



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

RAHISA HELENA DA SILVA

**GENÔMICA ESTRUTURAL, COMPARATIVA E FUNCIONAL DE GENES
CODIFICADORES DE HELICASES EXPRESSOS EM PINHÃO-MANSO (*Jatropha
curcas* L.) SOB ESTRESSE SALINO**

Recife

2024

RAHISA HELENA DA SILVA

**GENÔMICA ESTRUTURAL, COMPARATIVA E FUNCIONAL DE GENES
CODIFICADORES DE HELICASES EXPRESSOS EM PINHÃO-MANSO (*Jatropha
curcas* L.) SOB ESTRESSE SALINO**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutora em Genética e Biologia Molecular. Área de concentração: Biologia Molecular.

Orientador: Éderson Akio Kido

Coorientador: Antonio Félix da Costa

Recife

2024

Catálogo na Fonte:
Bibliotecária Natália Nascimento, CRB4/1743

Silva, Rahisa Helena.

Genômica estrutural, comparativa e funcional de genes codificadores de helicases expressos em pinhão-mansô (*Jatropha curcas* L.) sob estresse salino. / Recife. – 2024.

217f. : il., fig.

Orientador: Éderson Akio Kido.

Coorientador: Antônio Félix da Costa.

Tese (Doutorado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-graduação em Genética e biologia molecular, 2024.

Inclui referências.

1. Euphorbiaceae. 2. RNA helicase. 3. DNA helicase. 4. Bioinformática. 5. Transcriptômica. 6. Estresse abiótico. I. Kido, Éderson Akio. (Orient.). II. Costa, Antônio Félix da. (coorient.). III. Título.

572.8

CDD (22.ed.)

UFPE/CB – 2024-115

RAHISA HELENA DA SILVA

**Genômica Estrutural, Comparativa e Funcional De Genes Codificadores de
Helicases Expressos em Pinhão-Manso (*Jatropha curcas* L.) Sob Estresse
Salino**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutora em Genética e Biologia Molecular. Área de concentração: Biologia Molecular.

Aprovado em: 29/02/2024

BANCA EXAMINADORA:

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

Profª. Dra. Ana Christina Brasileiro Vidal (Examinador interno)
Universidade Federal de Pernambuco

Profª. Dra. Valesca Pandolfi (Examinador interno)
Universidade Federal de Pernambuco

Profº. Dr. Marcus Vinicius Loss Sperandio (Examinador externo)
Universidade Federal Rural de Pernambuco

Dr. Nataniel Franklin de Melo
Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA Semiárido (Examinador externo)

Dedico este trabalho a meu avô
Antônio Pedro, por sua bravura e
simplicidade.

AGRADECIMENTOS

À Universidade Federal de Pernambuco por ter sido minha casa ao longo dessa jornada e me proporcionar formação de alta de qualidade.

Ao Programa de Pós Graduação em Genética e Biologia Molecular (PPGBM) pela oportunidade.

À Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE) pela concessão da bolsa.

Ao Laboratório de Genética Molecular de Plantas (LGMP) pelo suporte com infraestrutura.

Ao meu orientador Professor Dr. Éderson Kido pela confiança, paciência e disponibilidade. Obrigada por tudo!

Ao meu coorientador Dr. Antonio Félix por ser sempre tão solícito e gentil.

Aos membros da banca examinadora Dra. Ana Christina, Dra. Valesca Pandolfi, Dr. Marcus Vinicius Loss e Dr. Nataniel Franklin por contribuírem de forma tão significativa para a qualidade deste trabalho.

À equipe do PPGBM por sua cordialidade e disponibilidade.

À coordenação do PPGBM formada pelas Professoras Dra. Ana Christina Brasileiro Vidal e Dra. Ana Maria Benko-Iseppon por sua atenção e zelo.

Ao meu esposo Dr. Manassés Daniel por ser meu maior incentivador, melhor amigo e por sempre enxergar o melhor em mim.

A meus pais Arnaldo e Luzinete, e minha irmã Thaysa que são a base da minha vida.

À equipe do LGMP, em especial, Dra. Bruna Brito, Dra. Elvia Oliveira, Dra. Francielly Negreiros e Dra. Vanessa Emanuelle. Obrigada por todo conhecimento compartilhado e por terem tornado a rotina muito mais leve. Teria sido muito mais difícil sem vocês!

A minhas queridas Carol, Luciele, Lilia e Sâmela, por todo carinho e amizade inestimáveis.

A Deus, por guiar meus caminhos e me manter firme a cada passo.

A todos, que direta ou indiretamente contribuíram para a conclusão deste trabalho.

Aqueles que sonham de dia estão
cientes de muitas coisas que
escapam aos que sonham apenas à
noite.

Edgar Allan Poe

RESUMO

O pinhão-mansô (*Jatropha curcas* L.) é uma oleaginosa não comestível com potencial para produção de biodiesel. Apesar de tolerante à seca, o pinhão-mansô é sensível a salinidade do solo, um dos principais estresses abióticos. Quando submetidas a estresses, o metabolismo das plantas recorre a uma variedade de grupos gênicos que irão atuar para manter o equilíbrio celular. Dentre esses grupos gênicos, as helicases têm sido associadas a vias de tolerância a estresses. Logo, o objetivo deste trabalho foi identificar e caracterizar os genes helicase de pinhão-mansô e avaliar sua resposta a salinidade. Foram identificados 130 genes codificadores de helicases para pinhão-mansô. Destes, 79 codificam RNA helicases, 25 DNA helicases e 26 remodeladores da cromatina. Todas as helicases de pinhão-mansô compartilham o *core* conservado, que consiste na presença dos motivos *Walker A* e *Walker B* dispostos em dois domínios centrais. As regiões promotoras desses genes apresentaram candidatos a elementos *cis* regulatórios importantes no contexto de respostas a estresses abióticos, tais como *Dof-type*, AP2-ERF, HD-ZIP, WRKY, bHLH dentre outros. Dados de RNA-Seq permitiram identificar 103 transcritos helicase diferencialmente expressos (p -value < 0.0001; FDR < 0.005) a estímulo de salinidade (150 mM de NaCl por três horas). Redes PPI evidenciaram uma variedade de processos envolvendo essas helicases incluindo vias de *splicing*, reparo do DNA, biogênese ribossomal, *turnover* de RNAs e proteínas, síntese de fosfoinosítídeos, transporte vesicular, manutenção de cloroplastos e mitocôndrias dentre outros. Estes resultados podem auxiliar o desenvolvimento de marcadores associados a tolerância a salinidade nos programas de melhoramento genético de plantas.

Palavras-chave: Euphorbiaceae; RNA helicase; DNA helicase; Bioinformática; Transcriptômica; Estresse abiótico.

ABSTRACT

Physic nut (*Jatropha curcas* L.) is a non edible oilseed that has potential for biodiesel production. Despite being drought tolerant, physic nut plants are sensitive to soil salinity, one of the main abiotic stresses. When subjected to stress, plant metabolism resorts to a variety of gene groups that will act to maintain cellular balance and thus ensure the survival of these organisms. Among these gene groups, helicases have been associated with stress tolerance pathways. The objective of this work was to identify and characterize the *Jatropha curcas* helicase genes and evaluate their response to salinity. Therefore, 130 genes encoding helicases were identified for *Jatropha curcas*. Of these, 79 encode RNA helicases, 25 DNA helicases and 26 chromatin remodelers. All *Jatropha curcas* helicases share the conserved core, which consists of the presence of the Walker A and Walker B motifs arranged in two central domains. The promoter regions of these genes presented candidates for important cis-regulatory elements in the context of responses to abiotic stresses, such as Dof-type, AP2-ERF, HD-ZIP, WRKY, bHLH among others. RNA-Seq data allowed the identification of 103 differentially expressed helicase transcripts (DEs; p-value < 0.0001; FDR < 0.005) under salinity stimulation (150 mM NaCl for three hours). Protein-protein interactions (PPI) networks have highlighted a variety of processes involving these helicases, including splicing pathways, chromatin remodeling, DNA repair, ribosomal biogenesis, RNA and protein turnover, phosphoinositide synthesis, vesicular transport, maintenance chloroplasts and mitochondria, among others. These results can help the development of markers associated with salinity tolerance in plant genetic improvement programs.

Keywords: Euphorbiaceae; RNA Helicase; DNA helicase; Bioinformatics; Transcriptomics; Abiotic stress.

LISTA DE ILUSTRAÇÕES

REVISÃO DE LITERATURA

Figura 1 -	Centro de origem e locais de ocorrência de <i>Jatropha curcas</i>	25
Figura 2 -	Características morfológicas de pinhão-manso (<i>J. curcas</i>).....	27
Figura 3 -	Principais limitações de <i>J. curcas</i> como cultura de biodiesel.....	28
Figura 4 -	Aplicabilidades de <i>Jatropha curcas</i>	32
Figura 5 -	Esquema da cascata de sinalização das espécies reativas de oxigênio (ROS), ativação da via MAPK e homeostase redox durante estresses abióticos.....	34
Figura 6 -	Representação esquemática dos efeitos da salinidade nas plantas.....	36
Figura 7 -	Representação gráfica de componentes da via SOS (<i>Salt Overly Sensitive</i>) que atuam na mantém a homeostase iônica sob estresse salino.....	37
Figura 8 -	Esquema das divisões da superclasse NTPases P-loop, com descrição do grupo ASCE (<i>Additional Strand Catalytic Glutamate</i>) no qual estão incluídas as superfamílias das helicases.....	38
Figura 9 -	Representação esquemática dos motivos conservados encontrados em helicases. Os motivos estão coloridos de acordo com sua atividade bioquímica principal: em roxo, ligação e hidrólise de ATP; em verde, coordenação entre sítios de ácidos nucleicos e sítios NTP; em rosa, ligação a ácido nucleico.....	39
Figura 10 -	Descrição dos motivos conservados em cada superfamília de helicase. Motivos A, B e R são comuns e os motivos em verde são únicos a determinada superfamília.....	40

Figura 11 -	Representação da subdivisão das helicases em famílias de acordo com cada uma das seis superfamílias descritas. Famílias de RNA helicase são mostradas em preto e DNA helicase em cinza.....	41
Figura 12 -	Participação das RNA helicases nos processos que envolvem o metabolismo de RNA numa célula.....	49
Figura 13 -	Fluxo de um <i>pipeline</i> para obtenção de um genoma vegetal. As ferramentas referentes a cada etapa são mencionadas ao lado.....	56
Figura 14 -	Diagrama ilustrativo da linha do tempo mostrando a evolução das principais técnicas de análise da expressão gênica.....	60
Figura 15 -	Representação esquemática das etapas relativas à obtenção de um transcriptoma RNA-Seq.....	61
Figura 16 -	Representação esquemática das etapas de análise de um RNA-Seq.....	63
Figura 17 -	Representação gráfica das fases da qPCR. A. <i>Baseline</i> ; B. Fase exponencial; C. Platô.....	64
ARTIGO 1 - ANÁLISE DA FAMÍLIA SNF2 DE PINHÃO-MANSO (<i>Jatropha curcas</i> L.) E PERFIL DE EXPRESSÃO EM RESPOSTA À SALINIDADE		
Figura 1 -	Estrutura éxon-íntron dos genes <i>JcSnf2</i> . A representação gráfica foi obtida a partir do GSDS (<i>Gene Structure Display</i>).....	78
Figura 2 -	Regiões promotoras dos genes <i>JcSnf2</i> (1.0 kb <i>upstream</i> a base +1) e presença de potenciais elementos <i>cis</i> regulatórios (p-value < 0,01; q-value < 0,01) preditos a partir do portal PlantRegMap.....	79

Figura 3 -	Árvore fenética da família Snf2 de pinhão-manso. A árvore foi construída pelo método <i>Neighbor joining</i> com <i>bootstrap</i> de 1.000 replicatas a partir do <i>software</i> MEGA11.....	80
Figura 4 -	Organização dos domínios das proteínas JcSnf2. A escala representa o tamanho em aminoácidos das proteínas. Os símbolos representam os diferentes domínios presentes nas sequências.....	82
Figura 5 -	Motivos conservados detectados nas sequências das proteínas JcSnf2 A. e sua distribuição de acordo com a classificação em subfamílias B. Os motivos são nomeados: Q, I, Ia, II, III, IV, V e VI; Os blocos conservados presentes entre as regiões dos motivos: E, F, A, G, H, B, J, C, K, D, L, M e N. Os motivos foram detectados pelo programa MEME e a visualização foi feita através do <i>software</i> TBtools.....	83
Figura 6 -	Modelos 3D obtidos a partir do Swissmodel para as proteínas JcSnf2. Os domínios conservados estão realçados em cada modelo; Domínios <i>core</i> SNF2 (N-terminal) e HELICc (C-terminal) estão representados em roxo e amarelo, respectivamente. Domínios auxiliares incluem RING (<i>greencyan</i>); HIRAN (<i>lighmagenta</i>); DUF3535 (<i>deepteal</i>); SAWADEE (<i>pink</i>); CHROMO (<i>green</i>); SANT (<i>green forest</i>); DUF1086 (<i>orange</i>); DUF1087 (<i>blue</i>); SnAC (<i>red</i>); HNHc (<i>purpleblue</i>).....	85
Figura 7 -	Clusterização hierárquica baseada nos valores de Log2FC dos transcritos RNA-Seq JcSnf2 diferencialmente expressos (p-value < 0.0001; FDR < 0.005) nas raízes do acesso Jc171 submetido a estímulo salino (150 mM de NaCl por três), bem como, respectiva modulação no acesso Jc183.....	

Figura 8 -	Expressão gênica relativa calculada a partir de dados qPCR de transcritos Snf2 (<i>JcSnf2</i>) expressos em raízes do acesso <i>Jc171</i> pinhão-mansó submetido a estímulo de salinidade (150 mM de NaCl por três horas). O eixo Y representa valores de expressão relativa e o eixo X os transcritos alvo; ‘*’: alvos com expressão validada. <i>Downregulation</i> (DR) é indicada em azul, <i>upregulation</i> (UR) em vermelho.....	88
Figura 9 -	Termos GO enriquecidos (p-value < 1.0e-16) associados as redes de interação PPI obtidas considerando os candidatos <i>JcSnf2</i> expressos nas bibliotecas RNA-Seq obtidas em ensaio de salinidade (150 mM de NaCl por 3 horas). Os termos representam A. processo biológico; B. função molecular; C. componente celular.....	90
Figura 10 -	Redes de interação proteína-proteína (PPI) construídas através do STRING utilizando <i>Arabidopsis</i> como modelo (<i>high score</i> > 0.7), considerando A. todo o conjunto de transcritos <i>JcSnf2</i> diferencialmente expressos nas bibliotecas RNA-Seq obtidas a partir de ensaio de salinidade (150 mM de NaCl) evidenciando dois clusters distintos que correspondem a remodelamento da cromatina (laranja) e reparo do DNA (amarelo); em B. C. e D. redes individuais considerando transcritos <i>JcSnf2</i> diferencialmente expressos a partir dos dados de qPCR.....	91
	ARTIGO 2 - CARACTERIZAÇÃO ESTRUTURAL DOS GENES CODIFICADORES DE DNA HELICASES DE PINHÃO-MANSO (<i>Jatropha curcas</i> L.)	
Figura 1 -	A. Candidatos a elementos <i>cis</i> -regulatórios detectados nos promotores (1 kb <i>upstream</i> ao sítio de início da transcrição) dos genes DNA helicase de pinhão-mansó; B. número de elementos <i>cis</i> detectados em cada promotor divididos em 14 tipos.....	117

Figura 2 -	Árvores fenéticas construídas pelo método <i>Neighbor Joining</i> (<i>bootstrap</i> de 1.000 replicatas) evidenciando as A. subfamílias MCM2 – MCM10 para a família MCM; B. subfamílias RuvBL1- RuvBL13 para a família RuvB; e C. subfamílias UvrD1 e UvrD2 para a família UvrD.....	119
Figura 3 -	Estruturas éxon-ínton dos genes DNA helicase de pinhão-mansão.....	120
Figura 4 -	Estrutura dos domínios conservados detectados nas DNA helicases <i>JcMCM</i> , <i>JcRuvBL</i> e <i>JcUvrD</i> de acordo com as respectivas subfamílias. Os domínios foram detectados através dos <i>softwares</i> SMART e InterProScan.....	124
Figura 5 -	Distribuição dos motivos conservados detectados nas DNA helicases <i>JcMCM</i> , <i>JcRuvBL</i> e <i>JcUvrD</i> de acordo com as respectivas subfamílias. Os motivos foram identificados através dos programas MEME e Jalview*.....	125
Figura 6 -	Estruturas secundárias detectadas nas proteínas codificadas por genes DNA helicase de pinhão-mansão para as famílias A. <i>JcMCM</i> ; B. <i>JcRuvBL</i> e C. <i>JcUvrD</i> . Retângulos verdes realçam os motivos conservados característicos de cada família.....	126
Figura 7 -	Modelos 3D gerados para DNA helicases MCM de pinhão-mansão através do Swiss-model. Estruturas MCM são divididas em domínios MCM N-terminal (NTD; <i>blue</i>), OB-fold (<i>orange</i>) e domínios C-terminal (CTD; <i>green</i>); MCM_lid (<i>purple blue</i>); Extensões NTE e CTE estão realçadas em vermelho.....	128
Figura 8 -	Modelos 3D gerados para DNA helicases RuvBL de pinhão-mansão através do Swiss-model. Estruturas dos domínios core estão divididas em RuvB_N (NTD; <i>blue</i>); AAA (CTD; <i>green</i>).....	130

Figura 9 -	Modelo 3D gerado para DNA helicase JcUvrD1 de pinhão-manso através do Swiss-model. Estruturas dos domínios core estão divididas em UvrD_N (1A; <i>blue</i>); UvrD_C (2A; <i>green</i>); Subdomínios 1B (<i>yellow</i>) e 2B (<i>orange</i>). Motivos conservados na fenda entre os domínios 1 e 2 estão realçados em vermelho.....	131
Figura 10 -	Rede de interação proteína-proteína (PPI) considerando as proteínas JcMCM. A rede foi construída com base nas informações disponíveis para <i>Arabidopsis</i> . O <i>score</i> de confiança foi definido como <i>high</i> (> 0.7). Evidências baseadas em <i>text mining</i> foram excluídas. A rede mostra parceiros relacionados com a replicação do DNA (rosa) e controle do ciclo celular (amarelo).....	134
Figura 11 -	Rede de interação proteína-proteína (PPI) considerando as proteínas JcRuvBL. A rede foi construída com base nas informações disponíveis para <i>Arabidopsis</i> . O <i>score</i> de confiança foi definido como <i>high</i> (> 0.7). Evidências baseadas em <i>text mining</i> foram excluídas. A rede mostra parceiros relacionados com remodelamento da cromatina (laranja); replicação do DNA (amarelo); transporte de proteínas a matriz do peroxissomo (roxo); desmontagem do complexo SNARE (verde); degradação de proteínas via ubiquitinação (azul); biogênese do cloroplasto (rosa).....	135
Figura 12 -	Rede de interação proteína-proteína (PPI) considerando a proteína JcUvrD1. A rede foi construída com base nas informações disponíveis para <i>Arabidopsis</i> . O <i>score</i> de confiança foi definido como <i>high</i> (> 0.7). Evidências baseadas em <i>text mining</i> foram excluídas. A rede mostra parceiros relacionados com a replicação do DNA (verde) e reparo do DNA (roxo).....	135

LISTA DE TABELAS

REVISÃO DE LITERATURA

Tabela 1 -	Propriedades medicinais exibidas por diferentes partes vegetativas de pinhão-manso.....	29
Tabela 2 -	Propriedades de pinhão-manso (<i>J. curcas</i>) como bio-pesticida contra diferentes patógenos e pragas vegetais.....	30
Tabela 3 -	Ferramentas utilizadas na etapa de genômica estrutural.....	57
Tabela 4 -	Informações gerais sobre as montagens disponíveis para <i>J. curcas</i> , incluindo: <i>status</i> da montagem; tecnologia de sequenciamento empregada; quantidade de genes codificadores de proteínas; cobertura; valores de N50 e L50 referentes aos <i>scaffolds</i> das respectivas montagens.....	58
ARTIGO 1 - ANÁLISE DA FAMÍLIA SNF2 DE PINHÃO-MANSO (<i>Jatropha curcas</i> L.) E PERFIL DE EXPRESSÃO EM RESPOSTA À SALINIDADE		
Tabela 1 -	Características dos genes codificadores das proteínas Snf2 de pinhão-manso (<i>JcSnf2</i>), incluindo nome, <i>locus</i> id, localização genômica, subfamília, tamanhos dos genes, da CDS e número de íntrons.....	77
Tabela 2 -	Características das proteínas Snf2 de pinhão-manso (<i>JcSnf2</i>), incluindo nome (quantidade de proteínas do <i>loci</i>), proteína id, subfamília, tamanhos em aminoácidos (aa) das proteínas primárias, peso molecular (MW), ponto isoelétrico (pI) e localização subcelular.....	81
Tabela 3 -	Expressão gênica relativa dos transcritos RNA-Seq <i>JcSnf2</i> expressos no acesso <i>Jc171</i> de pinhão-manso após exposição a 150 mM de NaCl por três horas. Dados foram obtidos a partir do <i>software</i> REST© (v.2.0.13).....	88

ARTIGO 2 - CARACTERIZAÇÃO ESTRUTURAL DOS
GENES CODIFICADORES DE DNA HELICASES DE
PINHÃO-MANSO (*Jatropha curcas* L.)

Tabela 1 -	Informações acerca dos genes codificadores de DNA helicase de pinhão-manso abrangendo as famílias MCM, RuvB e UvrD.....	115
Tabela 2 -	Características das proteínas codificadas pelos genes DNA helicase de pinhão-manso. MW: peso molecular; pI: ponto isoelétrico. Aa: tamanho em aminoácidos das proteínas; pI: ponto isoelétrico; MW: peso molecular; MP: membrana plasmática.....	121
Tabela 3 -	Estruturas secundárias detectadas pelo SOPMA a partir das proteínas das famílias DNA helicase <i>JcMCM</i> , <i>JcRuvB</i> e <i>JcUvrD</i>	127
Tabela 4 -	Parâmetros de qualidade obtidos para os modelos tridimensionais construídos pelo Swissmodel a partir das proteínas das famílias DNA helicase de pinhão-manso. Sm: similaridade; Cv: cobertura; Rf: Ramachandran favoured.....	132

LISTA DE ABREVIATURAS E SIGLAS

AA	Ácido ascórbico
AAA+	<i>ATPases Associated with diverse cellular Activities</i> (ATPases associadas com várias atividades moleculares)
ABA	Ácido abscísico
ABC	<i>ATP-binding cassette</i> (cassete de ligação ao ATP)
ABRE	<i>ABA responsive element</i> (Elemento ABA responsivo)
ACE	<i>Angiotensin converting enzyme</i> (Enzima conversora de angiotensina)
ASCE	<i>Additional Strand Catalytic Glutamate</i> (Glutamato catalítico adicional)
ASGAL	<i>Alternative Splicing Graph Aligner</i>
BACs	<i>Bacterial Artificial Chromosomes</i> (Cromossomo Artificial de Bactéria)
BER	<i>Base excision repair</i> (Reparo por Excisão de base)
BR	Brassinosteróides
BRAD	<i>The Brassicaceae Database</i>
BUSCO	<i>Benchmarking Universal Single-Copy Orthologs</i>
BWA	<i>Burrow Wheelers alignment</i>
CAT	Catalase
CBF	<i>C-repeat/DRE-Binding Factor</i>
CDD	<i>Conserved Domain Database</i>
CDPK	<i>Calcium-dependent protein kinase</i> (Proteína quinase dependente de cálcio)
ChromDB	<i>The Chromatin Database</i>
CK	Citocinina
CO	Monóxido de carbono
DBG	<i>De Bruijn Graph</i> (gráfico de Bruijn)
DRE	<i>Dehydration-responsive element</i> (Elemento responsivo à desidratação)
DSB	<i>Double stranded break</i> (Reparo de quebra na dupla fita)
dsRNA	<i>Double-stranded RNA</i> (RNA dupla fita)
ESTs	<i>Expressed sequence Tags</i> (Etiqueta de sequência expressa)

FAC	<i>Fatty acid composition</i> (Composição de ácidos graxos)
FPKM	<i>Fragments per kilobase of transcript per million mapped reads</i> (fragmentos por quilobase de transcritos por milhão de <i>reads</i> mapeadas)
GA	Giberelina
GHG	Gases do efeito estufa
GI	GIGANTEA
GO	<i>Gene ontology</i>
GR	Glutathione redutase
GSDS	<i>Gene Structure Display Server</i>
HIRAN	HIP116, Rad5p N-terminal
HMM	<i>Hidden Markov Model</i>
HR	Reparo de junções homólogas
HRDC	<i>RNase D C-terminal domain</i>
HT-SuperSAGE	<i>High-Throughput-SuperSAGE</i>
IAA	<i>Indoleacetic acid</i> (Ácido indolacético)
KEA 1, 2 e 3	Transportadores plastidiais de K ⁺
KEGG	<i>Kyoto encyclopedia of genes and genomes</i>
LongSAGE	<i>Long Serial Analysis of Gene Expression</i>
MAPKKK	<i>Mitogen-Activated Protein Kinase Kinase Kinase</i>
MATS	<i>Multivariate Analysis of Transcript Splicing</i>
MCM	<i>Mini chromosome maintenance</i>
MEGA	<i>Molecular Evolutionary Genetics Analysis</i>
MEME	<i>Multiple EM for Motif Elicitation</i>
MMS	Metil metanosulfonato
MSL8	<i>MscS-like 8</i>
NER	Reparo por excisão de nucleotídeo
NeVOmics	<i>Network-based Visualization for Omics</i>
NGS	<i>Next-generation sequencing</i> (Sequenciamento de nova geração)
NHEJ	Reparo de junções não-homólogas
NHX	trocador vacuolar de Na ⁺ /H ⁺
NMD	<i>nonsense mediated mRNA decay</i> (
MMR	<i>Mismatch repair</i> (Reparo de <i>mismatch</i>)

NTP	Nucleosídeos trifosfato
ORFs	<i>Open Reading Frames</i> (Quadro aberto de leitura)
OSCA1	<i>Induced Ca²⁺ Increase 1</i>
OsCAX	Trocador catiônico do tipo CAX
OstPC1	Proteína de canal de cálcio de dois poros
PCA	<i>Principal Component Analysis</i> (Análise de componente principal)
PDH45	<i>Pea DNA helicase</i>
PHD	<i>Plant homeodomain</i>
PlantCARE	<i>Plant Cis Acting Regulatory element</i>
PlantPAN	<i>Plant Promoter Analysis Navigator</i>
PlantProm	<i>Plant Promoter Database</i>
PlantsP	<i>The Plant Phosphorylation Database</i>
POX	Peroxidase
PTGS	<i>Post-transcriptional gene silencing</i> (Silenciamento gênico pós-transcricional)
QUAST	<i>Quality Assessment Tool for Genome Assemblies</i>
RBFs	<i>Ribosome biogenesis factors</i> (Fatores da biogênese do ribossomo)
RBR	<i>Ring between ring</i>
RGAP	<i>The Rice Genome Annotation Project</i>
RISC	<i>RNA-induced silencing complex</i> (Complexo de silenciamento induzido por RNA)
RNPs	Ribonucleoproteínas
ROS	<i>Reactive Oxygen Species</i> (Espécies reativas de oxigênio)
RPKM	<i>reads per kilobase of transcript per million mapped reads</i>
RQC	<i>RecQ C-terminal domain</i> (Domínio RecQ C-terminal)
SA	Ácido salicílico
SAGE	<i>Serial Analysis of Gene Expression</i>
SL	Strigolactona
SMART	<i>Simple Modular Architecture Research Tool</i>
SnRK2s	<i>Sucrose nonfermenting-1-related protein kinases 2</i> (Proteína quinase 2 relacionada à sacarose não fermentadora 1)
snRNPs	Ribonucleoproteínas nucleares
SOD	Superóxido dismutase

SOS	<i>Salt Overly Sensitive</i>
SOS2	Proteína quinase serina/treonina
SRA	<i>Sequencing Read Archive</i>
STAND	<i>Signal Transduction ATPases with Numerous Domains</i> (ATPases de transdução de sinal com vários domínios)
SuperSAGE	<i>Super Serial Analysis of Gene Expression</i>
TAIR	<i>The Arabidopsis Genome Initiative</i>
TBtools	<i>Toolkit for Biologists integrating various biological data</i>
TKM	<i>transcripts per million</i> (Transcritos por milhão)
Upf1	<i>Up-frameshift 1</i>
UPR	<i>unfolded protein response</i> (Via de resposta de proteínas desdobradas)

SUMÁRIO

1 INTRODUÇÃO.....	23
2 REVISÃO DE LITERATURA.....	25
2.1 O pinhão-manso.....	25
2.1.1 Taxonomia, origem e aspectos gerais.....	25
2.1.2 Morfologia.....	26
2.2 O multipotencial de <i>J. curcas</i>	27
2.2.1 Biodiesel.....	27
2.2.2 Propriedades medicinais.....	28
2.2.3 Bio-pesticida.....	29
2.2.4 Recuperação de solos degradados.....	31
2.3 Estresse abiótico.....	32
2.3.1 Salinidade.....	35
2.4 Helicases.....	37
2.4.1 Upf1-like.....	41
2.4.2 Pif1-like.....	42
2.4.3 UvrD/REP.....	42
2.4.4 MCM.....	43
2.4.5 RuvB.....	43
2.4.6 DEAD-box.....	44
2.4.7 Swi2/Snf2.....	45
2.4.8 RecQ-like.....	45
2.4.9 RecG-like.....	46
2.5 Helicases em plantas e envolvimento na tolerância a estresses.....	46
2.5.1 RNA helicases.....	48
2.5.2 DNA helicases.....	49
2.6 Genômica.....	50
2.6.1 Bancos de dados biológicos.....	51
2.6.2 Controle de qualidade do sequenciamento.....	52
2.6.3 Montagem e avaliação da qualidade.....	53
2.6.4 Repeat masking.....	54
2.6.5 Predição de genes e anotação.....	55
2.6.6 Análises <i>downstream</i>	56

2.7 Genoma de <i>Jatropha curcas</i>.....	57
2.8 Transcriptômica.....	58
2.8.1 RNA-Seq.....	60
2.8.2 qPCR.....	64
3 OBJETIVOS.....	66
3.1 Objetivo geral.....	66
3.2 Objetivos específicos.....	66
4 ARTIGO 1.....	67
5 ARTIGO 2.....	108
6 CONSIDERAÇÕES FINAIS	150
REFERÊNCIAS.....	152
APÊNDICE A.....	171
APÊNDICE B	205
APÊNCIDE C.....	213

1 INTRODUÇÃO

O pinhão-manso (*Jatropha curcas* L.) é uma oleaginosa não comestível, capaz de se desenvolver em solos de áreas com diferentes condições e regimes de chuva. Suas sementes são ricas em óleo com qualidades adequadas para produção de biodiesel. Além desta empregabilidade, a planta também pode ser utilizada como adubo, na recuperação de solos erodidos ou contaminados, como bio-pesticida e no consumo animal, se devidamente processada. Apesar disso, estresses bióticos e abióticos afetam negativamente o potencial produtivo de *J. curcas*.

Estresses abióticos são as principais ameaças à sobrevivência de plantas cultivadas, ocasionando quedas na produtividade e podendo levar à morte das plantas. Os efeitos desses estresses são exacerbados pelas mudanças climáticas, que elevam o desafio de obter culturas cada vez mais eficientes em sobreviver e produzir nessas condições extremas.

A salinidade do solo é um dos principais estresses causadores de perdas para a agricultura, sendo principalmente associada à presença de Na^+ e Cl^- . Ao serem absorvidos, desencadeiam desequilíbrio iônico e osmótico, acarretando em perda de água e do turgor celular. Apesar de tolerante a seca, o pinhão-manso é sensível a salinidade. Esse fator é agravado em solos marginais, geralmente ocupados por plantas não destinadas a alimentação humana, como o pinhão-manso.

A manutenção da homeostase celular das plantas sob estresses envolve múltiplos mecanismos de detecção e sinalização celular. Quando submetidas a condições adversas, esses organismos recorrem a reprogramação transcricional direcionada a minimizar os danos decorrentes dos estresses nas estruturas celulares. Os mecanismos envolvidos na mitigação desses danos atuam para garantir a limpeza de espécies reativas de oxigênio (ROS), a manutenção da integridade de membranas plasmáticas, equilíbrio do transporte de água e íons dentre outros processos.

O dinamismo da resposta envolve tanto mecanismos constitutivos, quanto induzíveis. Os mecanismos celulares de respostas a estresses ambientais envolvem diferentes sinalizações e ativação de grupos gênicos específicos, tais como fatores de transcrição (ex.: AP2-ERF, bZIP, WRKY, NAC, MYB etc), proteínas quinase e fosfatase, aquaporinas, fitorreguladores dentre outros. Além desses grupos gênicos, as helicases também têm sido apontadas como parte das linhas de defesa das plantas.

Helicases são proteínas codificadas por procariotos e eucariotos que utilizam a energia da hidrólise de ATP para promover o desenrolamento de ácidos nucleicos (DNA ou RNA), e de estruturas proteína-ácido nucleico. As helicases atuam na replicação, vias de reparo, recombinação, transcrição, biogênese ribossomal, tradução dentre outras, portanto, interferem no desenvolvimento das plantas. A regulação de processos relacionados ao metabolismo de ácidos nucleicos é essencial para garantir a manutenção do equilíbrio celular durante exposição a estresses.

Logo, helicases podem ser usadas como ferramentas biotecnológicas importantes para os programas de melhoramento genético de plantas. Portanto, o objetivo do presente trabalho foi caracterizar estruturalmente as famílias gênicas de helicases codificadas pelo genoma de *J. curcas* e analisar como essas famílias respondem a estímulo de salinidade (150 mM de NaCl por três horas) no transcriptoma RNA-Seq obtido de raízes de dois acessos contrastantes a salinidade.

2 REVISÃO DE LITERATURA

2.1 O pinhão-mansô

2.1.1 Taxonomia, origem e aspectos gerais

O pinhão-mansô (*Jatropha curcas* L.) pertence à família Euphorbiaceae, subfamília Crotonoideae, tribo Jatropeae. O nome *Jatropha* deriva do grego *jatros* (doutor) e *trophe* (comida) e designa o gênero que compreende cerca de 180 espécies (Dehgan, 1984). Sua ocorrência se estende de países tropicais até subtropicais (Figura 1). A presença da espécie em diferentes continentes é, provavelmente, devida a dispersão por navios portugueses através das Ilhas Cabo Verde e Guiné Bissau (Heller, 1996). É uma planta de fácil propagação, curto período de gestação e com elevada habilidade de se desenvolver em regiões com diferentes regimes de chuva.

Figura 1 - Centro de origem e locais de ocorrência de *Jatropha curcas* no mundo.



Fonte: Adaptado de Laviola *et al.*, (2017).

O mais antigo relato de uso de *J. curcas* vem dos Olmecas, povo que viveu no México de 3500-5000 anos atrás e foram os primeiros a colonizar a Mesoamérica (Dias *et al.*, 2012). Esta, e outras linhas de evidência, como a ocorrência exclusiva de acessos não-tóxicos na península de Yucatán, sugerem o México como provável centro de origem da espécie (Achten *et al.*, 2010). *J. curcas* é considerada como semi-domesticada, a diversidade genética presente dentro das populações da espécie varia de moderada a baixa (Maghuly *et al.*, 2015).

A planta foi inicialmente identificada como promissora para produção de biodiesel devido ao óleo presente nas suas sementes e por seu baixo custo de processamento (Koh *et al.*, 2011). A torta é o maior bioproduto do processo de obtenção do biodiesel e apresenta boa quantidade de proteínas e aminoácidos essenciais (Gomes *et al.*, 2018). Porém, a presença de componentes tóxicos dificulta a utilização desse produto na alimentação animal, sem que haja desintoxicação. A toxicidade está principalmente associada a ácidos hidroxi-octadecenóicos, ao invés de ésteres de forbol como era pensado (Wang *et al.*, 2020).

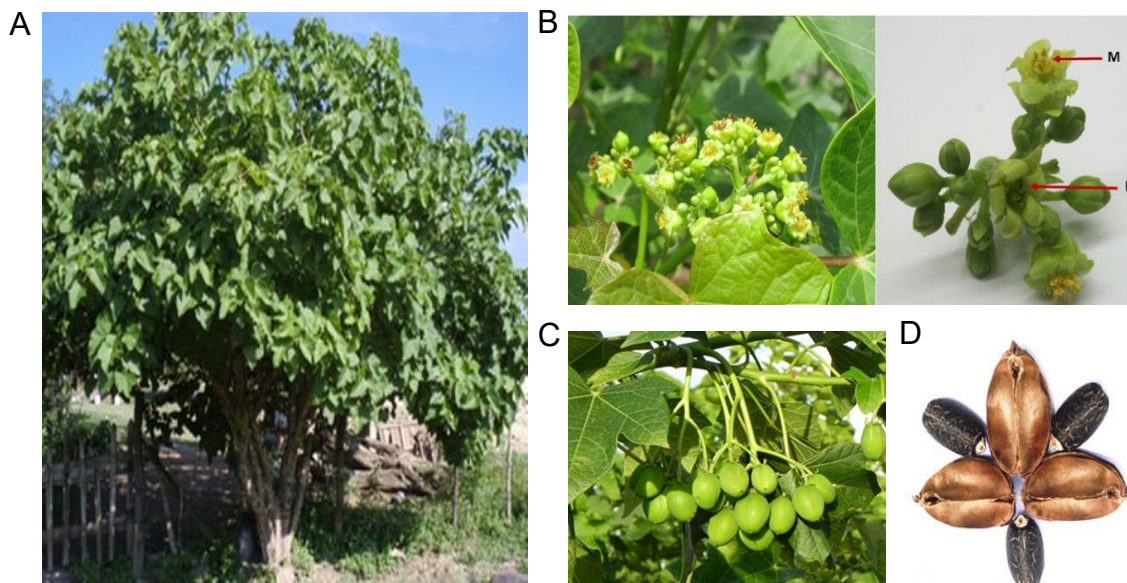
As sementes de *J. curcas* também abrigam inibidores de tripsina, lectinas e ácido fítico, que constituem fatores antinutricionais. Por sua adaptabilidade à solos não destinados a produção de alimento, é uma cultura com potencial para resolver questões relacionadas ao clima (Srivastava *et al.*, 2014). Ainda, comparada a outras matérias primas, *J. curcas* requer menores quantidades de insumos agrícolas, e não necessita de práticas agronômicas intensas (Ewunie *et al.*, 2021).

2.2.2 Morfologia

O pinhão-mansão é um arbusto perene que pode atingir 8-10 metros de altura (Carels, 2009) (Figura 2). Seu crescimento é articulado e apresenta ramos que contêm látex. As árvores são decíduas, com perda de suas folhas na estação seca. O caule é de aspecto liso, esverdeado e macio. Uma camada de cera reveste o tronco principal e ramos laterais (Alves *et al.*, 2010). Suas folhas possuem de cinco até sete lobos (Carels, 2009). O sistema radicular é formado por uma raiz principal que desenvolve quatro raízes laterais (Carels, 2009).

As plantas são monóicas com inflorescências protrândricas (Dehgan; Webster, 1984). As flores femininas (cerca de 20%) ocupam o centro dos ramos da inflorescência, enquanto as masculinas (cerca de 80%) aparecem em regiões axilares (Carels, 2009) (Figura 2). Após a polinização, que ocorre frequentemente por insetos da ordem Hymenoptera, a inflorescência produz frutos secos ovóides trilobulados que abrigam uma semente por lóculo (Carels, 2009) (Figura 2). A semente exibe coloração preta e no seu interior abriga o endosperma, de coloração branca e rico em óleo (Carels, 2009).

Figura 2 - Características morfológicas de pinhão-mansó (*J. curcas*). A. Planta adulta; B. Inflorescência; F [flor feminina]; M [flor masculina]; C. Fruto maduro; D. Fruto, em deiscência



Fonte: Adaptado de Alves *et al.*, (2010) e Montes *et al.*, (2016).

2.3 O multipotencial de *J. curcas*

2.3.1 Biodiesel

O aumento da população humana está cada vez mais pressionando a demanda por alimento e biocombustíveis. O maior desafio de conservação da natureza (promover agricultura em áreas marginais e uso eficiente de água e do solo), está intimamente relacionado com a produção de biocombustíveis. A utilização de culturas não-alimentares, como *J. curcas*, constitui uma alternativa sustentável para a produção de biodiesel de segunda geração (Kamel *et al.*, 2017).

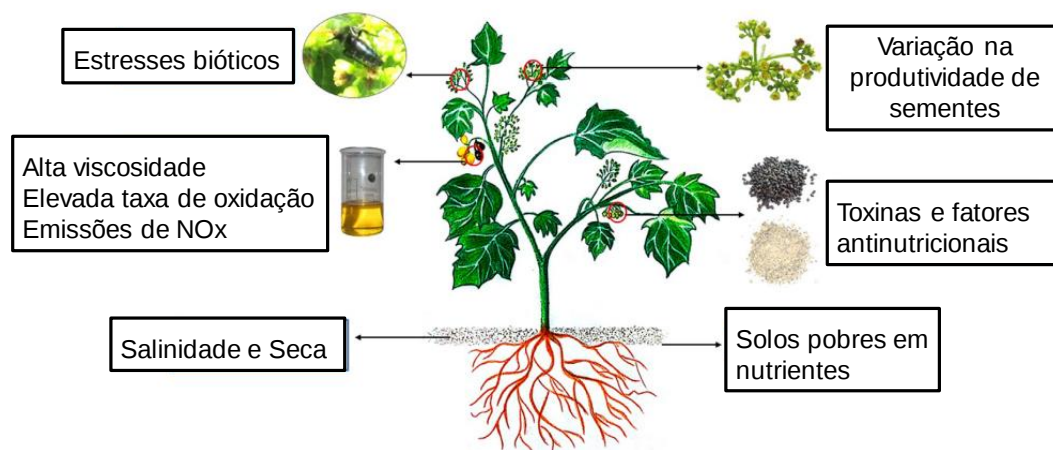
As sementes de *J. curcas* são conhecidas por sua elevada produção de óleo (783 kg/hectare) (Yang *et al.*, 2010). Em termos de composição, o óleo de *J. curcas* contém principalmente os ácidos oleico, linoleico, esteárico e palmítico que podem ser convertidos a ésteres de metil e então produzir biodiesel (Aransiola *et al.*, 2014). Além desses componentes, o óleo de *J. curcas* apresenta densidade, viscosidade, conteúdo de água, número de cetanos, estabilidade oxidativa e valor calorífico adequados (Kumar; Das, 2018). Logo, o potencial do óleo de *J. curcas* para produção de biodiesel é apontado por muitos estudos (Prabhu *et al.*, 2018; Gutiérrez-López *et al.*, 2022).

Apesar das muitas vantagens apresentadas pelo óleo de *J. curcas* para produção de biodiesel, o interesse e os investimentos vêm diminuindo por fatores

ambientais, sociais, econômicos e tecnológicos (Ewunie *et al.*, 2021). Os problemas incluem baixa produção de sementes, dificuldades de coleta, estresses, solo degradado dentre outros (Ewunie *et al.*, 2021).

Portanto, são necessários estudos futuros que visem desenvolvimento e otimização de tecnologias para modernizar a coleta e processamento de sementes, extração do óleo e sua conversão para biodiesel (Ewunie *et al.*, 2021). Além disso, para o cultivo economicamente viável são necessários mais estudos sobre requerimentos de água e nutrientes, bem como, suscetibilidade a doenças, pois a produção de sementes e qualidade do óleo são dependentes de fatores bióticos e abióticos (Soto *et al.*, 2018). A figura 3 resume os principais desafios da produção de biodiesel a partir de *J. curcas*.

Figura 3 - Principais limitações de *J. curcas* como cultura de biodiesel



Fonte: Adaptado de Mazumdar *et al.*, (2018).

2.3.2 Propriedades medicinais

As propriedades medicinais de *J. curcas* são conhecidas desde muito tempo. O atributo medicinal está relacionado a presença de alcaloides, taninos, flavonoides, esteroides, glicosídeos, saponinas, cumarinas e compostos fenólicos (Rahu *et al.*, 2021). Alguns metabólitos e peptídeos são únicos de *J. curcas*, como por exemplo, os diterpedóides Jatropholol e Jatropholonas A e B (Chen *et al.*, 1988), os peptídeos Curcaciclina A e B (Berg *et al.*, 1995; Auvin *et al.*, 1997) dentre outros.

Sendo assim, extratos obtidos a partir de *J. curcas* são empregados para tratar diarreias, cólicas, reumatismo, inflamações, micoses, eczemas, úlceras, queimaduras, parasitas intestinais, sangramento de gengivas, infecções bucais, e até filariose e malária (Upadhyay *et al.*, 2007; Prasad *et al.*, 2012). Além do conhecimento

tradicional, a composição química de *J. curcas* vem sendo pesquisada com o objetivo de investigar melhor suas propriedades (Tabela 1).

Tabela 1 - Propriedades medicinais associadas a diferentes partes vegetativas de pinhão-mansão.

Propriedade	Patógeno/doença	Extrato	Referência
Antifúngica	<i>Candida Albicans</i> e <i>C. Tropicalis</i>	Folha, caule e seiva	(Doughari; Abraham, 2021)
Antibacteriana e antioxidante	<i>Escherichia coli</i> , <i>Klebsiella pneumoniae</i> , <i>Staphylococcus aureus</i> , <i>Micrococcus luteus</i>	Óleo essencial	(Babahmad et al., 2018)
Antifúngica e antibacteriana	<i>S. aureus</i> , <i>Neisseria gonorrhoea</i> , <i>Pseudomonas aeruginosa</i> , <i>Escherichia coli</i> , <i>C. albicans</i> , e <i>Aspergillus flavus</i>	Raiz e látex	(Arekemase et al., 2011)
Antibacteriana	<i>K. pneumoniae</i> , <i>E. coli</i>	Caule e raiz	(Rahu et al., 2021)
Antibacteriana	<i>S. epidermidis</i>	Folha	(Ramadhana et al., 2023)
Antibacteriana	<i>S. aureus</i>	Folha e caule	(Oyetayo; Temenu 2023)
Antihiperglicêmico e Anti-hipertensivo	Diabetes e Hipertensão	Folha e caule	(Rodríguez-Guarcía et al., 2019); (Asuk et al., 2021)
Antiviral	HIV-1 e SARS-CoV-2	Semente	(Ferrão et al., 2022)
Antiplasmodial	<i>Plasmodium berghei</i>	Folha	(Airaodion; Ogbuagu 2020); (Akande et al., 2022)
Anti-inflamatória	Células da glia (neurodegeneração)	Folha	(Bastos et al., 2021)
Antitumoral	Hepatocarcinoma	Látex e semente	(Ahmed et al., 2020)
Antitumoral	Câncer do endométrio	Caule	(An et al., 2021)
Antifúngica	<i>Candida krusei</i>	Peptídeo sintético	(Souza et al., 2022)
Anti-hipertensiva	Hipertensão	Semente	(Segura-Campos et al., 2013);
Anti-tripanosomal	<i>Trypanosoma brucei-brucei</i>	Folha	(Khadija et al., 2022)

2.3.3 Bio-pesticida

O controle de doenças e pragas é frequentemente realizado através de inseticidas sintéticos. Essa prática colabora para a seleção de populações resistentes aos pesticidas, tornando o controle mais difícil e menos eficiente (Muniz et al., 2020).

Todas as partes de *J. curcas* são tóxicas e seu grau de toxicidade varia de acordo com o extrato, sua composição química, o ambiente e o modo de aplicação (Muniz et al., 2020). Extratos obtidos de *J. curcas* são fontes de metabólitos secundários com utilidade em muitas áreas, inclusive produção de bio-pesticidas com grande potencial para controle de diversos patógenos e pragas vegetais (Tabela 2).

Tabela 2 - Propriedades de pinhão-mansão (*J. curcas*) como bio-pesticida contra diferentes patógenos e pragas vegetais.

Patógeno/praga	Extrato	Referência
<i>Microtermes obesi</i> (cupim-do-trigo)	Raiz	(Sharad et al., 2013)
<i>Stephanitis pyrioides</i> (percevejo-azaleia)	Caule e raiz	(Sampson et al., 2023)
<i>Ceratitis capitata</i> (mosca-do-mediterrâneo)	Folha	(Silva et al., 2015)
<i>Cercospora coffeicola</i> (cercosporiose-do-café) e <i>Hemileia vastatrix</i> (ferrugem-do-cafeeiro)	Folha e casca	(Zaidan et al., 2021)
<i>Aphids crassivora</i> (pulgão-do-feijão-caupi), <i>Maruca testulalis</i> (broca de-vagem) e <i>Megalurothrips sjostedti</i> (tripes das flores de feijão-caupi)	Óleo	(Ahuchaogu et al., 2014)
<i>Meloidogyne incognita</i> (nematóide das galhas do quiabo)	Folha, raiz e semente	(Ogwudire et al., 2022)
<i>Bactrocera zonata</i> (mosca-do-pêssego) e <i>B. cucurbitae</i> (mosca-do-melão)	Casca e óleo	(Rampadarath et al., 2016)
<i>Odontotermes obesus</i>	Folha e caule	(Patel et al., 2022)
<i>Helicoverpa armigera</i> e <i>Plutella xylostella</i> (traça-das-crucíferas)	Casca de semente	(Ingle et al., 2017)
<i>Tetranychus urticae</i> (ácaro rajado)	Folha	(Premalatha et al., 2018)
<i>Callosobruchus maculatus</i> (caruncho-do-feijão)	Folha e semente	(Opuba et al., 2018); (Ifeanyieze et al., 2022)
<i>Phaeosariopsis personata</i> (doença da mancha tardia do amendoim)	Folha	(Francis, 2021)
<i>Acanthoscelides obtectus</i> (gorgulho-do-feijão)	Folha	(Tshimenga et al., 2018)
<i>Colletotrichum gloeosporioides</i> (antracnose)	Folha	(Kwodaga et al., 2019)

<i>Sclerotium cepivorum</i> (podridão branca da cebola)	Sementes	(Abd-Elbaky; Gharib 2021)
<i>Alternaria solani</i> (ferrugem precoce do tomate)	Óleo	(Dhaka <i>et al.</i> , 2022)
<i>Hypothenemus hampei</i> (broca-do-café)	Folha e caule	(Ohoueu <i>et al.</i> , 2023)
<i>Agrotis ipsilon</i> (lagarta-da-batata)	Óleo	(Bhattacharyya <i>et al.</i> , 2022)
<i>Spodoptera frugiperda</i> (lagarta-do-cartucho)	Óleo	(Mumpuni <i>et al.</i> , 2022)
<i>Macrotermes bellicosus</i>	Óleo	(Bessike <i>et al.</i> , 2023)

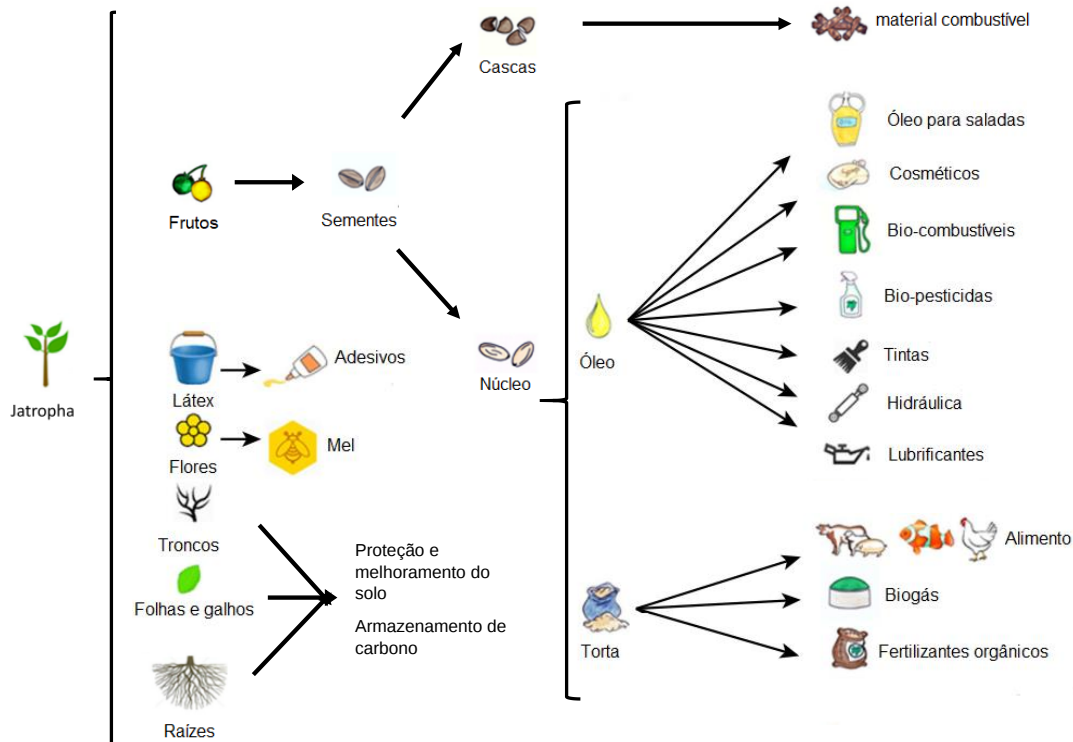
2.3.4 Recuperação de solos degradados

Sendo *J. curcas* uma pequena árvore, seu impacto positivo na qualidade de solos também é relatado. As raízes apresentam um ótimo desenvolvimento vertical, bem como, ancoramento profundo que são características promissoras para conter processos erosivos (Reubens *et al.*, 2011). O cultivo de plantas de *J. curcas* proporcionou melhoria na sua qualidade do solo por aumento de carbono, fósforo, potássio, macro e micro-agregados, da matéria orgânica e da resistência a erosão, além da regulação do pH de solos altamente ácidos e alcalinos (Baumert *et al.*, 2016; Mohammed *et al.*, 2021).

Alguns estudos mostraram que plantas de *J. curcas* também podem ser utilizadas para fitorremediação de solos contaminados. Plantas de *J. curcas* apresentam potencial para remediar solos contaminados com bauxita (Rahim *et al.*, 2019), metais pesados (Hg, Cu, Zn, Cr, Cd, Pb, Ni e As) (Awotedu *et al.*, 2019; Borah *et al.*, 2023; Mensah *et al.*, 2023) e derivados de petróleo (Nero; Nero, 2021).

Portanto, *J. curcas* apresenta potencial como fitorremediadora por suas características como: crescimento rápido elevada biomassa, habilidade de acumular metais pesados, ampla distribuição geográfica e toxicidade contra herbívoros (Mitra *et al.*, 2021). No entanto, são necessários mais estudos visando determinar a quantidade ótima de plantas empregadas e a duração do processo de fitorremediação, para assim melhorar os seus benefícios (Nero; Nero, 2021).

Além dos benefícios já citados, as partes de *J. curcas* também podem ser empregadas em outras finalidades. Destacam-se a produção de cosméticos, lubrificantes, tinta, óleo para cozinha, alimentação animal, mel, adesivos, biogás e fertilizantes orgânicos (Montes *et al.*, 2016). A figura 4 resume os principais produtos e potenciais aplicações de *J. curcas*.

Figura 4 – Aplicabilidades de *Jatropha curcas*

Fonte: Adaptado de Montes *et al.*, (2016).

2.3 Estresse abiótico

O crescimento e desenvolvimento de plantas são amplamente afetados por estresses abióticos, que representam um desafio principalmente considerando os impactos das mudanças climáticas (Kopecká *et al.*, 2023). Esses fatores restringem a produção agrícola, ameaçando a segurança alimentar global. Também, a população humana está estimada para crescer 2 bilhões nos próximos 30 anos e alcançar 9.7 bilhões em 2050, o que irá aumentar a pressão sobre a agricultura, já afetada pelo impacto das mudanças climáticas (ONU, 2023).

Devido a seu modo de vida sésstil, as plantas são continuamente expostas a fatores abióticos ou agentes estressores. Um estresse é definido como a qualquer fator que limite o desenvolvimento normal das plantas, que ao longo da sua evolução, desenvolveram mecanismos que as permitem sobreviver quando submetidas a essas condições. Esses fatores incluem: níveis extremos de luz (alta ou baixa luminosidade); temperatura; radiação; salinidade do solo; seca e inundação; metais pesados; excesso ou deficiência de nutrientes no solo; poluentes gasosos; fatores mecânicos, dentre outros (Kopecká *et al.*, 2023).

As mudanças climáticas causam maior frequência e intensidade desses estresses, além de conduzir a combinação desses fatores (Rivero *et al.*, 2022). Para mitigar os efeitos negativos dos estresses nas células, as plantas recorrem a três categorias gênicas: (1) genes que codificam proteínas regulatórias (ex.: fatores de transcrição, proteínas quinases e fosfatases); (2) genes associados ao transporte de água e íons (ex.: aquaporinas); (3) genes relacionados a proteção de membrana e de proteínas tais como chaperonas, *heat shock*, enzimas de detoxificação e osmoprotetores (Hossain *et al.*, 2016).

Uma das principais vias de resposta das plantas a estresses abióticos diz respeito a produção de espécies reativas de oxigênio (ROS - *Reactive Oxygen Species*). As maiores vias de produção de espécies reativas de oxigênio são as NADPH oxidases presentes nas membranas plasmáticas, as mitocôndrias, cloroplastos e peroxissomos, compartimentos celulares que apresentam forte fluxo de elétrons. As espécies reativas de oxigênio são: peróxido de hidrogênio (H_2O_2), radical superóxido ($O_2^{\bullet-}$), radical hidroxil ($OH^{\bullet}$) e oxigênio singlete (1O_2) (Sachdev *et al.*, 2021; Mansoor *et al.*, 2022).

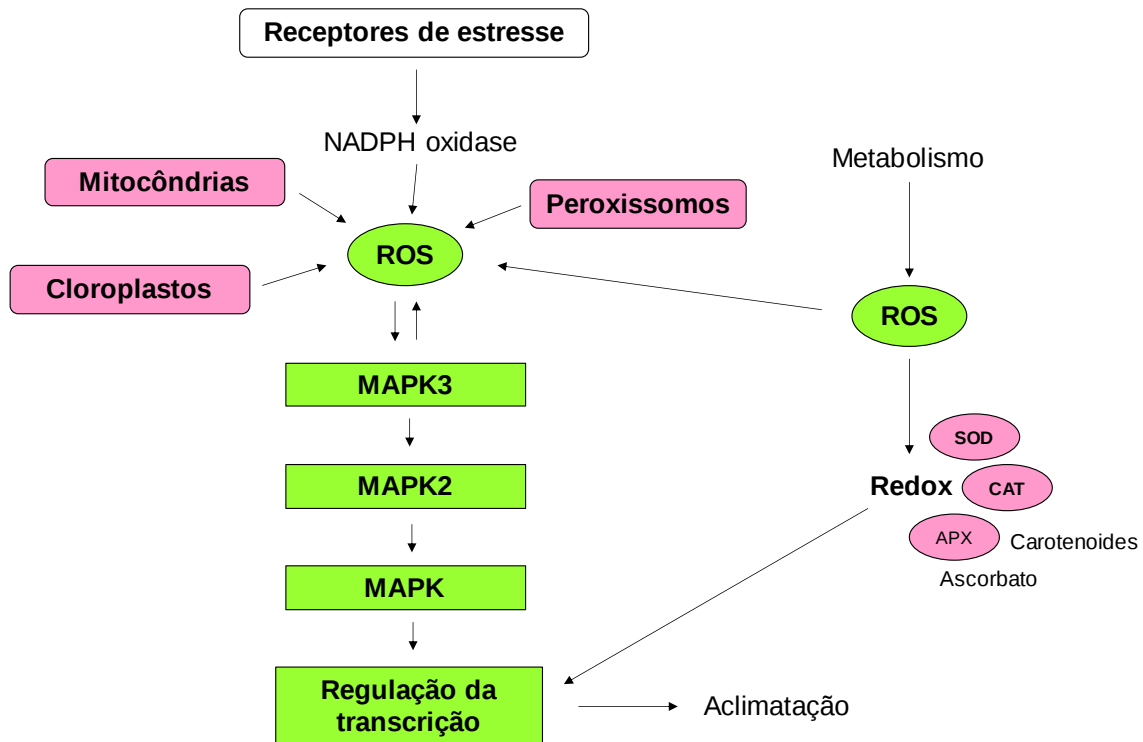
O cloroplasto é a fonte primária de ROS nas células. Condições adversas causam o fechamento dos estômatos e, conseqüentemente, reduzem a condutância estomatal, diminuindo a assimilação do CO_2 e resultando na formação de trigêmeos de clorofila ($^3Chl^*$). Essas moléculas reagem com o oxigênio molecular para produzir oxigênio singlete, desencadeando foto-oxidação. O estresse foto-oxidativo conduz a peroxidação lipídica, que danifica a membrana e afeta o foto-sistema II, acarretando morte celular. Danos em proteínas e no DNA também estão entre os impactos negativos de ROS (Sachdev *et al.*, 2021; Mansoor *et al.*, 2022).

A regulação da produção de ROS é feita por vias *scavenging* ou redox, rotas de defesa que abrangem antioxidantes enzimáticos e não-enzimáticos, e provêm proteção contra os efeitos negativos de ROS. Os principais antioxidantes enzimáticos são catalase (CAT), peroxidase (POX), superóxido dismutase (SOD), ascorbato peroxidase (APX) e glutathione redutase (GR). Dentre os não-enzimáticos estão os ácidos ascórbicos, alcaloides, flavonoides, carotenoides, α -tocoferol etc (Sachdev *et al.*, 2021; Mansoor *et al.*, 2022).

Mas além dos seus efeitos negativos, ROS também atuam como moléculas de transdução de sinal, e controlam diversas vias de aclimação da planta durante condições de estresse, principalmente quando se trata de H_2O_2 . Basicamente, elas

transportam o sinal do núcleo através de reações redox utilizando a via MAPK (proteína quinase mitógeno ativada) que resulta na ativação de uma série de fatores de transcrição estresse-responsivos (Devireddy *et al.*, 2021; Mansoor *et al.*, 2022) (Figura 5).

Figura 5 - Representação esquemática da cascata de sinalização das espécies reativas de oxigênio (ROS), ativação da via MAPK e homeostase redox durante estresses abióticos



Fonte: Adaptado de Mansoor *et al.*, (2022).

A adaptação das plantas a estresses é também associada a produção e acumulação de uma variedade de compostos orgânicos altamente solúveis. Esses compostos são conhecidos como osmoprotetores ou osmólitos, são de baixo peso molecular, eletricamente neutros e não-tóxicos. Ao contrário de compostos inorgânicos, os osmoprotetores mesmo em altas concentrações não apresentam toxicidade ao metabolismo das células (Mishra *et al.*, 2022; Mehta; Vyas, 2023).

Os osmoprotetores são classificados em três grupos: aminoácidos (ex.: prolina, glicina); betaínas (ex.: glicina betaína); álcoois de açúcar (ex.: Inositol, rafinose) e compostos quaternários de amônia (ex.: prolina betaína) (Mishra *et al.*, 2022). A principal função dos osmoprotetores é o ajuste osmótico das células, mantendo o turgor celular e atenuando a toxicidade dos íons. Além disso, também são

responsáveis pela proteção de estruturas celulares, de enzimas, da maquinaria fotossintética e pela limpeza de ROS (Mishra *et al.*, 2022).

Os fitorreguladores também exercem papel crucial na regulação de respostas de plantas a estresses abióticos. Dentre os principais fitorreguladores: auxinas, citocininas, giberelinas, ácido abscísico, ácido salicílico, etileno, jasmonato, brassinosteróides e strigolactonas (Zahid *et al.*, 2023). Desses, o ABA é um integrante central durante as respostas, sobretudo quando se trata de estresse osmótico. A síntese de ABA desencadeia a expressão de genes responsivos ao ABA e o fechamento estomatal, reduzindo, portanto, a perda de água via transpiração (Yamaguchi-Shinozaki; Shinozaki, 2006).

2.3.1 Salinidade

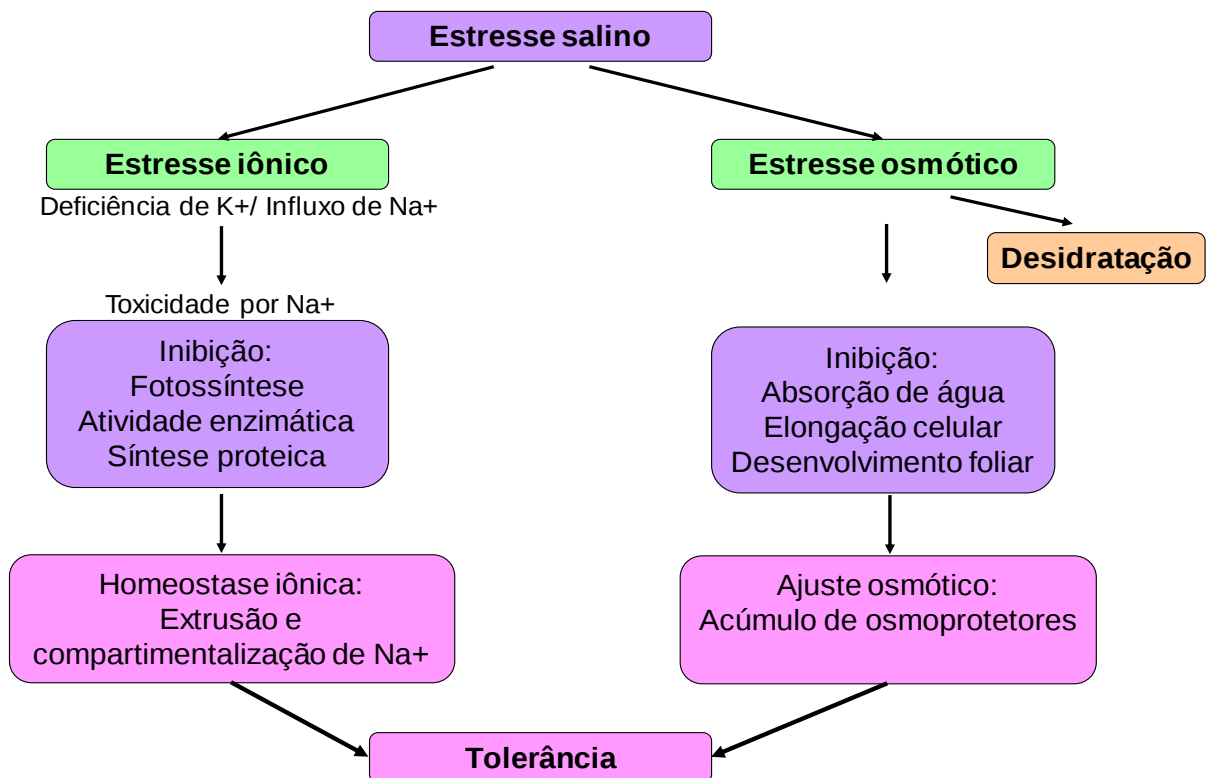
A salinidade afeta cerca de 20-33% das áreas cultivadas no mundo e seus efeitos estão previstos para alcançar 50% das áreas até 2050 (Machado; Serralheiro, 2017). A salinidade pode ser dividida em dois tipos: (1) A salinidade primária (ou natural), que ocorre através de processos geológicos, pedológicos ou hidrológicos, em áreas áridas ou semiáridas onde a evapotranspiração exerce papel considerável na pedogênese desses solos; (2) A salinidade secundária, que é decorrente de fatores humanos, principalmente devido a utilização de práticas de irrigação não adequadas (Safdar *et al.*, 2019).

A salinidade é associada com a presença de sais solúveis em água tais como Na^+ , K^+ , Cl^- , SO_4^{2-} . Plantas não reconhecem Na^+ como nutriente, sendo o principal íon relacionado com a salinidade do solo. Os efeitos da salinidade abrangem danos estruturais ao aparelho fotossintético, a membranas plasmáticas, diminuição da absorção de nutrientes e do conteúdo de clorofila, que resultam em inibição do crescimento e consequente queda na produtividade (Hao *et al.*, 2021).

O excesso de Na^+ frequentemente causa deficiência de K^+ , pois K^+ participa de uma miríade de processos fisiológicos e o Na^+ compete com seus sítios de absorção nas raízes. Além disso, Na^+ causa despolarização da membrana plasmática, conduzindo ao efluxo de K^+ ativado pelo acúmulo de H_2O_2 . Logo, sob salinidade, a deficiência de K^+ é severa. Portanto, manter o balanço citosólico de K^+/Na^+ é essencial para tolerância a salinidade (Hao *et al.*, 2021. Dos Santos *et al.*, 2022).

As respostas das plantas a salinidade podem ser agrupadas em duas etapas: (1) a primeira é considerada como rápida, é chamada de fase osmótica ou estresse osmótico, ocorre assim que os sais alcançam as raízes e diminuem a absorção de água pelas raízes; (2) a segunda é a fase iônica ou estresse iônico, descrita como lenta, consiste na absorção e sinalização de íons tóxicos, principalmente Na^+ (Dos Santos *et al.*, 2022) (Figura 6). Para lidar com seus efeitos negativos, plantas empregam estratégias de acumulação de osmólitos compatíveis, exclusão e compartimentalização de íons, limpeza de ROS e ativação de genes estresse-responsivos (Hao *et al.*, 2021) (Figura 6).

Figura 6 - Representação esquemática dos efeitos da salinidade nas plantas



Fonte: Adaptado de Mehta; Vyas (2022).

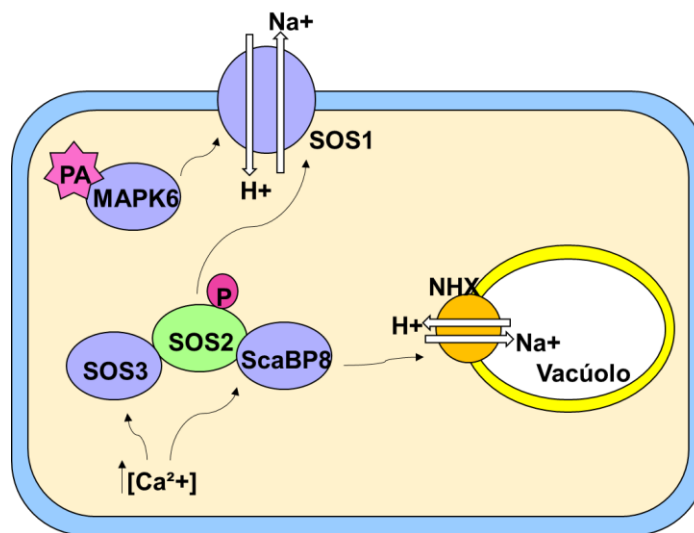
Ajuste osmótico - As SnRK2s (*sucrose nonfermenting-1-related protein kinases* 2) são ativadas pelo aumento do ABA e são componentes centrais da sua via de sinalização (Soma *et al.*, 2017). Raf-like quinases MAPKKK (*Mitogen-Activated Protein Kinase Kinase Kinase*) ativam a SnRK2 da subclasse I (Saruhashi *et al.*, 2015). Então, SnRK2 fosforila e ativa fatores de transcrição AREB/ABF para induzir a expressão de BAM1/AMY3 que irá degradar amido em açúcares, conduzindo a

acumulação de osmólitos compatíveis que estabilizam o potencial osmótico das células submetidas a estresse (Thalmann *et al.*, 2016).

Homeostase iônica – A manutenção do equilíbrio iônico é diretamente associada aos genes da via SOS (*Salt Overly Sensitive*) (Figura 7). Sob condições de estresse, o excesso de Na^+ conduz ao aumento citosólico de Ca^{2+} . As proteínas SOS3 e ScaBP8/CBL10 decodificam o sinal de Ca^{2+} e o traduzem para SOS2 (proteína quinase serina/treonina). SOS2 fosforila e ativa SOS1, um transportador de membrana Na^+/H^+ que expulsa Na^+ de volta ao espaço extracelular (Yang; Guo 2018).

