



UNIVERSIDADE ESTADUAL DE SANTA CRUZ
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA
MOLECULAR – PPGGBM

DALMA BRITO SANTOS

CARACTERIZAÇÃO MOLECULAR DA FAMÍLIA DE FATORES DE
TRANSCRIÇÃO bZIP EM
Citrus sinensis L.

ILHÉUS – BAHIA - BRASIL
FEVEREIRO, 2019

DALMA BRITO SANTOS

**CARACTERIZAÇÃO MOLECULAR DA FAMÍLIA DE FATORES DE
TRANSCRIÇÃO bZIP EM
Citrus sinensis L.**

Dissertação apresentada à Universidade Estadual de Santa Cruz (UESC) para obtenção do título de Mestra em Genética e Biologia Molecular.

Linha de pesquisa: Genética e Melhoramento Vegetal

Orientador: Prof. Dr. Márcio Gilberto Cardoso Costa

Co-orientadora: Dra. Luciana Rodriguez Camillo

**ILHÉUS – BAHIA - BRASIL
FEVEREIRO, 2019**

Dalma Brito Santos

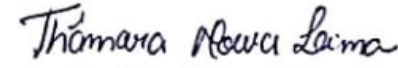
**CARACTERIZAÇÃO MOLECULAR DA FAMÍLIA DE FATORES DE TRANSCRIÇÃO
bZIP (BASIC LEUCINE ZIPPER) EM *Citrus sinensis***

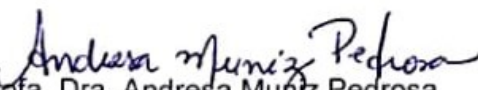
Dissertação apresentada à
Universidade Estadual de Santa Cruz,
como parte das exigências para
obtenção do título de Mestre em
Genética e Biologia Molecular.

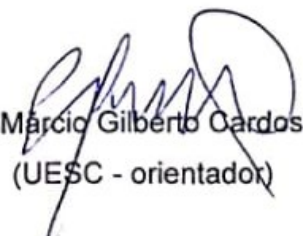
Área de Concentração: Genética e
Biologia Molecular.

APROVADA: 25 de fevereiro de 2019.


Profa Dra. Cláudia Fortes-Ferreira
(EMBRAPA)


Profa. Dra. Thâmara Moura Lima
(UESC)


Profa. Dra. Andresa Muniz Pedrosa
(UESC)


Prof. Dr. Marcio Gilberto Cardoso Costa
(UESC - orientador)

**Todos os lugares que toquei
Todo ser que amei
Todas as coisas estão em mim**

**Toda vez que a dor me chamou
Toda vez que o sorriso voltou
Eu chorei, sim e sorri**

**Tudo me tornou maior
Tudo me diminuiu para melhor.**

Crof

Dedico:

A todas as pessoas que passaram pela minha vida e de alguma forma deixaram um ensinamento que serviu para me colocar no lugar que hoje ocupo.

À minha família e amigos por não me deixarem fraquejar, sempre ofertando amor.

Ao meu universo musical e literário pelo acolhimento nos piores momentos.

À todas as grandes cientistas (mulheres), por me fazerem entender que ‘lugar de mulher é onde ela quiser’, inclusive fazendo ciência.

A mim, pela (auto) compreensão, afeto e respeito, afinal, “minha força de sonhar sou eu também”.

Gratidão!

Agradecimentos

À Universidade Estadual de Santa Cruz (UESC) por conceder a infraestrutura para o desenvolvimento da pesquisa.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular (PPGGBM) pela oportunidade de cursar o Mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudo.

À Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA – Mandioca e Fruticultura) pela parceria durante a execução do trabalho.

Ao Dr. Márcio G. C. Costa e a Dra. Luciana Rodrigues Camillo pela orientação/ensinamentos, e dedicação ao longo do mestrado.

Ao Laboratório de Cultura de Tecidos, principalmente a Gláucia Carvalho pela parceria e amizade ao longo dos experimentos, a Grazi Alcântara, Dilson e Thainá pela amizade e ensinamentos; e ao Laboratório de Biologia Molecular pelo suporte durante a pesquisa, principalmente a Natasha pela disponibilidade em ajudar (e pelo carinho, rs!)

Ao Núcleo de Biologia Computacional e Gestão de Informações Biotecnológicas (NBCGIB) que foi durante esse tempo uma segunda casa e ambiente de compartilhamento do conhecimento.

Ao grupo BioPREA (*Biology – Powerful researches with expertise in analysis*) por serem a minha família de amigos: meu pai, Edson Mário de Andrade; meu segundo pai, (Raner) José; minhas irmãs: Ceslaine Barbosa e Letícia Maróstica; e meu irmão, amigo de 9 anos Válder Magalhães. Aprendemos e ensinamos muito juntos... Gratidão! Amo vocês!

A turma 2017.1 de Pós-Graduação em Genética e Biologia Molecular da UESC! Foi um prazer viver esse período com vocês!

A minha família: mãe (Oderiana Alves de Brito Santos), pai (Agnaldo Barreto dos Santos) e irmã (Fenícia Brito Santos): obrigada pelo apoio, dedicação, afeto, cuidado de sempre. Mãe, pai e irmã chegamos aqui juntos!

A minha família Ilheense: Jullis (Julliana Castro) e Jacques (Jacqueline Fernandes)... Obrigada pelo afeto, lealdade e por me ensinarem a acreditar em mim. Como foi lindo vivenciar nosso crescimento e amadurecimento pessoal! Gratidão! Namastê! Axé!

Aos amigos da vida: Thaís Aline Silva, Lorena Pimentel, Ana Cláudia Oliveira, Paulo Henrique Gomes, Cátia Dias e Andresa Ramos vocês estavam ao meu lado em momentos muito importantes, crescemos e aprendemos muito juntxs! Obrigada por tudo!

ÍNDICE

EXTRATO	i
ABSTRACT	iii
1. INTRODUÇÃO.....	1
2. OBJETIVOS.....	4
2.1. Objetivo Geral.....	4
2.2. Objetivos Específicos.....	4
3. REVISÃO DE LITERATURA	5
3.1. Citros – Aspectos Gerais	5
3.2. Estresse abiótico em plantas – aspectos moleculares	6
3.3. Fatores de transcrição em plantas e sua relação com o estresse abiótico	8
3.4. Família bZIP (<i>Basic Leucine Zipper</i>) de fatores de transcrição	9
4. MATERIAL E MÉTODOS	14
4.1. Identificação dos genes da família bZIP em <i>Citrus sinensis</i> L.	14
4.2. Análise da estrutura dos genes e promotores	14
4.3. Alinhamento e análises filogenéticas	14
4.4. Propriedades físico-químicas das proteínas, análise dos motivos conservados e rede de interação	15
4.5. Material vegetal e tratamentos de déficit hídrico	15
4.6. Extração de RNA e síntese de cDNA	15
4.7. Análise de expressão de genes CsbZIP	16
5. RESULTADOS E DISCUSSÃO.....	17
5.1. Identificação e classificação dos genes CsbZIP	18
5.2. Caracterização da estrutura das proteínas CsbZIP	24
5.3. Padrões de expressão dos genes CsbZIP em diferentes tecidos	30
5.4. Expressão gênica de dois CsbZIPs em plantas submetidas ao déficit hídrico	32
6. CONCLUSÕES.....	34
REFERÊNCIAS BIBLIOGRÁFICAS	35
APÊNDICES	40

EXTRATO

SANTOS, Dalma Brito, M.S., Universidade Estadual de Santa Cruz, Ilhéus, fevereiro de 2019.

Caracterização molecular da família de fatores de transcrição bZIP em *Citrus sinensis*

L. Orientador: Prof. Dr. Márcio Gilberto Cardoso Costa. Co-orientadora: Dra. Luciana Rodriguez Camillo.

Fatores de transcrição (FTs) regulam muitas vias de sinalização associadas ao crescimento e desenvolvimento vegetal, e a resposta a estresses bióticos e abióticos por interação com elementos *cis*-reguladores nas regiões promotoras de genes alvo, ativando ou reprimindo sua expressão. Os bZIPs (*Basic Leucine Zipper*) são considerados uma das principais famílias de FTs associados à resposta ao estresse abiótico em plantas, com membros descritos nas mais diversas espécies de plantas. No entanto, os bZIPs não foram caracterizados em espécies cítricas. Portanto, nosso objetivo foi caracterizar a família de FTs bZIP em *Citrus sinensis* (CsbZIP), por meio de análises *in silico* e moleculares. Nós identificamos 34 genes que codificam para FTs bZIPs em *Citrus sinensis*, distribuídos em seis diferentes cromossomos e divididos em 12 grupos baseados em sua relação filogenética com *AtbZIPs* de *Arabidopsis*. Análises de motivos conservados e estrutura gênica foram realizadas para apoiar esta classificação. As regiões promotoras dos genes foram analisadas em busca de *cis*-elementos responsivos ao estresse (ABRE, DRE / CRT, MYBs, LTRE) e vários ABREs (elementos responsivos ao ácido abscísico (ABA)) foram encontrados. A partir das sequências de aminoácidos deduzidas, foram preditos os parâmetros físico-químicos das proteínas (peso molecular, ponto isoeletrico, *GRAVY*, localização subcelular). Além disso, verificou-se a presença de sítios de fosforilação em resíduos de Serina (Ser/S) e/ou Treonina (Thr/T), de acordo com o que é esperado para proteínas que são fosforiladas por quinases do tipo SnRKs (serina/treonina-quinase). Para entender com quais proteínas as CsbZIP interagem, criou-se uma rede de interação através do banco de dados STRING usando as proteínas de *Arabidopsis*, ortólogas as proteínas de *C. sinensis*. Observamos interações de proteínas do grupo A com proteínas envolvidas na resposta ao estresse abiótico, como proteínas quinases (SnRKs, CDKs), receptores de ABA (PYR/PYL/RCAR) e fosfatases (PP2C). Selecionamos dois genes *CsbZIPs* (*Cs8g06020* e *Cs3g23480*) ortólogos para os genes *ABF1* e *ABF2* de *Arabidopsis* e avaliamos sua expressão em plantas submetidas a diferentes níveis de déficit hídrico (controle, déficit hídrico moderado e severo e reidratado). Não foram observadas diferenças

significativas nos níveis de expressão dos genes entre os diferentes tratamentos. Nossos resultados são apresentados como a primeira descrição da família de genes bZIP de FTs em *Citrus*, e se destaca como base para futuros estudos que abordem o aumento da produtividade de citros conferindo resistência ao estresse através da modulação da expressão de proteínas bZIP específicas.

Palavras-chave: reguladores da transcrição; zíper de leucina; estresse abiótico; laranja doce.

ABSTRACT

SANTOS, Dalma Brito, M.S., Universidade Estadual de Santa Cruz, Ilhéus, February 2019.

Molecular characterization of the bZIP family of transcription factors in *Citrus sinensis*

L. Advisor: Prof. Dr. Márcio Gilberto Cardoso Costa. Co-advisor: Dr. Luciana Rodriguez Camillo.

Transcription factors (TFs) regulate several signaling pathways associated with plant growth and development, and the response to biotic and abiotic stresses through interaction with *cis*-regulatory elements in the promoter regions of target genes, activating or repressing their expression. bZIPs (Basic Leucine Zipper) are considered one of the main families of TFs associated with response to abiotic stress in plants, with members described in the most diverse plant species. However, bZIPs have not been characterized in citrus species. Therefore, our goal was to characterize the family of bZIP TFs in *Citrus sinensis* (CsbZIP), using *in silico* and molecular analysis. We have identified 34 genes coding for bZIPs TFs in *Citrus sinensis*, distributed in six different chromosomes and divided into 12 groups based on their phylogenetic relationship with AtbZIPs from *Arabidopsis*. Analysis of conserved motifs and gene structure were performed to support this classification. The promoter regions of the genes were assessed for *cis*-elements responsive to stresses (ABRE, DRE/CRT, MYBs, LTRE) and several ABREs (*abscisic acid (ABA)- responsive elements*) were found. From the deduced amino acid sequences, the proteins physico-chemical parameters were predicted (molecular weight, isoelectric point, GRAVY, subcellular localization). In addition, we found the presence of phosphorylation sites in Serine (Ser/S) and/or Threonine (Thr/T) residues, in accordance with what is expected of proteins that are phosphorylated by SnRKs kinases (serine/threonine kinase). In order to understand what CsbZIP proteins interact with, we created an interaction network through the STRING database using the *Arabidopsis* orthologs proteins to the proteins of *C. sinensis*. We observed interactions of proteins from group A with proteins involved in the abiotic stress response with kinase proteins (SnRKs, CDKs), PYR/PYL/RCAR ABA receptors and PP2C phosphatases. We selected two CsbZIPs genes (Cs8g06020 e Cs3g23480) orthologs to the ABF1 and ABF2 from *Arabidopsis* and assessed their expression in plants subjected to different levels of water deficit (control, moderate and severe water deficit and rehydrated). No significant differences were observed in the expression levels of the genes among the different treatments. Our results are set out as the first description of bZIP gene family of TFs in *Citrus*, and stands out as a base for future

studies addressing the increasing of citrus productivity by conferring resistance to stress through the modulation of the expression of specific bZIP proteins.

Key-words: transcription regulators; leucine zipper; abiotic stress; sweet orange.

1. INTRODUÇÃO

O gênero *Citrus* pertence à família Rutaceae e as plantas cítricas são muito apreciadas por seus benefícios para a saúde. A cultura citrícola é uma das mais cultivadas em regiões tropicais e subtropicais do mundo, sendo que a laranja doce (*Citrus sinensis* L.) destaca-se como a principal fruta do gênero, representando cerca de 70% da produção e do consumo total de citros. Além da laranja doce, outras frutas cítricas como as tangerinas (*C. reticulata*), as limas (*C. aurantifolia*) e limões (*C. limonum*) são cultivadas e consumidas extensivamente (Sharma *et al.*, 2017). O Brasil é considerado o maior produtor mundial de laranja doce, com produção estimada em 17 milhões de toneladas por ano (IBGE, 2018).

No entanto, o clima afeta a produtividade e a longevidade das plantas cítricas, pois estas desenvolvem-se melhor em regiões de clima ameno. Contudo, as plantas precisam sobreviver no seu ambiente natural, enfrentando condições desfavoráveis constantemente, como seca, salinidade elevada, temperaturas extremas, etc. Considerando a importância da citricultura brasileira, e diversos problemas relacionados ao clima que afetam fortemente a cultura, o déficit hídrico é uma das principais adversidades enfrentadas pelos citricultores, sendo importante salientar que as condições de estresse biótico/abiótico limitam o desenvolvimento e a produtividade das culturas em geral (Azevêdo, 2003; Hu *et al.*, 2016a).

Para enfrentar as adversidades ambientais, as plantas desenvolveram ao longo da sua evolução estratégias adaptativas integradas nos níveis molecular, celular e fisiológico, que favoreceram o seu crescimento, desenvolvimento, reprodução e produtividade (Baneerjee e Roychoudhury, 2017). Uma dessas estratégias, em nível molecular, ocorre por meio da expressão de genes induzidos por estresse, que codificam proteínas regulatórias, tais como os fatores de transcrição (FTs) (Sornaraj *et al.*, 2016).

Os fatores de transcrição (FTs) são proteínas capazes de regular inúmeras vias de sinalização em resposta ao crescimento e desenvolvimento vegetal, bem como aos estresses bióticos e abióticos por meio da interação com elementos *cis*-regulatórios de regiões promotoras de genes alvo, ativando ou reprimindo sua expressão (Li *et al.*, 2015). A modulação dos níveis e padrões de expressão dos FTs, contribui para a ativação de uma variedade de vias de sinalização (Huang *et al.*, 2010; Sornaraj *et al.*, 2016).

A família de fatores de transcrição bZIP é uma das principais famílias de FTs que atuam modulando a expressão de genes de resposta ao estresse em plantas. Além de resistência ao estresse, os FTs bZIPs atuam em muitos outros processos biológicos, incluindo o

desenvolvimento floral, maturação de sementes, fotomorfogênese, resistência a doenças e sinalização via ácido abscísico (ABA) (Liu *et al.*, 2017).

Os FTs bZIP são nomeados com base na sua característica comum, o domínio bZIP, que consiste em 60 a 80 aminoácidos, cercado por duas regiões funcionalmente distintas, uma região básica e um zíper de leucina (Leu). A região básica contém cerca de 18 resíduos de aminoácidos com um motivo invariante N-x₇-R/K-x₉, que é responsável pela localização nuclear e ligação ao DNA; o zíper de leucina é composto por várias repetições de sete leucinas ou outros aminoácidos hidrofóbicos como: isoleucina, valina, fenilalanina ou metionina, que são menos conservados e medeiam a homo e/ou heterodimerização (Jakoby *et al.*, 2002; Li *et al.*, 2015).

As proteínas bZIP apresentam a especificidade de ligação com *cis*-elementos regulatórios responsivos ao ácido abscísico (ABREs) de sequência de DNA contendo um núcleo ACGT. Esses FTs com domínio bZIP ligam-se preferencialmente a *cis*-elementos ABRE do tipo G-box (CACGTG), C-box (GACGTC) e A-box (TACGTA). Para interagir com o DNA, a porção N-terminal da região básica insere-se no sulco maior do DNA dupla fita e a metade C-terminal do zíper de leucina medeia a dimerização (Li *et al.*, 2015).

Basicamente, o mecanismo de ação do FT bZIP é iniciado durante eventos de estresse abiótico, como seca e alta salinidade, onde os níveis endógenos de ABA nas células vegetais aumentam, levando a expressão de genes de resposta ao estresse (Yoshida *et al.*, 2015). Grande parte dos mecanismos regulatórios relacionados ao estresse em plantas seguem uma via de resposta dependente de ácido abscísico (ABA), que é um fitormônio importante e reconhecidamente envolvido em diversos processos biológicos, incluindo a resposta ao estresse biótico e abiótico. A percepção do ABA pela célula vegetal se dá por meio de receptores citoplasmáticos específicos que promovem a transdução do sinal, por meio de uma cascata de fosforilação via proteínas quinases que fosforilam os FTs bZIPs, ativando-os. Após a ativação, os fatores de transcrição bZIP interagem com *cis*-elementos regulatórios, como os ABREs presente nas regiões promotoras de genes alvo, ocasionando o aumentando da expressão destes genes de resposta ao estresse (Banerjee e Roychoudhury, 2015).

