UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
REDE RENORBIO

GEORGE ANDRÉ DE LIMA CABRAL

MAPEAMENTO *IN SILICO* DE ELEMENTOS *CIS*-REGULATÓRIOS EM PROMOTORES DE GENES PARA FATORES DE TRANSCRIÇÃO DE *Jatropha curcas* L. (EUPHORBIACEAE) E PERFIS DE EXPRESSÃO DESTES GENES EM RESPOSTA AO ESTRESSE SALINO

RECIFE
2019

GEORGE ANDRÉ DE LIMA CABRAL

MAPEAMENTO *IN SILICO* DE ELEMENTOS *CIS*-REGULATÓRIOS EM PROMOTORES DE GENES PARA FATORES DE TRANSCRIÇÃO DE *Jatropha curcas* L. (EUPHORBIACEAE) E PERFIS DE EXPRESSÃO DESTES GENES EM RESPOSTA AO ESTRESSE SALINO

Tese de doutorado apresentado ao Programa de Pós-Graduação em Biotecnologia - RENORBIO do ponto focal da Universidade Federal de Pernambuco - UFPE, como parte dos requisitos para obtenção do grau de Doutor em Biotecnologia.

Área de Concentração: Biotecnologia na Agropecuária

Orientador: Profº. Dr. Ederson Akio Kido.

RECIFE

2019

Catalogação na Fonte:
Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

Cabral, George André de Lima
Mapeamento in silico de elementos cis-regulatórios em promotores de genes para fatores de transcrição jatropha curcas L. (Euphorbiaceae) e perfis de expressão destes genes em resposta ao estresse salino / George André de Lima Cabral. – 2020.

211 f. : il.

Orientador: Ederson Akio Kido.
Tese (doutorado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-graduação em Biotecnologia / RENORBIO, Recife, 2020.
Inclui referências.

1. Genética vegetal. 2. Interação genótipo-ambiente. 3. Pinhão-manso. I. kido Ederson Akio (Orientador). II. Título.

581.3. CDD (22.ed.) UFPE/CB-2020-082

GEORGE ANDRÉ DE LIMA CABRAL

MAPEAMENTO *IN SILICO* DE ELEMENTOS *CIS*-REGULATÓRIOS EM PROMOTORES DE GENES PARA FATORES DE TRANSCRIÇÃO DE *Jatropha curcas* L. (EUPHORBIACEAE) E PERFIS DE EXPRESSÃO DESTES GENES EM RESPOSTA AO ESTRESSE SALINO

Tese de doutorado apresentado ao Programa de Pós-Graduação em Biotecnologia - RENORBIO do ponto focal da Universidade Federal de Pernambuco - UFPE, como parte dos requisitos para obtenção do grau de Doutor em Biotecnologia.

Aprovado em: 02/08/2019.

BANCA EXAMINADORA

Prof. Dr. Éderson Akio Kido (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Valdir de Queiroz Balbino (Examinador Interno)
Universidade Federal de Pernambuco

Profª Drª. Valesca Pandolfi (Examinadora Interna)
Universidade Federal de Pernambuco

Profª. Drª. Flávia Figueira Aburjaile (Examinadora Externa)
Universidade Federal de Pernambuco

Profº Dr. Ricardo Yara (Examinador Externo)
Universidade Federal de Pernambuco

Dedico este trabalho aos meus pais João Cabral Filho e Gilvaneide de Lima Cabral, por acreditar e ajudar sempre nas minhas escolhas, pelos incentivos incessantes, fruto da confiança e do amor incondicional.

RESUMO

O pinhão manso (*Jatropha curcas* L. - Euphorbiaceae) é considerada uma excelente fonte de biodiesel disponível que apresenta algumas vantagens sobre as chamadas oleaginosas produtoras de biodiesel, como o alto conteúdo de óleo nas sementes (30-50%), relativa facilidade de conversão à biocombustível e ausência de competição com culturas de alimento, em função de ser uma espécie que ocorre de forma espontânea em diversas regiões tropicais de clima árido e semiárido. No entanto, para que *J. curcas* seja uma alternativa viável para a expansão da fronteira agrícola em diversas localidades do Nordeste brasileiro, dois problemas devem ser confrontados: a salinidade e a salinização dos solos. Solos salinos e sódicos podem ocorrer naturalmente em regiões semiáridas, e a irrigação nessas áreas, se feita de forma não apropriada, acarreta problemas de salinização dos solos, o que diminui a produção e a produtividade das culturas. Neste trabalho, após uma análise global de genes diferencialmente expressos em raízes de dois acessos de pinhão-manso sob estresse salino (150 mMol / 3h), membros de várias famílias de FTs foram identificados através da técnica de sequenciamento RNA-seq em um dos acessos, onde o perfil de expressão foi descrito e os FTs responsáveis por ampla quantidade de regulações em termos de alvos super-representativos foram destacados. Estes resultados auxiliaram na compreensão dos mecanismos de respostas à salinidade e tolerância em pinhão-manso, sugerindo grande potencial biotecnológico para disponibilização de valiosos recursos, como candidatos potenciais para transgenia, visando aumentar à eficácia da resposta à tolerância específica ao estresse estudado. Adicionalmente, uma análise na região regulatória 5'UTR de genes de fator de transcrição provenientes de ESTs de *J. curcas* revelou prováveis elementos *cis*-regulatórios envolvidos na regulação e expressão de genes de fatores de transcrição de *J. curcas*, o que permitiu gerar um mapa *in silico* de elementos *cis*-regulatórios, o que pode contribuir no desenvolvimento de ferramentas biotecnológicas inovadoras, tais como proposição de promotores sintéticos associados a genes responsivos a estresses, bem como no desenvolvimento de marcadores moleculares funcionais (baseados em mRNAs e cDNAs) ou mesmo genômicos (associados a potenciais polimorfismos), que possam discriminar genótipos potencialmente tolerantes, para uso futuro em seleção assistida.

PALAVRAS-CHAVE: genômica funcional. pinhão-manso. RNA-seq. salinidade.

ABSTRACT

Physic nut (*Jatropha curcas* L. - Euphorbiaceae) is considered an excellent source of available biodiesel that has some advantages over so-called biodiesel oilseeds, such as high oil content in seeds (30-50%), relative ease of conversion. biofuel and lack of competition with food crops, as it is a species that occurs spontaneously in various tropical regions of arid and semiarid climate. However, for *J. curcas* to be a viable alternative for the expansion of the agricultural frontier in several localities of Northeast Brazil, two problems must be confronted: salinity and soil salinization. Saline and sodic soils can occur naturally in semi-arid regions, and irrigation in these areas, if not properly done, causes problems of soil salinization, which decreases crop yield and productivity. In this work, after a global analysis of differentially expressed genes in roots of two *J. curcas* accessions (150 mMol / 3h), members of several FT families were identified by the RNA-seq sequencing technique in one of the accessions, where the expression profile was described and the TFs responsible for the large amount of overrepresentative targeting regulations were highlighted. These results helped to understand the mechanisms of response to salinity and tolerance in Jatropha curcas, suggesting great biotechnological potential for the availability of valuable resources, as potential transgenic candidates, aiming to increase the effectiveness of the stress tolerance response studied. Additionally, an analysis in the 5'UTR regulatory region of transcription factor genes from *J. curcas* ESTs to identify *cis*-regulatory elements was focused, revealing likely elements involved in the regulation and expression of *J. curcas* transcription factor genes, which allowed the generation of an *in silico* map of *cis*-regulatory elements, may contribute to the development of innovative biotechnological tools, such as the proposition of synthetic promoters associated with stress-responsive genes, as well as the development of functional molecular markers. (based on mRNAs and cDNAs) or even genomic (associated with potential polymorphisms) that may discriminate potentially tolerant genotypes for future use in assisted selection.

KEYWORDS: functional genomics. jatropha curcas. RNA-seq. salinity.

LISTA DE FIGURAS

Revisão Bibliográfica

Artigo I - *De novo* RNA-Seq transcriptome analysis of *J. curcas* accessions under salt stress

Artigo II - Physic nut transcription factors: identification and transcriptional modulation under salt stress

Artigo III - Dominant position of *cis*-acting regulatory elements in promoters of physic nut (Jatropha curcas L.) transcription factors genes

LISTA DE TABELAS

Artigo I - *De novo* RNA-Seq transcriptome analysis of *J. curcas* accessions under salt stress

Artigo II - Physic nut transcription factors: identification and transcriptional modulation under salt stress

Artigo III - Dominant position of *cis*-acting regulatory elements in promoters of physic nut transcription factors genes

SUMÁRIO

1 INTRODUÇÃO

O pinhão-manso (*Jatropha curcas* L.) é uma espécie perene, monóica, da família das Euforbiáceas (mamona, mandioca, seringueira, etc), de crescimento rápido, caducifólio, com frutos do tipo cápsula ovóide (trilocular, com uma semente por lóculo), cujas sementes apresentam potencial para produzir acima de 1.200 kg de óleo por hectare. Por ser um arbusto que vegeta espontaneamente em diversas regiões do Brasil, embora não seja originária do país, o pinhão-manso poderia contribuir para a expansão da fronteira agrícola, com a incorporação de áreas marginais, sem competir com a produção de alimentos. Entretanto, nas condições da região Nordeste, solos salinos e sódicos podem ocorrer naturalmente, onde a associação com a elevada evapotranspiração, a baixa pluviosidade e as propriedades físicas e físico-químicas de solos desfavoráveis, leva à toxidez das plantas, acarretando na diminuição do crescimento e da produtividade esperada.

Os fatores de transcrição (FTs) são importantes proteínas que regulam a expressão de genes alvos, cuja afinidade e seletividade dos FTs é influenciada pelas estruturas secundárias dos domínios de ligação ao DNA que contatam bases dos elementos cis-regulatórios presentes nos promotores dos genes. Atuando como moduladores da expressão gênica, os FTs direcionam vias regulatórias que definem estágios do desenvolvimento e respostas a estímulos ambientais e agentes estressores.

Pesquisas genéticas envolvendo diversas espécies vegetais têm mostrado que as plantas respondem aos estresses ambientais com alterações fisiológicas e metabólicas para uma maior tolerância, de acordo com os genes presentes no organismo em estudo e, principalmente, na maneira como eles são regulados. Além disso, muitos dos genes de respostas aos diferentes estresses em plantas são compartilhados, em função da sintenia, de tal forma, que um estudo nesta direção acabaria beneficiando outros em andamento nos programas de melhoramento genético ao redor do mundo. Neste sentido, explorar a variabilidade genética disponível, para identificar genótipos elites, portadores de genes favoráveis para produção, mesmo em condições não ideais, é objetivo do melhoramento genético, e isso requer novas metodologias.

Dentre as técnicas de estudos globais dos genes, a técnica RNA-seq, para estudos em transcriptômica, gera perfis transcricionais baseados em *reads* de cerca de 30 - 400 pb, e que são geradas a partir de RNAs poli A^+ de plantas em condições experimentais sob estresse *versus* controle (sem estresse). Neste trabalho, a técnica

RNA-seq foi aplicada em resposta ao estresse salino (150 mMol de NaCl). Para o pinhão-manso, atualmente, é disponibilizado na base de dados pública NCBI (http://www.ncbi.nlm.nih.gov/) cerca de 150 mil sequências nucleotídicas e quase 30 mil proteica, além de um genoma sequenciado também disponível, o que auxiliou na anotação dos fatores de transcrição e da prospecção de elementos *cis*-regulatórios em promotores dos respectivos genes. Essas regiões promotoras, que ficam 5' (*upstream*) do sitio de início da transcrição dos genes, são moduladores de expressão de genes reconhecidos pelos fatores de transcrição. Percebe-se, então, que o conhecimento desses elementos é essencial para o entendimento da regulação dos genes e é fundamental para interpretar e modelar as respostas da célula aos diversos estímulos, além de contribuir para o desenvolvimento de ferramentas biotecnológicas inovadoras, tais como a proposição de promotores sintéticos de plantas, associados a genes responsivos a estresses específicos.

2 OBJETIVO GERAL

Identificação de elementos cis-regulatórios em promotores de genes de fatores de transcrição de pinhão-manso (*J. curcas* L) e resposta destes genes, com base no transcriptoma RNA-seq, ao estresse de salinidade.

2.1 OBJETIVOS ESPECÍFICOS

• Realizar a prospecção de elementos *cis*-regulatórios em sequencias genômicas de pinhão-manso (*J. curcas*), a partir do ancoramento de ESTs disponíveis para a espécie, oriundos de banco público (NCBI).

• Identificar genes de fatores de transcrição associados à tolerância ao estresse salino, tendo como base a expressão diferencial do transcriptoma RNA-seq.

• Validar a expressão, a partir da técnica RT-qPCR, de genes para fatores de transcrição associados com a expressão diferencial RNA-seq, em resposta ao estresse salino.

3 REVISÃO BIBLIOGRÁFICA

3.1 O PINHÃO MANSO

3.1.1 Aspectos gerais da espécie e características botânicas

Carl Linnaeus foi o primeiro designar cientificamente o pinhão manso em 1753, a espécie *Jatropha curcas*, o qual pertence ao Reino Plantae, Divisão Embryophyta, Classe Spermatoporida, Ordem Malpighiales, Família Euphorbiaceae, Gênero Jatropha (PEIXOTO, 1973). Também conhecida popularmente como pinhão da índia, pinhão de purga, pinhão de cerca, pinhão branco, figo do inferno, pinhão paraguaio, pinhão croá, /pinhão dos barbados, purgante de cavalo, mandobi guaçu, medicineira, purgueira ou, simplesmente purga (TOMINAGA et al., 2007). A designação científica do gênero Jatropha deriva do grego “iatrós” (o doutor) e “trophé” (comida), o que sugere sua utilização na etnobotânica medicinal; o epíteto específico curcas é a designação vulgar da purgueira em Malabar, Índia (HELLER, 1996). O gênero Jatropha encontra-se representado por cerca de 175 espécies, distribuídas nas regiões semiáridas da América Tropical, Ásia e África (AUGUSTUS et al., 2002), sendo anteriormente dividido em dois subgêneros (Jatropha e Curcas), com 10 secções e 10 subsecções, adaptados às espécies Euroasiáticas, Africanas e Americanas (DEHGAN e WEBSTER, 1979). O subgênero Jatropha possui maior distribuição, tendo espécies encontradas na África, Índia, América do Sul, Antilhas, América Central e Caribe. Já o subgênero Curcas, com exceção do pinhão-manso está restrito ao México e regiões adjacentes (deserto do Saara, Arizona e Texas) (DEHGAN e WEBSTER, 1979). A família Euphorbiaceae é representada por aproximadamente 320 gêneros e 8.000 espécies, distribuindo-se principalmente nos trópicos e subtrópicos, em diferentes habitats (WEBSTER, 1994).

O pinhão manso é uma planta arbustiva, caducifólia, de crescimento rápido, cuja altura pode atingir mais de cinco metros e diâmetro do tronco de 20 cm, com poucas ramificações, mas pode chegar até cerca de 80 cm (CORTESÃO, 1956). A espécie é monóica, isto é, possui flores de sexo separado, com flores masculinas e femininas na mesma inflorescência e eventualmente com presença de flores hermafroditas (BRASIL, 1985) (Figura 1a-b). As flores são pequenas de coloração amarelo-esverdeadas, sendo as flores masculinas (Figura 1c-d), com dez estames, surgindo em maior número e localizadas nas extremidades superiores das ramificações, enquanto as flores femininas

(Figura 1e-f) são menos numerosas e encontram-se nas bases das ramificações, apresentando um pedúnculo longo, não articulado, com três células elípticas, ovário com três carpelos (Figura 5g), cada um com um lóculo que produz um óvulo, e com três estigmas bifurcados (DIAS et al., 2007). O período da florada é longo, e a polinização é feita por abelhas, formigas, moscas e outros insetos (JONGSCHAAP et al., 2007). As flores de uma mesma inflorescência abrem em dias diferentes, sendo que a floração das flores femininas ocorre primeiro que das flores masculinas da mesma inflorescência, o que favorece a polinização cruzada (SATURNINO et al., 2005). As folhas são verde-escuras e brilhantes, largas e alternas, em forma de palma, com três a cinco lóbulos e pecioladas, com nervuras esbranquiçadas e salientes na face inferior (ARRUDA et al., 2004). A filotaxia em espiral com os pecíolos longos e esverdeados, dos quais partem as nervuras divergentes. O pecíolo é longo e esverdeado, do qual partem as nervuras divergentes e o caule é liso, lenhoso, com medula desenvolvida, mas pouco resistente. O floema é formado por longos canais que se estendem até as raízes, nos quais circula o látex, suco leitoso que ocorre em abundância em qualquer ferimento da planta. O tronco é dividido desde a base, em ramos compridos (SATURNINO et al., 2005; TOMINAGA et al., 2007). Os frutos são do tipo cápsula ovóide, trilocular, com 1,5 a 3,0 cm de diâmetro, e são formados por um pericarpo ou casca dura e lenhosa, pesando entre 1,5 e 2,8g. No geral, 53 a 62% do peso do fruto é representado pelas sementes e 38 a 47% pela casca (HEIFFIG e CÂMARA, 2006). As sementes são ovaladas, endospérmicas, com tegumento rígido, quebradiço e de fratura resinosa. Debaixo do invólucro da semente existe uma película branca cobrindo a amêndoa, o endosperma ou albúmen, que é abundante e oleaginoso, com o embrião provido de dois largos cotilédones achatados e foliáceos. Quando secas, as sementes apresentam de 1,5 a 2,0 cm de comprimento e 1,0 a 1,5 cm de largura. O peso varia entre 0,551 e 0,797 g, dependendo da variedade e dos tratos culturais. A casca representa de 33,7 a 45% e a amêndoa de 55 a 66% do peso da semente. Na semente ainda é encontrado: 7,2% de água e 55,3% de açúcar, amido, albuminóides e materiais minerais, sendo 4,8% de cinzas e 4,2% de azoto e apresentando um teor de óleo que varia entre 33 e 38 % (DIAS et al., 2007). O sistema radicular do pinhão manso é do tipo pivotante, com uma raiz principal que atinge grandes profundidades, além de grande quantidade de raízes laterais, responsáveis pela nutrição da planta. De uma forma geral, a profundidade do sistema

radicular é equivalente à altura da planta, assim como o diâmetro de exploração de solo (AVELAR, 2006).

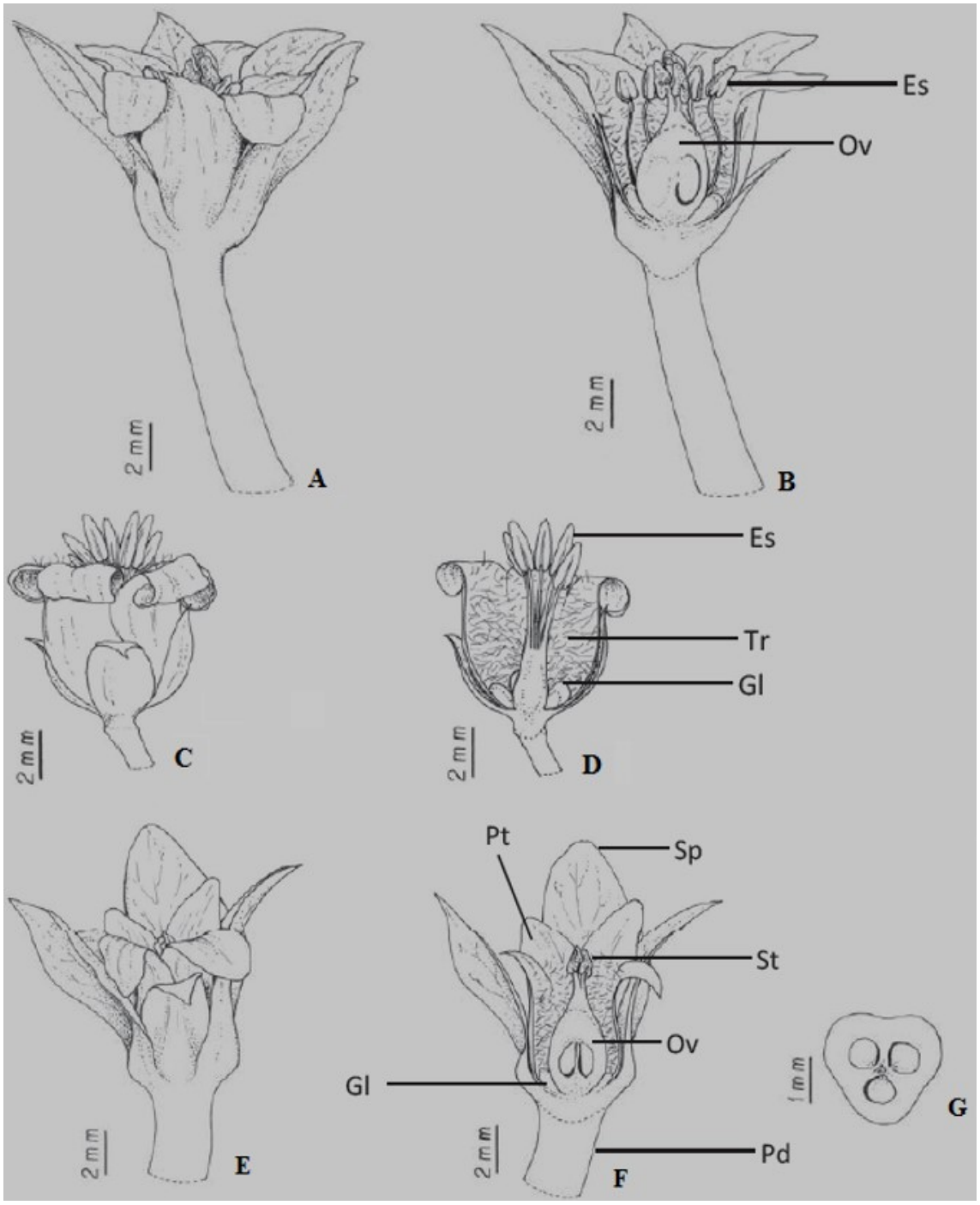


Figura 1. Tipos florais em *Jatropha curcas* L. Flores hermafroditas = A e B (Corte longitudinal). Flores masculinas = C e D (Corte longitudinal). Flores femininas = E e F (Corte longitudinal). G = Ovário (Corte transversal). Es-estame; Gl-glândula; Ov-ovário; Pt-pétala; Pd-pedúncolo; Sp-sépala; St-estigma; Tr-tricomas. Adaptado a partir de Brasileiro et al. (2012).

3.1.2 Centro de origem e distribuição geográfica

Vários cientistas tentaram identificar a origem da *Jatropha curcas*, mas ainda não existe uma definição precisa e as informações obtidas continuam controversas. Muitos autores relatam que o pinhão manso é originado do México e América Central, mas encontrado em abundância em áreas tropicais e subtropicais na África e Ásia, onde supostamente foi disseminado por navegantes portugueses para países latino-americanos, africanos e asiáticos pelas ilhas de Cabo Verde e Guiné-Bissau, durante a primeira metade do século XIX, quando a espécie já era um importante produto de exportação (RAO et al., 2008). Outros autores colocam que o pinhão manso é originado da América do Sul, e que foi introduzido nas Ilhas do Arquipélago de Cabo Verde em 1873, onde foi posteriormente disseminado para todas as regiões tropicais, ocorrendo em vários países como Índia, Cabo Verde, Malásia, Tailândia, Filipinas, além de algumas regiões do Brasil (BELTRÃO, 2005). Arruda et al. (2004) apontam que o pinhão manso possivelmente é originário do Brasil, tendo sido introduzido por navegadores portugueses nas ilhas do Arquipélago Cabo Verde e Guiné, de onde foi disseminado pelo continente Africano e atualmente está distribuído em todas as regiões tropicais do globo. Dias et al. (2012), com base em registros históricos e conceitos genéticos, sugeriram que o centro de origem e o de domesticação de *Jatropha curcas* é o México, e que não há registros do uso de *J. curcas* e seus produtos anteriores aos povos Olmeca do México, que viveram 3500-5000 anos. Além disso, a existência de genótipos não-tóxicos, que só existem neste país, juntamente com estudos de DNA, também sugerem fortemente que centro de domesticação da espécie é o México. Pecina-Quintero et al. (2014), com base em marcadores moleculares de DNA, confirmaram os relatos de Dias et al. (2012), demonstrando que o México detém grande diversidade genética de *J. curcas*, mais precisamente no estado de Chiapas, sendo o centro de origem mais provável, onde seu germoplasma possui características especiais não partilhadas com outros genótipos encontrados no resto do país. No entanto, o germoplasma mais domesticado é encontrado fora de Chiapas, sugerindo que os centros de origem e domesticação estão em estados diferentes, mais próximos da bacia do Golfo do México, como em Veracruz, Puebla, Hidalgo e Yucatan, uma vez que apresentam indivíduos com ausência ou baixo teor de ésteres de forbol, um derivado de diterpeno tetracíclico.

Devido a sua rusticidade, resistência a pragas e a longas estiagens, a distribuição geográfica de *J. curcas* é bastante vasta, adaptando-se a uma gama de condições edafoclimáticas muito variáveis e se desenvolvendo tanto nas regiões tropicais secas como nas zonas equatoriais úmidas, assim como pode povoar áreas de solos arenosos e pouco férteis, como em terrenos áridos e pedregosos, podendo suportar longos períodos de seca (ARRUDA et al., 2004). No Brasil, a espécie é encontrada desde o Maranhão até o Paraná, propagando-se com mais facilidade nos estados do Nordeste, assim como em Goiás e Minas Gerais (DRUMMOND et al., 2008).

3.1.3 A importância econômica do óleo do pinhão manso

A utilização do óleo extraído das sementes do pinhão manso é antiga. Em Cabo Verde, o óleo de *J. curcas* foi utilizado para iluminação de casas, e a partir de 1936, começou a exportação de sementes para Lisboa para a mesma finalidade (FERRÃO et al., 1983). No Brasil, a principal fonte de óleos vegetais atualmente é a soja, seguido do dendê (palma) e outras oleaginosas anuais como o girassol, a mamona, a canola, o algodão e o amendoim (PAZETO, 2013). No entanto, algumas espécies são indicadas como potencialidades para o futuro, como o pinhão-manso, que apresenta maior estabilidade oxidativa que o óleo de soja, menor viscosidade e densidade quando comparada ao da mamona e melhores propriedades a frio que o óleo de palma (TAPANES et al., 2008).

Em relação à produtividade, dependendo do espaçamento, a produção de sementes de pinhão manso pode passar dos 6.000 kg por hectare, sendo possível produzir mais de 2.000 kg por hectare de óleo (TOMINAGA et al., 2007). De acordo com este autor, com o aprimoramento do sistema de produção e melhoramento genético, o pinhão manso pode produzir acima de 4.000 kg de óleo por hectare, mas em função da região do plantio, idade da cultura, bem como da quantidade de chuva e da fertilidade do solo a produtividade do pinhão manso pode variar muito. O óleo do pinhão manso é composto por aproximadamente 80% de ácidos graxos insaturados e 20% de ácidos saturados, perfil adequado para a produção de biodiesel, sendo o ácido oléico o de maior abundância, seguido do ácido linoléico, palmítico e esteárico (LU et al., 2009). Aproximadamente 1% de compostos não-saponificáveis (tocoferois, fosfolipídeos, carotenoides e produtos oxidados) são encontrados no óleo (TAPANES et al., 2008). O óleo pode ser extraído e convertido em biodiesel, e os resíduos resultantes podem

também ser convertidos em energia. A torta contém aproximadamente 8% de óleo, o qual é reextraído com solventes orgânicos, geralmente hexano, sendo o farelo residual ensacado para aproveitamento como fertilizante natural (BRASIL, 1985).

O biodiesel é um éster de ácido graxo obtido a partir da reação química de transesterificação de qualquer triglicerídeo do óleo vegetal com um álcool de cadeia curta (metanol ou etanol), na presença de um catalisador ácido ou básico. Como resultado, obtém-se além do éster (biodiesel), a glicerina (TAPANES et al., 2008). A glicerina é utilizada como matéria-prima da indústria de cosméticos, produtos farmacêuticos, alimentação, bebidas, filmes de celulose, papel, sabões e resinas (OLIVÉRIO, 2006). Comparado ao óleo diesel derivado de petróleo, o biodiesel pode reduzir em 78% as emissões de gás carbônico, reduzir em 90% as emissões de fumaça e praticamente eliminar as emissões de óxido de enxofre (ACCARINI, 2006).

3.2 ESTRESSE SALINO

3.2.1 Extensão da salinidade e interferência sobre os solos

Cerca de 19% dos 230 milhões de hectares de terras irrigadas do globo terrestre são afetadas pela salinidade ou pela ocorrência de sodicidade (Organização das Nações Unidas para a Agricultura e Alimentação - FAO, 2008). Em regiões áridas e semiáridas do mundo, as concentrações de sais podem atingir valores elevados, prejudicando o solo e as plantas (MEDEIROS et al., 2010). Isso decorre principalmente devido às características edafo-climáticas dessas regiões, relacionadas com o intemperismo físico e químico das rochas, materiais geológicos e biológicos, e que apresentam elevados níveis de temperatura, evaporação e transpiração, além da baixa e irregular precipitação, que juntamente com o manejo inadequado da água, contribuem ainda mais para o aumento da salinidade local e consequente prejuízo às culturas (RENGASAMY, 2010). No Brasil, estima-se que há, aproximadamente, nove milhões de hectares com problemas de salinidade, sendo a maior parte desta área localizada nos perímetros irrigados do Nordeste (CRUZ et al., 2006). Estas áreas tem o uso da irrigação intensificada, uma vez que são regiões de alta produtividade agrícola, e a irrigação desempenha um papel fundamental para garantir a alta produtividade, mesmo em períodos de déficit hídrico (LIMA-JUNIOR et al., 2010).

Grande parte dos eventos de salinização e sodificação dos solos ocorrem devido a um processo natural conhecido como salinização primária, sem a interferência antropogênica (MUNNS, 2005). A intemperização química dos minerais e rochas da crosta terrestre é a principal fonte responsável pela liberação e distribuição de íons para o halomorfismo naturalmente induzido ao solo (DAKER, 1988). O aumento da concentração de sais na superfície dos solos em zonas áridas e semiáridas acontece por meio da ascensão dos íons por capilaridade, devido à alta demanda evaporativa e quanto mais alto for o nível freático, ou por serem transportados pelas águas de outros locais (SAHI et al., 2003). A água deposita seus sais excedentes na camada superficial do solo e, com o tempo, estes precipitam e se acumulam, tornando o solo salino (CHINNUSAMY et al., 2005). Além disso, a escassez da precipitação pluvial dificulta a lixiviação dos sais localizados na camada arável do solo (RIBEIRO et al., 2009). No entanto, um dos maiores problemas da salinidade, principalmente em regiões áridas e semiáridas, tem sido também ocasionado pelo processo conhecido como salinização secundária, que ocorre devido ao acumulo de sal proveniente da utilização incorreta de técnicas de irrigação associadas à drenagem deficiente e à presença de águas subsuperficiais ricas em sais solúveis localizadas em baixa profundidade (ASHRAF e FOOLAD, 2007). A aplicação de fertilizantes de forma excessiva e pouco parcelada ao longo do ciclo cultural pode ainda intensificar este processo (OLIVEIRA et al, 2010).

A caracterização dos solos quanto à salinidade é baseada na condutividade elétrica do extrato saturado do solo (CE_{ES}), percentual de sódio trocável (PST) e pH, ainda que sejam analisados outros parâmetros, como a capacidade de troca de cátions (CTC), razão de adsorção de sódio (RAS) e a proporção no solo de teores de íons de $Ca2^{+}$, $Mg2^{+}$, Na^{+}, Cl^{-}, SO^{2-}, HCO^{3-}, K^{+}, $CO3^{2-}$ e NO^{3-} (SANTOS e MURAOKA, 1997; Oliveira et al, 2010). Os solos são classificados como salinos quando apresentam pH inferior a 8,5, CE_{ES} superior a 4 dS m^{-1} e PST inferior a 15%; salinos-sódico quando possuem pH próximo de 8,5, CE_{ES} superior a 4 dS m^{-1} e PST superior a 15%; e sódicos quando possuem pH, em geral, na faixa de 8,5 a 10, CE_{ES} inferior a 4 dS m^{-1} e PST superior a 15% (SCHERER et al., 1996).

A interação eletroquímica existente entre os sais e as partículas constituintes do solo promove o efeito da salinidade sobre o mesmo, refletindo o grau de estabilização de seus agregados (SPERA et al., 2008). Neste processo, o acúmulo de cátions de menor valência sobre a superfície das partículas do solo (argilas) pode ocasionar a expansão

das mesmas a ponto de dispersá-las, em função do enfraquecimento de suas forças de ligação (LIMA, 1997). As partículas de argila, juntamente com a água de percolação, são eluviadas após serem dispersas, elevando a densidade do solo através da ocupação dos espaços porosos (SPERA et al., 2008). Limitações na disponibilidade de ar e na condutividade hidráulica para as plantas podem ser ocasionados pela redução da porosidade natural e maior adensamento do solo, além de interferir na atividade biológica de microrganismos e facilitar o processo de erosão (SANTI et al., 2002).

3.2.2 Mecanismos de tolerância ao efeito iônico e toxicidade

Altas concentrações de sais no solo causam estresse hiperiônico, e a consequência deste estresse, juntamente ao estresse osmótico, acaba por ser altamente danosa às plantas (ZHU, 2001). Um dos processos mais importantes para a tolerância das plantas em ambientes salinos é o controle do balanço iônico, determinado pelo desenvolvimento de mecanismos de acumulação de solutos, através da absorção de íons do meio externo, ou pela mobilização de seus próprios constituintes orgânicos (YAMAGUCHI e BLUMWALD, 2005). Em ambientes hipersalinos o transporte passivo de Na^+ para o interior celular é favorecido em função da formação de um gradiente eletroquímico, devido às elevadas concentrações extracelulares do íon, possibilitando o transporte passivo do íon para o interior celular (BLUMWALD, 2000). O potencial da membrana plasmática é dissipado pela entrada de Na^+ na célula, facilitando a entrada de Cl^- contra o gradiente eletroquímico, que é mediada por um transportador ativo (MUNNS, 2005).

Dois processos preponderantes que previnem a parada de crescimento das plantas em ambientes salinos são a compartimentalização dos íons de Na^+ no vacúolo e a exclusão destes íons excedentes da célula, constituindo assim mecanismos de tolerância à salinidade (TAIZ e ZEIGER, 2004). Mesmo ocorrendo o risco de toxidez iônica e/ou deficiência nutricional, o controle do armazenamento dos íons no vacúolo permite que a planta mantenha a turgescência, sem afetar os sistemas enzimáticos do citoplasma, e sem custos energéticos para síntese de solutos orgânicos (MARTINOIA et al., 1986). Ainda de forma alternativa à compartimentalização no vacúolo, os sais podem ser transportados para a parede celular, o que, no entanto, pode resultar na desidratação da célula (MUHLING e LAUCHLI, 2002).

Os processos de compartimentalização de Na^+ no vacúolo e eliminação de altas concentrações intracelulares deste íon requerem transporte dependente de energia, e são realizados com a regulação de proteínas transportadoras que integram a membrana plasmática e a membrana do vacúolo (MANSOUR et al., 2003). Tais proteínas estão envolvidas no controle do potencial de membrana e na transdução de sinais em plantas (ZIMMERMANN et al., 2001). As células vegetais tendem a manter alta razão entre os íons K^+/Na^+ no citosol em condições fisiológicas, onde os níveis de K^+ são relativamente altos e os níveis de Na^+ são baixos (MUNNS et al., 2006). O equilíbrio do K^+ e do Na^+ intracelular é importante para a atividade de muitas enzimas citosólicas, manutenção do potencial de membrana e do potencial osmótico que regula o volume celular (ZHU, 2003). A manutenção da homeostase do Na^+ e do K^+ torna-se ainda mais crucial sob estresse salino, pois os níveis de Na^+ e Cl^-, predominantes em ambientes salinos, aumentam nos tecidos, mas os níveis de K^+ e Ca^{2+} diminuem, devido à competição catiônica pelos sítios de adsorção localizados nas proteínas transportadoras de alta afinidade pelo íon K^+ ou daquelas de baixa afinidade, que são fortemente influenciadas pelo íon Ca^+ (MARSCHNER, 1995). Um dos principais efeitos nocivos das elevadas concentrações do íon Na^+ é o de deslocar o íon Ca^+ dos sítios de ligação das membranas plasmáticas intracelulares, com a consequente perda da seletividade iônica das células radiculares, impedindo assim homeostase de Ca^{2+} nestas células (MAATHUIS e AMATMANN, 1999). O íon potásio é um dos íons mais abundantes no meio intracelular, sendo um nutriente essencial para as plantas, pois contribui significativamente para a manutenção do potencial osmótico do vacúolo e turgor celular, é um ativador de reações enzimáticas vitais e é essencial na síntese de proteínas (TAKAHASHI et al., 2007). Na síntese proteica, há exigência do K^+ na ligação do tRNA aos ribossomos, a qual necessita de altas concentrações desse íon para ocorrer (BLAHA et al., 2000).

As proteínas transportadoras H^+-ATPases (V-ATPase) e as H^+-PPases (VPPase) estão entre as primeiras proteínas transportadoras de Na^+ identificadas em plantas, responsáveis por gerar um gradiente de prótons de H^+ no vacúolo e por auxiliar na manutenção de altas concentrações de K^+ e baixas concentrações de Na^+ no citosol (HASEGAWA et al., 2000). As V-ATPases são constituídas por um complexo multienzimático composto por 10 subunidades (RATAJCZAK, 2000). Já as V-PPases têm sido encontradas unicamente em plantas vasculares e algas (SAKAKIBARA et al.,

1996). Além da compartimentalização do Na^+, também é sugerido que a planta envia íons Cl^- excedentes no citoplasma para o interior do vacúolo através de transportadores antiporter de Cl^-/H^+ envolvidos nos processos de compartimentalização deste ânion, através do acoplamento e movimentação do íon a favor do gradiente de prótons (HASEGAWA et al., 2000). Transportadores de Cl^- têm sido isolados em algumas espécies e uma família denominada CLC tem sido mais estudada (BARBIER-BRYGOO et al., 2000). Uma proteína com participação no seqüestro intracelular de Na^+ para o interior do vacúolo, denominada NHX1, foi encontrada em *Arabdopisis thaliana* (NASS et al., 1997). O gene AtNHX1 que codifica esta proteína no tonoplasto atua na compartimentalização de Na^+ para dentro do vacúolo e sua expressão sofre aumentos consideráveis sob estresse mediado por NaCl e KCl, sugerindo uma rota na resposta à estresse osmótico e iônico (GAXIOLA et al., 1999). A atividade dessas proteínas antiporteres tem sido relacionada a outras espécies vegetais como: arroz (FUKUDA et al., 1999), algodão (WU et al., 2004), trigo (BRINI et al., 2005), *Atriplex gmelini* (Hamada et al., 2001) e *Brassica napus* (WANG et al., 2003). Outras proteínas identificadas em *Arabidopsis thaliana*, denominadas AKT1 (*Arabidopsis K^+ transporter 1*) e KAT1 (*K^+ Arabidopsis thaliana channel 1*), são seletivas para K^+ em concentrações fisiológicas normais dos íons Na^+ e K^+ no meio extracelular, porém, quando ocorre aumento nos níveis de Na^+ no meio externo, tais carreadores passam a transportar sódio em vez de potássio para dentro das células (BLUMWALD et al., 2000). HKT1 é outro transportador que foi inicialmente caracterizado de alta afinidade para a aquisição de K^+ (SCHACHTMAN e SCHROEDER, 1994). Posteriormente, esta proteína foi definida em uma variedade de espécies vegetais como um simporter para Na^+/K^+ (RUS et al., 2001). Foi demonstrando em ensaio que o transporte de K^+ era estimulado pela proteína transportadora HKT1 em baixas concentrações de Na^+ no meio externo e, ao contrário, o transporte de K^+ pela proteína era inibido com o aumento dos níveis de Na+ no meio extracelular, induzindo a aquisição de Na^+ para o meio intracelular (RUBIO et al., 1995). A família de proteínas transportadoras Kup/Hak (Kup - *K^+ Uptake* / Hak - *High Affinity K^+ Transporter*) apresentam maior afinidade ao K^+ que HKT1, no entanto, esses transportadores também são permeáveis a íons Na^+ em ambientes salinos (TAKAHASHI et al., 2007).

A via de sinalização SOS (*Salt Overly Sensitive*) é um dos principais componentes envolvidos na tolerância ao estresse salino (GUPTA e HUANG, 2014). É

composta por três proteínas (SOS1, SOS2 e SOS3) cuja interação culmina na exclusão de Na^+ do citoplasma e, ainda, influenciam na ativação de outros transportadores que atuam para o restabelecimento da homeostase iônica (MUNNS, 2005). O aumento transiente dos íons de cálcio vindos do apoplasto e carreados para o interior celular dispara uma cascata de sinalização que inicia o processo de adaptação celular ao estresse (LIU e ZHU, 1998). Quando os níveis de concetração intracelular de Ca^{2+} estão altos, esses íons se ligam à proteína SOS3 através de um domínio de ligação ao Ca^{2+} que altera sua conformação (Figura 2) e a habilita para interagir com a proteína SOS2, do tipo serina/treonina quinase, que possui um domínio catalítico amino-terminal e um domínio regulatório carboxiterminal (ISHITANI et al., 2000). A proteína SOS2 permanece em estado inativo por um sistema de auto-inibição até proteína SOS3 interagir com a SOS2 tornando-a uma proteína quinase ativa, em um sistema dependente de Ca^{2+} (ALBRECHT et al., 2001). No passo seguinte da via a proteína SOS2 fosforila a SOS1 ativando-a, que inicia um mecanismo de exclusão de Na^+ do citoplasma através de atividade antiporter de Na^+/H^+ (SHI et al., 2000). É proposto que o complexo SOS2-SOS3 influenciem na atividade de outros transportadores de íons além da ativação de SOS1, como na inativação da proteína HKT1 ou na inibição da expressão do gene que a codifica, e na ativação de NHX (antiporter Na^+/H^+ de vacúolo), que resulta no sequestro do excesso de íons Na^+, contribuindo, mais uma vez, para a homeostase iônica (MAHAJAN e TUTEJA, 2005).

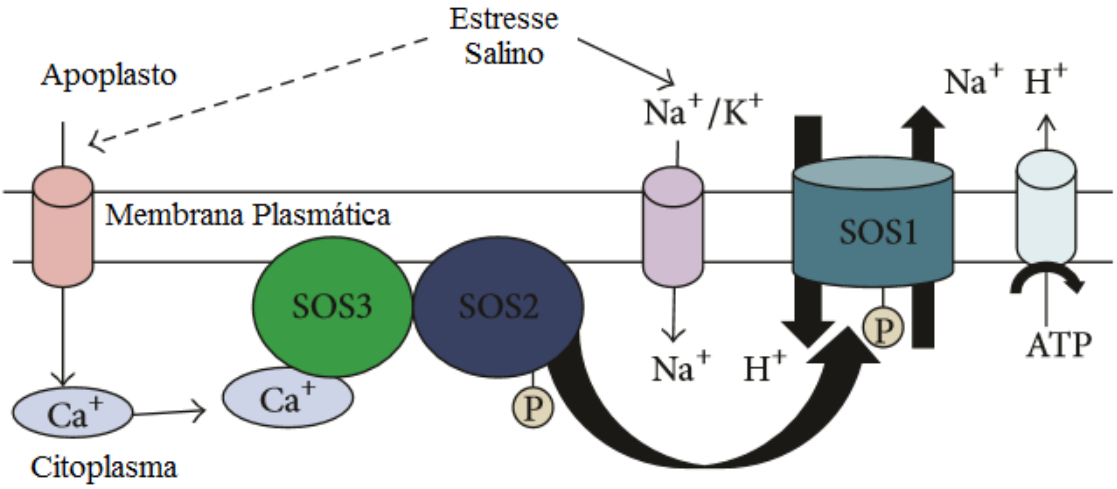


Figura 2. Modelo da regulação da homeostase iônica via SOS para respostas ao estresse salino. O estresse salino é percebido por receptores presentes na membrana plasmática da célula que inclui uma perturbação citosólica de Ca^{2+}, que interage com SOS3 e a conformação da proteína é modificada de uma maneira que a conecta com SOS2. Essa interação libera SOS2 de sua auto-inibição resultando em um complexo com SOS3 que fosforila SOS1, um antiporter Na^+/H^+, que promove o efluxo do excesso de Na^+. Adaptado a partir de Gupta e Huang (2014).

3.2.3 Mecanismos de tolerância ao estresse osmótico

A salinidade no solo influencia marcadamente o potencial hídrico do tecido radicular da planta, podendo causar interrupção no influxo de água e redução da turgência celular na epiderme da raiz (CRAMER e BOWNMAN, 1991). O ajusto osmótico é um mecanismo chave para a manutenção da absorção hídrica e do turgor celular (TURNER et al., 2007). Neste sentido, como resposta às mudanças no potencial osmótico do meio, as plantas sintetizam e acumulam metabólitos de baixo peso molecular, que são conhecidos como osmossolutos, osmoprotetores ou solutos compatíveis (MUNNS, 2005). Esses compostos atuam auxiliando no ajustamento osmótico, quando em altas concentrações na célula, e na proteção celular, quando em baixas concentrações (ASHAF e FOOLAD, 2007).

À medida que as células são expostas à salinidade, o movimento da água ocorre do compartimento com maior potencial hídrico para o compartimento de menor potencial hídrico, e o fluxo de carbono pode ser alterado para atender à biossíntese de solutos compatíveis e geração de energia necessária para esta biossíntese, aumentando a concentração intracelular, inibindo assim a perda de água pela célula e propiciando a sua aquisição (MAHAJAN e TUTEJA, 2005). Esta habilidade de tolerância à salinidade é claramente diferenciada entre espécies e também entre genótipos (MUNNS et al., 2002).

O termo “soluto compatível” faz referência a solutos que não são tóxicos para a célula vegetal e que são compatíveis com a atividade metabólica mesmo em altas concentrações no citosol (WYN JONES et al., 1977). Outras funções conferidas aos solutos compatíveis são a contribuição na regulação do pH citosólico, desintoxicação do excesso de NH^{4+}, e proteção de plantas removendo ROS gerados por estresse oxidativo (ZHU, 2001). Em adição, os solutos compatíveis em concentrações elevadas no citoplasma podem auxiliar no balanço osmótico quando os eletrólitos no citoplasma estão em menor concentração que no vacúolo, permitindo à célula manter um potencial de turgor positivo, requerido para a expansão celular e abertura de estômatos, e um efeito protetor nas proteínas e estruturas celulares quando os eletrólitos estão em concentração elevada no citoplasma (NABIL e COUDRET, 1995).

Dentre os principais grupos de solutos compatíveis destacam-se: os compostos poliidroxílicos, tais como glicose, frutose, sacarose, trealose ou rafinose; ácidos

orgânicos, tais como oxalato e malato; poliálcoois de cadeia linear (glicerol, manitol ou sorbitol) e cíclicos (inositol, pinitol, etc.); compostos amino-quaternários (glicina betaína, alanina betaína, prolina betaína), aminoácidos protéicos (prolina, arginina, glicina, serina, etc) e não-protéicos (citrulina, ornitina, hidroxiprolina, etc.), poliaminas (putrescina, espermidina e espermina), dentre outros compostos (ASHRAF e HARRIS, 2004).

A prolina é um dos osmólitos mais pesquisados nos estudos acerca de estresses abióticos em plantas cultivadas, e seu acúmulo em plantas sob estresse salino é conhecido como um mecanismo adaptativo ao estresse (HASANUZZAMAN et al., 2014). Sob condições de estresse salino, a prolina é produzida a partir do ácido glutâmico e seu teor pode apresentar valores 100 vezes maiores nas plantas submetidas a estresse quando comparadas às plantas controle, dependendo da espécie (VERBRUGGEN e HERMANS, 2008). Além disso, a acumulação da prolina é induzida pelo aumento da concentração dos íons Na^+ e Cl^-, pela diminuição da atividade de prolina desidrogenase (uma enzima catabólica de prolina) (PARIDA e DAS, 2005). As reações de síntese da prolina são realizadas pela ação das enzimas pirrolina-5-carboxilato sintetase (P5CS) e pirrolina-5-carboxilato redutase (P5CR) (DELAUNEY e VERMA, 1993).

Durante os estresses osmóticos, além de atuar como um osmorregulador e ser uma fonte de carbono e nitrogênio, a prolina auxilia na estabilização das estruturas das proteínas, do pH citosólico e na proteção da membrana plasmática, no sequestro de espécies reativas de oxigênio durante estresse oxidativo (SHARMA e DUBEY, 2005).

Entre os compostos amino-quaternários, a glicina betaína (GB) é produzida diante o estresse salino na tentativa de controlar a capacidade das células em reter água sem causar prejuízos às plantas (TUTEJA, 2007). Uma vez que a GB pode promover maior produção de vacúolos nas raízes, algumas plantas podem minimizar o efeito da salinidade na parte aérea, devido à produção de GB nas raízes, fazendo com que o Na^+ se acumule e não seja transportado para a parte aérea (ASHRAF e FOOLAD, 2007). Foi demonstrado que a aplicação exógena em plantas que não produzem GB promoveu melhor desenvolvimento sob estresse salino, diminuindo a concentração de Na^+ e aumentando a absorção de K^+, além de agir beneficamente na transdução de sinais e equilíbrio iônico (TUTEJA, 2007). No entanto, foi observado em alguns casos que a alta concentração de GB pode inibir o acúmulo de prolina e causar danos nos tecidos de

folhas sob altas concentrações de GB exógena, e por estes motivos, os efeitos da glicina betaína vão depender da concentração aplicada e da resposta de diferentes tipos de plantas ao seu acúmulo (AHMAD et al., 2013).

O aumento do processo de síntese de proteínas de membrana, denominadas aquaporinas, envolvidas no movimento da água entre os meios intracelular e extracelular, é outro mecanismo de ajuste osmótico (HASEGAWA et al., 2000). A regulação da quantidade das aquaporinas é afetada por estresses provocados por salinidade e seca, bem como sua distribuição no tonoplasto (SMART et al., 2001).

3.2.4 Mecanismos de tolerância ao estresse oxidativo

Além dos efeitos iônico e osmótico ocasionados pela salinidade, as plantas podem sofrer por estresse oxidativo, com produção de espécies reativas de oxigênio (*Reactive Oxygen Species* - ROS), tais como o peróxido de hidrogênio (H_2O_2), o radical superóxido (O_2^-) e o radical hidroxil (HO^-), os quais são altamente reativos e prejudiciais às células vegetais em concentrações elevadas, mas que, em níveis relativamente baixos, podem atuar como moléculas sinalizadoras de mecanismos que minimizam o estresse abiótico (AHMAD et al., 2013). Os danos citotóxicos das ROS produzidas durante o estresse salino podem ser traduzidos em diversos processos degenerativos, alterarando o metabolismo celular, atuando, por exemplo, na peroxidação de lipídios da membrana celular e pigmentos fotossintéticos, desnaturação de proteínas e morte celular programada (ALSCHER et al., 199). A nível de material genético, o excesso de ROS podem causar hipo ou hipermetilação do ácido nucleico, deleção e substituição de bases do DNA, alterações cromossômicas (aneuploidia e poliploidia) e rearranjo cromossômico (CASSELLS e CURY, 2001).

O estresse oxidativo tem sido constatado em plantas expostas a salinidade e também a outros tipos de injúrias, como a temperaturas extremas, seca, luz UV, herbicidas e exposição a ozônio (INZÉ e VAN MONTAGU, 1995). As espécies reativas de oxigênio podem ser produzidas nos peroxissomos, cloroplastos e mitocôndrias. Nos cloroplastos, uma vez que é local de produção de O_2, molécula que pode receber elétrons e formar ROS passando pelos fotossistemas, cuja eliminação rápida de ROS neste local é um mecanismo de proteção dos cloroplastos antes de chegarem até o estroma e prejudicar o ciclo de Calvin (GILL e TUTEJA, 2010). A fotorrespiração nos

peroxissomos, durante estresses abióticos, também é uma fonte produtora de ROS (MITTLER, 2002).

A oxidação lipídica origina diversos produtos secundários que agravam o dano oxidativo celular, entre eles o ácido malondialdeído (MDA), que é o principal e mais estudado produto da peroxidação lipídica, e considerado um marcador molecular para designar a peroxidação lipídica nas células de plantas submetidas a diferentes estresses abióticos, entre eles, a salinidade (DAVEY et al., 2005)

Plantas tolerantes à salinidade podem possuir um eficiente sistema de defesa contra o estresse oxidativo, através de enzimas antioxidativas que as protegem pela remoção efetiva de ROS (HOEKSTRA et al., 2001). Dentre as enzimas que participam dos sistemas de defesa das plantas destacam-se a catalase (CAT), a superóxido dismutase (SOD), a peroxidase (POX), o ascorbato peroxidase (APX), a glutationa peroxidase (GPX) e a redutase da glutationa (GR) (Breusegem et al., 2001). Dentre os metabólitos não enzimáticos que podem ser importantes nos sistemas de defesa destacam-se o ascorbato (AsA), o monodesidroascorbato (MDHA), a glutationa reduzida (GSH), o α-tocoferol e os carotenoides (MITTLER, 2002).

Nas plantas sob estresse salino, a atividade das enzimas antioxidantes como a CAT, POX, glutationa redutase (GR) e superóxido dismutase (SOD) aumentam, existindo correlação entre os níveis destas enzimas e a tolerância ao sal (SUN et al., 2007). Assim, o balanço entre as atividades das enzimas antioxidantes no metabolismo celular é crucial para a manutenção do equilíbrio dinâmico das ROS (MITTLER, 2002).

As SOD são metaloenzimas responsáveis pela dismutação do radical superóxido em oxigênio molecular e peróxido de hidrogênio H_2O_2 (ASADA et al., 1999). As SOD apresentam-se na forma de três isoenzimas nas plantas superiores (Fe-SOD, Mn-SOD e Cu/Zn-SOD), classificadas de acordo com o íon metálico presente no grupo prostético e estão distribuídas em diferentes compartimentos celulares (ARORA et al., 2002). A Fe-SOD é encontrada principalmente nos cloroplastos, a Mn-SOD nas mitocôndrias e peroxissomos e a Cu/Zn-SOD nos cloroplastos, citosol e, possivelmente, no espaço extracelular (ALSCHER et al., 2002).

O H_2O_2 produzido pela reação catalisada pelas SODs também é formado espontaneamente pela dismutação do radical superóxido (O_2^-), onde pode ser eliminado do metabolismo celular por ação das enzimas CAT, APX e POX (ARORA et al., 2002). A concentração deste elemento pode variar de 0,03 – 1μM (sem estresse) a 0.1 - 10μM

(sob estresse), bem como a quantidade de dias que isso pode ocorrer, dependendo da espécie vegetal e de uma série de outras variáveis (DEMIDCHIK, 2015). Já a produção do radical hidroxila (HO^-) é realizada através das reações "Haber-Weiss", a partir dos radicais superóxido (O_2^-) e peróxido de hidrogênio (H_2O_2) (SCANDALIOS, 1993).

As POX são enzimas amplamente distribuídas nas células, encontradas associadas às paredes celulares, membranas celulares, vacúolos e citoplasma (ASADA, 1992). São enzimas capazes de oxidar vários substratos na presença de H_2O_2 ou de hidroperóxidos orgânicos, e por apresentarem menor massa molecular (35 kDa) que as CAT, as POX apresentam maior mobilidade dentro dos diversos compartimentos celulares onde a sua ação é requerida (SIEGEL, 1993).

O ascorbato é um importante composto no processo antioxidativo, classificado como componente chave do sistema antioxidativo das plantas, e está relacionado aos estresses abióticos e bióticos (DIPIERRO et al., 2005). Encontra-se em alta concentração no interior das células, estando distribuído no citosol, mitocôndria, peroxissomos, apoplasto e, principalmente, no interior dos cloroplastos (MITTLER, 2002). A síntese de ascorbato ocorre em mitocôndrias e a manutenção da sua concentração em plantas submetidas à determinada condição de estresse, envolve complexa interação entre síntese, degradação, transporte e armazenamento no interior das células (CREISSEN et al., 1999). O ascorbato é requerido como doador de elétrons para que alguns sistemas de defesa enzimáticos possam atuar eliminando as ROS do interior das células (FOYER e NOCTOR, 2000). As APX, por exemplo, utilizam o ascorbato como doador de elétrons na reação de eliminação de H_2O_2 (ASADA, 1992). As APX são encontradas, principalmente, nos cloroplastos e no citosol, podendo também estar associadas às mitocôndrias, peroxissomos e ao apoplasto (ARORA et al., 2002).

O ciclo do ascorbato-glutationa é uma eficiente via de células vegetais que dispõem de H_2O_2 em determinados compartimentos onde estes metabólitos é produzido e não existe catalase presente, fazendo uso de antioxidantes não enzimáticos, como o ascorbato e a glutationa, este ciclo está bem demonstrado nos cloroplastos, citosol e nas mitocôndrias de nódulos radiculares (POLLE, 2001).

3.2.5 Efeito da salinidade na germinação, estabelecimento e crescimento das plantas

As respostas das plantas à salinidade são muitas vezes acompanhadas por alterações morfológicas e anatômicas (Taiz e ZEIGER, 2013). Uma vez que é possível as plantas serem adaptadas e ajustadas osmoticamente ao estresse salino, o gradiente de potencial poderá ser mantido, no entanto, associado a reduções na condutividade hidráulica das raízes e maior fechamento estomático, o que resulta em menores taxas de absorção e assimilação de CO_2 por unidade de área foliar, induzindo à redução da taxa líquida de fotossíntese e, consequentemente, em menores taxas de crescimento (AZAIZEH et al., 1992).

O efeito negativo da salinidade é provocado pela dificuldade na absorção de água pela semente, reduzindo a velocidade dos processos metabólicos e bioquímicos, além de atrasar ou inibir a germinação e interferindo no processo de embebição e no alongamento celular do embrião (HARTER et al., 2014). A captação de água durante o estabelecimento das plântulas acarreta a acumulação de íons, principalmente no eixo embrionário (ASHRAF, 2004). Os íons acumulados ao atingir concentrações tóxicas podem afetar vários processos fisiológicos e metabólicos dos tecidos embrionários, incluindo a divisão e diferenciação celular, a atividade de enzimas e a captação e distribuição de nutrientes, podendo ocasionar atraso da emergência das plântulas, da mobilização das reservas e diminuir a viabilidade das sementes (MISRA e DWIVEDI, 2004). Assim, o fato da semente germinar, verificado pela emissão da radícula, sob substrato salino, não garante que esta torne-se uma plântula apta para se desenvolver (KRZYZANOWSKI e FRANÇA-NETO, 2001).

Modificações no status hídrico impostas pela salinidade influenciam marcantemente as respostas na taxa de crescimento foliar de plantas submetidas a estresse salino (CRAMER e BOWMAN, 1991). Entre as alterações anatômicas decorrentes dos efeitos da salinização, destacam-se: a lignificação das paredes celulares, que provocam alterações na espessura dos tecidos; presença de estrias de Caspary, que atuam reduzindo o transporte de íons Na^+ pela via apoplástica; armazenamento de cristais de oxalato de cálcio nas células; e redução no número de células dos feixes vasculares, com desorganização na acomodação dessas células (HUNSCHE et al., 2010). Nas folhas, a salinidade reduz o crescimento e acelera a abscisão foliar (ZEKRI, 1991). Também ocorre o aumento da espessura do mesófilo foliar, em função do aumento no comprimento e no número de camadas de células paliçádicas e esponjosas,

além de aumentar a produção de fibras e cristais de oxalato de cálcio (PARIDA et al., 2004). Efeitos transientes na alteração do turgor foliar em plantas afetadas pela salinidade também geralmente são produzidos, embora nem sempre o turgor é reduzido (MUNNS, 1993). O turgor das folhas de variedades sensíveis ao sal é usualmente maior do que variedades relativamente tolerantes ao sal, presumivelmente, em consequência da baixa exclusão de sal pelas raízes de variedades sensíveis, promovendo altas concentrações de sal nas folhas, enquanto que as variedades tolerantes, que excluem maiores quantidade de sais pelas células das raízes, mantém menores concentrações de sais nas folhas (YANG et al., 1990).

A redução no potencial hídrico da solução nutritiva pode causar interrupção do influxo de água ou uma redução da turgescência celular na epiderme da raiz (CRAMER e BOWMAN, 1991). Em plantas salinizadas frequentemente ocorre o aumento da suberização da hipoderme e endoderme nas raízes, com formação de estrias de caspary bem desenvolvidas, como forma de reduzir o transporte de Na^+ pela via apoplástica (SHANNON e GILL, 1994). No entanto, esse mecanismo pode limitar a absorção de água e nutrientes pelas raízes (OLIVEIRA et al., 2010).

A inibição do crescimento das raízes, decorrente da salinidade, está associada com a expansão da parede celular, papel atribuído primordialmente para as peroxidases (POD) ligadas ionicamente à parede celular (LIN e KAO, 2001). Além da ação antioxidativa aumentada em função da implicação de diferentes tipos de estresses nos vegetais, as POD exercem importantes funções no crescimento, diferenciação, desenvolvimento e lignificação da parede celular, podendo, em alguns casos, ter o seu efeito acentuado quando associado a fatores bióticos e abióticos (MENEZES et al., 2004). O mecanismo de espessamento da parede celular, ocasionado pela incorporação de lignina, promove rigidez estrutural e resistência aos tecidos das plantas, o qual diminui a extensibilidade da parede celular em decorrência da formação de pontes difenil entre polímeros da parede, por ação das POD (POLLE et al., 1994). Neste sentido, a redução do crescimento das raízes sob estresse salino está diretamente associada com o aumento dos níveis de peróxido de hidrogênio (H_2O_2), essencial para a atividade das POD (LIN e KAO, 2001). É importante salientar a complexidade do processo de biossíntese de lignina, que envolve inúmeras etapas enzimáticas que compõem o metabolismo secundário nas plantas, o qual confere várias funções

fisiológicas para sobrevivência e adaptação a perturbações ambientais (STRACK, 1997).

A salinidade também influencia negativamente diferentes aspectos da reprodução, incluindo a polinização, florescência, desenvolvimento de frutos e produção de sementes (SHANNON et al., 1994). A germinação é afetada pela salinidade não apenas por dificultar a absorção de água em função da redução do potencial osmótico do substrato até às células do embrião, mas também por facilitar a entrada de íons em quantidades tóxicas nas sementes durante a embebição (MACHADO-NETO et al., 2006). Além disso, a redução da capacidade de germinar também é atribuída à redução das atividades enzimáticas da semente, em função do processo de hidratação, pois a água constitui a matriz onde ocorre a maioria dos processos bioquímicos e fisiológicos que resultam na protrusão da raiz primária (YE et al., 2005).

3.3 BIOINFORMÁTICA

3.3.1 Histórico

A bioinformática é considerada um campo de estudo para análise de sistemas biológicos através da informática (HOGEWEG, 2011). O termo "Bioinformática" foi utilizado primeiramente por Pauline Hogeweg em 1979 para estudos de processos de informática em estudos de biologia sistemacional (BARNES e GRAY, 2003). A partir da década de 1980, com o aprimoramento das técnicas de sequenciamento e de novas tecnologias, a bioinformática passou a fazer parte de todos os projetos biológicos como forma de armazenar e analisar grandes quantidades de dados, e com a apresentação dos resultados em interfaces acessíveis via web, tornando a pesquisa mais interativa e dinâmica (BAYAT, 2002). Desde então, a tecnologia da computação, matemática e biologia molecular se tornaram proveitosamente combinadas para responder a perguntas fundamentais nas ciências naturais, e em especial, para auxiliar no manejo da grande quantidade de dados gerado no sequenciamento de DNA, RNA e aminoácidos (HAGEN, 2000). Dada a riqueza das informações disponíveis, a bioinformática possui papel central, e integrador, permitindo que experimentos sejam feitos *in silico* e que o uso da bancada, para experimentos *in vitro* ou *in vivo*, sejam adiados, até que um foco maior seja dado à pesquisa seja considerada (MARTINS et al., 2009). No Brasil, a bioinformática foi utilizada pela primeira vez de maneira sistemática e em larga escala em 1997, no projeto de sequenciamento do genoma da bactéria *Xylella fastidiosa*,

causadora da doença do amarelinho nas laranjeiras (CIB, 2004). Atualmente, a bioinformática tem sido utilizada para a realização de diversos estudos, entre eles para a construção de banco de dados, mineração de dados, análises de sequências, identificação de genes e predição de suas funções, previsão da conformação tridimensional das proteínas, construção de árvores filogenéticas e modelos evolutivos, construção de bibliotecas genômicas, transcriptômicas e proteômicas, estudo de funções biológicas, entre outras (MATIOLI e FERNANDES, 2012).

3.3.2 Bancos de Dados de Bioinformática

Devido ao avanço dos recursos computacionais e da disponibilidade de genomas completos sequenciados de organismos, inclusive de plantas, houve uma crescente necessidade da disponibilidade de informações em bancos de dados biológicos (PEVZNER, 2004). Atualmente, existem uma série de bancos disponibilizados *on-line* em bases de dados biológicos. Os portais de bioinformática organizam os dados de diferentes organismos na forma de bancos de dados de genomas, transcriptomas, proteomas, estruturas tridimensionais, sequências intergênicas e peptídicas (JONES e PEVZNER, 2004).

Uma das principais bases de dados sobre informações biológicas é o GenBank do Centro Nacional para Informação Biotecnológica dos EUA (NCBI - *National Center for Biotechnology Information* - *www.ncbi.nlm.nih.gov*). O GenBank tem sido expandido para incluir dados de expressão de sequências (ESTs), dados de sequências e estrutura tridimensional de proteínas, coleção de anotações de sequências de DNA depositadas de diferentes organismos, informações sobre diferentes tipos de interações biológicas, taxonomia e literatura biomédica (BENSON et al., 2011).

O INSDC (*International Nucleotide Sequence Database Collaboration*) compõe uma colaboração internacional formada pelo o NCBI, o ENA (*European Nucleotide Archive*) e o DDBJ (*DNA Data Bank of Japan*), que mantém um conjunto de dados biológicos em constante sincronia (Barnes e Gray, 2003). O ENA é desenvolvido e mantido pelo EMBL-EBI (*European Bioinformatics Institute*) e contém informações sobre sequências nucleotídicas contabilizadas em cerca de 255,1 milhões (LESK, 2008). O DDBJ além de oferecer ferramentas de pesquisa de sequências de DNA e proteínas, fornece acesso a diversas ferramentas de análises de bioinformática (WHEELER et al., 2006).

Desenvolvida pelo NCBI na *Library of Medicine* (NLM), o PubMed é uma base de dados disponível na web do sistema de busca e recuperação Entrez do NCBI, que inclui o MEDLINE, Nucleotide Sequences, Protein Sequences, Macromolecular Structures e Whole Genomes (BARNES e GRAY, 2003). O serviço do Entrez fornece mecanismos de busca em diversos bancos de dados disponibilizados pelo NCBI, que integra literatura científica, bancos de dados de DNA e proteínas, estruturas tridimensionais de proteínas, genomas completos e informações taxonômicas de organismos conhecidos em seu sistema de busca (JONES e PEVZNER, 2004).

O Uniprot (*http://www.uniprot.org/*) é um repositório de sequências de proteínas com cobertura abrangente e abordagem sistemática para a anotação de proteínas. Integra, interpreta e padroniza os dados de literatura e de outros numerosos recursos para alcançar a mais abrangente base de dados sobre proteínas (SUZEK et al., 2007). No UniProt há dois módulos principais, o UniProtKB (*Protein Knowledgebase*), onde toda a anotação das sequências proteicas é encontrada, e o UniRef (*UniProt Reference Clusters*), que agrupa as sequências do UniProt nos níveis 50%, 90% e 100% de identidade (*The UniProt Consortium*, 2014; http://www.uniprot.org/). O UniProtKB é divido em dois módulos, o de sequências anotadas por métodos automáticos (TrEMBL) e o Swiss-Prot, onde as sequências passam por curadoria manual (APWEILER et al., 2010). As anotações do Swiss-Prot descrevem funções das proteínas, modificações pós-traducionais, domínios e sítios, estruturas secundárias, entre outras informações (SUZEK et al., 2007).

3.3.3 Bases de Dados de Fatores de Transcrição

PlnTFDB - *Plant Transcription Factor Database*

A base de dados de fatores de transcrição de plantas (PlnTFDB; http://plntfdb.bio.uni-potsdam.de/v3.0/) é um interativo portal que fornece conjuntos de putativos fatores de transcrição e outros elementos reguladores transcricionais em 19 espécies de plantas cujos genomas foram completamente sequenciados e anotados, onde para cada família de genes, é fornecida uma descrição básica que é complementada por referências bibliográficas e alinhamentos de sequências múltiplas de domínios protéicos (PÉREZ-RODRÍGUEZ et al., 2010).

DBD - *Transcription factor prediction database*

O DBD (*DNA-binding domain* - http://transcriptionfactor.org) é uma base de dados de sequências preditas de domínios de ligação ao DNA de fatores de transcrição para todos os proteomas disponíveis publicamente, em que são oferecidas novas opções de pesquisa, como busca por nomes de genes em organismos modelo (WILSON et al., 2008).

JASPAR

Disponível em http://jaspar.cgb.ki.se. o JASPAR é uma base de dados aberta para detecção, anotação e análise de potenciais perfis de ligação de fatores de transcrição de eucariotos, com complementação de uma interface para navegação, seleção de subconjuntos e um conjunto de ferramentas de programação para análise de genoma ampla e comparativa de regiões regulatórias (SANDELIN et al., 2004).

TRANSFAC

Base de dados de fatores de transcrição de organismos eucariotos e sítios de ligação de fatores de transcrição (TFBSs - *Transcription Factor Binding Sites*), o TRANSFAC (http://genexplain.com/transfac/) tem sido cuidadosamente mantido e curado para disponibilização de análises de sequências genômicas para potenciais sítios de ligação de fatores de transcrição, além de uma complexa plataforma *on line* (geneXplain) que possui uma interface padronizada com sistema de gerenciamento de fluxogramas para uma ampla gama de aplicações de bioinformática e biologia de sistemas (WINGENDER, 2008).

PlantTFDB - *Plant Transcription Factor Database*

Com um recurso de fornecimento bastante abrangente de fatores de transcrição e suas interações com genes alvo, a base de dados PlantTFDB (http://planttfdb.cbi.pku.edu.cn/) possui 320370 fatores de transcrição identificados em 165 espécies, além de promover uma abundante atualização de anotação funcional e evolutiva para fatores de transcrição identificados, a partir da geração de três novos tipos de anotação para a investigação de mecanismos funcionais: i) um conjunto de motivos de ligação não redundantes de fatores de transcrição derivados de experimentos; ii) múltiplos tipos de elementos *cis*-regulatórios identificados a partir de dados de sequenciamento de alto desempenho; iii) interações regulatórias curadas da literatura e inferidas pela combinação de motivos de ligação de fatores de transcrição e elementos *cis*-regulatórios (JIN et al., 2016).

3.3.4 Bases de Dados de Elementos *Cis*-Regulatórios

PlantCARE

A base de dados de elementos *cis*-regulatórios PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) utiliza motivos obtidos principalmente da literatura, sendo complementado com um número crescente de dados previsto pela própria base, fornecendo uma descrição geral de sítios específicos de ligação fatores de transcrição, além dos níveis de confiança pela evidência experimental, informação funcional e posição no promotor (LESCOT et al., 2002). Novos elementos *cis*-regulatórios identificados também podem ser enviados automaticamente à base, e em sequida adicionados no portal após curagem.

PLACE - *Plant cis-acting regulatory DNA elements*

Assim como na base PlantCARE, os motivos dos elemententos *cis*-regulatórios da base de dados PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/) são extraídos de artigos publicados proveniente de promotores de genes de plantas vasculares, além de suas variações nas diferentes espécies que são compilados de artigos publicados posteriormente (HIGO et al., 1999). As informações de cada motivo na base de dados PLACE contém, além da sequência do motivo, uma breve descrição e definição de cada elemento e o número de identificação no PubMed e de acesso no GenBank, para que o usuário possa acessar o resumo da literatura e a anotação na base de dados.

PlantPAN - *The Plant Promoter Analysis Navigator*

Além de uma base de dados, o PlantPAN (http://plantpan2.itps.ncku.edu.tw) é uma plataforma sistemática que fornece um recurso informativo para detecção, análise e reconstrução de vias de regulação transcricional em plantas, que incluem os próprios sítios dos elementos cis-regulatórios e outros importantes elementos reguladores, como ilhas CpG (*CpG islands*) e repetições em tandem localizadas nas regiões conservadas dos promotores dos genes (CHOW et al., 2016). Atualmente, a base de dados PlantPAN contém depositados 16960 fatores de transcrição e 1143 motivos de elementos *cis*-regulatórios no portal a partir de 76 espécies de plantas, além de atualização das informações das anotações depositadas, adicionamento de matrizes de fatores de transcrição verificadas experimentalmente e melhoramento das visualizações das vias de regulações transcricionais, através da incorporação de vários recursos e funções, tais como: i) curagem abrangente de informações de fatores de transcrição (condições de

resposta, genes alvo e seqüência dos motivos de ligação); ii) perfis de co-expressão de fatores de transcrição e seus genes alvos sob várias condições; iii) interações proteína-proteína entre fatores de transcrição e seus co-fatores; e iiii) vias de elementos *cis*-regulatórios alvos de fatores de transcrição.

