



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
LEANDRO RIBEIRO SILVEIRA

Identificação e caracterização de proteínas e genes relacionados à resposta ao estresse hídrico em
tomateiro (*Solanum lycopersicum* L.)

Monte Carmelo

2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
LEANDRO RIBEIRO SILVEIRA

Identificação e caracterização de proteínas e genes relacionados à resposta ao estresse hídrico em
tomateiro (*Solanum lycopersicum* L.)

Trabalho de Conclusão apresentado ao curso de
Agronomia da Universidade Federal de
Uberlândia, Campus Monte Carmelo, como
requisito necessário para a obtenção do grau de
Engenheiro Agrônomo.

Orientadora: Prof^ª Dra. Ana Carolina Silva
Siquieroli

Monte Carmelo
2023

LEANDRO RIBEIRO SILVEIRA

Identificação e caracterização de proteínas e genes relacionados à resposta ao estresse hídrico em
tomateiro (*Solanum lycopersicum* L.)

Trabalho de Conclusão apresentado ao curso de
Agronomia da Universidade Federal de
Uberlândia, Campus Monte Carmelo, como
requisito necessário para a obtenção do grau de
Engenheiro Agrônomo.

Monte Carmelo, 26 de junho de 2023

Banca Examinadora

Prof^a. Dr^a. Ana Carolina Silva Siquieroli
Orientadora

Prof. Dr. Flavio Tetsuo Sasaki
Membro da Banca

MSc. Ana Luísa Alves Ribeiro
Membro da Banca

Monte Carmelo
2023

SUMÁRIO

1. Introdução.....	5
2. Justificativa.....	6
3. Material e Métodos	7
3.1. Identificação e caracterização de proteínas envolvidos em respostas ao estresse hídrico em organismos modelo.....	7
3.2. Identificação de proteínas ortólogas	8
3.3. Análises de domínios conservados e sítios ativos.....	8
3.4. Análises filogenéticas	8
4. Resultados e Discussão.....	9
5. Conclusões.....	22
Referências	23

RESUMO: O estresse hídrico afeta o tomateiro nas diferentes fases do desenvolvimento, desde a germinação até a frutificação, influenciando diretamente a produção. Uma das principais respostas é a modificação da expressão de genes relacionados com a produção de enzimas importantes na via de síntese de moléculas antioxidantes, fatores de transcrição e outras proteínas envolvidas em processos regulatórios ao estresse hídrico. Este projeto teve como objetivo identificar e caracterizar, por meio de análises *in silico*, proteínas e genes envolvidos com os mecanismos de resposta ao estresse hídrico em tomateiro, a fim de se obter marcadores moleculares para seleção de genótipos resistentes. Foram identificadas diversas proteínas putativas, que por meio da revisão na literatura e bancos de dados especializados, foram comparadas com suas ortólogas e descritos os seus atributos. Os métodos de análise *in silico* empregados permitiram a identificação de proteínas putativas relacionadas com o estresse hídrico no tomateiro e que poderão ser utilizadas na identificação de genes de interesse. Assim, foi possível obter informações e ferramentas importantes a serem utilizadas em programas de melhoramento genético de tomateiro.

Palavras-chave: bioinformática; expressão gênica, melhoramento genético

1. Introdução

O tomateiro é considerado a principal olerícola produzida pelo Brasil, tendo recebido grande foco das entidades de pesquisas brasileiras, ocasionando grandes aumentos em produtividade nos últimos 20 anos, graças ao melhoramento genético, tecnificação de práticas agrícolas e uso racional de recursos (TRANI et al., 2015).

Dados do IBGE (2019) sobre a produção de tomate mostram que no ano de 2018 foram produzidos cerca de 4 milhões de toneladas, com previsão de aumento de 5,7% para a safra de 2019 e com aumento de eficiência em área de 3,5%.

Além da questão fitossanitária, se faz presente a necessidade de utilização racional de recursos básicos como a água. O estresse hídrico no tomateiro pode influenciar na altura da planta, número de folhas, taxa fotossintética, transpiração, condutância estomática e aumento da temperatura foliar, resultando em menor produção de frutos. Foi constatado reduções entre 1,8 e 2,7 vezes na produção de tomateiros submetidos ao estresse hídrico em comparação com plantas mantidas em condições normais (MORALES et al., 2015).

Sabendo-se que a água é um recurso que tem se tornado cada vez mais escasso em todo o mundo, exige-se da agricultura moderna sistemas de produção e tecnologias que reduzam a sua utilização. Em adição, o tomateiro é uma das hortaliças mais exigentes em consumo de água e, portanto, sua falta é mais expressiva (INCAPER, 2010).

As respostas ao estresse hídrico são modificações das expressões de genes relacionados com a produção de enzimas importantes na via de síntese de enzimas antioxidantes, fatores de transcrição e outras proteínas envolvidas nas respostas aos estresse hídrico (CARVALHO, 2008; MILLER et al., 2009; SAIBO et al., 2009). Análises de *microarray* permite a identificação de genes expressos ou silenciados (PINEDA et al., 2012).

A busca pela melhor compreensão da base molecular das respostas ao estresse hídrico em tomateiro é necessária para alcançar a eficiência e sucesso na seleção de marcadores (GERSZBERG & HNATUSZKO-KONKA, 2015).

Assim, este projeto objetivou identificar e caracterizar, por meio de análises *in silico*, genes e proteínas envolvidas com os mecanismos de resposta ao estresse hídrico em tomateiro, a fim de se obter marcadores moleculares para seleção de genótipos resistentes.

2. Justificativa

O estresse hídrico é uma das maiores restrições ambientais à produtividade das culturas agrícolas, sendo o tomateiro sensível a esta condição em diferentes estádios de desenvolvimento, desde a germinação até a frutificação, influenciando diretamente na produção (NAHAR et al., 2011).

O primeiro passo para a caracterização da resposta complexa ao estresse abiótico tem sido a descoberta de genes, a partir do sequenciamento de genomas, e o uso de plataformas computacionais para determinar padrões de expressão transcricional (ZHOU et al., 2010; CRAMER et al., 2011).

Assim, uma maneira de abordar o problema tem sido a aplicação da genômica e proteômica, obtendo uma visão da resposta ao estresse no nível da planta em geral. O processamento e análise, por meio de ferramentas de bioinformática, da grande quantidade de dados gerados possibilitará aumentar a eficiência dos programas de melhoramento genético, visando à obtenção de genótipos tolerantes à seca.

Diferentes estudos relacionados à transcrição e expressão gênica em tomateiro identificaram a ativação e a regulação de diferentes genes em condições de estresse hídrico. No entanto, a identificação de novos genes envolvidos na resposta e mecanismos de tolerância a esta condição são essenciais. Em adição, tem-se um aumento significativo na quantidade de genes de respostas iniciais regulados pelo estresse que codifica proteínas desconhecidas, indicando que existem ainda aspectos a serem descobertos sobre os mecanismos de tolerância a seca em tomateiro (GUPTA et al., 2009; PINEDA et al., 2012).