Após a conclusão do sequenciamento de muitos genomas eucarióticos, diversos membros de FTs da família bZIP foram identificados e/ou preditos em nível de genoma. A quantidade de FTs da família bZIP varia entre organismos. Jakoby *et al.* (2002) descreveram 75 genes bZIP em *Arabidopsis* que compartilhavam o mesmo domínio proteico e com alta similaridade de sequência da região básica, agrupando-os em dez diferentes grupos. No

entanto, como a classificação não se baseia puramente em filogenia, foi considerada parcialmente subjetiva. Dröge-Laser *et al.* (2018), publicaram uma atualização nos estudos com *A. thaliana* e descreveram 78 membros bZIP no genoma da espécie divididos em 13 subgrupos.

Genes bZIP foram descritos em diversas outras espécies vegetais, como em mandioca com 77 genes (Hu *et al.*, 2016a), 69 em tomate (Li *et al.*, 2015), 89 em arroz (Zhang *et al.*, 2014), 125 em milho (Wei *et al.*, 2012), entre outras espécies. Até o momento, genes bZIP não foram caracterizados em espécies cítricas e evidências crescentes indicam que as proteínas bZIP modulam as redes de sinalização para uma série de respostas que conferem resistência e tolerância ao estresse abiótico. Portanto, as investigações visando a caracterização molecular da família de fatores de transcrição bZIP em *C. sinensis* fornecerão uma melhor compreensão das redes biológicas que são mediadas pelos genes bZIP, oferecendo informações para estudos futuros que visam aumentar a produtividade citrícola ao conferir resistência ao estresse por meio de modulação e expressão de proteínas bZIP específicas.

2. OBJETIVOS

2.1. Objetivo Geral

Caracterizar em nível molecular a família gênica de fatores de transcrição bZIP envolvidos na resposta à tolerância à seca, por meio de análises *in silico* e moleculares.

2.2. Objetivos Específicos

- i. Identificar os genes da família gênica bZIP em bancos de informação genômica de citros (Phytozome/CAP) e caracterizar a estrutura genômica desses genes;
- ii. Caracterizar a estrutura das proteínas codificadas pelos genes *CsbZIPs*;
- iii. Analisar a relação filogenética entre os membros da família gênica bZIP em citros e com os membros de *Arabidopsis thaliana*, buscando propor uma classificação para os genes bZIP de laranja doce;
- iv. Caracterizar a expressão gênica de membros da família bZIP em plantas de citros submetidas ao déficit hídrico.

3. REVISÃO BIBLIOGRÁFICA

3.1. Citros – Aspectos gerais

Dentro das Angiospermas, o gênero *Citrus* pertence à subfamília *Aurantioideae*, família *Rutaceae* e ordem Sapindales. Sua origem e época de dispersão são incertas, com registros sugerindo que a laranja é oriunda do sul asiático, provavelmente China, há cerca de 4.000 anos (Neves, 2010). Estudo recente com base em análises genômicas, filogenéticas e biogeográficas também sugere a origem asiática, mais precisamente no sudeste dos Himalaias, em uma região que inclui a área leste de Assam, norte de Mianmar e sudoeste da China, na província de Yunnan ocidental (Wu *et al.*, 2018).

A expansão do cultivo dos citros ocorreu através do comércio de laranja entre as nações e das guerras. Na Idade Média, a laranja foi levada para a Europa pelos Árabes e em 1500, durante as expedições de Cristóvão Colombo, mudas de citros foram transportadas para o continente americano. No Brasil, os citros foram introduzidos logo no início da colonização e no território brasileiro encontrou boas condições edafoclimáticas que favoreceram seu estabelecimento (Neves, 2010).

Atualmente, a produção global de laranja no período compreendido entre 2016/17 foi de 53.3 milhões de toneladas. A previsão é que a produção global de laranja para 2017/18 caia 4 milhões de toneladas em relação ao ano anterior, devido ao clima desfavorável que afeta principalmente o Brasil e os Estados Unidos (USDA, 2018). O Brasil ocupa o segundo lugar na produção citrícola global e destaca-se como líder mundial na produção de laranja, produzindo cerca de 17 milhões de toneladas da fruta, seguido pela China que ocupa o segundo lugar na produção mundial, produzindo cerca de 7 milhões de toneladas e da União Europeia que produz, aproximadamente, 6 milhões de toneladas da fruta (USDA, 2018).

De acordo com dados obtidos em setembro de 2018 e disponibilizados pelo IBGE (2018), a produção nacional de laranja foi de 17.1 milhões de toneladas, em uma área de cerca de 627.5 ha., com rendimento médio de 27.3 Kg/ha. A produção nacional concentra-se na região Sudeste, com destaque para os estados de São Paulo, com produção estimada em 13 milhões de toneladas, e Minas Gerais, que produziu aproximadamente 863 mil toneladas da fruta. O Nordeste é a segunda maior região produtora, com produção estimada em aproximadamente 1,5 milhões de toneladas, sendo o estado da Bahia o principal produtor da região, com estimativa de produção para o ano de 2018 em 830 mil toneladas (IBGE, 2018).

Uma das mais importantes espécies de laranja cultivada no mundo é a laranja doce (*Citrus sinensis* cv. Valência), sendo empregada principalmente na produção de suco de

laranja. Xu *et al.*, (2013) realizaram um estudo onde sequenciaram o genoma de *C. sinensis*, espécie diploide, com nove pares de cromossomos ($n = 9$; $2n = 18$) e um tamanho médio de ~ 367 Mb. Por meio da reconstrução do cenário referente aos antigos eventos primários de origem da laranja doce, foi possível entender o alto grau de heterozigosidade presente no genoma da espécie, sugerindo que a laranja doce é um híbrido interespecífico entre toranja (*C. grandis*) e tangerina (*C. reticulata*) (Xu *et al.*, 2103).

Sobre a hibridação, esclareceu-se que ocorreu entre toranja (parental feminino) e tangerina (parental masculino), dando origem a um híbrido interespecífico que foi retrocruzado com tangerina originando a espécie *C. sinensis*. Esse evento pode ter acontecido cerca de 2.300 anos atrás, embora mudanças genéticas provavelmente ocorreram após esses eventos de hibridação. Nota-se que as características genéticas desse genótipo híbrido ainda estão preservadas na laranja doce atual, provavelmente por causa da sua natureza reprodutiva estritamente assexuada (apomixia por embrião nucelar) e seleção artificial e propagação por enxertia. Esse cenário explica por que a maioria das cultivares de laranja doce nos dias de hoje são geneticamente um biótipo altamente heterozigótico, com a diversificação ocorrendo principalmente por meio de mutações somáticas (Xu *et al.*, 2013).

3.2. Estresse abiótico em plantas – aspectos moleculares

Todos os fatores ambientais não vivos capazes de afetar negativamente ou prejudicar o crescimento e a produtividade das plantas são considerados estressores abióticos. As plantas estão constantemente expostas aos mais variados tipos de condições ambientais desfavoráveis (seca, salinidade, temperaturas extremas, alagamento, déficits nutricionais, etc.) que provocam alterações no metabolismo, crescimento e desenvolvimento vegetal, podendo, em última instância, levar à morte das plantas (Gong *et al.*, 2013; Khan *et al.*, 2018).

Estima-se que as tensões abióticas possam limitar a produção de culturas em até 70% e que muitas culturas realizam apenas 30% do seu potencial genético em relação ao seu rendimento. Logo, as alterações relacionadas ao estresse, no que se refere ao desenvolvimento, crescimento e produtividade da planta, geram grandes preocupações na agricultura, pois limitam o rendimento e podem resultar em grandes perdas econômicas (Khan *et al.*, 2014).

Entretanto, as plantas desenvolveram respostas adaptativas integradas nos níveis molecular, celular e fisiológico para manter o seu crescimento, sobrevivência e produtividade em situações desfavoráveis (Banerjee e Roychoudhury, 2017). Os mecanismos para resistir

ao estresse por desidratação podem ser de retardo ou tolerância. As plantas que investem em estratégias que retardam o estresse requerem mecanismos que isolem as suas células da condição de estresse, a fim de minimizar o impacto sofrido, mesmo que o estresse esteja presente no ambiente (Khan *et al.*, 2014); estratégia adotada pelas plantas decíduas que perdem suas folhas na estação seca.

Alternativamente, as plantas que toleram o estresse podem estabelecer uma nova homeostase metabólica em resposta e continuar a crescer e se desenvolver sem grandes problemas. As plantas tolerantes podem se adaptar ao estresse, o que exige ajustes hereditários na sua estrutura e/ou função, aumentando a sua capacidade de sobrevivência neste ambiente; ou podem se aclimatar ao estresse, mudando sua fisiologia de forma não hereditária ao longo da sua vida útil, essas modificações são alcançadas por meio de exposição gradual ao estresse (Khan *et al.*, 2014).

A detecção de uma condição estressante pode resultar em variações na expressão gênica, causando mudanças na composição do transcriptoma vegetal, proteoma e metaboloma. É por meio da detecção do estresse que as cascatas de sinalização são iniciadas e podem ativar canais iônicos, cascatas de quinase, fatores de transcrição, produção de espécies reativas de oxigênio (ROS) e acúmulo de hormônios vegetais, como o ácido abscísico (ABA). Esses sinais, em última análise, induzem a expressão de subconjuntos específicos de genes que levam à reação defensiva geral (Khan *et al.*, 2014).

Os genes expressos durante as diferentes condições de estresse abiótico podem seguir distintas vias de transdução de sinal, podendo ser ativados por vias ABA-dependentes, ABA-independentes ou exigir um sistema de comunicação cruzada (*cross-talk*), no caso de genes responsivos às duas condições. O ABA é o hormônio vegetal de maior importância na relação estabelecida entre as plantas e os diversos tipos de estresse, pois está intimamente relacionado com a tolerância às várias condições de estresse e a alteração dos padrões de expressão gênica durante essas adversidades (Oliveira *et al.*, 2010). O ABA apresenta um papel importante no que se refere ao enfrentamento das condições de estresse abiótico. Sabe-se que durante a exposição das plantas ao estresse abiótico, ocorre o aumento da concentração de ABA, este é considerado um dos primeiros sinalizadores de estresse nas células vegetais; entretanto, após o alívio da condição de estresse, as concentrações de ABA voltam à normalidade (Vishwakarma *et al.*, 2017).

3.3. Fatores de Transcrição em Plantas e sua relação com o estresse abiótico

Os fatores de transcrição (FTs) são uma das maiores classes de proteínas codificadas por genes eucarióticos, representando cerca de 8% do genoma (Pruneda-Paz *et al.*, 2014). São proteínas nucleares que, quando ativadas, interagem com regiões de DNA específicas (elementos de resposta) na região promotora de um gene e ativam, reprimem ou acentuam a transcrição desse gene, levando a uma síntese aumentada ou diminuída da proteína codificada (Nelson *et al.*, 2012).

São divididos em duas categorias: i) fatores de transcrição basais (FTBs), presentes em todas as células e que atuam na expressão de todos os genes transcritos pelo mesmo tipo de RNA-polimerase; e ii) fatores de transcrição específicos (ativadores, coativadores, repressores e correpressores da transcrição), responsáveis pela expressão de genes que codificam proteínas restritas para um determinado tipo e/ou momento celular (Zaha *et al.*, 2014). Os FTs apresentam domínios proteicos característicos, compostos por dois motivos estruturais, sendo um motivo de ligação ao DNA e um motivo responsável pela dimerização com outras proteínas. Esta organização estrutural possibilita o transporte para o núcleo, o reconhecimento das sequências reguladoras dos genes, além da ativação da transcrição e da interação com outras proteínas (Alberts *et al.*, 2008; Zaha *et al.*, 2014).

A interação molecular entre os fatores de transcrição e o DNA depende de um contato específico entre as superfícies das duas moléculas. O DNA em um cromossomo consiste em uma dupla hélice longa cujas torções formam os sulcos maiores e menores da molécula. Geralmente, as regiões regulatórias do DNA encontram-se dispostas no sulco maior, pois nestas regiões as bases nitrogenadas estão mais expostas facilitando o contato entre a proteína e o DNA (Alberts *et al.*, 2008). O FT reconhece uma sequência regulatória específica porque a superfície da proteína é extensivamente complementar às características químicas da dupla hélice que exhibe essa sequência. Cada fator de transcrição estabelece uma série de contatos com o DNA via ligações de hidrogênio, interações iônicas e interações hidrofóbicas. Embora cada contato individual seja fraco, os contatos estabelecidos ao longo da interface FT-DNA garantem a especificidade e a força da interação (Alberts *et al.*, 2008).

A sequência de DNA que pode ser reconhecida por um monômero é relativamente curta, cerca de 6-8 pares de nucleotídeos, podendo comprometer a especificidade da interação FT-DNA, dada a probabilidade de redundância dessa sequência curta no genoma. Portanto, muitos fatores de transcrição formam dímeros, pois esse arranjo permite duplicar o tamanho

da sequência *cis*-regulatória a ser reconhecida no DNA, logo, aumenta a afinidade e a especificidade da interação (Alberts *et al.*, 2008). A dimerização entre dois monômeros idênticos, forma um homodímero; entre monômeros diferentes, forma-se um heterodímero. Heterodímeros são vantajosos por permitir plasticidade de regulação, pois um monômero podendo dimerizar com outros diferentes monômeros gera diversas especificidades de ligação ao DNA (Alberts *et al.*, 2008); em resumo, um único monômero pode regular a expressão de diferentes genes a depender da sua interação com outros monômeros.

A maioria dos fatores de transcrição conhecidos são agrupados em famílias de acordo com o seu domínio de ligação ao DNA (Riechmann *et al.*, 2000). Várias famílias de FTs de plantas desempenham papéis significativos na tradução de sinais de estresse abióticos em mudanças na expressão gênica. As maiores famílias de FTs relacionados ao estresse abiótico em plantas são: *basic leucine zipper* (bZIP), *APETALA2/ethylene-responsive element binding factor* (AP2/ERF), *NAM/ATAF1/CUC2* (NAC), WRKY, MYB, *zinc finger* (ZFs) e *basic helix-loop-helix* (bHLH). Boa parte dos membros dessas famílias foram caracterizados com relação a seus papéis na regulação das respostas ao estresse abiótico, e a expressão ectópica de vários FTs dessas famílias resultou em melhor tolerância ao estresse das culturas (Lindemose *et al.*, 2013).

Wang *et al.* (2016) ressaltam que os FTs são bons candidatos à engenharia genética para criar culturas tolerantes ao estresse, por conta do seu papel como reguladores mestres de muitos genes responsivos ao estresse. Estudos recentes utilizando fatores de transcrição para melhorar a tolerância ao estresse abiótico em plantas por transgenia confirmam o potencial dos genes de fatores de transcrição para modular a expressão de genes responsivos ao estresse abiótico em plantas.

O estudo de um grande número de fatores de transcrição independentes e dependentes de ABA e seus *cis*-elementos pode fornecer uma melhor compreensão da base molecular da tolerância ao estresse, que envolve uma complexa rede de genes operando em coordenação (Agarwal e Jha, 2010).

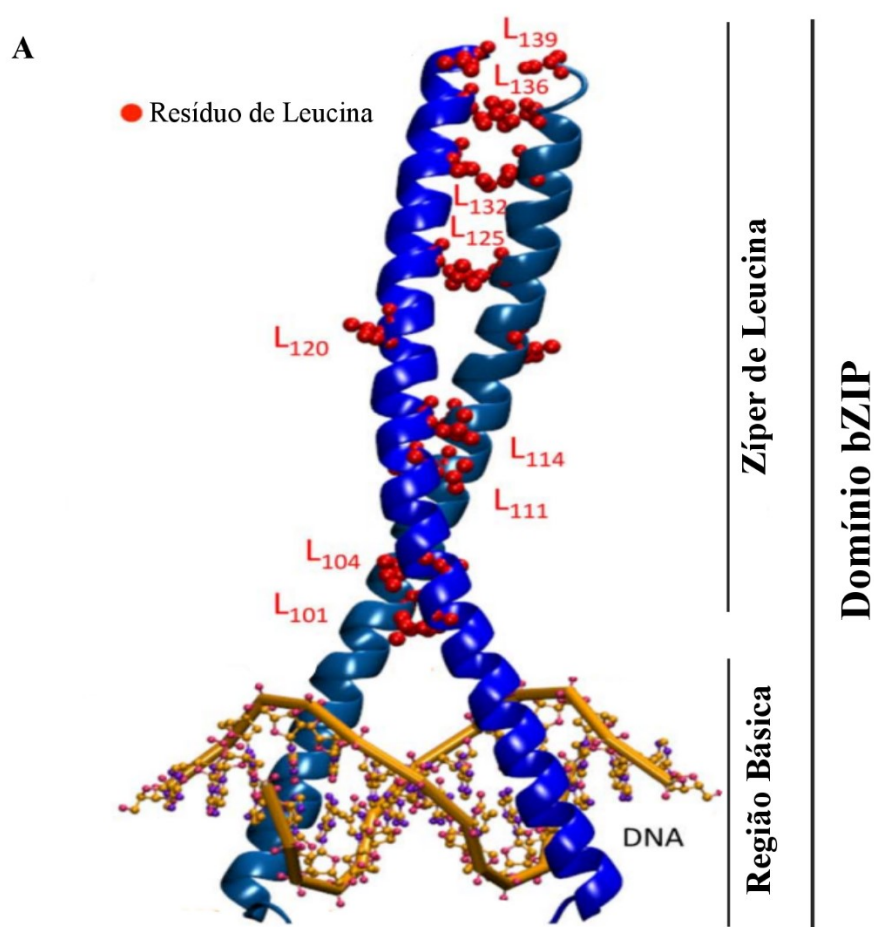
3.4. Família bZIP (*Basic Leucine Zipper*) de fatores de transcrição

Os fatores de transcrição bZIP (*Basic Leucine Zipper*) representam uma família multigênica grande e diversificada dentre os eucariotos, com membros descritos em muitas

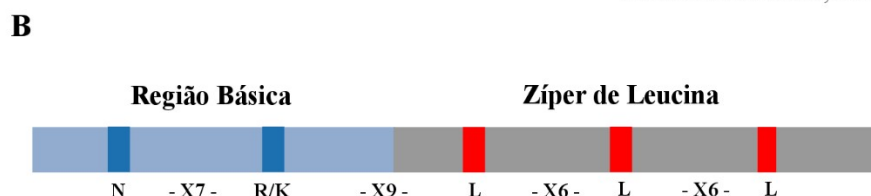
espécies vegetais: 77 genes descritos em mandioca (Hu *et al.*, 2016a), 69 em tomate (Li *et al.*, 2015), 78 em *Arabidopsis thaliana* (Dröge-Laser *et al.*, 2018), 121 em banana (Hu *et al.*, 2016b), 89 em arroz (Nijhawan *et al.*, 2008), 125 em milho (Wei *et al.*, 2012), entre outras culturas. Os FTs bZIP compreendem a maior família conhecida de proteínas de ligação ao DNA induzidas por ABA e podem regular inúmeros processos biológicos, como: diferenciação de órgãos e tecidos, defesa de patógenos, metabolismo energético, desenvolvimento floral, sinalização hormonal via ABA, resposta a estresse abiótico, alongamento celular (Jakoby *et al.*, 2002; Côrrea *et al.*, 2008; Pandey *et al.*, 2018), entre outros.