AGRIS - *Arabidopsis Gene Regulatory Information Server*

A base de dados AGRIS (http://agris-knowledgebase.org/) é um recurso de informação sobre sequências promotoras de genes de Arabidopsis, juntamente com fatores de transcrição e seus respectivos genes alvos (YILMAZ et al., 2011). Atualmente, a base AGRIS contém três bancos de dados: ArcisDB que consiste em um banco de aproximadamente 33000 regiões promotoras anotadas de genes de Arabdopsis, com uma descrição de seus elementos *cis*-regulatórios preditos e validados experimentalmente; o banco AtTFNet contém informação de aproximadamente 1770 fatores de transcrição agrupados em 50 famílias, baseadas na presença de domínios conservados específicos; e AtRegNet que é um banco que possui 19.013 interações diretas entre fatores de transcrição e genes alvos. Além das significativas contribuições na identificação de todo o conjunto de interações de fatores de transcrição e DNA, que são a chave para compreender as vias regulatórias que governam a expressão gênica de Arabidopsis, algumas ferramentas foram incorporadas na base de dados que incluem o conjunto completo de palavras de tamanho variando entre 5 e 15 letras presentes no genoma de Arabdopsis e a integração do banco AtRegNet com ferramentas de visualização.

3.3.5 Ferramentas Computacionais

Alinhamento de Sequências - BLAST

Para a busca de sequências em bancos de dados, é necessário utilizar uma ferramenta que reconheça as sequências mais similares, e assim facilitar a comparação e o alinhamento entre elas (SCHNEIDER, 2003). A homologia entre sequências pode ser analisada através do alinhamento obtido a partir de algorítmos e métodos estatísticos para availiar a significatividade (DURBIN et al., 1998). O programa BLAST (*Basic Local Alignemnt Search Tool*) usa métodos de similaridade local entre as sequências, através da localização, alinhamento e comparação de segmentos nucleotídicos ou peptídicos no banco de dados, e calcula a significância estatística da diferença entre as sequências (ALTSCHUL, 1997). O BLAST pode ser utilizado para a identificação de

um gene em um genoma ou de possíveis homólogos em um banco de dados, detecção de éxons, localização de domínios proteicos, predição de genes e proteínas, inferência de relações evolutivas e funcionais, bem como identificar membros de famílias gênicas (Barnes e GRAY, 2003; LESK, 2008). O resultado do BLAST é baseado em cálculos estatísticos após a obtenção de casamentos (*matches*) significativos entre as sequências, com a distribuição dos alinhamentos locais existentes e as pontuações (*scores*) atribuídas aos alinhamentos obtidos entre as sequências comparadas (ALTSCHUL, 1997; SCHNEIDER, 2003). O algoritmo processa os dados utilizando a heurística comparativa por pares de segmentos, com a finalidade de encontrar o menor número de falhas (*gaps*) e diferenças (*mismatches*) nos fragmentos resultantes. O *e-value* representa o número de alinhamentos com *scores* iguais ou melhores que seria de se esperar que ocorressem ao acaso, dado o tamanho da base de dados (LESK, 2008). Assim, quanto menor o *e-value*, melhor o alinhamento, de forma que, num banco de dados de grandes proporções, um *e-value* igual a zero significa que não há probabilidade de que um alinhamento entre duas sequências tenha ocorrido ao acaso (ALTSCHUL, 1997).

CLUSTAL X

Programa de execução de alinhamentos múltiplos biologicamente informativos a partir de sequências de bases nucléicas ou peptídicas alinhadas que dão origem a um determinado resultado (PROSDOCIMI et al., 2002), onde também é possível a visualização através de uma interface gráfica da relação filogenética das sequências em questão através de cladogramas (THOMPSON et al., 1997).

MEGA

O programa de análises genéticas de evolução molecular (*Molecular Evoluntionary Genetics Analysis*) foi desenvolvido para analisar caracteres evolutivamente informativos (SUDHIR et al., 1993) que permitem a análise de uma matriz de dados através de métodos utilizados para cálculo de distâncias genéticas, como distância de Kimura-2-parâmetros (KIMURA, 1980), distância de Tamura (TAMURA et al., 2011) e distância de Tajima-Nei (TAJIMA e NEI, 1984), além de disponibilizar os algoritmos para reconstruções fenéticas e filogenéticas através de geração de dendrogramas como o UPGMA (*Unweighted Pair Group Method with Arithmetic Means*) (SNEATH e SOKAL, 1973), NJ (*Neighbor-Joinning*) (SAITOU e NEI, 1987) e Máxima Parcimônia (FITCH, 1971).

TREE VIEW

É um programa para visualização de árvores filogenéticas capaz de ler diferentes formatos de arquivos, como Nexus, Phylip, Nona, Mega e ClustalW/X, além de gerar opções de visualização das árvores em diferentes formas, como em forquilha, radicalmente e em paralelo (PAGE, 1996).

CLUSTER

Um ambiente computacional e gráfico é fornecido pelo programa para a análise de sequências genômicas que incluem ferramentas de clusterização hierárquica (*Hierarchical Clustering*), análises de componentes principais (*Principal Component Analysis*), agrupamento de médias K (*K-Means Clustering*) e auto-organização de mapas (*Self Organizing Maps*) (EISEN et al., 1998).

3.4 TRANSCRIPTÔMICA

O transcriptoma de uma espécie pode ser entendido como produto da parte expressa do genoma, ou seja, o genoma funcional, sendo representado pelo conjunto de todos os transcritos derivados de genes produzidos em uma célula em determinado momento e condição fisiológica, os quais são compostos por uma coleção de moléculas de RNA codificadoras de proteínas e não codificadoras (ncRNA) (WANG et al., 2009). O transcriptoma fornece características sobre o padrão de expressão de um determinado organismo, tecido ou célula em questão, e sua análise é fundamental para compreender a função, estrutura e as interações dos genes envolvidos num determinado processo (MIR et al., 2004), além de constituir uma abordagem de alta performance para a investigação do comportamento de milhares de genes, permitindo também a inferência sobre a função de grupos de genes co-regulados (AMARAL et al., 2006). Entre os principais objetivos num estudo de transcriptômica destacam-se: a quantificação dos diferentes níveis de expressão de cada transcrito durante o desenvolvimento do organismo e sob diferentes condições; o catálogo de todos os elementos de transcrição, incluindo mRNAs, ncRNA e miRNA; a determinação da estrutura transcricional dos genes e do local de início das extremidades 5' e 3'; assim como os padrões de *splicing* e de outras modificações pós-transcricionais (WANG et al., 2009).

O estudo da expressão gênica em larga escala possibilitou gerar um grande volume de resultados, favorecendo o estudo da regulação gênica em função do grande

número de novos genes descobertos nos projetos genomas, desenvolvimento dos arranjos de DNA e pelas técnicas de análise de transcriptômica (NAGALAKSHMI et al., 2008). Consequentemente, a otimização de processos biológicos envolvidos pode ser encaminhada através do conhecimento global destes mecanismos e na modulação e/ou alteração dos padrões de expressão em uma determinada condição (HRDLICKOVA et al., 2016).

A quantificação da expressão gênica tem sido o foco do perfil transcriptômico tradicionalmente, contudo, é possível obter informações estruturais altamente resolvidas de populações de RNAs com o advento das tecnologias de sequenciamento em plataformas de alto rendimento (SHIRAKI et al., 2003). O fato da concentração relativa de transcritos de determinados genes ser diretamente proporcional ao seu nível de expressão permite, através da quantificação dos mesmos, inferir sobre os níveis de expressão destes genes em condições específicas (NAGALAKSHMI et al., 2008).

As primeiras tentativas de compreender os transcriptomas incluíram análises do RNA total em diferentes organismos, bem como a presença e a quantidade de transcritos de interesse (MOROZOVA et al., 2009). Os primeiros estudos baseados em genes utilizavam uma técnica de baixo rendimento denominada *Northern blot*, que usa radioatividade e grandes quantidades de RNA (ALWINE et al., 1977). Devido a elevada quantidade de RNA exigida pela técnica e à sua complexidade, o uso do método em estudos de expressão gênica foi reduzido e possibilitou o desenvolvimento de métodos baseados em PCR quantitativa de transcrição reversa (RT-qPCR), facilitando a detecção de transcritos de uma só vez usando uma baixa quantidade de RNA (SCHMID et al., 2005). Com o advento da técnica RT-qPCR, que permite quantificar o número de cópias de cDNA da sequência alvo, ficou possível analisar comparativamente o número de cópias do molde de DNA e a quantidade de RNAm que o gerou (AMARAL et al. 2006). Logo, tal estratégia tem permitido análises globais da expressão de genes de interesse, pois há uma relação direta entre o número de cópias e os níveis de expressão de determinados genes (BENDERS et al., 2005).

Entre as principais formas de estudo da expressão gênica para o estabelecimento de perfis transcricionais estão os arranjos de cDNA, cujas sequências são provenientes do método *Tag* de sequência expressa (*Expressed Sequence Tag* - EST) (ADAMS et al., 1993), que estuda a expressão de genes por sequenciamento parcial de clones complementares de DNA (cDNA), revelando assim a sequência e abundância de uma

população de RNAm expresso, para a identificação de regiões codificantes no DNA genômico (KING et al., 2011). A abundância dos diferentes transcritos está relacionada ao quantitativo de clones de cDNA correspondentes a cada um dos transcritos presentes em uma biblioteca, sendo, por isso, uma biblioteca genômica funcional (BOUCK e VISION, 2007). Esse método teve papel fundamental na identificação de novos genes em genomas na década de 1990 (HRDLICKOVA et al., 2016).

Entre 1990 e 2000 foram desenvolvidas tecnologias de sequenciamento de alto rendimento que tornaram as análises da expressão gênica mais apuradas para o estudo de transcriptomas (KASUGA et al., 2005), como por exemplo novas técnicas de hibridização subtrativa, Microarray, SAGE (*Serial Analysis of Gene Expression*) e cDNA-AFLP (*Amplified Fragment Length Polymorfism*, AFLP) (POGGELER et al., 2006). No entanto, estas tecnologias possuem algumas limitações, como a necessidade de conhecer previamente o genoma de estudo, intervalo dinâmico limitado (níveis de expressão) ou o risco de hibridação cruzada (WARD et al., 2012). A partir de 2005 tecnologias de sequenciamento de nova geração (*next generation sequencing*, NGS) começaram a ser utilizadas, baseadas em plataformas de sequenciamento massivo de cDNA, capazes de gerar informação sobre milhões de pares de bases em uma única corrida (ZHANG et al., 2012). Entre estas tecnologias, a tecnologia de RNA-seq permite organizar um quadro quase completo dos eventos transcriptômicos de uma amostra biológica (SHENDURE, 2008), pois pode evidenciar, mesmo se tratando de transcritos raros, o conjunto de sequências em um ponto específico (ZHONG et al., 2012). A tecnologia de RNA-seq tem superado limitações de outras tecnologias de ampla utilização, como os Microarrays, devido principalmente a necessidade de menores quantidades de RNA (WANG et al., 2009), além da possibilidade de identificar a estrutura e a quantificação exata dos níveis de expressão de éxons, assim como a caracterização de genes, localização precisa dos limites de transcrição, identificação das extremidades 5’ e 3’ dos genes e locais de *splicing* alternativo e variantes de *splicing* (ZHONG et al., 2012).

A abordagem aplicada na tecnologia RNA-seq consiste na purificação do RNAm, preparação e fragmentação de uma biblioteca de cDNA, obtida a partir do RNAm e no sequenciamento através de uma plataforma (Figura 3). Inicialmente, uma população de RNAm total ou fracionado é convertido, através de transcrição reversa, em bibliotecas de cDNA com tamanhos homogêneos. Adaptadores são adicionados em

uma ou ambas as extremidades dos fragmentos de cDNA e uma sequência curta é obtida a partir de cada molécula de cDNA, através de tecnologia de sequenciamento de alto desempenho. Em seguida, os *reads* de sequências resultantes podem ser alinhados com o genoma de referência, com o transcriptoma, por montagem *de novo* ou pelas estratégias combinadas. Os reads são utilizados para gerar um perfil de expressão para cada gene e são classificados em três tipos: *reads* exônicos, *reads* de junção e *reads* de final de cauda poly (A) (WANG et al., 2009). Quando cada transcrito for inteiramente coberto por um único contig, que representa uma região consenso de segmentos de DNA, a montagem de um transcriptoma pode ser considerada finalizada (MEYER et al., 2012).

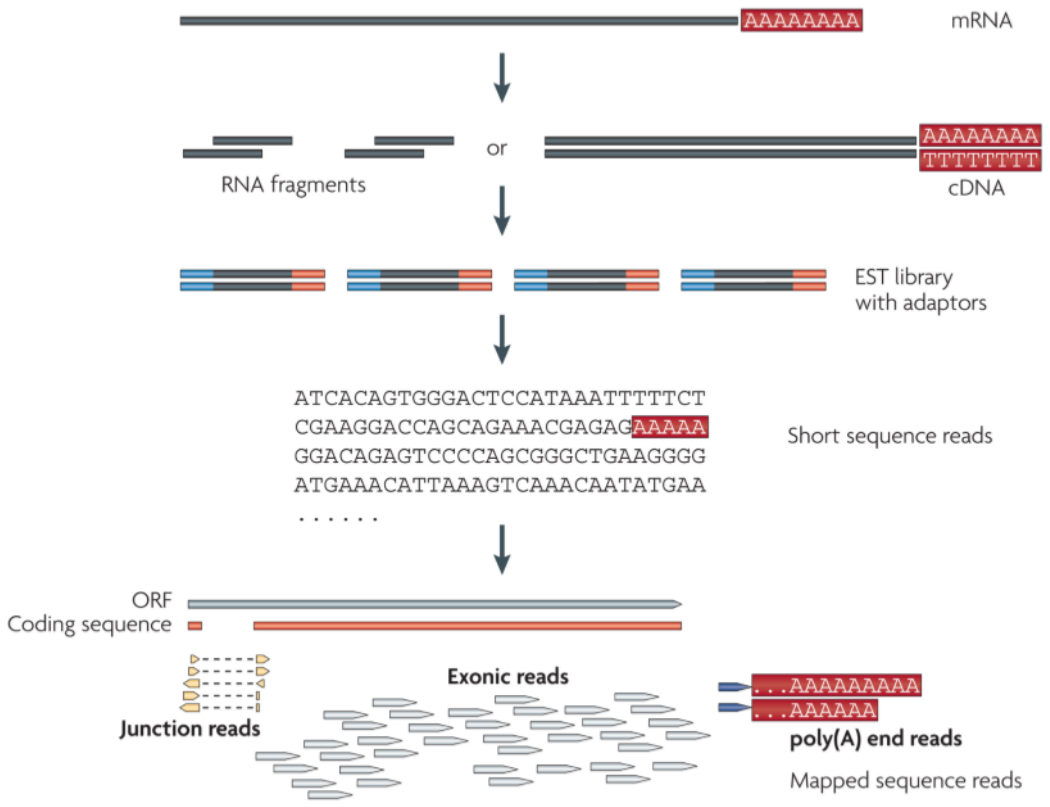


Figura 3. Esquema ilustrando a metodologia geral utilizada na técnica de sequenciamento RNA-seq. Imagem adaptada de Wang et al., 2009.

A abordagem RNA-seq encontra-se comercialmente disponível em algumas plataformas de NGS, como Illumina e Ion Proton System (WANG et al., 2009; BROWN et al., 2017). A escolha da plataforma mais apropriada depende das

particularidades de cada projeto, como a disponibilidade ou não de um genoma de referência, a longitude das leituras e o recurso financeiro (BARBA et al., 2014).

A plataforma Illumina, durante os últimos cinco anos, tem sido usada com maior frequência em diferentes projetos que envolvem sequenciamento, se tornando uma das opções mais recomendáveis em função de oferecer uma série de vantagens, como: custo, quantidade de nucleotídeos capazes de serem sequenciados numa mesma corrida, longitude das leituras geradas, precisão, tempo e rendimento de sequenciamento (LIU et al., 2012; BARBA et al., 2014). O sequenciamento na plataforma Illumina é realizado através da síntese de uma nova molécula utilizando DNA polimerase e nucleotídeos terminadores marcados com diferentes fluoróforos, muito semelhante ao que ocorre na técnica Sanger (HENSON et al., 2012). A inovação desta metodologia consiste na clonagem *in vitro* dos fragmentos em uma plataforma sólida de vidro, onde ocorre a amplificação das amostras por PCR, processo também conhecido como PCR de fase sólida (TURCATTI et al., 2008).

3.5 FATORES DE TRANSCRIÇÃO

Os fatores de transcrição (FTs) são importantes proteínas que regulam a expressão de genes alvos, e são compostos por pelo menos dois domínios distintos, sendo um domínio de ativação/repressão e um domínio ligação ao DNA, que se conecta a sequências específicas da região regulatória dos genes para controlar a atividade da RNA polimerase II (ZHENG et al., 2016).

Os genes que codificam FTs compreendem uma fração substancial dos genomas de eucariotos (TONOIKE et al., 1994). Em *Arabidopsis thaliana*, cerca de 1.700 genes para FT são sugeridos, o que representa mais de 5% dos 30.000 genes estimados para a espécie (RIECHMANN e RATCLIFFE, 2000). Nas células vegetais, alguns genes de FTs são expressos constitutivamente, enquanto outros genes respondem à estímulos específicos (LIU et al., 1999).

A afinidade e seletividade dos FTs é influenciada pelas estruturas secundárias dos domínios de ligação ao DNA, que contatam bases dos elementos *cis*-regulatórios presentes nos promotores dos genes (HUANG et al., 1996). Por exemplo, o domínio de ligação C-terminal do fator de transcrição Trilex GT2 (*GT2*-box-binding factor) em arroz perde a atividade quando resíduos de prolina são substituídos por outros aminoácidos, quebrando a hélice do domínio (NI et al., 1996).

Os FTs são regulados em seus domínios por, no mínimo, cinco diferentes mecanismos: controle da atividade por intermédio de quinases e fosfatases que fosforilam e desfosforilam domínios dos FTs; degradação proteolítica dos FTs ou de seus correguladores; modulação de interações proteína-proteína entre os FTs, correguladores e o complexo de transcrição basal; regulação da ligação do FT no sítio do DNA; e pela modificação da estrutura da cromatina (WHITMARSH e DAVIS, 2000). Estudos recentes indicam que complexos de fatores de transcrição se ligam a combinações específicas de histonas modificadas na cromatina, afetando o acesso de outros FTs e a iniciação da transcrição (MARTINO et al, 2009; LI et al., 2010).

Cada FT possui geralmente apenas um tipo de domínio de ligação ao DNA, ocorrendo em unidade ou em múltiplas cópias (LIU et al., 1999). Por exemplo, a maioria das proteínas Myb-relacionadas possuem dois domínios Myb de ligação ao DNA (BARANOWSKIJ et al., 1994). Adicionalmente, muitos FTs também podem se associarem entre si através de domínios idênticos e não idênticos, formando homoligômeros e heteroligômeros, respectivamente, o que pode afetar tanto a especificidade da ligação quanto a afinidade dos FTs nos sítios de ligação (VIGNALI et al., 2000). Uma vez que normalmente os heteroligômeros se formam a partir de proteínas com especificidades distintas de ancoramento ao DNA, a mistura e as combinações de diferentes proteínas, em vez das proteínas individuais, aumentam muito o repertório de resultados possíveis que essas proteínas podem modular, diversificando assim, a especificidade e a intensidade da ligação dos FTs ao DNA por este mecanismo combinatório (SHOGREN-KNAAK et al., 2006).

Outra consequência comum da oligomerização de FTs é o aumento da afinidade na ligação a um determinado sítio do DNA (OAKLEY et al., 2003). Por exemplo, a proteína homeodomínio de leveduras MATα2 pode se ligar no DNA como um monômero, assim como se heterodimerizar com outra proteína homeodomínio, a MATa1, ou ainda formar heterotetrâmeros com a proteína MADS box MCM1 (JOHNSON et al., 1998). Como resultado, o heterodímero MATα2/MATa1 e o heterotetrâmero MATα2/MCM1 ligam-se ao DNA com mais força, afinidade e especificidade de sequência do que o monômero MATα2 solitário (TAN e RICHMOND, 1998).

Tipicamente, os FTs reconhecem motivos curtos, de comprimento variando de 6 a 12 pb nas sequências de DNA, e esta baixa especificidade sugere que mecanismos

mais complexos, que não seja a simples afinidade dos FTs nas bases do DNA, estejam envolvidos no controle tanto da ocupação do sítio como no resultado funcional (SPITZ e FURLONG, 2012). De fato, foi constatado que 99,8% dos motivos de ligação encontrados no genoma humano não são ligados por seus respectivos TFs cognatos e, portanto, fica claro que a simples presença de um motivo não significa que este será reconhecido por seu TF correspondente (WANG, et al., 2012). Neste sentido, muitos estudos focalizam a compreensão dos princípios que governam o reconhecimento específico de motivos de DNA pelos FTs (BULYK et al., 2002).

A ocupação combinatória de múltiplos FTs em sítios adjacentes é uma ligação cooperativa frequentemente associada com interações proteína-proteína, e a resposta transcricional resultante pode produzir diversos efeitos de comutação, dependendo de como os FTs interagem uns com os outros (GIORGETTI et al., 2010). A importância da ligação cooperativa é refletida por observações em que muitos FTs não podem ocupar individualmente seus sítios no DNA nucleossomal (ADAMS e WORKMAN, 1995), mas interações cooperativas entre um grupo de FTs podem permitir a ligação no nucleossomo (ZARET e CARROLL, 2011).

Alguns mecanismos nos quais os FTs podem atuar cooperativamente existem, inclusive de forma indireta, reforçando assim a ocupação de cada um no sítio de ligação (SPITZ e FURLONG, 2012), por exemplo: dois ou mais FTs podem recrutar um cofator comum, ou diferentes componentes de um complexo multiproteico, como mediadores de transcrição da RNA Polimerase II (Fig. 4a); alguns FTs podem atuar cooperativamente na ativação do complexo de remodelagem da cromatina (Fig. 4b); um FT pode impedir o reposicionamento do nucleossomo ao permanecer ligado a um determinado sítio, regulando assim a ligação de outros FTs (Fig. 4c); alguns FTs podem induzir a flexão local no DNA, o que pode aumentar a afinidade de outros FTs nos sítios (Fig. 4d).

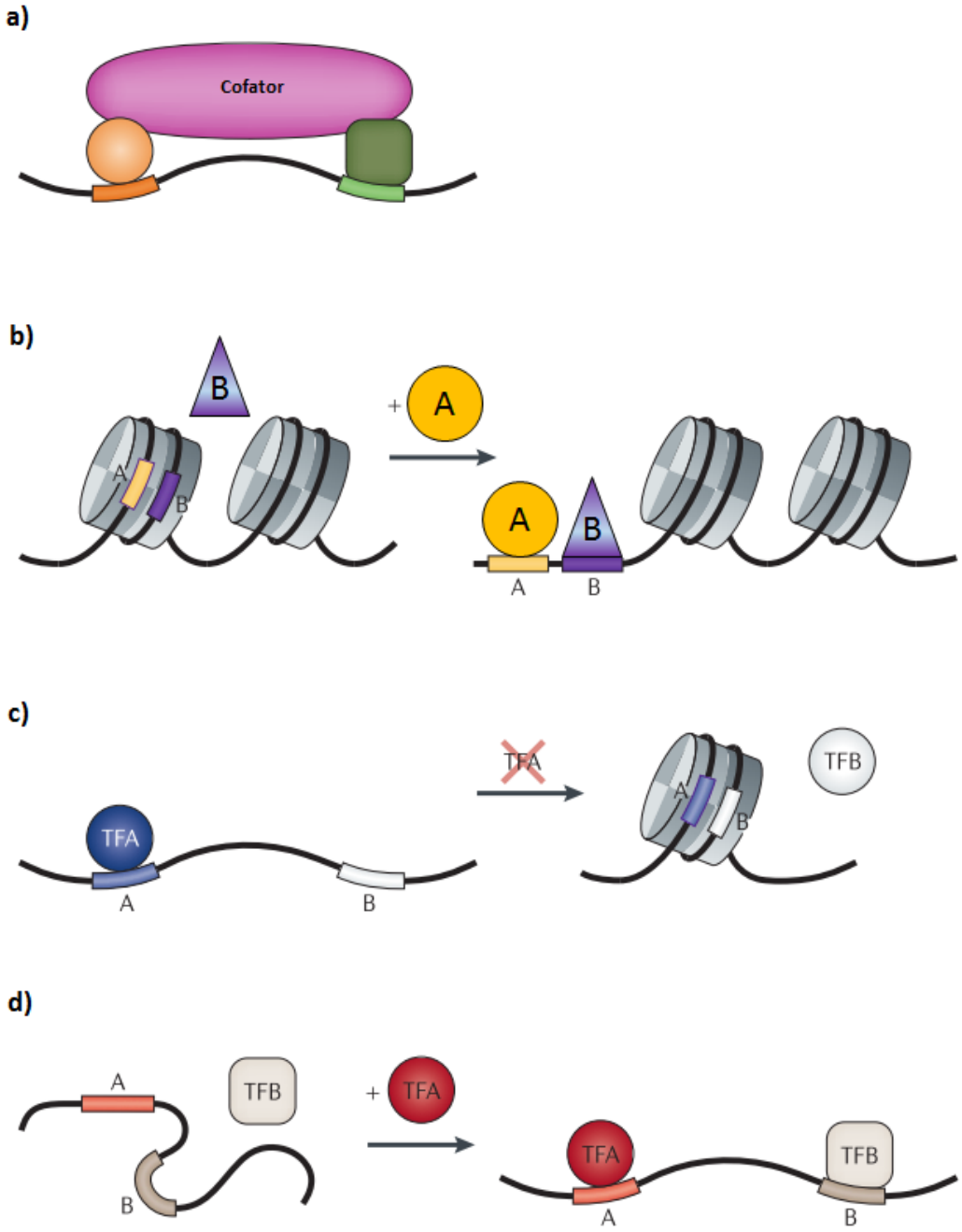


Figura 4. Mecanismos indiretos de cooperatividade entre fatores de transcrição. Adaptado a partir de Spitz e Furlong (2012).

O genoma de Arabidopsis contém grandes famílias de fatores de transcrição, como AP2/ERF, bHLH, MYB e MADS-box, que contém mais de 100 membros cada (RIECHMANN e RATCLIFFE, 2000; TOLEDO-Ortiz et al., 2003).

A classificação dos fatores de transcrição está estritamente relacionada com o mecanismo de ação, que é a ligação ao DNA e sua influência na transcrição, na qual depende de suas características estruturais (HEIM et al., 2003). As estruturas modulares das proteínas FTs, compreendidas pelos domínios, com regiões específicas responsáveis

pela ligação ao DNA e por produzirem o efeito inibitório ou estimulatório da transcrição, são frequentemente utilizadas como a fonte para esta classificação (BAILEY et al., 2003). Muitos domínios de ligação ao DNA dos fatores de transcrição de plantas têm caráter básico e contém resíduos de reconhecimento dos elementos *cis*-regulatórios que, na maioria dos casos, são altamente conservados (HUANG et al., 1996). Na base de dados de domínios conservados do NCBI (CDD - *Conserved Domain Database*, https://www.ncbi.nlm.nih.gov/cdd) encontram-se depositados para fatores de transcrição um total de 689 domínios, divididos em 262 superfamílias e 556 famílias. Além dos domínios curados pela própria base do NCBI, outros domínios também encontra-se anotados por bases externas, como SMART (*Simple Modular Architecture Research Tool*), Pfam (*Protein families*), COGs (C*lusters of Orthologous Groups of proteins*), TIGRFAM (*The Institute for Genomic Research's database of protein families*) e PRK [*PRotein K(c)lusters*].

Com base nas similaridades entre as estruturas dos domínios de ligação ao DNA, de multimerização para ativação/repressão, e de acordo com a caracterização do número e o espaçamento dos resíduos dos motivos conservados (RICHMANN e MEYEROWITZ, 1998), os fatores de transcrição podem ser classificados em diversas famílias, entre as principais, incluem:

AP2/ERF

Apetala2/Ethylene Response Factor compreende uma grande família de fatores de transcrição que são frequentemente envolvidos na regulação de processos de crescimento e desenvolvimento das plantas (DUBOUZET et al., 2003), e respostas a estresses bióticos (NAKANO et al., 2006) e abióticos das plantas (GUTTERSON e REUBER, 2004; LAKHWANI et al., 2016). A expressão de muitos genes AP2/ERF tem sido documentada demonstrando ser regulada por uma variedade de estímulos internos e externos (PRÉ et al., 2008), através de várias vias de sinalização envolvidas por fitormônios, como ácido jasmônico, ácido abscísico, ácido salicílico, etileno, decorrentes de infecção por patógenos e ferimento, e de respostas a outros tipos de estresses como salinidade, seca, frio e elevada temperatura (MCGRATH et al., 2005; LATA e PRASAD, 2011).

A família de fatores de transcrição AP2/ERF possui um domínio de ligação ao DNA altamente conservado, conhecido como domínio AP2, composto por três fitas folha-β antiparalelas e uma fita α-hélice (Figura 5), e constituído de 58-59 aminoácidos

envolvidos na ligação de alta afinidade em sequências alvo de DNA (ALLEN et al., 1998; JOFUKU et al., 2005). O domínio AP2 é essencial para a atividade dos FTs AP2/ERF de ligação ao elementos *cis*-regulatórios (KAGAYA et al., 1999), incluindo o motivo GCC-box, DRE (*dehydration responsive element*)/CRT (*C-repeat element*) (SUN et al., 2008), e motivos CAACA e TTG encontrados nas regiões promotoras dos genes alvos (WANG et al., 2015).

As proteínas AP2/ERF são classificadas de acordo com o número e sequência específica do domínio de ligação ao DNA (AP2, RAV e ERF): Membros da subfamília AP2 contém dois domínios AP2, enquanto que membros da subfamília RAV contém um domínio AP2 e um domínio adicional B3 de ligação ao DNA, e membros da subfamília ERF contém um domínio ERF e um domínio AP2 (RIECHMANN e MEYEROWITZ, 1998). Os genes da subfamília AP2 demonstram participar na regulação de processos de desenvolvimento, como florescência, diferenciação de células epidérmicas foliares e embriogênese (MOOSE e SISCO, 1996; BOUTILIER et al., 2002). O envolvimento de membros da subfamília RAV tem sido demonstrado em respostas aos hormônios etileno e brassinosteróides (ALONSO et al., 2003; HU et al., 2004). Muitas proteínas da subfamília ERF foram identificadas em diversas funções relacionadas a processos celulares, como transdução de sinal hormonal (STOCKINGER et al., 1997), regulação do metabolismo (LIU et al., 1998; GU et al., 2000), processos de desenvolvimento (YAMAMOTO et al.,1999), e respostas a estresses bióticos e abióticos (OHME-TAKAGI e SHINSHI, 1995; ZHANG et al., 2005).

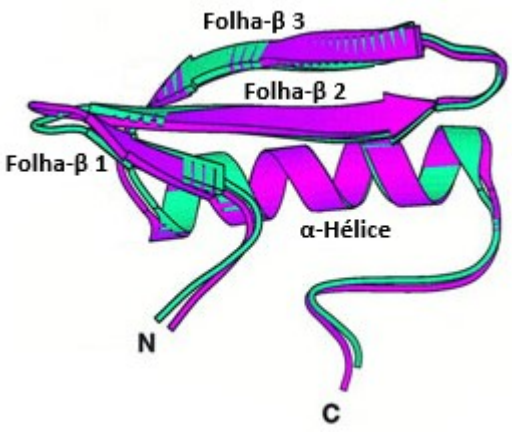


Figura 5. Estrutura 3D do domínio AP2, destacando as três folhas β-pregueadas antiparalelas, uma α-hélice e a indicação dos domínios N-terminal e C-terminal. Adaptado a partir de Allen et al. (1998).

MYB

Funcionalmente diversa e representada em todos os eucariotos, a família MYB contém membros de FTs que são envolvidos em uma variedade de processos específicos de plantas (ZHAO et al., 2008), incluindo morfogênese celular (ISHIDA et al., 2007), metabolismo secundário (LIPPOLD et al., 2009), diferenciação celular (BAUMANN et al., 2007) e resposta a estresses (NESI et al., 2001).

As proteínas MYB são agrupadas em quatro subfamílias, de acordo com os arranjos do domínio MYB de reconhecimento ao DNA (RAMSAY e GONDA, 2008). As subfamílias são compostas por uma, duas, três ou quatro repetições imperfeitas de hélice-volta-hélice (HTH - *helix-turn-helix*) que reconhecem motivos 5'-[GA]GATAA-3', localizados no sulco maior do DNA (YANHUI et al., 2006). Cada repetição é constituída por aproximadamente 50 aminoácidos, contendo resíduos de triptofano regularmente espaçados, formando um cluster de triptofano na estrutura HTH tridimensional (OGATA et al., 1994). A subfamília R0R1R2R3 contém quatro domínios MYB, enquanto a subfamília R1R2R3 e R2R3 contém três e dois domínios, respectivamente, e a subfamília R1-MYB contém apenas um domínio MYB (Figura 6) (DU et al., 2009). Cada proteína MYB contém duas regiões distintas, um domínio conservado N-terminal de ligação ao DNA e um domínio C-terminal consideravelmente diverso, responsável pela modulação da atividade regulatória da proteína (AMBAWAT et al., 2013).

Diferentes genes R2R3-MYB controlam o desenvolvimento da antera, incluindo AtMYB21, AtMYB24, AtMYB57, AtMYB108/BOS1, AtMYB35/TDF1, AtMYB80 e AtMYB99 (CHENG et al., 2009; MANDAOKAR e BROWSE, 2009), e AtMYB23 controla a iniciação do desenvolvimento de tricomas foliares em *Arabidopsis thaliana* (AMBAWAT et al., 2013). Proteínas MYB também são importantes na regulação do desenvolvimento e diferenciação da raiz (FENG et al., 2004; MU et al., 2009) e no controle da biossíntese de flavonoides e do metabolismo de fenilpropanóides (PAZ-ARES et al., 1987; RABINOWICZ et al., 1999). Estudos recentes demonstraram que genes MYB também são regulados pos-tranducionalmente por microRNAs, incluido AtMYB33, AtMYB35, AtMYB65 e AtMYB101, envolvidos no desenvolvimento do pólen e da antera, e que são alvos da família miR159 (ADDO-QUAYE et al., 2008).

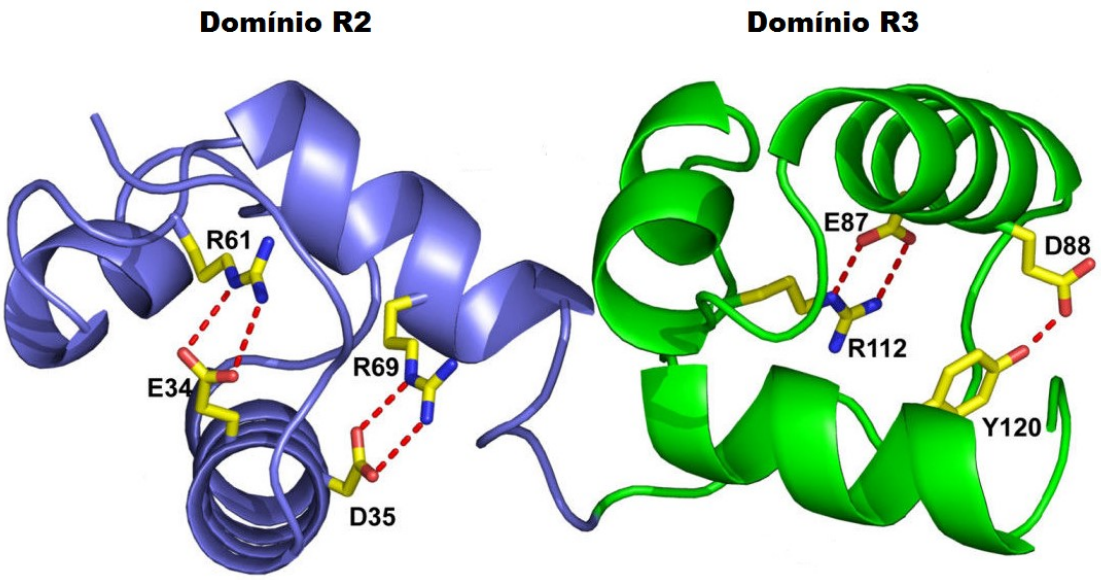


Figura 6. Estrutura dos domínios R2 e R3 de diferentes proteínas MYB. Em destaque, resíduos de aminoácidos de ligação ao DNA no domínio R2 (R61, R69 E34 e D35) e R3 (E87, D88, R112 e Y120). Adaptado a partir de Hichri et al. (2011).

NAC

A família NAC (NAM - *No Apical Meristem*), (ATAF - *Arabidopsis Transcription Activation Factor*) e (CUC - *Cup-Shaped Cotyledon*) compreende um dos grandes grupos de FTs específicos de plantas (PURANIK et al., 2011), representando cerca de 151 genes em arroz (*Oryza sativa*) (NURUZZAMAN et al., 2010), 152 em soja (*Glycine max*) (LE et al., 2011), 117 em *Arabidopsis thaliana* (OOKAA et al., 2003), 152 em Tabaco (RUSHTON et al., 2008), 163 em *Populus trichocarpa* (HU et al., 2010), 147 em *Setaria italica* (PURANIK et al., 2013), 145 em *Gossypium raimondii* (SHANG et al., 2013) e 167 em *Musa acuminate* (CENCI et al., 2014).

Tipicamente, um membro da família NAC contém um domínio N-terminal de ligação ao DNA conservado, compreendido por aproximadamente 160 resíduos de aminoácidos, que é dividido em cinco subdomínios (A, B, C, D e E) (AIDA et al., 1997; OOKA et al., 2003). Em contraste, o domínio C-terminal é uma região de atividade de ligação proteica altamente variável, e confere importantes funções de regulação da atividade transcricional (OLSEN et al., 2005; FANG et al., 2008). A dimerização entre domínios da região N-terminal NAC é comum, e pode funcionar na modulação específica de ligação ao DNA (Figura 7) (MÜLLER, 2001; ERNST et al., 2004).

A participação de proteínas NAC tem sido demonstrada em diversos processos de desenvolvimento, incluindo embriogênese (DUVAL et al., 2002), remobilização de nutrientes (WATERS et al., 2009), controle do ciclo celular (KIM et al., 2007;

WILLEMSEN et al., 2008), morfogênese floral (SABLOWSKI e MEYEROWITZ, 1998), desenvolvimento lateral da raíz (HE et al., 2005), desenvolvimento do meristema apical da folha (NIKOVICS et al., 2006), florescimento induzido por estresse (YOO et al., 2007; KIM et al., 2007), senescência foliar (UAUY et al., 2006) e sinalização hormonal (KIM et al., 2006). Além disso, numerosos FTs NAC têm sido associados com respostas de defesas a estresses bióticos e abióticos como, infecção (REN et al., 2000), seca (TRAN et al., 2004), salinidade (HE et al., 2005; ZHENG et al., 2009) e choque térmico (JIANG et al., 2006). Há também diversas evidências indicando que uma considerável porção de membros de FTs NAC têm função crucial nos processos de desenvolvimento do tecido vascular xilemático (ZHONG et al., 2006; YAMAGUCHI et al., 2008).

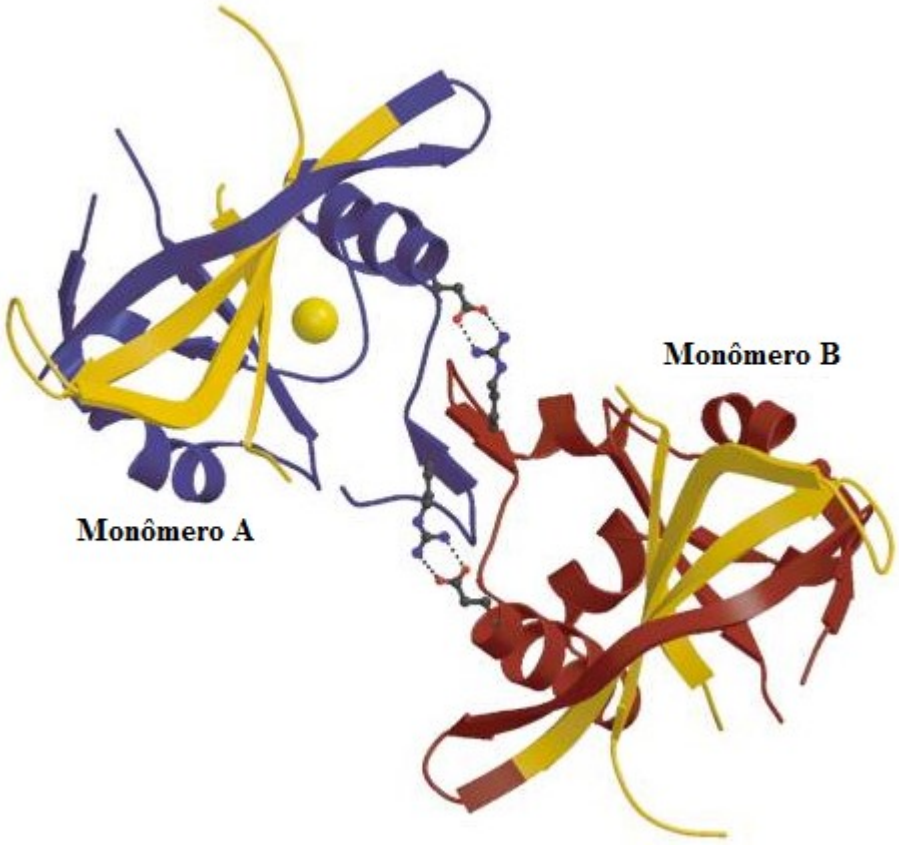


Figura 7. Dimerização funcional entre monômeros de proteínas NAC. Os dois monômeros são mostrados em azul e vermelho, assim como a região responsável pela ligação ao DNA, que é mostrada em amarelo. Adaptado a partir de Ernst et al. (2004).

bHLH

A família de FTs bHLH (*basic/helix-loop-helix*) é composta por proteínas que contém um domínio formado por aproximadamente 60 aminoácidos com duas regiões funcionalmente distintas (TOLEDO-ORTIZ et al., 2003). Uma região N-terminal

básica, com cerca de 15 aminoácidos, envolvida na ligação ao DNA, e uma região HLH C-terminal, com funções de dimerização e ativação transcricional (FERRE-D'AMARE et al., 1994), constituída principalmente por resíduos hidrofóbicos que formam duas α-hélices amfipáticas separadas por uma região de dobra variável em sequência e comprimento (ATCHLEY et al., 1999).

Análises estruturais de cristalografia demonstraram que as interações entre duas regiões HLH separadas por polipeptídeos promovem a formação de homodímeros e/ou heterodímeros, e que a região básica de cada monômero liga-se à metade da sequência de reconhecimento do DNA, formando dímeros simétricos (Figura 8) (MA et al., 1994; SHIMIZU et al., 1997). A sequência central do motivo de reconhecimento do DNA por FTs bHLH é um hexanucleotídeo consenso conhecido por E-box (5′-CANNTG-3′), sendo uma das mais comuns, entre os diferentes tipos de E-boxes, o palíndromo G-box (5′-CACGTG-3′) (TOLEDO-ORTIZ et al., 2003). Acredita-se que a dimerização e a variedade de diferentes E-boxes fornecem mecanismos pelos quais as proteínas bHLH geram uma diversidade de mecanismos de regulação transcricional (FAIRMAN et al., 1993).

Em plantas, as proteínas bHLH têm sido previamente descritas funcionalmente na regulação transcricional associada com a biossíntese de antocianina (WEISSHAAR e JENKINS, 1998), sinalização de fitocromos (ATCHLEY e FITCH, 1997), deiscência de frutos e desenvolvimento de células da epiderme e dos carpelos (MOL et al., 1998; HU et al., 2000).

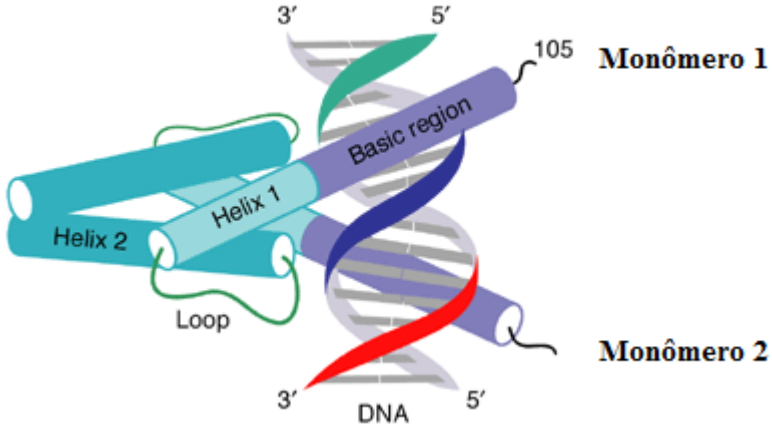


Figura 8. Complexo dimérico entre dois monômeros de domínio bHLH. Hélices reconhecimento de DNA representadas por cilindros. Adaptado a partir de Ma et al. (1994).

WRKY

Os FTs WRKY são conhecidos por regularem a expressão de genes alvos por ligação específica na região promotora contendo a sequência W-box [(T)TTGAC(C/T)] (EULGEM et al., 1999). Análises de ressonância magnética nuclear (NMR) demonstraram que o domínio WRKY é constituído por cerca de 60 resíduos de aminoácidos, que contém uma sequência motivo conservada WRKYGQK na região N-terminal e um domínio zinc finger do tipo C2H2 na região C-terminal (Figura 9) (EULGEM et al., 2000; YAMASAKI et al., 2008). Embora o domínio de ligação N-terminal seja altamente conservado, as estruturas globais das proteínas WRKY são altamente divergentes (WU et al., 2005; PARK et al., 2005).

As proteínas WRKY podem ser classificadas em três grandes grupos, com base no número de domínios WRKY e nas características do motivo zinc-finger (YAMASAKI et al., 2005). O grupo 1 contém dois domínios WRKY incluindo um motivo C2-H2 (C-X4–5-C-X22–23-H-X1-H), enquanto o grupo 2 contém um domínio WRKY incluindo um motivo C2-H2, e o grupo 3 possui um domínio WRKY com um motivo distinto zinc-finger C2-H-C (C-X7-C-X23-H-X1-C) (PARK et al., 2005; ZHOU et al., 2008).

Um número elevado de FTs WRKYs tem sido identificado em plantas (CHEN et al., 2011). Em Arabidopsis foram detectados 74 genes WRKYs (ULKER e SOMSSICH, 2004), cerca de 100 em arroz (SONG et al., 2010), 68 em sorgo (PANDEY e SOMSSICH, 2009), 197 em soja (SCHMUTZ et al., 2010), 80 em pinus (Liu e Ekramoddoullah, 2009), e cerca de 45 em cevada (MANGELSEN et al., 2008). Proteínas WRKY em plantas tem demonstrado envolvimento em respostas relacionadas ao desenvolvimento (XU et al., 2006), resistência a doenças (DONG et al., 2003) e estresses bióticos e abióticos (KARAM et al., 2002; CHEN e CHEN, 2002). Evidências adicionais também demonstraram que FTs WRKYs possuem envolvimento na senescência (ROBATZEK e SOMSSICH, 2002), na via de sinalização de giberelina (ZHANG et al., 2004), assim como na resposta combinada nas condições de estresse hídrico e choque térmico (RIZHSKY et al., 2002), assim como de seca e frio (MARE et al., 2004).

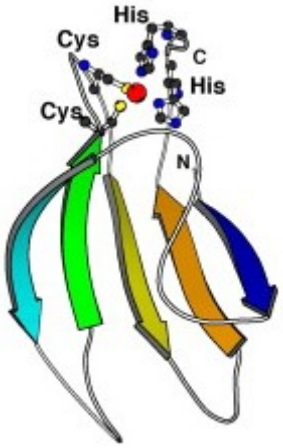


Figura 9. Estrutura do domínio WRKY. A estrutura consiste de cinco folhas-β antiparalelas do domínio N-terminal e resíduos conservados de Cys/His do domínio zinc finger C2H2 na região C-terminal. O íon de zinco é representado por uma esfera vermelha. As regiões N e C-terminal são representadas pelas letras N e C, respectivamente. Adaptado a partir de Yamasaki et al. (2008).

bZIP

Uma das mais diversas famílias de fatores de transcrição, bZIP é caracterizada pela presença de um domínio altamente conservado, composto por 60-80 aminoácidos e subdividida em duas regiões funcionalmente e estruturalmente distintas (WANG et al., 2015). A região básica compreende aproximadamente 16 resíduos de aminoácidos com motivos invariantes N-x7-R/K-x9, que é responsável pela ligação ao DNA em um pseudopalíndromo (AATGACTCAT/TACTGAGTA) e identificação da localização nuclear, enquanto a região *Leu zipper* é composta por repetições heptídicas de leucina ou outro aminoácido hidrofóbico identificado na região C-terminal, sendo responsável por mediar a homo e/ou heterodimerização de proteínas bZIP (Figura 10) (ELLENBERGER et al., 1992; PTASHNE e GANN, 1997).

Genes que codificam fatores de transcrição bZIP têm sido identificados extensivamente em plantas, incluindo Arabidopsis (JAKOBY et al., 2002), arroz (NIJHAWAN et al., 2008), sorgo (WANG et al., 2011), milho (WEI et al., 2012) e cevada (POURABED et al., 2015). Como outros FTs, os membros bZIP são expressos constitutivamente ou num órgão específico, de acordo com a fase de desenvolvimento (CHERN et al., 1996) ou ciclo celular (JAILLON et al., 2007), além de estarem envolvidos em vários processos biológicos, como na regulação da diferenciação de órgãos e tecidos (JAKOBY et al., 2002), embriogênese e maturação de sementes (IZAWA et al., 1994), e ainda na sinalização e repostas a estímulos bióticos e abióticos (RODRIGUEZ-URIBE e O'CONNELL, 2006; SCHMUTZ et al., 2010).

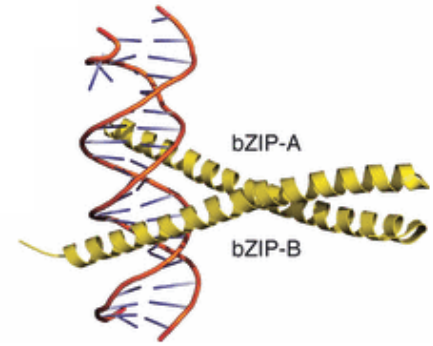


Figura 10. Ilustração da interação dimérica entre α-helices de monômeros bZIP. Adaptado a partir de Ellenberger et al. (1992).

HD-Zip

Encontradas em todos os organismos eucarióticos, os FTs da família HD-Zip são caracterizados por conterem um motivo LZ (*leucine zipper*) que atua como domínio de dimerização com outras proteínas, e um homeodomínio conservado HD, com aproximadamente 60 aminoácidos distribuídos em uma estrutura formada por três α-hélices (Figura 11), que permite a interação específica do domínio com o DNA (MOENS e SELLERI 2006; ARIEL et al. 2010). Em plantas, os FTs HD-Zip têm sido analisados em diferentes espécies, incluindo *Arabidopsis thaliana* (HENRIKSSON et al., 2005), *Oryza sativa* (AGALOU et al., 2008), *Populus trichocarpa* (HU et al., 2012), *Cucumis sativus* (FU et al., 2013) e *Glycine max* (CHEN et al., 2014).

As proteínas HD-Zip podem ser classificados em quatro subfamílias (HD-Zip I, II, III e IV), de acordo com um conjunto de características distintas, que incluem domínio conservado de ligação ao DNA, estrutura gênica, motivos conservados adicionais e funções biológicas (ARIEL et al., 2007). Em Arabidopsis, a subfamília HD-Zip I compreende 17 membros que codificam proteínas de tamanho similar (~35kDa), que ligam-se a sequências pseudopalíndromas [CAAT(A/T)ATTG] (HENRIKSSON et al., 2005). Alguns genes HD-Zip I são envolvidos na via de sinalização do ABA e sacarose, resposta a estresses abióticos, embriogênese e desenvolvimento foliar (HIMMELBACH et al., 2002; JOHANNESSON et al., 2003). A subfamília HD-Zip II consiste de nove membros que, da mesma forma que HD-Zip I, também reconhecem sequências pseudopalíndromas [CAAT(A/T)ATTG], e possuem um motivo denominado CPSCE, que contém aminoácidos conservados de Cys, Pro, Ser, Cys e Glu, localizados próximos ao domínio LZ (TRON et al., 2002; CIARBELLI et al., 2008). A maioria destes genes são principalmente envolvidos no desenvolvimento de órgãos mediado por

fitocromo, como na morfogênese foliar, além de resposta à auxina e alteração lumnítica (MORELLI e RUBERTI, 2002). Os cinco membros incluidos na subfamília HD-Zip III possuem três domínios adicionais (ELHITI e STASOLLA, 2009). MEKHLA é um domínio envolvido possivelmente na sinalização redox e luz (MUKHERJEE et al., 2006), enquanto o domínio START possui capacidade de ligação lipídica (SCHRICK et al., 2004), e SAD é um domínio de ativação transcricional (De CAESTECKER et al., 2000). A subfamília HD-Zip IV compreende 16 membros em Arabidopsis que possuem os mesmos domínios START e SAD de HD-Zip III (ARIEL et al., 2007; ELHITI e STASOLLA, 2009). No entanto, não contém o domínio MEKHLA (MUKHERJEE et al., 2006; CHEN et al., 2014). As proteínas HD-Zip IV ligam-se preferencialmente a motivos que contém uma sequência palindrômica denominada TAA core [GCATT(A/T)AATGC] (OHASHI et al., 2003; NAKAMURA et al., 2006), e estão envolvidas a uma série de respostas a estresses abióticos (YU et al., 2008; HARRIS et al., 2011) e no controle transcricional das células epidérmicas e subepidérmicas (JAVELLE et al., 2011; NADAKUDUTI et al., 2012).

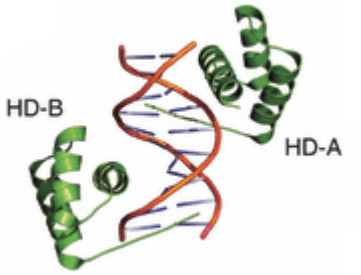


Figura 11. Estrutura tridiminensional de α-hélices, de um arranjo de duas proteínas HD-Zip, que podem se ligar ao suco maior do DNA. Adaptado a partir de Harris et al. (2011).

MADS-box

Proteínas da família de FTs MADS-box, cujo nome se refere a quatro membros identificados originalmente (MCM1, AG, DEFA e SRF), estão presentes em todos os genomas eucarióticos analisados até agora, com maior predominância de membros em genomas de espécies vegetais (SHORE e SHARROCKS., 1995; GRIMPLET et al., 2016). Análises estruturais de proteínas MADS-box têm demonstrado uma caracterização modular fundamental, denominada MICK, proveniente das regiões N-terminal e C-terminal, onde quatro domínios podem ser identificados (M - *MADS-box*; I - *intervening*; K - *keratin-like*; C - *C-terminal domains*), sendo o domínio M o mais conservado, com cerca de 56 aminoácidos, e com a função principal de ligação ao DNA

e um menor papel de dimerização (TONACO et al., 2006). O domínio I é menos conservado e tem uma função específica na formação de dímeros de ligação ao DNA (MASIERO et al., 2002), enquanto o domínio K, de aproximadamente 80 resíduos, contém repetições heptídicas em dobras de α-hélices anfipatícas que medeiam o processo de dimerização e formação de complexos multiméricos (EGEA-CORTINES et al., 1999). O domínio C-terminal, predominantemente hidrofóbico, é o menos conservado e responsável pela ativação transcricional através de interação proteica (KAUFMANN et al., 2005).

Os FTs MADS-box são envolvidos em diversas atividades biológicas de plantas, sendo identificados inicialmente como reguladores do desenvolvimento floral (SHORE e SHARROCKS, 1995; GRAMZOW et al., 2010). O modelo ABC tem sido proposto para explicar como é orquestrado a identificação do desenvolvimento floral em *Arabidopsis thaliana* por meio dos genes AG, AGL1-6, AP1-3 e PI, os quais atuam individualmente ou em combinação para ditar a ativação do desenvolvimento de uma região específica do meristema floral (FLANAGAN e Ma, 1994; SHORE e SHARROCKS, 1995). Em *Antirrhinium majus*, este modelo é caracterizado por meio dos genes MADS-box DEFA, GLO, PLE, SQA, DEFH24/49, enquanto em *Petunia hybrida* é intermediado por fbp-2 e pMADS1-2 (DAVIES e SCHWARZ-SOMMER, 1994). Adicionalmente, padrões de expressão de genes MADS-box sugerem que subconjuntos particulares de genes podem estar envolvidos no controle de outros processos, como embriogênese e desenvolvimento de frutos e raízes (SHORE e SHARROCKS, 1995).

Interações proteína-proteína entre FTs MADS-box e outros FTs parece ser comum, indicando que as especificidades de tais oligomerizações são essenciais na formação de um complexo regulatório transcricional específico, sendo um dos mecanismos para a seletividade da ativação de genes alvos (MASIERO et al., 2002). Em *Antirrhinum majus*, a formação de um complexo multimérico entre as proteínas MADS-box SQUA (*SQUAMOSA*), DEF (*DEFICIENS*) e GLO (*GLOBOSA*), envolvidas no controle da arquitetura floral, demonstrou alteração na afinidade e na força de ligação dos FTs em motivos CArG-box [CC(A/T)6GG] encontrados em promotores de genes alvos (Figura 12) (SOMMER et al., 1990; EGEA-CORTINES et al., 1999). A interação heteroligomérica do complexo DEF/GLO/SQUA via domínios C-terminal das proteínas apresentaram propriedades moleculares e fenotípocas distintas quando

comparadas ao heterodímero DEF/GLO e ao homodímero SQUA/SQUA (TRÖBNER et al., 1992; ZACHGO et al., 1995).

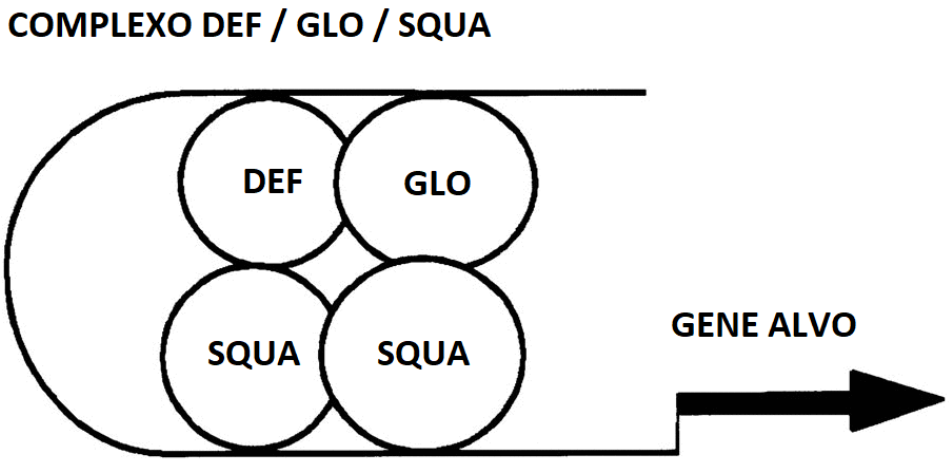


Figura 12. Modelo molecular do complexo multimérico DEF/GLO/SQUA baseado nos fenótipos e dados moleculares. A formação heteroligomérica das proteínas DEF/GLO e SQUA produzem um complexo de ligação ao DNA com maior força e afinidade nos genes alvos quando comparada com os respectivos homodímeros ou heterodímeros. Adaptado a partir de Egea-Cortines et al. (1999).

3.6 PROMOTORES REGULATÓRIOS

Promotor é um fragmento de DNA composto por motivos de elementos *cis*-regulatórios que são responsáveis por inicia e regular a transcrição de um gene (MITHRA et al., 2017). O estudo dos promotores dos genes é primordial para o entendimento global da regulação da expressão dos genes em plantas (HERNANDEZ-GARCIA e FINER 2014). Embora a participação de um determinado segmento de DNA na regulação da expressão gênica só poder ser demonstrada experimentalmente, o promotor dos genes é convencionalmente divido em três regiões (Figura 13): a) o promotor central (*Core promoter*), cuja localização encontra-se na posição entre +1 e -100 pb em referência ao sítio inicial de transcrição +1 TSS (*Transcription Start Site*), e possui um papel central na formação de complexos de pré-iniciação da transcrição, com elementos *cis*-regulatórios como TATA-box, BRE (*TFIIB Recognition Element*), Inr (*Initiator*) e MTE (*Motif Ten Element*), que são tipicamente encontrados nesta região; b) seguindo no sentido *upstream* do TSS, a região proximal (*Proximal promoter*) é localizada na posição entre -101 e -1000 pb, e funciona como um local de ancoramento para a maioria dos fatores de transcição; c) a região distal (*Distal promoter*), localizada entre -1001 e -3000 pb, contém elementos *cis*-regulatórios que são comumente conhecidos como estimuladores (*enhancers*) e silenciadores (*silencers*) ou repressores,

que promovem contribuem para ativação e repressão, respectivamente, da transcrição (LICHTENBERG et al., 2009). Além destes, os isoladores (*insulators*), através da ligação do seu fator de transcrição correspondente, impedem a interação entre um estimulador ou silenciador e um fator de transcrição cognato (Ashley et al., 2008).

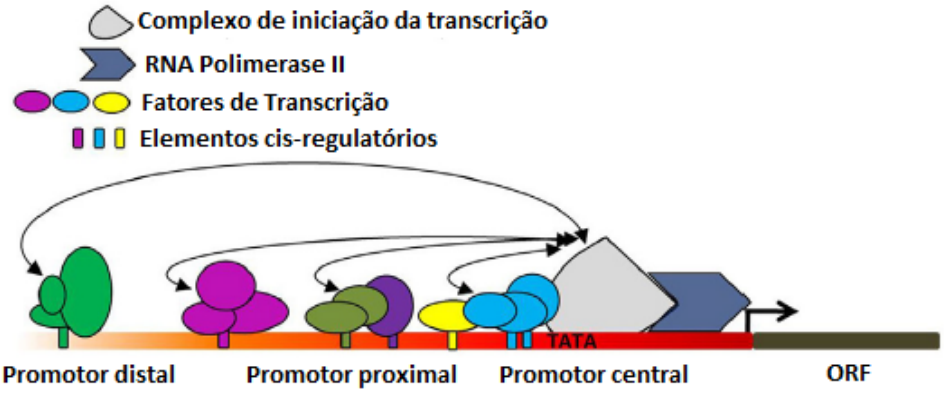


Figura 13. Modelo simplificado do posicionamento dos promotores de genes em eucariotos. Adaptado a partir de Hernandez-Garcia e Finer (2014). ORF (*Open Reading Frame*).

Os promotores de acordo com a aplicação na biotecnologia podem ser agrupados em quatro classes:

Promotores constitutivos – Regulam a expressão dos genes na maioria dos tecidos e durante todo o ciclo de vida das plantas, direcionando a expressão em muitos órgãos e tecidos e sob variadas condições, além de poderem atuar em níveis de expressão moderada quanto em alta expressão, como por exemplo, em tecidos meristemáticos de crescimento rápido e tecidos vasculares (HERNANDEZ-GARCIA e FINER, 2014). Os promotores constitutivos também são frequentemente utilizados na avaliação dos níveis de expressão de transgenes em vários tecidos e em todos os estágios de desenvolvimento das plantas (BENFEY e CHUA, 1990).

Promotores espaço-temporais – Regulam a expressão dos genes de forma restrita a determinadas células, tecidos e órgãos, assim como a certos estágios de desenvolvimento (HERNANDEZ-GARCIA e FINER 2014). Os promotores espaço-temporais mais frequentemente reportados são os específicos de genes induzidos durante o desenvolvimento da semente (KAWAKATSU e TAKAIWA, 2010), embora os específicos de tecidos de frutos têm merecido ateção especial devido à possibilidade do melhoramento do valor nutricional (COSGROVE, 2000). Os promotores espaço-temporais os específicos do polén e da antera são úteis para o controle da esterilidade

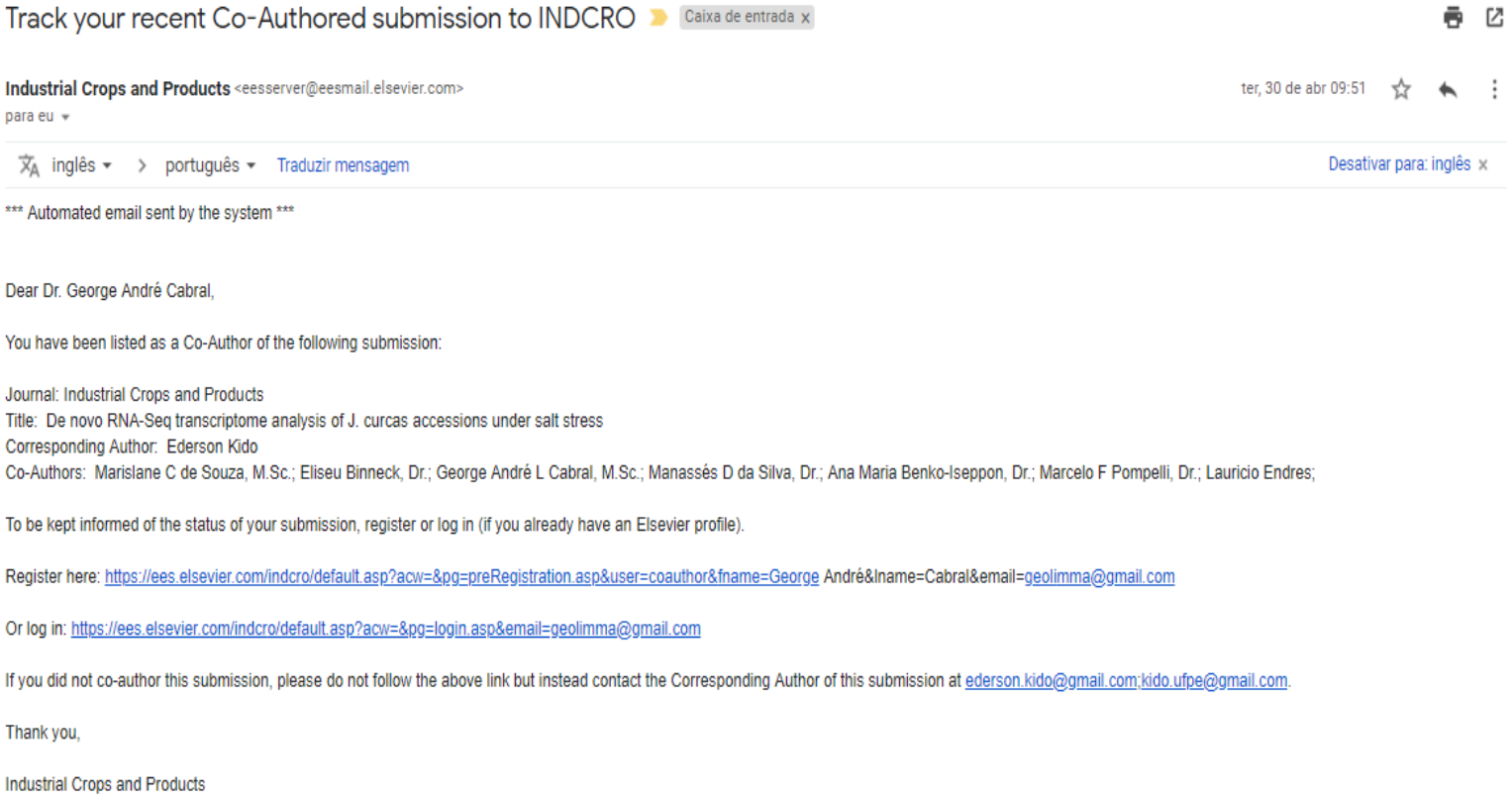
masculina, que é um importante traço na reprodução das plantas (PEREMARTI et al., 2010).

Promotores induzíveis – São responsíveis a estímulos ambientais, fornecendo uma regulação precisa da expressão dos genes através do controle externo, como resultado de estressse bióticos ou abióticos e sinais endógenos (HERNANDEZ-GARCIA e FINER, 2014).

Promotores sintéticos – São compostos por combinações de sequências em promotores centrais, proximais, distais, além de sequências intrônicas que possam conter regiões regulatórias, onde arranjos de elementos *cis*-regulatórios dentro de um promotor sintético podem resultar em uma expressão mais específica, intensa e precisa do transgene (RUSHTON et al., 2002).

4 RESULTADOS

4.1 *De novo* RNA-Seq transcriptome analysis of *J. curcas* accessions under salt stress



Track your recent Co-Authored submission to INDCRO — Caixa de entrada x

Industrial Crops and Products <eesserver@eesmail.elsevier.com> ter, 30 de abr 09:51

para eu

inglês › português Traduzir mensagem Desativar para: inglês x

*** Automated email sent by the system ***

Dear Dr. George André Cabral,

You have been listed as a Co-Author of the following submission:

Journal: Industrial Crops and Products
Title: De novo RNA-Seq transcriptome analysis of J. curcas accessions under salt stress
Corresponding Author: Ederson Kido
Co-Authors: Marislane C de Souza, M.Sc.; Eliseu Binneck, Dr.; George André L Cabral, M.Sc.; Manassés D da Silva, Dr.; Ana Maria Benko-Iseppon, Dr.; Marcelo F Pompelli, Dr.; Lauricio Endres;

To be kept informed of the status of your submission, register or log in (if you already have an Elsevier profile).

Register here: https://ees.elsevier.com/indcro/default.asp?acw=&pg=preRegistration.asp&user=coauthor&fname=George André&lname=Cabral&email=geolimma@gmail.com

Or log in: https://ees.elsevier.com/indcro/default.asp?acw=&pg=login.asp&email=geolimma@gmail.com

If you did not co-author this submission, please do not follow the above link but instead contact the Corresponding Author of this submission at ederson.kido@gmail.com;kido.ufpe@gmail.com.

Thank you,

Industrial Crops and Products

Marislane Carvalho Paz de Souza[1], Eliseu Binneck[2], George André de Lima Cabral[1], Manassés Daniel da Silva[1], Ana Maria Benko Iseppon[1], Marcelo Francisco Pompelli[3], Laurício Endres[4], Éderson Akio Kido[1*]

[1]Federal University of Pernambuco (UFPE), Bioscience Center, Department of Genetics, 50670-420 Recife, PE, Brazil
[2]Brazilian Agricultural Research Corporation (EMBRAPA), EMBRAPA-Soybean, 86001970 - Londrina, PR, Brazil
[3]Federal University of Pernambuco (UFPE), Bioscience Center, Department of Botany, 50670-901 - Recife, PE, Brazil
[4]Federal University of Alagoas (UFAL), Agricultural Sciences Center, 57072-970 - Maceió, AL, Brazil
* Corresponding author (e-mail address: kido.ufpe@gmail.com)

Abstract

Jatropha curcas (physic nut) is an oleaginous, non-food, small tree that represents an excellent source of biodiesel, among other products, including some medicinal ones. In this way, *J. curcas* could be an option for farmers in tropical and semi-arid regions, especially where salinity can compromise production and yield of crops. We generated and analyzed the root RNA-Seq transcriptome of two *J. curcas* accessions (salt-tolerant Jc183, and salt-sensitive Jc171) after salt-treatment (150 mM NaCl/three hours). The *de novo* transcriptome covering 101 Mb assembled 145,422 transcripts (126,343 UniGenes), and around half of them encoded predicted proteins. Comparing the total of differentially expressed genes (DEGs) by Jc183 (57) with that of Jc171(4,646), the intensive transcriptional effort of Jc171 stood out, probably trying to minimize the damages, some of them visually observed only in Jc171 leaves. The functional characterization of DEGs by the MapMan software, using *M. esculenta* genes as the reference, associated them with metabolic processes showing impacts on plant salt-stress responses. The associated processes covered the metabolisms of hormones, CHOs, lipids, amino acids, redox, and also secondary metabolites. Further, nine selected DEGs, including *S-adenosylmethionine synthase* (*SAM*), *carboxylesterase* (*CXE*), *Phenylalanine ammonia-lyase* (*PAL*), *homeobox-leucine zipper protein* (*HD-Zip*), *NAC TF* gene, *methionine-gamma lyase* (*MGL*), *S-adenosylmethionine-dependent methyltransferase* (*SAMe*), *peroxidase* (*PX*), and *xyloglucan endotransglucosylase* (*XTH*), were evaluated by RT-qPCR analysis, trying to validate their *in silico* expression. The developed functional molecular markers could be useful in marker-assisted selection process in breeding programs. Additionally, the data covering DEGs, provided as Supplementary information, may help to understand the molecular mechanisms involving *J. curcas* plants responding to salt stress, which is crucial for the development of salt-tolerant plants.

Keywords: Bioinformatic; Abiotic stress; Euphorbiaceae; Physic nut

Introduction

Physic nut (*Jatropha curcas* L.;) is an oilseed plant from Euphorbiaceae Family, with a widespread distribution, covering tropical and semiarid regions (Africa, Asia, and South America), and along all the Brazilian regions. Its occurrence combines different climates and soils conditions, including marginal agricultural production areas, and

those relatively deficient in nutrients (Wang et al., 2018; Beltrão e Oliveira, 2008). *J. curcas* seeds present expressive oil contents (40 - 50%), with a relatively high amount of unsaturated fatty acids, which are the preferred by the biodiesel industries (Pramanik, 2003; Wang et al., 2018; Grover et al., 2014).