A utilização dessa plataforma tecnológica acoplada ao Programa de Melhoramento Genético de Tomateiro da UFU, Campus Monte Carmelo, poderá contribuir com o desenvolvimento da cultura do tomateiro, principalmente em grandes áreas de produção, demonstrando o potencial da pesquisa.

3. Material e Métodos

A estratégia metodológica adotada no trabalho está representada na Figura 1:

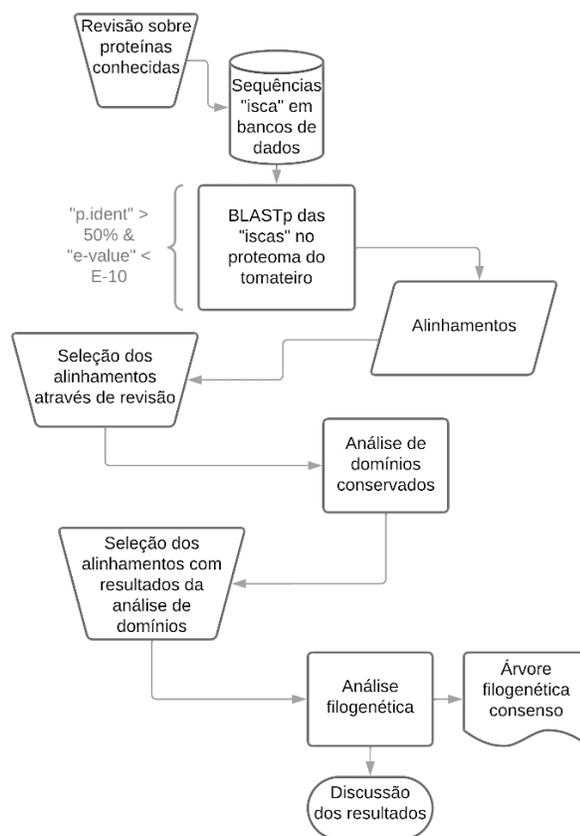


Figura 1. Fluxograma da metodologia adotada no trabalho.

3.1. Identificação e caracterização de proteínas envolvidos em respostas ao estresse hídrico em organismos modelo

Para a identificação dos prováveis genes foram realizadas buscas em banco de dados públicos de vegetais, por meio dos critérios de similaridade, palavras-chave e literatura. A busca teve como alvo genes relacionados a enzimas e fatores de transcrição envolvidos nas rotas metabólicas de resposta ao estresse hídrico, já citados na literatura em organismos-modelo, como *Arabidopsis thaliana* e *Coffea canephora*.

As sequências desses genes foram utilizadas como moldes para busca no banco de dados do *National Center for Biotechnology Information - NCBI* (SAYERS et al., 2021) através da

ferramenta *BLASTP* (ALTSCHUL, 1990), a fim de identificar as prováveis sequências ortólogas destes genes no genoma do tomateiro.

3.2. Identificação de proteínas ortólogas

Foi utilizado como “isca” para a identificação de prováveis proteínas relacionadas à resposta ao estresse hídrico em *Arabidopsis thaliana*, organismo modelo padrão utilizado para estudos vegetais, e o banco de dados gerado por Silva (2020), no qual foram identificados proteínas e fatores de transcrição relacionados com a resposta ao estresse hídrico no cafeeiro (*Coffea canephora*), com resultados promissores.

As sequências das prováveis proteínas envolvidos na resposta ao estresse hídrico encontradas no organismo modelo foram comparadas com o proteoma predito de tomateiro (*Solanum lycopersicum*) a partir do uso da plataforma *Basic Local Alignment Search Tool Protein - Blastp* (ALTSCHUL, 1990) com “e-value” menor ou igual a e^{-10} e “p.ident” maiores que 50%, com o intuito de melhorar a qualidade do alinhamento, sendo selecionadas as proteínas que apresentaram a melhor pontuação na combinação de cada “query” e “match” correspondente.

Todas as proteínas selecionadas, após os filtros, foram revisadas na literatura e anotações manuais e automáticas de bancos de dados como o site *PlantTFDB* (JIN et al., 2016), *UniProt* (THE UNIPROT CONSORTIUM, 2022) e *KEGG: Kyoto Encyclopedia of Genes and Genomes* (KANEHISA, 2000), procurando por indícios sobre suas funções no tomateiro.

3.3. Análises de domínios conservados e sítios ativos

Após a obtenção das sequências das prováveis proteínas preditas envolvidas na resposta ao estresse hídrico, as mesmas foram submetidas a uma análise de domínio conservado no site *HMMER - Biosequence analysis using profile hidden Markov Models* (<https://hmmer.org/>), com o objetivo de agregar indícios da conservação da função das proteínas ortólogas no tomateiro.

3.4. Análises filogenéticas

Para estabelecer as relações filogenéticas e a provável homologia entre as sequências obtidas nesse trabalho foi utilizando o software MrBayes, que utiliza o método bayesiano para o cálculo das árvores, realizada a análise com 1 milhão de simulações em bootstrap (RONQUIST; HUELSENBECK, 2003).

4. Resultados e Discussão

O cruzamento do banco de dados “isca” com o banco de dados de *A. thaliana* gerou 5.098 correspondências alinhadas com *S. lycopersicum*, que quando filtradas pelos parâmetros estabelecidos na metodologia e remoção de duplicatas, resultou em 70 proteínas que atendiam a todas as exigências inicialmente estabelecidas (Tabela 1).

Tabela 1. Proteínas selecionadas do alinhamento comparativo das sequências referência de *Arabidopsis thaliana* em *Solanum lycopersicum*.