O sequestro vacuolar de Na^+ também é uma importante estratégia para lidar com os efeitos da salinidade. SOS3/SOS2 atua na regulação positiva de NHX (trocaador vacuolar de Na^+/H^+), bem como da atividade das adenosinas trifosfatases de H^+ (PM H^+ ATPase). O trocaador NHX coordena o transporte de Na^+ do citoplasma para o vacúolo, sendo assim um importante ator na manutenção da homeostase K^+/Na^+ (Yang; Guo, 2018).

Figura 7 - Representação esquemática de componentes da via SOS (*Salt Overly Sensitive*) que atuam na mantêm a homeostase iônica sob estresse salino



Fonte: Adaptado de Yang; Guo (2018).

2.4 Helicases

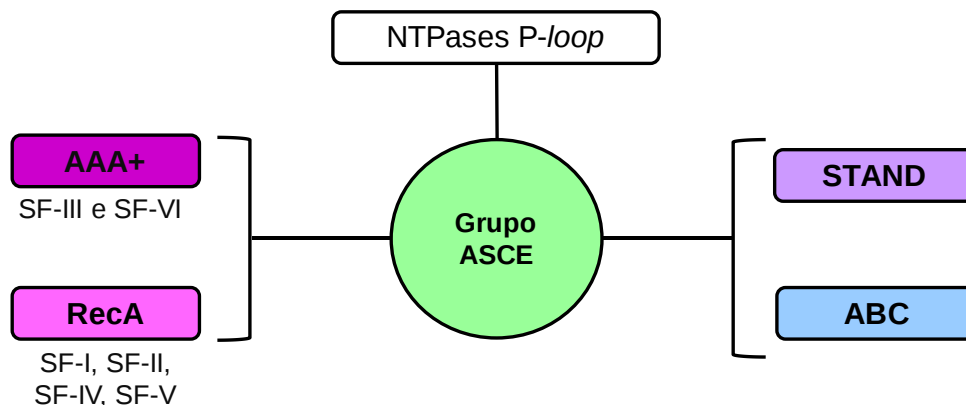
Helicases são enzimas especializadas em utilizar a energia livre da hidrólise de nucleosídeos trifosfato (NTP) para catalisar a separação de ácidos nucleicos. Essa atividade pode ser direcionada ao desenrolamento de DNA, RNA, duplex DNA-RNA e

proteínas ligadas a ácidos nucleicos. Além disso, as helicases estão envolvidas no desenrolamento de estruturas mais complexas como junções de Holliday, G-quadruplex e triplexes (Huselid; Bunting, 2020).

As helicases pertencem a superclasse de NTPases *P-loop*, um grupo muito diverso que compreende proteínas com aproximadamente 75 atividades enzimáticas distintas (Longo *et al.*, 2020). Essas proteínas são caracterizadas pela presença de uma dobra $\alpha\beta\alpha$ contendo dois elementos de sequência conservados, denominados Walker A e Walker B (Walker *et al.*, 1982).

Dentro das NTPases, as helicases pertencem ao grupo denominado ASCE (*Additional Strand Catalytic Glutamate*) (Leipe *et al.*, 2003). Da mesma forma, o grupo ASCE pode ser adicionalmente subdividido em classes denominadas ABC (*ATP-binding cassette*), STAND (*Signal Transduction ATPases with Numerous Domains*), AAA+ (*ATPases Associated with diverse cellular Activities*) e RecA (Leipe *et al.*, 2003) (Figura 8).

Figura 8 - Esquema das divisões da superclasse NTPases *P-loop*, com descrição do grupo ASCE (*Additional Strand Catalytic Glutamate*) que inclui as superfamílias das helicases



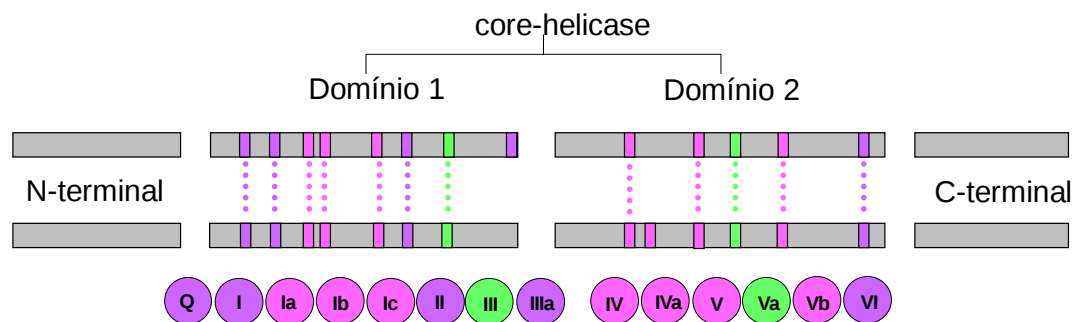
Fonte: Silva (2024).

A primeira helicase foi descrita em *E. coli* (Abdel-Monem *et al.*, 1976). Desde então, essas enzimas têm sido extensivamente pesquisadas. Existem três propriedades compartilhadas por quase todas as helicases, que são: (1) ligação a ácidos nucleicos, (2) ligação e hidrólise de NTP/dNTP e (3) desenrolamento e rearranjo de ácidos nucleicos (Passricha *et al.*, 2019).

No geral, o *core* das helicases é constituído de até 14 motivos conservados. Esses motivos são denominados: Q, I, Ia, Ib, Ic, II, III, IIIa, IV, IVa, V, Va, Vb, e VI e estão presentes ao longo de dois domínios (Fairman-Williams *et al.*, 2010) (Figura 9).

Esses motivos estão diretamente relacionados com as atividades bioquímicas dessas enzimas, que são basicamente: ligação e hidrólise de ATP, coordenação entre sítios de ácidos nucleicos e sítios de ligação a NTP e ligação a ácidos nucleicos.

Figura 9 - Representação esquemática dos motivos conservados encontrados em helicases. Os motivos estão coloridos de acordo com sua atividade bioquímica principal: em roxo, ligação e hidrólise de ATP; em verde, coordenação entre sítios de ácidos nucleicos e sítios NTP; em rosa, ligação a ácido nucleico



Fonte: Adaptado de Fairman-Williams *et al.*, (2010).

Com base na sequência, estrutura e mecanismos, as helicases são classificadas em seis superfamílias (SF-I – SF-VI) (Figura 10). As superfamílias SF-I e SF-II são monoméricas, enquanto que as demais (SF-III – SF-VI) são hexaméricas formadoras de anéis (Singleton *et al.*, 2007).

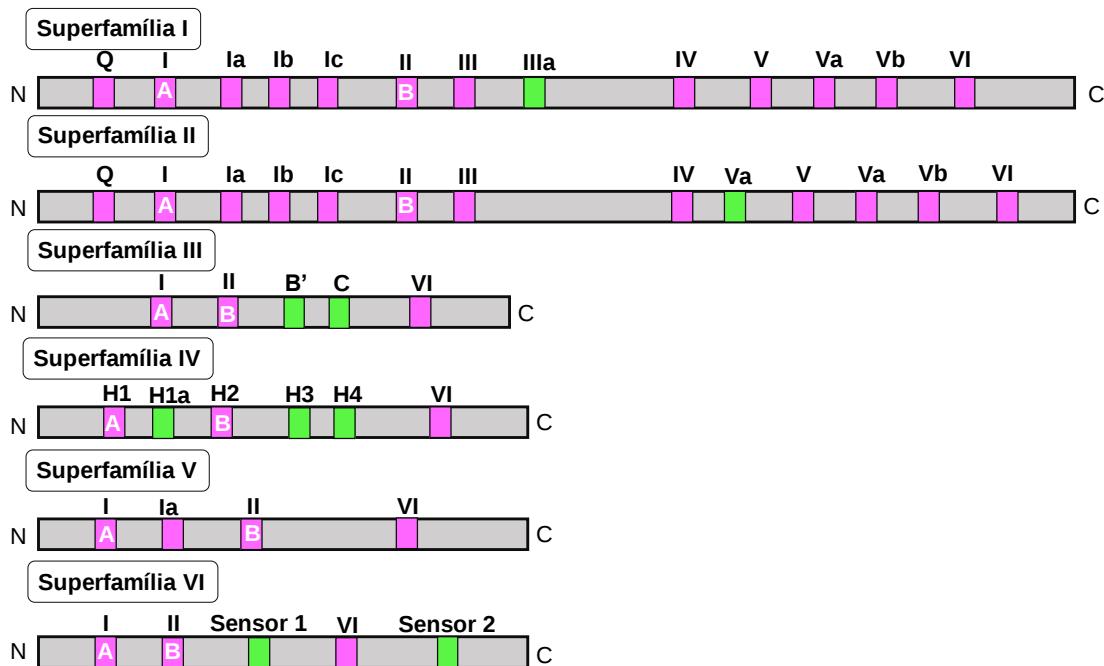
Helicases hexaméricas geralmente possuem múltiplos sítios de ligação a NTP e a presença de *loops* que se projetam em direção ao centro do anel (Singleton *et al.*, 2007). As superfamílias SF-I e SF-II compreendem a maior parte das DNA e RNA helicases, sendo a SF-II a maior superfamília. Cada superfamília é definida por um conjunto de motivos conservados, com alguns sendo compartilhados e outros únicos (Figura 10).

Os membros da superfamília SF-I possuem direcionalidade (ou polaridade) 3' - 5' ou 5' - 3' e podem ser divididos em SF-1A e SF-1B. Essas proteínas exibem até treze motivos conservados (Q, I, Ia, Ib, Ic, II, III, IIIa, IV, V, Va, Vb e VI) (Singleton *et al.*, 2007; Fairman-Williams *et al.*, 2010; Passricha *et al.*, 2019). A superfamília SF-II exibe preferencialmente polaridade 3' – 5' e alguns membros atuam como remodeladores da cromatina (ex.: Swi/Snf2). Os membros podem apresentar até treze motivos (Q, I, Ia, Ib, Ic, II, III, IV, IVa, V, Va, Vb e VI) (Singleton *et al.*, 2007; Fairman-Williams *et al.*, 2010; Passricha *et al.*, 2019).

Os membros da superfamília SF-III estão presentes em vírus (DNA e RNA) e possuem polaridade 3' – 5'. Apresentam cinco motivos (A (I), B (II), B', C, e R (VI)). Motivos B' e C são únicos dessa superfamília (Singleton *et al.*, 2007; Passricha *et al.*, 2019). A superfamília SF-IV inclui helicases replicativas de bactérias e bacteriófagos, com polaridade 5' - 3'. Exibem seis motivos (A (H1), H1a, B (H2), H3, H4 e R). Os motivos H1 e H2 são equivalentes aos motivos *Walker A* e *Walker B* (Singleton *et al.*, 2007; Passricha *et al.*, 2019).

A superfamília V é representada pelos fatores de terminação da transcrição em bactérias e possuem polaridade 5' – 3'. É caracterizada pela presença de quatro motivos conservados I (A), Ia, II (B), e VI (R) (Singleton *et al.*, 2007; Passricha *et al.*, 2019). A superfamília VI é também chamada de AAA+ (*ATPases Associated with various cellular Activities*). É definida pela presença de cinco motivos I (A), II (B), S1, VI (R) e S2 (Singleton *et al.*, 2007; Passricha *et al.*, 2019).

Figura 10 - Descrição dos motivos conservados em cada superfamília de helicase. Motivos A, B e R são comuns e os motivos em verde são únicos a determinada superfamília



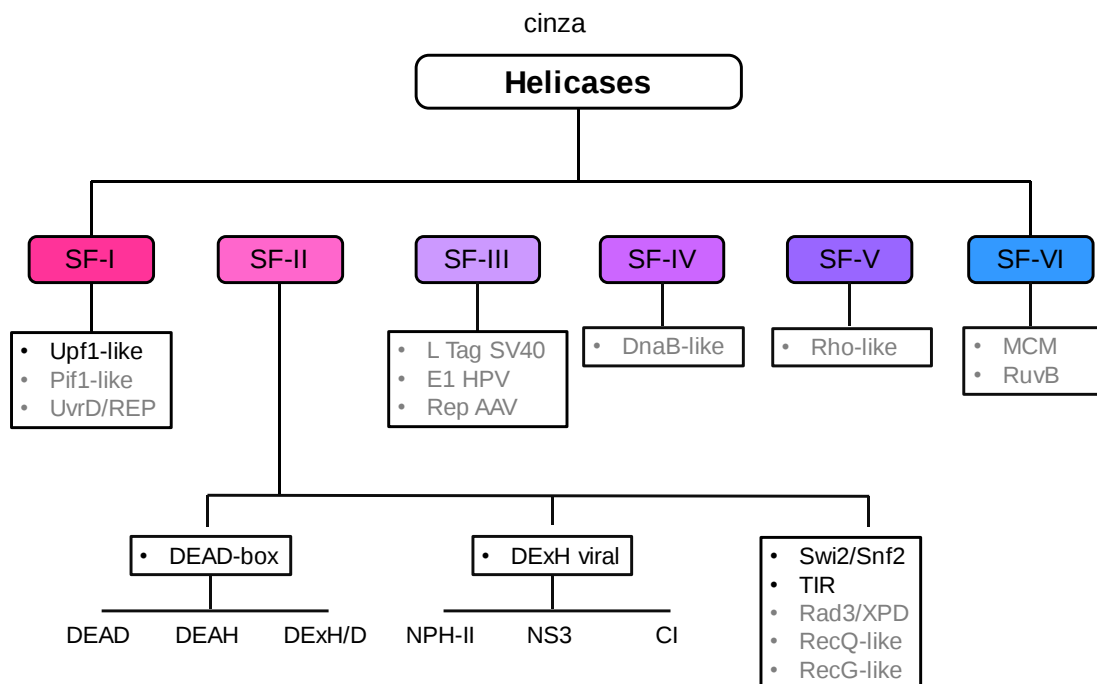
Fonte: Adaptado de Passricha *et al.*, (2019).

Com base em sua perspectiva estrutural, as superfamílias são adicionalmente subdividas em famílias (Figura 11). A superfamília SF-I compreende as famílias UvrD/Rep, Pif1-like e Upf1-like. A superfamília SF-II inclui as famílias DEAD-box

(subfamílias DEAD, DEAH e DExH/D), RecG-like, RecQ-like, Rad3/XPD, T1R, Swi2/Snf2, NS3/NPH-II e CI.

As superfamílias SF-III (SV40, EI HPV, Rep AAV), SF-IV (DnaB-like) e SF-V (Rho-like) incluem DNA helicases encontradas apenas em vírus e bactérias. A superfamília SF-VI, ou superfamília AAA+ (*ATPases Associated with various cellular Activities*), inclui as famílias de DNA helicase MCM (*mini chromosome maintenance*) e RuvB (Seraj *et al.*, 2018; Passricha *et al.*, 2019) (Figura 11).

Figura 11 - Representação da subdivisão das helicases em famílias de acordo com cada uma das seis superfamílias descritas. Famílias de RNA helicase estão exibidas em preto e DNA helicase em



Fonte: Adaptado de Seraj *et al.*, (2018).

2.4.1 Upf1-like

As helicases Upf1 (*Up-frameshift 1*) pertencem a superfamília SF-I, atuam como RNA e DNA helicases, e apresentam dois domínios clássicos RecA-like (Cheng *et al.*, 2007). São componentes centrais da via NMD (nonsense mediated mRNA decay), que envolve a montagem e desmontagem de fatores que medeiam a degradação de mRNA (Arciga-Reyes *et al.*, 2006), e também na regulação da atividade de telomerasas (Addinall *et al.*, 2011). A proteína Upf1 recruta Upf2 e Upf3 para formar o *core* do complexo NMD, que é conservado em eucariotos (Serin *et al.*,

2001). Os genes codificadores dessa helicase são apontados como essenciais para o desenvolvimento das plantas em processos tais como florescimento, germinação de sementes e senescência (Yoine *et al.*, 2006; Shi *et al.*, 2012; Gong *et al.*, 2019). Os genes Upf1 também são relacionados com a defesa vegetal (Shi *et al.*, 2012; May; Simon, 2021).

2.4.2 Pif1-like

A família Pif1-like foi primeiramente identificada em um *screening* genético para detecção de mutações envolvidas em alterações teloméricas (Schulz; Zakian, 1994). A família Pif1-like pertence a superfamília SF-I e inclui DNA helicases com uma região conservada de 400-500 aminoácidos que apresenta sete motivos conservados com regiões N e C terminais variando bastante (Bochman *et al.*, 2010). Os genes Pif1 possuem duas variantes, uma longa que é destinada ao núcleo e uma forma menor presente nas mitocôndrias (Schulz; Zakian, 1994). As funções dessas helicases variam entre os diferentes organismos, mas são tidas como essenciais para a manutenção dos genomas nuclear e mitocondrial, atuando no processamento de fragmentos de Okazaki (Pike *et al.*, 2009), regulação da atividade de telomerasas (Zhou *et al.*, 2000), na remoção de barreiras durante a replicação (Koc *et al.*, 2016), reparo de quebras (Wilson *et al.*, 2013) e manutenção da forquilha de rDNA (Ivessa *et al.*, 2000). A família Pif1-like é pequena, com apenas duas proteínas em humanos, leveduras e *E. coli* (Fairman-Williams *et al.*, 2010). O genoma de *Arabidopsis* codifica três proteínas Pif1-like (Bochman *et al.*, 2010). Em plantas, essas proteínas são pouco caracterizadas e não há relatos diretos de sua participação em vias de tolerância a estresses.

2.4.3 UvrD/REP

A família UvrD/REP abrange DNA helicases que pertencem a superfamília SF-I e são altamente similares a helicases PcrA de bactérias gram-positivas. Foi inicialmente caracterizada em *E. coli* e denominada DNA helicase II (Kumura; Sekiguchi, 1984). Helicases UvrD/REP possuem a estrutura clássica das helicases: dois domínios RecA-like, além de dois domínios de interação com o DNA (1B e 2B) e região C-terminal estendida (Manelyte *et al.*, 2009). Ao longo desses domínios, ocorrem onze motivos conservados (Tuteja *et al.*, 2013). Essas helicases fazem parte do complexo UvrABC, desenrolam o DNA durante o reparo por excisão de nucleotídeo

(NER) e o reparo de mismatch (MMR). Durante NER e MMR, UvrD é recrutada para deslocar bases danificadas e pares de bases com *mismatch*, dando prosseguimento as vias de reparo com as DNA polimerases e DNA ligases (Kunkel; Erie, 2005; Truglio *et al.*, 2006). Muito pouco é conhecido das helicases UvrD em plantas. Trata-se de uma família muito pequena, com apenas duas proteínas em humanos e leveduras, e quatro em *E. coli* (Fairman-Williams *et al.*, 2010). Os genomas de *Arabidopsis* e arroz codificam um gene cada (Tuteja *et al.*, 2013). Há apenas um relato da participação de UvrD na tolerância a seca em arroz, relacionado com o aumento do reparo do DNA e da atividade antioxidante (Zinati; Barati, 2018).

2.4.4 MCM

As proteínas MCM (*minichromosome maintenance*) foram primeiramente identificadas em mutantes de leveduras relacionados ao ciclo celular (Maine *et al.*, 1984). A família MCM inclui DNA helicases compreendidas na superfamília SF-VI, é formada por proteínas AAA+ que compartilham uma região conservada de aproximadamente 200 aminoácidos (Hanson; Whiteheart, 2005). Dentro da região conservada, mais precisamente no domínio AAA+, estão presentes os três motivos conservados característicos (Sarvajeet *et al.*, 2021). Essas proteínas formam um complexo de seis proteínas (MCM-2-7), que é conservado na maioria dos eucariotos (Fu *et al.*, 2011). O complexo MCM é essencial para a formação da forquilha de replicação. Brevemente, o complexo de pré-replicação (pré-RC) facilita o recrutamento do complexo MCM-2-7, este processo ativa a formação da forquilha durante a fase G1 (Tuteja *et al.*, 2011). Em plantas, até nove genes MCM (MCM-2-10) podem estar presentes. Os genes MCM-2-7 codificam as subunidades com atividade de helicase, e os demais (MCM8, MCM9 e MCM10) são co-reguladores de MCM-2-7 (Shultz *et al.*, 2007). Os genomas de *Arabidopsis*, *O. sativa*, *Brassica rapa* e *B. oleraceae* codificam nove, 10 e 12 proteínas MCM, respectivamente (Shammugan *et al.*, 2017; Sarvajeet *et al.*, 2021). Algumas proteínas do complexo MCM são apontadas como importantes reguladores de processos de desenvolvimento e também de resposta a estresse (Tuteja *et al.*, 2011; Dang *et al.*, 2011a; Dang *et al.*, 2011b; Herridge *et al.*, 2014).

2.4.5 RuvB

A família RuvB abrange DNA helicases incluídas na superfamília SF-VI. São proteínas AAA+ que apresentam um *core* contendo dois domínios abrangendo 200-250 resíduos de aminoácidos, esses domínios correspondem a um *core* $\alpha\beta\alpha$ e um domínio AAA+. No domínio AAA+ ocorrem os quatro motivos conservados característicos da família (Ammelburg *et al.*, 2006; Saifi *et al.*, 2019). Estão presentes de procariotos a eucariotos e são envolvidas em uma série de processos tais como: recombinação e reparo do DNA, transcrição, dobramento de proteínas, manutenção de telômeros, degradação de proteínas, progressão do ciclo celular, processamento de pré-rRNA dentre outros (Rotanova, 2006; Schumacher *et al.*, 2006; Nano; Houry, 2013). Há pouca informação sobre as helicases RuvB em plantas. O genoma de *O. sativa* possui dois genes RuvB caracterizados (*OsRuvBL1a* e *OsRuvBL1b*) (Saifi *et al.*, 2018). A RuvB *OIP30* de *Arabidopsis* foi detectada como substrato de uma CDPK (*calcium-dependent protein kinase*), proteínas envolvidas em vários processos, inclusive tolerância a estresses (Wang *et al.*, 2011). Há também alguns relatos da participação direta de genes RuvB em plantas expostas a estresses (Saifi *et al.*, 2018; Singh *et al.*, 2019; Saifi *et al.*, 2021).

2.4.6 DEAD-box

A família DEAD-box é constituída de RNA helicases pertencentes a superfamília SF-II e representa a maior de todas as famílias de helicases. Essas proteínas foram primeiramente identificadas em leveduras, com base em alinhamentos de homólogos do fator de iniciação da tradução eIF4A (Linder *et al.*, 1989). O *core* dessas proteínas abrange 350-400 aminoácidos e é formado por dois domínios RecA-like, ao longo dos quais estão presentes nove motivos conservados. As regiões N e C terminais são altamente variáveis em tamanho e composição de domínios (Cordin *et al.*, 2006; Fairman-Williams *et al.*, 2010). As proteínas DEAD-box atuam em todas as etapas que envolvem metabolismo da molécula de RNA, desde transcrição até a degradação de RNA (Rocak; Linder, 2004; Cordin *et al.*, 2006). Trata-se da família de helicases mais bem caracterizada em plantas, os genomas de *Arabidopsis*, *O. sativa*, *Z. mays*, *G. max*, *G. raymondii* e *S. lycopersicum* codificam 161, 149, 136, 213, 161 e 131 genes, respectivamente (Xu *et al.*, 2013; Chen *et al.*, 2014; Pandey *et al.*, 2019). Essas proteínas são conhecidas por estarem associadas a importantes vias de resposta a vários tipos de estresses e também por promover tolerância quando superexpressas (Sanan-Mishra *et al.*, 2005; Kant *et al.*, 2007; Li *et*

al., 2008; Nguyen *et al.*, 2018; Pandey *et al.*, 2019; Nawaz e Kang 2019; Lu *et al.*, 2020; Wang *et al.*, 2020; Wang *et al.*, 2022).

2.4.7 Swi2/Snf2

A família Swi2/Snf2 pertence a superfamília SF-II das helicases e inclui proteínas conhecidas principalmente como remodeladores da cromatina. A estrutura dessas proteínas inclui dois domínios RecA-like e cinco motivos conservados, além de até 14 blocos conservados entre esses motivos denominados de E-N (Flaus *et al.*, 2006). As regiões N e C terminais também abrigam domínios acessórios como PDH (*Plant homeodomain*), HIRAN (HIP116, Rad5p N-terminal), SANT dentre outros. As proteínas Swi2/Snf2 utilizam a energia da hidrólise do ATP para se mover ao longo do DNA, ao invés de desenrolá-lo, atuando assim como translocases. Essas proteínas alteram as interações DNA-histona conduzindo a mudanças na conformação, montagem e posição dos nucleossomos (Hopfner; Michaelis, 2007). Os genomas de *Arabidopsis* e *O. sativa* codificam 41 e 40 proteínas SNF2, respectivamente (Knizewski *et al.*, 2008; Hu *et al.*, 2013). Muitas dessas proteínas são apontadas como fatores essenciais em processos de desenvolvimento tais como senescência (Cho *et al.*, 2016), embriogênese (Sang *et al.*, 2012), formação de raízes (Yang *et al.*, 2015) e florescimento (Li *et al.*, 2015). Também estão diretamente relacionadas a regulação da metilação do DNA (Yang *et al.*, 2018). Além disso, também estão envolvidas nas respostas a estresse, como regulação da sinalização de ABA e na promoção de tolerância a salinidade e infecção fúngica, controlando o estado dos nucleossomos ao redor de sítios de início da transcrição de genes estresse responsivos (Han *et al.*, 2012; Yu *et al.*, 2018; Jakada *et al.*, 2019).

2.4.8 RecQ-like

As proteínas da família RecQ-like são DNA helicases conservadas em todos os domínios da vida, com papéis importantes na estabilidade do genoma. As helicases RecQ-like pertencem a superfamília SF-II, o *core* dessas proteínas possui de 300-450 aminoácidos é formado por três regiões conservadas, que são os domínios: helicase, RQC (*RecQ C-terminal domain*) e HRDC (*RNase D C-terminal domain*) (Vindigni; Hickson, 2009). Domínios RQC e HRDC são necessários para ligação e reconhecimento de estruturas de DNA complexas (Knoll; Puchta, 2011). As helicases RecQ-like são conhecidas por sua atividade no processamento de junções de

Holliday, que ocorrem durante a replicação, reparo do DNA ou recombinação, e as resolvem de maneira a evitar recombinação deletéria (Hong-Qing; Li, 2004; Hartung *et al.*, 2007; Kobbe *et al.*, 2008; Higgins *et al.*, 2011; Rohrig *et al.*, 2018). Não há muitas informações acerca dos genes *RecQ-like* em plantas, os genomas de *Arabidopsis* e *O. sativa* codificam sete membros da família (Hartung; Puchta, 2004). O nocaute de genes *RecQ*, bem como mutações, conduz a fenótipos de hiperrecombinação e letalidade (Hartung *et al.*, 2007).

2.4.9 RecG-like

A família *RecG-like* é constituída por DNA helicases que pertencem a superfamília SF-II. Como nas outras famílias, essas helicases são constituídas de dois domínios *RecA-like* no *core*, além de um domínio denominado “*wedge*”, associado com a especificidade de ligação a DNA seccionado e também para prevenir sua dissociação durante o processo de translocação (Briggs *et al.*, 2005; Fairman-Williams *et al.*, 2010). Essas proteínas atuam na migração de junções de *Holliday* (McGlyn; Lloyd, 1999) e na recuperação de forquilhas paradas (McGlyn; Lloyd, 2000). Em plantas, pouco se conhece sobre as helicases *RecG*. Até o momento, foram identificados e caracterizados um homólogo *RecG* nos genomas de *Arabidopsis* (Wallet *et al.*, 2015) e *Physcomitrella patens* (Odahara *et al.*, 2015). Essas proteínas foram apontadas como fatores chave para a manutenção da estabilidade dos genomas plastidial e mitocondrial. O nocaute desses genes causou defeitos de crescimento nas plantas e estruturas anormais presentes no cloroplasto e mitocôndria que conduziu a instabilidade das organelas (Wallet *et al.*, 2015; Odahara *et al.*, 2015). Não há relatos do envolvimento direto de genes *RecG* na resposta de plantas submetidas a estresses.

2.5 Helicases e seu envolvimento na tolerância a estresses

Helicases ocorrem no núcleo, cloroplasto e mitocôndria. Isto sugere o envolvimento dessas proteínas em vários aspectos do metabolismo das plantas, incluindo respostas a estresses (Tuteja *et al.*, 2001, Tuteja *et al.*, 2013). As helicases já foram descritas como componentes chave para tolerância a vários tipos de estresses em plantas. Dentre os genes helicases, *PDH45* (RNA helicase – DEAD-box) de ervilha é o mais bem caracterizado quanto a sua capacidade em promover tolerância a estresses em plantas, principalmente salinidade.

Sanan-Mishra *et al.*, (2005) foram os primeiros a mostrar que a superexpressão de *PDH45* em plantas de tabaco foi responsável por maior tolerância a salinidade. *PDH45* também induziu tolerância a salinidade e seca quando superexpresso em plantas de arroz (Sahoo *et al.*, 2012; Gill *et al.*, 2013; Nath *et al.*, 2015; Nath *et al.*, 2016), amendoim (Manjulatha *et al.*, 2014) e pimenta (Shivakumara *et al.*, 2017).

Esses autores detectaram que as plantas transgênicas que superexpressavam *PDH45* exibiram alta razão K^+/Na^+ , elevado conteúdo de clorofila e de prolina, além de maiores atividades de enzimas antioxidantes (SOD; APX; GR; GPX) e maiores conteúdos de glutatona e ascorbato, que atuam diretamente na limpeza de ROS. Sanan-Mishra *et al.*, (2005) apontam que o provável mecanismo pelo qual *PDH45* induz tolerância envolve a sua atuação em duas grandes vias: (1) Na tradução, pelo aumento e estabilização da síntese proteica e (2) Associação com complexos proteína-DNA para alterar a expressão gênica.

Gill *et al.*, (2013) detectaram que *PHD45* também interage diretamente com SOD, adenosina-5'-fosfosulfato-quinase, cisteína proteinase e eIF(4G), proteínas centrais na atividade antioxidante, maturação e degradação de proteínas e iniciação da tradução.

Ainda, Nath *et al.* (2015) propuseram que *PDH45* está provavelmente envolvido na deposição de barreiras hidrofóbicas apoplásticas e, conseqüente, inibindo o transporte de Na^+ ao caule e conferindo tolerância a salinidade em plantas de arroz e tabaco que superexpressavam esse gene. Plantas transgênicas que superexpressavam *PDH45* também exibiram maiores níveis de Ca^{2+} citosólico em suas raízes, além de menor acúmulo de Na^+ e H_2O_2 , quando submetidas a salinidade.

Essas plantas mostraram maior viabilidade celular e indução de importantes genes que atuam na homeostase iônica como: *SOS1* (transportador envolvido na remoção de Na^+ nas células), *OsAPX7* (ascorbato peroxidase 7), *OsCAX* (trocaador catiônico do tipo CAX), *OsTPC1* (proteína de canal de cálcio de dois poros) *OsNHX1* e *OsNHX2* (transportador vacuolar de Na^+/H^+), que foram *down-regulated* nas plantas selvagens (Nath *et al.*, 2016).

Outras helicases como *PDH47* (Vashisht *et al.*, 2005), *PsMCM6* (Dang *et al.*, 2011), *OsSUV3* (Sahoo *et al.*, 2014), *OsBAT1* (Tuteja; Sahoo 2015), *p68* (Tuteja *et al.*, 2014) e *OsXPB2* (Raikwar *et al.*, 2015) também foram associadas com tolerância a salinidade. Os mecanismos exatos de como as helicases, RNA ou DNA, provêm

tolerância a estresses ainda não foram totalmente elucidados e carecem de estudos adicionais.

2.5.1 RNA helicases

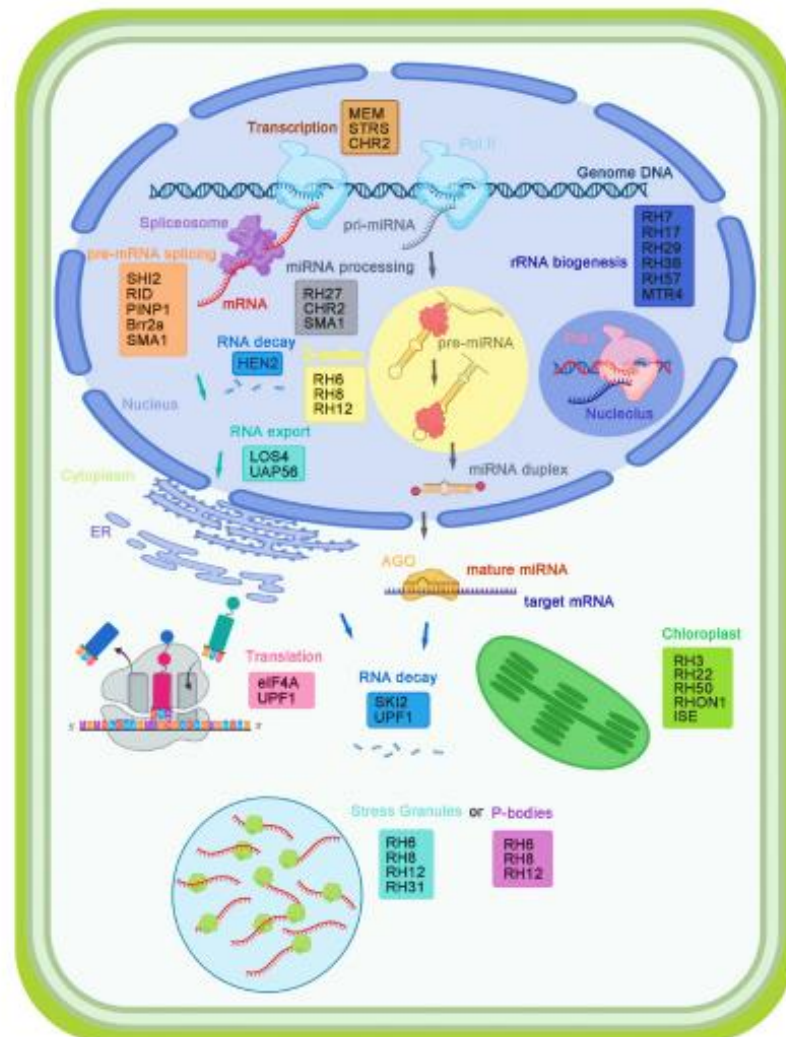
O metabolismo do RNA é diretamente afetado durante condições de estresse e pode acarretar em modificações na fisiologia da planta, podendo ter consequências letais para seu desenvolvimento (Pandey *et al.*, 2020). O envolvimento das RNA helicases no ajuste do metabolismo do RNA durante estresses tem sido amplamente pesquisado, devido a sua habilidade de remodelar estruturas ribonucleoproteicas (RNPs).

As RNA helicases são cruciais para transcrição, tradução, transporte de mRNA, silenciamento gênico pós-transcricional, *splicing* e biogênese do ribossomo (Cordin *et al.*, 2006; Pandey *et al.*, 2020) (Figura 12). A supressão da atividade de RNA helicases acarreta em efeitos graves em vários aspectos do desenvolvimento das plantas. A variedade de processos afetados pela atividade de RNA helicases enfatiza que a modulação de estruturas RNP é crucial para a regulação da expressão gênica e manutenção da homeostase celular (Linder *et al.*, 2009).

As habilidades das RNA helicases incluem a formação de pontes entre componentes da maquinaria transcricional, estabilizando-a nas regiões promotoras (Fuller-Pace, 2006); O desenrolamento de estruturas secundárias na 5' UTR do mRNA, garantindo maior eficiência da tradução (Svitkin *et al.*, 2020); O transporte através do complexo do poro nuclear (Chinnusamy *et al.*, 2008); Montagem do spliceossomo (Rocak; Linder, 2004); Reconhecimento de códons de terminação prematuros presentes em mRNAs e de proteínas aberrantes (Isken *et al.*, 2008); Processamento das subunidades pré-40S e pré-60S, mantendo a biogênese ribossomal dentre outras (Kalinina *et al.*, 2018).

Portanto, as RNA helicases funcionam como chaperonas de RNA, garantindo e facilitando o dobramento correto da molécula e, conseqüentemente, contribuindo para o funcionamento normal dos processos associados (Jarmoskaite; Russell, 2014).

Figura 12 - Participação das RNA helicases nos processos que envolvem o metabolismo de RNA numa célula.



Fonte: Adaptado de Li *et al.*, (2023).

2.5.2 DNA helicases

Como visto anteriormente, a produção de espécies reativas de oxigênio (ROS) é um mecanismo universal de resposta a todos os tipos de estresse. O excesso de ROS danos a estrutura do DNA. Caso não reparado, o dano ao DNA pode conduzir a efeitos citotóxicos ou genotóxicos, causando instabilidade genômica e até morte celular (Tuteja *et al.*, 2009). As vias de reparo do DNA são mecanismos que as células empregam para escapar dos danos e garantir a transmissão fidedigna da informação genética.

Alguns dos sistemas de reparo mais bem estudados são: o reparo por excisão de base (BER - *base excision repair*), reparo por excisão de nucleotídeo (NER - *nucleotide excision repair*), o reparo de *mismatch* (MMR - *mismatch repair*), reparo de

quebra na dupla fita (DSB – *double stranded break*) que compreende o reparo de junções homólogas (HR) e não-homólogas (NHEJ) (Tuteja *et al.*, 2001). Dentre os genes que codificam proteínas envolvidas nas maquinarias de reparo, estão as DNA helicases (Tuteja *et al.*, 2009).

DNA helicases são essenciais para o reparo pois desenrolam a área danificada tornando o DNA acessível as outras proteínas da maquinaria de reparo (Tuteja *et al.*, 2009). Essas proteínas também são requeridas em estágios iniciais do reparo, como por exemplo o complexo UvrAB (Rad3 ou XPD) que é necessária para a sensibilidade ao dano e também na remoção direta do DNA danificado, como no caso de UvrD (Evans *et al.*, 1997).

Em plantas, há vários trabalhos que associam as vias de reparo do DNA com tolerância a estresses (Hendrix *et al.*, 2020; Basu *et al.*, 2022), em contrapartida, há poucos relatos sobre a participação direta de DNA helicases nesses processos (Dang *et al.*, 2011; Raikwar *et al.*, 2015; Saifi *et al.*, 2018).

2.6 Genômica

Para entender a organização das bases das funções celulares em variados níveis, as ciências ômicas têm sido largamente aplicadas. As ômicas consistem numa abordagem integrativa que reúne dados de experimentos em larga escala, que incluem a genômica, transcriptômica, proteômica e metabolômica (Fukushima *et al.*, 2009). Neste tópico vamos abordar especificamente aspectos da genômica e ferramentas aplicadas no seu estudo em plantas e seu papel no entendimento da biologia desses organismos.

O avanço nas tecnologias de sequenciamento e das técnicas de bioinformática permitiram um grande aumento na disponibilidade de genomas de plantas, permitindo análises cada vez mais detalhadas de caracteres de importância econômica, tais como resistência a estresses bióticos e abióticos (Neale *et al.*, 2017). O maior requisito para estudos em escala genômica é a disponibilidade de um genoma de referência relativo à espécie a ser analisada.

Em plantas, o primeiro genoma de referência foi obtido de *Arabidopsis thaliana* (*The Arabidopsis Genome Initiative*, 2000). Cinco anos depois, o genoma de arroz foi completado usando cromossomos artificiais de bactérias (BACs - *Bacterial Artificial Chromosomes*) e sequenciamento de Sanger (Sasaki, 2005). Ao final de 2020, 1.031

genomas de referência ou *drafts* tinham sido disponibilizados para 788 espécies de plantas (Sun *et al.*, 2021).

O desenvolvimento das tecnologias de sequenciamento de nova geração (NGS – *Next-generation sequencing*) reduziram drasticamente o custo do sequenciamento (Sun *et al.*, 2021). Algumas das plataformas NGS incluem: 454 (Roche), SOLiD (*Life Technologies*), HiSeq (Illumina), MiSeq (Illumina) e Ion Torrent (*Life Technologies*). As tecnologias NGS permitem o sequenciamento mesmo com baixas concentrações de DNA, e são aplicadas no sequenciamento e re-sequenciamento de genomas, metagenômica, transcriptômica e na medicina pessoal (Kyriakidou *et al.*, 2018).

As técnicas NGS aumentaram a quantidade de informação disponível, mas a complexidade referente a montagem dos fragmentos *short read* gerados se tornou um desafio (Kyriakidou *et al.*, 2018). Esse fator é particularmente considerável quando se trata de genomas vegetais, que apresentam extensas regiões repetitivas, elevado tamanho dos genomas e poliploidia, que representam grandes dificuldades para a etapa de montagem (Kersey, 2019).

Logo, o desenvolvimento das tecnologias de sequenciamento de terceira geração, caracterizadas por gerarem *long reads*, foi essencial para a melhoria da acurácia e contiguidade dos genomas e serem montados e disponibilizados (Kersey *et al.*, 2019). Existem duas abordagens de tecnologias de sequenciamento de terceira geração: o sequenciamento *long-read* e a *long-range scaffolding* (Kersey *et al.*, 2019).

Dentre as técnicas de sequenciamento *long read*, a mais amplamente utilizada é a *Pacific Biosciences' Single Molecule Real-Time* (SMRT), com tamanho médio de *reads* de 20 Kb. A Illumina também desenvolveu a tecnologia *Synthetic Long-Reads* (SLR), com média de *reads* de até 10 Kb. Um outro exemplo é o *Minion* da Oxford Nanopore, que pode gerar *reads* de até 12 Kb (Schmidt *et al.*, 2017). A técnica de captura da conformação dos cromossomos (Hi-C) representa uma tecnologia de *long-range scaffolding*, e auxilia bastante na obtenção de montagens em escala de cromossomos (Lieberman-Aiden *et al.*, 2009).

2.6.1 Bancos de dados biológicos

Os bancos de dados são essenciais pois os dados gerados por tecnologias de sequenciamento se acumulam muito rápido. Existem vários bancos de dados que incluem coleções de genomas e de informações que foram especificamente

desenvolvidos para plantas. Dentre esses, podemos citar: Phytozome (<https://phytozome-next.jgi.doe.gov/>) (Goodstein *et al.*, 2012), EnsemblPlant (<http://plants.ensembl.org/>) (Bolser *et al.*, 2017), PlantGDB (<http://www.plantgdb.org/>) (Duvick *et al.*, 2007), PLAZA (<https://bioinformatics.psb.ugent.be/plaza>) (Van Bel *et al.*, 2018), Gramene (<http://www.gramene.org/>) (Gupta *et al.*, 2016), dentre outros.

Existem também os bancos especializados em determinadas espécies ou grupos como por exemplo: *Sol Genomics Network* (<https://solgenomics.net/>) (Fernandez-Pozo *et al.*, 2015), TAIR – *The Arabidopsis Information Resource* (<https://www.arabidopsis.org/>) (Rhee *et al.*, 2003), RGAP - *The Rice Genome Annotation Project* (<http://rice.uga.edu/>) (Kawahara *et al.*, 2013), BRAD – *The Brassicaceae Database* (<http://brassicadb.cn/>) (Cheng *et al.*, 2011), MaizeGDB (<https://www.maizegdb.org/>) (Lawrence *et al.*, 2004), dentre outros.

Há ainda bancos de dados específicos para determinadas informações genômicas tais como: PlantTFDB (<http://planttfdb.gao-lab.org/>) (Jin *et al.*, 2017), PlantProm – *Plant Promoter Database* (<https://ppdb.agr.gifu-u.ac.jp/>) (Shahmuradov *et al.*, 2003), ChromDB – *The Chromatin Database* (<http://www.chromdb.org/>) (Gendler *et al.*, 2008), PlantsP – *The Plant Phosphorylation Database* (<http://PlantsP.sdsc.edu>) (Gribskov *et al.*, 2001), PlantCARE – *Plant Cis Acting Regulatory element* (<http://bioinformatics.psb.ugent.be/webtools/plantcare>) (Lescot *et al.*, 2002), dentre outros.

Esses bancos de dados integram grandes volumes de dados sobre genótipos e fenótipos, constituindo um recurso valioso para melhoristas e a oportunidade para mineração de dados para descobrir e analisar novos genes candidatos associados a traços de interesse (Hu *et al.*, 2018).

2.6.2 Controle de qualidade do sequenciamento

O controle de qualidade é a primeira etapa na análise de dados de sequenciamento de alto rendimento. Nessa etapa, são avaliadas um conjunto de estatísticas em um arquivo de *reads*, que abrangem a contagem de frequência relativa de nucleotídeos em cada posição em um conjunto de leituras (Ewing; Green, 1998). Esse controle de qualidade também mede a frequência de contaminantes e de adaptadores, que não fazem parte da amostra de interesse.

A ferramenta FastQC é a mais amplamente empregada para checagem da qualidade de sequenciamento e permite visualização rápida dos dados evidenciando possíveis pontos problemáticos dentro dos dados (Andrews *et al.*, 2010). Após a checagem, os dados brutos devem ser pré-processados para remoção de sequências de baixa qualidade e de artefatos do sequenciamento, que podem facilmente comprometer análises posteriores.

Uma das ferramentas mais empregadas com esse objetivo é o Trimmomatic (Bolger *et al.*, 2014). As principais etapas realizadas são a trimagem ou remoção de sequências técnicas (ex.: adaptadores e *primers*) e filtragem da qualidade, que explora o *score* de qualidade de cada base (tipicamente valor de Phred maior ou igual a 25 ou 30).

2.6.3 Montagem e avaliação da qualidade

Existem duas abordagens básicas para a montagem de genomas. A primeira é a abordagem por referência, que utiliza como guia sequências já montadas de organismos taxonomicamente relacionados. A segunda é a abordagem *De novo*, que monta sequências de organismos que nunca foram sequenciados antes (Pop, 2009).

A duas abordagens não são mutuamente exclusivas, de modo que até em casos onde genomas de referência estão disponíveis, regiões que variam muito no genoma a ser sequenciado necessitam ser montadas *de novo* (Kyriakidou *et al.*, 2018). A abordagem comparativa é frequentemente utilizada no re-sequenciamento de genomas, para corrigir erros de montagem ou estender *contigs* já existentes na montagem e também para identificação de variantes (Kyriakidou *et al.*, 2018).

Na abordagem *de novo*, a sequência do genoma é construída a através da sobreposição das *reads* sequenciadas, frequentemente utilizando algoritmos baseados em gráficos. Essa abordagem é bastante dificultada no caso de *short reads*. Uma variedade de *softwares* montadores já foram desenvolvidos buscando obter a sequência completa de genomas. Dentre os montadores que montam por referência estão: RECORD (Buza *et al.*, 2015), AMOSmp (Pop *et al.*, 2004), ALLPATHS-LG (Gnerre *et al.*, 2011) e CAP3 (Huang; Madan, 1999).

Outros montadores para montagem *de novo* incluem o Velvet (Zerbino; Birney, 2008), ABySS (*Assembly By Short Sequences*) (Simpson *et al.*, 2009), SPAdes (Bankevich *et al.*, 2012) e MaSuRCA (*Maryland Super-Read Celera Assembler*) (Zimin

et al., 2013). Há também montadores especificamente desenhados para processamento de *long reads*, que incluem: FALCON (Chin *et al.*, 2016), NECAT (Chen *et al.*, 2020), Mecat2 (Xiao *et al.*, 2017) e Canu (Koren *et al.*, 2017), Raven (Vaser; Sikic, 2021) dentre outros.

Antes de iniciar o processo de anotação é fundamental entender a qualidade das sequências montadas, se há acurácia está baixa ou se existem erros de montagem. Existem várias métricas propostas para medir a qualidade da montagem (Wajid *et al.*, 2016). A mais conhecida é o N50, que consiste no tamanho do menor *contig/scaffold* em que metade da montagem está representada. Uma variante do N50 é o NG50, que constitui no tamanho do menor *contig/scaffold* que se encontra na metade de um genoma de referência.

Outra métrica muito utilizada é o L50, que representa quantos *contigs/scaffolds* foram necessários para alcançar metade do sequenciamento. Dentre os *softwares* empregados nesse processo podemos citar o QAST (*Quality Assessment Tool for Genome Assemblies*) (Gurevich *et al.*, 2013) e o GenomeQC (Manchanda *et al.*, 2020).

2.6.4 Repeat masking

Para lidar com a natureza altamente repetitiva dos genomas vegetais, as regiões de repetição são usualmente mascaradas. As sequências repetitivas abrangem três classes: (1) repetições locais, que incluem repetições simples e em *tandem*; (2) famílias de repetições dispersas, que compreendem os elementos transponíveis; (3) duplicações segmentais (Bao; Eddy, 2002). A presença de ORFs (*Open Reading Frames*) dentro de transposons (ex.: transposases e transcriptases reversas) pode confundir a predição de genes na montagem (Bao; Eddy, 2002).

As estratégias de identificação de elementos repetitivos abrangem duas grandes categorias: (1) detecção *de novo* e (2) detecção por similaridade. O processo de identificação ocorre pela busca de características estruturais tais como domínios associados a transposons utilizando bancos de dados de elementos repetitivos conhecidos, como o RepBase (<https://www.girinst.org/replib/>) (Bao *et al.*, 2015) e o TIGR (<https://www.hsls.pitt.edu/>) (Ouyang; Buell, 2004).

As ferramentas RepeatMasker (<https://www.repeatmasker.org/>) (Chen, 2004) e RepeatModeler2 (<http://www.repeatmasker.org/RepeatModeler/>) (Flynn *et al.*, 2020)

são amplamente empregadas para esse objetivo. RepeatMasker utiliza as sequências disponíveis em bibliotecas, tais como RepBase, para identificar por similaridade as famílias de repetições em determinado genoma. A descoberta de novas famílias de repetição espécie-específicas requer algoritmos para detecção *de novo* desses elementos, algumas das ferramentas são RECON (Bao; Eddy, 2002) e RepeatScout (Price *et al.*, 2005).

2.6.5 Predição de genes e anotação

A predição de genes é a etapa mais importante da anotação de um genoma. Esse processo é geralmente dividido em dois tipos de abordagem: (1) *ab initio* e (2) baseada em similaridade. Na predição *ab initio*, métodos computacionais são utilizados para identificar elementos chave da estrutura dos genes (ex.: TATA-box, sítios de *splicing*, éxons, íntrons, sítios de iniciação e terminação da tradução) (Ouyang *et al.*, 2009).

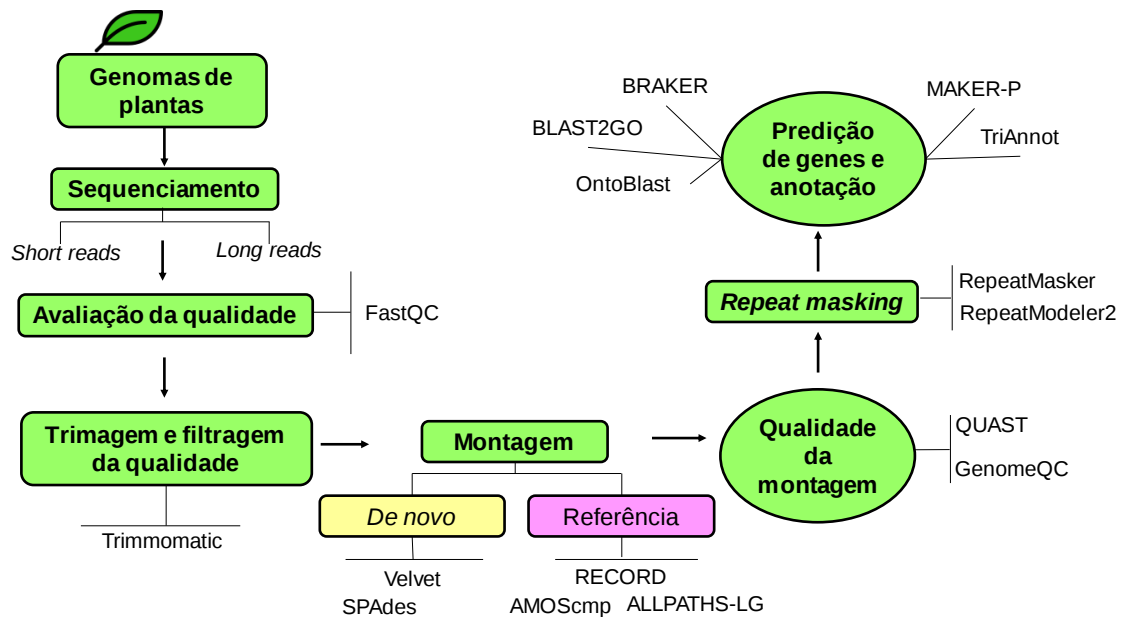
Dentre os *softwares* aplicados a predição de genes podemos citar FGENESH (Salamov; Solovyev, 2000), GeneMark.hmm (Lomsadze *et al.*, 2005) e GENESCAN (Burge; Karlin, 1997) que apresentam boa performance em plantas. Alguns *pipelines* automáticos também têm sido desenvolvidos para anotação de genomas vegetais como TriAnnot (Leroy *et al.*, 2012), Seqping (Chan *et al.*, 2017) e BRAKER (Hoff *et al.*, 2016).

A utilização de vocabulários controlados é uma estratégia muito empregada para obter uma eficiente anotação. *Gene Ontology* (GO - <http://www.geneontology.org>) é o esquema mais extensivamente empregado para descrição da função de produtos gênicos (Ashburner *et al.*, 2000). Essa descrição é feita com base em três ontologias independentes denominadas: processo biológico, função molecular e componente celular (Ashburner *et al.*, 2000).

Dentre as ferramentas e recursos que foram desenvolvidos para anotação funcional e mineração de novas sequências através de vocabulários controlados podemos citar: BLAST2GO (www.blast2go.org) (Conesa; Gotz, 2008), InterProScan (<https://www.ebi.ac.uk/interpro>) (Zdobnoz; Apweiler, 2001), OntoBlast (<http://functionalgenomics.de/ontogate>) (Zehetner, 2003), AmiGO (<http://amigo.geneontology.org/>) (Carbon *et al.*, 2009), dentre outros.

Logo, a obtenção de um genoma de referência de alta qualidade confere benefícios em análises genéticas posteriores e na tomada de decisões em programas de melhoramento. Dentre esses benefícios, as elevadas chances de encontrar mecanismos moleculares que atuam na variação de traços de interesse, tais como estresses bióticos e abióticos (Benevenuto *et al.*, 2019).

Figura 13 – Fluxo de um *pipeline* para obtenção de um genoma vegetal. As ferramentas referentes a cada etapa são mencionadas ao lado



Fonte: Silva (2024).

2.6.6 Análises *downstream*

Além dos bancos de dados biológicos, uma variedade de ferramentas é amplamente empregada em análises para identificação de genes e de seus ortólogos, estrutura gênica, análise filogenética/fenética, localização cromossômica, identificação de elementos *cis* regulatórios em promotores, dentre outros. Essa etapa, denominada de genômica estrutural é o ponto de partida do processo que visa compreender como as estruturas que compõem o genoma estão organizadas e como contribuem para o desempenho de suas funções. Na tabela 3 estão exemplificadas algumas das ferramentas mais utilizadas durante essa etapa.

Tabela 3 - Ferramentas utilizadas na etapa de genômica estrutural.

Ferramenta	Descrição	Referência
TBtools (<i>Toolkit for Biologists integrating various biological data-handling tools</i>)	Mais de 130 funções, que incluem: representação gráfica de motivos e domínios; mapeamento de genes em cromossomos; visualização de dados de expressão etc.	(Chen et al., 2020)
GSDS (<i>Gene Structure Display Server</i>)	Visualização de estruturas éxon-íntron.	(Hu et al., 2015)
PLANTCARE (<i>Plant Cis-Acting Regulatory Element</i>) e PlantPAN (<i>Plant Promoter Analysis Navigator</i>)	Identificação de sítios de ligação de fatores de transcrição (TFBSs) e outros elementos regulatórios.	(Chow et al., 2016); (Lescot et al., 2002)
OrthoFinder	Detecção de ortólogos.	(Emms; Kelly 2019);
McScanX (<i>Multiple Collinearity Scan Toolkit X</i>)	Detecção de eventos de duplicação.	(Wang et al., 2012)
MEGA (<i>Molecular Evolutionary Genetics Analysis</i>) e Clustal	Alinhamento de sequências e construção de árvores.	(Tamura et al., 2007); (Larkin et al., 2007)
Evolview e iTOL	Visualização e edição de árvores filogenéticas.	(Subramaniam et al., 2019)
SMART (<i>Simple Modular Architecture Research Tool</i>) e CDD (<i>Conserved Domain Database</i>)	Identificação e anotação de domínios de proteínas e visualização de suas arquiteturas.	(Letunic et al., 2021); (Marchler-Bauer et al., 2015)
PFAM (http://pfam.xfam.org/)	Banco de dados que abriga perfis HMM para identificação de famílias proteicas.	(Finn et al., 2016)
Jalview (https://www.jalview.org/)	Edição e anotação de alinhamentos múltiplos, e predição de estruturas secundárias.	(Waterhouse et al., 2009)
MEME (<i>Multiple Em for Motif Elicitation</i>)	Detecção, análise e visualização de assinaturas de sequência (motivos).	(Bailey et al., 2009)
ExPasy e JVirGel (http://www.jvirgel.de/)	Computam parâmetros físico-químicos a partir das sequências peptídicas.	(Gasteiger et al., 2005); (Hiller et al., 2006)

2.7 Genoma de *Jatropha curcas*

Em 2011, foi disponibilizado o primeiro genoma de referência para *J. curcas*, sequenciado pela combinação de Sanger junto com a plataforma 454 da Roche (Sato *et al.*, 2011). Desde então, foram disponibilizadas mais três montagens para o genoma de *J. curcas* com o intuito de prover informações para melhor entender o potencial da planta (Tabela 4). Jalali *et al.*, (2020) apresentaram a primeira montagem do genoma de *J. curcas* construída utilizando a tecnologia de *long-range scaffolding* Hi-C.

Apesar da utilização da metodologia de captura da conformação cromossômica, o nível de cromossomos não foi alcançado. Em contrapartida, a montagem apresenta diminuição de mais de 50% do número de *contigs* e *scaffolds* e aumento de mais de dez vezes do valor N50, comparado com a última montagem disponível (Kancharla *et al.*, 2019). Além disso, os autores conseguiram aumentar a quantidade de genes codificadores de proteínas, melhoraram a orientação dos *scaffolds*, o posicionamento dos genes no genoma e acurácia da montagem.

Tabela 4 - Informações gerais sobre as montagens disponíveis para *J. curcas*, incluindo: status da montagem; tecnologia de sequenciamento empregada; quantidade de genes codificadores de proteínas; cobertura; valores de N50 e L50 referentes aos *scaffolds* das respectivas montagens. As informações foram obtidas a partir do NCBI (*National Center for Biotechnology Information*; <https://www.ncbi.nlm.nih.gov/>).

Status	Tecnologia	Genes	Cobertura	L50	N50	Referência
<i>Scaffold</i>	Sanger; 454	9.870	89.5x	4.610	15.950	(Sato <i>et al.</i> , 2011)
<i>Scaffold</i>	Illumina	21.735	189x	96	746.835	(Wu <i>et al.</i> , 2015)
<i>Scaffold</i>	Illumina; PacBio	20.759	65x	439	169.453	(Kancharla <i>et al.</i> , 2019)
<i>Scaffold</i>	Illumina; Hi-C	22.588	190x	36	1.581.322	(Jalali <i>et al.</i> , 2020)

2.8 Transcriptômica

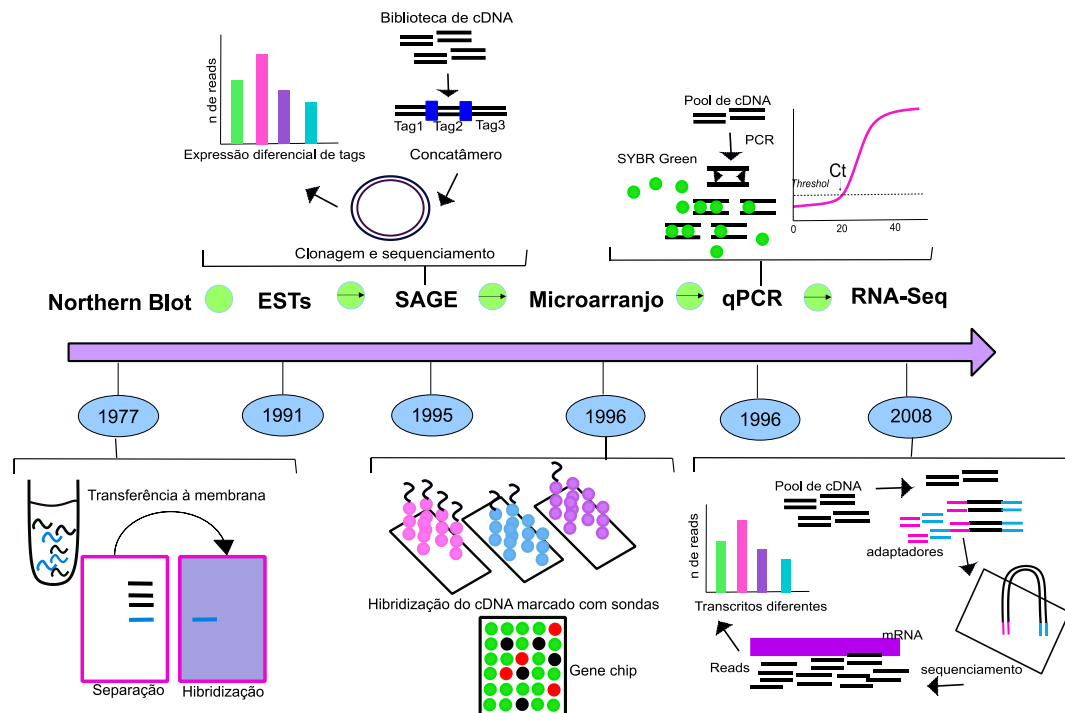
O transcriptoma compreende o conjunto de transcritos numa célula, tecido ou organismo em determinada condição fisiológica ou fase do desenvolvimento. A abordagem transcriptômica permite o estudo da população de transcritos (mRNA e RNAs não codificantes) e a obtenção de perfis de expressão gênica. Os estudos em transcriptômica são baseados em duas categorias: hibridização e sequenciamento (Shimkets, 2004).

Dentre as técnicas utilizadas para análise da expressão gênica estão: o microarranjo (*complementary DNA microarrays*) que é uma abordagem realizada por hibridização (Schena *et al.*, 1995); a PCR quantitativa em tempo real (qPCR) (Heid *et al.*, 1996); *Expressed sequence Tags* (ETs) (Adams *et al.*, 1991); SAGE (*Serial Analysis of Gene Expression*) (Velculescu *et al.*, 1995) e suas versões otimizadas pelo maior tamanho da *tag* LongSAGE 17-21 pb (*Long Serial Analysis of Gene Expression*) (Saha *et al.*, 2002), SuperSAGE 26 pb (*Super Serial Analysis of Gene Expression*) (Matsumura *et al.*, 2008) e HT-SuperSAGE (*High-Throughput-SuperSAGE*) (Matsumura *et al.*, 2010); e o RNA-Seq (Mortazavi *et al.*, 2008).

No microarranjo, as moléculas de cDNA marcadas com fluorescência são hibridizadas em sondas fixadas em uma superfície sólida (Schena *et al.*, 1995). A principal desvantagem da técnica é que sua aplicação é restrita a espécies que possuem genoma ou transcriptoma disponível para desenho das sondas. Na SAGE, e suas variantes, a geração de perfis de transcrição é baseado nas mudanças de frequência dos mRNAs que são capturados a partir da cauda poliA, que são representados por pequenas *tags*.

As *tags* são extraídas dos cDNAs por enzimas de restrição (ex.: EcoP15I), e seu tamanho é definido pela distância entre os sítios de reconhecimento e clivagem da enzima. A principal diferença da HT-SuperSAGE e suas versões anteriores, é a sua combinação com tecnologias de sequenciamento de alto rendimento, que permite a obtenção de mais dados com otimização do tempo (Matsumura *et al.*, 2008; Matsumura *et al.*, 2010).

Figura 14 – Diagrama ilustrativo da linha do tempo mostrando a evolução das principais técnicas de análise da expressão gênica



Fonte: Adaptado de Agarwal *et al.*, (2014)

2.8.1 RNA-Seq

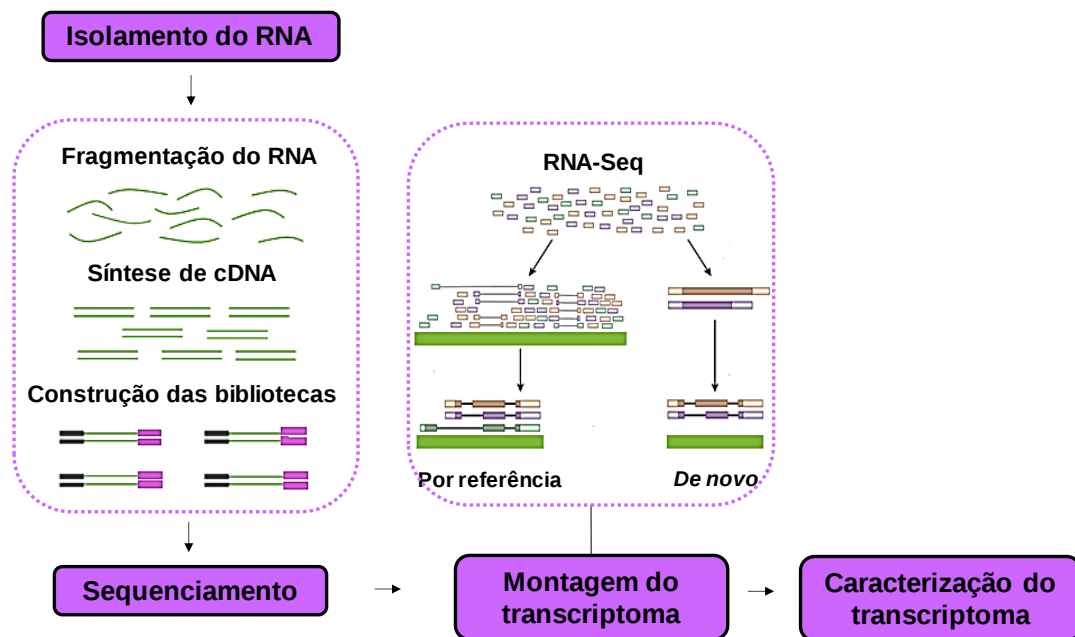
O RNA-Seq é a técnica de transcriptômica mais utilizada atualmente. Os dados gerados podem ser utilizados em vários tipos de análises tais como identificação de transcritos diferencialmente expressos entre tratamentos, análises de variantes e caracterização de *splicing* alternativo (Mortazavi *et al.*, 2008).

Por ser uma tecnologia de arquitetura aberta e por apresentar limites inferiores e superiores de detecção mais estendidos, a tecnologia possibilita a quantificação mais precisa da expressão diferencial e também a identificação de transcritos de baixa abundância. O RNA-Seq apresenta potencial na resolução de muitos problemas biológicos e é o mais adequado para eucariotos multicelulares (Maroli *et al.*, 2018).

As etapas de um experimento de RNA-Seq estão descritas na figura 15. Resumidamente, o processo inicia com a extração do RNA das amostras de interesse. Em seguida, esse RNA é fragmentado e convertido em cDNA através de um *primer* oligodT e uma transcriptase reversa. Em ambas as extremidades dos cDNAs são ligados adaptadores específicos e as bibliotecas de cDNA são então construídas e

sequenciadas. Após o sequenciamento e montagem do transcriptoma, a presença e quantidade de cada RNA pode ser calculada e posteriormente comparada com os dados de outra amostra. Também há a possibilidade de detectar e quantificar RNAs de todas as classes biológicas (Mortazavi *et al.*, 2008).

Figura 15 - Representação esquemática das etapas relativas ao RNA-Seq



Fonte: Adaptado de Maroli *et al.*, (2018).

Antes da montagem do transcriptoma, as *reads* sequenciadas precisam ser pré-processadas para remoção de quaisquer informações de natureza não biológica. Esse processo é essencial para obtenção de uma boa montagem, e consequentemente, um transcriptoma de boa qualidade. Para avaliação da qualidade e remoção das *reads* de baixa qualidade e de sequências de adaptadores, podemos citar os *softwares* FASTX-Toolkit (Gordon; Hannon, 2010) e rnaQUAST (Bushmanova *et al.*, 2016).

Após o processo de avaliação e filtragem da qualidade, as *reads* são usualmente mapeadas em um genoma ou transcriptoma disponível. Nessa etapa são principalmente avaliados a porcentagem de *reads* mapeadas (indicador da acurácia do sequenciamento e presença de DNA contaminante) e a uniformidade da cobertura das *reads* nos éxons (pode indicar baixa qualidade do RNA na amostra) (Conesa *et al.*, 2016). Ferramentas empregadas nessa etapa incluem: TopHat2 (Kim *et al.*, 2013), Bowtie2 (Langmead; Salzberg, 2012), STAR (Dobin *et al.*, 2013).

Após o alinhamento, o conteúdo GC e possíveis vieses no tamanho dos genes devem ser quantificados pois indicam a qualidade da purificação do RNA (Conesa *et al.*, 2016). Dentre os *softwares* comumente utilizado pode-se citar o EDA-Seq (Risso *et al.*, 2011) e NOISeq (Tarazona *et al.*, 2015). A última etapa de controle da qualidade se refere a checagem da reprodutibilidade entre as réplicas, que deve ser alta, de forma que as mesmas réplicas submetidas as mesmas condições devem clusterizar em uma análise de componente principal (PCA - *Principal Component Analysis*) (Conesa *et al.*, 2016).

Finalizado o controle de qualidade, a próxima etapa é a identificação ou reconstrução dos transcritos, que irá depender da disponibilidade, ou não, de sequência de referência. Como na montagem dos genomas, a montagem dos transcriptomas pode ser feita por dois tipos de abordagens, a *de novo* e a baseada em referência (genoma ou transcriptoma). Dentre as ferramentas empregadas para montagem de transcriptomas estão: Trinity (Grabherr *et al.*, 2011), SOAPdenovo-Trans (Xie *et al.*, 2014), Oases (Schulz *et al.*, 2012), Cufflinks (Trapnell *et al.*, 2010), BWA (*Burrow Wheelers alignment*) (Li; Durbin, 2009) dentre outras.

Seguindo a montagem do transcriptoma, a próxima etapa é a contagem de *reads* mapeadas em cada unidade codificante para posterior quantificação da expressão. Os valores referentes à abundância de transcritos em um RNA-Seq são FPKM (*fragments per kilobase of transcript per million mapped reads*); RPKM (*reads per kilobase of transcript per million mapped reads*) e TKM (*transcripts per million*) (Conesa *et al.*, 2016).

A quantificação da abundância pode ser realizada pelo *software* Cufflinks. Os valores de expressão obtidos devem ser comparados entre amostras. Os valores de RPKM, FPKM, e TPM normalizam a profundidade do sequenciamento, que é o fator mais importante quando se realiza a comparação (Conesa *et al.*, 2016). Os métodos mais utilizados para análise da expressão gênica diferencial são o edgeR (Robinson *et al.*, 2010) e o DESeq2 (Love *et al.*, 2014).

