHLNPAVTIAFGAASRLPLVQVEKFLSTCMATYVVSQ LIGSTLACLTLRV

```

loop B hélice 3

```

CcNIP1_1  IF---QDHFAGPLPHGS---DVQSLALEFIITFYLMFVICGVATDNRAIGELAGLAVGAT
CcNIP1_2  IFNGHQDHFSGTLPNGS---DVQSLVLEFIITFYLMFVISGVATDNRAIGELAGLAVGAT
CcNIP1_3  AFGEEHYQSVGTLPVGS---NIQSLLEFLITFYLMFAVSGVSTDERATPELAGIAVGAT
CcNIP4_1  MEDVASDDYFGTVPAGS---DGQSLAIEIIISFLLMFVISGVATDSRAIGNLAGITVGMT
CcNIP4_2  EYKSNEKNYFLWDEVQP---EMVELLFIFEQNASYILFCTNYFHPVSQIGELAGIATGMT
CcNIP2_1  IIR-PIEHVGTTSPSGT---TIQALVMEIVVTFSSMMFVTSAVATDTKAIGELAGIAVGSA
CcNIP5_1  VENPFMSGGVTVPSVG---NAQAFALEFIVTFILLFVVTVAVATDTRAVGELAGIAVGAT
CcNIP5_2  VEHPFMSGGVTVPSVG---NGQAFALEFVITFVLLFVVTVAVATDTRAVGELAGIAVGAT
CcNIP6_1  IEHPIMSGGVTVPSGE---YGQAFALEFIIGFNLMFVVTVAVATDTRAVGELAGIAVGAT
CcNIP7_1  VYGLQTDLVLT KPFHGR---SAAAFSVEFIATSMVVFLAAALANAPESVQNLSGFVVGGA
CcNIP3_1  IFHGQHDILPHTTQYRSSTSDLEAIAREFIITFILMFIICGAATDDRSNKALSGVATGVT

```

```

CcNIP1_1  VIVNVMIAGPISGSMSPAGRLGHAIVSSSNYKSLWVYILGPTAGAISGAWVYNIIRFTDK
CcNIP1_2  VLLNVMFAGPISGASMLNPARSLGHAIVSSSKYKSLWVYILGPTAGATGGAWVYNIIRFTDK
CcNIP1_3  VTVNSLLAGPISGASMLNPARSLGHAIVSNCTYGIWVYIVGPIAGALAGCWVYNIVRLKDK
CcNIP4_1  IMLNVFVAGPVS GASMLNPARSGHAIVKHVYKGLWVYIVGPIIGTILGAFVYNLIRYTDK
CcNIP4_2  ILLNVLVAGPVS GASMLNPARSIGHAIVMNEYKGLWIYTAGPLLGTIAGAFYTNLIRFTEK
CcNIP2_1  VCISILAGPISGSMNPARSIGHAIVASNYEYGIWVYVGPVCGTLLGAWSYNFIRVTDK
CcNIP5_1  VMLNILVAGPASGSMNPVRTLGHAIVAAGNYTSVWVYLVAPT LGALAGAGAYTMVKLQGE
CcNIP5_2  VMLNILVAGPSSGSMNPVRTLGHAIVAAGNYKSIWVYLLAPT LGALAGAAVYTFVKLQGD
CcNIP6_1  VMLNILIAGESTGASMLNPVRTLGHAIVANNFKGIWYLTAPVLGALAGAGTYSAVKLPDD
CcNIP7_1  IGLGVLTITGPVSGASMLNPARSLGHAIVSGDFAFLWIYLAAPT LGAIFGAILYRLIRLQGW
CcNIP3_1  MLFNSVNVICFFNWSIN-----ESWKKHWSNSI-----

```

loop E hélice 6

: * : : *

Continua...

