UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
CURSO DE MESTRADO

ESTUDO *IN SILICO*, PURIFICAÇÃO PARCIAL E PERFIL ELETROFORETICO DE PROTEÍNAS INATIVADORAS DE RIBOSSOMOS PRESENTE EM *Jatropha curcas* L.

FRANCIELLY CARVALHO DE OLIVEIRA

CRUZ DAS ALMAS - BAHIA
AGOSTO - 2020

ESTUDO *IN SILICO*, PURIFICAÇÃO PARCIAL E PERFIL ELETROFORETICO DE PROTEÍNAS INATIVADORAS DE RIBOSSOMOS PRESENTE EM *Jatropha curcas* L.

FRANCIELLY CARVALHO DE OLIVEIRA
Nutricionista
Faculdades Integradas de Patos, 2017

Dissertação apresentada ao Colegiado do Programa de Pós-Graduação em Ciências Agrárias da Universidade Federal do Recôncavo da Bahia, como requisito parcial para a obtenção do Título de Mestre em Ciências Agrárias (Área de Concentração: Fitotecnia).

Orientadora: Profa. Dra. Simone Alves Silva
Coorientadora: Profa. Dra. Jacqueline Araújo Castro
Coorientador: Prof. Dr. Angelo Gallotti Prazeres

CRUZ DAS ALMAS - BAHIA
AGOSTO – 2020

FICHA CATALOGRÁFICA

O48e Oliveira, Francielly Carvalho de.
Estudo in silico, purificação parcial e perfil eletroforético de proteínas inativadoras de ribossomos presentes em *Jatropha curcas* L. / Francielly Carvalho de Oliveira._ Cruz das Almas, Bahia, 2020.
122f.; il.

Orientadora: Profª. Drª. Simone Alves Silva.
Co-orientadora: Profª. Drª. Jacqueline Araújo Castro.
Co-orientadora: Prof. Dr. Angelo Gallotti Prazeres

Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, Mestrado em Ciências Agrárias.

1.Pinhão-manso – Melhoramento genético. 2.Pinhão-manso – Fitotecnia. 3.Eletroforese – Análise. I.Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas. II.Prazeres, Angelo Gallotti. III.Título.

CDD: 633.85

Ficha elaborada pela Biblioteca Universitária de Cruz das Almas - UFRB.
Responsável pela Elaboração - Antonio Marcos Sarmento das Chagas (Bibliotecário - CRB5 / 1615).
(os *dados para catalogação foram enviados pelo usuário via formulário eletrônico*).

UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
CURSO DE MESTRADO

ESTUDO IN SILICO, PURIFICAÇÃO PARCIAL E PERFIL ELETROFORETICO DE PROTEÍNAS INATIVADORAS DE RIBOSSOMOS PRESENTE EM *Jatropha curcas* L.

COMISSÃO EXAMINADORA DA DEFESA DA DISSERTAÇÃO DE FRANCIELLY CARVALHO DE OLIVEIRA

Realizada em 17 de Agosto de 2020

Profa. Dra. Simone Alves Silva
Doutora em Melhoramento Genético Vegetal – UFPEL
Universidade Federal do Recôncavo da Bahia - UFRB
Examinador Interno (Orientadora)

Dr. Ciro Ribeiro Filadelfo
Doutor em Ciências – CPqGM/FioCruz
Universidade Federal do Reconcavo da Bahia – UFRB
Examinador Externo

Dr. Bruno Silva Andrade
Doutor em Biotecnologia - UEFS
Universidade Estadual do Sudoeste da Bahia – UESB
Examinador Externo

DEDICATÓRIA

Ao Pai Eterno, criador de todas as coisas, Àquele que me deu a força e o saber.

Aos meus pais, Rosifrança de Carvalho e Francisco Gabriel.

AGRADECIMENTOS

À Deus, que capacitou, me guiou e não me deixou faltar nada. A Ele seja toda honra e toda glória.

À CAPES, pela concessão da bolsa, e ao Programa de Pós-Graduação em Ciências Agrárias da UFRB.

Agradeço a Professora Dra. Simone Alves, por me receber tão bem, pela orientação e por acreditar em minha capacidade de desenvolver esta pesquisa, a Professora Jacqueline Castro por toda paciência e orientação, ao Professor Angelo Gallotti Prazeres, pelas correções. Ao técnico do laboratório e amigo, Ciro Ribeiro, por estar sempre pronto a ajudar e auxiliar em todos os momentos da pesquisa. À professora Cecília Poy, por ceder seu laboratório para que parte da pesquisa fosse possível. Agradeço também a Raner José Santana pelo apoio.

Em geral, a todos os professores do Programa da Pós-Graduação, obrigada pela atenção e ensino.

Aos amigos que fiz no laboratório, Ciro, Leilane, Gean, Hilçana, Ilneide, Celízia, Laurenice e Rosenir, obrigada por todos os momentos de conversas, dos mais sérios aos mais descontraídos, eu aprendi com vocês em todos os momentos e espero ainda aprender muito mais.

A Patricia e Weliton, por terem me apresentado a Instituição e me recebido em Cruz das Almas, sou grata por tê-los comigo.

Por fim, a minha família por ter me ensinado que sempre se deve lutar por aquilo que sonhamos, e que mesmo tão distante, estiveram presentes, apoiando, consolando e sempre orando por mim. Meus pais, Rosifrança de Carvalho e Francisco Gabriel, minha irmã Francélia, e meu noivo, Michael Oliveira, espero um dia recompensá-los por todos esses dias distante. Amo vocês!

SUMÁRIO

Página

RESUMO

ABSTRACT

REFERENCIAL TEÓRICO 1

CAPÍTULO 1

IN SILICO CHARACTERIZATION OF THE GENIC FAMILY OF RIBOSOME INACTIVATING PROTEINS (RIPs) OF PHYSIC NUT (*Jatropha curcas* L.)

.....27

CAPÍTULO 2

SELEÇÃO DO MÉTODO DE EXTRAÇÃO PROTEICA PARA ANÁLISE ELETROFORÉTICA DE SEMENTES DE GENÓTIPOS DE PINHÃO-MANSO

.....59

CONSIDERAÇÕES FINAIS......78

APÊNDICES 79

LISTA DE TABELAS

Página

CAPÍTULO 1

Table 1. Characteristics of RIP genes present in genome of *Jatropha curcas* L, according *Jatropha Curcas* Database (JCDB)…….…….....34
Table 2. Characteristics of RIP proteins present in physic nut (*Jatropha curcas*) genome……...41
Table 3. Post-translational modifications of physic nut RIP proteins…….42
Table 4. Homology modeling using Swiss Model server of JcRIPs and validation of the quality of model obtained……..……47
Table 5. Result of molecular docking process between JcRIPs and their possible ligands……..…….48

CAPÍTULO 2

Tabela 1. Média de Concentrações de Proteínas nas amostras dos genótipos em estudo.....70

LISTA DE FIGURAS

Página

REVISÃO DE LITERATURA

Figura 1.: Imagens de plantas de Pinhão-manso no Banco Ativo de Germoplasma do Núcleo de Melhoramento Genético e Biotecnologia – NBIO, no Campo Experimental da UFRB, Campus de Cruz das Almas.....2
Figura 2. Sintomas de antracnose causados por isolados de *Colletotrichum* spp. em folhas e frutos de nozes; e morfologia dos conídios de ambos os isolados.....6
Figura 3. Sintomas da Podridão do Colo em *J. curcas* L.....8
Figura 4. Representação esquemática do sítio específico da ação das RIPs em ribossomos de ratos.....10
Figura 5. Representação esquemática da estrutura das Proteínas Inativadoras de Ribossomos.....11
Figura 6. Funções das Proteínas Inativadoras de Ribossomos presente em vegetais.....12

CAPÍTULO 1

Figure 1. Structure of the *JcRIPs* genes. The structure of the exons is represented shown in the green box and introns in yellow lines……35
Figure 2. Cis-elements present in the promoter region of physic nut RIP genes….38
Figure 3: Graph of gene expression of *JcRIPs* in different tissues under normal conditions……39
Figure 4. Alignment of multiple sequences of ribosome inactivating protein of *Jatropha curcas* L, and a RIP with known 3D structure (represented by the PDB:5DDZ code)……...45
Figure 5. Overlapping all JcRIPS proteins. The protein represented in violet is JcRIP12 which has two polypeptide chains……...46
Figure 6. 2D diagram of the main molecular coupling combinations. In the first, second and third columns the interaction of the ligands, respectively, with the proteins JcRIP1, JcRIP2 and JcRIP3……50
Figure 7. Coupling of ricin (PDB: 5DDZ) with dexamethasone……..51

CAPÍTULO 2

Figura 1. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de *J. curcas* L. obtidos por diferentes métodos de extração.....68
Figura 2. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de *J. curcas* L. em diferentes estágios de maturação.....69
Figura 3. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de genótipos de *J. curcas* L.....72
Figura 4. Fotografia das sementes dos genótipos usados na pesquisa.....73

ESTUDO IN SILICO, PURIFICAÇÃO PARCIAL E PERFIL ELETROFORETICO DE PROTEÍNAS INATIVADORAS DE RIBOSSOMOS PRESENTE EM *Jatropha curcas* L.

Autora: Francielly Carvalho de Oliveira
Orientadora: Dra. Simone Alves Silva

RESUMO: O pinhão-manso (*Jatropha curcas* L.) é uma espécie oleaginosa pertencente à Família Euphorbiaceae, reconhecida mundialmente por sua importância como fonte de biodiesel. Seu genoma já foi sequenciado e depositado em bancos de dados. Dentre os genes de maior interesse agronômico têm-se as Proteínas Inativadoras de Ribossomos (*Ribosome Inactivating Proteins* -RIP), pois, fazem parte da classe de genes envolvidos na defesa contra estresses ambientais causados por fitopatógenos (fungos, vírus e insetos). No presente estudo objetivou-se, caracterizar a família de genes RIP *J. curcas* L., incluindo a análise do promotor, modelagem e acoplamento proteína-ligante. Bem como, purificar parcialmente e caracterizar o perfil eletroforético das RIPs de sementes de pinhão-manso. A anotação existente JCDB permitiu a identificação de um total de doze genes RIP. Dos doze genes, onze foram classificados como RIPs do tipo 1 e um do tipo 2. Os elementos cis responsivos à luz, MYB e MYC foram identificados em todas as regiões promotoras. Em proporções menores, elementos cis que respondem a estresses (bióticos e abióticos) e hormônios. Todas as proteínas JcRIP foram modeladas e apresentaram boa qualidade. Foram testados compostos inibidores já conhecidos para RIPs, ácido fólico e ácido pteróico, e também a dexametasona e seus derivados, ambos apresentaram boa afinidade. A análise por eletroforese (SDS-PAGE) possibilitou selecionar o método de extração proteica com melhor resolução eletroforética, que foi com Tampão PBS. As concentrações de proteínas das amostras não divergiram estatisticamente. Pelo perfil eletroforético foi possível identificar as RIPs, com peso molecular de 28 kDa, em todos os genótipos avaliados. Foi observado um perfil de bandas diferente no genótipo UFRB05 e UFRB15. Quando associado com a qualidade das sementes armazenadas e teor de óleo o genótipo que mais se destacou foi o UFRB05.

Palavras-chave: Bioinformática, pinhão-manso, RIPs.

IN SILICO STUDY, PARTIAL PURIFICATION AND ELECTROPHORETIC PROFILE OF INACTIVATING RIBOSOME PROTEINS IN Jatropha curcas L.

Author: Francielly Carvalho de Oliveira
Advisor: Dr. Simone Alves Silva

ABSTRACT: The physic nut (*Jatropha curcas* L.) is an oilseed species belonging to Euphorbiaceae Family, recognized worldwide for its importance as a source of biodiesel. Its genome has already been sequenced and deposited in databases. Among genes of greatest agronomic interest are Ribosome Inactivating Proteins (RIP), as they are part of class of genes involved in defence against environmental stresses caused by phytopathogens (fungi, viruses and insects). The present study aimed to characterize RIP *J. curcas* L. gene family, including promoter analysis, modeling and protein-ligand coupling. As well as, partially purify and characterize electrophoretic profile of physic nut seeds. The existing JCDB annotation allowed identification of a total of twelve RIP genes. Of the twelve genes, eleven were classified as type 1 RIPs and one type 2. The light-responsive cis elements, MYB and MYC were identified in all promoter regions. In smaller proportions, cis elements that respond to stresses (biotic and abiotic) and hormones. All JcRIP proteins were modeled and showed good quality. Inhibitor compounds already known for RIPs, folic acid and pteroic acid were tested, as well as dexamethasone and its derivatives, both showed good affinity. The analysis by electrophoresis (SDS-PAGE) made it possible to select the method of protein extraction with the best electrophoretic resolution, which was with PBS buffer. The protein concentrations of the samples did not differ statistically. Through the electrophoretic profile, it was possible to identify RIPs, with a molecular weight of 28 kDa, in all evaluated genotypes. A different band profile was observed in UFRB05 and UFRB15 genotype. When associated with quality of stored seeds and oil content, genotype that stood out the most was UFRB05.

Keyword: Bioinformatics, physic nut, RIPs

REFERENCIAL TEÓRICO

1 PINHÃO-MANSO (*Jatropha curcas* L.)

Existem aproximadamente 170 espécies de plantas pertencentes ao gênero *Jatropha*, estando distribuídas em regiões tropicais e subtropicais da América e África, e grande parte destas são conhecidas por serem altamente tóxicas e alergênicas (KRISHNAN & PARAMATHMA., 2009). As espécies mais conhecidas do gênero são *J. podagrica*, conhecida na Nigéria como *lapalapa funfun* (AIYELAAGBE,et al., 2007); *J. multifida*, conhecida pelos nomes populares planta-coral, bálsamo, metiolate feridas (BUCH, ARANTES, & CAMPELO, 2008); e a *J. curcas* L., conhecida como pinhão-manso (MARTINS, et al., 2008).

A espécie *J. curcas* L. (pinhão-manso) também é conhecida por purgueira, pinha de purga e grão-de-maluco, é um arbusto de folhas pequenas, amarelo-esverdeadas, seu fruto é seco, formado por uma casca dura e lenhosa, inicialmente verde, passando por uma coloração amarela, marrom e preto, quando atinge a maturação final, possuindo 3 sementes lisas e escuras, apresentando propriedades medicinais e usadas também na produção de óleo (MARTINS, et al., 2008). É uma espécie com ciclo vegetativo perene, não necessitando de renovação anual para o plantio, tendo sua primeira produção oito meses após o plantio, possui elevado rendimento de sementes, consequentemente, de óleo, chegando a produzir cerca de 2.000 $kg.ha^{-1}$ de óleo ao ano, enquanto a soja tem rendimento estimado entre 500 – 700 $kg.ha^{-1}$ (DURÃES et al., 2009).

Devido às suas características físico-químicas favoráveis à produção de biodiesel e bioquerosene de aviação, com teores dos ácidos oleico e linoleico variando de 22 a 42%, o óleo de pinhão-manso é apontado como uma das principais fontes alternativas para a produção de biocombustíveis no Brasil (WANI et al., 2012; DURÃES et al., 2011; LAVIOLA et al., 2014). Com isso, tem aumentado o interesse dos estudiosos na cultura, visando uma maior produção de óleo, por meio de estudos de melhoramento genético.

Figura 1.: Imagens de plantas de Pinhão-manso no Banco Ativo de Germoplasma do Núcleo de Melhoramento Genético e Biotecnologia – NBIO, no Campo Experimental da UFRB, Campus de Cruz das Almas. Fonte: Autor (2019).

1.1 MELHORAMENTO GENÉTICO DA CULTURA

As estratégias de melhoramento adotadas para cada espécie de plantas estão relacionadas ao seu sistema de propagação. Geralmente, o melhoramento de plantas perenes compreende três: o pré-melhoramento, a partir da avaliação do germoplasma para o estabelecimento de populações melhoradas, levando em consideração o seu desempenho individual e informações genéticas; o melhoramento, com base no desenvolvimento de clones promissores ou na

recombinação de indivíduos superiores na população reprodutora afim de obter uma progênie promissora, seguindo para a avaliação em campo, quanto às características de interesse e posteriormente selecionada com base em seu valor genético; e por fim, a liberação das melhores variedades ou cultivares, a realização de testes de indivíduos previamente selecionados com base em seu potencial genético e a seleção destes para a liberação comercial (ALVES et al., 2015; LAVIOLA et al., 2017; SILVA; SANTOS, 2019).

Somente no ano de 2008, por meio da normativa nº 4, de 14 de janeiro do corrente ano, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), autorizou a inscrição da espécie no Registro Nacional de Cultivares (RNC) (BRASIL, 2008).

O pinhão-manso é uma cultura que ainda não está totalmente domesticada, porém, seu progresso e estudos de melhoramento tem evoluído nos últimos anos, apresentando um alto potencial para produção de bioenergia, produzindo sementes ricas em óleo e proteínas de alta qualidade. Com características bem atraentes e importantes para o seu melhoramento, como sua capacidade de adaptação e sobrevivência em ambientes quentes e secos, no entanto, o desnível de maturação de suas sementes, acarreta alguns problemas, como o maior tempo de mão-de-obra ocasionando maior custo para produção. Além disso, o surgimento de doenças, como a antracnose e a podridão do colo, causadas por fungos prejudicada a produção de sementes e boa parte das plantações (LAVIOLA, 2014; MONTES & MELCHINGER, 2016).

No Brasil, existem vários bancos ativos de pinhão-manso estabelecidos. Situados na Embrapa Algodão, com genótipos de diferentes Estados brasileiros, Embrapa Rondônia com 236 genótipos; Embrapa Agroenergia e Embrapa Cerrado, com genótipos de regiões do Brasil e México; e o Banco Ativo de Germoplasma (BAG) do Departamento de Fitotecnologia da Universidade Federal de Viçosa consiste em 78 acessos com ampla variabilidade, para os quais foram relatados anteriormente vários caracteres juvenis e teor de óleo (SPINELLI et al., 2010; FREITAS, et al., 2011; LAVIOLA et al., 2012; LAVIOLA, et al., 2017).

O Núcleo de Melhoramento Genético e Biotecnologia (NBIO), da Universidade Federal do Recôncavo da Bahia (UFRB), iniciou suas pesquisas com esta oleaginosa no ano de 2008, com acessos trazidos da Empresa Baiana de

Desenvolvimento Agrícola (EBDA), das cidades de Alagoinhas e Irará, e acessos provenientes de intercâmbio com a Universidade Federal de Viçosa (UFV) de diferentes regiões da Bahia. O tamanho efetivo da população foi de ~229 no momento da coleta e sofreu uma redução para ~200 após a propagação e introdução no BAG da UFRB. Nos primeiros estudos, já foi possível detectar a divergência e variabilidade genética entre as famílias de meio-irmãos (CARVALHO, 2010; BRASILEIRO et al., 2013; PESTANA-CALDAS et al., 2016).

Queiroz (2015), em seu estudo com 20 progênies de meio-irmãos de pinhão-manso, utilizando marcadores RAPD (*Random Amplified Polymorphic* DNA) revelou a existência de variabilidade genética entre as progênies avaliadas, das quais as mais divergentes foram: UFRB-PM 1, UFRB-PM 6, UFRB-PM 9, UFRB-PM 12, UFRB-PM 13, UFRB-PM 19, UFRB-PM 11 e UFRB-PM 8, e podem ser indicadas para combinações híbridas em programas de melhoramento da espécie, podendo as mesmas serem utilizadas como parentais em programas de melhoramento genético da espécie.

Ao avaliar a diversidade genética e o desempenho de clones de pinhão-manso desenvolvidos pelo mesmo programa, através de caracterização morfológica para seleção de variedades clonais, Almeida et al (2016), identificaram que os clones UFRBPR14 e UFRBPR15 foram mais divergentes, sendo: o número de ramos, a altura das plantas e o número de cachos as características que contribuíram para a divergência genética.

Ao realizar um estudo com oito destes clones, Farias (2018) detectou que os caracteres de produção de frutos (número de frutos) e óleo (produtividade de óleo) são os responsáveis pela detecção de variabilidade genética entre os clones, com destaque para UFRB05, UFRB13, UFRB14. Sugerindo que quando se retira o tegumento da semente é possível obter uma maior quantidade de óleo.

Atualmente, estão sendo desenvolvidas novas pesquisas nesta cultura, com o objetivo de identificar e controlar as doenças que tem causado tombamentos e morte de diversas plantas do BAG do NBIO.

1.2 PRINCIPAIS DOENÇAS DA CULTURA

Apesar de vários estudos trazerem o pinhão-manso como uma espécie resistente a pragas e doenças, a expansão de áreas cultivadas é acompanhada pelo surgimento de várias doenças, majoritariamente causada por fungos (MACHADO; PEREIRA, 2012).

A antracnose, uma das doenças de plantas mais importantes do mundo, acometendo diversas culturas, está presente em todas as áreas de cultivo do pinhão-manso, causada por espécies do gênero *Colletotrichum*, que ataca quase todos os órgãos da planta, especialmente as partes aéreas, apresentando sintomas como lesões necróticas de formato irregular, com colorações pardas a negras, que aparecem no centro e nas bordas da folha, podendo infectar também o fruto (Figura 2.) (STRANGE e SCOTT, 2005; DEAN et al., 2012; SANTOS et al., 2013; VELOSO et al., 2018).

Espécies de *Colletotrichum* apresentam-se como hemibiotróficos, obtendo nutrientes a partir de células vivas, progredindo para a segunda fase, onde se alimentam das células mortas, destruídas pelo fungo, sendo esta última induzida pela maturação e senescência dos frutos (PRUSKY e LICHTER, 2007).

No Brasil, a antracnose do pinhão-manso ocorre principalmente durante a estação chuvosa e é causada por duas espécies, *C. gloeosporioides* (Penz.) Sacc. e *C. capsici* (Syd.) Butler & Bisby, com sintomas semelhantes para ambos os patógenos (FREIRE; PARENTE, 2006).

No pinhão-manso, a antracnose causada pelo *C. capsisci* foi relatada e confirmada no México, em agosto de 2008. Foram observados sintomas de necrose foliar, podridão da coroa com morte apical de plântulas e queima das folhas causando 70% de desfolhamento. A identificação do fungo foi confirmada pelo Postulado de Koch e informações de dados moleculares como: amplificação de espaçadores internos transcritos (ITS - do inglês *Internal transcribed spacer*) com primers universais ITS1 / ITS4 e análises de RFLP-PCR (*Restriction Fragment Length Polymorphism - Polymerase Chain Reaction*) usando enzimas de restrição Alu I e Hae III (TORRES-CALZADA, et al., 2011).

Já a antracnose, causada pelo fungo *C. gloeosporioides,* foi relatada na Coreia no verão de 2010 a 2011, cujos sintomas eram o aparecimento de manchas marrons escuras nas folhas e a mumificação do fruto. O fungo foi identificado com base nas características morfológicas, no teste de patogenicidade e na

identificação molecular por meio da análise de alinhamento básico local (BLAST) da região ITS rDNA (ribosomal DNA), mostrando correspondência exata com sequências nucleotídicas de isolados de *C. gloeosporioides* descritos na China (número de acesso EU520076) e na Nova Zelândia (número de acesso EU149938) (KWON, et al., 2012).

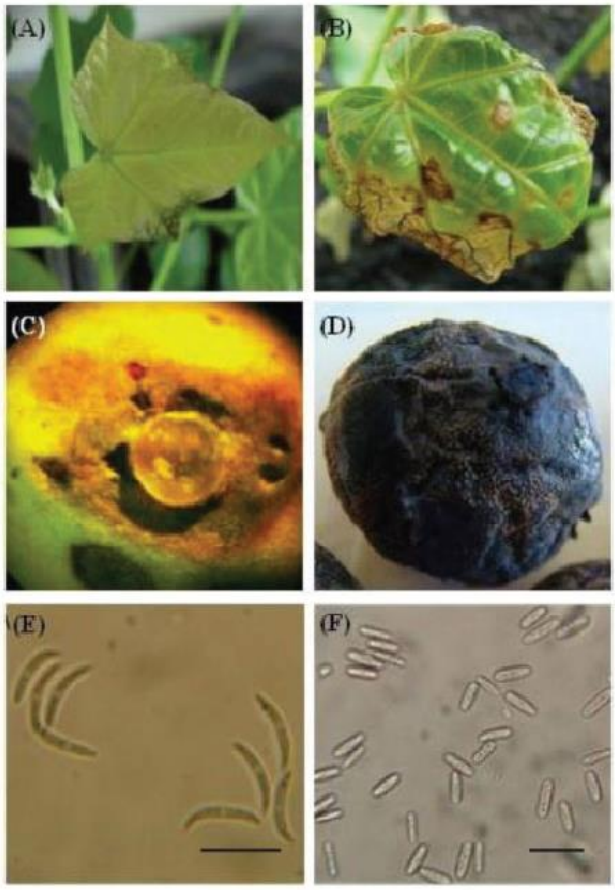


Figura 2. Sintomas de antracnose causados por isolados de *Colletotrichum* spp. em folhas e frutos de nozes; e morfologia dos conídios de ambos os isolados. (A) sintomas iniciais nas folhas; (B) sintomas avançados nas folhas; (C) desenvolvimento do sintoma inicial nos frutos a partir do local da inoculação; (D) fruto completamente necrótico e coberto pela esporulação laranja do patógeno; e (E) conídios do isolado P-01 (*C. capsici*), (F) conídios do P-02 isolado (*C. gloeosporioides*). Bar = 20 µm. FONTE: Imagens extraídas de Santos et al., (2013).

Uma outra fitomoléstia que tem se destacado na cultura do pinhão-manso, é a podridão do colo, que além de reduzir a produtividade, causa a morte súbita de plantas, provocando a podridão radicular e do colo, na qual os sintomas iniciais são murcha e amarelecimento das folhas, culminando com a queda das folhas e morte da planta (Figura 3.). No Brasil e na Índia, acreditava-se que esta doença estava

associada apenas ao *Lasiodiplodia theobromae* (Pat.) Griffon & Maubl. (LATHA et al. 2009; PEREIRA et al. 2009).

Espécies do gênero Botryosphaeriaceae são consideradas endofíticas em diversas plantas arbóreas, e normalmente conhecidas como oportunistas, e quando ocorre algum fator, a exemplo de ferimentos ou estresse hídrico, tornam-se patogênicos, causando sintomas em seus hospedeiros (SMITH, et al., 1996; TAYLOR, et al., 2005; GONÇALVES, et al., 2016).

Machado et al., (2014), a partir de estudos morfológicos e filogenéticos, baseados em regiões ITS, sequências de β-tubulina e fator de alongamento de tradução 1-α, identificaram nove espécies de *Botryosphaeriaceae* associadas a *Jatropha.* Essas espécies incluem *L. egyptiacae, L. pseudotheobromae, L. theobromae, Macrophomina phaseolina, Neoscytalidium hyalinum* e quatro *Lasiodiplodia* spp. que são propostas como novas espécies (*L. euphorbicola, L. jatrophicola, L.macrospora e L. subglobosa*), todas, exceto *M. phaseolina*, são patogênicas. Assim a podridão do colo e das raízes das plantas de pinhão-manso é causada por patógenos complexos.

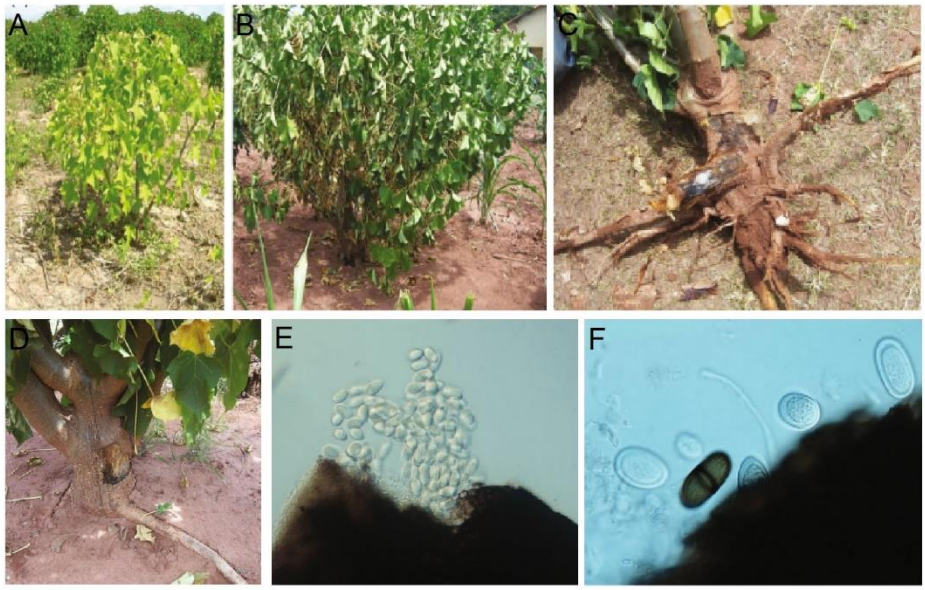


Figura 3. Sintomas da Podridão do Colo em *Jatropha curcas* L. A e B: amarelecimento e murchas das folhas; C e D: apodrecimento das raízes e colo; E e F: *Lasiodiplodia theobromae* em *Jatropha curcas*. (E) Conídios imaturos, hialinos, unicelulares sendo liberados do picnídio, tornando-se (F) marrom escuro, didiósporos e estriados com a idade. FONTE: Imagens extraídas de Pereira et al., (2009).

Além destas, outras patologias também foram relatadas em *J. curcas* L., tais como: oídio, causado pelo fungo *Pseudoidium jatrophae* (BRAUN & COOK, 2012); ferrugem, cujo agente etiológico é o *Phakopsora arthuriana* (DIAS et al., 2007); e o mosaico amarelo, promovida pelo vírus do mosaico (begomovírus), e seu vetor é a mosca branca (*Bemisia tabaci*) (NARAYANA et al., 2006).

Diante de tamanha diversidade de fitopatógenos capazes de causar danos ao pinhão-manso este apresenta uma diversidade de mecanismos de defesa que incluem as Proteínas Inativadoras de Ribossomos (RIPs), as quais podem conferir intoxicações à animais após a ingestão de sementes e folhas.

1.3 TOXICIDADE

A presença de antinutrientes nas sementes, tais como inibidores da tripsina, lectinas, fitatos e ésteres de forbol altamente tóxicos, inviabilizam o uso de *J. curcas* L. na alimentação animal (MAKKAR et al., 2008). Dessa forma, para tornar viável a utilização desta planta na alimentação animal, existem vários métodos e processos de destoxificação de sementes e frutos, visando à inativação dos esteres forbol e das proteínas tóxicas, exigindo assim, tratamentos químicos prévios, favorecendo a remoção de até 89% das substâncias tóxicas. No entanto, o aparelho digestório dos animais ainda pode ser intoxicado, devido a quantidades remanescentes de tais substâncias (RAKSHIT et al., 2008).

A presença de compostos alergênicos nas sementes de pinhão-manso, podem causar sintomas como: vermelhidão e coceira na mucosa, pelo simples contato com a pele ou olhos, tornando o manuseio incomodo, mas não perigoso (MENDONÇA; LAVIOLA, 2009).

O éster de forbol é o principal composto responsável pela elevada toxicidade das sementes desta cultura. No México, mesmo as variedades de pinhão-manso consideradas não tóxicas, apresentam uma quantidade de 0,11 $mg.g^{-1}$. Em um estudo realizado por Ferrari et al. (2010), evidenciou que o teor de ésteres de forbol, variou de 1,41 a 8,97 $mg.g^{-1}$ nas sementes, indicando a ocorrência de plantas com baixa e alta toxicidade (GUBITZ et al., 1999; MAKKAR et al., 1998; FERRARI et al., 2010).

Além dos esteres de forbol, o pinhão-manso possui uma classe de proteínas que conferem toxidez as suas sementes. A curcina, substância pertencente à classe das RIPs, é capaz de entrar nas células e se ligar a ribossomos, interrompendo a síntese proteica e causando a morte celular. Ela está presente em folhas e sementes de *J. curcas* L. (QIN et al., 2010; ENDO; TSURUGI, 1988).

1.3.1 PROTEÍNAS INATIVADORAS DE RIBOSSOMOS

O mecanismo molecular de ação das RIPs envolve a inibição da síntese de proteínas, por meio da ação N-glicosidase, a qual realiza a quebra da ligação N-glicosídica que liga a A4324 (Adenina na posição 4324) ao esqueleto de polifosfato do rRNA 28S, presente nos ribossomos eucariotos. Essa adenina está localizada em um *loop*, denominado “ricin/sarcin”, altamente conservado nas diferentes espécies e relatado como essencial para a interação dos fatores de alongamento nos ribossomos eucarióticos (28S rRNA) e procarióticos (23S rRNA) (Figura 4.) (ENDO, et al., 1987; FURUTANI, et al., 1992; GIRBÉS, et al., 2004).

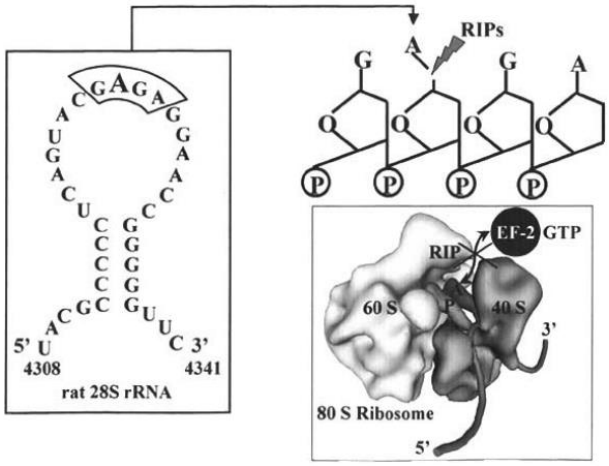


Figura 4. Representação esquemática do sítio específico da ação das RIPs em ribossomos de ratos. As RIPs clivam uma adenina do rRNA da subunidade ribossomal maior, impedindo a ligação dos fatores de alongamento. FONTE: Figura extraída de Girbés et al., (2004).

As Proteínas Inativadoras de Ribossomos (RIPs), podem ser encontradas em diversos tipos de plantas, bem como em fungos e bactérias (SHANG; ROUGÉ; VAN DAMME, 2016; MELTON-CELSA, 2014). São mais comuns em plantas da família: Caryophyllaceae, Sambucaceae, Euphorbiaceae, Cucurbitaceae,

Poaceae, Phytolaccaceae e Rosaceae (GIRBÉS et al., 2004; DOMASHEVSKIY E GOSS, 2015; SHANG et al., 2016). São classificadas em três tipos, RIPs tipo I, RIPs tipo II e RIPs tipo III (Figura 5).

As RIPs tipo 1 são compostas por um único domínio polipeptídico, de aproximadamente 30 kDa, sendo amplamente distribuídas. A primeira RIP tipo 1 a ser identificada, foi a proteína antiviral pokeweed (PAP) em *Phytolacca americana* L. (DALLAL e IRVIN, 1978), posteriormente, várias RIPs foram encontradas distribuídas em diversas espécies de plantas. Embora compartilhem resíduos de aminoácidos conservados no sítio ativo, a sequência geral e modificações pós-traducionais dos mesmos, são significativamente diferentes (MONZINGO E ROBERTUS, 1992; BARBIERI et al., 1993; MLSNA et al., 1993; HUSAIN et al., 1994; ZHU, et al., 2018).

As RIPs Tipo II possuem, em sua composição estrutural, duas cadeias polipeptídicas, em que a cadeia A é semelhante a cadeia das RIPs tipo I e a cadeia B é constituída por um domínio de lectina, ligadas por uma ponte dissulfeto. A ricina foi a primeira RIP tipo II a ser identificada e está presente nas sementes de mamona (*Ricinus Communis* L.), sendo a proteína mais conhecida entre as RIP tipo II. Este grupo de proteínas são mais citotóxicos do que as RIPs tipo I, e apresentam maior facilidade de adentrar nas células, devido a presença da cadeia de lectina (WALSH; MORGAN; HEY, 1991; HARTLEY, LORD, 2004; ZHU, et al., 2018).

Já as RIPs tipo III, têm um domínio N-terminal correspondente ao domínio A das demais RIPs, estando fundido a um domínio C-terminal com função ainda desconhecida. Foram caracterizados apenas a partir de milho (*Zea mays* L.) e cevada (*Hordeum vulgare*), e são bem menos conhecidas e relatadas do que as demais RIPs tipo I e tipo II (WALSH; MORGAN; HEY, 1991; REINBOTHE, S. et al., 1994; ZHU, et al., 2018).

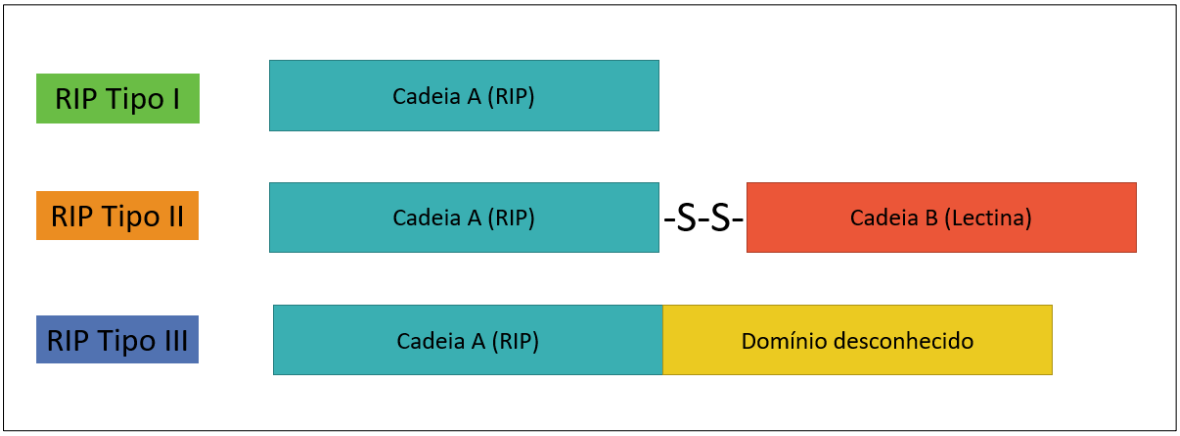


Figura 5. Representação esquemática da estrutura das Proteínas Inativadoras de Ribossomos. Tipo I, Tipo II e Tipo III**.** Cadeia A: domínio funcional RIP, Cadeia B: domínio de lectina. S-S: Ponte dissulfeto. Fonte: Autor (2020).

As RIPs apresentam mecanismos de entrada celular diferentes, estando de acordo com o tipo envolvido; onde as RIPs tipo 1, realizam a entrada na célula por meio de endocitose, mediada por receptores. Já as RIPs tipo 2, realiza a entrada na célula de forma mais facilitada, pois, apresentam uma cadeia de lectina que se liga a glicoproteínas ou glicolipídios da membrana celular, as quais medeiam sua entrada na célula (WANG, et al., 2016).

As RIPs parecem possuir distintos papéis entre determinadas espécies. Estão envolvidas na defesa da planta, com funções: antivirais, antifúngicas, inseticida, de proteção contra estresse e regulação do desenvolvimento, apresentando também funções de armazenamento e senescência programada (GIRBES et al., 2004). O papel antiviral das RIPs é um dos tópicos mais estudados por diversos autores. A PAP e outros tipos de RIPs de *Phytolacca,* foram relatadas como poderosos inibidores das infecções virais em plantas (PICARD et al., 2005) e animais (ISHAG et al., 2013). Lanzanova (2011), também propôs um papel antifúngico para os RIPs, na qual estudos realizados em diversas espécies, como: tabaco (*Nicotiana tabacum*), trigo (*Triticum* spp) e milho (*Z. mays* L.), confirmaram que a incorporação do gene b-32 do milho que expressa a proteína b-32, uma RIP do tipo 1, apresentou-se como uma ferramenta eficaz e confiável para o melhoramento e controle de doenças fúngicas das culturas (Figura 6.).

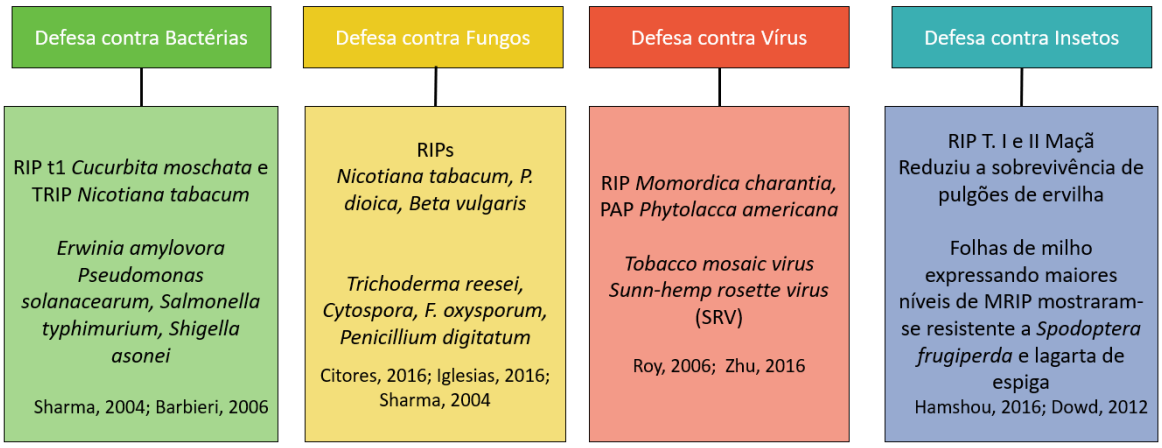


Figura 6. Funções das Proteínas Inativadoras de Ribossomos presente em vegetais. Fonte: SHARMA et al., (2004); BARBIERI, (2006); CITORES, (2016); IGLESIAS, (2016); ROY, (2006); ZHU, (2016); HAMSHOU, (2016); DOWD, (2012).

O papel e a via de atuação das RIPs na defesa da planta ainda não foram totalmente elucidados. Diversos estudos relatam a indução da expressão destas proteínas, após tratamentos com fitormônios, como: ácido jasmônico (JA), ácido salicílico (SA), ácido abscísico (ABA) ou estresses associados, a insetos, infecção por patógenos, frio, calor, seca, salinidade e ferimentos mecânicos (REINBOTHE et al., 1994 ; IGLESIAS et al., 2005; JIANG et al., 2008; QIN et al., 2009; TARTARINI et al., 2010; VIRGILIO, et al., 2010; NELLER et al., 2019).

As RIPs presentes no pinhão-manso, denominadas de curcina, podem ser encontradas nas sementes e em folhas. A curcina encontrada no endosperma, atinge seu pico de expressão no período embrionário maduro; já aquela presente nas folhas, é induzida por estresses bióticos e abióticos, tais como: baixa e alta temperatura, infecção fúngica e tratamento com fitormônios (QIN et al., 2010).

Muitos dos trabalhos supracitados foram realizados a partir de estudos utilizando a bioinformática, aplicando técnicas de análises estruturais dos genes, filogenia e estudos conformacionais das proteínas, para predizer de forma teórica, ou seja, *in sílico*, suas funcionalidades para posteriores estudos *in vitro* e *in vivo*.

1.4 BIOINFORMÁTICA E ANÁLISE *IN SILICO*

A bioinformática é considerada uma ciência interdisciplinar, baseada no uso da informática aplicada nos estudos da biologia, e abrange diversas áreas do

conhecimento, como a química, física, matemática e estatística, desenvolvendo métodos para armazenamento e recuperação de dados biológicos, elaboração e uso de ferramentas e algoritmos para a resolução de problemas biológicos. Permitindo que os cientistas consigam decifrar e gerenciar as vastas quantidades de dados disponíveis, aplicando princípios científicos e tecnologias da informação a dados complexos das ciências da vida (OJO; OMABE, 2010; ATTWOOD et al., 2011).

A origem do termo bioinformática se deu em 1970, por Ben Hesper, com o intuito de dar nome a sua pesquisa em que desenvolvia o estudo de processos informacionais em sistemas bióticos. Contudo, alguns algoritmos clássicos da bioinformática, como o alinhamento de sequências e filogenia, já haviam sido desenvolvidos na década de 60 (HOGEWEG, 2011).

