

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

**CARACTERIZAÇÃO GÊNICA E MODELAGEM POR
HOMOLOGIA DAS PROTEÍNAS RELACIONADAS À
PATOGENESE EM *Ricinus communis* L.**

HILÇANA YLKA GONÇALVES DE ALBUQUERQUE

**CRUZ DAS ALMAS – BAHIA
SETEMBRO - 2022**

**CARACTERIZAÇÃO GÊNICA E MODELAGEM POR HOMOLOGIA
DAS PROTEÍNAS RELACIONADAS À PATOGÊNESE EM *Ricinus
communis L.***

Hilçana Ylka Gonçalves de Albuquerque
Licenciatura em Ciências Biológicas
Universidade de Pernambuco, 2014

Tese 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 Doutora em Ciências Agrárias (Área de
Concentração: Fitotecnia)

Orientadora: Prof^a. Dr^a. Simone Alves Silva
Coorientadores: Dr^a. Jacqueline Araújo Castro
Dr. Carlos Augusto Dórea Bragança
Dr. Jair Wyzykowski

**CRUZ DAS ALMAS - BAHIA
SETEMBRO - 2022**

FICHA CATALOGRÁFICA

A345c	<p>Albuquerque, Hilçana Ylka Gonçalves de. Caracterização gênica e modelagem por homologia das proteínas relacionadas à patogênese em <i>Ricinus communis</i> L. / Hilçana Ylka Gonçalves de Albuquerque. Cruz das Almas, BA, 2022. 169f; il.</p> <p>Tese (Doutorado) – Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, Doutorado em Ciências Agrárias.</p> <p>Orientadora: Prof. Dr^a. Simone Alves Silva. Coorientadora: Prof. Dr^a. Jacqueline Araújo Castro. Coorientador: Prof. Dr. Carlos Augusto Dórea Bragança. Coorientador: Prof. Dr. Jair Wyzykowski.</p> <p>1.Essências e óleos essenciais – Mamona. 2.Óleos essenciais – Proteínas – Melhoramento genético. I.Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas. II.Título.</p> <p>CDD: 633.88</p>
-------	--

Ficha elaborada pela Biblioteca Universitária de Cruz das Almas - UFRB.
Responsável pela Elaboração –Antonio Marcos Sarmiento das Chagas
(Bibliotecário - CRB5 / 1615).


**UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS AMBIENTAIS EBIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AGRÁRIAS
CURSO DE DOUTORADO**

**CARACTERIZAÇÃO GÊNICA E MODELAGEM POR HOMOLOGIA
DAS PROTEÍNAS RELACIONADAS À PATOGÊNESE EM *Ricinus
communis L.***

Comissão Examinadora da Defesa de Tese de

Hilçana Ylka Gonçalves De Albuquerque

Realizada em 30 de setembro de 2022

Documento assinado digitalmente
 SIMONE ALVES SILVA
Data: 25/11/2022 13:17:05-0300
Verifique em <https://verificador.iti.br>

Prof. Dr^a. Simone Alves Silva
Universidade Federal do Recôncavo da Bahia - UFRB
Orientadora

Prof. Dr. Ricardo Franco Cunha Moreira
Universidade Federal do Recôncavo da Bahia - UFRB
Examinador Interno

Prof. Dr^a. Hellen Cristina da Paixão Moura
Universidade Federal do Recôncavo da Bahia - UFRB
Examinadora Interna

Prof. Dr^a. Dayanne Silva Monteiro de Almeida
Universidade Estadual de Santa Cruz - UESC
Examinadora Externa

Prof. Dr^a Rafaella de Lima Roque
Instituto Federal da Paraíba - IFPB
Examinadora Externa

DEDICATÓRIA

E se o aplauso eu receber, parabéns. No calvário irei Te gloriar.

À Deus eu dedico.

AGRADECIMENTOS

Meu tributo, meu reconhecimento por tudo quanto tens feito por mim. Por toda força que me trouxe até aqui. Minha eterna gratidão a Ti, Jesus Cristo.

Sou grata a minha família.

Sou grata a minha família científica, desde meu orientador da graduação, Dr. Francisco Pinheiro Neto e à Dr^a. Elaini Oliveira dos Santos Alves, que me acompanharam no início da minha vida acadêmica. Em especial aos meus orientadores do Mestrado, Dr. Eder Jorge de Oliveira e à Dr^a Ana Carla Brito, por todos os “tapas” que me levaram para frente e que muito contribuiu para o meu crescimento pessoal e profissional, e que também, me permitiu ser uma melhor estudante no doutorado. Tenho muito orgulho de ter sido orientada por vocês.

Agradeço a minha orientadora linda, humana, humilde e meu modelo de professora, Dr^a. Simone Alves Silva, por ter me recebido com muito carinho em seu grupo de pesquisa e por ter me permitido crescer enquanto profissional.

Aos meus coorientadores, Dr^a. Jacqueline Araújo Castro, Dr. Jair Wyzykowski e ao Dr. Carlos Augusto Dórea Bragança, também sou grata pelas contribuições científicas.

Aos integrantes do NBIO, Gean, Jacqueline, Monikuelly, Hélio, Francielly, Ilneide, Eduardo e a Ciro Filadelfo.

Também sou grata ao meu grupo de amigos mais chegados que um irmão, “Casa da Omma”, por serem o meu refúgio nos dias triste e por sempre me escutarem e me alegrarem, na pessoa de Janaína, AnaCarina, Deyse, Juliana Rosa; Raquel, Kennedy Oppa e a Vanessa, minha “Cumadre”.

Agradeço a Sávio, pela amizade ao longo da jornada científica.

Agradeço e à Josinilda e Josilândia minhas irmãs do coração, as quais posso contar sempre que precisar.

Agradeço à Universidade Federal do Recôncavo da Bahia e ao Programa de Pós-graduação em Ciências Agrárias pela estrutura e todo o programa.

Agradeço à CAPES pela concessão da bolsa de estudos.

SUMÁRIO

RESUMO	13
ABSTRACT	14
REVISÃO DE LITERATURA	15
Aspectos gerais sobre a composição e uso do óleo nas folhas e sementes e dados econômicos da cultura da mamona.....	15
Origem e diversidade genética na cultura da mamona.....	19
Recursos genéticos e Melhoramento na cultura da mamona.....	20
O patossistema da mamoneira x <i>Amphobotrys ricini</i>	24
Sistema Imune e Proteínas Relacionadas a Patogêneses – PRs.....	28
Caracterização Gênica e Modelagem de Proteínas por Homologia	32
REFERÊNCIAS	36
ARTIGO 1 – Caracterização das famílias gênicas codificantes de Proteínas Relacionada à Patogênese (PRs) em <i>Ricinus communis</i> L.	46
RESUMO	47
ABSTRACT	48
INTRODUÇÃO	49
MATERIAL E MÉTODOS	51
Análise <i>in silico</i> de genes RcPRs.....	51
Análise das regiões promotoras dos genes RcPRs.....	51
Análise <i>in silico</i> de proteínas PRs.....	52
Alinhamentos e filogenia da estrutura genética das RcPRs.....	52
RESULTADOS	53
Famílias gênicas PRs no genoma de <i>Ricinus communis</i> L.....	53
Análise das sequências promotoras dos genes RcPRs.....	56
Análise das proteínas das RcPRs.....	58
Relação filogenética.....	64
DISCUSSÃO	67
Família de genes PRs no genoma de <i>Ricinus communis</i> L.....	67
Análise da região promotora dos genes <i>RcPRs</i>	68
Análise das proteínas RcPRs.....	72
Relação filogenética.....	80
CONCLUSÃO	80
REFERÊNCIAS	82
ANEXOS	93
ARTIGO 2 - Resolução da Estrutura Proteica das Proteínas Relacionadas à Patogênese (PRs) via Modelagem por Homologia em <i>Ricinus communis</i> L.	151
RESUMO	152
ABSTRACT	153
INTRODUÇÃO	154
MATERIAL E MÉTODOS	155
Identificação das proteínas-PR no genoma de <i>Ricinus communis</i> L.....	155
Modelagem molecular.....	155
Predição de Sítios Ativos.....	156
RESULTADOS	156

Modelagem Molecular das RcPRs.....	156
Predição dos sítios ativos.....	161
DISCUSSÃO	163
Perspectivas Futuras.....	164
CONCLUSÃO	165
REFERÊNCIAS	166
ANEXOS	168
CONSIDERAÇÕES FINAIS	175

CARACTERIZAÇÃO GÊNICA E MODELAGEM POR HOMOLOGIA DAS PROTEÍNAS RELACIONADAS À PATOGÊNESE EM *Ricinus communis* L.

Autora: Hilçana Ylka Gonçalves de Albuquerque
Orientadora: Dra. Simone Alves Silva

RESUMO: O óleo oriundo da mamona (*Ricinus communis* L.) apresenta diversas aplicações como a produção de biocombustíveis e polímeros para fabricação de próteses ortopédicas, dentre outras 400 aplicabilidades diferentes. Assim, este trabalho objetivou realizar a caracterização gênica de 20 genes precursores das Proteínas Relacionadas a Patogêneses (PRs) em mamona e realizar a Modelagem por Homologia de cinco proteínas-PR. A caracterização gênica foi realizada em níveis de sequência nucleotídica e peptídica, incluindo a região promotora e análise filogenética, além da resolução da estrutura tridimensional das proteínas RcPR-1c, RcPR-4b; RcPR-5a; RcPR-6 e RcPR10-STH21 via modelagem por homologia. A caracterização gênica permitiu identificar que dos 20 genes RcPR, 10 são *RcPR-1*; três *RcPR-4*; cinco *RcPR-5*, uma *RcPR-6* e uma *RcPR10-STH21*, com éxons variável de 1 a 3 e íntrons de 1 a 2, em um painel de leitura aberto de 222 a 795 pb. A região de 1500pb a montante de cada gene PR apresentou elementos *cis*-reguladores relacionados a fatores de resposta internos e externos. O ponto isoelétrico e o peso molecular das proteínas com Peptídeo Sinal foi de 4,68 a 9,17 e 14,96 a 31,26 kDa e sem o PS foi de 4,58 a 9,35 e 12,67 a 28,83 kDa, respectivamente. O índice médio de hidropaticidade indicou presença de regiões hidrofóbicas nas proteínas RcPR-5a e RcPR-5e, e nas demais, regiões hidrofílicas. Já a modelagem por homologia permitiu descrever de forma detalhada as sequências, onde o grau de identidade variou de 49,37 a 76,23%, com uma semelhança de 44 a 56% e uma cobertura de 60 a 99%. A qualidade estereoquímica de acordo com o Diagrama de *Ramachandran* foi que mais de 90% dos resíduos de aminoácidos das cinco proteínas-PR encontram-se nas regiões favorecidas. Estes resultados permitirão entender o papel biológico das proteínas-PR no patossistema com *R. communis*, além de auxiliar no desenvolvimento de estratégias de controle a fitopatógenos invasores, à exemplo do *Amphobotrys ricini*.

Palavras-chave: Melhoramento genético, Mofo-cinzento, Patógenos

GENE CHARACTERIZATION AND HOMOLGY MODELING OF PROTEINS RELATED TO PATHOGENESIS IN *Ricinus communis* L.

Autora: Hilçana Ylka Gonçalves de Albuquerque
Orientadora: Dra. Simone Alves Silva

ABSTRACT: The oil from the castor bean (*Ricinus communis* L.) has several applications such as the production of biofuels and polymers for the manufacture of orthopedic prostheses, among other 400 different applications. This work aimed to carry out the gene characterization of 20 genes of the Pathogenesis Related Proteins (PRs) in castor bean and to carry out the Homology Modeling of five proteins-PR. Gene characterization was performed at nucleotide and peptide sequence levels, including the promoter region and phylogenetic analysis, in addition to the resolution of the three-dimensional structure of the proteins RcPR-1c, RcPR-4b; RcPR-5a; RcPR-6 and RcPR10-STH21 via homology modeling. The gene characterization allowed us to identify that of the 20 RcPR genes, ten RcPR-1; three RcPR4; five RcPR-5, one RcPR-6 and one RcPR10-STH21, with variable exons from 1 to 3 and introns from 1 to 2, in an open reading frame from 222 to 795 bp. The 1500bp region upstream of each PR gene showed cis-regulatory elements related to internal and external response factors. The isoelectric point and molecular weight of proteins with Signal Peptide ranged from 4.68 to 9.17 and 14.96 to 31.26 kDa and without PS ranged from 4.58 to 9 .35 and 12.67 to 28.83 kDa, respectively. The average hydropathicity index indicated the presence of hydrophobic regions in the RcPR-5a and RcPR-5e proteins, and in the others, hydrophilic regions. On the other hand, homology modeling allowed a detailed description of the sequences, where the degree of identity varied from 49.37 to 76.23%, with a similarity of 44 to 56% and a coverage of 60 to 99%. The stereochemical quality according to the Ramachandran Diagram was that more than 90% of the amino acid residues of the five PR proteins are found in the favored regions. These results will allow us to understand the biological role of PR-proteins in the pathosystem with *R. communis*, in addition to helping to develop strategies to control invading phytopathogens, such as *Amphobotrys ricini*.

Keywords: Plant breeding, Gray mold, Pathogens

REFERENCIAL TEÓRICO

Aspectos gerais sobre a composição e uso do óleo nas folhas e sementes e dados econômicos da cultura da mamona

A mamona (*Ricinus communis* L.) é uma espécie diplóide ($2n = 2x = 20$), pertencente à família Euphorbiaceae. *Linnaeus* identificou o gênero *Ricinus* em 1753 (ANJANI, 2012). Além de *R. Communis* L., dentre as Euphorbiaceae, destacam-se a mandioca (*Manihot esculenta* Crantz.), seringueira (*Hevea brasiliensis* L.) e o pinhão manso (*Jatropha curcas* L.), por serem de grande importância econômica e que se divergiram a aproximadamente 35 milhões de anos atrás (BREDESON et al., 2016), sendo a *R. communis* L. a única espécie do gênero *Ricinus* (RAMPRASAD; BANDOPADHYAY, 2010).

Os primeiros estudos publicados no meio acadêmico sobre a cultura da mamona datam de 55 anos atrás (CANVIN, 1963), e relatam a utilidade tanto das folhas como do óleo presente nas sementes. As folhas têm mostrado grande potencial como fonte de moléculas com propriedades inseticidas (ELIMAM et al., 2009; ROSSI et al., 2012), este dado é relevante, pois estas propriedades podem ser utilizadas em outras culturas, a exemplo da cultura do milho que no ano de 2006 teve cerca de 31% de perda em sua produção mundial devido ao ataque de insetos (ROSSI et al., 2012).

Ramos-López et al. (2010) mostraram que extratos hexânicos, metanólicos e de acetato de etila de folhas e sementes exerceram atividade inseticida em larvas de *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera: Noctuidae). Resultados satisfatórios também foram encontrados por Mushobozy et al. (2009), que observaram ação inseticida do óleo de mamona em *Zabrotes subfasciatus* (Boh.) (Coleoptera: Bruchidae).

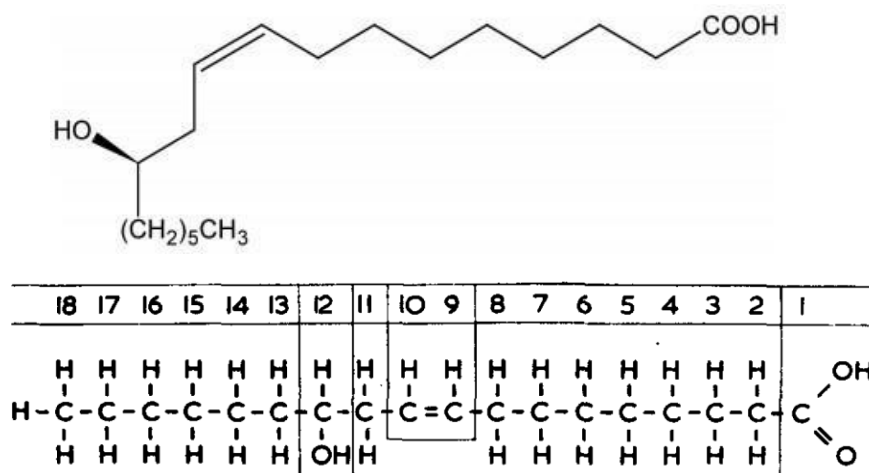
Além disso, o uso de inseticidas convencionais causa efeitos indesejáveis tanto para as espécies não-alvo quanto para o ambiente, e aceleram a seleção de resistência nos insetos (ROSSI et al., 2012). Dessa forma, uma das estratégias para o controle de tais pragas é o uso das moléculas inseticidas presentes nas folhas da mamona.

Quanto ao óleo presente nas sementes, de fato, é o componente de maior valor agregado a cultura, com predominância do ácido graxo ricinoleico em sua composição (~85%), possuindo propriedades químicas únicas ($C_{18}H_{34}O_3$), apresentando 18 carbonos ao longo de sua cadeia, com um grupo hidroxila no

átomo do carbono 12 e uma ligação dupla *cis* entre os carbonos 9 e 10, conforme visto na Figura 1 (ACHAYA, 1971), sendo esta posição na composição do ácido graxo com às ligações de hidrogênio que confere ao óleo elevado índice de viscosidade e estabilidade entre os demais óleos vegetais (OGUNNIYI, 2006). Além da presença do ácido ricinoleico, outros ácidos graxos em pequenas quantidades também compõe o óleo, como o linoleico (4,2%), oléico (3%), esteárico (1%), dihidroxiesteárico (0,7%), linolênico (0,3%), palmítico (1%) e eicosanóico (0,3%) (OGUNNIYI, 2006).

O óleo nas sementes da mamona apresentam alta lubricidade, especialmente sob condições de temperatura ambiente, o que permite a sua aplicabilidade em diversos segmentos na indústria (MUTLU; MEIER, 2010), na agricultura e na produção de medicamentos, particularmente em casos de certos exames radiológicos que requerem evacuação imediata e completa no intestino delgado, pois aumenta o movimento dos músculos intestinais, ajudando a expelir o bolo fecal e aliviar a constipação intestinal (ALLAN et al., 2008).

Figura 1. Composição química do ácido graxo ricinoléico.



Autor: ACHAYA, 1971.

Outro elemento importante na composição da semente é a presença da ricina, uma glicoproteína dimérica constituída de cadeias polipeptídicas A e B ligadas covalentemente por ligação dissulfeto (LAPPI et al., 1978), considerada uma proteína extremamente tóxica e um dos venenos naturais mais letais quando administrada via intravenosa, via oral ou quando inalada por meio de partículas menores, e até isso porque, apenas uma molécula de ricina da cadeia A é capaz

de inativar de modo irreversível aproximadamente mil ribossomos por minuto, prejudicando a síntese proteica e resultando na morte celular (OLSNES et al., 1975).

Estudos de toxicidade aguda da ricina foram realizadas em camundongos, a fim de determinar os valores de doses letais tanto por via intraperitoneal quanto via oral e os valores foram 2,4 a 36 µg/kg e 21 a 30 mg/kg, respectivamente (WORBS et al., 2011). Dessa forma, foi possível estimar que em seres humanos seja de 1 a 20 mg/kg, sendo muito mais tóxico se a exposição for por inalação (1-10 µg/kg) (WORBS et al., 2011), levando o paciente a apresentar diarreia seguida por fezes escuras, anorexia, dor abdominal, fraqueza no corpo, falta de apetite e câimbras como sintomas (SOUSA et al., 2017).

Após a descoberta do grau de letalidade que essa proteína apresenta, preocupações quanto à biossegurança mundial surgiram e por esta razão, países como os Estados Unidos não produzem extensivamente o óleo da mamona como também não incentivam o cultivo de *R. communis* (MUSSHOF et al., 2009).

Além disso, o Centro de Controle e Prevenção de Doenças (*Centers for Disease Control and Prevention* - CDC), caracterizou a ricina como um agente biológico de categoria B que pode ser usado para fins de bioterrorismo (MUSSHOF et al., 2009; HE et al., 2010; BOZZA et al., 2015), sobretudo, ainda são um dos maiores importadores do óleo e de seus coprodutos devido à versatilidade de aplicações que este óleo apresenta (MUTLU; MEIER, 2010).

Sucessivos métodos de desintoxicação foram desenvolvidos, a fim de inibir a ricina, contudo, o sucesso foi limitado (SOUSA et al., 2017). Procedimentos de fermentação com microrganismos, como *Paecilomyces variotii* e *Aspergillus niger*, foram capazes de desintoxicar durante o processo de produção de biodiesel (MADEIRA et al., 2011; FERNANDES et al., 2012).

Ademais, a desintoxicação também foi obtida por tratamento químico com compostos de cálcio (óxido de cálcio e hidróxido de cálcio), que do ponto de vista econômico, esses processos ainda não são práticos e eficientes o suficiente para serem utilizados em larga escala (FERNANDES et al., 2012).

Observa-se que a mamona desempenha um papel essencial na economia das regiões áridas e semiáridas do mundo, devido a sua adaptação a essas condições, como também em regiões de clima que apresentem temperaturas baixas e úmidas (FALASCA et al., 2012; CARRINO et al., 2020).

Na Índia, por exemplo, encontra-se a maior distribuição da diversidade

genética da espécie, que se encontra em todo território, onde formas perenes são encontradas crescendo em diversos *habitats*, como florestas, dunas de areia assentada, costa do mar, leitos de rios, minas de carvão a céu aberto, topos de montanhas, vales de montanhas, margens de estradas, barreiras de campo, trilhos de trem, até mesmo em áreas de lixões, terrenos baldios, quintais e muitas outras áreas em todo o país (ANJANI, 2012).

Além do mais, o germoplasma mantido nos centros de pesquisa indiano é aumentado principalmente por meio do cultivo de germoplasma em todo o país, devido a amplitude de material genético que precisa ainda ser explorado por toda a região.

Em relação a China, desde 1949 (CAO et al., 1997), pesquisas de exploração, coleta, preservação, caracterização e avaliação de recursos de germoplasma de diferentes culturas, por exemplo, são realizadas e como o país possui uma distribuição geográfica de oleaginosas ampla, à exemplo do pinhão manso (*Jatropha curcas* L.), soja (*Glycine max* L. LIU et al., 2019), segundo o *Ministry of Agriculture and Rural Affairs* e Academia de Ciências do Agronegócio da China – CAAS (https://www.cnabrazil.org.br/assets/images/China-Agricultural_v5.pdf) e da colza (*Brassica napus* L. LIU et al., 2021).

Desde então, esforços têm sido mantido para que oleaginosas como a mamoneira sejam amplamente produzidas por meio de um sistema de manejo mecanizado adaptado as variedades que são produzidas de baixo porte. (ANJANI, 2012).

Em relação aos aspectos econômicos da cultura, o continente asiático, possui a maior concentração de produção de óleo de mamona seguido das Américas, África e Europa (FAOSTAT, 2020). Onde a Índia destaca-se com uma produção de mais de 1.178 mil toneladas, seguida da China com uma produção de 158 mil de toneladas e o Brasil com uma produção de 67,24 mil de toneladas, ocupando assim, o terceiro lugar na produção mundial (FAOSTAT, 2020).

No Brasil, dentre as principais regiões produtoras, há destaque para a região Nordeste, responsável por mais de 90% de toda produção nacional. As estimativas para a safra 2021/22 é de aumento da produção em 21,9% em relação à safra anterior (CONAB, 2021). Entre os estados nordestinos produtores, a maior concentração está na Bahia, em relação à produção total do estado, cultivada principalmente no centro-norte do estado, tendo como polos produtores a microrregião de Irecê, Ourolândia, Lapão, São Gabriel, Cafarnaum e Canarana

(CONAB, 2021).

As condições climáticas favoráveis que estas regiões do estado da Bahia apresentam durante a época de plantio (novembro a janeiro), onde o solo apresenta umidade satisfatória e as precipitações pluviométricas são favoráveis, beneficiam uma boa germinação e um estande mais uniforme. Entretanto, essas condições edafoclimáticas também contribuem para a ocorrência de diversas pragas e doenças que dizimam as lavouras de plantio (FRACETTO et al., 2015).

O Brasil precisa investir em tecnologias que visem aumentar o rendimento da mamona e a obtenção de sementes de alta qualidade que incluam o desenvolvimento de novos materiais genéticos e melhorias no manejo de culturas que envolvam solo, qualidade de sementes, nutrição e controle de doenças e pragas, pois, são esses desafios que com prometem o cultivo da espécie, comprometendo a expansão tanto da produção quanto da produtividade.

Origem e diversidade genética na cultura da mamona

A mamona é um arbusto perene tropical originário da África oriental (MOSHKIN, 1986, ANJANI, 2012). O seu cultivo foi difundido pelas regiões tropicais e subtropicais do mundo (CHAN et al., 2010). No Brasil, foi trazida pelos portugueses e o clima tropical favoreceu sua expansão (SANTOS et al., 2007). As plantações comumente escapam do cultivo e são encontrados em locais como margem de estradas e de riachos, lotes abandonados e bordas de campos agrícolas, de tal forma que a espécie é muitas vezes considerada invasora em grande parte dos locais onde é introduzida (WEBER, 2003; ALLAN et al., 2008).

A propagação desta espécie ocorre por meio de sementes e suas plantas podem tanto realizar autopolinização (autógama) quanto apresentar polinização cruzada (alógama), assim, considerada uma cultura de reprodução mista, de modo que a maioria das cultivares consiste em populações de polinização aberta (FOSTER et al., 2010; OLIVEIRA et al., 2012).

Foster et al. (2010), ao estudarem a diversidade genética e a estruturação populacional de *R. communis* L. com base em marcadores do tipo *Single Nucleotide-Polymorphism* – SNP identificaram insatisfatórios níveis de diversidade, devido a natureza do marcador, pois detecta variação em um único nucleotídeo A, T, G ou C (ALLAN et al., 2008). Entretanto, tais resultados poderiam ser diferentes se usados marcadores com natureza multi-alélica e com maior nível de

polimorfismo, à exemplo dos *Simple Sequence Repeat* – SSR (CHAUDHARI et al., 2019). No entanto, para a espécie *R. communis* ainda não há estruturação genética definida.

Na mamona, outro fator que também contribui para a baixa diversidade existente é o insipiente conhecimento da diversidade na espécie, em relação a outras culturas de importância econômica e alimentícia como a mandioca (*Manihot esculenta* L.), sobretudo, esforços para sequenciar o genoma da espécie revelaram um genoma de tamanho moderado (~350Mb) organizado dentro de 10 cromossomos, mas, tanto quanto se sabe, ainda não há um mapa de ligação disponível para a espécie, devido à falta de informação genômica disponível (CHAN et al., 2010; FOSTER et al., 2010). Até o momento, não há atualizações disponíveis sobre o andamento de trabalhos de sequenciamento na espécie.

Contudo, espera-se que nos próximos anos estudos aprofundados na genética da mamoneira avancem rapidamente devido aos esforços já iniciados em sequenciar o genoma da espécie, atividade essencial em qualquer programa de melhoramento genético de espécies vegetais.

Recursos genéticos e Melhoramento na cultura da mamona

A priori, na mamona, houveram muitos incentivos para produção de biocombustíveis, no entanto, as demais aplicabilidades do óleo nos diferentes setores industriais, cosméticos e farmacêuticos, tem agregado um valor muito maior a cultura, e com isso o crescimento e avanço em pesquisas científicas (ANDRADE et al., 2017).

Os principais Bancos Ativos de Germoplasma - BAG de mamoneira estão localizados em 10 países contendo aproximadamente 11.300 acessos (SEVERINO et al., 2012), dos quais os principais estão na Unidade de Pesquisa de Manutenção de Germoplasma na Diretoria de Oleaginosas - Índia com 4.307 acessos, das quais 365 são coleções exóticas de 39 países (ANJANI, 2012); nos centros de pesquisas *Institute of Crop Science* e *Institute of Oil Crops Research* na China, com 1.689 e 1.652 acessos conservados, respectivamente (SEVERINO et al., 2012).

No Departamento de Agricultura dos Estados Unidos que contém 1.033 acessos *in vivo* e é mantida pela Unidade de Conservação de Recursos Fitogenéticos - PGRCU, em Griffin, GA, que tem como objetivo desenvolver

cultivares de mamona para produção de biodiesel utilizando genótipos oriundos de expedições de coletas internacionais (WANG et al., 2011); em um BAG com 424 acessos na Colômbia por Corpoica (SEVERINO et al., 2012) e no Brasil mantido pela Empresa Brasileira de Pesquisa e Agropecuária (EMBRAPA) junto com o Centro Nacional de Pesquisa de Algodão (CNPQ), em parceria, formando um único BAG com 620 acessos (SEVERINO et al., 2012) e o mantido na Universidade Federal do Recôncavo da Bahia - UFRB, com 203 linhagens puras, conservada em BAG na forma *ex situ* e em câmara refrigerada para estocagem de sementes (SILVA et al., 2017).

Os recursos disponíveis nestes BAGs internacionais são bastante diversos. Nos germoplasmas indianos foram avaliadas e identificadas várias características morfoagronômicas desejadas, por exemplo, plantas anãs, com floração e amadurecimento precoce (ANJANI, 2010; 2012). Muitos acessos também foram identificados possuindo resistência às principais pragas e doenças (ANJANI, 2010). Ainda, segundo Severino et al. (2012), esforços foram iniciados pela *Directorate of Oilseed Research* da Índia para desenvolver coleções nucleares com um número mínimo de genótipos, porém, representativos de mamona e validá-lo em diferentes condições de cultivo, diminuindo assim os custos financeiros, além do número de genótipos duplicados.

No repositório do Departamento de Agricultura dos Estados Unidos -USDA, foi encontrado por sementes variando de 370 a 610g na quantidade de óleo (WANG et al., 2011); e uma grande variação no conteúdo de ricina (PINKERTON et al., 1999), e alguns desses acessos foram usados para o desenvolvimento de '*Brigham*', uma cultivar com baixo teor de ricina (AULD et al., 2009).

Na Colômbia, segundo Severino et al. (2012), muitos esforços têm sido colocados para caracterizar todo material conservado no país. Já na China, apesar do cultivo de mamoneira ser feito há mais de 1.400 anos, a maioria das variedades popularizadas foram desenvolvidas no norte do país, mas expressou baixa resistência a doenças e tolerância ao calor e à umidade no sul da China. Entretanto, esforços têm sido mantidos em estudar a diversidade genética dos acessos de mamoneira do sul da China para fornecer uma base para proteção, pesquisa e utilização eficiente de todo recurso genético conservado (AGYENIM-BOATENG et al., 2019).

Os atuais programas de melhoramento da mamoneira, de maneira geral,

visam selecionar genótipos com maior resistência a estresses abióticos, tais como inibição de metais pesados, como o cádmio (Cd) bastante tóxico para plantas, animais e seres humanos (SHI; CAI 2009); *déficit hídrico* (BABITA et al., 2010); alto teor de sal no solo e na água de irrigação, resultando em alterações morfológicas, estruturais e metabólicas (FERRAZ et al., 2015) e ao metal alumínio (Al^{+3}), considerado tóxico no solo (PASSOS et al., 2015).

Busca-se também resistência a fatores bióticos, principalmente devido à baixa disponibilidade de cultivares resistentes a pragas e doenças, como as causadas por insetos, do tipo, cigarrinha (*Empoasca flavescens*F.); por bactérias como a mancha bacteriana causada por *Xanthomonas axonopodis* pv. *ricini*; ou por fungos como a podridão radicular causada pelo *Macrophomina phaseolina* (Tassi) Goid (PARMAR et al., 2018), a murcha causada pelo *Fusarium oxysporum* f. sp. *ricini* (SHAW et al., 2017) e o mofo-cinzento causado pelo fungo *Amphobotrys ricini* Buchw. considerada umas das doenças de maior importância econômica em *R. communis* L. (YAMUNA et al., 2015).

De acordo com Lima et al. (2012), no Brasil, os trabalhos de melhoramento genético com a mamoneira iniciaram-se pelo Instituto Agrônomo de Campinas – IAC, em 1936 (KRUG et al., 1943). Após um ano, em 1937, foram instalados ensaios de competição de genótipos de portes alto e anão objetivando à identificação de genótipos mais produtivos, bem como a realização de trabalhos de melhoramento com a cultura.

Na Bahia, os trabalhos com o melhoramento da espécie iniciaram-se na década de 60 pelo Instituto de Pesquisa e Experimentação Agropecuária do Leste – IPEAL, com sede na cidade de Cruz das Almas, com a extinção desse órgão, os trabalhos passaram então a ser conduzidos a partir de 1974 pela Empresa de Pesquisa Agropecuária da Bahia – EPABA, que mais tarde se tornou a Empresa Baiana de Desenvolvimento Agrícola - EBDA, que se manteve atuante em 417 municípios do Estado da Bahia com sede na cidade de Itaberaba, a qual passou a conduzir os experimentos a partir da década de 80 (NÓBREGA, 2008; CERQUEIRA, 2008).

A atuação da EBDA teve base no desenvolvimento da agropecuária baiana, com foco no fortalecimento da agricultura familiar, através da prestação de Assistência Técnica e Extensão Rural (ATER), Pesquisa e Classificação de Produtos de Origem Vegetal, fomento em agropecuária e agroindustrialização.

A empresa também teve como objetivo estimular e apoiar iniciativas de desenvolvimento rural sustentável, com atividades agrícolas e não agrícolas (LIMA et al., 2012).

Em 2011, foi lançada no mercado uma nova cultivar de mamoneira, a EBDA MPA 34, pela EBDA juntamente com a Secretaria da Agricultura (SEAGRI, 2011). Posteriormente, outras cultivares com características agronômicas bem adaptadas as condições edafoclimáticas de algumas regiões da Bahia foram lançadas. A linhagem de maior destaque foi a Sipeal 28 que apresenta porte médio, caule roxo, sem cera, frutos deiscentes, sementes de cor preta, floração do primeiro cacho com 55 dias, teor de óleo nas sementes de 47,3% e produtividade média em condições de sequeiro de 1300 kg/ha de bagas (BELTRÃO, 2006). Apesar dos resultados notórios que a EBDA prestou para o melhoramento genético da cultura da mamoneira, esta instituição teve suas atividades encerradas e a consequente extinção da Empresa Baiana de Desenvolvimento Agrícola S.A - EBDA, em liquidação por meio do DECRETO Nº 17.037 de Setembro de 2016.

Várias instituições de pesquisa contribuem com o melhoramento da mamona no Brasil, entre elas se destacam a Escola Superior de Agricultura Luiz de Queiroz – ESALQ, a Universidade Federal de Viçosa – UFV, a Empresa de Pesquisa Agropecuária de Minas Gerais – EPAMIG, a Empresa de Pesquisa Agropecuária do Ceará – EPACE e a Empresa Pernambucana de Pesquisa Agropecuária – IPA (LIMA et al., 2012), a Empresa Brasileira de Pesquisa e Agropecuária (EMBRAPA) em parceria com o Centro Nacional de Pesquisa de Algodão – CNPA (SEVERINO et al., 2012) e por fim, a Universidade Federal do Recôncavo da Bahia - UFRB, através do Núcleo de Melhoramento Genético e Biotecnologia - NBIO/UFRB (SILVA et al., 2017) que desenvolve trabalhos significativos e com reconhecimento pela comunidade científica nacional e internacional que objetivam o desenvolvimento de linhagens ou cultivares de mamoneira, subsidiados pela Petrobras Biocombustível e pela Agência Nacional do Petróleo e Gás natural e Biocombustíveis – ANP.

Desde 2005, o NBIO desenvolve trabalhos que vão desde o pré-melhoramento ao melhoramento genético em mamoneira com o objetivo de selecionar materiais promissores destinados a diversas aplicabilidades que o óleo apresenta, por meio do entendimento e utilização de ferramentas que descrevem de forma precisa o fenótipo, como atualmente, com o auxílio de ferramentas que permitem a identificação de moléculas de interesse com potencial biológico, desde

a unidade fundamental física e funcional da hereditariedade, os genes, até os seus produtos finais, como as enzimas e proteínas.

Atualmente, quatro novas linhagens com características agronômicas e tecnológicas superiores as comerciais, estão em processo de tramitação para serem lançadas como novas cultivares pelo Ministério da Agricultura, Pecuária e Abastecimento – MAPA, como resultado dos trabalhos desenvolvidos pelo Núcleo durante os 15 anos de pesquisa.

O NBIO também já desenvolveu 203 linhagens endogâmicas, obtidas por meio de cruzamentos dialélico parcial entre cinco cultivares parentais, EBDA MPA-17, Mirante 10, Sipeal 28, BRS 149 Nordestina e BRS 188 Paraguaçu desenvolvidas pela EBDA e Embrapa Algodão. Essas 203 linhagens constituem a base genética para os trabalhos científicos desenvolvidos por este grupo de pesquisa, conservados em BAG na forma *ex situ* e em câmara refrigerada a 8° C a 35% de umidade.

Para obtenção destas linhagens, foram desenvolvidos testes de hibridações (PASSOS et al., 2010), condução da população segregante por autofecundações para avanços até a geração F₆ (OLIVEIRA, 2011; SANTOS, 2013), confirmação da homozigose destas linhagens pela caracterização em nível molecular (MACHADO, 2011; MACHADO; SILVA, 2013), aplicação de descritores morfoagronômicos (OLIVEIRA, 2011; SANTOS, 2013; BRASILEIRO, 2014; SILVA et al., 2019), seguido de seleção de linhagens elites (SANTOS et al., 2019), quanto a resistência ao mofo cinzento (SOUSA, 2014) e comportamento germinativo e tolerância ao alumínio na presença de diferentes concentrações de ácido giberélico (ALMEIDA, 2014; PASSOS et al., 2015).

Acredita-se que com o surgimento de técnicas de menor custo, em relação as tradicionais, uma melhor manipulação dos recursos genéticos conservados internacionalmente na espécie disponíveis para comunidade acadêmica por meio de publicações científicas, garantam a ampliação da base genética da espécie, gerando subsídios para o melhoramento genético para diversas características de interesse, principalmente para resistência ao mofo-cinzento.

O patossistema da mamoneira x *Amphobotrys ricini*

Entre os patógenos causadores de doenças na cultura, há destaque para o fungo mitospórico *Amphobotrys ricini* (Buchw) Hennebert (sin. *Botrytis ricini* N. F.

Buchw.) (Teleomórfico *Botryotinia ricini* (G. H. Godfrey) Whetzel, agente etiológico da doença popularmente conhecida como mofo-cinzeno.

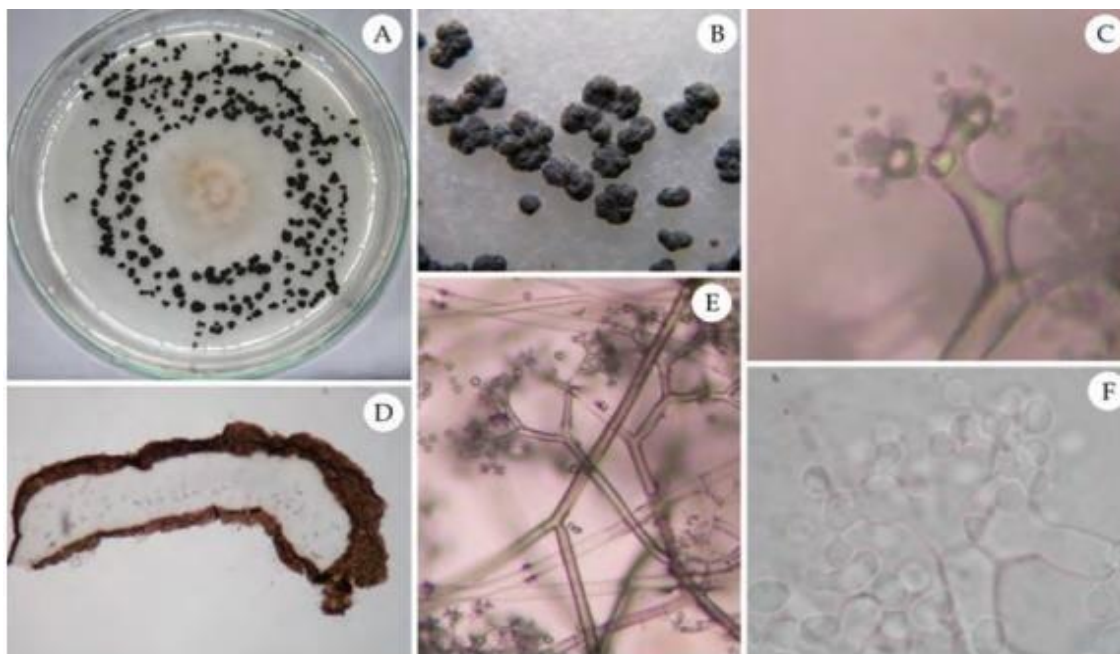
Diferentemente do que encontramos, normalmente, para nome de espécies, o mofo cinzeno foi descrito causado por diferentes anamorfos de *Botryotinia ricini*. A ambiguidade no anamorfo do patógeno do mofo cinzeno da mamona ainda existe como três anamorfos a saber, *Amphobotrys ricini*, *Botrytis ricini* e *Botrytis cinerea* estão associados a esta doença em literatura (BUCHWALD, 1949; HENNEBERT, 1973; SUNG-KEE et al., 2001). Entretanto, este patógeno pertence à divisão Ascomycota, ordem Helotiales e família Sclerotiniaceae.