P1

CcXIP2_1 INPIISFSAALVGLISMSRAIIYIIAQCGGVLGALALKAVVSSNIERTFSLGGCTLTVI
CcXIP2_2 INPIVSFSAGLIGLISMSRAIIYIIAQCGGVLGALALKAVVSSNIERTFSLAGCTLTVI
CcXIP1_1 INPIVSFSAALVGLISMSRAIIYIIAQCLGAVLGALALKAVVSGNIERTFSLGGCTLSIV
****:****:*****:*****:*****:*****:*****:*****:****:::
loop B hélice 3

CcXIP2_1 APGPQG-PVTIGLETAQAFWLEVFCSFVFLFASVWMAYDHRQRNSIGLIPVFSIIGIVLG
CcXIP2_2 APGPEG-PVTKGLEAAQAFWLEAFCSFVFLFASIWLFDHRQRNSIGMIRVFSIIGIVLG
CcXIP1_1 TPGPQGSPTIGLTAQAFWLEVFCSFVFLFASLWMAYDHRQKNSIGLVRVLAIVGTVLG
:***:* *** ** :*****:*****:***:****:****:***:***:****
hélice 4 hélice 5

H5 LE1 LE2 P2 P3 P4 P5

CcXIP2_1 LLVEFVSTTVTATKGYAGAGMNPARGFAAVVRGGHLWEEHWIFWAGPAVACFAFYVYTRI
CcXIP2_2 LLVEFVSTTVTATKGYAGAGMNPARGFAAVVRGGHLWEGHWIFWAGPAVACFAFYVYTLI
CcXIP1_1 LLVEFISTSVTAAGYGGAGINPARGFAAVVRGGHLWDGHWVFWAGPAVACFAFYVYTOI
****:***:***:***:***:*****:*****:***:*****:*****:***** *
loop E hélice 6

CcXIP2_1 IPNNHFHAEGFKHDFFGILRTISRLE
CcXIP2_2 IPDDHFHAEGYKHDFFGILRTISGLER
CcXIP1_1 IPNDHFHAQEFKHDFFGILRTVSGYNH
::**:*****:***:***:

REFERÊNCIAS

- ABRAMSON, J.; VARTANIAN, A. S. Watch water flow. **Science**, v.340, n.6138, p.1294-1295, 2013.
- AFZAL, Z. *et al.* The roles of aquaporins in plant stress responses. **Journal of Developmental Biology**, v.4, n.1, p.9, 2016.
- AGRE, P.; SASAKI, S.; CHRISPEELS, M. J. Aquaporins: a family of water channel proteins. **American Journal of Physiology-Renal Physiology**, v.265, n.3, p.F461, 1993.
- ALEXANDERSSON, E. *et al.* Whole gene family expression and drought stress regulation of aquaporins. **Plant Molecular Biology**, v.59, p.469-484, 2005.
- ALTSCHUL, S. F. *et al.* Basic local alignment search tool. **Journal of Molecular Biology**, v.215, n.3, p.403-410, 1990.
- ANDERBERG, H. I.; KJELLBOM, P.; JOHANSON, U. Annotation of *Selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants. **Frontier in Plant Science**, v.3, p.33, 2012.
- ANDERBERG, H.I.; DANIELSON, J. ÅH.; JOHANSON, U. Algal MIPs, high diversity and conserved motifs. **BMC Evolutionary Biology**, v.11, p.110, abr. 2011.
- BABICKI, S. *et al.* Heatmapper: web-enabled heat mapping for all. **Nucleic Acids Research**, v.44, n.W1, p.W147-W153, 2016.
- BAILEY, T. *et al.* Fitting a mixture model by expectation-maximization to discover motifs in biopolymers. *In*: PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY. **Proceedings [...]**. Menlo Park, California: CAAAI Press, 1994. p.28-36.
- BAILEY, T. L. *et al.* MEME: discovering and analyzing DNA and protein sequence motifs. **Nucleic Acids Research**, v.34, v.suppl-2, p.W369–W373, jul. 2006.
- 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, n.1, p.27, 2007.
- BEITZ, E. *et al.* Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons. **Proceedings of the National Academy of Sciences**, v.103, n.2, p.269-274, 2006.