A evolução da bioinformática tem oferecido avanços nas ciências “ômicas”, permitindo uma relação entre o genoma funcional e a informação codificada, especialmente nas anotações dos transcriptomas. Propondo novas formas de ciência, baseada na experimentação *in silico*, prevendo estruturas de proteínas e moléculas, realizando testes de interação, inibição ou excitação de moléculas, criando inibidores, moléculas de interferência, entre outras atividades (BADER; HOGUE, 2002; KEMMEREN et al., 2002; ESPINDOLA et al., 2010).

No entanto, os bancos de dados precisam ser alimentados e organizados em linguagem universal, afim de facilitar sua leitura. A partir da geração de novos dados e conhecimentos para que diversos produtos e soluções sejam desenvolvidos, correlacionando a bioinformática à biotecnologia, por exemplo, por meio da modelagem por homologia (ESPÍNDOLA et al., 2010).

1.4.1 MODELAGEM POR HOMOLOGIA

Experimentalmente, a determinação de estruturas proteicas pode ser alcançada por meio da difração de raios-X, considerada a técnica padrão ouro (NAKANO et al., 2012); por ensaios de ressonância nuclear magnética (RNM), contudo, essa técnica não se aplica a todos os tipos de proteínas (KAY, 2005); e também por crio-microscopia eletrônica (CALLAWAY, 2015). Esses protocolos são

muito trabalhosos e custosos, que podem levar anos para a obtenção do cristal de uma proteína (SANTOS-FILHO; ALENCASTRO, 2003).

Com o crescimento da tecnologia e a chegada da bioinformática, é possível que a estrutura de uma proteína seja determinada por meios não experimentais, como é o caso da modelagem molecular (KLEPEIS et al., 2005).

A modelagem por homologia de uma proteína (proteína-problema), baseia-se no conceito de evolução molecular, isto é, parte-se do princípio de que há semelhança entre as estruturas primárias desta proteína com a proteína-molde (CHOTIA & LESK, 1986).

A modelagem por homologia é importante em diversas áreas, como a biologia estrutural, bioquímica e biofísica, principalmente em estudos relacionados com os genomas. A modelagem apresenta um potencial imenso, mesmo com suas limitações, pois é capaz de acelerar o processo de elucidação de estruturas proteicas com custos reduzidos e em curto espaço de tempo, tornando-se uma ferramenta adequada para a predição teórica da estrutura de proteínas. Dentre as técnicas de predição de estrutura de proteínas *in sílico,* a modelagem por homologia foi eleita como o método mais acurado (SANTOS FILHO; ALENCASTRO, 2003; MOULT et al., 2011; CALIXTO, 2013).

Os métodos de modelagem de proteínas por homologia, implicam basicamente em quatro passos sucessivos: (1) identificação e seleção de proteínas-molde; (2) construção do modelo; (3) construção das coordenadas do modelo; (4) validação do modelo (CALIXTO et al, 2013).

No primeiro passo, surgem duas situações possíveis, (i) quando se conhece a família da proteína-problema, e (ii) quando não se sabe a que família a proteína-problema pertence. Para o primeiro caso, é suficiente selecionar a proteína molde diretamente do PDB (*Protein Data Bank*), banco de dados de informações sobre as estruturas 3D de proteínas, ácidos nucleicos e conjuntos complexos (proteína-ligante); e no segundo, faz-se necessário procurar de forma sistemática um ou mais moldes adequados em um banco de dados de estruturas primárias derivadas de proteínas armazenadas no PDB (SANTOS-FILHO, ALENCASTRO, FIGUEROA-VILLAR, 2001).

Atualmente, existem diversos servidores e programas capazes de gerar bons modelos estruturais de proteínas. O Swiss-model

(www.swissmodel.expasy.org/) é um dos servidores de modelagem de proteínas automatizado mais conhecido atualmente, e possui uma biblioteca de modelos com curadoria, atualizada semanalmente de acordo com a versão do PDB (BERMAN et al., 2002; BIASINI et al., 2014; WATERHOUSE et al., 2018).

A acurácia da estrutura, será diretamente proporcional ao grau de identidade com a proteína-molde. Quando o percentual de identidade oscila entre 30% e 50% ou mais, o modelo gerado apresentará de boa a ótima qualidade e com resultados precisos, se igualando a estruturas resolvidas por raios-X de baixa resolução (KOPP; SCHWEDE, 2004; FRENKEL; TRIFONOV, 2007; CALIXTO, 2013).

O alinhamento das sequências de resíduos, tem como objetivo ajustar resíduos estruturalmente equivalentes, levando em conta características estruturais comuns. O alinhamento pode ser simples, utilizando apenas um molde, e múltiplo, onde emprega-se todos os moldes homólogos disponíveis no PDB, sendo este mais confiável (TUCCINARDI et al., 2010).

O alinhamento local, BLAST (*Basic Local Alignment Search Tool* - http://www.ncbi.nlm.nih.gov/blast/), é a maneira mais comum de comparação entre sequências, pois gera uma lista com as proteínas de estruturas conhecidas que poderão ser empregues como molde, usando como entrada a proteína-alvo (WONG et al., 2011).

A construção das coordenadas do modelo, pode ser realizada por diversos programas computacionais, dentre eles o servidor SWISS-MODEL, um servidor de modelagem de homologia de estrutura de proteínas totalmente automatizado e acessível, por meio do servidor web ExPASy ou do programa DeepView (Swiss Pdb-Viewer). Esse é um dos servidores de modelagem de estrutura mais utilizados no mundo, gerando ~3.000 modelos por dia (~2 modelos por minuto) (GUEX; PEITSCH; SCHWEDE, 2009; WATERHOUSE et al., 2018).

A validação, é uma etapa essencial, e pode ser executada em diferentes níveis de organização estrutural, em que se deve avaliar a qualidade do empacotamento global da proteína, os possíveis erros estruturais em regiões localizadas e os parâmetros estereoquímicos (VYAS et al., 2012; RAMACHANDRAN et al., 1963; WIEDERSTEIN; SIPPL, 2007; CHEN et al., 2010). Atualmente, há uma variedade de programas e servidores gratuitos que podem ser

usados na validação de modelos tridimensionais gerados por homologia, dentre eles estão o Molprobity, ERRAT2, PROSHECK e ProSA.

O servidor ProSA, calcula os *scores* de qualidade geral e local da sequência do modelo comparando-o com outras proteínas depositadas no PDB que possuem um tamanho aproximado ao da proteína-alvo. Estes scores são plotados em gráficos, e espera-se um valor negativo para bons modelos (WEIDERSTEIN; SIPPL, 2007).

O gráfico de Ramachandran, é um método simples de validação, que julga a qualidade estrutural da proteína, baseada na distribuição dos ângulos diedros da cadeia principal, definindo os valores permitidos dos ângulos φ e ψ em estruturas tridimensionais, indicando os resíduos de aminoácidos que se encontram nas regiões energicamente mais favoráveis e desfavoráveis, prevendo conflitos estéricos e orientando na avaliação da qualidade dos modelos teóricos ou experimentais de proteínas (KLEIWEGT; JONES, 1996).

A partir da tecnologia de modelagem molecular é possível realizar estudos de ancoragem molecular, para a obtenção de diversos complexos proteína-ligante com a utilização de variados ligantes.

ANCORAGEM MOLECULAR

Técnicas de ancoragem molecular (*docking molecular*) fornecem estimativas *in sílico* de afinidade entre a proteína e o ligante, antes mesmo que esses sejam sintetizados, reduzindo os custos com gastos laboratoriais despendidos ao sintetizar e testar farmacologicamente determinadas substâncias. Esta importante ferramenta, é usada para filtrar compostos que não servem para serem designados como alvo e desenhar os possíveis candidatos que apresentariam uma boa interação com o sítio ativo do receptor (RODRIGUES et al., 2012).

As interações proteína-ligante, se dão através de forças intermoleculares do tipo dipolo-dipolo induzido (*Van Der Waals* ou *London*), interação eletrostática (Interação Coulombiana) e a formação de ligações de hidrogênio (PAGADALA; SYED; TUSZYNSKI, 2017). A partir de cada conformação espacial dos ligantes, são obtidas energias livres de ligação (proteína-ligante), em que a menor energia é

considerada a mais provável para justificar a conformação da interação (KITCHEN, 2004).

Os estudos de *docking* voltados para a agricultura, visam identificar agroquímicos específicos, capazes de atuar diretamente no organismo do patógeno, sem prejudicar o ambiente, como os polinizadores e o próprio homem.

REFERÊNCIAS BIBLIOGRÁFICAS

AIYELAAGBE, O.; ADESOGAN, K.; EKUNDAYO, O.; GLOER, J. Antibacterial diterpenoids from Jatropha podagrica Hook. **Phytochemistry.** v. 68 p. 2420-2425, 2007.

ALMEIDA, A.Q.; SILVA, S.A.; ALMEIDA, V.O.; SOUZA, D.R.; ARAÚJO, G.M. Genetic divergence and morpho-agronomic performance of *Jatropha curcas* L. clones for selection of clonal varieties. **Revista Caatinga**, v. *29*, p. 841-849, 2016.

ALVES, A.A.; LAVIOLA, B.G.; FORMIGHIERI, E.F.; CARELS, N. Perennial plants for biofuel production: bridging genomics and field research. **Biotechnology Journal**, v. 10, p. 505-507, 2015.

ATTWOOD, T.K.; GISEL, A.; ERIKSSON, N-E.; BONGCAM-RUDLOFF, E. Concepts, historical milestones and the central place of bioinformatics in modern biology: a European perspective. In: MAHDAVI, M.A. editor. **Bioinformatics-Trends and Methodologies.** Rijeka: Intech Online Publishers; 2011.

BADER, G. D.; HOGUE, C. W. V. Analyzing yeast protein–protein interaction data obtained from different sources. **Nature Biotechnology,** New York, v. 20, p. 991–997, 2002.

BARBIERI, L., BATTELLI, M. G., AND STRIPE, F. Ribosome inactivating protein from plants. **Biochimica et Biophysica Acta**. v. 1154, p. 237–282. 1993.

BERMAN, H.M.; BATTISTUZ, T.; BHAT, T.N.; BLUHM, W.F.; BOURNE, P.E.; BURKHARDT, K.; FENG, Z.; GILLILAND, G.L.; IYPE, L.; JAIN, S.; FAGAN, P.; MARVIN, J.; PADILLA, D.; RAVICHANDRAN, V.; SCHNEIDER, B.; THANKI, N.; WEISSIG, H.; WESTBROOK, J.D.; ZARDECKI, C. The Protein Data Bank. **Acta Crystallographica D. Biology**, v. 58, p. 899-907, 2002.

BIASINI, M.; BIENERT, S.; WATERHOUSE, A.; ARNOLD, K.; STUDER, G.; SCHMIDT, T.; KIEFER, F.; GALLO CASSARINO, T.; BERTONI, M.; BORDOLI, L.; SCHWEDE, T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, v. 42, p. 252-258, 2014.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa n. 4 de 14 de janeiro de 2008. **Diário Oficial da União,** p.4, 2008.

BRASILEIRO, B.P.; SILVA, S.A.; SOUZA, D.R.; SANTOS, P.A.; OLIVEIRA, R.S.; LYRA, D.H. Genetic diversity and selection gain in the physic nut (*Jatropha curcas*). **Genetics and Molecular Research.** v.3, p. 2341-2350, 2013.

BRAUN, U., & COOK, R. T. A.. Taxonomic Manual of the Erysiphales (Powdery Mildews). **CBS Biodiversity Series**. v. 11(11), p. 1-707, 2012.

BUCH, D. R., ARANTES, A. B., & CAMPELO, P. M. Verificação da atividade cicatrizante do exudato de folhas de *Jatropha multifida* L. **Revista Brasileira de Farmacognosia** , v. 2, p. 142-145, 2008.

CALIXTO, P.H.M. Aspectos gerais sobre a modelagem comparativa de proteínas. **Ciência Equatorial**. v. 3, p. 1-8, 2013.

CARVALHO, D.S. **Comportamento genético de progênies de meio– irmãos de pinhão manso no recôncavo baiano, Brasil**. 2010. 51f. (Mestrado em Ciências Agrárias) Universidade Federal do Recôncavo da Bahia, 2010.

CHEN, V.B.; ARENDALL, W.B.; HEADD, J.J.; KEEDY, D.A.; IMMORMINO, R.M.; KAPRAL, G.J.; MURRAY, L.W.; RICHARDSON, J.S.; RICHARDSON, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. **Acta Crystallogr Sect D Biol Crystallogr.** v. 66, p. 12-21, 2010.

CHOTHIA, C., & LESK, A. M. The relation between the divergence of sequence and structure in proteins. **The EMBO journal,** v. *5,* p. 823–826, 1986.

DALLAL, J. A.; IRVIN, J. D. Enzymatic inactivation of eukaryotic ribosomes by the pokeweed antiviral protein. **FEBS Letters**, v. 89, p. 257–259, 1978.

DEAN, R., VAN KAN, J.A.L., PRETORIUS, Z.A., HAMMOND-KOSACK, K.E., DI PIETRO, A., SPANU, P.D., RUDD, J.J., DICKMAN, M., KAHMANN, R., ELLIS, J. and FOSTER, G.D. The Top 10 fungal pathogens in molecular plant pathology. **Molecular Plant Pathology**, v. 13, p. 414-430, 2012.

DIAS, L. A. S., LEME, L. P., LAVIOLA, B. G., PALLINI, A., PEREIRA, O. L., DIAS, D. C. F. S., CARVALHO, M., MANFIO, SANTOS., SOUSA, L. C. A., OLIVEIRA, T. S., & PRETTI, L. A. Cultivo de pinhão manso (Jatropha curcas L.) para produção de óleo combustível. Viçosa, MG: **Editora UFV**. 2007.

DOMASHEVSKIY, A. V.; GOSS, D. J. Pokeweed antiviral protein, a ribosome inactivating protein: activity, inhibition and prospects. **Toxins**, v. 7, p. 274–298 2015

DURÃES, F. O. M., LAVIOLA, B. G., & ALVES, A. A. Potential and challenges in making physic nut (*Jatropha curcas* L.) a viable biofuel crop: the Brazilian perspective. **CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources**, v. 6, p. 43-60. 2011.

DURÃES, F.O.M, LAVIOLA, B.G., SUNDFELD, E., MENDONÇA, M., BHERING, L.L. Pesquisa, desenvolvimento e inovação em pinhão-manso para produção de biocombustíveis – Brasília, DF: **Embrapa Agroenergia**, v. 17 p. 2177-4439; 2009.

ENDO Y, MITSUI K, MOTIZUKI M, TSURUGI K. The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. **J Biol Chem**, v. 262, p. 5908-5912, 1987.

ENDO, Y.; TSURUGI, K. The RNA N-glycosidase activity of ricin A chain: mechanism of action of the toxic lectin ou eukaryotic ribosomes. **The Journal of Biological Chemistry**, v.263, p. 8735-8739, 1988.

ESPINDOLA F. C., CALÁBRIA, L. C. K.; REZENDE, A. A. A.; Boscolli Barbosa PEREIRA, B. B.; SANTANA, F. A.; AMARAL, I. M. R.; LOBATO, J.; FRANÇA, J. L.; MARIO, J. J.; FIGUEIREDO, L. B.; SANTOS-LOPES, L. P GOUVEIA, N. M.; NASCIMENTO, R.; TEIXEIRA, R. R.; REIS, T. A.; ARAÚJO, T. G. Recursos de bioinformática aplicados às ciências ômicas como genômica, transcriptômica, proteômica, interatômica e metabolômica. **Bioscience Journal:** Uberlândia, v. 26, n. 3, p. 463-477, 2010.

FARIAS, L.F**. Biologia floral, reprodutiva, visitantes florais e desempenho morfoagrônico de *Jatropha curcas* L. (Euphorbiaceae).** 2018. 94f. (Mestrado em Ciências Agrárias) Universidade Federal do Recôncavo da Bahia, 2018.

FERRARI, R. A.; CASARINI, M. B.; MARQUES, D. A.; SIQUEIRA, . J., Avaliação da composição química e de constituinte tóxico em acessos de pinhão-manso de diferentes origens. **Brasilian Journal of Food Technology**, v. 12, n. 4, p. 309-314, 2010.

FREIRE, F.C.O.; PARENTE, G.B. As doenças das Jatrophas (*Jatropha curcas L. e J. podagrica Hook*.) no Estado do Ceará. Fortaleza: **Embrapa Agroindústria Tropical**, 2006. 4p. (Embrapa Agroindústria Tropical. Comunicado Técnico, 120).

FREITAS, R.; MISSIO, R.; MATOS, F.; RESENDE, L.; DIAS, L. Genetic evaluation of Jatropha curcas: an important oilseed for biodiesel production. **Genetics and Molecular Research**, v.10, p. 1490-1498, 2011.

FRENKEL, Z. M.; TRIFONOV, E. N. Walking through the protein sequence space: towards new generation of the homology modeling. **Proteins**, v. 67, n. 2, p. 271-84, Mai., 2007.

FURUTANI, M.; KASHIWAGI, K.; ITO, K.; ENDO,Y.; IGARASHI, K. Comparison of the modes of action of a vero toxin (a Shiga-like toxin) from Escherichia coli, of ricin, and of α-sarcin, **Archives of Biochemistry and Biophysics**, v. 293, p. 140-146, 1992.

GIRBÉS, T., FERRERAS, J. M., ARIAS, F. J., AND STIRPE, F. Description, distribution, activity and phylogenetic relationship of ribosome-inactivating proteins

in plants, fungi and bacteria. **Mini-Reviews in Medicinal Chemistry**.v. 4, p. 461–476, 2004.

GONÇALVES, F.J.T.; FREIRE, F.C.O.; LIMA, J.S.; MELO, J.G.M.; CÂMARA, M.P.S. Patogenicidade de espécies de Botryosphaeriaceae endofíticas de plantas da Caatinga do estado do Ceará em manga e umbu-cajá. **Summa Phytopathologica**, v.42, n.1, p.43-52, 2016.

GUBITZ, G. M.; MITTELBACH, M.; TRABI, M. Exploitation of the tropical oil seed plant Jatropha curcas L. **Bioresource Technology**, v. 67, p. 73-82, 1999.

GUEX, N.; PEITSCH, M.C.; SCHWEDE, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. **Electrophoresis.** v. 30, p. 162-S173 2009.

HARTLEY, M. R.; LORD, J. M. Cytotoxic ribosome-inactivating lectins from plants. **Biochimica et Biophysica Acta,** v. 1701, p. 1-14, 2004.

HOGEWEG, P. The Roots of Bioinformatics in Theoretical Biology. **PLOS Computational Biology.** v. 7, e1002021, 2011.

HUSAIN, J., TICKLE, I. J., AND WOOD, S. P. Crystal structure of momordin, a type I ribosome inactivating protein from the seeds of Momordica charantia. **FEBS Letters**. v. 342, p. 154–158, 1994.

IGLESIAS, R.; LUCÍA CITORES, L.; RAGUCCI, S.; RUSSO, R.; MARO, A.D.; FERRERAS, J.M. Biological and antipathogenic activities of ribosome-inactivating proteins from *Phytolacca dioica* L., **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1860, p. 1256-1264, 2016.

IGLESIAS, R.; PÉREZ, Y.; TORRE, C.; FERRERAS, J.M.; ANTOLÍN, P.; JIMÉNEZ, P.; ROJO, M.A.; MÉNDEZ, E.; GIRBÉS, T. Molecular characterization and systemic induction of single-chain ribosome-inactivating proteins (RIPs) in sugar beet (*Beta vulgaris*) leaves, **Journal of Experimental Botany**, v. 56, p. 1675–1684, 2005.

ISHAG, H.Z.; LI, C.; HUANG, L.; SUN, M.X.; NI, B.; GUO, C.X.; MAO, X. Inhibition of Japanese encephalitis virus infection in vitro and in vivo by pokeweed antiviral protein. **Virus Research**. v. 171, p. 89-96, 2013

JIANG, S. Y., RAMAMOORTHY, R., BHALLA, R., LUAN, H. F., VENKATESH, P. N., CAI, M.; RAMACHANDRAN, S. Genome-wide survey of the RIP domain family in Oryza sativa and their expression profiles under various abiotic and biotic stresses. **Plant Molecular Biology.** v. 67, p. 603–614, 2008.

KAY, L.E. NMR studies of protein structure and dynamics**. Journal of Magnetic Resonance**, v. 173, p. 193-207, 2005.

KEMMEREN, P. et al. Protein interaction verification and functional annotation by integrated analysis of genome-scale data. **Molecular Cell**, St. Louis, v. 9, p.1133–1143, 2002.

KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; BAJORATH, J. Docking and scoring in virtual screening for drug discovery: methods and applications. **Nature Reviews**, v. 3, p. 935-949, 2004.

KLEIWEIGT, G.J.; JONES, T.A. Phi/Psi-chology: Ramachandran revisited. **Ways & Means**. v. 4, p. 1395-1400, 1996.

KLEPEIS, J.L.; WEI, Y.; HECHT, M.H.; FLOUDAS, C.A. Ab initio prediction of the three-dimensional structure of a de novo designed protein: a double-blind case study. **Proteins**. v. 58, p. 560-570. 2005.

KOPP, J.; SCHWEDE, T. The SWISSMODEL Repository of annotated threedimensional protein structure homology models. **Nucleic Acids Research**, v. 32, p. 230-4, 2004.

KRISHNAN, P. R.; PARAMATHMA, M. Potentials and Jatropha species wealth of India. **Current Science**. v. 97. p. 1000-1004, 2009.

KWON, J.H.; CHOI, O.; KIM, J.; KWAK, Y.S. First Report of Anthracnose Disease on *Jatropha curcas* Caused by *Colletotrichum gloeosporioides* in Korea. **Journal of Phytopathology**, v. 160, p. 255-257, 2012.

LANZANOVA, C.; TORRI, A.; MOTTO, M.; BALCONI, C. Characterization of the maize b-32 ribosome inactivating protein and its interaction with fungal pathogen development. **Maydica Journals**, v. 56, p. 1-11, 2011.

LATHA, P.; PRAKASAM, V.; KAMALAKANNAN, A.; GOPALAKRISHNAN, C.; RAGUCHANDER, T.; PARAMATHMA, M.; SAMIYAPPAN, R. First report of Lasiodiplodia theobromae (Pat.) Griffon & Maubl causing root rot and collar rot disease of physic nut (*Jatropha curcas* L.) in India. **Australas Plant Dis,** v. 4, p. 19–20, 2009.

LAVIOLA, B. G., SILVA, S. D. A., JUHÁSZ, A. C. P., ROCHA, R. B., OLIVEIRA, R. J. B., ALBRECHT, J. C., ROSADO, T. B. Desempenho agronômico e ganho genético pela seleção de pinhão-manso em três regiões do Brasil. **Pesquisa Agropecuária Brasileira**, v. 49, p. 356-363**,** 2014.

LAVIOLA, B.G.; ROCHA, R.B.; KOBAYASHI, A.K.; ROSADO, T.B.; BHERING, L.L. Genetic improvement of *Jatropha* for biodiesel production. **Ceiba**, v. 51, p. 1-10, 2012.

LAVIOLA, B.G.; RODRIGUES, E.V.; TEODORO, P.E.; PEIXOTO, L.A.; BHERING, L.L. Biometric and biotechnology strategies in Jatropha genetic breeding for biodiesel production. **Renewable and Sustainable Energy Reviews**, v. 76, p. 894-904, 2017.

MACHADO, A.R.; PEREIRA, O.L. Major Diseases of the Biofuel Plant, Physic Nut (Jatropha curcas), Biodiesel - Feedstocks, Production and Applications, Zhen Fang, **IntechOpen**, 2012

MACHADO, A.R.; PINHO, D.B.; PEREIRA, O.L. Phylogeny, identification and pathogenicity of the Botryosphaeriaceae associated with collar and root rot of the biofuel plant Jatropha curcas in Brazil, with a description of new species of Lasiodiplodia. **Fungal Diversity**, v. 67, p. 231–247, 2014.

MAKKAR, H. P. S.; ADERIBIGDE, A. O.; BECKER, K. Comparative evaluation of non-toxic varieties of Jatropha curcas for chemical composition, digestibility, protein degradability and toxic factors. **Food Chemistry,** v. 62, p. 207-215, 1998

MAKKAR, H.P.S.; FRANCIS, G.; BECKER, K. Protein concentrate from *Jatropha curcas* screw-pressed seed cake and toxic and antinutritional factors in protein concentrate. **Journal of the Science of Food and Agriculture**, v.88, p.42-1548, 2008.

MARTINS, C, C.; MACHADO, C. G.; CAVASINI, R., Temperatura e substrato para o teste de germinação de sementes de pinhão-manso, **Ciência agrotécnica**, v. 32, 2008.

MELTON-CELSA, A. R. Shiga toxin (Stx) classification, structure, and function. **Microbiology Spectrum**. v. 9, p. 1–13, 2014.

MENDONÇA, S.; LAVIOLA, B. G. Uso potencial e toxidez da torta de pinhão-manso. Brasília: **Embrapa Agroenergia**, Comunicado Técnico, n. 1. p. 8, 2009.

MLSNA, D., MONZINGO, A. F., KATZIN, B. J., ERNST, S., AND ROBERTUS, J. D. Structure of recombinant ricin A chain at 2.3 Å**. Protein Science.** v. 2, p. 429–435, 1993.

MONTES, J.M.; MELCHINGER, A.E. Domestication and Breeding of Jatropha curcas L. **Trends in Plant Science.** v.21, p. 1045-1057, 2016

MONZINGO, A. F., AND ROBERTUS, J. D. X-ray analysis of substrate analogs in the ricin A-chain active site. **Journal of Molecular Biology**. v. 227, p. 1136–1145. 1992.

MOULT J, FIDELIS K, KRYSHTAFOVYCH A, TRAMONTANO A. Critical assessment of methods of protein structure prediction (CASP) round IX. **Proteins**, v. 79, p. 1-5, 2011.

NAKANO, S.; TAKAHASHI, M.; SAKAMOTO, A.; MORIKAWA, H.; KATAYANAGI, K. X-ray crystal structure of a mutant assimilatory nitrite reductase that shows sulfite reductase-like activity.**Chemistry & Biodiversity,** v. 9, p. 1989-1999, 2012.

NARAYANA D.S.A., SHANKARAPPA K.S., GOVINDAPPA M.R., PRAMEELA H.A., RAO M.R.G., RANGASWAMY K.T. Natural occurrence of Jatropha mosaic virus disease in India. **Current Science** , v. 91, p. 584-586, 2006.

NELLER, K.C.M.; DIAZ, C.A.; PLATTS, A.E.; HUDAK, K.A. *De novo* Assembly of the Pokeweed Genome Provides Insight Into Pokeweed Antiviral Protein (PAP) Gene Expression**. Frontiers in Plant Science,** v. 10, p. 1002, 2019

OJO, O.O.; OMABE, M. Incorporating bioinformatics into biological science education in Nigeria: prospects and challenges. **Infection, Genetics and Evolution.** v. 11, p. 784-787, 2011.

PAGADALA, N.S.; SYED, K.; TUSZYNSKI, J. Software for molecular docking: a review. **Biophysical Reviews**, v. 9, p. 91-102, 2017

PEREIRA, O. L.; DUTRA, D. C.; DIAS, L. A. S. Lasiodiplodia theobromae is the causal agent of a damaging root and collar rot disease on the biofuel plant Jatropha curcas in Brazil. **Australasian Plant Disease Notes**, v. 4, p. 120-123, 2009.

PESTANA-CALDAS, C.N.; SILVA, S.A.; MACHADO, E.L; DE SOUZA, D.R; CERQUEIRA-PEREIRA, E.C.; SILVA, M.S. Genetic divergence through joint analysis of morphoagronomic and molecular characters in accessions of *Jatropha curcas*. **Genetics and Molecular Research.** v. 4, p. 1-11, 2016.

PICARD, D., KAO, C. C., AND HUDAK, K. A. Pokeweed antiviral protein inhibits brome mosaic virus replication in plant cells. *J.* **Journal of Biological Chemistry**.v. 280, p. 20069–20075, 2005.

PRUSKY, D., LICHTER, A. Activation of quiescent infections by postharvest pathogens during transition from the biotrophic to the necrotrophic stage. **FEMS Microbiology Letters**, v. 268, p. 1-8, 2007.

QIN, X., ZHENG, X., SHAO, C., GAO, J., JIANG, L., ZHU, X.; YAN, F.; TANG, L.; XU, Y.; CHEN, F. Stress-induced curcin-L promoter in leaves of *Jatropha curcas* L. and characterization in *transgenic tobacco*. **Planta.** v. 230, p. 387–395, 2009.

QIN, X.; SHAO, C.; HOU, P.; GAO, J.; LEI, N.; JIANG, L.; YE, S.; GOU, C.; LUO, S.; ZHENG, X.; GU, X.; ZHU, X.; XU, Y.; CHEN, F. Z. Different functions and expression profiles of curcin and curcin-L in Jatropha curcas. **Zeitschrift für Naturforschung C - De Gruyter,** v. 65, p. 355-62, 2010.

QUEIROZ, D.C. **Avaliação de progênies de meio-irmãos e identificação de clones promissores de pinhão-manso por meio de caracteres morfoagronômicos e moleculares**. 2015. 116f. (Doutorado em Ciências Agrárias) Universidade Federal do Recôncavo da Bahia, 2015.

RAKSHIT, K.D.; DARUKESHWARA, J.; RAJ, K.R. Toxicity stud-ies of detoxified *Jatropha meal* (*Jatropha curcas*) in rats. **Food and Chemical Toxicology**, v.46, p.3621-3625, 2008.

RAMACHANDRAN, G.N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal Molecular Biology.** v. 7, p. 95-99, 1963.

REINBOTHE, S.; REINBOTHE, C.; LEHMANN, J.; BECKER, W.; APEL, K.; AND PARTHIER, B. JIP60, a methyl jasmonate-induced ribosome-inactivating protein involved in plant stress reactions. **Proceedings of the National Academy of Sciences***.* v.91, p. 7012–7016, 1994

RODRIGUES, R. P.; MANTOANI, S. P.; DE ALMEIDA, J. R.; PINSETTA, F. R.; SEMIGHINI, E. P.; DA SILVA, V. B.; DA SILVA, C. H. P., Estratégias de Triagem Virtual no Planejamento de Fármacos, **Revista Virtual de Química**, v. 4, n. 6, p.739- 736, 2012.

SANTOS FILHO, O.A.; ALENCASTRO, R.B. Modelagem de proteínas por homologia. **Química Nova**, v. 26, p. 253-259, 2003.

SANTOS, G.R.; TOZZE JÚNIOR, H.J.; SÁ, D.A.C.; FURTADO, G.Q.; MASSOLA JÚNIOR, N.S. Etiology and pathogenicity of two different isolates of *Colletotrichum* spp. obtained from physic nut seeds. **Journal of Seed Science**, v. *35*, p. 139-146, 2013.

SANTOS-FILHO, O.A.; ALENCASTRO, R.B.; FIGUEROA-VILLAR, J.D. Homology modeling of wild type and pyrimethamine/cycloguanil-cross resistant mutant type Plasmodium falciparum dihydrofolate reductase. A model for antimalarial chemotherapy resistance. **Biophys Chem.** v. 91, p. 305-3, 2001.

SHANG, C.; ROUGÉ, P.; VAN DAMME, E. J. M. Ribosome inactivating proteins from rosacea. **Molecules**, v. 21, p. 2-17, 2016.

SILVA, S.A.; SANTOS, L.A. . Melhoramento Genético de Oleaginosas. In: CHAVES, L.H.G.; FRANCISCO, P.R.M.; VASCONCELOS, A.C.F. (Org.). **Oleaginosas e Hortaliças sob diferentes Manejo de Cultivo: Coletânea de Estudos.** 1ed.Campina Grande: EPGRAF, v. 1, p. 78-85, 2019.

SMITH, H.; WINGFIELD, M.J.; CROUS, P.W.; COUTINHO, T.A. *Sphaeropsis sapinea* and *Botryosphaeria dothidea* endophytic in *Pinus* spp. and *Eucalyptus* spp. in South Africa. **South African Journal of Botany**, v. 62, p. 86-88, 1996.

SPINELLI, V.M.; ROCHA, R.B.; RAMALHO, A.R.; MARCOLAN, A.L.; VIEIRA JÚNIOR, J.R.; FERNANDES, C.F.; MILITÃO, J.S.L.T.; DIAS, L.A.S. Componentes primários e secundários do rendimento de óleo de pinhão-manso. **Ciência Rural,** v.8, p. 1752-1758, 2010.

STRANGE, R. N., SCOTT, P.R. Plant Disease: A Threat to Global Food Security. **Annual Review of Phytopathology**, v. 43, p. 83-116, 2005.

TARTARINI, A., PITTALUGA, E., MARCOZZI, G., TESTONE, G., RODRIGUES-POUSADA, R. A., GIANNINO, D., SPANÒ, L. Differential expression of saporin genes upon wounding, ABA treatment and leaf development. **Physiologia Plantarum***.* v. 140, p. 141–152, 2010.

TAYLOR, A.; HARDY, G.E.J.; WOOD, P.; BURGESS, T. Identification and pathogenicity of Botryosphaeria species associated with grapevine decline in Western Australia. **Australasian Plant Pathology**, v. 34, p. 187-195, 2005.

TORRES-CALZADA, C.; TAPIA-TUSSELL, R,; NEXTICAPAN-GARCEZ, A.; MATIN-MEX, R.; QUIJANO-RAMAYO, A.; CORTÉS-VELÁZQUEZ, A.; HIGUERA-CIAPARA, I.;PEREZ-BRITO, D. First report of Colletotrichum capsici causing anthracnose in *Jatropha curcas* in Yucatan, Mexico**. New Disease Reports**, v. 23, p. 6-8. 2011.

TUCCINARDI, T.; BOTTA, M.; GIORDANO, A. MARTINELLI, A. Protein kinases: docking and homology modeling reliability**. Journal of Chemical Information and Modeling**. v. 50, p. 1432-1441, 2010.

VELOSO, J.S.; CÂMARA, M.P.; LIMA, W.G.; MICHEREF, S.J.; DOYLE, V.P. Why species delimitation matters for fungal ecology: Colletotrichum diversity on wild and cultivated cashew in Brazil. **Fungal Biology**. v. 122, p. 677–691, 2018.

VIRGÍLIO, M.; LOMBARDI, A.; CALIANDRO, R.; FABBRINI, M.S. Ribosome-inactivating proteins: from plant defense to tumor attack. **Toxins (Basel)**; v. 2, p. 2699-2737, 2010.

VYAS, V.K.; UKAWALA, R.D.; GHATE, M.; CHINTHA, C. Homology modeling a fast tool for drug discovery: current perspectives. **Indian Journal of Pharmaceutical Sciences**, v. 74, p. 1-17, 2012.

WALSH, T. A.; MORGAN, A. E.; HEY, T. D. Characterization and molecular cloning of a proenzyme form of a ribosome-inactivating protein from maize. Novel mechanism of proenzyme activation by proteolytic removal of a 2.8-kiloDalton internal peptide segment. **The Journal of Biological Chemistry,** v. 266, p. 23422–23427, 1991.

WANG, S.; LI, Z.; LI, S.; DI, R.; HO, C.-T.; YANG, G. Ribosome-inactivating proteins (RIPs) and their important health promoting property. **RSC Advances**, v.6, p. 46794–46805, 2016.

WANI, T. A.; KITCHLU, S.; RAM, G. Genetic variability studies for morphological and qualitative attributes among *Jatropha curcas* L. accessions grown under subtropical conditions of North India. **South African Journal of Botany**, v. 79, p. 102-105, 2012.

WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; STUDER, G.; TAURIELLO, G.; GUMIENNY, R.; HEER, F.T.; DE BEER, T.A.P.; REMPFER, C.; BORDOLI, L.; LEPORE, R.; SCHWEDE, T. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**. v. 46, p. 296-303, 2018.

WIEDERSTEIN, M.; SIPPL, M.J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research.** v. 35, p. 407-410, 2007.

WONG, W.C.; MAURER-STROH, S.; EISENHABER, F. Not all transmembrane helices are born equal: Towards the extension of the sequence homology concept to membrane proteins. **Biology Direct**, v. 6, p. 57, 2011.

ZHU, F.; ZHOU, Y.K.; JI, Z.L.; CHEN X.R. The Plant Ribosome-Inactivating Proteins Play Important Roles in Defense against Pathogens and Insect Pest Attacks. **Frontiers in Plant Science**. v. 9, p. 146, 2018.

CAPÍTULO 1

IN SILICO CHARACTERIZATION OF GENIC FAMILY OF RIBOSOME INACTIVATING PROTEINS (RIPs) OF PHYSIC NUT (*Jatropha curcas* L.)

In silico characterization of the genic family of ribosome inactivating proteins (RIPs) of physic nut (Jatropha curcas L.)

Abstract: RIPs can be found in various types of plants, fungi and bacteria. And they are part of a class of proteins involved in the plant's defense system. The present study aimed to characterize the family of RIP genes in *Jatropha curcas* L, to build the homology modeling of all described isoforms. In addition, molecular docking of two type I RIPs, one type II RIP and ricin was performed using different inhibitors. Existing annotation JCDB allowed for the identification of a total of twelve RIP genes. The JcRIP1 to JcRIP11 genes encode curcin, a type I RIP found in physic nut, while JcRIP12 is a gene similar to the genes that code for ricin, a type II RIP present in *Ricinus communnis* L. Cis-elements light-responsive, MYB and MYC were identified in all promoter regions of the JcRIPs genes. In smaller proportions, cis-elements were found that were responsive to stresses (biotic and abiotic) and hormones. The JcRIP3 gene is one with the highest expression in the tissues. In addition to seeds and leaves, there is also expression of RIP genes in stems, roots and flowers. The number of amino acid residues of physic nut RIP proteins ranged from 291 (JcRIP8) to 551 (JcRIP12). All proteins were predicted to follow secretory pathways and are located in extracellular environment. Eight of 12 JcRIPs have EAAR motif (Glu185-Ala186-Ala187-Arg188) well maintained. The YFF motif (Tyr99-Phe100-Phe101) was conserved in all JcRIPs studied here. All proteins of JcRIP family were modeled and showed good quality. Folic and pteroic acids were used as ligands because they are compounds already known as inhibitors of other RIPs, including curcin. In addition, dexamethasone and its derivatives were tested. The high conservation existing between residues of the JcRIP proteins with the ricin (5DDZ) allows dexamethasone and its derivatives to bind to all these proteins with affinity energy high, it indicates a potential application biotechnology of these compounds in blocking action of ribosome inactivating proteins.

Keywords: Bioinformatics, Proteomic, RIPs.

Introduction

Ribosome Inactivating Proteins (RIPs) are encoded by a family of nuclear genes and makes up a group of cytotoxin proteins that usually contain a RNA N-glycosidase domain, which irreversibly inactivate ribosome, thus inhibiting protein synthesis (LIU, 2017). RIPs can be found in several types of plants, fungi and bacteria (SHANG; ROUGÉ; VAN DAMME, 2016; MELTON-CELSA, 2014), are classified into: i) Type I, more widely distributed and composed of a single polypeptide chain of approximately 30kDa, identified initially in *Phytolacca americana* L.; ii) Type II, have two polypeptide chains, with chain A similar to type I, and the B chain which is slightly larger and has the lectin domain, both chains are linked by a disulfide bridge, the ricin protein is the best known among RIP type II and iii) Type III, are proteins that require proteolytic cleavage to become active, and are considered pro-RIPs (DALLAL e IRVIN, 1978; HARTLEY, LORD, 2004; WALSH; MORGAN; HEY, 1991). The different types of RIP are localized in plant leaves, seeds, flower, stem or roots. However, single-chain type I RIPs are significantly more common than their type II relatives (STIRPE, 2004).

Despite the fact that RIPs are abundant in the plant kingdom and are present in high levels in seeds of species of the families Caryophyllaceae, Cucurbitaceae, Euphorbiaceae and Phytolaccaceae (PURI et al., 2012), their role in nature is not yet completely understood, but studies report that RIP proteins are involved in plant defense, possessing antiviral, antifungal, insecticidal activity, acting in protection against stresses (biotic and abiotic) and in development regulation functions, also presenting functions storage and programmed senescence (GIRBES et al., 2004; ZHU et al., 2018). Moreover, diverse pharmacological attributes have been associated with RIPs, as immunosuppressive, anti-tumor, protein synthesis-inhibitory and anti-HIV (NG; CHAN; YEUNG, 1992).

J. curcas L. is an oleaginous species belonging to the Euphorbiaceae Family, well known for its multiple uses and medicinal value (THOMAS; SAH; SHARMA, 2008; PANDEY et al., 2012; DEBNATH, 2008), it has type I and II ribosome inactive proteins (RIPs) in its genome, called curcin. Felke was the first to isolated en 1914 the curcin toxic protein, from the seeds from *J.curcas* L.(FELKE, 1914). Recent work have demonstrated that curcin has antitumor activity, eliminating the core of the

cancer mass, avoiding the chance of future proliferation and recurrence, being an option cytocidal phytomolecule, able to curb the progression of cancer-metastasis (MOHAMED, et al., 2014).

Knowledge regarding the structure and function of plant RIP proteins has potentially useful applications in agriculture and medicine. Although the *J.curcas* L. species has gained prominence in recent years and several studies about this culture, the genome of the species has not yet been used to characterize the complete RIP gene family. The availability of data from the recent sequencing of the genome of *J. curcas* (http://jcdb.xtbg.ac.cn/) allowed for the genome-wide analysis of gene families as a pre-requisite for functional and/or pre-breeding studies. The present study aimed to characterize the family of *RIP* genes in *Jatropha curcas* at nucleotide and protein levels, including promoter analysis. The study also aimed to construct the homology modeling of all described isoforms. Moreover, the molecular docking of two type I RIPs, one type II RIP and ricin was performed using different inhibitors.

Material and methods

In silico analysis of RIPs genes and proteins

The identification and structural analysis of the RIPs genes were performed using the JCDB: *Jatropha Curcas* Database (http://jcdb.xtbg.ac.cn/). The conserved RIP domain and family protein were analyzed using the Pfam (EL-GEBALI et al., 2019). The predictions of the subcellular location of the protein was performed by the DeepLoc (ARMENTEROS et al., 2017). The prediction of the theoretical isoelectric point (pI), molecular weight (MW) and hidropathicity level (*Grand average of hydropathicity* – GRAVY) of the proteins were obtained using the program ProtParam (GASTEIGER et al., 2005). SignalP was used to identify the signal peptide. The NetPhos 3.1 Server (BLOM et al., 1999) and the NetNGlyc 1.0 Server (GUPTA, JUNG, BRUNAK, 2004) were used to identify putative phosphorylation sites (Ser/Thr/Tyr) and putative N-glycosylation sites (Asn-X-Ser/Thr type), respectively. We consider only N-glycosylation, which can affect the asparagine residue in the asparagine – X – serine / threonine sequence, where X can be any amino acid, except proline. Standard input parameters were used for all servers.

Analysis of the promoter regions and chromosomal locations of *JcRIP* genes

To identify the presence of the *cis*-regulatory elements in the promoter regions of the *JcRIP* genes, the 1500 bp upstream region from the translation start site of the genes was analyzed using the plantCARE software (LESCOT et al., 2002). The chromosomal locations of the *JcRIP* genes, as well as the number of introns and exons were obtained by screening the GFF3 deposited in the JCDB.

In silico *J. curcas* L. RIPs gene expression

The expression of the genes of the *J. curcas* L. RIPs gene in different tissues (leaf, flower, stem, root and fruit) was analyzed using conventional RNA-Seq data available on the JCDBTools Heatmap that employs the data Zhang et al., (2019). Plants were used in control conditions, without being subjected to stress, thus simulating expression in normal conditions.

Alignment and Molecular modeling

The alignment of JcRIPs proteins with ricin A chain (PDB ID: 5DDZ_A), resolved and deposited in PDB (FAN et al., 2016), was carried out with COBALT server (www.ncbi.nlm.nih.gov/tools/cobalt), in which the alignment is aided by a collection of restrictions of pairs derived from conserved domain database, protein motif database and local sequence similarity using RPS-BLAST, BLASTP and PHI-BLAST, respectively (PAPADOPOULOS; AGARWALA, 2007).