Até 1923 (GODFREY, 1923), acreditava-se que este patógeno era exclusivo da cultura da mamona, até que em 1980 foi registrado sua presença em outros hospedeiros (SOARES et al., 2012), incluindo registros recentes em *Acalypha* spp. (YU et al., 2012; COUTINHO et al., 2014).

Segundo Lima et al. (2001), o patógeno encontra-se disseminado em todas as áreas de cultivo de mamoneira dentro do território brasileiro, ocorrendo agressivamente sob condições climáticas favoráveis com temperatura em torno de 25 °C e umidade relativa do ar acima de 75%. A doença já afetou a produção de *R. communis* no país, avançando da inflorescência para cachos e frutos, reduzindo assim a produção de óleo.

Seus esporos (Figura 2) podem ser disseminados tanto pelo vento quanto por sementes (NEERGAARD, 1979), sendo, em alguns casos, disseminados por insetos também, pois são atraídos pela grande exsudação que o néctar nas flores apresenta (KIMATI, 1980).

Figura 2. Imagem microscópica do *A. Ricini*: A) e B) Esclerócio do fungo em meio de cultura; C) Células conidiogênicas mostrando a conidiogênese sincrônica; D) Corte transversal através de um esclerócio evidenciando sua forma plano-convexa; E) Ramo dicotômico dos conidióforos; F) Célula conidiogênica evidenciando os dentículos e conídios globosos.



Autor: D. J. Soares

Inicialmente, os sintomas são observados nas inflorescências e nos cachos de frutos como manchas acinzentadas ou azuladas. Nos frutos evoluem para manchas circulares ou elípticas, afundadas e de cor escura que podem resultar em ruptura ou apodrecimento da cápsula e dos grãos. As lesões aumentam e tornam-se uma praga marrom-acinzentada, com abundante massa de conídios sobre os tecidos das plantas e após a formação de esporos surge uma extensa área lesionada com um aspecto pulverulento cinza sob condições úmidas favoráveis ao desenvolvimento da doença (DANGE et al., 2005) (Figura 3).

Segundo Sussel (2009), o fator que contribui para que a umidade permaneça favorável, isto é, elevada são as flores da mamoneira que proporcionam uma grande quantidade de estames, próximos um do outro.

Oliveira Neto (2017) descreve que a presença de nectários nas inflorescências favorece a proliferação do patógeno, ao disponibilizarem açúcares que podem ser aproveitados como alimento pelo patógeno, a depender da fase de desenvolvimento dos racemos e das condições climáticas as perdas de

produção podem chegar a 100% (SUSSEL, 2009; SOARES, 2012).

Figura 3. Sintomas do mofo cinzento na inflorescência e racemo da mamona:
A e B) Sintomas em inflorescência jovens, antes da fertilização de flores femininas;
C a H) Sintomas em cápsulas em estágios distintos de desenvolvimento.



Autor: D. J. Soares

Apesar da importância econômica que este patógeno apresenta a cultura da mamona, ainda não há na literatura uma forma de controle eficiente da doença. Prazeres et al. (2017), testaram Azoxistrobina, Procimidone, Iprodiona e Captana, nas cultivares Mirante 10 e Sipeal 28 de mamoneira. Seus resultados mostraram que estes fungicidas apresentaram alta fungitoxidade e em condições de telado, reduziram a severidade do *A. ricini* e promoveram acréscimo nos caracteres diâmetro de caule e altura de plantas para os genótipos avaliados, porém, em campo não há dados de erradicação do fungo.

Entretanto, são necessários novos estudos que visem entender melhor tanto o mecanismo de atuação do patógeno quanto dos componentes que fazem parte do sistema de defesa *R. communis* L. ao patógeno *A. ricini* e por meio disso, sugerir novos métodos de controle a doença.

Resultados satisfatórios podem ser obtidos por meio da aplicação de

diferentes abordagens biotecnológicas em genótipos de *R. communis*, auxiliados pelo uso de ferramentas de bioinformática, que apresentam potencial para contribuir nos estudos que visam desenvolver genótipos resistentes ao *A. ricini*.

Sistema Imune e Proteínas Relacionadas a Patogêneses – PRs

As plantas são constantemente expostas a ataques de fitopatógenos que comprometem a sua sobrevivência, levando a uma redução significativa no rendimento anual de algumas culturas, representando uma séria ameaça à segurança alimentar (CRAMER et al., 2011; BOCCARDO et al., 2019).

Ao longo do processo evolutivo, as plantas precisaram desenvolver a capacidade de responder adequadamente a diferentes tipos de estresses, assim criaram mecanismos de reconhecimento e resposta que requerem a diferenciação de fitopatógenos e sua contenção nos tecidos infectados (URBACH et al., 2017; BOCCARDO et al., 2019).

Segundo Jones & Dangl (2006), os fitopatógenos usam diversas estratégias de ataque. As bactérias proliferam por meio dos espaços intercelulares (os apoplastos) após a entrada nos poros de gás ou água através dos estômatos e hidatódios, respectivamente; ou obtêm acesso por meio de ferimentos. Já os nematoides e pulgões se alimentam inserindo um estilete diretamente nas células vegetais.

Os fungos, por sua vez, desenvolveram a capacidade de identificar e responder com precisão aos organismos em interação, porque eles se interatuam com todos os principais grupos orgânicos: vírus, bactérias, protistas, amebas, animais e plantas (FREY-KLETT et al., 2011).

Após infectar a planta, o fungo pode crescer de maneira biotrófica, obtendo nutrientes da hospedeira sem provocar a morte das células. Nesse caso, alguns fungos produzem haustórios, estruturas especializadas na obtenção de nutrientes (vitaminas, aminoácidos, polissacarídeos), a partir das células hospedeiras (KNOGGE 1996; KAHMANN; BASSE, 2001; FERREIRA; MONTEIRO, 2006).

Contudo, em meio ao ataque pelo agressor, as plantas utilizam de diversos mecanismos presentes em seu sistema imune inato, de cada célula, e de sinais sistêmicos emanados dos locais de infecção para reconhecer e impedir a proliferação do fitopatógeno (JONES; DANGL 2006).

Neste momento, os elicitores fúngicos são reconhecidos imediatamente pela planta e nesta ocasião, os Padrões Moleculares Associados a Micróbios ou a

Patógenos, MAMPs e PAMPs (do inglês *Microbe-Associated Molecular Pattern and Pathogen-Associated Molecular Pattern*) são reconhecidos pelos Receptores de Reconhecimento de Padrões – PRRs (do inglês *Pattern Recognition Receptors*) que desencadeiam nas plantas respostas imunológicas que levam a ativação da via de defesa, conhecida como Imunidade Desencadeada por Padrões – PTI (do inglês *Pattern-Triggered Immunity*). Quando ausente, é ativada a Imunidade Ativada por Efetores – ETI (do inglês *Effector-triggered Immunity*) (JONES; DANGL 2006).

Após este reconhecimento, uma série de respostas imunológicas nas plantas são desencadeadas, como a produção de Espécies Reativas de Oxigênio - EROs, que atuam como agentes oxidantes danificando estruturas celulares, como lipídios membranares até o material genético dos oportunistas, acúmulo de ácido salicílico - AS, ácido jasmônico - AJ, Etileno – ET, auxinas e giberelinas, bem como, das fitoalexinas e das Proteínas Relacionadas à Patogênese (Proteínas-PR) (PARK et al., 2004; VAN LOON et al., 2006; KAUR et al., 2017; BOCCARDO et al., 2019).

As proteínas-PR foram identificadas pela primeira vez em folhas de uma cultivar de tabaco (*Nicotiana tabacum* L.) que produzia Resposta Hipersensível – RH à inoculação com o vírus do mosaico do tabaco (*Tabacco Mosaic Virus* - TMV), no início da necrose e aumentaram em quantidade com o passar do tempo. Após cerca de setes dias, elas se tornaram detectáveis em órgãos não inoculados da planta, onde o vírus estava quase ausente e a Resistência Adquirida Sistêmica - RAS foi provocada após uma nova inoculação (VAN LOON; VAN KAMMEN, 1970; GIANINAZZI et al., 1970).

Em 1990, o uso do termo proteínas-PR foi especificado para indicar proteínas induzidas por vários tipos de patógenos, bem como condições de estresse, como as provocadas por patógenos e aquelas induzidas pela aplicação de produtos químicos que imitam o efeito da infecção por patógenos ou induzem tensões semelhantes (VAN LOON et al., 1990).

Já o termo proteínas-PR *Like* foi proposto para acomodar proteínas e enzimas homólogas as PRs, deduzido da sequência de aminoácidos ou previsto pela sequência nucleotídica do cDNA correspondente ou gene, mas que são induzidos principalmente em um navegador específico do tecido (VAN LOON et al., 1994; VAN LOON, 1999).

Atualmente, as proteínas-PR foram agrupadas em 19 famílias (PR-1 a PR-

19 ARORA, 2020), onde cada família é composta por proteínas que compartilham sequências de aminoácidos, relações sorológicas e atividade enzimática ou biológica semelhantes, como glucanases (PR-2; TONON et al., 2002), osmotinas e taumatinas (PR-5; LIU J-J et al., 2010), inibidores de protease (PR-6; LALUK; MENGISTE, 2011), lisozimas (PR-8; HEITZ et al., 1994), peroxidase (PR-9; GHOSH, 2006), ribonucleases (PR10; AGARWAL et al., 2013) e quitinases (PR-3, PR-4, PR-8, PR-11; BERTINI et al., 2009). Podendo ainda apresentar atividade defensina (PR-12), tionina (PR-13), de transferência lipídica (PR-14) e de oxalato oxidase (PR-15; PR-16) (VAN LOON et al., 2006).

Das 19 famílias das PRs, duas delas (PR-1 e PR-19) ainda não tiveram suas funções totalmente elucidadas (VAN LOON et al., 2006; ARORA 2020). Além disso, a família das PR-1 é a mais abundante dentre as demais famílias nas plantas, atingindo entre 1 e 2% do total de proteínas foliares (EDREVA, 2005).

Ainda segundo Sels et al. (2008), as Proteínas-PR estão presentes em baixa concentração nas plantas e são induzidas em maiores quantidades em situações patológicas, não necessariamente implicadas em uma interação direta com fitopatógeno. Algumas destas proteínas acumulam-se localmente em resposta a infecções, levando a uma Resposta Hipersensível - RH, enquanto outras são induzidas sistemicamente, envolvidas na Resistência Adquirida Sistêmica - RAS que confere imunidade subsequente a um amplo espectro de patógenos, promovendo a formação de barreiras químicas (KINKEMA et al., 2000).

As proteínas-PR são consideradas os genes de assinatura dessas duas vias em muitas plantas e particularmente, a expressão aumentada dos genes precursores das PR-1 e PR-5 representam a ativação da via de sinalização do AS, por isso são consideradas como dependentes desse fitormônio. Em contraste, o aumento da expressão dos genes precursores das PR-3, PR-4 e PR-12 representam a ativação da via JA em *Arabidopsis thaliana* (BARI; JONE, 2009; ALI et al., 2018).

De forma geral, a ativação dessas vias de sinalização leva ao acúmulo das proteínas-PR, que por vez, minimizam a carga de fitopatógenos ou o aparecimento de doenças nos órgãos vegetais não infectados (ALI et al., 2018).

Quanto as propriedades bioquímicas específicas das proteínas-PR, como o ponto isoelétrico – pI, peso molecular e grand average of hydropathy - GRAVY classificam estas proteínas e permite predizer suas funções relacionadas a infecção nas plantas (BENHAMOU et al., 1991; VAN LOON, 1999; JEUN, 2000;

JEUN; BUCHENAUER, 2001).

As proteínas-PR são estabelecidas em todos os órgãos das plantas como folhas, caules, raízes e flores (VAN LOON, 1999), sendo particularmente abundantes nas folhas, onde podem atingir aproximadamente 5 a 10% do total das proteínas foliares, associadas também ao desenvolvimento e crescimento das plantas em situações não-patológicas. Um exemplo claro são as grandes concentrações das PR-1 encontradas nas raízes e folhas de plantas jovens que são expressas constitutivamente em algumas espécies, na ausência de ataques por fitopatógenos (DAFOE et al., 2010).

As proteínas-PR foram associadas ao desenvolvimento dos frutos (MCCOLLUM et al., 2006), identificadas nas secreções ovulares ou durante a senescência foliar (CAPELLI et al., 1997; O'LEARY et al., 2007). Além de serem descritas como proteínas associadas ao amadurecimento que se acumulam fortemente nos frutos durante o processo de maturação (TATTERSALL et al., 1997; KIM et al., 2002).

Ademais, as proteínas-PR também são ativadas mediante a estresses de origem abiótica nos vegetais, pois esse tipo de estresse resulta em alterações nos diferentes fitormônios que levam ao acúmulo e ativação das proteínas-PR. No entanto, a expressão mediada pelo estresse abiótico nesta classe de proteínas ainda não foi totalmente compreendida em nível molecular (ALI et al., 2017).

Entretanto, diversos trabalhos mostram o envolvimento das proteínas-PR regulando positivamente o aumento a tolerância sob condições de estresses salino e hídrico, como foi observado na cultura do arroz (*Oryza sativa* L.) (WU et al., 2016), outros mostram as PR-2 e PR-3 exercendo atividade anticongelante nas células, devido ao estresse causado por baixas temperaturas (JANSKA et al., 2010) e em resposta a estresse causado por ferida, que levaram ao acúmulo das PR-4 (KIM et al., 2014).

A quantidade de dados coletados sobre as proteínas-PR ampliou o conhecimento sobre a sua relevância nos desempenhos importantes da planta, como resistência a doenças, adaptação geral a ambientes estressantes e relacionadas ao crescimento/desenvolvimento (EDREVA, 2005).

Sobretudo, com a crescente demanda de controle, principalmente de doenças fúngicas em vegetais, diversas abordagens biotecnológicas e de bioinformática tem surgido, a fim de permitir uma melhor compreensão dos mecanismos que envolvem o sistema imune das plantas e o ataque de

fitopatógenos.

Em *Ricinus communis*, as proteínas-PR- já foram identificadas no genoma e depositadas em diferentes bancos públicos de dados genômicos, por meio disso, torna-se possível desenvolver estudos sobre os padrões de expressão em diferentes tecidos da planta, além de estudos funcionais dos seus genes e proteínas, que podem contribuir de forma significativa na elaboração de novas estratégias de manipulação desta classe de proteínas, e como resultado, ter uma melhor compreensão do seu papel no sistema imune da referida espécie.

Ademais, estes resultados podem gerar subsídios para até sugerir o envolvimento das proteínas-PR com outras moléculas também relacionadas ao processo de resposta ao reconhecimento de infecções causadas pelo *A. ricini*, causador do mofo-cinza em *R. communis*.

Em um futuro próximo, acredita-se que este conhecimento agregado a aplicação de diferentes abordagens biotecnológicas para controle do *A. ricini*, seja possível desenvolver nossas cultivares resistentes a este patógeno agressor que tem dizimado lavouras de plantações de mamoneira no país.

Caracterização Gênica e Modelagem de Proteínas por Homologia

Após o sequenciamento do genoma humano e a revolução que isso causou no campo da genética e da biologia molecular, permitindo conhecer desde os genes até as proteínas que eles codificam, incluindo seu funcionamento no organismo, desafios têm sido propostos, a fim de estudar também em nível molecular as espécies vegetais, principalmente as de importância agrícola e econômica.

Estas informações auxiliam na erradicação de pragas e doenças e permite reunir características de importância econômica em novos genótipos, levando a estabilidades dos vegetais para uso e consumo humano.

Com o aumento do sequenciamento genômico de organismos, pesquisadores tem acesso a milhões de sequências gênicas e proteicas em um ritmo sem precedentes.

Nesse contexto, sucessivos trabalhos de descrição, caracterização e análise de família gênicas envolvidas nos principais processos biológicos dos organismos tem se tornado cada vez maior, juntamente com a resolução da estrutura tridimensional – 3D de moléculas alvo, pois entende-se que este

conhecimento abrangente, tanto da variação gênica quanto da caracterização estrutural de proteínas, agregam informações valiosas que melhoram as estratégias de manipulação dessas moléculas, visto que constituem uma forma precisa, para descrever a estrutura, aspectos físicos e fundamentais de forma sistemática em cada molécula, sejam os próprios genes até o produto final pelos genes codificados, as proteínas.

Trabalhos de caracterização gênica foram desenvolvidos em diversas culturas, como no milho (*Zeamays*) (ZHU et al., 2019); na mandioca (*Manihot esculenta* Crantz) (ZOU; ZHANG 2019) e no tomate (*Solanum lycopersicum*) (AMAGAYA et al., 2020).

Em *R. communis*, a caracterização de famílias gênicas tem sido crescente nos últimos anos, como observa-se pela caracterização das famílias gênicas de superóxido dismutase em sementes (GOMES NETO et al., 2018); *one finger* – Dof, relacionados a fatores de transcrição (ZOU; ZHANG, 2019); *ARF*, *GH3* e *Aux/IAA*, relacionadas às auxinas (FENG et al., 2019); como também de genes relacionados à autofagia – ATG (HAN et al., 2020). Todas estas famílias gênicas estudadas em mamoneira estão relacionados ao crescimento, desenvolvimento e a estresses de origem biótica.

Em estudos de caracterização de famílias gênicas, os trabalhos de resolução de estrutura de proteínas por homologia têm ganhado importância nos últimos anos. Como observa-se na cultura do algodão (*Gossypium* spp.) (SHAHEEN et al., 2018; SARWAR et al., 2019); na mandioca (*Manihot esculenta* Crantz) (MARQUES et al., 2019) e tomate (*Solanum lycopersicum*) (TERRET-WELTER et al., 2020).

Em *R. communis* L., trabalhos significativos resolveram a estrutura de importantes proteínas por homologia, como as proteínas 2S albumina, Ric c1 e Ric c3, que exercem papel importante na resistência a insetos (NASCIMENTO et al., 2011); identificação de atividade antibacteriana de fitoquímicos de *R. communis* contra beta-lactamase de *Enterococcus faecalis* e *Staphylococcus aureus* através de estudos de modelagem por homologia e acoplamento molecular (UMAARASU et al., 2018) e resolução da estrutura tridimensional por homologia de enzimas envolvidas na biossíntese do óleo da mamoneira (YOUSAFI et al., 2019).

A magnitude de informações genéticas disponíveis hoje, têm propostos desafios, a fim de que toda informação seja descrita de forma confiável e utilizando

métodos alternativos com baixo custo, em relação a técnicas anteriores, porém reproduzíveis e viáveis, principalmente para o melhoramento genético de espécies em países em desenvolvimento, onde as instituições de pesquisa contam com recursos financeiros limitados.

A função de uma proteína quase sempre envolve movimentos e mudanças conformacionais, e um entendimento molecular de seu mecanismo requer uma descrição detalhada dos diferentes estados funcionais que a estrutura pode explorar dinamicamente.

Exemplos típicos disto incluem alterações conformacionais alostéricas após eventos de ligação, estados excitados intermediários em ciclos de reação, fenômenos de transporte e movimento (SCHWEDE, 2013). No entanto, a modelagem por homologia desempenha um papel importante na exploração dessas conformações alternativas e na descrição da dinâmica das transições (SCHWEDE, 2013).

Contudo, a resolução da estrutura tridimensional - 3D de moléculas alvo de proteínas, ainda é uma caracterização estrutural limitada pelo custo, tempo e desafios experimentais inerentes à determinação estrutural por cristalografia de raios-X e espectroscopia de ressonância magnética nuclear - RMN. Na ausência de estruturas determinadas experimentalmente, os modelos de estrutura de proteínas derivados computacionalmente tornam-se valiosos, pois podem gerar hipóteses testáveis para a interpretação espacial da dados (BAKER; SALI 2001).

Dessa forma, a modelagem comparativa de estruturas proteicas tem sido utilizada para produzir modelos de estrutura válidos para pelo menos um domínio, uma região com função bioquímica determinada, em mais da metade de todas as sequências conhecidas (WEBB; SALI 2016).

É notório o aumento desses estudos nas principais cultura de valor agrônômico e importância econômica, como observa-se na cultura do Algodão (*Gossypium* spp.), onde recentemente, Wang et al. (2019), identificam e analisam sistematicamente os genes catalases – CAT, que codificam proteínas metabolizadoras de espécies reativas de oxigênio – ERO, que desempenham papel crucial em uma variedade de respostas ao estresse biótico.

As informações geradas por estas abordagens de bioinformática, permitiu determinar as funções destes genes em escala genômica e, assim, compreender sobre o papel e sistema regulador molecular no metabolismo de ERO na espécie. Além disso, gerou subsídios para o avanço no estudo sobre os padrões de

expressão desta família gênica em diferentes tecidos na cultura do algodão e assim, investigar melhor as funções biológicas dessa família gênica e até sugerir o envolvimento dos genes CAT na resistência ao *Verticillium wilt* - VW, já que a produção de ERO ocorre de forma rápida como um dos processos em resposta ao reconhecimento de infecção por patógenos.

Atualmente, numerosos projetos genomas estão em andamento, sequenciando dezenas de milhares de nucleotídeos e disponibilizando tais informações em bancos de dados públicos para toda comunidade científica, pois nesses bancos de dados, como o *GenBank*-NCBI, são depositadas além de sequências de DNA, sequências de aminoácidos/proteínas, que são codificadas por genes e a partir disso é possível explorar a funcionalidade dos genes que compõe o genoma de muitas espécies-alvo (NORVELL; MACHALEK, 2000; HEINEMANN, 2000).

O *National Institute of General Medical Sciences* – NIGMS, anunciou a seu “*Protein Structure Initiative*”, que é projetado para organizar um grande esforço cooperativo na genômica estrutural, que provém dos resultados dos projetos de sequenciamento e das recentes determinações de estrutura proteicas.

Através desse projeto, o NIGMS procura desenvolver um recurso público para organizar e analisar estruturas proteicas e suas famílias. Este recurso ligará informações de sequências, estruturas e sua informação funcional e permitirá a previsão de estruturas desconhecidas por meio da modelagem de homologia que se baseia na identificação de genes homólogos cujas funções já são conhecidas (NORVELL; MACHALEK, 2000).

Além da magnitude dos estudos de natureza molecular e do grau de informações que estas abordagens no campo da bioinformática fornecem, ainda se faz necessário analisar o fenótipo resultante da expressão gênica, pois continua sendo uma atividade essencial para identificação e diferenciação de indivíduos. A análise do fenótipo permite estimar a diversidade, bem como, o desempenho agrônomo para orientar as estratégias de identificação de combinações parentais, que ao serem cruzados, possibilitam maior efeito heterótico na progênie e introgressão de genes desejáveis.

REFERÊNCIAS

- ACHAYA, K. T. Chemical Derivatives of Castor Oil. **Journal of the American Oil Chemists Society**, v. 48, p. 758-763, 1971.
- AGARWAL, P.; BHATT, V.; SINGH, R. M.; SOPORY, S.; CHIKARA, J. Pathogenesis-related gene, JcPR-10a from *Jatropha curcas* exhibit RNase and antifungal activity. **Molecular Biotechnology**, v. 2.; p. 412–425, 2013.
- AGYENIM-BOATENG, K. G.; LU, J.; SHI, Y.; ZHANG, D.; YIN, X. SRAP analysis of the genetic diversity of wild castor (*Ricinus communis* L.) in South China. **PLoS One**, v. 7, p. 1-16, 2019.
- ALLAN, G.; WILLIAMS, A.; RABINOWICZ, P. D.; CHAN, A. P.; RAVEL, J.; KEIM, P. Worldwide genotyping of castor bean germplasm (*Ricinus communis* L.) using AFLPs and SSRs. **Genetic Resources and Crop Evolution**, v. 55, p. 365-378, 2008.
- ALI, S.; MIR, Z. A.; BHAT, J. A.; CHANDRASHEKAR, P. K.; PAPOLU, S.; RAWAT, A. G. Identification and comparative analysis of *Brassica juncea* pathogenesis-related genes in response to hormonal, biotic and abiotic stresses. **Acta Physiologiae Plantarum**, v. 39, p.1-15, 2017.
- ALI, A.; KHAN, I. U.; JAN, M.; KHAN, H. A.; HUSSAIN, S.; NISAR, M.; CHUNG, W. S.; YUN, D-J. The High-Affinity Potassium Transporter EpHKT1;2 From the Extremophile *Eutrema parvula* Mediates Salt Tolerance. **Frontiers in Plant Science**, v. 9, p. 1-11, 2018.
- AULD, D. L.; ZANOTTO, M. D.; MCKEON, T.; MORRIS, J. B. Castor. **Oil Crops**, v. 4, p. 317-332, 2009.
- ALMEIDA, V. O. **Comportamento germinativo, armazenamento de sementes e tolerância ao alumínio da mamoneira**. Tese (Doutorado). Universidade Federal do Recôncavo da Bahia – UFRB, Cruz das Almas - BA, 79f, 2014.
- AMAGAYA, K.; SHIBUYA, T.; NISHIYAMA, M.; KATO, K.; KANAYAMA, Y. Characterization and Expression Analysis of the Ca²⁺/Cation Antiporter Gene Family in Tomatoes. **Plants**, v. 1, p. 1-16, 2020.
- ANDRADE, T. A.; ERRICO, M.; CHRISTENSEN, K. V. Influence of the reaction conditions on the enzyme catalyzed transesterification of castor oil: A possible step in biodiesel production. **Bioresource Technology**, v. 243, p. 366-374, 2017.
- ANJANI, K. Castor genetic resources: A primary gene pool for exploitation. **Industrial Crops and Products**, v. 35, p. 1-14, 2012.
- ANJANI, K. Extra-early maturing germoplasm for utilization in castor improvement. **Industrial Crop & Products**, v. 31, 139-144, 2010.
- ARORA, R.; KUMAR, A.; SINGH, I. K.; SINGH, A. Pathogenesis Related Proteins: a defensin for plants but an allergen for humans. **International Journal of**

Biological Macromolecules, v. 157, p. 659-672, 2020.

BABITA, M.; MAHESWARI, M.; RAO, L. M.; SHANKER, A. K.; RAO, D. G. Osmotic adjustment, drought tolerance and yield in castor (*Ricinus communis* L.) hybrids. **Environmental and Experimental Botany**, v. 69, p. 243–249, 2010.

BAKER, D.; SALI, A. Protein Structure Prediction and Structural Genomics. **Science**, v. 5540, p. 93-96, 2001.

BARI, R.; JONES, J. D. G. Role of plant hormones in plant defence responses. **Plant Molecular Biology**, v. 69, p. 473-488, 2009.

BELTRÃO, N. E. M. **A cadeia da Mamona no Brasil. Ênfase para o Segmento P & D: Estado da Arte, Demandas de Pesquisa e Ações Necessárias para o Desenvolvimento**. Campina Grande: Embrapa Algodão, 2006 (Documentos).

BENHAMOU, N.; GRENIER, J.; ASSELIN, A. Immunogold localization of pathogenesis-related protein P14 in tomato root cells infected by *Fusarium oxysporum* f. sp. *radicislycopersici*. **Physiological and Molecular Plant Pathology**, v. 38, p. 237-253, 1991.

BERTINI, L.; CAPORALE, C.; TESTA, M.; PROIETTI, S.; CARUSO, C. Structural basis of the antifungal activity of wheat PR4 proteins. **FEBS Letters**, v. 17, p. 2865–2871, 2009.

BOCCARDO, N. A.; SEGRETIN, M. E.; HERNANDEZ, I.; MIRKIN, F. G.; CHACÓN, O.; LOPEZ, Y.; BORRAS-HIDALGO, O.; BRAVO-ALMONACID, F. F. Expression of Pathogen-Related Proteins in transplastomic tobacco plants confers resistance to filamentous pathogens under field trials. **Nature Scientific Reports**, v. 9, n. 2791, p. 1-13, 2019.

BOZZA, W. P.; TOLLESON, W. H.; ROSADO, L. A. R.; ZHANG, B. Ricin detection: tracking active toxin. **Biotechnology Advances**, v. 33, p. 117–123, 2015.

BRASILEIRO, H. S. **Seleção de linhagens elites de mamoneira por meio de marcadores microssatélites e caracteres teor de óleo e peso de sementes por planta**. Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia - UFRB, Cruz das Almas – BA, 76f, 2014.

BREDESON, J. V.; LYONS, J. B.; PROCHNIK, S. E.; WU, G. A.; HA, C. M.; EDSINGER-GONZALES, E.; GRIMWOOD, J.; SCHMUTZ, J.; RABBI, I. Y.; EGESI, C.; NAULUVULA, P.; LEBOT, V.; NDUNGURU, J.; MKAMILO, G.; BART, R. S.; SETTER, T. L.; GLEADOW, R. M.; KULAKOW, P.; FERGUSON, M. E.; ROUNSLEY, S.; ROKHSAR, D. S. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. **Nature**, v. 34, p. 562-571, 2016.

CANVIN, D. T. Formation of oil in the seed of *Ricinus communis* L. Ln2. **Canadian Journal of Biochemistry and Physiology**, v. 41, p. 1879-1885, 1963.

CAO, Y., ZHENG, X.; BAI, J.; G, G. Distribution of the main crop germplasm resources in China. **Chinese Geographical Science**, v. 7, p. 310-316, 1997.

CAPELLI, N.; DIOGON, T.; GREPPIN, H.; SIMON, P. Isolation and characterization of a cDNA clone encoding an osmotin-like protein from *Arabidopsis thaliana*. **Gene**, v. 1, p. 51-56, 1997.

CARRINO, LINDA.; VISCONTI, D.; FIORENTINO, N.; FAGNANO, M. Biofuel Production with Castor Bean: A Win–Win Strategy for Marginal Land. **Agronomy**, v. 10, p. 1-22, 2020.

CHAGAS, H. A.; BASSETO, M. A.; ROSA, D. D.; ZANOTTO, M. D.; FURTADO, E. L. Diagrammatic scale to assess gray mold (*Amphobotrys ricini*) in castor bean (*Ricinus communis* L.). **Summa phytopathol**, v.36, p. 164-167, 2010.

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.; GEDIL, M.; STANKE, M.; HAAS, B. J.; WORTMAN, J. R.; FRASER-LIGGETT, C. M.; RAVEL, J.; RABINOWICZ, P. D. Draft genome sequence of the ricin-producing oilseed castor bean. **Nature Biotechnology**, v. 28, p. 951-956, 2010.

CHAUDHARI, B. A.; PATEL, M. P.; DHARAJIVA, D. T.; TIWARI, K. K. Assessment of Genetic Diversity in Castor (*Ricinus communis* L.) using Microsatellite Markers. **BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA**, v. 16, p. 61-69, 2019.

CHEN, S.; NELSON, M. N.; GHAMKHAR, K.; FU, T. Divergent patterns of allelic diversity from similar origins: the case of oilseed rape (*Brassica napus* L.) in China and Australia. **Genome**, v. 51, p. 1-10, 2008.

COUTINHO, F. F.; MACEDO, D. M.; BARRETO, R. W. First report of gray mold (*Amphobotrys ricini*) on copperleaf (*Acalypha wilkesiana*) in Brazil. **Plant Disease**, v. 98, p. 276, 2014.

CRAMER, G. R.; URANO, K.; DELROT, S.; PEZZOTTI, M.; SHINOZAKI, K. Effects of abiotic stress on plants: a systems biology perspective. **BMC Plant Biology**, v. 11, n. 163, p. 11-163, 2011.

DAFOE, N. J.; GOWEN, B. E.; CONSTABEL, C. P. Thaumatin-like proteins are differentially expressed and localized in phloem tissues of hybrid poplar. **BioMed Central Plant Biology**, v. 10, p. 1-11, 2010.

EDREVA, A. Pathogenesis-Related Proteinss: Research progress in the last 15 years. **General and Applied Plant Physiology**, v. 31, p. 105-124, 2005.

ELIMAM, A.M.; ELMALIK, K. H.; ALI, F. S. Larvicidal, adult emergence inhibition and oviposition deterrent effects of foliage extract from *Ricinus communis* L. against *Anopheles arabiensis* and *Culex quinquefasciatus* in Sudan. **Tropical Biomedicine**, v. 26, p. 130–139, 2009.

FALASCA, S.L.; ULBERICH, A.C.; ULBERICH, E. Developing an agro-climatic zoning model to determine potential production areas for castor bean (*Ricinus communis* L.). **Industrial Crops and Products**, v. 40, p. 185–191, 2012.

FENG, L.; LI, G.; HE, Z.; HAN, W.; SUN, J.; HUANG, F.; DI, J.; CHEN, Y. The ARF, GH3, and Aux/IAA gene families in castor bean (*Ricinus communis* L.):

Genome-wide identification and expression profiles in high-stalk and dwarf strains. **Industrial Crops and Products**, v. 141, p. 1-14, 2019.

FERNANDES, K. V.; DEUS-DE-OLIVEIRA, N.; GODOY, M. G.; GUIMARÃES, Z. A. S.; NASCIMENTO, V. V.; MELO, E. J. T.; FREIRE, D. M. G.; DANSA-PETRETSKI, M. MACHADO, O. L. T. Simultaneous allergen inactivation and detoxification of castor bean cake by treatment with calcium compounds. **Brazilian Journal of Medical and Biological Research**, v. 45, p. 1002–1010, 2012.

FERRAZ, R. L. S.; MAGALHÃES, I. D.; BELTRÃO, N. E. M.; MELO, A. S.; BRITO NETO, J. F.; ROCHA, M. S. Photosynthetic pigments, cell extrusion and relative leaf water content of the castor bean under silicon and salinity. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 19, p. 841-848, 2015.

FERREIRA, R. B.; MONTEIRO, S. Fungal pathogens: the battle for plant infection. **Critical Reviews in Plant Sciences**, v. 25, p. 505–524, 2006.

FOSTER, J. T.; ALLAN, G. J.; CHAN, A. P.; RABINOWICZ, P. D.; RAVEL, J.; JACKSON, P. J.; KEIM, P. Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). **BioMedCentral Plant Biology**, v. 10, p. 1-11, 2010.

FRACETTO, F. J. C.; FRACETTO, G. G. M.; FEIGL, B. J.; CERRI, C. C.; NETO, M. S. Greenhouse gas from castor bean production and their subproducts. **Revista Caatinga**, v. 28, p. 90-98, 2015.

FREY-KLETT, P.; BURLINSON, P.; DEVEAU, A.; BARRET, M.; TARKKA, M.; SARNIGUET, A. Bacterial-fungal interactions: hyphens between agricultural, clinical, environmental, and food microbiologists. **Microbiology and Molecular Biology Reviews**, v. 75, p. 583–609, 2011.

GHOSH, M. Antifungal properties of haem peroxidase from *Acorus calamus*. **Annals of Botany**, v. 6, p. 1145-1153, 2006.

GIANINAZZI, S.; MARTIN, c.; VALÉE, J.-C. Hypersensibilité aux virus, température et protéines solubles chez le *Nicotiana Xanthi* n.c. Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. **Comptes rendus de l'Académie des Sciences**, v. 270, p. 2383-2386, 1970.

GODFREY, G. H. Gray mold of Castor bean. **Journal Agricultural Research**, v. 23, p. 679-715, 1923.

GOMES NETO, V.; RIBEIRO, P. R.; DEL-BEM, L. E.; BERNAL, D. T.; CUNHA LIMA, S. T.; LIGTERINK, W.; FERNANDEZ, L. G.; CASTRO, R. D. Characterization of the superoxide dismutase gene family in seeds of two *Ricinus communis* L. genotypes submitted to germination under water restriction conditions. **Environmental and Experimental Botany**, v. 155, p. 453-463, 2018.

HAN, B.; XU, H.; FENG, Y.; XU, W.; CUI, Q.; LIU, A. Genomic Characterization and Expressional Profiles of Autophagy-Related Genes (ATGs) in Oilseed Crop Castor Bean (*Ricinus communis* L.). **International Journal of molecular Sciences**, v. 21, p. 1-15, 2020.

HE, X.; MCMAHON, S.; HENDERSON, T. D. II GRIFFEY, S. M. & CHENG, L. W. Ricin toxicokinetics and its sensitive detection in mouse sera or feces using immuno-PCR. **Plos One**, v. 5, p. 1-8, 2010.

HEINEMANN, U. Structural genomics in Europe: Slow start, strong finish? **Nature Structural Biology**, v. 7, p. 940-942, 2000.

HEITZ, T.; SEGOND, S.; KAUFFMANN, S.; GEOFFROY, P.; PRASAD, V.; BRUNNER, F.; FRITIG, B.; LEGRAND, M. Molecular characterization of a novel tobacco pathogenesis-related (PR) protein: a new plant chitinase/lysozyme. **Molecular Genetics and Genomics**, v. 2, p. 246–254, 1994.

JANSKA, A.; MARIK, P.; ZELENKOVA, S.; OVESNA, J. Cold stress and acclimation: what is important for metabolic adjustment?. **Plant Biology**, v. 12, p. 395-405, 2010.

JEUN, Y.-CH. Immunolocalization of PR-protein P14 in leaves of tomato plant exhibiting systemic acquired resistance against *Phytophthora infestans* induced by pretreatment with 3-aminobutyric acid and preinoculation with Tobacco necrosis virus. **Journal of Plant Diseases and Protection**, v. 107, p. 352-367, 2000.

JEUN, Y.-CH., H. BUCHENAUER. Infection structures and localization of the pathogenesis-related protein AP24 in leaves of tomato plant exhibiting systemic acquired resistance against *Phytophthora infestans* after pre-treatment with 3-aminobutyric acid or Tobacco necrosis virus. **Journal of Phytopathology**, v. 149, p. 141-153, 2001.

JONES, J. D.; DANGL, J. L. The plant immune system. **Nature**, v. 444, n. 16, p. 323–329, 2006.

KAHMANN, R.; BASSE, C. Fungal gene expression during pathogenesis-related development and host plant colonization. **Current Opinion in Microbiology**, v. 4, p. 374-380, 2001.

KAUR, A.; PATI, P. K.; PATI, A. M.; NAGPAL, A. K. In-silico analysis of cis-acting regulatory elements of Proteínas Relacionadas a Patogênese of *Arabidopsis thaliana* and *Oryza sativa*. **Plos One**, v. 9, n. 12, p. 1-24, 2017.

KIM, Y. J.; LEE, M. G.; JANG, W. S.; KWON, S. Y. K.; YANG, D. C. Cloning and characterization of pathogenesis related-protein 4 gene from *Panax ginseng* Russ. **Journal of Plant Physiology**, v. 61, p. 664-671, 2014.

KIM, Y. S.; PARK, J. Y.; KIM, K. S.; KO, M. K.; CHEONG, S. J.; OH, B. J. A thaumatin-like gene in nonclimacteric pepper fruits used as molecular marker in probing disease resistance, ripening, and sugar accumulation. **Plant Molecular Biology**, v. 2, p. 125-135, 2002.

KIMATI, H. Doenças da mamoneira – *Ricinus communis* L. In: GALLI, F. (coord). Manual de Fitopatologia. São Paulo: Agronômica Ceres, v. 2, p. 347-351, 1980.

KINKEMA, M.; FAN, W.; DONG, X. Nuclear Localization of NPR1 Is Required for Activation of PR Gene Expression. **The Plant Cell**, v. 12, n. 12, p. 2339-2350, 2000.

KNOGGE, W. Fungal infection of plants. **The Plant Cell**, v. 8, p. 1711-1722, 1996.