- BIENERT, G. P. *et al.* Maize plasma membrane aquaporin ZmPIP2; 5, but not ZmPIP1; 2, facilitates transmembrane diffusion of hydrogen peroxide. **Biochimica et Biophysica Acta (BBA)-Biomembranes**, v.1838, n.1, p.216-222, 2014.
- BIENERT, G. P. *et al.* Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. **The Plant Journal**, v.66, n.2, p.306-317, 2011.

BIENERT, M. D.; BIENERT, G. P. Plant aquaporins and metalloids. *In*: **Plant Aquaporins**. Cham: Springer, 2017. p.297-332.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Café no Brasil**. 2017. Disponível em: <http://www.agricultura.gov.br/assuntos/politica-agricola/cafe/cafeicultura-brasileira>. Acesso em: 20 fev. 2018.

CARVALHO, M. C. C. G.; SILVA, D. C. G. Next generation DNA sequencing and its applications in plant genomics. **Ciência Rural**, v.40, n.3, p.735-744, mar. 2010.

CHANG, S.; PURYEAR, J.; CAIRNEY, J. A simple and efficient method for isolating RNA from pine trees. **Plant Molecular Biology Reporter** v.11, n.2, p.113-116, jun. 1993.

CHARRIER, A.; BERTHAUD, J. Botanical classification of coffee. *In*: CLIFFORD, M. N.; WILSON, K. C. (ed.). **Coffee: botany, biochemistry and production of beans and beverage**. London: Croom Helm, 1985. p.13-47.

CHAUMONT, F. *et al.* Aquaporins constitute a large and highly divergent protein family in maize. **Plant Physiology**, v.125, n.3, p.1206-1215, 2001.

CHOU, K-C.; SHEN, H-B. Plant-mPLOC: a top-down strategy to augment the power for predicting plant protein subcellular localization. **Plos One**, v.5, n.6, p.e11335, 2010.

CONAB. **Acompanhamento da safra brasileira de café**. v.4 - Safra 2017,n.4 - quarto levantamento, dez. de 2017. Disponível em: http://www.sapc.embrapa.br/arquivos/consorcio/levantamento/conab_safra2017_n4.pdf. Acesso em: 21 ago. 2018.

DAMATTA, F. M.; RAMALHO, J. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal of Plant Physiology**, v.18, n.1, p.55-81, 2006.

DANIELSON, J. Å.H.; JOHANSON, U. Unexpected complexity of the Aquaporin gene family in the moss *Physcomitrella patens*. **BMC Plant Biology**, v.8, p.45, 2008.

DAVIS, A. P. *et al.* Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. **Botanical Journal of the Linnean Society**, v.167, n.4, p.357-377, 2011.

DE MAROTHY, M. T.; ELOFSSON, A. Marginally hydrophobic transmembrane α -helices shaping membrane protein folding. **Protein Science**, v.24, n.7, p.1057-1074, 2015.

DENOEUDE, F. *et al.* The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, v.345, n.6201, p.1181-1184, set. 2014.

DEOKAR, A. A.; TAR'AN, B. Genome-wide analysis of the aquaporin gene family in chickpea (*Cicer arietinum* L.). **Frontiers in Plant Science**, v.7, p.1802, 2016.

DEREEPER, A. *et al.* SNIPlay3: a web-based application for exploration and large scale analyses of genomic variations. **Nucleic Acids Research**, v.43, n.W1, p.W295–W300, jul. 2015.

DESHMUKH, R. *et al.* Intron gain, a dominant evolutionary process supporting high levels of gene expression in rice. **Journal of Plant Biochemistry and Biotechnology**, v.25, n.2, p.142-146, 2016.

DESHMUKH, R.; BÉLANGER, R. R. Molecular evolution of aquaporins and silicon influx in plants. **Functional Ecology**, v.30, n.8, p.1277-1285, 2016.