In Brazil, biodiesel production has increased in the last decade, mainly from soybean oil processing. However, soybean is a versatile crop, very important for human and animal nutrition. *J. curcas*, instead, not compete with the food market, and could be an option to be explored. Although, *J. curcas* is a relatively drought-tolerant plant, it is also considered a salt-sensitive one. Salinity in the soil is one of the leading causes of losses in agricultural production worldwide (FAO, 2015). It is estimated that globally 20% of all irrigated land are currently affected by salt stress (Taiz et al., 2017). Several factors, including soil composition, poor drainage, and inadequate irrigation could provoke salt stress. The combination of these factors became the Brazilian Northeast region potentially sensitive to salinity.

Plants in saline soils may grow irregularly due to osmotic stress caused by reduced water absorption, and the high concentration of ions, which interfere with the nutrient uptake, raising cytotoxicity (Munns, 2005). To understand the molecular mechanisms by which plants respond to salt stress is crucial to plant breeding programs. In this way, comparing the global expressed profiles of distinct accessions responding to salt NaCl is beneficial to uncover genes related to salt-tolerance mechanisms. Some transcriptomic approaches have been applied to *J. curcas* plants responding to abiotic stress, such as cold (Wang et al., 2013; Wang et al., 2014), flooding (Juntawong et al., 2014), drought (Cartagena et al, 2015; Sapeta et al., 2015; Zhang et al., 2015), and also salinity (Zhang et al., 2014). However, in the present study, plants of two Brazilian *J. curcas* accessions were exposed (three hours) to salt NaCl (150 mM), and RNA-Seq libraries generated of their roots allowed the comparison of gene expression profiles, aiming to develop functional molecular markers potentially useful to assist selection steps in breeding programs.

Materials and Methods

Plant materials and NaCl treatment

Two *J. curcas* accessions (Jc183 and Jc171) from the seed bank of the Semiarid Tropical Agricultural Research Center (Embrapa Agroenergia, Brasília, DF - Brazil), initially collected in different Brazilian regions (Supplementary Table S1), were selected for the salt assay, based on previous studies (Lozano-Isla et al., 2018). The Jc183 accession was considered the salt-tolerant, concerning the accession Jc171. The conducted salt assay followed a completely randomized experimental design with two accessions, two treatments (without salt or with NaCl, 150 mM, three hours of salt exposure), and three biological replicates of each accession. Seeds of homogeneous sizes and weights were sown (March 2016) in pots (50 L) containing washed sand (20 kg), being cultivated in greenhouse at UFAL/CECA (Rio Largo, AL, Brazil; geodesic coordinates: 09º28'02" S; 35º49'43"W; altitude: 127 m; climate: humid, metathermic, moderate water deficiency in the summer (December - March), and water excess in winter, according to Thornthwaite and Mather method (1955). After the first eophylls (10 DAG, days after germination), seedlings were thinned to one plant per pot. Plants were irrigated every three days with Hoagland nutrient solution (Epstein, 1972) with one-fifth strength. Seven days before the salt application (65 DAG), plants were irrigated daily with Hoagland nutrient solution full strength. On the day before the salt application, plants were irrigated at 16 h. The salt was applied at 9-10 h and consisted of NaCl (150 mM) added to the Hoagland solution. After salt exposure (three h), roots were collected, immediately frozen in liquid nitrogen, and stored (-80 °C) until RNA extraction.

RNA extraction and RNA-Seq libraries

Total RNAs were extracted from roots using SV Total RNA Isolation kit (Promega, USA), following the manufacturer's instructions. RNAs integrities were verified in RNA agarose gel (1.5% w/v), and the RNAs concentrations estimated in NanoDrop 2000 spectrophotometer (Thermo Scientific ™). RNAs showing absorbance ratio 260/280 nm close to 2.0, and a minimum of 50 µL RNA solution (80 ng/µL) were sent to ESALQ - Genomic Center (São Paulo University, Piracicaba, SP, Brazil) for the RNA-Seq libraries generation and sequencing. All the RNAs integrities were re-evaluated, using the Agilent 6000 Bioanalyzer (Agilent Technologies, CA, USA). The

total of 12 RNA-Seq libraries (two accessions x two treatments x biological triplicates) was generated following the Illumina TruSeq Stranded mRNA Sample Prep kit (Illumina Inc, CA, USA), "LS" Protocol. Libraries were sequenced on an Illumina HiSeq 2500 (100 bp, single-end reads), using flow Cell HiSeq run with the HiSeq SBS v4 chemistry.

De novo transcriptome assembly and DEGs identification

RNA-Seq raw sequence data were analyzed (FastQC v0.11.5) for reads qualities, before and after the initial filtering and trimming, using default parameters of the Trimmomatic tool (v.0.36; Bolger et al., 2014). Reads showing low quality, or unknown adapters and nucleotides, were excluded. Pairs of high-quality reads (Phred quality score, $Q \geq 30$ for all bases) were used for *de novo* transcriptome assembly performed with the Trinity software v.2.2.0 (Grabherr et al., 2011). A de Bruijn graph data structure represented the overlapping among the reads, and short reads with overlap regions were assembled into longer contigs. The longest transcripts in the cluster units were regarded as unigenes to eliminate redundant sequences. The alignment package Bowtie (v4.4.7; Langmead et al., 2009) was used to map reads back to unigenes. According to the comparison results, the expression levels were estimated employing RSEM (RNA-Seq by expectation maximization; Li and Dewey, 2011). Differences of an abundance of the unigene expression among the samples were represented using FPKM (Fragments Per Kilobase of transcript per Million mapped reads) method. Matrices of normalized FPKM values generated from the RSEM counts were used for the differential expression analyzes between the experimental conditions using the edgeR package (Robinson et al., 2010). Differentially expressed unigenes, henceforth, DEGs, for brevity, were determined based on *p-value* ≤ 0.0001, false discovery rate (FDR ≤ 0.005), and fold change (FC) based on Log_2 (FC) ≥ 1 (positive expression modulation) or ≤ -1 (negative modulation). FC is the ratio of the unigene abundance considering two RNA-Seq libraries.

Functional annotation of the assembled transcripts

Assembled transcripts (henceforth, transcripts, for brevity) were annotated using the BLASTx alignment (e-value $\leq 10^{-10}$) to various protein databases, including sets downloaded from *J. curcas* presented in the UniProtKB database

(http://www.uniprot.org/), and those including the taxonomically related species *Manihot esculenta* and *Ricinus comunnis* from the *Phytozome* portal (v.12.1.6; https://phytozome.jgi.doe.gov/pz/portal.html). Annotation contributions based on each dataset were observed by Venn diagrams (Oliveros, 2015). Also, a second round of functional annotation using Trinotate pipeline (https://trinotate.github.io/) were performed (BLASTx, e-value $\leq 10^{-5}$), against several databases, including: NCBI (non-redundant) protein database (Nr) (ftp://ftp.ncbi.nih.gov/blast/db/), UniProt/SwissProt database, Kyoto Encyclopedia of Genes and Genomes (KEGG), GO (Gene Ontology), eggNOG and InterproScan. Trinotate also provided a web-based graphical interface to support local user-based navigation of annotations and differential expression data (Bryant et al., 2017).

Metabolic pathways, and heatmaps

Metabolic pathways associated with DEGs were identified applying MapMan software (v.3.6.0; Thimm et al., 2004), using *M. esculenta* best hits from BLAST alignments. Unigenes were hierarchically clustered by the Cluster 3.0 software (https://cluster2.software.informer.com/3.0/), based on the FC values modulated in the comparison stressed *versus* negative control (henceforth, S *vs.* C, for brevity), being the clusters visualized as heatmaps using the JavaTreeview v.1.1 software (Saldanha, 2004).

Expression validation by RT-qPCR

The expression analysis by RT-qPCR assay validated DEGs representing candidates selected based on the functional annotation and their expression modulated on the contrast *S vs. C*. To this end, cDNAs from RNAs pre-treated with DNase were evaluated on real-time PCR Thermocycler LineGene 9600 (Bioer, Hangzhou, China), in reactions including biological and technical triplicates for each experimental treatment, negative controls, and two reference genes [actin (Tang et al., 2016), and β-tubulin (Xu et al., 2016)], tested adequately for this purpose. The proposed *primer* pairs (Supplementary Table S2) were designed (Primer 3 tool; Rozen and Skaletsky, 2000) based on the *J. curcas* transcripts and the following parameters: amplicon size (70 - 200 bp), melting temperature [50 - 80ºC, 70ºC (optimum)], and GC content (45 - 55%). *Primers* synthesized by Bioneer Corporation (South Korea) first amplified cDNAs in a conventional PCR test. After that, the RT-qPCR reactions (10 µL) included: 1 µL cDNA

(sample diluted 1/5), 5 µL SYBR™ Green (GoTaq ® qPCR Master Mix, Promega), 0.3 µL of each *primer* (5 µM) and 3.4 µL ddH_2O. The reactions followed the settings: initial denaturation of 95°C for 2 min, followed by 40 cycles of 95°C for 15 s, and 60°C for 60 s. The dissociation curves were obtained heating the amplicons from 65 to 95°C for 20 min after the RT-qPCR cycles. The LineGene software (v.1.1.10) estimated the Tm and Cq values, and the absolute and relative quantifications. Relative expression data were evaluated using the REST 2009 software (Relative Expression Software Tool v.2.0.13; Pfaffl et al., 2002), applying randomization test with 2,000 permutations, and testing the hypothesis of significant differences between the control and treatment groups. Also, the MIQE protocol was followed to increase the results reliabilities (*The Minimum Information for Publication of Quantitative Real-Time PCR Experiments*; Bustin et al., 2009).

Results and discussion

Visible damages on the J. curcas leaves

After three hours of salt exposure (150 mM NaCl), the salt-sensitive Jc171 accession presented visible damages on leaves, not observed in the salt-tolerant Jc183, which included, leaves slightly curved, wilted looking, and brown colored areas on the edges, progressing to necrosis (Fig.1). Walia et al., (2005) also noted visual damage of salinity stress on leaves of the sensitive rice cultivar IR29, in the form of necrosis at about one-third of leaf length. In the upland cotton (*G. hirsutum*) visible damage owing to salinity stress appeared on the leaves of salt-sensitive Nan Dan Ba Di Da Hua genotype, after 200 mM NaCl treatment (Peng et al., 2014); the same authors pointed that after 0.5 h, distinct wilting and dehydration were observed on leaves of the analyzed genotypes, and after four h, both genotypes showed more severe wilting. All mentioned damages can significantly compromise photosynthesis, influencing the growth and development of plants and their yields directly.

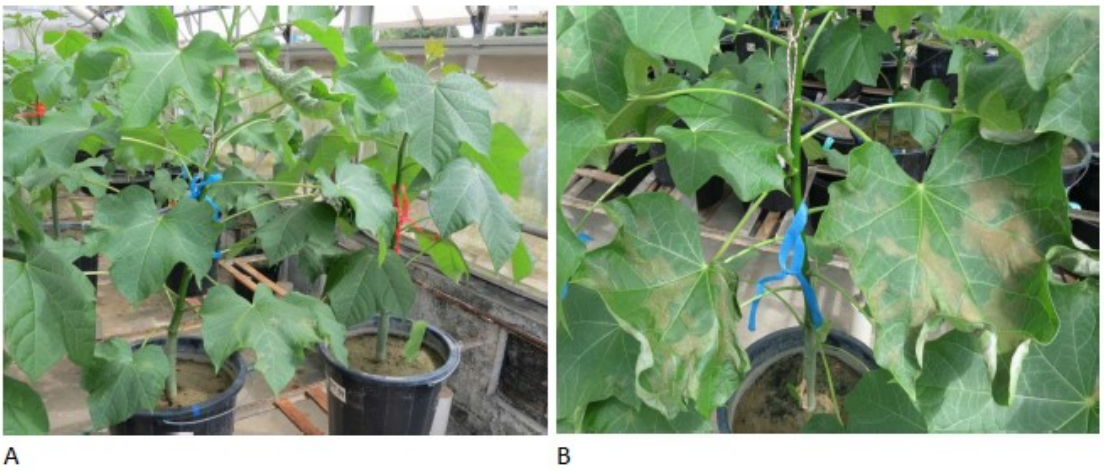


Figure 1. Aspects of *Jatropha curcas* leaves after three hours of exposition to Hoagland solution plus NaCl solution (150 mM). A) Both accessions: the salt-sensitive Jc171 (blue ribbon) and the salt-tolerant Jc183 (red ribbon) accession. B) Leaves (Jc171) showing visible damages: browning at the edges and brown spots.

The J. curcas transcriptome

The RNA-Seq libraries (12) generated with high-qualities RNAs (RIN > 8, Supplementary Fig. S1) after Illumina HiSeq 2500 sequencing provided for each accession (three biological replicates x two treatments) similar amounts of raw *reads* [113,668,090 (Jc183), and 124,618,733 (Jc171)]. Those *reads* showing good qualities (*Phred score* > 30), after filtering and trimming of adapters and low-quality bases, comprised 96,9 % (Jc183), and 96,3 % (Jc171) of the reads. The total of *reads* before and after the trimming step, concerning each library is shown in Supplementary Table S3. The *de novo* transcriptome generated based on the assembled transcripts covered 101 MB (based on unigenes, the size was 77 MB). The transcriptome also covered 145,422 transcripts (126,343 unigenes) and presented a GC percentage of 41.55. The N50 for transcripts comprised 1308 bp (based on unigenes, 993 pb), which is the maximum length where at least 50% total assembled sequence resides in contigs of at least that length; further details of the generated transcriptome are shown in Supplementary Table S4.

Functional annotation of the assembled transcripts

The first round of transcript (145,422) annotation, exploring *J. curcas* proteins (UniProtKB, 27,650 deposited sequences) and those from taxonomic-related species [*M. esculenta* (41,381) and *R. comunnis* (31,221), *Phytozome* database], in individually BLASTx analysis (*e-value* $\leq e^{-10}$), identified 27,363 transcripts encoding predicted

proteins observed in Euphorbiaceae family. Specifically, 25,668 transcripts associated directly with the *J. curcas* proteins, and other 1,695 transcripts with proteins from the two related species (Fig. 2A). Those transcripts encoding conserved proteins associated with the three species were 19,508 (Fig. 2A). From the set of 25,668 transcripts associated directly with the *J. curcas* proteins, only 9,288 transcripts (Fig. 2B) were properly annotated (protein name or gene function) based on those associated proteins. About the proteins properly annotated from the *J. curcas*-related species, another 14,642 transcripts were declared annotated (Fig. 2B). Considering only the sequence similarities with the *J. curcas* transcripts, the exclusive contribution of *R. cumunnis* exceeded that of *M. esculenta* (Fig. 2A), however, when considered the informative annotations, that position inverted (Fig. 2B).

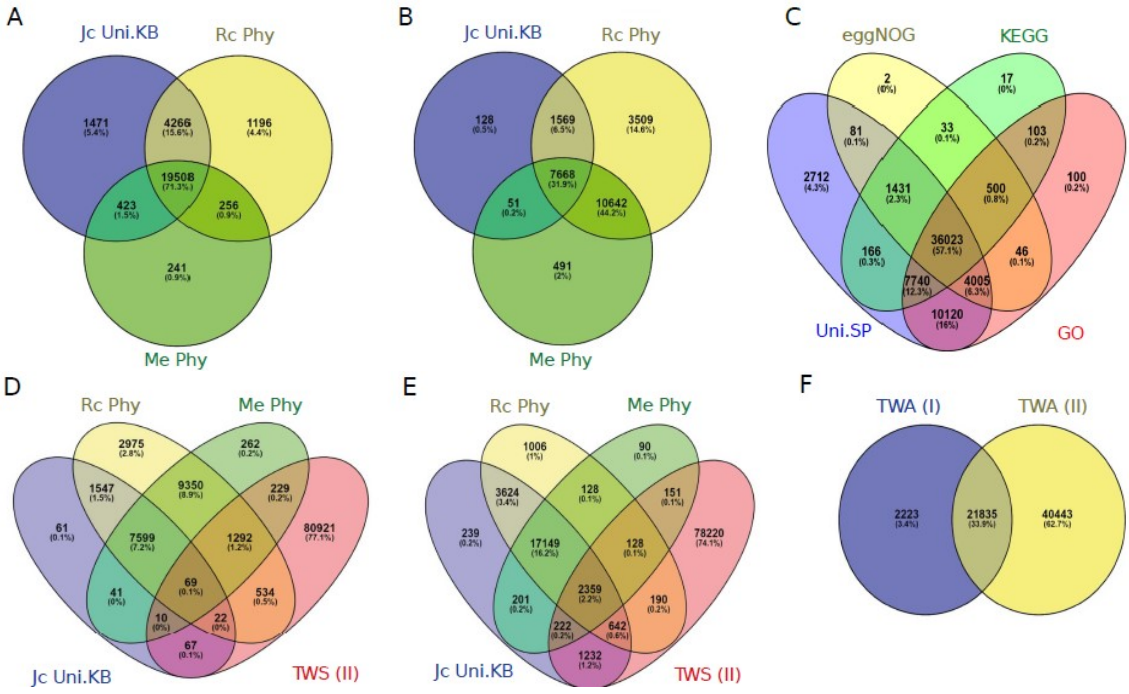


Figure 2. Venn diagrams comparing different results from the two annotation rounds of the RNA-Seq *J. curcas* transcripts: the first (I) round (BLASTx, *e-value* $\leq e^{-10}$) against proteins of *J. curcas*/UniProtKB (Jc UniP), *Ricinus communis*/Phytozome (Rc Phy) and *M. esculenta*/Phytozome (Me Phy), and the second (II) performing the Trinotate software (*e-value* $\leq e^{-5}$) against several protein databases (UniProt/SwissProt, KEGG, eggNOG, GO). A) The individual and shared contribution of the datasets used in I round, considering only similarities results. B) Comparison described before (A), but only considered the annotated results. C) The individual and shared contribution of the datasets used in the II round. D) Transcripts annotated by the I round compared with transcripts without similarities (TWS II) from the II round. E) transcripts encoding predicted proteins similar to those from the I round, regardless of whether there is functional annotation, compared with the transcripts without similarities (TWS II) from the II round. F) Transcripts annotated (TWA) by the two annotation rounds.

In short, concerning the 27,661 proteins from the UniProtKB database (January 2019), most of them (24,288) are annotated as *uncharacterized proteins*. The inclusion of the proteomes from *R. cumunnis* and *M. esculenta*, increased the annotation efficiency, especially *M. esculenta* (Phytozome database). The second round of annotation with the *Trinotate* pipeline (Bryant et al., 2017), considering a less stringent analysis (*e-value* $\leq e^{-05}$) and several databases, showed 63,079 transcripts encoding predicted proteins similar to those from UniProt/SwissProt (62,278 transcripts), eggNOG (42,121), KEGG (46,013), and Gene Ontology (58,637). The individual contribution of each database (Fig. 2C) highlighted the contribution of the Uniprot/SwissProt database, which is formed of manually cured sequences. However, 83,144 transcripts did not reach the required similarity threshold. Comparing the two annotation rounds, from the 83,144 transcripts failing to hit the II threshold (II round), some of them were previously annotated with *J. curcas* proteins or its related species (I round), but 80,921 remained non-associated (*transcripts without similarity* - TWS II, Fig. 2D). If discounting those transcripts showing acceptable similarities with the *J. curcas* and the related species (probably non-annotated), 78,220 transcripts remained without reaching the threshold (Fig. 2E).

In turn, the individual and overlapped contribution of the two annotation rounds are shown in Fig. 2F. Besides the 24,058 (21,835+2,223) transcripts annotated by I round analysis (Euphorbiaceae family members), another 40,443 transcripts were annotated by the II round analysis (Fig. 2F). In total, 64,501 transcripts were adequately annotated, while 80,921 were non-annotated (78,220 not reaching the similarity threshold after II rounds of annotation). Both sets are available for further researches.

The differentially expressed genes (DEGs) in response to the salt stimulus

The analysis of the expressed profiles of Jc183 and Jc171, comparing the respective contrast *S vs.C*, and considering the required thresholds (*p-value* ≤ 0.0001, FDR ≤ 0.005, Log_2 FC ≥ 1 or ≤ -1), identified 57 and 4,646 DEGs, respectively. The salt-sensitive Jc171 accession presented more transcriptional effort responding to the salt stimulus, trying to minimize the damages, as those visually observed in its leaves (Fig. 1). In a similar pattern, the stress response of the salt-sensitive rice IR29 genotype was characterized by a relatively large number of induced probe sets (in a GeneChip analysis, using the rice genome Affymetrix array), when compared to the salt-tolerant

FL478 (Walia et al., 2005). Another assay (Walia et al., 2007) using the same GeneChip, including two *japonica* rice lines (Agami and M103), besides the two *indica* lines mentioned above, revealed a strikingly large number of induced genes, in response to the applied salinity stress, by the sensitive lines (IR29 and M103) in relation to the tolerant ones.

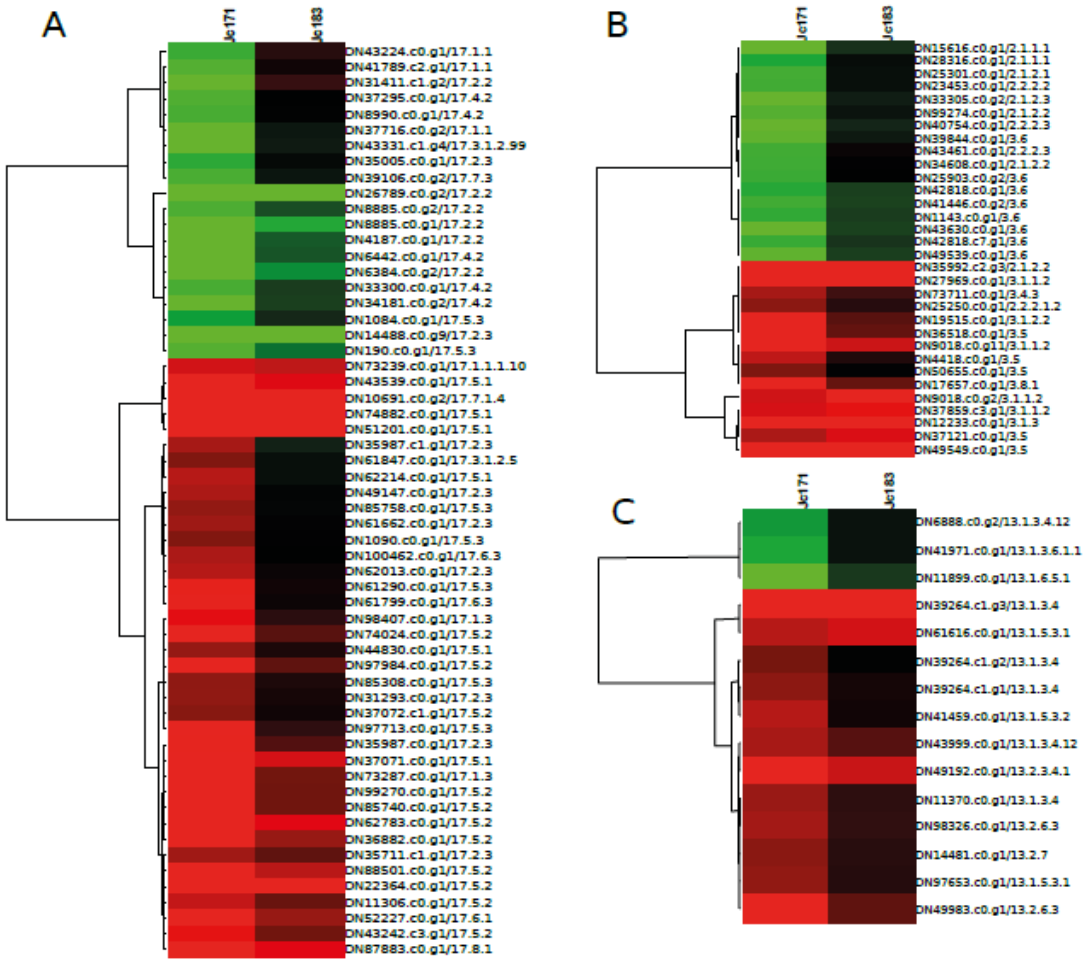
From the Jc183 DEGs (40 UR and 17 DR; Supplementary Table S5), 23 (13 UR and 10 DR) were exclusively DEGs by Jc183 (five non-annotated, three URs and two DR), and 34 were DEGs also detected by Jc171, 33 of them presenting the same regulation by both accessions (27 UR and six DR), and one divergence (DR/Jc183, and UR/Jc171 for DEG encoding Galactinol synthase 1). From the set of 34 shared DEGs, five remained non-annotated (four UR by both accessions, and one DR also by both accessions; Supplementary Table S5). Details of the Jc183 DEGs (Trinity ID, annotation, regulation, log_2FC, and the sequence in FASTA format), are available in the Supplementary Table S5. Concerning the Jc171 profile, from the 4,612 identified DEGs (2,753 UR and 1,859 DR), 1,296 UR DEGs remained non-annotated. The correspondent details covering the Jc171 DEGs were shown in the Supplementary Table S6.

The two distinct *J. curcas* expressed profiles could contribute to identify useful salt-related tolerance genes for plant breeding programs. A previous study conducted by plant physiologists (UFPE, Brazil) showed differences between the same accessions analyzed here, highlighting a better recovery capacity of Jc183 after saline stress (750 mM NaCl, 50 h) than the Jc171 accession, inclusive showing an earlier emergence of the juvenile leaf (Real-Cortes, personal communication).

Metabolic responses of the J. curcas accessions to the salt stimulus

A comparative MapMan analysis performed with 2,749 DEGs of both accessions, based on the correspondent *M. esculenta* best hits (BLAST alignments), covered 15 Jc183 DEGs (eight UR and seven DR), 2,711 Jc171 DEGs (1,338 UR and 1,373 DR), and 23 DEGs shared by both profiles, all of them presenting the same regulation (18 UR, and five DR). The analyzed DEGs, comprising more than half of all detected DEGs, identified MapMan bin codes (Thimm et al., 2004) associated to several metabolism pathways (CHOs, lipids, amino acids, phytohormones, and redox; Supplementary Table S7), as well as with secondary metabolites. The heatmaps represented by the expression modulated by each accession associated with the

MapMan bin codes are presented in the Fig. 3. In general, the induced genes grouped and separated from the repressed genes; the induced genes exceeded the repressed ones, and the expression modulated by Jc171 in response to the salt-stimulus, was more intense than that of Jc183, since some of the corresponded genes not modulated its expression or was n.d. in the Jc183 profile (Fig. 3). Details covering each accession and MapMan category are discussed below.


A
Jc171
Jc183
DN43224.c0.g1/17.1.1
DN41789.c2.g1/17.1.1
DN31411.c1.g2/17.2.2
DN37295.c0.g1/17.4.2
DN8990.c0.g1/17.4.2
DN37716.c0.g2/17.1.1
DN43331.c1.g4/17.3.1.2.99
DN35005.c0.g1/17.2.3
DN39106.c0.g2/17.7.3
DN26789.c0.g2/17.2.2
DN8885.c0.g2/17.2.2
DN8885.c0.g1/17.2.2
DN4187.c0.g1/17.2.2
DN6442.c0.g1/17.4.2
DN6384.c0.g2/17.2.2
DN33300.c0.g1/17.4.2
DN34181.c0.g2/17.4.2
DN1084.c0.g1/17.5.3
DN14488.c0.g9/17.2.3
DN190.c0.g1/17.5.3
DN73239.c0.g1/17.1.1.1.10
DN43539.c0.g1/17.5.1
DN10691.c0.g2/17.7.1.4
DN74882.c0.g1/17.5.1
DN51201.c0.g1/17.5.1
DN35987.c1.g1/17.2.3
DN61847.c0.g1/17.3.1.2.5
DN62214.c0.g1/17.5.1
DN49147.c0.g1/17.2.3
DN85758.c0.g1/17.5.3
DN61662.c0.g1/17.2.3
DN1090.c0.g1/17.5.3
DN100462.c0.g1/17.6.3
DN62013.c0.g1/17.2.3
DN61290.c0.g1/17.5.3
DN61799.c0.g1/17.6.3
DN98407.c0.g1/17.1.3
DN74024.c0.g1/17.5.2
DN44830.c0.g1/17.5.1
DN97984.c0.g1/17.5.2
DN85308.c0.g1/17.5.3
DN31293.c0.g1/17.2.3
DN37072.c1.g1/17.5.2
DN97713.c0.g1/17.5.3
DN35987.c0.g1/17.2.3
DN37071.c0.g1/17.5.1
DN73287.c0.g1/17.1.3
DN99270.c0.g1/17.5.2
DN85740.c0.g1/17.5.2
DN62783.c0.g1/17.5.2
DN36882.c0.g1/17.5.2
DN35711.c1.g1/17.2.3
DN88501.c0.g1/17.5.2
DN22364.c0.g1/17.5.2
DN11306.c0.g1/17.5.2
DN52227.c0.g1/17.6.1
DN43242.c3.g1/17.5.2
DN87883.c0.g1/17.8.1
B
Jc171
Jc183
DN15616.c0.g1/2.1.1.1
DN28316.c0.g1/2.1.1.1
DN25301.c0.g1/2.1.2.1
DN23453.c0.g1/2.1.2.2
DN33305.c0.g2/2.1.2.3
DN39274.c0.g1/2.1.2.2
DN40754.c0.g1/2.2.2.3
DN39844.c0.g1/3.6
DN43461.c0.g1/2.2.2.3
DN34608.c0.g1/2.1.2.2
DN25903.c0.g2/3.6
DN42818.c0.g1/3.6
DN41446.c0.g2/3.6
DN1143.c0.g1/3.6
DN43630.c0.g1/3.6
DN42818.c7.g1/3.6
DN49538.c0.g1/3.6
DN35992.c2.g3/2.1.2.2
DN27969.c0.g1/3.1.1.2
DN73711.c0.g1/3.4.3
DN25250.c0.g1/2.2.2.1.2
DN19515.c0.g1/3.1.2.2
DN36518.c0.g1/3.5
DN9018.c0.g11/3.1.1.2
DN4418.c0.g1/3.5
DN50655.c0.g1/3.5
DN17657.c0.g1/3.8.1
DN9018.c0.g2/3.1.1.2
DN37859.c3.g1/3.1.1.2
DN12233.c0.g1/3.1.3
DN37121.c0.g1/3.5
DN49549.c0.g1/3.5
C
Jc171
Jc183
DN6888.c0.g2/13.1.3.4.12
DN41971.c0.g1/13.1.3.6.1.1
DN11899.c0.g1/13.1.6.5.1
DN39264.c1.g3/13.1.3.4
DN61616.c0.g1/13.1.5.3.1
DN39264.c1.g2/13.1.3.4
DN39264.c1.g1/13.1.3.4
DN41459.c0.g1/13.1.5.3.2
DN43999.c0.g1/13.1.3.4.12
DN49192.c0.g1/13.2.3.4.1
DN11370.c0.g1/13.1.3.4
DN98326.c0.g1/13.2.6.3
DN14481.c0.g1/13.2.7
DN97653.c0.g1/13.1.5.3.1
DN49983.c0.g1/13.2.6.3

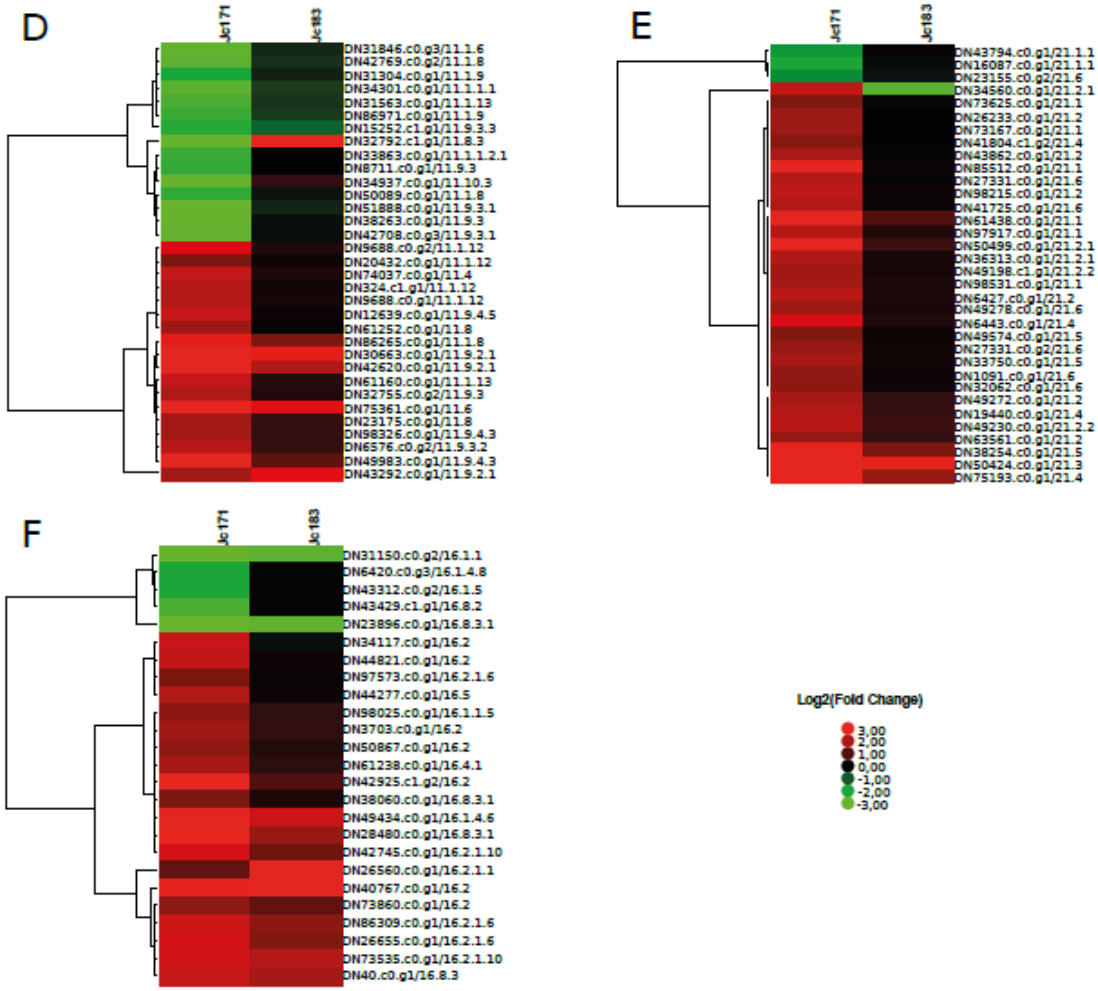


Figure 3. Heatmaps representing the gene expression profiles of Jc171 and Jc183 accessions, generated by hierarchical clustering analysis, and involving different metabolisms: A) hormone; B) CHO; C) Amino acid; D) Lipid; E) Redox; F) Secondary metabolites. The columns represent the expression modulated by the accessions after three hours of salt exposure (150 mM NaCl), in relation to the correspondent control without salt. The rows represent each *Jatropha curcas* RNA-Seq transcript and the annotated MapMan bin code. The clusters are on the left side. The up- and down-regulation of the transcripts are indicated in red and green, respectively, and the intensity of the colors increases with increasing expression differences based on the legend. The bin code description is provide in the Supplementary Table S7.

Phytohormone metabolism

The heatmap associated to the phytohormone metabolism (Fig. 3A) followed those general characteristics mentioned before: two groups; the induced group (Jc171) clustering more genes than the repressed one; expressed modulation by Jc171 more intense than those observed by Jc183.

Under normal conditions, the development of the root system architecture (RSA) of dicotyledonous plants, comprising the main root (MR) and the lateral roots (LRs), are under influence basically of auxin (AUX) and cytokinin (CK), two antagonistic phytohormones in some actions. However, under salt stress, the RSA development is influenced by abscisic acid (ABA), ethylene (ETH), jasmonic acid (JA), AUX, and

brassinosteroids (BR) (Julkowska and Testerink, 2015). Also, under osmotic stress conditions, ABA regulates root growth via an interacting hormonal network with cytokinin (CK), besides ETH and AUX (Rowe et al., 2016). According to Fig. 3A, some induced genes were associated to ABA (bin codes 17.1.1.1.10, 17.1.3, and 17.2.3), ETH (17.5.1, 17.5.2, and 17.5.3), AUX (17.2.3), JA (17.7.1.4, 17.7.3), BR (17.3.1.2.5), gibberellin (GA; 17.6.1, 17.6.3), and salicylic acid (SA; 17.8.1).

After a saline stimulus, the endogenous ABA level increase due to the action of dioxygenases (9-cis-epoxycarotenoid-dioxygenase; EC 1.13.11.51), which correspond to the bin code 17.1.1.1.10 (Fig. 3A), cleaving carotenoid precursors (Julkowska and Testerink, 2015). The ABA accumulation assists the plant acclimatization under stress, including stomatal closure, growth modulation, and synthesis of protective metabolites, some of them (e.g., proline, sugars, myo-inositol, polyamines) are osmoprotectant compounds that present important roles in ionic adjustment.

A JA action model in response to saline stress was presented by Riemann et al. (2015). Briefly, during osmotic stress, phytohormones are affected, some in a positive way [ABA, JA, and 12-OPDA (JA-precursor 12-oxo-phytodienoic acid)] and others, such as GA, in a repressed way. These phytohormones interfere with regulatory proteins that are crucial in metabolism. When JA is produced, the JAZ repressor is degraded, releasing MYC2, which is a versatile transcription factor (TF), also activated by ABA, for acting. In turn, with the reduced level of bioactive GA, DELLA proteins (which are GA repressors) accumulate and interact with JAZ, also releasing MYC2 from repression. MYC2, in turn, activates metabolic pathways, such as the secondary compounds for flavonoids and terpenoids (isoprenoids), some of them are also osmoprotectant-related compounds, helping plants to saline-stress acclimatization. Concerning the repressed genes (Jc171), the correspondent counterpart in Jc183 was non-modulated or n.d. In this group, CK transcripts stood out (bin code 17.4.2, Fig. 3A). CK is closely related to ABA metabolism, acting antagonistically in certain situations (Guan et al., 2014). Arabidopsis CK-deficient showed enhanced salt (250 mM NaCl) and drought tolerance, and the observed CK-downregulation was associated with cell membrane integrity and ABA hypersensitivity, rather than stomatal density and ABA-mediated stomatal closure (Nishiyama et al., 2011).

CHO metabolism

The heatmap representing major and minor CHO metabolism presented the two groups mentioned before, with the induced genes grouping apart from the repressed ones, but the induced genes numerically lower than the repressed ones (Fig. 3B). Some of the induced Jc183 genes modulated in a similar way of Jc171, while regarding the repressed Jc171 genes, the correspondent Jc183 almost not modulated its expression. Covering major CHO metabolism, the sucrose synthesis, represented by one of the main enzymes - Sucrose-phosphate synthase (bin code 2.1.1.1; DEG by Jc171; Fig. 3B), was repressed, reflecting the expected lower CO_2 fixation under saline stress. Still, in major CHO metabolism, the genes representing starch synthesis (2.1.2.1, 2.1.2.2; 2.1.2.3) was repressed (Fig. 3B), except by one induced *starch synthase* gene (both accessions). Starch is also an osmoprotective compound and thus protects macromolecules (membranes and proteins) from denaturing conditions (Singh et al., 2015). In turn, osmoprotectant compounds associated with the minor CHO metabolism [oligosaccharides of the raffinose family (3.1.1.2, 3.1.2.2; 3.1.3), and myo-inositol (3.4.3)], were induced by both accessions (Fig. 3B), probably helping to minimize salt-stress damages.

Amino acid metabolism

About amino acid metabolism (Fig. 3C), the heatmap followed the mentioned overall characteristics: induced and repressed genes (Jc171) grouped in independent clusters; the induced group clustering more genes; the Jc171 expression showing more expressive modulation than Jc183. From this set, the induced gene encoding methionine gamma-lyase (MGL; bin code 13.2.3.4.1, Fig. 3C) comprised one of the candidates selected to the RTq-PCR analysis. MGL converts methionine to 2-Ketobutyrate, a precursor in the Ile biosynthesis (Vijay and Jander, 2009), and studies covering drought-stress response pointed the amino acids Ile, Leu and Val increasing their abundances (Hildebrandt, 2018). Zhang et al. (2019) applying transcriptomic and metabolomic strategies to compare two contrasting sesame (*Sesamum indicum*) genotypes responding to salt stress (150 mM NaCl, different time points up to 24 h) observed many free amino acids accumulating higher in ST (salt-tolerant) than in SS (salt-sensitive) genotype, indicating this fact as a positive feature for withstanding salinity stress. The studied amino acids included: alanine, asparagine, aspartate, glutamate, glutamine, glycine, isoleucine, leucine, lysine, methionine, ornithine, phenylalanine, proline, serine,

threonine, tyrosine, and valine. Although substantial salt-induced accumulation was observed to arginine, glutamine, glycine, methionine, ornithine, phenylalanine, and tyrosine, most of the genes involved in the biosynthesis of these amino acids were up-regulated in both genotypes under salt stress.

Based on the heatmap (Fig. 3C), an induced expression also involved genes, such as *OASTL* (*cysteine synthase*, also *O-acetyl-serine* (*thiol*) *lyase*, bin code 13.1.5.3.1) and *SAT* (*serine acetyltransferase*, 13.1.5.3.2). The respective enzymes are involved with the cysteine biosynthesis. Transgenic plants overproducing SAT, OASTL or both enzymes present not only elevated levels of the respective products (cysteine and O-acetyl-serine, OAS) but also glutathione and other metabolites; in several cases, the transgenic plants were tolerant to the abiotic stresses (Sirko et al., 2004).

Lipids metabolism

About lipid metabolism, the heatmap followed most all of the general characteristic mentioned before: two groups (induced and repressed grouping apart), but in this case, almost the same amount of induced and repressed genes; more intense modulation by Jc171 when compared with Jc183 (Fig. 3D). A clear divergence in regulation (repression by Jc171 and induction by Jc183) involved the gene encoding SGT (UDP-glucose:sterol glucosyltransferase; DN32792_c1_g1, bin code 11.8.3, Fig. 1D). The SGT enzyme catalyzes the glycosylation of sterols to produce sterol glycosides. These glycosylated sterols play a crucial role in modulating the properties and function of cell membranes (Ramirez-Estrada et al., 2017). The mentioned authors correlated the expression of *SlSGT4* (tomato) with a marked increase in response to osmotic, saline and cold stress. Also, an induced *triacylglycerol lipase* gene (by both accessions) stood out (DN30663_c0_g1; bin code 11.9.2.1, Fig. 1D). In *A. thaliana*, a related gene (At2g31690) was among the most exclusively saline-stress induced in roots (Ma et al., 2006).

Redox metabolism

About the redox metabolism, the heatmap also presented the two primary groups, but almost all related genes were induced (Jc171), and again Jc171 modulated its expression more intensely than Jc183 (Fig. 1E). One divergence in gene regulation

(induced by Jc171 and repressed by Jc183) was observed (DN34560_co_g1, Fig. 1E) involving the bin code 21.2.2, which is related to ascorbate. Ascorbate and glutathione are potent antioxidants that react directly with singlet oxygen and superoxide radical (reactive oxygen species, ROS), or by detoxification of hydrogen peroxide (H_2O_2) (Foyer and Noctor, 2011). At the cellular level, the salt-stress alter the ionic homeostasis causing an imbalance in redox status in the cell, with subsequent high production of ROS, which is perceived by the antioxidant systems related to ascorbate and glutathione (bin code 21.2 in Fig. 1E), ferredoxin-thioredoxin reductase (21.1), superoxide dismutase (21.6), and ascorbate peroxidase (Foyer and Noctor 2009). In *Eutrema salsugineum*, the induction of ascorbate-glutathione in response to saline stress (300 mM NaCl) prevented the harmful production of singlet oxygen in the photosystem PSII (Wiciarz et al., 2017). In the roots of Arabidopsis under salt stress (150 mM NaCl), salt-induced changes in the cell redox status affected the meristem root, impacting auxin transport (Jiang et al., 2016). Besides the importance of the antioxidant systems and the ROS scavenging, details of the generated profiles need further studies.

Secondary metabolites process

Concerning the secondary metabolites process (Fig. 1F), the heatmap also presented the general characteristics mentioned before: two groups, separating the induced genes from the repressed ones (Jc171); the repressed genes comprising a smaller set; the Jc171 modulating its expression more intensively when compared to Jc183. Induced genes covered phenylpropanoids metabolism (bin code 16.2, Fig. 1F). In this group, the bin code 16.2.1.1, corresponding to the *PAL* gene (*phenylalanine ammonia lyase*) was selected to the RT-qPCR validation assay. The PAL enzyme is the entry point into the phenylpropanoid pathway, being a crucial enzyme that catalyzes the first step in that pathway and leads not only to the accumulation of phytoalexins but also contributes to the development of plants and their responses to biotic stresses (Zhang et al., 2013).

Gruber et al. (2009) reported global transcriptional changes in the secondary metabolism, including phenylpropanoid, flavonoid, and isoprenoid pathways, in Medicago root apexes responding to salt stress (100 mM NaCl, one hour after salt treatment). Also, it should be noticed that isoprenoids and phenylpropanoids are part of the antioxidant defense, and their representatives are orchestrated daily by drought-

stressed *Platanus acerifolia* plants during Mediterranean summers (Tattini et al., 2015). Additionally, salinity stress-induced genes involved in the flavonoid biosynthesis pathway in the salt-sensitive rice accession IR29 were observed but not in the salt-tolerant FL478 (Walia et al., 2005).

Selection of Jc183 DEGs and their expression validation by RT-qPCR

The hierarchical clustering of salt-tolerant Jc183 DEGs, and the corresponded expression by Jc171, allowed to select candidates to the RT-qPCR analysis (Fig. 4). The respective heatmap (Fig. 4) highlighted potential genes to be explored as a functional molecular marker for marker-assisted selection process in *J. curcas* breeding programs. From the cluster with UR Jc183 DEGs almost not modulated by Jc171, two candidates were selected [*S-adenosylmethionine-dependent methyltransferase* (*SAMe*), *Phenylalanine ammonia-lyase* (*PAL*). Another six candidates included UR Jc183 DEGs also induced by Jc171 [*S-adenosylmethionine synthase* (*SAM*), *peroxidase* (*PX*), *carboxylesterase* (*CXE*), *homeobox-leucine zipper protein* (*ATHB*), *TF of NAC family*, *methionine-gamma lyase* (*MGL*)]. Also, one DR Jc183 DEG (*xyloglucan endotransglucosylase*, *XTH*) was selected to the RT-qPCR validation assay.

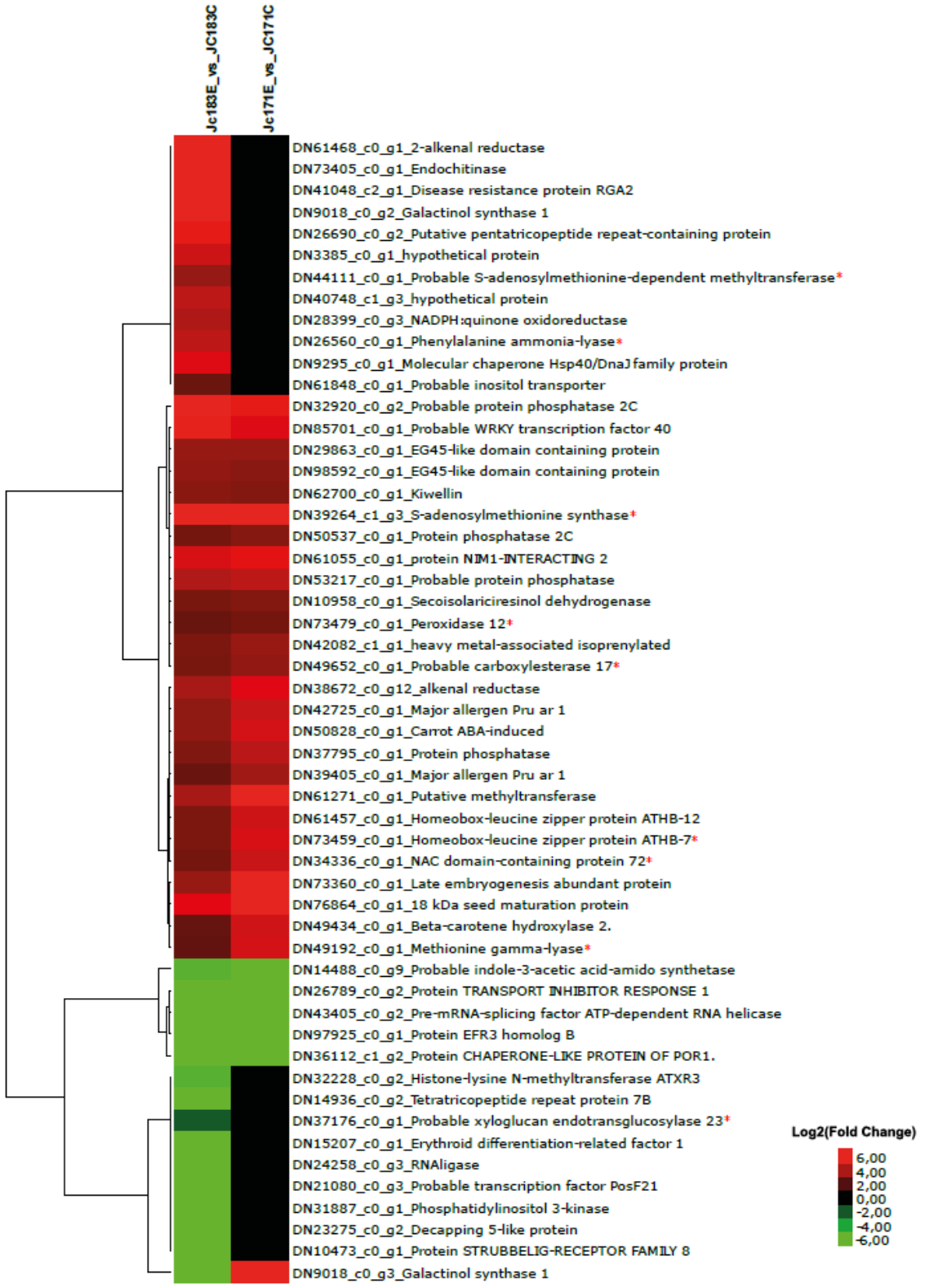


Figure 4. Heatmap representing the hierarchical clustering analysis of the 57 DEGs of the salt-tolerant Jc183, and the correspondent expressions by salt-sensitive Jc171 accession, after three hours of salt exposure (150 mM NaCl), based on the ratio of Log_2FC (Fold Change) values, concerning the abundances of the transcript in the stressed library, in relation to the negative control library. DEG: differentially expressed gene [*p-value* ≤ 0.0001; false discovery rate, FDR ≤ 0.005; Log_2 (FC) ≥ 1 (up-regulation, red) or ≤ -1 (down-regulation, green). On the right, the DEGs and its functional annotation. DEG with red asterisk were validated by RT-qPCR assay.

The *in silico* data to be validated by the RT-qPCR analysis, from each selected candidate gene, are shown in Table 1. The proposed primers pairs (candidates and references genes, Supplementary Table S2) amplified cDNA samples (except PX *primers* with Jc171 cDNAs), presenting a unique amplicon (data not shown). RT-qPCR parameters [amplification efficiency (E), *slope* (s), and correlation coefficient (R)], related to each *primer* pair and target, based on standard curves from serial dilution of root cDNAs samples (accessions and treatments), presented acceptable values (Supplementary Fig. S2), according the MIQE protocol (Bustin et al., 2009), aiming to ensure reliable relative qPCR expression data between samples. In general, the RT-qPCR results confirmed most of the *in silico* data, as described below.

Table 1. Selected genes and respective expressions by the *J. curcas* salt-tolerant (Jc183) and the salt-sensitive (Jc171) accessions, based on the *in silico* RNA-Seq analysis* and the RT-qPCR results**.

Method / **Accession**	**SAMe**	**PAL**	**SAM**	**PX**	**CXE**	**HD-Zip**	**NAC**	**MGL**	**XTH**
in silico									
Jc183	3.53/UR	4.44/UR	13.5/UR	2.49/UR	2.87/UR	2.95/UR	2.76/UR	2.36/UR	-2.08
Jc171	n.s	n.s	15.51/UR	2.76/UR	3.46/UR	5.09/UR	4.73/UR	4.99/UR	n.s
RT-qPCR									
Jc183	0.97/n.s	14.25/UR	12.30/UR	0.95/n.s	9.31/UR	5.03/UR	1.73/ n.s	15.71/UR	0.20/DR
Jc171	0.95/n.s	2.65/ UR	1.19/n.s	n.s	2.93/UR	18.69/UR	5.88/UR	8.16/UR	0.36/DR

UR: induced; DR: repressed; n.s: not significant at $p \leq 0.05$; SAMe: S-adenosylmethionine-dependent methyltransferase; PAL: Phenylalanine ammonia-lyase; SAM: S-adenosylmethionine synthase; PX: Peroxidase; CXE: Carboxylesterase; HD (Zip): Homeobox-leucine zipper; NAC: NAC transcript factor protein; MGL: Methionine-gamma lyase; XTH: Xyloglucan endotransglucosylase/hydrolase; * Log_2FC (FC: ratio of the abundances in the stressed library in relation to the respective control library);**Relative expression based on the REST software (v.2.0.13) (Pfaffl *et al*., 2002).

Carboxylesterase (CXE)

The induced DEG encoding CXE (EC 3.1.1.1), by both accessions (Table 1), confirmed its regulation in the RT-qPCR analysis (Fig. 5). CXEs are enzymes (α/β-hydrolase superfamily; Liu et al., 2014) that hydrolyze esters of short chain fatty acids. In plants, some of their biological functions are related to signal transduction and gene regulation (Lord et al., 2013). In plant signaling, CXEs activate phytohormones, such as SA and JA (Gershater and Edwards, 2007). The *ICME* (*isoprenylcysteine methylesterase*) gene, member of a small subfamily belonging to CXE family, encodes a protein involved in the *Arabidopsis* salt-response (200 mM NaCl), as a positive regulator of ABA signaling (Lan et al., 2010).

Homeobox-leucine zipper domain protein (HD-Zip)

The induced DEG (both accessions; Table 1) encoding ATHB-7 confirmed its UR regulation by RT-qPCR analysis (Fig. 5). The HD-Zip protein is a potential TF widely distributed in plants, playing roles in plant growth and response to abiotic stress (Shen et al., 2018). The overexpression of *ATHB-12* (*A. thaliana*) gene conferred salt tolerance (100 mM NaCl) in transgenic yeasts, regulating Na^+ exclusion and increasing NaCl tolerance (Shin et al., 2004). ATHB12 was involved in osmotic stress responses (Olsson et al., 2004). In transgenic tobacco, the expression of *CaHDZ12* (*Cicer arietinum*) transgene conferred salt stress tolerance (200 mM NaCl); in turn, the *CaHDZ12* silencing in chickpea plants led to a higher salt-sensitivity (Sen et al., 2017).

Methionine gamma-lyase (MGL)

The DEG encoding MGL (EC 4.4.1.11), induced by both accessions (Table 1), confirmed its expression by RT-qPCR analysis (Fig. 5). *Arabidopsis* plants responding to salt stress (100 mM NaCl), induced the *AtMGL* gene. MGL catalyzes the degradation of L-methionine to α-ketobutyrate, methanethiol, and ammonia. The α-Ketobutyrate is a precursor of isoleucine (Ile), an essential amino acid that also accumulates in plant cells under salt stress (Farhangi-Abriz and Ghassemi-Golezani, 2016).

Phenylalanine ammonia-lyase (PAL)

The induced Jc183 DEG encoding PAL (Table 1) confirmed its expression in the RT-qPCR assay (Fig. 5). PAL (EC 4.3.1.5) is a crucial enzyme in the phenylpropanoid pathway, catalyzing the deamination of L-phenylalanine (L-phe) to provide cinnamic acid, a precursor of secondary metabolites (Ibrahim et al., 2019). Phenolic acid, flavonoids, anthocyanins, lignins, and phytoalexins are derived from phenylpropanoids (Hsieh et al., 2010). Valifard et al. (2015) showed positive and significant correlations between the *PAL* gene induction, its enzymatic activity, and the phenolic content in *Salvia* species. The authors observed PAL activity increasing (42 - 45%) together to the phenolic content accumulation (35 - 43%) in the first six hours after the stress treatment (100 mM NaCl). Also, the induction of the *LjPAL* gene in *Lotus japonica* increased the PAL activity and the response to saline stress (150 mM NaCl) (Mrázová et al., 2017).

S-Adenosyl-L-methionine synthase (SAM)

The induced DEG encoding SAM (EC 2.5.1.6), by both accessions (Table 1), confirmed its expression by Jc183 (RT-qPCR; Fig. 5). SAM catalyzes the generation of

S-adenosylmethionine (Lindermayr et al., 2006), which is a Polyamine (PA) precursor. PAs, such as putrescine, spermidine, and spermine, are osmoprotectant compounds with relevant contributions on the osmotic adjustment of the cells under salt stress (Chen et al., 2018; Baniasadi et al., 2018). The positive correlation of PA accumulation with salt-tolerance of *A. thaliana* plants (Kasinathan and Wingler, 2004), probably reflected their roles on proteins and membranes stabilization, and the free radicals scavenging (Jang et al., 2012). Transgenic tobacco plants overexpressing *SAMS2* gene (from *Suadea salsa*) also presented increased PAs content, and salt-stress tolerance (200 mM NaCl) (Qi et al., 2010). A *SAM2* ortholog was downregulated in *J. curcas* plants after two h/100 mM NaCl but upregulated at seven days (Zhang et al., 2014). The authors highlighted the association of SAM with ETH biosynthesis from methionine.

NAC transcript factor protein

The induced DEG encoding the NAC TF (both accessions; Table 1) confirmed its regulation only by Jc171 (RT-qPCR, Fig. 5). The NAC superfamily (**N**AM, **A**FAT, and **C**UC) is one of the largest families of plant-specific TFs (Shao et al., 2015), playing crucial roles in abiotic stress responses, including salt, drought, and cold (Mao et al., 2014; Sakuraba et al., 2015). A *NAC TF* gene induced at two hours after the salt stress in *J. curcas* plants presented downregulation at seven days (Zhang et al., 2014). Transgenic rice plants overexpressing the *ONAC022* gene showed induction also two hours after salt-treatment (150 mM NaCl) (Hong et al., 2016). The ONAC022 acts as a transcriptional activator stress-responsive and plays a decisive role in drought and salt-stress tolerance modulating an ABA-mediated pathway. Positive regulation of NAC was associated with the synthesis and accumulation of proline, osmoprotective sugars, and LEA proteins (late embryogenesis abundant); all of these compounds play relevant roles in abiotic stress tolerance (Song et al., 2011).

Peroxidase (PX)

The induced DEG (both accessions, Table 1) encoding PX, not confirmed its expression by Jc183 in the RT-qPCR assay, and no amplicon was amplified with the proposed primers using Jc171 cDNAs (Fig. 5). PXs (EC 1.11.1.7) are oxidoreductase enzymes that catalyze the reduction of peroxides, such as hydrogen peroxide (H_2O_2), and the oxidation of organic and inorganic compounds (Chanwun et al., 2013). PX-ROS interactions are notable against harmful by-products of oxidative metabolism,

participating in cellular detoxification and free radical scavenging (Schaffer and Bronnikova, 2012).

S-adenosylmethionine-dependent methyltransferase (SAMe)

The gene encoding SAMe (Jc183 DEG, and n.s. by Jc171, Table 1) only confirmed the n.s. expression of Jc171, according to the RT-qPCR results (Fig. 5). In *J. curcas* plants under 150 mM NaCl, the *SAMe* expression showed the highest expression in root tissue (by semi-quantitative RT-PCR) after eight hours of salt exposure (Eswaran et al., 2012). SAMe is synthesized from ATP and methionine, a reaction catalyzed by methionine adenosyltransferase (SAM; Luka et al., 2009), also known as S-Adenosyl-L-methionine synthase (synonyms; https://www.brenda-enzymes.org/). SAM-binding methyltransferases utilize the methyl donor SAM as a cofactor to methylate proteins, small molecules, lipids, and nucleic acids (Martin and McMillan, 2002). SAMes play an essential role in cellular metabolism, transcription, signal transduction and detoxification (Hayashi et al., 2018). The overexpression of *SAMe* in *Hibiscus cannabinus* plants responding to NaCl (200 mM, six days) conferred salt-tolerance (Niu et al., 2016). The salt-induced *SAMe* gene (named *IbSIMT1*) in sweet potato (*Ipomoea batatas*) under saline stress (86 mM NaCl/ four weeks) showed induction in the first 12 hours of stress; the salt-tolerance was correlated with osmotic balance adjustment, membrane integrity and photosynthesis protection, and ROS detoxification (Liu et al. 2015). Probably, the three hours of salt-exposure time was not enough to show the *SAMe* induction by Jc183 accession.

Xyloglucan endotransglucosylase/hydrolase (XTH)

The gene encoding XTH was DR DEG by Jc183 (Table 1), and the DR expression was confirmed by RT-qPCR assay, also by Jc171 (Fig. 5). The xyloglucans (hemicellulosic polymers of dicotyledonous plants) bind to cellulose fibrils, whose interactions are modulated by expansin enzymes and XTHs (Malinowski et al., 2004). Thus, XTH (EC 2.4.1.207) is involved with wall remodeling and cell expansion. Under abiotic stress, XTH acts on the flexibility of tissues (leaf and root), conferring greater cell wall extensibility, and better plant adaptation (Tenhaken, 2015). In transgenic *Arabidopsis* plants responding to saline stress (100 mM NaCl), the *CaXTH* gene (*Capsicum annuum*) was induced after six days of salt treatment. Positive expression of

the *CaXTH* gene in pepper roots conferred a low reduction in root length of plants under salt stress (Cho et al., 2006). Depending on the abiotic stress, the cell wall is a target to be affected. In the present study, the wall remodeling and cell expansion in roots do not appear to be affected after three hours of J. *curcas* have been exposed to salt (150 mM NaCl), based on the *XTH* gene expression.

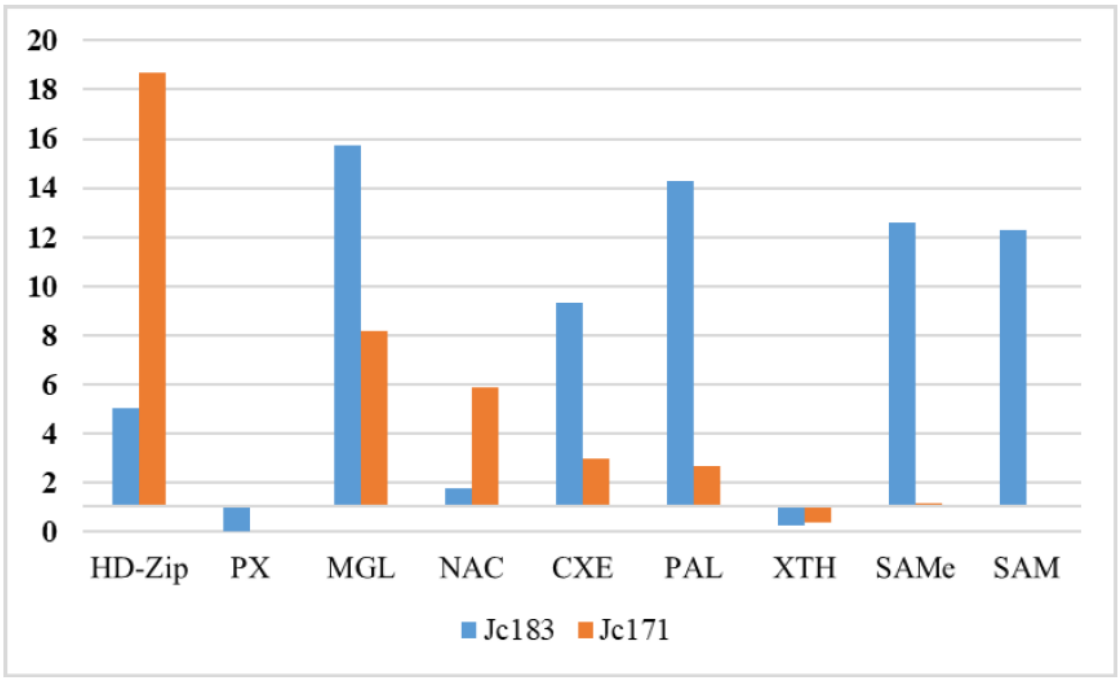


Figure 5. RT-qPCR results of selected Jc183 candidate genes, reference genes, and negative controls, performed with root cDNAs of *Jatropha curcas* Jc183 (salt-tolerant) and Jc171 (salt-sensitive) plants after three hours of salt-exposure. Expression values normalized by the reference genes *actin* and *beta-tubulin*, and the relative expression data calculated by the REST software (v.2.0.13) (Pfaffl *et al*., 2002). Genes: *HD-zip* (*homeobox-leucine*); *PX* (*Peroxidase*); *MGL* (*methionine-gamma-lyase*); *NAC Transcript factor*; *CXE* (*Carboxylesterase Probable 17*); *PAL* (*phenylalanine ammonia lyase*); *XTH* (*xyloglucan endotransglucosylase*); *SAMe* (*S-adenosylmethionine-dependent methyltransferase*); and *SAM* (*S-adenosylmethionine synthase 1*).

Conclusion

Sensitive plants in saline conditions use protection strategies, trying to stabilize photosystems, protect membranes and proteins, modulate the redox state, and provide detoxification of free radicals. The *de novo J. curcas* RNA-Seq transcriptome generated based on two accessions responding to salinity (150 mM NaCl, after three h), covered 101 MB and assembled 145,422 transcripts, around half of them encoding predicted proteins. Curiously, the salt-sensitive Jc171 accession induced more differentially expressed genes than the salt-tolerant Jc183. Based on Jc171 DEGs, the correspondent genes by Jc183, involving different metabolisms (phytohormone, CHO, lipid, amino acid, redox), and some secondary metabolites, presented their expressions almost not

modulated or even not detected. Based on the smaller number of Jc183 DEGs, the correspondent expressions by Jc171 was less similar. Although the intensive transcriptional effort of Jc171 inducing more DEGs in response to the salt stimulus, this effort was not enough to avoid the visible damages observed only in Jc171 leaves. In general, the upregulated genes numerically exceeded the downregulated ones, and despite some similar regulations presented by both accessions, the non-modulation by Jc183 suggest better control of the ionic homeostasis applying different salt-tolerance strategies from those observed in the Jc171 expressed profile. Some of the DEGs candidates involved in salt-stress tolerance with their expressions validated by RT-qPCR assays could be explored as functional molecular markers to be applied in marker-assisted selections, in *J. curcas* breeding programs. Thus, the present data not only uncovered genes related to salt-response but also promoted an overview of the molecular mechanisms underlying *J. curcas* salt-tolerance, providing potential functional molecular markers useful to the breeding programs. However, further studies are necessary to fully elucidate the molecular basis involving the development of plants under salt stress.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors contributions

LE, MFP, AMBI and EAK conceived and designed the experiments; MCPS, GALC, EB and MDS carried out the experiments; MCPS, GALC, MDS, and EAK analyzed the data; MCPS, and EAK wrote and revised the paper. All authors read and approved the final manuscript.

Acknowledgements

Authors are thankful to Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA SOJA, Londrina, PR - Brasil), and Universidade Federal de Alagoas (UFAL) for the facilities. This work was supported by Brazilian institutions (grants and fellowships): Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: 404357/2013-0, and 311894/2017-8).

Reference

Beltrão, N. D. M., Oliveira, M. I. P., 2008. Oleaginosas e seus óleos: vantagens e desvantagens para produção de biodiesel. Embrapa Algodão-Documentos (INFOTECA-E).

Bolger, A. M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120. https://doi:10.1093/bioinformatics/btu170

Bryant, D.M., Johnson, K., Di Tommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., Lee, T.J., Leigh, N.D., Kuo, T.H., Davis, F.G., Bateman, J., Bryant, S., Guzikowski, A.R., Tsai, S.L., Coyne, S., Ye, W.W., Freeman, R.M., Peshkin, L., Tabin, C.J., Regev, A., Haas, B.J., Whited, J.L., 2017. A Tissue-Mapped Axolotl De Novo Transcriptome Enables Identification of Limb Regeneration Factors. Cell Rep. 18, 762-776. htpps://doi:10.1016/j.celrep.2016.12.063

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Vandesompele, J., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical chemistry 55, 611-622. htpps://doi:10.1373/clinchem.2008.112797

Cartagena, J.A., Seki, M., Tanaka, M., Yamauchi, T., Sato, S., Hirakawa, H., Tsuge, T., 2015. Gene Expression Profiles in *Jatropha* Under Drought Stress and During Recovery. Plant Mol. Biol. Rep. 33, 1075-1087. https://10.1007/s11105-014-08150

Chanwun, T., Muhamad, N., Chirapongsatonkul, N., Churngchow, N., 2013. *Hevea brasiliensis* cell suspension peroxidase: purification, characterization and application for dye decolorization. AMB Express 3, 14. https://doi:10.1186/2191-0855-3-14

Chen, D., Shao, Q., Yin, L., Younis, A., Zheng, B., 2019. Polyamine Function in Plants: Metabolism, Regulation on Development, and Roles in Abiotic Stress Responses. Front. Plant Sci. 9, 1945. htpps://doi:10.3389/fpls.2018.01945

Cho, S., K., Kim, J.E., Park, Jong-A., P., Eom, T.J., Kim, W.T., 2006. Constitutive expression of abiotic stress-inducible hot pepper CaXTH3, which encodes a xyloglucan endotransglucosylase/hydrolase homolog, improves drought and salt tolerance in transgenic *Arabidopsis* plants. FEBS Letters 580, 3136-3144. https://doi10.1016/j.febslet.2006.04.062

Eswaran, N., Parameswaran, S., Anantharaman, B., Raja Krishna Kumar, G., Sathram, B., Johnson, T. S., 2012. Generation of an expressed sequence tag (EST) library from salt-stressed roots of *Jatropha curcas* for identification of abiotic stress-responsive genes. Plant Biol., 14, 428-437. https://doi:10.1111/j.1438-8677.2011.00529.x

Epstein, E., 1972. Mineral nutrition of plants: principles and perspectives.

FAO. Disponível em:<https: http://www.fao.org/docrep/003/T0234E/T0234E00.htm>. Acesso em 15 de Dez de 2018

Farhangi-Abriz, S., Ghassemi-Golezani, K., 2016. Improving amino acid composition of soybean under salt stress by salicylic acid and jasmonic acid. J. Appl. Bot. Food Qual. 89, 243-248. htpps://doi:10.5073/JABFQ.2016.089.031

Foyer, C.H., Noctor, G., 2009. Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. Antioxid. Redox Signaling 11, 861-905. https://doi.org/10.1089/ars.2008.2177

Foyer, C.H., Noctor, G., 2011. Ascorbate and glutathione: the heart of the redox hub. Plant physiol. 155, 2-18. htpps://doi:10.1080/02648725.2017.1340546

Gershater, M. C., Edwards, R., 2007. Regulating biological activity in plants with carboxylesterases. Plant Sci, 173(6), 579-588. https://doi.org/10.1016/j.plantsci.2007.08.008

Grabherr, M.G., Haas, B. J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, Z., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma. F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nature Biotechnol. https://doi:10.1038/nbt.1883

Grover, A., Kumari, M., Singh, S., Rathode, S.S., Gupta, S.M., Pandey, P., Gilotra, S., Kumar, D., Arif, M., Ahmed, Z., 2014. Analysis of *Jatropha curcas* transcriptome for oil enhancement and genic markers. Physiol Mol Biol Plants 20, 139-142. https://doi10.1007/s12298-013-0204-4

Gruber, V., Blanchet, S., Diet, A., Zahaf, O., Boualem, A., Kakar, K., Alunni, B., Udvardi, M., Frugier, F., Crespi, M., 2009. Identification of transcription factors involved in root apex responses to salt stress in *Medicago truncatula*. Mol. Genet. Genomics 281, 55-66. htpps://doi:10.1007/s00438-008-0392-8

Guan, C., Wang, X., Feng, J., Hong, S., Liang, Y., Ren, B., Zuo, J., 2014. Cytokinin antagonizes abscisic acid-mediated inhibition of cotyledon greening by promoting the degradation of abscisic acid insensitive5 protein in *Arabidopsis*. Plant Physiol. 3,1515-26. https://doi: 10.1104/pp.113.234740

Hayashi, T., Teruya, T., Chaleckis, R., Morigasaki, S., Yanagida, M., 2018. S-Adenosylmethionine Synthetase Is Required for Cell Growth, Maintenance of G0 Phase, and Termination of Quiescence in Fission Yeast. iScience 5, 38-51. https://doi.org/10.1016/j.isci.2018.06.011

Hildebrandt, T.M., 2018. Synthesis versus degradation: directions of amino acid metabolism during *Arabidopsis* abiotic stress response. Plant Mol. Biol. 98, 121-135. htpps://doi:10.1007/s11103-018-0767-0

Hong, Y., Zhang, H., Huang, L., Li, D., Song, F., 2016. Overexpression of a stress-responsive NAC transcription factor gene ONAC022 improves drought and salt tolerance in rice. Front. Plant Sci. 7, 4. htpps://doi:10.3389/fpls.2016.00004

Hsieh, L.S., Ma, G.J., Yang, C.C., Lee, P.D., 2010. Cloning, expression, site-directed mutagenesis and immunolocalization of phenylalanine ammonia-lyase in *Bambusa oldhamii*. Phytochemistry 71, 1999-2009. https://doi.org/10.1016/j.phytochem.2010.09.019

Ibrahim, W., Zhu, Y.M., Chen, Y., Qiu, C.W., Zhu, S., Wu, F., 2019. Genotypic differences in leaf secondary metabolism, plant hormones and yield under alone and

combined stress of drought and salinity in cotton genotypes. Physiol. Plant. 165, 343-355. https://doi.org/10.1111/ppl.12862

Jang, S.J., Wi, S.J., Choi, Y.J., An, G., Park, K.Y., 2012. Increased polyamine biosynthesis enhances stress tolerance by preventing the accumulation of reactive oxygen species: T-DNA mutational analysis of *Oryza sativa* lysine decarboxylase-like protein 1. Mol. cells 34, 251-262. htpps://doi:10.1007/s10059-012-0067-5

Jiang, K., Moe-Lange, J., Hennet, L., Feldman, L. J, 2016. Salt stress affects the redox status of *Arabidopsis* root meristems. . Front. Plant Sci. 7:81. doi:10.3389/fpls.2016.00081

Julkowska, M.M., Testerink, C., 2015. Tuning plant signaling and growth to survive salt. Trends Plant Sci. 9, 586-94. https://doi:10.1016/j.tplants.2015.06.008.