KRUG, C.A.; MENDES, P.T.; SOUZA, G.F. de. Melhoramento de mamoneira (*Ricinus comunis*, L.). III. Primeira série de ensaios de variedades (1937/38 – 1938/39). **Bragantia, Campinas**, v. 3, p. 85-122, 1943.

LALUK, K.; MENGISTE, T. The Arabidopsis extracellular unusual serine protease inhibitor functions in resistance to necrotrophic fungi and insect herbivory. **The Plant Journal**, v. 3, p. 480–494, 2011.

LAPPI, D. A.; KAPMEYER, W.; BEGLAU, J. M.; KAPLAN, N. O. The disulfide bond connecting the chains of ricin. **Proceedings of the National Academy of Sciences**, v. 75, p. 1096–1100, 1978.

LIMA, E. F.; ARAÚJO, A. E.; BATISTA, F. A. S. Doenças e seu controle. In: **O agronegócio da mamona no Brasil**. ed. Brasília: Embrapa Informação Tecnológica, 2001. p. 191-212.

LIMA, L. S. P.; OLIVEIRA, D. A.; PEIXOTO, L. S. Cultivares de Mamona com bom desempenho agrônômico para condições edafoclimáticas do Sertão Produtivo da Bahia. **VII CONNEPI**, v. 01, p. 1-7, 2012.

LIU, H.; ZHANG, C.; TIAN, H.; LI, L.; WANG, X.; QIU, T. Environmental and techno-economic analyses of bio-jet fuel produced from jatropha and castor oilseeds in China. **The International Journal of Life Cycle Assessment**, v. 26, p. 1071–1084, 2021.

LIU, J-J.; STURROCK, R.; EKRAMODDOULLAH, A. M. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. **Plant Cell Reports**, v. 5, p. 419–436, 2010.

LIU, X.; HU, P.; HUANG, M.; TANG, Y.; LI, Y.; LI, L.; HOU, X. The NFYC-RGL2 module integrates GA and ABA signalling to regulate seed germination in Arabidopsis. **Nature Communications**, v. 7, p. 1-14, 2016.

LIU, Q.; REN, T.; ZHANG, Y.; LI, X.; CONG, R.; WHITE, P. J.; LU, J. Yield loss of oilseed rape (*Brassica napus* L.) under nitrogen deficiency is associated with under-regulation of plant population density. **European Journal of Agronomy**, v. 103, p.80-89, 2019.

MACHADO, E. L. **Genotipagem, Teor de Óleo, Otimização do Método Gravimétrico e Desenvolvimento de Marcadores Microsatélite em (*Ricinus communis* L.)**. Tese (Doutorado em Ciências Agrárias). Universidade Federal do Recôncavo da Bahia -UFRB. Cruz das Almas – BA, 118 f, 2011.

MACHADO, E. L.; SILVA, S. A. Desenho e validação de iniciadores microsatélites SSR para mamoneira. **Pesquisa Agropecuária Brasileira, Brasília**, v. 28, p. 1457 - 1463, 2013.

MADEIRA, J. V. JR.; MACEDO, J. A.; MACEDO, G. A. Detoxification of castor bean residues and the simultaneous production of tannase and phytase by solid-state fermentation using *Paecilomyces variotii*. **Bioresource Technology**, v. 102, p. 7343–7348, 2011.

MARQUES, D. N.; SIQUEIRA, M. A.; GONÇALVES, E. C.; BARROS, N. L. F.; SOUZA, C. R. B. Homology modeling and molecular dynamics simulations of a cassava translationally controlled tumor protein (MeTCTP). **Plant Gene**, v. 19, p. 1-4, 2019.

MCCOLLUM, T. G.; DOOSTDAR, H.; MAYER, R. T.; MCDONALD, R. E. Characterization of chitinases and beta-1,3-glucanases in grapefruit flavedo during fruit development. **Physiologia Plantarum**, v. 99, p. 486-494, 2006.

MITSUHARA, I.; IWAI, T.; SEO, S.; YANAGAWA, H.; HIROSE, S.; OHKAWA, Y., OHASHI, Y. Characteristic expression of twelve rice PR1 family genes in response to pathogen infection, wounding, and defense-related signal compounds (121/180). **Molecular Genetics and Genomics**, v. 279, p. 415-427, 2008.

MOSHKIN, V. A. **Castor**. Amerind Publishing Co. PVT Ltd, New Delhi: Amerind Publishing, p. 315, 1986.

MUSSHOFF, F.; MADEA, B. Ricin poisoning and forensic toxicology. **Drug Testing and Analysis**, v. 1, p. 184–191, 2009.

MUSHOBOZY, D. M. K.; NGANILEVANU, G; RUHEZA, S.; SWELLA, G.B. Plant oils as common bean (*Phaseolus vulgaris* L.) seed protectants against infestations by the Mexican bean weevil *Zabrotes subfasciatus* (Boh.). **Journal of Plant Protection Research**, v. 49, p.35-39, 2009.

MUTLU, H.; MEIER, M. A. R. Castor oil as a renewable resource for the chemical industry. **European Journal of Lipid Science and Technology**, v. 112, p. 10-30, 2010.

NASCIMENTO, V. V.; CASTRO, H. C.; ABREU, P. A.; OLIVEIRA, A. E. A.; FERANDEZ, J. H.; ARAUJO, J. S.; MACHADO, O. L. T. In Silico Structural Characteristics and α -Amylase Inhibitory Properties of Ric c 1 and Ric c 3, Allergenic 2S Albumins from *Ricinus communis* Seeds. **Journal of Agricultural and Food Chemistry**, v. 59, p. 4814-4821, 2011.

NEERGAARD, P. Seed pathology. **Londres: McMillan**, v. 1, p. 839, 1979.

NIKI, T.; MITSUHARA, I.; SEO, S.; OHTSUBO, N., OHASHI, Y. Antagonistic effect of salicylic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounded mature tobacco leaves. **Plant and Cell Physiology**, v. 39, n. 5, p. 500-507, 1998.

NORVELL, J. C.; MACHALEK, A. Z. Structural genomics programs at the US National Institute of General Medical Sciences. **Nature Structural Biology**, v. 7, p. 931, 2000.

OGUNNIYI, D. S. Castor oil: A vital industrial raw material. **Bioresource Technology Barking**, v. 97, p.1086-1091, 2006.

OHASHI, Y.; MATSUOKA, M. Induction and secretion of Pathogenesis-Related Proteins by salicylate or plant hormones in tobacco suspension cultures. **Plant and Cell Physiology**, v. 28, p. 573-580, 1987.

O'LEARY, S. J.; POULIS, B. A.; VON ADERKAS, P. Identification of two thaumatin-like proteins (TLPs) in the pollination drop of hybrid yew that may play a role in pathogen defence during pollen collection. **Tree Physiology**, v. 12, p. 1649-59, 2007.

OLIVEIRA, I. J.; ZANOTTO, M. D.; KRIEGER, M. VENCOVSKY, R. Inbreeding depression in castor bean (*Ricinus communis* L.) progenies. **Crop Breeding and Applied Biotechnology**, v. 12, p. 269-276, 2012.

OLIVEIRA NETO, S. S. **CARACTERIZAÇÃO AGROMORFOLÓGICA E REAÇÃO AO MOFO CINZENTODE ACESSOS DE MAMONEIRA**. Dissertação (Mestrado em Agronomia - Agricultura) – Faculdade de Ciências Agrônômicas da Unesp Câmpus de Botucatu, Botucatu, p. 30. 2017.

OLIVEIRA, R. S.; SILVA, S. A.; BRASILEIRO, B. P.; MEDEIROS, E. P.; DOS ANJOS, E. V. A. Divergência genética em mamoneira utilizando a estratégia ward-mlm. **Revista Ciência Agrônômica**, v. 44, p. 564-570, 2013.

OLIVEIRA, R. S. **Avaliação de população segregante (F3) de mamoneira em condições do Recôncavo Baiano**. Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia, UFRB, Cruz das Almas, 40f, 2011.

OLSNES, S.; FERNANDEZ-PUENTES, C.; CARRASCO, L; VAZQUEZ, D. Ribosome inactivation by the toxic lectins abrin and ricin. **European Journal of Biochemistry**, v. 60, p. 281–288, 1975.

PARK, C. J.; KIM, K. J.; SHIN, R.; PARK, J. M.; SHIN, Y. C.; PAEK, K. H. Pathogenesis-related Protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. **The Plant Journal**, v. 37, n. 2. p. 186–198, 2004.

PARMAR, H.; KAPADIYA, K. J.; BHALIYA, C. M. Integrated management of root rot of castor (*Ricinus communis* L.) caused by *Macrophomina phaseolina* (Tassi) Goid. **International Journal of Chemical Studies**, v. 6, p. 849-851, 2018.

PASSOS, A. R.; SILVA, S. A.; SOUZA, C. S.; SOUZA, C. M. M.; FERNANDES, L. S. Parâmetros genéticos de caracteres agrônômicos em genótipos de mamoneira. **Pesquisa Agropecuária Brasileira**, v. 45, p. 709-714, 2010.

PASSOS; A. R.; SILVA; S. A.; HARTWIG; I.; HAWERROTH; M. C.; SANTOS; K. S.; SOUZA; C. M. M. Avaliação de cultivares de mamoneira para tolerância ao alumínio tóxico e insensibilidade ao ácido giberélico. **Magistra**, v. 27, p. 73-81, 2015.

PRAZERES, A. G.; SILVA, S. A.; SANTOS, A. S.; SILVA, R. O.; REIS, A. S. Reação de cultivares de Mamoneira com fungicidas e as características agrônômicas. **Agrotrópica**, v. 1, p. 75 - 84, 2017.

RAMOS-LÓPEZ, M. A.; PÉREZ, G.S.; RODRÍGUEZ-HERNÁNDEZ, C.; GUEVARA-FEFER, P.; ZAVALA-SÁNCHEZ, M.A. Activity of *Ricinus communis* (Euphorbiaceae) against *Spodoptera frugiperda* (Lepidoptera: Noctuidae). **African Journal of Biotechnology**, v. 9, p.1359-1365, 2010.

RAMPRASAD, R.; BANDOPADHYAY, R. Future of *Ricinus communis* after completion of the draft genome sequence. **Current Science**, v. 99, p. 1316-1318, 2010.

ROSSI, G. D.; SANTOS, C. D.; CARVALHO, G. A.; ALVES, D. S.; PEREIRA, L. L. S. CARVALHO, G. A. Biochemical Analysis of a Castor Bean Leaf Extract and its Insecticidal Effects Against *Spodoptera frugiperda* (Smith) (Lepidoptera: Noctuidae). **Neotropical Entomology**, v. 41, p. 503-509, 2012.

SANTAMARIA, M.; THOMSON, C. J.; READ, N. D.; LOAKE, G. J. The promoter of a basic PR1-like gene, AtPRB1, from *Arabidopsis* establishes an organ-specific expression pattern and responsiveness to ethylene and methyl jasmonate. **Plant Molecular Biology**, v. 47, p. 641-652, 2001.

SANTOS, L. A. **Caracterização e seleção de linhagens elites de mamoneira (*Ricinus communis* L.) por meio do desempenho morfoagronômico e molecular no Estado da Bahia**. Tese (Doutorado) - Universidade Federal do Recôncavo da Bahia - UFRB, Cruz das Almas – BA, 85 f, 2013.

SANTOS, L. A.; SILVA, S. A.; SOUZA, D. R.; ARAUJO, G. M.; ALMEIDA, V. O.; NEVES, C. G. Genetic parameters in agronomic characters for selection of superior lineages of *Ricinus communis* L. **Ciência Rural, Santa Maria**, v.4, p. 1- 6, 2019.

SANTOS, R. F.; KOURI, J.; BARROS, M. A. L.; MARQUES, F. M.; FIRMINO, P. T.; REQUIÃO, L. E. G. Aspectos econômicos do agronegócio da mamona. In: **O agronegócio da mamona no Brasil**. 2a ed. Campina Grande: Embrapa Algodão - Brasília, Embrapa Informação Tecnológica. 2007. p. 21-42, 2007.

SARWAR, M. W.; RIAZ, A.; NAHID, N.; QAHTANI, A. A.; AHMED, N.; NAWAZ-UL-REHMAN, M. S.; YOUNUS, A.; MUBIN, M. Homology modeling and docking analysis of β C1 protein encoded by Cotton leaf curl Multan betasatellite with different plant flavonoids. **Heliyon**, v. 5 p. 1-19, 2019.

SCHWEDE, T. Protein Modeling: What Happened to the "Protein Structure Gap"? **Structure**, v. 9, p. 1531-1540, 2013.

SEAGRI <<http://www.seagri.ba.gov.br/noticias/2011/12/05/ebda-lan%C3%A7a-nova-variedade-de-mamona>>. Acesso em 10 de fevereiro de 2020.

SELS, J.; MATHYS, J.; CONINCK, B. M. A.; CAMMUE, B. P. A.; BOLLE, M. F. C. Plant pathogenesis-related (PR) proteins: a focus on PR peptides. **Plant Physiology and Biochemistry**, v. 46, p. 941-950, 2008.

SEVERINO, L. S.; AULD, D. L.; BALDANZI, M.; CANDIDO, M. J. D.; CHEN, G.; CROSBY, W.; TAN, D.; HE, X.; LAKSHMAMMA, P.; LAVANYA, C.; MACHADO, O. L. T.; MIELKE, T.; MILANI, M.; MILLER, T. D.; MORRIS, J. B.; MORSE, S. A.;

NAVAS, A. A.; SOARES, D. J.; SOFIATTI, V.; WANG, M. L.; ZANOTTO, M. D.; A Review on the Challenges for Increased Production of Castor. **Agronomy Journal**, v. 104, p. 853-880, 2012.

SHAHEEN, T.; RAHMAN, M-U.; ABBAS, M. F.; ZAIB, P.; RAHMAN, M-U.; ULLAH, I.; QAMAR, M. T-U.; ARIF, A. Identification, characterization, homology modeling and protein-protein interactions of cotton (*Gossypium arboreum*) DREB gene. **International Journal of Agriculture and Biology**, v. 20, p. 1055-1061, 2018.

SHAW, R. K.; KADIRVEL, P.; SHAIK, M.; PRASAD, M. S. L.; PRASAD, R. D.; SENTHILVEL, S. Genetic characterization of resistance to wilt disease caused by *Fusarium oxysporum* f. sp. *ricini* in castor (*Ricinus communis* L.). **Plant Genetic Resources**, v. 16, p. 169-177, 2017.

SHI, G.; CAI, Q. Cadmium tolerance and accumulation in eight potential energy crops, **Biotechnology Advances**, v. 27, p. 555-561, 2009.

SILVA, A. R.; SILVA, S. A.; SANTOS, L. A.; SOUZA, D. R.; ARAUJO, G. M.; DANTAS, J. L. L.; LEITE, E. D.; DANTAS, A. C. V. L. Characterization and performance of castor bean lineages and parents at the UFRB germplasm bank. **Plos one**, v 14, p. -15, 2019.

SILVA, A. R.; SILVA, S. A.; SANTOS, L. A.; SOUZA, D. R.; ARAUJO, G. M.; MOREIRA, R. F. C. Genetic divergence among castor bean lines and parental strains using ward's method based on morpho-agronomic descriptors. **Acta Scientiarum Agronomy**, v. 39, p. 307-313, 2017.

SOARES, D. J. Gray mold of castor: a review. In: Cumagun CJR (ed) **Plant Pathology. InTech, Rijeka**, p. 219–240, 2012.

SOUSA, F. Q. **Prospecção fenotípica de linhagens elites de mamoneira resistentes ao mofo cinzento (*Amphobotrys ricini* (N.F. Buchw.) Hennebert)**. Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia - UFRB, Cruz das Almas – BA, 80 f, 2014.

SOUSA, N. L.; CABRAL, G. B.; VIEIRA, P. M.; BALDONI, A. B.; ARAGÃO, F. J. L. Bio-detoxification of ricin in castor bean (*Ricinus communis* L.) seeds. **Scientific Reports**, v. 7, p. 1-9, 2017.

SUSSEL, A. A. B. **Epidemiologia e manejo do mofo-cinzento-da-mamoneira**. Planaltina: EMBRAPA Cerrados, 2009. Documentos 241.

TATTERSALL, D. B.; VAN HEESWIJCK, R.; HOJ, P. B. Identification and characterization of a fruit-specific, thaumatin-like protein that accumulates at very high levels in conjunction with the onset of sugar accumulation and berry softening in grapes. **Plant Physiology**, v.3, p. 759-769, 1997.

TERRET-WELTER, Z.; BONNET, G.; MOURY, B.; GALLOIS, J-L. Analysis of tomato spotted wilt virus RNA-dependent RNA polymerase adaptative evolution and constrained domains using homology protein structure modelling. **Journal of General Virology**, v. 10, p. 1-10, 2020.

TONON, C.; GUEVARA, G.; OLIVA, C.; DALEO, G. Isolation of a potato acidic 39 kDa β -1,3-glucanase with antifungal activity against *Phytophthora infestans* and analysis of its expression in potato cultivars differing in their degrees of field resistance. **Journal Phytopathology**, v. 150, p.189–195, 2002.

UMAARASU, T.; PADMAVATHY, K.; THIRUNAVUKKARASU, D.; RAJESH, S. V.; SHANMUGAM, G. Molecular Docking Studies Of Ricinus Communis Phytochemicals Against Beta-Lactamase From *Enterococcus Faecalis* And *Staphylococcus Aureus*. **Asian Journal of Pharmaceutical and Clinical Research**, v. 11, p. 132-137, 2018.

URBACH, J. M.; AUSUBEL, F. M. The NBS-LRR architectures of plant R-proteins and metazoan NLRs evolved in independent events. **Proceedings of the National Academy of Sciences**, v.114, p. 1063–1068, 2017.

VAN LOON, L. C. The nomenclature of pathogenesis-related proteins. **Physiological and Molecular Plant Pathology**, v. 37, p. 229-230, 1990.

VAN LOON, L. C. Occurrence and properties of plant pathogenesis-related proteins. In: **Pathogenesis-related proteins in plants**. Eds. S.K. Datta, S. Muthukrishnan, CRC Press LLC, Boca Raton, 1999. p. 1-19.

VAN LOON, L. C.; PIERPOINT, W. S.; BOLLER, T.; CONEJERO, V. Recommendations for naming plant pathogenesis-related proteins. **Plant Molecular Biology Reporter**, v. 12, p. 245-264, 1994.

VAN LOON, L. C.; REP, M.; PIETERSE, C. M. J. Significance of inducible defense-related proteins in infected plants. **Annual Review of Phytopathology**, v. 44, n. 1, p. 135-162, 2006.

VAN LOON, L.C., E.A. VAN STRIEN. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. **Physiological and Molecular Plant Pathology**, v. 55, p. 85-97, 1999.

VAN LOON, L.C.; A. VAN KAMMEN. Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN" II. Changes in protein constitution after infection with tobacco mosaic virus. **Virology**, v. 40, p. 199- 211, 1970.

WANG, M. L.; MORRIS, J. B.; TONNIS, B.; PINNNOW, D.; DAVIS, J.; RAYMER, P.; PEDERSON, G. A. Screening of the Entire USDA Castor Germplasm Collection for Oil Content and Fatty Acid Composition for Optimum Biodiesel Production. **Journal of Agricultural and Food Chemistry**, v. 59, p. 9250-9256, 2011.

WANG, W.; CHENG, Y.; CHEN, D.; LIU, D.; HU, M.; DONG, J.; ZHANG, X.; SONG, L.; SHEN, F. The Catalase Gene Family in Cotton: Genome-Wide Characterization and Bioinformatics Analysis. **Cells**, v. 8, p. 1-27, 2019.

WEBB, B.; SALI, A. Comparative Protein Structure Modeling Using MODELLER. **Current Protocols in Bioinformatics**, v. 54, p. 1-55, 2016.

WEBER, E. The most complete global overview of invasive species in natural areas. **Diversity and Distributions**, v.10, p. 505, 2003.

WORBS, S.; KOHLER, 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. **Journal Toxins**, v. 3, p. 1332–1372, 2011.

WU, J.; KIM, S. G.; KANG, K. Y.; KIM, J-G.; PARK, S-R.; GUPTA, R.; KIM, Y. H.; WANG, Y.; KIM, S. T. Overexpression of a Pathogenesis-Related Proteins 10 Enhances Biotic and Abiotic Stress Tolerance in Rice. **The Plant Pathology Journal**, v. 6, p. 552-562, 2016.

YAMUNA, C.; VARMA, P. K.; PRASAD, R. D.; LAKSHMI, K. V. Morphological and molecular characterization of anamorph associated with gray mold of castor (Ricinus communis L.). **Journal of Oilseeds Research**, v. 32, p. 63-67, 2015.

YU, L.; ZHANG, J.; XU, F.; YANG, L.; LI, G. Q. First report of Amphobotrys ricini causing gray mold disease on Acalypha australis in Central China. **Plant Disease**, v. 96, p. 460. 2012.

YOUSAFI, Q.; AHMAD ALI, H.; HAMID, E.; KHAN, M. S. In silico comparative proteomic analysis of enzymes involved in fatty acid biosynthesis in castor bean (Ricinus communis) and soybean (Glycine max). **Turkish Journal of Botany**, v. 43, p. 1-26, 2019.

ZHU, T.; WU, S.; ZHANG, D.; LI, Z.; XIE, K.; AN, X.; MA, B.; HOU, Q.; DONG, Z.; TIAN, Y.; LI, J.; WAN, X. Genome-wide Analysis of Maize GPAT Gene Family and Cytological Characterization and Breeding Application of ZmMs33/ZmGPAT6 Gene. **Theoretical and Applied Genetics**, v. 7, p. 2137-2154, 2019.

ZOU, Z.; ZHANG, X. Genome-wide identification and comparative evolutionary analysis of the Dof transcription factor family in physic nut and castor bean. **PeerJ**, v. 7, p. 1-25, 2019.

ZOU, Z.; ZHU, J.; ZHANG, X. Genome-wide identification and characterization of the Dof gene family in cassava (Manihot esculenta). **Gene**, v. 687, p. 298-307, 2019.

ZUIN, C. G.; VIDIGAL FILHO, S. P.; KVITSCHAL, V. M.; GONÇALVESVIDIGAL, C. M.; COIMBRA, K. G. Genetic divergence among accesses of 25 sweet-cassava collected from Cianorte county, northwestern region of Paraná State. **Semina: Ciências Agrárias**, v. 30, p. 21-30, 2009.

ARTIGO 1**Caracterização das famílias gênicas codificantes de Proteínas
Relacionada à Patogênese (PRs) em *Ricinus communis* L.**

¹Artigo a ser ajustado para posterior submissão ao Comitê Editorial do periódico científico *Plus One*, em versão na língua inglesa.

Caracterização das famílias gênicas codificantes de Proteínas Relacionada à Patogênese (PRs) em *Ricinus communis* L.

RESUMO: As Proteínas Relacionadas à Patogênese (PRs) fazem parte de um mecanismo de defesa complexo que atua contra estresses bióticos ou abióticos em plantas. Assim, objetivou-se caracterizar famílias gênicas das PRs em *Ricinus communis*, chamadas neste trabalho de RcPRs, em níveis de sequência nucleotídica e peptídica, incluindo a região promotora e análise filogenética das proteínas RcPR-1c, RcPR-4b, RcPR-5a, RcPR-6 e RcPR10-STH21. A identificação e análise estrutural dos genes *RcPR* foi realizada no PHYTOZOME V 12.1, enquanto a análise da região promotora foi utilizado o servidor PlantCARE. Os domínios conservados e a superfamília proteica foram analisados no Pfam 32.0. A predição da localização subcelular, ponto isoelétrico (pI), peso molecular (MW) e índice médio de hidropaticidade (GRAVY) das proteínas foram obtidas usando os servidores DeepLoc e ProtParam. A identificação de locais de fosforilação e N-glicosilação foi dada a partir dos servidores NetPhos e NetNGlyc. O Clustal Omega e o MEME/MAST foram utilizados para identificação de aminoácidos e motivos conservados. Dos 20 genes *RcPRs*, 10 são *RcPR-1*; três *RcPR-4*; cinco *RcPR-5*, uma *RcPR-6* e uma *RcPR10-STH21*, com éxons variável de 1 a 3 e íntrons de 1 a 2, em um painel de leitura aberto (ORF) de 222 a 795 pb. A região de 1500pb a montante de cada gene PR apresentou elementos *cis*-reguladores relacionados a fatores de resposta internos e externos. O pI e o MW das proteínas com Peptídeo Sinal (PS) foi de 4,68 a 9,17 e 14,96 a 31,26 kDa e sem o PS foi de 4,58 a 9,35 e 12,67 a 28,83 kDa, respectivamente. O GRAVY indicou regiões hidrofóbicas nas proteínas RcPR-5a e RcPR-5e, e nas demais, regiões hidrofílicas. Quanto a filogenia, houve a formação de oito grupos, compostos majoritariamente por membros de uma mesma família. Estes dados fornecem informações das RcPRs para utilização funcional e contribuirá no desenvolvimento de estratégias para aumentar a resistência na espécie contra fitopatógenos na mamona.

Palavras-chaves: Análise *in silico*; *Cis*-elementos; Mamona

Characterization of the gene families encoding Pathogenesis-Related Proteins (PRs) in *Ricinus communis* L.

ABSTRACT: Pathogenesis-Related Proteins (PRs) are part of a complex defense mechanism that acts against biotic or abiotic stresses in plants. Thus, the objective was to characterize gene families of PRs in *Ricinus communis*, called RcPRs in this work, at nucleotide and peptide sequence levels, including the promoter region and phylogenetic analysis of the proteins RcPR-1c, RcPR- 4b, RcPR-5a, RcPR -6 and RcPR10-STH21. The identification and structural analysis of the RcPR genes was performed in PHYTOZOME V 12.1, while the analysis of the promoter region was used in the PlantCARE server. The conserved domains and the protein superfamily were analyzed in Pfam 32.0. The prediction of subcellular location, isoelectric point (pI), molecular weight (MW) and mean hydropathicity index (GRAVY) of proteins were obtained using DeepLoc and ProtParam servers. The identification of phosphorylation and N-glycosylation sites was given from the NetPhos and NetNGlyc servers. Clustal Omega and MEME/MAST were used to identify amino acids and conserved motifs. Of the 20 RcPRs genes, 10 are RcPR-1; three RcPR4; five RcPR-5, one RcPR-6 and one RcPR10-STH21, with variable exons from 1 to 3 and introns from 1 to 2, in an open reading frame (ORF) from 222 to 795 bp. The 1500bp region upstream of each PR gene showed cis-regulatory elements related to internal and external response factors. The pI and MW of proteins with Signal Peptide (PS) ranged from 4.68 to 9.17 and 14.96 to 31.26 kDa and without PS ranged from 4.58 to 9.35 and 12.67 to 28 .83 kDa, respectively. GRAVY indicated hydrophobic regions in the RcPR-5a and RcPR-5e proteins, and in the others, hydrophilic regions. As for phylogeny, there was the formation of eight groups, composed mostly by members of the same family. These data provide information on RcPRs for functional use and will contribute to the development of strategies to increase resistance in the species against phytopathogens in mammary glands.

Keywords: In silico analysis; Cis-elements; Castor bean

INTRODUÇÃO

Ao longo do processo evolutivo, as plantas desenvolveram diferentes mecanismos de defesa constitutiva e induzida contra condições ambientais adversas e ataque de microrganismos (TAKKEN; GOVERSE, 2012). No entanto, para que ocorra esta resposta a resistência contra fitopatógenos, receptores imunes nas plantas deverão ser ativados como resposta de defesa eficaz após a detecção de microrganismos (BOCCARDO et al., 2019).

Assim, a imunidade das plantas é desencadeada pela detecção dos Padrões Moleculares Associados a Micróbios - MAMPs (do inglês *Microbe-Associated Molecular Pattern*) e a Patógenos - PAMPs (do inglês *Pathogen-Associated Molecular Pattern*), que são reconhecidos pelos Receptores de Reconhecimento de Padrões – RRP (do inglês *Pattern Recognition Receptors*) do sistema imune inato das plantas. Estes padrões ativam a Imunidade Desencadeada por Padrões - PTI (do inglês *Pattern-Triggered Immunity*), e na ausência desse padrão, a detecção é feita pelos efetores de fitopatógenos que são reconhecidos pelos genes R e ativam a Imunidade Desencadeada por Efetores - ETI (do inglês *Effector-Triggered Immunity*) (JONES; DANGL 2006).

Este reconhecimento gera uma série de respostas, tais como a produção de Espécies Reativas de Oxigênio (ERO), ácido salicílico (AS), ácido jasmônico (AJ), auxinas e giberelinas; acúmulo de fitoalexinas e das Proteínas Relacionadas à Patogênese (Proteínas-PR) (PARK et al., 2004; VAN LOON et al., 2006; KAUR et al., 2017; BOCCARDO et al., 2019).

Segundo Sels et al. (2008), estas proteínas estão presentes em baixa concentração nas plantas e são induzidas em maiores quantidades em situações patológicas, mas não necessariamente implicadas em uma interação direta com fitopatógeno. Algumas acumulam-se localmente desencadeando a Resposta Hipersensível (RH), enquanto outras estão envolvidas na Resistência Adquirida Sistemicamente (RAS), promovendo a formação de barreiras químicas (KINKEMA et al., 2000).

Segundo Arora et al. (2020), as proteínas-PRs foram caracterizadas em 19 famílias (PR-1 a PR-19), apresentando diversas atividades enzimática como glucanases (PR-2; TONON et al., 2002), osmotinas e taumatinas (PR-5; LIU J-J et al., 2010), inibidores de protease (PR-6; LALUK; MENGISTE, 2011), lisozimas (PR-8; HEITZ et al., 1994), peroxidase (PR-9; GHOSH, 2006), ribonucleases

(PR10; AGARWAL et al., 2013) e quitinases (PR-3, PR-4, PR-8, PR-11; BERTINI et al., 2009). Podendo ainda apresentar atividade defensina (PR-12), tionina (PR-13), de transferência lipídica (PR-14) e de oxalato oxidase (PR-15; PR-16). Destas 19 famílias, duas delas (PR-1 e PR-17) ainda não tiveram suas funções totalmente elucidadas (VAN LOON et al., 2006; MENEZES et al., 2014).

Na cultura da mamona (*Ricinus communis* L.), assim como em outras culturas, existem as restrições de produções nas áreas de cultivo, decorrentes do ataque de fitopatógenos que desempenham um papel proeminente (PRASAD et al., 2019; SALMAN; ABU-NASER, 2019).

No Brasil, pode-se destacar o mofo-cinzento, causado pelo fungo *Amphobotrys ricini* Buchw., como sendo a principal doença da cultura da mamoneira, disseminado em todas as áreas de cultivo dentro do território brasileiro, ocorrendo agressivamente nas regiões de condições climáticas favoráveis, com temperatura em torno de 25°C e a umidade relativa do ar sempre acima de 75% (CHAGAS et al., 2010).

O fungo *A. Ricini* ataca as inflorescências da planta, cachos e frutos, reduzindo assim a produção de óleo, que é o componente de maior valor agregado a cultura, pela diminuição dos frutos colhidos e sua qualidade (CHAGAS et al., 2010).

Desta forma, a cidade de Cruz das Almas, localizada na região do Recôncavo da Bahia – Brasil, é uma região que apresenta as condições favoráveis para disseminação desta doença, com temperatura média anual de 24,5°C, 80% de umidade relativa e 1.250 mm de precipitação anual (ALBUQUERQUE et al., 2018).

Atualmente, o Melhoramento Genético é a melhor alternativa para controle de doenças e obtenção de variedades e/ou cultivares resistentes. Por isso, os programas de melhoramento têm buscado formas alternativas de avaliar os genomas das espécies.

Assim, com o advento da simulação computacional, através da bioinformática é possível identificar genes, prever estruturas tridimensionais de proteínas por homologia, identificar inibidores de enzimas, organizar e relacionar informações biológicas, agrupar proteínas homólogas e montar árvores filogenéticas, sem elevados custos financeiros, a partir dos resultados oriundos do sequenciamento genético (CASTRO et al., 2017; KOLAHİ et al., 2018).

Estes resultados fornecem informações confiáveis e valiosas que

contribuem no desenvolvimento de estratégias que visam aumentar a resistência na espécie contra fitopatógenos, auxiliando assim os programas de melhoramento, pelo melhor entendimento da interação planta-patógeno.

No entanto, na mamona, ainda não foi realizada este tipo de estudo com as Proteínas-PR. O sequenciamento e a anotação do genoma foram concluídos, porém, ainda em sua primeira versão (320 Mb) com tamanho relativamente pequeno, distribuído em 10 cromossomos. Até o momento não há mapa de ligação genético disponível para a espécie, assim o genoma encontra-se anotado em *scaffolds* (CHAN et al., 2010), no entanto, estas informações são suficientes para serem exploradas com uso de ferramentas de bioinformática.

Assim, o presente estudo objetivou caracterizar *in silico* as proteínas-PR em *Ricinus communis*, em nível de nucleotídeos e proteínas, incluindo a análise dos promotores. Espera-se que estes resultados proporcionem ao Melhoramento Genético da Mamona avanços significativos nas pesquisas que envolvam resistência, principalmente ao *Amphobotrys ricini*.

MATERIAL E MÉTODOS

Análise *in silico* de genes *RcPRs*

A identificação e análise estrutural dos genes *RcPRs* (íntrons/éxons) foram realizadas utilizando o Banco de Dados de Genomas PHYTOZOME V 12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>), por meio do acesso ao genoma de *Ricinus communis*. A análise do painel de leitura aberto – ORF (do inglês *Open Reading Frames*) foi realizada usando o servidor *ORFfinder* (<http://www.ncbi.nlm.nih.gov/orffinder/>), que procura quadros de ORF por meio da sequência de nucleotídeos obtidas da biblioteca de DNA em formato FASTA. Para a identificação da estrutura gênica, foi utilizado as sequências genômicas e codantes - CDS (do inglês *coding sequence*) em formato FASTA no *Gene Structure Display Server 2.0* (http://gsds.cbi.pku.edu.cn/Gsds_upload.php) (HU et al., 2015).

Análise das regiões promotoras dos genes *RcPRs*

Para identificar a presença dos elementos reguladores *cis* nas regiões promotoras dos genes *RcPRs*, foi adicionado a região a montante de 1500 pb a

partir do local inicial da tradução dos genes (*upstream/downstream*) e posteriormente analisados usando o programa plantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (LESCOT et al., 2002), por meio da sequência em formato FASTA.

Análise *in silico* de proteínas PRs

O ponto isoelétrico teórico - pI (do inglês *isoelectric point*), peso molecular - MW (do inglês *molecular weight*) e o índice de hidropaticidade média - GRAVY (do inglês *Grand average of hydropathicity*) foram identificados usando o servidor ExPASy ProtParam (<https://web.expasy.org/protparam/>) (GASTEIGER et al., 2005). A previsão da localização subcelular e do peptídeo sinal foi realizada com o auxílio dos servidores DeepLoc 1.0 (<http://www.cbs.dtu.dk/services/DeepLoc/>) e SignalP 1.1 (<http://www.cbs.dtu.dk/services/SignalP/>), respectivamente (ARMENTEROS et al., 2017; ARMENTEROS et al., 2019). Os servidores NetPhos 3.1 (<http://www.cbs.dtu.dk/services/NetPhos/>) (BLOM et al., 1999) e NetNGlyc 1.9 (<http://www.cbs.dtu.dk/services/NetNGlyc/>) foram usados para identificar possíveis locais de fosforilação (Ser/Thr/Tyr) e possíveis locais de N-glicosilação (Tipo Asn-X-Ser/Thr), respectivamente. Os domínios conservados e a superfamília proteica foram identificados pelo servidor Pfam v 32.0 (<https://pfam.xfam.org/>) (EL-GEBALI et al., 2019). Para o alinhamento e identificação dos aminoácidos conservados entre as proteínas-PR foi utilizado o servidor Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (SIEVERS et al., 2011). A análise dos motivos proteicos foi realizada utilizando o MEME/SUITE (<http://meme-suite.org/tools/meme>) (BAILEY et al., 2009). O número máximo de motivos foi ajustado para 20, o comprimento máximo dos motivos, a largura ótima, e os demais parâmetros foram usados como padrão. Todas estas análises foram realizadas pela inserção das sequências proteicas, no formato código de letras no servidor ExPASy ProtParam e formato FASTA nos demais respectivos servidores supracitados.

Alinhamentos e filogenia da estrutura genética das *RcPRs*

O alinhamento entre as 20 sequências de aminoácidos das Proteínas-PR de *Ricinus communis* com as sequências de *Nicotiana tabacum* L. (Tabaco) e *Arabidopsis thaliana* L. foram realizadas no servidor Clustal Omega

(<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (SIEVERS et al., 2011). A escolha de *N. tabacum* deve-se porque as proteínas-PR foram identificadas pela primeira vez em folhas de uma cultivar dessa espécie e *A. thaliana* por ser um organismo modelo para estudos de genética. O alinhamento entre as sequências proteicas foi realizado no programa MEGA 7 (KUMAR et al., 2016), posteriormente foi construída uma árvore filogenética pelo método estatístico *neighbor-joining*, que utiliza uma matriz de distância genética para unir em pares as sequências mais próximas, usando o modelo JTT. Para as sequências que não foram unidas em pares na etapa anterior é calculada uma nova distância para cada par. Esta etapa repete sucessivas vezes até que a árvore seja totalmente gerada (SAITOU; NEI 1987).

RESULTADOS

Famílias gênicas PRs no genoma de *Ricinus communis* L.

A anotação existente no Banco de Dados PHYTOZOME V 12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>) do genoma de *Ricinus communis* permitiu a identificação de 20 genes, chamados neste trabalho de *RcPR-1a*, *RcPR-1b.1*, *RcPR-1b.2*, *RcPR-1c*, *RcPR-1d*, *RcPR-1Be*, *RcPRB1-3f*, *RcPR-1g*, *RcPR-1h*, *RcPR-1i*, *RcPR-4a.1*, *RcPR-4a.2*, *RcPR-4b*, *RcPR-5a*, *RcPR-5b*, *RcPR-5c*, *RcPR-5d*, *RcPR-5e*, *RcPR-6* e *RcPR10-STH21*, classificados em cinco diferentes famílias: PR-1; PR-4; PR-5; PR-6 e PR-10 STH21.

Particularmente, nos genes *RcPR-1b* e *RcPR-4a* foram identificados *splicing* alternativo sofrido no RNA mensageiro, conforme registro no Banco de Dados PHYTOZOME V 12.1, pela ferramenta *nucleotide query to nucleotide db* - BLASTN (<https://phytozome.jgi.doe.gov/pz/portal.html>), em razão disto, as formas alternativas foram denominadas de genes *RcPR-1b.1* e *RcPR-1b.2*; *RcPR-4a.1* e *RcPR-4a.2*.