DIEHN, T. A. *et al.* Genome-wide identification of aquaporin encoding genes in *Brassica oleracea* and their phylogenetic sequence comparison to Brassica crops and *Arabidopsis*. **Frontiers in Plant Science**, v.6, p.166, 2015.

DYNOWSKI, M. *et al.* Molecular determinants of ammonia and urea conductance in plant aquaporin homologs. **FEBS Letters**, v.582, n.16, p.2458-2462, 2008.

EMRICH, S. J. *et al.* Gene discovery and annotation using LCM454 transcriptome sequencing. **Genome Research**, v.17, n.1, p.69-73, 2007.

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, n.4, p.263, 2007.

GALMES, J. *et al.* Aquaporin expression in response to different water stress intensities and recovery in Richter-110 (*Vitis* sp.): relationship with ecophysiological status. **Planta**, v.226, n.3, p.671-681, ago. 2007.

GASTEIGER, E. *et al.* Protein Identification and analysis tools on the ExPASy Server. In: WALKER, J. M. (ed): **The proteomics protocols handbook, humana**. 2005. p.571-607.

GOMES, D. *et al.* Aquaporins are multifunctional water and solute transporters highly divergent in living organisms. **Biochimica et Biophysica Acta (BBA) – Biomembrane**, v.1788, n.6, p.1213-1228, jun. 2009.

GOODSTEIN, D. M. *et al.* Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, v.40, n.D1, p.D1178-D1186, 2011.

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, n.1, p.134, 2009.

HOOIJMAIJERS, C. *et al.* Hydrogen peroxide permeability of plasma membrane aquaporins of *Arabidopsis thaliana*. **Journal of Plant Research**, v.125, n.1, p.147-153, 2012.

HOVE, R. M.; BHAVE, M. Plant aquaporins with non-aqua functions: deciphering the signature sequences. **Plant Molecular Biology**, v.75, n.4-5, p.413-430, 2011.

HU, B. *et al.* GSDS 2.0: An upgraded gene feature visualization server. **Bioinformatics**, v.31, n.8, p.1296-1297, Abr. 2015.

HU, H.; XIONG, L. Genetic engineering and breeding of drought-resistant crops. **Annual Review of Plant Biology**, v.65, p.715-741, 2014.

ISHIKAWA, F. *et al.* Novel type aquaporin SIPs are mainly localized to the ER membrane and show cell-specific expression in *Arabidopsis thaliana*. **FEBS Letters**, v.579, n.25, p.5814-5820, 2005.

JAHN, T. P. *et al.* Aquaporin homologues in plants and mammals transport ammonia. **FEBS Letters**, v.574, n.1-3, p.31-36, 2004.

JAVOT, H. *et al.* Role of a single aquaporin isoform in root water uptake. **Plant Cell**, v.15, p.509-522, fev. 2003.

JOHANSON, U. *et al.* The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants. **Plant Physiology**, v.126, n.4, p.1358-1369, 2001.

JOHNSON, K. D.; HERMAN, E. M.; CHRISPEELS, M. J. An abundant, highly conserved tonoplast protein in seeds. **Plant Physiology**, v.91, p.1006-1013, Nov. 1989.

KALDENHOFF, R. *et al.* Aquaporins and plant water balance. **Plant Cell Environment**, v.31, n.5, p.658-66, mai. 2008.

KAPILAN, R.; VAZIRI, M.; ZWIAZEK, J. J. Regulation of aquaporins in plants under stress. **Biological Research**, v.51, n.1, p.4, 2018.

KELLEY, L. A.; STERNBERG, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. **Nature Protocols**, v.4, p.363-371, fev. 2009.

KHAN, K. *et al.* Heterologous expression of two *Jatropha* aquaporins imparts drought and salt tolerance and improves seed viability in transgenic *Arabidopsis thaliana*. **Plos One**, v.10, n.6, p.e0128866, jun. 2015.