Juntawong, P., Sirikhachornkit, A., Pimjan, R., Sonthirod, C., Sangsrakru, D., Yoocha, T., Tangphatsornruang, S., Srinives, P., 2014. Elucidation of the molecular responses to waterlogging in *Jatropha* roots by transcriptome profiling. Front. Plant Sci. https://doi:10.3389/fpls.2014.00658

Kasinathan, V., Wingler, A., 2004. Effect of reduced arginine decarboxylase activity on salt tolerance and on polyamine formation during salt stress in *Arabidopsis thaliana*. Physiol. Plant. 121, 101-107. https://doi.org/10.1111/j.00319317.2004.00309.x

Lan, P., Li, W., Wang, H., Ma, W., 2010. Characterization, sub-cellular localization and expression profiling of the isoprenylcysteine methylesterase gene family in *Arabidopsis thaliana*. BMC plant boil. 10, 212. https://doi.org/10.1186/1471-2229-10-212

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. https://doi.org/10.1186/gb-2009-10-3-r25

Li, B., Dewey, C., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323. https://doi.org/10.1186/1471-2105-12-323

Lindermayr, C., Saalbach, G., Bahnweg, G., Durner, J., 2006. Differential inhibition of *Arabidopsis* methionine adenosyltransferases by protein S-nitrosylation. J. Biol. Chem. 281, 4285-4291. htpps://doi:10.1074/jbc.M511635200

Lord, C.C., Thomas, G., Brown, J.M., 2013. Mammalian alpha beta hydrolase domain (ABHD) proteins: Lipid metabolizing enzymes at the interface of cell signaling and energy metabolism. Biochimica et Biophysica Acta (BBA) – Mol. Cell Biol. L. 1831, 792-802. https://doi.org/10.1016/j.bbalip.2013.01.002

Liu, D., He, S., Song, X., Zhai, H., Liu, N., Zhang, D., Liu, Q., 2015. IbSIMT1, a novel salt-induced methyltransferase gene from *Ipomoea batatas*, is involved in salt tolerance. Plant Cell, Tissue Organ Cult. 120, 701-715. htpps:/doi:10.1007/s11240-014-0638-6

Liu, D., Wang, L., Zhai, H., Song, X., He, S., Liu, Q., 2014. A novel α/β-hydrolase gene IbMas enhances salt tolerance in transgenic sweetpotato. PloS one, 9, e115128. https://doi:10.1371/journal.pone.0115128

Lozano-Isla, F., Campos, M.L.O., Endres, L., Bezerra-Neto, E., Pompelli, M.F., 2018. Effects of seed storage time and salt stress on the germination of *Jatropha curcas* L. Ind. Crops Prod. 118, 214-224. https://doi.org/10.1016/j.indcrop.2018.03.052

Luka, Z., Mudd, S.H., Wagner, C., 2009. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. J. Biol. Chem. 284, 22507-22511 htpps://doi:10.1074/jbc.R109.019273

Ma, S., Gong, Q., Bohnert, H. J., 2006. Dissecting salt stress pathways. J. Exp. Bot. 57, 1097-107. htpps://doi:10.1093/jxb/erj098

Malinowski, R., Filipecki, M., Tagashira, N., Wiśniewska, A., Gaj, P., Plader, W., Malepszy, S., 2004. Xyloglucan endotransglucosylase/hydrolase genes in cucumber (*Cucumis sativus*) – differential expression during somatic embryogenesis. Physiol. Plant. 120, 678-685. htpps://doi:10.1111/j.0031-9317.2004.0289.x

Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., Li, J., Lam-Son, P.T., Qin, F., 2015. A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. Nat. Commun. 6, 8326. https://doi:10.1038/ncomms9326

Martin, J. L., McMillan, F. M., 2002. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. Curr. Opin. Struct. Biol. 12, 783-793. https://doi.org/10.1016/S0959-440X(02)00391-3

Mrázová, A., Belay, S. Assefa, Eliášová, A., Perez-Delgado, C., Kaducová, M., Betti, M., Vega, J.M., Paľove-Balang, P., 2017. Expression, activity of phenylalanine-ammonia-lyase and accumulation of phenolic compounds in *Lotus japonicus* under salt stress. Biologia 72, 36-42. https://doi:10.1515/biolog-2017-0001

Munns, R., 2005. Genes and salt tolerance: bringing them together. New Phytol. 167, 645-663. https://doi:10.1111/j.1469-8137.2005.01487.x

Nishiyama, R., Watanabe, Y., Fujita, Y., Le, D.T., Kojima, M., Werner, T., Vankova, R.,Yamaguchi-Shinozaki, K., Shinozaki, K., Kakimoto, T., Sakakibara, H., Schmülling, T., Tran, L. S., 2011. Analysis of cytokinin mutants and regulation of cytokinin metabolic genes reveals important regulatory roles of cytokinins in drought, salt and abscisic acid responses, and abscisic acid biosynthesis. Plant Cell 23, 2169-83. htpps://doi:10.1105/tpc.111.087395

Niu, X., Xu, J., Chen, T., Tao, A., Qi, J., 2016. Proteomic changes in kenaf (*Hibiscus cannabinus* L.) leaves under salt stress. Ind. Crops Prod. 91, 255-263. https://doi.org/10.1016/j.indcrop.2016.07.034

Oliveros, J.C., 2015. Venny. An interactive tool for comparing lists with Venn's diagrams. https://bioinfogp.cnb.csic.es/tools/venny/index.html

Olsson, A. S., Engström, P., Söderman, E., 2004. The homeobox genes ATHB12 and ATHB7 encode potential regulators of growth in response to water deficit in *Arabidopsis*. Plant Mol. Biol. 55, 663–677.

Peng, Z., He, S., Gong, W., Sun, J., Pan, Z., Xu, F., Lu, Y., Du, X., 2014. Comprehensive analysis of differentially expressed genes and transcriptional regulation

induced by salt stress in two contrasting cotton genotypes. BMC genomics, 15, 760. https://doi.org/10.1186/1471-2164-15-760

Pramanik, K., 2003. Properties and use of *Jatropha curcas* oil and diesel fuel blends in compression ignition engine. Renew. Energy. 28, 239-248. htpps://doi.org/10.1016/S0960-1481(02)00027-7

Pfaffl, M.W., Horgan, G.W., Dempfle, L., 2002. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic acids Res. 30, e36-e36.

Baniasadi, F., Saffari, V.R., Moud, A.A.M., 2018. Physiological and growth responses of *Calendula officinalis* L. plants to the interaction effects of polyamines and salt stress. Sci. hortic. 234, 312-317.

Qi, Y.C., Wang, F.F., Zhang, H., Liu, W.Q., 2010. Overexpression of *Suadea salsa* S-adenosylmethionine synthetase gene promotes salt tolerance in transgenic tobacco. Acta Physiol. Plant. 32, 263-269. htpps://doi:10.1007/s11738-009-0403-3

Ramirez-Estrada, K., Castillo, N., Lara, J.A., Arró, M., Boronat, A., Ferrer, A., Altabella, T. 2017. Tomato UDP-Glucose Sterol Glycosyltransferases: A Family of Developmental and Stress Regulated Genes that Encode Cytosolic and Membrane-Associated Forms of the Enzyme. Front. Plant Sci. https://doi:10.3389/fpls.2017.00984

Riemann, M., Dhakarey, R., Hazman, M., Miro, B., Kohli, A., Nick, P., 2015. Exploring Jasmonates in the Hormonal Network of Drought and Salinity Responses. Front. Plant Sci. 6, 1077. htpps://doi:10.3389/fpls.2015.01077.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140. https://doi:10.1093/bioinformatics/btp616

Rowe, J.H., Topping, J.F, Liu, J., Lindsey, K., 2016. Abscisic acid regulates root growth under osmotic stress conditions via an interacting hormonal network with cytokinin, ethylene and auxin. New Phytol. 1, 225-39. https://doi:10.1111/nph.13882

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132, 365–386. https://doi.org/10.1385/1-59259-192-2:365

Sakuraba, Y., Kim, Y.S., Han, S.H., Lee, B.D., Paek, N.C., 2015. The Arabidopsis transcription factor NAC016 promotes drought stress responses by repressing AREB1 transcription through a trifurcate feed-forward regulatory loop involving NAP. The Plant Cell 27, 1771-1787. https://doi:10.1105/tpc.15.00222

Saldanha, A.J., 2004. Java Treeview—extensible visualization of microarray data. Bioinformatics, 20, 3246-3248. https://doi.org/10.1093/bioinformatics/bth349

Sirko, A., Błaszczyk, A., Liszewska, F., 2004. Overproduction of SAT and/or OASTL in transgenic plants: a survey of effects. J. Exp. Bot. 55, 1881-1888. https://doi.org/10.1093/jxb/erh151

Sapeta, H., Lourenço, T., Lorez, S., Grumaz, C., Kirstahler, P., Barros, P.M., Costa, J.M., Sohn, K., Oliveira, M.M., 2016. Transcriptomics and physiological analyses

reveal co-ordinated alteration of metabolic pathways in *Jatropha curcas* drought tolerance. J. Exp. Bot. 67, 845–860, htpps://doi:10.1093/jxb/erv499

Sen, S., Chakraborty, J., Ghosh, P., Basu, D., Das, S. 2017. Chickpea WRKY70 regulates the expression of a Homeodomain-Leucine Zipper (HD-Zip) I transcription factor CaHDZ12, which confers abiotic stress tolerance in transgenic tobacco and chickpea. Plant Cell Physiol. 58, 1934-1952. htpps://doi:10.1093/pcp/pcx126 /

Shao, H., Wang, H., Tang, X., 2015. NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. Front. Plant Sci. 6, 902. htpps://doi:10.3389/fpls.2015.00902

Schaffer, W.M., Bronnikova, T.V., 2012. Peroxidase-ROS interactions. Nonlinear Dyn. 68, 413. https://doi.org/10.1007/s11071-011-0314-x

Shen, W., Li, H., Teng, R., Wang, Y., Wang, W., Zhuang, J., 2018. Genomic and transcriptomic analyses of HD-Zip family transcription factors and their responses to abiotic stress in tea plant (*Camellia sinensis*). Genomics. htpps://doi:10.1016/j.ygeno.2018.07.009

Shin, D., Koo, Y.D., Lee, J., Lee, H.J., Baek, D., Lee, S., Yun, D.J., 2004. Athb-12, a homeobox-leucine zipper domain protein from Arabidopsis thaliana, increases salt tolerance in yeast by regulating sodium exclusion. Biochem. Biophys. Res. Commun., 323, 534-540. https://doi.org/10.1016/j.bbrc.2004.08.127

Singh, A., Jha, S. K., Bagri, J., Pandey, G. K., 2015. ABA inducible rice protein phosphatase 2C confers ABA insensitivity and abiotic stress tolerance in *Arabidopsis*. PloS one 10, e0125168. https://doi.org/10.1371/journal.pone.0125168

Song, S.Y., Chen, Y., Chen, J., Dai, X.Y., Zhang, W.H., 2011. Physiological mechanisms underlying OsNAC5-dependent tolerance of rice plants to abiotic stress. Planta, 234, 331-345.

Taiz, L., Zeiger, E., Moller, I.M., Murphy, A., 2017. Fisiologia e desenvolvimento vegetal, six ed. Porto Alegre, Artmed.

Tang, Y., Qin, S., Guo, Y., Chen, Y., Wu, P., Chen, Y., Li, M., Jiang, H., Wu, G., 2016. Genome-Wide Analysis of the AP2/ERF Gene Family in Physic Nut and Overexpression of the JcERF011 Gene in Rice Increased Its Sensitivity to Salinity Stress. PLoSONE 11, e0150879. https//doi:10.1371/journal.pone.0150879

Tattini, M., Loreto, F., Fini, A., Guidi, L., Brunetti, C., Velikova, V., Gori, A., Ferrini, F., 2015. Isoprenoids and phenylpropanoids are part of the antioxidant defense orchestrated daily by drought-stressed *Platanus acerifolia* plants during Mediterranean summers. New Phytol. 207, 613-626. htpps://doi:10.1111/nph.13380

Tenhaken, R., 2015. Cell wall remodeling under abiotic stress. Front. Plant Sci. 5, 771. https://doi.org/10.3389/fpls.2014.00771

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., Selbig, J., Muller, L.A., Rhee, S.Y., Stitt, M., 2004. MAPMAN: a user – driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37, 914-939. https://doi:10.1111/j.1365-313X.2004.02016.x

Thornthwaite, C.W., Mather, J.R., 1955. The Water Balance. Drexel Institute ofTechnology [Philadelphia] Laboratory of Climatology, Centerton, N.J.Publications in Climatology.

Valifard, M., Mohsenzadeh, S., Niazi, A., Moghadam, A. Phenylalanine ammonia lyase isolation and functional analysis of phenylpropanoid pathway under salinity stress in *Salvia* species. Aust. J. Crop Sci. 9, 656-665.

Vijay, J., Georg, J. 2009. *Arabidopsis* Methionine γ-Lyase Is Regulated According to Isoleucine Biosynthesis Needs but Plays a Subordinate Role to Threonine Daeminase. Plant Physiol. 151, 367-378. htpps://doi:10.1104/pp.109.138651

Walia, H., Wilson, C., Condamine, P., Liu, X., Ismail, A.M., Zeng, L., Wanamaker, S. I., Mandal, J., Xu, J., Cui, X., Close, T. J., 2005. Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. Plant Physiol. 139, 822-35. https://doi.org/10.1104/pp.105.065961

Walia, H., Wilson, C., Zeng, L., Ismail, A. M., Condamine, P., Close, T.J., 2007. Genome-wide transcriptional analysis of salinity stressed japonica and indica rice genotypes during panicle initiation stage. Plant mol. biol. 63, 609-623. http://dx.doi.org/10.1007/s11103-006-9112-0

Wang, H., Gong, M., Xin, H., Tang, L., Dai, D., Gao, Y., Liu, C., 2018. Effects of chilling stress on the accumulation of soluble sugars and their key enzymes in *Jatropha curcas* seedlings. Physiol. Mol. Biol. Plants 24, 857-865. https://doi.org/10.1007/s12298-018-0568-6

Wang, H., Zou, Z., Wang, S., Gong, M., 2013. Global Analysis of Transcriptome Responses and Gene Expression Profiles to Cold Stress of *Jatropha curcas* L. PLoS ONE 8, e82817. https://doi:10.1371/journal.pone.0082817

Wang, H., Zou, Z., Wang, S., Gong, M., 2014. Deep sequencing-based transcriptome analysis of the oil-bearing plant Physic Nut (*Jatropha curcas* L.) under cold stress. Plant Omics 7, 178-187.

Wang, J., Meng, Y., Li, B., Ma, X., Lai, Y., Si, E., Yang, K., Xu, X., Shang, X., Wang, H., Wang, D., 2015. Physiological and proteomic analyses of salt stress response in the halophyte *Halogeton glomeratus*. Plant Cell Environ. 38, 655-669. https://doi:10.1111/pce.12428

Wiciarz, M., Niewiadomska, E., Kruk, J., 2018. Effects of salt stress on low molecular antioxidants and redox state of plastoquinone and P700 in *Arabidopsis thaliana* (glycophyte) and *Eutrema salsugineum* (halophyte). Photosynthetica, 56, 811-819. https://doi.org/10.1007/s11099-017-0733-0

Xu, G., Huang, J., Yang, Y., Yao, Y. A., 2016. Transcriptome analysis of flower sex differentiation in *Jatropha curcas* L. using RNA sequencing. PloS one, 11(2), e0145613. https://doi.org/10.1371/journal.pone.0145613

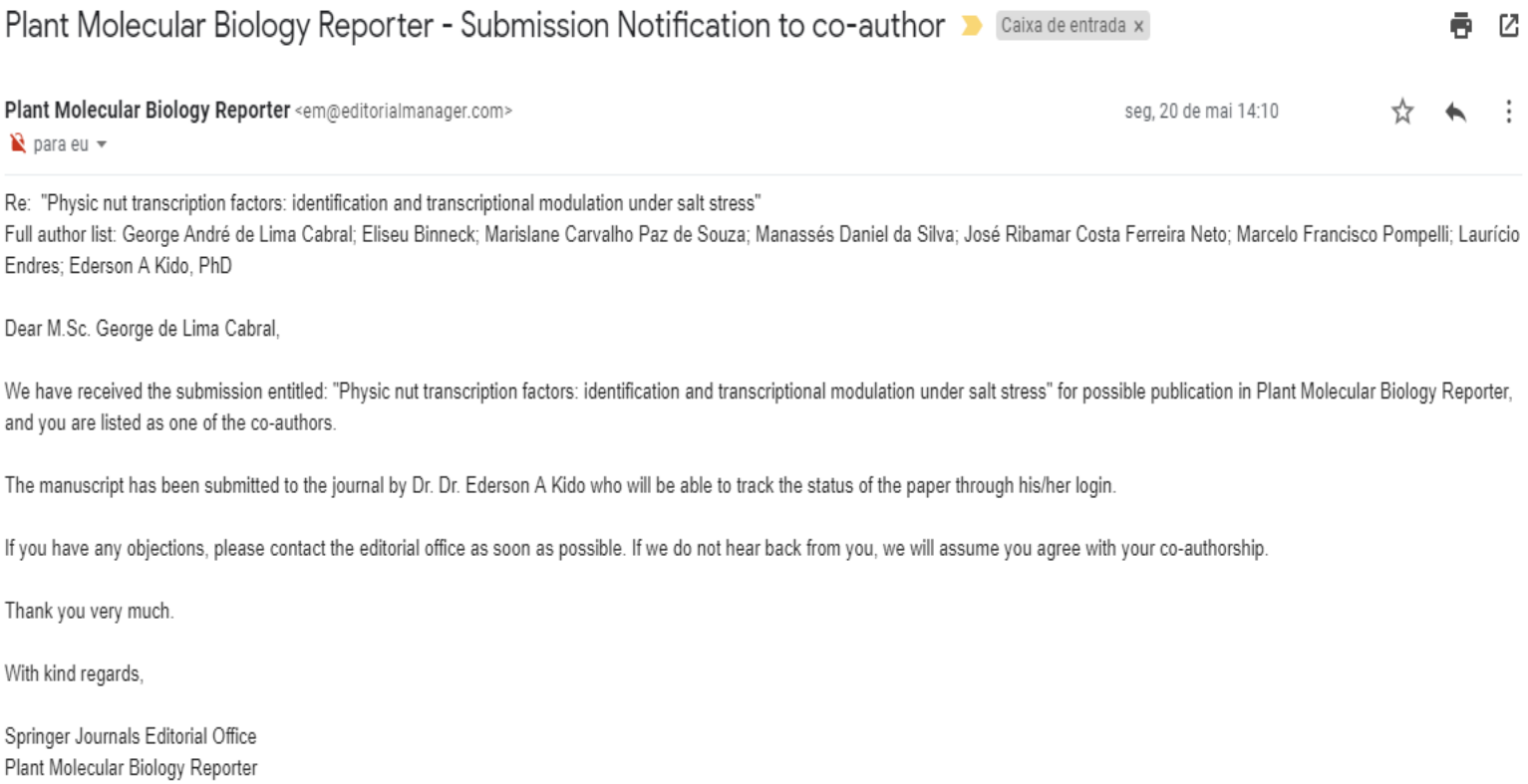
Zhang, C., Zhang, L., Zhang, S., Zhu, S., Wu, P., Chen, Y., Li, M., Jiang, H., Wu, G., 2015. Global analysis of gene expression profiles in physic nut (*Jatropha curcas* L.) seedlings exposed to drought stress. BMC Plant Biol. 15, 17. htpps://doi10.1186/s12870-014-0397-x

Zhang, L., Zhang, C., Wu, P., Chen, Y., Li, M., Jiang, H., Wu, G., 2014. Global Analysis of Gene Expression Profiles in Physic Nut (*Jatropha curcas* L.) Seedlings Exposed to Salt Stress. PLoS ONE 9(5): e97878. htpps://doi:10.1371/journal.pone.0097878

Zhang, X., Gou, M., Liu, C-J., 2013. *Arabidopsis* Kelch Repeat F-Box Proteins Regulate Phenylpropanoid Biosynthesis via Controlling the Turnover of Phenylalanine Ammonia-Lyase. Plant Cell. Plant Cell 25, 4994–5010. htpps://doi:10.1105/tpc.113.119644

Zhang, Y., Li, D., Zhou, R., Wang, X., Dossa, K., Wang, L., Zhang, Y., Yu, J., Gong, H., Zhang, X., You, J., 2019. Transcriptome and metabolome analyses of two contrasting sesame genotypes reveal the crucial biological pathways involved in rapid adaptive response to salt stress. BMC Plant Biol. 19, 66. https://doi.org/10.1186/s12870-019-1665-6

4.2 Physic nut transcription factors: identification and transcriptional modulation under salt stress



Plant Molecular Biology Reporter - Submission Notification to co-author Caixa de entrada ×

Plant Molecular Biology Reporter <em@editorialmanager.com> seg, 20 de mai 14:10

para eu

Re: "Physic nut transcription factors: identification and transcriptional modulation under salt stress"
Full author list: George André de Lima Cabral; Eliseu Binneck; Marislane Carvalho Paz de Souza; Manassés Daniel da Silva; José Ribamar Costa Ferreira Neto; Marcelo Francisco Pompelli; Laurício Endres; Ederson A Kido, PhD

Dear M.Sc. George de Lima Cabral,

We have received the submission entitled: "Physic nut transcription factors: identification and transcriptional modulation under salt stress" for possible publication in Plant Molecular Biology Reporter, and you are listed as one of the co-authors.

The manuscript has been submitted to the journal by Dr. Dr. Ederson A Kido who will be able to track the status of the paper through his/her login.

If you have any objections, please contact the editorial office as soon as possible. If we do not hear back from you, we will assume you agree with your co-authorship.

Thank you very much.

With kind regards,

Springer Journals Editorial Office
Plant Molecular Biology Reporter

George André de Lima Cabral[1], Eliseu Binneck[2], Marislane Carvalho Paz de Souza[1], Manassés Daniel da Silva[1], José Ribamar Costa Ferreira Neto[1], Marcelo Francisco Pompelli[3], Laurício Endres[4], Éderson Akio Kido[1]

[1]Federal University of Pernambuco (UFPE), Bioscience Center, Department of Genetics, 50670-420 Recife, PE, Brazil
[2]Brazilian Agricultural Research Corporation (EMBRAPA), EMBRAPA-Soybean, 86001970 - Londrina, PR, Brazil
[3]Federal University of Pernambuco (UFPE), Bioscience Center, Department of Botany, 50670-901 - Recife, PE, Brazil
[4]Federal University of Alagoas (UFAL), Agricultural Sciences Center, 57072-970 - Maceió, AL, Brazil
* Corresponding author (e-mail address: kido.ufpe@gmail.com)

Email addresses:
GALC: geolimma@gmail.com
EB: eliseubinneck@gmail.com
MCPS: maris.carvalho@hotmail.com
MDS: manassesdaniel@gmail.com
JRCFN: netocostaferreira@gmail.com
MFP: mfpompelli@gmail.com
LE: lauricioendres@hotmail.com
EAK: kido.ufpe@gmail.com

Abstract:

Physic nut (*Jatropha curcas*), a small oleaginous tree spontaneously occurring in arid and semi-arid tropical regions, is a sustainable and renewable energy source for biodiesel. However, the *J. curcas* yield in such areas should consider soil salinity and its consequences. Transcription factor (TF) proteins recognize *cis*-regulatory elements in promoters of genes regulating their expression. In the present work, differentially expressed genes (DEGs) encoding putative TFs in physic nut plants after three hours of NaCl (150 mM) exposition covered 23 TF families. The expressed profiles of members from AP2/ERF and NAC families basically presented induction after the salt treatment, while members of bHLH, FHY3-FAR1, and ARF families were repressed. The gene ontology (GO) enrichment analysis concerning the induced TF DEGs highlighted terms related to abiotic stress responses, while the repressed TF DEGs stood out terms highlighting the basal metabolism. In turn, the TF enrichment analysis predicted TFs over-represented targeting promoters of induced TF DEGs. The enriched TFs are good candidate as transgenes. Additionally, RT-qPCR analyses validated the up-regulation of six DEGs (*RAV1*, *ERF9*, *ZAT12*, *PTI5*, *MYB340*, and *BZIP4*) of eight candidates, suggesting the reliability of the expressed *J. curcas* TFoma after three hours of salt exposition (150 mM NaCl). The results help to understand the molecular basis of salinity stress response in physic nut plants. Also, provide valuable resources to select potential candidates for transgenic studies, as well as to develop functional molecular markers to assist selection steps in breeding programs.

Key words: *Jatropha curcas*; RNA-Seq; transcriptome; abiotic stress; salinity.

Introduction

Physic nut (*Jatropha curcas* L.), a perennial tropical plant belonging to the Euphorbiaceae family, has been an alternative source of renewable biofuel producer (Openshaw 2000). *J. curcas* presents some advantages over so-called oleaginous biodiesel plant producers, including the high (30–50%) oil content in the seeds allied to the relatively easily biofuel conversion (Deore and Johnson 2008), and no competition with human food destination. Since it is a spontaneous species occurring in arid and semi-arid tropical regions (Johnson et al. 2011), two problems need to be addressed: soil salinity and plant salt stress. Crop production in arid and semi-arid regions must consider natural saline/sodic soils, high plant evapotranspiration, low rainfall, and unfavorable

physical/physicochemical soil properties, which associated to irrigation problems leads to soil salinization, and its consequence in plant growth (Campos et al. 2012).

In plants growing in regions with low water availability, the high salt levels in soil solution reduce the osmotic potential in the root zone sufficiently to reduce water absorption (Dasgan et al. 2002), also the ions acting on protoplasm disturb the mineral plant nutrition (Munns 2002), limiting the plant growth (Gurgel et al. 2003).

Plants exposed to environmental stresses change their metabolisms according to the genes properly activated or repressed (Benko-Iseppon et al. 2005). Based on the set of transcription factor (TFs) activated by signal transduction in response to a stimulus, genes are expressed and the transcriptomes are reprogrammed. The TF proteins recognize the *cis*-regulatory elements (CRE) in the promoters of genes that will be expressed regulating that expression (Wang et al. 2009). In this way, TFs play a key role in biotic and abiotic stress responses, as well as in plant development (Riechmann et al. 2000), through the spatial and temporal regulation based on their targets (Zhang et al. 2011; Jin et al. 2014). Therefore, to characterize the TFoma after three hours of NaCl exposition (150 mM) helps to understand the transcriptional dynamics of *J. curcas* plants responding to the salt and also to improve the salt tolerance.

Materials and Methods

a) Plant material and the salinity assay

Two Brazilian physic nut accessions named Jc183 e Jc171 (Lozano-Isla et al. 2018) were carried out in a salt treatment assay with plants growing in greenhouse (March 2016) at the Agricultural Science Center/Federal University of Alagoas (UFAL/CECA, Rio Largo, AL, Brazil; geodesic coordinates 09º28'02"S; 35º49'43"W, altitude: 127 m). The classified climate, according to Thorthwaite and Mather (1955), is wet, megathermic, with moderate water deficiency in the summer (December to March) and some excess of water in the winter (July to September).

Homogeneous seeds (size and weight) of both accessions were sown in pots (50 L) filled with 20 kg of washed sand. From the first eophiles (5-10 days after germination, DAG), the seedlings were thinned, leaving only the most vigorous plant per pot. The plants were sampled in a completely randomized design with three biological replicates (two accessions x two treatments (with and without salt) x three biological replicates). During their cultivation, plants were irrigated (4 p.m.) every three

days with Hoagland nutrient solution (20% w/v) (Epstein, 1972). A week before the salt application (60 DAG), plants received Hoagland solution 100% (full strength) every day. To the salt application, a NaCl solution (150 mM) was added to the Hoagland solution, and plants were salt exposed (9 a.m.) for three hours. Plants irrigated only with the Hoagland solution comprised the negative control. After the NaCl exposure time, root samples were collected, immediately frozen in liquid nitrogen (N_2), being kept in -80ºC until RNA extraction.

b) RNA isolation, the RNA-Seq libraries, and its sequencing

Total RNA was isolated from the root samples using the SV Total RNA Isolation System (Promega). The RNA concentration was estimated by NanoDrop spectrophotometer (Thermo Scientific NanoDrop 2000), and the RNA quality assessed by absorbance ratios (OD 260/280 nm ≥ 1.9 and OD 260/230 nm ≥ 1.9), and agarose gel electrophoresis, 1.5% (w/v). Re-analysis of the RNAs integrities by the Agilent Bioanalyzer 2100 system (Santa Clara, CA, USA) identified samples with RIN (*RNA Integrated Number*) ≥ 9.0. High-quality RNAs were used to generate RNA-Seq libraries (12: two accessions x three biological replicates x two treatment), at the Genomic Center of the "Luiz de Queiroz" College of Agriculture (ESALQ/USP, Piracicaba, SP, Brazil). The RNA-Seq libraries were sequenced (2x100 bp *paired-end*) using the Illumina HiSeq2500 Platform (Eurofins MGW, Germany).

c) Transcriptome assembly and the transcript annotation

The quality data of the reads (base sequence quality and content) generated from the RNA-Seq *paired-end* libraries were visualized with the FastQC software (v.0.11.5), before and after adapter filtering and trimming (*paired-end*) steps using default parameters of the Trimmomatic tool (v.0.36; Bolger et al. 2014). After excluding reads showing low quality and those with unknown adapters and nucleotides, pairs of high-quality reads (*Phred* ≥ 30, all bases) were used for *de novo* transcriptome assembly performed with the Trinity 2.2.0 software (Grabherr et al. 2011). The expression levels of assembled transcripts and UniGenes were estimated by RSEM software (Li e Dewey 2011), and the alignment package Bowtie (v4.4.7; Langmead et al. 2009) was applied to map reads back to UniGenes. The normalized FPKM (*Fragment Per Kilobase of cDNA Per Million fragments mapped*) matrices were generated from the RSEM counts, which

were used for the differential expression analyses performed by edgeR package (Robinson et al. 2010).

Potential transcripts encoding TFs were identified by BLASTx alignments (*e-value* ≤ e^{-10}) against proteins sets downloaded (August 2018) from the databases: NCBI (*J. curcas*; https://www.ncbi.nlm.nih.gov/), Phytozome v.12 (*Ricinus communis* and *Manihot esculenta*; https://phytozome.jgi.doe.gov/pz/portal.html), and UniProtKB/SwissProt (http://www.uniprot.org/).

d) Identification of the differential expressed genes (DEGs)

The gene expression analysis between experimental samples detected differentially expressed genes (DEGs) as those UniGenes showing *p-value* ≤ 0.0001, FDR (*False Discovery Rate*) ≤ 0.005, and $Log_2FC \geq 1$ (classified as up-regulated, UR) or ≤ -1 (down-regulated, DR). Fold change (FC), based on Log_2FC values, was the ratio representing the modulation of the UniGene abundance in the stressed library compared to the negative control (henceforth, S *vs.* C, for brevity). The modulation of the gene expression data, after hierarchical clustering analysis performed by Cluster software (v.3.0; https://cluster2.software.informer.com/3.0/), generated heatmaps, visualized with the JavaTreeview software (v.1.1; http://jtreeview.sourceforge.net). The Venn diagrams were generated by online tool Venny (http://bioinfogp.cnb.csic.es/tools/venny/).

e) The GO and TF enrichment analyses

Together with the Gene Ontology analysis, a functional GO terms enrichment analysis identified those over-represented (Fisher's exact tests, *p-value* ≤ 0.01) based on the input file applied to the PlantRegMap tool (Plant Transcriptional Regulatory Map; http://plantregmap.cbi.pku.edu.cn; Jin et al. 2014). The input file corresponded to the set of genes encoding TFs, and individually covered the UR DEGs, the DR DEGs, and the non-DEGs (n.s.). A similar procedure was performed involving the TF enrichment analysis applying the respective tool also provided by the same database.

f) The gene expression validation by RT-qPCR assay

The gene expression of DEGs candidates encoding TFs [*RAP2*-3 (*ethylene-responsive transcription factor RAP2-3*), *RAV1*(*AP2/ERF and B3 domain-containing transcription factor RAV1*), *ERF9* (*ethylene-responsive transcription factor* 9),

DREB1H (*dehydration-responsive element-binding protein 1H*), *ZAT12* (*Zinc finger protein ZAT12*), *PTI5* (*pathogenesis-related genes transcriptional activator PTI5*), *MYB340* (*Myb-related protein 340*), and *BZIP4* (*basic leucine zipper 4*); Table S1] were analyzed in RT-qPCR assays. *Primer* pair were designed based on the correspondent RNA-Seq transcript using the *online* Primer 3 tool (Rosen and Skaletsky 2000), following some parameters: amplicon size (between 70 and 200 bp), melting temperature [50ºC (minimum), 70ºC (optimum) and 80ºC (maximum)], and GC content (45 - 55%). Proposed primers (**Supplementary Table S1**) were synthesized by Invitrogen Life Technologies (USA) and previously tested amplifying cDNAs in conventional PCR. After that, RT-qPCR reactions were performed in a real-time LineGene 9600 equipment (Bioer®, Hangzhou, China) using SYBR Green detection system. The PCR reaction (10 µL) included 5 µL of SYBR Green SuperMix (Applied Biosystems, Foster City CA, EUA), 1 µL of diluted cDNA (1/10), 0.3 µL of each *primer* (5 µM) and 3.4 µL ddH_2O. The reactions followed the settings: initial denaturation of 95°C for 2 min, followed by 40 cycles of 95°C for 15 s, and 60°C for 60 s. All RT-qPCR reactions were performed in 96-well plates with three biological, and three technical replicates, the negative controls, and two reference genes adequately tested for the present assays [*β-tubulin* and *actin* (Ma et al. 2016)]. The dissociation curves were obtained, heating the amplicons from 65 to 95°C for 20 min after the RT-qPCR cycles. The LineGene software (v.1.1.10) estimated the Cq values (quantification cycles), and the absolute and relative quantifications. The relative expression data evaluated by REST 2009 software (Relative Expression Software Tool v.2.0.13; Pfaffl et al. 2002) applied randomization test with 2,000 permutations and considered the hypothesis of significant differences between the control and treatment groups. The MIQE (*The Minimum Information for Publication of Quantitative Real-Time PCR Experiments*; Bustin et al. 2009) protocol was followed to assure data reliabilities.

Results

a) The *J. curcas de novo* transcriptome and the DEGs encoding TFs

The high-throughput sequencing of the *J. curcas* RNA-Seq libraries (12) of roots exposed to NaCl (150 mM, three hours) generated 238,286,823 raw *reads*. After removing adapters and trimming low quality bases, 230,140,599 high-quality *reads* (*Phred* ≥ 30, all bases; 96.58% of the *reads*) allowed the *de novo* transcriptome

assembly, with 145,422 transcripts (101 Mb) and 126,342 UniGenes (76 Mb), showing a GC% of 41.55, and the N_{50} comprising 1,308 bp for transcripts, and 993 pb for UniGenes. The global transcriptome will be not addressed in the presented report, only those transcripts encoding potential TFs.

Based on the BLASTx analysis (*e-value* e^{-10}) of the transcripts against protein databases from Euphorbiaceae species (Figure 1), 1,876 transcripts encoding TFs were identified. The inclusion of *R. communis* and *M. esculenta* (Phytozome database), both *J. curcas*-related species, besides increased the sequence similarities (**Figure 1**), also increased the efficiency of the annotation process.

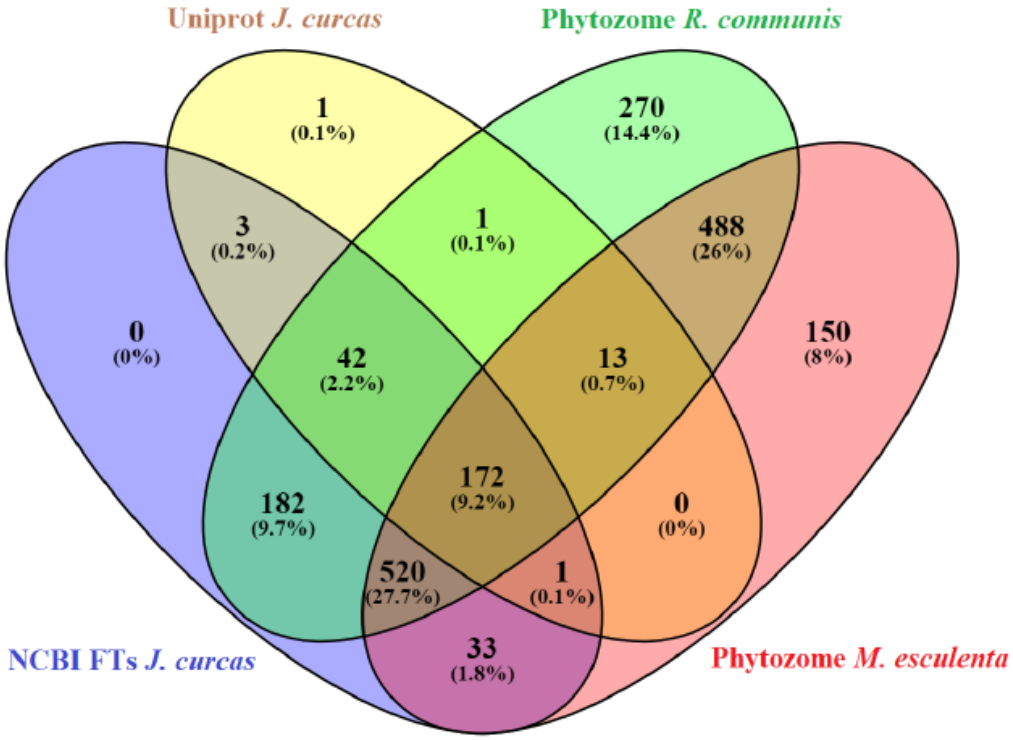


Figure 1. Venn diagram showing numbers of *Jatropha curcas* RNA-Seq transcripts (from roots of plants after three hours of NaCl exposition; 150 mM) encoding transcript factor proteins similar (*e-value* $\leq e^{-10}$) to those from different public proteins databases (NCBI, https://www.ncbi.nlm.nih.gov/.; Phytozome, https://phytozome.jgi.doe.gov/pz/portal.html; UniProt, https://www.uniprot.org/).

The declared DEGs from each accession [*p-value* $\leq$ 0.0001, FDR $\leq$ 0.005, $Log_2FC \geq 1$ (UR) or ≤ -1(DR)] in response to the salt-treatment was quite different (4,646 from Jc171, and 57 from Jc183), highlighting the great effort of the Jc171 trying to minimize visible damages only observed in Jc171 leaves (**Figure 2**). Based on that, the present investigation of TFs differentially expressed in response to the salt-treatment was restricted to the Jc171expressed profile.

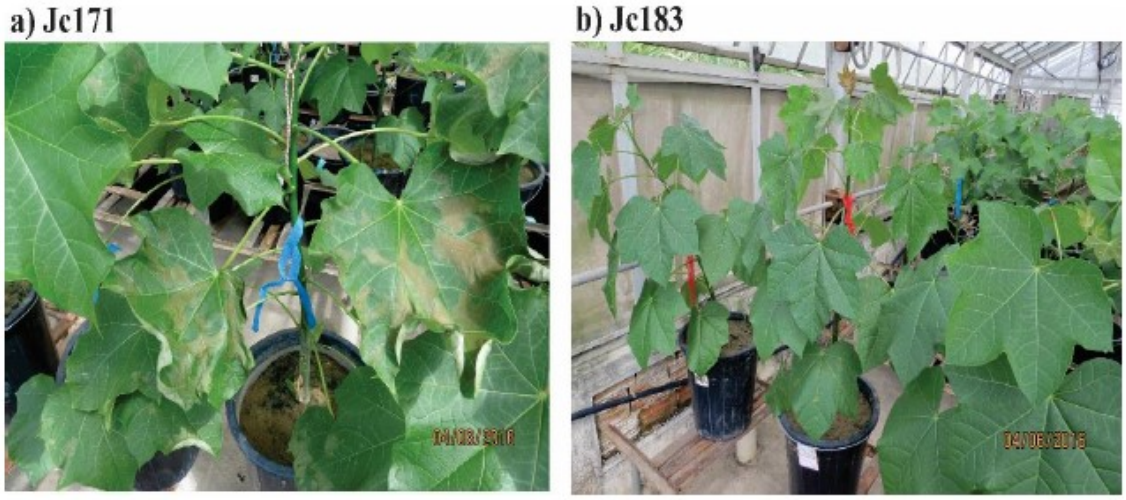


Figure 2. Visual damages on leaves of *Jatropha curcas* Jc171 accession, after three hours of NaCl exposition (150 mM), and not observed in the Jc183 accession.

From the Jc171 DEGs (4,646), 148 of them encoding TFs (78 UR and 70 DR; Table S4) encompassed 23 TF families (**Figure 3**).

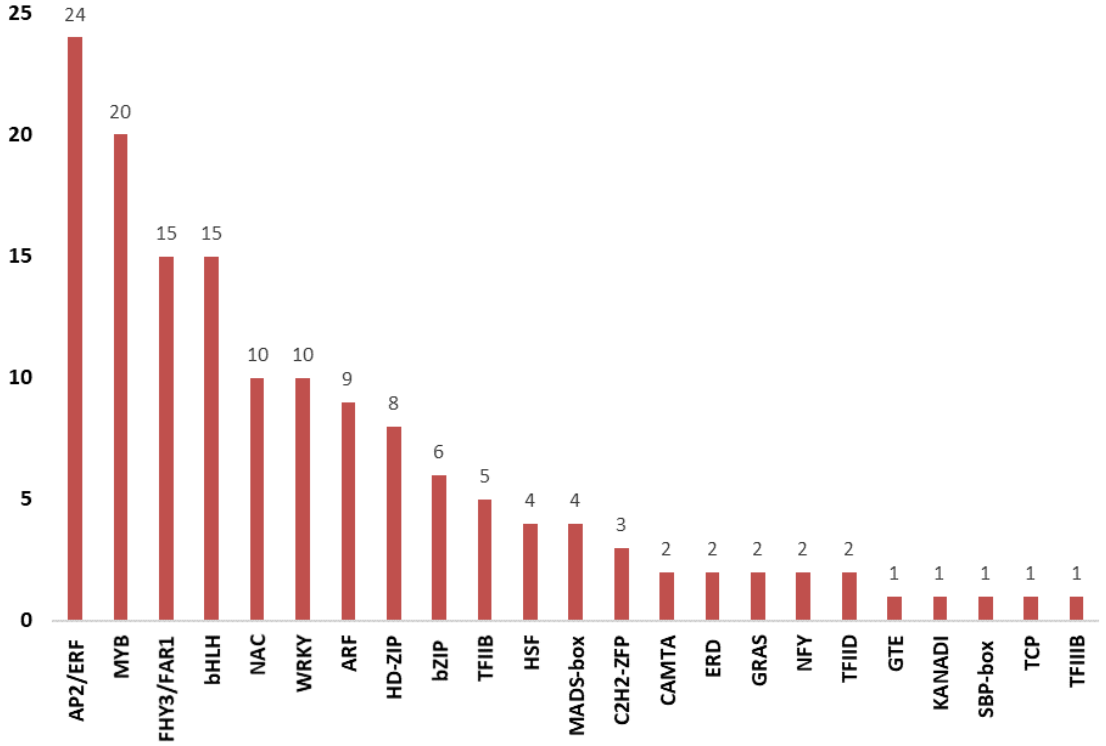


Figure 3. Transcription factor (TF) families associated with differentially expressed genes [DEG: *p-value* ≤ 0.0001, FDR ≤ 0.005, $Log_2FC \geq 1$ or ≤ -1] from *Jatropha curcas* Jc171accession after salt-treatment (150 mM NaCl). The TF family name is followed by the total of DEGs and the correspondent percentage.

The TF families outstanding in isoforms members were: AP2/ERF, MYB, bHLH, FHY3/FAR1, WRKY, NAC, ARF, HD-zip, and bZIP (**Figure 3**). Concerning a single TF family, the proportion of induced DEGs to the repressed DEGs was variable (**Figure 4**).

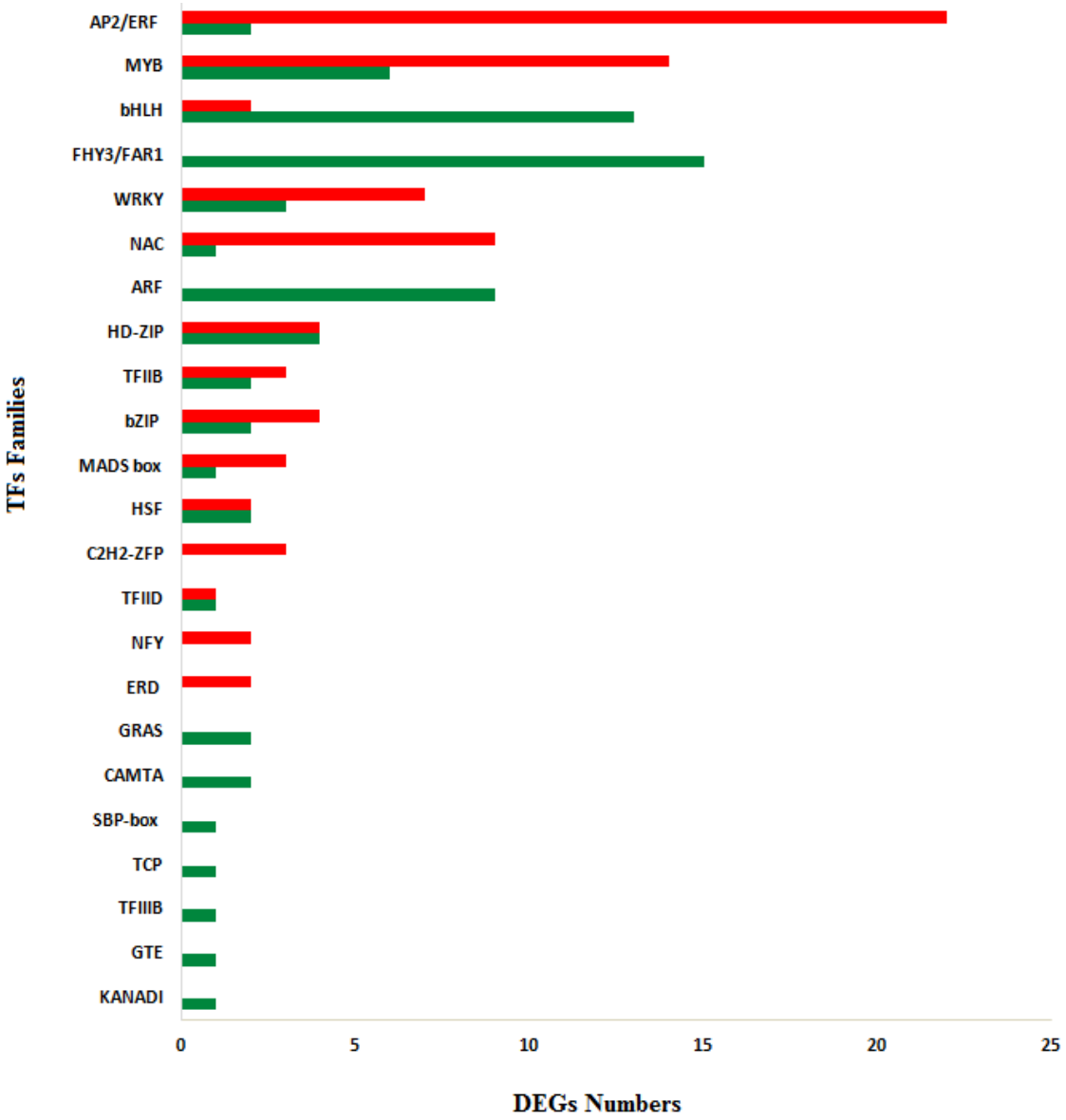


Figure 4. Families of transcription factors showing differentially expressed genes [DEG: *p-value* ≤ 0.0001, FDR ≤ 0.005, $Log_2FC \geq 1$ or ≤ -1] from Jc171accession after salt-treatment (150 mM NaCl): the induced DEGs are represented by red bars, and the repressed DEGs by the green bars.

TF families presenting more induced DEGs were AP2/ERF (22), MYB (14), NAC (9) and WRKY (7), while families showing more repressed DEGs were

FHY3/FAR1 (15), bHLH (13), ARF (9) and MYB (6) (**Figure 4**). Also, some TF families only presented UR or DR isoforms members (**Figure 4**).

c) The GO enrichment analysis

From the 78 induced DEGs encoding TFs, 71 identified by the PlantRegMap tool in the input gene list presented 128 enriched GO terms (*p-value* ≤ 0.01). From the 57 of 70 repressed DEGs (TFs) in the input gene list, the enriched GO terms were 143. When the input gene list involved 958 TFs codified by non-DEGs, the enriched GO terms were 435. All the enriched terms were distributed into the three main GO categories [Biological process (BP), Molecular function (MF), and Cellular Component (CC); **Supplementary Table S2**].

Comparing the three sets of enriched GO terms in a Venn diagram, six terms highlighted the UR DEGs codifying TFs (**Figure 5A**): *metabolic process* (GO:0008152), *death* (GO:0016265), *heterocyclic compound binding* (GO:1901363), *response to wounding* (GO:0009611), *cell death* (GO:0008219), and *organic cyclic compound binding* (GO:0097159). Eight enriched terms outstood the repressed DEGs (**Figure 5A**): *single-organism cellular process* (GO:0044763), *vegetative phase change* (GO:0010050), *cell proliferation* (GO:0008283), *single-organism process* (GO:0044699), *negative regulation of growth* (GO:0045926), *regulation of cell proliferation* (GO:0042127), *cell fate commitment* (GO:0045165), and *regulation of circadian rhythm* (GO:0042752).

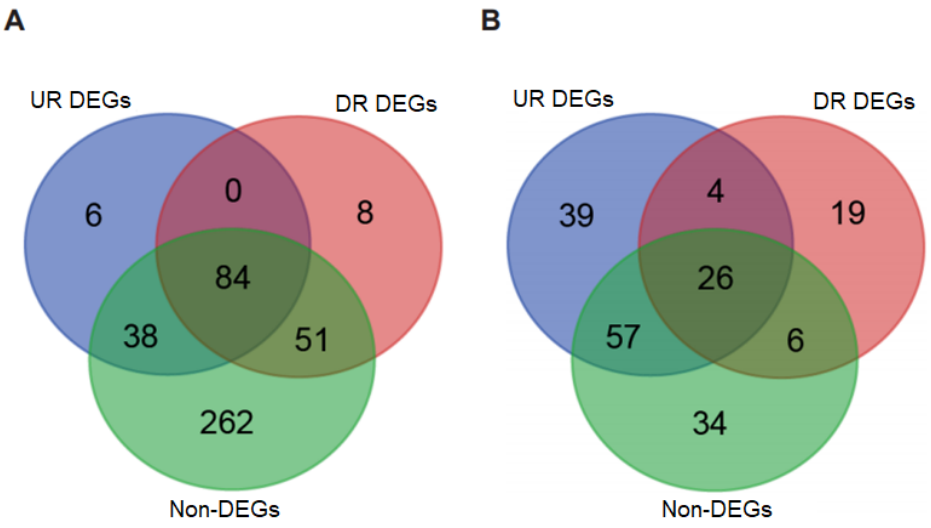


Figure 5. Venn diagram comparing the enriched GO terms (A) or the enriched TFs (B) identified by the respective PlantRegMap tool using individually different input gene list: the UR DEGs, the DR DEGs or the non-DEGs. DEG: differentially expressed gene (thresholds: *p-value* ≤ 0.0001, FDR ≤ 0.005, $Log_2FC \geq 1$ (UR, up-regulated) or ≤ -1 (DR, down-regulated).

d) TF enrichment analysis

Defining 72 UR DEGs codifying TFs, as the target genes in the input list, the TF enrichment tool predicted 1,164 regulations by 245 TFs; those enriched TFs counted 126. When DR DEGs comprised the input gene list, 58 target genes predicted 185 TFs and 729 regulations; from the predicted TFs, 55 were considered enriched. When 1,000 TFs non-DEGs (n.s.) comprehended the input gene set, 11,936 predicted regulations involved 302 TFs, being 123 the enriched TFs. All the enriched TFs, their TF families, and number of predicted targets genes presented in the input gene list (UR, DR or non-DEGs) are presented in the **Supplementary Table S3**. Comparing the three sets of enriched TFs in a Venn diagram, those enriched TFs interacting exclusively with promoters of UR DEGs (as their targets) were 39, exclusively with the DR DEGs were another 19 TFs, and those with the non-DEGs were another 34 TFs (**Figure 5B**).

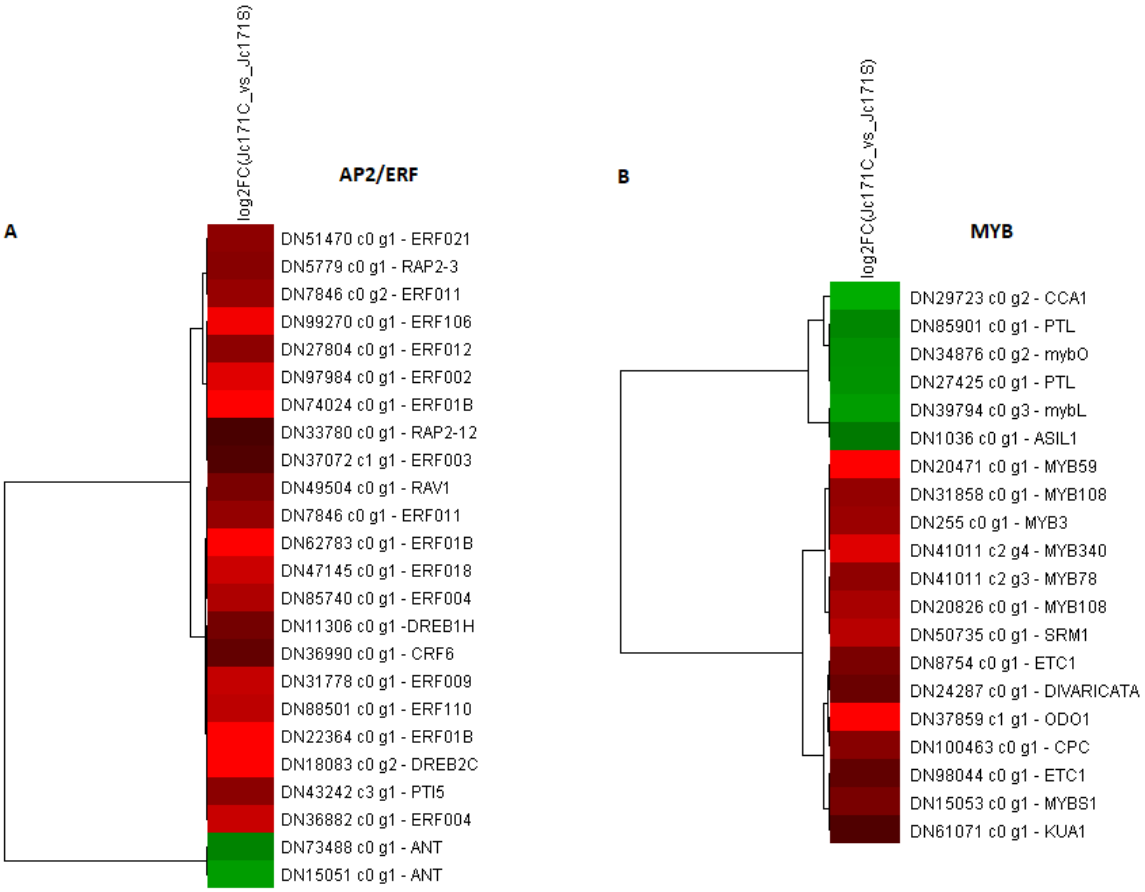
ERF family members (22) comprised most of the enriched TFs predicted interacting with the UR DEGs; also, ERF family (28) stood out with the non-DEGs; while Dof family members (9) interacted most with the DR DEGs. The distribution of the TF families by enriched FTs predicted targeting promoters of TF genes comprising the UR DEGs, the DR DEGs, or the non-DEGs, all codifying putative TFs expressed in roots of *J. curcas* after the salt exposition, is presented in the **Supplementary Table S4**, and the number or possible target genes in the **Supplementary Table S3**.

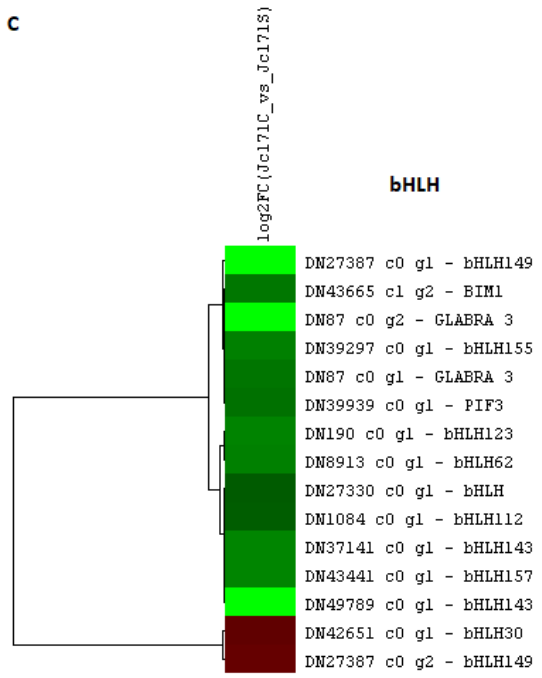
e) Expressed profiles of the most representative TF families associated to the salt response

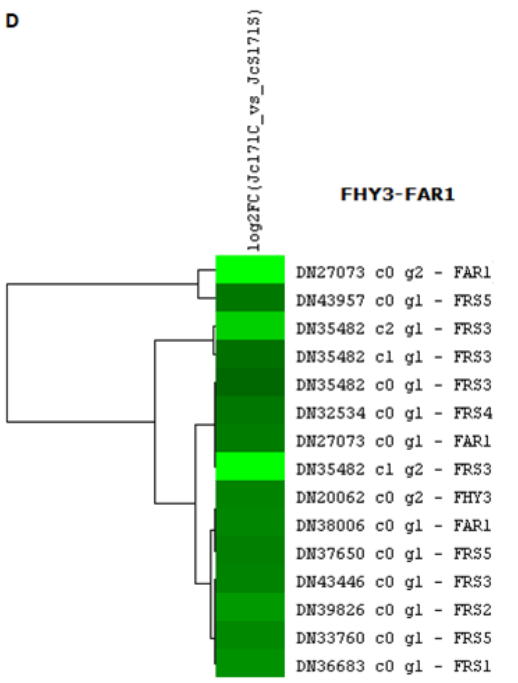
The heatmaps with the expressed profiles of members of the most representative TF families are shown in **Figure 6**, and the respective expression data modulated, based on the ratio of Log_2FC values expressed by Jc171 after the salt-treatment in relation to the negative control is provided in the **Supplementary Table S5**, together with the sequence fasta format.

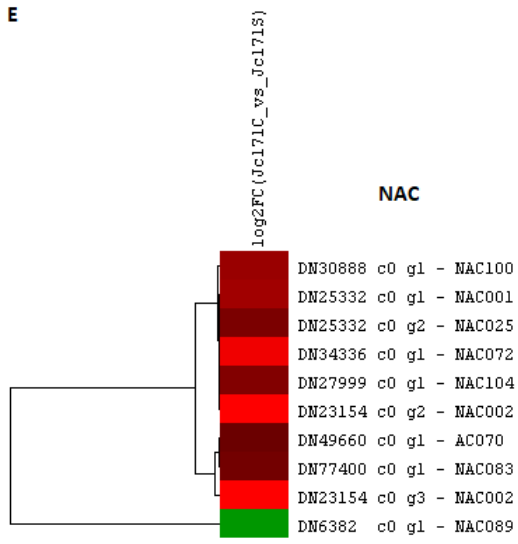
Almost all members of the AP2/ERF family were up-regulated after the salt stimulus (**Figure 6A**). A similar profile was observed with members of the NAC family (**Figure 6E**). Members of MYB (**Figure 6B**), WRKY (**Figure 6F**), and bZIP (**Figure**

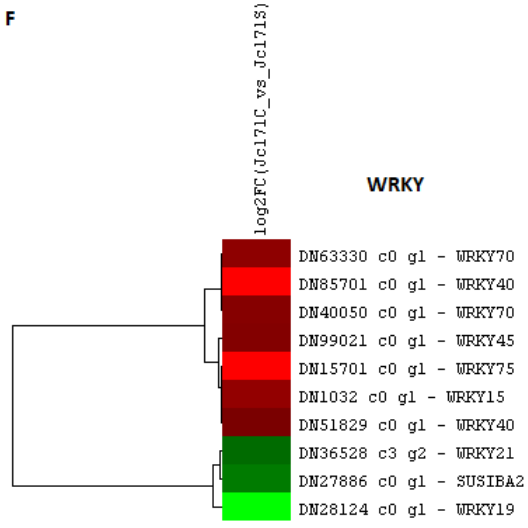
6H) family showed more up-regulation than down-regulation. A balance with up- and down-regulated members comprised the HD-ZIP family (**Figure 6I**). In turn, almost all members of the bHLH family presented down-regulation (**Figure 6C**), and the totality of the FHY3-FAR1 (**Figure 6D**) and ARF (**Figure 6G**) members exhibited down-regulation after the salt-treatment.




C
log2FC(Jc171C_vs_Jc171S)
bHLH
DN27387 c0 g1 - bHLH149
DN43665 c1 g2 - BIM1
DN87 c0 g2 - GLABRA 3
DN39297 c0 g1 - bHLH155
DN87 c0 g1 - GLABRA 3
DN39939 c0 g1 - PIF3
DN190 c0 g1 - bHLH123
DN8913 c0 g1 - bHLH62
DN27330 c0 g1 - bHLH
DN1084 c0 g1 - bHLH112
DN37141 c0 g1 - bHLH143
DN43441 c0 g1 - bHLH157
DN49789 c0 g1 - bHLH143
DN42651 c0 g1 - bHLH30
DN27387 c0 g2 - bHLH149


D
log2FC(Jc171C_vs_Jc171S)
FHY3-FAR1
DN27073 c0 g2 - FAR1
DN43957 c0 g1 - FRS5
DN35482 c2 g1 - FRS3
DN35482 c1 g1 - FRS3
DN35482 c0 g1 - FRS3
DN32534 c0 g1 - FRS4
DN27073 c0 g1 - FAR1
DN35482 c1 g2 - FRS3
DN20062 c0 g2 - FHY3
DN38006 c0 g1 - FAR1
DN37650 c0 g1 - FRS5
DN43446 c0 g1 - FRS3
DN39826 c0 g1 - FRS2
DN33760 c0 g1 - FRS5
DN36683 c0 g1 - FRS1


E
log2FC(Jc171C_vs_Jc171S)
NAC
DN30888 c0 g1 - NAC100
DN25332 c0 g1 - NAC001
DN25332 c0 g2 - NAC025
DN34336 c0 g1 - NAC072
DN27999 c0 g1 - NAC104
DN23154 c0 g2 - NAC002
DN49660 c0 g1 - AC070
DN77400 c0 g1 - NAC083
DN23154 c0 g3 - NAC002
DN6382 c0 g1 - NAC089


F
log2FC(Jc171C_vs_Jc171S)
WRKY
DN63330 c0 g1 - WRKY70
DN85701 c0 g1 - WRKY40
DN40050 c0 g1 - WRKY70
DN99021 c0 g1 - WRKY45
DN15701 c0 g1 - WRKY75
DN1032 c0 g1 - WRKY15
DN51829 c0 g1 - WRKY40
DN36528 c3 g2 - WRKY21
DN27886 c0 g1 - SUSIBA2
DN28124 c0 g1 - WRKY19

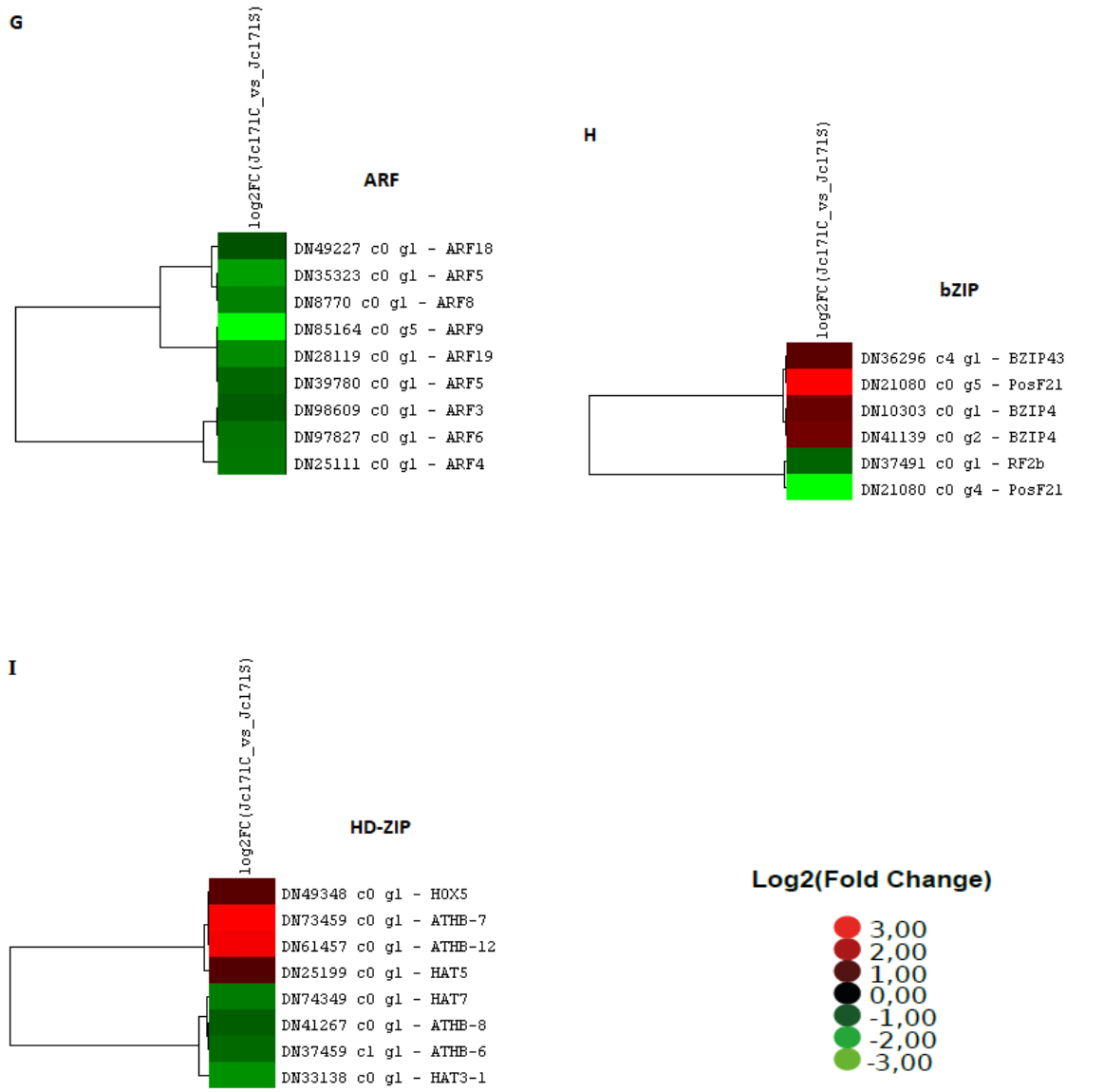


Figure 6. Heatmaps based on gene expression modulation of TFs family members identified in *Jatropha curcas* Jc171 roots after three hours of NaCl exposition (150 mM), in relation to the negative control without salt (ratio of Log_2FC values). The up- and down-regulation of the differentially expressed genes are indicated in red and green, respectively, and the intensity of the colors follow the legend.

f) Expression validation of TF DEGs by RT-qPCR analysis

Eight TF DEGs candidates were evaluated in RT-qPCR assays to confirm the *in silico* expressed profiles. The selected DEGs (*RAP2-3*, *RAV1*, *ERF9*, *DREB1H*, *ZAT12*, *PTI5*, *MYB340*, and *BZIP4*) and the reference genes (*β-tubulin* and *actin*) presented in the respective dissociation curves the unique expected amplicons (**Supplementary Figure S1**). The RT-qPCR parameters [amplification efficiency (E), *slope* (s), and correlation coefficient (R)] derived from standard curves generated using serial dilution of root cDNAs samples (accessions and treatments) and each *primer* pair presented acceptable values (**Table 1**), as those recommended following the MIQE protocol

(Bustin et al. 2009), aiming to ensure reliable relative qPCR data. In general, most of the RT-qPCR results (75 %) confirmed the *in-silico* gene expression (except for *RAP2-3* and *DREB1H*), suggesting the reliability of the expressed TFoma (**Figure 7**, and **Table 2**).

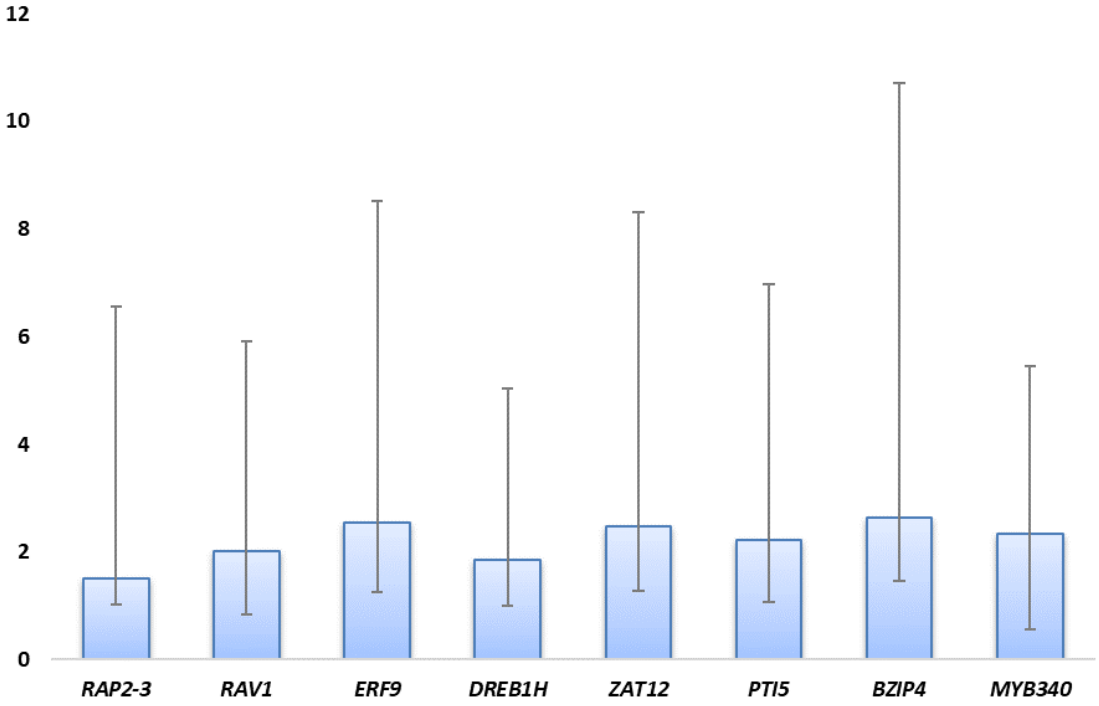


Figure 7. RT-qPCR results of eight candidate genes encoding TFs using cDNAs of *Jatropha curcas* root after three hours of NaCl exposition (150 mM). Expression data calculated by the REST software (v.2.0.13) (Pfaffl et al. 2002) considering biological and technical triplicates and *actin* and *β-tubulin* as the reference genes.

Table 1. RT-qPCR parameters [amplification efficiency (E), slope (S), correlation coefficient (R), and Y intercept] derived from the standard curves using serial dilution of *Jatropha curcas* root cDNAs samples (accessions and treatments) and each primer pair.

Gene (candidate/reference*)	E (%)	R	S	Y intercept
RAP2-3	91.51	-0.994	-3.54	27.78
RAV1	91.30	-0.998	-3.55	36.14
ERF9	105.45	-0.992	-3.20	34.68
DREB1H	98.69	-0.996	-3.35	33.64
ZAT12	104.71	-0.927	-3.21	33.88
PTI5	104.91	-0.915	-3.21	32.55
MYB340	109.78	-0.974	-3.11	34.12
BZIP4	91.03	-0.999	-3.56	32.59
*β-tubulin**	96.00	-0.986	-3.42	30.90
*Actin**	90.15	-0.998	-3.58	26.99

*reference gene: *actin* and *β-tubulin* (Ma et al., 2016).