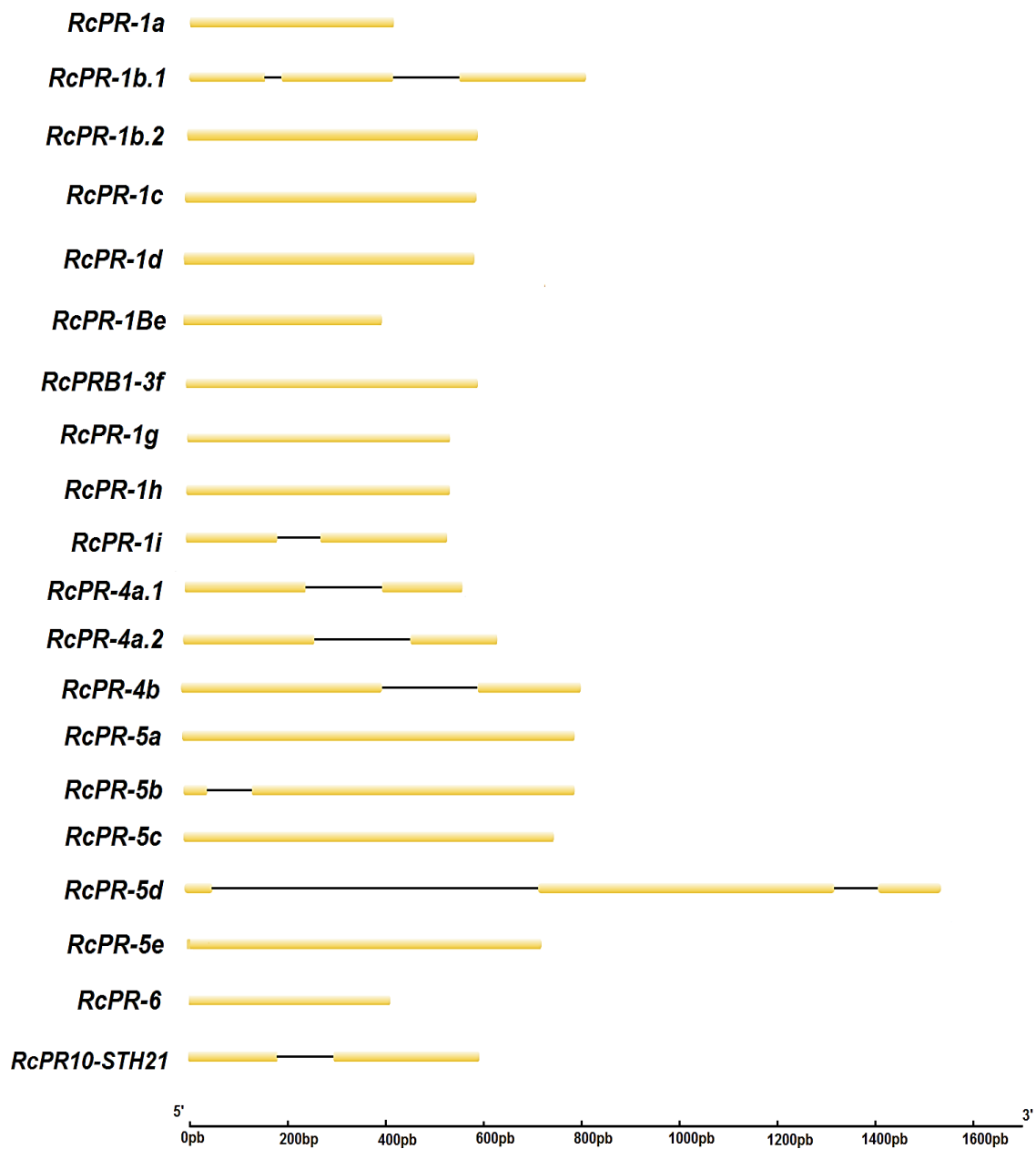
Dos 20 genes anotados no Banco de dados PHYTOZOME V 12.1 do genoma de *R. communis*, 10 pertencem a família PR-1 (*RcPR-1a*, *RcPR-1b.1*, *RcPR-1b.2*, *RcPR-1c*, *RcPR-1d*, *RcPR-1Be*, *RcPRB1-3f*, *RcPR-1g*, *RcPR-1h* e *RcPR-1i*); três pertencem a família PR-4 (*RcPR-4a.1*, *RcPR-4a.2*, *RcPR-4b*); cinco pertencem a família PR-5 (*RcPR-5a*, *RcPR-5b*, *RcPR-5c*, *RcPR-5d* e *RcPR-5e*); um pertence à família PR-6 (*RcPR-6*) e um à família PR-10 STH21 (*RcPR10-STH21*). Estes genes apresentaram número de éxons variável de 1 a

3 e íntrons de 1 a 2 (Figura 1), enquanto o painel de leitura aberto (ORF) variou de 222 a 795 pb (Tabela 1).

Tabela 1. Características dos genes PRs presentes no genoma de *Ricinus communis*, de acordo com PHYTOZOME V 12.1.

Nome do gene	Família Gênica	Gene ID	Localização (“Scaffold”)	ORF	Quantidade Éxons	Quantidade Íntrons
<i>RcPR-1a</i>	PR-1	30099.t000028	30099:134094..134579	486	1	0
<i>RcPR-1b.1</i>	PR-1	30099.t000029	30099:150206..150899	312	3	2
<i>RcPR-1b.2</i>	PR-1	30099.t000030	30099:151756..152238	483	1	0
<i>RcPR-1c</i>	PR-1	30099.t000031	30099:154297..155045	486	1	0
<i>RcPR-1d</i>	PR-1	30099.t000032	30099:161783..162283	501	1	0
<i>RcPR-1Be</i>	PR-1B	30099.t000036	30099:181211..181699	489	1	0
<i>RcPRB1-3f</i>	PRB1-3	27894.t000032	27894:426397..426912	516	1	0
<i>RcPR-1g</i>	PR-1	27766.t000007	27766:99107..99703	597	1	0
<i>RcPR-1h</i>	PR-1	27766.t000008	27766:111436..111975	540	1	0
<i>RcPR-1i</i>	PR-1	30099.t000027	30099:131507..131998	492	2	1
<i>RcPR-4a.1</i>	PR-4	29598.t000016	29598:67883..68473	300	2	1
<i>RcPR-4a.2</i>	PR-4	29598.t000017	29598:69822..70438	348	2	1
<i>RcPR-4b</i>	PR-4	29598.t000018	29598:71374..72176	501	2	1
<i>RcPR-5a</i>	PR-5	30170.t000403	30170:2181957..2182730	774	1	0
<i>RcPR-5b</i>	PR-5	30152.t000023	30152:873257..874051	795	2	1
<i>RcPR-5c</i>	PR-5	29669.t000034	29669:212759..213502	744	1	0
<i>RcPR-5d</i>	PR-5	30116.t000002	30116:61671..63370	222	3	2
<i>RcPR-5e</i>	PR-5	29669.t000043	29669:282612..283444	726	1	0
<i>RcPR-6</i>	PR-6	30099.t000037	30099:186152..186646	495	1	0
<i>RcPR10-STH21</i>	PR10-STH21	29726.t000222	29726:1239960..1240556	267	2	1

Figura 1. Estrutura gênica das *PRs* em *Ricinus communis* L. As barras amarelas representam os éxons e os íntrons estão representados em linhas.



Análise das sequências promotoras dos genes *RcPRs*

De acordo com a Figura 2 e a Suplementar 5 os elementos *cis*-regulatórios TATA, CAAT-box e de capacidade de resposta à luz estão presentes em todos os 20 genes analisados. Em menores proporções, foram encontrados elementos *cis* que respondem a estresses abióticos, bióticos e mecânicos, como resposta a seca, a água, a baixa temperatura, a deficiência nutricional, patogênese e seu ciclo celular, ao cádmio, indução anaeróbica e a lesão. Outros elementos *cis* em resposta ao crescimento e desenvolvimento da planta como, ativação do meristema, metabólicos *Zein* e local de ligação da proteína ao DNA em região rica em AT, expressão do endosperma, células do mesófilo, a controle de aspectos diferentes relacionados à fisiologia celular, biossíntese de flavonoides e controle do ciclo circadiano. Além dos elementos *cis* que respondem aos fitormônios, como etileno, giberelina, ácido salicílico, a auxina, ao ácido abscísico e ao indutor de resistência metil jasmonato – MeJA.

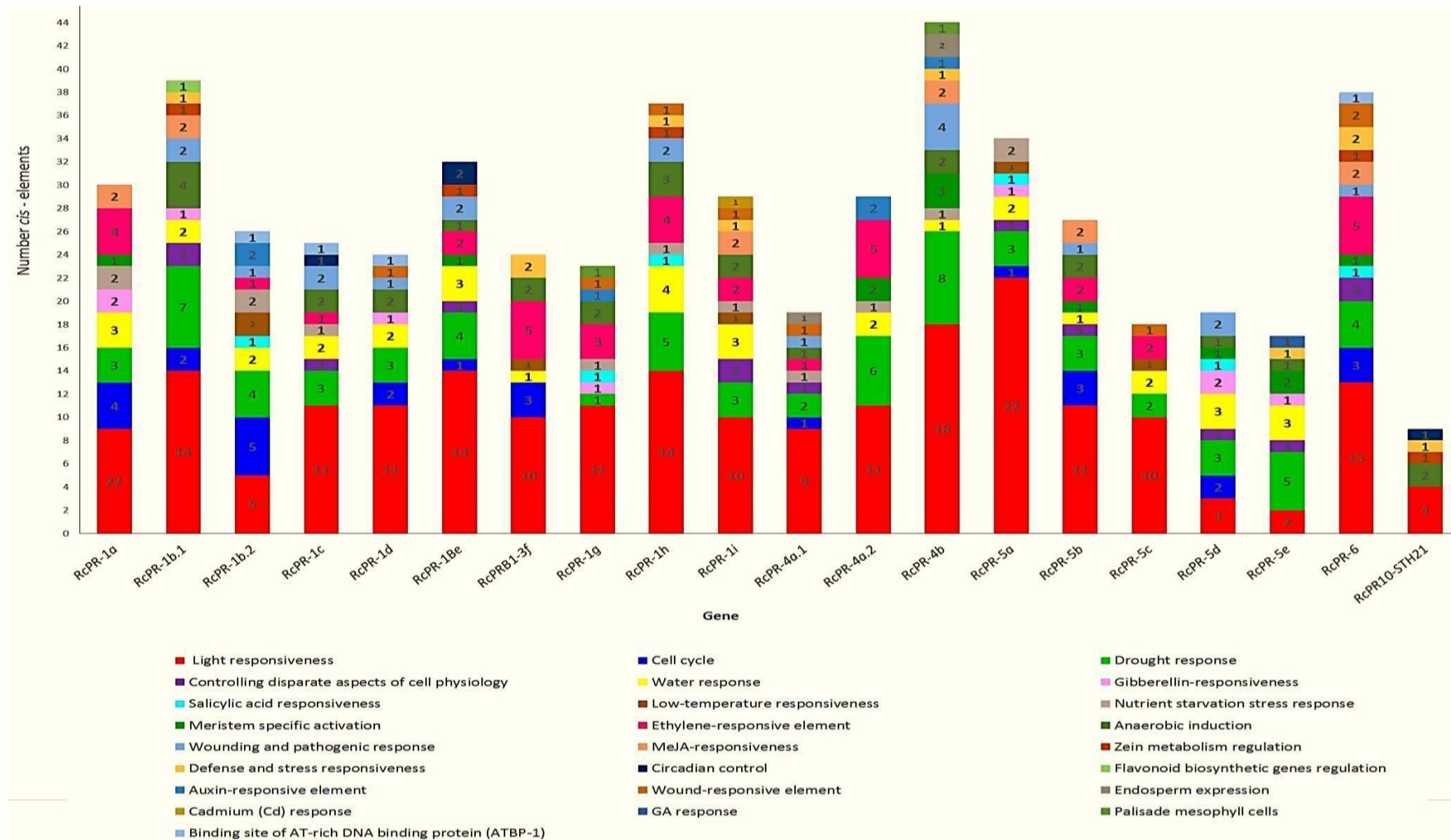


Figura 2. Elementos *cis* presentes na região promotora das Proteínas Relacionadas a Patogêneses em *Ricinus communis* L.

Análise das proteínas das RcPRs

O número de aminoácidos das proteínas *RcPRs* variou de 136 (RcPR-1i) a 290 (RcPR-5d) (Tabela 2). A predição do ponto isoelétrico (pI) das 17 sequências proteicas com peptídeo sinal (PS) variou de 4,68 (RcPR10- STH21) a 9,17 (RcPRB1-3f); e o peso molecular (MW) de 14,96 kDa (RcPR-1i) a 31,26 kDa (RcPR-5d). Sem o PS, o pI variou de 4,58 (RcPR-1d) a 9,35 (RcPRB1-3f); e o MW de 12,67 (RcPR-1i) a 28,83 kDa (RcPR-5d). O tamanho do PS variou de 20 a 26 resíduos de aminoácidos, com exceção das proteínas RcPR-1b.1; RcPR-5a e RcPR10- STH21 que não apresentaram o PS. O Índice de hidropaticidade média - GRAVY (do inglês *Grand average of hydropathicity*) foi positivo (+) indicando regiões hidrofóbicas (apolar) nas proteínas RcPR-5a e RcPR-5e(0,044 e 0,226), nas demais foram identificados uma hidropaticidade média negativa (-), regiões hidrofílicas (polar), variando de -0,451 (RcPRB1-3f) a -0,07(RcPR10- STH21) (Tabela 2).

Um total de 42,85%, das proteínas foram localizadas no espaço extracelular (55,48% a 98,64% de probabilidade) (Tabela 2), 38% nos lisossomos (99,99% de probabilidade), uma na membrana celular (85,81% de probabilidade), uma no citoplasma (68,14% de probabilidade) e uma no cloroplasto (27,40% de probabilidade) (Tabela 2).

Tabela 2. Características das Proteínas Relacionadas a Patogêneses presentes no genoma de *Ricinus communis*. pI: ponto isoelétrico; Mw: peso molecular em QuiloDalton (kDa); SP: peptídeo sinal; aa: resíduos de aminoácidos e GRAVY: grau médio de hidropaticidade (Continuação).

Proteína	ID Proteína	Tamanho Proteína	pI com/sem SP	Mw com/sem SP	Probabilidade de exportação (%)	Tamanho SP (aa)	GRAVY
RcPR-1a	30099.m001649	161	5,60/5,67	17,94/15,45	Lisossomo/Vacúolo 63,69	1-23	-0,365
RcPR-1b.1	30099.m001650	183	5,46	20,94	Cloroplasto 27,40	-	-0,093
RcPR-1b.2	30099.m001651	160	6,80/6,34	17,80/15,33	Lisossomo/Vacúolo 62,23	1-23	-0,366
RcPR-1c	30099.m001652	161	8,97/8,80	17,62/15,12	Lisossomo/Vacúolo 66,81	1-24	-0,231
RcPR-1d	30099.m001653	166	4,73/4,58	18,26/15,35	Lisossomo/Vacúolo 76,15	1-26	-0,249

Tabela 2. Características das Proteínas Relacionadas a Patogêneses presentes no genoma de *Ricinus communis*. pI: ponto isoelétrico; Mw: peso molecular em QuiloDalton (kDa); SP: peptídeo sinal; aa: resíduos de aminoácidos e GRAVY: grau médio de hidropaticidade (Conclusão).

Proteína	ID Proteína	Tamanho Proteína	pI com/sem SP	Mw com/sem SP	Probabilidade de exportação (%)	Tamanho SP (aa)	GRAVY
RcPR-1Be	30099.m001657	162	5,69/5,53	17,81/15,23	Lisossomo/Vacúolo 66,52	1-24	-0,446
RcPRB1-3f	27894.m000804	171	9,17/9,35	19,22/16,80	Lisossomo/Vacúolo 58,45	1-24	-0,451
RcPR-1g	27766.m000157	198	6,08/5,64	22,83/19,81	Extracelular 62,49	1-26	-0,429
RcPR-1h	27766.m000158	179	8,79/8,54	20,47/17,90	Extracelular 56,51	1-22	-0,45
RcPR-1i	30099.m001648	136	8,68/8,82	14,96/12,67	Lisossomo/Vacúolo 58,20	1-23	-0,347
RcPR-4a.1	29598.m000457	142	9,01/8,83	15,69/13,43	Extracelular 97,38	1-21	-0,111
RcPR-4a.2	29598.m000458	142	8,15/6,97	15,43/13,17	Extracelular 98,64	1-21	-0,013
RcPR-4b	29598.m000459	202	6,24/6,97	21,79/19,69	Extracelular 95,44	1-21	-0,306
RcPR-5a	30170.m013991	257	4,71	26,5	Extracelular 55,48	-	0,044
RcPR-5b	30152.m002402	234	9,02	25,45/23,21	Extracelular 95,46	1-20	-0,172
RcPR-5c	29669.m000825	247	5,17/5,14	26,92/24,68	Extracelular 91,32	1-20	-0,038
RcPR-5d	30116.m000371	290	5,89/5,90	31,26/28,83	Membrana celular 85,81	1-22	-0,067
RcPR-5e	29669.m000834	241	4,89/5,04	25,78/23,43	Extracelular 93,68	1-22	0,226
RcPR-6	30099.m001658	164	8,15/7,76	17,83/15,13	Lisossomo/Vacúolo 72,53	1-26	-0,191
RcPR10-STH21	29726.m004110	160	4,68	17,65	Citoplasma 68,14	-	-0,07

A proteína RcPR-5d apresentou 31 sítios de fosforilação (7Thre/19Ser/5Tyr); seguido da RcPR-5a com 28 sítios de fosforilação (9Thre/16Ser/3Tyr); a RcPR-5c com 26 sítios de fosforilação (7Thre/14Ser/5Tyr); a RcPR-5b com 25 sítios de fosforilação (14Thre/5Ser/6Tyr); a RcPR-1g com 24 sítios de fosforilação (9Thre/11Ser/4Tyr); as proteínas RcPRB1-3f e RcPR-5e com 23 sítios de fosforilação (7Thre/12Ser/4Tyr, 7Thre/15Ser/1Tyr, respectivamente); a RcPR-6 com 21 sítios de fosforilação (4Thre/11Ser/6Tyr); a RcPR-1b.1 com 20 sítios de fosforilação (3Thre/17Ser); as proteínas RcPR-1a e RcPR-4b com 18 sítios de fosforilação (5Thre/6Ser/7Tyr, 8Thre/7Ser/3Tyr, respectivamente); a RcPR-1d com 16 sítios de fosforilação (3Thre/6Ser/7Tyr); as

RcPR-1b.2 e RcPR-1h com 15 sítios de fosforilação (5Thre/4Ser/6Tyr, 5Thre/7Ser/3Tyr, respectivamente); a RcPR-1Be com 14 sítios de fosforilação (3Thre/7Ser/4Tyr); a RcPR-1c com 13 sítios de fosforilação (1Thre/9Ser/3Tyr); a RcPR-1i com 12 sítios de fosforilação (1Thre/7Ser/4Tyr); as RcPR10- STH21 e RcPR-4a.1 com 10 sítios de fosforilação (5Thre/4Ser/1Tyr, 7Thre/3Ser, respectivamente); e a RcPR-4a.2 com 9 sítios de fosforilação (7Thre/2Ser). Apenas as proteínas RcPR-5d; RcPR-5e; RcPR-5a; RcPR-1d; RcPR10- STH21; RcPR-1a; RcPR-6; RcPR-5c; RcPR-1ge RcPR-1iapresentaram sítios de N-glicosilação, sendo 10, 4, 2, 2, 2, 1, 1, 1, 1, 1 e 1 respectivamente (Tabela 3).

Tabela 3. Modificações Pós-traducionais em *Ricinus communis* L. das Proteínas Relacionadas a Patogêneses (Continuação).

Proteína	ID Proteína	Sítios de fosforilação	Sítios de N-glicosilação
RcPR-1a	30099.m001649	T ₁₀ ; T ₁₅ ; S ₂₆ ; Y ₃₀ ; T ₃₆ ; T ₅₈ ; Y ₅₉ ; S ₆₂ ; S ₆₆ ; Y ₇₃ ; Y ₈₄ ; Y ₁₀₃ ; Y ₁₀₅ ; S ₁₀₇ ; S ₁₀₉ ; Y ₁₂₀ ; T ₁₂₁ ; S ₁₂₈	N ₄₉
RcPR-1b.1	30099.m001650	S ₁₉ ; S ₂₄ ; T ₄ ; S ₆₀ ; S ₆₂ ; S ₆₇ ; S ₇₆ ; S ₈₀ ; S ₉₃ ; S ₁₀₁ ; T ₁₁₇ ; S ₁₂₁ ; S ₁₂₃ ; S ₁₃₀ ; T ₁₄₂ ; S ₁₄₇ ; S ₁₅₅ ; S ₁₅₈ ; S ₁₇₁ ; S ₁₇₇	-
RcPR-1b.2	30099.m001651	T ₁₈ ; S ₂₁ ; S ₂₆ ; Y ₃₀ ; Y ₅₉ ; T ₆₂ ; Y ₇₆ ; S ₈₉ ; Y ₁₀₄ ; Y ₁₀₆ ; S ₁₀₈ ; T ₁₂₇ ; T ₁₃₅ ; T ₁₄₂ ; Y ₁₄₃	-
RcPR-1c	30099.m001652	S ₃ ; S ₇ ; S ₁₁ ; S ₂₂ ; T ₄₈ ; S ₇₃ ; Y ₇₇ ; S ₈₄ ; S ₈₅ ; S ₈₉ ; Y ₁₀₄ ; Y ₁₂₀ ; S ₁₂₈	-
RcPR-1d	30099.m001653	S ₇ ; S ₁₇ ; Y ₃₃ ; T ₅₂ ; Y ₆₄ ; S ₇₇ ; Y ₈₁ ; S ₉₃ ; Y ₁₀₈ ; Y ₁₁₀ ; T ₁₁₆ ; Y ₁₂₄ ; S ₁₃₂ ; S ₁₃₄ ; T ₁₄₈ ; Y ₁₅₅	N ₅₀ ; N ₅₄
RcPR-1Be	30099.m001657	S ₇ ; T ₂₇ ; Y ₃₁ ; T ₄₈ ; S ₅₅ ; S ₇₃ ; Y ₇₇ ; S ₈₄ ; S ₈₉ ; T ₉₁ ; Y ₁₀₃ ; Y ₁₀₄ ; S ₁₂₈ ; S ₁₃₇	-
RcPRB1-3f	27894.m000804	S ₃ ; T ₉ ; S ₁₈ ; S ₂₀ ; S ₂₃ ; S ₂₄ ; S ₂₈ ; T ₃₂ ; T ₃₆ ; S ₅₉ ; S ₆₄ ; S ₆₈ ; Y ₈₆ ; S ₉₄ ; Y ₁₁₃ ; Y ₁₁₄ ; T ₁₂₀ ; T ₁₂₂ ; Y ₁₃₀ ; T ₁₃₁ ; S ₁₃₈ ; T ₁₄₉ ; S ₁₅₀	-
RcPR-1g	27766.m000157	S ₅ ; T ₂₅ ; S ₂₆ ; S ₂₈ ; S ₃₀ ; T ₃₅ ; T ₄₄ ; Y ₄₆ ; S ₄₉ ; S ₇₃ ; S ₈₃ ; T ₉₅ ; S ₁₀₅ ; Y ₁₁₈ ; S ₁₂₁ ; T ₁₂₆ ; S ₁₃₂ ; Y ₁₄₀ ; Y ₁₄₁ ; S ₁₄₇ ; T ₁₅₈ ; T ₁₆₄ ; T ₁₆₅ ; T ₁₈₃	N ₄₂

Tabela 3. Modificações Pós-traducionais em *Ricinus communis* L. das Proteínas Relacionadas a Patogêneses (Conclusão).

Proteína	ID Proteína	Sítios de fosforilação	Sítios de N-glicosilação
RcPR-1h	27766.m000158	S ₁₃ ; S ₁₅ ; T ₁₉ ; T ₂₃ ; S ₂₅ ; S ₄₂ ; T ₄₃ ; S ₈₉ ; Y ₉₃ ; S ₁₀₁ ; S ₁₁₆ ; Y ₁₂₉ ; Y ₁₃₇ ; T ₁₄₅ ; T ₁₅₄	-
RcPR-1i	30099.m001648	S ₁₈ ; S ₂₆ ; Y ₃₀ ; Y ₅₅ ; S ₅₉ ; S ₈₄ ; S ₈₆ ; Y ₉₅ ; T ₁₀₃ ; S ₁₀₄ ; Y ₁₂₅ ; S ₁₃₄	N ₄₉
RcPR-4a.1	29598.m000457	T ₆ ; S ₂₅ ; T ₃₀ ; S ₄₇ ; T ₅₂ ; T ₆₈ ; T ₉₀ ; T ₉₂ ; T ₉₇ ; S ₁₀₅	-
RcPR-4a.2	29598.m000458	T ₆ ; S ₂₅ ; T ₃₀ ; S ₄₇ ; T ₅₂ ; T ₆₈ ; T ₈₈ ; T ₉₂ ; T ₉₇	-
RcPR-4b	29598.m000459	S ₃ ; S ₄₀ ; Y ₄₂ ; T ₄₈ ; Y ₅₁ ; S ₅₃ ; S ₇₂ ; T ₇₇ ; T ₉₉ ; S ₁₀₇ ; Y ₁₁₂ ; T ₁₅₅ ; T ₁₂₇ ; T ₁₃₆ ; T ₁₄₄ ; T ₁₄₅ ; S ₁₅₃ ; S ₁₉₇	-
RcPR-5a	30170.m013991	T ₁₇ ; T ₄₉ ; S ₅₁ ; T ₅₇ ; T ₁₁₁ ; S ₁₁₉ ; S ₁₂₈ ; S ₁₂₉ ; Y ₁₃₅ ; S ₁₅₅ ; S ₁₈₀ ; S ₁₈₂ ; S ₁₈₄ ; S ₂₀₇ ; T ₂₀₈ ; T ₂₁₁ ; S ₂₁₃ ; T ₂₁₅ ; S ₂₁₈ ; S ₂₂₇ ; S ₂₃₀ ; Y ₂₃₃ ; S ₂₃₇ ; S ₂₃₈ ; T ₂₃₉ ; S ₂₄₃ ; Y ₂₄₇ ; T ₂₅₀	N ₂₈ ; N ₃₁
RcPR-5b	30152.m002402	T ₉ ; T ₃₂ ; T ₃₇ ; S ₃₉ ; T ₅₀ ; T ₇₄ ; T ₈₁ ; S ₁₀₂ ; T ₁₀₇ ; T ₁₁₂ ; S ₁₃₉ ; Y ₁₇₇ ; T ₁₇₉ ; Y ₁₈₂ ; Y ₁₈₈ ; S ₁₉₀ ; T ₁₉₃ ; T ₁₉₇ ; S ₂₀₀ ; T ₂₀₉ ; Y ₂₁₀ ; Y ₂₁₁ ; Y ₂₁₅ ; T ₂₂₁ ; T ₂₂₃	-
RcPR-5c	29669.m000825	T ₈ ; S ₁₅ ; T ₂₃ ; T ₄₃ ; Y ₄₄ ; Y ₈₀ ; Y ₈₅ ; T ₉₉ ; S ₁₀₅ ; S ₁₀₆ ; S ₁₁₇ ; Y ₁₂₂ ; S ₁₂₇ ; S ₁₃₇ ; S ₁₄₃ ; Y ₁₆₉ ; S ₁₉₀ ; S ₁₉₃ ; S ₁₉₈ ; S ₂₀₀ ; S ₂₀₂ ; S ₂₀₃ ; S ₂₀₅ ; T ₂₁₁ ; T ₂₁₉ ; T ₂₃₂	N ₆₁
RcPR-5d	30116.m000371	S ₆ ; S ₂₇ ; T ₂₈ ; S ₅₂ ; Y ₆₈ ; S ₉₇ ; Y ₁₀₁ ; S ₁₀₅ ; S ₁₀₉ ; T ₁₁₃ ; T ₁₁₈ ; S ₁₂₈ ; S ₁₄₀ ; S ₁₄₉ ; S ₁₆₆ ; S ₁₇₈ ; T ₁₈₅ ; Y ₁₈₈ ; Y ₁₉₄ ; S ₁₉₅ ; S ₂₀₆ ; S ₂₁₈ ; T ₂₂₆ ; S ₂₂₇ ; S ₂₃₂ ; Y ₂₃₆ ; T ₂₄₇ ; S ₂₅₀ ; S ₂₆₁ ; S ₂₆₅ ; T ₂₈₈	N ₄ ; N ₂₄ ; N ₅₀ ; N ₈₈ ; N ₁₀₃ ; N ₁₄₇
RcPR-5e	29669.m000834	T ₁₁ ; S ₁₄ ; T ₁₈ ; T ₂₆ ; S ₄₅ ; S ₄₈ ; S ₆₂ ; S ₈₅ ; T ₉₉ ; S ₁₁₈ ; Y ₁₂₇ ; S ₁₄₂ ; S ₁₄₈ ; S ₁₅₄ ; S ₁₅₅ ; S ₁₅₈ ; T ₁₉₉ ; T ₂₀₁ ; T ₂₀₆ ; S ₂₂₅ ; S ₂₂₇ ; S ₂₃₂ ; S ₂₃₄	N ₃₂ ; N ₆₄ ; N ₁₉₇ ; N ₂₀₇
RcPR-6	30099.m001658	S ₃ ; S ₇ ; S ₁₂ ; S ₂₉ ; Y ₃₃ ; T ₅₇ ; Y ₆₂ ; S ₇₆ ; S ₇₈ ; Y ₇₉ ; S ₈₇ ; S ₈₈ ; S ₈₉ ; S ₉₀ ; T ₉₂ ; T ₉₄ ; Y ₁₀₆ ; Y ₁₀₇ ; Y ₁₀₉ ; T ₁₂₄ ; S ₁₂₉	N ₁₀₈
RcPR10-STH21	29726.m004110	T ₇ ; S ₁₂ ; S ₁₇ ; S ₄₁ ; T ₅₃ ; Y ₆₇ ; T ₈₅ ; S ₁₀₉ ; T ₁₂₃ ; T ₁₂₈	N ₅₈

Os domínios funcionais das proteínas são unidades modulares que as caracterizam com uma função bioquímica, eles foram específicos para cada grupo de PRs segundo a base de dados de famílias proteicas Pfam. Assim, as RcPR-1 apresentaram o domínio PF00188 referente Família de Proteínas Secretória Ricas em Cisteína e PF13456 referente Família de Enzimas da Ribonuclease (RNase H); as RcPR-4 com o domínio PF00967 característico da Família *Barwin*; as RcPR-5 com o domínio PF00314 referente a Família Thaumatina; a RcPR-6 apresentou o domínio PF00188 também referente a Família de Proteínas Secretória Ricas em Cisteínas; a RcPR10-STH21 com o domínio PF00407 característico de membros da família PR-10 (Tabela 4).

Tabela 4. Características dos Domínios Proteicos conservados em *Ricinus communis* L. nas Proteínas Relacionadas a Patogênese. aa: resíduos de aminoácidos (Continuação).

Proteína	ID Proteína	Domínio Funcional (PFAM)	Intervalo conservado (aa)	Domínio Proteico conservado Família
RcPR-1a	30099.m001649	PF00188	31-149	Família de proteínas secretória rica em cisteína
RcPR-1b.1	30099.m001650	PF13456	133-172	Família de Enzimas da Ribonuclease (RNase H)
RcPR-1b.2	30099.m001651	PF00188	32-148	Família de proteínas secretória rica em cisteína
RcPR-1c	30099.m001652	PF00188	33-149	Família de proteínas secretória rica em cisteína
RcPR-1d	30099.m001653	PF00188	35-154	Família de proteínas secretória rica em cisteína
RcPR-1Be	30099.m001657	PF00188	32-147	Família de proteínas secretória rica em cisteína
RcPRB1-3f	27894.m000804	PF00188	42-159	Família de proteínas secretória rica em cisteína
RcPR-1g	27766.m000157	PF00188	64-186	Família de proteínas secretória rica em cisteína
RcPR-1h	27766.m000158	PF00188	50-167	Família de proteínas secretória rica em cisteína
RcPR-1i	30099.m001648	PF00188	62-124	Família de proteínas secretória rica em cisteína
RcPR-4a.1	29598.m000457	PF00967	24-138	Domínio <i>Barwin</i>
RcPR-4a.2	29598.m000458	PF00967	24-138	Domínio <i>Barwin</i>
RcPR-4b	29598.m000459	PF00967	68-186	Domínio <i>Barwin</i>

Tabela 4. Características dos Domínios Proteicos conservados em *Ricinus communis* L. nas Proteínas Relacionadas a Patogênese. aa: resíduos de aminoácidos (Conclusão).

Proteína	ID Proteína	Domínio Funcional (PFAM)	Intervalo conservado (aa)	Domínio Proteico conservado Família
RcPR-5a	30170.m013991	PF00314	39-253	Família Taumatina
RcPR-5b	30152.m002402	PF00314	27-234	Família Taumatina
RcPR-5c	29669.m000825	PF00314	26-247	Família Taumatina
RcPR-5d	30116.m000371	PF00314	35-241	Família Taumatina
RcPR-5e	29669.m000834	PF00314	29-241	Família Taumatina
RcPR-6	30099.m001658	PF00188	34-152	Família de proteínas secretória rica em cisteína
RcPR10-STH21	29726.m004110	PF00407	1-153	Família Bet v 1, típica em PR-10

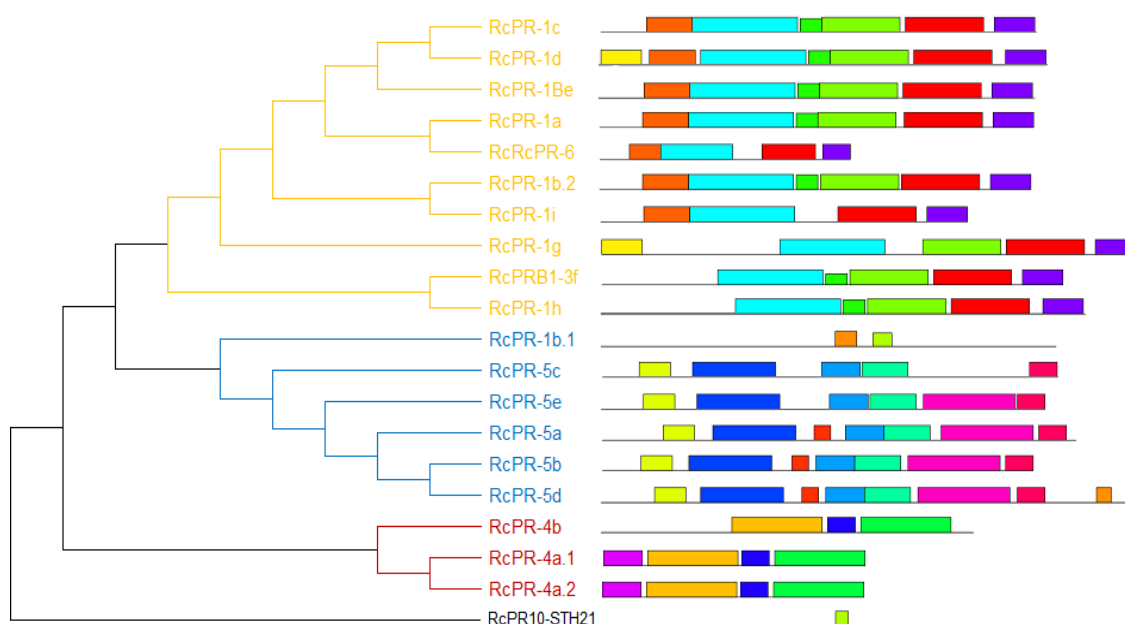
A identidade das sequências proteicas foi de 0% a 92,25%. (Tabela Suplementar 4), no qual o maior grau de identidade observado foi entre proteínas da mesma família, como observa-se na RcPR-5a e RcPR-5b (92,25%), seguidos da RcPR-1c e RcPR-1Be (69,57%), RcPR-4a.2 com a RcPR-4a.1 e RcPR-4b (67,61%), RcPR-1d e RcPR-1Be (67,70%), RcPR-1d e RcPR-1c (67,08%), RcPR-1b.2 e RcPR-1i (55,15%), RcPR-1a com RcPR-1b.2 e RcPR-1c (54,43%), RcPR-5a e RcPR-5b (53,88%), RcPR-1b.2 e RcPR-1Be (52,23%), RcPR-1b.2 e RcPR-1d (51,59%), RcPR-1g e RcPR-1h (51,12%), com exceção da RcPR-6 e RcPR-1c (66,46%), RcPR-6 e RcPR-1Be (58,39%) e da RcPR-1a com a RcPR-6 (56,25%), por apresentarem o mesmo domínio proteico funcional (Tabela 4). Ademais, não foi observado similaridade entre proteínas de diferentes famílias, (Figura Suplementar 4). As demais apresentaram graus de identidade inferior a 50%.

A análise no programa *Multiple Em for Motivo Elicitation- MEME Suite* identificou 20 sequências conservadas e dentre essas, três motivos proteicos com funções biológicas já descritas na literatura em proteínas-PRs.

Os motivos CNYx e PxGNxxxxxPY foram identificadas na mesma sequência conservada nas proteínas RcPR-1a, RcPR-1b.2, RcPR-1c, RcPR-1d, RcPR-1Be, RcPR-1-3f, RcPR-1g, RcPR-1h, RcPR-1i e na RcPR-6.

Além destas, foi possível identificar também o motivo GxGGxG, na sua forma reduzida GxGG, na proteína RcPR10- STH21 (Figura 3 e Figura Suplementar 7).

Figura 3: Árvore filogenética das 20 proteínas RcPRs e seus respectivos motivos proteicos conservados. Esta árvore foi construída a partir do método *Neighbor-Joining*, usando o modelo JTT. Os motivos conservados totalizaram 20 e foram identificados no servidor *MEME Suits*. O último motivo, de cor lilás, nas nove RcPR-1 e na RcPR-6, identificadas com as ramificações de cor amarela, contém as sequências CNYx e PxGNxxxxxPY e o menor motivo, de cor verde claro, presente na RCPR10-STH21 contém a sequência GxGG. Os demais motivos conservados são relacionados as funções de cada uma das cinco famílias (PR-1; PR-4; PR-5; PR-6 e PR-10) identificadas no genoma de *R. communis* L.



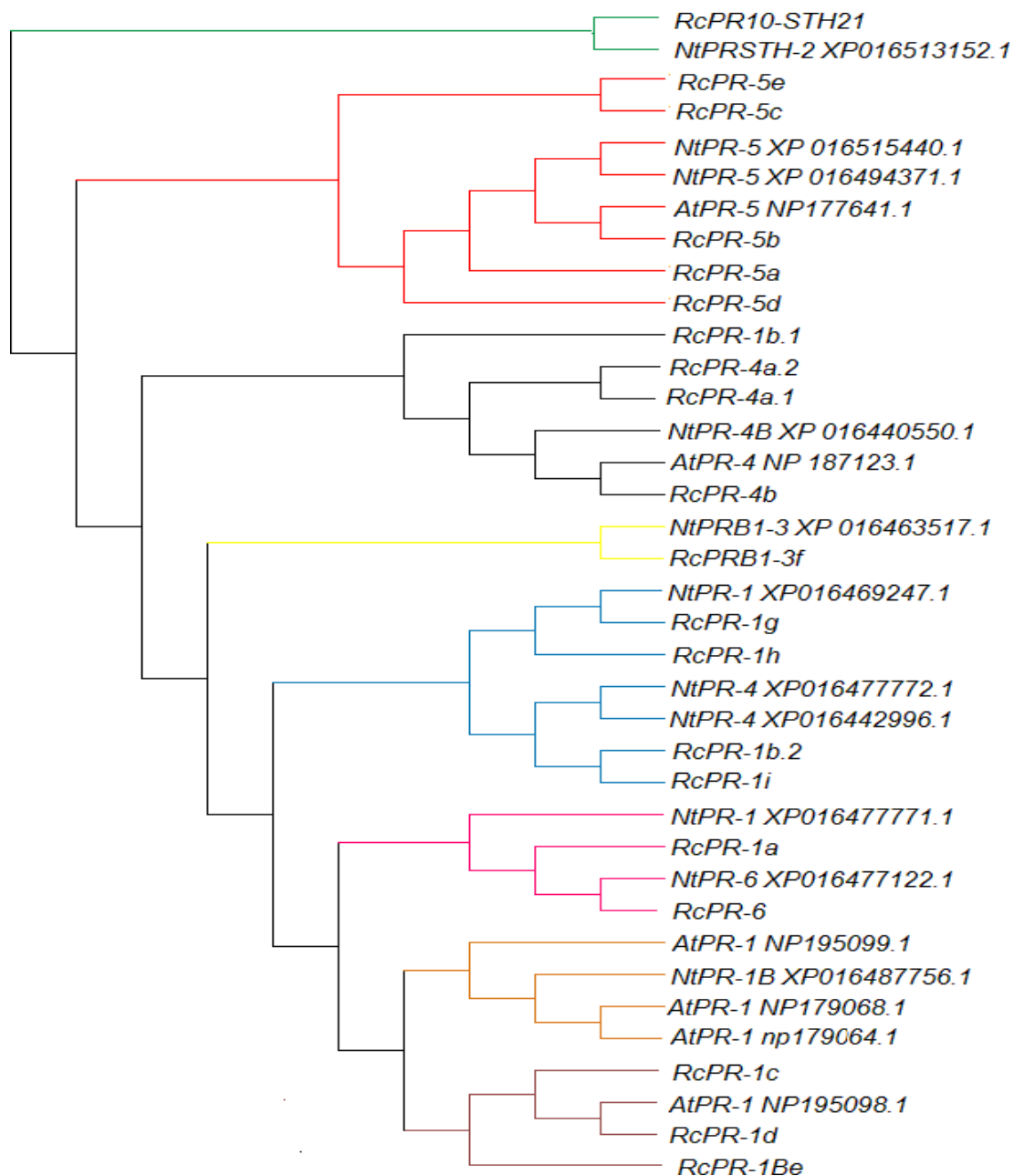
Relação filogenética

Uma outra árvore filogenética foi construída a partir das sequências proteicas RcPRs com as proteínas-PR de *Nicotiniana tabacum* L. (NtPR) e *Arabidopsis thaliana* L. (AtPR) (Figura 4). De acordo com a análise filogenética, as Proteínas-PR das três espécies analisadas, agruparam-se majoritariamente por famílias, em um total de oito grupos.