Uma das primeiras descrições de genes *bZIP* de fatores de transcrição em *Arabidopsis thaliana* foi realizada por Jakoby *et al.* (2002), que descreveram 75 genes *bZIP* divididos em 10 subgrupos (A-I e S), havendo três genes (*AtbZIP60*, *AtbZIP62* e *AtbZIP72*) que não foram agrupados, pois diferiam consideravelmente na estrutura dos motivos conservados. Uma atualização dos estudos de 2002 com *A. thaliana* foi realizada por Dröge-Laser *et al.* (2018), onde foram descritos 78 genes *bZIP* divididos em 13 grupos (A-M); das atualizações realizadas nesse estudo: adicionaram os subgrupos J, K e M para classificar os três genes que não foram agrupados por Jakoby *et al.* (2002), inseriram quatro novos genes no grupo E e eliminaram o gene *AtbZIP73* – subgrupo I, por se tratar de um pseudogene. Esses trabalhos com *A. thaliana* são importantes, pois os organismos modelo possuem mecanismos genéticos bem elucidados e comuns a um grande número de organismos relacionados, podendo ser utilizados como referência em novos estudos (Dröge-Laser *et al.*, 2018).

A família bZIP de FTs caracteriza-se pela presença de um domínio bZIP altamente conservado, composto por dois motivos, sendo: 1) uma região básica formada por aproximadamente 16 resíduos de aminoácidos contendo um sinal de localização nuclear, seguido por um motivo invariante (N-x7-R/K) de ligação ao DNA, na região N-terminal; e 2) uma região denominada zíper de leucina contendo uma repetição de leucinas ou outros aminoácidos hidrofóbicos (ex.: Ile, Phe, ou Val) a cada sete aminoácidos, posicionados exatamente nove aminoácidos em direção à região C-terminal, criando uma hélice anfipática (Figura 1) (Jakoby *et al.*, 2002; Wang *et al.*, 2016).



Fonte: Goudot *et al.*, 2011

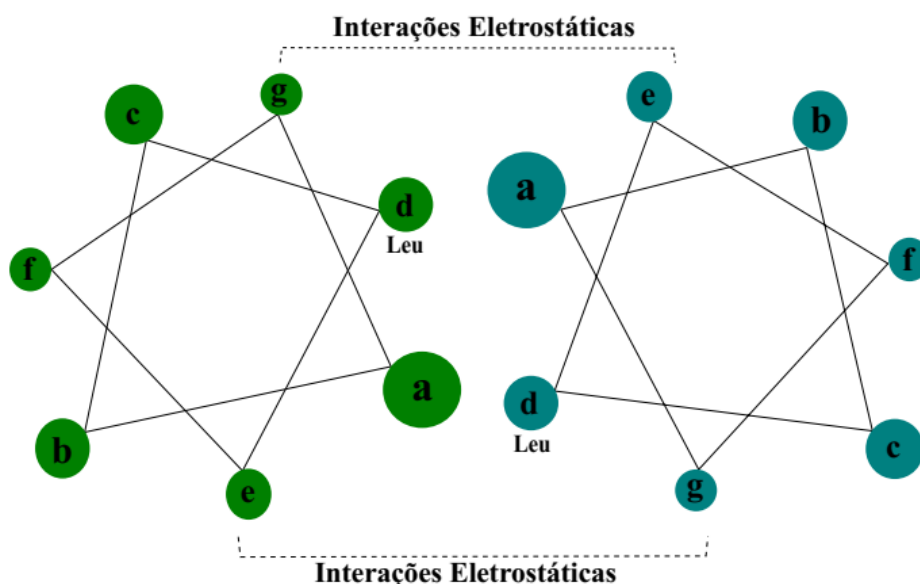


Fonte: Jakoby *et al.*, 2002

Figura 1. Estrutura do domínio bZIP. A) Estrutura terciária em α -hélice do domínio bZIP. B) O motivo referente à região básica está representado em azul e o motivo de zíper de leucina em cinza. Os resíduos altamente conservados estão destacados pelas barras azuis e vermelhas (uma sequência consenso é dada abaixo, onde: N = asparagina, R/K = arginina/lisina e L=leucina; X = outros resíduos de aminoácidos quaisquer).

Os fatores de transcrição bZIP interagem com o DNA e regulam a transcrição formando dímeros (heterodímeros/homodímeros) (Llorca *et al.*, 2014). A dimerização é possível devido à interação que se estabelece entre os resíduos de aminoácidos existentes na repetição de leucinas do motivo zíper de leucina que compõe o domínio bZIP. Os aminoácidos dentro desta região repetida (*heptad*) são classificados de *a-g*, com a leucina ocupando a posição *d*. As leucinas e os resíduos de aminoácidos apolares nas posições *a* e *d* garantem a

estabilidade do dímero. Os aminoácidos da posição *e* e *g* são principalmente aminoácidos hidrofóbicos e promovem a estabilidade do dímero por meio de interações eletrostáticas. Desta forma, os aminoácidos de um monômero, formam uma interface apolar, hidrofóbica, com o zíper de leucina de outro monômero (Figura 2) (Shen *et al.*, 2007).



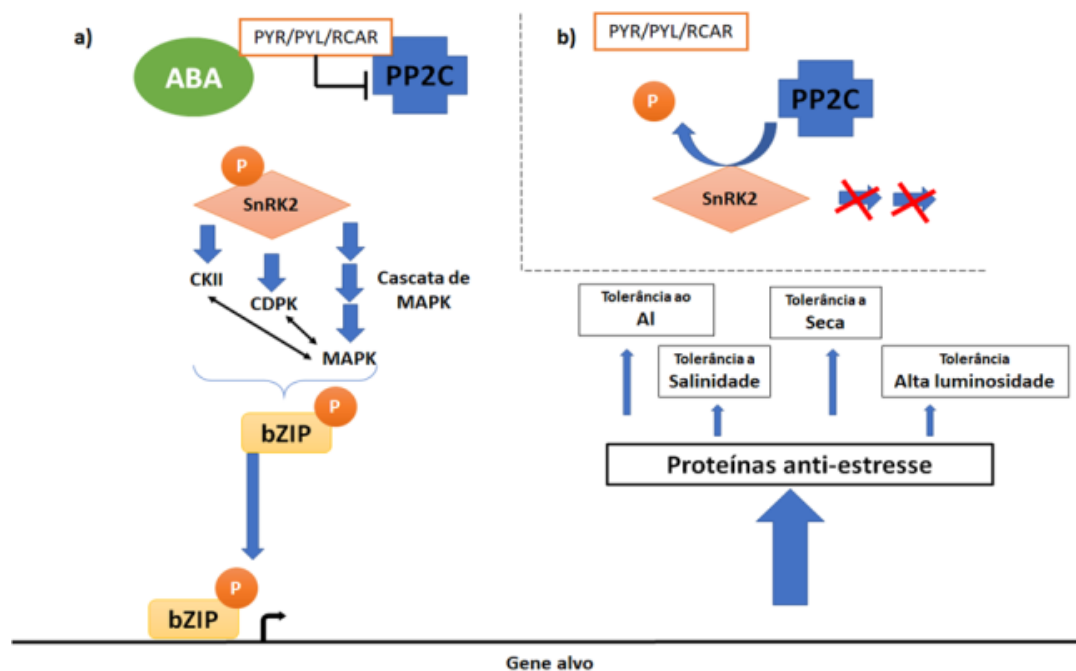
Fonte: Härmä *et al.*, 2018

Figura 2. Esquema ilustrando a dimerização entre os motivos de zíper de leucina. A posição dos resíduos de aminoácidos conservados do motivo zíper de leucina (indicados pelas letras: a, b, c, d, e, f, g) e as interações que estabilizam o dímero (indicadas pelas setas tracejadas).

Os fatores de transcrição bZIP seguem uma via de resposta ABA-dependente, logo, reconhecem um *cis*-elemento conservado de 8 pb denominado elemento responsivo ao ácido abscísico (ABRE, PyACGTGG/TC), sendo o ACGT o núcleo da sequência que também é referido como G-box. Geralmente, múltiplas cópias do *cis*-elemento ABRE estão presentes na região promotora de genes induzidos por estresse, podendo iniciar uma transcrição dependente de ABA (Banerjee e Roychoudhury, 2015).

Durante eventos de estresse, sabe-se que os níveis endógenos de ABA na célula vegetal aumentam, e este fitormônio atua como sinalizador da condição de estresse. Banerjee e Roychoudhury (2015) elucidaram o mecanismo de atuação do fator de transcrição bZIP, ABA-dependente, durante resposta ao estresse abiótico (Figura 3). Quando a célula vegetal apresenta altos níveis de ABA, a classe de receptores PYR/PYL/RCAR sequestra a fosfatase PP2C, impedindo a desfosforilação da proteína quinase SnRK2, que permanece ativa em seu estado fosforilado, iniciando uma cascata de fosforilação via proteínas quinases, como: MAPK (*mitogen activated protein kinases*); CDPK (*calcium-dependent protein kinase*) ou CKII (*casein kinase II*), que irão fosforilar o fator de transcrição bZIP. Fosforilado, o FT bZIP

interage com o *cis*-elemento na região promotora dos genes alvos, induzindo a transcrição destes. A proteína codificada age a jusante como um efetor crucial das respostas aos estresses abióticos (salinidade, seca, toxicidade ao Al, alta luminosidade, etc.). Quando a célula vegetal apresenta baixos níveis de ABA, a fosfatase PP2C permanece ativa e age desfosforilando a quinase SnRK2, impedindo a cascata de fosforilação que iria ativar o fator de transcrição (Banerjee e Roychoudhury, 2015).



Fonte: Banerjee e Roychoudhury, 2017

Figura 3. Ação do fator de transcrição bZIP, que segue via de transdução de sinal ABA-dependente durante resposta ao estresse abiótico. a) Na presença do sinal de estresse abiótico, ocorre o aumento da concentração de ABA na célula, este é percebido pelos receptores de ABA citoplasmáticos (PYR/PYL/RCAR) até a expressão de genes de resposta que irão codificar proteínas anti-estresse para induzir a tolerância ao estresse abiótico nas plantas; b) Na ausência da condição de estresse, a célula apresenta baixas concentrações de ABA, os receptores de ABA citoplasmático não impedem a atividade da fosfatase PP2C que atua desfosforilando a quinase SnRK2, evitando a cascata de fosforilação que culminaria com a ativação de genes de resposta.

Assim como outros FTs, os bZIP não desempenham apenas papéis essenciais no processo de desenvolvimento vegetal, mas também respondem aos diversos estresses abióticos, dentre eles: seca e salinidade. Alguns genes que atuam no aumento da tolerância a seca e a salinidade foram caracterizados em milho (Zhang *et al.*, 2011), arroz (Liu *et al.*, 2014a), trigo (Zhang L. *et al.*, 2015), algodão (Liang *et al.*, 2016) e *Poncirus trifoliata* (Huang *et al.*, 2010)

4. MATERIAL E MÉTODOS

4.1. Identificação dos genes da família bZIP em *Citrus sinensis* L.

Para identificar os genes codificantes para fatores de transcrição bZIP no genoma de laranja doce (*C. sinensis* L.), 78 sequências proteicas de fatores de transcrição bZIP de *Arabidopsis thaliana* (Dröge-Laser *et al.*, 2018) foram recuperadas na base de dados TAIR (<https://www.arabidopsis.org/>) e utilizadas para realizar BLASTs contra o genoma de *C. sinensis* disponível na base de dados Phytozome (<https://phytozome.jgi.doe.gov>), usando a ferramenta tBLASTn (nucleotídeo → proteína). Apenas as sequências não redundantes (com ID exclusivo do local) foram mantidas nas análises. A localização física dos genes *CsbZIPs* foi determinada confirmando a posição inicial dos genes em cada cromossomo disponível no banco de dados CAP (*Citrus sinensis* Annotation Project; <http://citrus.hzau.edu.cn/cgi-bin/orange/blast>) (Wang *et al.*, 2014). Para plotar os locos genéticos nos cromossomos de laranja doce, utilizou-se o *software* MapChart (Voorrips, 2002).

4.2. Análise da estrutura dos genes e promotores

A estrutura éxon-intron dos *CsbZIP* foi analisada usando o programa Gene Structure Display Server (GSDS) (<http://gsds.cbi.pku.edu.cn/>), comparando as sequências de codificação (CDS) com as suas sequências genômicas correspondentes disponíveis na base de dados Phytozome. Uma região de 1500 pb *upstream* ao sítio de início da tradução foi analisada por meio do *software* PLACE (<https://sogo.dna.affrc.go.jp>) (Higo *et al.*, 1999) para identificar os *cis*-elementos relacionados ao estresse abiótico.

4.3. Alinhamento e análise filogenética

Para entender as relações evolutivas entre os genes da família bZIP de laranja doce, as sequências de aminoácidos das proteínas *CsbZIP* e seus ortólogos foram alinhadas usando o programa ClustalW (Thompson *et al.*, 1994). A árvore filogenética foi construída usando o programa MEGA6 (Tamura *et al.*, 2013) e o método de agrupamento Neighbor-Joining (NJ) (Saitou e Nei, 1987) com repetições de 1000 *bootstraps*. As árvores foram visualizadas e editadas graficamente no *software* FigTree v1.4.3.

4.4. Propriedades físico-químicas das proteínas, análise dos motivos conservados e rede de interação

O peso molecular (MW), ponto isoelétrico (pI) e GRAVY (*grand average of hydropathy*) das proteínas CsbZIPs foram preditos pela ferramenta PROTPARAM (<http://web.expasy.org/protparam/>). Para as predições de sítios de fosforilação em resíduos de serina, treonina e serina, foi utilizado o servidor NetPhos 3.1 (<http://www.cbs.dtu.dk>). Os motivos conservados das proteínas codificadas pelos genes *CsbZIP* foram analisados usando o programa MEME Suite 5.0.3 (<http://meme-suite.org/index.html>). A rede de interação entre as bZIPs de *Arabidopsis thaliana* (AtbZIPs) foi construída usando o banco de dados STRING (<https://string-db.org>) por homologia as proteínas CsbZIP de laranja doce.

4.5. Material vegetal e tratamentos de déficit hídrico

Folhas de plantas submetidas ao déficit hídrico (Silva, 2018) foram utilizadas para as análises de qRT-PCR. As plantas são do genótipo híbrido ‘LCR x TR – 001’, porta enxerto resultante do cruzamento entre limoeiro ‘Cravo’ (*Citrus limonia* L. Osbeck) x trifoliata (*Poncirus trifoliata* L. Raf.), enxertadas com a copa de laranjeira ‘Valência’ (*Citrus sinensis* L. Osb.). As plantas foram submetidas a quatro diferentes regimes hídricos (controle, estresse moderado, estresse severo e reidratação). Os regimes hídricos foram definidos com base no potencial hídrico foliar (Ψ_L). Para o tratamento ‘controle’, a umidade do solo ficou situada próximo a capacidade de campo e o Ψ_L foi mantido entre -0,1 e -0,5 Mpa. Para o tratamento ‘estresse moderado’, houve a suspensão da irrigação e o estresse foi caracterizado quando as folhas atingiram valores de Ψ_L entre -2,5 e -3,5 MPa. Para o ‘estresse severo’, foi definido valores de Ψ_L menores que -3,5 MPa e para o tratamento ‘reidratado’, depois que as plantas atingiram Ψ_L abaixo de -3,5 MPa a umidade do solo foi elevada para a capacidade de campo (Silva, 2018).

4.6. Extração de RNA e síntese de cDNA

As amostras de RNA total foram isoladas utilizando o Kit RNAqueous (Ambion, Inc.),

seguindo as instruções do fabricante. A qualidade e a integridade do RNA extraído foi avaliada por análise em gel de agarose a 1,5%. As amostras foram tratadas com DNase após o processo de extração, usando o Kit Turbo DNA-free (Ambion, Inc.), seguindo as instruções do fabricante (1 µL de DNase). A reação de transcrição reversa foi realizada usando o kit RETROscript (Ambion Inc.), seguindo as instruções do fabricante.