Table 2. Selected putative transcript factor genes (DEGs) with the respective *in silico* expressions based on RNA-Seq data and their expression by RT-qPCR analysis with *Jatropha curcas* cDNAs from roots after three hours of NaCl exposition (150 mM).

Method	*RAP2-3*	*RAV1*	*ERF9*	*DREB1H*	*ZAT12*	*PTI5*	*BZIP4*	*MY340*
In silico *	2.665 (UR)	2.413 (UR)	3.390 (UR)	2.294 (UR)	3.659 (UR)	2.738 (UR)	2.072-2.260 (UR)	4.392 (UR)
RT-qPCR**	1.512 (n.s.)	2.023 (UR)	2.555 (UR)	1.849 (n.s.)	2.468 (UR)	2.225 (UR)	2.634 (UR)	2.346 (UR)

DEGs (differentially expressed genes: *p-value* ≤ 0.0001, false discovery rate (FDR ≤ 0.005), and fold change (FC) based on Log_2 (FC) ≥ 1 (up-regulated, UR) or ≤ -1 (down-regulated, DR). * Log_2FC values (FC: ratio of normalized transcript abundance observed in the stressed library in relation to the respective abundance in the control library). **Relative expression by REST software (v.2.0.13) (Pfaffl *et al*., 2002), UR (p ≤ 0.05) considering biological and technical triplicates, and *actin* and *β-tubulin* as the reference genes.

Discussion

Plant transcriptomic studies have been performed identifying TFs associated with abiotic stresses and their interaction with the transcriptional reprogramming activation in living cells (Seki et al. 2002). Since many TFs described in plant abiotic stress responses play a crucial role in stress tolerance processes (Lata et al. 2011), the expression modulated by TFs usually results in dramatic metabolic changes (Liu et al. 1999). Here, a *de novo* RNA-Seq transcriptome analysis uncovered the TFoma differentially expressed in roots of a *J. curcas* Brazilian Jc171accession plants after three hours of salt exposition (150 mM NaCl). From the assembled transcripts encoding TFs (1,876), almost 8% were differentially expressed after the salt stimulus (78 UR and 70 DR; thresholds: *p-value* ≤ 0.0001, FDR ≤ 0.005, Log_2FC ≥ 1 or ≤ -1). Jc171 seeds presented particular ability in germination, despite the presence of NaCl (50, 75, 100, and 150 mM) (Lozano-Isla et al. 2018).

a) The GO enrichment analysis

The enriched GO terms associated with UR or DR DEGs individualized the two distinct sets of genes encoding TFs.

The enriched GO terms based on the UR DEGs pointed out stress responses, as expected, but also cell death among others terms (Table S2). Programmed cell death

(PCD) is a critical process in eukaryotic cells (Lam 2008), mediating adaptive responses of plants to environmental stresses (Shabala 2009). The Jc171 accession after three hours of the salt exposition presented visible damages in the leaves, probable progressing to necrosis (**Figure 2**). The ionic imbalance induced by salt stress also promotes PCD (Katsuhara and Shibasaka 2000; Huh et al. 2002). Otherwise, the transcription factor MYB108 (also named BOS1, botrytis sensitive1), despite suppressing PCD dissemination in *A. thaliana* injury sites (Mengiste et al. 2003), was associated with oxidative stress, salinity, and water deficit responses. In the present work, two induced *MYB108-related* DEGs were identified (DN31858_c0_g1 and (DN20826_c0_g1; **Supplementary Table S5**).

In turn, the enriched GO terms associated with the DR DEGs highlighted more plant developmental activities, including *regulation of circadian rhythm* (Table S6), which is involved in adaptive responses to stresses (Hotta et al. 2007; Legnaioli et al. 2009; Grundy et al. 2015; Seo and Mas 2015). Metabolic and biochemical processes affected by circadian rhythms include photosynthesis and respiration (Kreps and Kay 1997), stomatal opening (McClung, 2001, water uptake by roots (Takase et al. 2011), and cellular Ca^{2+} levels oscillations (Johnson et al. 1995). The MYB-related TF CCA1 (circadian clock associated1) is a transcriptional activator strictly involved in circadian rhythm regulation, binding promoters of at least two genes (*Lhcb*, *light-harvesting chlorophyll a/b-protein*) encoding proteins related to the photosystem II (Wang and Tobin 1998). In the present work, the putative *CCA1* codified by the DEG DN29723_c0_g2 (**Supplementary Table S5**) could contributed with the expected photosynthesis inhibition of Jc171 after the salt stimulus. It is known that saline stress causes disturbances in photosynthesis, leading to a decrease in plant growth (Sudhir and Murthy 2004; Barhoumi et al. 2007), as salinity increases soil osmotic potential, reducing water uptake by root, and compromising root growth rates (Munns and Tester 2008) and growth of the leaves, reducing the photosynthesis (Julkowska and Testerink 2015).

b) The TF enrichment analysis

The enriched TFs predicted interacting with the UR DEGs codifying TFs were more comprehensive (broadly distributed into TF families) than those predicted based

on the DR DEGs (**Figure 5B**). The predicted TFs indicated an increased regulatory demand for particular TFs in the Jc171 salt-stress response.

Two enriched TF from the BBR-BPC family [BPC1 (BASIC PENTACYSTEINE1) and BPC6 (BASIC PENTACYSTEINE6)] stood out probable inducing 29 – 41 TFs after the salt stimulus, 25 of them shared by the two enriched TFs (**Supplementary Table S6**). The BPC1 is the regulator of the floral homeotic *STK* gene (*seedstick*), which controls tissue identity through the regulation of a wide range of processes (Kooiker et al. 2005), while BPC6 is a transcriptional regulator of *LHP1* (*like heterochromatin protein1*) gene, which is associated with a PRC (polycomb-repressive complex) component involved in plant epigenetic control by histone methylation (Schuettengruber and Cavalli 2013; Hecker et al. 2015).

Enriched TFs members of Dof family also presented meaningful interactions with the UR DEGs. The enriched Dof zinc finger proteins DOF3.1 and DOF3.6, predicting targeting 28 - 33 DEGs, shared 22 DEGs as their targets (**Supplementary Table S6**). The DOF3.6 was associated with plant growth and development, targeting genes induced by salicylic acid, such as *ORG1*, *ORG2*, and *ORG3* (*OBP3-responsive genes*; Kang et al. 2003). OBP3 is also a Dof transcription factor. Other enriched TFs from Dof family were DOF5.6 and DOF3.4, each one regulating 23 UR DEGs (13 shared targets; **Supplementary Table S6**). DOF5.6 acts on the regulation of vascular tissue development (Guo et al. 2009), while DOF3.4 is involved in the cell cycle regulation (Skirycz et al. 2008).

Among the enriched TF members of the AP2/ERF family, ERF1B (Ethylene-responsive transcription factor 1B) and ERF5 (Ethylene-responsive transcription factor 5) stood out, regulating 15 and 16 UR DEGs, respectively (12 shared targets, **Supplementary Table S6**). ERF1B was related to salinity tolerance in *Avicennia officinalis* (Krishnamurthy et al. 2017), while ERF5 was previously associated with drought and salinity responses in *Solanum lycopersicum* (Pan et al. 2012).

Concerning the C2H2 family, the enriched transcription factor IIIA (TFIIIA) stood out, targeting 23 induced TF DEGs (**Supplementary Table S6**); the *TFIIIA* gene over expression was associated with salt-tolerance in *Medicago truncatula* (De Lorenzo et al. 2007). In turn, the three enriched TFs from the WRKY family predicted interactions targeting only two-three UR DEGs (**Supplementary Table S6**), e.g., the

TF WRKY1, already associated with salinity and drought tolerance in *Triticum turgidum* (Mondini et al. 2012), were predicted targeting the promoters of *ERF1B* and *ZAT10* (Zinc finger protein ZAT10) DEGs (**Supplementary Table S6**).

c) The differentially expressed TFoma after the salt-treatment

Concerning almost 70 TF families identified in plants (Pérez-Rodríguez et al. 2010; Hong 2016), based on the DNA-binding domains (Riechmann et al. 2000), the identified *J. curcas* DEGs (148) encompassed 23 TF families with members displaying differential expression after the salt stimulus. Although the TF involvement in abiotic stress tolerance has been established (Reyes et al. 2004; Yanhui et al. 2006; Du et al. 2009; Yang et al. 2011; Cabello et al. 2012; Xie et al. 2012; Zhu et al. 2014; Zhang et al. 2014), and TFs families have been reported to orchestrate stress response pathways in plants, such as MYB, AP2/ERF, bZIP, MYC, NAC, HD-zip, and WRKY (Singh et al. 2002; Shameer et al. 2009), a comprehensive TFoma covering TFs differentially expressed in *J. curcas* roots after salt stimulus has not been presented. Until now, TF families presenting a *J. curcas* genome-wide analysis include WRKY (Xiong et al. 2013), NAC (Wu et al. 2015), MYB (Zhou et al. 2015), and AP2/ERF (Tang et al. 2016). In the mentioned reports, the authors exploring DGE (*Digital Gene Expression*) analysis identified TFs from those family in tissues (root, stem, leaf or seed) of plants under stress (drought, phosphate or nitrogen starvation, and salinity). The applied salt stress involved plants/seedlings under 100 mM NaCl for 2 hours or 2 - 7 days. Also, the gene expression of selected candidates when validated was performed using semi-quantitative RT-PCR (Xiong et al. 2013; Zhou et al. 2015; Tang et al. 2016) or RT-qPCR analysis (Wu et al. 2015; Tang et al. 2016).

In the present work, despite the wide distribution covering 23 TF families, some families were not detected presenting DEGs. Those families included ABI3VP1, LFY, SBP, Alfin-like, CCAAT, LIM, Sigma70-like, CPP, LOB, SRS, CSD, TAZ, ARR-B, DBP, mTERF, TCP, BBR/BPC, E2F-DP, Tify, BES1, EIL, TIG, BSD, FHA, NOZZLE, TUB, G2-like, OFP, ULT, GeBP, SAP, VARL, Dof, PBF-2-like, VOZ, GRF, PLATZ, YABBY, RWP-RK, HRT, S1Fa-like, Zn-clus, and C3H.

Besides not represented in the present TFoma, Dof family members have been associated with abiotic stress tolerance (Li et al. 2016; Wen et al. 2016), including salinity (Ma et al. 2015). The same has been reported with TCP (salinity: Zhou et al.

2013; Yin et al. 2018), and CCAAT family members (drought: Nelson et al. 2007; Kuromori et al. 2014). A broad RNA-Seq analysis with *Hippophae rhamnoides* plants under drought stress presented repressed TF members from the families ABI3VP1, Dof, YABBY, CCAAT, FHA, G2-like, and C3H, while the induced TFs involved members from the families mTERF, PLATZ, TUB, LIM, and Orphans (Ye et al. 2018).

The main results involving TF family members encoded by the Jc171 DEGs are highlighted below.

c.1.) AP2/ERF family

Members of the AP2/ERF family play fundamental roles in plant development and biotic or abiotic stress responses (Tang et al. 2017). Some AP2/ERF possible encoded by UR DEGs were:

- Ethylene-responsive transcription factor 3 (DN37072_c1_g1; **Supplementary Table S5**): the over-expression of *ERF3* gene was confirmed in plants under cold and drought stress (Cao et al. 2006b; Trujillo et al. 2008); also, its over-expression in wheat (*Triticum aestivum*) transgenic plants, positively regulated physiological adaptive response to salinity and drought tolerance, through increasing proline content, chlorophyll accumulation, and cell redox homeostasis regulation (Rong et al. 2014).
- Ethylene-responsive transcription factor11 (DN7846_c0_g1; **Supplementary Table S5**): the *ERF1* gene expression was modulated by jasmonic (JA; Dombrecht et al. 2007), gibberellic acid (GA; Liu and Hou 2018), and abiotic stresses, including cold (Vergnolle et al. 2005); the TF protein contains a repressor domain that interacts with dehydration-responsive element (DRE) in the promoter of the *ACS2/5* (*1-Aminocyclopropane-1-carboxylic acid synthase - ACS*) gene, affecting the ETH biosynthesis under increasing ABA levels (Li et al. 2011).
- Ethylene-responsive transcription factor 21 (DN51470_c0_g1; **Supplementary Table S5**): the ERF21 binds to the promoter of *RD29A* gene (Mitsuda et al. 2010), which is recognized to regulate mechanisms of perception and fast induction in water deficit situations (Yamaguchi-Shinozaki et al. 1993).
- Ethylene-responsive transcription factor ERF12 (DN27804_c0_g1; **Supplementary Table S5**): the *ERF12* (also known as *DREB26*) gene was highly responsive to salt treatment (200 mM NaCl), heat, and drought (Krishnaswamy et al. 2011); as a

DREB subfamily member (Guo et al. 2005) has an amphiphilic repression motive (Zhao et al. 2014), which is characteristic of repressor proteins that inhibit the expression of stress-related genes (Kazan 2006).

- Dehydration-responsive element-binding protein 1H (DN11306_c0_g1; **Supplementary Table S5**): as a DREB subfamily member, DREB1H plays a crucial role in plant development and gene expression mediated by abiotic stresses (Zhao et al. 2014). However, the RT-qPCR analysis not confirmed the DEG up-regulation (**Table 2**).
- Dehydration-responsive element-binding protein 2C (DN18083_c0_g2; **Supplementary Table S5**): the DREB2C is a transcriptional activator of genes, such as *COR15A* (*cold-regulated 15a*; salinity tolerance; Song et al. 2014), *HsfA3* (*heat shock factor a3*; heat stress response; Chen et al. 2010), *NCED9* (*9-cis-epoxycarotenoid dioxygenase 9*; ABA biosynthesis; Je et al. 2014a), *CYS4 (phytocystatin 4*; thermotolerance; Je et al. 2014b).
- Ethylene-responsive transcription factor 1B (DN74024_c0_g1; **Supplementary Table S5**): the ERF1B is related to the ETH signaling (Corbacho et al. 2013); the transcript up-regulation has been reported in plant responding to drought in soybean (Ferreira Neto et al. 2013), and tomato (Egea et al. 2018).
- AP2/ERF and B3 domain-containing transcription factor RAV1 (DN49504_c0_g1; **Supplementary Table S5**): the RAV1 presents roles in ABA signaling during seed germination, and the initial seedling development (Feng et al. 2014). The RT-qPCR analysis validated the DEG up-regulation (**Table 2**).
- Ethylene-responsive transcription factor 9 (DN31778_c0_g1; **Supplementary Table S5**): the *ERF9* gene is induced in leaves and roots at different stages of development under saline stress in tomato genotypes (Gharsallah et al. 2016). The RT-qPCR analysis confirmed the DEG up-regulation (**Table 2**).
- Ethylene-responsive transcription factor RAP2-3 (DN5779_c0_g1; **Supplementary Table S5**): the RAP2-3 modulates osmotic tolerance inducing genes like *PDC1* (*pyruvate decarboxylase1*), *SUS1* and *SUS4* (*sucrose synthases*) when associated to ABA signaling (Gibbs et al. 2015; Papdi et al. 2015). In this case, the RT-qPCR analysis not confirmed the DEG up-regulation (**Table 2**).
- Pathogenesis-related genes transcriptional activator PTI5 (DN43242_c3_g1; **Supplementary Table S5**): the PTI5 activates genes regulated by salicylic acid

(SA), such as *PR1* and *PR2* (pathogenesis-related genes) (Gu et al. 2002), which are involved in the systemic acquired resistance (SAR) process during phytopathogen infection (Ryals et al. 1996; Feys and Parker 2000). The RT-qPCR analysis confirmed the DEG up-regulation (**Table 2**).

c.2.) WRKY family

One of the primary TF groups involved in the control of biotic and abiotic stress responses (Ulker and Somssich 2004; Rushton et al. 2010). In the proposed TFoma, 10 members (7 UR and 3 DR) were identified, and some of them are presented below.

- WRKY transcription factor 40 (DN51829_c0_g1, and DN85701_c0_g1; **Supplementary Table S5**): the WRKY40 acts primarily on plant defense susceptibility but suffering influence from previous stresses (stressors may have an antagonistic, synergistic or additive effect on plant; Anderson et al. 2004; Asselbergh et al. 2008); this TF negatively modulates the expression of repressors of the JA signaling pathway (JAZ7, JAZ8, and JAZ10), taking part in the defense systems (Glazebrook 2005).
- WRKY transcription factor 70 (DN63330_c0_g1; **Supplementary Table S5**): the WRKY70 is a saline stress-response regulator interacting with another TF (Cys2/His2 zinc finger Zat7); both TFs presented involvement increasing salt tolerance (Ciftci-Yilmaz et al. 2007).
- WRKY transcription factor 45 (DN99021_c0_g1; **Supplementary Table S5**): the *WRKY45* gene expression is induced in ABA hormone-related response, and also in stress responses, including NaCl, dehydration, cold, heat, and pathogens infections (Yu and Qiu 2009).
- WRKY transcription factor 57 (DN40050_c0_g1; **Supplementary Table S5**): the WRKY57 interacts with promoters of genes, such as *RD29A* (Yamaguchi-Shinozaki and Shinozaki 1993) and *NCED3* (Chernys and Zeevaart 2000), assisting the plant adaptation to water stress tolerance, by increasing ABA levels (Finkelstein et al. 2002); the phytohormone ABA regulates essential processes (germination, seed dormancy, and stomatal behavior; Liotenberg et al. 1999); also, this TF affects *A. thaliana* germination under ABA influence and abiotic stress (osmotic, salinity and drought; Jiang et al. 2012).

c.3.) MYB family

Members of the MYB family have been investigated in biotic and abiotic stress responses (Denekamp and Smeekens 2003; Seo et al. 2009). In the proposed TFoma, 14 MYB members were associated to the DEGs (8 UR and 6 DR), including:

- Transcription factor MYB108 (DN20826_c0_g1; **Supplementary Table S5**): MYB108 regulates abiotic stresses responses (e.g., salinity, drought and cold) by the JA pathway and the ROS-mediated cellular signaling (Mengiste et al. 2003; Schmid et al. 2005).
- Transcription factor KUA1 (DN61071_c0_g1; **Supplementary Table S5**): KUA1 is a transcriptional repressor of genes encoding peroxidases (PRXs; Lu et al. 2014); PRXs also promote ROS generation, such as H_2O_2, which can cleave the polymers of the cell wall, restricting plant growth (Passardi et al. 2004).
- Myb-related protein 340 (DN41011_c2_g4; **Supplementary Table S5**): MYB340 activates the *PAL* gene (*phenylalanine ammonia-lyase*) transcription binding on its promoter (Moyano et al. 1996); the PAL enzyme is involved in the phenylpropanoid metabolism, and stresses (e.g., drought, and salinity) stimulating that metabolism (Cabane et al. 2012) generate precursors for lignin biosynthesis (Davin and Lewis 1992), which is also associated to stress tolerance (Liu et al. 2018). The RT-qPCR analysis confirmed the DEG up-regulation (**Table 2**).
- Transcription factor MYBS1 (DN15053_c0_g1; **Supplementary Table S5**): MYBS1 recognizes the TATCCA motif in promoters of genes (e.g., *α-amylase* gene), inducing its expression (Lu et al. 2002); however, during salt stress the α-amylase activity, degrading starch and releasing soluble sugar molecules, is reduced (Lin and Kao 1995; Othman and Al-Karaki 2006; Siddiqui and Khan 2011), affecting processes, such as germination and plant growth (Mei and Song 2008).
- Transcription factor SRM1 (Salt-Related MYB1; DN50735_c0_g1; **Supplementary Table S5**): SRM1 regulates the synthesis and signaling of ABA during germination and seed development in salinity conditions, activating the expression of the *NCED3*/*STO1* gene, which is a mediator of the ABA biosynthesis (Iuchi et al. 2001; Barrero et al. 2006).

- MYB-like transcription factor ETC1 (enhancer of try and cpc 1; DN98044_c0_g1; **Supplementary Table S5**): ETC1 acts as a negative regulator of trichome development, but also promoting an increase in the development of root hairs (Kirik et al. 2004).
- Transcription factor MYB59 (DN20471_c0_g1; **Supplementary Table S5**): TF involved in cell cycle regulation, and root growth (Mu et al. 2009); TF also responding to ETH, and JA (Razzaque et al. 2017).

c.4) HD-ZIP family

Members of the HD-Zip family play a significant role in plant growth and development, responding to several phytohormone stimuli, and stresses (Ge et al. 2015; Mao et al. 2016). In wheat (*Triticum aestivum*) plants, the salt-sensitive CS genotype presented 21 induced HD-Zip genes, while the salt-tolerant DK presented 18 (Yue et al. 2018). In the present TFoma, eight HD-Zip members and DEGs were identified (4 UR and 4 DR), and some of them stood out:

- Homeobox-leucine zipper protein HAT5 (DEG DN25199_c0_g1; **Supplementary Table S5**): TF associated with salt stress tolerance in *Thellungiella halophila* (halophytic plant; Wang et al. 2004).
- Homeobox-leucine zipper protein ATHB-12 (DN61457_c0_g1; **Supplementary Table S5**): in transgenic plants under drought conditions, ATHB12 and ATHB7 act as negative plant development regulators in response to the ABA levels (Olsson et al. 2004); the salinity induces ABA biosynthesis (Mahajan and Tuteja 2005), in response to the osmotic and water deficit stresses (Popova et al. 1995; He and Cramer 1996).
- Homeobox-leucine zipper protein ATHB-7 (DN73459_c0_g1; **Supplementary Table S5**): in tomato, the ectopic expression of *ATHB7* conferred drought tolerance (Mishra et al. 2012); also, it was strongly induced by drought and by ABA (Söderman et al. 1996).

c.5) NAC family

Members of the NAC (**N**AM, **A**TAF, and **C**UC) family present crucial roles in plant development (Kunieda et al. 2008; Ohtani et al. 2011) and stress responses

(Takasaki et al. 2015). From the induced DEGs, three NAC TFs stood out after the NaCl application.

- NAC domain-containing protein 72 (DN34336_c0_g1; **Supplementary Table S5**): the *AtNAC072* gene was induced by ABA (100 μM ABA), salinity (250 mM NaCl), and drought (Tran et al. 2004); NAC72 (*Poncirus trifoliata*) is the transcriptional repressor of *ADC* (arginine decarboxylase) gene (Wu et al. 2016), which enzyme is critical for the putrescine (Put) biosynthesis (Put is an osmoprotectant compound reducing oxidative damages in roots; Zhang et al. 2014); *NAC72* induction has been associated with stress level (Wu et al. 2016); the related TF binds to the CATGTG motif in promoters of genes, such as *ERD1* (*early responsive to dehydration stress 1*) gene, which protein (ClpA, ATP-dependent CLP protease ATP-binding subunit clpA; Tran et al. 2004) is essential for the maintenance of the chloroplast enzymatic apparatus (Sjögren and Clarke 2011).
- NAC domain-containing protein 100 (DN30888_c0_g1; **Supplementary Table S5**): NAC100 binds promoters of cell expansion-related genes, such as *CESA2* (*cellulose synthase2*), and *PIP* (*Plasma Membrane Intrinsic Protein*) aquaporins (Pei et al. 2013), which are gateways for cell membrane water exchange (Yaneff et al. 2015).
- NAC domain-containing protein 2 (DN23154_c0_g2, and DN23154_c0_g3; **Supplementary Table S5**): *NAC2* gene induction in roots of *A. thaliana* plants responding to saline stress (200 mM NaCl) was reported (He et al. 2005).

c.7) bZIP family

Members of the bZIP family mediate several biological processes, including energetic metabolism (Baena-González et al. 2007), cell expansion (Fukazawa et al. 2000), tissue and organ differentiation (Silveira et al. 2007), seed maturation and embryogenesis (Lara et al. 2003). bZIP members also participate in biotic (Thurow et al. 2005), and abiotic stress responses (Ji et al. 2018), including drought and salinity (Ying et al. 2012; Liu et al. 2014). Two bZIP members associated with three induced DEGs stood out:

- Basic leucine zipper 43 (DN36296_c4_g1; **Supplementary Table S5**): bZIP43 is a positive regulator of *bHLH109* gene (Nowak and Gaj 2016), which was

associated increasing LEA (late embryogenesis abundant) protein and enhancing plant stress tolerance (Nowak and Gaj 2016).

- Basic leucine zipper 4 [DN10303_c0_g1, and DN41139_c0_g2; **Supplementary Table S5**]: bZIP4 is also a positive regulator of the bHLH109 gene (Nowak and Gaj 2016), and the DEG (DN41139_c0_g2) up-regulation was confirmed by RT-qPCR results (**Table 2**).

c.8) C_2H_2-ZFP (C_2H_2 type Zinc Finger Protein) family

Members of the C_2H_2-ZFP family are involved in several biological processes (Gourcilleau et al. 2011), including growth mediation, plant development, and abiotic stress responses (Ding et al. 2016). In this case, the TF member stood out was:

- Zinc finger protein ZAT12 (DN26908_c0_g1; **Supplementary Table S5**): this TF regulate the expression of several oxidative-stress-response genes, including *APX* (*Ascorbate Peroxidase*), *CAT* (*Catalase*), *GR* (*Glutathione Reductase*), *POD* (*Guaiacol Peroxidase*) and *SOD* (*Superoxide Dismutase*) (Rizhsky et al. 2004; Davletova et al. 2005; Rai et al. 2012); the DEG up-regulation was confirmed the by RT-qPCR analysis (**Table 2**).

Conclusions

The present study represents the first TFoma differentially expressed in roots of *J. curcas* plants after salt stress (three hours of NaCl exposition, 150 mM), based on RNA-Seq de novo assembly strategy followed by gene expression validation in RT-qPCR assays. The proposed TFoma represents 148 DEGs (78 UR and 70 DR) codifying TFs encompassing 23 TF families. The GO enrichment analysis identifying those terms over-represented and exclusively associated with the UR or the DR DEGs, differentiated the two sets of DEGs. Enriched GO terms related to stress responses represented the UR DEGs, while GO terms more related to the basal metabolism represented the DR DEGs. The TF enrichment analysis stood out the most representative TFs predicted interacting with the sets of genes encoding TFs (UR, DR, and the non-DEGs). Predicted TFs regulating cognate TFs genes (as their target genes) were identified, as well as enriched TFs predicting interactions with more than 40 UR DEGs; some of them sharing more than 20 targets. These TFs are promising transgene candidates. The RT-qPCR analysis confirmed the *in silico* gene expression of 75% of

eight selected DEGs (from different TF families), and some of them could be functional molecular markers for marker-assisted selection on plant breeding programs, helping to develop *J. curcas* salt-tolerant accessions. The results help to understand the molecular mechanisms involved in *J. curcas* plants responding to salt-exposure.

Competitive interests

The authors declare that they have no competing interests.

Authors contributions

LE, MFP, EB, and EAK conceived and designed the experiments; GALC, MCPS, MDS, and JRCFN carried out the experiments; GALC, MCPS, MDS, and JRCFN analyzed the data; GALC and EAK wrote and revised the paper. All authors read and approved the final manuscript.

Acknowledgements

This work was supported (grants and fellowships) by the Brazilian agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 404357/2013-0, and 311894/2017-8), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE; IBPG-0271-5.01/14), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Authors are thankful to Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA SOJA, Londrina, PR - Brasil), Universidade Federal de Alagoas (UFAL), and Universidade Federal de Pernambuco (UFPE) for the infrastructure facilities.

References

Anderson JP, Badruzsaufari E, Schenk PM, Manners J, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K (2004) Antagonistic interaction between abscisic acid and JA-ETH signaling pathways modulates defense gene expression and disease resistance in Arabidopsis. Plant Cell, 16, 3460–3479. doi 10.1105/tpc.104.025833

Asselbergh B, Achuo AE, Hofte M, Van Gijegem F (2008) Abscisic acid deficiency leads to rapid activation of tomato defence responses upon infection with *Erwinia chrysanthemi*. Mol. Plant Pathol, 9, 11–24. doi 10.1111/j.1364-3703.2007.00437.x

Baena-González E, Rolland F, Thevelein JM, Sheen J (2007) A central integrator of transcription networks inplant stress and energy signalling. Nature, 448, 938–942. doi 10.1038/nature06069

Barhoumi Z, Djebali W, Chaibi W, Abdelly C, Smaoui A (2007) Salt impact on photosynthesis and leaf ultrastructure of *Aeluropus littoralis*. J Plant Res 120: 529–537. doi 10.1007/s10265-007-0094-z

Barrero JM, Rodríguez PL, Quesada V, Piqueras P, Ponce MR, Micol JL (2006) Both abscisic acid (ABA)-dependent and ABA-independent pathways govern the induction of NCED3, AAO3 and ABA1 in response to salt stress. Plant Cell Environ 29: 2000–2008. doi 10.1111/j.1365-3040.2006.01576.x

Benko-Iseppon AM, Soares-Cavalcanti NM, Nogueira ACW, Silva LCB, Silva RRM, Almeida PML (2005) Genes associados a estresses bióticos e abióticos em feijão-caupi [*Vigna unguiculata* (l) Walp] e outras angiospermas In: Araújo, EM, Simabukuro, EA, Nogueira, R.J.M.C. Lagoa. 11-498.

Bolger AM, Lohse M, Usadel B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30: 2114–2120. doi: 10.1093/bioinformatics/btu170

Bustin SA, BenesV, Garson JA, Hellemans J, Huggett J, Kubista M (2009) The MIQE guidelines: Minimum Information for publication of quantitative real-time PCR experiments Clin. Chem, 55, 611-622. doi: 10.1373/clinchem.2008.112797

Cabane M, Afif D, Hawkins S (2012) Lignins and abiotic stresses. Adv Bot Res, 61: 220-246. doi https://doi.org/10.1016/B978-0-12-416023-1.00007-0

Cabello JV, Arce AL, Chan RL (2012) The homologous HD-Zip I transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesis related and glucanase proteins. Plant J 69:141–153. doi: 10.1111/j.1365-313X.2011.04778.x

Campos, MLO, Hsie BS, Granja JAA, Correia RM, Silva SRS, Almeida-Cortez JS (2012) Photosynthesis and antioxidant activity mechanisms in *Jatropha* curcas L under salt stress. Braz J Plant Physiol, 24(1):55-67. doi http://dx.doi.org/10.1590/S1677-04202012000100008

Cao YF, Wu YF, Zhang Z, Song FM (2006) Overexpression of the rice EREBP-like gene OsBIERF3 enhances disease resistance and salt tolerance in transgenic tobacco. Physiological and Molecular Plant Pathology, 67. 202-211. doi https://doi.org/10.1016/j.pmpp.2006.01.004

Chen H, Hwang JE, Lim CJ, Kim DY, Lee SY, Lim CO (2010) Arabidopsis DREB2C functions as a transcriptional activator of HsfA3 during the heat stress response. Biochem Biophys Res Commun, 401, 238-244. doi: 10.1016/j.bbrc.2010.09.038

Chernys JT, Zeevaart JA (2000) Characterization of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in avocado. Plant Physiol, 124, 343-353. doi 10.1104/pp.124.1.343

Ciftci-Yilmaz S, Morsy MR, Song L (2007) The EAR-motif of the Cys2/His2-type zinc finger protein Zat7 plays a key role in the defense response of Arabidopsis to salinity stress. J Biol Chem, 282:9260–9268. doi 10.1074/jbc.M611093200

Corbacho J, Romojaro F, Pech JC, Latche A, Gomez-Jimenez MC (2013) Transcriptomic events involved in melon mature-Fruit abscission comprise the

sequential induction of Cell-Wall degrading genes coupled to a stimulation of endo and exocytosis. PLoS One, 8. doi e58363-10.1371/journal.pone.0058363.

Dasgan HY, Aktas H, Abak K, Cakmak, I (2002) Determination of screening techniques to salinity tolerance in tomatoes and investigation of genotypes responses. Plant Sci, 163:695-703. doi 10.1016/S0168-9452(02)00091-2

Davin LB, Lewis NG (1992) Phenylpropanoid metabolism: Biosynthesis of monolignols, lignans and neolignans, lignins and suberins. In Phenolic Metabolism in Plants, Stafford and R.K. Ibrahim, eds (New York: Plenum Press), 325-375.

Davletova S, Schlauch K, Coutu J, Mittler R (2005) The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in Arabidopsis. Plant Physiol. 139, 847–856. doi 10.1104/pp.105.068254

Denekamp M, Smeekens SC. Integration of wounding and osmotic stress signals determines the expression of the AtMYB102 transcription factor gene. Plant Physiol. 2003; 132(3): 1415-1423. doi 10.1104/pp.102.019273

Deore AC, Johnson TS (2008) High-frequency plant regeneration from leaf-disc cultures of *Jatropha curcas* L.: an important biodiesel. Plant Biotechnol Rep 2: 7-11. doi:10.1007/s11816-008-0042-y

De Lorenzo L, Merchan F, Blanchet S, Megiàs M, Frugier F, Crespi M, Sousa S (2007) Differential expression of the TFIIIA regulatory pathway in response to salt stress between *Medicago truncatula* genotypes. Plant Physiol, 145, pp. 1521-1532. doi 10.1104/pp.107.106146

Ding W, Wang Y, Fang W, Gao S, Li X, Xiao K (2016) TaZAT8, a C2H2-ZFP type transcription factor gene in wheat, plays critical roles in mediating tolerance to Pi deprivation through regulating P acquisition, ROS homeostasis and root system establishment. Physiol Plant 158(3):297-311. doi: 10.1111/ppl.12467.

Dombrecht B, Xue GP, Sprague SJ, Kirkegaard JA, Ross JJ, Reid JB, Fitt GP, Sewelam N, Schenk PM, Manners JM (2007) MYC2 differentially modulates diverse jasmonate-dependent functions in Arabidopsis. Plant Cell 19: 2225–2245. doi https://doi.org/10.1105/tpc.106.048017

Du H, Zhang L, Liu L, Tang XF, Yang WJ, Wu YM, Huang YB, Tang YX (2009) Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry (Mosc) 74: 1–11. doi 10.1134/S0006297909010015

Egea I, Albaladejo I, Meco V, Morales B, Sevilla A, Bolarin MC, Flores FB (2018) The drought-tolerant Solanum pennellii regulates leaf water loss and induces genes involved in amino acid and ethylene/jasmonate metabolism under dehydration. Scientific Reports, 12, 8(1):2791. doi: 10.1038/s41598-018-21187-2.

Epstein E (1972) Mineral nutrition of plants: principles and perspectives. New York: John Wiley e Sons. doi https://doi.org/10.1002/jpln.19721320211

Feng CZ, Chen Y, Wang C, Kong YH, Wu WH, Chen YF (2014) Arabidopsis RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of

ABI3, ABI4, and ABI5 during seed germination and early seedling development. Plant J, 80, 654–668. doi: 10.1111/tpj.12670

Ferreira Neto JRC, Pandolfi V, Guimaraes FCM, Benko-Iseppon AM, Romero C, De Oliveira Silva RL (2013) Early transcriptional response of soybean contrasting accessions to root dehydration. PLoS ONE 8:e83466. doi: 10.1371/journal.pone.0083466.

Feys BJ, Parker JE (2000) Interplay of signaling pathways in plant disease resistance. Trends Genet, 16, 449–455. doi https://doi.org/10.1016/S0168-9525(00)02107-7

Finkelstein RR, Gamplal SS, Rock CD (2002) Abscisic acid signaling in seeds and seedlings. Plant Cell, 14 (Suppl.), S15-S45. doi https://doi.org/10.1105/tpc.010441

Fukazawa J, Sakai T, Ishida S, Yamaguchi I, Kamiya Y, Takahashi Y (2000) Repression of shoot growth, a bZIP transcriptional activator, regulates cell elongation by controlling the level of gibberellins. Plant Cell. 12,901–915. doi 10.1105/tpc.12.6.901

Ge XX, Liu Z, Wu XM, Chai LJ, Guo WW (2015) Genome-wide identification, classification and analysis of HD-Zip gene family in citrus and its potential roles in somatic embryogenesis regulation. Gene, 574, 61–68. doi: 10.1016/j.gene.2015.07.079

Gharsallah C, Fakhfakh H, Grubb D, Gorsane F (2016) Effect of salt stress on ion concentration, proline content, antioxidant enzyme activities and gene expression in tomato cultivars. AoB Plants 8:plw055. doi: 10.1093/aobpla/plw055

Gibbs DJ, Conde JV, Berckhan S, Prasad G, Mendiondo GM, Holdsworth MJ (2015) Group VII ethylene response factors coordinate oxygen and nitric oxide signal transduction and stress responses in plants. Plant Physiol 169: 23–31. doi https://doi.org/10.1104/pp.15.00338

Gourcilleau D, Lenne C, Armenise C, Moulia B, Julien JL, Bronner G (2011) Phylogenetic study of plant Q-type C2H2 zinc finger proteins and expression analysis of poplar genes in response to osmotic, cold and mechanical stresses. DNA Res, 18: 77–92. doi: 10.1093/dnares/dsr001

Glazebrook J (2005) Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. Annu Rev Phytopathol, 43:205-227. doi 10.1146/annurev.phyto.43.040204.135923

Grabherr, MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-seq data without a 253 reference genome. Nature biotechnology, 29(7):644–52. 2011. doi 10.1038/nbt.1883

Grundy J, Stoker C, Carré IA (2015) Circadian regulation of abiotic stress tolerance in plants. Front Plant Sci 6:648. doi 10.3389/fpls.2015.00648

Gu YQ, Wildermuth MC, Chakravarthy S, Loh YT, Yang C, He X, Han Y, Martin GB (2002) Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in Arabidopsis. Plant Cell 14: 817–831. doi 10.1105/tpc.000794

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of Arabidopsis transcription factors. Bioinformatics 21:2568–2569. doi 10.1093/bioinformatics/bti334

Guo Y, Qin G, Gu H, Qu L-J (2009) Dof5.6/HCA2, a Dof transcription factor gene, regulates interfascicular cambium formation and vascular tissue development in Arabidopsis. Plant Cell 21:3518-3534. doi https://doi.org/10.1105/tpc.108.064139

Gurgel M, Fernandes P, Santos F, Gheyi H, Bezerra I, Nobre R (2003) Estresse salino na germinação e formação de porta-enxerto de aceroleira. Rev Bras Eng Agr Amb, v.7(1):31-6. doi http://dx.doi.org/10.1590/S1415-43662003000100006

He T, Cramer GR (1996) Abscisic acid concentrations are correlated with leaf area reductions in two salt-stressed rapid-cycling Brassica species. Plant and Soil. 179: 25–33.

He JX, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY (2005) AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J 44:903–916. doi 10.1111/j.1365-313X.2005.02575.x

Hecker A, Brand LH, Peter S, Simoncello N, Kilian J, Harter K, Gaudin V, Wanke D (2015) The Arabidopsis GAGA-Binding factor BASIC PENTACYSTEINE6 recruits the Polycomb-Repressive Complex1 component Like Heterochromatin Protein1 to GAGA DNA motifs. Plant Physiol 168: 1013– 1024. 2015. doi: 10.1104/pp.15.00409

Hong JC (2016) General Aspects of Plant Transcription Factor Families. In Gonzalez DH, editor. Plant Transcription Factors: Evolutionary, Structural and Functional Aspects, Academic Press, 35–56. doi https://doi.org/10.1016/B978-0-12-800854-6.00003-8

Hotta CT, Gardner MJ, Hubbard KE, Baek SJ, Dalchau N, Suhita D, Dodd AN, Webb AAR (2007) Modulation of environmental responses of plants by circadian clocks. Plant Cell Environ 30:333–349. doi: 10.1111/j.1365-3040.2006.01627.x

Huh G-H, Damsz B, Matsumoto TK, Reddy MP, Rus AM, Ibeas JI, Narasimhan ML, Bressan RA, Hasegawa PM (2002) Salt causes ion disequilibrium-induced programmed cell death in yeast and plants. The Plant Journal, 29, 649-659. doi 10.1046/j.0960-7412.2001.01247.x

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K (2001) Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. Plant J, 27: 325-333. doi https://doi.org/10.1046/j.1365-313x.2001.01096.x

Je J, Chen H, Song C, Lim CO (2014a) Arabidopsis DREB2C modulates ABA biosynthesis during germination. Biochem Biophys Res Commun, 452:91–98. doi 10.1016/j.bbrc.2014.08.052

Je J, Song C, Hwang JE, Chung WS, Lim CO. (2014b) DREB2C acts as a transcriptional activator of the thermo tolerance-related phytocystatin 4 (AtCYS4) gene. Transgenic Res 23: 109-123. doi: 10.1007/s11248-013-9735-2

Ji C, Mao X, Hao J, Wang X, Xue J, Cui H, Li R (2018) Analysis of bZIP transcription factor family and their expressions under salt stress in *Chlamydomonas reinhardtii*. Int J Mol Sci 19:2800. doi: 10.3390/ijms19092800

Jiang YJ, Liang G, Yu DQ (2012) Activated expression of WRKY57 confers drought tolerance in Arabidopsis. Mol Plant 5: 1375–1388. doi: 10.1093/mp/sss080

Jin J, Zhang H, Kong L, Gao G, Luo J (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res, 42:D1182–7. doi: 10.1093/nar/gkt1016

Johnson C, Knight M, Kondo T, Masson P, Sedbrook J, Haley A, Trewavas A (1995) Circadian oscillations of cytosolic and chloroplastic free calcium in plants. Science 269:1863–1865. doi 10.1126/science.7569925

Johnson T, Eswaran JN, Sujatha M (2011) Molecular approaches to improvement of *Jatropha curcas* Linn. as a sustainable energy crop. Plant Cell Rep 30: 1573-1591. doi:10.1007/s00299-011-1083-1.

Julkowska MM, Testerink C (2015) Tuning Plant Signaling and Growth to Survive Salt. Trends in Plant Science, 20(9):586-94. doi: 10.1016/j.tplants.2015.06.008

Kang H-G, Foley RC, Onate-Sanchez L, Lin C, Singh KB (2003) Target genes for OBP3, a Dof transcription factor, include novel basic helix-loop-helix domain proteins inducible by salicylic acid. Plant J 35:362-372. doi 10.1046/j.1365-313X.2003.01812.x

Katsuhara M, Shibasaka M (2000) Cell death and growth recovery of barley after transient salt stress. J. Plant Res. 113, 239–243. doi https://doi.org/10.1007/PL00013934

Kazan K (2006) Negative regulation of defence and stress genes by EAR-motif-containing repressors. Trends in Plant Science 11, 109–112. doi 10.1016/j.tplants.2006.01.004

Khavari-Nejad RA, Bujar M, Attaran E (2006) Evaluation of anthocyanin contents under salinity (NaCl) stress in *Bellis perennis* L. In Khan MA, Weber DJ (eds) Ecophysiology of high salinity tolerant plants, 127–134. doi 10.1007/1-4020-4018-0_8

Kirik V, Simon M, Huelskamp M, Schiefelbein J (2004) The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. Dev Biol, 5;268(2):506-13. doi 10.1016/j.ydbio.2003.12.037

Kooiker M, Airoldi CA, Losa A, Manzotti PS, Finzi L, Kater MM, Colombo L (2005) BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic Arabidopsis gene SEEDSTICK. Plant Cell, 17: 722-729. doi 10.1105/tpc.104.030130

Kreps JA, Kay SA (1997) Coordination of plant metabolism and development by the circadian clock. Plant Cell 9, 1235–1244. doi: 10.1105/tpc.9.7.1235

Krishnaswamy S, Verma S, Rahman MH, Kav NN (2011) Functional characterization of four APETALA2-family genes (RAP2.6, RAP2.6L, DREB19 and DREB26) in Arabidopsis. Plant Mol Biol, 75, 107–127. doi: 10.1007/s11103-010-9711-7

Krishnamurthy P, Mohanty B, Wijaya E, Lee D-Y, Lim T, Lin Q, Xu J, Loh CS. Kumar PP (2017) Transcriptomics analysis of salt stress tolerance in the roots of the mangrove *Avicennia officinalis*. Sci. Rep.7, 10031. doi:10.1038/s41598-017-10730-2

Kunieda T, Mitsud N, Ohme-Takagi M, Takeda S, Aida M, Tasaka M, Kondo M, Nishimura M, Hara-Nishimura I (2008) NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in Arabidopsis. Plant Cell 20: 2631–2642. doi: 10.1105/tpc.108.060160

Kuromori T, Mizoi J, Umezawa T, Yamaguchi-Shinozaki K, Shinozaki K (2014) Drought stress signaling network, in Molecular Biology. The Plant Sciences, v2, 383–409. doi: 10.1007/978-1-4614-7570-5_7

Lam E (2008) Programmed cell death in plants: orchestrating an intrinsic suicide program within walls. Critical Reviews in Plant Sciences, v27, 413-423. doi https://doi.org/10.1080/07352680802467744

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology, 10:R25. doi https://doi.org/10.1186/gb-2009-10-3-r25

Lara P, Oñatesánchez L, Abraham Z, Ferrándiz C, Díaz I, Carbonero P, Vicentecarbajosa J (2003) Synergisticactivation of seed storage protein gene expression in Arabidopsis by ABI3 and two bZIPs related to OPAQUE2. J Biol Chem, 278, 21003–21011. doi 10.1074/jbc.M210538200

Lata C, Prasad M (2011) Role of DREBs in regulation of abiotic stress responses in plants. J Exp Bot, 62, 4731–4748. doi: 10.1093/jxb/err210

Legnaioli T, Cuevas J, Mas P (2009) TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. EMBO J 28:3745–3757. doi: 10.1038/emboj.2009.297

Li L, White MJ, Macrae TH (1999) Transcription factors and their genes in higher plants functional domains, evolution and regulation. Eur J Biochem, 262: 247–257. doi 10.1046/j.1432-1327.1999.00349.x

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNASeq data with or without a reference genome. BMC Bioinformatics, 12: 323. doi: 10.1186/1471-2105-12-323

Li Z, Zhang L, Yu Y, Quan R, Zhang Z, Zhang H, Huang R (2011) The ethylene response factor AtERF11 that is transcriptionally modulated by the bZIP transcription factor HY5 is a crucial repressor for ethylene biosynthesis in Arabidopsis. Plant Journal 68: 88–99. doi: 10.1111/j.1365-313X.2011.04670.x

Li S, Xie Z, Hu C, Zhang J (2015) A review of auxin response factors (ARF) in plants. Front Plant Sci 7:1–14. doi:10.3389/fpls.2016.00047

Li H, Huang W, Liu ZW, Wang YX, Zhuang J (2016) Transcriptome-based analysis of Dof family transcription factors and their responses to abiotic stress in tea plant (*Camellia sinensis*). Int J Genomics, 5614142, 15p. doi 10.1155/2016/5614142

Lin CC, Kao CH (1995) NaCl stress in rice seedling: Starch mobilization and the influence of gibberellic acid on seedling growth. Bot. Bull. Academia Sinica, 36: 169-173.

Liotenberg S, North H, Marion-Poll A (1999) Molecular biology and regulation of abscisic acid biosynthesis in plants. Plant Physiol Biochem, 37, 341-350. doi https://doi.org/10.1016/S0981-9428(99)80040-0

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell 10:1391–1406. doi 10.1105/tpc.10.8.1391

Liu X, Hou X (2018) Antagonistic regulation of ABA and GA in metabolism and signaling pathways. Front Plant Sci, 9:251. doi: 10.3389/fpls.2018.00251

Liu C, Mao B, Ou S, Wang W, Liu L, Wu Y, Chu C, Wang X (2014) OsbZIP71, a bZIP transcription factor confers salinity and drought tolerance in rice. Plant Mol Biol, v84, 19–36. doi: 10.1007/s11103-013-0115-3

Liu Q, Luo L, Zheng L (2018) Lignins: Biosynthesis and Biological Functions in Plants. International journal of molecular sciences, 24;19(2). doi https://doi.org/10.3390/ijms19020335

Lozano-Isla F, Campos MLO, Endres L, Bezerra-Neto E (2018) Effects of seed storage time and salt stress on the germination of *Jatropha curcas* L. Industrial Crops & Products, 118, 214-224. doi https://doi.org/10.1016/j.indcrop.2018.03.052

Lopez-Molina L, Mongrand S, Chua NH (2001) A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the AB15 transcription factor in Arabidopsis. Proc Natl Acad Sci, 98: 4782–4787. doi 10.1073/pnas.081594298

Lopez-Molina L., Mongrand S, McLachlin DT, Chait BT, Chua NH (2002) ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. Plant J, 32: 317–328. doi https://doi.org/10.1046/j.1365-313X.2002.01430.x

Lu CA, Ho TH, Ho SL, Yu, SM (2002) Three novel MYB proteins with one DNA binding repeat mediate sugar and hormone regulation of alpha-amylase gene expression. Plant Cell 14, 1963–1980. doi 10.1105/tpc.001735

Lu D, Wang T, Persson S, Mueller-Roeber B, Schippers JH (2014) Transcriptional control of ROS homeostasis by KUODA1 regulates cell expansion during leaf development. Nat Commun, 5, 3767. doi 10.1038/ncomms4767

Ma J, Li MY, Wang F, Tang J, Xiong AS (2015) Genome-wide analysis of Dof family transcription factors and their responses to abiotic stresses in Chinese cabbage. BMC Genomics, 16(1):33. doi: 10.1186/s12864-015-1242-9

Ma R, Xu S, Zhao Y, Xia B, Wang R (2016) Selection and validation of appropriate reference genes for quantitative real-time PCR analysis of gene expression in Lycoris aurea. Front Plant Sci, 7, 536. doi: 10.3389/fpls.2016.00536

Mahajan S, Tuteja N (2005) Cold, salinity and drought stresses: An overview. Arch Biochem Biophys 444: 139–158. doi:10.1016/j.abb. 2005.10.018. PubMed: 16309626

Mao H, Yu L, Li Z, Liu H, Han R (2016) Molecular evolution and gene expression differences within the HD-Zip Transcription Factor Family of *Zea mays* L. Genetica, 144, 243–257. doi: 10.1007/s10709-016-9896-z

Mäser P, Eckelman B, Vaidyanathan R (2002) Altered shoot/root Na^+ distribution and bifurcating salt sensitivity in Arabidopsis by genetic disruption of the Na^+ transporter AtHKT1. FEBS Lett. 531, 157–161. doi https://doi.org/10.1016/S0014-5793(02)03488-9

Massonneau A, Condamine P, Wisniewski J-P, Zivy M, Rogowsky, PM (2005) Maize cystatins respond to developmental cues, cold stress and drought. Biochim Biophys Acta 1729:186–199. doi 10.1016/j.bbaexp.2005.05.004

McClung, CR (2001) Circadian rhythms in plants. Annu Rev Plant Physiol Plant Mol Biol, 52: 139-162. doi https://doi.org/10.1105/tpc.106.040980

McGrath KC, Dombrecht B, Manners JM, Schenk PM, Edgar CI, Maclean DJ, Scheible WR, Udvardi MK, Kazan K (2005) Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. Plant Physiol 139: 949–959. doi 10.1104/pp.105.068544

Mei Y, Song S (2008) Early morphological and physiological events occurring during germination of maize seeds. Agric Sci China, 7: 950-957. doi https://doi.org/10.1016/S1671-2927(08)60134-0

Meir S, Philosoph-Hadas S, Sundaresan S, Selvaraj KSV, Burd S (2010) Microarray analysis of the abscission-related transcriptome in the tomato flower abscission zone in response to auxin depletion. Plant Physiol 154: 1929–1956. doi: 10.1104/pp.110.160697

Mengiste T, Chen X, Salmeron J, Dietrich R (2003) The Botrytis Susceptible1 gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. Plant Cell, 15(11):2551-65. doi 10.1105/tpc.014167

Mishra KB, Iannacone R, Petrozza A, Mishra A, Armentano N, La Vecchia G, Trtílek M, Cellini F, Nedbal L (2012) Engineered drought tolerance in tomato plants is reflected in chlorophyll fluorescence emission. Plant Sci, 182: 79-86. doi: 10.1016/j.plantsci.2011.03.022

Mitsuda N, Ikeda M, Tanaka S, Takiguchi Y, Kondou Y, Yoshizumi T, Fujita M, Shinozaki K, Matsui M, Ohme-Takagi M (2010) Efficient yeast one-/two-hybrid screening using a library composed only of transcription factors in *Arabidopsis thaliana*. Plant Cell Physiol 51: 2145–2151. doi: 10.1093/pcp/pcq161

Moller IS, Gilliham M, Jha D, Mayo GM, Roy SJ, Coates JC, Haseloff J, Tester M (2009) Shoot Na^+ exclusion and increased salinity tolerance engineered by cell type-specific alteration of Na^+ transport in Arabidopsis. Plant Cell, 21, 2163–2178. doi: 10.1105/tpc.108.064568

Mondini L, Nachit M, Porceddu E, Pagnotta MA (2012) Identification of SNP mutations in DREB1, HKT1, and WRKY1 genes involved in drought and salt stress tolerance in durum wheat (*Triticum turgidum* L. var *durum*). OMICS 16 178–187. doi 10.1089/omi.2011.0081

Moyano E, Martinez-Garcia JF, Martin C (1996) Apparent redundancy in Myb gene function provides gearing for the control of flavonoid biosynthesis in Antirrhinum flowers. Plant Cell, 8, 1519–1532. doi https://doi.org/10.1105/tpc.8.9.1519

Mu RL, Cao YR, Liu YF, Lei G, Zou HF, Liao Y, Wang HW, Zhang WK, Ma B, Du JZ, Yuan M, Zhang JS, Chen SY (2009) An R2R3-type transcription factor gene AtMYB59 regulates root growth and cell cycle progression in Arabidopsis. Cell Res, 19: 1291-1304. doi: 10.1038/cr.2009.83

Munns, R (2002) Comparative physiology of salt and water stress. Plant Cell Environ, 25:239-250. https://doi.org/10.1046/j.0016-8025.2001.00808.x

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu Rev Plant Biol 59:651–681. doi: 10.1146/annurev.arplant.59.032607.092911

Nelson DE, Repetti PP, Adams TR, Creelman RA, Wu J, Warner DC (2007) Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres. Proc Natl Acad Sci USA, 104, 16450–16455. doi: 10.1073/pnas.0707193104

Nowak K, Gaj MD (2016) Stress-related function of bHLH109 in somatic embryo induction in Arabidopsis. Journal of Plant Physiology, 193, 119–126. doi: 10.1016/j.jplph.2016.02.012

Ohtani M, Nishikubo N, Xu B, Yamaguchi M, Mitsuda N, Goue N, Shi F, Ohme-Takagi M, Demura T (2011) A NAC domain protein family contributing to the regulation of wood formation in poplar. Plant J 67: 499–512. doi: 10.1111/j.1365-313X.2011.04614.x

Olsson AS, Engström P, Soderman E (2004) The Homeobox Genes Athb12 and Athb7 Encode Potential Regulators of Growth in Response to Water Deficit in Arabidopsis. Plant Mol Biol, 55: 663-677. doi 10.1007/s11103-004-1581-4

Openshaw L (2000) A review of Jatropha curcas: an oil plant of unfulfilled promise. Biomass Bioenergy 19: 1-15. doi:10.1016/S0961-9534(00)00019-2

Othman YG, Al-Karaki AR, Al-Horani A (2006) Variation in germination and ion uptake in barley genotypes under salinity conditions. World J Agric Sci, 2: 11-15.

Pan Y, Seymour GB, Lu C, Hu Z, Chen X, Chen G (2012) An ethylene response factor (ERF5) promoting adaptation to drought and salt tolerance in tomato. Plant Cell Rep 31(2): 349-60. doi: 10.1007/s00299-011-1170-3

Papdi C, Pérez-Salamó I, Joseph MP, Giuntoli B, Bögre L, Koncz C, Szabados L (2015) The low oxygen, oxidative and osmotic stress responses synergistically act through the ethylene response factor VII genes RAP2.12, RAP2.2 and RAP2.3. Plant J, 82: 772–784. doi: 10.1111/tpj.12848

Park DH, Lim PO, Kim JS, Cho DS, Hong SH, Nam HG (2003) The Arabidopsis COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. Plant J, 34:161-171. doi https://doi.org/10.1046/j.1365-313X.2003.01710.x

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci. 9(11):534-40. doi 10.1016/j.tplants.2004.09.002

Pei H, Ma N, Tian J, Luo J, Chen J, Li J, Zheng Y, Chen X, Fei Z, Gao J (2013) An NAC transcription factor controls ethylene-regulated cell expansion in flower petals. Plant Physiol, v163(2):775-91. doi: 10.1104/pp.113.223388

Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B (2010) PlnTFDB: updated content and new features of the plant transcription factor database, Nucleic Acids Res, v.38 p.D822-D827. doi: 10.1093/nar/gkp805

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30:e36. doi: 10.1093/nar/30.9.e36

Phuong ND, Hoi PX (2015) Isolation and characterization of a OsRap2.4A transcription factor and its expression in Arabidopsis for enhancing high salt and drought tolerance. Current Science, VOL. 108, n.1,10. doi http://dx.doi.org/10.1590/1678-4685-gmb-2016-0052

Popova LP, Stoinova ZG, Maslenkova LT (1995) Involvement of abscisic acid in photosynthetic process in *Hordeum vulgare* L. during salinity stress. Journal of Plant Growth Regulation, v14, n4, p211–218. doi 10.1007/BF00204914

Rai AC, Singh M, Shah K (2012) Effect of water withdrawal on formation of free radical, proline accumulation and activities of antioxidant enzymes in ZAT12-transformed transgenic tomato plants. Plant Physiology and Biochemistry 61, 108–114. doi: 10.1016/j.plaphy.2012.09.010

Razzaque S, Haque T, Elias SM, Rahman MS, Biswas S, Schwartz S, Ismail AM, Walia H, Juenger TE, Seraj ZI (2017) Reproductive stage physiological and transcriptional responses to salinity stress in reciprocal populations derived from tolerant (Horkuch) and susceptible (IR29) rice. Scientific Reports 7, 46138. doi 10.1038/srep46138

Reyes JC, Muro-Pastor MI, Florencio FJ (2004) The GATA family of transcription factors in Arabidopsis and rice. Plant Physiol, 134(4):1718–32. doi: 10.1104/pp.103.037788

Riechmann J, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290, 2105–2110. doi 10.1126/science.290.5499.2105

Rizhsky L, Davletova S, Liang H, Mittler R (2004) The zinc finger protein Zat12 is required for cytosolic ascorbate peroxidase 1 expression during oxidative stress in Arabidopsis . J Biol Chem 279, 11736–11743. doi 10.1074/jbc.M313350200

Robinson MD, McCarthy DJ, Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139-140. doi: 10.1093/bioinformatics/btp616

Rong W, Qi L, Wang A, Ye X, Du L, Liang H, Xin Z, Zhang Z (2014) The ERF transcription factor TaERF3 promotes tolerance to salt and drought stresses in wheat. Plant Biotechnol J, 12: 468–479. doi:10.1111/pbi.12153

Rozen S, Skaletsky HJ (2000) primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (eds, Krawetz S, Misener S), Humana Press, Totowa, New Jersey, 365–386. doi 10.1385/1-59259-192-2:365

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci 15: 247–258. doi: 10.1016/j.tplants.2010.02.006

Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD (1996) Systemic acquired resistance. Plant Cell, 8, 1809-1819. doi 10.1105/tpc.8.10.1809

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501–506. doi 10.1038/ng1543

Schuettengruber B, Cavalli G (2013) Polycomb domain formation depends on short and long distance regulatory cues. PLoS ONE 8(2). doi: 10.1371/journal.pone.0056531

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31 279–292. doi https://doi.org/10.1046/j.1365-313X.2002.01359.x

Seo PJ, Xiang F, Qiao M, Park JY, Lee YN (2009) The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in Arabidopsis. Plant Physiol 151: 275-289. doi https://doi.org/10.1104/pp.109.144220

SeoY, Park JB, Cho YJ, Jung C, Seo H, Park SK, Nahm B, Song J (2010) Overexpression of the ethylene-responsive factor gene BrERF4 from *Brassica rapa* increases tolerance to salt and drought in Arabidopsis plants. Molecules and Cells 30: 271– 277. doi: 10.1007/s10059-010-0114-z

Seo PJ, Mas P (2015) STRESSing the role of the plant circadian clock. Trends Plant Sci 20:230–237. doi :https://doi.org/10.1016/j.tplants.2015.01.001

Shabala S (2009) Salinity and programmed cell death: Unravelling mechanisms for ion specific signalling. J Exp Bot, 60, 709–711. doi 10.1093/jxb/erp013

Shameer K, Ambika S, Varghese SM, Karaba N, Udayakumar M, Sowdhamini R (2009) STIFDB — Arabidopsis stress responsive transcription factor dataBase. International Journal of Plant Genomics, vol. 2009, Article ID 583429, 8p. doi:10.1155/2009/583429

Shin D, Moon SJ, Han S, Kim BG, Park SR (2011) Expression of StMYB1R-1, a novel potato single MYB-like domain transcription factor, increases drought tolerance. Plant Physiol 155: 421–432. doi: 10.1104/pp.110.163634

Siddiqui ZS, Khan MA (2011) The role of enzyme amylase in two germinating seed morphs of *Halopyrum mucronatum* (L.) Stapf. in saline and non-saline environment. Acta Physiologiae Plantarum, 33: 1185-1197. doi 10.1007/s11738-010-0646-z

Silveira AB, Gauer L, Tomaz JP, Cardoso PR, Carmello GS, Vincentz M (2007) The Arabidopsis AtbZIP9 protein fused to the VP16 transcriptional activation domain alters leaf and vascular development. Plant Sci,172, 1148–1156. doi https://doi.org/10.1016/j.plantsci.2007.03.003

Singh KB, Foley RC, Onate-Sanchez L (2002) Transcription factors in plant defense and stress responses. Current Opinion in Plant Biology, v5, n5, pp430–436. doi https://doi.org/10.1016/S1369-5266(02)00289-3

Sjögren LL, Clarke AK (2011) Assembly of the chloroplast ATP-dependent Clp protease in Arabidopsis is regulated by the ClpT accessory proteins. Plant Cell 23: 322–332. doi: 10.1105/tpc.110.082321

Skirycz A, Radziejwoski A, Busch W, Hannah MA, Czeszejko J, Kwaśniewski M, Zanor MI, Lohmann JU, De Veylder L, Witt I, Mueller-Roeber B (2008) The DOF transcription factor OBP1 is involved in cell cycle regulation in *Arabidopsis thaliana*. Plant J 56 (5):779-92. doi: 10.1111/j.1365-313X.2008.03641.x

Söderman E, Mattsson J, Engström P (1996) The Arabidopsis homeobox gene ATHB-7 is induced by water deficit and by abscisic acid. Plant Journal 10, 375–381. doi 10.1046/j.1365-313X.1996.10020375.x

Song C, Je J, Hong JK, Lim CO (2014) Ectopic expression of an Arabidopsis dehydration-responsive element-binding factor DREB2C improves salt stress tolerance in crucifers. Plant Cell Rep, 33 (8): 1239-1254. doi 10.1007/s00299-014-1612-9.

Sudhir P, Murthy SDS (2004) Effects of salt stress on basic processes of photosynthesis. Photosynthetica 42:481–486. doi 10.1007/S11099-005-0001-6

Takasaki H, Maruyama K, Takahashi F, Fujita M, Yoshida T, Nakashima K, Myouga F, Toyooka K, Yamaguchi-Shinozaki K, Shinozaki K (2015) SNAC-As, stress-responsive NAC transcription factors, mediate ABA-inducible leaf senescence. Plant J 84: 1114–1123. doi: 10.1111/tpj.13067

Takase T, Ishikawa H, Murakami H, Kikuchi J, Sato-Nara K, Suzuki H (2011) The Circadian clock modulates water dynamics and aquaporin expression in Arabidopsis roots. Plant Cell Physiol 52:373–383. doi: 10.1093/pcp/pcq198

Tang Y, Liu K, Zhang J, Li X, Xu K, Zhang Y, Qi J, Yu D, Wang J, Li C (2017) JcDREB2, a physic nut AP2/ERF gene, alters plant growth and salinity stress responses in transgenic rice. Frontiers in Plant Science 8: 306. doi: 10.3389/fpls.2017.00306

Tang YH, Qin SS, Guo YL, Chen YB, Wu PZ, Chen YP, Li MR, Jiang HW, Wu GJ (2016) Genome-wide analysis of the AP2/ERF gene family in physic nut and

overexpression of the JcERF011 gene in rice increased its sensitivity to salinity stress. PLoS ONE 4;11(3). doi https://doi.org/10.1371/journal.pone.0150879

Thornthwaite CW, Mather JR (1955)The water balance. Centerton NJ: Drexel Institute of Technology - Laboratory of Climatology 104p Publications in Climatology vol VII, n.1.

Thurow C, Schiermeyer AS, Butterbrodt T, Nickolov K, Gatz C (2005) Tobacco bZIP transcription factorTGA2.2 and related factor TGA2.1 have distinct roles in plant defense responses and plant development. Plant J. 44, 100–113. doi 10.1111/j.1365-313X.2005.02513.x

Tran LSP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. The Plant Cell vol. 16, 2481-2498. doi https://doi.org/10.1105/tpc.104.022699

Trujillo LE, Sotolongo M, Menéndez C, Ochogavía ME, Coll Y, Hernández I, Borrás-Hidalgo O, Thomma BPHJ, Vera P, Hernández L (2008) SodERF3, a novel sugarcane ethylene responsive factor (ERF), enhances salt and drought tolerance when overexpressed in tobacco plants. Plant Cell Physiol 49: 512–525. doi: 10.1093/pcp/pcn025

Ulker B, Somssich IE (2004) WRKY transcription factors, From DNA binding towards biological function. Curr Opin Plant Biol 7: 491–498. doi: 10.1016/j.pbi.2004.07.012

Urano K, Maruyama K, Jikumaru Y, Kamiya Y, Yamaguchi-Shinozaki K, Shinozaki K (2017) Analysis of plant hormone profiles in response to moderate dehydration stress. Plant J 90(1):17–36. doi: 10.1111/tpj.13460

Vergnolle C, Vaultier MN, Taconnat L, Renou JP, Kader JC, Zachowski A, Ruelland E (2005) The cold-induced early activation of phospholipase C and D pathways determines the response of two distinct clusters of genes in Arabidopsis cell suspensions. Plant Physiol 139: 1217–1233. https://doi.org/10.1104/pp.105.068171

Wang Z-Y, Tobin EM (1998) Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell 93:1207–1217. doi.org/10.1016/S0092-8674(00)81464-6

Wang Z, Li P-H, Fredricksen M, Gong Z-Z, Kim C, Zhang C, Bohnert HJ, Zhu J-K, Bressan RA, Hasegawa PM (2004) Expressed sequence tags from *Thellungiella halophila*, a new model to study plant salt-tolerance. Plant Science 166, 609–616. doi.org/10.1016/j.plantsci.2003.10.030

Wang X, Haberer G, Mayer K (2009) Discovery of cis-elements between sorghum and rice using co-expression and evolutionary conservation. BMC Genomics v10, p1471-2164. doi: 10.1186/1471-2164-10-284

Wen C-L, Cheng Q, Zhao L, Mao A, Yang J, Yu S (2016) Identification and characterisation of Dof transcription factors in the cucumber genome. Sci. Rep. 6:23072. doi: 10.1038/srep23072

Wu QX, Chen J, Shi YP (2010) RP-HPLC and NMR study of antioxidant flavonoids in extract from *Gentiana piasezkii*. J Anal Chem 65: 298–304. dio: 10.1134/S1061934810030159

Wu Z, Xu X, Xiong W, Wu P, Chen Y, Li M (2015) Genome-wide analysis of the NAC gene family in physic nut (*Jatropha curcas* L.). PLoS ONE 10:e0131890. doi: 10.1371/journal.pone.0131890

Wu H, Fu B, Sun P, Xiao C, Liu JH (2016) A NAC transcription factor represses putrescine biosynthesis and affects drought tolerance. Plant Physiol 172, 1532–1547. doi.org/10.1104/pp.16.01096

Xie XB, Li S, Zhang RF, Zhao J, Chen YC, Zhao Q (2012) The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. Plant Cell Environ 35(11):1884–97. doi: 10.1111/j.1365-3040.2012.02523.x

Xiong W, Xu X, Zhang L, Wu P, Chen Y, Li M, Jiang H, Wu G (2013) Genome-wide analysis of the WRKY gene family in physic nut (*Jatropha curcas* L.). Gene, 524, 124–132. doi:10.1016/j.gene.2013.04.047

Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of the expression of a desiccation-responsive rd29 gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. Mol Gen Genet, 236, pp 331-340.

Yaneff A, Vitali V, Amodeo G (2015) PIP1 aquaporins: intrinsic water channels or PIP2 aquaporin modulators? FEBS Lett 589 3508–3515. doi.org/10.1016/j.febslet.2015.10.018

Yang CY, Hsu FC, Li JP, Wang NN, Shih MC (2011) The AP2/ERF transcription factor AtERF73/HRE1 modulates ethylene responses during hypoxia in Arabidopsis. Plant Physiology 156: 202–212 pmid:21398256. doi: 10.1104/pp.111.172486

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G (2006) The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol 60 107–124 101007/s11103-005-2910-y

Ye G, Ma Y, Feng Z, Zhang X (2018) Transcriptomic analysis of drought stress responses of sea buckthorn (*Hippophae rhamnoides* subsp. *sinensis*) by RNA-Seq. PloS one 13(8):e0202213. https://doi.org/10.1371/journal.pone.0202213

Ying S, Zhang DF, Fu J, Shi YS, Song YC, Wang TY, Li Y (2012) Cloning and characterization of a maize bZIP transcription factor, ZmbZIP72, confers drought and salt tolerance in transgenic Arabidopsis. Planta 235(2): 253-266. doi: 10.1007/s00425-011-1496-7

Yin Z, Li Y, Zhu W, Fu X, Han X, Wang J, Lin H, Ye W (2018) Identification, Characterization, and Expression Patterns of TCP Genes and microRNA319 in Cotton. Int J Mol Sci, 19 (11), 3655. https://doi.org/10.3390/ijms19113655

Yu DQ, Qiu YP (2009) Over-expression of the stress-induced OsWRKY45 enhances disease resistance and drought tolerance in Arabidopsis. Environ Exp Bot, 65 pp. 35-47. https://doi.org/10.1016/j.envexpbot.2008.07.002

Yue H, Shu D, Wang M, Xing G, Zhan H, Du X, Song W, Nie X (2018) Genome-Wide Identification and Expression Analysis of the HD-Zip Gene Family in Wheat (*Triticum aestivum* L.). Genes, 1; 9(2). doi: 10.3390/genes9020070

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, Luo J. (2011) PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. Nucleic Acids Res 39, D1114–D1117. doi: 10.1093/nar/gkq1141

Zhang Z, Wang, J, Zhang, R, Huang R (2012) The ethylene response factorAtERF98 enhances tolerance to salt through the transcriptional activation of ascorbic acid synthesis in Arabidopsis. Plant J 71, 273–287. doi: 10.1111/j.1365-313X.2012.04996.x

Zhang GW, Xu SC, Hu QZ, Mao WH, Gong YM (2014a) Putrescine plays a positive role in salt-tolerance mechanisms by reducing oxidative damage in roots of vegetable soybean. J Integr Agric 13: 349–357.

Zhang H, Liu Y, Wen F, Yao D, Wang L, Guo J (2014b) A novel rice C2H2-type zinc finger protein, ZFP36, is a key player involved in abscisic acid-induced antioxidant defence and oxidative stress tolerance in rice. J Exp Bot 65(20):5795-809 doi: 10.1093/jxb/eru313

Zhao T, Xia H, Liu J, Ma F (2014) The gene family of dehydration responsive element binding transcription factors in grape (*Vitis vinifera*): genome-wide identification and analysis, expression profiles, and involvement in abiotic stress resistance. Mol Biol Rep 41 1577–1590. doi: 10.1007/s11033-013-3004-6.

Zhou M, Li D, Li Z, Hu Q, Yang C, Zhu L, Luo H (2013) Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass. Plant Physiol 161 1375–1391. doi: 10.1104/pp.112.208702.

Zhu M, Chen G, Zhang J, Zhang Y, Xie Q, Zhao Z (2014) The abiotic stress-responsive NAC-type transcription factor SlNAC4 regulates salt and drought tolerance and stress-related genes in tomato (*Solanum lycopersicum*). Plant Cell Rep 33, 1851–1863. doi: 10.1007/s00299-014-1662-z

Zhou C, Chen Y, Wu Z, Lu W, Han J, Wu P, Li M, Jiang H, Wu G (2015) Genome-wide analysis of the MYB gene family in physic nut (*Jatropha curcas* L.) Gene 572:63–71. doi: 10.1016/j.gene.2015.06.072

4.3 Dominant position of cis-acting regulatory elements in promoters of physic nut (*Jatropha curcas* L.) transcription factors genes

George André de Lima Cabral[1], Marislane Carvalho Paz de Souza[2], Manassés Daniel da Silva[3], José Ribamar da Costa Ferreira Neto[4], Éderson Akio Kido[5]

Email addresses:
GALC: geolimma@gmail.com
MCPS: maris.carvalho@hotmail.com
MDS: manassesdaniel@gmail.com
JRCFN: netocostaferreira@gmail.com
EAK: ederson.kido@gmail.br

ABSTRACT

The physic nut (*Jatropha curcas* L.) is an oilseed crop that stands out among the others for its potential to produce biofuel along with its relevant capacity of growth in a wide range of soil and climatic conditions, sometimes even bearing abiotic stresses. During the plant development, transcription factors (TF), responds to complex signaling pathways. Therefore, according to the physiological processes involved, TFs modulate the transcription initiation rate of selected genes, by contacting the *cis*-acting regulatory elements (CARE) presented in the promoters of these genes. Thus, promoters and TFs are crucial components for specificity and intensity of gene expression regulation. In this study, CAREs were mapped by *in silico* analysis, focusing on the regulatory region (1.5 kb) *upstream* the TSS (transcription start site) of 106 potential TF genes previously identified and classified into 29 families, of *J. curcas* ESTs available in a public library (NCBI). Based on the CARE motifs, they were quantified, and their analyzed distribution along the promoter regions was subdivided into five segments (300 pb each). The prospection over the promoters detected a total of 20,158 motifs comprising 209 unique CAREs, most of them presenting 60-70 motifs. The multiple CARE motifs identified in some promoters of TF genes revealed possible gene regulation controlled by many TFs, involving cooperative actions of TFs and the target gene. Also, it was

observed auto-regulation of TF-encoding gene with the translated TF protein binding to the detected CARE motif. Based on the six TF families with the most prevalent motifs, some CAREs, were distributed along five prospected promoter segments, especially those associated with a constitutive developmental function. On the other hand, those related to stress responses and ABA stimulus were not uniformly distributed, showing a motif dominant position in some of those segments. Since the amount of CARE and motif location are important elements to be considered for modeling efficient synthetic promoters, these results may help *J. curcas* breeding programs, in developing tools, aiming to increase transcription efficiency of genes, responsive to abiotic stresses.