Destes, cinco agrupamentos foram formados por membros da família PR-1, no entanto, três desses grupos, alocaram também membros pertencentes as famílias PR-4 e PR-6. Além destes, um outro grupo foi formado por membros da PR-4 que também alocou um único membro pertencente à família das PR-1.

Já os membros da família das PR-5 uniram-se formando apenas um grupo, semelhante aos membros da família das PR-10 que também formaram apenas um grupo. Apesar da filogenia ter mostrado a presença de agrupamentos, pode-se observar que todos os membros estavam estreitamente próximos aos grupos constituídos. Nitidamente, os maiores grupos foram formados por membros pertencentes a família PR-1. De outro modo, o menor agrupamento foi formado pelos membros pertencentes a família das PR-10.

Figura 4. Árvore filogenética das Proteínas Relacionadas a Patogênese presentes em *Ricinus communis*, *Nicotiniana tabacum* e *Arabidopsis thaliana*. Análise foi realizada a partir de sequências proteicas pelo método *Neighbor-Joining*, inferida a partir de 1000 replicatas e usando o modelo JTT.



DISCUSSÃO

Família de genes PRs no genoma de *Ricinus communis* L.

A variação no número de genes PRs e sua distribuição em diferentes famílias gênicas é comumente observada entre as espécies.

Apesar da anotação científica de 19 famílias gênicas precursoras das proteínas-PR, como observado, essas famílias não necessariamente, estão presentes nos genomas vegetais em sua totalidade. Esse fenômeno acontece, possivelmente, devido as diferentes estratégias que cada espécie desenvolve ao longo do processo evolutivo para lidar com diferentes tipos de agressões sofridas pelo ataque de fitopatógenos, isto é, são ativadas apenas as famílias gênicas com funções biológicas correspondentes a cada agressor oportunista em cada espécie.

Além disso, o número de membros pertencentes a uma mesma família gênica é muito variável entre as espécies, como já dito. No genoma de *R. communis*, metade dos genes PRs, pertencem a família gênica das PR-1, muito provavelmente porque esta família é considerada a mais extensa entre as demais, e produzida em maior número nas plantas durante o ataque de fitopatógenos biotróficos na defesa intracelular mediante ao acúmulo do AS, pois a sua expressão tem sido usada há muito tempo como marcador de resistência a doenças mediadas por esse ácido que conduz a RAS (VAN LOON et al., 1994; VAN LOON; VAN STRIEN, 1999; BREEN et al., 2017).

Presumivelmente, os fitopatógenos de importância econômica em *Ricinus communis* são biotróficos e/ou hemibiotróficos, justificando assim, a grande maioria dos genes *RcPRs* serem pertencentes a família das PR-1.

Ademais, doze dos 20 genes analisados não apresentaram íntrons, o que tipicamente, indica que eles podem ser transcritos de maneira mais rápida que as demais sequências com regiões intrônicas, visto que a presença de íntrons pode resultar em atrasos na produção de produtos gênicos devido às etapas de processamento do transcrito (JEFFARES et al., 2008).

Contudo, estas regiões são de suma importância, pois abrigam os RNAs não traduzidos, bem como os elementos de controle de *splicing*, que regulam o processamento do RNA mensageiro (mRNA) e permitem o *splicing* alternativo, que leva a uma maior variabilidade de produtos gênicos, as proteínas (ZHU et al., 2009; WOODS et al., 2016).

Atualmente, estudos sobre a evolução dos íntrons descobriram que a sua perda é mais provável do que o ganho, durante o processo evolutivo (ROY et al., 2006; SHI; DU et al., 2020). Com base nisto, pode-se supor que a maioria dos genes *RcPRs* com regiões intrônicas sejam mais antigos que os genes que não apresentam essas regiões, pois ao longo do processo evolutivos essas regiões foram perdidas (SHI; DU et al., 2020).

Ademais, a presença do número de éxons e íntrons em genes *PRs* mostra-se ser pouco variável. De acordo com trabalhos anteriores, o número de éxons, geralmente, é variável de 1 a 2, e na presença de dois éxons a separação é feita por apenas uma região intrônica (MENEZES et al., 2014; JIANG et al., 2015; LU et al., 2018). Esta mesma variação também foi encontrada na maioria dos genes *RcPRs* analisados. No entanto, em apenas dois genes *RcPRs* foram identificados três éxons, separados por dois íntrons, sendo esses maiores que os demais genes *RcPRs* analisados (Figura 1).

Por meio disso, pode-se concluir que a ausência das regiões intrônicas na maioria dos genes *RcPRs* impossibilita a diversidade de transcritos a partir do mesmo gene ou até mesmo atuando como regulador de expressão gênica (micro-RNAs).

Análise da região promotora dos genes *RcPRs*.

A análise dos elementos *cis*-reguladores nas sequências promotoras de 1500pb, no montante do local de início da tradução de cada um dos 20 genes *RcPRs* analisados, permitiu identificar os elementos associados a várias respostas a estresses.

Ainda permitiu identificar elementos como o TATA e CAAT-box, dentre outros que evidenciam a presença de locais de acoplamento para fatores de transcrição, potenciadores e elementos repressores necessários para a expressão adequada dos genes *RcPRs* (WITTKOPP et al., 2012).

Hoje, sabe-se que numerosos genes em plantas são fortemente ativados em nível transcricional após o ataque de fitopatógenos (NASIR et al., 2017).

Com base nisto, a presença de elementos em resposta aos fitormônios ácido salicílico (AS), etileno (ET), giberelina (AG), ácido abscísico (ABA), metil-jasmônico (MeJA) e auxinas nos genes *RcPRs*, são de suma importância, pois são tidos como importantes fatores internos de sinais bioquímicos que medeiam

reações envolvidas na regulação desde o desenvolvimento/crescimento e florescimento de plantas (IONESCU et al., 2017).

Além disso, podem ser regulados simultaneamente em resposta a estresse bióticos, não somente em ocasiões em que a planta precisa se adaptar a ambientes externos adversos, mas também atuam como moduladores em complexas cascatas de sinalização para respostas imunológicas na planta (GRANT; JONES et al., 2009). Isto é, a presença destes elementos na região promotora dos genes *RcPRs* evidenciam que eles estão envolvidos em resposta a diferentes estresses, bióticos e abióticos.

Ademais, a presença destes elementos corroboram com as atuações desempenhadas pelas proteínas precursoras dos genes *PRs* nos vegetais, pois são envolvidas diretamente no sistema imunológico inato das plantas, onde alguns desses genes são induzidas em grandes quantidades, levando a uma resposta local, Hipersensível, como em outras ocasiões respondem a Resistência Adquirida Sistêmica – RAS (KINKEMA et al., 2000).

Sabe-se que alguns genes *PRs* são induzidos, particularmente, pelo acúmulo dos fitormônios AS (*PR-1*; *PR-2* e *PR-5*) e AJ (*PR-3*; *PR-4* e *PR-12*), onde a composição, quantidade e tempo desses elementos versáteis diferem significativamente dentro das espécies de plantas e dependem do estilo de vida do patógeno (biotróficos, hemibiotróficos ou necrotróficos), e do seu modo de infecção (DE VOS et al., 2005; ALI et al., 2018). Normalmente, o acúmulo de AS é dado na presença de fitopatógenos biotróficos, enquanto o acúmulo de AJ, como também do ET, medeiam a resistência a fitopatógenos necrotróficos, e a pragas herbívoras (BARI; JONE, 2009).

De forma geral, tanto o acúmulo de AS e AJ como do ET, respondem como sinais secundários após o ataque de fitopatógenos em resposta a RAS, produzindo uma barreira química em tecidos distantes do sítio primário da infecção (CHO; LEE, 2015; LIU et al., 2016). No entanto, ainda podem estar envolvidos em vias metabólicas primárias (CHO; LEE, 2015). O acúmulo tanto de AS como de AJ, também podem ativar genes através de vias seletivas, dependentes de fatores de transcrição, porém, normalmente, o acúmulo de AS, ativa a RAS de forma direta (ALI et al., 2018).

Zhang et al. (2013), ao analisarem o transcriptoma de *Gossypium barbadense* cv. Pima90-53, uma variedade de algodão inoculado com o fungo de solo, *Verticillium dahliae*, causador de uma doença vascular altamente

destrutiva na cultura, identificaram um total de 270 fatores de transcrição do tipo *ERF/EREBP*, elementos de ação cis associados ao ET. Os resultados mostraram que o ET regulou componentes que levaram a expressão de moléculas de sinalização em resposta de defesa. Além disso, a via do ET pode ser ativa tanto durante o desenvolvimento das plantas ou em resposta a demais estímulos, sejam bióticos ou abióticos (GRANT; JONES et al., 2009).

Além destes clássicos fitormônios, a presença do ABA, AG e de auxinas, também são conhecidas por modular o sistema inume das plantas, porém pouco se sabe sobre o modo de atuação (PIETERSE et al., 2012; OIDE et al., 2013).

Além dos estresses bióticos, os abióticos constituem umas das principais ameaças à agricultura, pois além de causarem perdas na produtividade, servem como pontos de entrada para diversos fitopatógenos oportunistas nos vegetais (ALI et al., 2018). Ainda segundo os mesmos autores, as mudanças climáticas ameaçam os sistemas de cultivo, favorecendo o aumento de cepas virulentas e de ampla gama de hospedeiros.

Por isso, estudar os mecanismos moleculares que envolvem os estresses bióticos, abióticos e mecânicos são de suma importância para o desenvolvimento de genótipos resistente e/ou tolerantes a esses múltiplos estresses. E sabe-se que os genes *PR* são ativos nestes diferentes tipos de estresses, apesar de não ser totalmente compreendido o mecanismo de atuação destes genes nos estresses abióticos e mecânicos.

Fundamentado a isto, estudos anteriores mostram que estresses, salínico e hídrico aumentam significativamente a expressão de genes *PR* em *Arabidopsis thaliana* e em plantas transgênicas no arroz (*Oryza sativa* L.) (SINGH et al., 2013; WU et al., 2016).

Também já foi comprovado que as famílias gênicas precursoras das PR-2 e PR-3 protegem os danos celulares devido ao estresse causados pelo frio e apresentam atividade anticongelante (JANSKA et al., 2010), e que há aumento significativos da expressão dos genes precursores das PR-4 após estresses causados pelo frio e salinidade, e ainda aqueles causados por feridas (KIM et al., 2014).

A presença destes diversos elementos *cis*-reguladores na região promotora dos genes *RcPRs*, só ressaltam o quanto as atividades biológicas relacionadas desde o crescimento e desenvolvimento e ataque de fitopatógenos estão estreitamente relacionados e que atuam de forma harmoniosa, e que esses

genes podem estar envolvidos em processos fisiológicos além da participação no mecanismo de defesa das plantas em *R. communis*.

Muitas PRs se acumulam em tecidos específicos, como flores, raízes, caules e sementes, em pontos bem definidos em seus ciclos de vida. Um exemplo claro são as grandes concentrações das PR-1 encontradas nas raízes e folhas de plantas jovens que são expressas constitutivamente em algumas espécies, na ausência de ataques por fitopatógenos (DAFOE et al., 2010).

Em *citrus*, as PRs com atividade quitinase e glucanase foram associadas ao desenvolvimento dos frutos (MCCOLLUM et al., 2006). Também foi relatado que algumas PRs do tipo Taumatina (PR-5), se acumulam nas secreções ovulares ou durante a senescência foliar (CAPELLI et al., 1997; O'LEARY et al., 2007). Além disso, as Taumatinas também foram descritas como proteínas associadas ao amadurecimento que se acumulam fortemente nos frutos durante o processo de maturação (TATTERSALL et al., 1997; KIM et al., 2002).

Com base nisto, fica evidente que os genes *PRs* além do papel notável na resistência, também estão envolvidos nos processos decrescimento e desenvolvimento dos vegetais (CAMPOS et al., 2007; HOU et al., 2012).

Em geral, estes elementos *cis*-regulatórios desempenham papéis essenciais na regulação gênica em diferentes tipos de células e isto sugere que os níveis de expressão de tais elementos podem ser alterados devido às variações nas condições em que se encontram as células em diferentes tecidos e órgãos.

Ainda segundo Kaur et al. (2017), identificar e compreender estes elementos também pode permitir alterar efetivamente o padrão de expressão de um gene de forma desejada.

Desta forma, o entendimento destes mais diversos tipos de elementos, bem como a sua identificação nos genes *RcPR* se faz uma abordagem poderosa, pois revela a multiplicidade funcional destes genes, fornece uma visão clara sobre a sua regulação e sobre seus efeitos combinatórios.

A partir deste conhecimento, em trabalhos futuros no genoma de *Ricinus communis*, será possível fazer inferências de forma sistêmica, também levará a uma melhor compreensão da regulação transcricional e da funcionalidade destes genes no genoma.

Ainda permitirá alterar efetivamente o padrão de expressão dos genes-*RcPR* de forma desejada, por meio da regulação do padrão de expressão, que

contribuirá para o desenvolvimento de genótipos resistentes as principais moléstias de importância econômica na espécie e de genótipos adaptativos a diferentes condições edafoclimáticas, situações de seca, estresse salínico e de frio, por exemplo. Além disso, estas informações podem ser empregadas em futuros estudos dos genes *PR* de outras espécies de plantas.

Análise das proteínas RcPRs

As Proteínas-PR são cruciais no mecanismo de defesa da planta, em geral, apresentam baixo peso molecular, ponto isoelétrico e grau médio de hidropaticidade versáteis, o que permitem desde a presença ao seu desempenho extra e intracelular, e interação com outras moléculas dispersas no meio (VAN LOON; VAN STRIEN, 1999; JIANG et al., 2015).

A família mais abundante neste trabalho, a das RcPR-1 com 10 membros, apresentaram pesos moleculares, menor do que se espera para as demais PRs, semelhante ao reportado na literatura (6-43 kDa) (TELLIS et al., 2017).

No presente trabalho, a caracterização desta família em dois subgrupos, ácidas e básicas, de acordo com o pI (Tabela 1), com regiões hidrofílicas, indicam que são proteínas solúveis em meio aquoso, contendo possivelmente aminoácidos polares sem carga.

Sabe-se que as Proteínas-PR básicas podem ser induzidas após o tratamento com ET e AJ, como observado nas proteínas NtPRB1b em *Nicotiana tabacum* L. e nas proteínas AtPRB1 em *Arabidopsis thaliana* L. (EYAL et al., 1992; NIKI et al., 1998; SANTAMARIA et al., 2001). Enquanto as proteínas ácidas geralmente são induzidas via AS, pois a elevada produção desse composto ativa a via de transdução de sinal molecular, conhecida como RAS (OHASHI et al., 1987; MITSUHARA et al., 2008).

Tellis et al. (2017), ao caracterizarem a família das PR-1 em leguminosas identificaram que esta família também apresentou GRAVY negativo (-), e pI que lhes permitiu diferenciar as PR-1 em ácidas e básicas, corroborando com os resultados do presente estudo. Isto dar-se porque esta família é tida como multigênica dentro dos genomas de plantas (BREEN et al., 2017).

Ademais, com base nos resultados da análise da região promotora, por meio da presença de elementos em resposta a fitormônios como o AS, pode-se sugerir que membros da família das proteínas RcPR-1 estejam envolvidos na

resposta a injúria pelo mecanismo de resistência RAS. Além disso, a expressão das PR-1, bem como das PR-2 tem sido rotineiramente usada como um marcador molecular da RAS (ALI et al., 2018).

Quanto à previsão de localização subcelular, os membros pertencentes ainda a esta família, direcionados a lisossomos/vacúolos, RcPR-1a, RcPR1b.2, RcPR-1c, RcPR-1d, RcPR-1Be, RcPR-1Be e RcPR-1i, podem estar envolvidos na degradação de compostos produzidos pelos fitopatógenos, advindos do meio extracelular, bem como de proteínas de membrana (BREEN et al., 2017).

Em trabalhos anteriores, na cultura do tabaco (*Nicotiana tabacum* L.) algumas proteínas PR-1 foram identificadas nos vacúolos de protoplastos (SESSA et al., 1995) e nos vacúolos de células especializadas, conhecidas como idioblastos de cristal (DIXON et al., 1991). Das proteínas RcPR-1, apenas a RcPR-1b.1 teve previsão de endereçamento para os cloroplastos e nela identificado o domínio conservado com atividade Ribonuclease, responsável por degradar RNA em componentes menores (Tabela 4).

Recentemente, Boccoardo et al. (2019) confirmaram a presença de proteínas-PR em cloroplastos de plantas transgênicas de tabaco (*Nicotiana tabacum* L.), bem como sua ação contra três fitopatógenos filamentosos com estilos de vida diferentes: um patógeno fúngico necrotrófico (*Rhizoctonia solani*), um oomiceto hemibiotrófico (*Phytophthora nicotianae*) e um obrigatório oomiceto biotrófico (*Peronospora hyoscyami* f.sp. *tabacina*).

Já as proteínas RcPR-1g e RcPR-1h, com probabilidade de endereçamento extracelular podem ter função análoga à outras já descritas no fluido extracelular, conferindo alta resistência à degradação por enzimas proteolíticas, responsáveis pela quebra de ligação peptídica entre aminoácidos de proteínas, pois o líquido extracelular ao redor da área necrótica apresenta um pH mais baixo e alta atividade proteolítica, logo, as proteínas devem ser bem adaptadas a estas condições para atuar neste meio (SESSA et al., 1995; DZHAVAKHIYA et al., 2007).

Tudo isto fica evidenciado pela presença do domínio proteico conservado nesta família, o de Proteínas Secretórias Ricas em Cisteínas (Tabela 4), pois as cisteínas formam pontes dissulfeto entre os grupos tiol, que além de proporcionar aumento da estabilidade da molécula, aumentam a sua resistência a atividade proteolítica, processo de degradação por enzimas proteases ou por digestão

intramolecular (GUO et al., 2005).

Assim, é possível que as proteínas RcPR-1 localizadas no meio extracelular, devido a sua estrutura, com pontes dissulfeto formadas pelos resíduos de cisteínas, lhes proporcionam estabilidade molecular para serem adaptadas para atuar nas áreas necróticas de planta infectadas por microrganismos oportunistas. Ainda neste sentido, as RcPR-1 extracelulares podem estar associadas na resposta imediata a infecção por fitopatógenos e tratamentos químicos, conforme Dzhavakhiya et al. (2007).

A abundância desta família tanto no espaço inter e intracelular sugere forte atuação durante a infecção, o que propicia uma função antimicrobiana bastante eficaz (BREEN et al., 2017).

Ainda segundo Breen et al. (2017), acredita-se que a maioria das PR-1 sejam secretadas e se acumulem no espaço extracelular/apoplástico, o que é facilitado pela presença do peptídeo de secreção N-terminal.

Ademais, a localização também pode estar associada ao fato da expressão das proteínas-PR serem dada por meio diferentes de vias de sinalização em diferentes compartimentos celulares que desencadeiam um processo por meio do qual uma célula converte um estímulo/sinal em outro estímulo/sinal, em resposta a RAS (DZHAVAKHIYA et al., 2007; ALI et al., 2018). Dessa forma, após a penetração do patógeno na planta, enzimas são liberadas dos tilacóides do cloroplasto, a exemplo das polifenoloxidasas, que têm a capacidade de oxidar compostos fenólicos para quinonas (MAYER; STAPLES 2002; THIPYAPONG et al., 2004). Os compostos fenólicos fazem parte do sistema de defesa bioquímico constitutivo das plantas e podem agir inibindo o crescimento bacteriano, limitando a translocação do patógeno, pois atuam como precursores na formação de barreiras físicas que compõe a primeira linha de defesa dos organismos (AYDEMIR, 2004).

Contudo, as quinonas são moléculas altamente reativas, e mais tóxicas aos microrganismos que o fenol original, podem se polimerizar levando a formação de pigmentos de cor marrom ou preta, responsáveis pelo escurecimento dos tecidos (KARSTEN et al., 2013).

Do mesmo modo, membros da família das RcPR-1 podem ser transportados por este sistema como estratégia de defesa para combater a invasões de fitopatógenos, pois assim como os compostos fenólicos são induzidos de forma constitutiva também são as proteínas-PR e outros

compostos, apesar da descrição separadamente, sabe-se que no sistema todos os compostos passivos, constitutivos e ativos e induzíveis são fatores de resistência que atuam de forma conjunta e harmoniosa em resposta a infecção. Duas das três proteínas RcPR-4 do presente trabalho, RcPR-4a.1 e RcPR-4a.2, apresentaram valores de peso molecular bem próximos ao reportado para membros desta mesma família na espécie de *Lycoris radiata* (LI et al., 2010).

Ainda de acordo com os mesmos autores, geralmente, as PR-4 apresentam baixo peso molecular, de 13-16 kDa, e são conhecidas por serem proteínas ácidas, podendo ser básicas também, que corroboram com os resultados deste trabalho, onde o pI variou de 6,24 a 9,01. Em contraste com outras famílias de PRs, as PR-4 constituem um dos grupos menos estudado e caracterizado em vegetais. Até o presente momento, estudos sobre esta família ainda são incipientes.

Os membros pertencentes ainda à esta família, apresentaram probabilidade de endereçamento para o meio extracelular, o que também foi observado em trabalhos anteriores, onde mostram algumas proteínas PR-4 com um domínio de extensão na região C-terminal sendo direcionadas a vacúolos, evidenciando que isoformas podem ser endereçadas tanto para o meio interno quanto para o meio externo da célula (MEINS et al., 1991; BERTINI et al., 2009; LU et al., 2012).

Ainda sobre esta família, sabe-se que podem atuar na parede celular de fitopatógenos, causando a degradação de suas estruturas, levando a inibição do crescimento fúngico (BORMANN et al., 1999; MENEZES et al., 2014).

Esta atuação é explicada pela presença do domínio proteico conservado, o domínio *Barwin*, em todas as RcPR-4, que é caracterizado pela presença de seis resíduos de cisteínas que formam três pontes dissulfeto, como já dito anteriormente, além de contribuir para estabilidade da molécula permite que as mesmas atuem no espaço extracelular, dessa forma, permitindo assim a sua atuação nas estruturas da parede celular de fitopatógenos (DZHAVAKHIYA et al., 2007).

Além disto, esta família também foi caracterizada por apresentar ação de RNase/DNase, conforme os resultados do trabalho de Menezes et al. (2014). Assim, além de degradar componentes da parede celular fúngica, pode também, degradar o material genético destes oportunistas.

As cinco proteínas pertencentes a família RcPR-5, apresentaram peso

superior; de 25,45 a 31,26 kDa; ao reportado em trabalhos anteriores, por Zhang et al. (2018) de 17,30 kDa na cultura do trigo (*Triticum ssp.*). A maioria destas proteínas possui um peso molecular na faixa de 20–26 kDa, enquanto as atípicas podem ser menores (17 kDa, aproximadamente) ou maiores com domínios e/ou funcionalidades adicionais, chegando a 70 kDa (ZHANG et al., 2018).

As proteínas RcPR-5 analisadas, que apresentam propriedades hidrofóbicas, a RcPR-5a e a RcPR-5e, e uma previsão de exportação para o meio extracelular, podem indicar possíveis associações com membranas de fungos, já que as PR-5 apresentam atividade antifúngica (EL-KEREAMY et al., 2011). Graças a maior quantidade de aminoácidos apolares nestas RcPR-5, é possível que estas proteínas estabeleçam interações hidrofóbicas com os lipídios presentes a membrana dos fitopatógenos, e por meio deste contato, podem impedir a passagem de água, causando instabilidade da membrana e dessa forma a inviabilidade do fitopatógeno em continuar seu processo de infecção na planta.

Entretanto, as demais RcPR-5 (RcPR-5b, RcPR-5c e RcPR-5d) com propriedades hidrofílicas, por apresentarem em sua maioria aminoácidos apolares, apresentaram probabilidade de exportação ao meio extracelular, o que já era esperado, apenas para a RcPR-5c, por apresentar um ponto isoelétrico ácido, pois sabe-se que as proteínas-PR, em geral, caracterizadas como ácidas normalmente são destinadas ao espaço extracelular, contudo as proteínas com características básicas podem ser também direcionadas ao meio extracelular (TAKEDA et al., 1991; ZHANG et al., 2018).

Já a previsão de localização da RcPR-5d, a única desta família destinada a membrana celular é sem dúvida de grande utilidade para o sistema de defesa, pois contribui para o aumento na permeabilidade da membrana plasmática do fitopatógeno, interrompendo a bicamada lipídica e criando poros transmembranares, exercendo assim a sua função antifúngica, isto é, reduzindo o crescimento e desenvolvimento dos fitopatógenos (EDREVA, 2005; EL-KEREAMY et al., 2011). Outras PR-5 já foram descritas com atividade osmotina, interferindo na via de transdução de sinal de patógenos para aumentar a suscetibilidade das células fúngicas (JI et al., 1998).

A única proteína pertencente à família das RcPR-6 analisada, apresentou um peso molecular de 17,83 kDa, típico para esta família, onde espera-se um peso variável de 8-20 kDa, aproximadamente, de predominância hidrofílica e

característica básica (VAN LOON et al., 1994; RAKWAL et al., 2001; SELS et al., 2008; MYAGMARJAV et al., 2017).

Sua probabilidade de direcionamento foi para lisossomos/vacúolos, porém, a proteína RcPR-6 é homóloga a mesma superfamília PF00188 das RcPR-1. Ambas compartilham do mesmo domínio proteico extracelular CAP ou SCP, contudo, ainda pouco se conhece sobre o seu papel biológico, sabe-se que se acumulam após a infecção por fitopatógenos, atuam como agentes antifúngicos e são inibidores de proteinase implicados na defesa contra insetos, microrganismos e nematoides.

Apenas uma proteína foi classificada na família das RcPR10-STH21. Inicialmente esta família foi identificada na cultura da batata (*Solanum tuberosum* L.) e apresentou um peso de 17,65 kDa, próximo ao apresentado por Dahal et al. (2009) de 17,20 kDa, como também o peso de outras RcPRs caracterizadas neste trabalho. Esta proteína apresentou propriedade hidrofílica e característica ácida como também foi observado em seis membros das PR-1 (RcPR-1a; RcPR-1b.1; RcPR-1b.2; RcPR-1d; RcPR-1Be; RcPR-1g), na RcPR-4b e em quatro membros das PR-5 (RcPR-5a; RcPR-5c; RcPR-5d e RcPR-5e) do presente trabalho.

Quanto a sua probabilidade de endereçamento para o citoplasma celular é comprovada porque esta família também é conhecida como Proteínas Relacionadas a Patogêneses Intracelular e apresenta atividade ribonuclease, corroborando assim com a predição de localização realizada neste trabalho (MELNIKOVA et al., 2017).

De forma geral, das 20 proteínas RcPR caracterizadas neste trabalho, apenas quatro não apresentaram peptídeo sinal. Em muitos casos, a capacidade de endereçamento de certas sequências sinalizadoras é confirmada por meio da fusão da sequência sinalizadora de uma proteína com outra proteína, mostrando que um peptídeo sinal pode endereçar mais de uma proteína ao seu destino final (NELSON, 2014).

Embora o principal papel das proteínas-PR sejam atuar no mecanismo de defesa juntamente com outros compostos orgânicos produzidos pelo mecanismo de defesa celular, elas claramente desempenham outros papéis no desenvolvimento da planta, como já descrito anteriormente (CAMPOS et al., 2007), justificando assim a sua ampla capacidade de endereçamento para compartimentos inter e intracelular.

A presença de múltiplos sítios de fosforilação nas 20 proteínas RcPR caracterizadas, evidenciam que nessas ocorrem alterações pós-traducionais, resultando na ativação, desativação ou até mesmo em modificações em suas funções, mediante a mudança nas interações com os aminoácidos adjacentes (COHEN et al., 2002). A presença destes múltiplos sítios de fosforilação em uma proteína podem ser independentes entre si, ou pode haver interdependências entre eles ou mesmo com outros componentes-chaves que também sejam essenciais na transdução de sinais no genoma de *R. communis* (COHEN, 2000).

Este tipo de modificação pós-traducional é a mais frequente nas proteínas dos seres vivos (KRUGER et al., 2006) e evidencia a evolução de um genoma (AMOUTZIAS et al., 2010; VLASTARIDIS et al., 2017). Um exemplo bem conhecido de modificações em suas funções via fosforilação é na proteína supressora de tumor conhecida por p53 com 18 locais diferentes de fosforilação já caracterizados, onde a ativação desta proteína pode levar à parada do ciclo celular, podendo até mesmo levar à morte celular (BATES; VOUSDEN, 1996).

Além dos sítios de fosforilação, os sítios de glicosilação também foram identificados nas proteínas RcPRs. Este tipo de modificação pós-traducional é conhecido como *N-glicosilação*, onde uma molécula de carboidrato é ligada a uma molécula de nitrogênio da cadeia lateral da asparagina (ligação do tipo N). Este tipo de ligação é exclusivo nos resíduos de asparagina (N), porém, é uma ligação tão importante quanto à fosforilação, envolvidas na modulação da estrutura e função das proteínas eucarióticas (O'CONNOR et al., 1999; MEDZIHRADESKY, 2008).

As estruturas de glicano ligadas às proteínas podem variar de acordo com a espécie e o tecido. Sabe-se que o evento de *N-glicosilação* desempenha uma função significativa em vários processos biológicos, como ligação ao receptor, reconhecimento imunológico celular, inflamação e patogenicidade (RABINOVICH et al., 2012; CHEN et al., 2014; SCOTT; PANIN, 2014; ALTER et al., 2018), e pode ser afetado por alterações fisiológicas (LANDBERG et al., 2000) e doenças (HIGAI et al., 2005).

No genoma de *R. communis*, em resumo ao que já foi dito, estas modificações nas proteínas RcPR podem determinar a atividade, localização se para o meio inter ou intracelular, além de interações com outras moléculas, principalmente as modificações de fosforilação, pois diversos sítios foram identificados nas 20 proteínas RcPR deste trabalho. Na sinalização, acredita-se

por exemplo, que estas modificações ainda podem ativar ou desativar cascatas de sinalização envolvidas produção destas proteínas no genoma de *R. communis* através da adição ou remoção de grupos fosfatos e dos de glicosilação.

A presença do motivo GxGGxG, também identificado nas espécies *Capsicum annuum*, *Panax ginseng* e *Saccharum arundinaceum*, indica a existência de um possível sítio ativo ao fosfato de RNA que pode estar correlacionado com atividade ribonuclease. Este motivo é rico em glicina e conhecido como loop P (loop de ligação ao fosfato) (BANTIGNIES et al., 2000; PARK et al., 2004; KIM et al., 2014; MOHAN et al., 2020).

Apesar de ser um motivo típico da família das PR-10, também pode ser encontrado em outras PRs (HOFFMANN-SOMMERGRUBER et al., 1997), tanto em sua forma completa já descrita, como na forma reduzida (GXG) observado na proteína STH2 em batata (*Solanum tuberosum*) (ELBEZ et al., 2002).

Chen et al. (2014) ao analisarem várias espécies mono e dicotiledôneas, identificaram o motivo PxGNxxxxxPY, na região do peptídeo conhecido como CAPE1. A fim de verificar a funcionalidade deste peptídeo, os autores selecionaram uma PR-1 de *Arabidopsis thaliana* que continha esse motivo (AtCAPE-PR-1) e experimentalmente testaram no tomate (*Solanum lycopersicum*). A aplicação do AtCAPE-PR-1 resultou no aumento da imunidade contra a infecção causada pelo *Pseudomonas syringae* pv. tomato, conhecido como Pst DC3000, uma mancha bacteriana. Assim, acredita-se que a presença deste motivo nas RcPR-1, contribuam para o aumento da imunidade contra o ataque de fitopatógenos bacterianos em *R. communis*.

Estes mesmos autores, ainda identificaram o motivo CNYx também localizado na região N-terminal do peptídeo CAPE1. Este motivo representa o local de clivagem CAPE1. Um estudo subsequente confirmou esta informação por meio de uma abordagem de mutação direcionada ao local. Contudo, o mecanismo de clivagem que este motivo exerce ainda permanece desconhecido (CHIEN et al., 2015).

Em geral, a presença destes motivos proteicos funcionais identificados a partir das sequências de aminoácidos, revelam a presença de um padrão de sequências conservadas que podem apresentar importância biológica, correspondente ao local ativo de uma RcPR ou de uma unidade estrutural, necessária para o dobramento adequado destas proteínas ou até mesmo domínios de estabilização.

Relação filogenética

A filogenia entre as proteínas-PR nas espécies *Ricinus communis*, *Nicotiniana tabacum* e *Arabidopsis thaliana*, mostrou a formação de agrupamentos, formados em sua maioria, por membros de uma mesma família gênica, porém, sem um distanciamento muito claro entre grupos. A formação de grupos com membros de uma mesma família dar-se devido à similaridade parcial nas sequências de aminoácidos com uma distância não muito claro, possivelmente devido ao *background* genético, em virtude desta classe de proteínas apresentarem a mesma função biológica em resposta as doenças fúngicas, bacterianas e virais.

Uma possível característica que poderia levar a segregação de membros de uma mesma família gênica de PRs em diferentes grupos seria pela análise entre espécies monocotiledôneas e dicotiledôneas, assim como foi observado por Mohan et al. (2020) que observaram membros da família gênica das PR-10 separadas nos dois clados, mono e dicotiledôneas.

CONCLUSÃO

Relatamos pela primeira vez a caracterização das famílias gênicas das Proteínas Relacionas a Patogênese em *Ricinus communis* L. Esta caracterização torna-se uma estratégia viável que proporciona uma melhor compreensão do mecanismo atual das Proteínas-PR, permitindo desenvolver estratégias direcionadas que visem aumentar a resistência da espécie contra agressores bióticos, especialmente na cultura da mamona onde este tipo de estudo ainda não é bem relatado.

A presença das cinco famílias das proteínas-PR em *R. communis*, distribuídas em 20 genes, evidencia que esta classe de proteínas está envolvida em um amplo espectro de respostas a diversas condições tanto patológicas quanto aquelas ligadas ao crescimento e desenvolvimento da planta.

Isto ficou fundamentado pela identificação de diferentes elementos *cis*-reguladores na região promotoras destes genes PRs, que categoricamente foram relacionados a fatores de transcrição, resposta a estresses de aspectos físicos, químicos, bem como pela presença dos principais elementos hormonais

como AJ, AS, ABA, MeJA, ET e AG, responsáveis por ativar as principais vias de sinalização na planta em resposta a ataques de fitopatógenos, permitindo a ativação de diferentes famílias gênicas com atividade biológica específica para contra-atacar o invasor oportunista.

Os direcionamentos previstos das proteínas *RcPRs* condisseram com os atributos intrínsecos de cada família gênica caracterizada, sendo a maioria direcionadas ao meio extracelular, a fim de impedir a formação de apressório fúngico, causando toxicidade a componentes da parede celular ou membrana plasmática dos oportunistas, resultando na sua morte celular.

A análise filogenética revelou que as proteínas-PRs nas três espécies analisadas, formou agrupamentos, composto majoritariamente por membros de uma família gênica, evidenciando que ao longo do processo evolutivo cada família manteve seu papel biológico específico.

AGRADECIMENTOS

Agradecemos a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudo que permitiu o desenvolvimento desta pesquisa.

REFERÊNCIAS

- MEDZIHRADESKY, K. F. Characterization of site-specific N-glycosylation. **Methods in molecular biology**, v. 446, p. 293-316, 2008.
- AGARWAL, P.; BHATT, V.; SINGH, R. M.; SOPORY, S.; CHIKARA, J. Pathogenesis-related gene, JcPR-10a from *Jatropha curcas* exhibit RNase and antifungal activity. **Molecular Biotechnology**, v. 2.; p. 412–425, 2013.
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**, 4th edition. Garland Science, New York 2002.
- ALI, S.; GANAI, B. A.; KAMILI, A. N.; BHAT, A. A.; MIR, Z. A.; BHAT, J. A.; TYAGI, A.; ISLAM, S. T.; MUSHTAQ, M.; YADAV, P.; RAWAT, S.; GROVER, A. Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. **Microbiological Research**, v. 212, p. 29-37, 2018.
- ALTER, G.; OTTENHOFF, T. H. M.; JOOSTEN, S. A. Antibody Glycosylation in Inflammation, Disease and Vaccination. **Seminars Immunology**, v. 39, p. 102-110, 2018.
- AMOUTZIAS, G. D.; HE, Y.; GORDON, J.; MOSSIALOS, D.; OLIVER, S., G.; PEER, Y., V. Posttranslational regulation impacts the fate of duplicated genes. **Proceedings of the National Academy of Sciences**, v. 107, p. 2967-2971, 2010.
- 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.
- ARMENTEROS, J. J. A.; TSIRIGOS, K. D.; SØNDERBY, C. K.; PETERSEN, T. N.; WINTHER, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN., H. SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature Biotechnology**, v. 37, p. 420-423, 2019.
- ARORA, R.; KUMAR, A.; SINGH, I. K.; SINGH, A. Pathogenesis Related Proteins: a defensin for plants but an allergen for humans. **International Journal of Biological Macromolecules**. v. 157, p. 659-672, 2020.
- AYDEMIR, T. Partial purification and characterization of polyphenol oxidase from artichoke (*Cynara scolymus* L.) heads. **Food Chemistry**, v. 87, p. 59-67, 2004.
- BANTIGNIES, B.; SÉGUIN, J.; MUZAC, I.; DÉDALDÉCHAMP, F.; GULICK, P.; IBRAHIM, R. Direct evidence for ribonucleolytic activity of a PR-10-like protein from white lupin roots. **Plant Molecular Biology**, v. 42, p. 871-881, 2000.
- BAILEY, T. L.; BODÉN, M. WHITINGTON, T.; MACHANICK, P. The value of position-specific priors in motivo discovery using MEME. **BioMed Central Bioinformatics**, v. 11, p. 1-14, 2010.
- BAILEY, T. L.; JOHNSON, J.; GRANT, C. E.; NOBLE, W. S. The MEME Suite.

Nucleic Acids Research, v. 43, p. 39-49, 2015.

BATES, S.; VOUSDEN, K. H. p53 in signaling checkpoint arrest or apoptosis. **Current opinion in genetics and development**, v. 6, p. 12-18, 1996.

BERTINI, L.; CAPORALE, C.; TESTA, M.; PROIETTI, S.; CARUSO, C. Structural basis of the antifungal activity of wheat PR4 proteins. **FEBS Letters**, v. 17, p. 2865–2871, 2009.

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.

BOCCARDO, N. A.; SEGRETIN, M. E.; HERNANDEZ, I.; MIRKIN, F. G.; CHACÓN, O.; LOPEZ, Y.; BORRAS-HIDALGO, O.; BRAVO-ALMONACID, F. F. Expression of Pathogen-Related Proteins in transplastomic tobacco plants confers resistance to filamentous pathogens under field trials. **Nature Scientific Reports**, v. 9, p. 1-13, 2019.

BORMANN, C.; RAPS, C.; BAIER, D.; JUNG, G.; SCHWARZ, H.; HÖRR, I.; BERGER, J. Characterization of a novel, antifungal, chitin-binding protein *Streptomyces tendae* Tu901 that interferes with growth polarity. **Journal of Bacteriology**, v. 181, p. 7421-7429, 1999.

BREEN, S.; WILLIAMS, S. J.; OUTRAM, M.; KOBE, B.; SOLOMON, P. S. Emerging Insights into the Functions of Pathogenesis-Related Proteins 1. **Trends in Plant Science**, v.22, p. 871-879, 2017.

CAPELLI, N.; DIOGON, T.; GREPPIN, H.; SIMON, P. Isolation and characterization of a cDNA clone encoding an osmotin-like protein from *Arabidopsis thaliana*. **Gene**, v. 1, p. 51-56, 1997.