To select the best 3-D templates for JcRIPs molecular modeling, the 12 proteins studied here were aligned with the Protein Data Bank (Pdb) using the PSIBLAST program (LASKOWSKI et al., 2005). Target 3-D structures were modeled using templates that presented the highest identity and coverage, starting from a minimum of 25% of identity amino acids in the alignment. Additionally, the minimum template resolution considered was 2.0 Å. The modeling was performed with a type 1 JcRIP protein, which showed a better model, good identity and coverage with the proteins deposited in the PDB. In addition to this, the JcRIP12 protein, which is a type II RIP, was modeled, because it has the peculiarity of not having a tyrosine that comprises the main residue of the RIPs binding site and, in addition, there are no modeling reports for a type II RIP. The predicted 3-D protein model was obtained using the SWISS-MODEL server (ARNOLD et al., 2006). The

stereochemical quality of both JcRIP models was calculated by Procheck (LASKOWSKI et al., 1993) and ProSA (WIEDERSTEIN; SIPPL, 2007). ProSA was used for the display of Z-score and energy plots. The visualization of the 3D model was obtained using PyMol (The PyMOL Molecular Graphics System, Version 2.3.3, LLC).

Molecular Docking

The docking was performed with two type I JcRIP protein (JcRIP1 e JcRIP3) which showed a better model, good identity and coverage with the proteins deposited in the PDB and greater expression in seeds, which is the commercial use portion in *J. curcas* L.. In addition to this, the JcRIP12 protein, a type II RIP, was also used for docking because it has the peculiarity of not having a tyrosine that comprises the main residue of the RIPs active site.

Ligands already known for RIPs were used, such as folic acid and pteroic acid, in addition to these, corticosteroids (dexamethasone and derivatives) were also used. Dexamethasone is a known anti-inflammatory and used in treatments, associated with other medications, in several cases of poisoning and intoxications (GAL et al., 2014; PINTO et al, 2017; MANDAL et al, 2020).

Before preforming the molecular docking between the ligand and the target protein, the ligands structure was downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and ZINC (STERLING; IRWIN, 2015) database in SMILES format. The ligands structure was converted into 3-D format using MarvinSketch v. 15.7.13.0, ChemAxon (www.chemaxon.com/), and saved in mol2 format. The AutoDockTools V1.5.6 (SANNER, 1999) was used to prepare the protein and ligands for docking calculations. First, polar hydrogens were added to the ligands structure and all torsions were checked; the ligand structure was then saved in PDBQT format. The amino acids of the active site were marked in order to get the grid box coordinates for the docking process. Afterward, the JcRiPs structure was saved in PDBQT format. AutoDock Vina (TROTT & OLSON, 2010) was used to docking and obtain the values of affinity energy (kCal / Mol) for each conformation of each ligand in its respective complex. All docking results were evaluated using PyMOL V1.7.4 [26] in order to check which ligand presents the best docking affinity energy. Additionally, Discovery Studio 4.0 was used to generate the 2-D map of the

interaction between protein and ligands. Then, the ligand that best fits the active site of the protein was selected.

Results and discussion

RIP gene family in the physic nut genome

The JCDB bank was chosen to search for gene sequences (version JatCur_1.0), as it has more detailed information, such as the location of the gene in the genome, coding regions and locations of exons. Existing annotation JCDB allowed for the identification of a total of twelve RIP genes, named here as *JcRIP1*, *JcRIP2*, *JcRIP3*, *JcRIP4*, *JcRIP5*, *JcRIP6*, *JcRIP7*, *JcRIP8*, *JcRIP9*, *JcRIP10*, *JcRIP11* and *JcRIP12* (Table 1).

The *JcRIP1* to JcRIP11 genes encode curcin, a type I RIP found in physic nut, while *JcRIP12* is a gene similar to the genes that code for ricin, a type II RIP present in *R. communis* L. (LORD, ROBERTS and ROBERTUS, 1994). The number of RIPs genes identified in species diferent is very variable, being 38 in *Oryza sativa ssp. japonica*, 34 in *Oryza sativa L. ssp. indica*, 28 in *Ricinus communis*, 10 in *Phytolacca americana*, 8 in *Cucurbita moschata* and *Cucurbita pepo*, 7 in *Citrullus lanatus* and *Cucurbita maximum*, 6 in *Cucurbita melo*, 5 in *Cucumis sativus*, 9 in *Lagenaria siceraria* and 18 in *Momordica charantia*, 2 in *Prunus persica* and 4 in the species *Malus domestica*, *Prunus mume*, *Pyrus bretschneideri* and *Pyrus communis* (WYTYNCK; ROUGÉ; VAN DAMME, 2017; CHAN et al., 2010; NELLER, et al., 2019; KOLE, et al., 2020; SHANG; ROUGÉ; VAN DAMME, 2016). *JcRIP1*, *JcRIP2*, *JcRIP3*, *JcRIP4*, *JcRIP5*, *JcRIP6*, *JcRIP8*, *JcRIP9* and *JcRIP11* are similar to cucurmosin, another type I RIP present in *Cucurbita moschata* (HOU et al., 2008), only JcRIP7 and JcRIP10 are homologous to luffaculin, RIP found in *Luffa acutangula* (HOU et al., 2007). Due to similarity of some RIPs with cucurmosin and luffaculin, they are identified on the JCDB as cucurmosin-like, luffaculin-like and ricin-like (Table 1).

The *J. curcas* L. genome is still being assembled, this way it was not possible to identify which chromosomes harbor the RIPs genes. However, the genes were in different scaffolds, what can indicate that they are on different chromosomes.

Table 1. Characteristics of the RIP genes present in the genome of *Jatropha curcas* L, according *Jatropha Curcas* Database (JCDB). RIP: Ribosome Inactivating Protein.

Name	Gene ID	Location ("*Scaffold*")	Protein-like	Quantity Exons	Quantity Introns
JcRIP1	JCDBG01176/LOC105635821	NW_012124076.1:8868..10287	Cucurmosin	2	1
JcRIP2	JCDBG01177/LOC105635830	NW_012124076.1:17905..19015	Cucurmosin	1	0
JcRIP3	JCDBG01178/LOC105635840	NW_012124076.1:102058..103441	Cucurmosin	2	1
JcRIP4	JCDBG01180/LOC105635895	NW_012124076.1:170314..171243	Cucurmosin	1	0
JcRIP5	JCDBG10548/LOC105635637	NW_012124293.1:18326..24861	Cucurmosin	2	1
JcRIP6	JCDBG01181/LOC105635850	NW_012124076.1:184499..190105	Cucurmosin	3	2
JcRIP7	JCDBG12049/LOC105637073	NW_012124368.1:38137..39481	Luffaculin	2	1
JcRIP8	JCDBG12048/LOC105637072	NW_012124368.1:19156..20031	Cucurmosin	1	0
JcRIP9	JCDBG12053/LOC105637076	NW_012124368.1:98496..99440	Cucurmosin	1	0
JcRIP10	JCDBG12050/LOC105637074	NW_012124368.1:53728..54495	Luffaculin	1	0
JcRIP11	JCDBG12054/LOC105637077	NW_012124368.1:108722..109639	Cucurmosin	1	0
JcRIP12	JCDBG12052/LOC105637075	NW_012124368.1:82430..84085	Ricin	1	0

Seven of the 12 RIP genes have no introns, the genes JcRIP5 and JcRIP6 have a much larger intronic region than the others, due to a large ambiguous region (NNNNNNNNNN...) in its gene sequences (Fig. 1, S.1 Fig.). According to Bolognesi (2016), most type 1 RIPs, except for Poaceae, are synthesized from genes without introns, which encode precursors with N-terminal and C-terminal extensions. They are synthesized in the endoplasmic reticulum (ER) and then they will be secreted or directed to the vacuole. The lack of introns or the small number (1 to 2) of them indicates that RIP genes can be quickly transcribed, since a larger number of introns can result in delays in the production of proteins due to the processing steps of the transcript (JEFFARES; PENKETT; BÄHLER. 2008).

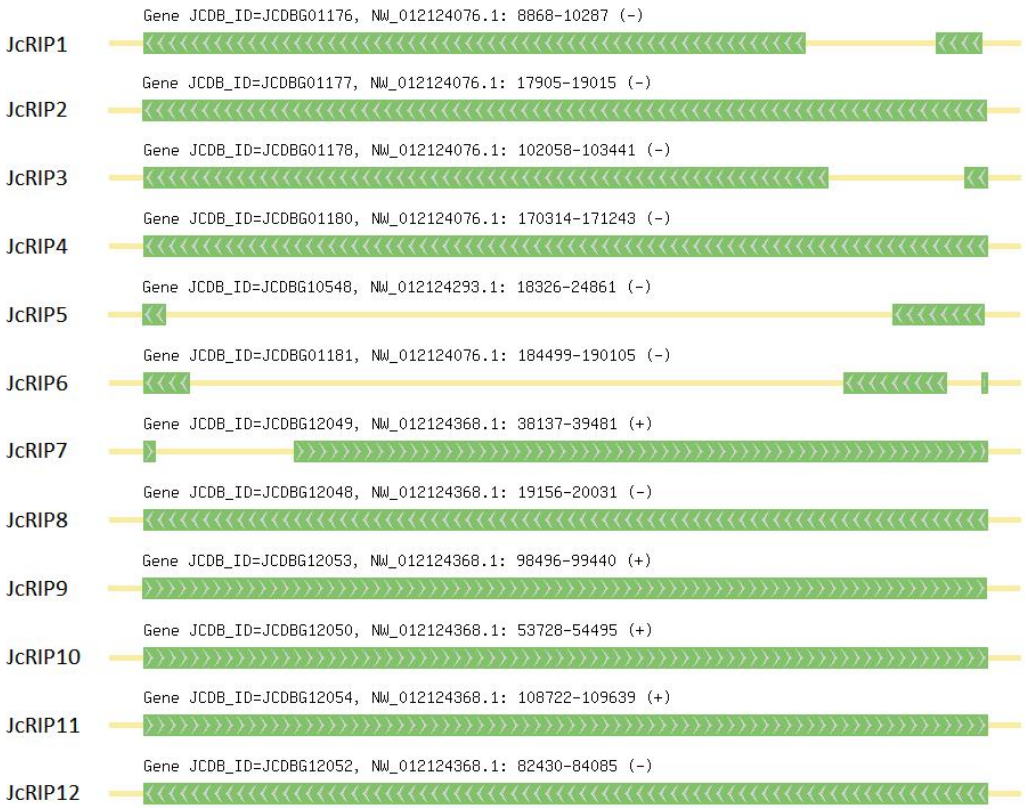


Figure 1. Structure of the *JcRIPs* genes. The structure of the exons is represented shown in the green box and introns in yellow lines.

Promoter sequence analysis of the physic nut *RIP* genes

A 1500bp fragment belonging to the upstream region of each *RIP* gene was analyzed to find plant-specific *cis*-elements using the PlantCARE database (S2 Fig.). The TATA and CAAT-box elements were found in all physic nut RIP promoter regions (S3 Fig); the other *cis*-elements varied between sequence promoters (Fig 2, S3 Fig.). Light-responsive cis-elements (G-Box, Box 4, ATC-motif, 3-AF1 binding site, MRE, GT1-motif, TCCC-motif, chs-CMA1a, AAGAA-motif, I-box, GA-box, AE-box, LAMP-element, chs-CMA2a, TCT- motif, ATCT- motif, ACE, GA-motif, GATA-motif) were abundant (5 to 17, according to the promoter) in all promoter sequences (Fig 1). Ribosome inactivating proteins (RIPs) participate in the strategies of defending plants against numerous potential invaders present in environments (XU; WANG; FAN, 2008). Light being an essential element for growth and a dominant factor in the response to stresses, many responses to biotic stress are specifically controlled by luminosity conditions (KARPINSKI, et al., 2003). Both in the response of plants to excess excitation energy (EEE) and infection by pathogens, similar

events occur: depletion of antioxidant pools, the chlorosis and necrosis of leaves, the rapid increase in the foliar concentrations of oxigen reactive species, the local and systemic defence response (SAA and SAR) and the induction of defence gene expression (KARPINSKI, et al., 2013).

Elements MYB (Myb, MYB, MYB-like sequence, MYB-binding site) and MYC (Myc, MYC) were also identified in all promoter regions of the JcRIPs genes, varying from 1 to 9 elements (Fig. 2, S3 Fig.). MYB and MYC are part of the main families of transcription factors that participate in the signaling network involved in the response of plants to environmental changes, as well as the signaling pathways of abscisic acid (ABA), being involved in the responses to different types of stress, as well as the development of the plant (SHESHADRI; NISHANTH; SIMON BINDU, 2016).

In smaller proportions, *cis*-elements were found that were responsive to *i*) biotic, abiotic or mechanical stresses (Fig. 2, S3 Fig.) such as drought (as-1 e MBS elements), low temperature (LTR element), anaerobiosis (ARE), wounds and pathogens (WRE3, Wun-motif, W box, Box S elements) and stress (STRE e TC-rich repeats elements); and *ii*) hormones or inducers such as methyl jasmonate (CGTCA-motif and TGACG-motif elements), gibberellin (GARE-motif, P-box and TATC-box elements), ethylene (ERE element), auxin (TGA element), abscisic acid (ABRE, ABRE3a e ABRE4 elements) and salicylic acid (TCA and TCA elements). Hormones are important regulators of defense responses to biotic and abiotic stresses, abscisic acid acts as a central regulator of many plant responses to environmental stresses (VERMA; RAVINDRAN; KUMAR, 2016; DAR, et al., 2017; BARI; JONES, 2009; SHINOZAKI; YAMAGUCHI-SHINOZAKI, 1997). Others *cis*-elements related to plant development such as zein metabolism, meristem expression and circadian control were also present in the promoters of the *JcRIP* genes (Fig 2).

This analysis suggests that the regulation of RIP genes expression is complex, as many motifs respond to different external or endogenous inductions. Studies show that the expression of RIPs is affected by environmental conditions, hormones and maturation stages. A ribosome inactivating protein known as PIP2, present in *Phytolacca insularis*, showed expression induced by drought and senescence, in addition, mechanical injuries, jasmonic acid and abscisic acid

caused an inducible systemic effect on the expression of PIP2 (SONG et al., 2000). Another RIP present in *Beta vulgaris* had its expression increased when the plant was subjected to treatments with salicylic acid and hydrogen peroxide, mediators of acquired resistance (IGLESIAS et al., 2008). The *PAP* gene encodes a type of ribosome inactivating protein (RIP) of the species *Phytolacca americana* L., CAO et al., (2011) conducted an *Agrobacterium*-mediated transformation to insert into cucumber plants (*Cucumis sativus* L.) the deficient *PAP* gene (*PacPAP*) that is capable to inhibit only viruses, without causing damage to the carrier plant. The results obtained indicated that the transformed plant showed better tolerance to pathogenic virus and fungi.

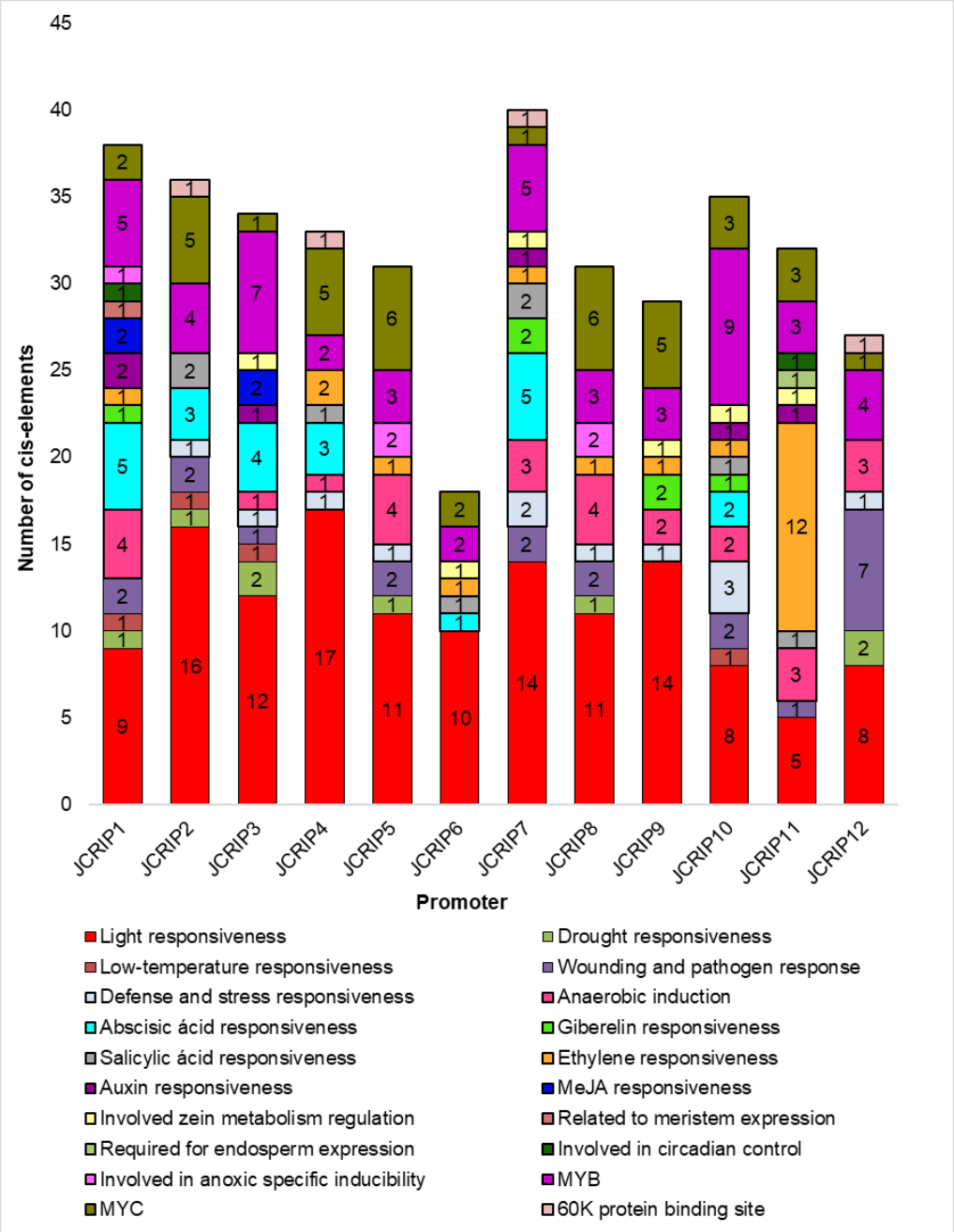


Figure 2. *Cis*-elements present in the promoter region of physic nut RIP genes. The cis-elements were analyzed in the upstream promoter region of the translation start site using the plantCARE database.

In silico JcRIP gene expression

The *JcRIP1* and *JcRIP3* genes showed higher levels of expression in seed and leaf, respectively, in addition to being expressed in all analyzed tissues. The genes *JcRIP7* and *JcRIP9* showed high expression in roots, while in seeds there was expression of the genes *JcRIP1*, *JcRIP3* and *JcRIP12*. In leaves, the most

expressed genes were JcRIP1, JcRIP2 and JcRIP3, with JcRIP3 being the most expressed. In the stem and in the flowers the levels of expression were lower, being that in these tissues the genes *JcRIP1*, *JcRIP2*, *JcRIP3* and *JcRIP6* showed subtly higher expression than the others. The JcRIP4 and JcRIP10 genes showed only a small expression in the leaves, and for the *JcRIP8* and *JcRIP11* genes, no expression was detected in the analyzed tissues (Fig. 3).

The graph of the promoter region indicated that the genes that had greater expression (*JcRIP1* and *JcRIP3*) have cis elements involved in the regulation mediated by Methyl-Jasmonate (CGTCA-motif and TGACG-motif elements) (Fig. 2 and Fig. 3). Jasmonic acid (JA) is part of an important group of phytohormones associated with the immune response of the plant (WASTERNACK; HAUSE, 2013). A RIP domain protein in barley (*Hordeum vulgare* L.), JIP60 (60kDa Jasmonate-Induced Protein), is involved in the immune response of the plant, and is expressed when the plant is exposed to stresses caused by the accumulation of jasmonate (REINBOTHE et al., 1994; PRZYDACZ et al., 2020).

Although *RIP* genes are expressed mainly in seeds, preventing animals from consuming it, and also in leaves, where they are involved in the defence against biotic and abiotic stresses (WU et al., 2017), it was evident that in roots, flowers and stems of *J. curcas* L. also occurs expression of RIPs genes.

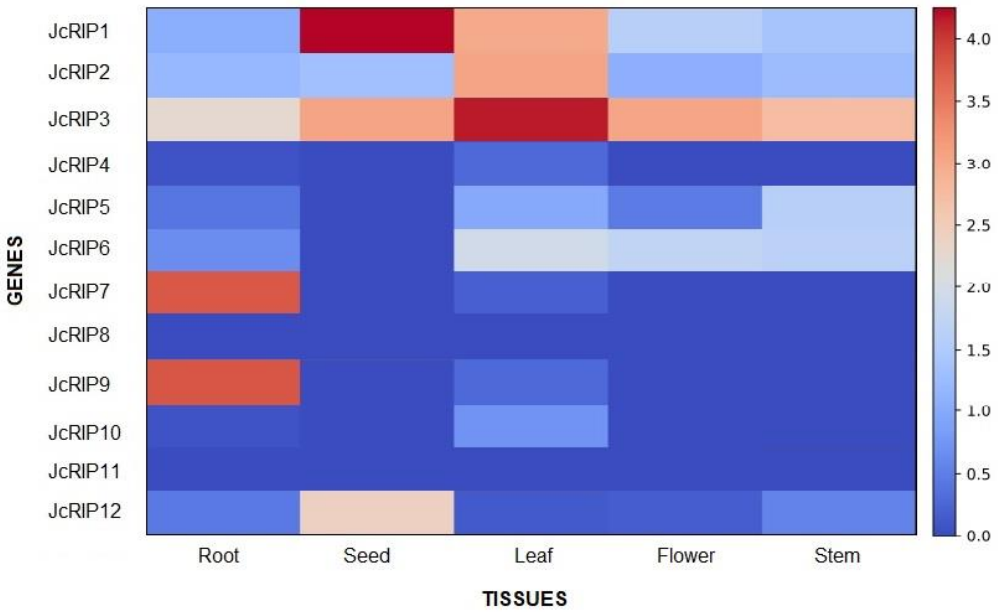


Figure 3. Graph of gene expression of *JcRIPs* in different tissues under normal conditions. The graph was obtained using the JCDB Heatmap tool (ZHANG et al., 2019).

Analysis of the physic nut RIP proteins

The number of amino acid residues of the physic nut RIP proteins ranged from 291 (JcRIP8) to 551 (JcRIP12) (Table 2). All proteins were predicted to follow secretory pathways and are located in the extracellular environment (59,63% to 90,92% probability; Table 2), they have a signal peptide sequence at the N-terminal end, containing between 22 and 28 amino acid residues. The fact that they have a signal peptide sequence is indicative that protein synthesis follows a secretion pathway, allowing the segregation of these cytotoxic proteins in extracellular compartments (READY et al., 1986). It is already well known that some RIPs have an initial sequence containing on average 24 amino acids necessary for the translocation of the protein to the endoplasmic reticulum (ER) (LAMB et al., 1985). In a study carried out with the RIP saporin (*Saponaria officinalis* L.), the results suggest that the translocation of the RIPs for ER can be increased in stressful situations, for example attacks of pathogens, selectively by the signal peptide (MARSHALL et al., 2011).

The PF00161 functional domain that characterizes the Ribosome Inactivating Proteins (RIPs) superfamily was found in all JcRIP proteins (Figure S4), it is the only domain present in 11 of the 12 studied proteins. However, the protein JcRIP12 presented in addition to the functional domain PF00161, two PF00652 domains, the latter being similar to the ricin lectin B chain, which has a specific binding site for plasmatic membrane glycoproteins, which facilitates the entry of this type of RIP into the cell and makes them more cytotoxic (SANDVIG; VAN DEURS, 1994; STEEVES, et al., 1999). In the Pfam Database there are 649 protein sequences containing only the RIP domain in its architecture, distributed among several species of plants, fungi, bacteria and viruses, which are characterized as Type I RIP. On the other hand, the same bank has only 23 strings with architecture like JcRIP12 (EL-GEBALI et al., 2019).

The hydropathicity (Grand average of hydropathicity – GRAVY) of the proteins ranged from -0,469 to -0,073 (Table 2), indicating that they have great affinity for water, a property expected for proteins destined for the extracellular environment. The prediction of the molecular weight of curcin proteins (JcRIP1 to JcRIP11) without signal peptide ranged from 25.89kDa (JcRIP10) to 35.0kDa (JcRIP9), and the Isoelectric point (pI) from 4.8 (JcRIP4) to 9.31 (JcRIP11), while

JcRIP12, a type II RIP, had a molecular weight of 58.22kDa and a pI of 5.47. Studies state that the molecular weight of type I RIPs is between 30kDa, while type II RIPs have molecular weight around 60kDa, due to the lectin B chain (WANG et al., 2016; ZHU et al., 2018).

Table 2. Characteristics of RIP proteins present in the physic nut (*Jatropha curcas*) genome. GRAVY: grand average of hydropathicity; Mw: molecular weight; pI: isoeletric point; SP: signal peptide.

Protein	ID	Protein size (aa)	SP size (aa)	Mw with/without SP (kDa)	pI with/without SP	Export probability Extracellular (%)	GRAVY with/without SP
JcRIP1	JCDBG01176	293	28	32,57/29,49	8,55/7,83	73,56	-0,05/ -0.146
JcRIP2	JCDBG01177	309	28	34,64/31,55	8,9/8,72	59,63	0,022/-0.073
JcRIP3	JCDBG01178	309	28	34,89/31,79	6,75/ 5,85	76,12	-0,102/-0.210
JcRIP4	JCDBG01180	309	28	34,86/31,85	4,95/4,8	84,31	-0,226/ -0.346
JcRIP5	JCDBG10548	308	28	34,91/31,70	5,99/5,33	73,01	-0,181/ -0.282
JcRIP6	JCDBG01181	302	28	34,22/30,99	5,7/5,16	80,84	-0,235/-0.335
JcRIP7	JCDBG12049	316	24	35,4/32,72	6,95/7,11	82,65	-0,290/ -0.420
JcRIP8	JCDBG12048	291	25	32,95/30,13	5,84/5,39	70,14	-0,327/ -0.469
JcRIP9	JCDBG12053	314	24	35,0/32,27	8,58/8,33	80,47	-0,110/-0,209
JcRIP10	JCDBG12050	255	22	28,39/25,89	5,99/5,99	75,62	-0,140/ -0.280
JcRIP11	JCDBG12054	305	24	33,96/31,19	9,25/9,31	90,92	-0,270/ -0.400
JcRIP12	JCDBG12052	551	23	60,88/58,22	5,76/5,47	99,99	-0.211/ -0.262

All proteins showed phosphorylation sites: JcRIP1 has 28 phosphorylation sites (8Thr / 15Ser / 5Tyr); JcRIP2 e JcRIP9 have 27 phosphorylation sites (7Thr / 14Ser / 6Tyr e 13 Thr / 10 Ser / 4 Tyr, respectively); JcRIP3 e JcRIP6 have 29 phosphorylation sites (7Thr / 16Ser / 6Tyr e 5Thr / 17Ser / 7Tyr, respectively); JcRIP4 has 30 phosphorylation sites (6Thr / 18Ser / 6Tyr); JcRIP5 e JcRIP10 have 26 phosphorylation sites (5Thr / 16Ser / 5Tyr e 14Thr / 6Ser / 6Tyr, respectively); JcRIP7 has 31 phosphorylation sites (13Thr / 12Ser / 6Tyr); JcRIP8 has 25 phosphorylation sites (9Thr / 7Ser / 9Tyr); JcRIP11 has 24 phosphorylation sites (7Thr / 13Ser / 4Tyr); e a JcRIP12 has 55 phosphorylation sites (18Thr / 31Ser / 6Tyr). The proteins JcRIP1, JcRIP2, JcRIP3, JcRIP5, JcRIP7, JcRIP9, JcRIP10,

JcRIP11 and JcRIP12 showed N-glycosylation site (3, 3, 4, 1, 1, 1, 2, 2 e 9 sites, respectively; Table 3; S.4 Fig.).

Table 3. Post-translational modifications of physic nut RIP proteins. RIP: Ribosome-inactivating protein (Proteínas Inativadoras de Ribossomos), S: Serine, Y: Tyrosine, T: Threonine, N: Asparagine. Underlined aminoacids are part of the signal peptide.

Protein	Phosphorylation sites	N-glycosylation sites
JcRIP1	S19, S27, S37, Y41, S45, T56, T57, S76, Y77, S78, S79, T88, T127, S128, S136, T146, S154, T156, S158, Y159, T186, S190, S191, Y215, S222, T227, S236, Y267	N7, N266, N274
JcRIP2	S19, S27, S37, Y41, S45, T46, Y62, S76, Y77, S78, S79, T88, T127, S128, S136, T146, S154, T156, S158, Y159, S191, Y215, T222, T227, S236, Y267, S294	N7, N266, N274
JcRIP3	S19, S27, Y41, T46, T56, T57, S76, Y77, S79, T88, S103, S126, S128, S134, S136, S158, Y159, Y183, T186, S191, Y215, S221, T222, S225, T227, S236, S251, Y267, S294	N7, N132, N266, N274
JcRIP4	S19, S27, S37, Y41, T46, Y62, T67, T72, S76, S79, T88, S126, S128, S134, S136, T146, S158, Y159, Y183, S191, S202, Y215, S225, T227, S236, S240, Y246, S251, S294, S302	-
JcRIP5	T6, S19, S27, S36, Y41, T48, T52, S76, S79, S88, S103, S126, S128, T146, Y160, Y184, S192, S203, Y216, S226, T228, S237, S249, Y265, S292, S300	N132
JcRIP6	T6, T14, S19, S27, S33, S36, Y41, T48, S75, S76, S79, S88, S126, S128, Y136, T146, S158, Y159, Y183, S191, S202, Y215, S225, T227, S236, S240, Y246, Y267, S294	-
JcRIP7	T13, T38, T48, S73, T75,T76, S77, S85, S88, Y93, S119, Y124, Y141, T144, T149, T150, Y158, S160, T170, T171, T184, T209, Y211, Y213, S223, S225, S243, S273, S291, S295, T299	N34
JcRIP8	T14, T43, T50, Y55, S64, T68, T70, S81, S83, Y88, Y111, Y136, Y152, S154, Y166, T178, T203, Y205, Y207, S228, S243, Y259, S283, T286, T291	-
JcRIP9	T9, T32, T42,Y54, S67, T69, T70, T79, T80, S82, Y87, S120, Y121, T134, T144, S150, S153, T177, Y206, T220, S227, S236, T251, S260, S265, S284, T292	N30
JcRIP10	T13, T36, T48, T50, T75, T76, S85, T86, S88, Y93, Y116, Y141, T144, T149, S152, Y158, S160, T170, T171, S180, T184, T189, T209, Y211, Y213, S232	N34, N97
JcRIP11	S30, T46, Y58, S70, S71, T73, T74, S83, T84, S86, Y91, S99, S124, T138, T147, S154, Y155, S157, S167, T181, S194, S210, S222, Y289	N28, N34
JcRIP12	T10, S15, T16, T20, S21, S31, Y49, S55, S62, S70, T77, S81, S90, S92, T94, T95, Y118, S142, S143, Y144, S150, S156, S170, T179, Y201, S209, T210, T221, S224, T234, S239, S277, S282, S284, S285, T292, T297, Y315, S316, S328, T337, T343, S346, T376, T407, Y413, S416, T425, S430, S468, T492, S497, S500, S501, S530	N38, N115, N131, N149, N354, N374, N383, N423, N443

Post-translational modifications, such as phosphorylation and glycosylation, play important roles in protein activation. In which phosphorylation, reversible binding of a phosphate group to a residue of ser / thr or tyr, is most prevalent covalent modification of proteins and has been described as having an important role in defence signaling cascades (PECK, 2003; DE LA FUENTE, BENTEM and HIRT, 2007). Glycosylation, covalent attachment of an oligosaccharide side chain to a protein, occurs quite frequently and can affect biological activity, folding and function of proteins, in addition to providing greater stability (VAN DER HOORN et al., 2005).

Conserved amino acids

Thirty-five amino acid residues are conserved in all JcRIPs sequences (Fig. 4). The catalytic site of RIPs is responsible for inactivating the ribosome and is composed of amino acid residues Y_{88}, Y_{131}, E_{185}, R_{188}, W_{219} (MONZINGO; ROBERTUS, 1992), they showed conservation in nine JcRIPs proteins, with the exception of JcRIP2, JcRIP5 and JcRIP12 proteins.

The JcRIP2 protein presented the amino acid A to replace E_{185}, while the JcRIP5 protein did not present the W_{219} residue. Although the two residues of Y_{88} and Y_{131} present in the catalytic site of the RIPs are fundamental in binding to the adenine of the rRNA (DI MARO et al., 2014), the protein JcRIP12 presented the amino acid G in place of the first residue Y_{88} (Fig. 4). Replacing a Trp with a Phe reduces catalytic activity, but does not abolish it, suggesting that Trp may play a role in forming the conformation of the active protein, but does not participate directly in the catalytic reaction (BRADLEY; MCGUIRE, 1990; FUNATSU et al., 1991). The replacement of Glu by an Ala, which is the case with JcRIP5, can reduce power by up to 18 times (SCHLOSSMAN et al., 1989).

The YFF motif (Tyr_{99}-Phe_{100}-Phe_{101}) was initially described in a RIP with antiviral potential belonging to the species *Trichosanthes kirilowii*, it is a potential antigenic site (AN et al, 2006) that has been conserved in all JcRIPs (Fig 3). Eight of the 12 JcRIPs have the EAAR motif (Glu_{185}-Ala_{186}-Ala_{187}-Arg_{188}) well maintained. Such residues, Glu_{185}, Ala_{186} and Arg_{188}, are well preserved in most of the RIPs already studied, and together with the others, compose the region of the RIPs catalytic site (FUNATSU et al., 1991; WANG; TURNER, 2000; BEGAM et al., 2006; ROY et al., 2006; MALETZKI et al., 2013; DANG, et al., 2017).

In studies of the N-glycosidase activity of ricin, a RIP protein that accumulates a large number of studies, it was observed that in the eukaryotic ribosomes, the Ricin toxin A chain (RTA), specifically and irreversibly cleaves the glycosidic bond of adenine at position 4324 in the 28S rRNA of the 60S subunit. This adenine belongs to a highly conserved sequence, 5'-AGUACGAGAGGA-3', which is about 12 nucleotides in length and is universally found in eukaryotic ribosomes, known as SRL (Sarcin-Ricin Loop) and is important for elongation factors binding during protein synthesis. This depurination caused by RTA quickly and completely

inactivates the ribosome, which results in toxicity, in turn, leads to inhibition of protein synthesis. Approximately 1500 ribosomes are inactivated by a single RTA molecule in the cytosol in one minute, which characterizes it as highly cytotoxic (OLSNES; KOZLOV, 2001; BRADBERRY et al., 2003; SPERTI et al., 1973; STURM; SCHRAMM, 2009).

There are several amino acid residues conserved at the active site of the RTA and other RIPs involved in the purification of ribosomal RNA, including the *J. curcas* L. RIPs. The exact mechanism of this event is still unknown, but the main amino acid residues identified as involved in the reaction include tyrosine (TYR) at positions 80 and 123, glutamic acid (GLU) at position 177 and arginine (ARG) at position 180. In particular, GLU_{177} and ARG_{180}, are involved in the catalytic mechanism, but not in the binding to the substrate (KIM; ROBERTUS, 1992; FRANKEL et al., 1990). Despite all JcRIPs have conserved active site amino acids, only JcRIP12 has the lectin chain that binds to plasmatic membrane glycoproteins resulting in easier entry into the cell and therefore greater toxicity (STEEVES, et al., 1999; SANDVIG; VAN DEURS, 1994).

Figure 4. Alignment of multiple sequences of ribosome inactivating protein of *Jatropha curcas* L, and a RIP with known 3D structure (represented by the PDB:5DDZ code). The invariant amino acids are in red color. In the black box the aminoacids of the active site described for the RIPs. The YFF and EAAR motif hatched in gray.

Molecular modeling of JcRIP proteins by homology

All proteins of the JcRIP family were modeled using the Swiss-model server. The models of the species *R. communis* L. (5DDZ, 1J1M, 4Q2V, 3RTI, 1RZO and 6OBG) were better for the six JcRIPs proteins, for the other JcRIPs proteins the best models were found in the species *Abrus precatorius* (1ABR), *Suregada multiflora*

(3KTZ), *Cucurbita moschata* (3BWH) and *Cinnamomum camphora* (2VLC) (Table 4). The identity values between the protein and the mold were from 31.56% (JcRIP9 and 5DDZ) to 52.4% (JcRIP12 and 2VLC), and the similarity between the amino acid sequences ranged from 36% (JcRIP5 and 4Q2V) to 44% (JcRIP12 and 2VLC), with a minimum coverage of 84% (JcRIP7 and 3KTZ) and a maximum of 99% (JcRIP12 and 2VLC), covering the entire RIP protein domain for all sequences. The QMEAN ranged from -3.48 (JcRIP9) to -1.82 (JcRIP10), and Ramachandran showed values from 90.61% (JcRIP4) to 96.76% (JcRIP11) of waste in favorable regions, including the residues from the active site (Table 4 S5 Fig.).

The general quality of all models (z-score) showed values within the range normally found in native proteins for this size (S6 Fig.). When aligned and overlapping, all proteins showed remarkably similar three-dimensional structures, only JcRIP12 showed greater divergence because it had two polypeptide chains (Fig. 5).

Figure 5. Overlapping all JcRIPS proteins. The protein represented in violet is JcRIP12 which has two polypeptide chains.

The three-dimensional structures of *Jc*RIPs proteins were not determined by X-ray crystallography and Nuclear Magnetic Resonance (NMR) spectroscopy, and so far only two studies have been reported that have performed molecular modeling and docking with curcin. One of them was published by Srivastava et al., (2011), who predicted the curcin 3D model (GenBank: ACO53803.1) through ab initio modeling, achieving a total of 84% of the amino acid residues in favorable regions

and 12.3% in regions allowed, and for docking only adenine was used as a ligand. The other was published more recently by Palanivel et al., (2019), using the Swiss Model to do homology modeling of a curcin type 1 (GenBank: ABZ04128.1), its generated model obtained 90.3% of the amino acids in regions, 11 pterin-derived ligands were used in the docking as possible inhibitors. In the present work, we obtained better quality models, this was possible due to software updates, creation of new tools and resolution of new three-dimensional structures deposited in the databases (Table 4, S5 and S6 Fig.).

Table 4. Homology modeling of proteins JcRIPs using the Swiss Model server and validation of the quality of the model obtained. Identity, similarity, coverage, GMQE, and Ramachandran values are shown in %. QMEAN: Qualitative Model Energy Analysis; GMQE: Global Model Quality Estimation.

Protein	Mold (PDB ID)	Identity	Seq Similarity	Coverage	Range	QMEAN	GMQE	RAMACHANDRAN	Overall model quality (z-score)
JcRIP1	5DDZ	35,51	37	92	17-263	-2,42	69	96,33	-7,27
JcRIP2	1J1M	37,25	38	88	18-266	-3,08	65	94,74	-6,72
JcRIP3	5DDZ	34,82	37	88	17-265	-3,41	64	95,95	-7,28
JcRIP4	1ABR	34,98	37	86	18-264	-2,68	63	90,61	-7,03
JcRIP5	4Q2V	33,06	36	88	17-263	-2,38	63	93,06	-7,16
JcRIP6	3RTI	34,41	37	90	18-264	-2,37	66	93,88	-6,8
JcRIP7	3KTZ	36,89	37	84	17-266	-2,58	61	95,16	-6,66
JcRIP8	1RZO	34,84	37	92	13-259	-2,38	68	95,51	-7,78
JcRIP9	5DDZ	31,56	36	84	15-263	-3,48	61	95,55	-7,49
JcRIP10	3BWH	36,27	38	88	22-233	-1,82	64	93,33	-6,85
JcRIP11	6OBG	33,74	37	86	15-263	-2,11	65	96,76	-7,89
JcRIP12	2VLC	52,4	44	99	11-528	-3,37	76	92,05	-10,34

Folic and pteroic acids were used as ligands because they are compounds already known as inhibitors of other RIPs, including curcin. In addition, dexamethasone and its derivatives were tested: hydrocortisone, prednisolone, methilprednisolone, flumetasone, betamethasone valerate, triamcinolone and fludrocortisone. To date, there are no reports of research using dexamethasone derivatives as ligands that inhibit RIPs.

The ligands/inhibitors already described for RIPs had good affinity with the catalytic site of JcRIPs proteins, folic acid had the same affinity energy value for JcRIP1 and JcRIP3, with -7.2 kcal/mol being the value, whereas for JcRIP12 the

value was -6.6. Pteroic acid had an affinity of -7.2 kcal/mol with JcRIP1 and -6.2 kcal/mol with JcRIP3, while JcRIP12 showed only -6.2 kcal/mol of affinity energy with both acids. As for dexamethasone and its derivatives, all showed high energy of affinity with the catalytic sites of the JcRIP proteins, except betamethasone valerate which presented values of -6.6, -6.8 and -6.2 respectively for JcRIP1, JcRIP2 and JcRIP12. Dexamethasone and hydrocortisone were the compounds that showed the highest affinity energy, with the first compound having a value of -8.0 kcal/mol for JcRIP1, and the second having a value of -8.1 kcal/mol for JcRIP12 (Tab. 5).

Table 5. Result of the molecular docking process between JcRIPs and their possible ligands. Docking performed with JcRIP1, JcRIP3 and JcRIP12, already known binders for RIPs and new drugs.

	JcRIP1	**JcRIP3**	**JcRIP12**
Ligands	**kcal/mol**	**kcal/mol**	**kcal/mol**
Folic acid	-7.2	-7.2	-6.6
Pteroic Acid	-7.2	-6.2	-6.2
Dexamethasone	-8.0	-8.1	-7.9
Hydrocortisone	-7.5	-7.5	-8.1
Prednisolone	-7.6	-7.6	-8.1
Methylprednisolone	-7.4	-7.4	-7.4
Flumethasone	-7.7	-7.8	-7.4
Betamethasone valerate	-6.6	-6.8	-6.2
Triamcinolone	-7.1	-7.2	-7.2
Fludrocortisone	-7.5	-7.5	-7.1

The present work presents an innovative character, because when verifying in silico that dexamethasone and its derivatives had high affinity with the JcRIP proteins, which in turn present the same conserved amino acids from the catalytic site of ricin and other RIPs, it indicates a potential application biotechnology of these compounds in blocking the action of ribosome inactivating proteins, acting as detoxifiers. Ricin is considered one of the most toxic substances known, and is even considered a potential chemical warfare agent (SOUSA et al., 2014). Its median lethal dose (LD50) for mice, rats and dogs is ± 1 µg / kg body weight (OLSNES; PIHL, 1982), whereas for human lethal oral dose it was estimated at 1 to 20 mg / kg body weight (approximately 8 seeds) (KLAIM et al., 1990; AUDI et al., 2005).

Over the years, innumerable cases of human and animal poisoning by ricin have been reported, in a summary of cases made by Worbs et al., (2011), a total of

887 human poisonings (oral and injectable) were identified, of which 13 were of these lethal. As for injectable poisoning, there were 6 reports, 5 of which were fatal. In animals, 187 cases of oral poisoning were identified, 12 of which were fatal. Poisoning by physic nut is rare when compared to castor bean. One case of mass accidental poisoning by eating physic nut seeds was recorded in 2014 in India, 23 people were hospitalized with gastrointestinal symptoms, such as nausea, vomiting, abdominal cramps, diarrhea and lethargy. Each victim consumed 3 to 9 seeds, on average, however the severity of clinical manifestations was not correlated with the number of seeds consumed (GUPTA et al., 2016).

Due to the toxicity of RIPs, their distribution in several plant species and the cases of intoxication that have already occurred, the incessant search for an effective inhibitor is crucial. The first inhibitor of ricin A chain, identified from virtual screening, was pteroic acid. However, efforts to improve the inhibitory activity of pteroic acid have been unsuccessful due to synthetic restrictions and a serious chemical disadvantage of pterine-based compounds, including their insolubility in water (YAN et al., 1997; JASHEWAY et al., 2011). Folic acid, which is also a ricin inhibitor, also has low water solubility of 1.6 mg/L at 25 ° C (O'NEIL, M.J. 2001).