ID ortólogo (<i>A. Thaliana</i>)	Proteína Referência (<i>S. lycopersicum</i>)	'p.ident'	'e-value'	'Score'
NP_567084.1	NP_001233963.2	0,60502	1,1E-143	409
NP_179396.1	NP_001233985.1	0,60811	7,65E-21	86,3
NP_567513.4	NP_001234198.1	0,8012	0	1144
NP_179396.1	NP_001234389.1	0,5	5,86E-38	129
NP_191911.1	NP_001234542.1	0,79228	0	899
NP_849777.1	NP_001234596.1	0,553	1,6E-130	383
NP_187713.1	NP_001234689.1	0,625	6,16E-46	158
NP_196107.1	NP_001234784.1	0,76346	0	924
NP_974009.1	NP_001238798.1	0,51572	1,18E-32	114
NP_188666.1	NP_001239044.1	0,57265	4,43E-46	153
NP_567773.1	NP_001306107.1	0,70769	7,7E-129	370
NP_001329320.1	NP_001333570.1	0,65328	0	657
NP_190400.1	NP_001333606.1	0,70459	0	746
NP_195994.3	NP_001351962.1	0,53774	1,38E-28	111
NP_001190159.1	XP_004228428.1	0,69945	3,45E-92	272
NP_850095.2	XP_004228494.1	0,81733	0	751
NP_191748.1	XP_004230017.1	0,50943	2,16E-65	202
NP_001324877.1	XP_004230106.1	0,51155	9,9E-103	303
NP_188286.1	XP_004230244.1	0,77976	2,84E-95	273
NP_001319253.1	XP_004230276.1	0,53333	1,66E-61	191
NP_176290.1	XP_004230475.1	0,86427	0	614
NP_195493.1	XP_004231956.1	0,5979	7,9E-102	298
NP_567564.1	XP_004233024.1	0,7725	0	1595
NP_001078705.1	XP_004233527.1	0,75486	2,2E-128	364
NP_565424.1	XP_004233742.1	0,6176	0	1473
NP_200128.1	XP_004234175.1	0,78378	4,1E-106	302
NP_568744.1	XP_004235358.1	0,7265	2,55E-55	168
NP_188478.1	XP_004235788.3	0,58849	0	550

NP_001323240.1	XP_004236023.1	0,65116	3,23E-82	241
NP_001331469.1	XP_004236734.1	0,52381	9,47E-95	293
NP_201280.1	XP_004237483.1	0,5	5,42E-37	136
NP_194150.1	XP_004238368.1	0,82353	0	1284
NP_178487.1	XP_004239010.1	0,8	0	1219
NP_181487.1	XP_004239787.1	0,65868	6,7E-161	456
NP_001030747.1	XP_004240193.1	0,83051	3,79E-31	114
NP_190246.1	XP_004240370.1	0,51289	1,63E-96	293
NP_567231.1	XP_004240547.1	0,60825	1,52E-29	101
NP_181546.1	XP_004240855.1	0,79348	8,14E-46	142
NP_182191.1	XP_004240976.1	0,59649	1,16E-37	132
NP_001318958.1	XP_004241580.1	0,95833	6,69E-10	57,4
NP_174534.1	XP_004241969.1	0,57895	1,41E-14	65,5
NP_567137.1	XP_004242329.1	0,59684	5,4E-105	306
NP_850146.1	XP_004242480.1	0,51793	0	680
NP_179685.1	XP_004243339.1	0,50204	2,38E-59	195
NP_172003.1	XP_004243853.1	0,5034	3,4E-101	301
NP_563626.1	XP_004245523.1	0,82424	3,3E-104	297
NP_176833.1	XP_004245530.1	0,56667	1,63E-17	80,5
NP_001190812.1	XP_004245586.1	0,68	1,78E-20	85,5
NP_182150.1	XP_004245929.1	0,5875	1,1E-23	101
NP_001326558.1	XP_004246319.1	0,77273	1,12E-42	135
NP_565702.1	XP_004247220.1	0,5514	2,8E-157	451
NP_191244.1	XP_004249361.1	0,56151	2,38E-97	290
NP_565928.1	XP_004249671.1	0,62234	2,79E-81	241
NP_174807.1	XP_004250396.1	0,78782	0	805
NP_001318754.1	XP_004250562.3	0,70068	1,18E-66	205
NP_850095.2	XP_004250885.1	0,82831	0	763
NP_195215.2	XP_004250950.1	0,74085	0	2075
NP_175818.1	XP_004251681.1	0,68182	6,2E-38	135
NP_180090.6	XP_004251765.1	0,55584	4,5E-129	395
NP_198513.1	XP_010315726.1	0,54651	1E-92	273
NP_190334.1	XP_010316809.1	0,85	6,91E-34	124
NP_182274.1	XP_010318408.1	0,5	9,36E-33	120
NP_001326054.1	XP_010318851.1	0,52	0	576
NP_563933.1	XP_010320856.1	0,54472	1,72E-34	131
NP_001320130.1	NP_001306107.1	0,5	2,35E-19	90,1
NP_001031785.2	XP_010325920.1	0,61111	1,05E-17	85,1
NP_196160.1	XP_010327719.1	0,57143	1,53E-43	154
NP_179744.1	XP_019068107.1	0,53425	5,02E-11	57,4
NP_196294.1	XP_019071681.1	0,54945	3,62E-24	90,1
NP_182150.1	XP_025887866.1	0,50234	6,7E-122	362

Com informações das 70 proteínas foi possível realizar buscas na literatura e banco de dados especializados, para relacionar as informações estatísticas anotadas de forma automática pelas próprias plataformas com as sequências de proteínas já descritas, na mesma espécie e espécie referência (*A. thaliana*). Com as informações de cada proteína dos bancos de dados UniProt KEGG: Kyoto Encyclopedia of Genes and Genomes e Plant Transcription Factor Database foram identificadas 26 candidatas, sendo proteínas e mais especificamente, fatores de transcrição, que demonstravam indícios de relação com a resposta ao estresse hídrico no tomateiro (Tabela 2).

Tabela 2. Revisão em bancos de dados especializados de informações anotadas das sequências referência de proteínas putativas encontradas em tomateiro.

Proteína referência (<i>S. lycopersicum</i>)	Família de fator de transcrição	Nome do gene putativo	Descrição na literatura	Tamanho (aa)	Processos anotados nos bancos de dados
NP_001233985.1	-	H1	histone H1	271	Resposta ao estresse salino, etileno, ácido jasmônico, íon de cádmio
NP_001234389.1	MYB	LOC544142	H1 histone-like protein	207	Resposta ao estresse salino, etileno, ácido jasmônico, íon de cádmio
NP_001234596.1	bZIP	AREB1	ABA-response element binding factor AREB1	447	Resposta à privação de água, resposta ao estresse salino, via de sinalização ativada por ácido abscísico, via de sinalização mediada por glicose
NP_001234689.1	ERF	DREB1	dehydration responsive element binding 1	300	Resposta ao ácido abscísico, aclimação térmica
NP_001239044.1	ERF	ERF7	ethylene response factor F.2	240	Fator de transcrição que se liga a GCC, elemento promotor relacionado à patogênese. Envolvido na regulação da expressão genética por fatores de estresse e por componentes de vias de transdução de sinal de estresse. Provavelmente age como um repressor transcricional e pode regular outros AtERFs (por similaridade).