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.; GEDIL, M.; STANKE, M.; HAAS, B. J.; WORTMAN, J. R.; FRASER-LIGGETT, C. M.; RAVEL, J.; RABINOWICZ, P. D. Draft genome sequence of the ricin-producing oilseed castor bean. **Nature Biotechnology**, v. 28, p. 951-956, 2010.

CHEN, Y-L.; LEE, C-Y.; CHENG, K-T.; CHANG, W-H.; HUANG, R-N.; NAM, H. G.; CHEN, Y-R. Quantitative Peptidomics Study Reveals That a Wound-Induced Peptide from PR-1 Regulates Immune Signaling in Tomato. **The Plant Cell**, v. 26, p. 4135–4148, 2014.

CHEN, X-L.; SHI T.; YANG, J.; SHI, W.; GAO, X.; CHEN, D.; XU, X.; XU, J-R.; TALBOT, N. J.; PENG, Y-L. N-Glycosylation of Effector Proteins by an α -1,3-Mannosyltransferase Is Required for the Rice Blast Fungus to Evade Host Innate Immunity. **The Plant Cell**, v. 26, p. 1360-1376, 2014.

CHIEN, P-S.; NAM, H. G.; CHEN, Y-R. A salt-regulated peptide derived from the CAP superfamily protein negatively regulates salt-stress tolerance in *Arabidopsis*. **Journal of Experimental Botany**, v. 66, p. 5301–5313, 2015.

CHO, M. H.; LEE, S. W. Phenolic phytoalexins in rice: biological functions and

biosynthesis. **International Journal of Molecular Sciences**, v. 12 p. 29120–29133, 2015.

CHOUDHARY, V.; SCHNEITER, R. Pathogen-Related Yeast (PRY) proteins and members of the CAP superfamily are secreted sterol-binding proteins. **PNAS**, v. 42, p. 168882-16887, 2012.

COHEN, P. The origins of protein phosphorylation. **Nature Cell Biology**, v. 4, p.127-130, 2002.

COHEN, P. The regulation of protein function by multisite phosphorylation—a 25 year update. **Trends in Biochemical Sciences**, v. 25, p. 596-601, 2000.

DAFOE, N. J.; GOWEN, B. E.; CONSTABEL, C. P. Thaumatin-like proteins are differentially expressed and localized in phloem tissues of hybrid poplar. **BioMed Central Plant Biology**, v. 10, p. 1-11, 2010.

DAHAL, D.; HEINTZ, D.; VAN DORSSELAER, A.; BRAUN, H-P.; WYDRA, K. Pathogenesis and stress related, as well as metabolic proteins are regulated in tomato stems infected with *Ralstonia solanacearum*. **Plant Physiology and Biochemistry**, v. 47, p. 838-846, 2009.

DELANO, W. L. PyMOL: An Open-Source Molecular Graphics Tool. **Ccp4 Newslett Protein Crystallogr**, v. 40, p. 1-9, 2002.

DIXON, D. C.; CUTT, J.R.; KLESSIG, D. F. Differential targeting of the tobacco PR-1 Pathogenesis-Related Proteinss to the extracellular space and. **The Embo Journal**, v. 10, p. 1317-1324, 1991.

DZHAVAKHIYA, V. G.; OZERETSKOVSKAYA, O. L.; ZINOVYEVA, S. V. Chapter 10 - Immune response. **Comprehensive and Molecular Phytopathology**, p. 265-314, 2007.

EDREVA, A. Pathogenesis-Related Proteinss: Research progress in the last 15 years. **General and Applied Plant Physiology**, v. 31, p. 105-124, 2005.

ELBEZ, M.; KEVERS, C.; HAMDI, S.; RIDEAU, M.; PETIT-PALY, G. Les protéines de pathogenèse PR-10 des végétaux. **Acta Botanica Gallica**, v. 4, p. 415-444, 2002.

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

EL-KEREAMY, A.; EL-SHARKAWY, I.; RAMAMOORTHY, R.; TAHERI, A.; ERRAMPALLI, D.; KUMAR, P.; JAYASANKAR, S. *Prunus domestica* Pathogenesis-Related Protein-5 Activates the Defense Response Pathway and Enhances the Resistance to Fungal Infection. **Plos One**, v. 3, p. 1-11, 2011.

GAMIR, J.; DARWICH, R.; VN1T HOF, P.; CHOUDHARY, V.; STUMPE, M.;

SCHNEITER, R.; MAUCH, F. The sterol-binding activity of PATHOGENESIS-RELATED PROTEIN 1 reveals the mode of action of an antimicrobial protein. **The Plant Journal**, v. 89, p. 502-509, 2017.

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 (2005), p. 571-607.

GIANINAZZI, S.; MARTIN, c.; VALÉE, J.-C. Hypersensibilité aux virus, température et protéines solubles chez le *Nicotiana Xanthi* n.c. Apparition de nouvelles macromolécules lors de la répression de la synthèse virale. **Comptes rendus de l'Académie des Sciences**, v. 270, p. 2383-2386, 1970.

GHOSH, M. Antifungal properties of haem peroxidase from *Acorus calamus*. **Annals of Botany**, v. 6, p. 1145-1153, 2006.

GOODWIN, S. B.; SUJKOWSKI, L. S.; FRY, W. E. Widespread Distribution and Probable Origin of Resistance to Metalaxyl in Clonal Genotypes of *Phytophthora infestans* in the United States and Western Canada. **Phytopathology**, v. 86, p. 793–800, 1996.

GRANT, M. R.; JONES, J. D. Hormone (dis) harmony moulds plant health and disease. **Science**, v. 324, p. 750-752, 2009.

GUEx, N.; PEITSCH, M. C.; SCHWEDE, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. **Electrophoresis**, v. 1, p. 162-173, 2009.

GUO, M.; TENG, M.; NIU, L.; LIU, Q.; HUANG, Q.; HAO, Q. Crystal structure of the cysteine-rich secretory protein stecrisp reveals that the cysteine-rich domain has a K⁺ channel inhibitor-like fold. **The Journal of Biological Chemistry**, v. 13, p. 12405–12412, 2005.

HA, C. D.; DUNG, L. T. VAI TRÒ CỦA YẾU TỐ ĐIỀU HÒA CIS TRONG ĐÁPỨNG CỦA THỰC VẬT VỚI CÁC ĐIỀU KIỆN BẤT LỢI. **Academia Journal of Biology**, v.3. p. 370-383, 2015.

HEITZ, T.; SEGOND, S.; KAUFFMANN, S.; GEOFFROY, P.; PRASAD, V.; BRUNNER, F.; FRITIG, B.; LEGRAND, M. Molecular characterization of a novel tobacco pathogenesis-related (PR) protein: a new plant chitinase/lysozyme. **Molecular Genetics and Genomics**, v. 2, p. 246–254, 1994.

HIGAI, K.; AOKI, Y.; AZUMA, Y.; MATSUMOTO, K. Glycosylation of site-specific glycans of alpha1-acid glycoprotein and alterations in acute and chronic inflammation. **Biochimica Biophysica Acta**, v. 1725, p. 128-135, 2005.

HOFFMANN-SOMMERGRUBER, K.; VANEK-KREBITZ, M.; RADAUER, C.; WEN, J.; FERREIRA, F.; SCHEINER, O.; BREITENEDER, H. Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. *Gene*, v. 197, p. 91-100, 1997.

HOU, M.; XU, W.; BAI, H.; LIU, Y.; LI, L.; LIU, L.; LIU, B.; LIU, G. Characteristic Expression of Rice Pathogenesis-Related Proteinss in Rice Leaves During Interactions With *Xanthomonas Oryzae* Pv. *Oryzae*. **Plant Cell Reports**, v. 5, p. 895-904, 2012.

HU, B.; JIN, J.; GUO, A-Y.; ZHANG, H.; LUO, J.; GAO, G. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics**, v. 31, p. 1296-1297, 2015.

IONESCU, I. A.; LOPEZ-ORTEGA, G.; BUROW, M.; BAYO-CANHA, A.; JUNGE, A.; GERICKE, O.; MOLLER, B.; SANCHEZ-PEREZ, R. Transcriptome and Metabolite Changes during Hydrogen Cyanamide-Induced Floral Bud Break in Sweet Cherry. **Frontiers in Plant Science**, v. 8, p. 1-17, 2017.

JANSKA, A.; MARIK, P.; ZELENKOVA, S.; OVESNA, J. Cold stress and acclimation: what is important for metabolic adjustment? **Plant Biology**, v. 12, p. 395-405, 2010.

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

JIANG, L.; WU, J.; FAN, S.; LI, W.; DONG, L.; CHENG, Q.; XU, P.; ZHANG, S. Isolation and Characterization of a Novel Pathogenesis-Related Protein Gene (GmPRP) with Induced Expression in Soybean (*Glycine max*) during Infection with *Phytophthora sojae*. **Plos one**, v. 6, p. 1-19, 2015.

JONES, J. D.; DANGL, J. L. The plant immune system. **Nature**, v. 444, p. 323–329, 2006.

KARSTEN, J.; CHAVES, D. V.; FINGER, F. L.; BARBOSA, J. G. Papel e caracterização da polifenoloxidase em hastes de ave-do-paraíso (*Strelitzia reginae* Ait.). **Revista Brasileira de Horticultura Ornamental**, v. 19, p. 161-171, 2013.

KAUR, A.; PATI, P. K.; PATI, A. M.; NAGPAL, A. K. In-silico analysis of cis-acting regulatory elements of Pathogenesis-Related Proteinss of *Arabidopsis thaliana* and *Oryza sativa*. **Plos One**, v. 9, p. 1-24, 2017.

KIM, Y-J.; JANG, M-G.; LEE, H-J.; JANG, G-H.; SUKWEENADHI, J.; KWON, W-S.; YANG, D-C. Functional characterization of the pathogenesis-related protein family 10 gene, PgPR10-4, from *Panax ginseng* in response to environmental stresses. **Plant Cell**, v. 118, p. 531–543, 2014.

KIM, Y. J.; LEE, H. J.; JANG, M. G.; KWON, W. S.; KIM, S. Y.; YANG, D. C. Cloning and characterization of pathogenesis related-protein 4 gene from *Panax ginseng*. **Russian Journal of Plant Physiology**, v. 61, p. 664-671, 2014.

KIM, Y. S.; PARK, J. Y.; KIM, K. S.; KO, M. K.; CHEONG, S. J.; OH, B. J. A thaumatin-like gene in nonclimacteric pepper fruits used as molecular marker in probing disease resistance, ripening, and sugar accumulation. **Plant Molecular Biology**, v. 2, p. 125-135, 2002.

KINKEMA, M.; FAN, W.; DONG, X. Nuclear Localization of NPR1 Is Required for Activation of PR Gene Expression. **The Plant Cell**, v. 12, p. 2339-2350, 2000.

KRUGER, R.; KUBLER, D.; PALLISSÉ, R.; BURKOVSKI, A.; LEHMANN, W. D. Protein and proteome phosphorylation stoichiometry analysis by element mass spectrometry. **Analytical Chemistry**, v. 78, p. 1987-1994, 2006.

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution**, v. 7, p. 1870-1874, 2016.

LALUK, K.; MENGISTE, T. The Arabidopsis extracellular unusual serine protease inhibitor functions in resistance to necrotrophic fungi and insect herbivory. **The Plant Journal**, v. 3, p. 480–494, 2011.

LANDBERG, E.; HUANG, Y.; STROMQVIST, M.; MECHREF, Y.; HANSSON, L.; LUNDBLAD, A.; NOVOTNY, M. V.; PAHLSSON, P. Changes in glycosylation of human bile-salt-stimulated lipase during lactation. **Archives of Biochemistry and Biophysics**, v. 377, p. 246-254, 2000.

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

LESCOT, M.; DH AIS, P.; THIJIS, G.; MARCHAL, K.; MOREAU, Y.; VAN PEER, Y.; ROUZ, 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. 1, p. 325-327, 2002.

LI, X.; XIA, B.; JIANG, Y.; WU, Q.; WANG, C.; HE, L. PENG, F.; WANG, R. A new Pathogenesis-Related Proteins, LrPR4, from *Lycoris radiata*, and its antifungal activity against *Magnaporthe grisea*. **Molecular Biology Reports**, v. 37, p. 995-1001, 2010.

LI, X-L.; FAN, S.; HU, W.; LIU, G.; WEI, Y.; HE, C.; SHI, H. Two Cassava Basic Leucine Zipper (bZIP) Transcription Factors (MebZIP3 and MebZIP5) Confer Disease Resistance against Cassava Bacterial Blight. **Frontiers in Plants Science**, v. 8, p. 1-11, 2017.

LIU, J-J.; STURROCK, R.; EKRAMODDOULLAH, A. M. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. **Plant Cell Reports**, v. 5, p. 419–436, 2010.

LIU, X.; HUANG, M.; FAN, B.; BUCKLER, E. S.; ZHANG, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome Wide Association Studies. **Plos Genetics**, v. 2, p. 1-24, 2016.

LU, H-C.; LIN, J-H.; CHUA, A. C. N.; CHUNG, T-Y.; TSAI, I-C.; TZEN, J. T. C.; CHOU, W-M. Cloning and Expression of Pathogenesis-Related Proteins 4 From Jelly Fig (*Ficus Awkeotsang* Makino) Achenes Associated With Ribonuclease, Chitinase and Anti-Fungal Activities. **Plant Physiology and Biochemistry**, v. 56, p. 1-13, 2012.

LU, S.; EDWARDS, M. C. Molecular Characterization and Functional Analysis of PR-1-Like Proteins Identified from the Wheat Head Blight Fungus *Fusarium graminearum*. **Phytopatology**, v. 108, p. 510-520, 2018.

MAYER, A. M.; STAPLES, R. C. LACCASE: new functions for an old enzyme. **Phytochemistry**, v. 6, p. 551-565, 2002.

MCCOLLUM, T. G.; DOOSTDAR, H.; MAYER, R. T.; MCDONALD, R. E. Characterization of chitinases and beta-1,3-glucanases in grapefruit flavedo during fruit development. **Physiologia Plantarum**, v. 99, p. 486-494, 2006.

MEINS, F.; NEUHAUS, J. M.; JR.; STICHERT, L.; BOLLER, T. A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. **Plant Biology**, v. 88, p. 10362-10366, 1991.

MELNIKOVA, D. N.; FINKINA, E. I.; BOGDANOV, I. V.; OVCHINNIKOVA, T. V. Plant Pathogenesis-Related Proteinss PR-10 and PR-14 as Components of Innate Immunity System and Ubiquitous Allergens. **Current Medicinal Chemistry**, v. 24, p. 1772-1787, 2017.

MELO, F.; FEYTMANS, E. Assessing Protein Structures with a Non-local Atomic Interaction Energy". **Journal of Molecular Biology**, v. 277, p. 1141-115, 1998.

MENEZES, S. P.; SILVA, E. M. A.; LIMA, E. M.; SOUSA, A. O.; ANDRADE, B. S.; LEMOS, L. S. L.; GRAMACHO, K. P.; GESTEIRA, A. S.; PIROVANI, C. P.; MICHELI, F. The Pathogen-Related ProteinsPR-4b from *Theobroma cacao* presents RNase activity, Ca²⁺ and Mg²⁺ dependent-DNase activity and antifungal action on *Moniliophthora perniciosa*. **BioMed Central Plant Biology**, v. 14, p. 1-21, 2014.

MITSUHARA, I.; IWAI, T.; SEO, S.; YANAGAWA, H.; HIROSE, S.; OHKAWA, Y.; OHASHI, Y. Characteristic expression of twelve rice PR1 family genes in response to pathogen infection, wounding, and defense-related signal compounds (121/180). **Molecular Genetics and Genomics**, v. 279, p. 415-427, 2008.

MOHAN, C.; SANTOS JUNIOR, C. D.; CHANDRA, S. In silico characterisation and homology modelling of a pathogenesis-related protein from *Saccharum arundinaceum*. **Archives of Phytopathology and Plant Protection**, v. 56, p. 199-216, 2020.

MUTHUKRISHNAN, S.; LIANG, G. H.; TRICK, H. N.; GILL, B. S. Pathogenesis-related proteins and their genes in cereals. **Plant Cell, Tissue and Organ Culture**, v. 64, p. 93-114, 2001.

MYAGMARJAV, D.; SUKWEENADHI, J.; KIM, Y. J.; JANG, M. G.; RAHIMI, S.; SILVA, J.; CHOI, J. Y.; MOHANAN, P.; KWON, W. S.; KIM, C. G.; YANG, D.-C. Molecular characterization and expression analysis of pathogenesis related protein 6 from *Panax ginseng*. **Russian Journal of Genetics**, v. 53, p. 1211-1220, 2017.

NASIR, F.; TIAN, L.; CHANG, C.; LI, X.; GAO, Y.; TRAN, L. P.; TIAN, C. Current

understanding of pattern-triggered immunity and hormone-mediated defense in rice (*Oryza sativa*) in response to *Magnaporthe oryzae* infection. **Seminars in Cell and Developmental Biology**, v. 83, p. 95-105, 2017.

NELSON, D. L.; COX, M. M. *Princípios de bioquímica de Lehninger*. 6. ed. Porto Alegre: Artmed, 2014.

NIKI, T.; MITSUHARA, I.; SEO, S.; OHTSUBO, N., OHASHI, Y. Antagonistic effect of salicylic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounded mature tobacco leaves. **Plant and Cell Physiology**, v. 39, n. 5, p. 500-507, 1998.

NISHI, H.; SHAYTAN, A.; PANCHENKO, A. R. Physicochemical mechanisms of protein regulation by phosphorylation. **Frontiers in Genetics**, v. 5, p. 1-10, 2014.

NUSRAT, Y.; SALEEM, M.; NAS, M.; GUL, R.; REHMAN, H. M. Molecular Characterization, Structural Modeling, and Evaluation of Antimicrobial Activity of Basrai Thaumatin-Like Protein against Fungal Infection. **BioMed Research International**, v. 2017, p. 1-10, 2017.

O'CONNOR, S. E.; IMPERIALI, B. Conformational Switching by Asparagine-Linked Glycosylation. **Journal of the American Chemical Society**, v. 3, p. 643-649, 1999.

OHASHI, Y.; MATSUOKA, M. Induction and secretion of Pathogenesis-Related Proteinss by salicylate or plant hormones in tobacco suspension cultures. **Plant and Cell Physiology**, v. 28, p. 573-580, 1987.

OIDE, S.; BEJAI, S.; STAAL, J.; GUAN, N.; KALIFF, M.; DIXELIUS, C. (2013). A novel role of PR2 in abscisic acid (ABA) mediated, pathogen-induced callose deposition in *Arabidopsis thaliana*. **New Phytologist**, v 4, p. 1187–1199, 2013.

O'LEARY, S. J.; POULIS, B. A.; VON ADERKAS, P. Identification of two thaumatin-like proteins (TLPs) in the pollination drop of hybrid yew that may play a role in pathogen defence during pollen collection. **Tree Physiology**, v. 12, p. 1649-59, 2007.

OSMARK, P.; BOYLE, B.; BRISSON, N. Sequential and structural homology between intracellular Pathogenesis-Related Proteinss and a group of latex proteins. **Plant Molecular Biology**, v. 38, p. 1243–1246, 1998.

PARK, C. J.; KIM, K. J.; SHIN, R.; PARK, J. M.; SHIN, Y. C.; PAEK, K. H. Pathogenesis-Related Proteins 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. **The Plant Journal**, v. 37, p. 186–198, 2004.

PATTERSON, M. C. Metabolic mimics: the disorders of N-linked glycosylation. In: *Seminars in pediatric neurology*. **Seminars in Pediatric Neurology**, v. 12, p. 144-151, 2005.

PIETERSE, C. M. J.; VAN-DER DOES, D.; ZAMIOUDIS, C.; LEON-REYES, A.; VAN WEES, S. C. M. Hormonal modulation of plant immunity. **Annual Review of Cell and Developmental Biology**, v. 28, p. 489-521, 2012.

SIPPL, M. J. Recognition of errors in three-dimensional structures of proteins. **Proteins**, v. 17, p. 355-362, 1993.

RABINOVICH, G. A.; VAN KOOYK, Y.; COBB, B. A. Glycobiology of immune responses. **Annals of the New York Academy of Sciences**, v. 1253, P. 1-20, 2012.

RAKWAL, R; AGRAWAL, G. K.; JWA, N-S. Characterization of a rice (*Oryza sativa* L.) Bowman–Birk proteinase inhibitor: tightly light regulated induction in response to cut, jasmonic acid, ethylene and protein phosphatase 2A inhibitors, **Gene**, v. 263, p. 189-198, 2001.

ROY, S. W.; GILBERT, W. The evolution of spliceosomal introns: patterns, puzzles and progress. **Nature Reviews Genetics**, v. 3, p. 211–221, 2006.

SAITOU, N.; NEI, M. The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. **Molecular Biology and Evolution**, v. 4, p. 406-425, 1987.

SANTAMARIA, M.; THOMSON, C. J.; READ, N. D.; LOAKE, G. J. The promoter of a basic PR1-like gene, AtPRB1, from *Arabidopsis* establishes an organ-specific expression pattern and responsiveness to ethylene and methyl jasmonate. **Plant Molecular Biology**, v. 47, p. 641-652, 2001.

SCOTT, H.; PANIN, V. M. The role of protein N-glycosylation in neural transmission. **Glycobiology**, v. 24, p. 407-417, 2014.

SELS, J.; MATHYS, J.; CONINCK, B. M. A.; CAMMUE, B. P. A.; BOLLE, M. F. C. Plant pathogenesis-related (PR) proteins: a focus on PR peptides. **Plant Physiology and Biochemistry**, v. 46, p. 941-950, 2008.

SESSA, G.; FLUHR, R.; RAZ, V.; YANG, X. Q.; EYAL, Y. Dark induction and subcellular localization of the pathogenesis-related PRB-1b protein. **Plant Molecular Biology**, v. 28, p. 537-547, 1995.

SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T. J.; KARPLUS, K.; LI, W.; LOPEZ, H.; MCWILLIAM, H.; REMMERT, M.; SODING, J.; THOMPSON, J. D.; HIGGINS, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. **Molecular Systems Biology**, v. 7, p. 1-6, 2011.

SHI, J.; DU, X. Identification, characterization and expression analysis of calmodulin and calmodulin-like proteins in *Solanum pennellii*. **Nature**, v. 10, p. 1-17, 2020.

SINGH, N. K.; KUMAR, K. R. R.; KUMAR, D.; SHUKLA, P. KIRTI, P. B. Characterization of a Pathogen induced thaumatin-like protein gene AdTLP from *Arachis degoi*, a wild peanut. **Plos One**, v. 12, p. 1-18, 2013.

SVENSSON, B.; SVENDSEN, I. B.; HOEJRUP, P.; ROEPSTORFF, P.; LUDVIGSEN, S.; POULSEN, F. M. Primary structure of barwin: a barley seed protein closely related to the C-terminal domain of proteins encoded by wound-induced plant genes. **Biochemistry**, v. 31, p. 8767-8770, 1992.

TAKEDA, S.; SATO, F.; IDA, K.; YAMADA, Y. Nucleotide Sequence of a cDNA for Osmotin-Like Protein from Cultured Tobacco Cells. **Plant Physiology**, v. 2, p. 844-846, 1991.

TAKEMOTO, D.; FURUSE, K.; KAWAKITA, K.; DOKE, N. Identification of chitinase and osmotin-like protein as actin-binding proteins in suspension-cultured potato cells. **Plant & Cell Physiology**, v. 38, p. 441-448, 1997.

TAKKEN, F. L.; GOVERSE, A. How to build a pathogen detector: structural basis of NBLRR function. **Current Opinion in Plant Biology**, v. 4, p. 375–384, 2012.

TATTERSALL, D. B.; VAN HEESWIJCK, R.; HOJ, P. B. Identification and characterization of a fruit-specific, thaumatin-like protein that accumulates at very high levels in conjunction with the onset of sugar accumulation and berry softening in grapes. **Plant Physiology**, v.3, p. 759-769, 1997.

TELLIS, M.; MATHUR, M.; GURJAR, G.; KADDOO, N.; GUPTA, V. Identification and functionality prediction of Pathogenesis-Related Proteins 1 from legume family. **Proteins**, v. 85, p. 2066-2080, 2017.

THIPYAPONG, P.; HUNT, M. D.; STEFFENS, J. C. Antisense downregulation of polyphenol oxidase results in enhanced disease susceptibility. **Planta**. v. 1, p. 105-117, 2004.

TONON, C.; GUEVARA, G.; OLIVA, C.; DALEO, G. Isolation of a potato acidic 39 kDa β -1,3-glucanase with antifungal activity against *Phytophthora infestans* and analysis of its expression in potato cultivars differing in their degrees of field resistance. **Journal Phytopathology**, v. 150, p.189–195, 2002.

VAN LOON, L. C.; PIERPOINT, W. S.; BOLLER, T.; CANEJERO, V. Recommendations for naming plant Pathogenesis-Related Proteins. **Genetic Resources**, v. 12, p. 245-264, 1994.

VAN LOON, L. C.; STRIEN, E. A. The families of Pathogenesis-Related Proteins, their activities, and comparative analysis of PR-1 type proteins. **Physiological and Molecular Plant Pathology**, v. 55, p. 85-97, 1999.

VAN LOON, L. C.; REP, M.; PIETERSE, C. M. J. Significance of inducible defense-related proteins in infected plants. **Annual Review of Phytopathology**, v. 44, p. 135-162, 2006.

VAN LOON, L.C.; VAN KAMMEN, A. Polyacrylamide disc electrophoresis of the soluble leaf proteins from *Nicotiana tabacum* var. "Samsun" and "Samsun NN" II. Changes in protein constitution after infection with tobacco mosaic virus. **Virology**, v. 40, p. 199- 211, 1970.

VLASTARIDIS, P.; KYRIAKIDOU, P.; CHALLOTIS, A.; PEER, Y. V.; OLIVER, S. G.; AMOUTZIAS, G. D. Estimating the total number of phosphoproteins and phosphorylation sites in eukaryotic proteomes. **Gigascience**, v. 6, p. 1-11, 2017.

WARNER, S. A. J.; SCOTT, R. DRAPER, J. Characterisation of a woundinduced transcript from the monocot asparagus that shares similarity with a class of

intracellular pathogenesis-related (PR) proteins. **Plant Molecular Biology**, v. 19, p. 555–561, 1992.

WITTKOPP, P. J.; KALAY, G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. **Nature Reviews Genetics**, v. 13, p. 59-69, 2012.

WOODS, T.; PREEPREM, T.; LEE, K.; CHANG, W.; VIDA KOVIC, B. Characterizing exons and introns by regularity of nucleotide strings. **Biology Direct**, v. 6, p. 1-17, 2016.

WU, J.; KIM, S. G.; KANG, K. Y.; KIM, J-G.; PARK, S-R.; GUPTA, R.; KIM, Y. H.; WANG, Y.; KIM, S. T. Overexpression of a Pathogenesis-Related Proteins 10 Enhances Biotic and Abiotic Stress Tolerance in Rice. **The Plant Pathology Journal**, v. 6, p. 552-562, 2016.

XIE, Y. R.; CHEN, Z. Y.; BROWN, R. L.; BHATNAGAR, D. Expression and functional characterization of two Pathogenesis-Related Proteins 10 genes from *Zea mays*. **Journal of Plant Physiology**. v. 167, p. 121–130, 2010.

YAMUNA, C.; VARMA, P. K.; PRASAD, R. D.; LAKSHMI, K. V. Morphological and molecular characterization of anamorph associated with gray mold of castor (*Ricinus communis* L.). **Journal of Oilseeds Research**, v. 32, p. 63-67, 2015.

ZHANG, J.; WANG, F.; LIANG, F.; ZHANG, Y.; MA, L. WANG, H.; LIU, D. Functional analysis of a pathogenesis-related thaumatin-like protein gene TaLr35PR5 from wheat induced by leaf rust fungus. **BioMed Central Plant Biology**, v. 18, p. 1-11, 2018.

ZHANG, Y.; WANG, X. F.; DING, Z. G.; MA, Q.; ZHANG, G. R.; ZHANG, S. L.; LI, Z. K.; WU, L. Q.; ZHANG, G. Y.; MA, Z. Y. Transcriptome Profiling of *Gossypium Barbadense* Inoculated With *Verticillium Dahliae* Provides a Resource for Cotton Improvement. **BioMed Central Genomics**, v. 14, p. 1-18, 2013.

ZHU, L.; ZHANG, Y.; ZHANG, W.; YANG, S.; CHEN, J. Q.; TIAN, D. Patterns of exon-intron architecture variation of genes in eukaryotic genomes. **BioMed Central Genomics**, v. 1, p. 1-12, 2009.

ANEXOS

Dados Suplementares 1. Sequências promotoras dos genes *Ricinus communis* L. Proteínas Relacionadas a Patogênese – RcPR com 1500pb a montante.

Gene RcPR-1a

>30099.t000028

TTAGGAATACAATACTTACCAGCTGCAATCTTTAATTTCCACCGATCATCATATTTTAAAAATAGAAACAA
ACACGGCCTCACAAAGTCATCTCTCTCGTAAGAACAAGATCATGTACATTTCAAGCTATGACCATAAAC
TGTCAAAACATTGAATTTAGACAACACTACAGTCATTTCTTTTCACTGCTGTAAAAATGTAATATTTAA
ATTTATTTAAAAATAATTTAAAAATTTTAAATAAAGATTTTAAATTCAAATCTTGGCCTGCCTCTGCGTTAA
ACGAGTTTCACAGTCATATCTTAATTTACATTTGGAGTACTTTCTTATGCATGAGATGAAAGAGTTACCA
TAACTGACATAAGCATGACAAGGTTTCCCTACCACATAATTAAGTCTGTTGTGAAACATCATAATTTGCA
TGGTCCCATTAGTATTATGTGTCTGATAATTGACTGTCTTCCCTGTGTTGGCAGGTTATTTATTGTAACC
TTATATTATATGATATGACAGTGACTTGGGTGTTTACATGATTTCACTATTTGGCGCGTGAATTTTTATT
TATTTGTTATTTTAAATCCTTTTACTCTATTTATCTTTTTCTTTTCTGATAACAGGTACAAAGAGTAATA
ACCTAATCTATAATCACCCTCATACCAAGTTGAACAAGAAATTCATTTCTGTTGATACTTCCCTTAATT
AATGGCTTAAAACTGACAAATAACCTAACTGTACTATGATTTGTAGAAGAAATTTAAACCGATAACAGT
AAACTTGACAAGTGGCAAATAACCCAATTACCATGATTTCTAAGAATCTATTCGTCTTGTGAATCTATAG
AATTCAGATAATGCCAAATAAAGCATCTTCTTCAATTATTTTCTCATCTTATACAATGTCAAGCATT
GCTACAAGCCAAGAGAAATGTTCCATAAAATAGAAGAACTGGTCGAGAAATTAAGATTTCTTTGTCCCT
TCTAAGTTTATAGATATGAATCTATTTATCCAAATCTTGGCAATAAATATGCACATTTAATATTATA
ATTTAATAATAAAATAAATTCAGTATTTACAATTTAAATATGATTATATCCAATCTATTTCTAATCCAC
ATAATGACGTATTTGATATTTCTAATTTGCTTGATTAATTTAAATATTATATAAATAATAAATGATACTT
AATATTACTTTATAAAGTGACTAAGGGCTCAAAACAAAATATGCAATTTATTTTATTCTTTAAAGATTTGT
TCATGAACAGACATCATTGTGATTTTTCTTTTTCTACAAATTTAGAAAAGCTTTCACCTAGGCAGTTAAA
CTAAGTCTTTAACGTGCAAAATAAAGAAAGATGTAAAGACGACAAAACCTAGCCATCAGACTATGCTATTTA
GTCATCTTCTTATAAAATACCCAAGTTTTTGTCTCCATCTTAATCAAGAATCTTAAACCTCTCCATT
TCCATGCCCTGCATAGAGATTTTCATAAAA

Gene RcPR-1b.1

>30099.t000029

TTTCTCTGCGGCCTGAATTATAAATTTTAAACAAAACCTTTCATTCAAAGATTAATAAAAAACAAACATAGT
TTTTAAGGATCAAATGCCAAAAGATCTTTAAGGGCCGATACAGGCTTTGCATAAACTCTAATATTTTTCG
TGTGCTCTTAATTAGAAATCATCAAGACCTTAAGAATCCAGAGTTGGTTTTTCTTGTCTTAACTTTGCTAT
CATTATCTCTGAATTTAGCAAGGCTAGGCAAAATTAATAAAGGGAAAACGATCCTAGATACCCCAACCT
TTAATCATTTTTCGCGAATCTATCTCACCTTTAAAACGATCAATTTGGTCCACCATGTTTGGCTCTTGT
ATCATTCCTACCCACCGTGTTTTGGCGTGTGCCAGTAATGACATGGCGTCCACACGTGGCTCCACATGT
TCGTCTACGTGGCTGAGCATAATGAAGGACCTCAAAAACCTCCCGTTCCTCCTAACATTTCCCAATTTCT
AATTTCTAAGGTTTACAATTTCTAGGGTTTATTTGTTAACTAAATGCTGCAGTGTTTGCAATGGGAAAT
TCTGCAAGAACCCAAATCTTATCGTAGCTGAAGACTTCTTTTCTGGATTAAACATCGCAAAAAGATATTG
CAAACCAAATGGAGTAAATGTCACTCTCTTAAATGTCGATCAAATTCAGGCCTTAATACACTTGGTAT
TTCTCTTGTCTGCTTGTACTATGCAGCAAATGGTGGCTTAAACCTCCGCACACTCACCTCGAGCAACA
GAAATCTTATGGTCATTGAAGGAACCTTTATGTGACTTTGTTACATCAAACCTAATCGCTTAGTCT
CAAAGTGTACACCCAGGAGATGTTTTCGTGTTCCCAATTTGGTCTAATTCACCTCCAGTTCAATCTTGC
AAAGAAGAAATGCAGTAGCTTTTGTGTTTGGAGCAGTCAAAATCCTGGCGTCATTACTATAGCTAATGCA
GTGTTTGGACCTAACCACCTATTAATCTGCTATTTTGGCTATTTAGCAATTTGGTATTGTTGGAAGTTA
TTAGTTATATTGTAACGTAAATCTTAGAAATTAGAACTAGGGATAGAATGGGGAGAAAACGATGAGTT
TTGAGGTACTTCAATATCTTCAGCCACGTAGACAAACATGTAGAGCTACTTGGACACCATATCATTAGTG
GCATACGCCAAAACCGGAGTGGATAGAAATGATACAAGACGCAAAACATGGTGGACCAAATGATACGTTT
TAAAGGTGAGGGTAGATTTCGCAAAAACGGTTAAAGGTGCGGGTATCTGAGATCGTTTTTCTCTTAAATAAA
TAAGTAATTATAGCTATGTGAGACTTTCTTCTGCATAGCATGTAAAATCTTTGTCACTGCTTTTGGGGC
ATTCTTCAGTGGGAACCTGCATGTTTTCTTGCACACTGGTTTTCTTTCTTATTCAAATTTAATTTAG
AACAATCTTTATATAAGCTTCTTTTGATTC

Gene RcPR-1b.2

>30099.t000030

GTCCATCAACCACTTGGTTGCTGTTCTTGGAGATGGGCGAGAAGGAGATGAAAGGCAGCTGAGTGGCTGC
AAGTTCACTGATTTCCCTCAAGAGCCATTTCCATTTACCCAAGACCCAACCAACTGTATTGTGCGAGTCA
GACTCCTTAATTTGTTTTCGAAGTTTCTGAAATTTGCTGCTCGAGGTAGTTAGCTGTATAGCCTTCAGGATTG
CGTGAAGCTAAGCTTCGTTGGATTCTTTAATTCAGTAGGACAAGAGAACAAGGAAATGAAAAATGGCCT
TCATTACTACGGAGGACGCCACCAATCCATGGGGGAGGCGTAGACGTTCCACTTAACCGAAGAGGCAGCT
GGTGCTTTCCAGACAACATTCACAGGAGACTTTCTTCTGATTCCCAGTCTTTCGCACCCATTAAGGAGA
AGACGAGTTGACTGCCGGAGTAGGGGAAGGCTGGCTCTACAGCATTGATCCATGAGGAAAAGCCATAATA

TGATTAAGTCAAACATTTACCCCAATTCGGCGTACAAGTGTGAAGGTAATGTCGTTTTCGGAATAACCA
 GAGGCTCCATAGGGAGGAGAAGAGAAGAGATTGACAGAAATTACACTGAAAGCGGCCCTTCAGCAAAAAAT
 AGGCCATTGCTCATAAAGATCGGAAAGGATTTGGGTGCACAACATTGATATCCCCACCAGCGACCAAC
 TGCACCAAAAATTTCCAAGTGATATTGCAATAATAATAGATTTAATTATGAATTGTCAAGATGATAATAAA
 TTTTTAAGATATTTTCGTTAATTATTGAATAAGTGACATATATTTCAAATATTTTATATTTGTAAATGA
 AAAAATAACATATATTTTAAATTTTATAATTATTTTATATGTATTTTTTATCCAATAATAAAACGTTATA
 TACTGAATAGTAAAGAAAAATTTCTATAAGATAAAAGTAATAACAATTTGTAGTTCAGTTTTACAATAGGT
 GCCACCTTTTTAAATAAAAATGAAAGTAAAGTAGAGCAATAAGCATAACATTATAGGCTAAAGAAAAATGA
 AAAAATCTCTGTTTTTTACCTTACAATGTCTGTAGATTAGGGATAATTAAACTCTGTACCCACC
 GCCATATTATCATAAATTCATAATCATTGATTTGTTTTTTCATATTTCTCATATGCAACGGTTTTGTTTAG
 AATCATGCAATTTCTATATCACCTGAAATATGATTTGAAAATACAAATGTACGCACATTCACCTGTCAA
 CCAGAAATAAAGGCAGATCAGATCCTACAGTTTACCAATGTAGAAACTTAATTTCTATCCAATAAATCT
 GTAATATTACTCACATCCACTTACTTAGTTGCTAAATTTAGCATCGTAAACTGAGCCATAGGCATGATTT
 CCCCTATAAATATACAAGTTAATCTAGAGTGCACCTCTGCTGAGAATCTCAATCTTTTCTACTCTCTGTCT
 CCCTAGCTTTGCCCCGGAGAGACAGAAAAAG

Gene RcPR-1c

>30099.t000031

TTCTTGTCTCGCTATTCCTTATTTTCATGATTTACTCCATTTGTTATATCATATAGATCACATATGTAT
 AAAAGATAGTATGTCATTCCAAAGATTAATAAAAATTTTGACTCTTCTTGGGAGACATGTCTCTGGAA
 TTTCTCACAGATGACTTTGGTTTATGGTAATCCAACCTGGGGTTACAATCAATAAAGGCTAAAGTTAAAAT
 TTGAGTAACCTTTTAAACTTTTAGCTCTAGCTCAAATAATAAGAATTTTAAATTTTAAAAACAAAGAT
 ATGGAGCTCAATTCAGTTAATAAATTTAGCTTTTATGATAAGAGCACTAGGATGATTATTGTATTCTCAT
 CATTTTATTTTTTCAATATATAGAAGTCAATATCTGGATTCTAATATTTCAATTAGATAAATGGCCATA
 GGCTTTAAATATACACTGAATTATAAATTTATTTAACTATTAATTTTAAATAAATTTGCATATCGCTTTT
 ATACGATTATTGGTATTGGTTTGGCACATTTAAATGTTAAAATTTTTTAAATATATATATATATAATAT
 TTAATAAATTTAAATATTTAATTATTAATTTTAAATAAATTTACACACTATTTTATAATTATTTTAAATA
 ATGATAGATATCATTTTTCAGTTTTACATTGATATTTCTTTCCTTTACTGTTTGCATTTTAAATACATGAA
 ATGGTACATCTCCACCATTCTAATTAACAAACAGTTTCCATAAGAAGTGAAGAAAATCATTTTCTTTC
 CAATTTTCCAGTATCCTTTTTTGCAATGTAGTTAATAGAAAATTTAATTTATCCATTCAAGAAAATCAA
 CTTATCCTTGAGTAACTCTTGTACATTTCCGAAGTCTTAATTAATTCAGTGTCCACTCGATATGAAGT
 GACATCAATGTGACTTTTCCATGGAAGTGTCTTAGAATAACTAACTACAGTTTCCCTCCATTCAATACAAGT
 AATTAATTTTTTTTTTCCACGCCTTTCTTTTCAATGTATAGAATTTCCACACCCTTACGCCTATAAG
 TACCAATCAACATGTTATGGCTTTTCCACCATTTGCCTAACTTTCTTGTGTATTGTAAAACAAAAAGG
 AAAACAAAATGTTATTATTAATAAAGAAAAATGTTATTATTTGGGAATGAAGGTTTCGATGGCTTTAATAC
 ACGTAGATGCCCTTAGGAAGATGATATCTCTAGCTTTAAGAATTTGTTATATTATTAATAAATCTT
 GTTGTTCATTCAAGCCAAATGACCACAAGGTGGAGAATAGAATTTTGAAGCCAGCATTTTCTTTGAA
 ACTTATTGCAAATAAATGAAACCCAAAAGTTGAATAAAGTGCAGAAATGCTGTGGGATACACCACCAT
 TCCATTTCTTAGAAGGACAATGGCATAATTTAGGCAAAATGGGCATATTTGATTTCTTCTCAGAGATTAT
 TTCTACAATTAATACCAGTATTTATTTTCA