Dexamethasone and its derivatives, in addition to presenting good interaction results when coupled with the RIP proteins studied here, have high solubility in water at 25 ° C, being 89 mg/L for dexamethasone and 320 mg/L for hydrocortisone (YALKOWSKY; DANNENFELSER, 1992; YALKOWSKY et al., 2003).

The 2D diagram of the interactions between JcRIP1 and JcRIP3 with the binding compounds showed interactions through non-covalent hydrogen bonds, van der walls, π interactions (π-cation, π-Sigma, π-Alkyl and Stacking-π-π). In both, the Tyr_{90} residue, which is part of the RIPs catalytic site, appeared interacting with all the tested compounds, with different types of strength, among them the π-π stacking, type of interaction that occurs between the tyrosine aromatic rings ligand, facilitating protein-ligand binding (Fig. 6).

In the 2D diagram of JcRIP12 with the ligands there were no π-π stacking interactions, as it does not have the tyrosine that performs this interaction with the ligand, while the other amino acids from the active site were all present in the interactions. In some protein-ligand complexes, such as JcRIP12 linked to hydrocortisone and prednisolone, the Glu_{195} residue showed an attractive charge

bond. And Glu_{230} showed a halogen bond (fluorine) in the last three protein-ligand complexes (flumetasone, fludrocortisone and triamcinolone).

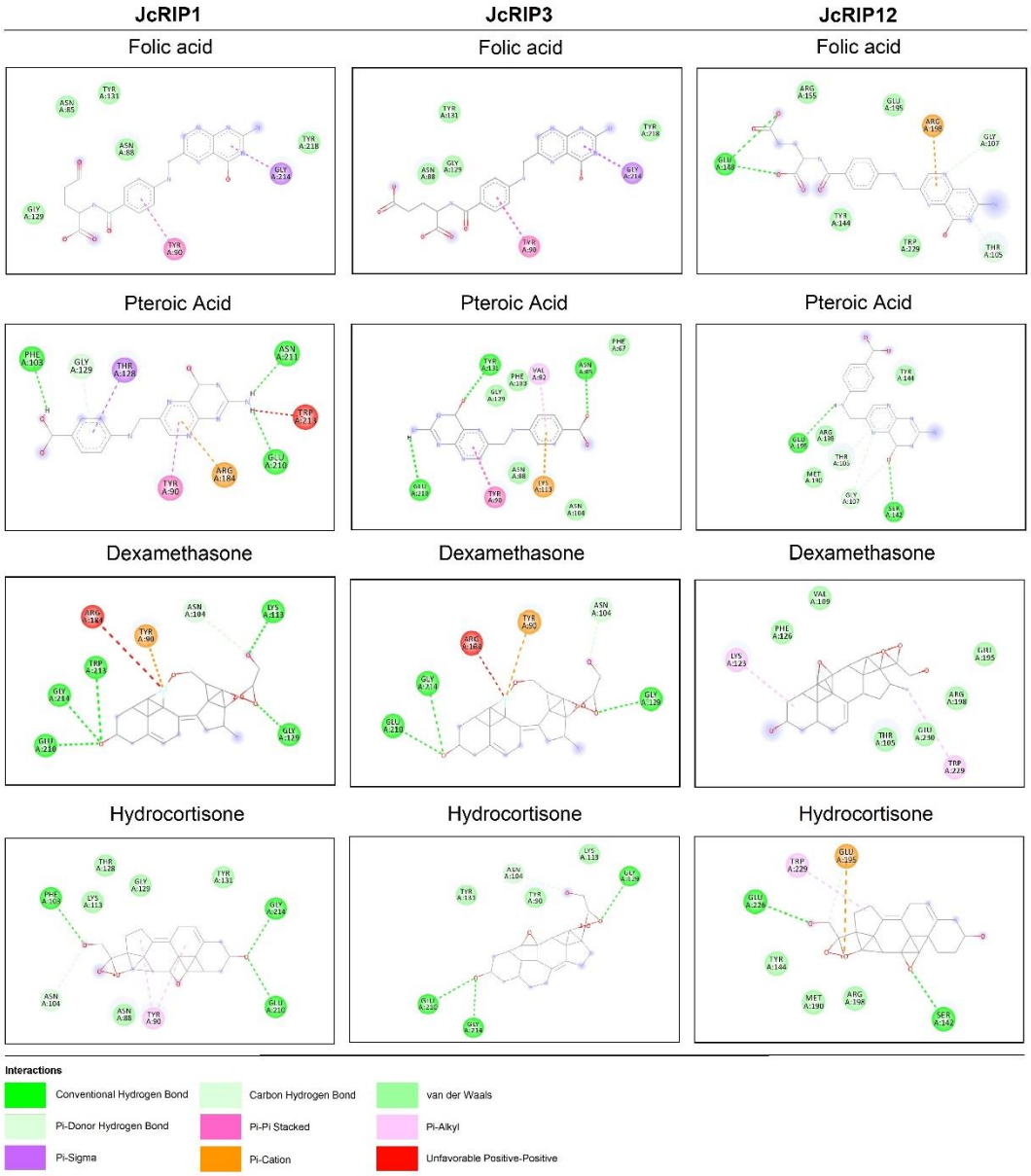


Figure 6. 2D diagram of the main molecular coupling combinations. In the first, second and third columns the interaction of the ligands, respectively, with the proteins JcRIP1, JcRIP2 and JcRIP3.

In view of the results found in the coupling of the JcRIPs with dexamethasone, the interaction energy test between the ricin A chain (RTA) (PDB code: 5DDZ: A) and dexamethasone was performed, an interaction free energy value was obtained -9.2 kCal/mol. The drug fits very well in the catalytic pocket of ricin, and Tyr88, an important residue of the catalytic site, interacts strongly with the

compound forming three Alkyl bonds (Fig. 7). These data prove the potential use of dexamethasone, a drug widely used to treat other diseases and whose contraindications are already well known, as a ricin inhibitor.

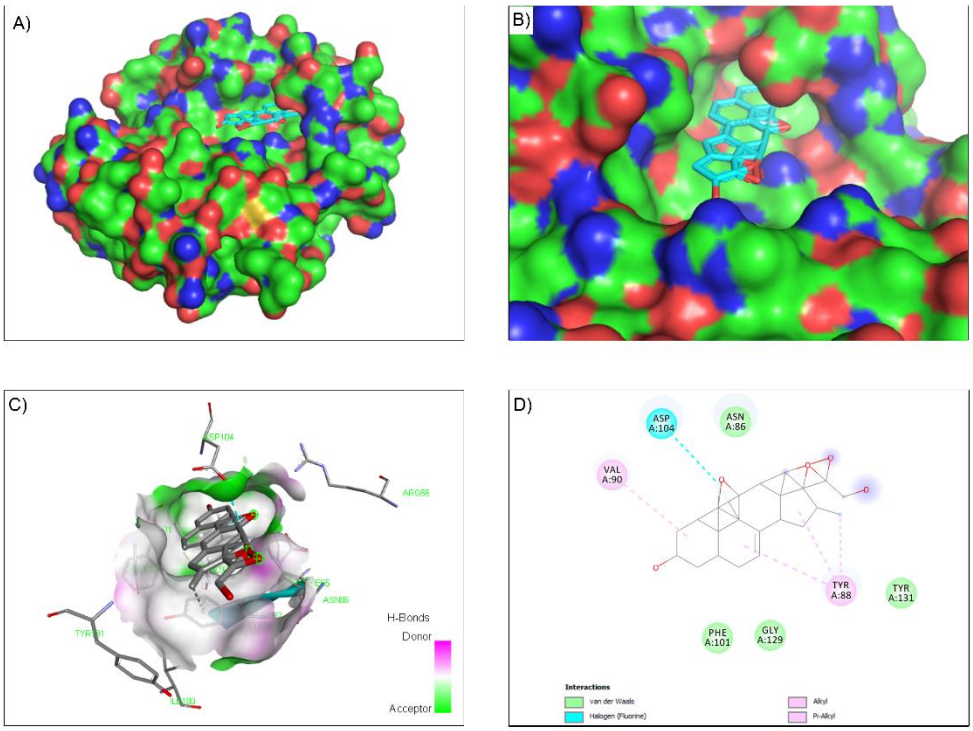


Figure 7. Coupling of ricin (PDB: 5DDZ) with dexamethasone. A and B) 3D structure of ricin linked to dexamethasone. C) Detail of the ricin catalytic pocket and D) 2D diagram showing the interactions between the ricin amino acids and dexamethasone.

Conclusion

The present study is the first to characterize the gene family of ribosome inactivating proteins (RIPs) of *Jatropha curcas* L. Twelve members have been identified, eleven of them are type 1 RIPs and only one is classified as type 2 RIP. According to the cis elements in the promoter region, the *J. curcas* L. RIPs genes can be regulated by several external and internal factors, possibly involved involved in the defense of the plant, since they all had elements MYB and MYC that are part of important classes of transcription factors.

According to the expression profiles in the different tissues, the JcRIP3 gene is the one with the highest expression in the tissues. In addition to seeds and leaves,

there is also expression of RIP genes in stems, roots and flowers. Post-translational changes such as phosphorylation and glycosylation have also been predicted.

The analysis of protein structures showed that the eleven type I JcRIPs have a single polypeptide chain, formed by a RIP domain containing an average of 8 alpha-helix and 8 beta-sheet, while the only type II RIP presented two polypeptide chains, the first presenting the RIP domain , containing 8 alpha-helix and 7 beta-sheet, and the second presenting the lectin domain and 2 alpha-helix and 20 beta-sheet.

These data are important for the functional study of RIPs genes / proteins of *J. curcas* L., in addition, the high conservation existing between the residues of the JcRIP proteins with the RTA (5DDZ) allows dexamethasone and its derivatives to bind to all these proteins with high affinity energy, thus being possible to indicate in an unprecedented way the use of these drugs in biotechnological approaches that lead to the inhibition of ribosome inactivating proteins of *J. curcas* L. and *R. communis* L.

References

AN, Q.; WEI, S.; MU, S.; ZHANG, S.; LEI, Y.; ZHANG, W.; JIA, N.; CHENG, X.; FAN, A.; LI, Z.; XU, Z. Mapping the antigenic determinants and reducing the immunogenicity of trichosanthin by site-directed mutagenesis. **Journal of Biomedical Science.** v. 13, p. 637-643, 2006.

ARMENTEROS, J.J.A.; SONDERBY, C.K.; SONDERBY, S.K.; NIELSEN, H.; WINTHER, O. DeepLoc: prediction of protein subcellular localization using deep learning. **Bioinformatics**, v. 33, p. 3387–3395, 2017.

ARNOLD, K.; BORDOLI, L.; KOPP, J.; SCHWEDE, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling**. Bioinformatics.** v. 22, p. 195-201, 2006.

AUDI, J.; BELSON, M.; PATEL, M.; SCHIER, J.; OSTERLOH, J. Ricin Poisoning. **JAMA***,* v. 294, p. 2342, 2005.

BARI, R., JONES, J.D.G. Role of plant hormones in plant defence responses. **Plant Mol Biol,** v. 69, p. 473–488, 2009.

BEGAM, M.; KUMAR, S.; ROY, S.; CAMPANELLA, J.J.; KAPOOR, H.C. Molecular cloning and functional identification of a ribosome inactivating/antiviral protein from

leaves of post-flowering stage of Celosia cristata and its expression in E. coli. **Phytochemistry**, v. 67, p. 2441–2449, 2006.

BLOM. N.; GAMMELTOFT, S.; BRUNAK, S.; Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. **Journal of Molecular Biology,** v. 294, p. 1351-1362, 1999.

BOLOGNESI, A.; BORTOLOTTI, M.; MAIELLO, S.; BATTELLI, M.G.; POLITO, L. Ribosome-Inactivating Proteins from Plants: A Historical Overview. **Molecules**, v. 21, p. 1627, 2016.

BRADBERRY, S.M.; DICKERS, K.J.; RICE, P.; GRIFFITHS, G.D.; VALE, J. A. Ricin Poisoning. **Toxicological Reviews**, v. 22, p. 65-70, 2003.

BRADLEY, J.L.; MCGUIRE, P.M. Site-directed mutagenesis of ricin A chain Trp 211 to Phe**. International Journal of Peptide and Protein,** v. 35, p. 365-366, 1990.

CAO,B.; LEI, J.; CHEN, G.; CAO, P.; LIU, X.; CHEN, Q.; WEI, X. Testing of disease-resistance of pokeweed antiviral protein gene (*PacPAP*) in transgenic cucumber (*Cucumis sativus*). **African Journal of Biotechnology**. v. 10, p. 6883-6890, 2011.

CHAN, A.P.; CRABTREE, J.; ZHAO, Q.; LORENZI, H.; ORVIS, J.; PUIU, D.; MELAKE-BERHAN, A.; JONES, K.M.; REDMAN, J.; CHEN, G.; CAHOON, E.B.; GEDIL, M.; STANKE, M.; HAAS, B.J.; WORTMAN, J.R.; FRASER-LIGGETT, C.M.; J.; PABLO D RABINOWICZ, P.D. Draft genome sequence of the oilseed species *Ricinus communis*. **Nature Biotechnology,** v. 28, p. 951–956, 2010.

DALLAL, J. A.; IRVIN, J. D. Enzymatic inactivation of eukaryotic ribosomes by the pokeweed antiviral protein. **FEBS Letters,** v. 89, p. 257–259, 1978.

DANG, L.; ROUGÉ, P.; DAMME, E. Expression Analysis of Ribosome-inactivating Proteins From Cucumber. **Progress in Biochemistry and Biophysics**, v. 44, p. 929-941, 2017.

DAR, N.; AMIN, I.; WANI, W.; WANI, S.; SHIKARI, A.; WANI, S.; MASOODI, K. Abscisic acid: A key regulator of abiotic stress tolerance in plants. **Plant Gene,** v. 11, p. 106- 111, 2017.

DEBNATH, M.; BISEN, P. *Jatropha Curcas L.,* A Multipurpose Stress Resistant Plant with a Potential for Ethnomedicine and Renewable Energy*.* **Current Pharmaceutical Biotechnology**, v. 9, p. 288–306, 2008.

DE LA FUENTE, V.; BENTEM, S.; HIRT, H. Using phosphoproteomics to reveal signalling dynamics in plants. **Trends in Plant Science**, v. 12, p. 404–411, 2007

DI MARO, A.; CITORES, L.; RUSSO, R.; IGLESIAS, R.; FERRERAS, J.M. Sequence comparison and phylogenetic analysis by the maximum likelihood method of ribosome-inactivating proteins from angiosperms. **Plant Molecular Biology**, v. 85, p. 575–588, 2014.

EL-GEBALI, J.S; MISTRY, A.; BATEMAN, S. R.; EDDY, A.; LUCIANI, S. C.; POTTER, M.; QURESHI, L. J.; RICHARDSON, G. A.; SALAZAR, A.; SMART, E. L. L.; SONNHAMMER, L.; HIRSH, L.; PALADIN, D.; PIOVESAN, S. C. E.; TOSATTO, FINN, R. D. The Pfam protein families database in 2019. **Nucleic Acids Research,** v. 47, p. 427-432, 2019.

FAN, X.; ZHU, Y.; WANG, C.; NIU, L.; TENG, M.; LI, X. Structural insights into the interaction of the ribosomal P stalk protein P2 with a type II ribosome-inactivating protein ricin. **Scientific Reports**, v. 6, n. 37803, p. 1-10, 2016.

FELKE, J. The poisonous principles of the seeds of *Jatropha curcas* Linn. **Landw Versuchsw**, v. 82, p. 427-430, 1914.

FRANKEL, A.; WELSH, P.; RICHARDSON, J.; ROBERTUS, J.D. Role of arginine 180 and glutamic acid 177 of ricin toxin A chain in enzymatic inactivation of ribosomes. **Molecular and Cellular Biology**, v. 10, p. 6257-6263,1990.

FUNATSU, G.; ISLAM, M.; MINAMI, Y.; SUNG-SIL, K.; KIMURA, M. Conserved amino acid residues in ribosome-inactivating proteins from plants. **Biochimie**, v. 73, p. 1157–1161, 1993.

GAL, Y.; MAZOR, O.; ALCALAY, R.; SELIGER, N.; AFTALION, M.; SAPOZNIKOV, A.; FALACH, R.; KRONMAN, C.; SABO, T. Antibody/doxycycline combined therapy for pulmonary ricinosis: Attenuation of inflammation improves survival of ricin-intoxicated mice, **Toxicology Reports**, v. 1, p. 496-504, 2014.

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) John M. Walker (ed): **The Proteomics Protocols Handbook**, **Humana Press**., p. 571-607, 2005.

GIRBÉS, T.; FERRERAS, J. M.; ARIAS, F. J.; AND STIRPE, F. Description, distribution, activity and phylogenetic relationship of ribosome-inactivating proteins in plants, fungi and bacteria. **Mini-Reviews in Medicinal Chemistry**. v. 4, p. 461–476. 2004.

GUPTA, A.; KUMAR, A.; AGARWAL, A.; OSAWA, M.; VERMA, A. Acute accidental mass poisoning by *Jatropha curcas* in Agra, North India, **Egyptian Journal of Forensic Sciences**, v. 6, p. 496-500, 2016.

GUPTA, R. J.; BRUNAK, S. NetNGlyc – Prediction of n-linked glycosylation sites in human proteins. **In preparation,** 2004.

HARTLEY, M. R.; LORD, J. M. Cytotoxic ribosome-inactivating lectins from plants. **Biochimica et Biophysica Acta,** v. 1701, p. 1-14, 2004.

HOU, X.; CHEN, M.; CHEN, L.; MEEHAN, E. J.; XIE, J.; HUANG, M. X-ray sequence and crystal structure of luffaculin 1, a novel type 1 ribosome-inactivating protein. **BMC structural biology**, v. 7, p. 29, 2007.

HOU, X.; MEEHAN, E.J.; XIE, J.; HUANG, M.; CHEN, M.; CHEN, L. Atomic resolution structure of cucurmosin, a novel type 1 ribosome-inactivating protein from the sarcocarp of Cucurbita moschata. **Journal of Structural Biology**, v. 164, p. 81-87, 2008.

IGLESIAS, R.;PÉREZ, Y.; CITORES, L.; FERRERAS, J.; MÉNDEZ, E.; GIRBES, T. Elicitor-dependent expression of the ribosome-inactivating protein beetin is developmentally regulated. **Journal of experimental botany**. v. 59, p. 1215-23, 2008.

JASHEWAY, K.; PRUET, J.; ANSLYN, E.V.; ROBERTUS, J.D. Structure-Based Design of Ricin Inhibitors. **Toxins,** v. 3, p. 1233-1248, 2011.

JEFFARES, D.C.; PENKETT, C.J; BÄHLER, J. Rapidly regulated genes are intron poor. **Trends in Genetics**. v. 24, p. 375-378, 2008.

KARPIŃSKI, S.; GABRYS, H.; MATEO, A.; KARPINSKA, B.; MULLINEAUX, P.M. Light perception in plant disease defence signalling. **Current opinion in plant biology**, v. 6, p. 390-396, 2003.

KARPIŃSKI, S., SZECHYŃSKA-HEBDA, M., WITUSZYŃSKA, W. and BURDIAK, P. Light acclimation, retrograde signalling, cell death and immune defences in plants. **Plant, Cell & Environment**, v. 36, p. 736-744, 2013.

KIM, Y.; ROBERTUS, J.D. Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography, **Protein Engineering, Design and Selection,** v. 5, p. 775–779, 1992.

KLAIM GJ, JAEGER JJ. Castor Seed Poisoning in Humans: A Review: Technical Report, **Letterman Army Institute of Research**, 1990.

KOLE, C., MATSUMURA, H., & BEHERA, T. K. (Eds.). The Bitter Gourd Genome. **Compendium of Plant Genomes,** 2020.

LAMB F.I.; ROBERTS, L.M.; LORD J.M. Nucleotide sequence of cloned cDNA coding for preproricin. **European Journal of Biochemistry**, v. 148, p. 265-270, 1985.

LASKOWSKI, R.A., MACARTHUR, M.W., MOSS, D.S. AND THORNTON, J.M. *PROCHECK* : a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography,** v. 26, p. 283-291, 1993.

LESCOT, M.; DÉHAIS, P.; THIJS, G.; MARCHAL, K.; MOREAU, Y.; VAN DE PEER, Y.; ROUZÉ, P.; ROMBAUTS, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic Acids Research**, v. 30, p. 325-327, 2002.

LIU, W.Y. Research on ribosome-inactivating proteins from angiospermae to gymnospermae and cryptogamia. **American Journal of Translational Research**, v. 9, p. 5719-5742, 2017.

LORD, J.M.; ROBERTS, L.M.; ROBERTUS, J.D. Ricin: structure, mode of action, and some current applications. **The FASEB Journal**, v. 8, p. 201-208, 1994.

MALETZKI, C.; LINNEBACHER, M.; SAVAI, R.; HOBOHM, U. Mistletoe lectin has a shiga toxin-like structure and should be combined with other Toll-like receptor ligands in cancer therapy. **Cancer Immunology, Immunotherapy,** v. 62, p. 1283–1292, 2013.

MANDAL, S.; VERMA, N.; BORA, N.S.; DEY, P.; ISLAM, J.; DWIVEDI, S.K.; CHATTOPADHYAY, P. Exploration of therapeutic role of montelukast and dexamethasone combination against paraquat induced inhalational toxicity, **Inhalation Toxicology**, v. 32, p. 299-310, 2020.

MARSHALL, R.S.; D'AVILA, F.; DI COLA, A.; TRAINI, R.; SPANÒ, L.; FABBRINI, M.S.; CERIOTTI, A. Signal peptide-regulated toxicity of a plant ribosome-inactivating protein during cell stress. **Plant Journal,** v. 65, p. 218-229, 2011.

MELTON-CELSA, A. R. Shiga toxin (Stx) classification, structure, and function. **Microbiology Spectrum**. v. 9, p. 1–13, 2014.

MOHAMED, M.S.; VEERANARAYANAN, S.; POULOSE, A.C.; NAGAOKA, Y.; MINEGISHI, H.; YOSHIDA, Y.; MAEKAWA, T.; KUMAR, D.S. Type 1 ribotoxin-curcin conjugated biogenic gold nanoparticles for a multimodal therapeutic approach towards brain cancer, **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1840, p. 1657-1669, 2014.

MONZINGO, A. F., AND ROBERTUS, J. D. X-ray analysis of substrate analogs in the ricin A-chain active site. **Journal of Molecular Biology**. v. 227, p. 1136–1145. 1992.

NELLER, K.C.M.; DIAZ, C.A.; PLATTS, A.E.; HUDAK, K.A. De novo Assembly of the Pokeweed Genome Provides Insight Into Pokeweed Antiviral Protein (PAP) Gene Expression. **Frontiers in Plant Science**. v. 10, p. 1002. 2019.

NG, T. B.; CHAN, W. Y.; YEUNG, H. W. Proteins with abortifacient, ribosome inactivating, immunomodulatory, antitumor and anti-AIDS activities from cucurbitaceae plants. **General Pharmacology**. v. 23, p. 575–590, 1992.

OLSNES, S.; PIHL, A. CHAPTER 3 - Toxic lectins and related proteins, Editor(s): COHEN, P.; HEYNINGEN, S.V. **Molecular Aspects of Cellular Regulation, Elsevier**, v. 2, p. 51-105, 1982.

OLSNES, S.; KOZLOV, J.V. Ricin. **Toxicon,** v. 39, p. 1723-1728, 2001.

O'NEIL, M.J. (ed.). **The Merck Index - An Encyclopedia of Chemicals, Drugs, a nd Biologicals**, p. 748, 2001.

PALANIVEL, H.; EASWARAN, M.; MEENA, A; CHANDRASEKARAN, S.; KADER, M.A.; MURALI, A. Structural dynamics and modeling of curcin protein: docking against pterin derivatives. **SN Applied Sciences,** v.1, p. 812, 2019.

PANDEY, V.C.; SINGH, K.; SINGH, J.S.; KUMAR, A.; SINGH, B.; SINGH, R.P. *Jatropha curcas*: A potential biofuel plant for sustainable environmental development. **Renewable and Sustainable Energy Reviews**, v. 16, p. 2870-2883, 2012.

PAPADOPOULOS, J.S.; AGARWALA, R. COBALT: constraint-based alignment tool for multiple protein sequences, **Bioinformatics,** v. 23, p. 1073-79, 2007.

PECK, S.C. Phosphoproteomics in Arabidopsis: moving from empirical to predictive science. **Journal of Experimental Botany**, v. 57, p. 1523– 1527, 2006.

PINTO, A.; CANGELOSI, A.; GEOGHEGAN, P.A.; GOLDSTEIN, J. Dexamethasone prevents motor deficits and neurovascular damage produced by shiga toxin 2 and lipopolysaccharide in the mouse striatum. **Neuroscience**, v. 344, p 25-38, 2017.

PRZYDACZ, M.; JONES, R.; PENNINGTON, H.G.; BELMANS, G.; BRUDERER, M.; GREENHILL, R.; SALTER, T.; WELLHAM, P.A.D.; COTA, E.; SPANUL, P.D. Mode of Action of the Catalytic Site in the N-Terminal Ribosome-Inactivating Domain of JIP60. **Plant Physiology,** v. 183, p. 385-398, 2020.

PURI, M.; KAUR, I.; PERUGINI, M.A.; GUPTA; R.C. Ribosome-inactivating proteins: current status and biomedical applications. **Drug Discovery Today**, v. 17, p. 774-783, 2012.

READY, M. P., BROWN, D. T., AND ROBERT US, J. D. Extracellular localization of pokeweed antiviral protein. **Proceedings of the National Academy of Sciences**, 1986.

REINBOTHE, S.; REINBOTHE, C.; LEHMANN, J.; BECKER, W.; APEL, K.; PARTHIER, B. JIP60, a methyl jasmonate-induced ribosome-inactivating protein involved in plant stress reactions. **Proceedings of the National Academy of Sciences,** v. 91, p. 7012-7016, 1994.

ROY, S.; SADHANA, P.; BEGUM, M.; KUMAR, S.; LODHA, M.L.; KAPOOR, H.C. Purification, characterization and cloning of antiviral/ribosome inactivating protein from Amaranthus tricolor leaves. **Phytochemistry**, v. 67, p. 1865–1873, 2006.

SANDVIG, K.; VAN DEURS, B. Endocytosis and intracellular sorting of ricin and Shiga toxin, **FEBS Letters,** v. 346, p. 99-102, 1994.

SANNER, M.F. Python: A Programming Language for Software Integration and Development. **Journal of Molecular Graphics and Modelling**, v. 17, p. 57-61, 1999.

SCHLOSSMAN, D.; WITHERS, D.; WELSH, P.; ALEXANDER, A.; ROBERTUS, J.; FRANKEL, A. Role of glutamic acid 177 of the ricin toxin A chain in enzymatic inactivation of ribosomes. **Molecular and Cellular Biology**. v. 9, p. 5012–5021, 1989.

SHANG, C.; ROUGÉ, P.; VAN DAMME, EJM. Ribosome Inactivating Proteins from Rosaceae. **Molecules**, v. 21, p. 1105, 2016.

SHESHADRI, S.A.; NISHANTH, M.J.; SIMON BINDU. Stress-Mediated cis-Element Transcription Factor Interactions Interconnecting Primary and Specialized Metabolism in planta**. Frontiers in Plant Science**, v. 7, p. 1725, 2016.

SHINOZAKI, K.; YAMAGUCHI-SHINOZAKI, K. Gene Expression and Signal Transduction in Water-Stress Response. **Plant physiology**, v. 115, p. 327–334, 1997.

SONG, S-K.; CHOI, Y.; MOON, Y. H.; KIM, S-G.; CHOI, Y. D.; LEE, J. S. Systemic induction of a Phytolacca insularis antiviral protein gene by mechanical wounding, jasmonic acid, and abscisic acid. **Plant Molecular Biology**, v. 43, p. 439-450, 2000.

SOUSA, R.B.; OLIVEIRA, S.E.M.; SANTOS, M.C.; LIMA, K.S.C.; LIMA, A.L. Ricina e a Convenção para Proibição de Armas Químicas no Brasil. **Revista Virtual de Química**, v. 6, p. 744-760, 2014.

SPERTI, S.; MONTANARO, L.; MATHIOLI, A.; STIRPE, F. Inhibition by Ricin of Protein Synthesis in vitro: 60S Ribosomal Subunit as the Target of the Toxin. **Biochemical Journal**, v. 136, p. 813-815, 1973.

SRIVASTAVA, M.; GUPTA, S.K.; ABHILASH, P.C.; SINGH, N. Structure prediction and binding sites analysis of curcin protein of *Jatropha curcas* using computational approaches. **Journal of Molecular Modeling**, v. 18,p. 2971-2979, 2011.

STEEVES, R. M.; DENTON, M. E.; BARNARD, F. C.; HENRY, A.; AND LAMBERT, J. M. Identification of three oligosaccharide binding sites in ricin. **Biochemistry***,* v. 38, p. 11677–11685, 1999.

STERLING, T.; IRWIN, J.J. ZINC 15 – Ligand Discovery for Everyone. **Journal of Chemical Information and Modeling,** v. 55, p. 2324-2337, 2015.

STIRPE, F. Ribosome-inactivating proteins. **Toxicon**, v. 44, p. 371-383, 2004.

STURM, M.B.; SCHRAMM, V.L. Detecting Ricin: Sensitive Luminescent Assay for Ricin A-Chain Ribosome Depurination Kinetics. **Analytical Chemistry,** v. 81, p. 2847–2853, 2009.

THOMAS, R.; SAH, N.K.; SHARMA, P.B. Therapeutic biology of *Jatropha curcas*: a mini review. **Current Pharmaceutical Biotechnology.** v. 9, p. 315-324, 2008.

TROTT, O.; OLSON, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**. v. 31, p. 455-461, 2010.

VAN DER HOORN, R.A.L.; WULFF, B.B.H.; RIVAS, S.; DURRANT, M.C.; VAN DER PLOEG, A.; DE WIT, P.J.G.M.; JONES, J.D.G. Structure-function analysis of Cf-9, a receptor-like protein with extracytoplasmic leucine-rich repeats. **Plant Cell**, v. 17, p. 1000–1015, 2005.

VERMA, V., RAVINDRAN, P. & KUMAR, P.P. Plant hormone-mediated regulation of stress responses. **BMC Plant Biol,** v. 16, p. 86, 2016.

WALSH, T. A.; MORGAN, A. E.; HEY, T. D. Characterization and molecular cloning of a proenzyme form of a ribosome-inactivating protein from maize. Novel mechanism of proenzyme activation by proteolytic removal of a 2.8-kiloDalton internal peptide segment. **The Journal of Biological Chemistry,** v. 266, p. 23422–23427, 1991.

WASTERNACK, C.; HAUSE, B. Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. **Annals of Botany**. v. 111, p. 1021-1058, 2013.

WANG, P.; TURNER, N.E. Virus resistance mediated by ribosome inactivating proteins. **Advances in Virus Research,** p. 325–355, 2000.

WANG, S.; LI, Z.; LI, S.; DI, R., HO, CT.; YANG, G. Ribosome-inactivating proteins (RIPs) and their important health promoting property. **RSC Advances**. 2016.

WIEDERSTEIN, M.; SIPPL, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, **Nucleic Acids Research***,* v. 35, p. 407–410, 2007.

WORBS, S.; KÖHLER, K.; PAULY, D.; AVONDET, M.-A.; SCHAER, M.; DORNER, M.B.; DORNER, B.G. Ricinus communis Intoxications in Human and Veterinary Medicine—A Summary of Real Cases. **Toxins**, v. 3, p. 1332-1372, 2011.

WU, L.; GOH, M.L.; TIAN, D.; GU, K.; HONG, Y.; YIN, Z. Isolation and characterization of curcin genes with distinct expression patterns in leaves and seeds of *Jatropha curcas* L., **Plant Gene,** v. 9, p. 34-44, 2017.

WYTYNCK, P., ROUGÉ, P., VAN DAMME, EJM. Genome-wide screening of Oryza sativa ssp. japonica and indica reveals a complex family of proteins with ribosome-inactivating protein domains. **Phytochemistry,** v. 143, p 87-97, 2017.

XU, J.; WANG, H.; FAN, J. Expression of a ribosome-inactivating protein gene in bitter melon is induced by Sphaerotheca fuliginea and abiotic stimuli. **Biotechnology Letters,** v. 30, p. 569, 2008.

YALKOWSKY, S.H.; DANNENFELSER, R.M**. Aquasol Database of Aqueous Solubility**, College of Pharmacy, Univ. of Ariz, Tucson, AZ, , v. 5, 1992.

YALKOWSKY, S.H., HE, YAN. Handbook of Aqueous Solubility Data: An Extensiv e Compilation of Aqueous Solubility Data for Organic Compounds Extracted from t he AQUASOL dATAbASE. **CRC Press LLC**, Boca Raton, FL., p. 1183, 2003.

YAN, X.; HOLLIS, T.; SVINTH, M.; DAY, P.; MONZINGO, A.F.; MILNE, G.W.; ROBERTUS, J.D. Structure-based identification of a ricin inhibitor. **Journal of Molecular Biology.** v. 266, p. 1043–1049, 1997.

ZHANG, X.; PAN, B.Z.; CHEN, M.; CHEN, W.; LI, J.; XU, Z.F.; LIU, C. JCDB: a comprehensive knowledge base for *Jatropha curcas*, an emerging model for woody energy plants. **BMC genomics,** v. 20, p. 958, 2019.

ZHU, F.; ZHOU, YK; JI, ZL; CHEN XR. The Plant Ribosome-Inactivating Proteins Play Important Roles in Defense against Pathogens and Insect Pest Attacks. **Frontiers in Plant Science**. v. 9, p. 146, 2018.

CAPÍTULO 2

SELEÇÃO DO MÉTODO DE EXTRAÇÃO PROTEICA PARA ANÁLISE ELETROFORÉTICA DE SEMENTES DE *Jatropha curcas* L.[1]

[1]Artigo a ser ajustado para posterior submissão ao Comitê Editorial do periódico científico Pesquisa Agropecuária Brasileira, em versão na língua portuguesa.

Seleção do método de extração proteica para análise eletroforética de sementes de *Jatropha curcas* L.

Selection of protein extraction method for electrophoretic analysis of seeds of *Jatropha curcas* L.

Resumo – O objetivo do estudo foi extrair e purificar parcialmente as proteínas presentes nos tecidos de reserva de sementes de *Jatropha curcas* L., selecionar o melhor método de extração para caracterizar o perfil eletroforético proteico de genótipos pertencentes ao Núcleo de Melhoramento Genético e Biotecnológico (NBIO), priorizando as proteínas inativadoras de ribossomos (*Ribosome inactivating proteins* -RIPs). O ensaio de padronização foi conduzido em quatro modos diferentes, os quais distinguiram na forma de extração das proteínas e nas soluções extratoras, sendo estas: I) Tampão Acetato de Sódio (0,45 M) e II), HCl (0,1%)/NaCl (0,6 M), III) para a extração com NaOH (1M) e IV) Tampão fosfato-salino (PBS). A análise por eletroforese (SDS-PAGE) possibilitou selecionar o método de extração proteica com melhor resolução eletroforética, que foi com PBS de albúmen das sementes armazenadas. As concentrações de proteínas das amostras variaram de 2,75 a 5,15 mg/mL entre os genótipos, mas não divergiram estatisticamente entre si pelo teste de *Tukey* a 5% de probabilidade. Pelo perfil eletroforético foi possível identificar as RIPs, com peso molecular de 28 kDa, em todos os genótipos avaliados. Além das RIPs observou-se outras possíveis proteínas de reserva, e supõe-se que seja a globulina 11S. Foi observado um perfil de bandas diferente no genótipo UFRB05 e UFRB15. Quando associado com a qualidade das sementes armazenadas e teor de óleo o genótipo que mais se destacou foi o UFRB05.

Palavras chave – Pinhão-manso, RIPs, SDS-PAGE

Abstract - The objective study was to extract and partially purify the proteins present in seed reserve tissues of *Jatropha curcas* L., select best extraction method to characterize protein electrophoretic profile of genotypes belonging to the Genetic and Biotechnological Improvement Center (NBIO), prioritizing ribosome inactivating proteins (RIPs). The standardization test was conducted in four different ways, which distinguished in form of protein extraction and in the extraction solutions, which are: I) Sodium Acetate Buffer (0.45 M) and II), HCl (0.1%) / NaCl (0.6 M), III) for extraction with NaOH (1M) and IV) Phosphate-saline buffer (PBS). The analysis by electrophoresis (SDS-PAGE) made it possible to select method of protein extraction with best electrophoretic resolution, which was with albumen PBS from stored seeds. The protein concentrations of samples varied from 2.75 to 5.15 mg / mL between genotypes, but did not differ statistically from each other by Tukey test at 5% probability. Through electrophoretic profile, it was possible to identify RIPs, with a molecular weight of 28 kDa, in all evaluated genotypes. In addition to the RIPs, other possible reserve proteins were observed, and it is assumed to be 11S globulin. A different band profile was observed in UFRB05 and UFRB15 genotype. When associated with quality of stored seeds and oil content, genotype that stood out the most was UFRB05.

Key words – Physic nut, RIPs, SDS-PAGE

Introdução

O pinhão-manso (*Jatropha curcas* L.), conhecido também como pinha de purga, purgueira e grão-de-maluco, dentre outros, é uma oleaginosa, pertencente à família *Euphorbiaceae* (ARRUDA, et al., 2004; ABDELGADIR; VAN STADEN, 2013). Um arbusto com folhas pequenas, amarelo-esverdeadas, apresentando frutos secos, com sementes lisas e escuras, formado por uma casca dura e lenhosa, que se desenvolve bem em climas tropicais e sub-tropicais. Atualmente, destaca-se como uma das principais fontes alternativas renováveis para a produção de biocombustíveis no Brasil, por seu óleo possuir alta fluidez e composição (MARTINS et al., 2008; DURÃES et al., 2011; LAVIOLA et al., 2014; WANI et al., 2012).

Além do biodiesel, o pinhão-manso apresenta outras diversas aplicabilidades, podendo ser utilizado pela medicina popular, para o tratamento da paralisia, reumatismo, gota, infecções de pele, bem como pela indústria, na fabricação de tintas e sabão, na agricultura, como cerca viva e na recuperação de terrenos baldios (DURÃES et al, 2009; LAVIOLA et al. 2014; KRISHNAN & PARAMATHMA., 2009).

Seu desempenho agronômico de produtividade varia entre 1.328 a 1.543 Kg ha^{-1} em diferentes regiões, possui maior adaptabilidade, podendo ser cultivada em solos com baixa fertilidade e não irrigados, pois, é considerada uma cultura tolerante à seca (LAVIOLA et al., 2014; LAVIOLA et al., 2017; PU et al., 2011).

De forma geral, a diversidade genética do pinhão-manso ainda precisa ser devidamente explorada, principalmente pela crescente demanda mundial para produção de biocombustíveis renováveis em substituição dos combustíveis tradicionais a base de petróleo (TEIXEIRA et al., 2010; ALMEIDA et al., 2016).

Até o momento, não há cultivares comerciais, pois ainda é uma cultura em processo de domesticação e os estudos de melhoramento genético da espécie ainda estão em fase de desenvolvimento, principalmente, devido a não uniformidade na maturação e colheita dos frutos, além disso, a espécie também tem sofrido com perdas devido a ataques de diversos fitopatógenos (OLIVEIRA et al., 2013; BRASILEIRO et al., 2012; DURÃES et al., 2011; LAVIOLA et al., 2014).

No Núcleo de Melhoramento Genético e Biotecnologia (NBIO) da Universidade Federal do Recôncavo da Bahia (UFRB), Cruz das Almas – Bahia, Brasil, encontra-se o Programa de Melhoramento Genético de *J. curcas* L. O qual mantém clones oriundos de cruzamentos entre meio irmãos por meio de propágulos vegetativos (estacas), conservando os recursos genéticos da espécie e selecionando genótipos com teor de óleo favorável para posterior indicação de uma nova(s) cultivar(es) para produção de biocombustível (PESTANA-CALDAS et al., 2016; ALMEIDA et al., 2016).

Os genótipos obtidos do NBIO são frutos de intensos trabalhos de diversidade genética por meio da caracterização morfoagronômicas e desempenho produtivos dos acessos clonais (PESTANA-CALDAS et al., 2016). Atualmente, o programa de melhoramento desenvolve novos estudos que visam analisar características genotípicas, pois as sementes de *J. curcas* L., além de serem ricas em óleo, possuem também um alto teor de proteínas. No entanto, devido à presença de fatores antinutricionais, como ésteres de forbol e a curcina, uma proteína tóxica inativadora de ribossomos, impede a utilização desta cultura como fonte alimentícia animal (MAKKAR et al., 1997; FERRARI et al., 2009).

A padronização de um método mais simples de extração de proteínas em sementes e estudos do perfil eletroforético de proteínas SDS-PAGE (*Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis*) em diferentes genótipos podem ser importantes para o melhoramento dessa espécie, pois apesar de antigo é o método mais utilizado para análises qualitativas de proteínas, prático e mais barato dentre as técnicas de proteômica.

Com o intuito de aumentar a base de conhecimentos sobre a referida espécie o presente estudo teve como objetivo extrair e purificar parcialmente as proteínas expressas de sementes de *J. curcas* L., selecionar o melhor método de extração e estágio de maturação, e caracterizar o perfil eletroforético proteico de genótipos pertencentes ao Programa de Melhoramento Genética do NBIO, priorizando as proteínas inativadoras de ribossomos.

Material e Métodos

O estudo foi realizado no laboratório do Núcleo de Melhoramento Genético e Biotecnologia (NBIO) da Universidade Federal do Recôncavo da Bahia (UFRB). As sementes foram obtidas do banco de sementes armazenados na câmara fria (-8 ºC, 34-36% de Umidade) do NBIO e do campo experimental do Campus da UFRB, no município de Cruz das Almas-BA, situado na região fisiográfica do Recôncavo Baiano, apresentando as coordenadas geográficas de 12º40'39'' latitude sul, 39º06'23'' de longitude oeste de *Greenwich* e altitude média de 220 m. Segundo a classificação de Köeppen, o clima é do tipo sub úmido, com pluviosidade média anual de 1.170 mm, variando entre 900 e 1.300 mm, sendo os meses de março a agosto os mais chuvosos. A temperatura média anual é de 24,1ºC e o solo é classificado como Latossolo Amarelo Álico Coeso, de textura argilosa e relevo plano (UFRB, 2011).

As sementes oriundas do banco de sementes foram coletadas no ano de 2017, armazenadas e identificadas em tubos de plástico em câmara fria (-8 ºC, 34-36% de Umidade). Já as sementes do campo experimental, foram coletadas e higienizadas para o procedimento de extração no mesmo dia da coleta.

O ensaio de seleção do método foi conduzido em quatro modos diferentes, os quais distinguiram na forma de extração das proteínas e nas soluções extratoras, o tampão de amostra e a concentração do gel foram os mesmos para todos os métodos.

Seguindo o método de Stephan et al. (2010) com algumas alterações, foram utilizadas 6 g da torta desengordurada, obtidas a partir da extração de óleo, com hexano ≥ 99%, pelo método *Soxhlet* (1879) com adaptações, e 30 mL das seguintes soluções extratoras: I) Tampão Acetato de Sódio (0,45 M) e II), HCl (0,1%)/NaCl (0,6 M), as amostras foram maceradas, homogeneizadas em Vórtex (Kylin-Sino Instrumentos de Laboratório CO., LTD.) durante 5 minutos e mantidas sob agitação magnética, sem refrigeração, por 55 minutos, em seguida, foram filtradas em filtro de papel (Mellita) e armazenadas a 4°C; III) para a extração com NaOH (1M) foi utilizado 1,0 g da torta desengordurada, macerada e mantida em banho-maria a 65°C por 10 minutos. Para obtenção do precipitado proteico, a amostra foi filtrada em filtro de papel (Mellita), e centrifugada a 1.000 rpm por 10 minutos, e na sequência, retirado o sobrenadante.