4.7. Análise da expressão de genes *CsbZIPs*

Oligonucleotídeos específicos foram desenhados (Apêndice 1) para amplificar dois genes (Cs8g06020 e Cs3g23480) que codificam fatores de transcrição bZIP, baseando-se nas suas sequências disponíveis no banco de dados Phytozome (<https://phytozome.jgi.doe.gov>). As reações de amplificação foram realizadas em termociclador da Stratagene Mx 3005P (Agilent Technologies), contendo o software MxPro-Mx 3005P, usando o kit Maxima™ SYBR Green/ROX qPCR Master Mix (Thermo Scientific). Como controle endógeno (gene de referência) utilizou-se o gene da gliceraldeído-3-fosfato desidrogenase C2 (GAPC2) (Mafra *et al.*, 2012).

As reações foram feitas em três réplicas, com volume utilizado na reação de 25 µL, contendo: 100 ng de cDNA, 1 µL de cada primer específico (R+F) 10 µmol.L⁻¹ e 11 µL de Maxima® SYBR Green/ROX qRT-PCR Master Mix (2X) (Fermentas). Para testar a especificidade dos oligonucleotídeos, os produtos foram analisados por curva de dissociação. A ciclagem de amplificação foi realizada nas seguintes etapas: 50°C por 2 minutos, 95°C por 10 minutos, 95°C por 15 segundos e 60°C por 1 minuto. Essas etapas foram repetidas por 40 ciclos. Após a ciclagem de amplificação dos genes, foi realizada a dissociação dos produtos para confirmar a formação de produtos únicos e específicos, neste caso as etapas realizadas apresentaram as seguintes características: 95 °C por 15 segundos, 60 °C por 1 minuto e 95 °C por 1 segundo.

Para quantificação da expressão utilizou-se o método $2^{-\Delta\Delta C_t}$ (Livak e Schmittgen, 2001), considerando a média dos valores de C_t das três réplicas realizadas. Os resultados foram submetidos à análise de variância ($P \leq 0,05$) e teste t de Student. Todas as análises estatísticas foram realizadas no ambiente R.

Para gerar um *heatmap* com os dados de RNA-seq obtidos no banco de dados de laranja doce do CAP, foi realizada uma análise de *hierarchical clustering*. Essas análises foram desenvolvidas em ambiente R com a função *Heatmap* do pacote *ComplexHeatmap* (Gu,

2016). Os valores de expressão foram convertidos para escala de *z-score*, com base nas linhas. Para o cálculo de correlação, foi utilizada a correlação de Pearson e para o agrupamento foi empregado o método *complete linkage*.

5. RESULTADOS E DISCUSSÃO

5.1. Identificação e classificação dos genes *CsbZIP*

Um total de 34 genes *bZIPs* (*CsbZIPs*) não redundantes foram identificados no genoma de laranja doce (Figura 4). Comparativamente, a família *bZIP* de fatores de transcrição de laranja doce é menor em relação a outras espécies vegetais, como a videira com 55 genes descritos (Liu *et al.*, 2014b), mandioca com 77 (Hu *et al.*, 2016a), banana com 121 (Hu *et al.*, 2016b), milho com 125 (Wei *et al.*, 2012), *Arabidopsis thaliana* com 78 (Dröge-Laser *et al.*, 2018) e tomate com 69 (Li *et al.*, 2015).

Dos 34 genes *CsbZIPs* identificados, 31 foram distribuídos em seis dos nove cromossomos da espécie: chr 1, 3, 5, 6, 7 e 8. Os outros três genes identificados (orange1.1t04546, orange1.1t00453 e orange1.1t01674) encontram-se em *scaffolds*, pois o genoma de laranja doce foi sequenciado recentemente (Xu *et al.*, 2013) e os cromossomos ainda são rascunhos com genes a serem anotados. Para fins didáticos, os genes distribuídos em *scaffolds* foram incluídos em um cromossomo desconhecido (*Chr Un*). Maior abundância de genes *CsbZIPs* foi observada nos cromossomos 3 e 7, contendo 10 e 8 genes, respectivamente. O cromossomo 5 apresentou apenas um gene *CsbZIP* (Cs5g23040), enquanto que os cromossomos 1, 6 e 8 apresentaram 3, 5 e 4 genes cada, respectivamente (Figura 4).

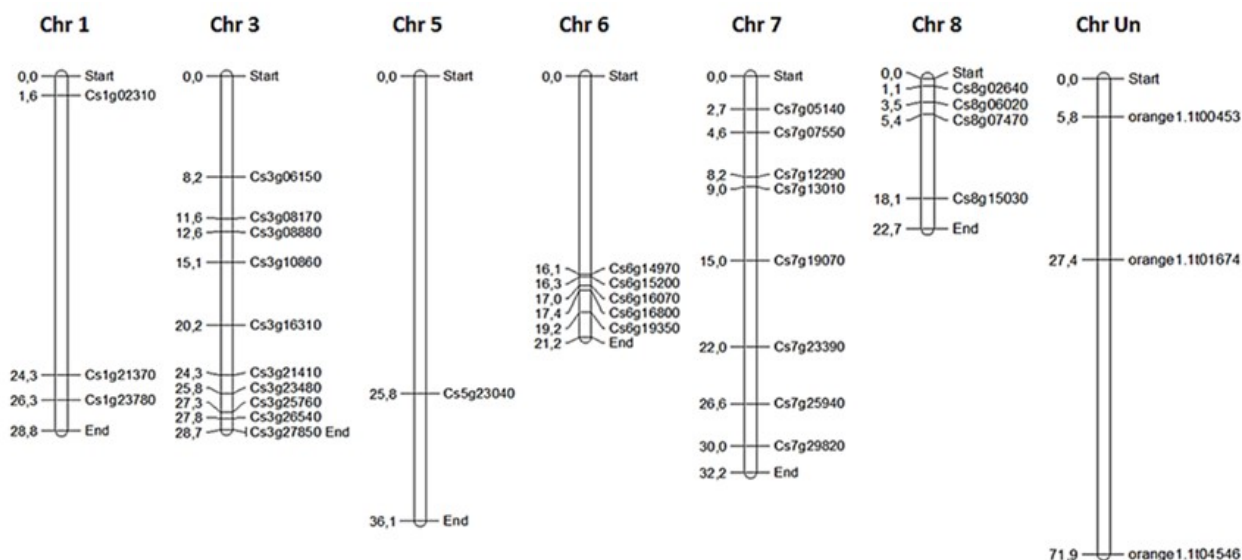


Figura 4. Localização cromossômica dos genes codantes para fatores de transcrição bZIP em laranja doce. Distribuição e posição dos genes *CsbZIPs* em 6 dos 9 cromossomos de *C. sinensis*; três genes que não foram distribuídos em cromossomos verdadeiros da espécie foram distribuídos em um cromossomo desconhecido (Chromosome Unknown = ChrUn). Os números à esquerda dos cromossomos referem-se ao início dos genes *CsbZIPs* em Mb (megabases). À direita está o código identificador dos locos de acordo com classificação disponível na base de dados CAP (*Citrus sinensis* Annotation Project).

Em busca de *cis*-elementos responsivos ao estresse nos genes *CsbZIP*, realizou-se a análise da região promotora (1,5kb *upstream*) usando a base de dados PlantCARE. Como é possível observar na Figura 5, a região promotora de todos os genes *CsbZIPs* de laranja doce apresentaram no mínimo 1 e no máximo 20 *cis*-elementos estresse responsivos, como DRE (*dehydration-responsive element*)/CRT (*C-repeat*), MBS (*MYB binding site involved in drought-inducibility*), LTRE (*low temperature-responsive element*) e ABRE (*abscisic acid (ABA)- responsive element*), exceto o gene Cs8g15030, que não apresentou estes *cis*-elementos.

O *cis*-elemento ABRE responsivo ao ABA foi o mais abundante, presente na região promotora de 22 genes *CsbZIPs*, sendo que o gene Cs6g15200 possui 20 ABREs, os genes Cs1g23780 e Cs5g23040 possuem 16 ABREs e o gene Cs3g21410 têm 9 ABREs em suas regiões promotoras. Espera-se que os genes com maior quantidade de *cis*-elementos ABREs respondam ao estresse abiótico, no entanto, duas sequências ABREs são suficientes para desencadear resposta ao estresse abiótico dependente de ABA (Yamaguchi-Shinozaki e Shinozaki, 2005). Uno *et al.* (2000), observaram que a ocorrência de dois *cis*-elementos ABREs na região promotora do gene *rd29B* de *Arabidopsis*, gene responsivo a desidratação e induzido por ABA, é suficiente para a sua expressão em resposta a desidratação.

CIS-ELEMENTOS ESTRESSE RESPONSIVOS

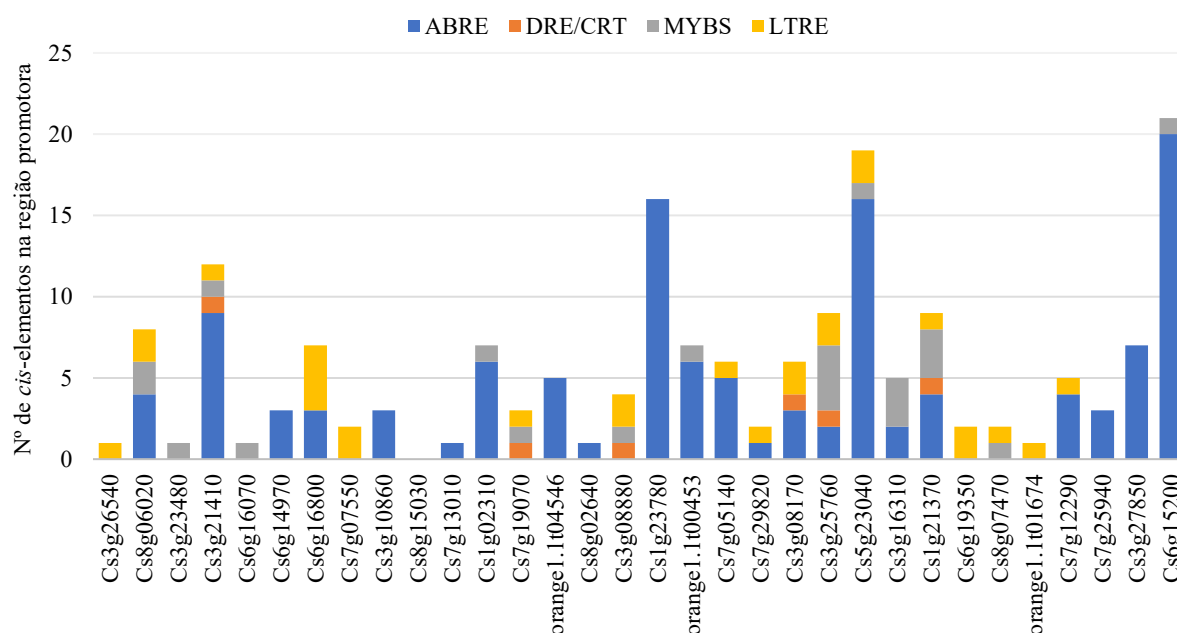


Figura 5. Análise de cis-elementos estresse-responsivos: ABRE (ACGTG), MYBs (TAACTG), DRE/CRT (G/ACCGAC) e LTRE (CCGAC) na região promotora dos genes bZIP de laranja doce. A região analisada encontra-se 1,5 Kb upstream ao sítio de início da transcrição dos genes CsbZIP usando a base de dados PLACE.

Buscando propor uma classificação para CsbZIPs com base na classificação de *Arabidopsis*, que apresenta 78 genes AtbZIPs descritos, classificados em 13 grupos (A-M e S), de acordo com a homologia da região básica e a similaridade dos motivos conservados (Dröge-Laser *et al.*, 2018), realizou-se o alinhamento da região do domínio bZIP de 77 AtbZIPs (exceto o gene *AT2G13130-73* por se tratar de um pseudogene) e das 34 CsbZIPs. É possível observar na filogenia (Figura 6) que as 34 CsbZIPs foram classificadas em 12 grupos, sendo que nenhum CsbZIP foi classificado dentro do grupo M. Também é possível observar que o gene *AT5G07160-72*, o único membro do grupo M de *Arabidopsis*, não foi agrupado com nenhum outro gene de laranja doce.

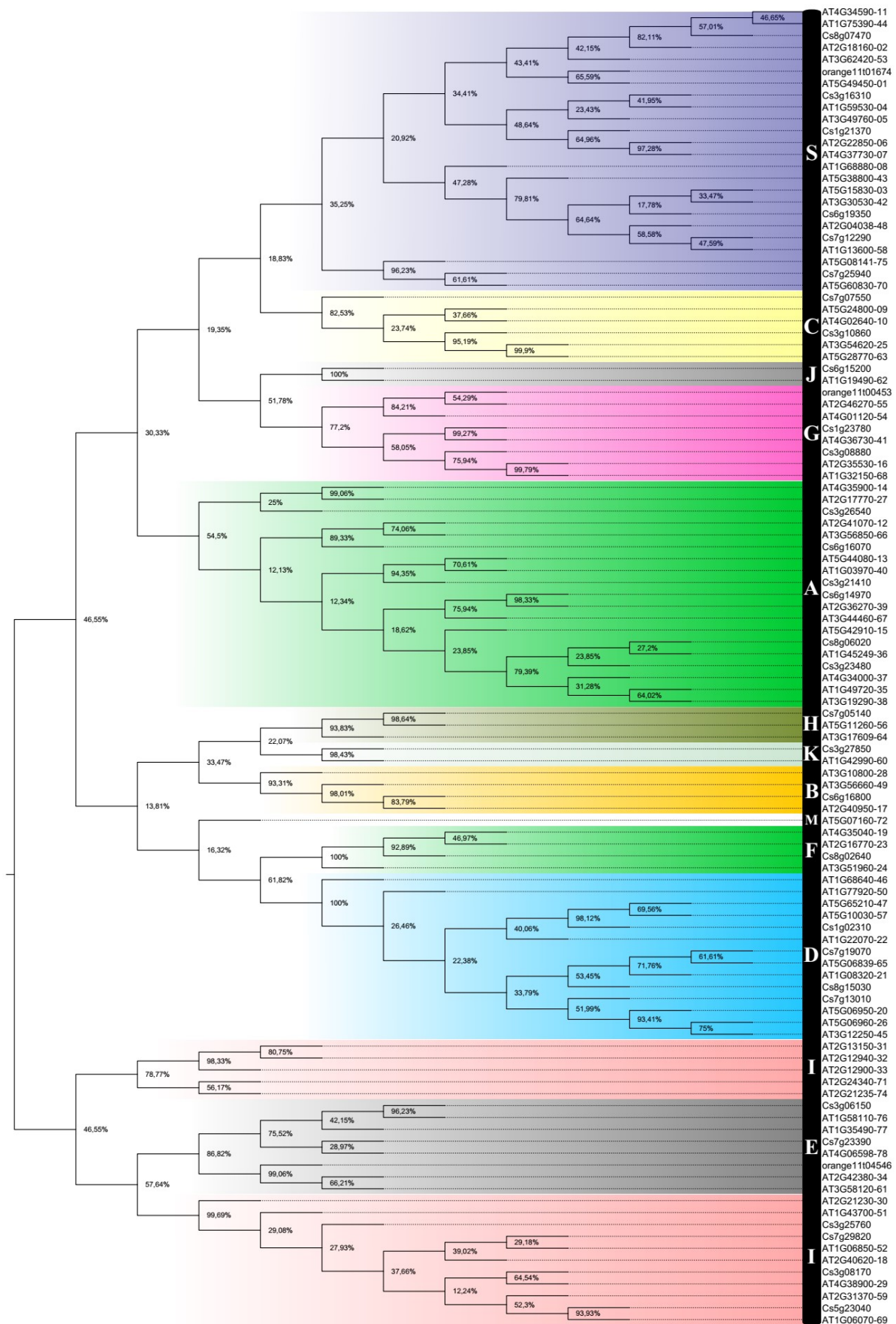


Figura 6. Relação filogenética entre 77 AtbZIPs e 34 CsbZIPs. A árvore filogenética foi baseada no alinhamento do domínio bZIP das sequências proteicas de *Arabidopsis thaliana* e *C. sinensis*. Dendrograma construído por meio do software MEGA 6, usando o método de agrupamento Neighbor-Joining e 1000 repetições de *bootstraps*. Os valores de *bootstrap* são indicados em cada nó. As proteínas CsbZIP são agrupadas em 12 clados distintos.

Buscando refinar a classificação dos CsbZIP e conhecer a estrutura éxon/intron dos genes foi utilizado o software *Gene Structure Display Server 2.0* comparando as sequências codantes (CDS) com as suas respectivas sequências genômicas de acordo com os modelos dos genes de laranja doce anotados na base de dados do *Phytozome* (Figura 7).