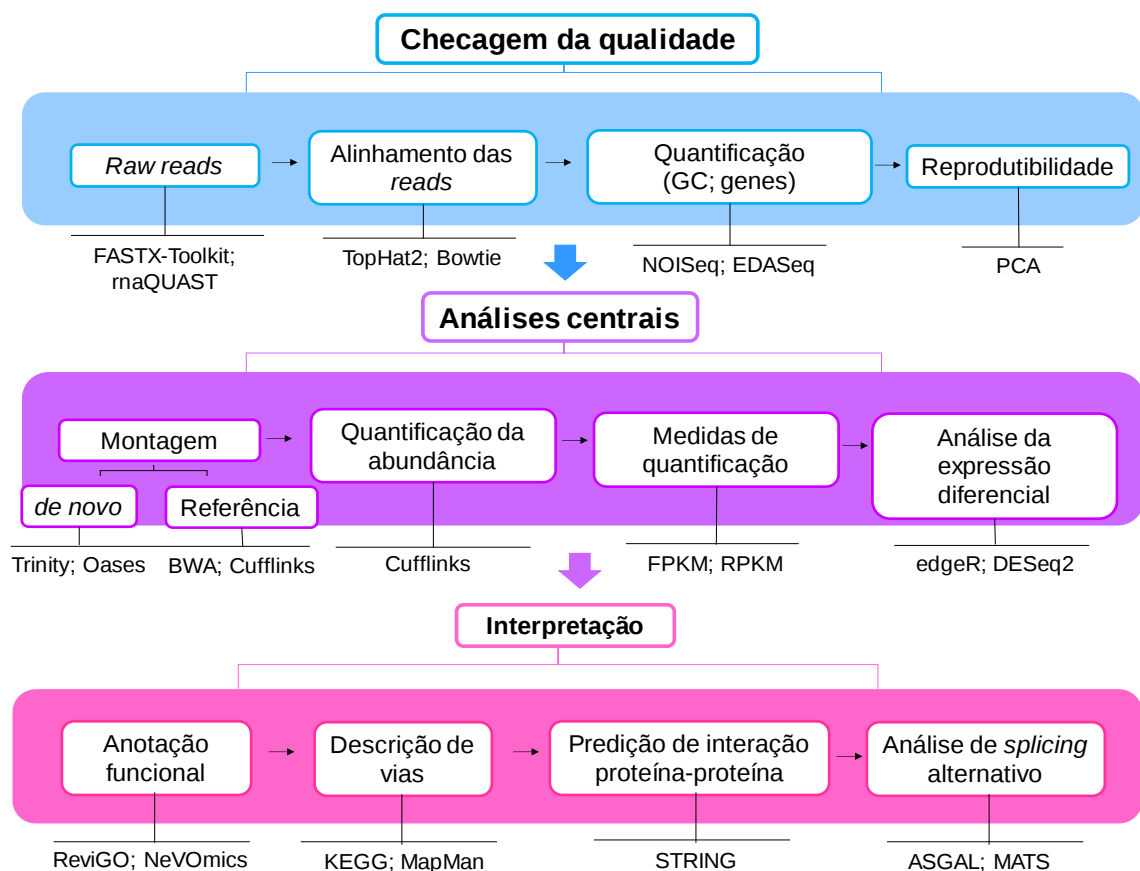
Essas ferramentas utilizam esses dados de quantificação para determinar quais genes têm diferenças estatisticamente significantes, provendo informações sobre o nível de expressão e a magnitude da diferença de expressão entre as amostras (Conesa *et al.*, 2016). A análise da expressão gênica diferencial pode fornecer informações valiosas sobre mecanismos genéticos que podem estar associados a diferenças fenotípicas tais como padrões de crescimento de desenvolvimento e

tolerância a estresses bióticos e abióticos (Conesa *et al.*, 2016). Um esquema ilustrando as etapas de análise de um transcriptoma RNA-Seq está descrito na figura 16.

A principal aplicação do RNA-Seq é a análise da expressão diferencial. A última etapa nos estudos em transcriptômica é a caracterização molecular das funções ou vias nos quais os genes diferencialmente expressos identificados estão envolvidos (Conesa *et al.*, 2016, Muhammad *et al.*, 2020).

Dentre as ferramentas utilizadas nessa etapa podemos citar: KEGG (*Kyoto encyclopedia of genes and genomes*) (Kanehisa *et al.*, 2017); ReviGO (Supek *et al.*, 2011); NeVOmics (*Network-based Visualization for Omics*) (Zúniga-Léon *et al.*, 2018); MapMan (Thimm *et al.*, 2004); STRING (Szklarczyk *et al.*, 2021); Para a detecção de eventos de *splicing* alternativo pode-se empregar as ferramentas MATS (*Multivariate Analysis of Transcript Splicing*) (Shen *et al.*, 2014); ASGAL (*Alternative Splicing Graph Aligner*) (Denti *et al.*, 2018) dentre outras.

Figura 16 - Representação esquemática das etapas de análise de um RNA-Seq



Fonte: Adaptado de Muhammad *et al.*, (2020).

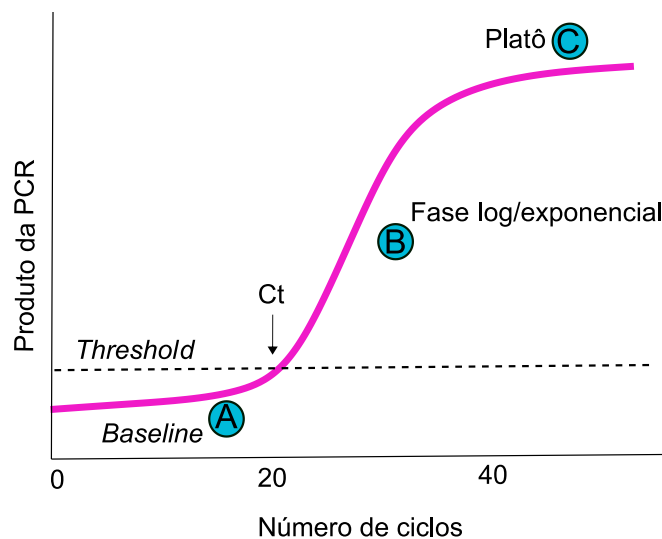
2.8.2 PCR quantitativa em tempo real (qPCR)

A qPCR é uma tecnologia rápida e altamente sensível de detecção, considerada a técnica padrão ouro para análise da expressão gênica (Segundo-Val *et al.*, 2016). A qPCR constitui uma evolução da PCR convencional, desenvolvida por Kary Mullis (Mullis *et al.*, 1987). São muitas as aplicações da técnica, tais como ensaios de detecção de cópias de genes (ex.: HIV, SARS-Cov-2, HPV etc.), genotipagem, controle de qualidade industrial, análise de variantes de *splicing*, medicina forense dentre outras (Heid *et al.*, 1996).

A reação de qPCR progride através de três fases principais: (1) *Baseline*; (2) fase exponencial; (3) platô (Figura 17). Nos ciclos iniciais, valores de fluorescência não são suficientes para serem detectados pelo equipamento e permanecem no *baseline*. Quando os substratos (ex.: dNTPs) e a atividade enzimática passam a ser limitantes, a reação alcança o platô. Um *threshold* é atribuído com base na variabilidade do *baseline* (Heid *et al.*, 1996; Arya *et al.*, 2005)

O ponto em que a fluorescência excede o *threshold* é utilizado para definir o C_t (*cycle threshold*). O C_t é relacionado a quantidade do alvo, quanto menor (mais inicial) o C_t maior a quantidade de RNA ou DNA na amostra. O C_t é um princípio básico da qPCR, sempre ocorre durante a fase exponencial e é um componente essencial da reação (Arya *et al.*, 2005; Segundo-Val *et al.*, 2016).

Figura 17 – Representação gráfica das fases da qPCR. A. *Baseline*; B. Fase exponencial; C. Platô



Fonte: Adaptado de Arya *et al.*, (2005).

Os principais detectores utilizados na qPCR são o SYBR Green® e as sondas TaqMan®. O SYBR Green® é um corante que se intercala ao DNA dupla-fita. À medida que polimerase amplifica a sequência alvo, mais *amplicons* são gerados e consequentemente ocorre aumento da fluorescência pela ligação do SYBR Green, permitindo assim a quantificação do produto. Curvas de *melting* são requeridas para determinar a especificidade da ligação do SYBR Green® ao DNA (Arya *et al.*, 2005; Segundo-Val *et al.*, 2016).

A TaqMan® utiliza uma Taq polimerase que possui atividade de exonuclease 5'-3'. As sondas TaqMan® são marcadas com um repórter fluorescente e um quencher, que são mantidos próximos até que a hibridização ao alvo ocorra. Esses sítios são hidrolisados pela atividade exonuclease da Taq polimerase, que resulta na liberação de fluorescência pelo repórter (Segundo-Val *et al.*, 2016).

As análises de qPCR podem ser do tipo absoluta (ex.: número de cópias de um RNA específico por amostra) e relativa (ex.: amostra 1 tem três vezes mais mRNA de um gene específico do que a amostra 3). Para quantificação absoluta, uma curva padrão do gene de interesse é requerida, de forma que uma diluição seriada de uma amostra conhecida é submetida à amplificação (Arya *et al.*, 2005).

Na quantificação relativa, o método mais utilizado é o $2^{-\Delta\Delta CT}$ (Livak *et al.*, 2001). Esse método preconiza a presença de um controle endógeno, um gene que é expresso de forma constante entre as amostras. Assim, o método compara as diferenças entre o gene de interesse e o controle endógeno, permitindo avaliar se um determinado gene é mais ou menos expresso numa amostra em relação a outra (Livak *et al.*, 2001).

A principal vantagem da qPCR é que a quantificação do produto é feita em tempo real durante cada ciclo, enquanto que na PCR convencional os produtos necessitam ser quantificados em gel de eletroforese. Logo, a qPCR elimina etapas de processamento pós-reação, diminuindo erros experimentais e possíveis contaminações (Heid *et al.*, 1996).

Ainda, na qPCR a quantificação é realizada durante a fase *log* de amplificação do produto. Na PCR convencional o produto é quantificado na fase de platô, que na prática apresentará diferentes níveis para diferentes amostras. Após um alto número de ciclos essa diferença é amplificada, conduzindo a erros de quantificação (Heid *et al.*, 1996; Arya *et al.*, 2005; Segundo-Val *et al.*, 2016).

3 OBJETIVOS

3.1 Objetivo geral

Estudar a genômica funcional, estrutural e comparativa de genes codificadores de helicases de *J. curcas* e avaliar resposta ao estresse salino, visando identificar genes candidatos para obtenção de plantas mais tolerantes a salinidade.

3.2 Objetivos específicos

1. Identificar e anotar os genes codificadores das famílias helicase no genoma de *J. curcas*;
2. Executar análises fenéticas e estruturais para classificar e caracterizar as famílias gênicas de helicases;
3. Estudar a ortologia de genes helicase entre os genomas de *M. esculenta* e *R. communis*;
4. Mapear de elementos *cis*-reguladores nos promotores dos genes codificadores de helicases de *J. curcas*;
5. Analisar a expressão e a modulação de transcritos helicase, oriundos das bibliotecas RNA-Seq de acessos contrastantes de *J. curcas*, produzidos em resposta ao estresse salino em relação ao respectivo controle negativo;
6. Propor e testar 20 pares de oligonucleotídeos iniciadores para transcritos alvos para validação de dados via qPCR;

4 ARTIGO 1

Artigo a ser submetido na revista Genetics and Molecular Biology

ANÁLISE DA FAMÍLIA SNF2 DE PINHÃO-MANSO (*Jatropha curcas* L.) E PERFIL DE EXPRESSÃO EM RESPOSTA À SALINIDADE

Running title: Família Snf2 de pinhão-manso

Rahisa Helena da Silva¹; Manassés Daniel da Silva¹; Bruna de Brito Souza¹; Antonio Félix da Costa²; Éderson Akio Kido¹

¹Departamento de Genética, Centro de Biociências, Programa de Pós-graduação em Genética e Biologia Molecular -PPGBM, Universidade Federal de Pernambuco, Recife, PE, Brasil.

²Instituto Agrônômico de Pernambuco - IPA, Recife, PE, Brasil.

Corresponding author: kido.ufpe@ufpe.br

Resumo

Estresses ambientais constituem grave ameaça a sobrevivência das plantas e são responsáveis por perdas consideráveis para a agricultura. Plantas de pinhão-manso são conhecidas como atrativas para produção de biodiesel, porém seu potencial é reduzido por ser uma espécie sensível a salinidade. As respostas adaptativas a estresses como a salinidade são governadas por mudanças transcricionais, que são principalmente conduzidas através da regulação da dinâmica da cromatina, que vai determinar a acessibilidade da maquinaria transcricional ao DNA. O remodelamento da cromatina é regulado pelas proteínas da família Snf2. Neste trabalho toda a família Snf2 de *J. curcas* (*JcSnf2*) foi analisada quanto a seus padrões estruturais e resposta a salinidade. As proteínas foram classificadas nas subfamílias descritas. De acordo com os agrupamentos obtidos, cada subfamília apresentou distribuições específicas de domínios e motivos conservados. As subfamílias do grupo Snf2-Like exibiram maiores variações em termos de ocorrência de motivos, domínios e estruturas éxon-íntron. As proteínas *JcSnf2* possuem o *core* catalítico comum relativo ao grupo e parecem se diversificar através de alterações na ocorrência dos blocos menos conservados (E-N) presentes entre os motivos detectados. Modelos tridimensionais obtidos remontaram ao dobramento das helicases de SF-II, indicando que essas proteínas são compatíveis com suas atividades. Regiões promotoras dos genes exibiram grande número de candidatos a elementos *cis*-regulatórios tais como Dof, MIKC_MADS e AP2, demonstrando que esses genes podem participar de rotas de resposta a estresses abióticos. Nas bibliotecas RNA-Seq analisadas, foram detectados 35 transcritos *JcSnf2* diferencialmente expressos ao estímulo de salinidade (150 mM de NaCl por 3 horas) e validações por qPCR permitiram confirmar a expressão de quatro desses candidatos. Redes PPI obtidas indicaram que as *JcSnf2* expressas nessas bibliotecas

possivelmente atuam através da substituição da variante de histona H2A.Z para regular a acessibilidade da cromatina a maquinaria da RNA polimerase II e ao reparo do DNA, aspectos importantes da resposta a estresse. Este é o primeiro relato da família Snf2 em pinhão-mansão, o conhecimento gerado irá auxiliar futuras investigações para o melhor entendimento sobre a participação dessas proteínas nas respostas da planta a salinidade.

Palavras-chave: Helicase; Bioinformática; Transcriptômica; Estresse abiótico

Introdução

Mudanças climáticas expõem as plantas a extremos de condições ambientais que afetam negativamente seu crescimento e desenvolvimento, por alterar regimes de chuvas, concentrações de CO₂ e distribuição de pragas (Chaudhry e Sidhu, 2022). Atualmente, cerca de um terço das terras aráveis no mundo é afetado pela salinidade (Santos et al., 2022). A salinidade conduz a toxicidade iônica devido a, dentre outros fatores, desbalanço na razão K⁺/Na⁺ (Assaha et al., 2017). Além disso, o estresse oxidativo, provocado pelo acúmulo de espécies reativas de oxigênio (ROS), também acompanha a salinidade (Sachdev et al., 2021). Portanto, o entendimento dos aspectos das respostas das plantas a salinidade, associado com a utilização de práticas de agricultura sustentável, é essencial para obter culturas produtivas sob o futuro cenário de mudança climática (Hernández, 2019).

O pinhão-mansão (*Jatropha curcas* L.) é uma planta não alimentar que pertence à família Euphorbiaceae considerada como fonte promissora de biodiesel (Achten et al., 2008). São arbustos que se desenvolvem facilmente em áreas não direcionadas a agricultura convencional (Openshaw 2000). Esses atributos tornam *J. curcas* uma alternativa sustentável para a produção de biodiesel, visto que o aumento da população humana pressiona a demanda por alimento e biocombustíveis (Kamel et al., 2017). Além disso, as plantas de pinhão-mansão também podem ser utilizadas para resolver questões ambientais como controle da erosão e recuperação de solos por biorremediação (Sani et al., 2022). Porém, as plantas são sensíveis a salinidade, principalmente em estágios iniciais de germinação (Lozano-Isla et al., 2018). Logo, a utilização de abordagens ômicas, tais como transcriptômica, para estudo e proposição de genes-candidatos provê informação útil sobre mecanismos moleculares

associados a vias de aclimação e defesa contra a salinidade, além de alvos biotecnológicos (Hernández, 2019).

O ajuste fino da expressão gênica durante o desenvolvimento e exposição a fatores externos é diretamente associado ao contexto da cromatina (Raxwal et al., 2020; Halder et al., 2022). As proteínas da família Snf2 (*Sucrose nonfermentable*) são membros de complexos de remodelamento da cromatina. Sua atividade ATPase provê energia para translocar o complexo e facilitar a ligação de proteínas ativadoras ao DNA nucleossomal, essencial para a regulação da dinâmica da cromatina (Liu et al., 2017; Lin et al., 2020). A família Snf2 está compreendida na superfamília II das helicases (SF-II), e estruturalmente se define pela presença de regiões conhecidas como “motivos helicase”, nomeados Q, I, Ia, II, III, IV, V e VI dispostos ao longo de dois domínios RecA-Like (Eisen et al., 1995; Flaus et al., 2006).

Nos genomas das plantas, a família Snf2 mostra variabilidade em número de genes (Bargsten et al., 2013). Em *Arabidopsis* e arroz, foram identificados 41 e 40 membros da família Snf2, respectivamente (Knizewski et al., 2008; Hu et al., 2013). Análises funcionais têm indicado que muitas dessas proteínas podem ter papéis importantes durante processos de desenvolvimento e resposta a estresses bióticos e abióticos (Hu et al., 2013; Cho et al., 2016; Song et al., 2021). O remodelador da cromatina *PICKLE* (*PKL*) é um regulador positivo de genes de resposta ao frio e plantas mutantes *pkl* são hipersensíveis ao frio e a salinidade (Yang et al., 2019).

O objetivo deste trabalho foi identificar e investigar os padrões estruturais dos genes da família Snf2 de *J. curcas*, além de avaliar como a família responde a estímulo de salinidade. Sendo assim, foram identificados um total de 26 genes codificadores de proteínas Snf2 no genoma de pinhão-mansão. De acordo com suas características estruturais, esses genes foram classificados nas subfamílias correspondentes. Os

padrões de domínios, motivos conservados, estrutura gênica também foram analisados. Modelos tridimensionais dos candidatos também foram propostos. As regiões promotoras desses genes foram analisadas quanto a presença de elementos *cis*-regulatórios. Para avaliar a resposta das Snf2 de pinhão-manso a salinidade, dados de RNA-Seq de raízes de acessos de *J. curcas* contrastantes foram empregados. Este é o primeiro relato *genome-wide* da família Snf2 para pinhão-manso. O conhecimento gerado pode auxiliar investigações futuras sobre essa família gênica em Euphorbiaceae, além do desenvolvimento de marcadores moleculares funcionais úteis para os programas de melhoramento de *J. curcas*.

Metodologia

Identificação dos membros da família Snf2 em *J. curcas* (*JcSnf2*) e estruturas éxon-íntron

O proteoma traduzido do pinhão-manso (GCF_014843425.1_RJC1_Hi-C_protein.faa) foi obtido do NCBI (<https://www.ncbi.nlm.nih.gov>). Para busca das sequências candidatas, foi aplicado o *Hidden Markov Model* (HMM). O perfil HMM SNF2-rel_dom (PF00176) do banco de dados de famílias proteicas do Pfam (<http://pfam.xfam.org/>) (Finn et al., 2015) foi utilizado como *query* para busca nas sequências de pinhão-manso, utilizando parâmetros *default* com um *e-value cut-off* de $< 1.0e^{-5}$. As potenciais sequências foram submetidas ao CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Marchler-Bauer et al., 2015) e ao SMART (<http://smart.embl-heidelberg.de/>) (Letunic et al., 2021) para identificação de domínios conservados. As estruturas éxon-íntron dos genes foram detectadas por comparação da sequência codificante (CDS) com a sequência genômica através do *Gene Structure Display Server* v2.0 (GSDS, <http://gsds.gao-lab.org/>) (Hu et al., 2015).

Análise dos promotores

As regiões promotoras (1.0 kb *upstream* ao sítio de início da transcrição) dos genes *JcSnf2* foram obtidas pelo painel *Genome Browser* do NCBI (<https://www.ncbi.nlm.nih.gov/genome/gdv/browser/>). Cada região promotora foi analisada pelo PlantRegMap (*Plant Transcriptional Regulatory Map* - <http://plantregmap.gao-lab.org/>) (Tian et al., 2020) utilizando a ferramenta FIMO (*Find Individual Motif Occurrences*) para identificar potenciais sítios de ligação de fatores de transcrição. A identidade dos potenciais elementos *cis*-regulatórios foi associada com motivos que apresentaram *p*-value *cut-off* < 10^{-2} e *q*-value *cut-off* < 10^{-2} . O software TBTools foi utilizado para visualização dos resultados (Chen et al., 2020).

Alinhamento de sequências, análise fenética e classificação

A partir da classificação proposta por Hu et al., (2013), sequências *Snf2* de *Arabidopsis* e *O. sativa* foram obtidas a partir dos bancos de dados do TAIR - *The Arabidopsis Information Resource* (<https://www.arabidopsis.org/>) e do iTAK - *Plant Transcription factor & Protein Kinase Identifier and Classifier* (<http://itak.feilab.net/cgi-bin/itak/index.cgi>), respectivamente. Estas sequências, juntamente com as sequências *JcSnf2* identificadas, foram alinhadas pelo algoritmo Muscle utilizando parâmetros *default* e a árvore foi construída pelo método *Neighbor joining* com *bootstrap* de 1.000 replicatas no software MEGA11 (Tamura et al., 2021). A árvore foi visualizada através da ferramenta online iTOL v6 – *Interactive Tree Of Life* (<https://itol.embl.de/>) (Letunic e Bork, 2019).

Características estruturais das proteínas *JcSnf2*

As proteínas *JcSnf2* identificadas foram analisadas pelo ExPASy (<http://web.expasy.org/protparam/>) (Gasteiger et al., 2005) para calcular o número de

aminoácidos, peso molecular (MW) e o ponto isoelétrico (pI). A localização subcelular foi predita através do *software* LOCALIZER v1.0.4 (Sperschneider et al., 2017). O programa MEME v5.3.3 (<https://meme-suite.org/meme/>) (Bailey, 2015), foi utilizado para identificar motivos conservados nas sequências, de acordo com os seguintes parâmetros: anr (*any numbers of repetitions*); número máximo de motivos: 20; tamanho dos motivos: 6-50 resíduos de aminoácidos.

Predição de estruturas secundárias e modelagem 3-D

A modelagem por homologia de candidatos *JcSnf2* foi conduzida utilizando a ferramenta *web* SWISS-MODEL (<https://swissmodel.expasy.org/>) (Waterhouse et al., 2018), com base no *template* PDB mais similar. Os modelos 3D gerados foram visualizados e editados utilizando o programa PyMOL v2.5 (<https://pymol.org/>) (DeLano, 2002). A acurácia dos modelos gerados foi avaliada através de gráficos Ramachandran. SOPMA foi utilizado como método estrutural secundário (*Self-Optimized Prediction Method with Alignment*; https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon e Deléage, 1995).

Perfil de expressão de *JcSnf2* em resposta à salinidade

Transcritos RNA-Seq foram alinhados, em análise BLASTx (e-value *cut off* e^{-10}), com as proteínas *JcSnf2* anteriormente identificadas. Dados RNA-Seq foram oriundos de um ensaio compreendendo dois acessos brasileiros de *J. curcas* com respostas diferentes ao NaCl, de fenótipo tolerante (*Jc183*) e de menor tolerância (*Jc171*) (Lozano-Isla et al., 2018), cujas raízes tinham sido expostas (3h) ao NaCl (150 mM) incorporado à solução de Hoagland. Informações acerca do ensaio de salinidade, extração do RNA e montagem do transcriptoma estão descritos em Souza et al., (2020). Transcritos *JcSnf2* foram declarados diferencialmente expressos de acordo

com $p\text{-value} < 0.0001$; *false discovery rate* (FDR) < 0.005 , sendo os respectivos valores de *fold change* (FC), baseados em $(FC) \geq 1$, informativos de modulação positiva da expressão (UR, upregulated) e ≤ -1 , da modulação negativa da expressão (DR, downregulated). A clusterização hierárquica dos transcritos *JcSnf2* foi feita com base nos valores FC utilizando o *software* Cluster v3.0 (<https://cluster2.software.informer.com/3.0/>), sendo a visualização feita através do *software* Java TreeView v1.1 (Saldanha, 2004).

Desenho de *primers* e validação da expressão por qPCR

Primers para qPCR foram desenhados a partir dos transcritos DEs, através da ferramenta Primer3 (Rozen e Skaletsky, 2000), com ajustes dos seguintes parâmetros: tamanho do *amplicon* (70-200 bp), conteúdo GC (45-55%), temperatura de dissociação (50-80°C), e CG *clamp* (um). As especificidades dos amplicons foram analisadas em curvas de dissociação geradas entre 65-95°C (20 min.), após 40 ciclos de qPCR. A eficiência de amplificação (Rasmussen, 2001) de cada alvo foi derivada de curvas de calibração feitas com cDNAs dos acessos, em diluições seriadas (sem diluição, 10^{-1} , 10^{-2} , 10^{-3} e 10^{-4}). As reações de qPCR foram feitas em termociclador LineGene 9660 (Bioer®), utilizando SYBR Green como sistema de detecção, em volume final de 10 µL, sendo um µL de cDNA da amostra (1:10), cinco µL de GoTaq® qPCR Master Mix (Promega), dois µL de água ultra-pura, e um µL de cada primer (0.05 µM). As reações de qPCR foram feitas em triplicata técnica e biológica, presença de controle negativo, e uso de dois genes de referência (Actina e β Tubulina), analisados em ensaio para esta finalidade e com as amostras do ensaio em questão. As reações seguiram a programação: desnaturação inicial a 95°C (2 min.), seguido por 40 ciclos de 95°C (15 s) e 60°C (60 s). Ciclos de quantificação (Cq), temperatura de dissociação, e quantificação absoluta e relativa, foram estimados com o *software*

proprietário da Bioer®. Dados de expressão foram calculados com o *software* REST© (v 2.0.13) (Pfaffl et al., 2002), com aplicação de teste de randomização com 10.000 permutações; além de testar a hipótese de diferenças significativas entre os grupos controle e tratamento. Para aumentar a confiabilidade dos resultados de qPCR, seguimos o protocolo MIQE (*The Minimum Information for Publication of Quantitative Real-Time PCR Experiments*; Bustin et al., 2009).

Redes de interação proteína-proteína (PPI) de *JcSnf2*

Redes PPI foram previstas a partir do *software* STRING v12.0 (Szklarczyk et al., 2023) contemplando as proteínas codificadas pelas *JcSnf2* declaradas DEs. O score de confiança da interação foi definido como 0.7 (*high*), utilizando o genoma de *Arabidopsis* como modelo. As evidências baseadas em *text-mining* foram excluídas. Termos GO enriquecidos na rede ($p\text{-value} < 1.0\text{e-}16$) foram visualizados através da ferramenta ReViGO (*Reduce and Visualize Gene Ontology*; <https://revigo.irb.hr/>; Supek et al., 2011). Redes PPI geradas foram visualizadas e editadas utilizando Cytoscape v3.10.1 (Shannon et al., 2003).

Resultados

Identificação dos membros da família *Snf2* em pinhão-mansão (*JcSnf2*) e estruturas éxon-intron

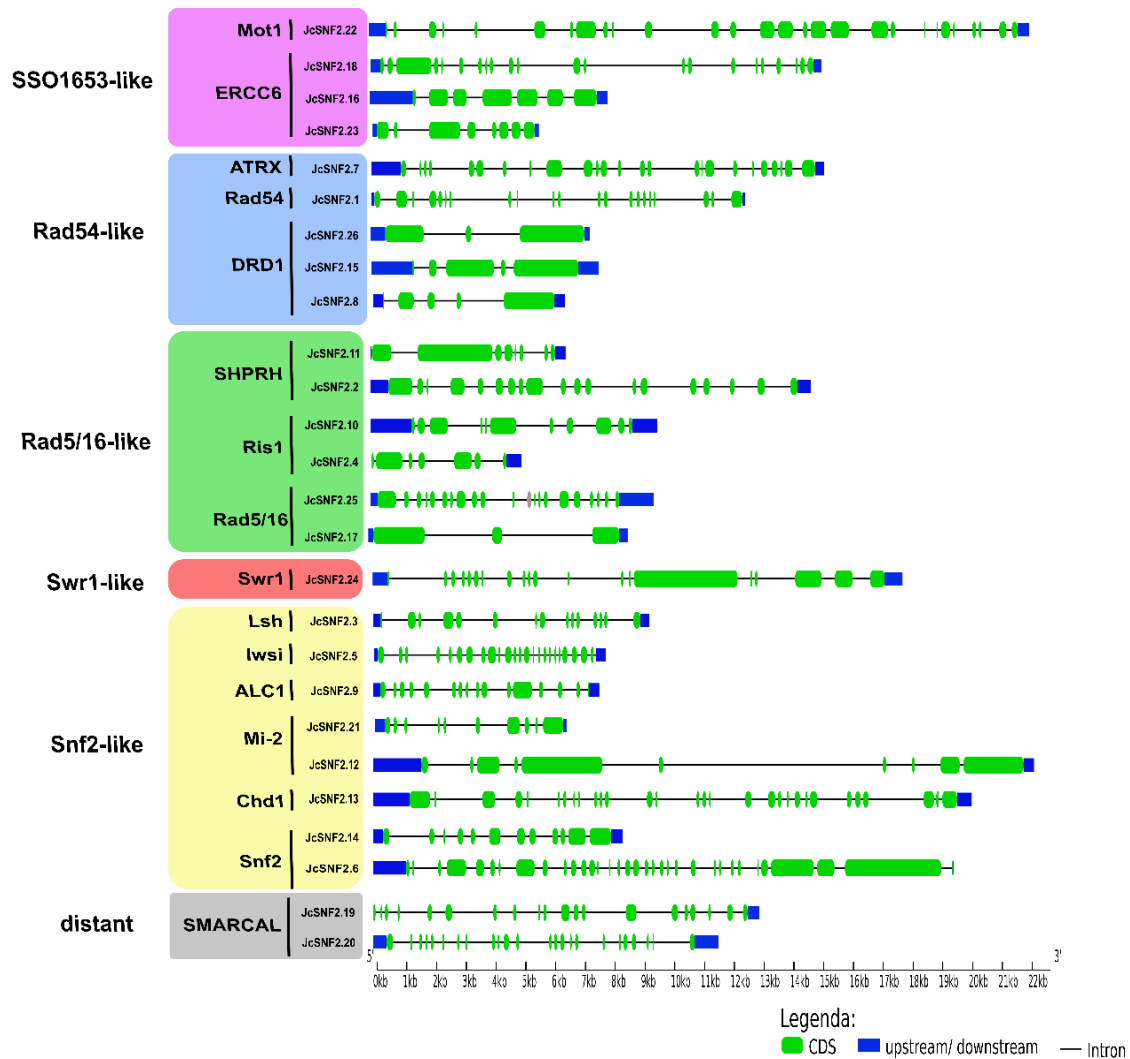
A abordagem *Hidden Markov model* (HMM) permitiu identificar 26 genes codificadores de 60 proteínas *Snf2* no genoma de referência da espécie (Jalali et al., 2020). Com base na posição nos *scaffolds*, estes genes foram nomeados *JcSnf2.1* – *JcSnf2.26*. As informações sobre os membros da família gênica estão descritas na tabela 1, incluindo tamanho da sequência codificante (CDS), tamanho do gene, tamanho da proteína primária e número de íntrons. O tamanho dos genes *JcSnf2*

variou de 4.916 até 22.616, e as regiões CDS de 2.085 até 10.230 (Tabela 1). Os genes *JcSnf2* apresentam de dois até 33 íntrons (Figura 1). Nenhum padrão de estrutura gênica muito específico foi observado dentro das subfamílias, apenas que ambos os dois genes da subfamília Mi-2 (*JcSnf2.21*; *JcSnf2.12*) apresentam nove íntrons cada (Figura 1). No geral, os genes mais ricos em íntrons foram os das subfamílias do grupo *Snf2-like* (9 – 33 íntrons), em comparação com as demais subfamílias. Os genes das subfamílias DRD1 (2 e 4 íntrons) e Ris1 (10 e 6 íntrons) incluíram os genes com menores quantidades de íntrons (Figura 1). Mesmo dentro da mesma subfamília, os tamanhos e as quantidades de éxons e íntrons variaram entre os genes

Tabela 1 - Características dos genes codificadores das proteínas Snf2 de pinhão-mansão (*JcSnf2*), incluindo nome, *locus* id, localização genômica, subfamília, tamanhos dos genes, da CDS e número de íntrons.

Nome	Gene id	Localização genômica	subfamília	gene (nt)	cds (nt)	n° de íntrons
<i>JcSNF2.6</i>	LOC105631838	50752-70638	Snf2	19887	10230	33
<i>JcSNF2.12</i>	LOC105641064	261298-283913	Mi-2	22616	7047	9
<i>JcSNF2.24</i>	LOC105640996	2715771-2731452	Swr1	15682	6147	19
<i>JcSNF2.22</i>	LOC105632541	2248410-2267955	Mot1	19546	5895	27
<i>JcSNF2.13</i>	LOC105635949	292349-312843	Chd1	20495	5286	29
<i>JcSNF2.2</i>	SNF2	13657-28731	SHPRH	15075	5052	18
<i>JcSNF2.7</i>	LOC105647763	50821-66318	ATRX	15498	4560	24
<i>JcSNF2.21</i>	LOC105637572	1787300-1805252	Mi-2	5665	4467	9
<i>JcSNF2.11</i>	LOC105631220	143736-150416	SHPRH	6681	4188	7
<i>JcSNF2.4</i>	LOC105649495	36469-41598	Ris1	5130	4083	6
<i>JcSNF2.15</i>	LOC105645644	697347-704072	DRD1	6726	3756	4
<i>JcSNF2.19</i>	LOC105640590	1300383-1311814	SMARCAL	11432	3684	19
<i>JcSNF2.16</i>	LOC105644519	772876-779914	ERCC6	7039	3684	6
<i>JcSNF2.14</i>	LOC105631950	302424-310954	Snf2	8531	3369	11
<i>JcSNF2.18</i>	LOC105640183	1053310-1066672	ERCC6	13363	3345	21
<i>JcSNF2.26</i>	LOC105636961	6737550-6744041	DRD1	6492	3213	2
<i>JcSNF2.10</i>	LOC105631165	96536-106346	Ris1	9811	3201	10
<i>JcSNF2.5</i>	LOC105631817	37569-45491	Iswi	7923	3198	24
<i>JcSNF2.25</i>	LOC105637347	5634387-5642778	Rad5/16	8392	3111	20
<i>JcSNF2.1</i>	LOC105629844	8339-21149	Rad54	12811	2808	20
<i>JcSNF2.8</i>	LOC105630119	53606-60165	DRD1	6560	2745	4
<i>JcSNF2.23</i>	LOC105646917	2624595-2629510	ERCC6	4916	2634	7
<i>JcSNF2.17</i>	LOC105634476	817494-825167	Rad5/16	7674	2616	2
<i>JcSNF2.9</i>	LOC105643213	92120-99866	ALC1	7747	2604	15
<i>JcSNF2.3</i>	LOC105633941	15300-24750	Lsh	9451	2259	14
<i>JcSNF2.20</i>	LOC105643644	1420256-1430482	SMARCAL	10227	2085	23

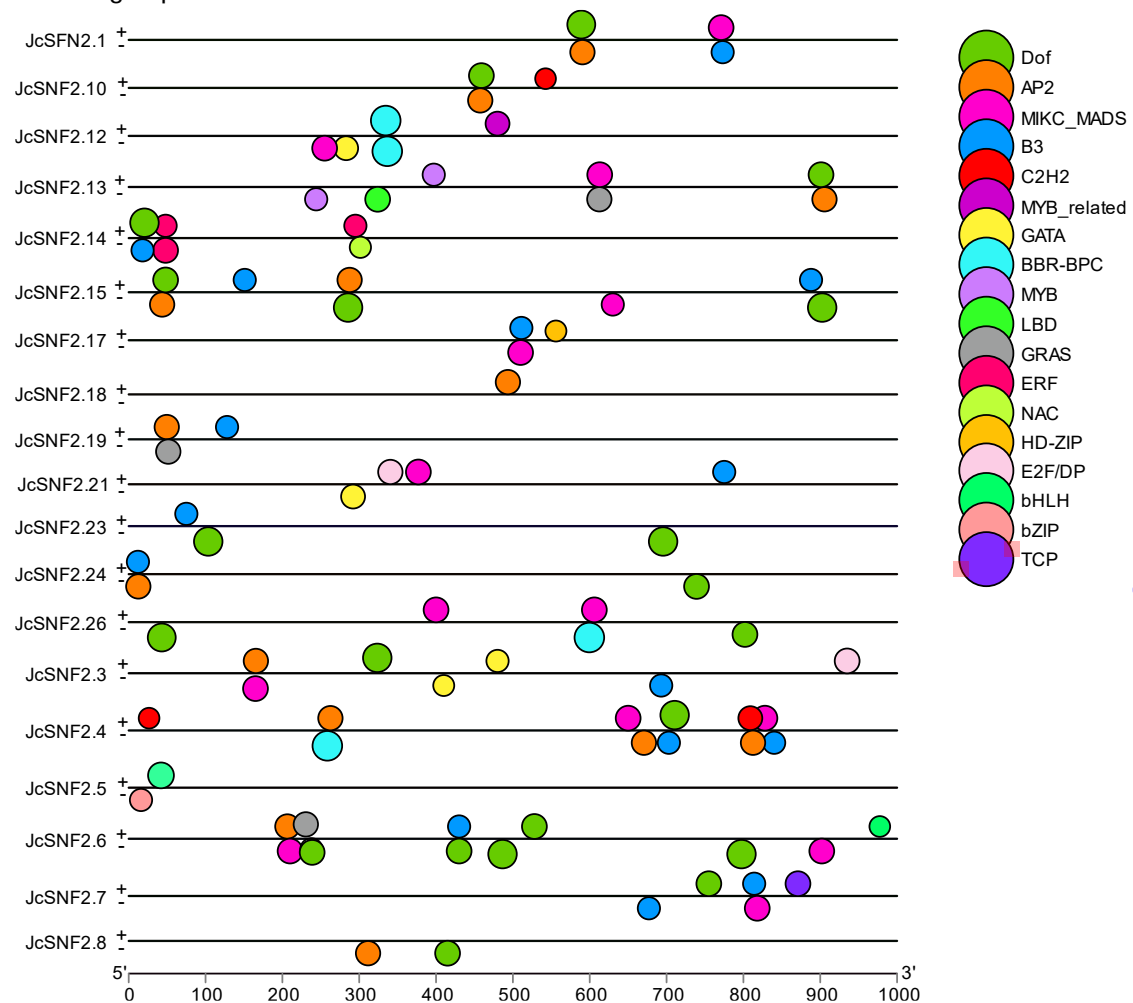
Figura 1 - Estrutura éxon-íntron dos genes *JcSnf2*. A representação gráfica foi obtida a partir do GSDS (Gene Structure Display).



Análise da região promotora dos genes *JcSnf2*

As potenciais regiões promotoras (1.0 kb *upstream* ao sítio de iniciação da transcrição; base +1) dos genes *JcSnf2* foram analisadas quanto à presença de elementos cis-regulatórios. A partir da análise foi possível prever sítios de ligação de fatores de transcrição associados com dezoito famílias diferentes (p -value < 0,01; q -value < 0,01) (Figura 2). Dentre essas famílias, as mais representativas foram Dof, AP2, B3 e MIKC_MADS. Os fatores de transcrição das famílias detectadas são conhecidos por sua importância na regulação de processos do metabolismo vegetal, desde crescimento e desenvolvimento até respostas a estresses bióticos e abióticos.

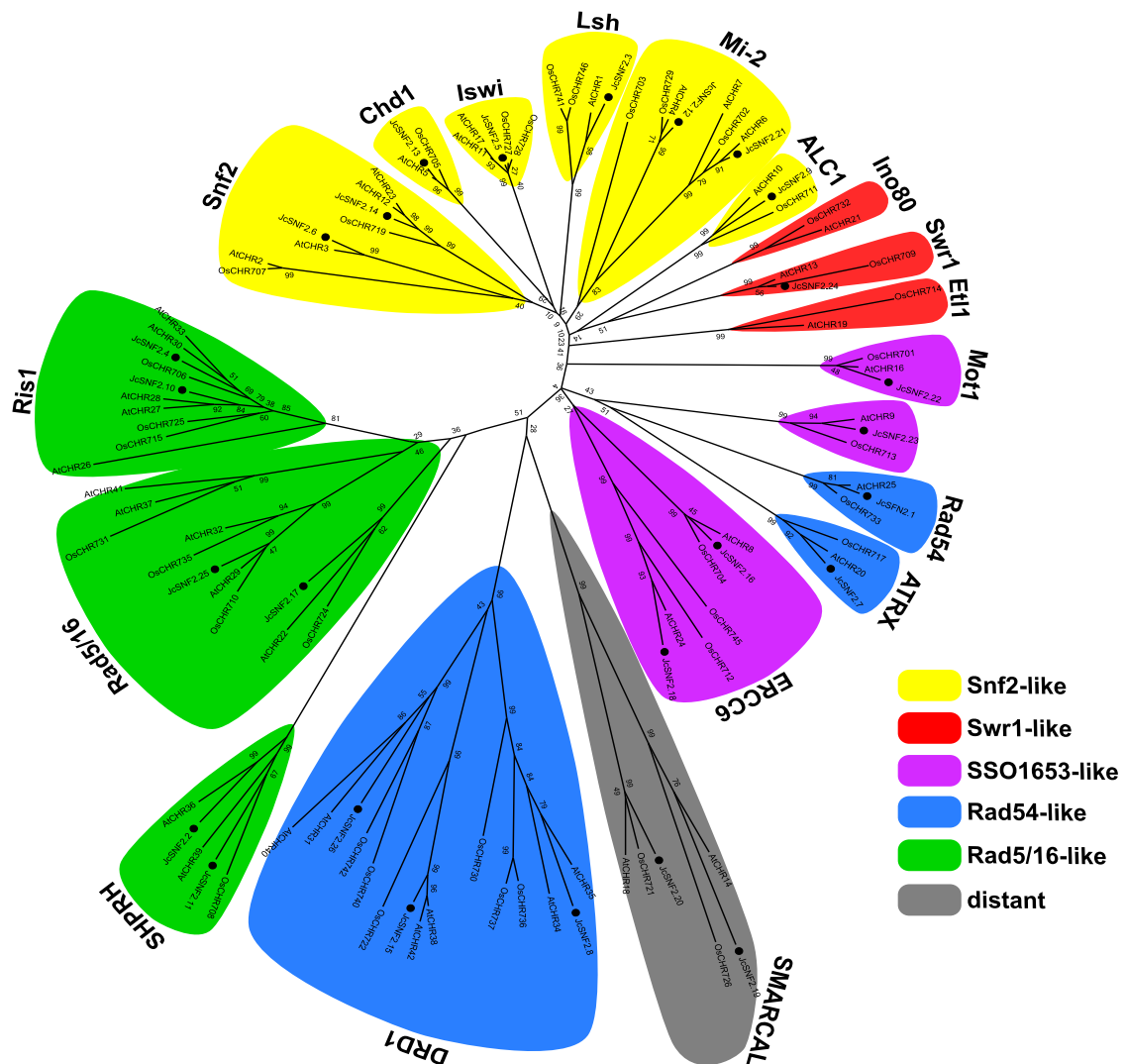
Figura 2 - Regiões promotoras dos genes *JcSnf2* (1.0 kb *upstream* a base +1) e presença de potenciais elementos cis regulatórios (p -value < 0,01; q -value < 0,01) preditos a partir do portal PlantRegMap.



Análise fenética

As proteínas Snf2 são classificadas em seis grupos com base na região ATPase/helicase (Flaus et al., 2006). A árvore fenética permitiu a classificação das *JcSnf2* nas 18 subfamílias incluídas nos seis grandes grupos previamente descritos (Figura 3). Os grupos *Snf2-like* e *Rad5/16-like* abrangeram as maiores quantidades de proteínas, nove e seis, respectivamente (Figura 3). A quantidade de proteínas em cada subfamília variou pouco, com as subfamílias ERCC6, DRD1 e Mi-2 apresentando três proteínas; SHPRH, Ris1, Rad5/16, Snf2 e SMARCAL com duas proteínas cada e Mot1, ATRX, Rad54, Swr1, Lsh, Iwsi, ALC1 e Chd1 com uma proteína cada (Figura 3).

Figura 3 - Árvore fenética da família Snf2 de pinhão-manso. A árvore foi construída pelo método *Neighbor joining* com *bootstrap* de 1.000 replicatas a partir do software MEGA11.



Propriedades e arquitetura dos domínios das proteínas *JcSnf2*

O tamanho das proteínas em aminoácidos variou de 694 até 3.409 resíduos. O peso molecular variou de 77.61 até 260.84 KDa e o ponto isoelétrico de 5.05 até 8.53 (Tabela 2). A análise de predição da localização subcelular revelou todas as proteínas *JcSnf2* localizadas no núcleo e uma proteína predita ao núcleo e mitocôndria (*JcSnf2.22*). Foram detectados 16 domínios nas sequências. Esses domínios apresentam organização típica das helicases da SF2, dois domínios *core* (SNF2_N e

HELICASE_C) acompanhados de domínios auxiliares: RING; PHD; HIRAN; SAWADEE; HNHc; DUF3535; DUF1087; SANT; DUF1086; CHROMO; DUF4208; SnAC; QLQ; AT_hook (Figura 4).

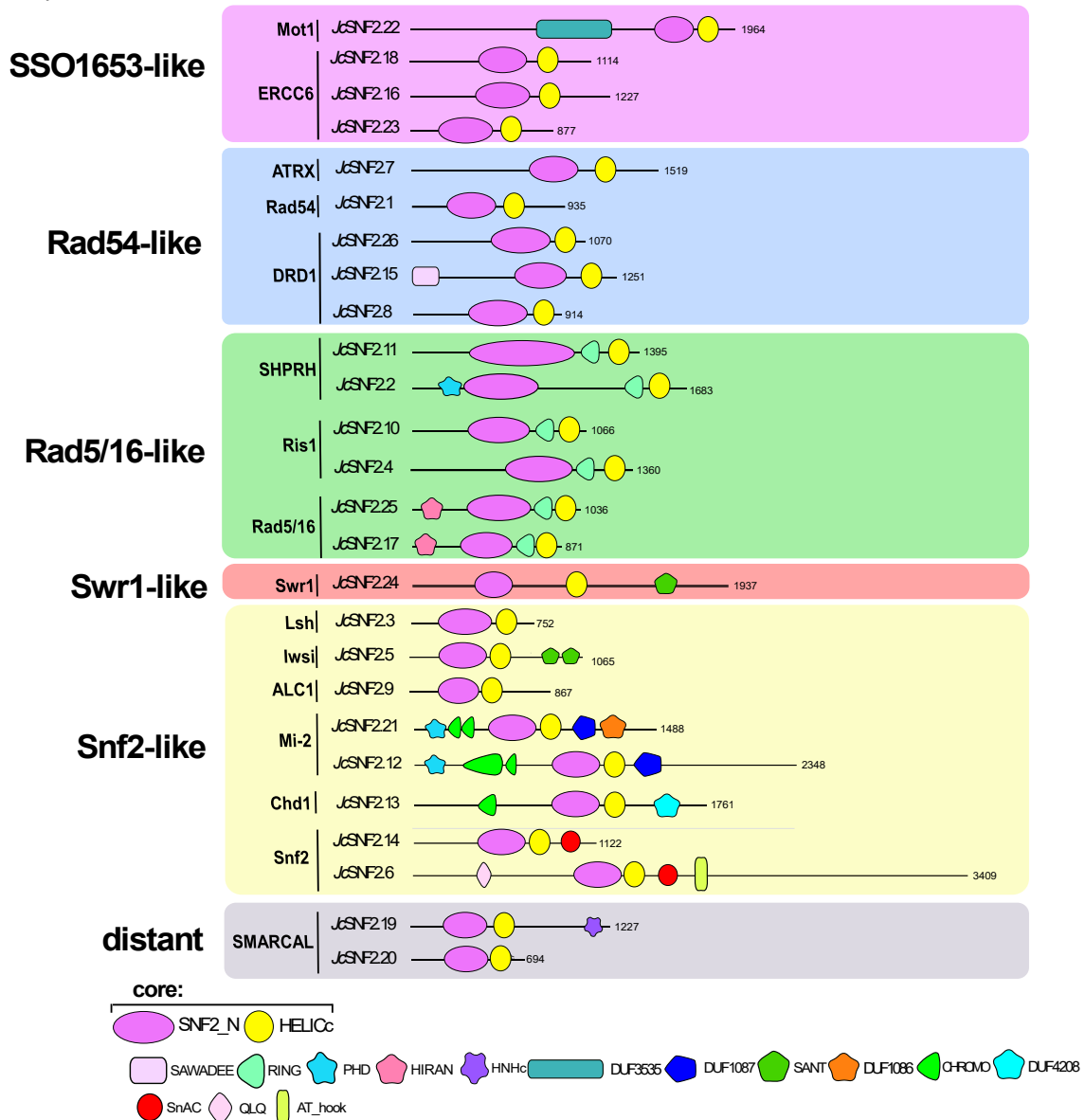
Tabela 2 - Características das proteínas Snf2 de pinhão-mansão (*JcSnf2*), incluindo nome (quantidade de proteínas do *loci*), proteína id, subfamília, tamanhos em aminoácidos (aa) das proteínas primárias, peso molecular (MW), ponto isoelétrico (pI) e localização subcelular.

Nome	Proteína	Subfamília	aa	MW	pI	Localização subcelular
<i>JcSNF2.6</i> (1)	XP_037495642.1	Snf2	3409	129,06	6,71	Núcleo
<i>JcSNF2.12</i> (5)	XP_037493788.1	Mi-2	2348	260,84	7,97	Núcleo
<i>JcSNF2.24</i> (3)	XP_037493417.1	Swr1	1937	220,70	5,11	Núcleo
<i>JcSNF2.22</i> (1)	XP_037492536.1	Mot1	1964	218,41	5,96	Núcleo e Mitocôndria
<i>JcSNF2.13</i> (4)	XP_012074478.1	Chd1	1761	200,52	5,58	Núcleo
<i>JcSNF2.2</i> (1)	XP_012091123.1	SHPRH	1683	190,78	6,32	Núcleo
<i>JcSNF2.7</i> (9)	XP_012089372.1	ATRAX	1519	172,10	5,44	Núcleo
<i>JcSNF2.21</i> (1)	XP_037493511.1	Mi-2	1488	169,69	5,05	Núcleo
<i>JcSNF2.11</i> (3)	XP_012068651.2	SHPRH	1395	156,93	8,52	Núcleo
<i>JcSNF2.4</i> (2)	XP_037492921.1	Ris1	1360	149,42	5,18	Núcleo
<i>JcSNF2.15</i> (1)	XP_012086690.1	DRD1	1251	143,63	8,18	Núcleo
<i>JcSNF2.19</i> (1)	XP_037493316.1	SMARCAL	1227	138,63	7,01	Núcleo
<i>JcSNF2.16</i> (2)	XP_012085290.1	ERCC6	1227	137,86	6,02	Núcleo
<i>JcSNF2.14</i> (3)	XP_012069573.1	Snf2	1122	129,06	6,71	Núcleo
<i>JcSNF2.18</i> (1)	XP_012079819.1	ERCC6	1114	125,74	6,10	Núcleo
<i>JcSNF2.26</i> (2)	XP_012075760.1	DRD1	1070	120,42	5,45	Núcleo
<i>JcSNF2.10</i> (3)	XP_012068570.1	Ris1	1066	118,35	6,76	Núcleo
<i>JcSNF2.5</i> (1)	XP_012069394.1	lswi	1065	123,78	5,93	Núcleo
<i>JcSNF2.25</i> (4)	XP_012076178.1	Rad5/16	1036	116,08	7,22	Núcleo
<i>JcSNF2.1</i> (1)	XP_012066883.1	Rad54	935	103,95	6,14	Núcleo
<i>JcSNF2.8</i> (2)	XP_012067210.1	DRD1	914	103,03	6,27	Núcleo
<i>JcSNF2.23</i> (1)	XP_012088256.1	ERCC6	877	99,20	8,21	Núcleo
<i>JcSNF2.17</i> (1)	XP_012072730.1	Rad5/16	871	97,28	8,52	Núcleo
<i>JcSNF2.9</i> (1)	XP_012083688.1	ALC1	867	98,08	6,47	Núcleo
<i>JcSNF2.3</i> (2)	XP_012072038.1	Lsh	752	86,17	5,84	Núcleo
<i>JcSNF2.20</i> (4)	XP_037494778.1	SMARCAL	694	77,61	8,53	Núcleo

Alguns grupos e subfamílias exibiram diferenças quanto à arquitetura e presença de domínios. A presença de domínios auxiliares foi marcante nas subfamílias do grupo *Snf2-like* (Lsh; lswi; ALC1; Mi-2; Chd1; Snf2) em comparação com as demais subfamílias (Figura 4). O domínio HIRAN foi único das proteínas classificadas como Rad5/16 (Figura 4). O domínio RING esteve restrito as subfamílias do grupo Rad5/16-

like: SHPRH, Ris1 e Rad5/16 (Figura 3). O domínio SANT foi específico de *lws1*, assim como o domínio CHROMO das subfamílias Chd1 e Mi-2 e assim por diante (Figura 4).

Figura 4 - Organização dos domínios das proteínas *JcSnf2*. A escala representa o tamanho em aminoácidos das proteínas. Os símbolos representam os diferentes domínios presentes nas sequências.

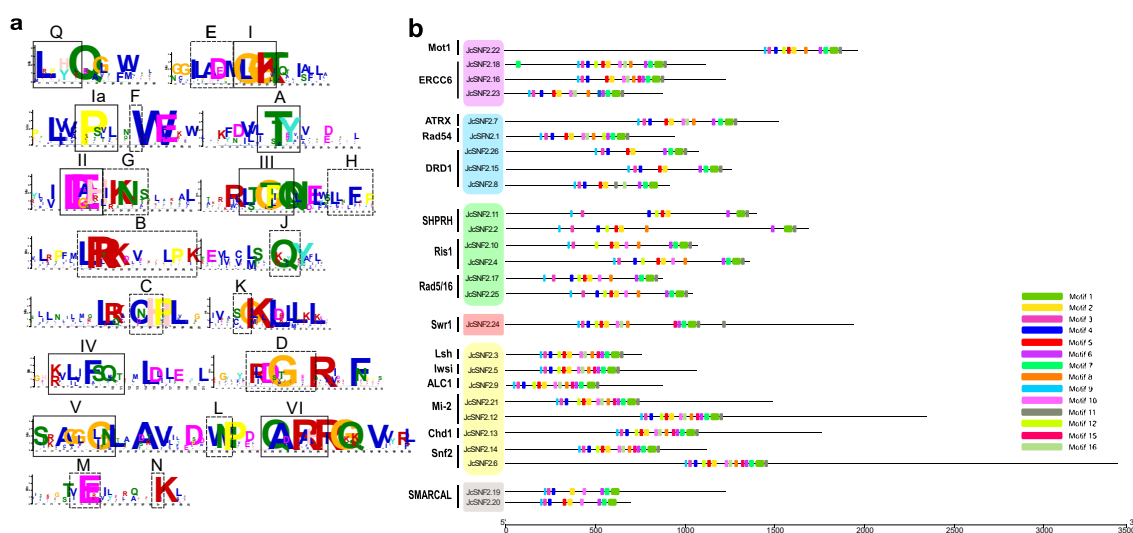


Motivos conservados nas proteínas *JcSnf2*

Todos os oito motivos característicos: Q, I, Ia, II, III, IV, V e VI, além dos blocos menos conservados: E-N estiveram presentes nas proteínas *JcSnf2* (Figura 5). As sequências das proteínas *JcSnf2* mostraram pouca variação na composição de

motivos conservados, com os blocos menos conservados mostrando alguns padrões relacionados com subfamílias específicas. Ainda assim, nas sequências do grupo *Snf2-like* foram detectados todos os motivos e blocos conservados que definem a família (Figura 5). As sequências da subfamília SMARCAL exibiram os motivos clássicos das helicases, enquanto que as regiões conservadas de todo o bloco entre os motivos III e IV (H, B, J, C e K) estiveram ausentes (Figura 5). As demais famílias exibiram pequenas divergências, como a falta dos blocos J e K nas subfamílias SHPRH, Ris1 e Rad5/16, bem como a ausência dos blocos C, K e A nas sequências da subfamília DRD1 (Figura 5). Proteínas *JcSnf2* possuem as características estruturais esperadas e se diversificam através de alterações na ocorrência dos blocos menos conservados E-N.

Figura 5 - Motivos conservados detectados nas sequências das proteínas *JcSnf2* A. e sua distribuição de acordo com a classificação em subfamílias B. Os motivos são nomeados: Q, I, Ia, II, III, IV, V e VI; Os blocos menos conservados presentes entre as regiões dos motivos: E, F, A, G, H, B, J, C, K, D, L, M e N. Os motivos foram detectados pelo programa MEME e a visualização através do software TBtools.



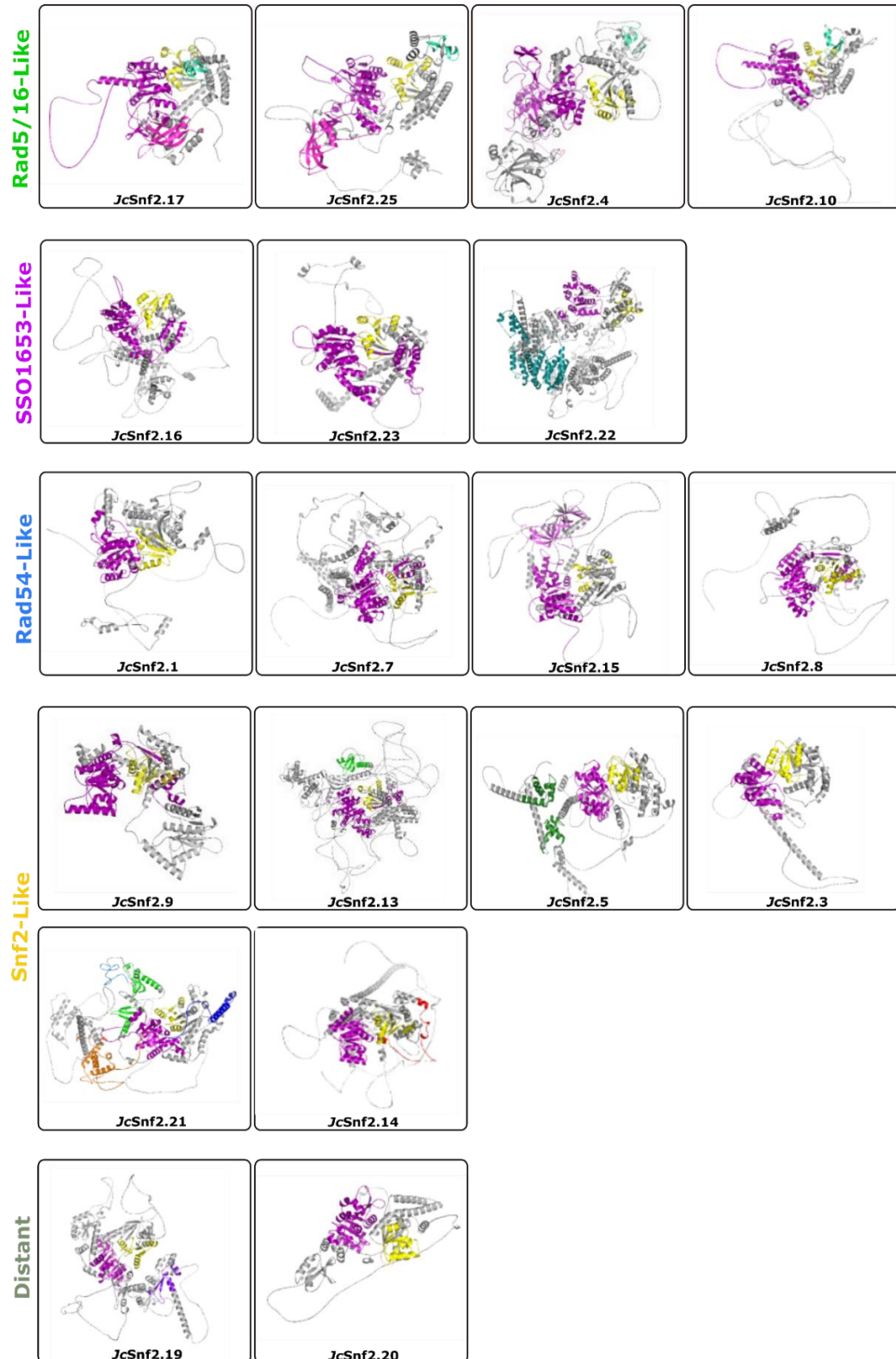
Estruturas secundárias e modelagem tridimensional de *JcSnf2*

Estruturas secundárias, preditas das sequências peptídicas de *JcSnf2* através da ferramenta SOPMA (Tabela suplementar 1), indicaram uma composição geral

formada por 27,35 – 59,98% de α -helice, por 6,62 – 15,38% de β -sheet, por 2,57 – 5,73% de β -turn; e por 29,26 – 60,52% de *random coil*. Mais especificamente, membros do grupo Rad5/16-Like, em média, 37,23% de α -helice e 12,46 de β -sheet; no grupo SSO1653-Like, em média 46,76% de α -helice e 10,00% de β -sheet; no grupo Rad54-Like, média de 38,59% de α -helice e 12,69 de β -sheet; no grupo Snf2-Like, média de 44,40% de α -helice e 10,24% de β -sheet; no grupo *Distant*, 43,35% de α -helice e 12,37% de β -sheet. As proteínas Snf2 de pinhão-mansão exibem estruturas similares quanto à ocorrência de estruturas β -sheet, com algumas diferenças nas quantidades de α -helice.

As proteínas JcSnf2 (26, representativas dos *loci*) foram submetidas ao Swiss-model para modelagem tridimensional. De 19 modelos com GMQE (*Global Model Quality Estimate*) a partir de 0.60, todos mostraram cobertura > 90%. As estruturas apresentaram 80,85 – 93,32% dos resíduos nas regiões permitidas em gráficos de Ramachandran e o QMEANDisCo global score variou de 0,42 – 0,60 (Tabela suplementar 2; Figura suplementar 1). Os modelos obtidos, realçando os domínios conservados, estão ilustrados na Figura 6. Os modelos mostram a estrutura típica da superfamília SF2, dois domínios globulares centrais (*core*) constituídos por β -sheets envoltas por α -hélices (Figura 6), demonstrando que são aparentemente compatíveis com o desempenho de suas atividades.

Figura 6 - Modelos 3D obtidos a partir do Swissmodel para as proteínas JcSnf2. Os domínios conservados estão realçados em cada modelo; Domínios *core* SNF2 (N-terminal) e HELICc (C-terminal) estão representados em roxo e amarelo, respectivamente. Domínios auxiliares incluem RING (*greencyan*); HIRAN (*lighmagenta*); DUF3535 (*deepteal*); SAWADEE (*pink*); CHROMO (*green*); SANT (*green forest*); DUF1086 (*orange*); DUF1087 (*blue*); SnAC (*red*); HNHc (*purpleblue*).



Análise do perfil de expressão de *JcSnf2* em resposta à salinidade e validação por qPCR

Alinhamentos BLASTx (e-value *cut-off* e^{-10}) entre as proteínas *JcSnf2* e os transcritos RNA-Seq, permitiram identificar um total de 85 transcritos *JcSnf2* associados com 33 proteínas e 26 genes *JcSnf2*. Esses transcritos tiveram seus domínios confirmados pelas ferramentas CDD e SMART. No acesso *Jc171* (menos tolerante) foram detectados 45 transcritos, destes 35 foram DEs (p -value < 0.0001 ; FDR < 0.005) e 11 não significativos, considerando o estímulo aplicado (150 mM de NaCl por três horas). Para o acesso *Jc183* (mais tolerante) nenhum transcrito *JcSnf2* foi diferencialmente expresso, considerando o *cut-off* e o estímulo aplicados. As informações sobre os transcritos do acesso *Jc171* estão descritas na tabela suplementar 5. Dos transcritos *JcSnf2* DEs, todos os 35 foram DR ($FC \leq 1$; modulação negativa da expressão) (Figura 7).

Dos pares de *primers* propostos para os transcritos *JcSnf2* para os ensaios de qPCR, e que amplificaram amostras de cDNAs do estudo (Tabela suplementar 3), oito pares mostraram valores adequados de eficiência de amplificação (E), slope (s) e coeficiente de correlação (R) (Figura suplementar 2; Tabela 3). Curvas de *melting* mostraram as especificidades de *amplicons* para esses candidatos (Figura suplementar 3). Para quatro desses candidatos, os resultados de qPCR confirmaram a expressão DR *in silico*, enquanto que para outros dois candidatos os resultados de qPCR mostraram expressão oposta, mostrando haver indução (Tabela 3; Figura 8).

Figura 7 - Clusterização hierárquica baseada nos valores de Log2FC dos transcritos RNA-Seq *JcSnf2* diferencialmente expressos (p-value < 0.0001; FDR < 0.005) nas raízes do acesso *Jc171* submetido a estímulo salino (150 mM de NaCl por três horas), bem como, respectiva modulação no acesso *Jc183*.

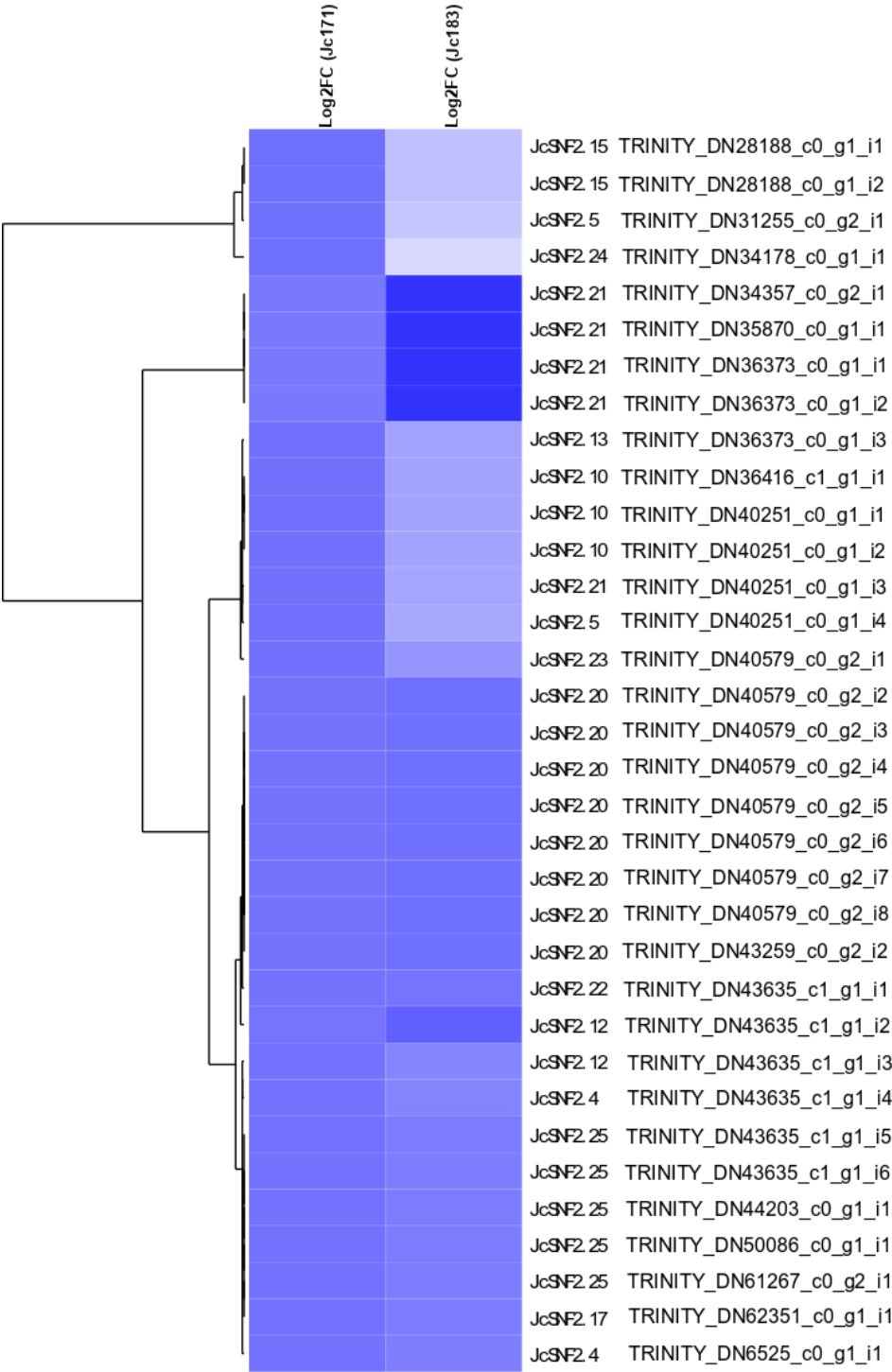
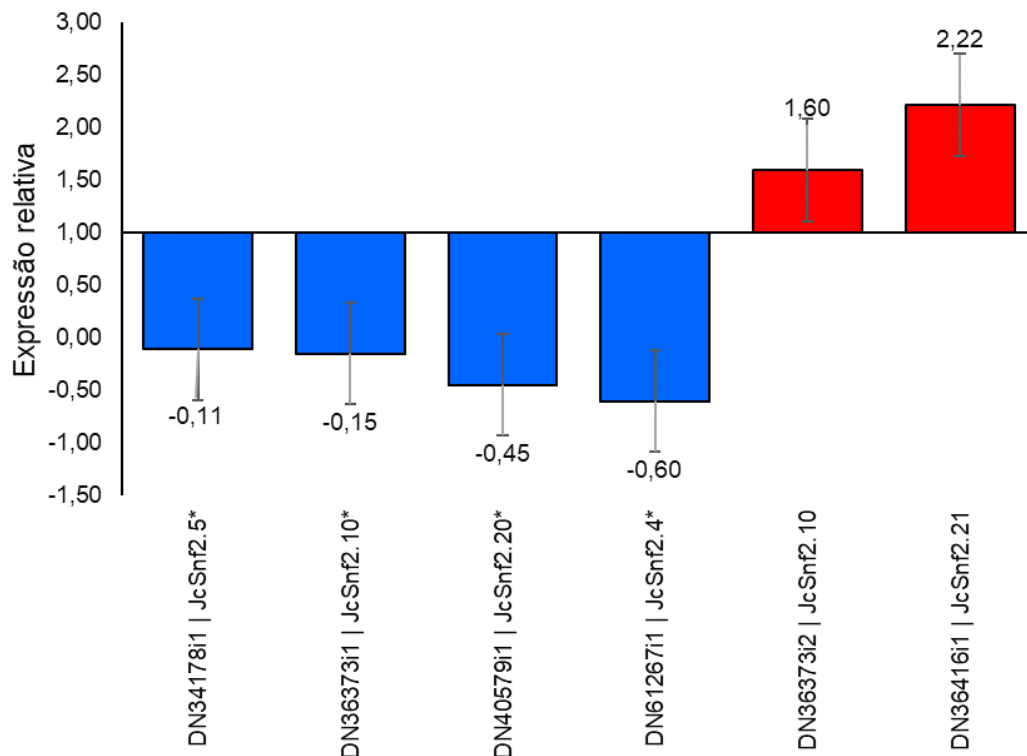


Tabela 3 - Expressão gênica relativa dos transcritos RNA-Seq *JcSnf2* expressos no acesso *Jc171* de pinhão-mansão após exposição a 150 mM de NaCl por três horas. Dados foram obtidos a partir do software REST© (v.2.0.13). DR: *Downregulation*; UR: *Upregulation*. * Genes de referência.

Transcrito	Nome	Eficiência	Std error	95% CI	Expressão	Resultado
DN34178i1	<i>JcSnf2.5</i>	108.10	0,007 - 0,751	0,003 - 1,396	-0,11	DR
DN36373i1	<i>JcSnf2.10</i>	98.32	0,030 - 0,575	0,002 - 1,671	-0,15	DR
DN36416i1	<i>JcSnf2.21</i>	91.94	1,017 - 6,299	0,617 - 8,288	2,22	UR
	Actina*	0.90			1,60	
	β -Tubulina*	0.96			0,63	
DN40579i1	<i>JcSnf2.20</i>	96.92	0,145 - 1,257	0,062 - 3,123	-0,45	DR
DN61267i1	<i>JcSnf2.4</i>	101.54	0,343 - 1,070	0,205 - 1,293	-0,60	DR
	Actina*	0.90			1,05	
	β -Tubulina*	0.96			0,96	
DN36373i2	<i>JcSnf2.10</i>	98.00	1,161 - 2,242	0,864 - 2,930	1,60	UR
	Actina*	0.90			1,10	
	β -Tubulina*	0.96			0,91	

Figura 8 - Expressão gênica relativa calculada a partir de dados qPCR de transcritos *Snf2* (*JcSnf2*) expressos em raízes do acesso *Jc171* pinhão-mansão submetido a estímulo de salinidade (150 mM de NaCl por três horas). O eixo Y representa valores de expressão relativa e o eixo X os transcritos alvo; '*': alvos com expressão validada. *Downregulation* (DR) é indicada em azul, *upregulation* (UR) em vermelho.