KONG, W. *et al.* Genome-wide identification and characterization of aquaporin gene family in *Beta vulgaris*. **PeerJ**, v.5, p.e3747, set. 2017.

KROGH, A.; RAPACKI, K. **Servidor TMHMM, v. 2.0**. Cbs. dtu. Dk. 2016.

KUMAR, S. *et al.* MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular Biology and Evolution**, v.35, n.6, p.1547-1549, 2018.

LASHERMES, P. *et al.* Molecular characterisation and origin of the *Coffea arabica* L. genome. **Molecular and General Genetics MGG**, v.261, n.2, p.259-266, 1999.

LASHERMES, P. *et al.* Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interespecific hybrid *C. arabica* x *C. canephora*. **The Journal of Heredity**, v.91, n.1, p.81-85, fev. 2000.

- LEROY, T. *et al.* Developing core collections to optimize the management and the exploitation of diversity of the coffee. *Coffea canephora*. **Genetica**, v.142, p.185-199, 2014.
- LI, G.; SANTONI, V.; MAUREL, C. Plant aquaporins: roles in plant physiology. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v.1840, n.5, p.1574-1582, 2014.
- LI, T. *et al.* *Arabidopsis thaliana* NIP7; 1: an anther-specific boric acid transporter of the aquaporin superfamily regulated by an unusual tyrosine in helix 2 of the transport pore. **Biochemistry**, v.50, n.31, p.6633-6641, 2011.
- LIMA, A. L. S. *et al.* Photochemical responses and oxidative stress in two clones of *Coffea canephora* under water deficit conditions. **Environmental and Experimental Botany**, v.47, n.3, p.239-247, 2002.
- LIMA, E. A. *et al.* The multi-resistant reaction of drought-tolerant coffee 'Conilon clone 14' to *Meloidogyne* spp. and late hypersensitive-like response in *Coffea canephora*. **Phytopathology**, v.105, n.6, p.805-814, 2015.
- LIN, C. *et al.* Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. **Theoretical and Applied Genetics**, v.112, n.1, p.114-130, 2005.
- LIU, Q. *et al.* Divergence in function and expression of the NOD26-like intrinsic proteins in plants. **BMC Genomics**, v.10, n.1, p.313, 2009.
- LIU, Q.; ZHU, Z. Functional divergence of the NIP III subgroup proteins involved altered selective constraints and positive selection. **BMC Plant Biology**, v.10, n.1, p.256, 2010.
- LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. **Methods**, v.25, n.4, p.402-408, 2001.
- LOQUÉ, D. *et al.* Tonoplast intrinsic proteins AtTIP2; 1 and AtTIP2; 3 facilitate NH₃ transport into the vacuole. **Plant Physiology**, v.137, n.2, p.671-680, 2005.
- MA, J. F. *et al.* A silicon transporter in rice. **Nature**, v.440, n.7084, p.688, 2006.
- MAESHIMA, M. Tonoplast transporters: organization and function. **Annual Review of Plant Biology**, v.52, n.1, p.469-497, 2001.
- MAHAJAN, S.; TUTEJA, N. Cold, salinity and drought stresses: an overview. **Biochemistry Biophysics**, v.444, p.139-158, dez. 2006.
- MANAVALAN, L. P.; NGUYEN, H. T. Drought tolerance in crops: Physiology to genomics. **Plant Stress Physiol**, v.1, n.23, p.10.1079, 2012.
- MAO, Z.; SUN, W. *Arabidopsis* seed-specific vacuolar aquaporins are involved in maintaining seed longevity under the control of ABSCISIC ACID INSENSITIVE 3. **Journal of Experimental Botany**, v.66, n.15, p.4781-4794, ago. 2015.

MARRACCINI, P. *et al.* Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. **Journal of Experimental Botany**, v.63, n.11, p.4191-4212, 2012.

MARRACCINI, P. *et al.* RBCS1 expression in coffee: coffee orthologs, *Coffea arabica* homeologs, and expression variability between genotypes and under drought stress. **BMC Plant Biology**, v.11, n.1, p.85, 2011.