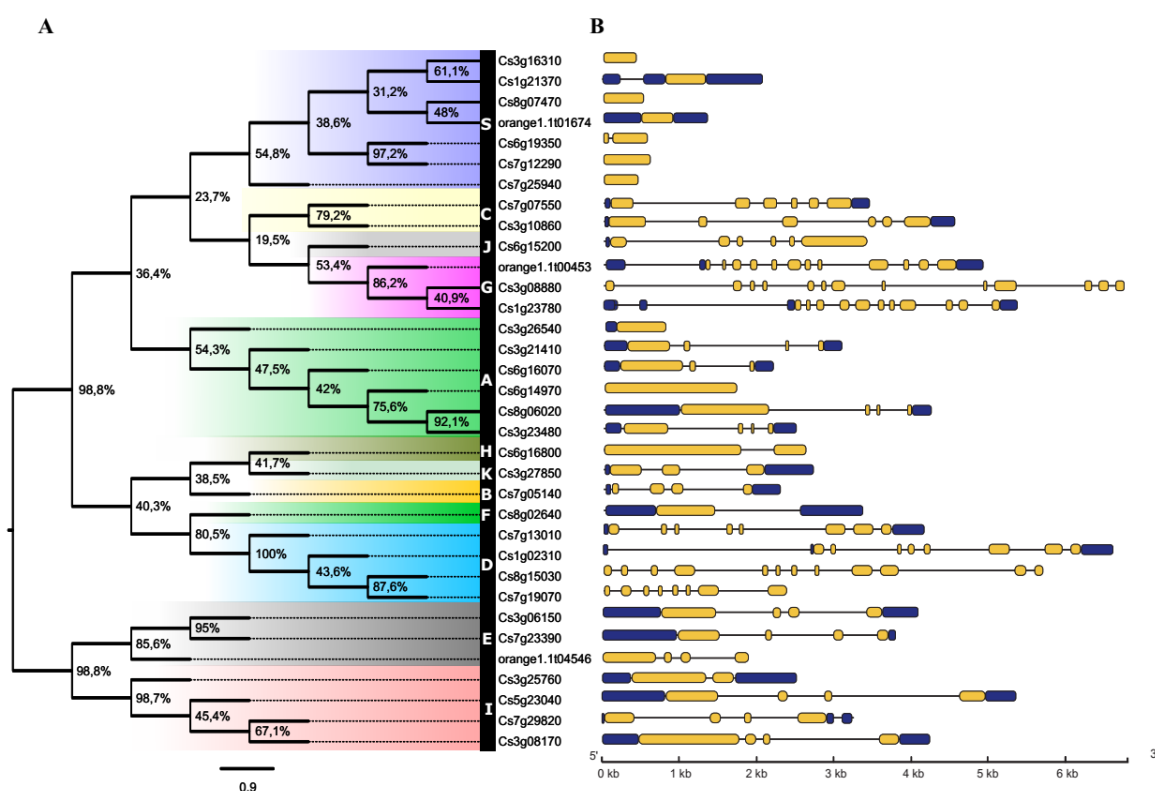


Figura 7. Relação filogenética e estrutura gênica dos CsbZIPs. A) Árvore filogenética produzida no software MEGA 6, usando o método de agrupamento Neighbor-Joining com réplicas de 1000 bootstrap. B) Estrutura éxon/intron dos genes *CsbZIPs*. As caixas azuis representam às regiões 5' e 3' não traduzidas (UTRs), as caixas amarelas representam os éxons e as linhas pretas os íntrons.

A organização éxon-ínton de uma família gênica carrega uma marca evolutiva desta família (Wei *et al.*, 2012) e sabe-se que os íntrons presentes em genes eucarióticos fornecem uma vantagem evolutiva pelo aumento da diversidade de proteínas geradas por consequência do embaralhamento de éxons (*exon shuffling*) e *splicing* alternativo. Estas informações podem ser úteis para fazer inferências quanto a evolução dos genes de uma família gênica (Gorlova *et al.*, 2014).

Como é possível observar na Figura 7, o número de íntrons nos genes *CsbZIPs* variam de 0 a 12 e o número de éxons variam de 1 a 13. Seis genes não apresentam íntrons, representando 17,6% do total de genes identificados. Os genes classificados dentro dos grupos D e G são os que apresentam o maior número de íntrons (7 a 12 íntrons) e os genes do grupo S são os que apresentam o menor número de íntrons, com genes com apenas 1 íntron ou que não apresentam íntrons. No entanto, estes genes são ortólogos a genes de *Arabidopsis* classificados dentro do grupo S, sendo a presença de genes com poucos ou nenhum íntron e o tamanho aproximado das proteínas em torno de 20 KDa as principais características do grupo, com base na classificação realizada por Dröge-Laser *et al.* (2018) em *Arabidopsis thaliana*. O grupo S de *Arabidopsis* abriga o maior número de genes *AtbZIPs*, com 17 genes no total. Da mesma forma, o grupo S de laranja doce também apresentou o maior número de genes, com 7 *CsbZIPs* classificados (Figura 7).

De acordo com a descrição de Dröge-Laser *et al.* (2018) sobre o grupo S de *A. thaliana*, outra característica do grupo é a heterodimerização preferencial com proteínas do grupo C, sendo referidas como a rede C/S₁-bZIP, estando relacionados com a sinalização diante do déficit nutricional, ajudando a promover a homeostase energética (Dröge-Laser e Weiste, 2018). É interessante notar, com base na filogenia, que os genes dos grupos S e C de laranja doce encontram-se dispostos em clados próximos e desse modo supõe-se que eles compartilham algum traço evolutivo que pode explicar a formação preferencial de heterodímeros entre as proteínas codificadas por esses genes.

O segundo maior grupo observado em laranja doce foi o grupo A, com seis membros representantes (Cs3g21410, Cs6g16070, Cs3g26540, Cs6g14970, Cs8g06020 e Cs3g23480). O grupo A de *Arabidopsis* também é o segundo com maior número de genes, com 13 *AtbZIPs* classificados que codificam para proteínas regulatórias que desempenham atuação central na via de sinalização ABA-dependente, como: AREBs/ABFs (fatores de ligação ao elemento responsivo ao ABA), DPBFs (fatores de ligação ao promotor DC3), que é expresso durante a embriogênese e também pode ser induzido por ABA, sendo que os genes DC3 pertencem a classe das proteínas LEA (*late embryogenesis abundant*) (Kim *et al.*, 1997), FDs, fatores de transcrição bZIP que atuam promovendo a transição floral e induzindo o desenvolvimento floral (Abe *et al.*, 2005) e GBFs (fator de ligação ao *cis*-elemento G-box).

Os membros do grupo A de *Arabidopsis* são induzidos transcricionalmente por estresses abióticos, como dessecação, frio, salinidade ou seca, e pós-transcricionalmente ativados por ABA, sendo capazes de induzir respostas adaptativas para neutralizar os efeitos

do déficit hídrico nos tecidos vegetais (Dröge-Laser *et al.*, 2018; Banerjee e Roychoudhury, 2017). Os seis genes do grupo A de laranja doce são ortólogos a genes de *Arabidopsis* conhecidos por atuar na modulação da resposta relacionada ao aumento da tolerância/resistência a estresses abióticos, como seca e salinidade (Dröge-Laser *et al.*, 2018). Para observar a interação entre as proteínas bZIP de laranja doce (CsbZIP) do grupo A com outras proteínas, construiu-se uma rede de interação com base na homologia entre as proteínas bZIP de laranja doce e de *Arabidopsis* (Figura 8).

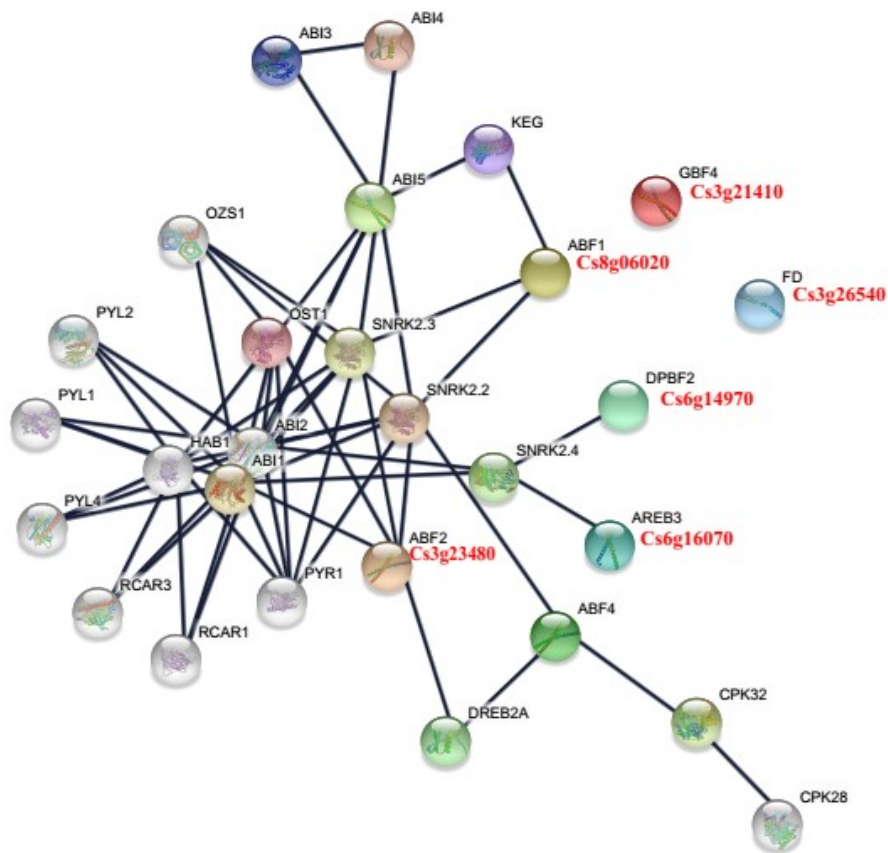


Figura 8. Rede de interação de proteínas bZIP em *A. thaliana* e *C. sinensis*. A rede foi construída por meio do banco de dados STRING, com alta confiança (score > 0,9). As proteínas bZIPs do grupo A de *C. sinensis*, ortólogas aos de *Arabidopsis* estão nomeados na cor vermelha.

A rede de interação entre as proteínas foi estabelecida com alta confiança (score > 0,9), envolvendo as seis CsbZIPs do grupo A e outras 22 proteínas interativas, incluindo proteínas da família PYR/PYL/RCAR, que são receptores citoplasmáticos de ABA, SnRKs (proteínas serina-treonina quinase) e CPKs (proteínas quinases dependentes de cálcio). Estas proteínas estão envolvidas na via de resposta ao estresse abiótico dependente de ABA (Banerjee e Roychoudhury, 2017).

Wang *et al.* (2018) ao analisar a rede de interação envolvendo 24 bZIPs de gergelim (SibZIPs) e outras 10 proteínas interativas, observaram a interação de oito SibZIPs do grupo

A com proteínas que participam da via de sinalização de ABA, como SnRKs e PYL Os autores consideram estes resultados importantes para uma melhor compreensão da função das SibZIPs com base em interações proteína-proteína experimentalmente validadas.

Sabendo que o estresse abiótico, principalmente a seca, representa o principal problema entre os citricultores e com base nestes resultados, podemos supor que os membros do grupo A de laranja doce podem ser considerados alvos potenciais a serem explorados no melhoramento genético de citros.

5.2. Caracterização da estrutura das proteínas CsbZIP

Os parâmetros físico-químicos das proteínas CsbZIP foram preditos e encontram-se disponíveis na Tabela 1, incluindo localização cromossômica dos genes que codificam as CsbZIPs, peso molecular, ponto isoelétrico (pI), índice de hidropaticidade (GRAVY), e localização subcelular.

As proteínas CsbZIP variam de 727 (78,24 KDa) a 140 (16,32 KDa) resíduos de aminoácidos de comprimento, com uma média de 344 aminoácidos (aa). O tamanho das proteínas CsbZIP é equivalente as proteínas bZIP de arroz, com 311 aa em média (Nijhawan *et al.*, 2008), tomate com 318 aa em média (Li *et al.*, 2015) e *Arabidopsis* com 321 aa em média (Jakoby *et al.*, 2002). O ponto isoelétrico das proteínas CsbZIP também variou entre 4,70 (Cs3g27850) a 10,18 (Cs7g25940). De acordo com o resultado referente ao índice de hidropaticidade (GRAVY - *grand average of hydropathicity*), todas as CsbZIPs apresentaram *gravy* negativo, indicando que são proteínas hidrofílicas. A predição da localização subcelular foi realizada e as CsbZIP foram consideradas proteínas nucleares, resultado esperado por se tratar de fatores de transcrição com atuação nuclear (Alberts *et al.*, 2008) (Tabela 1).

Tabela 1. Caracterização físico-química das proteínas CsbZIP codificadas em laranja doce

Grupo	Loco <i>C. sinensis</i> (CAP)	Ortólogos <i>A. thaliana</i>	Localização cromossômica	Peso molecular (kDa)	pI	GRAVY	Localização Subcelular
A	Cs3g26540	At4g35900 - FD	chr3: 27,810,821 – 27.812,224	221 (24.17kDa)	9.35	-0.638	nucl
A	Cs8g06020	At1g49720 - ABF1	chr8: 3,525,102 - 3,530,891	448 (48.55kDa)	9.64	-0.681	nucl
A	Cs3g23480	At1g45249- ABF2/AREB1	chr3: 25,752,151 – 25,756,101	427 (45.51kDa)	9.11	-0.578	nucl
A	Cs3g21410	At1g03970 - GBF4	chr3: 24,321,477 - 24,325,272	245 (27.19kDa)	5.21	-0.589	nucl
A	Cs6g16070	At3g56850 - AREB3	chr6: 17,007,268 - 17,009,944	322 (36.02kDa)	6.41	-0.859	nucl
A	Cs6g14970	At3g44460 - DPBF2	chr6: 16,096,707- 16,097,876	456 (49.60kDa)	8.87	-0.740	nucl
B	Cs7g05140	At5g11260 - HY5	chr7: 2,692,955 - 2,695,598	168 (18.48kDa)	9.56	-1.146	nucl
C	Cs7g07550	At5g24800 - BZO2H2	chr7: 4,591,494 - 4,595,742	401 (43.57kDa)	6.27	-0.724	nucl
C	Cs3g10860	At5g28770 - BZO2H3	chr3: 15,091,515 - 15,097,658	466 (50.50kDa)	9.20	-0.750	nucl
D	Cs8g15030	At1g08320 - TGA9	chr8: 18,146,349 - 18,153,047	508 (56.52kDa)	6.99	-0.573	nucl
D	Cs7g13010	At1g68640 - TGA8/PAN	chr7: 8,956,683 - 8,961,880	357 (39.51kDa)	7.17	-0.403	nucl
D	Cs1g02310	At5g10030 - TGA4/OBF4	chr1: 1,637,432 - 1,643,794	359 (40.82kDa)	7.07	-0.487	nucl
D	Cs7g19070	At5g06839 - TGA10	chr7: 15,041,228 - 15,046,841	332 (36.79kDa)	7.69	-0.492	nucl
E	orange1.1t04546	At2g42380	chrUn: 71,851,340- 71,853,925	363 (40.90kDa)	6.51	-1.064	nucl
E	Cs3g06150	At1g58110	chr3: 8,195,074 - 8,199,483	374 (41.79kDa)	6.96	-0.901	nucl
E	Cs7g23390	At5g04840	chr7: 22,022,476 - 22,027,873	307 (34.12kDa)	8.57	-0.536	nucl
F	Cs8g02640	At2g16770	chr8: 1,064,352 - 1,067,913	268 (29.21kDa)	6.07	-0.705	nucl
G	Cs3g08880	At2g35530	chr3: 12,618,230 - 12,624,739	453 (49.10kDa)	9.53	-0.789	nucl
G	Cs1g23780	At4g36730 - GBF1	chr1: 26,304,060 - 26,310,197	353 (37.29kDa)	5.21	-0.883	nucl
G	orange1.1t00453	At2g46270 - GBF3	chrUn: 5,757,053 - 5,762,358	421 (44.65kDa)	8.69	-0.749	nucl
H	Cs6g16800	At2g40950	chr6: 17,426,355 - 17,429,782	727 (78.24kDa)	6.62	-0.484	nucl
I	Cs7g29820	At2g40620	chr7: 29,951,237 - 29,954,735	332 (37.03kDa)	5.44	-0.867	nucl
I	Cs3g08170	At4g38900	chr3:11,632,304 - 11,636,660	597 (64.85kDa)	6.49	-0.923	nucl
I	Cs3g25760	At1g06850	chr3: 27,284,952 - 27,287,970	424 (46.79kDa)	6.37	-0.446	nucl
I	Cs5g23040	At1g06070	chr5: 25,774,957 - 25,780,092	426 (46.48kDa)	6.01	-0.823	nucl
J	Cs6g15200	At1g19490	chr6: 16,285,261 - 16,290,793	488 (53.40kDa)	7.61	-0.663	nucl

K	Cs3g27850	At1g42990	chr3: 28,682,837 - 28,685,846	298 (33.48kDa)	4.70	-0.616	nucl
S	Cs3g16310	At1g59530	chr3: 20,172,057 - 20,173,327	140 (16.32kDa)	5.65	-0.789	nucl
S	Cs1g21370	At3g49760	chr1: 24,283,091 - 24,284,613	185 (21.57kDa)	10.01	-0.926	nucl
S	Cs6g19350	At3g30530	chr6: 19,212,187 - 19,212,971	169 (19.49kDa)	5.78	-0.699	nucl
S	Cs8g07470	At1g75390	chr8: 5,367,723 - 5,369,527	172 (19.24kDa)	5.83	-0.623	nucl
S	orange1.1t01674	At3g62420	chrUn: 27,396,492 -27,398,485	145 (16.63kDa)	6.60	-0.694	nucl
S	Cs7g12290	At1g13600	chr7: 8,221,855 - 8,222,749	201 (23.00kDa)	6.30	-0.768	nucl
S	Cs7g25940	At5g60830	chr7: 26,550,258 - 26,550,872	153 (17.29kDa)	10.18	-0.595	nucl

Fatores de transcrição da família bZIP seguem via de regulação ABA-dependente e são ativados via fosforilação, sendo a modificação mais comum para induzir a ativação de FTs bZIPs. Os FTs bZIP são ativados por uma via de sinalização mediada por ABA, e após a percepção desse fitormônio pelo complexo PYL/PYR/RCAR, o sinal é transduzido através de proteínas do tipo Serina/Treonina quinases, como as SnRKs que fosforilam os FTs bZIPs (Banerjee e Roychoudhury, 2017). Com base nessas informações, foi realizada a predição de sítios de fosforilação em resíduos de serina (S), treonina (T) e tirosina (Y) por meio do *software* NetPhos 3.1 (Blom, N. *et al.*, 1999) nas 34 proteínas CsbZIPs de laranja doce.

O resultado obtido condiz com o esperado para fatores de transcrição bZIP, fosforilados por Serina/Treonina quinases (Figura 9). Todas as proteínas CsbZIP apresentaram sítios de fosforilação em resíduos de serina (S) e treonina (T). Os genes Cs6g16800 e Cs3g08170 possuem 72 e 68 sítios de fosforilação em resíduos de serina, respectivamente. Os genes Cs3g10860 e Cs6g14970 apresentam 20 e 18 sítios de fosforilação em resíduos de treonina. Seis proteínas, Cs3g26540, Cs3g16310, Cs1g21370, Cs6g19350, Cs7g12290 e Cs7g25940, não apresentaram sítios de fosforilação em resíduos de tirosina (Y). Maiores informações sobre os sítios de fosforilação encontram-se no Apêndice 3.