Redes PPI para os alvos *JcSnf2* expressos em resposta à salinidade

Na rede PPI, foram observados dois clusteres distintos e 201 termos GO enriquecidos (p -value $< 1.0e-16$). Termos GO compreendem aqueles associados com regulação epigenética da expressão gênica tais como modificação de histona (GO:0016570), histona *exchange* (GO:0043486), acetilação de histona (GO:0016573), regulação da modificação de histona (GO:0031056), organização do nucleossomo (GO:0034728) dentre outros. Além desses termos, também estiveram presentes reparo por excisão de nucleotídeo acoplado à transcrição (GO:0006283), resposta celular a estímulo (GO:0051716), resposta celular a estresse (GO:0033554) e resposta a estresse (GO:0006950) (Figura 9). Os dois clusteres observados compreendem parceiros funcionais de *JcSnf2* que atuam na transcrição pela RNA polimerase II e remodelamento da cromatina (Figura 10). Além disso, de acordo com a rede, esses dois grandes processos estão associados com o reparo do DNA, sendo possível observar proteínas como a DNA helicase RuvB (RIN1) e a endonuclease UVH3 (Figura 10). Uma fosfatidil Inositol quinase (F4JPL2) também foi detectada como parceiro funcional de *JcSnf2* (Figura 10).

Ainda, em redes PPI individuais (Figura 10B, C e D) construídas a partir dos transcritos diferencialmente expressos através dos dados de qPCR, foi possível observar os candidatos *JcSnf2* interagindo diretamente com parceiros funcionais específicos que incluem proteínas ARPs (*Actin-Related Protein*) (ex.: *ARP6*), remodeladores da cromatina (SYD), proteínas HMGB (ex.: *HMBG10* e *HMBG9*), proteínas contendo a repetição WD-40 e um fator de transcrição NF-Y (*NF-YB11*). Esses resultados confirmam as informações já descritas para a família, e indicam que as *JcSnf2* fazem parte do componente epigenético das respostas de *J. curcas* a estresses

Figura 9 - Termos GO enriquecidos (p -value $< 1.0e-16$) associados às redes de interação PPI obtidas considerando os candidatos *JcSnf2* expressos nas bibliotecas RNA-Seq obtidas em ensaio de salinidade (150 mM de NaCl por 3 horas). Os termos representam A. processo biológico; B. função molecular; C. componente celular.

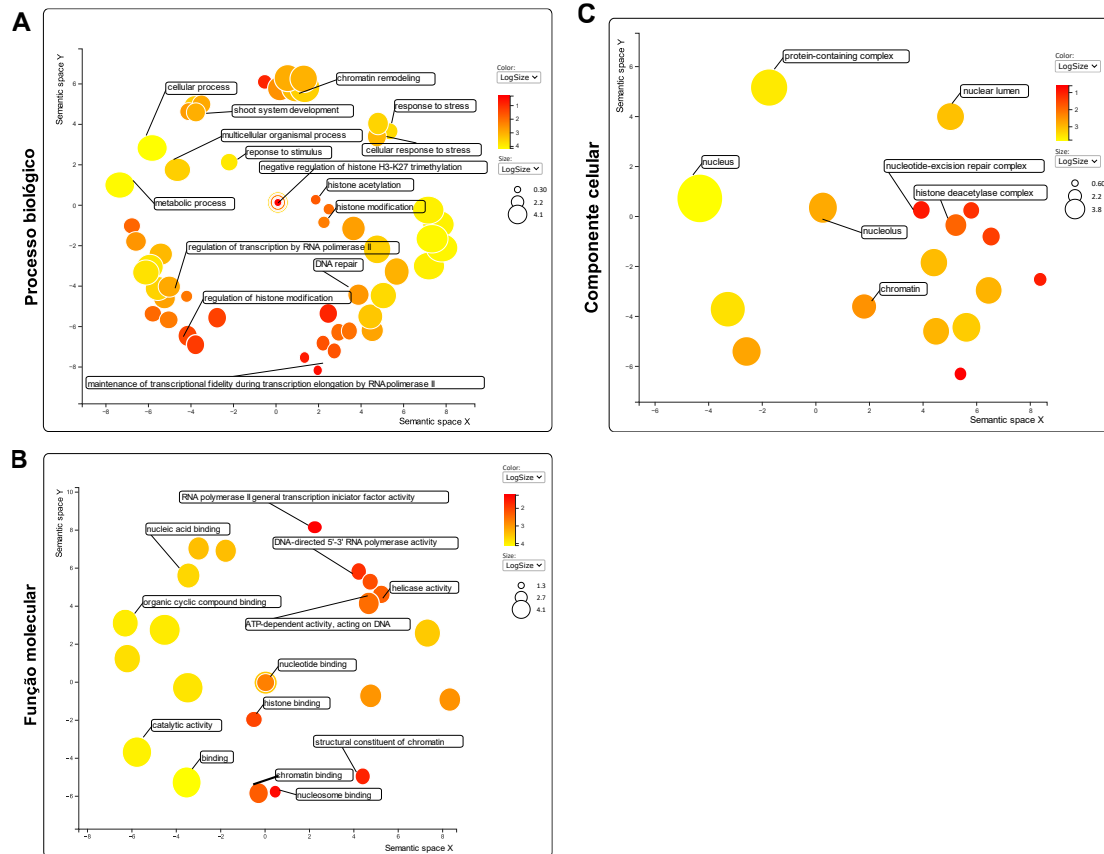
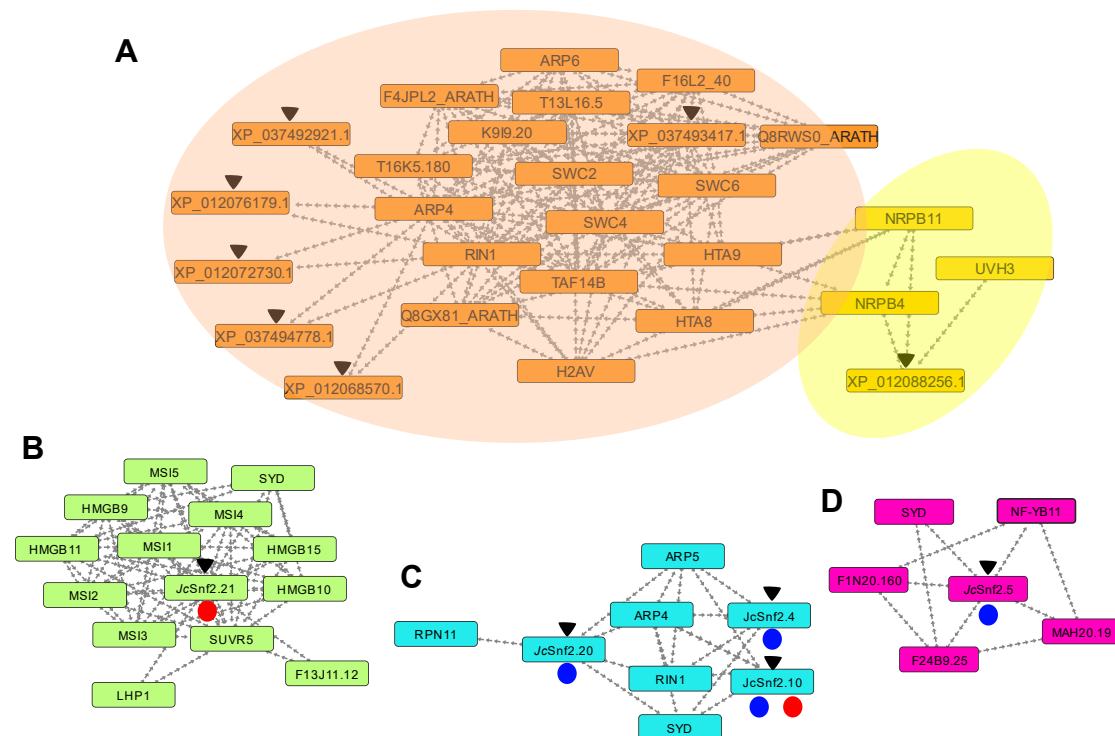


Figura 10 - Redes de interação proteína-proteína (PPI) construídas através do STRING utilizando *Arabidopsis* como modelo (*high score* > 0.7), considerando A. todo o conjunto de transcritos *JcSnf2* diferencialmente expressos nas bibliotecas RNA-Seq obtidas a partir de ensaio de salinidade (150 mM de NaCl) evidenciando dois *clusters* distintos que correspondem a remodelamento da cromatina (laranja) e reparo do DNA (amarelo); em B. C. e D. redes individuais considerando transcritos *JcSnf2* diferencialmente expressos a partir dos dados de qPCR.



Discussão

Como organismos sésseis, as plantas possuem grande habilidade de adaptar sua expressão gênica e manter a homeostase celular para garantir sua sobrevivência, principalmente quando expostas a estresses ambientais. Mecanismos epigenéticos tais como metilação do DNA e modificações em histonas fazem parte das respostas de plantas a estresses (Asensi-Fabado et al., 2017; Chang et al., 2020). Neste trabalho, a família Snf2 de *Jatropha curcas* L. (*JcSnf2*) foi investigada quanto a estruturas éxon-intron, regiões promotoras, classificação, arquitetura de domínios e

motivos conservados além de resposta a salinidade a partir de bibliotecas RNA-Seq obtidas de raízes de acessos contrastantes ao estresse.

A família Snf2 já foi predita e analisada em *Arabidopsis thaliana* (41; Knizewski et al., 2008), *Oryza sativa* (48; Hu et al., 2013), *Solanum lycopersicum* (45; Zhang et al., 2019), *Hordeum vulgare* (38; Chen et al., 2022) *Glicine max* (66; Wang et al., 2023) e *Triticum aestivum* (112; Fahad et al., 2023). As variações observadas quanto a tamanho de proteínas, ponto isoelétrico e peso molecular estão de acordo com o reportado por esses autores. A maioria das proteínas JcSnf2 tiveram localização subcelular predita para o núcleo, condizendo com as atividades exercidas no remodelamento da cromatina, transcrição e reparo do DNA. Ainda, a proteína JcSnf2.22 (XP_037492536.1) foi predita para dois compartimentos, núcleo e mitocôndria. Não há relatos de proteínas Snf2 na mitocôndria, o que pode indicar nova função dentro de *J. curcas*.

Candidatos a elementos cis regulatórios foram abundantes nas potenciais regiões promotoras dos genes *JcSnf2*. Esses elementos cis puderam ser associados com 18 famílias de fatores de transcrição diferentes (p -value < 0,01; q -value < 0,010). Dentre essas famílias estão Dof, AP2, MIKC_MADS, B3, C2H2 e MYB. A regulação da expressão gênica mediada por fatores de transcrição (FTs) é fundamental para garantir adaptabilidade das plantas a estresses ambientais. Por exemplo, plantas de *Arabidopsis* superexpressando o gene de milho *ZmMYB3R* foram mais tolerantes a seca e a salinidade, apresentando maiores atividades de SOD, CAT e POD além de apresentarem boa regulação da abertura estomatal (Wu et al., 2018). O gene AP2 de tabaco *NtERF172* conferiu tolerância a seca nas plantas que o superexpressavam (Zhao et al., 2020). As demais famílias de fatores de transcrição detectadas também estão diretamente associadas a importantes aspectos da resposta a estresse, tais

como indução de componentes da via de síntese de oligossacarídeos da família rafinose (RFOs) pelo FT *Dof17d* (Wang et al., 2021), ativação da expressão de genes DRE/CRT pelo FT *StERF94* (Charfeddine et al., 2019) e regulação da acumulação de isoflavonas pelo FT C2H2 *GmZFP7* (Feng et al., 2023). Portanto, esses resultados indicam que os genes *JcSnf2* podem responder a vias de defesa a estresses associadas com esses FTs.

Estruturas gênicas constituem um aspecto importante da evolução dos genomas de eucariotos, com muitas características sendo linhagem específicas (Li et al., 2017). No geral, os genes *JcSnf2* apresentam grandes variações nos tamanhos de íntrons e éxons, essa característica é associada com diferentes tipos de *splicing* alternativo (Kim et al., 2007). Estruturas éxon-íntron dos genes *JcSnf2* mostram poucos padrões relacionados as subfamílias, o que também ocorre em *S. lycopersicum* (Zhang et al., 2019), *H. vulgare* (Chen et al., 2022) e soja (Wang et al., 2023). Assim como detectado em *J. curcas*, os genes do grupo *Snf2-like* nessas espécies também apresentam estruturas gênicas mais complexas, como maior número de íntrons e maiores diferenças nos tamanhos de éxons e íntrons. Igualmente, os genes das subfamílias DRD1 incluíram os genes com menores quantidades de íntrons. Estes resultados corroboram com a maior presença de domínios auxiliares na subfamília *Snf2*, visto que genes com maior número de íntrons oferecem maiores oportunidades para a maquinaria de *splicing* gerar diferentes isoformas (Jo e Choi, 2015). Membros de *Snf2-like*, provavelmente, estão envolvidas em funções mais diversas, comparado com as demais subfamílias que apresentam estruturas mais simples.

As proteínas *JcSnf2* compartilham o *core* catalítico comum relativo à família, mas cada subfamília apresenta algumas especificidades na ocorrência de domínios auxiliares e presença de blocos conservados entre os motivos. Além da presença dos

oito “motivos helicase”, as proteínas Snf2 apresentam uma série de regiões “menos conservadas” posicionadas entre esses motivos (nomeadas [E-N]). Essa característica estrutural é decorrente da natureza de atuação da família, visto que são componentes de complexos multiproteicos, esses motivos adicionais funcionam na adaptação das Snf2, no estabelecimento de diferentes interações com diferentes propósitos (Flaus et al., 2006). Isso se reflete na arquitetura dos domínios presentes em *JcSnf2*, visto que as subfamílias apresentam domínios exclusivos, como por exemplo, o domínio RING (*Really Interesting New Gene*) nas subfamílias do grupo Rad5/16-like e os domínios SANT (Swi3, Ada2, N-Cor e TFIIIB) e CHROMO (*CHR*romatin *O*rganization *M*Odifier) nas subfamílias do grupo Snf2-like.

Esses domínios estão envolvidos em processos tais com degradação de proteínas via proteassomo 26S (RING) (Sun et al., 2019), interação entre caudas de histonas e DNA (SANT) (Boyer et al., 2004), ligação e reconhecimento a histonas metiladas (CHROMO e PHD; *Plant Homeodomain*) (Flanagan et al., 2005; Taverna et al., 2006;), e reconhecimento de regiões danificadas no DNA (HIRAN; HIP116, Rad5p N-terminal) (Korzhnev et al., 2016).

Modelos 3D gerados estão de acordo com o *hallmark* da família, visto que mostram a presença dos dois domínios RecA-like constituídos por β -sheets envoltas por α -helices. Essas estruturas estão de acordo com os modelos já disponibilizados para a família (Xia et al., 2016; Yan et al., 2016). Dados de estruturas secundária e modelagem 3D reforçam que as Snf2 de *J. curcas* são compatíveis com o desempenho de suas atividades de ligação a ácidos nucleicos, hidrólise de NTPs e remodelamento da cromatina.

Alguns trabalhos relatam o envolvimento dos membros da família Snf2 em respostas a estresses abióticos. Hu et al., (2013) verificaram através de qPCR que em

plantas de arroz expostas a seca (sementes secas em papel de filtro; 0, 1, 2 e 4 horas), os genes *CHR720* e *CHR728* tiveram expressão aumentada, enquanto que a expressão de *CHR735* foi diminuída sob seca e também sob salinidade (300 mM de NaCl; 0, 1, 2 e 4 horas). Quando Zhang et al., (2019) analisaram os padrões de expressão dos genes *SICHR* de tomate através de qPCR, os autores observaram que sob tratamentos de ácido salicílico (1 mM; 4 horas) e frio (4 °C; 4 horas), a maioria dos genes foi reprimida, enquanto que sob ABA (50 µM; 4 horas) e sal (200 mM; 4 horas) a maioria foi induzida.

Em plantas de arroz, mais 55% dos genes *Snf2* foram induzidos sob seca (0, 1, 3, 6 e 12 horas) e salinidade (200 mM; 0, 1, 3, 6 e 12 horas), com destaque para os genes (*OsCHR705*, *OsCHR706*, *OsCHR710*, *OsCHR714*, *OsCHR721*, *OsCHR726* e *OsCHR737*). Ainda, plantas mutantes para o gene *OsCHR726* exibiram fenótipo hipersensível a salinidade (Guo et al., 2022). Plantas de *Arabidopsis* mutantes para o gene *PKL* (*PICKLE*) exibiram fenótipo de sensibilidade quando submetidas ao frio (4 °C;) e a salinidade (150 mM de NaCl e 200 mM de manitol; 6 dias) (Yang et al., 2019). O gene *MdRAD5B* regula a tolerância a seca em maçã, esse gene foi visto controlando a expressão de 466 genes de resposta ao estresse por modular a deposição da marca H3K27me3, associada a regulação negativa de genes (Song et al., 2023).

Em *Arabidopsis*, o gene *CH23* se liga a proteína responsiva ao estresse Di19, e os duplo mutantes *chr12 chr23* foram tolerantes a seca (supressão de rega; 18 até 21 dias) (Sun et al., 2022). O gene *ZmCHB101* se associa com a RNA polimerase II e foi descrito regulando a acessibilidade ao DNA em sítios de início da transcrição de genes estresse-responsivos. Plantas de milho transgênicas (RNAi) para esse gene foram sensíveis a seca, salinidade e estresse osmótico (Yu et al., 2018). Assim é possível observar que os padrões de expressão da família variam bastante de acordo com o

estresse aplicado, indicando que esses genes desempenham diferentes funções sob diferentes estímulos ambientais.

Nas bibliotecas RNA-Seq em estudo, todos os transcritos detectados para o acesso *Jc171* (menos tolerante) foram reprimidos sob tratamento de salinidade (150 mM de NaCl por 3 horas). Após o tratamento, o acesso *Jc171* apresentou danos foliares, ex.: clorose, decorrentes da aplicação do tratamento salino, o que não foi observado no acesso *Jc183* (mais tolerante) (Souza et al., 2020). Logo, a família *Snf2* de *J. curcas* exibiu um padrão de resposta à salinidade dependente do genótipo analisado, o que pode auxiliar a elucidar o comportamento do acesso além de prover candidatos para análises funcionais visando obter plantas mais tolerantes ao estresse.

Na rede PPI foram observados dois clusteres associados com a transcrição e com o remodelamento da cromatina, acoplados com reparo do DNA. A regulação da transcrição de genes alvo no nível transcricional é crucial em vias de crescimento e desenvolvimento além de aclimação durante condições estresses em plantas. Durante a transcrição, a via de reparo por excisão de nucleotídeo (*transcription-coupled nucleotide excision repair*; TC-NER) funciona para garantir a fidelidade da transmissão da informação genética, removendo lesões do DNA e mantendo a estabilidade do genoma (Moreno et al., 2023). Termos GO associados com esse processo foram enriquecidos na rede PPI (ex.: GO:0006283; GO:0006289; GO:0001193). Logo, os candidatos *JcSnf2* atuam regulando a dinâmica da cromatina para alterar o acesso ao DNA durante os processos de transcrição e reparo.

Uma fosfatidil inositol quinase (*F4JPL2*) também foi detectada como parceiro funcional de *JcSnf2* na rede PPI. Essas proteínas são componentes importantes durante a aclimação e defesa de plantas, pois orquestram a dinâmica espacial e

temporal de fosfoinosítídeos (PIs), moléculas que atuam na regulação de processos como a divisão celular e tráfego através da membrana (Lin et al., 2022). A ativação de PM H⁺-ATPases e transportadores Na⁺/H⁺, proteínas requeridas para manutenção da homeostase iônica durante condições de salinidade, foi associada a adição de fosfatidis PI4P em plantas de *Arabidopsis* (Yang et al., 2021). A fosfatidil Inositol 3-quinase (PI3K) promove a autofagia em células de *Arabidopsis* (Gomez et al., 2022), a autofagia é uma rota de degradação para reciclagem de material celular de células danificadas, de fato plantas incapazes de realizar esse processo apresentam baixa sobrevivência a estresses abióticos (Chen et al., 2021). Plantas de *Arabidopsis* triplo mutantes para genes codificadores de fosfatidil inositol quinases (*pip5k7pip5k8pip5k9*) apresentam crescimento prejudicado de raízes sob condições de estresses hiperosmótico (manitol e NaCl) (Kuroda et al., 2021).

Nas redes individuais, foi possível observar os candidatos *JcSnf2* interagindo diretamente com parceiros funcionais específicos que incluem proteínas ARPs (*Actin-Related Protein*) (ex.: *ARP6*), remodeladores da cromatina (*SYD*), proteínas HMGB (ex.: *HMBG10* e *HMBG9*), proteínas contendo a repetição WD-40 e um fator de transcrição NF-Y (*NF-YB11*).

As proteínas *JcSnf2.4*, *JcSnf2.10* e *JcSnf2.20* interagem diretamente com proteínas que realizam a troca de variantes de histona. As proteínas responsáveis por esse processo incluem *ARP4* e *ARP6*. Proteínas ARP (*Actin-Related Protein*) participam de muitas funções celulares. No citoplasma, essas proteínas regulam a montagem de filamentos de actina e o tráfego de vesículas. O citoesqueleto é um componente central na sinalização celular durante exposição a estresses (Chun et al., 2021). No núcleo, as ARPs são subunidades de complexos de remodelamento da cromatina, por exemplo a *ARP6* insere a variante H2A.Z aos nucleossomos (Kumar

e Widge, 2010). A expressão ectópica do gene *ARP6* de abacaxi em *Arabidopsis* confere tolerância a salinidade, seca e deficiência nutricional (Jakada et al., 2023), além disso, plantas de *Arabidopsis* mutantes para esse mesmo gene foram menos tolerantes a salinidade (Do et al., 2023). Três isoformas do transcriptoma em estudo associadas com essas proteínas tiveram expressão DR confirmada por qPCR e uma isoforma teve expressão UR.

JcSNF2.5 interage diretamente com um componente do complexo remodelamento da cromatina SYD e com um fator de transcrição NF-Y. Remodeladores SYD que já foram descritos promovendo a expressão de múltiplos genes de defesa das vias de sinalização do ácido jasmônico e etileno, sendo requerido para resistência contra o patógeno *Botrytis cinerea* (Walley et al., 2008). Fatores de transcrição NF-Y (*Nuclear factor Y*) desempenham papéis cruciais nas respostas de plantas a estresses. A superexpressão do gene *CdtNF-YC1* em plantas de *Paspalum* resultou em maior tolerância a seca e a salinidade (Wu et al., 2018). Plantas de *Arabidopsis* superexpressando o gene *GmNFYA5* foram mais tolerantes a seca, essas plantas mostraram com altos níveis de transcritos de genes de resposta ao ABA (ex.: *RD29A*, *DREB1A* e *DREB2A*) e de fatores de transcrição bZIP e WRKY (Ma et al., 2020). Uma isoforma associada com a proteína *JcSnf2.5* teve sua expressão reprimida confirmada por qPCR.

JcSnf2.21 interage diretamente com proteínas HMGB e proteínas contendo a repetição WD40. Os genes dessas famílias também já são descritos quanto suas funções durante exposição de plantas a estresses. A superexpressão de *OsHMGB707* em plantas de arroz conferiu a seca por afetar a expressão de genes, como proteínas *Heat Shock* e fatores de transcrição DREB (Xu et al., 2021). A expressão ectópica de *AtHMGB15* aumentou a resposta ao frio, mutantes *athmgb15* exibiram decréscimo da

expressão de genes responsivos ao frio (Mallik et al., 2020). A expressão heteróloga de uma proteína WD-40 da halófita *Limonium bicolor* em *Arabidopsis thaliana* conferiu tolerância a salinidade, com plantas exibindo maiores conteúdos de prolina e glutathione S-transferase, além de menores quantidades de MDA e Na⁺ (Yuan et al., 2019). Uma isoforma do transcriptoma em estudo associada com *JcSnf2.21* teve expressão UR na qPCR.

Os resultados das redes PPI indicam que as *JcSnf2* podem participar de vias importantes para regulação da expressão gênica durante condições de estresse regulando a compactação e composição de nucleossomos, principalmente através da troca da variante de histona H2A.Z, e assim alterando a acessibilidade ao DNA para as maquinarias de transcrição e de reparo do DNA. Sendo assim, considerando a repressão de parceiros *JcSnf2*, é possível inferir que parte do componente epigenético de *Jc171* não está sendo bem regulado para lidar com a salinidade. Esses resultados serão úteis para futuras caracterizações com o intuito de melhorar o entendimento das respostas de *Jc171* sob estresses.

Conclusões

Neste trabalho, toda a família Snf2 de *J. curcas* (*JcSnf2*) teve seus padrões estruturais e perfil transcricional em resposta à salinidade analisados. As proteínas foram classificadas nas subfamílias descritas, o que permitiu observar que cada subfamília apresenta composições específicas de domínios e motivos conservados. Dentre as subfamílias analisadas, as do grupo Snf2-Like são mais diversas em termos de ocorrência de motivos, presença de domínios auxiliares e estruturas éxon-íntron. Todas as proteínas compartilham o *core* catalítico comum característico e parecem se diversificar através de alterações na ocorrência dos blocos menos conservados (E-N),

dispostos nas regiões entre os motivos detectados. Modelos tridimensionais obtidos remontam ao *hallmark* típico das helicases de SF-II, indicando que essas proteínas são provavelmente compatíveis com suas atividades. Regiões promotoras dos genes exibiram importantes candidatos a elementos cis-regulatórios associados com famílias de fatores de transcrição tais como Dof, MIKC_MADS e AP2, demonstrando que podem participar de rotas de resposta a estresses abióticos. Nas bibliotecas RNA-Seq analisadas, o perfil de expressão de *JcSnf2* foi totalmente reprimido no acesso *Jc171*. A expressão de alguns desses candidatos foi confirmada através de validações por qPCR. Redes PPI evidenciaram que as *JcSnf2* expressas possivelmente atuam remodelando a estrutura da cromatina, principalmente através da troca de variantes de histona, para regular a acessibilidade da cromatina a maquinaria da RNA polimerase II e ao reparo do DNA, que constituem aspectos importantes da resposta a estresse. Porém, a repressão desses candidatos pode ter influenciado no fenótipo de sensibilidade exibido por *Jc171*. Os dados sugerem que as *JcSnf2* podem ter papel significativo nas respostas de plantas a estresses e podem ser úteis para o desenvolvimento de marcadores moleculares funcionais para serem utilizados em programas de melhoramento de plantas. Além disso, os alvos com expressão induzida via qPCR *JcSnf2.10* e *JcSnf.21* podem ser candidatos em futuros ensaios funcionais para o melhor entendimento sobre suas participações nas respostas de plantas a salinidade e também desenvolver novas variedades (ex.: transgenia) com maior tolerância a estresses.

Referências

- Achten WM, Verchot L, Franken YJ, Mathijs E, Singh VP, Aerts R, and Muys B (2008) *Jatropha* bio-diesel production and use. *Biomass and bioenergy*, 32(12), 1063-1084.

- Asensi-Fabado MA, Amtmann A and Perrella G (2017) Plant responses to abiotic stress: The chromatin context of transcriptional regulation. *Biochim Biophys Acta - Gene Regul Mech* 1860:106–122.
- Assaha DVM, Ueda A, Saneoka H, Al-Yahyai R and Yaish MW (2017) The role of Na⁺ and K⁺ transporters in salt stress adaptation in glycophytes. *Front Physiol.* doi: 10.3389/fphys.2017.00509
- Bargsten JW, Folta A, Mlynárová L and Nap JP (2013) Snf2 family gene distribution in higher plant genomes reveals DRD1 expansion and diversification in the tomato genome. *PLoS One.* doi: 10.1371/journal.pone.0081147
- Becker PB and Hörz W (2002) Atp-dependent nucleosome remodeling. *Annu Rev Biochem* 71:247–273.
- Bhadouriya SL, Mehrotra S, Basantani MK, Loake GJ and Mehrotra R (2021) Role of Chromatin Architecture in Plant Stress Responses: An Update. *Front Plant Sci.* doi: 10.3389/fpls.2020.603380
- Bailey TL, Johnson J, Grant CE and Noble WS (2015) The MEME suite. *Nucleic acids research*, 43(W1), W39-W49.
- Boyer LA, Latek RR and Peterson CL (2004) The SANT domain: a unique histone-tail-binding module?. *Nature reviews Molecular cell biology*, 5(2), 158-163.
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL et al. (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622.
- Chang YN, Zhu C, Jiang J, Zhang H, Zhu JK and Duan CG (2020) Epigenetic regulation in plant abiotic stress responses. *J Integr Plant Biol* 62:563–580.
- Charfeddine M, Samet M, Charfeddine S, Bouaziz D and Gargouri Bouzid R (2019) Ectopic Expression of StERF94 Transcription Factor in Potato Plants Improved Resistance to *Fusarium solani* Infection. *Plant Mol Biol Report* 37:450–463.
- Chaudhry S and Sidhu GPS (2022) Climate change regulated abiotic stress mechanisms in plants: a comprehensive review. *Plant Cell Rep.* doi: 10.1007/s00299-021-02759-5
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y and Xia R (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* 13:1194–1202.
- Chen G, Mishina K, Zhu H, Kikuchi S, Sassa H, Oono Y and Komatsuda T (2023) Genome-Wide Analysis of Snf2 Gene Family Reveals Potential Role in Regulation of Spike Development in Barley. *Int J Mol Sci.* doi: 10.3390/ijms24010457
- Chen G, Mishina K, Zhu H, Kikuchi S, Sassa H, Oono Y and Komatsuda T (2023) Genome-Wide Analysis of Snf2 Gene Family Reveals Potential Role in Regulation of Spike Development in Barley. *Int J Mol Sci.* doi: 10.3390/ijms24010457

- Chen H, Dong J and Wang T (2021) Autophagy in Plant Abiotic Stress Management. *Int J Mol Sci* 22.
- Cho EJ, Choi SH, Kim JH, Kim JE, Lee MH, Chung BY, Woo HR and Kim JH (2016) A mutation in plant-specific SWI2/SNF2-Like chromatin-remodeling proteins, DRD1 and DDM1, delays leaf senescence in *Arabidopsis thaliana*. *PLoS One*. doi: 10.1371/journal.pone.0146826
- Chun HJ, Baek D, Jin BJ, Cho HM, Park MS, Lee SH, Lim LH, Cha YJ, Bae DW, Kim ST et al. (2021) Microtubule dynamics plays a vital role in plant adaptation and tolerance to salt stress. *Int J Mol Sci*. doi: 10.3390/ijms22115957
- DeLano WL (2002) Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr*, 40(1), 82-92.
- Do BH, Hiep NT, Lao TD and Nguyen NH (2023) Loss-of-Function Mutation of ACTIN-RELATED PROTEIN 6 (ARP6) Impairs Root Growth in Response to Salinity Stress. *Mol Biotechnol* 65:1414–1420.
- Dos Santos TB, Ribas AF, de Souza SGH, Budzinski, IGF, and Domingues, DS (2022) Physiological responses to drought, salinity, and heat stress in plants: a review. *Stresses*, 2(1), 113-135.
- Eisen JA, Sweder KS and Hanawalt PC (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions.
- Fahad M, Liu C, Shen Y, Sajid M and Wu L (2023) Genomic exploration and functional insights into the diversification of the Snf2 gene family in subgenomes of *Triticum aestivum*. *Crop Des* 100050.
- Feng Y, Zhang S, Li J, Pei R, Tian L, Qi J, Azam M, Agyenim-Boateng KG, Shaibu AS, Liu Y et al. (2023) Dual-function C2H2-type zinc-finger transcription factor GmZFP7 contributes to isoflavone accumulation in soybean. *New Phytol* 237:1794–1809.
- Ferreira OJM, Holanda FSR, Pedrotti A, Vidal Santos LD and Silva-Mann R (2022) Root System of *Jatropha curcas* Provides Resistance and Strength to the Soil. *Communications in Soil Science and Plant Analysis*, 53(22), 2955-2967.
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A et al. (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285.
- Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim Y, Minor W, Rastinejad F and Khorasanizadeh S (2005) Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 438:1181–1185.
- Flaus A, Martin DMA, Barton GJ and Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* 34:2887–2905.
- Flaus A and Owen-Hughes T (2011) Mechanisms for ATP-dependent chromatin remodelling: The means to the end. *FEBS J* 278:3579–3595.

- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD and Bairoch A (2001) Protein Analysis Tools on the ExPASy Server 571–579. In: The Proteomics Protocols Handbook. Protein Identification and Analysis Tools on the ExPASy Server.
- Gentry M and Hennig L (2014) Remodelling chromatin to shape development of plants. *Exp Cell Res* 321:40–46.
- Geourjon C and Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments.
- Gomez RE, Chambaud C, Lupette J, Castets J, Pascal S, Brocard L, Noack L, Jaillais Y, Joubès J and Bernard A (2022) Phosphatidylinositol-4-phosphate controls autophagosome formation in *Arabidopsis thaliana*. *Nat Commun*. doi: 10.1038/s41467-022-32109-2
- Gorbalenya AE, Koonin E V, Donchenko AP and Blinov VM (2002) A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination.
- Guo M, Zhao H, He Z, Zhang W, She Z, Mohammadi MA, Shi C, Yan M, Tian D and Qin Y (2022) Comparative Expression Profiling of Snf2 Family Genes During Reproductive Development and Stress Responses in Rice. *Front Plant Sci*. doi: 10.3389/fpls.2022.910663.
- Halder K, Chaudhuri A, Abidin MZ, Majee M and Datta A (2022) Chromatin-Based Transcriptional Reprogramming in Plants under Abiotic Stresses. doi: 10.3390/plants.
- Hara T, Katoh H, Ogawa D, Kagaya Y, Sato Y, Kitano H, Nagato Y, Ishikawa R, Ono A, Kinoshita T et al. (2015) Rice SNF2 family helicase ENL1 is essential for syncytial endosperm development. *Plant J* 81:1–12.
- Hernández JA (2019) Salinity tolerance in plants: Trends and perspectives. *Int J Mol Sci*. doi: 10.3390/ijms20102408.
- Hu B, Jin J, Guo AY, Zhang H, Luo J and Gao G (2015) GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31:1296–1297.
- Hu Y, Zhu N, Wang X, Yi Q, Zhu D, Lai Y and Zhao Y (2013) Analysis of rice Snf2 family proteins and their potential roles in epigenetic regulation. *Plant Physiol Biochem* 70:33–42.
- Jakada BH, Fakher B, Yao LA, Wang X, Aslam M and Qin Y (2023) Ectopic Expression of Pineapple Actin-Related Protein 6 (AcARP6) Regulates Flowering and Stress Responses in *Arabidopsis*. *J Plant Growth Regul* 42:6461–6473.
- Jalali S, Kancharla N, Yepuri V and Arockiasamy S (2020) Exploitation of Hi-C sequencing for improvement of genome assembly and in-vitro validation of differentially expressing genes in *Jatropha curcas* L. *3 Biotech*. doi: 10.1007/s13205-020-2082-0
- Jo B-S and Choi SS (2015) Introns: The Functional Benefits of Introns in Genomes. *Genomics Inform* 13:112.

- Kamel DA, Farag HA, Amin NK, Zatout A A and Ali RM (2018) Smart utilization of jatropha (*Jatropha curcas* Linnaeus) seeds for biodiesel production: Optimization and mechanism. *Industrial crops and products*, 111, 407-413.
- Kim E, Magen A and Ast G (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* 35:125–131.
- Knizewski L, Ginalski K and Jerzmanowski A (2008) Snf2 proteins in plants: gene silencing and beyond. *Trends Plant Sci* 13:557–565.
- Korzhnev DM, Neculai D, Dhe-Paganon S, Arrowsmith CH and Bezsonova I (2016) Solution NMR structure of the HLTF HIRAN domain: a conserved module in SWI2/SNF2 DNA damage tolerance proteins. *J Biomol NMR* 66:209–219.
- Kumar SV and Wigge PA (2010) H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. *Cell* 140:136–147.
- Kuroda R, Kato M, Tsuge T and Aoyama T (2021) Arabidopsis phosphatidylinositol 4-phosphate 5-kinase genes PIP5K7, PIP5K8, and PIP5K9 are redundantly involved in root growth adaptation to osmotic stress. *Plant J* 106:913–927.
- Letunic I and Bork P (2021) Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296.
- Letunic I, Khedkar S and Bork P (2021) SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Res* 49:D458–D460.
- Li X-Y, Wang C, Nie P-P, Lu X-W, Wang M, Liu W, Yao J, Liu Y-G and Zhang Q-Y (2011) Characterization and expression analysis of the SNF2 family genes in response to phytohormones and abiotic stresses in rice.
- Li Y, Xu Y and Ma Z (2017) Comparative Analysis of the Exon-Intron Structure in Eukaryotic Genomes. *Yangtze Med* 01:50–64.
- Lin A, Du Y and Xiao W (2020) Yeast chromatin remodeling complexes and their roles in transcription. *Curr Genet* 66:657–670.
- Liu M, Pan Z, Yu J, Zhu L, Zhao M, Wang Y, Chen P, Liu C, Hu J, Liu T et al. (2022) Transcriptome-wide characterization, evolutionary analysis, and expression pattern analysis of the NF-Y transcription factor gene family and salt stress response in *Panax ginseng*. *BMC Plant Biol.* doi: 10.1186/s12870-022-03687-6.
- Liu X, Li M, Xia X, Li X and Chen Z (2017) Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature* 544:440–445.
- Lozano-Isla F, Campos MLO, Endres L, Bezerra-Neto E and Pompelli MF (2018) Effects of seed storage time and salt stress on the germination of *Jatropha curcas* L. *Ind Crops Prod* 118:214–224.
- Ma XJ, Yu TF, Li XH, Cao XY, Ma J, Chen J, Zhou Y Bin, Chen M, Ma YZ, Zhang JH et al. (2020) Overexpression of GmNFYA5 confers drought tolerance to transgenic Arabidopsis and soybean plants. *BMC Plant Biol.* doi: 10.1186/s12870-020-02337-z
- Mallik R, Prasad P, Kundu A, Sachdev S, Biswas R, Dutta A, Roy A, Mukhopadhyay J, Bag SK and Chaudhuri S (2020) Identification of genome-wide targets and DNA

- recognition sequence of the Arabidopsis HMG-box protein AtHMGB15 during cold stress response. *Biochim Biophys Acta - Gene Regul Mech.* doi: 10.1016/j.bbagr.2020.194644.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI et al. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226.
- Rasmussen R (2001) Quantification on the LightCycler. In *Rapid cycle real-time PCR: methods and applications* (pp. 21-34).
- Rozen S and Skaletsky H Primer3 on the WWW for General Users and for Biologist Programmers.
- Moreno NN, Olthof AM and Svejstrup JQ (2023) Transcription-Coupled Nucleotide Excision Repair and the Transcriptional Response to UV-Induced DNA Damage. doi: 10.1146/annurev-biochem-052621.
- Openshaw K (2000) A review of *Jatropha curcas*: an oil plant of unfulfilled promise. *Biomass and bioenergy*, 19(1), 1-15.
- Pfaffl MW, Horgan GW and Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR.
- Raxwal VK, Ghosh S, Singh S, Katiyar-Agarwal S, Goel S, Jagannath A, Kumar A, Scaria V and Agarwal M (2020) Abiotic stress-mediated modulation of the chromatin landscape in *Arabidopsis thaliana*. *J Exp Bot* 71:5280–5293.
- Sachdev S, Ansari SA, Ansari MI, Fujita M and Hasanuzzaman M (2021) Abiotic stress and reactive oxygen species: Generation, signaling, and defense mechanisms. *Antioxidants* 10:1–37.
- Saldanha AJ (2004) Java Treeview - Extensible visualization of microarray data. *Bioinformatics* 20:3246–3248.
- Sani, A., Danturai, S., Mohamed, M., Abubakar, M., Kafinga, M., Aliyu, R., & Garba, M. (2022). Phytoremediation of heavy metals contaminated agricultural soil planted to *Jatropha curcas*. *International Journal of Energy and Environmental Research*, 10(3), 3.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.
- Shen X, Mizuguchi G, Hamiche A and Wu C (2000) A chromatin remodelling complex involved in transcription and DNA processing.
- Song Y, He J, Guo J, Xie Y, Ma Z, Liu Z, Niu C, Li X, Chu B, Tahir MM et al. (2023) The chromatin remodeller MdRAD5B enhances drought tolerance by coupling MdLHP1-mediated H3K27me3 in apple. *Plant Biotechnol J.* doi: 10.1111/pbi.14210.
- Song ZT, Liu JX and Han JJ (2021) Chromatin remodeling factors regulate environmental stress responses in plants. *J Integr Plant Biol* 63:438–450.

- Souza MCP, da Silva MD, Binneck E, de Lima Cabral GA, Benko Iseppon AM, Pompelli MF, Endres L and Kido ÉA (2020) RNA-Seq transcriptome analysis of *Jatropha curcas* L. accessions after salt stimulus and unigene-derived microsatellite mining. *Ind Crops Prod.* doi: 10.1016/j.indcrop.2020.112168
- Sperschneider J, Catanzariti AM, Deboer K, Petre B, Gardiner DM, Singh KB, Dodds PN and Taylor JM (2017) LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci Rep.* doi: 10.1038/srep44598.
- Sun J, Sun Y, Ahmed RI, Ren A and Xie M (2019) Research progress on plant RING-finger proteins. *Genes (Basel).* doi: 10.3390/genes10120973
- Sun Y, Wu L, Li H, Liu Y, Zhang Q, Guo T, Sang Y and Mu C (2022) Chromatin Remodeling ATPases CHR12 and CHR23 Affect Apical Meristem Development and Play an Important Role during Response to Drought Stress in *Arabidopsis*. *Russ J Plant Physiol.* doi: 10.1134/S1021443722060255.
- Supek F, Bošnjak M, Škunca N and Šmuc T (2011) Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One.* doi:10.1371/journal.pone.0021800
- Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S et al. (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 51:D638–D646.
- Tamura K, Stecher G and Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* 38:3022–3027.
- Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BTT et al. (2006) Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3 HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Mol Cell* 24:785–796.
- Tian F, Yang DC, Meng YQ, Jin J and Gao G (2020) PlantRegMap: Charting functional regulatory maps in plants. *Nucleic Acids Res* 48:D1104–D1113.
- Tsukiyama T, Daniel C, Tamkun J and Wu C (1995) /SWI/, a Member of the SWI2/SNF2 ATPase Family, Encodes the 140 kDa Subunit of the Nucleosome Remodeling Factor.
- Walley JW, Rowe HC, Xiao Y, Chehab EW, Kliebenstein DJ, Wagner D and Dehesh K (2008) The chromatin remodeler SPLAYED regulates specific stress signaling pathways. *PLoS Pathog.* doi: 10.1371/journal.ppat.1000237
- Wang J, Sun Z, Liu H, Yue L, Wang F, Liu S, Su B, Liu B, Kong F and Fang C (2023) Genome-Wide Identification and Characterization of the Soybean Snf2 Gene Family and Expression Response to Rhizobia. *Int J Mol Sci.* doi: 10.3390/ijms24087250
- Wang Z, Wang Y, Tong Q, Xu G, Xu M, Li H, Fan P, Li S and Liang Z (2021) Transcriptomic analysis of grapevine Dof transcription factor gene family in response to cold stress and functional analyses of the VaDof17d gene. *Planta.* doi: 10.1007/s00425-021-03574-8

- Wu J, Jiang Y, Liang Y, Chen L, Chen W and Cheng B (2019) Expression of the maize MYB transcription factor ZmMYB3R enhances drought and salt stress tolerance in transgenic plants. *Plant Physiol Biochem* 137:179–188.
- Wu X, Shi H and Guo Z (2018) Overexpression of a NF-YC gene results in enhanced drought and salt tolerance in transgenic seashore paspalum. *Frontiers in Plant Science*, 9, 1355.
- Xia X, Liu X, Li T, Fang X and Chen Z (2016) Structure of chromatin remodeler Swi2/Snf2 in the resting state. *Nat Struct Mol Biol* 23:722–729.
- Xu K, Chen S, Li T, Yu S, Zhao H, Liu H and Luo L (2021) Overexpression of OsHMGB707, a High Mobility Group Protein, Enhances Rice Drought Tolerance by Promoting Stress-Related Gene Expression. *Front Plant Sci*. doi: 10.3389/fpls.2021.711271.
- Yan L, Wang L, Tian Y, Xia X and Chen Z (2016) Structure and regulation of the chromatin remodeller ISWI. *Nature* 540:466–469.
- Yang R, Hong Y, Ren Z, Tang K, Zhang H, Zhu JK and Zhao C (2019) A role for PICKLE in the regulation of cold and salt stress tolerance in Arabidopsis. *Front Plant Sci*. doi: 10.3389/fpls.2019.00900.
- Yang Y, Han X, Ma L, Wu Y, Liu X, Fu H, Liu G, Lei X and Guo Y (2021) Dynamic changes of phosphatidylinositol and phosphatidylinositol 4-phosphate levels modulate H⁺-ATPase and Na⁺/H⁺ antiporter activities to maintain ion homeostasis in Arabidopsis under salt stress. *Mol Plant* 14:2000–2014.
- Yu X, Meng X, Liu Y, Li N, Zhang A, Wang TJ, Jiang L, Pang J, Zhao X, Qi X et al. (2018) The chromatin remodeler ZmCHB101 impacts expression of osmotic stress-responsive genes in maize. *Plant Mol Biol* 97:451–465.
- Yuan F, Leng B, Zhang H, Wang X, Han G and Wang B (2019) A WD40-Repeat Protein From the Recretohalophyte *Limonium bicolor* Enhances Trichome Formation and Salt Tolerance in Arabidopsis. *Front Plant Sci*. doi: 10.3389/fpls.2019.01456.
- Zhang D, Gao S, Yang P, Yang J, Yang S and Wu K (2019) Identification and expression analysis of Snf2 family proteins in tomato (*Solanum lycopersicum*). *Int J Genomics*. doi: 10.1155/2019/5080935.
- Zhao Q, Hu RS, Liu D, Liu X, Wang J, Xiang XH and Li YY (2020) The AP2 transcription factor NtERF172 confers drought resistance by modifying NtCAT. *Plant Biotechnol J* 18:2444–2455.

5 ARTIGO 2**CARACTERIZAÇÃO ESTRUTURAL DOS GENES CODIFICADORES DE DNA
HELICASES DE PINHÃO-MANSO (*Jatropha curcas* L.)**

Running title: DNA helicases de pinhão-manso.

Rahisa Helena da Silva¹; Manassés Daniel da Silva¹; Bruna de Brito Souza¹; Antonio Félix da Costa²; Éderson Akio Kido¹

¹Departamento de Genética, Centro de Biociências, Programa de Pós-graduação em Genética e Biologia Molecular - PPGGBM, Universidade Federal de Pernambuco, Recife, PE, Brasil.

²Instituto Agrônômico de Pernambuco, IPA, Recife, PE, Brasil.

Corresponding author: kido.ufpe@ufpe.br

Resumo

DNA helicases são proteínas motoras caracterizadas por desenrolar estruturas de DNA. Essas enzimas estão diretamente envolvidas na manutenção da estabilidade genômica por atuarem em vias de reparo de lesões no DNA. Pinhão-mansão, uma oleaginosa atrativa para produção de biodiesel, carece de informações sobre as helicases codificadas pelo seu genoma. Neste trabalho, foram identificados 25 genes codificadores de DNA helicases no genoma de pinhão-mansão. Esses genes correspondem às famílias MCM [SF-VI], RuvB [SF-VI] e UvrD [SF-I]. As proteínas codificadas por esses genes foram classificadas nas subfamílias correspondentes: MCM (quatro subfamílias: Mcm2 [dois genes], Mcm5 [um gene], Mcm6 [dois genes] e Mcm8 [dois genes]); RuvB (11 subfamílias: RuvBL1 – RuvBL11); e UvrD (duas subfamílias: UvrD1 e UvrD2). As análises de arquitetura de domínios e distribuição de motivos conservados ressaltaram a presença do *core* das helicases constituído por dois domínios (um helicase N-terminal e um ATPase C-terminal) e o compartilhamento dos motivos *Walker A* e *Walker B*. Também foram observadas algumas especificidades quanto à arquitetura dos domínios e motivos de acordo com as subfamílias analisadas. DNA helicases *JcRuvBL* são altamente diversificadas estruturalmente, enquanto que *JcMCM* e *JcUvrD* apresentam poucas variações estruturais. Modelos 3D reforçaram o dobramento típico esperado para essas proteínas. Os resultados indicam que essas DNA helicases possuem todas as características estruturais compatíveis com suas atividades. As regiões promotoras dos genes exibiram candidatos a elementos *cis* tais como MYB, MYC e ABRE, demonstrando que possivelmente estão associados com rotas de defesa e aclimação contra estresses. Redes de interação proteína-proteína (PPI) mostraram a associação dessas proteínas com vias importantes para a manutenção da homeostase celular em plantas.

submetidas a estresses tais como: reparo do DNA e checagem do ciclo celular, *turnover* de proteínas, tráfego intracelular em vesículas e biogênese de cloroplastos e peroxissomos.

Palavras chave: Estabilidade genômica; Oleaginosa; MCM; RuvB; UvrD; Estresse abiótico.

Introdução

A manutenção da integridade do genoma é crucial a todos os organismos vivos pois garante a fidelidade de transmissão da informação genética. Lesões no DNA podem ser reparadas por múltiplas vias e as DNA helicases são componentes centrais (Tuteja et al., 2009; Nisa et al., 2019). Com base da formação ou não de anéis oligoméricos, helicases são agrupadas em seis superfamílias (SF1-6) (Singleton et al., 2007; Passricha et al., 2019). Em eucariotos, famílias de DNA helicases estão compreendidas nas superfamílias SF-I, SF-II e SF-VI. Essas famílias correspondem a UvrD, Pif1 (*Petite Integration Factor 1*), RecG, Rad3, MCM (*mini-chromosome maintenance*) e RuvB [*Recombination UVB (sensitivity) like*] (Seraj et al., 2018).

Em plantas, famílias de DNA helicase já foram descritas em *Arabidopsis* (nove genes MCM; um gene UvrD), *Oryza sativa* (10 genes MCM; um gene UvrD) (Tuteja e Tuteja, 2013; Singh et al., 2021), *Brassica oleraceae* e *Brassica rapa* (12 genes MCM cada) (Shanmugam et al., 2017), e *Triticum aestivum* (RuvBL, nove genes) (Chaudhary et al., 2023). Alguns desses trabalhos relatam a participação desses genes em respostas de plantas submetidas a estresses (Dang et al., 2011; Saifi et al., 2018; Chaudhary et al., 2023; Singh et al., 2020; Saifi et al., 2023). O gene *MCM6* de ervilha foi descrito promovendo tolerância a salinidade em plantas de tabaco que o superexpressavam (Dang et al., 2011). Sendo assim, a exploração de genes codificadores de DNA helicases pode ser uma alternativa biotecnológica, pois a manutenção da estabilidade genômica provavelmente desempenha importantes papéis na tolerância de plantas a estresses (Nisa et al., 2019).

Pinhão-mansô (*Jatropha curcas* L.) é uma oleaginosa não comestível que pertence à família Euphorbiaceae. As plantas são atrativas devido a grande quantidade de óleo armazenado em suas sementes, além de seu fácil

desenvolvimento em áreas marginais (Achten et al., 2008), podendo também ser utilizadas no controle de erosão e contaminações do solo (Nero e Neto, 2021; Ferreira et al., 2022). O potencial de produção de biodiesel do pinhão-mansão é negativamente afetado devido a presença de estresses abióticos devido ao fato da espécie ser sensível a salinidade (Lozano-Isla et al., 2018). Portanto o estudo de potenciais genes candidatos associados com rotas de defesa é crucial para obtenção de variedades mais tolerantes.

Neste estudo, foram identificados um total de 25 genes codificadores de DNA helicases no genoma de pinhão-mansão. Os padrões de domínios, motivos conservados e estrutura gênica foram analisados. Modelos tridimensionais foram obtidos considerando os candidatos identificados. As regiões promotoras dos genes também foram analisadas quanto a presença de candidatos a elementos *cis*-regulatórios. Redes de PPI também foram construídas visando melhor entender os mecanismos e funções dessas DNA helicases. O conhecimento gerado pode auxiliar investigações futuras sobre essas famílias gênicas em plantas submetidas a estresses.

Metodologia

Identificação dos membros das famílias de DNA helicases em *J. curcas* e estruturas éxon-ínton

Para busca das helicases candidatas, foi aplicado o *Hidden Markov Model* (HMM). Os perfis HMM MCM (PF00493), RuvB (PF05496) e UvrD (PF00580) do banco de dados de famílias do Pfam ([http:// pfam.xfam.org/](http://pfam.xfam.org/)) (Finn et al., 2015) foram utilizados. O proteoma putativo do pinhão-mansão (GCF_014843425.1_RJC1_Hi-C_protein.faa) foi obtido do NCBI (<https://www.ncbi.nlm.nih.gov>) e os perfis HMM utilizados como

query para busca nessas sequências utilizando parâmetros *default* com um *e-value cut-off* de $< 1.0 \times 10^{-5}$ através do software HMMER v3.3.2. As potenciais sequências foram submetidas a identificação de domínios conservados através do CDD (*Conserved Domain Database* - <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Marchler-Bauer et al., 2015), e InterPro (<https://www.ebi.ac.uk/interpro>) (Quevillon et al., 2005). *Gene Structure Display Server* v2.0 (GSDS, <http://gsds.gao-lab.org/>) (Hu et al., 2015) foi utilizado para obter as estruturas éxon-íntron dos genes a partir da comparação da sequência codificante (CDS) com a sequência genômica.

Análise das regiões promotoras

A partir do painel *Genome Browser* do NCBI (<https://www.ncbi.nlm.nih.gov/genome/gdv/browser/>), as regiões promotoras (1.0 kb *upstream* ao sítio de início da transcrição) dos genes DNA helicase foram recuperadas. Cada região promotora foi analisada pela ferramenta *web* PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/0>) para detecção de potenciais sítios de ligação de fatores de transcrição (Rombauts et al., 1999).

Caracterização estrutural das sequências

As proteínas identificadas tiveram seus parâmetros físico-químicos detectados pelo ExPASy (<http://web.expasy.org/protparam/>) (Gasteiger et al., 2005). A localização subcelular foi predita através da ferramenta LOCALIZER v1.0.4 (Sperschneider et al., 2017). Os programas Jalview v2.11.1.7 e MEME v5.3.3 (<https://meme-suite.org/meme/>) (Bailey et al., 2015) e foram utilizados para identificar motivos conservados nas sequências. Os seguintes parâmetros foram aplicados na busca pelo MEME: *anr* (*any numbers of repetitions*); número máximo de motivos: 05 (MCM); 07 (RuvB); tamanho dos motivos: 6-50 resíduos de aminoácidos. A ferramenta TBtools

v1.098693 (Chen et al., 2020) foi utilizada para visualizar os padrões estruturais dos motivos detectados.

Análise fenética

Com o objetivo de classificar as famílias DNA helicase, sequências relativas as mesmas famílias DNA helicases nas espécies: *Manihot esculenta* (39), *Ricinus communis* (36), *Populus trichocarpa* (34) foram obtidas através de buscas HMM nos proteomas dessas espécies oriundos do banco de dados do Phytozome (<https://phytozome-next.jgi.doe.gov/>) além de sequências previamente caracterizadas de *Brassica rapa* (41; <http://brassicadb.org/brad/>), *Oryza sativa* (26; <https://rapdb.dna.affrc.go.jp/>) e *Arabidopsis thaliana* (27; <https://www.arabidopsis.org/>). As sequências das proteínas foram alinhadas pelo programa ClustalX (Larkin et al., 2007). As árvores foram construídas pelo método *Neighbor Joining* utilizando *bootstrap* de 1.000 replicatas e visualizadas pela ferramenta online iTOL v5 (*Interactive Tree Of Life*; <https://itol.embl.de/>) (Letunic e Bork, 2021).

Predição de estruturas secundárias e modelagem 3-D

A modelagem por homologia dos candidatos foi construída através da ferramenta web Swissmodel (<https://swissmodel.expasy.org/>) (Waterhouse et al., 2018) baseada na estrutura do *template* PDB mais similar. A acurácia dos modelos propostos foi avaliada utilizando gráficos Ramachandran gerados no próprio Swissmodel. SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) foi utilizado como método estrutural secundário. Os *softwares* Ali2D (<https://toolkit.tuebingen.mpg.de/tools/ali2d>) (Alva et al., 2016) e 2SSd

(<http://genome.lcqb.upmc.fr/2dss/>) foram utilizados para visualização das estruturas secundárias (Lotun et al., 2019).

Rede de interação proteína-proteína (PPI)

STRING (v11.5) (Szklarczyk et al., 2023) foi utilizado para determinar a rede PPI das DNA helicases identificadas. O *score* de confiança da interação foi definido como 0.7 (*high*), baseado em interações oriundas de experimentos em *Arabidopsis*. As evidências baseadas em *text-mining* foram excluídas. Cytoscape v3.10.1 (Shannon et al., 2003) foi utilizado para visualização e edição das redes PPI.

Resultados

Identificação dos genes DNA helicase em pinhão-mansão

As buscas HMM resultaram em 25 genes (43 proteínas) codificadores de DNA helicase a partir do genoma de pinhão-mansão depositado no NCBI (GCF_014843425.1_RJC1_Hi-C). Destes, sete genes (15 proteínas) correspondem a família de DNA helicases MCM (*JcMCM*), 17 (23 proteínas) a família RuvB (*JcRuvBL*) e um gene (05 proteínas) a família UvrD (*JcUvrD*). As informações sobre os membros das famílias gênicas estão descritas na tabela 1, incluindo localização genômica, tamanho da ORF, tamanho do gene e número de íntrons.

Tabela 1 - Informações acerca dos genes codificadores de DNA helicase de pinhão-mansão abrangendo as famílias MCM, RuvB e UvrD.

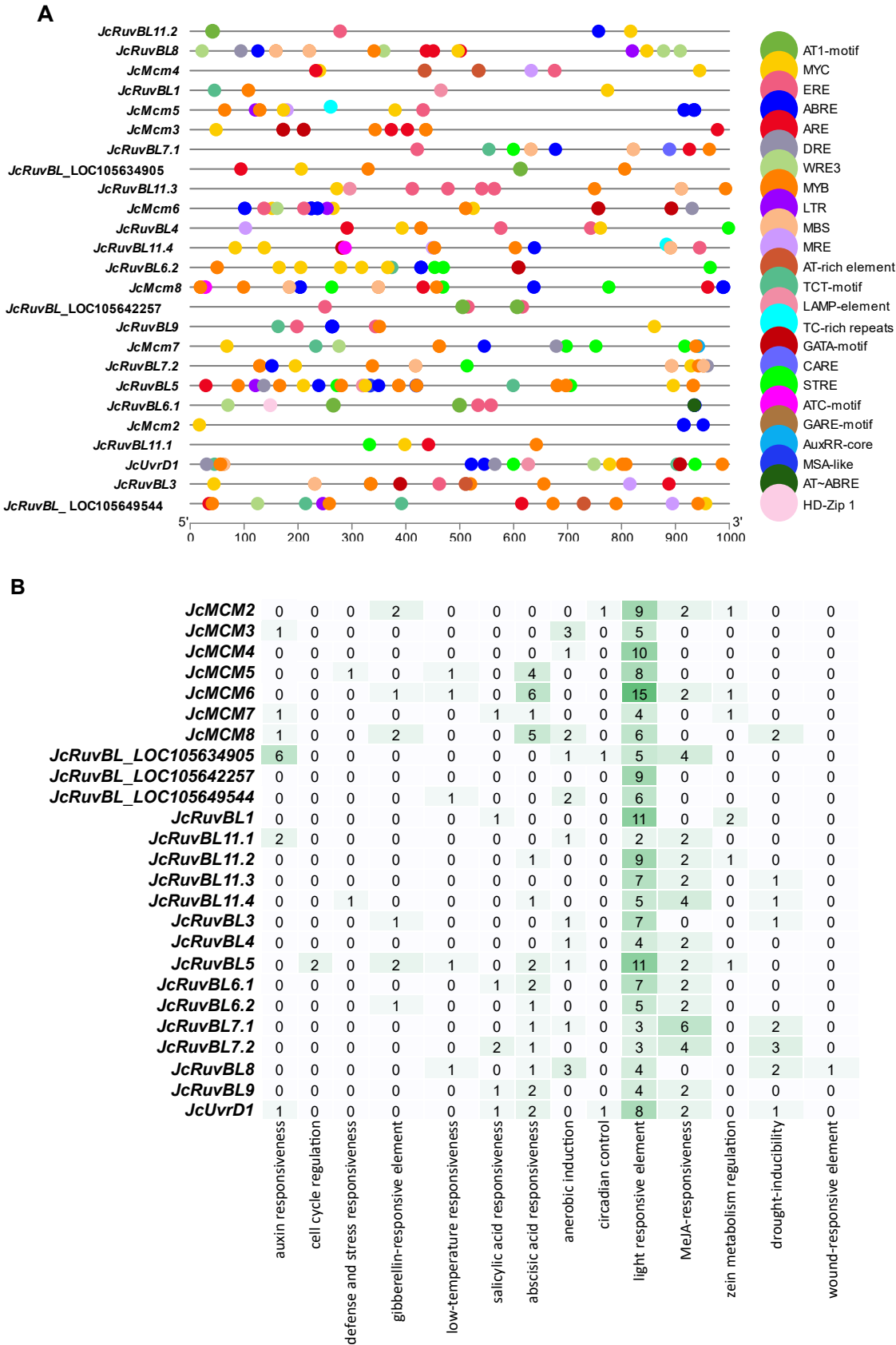
Gene	Nome	Localização genômica	Gene (nt)	CDS (nt)	Nº de íntrons
LOC105633466	<i>JcMcm5</i>	112069-117633	5564	2184	17
LOC105630109	<i>JcMcm4</i>	207155-214433	7278	2502	17
LOC105642618	<i>JcMcm7</i>	1740549-1746882	6333	2157	14
LOC105633977	<i>JcMcm3</i>	488753-494550	5797	2310	15
LOC105636763	<i>JcMcm6</i>	1465029-1470464	5435	2514	17
LOC105642169	<i>JcMcm8</i>	13453-19569	6116	2091	14
LOC105647274	<i>JcMcm2</i>	8570-15231	6661	2826	16
LOC105629586	<i>JcRuvBL8</i>	656170-683861	27691	2358	14
LOC105634905	<i>JcRuvBL</i> _LOC105634905	672622-684295	11673	2898	13

LOC105649544	<i>JcRuvBL</i> _LOC105649544	326985-338307	11322	2223	20
LOC105635574	<i>JcRuvBL11.3</i>	1735112-1745493	10381	1362	1
LOC105647925	<i>JcRuvBL11.1</i>	27987-38190	10203	3141	11
LOC105639509	<i>JcRuvBL11.4</i>	1032846-1042176	9330	2499	9
LOC105642473	<i>JcRuvBL9</i>	320360-328154	7794	2832	15
LOC105629294	<i>JcRuvBL11.2</i>	717138-724664	7526	1848	11
LOC105633993	<i>JcRuvBL7.1</i>	144014-151304	7290	1224	10
LOC105641200	<i>JcRuvBL6.2</i>	2225363-2231825	6462	1185	4
LOC105645734	<i>JcRuvBL6.1</i>	45168-51228	6060	2511	15
LOC105642257	<i>JcRuvBL</i> _LOC105642257	439325-444485	5160	2418	8
LOC105638964	<i>JcRuvBL4</i>	4052179-4057075	4896	1032	10
LOC105643096	<i>JcRuvBL7.2</i>	208308-213157	4849	1482	12
LOC105630761	<i>JcRuvBL1</i>	23641-28094	4453	1392	11
LOC105649259	<i>JcRuvBL3</i>	302797-304785	1988	1602	0
LOC105648667	<i>JcUvrD1</i>	16586089-16596072	9983	2928	22
LOC105645524	<i>JcRuvBL5</i>	399614-401320	1706	1419	1

Análise dos promotores dos genes DNA helicase

As potenciais regiões promotoras (1.0 kb) dos genes foram analisadas quanto a presença elementos *cis*-regulatórios. Foram detectados 22 tipos de elementos *cis* nessas regiões (Figura 1). De acordo com a análise, os elementos mais frequentes são os da família MYB (*v-myb avian myeloblastosis viral oncogene homolog*) ABRE (*ABA-responsive element*), MYC (*myelocytomatosis oncogenes*) e ERE (*ethylene responsive element*). Outros candidatos incluindo LTR (*low temperature response*), DRE (*dehydration-responsive element*), GATA (*GATA-binding factor*) e HD-ZIP (*homeodomain leucine zipper*) também foram detectados (Figura 1). Esses elementos *cis* puderam ser divididos em 14 tipos diferentes, destacando resposta à luz, resposta a MeJA (metiljasmonato) e resposta ao ABA (Figura 1). Esses candidatos são descritos como componentes chave para regulação da expressão gênica durante exposição de plantas a estresses.

Figura 1 – A. Candidatos a elementos cis-regulatórios detectados nos promotores (1 kb *upstream* ao sítio de início da transcrição) dos genes DNA helicase de pinhão-mansó; B. número de elementos cis detectados em cada promotor divididos em 14 tipos.

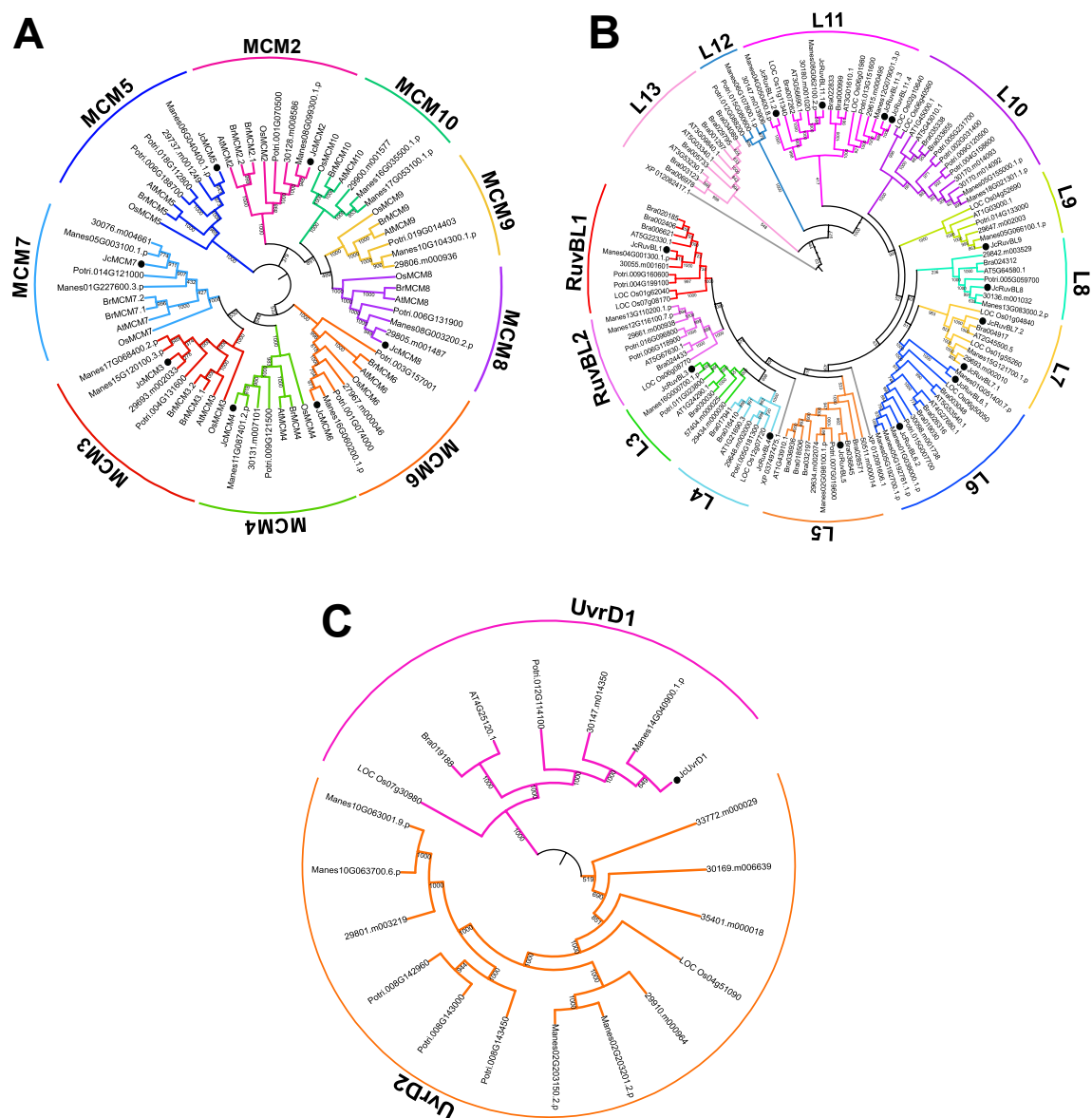


Análise fenética

Com o objetivo de classificar as DNA helicases de *J. curcas*, foram construídas árvores fenéticas considerando sequências de DNA helicases das famílias MCM, RuvB e UvrD. Apenas as sequências das proteínas primárias de cada *loci* foram consideradas. As DNA helicases de *J. curcas* exibiram nove, treze e dois subgrupos para as famílias MCM, RuvB e UvrD, respectivamente (Figura 2). Genomas de plantas possuem até nove cópias distintas para a família MCM, que codifica um complexo helicase hexamérico formado pelas subunidades centrais MCM2-7, enquanto que as subunidades MCM8-MCM10 são co-reguladores de MCM2-7 (Shultz et al., 2007). A partir das informações disponíveis para *Arabidopsis*, *Oryza sativa* e *Brassica rapa*, as proteínas JcMCM foram classificadas nas subfamílias JcMCM2, JcMCM3, JcMCM4, JcMCM5, JcMCM6, JcMCM7 e JcMCM8 (Figura 2).

Não há classificação proposta para as famílias RuvB e UvrD. No presente trabalho, a nomenclatura designada para as famílias RuvB (RuvBL1-RuvBL13) e UvrD (UvrD1-UvrD2) foi feita inicialmente com base na observação dos subgrupos formados nas árvores (Figura 2) e nos padrões de distribuição de domínios e motivos conservados, comentados a seguir. Três sequências JcRuvBL (JcRuvBL_LOC105634905; JcRuvBL_LOC105649544; JcRuvBL_LOC105642257) não foram classificadas em nenhuma subfamília, mas foram consideradas pois apresentam as características estruturais esperadas.

Figura 2 - Árvore fenéticas construídas pelo método *Neighbor Joining* (bootstrap de 1.000 replicatas) evidenciando A. subfamílias MCM2 – MCM10 para a família MCM; B. subfamílias RuvBL1- RuvBL13 para a família RuvB; e C. subfamílias UvrD1 e UvrD2 para a família UvrD.