KEYWORDS: *cis*-acting regulatory element, transcription factors, genomic, *in silico* promoter analysis, Euphorbiaceae.

INTRODUCTION

The physic nut (*Jatropha curcas* L.) is an oilseed crop from the Euphorbiaceae family, which stands out among others due to its potential for the biofuel production (Nahar and Sunny, 2011). Since *J. curcas* presents relative rusticity, pest resistance, and drought tolerance, the geographical distribution is quite broad, adapting to a wide range of climatic conditions in the tropical and equatorial regions, also tolerating not very fertile soils (Arruda et al., 2004). During the crop cultivation, plants may be affected by abiotic stresses, of which drought and salinity are the most damaging, generally provoking economic lost. Plants when trying to tolerate these stresses modulate their transcriptomes activating or repressing genes. In this process, the transcription factors (TFs) are the primary proteins regulating gene expression. TFs have two primary domains, the activation/ repression domain, and the DNA binding domain, connecting specific sequences of the target genes to the RNA polymerase II (Zhang et al., 2016). Both domains operate together, through complex signaling pathways, responding to internal and external stimuli, modulating the transcription initiation rate of target genes involved in different physiological and biochemical processes (Franco-Zorrilla et al., 2014).

A substantial fraction of the eukaryotic genomes comprised TF genes classified into TF families. In *Arabidopsis thaliana* about 1,700 TFs are known, representing more than 5% of the estimated total genes (Riechmann et al., 2000). Some of those TF families, such as AP2/ERF, bHLH, MYB, and MADS, comprise more than 100 members each. The TF influence on gene transcription depends on their structural features, but its classification into families is strictly related to the action mechanism, which reflects the DNA binding domain (Parenicova et al., 2003). The TF affinity/ selectivity is influenced by the DNA binding domains secondary structures, contacting bases of the *cis*-acting regulatory elements (CAREs) in the promoters of the target genes, providing essential conditions for gene expression (Brivanlou & Darnell, 2002). The promoter is one of the most critical components of the regulation, specificity, and intensity of expression of native and transgenic genes in plants (Hernandez-Garcia & Finer, 2014).

The gene expression pattern may be influenced not only by the presence or absence of the CAREs individually but mainly by the combinatorial control between the TFs (Pilpel et al., 2001). A single TF rarely controls the expression of a gene; however,

precise combinations of TFs are essential for gene expression in higher organisms (Bhattacharjee et al., 2013). Thus, TFs and CAREs are an integral part of the regulation of gene expression, and both are essential for interpreting and modeling cellular responses to perceived stimuli (Wang et al., 2009), being involved in almost all aspects of cellular activity (Xiong et al., 2005). With the increasing advances in obtaining genomic sequences from plants and the development of bioinformatics tools, much information is now available for selection and identification of new plant promoters and their regulatory components (Zhang et al., 2015). The ESTs (*Expressed Sequence Tag*) is one of the most common sequences available in biomolecule public databases (Reverter et al., 2004). The analysis of expressed gene sequences by bioinformatics tools has been a valuable resource for the functional investigation of target genes and its supposed regulation (Won et al., 2009). Also, there are some databases with tools covering plant TFs and CAREs, such as PLACE (Higo et al., 1999), PlantCARE (Lescot et al., 2002), PlantTFDB (Perez-Rodriguez et al., 2009), iTAK (Zheng et al., 2016), GTRD (Yevshin et al., 2017), among others. Trying to understand the CAREs related to *J. curcas* TF genes, the present work sought to identify those genes derived from available public ESTs (*Genbank*), mapped them into genomic sequences, and prospected the CARE motifs over segments covering potential promoter regions.

METHODS

The *J. curcas* biomolecule sequences, the EST annotation process, and the Gene Ontology analyzes

J. curcas sequences (ESTs and genomics) were downloaded from The National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) and the *Jatropha Genome Database* (TFp://TFp.kazusa.or.jp/pub /jatropha/). The downloaded *J.curcas* ESTs, trying to avoid redundancies, were clustered with the online tool EGassembler (http://www.genome.jp/tools/egassembler/; Masoudi-Nejad et al., 2006), applying the identity *cut-off* of 0.95. The resulted contigs and singlets were annotated, providing the *J. curcas* TF identification, based on BLASTx similarities (e-value < e^{-10}) with the UniProtKB/Swiss-Prot cured protein sequences from the UniProt Knowledgebase (http://www.uniprot.org/).

The gene ontology (GO) analyzes, and the functional enrichment of GO terms (Threshold: *p-value* ≤ 0.01) was performed by the online tool PlantRegMap

(http://plantregmap.cbi.pku.edu.cn; Jin et al., 2015) applying the UniProt IDs of the proteins related to the identified *J. curcas* TFs. From the enriched GO terms, hierarchical clustering was carried out, with summarization and removal of GO terms redundancies, applying the REViGO algorithm (http://revigo.irb.hr/; Supek et al., 2011).

The promoter regions of *J. curcas* TFs genes and the CARE prospection

The contigs e singlets previously identified were aligned via BLASTx (e-value < e^{-10}) with *J. curcas* protein sequences from the RefSeq database (NCBI). The respective accession numbers related to the RefSeq TF sequences were identified in the *J. curcas* genome (organism txid180498; https://www.ncbi.nlm.nih.gov/projects/sviewer/) using the NCBI Sequence Viewer tool, aiming to map the TSS (transcriptional start site) of each predicted TF gene. The region delimited up to -1500 bp *upstream* (5' end) of the TSS was considered the promoter region of each gene, where the CAREs were prospected applying the NewPLACE tool (http://www.dna.affrc.go.jp/PLACE/; Higo et al., 1999). The potential promoters were partitioned into five segments (300 bp each) to facilitate the CARE prospection. The promoter region of eight contitutive genes (GC) [*apx* (ascorbate peroxidase; Park et al., 2010a), *pgd1* (phosphogluconate dehydrogenase decarboxylating 1; Park et al., 2010a), *ubia* (polyubiquitin-A; Cornejo et al., 1993), *ubib* (polyubiquitin-B; Wang et al., 2003), *act2* (actin-2; He et al., 2009), *act7* (actin-7; McElroy et al., 1991), *tuba2* (tubulin alpha-2 chain; Jeon et al., 2000) and *cc1-2* (cytochrome c1-2; Jang et al., 2002)] were analized for comparision with the distribution of CAREs found in promoters of TFs. The identified CAREs were analyzed considering the promoter segments, and the CAREs shared by these sets determined using the Venny tool (http://bioinfogp.cnb.csic.es/tools/venny/).

RESULTS

The *J. curcas* TF genes

The *J. curcas* ESTs (46,874) from the NCBI database comprised 19,959 unique sequences (6,121 contigs and 13,838 singletons). From these sequences, 335 were similar in BLASTx analysis (e-value < e^{-10}) with cured TF proteins from the UniprotSP database, and 222 of them were positioned by similarities (BLASTn, e-value < e^{-20}) in the *J. curcas* genome from the *Jatropha Genome Database*. The manual curation allowed classifying 106 putative TFs genes into 29 TF families (Table 1). The TF

families and the identified members (respective EST, gene and protein IDs) are presented in Table 1. The represented TF families and respective numbers of members were: AP2/ERF (16), MYB (11), bHLH (nine), HD-Zip (eight), GATA (six), NFY (six), bZIP (five), TCP (five), HSF (four), LAV (four), TAF (three), NAC (three), WRKY (three), CAMTA (two), GBF (two), GRF (two), MADS-box (two), SWI/SNF (two), Trihelix (two), ARF (one), C2H2 (one), CNOT2/3/5 (one), GRAS (one), GroTLE (one), GTF (one), SEU (one), SHH2 (one), SPT4 (one), and WHIRLY (one). The relative distribution of the TF families is shown in the Figure 1.

Insert Table 1

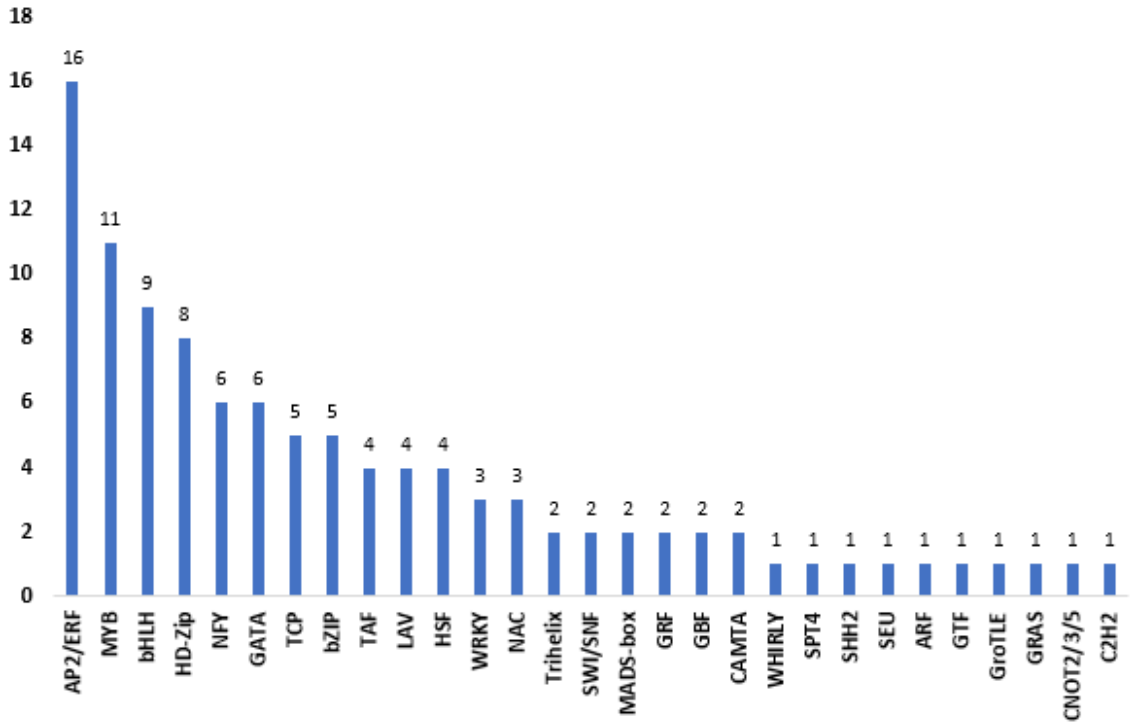


Figure 1. Transcription factors family distribution based on *J. curcas* EST annotations.

A general functional view of these putative TF proteins encoded by the *J. curcas* ESTs using a GO enrichment analysis (REViGO tool; Supek et al., 2011) pointed 138 terms distributed in the main categories: Biological Process (114), Molecular Function (15) and Cell Component (nine). From the 114 enriched GO terms related to Biological Process, 46 (40%) represented 64 - 71 putative TF genes (Table S1). The enriched Biological Processes GO terms showing the most significant *p-values* (*p-value* <0.01) included (Figure 2): *heterocycle metabolic process*, *response to endogenous stimulus*,

organic cyclic compound metabolic process, *biosynthetic process*, *biological regulation*, *developmental process*, *nitrogen compound metabolic process*, *cellular process*, *positive regulation of nucleic acid-templated transcription*, *response to stimulus*, *primary metabolic process*, *organic substance metabolic process*, *multicellular organismal process* and *anatomical structure development*. According to the Figure 2, most of the enriched GO terms clustered with greater semantic similarities, and *regulation* (36.8%) and *response* (22.8%) were the central keywords, reinforcing the regulatory role of TFs in cellular activity processes, and the TF responsiveness to environmental stimuli.

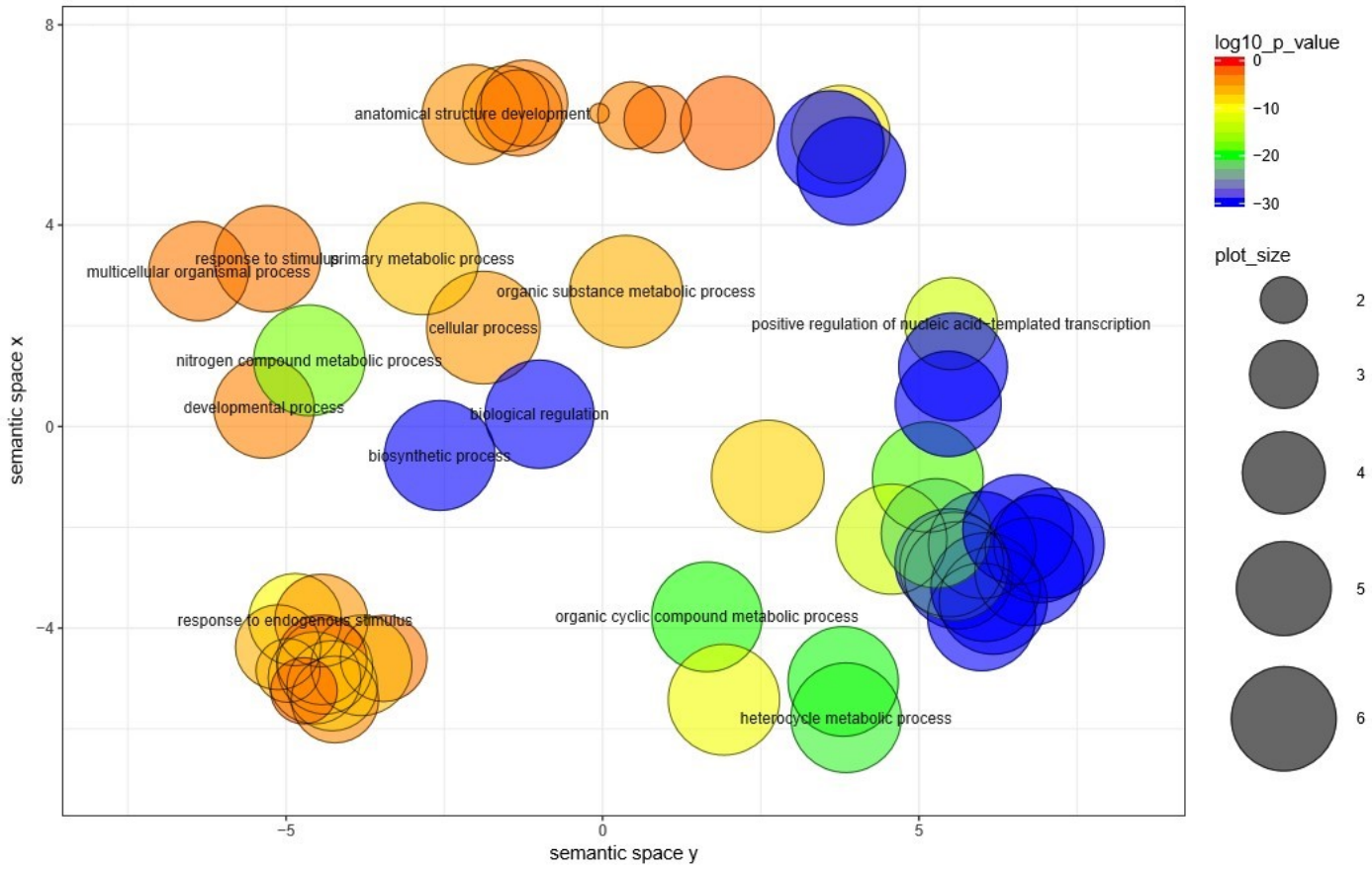


Figure 2. Gene ontology analysis (http://revigo.irb.hr) of *J. curcas* ESTs annotated for transcription factors (p-value <0.01), relative the Biological Process categories. The color of an individual bubble indicates the value of p-value. The size of the bubble representing the GO term shows the frequency of the term in the annotation database.

Concerning the 15 enriched GO terms related to Molecular Function, seven (47%) represented 46 - 84 putative TF genes (Table S1). The GO enrichment (Figure 3) highlighted the terms: *binding* and its relatives (*DNA binding*, *sequence-specific DNA binding*, *nucleic acid binding transcription factor activity*), *transcription factor activity,* and *protein dimerization activity*. All of those GO terms representing functional actions

involving TFs. Also, as expected, about the Cell Component GO categories, the *nucleus* was the location showing the most relevant *p-value* (Table S1).

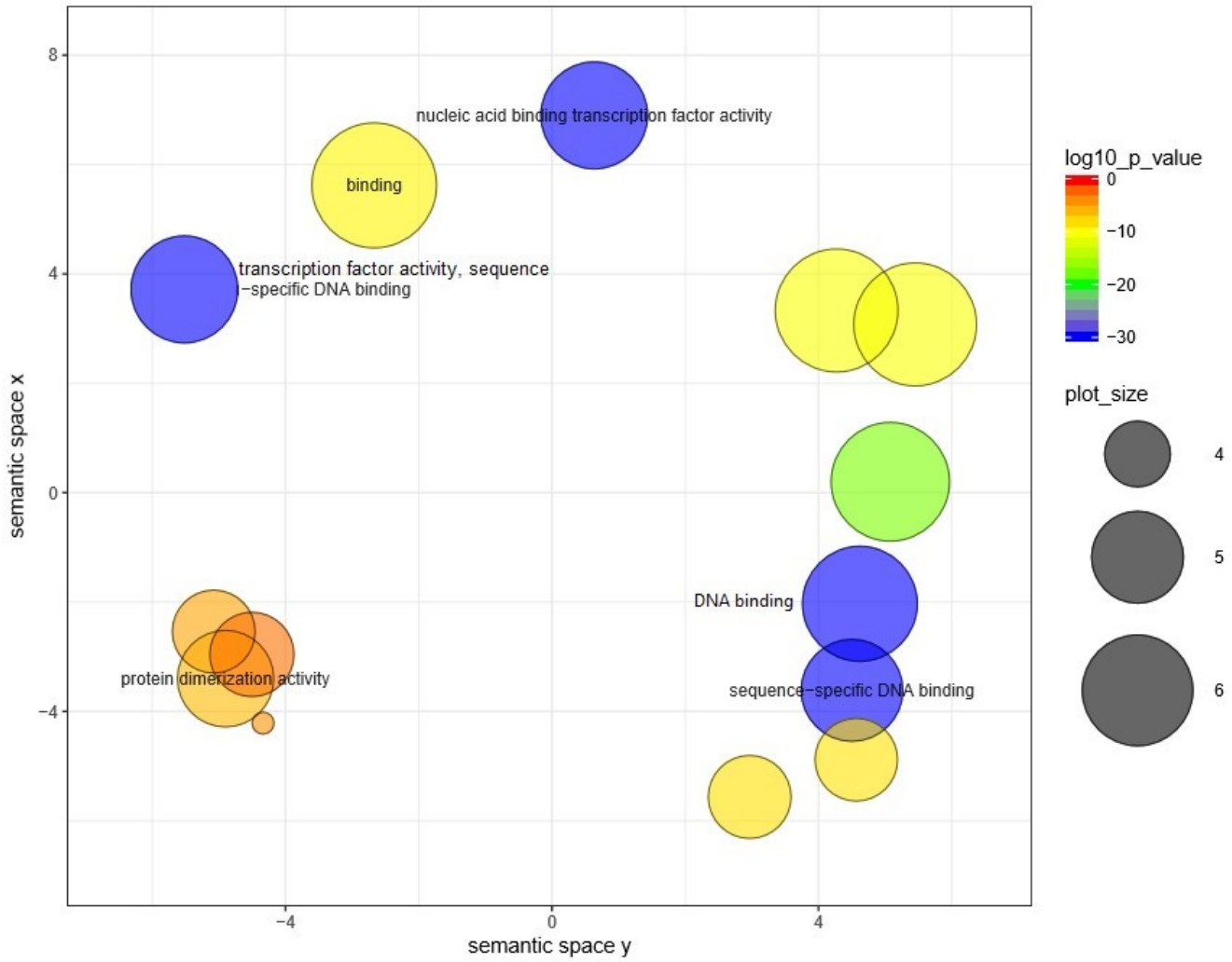


Figure 3. Gene ontology analysis (http://revigo.irb.hr) of *J. curcas* ESTs annotated for transcription factors (p-value <0.01), relative the Molecular Function categories. The color of an individual bubble indicates the value of p-value. The size of the bubble representing the GO term shows the frequency of the term in the annotation database

The most prospected CAREs in the promoters of the *J. curcas* TF genes

The CARE prospection covering 1,500 bp in the potential promoter regions of the putative 106 TF genes detected 20,158 motifs, representing 209 unique CAREs (almost 52% of the 469 CAREs available in the newPLACE database). Basically, 48 - 91 CARE motifs were detected by TF promoter, but most of the promoters (around

60%) presented 60 - 70 motifs. The TF genes showing higher numbers of CAREs in their respective promoters were *erf012* (91), *nf-yc2* (86), *divaricata* (85; ref. gene LOC105632398), *nf-yc3* (82), and *hat3* (81). From the 209 detected CAREs, only two were presented in all TF promoters (CACTFTPPCA1 and CAATBOX1; Table 2). The 25 most prevalent CAREs (Table 2), representing 13,156 of the 20,158 detected motifs (around 65%) were identified in at least 84 of the 106 TF promoters.

Insert Table 2

According to the Table 2, the most prevalent CARE (detected in almost all TF promoters analyzed; 105) was DOFCOREZM (1,340 AAAG motifs). This motif is the core site for binding of Dof proteins. TF genes showing expressive DOFCOREZM motifs (25 – 29; Table S2) in their potential promoters were: *zhd4* (29), *gamyb* (28), *tcp3* (27), *hdg2* (26), and *hat4* (25). Considering the motif distributed along the promoter segments (S_I – S_V; Table S2), the TF gene *gt-3b* presented 61% of them in S_I (-1 to -300 bp). Relevant motif concentration in S_I was also perceived in the promoters of the TF genes *hdg2* and *gata26* (around 50%; Table S2). In the others segments of these promoters, the motif concentration was lower or even not detected. But, considering others TF promoters, some different **m**otif **d**ominant **p**osition (MDP) were verified: S_{II} (-301 to -600 pb; TF genes: *erf4* and *arf5*); S_{III} (-601 to -900 pb; TF genes: *myb25* and *pat1*); S_{IV} (-901 to -1200 pb; TF genes: *dpbf3* and *hsfc*); S_V (-1201 to -1500 pb: TF genes: *tcp3* and *zhd4*).

Another CARE well detected and in almost all analyzed promoters (104; Table 2) was GT1CONSENSUS [motif GRWAAW (R = A/G; W = A/T)]. This CARE corresponds to GT-1 binding site of TFs from the Trihelix family. TF genes showing in their promoters expressive numbers of GRWAAW motifs were (Table S2): *erf003* (18), *erf004* (17), *tcp2* (16), *dpbf3* (16), *athb-12* (16) and *hsfc1* (16). Some TF genes showed the MDP (Table S2) in S_I [*rf2b* (50% of the motifs in its promoter), *bhlh121* (46%), and *gt-3b* (41%)]. But, another TF genes presented the MDP (Table S2) in S_{II} [*athb-12 (*56%), *bhlh95* (50%), and *bhlh147* (46%)], S_{III} [*hat7* (66%), *taf4b* (46%), and *gt-1* (45%)], S_{IV} [*myb86* (ref. gene LOC105642721; 53%), *seuss* (45%), *dpbf2* and *zhd4* (both with 41%)], or S_V [*erf012* (75%), *divaricata* (ref. gene LOC105631171; 60%), and *tcp3* (46%)]. Thus, distinct TF promoters showed the MDP in different segments.

However, the highest motif concentration (75%) in S_V (*erf012*; Table S2) was due the fact that the respective motif was not detected in the first three segments.

Another CARE well represented (detected in the 106 analyzed promoters; Table 2), was CACTTFPPCA1 (motif YACT). Expressive motif abundance was observed in the promoters of the TF genes (Table S2): *erf002* (19), *not3* (18), *myb86* (ref. gene LOC105642721; 17), *tcp9* (16), and *tcp23* (16). Concerning the MDP, few TF genes showed a YACT MDP in S_I [*nf-yb9* (41%)], S_{II} [*gtag* (63%; almost not detected in the others segments)], and S_{III} [*erf034* (45%)], but a distal MDP was well represented in S_{IV} [genes *nac25* (45%), *divaricata* (ref. gene LOC105631171; 54%), *mybB44* (67%), *val2* (50%), and *bhlh95* (45%)], and S_V [genes *fus3* (54%), *myb308* (40%), *myb86* (ref. gene LOC105642721; 47%), and *hat3* (40%)]. Thus, the distal segments seem to concentrate the YACT motifs.

The CARE ARR1AT (motif NGATT; the ARR1-binding site), presented in 105 putative TF genes (Table 2), was detected (14 – 16 motifs) in the promoters of the genes (Table S2): *taf1* (16), *hat4* (15), *erfF011* (14), *erfF084* (14), *val1* (14), and *nf-ya9* (14). The NGATT MPD detection (Table S2) was, basically, in S_I [genes *val1* (71%), *swi3b* (63%), and *bzip16* (50%)], and S_{IV} [genes *erf073* (63%), *hat4* (50%), and *ail6* (45%)].

The ROOTMOTITFAPOX1 (motif ATATT), another CARE well detected and presented in almost all analyzed promoters (105; Table 2), showed expressive concentration (Table S2) in the promoters of the TF genes *gbf1* (21 motifs), *dpbf2* (21), *swi3d* (20), and *gtaI* (19). Concerning the ATATT concentration, the MDP (Table S2) was verified in S_{IV} [genes *gbf1*, *hsfc1*, and *erf010* (90, 63, and 41%, respectively)] and S_V [gene *bhlh79* (55%)].

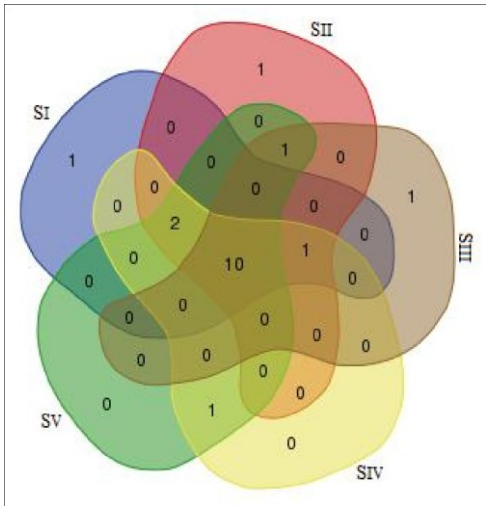
In turn, the GATABOX (motif GATA), also detected in almost all potential TF promoters, showed relevant motif concentration (Table S2) in the promoters of the genes *myb1r1* (16), *abi3* (14), *erf005* (13), *grf7* (13), and *gt-3b* (12), while the GATA MDP (Table S2) was detected basically in S_I [*erf005* (38%), *myb1r1* (31%)], S_{III} [*leunig* (54%)], and S_V [*abi3* (43%)].

In general, based on the results concerning those most detected CAREs presented in almost all promoters analyzed, the characteristic motifs were not uniformly distributed along the entire promoters, being the motifs concentrated in few segments, showing a variable MDP according to the TF gene/ promoter.

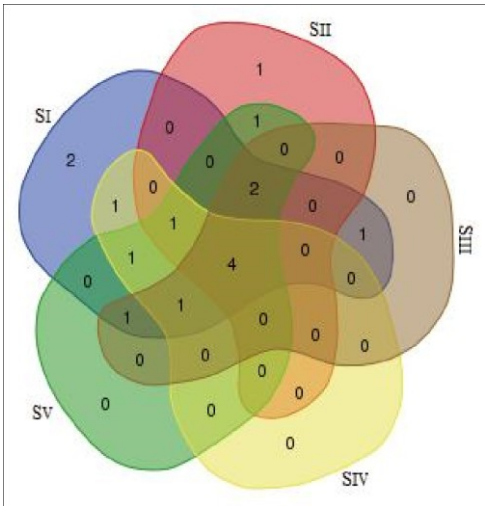
The CARES distribution on promoter segments according to the TF families' members and CG genes

Considering the six TF families showing the most frequently detected motifs (AP2/ERF, bHLH, bZIP, MYB, HD-ZIP, and HSF), an analysis was performed trying to find a CARE profile shared by promoters of genes from one of those TF families. Based on the Venn diagrams, using the motifs detected over each promoter segment from genes members of the mentioned TF family, the results showed that the majority of the different CARE motifs were qualitatively identified in the five promoter segments, being only a few motifs observed in a particular sector (Figure 4). Some CAREs commonly noted in the five segments were ROOTMOTIFTAPOX1, GT1CONSENSUS, CAATBOX1, ARR1AT, and CACTFTPPCA1 (Table S3). The mapping of these motifs considering the promoter segments and the TF family is provided in the Table S4. When this CARE distribution is compared with those comprising FT genes relative to some stress-responsive and ABA-related genes (Table S3), the representative CARE motifs were not uniformly distributed along the promoters, being the MDPs observed in some segments accord the TF family. For instance, the CARE DRECRTCOREAT (Table S4) presented MDP in S_I (AP2/ERF and bZIP TF families), S_V (HSF family), S_{IV} and S_V (HD-Zip family), and S_I, S_{III}, S_{IV}, and S_V (MYB TF family). In turn, the CARE ABREZMRAB28 (Table S4) showed the MDP in S_I (bHLH family), S_{II} (AP2/ERF family), and S_{III} (HD-Zip family), while the ABREATCONSENSUS (Table S4) showed the MDP in S_I (bHLH family), S_{II} (MYB family), and S_{IV}, S_V (AP2/ERF family). On the other hand, the CARE ABREMOTIFAOSOSEM presented the MDP mapped only in S_{II} (bHLH and bZIP families), while the ABREOSRAB21 showed the MDP only in S_{III} (AP2/ERF and HDZip families), and the MYCATRD22 showed the MDP only in S_{IV} (HD-Zip family), and S_I, S_{IV} (bHLH family).

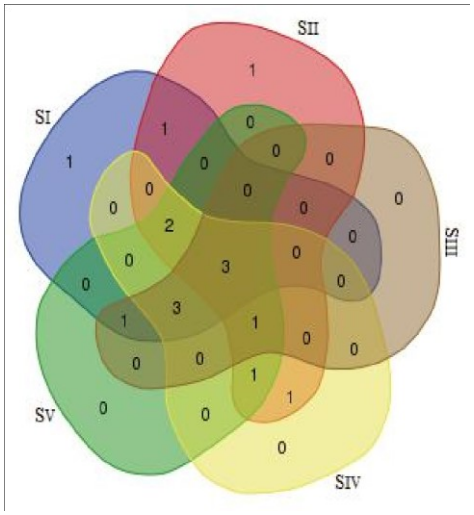
AP2/ERF



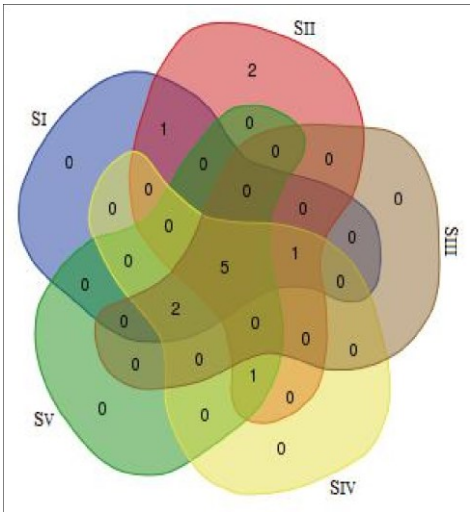
bHLH



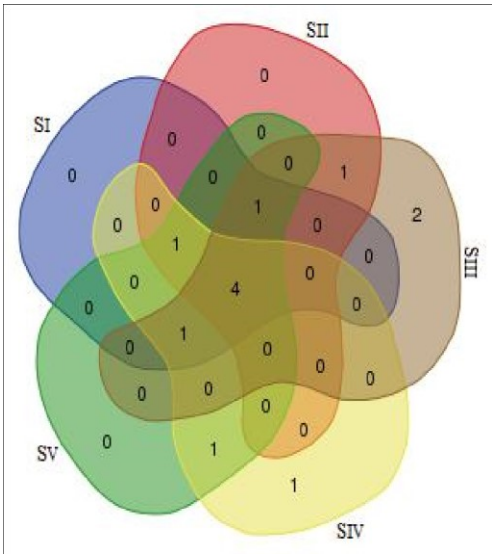
bZIP



MYB



HD-ZIP



HSF

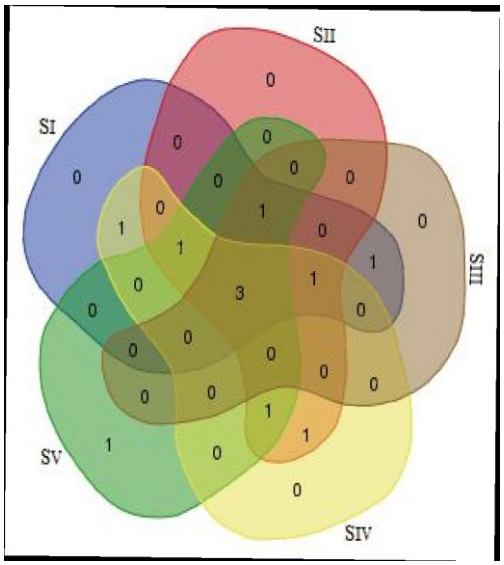


Figure 4. Venn diagram showing the distribution of cis-acting regulatory elements (CAREs) mapped along the segments (SI – SV) of the promoter regions of *J. curcas* trancription factors genes delimited from the start transcription site (STS) to the upstream -1,500 bp: SI (-1 to -300 bp), SII (-301 to -600 bp), SIII (-601 to -900 bp), SIV (-901 to -1200 bp), SV (-1201 to -1500 bp).

The most abundant CAREs foud in the promoters of the constitutive genes were DOFCOREZM, CACTFTPPCA1, GATABOX, MARTBOX, CAATBOX1, ROOTMOTIFTAPOX1, POLASIG3, GT1CONSENSUS, ARR1AT, ACGTATERD1 and TAAAGSTKST1 (Table S5). In all, 141 CAREs were foud in the promoter region of the constitutive genes analyzed (Table S6), where 136 CAREs had a distributed presence between the promoters of TFs and CG genes, 73 CAREs found exclusively in the set of promoters of TFs, and only five CAREs found exlcusively in the promoters of CG (Figure 5). Most CAREs were evenly distributed, and with a few motif units across the sectors of the CG promoters. However, in the promoter region of certain CG, some CAREs have concentrated on segments determined. For example, DOFCOREZM presente the MDP in S_{III} (*apx*) and in S_V (*act2*), CACTFTPPCA1 in S_{IV} (*pgd1*) and S_I (*ubia*), MARTBOX in S_{IV} (*ubia* e *ubib*) and S_{III} (*act2*), and CAATBOX1 in S_{II} (*act2*) (Table S5). Among the CAREs involved in the ABA response, MYCATRD22 was found in S_{II} (*pgd1*) and S_V (*unib*), ABRERATCAL in S_V (*tuba2*) and ABRELATERD1 in S_I (*apx*), S_{III} (*pgd1* e *act2*) and S_V (*tuba2*) (Table S7).

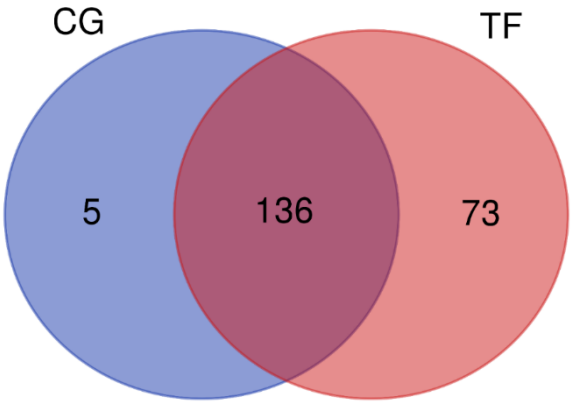


Figure 5. Venn diagram showing the distribution of cis-acting regulatory elements (CAREs) mapped along the segments of the promoter regions of *J. curcas* trancription factors and constitutive genes.

DISCUSSION

The set of *J. curcas* TF genes

From the putative TFs identified from the *J. curcas* ESTs/ contigs, 106 were manually cured and classified into 29 TF families. The most abundant TF family based on the identified genes belongs to the AP2/ERF family (16 members). This TF family regulates different development processes and responses to plant stresses (Lakhwani et al., 2016). Also, AP2/ERF is the largest family of plant-specific TFs (Yamasaki *et al.*, 2013), comprising 147 AP2/ERF identified members in *Arabidopsis* (Nakano et al., 2006), 170 in rice (Rashid et al., 2012), 291 in cabbage (Song et al., 2013), 116 in bamboo (Wu et al., 2015), and 173 in willow (Rao et al., 2015). AP2/ERF proteins are classified according to the number of specific sequences of the DNA-binding domain (AP2, RAV, and ERF): the AP2 group presents two AP2 domains; the RAV group contains an AP2 domain and a B3 domain, and the ERF group, an ERF domain and an AP2 domain (Gu et al., 2017).

The majority of AP2/ERF identified members comprised the ERF (Ethylene Response Factor) subfamily (13 candidates). This relative proportion is in agreement with the observed to the others plant species. From ERF subfamily, 122 members have been identified in Arabidopsis and 139 in rice (Nakano et al., 2006). TFs from ERF subfamily bind specifically to strictly conserved DNA motif sequences named GCCbox (GCCGCC), which are responsible for modulating the transcription of genes bearing this CARE in their promoters (Ohme-Takagi and Shinshi, 1995;). ERF proteins are involved in different stages of plant development, from seed germination to growth and plant development (Büttner and Singh, 1997), as well as in plant responses to biotic and abiotic stresses, through signaling pathways influenced by phytoregulators, including ABA, salicylic acid (Onate-Sanchez and Singh, 2002), jasmonic acid (Brown et al., 2003), and ethylene (Fujimoto et al., 2000). Among the AP2/ERF members identified in the present analysis, the following members stand out:

a) *ERF002* (ref. gene LOC105648655): ERF (Class IV, conserved N-terminal domain MCGGAAI/L) regulator of genes involved in ethylene biosynthesis (Tournier et al., 2003); its overexpression resulted in premature tomato seed germination (Joseleau et al., 1992);

b) *ERF004* (ref. gene LOC105637492): regulator of floral organ development, also acting on cell wall elongation (Shi et al., 2011); the expression pattern was associated with anther dehiscence (Aharoni et al., 2004);

c) *ERF005* (ref. gene LOC105632315): ERF acting in the innate immunity of the plant by activation of plant defense pathways in response to pathogens (Moffat et al., 2012); it was associated with defense genes responding to ethylene, salicylic acid, and jasmonic acid signaling;

d) *ERF009* (ref. gene LOC105646903): ERF (Class II, transcriptional inhibitor; Fujimoto et al., 2000) acting on plant defense mechanisms against necrotic fungi, mediating the DEAR1-dependent (DREB and EAR Motif Protein 1) ethylenesignaling pathways; a transcriptional repressor of the gene encoding DREB (Dehydration-Responsive Element Binding Protein; Maruyama et al., 2013);

e) *ERF010* (ref. gene LOC105639746): ERF (Class II) acting on transcription regulation through elements responsible for dehydration (DREs), in drought or cold induced gene promoters (Ohta et al., 2001);

f) *ERF011* (ref. gene LOC105630926): ERF (Class II) acting as transcriptional repressor with altered expression under abiotic stresses (K^+ depletion, cold, UV, and H_2O_2) (Lee et al., 2005); under increasing ABA levels, it negatively regulates ethylene biosynthesis by repressing involved genes (Li et al., 2011); also observed responding to jasmonic acid (Dombrecht et al., 2007);

g) *ERF012* (ref. gene LOC105649572): ERF (Class II) presenting an essential domain for activation of repression in target motifs (GCC-box); it was involved in the regulation of seed desiccation tolerance (González-Morales et al., 2016), and in the regulation of other ERFs (Ohta et al., 2001);

h) *ERF018* (ref. gene LOC105638000): it is another example of TF synergy; the overexpression of selected ethylene responsive ERFs, including ERF018, presented involvement in the regulation of stem diameter growth and wood chemotype of *Populus trichocarpa* (Vahala et al., 2013);

i) *ERF073* (ref. gene LOC105642565): TF strongly induced by ethylene (*Arabidopsis*), without responses to cold and dehydration (Licausi et al., 2010); it showed modulation of ethylene responses under hypoxic conditions (Yang et al., 2011);

The second TF family in the number of members identified (11; the MYB family) is also involved in several plant-specific processes, including cell morphogenesis, secondary metabolism, cell differentiation and stress response (Ambawat et al., 2013). The MYB family comprises four subfamilies, according to the MYB domain arrangements of the DNA-binding domains (Dubos et al., 2010; Chen et al., 2016). Each subfamily is composed of one, two, three or four imperfect repeats of helix-loop-helix that recognize the 5'-[GA]GATAA-3' motifs located in the major groove of the DNA (Yanhui et al., 2006). The subfamily R0R1R2R3 contains three MYB domains, while the R2R3 and R1R2R3 subfamilies have respectively two and three domains (Shin et al., 2011), and the subfamily R1-MYB, only one MYB domain (Du et al., 2009). The MYB family also is a significant TF family with more than 100 members in pear (Li et al., 2016), and up to more than 200 in cotton (He et al., 2016). Several MYB TF members were reported induced during the early stages under stressed conditions, such as stress osmotic (Lippold et al. 2009), ABA treatments (Jung et al., 2008), and high salinity, drought, and low temperature (Baldoni et al., 2015). In our analysis, the identified MYB members were:

a) *MYB1R1A* (ref. gene LOC105640456): TF activating genes responding to dehydration in potato (Shin et al., 2011); when over-expressed in chickpea under drought stress interacted by the C-terminal domain with galactosyltransferase, two protein kinases and ABA-responsive protein (ABR17-like; Ramalingam et al., 2015);

b) *MYB6* (ref. gene LOC105636993): TF regulator of innate immunity responses mediated by MLA (Leucine-Rich Repeat Protein), where the active state of MLA releases MYB6 from the WRKY1 repressor, and stimulates its binding to cognate *cis* elements to initiate resistance signaling to the biotic stress caused by *Blumeria graminis*, in barley (Chang et al., 2013);

c) *MYB25* (ref. gene LOC105633683): TF involved in the regulation of the process affecting cotton fiber cell differentiation (Walford et al., 2011);

d) *MYB44* (ref. gene LOC105639771): TF involved in the primary mechanism of stomatal closure, triggering the ABA signaling pathway, besides repressing the expression of a set of genes, favoring the tolerance processes to salinity and drought stresses (Jung et al., 2008);

e) *MYB86* (ref. genes LOC105642721 and LOC105631441): TF acting as a transcriptional repressor (Hemm et al., 2001); it is involved in the regulation of UV protection, cold tolerance, phenylpropanoid pathway (Jin et al., 2000, Vannini et al., 2004); and also in the osmotic adjustment (Park et al., 2010b);

f) *MYB308* (ref. gene LOC110008593): a potential repressor of the phenylpropanoid pathway after the cinnamic acid synthesis, through competitive inhibition with activating factors (Tamagnone et al., 1998);

Another TF family based on the number of members detected was bHLH (nine members); some of them were:

a) *ICE1* (*Inducer of CBF Expression 1*; ref. gene LOC105630202): TF that controls the gene expressing *CBF3* [C-repeat (CRT)-Binding Factor-3), one of the TFs that bind to *cis* elements CRT/DRE [C-repeat (CRT)/dehydration-responsive element] to activate the transcription of genes responding to cold stress (Liu et al., 1998);

b) *bHLH51* (ref. gene LOC105645984): a positive regulator of hypocotyl and root elongation of *Arabidopsis* seedlings responding to methyl jasmonate (Chen et al., 2015);

c) *bHLH63* (ref. gene LOC105630155): also known as *CIB1* (*cryptochrome-interacting basic-helix-loop-helix 1*), this TF interacts with CRY2 (*cytochrome 2*) through blue light mediation to regulate transcription and floral initiation in *Arabidopsis* (Liu et al., 2008);

d) *bHLH79* (ref. gene LOC105636735): TF induced during the stages of inflorescence development in *Syringa oblata* (Zengh et al., 2015);

e) *bHLH93* (ref. gene LOC105630754): TF acting as a regulator of flowering time in *Arabidopsis* plants under specific photoperiod conditions (Sharma, 2011);

f) *bHLH121* (ref. gene LOC105629018): TF responsible for activation and expression of the potassium transporters HAK5 under K^+ depletion conditions (Clouse and Sasse, 1998);

g) *bHLH147* (ref. gene LOC105643560): TF regulating negatively the brassinosteroid signaling (BRs), which involves many aspects of the growth and development of plants (Clouse and Sasse, 1998).

About the HD-Zip family, some of the featured TF members were:

a) *ATHB-6* (ref. gene LOC105630651): TF responsible for regulating diverse plant responses, such as stomatal closure, seed dormancy and inhibition of vegetative

growth; it is induced by ABA during drought stress, functioning as a negative regulator of the ABA signaling pathway (Söderman et al., 1999);

b) A*THB-54* (ref. gene LOC105649782): TF highly expressed in leaves, stems, and flowers of *Arabidopsis* under water deficit, different light conditions, or in responses to ABA (Henriksson et al., 2005);

c) *HB-12* (ref. gene LOC105642795): TF regulating negatively the expression of GA20ox1 gene (*gibberellin 20-oxidase 1*), which enzyme is crucial in the gibberellin synthesis (Son et al., 2010); it is induced by drought and exogenous ABA stimulus, acting together with *ABHB-7* as negative regulators of plant growth under these conditions (Ré et al., 2014);

d) *HAT3* (ref. gene LOC105643198): TF presenting induction in response to luminosity changes (Bou-Torrent et al., 2012); it is positively regulated by SPT (Spatula), a crucial TF in the regulation and development of carpel tissues (Alvarez and Smyth, 1999);

e) *HAT4* (ref. gene LOC105632304): also known as *ATHB-2* (*A. thaliana homeobox 2*), this TF is a plant morphogenesis regulator during light signals responses and phytochrome signaling (Carabelli et al., 1993);

f) *HAT5* (ref. gene LOC105640489): also called *ATHB-1* (*A. thaliana homeobox 1*), this TF is a transcriptional regulator involved in leaf and hypocotyl development (Aoyama et al., 1995); its expression is regulated by PIF1 (Phytochromeinteracting Factors; TF from the bHLH family), which negatively regulates the expression of genes involved in the chlorophyll biosynthesis pathway in the dark, besides acting directly and indirectly in the gibberellins biosynthesis (Castillon et al., 2007);

g) *HAT7* (*ATHB-3*; ref. gene LOC105642825): this TF inhibits the primary root development in *Arabidopsis* without affecting the secondary development (Hanson, 2000); it is a repressor of *ATHB-13*, another HD-Zip TF related to the cold tolerance, mediating the activation of proteins capable of stabilizing cell membranes and inhibiting growth under low-temperature conditions (Cabello et al., 2012);

These mentioned putative TFs composed the set of 106 *J. curcas* TFs (classified into 29 families) mapped in the *J. curcas* genome. Taking into account, the transcripts, the loci, and the genes, together with the scientific literature showing their action on plant development or regulating genes responding to environmental stresses, these data is a valuable resource to be exploited by *J. curcas* breeders around the world.

The CARES most prevalent in promoters of *J. curcas* TF genes

Few studies addressing the global TF set expressed by crops have been published (Riechmann et a
TFs genes responding to stimulus, the prospection of CAREs were carried out over some promoter regions. From those more prevalent CAREs based on the motifs detected (Table 2), the following should be mentioned:

a) DOFCOREZM: the representative motif (AAAAG) is the specific binding domain of Dof protein, which is a plant TF associated with seed germination (Papi et al., 2002), phytochrome signaling (Park et al., 2003), light responses (Yanagisawa 2000), phytoregulator responses (auxin and gibberellins; Kisu et al., 1998;), and drought (Zhang et al., 1995); this motif was detected in promoters of 36 Dof genes (*Cucumis sativus*), suggesting that these TFs can be regulated by themselves (Wen et al., 2016); such motifs were also observed in promoters of genes encoding PEPC (*phosphoenolpyruvate carboxylase*), and GST (*glutathione S-transferase*) (Yanagisawa, 2000); in promoters of 141 ERF genes (*Brachypodium distachyon*), they ranged from 2 to 24 motifs per promoter (Cui et al., 2016); this CARE was also the most abundant in promoters of 50 genes preferentially expressed during the meiosis of *A. thaliana* pollen grains (Li et al., 2014);

b) GT1CONSENSUS: the respective motif [GRWAAW (R=A/G; W=A/T)] is the Trihelix GT-1 binding site found in promoters of light-regulated genes (Liu et al., 2011), highly expressed genes involved with reproductive cells development (Sharma et al., 2011), phosphate transporter genes (Liu et al., 2011); concerning 141 promoters of *Brachypodium distachyon* ERF genes, they ranged up to 17 motifs per promoter (Cui et al., 2016);

c) ARR1AT: the considered motif (NGATT) is the TF Myb-*At*ARR1 binding site; this TF act as repressors in cytokine regulatory pathway, which phytoregulator (cytokine) is recognized for acting on the regulation of stress-induced leaf senescence (Büchert et al., 2011); the expression of the *pph* gene (*pheophytin pheophorbide hydrolyase*), which encodes the enzyme responsible for catalyzing chlorophyll degradation in foliar senescence, was suppressed by FT ARR1, repressing the chlorophyll degradation process (Zhang et al., 2016);

d) CACTTFPPCA1: the representative motif (YACT) is detected in the promoter of the gene codifying PEP carboxylase (*ppcA1*; *phosphoenolpyruvate carboxylase*;

Gowik et al., 2004), which catalyzes the reaction between PEP and CO_2, in the photosynthesis process (Izui et al., 2004) - PEP has a crucial function in stomatal opening processes (Britto and Kronzucker, 2005); this is one of the most observed motifs along the promoter of the gene coding GST (glutathione-S-transferases) in *Salicornia brachiata*, which presents functions in plant growth and development, besides hormone signaling responding to abiotic stresses (Tiwari et al., 2016).

These few examples showed that the CARE motif and the respective TF carrying the necessary binding site domain are crucial elements in the expression of target genes. Therefore, the distribution of motifs detected in promoters of genes could improve the understanding of the regulatory role implicit by particular CAREs.

The qualitative and quantitative aspect of the CARES in promoter regions

Some motifs could be very prevalent in promoters of a particular set of genes. For instance, the motif YACT (CARE CACTTFPPCA1) was the most common in promoters of rice genes encoding metal ion transport proteins (*Os*Nramp), with more than 100 copies detected at -600 and -1000 bp of the promoter region of seven *Os*Nramp genes (Bervald, 2009), also this motif presented around 90 copies over the promoter regions of rice phosphate transporter genes (Hatorangan et al., 2009). In this way, the number of copies and the spacing between specific CAREs should be crucial for the modeling of relevant synthetic promoters. In addition, concerning TF genes and their involvement with stress responsive genes, the presence of multiple CAREs in promoters of these TF genes potentiate combinations of responses to the same environmental stimulus or even to different stimuli. Thus, although TFs belonging to the ERF subfamily (APF2/ERF family) mediate ethylene responses, they include members that respond to biotic and abiotic stresses (Dey and Vlot, 2015). For instance, ERF1 binds to CAREs in promoters of different genes that regulate responses to pathogens (De Boer et al. 2011) or tolerance to drought, salinity or temperature stress (Cheng et al., 2013). Also, TFs from the AP2/ERF family, in addition to the ability to bind to GCC-box (AGCCGCC motif) in promoters of particular genes (Fujimoto et al. 2000), have the property of binding to RCCGCC motifs of DRE/CRT elements (*dehydration responsive element/C-repeat*) (Wang et al. 2012). The induced expression of the TF DREB2A (DREB subfamily) in *Arabidopsis* under drought stress is resulted, at least in part, by the presence of DRE/CRT elements (Liu et al., 1998).

Still concerning drought tolerance, a total of 57 bHLH TFs were identified as genes responsive to drought (Li et al., 2016). Thus, TF members of the bHLH family would be good candidates to increase the plant stress tolerance, due to their performance by maintaining turgid leaves in stressed environments (Castilhos et al., 2014). Individually, the TF bHLH122 would be a positive regulator considering the expression of genes related to drought stress signaling (Liu et al., 2014). Otherwise, the TF bHLH092 would respond to osmotic stress (Jiang et al., 2009) and regulate the circadian rhythms in *Arabidopsis* (Hanano et al., 2008). Likewise AP2/ERF and bHLH families, studies revealed that MYB family members, such as *At*MYB2 (Urao et al., 1993), *At*MYB44 (Jung et al., 2008), *At*MYB60 (Cominelli et al., 2005) and *At*MYB61 (Liang et al., 2005), are induced by water deficit stress in *Arabidopsis*, through involvement in the regulation of stomatal opening (Urao et al., 1993).

The CAREs that occurred only in the promoters of CGs (ZDNAFORMINGATCAB1, GARE2OSREP1, PYRIMIDINEBOXOS, GMHDLGMVSPB and BS1EGCCR) are linked to the regulation of genes related to metabolic processes of development that undergo environmental and hormonal influences. For example, CARE BS1EGCCR, of AGCGGG motif, is a tissue-specific vascular system found in the promoter of the *ccr* gene (Cinnamoyl-CoA reductase), which encodes the first enzyme responsible for catalyzing the biosynthesis of lignin monomeric units (Lacombe et al., 2000). ZDNAFORMINGATCAB1 (ATACGTGT), also called Z-box, is an element involved in the transcriptional activity mediated by luminosity (LRE - light-responsive element) which responds to a broad spectrum of light (Yadav et al., 2005), found in promoters of genes related to development, such as *cab1* (chlorophyll a/b binding protein gene), *cop1* and *hy5* (Ang et al., 1998; Yadav et al., 2002). GMHDLGMVSPB (CATTAATTAG) was found in the promoter of a member encoding a family protein CesA (cellulose synthase), composed of catalytic subunits of a large protein complex responsible for the deposition of cellulose in the cell wall (Creux et al., 2008). PYRIMIDINEBOXOS (CCTTTT) is considered a responsive element to Gibberellic acid (GARE - Gibberellin-respons cis-element) which was found in the promoter of the genes *RAmy1A* in rice and *Amy2/32b* and barley, which encode α-amylase during the development of the endosperm (Roger et al., 1994; Morita et al., 1998). GARE2OSREP1 (TAACGTA) is also a gibberellin-responsive element (GARE - *Gibberellin-responsive element*) found in the promoter region of a gene encoding REP-1

(cystein proteinase), an enzyme responsible for the degradation of glutein in the endosperm in rice (Sutoh & Yamauchi, 2003). During germination of cereal seeds, gibberellins (GAs) are synthesized in the embryo and induce the expression of hydrolytic enzymes, such as α-amylases and proteinases, to degrade and mobilize starch and protein allocation for energetic storage of the embryo for growth (Jacobsen and Beach, 1985; Fincher, 1989).

Concerning the presence of multiple CAREs in TF gene promoters, considering those TF families showing the most prevalent CAREs in the present work (AP2/ERF, bHLH, bZIP, MYB, HD-Zip and HSF), also considering those TF responding to environmental stress or the ABA stimulus, some of the shared CAREs comprised DRECRTCOREAT, ABRERATCAL, and MYCATRD22. The MYCATRD22, a dehydration responsive element, was a CARE detected in the present work in promoters of several TF genes from different families [AP2/ERF (eight members), MYB (four) and bHLH (two)]. The representative motif of this CARE (CACATG) is a target of two TFs [rd22BP1 (MYC) e ATMYB2 (MYB)], which are transcription activators of *rd22* (*responsive to desiccation 22*) gene, in response to drought or ABA stimulus (Abe et al., 1997). This CARE was found in promoters of 159 ERF genes in rice (Pegoraro et al., 2013), and in promoters of 25 genes with significant expressions in response to osmotic stress in leaves of *Arabidopsis* (Vandepoele et al., 2009). Superimposed on the motif CACATG in the *rd22* gene promoter was identified the motif CANNTG, representing the CARE MYCCONSENSUSAT, which is also a water stress responsive recognition site, besides a regulatory function was associated with the TF CBF3 responding to drought (Abe et al., 2003).

Another CARE identified in the present study, in promoters of genes belonging to several FT families, was MYCATERD, presented in AP2/ERF (11 members), MYB (three) and NFY (two). This CARE presents the motif CATGTG, a consensus sequence recognized by the TF NAC, which is required for the expression of *erd1* (*early responsive to dehydration stress 1*) gene. This gene encodes ClpA (*ATP binding subunit of the caseinolytic ATP-dependent protease*), which interact with the chloroplastlocalized ClpP protease to facilitate proteolysis, being related to the embryonic and leaf development in *Arabidopsis* (Akita et al., 1997). Microarray analyses showed induction of NAC genes in transgenic *Arabidopsis* plants under water deficit stress, increasing the drought tolerance (Tran et al., 2004).

Still in the present work, the CARE ABRERATCAL was detected in promoters of members of the AP2/ERF (five members), bHLH (five), MYB (two), NAC (two), HD-Zip (two), HSF (two), NFY (two), besides at least one member from the families ARF, TGA, GATA, LEUNIG and VAL. The CARE motif ACGCG (G/T/C) corresponds to a binding site of TF from the CAMT family (*Calmodulin-binding transcription activator*), whose members mediate the transcriptional activation of genes in response to transient cytosolic Ca^{2+} signals (Kaplan et al., 2006). In plants, transient Ca^{2+} signals mediate responses to environmental stresses, including salinity, drought and cold, suggesting a relevant role in the initial phase of the signal transduction pathway of these stresses (Knight, 2000). The Ca^{2+} signaling triggered in *Arabidopsis* is recognized by activating class I genes of TCP factors through the binding sites in the promoter region (Whalley et al., 2011). TCP proteins regulate different aspects of plants development by molecular signaling pathways involving the phytoregulators cytokinin, auxin and jasmonic acid, besides involvement in defense signaling pathways, stress acclimatization, and interaction with other TFs (Aguilar-Martinez et al., 2013). Also, the activity of some TF bHLH proteins can be suppressed by an increase in intracellular concentration of Ca^{2+} (Hermann et al., 1998), and it has been confirmed that the Ca^{2+} ion binds to the *At*NG1, TF (bHLH family) that recognizes E-boxes motif present in promoter regions of many saline stressinduced genes (Kim and Kim, 2006).

Another CARE with motif identified in promoters of genes belonging to several FT families (AP2/ERF, bHLH, bZIP, HD-Zip, HSF, NAC, GATA, MYB, TAF, and NFY), was DRECRTCOREAT. The related motif RCCGCA (R = G/A) has been described in promoters of rice genes encoding *Os*DREB1A and *Os*DREB2A, which are critical TFs in the activation of the expression of genes responsive to salinity, drought, and cold (Dubouzet et al., 2003). TFs DREB1A bind efficiently in both the GCCGAC and ACCGAC motifs and the DREB1A overexpression in transgenic *Arabidopsis* plants showed functional results in drought and cold tolerance (Qin et al. 2004).

In a similar way, CAREs responding to ABA [ABREs (*ABA-Responsive Elements*; ACGTGG/TC)], such as ABREATCONSENSUS (YACGTGGC – Y = C/T), ABREZMRAB28 (CCACGTGG), ABREMOTIFAOSOSEM (TACGTGTC), and ABREOSRAB21 (ACGTSSSC – S = C/G), were detected in promoters of genes from several TF families, including AP2/ERF (four members), and others families with at

least one identified TF member (bHLH, HSF, HD-Zip, MYB, NF-YB, and TGA). ABREs are binding elements of ABFs (*ABA-responsive elements binding factors*), such as ABF3 and ABF4, and positive regulators of NYE1 transcripts, whose protein act crucially on chlorophyll degradation during the maturation and senescence of vegetative organs in *Arabidopsis*, via ABA signaling (Choi et al., 2000). ABA, in turn, is a phytoregulator that play essential roles in plant growth and development (Leung and Giraudat, 1998), mediating physiological processes and adaptive responses to various adverse environmental conditions, such as salinity and drought (Uno et al., 2000). In the promoter region of the *ERF002* gene in tomato (*Solanum lycopersicum*) three ABRE motifs (positions -929, -1.183 and -1.194 bp) were associated with the germinative development process, through loss of inhibition typically influenced by ABA, favoring seeds germination (Pirrello et al., 2006). Therefore, the proper interpretation of CARE motifs in specific TF promoters may contribute to the appropriate modeling of synthetic promoters responsive to stress events to which the plants are exposed. For example, recently the construction of a synthetic promoter used the artificial design of the element 2xW-box (2 x TTGAC), of the WBOXATNPR1 CARE (Zhu et al., 2015), a target element of WRKY TFs specifically induced by salicylic acid (SA) (Chen & Chen, 2000) and essential in the expression of *npr1* (nonexpressor of PR genes 1), a pathogenicity-related defense gene (Yu et al., 2001). The synthetic element *6xABRE* (6 X TACGTGTC), CARE motif ABREMOTIFAOSOSEM, was used for the construction of a synthetic promoter in the gene *rd29a* (Response to Dehydration29a) (Wu et al., 2018), recognized by regular mechanisms of perception and rapid response in situations of drought, cold and salinity (Yamaguchi-Shinozaki et al., 1993). The generated 6xABRE reporters have been demonstrated as useful tools in the temporal and quantitative analysis of the transcriptional response of rd29a mediated by treatment of the hormone ABA and saline stress on roots of *Arabidopsis thaliana* (Wu et al., 2018). The increase in the transcriptional stimulus of BZIP TFs *HSBF* (hex-1-specific binding factor) and *ASF-1* (activation sequence factor-1), that participate in epigenetic regulation via histone modification in response to water stress and the presence of ABA, was conferred both on young leaves with mature tobacco, from the construction of the synthetic element *hex-1* (histone H3), which contains the motif YACGTGGC (ABREATCONSENSUS) (Lam & Chua, 1991). The artificial combination of the ATATT motif, belonging to CARE ROOTMOTIFTAPOX1, was used for the synthesis of a green tissue specific

promoter (Green tissue-specifically Expressed AT-rich element - GEAT) in rice, whose specificity of expression has been demonstrated in several photosynthetic tissues, such as leaf, sheath, panicle and stem (Wang et al., 2015). The construction of four synthetic copies in *Arabidopsis thaliana* of CARE DRECRTCOREAT (GCCGAC), essential for the transcriptional activation of ABA-independent pathways responsive to cold, drought and salinity stress (Yamaguchi-Shinozaki & Shinozaki, 1994), demonstrated an efficient response in the quantitative increase of expressed transcripts of the gene *cor15a* (Kim et al., 2002), a cold-responsive gene (COR) involved in the cold acclimatization of adult plants and tolerance to osmotic stress during the early stages of plant development (Wang & Hua, 2009; Liu et al., 2014). Interestingly, in the latter study, it has been shown that when only two copies of GCCGAC motifs are fused in the promoter, the transcriptional response to cold tolerance is not effective. Therefore, the importance of understanding the functional elements responsible for directing gene expression is crucial for the construction of synthetic promoters of interest, in order to enable the expression of potential transgenes (Venter, 2007). Moreover, the selection, number of copies and spacing of CARE motifs determine the force and spatiotemporal pattern of the synthetic promoters (Liu and Stewart, 2016). Depending on the elements included synthetic promoters can generate constitutive, spatiotemporal inductive combinations (Hernandez-Garcia & Finer, 2014). A synthetic promoter referred to as DR5, consisting of 11 bp tandem repeats of auxin-responsive TGTCTC elements (CARE AuxRE), behaved as in vitro ARF1 TF binding sites in promoters of the *GH3* gene in soybean, showing greater auxin responsiveness than the natural composition (Ulmasov et al., 1997).

A transient expression analyzes with the gusA reporter gene system comparing two copies of the ACGT motif differently influencing the expression of the gene encoding PP2C (*protein phosphatase 2C*) in response to salicylic acid or abscisic acid in tobacco demonstrated the spacing importance between motifs (Mehrotra and Mehrotra, 2010). In the case, two copies of ACGT motifs on synthetic promoters resulted in inducibility to salicylic acid when separated by five bp, but when separated by 25 bp allowed the promoter of the *PP2C* gene to be induced by abscisic acid and not by salicylic acid. Thus, synthetic promoters differing from native promoters can provide expression profiles that do not exist in nature and therefore present applicability as a tool in plant breeding programs (Hernandez-Garcia & Finer, 2014).

CONCLUSIONS

The *in silico* investigation of TF genes derived from *J. curcas* ESTs public available (NCBI), with the mapping of the respective genes over the *J. curcas* genome, provided potential promoter regions (1.5 kb *upstream* of the TSS of the TF genes), in which CAREs prospected over five segments. The motifs detection allied to respective quantification of them in the promoter segments could help to understand the CAREs influencing the TF gene expression, especially those responding to environmental stimuli. The results revealed CAREs involved in the regulation and expression of *J. curcas* TF genes, and their interactions with cognate TFs, as well as to anticipate cooperative actions involving genes and TFs, allowing a better understanding of the transcriptional regulation of these genes during normal plant development or under stress conditions. The CAREs, especially those associated with a constitutive and developmental function, based on the six TF families showing the most curent motifs, were distributed along the promoter segments, while most of which related to stress responses and ABA stimulus, were not equally distributed in the prospected sectors. The importance of this finding still requires further studies and may help in the development of innovative biotechnology tools, including the proposition of synthetic promoters, aimed to increase the transcription efficiency of *J. curcas* genes, mainly those responsive to stress, which would be useful for the plant breeding programs.

Competitive interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the Brazilian agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 404357/2013-0 and 309779/2014-6), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE; IBPG-0271-5.01/14) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERENCES

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Lwasaki, T., Hosokawa, D., Shinozaki, K., 1997. Role of Arabidopsis MYC and MYB Homologs in Drought-and Abscisic Acid-Regulated Gene Expression. Plant Cell 9, 1859–1868. https://doi.org/10.1105/tpc.9.10.1859

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., 2003. Transcriptional Activators in Abscisic Acid Signaling. Society 15, 63–78. https://doi.org/10.1105/tpc.006130.salt

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K. Yamaguchi-Shinozaki, K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell 15, 63–78. doi 10.1105/tpc.006130

Aguilar-Martínez, J.A., Sinha, N., 2013. Analysis of the role of Arabidopsis class I TCP genes AtTCP7, AtTCP8, AtTCP22, and AtTCP23 in leaf development. Front. Plant Sci. 4, 1–13. https://doi.org/10.3389/fpls.2013.00406

Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., Arkel, G. van, Pereira, A., 2004. The SHINE Clade of AP2 Domain Transcription Factors Activates Wax Biosynthesis, Alters Cuticle Properties, and Confers Drought Tolerance when Overexpressed in Arabidopsis. Plant Cell 16, 2463–2480. https://doi.org/10.1105/tpc.104.022897

Akita, M., Nielsen, E., Keegstra, K., 1997. Identification of Protein Transport Complexes in the via Chemical Cross-Linking Chloroplastic Envelope Membranes. J. Cell Biol. 136, 983–994.

Allen, R.D., Bernier, F., Lessard, P.A., Beachy, R.N. 1989. Nuclear factors interact with a soybean beta-conglycinin enhancer. Plant Cell 1, 623–631.

Allen, R.S., Li, J., Stahle, M.I., Dubroue, A., Gubler, F., Millar, A.A., 2007. Genetic analysis reveals functional redundancy and the major target genes of the Arabidopsis miR159 family. Proc. Natl. Acad. Sci. 104, 16371–16376. https://doi.org/10.1073/pnas.0707653104

Alvarez, J., Smyth, D.R., 1999. CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. Development 126, 2377–86.

Ambawat, S., Sharma, P., Yadav, N.R., Yadav, R.C., 2013. MYB transcription factor genes as regulators for plant responses: An overview. Physiol. Mol. Biol. Plants 19, 307–321. https://doi.org/10.1007/s12298-013-0179-1

Ang, L.-H., Chattopadhyay. S., Wei, N., Oyama, T., Okada, K., Batschauer, A., Deng, X.-W. 1998. Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. Mol Cell, 1, 213–222.

Aoyama, T., Dong, C.-H., WU, Y., Carabelli, M., Sessa, G., Ruberti, I., Morelli, G., Chua, N.-H., 1995. Ectopic Expression of the Arabidopsis Transcriptional Activator Athb-1 Alters Leaf Cell Fate in Tobacco. Plant Cell 7, 1773–1785. https://doi.org/10.1105/tpc.7.11.1773

Arruda, F.P. de, Beltrão, N.E. de M., Andrade, A.P. de, Pereira, W.E., Severino, L.S., 2004. CULTIVO DE PINHÃO MANSO (Jatropha curca L.) COMO ALTERNATIVA PARA O SEMI-ÁRIDO NORDESTINO. Rev. bras. ol. fibros 8, 789–799.

Baldoni, E., Genga, A., Cominelli, E., 2015. Plant MYB transcription factors: Their role in drought response mechanisms. Int. J. Mol. Sci. 16, 15811–15851. https://doi.org/10.3390/ijms160715811

Bate, N., Twell, D. 1998. Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol. Biol. 37(5), 859–869.

Bervald, C.M.P., 2009. Genômica comparativa em gramíneas. Universidade Federal de Pelotas.

Bhattacharjee, S., Renganaath, K., Mehrotra, R., Mehrotra, S., 2013. Combinatorial control of gene expression. Biomed Res. Int. 2013, 1–11. https://doi.org/10.1038/nsmb820

Bou-Torrent, J., Salla-Martret, M., Brandt, R., Musielak, T., Palauqui, J.-C., MartínezGarcía, J.F., Wenkel, S., 2012. ATHB4 and HAT3, two class II HD-ZIP transcription factors, control leaf development in Arabidopsis. Plant Signal. Behav. 7, 1382–1387. https://doi.org/10.4161/psb.21824

Britto, D.T., Kronzucker, H.J., 2005. Nitrogen acquisition, PEP carboxylase, and cellular pH homeostasis: New views on old paradigms. Plant, Cell Environ. 28, 1396–1409. https://doi.org/10.1111/j.1365-3040.2005.01372.x

Brivanlou, A.H., Darnel Junior, J.E., 2002. Signal transduction and the control of gene expression. Science. 295:813–818. https://doi: 10.1126/science.1066355

Brown, R.L., Kazan, K., McGrath, K.C., Maclean, D.J., Manners, J.M., Cooperative, 2003. A Role for the GCC-Box in Jasmonate-Mediated Activation of the PDF1.2 Gene of Arabidopsis. Plant Physiol. 132, 1020–1032. https://doi.org/10.1104/pp.102.017814

Buchel A.S., Brederode F.T., Bol J.F., Linthorst H.J.M. 1999. Mutation of GT-1 binding sites in the Pr-1A promoter influences the level of inducible gene expression in vivo. Plant Mol. Biol. 40, 387–396.