Gene RcPR-1d

>30099.t000032

ATCTTTGGTTTTCACAGATGTTATAGATAGTGCAGGATTTCTTTGATGAATGAGAGAAATCAATGAAGAAA
 AATGTTTGAATCAACAGTTAAAGCTTCTATGTATTTGTAGACTTATTGTAGCAAGTTCTTTGCTTTAATA
 AAGAAATGAGTTATTTAAGTTTGGTTATAGTATCTTAAGAAAATGTTCCAAACAAGATATCTCTCATTTA
 ATTTGGGTGCAATAAATGAAATTTTAAATTAATAACTAATTGAGTTATTTCTCTTTTGATTATTTTTTCTT
 TTTATAATAACAAAATGGATCCTTCCAGCATATAAAATAAATTTCAACGGATTTTATTACTATAACCT
 TCCCAACTACAATTTTACATGTTTAGCAAAAGGTTAATTTATATATTGGACTAGTAGGCCAAAAGTGAA
 GATTAACGGTTTCAAATCTCAGCTCAAACAGGAGAAATATAATGTGCCCTTGATACATTTTATGTATTTA
 GATTATGGGTCTGAGTTTGGGGAGATAATGATATAAATATGCCACAATCTAAATACATAAAAATGTATCAA
 GTACATTTTAATTATAAAAATTTAATTTGTTTCGTTTTAGTTATTATTTCAATTAACAATTTAGTCACATTACA
 GTAGAAAATGAATGTACATTTTTTCTTATTAAAGTTACAACCTAATACTATTACTGAACCTTTGCATTT
 AGTTTCATTTTAGTTACAAGTTTTAATTTGTTTCATTTAGTCAATTCATTAACAATTTAGTCATATTGT
 CAGTAGAAAATTAGCATGTACATTTTTTCTTATTTAATTTACTCTTTTTTATATTGATTGTTTTTGGCACGT
 GTTAATATTTTATTTGTAAAGTACTAAAATTAATAAATAAATAAATTAACCTGGATAGCTGAAATGAAAAAAA
 ATTACAGTTTATGACTAAAATAAACTAAAATGAAAATTTGTAGAGATGATAAATTTGCTATCAGCTCTT
 TGGGAGTCTAAGAATGATTAGTCCGGTTCAATTTGGAATAAGATTCTGATGTCTAAAAGAATTCATGGGT
 CTTATTGTTAATTATATGTAATTTAGGTTCTATTAGATAGCTTGGGGGTAAAATTTATGGCTAATTTAAG
 TTCTATTAGCTGGGCAATTAGCGAGTTAATCTGCTAATTTATGTACTTATTAGTTTAGGTTTAAAGGTTT
 TTAGGTTCAATTTATATGAAAAAAAATTCACCCAAAACCTCAGATGGAAGAATTATAAAAATTTAGAAAATA
 TTTATTTTACTGGAATATTTCTTTTATACATTTTCGACTTTTTTTTTGTTAGTTGTTTTTTTATACATTA
 CAACTAATCATATATAACATATAATGAAATAACAGGATTTTATCTCCCTTGCATGTTACATAAATTCATT
 ATTATATGTCATATAGTGATTGATTATAGTATATAAAAAGGATTATACGTCCAAATTCAAATGAAATATT

ATAATCTATCATAAATCATAATTTAGATAG

Gene RcPR-1Be

>30099.t000036

TCATGAATTGGATATGAAGAACATTTGAACCAACAAAATTTCAAATTTGATGATTAATTACACGCCTATAA
 TTTAAATTAATATTTTATTTAGTTAATTAATTGAATCTATGTAGCCAAAAGTTAATAATGTTTTAGTTA
 ATTGGTCTAGTTAAGATTTCAACTAAAAATTCATAACTATAAAAAATGATTCAGTACCAATTAATTTGC
 ACCAAATTAATATGTTATTTACTACTTGATTGCAAAAAGGAAAAAATAAAAAATAGAAGAGCACACACCA
 TTTGAAGATAGTGTACTTAGAATCTTACTCTTCATAATGATTTTCCATTGGCCAAGAATTTTTGTTGGGA
 TGAGTAGCTAGTTGATACTGTGGAGCCTAGTTTTATTTCTAAAAGCATATATATATATATATATATATA
 TATAAATTAGAGTGTGTTTATATAGCTCTTAACAGTGACAAAATAACTATTTTAAAAATTTAAAACAAC
 GCATACTTAAAGTTTAAATTCAAAATTTTAGTTAAGAGCTTTAACGTAACATATATATTTAAAATTTGAAT
 CAAAGCCTTAATTAACATATAAAAACTTTTATTATTTTCATGTACGTGTCCAAGCTTAATGCCTATAAAT
 ACCCATCAAACATATCATCTATATTTCACTCCACAAAGAAAGAAAAACAAAATGATATCGCACAAAGA
 CTTCACTAGCTTTTCATTCCTCTGATTGTTAGCTTAGCCCTAATAATCCTTCCCTCCACTTATGCAAACT
 CAGCTCGATTAGGTTGTGTAAGCAAGGTGTAGCAATGAACGAACCTCATTACTTGTAACTATGATCC
 TCCCGGTAACTATATTTGGCAGCGATCTAACCAAGTACTTTTAGTTTATAGTGTACAGCTTCCATTACC
 TGAAGCTGAAGCTGTTCTAATGCAATATATTTATATGTCTTGGCATCCAGCAATAAAAAATGAAATATGAT
 CTCTTATAATATCTATAAATTTGCCACTTCTGTGTTTTTGCATCTTTGATCCAGATATACATGTCTATAT
 AACAAACATTTTCCAAAATAAGAAAAAGTTTCAGTTAATCGGCTCTAACTGGATTAAGATTTGAGACCA
 TACAGCTTTGAAGAATTTGACCTATATATGATTATTTAATTTGCTCGTTTTACTAGAACACAAAATATATT
 CTTCTTTTCTTTCTTTCTTTGCTAAATAAATCAATGAAAGTTTAGAAATCGCAAAAATAATTTGGTGAAGAA
 TATGTTAGTTCGATAATAACTTAATAGTTGTGCTATTTGTTGATTTGCATAATAGTTTAAAGAAATATGTTA
 TTTGATAGGCCTGTGATTTAGGATGATAATAAGGTCAATGACAGCATGTGAATGTGCAGTCTACTGCAG
 TTCCAGAAGTAACTACTAGTACTGCATATTCAAAATAGTTCTGAACATTCCTATAAATACCCATCAAAGT
 TTTTCGCTTGATACATAACAAAAACCACAAA

Gene RcPRB1-3f

>27894.t000032

TGAGAGAAGAAGAAGAAGAAAATGGCGATGAGGGAAAAGAGAAGGAAAAAGGGAAAATCCAATTTTT
 ATGAAATCAAGGCTTAGGAAAAGGGTTTTGAGGTATCTTAGTAATTTGTAATGATTTTTTATTAGCAAAA
 TCAAAAAGAACTAGGAGATTTTAAAGAAGGAAGATCAAAGATGGAGTTTTAAAGTCCCCTAAGAGA
 TGGAATAAAAAATAAATTTGCTTGAGACTCTTGGGATTTGTTTGGCTTCAATTTTCTAATACCATACGTCA
 TATGCATGATATAGACTTTTGGACAGAAATTTTTTAAACGTTTTTAAAGATTAGATTGGATAAAGGAAA
 TGAAGATATGAAGTGAAGAAAAGTTTCAATATATAGCCTTGGCCATGGGCTAATGTATATATCAATGATA
 TTTTTAGATTGTACAAAATTTCTCTATAAGAATTATATTTTTATTTTAAATATATTTCCGAGGATGAGGTTT
 TTATTTTTGAAAAATTTCAAATTTTAAAAAAAATTAATAATTATTTATTTATTTGTTGTTATTTTTTTG
 AAAGTCTTAAAGATTTTTAGAAGATTTTTAGAAAGGAAAAATCAATAGCTAAGAAAGTTGAACTGAGGTT
 TGATTATAACTTAAAAGTTAAATACTTAAAGAAAACAAAAAGTAAAAAATACATGTAAATACTTAATTT
 CTATAATTATTTAATTTTTCTTATACAAATGTAAGTAATCTTTTTTAAATTTCTGTCTTAATTTATA
 ATTATTTATTTATATAGCTTTTATTAAAAAATATTATCATTTGGCTATTTATCAATTTATATTAATTTA
 ATAAAATTTGTTAAAACATAACTTACACTATATTTTCATATCTATAAATAAAAAATCTCTCAAACGTTTCG
 TGTCTTGATTTGAAAAAGAAATTCGTATATTGTCCATTTTAAATTTTTGTTTCTACACAATTTAAGTGG
 TATATATCCTTTTTAATAAAAATTTTTTAGAAAGAAATTTAGATAGACCCAGTCTATCACATCATCCA
 TAAATTAATAAGTTACATAATGAAATATGACTTTATTTATTAATTATTACATTTTTTATTGATTGGATTT
 TAATTATGAGTTGTTAATGGTAATAAATAATTCAACCAATAAAAAATGTAACAATTAATAAATAAAGTAC
 GGAAGTACGATCTTATTTTTAGATTAAGAATTCCTGAAAACATTAAGGATCCGAGTGGCAACATTTGTA
 GCATGAAAAATTCAGTTTAGAGAAAAGAAAAAGAAAATAGTCAGAGCTGAATACCATCATCCTTCACAT
 TCACTAACATGTTTTCCACCTCAAATCTCCTTCAAGTCCAAGCCTTAGCATGTGAAATGATTCCTTTTCC
 GCATGAAATTTATGCTTATGCATGATTGATGTATCCAGTATATAGATAGACCCATTAATATCAATAG
 ATCAATAGAAATTTTCTTAGATCAGAGCC

Gene RcPR-1g

>27766.t000007

GGAAGTGCCTGGGAAGGAGGGAAGGTTGAACCAAGGAATTGGCAGAGCTTAGCAAGAGTATATATGTGGG
 AAGACCTGAGACCTCTTTGAGGTGACTTTACGTTATGTCATCCAAGTTAGACCTAAGTCTCCTCGGCTT
 GCTAATCTTACCAAGTGAATCTTTTCTTTTCCCAATAGAAATGATAATTCATTTCAAATCAAGCTC
 ATTGGTGTAAAAGACACATATTTATATCTCCAACACTATTCTCACACACTTGCTCCAAAATAAGCCA
 AGACGTAGAATTACTTACCAATAGGTATACAAATATAAATTTATATACACCTGACATCTTAAATGACTGC
 TTTTTATTAACCAAGCTAATTTATTAATAAAAAAATTAATAAACTTTAATAAATTTATATAATTTCTTTT
 AAATATGAAATTTGAATAAACAAACCCAGATTGCAACAAAAGATGAAAACAACACATGAGTTTCCTAAT
 TCCTTATATAAAGTTGCGGTGCCACGTTTGTGTTCTAGATTCCTCTTTGGGTCGCAATGTATATAATAC
 CAACAATAATAGTTACGGGTGTAACATAAACCCAGCTCAGGAGCCATTTAATCAATGCACTCTATTA
 TGTACCTTTAAACAAAAGGGACAGAAAATTTGCAATACCCACTACCAATCACCTTTCTTTTGTACAAA

CAACACCTTCTAATTCCTTCTCTGCGAAATTTTATTCCCATTTTAAACAATGATTGGAACCTTTTCATCCA
CACACACTATCACTGCGTTTGCCTTCCCTTGACATTTAAACTTTTTCTTTATATAAAATTTTATTGGCCAT
ATGCCACCACAAATCATATCTAAATTTGATACATGATAAATATGGACCAGTAGATGAATTAGTATATTTT
CAAATACAAATTCGGGTATAAGTTCATATGATCTTGAGTTAATTCATAAATTAATGTTTTGGATATTAAG
AATTTGAATTATAGATTTTAAAGTTATGACAAAAGAAATTTGTTTGAAAACCATGTATATTTATTGTAGA
AACAAGAATCTGAAACAAGGTTAATTCAGTGGTAAGAATATTGATTTTCAGAAGTGTGTTTTATTTT
TATATTAATTAACAAAAATTTATACTAGTACTCAACAATATATTTTAAAAAAAATTAATATCTCAAATAT
ACATATTTAAGAAAGTAACAGCAACAACCTGATAAACTGCCATGACAGAGTGCAGCGGTAAACGAGA
AATAATATTTAAGACCAACAAATTTGAAAAATAATGTACAATCATGACCTTAGATACCGATAAAAATAA
ACCCATACCACAAACCGCCAGATTTAAAAAACCAATTTAAAATGTCATGGCAGATCTTTTAAGCATGA
ACACACATTTGTGCTCTTCTATATAAAGGCACAGCCCCTCTAAAATTAGCGCCACAACTAATACGCTTC
CCTAAAAGAAGAAGAAAAGAAAAAGAAAAG

Gene RcPR-1h

>27766.t000008

ACTTAACTGTTTCATGTGTTAAATAAAACCATTTTTGCAAGTCTGGAAATATGGAACCAAGATACAGTTA
AATTCCTTATAATTAATATTTAGTCAATTGAGTGTTTTGATGATAAAGTGTTTTTCCCTTGAAATCA
AATTTCTGTATGATGAGTATATAAATTTTAAAAGATAACATGGATTTTAAAGCCAATTAATAAAAAGTA
ATATTCGCATATATATATAGCAATGCTGCAACTCTAGTGCCTGCATAGTTTTGGCTTCAGTTTTCTGG
CATTTAAGTCTCGGCTCATCTTTGCCGATAGGGGTATCTACCTGCTTTTTTTTTTTGTTTTCTGCTTGT
TCAGATAGTAATGATTGTTAAACAAAGTGAACCTGTTAAAAATCAAAGGTTAAAAAGTGCCAATTTAAGCCCT
ATGAATTTCAAGACAATGGCCAATTTGCTTTTATTTTAAATATGATGATCAATTTATTCAGTAATTCGGCA
AAAATAGTCAATTCACCAAATCAGAAGATTTAACCAATAAAAAATTTAACGAACATATTTTAAATTT
AATTA AAAATTTAACTGAAAAATTTTTATACATGCTAACATATTTTGACTGGTAAAAATTTTTTAA
AACCAGTTATACTAACATATCATAAATTTAACATTGCTAAAAATTTTTACATTACGATATTTTATAT
TTTCTATACAAATCAGAAATTTTATGTAATATGCAACTCTATTTTTAGTTTTAAATATCAATATTTTTG
ATAATATATTAACATAAATTTATAAATATGATAAATTAATATTGTTATAAAAAATAAATTAATTTTATA
TAAATTAATATAATTTTAGAAATATTAATAATAAATTTCTATAACTATTAAGTCATACTTATATGTCTA
ATTATTTTTTTGTTAAATAATATTAGTAAATTACATTAATAATAAATAATTTGACTTTAACTTAATAGATA
TCTATTTCTATGTAATTAGCATTTAATATGAAGTTCAATGACCAAATTTGAACTTTTCTTATTTAGTAACT
CCTTTGATTGAGTTTAAATGACCATTTTTTTTTAAATTTGCCGAATAAATTTGACTAAAAATTTAAAATAT
GAGTTAACTGTTTCATTTGTTTCAAATTTAAATGGCAAATTTGGCATTTTAGCCTAAAATTTAAAGGTGTCTT
AGCATCTCCCTCCTAAAATAATCTAATTTGACCGGATACTTTGGATGTTGAGGAAGATATTAGTTCTAAT
CTAATTTCAAGACCGGAAACGAGAGACACATGGTTGAGATACAATTTTACCTGAAGGAAGCTGCAAG
TTGCAACAAAACCTCGCAGCCCAATTTGCCATTTGCATTAATACAAGGGAAAGTTACATATATATTTTACA
GCCCAAAATCTTCATGTTTATTTATACGGTCACGAAGCTGAAAGAATTTTGTTTATACGATTTTAAAAC
TTCAGCTCTTTCTTACTAAAATTAGTTATT

Gene RcPR-1i

>30099.t000027

CCTAGTTCACCTTGACATACTCTTCATCAGGGTCACTGCACAGCAAAACAAGAGGACAATGATCAACTTT
GATAAATAAAAAATGAAAAGGAAAAAACCTGATGATCCGAACTCATAGAAATCCCAAAATCCAAA
GCCAGTAGACTATGATACGGTACCTGTTGTTAAGGGTTGAGTTAGATAATTTGCACTGAGGTTGGAGCCAT
ATATGCAAACTAAGCTCTTAGCTGCAAGGAAGCTGTAGGAAAAGACTTGCTAAATATACGCAACCTTGAA
GGAAGTTGTAATTTTTTAGTAAGAGAGTAATTTCCAATAGTTTGAATAGCAACTGTAGTTGAGAATAAAA
CAAAATAAATTTTGTACATCTGAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA
TAATAATACATGAGATTTCTTACCTAGGCCTGGTGTATACACCATAACTAATAGGAGAGTGACCCGTATGT
ATGAGTGATTTATATTTTGGTATTGGATGATTGACACATACCCCATTTAATATGTAATTTATCGAT
TTAGTGTATAGGTGAGGACATACATCAAGTCTAGATAAAAAAGATTTCTAATATTTTTAGAATTTCTATT
TGGTATATTCGCCGATAGCTATCTCTTACGGCACAACATATATAAAAAGTTTGAATCGATAACTAGTTCA
ATTTTAAATCAATTGCTAATTTGATCAGTTCTATTTTAAATTTTTTTTTATTTAGCTGATAACTGATTTGATT
TATTTATTTATTTGGATCCTATTATTTTTATTTATATAAATTTTTATATTTTATAAATAATTTATTTGTA
AGTTTTTTAATAGTTAAAAAATATAAATTTATAAATTTACGATATTTAAAAATTTAAAAATTTTATTAGTA
GAATTTAAATTTGAATTTCTACTTTTTAAATTAATTTTTTAAATATTTTAAATAAAAAATATAATTTAGTAA
AAAAAGTAACTATAATATTTTAGTTATTTATAAATTTTTTTATTAATTTATATTTGAATATAATAAAA
ATAAATAAAAAATATTTAAAAAATTTACTCTCTCAATAGGCATTTAACATATTAATAACAATCACTA
ATAAATCTTATTAATGGTTTACTTTTTATGGCATAAAAACAAACCTCGTACCAGTTAGTTATATTTATC
AAGCAACTCAAAGAAAAGTAATATGATTAAGTTATCGGAGTAAAATTTATTTGATTTTATTCATGACATAA
ATTATAAATTAATTTGGGCCCCTTTATTTCTTAGACAAGAAACTTATTTGTATCCAAACATTAACATCTA
GTATTAACATCCATCAGTTTGTGTGCTAAATTTGTAGTTGTGAACGAATTTAAGGATAAGTTCTATT
TTTTTTTTTTCGTTTTTTGTCCTTAAAGTCTCTCTACCTACAAATACGTCAAATTAACCATATAAGAAC
CTCAGGTGTTCTTTGAAGAGAAAATAGAAA

Gene RcPR-4a.1

>29598.t000016

ATGAATTCACCCATACACCTTTGTAGAAGCATTACTTACTACCAATTAGAATAAGCTTCTACCATAAA
 TAAAAATTTCTGCATTAGAAAATCTTGGCTGGTCTTAGAATAATGACAAAGATTAATACAAACAAGGGTTA
 TTCCCAATTGCCTTCATCCTTGAATTCATGCAAGTAACTTTCAAATCTTATTCATTTTTAAAAAGTTTTTA
 AGACTAATACTTGGAGCATATAGACCCACCAATATAACCATTTAAAAAAAATAACATAATATAATTATG
 TGTTAGTATCAGTATCATATCTAATATCTATATATGATATTGACACTCTGTCTTGTAGAAAGTTCTGTAT
 TTTTTAATTTATAAATAGTTAATTTATTTTTTATTTTCGAAATTGATTCATTAATGTTTGATAAGAAAA
 ATAATTTAAACATAAATGAATCTTTTTATGTGACAACAAATTAAGATTACTAAACAAGTTAATAATAG
 TAGTAACAATAACAATGCAATTAAGGATAAATGTCAAAGCATTACATACATAAACAAACATTTCCAGGCTT
 TACAGTTTGAAGTTACATTTCTACAGATGTCTCTAGTGGAACTGGAAACATTTCTGGAGCTTTGATAAC
 TTAAAGGGACAAAGCAGAATCATCATCCTCATCAAGTGTGAAGTGAAACAGTGTAGAGTGACCAAATC
 AAAGCAAACACGCCAAGAAGTATCCATCCAAGAAGTTGTTGCTCAATCCAATGGAAGTCCAGTTCCTT
 CTGTGCTCATTCTTTCATCCACAAGTGCCATTGCTGGTCTGCGGTTGCAGCCACCGCTGATGCTGCTGC
 CACCATTGAGGCTACCATGCTTGTGTTTCCAGGCTCTTGTCTTTGCTCTGCCTTCTCTTATGCTGCAC
 CTAACCTGCACCTCTTGTAGCCATTACTGGCAACCTTATATCAAGAACACATTGAACATAAATTAATTA
 CATCCGAATTAATGGAATTAATTAATTTAAAAACCTGCACTATGCGCACATTTACTATCAAACATACT
 TTAATAATGATGATTTGTGATATACATCGTATTTTATGCATATGTTTTCCGAATTTACTTTTCATTTA
 TATTAATGGTCACTATATATGAACAATATATATATATAATAATAGATATTTAATGACTCAAAGCTA
 ATTAGAAAAGTTTACTTCTCTCATTATATTAATGTTGGAGTTCACATTTACTCAAATCGAAAGGGATGT
 ACGGATCTAAACATGGGTATTGTGATTCTCCAATGACTTTCAGATGGAACATAAATTTGATCAGCTGGCAC
 AGTTCAAGAATTAAGCTCGATTATTAAGCTACTGAGTTAAGAGACATAGTAATCTAAAAATCTATGATAT
 ATTATATAAGACGGCAGTGAATAATCATAATTCTTAGTAAAAGAATGAAGGATGAAGTGTACCAAGAA
 TAGGAGTTGCCTGAATCCTGGTTTTGTTGA

Gene RcPR-4a.2

>29598.t000017

ATGAATTCACCATATACACCTTTGAAGAAGCATCACTTGTCTACCACCTTAGAATAAGCTTCTACCATAAA
 TAAATATTTTCATGCATTGGAAAATCTTGGCTCGGAATAATGACAAAGATTAATACAAACATGGATTA
 TTCCCAATTGCCATCATCCTTGAATTAATGCAAATAAACTTTCTAATCTCATTCCTTTTAGTGACTTATG
 ACATTTGACTTAAACGCGCGGTGTGAACATATTTATATAGTCATTTTAAAAATAACTATTTCTGGAAAGA
 TCTAAGACATGATACTATTTCTTTGTTTGTGTTAATATCATATTTGATATGTGTTTCGACATAGACTCTA
 TCCTGTTAGAAATGTTAATGTCTCGTCGTTTATTGAATAGTTAAATATTTCTTTTTTAAAAATGATTCAT
 GTAGTGAATGATAAAAAATAAATTTATTAAGATTCACTCAACAAGTTAATAACAGTAGTAACAATAAT
 AATGCAATTAAGGATAAATGTCAAAGCTCCAGGCTTTACAATTTGAAGTTACATTCTACTAGTGTCTCT
 AGTGGAAACTGAAAGCATTTCTAGAGCTTTAATCACTTAAAGGGACAAAGCAGAATCATCATCCGTAAAT
 TAATATTCACATGCGATACATAAAAAATAGAAAAATAATTAATTTAAAGACTTGGCAGAAATGCACACATT
 CTACTATGAAAACATATATTTCAAAATATGATAATTGTAAAAGTGAAGATGACATATTTGTATTTACAT
 TTATACCTGAATTTCACTTTTGGCTCAAGGCACAGTTCCACAACATCATATATATCCGTAATTTTCAGCA
 CGCTTTGTAATGTTATTAGAAAATTTCAATATGAAGTACATTTTTTTATTTTAAATATACCCCAT
 AATTTGTTTGAATATGTTATATCAACTGGAAGAATTACATCTAATTTCTTTAAGTCTGGATTCCAATAA
 ACTTTAATTAATGCTACTAAATGTATACAAAAATTCAGCTTTACTATGAAGATTTGGTTAATTACCATT
 AGAGTCGTAGGGGGAATAATGATCTTATCTTAGAAGAGTTTACTTTTTCAGGCATTCATTGCCGCTCCTG
 GCAAATACACAATGCATCCCAATGCCACAATTAATGTTTAGTGTCTATTTTGGCTTGCATTCATGTG
 CCGAGTATGCACAAATGAACCTCGACTTGTAAACAGTTCAGTCTGTTAGTCCCTGTTTTCAAATCCCCAA
 TTTCCACATTTTAAACTTACTTTTTAGTATTTTCTCCTATAAATTCATGAATTCTCTCATGAAACAAAC
 CCAATACAGACGAATAGAATGGGGTAAGGGTAACTGTGTGCATCGTGTCTTAGCCTTTCTCATAGCTA
 GAGCCACTTGCCAAAGTGGCTCTAATGTGAGAGCTACATATCATTCTACAATCCTGCCAAAATTAATTG
 GGACTTGAGAGCTGCCAGTGCTTACTGTGC

Gene RcPR-4b

>29598.t000018

AGATAATTTTATGTTATGGTAATAAACTTCGCAAGCAAGGCTAGCATTGCACATGTGTAGCTTGAGGACT
 TGAAAGTAACTTATATCTTTATGAGTCAATGGCAGCAATTATTGGCTACTGAAAATAGAAAGGGTTCAT
 GCATTTCAAATAATTTGGAGAAGATGCATATAGCTTCTGGAAATTAAGAAGCAAAAAGAAAATTAACAGC
 CCACCTTTCTTACTACGTTAATTTCCACGTCGTTCAAGGTCAAGAAAATGAATCAGCTTGCAAGATT
 GATAAAGTAAAAAACAAGAATATAAAGACAATAAATTTATGTCTCAAGAACTTAACAGAATAAAGGAA
 AATATTAATTAATTAATTAATGTCATGGTTTATAAATTTTATAGTAATCCATGCTTTGCTACTTGAGAGGG
 CTATGAAGATTTATAGTATATTAAGTAAAAAGATCTTTGAGTCAGATGGCAGTGATTTAAGAATCTCA
 TCCTCCCATAAATATTTTAACTCCATTTATATCTAAAATTTTATATTTTCCGCGAGGTACAATCCATA
 TCCGCATATTTTTCAGCAGTTCTGAGAAAATATACAGAGGAATTGTTTCATCCACATTATAACCAATTAT
 GAACGGAAAAGAAATGATCTTATCTCAATAGAGTTTTACTTTTCTAACATTCATGCCGTATCCTGACTA
 ATACACAATGCATCCATCAATGCCAGAATTAATGTTAAGTGTCTATTTTGGCTTGCATTCATGTGCTGA
 CTATGGACAAAATGAACCTCGACTTATAACCATTCCAGTCGATAGTCCCTGTTCAAATCCCCAATTTCCA

CGTTCAAACCTNNNGTTTCAAACCTACTTTTTCAGTGAATTTCTTCTATAAAATTCAGATTTCTTTCACAA
 ATCTAACCCAATACAGACGCATAGAATGGGGACAAGAGTAACTGTGTGCATTGTGTTCTTAGCCTTTCTC
 ATAGCTAGAGCCACTTGCCAAAGTGGCTCTAATGTGAGAGCTACCTATCATTTCTACAATCCGGCCAAA
 TTAATTGGGACTTGAGAGCTGCCAGTGTCTTCTGCGCTACCTGGGATGCCGACAAGCCGCTCGCATGGCG
 TCAAAAGTATGGATGGACGGCCTTCTGTGCATCCGGTGGGTCAAGGTCAAGCGGCCTGTGGCAAATGCTTG
 AGGGTAAGAAAACATCTATGATCAAAGTAAAACCTTATATTTGAGTGTCTTGTCTTGTTCGTAGGGCAGG
 GCGGCAAGAGTCCCCTCTAGTAATCTATTCATTTTTCTGAAAAATTCAGCAAATATTTATTCTGGA
 TTTATTTAATCAAATTCATGTTAATTCGCTGTTAAAATGATATTGTTTCAGGTGACAAATACTGGAACAG
 GAGCTCAAGTGACTGTCCGGTAGTTGATCAGTGTAGCAACGGAGGGTTGGACTTGGATGAAGGCGTTTT
 CCGGCAGATAGACACCGATGGAAGAGGCAT

Gene RcPR-5a

>30170.t000403

TAGTTCATATATCTGTTAAGCCTAAATGAAAACGCTTAAGGACAACAATTCATCAATGATCCTGAAAGGC
 AAACGCTAAATAGCAAACAATTTTTATCCCATGATTGTGGGTGGGAATAATTACTTAATAATTAATATGA
 GCGGGAAATTAATTTGGAAAATGGAGAAGTAAAAGTAGTGAATCTAGGGGAGTGGAGGAAGCACGTGCGGT
 GTTTAAGTGCCTAAGGACGTGTGAGAGTGGGAGGTTGTGCTTTGCTGATGTTTGCCACGTGAGGACTCTT
 TTTTTTTTTTTAGGGGCACCACCTGGTTTCTCAGTGACTTGATACCCATTTCCGCAATTTCTGCTAAT
 AATCTTAATAATTAATGCTATTACTTATTTATTTTTATAAATTAATTAATTAATTAATTAATTAATTAAT
 TGCCCTTAGTTTTTCAATTAATGTTGTTCAATTTAGTCCCAAAAAAAGTTATCTTTTGGCTTCATTTGGATACT
 TAATTTGATTTGCTGGAGCATTATTAATCACTCGAGGTAAATGTTATTTTCTAGGAAGAGATACCAGAAAT
 GGAGACGACCTTAATTAGTGCAAACGGCTAAAACACATGGAGATTGCCCATCACTTTTGCCTTATAACAAA
 ATTAACCCTAATGTACACCAAATTTACATGAATCTTATTTCAATTTAATCCCTTGATCATGAAAAGGA
 TATGGTAACGGCCTGAAAAAATGGAGCGGAATCTACCAATATCTCAATCAATTAATTTTGTTCCT
 CGTTACTATTAGCTTCAGTAAGTTCAATATTTCTAGAAAACGTGATTATGTGAATTAACAAAGGACATGA
 GAAATATACAAAAGAAAGGAGGCGCAGGTAATATTTAATACTTGAAAGAAAATGTTTATATTCATAA
 GTATTGCTCACACAAATGGACACATTTTTTGGAGCATGAACAGGAAGCTAAAAATAAGAAAGAAAGAAAC
 ATAAAACATAAAGAGTAACCCATATAATTAATTTGATGGTCAGAAAGATAAATAAGAAATCTTAAAGTGG
 TCACGAAAAGATAGCAAGCAATAAAAAATGGAAAGAGCAAAAGACGAGTGCCCAAAACCAAAAAATATAAA
 TTCAGAGCTCTCTCTATAATGACTTCCATGACTCCTCTACCTACAACCACTCTAACCTGCCTTCTTTTAT
 CATTCCCTCTCAATTTGCTTCTCCACAACCTTCTCTTCTTCTTCACTATTACCCGCAAAAAGATTTCTAA
 GAGAAAATGGCGACTAAGATTAGCTCATGTTGCTGTTATCTTCTTCTACTTCTTCTCTTTAGGTAAT
 ATTCAGTTACTTTACACGCATTAATATATAAAATAGCCCTTCTTCTCTTCTTCTTCTTCTTCTTCTTCT
 GGTGATTGTTCAATTTAATTTTTCAAGAAATATTGGAATCTCTTCTTCTTCTGGGTATTTTGATTCTTGTT
 TTCGGAAAACATTATTGGAATTTCAAAGTC

Gene RcPR-5b

>30152.t000023

AACATAAACTAAATAGTTAATAGAGCATATATATTAATATATTTTATTAATATTAATTTAACAAATGTA
 AATATAATAACTGTCTTAGCATATATAAAAGAGATGTTACGGCAATGGCGCCGTCCGGTCCGGTCCCA
 GTTCCAGTTCCAATCAATTTTTGAAAAATAAAAAATTAATTAATAAATTAACCTTTTACAAGTTGAAAT
 TGAATTAGATAATTTAGCTCCGTATGAGGAGAAACCAATCTGAACCAGTCTTAGTTCTAAGCTGGTTC
 CGATTTTCTTTTATGATTCAATTTATCTAATTTATAAATTTTTTACCTGAATGTTTAAACTAATATAA
 ATAATGTTAGAATATATGAAATAAAAAAATAAAAAAATCATATTTAAATCCTAAATTCATTAGTAAATA
 AAAATAAAAAATAAAGTATAGAAGCATAAAGAGAAAGAAAATGTAGAATCAGTTAAGAAAAACTTAAATG
 GTTTCAATACTTTCTCATGAAAATAAATTTTTTTATTTTTTCTTGAAGACTGTTTATAGGTAAAATTTA
 TAAAATTACATAACATTTAGACTTTTTTTACAATTAACAAAAAATCATTAATAAAGAAAAATAAAAA
 GTATAAATAGATTAACAGGAGTTAAAACATAAATGAACAAAAAATTAATCTATAGAAAAGAAAAGAGA
 ATATAAATTTCTAATGTAACCTCAATTTAGTAAATAAAGAAAAATAAAAAATTTAGACAAATTTCAAATC
 TAATATATATAAAGAAATATCATATGAGAAAAATAAATAAATTAATTAATTAATTAATAAATAAATAAAT
 ATATATGTAATAAATAAATTTAAAAAATTAATTTGTTATAAATTAATAAATAAAGAGAAAAGAAATGAT
 TTTTCTAAATTAGTATTTATAAAAAAGTAAACAAATATCATTTTTTATGGCAAGTTGTAGTTAAATCAT
 TTCTTCGTTTCAGTTTTTTAAAAATATATATATATAGATTGCTTATGAGAGCTGCTTGTAGTGTCAATTT
 ATTTAATTTAATTTGTTTAAATTTTTTTTACTAATCTAGTAGAGTAACTCTGTCTGATAACTTCATTTAAAA
 AAAAAAGGAACATAAATGTCCTGAGAAGTTGCAGTTTTACAAGAATGACTTAACCATTTGATCAGTTCT
 AGCATAGAAAAGGTAAGGGATGTTGGAGCTTTGGTTACCTGTCAATTTGTCACAAAATTTACACTTGTCAA
 TACACAACTTGTACTAAGTTGGAATGCAACTTACCACACTATATAAGGTGATTTGTTGAACCTCCCTC
 TACAAACATGCCAATCTTCCACTCCCTTGTATATAAAACATAATAAGAGAAGGAGAAATCTCATTTAT
 CAAATTTCAATCTTAGCATCTACTCCAAATCAACAATCAGAACCCAGAACTGGAAAGAAAATTAAGA
 AGAAGAAGAAGAAGAATCATTGCACCTCAC

Gene RcPR-5c

>29669.t000034

TACTTAATGATGATATCTATTTACAGAATTTACAGATTAGGGATGATTTAAAACGCACTGGCCATTACAT

GTATGCTATCTGGAATTTCTTGGAGATATAACCTTACTGTTCTTTTTCTTATGTTCTAGTGAGGGACT
 AGTTAACTTTTAGTGTTTTTTCTTAATTAATAAACACCCATTGGGGATGGAACCTCAATTAATTAACTTTT
 TAGTGTTCGATTATGTTAATAATCGGATACCTATATGCACATAAAATTTAATTGATATATAAGAAATTT
 TTAGATAGATTTCGTATAATTAATAAAATTAATAATAAATACTCATATATTTTACAATTAATTTATATGTTT
 AGGCATTTAATAGATAGAAATCTAGTGTTTATAAAAAAAGTTCAGCAAAAAGCACTAAGACATCAACAA
 TCCATCATTATGGCCAAATTACATAATTAATTATAACCATATATTGTGCTATGGCATACTTCCATAGCAT
 ACATCACATATCACGCATAATTAGGGGATAAAAGTAATTATAAAAGTTCGTCAATCCATAATTCATCAA
 ACTCAGAAACTAAGATTCAGAAAGAAAAATAAAAAAAGAAAAACATAAGCACTATATCTAAAATCTAAC
 TAAGGCCTAAATTGAATGTTGTTTATTCCTTCTTTGGGCAATCTCAGTGAAGACAGACAACCTTTTCTT
 TGAAGATTTAATATAAAGAACTTTGTCAATGTAGGCGTCAACTTCTTTACTAGTCGCAGGTTGTTTATTA
 CCTGGCAAGCGAACTTCAATCTTGTGCATCCATGGGCGTCCAGGACAAATCAGAAAAGTTTCCCGAAGCGG
 AAGTAACAACCTTCTTAAAGTCTGGACCTTTGGCGTTGTCTTTGATGGTGGTCTTGTCTGCTTTCTGCCAT
 AATTAATAGCTTAATGTGGTCTTGGCTAGAATAAACTTGAACATAAATACATCCCTTTATGCTTGTAACTT
 ATATACACATTGTAGCCTGTCAATAAACAAAACCAATAAGAAGAGGATTACTAATCCTAATTAGAAGTGA
 CTTGGAATTTCTACTAATTAACCTTCTCTATTTAGGAGTAGAAAAGGGAATCATCCGTTTTCTGTCCCT
 TTTTAAGAATGAAATTCCTATAAGAATTAAGTCAATGAAACAATCCTAATTAAGCCAAGCCGAATATTTA
 AAAAAAAAATGAAGAACTTCAAAGGCTAGCTAACATTATCAAAGCTCATAGCCTCATATCCCTATTTT
 TAAGCCTTTCAGAGTAGACCTGCACAGATGCCTGGATCTTAACAATTCCTCCTTTACAAACACAAAAAT
 ATTTGGTCAAGAATATACCAAGTTAAACCTTTTGTATCTTTACGGATGTTTGTGGTAAGAGTAGCATGC
 ATTAATAATTACAATATGGAACAGCTTGAGATGTTGAACATAAATCAAGCGGACGTGAAGTTGCTGAG
 ACTAGATTTACTAGGTCAGGCTCTTCCCTCA

Gene RcPR-5d

>30116.t000002

AGGTGTTCTAGTACTAATTTGTTTTAAGAGTCACTCTAGTGTAAACAATGAAAAAGGAACAATATACGACGG
 AGATTAAGAATAAAAGAATAATCAAGTAATAATTAAGAAATATGTAGTTAGACTTGAAACATATATGA
 AACTTTTTAAAAAAGTTATATTTGCATAATCTTTATCATAATATTATTAAGTGGCTTAAGAATGTTTA
 GTCTTTGATCACTTGGAGTTTAGAGTTTGGATTCTTTTTATAATAAAAAATCTTATTACAATATATAACT
 GATATCTACTTTCAGATATTTGAAGTGTATAAATTTTTTAGGAGGAAACAATAAATTTACAGAAAAAT
 TAAATAAATATTTGTTTATTTATCGGACTGTAGGACCAACTGGTATTTCTGTTTATATTTGTAACCGATTCT
 AACAATCTTAAAAAATCTTAAATTAATAAAAATTATAAACCGGTCAATTGGTCACTGAAGCAATATGA
 TAGCCAAATTTTTTATTTTTATTTTTTTTTATGTTGGCATTGAACCTATCAAGATAGGTAATAAAGT
 AGGCATAAAAGATGGCTAACACACAAGAGCAATTACCTTGGCAAGTTTGTCTAGACATTTCTGCCACT
 GGCAATCTTAAACATAACAAAAAAGGTGATGCCTCCTGACTCTTGAGCGACCATTATAATATCTTTTCA
 GATGGAAACTACCATTATTGATTAAATATTTTCAAACACTACAAGAAGACAATTAGAATCATGTTTCTGC
 TACACATTAATATGTATATGTTTCATATACTTGCATCTATTTGGTTTCTTTATACTATAAAAATCCCTCA
 AAAGACTAAAACCTAATTTAGGTCAACATAATATAGTAAAATTAATACCCAAAAAATCTGTTTCTCAAT
 CCTTTTTATTTTTATACATTTACTGCATTTGCCATCTATAGAGCAGGAAAAACAGAAATTTCTGTGAAAGTT
 AGAAATGCAAAACATACTTAGTTAAAACATTATTTCAAGTTAAAATATATACCTACTGTCAGATTAAGCT
 TAAGCTTATGGCTGCAAAAGAAATGGTCAGCTTTGATGTTAATTAAGATGGGATCTTGTAGCAGGATGTA
 AACCCCAATCTAACTGAACTTTAAGACACTAACAAAATCAAATCATGATTAAGAGACTTGCTTTTGTCA
 CAAAATCAGTATACATATTGACAATAAAGGATAATAGCCAACTAAGGATCAAAGAATTGGCCAAAAG
 AAAAAAGGTTTAGCAATTAATTAAGAAATTAGAATAATTTGTTGGGAGGAAGGTATCCATTGGTTGGA
 CATGCAATTAAGCAAGGAAGCCTTTTGTTTTTGGGCTGTCTCTGTTTGTAGATTCCTCCCTTCTACACG
 TATTCCTTCTCCTTTCCCTTTATCCCATGCACTCATATAAAATCTTCACTCACTTCTAATGGCTTTTTCAC
 TTCATGTATACACTAAGCTGCAGAGGGCAC

Gene RcPR-5e

>29669.t000043

TTCGATATCCCACTACTGATTCTAATCAATATAAGGCATCAAGTTGCCCTTCAGATTGAACAGGATTTGC
 CGTAGTGAGTTGGCAAATCAACAAGCTCAGATTTTCGTTGCTTATCCGAGTATGGTAATACAAATTTAACA
 GTTGGCAGCCAACGAATTTATGCTATGATGTTCAAGCGGCGCCCTACAGCTACTGCTGATGGTTCTGT
 TTATTGAAGTACGTTTCTGTGTCATCAATCTCATTACTTGGTTAGCTTTTGTAAAACCTTCTTACACG
 GTGTTTTCCATTTACAAGAAATACACTTATATGATGAATATAACTAGATATTTGGTTATATGCATACCA
 ACTGTTTTATACCTTTACCTGTTTATGTAAGTTCATGGATGTTGCTTTTTCTGATCCATTTGTAAAATAT
 ACTGCAAGCACATAATTAATAATACACACTAGAGAATGCTTTATTTACAGAATCAACATCAATGGCAATC
 TTTACAGCTCTTTAGAGTCAAGATTCTGAACTAATGTTTCCCTAAACCCTATTCAATTTTCAAGTTCACTAT
 AGCCTTGAGAAATTCATCAAAGAAAACTCCACCCTAATAAAACATGATTAGTAGCTTCAACAGGTTTT
 TCCTCAAACGAAAATCTCAAGATATCATAATAGACCAAAATTAATAATTTGTGATGCAGCAAGAAAACCA
 TAAATAGAAAAATAATCAAAAATCACCAGCAGCTCATAAATTTCCCACTAATTTGGAACCTTAGATTATAT
 CAGTGTTCAGAACAATGAAAACCTGTTTCAAGAAGAGGTTTAATACTAAGAGAACCTAATCAAATGACT
 CGTATGTTGAAAATATCAATTTCAAAGCGTCTTCTTGACAGGAACCTGCCCTCCCTTCATATPATCCTTG
 TTAATAAAATTAATAGTCATATAGCAATCTAGTACCTGCCACTTATTCAGTACCAGGTTATTTTTAGAAT
 TCCAACATGATTTTGTCTAAAAGATTTAGTTTTCATAGTTATTATAATTTACCTGCCTCCCTTCAATGTAA

Dados Suplementares 2. Sequências dos genes. *Ricinus communis* L. Proteínas Relacionadas a Patogênese – RcPR. Depositados no *Phytozome*. 5' e 3'UTR indicados nas cores verde e rosa, respectivamente. Exons indicado na cor azul e Introns indicado sem cor.