NP_001306107.1	NAC	JA2L	NAC domain-containing protein JA2L	356	Resposta de defesa, resposta à privação de água, resposta ao ácido abscísico
NP_001351962.1	C2H2	HDT3	histone deacetylase	306	Especificação de polaridade do eixo adaxial/abaxial, processo de dormência de sementes
XP_004230017.1	HD-ZIP	LOC101262661	homeobox-leucine zipper protein ATHB-12	239	Resposta à privação de água, via de sinalização ativada por ácido abscísico
XP_004231956.1	HD-ZIP	LOC101246982	homeobox-leucine zipper protein HAT22	268	Resposta à citocinina
XP_004237483.1	ERF	LOC101263697	ethylene-responsive transcription factor ABR1-like	303	*
XP_004240193.1	MYB	LOC101265029	transcription factor MYB14-like	218	Regula a tolerância ao congelamento, afetando a expressão dos genes da CBF envolvidos na seca e na tolerância ao sal. Pode aumentar os níveis de expressão de genes envolvidos na biossíntese e sinalização do ácido abscísico (ABA), bem como aquelas proteínas protetoras do estresse codificadas (PubMed:19161942)
XP_004240976.1	HD-ZIP	LOC101264731	homeobox-leucine zipper protein ATHB-12	225	*
XP_004241580.1	MYB	LOC101267793	myb-related protein 306	334	Resposta à privação de água, estresse salino, ácido abscísico, ácido salicílico
XP_004242480.1	B3	LOC101245071	B3 domain-containing transcription repressor VAL1-like	845	Regulação positiva da germinação de sementes, regulação do amadurecimento das sementes
XP_004245530.1	WRKY	LOC101261141	probable WRKY	339	*

XP_004245929.1	bZIP	LOC101253816	transcription factor 53 bZIP transcription factor 16 isoform X2 ABSCISIC ACID-	409	Regulação do processo metabólico de peróxido de hidrogênio
XP_004249361.1	bZIP	LOC101259896	INSENSITIVE 5-like protein 2	350	Resposta ao ácido abscísico
XP_004250562.3	NF-YB	LOC101252139	nuclear transcription factor Y subunit B-6	222	*
XP_004251681.1	NF-YB	LOC101247294	nuclear transcription factor Y subunit A-7- like isoform X1	283	Fatores beta-Scaffold com pequenos contatos de ranhura. Fatores heteroméricos CCAAT
XP_004251765.1	ARR-B	LOC101246797	two- component response regulator ORR24-like	708	Resposta à privação de água / regulação do crescimento meristemático / regulação do processo biossintético de clorofila / regulação do processo metabólico de antocianina / desenvolvimento do sistema brotação / desenvolvimento primário raiz / regulação da via de sinalização ativada por citocinina / regulação do crescimento das sementes
XP_010316809.1	Dof	LOC104645845	dof zinc finger protein DOF1.5	163	Regulação negativa da manutenção populacional de células-tronco
XP_010318408.1	ERF	LOC543910	JERF3 isoform X1	397	Detecção de hipóxia
XP_010325918.1	bZIP	LOC104649075	protein ABSCISIC ACID- INSENSITIVE 5 isoform X1	426	Resposta à privação de água/ frio / estresse salino / ácido abscísico / giberelina / maturação do pólen / regulação negativa da germinação de sementes / resposta ao desenvolvimento de quitina e sementes

XP_010325920.1	bZIP	LOC104649075	protein ABSCISIC ACID- INSENSITIVE 5 isoform X2	415	Resposta à privação de água / frio / estresse salino / ácido abscísico / giberelina / maturação do pólen / regulação negativa da germinação de sementes / resposta ao desenvolvimento de quitina e sementes
XP_010327719.1	ERF	LOC101245410	dehydration- responsive element- binding protein 2A transcriptional	406	*
XP_025887866.1	bZIP	LOC101249493	activator TAF- 1 isoform X1	504	Fatores de ligação ao nucleotídeo G da planta / resposta ao ácido abscísico

* Informações anotadas de forma automática insuficientes.

A busca por informações contidas nos bancos de dados especializados sobre as proteínas encontradas por meio do alinhamento comparativo (Tabela 2) e identificação de sítios conservados (Tabela 3), retornaram diversas informações sobre suas possíveis funções no tomateiro. Foram selecionadas famílias de fatores de transcrição que possuíam maiores indícios da relação com a resposta ao estresse hídrico de acordo com as informações disponíveis. A análise dos sítios conservados permitiu relacionar a sua função putativa com as proteínas já caracterizadas na literatura como atuantes na resposta ao estresse hídrico (Tabela 3).

Tabela 3. Identificação de sítios conservados nas proteínas putativas identificadas.

'Query'	PSSM-ID	Início do sítio (aa)	Fim do sítio (aa)	E-Value	Nome curto	Superfamília
Q#1 - NP_001233985.1	197772	54	118	7,42E-20	H15	cl00073
Q#2 - NP_001234389.1	197772	51	117	1,02E-11	H15	cl00073
Q#3 - NP_001234596.1	269855	369	423	9,46E-25	bZIP_plant_BZIP4 6	cl21462
Q#4 - NP_001234689.1	197689	81	139	3,83E-28	AP2	cl00033
Q#5 - NP_001239044.1	197689	38	98	1,81E-15	AP2	cl00033
Q#6 - NP_001306107.1	367055	16	139	2,67E-32	NAM	cl03558

NP_001351962.1	Q#7 -	375342	3	93	1,65E-15	NPL	cl03870
XP_004230017.1	Q#8 -	365835	46	96	9,73E-13	Homeobox	cl00084
XP_004230017.1	Q#8 -	366961	98	137	9,43E-12	HALZ	cl23880
XP_004231956.1	Q#9 -	128634	179	222	7,06E-18	HALZ	cl23880
XP_004231956.1	Q#9 -	197696	123	177	1,05E-10	HOX	cl00084
XP_004237483.1	Q#10 -	197689	144	195	7,36E-25	AP2	cl00033
XP_004240193.1	Q#11 -	178751	5	116	1,12E-57	PLN03212 superfamily	-
XP_004240976.1	Q#12 -	365835	23	73	1,64E-12	Homeobox	cl00084
XP_004240976.1	Q#12 -	366961	75	116	9,37E-08	HALZ	cl23880
XP_004241580.1	Q#13 -	178751	5	111	1,04E-43	PLN03212 superfamily	-
XP_004242480.1	Q#14 -	197383	315	416	3,28E-19	B3_DNA	cl15242
XP_004242480.1	Q#14 -	369393	547	589	1,48E-08	zf-CW	cl06504
XP_004245530.1	Q#15 -	367337	118	178	6,56E-32	WRKY	cl03892
XP_004245929.1	Q#16 -	374656	137	272	7,06E-43	MFMR_assoc	cl24974
XP_004245929.1	Q#16 -	369512	1	98	1,09E-30	MFMR	cl06693
XP_004245929.1	Q#16 -	197664	305	369	1,84E-20	BRLZ	cl21462
XP_004245929.1	Q#16 -	388540	342	398	0.0020836 2	MreC superfamily	-
XP_004249361.1	Q#17 -	269855	280	325	2,56E-19	bZIP_plant_BZIP4 6	cl21462
XP_004250562.3	Q#18 -	366318	84	148	3,18E-22	CBFD_NFYB_HM F	cl23735
XP_004251681.1	Q#19 -	366892	147	202	1,32E-29	CBFB_NFYA	cl02165
XP_004251765.1	Q#20 -	381121	26	140	5.31e-64	REC_typeB_ARR- like	cl19078
XP_004251765.1	Q#20 -	130620	208	263	8,71E-17	myb_SHAQKYF	cl21498
XP_010316809.1	Q#21 -	367145	54	110	1,94E-32	zf-Dof	cl03664