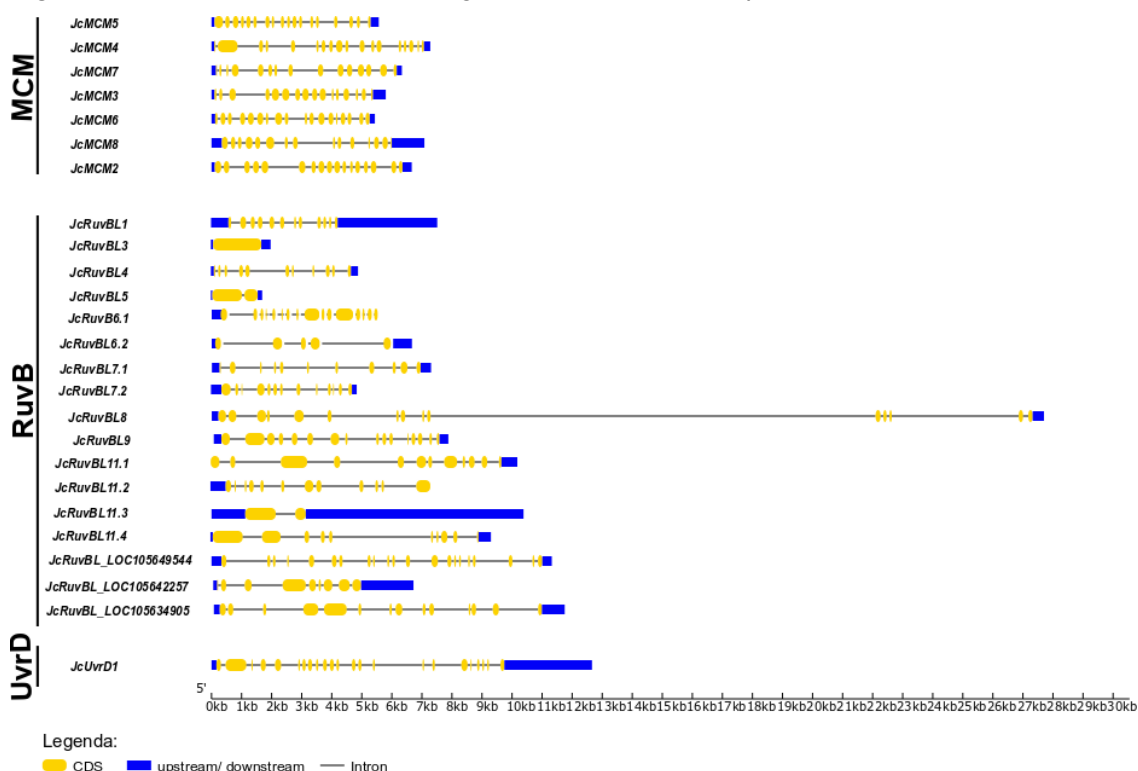


Estrutura gênica

Nas DNA helicases, o número de íntrons nos genes *JcMCM* variou de 14 até 17 e de 0 até 20 nos genes *JcRuvB* (Tabela 1; Figura 3, 4 e 5). Apenas o gene *JcRuvBL3* não apresentou íntrons (Figura 3). Os genes das famílias *JcMCM* apresentam estruturas gênicas mais simples, com pouca variação de tamanho de genes (5.500 até quase 7.200 pb) e tamanhos de éxons e íntrons. Em contrapartida,

os genes *JcRuvBL* apresentam grandes variações nos tamanhos de genes (1.706 até quase 28.000 pb) e nos tamanhos de éxons, íntrons e de regiões UTR (Tabela 1; Figura 3). O gene *JcUvrD1* possui 22 íntrons e aproximadamente 10.000 pb (Figura 3). Nenhum padrão específico de estrutura gênica foi evidente dentro das subfamílias dessas DNA helicases.

Figura 3 - Estruturas éxon-íntron dos genes DNA helicase de pinhão-mansô.



Características das proteínas DNA helicase

As proteínas codificadas pelos genes DNA helicase de pinhão-mansô identificados foram caracterizadas quanto ao peso molecular (MW), ponto isoelétrico (pI) e localização subcelular. As informações estão descritas na tabela 2. O tamanho das proteínas variou de 669 até 941 nas *JcMCM*, de 342 - 1.046 nas *JcRuvB*, e em *JcUvrD* de 1061 - 1155 aminoácidos. Os valores pI variaram de 5,17 até 8,34 em *JcMCM*, de 5,05 - 9,29 em *JcRuvBL* e de 7,10 - 8,67 em *JcUvrD*. Da mesma forma, o

peso molecular nas *JcMCM* variou de 73,63 - 106,95 KDa, nas *JcRuvBL* de 37,01 - 115,39 KDa e em *JcUvrD* de 119,58 – 130,96 KDa (Tabela 2).

A função biológica de uma proteína é influenciada diretamente pela localização subcelular. De acordo com a análise, as DNA helicases de pinhão-mansão foram preditas para até cinco compartimentos celulares distintos. Para as DNA helicases *JcMCM* (15) e *JcUvrD* (05), todas as proteínas tiveram localização predita para o núcleo; as *JcRuvBL* tiveram localização predita para núcleo (09), citoplasma (07), mitocôndria (04), membrana plasmática (2) e cloroplasto (01) (Tabela 2). Helicases de pinhão-mansão da família RuvBL são altamente variáveis em termos de massa molecular, ponto isoelétrico e localização subcelular, enquanto que as famílias MCM e UvrD variam menos com relação a esses parâmetros.

Tabela 2 - Características das proteínas codificadas pelos genes DNA helicase de pinhão-mansão. MW: peso molecular; pI: ponto isoelétrico. aa: tamanho em aminoácidos das proteínas; pI: ponto isoelétrico; MW: peso molecular; MP: membrana plasmática.

Nome	Proteína	aa	MW (KDa)	pI	Localização Subcelular
<i>JcMcm5</i>	XP_012071446.1	727	80,80936	8.34	Núcleo
<i>JcMcm4</i>	XP_012067200.1	833	92,57524	6.02	Núcleo
<i>JcMcm7</i>	XP_012082889.1	718	80,8644	5.72	Núcleo
<i>JcMcm3</i>	XP_012072088.1	769	85,47919	6.15	Núcleo
<i>JcMcm6</i>	XP_012075497.1	838	93,21738	5.74	Núcleo
<i>JcMcm8</i>	XP_037494357.1	669	73,62712	8.31	Núcleo
<i>JcMcm2</i>	XP_037496346.1	941	106,95305	5.17	Núcleo
<i>JcRuvBL8</i>	XP_037491287.1	785	88,34255	6.84	Cloroplasto
<i>JcRuvBL_LOC105634905</i>	XP_037497475.1	965	106,22746	6.47	MP
<i>JcRuvBL_LOC105649544</i>	XP_012091606.1	740	81,4484	6.18	Citoplasma
<i>JcRuvBL11.3</i>	XP_037496430.1	841	93,81385	6.52	Núcleo
<i>JcRuvBL11.1</i>	XP_037491077.1	1046	115,39476	5.94	Núcleo
<i>JcRuvBL11.4</i>	XP_037497874.1	758	85,05114	6.20	Núcleo
<i>JcRuvBL9</i>	XP_012082698.1	943	103,21847	6.35	Núcleo
<i>JcRuvBL11.2</i>	XP_012066248.1	615	67,11649	7.63	Mitocôndria
<i>JcRuvBL7.1</i>	XP_012072108.1	407	45,57673	6.08	Núcleo
<i>JcRuvBL6.2</i>	XP_012081080.1	394	44,0465	5.52	Citoplasma
<i>JcRuvBL6.1</i>	XP_020539541.1	836	93,22509	7.57	Mitocôndria
<i>JcRuvBL_LOC105642257</i>	XP_012082417.1	805	89,59561	5.16	Citoplasma
<i>JcRuvBL4</i>	XP_037492311.1	342	38,21831	8.19	MP

<i>JcRuvBL7.2</i>	XP_012083544.1	493	54,68507	9.29	Núcleo
<i>JcRuvBL1</i>	XP_012068094.1	458	50,19896	5.76	Citoplasma
<i>JcRuvBL3</i>	XP_012091241.1	533	58,38265	7.21	Núcleo
<i>JcUvrD1</i>	XP_020540797.1	1061	119,58181	7.10	Núcleo
<i>JcRuvBL5</i>	XP_020539496.1	472	54,70079	9.26	Núcleo

Domínios e motivos conservados

Domínios e motivos conservados estão diretamente associados às especificidades das funções de uma proteína. O *core* das helicases é formado por até quatorze motivos conservados distribuídos ao longo de dois domínios centrais (Singleton et al., 2007). As DNA helicases MCM e RuvB pertencem a SF-VI (ou AAA+) e as helicases UvrD a SF-I. De modo geral, as helicases de pinhão-mansão apresentaram dois domínios compondo a estrutura do *core*, um N-terminal e um C-terminal com atividades helicase e ATPase, respectivamente. Todas compartilham os motivos I (*Walker A*) e II (*Walker B*) e cada família exibe motivos específicos (Figura 5).

Nas sequências *JcMCM*, foram detectados os domínios MCM_N e MCM_OB, que correspondem a região N-terminal (NTD). Na região C-terminal (CTD) foram detectados os domínios MCM (AAA+, ATPase) e MCM_lid. Os três motivos conservados característicos estiveram dispostos na região que compreende o domínio MCM (AAA+), os três motivos conservados são nomeados: *Walker A* (GDPGTAKSQL), *Walker B* (DEFDKM) e dedo de arginina (SFRD) (Figura 4 e 5). Não foi observada nenhuma especificidade com relação aos padrões de organização dos domínios e motivos conservados nessas sequências (Figura 5).

Nas sequências *JcRuvBL* os dois domínios que representam o *core* são relativos a RuvB_N terminal e AAA+ C-terminal (ATPase). A maior parte das

sequências não exibem essa estrutura, mas possuem domínio o domínio C-terminal AAA+ que as definem como membros da família. Domínios auxiliares tais como CDC48_N, MIT, MgSA e PEX1_N também foram detectados (Figura 4). Os motivos *Walker A* (GXXGXGK[T/S]), *Walker B* (DEXH), Sensor I (XATNR), Sensor II (TXLRX) e dedo de arginina também estiveram presentes nas sequências *JcRuvBL* (Figura 6). Foi possível observar alguns padrões de ocorrência dos domínios e motivos conservados nas *JcRuvBL*, como por exemplo as subfamílias *JcRuvBLL11* e *JcRuvBL6* (Figura 4 e 5).

Da mesma forma, as *JcUvrD* identificadas possuem um domínio N-terminal (helicase) e C-terminal (ATPase) e todos os 14 motivos conservados próprios da família, nomeados Q, Ia, Ic, Id, II, III, IV, IVa, IVb, IVc, V, Va, VI e VIa (Figura 4 e 5). Os resultados indicam que as proteínas codificadas pelos genes DNA helicase de pinhão-mansão exibem o *core* estrutural característico das helicases, essencial para o desempenho de suas funções.

Figura 4 - Estrutura dos domínios conservados detectados nas DNA helicases JcMCM, JcRuvBL e JcUvrD de acordo com as respectivas subfamílias. Os domínios foram detectados através dos softwares SMART e InterProScan.

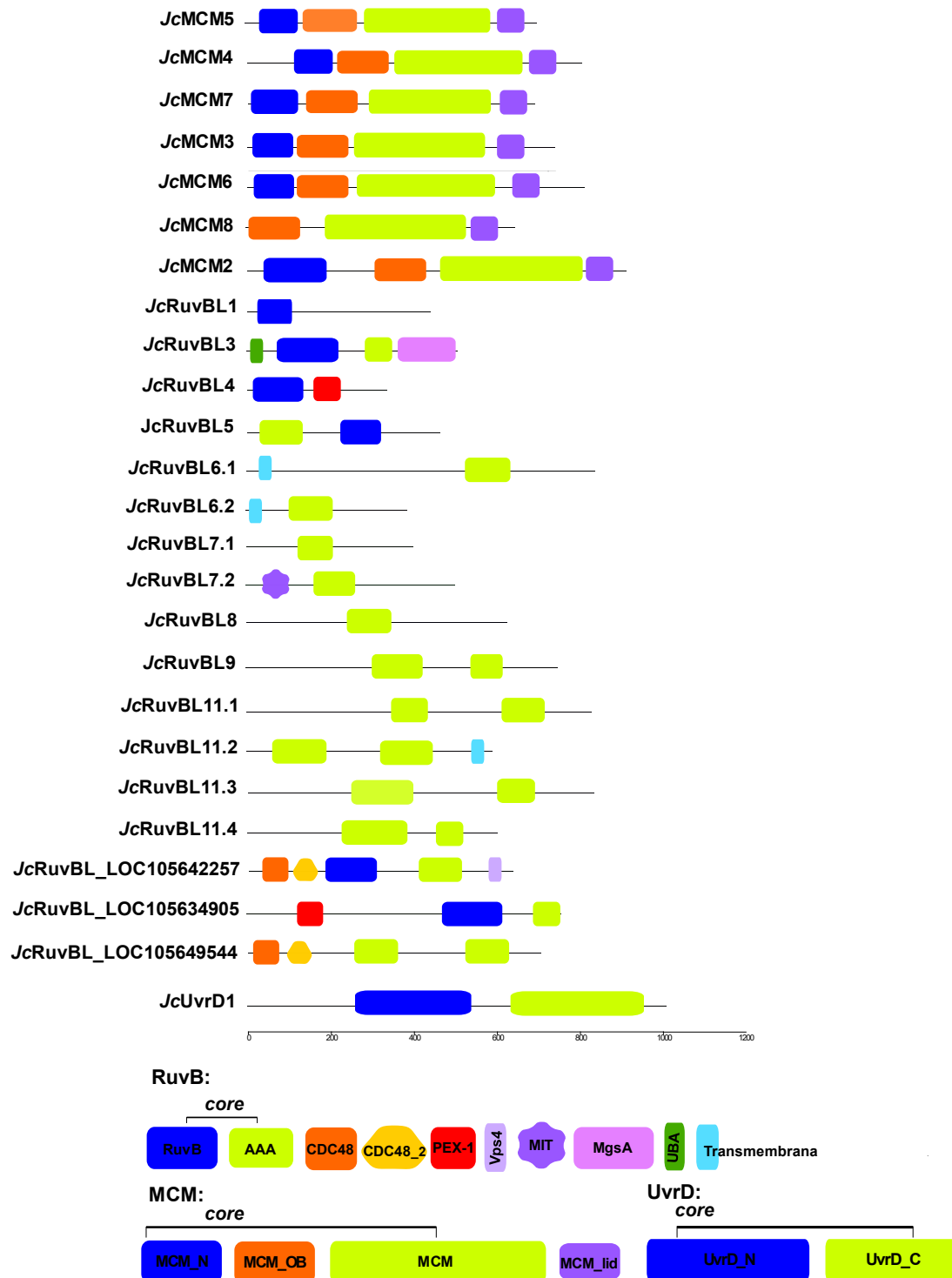
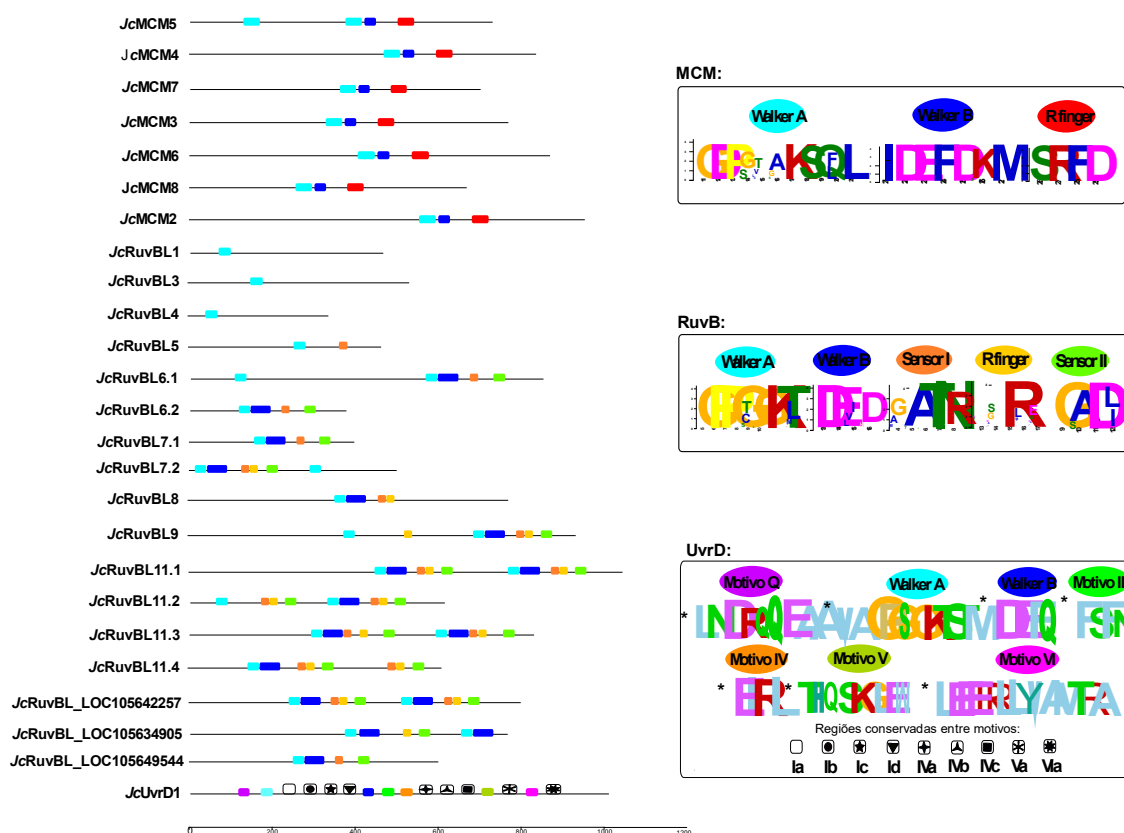


Figura 5 - Distribuição dos motivos conservados detectados nas DNA helicases *JcMCM*, *JcRuvBL* e *JcUvrD* de acordo com as respectivas subfamílias. Os motivos foram identificados através dos programas MEME e Jalview*.



Estruturas secundárias

As estruturas secundárias nas sequências peptídicas das helicases de pinhão-mansão foram preditas através da ferramenta SOPMA (Tabela 3). Essas helicases contêm em média 44,11% de α -hélice e 13,55% de folha- β em *JcMCM*, 46,02% de α -hélice e 8,90% de folha- β em *JcRuvBL* e 47,17% de α -hélice e 9,38% de folha- β em *JcUvrD*. Motivos conservados foram localizados, em sua maioria, nas regiões de transição entre folha- β e α -hélice (Figura 6). Apesar de suas diferenças em composição de domínios e motivos, DNA helicases de pinhão-mansão são altamente similares quanto à presença de estruturas secundárias em suas proteínas.

Figura 6 - Estruturas secundárias detectadas nas proteínas codificadas por genes DNA helicase de pinhão-mansão para as famílias A. *JcMCM*; B. *JcRuvBL* e C. *JcUvrD*. Retângulos verdes realçam os motivos conservados característicos de cada família.



Tabela 3 - Estruturas secundárias detectadas através da ferramenta SOPMA a partir das proteínas das famílias DNA helicase *JcMCM*, *JcRuvB* e *JcUvrD*.

Nome	α -hélice	Folha- β	Beta turn	Random coil
<i>JcMcm5</i>	41.95%	16.23%	5.78%	36.04%
<i>JcMcm4</i>	39.26%	14.17%	5.52%	41.06%
<i>JcMcm7</i>	49.44%	13.93%	4.74%	31.89%
<i>JcMcm3</i>	44.08%	12.22%	4.16%	39.53%
<i>JcMcm6</i>	46.42%	13.01%	4.18%	36.40%
<i>JcMcm8</i>	42.00%	13.90%	5.53%	38.57%
<i>JcMcm2</i>	45.59%	11.37%	3.83%	39.21%
<i>JcRuvBL8</i>	44.20%	10.45%	0.00%	45.35%
<i>JcRuvBL_LOC105634905</i>	34.40%	10.88%	0.00%	54.72%
<i>JcRuvBL_LOC105649544</i>	43.92%	12.84%	0.00%	43.24%
<i>JcRuvBL11.3</i>	48.39%	6.18%	0.00%	45.42%
<i>JcRuvBL11.1</i>	40.25%	6.50%	0.00%	53.25%
<i>JcRuvBL11.4</i>	43.67%	5.28%	0.00%	51.06%
<i>JcRuvBL9</i>	37.22%	9.54%	0.00%	53.23%
<i>JcRuvBL11.2</i>	48.13%	10.24%	0.00%	41.63%
<i>JcRuvBL7.1</i>	40.29%	7.37%	0.00%	52.33%
<i>JcRuvBL6.2</i>	55.84%	6.60%	0.00%	37.56%
<i>JcRuvBL6.1</i>	48.80%	8.37%	0.00%	42.82%
<i>JcRuvBL_LOC105642257</i>	43.11%	11.55%	0.00%	45.34%
<i>JcRuvBL4</i>	61.99%	7.02%	0.00%	30.99%
<i>JcRuvBL7.2</i>	50.30%	6.09%	0.00%	43.61%
<i>JcRuvBL1</i>	52.62%	11.14%	0.00%	36.24%
<i>JcRuvBL3</i>	39.21%	8.07%	0.00%	52.72%
<i>JcRuvBL5</i>	50.00%	13.14%	0.00%	36.86%
<i>JcUvrD1</i>	43.54%	8.95%	0.00%	47.50%

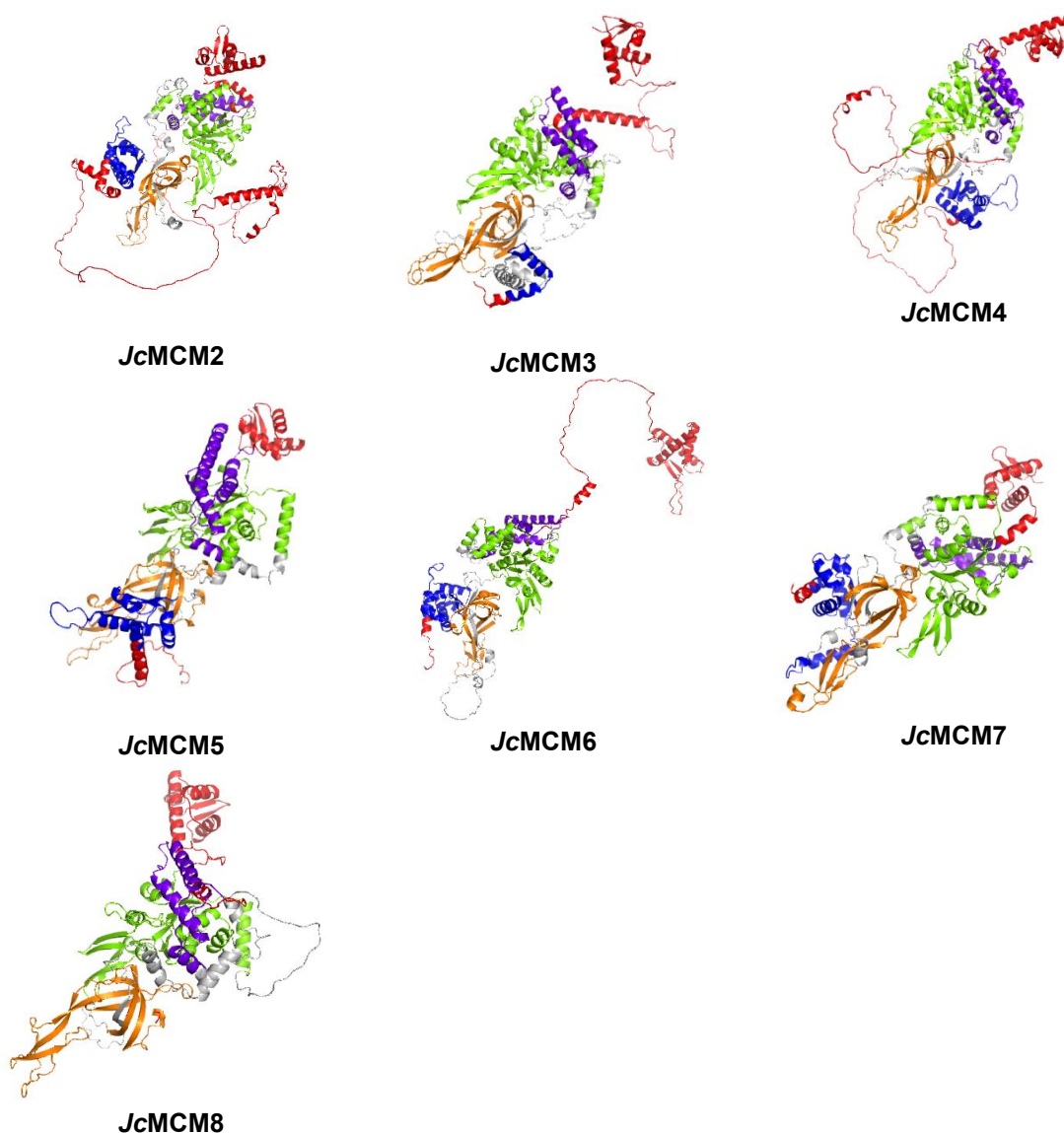
Modelagem por homologia

Todas as DNA helicases de pinhão-mansão foram submetidas ao Swiss-model para construção de estruturas tridimensionais. Apenas foram considerados modelos com GMQE (*Global Model Quality Estimate*) acima de 0.60, totalizando 25 estruturas, das quais 24 apresentam > 90% de cobertura. Os modelos apresentaram 84,30 – 96,38% dos resíduos nas regiões permitidas e o QMEANDisCo global score variou de 0,44 – 0,83 (Tabela 4). Os modelos para cada família helicase de pinhão-mansão estão ilustrados nas figuras 7, 8 e 9.

Todas as estruturas propostas para as *JcMCM* compartilham a estrutura em monômero típica mostrando dois domínios distintos: um domínio N-terminal (NTD) um

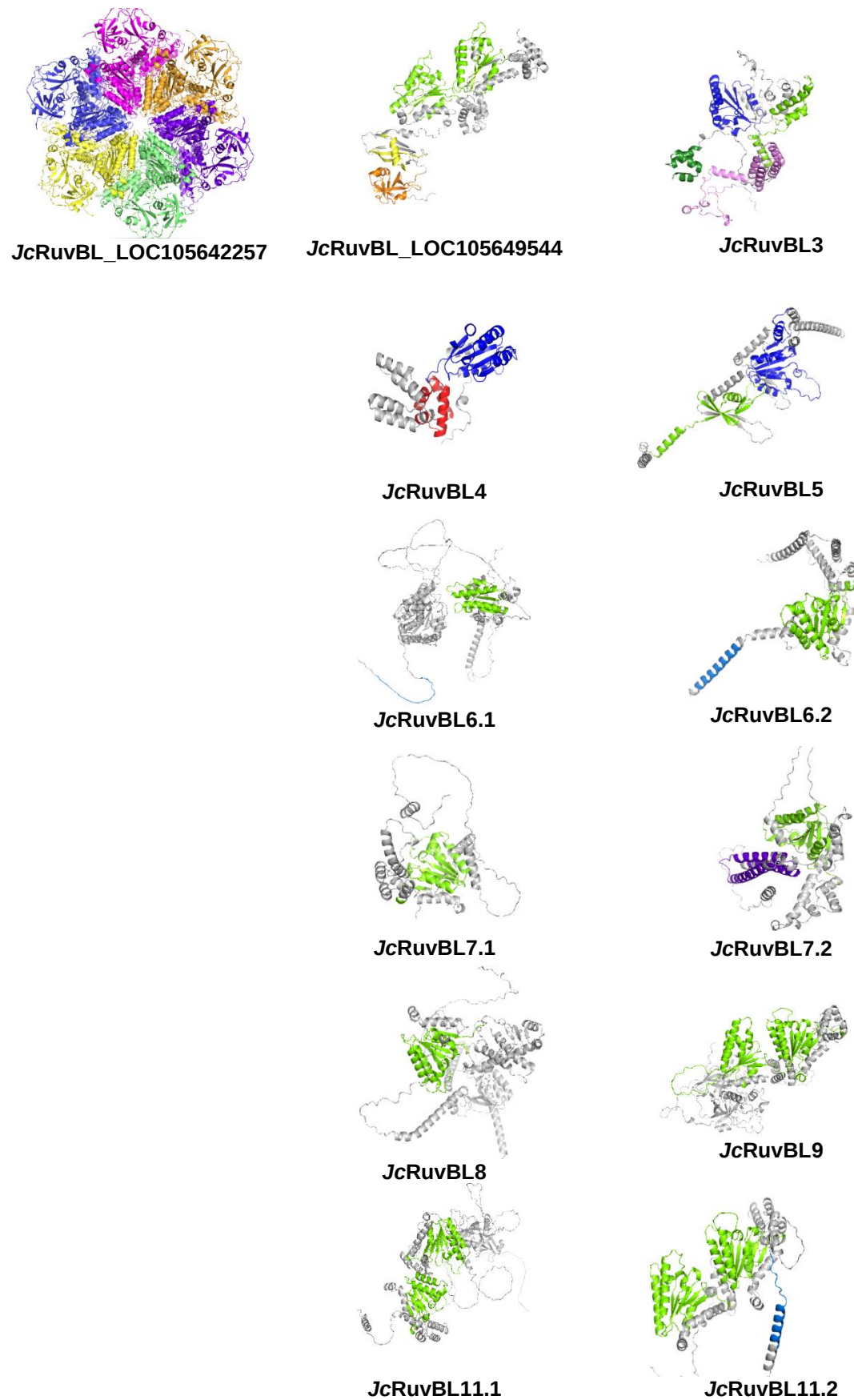
domínio C-terminal AAA+ (CTD) (Figura 7). Domínios OB-*fold* exibiram longas folhas- β que se enrolam formando uma estrutura semelhante a um β -barril (Fletcher et al, 2003). Ainda, estiveram presentes entre os domínios NTD e CTD, *linkers* flexíveis além de extensões N e C-terminal, NTEs e CTEs, respectivamente, que variam em extensão de acordo com cada proteína (Figura 7).

Figura 7 - Modelos 3D gerados para DNA helicases MCM de pinhão-mansó através do Swiss-model. Estruturas MCM são divididas em domínios MCM N-terminal (NTD; *blue*), OB-*fold* (orange) e domínios C-terminal (CTD; *green*); MCM_lid (*purple blue*); Extensões NTE e CTE estão realçadas em vermelho.



As estruturas tridimensionais para as proteínas JcRuvB mostraram o dobramento em monômero que reflete as helicases de SF-VI do grupo (HEC; *helicases and clamp loaders*) (Ammelburg et al., 2006) (Figura 8). Domínios *core* exibiram estrutura em “formato de triângulo” em um dobramento α - β - α que diz respeito a estruturas folha- β rodeados por α -hélices de cada lado (Matias et al., 2006). Ainda, um modelo hexamérico que apresentou a estrutura de um anel contendo um canal central, foi obtido para a proteína (JcRuvBL_LOC105642257) (Figura 8).

Figura 8 - Modelos 3D gerados para DNA helicases RuvBL de pinhão-mansó através do Swiss-model. Estruturas dos domínios *core* estão divididas em RuvB_N (NTD; *blue*); AAA (CTD; *green*);



Igualmente, a proteína JcUvrD1 mostra o dobramento característico da superfamília SF-I, com dois domínios globulares relativos ao *core*, cada qual contendo estruturas folhas- β envolvidas por α -hélices (Figura 9). A partir da localização dos motivos, foi possível distinguir em cada domínio N- *core* (1A) e C-*core* (2A) dois subdomínios nomeados 1B e 2B (Velankar et al., 1999; Singleton et al., 2007).

Figura 9 - Modelo 3D gerado para DNA helicase JcUvrD1 de pinhão-mansão através do Swiss-model. Estruturas dos domínios *core* estão divididas em UvrD_N (1A; *blue*); UvrD_C (2A; *green*); Subdomínios 1B (*yellow*) e 2B (*orange*). Motivos conservados na fenda entre os domínios 1 e 2 estão realçados em vermelho.

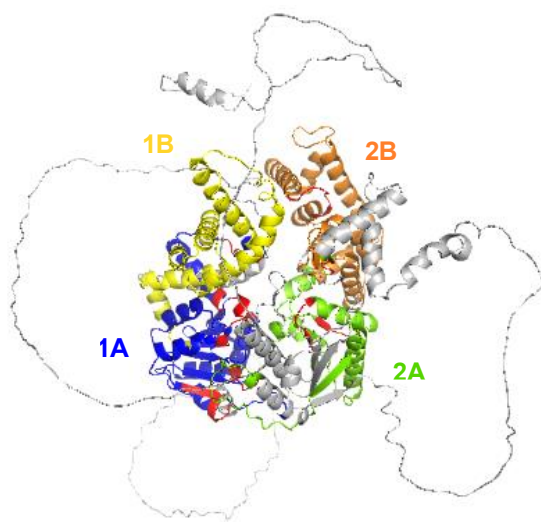


Tabela 4 - Parâmetros de qualidade obtidos para os modelos tridimensionais construídos pelo Swissmodel a partir das proteínas das famílias DNA helicase de pinhão-mansão. Sm: similaridade; Cv: cobertura; Rf: Ramachandran *favoured*.

Nome	GMQE	Identidade	Sm	Cv	Rf	QMEANDisCo
JcRuvBL6.1	0.69	81.38%	0.54	0,99	87,89%	0,45
JcRuvBL8	0.69	83.27%	0.56	0,97	91,57%	0,48
JcUvrD1	0.65	70.71%	0.51	0,99	84,30%	0,50
JcRuvBL_LOC105634905	0.68	70.66%	0.51	0,99	91,48%	0,51
JcRuvBL9	0.70	70.93%	0.51	0,97	90,22%	0,52
JcRuvBL11.4	0.66	85.87%	0.56	0,99	90,21%	0,53
JcRuvBL11.3	0.68	70.01%	0.50	0,97	89,14%	0,57
JcRuvBL7.2	0.78	75.88%	0.53	0,98	90,22%	0,57
JcRuvBL7.1	0.74	84.44%	0.56	1	91,85%	0,59
JcMcm8	0.78	85.87%	0.56	0,98	93,85%	0,59
JcRuvBL5	0.78	100.00%	0.62	1	96,38%	0,59
JcRuvBL3	0.80	72.47%	0.52	0,98	93,97%	0,60
JcRuvBL6.2	0.79	100.00%	0.60	1	95,92%	0,60
JcRuvBL11.2	0.78	80.56%	0.54	1	93,64%	0,60
JcMcm4	0.74	82.26%	0.55	0,99	90,84%	0,63
JcRuvBL_LOC105649544	0.83	89.73%	0.57	1	97,70%	0,64
JcMcm6	0.75	82.81%	0.55	0,99	92,58%	0,64
JcMcm3	0.77	84.90%	0.56	1	93,87%	0,65
JcRuvBL1	0.76	73.63%	0.53	0,99	95,28%	0,70
JcRuvBL4	0.75	81.48%	0.55	0,87	95,62%	0,70
JcMcm5	0.78	81.77%	0.55	1	97,06%	0,71
JcRuvBL_LOC105642257	0.69	78.62%	0.55	0,99	92,57%	0,72
JcRuvBL11.1	0.66	76.25%	0.52	0,99	87,50%	—
JcMcm7	0.81	90.38%	0.58	1	97,06%	—
JcMcm2	0.72	87.97%	0.58	0,99	89,10%	—

Rede de interações PPI das DNA helicases de pinhão-mansão

Com o objetivo de entender melhor os mecanismos moleculares associados com as DNA helicase de pinhão-mansão, redes PPI considerando cada família identificada foram construídas.

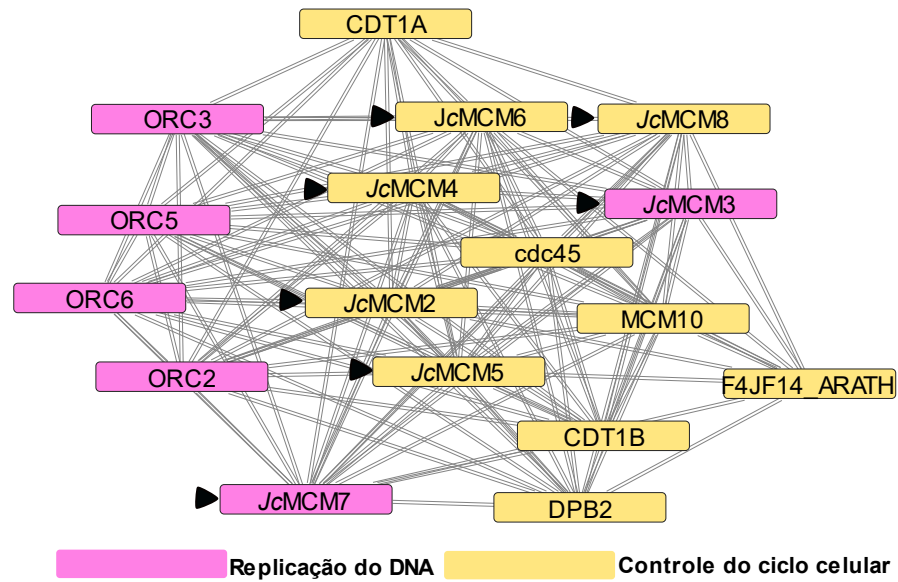
A rede PPI construída para as proteínas JcMCM evidencia dois *clusters* que correspondem a (1) replicação do DNA e (2) regulação do ciclo celular (Figura 10). Termos GO enriquecidos ($< 1.0e-16$) associados com esses processos incluem

replicação do DNA (GO:0006260), montagem do complexo de pré-iniciação da replicação do DNA (GO:0071163), progressão do ciclo celular (GO:0022402), reparo de quebra na dupla fita induzido pela replicação (GO:0000727) dentre outros (Tabela suplementar 1).

A rede PPI para as proteínas *JcRuvBL* evidencia seis *clusters* distintos que estão relacionados com: (1) remodelamento da cromatina e reparo do DNA; (2) replicação do DNA; (3) degradação de proteínas via proteassomo 26S; (4) desmontagem do complexo SNARE de transporte mediado por vesículas; (5) biogênese do cloroplasto; (6) transporte de proteínas a matriz do peroxissomo (Figura 11). Termos GO enriquecidos ($< 1.0e-16$) associados com esses processos estão descritos na tabela suplementar 1.

A rede PPI construída considerando a proteína *JcUvrD1* evidencia dois *clusters* associados com (1) reparo do DNA e (2) replicação do DNA (Figura 12). Termos GO enriquecidos associados com esses processos incluem regulação da replicação do DNA (GO:0006275), reparo do DNA (GO:0006281); processividade da DNA polimerase (GO:0030337) dentre outros (Tabela suplementar 1).

Figura 10 - Rede de interação proteína-proteína (PPI) considerando as proteínas JcMCM. A rede foi construída com base nas informações disponíveis para *Arabidopsis*. O score de confiança foi definido como *high* (> 0.7). Evidências baseadas em *text mining* foram excluídas. A rede mostra parceiros relacionados com a replicação do DNA (rosa) e controle do ciclo celular (amarelo).



Discussão

Plantas são constantemente expostas a agentes externos e internos, tais como excesso de radiação solar, que frequentemente causam danos ao DNA e conduzem a estresse genotóxico. O dano ao DNA diminui a estabilidade do genoma e, conseqüentemente, o crescimento e a produtividade (Tuteja et al., 2009). Vias de resposta ao dano no DNA (DDR; *DNA-damage responses*) incluem principalmente quinases *checkpoint* e DNA helicases. Essas vias permitem que o DNA danificado seja reparado, garantindo a fidelidade da transmissão da informação genética (Tuteja et al., 2009; Nisa et al., 2019). Caso não reparado, o dano ao DNA pode conduzir a redução da síntese proteica, destruição das membranas plasmáticas, deterioração do aparato fotossintético e até morte celular (Tuteja e Tuteja, 2013).

As características das famílias gênicas de DNA helicases ainda são desconhecidas em *J. curcas*, bem como, nas euforbiáceas mais próximas, como *M. esculenta* e *R. communis*. Neste estudo, as famílias de genes codificadores de DNA helicases em *J. curcas* foram analisadas a despeito de sua classificação, padrão de domínios, motivos conservados, estrutura gênica, região promotora, modelos 3D e interações em redes PPI.

Todas as DNA helicases analisadas compartilham a estrutura característica do *core* que corresponde aos motivos *Walker A* e *Walker B* e dois domínios centrais (N-terminal helicase e C-terminal ATPase). A alta conservação dos motivos *Walker A* e *Walker B* está relacionada com a atividade de coordenação e entre ligação e hidrólise de trifosfatos, e reflete a principal pressão evolutiva no sítio ativo dessas enzimas (Leipe et al., 2002).

As DNA helicases *JcMCM* e *JcUvrD* apresentam pouca variação quanto a seus padrões de estrutura gênica e características de suas proteínas tais como organização de domínios e motivos conservados, além de todas terem sido preditas a estarem localizadas no núcleo. Essas características são provavelmente decorrentes da conservação da função dessas helicases. DNA helicases MCM formam um complexo hexamérico que atua na replicação do DNA e DNA helicases UvrD são componentes do sistema de reparo de *mismatch* (MMR; *mismatch repair*), ambos sistemas conservados de eubactéria a eucariotos (Maiorano et al., 2006; Tuteja et al., 2009).

Em contrapartida, as DNA helicases *JcRuvBL* exibem grande variação quanto estruturas de éxon-ínton dos seus genes, ocorrência e padrão de domínios e motivos conservados, além de maior presença de domínios auxiliares tais como o domínio MIT (*microtubule interacting and trafficking*), UBA (*ubiquitin-associated*), CDC48 (*Cell Division Control Protein 48*) dentre outros. Ainda, as proteínas *JcRuvBL* foram preditas a estarem localizadas em cinco compartimentos celulares distintos. Essas evidências apoiam o envolvimento dessas proteínas em múltiplas funções celulares (Saifi et al., 2019).

As estruturas tridimensionais para as proteínas DNA helicases *JcMCM*, *JcRuvB* e *JcUvrD* mostraram os dobramentos característicos relacionados as superfamílias correspondentes (Amelburg et al., 2006; Matias et al., 2006; Tuteja et al., 2013). Estruturas obtidas para as helicases *JcMCM* e *JcRuvBL* exibiram características diretamente associadas com a formação dos anéis hexaméricos, a conformação correspondente com a atividade dessas enzimas, tais como: NTEs, CTEs e *linkers* flexíveis nas *JcMCM* (Fletcher et al., 2003; Brewster et al., 2008; Singh et al., 2021). Da mesma forma, o dobramento observado em *JcUvrD1* é associado com a dinâmica da formação de fendas que correspondem aos estados “aberto” e “fechado” dos

domínios *core* (1A; 2A) e dos subdomínios (1B; 2B) durante a interação com o DNA (Singleton et al., 2007). Sendo assim, os resultados indicam que as DNA helicases de pinhão-mansão possuem estruturas relacionadas ao desempenho das suas funções.

A ativação ou repressão da expressão gênica é controlada pela ligação de fatores de transcrição a elementos cis presentes nas regiões promotoras dos genes. A análise das regiões promotoras dos genes DNA helicase de pinhão-mansão revelou que os elementos estão principalmente relacionados com a resposta a luz, ao metiljasmonato e ao ácido abscísico. A luz pode modular respostas a estresses em plantas. Dentre os efeitos causados pelo excesso de luz estão a acumulação de espécies reativas de oxigênio (ROS) e a fotoinibição, que conduzem a estresse genotóxico (Roeber et al., 2021). DNA helicases são componentes centrais de vias de reparo do DNA. O reparo do DNA lesionado é essencial para manutenção da estabilidade genômica, garantindo o crescimento e a produtividade das plantas (Prajapati et al., 2023).

O metiljasmonato (MeJA) é um composto volátil que atua como mensageiro intracelular em mecanismos de defesa. Sob ataque de patógenos ou estresses abióticos, plantas aumentam a produção de MeJA. Esses compostos são liberados no ar e são reconhecidos por plantas vizinhas, que reforçam suas produções de MeJA para lidar com estresses (Gunjegaonkara e Shanmugarajam, 2019). Plantas de arroz apresentaram maior tolerância a salinidade quando tratadas com aplicações de MeJA (Hussain et al., 2022). O gene *PtMAPK3-1* de *Populus trichocarpa* é induzido por MeJA e atua na maquinaria antioxidante das plantas (Wu et al., 2023).

O ácido Abscísico (ABA) é um hormônio que regula diversos aspectos do crescimento e desenvolvimento de plantas. A via de sinalização do ABA é crítica para adaptação a estresses ambientais. Diversos aspectos da resposta estão relacionados

a essa via de sinalização tais como regulação estomatal, sinalização de cálcio, ubiquitinação de proteínas e arquitetura de raízes (Kumar et al., 2019; Aslam et al., 2022). Além disso, o ABA é considerado como sinal de estresse de longa distância (Ko e Helariutta, 2017). Fatores de transcrição controlam a expressão gênica e consequentemente muitos processos fisiológicos relacionados com respostas de plantas a fatores externos, logo, os resultados indicam que os genes DNA helicase de pinhão-mansão podem responder e participar de importantes rotas de resposta a estresses ambientais.

De acordo com a rede PPI, DNA helicases *JcMCM* e *JcUvrD* participam na replicação do DNA, checagem do ciclo celular e reparo do DNA. A manutenção da integridade do genoma requer a interação entre componentes de vias de reparo do DNA e a checagem do ciclo celular, pois o reparo depende da pausa na divisão celular (Hu et al., 2016). DNA helicases são componentes chave das vias de reparo por abrir a área danificada no DNA e permitir a entrada da maquinaria de reparo (Tuteja et al., 2009). Esses processos estão diretamente conectados com estresses ambientais, e já foram associados com tolerância a seca (Poku et al., 2021), salinidade (Okumura et al., 2021), calor (De Jaeger-Braet et al., 2022) e resistência a patógenos (Bourbousse et al., 2018). Por exemplo, a superexpressão da DNA topoisomerase *MtTdp2a* de *Medicago truncatula* conferiu maior tolerância aos estresses osmótico e foto-oxidativo (Confaloneri et al., 2014).

Parceiros funcionais de *JcRuvBL* incluem proteínas envolvidas em diversos processos que abrangem replicação e reparo do DNA, transporte de proteínas a matriz do peroxissomo, biogênese do cloroplasto, desmontagem do complexo SNARE e degradação de proteínas via ubiquitinação. Durante exposição a estresses, a manutenção da homeostase de peroxissomos é crucial para funcionamento normal

das células (Su et al. 2019). O transporte de proteínas a matriz do peroxissomo é uma das etapas importantes referentes a biogênese dessas organelas, que não possuem material genético, com todo seu conteúdo proteico codificado no núcleo (Hasan et al., 2013). Peroxissomos são um dos principais sítios de geração e eliminação de H_2O_2 , pois abrigam muitas enzimas incluindo catalases, ascorbato peroxidases e peroxirredoxinas o que torna essa organela crucial na coordenação de respostas a estresses (Schrader e Fahimi, 2006; Hasan et al., 2013).

Um outro processo relacionado com proteínas JcRuvBL é a biogênese de cloroplastos. Cloroplastos têm atuação relevante na adaptação de plantas a estresses (Zhang et al., 2020). Essas organelas são responsáveis pela fotossíntese e produção de compostos metabólicos tais como aminoácidos, vitaminas, fitohormônios, metabólitos secundários dentre outros (Chan et al., 2016; Zhang et al., 2020). Além disso, os cloroplastos servem como sensores de mudanças externas. Sob condições de estresse, cloroplastos auxiliam na otimização da expressão gênica nuclear, esse processo é conhecido como sinalização retrógrada, e é essencial para manutenção de processos fisiológicos (Chan et al., 2016).

Proteínas JcRuvBL também estiveram relacionadas com a desmontagem do complexo SNARE (*soluble N-ethylmaleimide-sensitive factor attachment protein receptors*). O complexo SNARE consiste em uma maquinaria de tráfego intracelular em vesículas que funciona para comunicação entre diferentes compartimentos celulares (Jahn et al., 2006). A função desse complexo é requerida para respostas a estresses bióticos e abióticos (Kwon et al., 2020). Por exemplo, o complexo foi descrito como responsável pela entrega de canais de K^+ e aquaporinas a membrana plasmática (Hachez et al., 2014; Zhang et al., 2015).

A degradação de proteínas dependente de ubiquitinação via proteassomo 26S (UPS) também foi associada com as proteínas *JcRuvBL* através da rede PPI. A ligação de moléculas de ubiquitina a proteínas alvo está relacionada com abundância, tráfego e localização (Stone, 2014). As mudanças no conteúdo proteico das células são apoiadas pelo sistema UPS, e são requeridas para processos de crescimento e desenvolvimento além de respostas a fatores externos (Ali, 2020). Por exemplo, o sistema UPS atua regulando a abundância de fatores de transcrição tais como *TaWRKY29* (Meng et al., 2023) e *R2R3-MYB* (Beathard et al., 2021). Além disso, a atividade do proteassomo 26S é associada com a degradação de proteínas com erro de dobramento, frequentemente produzidas sob condições de estresse e causam estresse proteotóxico, impedindo o funcionamento normal das células (Ali, 2020).

Os resultados confirmam e dão ênfase a variedade de funções exercidas pelas proteínas RuvB. Em destaque, as *JcRuvBL* parecem contribuir diretamente para a sinalização e cooperação entre organelas.

Há poucos relatos do envolvimento direto dessas famílias gênicas nas respostas de plantas submetidas a estresses. A subunidade codificada pelo gene *MCM6* de ervilha foi descrita promovendo tolerância a salinidade em plantas de tabaco. Plantas transgênicas mostraram maior conteúdo de clorofila e maiores taxas fotossintéticas comparado com plantas selvagens (Dang et al., 2011). Saifi et al., (2023) analisaram perfil de expressão de genes *RuvBL* de trigo em resposta a seca e ao calor (1 e 6 h). Os autores observaram que os genes *TaRuvBL1b-3B* e *TaRuvBL2a-2D* foram significativamente induzidos ao calor na cultivar tolerante a partir de dados qPCR. O gene *OsRuvBL1a* foi induzido em plantas de arroz submetidas a salinidade (100 mM), frio (4°C) e calor (45°C) por 1, 2, 12, 24, 30 e 36 h (Saifi et al., 2018). Plantas transgênicas de arroz superexpressando *OsRuvBL1a* foram mais tolerantes a

salinidade (Saifi et al., 2023). O gene *OsRuvB* também foi descrito conferindo tolerância a salinidade quando superexpresso em plantas de *Cajanus cajan* (Singh et al., 2023).

A despeito da importância dos processos relacionados com essas DNA helicases, os resultados obtidos podem auxiliar na escolha e proposição de genes candidatos a serem utilizados em ensaios de caracterização funcional com o objetivo de obter plantas mais tolerantes a estresses.

Conclusões

Neste trabalho foram identificados 25 genes codificadores de DNA helicases no genoma de pinhão-manso. As análises estruturais conduzidas ressaltaram que as DNA helicases analisadas compartilham a estrutura clássica relativa às helicases, indicando que são provavelmente compatíveis com o desempenho das suas atividades. DNA helicases *JcMCM* e *JcUvrD* exibiram pouca variação em relação aos padrões estruturais analisados, o que pode ser refletido no direcionamento dessas proteínas a funções muito específicas e conservadas, que são replicação e reparo do DNA. De fato, nas redes PPI construídas considerando as proteínas *JcMCM* e *JcUvrD* os parceiros funcionais identificados estão relacionados com esses dois processos acima mencionados. Por outro lado, as DNA helicases *JcRuvBL* são muito diversificadas em termos de estrutura éxon-íntron dos seus genes, padrões de domínios e motivos conservados e localização subcelular. A grande variação observada para essas características apoia o envolvimento dessas proteínas em múltiplas funções celulares, que corrobora com a variedade de processos observados na rede PPI construída com base nas proteínas dessa família. Ainda, os resultados indicam que as *JcRuvBL* parecem contribuir diretamente para a sinalização e cooperação entre organelas. Em virtude da importância dos processos relacionados

com essas DNA helicases, os resultados obtidos podem auxiliar a utilização desses alvos como ferramentas para obtenção de plantas com melhor performance quando sob estresses ambientais.

Referências

- Achten WMJ, Verchot L, Franken YJ, Mathijs E, Singh VP, Aerts R and Muys B (2008) *Jatropha* bio-diesel production and use. *Biomass and Bioenergy* 32:1063–1084.
- Ali S (2020) Protective Roles of Cytosolic and Plastidal Proteasomes on Abiotic Stress and Pathogen Invasion. 1–17.
- Alva V, Nam S, Johannes S, Lupas AN, Mpi T and Toolkit B (2016) The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure. 44:410–415.
- Ammelburg M, Frickey T and Lupas AN (2006) Classification of AAA + proteins. 156:2–11.
- Aslam MM, Waseem M, Jakada BH, Okal EJ, Lei Z, Sohaib H, Saqib A, Yuan W, Xu W and Zhang Q (2022) Mechanisms of Abscissic Acid-Mediated Drought Stress Responses in Plants.
- Bailey TL, Johnson J, Grant CE and Noble WS (2015) The MEME suite. *Nucleic acids research*, 43(W1), W39-W49.
- Bajaj I and Rohde M (2021) Light acts as a stressor and influences abiotic and biotic stress responses in plants. 645–664.
- Beathard C, Mooney S, Al-saharin R, Goyer A and Allen RD (2021) Characterization of *Arabidopsis thaliana* R2R3 S23 MYB Transcription Factors as Novel Targets of the Ubiquitin Proteasome-Pathway and Regulators of Salt Stress and Abscissic Acid Response. 12:1–15.
- Bochman ML and Schwacha A (2009) The Mcm Complex: Unwinding the Mechanism of a Replicative Helicase. 73:652–683.
- Bourbousse C, Vegesna N and Law JA (2018) SOG1 activator and MYB3R repressors regulate a complex DNA damage network in *Arabidopsis*. 115:12453–12462.
- Brewster AS, Slaymaker IM, Afif SA and Chen XS (2010) Mutational analysis of an archaeal minichromosome maintenance protein exterior hairpin reveals critical residues for helicase activity and DNA binding. 0–7.
- Chan KX, Phua SY, Crisp P, Mcquinn R and Pogson BJ (2016) Learning the Languages of the Chloroplast: Retrograde Signaling and Beyond. doi: 10.1146/annurev-arplant-043015-111854
- Chaudhary J, Gautam T, Gahlaut V, Singh K, Kumar S and Batra R (2023) Identification and characterization of RuvBL DNA helicase genes for tolerance against abiotic

- stresses in bread wheat (*Triticum aestivum* L.) and related species. *Funct Integr Genomics* 23:1–24.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y and Xia R (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* 13:1194–1202.
- Confalonieri M, Fae M, Balestrazzi A, Dona M, Valassi A, Giraffa G and Carbonera D (2014) Enhanced osmotic stress tolerance in *Medicago truncatula* plants overexpressing the DNA repair gene *MtTdp2* a (tyrosyl-DNA. 187–203.
- Dang HQ, Tran NQ, Gill SS, Tuteja R and Tuteja N (2011) A single subunit MCM6 from pea promotes salinity stress tolerance without affecting yield. 19–34.
- De Jaeger-Braet J, Krause L, Buchholz A and Schnittger A (2022) Heat stress reveals a specialized variant of the pachytene checkpoint in meiosis of *Arabidopsis thaliana*. *The Plant Cell*, 34(1), 433-454.
- Emanuelle S, Doblin MS, Gooley PR and Gentry MS (2018) Biochemical and Biophysical Research Communications The UBA domain of SnRK1 promotes activation and maintains catalytic activity. *Biochem Biophys Res Commun* 497:127–132.
- Ferreira OJM, Holanda FSR, Pedrotti A, Vidal Santos LD and Silva-Mann R (2022) Root System of *Jatropha curcas* Provides Resistance and Strength to the Soil. *Communications in Soil Science and Plant Analysis*, 53(22), 2955-2967.
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A et al. (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285.
- Fletcher RJ, Bishop BE, Leon RP, Sclafani RA, Ogata CM and Chen XS (2003) The structure and function of MCM from. 10:160–168.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD and Bairoch A Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server.
- Gunjegaonkar SM and Shanmugarajan TS (2019) Molecular mechanism of plant stress hormone methyl jasmonate for its anti-inflammatory activity. *Plant Signal Behav.* doi: 10.1080/15592324.2019.1642038
- Hachez C, Laloux T, Reinhardt H, Cavez D, Degan H, Grefen C and Chaumont, F (2014) *Arabidopsis* SNAREs SYP61 and SYP121 coordinate the trafficking of plasma membrane aquaporin PIP2; 7 to modulate the cell membrane water permeability. *The Plant Cell*, 26(7), 3132-3147.
- Hasan S, Platta HW and Erdmann R (2013) Import of proteins into the peroxisomal matrix. 4:1–12. Hu B, Jin J, Guo AY, Zhang H, Luo J and Gao G (2015) GS

- 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31:1296–1297.
- Hu Z, Cools T and Veylder L De (2016) Mechanisms Used by Plants to Cope with DNA Damage. doi: 10.1146/annurev-arplant-043015-111902
- Hu B, Jin J, Guo AY, Zhang H, Luo J and Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8), 1296-1297.
- Hussain S, Zhang R, Liu S, Li R, Wang Y, Chen Y and Dai Q (2022) Methyl Jasmonate Alleviates the Deleterious Effects of Salinity Stress by Augmenting Antioxidant Enzyme Activity and Ion Homeostasis in Rice (*Oryza sativa* L.). *Agronomy*, 12(10), 2343.
- Iwaya N, Kuwahara Y, Fujiwara Y, Goda N, Tenno T, Akiyama K, Mase S, Tochio H, Ikegami T, Shirakawa M et al. (2010) A Common Substrate Recognition Mode Conserved between Katanin p60 and VPS4 Governs Microtubule Severing and Membrane Skeleton Reorganization * □. *J Biol Chem* 285:16822–16829.
- Jahn R and Scheller RH (2006) fusion. 631–643. José O, Ferreira M, Sandro F, Holanda R, Diego L,
- Santos V and Silva-mann R (2022) Communications in Soil Science and Plant Analysis Root System of *Jatropha curcas* Provides Resistance and Strength to the Soil Root System of *Jatropha curcas* Provides Resistance and Strength to. *Commun Soil Sci Plant Anal* 53:2955–2967.
- Ko D and Satiety LN (2017) Minireview Shoot – Root Communication in Flowering Plants. 973–978.
- Krause L and Buchholz A (2022) Heat stress reveals a specialized variant of the pachytene checkpoint in meiosis of *Arabidopsis thaliana*. 433–454.
- Kumar M, Kesawat MS, Ali A, Lee S, Gill SS and Kim HU (2019) Integration of Absciscic Acid Signaling with Other Signaling Pathways in Plant Stress Responses and Development. 1–20.
- Ko D and Helariutta Y (2017) Shoot–root communication in flowering plants. *Current Biology*, 27(17), R973-R978.
- Kwon C, Lee J and Yun HS (2020) Molecules and Cells SNAREs in Plant Biotic and Abiotic Stress Responses. 25:501–508.
- Larkin MA, Blackshields G, Brown NP, Chenna R, Mcgettigan PA, Mcwilliam H, Valentin F, Wallace IM, Wilm A, Lopez R et al. (2007) Clustal W and Clustal X version 2. 0. 23:2947–2948.
- Leipe DD, Wolf YI, Koonin E V and Aravind L (2002) Classification and Evolution of P-loop GTPases and Related ATPases. doi: 10.1006/jmbi.2001.5378
- Letunic I and Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic acids research*, 49(W1), W293-W296.

- Lotun DP, Cochard C, Vieira FRJ and Bernardes JS (2019) 2dSS: a web server for protein secondary structure visualization.
- Maiorano D, Lutzmann M and Me M MCM proteins and DNA replication ´ chali. doi: 10.1016/j.ceb.2006.02.006
- Lozano-Isla F, Campos ML, Endres L, Bezerra-Neto E and Pompelli MF (2018) Effects of seed storage time and salt stress on the germination of *Jatropha curcas* L. *Industrial Crops and Products*, 118, 214-224.
- Mamaeva A, Taliansky M, Golub N, Fesenko I, Filippova A and Love AJ (2020) The role of chloroplast protein remodeling in stress responses and shaping of the plant peptidome. doi: 10.1111/nph.16620
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI et al. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226.
- Matias PM, Gorynia S, Donner P and Arme M (2006) Crystal Structure of the Human AAA ½ Protein RuvBL1 *. 281:38918–38929.
- Meng Y, Lv Q, Li L, Wang B, Chen L, Yang W and Li X (2023) E3 ubiquitin ligase TaSDIR1-4A activates membrane-bound transcription factor TaWRKY29 to positively regulate drought resistance. *Plant Biotechnology Journal*.
- Nero BF and Nero BF (2021) Phytoremediation of petroleum hydrocarbon-contaminated soils with two plant species: *Jatropha curcas* and *Vetiveria zizanioides* at Ghana Manganese Company Ltd. *Int J Phytoremediation* 23:171–180.
- Nisa M, Huang Y, Benhamed M and Raynaud C (2019) The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. 10:1–12.
- Okumura T, Nomoto Y, Kobayashi K, Suzuki T, Takatsuka H and Ito M (2021) MYB3R - mediated active repression of cell cycle and growth under salt stress in *Arabidopsis thaliana*. *J Plant Res* 134:261–277.
- Oryza R and Dai Q (2022) Methyl Jasmonate Alleviates the Deleterious Effects of Salinity Stress by Augmenting Antioxidant Enzyme Activity and Ion.
- Panda S, Isbatan A and Adami GR (2008) Modification of the ATM / ATR directed DNA damage response state with aging and long after hepatocyte senescence induction in vivo. 129:332–340.
- Passricha N, Saifi SK, Gill SS, Tuteja R and Tuteja N (2019) Role of plant helicases in imparting salinity stress tolerance to plants. In *Helicases from all domains of life*. 39-52.
- Peng Z, Staub JM, Serino G, Kwok SF, Kurepa J, Bruce BD, Vierstra RD, Wei N and Deng X (2001) The Cellular Level of PR500, a Protein Complex Related to the 19S Regulatory Particle of the Proteasome, Is Regulated in Response to Stresses in Plants. 12:383–392.

- Physiology P, Protein HS, Protein B, Non-atpase R-TP and Non-atpase RP (2019) Heat Shock Proteins Support Refolding and Shredding of Misfolded Proteins. 180:1777–1778.
- Pisani FM, Ladenstein R, Liu W and Pucci B (2008) Structural analysis of the *Sulfolobus solfataricus* MCM protein N-terminal domain y. 36:3235–3243.
- Poku SA, Chukwurah PN, Aung HH and Nakamura I (2021) Knockdown of GmSOG1 Compromises Drought Tolerance in Transgenic Soybean Lines. 18–36.
- Prajapati P, Gupta P and Kharwar RN (2023) Nitric oxide mediated regulation of ascorbate- glutathione pathway alleviates mitotic aberrations and DNA damage in *Allium cepa* L. under salinity stress. *Int J Phytoremediation* 25:403–414.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R, Bucher P, Cerutti L and Res A (2005) InterProScan: protein domains identifier. 33:116–120.
- Roeber VM, Bajaj I, Rohde M, Schmölling T and Cortleven A (2021) Light acts as a stressor and influences abiotic and biotic stress responses in plants. *Plant, Cell & Environment*, 44(3), 645-664.
- Rombauts S, Déhais P, Montagu M Van and Rouzé P (1999) PlantCARE , a plant cis-acting regulatory element database. 27:295–296.
- Rycke R De, Inzé D, Blatt MR, Russinova E and Chaumont F (2014) Arabidopsis SNAREs SYP61 and SYP121 Coordinate the Trafficking of Plasma Membrane Aquaporin PIP2; 7 to Modulate the Cell Membrane Water Permeability. 26:3132–3147.
- Saifi SK, Passricha N, Tuteja R, Nath M, Gill R, Singh S and Tuteja N (2023) Plant Science OsRuvBL1a DNA helicase boost salinity and drought tolerance in transgenic indica rice raised by in planta transformation. 335:1–17.
- Saifi SK, Passricha N, Tuteja R and Tuteja N (2019) Chapter 14. An Overview of AAA+ Superfamily Proteins Associated Helicases. 247–264.
- Saifi SK, Passricha N, Tuteja R and Tuteja N (2018) Stress-induced *Oryza sativa* RuvBL1a is DNA-independent ATPase and unwinds DNA duplex in 3 ' to 5 ' direction. 669–684.
- Schrader M and Fahimi HD (2006) Peroxisomes and oxidative stress. 1763:1755–1766.
- Seraj ZI, Elias SM, Biswas S and Tuteja N (2018) Helicases and Their Importance in Abiotic Stresses. doi: 10.1007/978-3-319-90318-7
- Shanmugam A, Hasan A, Robin K, Thamilarasan SK and Vijayakumar H (2017) Genome-wide Characterization and Stress-responsive Expression Profiling of MCM Genes in *Brassica oleracea* and *Brassica rapa*. 472–484.

- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.
- Singh R, Sharma S, Kharb P and Saifi S (2020) OsRuvB transgene induces salt tolerance in pigeon pea. doi: 10.1080/17429145.2020.1722267
- Singh S, Chahar P, Macovei A, Yadav S, Ansari AA, Tuteja N and Gill R (2021) Comparative genomic analysis reveals evolutionary and structural attributes of MCM gene family in *Arabidopsis thaliana* and *Oryza sativa*. 327:117–132.
- Singleton MR, Dillingham MS and Wigley DB (2007) Structure and Mechanism of Helicases and Nucleic Acid Translocases. doi: 10.1146/annurev.biochem.76.052305.115300
- Sperschneider J, Catanzariti A, Deboer K, Petre B, Gardiner DM, Singh KB, Dodds PN and Taylor JM (2017) LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Nat Publ Gr* 1–14.
- Stone SL (2014) The role of ubiquitin and the 26S proteasome in plant abiotic stress signaling. 5:1–10.
- Su T, Li W, Wang P and Ma C (2019) Dynamics of Peroxisome Homeostasis and Its Role in Stress Response and Signaling in Plants. doi: 10.3389/fpls.2019.00705
- Shultz RW, Tatineni VM, Hanley-bowdoin L and Thompson WF (2007) Genome-Wide Analysis of the Core DNA Replication Machinery in the Higher Plants. 144:1697–1714.
- Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S et al. (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 51:D638–D646.
- Tuteja N, Ahmad P, Panda BB and Tuteja R (2009) Mutation Research / Reviews in Mutation Research Genotoxic stress in plants: Shedding light on DNA damage, repair and DNA repair helicases. 681:134–149.
- Tuteja R and Tuteja N (2013) Plant Physiology and Biochemistry Analysis of DNA repair helicase UvrD from *Arabidopsis thaliana* and *Oryza sativa*. 71:254–260.
- Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB and William S (1999) Crystal Structures of Complexes of PcrA DNA Helicase with a DNA Substrate Indicate an Inchworm Mechanism. 97:75–84.
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, Beer TAP De, Rempfer C, Bordoli L et al. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. 46:296–303.
- Wu J, Liang X, Yan H, Lin M and Lan Y (2023) Comprehensive analysis of MAPK gene family in *Populus trichocarpa* and physiological characterization of PtMAPK3-1 in response to MeJA induction. 1–15.

- Yang W (2010) Lessons Learned from UvrD Helicase: Mechanism for Directional Movement *. doi: 10.1146/annurev.biophys.093008.131415
- Yu F, Wu Y and Xie Q (2016) Ubiquitin – Proteasome System in ABA Signaling: From Perception to Action. MOLP 9:21–33.
- Zhang B, Karnik R, Wang Y, Wallmeroth N, Blatt MR and Grefen C (2015) The Arabidopsis R-SNARE VAMP721 Interacts with KAT1 and KC1 K + Channels to Moderate K + Current at the Plasma Membrane. 27:1697–1717.
- Zhang Y, Zhang A, Li X and Lu C (2020) The Role of Chloroplast Gene Expression in Plant Responses to Environmental Stress.

6 CONSIDERAÇÕES FINAIS

Foram detectados 130 genes codificadores de helicases para pinhão-manso. Esses genes estão distribuídos nas famílias DEAD-box RNA helicases (SF-II) e DNA helicases MCM (SF-VI), RuvB (SF-VI) e UvrD (SF-I), além da família de remodeladores da cromatina Snf2 (SF-II). Análises de caracterização estrutural dessas famílias gênicas permitiram detectar padrões estruturais típicos das helicases. Basicamente, todas as helicases de pinhão-manso compartilham a estrutura do *core* característico que inclui a presença dos motivos conservados *Walker A* e *Walker B* distribuídos ao longo de dois domínios centrais. No geral, helicases de pinhão-manso são altamente diversificadas em termos de estrutura gênica, ocorrência de domínios auxiliares, distribuição de motivos conservados e localização subcelular. Estes resultados demonstram que são proteínas versáteis que atuam em uma variedade de processos biológicos. Modelos tridimensionais construídos para essas proteínas apresentaram características estruturais diretamente relacionadas com as funções exercidas por essas enzimas, apoiando a formação de anéis oligoméricos para helicases de SF-VI e manutenção de estados “aberto” e “fechado” para SF-I e SF-II. Sob o estímulo de salinidade analisado (150 mM de NaCl por 3 horas), plantas de *Jc171* apresentaram perfil de expressão de genes helicase quase totalmente reprimido, com DNA helicases apresentando expressão não significativa. Este resultado foi confirmado para a maioria dos transcritos submetidos a ensaios de validação por qPCR. Redes de interação proteína-proteína considerando as famílias gênicas analisadas mostraram parceiros funcionais envolvidos em uma gama de processos. Dentre os processos associados estão: remodelamento da cromatina; reparo do DNA; *turnover* de RNA e de proteínas; biogênese do ribossomo e de organelas como cloroplastos, mitocôndrias e peroxissomos; tráfego intracelular

através de vesículas; dentre outros. Logo, em virtude da importância dessas helicases no metabolismo de plantas submetidas a estresses, esse perfil de expressão reprimido pode sinalizar a menor habilidade de *Jc171* em modular seu transcriptoma quando submetido a salinidade. Contudo, mais análises são necessárias a fim de melhor entender o envolvimento dessas proteínas nas respostas de *J. curcas* a salinidade. Este é o primeiro relato da descrição de genes codificadores de helicases (DNA e RNA) em *J. curcas*. Os candidatos aqui propostos são promissores para ensaios de caracterização funcional para obtenção de variedades mais tolerantes.

REFERÊNCIAS

- ABDEL-MONEM, Mahmoud; DÜRWARD, Hildegard; HOFFMANN-BERLING, Hartmut. Enzymic unwinding of DNA: 2. Chain Separation by an ATP-Dependent DNA unwinding enzyme. **European Journal of Biochemistry**, v. 65, n. 2, p. 441-449, 1976.
- ABD-ELBAKY, Ahmed A.; GHARIB, Hesham A. Effect of *Jatropha curcas* and *Taxodium distichum* extracts on *Sclerotium cepivorum* the cause of onion white rot. **Egyptian Journal of Agricultural Research**, v. 99, n. 4, p. 397-410, 2021.
- ADAMS, Mark D. et al. Complementary DNA sequencing: expressed sequence tags and human genome project. **Science**, v. 252, n. 5013, p. 1651-1656, 1991.
- AGARWAL, Pinky et al. Expanding frontiers in plant transcriptomics in aid of functional genomics and molecular breeding. **Biotechnology Journal**, v. 9, n. 12, p. 1480-1492, 2014.
- ACHTEN, Wouter MJ et al. Towards domestication of *Jatropha curcas*. **Biofuels**, v. 1, n. 1, p. 91-107, 2010.
- AHMED, A. M. et al. Anti-cancer activity of curcin and latex isolated from *jatropha* plant (*Jatropha curcas* L.). **Journal of Agricultural Chemistry and Biotechnology**, v. 11, n. 11, p. 339-344, 2020.
- ASHBURNER, Michael et al. Gene ontology: tool for the unification of biology. **Nature genetics**, v. 25, n. 1, p. 25-29, 2000.
- AIRAODION, Augustine Ikhueoya; OGBUAGU, Emmanuel Onyebuchi. Antiplasmodial potential of ethanolic leaf extract of *Jatropha curcas* against *Plasmodium berghei*. **Asian Journal of Medical Principles and Clinical Practice**, v. 3, n. 3, p. 29-36, 2020.
- AREKEMASE, M. O.; KAYODE, R. M. O.; AJIBOYE, A. E. Antimicrobial activity and phytochemical analysis of *Jatropha curcas* plant against some selected microorganisms. **International Journal of Biology**, v. 3, n. 3, p. 52, 2011.
- ARYA, Manit et al. Basic principles of real-time quantitative PCR. **Expert review of molecular diagnostics**, v. 5, n. 2, p. 209-219, 2005.
- AKANDE, O. A. et al. In-vivo anti-plasmodial activity of methanolic leaf extracts of *Jatropha curcas* on *Plasmodium berghei* (NK 65) infected mice. **Journal of Phytomedicine and Therapeutics**, v. 20, n. 2, p. 702-713, 2022.
- ALVES, José Maria Arcanjo et al. Pinhão-mansão: uma alternativa para produção de biodiesel na agricultura familiar da Amazônia brasileira. **Revista Agro@mbiente online**, v. 2, n. 1, p. 57-68, 2008.
- AMMELBURG, Moritz; FRICKEY, Tancred; LUPAS, Andrei N. Classification of AAA+ proteins. **Journal of structural biology**, v. 156, n. 1, p. 2-11, 2006.