MARRACCINI, P. *et al.* Differentially expressed genes and proteins upon drought acclimation in tolerant and sensitive genotypes of *Coffea canephora*. **Journal of Experimental Botany**, v.63, n.11, p.4191-4212, 2012.

MARTINS, C.P.S. *et al.* Genome-wide characterization and expression analysis of major intrinsic proteins during abiotic and biotic stresses in sweet orange (*Citrus sinensis* L. Osb.). **Plos One**, v.10, p.e0138786, 2015.

MARTRE, P. *et al.* Plasma membrane aquaporins play a significant role during recovery from water deficit. **Plant Physiology**, v.130, n.4, p.2101-2110, dez. 2002.

MATIELLO, J. B. **Café conilon: como plantar, tratar, colher, preparar e vender**. Rio de Janeiro: MM Produções Gráficas, 1998.

MAUREL, C. *et al.* Aquaporins in plants. **Physiological Reviews**, v.95, n.4, p.1321-1358, Set. 2015.

MAUREL, C. *et al.* Molecular physiology of aquaporins in plants. *In: International review of cytology*. Academic Press, 2002. p.105-148.

MAUREL, C. *et al.* The vacuolar membrane protein gamma-TIP creates water specific channels in *Xenopus* oocytes. **The EMBO Journal**, v.12, n.6, p.2241-2247, 1993.

MAUREL, C.; VERDOUCQ L.; LUU, D-T.; SANTONI, V. Plant aquaporins: Membrane channels with multiple integrated functions. **Annual Review of Plant Biology**, v.59, p.595-624, 2008.

MENDES, A. N. G.; GUIMARÃES, R. J. **Cafeicultura empresarial (produtividade e qualidade - genética e melhoramento do cafeeiro)**. Lavras: UFLA/FAEPE, 1996. 99 p.

MENEGASSI, A. **Identificação, caracterização estrutural e funcional de aquaporinas de soja (*Glycine max*) e seu envolvimento no metabolismo de ureia**. 2015. 100f. Tese (Doutorado) - Universidade Federal do Rio Grande do Sul. Centro de Biotecnologia do Estado do Rio Grande do Sul, 2015.

MITANI, N.; YAMAJI, N.; MA, J. F. Characterization of substrate specificity of a rice silicon transporter, Lsi1. **Pflügers Archiv-European Journal of Physiology**, v.456, n.4, p.679-686, 2008.

MITANI-UENO, N. *et al.* The aromatic/arginine selectivity filter of NIP aquaporins plays a critical role in substrate selectivity for silicon, boron, and arsenic. **Journal of Experimental Botany**, v.62, n.12, p.4391-4398, 2011.

- MONTAGNON, C. *et al.* Amélioration génétique du caféier *Coffea canephora* Pierre: connaissances acquises, stratégies et perspectives. **Cahiers Agricultures**, v.21, p.143-153, 2012.
- MÜLLER, C.; HOFMANN, T. Screening of raw coffee for thiol binding site precursors using “in bean” model roasting experiments. **Journal of Agricultural and Food Chemistry**, v.53, n.7, p.2623-2629, 2005.
- MURATA, K. *et al.* Structural determinants of water permeation through aquaporin-1. **Nature**, v.407, n.6804, p.599, 2000.
- NOIROT, M. *et al.* Genome size variations in diploid African *Coffea* Species. **Annals of Botany**, v.92, n.5, p.709-714. nov. 2003.
- ORGANIZAÇÃO INTERNACIONAL DO CAFÉ. **Embrapa**: mercado de café encerra 2017/18 em superávit. 2019. Disponível em: <https://icocoffeeorg.tumblr.com/post/178884744500/coffee-market-ends-201718-in-surplus>. Acesso em: mai. 2019.
- PARK, W. *et al.* Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). **BMC Plant Biology**, v.10, n.1, p.142, 2010.