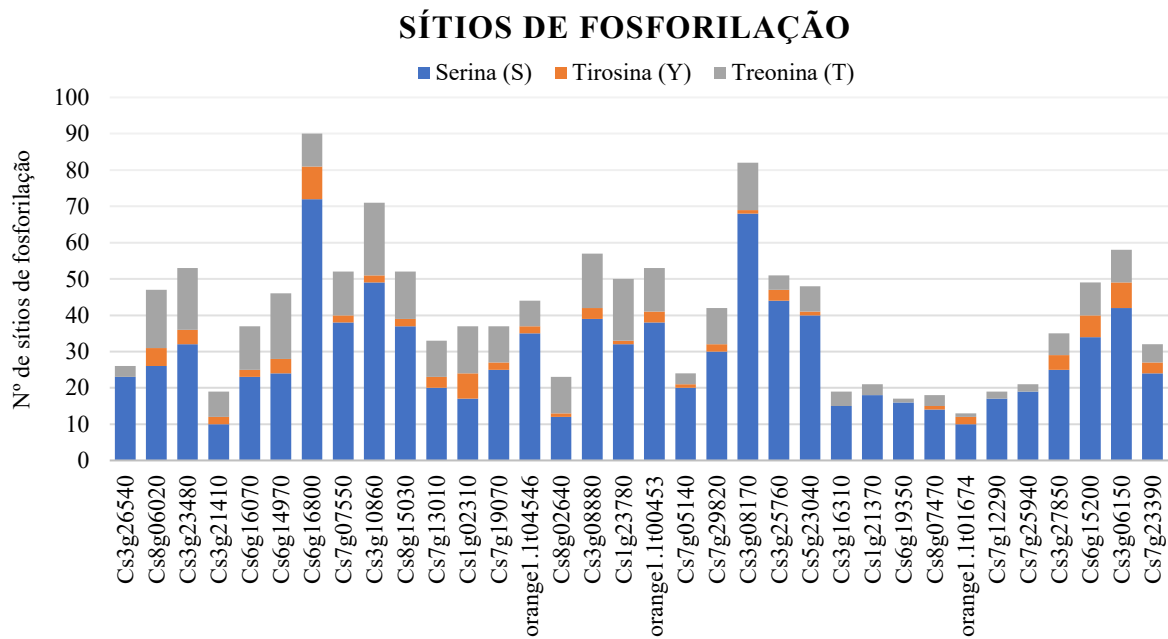


Figura 9. Predição de sítios de fosforilação. Sítios de fosforilação em resíduos de serina (S), treonina (T) e tirosina (Y) nas 34 proteínas CsbZIPs de *C. sinensis*. Análise realizada por meio do *software* NetPhos3.1.

Para investigar a diversidade estrutural e prever funções das proteínas CsbZIP, um total de 20 motivos conservados foram preditos por meio do *software* MEME (Figura 10). Assim como nas análises realizadas por Wang *et al.* (2018), que predisseram 20 motivos conservados nas 63 proteínas bZIPs de gergelim (SibZIP), resultado que contribuiu para a compreensão das relações evolutivas e padrões de conservação dos motivos nas proteínas bZIPs.

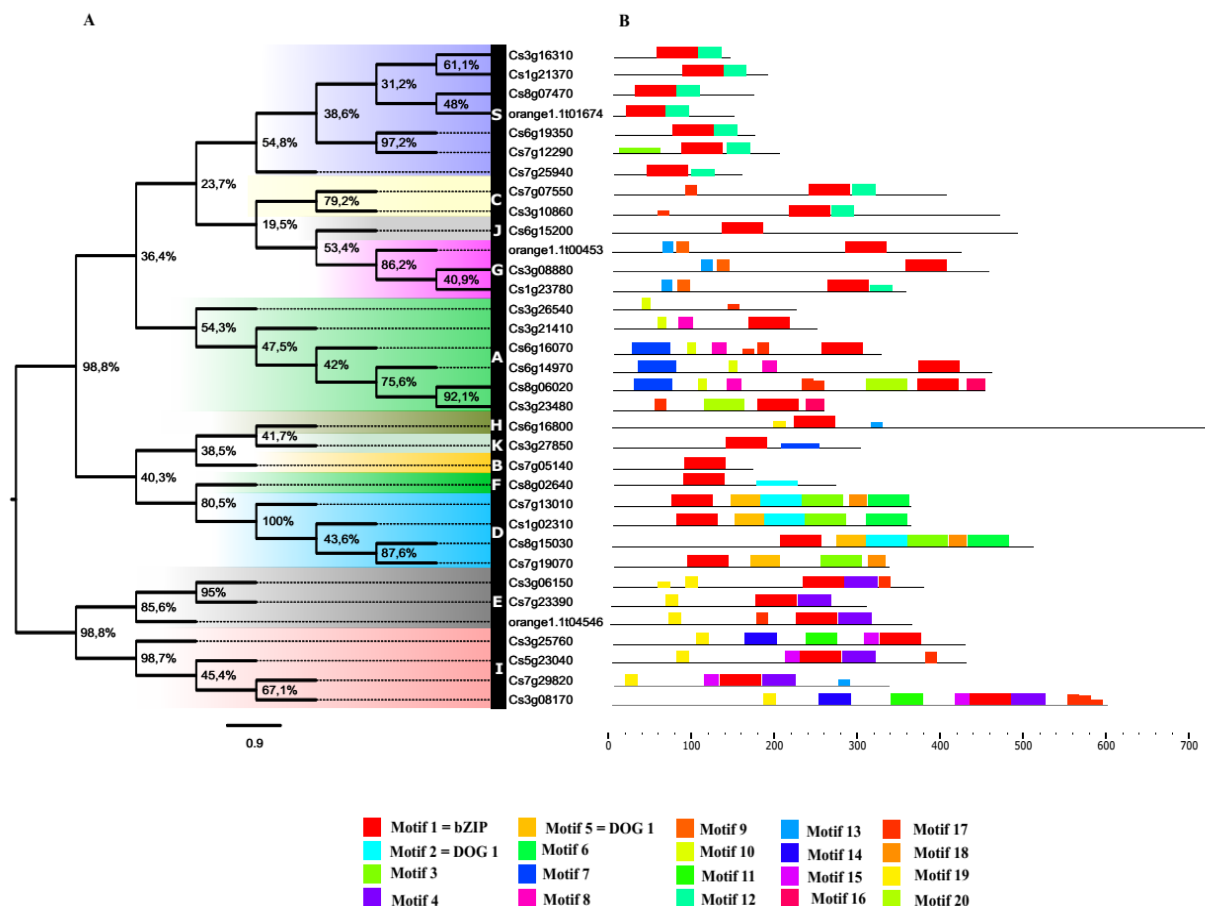


Figura 10. Relação filogenética e motivos conservados das proteínas CsbZIPs. A) Árvore filogenética produzida no software MEGA 6, usando o método de agrupamento Neighbor-Joining com réplicas de 1000 bootstraps; B) Representação esquemática de 20 motivos conservados nas proteínas CsbZIPs identificados pelo software MEME. Cada motivo é indicado por uma cor. Os detalhes individuais de cada motivo encontram-se disponíveis no Apêndice 4.

Os motivos preditos para as 34 CsbZIPs foram identificados por meio do banco de dados de domínio conservado (CDD) do NCBI. Apenas três motivos foram identificados, sendo o motivo 1 anotado como o domínio bZIP, caracterizado pela sequência invariante: N-x₇-R/K-x₉-L-x₆-L-x₆-L, e os motivos 2 e 5 foram anotados como o domínio DOG1 (*Delay of Germination* 1). Sabe-se que DOG1 está envolvido na regulação da dormência de sementes em resposta à temperatura e também tem sido associado com fenótipos de florescimento em *Arabidopsis*. No entanto, seu mecanismo de ação ainda não foi elucidado (Bentsink *et al.*, 2006; Huo *et al.*, 2016).

É possível notar que a composição de motivos das proteínas CsbZIP de cada grupo é semelhante e alguns motivos sempre aparecem associados a outros, como o motivo 1 e 12, facilmente observados no grupo S. O motivo 1, trata-se do domínio bZIP e o motivo 12, que não foi anotado, é composto por uma repetição de leucinas a cada sete aminoácidos, o que o caracteriza como uma continuação do zíper de leucina comum ao domínio bZIP. O motivo 8

está associado ao motivo 10, presentes nas proteínas do grupo A, e o motivo 19 é característico das proteínas dos grupos E e I, que estão em clados próximos. Sendo assim, o compartilhamento de motivos conservados faz-se útil para dar suporte à classificação em grupos, proposta para os fatores de transcrição bZIP de laranja doce.

Como esperado, todas as proteínas CsbZIPs apresentaram o motivo 1, correspondente ao domínio bZIP, exceto a proteína codificada pelo gene Cs3g26540, que após uma análise da sua sequência proteica, verificou-se a presença do motivo região básica (N-x₇-R/K-x₉) e de um resíduo de prolina no início do motivo zíper de leucina (Figura 11), que é responsável pela dimerização. A proteína ortóloga de *A. thaliana* (At4g35900) apresenta o domínio bZIP completo e não possui os resíduos de prolina. Sendo assim, caso a proteína codificada por Cs3g26540 seja funcional, pode não desempenhar funções similares a de outros bZIPs conhecidos.

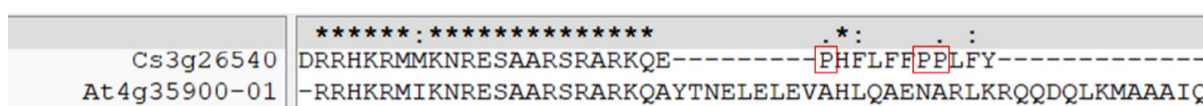


Figura 11. Alinhamento da região do domínio bZIP codificado pelo loco Cs3g26540 de *C. sinensis* e seu respectivo ortólogo em *A. thaliana* (AT4G35900). Os resíduos de prolina no início do motivo zíper de leucina estão destacados pelos retângulos vermelhos.

Dois genes de *Arabidopsis* que codificam para fatores de transcrição bZIP (*AtbZIP34* e *AtbZIP61*) apresentam um resíduo de prolina na terceira repetição de leucinas do motivo zíper de leucina, de acordo com a descrição de Shen *et al.* (2007). No entanto, esses genes são incapazes de homodimerizar para atuar na regulação da expressão gênica (Shen *et al.* 2007).

Buscando compreender melhor essa incapacidade das AtbZIPs 34 e 61 que possuem o resíduo de prolina em formar homodímeros, Shen *et al.* (2007) obtiveram dois mutantes, denominados AtbZIP34m e AtbZIP61m, cujo resíduo de prolina foi substituído por uma alanina e puderam observar que os mutantes foram capazes de formar homodímeros. Desse modo, evidenciaram que o resíduo de prolina da região do zíper de leucina interfere na formação de homodímeros, por distorcer a estrutura em α -hélice dos monômeros (Shen *et al.*, 2007). Com base nestas informações, pode-se inferir que o gene Cs3g26540 de laranja doce, classificado no grupo A e ortólogo da proteína FD de *Arabidopsis*, um fator de transcrição conhecido por promover a floração (Abe *et al.*, 2005), provavelmente codifica uma proteína funcional, capaz de heterodimerizar e atuar na regulação da expressão gênica.

5.3. Padrões de expressão dos genes CsbZIP em diferentes tecidos

Para compreender os padrões transcricionais e os possíveis papéis dos genes CsbZIP, os dados de RNA-seq de quatro tecidos (folha, flor, fruto e calo) foram obtidos no banco de dados do CAP (*Citrus sinensis* Annotation Project) (Wang *et al.*, 2014). Em seguida gerou-se o *heatmap* baseado nos valores de RPKM de cada gene (Figura 12). É possível observar a formação de três *clusters* que apresentam padrão de expressão diverso entre si. O *cluster* 1 agrupou 12 genes com maiores níveis de expressão nas folhas e frutos; o *cluster* 2 agrupou 11 genes com maiores níveis de expressão nas flores; e o *cluster* 3 também agrupou 11 genes, porém com maiores níveis de expressão em calos.

Todos os três genes do grupo E encontram-se presentes no *cluster* 1, com maiores níveis de expressão nos tecidos foliares e nas flores. No entanto, os genes Cs3g06150 e Cs7g23390 do grupo E apresentaram também alto nível de expressão nos tecidos de calos. Os membros do grupo E de *Arabidopsis* ainda não foram caracterizados funcionalmente, apenas o AtbZIP34 foi parcialmente caracterizado, apresentando ligação com a germinação do grão de pólen e o crescimento do tubo polínico (Giblová *et al.*, 2009; Dröge-Laser *et al.*, 2018).

Os genes dos grupos C e S apresentaram altos níveis de expressão nos tecidos foliares, nas flores e frutos, e dispõem-se nos *clusters* 1 e 2. Esses genes apresentam heterodimerização preferencial entre si, conhecidos como a rede C/S₁, e atuam regulando a resposta à deficiência nutricional (Dröge-Laser e Weiste, 2018). Os genes do grupo I apresentaram maior nível de expressão nos tecidos de calos, de acordo com a sua atuação na formação de calos induzido por auxina, função desempenhada pelo gene AtbZIP59 membro do grupo I de *A. thaliana* (Xu *et al.*, 2018).

Membros dos grupos A e D foram observados nos três *clusters*, logo, apresentaram altos níveis de expressão nos quatro tecidos analisados (folha, flores, frutos e calos). Os genes do grupo D de *C. sinensis* são ortólogos aos fatores TGA de *Arabidopsis*. Jakoby *et al.* (2002) descrevem que os genes do grupo D de *Arabidopsis* participam de dois processos biológicos diferentes: defesa contra patógenos e desenvolvimento. Em resposta ao ataque de patógenos, o ácido salicílico induz a expressão de genes relacionados à patogênese (*PRs*) por toda a planta e acredita-se que os fatores TGA regulam a indução sistêmica por interagir com *cis*-elementos presentes na região promotora dos genes *PRs*. Além disso, alguns fatores TGA controlam a formação de órgãos florais e já foi relatada função relacionada à captação de nitrato das raízes (Jakoby *et al.*, 2002; Dröge-Laser *et al.*, 2018). A maioria das bZIPs desse grupo aumentam a

capacidade da planta em sobreviver a agentes patogênicos e estresses xenobióticos (Dröge-Laser *et al.*, 2018).

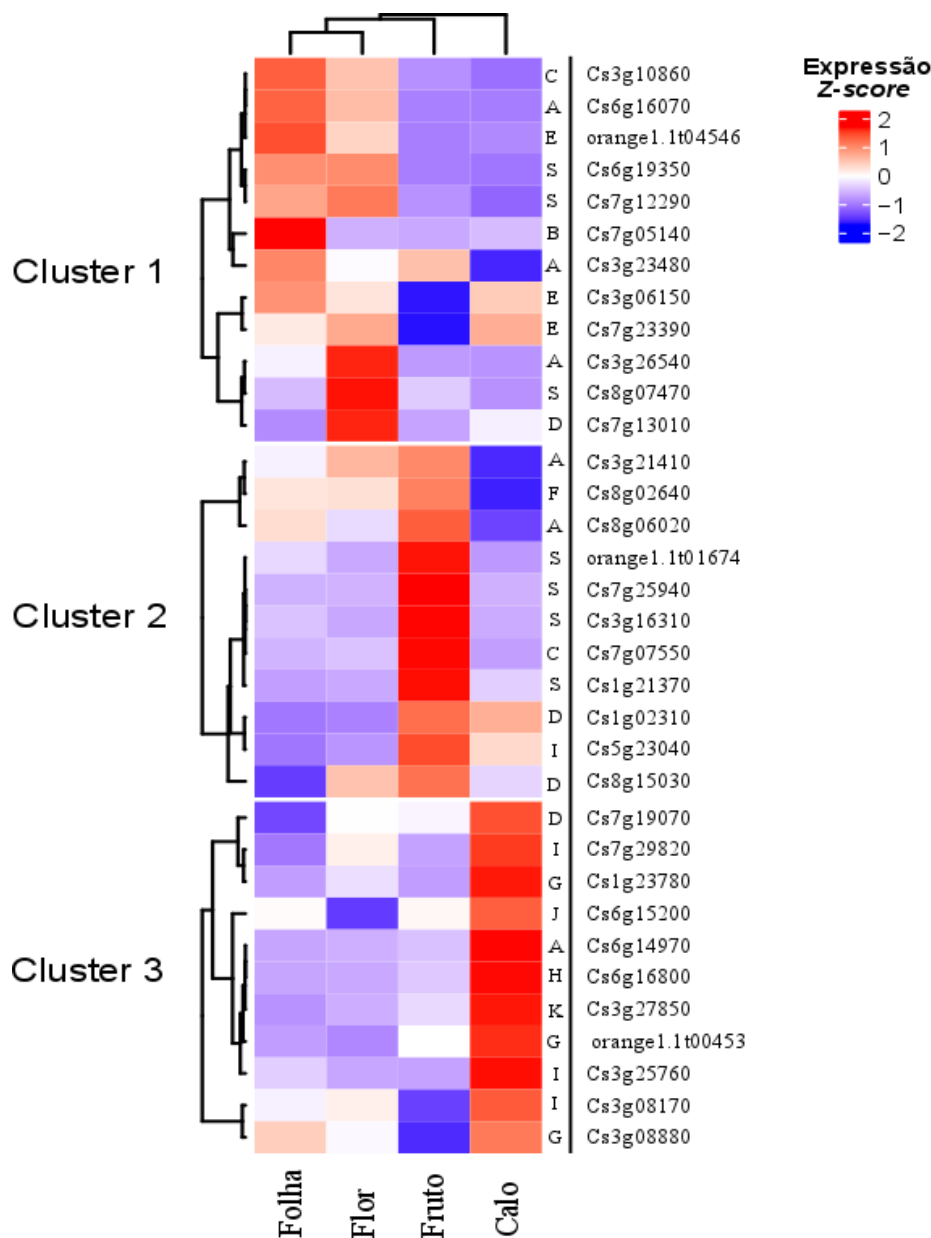


Figura 12. Perfil de expressão dos genes *CsbZIPs* em diferentes tecidos de laranja doce. Dados de RNA-seq de quatro tecidos (folha, flor, fruto, calo) usados para analisar o padrão de expressão dos genes *CsbZIPs*. O nível de expressão é exibido de acordo com a escala de cores. As letras posicionadas a esquerda do nome dos genes *CsbZIPs* indicam os grupos que cada gene pertence, com base na classificação proposta.