AN, Junxia; LI, Lifei; ZHANG, Xuehong. Curcusone C induces apoptosis in endometrial cancer cells via mitochondria-dependent apoptotic and ERK pathway. **Biotechnology Letters**, v. 43, p. 329-338, 2021.

ARANSIOLA, Elizabeth Funmilayo et al. A review of current technology for biodiesel production: State of the art. **Biomass and bioenergy**, v. 61, p. 276-297, 2014.

AHUCHAOGU, C. E.; OJIAKO, F. O.; KABEH, J. D. Evaluation of Jatropha curcas Lam. extracts in the control of some field insect pests of cowpea (*Vigna unguiculata* L. Walp). **International Journal of Scientific & Technology Research**, v. 3, n. 3, p. 136-141, 2014.

ASUK, Atamgba Agbor et al. Jatropha curcas Leaf extract and fractions attenuate hyperglycemia, tissue oxidation, and kidney dysfunction in diabetic rats. **Journal of Biochemical Technology**, v. 12, n. 3-2021, p. 84-90, 2021.

ANDREWS, Simon et al. FastQC: a quality control tool for high throughput sequence data. 2010.

AUVIN, Catherine et al. Curcacycline B, a cyclic nonapeptide from *Jatropha curcas* enhancing rotamase activity of cyclophilin. **Tetrahedron Letters**, v. 38, n. 16, p. 2845-2848, 1997.

AWOTEDU, Olamilekan L.; OGUNBAMOWO, Paul O. Comparative heavy metal uptake and phytoremediation potential of three jatropha species. **Environment & Ecosystem Science**, v. 3, n. 2, p. 26-30, 2019.

BABAHMAD, Rachid Ait et al. Chemical composition of essential oil of *Jatropha curcas* L. leaves and its antioxidant and antimicrobial activities. **Industrial crops and products**, v. 121, p. 405-410, 2018.

BHATTACHARYYA, Badal et al. Effect of physic nut oil against potato cutworm, *Agrotis ipsilon*. **Journal of Entomological Research**, v. 46, n. 1, p. 73-76, 2022.

BAILEY, Timothy L. et al. MEME SUITE: tools for motif discovery and searching. **Nucleic acids research**, v. 37, n. suppl_2, p. W202-W208, 2009.

BANKEVICH, Anton et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of computational biology**, v. 19, n. 5, p. 455-477, 2012.

BAO, Weidong; KOJIMA, Kenji K.; KOHANY, Oleksiy. Repbase Update, a database of repetitive elements in eukaryotic genomes. **Mobile Dna**, v. 6, p. 1-6, 2015.

BAO, Zhirong; EDDY, Sean R. Automated de novo identification of repeat sequence families in sequenced genomes. **Genome research**, v. 12, n. 8, p. 1269-1276, 2002.

BASU, Udita et al. The DNA replication, repair, and recombination pathway genes modulating yield and stress tolerance traits in chickpea. **Plant Molecular Biology Reporter**, p. 1-17, 2021.

BASTOS, Eduardo Muniz Santana et al. Anti-inflammatory activity of *Jatropha curcas* L. in brain glial cells primary cultures. **Journal of Ethnopharmacology**, v. 264, p. 113201, 2021.

BAUMERT, Sophia; KHAMZINA, Asia; VLEK, Paul LG. Soil organic carbon sequestration in *Jatropha curcas* systems in Burkina Faso. **Land Degradation & Development**, v. 27, n. 8, p. 1813-1819, 2016.

BENEVENUTO, Juliana et al. How can a high-quality genome assembly help plant breeders?. **Gigascience**, v. 8, n. 6, p. giz068, 2019.

VAN DEN BERG, A. J. J. et al. Curcacycline A—a novel cyclic octapeptide isolated from the latex of *Jatropha curcas* L. **Febs Letters**, v. 358, n. 3, p. 215-218, 1995.

BESSIKE, Josias Georges et al. Evaluation of the potentials of *Jatropha curcas* seed oil and in combination with leaf extracts of *Cymbopogon citratus*, *Ocimum basilicum*, and *Eucalyptus globulus* as wood preservatives against *Macrotermes bellicosus* termites. **Industrial Crops and Products**, v. 195, p. 116205, 2023.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.

BORAH, Pallabi et al. Phytoremediation of nickel and zinc using *Jatropha curcas* and *Pongamia pinnata* from the soils contaminated by municipal solid wastes and paper mill wastes. **Environmental Research**, v. 219, p. 115055, 2023.

BURGE, Chris; KARLIN, Samuel. Prediction of complete gene structures in human genomic DNA. **Journal of molecular biology**, v. 268, n. 1, p. 78-94, 1997.

BUSHMANOVA, Elena et al. rnaQUAST: a quality assessment tool for de novo transcriptome assemblies. **Bioinformatics**, v. 32, n. 14, p. 2210-2212, 2016.

BUZA, Krisztian et al. RECORD: reference-assisted genome assembly for closely related genomes. **International journal of genomics**, v. 2015, 2015.

CARBON, Seth et al. AmiGO: online access to ontology and annotation data. **Bioinformatics**, v. 25, n. 2, p. 288-289, 2009.

CARELS, Nicolas. *Jatropha curcas*: a review. **Advances in botanical research**, v. 50, p. 39-86, 2009.

CHAN, Kuang-Lim et al. Seqping: gene prediction pipeline for plant genomes using self-training gene models and transcriptomic data. **BMC bioinformatics**, v. 18, p. 1-7, 2017.

CHEN, Ying et al. Fast and accurate assembly of Nanopore reads via progressive error correction and adaptive read selection. **BioRxiv**, p. 2020.02. 01.930107, 2020..

CHEN, Nansheng. Using Repeat Masker to identify repetitive elements in genomic sequences. **Current protocols in bioinformatics**, v. 5, n. 1, p. 4.10. 1-4.10. 14, 2004.

CHENG, Feng et al. BRAD, the genetics and genomics database for Brassica plants. **BMC plant biology**, v. 11, p. 1-6, 2011.

CHEN, Chengjie et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. **Molecular plant**, v. 13, n. 8, p. 1194-1202, 2020.

CHEN, M. J.; HOU, L. L.; ZHANG, G. W. The diterpenoids from *Jatropha curcas* L. 1988.

CHIN, Chen-Shan et al. Phased diploid genome assembly with single-molecule real-time sequencing. **Nature methods**, v. 13, n. 12, p. 1050-1054, 2016.

CHINNUSAMY, Viswanathan; GONG, Z.; ZHU, J.-K. Nuclear RNA export and its importance in abiotic stress responses of plants. **Nuclear Pre-Mrna Processing in Plants**, p. 235-255, 2008.

CHOW, Chi-Nga et al. PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. **Nucleic acids research**, v. 44, n. D1, p. D1154-D1160, 2016.

CONESA, Ana et al. A survey of best practices for RNA-seq data analysis. **Genome biology**, v. 17, p. 1-19, 2016.

ANA CONESA, S. G. Blast2GO: a comprehensive suite for functional analysis in plant genomics. **Int J Plant Gen**, v. 2008, p. 1-12, 2008.

CORDIN, Olivier et al. The DEAD-box protein family of RNA helicases. **Gene**, v. 367, p. 17-37, 2006.

DANG, Hung Quang et al. A single subunit MCM6 from pea promotes salinity stress tolerance without affecting yield. **Plant molecular biology**, v. 76, p. 19-34, 2011.

DHAKA, Sunita; BHATT, Jayant; CHOUDHARY, Anand. Ecofriendly Management of Early Blight of Tomato Using Plant Extracts and Oil Cakes. **Biopesticides International**, v. 18, n. 1, 2022.

DEHGAN, Bijan. Phylogenetic significance of interspecific hybridization in *Jatropha* (Euphorbiaceae). **Systematic Botany**, p. 467-478, 1984.

DENTI, Luca et al. ASGAL: aligning RNA-Seq data to a splicing graph to detect novel alternative splicing events. **BMC bioinformatics**, v. 19, p. 1-21, 2018.

DIAS, LA dos S.; MISSIO, Robson Fernando; DIAS, D. Cunha FS. Antiquity, botany, origin and domestication of *Jatropha curcas* (Euphorbiaceae), a plant species with potential for biodiesel production. 2012.

DOBIN, Alexander et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics**, v. 29, n. 1, p. 15-21, 2013.

DOS SANTOS, Tiago Benedito et al. Physiological responses to drought, salinity, and heat stress in plants: a review. **Stresses**, v. 2, n. 1, p. 113-135, 2022.

DOUGHARI, James Hamuel; ABRAHAM, Martha. Antifungal activity of *Jatropha Curcas* Linn on *Candida albicans* and *Candida tropicalis* associated with neonatal and infantile infections in Yola, Nigeria. **Am. J. Agric. Biol. Sci**, v. 16, p. 19-32, 2021.

DUVICK, Jon et al. PlantGDB: a resource for comparative plant genomics. **Nucleic acids research**, v. 36, n. suppl_1, p. D959-D965, 2007.

EMMS, David M.; KELLY, Steven. OrthoFinder: phylogenetic orthology inference for comparative genomics. **Genome biology**, v. 20, p. 1-14, 2019.

EVANS, Elizabeth et al. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. **The EMBO journal**, 1997.

EWING, Brent et al. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. **Genome research**, v. 8, n. 3, p. 175-185, 1998.

EWUNIE, Gebresilassie Asnake et al. Factors affecting the potential of *Jatropha curcas* for sustainable biodiesel production: A critical review. **Renewable and Sustainable Energy Reviews**, v. 137, p. 110500, 2021.

FAIRMAN-WILLIAMS, Margaret E.; GUENTHER, Ulf-Peter; JANKOWSKY, Eckhard. SF1 and SF2 helicases: family matters. **Current opinion in structural biology**, v. 20, n. 3, p. 313-324, 2010.

FERNANDEZ-POZO, Noe et al. The Sol Genomics Network (SGN)—from genotype to phenotype to breeding. **Nucleic acids research**, v. 43, n. D1, p. D1036-D1041, 2015.

DE SOUSA FERRÃO, Elvino Jos&e; JANEQUE, Edilson Armando De Germano. Anti-viral compounds from *Jatropha curcas* seed extract with anti-HIV-1 and anti-SARS-CoV-2 action. **African Journal of Pharmacy and Pharmacology**, v. 17, n. 1, p. 1-9, 2023.

FINN, Robert D. et al. The Pfam protein families database: towards a more sustainable future. **Nucleic acids research**, v. 44, n. D1, p. D279-D285, 2016.

FLYNN, Jullien M. et al. RepeatModeler2 for automated genomic discovery of transposable element families. **Proceedings of the National Academy of Sciences**, v. 117, n. 17, p. 9451-9457, 2020.

FRANCIS, Magreth et al. Phytochemical analysis and in vitro antifungal evaluation of *Jatropha curcas* against Late Leaf Spot disease on groundnut. 2021.

FRÖSCHLE, M.; HORN, H.; SPRING, O. Characterization of *Jatropha curcas* honeys originating from the southern highlands of Madagascar. **LWT**, v. 93, p. 525-533, 2018.

FUKUSHIMA, Atsushi et al. Integrated omics approaches in plant systems biology. **Current opinion in chemical biology**, v. 13, n. 5-6, p. 532-538, 2009.

FULLER-PACE, Frances V. DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. **Nucleic acids research**, v. 34, n. 15, p. 4206-4215, 2006.

GENDLER, Karla; PAULSEN, Tara; NAPOLI, Carolyn. ChromDB: the chromatin database. **Nucleic acids research**, v. 36, n. suppl_1, p. D298-D302, 2008.

GASTEIGER, Elisabeth et al. **Protein identification and analysis tools on the ExPASy server**. Humana press, 2005.

GNERRE, Sante et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. **Proceedings of the National Academy of Sciences**, v. 108, n. 4, p. 1513-1518, 2011.

GILL, Sarvajeet Singh et al. A DESD-box helicase functions in salinity stress tolerance by improving photosynthesis and antioxidant machinery in rice (*Oryza sativa* L. cv. PB1). **Plant molecular biology**, v. 82, p. 1-22, 2013.

GOODSTEIN, David M. et al. Phytozome: a comparative platform for green plant genomics. **Nucleic acids research**, v. 40, n. D1, p. D1178-D1186, 2012.

GOMES, Taisa G. et al. Current strategies for the detoxification of *Jatropha curcas* seed cake: a review. **Journal of agricultural and food chemistry**, v. 66, n. 11, p. 2510-2522, 2018.

GORDON, Assaf. Fastx-toolkit. **FASTQ/A short-reads preprocessing tools.**, v. 5, 2010.

GRABHERR, Manfred G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature biotechnology**, v. 29, n. 7, p. 644-652, 2011.

GRIBSKOV, Michael et al. PlantsP: a functional genomics database for plant phosphorylation. **Nucleic acids research**, v. 29, n. 1, p. 111-113, 2001.

GUPTA, Parul et al. Gramene database: navigating plant comparative genomics resources. **Current plant biology**, v. 7, p. 10-15, 2016.

GUREVICH, Alexey et al. QUAST: quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072-1075, 2013.

GUTIÉRREZ-LÓPEZ, Adriana N. et al. Green and fast biodiesel production at room temperature using soybean and *Jatropha curcas* L. oils catalyzed by potassium ferrate. **Journal of Cleaner Production**, v. 372, p. 133739, 2022.

HAO, Shanhu et al. A review on plant responses to salt stress and their mechanisms of salt resistance. **Horticulturae**, v. 7, n. 6, p. 132, 2021.

HEID, Christian A. et al. Real time quantitative PCR. **Genome research**, v. 6, n. 10, p. 986-994, 1996.

HENDRIX, Sophie et al. Suppressor of gamma response 1 modulates the DNA damage response and oxidative stress response in leaves of cadmium-exposed *Arabidopsis thaliana*. **Frontiers in Plant Science**, v. 11, p. 508457, 2020.

HELLER, Joachim. **Physic nut, *Jatropha curcas* L.** Bioversity International, 1996.

HILLER, Karsten et al. JVirGel 2.0: computational prediction of proteomes separated via two-dimensional gel electrophoresis under consideration of membrane and secreted proteins. **Bioinformatics**, v. 22, n. 19, p. 2441-2443, 2006.

HOFF, Katharina J. et al. BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. **Bioinformatics**, v. 32, n. 5, p. 767-769, 2016.

HOSSAIN, Mohammad Anwar et al. **Drought stress tolerance in plants, Volume.** 2016.

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

HU, Haifei; SCHEBEN, Armin; EDWARDS, David. Advances in integrating genomics and bioinformatics in the plant breeding pipeline. **Agriculture**, v. 8, n. 6, p. 75, 2018.

HUANG, Xiaoqiu; MADAN, Anup. CAP3: A DNA sequence assembly program. **Genome research**, v. 9, n. 9, p. 868-877, 1999.

HUSELID, Eric; BUNTING, Samuel F. The regulation of homologous recombination by helicases. **Genes**, v. 11, n. 5, p. 498, 2020.

HURGOBIN, Bhavna. Short read alignment using SOAP2. **Plant Bioinformatics: Methods and Protocols**, p. 241-252, 2016.

INGLE, Krishnananda P. et al. Screening of insecticidal activity of *Jatropha Curcas* (L.) against diamond back moth and *Helicoverpa Armigera*. **Seed**, v. 5, n. 80, p. 20, 2017.

IFEANYIEZE, F. O. et al. Seed powder extract of physic nut (*Jatropha curcas*) as a biopesticide for weevils (*Callosobruchus maculatus*) in stored cowpea (*Vigna unguiculata*). **African Journal of Science, Technology, Innovation and Development**, v. 14, n. 7, p. 1864-1869, 2022.

ISKEN, Olaf et al. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. **Cell**, v. 133, n. 2, p. 314-327, 2008.

JALALI, Saakshi et al. Exploitation of Hi-C sequencing for improvement of genome assembly and in-vitro validation of differentially expressing genes in *Jatropha curcas* L. **3 Biotech**, v. 10, p. 1-9, 2020.

JARMOSKAITE, Inga; RUSSELL, Rick. RNA helicase proteins as chaperones and remodelers. **Annual review of biochemistry**, v. 83, p. 697-725, 2014.

JIN, Jinpu et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. **Nucleic acids research**, p. gkw982, 2016.

KHADIJA, Muktar Sukola et al. Phytochemical Analysis and Anti-trypanosomal Efficacy of the Methanol Leaf Extract of *Jatropha curcas*.

KALININA, Natalia O. et al. The multiple functions of the nucleolus in plant development, disease and stress responses. **Frontiers in plant science**, v. 9, p. 132, 2018.

KAMEL, Dena A. et al. Smart utilization of jatropha (*Jatropha curcas* Linnaeus) seeds for biodiesel production: Optimization and mechanism. **Industrial crops and products**, v. 111, p. 407-413, 2018.

KANCHARLA, Nagesh et al. De novo sequencing and hybrid assembly of the biofuel crop *Jatropha curcas* L.: identification of quantitative trait loci for geminivirus resistance. **Genes**, v. 10, n. 1, p. 69, 2019.

KANEHISA, Minoru et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. **Nucleic acids research**, v. 45, n. D1, p. D353-D361, 2017.

KAWAHARA, Yoshihiro et al. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. **Rice**, v. 6, p. 1-10, 2013.

KERSEY, Yoshihiro et al. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. **Rice**, v. 6, p. 1-10, 2013.

KIM, Daehwan et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome biology**, v. 14, p. 1-13, 2013.

KOH, May Ying; GHAZI, Tinia Idaty Mohd. A review of biodiesel production from *Jatropha curcas* L. oil. **Renewable and sustainable energy reviews**, v. 15, n. 5, p. 2240-2251, 2011.

KOPECKÁ, Romana et al. Abiotic stress in crop production. **International Journal of Molecular Sciences**, v. 24, n. 7, p. 6603, 2023.

KOREN, Sergey et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. **Genome research**, v. 27, n. 5, p. 722-736, 2017.

KUMAR, Rajneesh; DAS, Niranjan. Seed oil of *Jatropha curcas* L. germplasm: Analysis of oil quality and fatty acid composition. **Industrial crops and products**, v. 124, p. 663-668, 2018.

KWODAGA, Joseph Kwowura; SOWLEY, Elias Nortaa Kunedeb; BADII, Benjamin Kongyeli. Antifungal activity of plant extracts against *Colletotrichum gloeosporioides* (Penz.) the causative agent of yam anthracnose disease. **Archives of Phytopathology and Plant Protection**, v. 52, n. 1-2, p. 218-233, 2019.

KYRIAKIDOU, Maria et al. Current strategies of polyploid plant genome sequence assembly. **Frontiers in plant science**, v. 9, p. 385637, 2018.

LANGMEAD, Ben; SALZBERG, Steven L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, v. 9, n. 4, p. 357-359, 2012.

LARKIN, Mark A. et al. Clustal W and Clustal X version 2.0. **bioinformatics**, v. 23, n. 21, p. 2947-2948, 2007.

LAVIOLA, Bruno Galvêas et al. Desempenho agrônômico e ganho genético pela seleção de pinhão-mansão em três regiões do Brasil. **Pesquisa Agropecuária Brasileira**, v. 49, n. 5, p. 356-363, 2014.

LAWRENCE, Carolyn J. et al. MaizeGDB, the community database for maize genetics and genomics. **Nucleic acids research**, v. 32, n. suppl_1, p. D393-D397, 2004.

LEIPE, Detlef D.; KOONIN, Eugene V.; ARAVIND, LJJOMB. Evolution and classification of P-loop kinases and related proteins. **Journal of molecular biology**, v. 333, n. 4, p. 781-815, 2003.

LERROY, Philippe et al. TriAnnot: a versatile and high performance pipeline for the automated annotation of plant genomes. **Frontiers in plant science**, v. 3, p. 5, 2012.

LESCOT, Magali et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic acids research**, v. 30, n. 1, p. 325-327, 2002.

LETUNIC, Ivica; KHEDKAR, Supriya; BORK, Peer. SMART: recent updates, new developments and status in 2020. **Nucleic acids research**, v. 49, n. D1, p. D458-D460, 2021.

LETUNIC, Ivica; BORK, Peer. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. **Nucleic acids research**, v. 47, n. W1, p. W256-W259, 2019.

LI, Heng; DURBIN, Richard. Fast and accurate short read alignment with Burrows–Wheeler transform. **bioinformatics**, v. 25, n. 14, p. 1754-1760, 2009.

LI, Xindi et al. Functions and mechanisms of RNA helicases in plants. **Journal of Experimental Botany**, v. 74, n. 7, p. 2295-2310, 2023.

LIEBERMAN-AIDEN, Erez et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. **science**, v. 326, n. 5950, p. 289-293, 2009.

LINDER, Patrick; OWTTRIM, George W. Plant RNA helicases: linking aberrant and silencing RNA. **Trends in plant science**, v. 14, n. 6, p. 344-352, 2009.

LIVAK, Kenneth J.; SCHMITTGEN, Thomas D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta CT$ method. **methods**, v. 25, n. 4, p. 402-408, 2001.

LOMSADZE, Alexandre et al. Gene identification in novel eukaryotic genomes by self-training algorithm. **Nucleic acids research**, v. 33, n. 20, p. 6494-6506, 2005.

LONGO, Liam M. et al. On the emergence of P-Loop NTPase and Rossmann enzymes from a Beta-Alpha-Beta ancestral fragment. **Elife**, v. 9, p. e64415, 2020.

LOVE, Michael I.; HUBER, Wolfgang; ANDERS, Simon. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome biology**, v. 15, p. 1-21, 2014.

MAGHULY, Fatemeh et al. Geographic origin is not supported by the genetic variability found in a large living collection of *Jatropha curcas* with accessions from three continents. **Biotechnology Journal**, v. 10, n. 4, p. 536-551, 2015.

MANCHANDA, Nancy et al. GenomeQC: a quality assessment tool for genome assemblies and gene structure annotations. **BMC genomics**, v. 21, p. 1-9, 2020.

MANJULATHA, M. et al. Overexpression of a pea DNA helicase (PDH45) in peanut (*Arachis hypogaea* L.) confers improvement of cellular level tolerance and productivity under drought stress. **Molecular biotechnology**, v. 56, p. 111-125, 2014.

MACHADO, Rui Manuel Almeida; SERRALHEIRO, Ricardo Paulo. Soil salinity: effect on vegetable crop growth. Management practices to prevent and mitigate soil salinization. **Horticulturae**, v. 3, n. 2, p. 30, 2017.

MARCHLER-BAUER, Aron et al. CDD: NCBI's conserved domain database. **Nucleic acids research**, v. 43, n. D1, p. D222-D226, 2015.

MAROLI, Amith S. et al. Omics in weed science: a perspective from genomics, transcriptomics, and metabolomics approaches. **Weed Science**, v. 66, n. 6, p. 681-695, 2018.

MANSOOR, Sheikh et al. Reactive oxygen species in plants: from source to sink. **Antioxidants**, v. 11, n. 2, p. 225, 2022..

MATSUMURA, Hideo et al. SuperSAGE: a modern platform for genome-wide quantitative transcript profiling. **Current Pharmaceutical Biotechnology**, v. 9, n. 5, p. 368-374, 2008.

MATSUMURA, Hideo et al. High-throughput SuperSAGE for digital gene expression analysis of multiple samples using next generation sequencing. **PloS one**, v. 5, n. 8, p. e12010, 2010.

MAZUMDAR, Purabi et al. An update on biological advancement of *Jatropha curcas* L.: New insight and challenges. **Renewable and Sustainable Energy Reviews**, v. 91, p. 903-917, 2018.

MEHTA, Dhruvisha; VYAS, Suhas. Comparative bio-accumulation of osmoprotectants in saline stress tolerating plants: A review. **Plant Stress**, p. 100177, 2023.

MENSAH, Martin Kofi et al. Immobilization effects of co-pyrolyzed neem seed mixed with poultry manure on potentially toxic elements in soil and the phytoremediation potentials of native *Manihot esculenta* and *Jatropha curcas* in ensuring sustainable land use. **Environmental Monitoring and Assessment**, v. 195, n. 6, p. 793, 2023.

MITRA, Sudip et al. A review on environmental and socioeconomic perspectives of three promising biofuel plants *Jatropha curcas*, *Pongamia pinnata* and *Mesua ferrea*. **Biomass and Bioenergy**, v. 151, p. 106173, 2021.

MISHRA, Udit Nandan et al. Emerging Roles of Osmoprotectants in Response to Multiple Abiotic Stress Tolerance in Plants. In: **Omics Analysis of Plants under Abiotic Stress**. Apple Academic Press, 2022. p. 179-206.

MISHRA, Ajay Shankar et al. Antimicrobial Activities of different plant extracts of *Jatropha Curcas*. **Bulletin of Environmental and Scientific Research**, v. 4, n. 1-2, p. 21-28, 2015.

MOHAMMED, Abdullahi Tanko et al. Soil fertility enrichment potential of *Jatropha curcas* for sustainable agricultural production: A case study of Birnin Kebbi, Nigeria. **Annals of the Romanian Society for Cell Biology**, p. 21061-21073, 2021.

MONTES, Juan M.; MELCHINGER, Albrecht E. Domestication and Breeding of *Jatropha curcas* L. **Trends in plant science**, v. 21, n. 12, p. 1045-1057, 2016.

MORTAZAVI, Ali et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature methods**, v. 5, n. 7, p. 621-628, 2008.

MUHAMMAD, Isiaka Ibrahim et al. RNA-seq and ChIP-seq as complementary approaches for comprehension of plant transcriptional regulatory mechanism. **International journal of molecular sciences**, v. 21, n. 1, p. 167, 2019.

MULLIS, Kary et al. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In: **Cold Spring Harbor symposia on quantitative biology**. Cold Spring Harbor Laboratory Press, 1986. p. 263-273.

MUMPUNI, Restu Puji; SHOLICHUDIN, Akhmad. The Efficacy of Several Types of Organic Pesticides against Mortality of Armyworm (*Spodoptera frugiperda*). **Advances in Science and Technology**, v. 112, p. 113-118, 2022.

MUNIZ, D. R. et al. Biocide potential of *Jatropha curcas* L. extracts. **Journal of Biology and Life Science**, v. 11, n. 2, p. 138-154, 2020.

NATH, Manoj et al. PDH45 overexpressing transgenic tobacco and rice plants provide salinity stress tolerance via less sodium accumulation. **Plant signaling & behavior**, v. 10, n. 4, p. e992289, 2015.

NATH, Manoj et al. PDH45 transgenic rice maintain cell viability through lower accumulation of Na⁺, ROS and calcium homeostasis in roots under salinity stress. **Journal of plant physiology**, v. 191, p. 1-11, 2016.

NEALE, David B. et al. Novel insights into tree biology and genome evolution as revealed through genomics. **Annual Review of Plant Biology**, v. 68, p. 457-483, 2017.

NERO, Bertrand Festus. Phytoremediation of petroleum hydrocarbon-contaminated soils with two plant species: *Jatropha curcas* and *Vetiveria zizanioides* at Ghana Manganese Company Ltd. **International Journal of Phytoremediation**, v. 23, n. 2, p. 171-180, 2021.

OGWUDIRE, V. E. et al. Phyto-toxicity of *Jatropha curcas* L. Root Extracts on the hatchability of Root-knot Nematode *Meloidogyne incognita*. **Futo Journal Series (FUTOJNLS)**. e-ISSN, p. 2476-8456.

OHOUEU, Ehouman Jean Brice et al. Effect of aqueous extracts of *Azadirachta indica* A. Juss, *Jatropha curcas* L. and *Moringa oleifera* Lam. on coffee berry borer (*Hypothenemus hampei* F.; Coleoptera: Scolytidae) in laboratory. **International Journal of Biological and Chemical Sciences**, v. 16, n. 5, p. 2289-2301, 2022.

OPUBA, Sinkopere Kenneth et al. Insecticidal and anti-ovipositional activities of the leaf powder of *Jatropha curcas* (L.)(euphorbiaceae) against *Callosobruchus maculatus* (F.)(Coleoptera: Chrysomelidae). **Animal Research International**, v. 15, n. 1, p. 2971–2978-2971–2978, 2018.

ONU, Organização das Nações Unidas. Disponível em: <<https://brasil.un.org/>>. Acesso em 16 Jul. 2023.

OUYANG, Shu; BUELL, C. Robin. The TIGR Plant Repeat Databases: a collective resource for the identification of repetitive sequences in plants. **Nucleic acids research**, v. 32, n. suppl_1, p. D360-D363, 2004.

OUYANG, Shu et al. Plant genome annotation methods. **Plant Genomics: Methods and Protocols**, p. 263-282, 2009.

OYETAYO, Victor Olusegun; TEMENU, O. E. Comparative Study on The Phytochemical, Antistaphylococcal and Antioxidant Properties of the Stem Bark of *Jatropha curcas* L and *Jatropha gossypifolia* L: Comparative study on two *Jatropha* species. **Sultan Qaboos University Journal for Science [SQUJS]**, v. 28, n. 1, p. 17-24, 2023.

PANDEY, Saurabh et al. Linking the plant stress responses with RNA helicases. **Plant science**, v. 299, p. 110607, 2020.

PATEL, Krupal Kanubhai; NARASIMHACHARYA, A. V. R. L. Potentiality of phytochemical compounds isolated and identified from *Jatropha curcas* L. against *Odontotermes obesus* Rambur. **Journal of Natural Pesticide Research**, v. 2, p. 100009, 2022.

POP, Mihai. Genome assembly reborn: recent computational challenges. **Briefings in bioinformatics**, v. 10, n. 4, p. 354-366, 2009.

POP, Mihai et al. Comparative genome assembly. **Briefings in bioinformatics**, v. 5, n. 3, p. 237-248, 2004.

PRABHU, A.; VENKATA RAMANAN, M.; JAYAPRABAKAR, J. Production, properties and engine characteristics of *Jatropha* biodiesel—a review. **International Journal of Ambient Energy**, v. 42, n. 15, p. 1810-1814, 2021.

PRASAD, DM Reddy et al. *Jatropha curcas*: Plant of medical benefits. **Journal of medicinal plants research**, v. 6, n. 14, p. 2691-2699, 2012.

PREMALATHA, K. et al. Acaricidal activity of plant extracts on two spotted spider mite, *Tetranychus urticae* Koch (Acari: Tetranychidae). **Journal of Entomology and Zoology Studies**, v. 6, n. 1, p. 1622-1625, 2018.

PRICE, Alkes L.; JONES, Neil C.; PEVZNER, Pavel A. De novo identification of repeat families in large genomes. **Bioinformatics**, v. 21, n. suppl_1, p. i351-i358, 2005.

PROJECT, International Rice Genome Sequencing. The map-based sequence of the rice genome. **Nature**, v. 436, n. 7052, p. 793-800, 2005.

RAIKWAR, Shailendra et al. Emerging importance of helicases in plant stress tolerance: characterization of *Oryza sativa* repair helicase XPB2 promoter and its functional validation in tobacco under multiple stresses. **Frontiers in plant science**, v. 6, p. 1094, 2015.

RAHIM, F. A. A.; HAMID, T. H. T. A.; ZAINUDDIN, Z. *Jatropha curcas* as a potential plant for bauxite phytoremediation. In: **IOP Conference Series: Earth and Environmental Science**. IOP Publishing, 2019. p. 012006.

RAHU, Muhammad Idrees et al. Determination of antimicrobial and phytochemical compounds of *Jatropha curcas* plant. **Saudi journal of biological sciences**, v. 28, n. 5, p. 2867-2876, 2021.

RAMADHANA, Syahrul et al. Physicochemical study of castor leaf (*Jatropha curcas* L.) and papaya leaf (*Carica papaya* L.) extract and their application as photosensitizer agents in antimicrobial photodynamic therapy system to *Staphylococcus epidermidis* biofilm. In: **AIP Conference Proceedings**. AIP Publishing, 2023.

RAMPADARATH, Sillma; PUCHOOA, Daneshwar; JEEWON, Rajesh. *Jatropha curcas* L: Phytochemical, antimicrobial and larvicidal properties. **Asian Pacific Journal of Tropical Biomedicine**, v. 6, n. 10, p. 858-865, 2016.

REUBENS, Bert et al. More than biofuel? *Jatropha curcas* root system symmetry and potential for soil erosion control. **Journal of Arid Environments**, v. 75, n. 2, p. 201-205, 2011.

RHEE, Seung Yon et al. The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. **Nucleic acids research**, v. 31, n. 1, p. 224-228, 2003.

CARVALHO, Carlos Roberto et al. Genome size, base composition and karyotype of *Jatropha curcas* L., an important biofuel plant. **Plant Science**, v. 174, n. 6, p. 613-617, 2008.

RIVERO, Rosa M. et al. Developing climate-resilient crops: improving plant tolerance to stress combination. **The Plant Journal**, v. 109, n. 2, p. 373-389, 2022.

ROBINSON, Mark D.; MCCARTHY, Davis J.; SMYTH, Gordon K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **bioinformatics**, v. 26, n. 1, p. 139-140, 2010.

ROCAK, Sanda; LINDER, Patrick. DEAD-box proteins: the driving forces behind RNA metabolism. **Nature reviews Molecular cell biology**, v. 5, n. 3, p. 232-241, 2004.

RODRÍGUEZ-GARCÍA, Cecilia Mónica et al. Antioxidant, antihypertensive, anti-hyperglycemic, and antimicrobial activity of aqueous extracts from twelve native plants of the Yucatan coast. **PLoS One**, v. 14, n. 3, p. e0213493, 2019.

SACHDEV, Swati et al. Abiotic stress and reactive oxygen species: Generation, signaling, and defense mechanisms. **Antioxidants**, v. 10, n. 2, p. 277, 2021.

SAHA, Saurabh et al. Using the transcriptome to annotate the genome. **Nature biotechnology**, v. 20, n. 5, p. 508-512, 2002.

SAHOO, Ranjan Kumar et al. OsSUV3 transgenic rice maintains higher endogenous levels of plant hormones that mitigates adverse effects of salinity and sustains crop productivity. **Rice**, v. 7, p. 1-3, 2014.

SAHOO, Ranjan Kumar; GILL, Sarvajeet Singh; TUTEJA, Narendra. Pea DNA helicase 45 promotes salinity stress tolerance in IR64 rice with improved yield. **Plant signaling & behavior**, v. 7, n. 8, p. 1042-1046, 2012.

SAFDAR, Hamna et al. A review: Impact of salinity on plant growth. **Nat. Sci**, v. 17, n. 1, p. 34-40, 2019.

SAIFI, Shabnam K. et al. Stress-induced *Oryza sativa* RuvBL1a is DNA-independent ATPase and unwinds DNA duplex in 3' to 5' direction. **Protoplasma**, v. 255, p. 669-684, 2018.

SAIFI, Shabnam K. et al. Stress-induced *Oryza sativa* RuvBL1a is DNA-independent ATPase and unwinds DNA duplex in 3' to 5' direction. **Protoplasma**, v. 255, p. 669-684, 2018.

SANAN-MISHRA, Neeti et al. Pea DNA helicase 45 overexpression in tobacco confers high salinity tolerance without affecting yield. **Proceedings of the National Academy of Sciences**, v. 102, n. 2, p. 509-514, 2005.

SAN SEGUNDO-VAL, Ignacio; SANZ-LOZANO, Catalina S. Introduction to the gene expression analysis. **Molecular genetics of asthma**, p. 29-43, 2016.

SAMPSON, B. J. et al. Insecticidal Activity of *Jatropha* Extracts Against the Azalea Lace Bug, *Stephanitis pyrioides* (Hemiptera: Tingidae). **Journal of Economic Entomology**, v. 116, n. 1, p. 192-201, 2023.

SARUHASHI, Masashi et al. Plant Raf-like kinase integrates abscisic acid and hyperosmotic stress signaling upstream of SNF1-related protein kinase2. **Proceedings of the National Academy of Sciences**, v. 112, n. 46, p. E6388-E6396, 2015.

SATO, Shusei et al. Sequence analysis of the genome of an oil-bearing tree, *Jatropha curcas* L. **DNA research**, v. 18, n. 1, p. 65-76, 2011.

SCHENA, Mark et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science**, v. 270, n. 5235, p. 467-470, 1995.

SCHMIDT, Maximilian H.-W. et al. De novo assembly of a new *Solanum pennellii* accession using nanopore sequencing. **The Plant Cell**, v. 29, n. 10, p. 2336-2348, 2017.

SCHULZ, Marcel H. et al. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. **Bioinformatics**, v. 28, n. 8, p. 1086-1092, 2012.

SEGURA-CAMPOS, Maira R. et al. Effect of *Jatropha curcas* peptide fractions on the angiotensin I-converting enzyme inhibitory activity. **BioMed Research International**, v. 2013, 2013.

KUMAR, Vinay et al. **Salinity Responses and Tolerance in Plants, Volume**. Springer, Switzerland, 2018.

SHAHMURADOV, Ilham A. et al. PlantProm: a database of plant promoter sequences. **Nucleic acids research**, v. 31, n. 1, p. 114-117, 2003.

SHEN, Shihao et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. **Proceedings of the National Academy of Sciences**, v. 111, n. 51, p. E5593-E5601, 2014.

SHIMKETS, Richard A. (Ed.). **Gene expression profiling**. Springer Science & Business Media, 2008.

SHIVAKUMARA, Tagginahalli N. et al. Overexpression of Pea DNA Helicase 45 (PDH45) imparts tolerance to multiple abiotic stresses in chili (*Capsicum annuum* L.). **Scientific reports**, v. 7, n. 1, p. 2760, 2017.

SIMPSON, Jared T. et al. ABySS: a parallel assembler for short read sequence data. **Genome research**, v. 19, n. 6, p. 1117-1123, 2009.

SILVA, Débora Mello da; BUENO, Adeney de Freitas. Organic products selectivity for *Trichogramma pretiosum* (Hymenoptera: Trichogrammatidae). **Arquivos do Instituto Biológico**, v. 82, p. 01-08, 2015.

SINGLETON, Martin R.; DILLINGHAM, Mark S.; WIGLEY, Dale B. Structure and mechanism of helicases and nucleic acid translocases. **Annu. Rev. Biochem.**, v. 76, p. 23-50, 2007.

SOMA, Fumiyuki et al. ABA-unresponsive SnRK2 protein kinases regulate mRNA decay under osmotic stress in plants. **Nature Plants**, v. 3, n. 1, p. 1-8, 2017.

SOTO, Iria et al. Why do farmers abandon jatropha cultivation? The case of Chiapas, Mexico. **Energy for Sustainable Development**, v. 42, p. 77-86, 2018.

SHARAD VERMA, Sharad Verma et al. Determination of phytocomponents of *Jatropha curcas* root by GC-MS analysis and their termiticidal activity. 2013.

SRIVASTAVA, Pankaj; SHARMA, Yogesh K.; SINGH, Nandita. Soil carbon sequestration potential of *Jatropha curcas* L. growing in varying soil conditions. **Ecological engineering**, v. 68, p. 155-166, 2014.

SOUZA, Larissa AL et al. JcTI-Pepl, a synthetic peptide bioinspired in the trypsin inhibitor from *Jatropha curcas*, presents potent inhibitory activity against *C. krusei*, a neglected pathogen. **Biochimie**, v. 200, p. 107-118, 2022.

SUBRAMANIAN, Balakrishnan et al. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. **Nucleic acids research**, v. 47, n. W1, p. W270-W275, 2019.

SUN, Yanqing et al. Twenty years of plant genome sequencing: achievements and challenges. **Trends in Plant Science**, v. 27, n. 4, p. 391-401, 2022.

SUPEK, Fran et al. REVIGO summarizes and visualizes long lists of gene ontology terms. **PloS one**, v. 6, n. 7, p. e21800, 2011.

SVITKIN, Yuri V. et al. The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5'secondary structure. **Rna**, v. 7, n. 3, p. 382-394, 2001.

SZKLARCZYK, Damian et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. **Nucleic acids research**, v. 49, n. D1, p. D605-D612, 2021.

TAMURA, Koichiro et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. **Molecular biology and evolution**, v. 24, n. 8, p. 1596-1599, 2007.

TARAZONA, Sonia et al. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. **Nucleic acids research**, v. 43, n. 21, p. e140-e140, 2015.

THALMANN, Matthias et al. Regulation of leaf starch degradation by abscisic acid is important for osmotic stress tolerance in plants. **The Plant Cell**, v. 28, n. 8, p. 1860-1878, 2016.

THIMM, Oliver et al. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. **The Plant Journal**, v. 37, n. 6, p. 914-939, 2004.

TRAPNELL, Cole et al. Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. **Nature biotechnology**, v. 28, n. 5, p. 511, 2010.

TSHIMENGA, Karl et al. Evaluation of antifungal activity of Moringa oleifera and Jatropha curcas extracts as a natural fungicide against adults of Acanthoscelides Obtectus. **International Journal of Engineering and Applied Sciences**, v. 5, n. 7, p. 257194, 2018.

TUTEJA, Narendra et al. Genotoxic stress in plants: shedding light on DNA damage, repair and DNA repair helicases. **Mutation Research/Reviews in Mutation Research**, v. 681, n. 2-3, p. 134-149, 2009.

TUTEJA, Narendra et al. OsBAT1 augments salinity stress tolerance by enhancing detoxification of ROS and expression of stress-responsive genes in transgenic rice. **Plant molecular biology reporter**, v. 33, p. 1192-1209, 2015.

TUTEJA, Narendra et al. O s SUV 3 dual helicase functions in salinity stress tolerance by maintaining photosynthesis and antioxidant machinery in rice (O ryza sativa L. cv. IR 64). **The Plant Journal**, v. 76, n. 1, p. 115-127, 2013.

TUTEJA, Narendra et al. Molecular mechanisms of DNA damage and repair: progress in plants. **Critical reviews in biochemistry and molecular biology**, v. 36, n. 4, p. 337-397, 2001.

TUTEJA, Narendra et al. Pea p68, a DEAD-box helicase, provides salinity stress tolerance in transgenic tobacco by reducing oxidative stress and improving photosynthesis machinery. **PloS one**, v. 9, n. 5, p. e98287, 2014.

UPADHYAY, B. et al. Traditional uses of medicinal plants among the rural communities of Churu district in the Thar Desert, India. **Journal of ethnopharmacology**, v. 113, n. 3, p. 387-399, 2007.

VAN BEL, Michiel et al. PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. **Nucleic acids research**, v. 46, n. D1, p. D1190-D1196, 2018.

VASER, R.; ŠIKIC, M. Raven: a de novo genome assembler for long reads. *bioRxiv*, bR. 2021.

VASHISHT, Ajay Amar et al. Cold-and salinity stress-induced bipolar pea DNA helicase 47 is involved in protein synthesis and stimulated by phosphorylation with protein kinase C. **The Plant Journal**, v. 44, n. 1, p. 76-87, 2005.

VELCULESCU, Victor E. et al. Serial analysis of gene expression. **Science**, v. 270, n. 5235, p. 484-487, 1995.

WAJID, Bilal et al. The A, C, G, and T of genome assembly. **BioMed research international**, v. 2016, 2016.

WALKER, John E. et al. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. **The EMBO journal**, v. 1, n. 8, p. 945-951, 1982.

WANG, Xing-Hong et al. Hydroxy-octadecenoic acids instead of phorbol esters are responsible for the *Jatropha curcas* kernel cake's toxicity. **Communications biology**, v. 3, n. 1, p. 228, 2020.

WANG, Yupeng et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. **Nucleic acids research**, v. 40, n. 7, p. e49-e49, 2012.

WATERHOUSE, Andrew M. et al. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, v. 25, n. 9, p. 1189-1191, 2009.

WU, Pingzhi et al. Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. **The Plant Journal**, v. 81, n. 5, p. 810-821, 2015.

XIAO, Chuan-Le et al. MECAT: fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. **nature methods**, v. 14, n. 11, p. 1072-1074, 2017.

XIE, Yinlong et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. **Bioinformatics**, v. 30, n. 12, p. 1660-1666, 2014.

YAMAGUCHI-SHINOZAKI, Kazuko; SHINOZAKI, Kazuo. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. **Annu. Rev. Plant Biol.**, v. 57, p. 781-803, 2006.

YANG, Yongqing; GUO, Yan. Unraveling salt stress signaling in plants. **Journal of integrative plant biology**, v. 60, n. 9, p. 796-804, 2018.

YANG, Cheng-yuan et al. Selection of high-oil-yield seed sources of *Jatropha curcas* L. for biodiesel production. **Biofuels**, v. 1, n. 5, p. 705-717, 2010.

ZAIDA, Iasmine N. Ramos et al. Biocide potential of extracts of *Jatropha curcas* L. on fungi *Hemileia vastatrix* and *Cercospora coffeicola*: causal agents of two main diseases of the coffee trees. 2021.

ZAHID, Ghassan et al. An overview and recent progress of plant growth regulators (PGRs) in the mitigation of abiotic stresses in fruits: A review. **Scientia Horticulturae**, v. 309, p. 111621, 2023.

ZDOBNOV, Evgeni M.; APWEILER, Rolf. InterProScan—an integration platform for the signature-recognition methods in InterPro. **Bioinformatics**, v. 17, n. 9, p. 847-848, 2001.

ZEHEBNER, Gunther. OntoBlast function: From sequence similarities directly to potential functional annotations by ontology terms. **Nucleic acids research**, v. 31, n. 13, p. 3799-3803, 2003.

ZERBINO, Daniel R.; BIRNEY, Ewan. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome research**, v. 18, n. 5, p. 821-829, 2008.

ZIMIN, Aleksey V. et al. The MaSuRCA genome assembler. **Bioinformatics**, v. 29, n. 21, p. 2669-2677, 2013.

ZÚÑIGA-LEÓN, Eduardo; CARRASCO-NAVARRO, Ulises; FIERRO, Francisco. NeVOmics: an enrichment tool for gene ontology and functional network analysis and visualization of data from OMICs technologies. **Genes**, v. 9, n. 12, p. 569, 2018.

APÊNDICE A - DEAD-Box RNA Helicase Family in Physic Nut (*Jatropha curcas* L.): Structural Characterization and Response to Salinity

Article

DEAD-Box RNA Helicase Family in Physic Nut (*Jatropha curcas* L.): Structural Characterization and Response to Salinity

Rahisa Helena da Silva ¹, Manassés Daniel da Silva ¹, José Ribamar Costa Ferreira-Neto ², Bruna de Brito Souza ¹, Francielly Negreiros de Araújo ¹, Elvia Jéssica da Silva Oliveira ¹, Ana Maria Benko-Iseppon ^{2,*}, Antonio Félix da Costa ³ and Éderson Akio Kido ¹

¹ Plant Molecular Genetics Laboratory, Genetics Department, Center of Biosciences, Federal University of Pernambuco, Recife CEP 50670-901, PE, Brazil; rahisa.silva@ufpe.br (R.H.d.S.); manasses.dsilva@ufpe.br (M.D.d.S.); bruna.britosouza@ufpe.br (B.d.B.S.); francielly.negreiros@ufpe.br (F.N.d.A.); elvia.oliveira@ufpe.br (E.J.d.S.O.); ederson.kido@ufpe.br (É.A.K.)

² Plant Genetics and Biotechnology Laboratory, Genetics Department, Center of Biosciences, Federal University of Pernambuco, Recife CEP 50670-901, PE, Brazil; joseribamar.ferreiraneto@ufpe.br

³ Agronomic Institute of Pernambuco—IPA, Recife CEP 50761-000, PE, Brazil; felix.antonio@ipa.br

* Correspondence: ana.iseppon@ufpe.br

Citation: da Silva, R.H.; Silva, M.D.d.; Ferreira-Neto, J.R.C.; Souza, B.d.B.; de Araújo, F.N.; Oliveira, E.J.d.S.; Benko-Iseppon, A.M.; da Costa, A.F.; Kido, É.A. DEAD-Box RNA Helicase Family in Physic Nut (*Jatropha curcas* L.): Structural Characterization and Response to Salinity. *Plants* **2024**, *13*, x. <https://doi.org/10.3390/xxxxx>

Academic Editors: Dezhi Wu, Peng Zhang and Hanhua Tong

Received: 19 January 2024

Revised: 16 March 2024

Accepted: 18 March 2024

Published: date



Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Helicases, motor proteins present in both prokaryotes and eukaryotes, play a direct role in various steps of RNA metabolism. Specifically, SF2 RNA helicases, a subset of the DEAD-box family, are essential players in plant developmental processes and responses to biotic and abiotic stresses. Despite this, information on this family in the physic nut (*Jatropha curcas* L.) remains limited, spanning from structural patterns to stress responses. We identified 79 genes encoding DEAD-box RNA helicases (*JcDHX*) in the *J. curcas* genome. These genes were further categorized into three subfamilies: DEAD (42 genes), DEAH (30 genes), and DExH/D (seven genes). Characterization of the encoded proteins revealed a remarkable diversity, with observed patterns in domains, motifs, and exon–intron structures suggesting that the DEAH and DExH/D subfamilies in *J. curcas* likely contribute to the overall versatility of the family. Three-dimensional modeling of the candidates showed characteristic hallmarks, highlighting the expected functional performance of these enzymes. The promoter regions of the *JcDHX* genes revealed potential *cis*-elements such as Dof-type, BBR-BPC, and AP2-ERF, indicating their potential involvement in the response to abiotic stresses. Analysis of RNA-Seq data from the roots of physic nut accessions exposed to 150 mM of NaCl for 3 h showed most of the *JcDHX* candidates repressed. The protein–protein interaction network indicated that *JcDHX* proteins occupy central positions, connecting events associated with RNA metabolism. Quantitative PCR analysis validated the expression of nine DEAD-box RNA helicase transcripts, showing significant associations with key components of the stress response, including RNA turnover, ribosome biogenesis, DNA repair, clathrin-mediated vesicular transport, phosphatidyl 3,5-inositol synthesis, and mitochondrial translation. Furthermore, the induced expression of one transcript (*JcDHX44*) was confirmed, suggesting that it is a potential candidate for future functional analyses to better understand its role in salinity stress tolerance. This study represents the first global report on the DEAD-box family of RNA helicases in physic nuts and displays structural characteristics compatible with their functions, likely serving as a critical component of the plant’s response pathways.

Keywords: genomics; transcriptomics; abiotic stress; RNA helicase; Euphorbiaceae; oilseed; qPCR

1. Introduction

The physic nut (*Jatropha curcas* L.), a small, inedible shrub belonging to the Euphorbiaceae family, is notable for its seeds containing a substantial amount of oil. The quality of this oil makes it a promising source for biodiesel generation, presenting a potential solution to the issues associated with fossil fuel use [1]. Moreover, these plants thrive in marginal areas and play a vital role in controlling erosion and revitalizing soils contaminated by heavy metals [2,3].

A primary challenge in modern agriculture is the increasing salinization of soils, which is a critical abiotic stress that causes substantial production loss. Salinity disrupts the ionic and osmotic balance of plants, hindering water and nutrient absorption and ultimately leading to the excessive production of reactive oxygen species (ROS) and damage to membranes, proteins, and organelles [4].

RNA helicases are universal enzymes that use ATP hydrolysis energy to unwind RNA strands and are essential components of various steps of RNA metabolism [5]. Categorized into six superfamilies (SF1-6), the

majority fall into SF2, further divided into three subfamilies: DEAD, DEAH, and DExH/D, commonly referred to as DEAD-box, based on variations in the Asp-Glu-Ala-Asp (DEAD) motif [6]. DEAD-box RNA helicases are the largest family of helicases, characterized by the presence of nine conserved motifs (Q, I, Ia, II, III, IV, V, and VI) that are known to be involved in helicase and ATPase activities [5,6]. DEAD-box proteins also contain N- and C-terminal extensions that differ in domain composition and, in some cases, target the proteins to specific substrates via protein–protein interactions [5,6].

Several plant species, including *Arabidopsis thaliana*, *Oryza sativa*, *Gossypium raimondii*, *Solanum lycopersicum*, *Vitis vinifera*, and *Triticum aestivum*, have reported *DHX* genes encoding DEAD-box RNA helicases [7–12]. These *DHX* genes play multiple roles in RNA metabolism and have been implicated in responses to environmental stresses. For example, the *BrRH22* gene, which exhibits RNA chaperone activity, has been linked to drought and salinity tolerance in transgenic *A. thaliana* plants [13]. Similarly, *SIDEAD30* and *SIDEAD31*, which are responsive to drought and salinity, conferred tolerance to these abiotic stresses when overexpressed in tomato [14]. Overexpression of the *AvDH1* gene in cotton provided salinity tolerance and reduced oxidative stress [15], whereas overexpression of the *OsRH58* gene enhanced salinity and drought tolerance in transgenic *A. thaliana* plants [16].

Understanding plant responses to abiotic stresses is crucial, given the anticipated increase in the frequency and intensity of environmental stresses due to ongoing climate change. This poses a threat to food security in several populations in addition to biodiversity loss. In this study, the *J. curcas* DEAD-box RNA helicase family was comprehensively analyzed, including characterization of exon–intron structure patterns, conserved domains and motifs, potential secondary structures, and 3D models of DEAD-box subfamily candidates. Additionally, this study explored the promoter regions of related genes for the presence of cis-regulatory elements (CREs) and evaluated the responses of *DHX* transcripts expressed in *J. curcas* plants after three hours of exposure to NaCl (150 mM). This study marks the first genome-wide report of the DEAD-box RNA helicase family in *J. curcas*, offering insights that may contribute to future investigations of this gene family in related species, particularly those with current breeding programs for developing plant tolerance to abiotic stresses.

2. Material and Methods

2.1. Identification of Putative JcDHX Proteins and Genes

The identification of *JcDHX* (*Jatropha curcas* *DHX* genes) potential proteins commenced with a comprehensive exploration of the physic nut proteome (RJC1_Hi-C_protein.faa) associated with the *J. curcas* reference genome (NCBI RefSeq assembly GCF_014843425.1, Genome assembly RJC1_Hi-C: 282,312 Mb, 22,718 genes, 29,586 proteins), accessible on the NCBI website (<https://www.ncbi.nlm.nih.gov> (accessed on 25 November 2020)). The Hidden Markov Model (HMM) played a pivotal role in this process, employing the HMM DEAD profile (PF00270) sourced from the Pfam protein family database, according to InterPro v97.0 (<https://www.ebi.ac.uk/interpro/> (accessed on 5 April 2021); Ref. [17]). Standard parameters and an e-value cut-off of $<1.0^{-5}$ were applied for domain annotation of each protein sequence. Furthermore, the *JcDHX* candidates were scrutinized through CDD v3.13 (Conserved Domain

Database—<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> (accessed on 7 April 2021)) [18] and SMART v9 (Simple Modular Architecture Research Tool—<http://smart.embl-heidelberg.de/> (accessed on 20 April 2021)) [19]. This dual-validation approach bolstered the identification of conserved domains within candidate proteins. Putative *JcDHX* proteins were associated with genes in the *J. curcas* reference genome.

2.2. Promoter Analysis

To explore the regulatory landscape of *JcDHX* genes, we retrieved the promoter regions situated 1000 base pairs upstream of the transcription initiation site from the NCBI Genome Browser panel (<https://www.ncbi.nlm.nih.gov/genome/gdv/browser/> (accessed on 9 July 2021)) of the *J. curcas* reference genome (Genome assembly RJC1_Hi-C). Subsequently, each promoter region was analyzed using the MEME v5.3.3 program (<https://meme-suite.org/meme/> (accessed on 2 August 2021)) [20]. For each identified motif, the software reported the corresponding e-value. Motifs with an e-value < 0.05 are considered significant. The maximum number of motifs analyzed in this study for a single *JcDHX* promoter region was 10, and the extension ranged from 6 to 50 nt. To identify shared motifs and gain insights into their significance, the Tomtom software v5.3.3 (<https://meme-suite.org/meme/tools/tomtom> (accessed on 2 August 2021)) [21], in conjunction with the JASPAR database (JASPAR2018_CORE_plants_non-redundant), was employed. This collaborative approach aimed to recover the identity of the identified motifs, emphasizing those most representative, with a *p*-value cut-off of $<10^{-2}$.

2.3. Structural Characterization of the *JcDHX* Candidates and Potential Subcellular Localization

The prediction of exon–intron structures for the genes encoding candidate *JcDHX* proteins involved a comparative analysis of the coding sequence (CDS) with the genomic sequence. This was accomplished using the Gene Structure Display Server v2.0 (GSDS, <http://gsds.gao-lab.org/> (accessed on 10 November 2021)) [22]. Simultaneously, the physicochemical attributes of the candidate *JcDHX* proteins were systematically characterized using the ExPASy tool (<http://web.expasy.org/protparam/> (accessed on 28 April 2021)) [23]. Concurrently, predictions regarding subcellular localization were made using the CELLO v2.5 tool (<http://cello.life.nctu.edu.tw/> (accessed on 27 December 2021)) [24]. To unveil potentially conserved motifs within the sequences, the MEME v5.3.3 program (<https://meme-suite.org/meme/> (accessed on 7 July 2021)) [20] was employed. This analysis adhered to the following specific parameters: anr (any number of repetitions), a maximum number of motifs set at 10, and a motif size spanning 6–50 amino acid residues.

2.4. Phenetic Analysis and Orthology

To classify *JcDHX*, a phenetic analysis was conducted, encompassing *DHX* protein sequences sourced from *S. lycopersicum* (63) and *A. thaliana* (69). These annotated sequences were obtained from the Sol Genomics Network (<https://solgenomics.net/> (accessed on 23 July 2021)) and The *Arabidopsis* Information Resource (TAIR; <https://www.arabidopsis.org/> (accessed on 29 July 2021)), respectively. Alignment of the protein

sequences was performed using the ClustalX v2.1 program [25], and subsequent clustering was achieved using the Neighbor-Joining method with bootstrap analysis comprising 1000 replicates. The resulting clusters were visualized using the online tool iTOL v5 (<https://itol.embl.de/> (accessed on 29 March 2022)) [26]. To scrutinize the conservation of the JcDHX family across different species, orthologs were inferred using the Bidirectional Best-Hit (BBH) method [27]. The analysis encompassed species from the Euphorbiaceae family, namely *Manihot esculenta* (GCF_001659605.1; *Manihot esculenta_v6_protein.faa*), *Ricinus communis* (GCF_000151685.1; *JCVI_RCG_1.1_protein.faa*), and *Hevea brasiliensis* (GCF_001654055.1; *ASM1654_05v1_protein.faa*). Additionally, representatives of *Populus trichocarpa* (GCF_000002775.4; *Pop_tri_v3_protein.faa*), *S. lycopersicum* (GCF_000188115.4; *SL3.0_protein.faa*), and *A. thaliana* (Araport11_genes.201606.pep.fasta), were included in the analysis to provide a broader comparative context.

2.5. Secondary Structure Elements and 3D Modeling

We performed a secondary structural analysis using the SOPMA method (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html (accessed on 23 October 2023)) to predict secondary structure elements (SSEs) [28]. Multiple alignment of the JcDHX DEAD subfamily candidates was conducted using ClustalX v2.1 [25] and viewed using the Jalview v2.11.3.2 software [29]. This alignment was used as an input file to generate a 2D alignment with the Ali2D tool (<https://toolkit.tuebingen.mpg.de/tools/ali2d> (accessed on 9 September 2023)) [30], followed by graphical visualization provided by the 2dSS tool (<http://genome.lcqb.upmc.fr/2dss/> (accessed on 10 September 2023)) [31].

Homology modeling of the JcDHX candidates hinged upon the structure of the most similar PDB templates, facilitated by the SWISS-MODEL web tool (<https://swissmodel.expasy.org/> (accessed on 22 September 2023)) [32]. Subsequent scrutiny and adjustments of the generated 3D models were performed using the PyMOL v2.5 program (<https://pymol.org/> (accessed on 1 September 2023)) [33]. Model accuracy was assessed using Ramachandran plots.

2.6. RNA-Seq Analysis of JcDHX Candidates and Gene Expression Validation by qPCR

We performed an in silico analysis of JcDHX expression using RNA-Seq data from two distinct Brazilian accessions of *J. curcas* that displayed different NaCl tolerance phenotypes, with Jc183 manifesting tolerance and Jc171 exhibiting a less tolerant phenotype [34]. The transcriptome data used were previously sequenced and analyzed by our group [35]. The experimental salt assay followed a completely randomized design with two accessions, two treatments (without salt or with NaCl, 150 mM, three-hour salt exposure), and three plants (half-siblings) of each accession simulating biological replicates. After salt exposure, we collected the roots, which were immediately frozen in liquid nitrogen and stored (−80 °C) until RNA extraction. A total of 12 RNA-Seq libraries (two accessions × two treatments × three plants each accession) were generated following the LS protocol of the Illumina TruSeq Stranded mRNA Sample Prep kit (Illumina, Inc., San Diego, CA, USA). Libraries were sequenced on an Illumina HiSeq 2500 (paired-end 100 bp reads). The de novo transcriptome covered 101 MB and 145,422 assembled transcripts with a

GC content of 41.55%, and the N50 reached 1308 bp. The average aligned fraction was 0.97, with 84,534 transcripts with at least one significant alignment to the reference genome (GenBank assembly accession number: GCA_000696525.1) [35]. Comprehensive details regarding the salinity assay, RNA extraction, RNA-Seq libraries, and transcriptome assembly procedures are outlined by Souza et al. [35]. In order to identify DHX candidates in this transcriptome, RNA-Seq transcripts were subjected to BLASTx analysis (e-value cut-off: e^{-10}), aligning them with previously identified *JcDHX* proteins. Each transcript sequence had its domains confirmed using the CDD [18] and SMART [19] tools. TransDecoder (<https://github.com/TransDecoder/TransDecoder> (accessed on 17 August 2021)) [36] was used to obtain peptide sequences from *JcDHX* RNA-Seq transcripts. RNAsamba (<https://rnasamba.lge.ibi.unicamp.br/> (accessed on 6 February 2024)) [37] was used to describe the coding probabilities of identified transcripts. To evaluate whether RNAs classified as non-coding by RNAsamba could be assigned to protein families, we compared transcripts translated to a UniProtKB/Swiss-Prot from *Arabidopsis thaliana* using BLASTp (e-value cut-off: e^{-10}) [38]. Finally, the translated ORFs were compared to the Pfam protein family database (version 32.0) using the *hmmsearch* command from the HMMER suite. Transcripts not translated by TransDecoder had open reading frames (ORFs) predicted by ORF Finder (<https://www.ncbi.nlm.nih.gov/orffinder/> (accessed on 7 February 2024)) [39]. *JcDHX* candidates identified as differentially expressed transcripts (DETs), representing the Trinity unigenes (unique assembled transcripts), exhibited a *p*-value < 0.0001 and false discovery rate (FDR) < 0.005 for de novo transcriptome assembly. The fold change (FC) values discriminated positively (UR, upregulated; $FC \geq 1$) from the negative modulation of expression (DR, downregulated; $FC \leq -1$). FC values represent the ratio of transcript abundance considering their presence in two compared RNA-Seq libraries with the untreated library as the reference sample. Hierarchical clustering of the *JcDHX* transcript, based on the Cluster v3.0 software (<https://cluster2.software.informer.com/3.0/> (accessed on 11 March 2024)), considered the FC values, and the visualization was performed using the Java TreeView v1.1 software [40].

For the validation of the *in silico* expression of selected DETs, we designed primers for qPCR assays using the Primer3 tool [41], with some adjustments for amplicon size (70–200 bp), GC content (45–55%), dissociation temperature (50–80 °C), and CG clamp (one). The specificity of amplicons was assessed using dissociation curves generated between 65 and 95 °C (20 min) after 40 qPCR cycles. To ensure robust quantification, the amplification efficiency [42] for each primer pair was determined using calibration curves established with cDNAs from the respective accessions in serial dilutions (no dilution, 10⁻¹, 10⁻², 10⁻³, and 10⁻⁴). qPCR reactions were conducted in a LineGene 9660 thermocycler (Bioer®, Hangzhou, China using SYBR Green as the detection system. The final reaction volume was 10 µL, comprising 1 µL of sample cDNA (1:10), 5 µL of GoTaq® qPCR Master Mix (Promega®, Wisconsin, EUA), 2 µL of ultrapure water, and 1 µL of each primer (0.05 µM). The qPCR reactions were executed in technical and biological triplicates, featuring a negative control, and included two reference genes (actin and β-tubulin). These reference genes were selected through specific assays using samples from the same experimental set. The reaction protocol included an initial denaturation step at 95 °C (2 min), followed by 40 cycles of 95 °C (15 s)

and 60 °C (60 s). Quantification cycles (Cq), dissociation temperature, and both absolute and relative quantification were determined using the Bioer® proprietary software. Expression data were further analyzed using the REST software (v. 2.0.13) [43].

2.7. Protein–Protein Interaction (PPI) Networks for JcDHX Candidates

To explore the potential interactions of JcDHX proteins codified by DETs, we used the STRING software (v11.5) [44] to construct PPI networks. A confidence score of >0.7 (high) was set for interaction, relying on experiment-derived interactions in *Arabidopsis*. Text-mining-based evidence was excluded to ensure reliability of the predicted PPI networks. Cytoscape v3.10.1 [45] was used to visualize and editing protein–protein interaction (PPI) networks.

3. Results

3.1. Identification of JcDHX Genes and Their Gene Structures

The *J. curcas* reference genome (GCF_014843425.1_RJC1_Hi-C) contains 79 genes encoding 146 DHX proteins. DHX genes, designated JcDHX1 to JcDHX79 based on their respective scaffold positions, are shown in Table S1, with details including the genomic location, locus ID, gene length (nt), intron count, ORF length (nt), expected protein length (aa), and DHX subfamily.

Gene structure is a critical factor for understanding the evolution of gene families. The exon–intron structures of JcDHX genes, comprising 42 DEAD and 37 DEAH/DExH/D subfamily members, are shown in Figure 1. Our analysis revealed a lack of discernible patterns within each subfamily and considerable variability even among members of the same subfamily (Figure 1). The intron investigation revealed a broad range of variation, detecting no introns to a maximum of 30 introns (Table S1). In general, DEAD subfamily members exhibited simpler gene structures, with sizes ranging from approximately 1500 nt to almost 18,000 nt, and a reduced number of introns, including those without introns (e.g., JcDHX6, JcDHX44, JcDHX33, and JcDHX14; Table S1 and Figure 1). Conversely, DEAH and DExH/D subfamily members exhibited greater lengths, spanning approximately 3200 nt to almost 39,000 nt, and more complex gene structures, with a potential for up to 30 introns (Table S1 and Figure 1). This comprehensive exploration of the JcDHX gene family provides valuable insights into structural diversity and sheds light on the distinct characteristics within and between subfamilies.

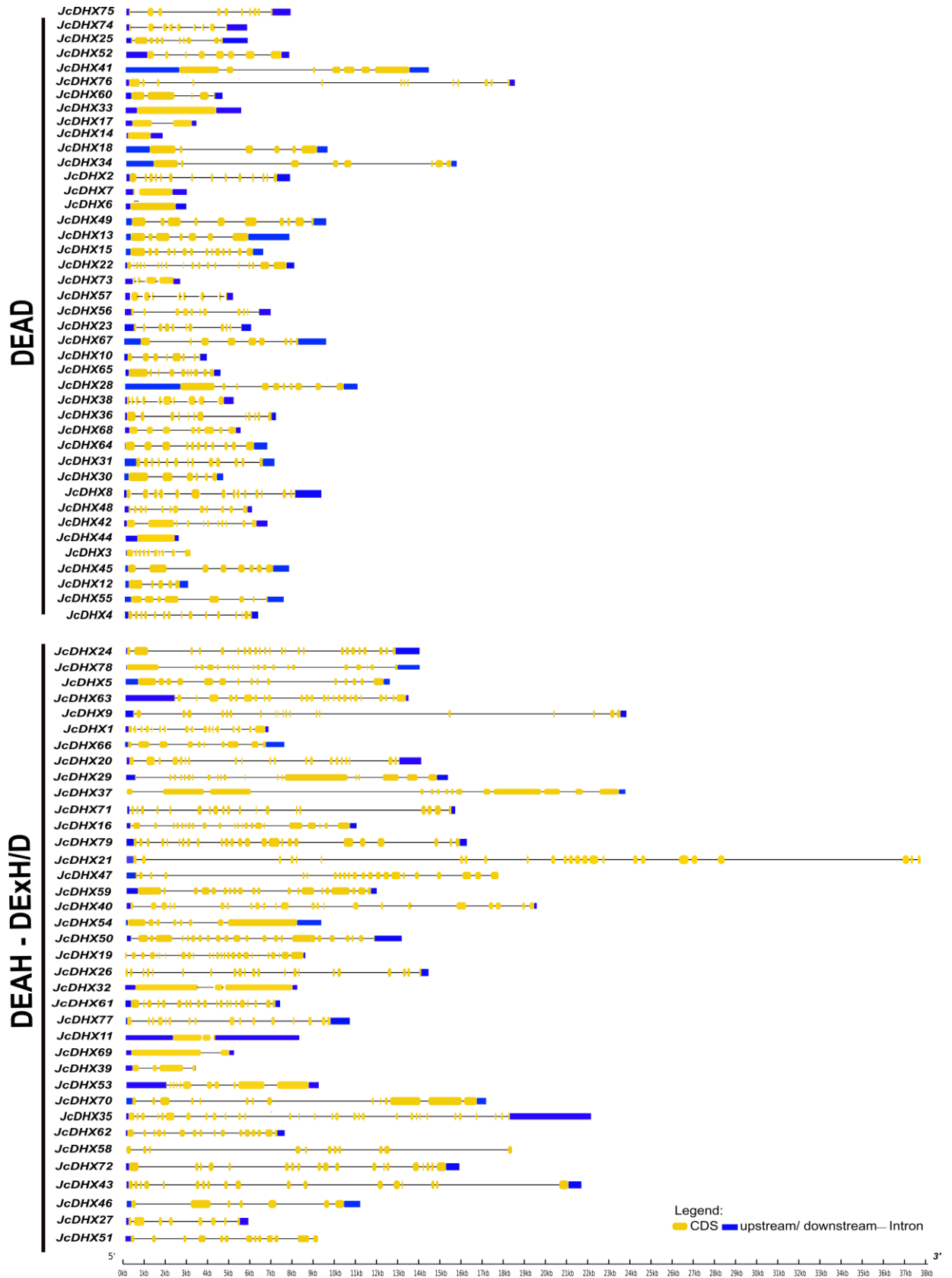


Figure 1. Exon–intron structures of *JcDHX* genes. The graphical representation was generated using the Gene Structure Display (GSDS).

3.2. Analysis of *JcDHX* Gene Promoter Regions

Analysis of *cis*-regulatory elements (CREs) in promoter regions (1.0 kb) of the 79 putative *JcDHX* genes of the *J. curcas* RJC1_Hi-C genome identified potential TFs possibly interacting with these genes. Eight motifs were detected using the MEME program (e-value < 0.05). Seven of the top eight detected CRE motifs (p -value < 10^{-2}) were associated with TF members of the Dof-type, BBR-BPC, HD-ZIP, WRKY, Myb-related, and bHLH families (Figure 2; Table S2). For each enriched CRE motif, details such as the MEME logo, JASPAR IDs, and e-values are shown in Figure S1. CREs distributed along the promoter regions are shown in Figure S2. Notably, the most prevalent TFs associated with the analyzed gene promoters belonged to the Dof-type and BBR-BPC families. Members of these TF families are key regulatory players involved in various cellular processes, including those involving plants' abiotic stress responses. This comprehensive examination enhances our understanding of the regulatory elements governing the expression of *JcDHX* genes and provides valuable insights concerning their potential assistance in plant responses to stress.