- PEREZ-MARTIN, A. *et al.* Regulation of photosynthesis and stomatal and mesophyll conductance under water stress and recovery in olive trees: correlation with gene expression of carbonic anhydrase and aquaporins. **Journal of Experimental Botany**, v.65, n.12, p.3143-3156, jul. 2014.
- POMMERENIG, B.; DIEHN, T. A.; BIENERT, G. P. Metalloido-porins: essentiality of nodulin 26-like intrinsic proteins in metalloid transport. **Plant Science**, v.238, p.212-227, 2015.
- POSTAIRE, O. *et al.* A PIP1 aquaporin contributes to hydrostatic pressure-induced water transport in both the root and rosette of *Arabidopsis*. **Plant Physiology**, v.152, p.1418-1430, 2010.
- QIAN, W. *et al.* Genome-wide identification and expression analysis of aquaporins in salt cress (*Eutrema salsugineum*). **PeerJ Preprints**, 2019.
- RAMAKERS, C. *et al.* Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, v.339, n.1, p.62- 66, mar. 2003.
- REDDY, P. S. *et al.* Genome-wide identification and characterization of the aquaporin gene family in *Sorghum bicolor* (L.) **Plant Gene**, v.1, p.18-28, 2015.
- REGON, P. *et al.* Genome-wide comparative analysis of tonoplast intrinsic protein (TIP) genes in plants. **Functional & Integrative Genomics**, v.14, n.4, p.617-629, 2014.
- REIZER, J.; REIZER, A.; SAIER, M. H. The MIP family of integral membrane channel proteins: sequence comparisons, evolutionary relationships, reconstructed pathway of evolution, and proposed functional differentiation of the two repeated halves of the

proteins. **Critical Reviews in Biochemistry and Molecular Biology**, v.28, n.3, p.235-257, 2008.

REUSCHER, S. *et al.* Genome-wide identification and expression analysis of aquaporins in tomato. **Plos One**, v.8, n.11, p.e79052, 2013.

SAKURAI, J. *et al.* Identification of 33 rice aquaporin genes and analysis of their expression and function. **Plant Cell Physiology**, v.46, n.9, p.1568-1577, set. 2005.

SANTOS, A. V. *et al.* Reaction of *Coffea canephora* clones to the root knot nematode, *Meloidogyne incognita*. **African Journal of Agricultural Research**, v.12, n.11, p.916-922, 2017.

SANTOS, A. B.; MAZZAFERA, P. Aquaporins and the control of the water status in coffee plants. **Theoretical and Experimental Plant Physiology**, v.25 n.2, p.79-93, 2013.

SANTOS, T. B. *et al.* Galactinol synthase transcriptional profile in two genotypes of *Coffea canephora* with contrasting tolerance to drought. **Genetics and Molecular Biology**, v.38, n.2, p.182-190, 2015.

SHAAR-MOSHE, L.; BLUMWALD, E.; PELEG, Z. Unique physiological and transcriptional shifts under combinations of salinity, drought and heat. **Plant Physiology**, p.17.00030, mar. 2017.

SHEKOOFA, A.; SINCLAIR, T. Aquaporin activity to improve crop drought tolerance. **Cells**, v.7, n.9, p.123, 2018.

SHELDEN, M. C. *et al.* Identification and functional characterisation of aquaporins in the grapevine, *Vitis vinifera*. **Functional Plant Biology**, v.36, n.12, p.1065-1078, 2009.

SHIVARAJ, S. M. *et al.* Identification and characterization of aquaporin genes in *Arachis duranensis* and *Arachis ipaensis* genomes, the diploid progenitors of peanut. **BMC Genomics**, v.20, n.1, p 222, 2019.

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

SOUZA, F. F. *et al.* **Aspectos gerais da biologia e da diversidade genética de *Coffea canephora***. Embrapa Café - ALICE, 2015.

SOUZA, F. F. *et al.* **Características das principais variedades de café cultivadas em Rondônia**. Porto Velho: Embrapa Rondônia, 2004. 21 p.