Conforme Lin et al (2010), com adaptações, para a extração com tampão fosfato-salino (PBS), foi utilizado 10mL da solução extratora para 2 g de sementes (tegumento e albúmen), a amostra foi macerada, homogeneizadas em Vórtex (Kylin-Sino Instrumentos de Laboratório CO., LTD.) durante 5 minutos e mantida sob agitação magnética por 55 minutos. Em seguida, foram filtradas e centrifugadas, em centrífuga refrigerada (Vision), à 12.000 rpm por 20 minutos, coletou-se a fase intermediária e estas foram mantidas a 4°C por aproximadamente 12 horas na presença de sulfato de amônio ($(NH_4)_2SO_4$) a 30% e 60% para a precipitação.

Adiante, as amostras passaram novamente por centrifugação a 12.000 rpm por 20 minutos, o sobrenadante foi retirado e acrescentado 750 µL de PBS. Após todas essas etapas as amostras foram dialisadas, por meio de cassetes de diálise (Slide-A-Lyzer) de 10,000 MWCO com capacidade de 3-12 mL.

Após escolha do melhor método foram realizadas extrações em diferentes estágios de maturação da semente e de partes (tegumento e albúmen) separadas. Dentre estes, sementes verdes, maduras (amareladas), secas diretamente da árvore e sementes armazenadas em câmara fria. Só foi possível separar o tegumento e albúmen das sementes secas, nos outros estágios foram utilizadas sementes completas.

Após a seleção do método de extração e eletroforese, foi realizado a extração de proteínas de diferentes genótipos, com o objetivo de caracterizar o perfil eletroforético dos mesmos. Para isto foram utilizados seis genótipos (UFRB03, UFRB05, UFRB09, UFRB11, UFRB13, UFRB15) provenientes do Banco de Germoplasma do NBIO da Universidade Federal do Recôncavo da Bahia – UFRB. O critério de escolha dos genótipos se deu pela disponibilidade de sementes armazenadas.

A quantificação das proteínas foi realizada utilizando o kit de Albumina da LabTest. A leitura das absorbâncias foi realizada em espectrofotômetro Sf325nm (Tecnal©) com comprimento de onda de 630nm, seguindo as instruções de uso do kit de Albumina.

Para obtenção de uma curva de calibração, foi utilizado como proteína padrão, a albumina de soro bovino (BSA) (Sigma-Aldrich, St. Louis, MO, EUA), em concentrações crescentes de: 0,15 a 1,2 mg/mL, diluídas em água ultrapura (Milli-

Q). Que posteriormente, foi misturada com o reagente verde de bromocresol e após 2 minutos procedeu-se às medidas de absorbâncias.

Para a leitura no espectrofotômetro (Tecnal©), as amostras dos genótipos foram diluídas a 1/10 em água destilada, para que as concentrações ficassem numa faixa em que o equipamento fosse capaz de realizar a leitura e que ficassem dentro da curva de calibração. Em seguida, foram misturadas com o reagente, sendo 1 mL de amostra e 100 µL do reagente e então, submetidas à leitura no espectrofotômetro. Na sequência, o valor encontrado foi multiplicado por dez para obtenção da concentração da amostra pura (100%).

A curva de calibração, foi construída utilizando o programa Excel pelo método de regressão linear, a partir da medida de absorbância de crescentes concentrações da proteína padrão para posterior análise do coeficiente de correlação (R^2). O teste estatístico de Tukey foi realizado pelo programa R (R Development Core Team, 2020) à 5% de probabilidade. Após isso, as amostras de todos os genótipos foram diluídas para 2 mg/mL.

A análise por eletroforese (SDS-PAGE), foi realizada seguindo o método de Laemmli (1970) com modificações, nas proporções de 5% e 10% para os géis de empilhamento e de separação, respectivamente. As amostras foram dissolvidas em tampão Tris-HCl 80 mM pH 6,8 contendo SDS 2%, glicerol 10% e azul de bromofenol (0,1%). Foram aplicados 25 µL de amostra em cada poço e 10 µL do padrão de proteínas. A corrida foi realizada a 72 V até que as amostras passassem do gel de empilhamento para o de separação, e 135 V até o final do gel de resolução. Como marcadores de massa molecular, utilizou-se os marcadores de proteínas da Bio-Rad (10 a 250 KDa) e Novex® Pre-Stained (3,5 a 260 KDa). Os géis foram corados com Azul de Coomassie Brilhante R-250 (VETEC).

Resultados e Discussão

A análise dos extratos proteicos de *J. curcas* L.*,* em SDS-PAGE, sob diferentes modos de extração, usando sementes armazenadas, mostrou um perfil de bandas diferente. Evidenciando o poder de extração de cada método. Tecnicamente, as bandas proteicas presentes na faixa entre 20 a 40 KDa

(kiloDalton), dentro do retângulo na figura, sugerem a presença das RIPs (Figura 1).

De acordo com a Figura 1, é possível observar que o perfil eletroforético, a partir de extratos obtidos pelo método de extração utilizando o Hidróxido de Sódio (NaOH), não apresentou nenhuma banda. Apesar dos extratos obtidos por Tampão Acetato e Ácido Clorídrico (HCl) apresentaram bandas mais fortes, a utilização do tampão fosfato-salino (PBS) como solução extratora foi a que apresentou melhor resolução das bandas, com isto, optou-se por utilizar o PBS como método de extração para as próximas etapas (Figura 1).

Além de apresentar uma melhor resolução das bandas no perfil eletroforético, a extração de proteínas utilizando o PBS é mais simples e fácil, pois não é necessário que aconteça o desengorduramento das sementes, reduzindo assim o custo e tempo, além de preservar as proteínas, visto que, na extração de óleo utilizando o método de extração *Soxhlet,* as amostras são submetidas a altas temperaturas, evitando que estas sejam desnaturadas e percam suas funções, o que seria bem mais interessante para testes contra fungos.

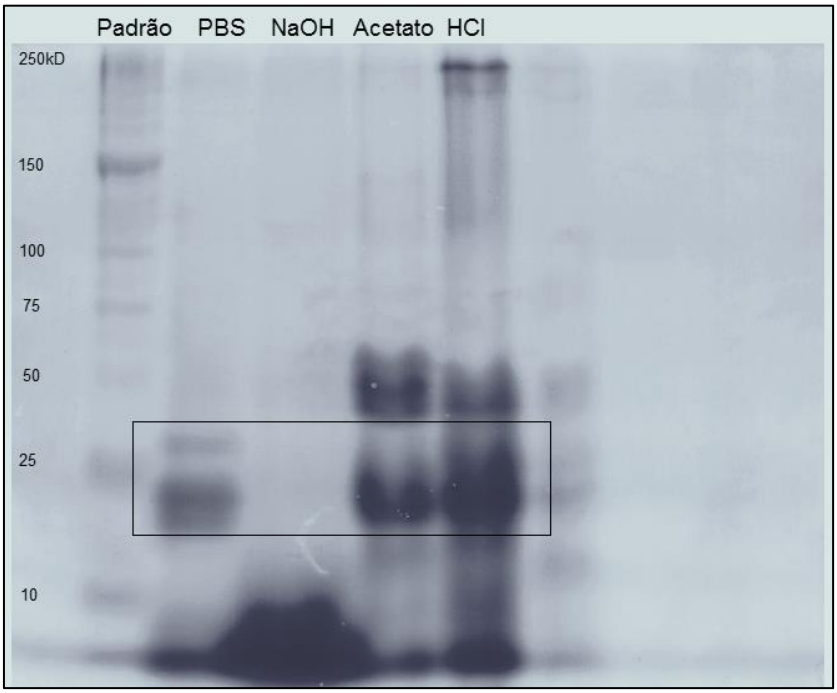


Figura 1. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de *J. curcas* L**. obtidos por diferentes métodos de extração.** Padrão marcador de proteínas da Bio-Rad (10 a 250 KDa); PBS: *Phosphate-bufferid saline*. NaOH: Hidróxido de Sódio. HCl: Ácido Clorídrico. kDa: kiloDalton. **Fonte:** Dados da Pesquisa.

O segundo teste foi realizado com diferentes estágios de maturação das sementes (verdes, amarelas/maduras e secas colhidas da planta e do chão, utilizando sementes inteiras (albúmen + tegumento) para todos os tipos de sementes, sendo testado também somente o albúmen e apenas o tegumento para sementes secas e colhidas na planta e no chão. Nos perfis eletroforéticos, observou-se que o rastro de bandas só apareceu nas extrações de albúmen das sementes colhidas no chão, com o método selecionado com PBS. Com isso, supõe-se que estas proteínas são mais expressas quando atingem a maturação completa e/ou quando são expostas a infecções causadas por fungos (Figura 2). Corroborando com Qin et al (2010), que relata que a curcina presente nas sementes de *J. curcas* L. atinge seu pico de expressão durante o período embrionário maduro.

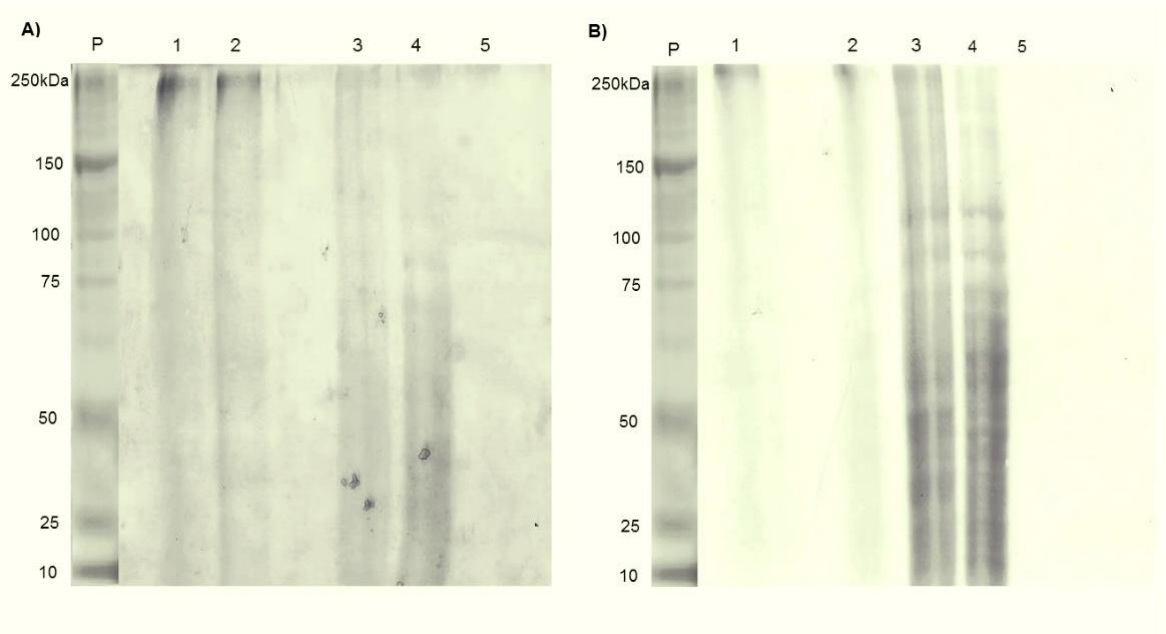


Figura 2. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de *J. curcas* L. em diferentes estágios de maturação. Os extratos proteicos foram obtidos utilizando tampão PBS. A) 1- Sementes de frutos verdes (albúmen + tegumento); 2- Sementes de frutos maduros (albúmen + tegumento); 3- Sementes colhidas na planta (albúmen + tegumento); 4- Sementes colhidas na planta (albúmen); 5- Sementes colhidas na planta (tegumento); B) 1- Sementes de frutos verdes (albúmen + tegumento); 2- Sementes de frutos maduros (albúmen + tegumento); 3- Sementes colhidas no chão (albúmen + tegumento); 4- Sementes colhidas no chão (albúmen); 5- Sementes colhidas no chão (tegumento); P: Padrão de Peso molecular, marcador de proteínas da Bio-Rad (10 a 250 kDa). **Fonte:** Dados da Pesquisa.

As concentrações de proteínas das amostras não divergiram estatisticamente entre si pelo teste *Tukey* a 5% de probabilidade (R Development Core Team, 2020). A maior concentração foi conferida ao genótipo UFRB15 e a mais baixa foi a do genótipo UFRB11. Apesar destas não divergirem, futuramente, tal caráter poderá ser correlacionado com os demais na busca de genótipos superiores (Tabela 1.).

Para a análise do perfil eletroforético das proteínas as amostras foram diluídas para a concentração de 2mg/mL para que não houvesse influência na identificação das bandas entre os genótipos com diferentes concentrações de proteínas.

Tabela 1. Média de Concentrações de Proteínas nas amostras dos genótipos em estudo.

Genótipos	Absorbâncias	Média de concentração em 10% (mg/ml)	Média de concentração total (mg/ml)
UFRB03	0,304	0,428	4,283
UFRB05	0,282	0,367	3,672
UFRB09	0,294	0,400	3,996
UFRB11	0,249	0,275	2,754
UFRB13	0,325	0,488	4,877
UFRB15	0,335	0,515	5,155

Fonte: Dados da Pesquisa.

Na caracterização do perfil eletroforético dos genótipos foi possível observar algumas diferenças no padrão de bandas (Figura 3). Da mesma forma observado em estudos de perfil proteico entre espécies do gênero *Capsicum* obtendo indicativas de diversidade genética, mostrando a eficácia da eletroforese de sementes e folhas para a caracterização, identificação e diferenciação das plantas (OLATUNJI; MORAKINYO, 2015).

Todos os genótipos apresentaram em seu perfil eletroforético bandas com peso molecular de ~ 20 e 30 kDa (Figura 3). Lin et al. (2010), ao extrair e purificar proteínas de sementes de pinhão-manso, obteve extratos com massa molecular semelhante ao reportado no presente estudo, de 28,2 kDa, identificando-a como curcina, proteína inativadora de ribossomos. Com base nisto, sugere-se que a proteína extraída de fato foi a que esperava-se obter.

Esta classe de proteínas é considerada como tóxica, e atuam impedindo o alongamento da síntese proteica, por meio da clivagem de uma adenina localizada

na alça do RNA ribossomal, podendo ocasionar a morte das células (LIU, 2017). E possuem diferentes papéis no sistema de defesa das plantas, atuando contra estresses bióticos e abióticos, porém, o mecanismo de atuação destas proteínas na defesa ainda não foi totalmente estabelecido, no entanto, diversos estudos mostram que as RIPs possuem papeis antifúngicos, antibacterianos, antiviral e inseticida (IGLESIAS et al., 2016; ZHU et al., 2018).

Estudos *in vitro* que comprovam estas funções antifúngicas e antivirais já foram realizados com diversas espécies, e as plantas transgênicas que receberam genes RIPs apresentaram maior tolerância a fungos e vírus (VANDENBUSSCHE et al., 2004; CHOPRA; SAINI, 2014). Por exemplo, um estudo realizado com tabaco (*Nicotiana tabacum*) transgênico, contendo um gene RIP de *J. curcas* L. mostrou maior tolerância ao vírus do mosaico do tabaco (TMV) e um patógeno fúngico *Rhizoctonia solani*, atrasando o desenvolvimento de sintomas sistêmicos de TMV e reduzindo os danos causados pela doença fúngica (HUANG et al., 2008).

Sugere-se que quando as plantas passam por algum tipo de estresse, as RIPs são liberadas de seus compartimentos e inativam o maquinário da síntese de proteínas, levando a morte celular programada, impedindo a proliferação do patógeno (NIELSEN; BOSTON, 2001; AKKOUH et al., 2015).

O genótipo UFRB05 apresentou um perfil de bandas mais fraco na região entre 50 kDa e 60 kDa, apresentando também uma banda mais forte entre 30 kDa e 40 kDa, e uma banda entre 20 e 30 kDa que não apareceu em nenhum outro genótipo. Pode-se sugerir que este genótipo tenha apresentado, além da RIP tipo 1, de 30 kDa, uma RIP tipo 2, com duas cadeias polipeptídicas (Figura 3).

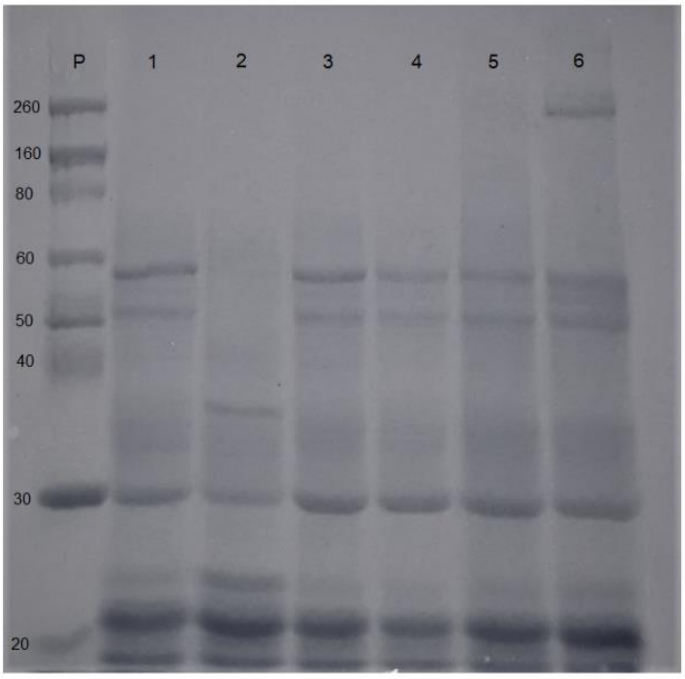


Figura 3. Perfil eletroforético em gel unidimensional SDS-PAGE das frações provenientes de extratos proteicos de sementes de genótipos de *J. curcas* L. Concentração das amostras: 2mg/mL. P: Padrão de Peso Molecular de proteínas Novex Pre-Stained (3.5 a 260 KDa) Genótipos: 1 – UFRB3; 2 – UFRB5; 3 – UFRB9; 4 – UFRB11; 5 – UFRB13; 6 – UFRB15. **Fonte:** Dados da Pesquisa.

Como as sementes utilizadas na pesquisa foram coletadas e armazenadas em câmara fria, em anos anteriores (2016/2017), mesmo em condições controladas, alguns genótipos apresentaram muitas contaminações fúngicas (Figura 4.).

No entanto, o genótipo UFRB05 foi o que apresentou sementes em melhores condições de uso, com poucas contaminações por fungos (Figura 4.). Com isso, supõe-se que estas proteínas presentes neste genótipo podem estar associadas a resistência da semente, garantindo uma melhor qualidade durante seu armazenamento. No estudo de Farias (2018), trabalhando com os mesmos genótipos, verificou-se que o genótipo UFRB05 foi um dos superiores quanto ao teor de óleo.

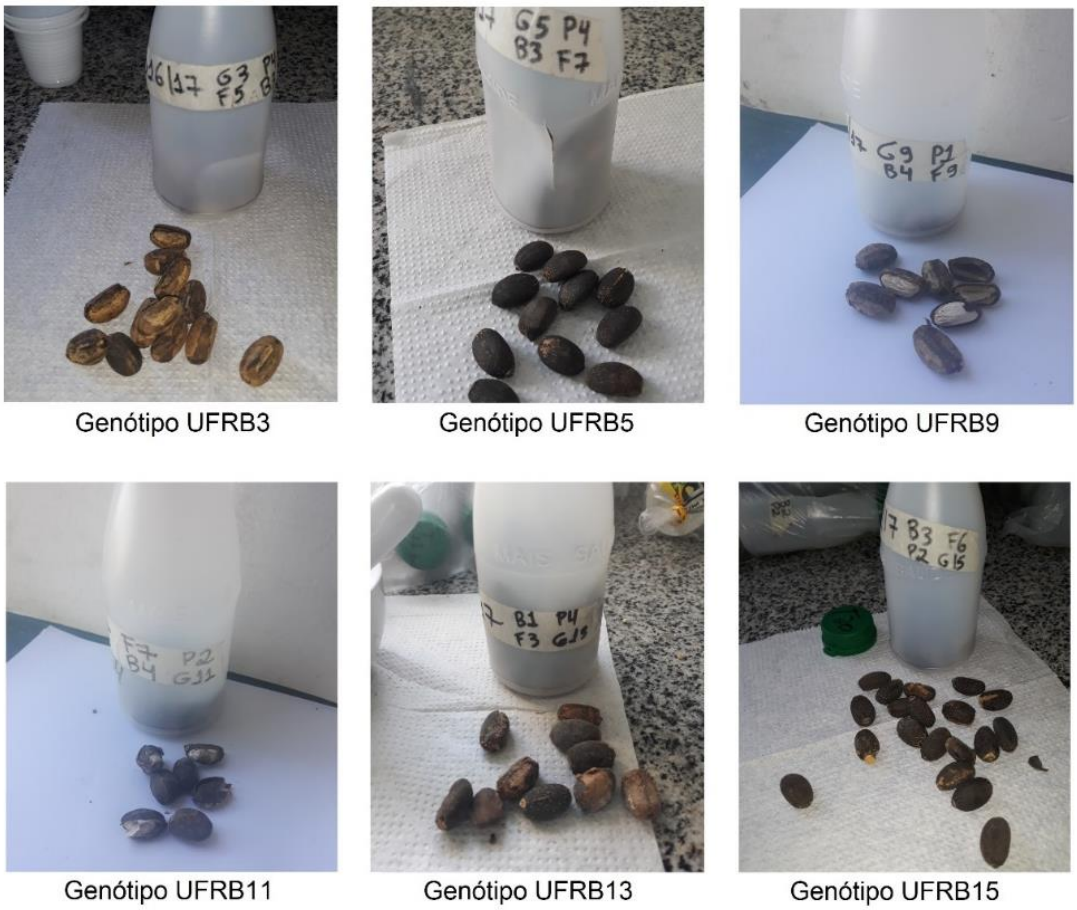


Figura 4. Fotografia das sementes dos genótipos usados na pesquisa. **Fonte:** Dados da Pesquisa.

O pinhão-manso possui em seu genoma uma RIP Tipo II, uma proteína semelhante a ricina, mais citotóxicas do que as RIPs tipo I, com peso molecular de ~60 kDa, tais proteínas são compostas por 2 cadeias polipeptídicas, sendo a cadeia A de domínio RIP e a cadeia B semelhante à lectina, que ao passarem pelo processo de desnaturação, ocorre a quebra das pontes dissulfeto, devido a atuação do dodecilsulfato de sódio (SDS), e estas aparecem separadas com bandas de peso molecular de 30 e 34 kDa (SEHGAL et al., 20010).

Supõe-se que as bandas que aparecem entre 50 e 60 kDa nos géis, sejam as Globulinas, mais especificamente a Globulina 11S, uma proteína de reserva já relatada em diversas pesquisas com perfil eletroforético de proteínas de sementes, a qual também está presente no genoma do pinhão-manso, com massa molecular de ~53,63 kDa (CARVALHO; GARCIA; FARFÁN, 2008; CAVALCANTI; BORA, 2010; LQARI, et al., 2004).

O genótipo UFRB15 também apresentou diferenças significativas, neste caso ele exibiu uma banda de 260 kDa, ainda não descrita na literatura.

Não foi possível purificar totalmente e eliminar as demais proteínas de armazenamento presentes nas sementes, para que ficasse apenas a curcina. Para que isto acontecesse seria necessário que as amostras passassem por mais algumas etapas. Aqui realizou-se uma purificação parcial, seguindo até a etapa da diálise, a qual remove o excesso de sais (íons) do tampão da amostra que atrapalham na caracterização das proteínas e as proteínas de peso molecular menores que 10 kDa.

Tal estudo será importante para direcionar as demais pesquisas que estão sendo desenvolvidas pelo Núcleo. Já foram realizados alguns testes com fungos fitopatogênicos de interesse da cultura, como o *Lasiodiploidia theobromae* e *Colletotrichum gloesporioides*, porém ainda se encontram em andamento. Nas próximas etapas do estudo pretende-se realizar novos testes com novos métodos, para posterior uso destas RIPs, de forma purificada, como biopesticidas conforme já descrito para outras espécies também (SUHARTI; DJAM'AN, 2019). Além disso, este estudo é mais um passo, dentre tantas etapas em um programa de melhoramento genético, para obtenção de nova(s) cultivar(es).

Conclusão

O melhor método de extração é com adição do tampão PBS, expressando melhor resolução de bandas, além de ser de fácil execução e baixo custo.

O perfil eletroforético possui variabilidade no perfil de bandas entre os genótipos. Quando associado com a qualidade das sementes armazenadas e teor de óleo o genótipo UFRB05 destaca-se.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos para realização do presente estudo. E ao Programa de Pós Graduação em Ciências Agrárias da Universidade Federal do Recôncavo da Bahia – UFRB.

Referências

ABDELGADIR H.A.; VAN STADEN,J. Ethnobotany, ethnopharmacology and toxicity of *Jatropha curcas* L. (Euphorbiaceae): a review. **South African Journal of Botany**. v. 88, p. 204-218, 2013.

AKKOUH, O.; NG, T.B.; CHEUNG, R.C.F.; WONG, J.H.; PAN, W.; NG, C.C.W.; SHA, O.; SHAW, P.C.; CHAN, W.Y. Biological activities of ribosome-inactivating proteins and their possible applications as antimicrobial, anticancer, and anti-pest agents and in neuroscience research. **Applied Microbiology and Biotechnology,** v. 99, p. 9847–9863, 2015.

ALMEIDA, A.Q.; SILVA, S.A.; ALMEIDA, V.O.; SOUZA, D.R.; ARAÚJO, G.M. Genetic divergence and morpho-agronomic performance of *Jatropha curcas* L. clones for selection of clonal varieties. **Revista Caatinga**, v. *29*, p. 841-849, 2016.

ARRUDA, F.P.; BELTRÃO, N.E.M.; ANDRADE, A.P.; PEREIRA, W.E.; SEVERINO, L.S. Cultivo de pinhão-manso (*Jatropha curcas* L.) como alternativa para o cenário nordestino. **Revista Brasileira de engenharia agrícola e ambiental**, v. 8, p.789-799, 2004.

BRASILEIRO, B.P.; SILVA, S.A.; SOUZA, D.R.; OLIVEIRA, R.S.; SANTOS, P.A. Conservação de pinhão-manso (*Jatropha curcas* L.) no estado da Bahia, Brasil. **Magistra**, v. 24, p. 286-292, 2012.

CARVALHO, A.V.; GARCÍA, N.H.P.; FARFÁN, J.A. Proteínas da semente de cupuaçu e alterações devidas à fermentação e à torração. **Ciencia e tecnologia de alimentos**, v. 28, p. 986-993, 2008.

CAVALCANTI, M.T.; BORA, P.S. Análise das proteínas e estudo reológico dos isolados proteicos das amêndoas da faveleira (Cnidosculus phyllacanthus (Mart.) Pax. et K. Hoffm.) com e sem espinhos. **Revista do Instituto Adolfo Lutz**, v. 69, p. 243-51, 2010.

CHOPRA, R., SAINI, R. Transformation of Blackgram (Vigna mungo (L.) Hepper) by Barley Chitinase and Ribosome-Inactivating Protein Genes Towards Improving Resistance to Corynespora Leaf Spot Fungal Disease. **Biotechnology and Applied Biochemistry,** v. 174, p. 2791–2800, 2014.

DURÃES, F.O.M, LAVIOLA, B.G., SUNDFELD, E., MENDONÇA, M., BHERING, L.L. Pesquisa, desenvolvimento e inovação em pinhão-manso para produção de biocombustíveis – Brasília, DF. **Embrapa Agroenergia**, v. 17 p. 2177-4439, 2009.

DURÃES, F.O.M., LAVIOLA, B.G.; ALVES, A.A. Potential and challenges in making physic nut (*Jatropha curcas* L.) a viable biofuel crop: the Brazilian perspective. **CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources**, v. 6, p. 43-60, 2011.

FARIAS, L.F**. BIOLOGIA FLORAL, REPRODUTIVA, VISITANTES FLORAIS E DESEMPENHO MORFOAGRÔNICO DE *Jatropha curcas* L.** (Euphorbiaceae). 2018. 94f. (Mestrado em Ciências Agrárias) Universidade Federal do Recôncavo da Bahia, 2018.

FERRARI, R. A.; CASARINI, M. B.; MARQUES, D. D. A.; SIQUEIRA, W. J. Avaliação da composição química e de constituinte tóxico em acessos de pinhão-manso de diferentes origens. **Brazilian Journal of Food Technology**, v. 12, p. 309-314, 2009.

HUANG, M.; HOU, P.; WEI, Q.; XU, Y. A ribosome-inactivating protein (curcin 2) induced from *Jatropha curcas* can reduce viral and fungal infection in transgenic tobacco. **Plant Growth Regulation**, v. 54, p. 115–123, 2008.

IGLESIAS, R.; LUCÍA CITORES, L.; RAGUCCI, S.; RUSSO, R.; MARO, A.D.; FERRERAS, J.M. Biological and antipathogenic activities of ribosome-inactivating proteins from *Phytolacca dioica* L., **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1860, p. 1256-1264, 2016.

KRISHNAN, P. R.; PARAMATHMA, M. Potentials and *Jatropha* species wealth of India. **Current Science**. v. 97. p. 1000-1004, 2009.

LAEMMLI U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227p. 680-5, 1970.

LAVIOLA, B.G.; RODRIGUES, E.V.; TEODORO, P.E.; PEIXOTO, L.A.; BHERING, L.L. Biometric and biotechnology strategies in *Jatropha* genetic breeding for biodiesel production. **Renewable and Sustainable Energy Reviews**, v. 76, p. 894-904, 2017.

LAVIOLA, B.G.; SILVA, D.A.S.; JUHASZ, A.C.P.; ROCHA, R.B.; OLIVEIRA, R.J.P.; ALBRECHT, J.C.; ALVES, A.A.; ROSADO, T.B. Desempenho agronômico e ganho genético pela seleção de pinhão-manso em três regiões do Brasil. **Pesquisa Agropecuária Brasileira**, v. 49, p. 356-363, 2014.

LIN, J.; ZHOU, X.; WANG, J.; JIANG, P.; TANG, K. Purification and characterization of curcin, a toxic lectin from the seed of *Jatropha curcas*. **Preparative Biochemistry and Biotechnology**, v. 40, p. 107–18, 2010.

LIU, W.Y. Research on ribosome-inactivating proteins from angiospermae to gymnospermae and cryptogamia. **American Journal of Translational Research**, v. 9, p. 5719-5742, 2017.

LQARI, H.; PEDROCHE, J.; GIRÓN-CALLE, J.; VIOQUE, J.; MILLÁN, F. Purification and partial characterization of storage proteins in *Lupinus angustifolius* seeds. **Grasas y Aceites**, v. 55, p. 202, 2004.

MAKKAR, H. P. S.; BECKER, K.; SPORER, F.; WINK, M. Studies on nutritive potential and toxic constituents of different provenances of . *Jatropha curcas,* **Journal of Agricultural and Food Chemistry**, v. 45, p. 3152-3157, 1997.

MARTINS, C.C.; MACHADO, C.G.; CAVASINI, R., Temperatura e substrato para o teste de germinação de sementes de pinhão-manso, **Ciência agrotécnica**, v. 32, n. 3, 2008.

NIELSEN, K., BOSTON, R.S. Ribosome-inactivating proteins: A plant perspective. **Annu. Rev. Plant Physiol. Plant Mol. Biol.** v. 52. p. 785–816, 2001.

OLATUNJI, T.;MORAKINYO, J. Crude Protein Profiling of Varieties of *Capsicum annuum* and *Capsicum frutescens* using SDS-PAGE. **IOSR Journal of Pharmacy and Biological Sciences**, v. 10, p. 64-71, 2015.

OLIVEIRA, V.D.; RABBANI, A.R.C.; SILVA, A.V.C.; LEDO, A.S. Genetic variability in physic nuts cultivated in Northeastern Brazil. **Ciencia Rural**, v. 43, p. 978-984, 2013.

PESTANA-CALDAS, C.N.; SILVA, S.A.; MACHADO, E.L; DE SOUZA, D.R; CERQUEIRA-PEREIRA, E.C.; SILVA, M.S. Genetic divergence through joint analysis of morphoagronomic and molecular characters in accessions of *Jatropha curcas*. **Genetics and Molecular Research,** v. 4, p. 1-11, 2016.

PU Y.; TREASURE T.; GONZALEZ R.W.; VENDITTI R.; JAMEEL H. Autohydrolysis pretreatment of mixed hardwoods to extract value prior to combustion. **BioResources**, v. 4, p. 4856–70, 2011.

R Development Core Team. R: **A Language and Environment for statistical Computing, reference index version 3.6.3.** Vienna, Austria: R foundation for statistical Computing Retrieved from http://www. R-project.org. 2020.

SEHGAL, P.; KHAN, M.; KUMAR, O.; VIJAYARAGHAVAN, R. Purification, characterization and toxicity profile of ricin isoforms from castor beans. **Food and Chemical Toxicology***,* v. 48, p. 3171–3176, 2010.

STEPHAN, M.P.; SILVA, B.M.; AZEVEDO, T.L.; ASCHERI, J.L.M. Metodologia de extração de proteínas em torta de mamona e pinhão manso para análise por eletroforese (SDS-PAGE). **Embrapa Agroindústria de Alimentos**: Rio de Janeiro, RJ, 2010.

SUHARTI, T.; DJAM'AN, D.D. POTENSI RIP (RIBOSOME INACTIVATING PROTEIN) YANG BERASAL DARI TUMBUHAN SEBAGAI BIOPESTISIDA. **Buletin Eboni,** v.1, p. 33-39, 2019.

TEIXEIRA, L.S.G.; COUTO, M.B; SOUZA, G.S.; ANDRADE FILHO, M.; ASSIS, J.C.R.; GUIMARÃES, P.R.B.; PONTES, L.A.M.; ALMEIDA, S.Q.; TEIXEIRA, J.S.R.

Characterization of beef tallow biodiesel and their mixtures with soybean biodiesel and mineral diesel fuel, **Biomass and Bioenergy**, v. 34, p. 438-441, 2010.

UFRB. Cruz das Almas. **Universidade Federal do Recôncavo da Bahia.** Quarta, 28 dezembro 2011. Disponível em < https://www.ufrb.edu.br/pgea/noticias/17-cruz-das-almas> Acesso em: 24 de julho de 2020.

VANDENBUSSCHE, F.; PEUMANS, W.J.; DESMYTER, S.; PROOST, P.; CIANI, M.; VAN DAMME, E.J. The type 1 and type 2 ribosome-inactivating proteins from Iris confer transgenic tobacco plants local but not systemic protection against viruses. **Planta,** v. 220, p. 211-221, 2004.

WANI, T.A.; KITCHLU, S.; RAM, G. Genetic variability studies for morphological and qualitative attributes among *Jatropha curcas* L. accessions grown under subtropical conditions of North India. **South African Journal of Botany**, v. 79, p. 102-105, 2012.

ZHU, F.; ZHOU, Y.K.; JI, Z.L.; CHEN X.R. The Plant Ribosome-Inactivating Proteins Play Important Roles in Defense against Pathogens and Insect Pest Attacks. **Frontiers in Plant Science**, v. 9, p. 146, 2018.

CONSIDERAÇÕES FINAIS

As análises *in sílico* e em bancada, das proteínas inativadoras de ribossomos presentes no pinhão-manso, realizadas no presente trabalho contribuíram para a caracterização desta família gênica de proteínas e para o estabelecimento do melhor protocolo de obtenção do extrato proteico parcialmente purificado.

Foram encontrados doze membros distribuídos no genoma, sendo onze caracterizada como RIP tipo I e uma como RIP tipo II, os elementos da região promotora mostra que esses genes podem ser regulados por diversos fatores, a análise de expressão *in sílico* prevê que estes genes são expressos nos diferentes tecidos da planta. Permitiu realizar modelagem e acoplamento molecular das RIPs com novos possíveis inibidores. Este estudo foi o primeiro a testar *in sílico* a dexametasona e seus derivados como inibidores de RIPs. Tais resultados podem servir posteriormente para estudos com abordagens biotecnológicas que visam a inativação de proteínas pertencentes a esta classe.

O estudo de extração de proteínas de sementes permitiu identificar diferenças no perfil de bandas entre os genótipos. Os extratos proteicos parcialmente purificados foram obtidos por meio da extração com PBS, um método de extração que não desnatura as proteínas, apresentando melhor resolução de bandas, além de ser de fácil execução e baixo custo.

Esta pesquisa é mais um passo, dentre tantas etapas em um programa de melhoramento genético, para obtenção de nova(s) cultivar(es). E posteriormente, pretende-se realizar estudos a fim de testar o potencial antifúngico ou fungiostático desses extratos proteicos.

APÊNDICES

Apêndice 1: S1 Figure. Nucleotide sequences of RIP from *Jatropha curcas* L. from the JCDB: Jatropha curcas Database. Exons are indicated green and intron are indicated in red.

>JcRIP1|JCDBG01176|NW_012124076.1:8868..11787
CTATAAAAGCAGGTGATGGGGGAAGGCAAAAGACCATCTCTTGCTCTCTTCTTCTTTACTTCCCCGTTTGCTCAGTTGCT
TTCTTTGTAAGTAATATAGAAGCCTCTGCCCTTCTTTGTTGTTGACAAATTCCATTTTTTGTTTTACTATTAGCATGTTA
ATTTCTAGCTTCTGGAAATGAGTTTATTATCCTTTATATGATAAACTTGTGACCATTCTATCTCTTTTTAATTATTTTTA
TAATTTTATGCAATTCTATTAAAATAATCGTATTCGTATAATGATATTTGTGTTTCTTCATACAACTGGACAGGTGAAAT
CAATATGAAAGGTGGAAAGATGAATCTCTCCATTATGGTGGCTGCATGGTTTTGCTGGAGTAGTATTATATTCGGATGGG
CATCGGCTAGGGAAATAGTTTGTCCATTCTCATCAAACCAAAACTACAAAGCTGGTTCCCCTCCAACTTTAACCATTACT
TATGACGCTACTACTGATAAGAAAAACTACGCCCAGTTCATTAAAGATCTAAGAGAAGCATTTGGCTTCAGTTATTCAAG
CCATGAAATACCAGTCTTACGGGCCACAGTTGCTCCAAATCAGAAATTTATTGTAGCCAAAGACATAAATGTAGCGAATT
TAGAAGTATCATTAGGATTAAACGTCGTTAATGCGTATTTAGTGGGTTATAAGGTAGGAGGTACTTCCTATTTCTTTAAC
GATCCGGAATCTTTGGCTGATGCAAAAACATATCTTTTCACAGACACAAAGCAACAAACGCTATCATTTACTGGTAGCTA
TGCAGATTTTCTATCTAGGGCAAACGTACACAGAGAGGATGTGGATTTAGGGGTGCAGGCATTAGATAATTACATATATA
CACTTGAAAAAAGTTCAAAGCCAGCAGACATTGCTAAACCTCTAGTTGGTTTTATCGAAGTGGTTCCAGAGGCAGCAAGA
TTCAAATATATTGAGAAAAAAGTATTAAGTCAAATTAGCAAAACCTTTAGGCCGGGTGGTGACATAATTAGCCTTGAGAA
CAACTGGGGAGACCTCTCTTATCAAATACAGAAATGTGTAAATGGTGTATTTCTGAAGCCAGTTCAATTACAACGTGAAA
ACTATACCAATATCCTAGTGAACAATGTCACCCAAGTAGCAGGTGTCATGGGAGTCTTGTTGAATGCAGTCAATTACAAA
GTCTGAATGGAAGAAATTATTTTCAACTACCAAAAGTGGCTGCCATGGCTTTAATCCTACTTTTGCTCTATATATAGAGT
AGCATAAATAAAGGACAACAAATTTATTATTATTGTTGCTAATGCTATATGCTATTTCCCTGTAATATCCTCATCTTTCC
AATGTATGAATATGATGATGAATTATATATGACAAATAAAGTTTCTGCTAGTTCTTAATT

>JcRIP2|JCDBG01177|NW_012124076.1:17904..20514
TATGAAAGGTGGCAAGATGAACCTCTCTATTATGGTGGCTGCATGGTTTTGCTGGAGTTGTATTATATTCGGATGGGCAT
CGGCTAGGGAAATAGGTTGTCCCTTCTCATCAAACCAAAACTACAAAGCTGGTTCCACTCCAACTTTAACCATTACTTAT
GACGCTGCTGCTGATAAGAAAAACTACGCCAACTTCATTAGAGATCTAAGAGAAGCATTTGGCTTCAGTTATTCAAGCCA
TGAAATACCAGTCCTACGGGCCACGGTTGCTGCAAATCAGAAATTTATTGTAGCCAAAGTCATAAAGGTAGCGAATTTAG
AAGTATCATTAGGATTAAACGTCGTTAATGCATATTTAGTGGCTTATAAGGCAGGAGGTACATCCTATTTCTTTAACGAT
CCCGAATCTTTGGCTGATGCAAAAAAATATCTTTTCACAGACACAAAGCAACAAACGCTATCATTTACTGGTAGCTATGC
AGATTTTCTATCTAGGGCAAACGTACACAGAGAGGATGTGGATTTAGGGGTGCTGGCATTAGATAATTACATATATATAC
TTCACAAAAGTTCTCAACCAGCAGACATTGCTAAACCTCTAGTTGGTTTTATCGAAATGGTTCCAGCGGCAGCAAGATTC
AAATATATTGAGAAAAAAGTATTAACTCAAATTAGCGAAACCTTTAGGCCGCGTGGTGTCATAATTAGCCTTGAGAACAA
CTGGGGAGACCTCTCTTATCAAATACAGAAATCTGTAAATGGTATATTTCTGAAGCCAGTTCAATTGCAACGTGAAAACT
ATACCAATATCCTAGTGAACAATGTCACCCAAGTAACAGGTCTCATGGGAGTCTTGTTGAATGCAGTCAATTACAAAGTC
TCAATGGAAGAAATTATTTTCAACGACCAAAAGTGGCTGCCATGGCTTTAATCCTACTTTTGCTCTATATATAGTAGCAT
AAATAAAGGACAACAAATTTAGTATTATTGTTGTTGTCCAAACATGTTGCCAATGATATATGCTATTTCCCTGTAATATC
CTCATCTTTCCAATGTATGAATATGATGATGTATTATATATGACAAATAAAGTTTCTACTAGTTCTTAATT

>JcRIP3|JCDBG01178|NW_012124076.1:102057..104940
TTCTTCTTTACTTCCCCGTTTGCTCAGTTGCTTTCTTTGTAAGTAATATTGAAGCCTCTGCCCTTCTTTTTTGTTGACAA
ATTCCATTTTTTTGTTTTACTAATAGCATGTTAATTTCTAGCTTCTGGAAATGAGTTTATTATACTTTATATGATAAACT
TGTGACCATTCTATCTCTTTTTAATCATTTTTATAATTTTATGCAAATCTATTATAATAATCGTATTCGTATAATGATAT
TTGTGTTTCTTCATACAACTGGACAGGTGAAATCAATATGAAAGGTGGCAAGATGAACCTCTCTATTATGGTGGCTGCAT
GGTTTTGCTGGAGTTGTATTATATTCGGATGGGCATCGGCTAGGGAAATAGTTTGTCCATTCTCATCAAACCAAAACTAC
AAAGCTGGTTCCACTCCAACTTTAGCCATTACTTATGACGCTACTACTGATAAGAAAAACTACGCCCAGTTCATTGAAGA
TCTAAGAGAAGCATTTGACTTCAGTTATTTAAGCCATAAAATACCAGTCTTACGGGCCACGGTTGCTGCAAATCAGAAAT
TTATTGTAGCCAAAGTCATAAATTCTGGGGACATAGAAGTATCAGTAGGATTAAACGTCATTAATGCATATCTAGTGGCT
TATAAGGCAGGAAGTAATTCCTATTTCTTTAACGATTCGGAATCTTTGGCTGATGCAAAAAAAAATCTTTTCACAGACAC
AAACCAACAAACACTAGCATTTACTGGTAGCTATGCAGATTTTGAATCTAGGGCAAAGTTACATAGAGAGGAAGTGGATT
TAGGAGTGGTGGCATTGGATAATTACGTATATACACTTGAAAAAAGTTCTCAGCCAGCAGACATTGCTAAACCTCTAGTT
GGTTTTATCGAAATGGTTCCAGAGGCAGCAAGATTCAAATATATTGAGAAAAAAAATATCAACTCAAATTAGCAAAACCTT
TAGGCCGCGTGGTGACATAATTAGCCTTGAGAACAACTGGGGAGACCTCTCTTATCAAATACAGAAATCTGTTGATGATG
TATTTCTGAAGCCAGTTCAATTGCAACGTGAAAACTATACCAATATCCTAGTGAACAATGTCACCCAAGTAAAAGGTCTC
ATGGGAGTCTTGTTGAATGCAGTCAATTACAAAGTCTCAATGGAAGAAATTATTTTCAACTACCAAAAGTGGCTGCCATG
GCTTTAATCCTACTTTTGCTCTATATATAGTAGCATAAATAAAGGACAACAAATTTAGTATTATTGTTGTTGTCCAAACA
TGTTGCTAATGATATATGCTCTTTCCCTGTAATATCCTCGTCTTTCCAATGTATGAATATGATGATGAATTATATATGAC
AAATAAAGTTTCTACTAGTTCTTA