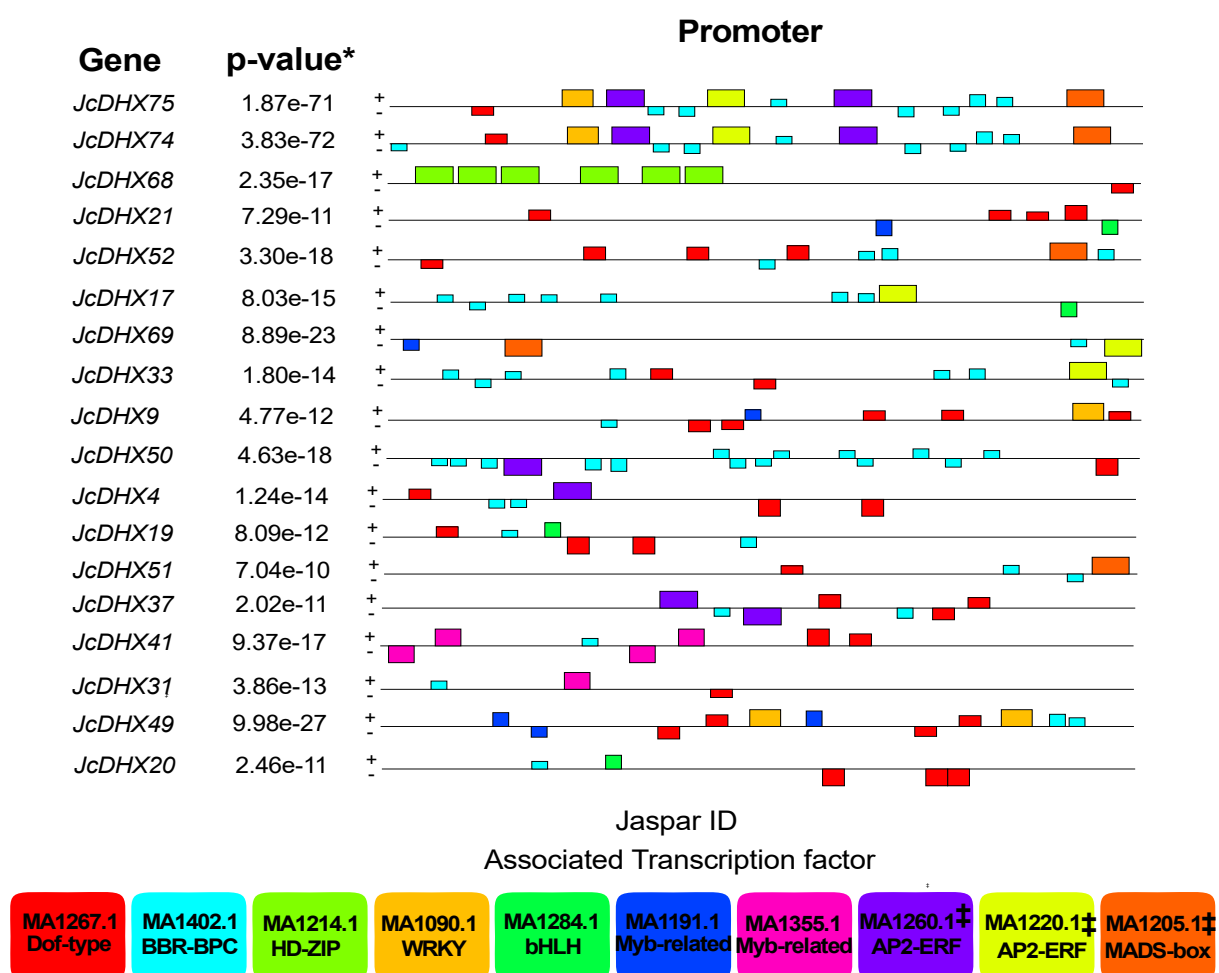


Figure 2. Distribution of candidate *cis*-regulatory elements (CREs) in the promoter set of *Jatropha curcas* DHX genes. The colored boxes present information about the CREs and the associated transcription factors with JASPAR IDs. MEME's combined p -value *, representing the probability of a random sequence matching the motif under test with a score greater than or equal to that found in the sequence under test; ‡ statistical significance below the considered cut-off (e-value < 0.05; p -value < 0.01).

3.3. Orthology Analysis of JcDHX Genes

The orthology analysis provided insights into the evolutionary relationships of the *DHX* gene family across the plant species analyzed. The orthology analysis, encompassing bidirectional comparisons of the 79 *JcDHX* genes with sequences of each examined species, according to the BBH methodology, identified substantial ortholog amounts: 74 (with *M. esculenta*), 74 (with *H. brasiliensis*), 73 (with *R. communis*), 72 (with *P. trichocarpa*), 74 (with *S. lycopersicum*), and 70 (with *A. thaliana*). The three Euphorbiaceae species (*M. esculenta*, *R. communis*, and *H. brasiliensis*) shared orthologs of 73 *JcDHX* genes. In turn, three species outside the Euphorbiaceae family (*P. trichocarpa*, *S. lycopersicum*, and *A. thaliana*) shared orthologs of 70 *JcDHX* genes. Across all six analyzed species, shared orthologs of the 69 *JcDHX* genes were identified (Table S3 and Figure S3). Interestingly, no orthologs of five *JcDHX* genes were detected in the six analyzed species. The genes without detected orthologs were *JcDHX11*, *JcDHX14*, *JcDHX34*, and *JcDHX75* (from the DEAH subfamily), as well as *JcDHX58* (DExH subfamily) (Table S3).

3.4. Phenetic Analysis of JcDHX Proteins

DEAD-box helicases are primarily categorized into subfamilies (DEAD, DEAH, and DExH/D) based on variations in motif II (D-E-A-D). To enhance the classification of potential *JcDHX* proteins into subfamilies, we conducted a phenetic analysis, covering the largest protein translated from each *JcDHX* gene together with curated DEAD-box RNA helicases from *A. thaliana* and *S. lycopersicum* [7,10]. The resulting phenetic tree grouped *JcDHX* proteins into ten distinct subgroups (Figure 3).

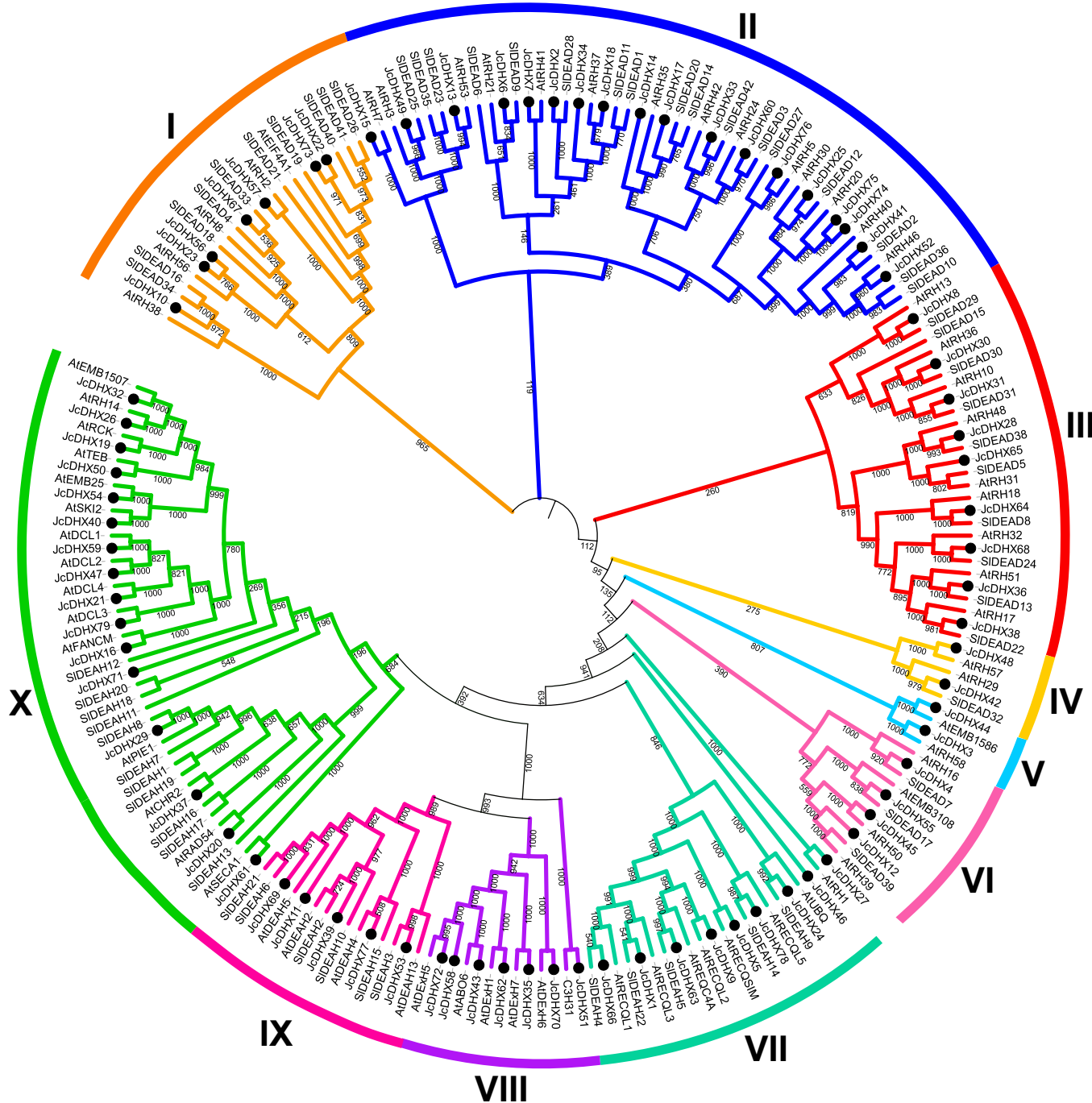


Figure 3. Phenetic tree generated by applying the Neighbor-Joining method (bootstrap of 1000 replicates), considering DEAD-box sequences from *Jatropha curcas*, *Arabidopsis thaliana*, and *Solanum lycopersicum*. Subgroups I to VI correspond to the DEAD subfamily; subgroups VII, IX, and X represent the DEAH subfamily, while subgroup VIII corresponds to DExH/D.

Evaluation of sequences covering motif II revealed that 42 *JcDHX* proteins from subgroups I, II, III, IV, V, and VI (Figure 3) comprised the DEAD subfamily, whereas subgroups VII, IX, and X (Figure 3) were assigned to the DEAH subfamily (30 members), and subgroup VIII comprised the DExH/D subfamily (seven members) (Figure 3). This comprehensive analysis supports the validity of the adopted classification for *JcDHX* proteins.

3.5. Conserved Domains and Motifs in *JcDHX* Proteins

Conserved domains and motifs are crucial for defining the functional specificity of a protein. Among the 146 putative DHX proteins, the structural core of the family, comprising the DEAD-box (N-terminal) and Helicase_C (C-terminal) domains, was observed. In essence, *JcDHX* proteins have a conserved helicase core essential for their anticipated functions.

In addition, auxiliary domains flanking the N- and C-terminal regions were detected, contributing to the functional diversity of DHX proteins (Figure 4). A distinctive feature of DHX proteins is the presence of up to nine conserved motifs (Q, I, Ia, Ib, II, III, IV, V, and VI), which were detected in the putative *JcDHX* proteins, showing the conservation of residues and their sequential arrangement (Figure 4; Table S4).

Additionally, the presence and variations in domains and motifs serve as indicators of the DHX subfamily classification. The DEAD subfamily sequences exhibited a canonical structure, with all nine motifs coupled, with a minimal presence of auxiliary domains (Figure 4). Conversely, the DEAH and DExH/D subfamilies displayed notable diversity in motif presence, accompanied by a greater number of auxiliary domains, such as RecQ, HRDC, HA2, OB_NTP_Bind, DSRM, Dicer, PAZ, RIBOc, R3H, POLAc, rRNA_pro-arch, DSHCT, AAA, HTH_40, SecA_DEAD, SecA_SW, HSA, SANT, DUF1998, and Sec63 (Figure 4). This comprehensive analysis provides valuable insights into the structural nuances that underlie functional diversity within the *JcDHX* protein family.

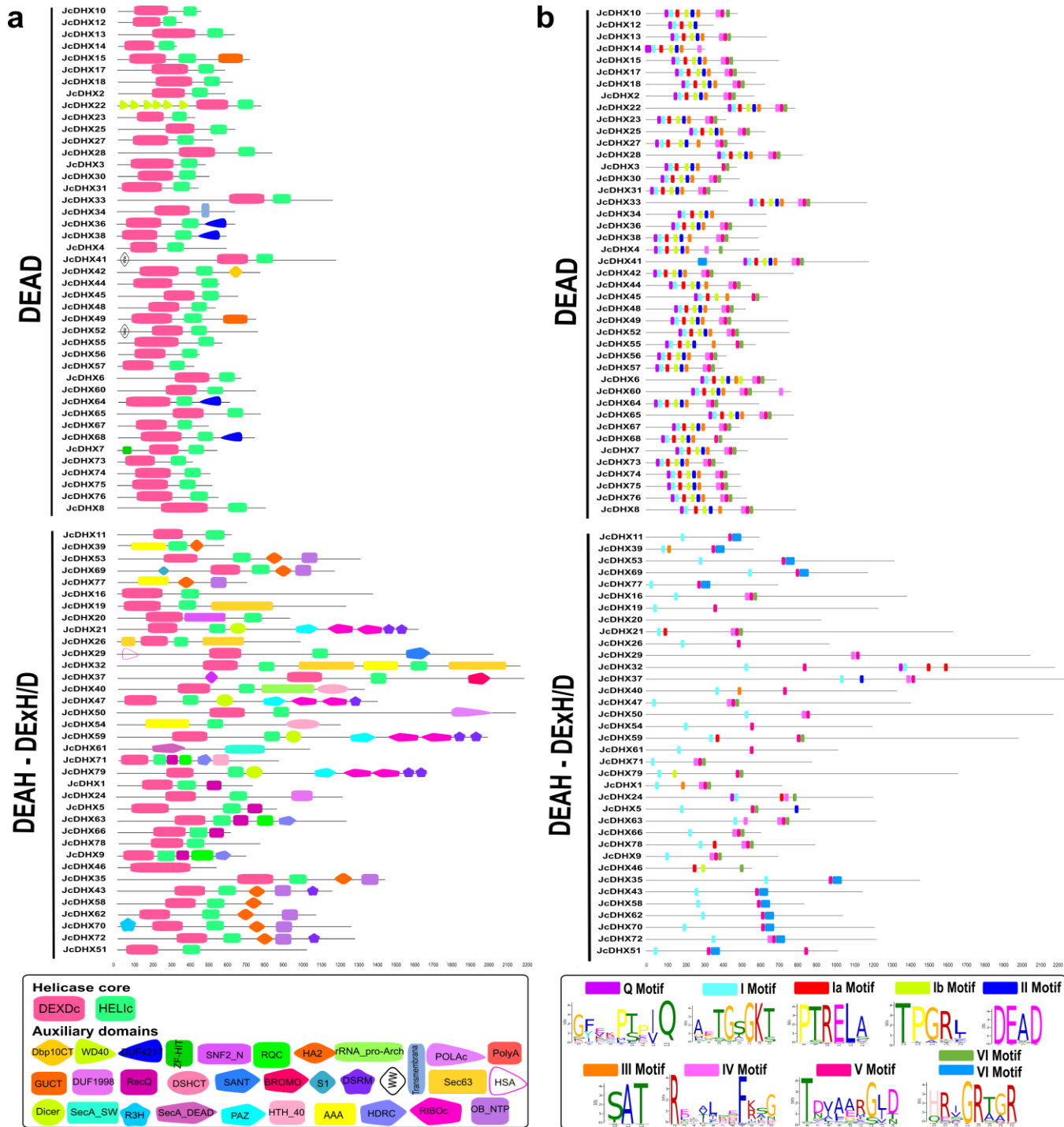


Figure 4. Structure of (a) domains and (b) conserved motifs detected in *JcDHX* proteins according to the DEAD, DEAH, and DEXH/D subfamilies. The conserved domains were identified using the SMART software, and the motifs were detected by the MEME program.

3.6. Physicochemical Characteristics and Subcellular Localization of *JcDHX* Proteins

The comprehensive characterization of *JcDHX* proteins was extended to their physicochemical attributes, including protein size (aa), molecular weight (MW), isoelectric point (pI), and potential subcellular localization (Table S5). Regarding protein size, the *JcDHX* proteins exhibited a substantial range, varying from 317 aa (*JcDHX14*) to 2247 aa (*JcDHX37*).

The isoelectric point (pI) values spanned from 5.19 (*JcDHX54*) to 9.97 (*JcDHX25*), reflecting the diverse charge characteristics of these proteins. In terms of molecular weight, the *JcDHX* proteins showcased considerable variation, with weights ranging from 34.76 KDa (*JcDHX14*) to 251.62 KDa (*JcDHX37*). Such variability in *JcDHX* proteins across the mentioned parameters underscores the complexity of this protein family.

Prediction analysis of subcellular localization, a crucial factor influencing biological function by regulating access to specific molecular partners, revealed the likelihood of *JcDHX* proteins being distributed across five distinct cellular compartments. Predominantly, these proteins were predicted to be present in the nucleus (94), cytoplasm (33), chloroplast (nine), mitochondria (seven), and plasma membrane (three) (Table S5). This multifaceted analysis provided valuable insights into the diverse roles of these proteins in different cellular environments.

3.7. Prediction of Secondary Structure Elements in *JcDHX* Proteins

The forecasted secondary structure elements predicted from the *JcDHX* protein sequences, employing the SOPMA tool (Table S6), revealed a prevailing composition ranging from 17.52% to 60.68% α -helices, 6.84% to 24.37% β -sheets, 0.00% to 11.06% β -turns, and 29.44% to 63.11% random coils. Specifically, the DEAD subfamily members exhibited an average of 40.75% α -helices and 13.47% β -sheets, whereas the DEAH subfamily displayed an average of 40.53% α -helices and 13.35% β -sheets, and the DExH/D subfamily showed an average of 40.07% α -helices and 12.61% β -sheets.

Distinctive structural features, referred to as “caps,” have been identified in the *J. curcas* DEAD subfamily proteins. These caps, situated atop domain “1” and depicted as green triangles (Figure 5), are composed of a β -sheet and two α -helices located just above the Walker A motif. Additionally, conserved motifs, such as motifs I (Walker A) and II (Walker B), primarily reside in the transition regions between the β -sheets and α -helices, as indicated by the red lines. These motifs are shared among the *JcDEAD* proteins. Therefore, our *JcDEAD* candidates exhibit similar structural elements, including a conserved cap structure above the Walker A motif, which reinforces their compatibility with their functional roles. The graphical representation of all predicted SSEs (Figure S4) considers the multiple alignment (ClustalX v2.1) of 42 *JcDEAD* subfamily proteins, as visualized by the Jalview software (Figure S5).

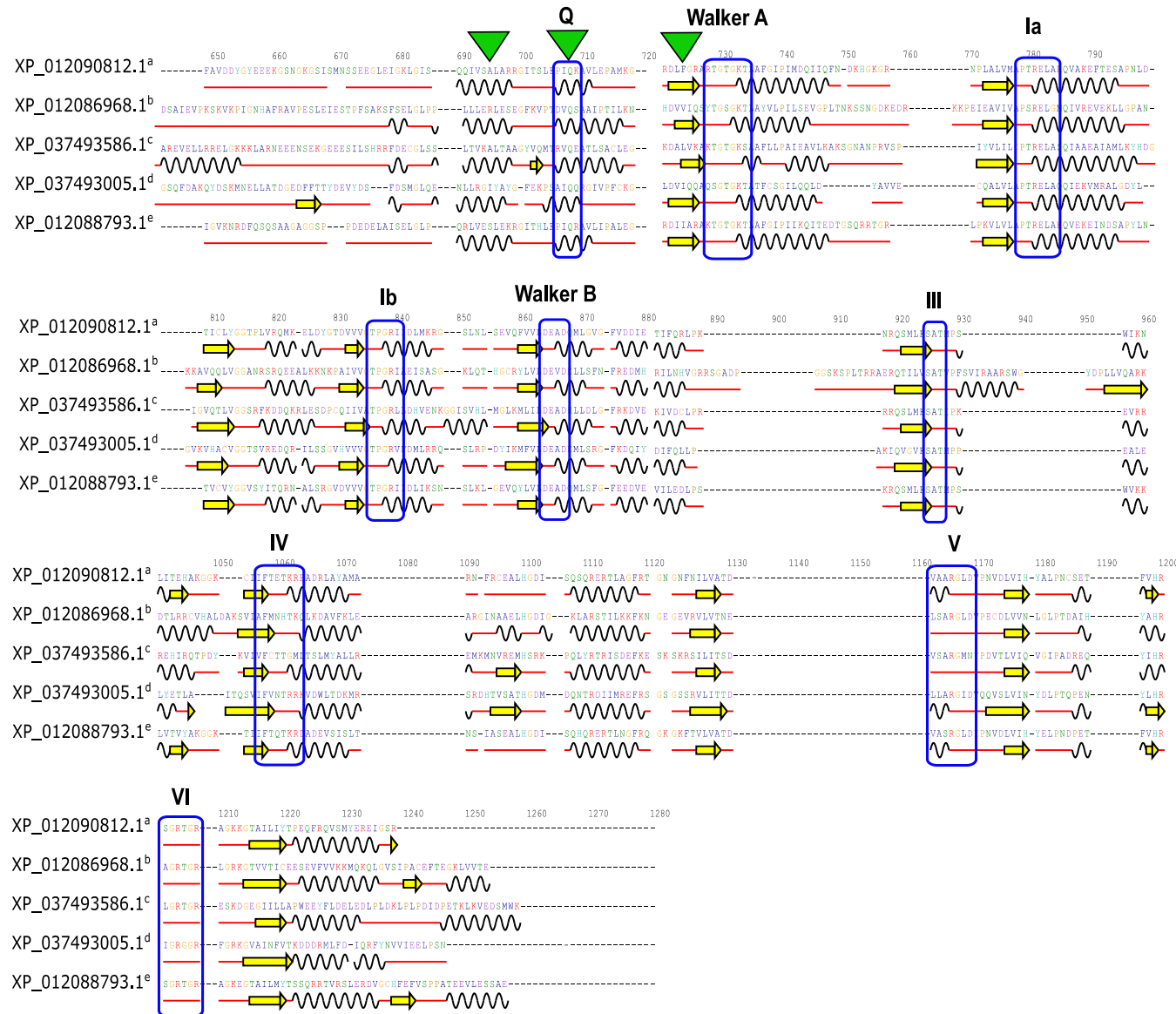


Figure 5. Schematic representation of secondary structure elements in DEAD subfamily proteins encoded by *J. curcas* RNA helicase genes. Curved black lines represent α -helices, horizontal yellow arrows represent β -sheets, and red lines represent the transition regions between structures. Dotted lines indicate the amino acids present in other proteins within the analysis. Blue rectangles highlight conserved motifs in the sequences, whereas green triangles emphasize the β -sheets and α -helices that make up the cap structure upstream of motif I (Walker A), encompassing motif Q. Superscript letters (a, b, c, d, e) represent *JcDHX*13, 44, 28, 22, and 49, respectively. This figure illustrates the first 5 out of 42 protein sequences from the graphical representation view of the 2D alignment provided by the 2dSS tool after Ali2D analysis (Figure S4) based on the multiple alignment data generated by ClustalX v2.1 (Figure S5).

3.8. Homology Modeling of *JcDHX* Candidates

A total of 79 *JcDHX* proteins, representing distinct *JcDHX* genes, underwent 3D modeling using Swiss-Model. Of these 79 candidates, 71 exhibited a Global Model Quality Estimate (GMQE) exceeding 0.60, with 68 achieving coverage exceeding 90%. The resulting structures displayed 80.72% to 98.05% of residues within permissible regions in Ramachandran plots, and the QMEANDisCo global scores ranged from 0.44 to 0.83 (Table S7; Figure S6). The most optimal model from each subfamily, *JcDHX*73

(DEAD), *JcDHX26* (DEAH), and *JcDHX58* (DExH/D), is depicted in Figure 6, emphasizing the conserved motifs and domains. Within N-terminal domain 1 (DEAD-box), motifs Q, I, Ia, Ib, II, and III were arranged, while in C-terminal domain 2 (HELIC), motifs IV, V, and VI were positioned (Figure 6). All models exhibited two central globular domains (core), characterized by enveloped β -sheets surrounded by α -helices, which is a hallmark of the SF2 superfamily (Figures 6 and S7).

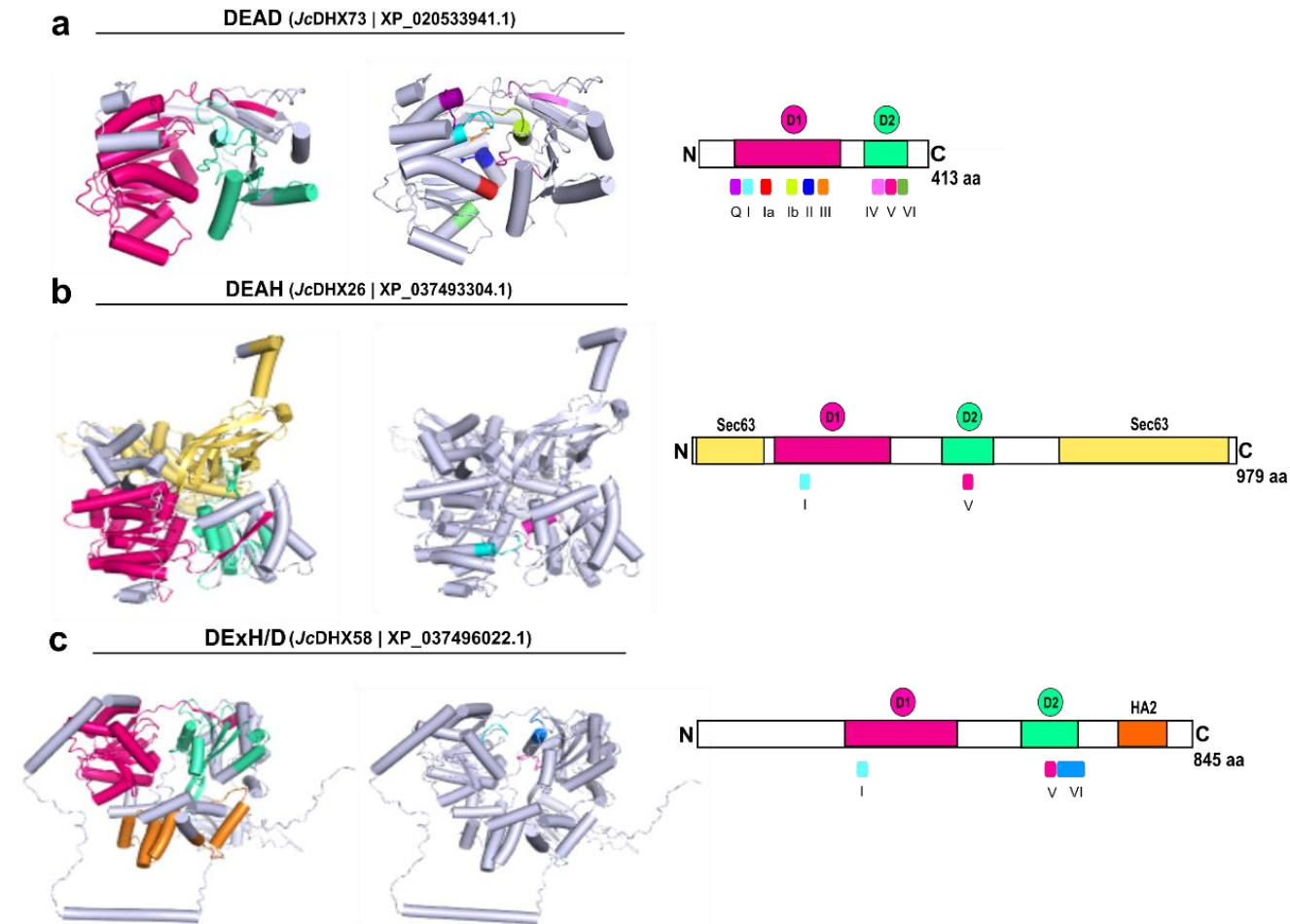


Figure 6. Best 3D models generated for *Jatropha curcas* DEAD-box helicases using the Swiss-model [(a) DEAD, (b) DEAH, and (c) DExH/D subfamilies], highlighting the presence of conserved domains and motifs. The N-terminal DEAD-box core domains (D1) are colored in pink, with the C-terminal HELIC (D2) in cyan green.

3.9. In Silico Expression of *JcDHX* Candidates and the qPCR Assay

Considering the *JcDHX* proteins uncovered from the RJC1_Hi-C reference genome and the RNA-Seq data of the two *J. curcas* accessions following a 3 h exposure to NaCl (150 mM), the BLASTx analysis (e-value cut-off e^{-10}) revealed 384 transcripts associated with 94 non-redundant proteins (the best hits) and 76 *JcDHX* genes. All *JcDHX* candidates exhibited the expected domains identified using the CDD and SMART tools.

The coding potential of each *JcDHX* transcript was initially detected for 234 transcripts using the SAMBA tool. Another 135 transcripts encoded ORFs translated by TransDecoder, and these ORFs in the BLASTp analysis showed similarities with cured proteins from the

UniProtKB/Swiss-Prot database. Similarly, we applied the ORFfinder tool, with a total of 369 *JcDHX* transcripts encoding 308 potentially functional proteins (Table S8). This result emphasizes the quality of the analyzed transcriptome [34]. These proteins have conserved domains associated with different processes, such as splicing, rRNA processing, translation, DNA repair, chromatin organization, post-transcriptional gene regulation, and mRNA export (Table S8).

Concerning the transcriptomic profile of salt-tolerant *Jc183*, this accession did not display any DET (p -value < 0.0001; FDR < 0.005) encoding the DHX protein. Concerning the *Jc171* accession (the less salt-tolerant phenotype), from 120 *JcDHX* transcript isoforms, it comprised 30 repressed DETs [DEAH (20), DEAD (10), and DEXH/D (one) subfamilies] and one induced DET (DEAD) (Figure 7). Detailed information about the *JcDHX* transcript isoforms expressed by the *Jc171* accession is outlined in Tables S8 and S9.

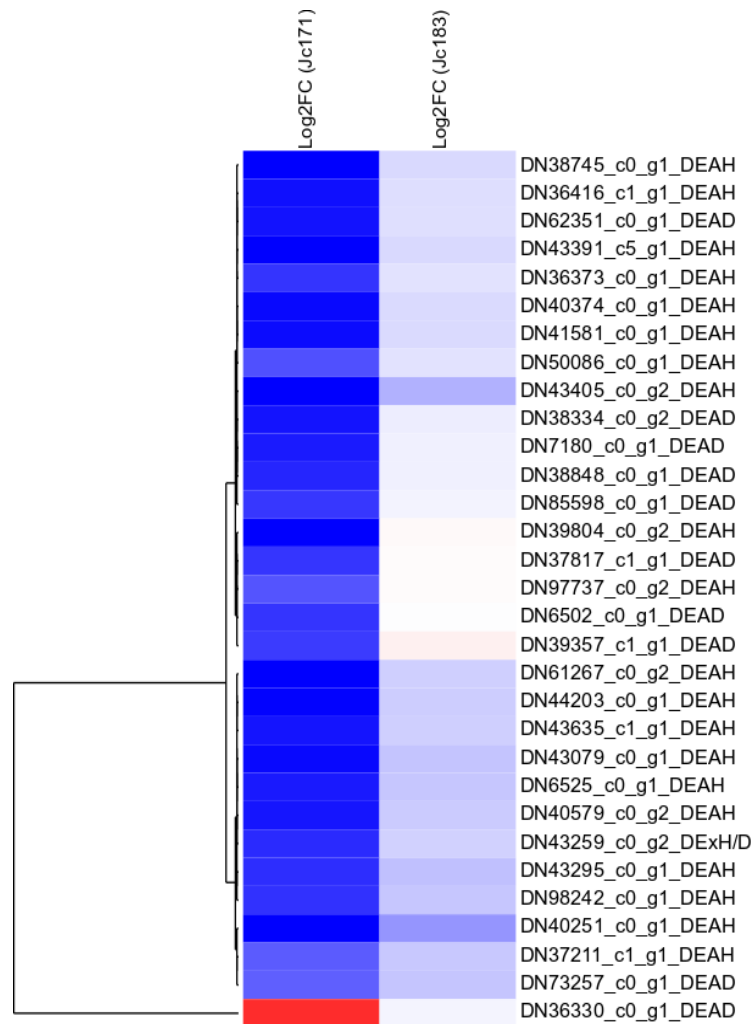


Figure 7. Hierarchical clustering based on Log2FC values of differentially expressed DEAD-box RNA-Seq assembled transcripts (p -value < 0.0001; FDR < 0.005) in the roots of accession *Jc171* subjected to saline stimulation (150 mM of NaCl for 3 h), as well as the respective modulation in *Jc183* accession.

Based on the qPCR analysis, out of the 20 proposed primer pairs that successfully amplified the cDNA samples (Table S10), only 13 primer pairs exhibited suitable amplification efficiency (E), slope (s), and correlation

coefficient (R) values (Figure S8; Table 1). From the thirteen primer pairs that presented melting curves showing the specificity of the amplicons (Figure S9), nine *JcDHX* candidates confirmed the in silico expression patterns in the qPCR assay, while four other candidates showed a different expression (Table 1; Figure 8).

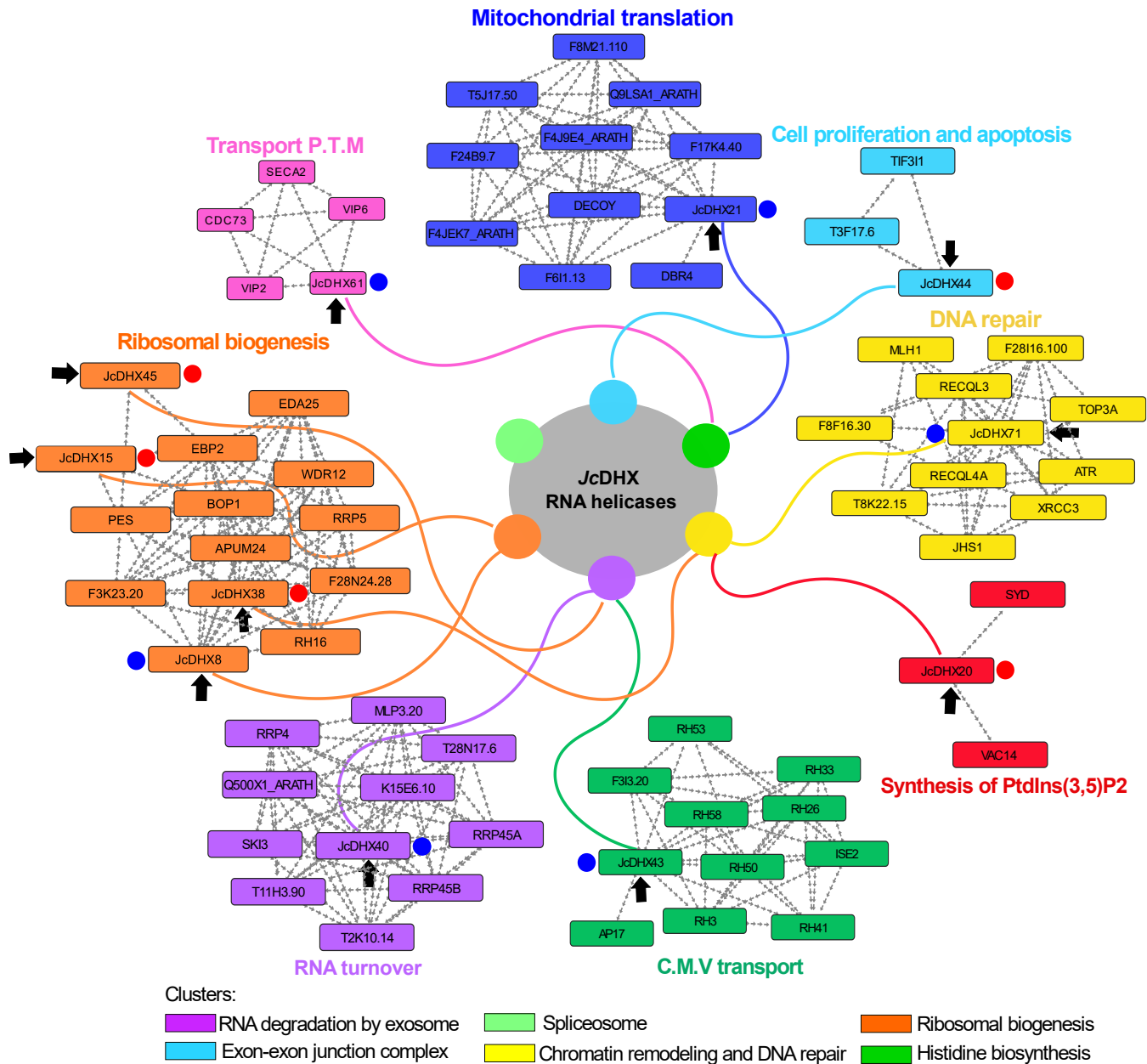


Figure 8. Protein-protein interaction (PPI) network proposed by STRING considering differentially expressed DEAD-box candidates (*J. curcas* RNA-Seq transcripts) also analyzed by qPCR. In the center, there is a representation of the clusters highlighted considering the entire set of expressed *JcDHX*s. Arrows indicate the *JcDHX* in the networks. Blue and red circles next to the names of each protein report the qPCR result: blue for downregulation and red for upregulation. *Arabidopsis thaliana* was used as a reference. The confidence score was >0.7 (high confidence). “Transport P.T.M.” represents transport of proteins across the thylakoid membrane, and “C.M.V. transport” represents clathrin-mediated vesicular transport.

Table 1. In silico results and relative gene expression (qPCR) of RNA-Seq *JcDHX* transcripts of *J. curcas* accession *Jc171* after exposure to 150 mM of NaCl for three

hours. Data analysis performed using REST© software (v.2.0.13). DR: downregulation; UR: upregulation.

RNA-Seq Transcript	Gene	Efficiency (%)	Relative Expression	Std. Error	95% C.I.	P(H1)	Result	
							In Silico	qPCR
DN43295_c0_g1_i2	<i>JcDHX43</i>	102.34	-0.54	0.237–1.111	0.105–1.999	0.047	DR	DR
DN43259_c0_g2_i2	<i>JcDHX43</i>	95.76	-0.22	0.083–0.694	0.033–2.514	0.000	DR	DR
DN39804_c0_g2_i2	<i>JcDHX8</i>	108.19	-0.30	0.051–1.709	0.024–2.832	0.043	DR	DR
DN43635_c1_g1_i2	<i>JcDHX71</i>	99.72	-0.13	0.012–1.999	0.001–3.936	0.032	DR	DR
DN36330_c0_g1_i1	<i>JcDHX44</i>	102.98	3.50	1.451–14.283	1.043–30.645	0.000	UR	UR
DN97737_c0_g2_i1	<i>JcDHX40</i>	101.48	-0.49	0.346–0.686	0.236–0.840	0.000	DR	DR
DN43391_c5_g1_i3	<i>JcDHX40</i>	97.43	-0.51	0.380–0.704	0.296–0.867	0.000	DR	DR
DN39804_c0_g2_i1	<i>JcDHX21</i>	90.38	-0.60	0.345–1.059	0.214–1.426	0.014	DR	DR
DN41581_c0_g1_i2	<i>JcDHX61</i>	108.16	-0.68	0.398–1.046	0.282–2.043	0.031	DR	DR
DN40374_c0_g1_i2	<i>JcDHX20</i>	109.42	1.97	1.135–3.584	0.677–5.953	0.004	DR	UR
DN62351_c0_g1_i1	<i>JcDHX38</i>	95.05	1.78	1.273–2.527	0.785–3.596	0.000	DR	UR
DN7180_c0_g1_i1	<i>JcDHX45</i>	91.69	5.56	0.966–28.716	0.185–79.341	0.012	DR	UR
DN85598_c0_g1_i1	<i>JcDHX15</i>	98.48	1.42	0.942–2.299	0.573–3.147	0.042	DR	UR

3.10. Protein–Protein Interaction Network

The STRING PPI network (high score > 0.7) based on *A. thaliana* orthologs (Figure 8) of *JcDHX* proteins encoded by DETs provided a better understanding of the *Jc171* salt-response profile. The predicted clusters (Figure S10) indicated that *JcDHX* proteins are part of a complex network in which components of the exosome, spliceosome, and ribosome biogenesis are interconnected. For instance, EIF243 (DEAD-box, *JcDHX33*) probably acts as a binding platform between the components involved in mRNA degradation and transport, as well as RNA processing (Figure S10). In addition, Ski2 (DEAD-box *JcDHX40*), a component of the exosome complex, interacts with Ski3, a protein associated with histidine biosynthesis, while clusters are also associated with chromatin remodeling with DNA repair (Figure S10).

The biological processes involving each protein codified by the *JcDHX* candidates performed in the qPCR assays (Table S11) highlighted the relevance of mitochondrial translation, phosphatidylinositol 3,5-bisphosphate synthesis, thylakoid membrane protein transport, clathrin-mediated vesicular protein transport, cell proliferation, and apoptosis (Figure 8) as those involving *JcDHX* proteins helping to regulate specific aspects of plants responding to abiotic stress.

4. Discussion

4.1. Comprehensive Analysis of the DEAD-Box RNA Helicase Family in *J. curcas* Genome

RNA helicases from the DEAD-box family play pivotal roles in biological processes in prokaryotes and eukaryotes [46,47]. Although extensively studied in the *A. thaliana* model plant or economically relevant crops, such as rice, tomato, cotton, corn, and soybean [7,9,10,48], the *DHX*

gene family remains unexplored in *J. curcas* and closely related Euphorbiaceae, including *H. brasiliensis*, *M. esculenta*, and *R. communis*.

Thus, we performed a comprehensive analysis of the DEAD-box family in *J. curcas*, covering classification, gene structure, gene orthologs, and protein characterization, including physicochemical parameters, subcellular localization, conserved domains/motifs, secondary structures, and 3D modeling, together with RNA-Seq analysis of *JcDHX* transcripts of two *J. curcas* accessions, after 3 h of roots exposed to NaCl (150 mM). Understanding the *JcDHX* gene family, based on its gene and protein structures, classification, and evolutionary aspects, will provide insights into its potential roles in *J. curcas* plants in response to salinity.

The 79 putative *JcDHX* genes identified corresponded to 0.35% of the *J. curcas* genes that encode proteins. Similar representativeness was observed in plant species, such as *S. lycopersicum*, *O. sativa*, *Z. mays*, *G. max*, and *Gossypium raimondii* [8–10,48]. The structural organization of the 79 *JcDHX* genes revealed a variable number of introns, ranging from 0 to 30 introns. In general, *JcDHX* genes are intron-rich, with members of the DEAH and DExH/D subfamilies exhibiting more complex structures than members of the DEAD subfamily. This pattern aligns with the findings in other plant species, suggesting a conserved characteristic of the family's genes. Furthermore, considering that genes with multiple introns increase the versatility of the proteome [49], DEAH and DExH/D genes are probably more efficient in producing different isoforms. This corroborates the greater variety of auxiliary domains present in the DEAH and DExH/D subfamilies, as discussed below.

The orthologs of *JcDHX* genes identified in close-related species (*M. esculenta*, *R. communis*, *H. brasiliensis*) and also outside the Euphorbiaceae family (*P. trichocarpa*, *S. lycopersicum*, and *A. thaliana*) showed high gene conservation (87.3%), underscoring their importance in the plant metabolism. The shared orthologs across the studied plant species highlighted the relevance of these RNA helicases in plant biological processes. Notably, *J. curcas* and its taxonomically related species could benefit from *A. thaliana* and *S. lycopersicum* orthologs, given their well-described gene functions, especially in plants exposed to abiotic stress. Further investigations into the functional aspects of *JcDHX* genes in stress responses could pave the way for enhancing the resilience of *J. curcas* and related species.

The physicochemical characteristics and subcellular localization of the *JcDHX* proteins underscore their functional diversity. These helicases predominantly localize to the nucleus, aligning with their primary role in nucleic acid-related processes such as ribosome biogenesis and the transport of mRNAs from the nucleus to the cytoplasm [50,51]. However, predictions also indicate the presence of *JcDHXs* in the cytoplasm, where they serve as components of exosomes [52], chloroplasts, and mitochondria, playing roles in gene expression within these organelles [53], and in the plasma membrane, participating in the secretion of proteins through the endoplasmic reticulum [54].

All nine characteristic conserved motifs were identified in *JcDHX* proteins. The motifs are named Q, I, Ia, Ib, II, III, IV, V, and VI. These signatures are directly related to the biochemical activities of helicases. The Q motif acts as a regulator of ATPase activity [55]; motif I (also known as Walker A) is involved in binding to NTP motifs; Ia and Ib are required for RNA binding [56]; motif II (also known as Walker B) is responsible for coordinating Mg^{2+} ions, essential for ATP hydrolysis [57]; motif III

participates in linking helicase and ATPase activities; motif IV does not have a well-defined consensus and may be functionally connected with motifs V and VI; motif V (in conjunction with Ia, Ib, and IV) acts in binding to RNA and regulating ATP hydrolysis; and motif VI, which interacts with motifs II and III, has been described as important for both ATP hydrolysis and RNA binding [6,55–57].

The versatility of *JcDHX* proteins, as reflected in the variety of auxiliary domains that regulate core helicase activity (DEXDc and HELIc), contributes to their multifaceted functions, such as those associated with gene silencing (DSRM, Dicer, RIBOc), recombination and repair (RecQ-Zn-bind), chromatin remodeling (SANT, SNF2_N), translocation through the membrane (SecA_DEAD), and the endoplasmic reticulum (Sec63). In addition, different domains assist interactions with nucleic acids, such as HDRC, R3H, and HAS. These domains provide functional diversity and specificity for the catalytic reactions of *JcDHX* proteins [58,59]. Understanding the characteristics of DHX proteins within the same subfamily is crucial as they may perform similar functions in different species.

The phenetic analysis based on motif II variations (D-E-A-D) supports the classification of *JcDHX* genes into three known subfamilies [59]: DEAD (42 members), DEAH (30), and DExH/D (seven). In plants, the number and composition of subfamilies vary significantly across species, reflecting the high diversity of the *DHX* genes [9–12]. Conserved motifs crucial for protein interactions and functional similarities were identified in *JcDHX* candidates, with variations observed in motif VI between the DEAD and DEAH subfamilies. These variations require further study. The DEAH and DExH/D subfamilies showed more considerable motif variations and a higher occurrence of auxiliary domains compared to the DEAD subfamily, reinforcing their genetic diversity and functionalities [8–12].

The differences in the conserved motifs contribute to structural variations, promoting the flexibility required for diverse activities [58,59]. These structural changes favor the emergence of specialized proteins with distinct cellular activities, suggesting that the DEAH and DExH/D subfamilies may participate in more specific pathways than the more basic functions of the DEAD subfamily members. The versatility of *J. curcas* DEAD-box RNA helicases likely stems from their diversity within the DEAH and DExH/D subfamilies.

Typical secondary structures of DEAD-box helicases include a β -sheet and two α -helices, forming a “cap” upstream of motif I. These structures are associated with the presence of the Q motif, which acts as a regulator of ATPase activity [55]. The 3D models confirmed the presence of two core domains comprising β -sheets surrounded by α -helices, with motifs positioned in the cleft between the two domains. This characteristic resembles the RecA-like ATPase folding pattern and is consistent with the known structures in the family [57]. These structural characteristics confirm that *J. curcas* DEAD-box helicases are well suited for nucleic acid binding, NTP hydrolysis, and strand unwinding activities [55,57,59,60].

4.2. Regulatory Landscape of *JcDHX* Genes: Insights into Cis-Regulatory Elements

Cis-regulatory elements (CREs) within gene promoters are pivotal for gene expression regulation by serving as recognition sites for TFs. In the analysis of the promoter regions of *JcDHX* genes, a diverse array of

TFs spanning different families (Dof-type, BBR-BPC, HD-ZIP, AP2-ERF, WRKY, bHLH, and Myb-related) were identified, with particular emphasis on Dof-type TFs. These TFs, prominently featured in plant responses to abiotic stress [61–63], have demonstrated key roles in orchestrating stress-responsive gene expression. Dof-type TFs, which are characterized by a DNA-binding zinc finger, have emerged as critical regulators of plant environmental stress responses. Examples include *MtDof32* in *A. thaliana*, which confers enhanced tolerance to osmotic and salt stress [64], and *GhDof1* in cotton, contributing to salinity and cold tolerance [65]. The overexpression of the cotton *GhDof1.7* gene in *A. thaliana* transgenic plants exhibited increased salinity tolerance, accompanied by reduced ROS accumulation and elevated activities of superoxide dismutase (SOD) and catalase (CAT) enzymes [66]. These findings underscore the potential of Dof-type TFs to modulate plant responses to abiotic stress. Thus, the association of Dof-type TFs with CREs on the promoters of *JcDHX* genes also suggests the involvement of *JcDHX* genes in the broader context of stress-responsive genes.

4.3. Differential Regulation of *JcDHX* Genes in *J. curcas* under Salt Stress

In order to decipher the expression profile of *JcDHX* transcripts in *J. curcas* roots after 3 h of NaCl (150 mM) exposure, RNA-Seq analysis revealed a distinctive response of the *Jc171* accession, the less salt-tolerant phenotype. In the biological assay, after 3 h of NaCl exposure, only *Jc171* plants showed visible leaf damage, whereas *Jc183* showed a salt-tolerant phenotype [35]. Concerning the respective RNA-Seq transcriptomes, *Jc183* almost did not modulate its transcriptome significantly after salt treatment [35]. On the other hand, the present study identified, from the *Jc171* salt-response profile, 116 *JcDHX* transcripts, with 68 (58.6%) of them declared DETs (related to 22 *JcDHX* genes) and 67 DETs showing a remarkable downregulated response (DR).

In general, analysis of the RNA-Seq transcript isoforms identified six auxiliary domains that were not predicted in our reference sequences (Table S8): DBINO, RING, PHD, HIRAN, CHROMO, and PWI. The DBINO domain (DN62351_c0_g1_i1) is related to DNA binding activities [67], the RING domain (DN36373_c0_g1_i1) is related post-translational modification in proteins [68], the PHD domain (DN36416_c1_g1_i1) to the recognition of methylated histones [69], the HIRAN domain (DN43635_c1_g1_i1) to recognition of damaged DNA [70], the CHROMO domain (DN6525_c0_g1_i1) interacts with methylated histones [71], and the PWI domain (DN43405_c0_g1_i1) with the processing of pre-RNAs [72]. The combination of these domains was also observed in DN43635_c1_g1_i2, which presented the HIRAN and RING domains, and whose expression was validated by qPCR (Table 1). Protein isoforms presenting these unpredicted domains can add functional diversity to the helicases expressed by the studied accessions after salinity exposure, highlighting, for example, the interaction with histone methylation that could be important in chromatin remodeling and gene expression.

Regarding the subcellular location of the proteins encoded by the isoforms, 12 isoforms pointed to the extracellular region, which was not predicted by the reference proteins. Interestingly, of the 12 isoforms, 10 were expressed only by *Jc183* (Table S8). Helicases in the extracellular region have already been reported in *Arabidopsis* defense against the pathogenic fungus *Botrytis cinerea*, which is correlated with the transport

of small RNAs (sRNAs) in extracellular vesicles [73]. Such correlations concerning abiotic stress have not yet been reported.

Comparing the different isoforms, especially those related to DETs, differences in subcellular localization were observed less than changes in relation to domains. Two isoforms of the assembled transcript, DN41581_c0_g1, for example, predicted to act in the nucleus, showed differences in terms of the DEXDc/HELICc/Dicer/PAZ/RIBOc domains (Table S8). One of the isoforms (DN41581_c0_g1_i2) had three more domains (PAZ/RIBOc/RIBOc) than the other (DN41581_c0_g1_i1). The presence of more domains reinforce function or provide functional variability. The PAZ domains (Piwi/Argonaute/Zwille), RIBOc, and Dicer were related to post-transcriptional gene silencing [74], and one of the isoforms, DN41581_c0_g1_i2, was validated in the qPCR assay. In turn, the three isoforms of the assembled transcript DN43635 (Table S8) showed differences both in subcellular localization (nucleus and cytoplasm) and in the detected domains (DEXDc/HELICc/HIRAN/RING), with one of the isoforms (DN43635_c1_g1_i2) showing an extra HELICc domain. The gene expression of this isoform was validated by qPCR. Furthermore, the DN62351_c0_g1_i1 isoform, induced in the qPCR assay, presented a DBINO domain that was not detected in the reference protein (Table S8). These variations were also observed in assembled transcripts that did not show differential expression after saline exposure. For example, the products encoded by the six isoforms of the assembled transcript DN42995_c0_g2, which would be addressed to the nucleus, chloroplast, or cytoplasm, showed variations in the distribution of the HELICc and HA2 domains, which are related to helicase activity and nucleic acid binding, respectively. Overall, this sample provides insights into the diversity of DHX transcripts and protein isoforms that can be expressed by the two studied accessions, as well as their involvement in several important metabolic processes in plant metabolism.

DHX genes respond to abiotic stress in plants. qPCR results from 42 DHX genes in tomato leaves under salinity, drought, cold, and heat stress conditions revealed 14 genes induced across all scenarios, with five genes (*SIDEAD24*, *SIDEAD32*, *SIDEAD34*, *SIDEAD35*, and *SIDEAD42*) significantly induced during salinity exposure [75]. In *Arabidopsis*, *AtRH9* and *AtRH25* were induced under cold stress and repressed under drought and salinity conditions [76]. Further corroborating these findings, qPCR analyses in tomato plants demonstrated the induction of *SIDEAD30* and *SIDEAD31* genes exposed to salinity [14], as well as *SIDEAD25* and *SIDEAH15* genes exposed to 200 mM of NaCl [10]. Additionally, various genes (*SIDEAD23*, *SIDExD/H9*, and *SIDEAD35*) were induced in tomato leaves under drought, salinity, heat, and cold stress [10]. More specific studies are necessary to explore the observed gene expression profiles efficiently.

The scientific literature supports that some DHX genes stand as promising candidates for biotechnological exploration, aiming to develop stress-tolerant genotypes. Some successful cases comprise transgenic tomato plants overexpressing the *SIDEAD31* gene, showcasing heightened salinity and moderate drought tolerance [14]; the *AtRH17* gene in *A. thaliana* plants conferring salinity tolerance [77]; the wheat *TaDEAD-57-3B* gene, improving proline and chlorophyll levels and enhancing drought and salinity tolerance in *A. thaliana* [12]; the pea DEAD-box *Psp68*, improving rice salinity tolerance, marked by reducing the MDA levels [78]; and the overexpression of the DEAD-box *BrDHC1* in

Brassica rapa, increasing the drought tolerance by enhancing water retention, chlorophyll content, and activities of antioxidant enzymes [79].

However, some DEAD-box proteins are negative regulators during abiotic stresses. In grape plants (*Vitis vinifera* L.), leaf RNA-Seq data showed more than 70% repressed DEAD-box genes (28 of 40) in response to drought (1, 2, 4, 8, and 24 h of irrigation suppression) [11]. *Arabidopsis thaliana* plants overexpressing *VviDEADRH25a* presented higher drought sensitivity [11]. In turn, the *Arabidopsis* DEAD-box *STRS1* and *STRS2* genes were repressed under various abiotic stresses (heat, drought, salinity—200 mM NaCl: 1, 3, 6, 12, 24, 48 h) and they acted by attenuating the expression of the transcriptional activators DREB1A/CBF3, DREB2A, and RD29A, which function in both ABA-dependent and ABA-independent pathways [80].

Our RNA-Seq study only pointed out the nuances of *JcDHX* genes expressed in the salt response of *Jc171*, offering insights into its acclimatization strategies to salt stress, since *Jc183* did not modulate its *JcDHX* genes significantly after the salt treatment, indicating a distinct adaptive response. The tolerant accession *Jc183* showed little modulation of its transcriptome, with only 57 DETs in previous works [35,81], indicating that it is not only the gene category of DHX helicases that has no modulation, being in agreement with the general response of the accession. Furthermore, *Jc183* has already been described to support up to 750 mM of NaCl, showing rapid recovery from salinity after the alleviation of the saline conditions in the soil, and scanning electron microscopy revealed that the stomata of *Jc183* are smaller and have a higher stomatal index compared to those of the *Jc171* genotype [82]. On the other hand, *Jc171* has also been reported as salt-tolerant after exposure to some levels of salinity [34], and our transcriptome analysis indicated most of the identified DHX candidates as repressed. In turn, the observed repression may not have contributed to the salt tolerance phenotype of *Jc171*, or it at least did not efficiently explore the *DHX* genes. Thus, this differential salt response showed by the two accessions underscores a genotype-dependent pattern of the *DHX* gene families after salinity exposure.

4.4. Unraveling Functional Networks with *JcDHX* Proteins in Salinity Response

Observing the predicted PPI network, our *JcDHX* candidates and pivotal functional partners revealed distinct clusters pointing to crucial cellular processes, such as DNA repair and chromatin remodeling, splicing, RNA degradation, ribosome biogenesis, and histidine biosynthesis. Inside the PPI network, a *JcDHX37* ortholog (*AtDEAH11*) assumed a central position. Together with other proteins, they underscored the significance of RNA metabolism, with a particular emphasis on rRNA processing, alternative splicing, and mRNA degradation. Alternative splicing (AS), a regulatory process impacting diverse physiological aspects, also spotlighted abiotic stress responses, including salinity. An RNA-Seq of *A. thaliana* plants from NaCl-treated seeds disclosed that 49% of genes with introns underwent splicing alterations, with 10% undergoing differential alternative splicing (SAD) [83]. In addition, after salinity exposure, plants of *Gossypium davidsonii* exhibited a significant increase (32%) in genes undergoing splicing alterations [84], while a crucial role in the stress response was played particularly by the *SHI2* gene, with its DEAD-box activity in splicing cold

response genes [85]. Another distinct cluster highlighted the biosynthesis of histidine, an amino acid crucial for plant growth, development, and responses to environmental factors, including salinity. A treatment with histidine in corn plants, when exposed to salinity, exhibited enhanced tolerance and increased activities of antioxidant enzymes after salt stress [86], while the induction of histidine biosynthesis enzymes in tomato leaves under heat and flooding stresses further underscored its importance [87].

In addition, regarding the particular contribution of each *JcDHX* DET in the *Jc171* salt-response profile, according to the predicted PPI networks, some processes highlighted spanned DNA repair, RNA turnover, ribosomal biogenesis, mitochondrial translation, protein transport across the thylakoid membrane, clathrin-mediated protein vesicular transport, phosphatidyl 3,5-bisphosphate (PtdIns(3,5)) synthesis, and cell proliferation/apoptosis.

Considering the implications of ionic and osmotic stresses on ROS accumulation and the ensuing damage to DNA, unrepaired DNA damage can lead to genomic instability, disrupting cellular functions and potentially resulting in cell death [88]. In this context, the PPI network revealed interactions of *JcDHX71* with DNA repair and cell cycle checkpoint partners, emphasizing the role of this helicase in maintaining genomic integrity during salinity stress. Additionally, DNA repair pathways, integral to cell cycle checkpoints, play a crucial role in plant development and stress adaptation [89,90]. Acting as damage sensors, checkpoint proteins intervene in cell cycle arrest, allowing DNA lesion repair to ensure normal cellular functioning [91]. However, the downregulation of candidate *JcDHX71*, in both in silico and qPCR analyses, underlines its unique regulatory pattern of *Jc171* after salinity exposure.

In RNA turnover, the interplay between the RNA exosome machinery, responsible for 3'-5' degradation and the processing of various RNA classes, and DEAD-box RNA helicases influences stress granules (SGs) and processing bodies (P-bodies) dynamics [92,93]. These cytosolic ribonucleoprotein complexes are stimulated by various stresses, including oxidative stress, and modulate mRNA translation, storage, and degradation, conserving energy for translational machinery during stress conditions [94]. The predicted PPI network implicated *JcDHX40* in the RNA exosome pathway, and both in silico and qPCR analyses showed its downregulation in *Jc171* after salt stress.

Several DEAD-box RNA helicases contribute to ribosomal biogenesis, a fundamental process involving rRNA maturation and assembly with ribosomal proteins [95,96]. However, salinity-induced nucleolar stress affects pre-rRNA accumulation, leading to nucleolar cavity formation and the activation of apoptotic pathways [97]. The predicted PPI network highlighted interactions of *JcDHX38*, *JcDHX8*, *JcDHX15*, and *JcDHX45* with partners involved in rRNA processing. Besides the downregulation detection in silico data of the four candidates, the qPCR results only confirmed the downregulation of one of them, reminiscent of the essential role of DEAD-box helicases in maintaining ribosomal homeostasis.

The eIF3 complex, a key player in translation initiation, controls cyclin-dependent kinases (CDKs) linked to cell proliferation, cell cycle progression, and programmed cell death (PCD) [98,99]. Crucial for maintaining cellular homeostasis, PCD is part of the plant's defense

against stress [100]. Salinity-induced PCD and autophagy in halophyte cells underscore its role in stress responses [101]. The PPI network predicted interactions of *JcDHX44* with proteins of the eIF3 complex, and this candidate was upregulated in both in silico and qPCR analyses.

Mitochondria, pivotal in energy production and cell signaling, undergo stress-induced alterations affecting mitochondrial translation and protein transport [102,103]. Disruption in mitochondrial processes triggers the mitochondrial unfolded protein response pathway, essential for restoring mitochondrial homeostasis [104,105]. The PPI network implicated *JcDHX21* in mitochondrial translation interactions. Besides the downregulation of *JcDHX21* in both in silico and qPCR analyses, the constitutive expression of *OsSUV3*, a mitochondria-localized DEAD-box helicase, enhanced salinity tolerance in rice plants [106], emphasizing the potential of *JcDHX21* in a potential salinity response.

Under osmotic stress conditions, alterations in plasma membrane balance are crucial. Clathrin-mediated transport vesicles play a pivotal role in protein trafficking between membrane systems. This system is directly linked to the abundance and localization of aquaporins (type PIP) in the membrane, crucial for water absorption by roots during salinity conditions [107]. Additionally, the clathrin system is implicated in stomatal function in *A. thaliana* [108] and ROS accumulation under salinity stress [109]. The PPI network predicted interactions of *JcDHX43* with partners involved in clathrin-mediated vesicular protein transport; this candidate also exhibited downregulation in both in silico and qPCR data.

During osmotic stress in plants, inositol phospholipids play diverse signaling roles in cells, with rapid accumulation of phosphatidyl 3,5-bisphosphate (PtdIns [3,5]) [110], which is essential for normal vacuole function, since its depletion results in aberrant vacuoles [111]. The PtdIns(3,5) metabolic pathway has been associated with critical aspects of stress response, including stomatal conductance [112], vacuolar convolution [113], and activation of V-ATPases [114]. The PPI network predicted interactions of *JcDHX20* with proteins involved in PtdIns(3,5) synthesis. *JcDHX20* was downregulated in silico and upregulated in the qPCR assay.

In summary, predicted interactions of *JcDHX* candidates showcase central roles in the plant's adaptive strategies against salinity exposure. However, our *JcDHX* candidates were not sufficiently induced to confer a positive salt response of the *Jc171* accession, diminishing the impact of these processes highlighted by the PPI networks. Therefore, the lower-tolerance phenotype must also be associated with the impairment of these processes due to the repression of these *JcDHX* candidates. Additionally, the upregulation of certain candidates initially identified as repressed but confirmed as upregulated by qPCR assays may underscore the relevance of DHX proteins in *Jc171*'s salt response, thereby deepening our understanding of these proteins.

5. Conclusions

This study comprehensively obtained and analyzed the entire family of DEAD-box RNA helicases from *Jatropha curcas*. A total of 79 *JcDHX* genes were identified, a quantity proportional to that observed in other species. Orthology analyses involving *M. esculenta*, *R. communis*, *H. brasiliensis*, *P. trichocarpa*, *A. thaliana*, and *S. lycopersicum* highlighted the conserved nature of the family and its significance across these species. The proteins were classified based on known subfamilies (DEAD, DEAH,

and DExH/D), and our observations revealed that the pattern of domains, motifs, and exon–intron structures reflected this classification. Helicases from the DEAD subfamily exhibited the classic structure composed of the nine characteristic conserved motifs, featuring small C- and N-terminal extensions and very few auxiliary domains, in addition to simpler gene structures. In turn, the DEAH and DExH/D helicases are more complex in terms of both domains and gene structure, although they have shown notable variations in the presence of conserved motifs. The three-dimensional models of *JcDHX* generated are consistent with the functions performed by these enzymes. Taken together, the analyses indicate that *Jatropha curcas* DEAD-box RNA helicases constitute a highly structurally diverse family contributing to the execution of a variety of functions. The presence of candidates for cis-regulatory elements (CREs) in the promoters of *JcDHX*, associated with important transcription factors such as Dof-type, BBR-BPC, HD-ZIP, and bHLH, demonstrates that these genes can participate in crucial pathways during environmental stresses. The repertoire of transcripts in *JcDHX* was modulated in response to salinity stimulation, particularly in the *Jc171* accession. Protein–protein interaction networks revealed significant functional partners of *JcDHX* expressed in response to salinity, demonstrating their crucial role in plants under stress. Considering all of the results and acknowledging the significance of DEAD-box RNA helicases in RNA metabolism, the observed transcriptional response may be associated with the reduced capacity of *Jc171* to cope with the applied salinity stimulus. This study contributes valuable insights into the structural and functional aspects of *JcDHX* helicases, providing a foundation for future functional characterizations of these genes concerning their involvement in responses to abiotic stresses.

Supplementary Materials: The following supporting information can be downloaded at www.mdpi.com/xxx/s1: Figure S1: MEME logo of the motifs detected in the promoter regions (1000 bp upstream the TSS—transcription start site) of the *JcDHX* genes. For each motif, the e-value and JASPAR ID are reported; Figure S2: Distribution of candidate cis-regulatory elements (CREs) (p -value < 0.01) in the promoter regions (1.0 kb) of *JcDHX* genes. Colored boxes provide information about CREs, JASPAR IDs, and associated transcription factors. *Combined MEME p -value; ‡ statistical significance below the considered cut-off (e-value < 0.05; p -value < 0.01); Figure S3: Graphical representation of (a) the number of identified *JcDHX* orthologs and (b) a Venn diagram illustrating the sharing of these genes among the analyzed species; Figure S4: Graphical representation of predicted secondary structure elements in the proteins of 42 *J. curcas* RNA helicase genes from the DEAD subfamily. A 2D alignment view was generated using the 2DSS tool and Ali2D analysis, considering multiple sequence alignment data (ClustalX v2.1) (Figure S5). Multiple alignment of 42 proteins from *J. curcas* RNA helicase genes of the DEAD subfamily. This data alignment (ClustalX software v2.1) was employed in the Ali2D analysis to predict secondary structure elements; Figure S6: Ramachandran plots generated by Swiss-model from models built for *JcDHX*; Figure S7: Three-dimensional homology models proposed for *JcDHX* from the Swiss-model. The structures were visualized and edited using PyMOL software; Figure S8: Standard curves demonstrating the efficiency of the *JcDHX* primer pairs. The data were generated from serial dilutions made with cDNAs from the cultivars (undiluted, 1:10, 1:100, 1:1000, and 1:10,000); Figure S9: Dissociation curves generated between 65 and 95 °C, demonstrating the specificity of the products amplified by the *JcDHX* primers; Figure S10: Protein–protein interaction network constructed by STRING (high confidence > 0.70), considering all differentially expressed *JcDHX* transcripts (p -value < 0.0001; FDR < 0.005) in RNA-Seq libraries from *Jatropha curcas* roots (*Jc171*)

subjected to saline stimulation (150 mM of NaCl for 3 h). Six clusters were highlighted from the set, with proteins (nodes) colored with the same colors predicted to be components of the same cluster. Dashed lines indicate interactions between distinct clusters; Table S1: Characteristics of *Jatropha curcas* genes (*JcDHX*) encoding DEAD-box RNA helicases: name (ID), subfamilies, locus ID, genomic location, gene length, number of introns, CDS length, primary protein at each locus, and transcript variants; Table S2: Motifs detected as candidate cis-regulatory elements (CREs) by the MEME and TOMTOM programs coupled to the JASPAR database; Table S3: Orthologs to genes encoding DEAD-box helicases from the *Jatropha curcas* RJc1_Hi-C genome (GCF_014843425.1); Table S4: Conserved motifs detected by the MEME program in *Jatropha curcas* DEAD-box helicases (*JcDHX*). Motif VI had two variations detected (MEME-1 and MEME-7); Table S5: Physicochemical characteristics of proteins encoded by *JcDHX* genes, their sequences and predicted subcellular localization. 'aa': amino acids; 'pI': isoelectric point; 'MW': molecular weight 'kDa': kilodalton; Table S6: Secondary structures detected in *JcDHX* proteins performing the SOPMA tool; Table S7: Quality parameters obtained for the 3D models generated for the *Jatropha curcas* DEAD-box RNA helicases (*JcDHX*) using the Swiss-model; Table S8: DEAD-box transcripts (*JcDHX*) identified in RNA-Seq libraries (BLASTx e-value cut-off e^{-10}) for accessions *Jc171* and *Jc183* after exposure to 150 mM of NaCl for 3 h. For each transcript, it is possible to access the transcript isoform; *J. curcas* accession; Trinity unigene (‡ corresponds to differentially expressed unigene); regulation (DEG); classification regarding the coding nature ('*' corresponds to a transcript verified manually using the ORF finder) transcript sequence; annotation; domain CDD; encoded protein; predicted protein (aa); domain (SMART); subcellular location; BLASTx e-value and information about of the *J. curcas* RJc1_Hi-C genome-associated proteins; Table S9: Information on each assembled transcript in different samples. For each assembled transcript, it is possible to access the subfamily, regulation, Log2FC, *p*-value, FDR, and normalized FPKM data in the *Jc171* transcripts that are also expressed in *Jc183*. "n.d" corresponds to *Jc183* transcripts that were not expressed in *Jc171*; Table S10: Primers designed for *JcDHX* transcripts identified in RNA-Seq libraries for accession *Jc171* after exposure to 150 mM of NaCl for 3 h. 'Tm': Melting temperature; Table S11: Data on protein–protein interaction networks constructed by STRING (High score > 0.7) from *JcDHX* associated with transcripts differentially expressed in qPCR.

Author Contributions: R.H.d.S.: methodology, formal analysis, investigation, validation, data curation, visualization, and writing (original draft, review, and editing). M.D.d.S., J.R.C.F.-N. and B.d.B.S.: methodology, formal analysis, and investigation. F.N.d.A. and E.J.d.S.O.: methodology and formal analysis. A.M.B.-I.: methodology, resources. A.F.d.C.: project administration and writing review. É.A.K.: initial conceptualization, funding acquisition, project administration, and writing (original draft, review, and editing). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported (grants and fellowships) by the following Brazilian agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 404357/2013-0; CNPq 311894/2017-8), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Universidade Federal de Pernambuco (UFPE/PROPG).

Data Availability Statement:

Acknowledgments: The authors are grateful to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES), and Universidade Federal de Pernambuco (UFPE).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AAA (ATPases associated with a variety of cellular activities); AP2-ERF (Apetala 2-Ethylene Response Factor); AS (alternative splicing); ATP (Adenosine TriPhosphate); BBR-BPC (Barley B Recombinant/Basic PentaCysteine); bHLH (basic/helix-loop-helix); BROMO (Bromodomain); C.M.V (clathrin-mediated vesicular transport); SAD (differential alternative splicing); CAT (catalase); CDD (Conserved Domain Database); CDS (coding sequence); CREs (cis-regulatory elements); DEAD (Asp-Glu-Ala-Asp); DETs (differentially expressed transcripts); Dof (DNA-binding One Zinc Finger); DR (downregulated); DSHCT (DOB1/SK12/helY-like C-terminal domain); DSRM (Double-Stranded RNA-binding Motif); DUF (Domain of unknown function); FC (fold change); FDR (false discovery rate); GMQE (Global Model Quality Estimate); GSDS (Gene Structure Display Server); GUCT (Gu C-terminal domain); HA2 (Helicase associated 2); HAS (helicase/SANT-associated); HD-ZIP (Homeodomain-leucine zipper); HMM (Hidden Markov Model); HRDC (helicase and RNaseD C-terminal); HTH_40 (helix-turn-helix); iTOL (Interactive Tree of Life); JcDHX (*Jatropha curcas* DEAD-box); MDA (Malondialdehyde); MEME (Multiple Em for Motif Elicitation); MYB (myeloblastosis); NCBI (National Center for Biotechnology Information); NTP (nucleoside triphosphate); P.T.M. (transport of proteins across the thylakoid membrane); PAZ (Piwi Argonaut and Zwillig); PFAM (protein family); POLAc (DNA polymerase A); PIP (plasma membrane intrinsic proteins); PPI (protein-protein interaction); PtdIns(3,5) (phosphatidyl 3,5-bisphosphate); RIBOc (ribonuclease); ROS (reactive oxygen species); RQC (RecQ C-terminal); SANT (SWI3, ADA2, N-CoR and TFIIB); SF2 (Superfamily 2); SGs (stress granules); SMART (Simple Modular Architecture Research Tool); SNF2 (Sucrose non-fermenting); SOD (superoxide dismutase); TAIR (The Arabidopsis Information Resource); TF (transcription factor); UR (upregulated); WD40 (WD [Trp-Asp] or beta-transducin repeats).