STURN, A. *et al.* Genesis: cluster analysis of microarray data. **Bioinformatics**, v.18, n.1, p.207-8, jan.2002.

SUN, H. *et al.* Genome-wide identification and characterization of aquaporin gene family in moso bamboo (*Phyllostachys edulis*). **Molecular Biology Reports**, v.43, n.5, p.437-450, mai. 2016.

SUN, H. *et al.* Whole-genome and expression analyses of bamboo aquaporin genes reveal their functions involved in maintaining diurnal water balance in bamboo shoots. **Cells**, v.7, n.11, p.195, 2018.

TAKANO, J. *et al.* The *Arabidopsis* major intrinsic protein NIP5; 1 is essential for efficient boron uptake and plant development under boron limitation. **The Plant Cell**, v.18, n.6, p.1498-1509, 2006.

TANAKA, M. *et al.* NIP6; 1 is a boric acid channel for preferential transport of boron to growing shoot tissues in *Arabidopsis*. **The Plant Cell**, v.20, n.10, p.2860-2875, 2008.

TAO, P. *et al.* Genome-wide identification and characterization of aquaporin genes (AQPs) in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). **Molecular genetics and genomics**, v.289, n.6, p.1131-1145, 2014.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. *ClustalW*—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids**, v.22, n.22, p.4673-80, nov. 1994.

TÖRNROTH-HORSEFIELD, S. *et al.* Structural mechanism of plant aquaporin gating. **Nature**, v.439, n.7077, p.688, 2006.

VENKATESH, J.; YU, J. W.; PARK, S. W. Genome-wide analysis and expression profiling of the *Solanum tuberosum* aquaporins. **Plant Physiology and Biochemistry**, v.73, p.392-404, 2013.

VIEIRA, L. G. E. *et al.* Brazilian coffee genome project: an EST-based genomic resource. **Brazilian Journal of Plant Physiology**, v.18, n.1, p.95-108, 2006.

VOORRIPS, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. **Journal of Heredity**, v.93, n.1, p.77–78, Jan. 2002.

WALLACE, I. S.; ROBERTS, D. M. Distinct transport selectivity of two structural subclasses of the nodulin-like intrinsic protein family of plant aquaglyceroporin channels. **Biochemistry**, v.44, n.51, p.16826-16834, 2005.

WALLACE, I. S.; ROBERTS, D. I. M. Homology modeling of representative subfamilies of *Arabidopsis* major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. **Plant Physiology**, v.135, n.2, p.1059-1068, 2004.

WEI, Q. *et al.* Genome-wide identification and characterization of sweet orange (*Citrus sinensis*) aquaporin genes and their expression in two citrus cultivars differing in drought tolerance. **Tree Genetics & Genomes**, v.15, n.2, p.17, 2019.

WUDICK, M. M.; LUU, D. T.; MAUREL, C. A look inside: localization patterns and functions of intracellular plant aquaporins. **New Phytologist**, v.184, n.2, p.289-302, Out. 2009.

YU, Q. *et al.* Micro-collinearity and genome evolution in the vicinity of an ethylene receptor gene of cultivated diploid and allotetraploid coffee species (*Coffea*). **The Plant Journal**, v.67, n.2, p.305-317, 2011.

ZOU, Z. *et al.* Genome-wide identification of rubber tree (*Hevea brasiliensis* Muell. Arg.) aquaporin genes and their response to ethephon stimulation in the laticifer, a rubber-producing tissue. **BMC Genomics**, v.16, n.1, p.1, dez. 2015a.

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

ZOU, Z.; YANG, J. Genome-wide comparison reveals divergence of cassava and rubber aquaporin family genes after the recent whole-genome duplication. **BMC Genomics**, v 20, n.1, p.380, 2019.

ZUPIN, M. *et al.* Drought-induced expression of aquaporin genes in leaves of two common bean cultivars differing in tolerance to drought stress. **Journal of Plant Research**, v.130, n.4, p.735-745, 2017.