Os genes do grupo A, como já relatado, incluem os fatores de ligação aos elementos responsivos ao ABA (ABFs) que atuam modulando a expressão de genes relacionados com o aumento da tolerância/resistência a estresses abióticos, como seca e salinidade (Dröge-Laser *et al.*, 2018). Com base nestas informações e nos resultados obtidos, a presença de genes dos

grupos A e D nos três *clusters*, ou seja, sendo expressos em folhas, flores, frutos e calos é compreensível, pois o ataque por patógenos e os estresses abióticos não se limitam a apenas um tecido específico. Todos os tecidos vegetais podem ser afetados e consequentemente devem desencadear respostas à tais eventos de estresse.

5.4. Expressão gênica de dois *CsbZIPs* submetidos a condições de déficit hídrico

Para explorar o possível envolvimento dos genes *CsbZIPs* com a resposta à deficiência hídrica, foi analisado o perfil de expressão de dois genes em plantas submetidas ao déficit hídrico, provenientes do experimento de doutorado de Matheus Carvalho Silva (Silva, 2018). Os genes escolhidos para a análise de expressão foram o Cs8g06020 e o Cs3g23480, pois são genes do grupo A, ortólogos aos genes *ABF1* e *ABF2* de *Arabidopsis thaliana*.

Os genes *ABF* (fatores de ligação ao *cis*-elemento ABRE) de *Arabidopsis* estão reconhecidamente envolvidos com a resposta ao estresse hídrico, pois codificam fatores de transcrição bZIP ativados por serina/treoninas quinases como SnRKs que interagem com *cis*-elementos responsivos ao ABA, como os ABREs presente na região promotora de genes induzidos por ABA (Yoshida *et al.*, 2015).

Com base nos resultados, observa-se que os genes foram expressos nas plantas expostas ao estresse moderado e apresentaram expressão reprimida nas plantas expostas à condição de reidratação, quando comparadas às plantas do tratamento controle. No tratamento de estresse severo, a expressão do gene *Cs8g06020* (*ABF1*) foi reprimida, enquanto o gene *Cs3g23480* (*ABF2*) apresentou expressão em relação ao controle (Figura 13).

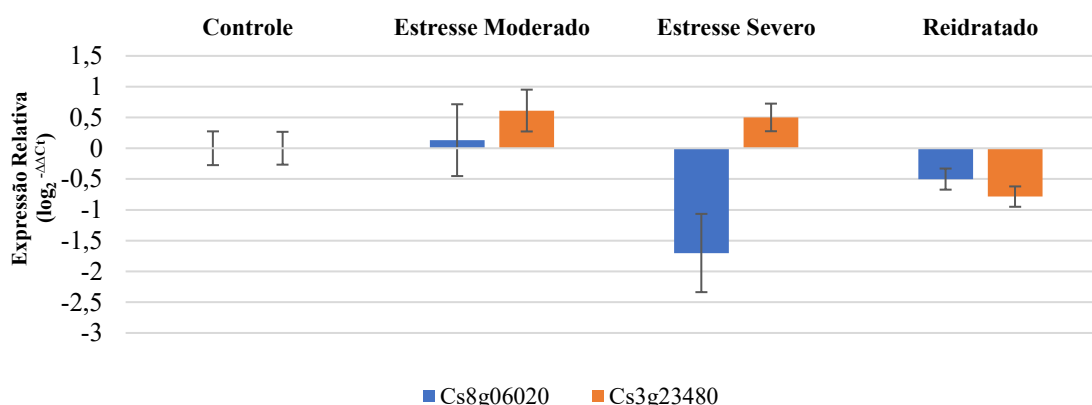


Figura 13. Perfil de expressão dos genes Cs8g06020 e Cs3g23480 de *C. sinensis* em resposta ao déficit hídrico. Razões ($\log_2^{-\Delta\Delta C_t}$) referentes aos níveis de mRNA entre plantas submetidas ao déficit hídrico e controles para dois genes *CsbZIP* (*Cs8g06020* e *Cs3g23480*) via qPCR. O gene utilizado como controle endógeno foi o *GAPC2*.

No entanto, diferenças entre os níveis de expressão do controle em relação aos tratamentos de estresse hídrico (estresse moderado, severo e reidratado) não foram consideradas significativas com base no teste de média aplicado ($p\text{-value} \leq 0,05$). Os baixos níveis de expressão devem-se, possivelmente, aos níveis basais de expressão que normalmente estes genes apresentam, como foi observado por Yoshida *et al.* (2015), que relataram que o gene *ABF1* de *A. thaliana* foi claramente induzido por tratamentos de seca, alta salinidade e ABA, no entanto, seus níveis de expressão foram considerados baixos mesmo sob condições de estresse.

Considerando a possível atuação dos genes Cs8g06020 (*ABF1*) e Cs3g23480 (*ABF2*) nas vias de resposta ao estresse abiótico, uma potencial abordagem a ser considerada no melhoramento de citros é a superexpressão desses genes sob regulação de um promotor constitutivo.

Fujita *et al.* (2005) consideraram a expressão do gene *ABF2* em *Arabidopsis* insuficiente para induzir a expressão de genes responsivos ao estresse; porém, quando este gene foi superexpresso sob controle de promotor constitutivo (35S), as plantas de *A. thaliana* apresentaram maior sensibilidade ao ABA e maior tolerância a seca. Além disso, o gene *ABF2* foi responsável por regular a expressão de oito genes estresse responsivos com dois ou mais *cis*-elementos ABRE em suas regiões promotoras.

6. CONCLUSÃO

O genoma de laranja doce codifica para uma família relativamente pequena de fatores de transcrição da classe bZIP, em comparação àquelas descritas para outras espécies vegetais. As análises de motivos conservados das proteínas CsbZIPs e da estrutura gênica foram utilizadas para dar suporte à classificação proposta no presente estudo. Os membros do grupo A podem ser considerados alvos potenciais para futuros estudos funcionais, por serem ortólogos a genes de *Arabidopsis* responsivos ao estresse abiótico e por interagirem com proteínas específicas da via de resposta dependente de ABA. Os resultados obtidos servirão como base para estudos futuros que visam aumentar a produtividade citrícola ao conferir resistência ao estresse por meio da modulação da expressão de proteínas bZIP específicas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Abe, M. et al. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot Apex. **Science**, v. 309, n. 9, p. 1052-1057, 2005.
2. Agarwal, P. K.; Jha, B. Transcription factors in plants and ABA dependent and independent abiotic stress signalling. **Biologia Plantarum**, v. 54, p. 201–212, 2010.
3. Alberts, B. *et al.* **Molecular biology of the cell**. 6^a ed. New York: Garland Science Taylor & Francis Group, 2008.
4. Augustine, S. M. *et al.* Overexpression of *EaDREB2* and pyramiding of *EaDREB2* with the pea DNA helicase gene (PDH45) enhance drought and salinity tolerance in sugarcane (*Saccharum* spp. hybrid). **Plant Cell Reports**, v. 34, n. 2, p. 247-263, 2014.
5. Azevêdo, C. L. Sistemas de Produção de Citros para o Nordeste. **Embrapa Mandioca e Fruticultura** - Versão eletrônica, ISSN: 1678-8796. Disponível em: <https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Citros/CitrosNordeste/clicli.html>; 2003.
6. Banerjee, A.; Roychoudhury, A. Absciscic-acid-dependent basic Leucine zipper (bZIP) transcription factors in plant abiotic stress. **Protoplasma**, v. 254, n. 1, p. 3-16, 2017.
7. Bentsink L., Jowett J., Hanhart CJ, Koornneef M. Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. **Plant Biology**, v.103, n. 45, p. 17042-17047, 2006.
8. Blom, N.; *et al.* Sequence and structure based prediction of eukaryotic protein phosphorylation sites. **Journal of Molecular Biology**, v. 294, n. 5, p. 1351-1362, 1999.
9. Côrrea, L. G. G. *et al.* The role of bZIP transcription factors in green plant evolution: Adaptive features emerging from four founder genes. **Plos One**, v. 3, n. 8, 2008.
10. Dröge-Laser, W., *et al.* The *Arabidopsis* bZIP transcription factor Family – an update. **Current Opinion in Plant Biology**, v. 45, p. 36-49, 2018.
11. Dröge-Laser, W.; Weiste, C. The C/S1 bZIP network: A regulatory hub orchestrating plant energy homeostasis. **Trends in Plant Science**, v. 23, n. 5, p. 422-433, 2018.
12. Estrada-Melo, A. *et al.* Overexpression of an ABA biosynthesis gene using a stress-inducible promoter enhances drought resistance in petunia. **Horticulture Research**, v. 2, n. 4, 2015.
13. Gibalová, A. *et al.* AtbZIP34 is required for *Arabidopsis* pollen wall patterning and the control of several metabolic pathways in developing pollen. **Plant Molecular Biology**, v. 70, p. 581-601, 2009.

14. Gong, Y. *et al.* Abiotic stress in plants. **Agricultural Chemistry**, p. 113-152, 2013.
15. Gorlova, O. *et al.* Genes with a large intronic burden show greater evolutionary conservation on the protein level. **BMC Evolutionary Biology**, v. 14, p. 1-7, 2014.
16. Gu, Z. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. **Bioinformatics**, 2016.
17. Härma, H. *et al.* Toward universal protein post-translational modification detection in high throughput format. **Chemical Communications**, v. 54, n. 23, p. 2910-2913, 2018.
18. Higo, K. *et al.* Plant cis-acting regulatory DNA elements (PLACE) database. **Nucleic Acids Research**, v. 27, n. 1, p.297-300, 1999.
19. Hu, W. *et al.* Genome-wide characterization and analysis of bZIP transcription factor gene family related to abiotic stress in cassava. **Scientific Reports**, v. 6, n. 1, 2016a.
20. Hu, W. *et al.* Genome-wide analyses of the bZIP Family reveal their involvement in the development, ripening and abiotic stress response in banana. **Scientific Reports**, v. 6, n. 7, 2016b.
21. Huang, X. *et al.* Overexpression of PtrABF gene, a bZIP transcription factor isolated from *Poncirus trifoliata*, enhances dehydration and drought tolerance in tobacco via scavenging ROS and modulating expression of stress-responsive genes. **BMC Plant Biology**, v. 10, n. 1, p. 230, 2010.
22. Huo, H. Delay of Germination1 (DOG1) regulates both seed dormancy and flowering time through microRNA pathways. **Proceedings of the National Academy of Sciences of the USA (PNAS)**, v. 1, n. 26, p. 1-8, 2016.
23. IBGE. **Instituto Brasileiro de Geografia e Estatística**. Disponível em: <https://www.ibge.gov.br>. Acesso em: novembro, 2018.
24. Jakoby, M. *et al.* bZIP transcription factors in *Arabidopsis*. **Trends in Plant Science**, v. 7, n. 3, p. 106-111, 2002.
25. Khan, P. *et al.* Abiotic stress tolerance in Plants: Insights from Proteomics. **Emerging Technologies and Management of Crop Stress Tolerance**, v. 2, p. 23-68, 2014.
26. Khan, S. *et al.* Revisiting the role of plant transcription factors in the battle against abiotic stress. **International Journal of Molecular Sciences**, v. 19, n. 6, 2018.
27. Nelson, D.; Cox, M. **Lehninger Principles of Biochemistry**. 6th edition. W. H. Freeman and Company, New York, 2012.

28. Lindemose, S. *et al.* Structure, function and networks of transcription factors involved in abiotic stress responses. **International Journal of Molecular Sciences**, v. 14, n. 3, p. 5842-5878, 2013.
29. Llorca, C. M. *et al.* bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. **Frontiers in Plant Science**, v. 5, n. 4, p. 1-14, 2014.
30. Li, D. *et al.* Genome-wide systematic characterization of the bZIP transcriptional factor family in tomato (*Solanum lycopersicum* L.). **BMC Genomics**, v. 16, n. 1, p. 771, 2015.
31. Liang, C. *et al.* GhABF2, a bZIP transcription factor, confers drought and salinity tolerance in cotton (*Gossypium hirsutum* L.). **Scientific Reporter**, v. 6, n. 35040, p. 1-14, 2016.
32. Liu, C. *et al.* OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. **Plant Molecular Biology**, v. 84, p. 19-36, 2014a.
33. Liu, H. *et al.* Genome-wide analysis and Evolution of the bZIP transcription factor gene Family in six *Fragaria* species. **Plant Systematics Evolution**, p. 1-13, 2017.
34. Liu, J. *et al.* Genome-wide analysis and expression. Profile of the bZIP transcription factor gene Family in grapevine (*Vitis vinifera*). **BMC Genomics**, v. 15, n. 1, p. 1- 18, 2014b.
35. Livak, K. J.; e Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method. **Methods**, v. 25, p. 402-408, 2001.
36. Mafra, V; et al. Reference Genes for Accurate transcript normalization in *Citrus* genotypes under different experimental conditions. **Plos One**, v. 7, p. 1-11, 2012.
37. Neves, M. F. *et al.* **O retrato da citricultura brasileira**. Ribeirão Preto: Markestrat (Centro de Pesquisa e Projetos em Marketing e Estratégia - Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo), p. 138, 2010.
38. Nijhawan, A. *et al.* Genomic survey and gene expression. Analysis of the basic Leucine zipper transcription factor Family in rice. **Plant Physiology**, v. 146, n. 2, p. 333-350, 2008.
39. Oliveira, A. C. Ácido abscísico e o estresse abiótico. Pelotas: **Embrapa Clima Temperado - Documentos**, 307; p. 30, 2010.
40. Pandey, A. S. *et al.* A rice bZIP transcription factor, *OsbZIP16*, regulates abiotic stress tolerance when over-expressed in *Arabidopsis*. **Journal of Plant Biochemistry and Biotechnology**, v. 8, n. 4, p. 393-400, 2018.
41. Pruneda-Paz, J. *et al.* A genome-scale resource for the functional characterization of *Arabidopsis* transcription factors. **Cell Reports**, v. 8, n. 2, p. 622-632, 2014.

42. Riechmann, J. L. *et al.* *Arabidopsis* transcription factors: Genome-wide comparative analysis among Eukaryotes. **Science**, v. 290, p. 2105-2110, 2000.
43. Saitou, N.; Nei, M. The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p.406- 425, 1987.
44. Sharma *et al.* Converting citrus wastes into value-added products: Economic and environmently friendly approaches. **Nutrition**, v. 34, p. 29-46, 2017.
45. Shen, H. *et al.* A conserved proline residue in the Leucine zipper region of *AtbZIP34* and *AtbZIP61* in *Arabidopsis thaliana* interferes whit the formation of homodimer. **Biochemical and Biophysical Research Communications**, v. 362, n. 2, p. 425-430, 2007.
46. Silva, M. de C. **Respostas fisiológicas de laranjeira ‘valencia’ sob diferentes porta-enxertos cítricos em função da variação da umidade do solo.** Tese (Programa de Pós-Graduação em Produção Vegetal) – Universidade Estadual de Santa Cruz (UESC), Ilhéus – BA, 2018.
47. Sornaraj, P. *et al.* Basic leucine zipper (bZIP) transcription factors involved in abiotic stresses: A molecular model of a wheat bZIP factor and implications of its structure in function. **Biochimica et Biophysica Acta – General Subjects**, v. 1860, n. 1, p. 46-56, 2016.
48. Tamura, K., *et al.* MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725-2729, 2013.
49. Thompson, J. D. *et al.* CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p. 4673-4680, 1994.
50. United States Department of Agriculture (USDA). Citrus: Word markets and Trade. Disponível em: <https://www.fas.usda.gov/data/citrus-world-markets-and-trade>. 2018.
51. Uno, Y. *et al.* *Arabidopsis* basic Leucine zipper transcription factors involved in na abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. **Proceedings of the National Academy of Sciences of the USA (PNAS)**, v. 97, p. 11632–11637, 2000.
52. Vishwakarma, K. *et al.* Absciscic acid signaling and abiotic stress tolerance in plants: A review on current knowledge and future prospects. **Frontiers in Plant Science**, v. 8, n. 2, p. 1-12, 2017.
53. Voorrips, R.E. MapChart: Software for the graphical presentation of linkage maps and QTLs. **The Journal of Heredity**, v. 93, n. 1, p. 77-78, 2002.
54. Wang, J. *et al.* *Citrus sinensis* Annotation Project (CAP): a comprehensive database for sweet orange genome. **PLoS One**, v. 9, n. 1, p. 1-7, 2014.