>JcRIP4|JCDBG01180|NW_012124076.1:170314..172743
TATGAAAGGTGGAAACATGAAGCTCTGCATTATGGTGGCTGCATGGTTTTGCTGGAGTACTATTATATTTGGATCGGCAT
CCGCTAGGGAAAGAGCTTGGCCATTCTCATCAAACAACAACTACGCAGCTGATTCCACTCCAACTTTAACCATTACTTAT
ATTCCTGATGAGGATGAGCAAAACTACGCCAAGTTTATTACAGATCTAAGAGAAACATTTGGCTCTAGCGGTTTAAGCCA
TGGAATACCAGTCTTACGGGCCACAGTTGCTGCAAATCAGAAATTTTTTGTAGCCAAAGTCATAAATGCTGGGGATATAG

AAGTATCAGTAGGATTAAACGTCATTAATGCATATCTAGTGGCTTATAAGGTAGGAAGTAATTCCTATTTCTTTAACGAT
TCGGAATCTTTGGCTGATGCAAAAAAATATCTTTTCACAGACACAAAGCAACAAACACTAGCATTTACTGGTAGCTATGC
AGATTTTGAATCTAGGGCAAAGGTACATAGAGAGGAAGGGGATTTAGGAGTGGTGGCATTAGATAATTACATATATGATC
TTCAAAAAAGTTCTCAACCAGCAGACATTGCTAAACCTCTAGTTTCTTTTATCCAAATGGTTTCAGAGGCAGCCCGATTC
AAATATATTGAGAACAAAGTATTAGATCAAATTAGCCAAACATTTAGGCCGCGCGGTGACATACTTAGCCGGGAGAACAG
CTGGGACGACCTATCTTATCAAATACAGAAATCTGTAAATGATGTATTTCTCAACCCAGTTCAATTGCAACTTGAAGACT
ATTCCTTTTACCAAGTGAACAATGTCAACCAAGTAAAAGATGATATGGGAATCCTGTACAATGAAGCCAATCACAAAGTC
TCCATGGAAGAAATTATTATCAACAGCCAAAAGTGGCTGCCATTGCTTTA

>JcRIP5|JCDBG10548|NW_012124293.1:18323..26358
TATGAAACGGGGAAACACGAAGCTCTGCATTATGGTGGCCGCATGGTTTTGCTGGAGTACTATTATATTTGGATGGGTAT
CGGCTAGGGAAATAGATTGTCCATTCTCATCAAACGACCAGTACACAGCTGATTCCGTTCCAACTTTAACCATTACTTAT
GATGCTGATGAAGATAAGCAAAACTACGCCAAGTTTATTAAAGATCTAATACAAGCATTTGGCTCCAGCGGTTTAAGCCA
CGGAATACCAGTCTTACGGGCCTCAGGGGCTGCAAATCAGAAATTTTTTGTAGCCAAAGTCATAAATTCTGGGGATATAG
AAGTATCAGTAGGATTAAACGTCATTAATGCATATCTAGTGGCTTATAAGGTAGGAAGTAAATCCTATTTCTTTAACGAT
ACGAATTTTTTGGCCGATGCAAAAAAATATCTTTTCACAGACATACAGCAACAACCAGAATTACCATTTACTGGTAACTA
TGCAGATTTTATATCTAAGGCAAAGGTACATAGAGAGGAAGTGGATTTAGGAGTGGTGGCATTAGATAATTACATATATG
ATCTTCAAAAAAGTTCTCTACCAGCAGACATTGCTAAACCTCTAGTTTCTTTTATCCAAATGGTTTCAGAGGCAGCCCGA
TTCAAATATATTGAGAACAAAGTATTAGATCAAATTAGCCAAACATTTAGGCCCCGCGGCGACATACTTAGCCGGGAAAA
CTGTTATAAATTAATCTCAAATAATTCCACAATAACCTAGATCGTGCAAACCACTAGATTCAACAAAATAGATTGAATAA
GAAAAGGTTGCGTAACTGACATCGTTGTTCGTCGATCTTGATGTGCAGAATTAGGTTTGCAGCCTTCCAGAGAATTGAAG
AATCGATTTAGCGGAAGCAGTACGAAAGGAACGTAAAACTGAATCGTATTCAACTCCTCTGGGGACCACGACCTTCATAT
ACCTGTAGCGTTAGTACAATTCCTATTTCGGTCCTTTAACGAAATAGAAATTATATAATGGTCTCTACACTTAATATCTA
TAATTAGTTGATACCTCAATAGTCAATTATTCACAATCAAGACCACATTTGAAATTGACTTTACAATTGGACTGCTGTAC
ATATTTAATTTTACATATTAAGCCCACAATAGAAAATTAAATCCAACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

GGATTCATGATAACTAGTATCAATGTCTAACATATATAGATTATCACATTCATTCAAGCAACCAGTACCCAAAACTTTTG
AATTCACAGATAAAGTAAATACACTGCCTCCAAAAGAACACTGAAAACCAGATTTGTCCAAAAAGGAAACAGAAATCAAA
TTCCTTCTAAAAGAGGGAATAACAAAAGTATCTATAAGATCCAAAAAATAACCAGTACTTAAACATAATCTAAAATGTCC
AATTGCCTCTACTTCAACCGGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNN

TTTGGGATCTTGTTCAATTACCTGAAGGAAAGAAACCCATTGGTTGCAAATGGATATTTAAAACCAAAAGGGATTCTCAA
GGTAATGTGGAAAGATATAAAGCACGTCTTGTAGCAAAAGGCTTTACTCAAAAGGAAGGGATTGACTATAAAGAAACCTT
CTCTCCAGTTTCTTCGAAGGATTCTTTTAGGACAATCATGGCCTTAGTTGCAAATCTTGATCTTGAGCTTCATCAAATGG
ATGTTAAGACTGCGTTTCTCAATGGCGACATTGATGAAACAATATATATGGTGCAACCAGAAAATTTTGTAATTGGAGAT
CCAAAACATCTGGTATGCAAATTAAAGAAATCCATTTATGGACTCAAGCAAGCTTCTCGTCAATGGTACTTTAAGTTCAA
CCAAGTAATTGTCTCGTTTGGCTTTGAGGCAAATGCTGTTGAAAATTGTGTGTACCATAAGTTCAGTGGGAGTAGGTTTA
TCTTTTTGGTCTTATATGTGGATGACATACTTCTTGCTACTAATTGTAAAGTCTTGCTGCATGAAACCAAAGTTTTTCTT
TCAAAGCATTTTGAGATGAAAGATTTAGGTGAAGCTTCTTTTGTTTTAGGGATTCAGATACATCGAGACCGTTCTCGAGG
TATTCTTGGATTATCACAAAAGAGCTATATTGATAAAGTGCTCAAGAGATTTAGTATGCAGAATTCTAGACCATTGGATA
CCCCTATTGCAAAAGGTGATAAATTTAGTAATCAACAATGTCCAAAGAGTGAACTTGAAATTCAGGAAATGCAAAAGGTT
CCTTATGCGTCAGCTGTAGGAAGTCTTATGTATGTTCAAGTTTGTACTCGTCCAGACATTTCTTTTATAGTTGGCATGTT
AGGTAGATATCTCAGCAATCCAGGAATGGATCATTGGAAAGCAGCCAAAAGAGTCATGAGATATCTTCAGAGGACTAAGG
GACATATGCTCACATATAGGAGATCAGATCAGCTTGAGATCATTGGGTATACTGATTCTGATTTCGCTGGATGCCAAGAT
AGTAGAAAATCCACATCTGGCTACGTTTTTCTATTGGCTGGCGGTGCTATTTCTTGGAAGAGTGTGAAACAAACTCTAGT
AGCAACCTCCACCATGGAAGCTGAATTTGTAGCATCTTATGAGGCATCCAATCAAGCTATTTGGCTGAGAAACCTTGTCA
CTGGGCTGCGTGTAGTAGGTGGAATAGAAAGACCACTGAAGTTGTATTGTGACAGCAGGTCAGCCGTTTTATTCTCCAAC
AACAACAGGAGCTCTTCTCGATCAAAACACATTGAGATTAAGTTCTTGGCTGTGAAGGAGAGGGTACAGAGTGGACAGAT
ATCCTTAGAGCATATTGGAACAAACTCTATGATTGCGGATCCGCTTACAAAGGCAGTGATACCGAAAGTTTTCCATGAGC
ACACCGCTCATATGGGTATCGTTTTGTTTGAGACTGTCTAGTTTTAGTGGGAGTTTGCTCTTTATGCATACATTATGTTT
TTGAATGCTCTATTTTATTTTGGATATGTTTTATTATGAAATTTTGTTGAGTACTCATTTTGGTTGATGGTTACACTCTG
TATTATTATTATTGTTTTGCTTGATCTCACTTTTGTTGGTGACCAGTTGGAAACAGACATGATGATCACATTACATGTAG
TTTTCATGCTACACTTCTATAATGAATCTATGTCATTTAGGGGTATTGGTAATGTAGTTGATGATGAACTTAGTCATGAG
ATCTATGGCGAATGTTTCTTCAGCTCTTTATCAATATAACTAATGGACGAGATTGTTTTTGGATACCTTGTATGACTGCG
TTTGTTGAGCTCATTTTGGTGTATGTATTTTATTTAACATATGTTCGCCCAAGTGGGAGATTGTTGGATTTAATTTTCTA
TTGTGGGCTTAATATGTAAAATTAAATATGTACAGCAGTCCAATTGTAAAGTCAATTTCAAATGTGGTCTTGATTGTGAA
TAATTGACTATTGAGGTATCAACTAATTATAGATATTAAGTGTAGAGACCATTATATAATTCCTATTTCGTTAAAGGACC
GAAATAGGAATTGTACTAACCCTACAGGTATATAAAGGTCGTGGTCCCCAGAGGAGTTGAATACGATTCAGTTTTACGTT
CCTTTCGTACTGCTTCTGCAAAATCGATTCTTCAATTCTCTGGAAGGCTGCAAACCTAATTCTATACATCAAGATCGACG
AACAACGATGTCAGGTACGCAACCTTTTCTTATTCAATCTATTTTGTTGAATCTAGTGGTTTGCACGATCTAGGTTATTG
TGGAATTATTTGAGATTAATTCTTAACACAAATTAGCCAAAAATTTAGGCCGCGCGGTGACATACTTAGCCGGGAGAACA
GCTGGGACGACCTATCTTATCAAATACAGAAATCTGTAAATAATGTATTTCTCAACCCAGTTCAATTGCAACGTGAAGAC
TATTCCTTTTACCAAGTGAACAATGTCAACCAAGTAAAAGATGATATGGGAATCCTGTACAATGAAGCCAATCACAAAGT
CTCCACGGAAGAAATTATTATCAACAGCTAAAAGTGGCTGCCATTGCTTTAATAGTACTATGGTGTAAATATAGTAGCAT
CAATAAA

>JcRIP7|JCDBG12049|NW_012124368.1:36636..39480
ACTTGCTTGTAAGTCAATACTATGAAGCTTTTCTCAAAAAACTTTTTTTAACATCTCAACTGCAATTTCTGCTATTTCTG
TTCATTAAATTATAGTACTCTTGAAAACCTTTTTAGTTCTAGCAACTTGAATTTAATAATATTTGTTCTAATTTAATTGC
TCAAGCTCTTTTAATCGTTGGATTTTCAGTTCAGTTAGCAATCTTTTGTTGATATAAACATGTGCATTGATTTTTGACAG
GGGAAATATGGAAGGAAATATGAAGGTTTGTGTTGTGGTGGCGACATGGTTGTGCTGGACTATGATATTTGGATTGGCTA
GGGTTATTCACCCATCAGCAATCCATAACTGCACAGTCACAGTCGATGAAATTCCCTCTGTGAGCTTTACTATTACTCGC
ATCCCCGGTGACGATACCGCTGGCTACAAGCAGTTTATGGTAGATCTTCGAGAGAAATTGGCTTCCGGTACTACTAGCCA
TGGAGTACCAGTGTTACGCTCCACAGCTTCAAGAGAAGCCAAATATCTTCTAGTCAATATTATAAATTCAGGGAATAAGG
AAATCACATTAGGATTGAATGTTATTAGCGCATATATTTTGTCCTATAAGGTGGGATATAATTCCTATTTCTTTAAAGAC
AAAGCTGAACTGAAGGATGCGCAAAAATACCTTTTCACAGACACAAGGCAAACCACACTCGACAAATTTTCTGGTAACTA
TGACTCTTTCAAAGCTGAAGGAGGAGATCGAGAAACAACCGATTTAGGAATCGGGCAATTAGACTCTCATATTTATACAC
TTCATAAATCGACATTGCCAAAAGATATCGCTAAACCTCTTGTTTGTATTATTCAAATGGTTTCAGAAGCGACCCGATAC
AAATATATTGAGAATAAGATAATAGATAAGATTAGTGGAAGTTTTAGGCCGAAGCTTGACACAATCACTCGTGAGAACAA
ATGGGAAGACCTTTCTGATGGAATTGAAAATGCTGATGCGAAGGGAAATTTTAAAACAGAAGTTAGATTGCAGAAGGAAG
ATGGTAAGGAGGATATTATATCCAGCGTTAATCAAATAATAGGTGATATGGGAATTTTGCTGTATCAAAAGAAGAAGAGC
TACAATCCAAGCTTTGGACAAACAACATTTGGAAACCTTATCACAACAAAATCAAATTGCTGCCTCGACTTTGATTAATA
TTACTTATATCTGTATTATGTCAAAATAAAGGACATGTCATATTATTACTCTTGTTGTCTAAATGTTCTCCTATTTTCGT
CAATTTCCAGATTTGATCCGTGATTAGTAATTGATGTGAGGATCAATTTGGATATATGTGAGTTT

>JcRIP8|JCDBG12048|NW_012124368.1:19157..21532
TATGAAAGGCAATATGAACAAGGTGTTGATTGCGCTAGTAACATGGTTTTGCTGGACTGTGGTGTTTGGATTGGCTAGAG
TTTACCCGTTACAAACAAATGGTAATTCTGTCGCCATAGCCAGCTTTACCATTACTCGGTATGGCGAAACAGATAAAGCT
GACTATTTACAATTCATGGAAGATATACGAAGTAAATTAGGTACTGGTACTCCAAGTCATGGAATTCCAGTATTACCTGT
CTCAGTttctaaaaatcaaaagtataCTACAGTGGAACTTATAAATTCAGGGAACAAGAAAGTAACAATAGGTTTAAATG
TGCTGGACCTTTATGTGGTTGCTtacaaattagaaaatcaaaattcctATTTCTTCAACGATCCAAAATTCAAGGACGCA
ACCGACTATCTTTTCAAAGATACAAAACAAAACCCACTTAAAATCAGTGGTAACTATGAATCCTTGAAAACCCAAGGAGG
TGATAGAGAGACAGTTTATTTAGGGGTTGGGCAATTAGACTCTTTCATTTATACACTCTACGAATCGACATTGCCAAAAG
ATATTGCTAAGCCTCTTGTTTGTATTATTCAAATGGTTTCAGAAGCGACACGATACAAATATATTGAGAATGAAGTGGTG
AAACATATTAGCGACCGCTTTCTGCCAAAGGGCGACATAATCAGCCGTGAGAACAAATGGCAGGACCTTTCTGAAGCAAT
TCAAAAATCTGTAGCAGATGAATTTACAAAACCAGTTCAATTGCAAGCTCCAGATTATAATGTGTACAATGTGTACCGTG
TCCAACAAGTGAAAGACGATATGGGACTCTTGCTCAACCAAGCCAATTCCGGCGAGACTGAAGACATTACTACCTA

>JcRIP9|JCDBG12053|NW_012124368.1:96996..99440
TGGAAGGAAAGATGAAGGTCTGGGTTGTGGTGGCGACATGGTTGAGCTGGACTGTTATTTTTGGATTGGCTAGGTTTATT
TGCCCATTGGCAATCCATAACCATACAGTCGATGCAATTCCCTCTGTAAGCTTTACTATTAGTCGTATCCCCGATGATGA
TAAGACTGGCTACAAACAACTTATGGTCGATCTTCGAAACAAGTTGGCTTCCGGTACAACCAGCAACGGAGTACCAGTTT
TACGCACCACAGCTTCAAAAGAAGCCAAATATCTTTTAGTCAATATAATAAATACAGGAAATAAGGAAATCACATTAGGA
TTAAATGTTATTAACGCATATGTTTTGGCCTATAAAGTGGGAGATAATTCCTATTTCTTTAACGATCCAACAGAATTGAA
GGACGCACAAACTTATCTTTTCAAAGACACAAAGCAAAACACCATAAAAATGACAGGTAGCTACGACTCTCTCAAAGCTC
AAGGAGGAGATCGAGAAACAATGGATATAGGAATTGGGCAATTAGACTCTCACATTTACACACTCCATAAATCGACAGCG
TTAAAAGACATTGCTAAGCCTCTTGTTTGCATTATTCAAATGGTTTCAGAAGCTGCCCGATTTAAATACATTGAGAAGAA
AATAATAGATGAAGTTGAGGGAGGTTTTACGCCCAAACTTGATGTAATCAGTCGCGAGAACAACTGGGGAGGCCTTTCAG
AAGGAATTGAAAATGCTGACAAAAATGGAAAGTTCAAAACGACAGTTCGATTGCAAAACGAAGATAGTTCAGCGAAGGTT
ATTTCTAAGGTTGATGAAATAATAGTAGAGATGGGAGTTTTGTTGTATGTAAAGAAGAAAAGCTACATTCCAAGCTTTGA
ACAAACAATATTTGGAATTGGAAACCTTATCCTCAACCAAATCAACTTGCTGCCTCGGCTTTAAT

>JcRIP10|JCDBG12050|NW_012124368.1:52227..54494
TGGAAGGAAATATGAAGGTTTGGGTTGTGGTGGCGACATTGTTGTGCTGGACTATTATATTTGGATTGGATAGGGTTATT
CACCCATCAGCAATCCATAACTGCACAGTCATAGTCGATGAAATTCCCTTTGTGAGCTTTACTATTACTGGTGACCCCGG
TGACGATGCCGCTGGCTACAAGCAGTTTATGATAGATCTTCGAGAGAAATTGGCTTCCGGTACTACTAGCAATGGAGTAC
CAGTGTTACGCTCCACAGCTTCAAACGAAGCCAAATATCTTCTAGTCAATATTACAAATTCAGGGAATAAGGAAATCACG
TTAGGATTGAATGTTATTAGTGCATATATTTTAGCCTATAAGGTGGGAGTTAATTCCTATTTCTTTAAAGACAGATCTGA
ACTGAAGGATGCGAAAAAATACCTTTTCACTGGCACAACGCAAACCATACTCAGTAAATTTTCAGGTAACTATCACTCTT
TGAAAGATGAAGGAGGAGATCGAGAAACAACCGATTTAGGAATCCAGCAATTAGACTCTCATATTTATACACTTCATAAA
TCGACATTGCCAAAAGAAATTGCTAAACCTCTTGTTTGTATTATTCAACTGGTTTCAGAAGCTACCCGATACAAATATGT
TGAGAATAAGATAATAGATGAGATTATTGTAGGTTTTAGGCCGAAGCTTGACTCAATCACTCGTGAGAACAACTGGGGTG
ACCTTTCTAAAGGAATTGAAAACGCTGATGAGAAGGGAAATTTTTAAA

>JcRIP11|JCDBG12054|NW_012124368.1:107222..109639
TGGAAGGAAATATGAAGGTCTGGCTTGTAGTGGCAACATGGTTATGCTGGACTATTATATTTGGATTGGCTAGGGTTATT
AACCCATCTGCAATTCATAACTACACAGCCGATGCAATTCCCTCTGTAAGCTTTACCATTACTCGTATCCCCGGTGATGA
TAAGACTGACTACAAGCAGCTTATGGTAGATCTTCGAAAAAAATTGTCTTCTGGCACTACTAGCAATGGAGTACCAGTGT
TACGCTCCACAGCTTCAAAAGAAGCCAAATATCTTTTAGTCAATATAATAAATTCAGGGAAGAAGGAAATCACGTTAGGA
TTAAATGTTATTAATGCATATGTTTTGGCTTATAAAGTGGGAGATAAATCTTATTTCTTTAACGATCCGACTGAATTGAA
GGATGCACAAACTCATCTTTTCAAAGACACAAAGCAAACCGCCATAAAAATTACAGGTAGCTACGACTCTCTCAAAGCTC
AAGGAGGAGATCGAGAATCAGTAGATTTAGGAATCGGGCAATTAGACTCTCACATCTATACACTCCATAAATCGACAGCG
TTAAAAGACGTTGCTAAGTCTCTTGTTTGTATTATTCAGATGGTTTCAGAAGCAGCCCGATTTAAGTCTATTGAGAATAA
GATAGTAGATAAGATTGATGGAAGTTTTAAACCGAAGCTTGATATAATCACTCGTGAGAACAACTGGGGTGACCTTTCTG
AAGGAATTCAAAACGCTGATAAGAAGGGAAATTTTAAAACAAAAGTTAGATTGCAGAAGGAAGATGGTAAGGAGGATATT
ATATCTAACGTTAATCAAATAATAGGAGAGATGGGAATTTTGCTGTATAAGAAGAAGAAGATCTACAATATTCCAAGTTT
TGGACAAACAAATTTTGGAAACCTTATCCAGAACTAAG

>JcRIP12|JCDBG12052|NW_012124368.1:82429..85584
CATGAAGGTATGGACATGGCTTTGTTGGACTGTAATATTTGCATCAACTTGTATCTGCACATCAAAAGCGGAAGATGGTA
CAAACGACAATAGCCTTGGCTATCCGACGCTGAATTTCACTATCAACGTCAAAGACAAGGAGGGCTATACAAGATTTCTT
CAATCTGTAAGAGATCAATTGACAAGCGGAAACCAAATCCATGGAATATCGGTATTGCCCGATATATCCACTTTGCCTGT
TTCGCAACGGTTTCTTTTACTCAAACTTTCAAGTTCAGCTACAACACCCATTACATTAGCTCTGGATGTTGCCACTGCAG
GTGTAGTTGCTTATGGGTTTCAAAATCAATCCTACTTCTTTAATGATAAGCTTGCATTTTCTAATCTTTTCAATGACACA
AAGCAAAATACGCTTCCATTTGGAAGTAGCTATGGTGAGCTTGAAAATAGTACAGGTATGCCAAGAAGTAAAATTGATCT
GGGACTATTGCCACTAAATGAAGCTGTCAGCAACATTGTTGCTAACAACATTAAGACAGAGGACTTGGCTTCCTGCTTGA
TAGTTGTTATGCAAATGGTTTCAGAAGCAGCGAGATTCAGATACATCGAACATCAAGTGCGTTGGAGCACTGAAAAAGGG
CAAAGATTTCGCCCAGATGGTACCATCTACAGCCTTGAGAACAACTGGGAAGCTCTCTCAACTGCAGTTCAAGAATCTAG
AGGAGGAGTTTTTTCCGAGCCAGTTCAACTACAAAGGTTGAACGTTAAGTTCAATTTGGATAGCGTGACTAGAGAATTAG
TTGCCAACTTGGGGCTCATGTTATTTATCAGTTGCAAAGATAATTCAACCTCATCTGGTGGTGGTGGTTGTGCAACTGAT
TTTGAGCCTACAGTGCGTATTCTTGGTCGAAATGGTCTATGTCTTGATGTGAGCGATGAAATTTATAGTGATGGAAACAA
AATTCATATGTGGCCATGCAAATCTAAACCTGCAGCCAATCAACTATGGACCTTGAAAAGAGATGGAACACTGCAATCTA
ATGGGAAGTGCTTGACGGCTAATGGGTCTAATCCAGCAGGAAATAATGTAGTGATTTATGATTGTCAAACTGCAATGCAT
AATGCTACTCAGTGGGCAATATGGGACAATGGAACCATCATAAACCCTCAATTTGGGCTAGTTTTGACGGTAAATTCAGG
AGATAAAGGTGCATTATTAACGTTAGAGAAAAACATTTATGCTTCTAGTCAAGGTTTTCTTGCCACTAATAATACACAAC
CTTTTCGCTCACCCATTCTTGGGCAAAATGATCTTTGCTTGCAAACAAACGGAACAAAAGTATGGGTTGTGGAGTGTGTA
AGTAACAGAACAGAGCAAAAATGGGCTTTGTATGCAGATGGTTCGATACGGCCACAACAGAATAAAGATGACTGCCTTAC
TTGCAATAAATTAGATGCCAAAGGAACGATCTTTACTATCGATTCTTGTTCCTTGGCATCGTCTGGTCAGCGTTGGGCGT
TTCGAGACGATGGCATAATTTTCAATTTACAGAATGGAATGGTAATGGATGTTAAAAAGATAGATCCAAGCCTTGAGGAA
GTAATCATTTGGCCATTCAATGGAGGACGCAACCAGAGATGGTTACCGCTACTTTG

Apêndice 2: S2 Figure. Promotor sequence of the physic nut JcRIPs genes**.** (1500 bp upstream).

>JcRIP1|JCDBG01176|NW_012124076.1:8868..11787|-
GGGAAACTGCCTTCCCGGCCCTTTATTTTGTCAATTTTTTTTCTTTATTATTATTATTATTACTATTAGCAACCAAAC
GACAGTTTCATAAGATGTAAAATAAAGACAAACAACAACATATCCCAATGAAAAAGAATTATTGAAAATCCAAGGAAA
TGCCGAGTCATAGCGGGTGTGGGGCGCTGATTTTGATGACCGATGATTAGAAAAAACAAAAATTGATAAAAGACAAAC
GAGAAATATGACACGTTTAGAAGTAATTTCAAGCTTTTGAACACTATCATCTTTAGTGTTTTAAGCTCATCAAAGAAA
GATTACTCAAGATTTATTCAACATAAAAAAGGAAAAAATAAATGCATAAGAAGGGAGGAAATCCATATTGAAATCCAA
AAAAATCAAAGGTTTTTGTAATACCCCGAATTTTTATTTTCGAAACTTTCGAATTTTTATTAAATCAAAAATCTTGAT
TTTCAACTATATGGGCTATAAATACCAAAAATCTCAATTCGAGCCCAAGTTCGCGAGTTGGGCCTAAGCCCAAATATA
TCATTGGAATTTTCAATTTGAGAACAGTTAAAATTTTCGAGAAATAATATGAAACTCTAACAATTAATCTATTATGTA
TATATATAAATTAATTTCCCAAAAAAATCAAATATGATTACTAATTATATTTTTATTTTTATGGAGGAAATGTTACGA
AACCTAATTAAAATAATCAAAATAAATCAATCCGGATCCGGAAATCAAATTAATTTTTCGAAAACCACACAAATAAAT
TGAAAACCAAGTGAAAAACATACAAACCGAGCAAAATCAAGCCAAATCGTTGAATCAGAACAAATCAGTTCGAATCGA
ACCAGAATGATCTAGAAACCGACTTGGTTTCTTTTTCTTCCTCTCTTCTCCTTCTTCTTCTCCAGTCACGCGACAGCC
TCCTCCTCCCTTGTTCACTGTGTTTTCCGGCCATCGACCACAGATTCCGACAACCAAGCACCCTAACTGCAAGTGCTC
GACGAGAGGAACCCACCTATGGTCTCCGTTCACGTAGTAGACGCCGGAAGCGAAAGGAATCATGCTTGAAGTGACGGT
ACCCTCCGATTTCTATTCGCTGGCGATTCTCGCCGTTTCGGGCAAATATAGCCAGCGAGAGGACCTTCGATCGACTCC
TCTTGGTCTTGCGTGTGTTTTCCATCCCTTGGTTCCGATGATTGGGCTCCGGTGGTGAAACCACCACTAATGGTGGCA
TTAAGCTTCGATCGGCGTTGGTTGGGCCAAACGGACCGTAGACTCAACAATTGGAGCATGGATTTGAGTTCTTGGTGA
AATTTTGATTTTAAAGGTGTATGAAGTGCATGATTCTGGGCAAGAACGACTTTCTGGCCATCGTGCACGGGGGCTAGT
GGCGGCGGTCCGGCGACACTCCGGCGAGGCCCCGAGAGACAGAATTCTCTTCTCTTCGCGCTGATAATGGAATCGTCG
CCCAATTCTCAATTGGTT

>JcRIP2|JCDBG01177|NW_012124076.1:17904..20514
TATTGATTTCACCTGTCCAGTTGTATGAAGAAACACAAATATCATTATACGAATACGAT
TATTTTAATAGCTTTGCATAAAATTATAAAAATAATTAAAAAGAGATAGAATGTTCACAAGTTTATCATATAAAGGAT
AATAAAGTCATTTCCAGAAGCTAGAAATTAACATGCTATTAGTAAAACAAAAAATGGAATTTGTCAACAAAAAAGAAG
GGCAGAGGCTTCACTATTACTTACAAAGAAAGCGAGAGATGGTCTTTTGCCTTCCCCCATCACCTGCTTTTATAGGGA
AACTGCCCTTTATTTTGTCAATTTTTTTTCTTTATTATTATTATTATTATTATTATTAGCTACCAAACGACAGTT
TCATAAGATGTAAAATAAAGACAAACAACAACAAATCCCAATGAAAAAGAATTATTGAAAATGCAAGGAAATGCCGCG
TCATAGCGGGTGTGGGGCGCTGATTTGATGACCGTTGATTAAAAAAAAACCAAAATTGATAAACCAAAATATGACACG
TTTAGAATTTTTTTTTTATTATTATTTTTTTATTGGTGAAACCTTAGATGAAAAATAATGAAAGAAATGTAACCCCAA
AAAAAAAAAGTAATGAAAGAAAAGAAAACAAAGTCTTCTATTCCAATATTTATTTGTAAAAAGGGAAAATAATAAAAA
GGAAACTATATTTTTTAAAAAAATTAAATTTATTAAAAATTAGCATTATCAAACCAACAAACTAAACTCAAACTCAAA
TCATCTGTTTCTCTAAGGCCCATCTTAAGCTTAGTATTGCTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAA
AAAACTTGGAAACTCTTTAATAAAGATTATAGCTTAAGTGGGTCATAATAGTTCCATAATAAACATATGATTTTGACA
AAAGACATTAATCTTCTAGGTAAAGACCAATCTATTTTCAAACAGAATGATAAAATAAAATTTTTGAACTAGTGTCTT
GAAGTGTCTTTGAGAAAACTTGCTACCTCCTATAGAAGAGTTTGAAATTATCAAATTTTAATGGAGGAAGGTGGTGCT
ACCAAGATCTATTGAAAAAAAAAATTATCAGCATTCAAAGAATTGAAGTTTTCGACTGCAATCAATAAATGATTCATC
TTATATTTTAGATTATCAAGAAAGAAAAATCACAAACTACGTAAAAACTCTTTAATAATGAGAAATAACTTCATCAAA
GGAGAGAAAACGTAACAAATCAAAATTGGACAACCCCACTAGGAGGATATCAAAACATTGAAGGTCTGTTAACTATGA
TAGATATGAGAAATAAAAATTAAGATTTGTTTTAAACCATCCTACTCTAGAAAGAAAACTAAATAAAGATAAAGAAAA
CCGATGAAAAGAAAACGAAATCCTAATCATGCTCTTCACTATATTAAGGATGCCCTAGATGCTAGATTATTTTTCCAA
TCTAATTTTTCTTTTAGCTTAAAGCAGAAAAGTCACC

>JcRIP3|JCDBG01178|NW_012124076.1:102057..104940
AAGAGAGCGAGAGATGGTCTTTTGCCTTCCCCCATCACCTGCTTTTATAGGGAAACTGCCCTTTATTTTGTCAATTTT
TTTTTCTTTATTATTATTATTTTTTATTATTATTTTGATTATTAGCTACCAAACGACAGTTTCATAAGATGTAAAAAT
AAAGACAAACAATAACAACTGTCAAAAAAAAAAAAAAAAAAGAAAGACAAACAATAACAAATCCCAATGAAAAAGAATT
ATTGAAAATCCAAGGAGAAAAAACAAAAATTGATAAATGCCGCGTCATAGCGGGTGTGGGGCGCTGATTTGATGACCG
GTGATTAGAAAAAACAAAAATTGATAAAAGACAAACGAGAAATATGACACGTTTATAATTTTTTTTATTATTATTTTT
TTATTGGTGAAACCTTAGATGAAAAATAAGGAAAGAAATGTAACCAAAAAAAAAAAAAAAGTAATGAAAGAAAAGAAAA
CAAAGTCTTCTATTCCAACATTTATTTGTAAAAAGGGAAAATAATAAAAAGGAAACTATAATTTTTTAAAAAATTAAA
TTTATTAAAAATTAGCATTATCAAACCAACAAACTAAACTCAAACTCAAATCATCTAGTTTTTCTAACGCTCATCTTA
AGCTTAGTATTGATTTATTTATTTATTTATTTATTTTTATTTTTATTTTTTTACAAACTTCTCAAGCCTTTAATAAAG
ATTATAGCTTAAGTAGGTCATAATATTTCCATAACCAATGTATGATTTGGTACATTAATCTTCTAGGTAAAGACCAAT
CTATTTTCAAATAGAATGATAAAATAAATTTTTTGAACTAGTGTCTTGAACTTTGTCCTTGAGAAAACTTGTTACCTC
CTATAGAAGAGTTTGAAATTATCAAATTTTAACTTAGGAAGGTGGTGCTACCAAGATCTATTGAAAAAAGAAATTATC
AGCATTCGAAGAATTAAAGTCTCCACTGCAATCAATAAATGATTCATCTTATATTTTAGACTTTTAAGAAAGAAAAAT
CACAAATCACGTAAAAACTCTTTAATAATGAGAAATAACTTCATCAAATGAGAGAAAACGTAACAAATCAAAATTGGA
CAACCCCACTAGGAGGGTATCAAAAACATCGAAGGTCTATTAACTATGATAGATCCGAAAAATAAAAACTAAGATTTG
TTTTAAACGATCCTGCTCTAGAAAGAAAAGTAAATAAAGATAAATTTAAACGATCCTACTCTAGAAAGAAAACTAAAT
AAAGATAAAGAAAACCGATGGAAAAAAAAAAATCCTAATCATGCTCTTCACTACATTAAGGATGCCCTAGATGCTAGA

TTATTTTTCCCATCTAATTTTTCTTTTAGCTTAAAGCATCTAATGGTTACGACTGGGCCGGCTGTTTAGACACACACA
AATCTATTACAAGTTATTGTGTTTACCTTAGATAATGTAATTAGTCTGCTAGAACTTAACAAACAGTCAGACTTTTTT
ACGTTCCTCATCAAAAGC

>JcRIP4|JCDBG01180|NW_012124076.1:170314..172743
ATTGATTTCACCTGTCCAAATGCTTAATTAATTATCATAGAAATAGAAACATTTGCATGAAGAAAAACAGATATCATT
ATATAAATATAGTTATTTAATAAAATTGCATGAAATTATAAAAATAATTAAGAAAAGAAAGAATGATCGTAGTTTTTA
ACATTTTGGGTAAAGGCTGTTAGTGAAAGAAGAAATGGGAAGTTAATGTTAGAGGAATTTATCAAAAAGATAACTTAC
AAAGAGCGCAAGAGCAAGAAGGAATTAATGAAGAAGAGAGCAAGAGATGGTGTTTGGCCTCTCCCTTATAACCTGCTT
TTATAGGGAAACCGCTCTCACTTTCTTTTCTCAATTTTTTTTCTTTATTATTATTATTATTATCATCATTGAAAATCC
AAGCAGATTCTGTTCTCATAGCACGTTGTGTGGCGCTGATTTGTTGATAAATGATTAGAAAAAAATAAAAGTTAATTG
ATAAAAGACAAACGAACAAGGAATATGCTCATTTAATGTAATGTTTGGACATTATGTCCTCTACTGGTGCATTAATTA
ACAACGAGATGAAGAGATTGCTAATTAATTATTACAAATTAGTAGAAGTTAGTTGATTTAAAATTTAAAATTAAATAT
AAAATAAAAGCCAAATCAATTGCATTTAATTCATATTGACATGGCTTCATAGGATAGAGGCAATATTGTGCTTTTGTA
TATTTACTAGTCATGTACTCCACCACGTGCAGATTATATAAATTTATATCTTTTATTGTCTTAAATAGAAAATTTTTC
AGAAAATTCTTGTAAAATTATTCGCTTAGTCACCTCTTATAGATAACAATCCAAAAAACATAGAACACATAAACTATG
AATAAGTGATAGTAATAAAATGTCTACAATATTAATATCTTAATCAATTTCAACAAAAAATATGATATCGGGAACCAT
GATTAGTGAATAAATTAATATTCAAACATGAGCTAAAAATATAAAAGAAGTGAGGAATGTCAAAAAGCCTAGTGTCAA
TGCATATTGCACTCATATAATAATTAAATACTAAAAAGGATCAAAAGATTCAATAAAATGATAAAAAAAAAAAGCTAAA
GCAAGAGAGAGAAAATTAGATACTCTATTTTTTTTTTTATAATAAGAGATACGCATTTTTAGTTTATAGTATATCCAT
ATTATTTAATAGGGTATTCAAAACTTTAATTAATACCTTTAGAATTTTCAAGCTATAAACTAAAAATTTAATTAGCAT
AATAAATTATCATAGTTAATAGGTATTGAAATAATCTACACTCACATATTTTTCCCTTGCCAAAATCAGCACTTCAAA
ACTCTATATTTATGCCATGTCAGTTTTTATGGTTTTGTTTTTATATATTTATATAGATATAGATATAAATAGATAGAT
AGATAGATATGATTTAATAACAATTTAACCCCCTAATTAAACTGTTACTTAACCATGTGACACTCTTAAATTTTATTT
TAATCCACCAAACACTTC

>JcRIP5|JCDBG10548|NW_012124293.1:18323..26358
CATATTGATTTCACCTGTCAAAATGCTTAATTAATTATCATACAAATAAAAACATTTGCATGAAGAAACACAGACATC
ATTCTAAAAATATACTTATTTAAGAAAATTGCATAAAATTATAAAAATAATTAATAAGAGAAAGAATGGGAAATTAAC
GTTAAGAGGAATTTATCAAAAAAATAACTTACAAAGAGAGCAACTGAGCAAGAAGGAATTAATGAAGAAGAGAGCAAG
AGATGGTGTTTAGCCTCTCCCTTAACCTGCTTTTATAGGTGCTCTCGCTTTCTTTTCTTAATCTTTTTCTTTATTTTT
ATTTTCTCATTGTTAATAATATTATTGAAAATCCTATTTGCAGAATTTCCAGCAATCCTATTTGCAATGGAAGTCTTC
TCTTCCTCGCATCTGAATTCTTTTCCCTTGCTAGTTGAGTCTGATTCAAAAAATGCGATTGGTTGGATGCTGAACACA
AATAATAGACTGTGGCATTGCTGAAAAAAATTTAACAGGTTGGATTGTTTGATCTCATCTTTAGTACACATTGCTTTT
TGTCATATTAAAAGAAGCAAGAATAATTTTGCAGATATATTGGCTAAAAAAGGAGTTTATCCACAGAGTGATTTTTTG
GCTTGGATATGAAGTTTTTGTATTTGGTCTACGTAAACTATTTATTCTCTTATCTCAAGAGAAGAAATAATTGTCTAT
TATGAACTTTGTACTGTTTTGATTTTCAGATTTTTTACTTAACATTTTCGGGTTAGGTTGGGTTTGTTCAGAGTGAGC
ATATTGGGTCGAGATTAGATATCTGAAAATATTAAGTTTTTTTTTAGATATATAAAAATTAATAAAATAATTAAATTGC
CTATAAAAAAAATATTATTATTGAAAATCCAGGGTGACACTGATTTGTTGACCAATGATTAGAAAAAAAAAACACGAGT
TAATTGATAAAAGACAAACGAACCAGGAATTATGCCCATTTTTATATTTAATGTAATGCATTAATTACTGATAAATAT
TTGTATAATCTGACATGATAATATATATATCTCTTATGCATCTCTTTTTTATATGTGAGGTTTATCTAAAGCTGATTC
ATATATATCATATTATATTTTCTAACTAATTATTAATATCAATTGAATAAAAAAAAAAAAAAAAAGATAGACGTGCTA
GGTTTGTTTTACAAGTCAAATACATTTTTCCACGTGAAAAAATTGATAAAAATGTTTGTATTCAATAATTTTTATGTG
AAAATCAAGATATAACCTATTGATAAATATGTGTATAATCTAGCAATACAGTTAGATTTGAATTCTTTCTTCCAATGC
TTAAAAATATATATTTCAATAATTCTCATATTAAGAAATTTACAAATTTGATACAGAATATCTACTGTTAAAAATCTC
ACATCGATTAAATAATCAATTCATAAAAAGCATATAATGAAGACAAGCTAGCTAAATAATTAATTATTTATAACATTT
TAGCATTGGTTGGCCAAA

>JcRIP6|JCDBG01181|NW_012124076.1:184500..191606
AAGAGAGCAAGAGATGGTGTTTAGCCTCTCCCTTATAACCTGCTTTTATAGGGAACTGCTCACACTTTCTTTTCTCAA
TTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTGAAAATCTAAGTAGATGCCGTACTCATAG
CACGTTGTGTGGCTCTGATTTGTTGATAAATGATAGGAAAAAATAAAAGTTGATTGATAAAAGACAAACGAACGAGGA
ATATGCTCGTTTAATGTAATGTTTGGACATTCTGTCCTCTACTGGTGCATTAATAAACAAAAAGATGAAGAAAGTGCT
AATTAATTATTACAAATTAGTAGAAGTTAGTTGATTTGAAATTTAAAATTAAATACAAAATAGAAGCCAAATCATTTG
CATTTAATTCATATTGACATGTCTTCATAGGATAGAAGCAATGTTGTGCTTTTGTATATATTTACTAGTCATATATCC
ACTATGCACAGATTATATAAATTTATATTTTCTATTGTCTTAAATATAAAATTTTTCAGAAAATTCTTATAAAATTAT
TCGCTTATAAATAACAACCCTAAAAATATAGAACACATAAATTATGAATATGTTATAGTAATAAAATATCTATAGTAT
TAATATCTTAATCAATTTCAACAAAAAATATGATATTGAGAACCATAATTAGTGAATAAATTAATATCCAAACATGAG
CCAAAAATATAAAGGAAGCGAGCAATGTCAAAAAGCCTAGTATCAATGGATATTGCCAACACTCATATAATAATTAAA
TACTAAAAAAGATCAAAAGATTCCATAAAATGATAAAAAAAAAAAAAGATAAAACAAGAGAGAGAAAATTAGATACTGC
TTTTTTTTTTTTTTTTTTTTGATAATAAGAGATACGTATTATTAGTTTATAGTATATCCATATTATTAAATAGGGTATT
CAAAACTTCAATTGTAAAGTTTAATTAATACCTTCAGAATTTTCAAACTATAAAATAAAAATTTAATTAGCATAATAA
ATTATCATAGTTAATCGGTATTGAAATAACCTACAATCACATATTTTTCCCTTGCCAAAATCAGCACTTCAAAACTCT
ATATTTATGCCATGTCAGCTTTAAGGGTTTTGTTTTTATATATTTATATAGATATAGATAAATAAAAATGATTTAATA
ATAATTTCACACCCTAATTAAGCTATTACTTTAACCATATGACACCTTTAAATTTTATTTTAATCCACAAAACACTTC
AATTTATAAAATAGTAAACACTTACAAAATTATTTTATTAAAACACTTCACTTTGTAAAATTAAATCAATTAAACATT