Büchert, A.M., Civello, P.M., Martínez, G.A., 2011. Chlorophyllase versus pheophytinase as candidates for chlorophyll dephytilation during senescence of broccoli. J. Plant Physiol. 168, 337–343. https://doi.org/10.1016/j.jplph.2010.07.011

Büttner, M., Singh, K.B., 1997. Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. Proc. Natl. Acad. Sci. U. S. A. 94, 5961–5966. https://doi.org/10.1073/pnas.94.11.5961

Cabello, J. V., Arce, A.L., Chan, R.L., 2012. The homologous HD-Zip i transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesisrelated and glucanase proteins. Plant J. 69, 141–153. https://doi.org/10.1111/j.1365313X.2011.04778.x

Carabelli, M., Sessa, G., Baima, S., Morelli, G., Ruberti, I., 1993. The Arabidopsis Athb-2 and -4 genes are strongly induced by far-red-rich light. Plant J. https://doi.org/10.1046/j.1365-313X.1993.04030469.x

Castilhos, G., Lazzarotto, F., Spagnolo-Fonini, L., Bodanese-Zanettini, M.H., MargisPinheiro, M., 2014. Possible roles of basic helix-loop-helix transcription factors in adaptation to drought. Plant Sci. 223, 1–7. https://doi.org/10.1016/j.plantsci.2014.02.010

Castillon, A., Shen, H., Huq, E., 2007. Phytochrome Interacting Factors: central players in phytochrome-mediated light signaling networks. Trends Plant Sci. 12, 514–521. https://doi.org/10.1016/j.tplants.2007.10.001

Chang, C., Yu, D., Jiao, J., Jing, S., Schulze-Lefert, P., Shen, Q.-H., 2013. Barley MLA Immune Receptors Directly Interfere with Antagonistically Acting Transcription Factors to Initiate Disease Resistance Signaling. Plant Cell 25, 1158–1173. https://doi.org/10.1105/tpc.113.109942

Chen, C., Chen, Z. 2000. Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. Plant Mol Biol. 42(2):387-96

Chen, B., Niu, F., Liu, W.Z., Yang, B., Zhang, J., Ma, J., Cheng, H., Han, F., Jiang, Y.Q., 2016. Identification, cloning and characterization of R2R3-MYB gene family in canola (Brassica napus L.) identify a novel member modulating ROS accumulation and hypersensitive-like cell death. DNA Res. 23, 101–114. https://doi.org/10.1093/dnares/dsv040

Chen, H., Ahmad, M., Rim, Y., Lucas, W.J., Kim, J.Y., 2013. Evolutionary and molecular analysis of Dof transcription factors identified a conserved motif for intercellular protein trafficking. New Phytol. 198, 1250–1260. https://doi.org/10.1111/nph.12223

Chen, R., Li, Q., Tan, H., Chen, J., Xiao, Y., Ma, R., Gao, S., Zerbe, P., Chen, W., Zhang, L., 2015. Gene-to-metabolite network for biosynthesis of lignans in MeJAelicited Isatis indigotica hairy root cultures. Front. Plant Sci. 6, 1–15. https://doi.org/10.3389/fpls.2015.00952

Cheng, M.-C., Liao, P.-M., Kuo, W.-W., Lin, T.-P., 2013. The Arabidopsis ETHYLENE RESPONSE FACTOR1 Regulates Abiotic Stress-Responsive Gene Expression by Binding to Different cis-Acting Elements in Response to Different Stress Signals. Plant Physiol. 162, 1566–1582. https://doi.org/10.1104/pp.113.221911

Choi, H.-I., Hong, J.-H., Ha, J.-O., Kang, J.-Y., Kim, S.Y., 2000. ABFs, a family of ABA responsive element binding factors. J. Biol. Chem. 275, 1723–1730. https://doi.org/10.1074/jbc.275.3.1723

Clouse, S.D., Sasse, J.M., 1998. BRASSINOSTEROIDS: Essential Regulators of Plant Growth and Development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 427–451. https://doi.org/10.1146/annurev.arplant.49.1.427

Cominelli, E., Galbiati, M., Vavasseur, A., Conti, L., Sala, T., Vuylsteke, M., Leonhardt, N., Dellaporta, S.L., Tonelli, C., 2005. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. Curr. Biol. 15, 1196–1200. https://doi.org/10.1016/j.cub.2005.05.048

Cornejo, M.J., Luth, D., Blankenship, K.M., Anderson, O.D. and Blechl, A.E. 1993. Activity of a maize ubiquitin promoter in transgenic rice. Plant Mol. Biol. 23, 567–81.

Creux, N.M., Ranik, M., Berger, D.K., Myburg, A.A. 2008. Comparative analysis of orthologous cellulose synthase promoters from Arabidopsis, Populus and Eucalyptus: evidence of conserved regulatory elements in angiosperms, New Phytol. v.179, 722-737

Cserhati, M. 2015. Motif content comparison between monocot and dicot species. Genom Data. 17; 3:128-36. doi: 10.1016/j.gdata.2014.12.006

Cui, L., Feng, K., Wang, M., Wang, M., Deng, P., Song, W., Nie, X., 2016. Genomewide identification, phylogeny and expression analysis of AP2/ERF transcription factors family in Brachypodium distachyon. BMC Genomics 17, 1–19. https://doi.org/10.1186/s12864-016-2968-8

De Boer, K., Tilleman, S., Pauwels, L., Vanden Bossche, R., De Sutter, V., Vanderhaeghen, R., Hilson, P., Hamill, J.D., Goossens, A., 2011. APETALA2/ETHYLENE RESPONSE FACTOR and basic helix-loop-helix tobacco transcription factors cooperatively mediate jasmonate-elicited nicotine biosynthesis. Plant J. 66, 1053–1065. https://doi.org/10.1111/j.1365-313X.2011.04566.x

Dey, S., Corina Vlot, A., 2015. Ethylene responsive factors in the orchestration of stress responses in monocotyledonous plants. Front. Plant Sci. 6, 1–7. https://doi.org/10.3389/fpls.2015.00640

Dombrecht, B., Xue, G.P., Sprague, S.J., Kirkegaard, J.A., Ross, J.J., Reid, J.B., Fitt, G.P., Sewelam, N., Schenk, P.M., Manners, J.M., Kazan, K., 2007. MYC2 Differentially Modulates Diverse Jasmonate-Dependent Functions in Arabidopsis. Plant Cell 19, 2225–2245. https://doi.org/10.1105/tpc.106.048017

Du, H., Zhang, L., Liu, L., Tang, X.F., Yang, W.J., Wu, Y.M., Huang, Y.B., Tang, Y.X., 2009. Biochemical and molecular characterization of plant MYB transcription factor family. Biochem. 74, 1–11. https://doi.org/BCM74010005 [pii]

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., Lepiniec, L., 2010. MYB transcription factors in Arabidopsis. Trends Plant Sci. 15, 573–581. https://doi.org/10.1016/j.tplants.2010.06.005

Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, S., Seki, M., Shinozaki, K., Yamaguchi Shinozaki, K., 2003. OsDREB genes in rice, Oryza sativa L., encode transcription activators that function in drought, high salt and cold responsive gene expression. Plant J. 33, 751–763.

Elmayan T., Tepfer M. 1995. Evaluation in tobacco of the organ specificity and strength of the rol D promoter, domain A of the 35S promoter and the 35S2 promoter. Transgenic Res. 4, 388–396.

Fehlberg, V., Vieweg, M.F., Dohmann, E.M.N., Hohnjec, N., Pühler, A., Perlick, A.M., Küster, H. 2005. The promoter of the leghaemoglobin gene VfLb29: Functional analysis and identification of modules necessary for its activation in the infected cells of root nodules and in the arbuscule-containing cells of mycorrhizal roots. Journal of Experimental Botany. 413:799–806; https://doi.org/10.1093/jxb/eri074

Filiz, E., Ozyigit, I.I., Vatansever, R. 2015. Genome-wide identification of galactinol synthase (GolS) genes in *Solanum lycopersicum* and *Brachypodium distachyon*. Comput Biol Chem. 58:149-57. doi: 10.1016/j.compbiolchem.2015.07.006.

Fincher, G.B. 1989. Molecular and celluar biology associated with endosperm mobilization in germinating cereal grains. Annu. Rev. Plant Physiol. 40, 305–346.

Franco-Zorrilla, J.M., López-Vidriero, I., Carrasco, J.L., Godoy, M., Vera, P., Solano, R., 2014. DNA-binding specificities of plant transcription factors and their potential to define target genes. Proc. Natl. Acad. Sci. 111, 2367–2372. https://doi.org/10.1073/pnas.1316278111

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., Ohme-Takagi, M., 2000. Arabidopsis Ethylene-Responsive Element Binding Factors Act as Transcriptional Activators or Repressors of GCC Box-Mediated Gene Expression. Plant Cell 12, 393–404. https://doi.org/10.1105/tpc.12.3.393

González-Morales, S.I., Chávez-Montes, R.A., Hayano-Kanashiro, C., Alejo-Jacuinde, G., Rico-Cambron, T.Y., de Folter, S., Herrera-Estrella, L., 2016. Regulatory network analysis reveals novel regulators of seed desiccation tolerance in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. 113, E5232–E5241. https://doi.org/10.1073/pnas.1610985113

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., Westhoff, P., 2004. cis-Regulatory Elements for Mesophyll-Specific Gene Expression in the C 4 Plant *Flaveria trinervia*, the Promoter of the C 4 Phosphoenolpyruvate Carboxylase Gene. Plant Cell 16, 1077–1090. https://doi.org/10.1105/tpc.019729.

Gu, C., Guo, Z.H., Hao, P.P., Wang, G.M., Jin, Z.M., Zhang, S.L. Multiple regulatory roles of AP2/ERF transcription factor in angiosperm. 2017. Bot Stud.58(1):6. doi: 10.1186/s40529-016-0159-1

Hanano, S., Stracke, R., Jakoby, M., Merkle, T., Domagalska, M.A., Weisshaar, B., Davis, S.J., 2008. A systematic survey in Arabidopsis thaliana of transcription factors that modulate circadian parameters. BMC Genomics 9, 1–13. https://doi.org/10.1186/1471-2164-9-182

Hanson, J., 2000. Functional characterization of the pointed cotyledon subclass of HDZip genes in Arabidopsis thaliana. UPPSALA.

Hartmann, U., Sagasser, M., Mehrtens, F., Stracke, R., Weisshaar, B. 2005. Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. Plant Mol. Biol. 57, 155–171. doi 10.1007/s11103-004-6910-0

Hatorangan, M.R., Sentausa, E., Wijaya, G.Y., 2009. In silico identification of cisregulatory elements of phosphate transporter genes in rice (Oryza sativa L.). J. Crop Sci. Biotechnol. 12, 25–30. https://doi.org/10.1007/s12892-008-0054-8

Heidecker, G., Messing, J. 1986. Structural analysis of plant genes. Annual Review of Plant Physiology, 37, 439-466.

Hemm, M.R., Herrmann, K.M., Chapple, C., 2001. AtMYB4: A transcription factor general in the battle against UV. Trends Plant Sci. 6, 135–136. https://doi.org/10.1016/S1360-1385(01)01915-X

Henriksson, E., Olsson, A.S.B., Johannesson, H., Johansson, H., Hanson, J., Engström, P., Söderman, E., 2005. Homeodomain leucine zipper class I genes in Arabidopsis.

Expression patterns and phylogenetic relationships. Plant Physiol. 139, 509–518. https://doi.org/10.1104/pp.105.063461.et

Hermann, S., Saarikettu, J., Onions, J., Hughes, K., Grundström, T., 1998. Calcium regulation of basic helix-loop-helix transcription factors. Cell Calcium 23, 135–142. https://doi.org/10.1016/S0143-4160(98)90112-9

Hernandez-Garcia, C.M., Finer, J.J., 2014. Identification and validation of promoters and cis-acting regulatory elements. Plant Sci. 217–218, 109–119. https://doi.org/10.1016/j.plantsci.2013.12.007

Higo, K., Ugawa, Y., Iwamoto, M., Korenaga, T., 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 27, 297–300. https://doi.org/10.1093/nar/27.1.297

Izui, K., Matsumura, H., Furumoto, T., Kai, Y., 2004. PHOSPHO *ENOL* PYRUVATE CARBOXYLASE: A New Era of Structural Biology. Annu. Rev. Plant Biol. 55, 69–84. https://doi.org/10.1146/annurev.arplant.55.031903.141619

Jacobsen, J.V., Beach, L.R. 1985. Control of transcription of a-amylase and rRNA genes in barley aleurone protoplasts by gibberellin and abscisic acid. Nature, 316, 275–277.

Jang, I.C., Choi, W.B., Lee, K.H., Song, S.I., Nahm, B.H., Kim, J.K. 2002. High-level and ubiquitous expression of the rice cytochrome c gene OsCc1 and its promoter activity in transgenic plants provides a useful promoter for transgenesis of monocots. Plant Physiology. 129:1473–1481. doi 10.1104/pp.002261

Jeon, J-S., Lee, S., Jung, K-H., Jun, S-H., Kim, C., An, G. 2000. Tissue-preferential expression of a rice α-tubulin gene, OsTubA1, mediated by the first intron. Plant Physiology.123: 1005–1014. doi 10.1104/pp.123.3.1005

Jiang, Y., Yang, B., Deyholos, M.K., 2009. Functional characterization of the Arabidopsis bHLH92 transcription factor in abiotic stress. Mol. Genet. Genomics 282, 503–516. https://doi.org/10.1007/s00438-009-0481-3

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Toralli, C., Weisshaar, B., Martin, C., 2000. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. EMBO J. 19, 6150–6161. https://doi.org/10.1093/emboj/19.22.6150

Jin, J., He, K., Tang, X., Li, Z., Lv, L., Zhao, Y., Luo, J., Gao, G., 2015. An arabidopsis transcriptional regulatory map reveals distinct functional and evolutionary features of novel transcription factors. Mol. Biol. Evol. 32, 1767–1773. https://doi.org/10.1093/molbev/msv058

Joseleau, J.P., Comtat, J., Ruel, K. 1992. Chemical structure of xylans and their interaction in the plant cell walls. In: Visser, J., Beldman, G., Kusters-Van Someren, M. A., Voragen, A. G. J. Xylans and xylanases: Progress in Biotechnology. New York: Elsevier, p. 1-15.

Jung, C., Seo, J.S., Han, S.W., Koo, Y.J., Kim, C.H., Song, S.I., Nahm, B.H., Choi,

Y.D., Cheong, J.-J., 2008. Overexpression of AtMYB44 Enhances Stomatal Closure to Confer Abiotic Stress Tolerance in Transgenic Arabidopsis. Plant Physiol. 146, 623–635. https://doi.org/10.1104/pp.107.110981

Kaplan, B., Davydov, O., Knight, H., Galon, Y., Knight, M.R., Fluhr, R., Fromm, H., 2006. Rapid Transcriptome Changes Induced by Cytosolic Ca2+ Transients Reveal ABRE-Related Sequences as Ca2+-Responsive cis Elements in Arabidopsis. Plant Cell 18, 2733–2748. https://doi.org/10.1105/tpc.106.042713

Kim, H.J., Kim, Y.K., Park, J.Y., Kim, J. 2002. Light signaling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in Arabidopsis thaliana. Plant J., 29, pp. 693-704

Kim, J., Kim, H.Y., 2006. Functional analysis of a calcium-binding transcription factor involved in plant salt stress signaling. FEBS Lett. 580, 5251–5256. https://doi.org/10.1016/j.febslet.2006.08.050

Kisu, Y., Ono, T., Shimofurutani, N., Suzuki, M., Esaka, M.1998. Characterization and expression of a new class of zinc finger protein that binds to silencer region of ascorbate oxidase gene. Plant Cell Physiol. 39:1054–1064

Lacombe, E., Doorsselaere Van, J., Boerjan, W., Boudet, A. M., Grima-Pettenati, J. 2000. Characterization of cis-elements required for vascular expression of the Cinnamoyl CoA Reductase gene and for protein-DNA complex formation The Plant Journal, 23, 663-676

Lakhwani, D., Pandey, A., Dhar, Y.V., Bag, S.K., Trivedi, P.K., Asif, M.H., 2016. Genome-wide analysis of the AP2/ERF family in Musa species reveals divergence and neofunctionalisation during evolution. Sci. Rep. 6, 1–17. https://doi.org/10.1038/srep18878

Lam, E., Chua, N.-H. 1991. Tetramer of a 21 base pair synthetic element confers seed expression and transcriptional enhancement in response to water stress and abscisic acid. J Biol Chem. 266: 17131-17135

Lee, B., Henderson, D.A., Zhu, J.-K., 2005. The Arabidopsis Cold-Responsive Transcriptome and Its Regulation by ICE1. Plant Cell 17, 3155–3175. https://doi.org/10.1105/tpc.105.035568

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Peer, Y. Van de, Rouzé, P., Rombauts, S., 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30, 325–327. https://doi.org/10.1093/nar/30.1.325

Leung J., Giraudat J. 1998. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49:199–222.

Liang, Y.K., Dubos, C., Dodd, I.C., Holroyd, G.H., Hetherington, A.M., Campbell, M.M., 2005. AtMYB61, an R2R3-MYB transcription factor controlling stomatal aperture in Arabidopsis thaliana. Curr. Biol. 15, 1201–1206. https://doi.org/10.1016/j.cub.2005.06.041

Licausi, F., Van Dongen, J.T., Giuntoli, B., Novi, G., Santaniello, A., Geigenberger, P., Perata, P., 2010. HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in Arabidopsis thaliana. Plant J. 62, 302–315. https://doi.org/10.1111/j.1365-313X.2010.04149.x

Lippold, F., Sanchez, D.H., Musialak, M., Schlereth, A., Scheible, W.-R., Hincha, D.K., Udvardi, M.K., 2009. AtMyb41 Regulates Transcriptional and Metabolic Responses to Osmotic Stress in Arabidopsis. Plant Physiol. 149, 1761–1772. https://doi.org/10.1104/pp.108.134874

Liu, F., Chang, X.J., Ye, Y., Xie, W.B., Wu, P., Lian, X.M., 2011. Comprehensive sequence and whole-life-cycle expression profile analysis of the phosphate transporter gene family in rice. Mol. Plant 4, 1105–1122. https://doi.org/10.1093/mp/ssr058

Liu, H.Y., Badarinarayana, V., Audino, D.C., Rappsilber, J., Mann, M., Denis, C.L., 1998. The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. EMBO J. 17, 1096–1106. https://doi.org/10.1093/emboj/17.4.1096

Liu, L.-J., Zhang, Y.-C., Li, Q.-H., Sang, Y., Mao, J., Lian, H.-L., Wang, L., Yang, H.-Q., 2008. COP1-Mediated Ubiquitination of CONSTANS Is Implicated in Cryptochrome Regulation of Flowering in Arabidopsis. Plant Cell 20, 292–306. https://doi.org/10.1105/tpc.107.057281

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Kazuko Yamaguchi-Shinozaki, K.S., 1998. Two Transcription Factors, DREB1 and DREB2, with an EREBP/AP2 DNA Binding Domain Separate Two Cellular Signal Transduction Pathways in Drought- and Low-Temperature-Responsive Gene Expression, Respectively, in Arabidopsis. Plant Cell 10, 1391–1406. https://doi.org/10.1105/tpc.10.8.1391

Liu, W., Tai, H., Li, S., Gao, W., Zhao, M., Xie, C., Li, W.X., 2014. bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. New Phytol. 201, 1192–1204. https://doi.org/10.1111/nph.12607

Liu, H.,Yu, X., Li, K., Klejnot, J.,Yang, H., Lisiero, D., Lin, C. 2008. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. Science. Dec 5,322(5907):1535-9. https://doi: 10.1126/science.1163927

Liu, D., Li, W., Chen, J., Hou, L. 2014. Expression analysis and functional characterization of a cold-responsive gene COR15A from Arabidopsis thaliana. Acta Physiol Plant 36: 2421– 2432

Liu, W., Stewart Junior, C.N. 2016. Plant synthetic promoters and transcription factors Curr Opin Biotechnol, 37:36–44. http://dx.doi.org/10.1016/j.copbio.2015.10.001

Loke, J.C., Stahlberg, E.A., Strenski, D.G., Haas, B.J., Wood, P.C., Li, Q.Q. Compilation of mRNA polyadenylation signals in Arabidopsis revealed a new signal element and potential secondary structures. Plant Physiol. 2005; 138:1457–1468. doi https://doi.org/10.1104/pp.105.060541

Maruyama, Y., Yamoto, N., Suzuki, Y., Chiba, Y., Yamazaki, K. ichi, Sato, T., Yamaguchi, J., 2013. The Arabidopsis transcriptional repressor ERF9 participates in resistance against necrotrophic fungi. Plant Sci. 213, 79–87. https://doi.org/10.1016/j.plantsci.2013.08.008

Masoudi-Nejad, A., Tonomura, K., Kawashima, S., Moriya, Y., Suzuki, M., Itoh, M., Kanehisa, M., Endo, T., Goto, S., 2006. EGassembler: Online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. Nucleic Acids Res. 34, 459–462. https://doi.org/10.1093/nar/gkl066

McElroy D., Blowers A.D., Jenes B., Wu R. 1991. Construction of expression vectors based on the rice actin1 (Act1) 5′ region for use in monocot transformation. Molecular and General Genetics. 231:150–160.

Mehrotra, R., Mehrotra, S. 2010. The location of the ACGT motifs in the promoter region of the gene was comprised from 1 kbp of the translation initial site (ATG). J Plant Physiol, 167, 1214-1218.

Moffat, C.S., Ingle, R.A., Wathugala, D.L., Saunders, N.J., Knight, H., Knight, M.R., 2012. ERF5 and ERF6 play redundant roles as positive regulators of JA/Et-mediated defense against botrytis cinerea in arabidopsis. PLoS One 7, 1–11. https://doi.org/10.1371/journal.pone.0035995

Morita, A., Umemura, T., Kuroyanagi, M., Futsuhara, Y., Perata, P., Yamaguchi, J. 1998. Functional dissection of a sugar-repressed a-amylase gene (RAmy1 A) promoter in rice embryos. FEBS Lett. 423, 81–85.

Nahar, K., Sunny, S.A., 2011. Extraction of biodiesel from a second generation energy crop (Jatropha curcas L.) by transesterification process. J. Environ. Sci. Technol. https://doi.org/10.3923/jest.2011.498.503

Nishiuchi, T., Shinshi, H., Suzuki, K. 2004. Rapid and transient activation of transcription of the ERF3 gene by wounding in tobacco leaves: Possible involvement of NtWRKYs and autorepression. Journal of Biological Chemistry 279:55355–55361; https://doi.org/10.1074/jbc.M409674200

Ohme-Takagi, M., Shinshi, H., 1995. Ethylene-Inducible DNA Binding Proteins That Interact with an Ethylene-Responsive Element. Plant Cell 7, 173–182. https://doi.org/10.1105/tpc.7.2.173

Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., Ohme-Takagi, M., 2001. Repression Domains of Class II ERF Transcriptional Repressors Share an Essential Motif for Active Repression. Plant Cell 13, 1959–1968. https://doi.org/10.1105/tpc.13.8.1959

Papi, M., Sabatini, S., Altamura, M.M., Hennig, L., Scha, E., Costantino, P., Vittorioso, P., 2002. Inactivation of the Phloem-Specific Dof Zinc Finger Gene. Plant Physiol. 128, 411–417. https://doi.org/10.1104/pp.010488.1

Parenicova, L., De Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C., Colombo, L. 2003. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell. 15:1538–1551. https://doi: 10.1105/tpc.011544

Park, D.H., Lim, P.O., Kim, J.S., Cho, D.S., Hong, S.H., Nam, H.G., 2003. The Arabidopsis COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. Plant J. 34, 161–171. https://doi.org/10.1046/j.1365313X.2003.01710.x

Park, H.C., Kim, M.L., Kang, Y.H., Jeon, J.M., Yoo, J.H., Kim, M.C. 2004. Pathogen- and NaCl-induced expression of the scam-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. Plant Physiology 135:2150–2161; https://doi.org/10.1104/pp.104.041442

Park, S.H., Yi, N., Kim, Y.S., Jeong M.H., Bang, S.W., Choi, Y.D., Kim, J.K. 2010a. Analysis of five novel putative constitutive gene promoters in transgenic rice plants. J. Exp. Bot., 61, 2459-2467. doi:10.1093/jxb/erq076

Park, M.R., Yun, K.Y., Mohanty, B., Herath, V., Xu, F., Wijaya, E., Bajic, V.B., Yun, S.J., de Los Reyes, B.G. 2010b. Supra-optimal expression of the cold-regulated OsMyb4 transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. Plant, Cell Environ. 33, 2209–2230. https://doi.org/10.1111/j.1365-3040.2010.02221.x

Perez-Rodriguez, P., Riano-Pachon, D.M., Correa, L.G.G., Rensing, S. A., Kersten, B., Mueller-Roeber, B. 2009. PlnTFDB: updated content and new features of the plant transcription factor database. Nucleic Acids Research. https://doi: 10.1093/nar/gkp805

Pegoraro, C., Farias, D. da R., Mertz, L.M., Santos, R.S. dos, Maia, L.C. da, Rombaldi, C.V., Oliveira, A.C. de, 2013. Ethylene response factors gene regulation and expression profiles under different stresses in rice. Theor. Exp. Plant Physiol. 25, 261–274. https://doi.org/10.1590/S2197-00252013000400004

Pilpel, Y., Sudarsanam, P., Church, G.M., 2001. Identifying regulatory networks by combinatorial analysis of promoter elements. Nat. Genet. 29, 153–159. https://doi.org/10.1038/ng724

Pirrello, J., Jaimes-Miranda, F., Sanchez-Ballesta, M.T., Tournier, B., Khalil-Ahmad, Q., Regad, F., Latché, A., Pech, J.C., Bouzayen, M., 2006. Sl-ERF2, a tomato ethylene response factor involved in ethylene response and seed germination. Plant Cell Physiol. 47, 1195–1205. https://doi.org/10.1093/pcp/pcj084

Plesch, G., Ehrhardt, T. Mueller-Roeber, B. 2001. Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. The Plant Journal, 28, 455-464. doi https://doi.org/10.1046/j.1365-313X.2001.01166.x

Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y.-Q., Shinozaki, K., Yamaguchi-, K., Shinozaki, 2004. Cloning and Functional Analysis of a Novel DREB1/CBF Transcription Factor Involved in Cold-Responsive Gene Expression in Zea mays L. Plant Cell Physiol. 45, 1042–1052. https://doi.org/10.1093/pcp/pch118

Quinn J.M., Barraco P., Eriksson M., Merchant S. 2000. Coordinate copper- and oxygen-responsive Cyc6 and Cpx1 expression in Chlamydomonas is mediated by the same element. Journal of Biological Chemist. 275, 6080-6089. doi 10.1074/jbc.275.9.6080

Ramalingam, A., Kudapa, H., Pazhamala, L.T., Garg, V., Varshney, R.K., 2015. Gene Expression and Yeast Two-Hybrid Studies of 1R-MYB Transcription Factor Mediating Drought Stress Response in Chickpea (Cicer arietinum L.). Front. Plant Sci. 6, 1117. https://doi.org/10.3389/fpls.2015.01117

Rao, G., Sui, J., Zeng, Y., He, C., Zhang, J., 2015. Genome-wide analysis of the AP2/ERF gene family in Salix arbutifolia. FEBS Open Bio 5, 132–137. https://doi.org/10.1016/j.fob.2015.02.002

Rashid, M., Guangyuan, H., Guangxiao, Y., Hussain, J., Xu, Y., 2012. AP2/ERF transcription factor in rice: Genome-wide anvas and yntenic relationships between monocots and udicots. Evol. Bioinforma. 2012, 321–355. https://doi.org/10.4137/EBO.S9369

Ré, D.A., Capella, M., Bonaventure, G., Chan, R.L., 2014. Arabidopsis AtHB7 and AtHB12 evolved divergently to fine tune processes associated with growth and responses to water stress. BMC Plant Biol. 14, 1–14. https://doi.org/10.1186/14712229-14-150

Reverter, A., McWilliam, S.M., Barris, W., Dalrymple, B.P., 2005. A rapid method for computationally inferring transcriptome coverage and microarray sensitivity. Bioinformatics 21, 80–89. https://doi.org/10.1093/bioinformatics/bth472

Reyes, J.C., Muro-Pastor, M.I., Florencio, F.J. 2004. The GATA family of transcription factors in Arabidopsis and rice. Plant Physiol. 134, 1718–1732. doi 10.1104/pp.103.037788

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R. 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105–2110. https://doi: 10.1126/science.290.5499.2105

Rogers, J.C., Lanahan, M.B., Rogers, S.W. 1994. The cis-acting gibberellin response complex in high-pI a-amylase gene promoters. Plant Physiol 105:151–158.

Rogers H. J., Bate N., Combe J., Sullivan J., Sweetman J., Swan C. 2001. Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene g10. Plant Mol. Biol. 45, 577–585.

Sandal, N.N., Bojsen, K., Marcker, K.A. 1987. A small family of nodule specific genes from soybean. Nucleic Acids Res. 15(4): 1507–1519. doi: 10.1093/nar/15.4.1507

Sakai, H., Aoyama, T., Oka, A. 2000. Arabidopsis ARR1 and ARR2 response regulators operate as transcriptional activators. Plant J. 24, 703–711.

Sharma, N., Russell, S.D., Bhalla, P.L., Singh, M.B., 2011. Putative cis-regulatory elements in genes highly expressed in rice sperm cells. BMC Res. Notes 4, 319. https://doi.org/10.1186/1756-0500-4-319

Sharma, N. 2011. Role of BHLH93 in Controlling Flowering Time in Arabidopsis Thaliana. Dissertation (Doctor of Philosophy)- Faculty of the Graduated School, University of Texas, Austin. 296 pg.

Shi, J.X., Malitsky, S., de Oliveira, S., Branigan, C., Franke, R.B., Schreiber, L., Aharoni, A., 2011. SHINE transcription factors act redundantly to pattern the archetypal surface of arabidopsis flower organs. PLoS Genet. 7. https://doi.org/10.1371/journal.pgen.1001388

Shin, D., Moon, S.-J., Han, S., Kim, B.-G., Park, S.R., Lee, S.-K., Yoon, H.-J., Lee, H.E., Kwon, H.-B., Baek, D., Yi, B.Y., Byun, M.-O., 2011. Expression of StMYB1R-1, a Novel Potato Single MYB-Like Domain Transcription Factor, Increases Drought Tolerance. Plant Physiol. 155, 421–432. https://doi.org/10.1104/pp.110.163634

Shirsat A., Wilford N., Croy R. Boulter D. 1989. Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol. Gen. Genet. 215, 326–331.

Simpson S.D., Nakashima K., Narusaka Y., Seki M., Shinozaki K., Yamaguchi-Shinozaki K. 2003. Two different novel cis-acting elements of erd1, a clpA homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. Plant J. 33, 259–270.

Son, O., Hur, Y.S., Kim, Y.K., Lee, H.J., Kim, S., Kim, M.R., Nam, K.H., Lee, M.S., Kim, B.Y., Park, J., Park, J., Lee, S.C., Hanada, A., Yamaguchi, S., Lee, I.J., Kim, S.K., Yun, D.J., Sderman, E., Cheon, C.I., 2010. ATHB12, an ABA-inducible homeodomainleucine zipper (HD-Zip) protein of arabidopsis, negatively regulates the growth of the inflorescence stem by decreasing the expression of a gibberellin 20-oxidase gene. Plant Cell Physiol. 51, 1537–1547. https://doi.org/10.1093/pcp/pcq108

Song, X., Li, Y., Hou, X. 2013. Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (Brassica rapa ssp. pekinensis). BMC Genomics 14:573.

Supek, F., Bošnjak, M., Škunca, N., Šmuc, T., 2011. Revigo summarizes and visualizes long lists of gene ontology terms. PLoS One 6. https://doi.org/10.1371/journal.pone.0021800

Sutoh, K., Yamauchi. D. 2003. Two cis-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds. Plant J. 34:635–45

Tamagnone, Merida, Parr, Mackay, Culianez-Macia, Roberts, Martin, 1998. The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. Plant Cell 10, 135–54. https://doi.org/10.1105/tpc.10.2.135

Terzaghi, W.B., Cashmore, A.R. 1995. Light-regulated transcription. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46, 445–474.

Tiwari, V., Patel, M.K., Chaturvedi, A.K., Mishra, A., Jha, B., 2016. Functional characterization of the tau class glutathione-S-transferases gene (SbGSTU) promoter of Salicornia brachiata under salinity and osmotic stress. PLoS One 11, 1–20. https://doi.org/10.1371/journal.pone.0148494

Tournier, B., Sanchez-Ballesta, M.T., Jones, B., Pesquet, E., Regad, F., Latché, A., Pech, J.C., Bouzayen, M., 2003. New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. FEBS Lett. 550, 149–154. https://doi.org/10.1016/S0014-5793(03)00757-9

Tran, L.P., Nakashima, K., Sakuma, Y., Simpson, S.D., Fujita, Y., Maruyama, K., Fujita, M., Seki, M., Shinozaki, K., Yamaguchi-shinozaki, K., 2004. Isolation and

Functional Analysis of Arabidopsis Stress-Inducible NAC Transcription Factors That Bind to a Drought-Responsive cis -Element in the early responsive to dehydration stress 1 Promoter. Plant Cell 16, 2481–2498. https://doi.org/10.1105/tpc.104.022699.Shinozaki

Ulmasov, T., Murfett, J., Hagen, G., Guilfoyle, T.J. 1997. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant Cell 9:1963–1971

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., Yamaguchi-Shinozaki, K., 2000. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. Proc. Natl. Acad. Sci. U. S. A. 97, 11632–11637. https://doi.org/10.1073/pnas.190309197

Urao, T., Kazuko Yamaguchi-Shinozaki, S.U., Shinozaki, K., 1993. An Arabidopsis myb Homolog Is Induced by Dehydration Stress and Its Gene Product Binds to the Conserved MYB Recognition Sequence. Plant Cell 5, 1529–1539. https://doi.org/10.1105/tpc.5.11.1529

Vahala, J., Felten, J., Love, J., Gorzsás, A., Gerber, L., Lamminmäki, A., Kangasjärvi, J., Sundberg, B., 2013. A genome-wide screen for ethylene-induced Ethylene Response Factors (ERFs) in hybrid aspen stem identifies ERF genes that modify stem growth and wood properties. New Phytol. 200, 511–522. https://doi.org/10.1111/nph.12386

Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L., Van de Peer, Y., 2009. Unraveling Transcriptional Control in Arabidopsis Using cis-Regulatory Elements and Coexpression Networks. Plant Physiol. 150, 535–546. https://doi.org/10.1104/pp.109.136028

Vannini, C., Locatelli, F., Bracale, M., Magnani, E., Marsoni, M., Osnato, M., Mattana, M., Baldoni, E., Coraggio, I., 2004. Overexpression of the rice Osmyb4 gene increases chilling and freezing tolerance of Arabidopsis thaliana plants. Plant J. 37, 115–127. https://doi.org/10.1046/j.1365-313X.2003.01938.x

Venter, M. 2007. Synthetic promoters: genetic control through cis engineering. Trends Plant Sci, 12(3), 118-12

Yadav, V., Kundu, S., Chattopadhyay, D., Negi, P., Wei, N., Deng, X.W., Chattopadhyay, S. 2002. Light regulated modulation of Z-box containing promoters by photoreceptors and downstream regulatory components, COP1 and HY5, in Arabidopsis. Plant J. 31, 741–753.

Yadav, V., Mallappa, C., Gangappa, S.N., Bhatia, S., Chattopadhyay, S. 2005. A basic helixloop-helix transcription factor in Arabidopsis, MYC2, acts as a repressor of blue light–mediated photomorphogenic growth. The Plant Cell.17(7):1953-66.

Walford, S.A., Wu, Y., Llewellyn, D.J., Dennis, E.S., 2011. GhMYB25-like: A key factor in early cotton fibre development. Plant J. 65, 785–797. https://doi.org/10.1111/j.1365-313X.2010.04464.x

Wang, J., Oard, J.H. 2003. Rice ubiquitin promoters: deletion analysis and potential usefulness in plant transformation systems. Plant Cell Reports. 22:129–134. doi 10.1007/s00299-003-0657-y

Wang, Y., Hua, J. 2009. A moderate decrease in temperature induces COR15a expression through the CBF signaling cascade and enhances freezing tolerance. Plant J. 60, 340–349.

Wang, Y., Wan, L., Zhang, L., Zhang, Z., Zhang, H., Quan, R., Zhou, S., Huang, R., 2012. An ethylene response factor OsWR1 responsive to drought stress transcriptionally activates wax synthesis related genes and increases wax production in rice. Plant Mol. Biol. 78, 275–288. https://doi.org/10.1007/s11103-011-9861-2

Wang, R., Zhu, M., Ye, R., Liu, Z., Zhou, F., Chen, H., Lin, Y. 2015. Novel green tissue-specific synthetic promoters and cis-regulatory elements in rice. Scientific Reports. 5: 18256. doi: 10.1038/srep18256

Wen, C.L., Cheng, Q., Zhao, L., Mao, A., Yang, J., Yu, S., Weng, Y., Xu, Y., 2016. Identification and characterisation of Dof transcription factors in the cucumber genome. Sci. Rep. 6, 1–11. https://doi.org/10.1038/srep23072

Whalley, H.J., Sargeant, A.W., Steele, J.F.C., Lacoere, T., Lamb, R., Saunders, N.J., Knight, H., Knight, M.R., 2011. Transcriptomic Analysis Reveals Calcium Regulation of Specific Promoter Motifs in *Arabidopsis*. Plant Cell 23, 4079–4095. https://doi.org/10.1105/tpc.111.090480

Won, S.-K., Lee, Y.-J., Lee, H.-Y., Heo, Y.-K., Cho, M., Cho, H.-T., 2009. cisElement- and Transcriptome-Based Screening of Root Hair-Specific Genes and Their Functional Characterization in Arabidopsis. Plant Physiol. 150, 1459–1473. https://doi.org/10.1104/pp.109.140905

Wu, H., Lv, H., Li, L., Liu, J., Mu, S., Li, X., Gao, J., 2015. Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of DREB genes in moso bamboo (Phyllostachys edulis). PLoS One 10, 1–22. https://doi.org/10.1371/journal.pone.0126657

Wu, R., Duan, L., Pruneda-Paz, J. L., Oh, D. H., Pound, M., Kay, S. 2018. The 6xABRE Synthetic Promoter Enables the Spatiotemporal Analysis of ABA462 Mediated Transcriptional Regulation. Plant Physiol. 177, 1650–1665.

Xiong, Y., Liu, T., Tian, C., Sun, S., Li, J., Chen, M., 2005. Transcription factors in rice: A genome-wide comparative analysis between monocots and eudicots. Plant Mol. Biol. 59, 191–203. https://doi.org/10.1007/s11103-005-6503-6

Yamaguchi-Shinozaki K, Shinozaki K. 1993. Characterization of the expression of a desiccation-responsive rd29 gene of Arabidopsis thaliana and analysis of its promoter in transgenic plants. Mol. Gen. Genet. 236:331–40

Yamaguchi-Shinozaki, K., Shinozaki, K. 1994. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. The Plant Cell Online, 6, 251-264. doi:10.1105/tpc.6.2.251

Yanagisawa, S. 2000. Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. Plant J. 21, 281–288. https://doi.org/10.1046/j.1365-313x.2000.00685.x

Yang, C.-Y., Hsu, F.-C., Li, J.-P., Wang, N.-N., Shih, M.-C., 2011. The AP2/ERF Transcription Factor AtERF73/HRE1 Modulates Ethylene Responses during Hypoxia in Arabidopsis. Plant Physiol. 156, 202–212. https://doi.org/10.1104/pp.111.172486

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., Zhangliang, C., Hongya, G., Li-Jia, Q., 2006. The MYB transcription factor superfamily of Arabidopsis: Expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. 60, 107–124. https://doi.org/10.1007/s11103-005-2910-y

Yevshin, I., Sharipov, R., Valeev, T., Kel, A., Kolpakov, F., 2017. GTRD: A database of transcription factor binding sites identified by ChIP-seq experiments. Nucleic Acids Res. 45, D61–D67. https://doi.org/10.1093/nar/gkw951

Yanagisawa S., Schmidt R. J. 1999. Diversity and similarity among recognition sequences of Dof transcription factors. Plant J. 17(2), 209–214.

Yu, D., Chen, C., Chen, Z. 2001. Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell. Jul;13(7):1527-40. https://doi.org/10.1146/annurev.phyto.42.040803.140421

Zhang, B., Chen, W., Foley, R.C., Büttner, M., Singh, K.B., 1995. Interactions between Distinct Types of DNA Binding Proteins Enhance Binding to ocs Element Promoter Sequences. Plant Cell 7, 2241–2252. https://doi.org/10.1105/tpc.7.12.2241

Zhang, Z.L., Xie, Z., Zou, X., Casaretto, J., David, H.T., Shen, Q.J. 2004. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol. 134(4): 1500–1513. doi 10.1104/pp.103.034967

Zhang, N., McHale, L.K., Finer, J.J. 2015. Isolation and characterization of "GmScream" promoters that regulate highly expressing soybean (Glycine max Merr.) genes. Plant Sci. 241:189–98. https://doi: 10.1016/j.plantsci.2015.10.010

Zhang, J., Yu, G., Wen, W., Ma, X., Xu, B., Huang, B., 2016. Functional characterization and hormonal regulation of the PHEOPHYTINASE gene LpPPH controlling leaf senescence in perennial ryegrass. J. Exp. Bot. 67, 935–945. https://doi.org/10.1093/jxb/erv509

Zheng, Y., Jiao, C., Sun, H., Rosli, H.G., Pombo, M.A., Zhang, P., Banf, M., Dai, X., Martin, G.B., Giovannoni, J.J., Zhao, P.X., Rhee, S.Y., Fei, Z., 2016. iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. Mol. Plant 9, 1667–1670. https://doi.org/10.1016/j.molp.2016.09.014

Zhu, Z., Gao, J., Yang, J.X., Wang, X.Y., Ren, G.D., Ding, Y.L., Kuai, B.K. 2015. Synthetic promoters consisting of defined cis-acting elements link multiple signaling pathways to probenazole-inducible system. J Zhejiang Univ Sci B. Apr;16(4):253-63. doi: 10.1631/jzus.B1400203

Table 1. Transcrption factor identify *in silico* characterized by identificate acess of the proteins and of the ESTs, (BlastX, e-value e^{-10}), famíly and genic identification of the sequence reference (RefSeq) of the NCBI.

Alinhament (BlastX, e-value^{-10})					
EST	**Protein ID**	**e-value**	**Transcription Factor**	**Family**	**Gene ID**
GT981049.1	XP_012081678.1	$1e^{-141}$	AP2-like ethylene-responsive transcription factor AIL6	AP2/ERF	105641695
GW877927.1	XP_012086183.1	$2e^{-113}$	AP2-like ethylene-responsive transcription factor At2g41710	AP2/ERF	105645242
Contig3256	XP_012085806.1	$6e^{-45}$	AP2-like ethylene-responsive transcription factor BBM	AP2/ERF	105644909
Contig855	XP_012090506.1	$9e^{-13}$	Ethylene-responsive transcription factor 2 (ERF002)	AP2/ERF	105648655
Contig986	NP_001295747.1	$5e^{-75}$	Ethylene-responsive transcription factor 3 (ERF003)	AP2/ERF	105633012
Contig3303	NP_001295676.1	$6e^{-19}$	Ethylene-responsive transcription factor 4 (ERF004)	AP2/ERF	105637492
GW617574.1	XP_012070057.1	$4e^{-137}$	Ethylene-responsive transcription factor 5 (ERF005)	AP2/ERF	105632315
GW880806.1	XP_012088236.1	$1e^{-108}$	Ethylene-responsive transcription factor 9 (ERF009)	AP2/ERF	105646903
FM890952.1	XP_012079279.1	$4e^{-12}$	Ethylene-responsive transcription factor 10 (ERF010)	AP2/ERF	105639746
Contig8323	XP_012068309.1	$3e^{-106}$	Ethylene-responsive transcription factor 11 (ERF011)	AP2/ERF	105630926
Contig9625	XP_012091650.1	$1e^{-21}$	Ethylene-responsive transcription factor 12 (ERF012)	AP2/ERF	105649572
Contig4551	XP_012077090.1	$2e^{-84}$	Ethylene-responsive transcription factor 18 (ERF018)	AP2/ERF	105638000
GW877349.1	XP_012083682.1	$6e^{-11}$	Ethylene-responsive transcription factor 34 (ERF034)	AP2/ERF	105643210
Contig4984	XP_012088313.1	$2e^{-16}$	Ethylene-responsive transcription factor 54 (ERF054)	AP2/ERF	105646962
Contig1625	XP_012082821.1	$2e^{-142}$	Ethylene-responsive transcription factor 73 (ERF073)	AP2/ERF	105642565
GT978427.1	XP_012073423.1	$1e^{-14}$	Ethylene-responsive transcription factor 84 (ERF084)	AP2/ERF	105635041
GW611399.1	XP_012073833.1	$2e^{-157}$	Auxin response factor 5 (ARF5)	ARF	105635367
GT973126.1	XP_012087135.1	$1e^{-53}$	Transcription factor bHLH51	bHLH	105645984
Contig1769	XP_012067264.1	0	Transcription factor bHLH63	bHLH	105630155
GW876147.1	XP_012075467.1	$4e^{-104}$	Transcription factor bHLH79	bHLH	105636735
GW881219.1	XP_012068085.1	$3e^{-25}$	Transcription factor bHLH93	bHLH	105630754
GT978533.1	XP_012078963.1	$9e^{-157}$	Transcription factor bHLH95	bHLH	105639494
GW878262.1	XP_012065925.1	$2e^{-93}$	Transcription factor bHLH121	bHLH	105629018
Contig1890	XP_012084104.1	$3e^{-43}$	Transcription factor bHLH147	bHLH	105643560
Contig4487	NP_001306848.1	$6e^{-116}$	Transcription factor ICE1	bHLH	105630202
Contig2723	XP_020533722.1	$6e^{-113}$	Transcription factor SPATULA (SPT)	bHLH	105630933
FM896789.1	XP_012079779.1	$2e^{-103}$	Abscisic acid-insensitive 5-like protein 1 (DPBF2)	bZIP	105640154
GT971719.1	NP_001295728.1	$1e^{-90}$	Abscisic acid-insensitive 5-like protein 2 (DPBF3)	bZIP	105644237
Contig5691	XP_012080771.1	$6e^{-99}$	bZIP transcription factor 16 (BZIP16)	bZIP	105640957
Contig1257	XP_012091145.1	$4e^{-13}$	Transcription factor RF2b	bZIP	105649181
FM890713.1	XP_012078126.1	$2e^{-120}$	Transcription factor TGA3	bZIP	105638859

FM888151.1	XP_020534070.1	$6e^{-74}$	Zinc-finger homeodomain protein 4 (ZHD4)	C2H2	105631573
GW875656.1	XP_012084590.1	$4e^{-83}$	Calmodulin-binding transcription activator 2 (CMTA2)	CAMTA	105643958
GT972837.1	XP_012084038.1	$4e^{-57}$	Calmodulin-binding transcription activator 5 (CMTA5)	CAMTA	105643515
GW618109.1	XP_020541430.1	$3e^{-135}$	General negative regulator of transcription subunit 3 (NOT3)	CNOT2/3/5	105650400
GW877859.1	XP_012092669.1	$3e^{-19}$	GATA transcription factor 1 (GATA1)	GATA	105650386
Contig4530	XP_012083638.1	$3e^{-38}$	GATA transcription factor 4 (GATA4)	GATA	105643177
Contig907	XP_012067181.1	$1e^{-52}$	GATA transcription factor 5 (GATA5)	GATA	105630095
Contig859	XP_012070091.1	$1e^{-27}$	GATA transcription factor 26 (GATA26)	GATA	105632343
FM894487.1	XP_012093243.1	$4e^{-53}$	GATA zinc finger domain-containing protein 7 (gtaG)	GATA	105650892
Contig1262	XP_012093246.1	$5e^{-55}$	GATA zinc finger domain-containing protein 9 (gtaI)	GATA	105650895
JK613187.1	XP_012079074.1	$7e^{-77}$	G-box-binding factor 1 (GBF1)	GBF	105639587
Contig2777	XP_012067143.1	$4e^{-153}$	G-box-binding factor 4 (GBF4)	GBF	105630069
Contig4454	XP_012081428.1	0	Scarecrow-like transcription factor PAT1	GRAS	105641485
GT972943.1	XP_012079593.1	$6e^{-123}$	Growth-regulating factor 2 (GRF2a2)	GRF	105639999
Contig1338	XP_012086689.2	$5e^{-77}$	Growth-regulating factor 7 (GRF7)	GRF	105645559
Contig45	XP_020535325.1	$9e^{-121}$	Transcriptional corepressor LEUNIG	GroTLE	105635776
GW612122.1	XP_012081491.1	$2e^{-136}$	General transcription factor IIE subunit 1	GTF	105641498
GW614059.1	XP_012067920.1	$3e^{-115}$	Homeobox-leucine zipper protein ATHB-6	HD-Zip	105630651
Contig3495	XP_012083130.1	$6e^{-60}$	Homeobox-leucine zipper protein ATHB-12	HD-Zip	105642795
Contig1993	XP_012091954.1	$4e^{-149}$	Homeobox-leucine zipper protein ATHB-54	HD-Zip	105649782
Contig4894	XP_012083669.1	$2e^{-69}$	Homeobox-leucine zipper protein HAT3	HD-Zip	105643198
Contig420	XP_012070041.1	0	Homeobox-leucine zipper protein HAT4	HD-Zip	105632304
Contig3931	XP_012080202.1	$2e^{-166}$	Homeobox-leucine zipper protein HAT5	HD-Zip	105640489
Contig3415	XP_012083177.1	$8e^{-73}$	Homeobox-leucine zipper protein HAT7	HD-Zip	105642825
GT971697.1	XP_012086347.1	$6e^{-43}$	Homeobox-leucine zipper protein HDG2	HD-Zip	105645373
Contig2217	XP_020536555.1	$7e^{-110}$	Heat shock factor protein HSF30	HSF	105638131
GW611053.1	XP_012081059.1	$2e^{-135}$	Heat shock factor protein HSF8	HSF	105641182
Contig3196	XP_012076960.1	$5e^{-98}$	Heat stress transcription factor A-2 (HSFA2)	HSF	105637902
Contig2804	XP_012076680.1	$2e^{-129}$	Heat stress transcription factor C-1 (HSFC1)	HSF	105637712
GT969949.1	XP_012085841.1	$2e^{-111}$	B3 domain-containing transcription factor ABI3	LAV	105644942
Contig4908	XP_020533551.1	$6e^{-70}$	B3 domain-containing transcription factor FUS3	LAV	105630373
FM895492.1	XP_012078073.1	$2e^{-122}$	B3 domain-containing transcription repressor VAL1	LAV	105638817
Contig1905	XP_012091051.1	$3e^{-42}$	B3 domain-containing transcription repressor VAL2	LAV	105649107
GW615638.1	XP_012088064.1	$4e^{-83}$	Developmental protein SEPALLATA 1	MADS-box	105646752
Contig2623	XP_012067820.1	$6e^{-47}$	MADS-box protein CMB1	MADS-box	105630574

GW879514.1	XP_012071151.1	$9e^{-25}$	Transcription factor APL	MYB	105633182
GW880385.1	XP_012089895.1	$1e^{-98}$	Myb-related protein 308 (MYB308)	MYB	110008593
Contig2137	XP_012068582.1	$3e^{-156}$	Transcription factor DIVARICATA	MYB	105631171
Contig1844	XP_012070162.1	$3e^{-132}$	Transcription factor DIVARICATA	MYB	105632398
Contig4593	XP_012068751.1	$4e^{-17}$	Transcription factor GAMYB	MYB	105631288
Contig2003	XP_012080161.1	0	Transcription factor MYB1R1	MYB	105640456
GW615075.1	XP_012071707.1	$6e^{-135}$	Transcription factor MYB25	MYB	105633683
Contig963	XP_012079312.1	$1e^{-145}$	Transcription factor MYB44	MYB	105639771
GW878023.1	XP_012083029.1	$2e^{-91}$	Transcription factor MYB86	MYB	105642721
GW875559.1	XP_012068953.1	$8e^{-123}$	Transcription factor MYB86	MYB	105631441
Contig1187	XP_012075785.1	$7e^{-130}$	Transcription repressor MYB6	MYB	105636993
Contig2221	XP_012070614.1	0	NAC domain-containing protein 18 (ANAC018)	NAC	105632781
GW614724.1	XP_012091551.1	$1e^{-116}$	NAC transcription factor 25 (AtNAC025)	NAC	105649499
Contig808	NP_001306860.1	$8e^{-129}$	NAC transcription factor 29 (AtNAC029)	NAC	105650420
GT973652.1	XP_012077008.1	$6e^{-108}$	Nuclear transcription factor Y subunit A-3 (NF-YA3)	NFY	105637932
Contig1324	XP_012086120.1	0	Nuclear transcription factor Y subunit A-9 (NF-YA9)	NFY	105645194
GW875226.1	XP_012081482.1	$5e^{-28}$	Nuclear transcription factor Y subunit B-3 (NF-YB3)	NFY	105641516
FM894510.1	XP_012071003.1	$2e^{-29}$	Nuclear transcription factor Y subunit B-9 (NF-YB9)	NFY	105633082
Contig3806	NP_001295673.1	$3e^{-119}$	Nuclear transcription factor Y subunit B-6 (NF-YB6)	NFY	105636210
Contig3340	XP_012087361.1	$3e^{-139}$	Nuclear transcription factor Y subunit C-2 (NF-YC2)	NFY	105646169
Contig3046	XP_012066097.1	$4e^{-160}$	Transcriptional corepressor SEUSS	SEU	105629170
Contig107	XP_012074200.1	$4e^{-145}$	Protein SAWADEE HOMEODOMAIN HOMOLOG 2	SHH2	105635724
GW880065.1	XP_012065876.1	$2e^{-60}$	Transcription elongation factor SPT4 homolog 2	SPT4	105628974
Contig2200	XP_012073569.1	$2e^{-131}$	SWI/SNF complex subunit SWI3B	SWI/SNF	105635174
GW613857.1	XP_012081764.1	$8e^{-87}$	SWI/SNF complex subunit SWI3D	SWI/SNF	105641778
GT981631.1	XP_020534043.1	$3e^{-103}$	Transcription initiation factor TFIID subunit 1 (TAF1)	TAF	105631625
GT973402.1	XP_012089643.1	$4e^{-24}$	Transcription initiation factor TFIID subunit 15 (TAF15)	TAF	105648008
GW610999.1	XP_012068247.1	$1e^{-140}$	Transcription initiation factor TFIID subunit 4b (TAF4B)	TAF	105630869
Contig3245	XP_012068877.1	$1e^{-128}$	Transcription initiation factor TFIID subunit 9 (TAF9)	TAF	105631389
GW875662.1	NP_001306852.1	$6e^{-127}$	Transcription factor TCP2	TCP	105634014
GT972269.1	XP_012074215.1	$7e^{-124}$	Transcription factor TCP3	TCP	105635738
GW878450.1	XP_012087026.1	$4e^{-50}$	Transcription factor TCP9	TCP	105645895
Contig3143	XP_012080249.1	$7e^{-36}$	Transcription factor TCP14	TCP	105640523
Contig4590	XP_012084808.1	$2e^{-36}$	Transcription factor TCP23	TCP	105644117
Contig918	XP_012085608.1	$3e^{-109}$	Trihelix transcription factor GT-1	Trihelix	105644757

JK610975.1	XP_012087841.1	$8e^{-62}$	Trihelix transcription factor GT-3b	Trihelix	105646582
Contig4446	XP_012091932.1	$5e^{-145}$	Single-stranded DNA-binding protein WHY1, chloroplastic	WHIRLY	105649768
Contig1315	XP_012076351.1	$5e^{-117}$	Probable WRKY transcription factor 17 (WRKY17)	WRKY	105637490
Contig2829	XP_012071503.1	0	Probable WRKY transcription factor 32 (WRKY32)	WRKY	105633511
Contig4153	XP_012091128.1	$1e^{-11}$	Probable WRKY transcription factor 70 (WRKY57)	WRKY	105649166

Table 2. The 25 more prevalent CAREs, based on the number of motifs and FT genes showing the motif in the respective promoter, the motif sequence, and some details from the reference associated.

CARE	Motifs detected	TF Genes	Motif Sequence	Functions	Reference
DOFCOREZM	1340	105	AAAG	Core site required for binding of Dof proteins	Yanagisawa and Schmidt (1999)
GT1CONSENSUS	992	104	GRWAAW	Consensus GT-1 binding site in many light-regulated genes	Buchel et al. (1999)
CAATBOX1	978	106	CAAT	Seed storage protein gene	Shirsat et al. (1989)
CACTFTPPCA1	973	106	YACT	Mesophyll-specific gene expression in the C4 plant	Gowik et al. (2004)
ARR1AT	858	105	NGATT	Response regulators operate as transcriptional activators	Sakai et al. (2000)
ROOTMOTIFTAPOX1	858	105	ATATT	Motif found in promoter of *rolD* gene	Elmayan and Tepfer (1995)
GATABOX	708	105	GATA	Conserved GATA motif	Reyes et al. (2004)
POLLEN1LELAT52	645	104	AGAAA	Binding site required for pollen specific expression	Bate and Twell (1998)
MYCCONSENSUSAT	553	105	CANNTG	Recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis	Abe et al. (2003)
EBOXBNNAPA	537	105	CANNTG	Control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes	Hartmann et al. (2005)
POLASIG1	526	102	AATAAA	Poly A signal found in legA gene of pea, rice alpha-amylase. Near upstream elements (NUE) in Arabidopsis	Loke et al. (2005)
MARTBOX	494	86	TTWTWTTWTT	Most common element in flowering plants and is suggested to play role in transcriptional regulation	Cserhati (2015)
POLASIG3	446	96	AATAAT	Plant polyA signal, consensus sequence for plant polyadenylation signal	Heidecker and Messing (1986)
GTGANTG10	429	100	GTGA	Regulation of late pollen genes	Rogers et al. (2001)
WRKY71OS	377	97	TGAC	A core W-box, binding site of rice WRKY71, a transcriptional repressor of the gibberellin signaling pathway	Zhang et al. (2004)
SEF4MOTIFGM7S	295	98	RTTTTTR	Sequence found in promoter region of beta-conglycinin (7S globulin) gene, SEF4 (soybean embryo factor 4) binding motif	Allen et al. (1989)

IBOXCORE	294	100	GATAA	Conserved sequence upstream of light-regulated genes of both monocots and dicots	Terzaghi and Cashmore (1995)
TAAAGSTKST1	267	101	TAAAG	Motif found in promoter of KST1 gene; Target site for StDof1 protein controlling guard cell-specific gene expression	Plesch et al. (2001)
GT1GMSCAM4	259	93	GAAAAA	Plays a role in pathogen- and salt-induced CaM-4 gene expression	Park et al. (2004)
OSE2ROOTNODULE	242	88	AAAGAT	One of the motifs of organ-specific elements characteristic of the promoters activated in infected cells of root nodules	Fehlberg et al. (2005)
NODCON2GM	242	88	CTCTT	One of two putative nodulin consensus sequences	Sandal et al. (1987)
CURECORECR	232	95	GTAC	CuRE (copper-response element) involved in oxygen-response of Cyc6 and Cpx1 genes in Chlamydomonas	Quinn et al. (2000)
ACGTATERD1	218	84	ACGT	Required for etiolation-induced expression of erd1 (early responsive to dehydration) in Arabidopsis	Simpson et al. (2003)
WBOXNTERF3	200	88	TGACY	Maybe involved in activation of a transcriptional repressor ERF3 gene by wounding in tobacco	Nishiuchi et al. (2004)
POLASIG2	193	90	AATTAAA	Poly A signal potential candidate for stable transcripts generated *SlGolS3* and *SlGolS4* genes in *Solanum lycopersicum*	Filiz et al. (2015)

5 CONSIDERAÇÕES FINAIS:

O presente estudo permitiu verificar a resposta global da modulação de genes expressos em raízes de pinhão-manso (*Jatropha curcas* L.) envolvidos na resposta ao estresse de salinidade (150mMol de NaCl / 3h). Estas análises formam o primeiro passo para identificação de interações regulatórias destes genes durante o desenvolvimento ou sob condições de estresse. Foram detectados representantes de várias famílias de FTs a partir de transcritos identificados por RNA-seq e prováveis genes de FTs que controlam a expressão desses transcritos RNA-seq em resposta ao estresse aplicado, destacando os FTs responsáveis por ampla quantidade de regulações, em termos de alvos super-representativos nos genes induzidos, reprimidos e não diferencialmente expressos. Assim, a realização da análise funcional, consequente da reprogramação transcricional do conjunto de FTs expressos em raízes pinhão-manso aqui apresentados, pode sugerir grande potencial biotecnológico, servindo para disponibilização de valiosos recursos, como candidatos potenciais para transgenia, visando aumentar à eficácia da resposta à tolerância específica ao estresse estudado. Estes resultados auxiliam na compreensão dos mecanismos moleculares de respostas à salinidade e tolerância em pinhão-manso.

Adicionalmente, foi realizado a investigação *in silico* de genes de FTs derivados do banco público (NCBI) de ESTs, com o mapeamento dos seus respectivos elementos cis-regulatórios [1500 pb upstream do sítio inicial de transcrição (TSS)]. Os CAREs associados especificamente com função constitutiva e de desenvolvimento, foram detectados ao longo da região promotora dos genes FTs, enquanto os CAREs responsíveis a estresses e estímulos de ABA são distribuídos de forma pontual ao longo dos setores. Desta forma, a identificação da localização dos motivos com suas respectivas quantificações pode ajudar a compreender a influência dos CAREs na expressão dos genes, durante o desenvolvimento normal da planta e especialmente pela resposta a estímulos ambientais. Logo, a importância desta abordagem se apoia na proposta de desenvolvimento de promotores sintéticos, que possam aumentar a eficiência da transcrição de genes FTs, principalmente aos de respostas a estresses, os quais podem ser utilizados em programas de melhoramento de plantas.

REFERÊNCIAS

ADAMS, C.C.; WORKMAN, J.L. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. **Mol. Cell. Biol**., 15, 1405-1421. 1995.

ADAMS, M.D.; SOARES, M.B.; KERLAVAGE, A.R.; FIELDS, C.; VENTER, J.C. Rapid cDNA sequencing (*expressed sequence tags*) from a directionally cloned human infant brain cDNA library. **Nat. Genet.**, 4: 373–380. 1993.

ADDO-QUAYE, C.; ESHOO, T.W.; BARTEL, D.P.; AXTELL, M.J. Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. **Curr. Biol.** 18: 758–762. 2008.

AGALOU, A.; PURWANTOMO, S.; OVERNÄS, E.; JOHANNESSON, H.; ZHU, X.; ESTIATI, A.; DE KAM, R.J.; ENGSTRÖM, P.; SLAMET-LOEDIN, I.H.; ZHU, Z.; WANG, M.; XIONG, L.; MEIJER, A.H.; OUWERKERK, P.B. A genome-wide survey of HD-Zip genes in rice and analysis of drought-responsive family members. **Plant Molecular Biology**, v.66, 87-103, 2008.

AHMAD, P.; AZOOZ, M.M.; PRASAD, M.N.V. **Salt Stress in Plants**: Signalling, Omics and Adaptations Science, Springer Science e Business Media, New York, USA. 2013.

AIDA, M.; ISHIDA, T.; FUKAKI, H.; FUJISAWA, H.; TASAKA, M. Genes involved in organ separation in Arabidopsis: An analysis of the cup-shaped cotyledon mutant. **Plant Cell**, 9, 841–857. 1997.

ALBRECHT, V.; RITZ, O.; LINDER, S.; HARTER, K.; KUDLA, J. The NAF domain defines a novel protein-protein interaction module conserved in Ca^{2+} regulated kinases. **EMBO J**. 20: 1051–1063. 2001.

ALLEN, M.D.; YAMASAKI, K.; OHME-TAKAGI, M.; TATENO, M.; SUZUKI, M. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. **EMBO Journal**, 17, 5484–5496. 1998.

ALONSO, J.M.; STEPANOVA, A.N.; LEISSE, T.J. Genome-wide insertional mutagenesis of Arabidopsis thaliana. **Science**, 301. 653-657. 2003.

ALSCHER, R.G., N. ERTURK AND L.S. HEATH. Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. **J. Exp. Biol.**, 53: 1331-1341. 2002.

ALSCHER, R.G.; DONAHUE, J.L.; CRAMER, C.L. Reactive oxygen species and antioxidants: relationship in green cells. **Physiol. Plant**., 100:224-233. 1997.

ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res**., 25:3389–3402. 1997.

ALWINE, J.C. KEMP, D.J. STARH, G.R. Method for the detection of specific RNAs in agarose gels by transfer to diazobenyloxymethyl paper and hybridization with DNA probes. **Proc. Nad. Acad. Sci**., USA, 74,5350-5354. 1977.

AMARAL, A.M.; BAPTISTIA, J.C.; WINCK, F.V.; HOMEM, R.A.; MACHADO, M.A. Plataformas tecnológicas no estudo da bactéria causadora do cancro cítrico: genômica, transcriptômica e proteômica. **Laranja**. v. 27, n. 2, p. 355-372, 2006.

AMBAWAT, S.; SHARMA, P.; YADAV, N.R.; YADAV, R.C. MYB transcription factor genes as regulators for plant responses: an overview. **Physiol. Mol. Biol. Plants**., 19: 307–321. 2013.

APWEILER, R.; ARMSTRONG, R.N.; BAIROCH, A.; CORNISH-BOWDEN, A.; HALLING, P.J.; HOFMEYR, J.H.S.; KETTNER, C.; LEYH, T.S.; ROHWER, J.; SCHOMBURG, D.; STEINBECK, C.; TIPTON, K.F. A large-scale protein-function database. **Nat. Chem. Biol**. 6, 785. 2010.

ARIEL, F., DIET, A., VERDENAUD, M., GRUBER, V., FRUGIER, F., CHAN, R. Environmental regulation of lateral root emergence in *Medicago truncatula* requires the HD-Zip I transcription factor HB1. **Plant Cell**, 22: 2171–2183. 2010.

ARIEL, F.D.; MANAVELLA, P.A.; DEZAR, C.A.; CHAN, R.L. The true story of the HD-Zip family. **Trends in Plant Science**, 12: 419–426. 2007.

ARORA, A.; SAIRAM, R.K.; SRIUASTAVA, G.C. Oxidative stress and antioxidative system in plants. **Current Science**, 82: 1227–1238. 2002.

ARRUDA, F.P.; BELTRÃO, N.E.M.; ANDRADE, A.P.; PEREIRA, W.E.; SEVERINO, L.S. Cultivo de pinhão-manso (*Jatropha curcas* L.) como alternativa para o semiárido nordestino. **R. Bras. Oleag. Fibrosas**, 8:789-799, 2004.

ASADA, K. Ascorbate peroxidase: a hydrogen peroxide-scavenging enzyme in plants. **Physiol Plant.**, 85: 235–241. 1992.

ASADA, K. The water–water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. **Annual Review of Plant Physiology and Plant Molecular Biology**, 50, pp. 601-639. 1999.

ASHLEY, M.B.; DORMAN, E.R.; CORCES, V.G. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. **Mol Cell**, 32(1):1–9, 2008.

ASHRAF, M. Some important physiological selection criteria for salt tolerance in plants. **Flora**., 199:361-376. 2004.

ASHRAF, M.; FOOLAD, M.R. Roles of glycine betaine and proline in improving plant abiotic stress resistance. **Environmental and Experimental Botany**, v.59, n.2, p.206-216, 2007.

ASHRAF, M.; HARRIS, J.C. Potential biochemical indicators of salinity tolerance in plants. **Plant Science**, 166: 3–16. 2004.

ATCHLEY, W.R.; FITCH, W.M. A natural classification of the basic helix-loop-helix class of transcription Factors. **Proc. Natl. Acad. Sci.** USA. 94:5172–5176. 1997.

ATCHLEY, W.R.; THERHALLE, W.; DRESS, A. Positional dependence, cliques and predictive motifs in the bHLH protein domain. **J. Mol. Evol.**, 48, 501–516. 1999.

AVELAR, R.C.; DEPERON J.R.; CARVALHO J.P.F. Produção de mudas de pinhão-manso (*Jatropha curcas*) em tubetes. In: CONGRESSO DA REDE BRASILEIRA DE TECNOLOGIA DE BIODIESEL, 1., 2006, Brasília, DF. **Anais**.... Brasília, DF: ABIPTI, 2006. p.137-139.

AZAIZEH, H.; GUNSE, B.; STEUDLE, E. Effects of NaCl and $CaCl_2$ on water transport across root cells of maize (*Zea mays* L.) seedlings, **Plant Physiol**. 99, 886-894. 1992.

BAILEY, P.C.; MARTIN, C.; TOLEDO-ORTIZ, G.; QUAIL, P.H.; HUQ, E.; HEIM, M.A.; JAKOBY, M.; WERBER, M.; WEISSHAAR, B. Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana. **Plant Cell**., 15: 2497–2501. 2003.

BARANOWSKIJ, N.; FROHBERG, C.; PRAT, S.; WILLMITZER, L. A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. **EMBO J.**, 13, 5383–5392. 1994.

BARBA, M.; CZOSNEK, H.; HADIDI, A. Historical perspective, development and applications of next-generation sequencing in plant virology. **Viruses**., v. 6, n. 1, p. 106-36. 2014.

BARBIER-BRYGOO, H.; VINAUGER, M.; COLCOMBET, J.; EPHRITIKHINE, G.; FRACHISSE, J.; MAUREL, C. Anion channels in higher plants: functional characterization, molecular structure and physiological role. **Biochim. Biophys. Acta.** 1465: 199–218. 2000.

BARNES, M.R.; GRAY, I.C. **Bioinformatics for Geneticists**, ISBN: 0-470-84393-4. 2003.350p.

BARNES, M.R.; GRAY, I.C. **Bioinformatics for geneticists**. 1st ed. New York. Wiley, 2003. 422pp.

BAUMANN, K.; RODRIGUEZ, M.P.; BRADLEY, D.; VENAIL, J.; BAILEY, P.; JIN, H.; KOES, R.; ROBERTS, K.; MARTIN, C. Control of cell and petal morphogenesis by R2R3 MYB transcription factors. **Development**., 134:1691–1701. 2007.

BAYAT, A. Science, Medicine and the Future: Bioinformatics. **British Medical Journal**, 324, 1018-1022. 2002.

BELTRÃO, N.E.M. Agronegócio das oleaginosas no Brasil. **Inf. Agropec**., 26:44-78, 2005.

BENDERS, G.A.; POWELL, B.C.; HUTCHISON, C.A. Transcriptional analysis of the conserved ftsZ gene cluster in *Mycoplasma genitalium* and *Mycoplasma pneumoniae*. **J. Bacteriol.**, 187: 4542–4551. 2005.

BENFEY, P.N.; CHUA, N.H. The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. **Science**., 250, 959–966. 1990.

BENSON, D.A.; KARSCH-MIZRACHI, I.; LIPMAN, D.J.; OSTELL, J.; SAYERS, E.W. GenBank. **Nucleic Acids Research**, 39, D32–37. 2011.

BLAHA, G.; STELZL, U.; SPAHN, C.M.T.; AGRAWAL, R.K.; FRANK, J.; NIERHAUS, K.H. Preparation of functional ribosomal complexes and effect of buffer conditions on tRNA positions observed by cryoelectron microscopy. **Methods in Enzymology,** 317, 292–306. 2000.

BLUMWALD, E. Sodium transport and salt tolerance in plants. **Curr. Opin. Cell Biol**. 12, 431–434. 2000. doi: 10.1016/S0955-0674(00) 00112-5

BLUMWALD, E.; AHARON, G.S.; APSE, M.P. Sodium transport in plant cells. **Biochim. Biophys. Acta.** 1465, 140–151. 2000.

BOUTILIER, K.; OFFRINGA, R.; SHARMA, V.K.; KIEFT, H.; OUELLET, T.; ZHANG, L.; HATTORI, J.; LIU, C.-M.; VAN LAMMEREN, A.A.M.; MIKI, B.L.A. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. **Plant Cell**., 14: 1737–1749. 2002.

BRASIL. Ministério da Indústria e do Comércio. Secretária de Tecnologia Industrial. **Produção de combustíveis líquidos a partir de óleos vegetais**. Brasília: STI/CIT, 1985. 364p.

BREUSEGEM, F.V.; VRANOVÁ, E.; DAT, J.F.; INZÉ, D. The role of active oxygen species in plant signal transduction. Plant Science 161:405-414. 2001.

BRINI, F.; GAXIOLA, R.A.; BERKOWITZ, G.A.; MASMOUDI, K. Cloning and characterization of a wheat vacuolar cation/proton antiporter and pyrophosphatase proton pump. **Plant Physiol. Biochem**. 43, 347–354. 2005.

BROWN, R.B.; MADRID, N.J.; SUZUKI, H.; NESS, S.A. Optimized approach for Ion Proton RNA sequencing reveals details of RNA splicing and editing features of the transcriptome. **PLoS One**. 2017; 12:e0176675. https://doi.org/10.1371/journal.pone.0176675

BULYK, M.L.; JOHNSON, P.L. CHURCH, G.M. Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. **Nucleic Acids Res**., 30: 1255–1261. 2002.

CASSELLS, A.C.; CURY, R.F. Oxidative stress and physiological, epigenetic and genetic variability in plant tissue culture: implications for micropropagators and genetic engineers. **Plant Cell, Tissue and Organ Culture**, v. 64, p. 145-157, 2001.

CENCI, A.; GUIGNON, V.; ROUX, N.; ROUARD, M. Genomic analysis of NAC transcription factors in banana (*Musa acuminata*) and definition of NAC orthologous groups for monocots and dicots. **Plant Mol. Biol.**, 85, 63–80. 2014.

CHEN, C.; CHEN, Z. Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced Arabidopsis transcription factor. **Plant Physiol.**, 129. 706-716. 2002.

CHEN, L.; SONG, Y.; LI, S.; ZHANG, L.; ZOU, C.; YU, D. The role of WRKY transcription factors in plant abiotic stresses. **Biochim Biophys Acta**, 1819(2):120-8. 2012.

CHEN, X.; CHEN, Z.; ZHAO, H.; ZHAO, Y.; CHENG, B.; XIANG, Y. Genome-wide analysis of soybean HD-Zip gene family and expression profiling under salinity and drought treatments. **PLoS One**, 9:e87156. 2014.

CHERN, M.S.; EIBEN, H.G.; BUSTOS, M.M. The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. **Plant J**., 10: 135–148. 1996.

CHOW, C.N.; ZHENG, H.Q.; WU, N.Y.; CHIEN, C.H.; HUANG, H.D.; LEE, T.Y.; CHIANG-HSIEH, Y.F.; HOU, P,F.; YANG, T.Y.; CHANG, W.C. PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. **Nucleic Acids Res**, 44: D1154–D1160. 2016.

CIARBELLI, A.R.; CIOLFI, A.; SALVUCCI, S.; RUZZA, V.; POSSENTI, M.; CARABELLI, M.; FRUSCALZO, A.; SESSA, G.; MORELLI, G.; RUBERTI, I. The Arabidopsis Homeodomain-leucine Zipper II gene family: diversity and redundancy. **Plant Mol. Biol**., 68, 465-478. 2008.

CIB – CONSELHO DE INFORMAÇÃO SOBRE BIOTECNOLOGIA. **As idéias e os avanços da biotecnologia**. 2004. Disponível em http://www.cib.org.br. Acesso em: 05 mar. 2017.