Q#22 - XP_010318408.1	197689	115	172	2,19E-28	AP2	cl00033
Q#23 - XP_010325918.1	269855	342	396	1,15E-18	bZIP_plant_BZIP4 6	cl21462
Q#24 - XP_010325920.1	269855	342	383	4,09E-18	bZIP_plant_BZIP4 6	cl21462
Q#25 - XP_010327719.1	197689	69	138	1,08E-26	AP2	cl00033
Q#26 - XP_025887866.1	369512	81	172	1,28E-53	MFMR	cl06693
Q#26 - XP_025887866.1	374656	210	341	1,96E-35	MFMR_assoc	cl24974
Q#26 - XP_025887866.1	333897	361	423	5,35E-20	bZIP_1	cl21462

A análise de sítios conservados permitiu obter maiores evidências, além da alta similaridade da proteína, de que há famílias de domínios que possuem indícios de ter as mesmas funções de proteínas utilizadas como referência conhecidas e relacionadas às respostas ao estresse hídrico em *C. canephora* e *A. thaliana*.

O cálculo de probabilidade de semelhança leva em conta a conformação tridimensional, baseada em cálculos termodinâmicos realizados pela plataforma “*HMMER*”, o que corrobora com a hipótese de a proteína ser ortóloga, com função similar às referenciadas nas espécies modelo conhecidas.

Foram identificadas 11 famílias de fatores transcricionais (HD-ZIP, B3, WRKY, NF-YB, Dof, C2H2, MYB, bZIP, ERF, NAC e ARR-B) e um grupo de histonas, sendo descartada após revisão individual por não apresentar relação direta com o estresse hídrico no tomateiro.

O dendrograma consenso gerado de forma automática pelo software “*MrBayes*” resultante da análise bayesiana de filogenia de um milhão de simulações (Figura 2) apresentou ‘valor p’ de 0,0002, indicando grande confiabilidade no posicionamento das proteínas. Foi demonstrada a similaridade entre os fatores transcricionais da mesma família.

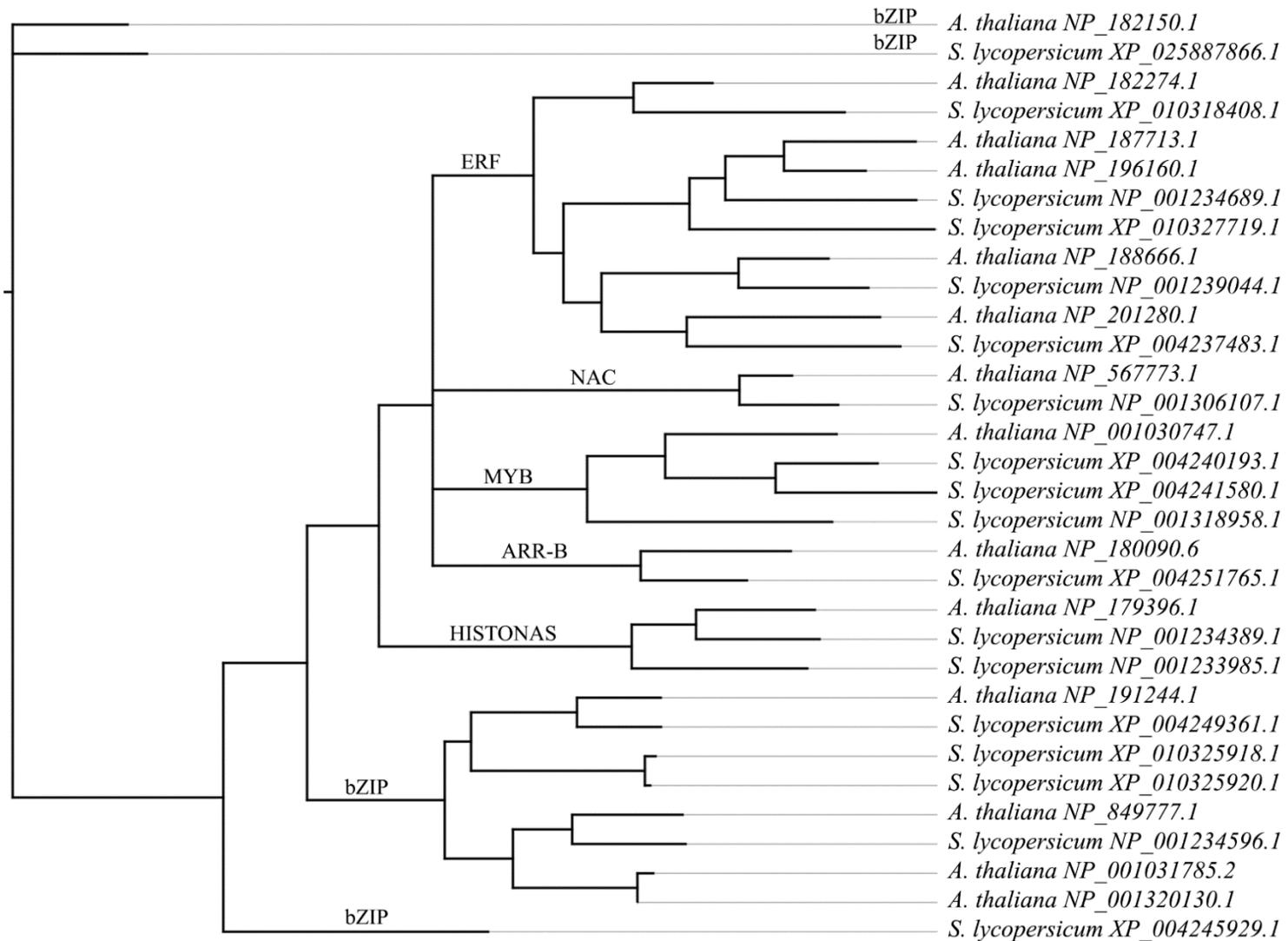


Figura 2. Dendrograma consenso das proteínas putativas relacionadas ao estresse hídrico no *S. lycopersicum* e *A. thaliana*.

Tabela 4. Comparação descritiva entre as proteínas ortólogas de *A. thaliana* e as proteínas putativas em *S. lycopersicum*.