TATTAAATAGTTAAAAAATAAAAATTAGTGCATTAAAAGTTATGCTAACTATACATCAAACTAATATTTTGCATATAA
AAGTTAAAAATTCTATACTATCAAAATTTATAAATTTAAGTGCACAATATATTATTTACAACGTTAAAGTTCAGAATA
CCGTAAAATTAAAATTAT

>JcRIP7|JCDBG12049|NW_012124368.1:36636..39480
TACAGAGAGAGAGTGTGTGTATTTGTAGTGAGAAGAGAAACTACGACTTTAAATCAAGTTTTCTTACATTAAAGAGAG
AAATAATTTCTACGAGGAAGGGAAACAACAACTTTGAATGAAGGCTTTTCTTTTAGGTTGCAGATAGAGCGGCAATAG
GACATTTTCTAGATATAAGAGGACTAGGGGTGTTAATTATTAACCGCCAAATTAAATCGAACCAAACTAGAGCAGTTT
AATTCAATTTGGTTACTATGAGAAAAAAGTTTGGTTTATATATTTTGAAAAAAAGAGCGATTTACAGTTCAGTTTTAA
TTTTTTTATTTAAAGAATTTAAATAAAACTGTAAGTCGTATAATAATATATAACATAGCAAGCAAACAGATTACGAAT
TATAATCTACGAAATAATAAAACTGTTCTCTTCAACTTAATTTACCTACTATTCTCCGGGCATTATTCACTTAATTAC
TTATATTGAGCATAAACCTTAATTAGTTATATATAAGTTTGCAACTTCAACAGAATTAATATGTCTTTTATTTAGTTG
AGTTGAATTTGATAATTTAAAAATACTTTAGTTTCAGTTTGAAAATTCTACAAATCGTAATTATCAGTTTGGTTTGAA
AATCCATTTAAACTGTCTAAACTAGACCGTGTATATCCCTAAAGAGGAATAAGGAGAAAGTTTACTATAATATTAAAT
CGAGCCTTAATTATCCATAGAAAAGTCAATTAACAAAATGTTTCCCTAGTCAAAAGGAACATGATTTCCCTCATTCAA
AGAGGAAAACATTTTCTAGAAATAAGTCTAATGGTAAGGACTTAAAACGTGAAGTGTTGTATGAACGTGCATGGTTGA
GATAGTTGGACTTGGAACATAAAATGTGAATTGTAGCTTTTGTATGTGAATTATTGAAAACAATAGAGAATTAAAATG
TGAATTGTGATTTTTACGTGAATTCTTAACCTAATACCATATTTTTTTTAAGTTTTAGTTTTGCTAAGGTGTATTATG
TATAACATAAAATAAGTCTAATTTCTCACTTAAAATAGAAAAAAAAAATTAACTCATTTAGTTTATATAATAACTTGAA
AACATAAAAATGTTTGATTTACATAATCATTTATAGTAAATAGATTATTTATTTTCCATTTTGATACACATTTAAATA
TTGTAATTTTATAATTTCAGATTTTAATATGTAATAATTTATTACATTTTTTAAGAAATATTTAGATTTCATAAAAAA
AATAAAGCAAAATACCTATAGTACGATTCCCTCTTATTCTTTTTCTCAAATCAAACGAGTACTATCTGTATTATAATC
TGATCCAAAGGCATCAAGTTTCTTTTTTCCAGTAATTCCTCCACCGCCATGAATAAAACGACAGAAAGAACAAAAAAA
ACAAACAAATAAGAGAGAGAAAAAAAAATGGGACCAGTTCCCTCTCCTATAAAAGCATATGATTTACTATGAAGAAAC
CATCTCATCCTCACTTCT
>JcRIP8|JCDBG12048|NW_012124368.1:19157..21532
TTTTCTACCTGCCAAAAATGATTAAACCATATTTAAAATGAATTAAACCAATAAATGTAATAAAAAAAATCAAGAAAG
AAAAGGCCTCGTGTGATTCATCAAGATGGATTGCATTTCTAGCAGGGAAAACTTGGGGAATATAAATAGTCATAAAAT
TAAAACTAATCTATTTTGAGGAGGGCAGAATATAAGAATGAAACAATATTTAATAAAATTGAATTAGAAATTTCAAAG
GAAAAAAACTTGCCAGGGGCTAGACCCTCCCACATAAAGCTTGCTTCTGTGTTCTAGACTAAGTTCATTTGTGCAGTC
AAGGGTTTAAACCTAAAATCCATATTCAAATGATTTTACTCGAATACAATTGTCTAATATCACCTCTCACATGTGCCA
AAAAAAAAAAAAAAAAAAGAAAAGAGAAAACCCCCGACAACAACGTATACACCAGAAAGAAATCCTTTATCATTCTTT
ACTTTTCTTTCTTTTAAGAAATTAACATATCTTCAGTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAATTTCAGAACTTGAAGTCTAGAACCAGCCAACTCAGTTGGACAGTTTCAAATTCTGAATCAATAAGC
TACAAATAATTATAAATTTAATTTATCAAAAGAATTTGAGAAAAGCTTCATTACTTACAAAGCAAGTGGGGAAGCAAA
TAGGAAGAGATGGTTTTAATTCTAGTGTTGGATTCATCAGCTTTTATATTGGGGAACGGGTCCCCCGGTTCTCATTTC
TCTTATTCTTTTTTTTTTCCTTAATTGCTTTGTTTCTTTCTTTTTTAATCCTAGAAGATATTTGTTCACGGTGATTAA
ATTACAAATTAAAACAAGAATGTTGGTTGTTGGAATTCCGGGGCAGATGATCATTTTTGTTTTCAAATTTCTTTTTCT
TTTTAAATTTATTAAGCTTGAGATTAGTATTTTTTTATTTATTTAACAATAAAGTGATGTATTGTGTCTCTTTAGTCA
TATTAATTATTAGTTTATTGATGCAATGGTTTAAATTAGAAATGTCCGATAAAAATAACTGTTTAATATAAAATAAAA
AAAAAAGAATTAGTTATTTAATTAGATAAACGCTTAACGTTTCATAAAAAAAAACTACCTCAATTTCTAAAAACTATTA
TAAAAACTATATTATAGAACATGCAATTATATATAGAGTTTTAACTTTATCAAAATATCAAAAGCTAGCTCAAGTCTT
CTTGCAAATTGGGGCATCATCTTAGGTTCTGTTTATTTCTAGAAAAGAACTTACCTTTGAAAAATATATTTTGAGGAA
TTAAATTTTAACTAAAAATCTTTTGCGAACGGGGCCATAGAGACTAATTTTTTCTTGTTGATCAAAGATAAATCATTG
GGTACAAAGCGTTTTCTTCTCAAGTGAGGCGGTTGTTGCTGCAGATAGACAAGGCATTTAATGGCCTCGTTTAGTCCC
TTCTGATGCTTGATGTTT

>JcRIP9|JCDBG12053|NW_012124368.1:96996..99440
GCAAAATTACATGAAATAAACACAAAAAAAAAAAAAACATGAAAAACTGTTCAATCAATTGAATCAAGCTAATTAATTT
ACATATATAAAAGCAAAAGGGATTATATGGATATCGAAAATGTAACATCTCTTTCCATTTTCCTTTACTTTTTTGAAC
AATTATGCACCACACACACTCTTGTTATCTGCAAGATTTTTATCTTTTCCCAAATAAGTTTGGATGAAGTATAATTGT
TGTATCTTCATAATTGCAATTTGAGATTTGGACTTTTCTATTTCACCTTCGTTTTCCCATTAATATGCACCCTTTCAA
CAGCTGCCTTTGTCTTTTGTTTATCTTAATTTTTGGAAAAATGATATACTATTTTTTCTATATATATATGTCACTTTA
AGATTTTTATACCATTATTAATAAATTAATATTTCATCTAAAATGATTTATATAAACATATAATTTATAAGGTCTAAA
TATTTATGGGGCAACATGTTAGAATATGAAAGTAATGACATTTTATTTTTTCATTATAATTATTATCTTTTATATTTA
AATAAAAATTAATTGTTCTTTATTAAATAATAGAGTTACAATTAAATTAATAAGATTTCATTATACTTAATATATAAT
AAATATAAAATTTTTTATTAAATAGTTTTTATTCTATCAAGAAAAACTAATAGAGTCAAAAATTAATTAAATAATGTA
ATTATATTAACAAATAGAATTTAAAAGATTAAATATTCTTATACTTCATTTAATTTGAATATATGAGAATTTAAAAGA
GAAAAAATAAATAATTACTAACTAAATCTGCCAATTGAATAGCAATAATTTGCTTACAATAATTAATTGTGTTAATTA
ATATTTCAGTTCAATATTTCAGAATTTAAAAAAGAATCTATAATTTTACCTTTCATCATTTAATAAAAAATAGATAAG
CAAACTATTTAAATAAAGCAACACAAATATAAATTAAAATAATTTTCTTGAAAATTACATTCAATTTAAATAAAAAAA
TAGAAAATTATACAATCCGCACATCGCATAGGTACAGAACTAGTTTATATGTATATATGAACAGAAACAATATTAAAT
TTAATCAAATCTATGCAATAAATTAAAGTAAAATAAAGATAACTAATGTTCCATCCCTCTCTTATCCCATTTGGCAAA
CAAAACCAGTCTTATTTGTGTTCTTACATATTACAACCAGGCATCAAGATTCTTTTCAACTAATTATTTAGCCGCCAG
GAATAAATTGACTTAAAAAAAAAAAAAAAAAAAATCTCTAAAATGGAGAGAGGAGCAAATGGACTATCTTCTATAAAAG
CAGATGATATGCAACACTTAAAAGAAACCATCTATTCCTTACTTTGTAAGTAATTATGAAGCTTCTGCATTTTCTACT

GTTTTTGTTGCCTCATTACATGCTTCATTTTAAAGTAATTAAGTTACCTATCTTTTGTTGGTATAAACATCTCTAATG
ATTTTTGACAGGTGAAAT

>JcRIP10|JCDBG12050|NW_012124368.1:52227..54494
TTCCATTTAAGTGGTCACAACACAACAGTTTATACTATAGTAAAAGGCATTTCATTTTCTGAAACTGATCCAGCAAAC
ATAGAATTTATAAAAATATAAACATTTTTTGAAAGGTTGTACAGAGAGAGAGTGTGTGTATTTGTAGTGAGAAGAGAA
ACTACGACTTTAAATCAAGTTTTCTTACATTAAAGAGAGAAATAATTTCTGCGAGGAGGGGAAACTTAATTAGATATA
TATAAGTTTGAAACTTCAACAGAATTAATATATCTTTTATTTATTTGTAACTTAAAATAATTTAGTTTCGGTTTGAAA
ATCCATTTAAACTGCCTAAACTAGACCGTGTATATGCCTAGAGACGACTAAAGAGAAAGTTTACTATAATATTAAATC
GAGCCTTAATTATATTTAGAAAATTCAATTAACAAAATGTTTTCCTAGTCCAAAAGGAAAATGATTTCCCTTATTCAA
GGACGAAAACATTTTCTAGAAATCTAGATATTCTTATAGTCTTATAATTTCAACTAATTTCCAAATACTGTATGTGAA
TTGTTCATGGTGGTAAGGACTTAAAACGTGAAGTGTTGCATGAACGTGCATGGTTGAAGATAGTTGGACTTGAAACAT
AAAACAGTGAATTGTAGCTTTTGTATGTGAATTATTGAAAACAATTGGGAATTAAAATGTGAATTGTGATTTTTTACG
TAAATTCTTAACCTAATACCATTTTTTTTTTTTAGTTTTAGTTTTGCTAAGGTGTATTATGTATTAACATAAAATAAG
TCTACTTTCTCAATTTAAATAGAAAAAAAAATTAACTCATTTAGTTTATATAATAACTTGAAAACATAAAATTGTTTG
GTTTATATAATCATTTATAGTAAATAGATTATTTATTTTTCATTTTGATACACATTTAAATATTATAATTTTATAATT
TCACATTTTAATATGTAATAATTTATTACATTGTTTAAGAAATATTTAGATTTCATAAAAAAAAAAAAAAAAAGCAAAA
TACCTTTACTACGATTCCCTCTTATTCTTTTTCTCAAATCAAACGAGTACTATATACATTATAATCTGATCCAAAGGC
ATCGAGTTTCTTTTTTCCAGTAATTCCTCCGCCGCCATGAATAAAACGACGGAAAGAACAAAAAAACAACAAATAAGA
GAGAGAAAAAAAAAAATGAGACCAGTTCCCTCTCCTATAAAAGCATATGATTTACTATGAAGAAACCATCTCATCCTCA
CTTCTTCACTTGCTTGTAAGTCAATACAAAAAGCTTTTCTCAAAAAATTTTTTTAAACATCTCTACTGCATTTTTCGC
TATTTCTGTTCATTAAATTATTGTACTCTTGAAAACCTTTTTAGTTCTAGCAACTTGAATTTAATAATATTTGTTCTA
ATTTAATTGCTCAAGCTCTTTTAACTGTTGGATTTTCATTTCAGTTAGCAATCTTTTGTTGATATAAACATGTGCATT
GATTTTTGGCAGGGGAAA
>JcRIP11|JCDBG12054|NW_012124368.1:107222..109639
TTTTAGTTAATTATTAGTAGAGAATATTAGTTTAGATTTTTGTTTTGCCATTTTTTAAGCATGAAATTTTAGTTAGGG
TTTGTTGAAATTATACTTAAGTTAGATATCTTATAAAAATAGGTTTTATTCGTTACGAGGTCATTTTATTTTTAAGAA
ATAATTACTTGTTGTCAATATTATATAAATAAATAAAACGTATTTTTTTTTAAGATGAATATATATATTTTTTATTTTA
ATATTTACATAAGTAAATATAAATATTTATTTATTTATAAGTGGTGAGTGTGCGTACCACAAAGATAACATTTTATTC
ATTAAATTGGTACAATTTGAAATAGGATGCTATCGTAATTAAAGTTGAAATTTAAAATTTAGAGTAATAAGGATTAAA
AAAATTAGATTTAGAGTACAATTTTTTTAATGATTTAGTCTAGAAAAAAATAACACAACCCAAAATTTATCTGTAGAG
TTATTCTTGTTCCAATGATTAGATGACAAAGCAAATGATCCCTAACCTAAATAACAAGAATGATGGCTTGGTTCCCAG
GTTTCATTGTAATATATATTACTGTGGATAGTACAGATTGGTGGGCATATCAGTGAAGGTACTCTTATTTATCACACT
GCTGCTGGGTAATTGAATTTTTCAATAAACACCGTAATTAAATTAGGGATTAGAAAAATAATCCCTAATTTACTCAGA
TGAGCAAGTAGAGGCATGGAAAAAGGTTGTCGGTGCAGTTCACTAAGGAAGCTTCATTTTCTGTCAATTATGATACAC
TGGTCGTGCGTGCGGGACACAGATGAGTCAATTGATATTCATACGAGAGAAGGGAAAAATATCTGTCGTTTTAGAATG
AAGAGGAAAAATTTCCCTGTCCTCCCAGATGACAAGCAAAGGCGGACAGTTTCCTTTGATCGTCTCAGAAAAAAGGCT
GTAATGGTATGTCGAGTCTTGTAAACCAAAGGCTGTAATGGCATGTCGAGTCTTGTAAACCATGGAATGGACGCTTCA
ATATTAAATCACTAAATCTGAAATTGAAATATAATATGTATATAGGCATAACAAAGGGTAAAAAATAATCAGTATTTA
CTTTAACAAATTCATTATAACAAAAATAATAAAAAGACAATAATATATATTCTCAATTACAAAAAATAACTTTATGGG
ATTCACTTGGTGTATGGGCATCAAGATTACTTTCTAGTAATCATTCAGCCTCCATTTCATAAAATGACTAAAAAAATG
AGAGATGAGAGAATGGCACCAGTTCCAATCTCCTATAAAAGCAGATGATATGAAACACTAGGAGAAACCATCTCTTCC
TCGCTTCTTTTCTGTTTTGTAAGTATTTATGAAGCTTTTCTCAAAATATTTCCGAACATCTCTACTGCATTTTCTACT
GTTTTTGTTGTCTTATTACATGCTTCATCTGTAAGTAACTATCTTATCTATCTTTTGTTGATATAAACATTTGTAATA
ATTTTTGACAGGTGAAAT
>JcRIP12|JCDBG12052|NW_012124368.1:82429..85584
TGTTACCTTTAGTTCTCAGAGTGTAGAACATTTACAATTTTTATCCACCTGTCAACAAAAGAAAAGAAGATGCTCAAA
GTTAATCTAGCTTATTCATCAATAATTAACAAAATAATCAAACTAATTTCATGAAGAGCAAATTAGGAAAATAAAATT
TATTTTTTTATATATAACGCACACAAGCACACAAGTACAAATTATAAATCAAATTCAGGTTAAACTGAATTGAACCGG
TAAAACCTAAATTGGTTTGAGTGGTTGATGATTCACTTAGTTGAATTTGACCATTTTCAAATAATTGAGAAATTCGAA
TTGAATCGGTTAGTCTAATCTTAGTTTAAAAACCATAGATAAGAGGAAAAATACTTCCAAAATAAAATAAATGTTATA
GGAATGTTCTAATCTGAAAACCATTGAATTAGAATTTGAAGAAAATAGAAACATAATAATATATCTAAACTCAACTAC
TACAAATTCTTAAATGAGGAGGTAAATTACTTCTAAAATAAGATAATAAAAAAGTGCTACTATAAATACGTTCACATA
TGAGAAAGAGTTATTTTATTATCTAATTAGAATTTAAATAAAATAGAGACATACTTTATATAAAACTCAAGTACATTC
AAATTCTATAAAAGAGTATAAGTTGATCGATTTTTGACTCTAGTCGGTAATAGCTTTTCAAGTAATTTTTAGTATTTT
GACTAGTCAAAATGCTAAATGCTATTTTTTATAGTATTTAATAATATAGCATTTGAAATACTCAGGGAATTAAAGCTT
ATTACTTTACGTTGTTCATACTAACTAGTAGTTGAGCACCATAGTTATCCTAATAACCATTAGTTAGTTGATGCACAT
TAAAATATAAAATCAATATTAATAATCTACATAATTATTTTCAGATGCTATGTATCATAATAAATAATTATTAGATAT
TATTATACCAAAATGTAATATTTTTTATTTATATAACATATCAAAATGAAAAATAAAATATACACACACCACATATAT
AAACTATTAATGTTTAAAAATATATAATTTTGCTAGTAATTATTTTTATAAATTATATATATATATATATATATATAT
CCGCTAATAATTGATAAAAAAAATTTGTTAAATACAGTTGATAGAATGAACAAATTATTATGATTTCTACCGGTTATT
AACTCCTAATTAAAACATTAACCAAAAAATTACAGAAATCAGTTGAACTAAATACAGCTCTTTATTTCATGCTAATTG
CAATAGAGAATAGCTAGTAATAGTAAATGTATTTAAAAAAAAAAAAAAAAATTAAGTATTGTGGTTCAATGAACGAAAA
TTGGGAACCCTATCGAGCTTATAGGGGTCAACAAGGACTCTACATATATAAGAATTGAGAAAAGCTTCATTACTTACA
AAACAAGTGAAGAAGCAAAATAGATAGAGGTGGCTTAATTCTAGGATTGCATATTCATCACCTTTTATATTGGGTAAA
TGGGTTCCTGTTTCCTTT

Apêndice 3: S3 Figure. List of the cis-elements found in the promotor regions of the physic nut RIP genes. The number of elements by function was used to obtain the Figure 2 (excepted the categories indicated by *)

Gene	JcRIP1				
	Name	**function**	**Position**	**Number of elements**	**Total**
JcRIP1	3-AF1 binding site	light responsive element	895 -	1	9
	ATC-motif	part of a conserved DNA module involved in light responsiveness	311 -	1	
	Box 4	part of a conserved DNA module involved in light responsiveness	608 +, 750 -, 633 +	3	
	G-Box	cis-acting regulatory element involved in light responsiveness	245 +	1	
	G-box	cis-acting regulatory element involved in light responsiveness	245 -, 1044 -	2	
	MRE	MYB binding site involved in light responsiveness	702 +	1	
	ARE	cis-acting regulatory element essential for the anaerobic induction	763 +, 882 -, 783 +, 1227 +	4	4
	GC-motif	enhancer-like element involved in anoxic specific inducibility	1393 -	1	1
	LTR	cis-acting element involved in low-temperature responsiveness	1127 -	1	1
	Circadian	cis-acting regulatory element involved in circadian control	538 +	1	1
	TGA-element	auxin-responsive element	75 +, 1370 +	2	2
	CAT-box	cis-acting regulatory element related to meristem expression	1401 -	1	1
	CGTCA-motif	cis-acting regulatory element involved in the MeJA-responsiveness	1085 -	1	2
	TGACG-motif	cis-acting regulatory element involved in the MeJA-responsiveness	1085 +	1	
	TATC-box	cis-acting element involved in gibberellin-responsiveness	118 +	1	1
	ERE	ethylene-responsive element	1333 +	1	1
	ABRE	cis-acting element involved in the abscisic acid responsiveness	1281 -, 245 -, 1044 -	3	5
	ABRE3a	cis-acting element involved in the abscisic acid responsiveness	1044 -	1	
	ABRE4	cis-acting element involved in the abscisic acid responsiveness	1044 +	1	
	MYB	transcription factor MYB	69 +, 1266 -, 986 +	3	5
	Myb	transcription factor MYB	570 -, 999 +	2	
	MYC	transcription factor MYC	1295 -, 1491 -	2	2
	WRE3	wounding and pathogen response	1026 +	1	2
	WUN-motif	wounding and pathogen response	255 -	1	
	as-1	drought response	1085 +	1	1

	CAAT-box*	common cis-acting element in promoter and enhancer regions	31 +, 122 +, 123 +, 138 -, 217 -, 378 -, 502 +, 538 +, 548 -, 559 +, 561 -, 606 +, 652 +, 729 +, 747 +,, 771 +, 778 -, 822 +, 840 +, 1134 +, 1210 -, 1261 -, 1295 +, 1309 -, 1483 +, 1484 +, 1491 +, 1493 -	29	47
	TATA-box*	core promoter element around -30 of transcription start	403 -, 626 +, 622 +, 668 -, 624 +, 1138 -, 484 +, 627 +, 623 +, 669 +, 620 -, 667 +, 625 +, 475 +, 541 +	15	
	AT~TATA-box*		622 +, 624 +, 626 +	3	
	Unnamed__4*		365 -, 1095 +, 935 +, 1422 +, 905 +, 1216 +, 941 +, 1166 +, 686 -, 1300 -, 938 +, 917 +, 1037 +	13	13
Gene	**JcRIP2**				
JcRIP2	**Name**	**function**	**Position**	**Number of elements**	**Total**
	Box 4	part of a conserved DNA module involved in light responsiveness	30 +, 996 -, 128 +, 1463 -, 214 +, 1124 -, 837 -	7	16
	G-Box	cis-acting regulatory element involved in light responsiveness	1201 -	1	
	G-box	cis-acting regulatory element involved in light responsiveness	1201 -, 1162 -	2	
	GT1-motif	light responsive element	256 -	1	
	MRE	MYB binding site involved in light responsiveness	755 -	1	
	TCCC-motif	part of a light responsive element	250 +	1	
	chs-CMA1a	part of a light responsive element	737 +	1	
	AAGAA-motif	part of a module for light response	138 +, 1311 -	2	
	TCA-element	cis-acting element involved in salicylic acid responsiveness	1153 -	1	1
	TC-rich repeats	cis-acting element involved in defense and stress responsiveness	1109 +	1	1
	LTR	cis-acting element involved in low-temperature responsiveness	748 -	1	1
	MBS	MYB binding site involved in drought-inducibility	197 +	1	1
	ABRE	cis-acting element involved in the abscisic acid responsiveness	1163 +, 1202 +, 1201 -	3	3
	Myc	transcription factor MYC	133 -, 1044 +, 670 +	3	5
	MYC	transcription factor MYC	53 +, 1132 -	2	
	Myb	transcription factor MYB	197 +, 1297 -	2	4
	MYB	transcription factor MYB	450 -, 1489 -	2	
	W box	wounding and pathogen response	905+	1	2
	WUN-motif	wounding and pathogen response	1359 -	1	
	TCA	cis-acting element involved in salicylic acid responsiveness	218 -	1	2
	Unnamed__1	60K protein binding site	1200 -	1	1

	TATA~box*	core promoter element around -30 of transcription start	89 +, 1035 -, 1018 -, 1036 -, 116 +, 1037 -, 1016 -, 1038 -, 117 -, 1039 -, 118 +, 1040 -, 264 -, 1060 -, 265 -, 1061 -, 266 -, 1062 -, 267 -, 1063 -, 268 +, 1064 -, 581 +, 1088 -, 582 +, 1093 -, 1094 -, 1095 -, 1096 -, 1104 +, 1105 -, 1106 -, 1258 +, 1259 -, 1280 -, 1282 -, 1333 -, 1334 -, 640 -, 1335 -, 1336 -, 827 -, 828 -, 829 +, 830 -, 1436 +, 1437 -, 858 +, 860 -, 1469 -, 976 -, 1470 -, 977 -, 1471 -, 978 -, 1472 -, 979 -, 1479 -	58	104
	AT~TATA-box*		828 -, 1334 -, 1036 -, 1038 -, 1094 -	5	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	4 -, 43 +, 54 -, 106 -, 321 -, 336 -, 348 -, 365 +, 372 -, 377 +, 448 -, 467 +, 485 -, 646 -, 694 -, 783 -, 854 -, 877 -, 900 -, 909 +, 910 +, 939 -, 1013 -, 1132 +, 1134 -, 1187 +, 1212 -, 1233 +, 1267 -, 1292 +, 1304 -, 1320 +, 1321 +, 1343 +, 1369 +, 1372 -, 1421 +, 1487 -	41	
	Unnamed__2*		752 -, 757 -	2	2
	Unnamed__4*		251 +, 598 -	2	2
	Unnamed__6*		1036 +	1	1
	chs-Unit 1 m1*		750 -	1	1

Gene	**JcRIP3**				
	Name	**function**	**Position**	**Number of elements**	**Total**
JcRIP3	AT1-motif	part of a light responsive module	93 +	1	12
	Box 4	part of a conserved DNA module involved in light responsiveness	756 -	1	
	G-Box	cis-acting regulatory element involved in light responsiveness	360 +	1	
	G-box	cis-acting regulatory element involved in light responsiveness	360 -, 1022 -	2	
	I-box	part of a light responsive element	731 +	1	
	AAGAA-motif	part of a module for light response	1234 +, 421 +, 1004 +, 1252 +, 455 +, 1191 +	6	
	TGACG-motif	cis-acting regulatory element involved in the MeJA-responsiveness	277 -	1	2
	CGTCA-motif	cis-acting regulatory element involved in the MeJA-responsiveness	277 +	1	
	ABRE	cis-acting element involved in the abscisic acid responsiveness	360 -, 1022 -	2	1
	ARE	cis-acting regulatory element essential for the anaerobic induction	569 +	1	1
	LTR	cis-acting element involved in low-temperature responsiveness	1147 +	1	1
	MBS	MYB binding site involved in drought-inducibility	172 +	1	1

	O2-site	cis-acting regulatory element involved in zein metabolism regulation	592 -	1	1
	TC-rich repeats	cis-acting element involved in defense and stress responsiveness	1234 -	1	1
	TGA-element	auxin-responsive element	130 +	1	1
	ABRE3a	cis-acting element involved in the abscisic acid responsiveness	1022 -	1	4
	ABRE4	cis-acting element involved in the abscisic acid responsiveness	1022 +	1	
	MYB	Transcription factor MYB	431 +, 1370 -, 734 +	3	
	MYB-like sequence	transcription factor MYB	431 +, 1370 -, 734 +	3	6
	Myb	transcription factor MYB	172 +	1	7
	MYC	transcription factor MYC	1059 -	1	1
	WRE3	wounding and pathogen response	898 -	1	1
	as-1	drought response	277 -	1	1
	Unnamed__4*		248 -, 957 +, 856 +, 1104 -	4	
	TATA-box*	core promoter element around -30 of transcription start	42 -, 636 -, 665 -, 705 +, 46 +, 703 +, 388 -, 985 -, 44 -, 698 +, 366 +, 860 -, 364 -, 704 -, 525 +, 986 -, 43 -, 45 -	18	49
	CAAT-box*	common cis-acting element in promoter and enhancer regions	72 +, 166 +, 207 +, 213 +, 219 +, 220 +, 235 -, 263 -, 301 -, 332 -, 393 -, 492 -, 593 +, 633 -, 737 +, 738 +, 747 -, 776 +, 777 +, 778 +, 881 +, 918 -, 965 +, 969 +, 1017 +, 1059 +, 1078 +, 1087 -, 1166 -, 1403 + 1420 -	31	

Gene	**JcRIP4**				
	Name	**function**	**Position**	**Number of elements**	**Total**
JcRIP4	Box 4	part of a conserved DNA module involved in light responsiveness	27 +, 889-, 539 +, 1199 -, 258 +, 950 -, 570+	7	17
	G-Box	cis-acting regulatory element involved in light responsiveness	412 +, 726 +	2	
	G-box	cis-acting regulatory element involved in light responsiveness	726 +	1	
	GA-motif	part of a light responsive element	819 +	1	
	GT1-motif	light responsive element	1429 -, 1453 -	2	
	TCCC-motif	part of a light responsive element	294 +	1	
	chs-CMA1a	part of a light responsive element	1449 +	1	
	AAGAA-motif	part of a module for light response	181 +, 135 +	2	
	ABRE	cis-acting element involved in the abscisic acid responsiveness	412 -, 727 +, 726 +	3	3
	ARE	cis-acting regulatory element essential for the anaerobic induction	1356 -	1	1
	ERE	ethylene-responsive element	603 -, 610 -	2	2
	MYB	transcription factor MYB	1454 +	1	2

	MYB-like sequence	transcription factor MYB	1454 +	1	
	MYC	transcription factor MYC	17 -, 641 +, 50 +, 1468 +	4	5
	Myc	transcription factor MYC	1135 -	1	
	TCA	cis-acting element involved in salicylic acid responsiveness	262 -	1	1
	STRE	Stress Response Element	1434 -	1	1
	Unnamed__1	60K protein binding site	725 -	1	1
	TATA-box*	core promoter element around -30 of transcription start	77 -, 739 +, 78 +, 744 -, 76 +, 738 +, 79 +, 745 -, 80 -, 746 -, 86 +, 747 +, 113 +, 817 -, 114 -, 818 -, 115 +, 975 +, 300 -, 976 -, 301 +, 1029 +, 310 -, 1030 -, 311 -, 312 -, 313 -, 314 +, 621 +, 1127 -, 1128 -, 1129 -, 1130 -, 1155 -, 1156 -, 1157 -, 1162 -, 1224 -, 1331 -, 622 +, 696 -, 1365 -, 699-, 1366 -, 701 +, 1367 -, 735 +, 1368 -, 736 -, 1369 -, 737 +, 1370 -, 1372 -, 1373 -, 1374 -, 1375 -, 1376 -, 1378 -, 1384 -, 1389 +, 1390 -	61	93
	TATA*		622 +	1	
	AT~TATA-box*		737 +, 1374 -, 1366 -, 1376 -, 1368 -, 78 +	6	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	1 -, 17 +, 51 -, 103 -, 344 +, 380 -, 429 -, 465 -, 563 -, 581 +, 636 +, 641 +, 643 -, 659 -, 685 +, 689 -, 757 -, 827 +, 885 +, 903 +, 1012 +, 1020 -, 1063 +, 1273 -, 1425 +	25	
	Unnamed__4*		295 +, 720 +	2	2
	Unnamed__6*		888 +, 1369 -	2	2

Gene	**JcRIP5**				
JcRIP5	**Name**	**function**	**Position**	**Number of elements**	**Total**
	AAGAA-motif	part of a module for light response	74 +, 445 +, 473 -, 814 -	4	11
	AE-box	part of a module for light response	810 -	1	
	LAMP-element	part of a light responsive element	455 +, 1215 +	2	
	Box 4	part of a conserved DNA module involved in light responsiveness	1016 -	1	
	MRE	MYB binding site involved in light responsiveness	322 +, 1270 -	2	
	chs-CMA2a	part of a light responsive element	1424 -	1	
	ARE	cis-acting regulatory element essential for the anaerobic induction	45 +, 713 -, 24 +, 1041 -	4	4
	GC-motif	enhancer-like element involved in anoxic specific inducibility	421 +, 764 +	2	2
	MBS	MYB binding site involved in drought-inducibility	591 -	1	4
	MYB	transcription factor MYB	882 -	1	

	Myb	transcription factor MYB	591 -, 1070 +	2	
	MYC	transcription factor MYC	300 +, 382 +, 380 -, 339 -, 359 +	5	6
	Myc	transcription factor MYC	779 +	1	
	STRE	Stress Response Element	249 +	1	1
	ERE	ethylene-responsive element	32 -	1	1
	WUN-motif	wounding and pathogen response	55 -, 923 +	2	2
	CAAT-box*	common cis-acting element in promoter and enhancer regions	49 +, 214 -, 200 +, 301 -, 48 +, 339 +, 108 -, 359 +, 361 -, 605 +, 617 +, 627 +, 658 -, 699 +, 750 -, 804 -, 840 -, 863 +, 922 +, 983 +, 997 -, 1031 -, 1038 +, 1152 +, 1194 +, 1253 +, 1254 -, 1256 -, 1401 -	29	64
	TATA-box*	core promoter element around -30 of transcription start	30 +, 1196 +, 187 +, 1200 -, 186 +, 1198 -, 1183 -, 1312 -, 138 +, 1197 -, 436 +, 1202 -, 139 +, 1199 -, 1182 -, 1313 -, 505 +, 633 +, 634 -, 635 +, 744 -, 745 -, 746 -, 747 +, 1060 +, 1080 +, 1081 -, 1167 +, 1168 -, 1169 -, 1178 -, 1181 +	32	
	TATA*		1081 +	1	
	AT~TATA-box*		1198 -, 1200 -	2	
	Unnamed__2*		765 +, 896 -	2	2
	Unnamed__4*		176 -, 261 +	2	2

Gene	**JcRIP6**				
JcRIP6	**Name**	**function**	**Position**	**Number of elements**	**Total**
	Box 4	part of a conserved DNA module involved in light responsiveness	283 +, 960 -, 623 +, 314 +, 684 +	5	10
	G-Box	cis-acting regulatory element involved in light responsiveness	157 +	1	
	GA-motif	part of a light responsive element	1146 +	1	
	GATA-motif	part of a light responsive element	188 +	1	
	GT1-motif	light responsive element	1201 -	1	
	TCCC-motif	part of a light responsive element	27 +	1	
	O2-site	cis-acting regulatory element involved in zein metabolism regulation	1102 -	1	1
	TCA-element	cis-acting element involved in salicylic acid responsiveness	293 -	1	1
	MYB	transcription factor MYB	1202 +	1	2
	MYB-like sequence	transcription factor MYB	1202 +	1	
	MYC	transcription factor MYC	945 -, 385 +	2	2
	ERE	ethylene-responsive element	354 -	1	1
	ABRE	cis-acting element involved in the abscisic acid responsiveness	157 -	1	1

	CAAT-box*	common cis-acting element in promoter and enhancer regions	76 +, 125 -, 174 -, 209 -, 325 +, 346 -, 380 +, 386 -, 403 -, 429 +, 502 -, 637 +, 659 -, 725 +, 746 +, 754 -, 756 +, 945 +, 947 -, 1034 -, 1048 +, 1248 +, 1315 +, 1449 +	24	
	TATA*		513 +, 985 +, 536 +, 1253 +	4	
	TATA-box*	core promoter element around -30 of transcription start	33 -, 34 +, 43 -, 44 -, 45 -, 46 -, 47 +, 366 +, 440 -, 443 -, 445 +, 446 +, 447 +, 461 +, 462 +, 480 +, 481 -, 482 +, 483 +, 484 +, 486 +, 489 -, 490 -, 491 -, 492 +, 512 +, 513 +, 532 -, 535 -, 536 +, 551 -, 552 +, 573 +, 599 -, 600 +, 617 +, 709 +, 710 +, 767 +, 768 -, 903 -, 904 - , 905 -, 910 -, 985 -, 1092 -, 1126 - , 1127 - 1128 -, 1129 -, 1130 -, 1131 -, 1133 -, 1134 -,1135 -, 1136 -, 1137 -, 1139 -, 1145 -, 1250 -, 1251 -, 1252 -, 1253 -, 1271 +, 1376 -, 1399 +, 1400 -, 1418 -, 1430 -, 1431 -, 1432 -, 1433 -, 1451 -, 1452 -	74	108
	AT~TATA-box*		445 +, 1137 -, 1127 -, 1135 -, 482 +, 1129 -	6	
	Unnamed__4*		28 +	1	1
	Unnamed__6*		1130 -, 622 +	2	2
Gene	**JcRIP7**				
JcRIP7	**Name**	**function**	**Position**	**Number of elements**	**Total**
	ACE	cis-acting element involved in light responsiveness	568 -	1	14
	AE-box	part of a module for light response	1342 -	1	
	Box 4	part of a conserved DNA module involved in light responsiveness	523 +	1	
	G-Box	cis-acting regulatory element involved in light responsiveness	826 -, 844 -, 951 +	3	
	GT1-motif	light responsive element	195 -, 962 -, 196 -	3	
	MRE	MYB binding site involved in light responsiveness	131 -, 964 +	2	
	TCT-motif	part of a light responsive element	62 +	1	
	ATCT-motif	part of a conserved DNA module involved in light responsiveness	1323 +	1	
	AAGAA-motif	part of a module for light response	1390 +	1	
	ABRE	cis-acting element involved in the abscisic acid responsiveness	827 +, 952 +, 845 +	3	5
	ABRE3a	cis-acting element involved in the abscisic acid responsiveness	951 +	1	
	ABRE4	cis-acting element involved in the abscisic acid responsiveness	951 -	1	
	GARE-motif	gibberellin-responsive element	516 -	1	2
	P-box	gibberellin-responsive element	753 -	1	
	MYB	transcription factor MYB	244 -, 852 -, 516 +	3	5
	MYB-like sequence	transcription factor MYB	244 -	1	

	Myb-binding site	transcription factor MYB	516 +	1	
	Myc	transcription factor MYC	1414 -	1	1
	ARE	cis-acting regulatory element essential for the anaerobic induction	266 -, 1479 +, 616 -	3	3
	O2-site	cis-acting regulatory element involved in zein metabolism regulation	1482 -	1	1
	ERE	ethylene-responsive element	1236 +	1	1
	STRE	Stress Response Element	182 +	1	1
	TC-rich repeats	cis-acting element involved in defense and stress responsiveness	58 +	1	1
	TCA-element	cis-acting element involved in salicylic acid responsiveness	973 +, 1426 -	2	2
	TGA-element	auxin-responsive element	1383 +	1	1
	Unnamed__1	60K protein binding site	327 +	1	1
	WUN-motif	wounding and pathogen response	1199 -, 1209 +	2	2
	TATA-box*	core promoter element around -30 of transcription start	169 +, 170 +, 267 -, 269 -, 270 -, 271 +, 272 +, 273 +, 319 +, 330 -, 351 +, 358 +, 359 +, 360 +, 361 +, 389 +, 390 -, 391 +, 469 -, 470 +, 495 -, 496 +, 497 +, 498 +, 499 +, 500 +, 654 -, 656 +, 690 +, 896 -, 1013 -, 1015 -, 1075 -, 1076 -, 1077 -, 1078 +, 1079 -, 1108 -, 1121 -, 1122 -, 1123 -, 1124 -, 1163 -, 1177 -, 1178 -, 1179 -, 1180 -, 1265 -, 1318 +, 1319 -, 1320 -, 1450 +, 1452 -	53	77
	AT~TATA-box*		269 -, 1075 -, 359 +, 1077 -, 271 +, 498 +, 496 +	7	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	21 -, 729 +, 241 -, 940 -, 204 +, 911 -, 553 -, 1295 +, 151 +, 888 -, 473 -, 1170 -, 239 +, 919 +, 597 +, 1410 +	16	
	TATA*		1176 -	1	
	Unnamed__10*		87 -	1	1
	Unnamed__12*		87 -	1	1
	Unnamed__14*		87 -	1	1
	Unnamed__4*		444 +, 1365 +, 677 -, 1448 +	4	4
	Unnamed__8*		87 -	1	1

Gene	**JcRIP8**				
JcRIP8	**Name**	**function**	**Position**	**Number of elements**	**Total**
	AE-box	part of a module for light response	810 -	1	11
	LAMP-element	part of a light responsive element	455 +, 1215 +	2	
	chs-CMA2a	part of a light responsive element	1424 -	1	

	Box 4	part of a conserved DNA module involved in light responsiveness	1016 -	1	
	MRE	MYB binding site involved in light responsiveness	322 +, 1270 -	2	
	AAGAA-motif	part of a module for light response	74 +, 445 +, 473 -, 814 -	4	
	MBS	MYB binding site involved in drought-inducibility	591 -	1	4
	MYB	transcription factor MYB	882 -	1	
	Myb	transcription factor MYB	591 -, 1070 +	2	
	MYC	transcription factor MYC	300 +, 382+, 380 -, 339 -, 359 +	5	6
	Myc	transcription factor MYC	779 +	1	
	ERE	ethylene-responsive element	32 -	1	1
	STRE	Stress Response Element	249 +	1	1
	GC-motif	enhancer-like element involved in anoxic specific inducibility	421 +, 764+	2	2
	ARE	cis-acting regulatory element essential for the anaerobic induction	45 +, 713 -, 24 +, 1041 -	4	4
	WUN-motif	wounding and pathogen response	55 -, 923 +	2	2
	AT~TATA-box*		1198 -, 1200 -	2	64
	TATA*		1081 +	1	
	TATA-box*	core promoter element around -30 of transcription start	30 +, 1196 +, 187 +, 1200 -, 186 +, 1198 -, 1183 -, 1312 -, 138 +, 1197 -, 436 +, 1202 -, 139 +, 1199 -, 1182 -, 1313 -, 505 +, 633 +, 634 -, 635 +,744 -, 745 -, 746 -, 747 +, 1060 +, 1080 +, 1081 -, 1167 +, 1168 -, 1169 -, 1178 -, 1181 +	32	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	49 +, 214 -, 200 +, 301 -, 48 +, 339+, 108 -, 359 + 361 -, 605 +, 617 +, 627 +, 658 -, 699 +, 750 -, 804 -, 840 -, 863 +, 922 +,983 +, 997 -, 1031 -, 1038 +, 1152 +, 1194 +, 1253 +, 1254 -, 1256 -, 1401 -	29	
	Unnamed__4*		176 -, 261 +	2	2
	Unnamed__2*		765 +, 896 -	2	2

Gene	**JcRIP9**				
JcRIP9	**Name**	**function**	**Position**	**Number of elements**	**Total**
	3-AF1 binding site	light responsive element	1146 -	1	14
	AE-box	part of a module for light response	1077 +	1	
	Box 4	part of a conserved DNA module involved in light responsiveness	71 +, 686 +, 554 +, 855 -, 407 +, 842 -, 592 +, 293 +, 415 +	9	
	AT1-motif	part of a light responsive module	547 -	1	
	GA-motif	part of a light responsive element	928 +	1	
	TCT-motif	part of a light responsive element	1192 +	1	