References

1. Achten, W.M.J.; Verchot, L.; Franken, Y.J.; Mathijs, E.; Singh, V.P.; Aerts, R.; Muys, B. *Jatropha* Bio-Diesel Production and Use. *Biomass Bioenergy* **2008**, *32*, 1063–1084. <https://doi.org/10.1016/j.biombioe.2008.03.003>.
2. Reubens, B.; Achten, W.M.J.; Maes, W.H.; Danjon, F.; Aerts, R.; Poesen, J.; Muys, B. More than Biofuel? *Jatropha curcas* Root System Symmetry and Potential for Soil Erosion Control. *J. Arid Environ.* **2011**, *75*, 201–205. <https://doi.org/10.1016/j.jaridenv.2010.09.011>.
3. Mensah, M.K.; Drebenstedt, C.; Ola, I.M.; Hoth, N.; Dampney, F.G.; Wiafe, E.D. Immobilization Effects of Co-Pyrolyzed Neem Seed Mixed with Poultry Manure on Potentially Toxic Elements in Soil and the Phytoremediation Potentials of Native Manihot Esculenta and *Jatropha curcas* in Ensuring Sustainable Land Use. *Environ. Monit. Assess.* **2023**, *195*, 793. <https://doi.org/10.1007/s10661-023-11430-3>.
4. Mushtaq, Z.; Faizan, S.; Gulzar, B. Salt Stress, Its Impacts on Plants and the Strategies Plants Are Employing against It: A Review. *J. Appl. Biol. Biotechnol.* **2020**, *8*, 81–91. <https://doi.org/10.7324/JABB.2020.80315>.
5. De La Cruz, J.; Kressler, D.; Linder, P.; Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem. Sci.* **1999**, *24*, 192–198. [https://doi.org/10.1016/s0968-0004\(99\)01376-6](https://doi.org/10.1016/s0968-0004(99)01376-6).
6. Cordin, O.; Banroques, J.; Tanner, N.K.; Linder, P. The DEAD-Box Protein Family of RNA Helicases. *Gene* **2006**, *367*, 17–37. <https://doi.org/10.1016/j.gene.2005.10.019>.
7. Umate, P.; Tuteja, R.; Tuteja, N. Genome-Wide Analysis of Helicase Gene Family from Rice and Arabidopsis: A Comparison with Yeast and Human. *Plant Mol. Biol.* **2010**, *73*, 449–465. <https://doi.org/10.1007/s11103-010-9632-5>.
8. Xu, R.; Zhang, S.; Lu, L.; Cao, H.; Zheng, C. A Genome-Wide Analysis of the RNA Helicase Gene Family in *Solanum Lycopersicum*. *Gene* **2013**, *513*, 128–140. <https://doi.org/10.1016/j.gene.2012.10.053>.
9. Chen, J.; Zhang, Y.; Liu, J.; Xia, M.; Wang, W.; Shen, F. Genome-Wide Analysis of the RNA Helicase Gene Family in *Gossypium Raimondii*. *Int. J. Mol. Sci.* **2014**, *15*, 4635–4656. <https://doi.org/10.3390/ijms15034635>.
10. Pandey, S.; Muthamilarasan, M.; Sharma, N.; Chaudhry, V.; Dulani, P.; Shweta, S.; Jha, S.; Mathur, S.; Prasad, M. Characterization of DEAD-Box Family of RNA Helicases in Tomato Provides Insights into Their Roles in

- Biotic and Abiotic Stresses. *Environ. Exp. Bot.* **2019**, *158*, 107–116. <https://doi.org/10.1016/j.envexpbot.2018.11.018>.
11. Yang, S.-d.; Guo, D.-l.; Pei, M.-s.; Wei, T.-l.; Liu, H.-n.; Bian, L.; Yu, K.-k.; Zhang, G.-h.; Yu, Y.-h. Identification of the DEAD-Box RNA Helicase Family Members in Grapevine Reveals That VviDEADRH25a Confers Tolerance to Drought Stress. *J. Integr. Agric.* **2022**, *21*, 1357–1374. [https://doi.org/10.1016/S2095-3119\(21\)63870-4](https://doi.org/10.1016/S2095-3119(21)63870-4).
 12. Ru, J.-N.; Hou, Z.-H.; Zheng, L.; Zhao, Q.; Wang, F.-Z.; Chen, J.; Zhou, Y.-B.; Chen, M.; Ma, Y.-Z.; Xi, Y.-J.; et al. Genome-Wide Analysis of DEAD-Box RNA Helicase Family in Wheat (*Triticum aestivum*) and Functional Identification of TaDEAD-Box57 in Abiotic Stress Responses. *Front. Plant Sci.* **2021**, *12*, 797276. <https://doi.org/10.3389/fpls.2021.797276>.
 13. Nawaz, G.; Lee, K.; Park, S.J.; Kim, Y.O.; Kang, H. A Chloroplast-Targeted Cabbage DEAD-Box RNA Helicase BrRH22 Confers Abiotic Stress Tolerance to Transgenic Arabidopsis Plants by Affecting Translation of Chloroplast Transcripts. *Plant Physiol. Biochem.* **2018**, *127*, 336–342. <https://doi.org/10.1016/j.plaphy.2018.04.007>.
 14. Zhu, M.; Chen, G.; Dong, T.; Wang, L.; Zhang, J.; Zhao, Z.; Hu, Z. SIDEAD31, a Putative DEAD-Box RNA Helicase Gene, Regulates Salt and Drought Tolerance and Stress-Related Genes in Tomato. *PLoS ONE* **2015**, *10*, e0133849. <https://doi.org/10.1371/journal.pone.0133849>.
 15. Chen, J.; Wan, S.; Liu, H.; Fan, S.; Zhang, Y.; Wang, W.; Xia, M.; Yuan, R.; Deng, F.; Shen, F. Over Expression of an *Apocynum venetum* DEAD-Box Heli Case Gene (*AvDH1*) in Cotton Confers Salinity Tolerance and Increases Yield in a Saline Field. *Front. Plant Sci.* **2016**, *6*, 1227. <https://doi.org/10.3389/fpls.2015.01227>.
 16. Nawaz, G.; Kang, H. Rice OsRH58, a Chloroplast DEAD-Box RNA Helicase, Improves Salt or Drought Stress Tolerance in *Arabidopsis* by Affecting Chloroplast Translation. *BMC Plant Biol.* **2019**, *19*, 17. <https://doi.org/10.1186/s12870-018-1623-8>.
 17. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B.L.; Salazar, G.A.; Bileschi, M.L.; Bork, P.; Bridge, A.; Colwell, L.; et al. InterPro in 2022. *Nucleic Acids Res.* **2023**, *51*, D418–D427. <https://doi.org/10.1093/nar/gkac993>.
 18. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; et al. CDD: NCBI's Conserved Domain Database. *Nucleic Acids Res.* **2015**, *43*, D222–D226. <https://doi.org/10.1093/nar/gku1221>.
 19. Letunic, I.; Khedkar, S.; Bork, P. SMART: Recent Updates, New Developments and Status in 2020. *Nucleic Acids Res.* **2021**, *49*, D458–D460. <https://doi.org/10.1093/nar/gkaa937>.
 20. Bailey, T.; Elkan, C. *Fitting a Mixture Model by Expectation Maximization*; AAAI Press: Cambridge, MA, USA, 1994; pp. 3–9.
 21. Gupta, S.; Stamatiyannopoulos, J.A.; Bailey, T.L.; Noble, W.S. Quantifying Similarity between Motifs. *Genome Biol.* **2007**, *8*, R24. <https://doi.org/10.1186/gb-2007-8-2-r24>.
 22. Hu, B.; Jin, J.; Guo, A.Y.; Zhang, H.; Luo, J.; Gao, G. GSDB 2.0: An Upgraded Gene Feature Visualization Server. *Bioinformatics* **2015**, *31*, 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>.
 23. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, USA, 2005; pp. 571–608.
 24. Yu, C.-S.; Chen, Y.-C.; Lu, C.-H.; Hwang, J.-K. Prediction of Protein Subcellular Localization. *Proteins* **2006**, *64*, 643–651. <https://doi.org/10.1002/prot.21018>.
 25. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; Mcgettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X Version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>.
 26. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. <https://doi.org/10.1093/nar/gkab301>.
 27. Tatusov, R.L.; Koonin, E.V.; Lipman, D.J. A Genomic Perspective on Protein Families. *Science* **1997**, *278*, 631–637. <https://doi.org/10.1126/science.278.5338.631>.
 28. Geourjon, C.; Deleage, G. SOPMA: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments. *Bioinformatics* **1995**, *11*, 681–684.
 29. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25*, 1189–1191.
 30. Alva, V.; Nam, S.Z.; Soding, J.; Lupas, A.N. The MPI Bioinformatics Toolkit as an Integrative Platform for Advanced Protein Sequence and Structure Analysis. *Nucleic Acids Res.* **2016**, *44*, W410–W415. <https://doi.org/10.1093/nar/gkw348>.
 31. Lotun, D.P.; Cochard, C.; Vieira, F.R.J.; Bernardes, J.S. 2dSS: A Web Server for Protein Secondary Structure Visualization. *bioRxiv* **2019**, 649426. <https://doi.org/10.1101/649426>

32. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. <https://doi.org/10.1093/nar/gky427>.
33. DeLano, W.L. Pymol: An Open-Source Molecular Graphics Tool. *CCP4 Newsl. Protein Crystallogr.* **2002**, *40*, 82–92.
34. Lozano-Isla, F.; Campos, M.L.O.; Endres, L.; Bezerra-Neto, E.; Pompelli, M.F. Effects of Seed Storage Time and Salt Stress on the Germination of *Jatropha curcas* L. *Ind. Crops Prod.* **2018**, *118*, 214–224. <https://doi.org/10.1016/j.indcrop.2018.03.052>.
35. Souza, M.C.P.; da Silva, M.D.; Binneck, E.; de Lima Cabral, G.A.; Iseppon, A.M.B.; Pompelli, M.F.; Endres, L.; Kido, E.A. RNA-Seq Transcriptome Analysis of *Jatropha curcas* L. Accessions after Salt Stimulus and Unigene-Derived Microsatellite Mining. *Ind. Crops Prod.* **2020**, *147*, 112168.
36. Haas, B.; Papanicolaou, A.; Yassour, M. TransDecoder. 2017. Available online: <https://github.com/TransDecoder/TransDecoder> (accessed on 17 August 2021).
37. Camargo, A.P.; Sourkov, V.; Pereira, G.A.G.; Carazzolle, M.F. RNAsamba: Neural Network-Based Assessment of the Protein-Coding Potential of RNA Sequences. *NAR Genom. Bioinform.* **2020**, *2*, lqz024.
38. Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bairoch, A. UniProtKB/Swiss-Prot: The Manually Annotated Section of the UniProt KnowledgeBase. In *Plant Bioinformatics: Methods and Protocols*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 89–112.
39. Wheeler, D.L.; Church, D.M.; Federhen, S.; Lash, A.E.; Madden, T.L.; Pontius, J.U.; Schuler, G.D.; Schriml, L.M.; Sequeira, E.; Tatusova, T.A. Database Resources of the National Center for Biotechnology. *Nucleic Acids Res.* **2003**, *31*, 28–33.
40. Saldanha, A.J. Java Treeview—Extensible Visualization of Microarray Data. *Bioinformatics* **2004**, *20*, 3246–3248.
41. Rozen, S.; Skaletsky, H. Primer3 on the WWW for General Users and for Biologist Programmers. *Methods Mol. Biol.* **2000**, *132*, 365–386. <https://doi.org/10.1385/1-59259-192-2:365>.
42. Rasmussen, R. Quantification on the LightCycler. In *Rapid Cycle Real-Time PCR: Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 21–34.
43. Pfaffl, M.W.; Horgan, G.W.; Dempfle, L. Relative Expression Software Tool (REST) for Group-Wise Comparison and Statistical Analysis of Relative Expression Results in Real-Time PCR. *Nucleic Acids Res.* **2002**, *30*, e36. <https://doi.org/10.1093/nar/30.9.e36>.
44. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S. The STRING Database in 2023: Protein–Protein Association Networks and Functional Enrichment Analyses for Any Sequenced Genome of Interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646.
45. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504.
46. Linder, P.; Lasko, P.F.; Ashburner, M.; Leroy, P.; Nielsen, P.J.; Nishi, K.; Schnier, J.; Slonimski, P.P. Birth of the DEAD Box. *Nature* **1989**, *337*, 121–122.
47. Aubourg, S.; Kreis, M.; Lecharny, A. The DEAD Box RNA Helicase Family in Arabidopsis Thaliana. *Nucleic Acids Res.* **1999**, *27*, 628–636. <https://doi.org/10.1093/nar/27.2.628>.
48. Xu, R.; Zhang, S.; Huang, J.; Zheng, C. Genome-Wide Comparative in Silico Analysis of the RNA Helicase Gene Family in Zea Mays and Glycine Max: A Comparison with Arabidopsis and Oryza Sativa. *PLoS ONE* **2013**, *8*, e78982.
49. Jo, B.-S.; Choi, S.S. Introns: The Functional Benefits of Introns in Genomes. *Genom. Inform.* **2015**, *13*, 112.
50. Kammel, C.; Thomaier, M.; Sørensen, B.B.; Schubert, T.; Längst, G.; Grasser, M.; Grasser, K.D. Arabidopsis DEAD-Box RNA Helicase UAP56 Interacts with Both RNA and DNA as Well as with MRNA Export Factors. *PLoS ONE* **2013**, *8*, e0060644. <https://doi.org/10.1371/journal.pone.0060644>.
51. Huang, C.K.; Shen, Y.L.; Huang, L.F.; Wu, S.J.; Yeh, C.H.; Lu, C.A. The DEAD-Box RNA Helicase AtRH7/PRH75 Participates in Pre-RRNA Processing, Plant Development and Cold Tolerance in Arabidopsis. *Plant Cell Physiol.* **2016**, *57*, 174–191. <https://doi.org/10.1093/pcp/pcv188>.
52. Lange, H.; Zuber, H.; Sement, F.M.; Chicher, J.; Kuhn, L.; Hammann, P.; Brunaud, V.; Bérard, C.; Bouteiller, N.; Balzergue, S.; et al. The RNA Helicases AtMTR4 and HEN2 Target Specific Subsets of Nuclear Transcripts for Degradation by the Nuclear Exosome in Arabidopsis Thaliana. *PLoS Genet.* **2014**, *10*, e1004564. <https://doi.org/10.1371/journal.pgen.1004564>.
53. Nawaz, G.; Kang, H. Chloroplast- or Mitochondria-Targeted DEAD-Box RNA Helicases Play Essential Roles in Organellar RNA Metabolism and Abiotic Stress Responses. *Front. Plant Sci.* **2017**, *8*, 871. <https://doi.org/10.3389/fpls.2017.00871>.
54. Jermy, A.J.; Willer, M.; Davis, E.; Wilkinson, B.M.; Stirling, C.J. The Brl Domain in Sec63p Is Required for Assembly of Functional Endoplasmic Reticulum Translocons. *J. Biol. Chem.* **2006**, *281*, 7899–7906. <https://doi.org/10.1074/jbc.M511402200>.

55. Cordin, O.; Tanner, N.K.; Doere, M.; Linder, P.; Banroques, J. The Newly Discovered Q Motif of DEAD-box RNA Helicases Regulates RNA-binding and Helicase Activity. *EMBO J.* **2004**, *23*, 2478–2487.
56. Rogers, G.W., Jr.; Komar, A.A.; Merrick, W.C. EIF4A: The Godfather of the DEAD Box Helicases. *Prog. Nucleic Acid Res. Mol. Biol.* **2002**, *72*, 307–331.
57. Story, R.M.; Li, H.; Abelson, J.N. Crystal Structure of a DEAD Box Protein from the Hyperthermophile *Methanococcus Jannaschii*. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1465–1470.
58. Tanner, N.K.; Linder, P. DEXD/H Box RNA Helicases: From Generic Motors to Specific Dissociation Functions. *Mol. Cell* **2001**, *8*, 251–262.
59. Caruthers, J.M.; McKay, D.B. Helicase Structure and Mechanism. *Curr. Opin. Struct. Biol.* **2002**, *12*, 123–133.
60. Fairman-Williams, M.E.; Guenther, U.-P.; Jankowsky, E. SF1 and SF2 Helicases: Family Matters. *Curr. Opin. Struct. Biol.* **2010**, *20*, 313–324.
61. Xie, Z.; Nolan, T.M.; Jiang, H.; Yin, Y. AP2/ERF Transcription Factor Regulatory Networks in Hormone and Abiotic Stress Responses in Arabidopsis. *Front. Plant Sci.* **2019**, *10*, 228. <https://doi.org/10.3389/fpls.2019.00228>.
62. Li, C.; Yan, C.; Sun, Q.; Wang, J.; Yuan, C.; Mou, Y.; Shan, S.; Zhao, X. The BHLH Transcription Factor *AhbHLH112* Improves the Drought Tolerance of Peanut. *BMC Plant Biol.* **2021**, *21*, 540. <https://doi.org/10.1186/s12870-021-03318-6>.
63. Li, Y.; Yang, Z.; Zhang, Y.; Guo, J.; Liu, L.; Wang, C.; Wang, B.; Han, G. The Roles of HD-ZIP Proteins in Plant Abiotic Stress Tolerance. *Front. Plant Sci.* **2022**, *13*, 1027071. <https://doi.org/10.3389/fpls.2022.1027071>.
64. Guo, T.; Wang, S.; Zhang, T.; Xu, L.; Li, Y.; Chao, Y.; Han, L. Expression of the Medicago Truncatula MtDof32 Transcription Factor Regulates Plant Growth and Enhances Abiotic Stress Tolerances in Transgenic Arabidopsis. *Environ. Exp. Bot.* **2021**, *183*, 104339. <https://doi.org/10.1016/j.envexpbot.2020.104339>.
65. Su, Y.; Liang, W.; Liu, Z.; Wang, Y.; Zhao, Y.; Ijaz, B.; Hua, J. Overexpression of GhDof1 Improved Salt and Cold Tolerance and Seed Oil Content in *Gossypium Hirsutum*. *J. Plant Physiol.* **2017**, *218*, 222–234. <https://doi.org/10.1016/j.jplph.2017.07.017>.
66. Li, Y.; Tian, M.; Feng, Z.; Zhang, J.; Lu, J.; Fu, X.; Ma, L.; Wei, H.; Wang, H. GhDof1. 7, a Dof Transcription Factor, Plays Positive Regulatory Role under Salinity Stress in Upland Cotton. *Plants* **2023**, *12*, 3740.
67. Bakshi, R.; Prakash, T.; Dash, D.; Brahmachari, V. In Silico Characterization of the INO80 Subfamily of SWI2/SNF2 Chromatin Remodeling Proteins. *Biochem. Biophys. Res. Commun.* **2004**, *320*, 197–204.
68. Sun, J.; Sun, Y.; Ahmed, R.I.; Ren, A.; Xie, M. Research Progress on Plant RING-Finger Proteins. *Genes* **2019**, *10*, 973.
69. Jain, K.; Fraser, C.S.; Marunde, M.R.; Parker, M.M.; Sagum, C.; Burg, J.M.; Hall, N.; Popova, I.K.; Rodriguez, K.L.; Vaidya, A. Characterization of the Plant Homeodomain (PHD) Reader Family for Their Histone Tail Interactions. *Epigenet. Chromatin* **2020**, *13*, 3.
70. Iyer, L.M.; Babu, M.; Aravind, L. The HIRAN Domain and Recruitment of Chromatin Remodeling and Repair Activities to Damaged DNA. *Cell Cycle* **2006**, *5*, 775–782.
71. Du, J.; Zhong, X.; Bernatavichute, Y.V.; Stroud, H.; Feng, S.; Caro, E.; Vashisht, A.A.; Terragni, J.; Chin, H.G.; Tu, A. Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. *Cell* **2012**, *151*, 167–180.
72. Zhan, X.; Qian, B.; Cao, F.; Wu, W.; Yang, L.; Guan, Q.; Gu, X.; Wang, P.; Okusolubo, T.A.; Dunn, S.L. An Arabidopsis PWI and RRM Motif-Containing Protein Is Critical for Pre-mRNA Splicing and ABA Responses. *Nat. Commun.* **2015**, *6*, 8139.
73. He, B.; Cai, Q.; Qiao, L.; Huang, C.-Y.; Wang, S.; Miao, W.; Ha, T.; Wang, Y.; Jin, H. RNA-Binding Proteins Contribute to Small RNA Loading in Plant Extracellular Vesicles. *Nat. Plants* **2021**, *7*, 342–352.
74. Fukudome, A.; Fukuhara, T. Plant Dicer-like Proteins: Double-Stranded RNA-Cleaving Enzymes for Small RNA Biogenesis. *J. Plant Res.* **2017**, *130*, 33–44.
75. Cai, J.; Meng, X.; Li, G.; Dong, T.; Sun, J.; Xu, T.; Li, Z.; Han, Y.; Zhu, M. Identification, Expression Analysis, and Function Evaluation of 42 Tomato DEAD-Box RNA Helicase Genes in Growth Development and Stress Response. *Acta Physiol. Plant.* **2018**, *40*, 94. <https://doi.org/10.1007/s11738-018-2665-0>.
76. Kim, J.S.; Kim, K.A.; Oh, T.R.; Park, C.M.; Kang, H. Functional Characterization of DEAD-Box RNA Helicases in Arabidopsis Thaliana under Abiotic Stress Conditions. *Plant Cell Physiol.* **2008**, *49*, 1563–1571.
77. Nguyen, L.V.; Seok, H.Y.; Woo, D.H.; Lee, S.Y.; Moon, Y.H. Overexpression of the DEAD-Box RNA Helicase Gene AtRH17 Confers Tolerance to Salt Stress in Arabidopsis. *Int. J. Mol. Sci.* **2018**, *19*, 3777. <https://doi.org/10.3390/ijms19123777>.
78. Banu, M.S.A.; Huda, K.M.K.; Harun-Ur-Rashid, M.; Parveen, S.; Tuteja, N. A DEAD Box Helicase Psp68 Positively Regulates Salt Stress Responses in Marker-Free Transgenic Rice Plants. *Transgenic Res.* **2023**, *32*, 293–304. <https://doi.org/10.1007/s11248-023-00353-x>.
79. Cao, G.; Gu, H.; Jiang, W.; Tian, Z.; Shi, G.; Chen, W.; Tian, B.; Wei, X.; Zhang, L.; Wei, F.; et al. BrDHC1, a Novel Putative DEAD-Box Helicase Gene, Confers Drought Tolerance in Transgenic Brassica Rapa. *Horticulturae* **2022**, *8*, 707. <https://doi.org/10.3390/horticulturae8080707>.

80. Kant, P.; Kant, S.; Gordon, M.; Shaked, R.; Barak, S. STRESS RESPONSE SUPPRESSOR1 and STRESS RESPONSE SUPPRESSOR2, Two DEAD-Box RNA Helicases That Attenuate Arabidopsis Responses to Multiple Abiotic Stresses. *Plant Physiol.* **2007**, *145*, 814–830.
81. de Lima Cabral, G.A.; Binneck, E.; de Souza, M.C.P.; da Silva, M.D.; Costa Ferreira Neto, J.R.; Pompelli, M.F.; Endres, L.; Kido, É.A. First Expressed TFome of Physic Nut (*Jatropha curcas* L.) After Salt Stimulus. *Plant Mol. Biol. Report.* **2020**, *38*, 189–208. <https://doi.org/10.1007/s11105-019-01187-w>.
82. Silva-Santos, L.; Corte-Real, N.; Dias-Pereira, J.; Figueiredo, R.C.B.Q.; Endres, L.; Pompelli, M.F. Salinity Shock in *Jatropha curcas* Leaves Is More Pronounced during Recovery than during Stress Time. *bioRxiv* **2018**, 378208. <https://doi.org/10.1101/378208>.
83. Ding, F.; Cui, P.; Wang, Z.; Zhang, S.; Ali, S.; Xiong, L. Genome-Wide Analysis of Alternative Splicing of Pre-mRNA under Salt Stress in Arabidopsis. *BMC Genom.* **2014**, *15*, 431. <https://doi.org/10.1186/1471-2164-15-431>.
84. Zhu, G.; Li, W.; Zhang, F.; Guo, W. RNA-Seq Analysis Reveals Alternative Splicing under Salt Stress in Cotton, *Gossypium davidsonii*. *BMC Genom.* **2018**, *19*, 73. <https://doi.org/10.1186/s12864-018-4449-8>.
85. Wang, B.; Chai, H.; Zhong, Y.; Shen, Y.; Yang, W.; Chen, J.; Xin, Z.; Shi, H. The DEAD-Box RNA Helicase SHI2 Functions in Repression of Salt-Inducible Genes and Regulation of Cold-Inducible Gene Splicing. *J. Exp. Bot.* **2020**, *71*, 1598–1613. <https://doi.org/10.1093/jxb/erz523>.
86. Ji, H.; Yang, G.; Zhang, X.; Zhong, Q.; Qi, Y.; Wu, K.; Shen, T. Regulation of Salt Tolerance in the Roots of Zea Mays by L-Histidine through Transcriptome Analysis. *Front. Plant Sci.* **2022**, *13*, 1049954. <https://doi.org/10.3389/fpls.2022.1049954>.
87. Lin, H.H.; Lin, K.H.; Syu, J.Y.; Tang, S.Y.; Lo, H.F. Physiological and Proteomic Analysis in Two Wild Tomato Lines under Waterlogging and High Temperature Stress. *J. Plant Biochem. Biotechnol.* **2016**, *25*, 87–96. <https://doi.org/10.1007/s13562-015-0314-x>.
88. Prajapati, P.; Gupta, P.; Kharwar, R.N.; Seth, C.S. Nitric Oxide Mediated Regulation of Ascorbate-Glutathione Pathway Alleviates Mitotic Aberrations and DNA Damage in *Allium cepa* L. under Salinity Stress. *Int. J. Phytoremediat.* **2023**, *25*, 403–414. <https://doi.org/10.1080/15226514.2022.2086215>.
89. Qi, F.; Zhang, F. Cell Cycle Regulation in the Plant Response to Stress. *Front. Plant Sci.* **2020**, *10*, 1765. <https://doi.org/10.3389/fpls.2019.01765>.
90. Chirinos-Arias, M.C.; Spampinato, C.P. Role of the Mismatch Repair Protein MSH7 in Arabidopsis Adaptation to Acute Salt Stress. *Plant Physiol. Biochem.* **2021**, *169*, 280–290. <https://doi.org/10.1016/j.plaphy.2021.11.029>.
91. Branzei, D.; Foiani, M. Regulation of DNA Repair throughout the Cell Cycle. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 297–308. <https://doi.org/10.1038/nrm2351>.
92. Chantarachot, T.; Bailey-Serres, J. Polysomes, Stress Granules, and Processing Bodies: A Dynamic Triumvirate Controlling Cytoplasmic mRNA Fate and Function. *Plant Physiol.* **2018**, *176*, 254–269. <https://doi.org/10.1104/pp.17.01468>.
93. Liu, Y.; Liu, S.; Shi, H.; Ma, J.; Jing, M.; Han, Y. The TSN1 Binding Protein RH31 Is a Component of Stress Granules and Participates in Regulation of Salt-Stress Tolerance in Arabidopsis. *Front. Plant Sci.* **2021**, *12*, 804356. <https://doi.org/10.3389/fpls.2021.804356>.
94. Kearly, A.; Nelson, A.D.L.; Skirycz, A.; Chodasiewicz, M. Composition and Function of Stress Granules and P-Bodies in Plants. *Semin. Cell Dev. Biol.* **2024**, *156*, 167–175. <https://doi.org/10.1016/j.semcdb.2022.11.008>.
95. Liu, Y.; Imai, R. Function of Plant DEXD/H-Box RNA Helicases Associated with Ribosomal RNA Biogenesis. *Front. Plant Sci.* **2018**, *9*, 125. <https://doi.org/10.3389/fpls.2018.00125>.
96. Panse, V.G.; Johnson, A.W. Maturation of Eukaryotic Ribosomes: Acquisition of Functionality. *Trends Biochem. Sci.* **2010**, *35*, 260–266.
97. Kalinina, N.O.; Makarova, S.; Makhotenko, A.; Love, A.J.; Taliansky, M. The Multiple Functions of the Nucleolus in Plant Development, Disease and Stress Responses. *Front. Plant Sci.* **2018**, *9*, 132. <https://doi.org/10.3389/fpls.2018.00132>.
98. Chen, X.; Wang, Y.; Li, J.; Jiang, A.; Cheng, Y.; Zhang, W. Mitochondrial Proteome during Salt Stress-Induced Programmed Cell Death in Rice. *Plant Physiol. Biochem.* **2009**, *47*, 407–415. <https://doi.org/10.1016/j.plaphy.2008.12.021>.
99. Wang, W.; Xu, M.; Liu, X.; Tu, J. The Rice Eukaryotic Translation Initiation Factor 3 Subunit e (OseiF3e) Influences Organ Size and Pollen Maturation. *Front. Plant Sci.* **2016**, *7*, 1399. <https://doi.org/10.3389/fpls.2016.01399>.
100. Valandro, F.; Menguer, P.K.; Cabreira-Cagliari, C.; Margis-Pinheiro, M.; Cagliari, A. Programmed Cell Death (PCD) Control in Plants: New Insights from the Arabidopsis Thaliana Deathosome. *Plant Sci.* **2020**, *299*, 110603. <https://doi.org/10.1016/j.plantsci.2020.110603>.
101. Wang, J.; Li, X.; Liu, Y.; Zhao, X. Salt Stress Induces Programmed Cell Death in *Thellungiella Halophila* Suspension-Cultured Cells. *J. Plant Physiol.* **2010**, *167*, 1145–1151. <https://doi.org/10.1016/j.jplph.2010.03.008>.
102. Ng, S.; De Clercq, I.; Van Aken, O.; Law, S.R.; Ivanova, A.; Willems, P.; Giraud, E.; Van Breusegem, F.; Whelan, J. Anterograde and Retrograde Regulation of Nuclear Genes Encoding Mitochondrial Proteins during Growth, Development, and Stress. *Mol. Plant* **2014**, *7*, 1075–1093. <https://doi.org/10.1093/mp/ssu037>.

103. Wang, X.; Auwerx, J. Systems Phytohormone Responses to Mitochondrial Proteotoxic Stress. *Mol. Cell* **2017**, *68*, 540–551.e5. <https://doi.org/10.1016/j.molcel.2017.10.006>.
104. Depaepe, T.; Hendrix, S.; van Rensburg, H.C.J.; Van den Ende, W.; Cuypers, A.; Van Der Straeten, D. At the Crossroads of Survival and Death: The Reactive Oxygen Species–Ethylene–Sugar Triad and the Unfolded Protein Response. *Trends Plant Sci.* **2021**, *26*, 338–351.
105. Zou, Y.; Liu, Z.; Bai, J.; Zhou, Y.; Lu, D. Mitochondrial Proteotoxic Stresses Activate Absciscic Acid Signaling in Plants. *Environ. Exp. Bot.* **2023**, *205*, 105134. <https://doi.org/10.1016/j.envexpbot.2022.105134>.
106. Tuteja, N.; Sahoo, R.K.; Garg, B.; Tuteja, R. OsSUV3 Dual Helicase Functions in Salinity Stress Tolerance by Maintaining Photosynthesis and Antioxidant Machinery in Rice (*Oryza sativa* L. Cv. IR64). *Plant J.* **2013**, *76*, 115–127. <https://doi.org/10.1111/tpj.12277>.
107. Gómez-Méndez, M.F.; Vera-Estrella, R.; Amezcua-Romero, J.C.; Rosas-Santiago, P.; Hernández-Domínguez, E.E.; de Luna-Valdez, L.A.; Pantoja, O. *Mesembryanthemum crystallinum* Plasma Membrane Root Aquaporins Are Regulated via Clathrin-Coated Vesicles in Response to Salt Stress. *bioRxiv* **2022**. <https://doi.org/10.1101/2022.03.16.484672>.
108. Larson, E.R.; van Zelm, E.; Roux, C.; Marion-Poll, A.; Blatt, M.R. Clathrin Heavy Chain Subunits Coordinate Endo- and Exocytic Traffic and Affect Stomatal Movement. *Plant Physiol.* **2017**, *175*, 708–720. <https://doi.org/10.1104/pp.17.00970>.
109. Lee, J.; Hanh Nguyen, H.; Park, Y.; Lin, J.; Hwang, I. Spatial Regulation of RBOHD via AtECA4-Mediated Recycling and Clathrin-Mediated Endocytosis Contributes to ROS Accumulation during Salt Stress Response but Not Flg22-Induced Immune Response. *Plant J.* **2022**, *109*, 816–830. <https://doi.org/10.1111/tpj.15593>.
110. Meijer, H.J.G.; Divecha, N.; Van Den Ende, H.; Musgrave, A.; Munnik, T. Hyperosmotic Stress Induces Rapid Synthesis of Phosphatidyl-d-Inositol 3,5-Bisphosphate in Plant Cells. *Planta* **1999**, *208*, 294–298. <https://doi.org/10.1007/s004250050561>.
111. Compton, L.M.; Ikononov, O.C.; Sbrissa, D.; Garg, P.; Shisheva, A. Active Vacuolar H⁺ ATPase and Functional Cycle of Rab5 Are Required for the Vacuolation Defect Triggered by Ptdins(3,5)P₂ Loss under PIKfyve or Vps34 Deficiency. *Am. J. Physiol.—Cell Physiol.* **2016**, *311*, C366–C377. <https://doi.org/10.1152/ajpcell.00104.2016>.
112. Bak, G.; Lee, E.J.; Lee, Y.; Kato, M.; Segami, S.; Sze, H.; Maeshima, M.; Hwang, J.U.; Lee, Y. Rapid Structural Changes and Acidification of Guard Cell Vacuoles during Stomatal Closure Require Phosphatidylinositol 3,5-Bisphosphate. *Plant Cell* **2013**, *25*, 2202–2216. <https://doi.org/10.1105/tpc.113.110411>.
113. Pérez Koldenkova, V.; Hatsugai, N. Vacuolar Convolution: Possible Mechanisms and Role of Phosphatidylinositol 3,5-Bisphosphate. *Funct. Plant Biol.* **2017**, *44*, 751–760. <https://doi.org/10.1071/FP16443>.
114. Li, S.C.; Diakov, T.T.; Xu, T.; Tarsio, M.; Zhu, W.; Couoh-Cardel, S.; Weisman, L.S.; Kane, P.M. The Signaling Lipid PI(3,5)P₂ Stabilizes V1-Vo Sector Interactions and Activates the V-ATPase. *Mol. Biol. Cell* **2014**, *25*, 1251–1262. <https://doi.org/10.1091/mbc.E13-10-0563>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

APÊNDICE B – ARTIGO 1

Tabela suplementar 1. Estruturas secundária preditas nas proteínas JcSnf2 a partir do SOPMA.

Proteína	Subfamília	Grupo	α -helice	β -sheet	Beta turn	Random coil
JcSNF2.20	SMARCAL	Distant	45.53%	13.98%	3.89%	36.60%
JcSNF2.19	SMARCAL	Distant	41.16%	10.76%	4.73%	43.36%
JcSNF2.4	Ris1	Rad5/16	27.35%	11.25%	2.57%	58.82%
JcSNF2.10	Ris1	Rad5/16	34.05%	11.16%	3.28%	51.50%
JcSNF2.11	SHPRH	Rad5/16	36.56%	14.34%	5.52%	43.58%
JcSNF2.17	Rad5/16	Rad5/16	39.61%	15.38%	3.90%	41.10%
JcSNF2.25	Rad5/16	Rad5/16	40.25%	11.97%	3.28%	44.50%
JcSNF2.2	SHPRH	Rad5/16	45.57%	10.64%	3.21%	40.58%
JcSNF2.7	ATRX	Rad54	45.16%	10.07%	5.73%	39.04%
JcSNF2.15	DRD1	Rad54	38.05%	13.59%	4.24%	44.12%
JcSNF2.8	DRD1	Rad54	36.98%	14.88%	3.61%	44.53%
JcSNF2.26	DRD1	Rad54	34.58%	13.55%	5.23%	46.64%
JcSNF2.1	Rad54	Rad54	38.18%	11.34%	5.56%	44.92%
JcSNF2.9	ALC1	Snf2	44.75%	13.96%	4.61%	36.68%
JcSNF2.13	Chd1	Snf2	42.87%	8.97%	3.86%	44.29%
JcSNF2.5	Iswi	Snf2	54.55%	10.33%	3.85%	31.27%
JcSNF2.3	Lsh	Snf2	53.46%	12.37%	4.92%	29.26%
JcSNF2.21	Mi-2	Snf2	46.30%	9.14%	4.64%	39.92%
JcSNF2.12	Mi-2	Snf2	28.19%	8.01%	3.28%	60.52%
JcSNF2.14	Snf2	Snf2	54.10%	8.73%	4.10%	33.07%
JcSNF2.6	Snf2	Snf2	30.95%	10.38%	4.19%	54.47%
JcSNF2.23	ERCC6	SSO1653	45.50%	11.29%	4.90%	38.31%
JcSNF2.18	ERCC6	SSO1653	42.91%	13.11%	4.76%	39.23%
JcSNF2.16	ERCC6	SSO1653	41.65%	8.96%	4.56%	44.82%
JcSNF2.22	Mot1	SSO1653	56.98%	6.62%	3.26%	33.15%
JcSNF2.24	Swr1	Swr1	42.59%	10.69%	3.77%	42.95%

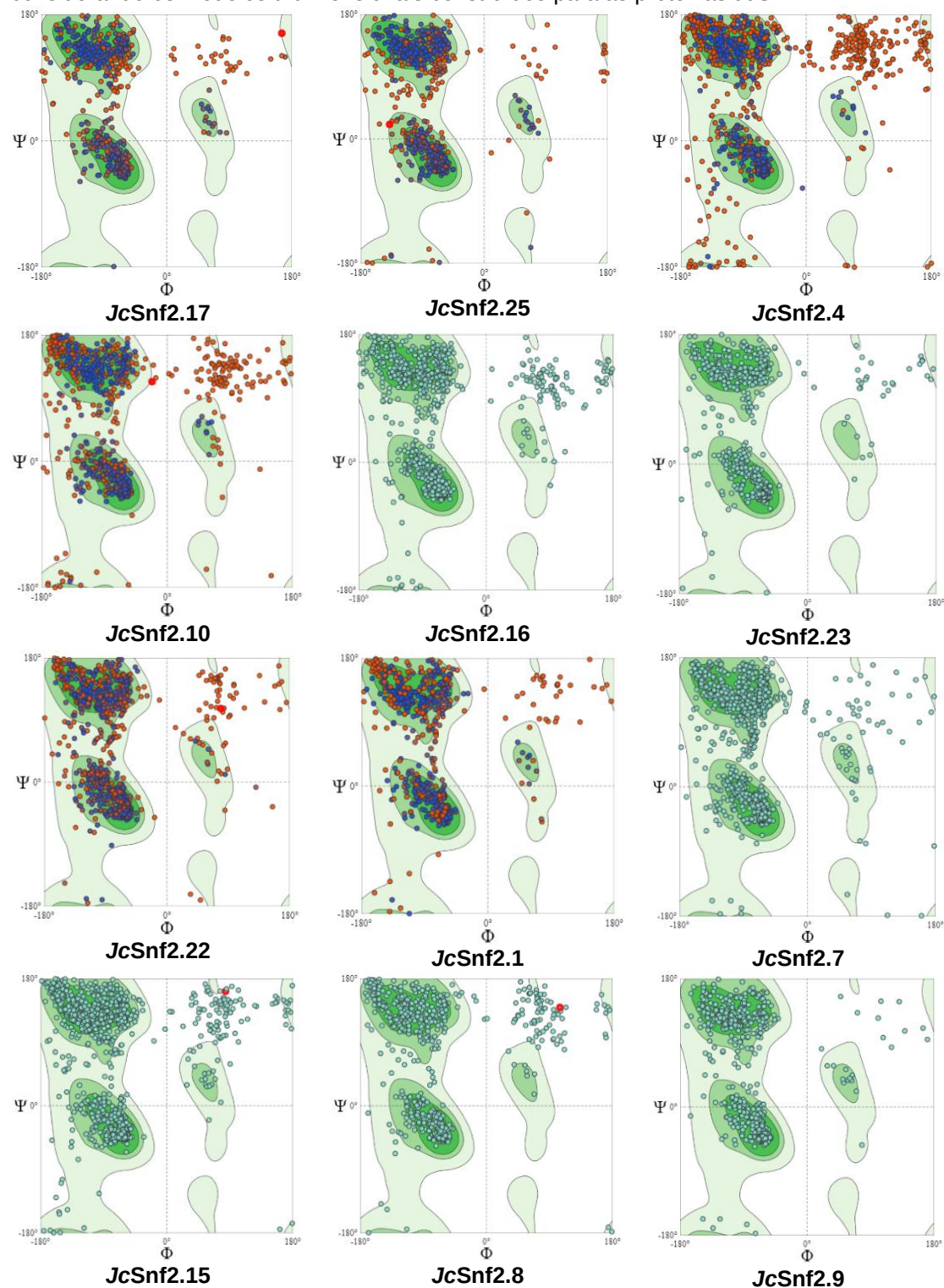
Tabela suplementar 2. Parâmetros de qualidade obtidos a partir do Swissmodel para os modelos construídos para as proteínas *JcSnf2*.

Proteína	Subfamília	GMQE	Identidade	Similaridade	Cobertura	Ramachandran Favoured	QMEANDisCo global
JcSNF2.9	ALC1	0.74	76.77%	0.53	0.99	93.28%	0.60
JcSNF2.7	ATRX	0.61	72.00%	0.52	0.95	88.92%	0.51
JcSNF2.13	Chd1	0.60	79.08%	0.54	0.99	77.43%	0.47
JcSNF2.15	DRD1	0.64	100.00%	0.62	1.00	83.91%	0.47
JcSNF2.8	DRD1	0.65	100.00%	0.61	1.00	83.88%	0.42
JcSNF2.16	ERCC6	0.62	73.31%	0.52	0.99	85.47%	0.53
JcSNF2.23	ERCC6	0.71	74.48%	0.53	0.98	91.30%	0.56
JcSNF2.5	Iswi	0.72	93.88%	0.59	1.00	93.32%	0.59
JcSNF2.3	Lsh	0.69	81.38%	0.55	1.00	91.07%	0.59
JcSNF2.21	Mi-2	0.64	78.78%	0.55	0.96	85.39%	0.52
JcSNF2.22	Mot1	0.69	74.86%	0.52	0.99	92.55%	0.55
JcSNF2.17	Rad5/16	0.69	100.00%	0.61	1.00	89.87%	0.57
JcSNF2.25	Rad5/16	0.70	75.66%	0.53	0.99	89.65%	0.56
JcSNF2.1	Rad54	0.68	80.47%	0.55	0.99	87.34%	0.57
JcSNF2.10	Ris1	0.60	74.50%	0.52	0.98	80.85%	0.45
JcSNF2.11	SHPRH	0.65	74.69%	0.54	0.86	87.74%	0.46
JcSNF2.19	SMARCAL	0.65	70.40%	0.51	0.95	90.66%	0.51
JcSNF2.20	SMARCAL	0.78	76.15%	0.53	0.97	91.16%	0.54
JcSNF2.14	Snf2	0.68	89.10%	0.58	1.00	85.36%	0.56

Tabela suplementar 3. Pares de *primers* relativos a transcritos RNA-Seq *JcSnf2* diferencialmente expressos em resposta a salinidade a partir dos dados *in silico* e de qPCR.

Transcrito RNA-Seq	Produto (pb)	Tm produto	Primer (L)	Tm	%GC	Primer (R)	Tm	%GC
Jc_DN34178_c0_g1_i1	116	70.28	GGTGAAGACGTTGAAGAGGATG	61.05	50.00	TGTGCGTCCAATATTTCTG	59.54	45.00
Jc_DN36373_c0_g1_i1	108	72.57	ACTGGGAATGATGAGAGACTGG	60.51	50.00	TCTCAGAAGAGGCACAGACAAC	59.65	50.00
Jc_DN36416_c1_g1_i1	146	72.64	CTGAGCCCTTCTTTAGGAACAC	59.41	50.00	TATGACTGGAAGGAGTTTGCAC	59.24	45.45
Jc_DN40579_c0_g2_i1	83	70.46	ACCTCCAGCTTTGATAGACAGC	59.91	50.00	CAGCATCAAGGCAAGCATTAG	60.91	47.62
Jc_DN61267_c0_g2_i1	119	71.40	CAAGCTGCAACCATGTTTAGTC	59.81	45.45	ACAATGTCAGTTGTTGCCAGAG	60.22	45.45
Jc_DN36373_c0_g1_i2	80	71.88	TTCTATACCCGGCTAGAAGCTG	59.90	50.00	GCATAATTCTGGTTCACTGTGC	59.64	45.45

Figura suplementar 1. Gráficos de Ramachandran gerados a partir do Swissmodel considerando os modelos tridimensionais construídos para as proteínas JcSnf2.



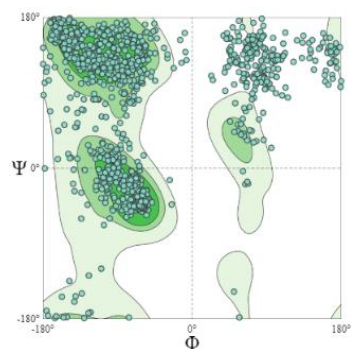
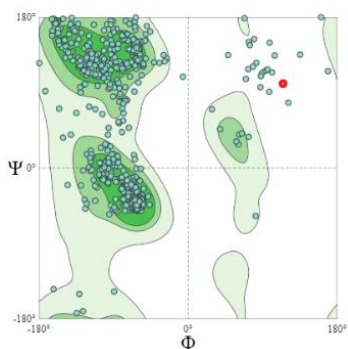
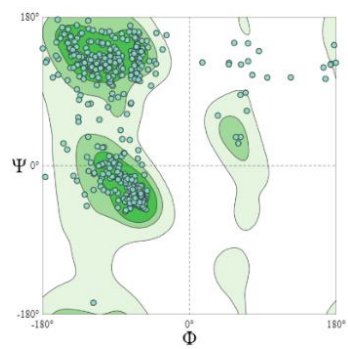
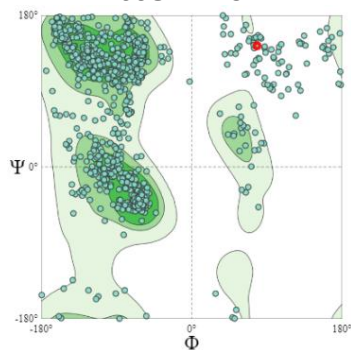
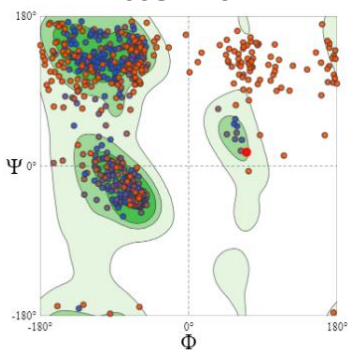
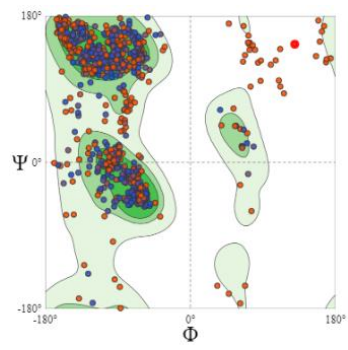
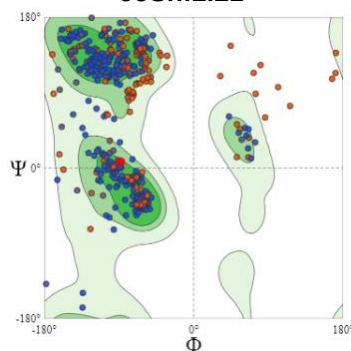
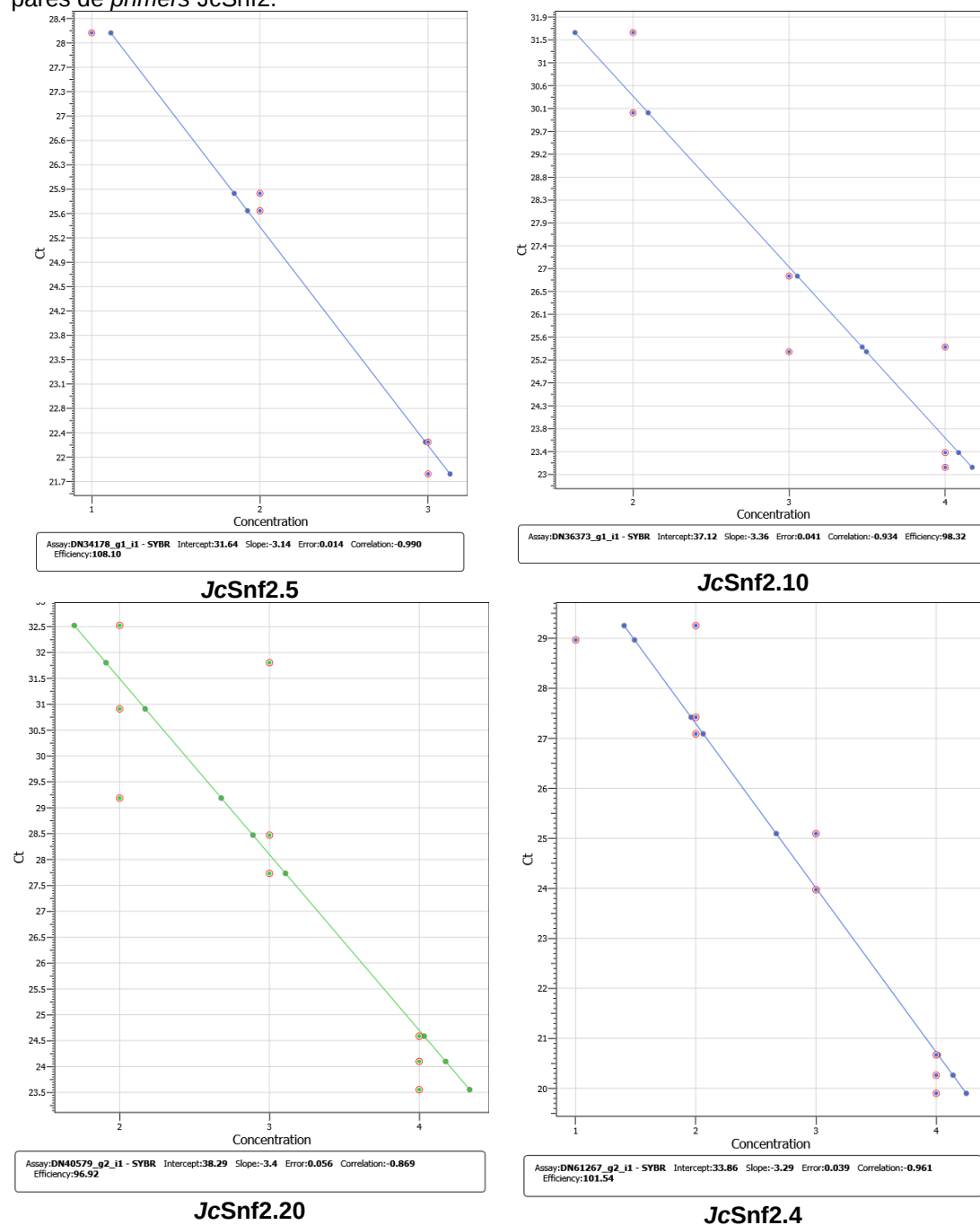
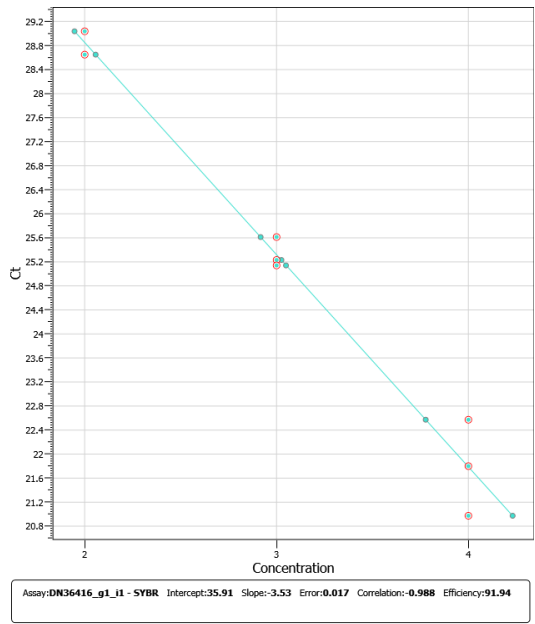
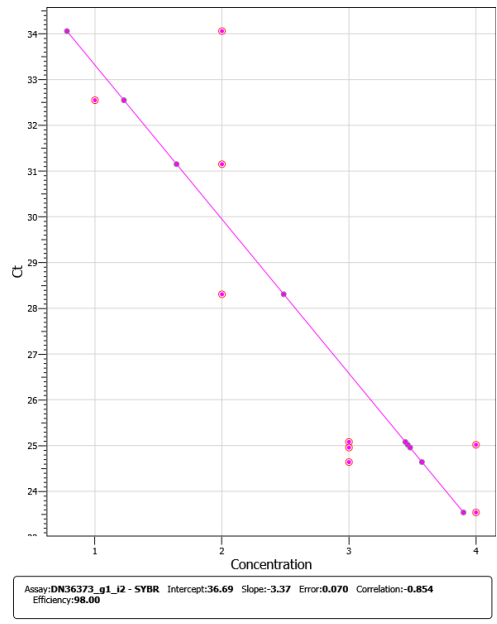
**JcSnf2.13****JcSnf2.5****JcSnf2.3****JcSnf2.21****JcSnf2.14****JcSnf2.19****JcSnf2.20**

Figura suplementar 2. Curvas padrão geradas a partir de diluições seriadas feitas com cDNAs das cultivares (sem diluição, 10-1, 10-2, 10-3 e 10-4) demonstrando a eficiência dos pares de *primers* JcSnf2.



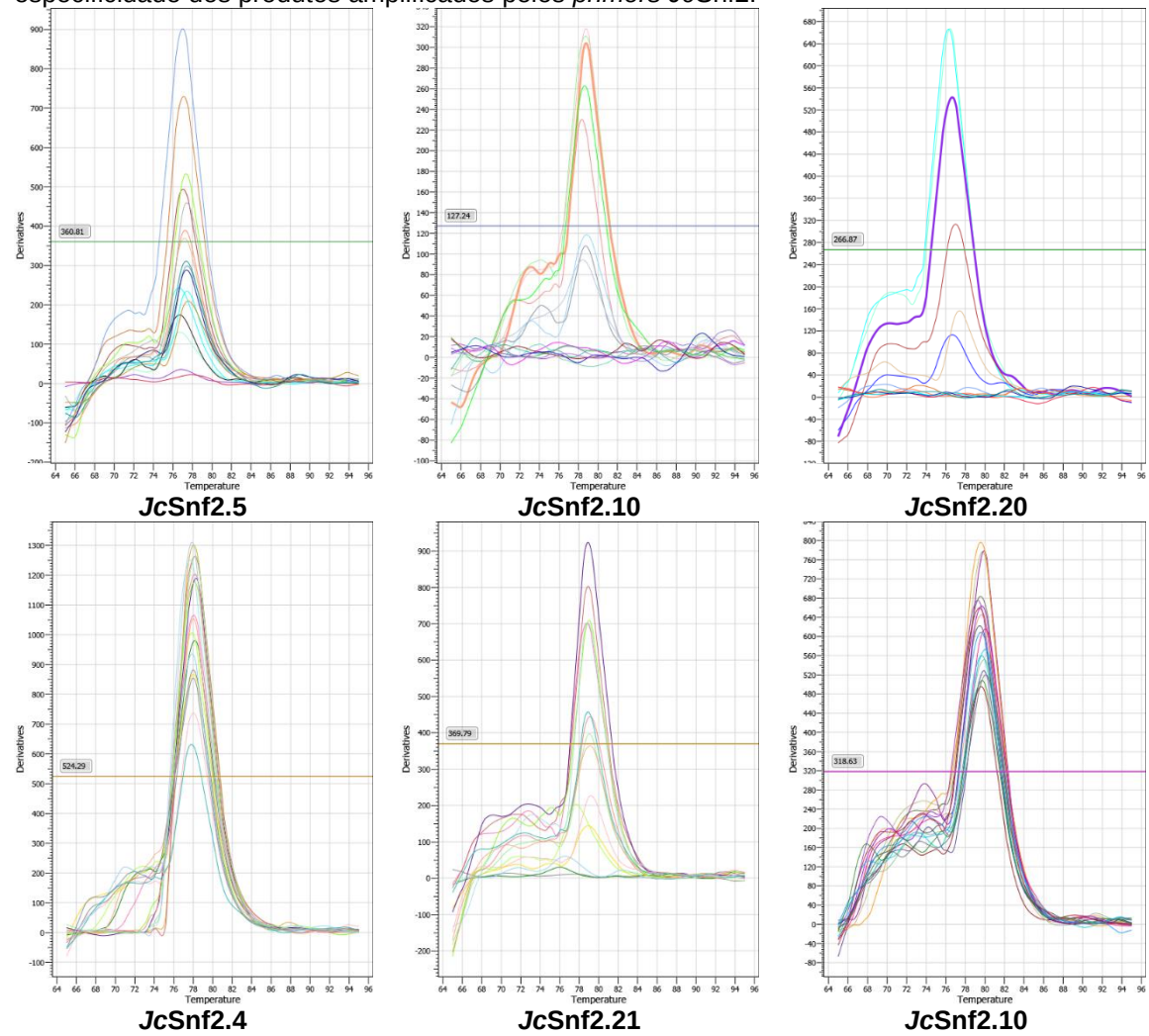


JcSnf2.21



JcSnf2.10

Figura suplementar 3. Curvas de dissociação geradas entre 65-95°C evidenciando a especificidade dos produtos amplificados pelos *primers* *JcSnf2*.



APÊNDICE C – ARTIGO 2

Tabela suplementar 1. Termos GO enriquecidos ($< 1.0e-16$) associados às redes de interação proteína-proteína (PPI) construídas considerando as DNA helicases de pinhão-mansão (*J. curcas*) a partir do bando de dados do STRING

Termo ID	Descrição	<i>observed gene count</i>	<i>background gene count</i>	<i>strength</i>	<i>false discovery rate</i>
GO:0006260	DNA replication	16	147	2.25	1.07e-31
GO:0006261	DNA-templated DNA replication	15	113	2.33	1.75e-30
GO:0006270	DNA replication initiation	11	24	2.87	5.85e-26
GO:0006259	DNA metabolic process	17	614	1.65	1.60e-25
GO:0000727	Double-strand break repair via break-induced replication	7	12	2.97	4.00e-16
GO:0033260	Nuclear DNA replication	7	14	2.91	7.90e-16
GO:0007049	Cell cycle	12	598	1.51	3.45e-14
GO:0000724	Double-strand break repair via homologous recombination	8	92	2.15	2.10e-13
GO:0006268	DNA unwinding involved in DNA replication	6	19	2.71	1.17e-12
GO:0006281	DNA repair	10	387	1.62	1.91e-12
GO:1902969	Mitotic DNA replication	5	10	2.91	3.14e-11
GO:0000278	Mitotic cell cycle	8	250	1.71	2.61e-10
GO:0051276	Chromosome organization	9	545	1.43	2.23e-09
GO:0022402	Cell cycle process	8	411	1.5	1.13e-08
GO:0071163	DNA replication preinitiation complex assembly	3	3	3.21	6.47e-07
GO:0048229	Gametophyte development	7	449	1.4	8.92e-07
GO:1902975	Mitotic DNA replication initiation	3	5	2.99	1.67e-06
GO:0051716	Cellular response to stimulus	12	3214	0.78	3.31e-06
GO:0006271	DNA strand elongation involved in DNA replication	3	16	2.48	2.48e-05
GO:0007275	Multicellular organism development	10	2344	0.84	2.90e-05
GO:0016043	Cellular component organization	10	2788	0.76	0.00013
GO:0030174	Regulation of DNA-templated DNA replication initiation	2	3	3.03	0.00040
GO:0050896	Response to stimulus	13	6206	0.53	0.00041
GO:0090329	Regulation of DNA-templated DNA replication	3	47	2.01	0.00041
GO:0009555	Pollen development	4	343	1.28	0.0049
GO:0000076	DNA replication checkpoint signaling	2	16	2.31	0.0051
GO:0003697	Single-stranded DNA binding	9	87	2.22	4.00e-15
GO:0003677	DNA binding	16	2517	1.01	6.12e-13
GO:0003688	DNA replication origin binding	6	12	2.91	6.12e-13

GO:0003678	DNA helicase activity	7	53	2.33	2.51e-12
GO:0016887	ATP hydrolysis activity	7	401	1.45	9.18e-07
GO:0070182	DNA polymerase binding	3	6	2.91	5.27e-06
GO:0005488	Binding	17	11632	0.37	7.70e-05
GO:0005524	ATP binding	8	2421	0.73	0.0063
GO:0004693	Cyclin-dependent protein serine/threonine kinase activity	2	33	1.99	0.0218
GO:0042555	MCM complex	7	10	3.05	1.62e-16
GO:0140513	Nuclear protein-containing complex	13	947	1.35	1.12e-13
GO:0000347	THO complex	6	14	2.84	2.68e-13
GO:0000228	Nuclear chromosome	7	112	2.0	1.13e-10
GO:0005634	Nucleus	17	6070	0.66	1.47e-09
GO:0005664	Nuclear origin of replication recognition complex	4	7	2.97	5.53e-09
GO:0032991	Protein-containing complex	14	3222	0.85	5.81e-09
GO:0031981	Nuclear lumen	8	1028	1.1	7.18e-06
GO:0031261	DNA replication preinitiation complex	2	7	2.66	0.00072
GO:0043596	Nuclear replication fork	2	22	2.17	0.0045
GO:0006338	Chromatin remodeling	16	134	1.7	1.04e-18
GO:0006325	Chromatin organization	19	347	1.36	3.96e-17
GO:0051276	Chromosome organization	21	545	1.21	1.88e-16
GO:0006996	Organelle organization	30	1749	0.85	1.32e-15
GO:0006281	DNA repair	17	387	1.26	7.31e-14
GO:0016043	Cellular component organization	33	2788	0.69	3.70e-13
GO:0071840	Cellular component organization or biogenesis	34	3168	0.65	1.66e-12
GO:0006259	DNA metabolic process	18	614	1.09	3.89e-12
GO:0016573	Histone acetylation	9	56	1.83	2.46e-11
GO:0015031	Protein transport	18	814	0.96	2.49e-10
GO:0006886	Intracellular protein transport	15	504	1.09	4.16e-10
GO:0065002	Intracellular protein transmembrane transport	9	87	1.63	5.23e-10
GO:0033554	Cellular response to stress	19	1242	0.8	1.76e-08
GO:0043486	Histone exchange	5	10	2.32	4.43e-08
GO:0044260	Cellular macromolecule metabolic process	25	2581	0.61	9.33e-08
GO:0009987	Cellular process	56	13552	0.24	3.51e-07
GO:0043933	Protein-containing complex organization	13	632	0.93	6.31e-07
GO:0006605	Protein targeting	9	230	1.21	9.55e-07
GO:0045037	Protein import into chloroplast stroma	5	26	1.9	1.71e-06

GO:0006261	DNA-templated DNA replication	7	113	1.41	2.39e-06
GO:0072594	Establishment of protein localization to organelle	9	273	1.14	3.59e-06
GO:0016558	Protein import into peroxisome matrix	4	11	2.18	5.81e-06
GO:0043967	Histone H4 acetylation	4	17	1.99	2.10e-05
GO:0035494	SNARE complex disassembly	3	3	2.62	2.86e-05
GO:0019219	Regulation of nucleobase-containing compound metabolic process	21	2622	0.52	5.24e-05
GO:0043170	Macromolecule metabolic process	34	6508	0.34	9.74e-05
GO:0000492	Box C/D snoRNP assembly	3	6	2.32	0.00011
GO:0051171	Regulation of nitrogen compound metabolic process	23	3272	0.47	0.00011
GO:0080090	Regulation of primary metabolic process	23	3346	0.46	0.00016
GO:0006357	Regulation of transcription by RNA polymerase II	10	608	0.84	0.00019
GO:0031323	Regulation of cellular metabolic process	23	3384	0.45	0.00019
GO:0019222	Regulation of metabolic process	24	3735	0.43	0.00026
GO:0060255	Regulation of macromolecule metabolic process	23	3486	0.44	0.00029
GO:0006950	Response to stress	24	3820	0.42	0.00037
GO:0006807	Nitrogen compound metabolic process	34	7251	0.29	0.00099
GO:0032984	Protein-containing complex disassembly	4	56	1.47	0.0010
GO:0051052	Regulation of DNA metabolic process	5	131	1.2	0.0015
GO:0051716	Cellular response to stimulus	20	3214	0.41	0.0033
GO:0006508	Proteolysis	11	1107	0.62	0.0047
GO:0044238	Primary metabolic process	36	8604	0.24	0.0056
GO:0016560	Protein import into peroxisome matrix, docking	2	4	2.32	0.0059
GO:0048583	Regulation of response to stimulus	9	780	0.68	0.0073
GO:0006355	Regulation of transcription, DNA-templated	16	2386	0.45	0.0089
GO:0050789	Regulation of biological process	28	5971	0.29	0.0089
GO:0006272	Leading strand elongation	2	6	2.14	0.0099
GO:0043981	Histone H4-K5 acetylation	2	7	2.08	0.0126
GO:0019985	Translesion synthesis	2	8	2.02	0.0153
GO:0010556	Regulation of macromolecule biosynthetic process	17	2796	0.4	0.0155
GO:0050794	Regulation of cellular process	25	5300	0.29	0.0209
GO:0031326	Regulation of cellular biosynthetic process	17	2880	0.39	0.0216
GO:0033314	Mitotic DNA replication checkpoint signaling	2	10	1.92	0.0216
GO:0071704	Organic substance metabolic process	36	9277	0.21	0.0239
GO:0050896	Response to stimulus	27	6206	0.26	0.0368
GO:0010468	Regulation of gene expression	17	3061	0.36	0.0395

GO:0000077	DNA damage checkpoint signaling	2	16	1.72	0.0447
GO:0016887	ATP hydrolysis activity	24	401	1.4	1.93e-23
GO:0140657	ATP-dependent activity	26	778	1.14	7.46e-20
GO:0005524	ATP binding	28	2421	0.68	1.76e-10
GO:0016787	Hydrolase activity	30	3284	0.58	3.83e-09
GO:0016464	Chloroplast protein-transporting ATPase activity	5	9	2.36	2.67e-08
GO:0019905	Syntaxin binding	5	14	2.17	1.23e-07
GO:0004176	ATP-dependent peptidase activity	5	37	1.75	7.49e-06
GO:0005488	Binding	49	11632	0.24	2.06e-05
GO:0003678	DNA helicase activity	5	53	1.59	3.22e-05
GO:0005483	Soluble NSF attachment protein activity	3	4	2.49	5.04e-05
GO:0004222	Metalloendopeptidase activity	5	62	1.53	6.22e-05
GO:0008094	ATP-dependent activity, acting on DNA	6	128	1.29	8.76e-05
GO:0042393	Histone binding	5	114	1.26	0.00094
GO:0043167	Ion binding	32	6562	0.31	0.0011
GO:0030337	DNA polymerase processivity factor activity	2	2	2.62	0.0030
GO:1901363	Heterocyclic compound binding	34	7692	0.26	0.0041
GO:0097159	Organic cyclic compound binding	34	7733	0.26	0.0045
GO:0005200	Structural constituent of cytoskeleton	3	31	1.6	0.0061
GO:0005515	Protein binding	17	2556	0.44	0.0070
GO:0008233	Peptidase activity	8	609	0.74	0.0085
GO:0004843	Cysteine-type deubiquitinase activity	3	67	1.27	0.0446
GO:1904949	ATPase complex	27	70	2.21	2.44e-47
GO:0097346	INO80-type complex	22	37	2.39	1.86e-41
GO:0000228	Nuclear chromosome	23	112	1.93	7.82e-35
GO:0005694	Chromosome	29	406	1.47	1.20e-32
GO:0000785	Chromatin	24	214	1.67	7.07e-31
GO:0031011	Ino80 complex	14	28	2.32	3.41e-25
GO:0000812	Swr1 complex	12	13	2.58	7.97e-24
GO:0032991	Protein-containing complex	42	3222	0.73	3.82e-21
GO:0043232	Intracellular non-membrane-bounded organelle	34	1898	0.87	3.45e-20
GO:0140513	Nuclear protein-containing complex	26	947	1.06	4.70e-19
GO:1902494	Catalytic complex	28	1222	0.98	9.37e-19
GO:0031981	Nuclear lumen	26	1028	1.02	2.94e-18
GO:0035267	NuA4 histone acetyltransferase complex	8	19	2.24	7.97e-14

GO:0043229	Intracellular organelle	61	12866	0.3	1.71e-13
GO:0043231	Intracellular membrane-bounded organelle	60	12391	0.3	2.31e-13
GO:0005622	Intracellular anatomical structure	63	14335	0.26	3.53e-13
GO:0005634	Nucleus	42	6070	0.46	2.19e-11
GO:0005654	Nucleoplasm	14	493	1.07	4.53e-10
GO:0062091	Ycf2/FtsHi complex	5	7	2.47	1.67e-09
GO:0005663	DNA replication factor C complex	4	10	2.22	9.03e-07
GO:0097255	R2TP complex	3	3	2.62	7.75e-06
GO:0005778	Peroxisomal membrane	4	43	1.59	0.00014
GO:0043626	PCNA complex	2	2	2.62	0.00087
GO:0140535	Intracellular protein-containing complex	9	716	0.72	0.0014
GO:0009706	Chloroplast inner membrane	4	91	1.26	0.0020
GO:1905368	Peptidase complex	4	109	1.18	0.0036
GO:0005730	Nucleolus	7	515	0.75	0.0055
GO:0031201	SNARE complex	3	54	1.36	0.0074
GO:0009579	Thylakoid	7	586	0.7	0.0111
GO:0016514	SWI/SNF complex	2	14	1.77	0.0134
GO:0031969	Chloroplast membrane	5	295	0.85	0.0152
GO:0000124	SAGA complex	2	17	1.69	0.0182
GO:0031090	Organelle membrane	15	2584	0.38	0.0198
GO:0005777	Peroxisome	5	333	0.8	0.0236
GO:0005829	Cytosol	15	2641	0.37	0.0236
GO:0044815	DNA packaging complex	3	86	1.16	0.0236
GO:0009534	Chloroplast thylakoid	6	517	0.68	0.0270
GO:0009524	Phragmoplast	3	102	1.09	0.0342
GO:0006281	DNA repair	4	387	1.75	0.0012
GO:0006272	Leading strand elongation	2	6	3.26	0.0014
GO:0019985	Translesion synthesis	2	8	3.14	0.0017
GO:0006298	Mismatch repair	2	20	2.74	0.0036
GO:0006275	Regulation of DNA replication	2	53	2.32	0.0188
GO:0030337	DNA polymerase processivity factor activity	2	2	3.74	0.00051
GO:0043626	PCNA complex	2	2	3.74	0.00016