ID da proteína ortóloga em <i>A. Thaliana</i>	ID da proteína putativa em <i>S. lycopersicum</i>	Família do fator de transcrição em <i>A. thaliana</i>	Família do fator de transcrição em <i>S. lycopersicum</i>	Número de aa em <i>A. thaliana</i>	Número de aa em <i>S. lycopersicum</i>	Função em <i>A. thaliana</i>	Função em <i>S. lycopersicum</i>
NP_001030747.1	XP_004240193.1	MYB	MYB	228	218	resposta ao frio, resposta ao estresse salino, resposta à privação de água	regula a tolerância ao congelamento, envolvido na tolerância à seca e ao sal / regulação do ácido abscísico, bem como aqueles que codificam proteínas protetoras de estresse
NP_001031785.2	XP_010325920.1	bZIP	bZIP	449	415	via de sinalização ativada por ácido abscísico, resposta ao ácido abscísico, resposta ao estresse salino, resposta à privação de água	privação de água / ácido abscísico /
NP_001318958.1	XP_004241580.1	MYB	MYB	197	334	resposta à privação de água, resposta ao ácido abscísico, resposta ao frio, resposta ao sal	resposta à privação de água, estresse salino, ácido abscísico
NP_001320130.1	XP_010325918.1	bZIP	bZIP	454	426	via de sinalização ativada por ácido abscísico, resposta ao ácido abscísico, resposta ao estresse	resposta ao frio / privação de água / estresse salino / ácido abscísico

						salino, resposta à privação de água	
NP_179396.1	NP_001233985.1	Histona* ⁴	Histona* ⁴	167	271	resposta à privação de água	resposta ao estresse salino, etileno
NP_179396.1	NP_001234389.1	ARR-B	ARR-B	167	207	desenvolvimento da raiz primária, regulação do meristema da raiz crescimento, resposta à privação de água	resposta ao estresse salino, etileno
NP_180090.6	XP_004251765.1	ARR-B	ARR-B	596	708	resposta à privação de água e frio	resposta à privação de água
NP_182150.1	XP_004245929.1	bZIP	bZIP	382	409	resposta à anoxia	regulação do processo metabólico do peróxido de hidrogênio
NP_182150.1	XP_025887866.1	bZIP	bZIP	382	504	resposta à privação de água	resposta ao ácido abscísico
NP_182274.1	XP_010318408.1	ERF	ERF	171	397	resposta ao ácido abscísico, resposta à privação de água	detecção de hipóxia
NP_187713.1	NP_001234689.1	ERF	ERF	330	300	regulação da resposta ao ácido abscísico	resposta ao ácido abscísico, aclimatação ao calor
NP_188666.1	NP_001239044.1	ERF	ERF	244	240	resposta à privação de água	resposta ao estresse, pode regular outros AtERFs

NP_191244.1	XP_004249361.1* ¹	bZIP	bZIP	297	350	resposta ao ácido abscísico, resposta ao estresse osmótico	resposta ao ácido abscísico
NP_196160.1	XP_010327719.1* ¹	ERF	ERF	335	406	Informação indisponível* ³	Informação indisponível* ³
NP_201280.1	XP_004237483.1	ERF	ERF	391	303	resposta ao ácido abscísico, resposta à privação de água	Informação indisponível* ³
NP_567773.1	NP_001306107.1	NAC	NAC	297	356	informação indisponível* ³	resposta de defesa / resposta à privação de água, resposta ao ácido abscísico
NP_849777.1	NP_001234596.1	bZIP	bZIP	416	447	via de sinalização ativada por ácido abscísico, resposta ao ácido abscísico, resposta ao estresse salino, resposta à privação de água	resposta à privação de água, resposta ao estresse salino, via de sinalização ativada por ácido abscísico

*¹ Proteína desconsiderada nesta análise devido à baixa similaridade *² Família não identificada nos bancos de dados especializados. *³ Proteínas anotadas automaticamente pelos bancos de dados especializados devido à similaridade com outras previamente descritas em outros organismos. *⁴ Proteínas do tipo histona.

Após a comparação descritiva entre as proteínas ortólogas de *A. thaliana* e as proteínas putativas em *S. lycopersicum* (Tabela 4) foi possível observar que pertencem ao mesmo grupo ou família de fatores transcricionais. A maioria possui tamanhos e descrições similares, o que reforça o fato de estarem relacionadas com a resposta ao estresse hídrico no tomateiro.

As análises de alinhamento básico local, análise de domínio conservado e filogenéticas revelam que a maior parte das proteínas putativas encontradas são conservadas a nível de aminoácidos, família e sítios ativos entre *S. lycopersicum* e *A. thaliana*, além de apresentarem um ‘valor e’ bem reduzido, o que aumenta as probabilidades de os alinhamentos não serem ao acaso. Após revisões na literatura foram desconsideradas aquelas que não possuíam informações suficientes, ou não apresentavam descrição relacionando-as ao estresse hídrico.

Das proteínas selecionadas, foram obtidos dois fatores de transcrição da família MYB, cinco bZIP, quatro ERF, uma NAC, duas ARR-B, assim como uma proteína do tipo histona. Todas estas proteínas atenderam os critérios inicialmente estabelecidos neste trabalho e foram revisadas individualmente nos bancos de dados especializados.

A família de fatores transcricionais ERF/AP2 já foi descrita em trabalhos semelhantes como o de Silva (2020) em *C. canephora* por suas relações com o estresse hídrico no cafeeiro. Outros autores chegaram a conclusões semelhantes sobre a função desta família de fatores transcricionais estar relacionada com a resposta ao estresse hídrico no tomateiro (SHARMA et al 2010), arroz (SHARONI et al. 2011), soja (ZHANG et al. 2008) e pimenta (*Capsicum annuum* L.) (JIN et al., 2018).

He et al. (2020) descreveram funções relacionadas à resposta a diversos estresses bióticos e abióticos como estresse salino, estresse hídrico, resistência a *Sclerotinia sclerotiorum* e *Pseudomonas syringae* quando em diferentes níveis de expressão da proteína GmbZIP19 em soja, uma representante da família bZIP. Yang et al. (2020) concluíram que diversos fatores transcricionais bZIP desempenham papéis importantes na resistência de plantas. Kang et al. (2019) também encontraram proteínas bZIP atuando na resposta ao estresse em *A. thaliana* e Yang et al. (2019) identificaram os mesmos resultados na cultura do arroz.

São reportadas na literatura as correlações entre a expressão de proteínas MYB e suas relações com a resposta a diversos estresses em soja (PEREIRA et al., 2011), trigo (LI et al., 2019), *A. thaliana* (ZHAO et al., 2018) e arroz (TANG et al., 2019). De acordo Baldoni et al. (2015) este comportamento da família MYB se repete em diversas outras espécies e tem se tornado cada vez

mais importante na aplicação prática em programas de melhoramento, principalmente na tolerância ao estresse.