	P-box	gibberellin-responsive element	92 -	1	2
	TATC-box	cis-acting element involved in gibberellin-responsiveness	1155 +	1	
	MYB	transcription factor MYB	310 +, 1204 +	2	3
	Myb-binding site	transcription factor MYB	310 +	1	
	MYC	transcription factor MYC	812 -, 1160 +, 55 +, 1303 -	4	5
	Myc	transcription factor MYC	1150 +	1	
	O2-site	cis-acting regulatory element involved in zein metabolism regulation	1329 +	1	1
	ARE	cis-acting regulatory element essential for the anaerobic induction	1173 +, 1351 +	2	2
	ERE	ethylene-responsive element	1431 +	1	1
	TC-rich repeats	cis-acting element involved in defense and stress responsiveness	1188 +	1	1
	AT~TATA~box*		371 +, 437 -, 373 +, 617 +, 82 +, 1066 -, 375 +, 439 +	8	121
	CAAT-box*	common cis-acting element in promoter and enhancer regions	51 +, 55 +, 57 -, 56 +, 207 +, 230 -, 247 -, 251 +, 253 -, 260 -, 558 -, 285 +, 713 +, 755 -, 811 +, 812 +, 814 -, 823 +, 828 -, 837 +, 846 -, 870 +, 960 +, 998 +, 1027 +, 1082 +, 1098 +, 1108 +, 1161 -, 1184 -, 1255 -, 1303 +	32	
	TATA*		628 +	1	
	TATA-box*	core promoter element around -30 of transcription start	81 +, 539 +, 538 -, 541 +, 82 +, 543 -, 606 +, 607 -, 608 +, 537 -, 616 +, 617 +, 618 +, 619 +, 627 +, 628 +, 703 +, 704 -, 705 +, 738 -, 741 -, 742 +, 761 -, 762 -, 897 -, 942 +, 944 -, 963 +, 964 -, 1001 -, 1021 =, 1022 -, 1023 -, 1058 -, 1059 -, 1060 -, 1064 -, 1066 -, 1067 -, 1068 -, 1319 -, 1466 -, 83 +, 84 +, 100 +, 101 -, 102 +, 226 +, 357 +, 371 +, 372 +, 373 +, 374 +, 375 +, 376 +, 377 +, 392 -, 395 -, 396 -, 397 -, 398 +, 435 -, 436 -, 437 -, 438 -, 439 +, 440 +, 441 +, 448 +, 449 +, 453 -, 454 -, 455 -, 456 +, 521 +, 522 -, 523 +, 536 -	78	
	Unnamed_4*		1292 -, 1299 -	2	2
	Unnamed_6*		391 -, 1466 +	2	2

Gene	**JcRIP10**				
	Name	**function**	**Position**	**Number of elements**	**Total**
JcRIP10	ATCT-motif	part of a conserved DNA module involved in light responsiveness	1077 +	1	8
	G-Box	cis-acting regulatory element involved in light responsiveness	571 -, 589 -	2	
	Box 4	part of a conserved DNA module involved in light responsiveness	258 +	1	
	GT1-motif	light responsive element	710 -	1	
	MRE	MYB binding site involved in light responsiveness	712 +	1	

	TCT-motif	part of a light responsive element	179 +	1	
	AAGAA-motif	part of a module for light response	1144 +	1	
	ABRE	cis-acting element involved in the abscisic acid responsiveness	572 +, 590 +	2	2
	ARE	cis-acting regulatory element essential for the anaerobic induction	857 -, 1232 +	2	2
	GARE-motif	gibberellin-responsive element	251 -	1	1
	LTR	cis-acting element involved in low-temperature responsiveness	300 -	1	1
	O2-site	cis-acting regulatory element involved in zein metabolism regulation	1235 -	1	1
	TC-rich repeats	cis-acting element involved in defense and stress responsiveness	175 +	1	1
	TCA-element	cis-acting element involved in salicylic acid responsiveness	721 +	1	1
	TGA-element	auxin-responsive element	1137 +	1	1
	MYC	transcription factor MYC	666 +, 1473 +	2	3
	Myc	transcription factor MYC	1166 -	1	
	MYB	transcription factor MYB	23 +, 597 -, 251 +, 1429 -	4	9
	Myb	transcription factor MYB	1426 +, 1446 -	2	
	Myb-binding site	transcription factor MYB	23 +, 1429 -, 251 +	3	
	ERE	ethylene-responsive element	986 +	1	1
	STRE	Stress Response Element	213 +, 1493 +	2	2
	WUN-motif	wounding and pathogen response	949 -, 959 +	2	2
	Unnamed__4*		211 -, 1201 +, 1119 +	3	3
	Unnamed__6*		228 -	1	1
	AT~TATA-box*		231 +, 862 -, 825 -, 1065 -, 233 +, 860 -, 827 -	7	94
	CAAT-box*	common cis-acting element in promoter and enhancer regions	138 -, 277 -, 416 +, 530 +, 546 -, 635 -, 658 -, 666 +, 668 -, 687 -, 791 +, 851 -, 966 -, 1049 +, 1162 +, 1270 +, 1346 -, 1395 -, 1410 -, 1453 +, 1480 -	21	
	TATA*		926 -	1	
	TATA-box*		29 -, 30 -, 31 +, 36 +, 84 -, 85 -, 86 -, 87 +, 94 +, 95 +, 229 -, 230 +, 231+, 232 +, 233 +, 234 +, 235 +, 262 +, 263 +, 341 -, 343 +, 377 +, 399 +, 400 -, 401 +, 499 -, 502 -, 503 +, 510 -, 511 +, 643 -, 794 -, 825 -, 826 -, 827 -, 828 +, 829 -, 858 -, 860 -, 861 -, 862 -, 863 +, 864 -, 871 -, 872 -, 873 -, 874 -, 913 -, 920 +, 921 -, 922 -, 927 -, 928 -, 929 -, 930 -, 1065 -, 1067 -, 1072 +, 1073 -, 1074 -, 1203 +, 1205 -	65	

Gene		JcRIP11			
JcRIP11	**Name**	**function**	**Position**	**Number of elements**	**Total**
	GA-motif	part of a light responsive element	1448 -	1	5
	G-box	cis-acting regulatory element involved in light responsiveness	783 -	1	
	ATC-motif	part of a conserved DNA module involved in light responsiveness	1194 -	1	
	GATA-motif	part of a light responsive element	296 +	1	
	MRE	MYB binding site involved in light responsiveness	512 +	1	
	GCN4_motif	cis-regulatory element involved in endosperm expression	804 +	1	1
	ARE	cis-acting regulatory element essential for the anaerobic induction	959 +, 1313 +, 993 +	3	3
	TGA-element	auxin-responsive element	845 -	1	1
	O2-site	cis-acting regulatory element involved in zein metabolism regulation	1292 +	1	1
	circadian	cis-acting regulatory element involved in circadian control	294 +	1	1
	TCA	cis-acting element involved in salicylic acid responsiveness	856 -	1	
	CTAG-motif		1305 +	1	
	ERE	ethylene-responsive element	1224 +, 363 -	2	
	MYC	transcription factor MYC	500 -, 1472 +, 809 -	3	
	WUN-motif	wounding and pathogen response	1417 +	1	
	TATA-box*	core promoter element around -30 of transcription start	89 -, 269 -, 90 +, 270 +, 88 +, 268 -, 108+, 558 +, 109 -, 559 +, 110 +, 560 +, 176 +, 561 +, 177 -, 1043 +, 178 +, 1044 -, 179 +, 1051 -, 180 +, 1053 -, 214 +, 1055 -, 215 +, 1106 +, 216 +, 1107 -, 217 +, 1108 -, 218 +, 1135 -, 219 +, 1136 -, 249 -, 1137 -, 251 +, 1138 -, 252 +, 1150 +, 266 -, 1280 +, 267 -, 1282 -, 1341 -, 1465 +, 1466 -	47	
	AT~TATA-box*		215 +, 1053 -, 559 +, 1136 -, 217 +, 178 +	6	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	325 +, 767 +, 481 +, 1130 +, 480 +, 1037 -, 636 -, 1473 -, 318 -, 811 -, 552 -, 1273 -, 409 +, 1099 +, 647 +, 1146 +, 172 +, 809 +, 500 +, 1274 +, 327 -, 1013 +, 583 -	23	
	Unnamed__4*		880 +, 1278 +, 1220 +, 1309 -	4	
	Unnamed__6*		252 +	1	

Gene	JcRIP12				
JcRIP12	**Name**	**function**	**Position**	**Number of elements**	**Total**
	Box 4	part of a conserved DNA module involved in light responsiveness	975 -, 1019 -	2	8
	GA-motif	part of a light responsive element	346 +	1	
	GATA-motif	part of a light responsive element	1333 -	1	

	GT1-motif	light responsive element	1186 -, 1187-, 213 +	3	
	MRE	MYB binding site involved in light responsiveness	237 +	1	
	ARE	cis-acting regulatory element essential for the anaerobic induction	246 -, 407 +, 341 +	3	3
	Unnamed__1	60K protein binding site	575 +	1	1
	MBS	MYB binding site involved in drought-inducibility	1209 -, 1126 -	2	2
	MYB	transcription factor MYB	833 +, 1188 +, 255 -	3	7
	MYB-like sequence	transcription factor MYB	1188 +, 833 +	2	
	Myb	transcription factor MYB	1126 -, 1209 -	2	
	MYC	transcription factor MYC	751 +	1	1
	STRE	Stress Response Element	1348 +	1	1
	box S	wounding and pathogen response	1432 -	1	7
	W box	wounding and pathogen response	280 +, 1351 -	2	
	WRE3	wounding and pathogen response	44 +, 1431 -	2	
	WUN-motif	wounding and pathogen response	491 +, 684 -	2	
	CAAT-box*	common cis-acting element in promoter and enhancer regions	1101 -, 34 +, 291 +, 1323 -, 205 +, 423 -, 1244 -, 1472 -, 1114 -, 1379 -, 136 +, 311 -, 1304 -, 1325 -, 244 -, 623 +, 871 +, 97 +, 1449 -, 297 -, 1248 +, 224 -, 1142 +, 470 +, 1313 +, 193 +, 412 -, 278 -, 752 -	29	140
	TATA*		863 +	1	
	TATA-box*	core promoter element around -30 of transcription start	164 +, 163 -, 165 +, 746 +, 161 -, 166 +, 862 +, 863 -, 937 +, 162 -, 167 +, 938 -, 168 +, 731 +, 196 +, 730 -, 197 -, 939 -, 198 +, 961 -, 385 -, 962 -, 386 +, 963 -, 448 +, 964 -, 449 +, 965 -, 528 + ,966 +, 578 -, 967 -, 599 -, 994 -, 600 -, 1008 -, 601 -, 1009 -, 602 +, 1010 +, 603 +, 1011 -, 604 +, 1033 -, 630 +, 1034 -, 640 +, 1035 +, 727 -, 1036 -, 728 -, 1057 -, 729-, 1058 -, 1059 -, 1060 -, 1065 +, 1066 -, 1067 -, 1068 -, 1069 -, 1070 -, 1071 -, 1072 -, 1073 -, 1074 -, 1075 -, 1076 -, 1077 -, 1078 -, 1079 -, 1080 -, 1081 -, 1082 -, 1083 -, 1084 -, 1085 -, 1086 -, 1087 -, 1277 +, 1344 -, 1345 -, 1369 -, 1370 -, 1371 +, 1372 -, 1400 +, 1466 -, 1467 -, 1468 -, 1469 -	90	
	AT~TATA-box*		162 -, 164 +, 166 +, 600 +, 602 +, 963 -, 965 -, 1009 -, 1034 -, 1067 -, 1069 -, 1071 -, 1073 -, 1075 -, 1077 -, 1079 -, 1081 -, 1083 -, 1085 -, 1370 -	20	
	Unnamed__4*		485 -, 1172 +	2	2
	CTAG-motif*		1041 -	1	1

Apêndice 4: S4 Figure. Amino acid sequences of JcRIP from *Jatropha curcas* L. The signal peptide is underlined. The PF00161 functional domain is indicated in gray and the PF00652 functional domains in yellow. Putative phosphorylation sites are indicated in blue. Putative glycosylation sites are indicated in red. Conserved domains are squared.

>JcRIP1
MKGGKMNLSIMVAAWFCWSSIIFGWASAREIVCPFSSNQNYKAGSPPTLTITYDATTDKKNYAQFIKDLREAFGFSYSSH
EIPVLRATVAPNQKFIVAKDINVANLEVSLGLNVVNAYLVGYKVGGTSYFFNDPESLADAKTYLFTDTKQQTLSFTGSYA
DFLSRANVHREDVDLGVQALDNYIYTLEKSSKPADIAKPLVGFIEVVPEAARFKYIEKKVLSQISKTFRPGGDIISLENN
WGDLSYQIQKCVNGVFLKPVQLQRENYTNILVNNVTQVAGVMGVLLNAVNYKV

>JcRIP2
MKGGKMNLSIMVAAWFCWSCIIFGWASAREIGCPFSSNQNYKAGSTPTLTITYDAAADKKNYANFIRDLREAFGFSYSSH
EIPVLRATVAANQKFIVAKVIKVANLEVSLGLNVVNAYLVAYKAGGTSYFFNDPESLADAKKYLFTDTKQQTLSFTGSYA
DFLSRANVHREDVDLGVLALDNYIYILHKSSQPADIAKPLVGFIEMVPAAARFKYIEKKVLTQISETFRPRGVIISLENN
WGDLSYQIQKSVNGIFLKPVQLQRENYTNILVNNVTQVTGLMGVLLNAVNYKVSMEEIIFNDQKWLPWL

>JcRIP3
MKGGKMNLSIMVAAWFCWSCIIFGWASAREIVCPFSSNQNYKAGSTPTLAITYDATTDKKNYAQFIEDLREAFDFSYLSH
KIPVLRATVAANQKFIVAKVINSGDIEVSVGLNVINAYLVAYKAGSNSYFFNDSESLADAKKNLFTDTNQQTLAFTGSYA
DFESRAKLHREEVDLGVVALDNYVYTLEKSSQPADIAKPLVGFIEMVPEAARFKYIEKKISTQISKTFRPRGDIISLENN
WGDLSYQIQKSVDDVFLKPVQLQRENYTNILVNNVTQVKGLMGVLLNAVNYKVSMEEIIFNYQKWLPWL

>JcRIP4
MKGGNMKLCIMVAAWFCWSTIIFGSASARERAWPFSSNNNYAADSTPTLTITYIPDEDEQNYAKFITDLRETFGSSGLSH
GIPVLRATVAANQKFFVAKVINAGDIEVSVGLNVINAYLVAYKVGSNSYFFNDSESLADAKKYLFTDTKQQTLAFTGSYA
DFESRAKVHREEGDLGVVALDNYIYDLQKSSQPADIAKPLVSFIQMVSEAARFKYIENKVLDQISQTFRPRGDILSRENS
WDDLSYQIQKSVNDVFLNPVQLQLEDYSFYQVNNVNQVKDDMGILYNEANHKVSMEEIIINSQKWLPLL

>JcRIP5
MKRGNTKLCIMVAAWFCWSTIIFGWVSAREIDCPFSSNDQYTADSVPTLTITYDADEDKQNYAKFIKDLIQAFGSSGLSH
GIPVLRASGAANQKFFVAKVINSGDIEVSVGLNVINAYLVAYKVGSKSYFFNDTNFLADAKKYLFTDIQQQPELPFTGNY
ADFISKAKVHREEVDLGVVALDNYIYDLQKSSLPADIAKPLVSFIQMVSEAARFKYIENKVLDQISQTFRPRGDILSREN
FLSYQIQKSVNDVFLNPVQLQREDYSFYQVNNVNQVKDDMGILYNEANHKVSMEEIINNSKKWLPLLK

>JcRIP6
MKRGNTKLCIMVATWFCWSTIIFGWVSAREIDSPFSSNDQYTADSVPTLTITYDADKDKQNYAKFIKDLRQAFGSSGLSH
GIPVLRASVAANQKFFVAKVINAGDIEVSVGLNVINAYLVAYKVGSNSYFFNDPQYLADAKKYLFTDTKQQTLAFTGSYA
NFESFAKVHREQVDLGVVALDNYIYDLQKSSLPADIAKPLVSFIQMVSEAARFKYIENKVLDQISQTFRPRGDILSRENS
WDDLSYQIQKSVNNVFLNPVQLQREDYSFYQVNNVNQVKDDMGILYNEANHKVSTEEIIINS

>JcRIP7
MEGNMKVCVVVATWLCWTMIFGLARVIHPSAIHNCTVTVDEIPSVSFTITRIPGDDTAGYKQFMVDLREKLASGTTSHGV
PVLRSTASREAKYLLVNIINSGNKEITLGLNVISAYILSYKVGYNSYFFKDKAELKDAQKYLFTDTRQTTLDKFSGNYDS
FKAEGGDRETTDLGIGQLDSHIYTLHKSTLPKDIAKPLVCIIQMVSEATRYKYIENKIIDKISGSFRPKLDTITRENKWE
DLSDGIENADAKGNFKTEVRLQKEDGKEDIISSVNQIIGDMGILLYQKKKSYNPSFGQTTFGNLIXNKIKLLPRL

>JcRIP8
MKGNMNKVLIALVTWFCWTVVFGLARVYPLQTNGNSVAIASFTITRYGETDKADYLQFMEDIRSKLGTGTPSHGIPVLPV
SVSKNQKYTTVELINSGNKKVTIGLNVLDLYVVAYKLENQNSYFFNDPKFKDATDYLFKDTKQNPLKISGNYESLKTQGG
DRETVYLGVGQLDSFIYTLYESTLPKDIAKPLVCIIQMVSEATRYKYIENEVVKHISDRFLPKGDIISRENKWQDLSEAI
QKSVADEFTKPVQLQAPDYNVYNVYRVQQVKDDMGLLLNQANSGETEDITT

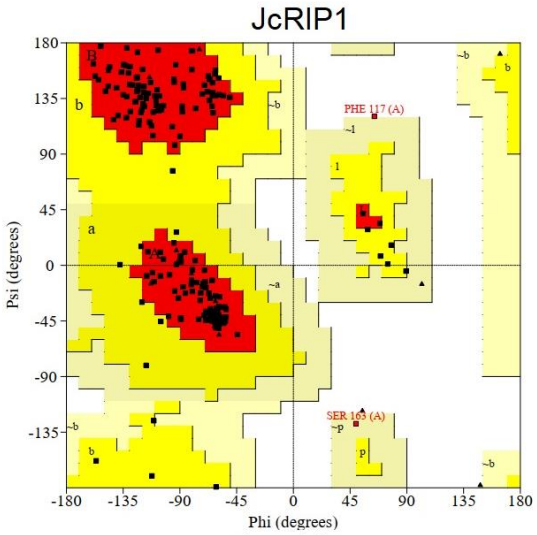
>**JcRIP9**
MEGKMKVWVVVATWLSWTVIFGLARFICPLAIHNHTVDAIPSVSFTISRIPDDDKTGYKQLMVDLRNKLASGTTSNGVPV
LRTTASKEAKYLLVNIINTGNKEITLGLNVINAYVLAYKVGDNSYFFNDPTELKDAQTYLFKDTKQNTIKMTGSYDSLKAQG
GDRETMDIGIGQLDSHIYTLHKSTALKDIAKPLVCIIQMVSEAARFKYIEKKIIDEVEGGFTPKLDVISRENNWGGLSEGIE
NADKNGKFKTTVRLQNEDSSAKVISKV DEIIVEMGVLLYVKKKSYIPSFEQTIFGIGNLILNQINLLPRL

>**JcRIP10**
MEGNMKVWVVVATLLCWTIIFGLDRVIHPSAIHNCTVIVDEIPFVSFTITGDPGDDAAGYKQFMIDLREKLASGTTSNGV
PVLRSTASNEAKYLLVNITNSGNKEITLGLNVISAYILAYKVGVNSYFFKDRSELKDAKKYLFTGTTQTILSKFSGNYHS
LKDEGGDRETTDLGIQQLDSHIYTLHKSTLPKEIAKPLVCIIQLVSEATRYKYVENKIIDEIIVGFRPKLDSITRENNWG
DLSKGIENADEKGNF

>JcRIP11
MEGNMKVWLVVATWLCWTIIFGLARVINPSAIHNYTADAIPSVSFTITRIPGDDKTDYKQLMVDLRKKLSSGTTSNGVPV
LRSTASKEAKYLLVNIINSGKKEITLGLNVINAYVLAYKVGDKSYFFNDPTELKDAQTHLFKDTKQTAIKITGSYDSLKA
QGGDRESVDLGIGQLDSHIYTLHKSTALKDVAKSLVCIIQMVSEAARFKSIENKIVDKIDGSFKPKLDIITRENNWGDLS
EGIQNADKKGNFKTKVRLQKEDGKEDIISNVNQIIGEMGILLYKKKKIYNIPSFGQTNFGNLIQN

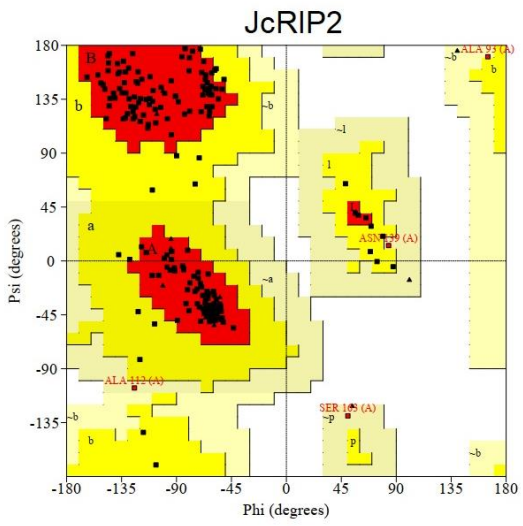
>JcRIP12
MKVWTWLCWTVIFASTCICTSKAEDGTNDNSLGYPTLNFTINVKDKEGYTRFLQSVRDQLTSGNQIHGISVLPDISTLPV
SQRFLLLKLSSSATTPITLALDVATAGVVAYGFQNQSYFFNDKLAFSNLFNDTKQNTLPFGSSYGELENSTGMPRSKIDL
GLLPLNEAVSNIVANNIKTEDLASCLIVVMQMVSEAARFRYIEHQVRWSTEKGQRFRPDGTIYSLENNWEALSTAVQESR
GGVFSEPVQLQRLNVKFNLDSVTRELVANLGLMLFISCKDNSTSSGGGGCATDFEPTVRILGRNGLCLDVSDEIYSDGNK
IHMWPCKSKPAANQLWTLKRDGTLQSNGKCLTANGSNPAGNNVVIYDCQTAMHNATQWAIWDNGTIINPQFGLVLTVNSG
DKGALLTLEKNIYASSQGFLATNNTQPFRSPILGQNDLCLQTNGTKVWVVECVSNRTEQKWALYADGSIRPQQNKDDCLT
CNKLDAKGTIFTIDSCSLASSGQRWAFRDDGIIFNLQNGMVMDVKKIDPSLEEVIIWPFNGGRNQRWLPLL

Apêndice 5: S5 Figure. Modeling validation of the JcRIPs structure using the Ramachandran plot.



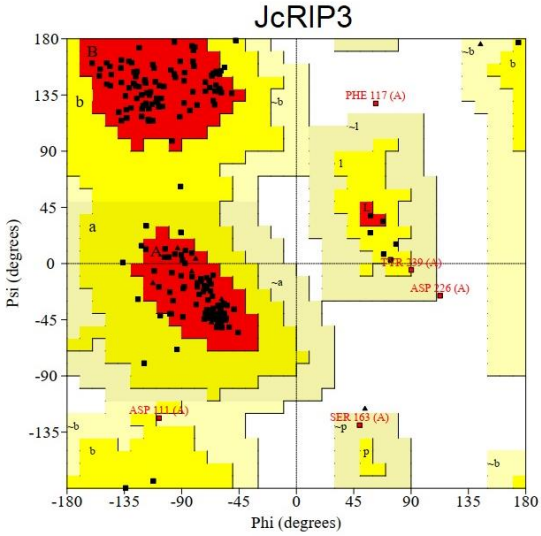
Plot statistics

Residues in most favoured regions [A,B,L]	202	91.4%
Residues in additional allowed regions [a,b,l,p]	17	7.7%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.5%
Residues in disallowed regions	1	0.5%
Number of non-glycine and non-proline residues	221	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	14	
Number of proline residues	10	
Total number of residues	247	



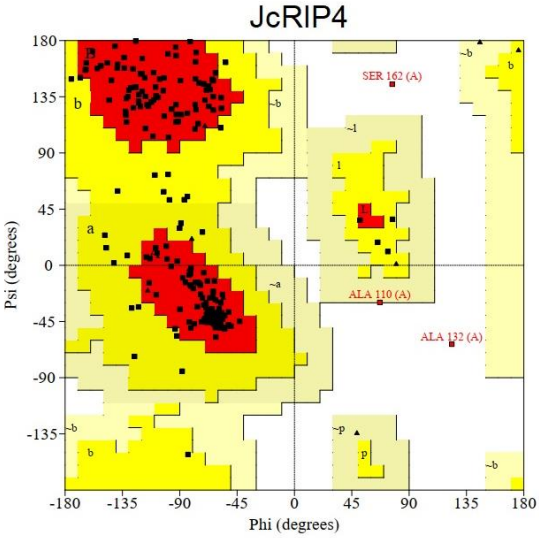
Plot statistics

Residues in most favoured regions [A,B,L]	203	89.4%
Residues in additional allowed regions [a,b,l,p]	20	8.8%
Residues in generously allowed regions [~a,~b,~l,~p]	4	1.8%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	227	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	12	
Number of proline residues	8	
Total number of residues	249	



Plot statistics

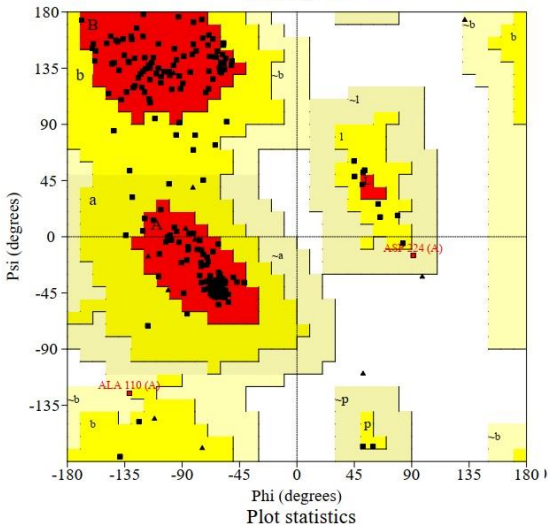
Residues in most favoured regions [A,B,L]	206	89.6%
Residues in additional allowed regions [a,b,l,p]	19	8.3%
Residues in generously allowed regions [~a,~b,~l,~p]	3	1.3%
Residues in disallowed regions	2	0.9%
Number of non-glycine and non-proline residues	230	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	10	
Number of proline residues	7	
Total number of residues	249	



Plot statistics

Residues in most favoured regions [A,B,L]	197	86.8%
Residues in additional allowed regions [a,b,l,p]	27	11.9%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.4%
Residues in disallowed regions	2	0.9%
Number of non-glycine and non-proline residues	227	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	11	
Number of proline residues	7	
Total number of residues	247	

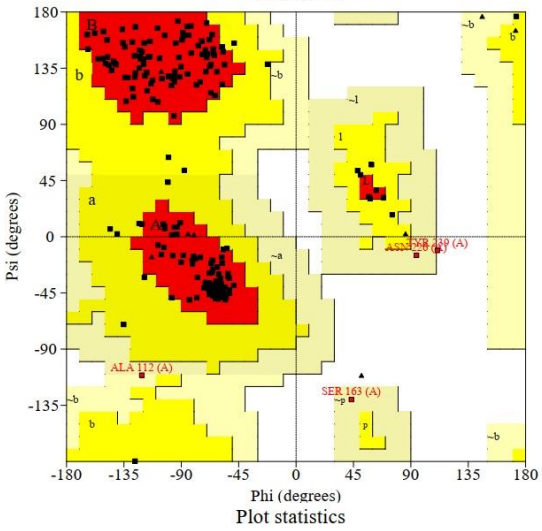
JcRIP5



Plot statistics

Residues in most favoured regions [A,B,L]	195	86.7%
Residues in additional allowed regions [a,b,l,p]	28	12.4%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.9%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	225	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	12	
Number of proline residues	8	
Total number of residues	247	

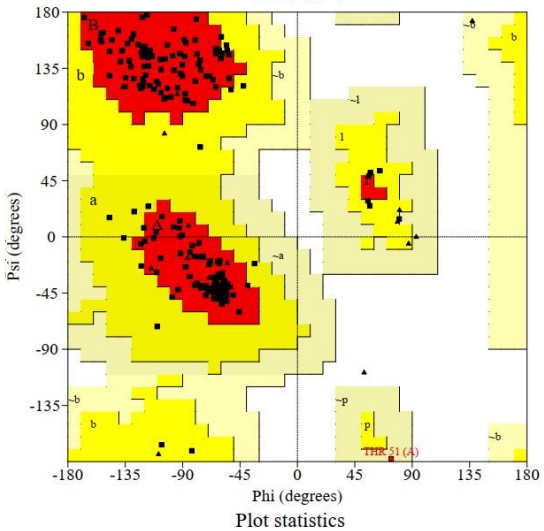
JcRIP6



Plot statistics

Residues in most favoured regions [A,B,L]	206	90.4%
Residues in additional allowed regions [a,b,l,p]	18	7.9%
Residues in generously allowed regions [~a,~b,~l,~p]	3	1.3%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	228	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	10	
Number of proline residues	7	
Total number of residues	247	

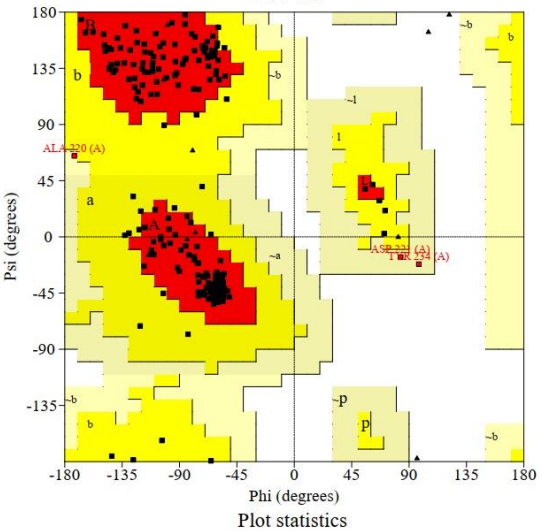
JcRIP7



Plot statistics

Residues in most favoured regions [A,B,L]	203	90.6%
Residues in additional allowed regions [a,b,l,p]	20	8.9%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.4%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	224	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	18	
Number of proline residues	6	
Total number of residues	250	

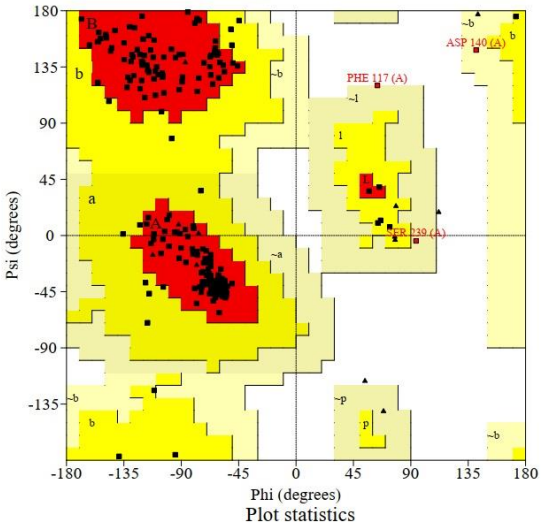
JcRIP8



Plot statistics

Residues in most favoured regions [A,B,L]	197	88.7%
Residues in additional allowed regions [a,b,l,p]	22	9.9%
Residues in generously allowed regions [~a,~b,~l,~p]	3	1.4%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	222	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	14	
Number of proline residues	10	
Total number of residues	247	

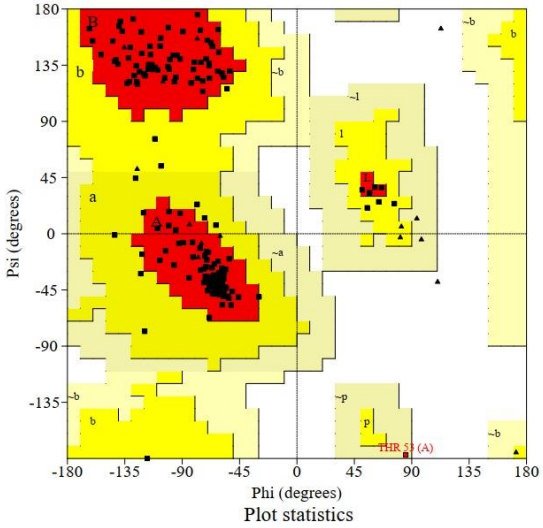
JcRIP9



Plot statistics

Residues in most favoured regions [A,B,L]	200	89.7%
Residues in additional allowed regions [a,b,l,p]	20	9.0%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.9%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	223	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	18	
Number of proline residues	6	
Total number of residues	249	

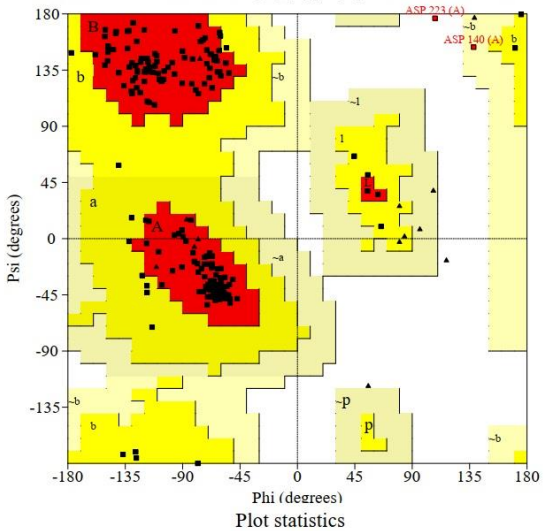
JcRIP10



Plot statistics

Residues in most favoured regions [A,B,L]	170	90.4%
Residues in additional allowed regions [a,b,l,p]	17	9.0%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.5%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	188	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	17	
Number of proline residues	5	
Total number of residues	212	

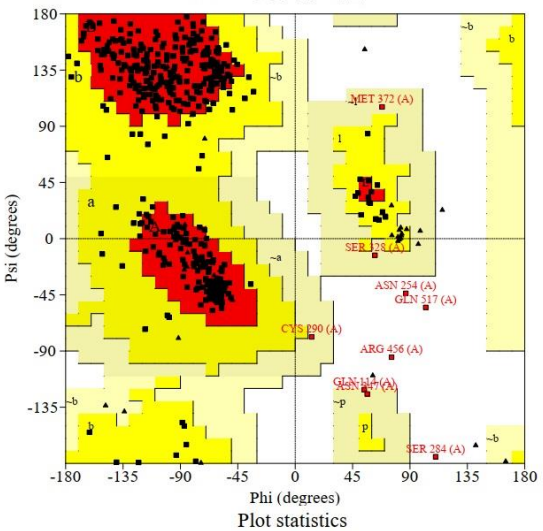
JcRIP11



Plot statistics

Residues in most favoured regions [A,B,L]	204	91.1%
Residues in additional allowed regions [a,b,l,p]	18	8.0%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.4%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	224	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	18	
Number of proline residues	5	
Total number of residues	249	

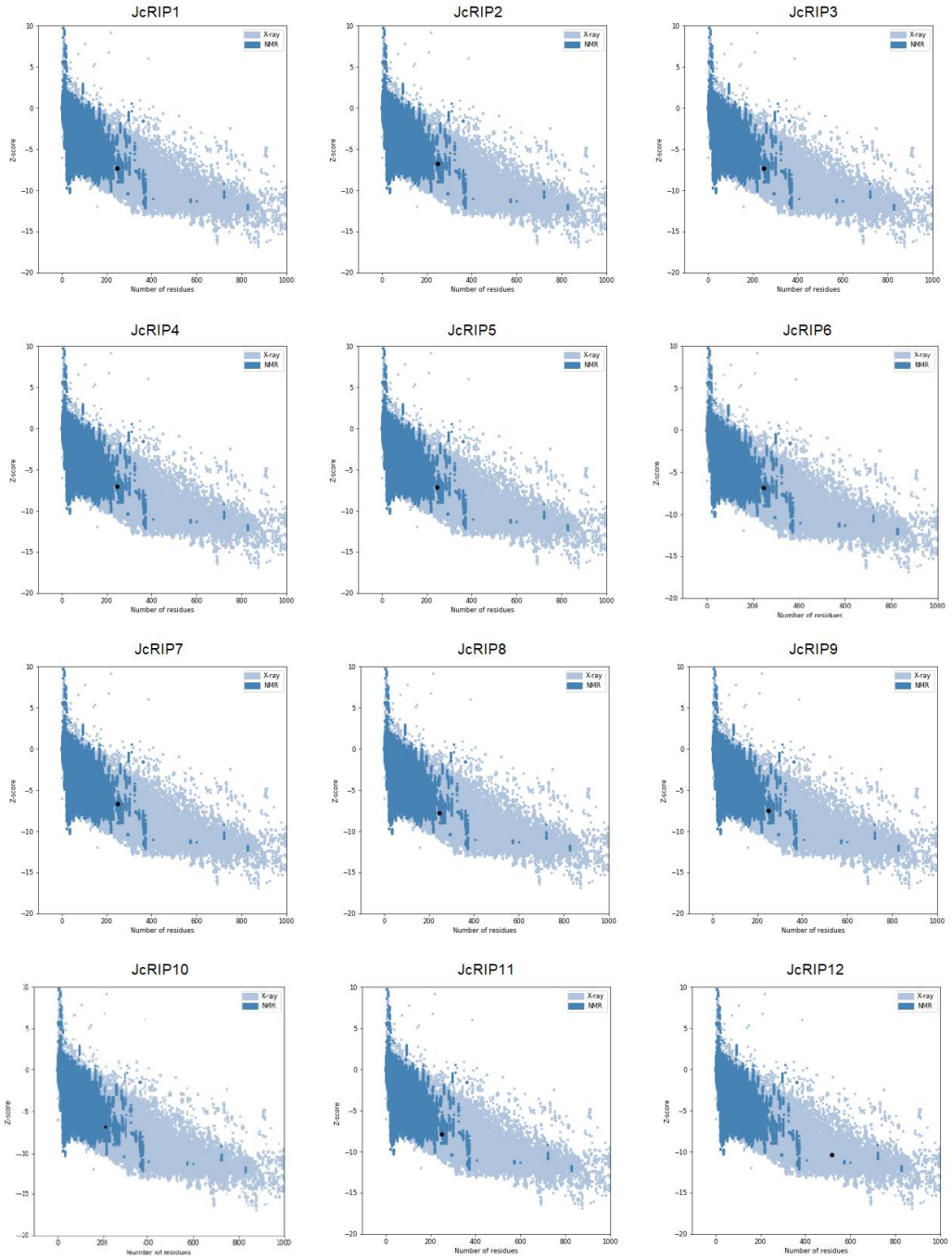
JcRIP12



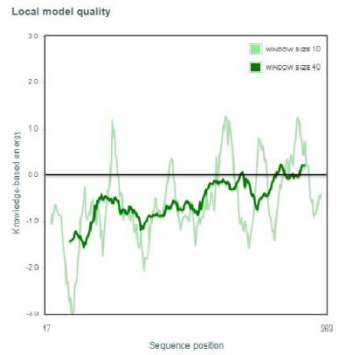
Plot statistics

Residues in most favoured regions [A,B,L]	392	85.8%
Residues in additional allowed regions [a,b,l,p]	56	12.3%
Residues in generously allowed regions [~a,~b,~l,~p]	5	1.1%
Residues in disallowed regions	4	0.9%
Number of non-glycine and non-proline residues	457	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	39	
Number of proline residues	20	
Total number of residues	518	

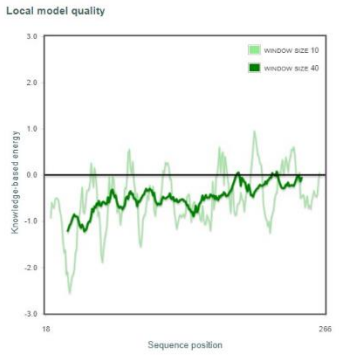
Apêndice 6: S6 Figure. Modeling validation of the JcRIPs structure using the ProSA analysis. The z-score indicates overall model quality. Its value is displayed in a plot that contains the z-scores of all experimentally determined protein chains in current PDB. The second plot shows local model quality by plotting energies as a function of amino acid sequence position. In general, positive values correspond to problematic or erroneous parts of the input structure.



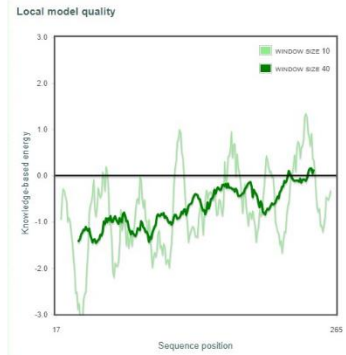
JcRIP1



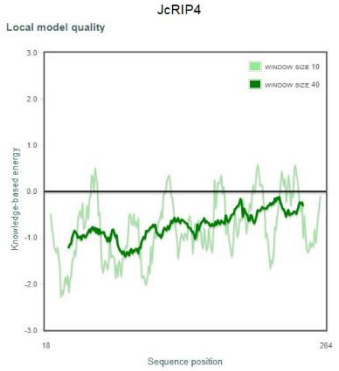
JcRIP2



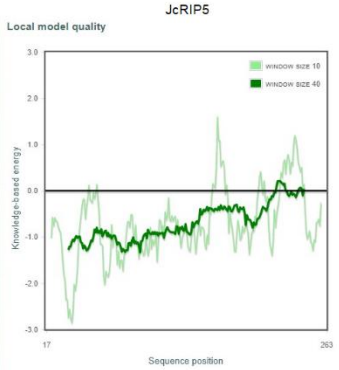
JcRIP3



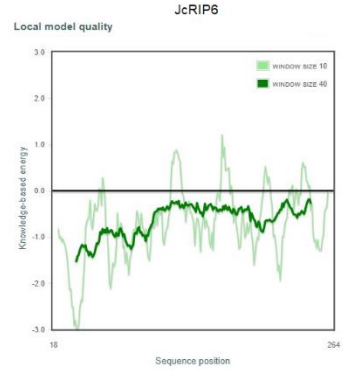
JcRIP4



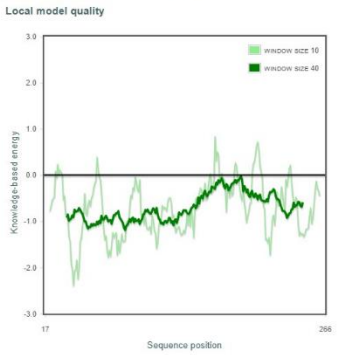
JcRIP5



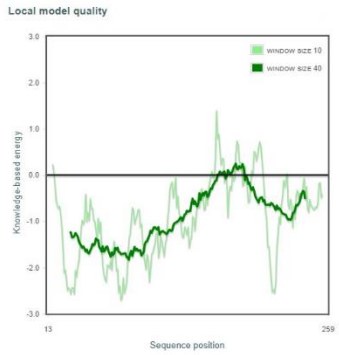
JcRIP6



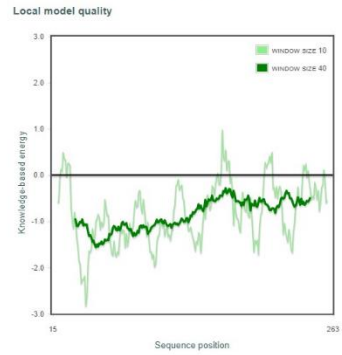
JcRIP7



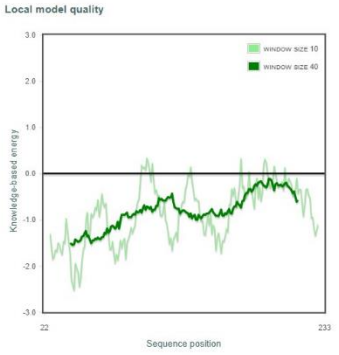
JcRIP8



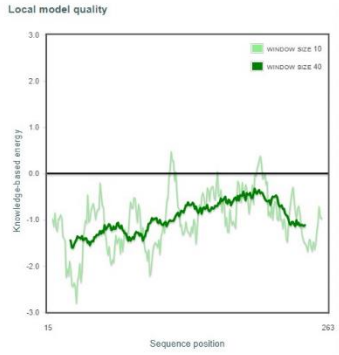
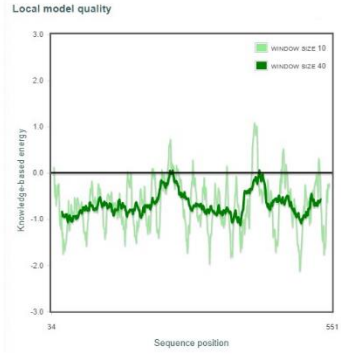
JcRIP9



JcRIP10



JcRIP11



JcRIP12