CORTESÃO, M. **Culturas tropicais**: plantas oleaginosas. Lisboa: Clássica, 1956. 231p.

COSGROVE, D.J. Loosening of plant cell walls by expansins, **Nature**, 407, 321–326. 2000.

CRAMER, G.R.; BOWMAN, D.C. Kinetics of maize leaf elongation. I. Increased yield threshold limits short-term, steady-state elongation rates after exposure to salinity. **Journal of Experimental Botany,** 42: 1417-1426. 1991.

CRAMER, G.R.; BOWMAN, D.C. Kinetics of maize leaf elongation. I. Increased yield threshold limits short-term, steady-state elongation rates after exposure to salinity, **J. Exp. Bot.** 42, 1417-1426. 1991.

CREISSEN, G.; FIRMIN, J.; FRYER, M.; KULAR, B.; LEYLAND, N.; REYNOLDS, H.; PASTORI, G.; WELLBURN, F.; BAKER. N.; WELLBURN, A. Elevated glutathione biosynthetic capacity in the chloroplasts of transgenic tobacco plants paradoxically causes increased oxidative stress. **Plant Cell.**, 11: 1277–1292. 1999.

CRUZ, C.A.F; PELACANI,C.R.; COELHO, E.F.; CALDAS, R.C; ALMEIDA, A.Q.; QUEIROZ, J.R. Influencia da salinidade sobre o crescimento, absorção e distribuição de sódio, cloro e macronutrientes em plântulas de maracajuzeiro amarelo. **Bragatinga**, Campinas, v.65,n 2,p. 275-284, 2006.

DAKER, A. **Irrigação e Drenagem**: A Água na Agricultura. 7 ed. , Rio de Janeiro: Freitas Bastos. 1988, v. 3, 453p.

DAVEY, M.W.; STALS, E.; PANIS, B.; KEULEMANS, J.; SWENNEN, R.L. High-throughput determination of malondialdehyde in plant tissues. **Ann. Biochem**., 347:201–207. 2005.

DAVIES, B.; SCHWARZ-SOMMER, Z. Control of floral organ identity by homeotic MADS-box transcription factors, in results and problems in cell differentiation (Nover, L. ed.) **Springer-Verlag**. Berlin Heidelberg. vol. 20, pp.235 -258, 1994.

DE CAESTECKER, M.P.; YAHATA, T.; WANG, D.; PARKS, W.T.; HUANG, S.; HILLI, C.S.; SHIODA, T.; ROBERTS, A.B.; LECHLEIDER, R.J. The Smad 4 activation domain (SAD) is a proline-rich, p300-dependent transcriptional activation domain. **J. Biol. Chem.**, 3:2115-22. 2000.

DEHGAN, B.; WEBSTER, G.L. **Morphology and infrageneric relationships of the genus Jatropha (Euphorbiaceae)**. University of California Press, Berkeley, 1979. p. 73.

DELAUNEY, A.; VERMA, D. Proline biosynthesis and osmoregulation in plants. **Plant Journal**, v.4, p.215-223, 1993.

DEMIDCHIK, V. Mechanisms of oxidative stress in plants: from classical chemistry to cell biology. **Environ. Exp. Bot**., 109:212–228. 2015.

DIAS, L.A.S.; LEME, L.P.; LAVIOLA, B.G.; PALLINI FILHO, A.; PEREIRA, O.L.; CARVALHO, M.; MANFIO, C.E.; SANTOS, A.S.; SOUSA, L.C.A.; OLIVEIRA, T.S. e DIAS, D.C.F.S. **Cultivo de pinhão-manso (*Jatropha curcas* L.) para produção de óleo combustível**. Viçosa, MG, 2007. v.1. 40p.

DIAS, L.A.S.; MISSIO, R.F.; DIAS, D.C.F.S. Antiquity, botany, origin and domestication of *Jatropha curcas* (Euphorbiaceae), a plant species with potential for biodiesel production. **Genetics and Molecular Research**, v.11, n.3, p.2719-2728, 2012. http://www.geneticsmr.com/issue/11/13262

DIPIERRO, N.; MONDELLI, D.; PACIOLLA, C.; BRUNETTI, G.; DIPIERRO, S. Changes in the ascorbate system in the response of pumpkin (*Cucurbita pepo* L.) roots to aluminium stress. **J. Plant Physiol**. 162:529-536. 2005.

DONG, J.; CHEN, C.; CHEN, Z. Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. **Plant Mol. Biol**., 51, 21–37. 2003.

DRUMOND, M. A.; PIRES, I. E.; OLIVEIRA, V. R.; OLIVEIRA, A. R.; ALVAREZ, I. A. Produção e distribuição de biomassa de espécies arbóreas no Semi-Árido brasileiro. **Revista Árvore**, v. 32, n. 4, p. 665-669, 2008.

DU, H.; FENG, B.-R.; YANG, S.-S.; HUANG, Y.-B.; TANG, Y.-X. . The R2R3-MYB transcription factor gene family in maize. **PLoS One**. 7:e3746. 2012.

DUBOUZET, J.G.; SAKUMA, Y.; ITO, Y.; KASUGA, M.; DUBOUZET, E.G.; MIURA, S.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. **Plant J.**, 33: 751–763. 2003.

DURBIN, R.; EDDY, S.R.; KROGH, A.; MITCHISON, G.J. **Biological Sequence Analysis**: Probabilistic Models of Proteins and Nucleic Acids. Cambridge UK: Cambridge University Press. 1998.

DUVAL, M.; HSIEH, T.F.; KIM, S.Y.; THOMAS, T.L. Molecular characterization of AtNAM: A member of the Arabidopsis NAC domain superfamily. **Plant Mol. Biol**., 50, 237–248. 2002.

EGEA-CORTINES, M.; SAEDLER, H.; SOMMER, H. Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. **EMBO J.** 18: 5370–5379. 1999.

EISEN, M.B.; SPELLMAN, P.T.; BROWN, P.O.; BOTSTEIN, D. Cluster analysis and display of genome-wide expression patterns. **Proc. Natl. Acad. Sci**., USA 95:14863-14868. 1998.

ELHITI, M.; STASOLLA, C. Structure and function of homodomain-leucine zipper (HD-Zip) proteins. **Plant Signaling and Behavior**, v.4, p.86-88, 2009.

ELLENBERGER, T.E.; BRANDL, C.J.; STRUHL, K.; HARRISON, S.C. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. **Cell**, 71, 1223–1237. 1992.

ERNST, H.A..; OLSEN, A.N.; LARSEN, S.; LO LEGGIO, L. Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. **EMBO Rep**., 5, 297–303. 2004.

EULGEM, T.; RUSHTON, P.J.; ROBATZEK, S.; SOMSSICH, I.E. The WRKY superfamily of plant transcription factors. **Trends Plant Sci**., 5, 199–206. 2000.

EULGEM, T.; RUSHTON, P.J.; SCHMELZER, E.; HAHLBROCK, K.; SOMSSICH, I.E. Early nuclear events in plant defense: rapid gene activation by WRKY transcription factors. **EMBO J**., 18 4689–4699. 1999.

FAIRMAN, R.; BERAN-STEED, R.K.; ANTHONY-CAHILL, S.J.; LEAR, J.D.; STRAFFORD, W.F.; DEGRADO, W.F.; BENFIELD, P.A.; BRENNER, S.L. Multiple oligomeric states regulate the DNA binding of helix-loop-helix peptides. **Proc. Natl. Acad. Sci**. USA 90, 10429–10433. 1993.

FANG, Y.; YOU, J.; XIE, K.; XIE, W.; XIONG, L. Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. **Mol. Genet. Genomics**, 280, 547–563. 2008.

FAO - **Food and Agriculture Organization**. Production: Values 1960-2008. In: The State of World Fisheries and Aquaculture. Aquaculture in China and Asia. Rome: FAO. 2008.

FENG, C.; ANDRESSSON, E.; MASLAK, A.; MOCK, H.P.; MATTSSON, O.; MUNDY, L. Arabidopsis MYB68 in development and responses to environmental cues. **Plant Sci**., 167:1099–1107. 2004.

FERRÃO, J.E.M.; FERRÃO, A.M.B.C.; PATRÍCIO, M.T.S. Purgueira da Ilha do Fogo. Composição da semente, algumas características da gordura. **Garcia de Orta, Sér. Est. Agron**. 10: 175-178. 1983.

FERRE-D'AMARE, A.R.; POGNONEC, P.; ROEDER, R.G.; BURLEY, S.K. Structure and function of the b/HLH/Z domain of USF. **EMBO J**., 13, 180–189. 1994.

FITCH, W.M. Toward defining the course of evolution: Minimum change for a specific tree topology. **Syst. Zool.,** 20: 406-416, 1971.

FLANAGAN, C. A.; MA, H. Spatially and temporally regulated expression of the MADS-box gene AGL2 in wild-type and mutant Arabidopsis flowers. **Plant Mol. Biol**. 26, 581-595. 1994.

FOYER, C.H.; NOCTOR, G. Oxygen processing in photosynthesis: regulation and signalling. **New Phytol**., 146: 359–388. 2000.

FU, R.; LIU, W.; LI, Q.; LI, J.; WANG, L.; REN, Z. Comprehensive analysis of the homeodomain-leucine zipper IV transcription factor family in *Cucumis sativus*. **Genome**, 56, 395–405. 2013.

FUKUDA, A.; NAKAMURA, A.; TANAKA, Y. Molecular cloning and expression of the Na/H exchanger gene in *Oryza sativa*. **Biochim. Biophys. Acta.** 1446, 149–155. 1999.

GAXIOLA, R.A.; RAO, R.; SHERMAN, A.; GRISAFI, P.; ALPER, S.L.; FINK, G.R. The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. **Proc. Natl. Acad. Sci**. USA 96, 1480–1485. 1999.

GILL, S.S.; TUTEJA, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant Physiol. Biochem**., 48: 909-930. 2010.

GIORGETTI, L.; SIGGERS, T.; TIANA, G.; CAPRARA, G.; NOTARBARTOLO, S.; CORONA, T.; PASPARAKIS, M.; MILANI, P.; BULYK, M.L.; NATOLI, G. Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. **Mol. Cell**, 37, pp. 418-428. 2011.

GRAMZOW, L.; THEISSEN, G. A hitchhiker's guide to the MADS world of plants. **Genome Biol**., 11(6):214. 2010.

GRIMPLET, J.; MARTÍNEZ-ZAPATER. J.M.; CARMONA, M.J. Structural and functional annotation of the MADS-box transcription factor family in grapevine. **BMC Genomics**. 2016;17:80. doi: 10.1186/s12864-016-2398-7.

GU, Y-Q.; YANG, C.; THARA, V.K.; ZHOU, J.; MARTIN, G.B. Pti4 is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. **Plant Cell**. 12: 771–785. 2000.

GUPTA, B.; HUANG, B. Mechanism of salinity tolerance in plants: physiological, biochemical, and molecular characterization. **Int. J Genomics.** 2014:18. 2014. doi: 10.1155/2014/701596

HAGEN, J.B. The origins of bioinformatics. **Nat. Rev. Genet**., 231-236. 2000.

HAMADA, A.; SHONO, M.; XIA, T.; OHTA, M.; HAYASHI, Y.; TANAKA, A.; HAYAKAWA, T. Isolation and characterization of a Na^+ /H^+ antiporter gene from the halophyte Atriplex gmelini. **Plant Mol. Biol**. 46 35–42. 2001.

HARRIS, J.C.; HRMOVA, M.; LOPATO, S.; LANGRIDGE, P. Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. **New Phytologist**, 190, 823–837. 2011.

HARTER, L.S. H.; HARTER, F.S.; DEUNER, C. Effect of salinity on physiological performance of mogango seeds and seedlings. **Horticultura Brasileira**, v.32, n.1, p.80-85, 2014.

HASANUZZAMAN, M.; ALAM, M.M.; RAHMAN, A.; HASANUZZAMAN, M.; NAHAR, K.; FUJITA, M. Exogenous proline and glycine betaine mediated upregulation of antioxidant defense and glyoxalase systems provides better protection against salt-induced oxidative stress in two rice (*Oryza sativa* L.) varieties. **Biomed. Res. Int**. 757219. 2014. doi: 10.1155/2014/757219

HASEGAWA, P.M.; BRESSAN, R.A.; ZHU, J.K.; BOHNERT, H.J. Plant cellular and molecular responses to high salinity. **Annual Review of Plant Physiology and Plant Molecular Biology.** 51, 463–499. 2000.

HE, X.J.; MU, R.L.; CAO, W.H.; ZHANG, Z.G.; ZHANG, J.S.; CHEN, S.Y. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. **Plant J**., 44, 903–916. 2005.

HE, X.J.; MU, R.L.; CAO, W.H.; ZHANG, Z.G.; ZHANG, J.S.; CHEN, S.Y. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. **Plant J**., 44:903–916. 2005.

HEIFFIG, L.S.; CÂMARA, G.M.S. Potencial da cultura do pinhão-manso como fonte de matériaprima para o Programa Nacional de Produção e Uso do Biodiesel. In: CÂMARA, G.M.S.; HEIFFIG, L.S. (Coord.) **Agronegócio de Plantas Oleaginosas: matérias-primas para biodiesel**. Piracicaba: ESALQ/USP/LPV, 2006. p. 105-121.

HEIM, M.A.; JAKOBY, M.; WERBER, M.; MARTIN, C.; WEISSHAAR, B.; BAILEY, P.C. The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. **Mol. Biol. Evol**., 20: 735–747. 2003.

HELLER, J. **Physic nut (*Jatrhopha curcas* L**.). Promiting the conservation and use of underutilized and neglected crops 1. IBPGR 161. Roma, IBPGR, 1996. 66p.

HENRIKSSON, E.; OLSSON, A.S.B.; JOHANNESSON, H.; HANSON, J.; ENGSTROM, P.; SODERMAN, E. Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. **Plant Physiology**, v.139, p.509-518, 2005.

HENRIKSSON, E.; OLSSON, A.S.B.; JOHANNESSON, H.; JOHANSSON, H.; HANSON, J.; ENGSTROM, P.; SODERMAN, E. Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. **Plant Physiol**., 139(1), 509–518. 2005.

HENSON, J.; TISCHLER, G.; NING, Z. Next-generation sequencing and large genome assemblies. **Pharmacogenomics**., 13: 901-915. 2012.

HERNANDEZ-GARCIA, C.M.; FINER, J.J. Identification and validation of promoters and cis-acting regulatory elements. **Plant Science**., 217: 109-119. 10.1016/j.plantsci.2013.12.007. 2014.

HICHRI, I.; BARRIEU, F.; BOGS, J.; KAPPEL, C.; DELROT, S.; LAUVERGEAT, V. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. **J. Exp. Bot.**, 62, 2465–2483. 2011.

HIGO, K.; UGAWA, Y.; IWAMOTO, M.; KORENAGA, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. **Nucleic Acids Res**, 27: 297–300. 1999.

HIMMELBACH, A.; HOFFMANN, T.; LEUBE, M.; HÖHENER, B.; GRILL, E. Homeodomain protein ATHB6 is a target of the protein phosphatase ABI1 and regulates hormone responses in Arabidopsis. **EMBO Journal**, 21, 3029–3038. 2002.

HOEKSTRA, P.A.; GOLOVINA, E.A.; BUITINK, J. Mechanisms of plant dessication tolerance. **Trends Plant Sci.** 6:431-438. 2001.

HOGEWEG, P. The Roots of Bioinformatics in Theoretical Biology. **PLoS Comput Biol**., 7(3): e1002021. 2011.

HRDLICKOVA, R.; TOLOUE, M.; TIAN, B. RNA-Seq methods for transcriptome analysis. Wiley Interdiscip. **Rev. RNA.**, 8: 2016.

HU, J.; REDDY, V.S.; WESSLER, S.R. The rice R gene family: two distinct subfamilies containing several miniature inverted-repeat transposable elements. **Plant Mol Biol**, 42, 667-678. 2000.

HU, R.; CHI, X.; CHAI, G.; KONG, Y.; HE, G.; WANG, X.; SHI, D.; ZHANG, D.; ZHOU, G. Genome-wide identification, evolutionary expansion, and expression profile of homeodomain-leucine zipper gene family in poplar (*Populus trichocarpa*). **PLoS One**, 7: e31149. 2012.

HU, R.; QI, G.; KONG, Y.; KONG, D.; GAO, Q.; ZHOU, G. Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. **BMC Plant Biol.**, 10:145. 2010.

HU, Y.X.; WANG, Y.X.; LIU, X.F.; LI, J.Y. Arabidopsis RAVI is down-regulated by brassinosteroid and may act as a negative regulator during plant development. **Cell Research**, 14. 8-15. 2004.

HUANG, H.; TUDOR, M.; SU, T.; ZHANG, Y.; HU, Y.; MA, H. DNA binding properties of two Arabidopsis MADS domain proteins: bindig consensus and dimmer formation. **Plant Cell**., 8:81-94. 1996.

HUNSCHE, M.; BÜRLING, K.; SAIED, A.S.; SCHMITZ-EIBERGER, M.; SOHAIL, M.; GEBAUER, J.; NOGA, G.; BUERKERT, A. Effects of NaCl on surface properties, chlorophyll fluorescence and light remission, and cellular compounds of *Grewia tenax* (Forssk.) Fiori and *Tamarindus indica* L. leaves. **Plant Growth Regulation**, v. 61, n. 03, p. 253-263, 2010.

INZÉ. D.; VAN MONTAGU, M. Oxidative stress in plants. **Current Opninion in Biotechnology**, v.6, n. 02, p 153-158, 1995.

ISHIDA, T.; HATTORI S.; SANO, R. Arabidopsis TRANSPARENT TESTA GLABRA2 is directly regulated by R2R3-MYB transcription factors and is involved in regulation of GLABRA2 transcription in epidermal differentiation. The Plant Cell 19, 2531–2543. 2007.

ISHITANI, M.; LIU, J.; HALFTER, U.; KIM, C.S.; M.W.; ZHU, J.K. SOS3 function in plant salt tolerance requires myristoylation and calcium-binding. **Plant Cell.** 12, 1667–1677. 2000.

IZAWA, T.; FOSTER, R.; NAKAJIMA, M.; SHIMAMOTO, K.; CHUA, N.H. The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development. **Plant Cell**, 6: 1277-1287. 1994.

JAILLON O.; AURY JM, NOEL B, POLICRITI A, CLEPET C, CASAGRANDE A, CHOISNE N, AUBOURG S, VITULO N, JUBIN C, VEZZI A, LEGEAI F, HUGUENEY P, DASILVA C, HORNER D, MICA E, JUBLOT D, POULAIN J, BRUYÈRE C, BILLAULT A, SEGURENS B, GOUYVENOUX M, UGARTE E, CATTONARO F, ANTHOUARD V, VICO V, DEL FABBRO C, ALAUX M, DI GASPERO G, DUMAS V, FELICE N, PAILLARD S, JUMAN I, MOROLDO M, SCALABRIN S, CANAGUIER A, LE CLAINCHE I, MALACRIDA G, DURAND E, PESOLE G, LAUCOU V, CHATELET P, MERDINOGLU D, DELLEDONNE M, PEZZOTTI M, LECHARNY A, SCARPELLI C, ARTIGUENAVE F, PÈ ME, VALLE G, MORGANTE M, CABOCHE M, ADAM-BLONDON AF, WEISSENBACH J, QUÉTIER F, WINCKER P. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. **Nature**, 449: 463–467. 2007.

JAKOBY, M.; WEISSHAAR, B.; DROGE-LASER, W.; VICENTE-CARBAJOSA, J.; TIEDEMANN, J.; KROJ, T.; PARCY, F. bZIP transcription factors in Arabidopsis. **Trends Plant Sci**., 7,106 -111. 2002.

JAVELLE, M.; VERNOUD, V.; ROGOWSKY, P.M.; INGRAM, G.C. Epidermis: the formation and functions of a fundamental plant tissue. **New Phytol**, 189: 17–39. 2011.

JIANG, H.; BETANCOURT, L.; SMITH, R.G. Ghrelin amplifies dopamine signaling by cross talk involving formation of growth hormone secretagogue receptor/dopamine receptor subtype 1 heterodimers. **Mol Endocrinol**, 20:1772–85. 2006.

JIN, J.; ZHANG, H.; KONG, L.; GAO, G.; LUO, J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. **Nucleic Acids Res**. 42:D1182–D1187. 2013.

JOFUKU, K.D.; OMIDYAR, P.K.; GEE, Z.; OKAMURO, J.K. Control of seed mass and seed yield by the floral homeotic gene APETALA2. **Proc. Natl. Acad. Sci.** U.S.A. 102, 3117–3122. 2005.

JOHANNESSON, H..; WANG, Y.; HANSON, J.; ENGSTRÖM, P. The Arabidopsis thaliana homeobox gene ATHB5 is a potential regulator of abscisic acid responsiveness in developing seedlings**. Plant Molecular Biology**, 51, 719–729. 2003.

JOHNSON, P.R.; SWANSON, R.; RAKHILINA, L.; HOCHSTRASSER, M. Degradation signal masking by heterodimerization of MATalpha2 and MATa1 blocks their mutual destruction by the ubiquitin-proteasome pathway. **Cell**, 94, 217–227. 1998.

JONES, N.C.; PEVZNER, P.A. **An introduction to bioinformatics algorithms**. Massachusetts Institute of Technology. MIT Press books, 2004.

JONGSCHAAP, R.E.E.; W.J. CORRÉ; BINDRABAN, P.S.; BRANDENBURG, W.A.; 2007. **Claims and Facts on *Jatropha curcas* L**. Global Jatropha curcas evaluation, breeding and propagation programme. Plant Research International B.V., Wageningen, the Netherlands, Report 158, 42 pp.

KARAM, B.S.; RHONDA, C.F.; LUIS, O.S. Transcription factors in plant defense and stress response. **Current Opinion in Plant Biology**, 5. 430–436. 2002.

KASUGA, T.; TOWNSEND, J.P.; TIAN, C.G.; GILBERT, L.B.; MANNHAUPT, G.; TAYLOR, J.W.; GLASS, N.L. Long-oligomer microarray profiling in Neurospora crassa reveals the transcriptional program underlying biochemical and physiological events of conidial germination. **Nucleic Acids Res.**,33, 6469–6485. 2005.

KAUFMANN, K.; MELZER, R.; THEISSEN, G. MIKC-type MADS-domain proteins: Structural modularity, protein interactions and network evolution in land plants. **Gene**, 347: 183–198. 2005.

KAWAKATSU, T.; TAKAIWA, F. Cereal seed storage protein synthesis: fundamentalprocesses for recombinant protein production in cereal grains. **Plant Biotech-nology Journal**., 8, 939–953. 2010.

KIM, H.J.; RYU, H.; HONG, S.H.; WOO, H.R.; LIM, P.O.; LEE, I.C.; SHEEN, J.; NAM, H.G.; HWANG, I. Cytokinin-mediated control of leaf longevity by AHK3 through phosphorylation of ARR2 in Arabidopsis. **Proceedings of the National Academy of Sciences**, USA 103, 814–819. 2006.

KIM, S.G.; KIM, S.Y.; PARK, C.M. A membrane-associated NAC transcription factor regulates salt-responsive flowering via FLOWERING LOCUS T in Arabidopsis. **Planta**, 226, 647–654. 2007.

KIM, S.Y.; KIM, S.G.; KIM, Y.S.; SEO, P.J.; BAE, M.; YOON, H.K.; PARK, C.M. Exploring membrane-associated NAC transcription factors in Arabidopsis: implications for membrane biology in genome regulation. **Nucleic Acids Res**., 35, 203–213. 2007.

KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **J. Mol. Evol.**, 16, 111–120. 1980.

KING, A.J.; LI, Y.; GRAHAM, I.A. Profiling the developing *Jatropha curcas* L. seed transcriptome by pyrosequencing. **BioEnergy Res**.,4:211-221. 2011.

KRZYZANOSWSKI, F.C.; FRANÇA-NETO, J.B. Vigor de sementes. **Informativo ABRATES**, Londrina, v.11, n.3, p.81-84, 2001.

LATA, C.; PRASAD, M. Role of DREBs in regulation of abiotic stress responses in plants. **J. Exp. Bot**. 62, 4731–4748. 2011.

LE, D.T.; NISHIYAMA, R.; WATANABE, Y.; MOCHIDA, K.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. **DNA Res.**, 18, 263–276. 2011.

LESCOT, M.; DÉHAIS, P.; THIJS, G.; MARCHAL, K.; MOREAU, Y.; VAN DE PEER, Y.; ROUZÉ, P.; ROMBAUTS, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic Acids Res**, 30: 325–327. 2002.

LESK, A.M. **Introdução à Bioinformática**. Artmed, 2ª edição. Editora Artmed, Porto Alegre, Brasil. 2008.

LI, G.; MARGUERON, R.; KU, M.; CHAMBON, P.; BERNSTEIN, B.E.; REINBERG, D. Jarid2 and PRC2, partners in regulating gene expression. **Genes Dev**. 24(4):368–380. 2010.

LICHTENBERG, J.; YILMAZ, A.; WELCH, J.; KURZ, K.; LIANG, X.; DREWS, F.; ECKER, K.; LEE, S.; GEISLER, M.; GROTEWOLD, E.; WELCH, L. The word landscape of the non-coding segments of the Arabidopsis thaliana genome. **BMC Genomics**, 10:463. 2009.

LIMA JUNIOR, J.A.; PEREIRA, G. M.; GEISENHOFF, L. O.; SILVA, W. G.; SOUZA, R. O. R. M.; VILAS BOAS, R. C. Economic viability of a drip irrigation system on carrot crop. **Revista de Ciências Agrarias**, Belém, v. 57, n. 1, p. 15-21. 2014.

LIMA, L.A. Efeitos de sais no solo e na planta. **Anais do XXVI Congresso Brasileiro de Engenharia Agrícola** – Manejo e Controle da Salinidade na Agricultura Irrigada, Campina Grande: UFPB, 1997. pg. 113–133.

LIN, C.C.; KAO, C.H. Regulation of ammonium-induced proline accumulation in detached rice leaves. **Plant Growth Regulation**, Dordrecht, v.35, n.3, p. 69-74, 2001.

LIPPOLD, F.; SANCHEZ, D.H.; MUSIALAK, M.; SCHLERETH, A.; SCHEIBLE, W.R.; HINCHA, D.K.; UDVARDI, M.K. AtMyb41 regulates transcriptional and metabolic responses to osmotic stress in Arabidopsis. **Plant Physiol**,. 149:1761–1772. 2009.

LIU, J.; ZHU, J.K. A calcium sensor homolog required for plant salt tolerance. **Science**. 280, 1943–1945. 1998.

LIU, J.J.; EKRAMODDOULLAH, A.K.M. Identification and characterization of the WRKY transcription factor family in *Pinus monticola*. **Genome**, 52. 77-88. 2009.

LIU, L. LI, Y.; LI, S.; HU, N.; HE, Y.; PONG, R.; LIN, D.; LU, L.; LAW, M. Comparison of next-generation sequencing systems. **J Biomed Biotechnol**, v. 2012, p. 251364, 2012.

LIU, Q.; KASUGA, M.; SAKUMA, Y.; ABE, H.; MIURA, S.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. **Plant Cell.** 10. 1391–1406. 1998.

LIU. L.; WHITE, M.J.; MACRAE, T.H. Transcription factors and their genes in higher plants functional domains, evolution and regulation. **Eur. J. Biochem**., 262: 247–257. 1999.

LU, H.; LIU, Y.; ZHOU, H.; YANG, Y.; CHEN, M.; LIANG, B. Production of biodiesel from *Jatropha curcas* L. oil. **Computers and Chemical Engineering**, v. 33, p. 1091-1096, 2009.

MA, P.C.M.; ROULD, M.A.; WEINTRAUB, H.; PABO, C.O. Crystal structure of MyoD bHLH domain-DNA recognition and implications for transcriptional activation. **Cell**, 77, 451–459. 1994.

MAATHUIS, F.J.M.; AMTMANN, A. K^+ Nutrition and Na^+ toxicity: The basic of cellular K^+ /Na^+ ratios. **Annals Botany**. 84: 123-133. 2004.

MACHADO NETO, N.B.; CUSTÓDIO, C.C.; COSTA, P.R.; DONÁ, F.L. Deficiência hídrica induzida por diferentes agentes osmóticos na germinação e vigor de sementes de feijão. **Revista Brasileira de Sementes**, v. 28, n. 1, p. 142-148, 2006.

MAHAJAN, S.; TUTEJA, N. Cold, salinity and drought stresses: An overview. **Arch. Biochem. Biophys.** 444, 139–158. 2005.

MANGELSEN, E.; KILIAN, J.; BERENDZEN, K.W.; KOLUKISAOGLU, U.H.; HARTER, K.; JANSSON C.; WANKE, D. Phylogenetic and comparative gene expression analysis of barley (*Hordeum vulgare*) WRKY transcription factor family reveals putatively retained functions between monocots and dicots. **BMC Genomics**, 9 2008.

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

MARE, C.; MAZZUCOTELLI, E.; CROSATTI, C.; FRANCIA, E.; STANCA, A.M. CATTIVELLI, L. Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. **Plant Mol. Biol**., 55: 399–416. 2004.

MARSCHNER, H. **Mineral nutrition of higher plants**. San Diego: Academic Press, 1995. 889 p.

MARTINO, F.; KUENG, S.; ROBINSON, P.; TSAI-PFLUGFELDER, M.; VAN LEEUWEN, F.; ZIEGLER, M.; CUBIZOLLES, F.; COCKELL, M.M.; RHODES, D.; GASSER, S.M. Reconstitution of yeast silent chromatin: multiple contact sites and O-AADPR binding load SIR complexes onto nucleosomes *in vitro*. **Mol Cell,** 33: 323–334. 2009.

MARTINOIA, E.; SCHRAMM, M. J.; KAISER, G.; KAISER, W. M.; HEBER, U. Transport of anions in isolated barley vacuoles: I. Permeability to anions and evidence for a Cl^- uptake system. **Plant Physiology**, v.80, p.895-901, 1986.

MARTINS, W.S.; LUCAS, D.C.; NEVES, K.F.; BERTIOLI, D.J. WebSat - a web software for microsatellite marker development. **Bioinformation**, 3: 282-283. 2009.

MASIERO, S.; IMBRIANO, C.; RAVASIO, F.; FAVARO, R.; PELUCCHI, N.; GORLA, M.S.; MANTOVANI, R.; COLOMBO, L.; KATER, M.M. Ternary complex formation between MADS-box transcription factors and the histone fold protein NF-YB. **J. Biol. Chem**., 277, 26429–26435. 2002.

MATIOLI, S.R.; FERNANDES, F.M.C. **Biologia molecular e evolução**. 2. ed. Ribeirão Preto: Holos Editora, 2012. 256p.

MCGRATH, K.C.; DOMBRECHT, B.; MANNERS, J.M.; SCHENK, P.M.; EDGAR, C.I.; MACLEAN, D.J.; SCHEIBLE, W.R.; UDVARDI, M.K.; KAZAN, K. Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. **Plant Physiol** 139:949–959. 2005.

MEDEIROS, J.F.; NASCIMENTO, I.B.; GHERY, H.R. Manejo do solo-água-planta em área afetadas por sais. In: GHEYI, H. R.; DIAS, N. S.; LACERDA, C. F. Manejo da salinidade na agricultura: Estudos básicos e aplicados. Fortaleza: **Instituto Nacional de Ciência e Tecnologia em Salinidade**, p. 280-302, 2010.

MENEZES, S.M.; TILLMANN, M.A.A.; DODE, L.B.; VILLELA, F.A. Detecção de soja geneticamente modificada tolerante ao glifosato por métodos baseados na atividade de enzimas. **Revista Brasileira de Sementes**, v. 26, n. 2, p. 150-155, 2004.

MEYER, E.; LOGAN, T.L.; JUENGER, T.E. Transcriptome analysis and gene expression atlas for *Panicum hallii* var. filipes, a diploid model for biofuel research. **Plant J.** 70: 879–890. 2012.

MIR, L.; MOREIRA FILHO, C.A.; MENCK, C.F.M. **Genômica**. Editora Atheneu. 2006, 1114p.

MISRA, N.; DWIVEDI, U.N. Genotypic difference in salinity tolerance of green gram cultivars. **Plant Sci**., 166, 1135–1142. 2004.

MITHRA, S.V.A.; KULKARNI, K.; SRINIVASAN, R. Plant Promoters: Characterization and Applications in Transgenic Technology. Springer, Singapore. **Plant Biotechnology**: Principles and Applications. p.117-172. 2017. DOI https://doi.org/10.1007/978-981-10-2961-5_5

MITTLER, R. Oxidative stress, antioxidants and stress tolerance. **Trends Plant Sci.**, 7:405-410. 2002.

MOENS, C.B.; SELLERI, L. Hox cofactors in vertebrate development. **Dev. Biol.**, 291,193 -206. 2006.

MOL J., GROTEWOLD E., KOES R. How genes paint flowers and seeds. **Trends Plant Sci**., 3, 212-217. 1998.

MOOSE, S.P.; SISCO, P.H. Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. **Genes Dev.** 10, 3018–3027. 1996.

MORELLI, G.; RUBERTI, I. Light and shade in the photocontrol of Arabidopsis growth. **Trends Plant Sci**., 7: 399–404. 2002.

MOROZOVA, O.; HIRST, M.; MARRA, M.A. Applications of new sequencing technologies for transcriptome analysis. **Annu Rev Genomics Hum Genet**., 10, 135–151. 2009.

MU, R.L.; CAO, Y.R.; LIU, Y.F.; LEI, G.; ZOU, H.F.; LIAO, Y.; WANG, H.W.; ZHANG, W.K.; MA, B.; DU, J.Z.; YUAN, M.; ZHANG, J.S.; CHEN, S.Y. An R2R3-type transcription factor gene AtMYB59 regulates root growth and cell cycle progression in Arabidopsis. **Cell Res**., 19:1291–1304. 2009.

MÜHLING, K.H.; LAUCHLI, A. Effect of salt stress on growth and cation compartmentation in leaves of two plants species differing in salt tolerance. **Journal of Plant Physiology**, Stuttgart, v.159, p.137-146, 2002.

MUKHERJEE, K.; BURGLIN, T.R. MEKHLA, a novel domain with similarity to PAS domains, is fused to plant homeodomain lecine zipper III proteins. **Plant Physiol**., 140,1142 -1150. 2006.

MULLER, C.W. Transcription factors: global and detailed views. **Curr. Opin. Struct. Biol.**, 11: 26–32. 2001.

MUNNS, R. Genes and salt tolerance: bringing them together. **New Phytologist**, v.167, n.3, p.645-663. 2005.

MUNNS, R. Physiological processes limiting plant growth in saline soil: some dogmas and hypotheses. **Plant, Cell e Environment**., 16: 15–24. 1993.

MUNNS, R.; HUSAIN, S.; RIVELLI, A.R.; JAMES, R.A.; CONDON, A.G.; LINDSAY, M.P.; LAGUDAH, E.S.; SCHACHTMAN, D.P.; HARE, R.A. Avenues for increasing salt tolerance of crops, and the role of physiologically-based selection traits. **Plant and Soil**, 247,93–105. 2002.

MUNNS, R.; JAMES, R.A.; LAUCHLI, A. Approaches to increasing the salt tolerance of wheat and other cereals. **Journal of Experimental Botany**, Oxford, v.57, n.5, p.1025-1043, 2006.

NABIL, M.; COUDRET, A. Effects of sodium chloride on growth, tissue elasticity and solute adjustment in two *Acacia nilotica* subspecies. **Physiologia Plantarum**, 93:217–224. 1995.

NADAKUDUTI, S.S.; POLLARD, M.; KOSMA, D.K.; ALLEN, C.; OHLROGGE, J.B.; BARRY, C.S. Pleiotropic phenotypes of the sticky peel mutant provide new insight into the role of CUTIN DEFICIENT2 in epidermal cell function in tomato. **Plant Physiol.** 159:945–960. 2012.

NAGALAKSHMI, U.; WANG, Z.; WAERN, K.; SHOU, C.; RAHA, D.; GERSTEIN, M.; SNYDER, M. The transcriptional landscape of the yeast whole genome defined by RNA sequencing. **Science**., 320:1344-1349. 2008.

NAKAMURA, M.; KATSUMATA, H.; ABE, M.; YABE, N.; KOMEDA, Y.; YAMAMOTO, K.T.; TAKAHASHI, T. Characterization of the class IV homeodomain-Leucine Zipper gene family in Arabidopsis. **Plant Physiol**., 141, 1363–1375. 2006.

NAKANO, T.,; SUZUKI, K.; FUJIMURA, T.; SHINSHI, H. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. **Plant Physiol.** 140, 411–432. doi: 10.1104/pp.105.073783. 2006.

NASS, R.; CUNNINGHAM, K.W.; RAO, R. Intracellular sequestration of sodium by a novel Na^+/H^+ exchanger in yeast is enhanced by mutations in the plasma membrane H^+-ATPase. Insights into mechanisms of sodium tolerance. **J. Biol Chem.** 272:26145-26152. 1997.

NESI, N.; JOND, C.; DEBEAUJON, I.; CABOCHE, M. LEPINIEC, L.The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. **Plant Cell**, 13, 2099–2114. 2001.

NI, M.; DEHESH, K.; TEPPERMAN, J.M.; QUAIL, P.H. GT-2: In vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity. **Plant Cell**., 8: 1041–1059. 1996.

NIJHAWAN, A.; JAIN, M.; TYAGI, A.K.; KHURANA, J.P. Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. **Plant Physiol**., 146: 333-350. 2008.

NIKOVICS, K.; BLEIN, T.; PEAUCELLE, A.; ISHIDA, T.; MORIN, H.; AIDA, M.; LAUFS, P. The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. **The Plant Cell**, 18, 2929–2945. 2006.

NURUZZAMAN, M.; MANIMEKALAI, R.; SHARONI, A.M.; SATOH, K.; KONDOH, H.; OOKA, H. Genome-wide analysis of NAC transcription factor family in rice. **Gene**, 465, 30–44. 2010.

OAKLEY, F.; MANN, J.; RUDDELL, R.G.; PICKFORD, J.; WEINMASTER, G.; MANN, D.A. Basal expression of IkappaBalpha is controlled by the mammalian transcriptional repressor RBP-J (CBF1) and its activator Notch1. **J. Biol. Chem**., 278, 24359-24370. 2003.

OGATA, K.; MORIKAWA, S.; NAKAMURA, H.; SEKIKAWA, A.; INOUE, T.; KANAI, H.; SARAI, A.; ISHII, S.; NISHIMURA, Y. Solution structure of a specific DNA complex of the Myb DNAbinding domain with cooperative recognition helices. **Cell**, 79, 639–648. 1994.

OHASHI, Y.; OKA, A.; RODRIGUES-POUSADA, R.; POSSENTI, M.; RUBERTI, I.; MORELLI, G.; AOYAMA, T. Modulation of phospholipid signaling by GLABRA2 in root-hair pattern formation. **Science**, 300, 1427–1430. 2003.

OHME-TAKAGI, M.; SHINSHI, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. **Plant Cell**. 7, 173–182. 1995.

OLIVEIRA, A.B.; GOMES-FILHO,E.; ENÉAS-FILHO,J. O problema da salinidade na agricultura e as adaptações das plantas ao estresse salino. **Enciclopédia Biosfera**, Centro Científico Conhecer - Goiânia, v.6, n.11, p.1-16. 2010.

OLIVEIRA, F.A.; OLIVEIRA, F.R.A.; CAMPOS, M.S.; OLIVEIRA, M.K.T.; MEDEIROS, J.F.; SILVA, O.M.P. Interação entre salinidade e fontes de nitrogênio no desenvolvimento inicial da cultura do girassol. **Revista Brasileira de Ciências Agrárias**, v.5, p.479-484, 2010.

OLSEN, A.N.; ERNST, H.A.; LO LEGGIO, L.; SKRIVER, K. NAC transcription factors: structurally distinct, functionally diverse. **Trends Plant Sci**., 10, 1360–1385. 2005.

OOKA, H.; SATOH, K.; DOI, K.; NAGATA, T.; OTOMO, Y. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. **DNA Res.**, 10: 239–247. 2003.

PAGE, R.D. TreeView: An application to display phylogenetic trees on personal computers. **Comput. Appl. Biosci.**, 12: 357–358. 1996.

PANDEY, S.P.; SOMSSICH, I.E. The role of WRKY transcription factors in plant immunity. **Plant Physiol**., 150, 1648-1655. 2009.

PARIDA A.K., DAS A.B., MOHANTY P. Defense potentials to NaCl in a mangrove, *Bruguiera parviflora*: differential changes of isoforms of some antioxidative enzymes. **J. Plant Physiol**., 161: 531–542. 2004.

PARIDA, A.K.; DAS, A.B. Salt tolerance and salinity effect on plants: a review. **Ecotoxicol. Environ. Saf.**, 60: 324–349. 2005.

PARK,C.Y.; LEE, J.H.; YOO, J.H.; MOON, B.C.; CHOI, M.S.; KANG, Y.H., LEE, S.M., KIM, H.S.; KANG, K.Y. CHUNG, W.S. WRKY Group IId transcription factors interact with calmodulin. **FEBS Lett.**, 579, 1545-1550. 2005.

PAZ-ARES, J.; GHOSAL, D.; WIENAND, U.; PETERSON, P.; SAEDLER, H. The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb protooncogene products and with structural similarities to transcriptional activators. **EMBO J.**, 6:3553.1987.

PAZETO, M.S.R. **Estudo da diversidade genética entre acessos de Jatropha spp. por meio de caracteres morfológicos e marcadores moleculares ISSR**. Jaboticabal, 2013. 73p. Tese de Doutorado, Universidade Estadual Paulista. 2013.

PECINA-QUINTERO, V.; ANAYA-LÓPEZ, J. L.; ZAMARRIPA-COLMENERO, A.; NÚÑEZ-COLÍN, C. A.; MONTES-GARCÍA, N.; SOLÍS-BONILLA, J. L.; JIMÉNEZ-BECERRIL, M. F. Genetic structure of Jatropha curcas L. in Mexico and probable centre of origin. **Biomass and Bioenergy** , v. 60, p. 147-155, 2014.

PEIXOTO, A.R. **Plantas oleaginosas arbóreas**. São Paulo: Nobel, 1973. 284p.

PEREMARTI, A.; TWYMAN, R.M.; GOMEZ-GALERA, S.; NAQVI, S.; FARRE, G.; SABALZA, M.; MIRALPEIX, B.; DASHEVSKAYA, S.; YUAN D.W.; RAMESSAR, K.; CHRISTOU, P.; ZHU, C.F.; BASSIE, L. CAPELL, T. Promoter diversity in multigene transformation. **Plant Mol Biol**, 73:363–378. 2010.

PÉREZ-RODRÍGUEZ, P.; RIAÑO-PACHÓN, D.M.; CORRÊA, L.G.G.; RENSING, S.A.; KERSTEN, B.; MUELLER-ROEBER, B. PlnTFDB: updated content and new features of the plant transcription factor database. **Nucleic Acids Res**, 38: D822–D827. 2010.

PEVZNER, P. A. Educating biologists in the 21stcentury: bioinformatics scientists versus bioinformatics technicians. **Bioinformatics**, 20, 2159–2161. 2004.

POGGELER, S.; NOWROUSIAN, M.; RINGELBERG, C.; LOROS, J.J.; DUNLAP, J.C.; KUCK, U. Microarray and real-time PCR analyses reveal mating type-dependent gene expression in a homothallic fungus. **Mol Genet Genomics**., 275, 492–503. 2006.

POLLE, A. Dissection of the superoxide dismutase ascorbate-glutathione pathway by metabolic modeling: computer analysis as a step towards flux analysis. **Plant. Physiol.**, 126: 445–462. 2001.

POLLE, A.; OTTER, T.; SEIFERT, F. Apoplastic peroxidases and lignification in needles of Norway spruce (Picea abies L.). **Plant Physiology**, 106, 53–60. 1994.

POURABED, E; GHANE, G.F; SOLEYMANI, M.P.; RAZAVI S.M. Basic leucine zipper family in barley: genome-wide characterization of members and expression analysis. **Mol. Biotechnol**., 57: 12-26. 2015.

PRÉ, M.; ATALLAH, M.; CHAMPION, A.; DE VOS, M.; PIETERSE, C.M.J.; MEMELINK, J. The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. **Plant Physiology** 147: 1347–1357. 2008.

PROSDOCIMI, F.; CERQUEIRA, G.C.; BINNECK, E.; SILVA, A.F.; REIS, A.N.; JUNQUEIRA, A.C.; SANTOS, A.C.F.; NBANI, J.A.; WUST, C.I.; CAMARGO FILHO, F.; KESSEDJIAN, J.L.; PETRETSKI, J.H.; CAMARGO, L.P.; FERREIRA, R.G.M.; LIMA, R.P.; PEREIRA, R.M.; JARDIM, S.; SAMPAIO, V.S.; FLATSCHART, A.V.F. Bioinformática: manual do usuário. **Revista Biotecnologia Ciência e Desenvolvimento**, v. 5, n. 29, nov.;dez.; 2002.

PTASHNE, M.; GANN, A. Transcriptional activation by recruitment. **Nature**, 386, 569-577. 1997.

PURANIK, S.; BAHADUR, R.P.; SRIVASTAVA, P.S.; PRASAD, M. Molecular cloning and characterization of a membrane associated NAC family gene, SiNAC from foxtail millet [*Setaria italica* (L.) P. Beauv]. **Mol. Biotechnol**., 49, 138–150. 2011.

RABINOWICZ, P.D.; BRAUN, E.L.; WOLFE, A.D.; BOWEN, B.; GROTEWOLD, E. Maize R2R3 Myb genes: Sequence analysis reveals amplification in the higher plants. **Genetics**, 153. 427–444. 1999.

RAMSAY, R.G.; GONDA, T.J. MYB function in normal and cancer cells. **Nat. Rev. Cancer,** 8: 523–534. 2008.

RAO, G.R.; KORWAR, G.R.; SHANKER, A.K.; RAMAKRISHNA, Y.S. Genetic associations, variability and diversity in seed characters, growth, reproductive phenology and yield in *Jatropha curcas* (L.) accessions. **Trees**, v. 22, p. 697 - 709, 2008.

RATAJCZAK, R. Structure, function and regulation of the plant vacuolar $H(^{+})$-translocating ATPase. **Biochim Biophys Acta** 1465: 17–36. 2000.

REN, T.; QU, F.; MORRIS, T.J. HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. **Plant Cell**, 12, 1917–1925. 2000.

RENGASAMY, P. Soil processes affecting crop production in salt-affected soils. **Functional Plant Biology**, Victoria, v. 37, p. 613–620, 2010.

RIBEIRO, R.V.; MACHADO, E.C.; SANTOS, M.G.; OLIVEIRA, R.F. Photosynthesis and water relations of well-watered orange plants as affected by winter and summer conditions. **Photosynthetica**, v.47, p.215-222, 2009.

RICHMANN, J.L.; MEYEROWITZ, E.M. The Ap2/EREBP famly as plant transcription factors. **Biol. Chem**., 379, 633-646. 1998.

RIECHMANN, J.L.; RATCLIFFE, O.J. A genomic perspective on plant transcription factors. **Curr. Opin. Plant. Biol**., 3: 423-434. 2000.

RIZHSKY, L.; LIANG, H.; MITTLER, R. The combined effect of drought stress and heat shock on gene expression in tobacco. **Plant Physiol.**, 130: 1143–1151. 2002.

ROBATZEK, S.; SOMSSICH, I.E. Targets of AtWRKY6 regulation during plant senescence and pathogen defense. **Genes Dev.**, 16: 1139–1149. 2002.

RODRIGUEZ-URIBE, L.; O'CONNELL, M.A. A root-specific bZIP transcription factor is responsive to water deficit stress in tepary bean (*Phaseolus acutifolius*) and common bean (*P. vulgaris*). **J. Exp. Bot.** 57: 1391-1398. 2006.

RUBIO, F.; GASSMAN, W.; SCHROEDER, J.I. Sodium-driven potassium uptake by the plant potassium transporter HKT1 and mutations conferring salt tolerance. **Science** 270: 1660–1663. 1995.

RUS, A.; YOKAI, S.; SHARKHUU, A.; REDDY, M.; LEE, B.; MATSUMOTO, T.K.; KOIWA, H.; ZHU, J.K.; BRESSAN, R.A.; HASEGAWA, P.M. AtHKT1 is a salt tolerance determinant that controls Na entry into plant roots. **Proc Natl Acad Sci** USA 98: 14150–14155. 2001.

RUSHTON, P.J.; BOKOWIEC, M.T.; LAUDEMAN, T.W.; BRANNOCK, J.F.; CHEN, X.; TIMKO, M.P. TOBFAC: the database of tobacco transcription factors. **BMC Bioinformatics**, 9: 53. 2008.

RUSHTON, P.J.; REINSTADLER, A.; LIPKA, V.; LIPPOK, B.; SOMSSICH, I.E.; Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling, **Plant Cell,** 14, 749–762. 2002.

SABLOWSKI, R.W.; MEYEROWITZ, E.M. A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. Cell. 92: 93-Smalle J, Vierstra RD. 2004. The ubiquitin 26S proteasome proteolytic pathway. **Annu Rev Plant Biol**., 55:555-90. 1998.

SAHI, C.; AGARWAL, M.; REDDY, M.K.; SOPORY, S.K.; GROVER, A. Isolation and expression analysis of salt stress-associated ESTs from contrasting rice cultivars using a PCR-based subtraction method. **Theor Appl Genet**, v. 106 p. 620-628, 2003.

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

SAKAKIBARA, Y.; KOBAYASHI, H.; KASAMO, K. Isolation and characterization of cDNAs encoding vacuolar H1-pyrophosphates isoforms from rice (*Oryza sativa* L.). **Plant Mol Biol.** 31: 1029–1038. 1996.

SANDELIN, A.; ALKEMA, W.; ENGSTRÖM, P.; WASSERMAN, W.; LENHARD, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. **Nucleic Acids Res**., 32 (Database issue), D91–D94. 2004.

SANTI, G.R.; REINERT, D. J.; REICHERT, J. M.; SEQUINATTO, L.; OSÓRIO FILHO, B.; KUNZ, M.; FONTINELLI, F. Características físicas do solo da microbacia hidrográfica de Cândido Brum – Arvorezinha-RS.. In: REUNIÃO BRASILEIRA DE

MANEJO E CONSERVAÇÃO DE SOLO E ÁGUA, 14, Cuiabá, 2002. **Anais...** UFMT: Cuiabá, 2002.

SANTOS, R.V.; MURAOKA, T. Interações salinidade e fertilidade do solo. In: GHEYI, H.R.; QUEIROZ, J.E.; MEDEIROS, J.M. **Manejo e controle da salinidade na agricultura**. Campina Grande: UFPB, SBEA. 1997. p.289-315.

SATURNINO, H.M.; PACHECO, D.D.; KAKIDA, J.; TOMINAGA, N.; GONÇALVES, N.P. Cultura do pinhão-manso (Jatropha curcas L.). **Inf. Agropec**., 26:44-78, 2005.

SCANDALIOS, J.G.; GUAN, L.; POLIDOROS, A.N. Catalases in plants: gene structure, properties, regulation, and expression. In: **Oxidative stress and the molecular biology of antioxidant defenses**. Cold Spring Harbor Laboratory Press, U S A, vol 34, pp 343–406. 1997.

SCHACHTMAN, D.P.; SCHROEDER, J.I. Structure and transport mechanism of a high-affinity potassium uptake transporter from higher plants. **Nature**. 370: 655–658, 1994.

SCHERER, T.M.; SEELIG, B.; FRANZEN, D. **Soil, water and plant characteristics important to irrigation**. NRCS Publications, North Dakota Irrigation Guide, County Soil Survey Report , NDSU Extension Service, North Dakota. 1996.

SCHMID, M.; HALBWACHS, M.; WEHRLI, B.; WÜEST, A. Weak mixing in Lake Kivu: New insights indicate increasing risk of uncontrolled gas eruption. **Geochem. Geophys. Geosys**., 6, Q07009, p.11. 2005.

SCHMUTZ, J.; CANNON, S.B.; SCHLUETER, J.; MA, J.; MITROS, T.; NELSON, W. Genome sequence of the palaeopolyploid soybean. **Nature**, 463, 178–183. 2010.

SCHMUTZ, J.; CANNON, S.B.; SCHLUETER, J.; MA, J.; MITROS, T.; NELSON, W. Genome sequence of the palaeopolyploid soybean. **Nature**, 463, 178–183. 2010.

SCHNEIDER, H. **Métodos de análise filogenética**: um guia prático. 2ª ed. Ribeirão preto. Holos. 2003. 114p.

SCHRICK, K.; NGUYEN, D.; KARLOWSKI, W.M.; MAYER, K.F. START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. **Genome Biology**, 5, R41. 2004.

SHANG, H.; LI, W.; ZOU, C.; YUAN, Y. Analyses of the NAC transcription factor gene family in *Gossypium raimondii* Ulbr.: chromosomal location, structure, phylogeny, and expression patterns. **J. Integr. Plant Biol**., 55, 663–676. 2013.

SHANNON, M.C.; GRIEVE, C.M.; FRANCOIS, L.E. Whole-plant response to salinity. In: Wilkinson, R.E. (Ed.), **Plant-Environment Interactions**. Marcel Dekker, New York, pp. 199±244. 1994.

SHARMA, P.C.; GILL, K.S. Salinity-induced effect on biomass, yield, yield-attributing characters and ionic contents in genotypes of Indian mustard (*Brassica juncea*). **Indian J. Agric. Sci.** 64, 785±788. 1994.

SHARMA, P; DUBEY, R.S. Lead toxicity in plants. Braz. **J. Plant Physiol**., 17:35-52. 2005.

SHENDURE, J. The beginning of the end for microarrays? **Nature Methods**, v. 5, n. 7, p.585-587, 2008.

SHI, H.; ISHITANI, M.; KIM, C.; ZHU, J.K. The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na^+/H^+ antiporter. **Proc. Natl. Acad. Sci**. USA 97: 6896–6901. 2000.

SHIMIZU, T.; TOUMOTO, A.; IHARA, K.; SHIMIZU, M.; KYOGOKU, Y.; OGAWA, N.; OSHIMA, Y.; HAKOSHIMA, T. Crystal structure of PHO4 bHLH domain-DNA complex: Flanking base recognition. **EMBO J**. 16, 4689–4697. 1997.

SHIRAKI, T.; KONDO, S.; KATAYAMA, S.; WAKI, K.; KASUKAWA, T.; KAWAJI, H.; KODZIUS, R.; WATAHIKI, A.; NAKAMURA, M.; ARAKAWA, T. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. **Proc Natl Acad Sci**., 100: 15776–15781. 2003.

SHOGREN-KNAAK, M.; ISHII, H.; SUN, J.M.; PAZIN, M.J.; DAVIE, J.R.; PETERSON C.L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. **Science**, 311, 844-847. 2006.

SHORE P.; SHARROCKS, A.D. The MADS-box family of transcription factors. **Eur. J. Biochem**. 229:1–13. 1995.

SIEGEL, B. Plant peroxidases- an organismic perspective. **Plant Growth Regulation**., 12: 303-312. 1993.

SMART, L.B.; MOSKAI, W.A.; CAMERON, K.D.; BENNET, A.B. MIP genes are down-regulated under drought stress in *Nicotiana glauca*. **Plant and Cell Physiology**, 42: 686–693. 2001.

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

SOMMER, H.; BELTRAN, J.; HUIJSER, P.; PAPE, H.; LONNIG, W.; SAEDLER, H; SCHWARZ-SOMMER, Z. Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. **EMBO J**., 9, 605–613. 1990.

SONG, Y.; AI, C.R.; JING, S.J.; YU, D.Q. Research progress on function analysis of rice WRKY gene. **Rice Sci**. 17: 60–72. 2010.

SPERA, S.T.; DENARDIN, J.E.; ESCOSTEGUY, P.A.V.; SANTOS, H.P.; FIGUEROA, E.A. Dispersão de argila em microagregados de solo incubado com calcário. **Revista Brasileira de Ciência do Solo**, v. 32, n. especial, p.2613-2620, 2008.

SPITZ, F.; FURLONG, E.E. Transcription factors: from enhancer binding to developmental control. **Nat Rev Genet**., 13:613–626. 2012.

STOCKINGER, E.J.; GILMOUR, S.J.; THOMASHOW, M.F. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the

C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. **Proceedings of the National Academy of Sciences**. USA. 94. 1035–1040. 1997.

STRACK, D. Phenolic metabolism, p.387-416. In P.M. Dey e J.B. Harborne (eds.), **Plant biochemistry**. London, Academic Press, 554p. 1997.

SUDHIR, K., TAMURA, K., AND NEI, M. **MEGA**: Molecular Evolutionary Genetics Analysis, version 1.01. The Pennsylvania State University, University Park, PA. 1993.

SUN, B.T.; JING, Y.; CHEN, K.M.; SONG, L.L.; CHEN, F.J.; ZHANG, L.X. Protective effect of nitric oxide on iron deficiency-induced oxidative stress in maize (*Zea mays*). **Journal of Plant Physiology**, 164, pp. 536-543. 2007.

SUN, S.; YU, J.P.; CHEN, F.; ZHAO, T.J., FANG, X.H.; LI, Y. TINY, a dehydration-responsive element (DRE)-binding protein-like transcription factor connecting the DRE- and ethylene-responsive elementmediated signaling pathways in Arabidopsis. **J. Biol. Chem.** 283, 6261–6271. 2008.

SUZEK,B.E.; HUANG, H.; MCGARVEY, P.; MAZUMBER, R.; WU, C.H. UniRef: comprehensive and non-redundant UniProt reference clusters. **Bioinformatics**, 23, 1282–1288. 2007.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**, Porto Alegre, ARTMED. 5º ed., 2013, 918p.

TAIZ, L.; ZEIGER, E. **Plant Physiology**. 3.ed. Porto Alegre: Artmed, 2004. 719p.

TAJIMA, F.; NEI, M. Estimation of evolutionary distance between nucleotide sequences. **Mol Biol Evol.,** 1984; 1:269-285.

TAKAHASHI, R.; LIU, S.; TAKANO, T. Cloning and functional comparison of a high-affinity K+ transporter gene PhaHKT1 of salt-tolerant and salt-sensitive reed plants. **J. Exp. Bot.** 58, 4387–4395. 2007. doi: 10.1093/jxb/erm306

TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol. Biol. Evol.**, 28, 2731–2739. 2011.

TAN, S.; RICHMOND, T.J. Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. **Nature**, 39, 660-666. 1998.

TAPANES, N.C.O.; GOMEZ, A.D.A., CARNEIRO, M.J.W.; ANTUNES, C.O.A. Transesterification of *Jatropha curcas* oil glycerides: Theoretical and experimental studies of biodiesel reaction. **Fuel**, 87, 2286 – 2295. 2008.

TERAUCHI, R.; MATSUMURA, H.; KRÜGER, D. H.; KAHL, G. SuperSAGE: The most advanced transcriptoma technology for functional genomics. In: KAHL, G.; MEKSEM, K. (eds.). **The handbook of plant functional genomics**. 1ª ed. Weinheim: Wiley-VCH.

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.G. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res**., 25: 4876–4882. 1997.

TOLEDO-ORTIZ, G.; HUQ, E.; QUAIL, P.H. The Arabidopsis basic/helix-loop-helix transcription factor family. **Plant Cell**, 15(8): 1749-70. 2003.

TOMINAGA, N.; KAKIDA, J.; YASUDA, E.K. **Cultivo de pinhão-manso para produção de biodiesel**. Viçosa, MG, CPT, 2007. 220p.

TONACO, I.A.N.; BORST, J.W.; DE VRIES, S.C.; ANGENENT, G.C.; IMMINK, R.G.H. In vivo imaging of MADS-box transcription factor interactions. **J. Exp. Bot.**, 57: 33–42. 2006.

TONOIKE, H.; HAN, I.S.; JONGEWAARD, I.; DOYLE, M.; GUILTINAN, M.; FOSKET, D.E.Hypocotyl expression and light downregulation of the soybean tubulin gene, tubB1. **Plant J**., 5: 343–351. 1994.

TRAN, L.S.; NAKASHIMA, K.; SAKUMA, Y.; SIMPSON, S.D.; FUJITA, Y.; MARUYAMA, K.; FUJITA, M.; SEKI, M.; SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Isolation and functional analysis of Arabidopsis stress inducible NAC transcription factors that bind to a drought responsive cis-element in the early responsive to dehydration stress 1 promoter. **Plant Cell**, 16:2481–2498. 2004.

TRÖBNER, W.; RAMIREZ, L.; MOTTE, P.; HUE, I.; HUIJSER, P.; LONNIG, W.; SAEDLER, H.; SOMMER, H.; SCHWARZ-SOMMER, Z. GLOBOSA: a homeotic gene which interacts with DEFICIENS in the control of Antirrhinum floral organogenesis. **EMBO J.**, 11, 4693–4704. 1992.

TRON, A.E.; BERTONCINI, C.W.; CHAN, R.L.; GONZALEZ, D.H. Redox regulation of plant homeodomain transcription factors. **J. Biol. Chem.**, 277, 34800-34807. 2002.

TURCATTI, G.; ROMIEU, A.; FEDURCO, M.; TAIRI, A.-P. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. **Nucleic Acids Res.**, v.36, e25, 2008.

TURNER, N.C.; ABBO, S.; BERGER, J.D.; CHATURVEDI, S.K.; FRENCH, R.J.; LUDWIG, C.; YADAV, H.S. Osmotic adjustment in chickpea (Cicer arietinum L.) results in no yield benefit under terminal drought. **Journal of Experimental Botany**, 58, 187–194. 2007.

TUTEJA, N. Abscisic acid and abiotic stress signaling. **Plant Signal Behav.**, 2:135–138. 2007.

UAUY, C.; DISTELFELD, A.; FAHIMA, T.; BLECHL, A.; DUBCOVSKY, J. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. **Science**, 314:1298–301. 2006.

ULKER, B.; SOMSSICH, I.E. WRKY transcription factors: from DNA binding towards biological function. **Curr. Opin. Plant. Biol.**, 7: 491–498. 2004.

VERBRUGGEN, N.; HERMANS, C. Proline accumulation in plants: a review. **Amino Acids**, 35: 753 759. 2008.

VIGNALI, M. HASSAN, A.H. NEELY, K.E. WORKMAN, J.L. ATP-dependent chromatin-remodeling complexes **Mol. Cell. Biol.**, 20, pp. 1899-1910. 2000.

WANG, J.; ZHOU, J.; ZHANG, B.; VANITHA, J.; RAMACHANDRAN, S.; JIANG, S.Y. genome-wide expansion and expression divergence of the basic leucine zipper transcription factors in higher plants with an emphasis on sorghum. **J. Integr. Plant Biol.** 53, 212–231. 2011.

WANG, J.; ZHUANG, J.; IYER, S.; LIN, X.; WHITFIELD, T.W.; GREVEN, M.C.; PIERCE, B.G.; DONG, X.; KUNDAJE, A.; CHENG, Y. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. **Genome Res**., 22: 1798–1812. 2012.

WANG, J.; ZUO, K.; WU, W.; SONG, J.; SUN, X.; LIN, J.; LI, X.; TANG, K. Molecular cloning and characterization of a new Na+/H+ antiporter gene from *Brassica napus*. **DNA Seq**. 14 351–358. 2003.

WANG, L.; WANG, C.; QIN, L.; LIU, W.; WANG, Y. ThERF1 regulates its target genes via binding to a novel cis-acting element in response to salt stress. **J. Integr. Plant Biol.** 57, 838–847. 2015.

WANG, Z.; CHENG, K.; WAN, L.; YAN, L.; JIANG, H.; LIU, S. Genome-wide analysis of the basic leucine zipper (bZIP) transcription factor gene family in six legume genomes. **BMC Genomics**, 16:1053. 2015.

WANG, Z.; Gerstein; M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nat. Rev. Genet**., 10, 57–63. 2009.

WARD, J. A.; PONNALA, L.; WEBER, C. A. Strategies for transcriptome analysis in nonmodel plants. **American Journal of Botany**, v. 99, n. 2, p. 267-276, 2012.

WATERS, B.M.; UAUY, C.; DUBCOVSKY, J.; GRUSAK. M.A. Wheat (Triticum aestivum) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. **Journal of Experimental Botany**, 60, 4263–4274. 2009.

WEBSTER, G.L. Classification of the Euphorbiaceae. **Annals of the Missouri Botanical Garden**, v. 81, n. 1, p. 3 - 32, 1994.

WEI, K.; CHEN, J.; WANG, Y.; CHEN, Y.; CHEN, S.; LIN, Y. Genome wide analysis of bZIP-encoding genes in maize. **DNA Res.** 19, 463–476. 2012.

WHEELER, D.L. BARRETT, T. BENSON, D.A.; BRYANT, S.H.; CANESE, K.; CHETVERNIN, V.; CHURCH, D.M.; DICUCCIO, M.; EDGAR, R.; FEDERHEN, S.; GEER, L.Y.; HELMBERG, W.; KAPUSTIN, Y.; KENTON, D.L.; KHOVAYKO, O.; LIPMAN, D.J.; MADDEN, T.L.; MAGLOTT, D.R.; OSTELL, J.; PRUITT, K.D.; SCHULER, G.D.; SCHRIML, L.M.; SEQUEIRA, E.; SHERRY, S.T.; SIROTKIN, K.; SOUVOROV, A.; STARCHENKO, G.; SUZEK, T.O.; TATUSOV, R.; TATUSOVA, T.A.; WAGNER, L.; YASCHENKO, E. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Res**., 34, D173–180. 2006.

WHITMARSH, A.J.; DAVIS, R.J. Regulation of transcription factor function by phosphorylation. **Cell. Mol**. **Life Sci.,** 57, 1172±1183. 2000.

WILLEMSEN, V.; BAUCH, M.; BENNETT, T.; CAMPILHO, A.; WOLKENFELT, H.; XU, J.; HASELOFF, J.; SCHERES, B. The NAC domain transcription factors FEZ

and SOMBRERO control the orientation of cell division plane in Arabidopsis root stem cells. **Developmental Cell**, 15, 913–922. 2008.

WILSON, D.; CHAROENSAWAN, V.; KUMMERFELD, S.K.; TEICHMANN, S.A. DBD – taxonomically broad transcription factor predictions: new content and functionality. **Nucleic Acids Res**, 36: D88–D92. 2008.

WINGENDER, E. The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. **Bioinformatics**, 9, 326–333. 2008.

WU, C.A.; YANG, G.D.; MENG, Q.W.; ZHENG, C.C. The cotton GhNHX1 gene encoding a novel putative tonoplast $Na(^{+})/H(^{+})$ antiporter plays an important role in salt stress. **Plant Cell Physiol**. 45, 600–607. 2004.

WU, K.L.; GUO, Z.J.; WANG, H.H.; LI, J. The WRKY family of transcription factors in rice and Arabidopsis and their origins. **DNA Res.**, 12,9-26. 2005.

WYN JONES, R.G.,; STOREY, R.; LEIGH, R.A.; AHMAD, N.; POLLARD, A. A hypothesis on cytoplasmic osmoregulation. In: Regulation of cell membrane activities in plants, pp. 121-136, 1977. MarrY, E., Ciferri, O., eds. Elsevier/**North-ttolland Biomedical Press**, Amsterdam.

XU, X.; CHEN, C.; FAN, B.; CHEN, Z. Physical and functional interactions between pathogen-induced Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. **Plant Cell**, 18 1310–1326. 2006.

YAMAGUCHI, M.; KUBO, M.; FUKUDA, H.; DEMURA, T. VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. **Plant J**., 55: 652–664. 2008.

YAMAGUCHI, T.; BLUMWALD, E. Developing salt tolerant crop plants: challenges and opportunities. **Trends Plant Sci**. 10: 615–620. 2005.

YAMAMOTO, S.; SUZUKI, K.; SHINSHI, H. Elicitor-responsive, ethylene-independent activation of GCC box–mediated transcription that is regulated by both protein phosphorylation and dephosphorylation in cultured tobacco cells. **Plant J.** 20, 571–579. 1999.

YAMASAKI, K.; KIGAWA, T.; INOUE, M.; TATENO, M.; YAMASAKI, T.; YABUKI T.; AOKI, M.; SEKI, E.; MATSUDA, T.; TOMO, Y. Solution structure of an Arabidopsis WRKY DNA binding domain. **Plant Cell**, 17, 944-956. 2005.

YAMASAKI, K.; KIGAWA, T.; INOUE, M.; WATANABE, S.; TATENO, M.; SEKI, M.; SHINOZAKI, K.; YOKOYAMA, S. Structures and evolutionary origins of plant-specific transcription factor DNA-binding domains. **Plant Physiol Biochem**, 46 394–401. 2008.

YANG, Y.W.; NEWTON, R.J.; MILLER, F.R. Salinity tolerance in sorghum 1. Whole plant response to sodium chloride in S. bicolor and S. helepense; **Crop Science**; 30:775–781; 1990.

YANHUI, C.; XIAOYUAN, Y.; KUN, H.; MEIHUA, L.; JIGANG, L.; ZHAOFENG, G.; ZHIQIANG, L. The MYB transcription factor superfamily of Arabidopsis:

expression analysis and phylogenetic comparison with the rice MYB family. **Plant Mol. Biol**. 60, 107–124. 2006.

YE, Y.; TAM, N.F.Y.; LU, C.Y.; WONG, S.H. Effects of salinity on germination, seedling growth and physiology of three salt secreting mangrove species. **Aquat. Bot**. 83: 193–205. 2005.

YILMAZ, A.; MEJIA-GUERRA, M.K.; KURZ, K.; LIANG, X.; WELCH, L.; GROTEWOLD, E. AGRIS: The Arabidopsis Gene Regulatory Information Server, an update. **Nucleic Acids Res**. 39: D1118–D1122. 2011.

YOO, S.Y.; KIM, Y.; KIM, S.Y.; LEE, J.S.; AHN, J.H. Control of flowering time and cold response by a NAC-Domain protein in arabidopsis. **PLoS ONE**. 2:e642. 2007.

YU, H.; CHEN, X.; HONG, Y. Y., WANG, Y., XU, P., KE, S.D. Activated expression of an Arabidopsis HD-START protein confers drought tolerance with improved root system and reduced stomatal density. **Plant Cell**, 20, 1134–1151. 2008.

ZACHGO, S.; DE ANDRADE SILVA, E.; MOTTE, P.; TROBNER, W.; SAEDLER, H.; SCHWARZ-SOMMER, Z. Functional analysis of the Antirrhinum floral homeotic DEFICIENS gene in vivo and in vitro by using a temperature-sensitive mutant. **Development**, 121, 2861–2875. 1995.

ZARET, K.S.; CARROLL, J.S. Pioneer transcription factors: establishing competence for gene expression. **Genes Dev**., 25: 2227–2241. 2011.

ZEKRI, M. Effects of sodium chloride on growth and physiology of sour orange and *Cleopatra mandarin* seedlings. **Scientia Hort**. 47: 305-316. 1991.

ZHANG, H.; HAN, B.; WANG, T.; CHEN, S.; LI H. Mechanisms of plant salt response: insights from proteomics. **J. Proteome Res.,** 11: 49–67. 2012.

ZHANG, X.; ZHANG, Z.; CHEN, J.; CHEN, Q.; WANG, X.C. HUANG, R.Expressing TERF1 in tobacco enhances drought tolerance and abscisic acid sensitivity during seedling development. **Planta**, 222:494–501. 2005.

ZHANG, Z.-L.; XIE, Z.; ZOU, X.; CASARETTO, J.; HO, T.-H.D.; SHEN, Q.J. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. **Plant Physiol**., 134, 1500–1513. 2004.

ZHAO, M.; MOROHASHI, K.; HATLESTAD, G.; GROTEWOLD, E.; LLOYD, A. The TTG1–bHLH–MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. **Development**, 135, 1991–1999. 2008.

ZHENG, X.; CHEN, B.; LU, G.; HAN, B. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. **Biochem Biophys Res Commun**, 379:985–989. 2009.

ZHENG, Y.; JIAO, C.; SUN, H.; ROSLI, H.G.; POMBO, M.A.; ZHANG, P.; BANF, M.; DAI, X.; MARTIN, G.B.; GIOVANNONI, J.J.; ZHAO, P.X.; RHEE, S.Y.; FEI, Z. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. **Molecular Plant**, 9:1667-1670. 2016.

ZHONG, H.; GUO, Q.Q..; CHEN, L.; REN, F.; WANG, Q.Q.; ZHENG, Y. Two Brassicanapus genes encoding NAC transcription factors are involved in response to high-salinity stress. **Plant Cell Rep**. 31,1991–2003.

ZHONG, R.; DEMURA, T.; YE, Z.H. SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. **Plant Cell**, 18, 3158–3170. 2006.

ZHOU, Q.Y.; TIAN, A.G.; ZOU, H.F.; XIE, Z.M.; LEI, G.; HUANG, J.; WANG, C.M.; WANG, H.W.; ZHANG, J.S.; CHEN, S.Y. Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. **Plant Biotechnol J.**, 6:486–503. 2008.

ZHU, J.K. Regulation of ion homeostasis under salt stress. **Current Opinion in Plant Biology**, v. 6, n. 5, p. 441-445, 2003.

ZHU, J-K. Plant salt tolerance. **Trends in Plant Science**, v. 6, n. 2, p. 66-71, 2001.

ZIMMERMANN, S.; HARTJE, S.; EHRHARDT, T.; PLESCH, G.; MUELLER-ROEBER, B. The K+ channel SKT1 is co-expressed with KST1 in potato guard cells–both channels can co-assemble via their conserved KT domains. **Plant J**. 28: 517–527. 2001.