Estudos recentes indicam relação entre fatores de transcrição da família NAC com a resposta ao estresse hídrico e oxidativo em tomate (THIRUMALAIKUMAR et al., 2017) e arroz (YUAN et al., 2019). Com a introdução do gene AmNAC11 de *Ammopiptanthus mongolicus* em *A. thaliana*, Pang et al. (2019) relataram aumento significativo na resistência a seca e frio. Hoang et al. (2019) reforçaram a importância do maior entendimento das ações destas proteínas para o desenvolvimento de novos transgênicos mais resistentes a condições de escassez de água e/ou salinidade.

Em *A. thaliana*, estudos como de Huang et al. (2018) e Nguyen et al. (2016) reportaram os fatores de transcrição do tipo ARR (A, B e C) como responsáveis pela regulação do ácido abscísico e citocinina durante a resposta ao estresse hídrico.

Os métodos de análise *in silico* empregados neste trabalho permitiram a identificação de proteínas putativas relacionadas com o estresse hídrico no tomateiro. Estas proteínas estão sendo utilizadas nas análises para identificação de genes e proteínas de interesse em parceria com o LBAM (Laboratório de Bioinformática e Análises Moleculares) da UFU, Campus Patos de Minas.

Sugere-se que outros trabalhos que busquem realizar análises similares às aqui apresentadas utilizem uma seleção de proteínas mais restrita, com ‘e-values’ menores que 10^{-20} e ‘p-ident’ maiores que 60 % para que se evitem falsos positivos. Em adição, torna-se importante a limitação a um tipo de proteína ou família de fator transcricional para maior entendimento individual e do grupo estudado.

5. Conclusões

O alinhamento comparativo, análise de sítios conservados e análise filogenética possibilitou a obtenção de proteínas putativas em *S. lycopersicum* relacionadas com a resposta ao estresse hídrico.

Os resultados apresentados dão maiores indícios para a possibilidade da construção de marcadores moleculares a partir das sequências dos genes codificadores e RNA mensageiros, tornando-se importantes ferramentas para seleção de genótipos em programas de melhoramento

vegetal, assim como a identificação de genes potenciais a serem utilizados na transformação genética das espécies.

Referências

- ALTSCHUL, Stephen F.; GISH, Warren; MILLER, Webb; MYERS, Eugene W.; LIPMAN, David J.. Basic local alignment search tool. **Journal Of Molecular Biology**, [S.L.], v. 215, n. 3, p. 403-410, out. 1990. Elsevier BV. [http://dx.doi.org/10.1016/s0022-2836\(05\)80360-2](http://dx.doi.org/10.1016/s0022-2836(05)80360-2).
- BALDONI, E.; GENGA, A.; COMINELLI, E. Plant MYB Transcription Factors: Their Role in Drought Response Mechanisms. **International Journal of Molecular Sciences**, 13 Jul. 2015. v. 16, n. 7, p. 15811–15851. Disponível em: <<http://www.mdpi.com/1422-0067/16/7/15811>>.
- CARVALHO, M. H. C. de. Drought stress and reactive oxygen species. **Plant Signal Behav**, v. 3, n. 3, p. 156-160, 2008.
- CRAMER, G. R.; URANO, K.; DELROT, S.; PEZZOTTI, M. Y SHINOZAKI, K. Effects of abiotic stress on plants: a systems biology perspective BMC. **Plant Biology**, 2011, vol. 11, pp. 163-176.
- GERSZBERG, A.; HNATUSZKO-KONKA, K. Tomato (*Solanum Lycopersicum*) in the service of biotechnology. **Plant Cell. Tissue Organ Cult.**, v.120, p. 881–902, 2015.
- GUPTA, V.; MATHUR, S.; SOLANKE, A. U.; SHARMA, M. K.; KUMAR, R.; VYAS, S.; KHURANA, P.; KHURANA, J. P.; TYAGI, A. K. Y SHARMA, A. K. Genome analysis and genetic enhancement of tomato. **Critical Reviews Biotech**, v. 29, n. 2, p. 152-181, 2009.
- HE, Q. *et al.* A Soybean bZIP Transcription Factor GmbZIP19 Confers Multiple Biotic and Abiotic Stress Responses in Plant. **International Journal of Molecular Sciences**, 1 Jul. 2020. v. 21, n. 13, p. 4701. Disponível em: <<https://www.mdpi.com/1422-0067/21/13/4701>>.
- HOANG, X. L. T. et al. NAC Transcription Factors in Drought and Salinity Tolerance. **BMC Plant Biology**, v. 19, p. 351–366, 2019. Disponível em: <<https://doi.org/10.1186/s12870-019-1883-y>>.
- HUANG, X. *et al.* The Antagonistic Action of Abscisic Acid and Cytokinin Signaling Mediates Drought Stress Response in Arabidopsis. **Molecular Plant**, jul. 2018. v. 11, n. 7, p. 970–982. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1674205218301606>>.
- IBGE. Área, Produção e Rendimento Médio - Confronto das Safras de 2018 e das Estimativas para 2019 - Brasil. **Levantamento Sistemático da Produção Agrícola - LSPA**, 2019. Disponível em: <[ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_\[mensal\]/Tabelas_xls/2019/lspa_201902_02.xls](ftp://ftp.ibge.gov.br/Producao_Agricola/Levantamento_Sistematico_da_Producao_Agricola_[mensal]/Tabelas_xls/2019/lspa_201902_02.xls)>. Acesso em: 28 Março 2019.

INCAPER. **Tomate**. Vitória, ES: Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural, 2010.

JIN, Jinpu; TIAN, Feng; YANG, De-Chang; MENG, Yu-Qi; KONG, Lei; LUO, Jingchu; GAO, Ge. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. **Nucleic Acids Research**, [S.L.], v. 45, n. 1, p. 1040-1045, 24 out. 2016. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkw982>.

JIN, J. H. *et al.* Genome-wide identification of the AP2/ERF transcription factor family in pepper (*Capsicum annuum* L.). **Genome**, set. 2018. v. 61, n. 9, p. 663–674. Disponível em: <<http://www.nrcresearchpress.com/doi/10.1139/gen-2018-0036>>.

KANEHISA, M.. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic Acids Research**, [S.L.], v. 28, n. 1, p. 27-30, 1 jan. 2000. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/28.1.27>.

KANG, C. *et al.* A novel sweetpotato bZIP transcription factor gene, IbbZIP1, is involved in salt and drought tolerance in transgenic Arabidopsis. **Plant Cell Reports**, 10 Nov. 2019. v. 38, n. 11, p. 1373–1382. Disponível em: <<http://link.springer.com/10.1007/s00299-019-02441-x>>.