55. Wang, H. *et al.* Recent advances in utilizing transcription factors to improve plant abiotic stress tolerance by transgenic technology. **Frontiers in Plant Science**, v. 7, n. 2, p. 1-13, 2016.
56. Wang, Y. *et al.* Identification and characterization of the bZIP transcription factor Family and its expression in response to abiotic stresses in sesame. **PLoS One**, v. 13, n. 7, 2018.
57. Wei, K. *et al.* Genome-wide analysis of bZIP-encoding genes in maize. **DNA Research**, v. 19, n. 6, p. 463-476, 2012.
58. Wu, M. *et al.* A moso bamboo WRKY gene *PeWRKY83* confers salinity tolerance in transgenic *Arabidopsis* plants. **Scientific Reports**, v. 7, n. 1, p. 1-16, 2017.
59. Wu, G. A. Genomics of the origin and evolution of Citrus. **Nature**, v. 554, n. 7692, p. 311-316, 2018.
60. Xu, Q. *et al.* The draft genome of sweet Orange (*Citrus sinensis*). **Nature Genetics**, v. 45, n. 1, p. 59-66, 2013.
61. Xu, C., *et al.* Control of auxin-induced callus formation by bZIP59– LBD complex in *Arabidopsis* regeneration. **Nature Plants**, v. 4, p. 108-115, 2018.
62. Yamaguchi-Shinozaki, K.; Shinozaki, K. Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. **Trends in Plant Science**, v. 10, n. 2, p. 88-94, 2005.
63. Yoshida, T. *et al.* Four *Arabidopsis* AREB/ABF transcription factors function predominantly in gene expression downstream of SnRK2 kinases in abscisic acid signalling in response to osmotic stress. **Plant, Cell and Environment**, v. 38, n. 1, p. 35-49, 2015.
64. Zaha, A.; Ferreira, H. B.; Passaglia, L. M. P. (Org.). **Biologia molecular básica**. 5. ed. Porto Alegre: ArtMed, 2014.
65. Zhang, L. *et al.* A novel wheat bZIP transcription factor, TabZIP60, confers multiple abiotic stress tolerances in transgenic *Arabidopsis*. **Physiologia Plantarum**, v.153, p. 538–554, 2015.
66. Zhang, X. *et al.* Maize *ABP9* enhances tolerance to multiple stresses in transgenic *Arabidopsis* by modulating ABA signaling and cellular levels of reactive oxygen species. **Plant Molecular Biology**, v. 75, p.365–378, 2011.
67. Zhang, Z. *et al.* Roles of the bZIP gene Family in rice. **Genetics and Molecular Research**, v. 13, n. 2, p. 3025-3036, 2014.

APÊNDICES

Apêndice 1 - Oligonucleotídeos empregados nas análises de qRT-PCR.

Gene		Sequência
Cs8g06020*	F	5' - CTGTGTTGGGTGAAAAATTGG - 3'
	R	3' - CTTCACAACCATCCAAACTGAA - 5'
Cs3g23480**	F	5' - CAGCAGACTTTGGGGGAAAT - 3'
	R	3' - ACCACCACCAAACTGCAAAT - 5'
GAPC2	F	5'-TCTTGCCTGCTTTGAATGGA - 3'
	R	5' - TGTGAGGTCAACCACTGCGACAT - 3'

* *CsbZIP* ortólogo a *ABF1* de *Arabidopsis thaliana*;

** *CsbZIP* ortólogo a *ABF2* de *Arabidopsis thaliana*.

Apêndice 3. Sítios de fosforilação preditos para as proteínas codificadas pelos genes *CsbZIPs*.

Sítios de Fosforilação					
Gene	Tamanho do polipeptídeo	Serina (S)	Tirosina (Y)	Treonina (T)	Quantidade de sítios de fosforilação
Cs3g26540	221	14, 15, 16, 18, 20, 21, 22, 23, 25, 26, 28, 29, 37, 46, 49, 91, 125, 145, 159, 172, 173, 200, 204	–	32, 62, 187	23 ^S , 3 ^T
Cs8g06020	448	3, 14, 31, 34, 53, 82, 94, 102, 129, 163, 176, 200, 201, 204, 208, 219, 269, 281, 295, 327, 333, 340, 345, 346, 360, 384	7, 15, 170, 246, 352	13, 36, 43, 65, 96, 100, 105, 137, 142, 264, 290, 325, 341, 392, 443, 445	26 ^S , 5 ^Y , 16 ^T
Cs3g23480	427	4, 26, 29, 37, 38, 45, 57, 86, 94, 111, 119, 126, 128, 130, 181, 188, 207, 214, 229, 268, 293, 299, 306, 310, 311, 312, 314, 316, 329, 349, 353, 424	171, 226, 318, 367	31, 61, 86, 88, 92, 135, 140, 178, 224, 225, 246, 296, 331, 361, 365, 372, 422	32 ^S , 4 ^Y , 17 ^T
Cs3g21410	245	17, 35, 49, 65, 95, 114, 127, 176, 180, 242	187, 217	13, 14, 18, 29, 81, 110, 213,	10 ^S , 2 ^Y , 7 ^T
Cs6g16070	322	8, 21, 27, 38, 49, 58, 73, 79, 87, 103, 117, 161, 173, 201, 222, 224, 225, 231, 243, 264, 268, 284, 319	275, 312	3, 5, 32, 81, 90, 116, 122, 138, 155, 235, 250, 276	23 ^S , 2 ^Y , 12 ^T
Cs6g14970	456	10, 34, 35, 51, 57, 66, 118, 124, 126, 130, 147, 161, 162, 224, 271, 282, 286, 297, 326, 329, 342, 382, 386, 453	255, 315, 393, 435	28, 40, 47, 65, 80, 109, 114, 132, 141, 179, 184, 200, 201, 236, 288, 311, 394, 436	24 ^S , 4 ^Y , 18 ^T
Cs6g16800	727	2, 13, 31, 39, 40, 64, 73, 85, 96, 99, 122, 126, 131, 135, 139, 143, 146, 148, 151, 152, 159, 177, 203, 204, 205, 207, 234, 238, 265, 268, 276, 282, 284, 323, 344, 349, 354, 356, 359, 366, 398, 404, 424, 427, 451, 456, 462, 472, 478, 498, 502, 508, 512, 516, 529, 545, 559, 562, 573, 594, 606, 616, 617, 661, 667, 669, 670, 692, 695, 697, 707, 727	28, 61, 116, 119, 200, 245, 401, 420, 710	56, 191, 259, 346, 361, 416, 520, 630, 682	72 ^S , 9 ^Y , 9 ^T

Cs7g07550	401	16, 25, 27, 36, 37, 93, 113, 132, 134, 139, 140, 143, 153, 155, 156, 160, 187, 188, 190, 191, 195, 210, 225, 229, 233, 281, 298, 301, 312, 315, 346, 357, 359, 362, 363, 365, 370, 393	259, 269	42, 44, 46, 48, 117, 144, 149, 150, 169, 171, 348, 350	38 ^S , 2 ^Y , 12 ^T
Cs3g10860	466	18, 42, 47, 52, 58, 59, 64, 66, 82, 83, 86, 90, 132, 135, 151, 166, 179, 181, 184, 187, 191, 202, 204, 219, 234, 236, 237, 241, 269, 273, 277, 293, 310, 354, 359, 361, 363, 364, 368, 370, 373, 398, 409, 412, 428, 437, 443, 456, 462	145, 313	2, 13, 17, 24, 31, 110, 139, 173, 220, 221, 233, 250, 253, 286, 307, 329, 340, 345, 372, 424	49 ^S , 2 ^Y , 20 ^T
Cs8g15030	508	13, 20, 101, 102, 104, 108, 110, 115, 121, 122, 128, 130, 137, 140, 148, 155, 169, 170, 188, 197, 199, 221, 251, 288, 292, 299, 312, 327, 340, 362, 390, 399, 471, 472, 476, 484, 501	228, 310	73, 83, 97, 165, 183, 207, 234, 339, 382, 440, 449, 481, 491	37 ^S , 2 ^Y , 13 ^T
Cs7g13010	357	11, 29, 35, 59, 63, 86, 100, 132, 161, 165, 166, 251, 259, 268, 276, 277, 282, 342, 348, 349	93, 145, 340	6, 32, 34, 36, 40, 67, 194, 208, 281, 326	20 ^S , 3 ^Y , 10 ^T
Cs1g02310	359	3, 32, 54, 58, 61, 93, 130, 140, 145, 187, 229, 256, 265, 308, 322, 348, 349,	8, 68, 100, 127, 189, 206, 289	6, 10, 25, 57, 106, 142, 198, 214, 274, 295, 326, 327, 359	17 ^S , 7 ^Y , 13 ^T
Cs7g19070	332	11, 19, 20, 33, 38, 39, 43, 46, 58, 59, 78, 79, 81, 107, 120, 157, 160, 233, 256, 275, 284, 293, 302, 304, 324	114, 200	28, 35, 63, 88, 93, 126, 137, 221, 276, 325	25 ^S , 2 ^Y , 10 ^T
orange1.1t04546	363	12, 17, 24, 28, 71, 72, 81, 83, 85, 104, 107, 121, 122, 123, 139, 146, 148, 155, 156, 159, 166, 198, 199, 206, 207, 208, 219, 239, 243, 252, 257, 266, 269, 287, 348	20, 250	152, 160, 210, 211, 215, 259, 340	35 ^S , 2 ^Y , 7 ^T
Cs8g02640	268	9, 28, 53, 54, 71, 76, 83, 112, 161, 206, 218, 253	182	64, 70, 77, 80, 84, 88, 110, 250, 255, 262	12 ^S , 1 ^Y , 10 ^T
Cs3g08880	453	44, 52, 90, 100, 101, 144, 148, 154, 168, 176, 187, 190, 193, 208, 217, 219, 222, 224, 228, 231, 235, 239, 243, 249, 255, 259, 260, 279, 286, 324, 339, 342, 360, 364, 368, 392, 413, 429, 446	118, 125, 216	11, 21, 22, 23, 26, 39, 73, 74, 76, 120, 165, 227, 325, 440, 443	39 ^S , 3 ^Y , 15 ^T

Cs1g23780	353	9, 17, 18, 27, 32, 55, 65, 119, 123, 139, 141, 147, 149, 150, 153, 155, 158, 159, 162, 178, 199, 240, 268, 272, 276, 295, 309, 317, 321, 327, 346, 353	28	3, 8, 12, 15, 19, 25, 44, 52, 57, 75, 97, 99, 105, 126, 188, 293, 344	32 ^S , 1 ^Y , 17 ^T
orange1.1t00453	421	9, 12, 16, 17, 21, 51, 55, 97, 107, 122, 143, 148, 151, 161, 163, 168, 172, 182, 187, 204, 227, 229, 238, 258, 261, 289, 293, 297, 309, 314, 321, 324, 330, 370, 376, 381, 399, 412	48, 62, 74	106, 112, 169, 176, 191, 200, 205, 214, 221, 243, 367, 383	38 ^S , 3 ^Y , 12 ^T
Cs7g05140	168	7, 8, 12, 13, 16, 17, 18, 21, 22, 23, 36, 56, 62, 82, 100, 127, 134, 153, 158, 167	111	6, 135, 151	20 ^S , 1 ^Y , 3 ^T
Cs7g29820	332	3, 23 25, 39, 43, 53, 60, 84, 94, 96, 102, 104, 105, 109, 134, 142, 146, 153, 155, 172, 186, 276, 288, 292, 313, 318, 322, 324, 326, 328	29, 153	100, 163, 176, 182, 183, 187, 190, 229, 260, 277	30 ^S , 2 ^Y , 10 ^T
Cs3g08170	597	5, 21, 22, 26, 27, 28, 29, 58, 68, 69, 86, 90, 103, 107, 115, 117, 120, 124, 132, 134, 139, 141, 143, 149, 153, 161, 163, 168, 191, 193, 200, 207, 211, 223, 242, 249, 259, 282, 283, 299, 302, 311, 312, 318, 319, 326, 334, 352, 359, 361, 364, 375, 381, 390, 392, 399, 403, 412, 445, 449, 458, 475, 485, 532, 537, 549, 560, 576	456	13, 35, 111, 142, 150, 201, 285, 304, 353, 354, 431, 466, 591	68 ^S , 1 ^Y , 13 ^T
Cs3g25760	424	2, 5, 25, 29, 35, 36, 44, 49, 51, 54, 56, 60, 66, 76, 80, 83, 97, 109, 110, 111, 118, 132, 168, 191, 192, 201, 204, 206, 213, 214, 220, 227, 228, 232, 241, 257, 260, 281, 286, 335, 339, 348, 374, 424	223, 261, 346	125, 194, 249, 422	44 ^S , 3 ^Y , 4 ^T
Cs5g23040	426	6, 8, 16, 20, 21, 23, 38, 39, 50, 52, 55, 63, 67, 88, 113, 115, 124, 139, 162, 165, 166, 174, 185, 187, 191, 193, 201, 203, 208, 211, 216, 240, 244, 268, 367, 406, 409, 417, 422, 424	251	92, 118, 175, 192, 261, 356, 416	40 ^S , 1 ^Y , 7 ^T
Cs3g16310	140	30, 35, 37, 39, 40, 43, 49, 60, 64, 68, 86, 116, 126, 134, 139	—	80, 115, 123, 137	15 ^S , 4 ^T
Cs1g21370	185	24, 41, 47, 48, 49, 50, 51, 63, 65, 73, 85, 89, 93, 119, 141, 151, 181, 182	—	27, 134, 166	18 ^S , 3 ^T

Cs6g19350	169	5, 17, 19, 41, 44, 45, 47, 48, 51, 53, 63, 76, 80, 84, 151, 152	–	52	16 ^S , 1 ^T
Cs8g07470	172	4, 5, 8, 9, 12, 20, 22, 40, 44, 48, 92, 107, 139, 165	84	7, 81, 102	14 ^S , 1 ^Y , 3 ^T
orange1.1t01674	145	9, 10, 12, 14, 31, 35, 39, 71, 98, 101	18, 75	93	10 ^S , 2 ^Y , 1 ^T
Cs7g12290	201	48, 60, 62, 64, 69, 71, 92, 96, 100, 133, 135, 163, 164, 170, 188, 192, 196	–	59, 70	17 ^S , 2 ^T
Cs7g25940	153	4, 7, 12, 14, 24, 28, 30, 33, 35, 46, 58, 62, 66, 92, 110, 118, 124, 149, 151	–	27, 43	19 ^S , 2 ^T
Cs3g27850	298	24, 30, 33, 37, 39, 47, 57, 65, 79, 82, 101, 105, 120, 137, 153, 168, 173, 195, 211, 223, 242, 246, 267, 283, 295	124, 160, 170, 20	18, 91, 127, 207, 235, 290	25 ^S , 4 ^Y , 6 ^T
Cs6g15200	488	27, 51, 58, 70, 79, 104, 109, 117, 123, 146, 169, 174, 206, 224, 225, 251, 256, 273, 280, 281, 311, 313, 331, 338, 340, 341, 342, 383, 407, 408, 412, 433, 441, 454	115, 186, 298, 336, 364, 372	49, 86, 128, 150, 162, 191, 211, 240, 265, 275, 288, 295, 345, 400, 402, 429, 474	34 ^S , 6 ^Y , 9 ^T
Cs3g06150	374	4, 8, 13, 26, 30, 32, 36, 56, 63, 64, 66, 74, 95, 96, 97, 99, 119, 126, 129, 144, 175, 187, 197, 200, 213, 215, 216, 217, 223, 228, 230, 262, 271, 285, 297, 336, 337, 338, 343, 344, 355, 363	35, 102, 113, 121, 137, 147, 252	16, 62, 85, 141, 198, 199, 234, 268, 342	42 ^S , 7 ^Y , 9 ^T
Cs7g23390	307	19, 27, 34, 42, 43, 45, 46, 74, 76, 78, 108, 113, 120, 125, 139, 141, 158, 163, 231, 235, 268, 272, 275, 290	35, 198, 255	128, 180, 183, 279, 283	24 ^S , 3 ^Y , 5 ^T

APÊNDICE 4 – Motivos conservados preditos pelo software MEME Suite para as 34 proteínas CsbZIPs. O tamanho das letras indica o grau de conservação do aminoácido em cada posição.

Motivo 1 = bZIP



Motivo 2 = DOG1



Motivo 3



Motivo 4



Motivo 5 = DOG1



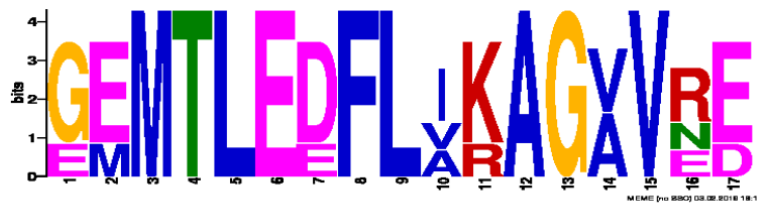
Motivo 6:



Motivo 7



Motivo 8



Motivo 9



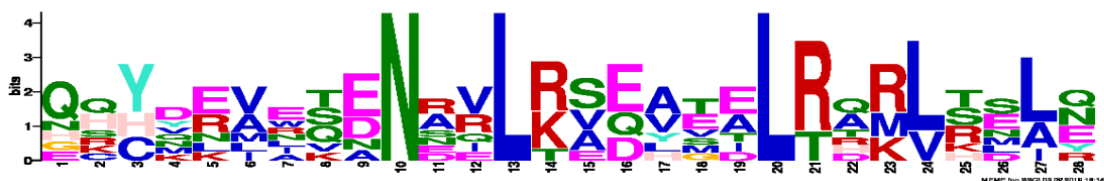
Motivo 10



Motivo 11



Motivo 12



Motivo 13



Motivo 14



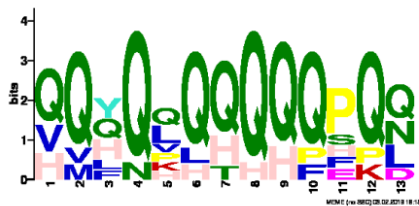
Motivo 15



Motivo 16



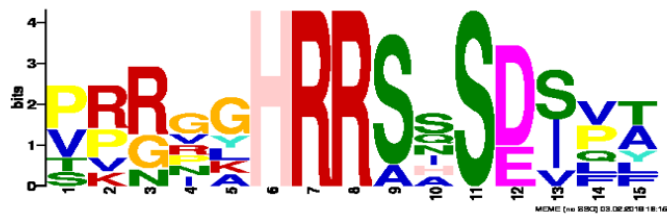
Motivo 17



Motivo 18



Motivo 19



Motivo 20