LI, Y. *et al.* MYB-CC transcription factor, TaMYBsm3, cloned from wheat is involved in drought tolerance. **BMC Plant Biology**, 15 dez. 2019. v. 19, n. 1, p. 143. Disponível em: <<https://bmcpantbiol.biomedcentral.com/articles/10.1186/s12870-019-1751-9>>.

MILLER, G.; SUZUKI, N.; CIFTCI-YILMAZ, S. Y MITTLER, R. Reactive oxygen species homeostasis and signaling during drought and salinity stresses. **Plant Cell Environ**, v. 33, n. 4, p. 453-67, 2009.

MORALES, R. G. F. *et al.* Caracterização do Tomateiro Submetido ao Déficit Hídrico. **Scientia Agraria**, Curitiba, v. 16, n. 1, p. 09-17, Jan/Fev 2015. ISSN 1519-1125.

NAHAR, K. Y ULLAH, S. M. GRETZMACHER, R. Influence of soil moisture stress on height, dry matter and yield of seven tomato cultivars. **Canadian J. Scientific Industrial Res.**, v. 2, n. 4, p. 160-163, 2011.

NGUYEN, K. H. *et al.* Arabidopsis type B cytokinin response regulators ARR1, ARR10, and ARR12 negatively regulate plant responses to drought. **Proceedings of the National Academy of Sciences**, 15 mar. 2016. v. 113, n. 11, p. 3090–3095. Disponível em: <<http://www.pnas.org/lookup/doi/10.1073/pnas.1600399113>>.

PANG, X. *et al.* *Ammopiptanthus mongolicus* stress-responsive NAC gene enhances the tolerance of transgenic Arabidopsis thaliana to drought and cold stresses. **Genetics and Molecular Biology**, set. 2019. v. 42, n. 3, p. 624–634. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-47572019000400624&tlng=en>.

PEREIRA, S. S. *et al.* Transcription factors expressed in soybean roots under drought stress. **Genetics and Molecular Research**, 2011. v. 10, n. 4, p. 3689–3701. Disponível em: <<http://www.funpecrp.com.br/gmr/year2011/vol10-4/pdf/gmr1475.pdf>>.

PINEDA, B.; GARCÍA-ABELLÁN, J. O.; ANTÓN, T. *et al.* Tomato: Genomic Approaches for Salt and Drought Stress Tolerance. **Improving Crop Resistance to Abiotic Stress**, v. 1 e 2, p. 1085-1120, 2012.

RONQUIST, F.; HUELSENBECK, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. **Bioinformatics**, v. 19, n. 12, p. 1572–1574, 2003.

SAIBO, N. J. M.; LOURENÇO, T. Y OLIVEIRA, M. M. Transcription factors and regulation of photosynthetic and related metabolism under environmental stresses. **Annals of Botany**, v. 103, p. 609-623, 2009.

SAYERS, Eric W; BOLTON, Evan e; BRISTER, J Rodney; CANESE, Kathi; CHAN, Jessica; COMEAU, Donald C; CONNOR, Ryan; FUNK, Kathryn; KELLY, Chris; KIM, Sunghwan. Database resources of the national center for biotechnology information. **Nucleic Acids Research**, [S.L.], v. 50, n. 1, p. 20-26, 1 dez. 2021. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkab1112>. Disponível em: <http://academic.oup.com/nar/article/50/D1/D20/6447242?login=false>. Acesso em: 02 jul. 2023.

SHARMA, R. KUMAR, A. SOLANKE, R. SHARMA, A. TYAGI, A. SHARMA, Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato, **Mol. Genet. Genomics**, v. 284, p. 455-475, 2010.

SHARONI, A. M. *et al.* Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice. **Plant and Cell Physiology**, fev. 2011. v. 52, n. 2, p. 344–360. Disponível em: <<https://academic.oup.com/pcp/article-lookup/doi/10.1093/pcp/pcq196>>.

SILVA, W. G. Da. **Identificação e caracterização dos fatores transcricionais da família AP2 / ERF envolvidos com os mecanismos associados ao estresse hídrico em *Coffea Canephora***. Patos de Minas: Universidade Federal de Uberlândia, 2020.

TANG, Y. *et al.* Overexpression of a MYB Family Gene, OsMYB6, Increases Drought and Salinity Stress Tolerance in Transgenic Rice. **Frontiers in Plant Science**, 18 fev. 2019. v. 10. Disponível em: <<https://www.frontiersin.org/article/10.3389/fpls.2019.00168/full>>.

THE UNIPROT CONSORTIUM. **UniProt**: the universal protein knowledgebase in 2023. **Nucleic Acids Research**, [S.L.], v. 51, n. 1, p. 523-531, 21 nov. 2022. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkac1052>.

THIRUMALAIKUMAR, V. P. *et al.* NAC transcription factor JUNGBRUNNEN1 enhances drought tolerance in tomato. **Plant Biotechnology Journal**, fev. 2018. v. 16, n. 2, p. 354–366. Disponível em: <<http://doi.wiley.com/10.1111/pbi.12776>>.

TRANI, P. E. *et al.* Calagem e Adubação do Tomate de Mesa. **Boletim Técnico IAC**, Campinas, n. 215, 2015. ISSN 1809-7936.

YANG, S. *et al.* A stress-responsive bZIP transcription factor OsbZIP62 improves drought and oxidative tolerance in rice. **BMC Plant Biology**, 17 dez. 2019. v. 19, n. 1, p. 260. Disponível em: <<https://bmcpantbiol.biomedcentral.com/articles/10.1186/s12870-019-1872-1>>.

YANG, Y. *et al.* The Soybean bZIP Transcription Factor Gene GmbZIP2 Confers Drought and Salt Resistances in Transgenic Plants. **International Journal of Molecular Sciences**, 20 Jan. 2020. v. 21, n. 2, p. 670. Disponível em: <<https://www.mdpi.com/1422-0067/21/2/670>>.

YUAN, X. *et al.* Rice NAC transcription factor ONAC066 functions as a positive regulator of drought and oxidative stress response. **BMC Plant Biology**, 25 dez. 2019. v. 19, n. 1, p. 278. Disponível em: <<https://bmcpantbiol.biomedcentral.com/articles/10.1186/s12870-019-1883-y>>.

ZHANG, G. *et al.* Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.). **Journal of Experimental Botany**, Nov. 2008. v. 59, n. 15, p. 4095–4107. Disponível em: <<https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/ern248>>.

ZHAO, Y. *et al.* The Wheat MYB Transcription Factor TaMYB31 Is Involved in Drought Stress Responses in Arabidopsis. **Frontiers in Plant Science**, 28 set. 2018. v. 9. Disponível em: <<https://www.frontiersin.org/article/10.3389/fpls.2018.01426/full>>.

ZHOU, L.; LIU, Y.; LIU, Z.; KONG, D.; DUAN, M. Y LUO, L. Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. **Journal of Experimental Botany**, v. 61, n. 15, p. 4157- 4168, 2010.