Gene RcPR-1a

>30099.t000028 | 30099:134094..134579

ATGTATTTGAACGACTTCTTTGCAGCCACAGTTATGGTGGGAACAGTCCTAGTTCACATCTCTCTAGCCC
AAAATTTCTCCCAAGACTACGTAAATGCACATAACACAGTTCGTGCAGAGATAGGAGTCGGCCCAATTAC
TTGGAACAAAACAGTAGCAGCCTATGCTCAAACCTACGCAAATTCAGAATCGAAAAGCTGTGAATTTGAG
CATTATATGGACCCTATGGTGAAAACATTCAGAAAGGTTATGGTAACTTGAACGGGGTTGATGCAGTGA
ACATGTGGGTGAGTGAGAAGCCTAACTATGATTACGGATCAAATTTCTTGCAAAAGGTGGGGAAGATGAGTG
CTTGCATACTCAAGTTGTTTGGCGAAACCTCAGTTCATCTTGGATGTGGTAGGGCTAAGTGCAAGACC
GGATGGTGGTTTGTAACTTGTAAATATCATCTCTGTTGGGAATATGAGGGTCAGCGTCCCTTTCTGA

Gene RcPR-1b.1

>30099.t000029 | 30099:150206..150899

ATGAGCAATGGCCTATTTTTGCTGAAGGCCGCTTTCAGTGAATTTCTGTCAATCTTCTCTTCTCCT
CCCTATGGAGCCTCTGGTTATTCGAAACGACATTACCTTCAACACTTGTACGCCGAATTGGGGTAAATG
TTTTGACTTAATCATTATTAGGCTTTTCTCATGGATCAATGCTGTAGAGCCAGCCTTCCCTACTCCGG
CAGTCAACTCGTCTTCTCCTAATGGGTGCGAAAGACTGGGAATCAAGAAGAAAGTCTCCTGTGAATGTT
GTCTGGAAAGCACCAGCTGCCTCTTCGGTTAAGTGGAACGCTACGCCCTCCCCATGGATTGGTGGCGTC
CTCCGTAGTAATGAAGGCCATTTTTTCAATTCCTTGTCTCTTGTCTACTGGAATTAAGAATCCAACGA
AGCTTAGCTTCAGCAATCTGAAGGCTATACAGCTAACTACCTCGAGCAGCAATTTTCAGAACTTCGAA
ACAATTAAGGAGTCTGACTCGACAATACAGTTGGTTGGGTCTTGGGTGAAATGGAGAATGGCTCTTGAGG
AAATCAGTGAACCTGCGAGCACTCAGCTGCCTTTTCATCTCTTCTCGCCCATCTCAAGAACAGGCAACCA
AGTGGTTGATGGACTTGCCAAATCTGCAGTCGAAAGAACGCTGACTTTGTGGCCTGGCTGTGA

Gene RcPR-1b.2

>30099.t000030 | 30099:151756..152238

ATGGGGTTCAACAAGTTCATAGTAGCTCTATGTTTAAATGACATTAACCATAACTCCTGTCTCCTTAGCCC
AGAATTCAGAAAGGACTATGTTGCTGCTCATAATGAGGTACGTGCTGAACTAGGGCTGGGTTCCGGTTCG
CTGGAACGAAAATGGCAGCTGTATGCAGTAAATATATACAACCAAGGTTGAAACCTGCATACTAGAG
CATTCCAATGGCCCTTATGGAGAGAACCCTGGCGAAAGGTTCCGGTGAAGGCTTCAGCGGTGTGGACGCAG
TGAAATTTGGGCAGATGAGAAGCCTAACTACGACTACCTATCTAACTCATGTGCTGGTGGTATGTGTGG
ACATTACACTCAGATCATTTGGCGGGACACCAAGAGATTGGTTGTGCTAAAATAAGTGAAGGATGGC
TGGACTTACATTAGCTGCAACTATGACCCTCCAGGCAACTATATTTGGGGAGCGTCCCTTTCTGA

Gene RcPR-1c

>30099.t000031 | 30099:154297..155045

GAAAAAAAATGTTGTCACTCAAGATTTTCAGCAGCCCTTTCTTTTCTCATTAGCTTGTCTTAATCCTCT
CCTCTCATGCCAGAACTTACCACAAGACTACATCAATGCCATAACGCAGCTCGTGCAGCTGTAGGTGT
TGGGCCATGACTTGGGACAACACAGTAGCTGCCTATGCACAGAACTACGCCAACCAACGCATCAACGAT
TGCAGGCTTGTGCATTACAGTGGACGTTACGGGGAGAATTCATGATGGAGCAGTGGTGACCTTTCTGGGA
CAGCTGCTGTGAAGTTGTGGGTGATGAGAAGGCATTTCTATAATTATAAATTCGAATCTTGTGCTGCAGG
GCAGCAGTGTGGGCATTACACTCAGGTGGTGTGGCGAAACCTCGGTTCCGCTTGGGTTGTGCTAAGGTTAGG
TGCAGAAATGGAGGCATTTCACTTGAACCTATGATCCTCCTGGCAACTTTGTTAACCAGCGTCCCTT
ATTAGATATTCATGTTATCATAGCTTGCCTTTGTGATGCTTTTCGATTTCTCTATGAATAAGTCTAAAAG
AACTCGAGATTTCTGAGAAAATTATCGAGTTTCTACTCCTGTTTGGATGACAATGGGAGGATTACAGA
CAGAATAAGGAGAATAGATTATGAAACAAATTATAAAATTTGATTTTCATGCATGACTTGTTCATGAATA
TAAGTGAATCCGTTTCTCTATCAATAGTTTGTATTAACCTTTGCACCTCA

Gene RcPR-1d

>30099.t000032 | 30099:161783..162283

ATGGCGATGCCAAGGTTTCCCTACCTCTCTTTTATTTTCATGATCATGAGCTTAGCCCTACTCTTCTCCT
CTAATGCCCAAGACACGAACCAAGACTACATCAATTTCTACAATGATGCTCGTGTGCTGTAGGTGCTGG
ACTCGGCAACATGACTTGAACGAAACTGTAGCAGATTATGCACGAGATTATGCTAACCAACGATTTGCT
GATTGCAATCTTGTGCATTCGGATGGACCTTATGGGGAGAATCTTGCATGGGGTAGCGGTGATCTTTCTG
GCTTGGATGCTGTGAGAAATGTGGGTTGATGAGAAGGCATTTCTACGATTACAACCTCCAATACTTGCAGTGG
AGGCCAGCAATGTGGGCACTATACTCAGGTGGTTTGGCGGACTCCATTTCCCTTAGGATGTGCCAAGGTC

ACTTGCAACAATGGGTTAGGAACCTTAATTAAGTACTGTAACCTATTACCCCTCGGCAATGTTATTGGCCAGC
GACCCTATTAA

Gene RcPR-1Be

>30099.t000036 | 30099:181211..181699

ATGAAGCAGTTCGATCCCTCGTTAGCTCTTGTGGTATCATCTCTATAGCCCTAATCTTTTCATCTCATG
CCCAGGACACACAACAAGACTATCTTAATGCACACAACCTCAGCCAGAGGAGATGTTGGCCTCGGACCCCT
GACCTGGGATGACAAGGTTGCATCCTATGCACAGCACCACGCCAATCAACACAAAAGTGACTGCAGCCTT
GTGCACTCTGAAGGCCCTTATGGCGAGAAGTTGGCATGGAGCAGTGGTGACTTATCAGGAACAGATGCAG
TAAAATGTGGATAGATGAGAAGCCTTACTATGATTACAATTCTAATTCCTGTGCTTTCAGGACAGATATG
CGGGCATTACACTCAGTTCGTTTGGCGCAGTTCTGTTTCGATTGGGATGTGCTAAGGTTAGTTGCAACAAT
GGAGGAACATTCATCGGATGCAACTATGATCCTCCTGGTAACTATATTGGGCAGAGACCTTACTGGTGA

Gene RcPRB1-3f

>27894.t000032 | 27894:426397..426912

ATGTTATCATGTGGGGCACTAGGAACCTTCTTGCTGTTCTCTTGCTAGCTTCAGCCTCTGAAGTTTCTT
CTGGGCGGAGTTTCATTAATAAGAACCACAAAGAGAGACAATTAAGCAATATCTAAAACCCACAAACAGGGA
AAGACCAAAGCTAGGCCTTCGTCTCTAAAATGGAGCAACAAGCTCGCAAGCTTTGCGTCTTCATGGGCA
CACCAGCGACAAGGAGACTGTGCGTTACTTCATTCAATAGCAACTACGGGGAGAACCCTGTTTTGGGGCA
GCGGAAAGGATTGAAACCAGGTGATGCAGTGGCAGCATGGGCAGAGGAGAAGTGCTACTACAATCATAA
TACAAACTTGCCTAAGAATAAAGACTGCCTACATTATACCCAGATGGTTTGGAGGCAAAGTTTGAAG
GTTGGTTGTGCAAGAGTTATCTGCACAAGTGGAGATACGTTGATCACATGTAACCTATGACCCTCATGGCA
ATGTCATCGGTGAAAGGCCCTTTTGA

Gene RcPR-1g

>27766.t000007 | 27766:99107..99703

ATGAAGCCTCATTCCTTCTCTTGCTCTTCATCTTCTCCTCATCACCATATTCATAACCAACAATCTTCTGG
TGACATCTCAATCACCTTCTCCGCCAGCCGAAACACCGCTAAACAAGCGGAACAATGACACCATTTACAA
GGTCTCCAAGCAGCTATGCTGGGGTGCATTGGGGAGGCACTCCAGTTCCTGTACGCGCACAACTAGTG
AGGGCATCTAAGTGGGAGCTGCCACTGACATGGGACTCCCAGCTAGAGAGGTACGCCAGGTGGTGGGCTG
GAACAAGAAAACAGGACTGCCAGTTGGAACACTCGTTTCCAGAAGGTGATTTCAAGCTGGGGGAGAACAT
ATACTGGGGAAGTGGCACAGCCTGGACGCCAAGGGATGCTGTGATGATGGGCGAGTGAGGAGAAGTAC
TATACCTACGCAACAACTCATGTGAAGAGGGACAAATGTGTGGCCACTATACACAGATTGTGTGGAAGA
CGACAAGGAGAATTGGGTGTGCTCGTGTGTTTGTGATGATGGAGATGTGTTTATGACATGTAATTATGA
TCCTCCTGGTAATTATATTGGGGAAAAGCCTTACTAA

Gene RcPR-1h

>27766.t000008 | 27766:111436..111975

ATGCGTACTTACTGTTATTATTGTGTTGCTATCATCTCCATTTCTCTCTTCTCACAACAAGAGCCACTC
AATCTTCTGCTGCTGCTGAGTCTGCCATTGCCAATCATCAGCATGCCTACTTGAGCACTGCCAACCAATT
CTTGGGTCCCACAACGCCGCCCGTGTGCTGCACTAAGAATGCCTCCACTGATCTGGGATACTAGGCTAGCA
AATTATGCGCAGTGGTATGCTAACCAGAGGAGGTTGACTGCGACCTGAGGCACTCGAATGGACCTTATG
GTGAGAACATTTTCTGGGGGAGCGGTACTGGGTGGACACCGGCTCAGGCTGTCACCGCATGGGTCTCGGA
GCGGAAGTGGTATAATTACTGGTCTAATTCTTGCTACGGACATCAAGAATGTGGGCACATACTCAGATT
GTTTGGAGAAAACAAGAAGAATTGGATGTGCTAAAGTGACTTGCTCTGATGATTTGGGAGTTTTTCATGA
CTTGCAATTATGATCCTCCTGGTAACTACATTGGAGAAAGGCCCTTACTGA

Gene RcPR-1i

>30099.t000027 | 30099:131507..131998

ATGGGTTTACCAATTTGGTAATAGCTCTATGTTTGGTGGCATCAGCCCTATCTCCAACATGTCTAGCTC
GAAACTCCCAGGAGACTACCTTGATGCACATAATGCAGTTCGAGCAGAGGTAGGCGTTGACCCACTTGT
ATGGAACAAGACATTGGCAGACTATGCCAAGGTTATGCAAAAACAAAAATTAATAATTTTGTGTATGAG
CATTCAAATTAATCTTCAATTTGGTTATGGTGAATGTTTGGCAGCAAGTCTCAACCCAACTCAATGCCA
AAGATGCAGTGAAAATATGGCAGATGAGAAGAAATTTTATGATCGTAAATCAAATTCCTGTAAAGGTGG
TGAATGTGGGCATTACACTCAAATCGTTTGGCATGACACAAGTCAAGTGGGTTGTGCTAGAGTTAAGTGC
AAAAATGGGCATACATTTATCAGCTGCAACTATTATCCCATAGGGAATGTTCAAGGTCAAAGCCCTTATT
GA

Gene RcPR-4a.1

>29598.t000016 | 29598:67883..68473 reverse

ATGGGGGTAAGGGTAACTGTGTGCATCGTGTCTTAGCCTTCTCATAGCTAGAGCCACTTGCCAAAGTG
GCTCTAATGTGAGAGCTACATATCATTTCTACAATCCTGCCAAAATTAATGGGACTTGAGAGCTGCCAG
TGCTTACTGTGCTACCTGGGATGCCGATAAGCCACTCGCATGGCGTCAAAGTATGGATGGACGGCCTTC
TGTCATCCAGTGGGTCAAGGTCAAGCAGCTTGTGGCAGATGCTTAAGGTTAAGAAACATTTATATGATGA

AAGTGAACCTTGTATTTGAGTGAGCAGGGCGGCCAAGAGTCCCCTCCTAATAATCTAAAAAATTCTG
 GCGAATACTTATTCTAGATTTTCTTGAATAAAATTTACCTTAACTCATTGTTGAGATGATATATTTTCAG
 GTGACAAATACTGGAACAGAGCTCAAGTGACTGTGAGAAATAGTTGATCAGTGTAGCAACGAGGGTTGG
 ACTTGGATGAAGGCGTTTTCCGGCAGATAGACACCGATGGAAAAGGCTATGCTCGAGGCCACCTCATTGT
 TAACTATCAATTTCGTAATTTGTGGCGATTAG

Gene RcPR-4a.2

>29598.t000017 | 29598:69822..70438

ATGGGGACAAGAGTAACCTGTGTGCATTGTGTTCTTAGCCTTCTCATAGCTAGAGCCACTTGCCAAAGTG
 GCTCTAATGTGAGAGCTACCTATCATTCTACAATCCGGCCAAATTAATGGGACTTGAGAGCTGCCAG
 TGCTTTCTGCGCTACCTGGGATGCCGACAAGCCGCTCGCATGGCGTCAAAAGTATGGATGGACGGCCTTC
 TGTTCATCCGGTGGGTCAAGGTCAAGCGGCCTGTGGCAAAATGCTTGAGGGTAAGAAACATCTATGATCAAA
 GTGAACCTTATATTTGAGTGTCTTGTCTTGTTCGTAGGGCAGGGCGGCCAAGAGTCCCCTCCTAGTA
 ATCTATTCATTTTTCTGAAAAATCTAGCAAATATTTATCTGGATTTATTTAATCAAAATTCATGTTAA
 TTCGCTGTTAAATGATATTGTTTCAGGTGACAAATACTGGAACAGGAGCTCAAGTGACTGTCGGGTAGT
 TGATCAGTGTAGCAACGGAGGGTTGGACTTGATGAAGGCGTTTTCCGGCAGATAGACACCGATGGAAGA
 GGCATTGCTCAAGGCCACCTTATTGTTAGCTATCAGTTTGTAAATTTGTGGTGATTAG

Gene RcPR-4b

>29598.t000018 | 29598:71374..72176

ATGGGAAGTGTATTGATGTGCTTATGGTTGTGCTACTATGTTCACTAGGTGTTGCATTTGGTGAGCAAT
 GTGGCAGGCAAGCAGGCGGTAAGCTTTGCCCTAATAACCTATGTTGTAGCCAGTATGGATGGTGTGGCTC
 TACTGATGACTATTGCTCACCTTCCAAAACTGCCAAAGCAATTGCAAAGGTGGTGGTGGCGGTGAAAGT
 GCTTCTAATGTCCGTGTACATAACCTTCTATAATCCTCAGCAACATGGATGGGATCTGAATGCAGTTA
 GTGCATATTGTTCTACATGGGATGCTAACAAGCCTTATTCATGGAGGAGCAAGTATGGGTGGACTGCGTT
 CTGCGGTCTGCTGGACCTCGCGGACAAACCTCCTGTGGCAAGTGTGAGGTAAGAAAATTAAGAAAT
 TGCACAGTTTTTTAATCGCTCAAAACAACATTATTTCTGTTTAATAAAAATCGGTGCACTCGGATGTAC
 GAGCACACTGAATCATAGTTTGTATTTAATCTATTCCCTGATAATTAAGATTAATGTTATATATGCCTT
 ATTTATGTGTATACATATATATGTGTGTGTGTGTAGGTGACAAATACTAGGACTGGAGCTGCGACAACAG
 TGAGAATAGTTGATCAATGCAGCAATGGAGGACTGGATTTGGATGTGAATGTCTTTTCGTCAATTGGACAC
 AGACGGTATTGGATACCAGCAAGGTCACCTTATAGTGAATTAACAGTTTGTAACTGTGGAGATTCTTTC
 AACCTCTGTTTTCTATTATGGATGATCAGTAA

Gene RcPR-5a

>30170.t000403 | :2181957..2182730

ATGCATTACAAACTTATAAAATCTATACTAGCTATTATAGTGTGTGGCACAACCTTATATTTCTGTGTCA
 AATTTGCAGGTAACCTCAGCAATGCCACGTTTTTCACTCTCAAAACCATTTGATAGCTACATGATATGGCC
 CGGAACCTCTCCGGAACCGGCGCTGGAACCTCTCGGAGACGGTGGCTTTGCATTGCCACCGGGCTCATCA
 ATCCAACCTCAAGCTCCACCCACTTGGTCTGGCAGGTTCTGGGCTCGAACTGGGTGCACATTTGATGATT
 TAGGCAATGGCAAGTGTGACACCGGTGACTGCGGAGGCGCCTGAAGTGCACCGGCGGTGGAGCTCCCC
 TGTCAGCCTGGCTGAGTTCATATTGCACCAAGCTCTAATGAGAAGGATTTCTATGATGTTAGCCTAGTC
 GATGGATATAACGTTGGCTTAGGTGTCAAGGCATTGGGCGGGTCAAGTGTGCAATATGCGGGTTGTG
 TTGCGGATCTTAATGGGAATTGCCCGCCAGAGCTACGGGTTACGGATTCAGTTTCGGGTTGCGTTGTCGC
 TTGCAAGAGTGCTTGTGACGATTTAATGCGCCGAGTTTTGTTGACCGGCGATCATTGCACACCGCAG
 ACTTGCTCACCACAGTACTCGGCGTTGTTCAAGAACGCATGCCCTTCGGCTTATAGTTATGCTTATG
 ATGATGCATCAAGTACTTGCACCTGTTCCGGATCCGATTACATGATTACATTTTGTCCAACCTGGGTCTGG
 ATGA

Gene RcPR-5b

>30152.t000023 | 30152:873257..874051

ATGGCAAATGGCAGTTTCTTCTCACCTTGGTTTTGATATCATTCTTAGGTAACAAGAATAGTATGCTTC
 CCTTTTCTATAGTAATGATGTTTTGTCCATTATGTAATGATGTTAAGAGTCTGGGAATTTTTTTTGCAGG
 CAGTGAAGCTGCAATTTTACCCTCCAAAATAATGCAGGTCAACCATATGGCCAGGAACCTTTTCAGGT
 AAAGGACCTTTACTAATCAATGGTGGAACTAGATTAATAAACCAGGTCGAACAATGAGGATTAATGTACCTC
 ACAAGTGGTCAGGCCGCTTCTGGGCTCGTACCGGATGTAAGTTTGTATCAAACTGGCAAGGGAAAATGCGC
 CACCGGTGACTGTGGCGGCTTGAACATTTGTGGTGGGGCTGGCAGTGTACCACCAGTCACACTAGCTGAG
 TTCACCTTGGATGACCTAAACTTCTACGATGTTAGCCTTGTGATGGCTACAACATCCCTATTTCCATCT
 ACCCTTATGGTGGCTCTGGCGACTGCAAGAAGGTAGAGTGCTTTTCAAACCTAAATCCGAGATGCCCCAA
 GGATCTACAAGTGATAAATAATGGCAAGATTGTTGCCGCAAGAGTGCCGTCACAGCATAACAATACTCCA
 AAATATTGTTGCAGTGGTGCCTATGGTTCTCCTCAGACTTGAAGCCTACAGGTTTTTTCAGTATATTCA
 AGGCCGCTTGCTTACATATTACAGTTATGCATATGATGATGCAACAAGCACATTTACATGCAAAGACGC
 TAATTATAGAATTAAGTTTTGTTGA

Gene RcPR-5c

>29669.t000034 | 29669:212759..213502

ATGGCGTTATGGAGTTTTTTTACTCTGATTTGTATTATCCTATCAGGTGTTGAAGGTAGGGTTATCACCC
 TGGAAAACAAAATGCAAGGACACTATATGGCCTGGAATTTTAATAGAAGAAGGATCAACTTACCCTCAAGA
 ACTGGAGGGTGATCTCGTATTAAGACCGGGACAGGCATTCAATGTCACAGTACCACAACATTTGGACGATA
 TCTGTATGGGGAGGTAATGGGTGCTCATATAATAGCGTTGGTTATGGAACATGTGTCACCCGGTGGCTGTG
 ACAGGCAAAGATGCACCCGAGAACGCATCCCCTCATCACCAGTACATTAGTTGTACTTTACTTTGGATAG
 TCAACGAGATGTTTACGGAGTTGACGTTTCTGGTGGCTTCAACATACCGATTTAGTATCGCCATATGGC
 AAAGGTTCTGCAGAATGTAAGGCTGCAAGTTGCCCTACAGACGTGAACAGGATTTGCCCAGGAGTGTTC
 AAATGAAATCAAATTATCATGTCTGATGATGTAAGAAGCGATGACTGTCTGTACGTGACAGGCCGTTGAACG
 TTGTTGCTTAACAGCAGTACTGCTATTACTTCTAATTCTGATAGCTCGCATTAGGCACCTGCATGTGG
 ACCGATTATCTGATGGCGCTCAGAACCGCTTGTCTACTGCTACTCTTAATGGTTTTGACTGGACAAATA
 TGTTCGAGTGTGATGGACCTAATTATTTAATTAGCTTTTGTGTTGA

Gene RcPR-5d

>30116.t000002 | 30116:61671..63370

ATGGCTCAAAATCCCAGTTTGCTTCTGCTTGCTCTCTTCTTCTTACTATTACATCAGGTAACA
 TTATCTAATCTTTGCTTAACTGTTGTCATGGAGTGGAGCTCAACATTTTTCTTTTCTTTTATAA
 CCTTCTATTGCTTCGTTAACACCAACGATCATTGTTAAAAATAAATTTCTTATGTATATTTTTAATAT
 TTAATATTTATTTTTATTTTTATATTTAAATATATTTGTTAATATTTAATACACTATTTACAAAAA
 ATCTAAGTTATAGGCTCTAGTAACATTTAGGCACCCACTATTTTCTATTGACATATATATATATATAT
 ATATATATATATATATATATAAGTTTGATAGTTATGTAATTAATTTGATAGTCAAATAATGACTAAT
 TAAGTAATTTTTATGTGAGGCTTTTATTGCTTTAATTACTTTTTATCTGTTAATGAGTTATACCAAT
 AATCATGTTTTAACTACATTTACTACTTTAAAAATATAAAAATAGAATGAATCAGTGGACCGTTATGAAA
 TATAAATTTTAGCACCATACTCGAGTTATGAAATATAATTTTAGCATCATAATGTAATTAATTTGATA
 GTTAAATAATGACTAATTTAGGCACCCACTATTTTTTTTATGACATTTGATCATACTGGACCGTTTCGCA
 AGTCCAATAGATACTGGAGTTGGTATATTTCTTACTTGCCAAGACATCTCTGAGTTGTCAGCATCTTGA
 CAAGCTACTTTTGCCTTTGACAGGGCAAATGCTTCAAGTACAAAACTTTCACATTGCTTAATAACTG
 CAAAGAGACAATTTGGCCTGGAATAATAACAAAAGGTAATGACAGTGATGGAAAGGTTTTCCCCTGAAG
 CCTGGCCAAACCAATCTTCATTATGCACCAATCGGCTGGAGTGGTTCGTATATGGGCTAGAACTGGTTGTA
 ATTTTGATAAAAATGGCACTGGTACATGTCAAACAGGCAGTTGTGGCGCCTATCTGAATTGTTTCAGGACC
 AAGTAGCCACCTAACACAATGCTGAATTCACCCTGGAGATAATGACTTTTTTCGATATTAGCCTTGTG
 AACGGGTTAATTTGCCTGTAGTGATTTACCTCTAACAGTAAAGGAAATGTAGCATTGCTGGTTGTG
 ATGGAGATTTAAGGCAGAATTTGCCCGCCGAGCTCTCCGTCAGGTCGGATGGGAAGGTCATTTGCTTGGCC
 GAGCCGTGCGACGTGTTCAACACTGATGCGTACTGTTGTAGAGGAGCATAAGTATCCTGAGGCATGT
 GCACCTCAAACATTTCAAGGATTTTTAAGCAAGTCTGCCCTGCATCGTCTAGCTATGCATTTGACAGGG
 ATGCTACAGGCATCATAACGTGCTCTGCATCAGATTATGTTGTTGCCTTTTGTGCATCAAGTTAAGCTTT
 TATCAGTTGTCCTTTAGTAGCTTAATAGTAATAATGGAAGTATATATTTATTTATTTACTGATTGTTTTAT
 TTTGTGGACGGATTTTAGGAATCAGACTTTCTGTTCTTACCATGACAATGAGGTTTCAGTGCATAACATCA
 GCAAATGACTCAAAGCACTTCTGCCCATGGCTGGTGGATCTGATGGTGACATTGCCCTTGGCATTTA
 TTTTGCAAACCAAGATTTAG

Gene RcPR-5e

>29669.t000043 | 29669:282612..283444

ATGGAACTTTGCTCGCATTTTATCCTTCTAAGTTTGTATTCTTTGTTAGCTACAGGCCTGATGCCAGGG
 TAATTACCTGGTGAACAACTGTAACACTACCATATGGCCTGGTTTTCAAGTTGTAATGGCAAGTCAACG
 GTCTGAATTTATCGAGGATGGTTTCGAGTTAAGACCTGGACAATCAGCAAATCTTACAGCACCAGCAGTT
 TGGTCACTCTACATCTGGGGTCTAGTGGGTGCTCATTTCAACAGCCTCGGCTATGGTGGATGTGTCAGT
 GTGACTGTGGCGGCACACTGAGATGTCCCTAGGACGCTTCCCTAACCTACCAATTACGATGGTCAACT
 TTCCCTGAATAATAATTTGTTGACACTTATGGAGTTTATCTTGGTGGGGGTTAATATTCCTGTCTCT
 ATATCGCCTTATGGACCTGGTTCTGATCAATGCCAGGAGTCAAGTTGCCCTTTCGGACGTGAACAGCATT
 GCCCGCTGAGTTGCAAGTAAGAGTAAATAATATGATTTGTAGCATGCAAAAAGTGCCTGCCTTGCATTCAA
 CAGGCCGTAATTTCTGTTGCTTAAATACCAATGGTACGAACTGCTTGCAGCCAACGAATATACGACC
 ATGTTTCAGGGAGTATGTCTGCAGCTAGCGCTGATGGCTCTAGTCAGAGTATATTTAAGTGCAGCGGAT
 CTCAGTTCTTCTGCTAGCTTTTGTGAGACTGCTTATTGATCAACATAATAAAAATATATATGTACCAACAC
 ATATTTGTATTTCTGCAATGGCCATTTATGATTGGAGAATCTTAGTACAAACTCTATGGCGGG

Gene RcPR-6

>30099.t000037 | 30099:186152..186646

ATGGGTTCTGTTAAAGATTTACAGGTTCTGTTTCTGCTCATGGCTTAGCCTTGGTGGTGCACCCCT
 GCCATGCCAAAACACTCACAACAAGACTACCTTGATGCCACAATGCTGCTCGTGCACAAGTAGGTGTAGC
 CAATATAGTATGGGACAATACTGTAGCAACCTATGCTCAAACATATGCCAACTCAAGAATGGTACTGC
 AATCTTGTTCATTCTAGCGGATCCTACGGCGAAAACCTTAGCCAAAGGCAGCAGGTTCACTTACAGGAA
 CTGCAGCAGTGAACCTGTGGGTGGCGGAGAAGCCTTACTACAACACTACACTTCCAATTCCTGTACTGGGG

TCAACAATGCTTGCACTATACTCAAGTGGTTTGGAGCAAATCGGTTTCGTCTAGGGTGCGCCAGGGTCCAG
TGTACTAATGGCTGGTGGTTCGTTACTTGCAACTATGATCCCCCTGGCAATTATATTGGTCAGAAACCTT
ACTAA

Gene RcPR10-STH21

>29726.t000222 | 29726:1239960..1240556

ATGGGTGTATCTACCTTCACACAAGAATTCACCTCCCCAGTTGCTCCATCACGTATGTTCAAGGCCTTAA
TCCTAGATTCTAACATTTTATGCCAAAGCTCCTACCCCAATTTATTAAGAGTGGTGGTATGTTATCCAAG
TGGCGATGGCGCTGGGACCATTGAACAAGTCAACTTCACTGAAGGTATCTCACTTTACATGTTCTTTTTT
ATCTCTCTTTTTCTTTTTAACTCTTATTTTCTCATTTCTTGGGATTTGGTTATTGATATATAATGATTT
GATGCATATGTTATGCAGCTAGTGACTTCAAGTATATCAAGAATAGAATTGAAGAGCTAGACAAAGACAA
TCTTGTGTGCAAGTATACTCTGATTGAGGAGGTCCACTCGGGGACAAGCTTGATTCTATGCTTATGAG
GTCAAATTTGAGGCAGCCAGTGATGGAGGCTGCATCTGTAAGATCACAAGCAACTATTTTACTCTTGGAA
AATTTACTATTAAGAAGAAGATATTGAGGATGGCAAGGGCAAAGCTCTAGGAATATACAGAGTCGTTGA
AGCTTACCTCTTGGAGAACCCTCAAGTTTATGCTTAA

Dados Suplementares 3: Sequências de aminoácidos das Proteínas Relacionadas a Patogênese em *Ricinus communis* L. Peptídeo sinal: sublinhado; Domínio Funcional Conservado: hachurada em cinza, Sítios de Fosforilação (Serina, Treonina e Tirosina) e em azul e os Sítios de N-glicosilação (Asparagina-x-Serina/Treonina), em que x pode ser qualquer aminoácido, exceto a prolina) em vermelho.

RcPR-1a

>30099.m001649

MYLNDFFAATVMVGTVLVHISLAQNSPQDYVNAHNTVRAEIGVGPITWNKTVAAYAQTYANSRIESCEFE
HSYGPYGENIAEGYGNLNGVDVAVNMWVSEKPNYDYGNSCKGGEDECLHYTQVVWRNSVHLGCGRAKCKT
GWWFVTCNYHPVGNIEGQRPF

RcPR-1b.1

>30099.m001650

MSNGLFFAEGRFQCNFCQSLLFSSLWLSLWLFNRNDITFNTCTPNWGFSSWINAVEPAFPYSGSQLVFSLMG
AKDWE~~SRRK~~SPVNVVWKA~~PAASSVKWNVYA~~SPMDWRRPPYNFRNFETIKE~~SD~~STIQLVG~~SVWKWRMALEE~~
ISELAATQLPFISF~~SP~~ISRTGNQVVDGLAK~~SA~~VERT~~S~~DFVAWL

RcPR-1b.2

>30099.m001651

MGFNKFIVALCLMTLTI~~TPV~~SLAQNSQKDYVAAHNEVRAELGLGPVRWNEKLALYARKYIQTKVETCILE
HSNGPYGENLAKGSGEGFS~~GVD~~AVKLW~~ADEKPNYDYL~~SN~~SCAGG~~MC~~GHYTQ~~IWRD~~TKE~~IGCAK~~TKCKDG~~
W~~TY~~IS~~CNY~~D~~PPGN~~YIGERPF

RcPR-1c

>30099.m001652

MLSLKISAA~~LS~~FLISLSLILS~~SHAQ~~NLPQDYINAHNAARA~~AVGVG~~PM~~T~~WDNTVAAYAQN~~YANQR~~INDCRL
VH~~SSGR~~YGENIAW~~SSGDL~~SGTAAV~~KLWVDEKAF~~YN~~YNSNT~~CAAGQ~~CGHY~~TQVVWRNSVRLGCAK~~VRCRN~~
GGTFITCNYD~~PPGN~~FVNQRPY

RcPR-1d

>30099.m001653

MAMPKV~~SLPLFY~~F~~MIMS~~LALLFSSNAQDTNQDYINSHNDARVAVGAGLGN~~MT~~WNETVADYARD~~YANQRIA~~
DCNLVH~~SDGP~~YGENLAWGSGDL~~SGLDAVR~~MWVDEKAF~~YD~~YNSNTCTGGQ~~CGHY~~TQVVWRD~~S~~ISL~~GCAV~~
TCN~~NGLG~~T~~LITCNY~~Y~~PPGN~~VIGQRPY

RcPR-1Be

>30099.m001657

MKQFDPSLALVCIISIALIFSSHAQDTQQDYLN~~AHNS~~ARGDVGLG~~PLT~~WDDKVA~~SYAQ~~H~~HANQ~~H~~KSDCSL~~

VHSEGPYGENLAWSSGDLSTDAVKMWEKPYDYNSNSCASCQICGHYTQVVWRSVRLGCAKVSCNNGG
TFIGCNYDPPGNYIGORPYW

RcPRB1-3f

>27894.m000804

MLSCGALGTFLFLLLASASEVSSGRSSLIRTQRETIKQYLKPHNRERAKLGLRPLKWSNKLASFASSWA
HQRQGDCELLHSNSNYGENLFWGSGKDWKPGDAVAWAEKCYNHNTNTCTKNKDCLHYTQMVWRQSLK
VGCARVICTSGDTLITCNYDPHGNVIGERPF

RcPR-1g

>27766.m000157

MKPHSFLLLFIFLITIFITNNLLVTSQSPSPAETPLNKRNDTIYKVSQQLCWGCIGEALQFLYAHNLV
RASKWELPLTWDSQLERYARWWAGTRKQDCQLEHSFPEGDFKLGENIYWGSGTAWTPRDAVSAWASEEKY
YTYATNSCEEQMCQGHYTQIVWKTRRIGCARVVCDDGDVFMTCNYDPPGNYIGEKPY

RcPR-1h

>27766.m000158

MRTYCYCVVAIIISISLFLTTRATQSSAAEAIAIHQHAYLSTANQFLGPHNAARAALRMPPLIWDTRLA
NYAQWYANQRRFDCDLRHSNGPYGENIFWGSWTGTPAQAVTAWVSEKWKYNYWSNSCYGHQECGHYTQI
VWRKTRRIGCAKVTCSDDLGVFMTTCNYDPPGNYIGERPY

RcPR-1i

>30099.m001648

MGFTNLVIALCLVASALSPTCLARNRSGDYLDAHNAVRAEVGDPLVWNKTLDYAKASQPQNFNAKDAV
KIWADEKKFYDRKSNCKGGEQGHYTQIVWHDTSQVGCARVKCKNGHTFISCNYPYPIGNVQGGQSPY

RcPR-4a.1

>29598.m000457

MGVRVTVCIVFLAFLIARATCQSGSNVRATYHFYNPAKINWDLRAASAYCATWDADKPLAWRRKYGWTAF
CHPVGQQAACGRCLRVNTGTTRAQVTVRIVDQCSNGGLDLDEGVFRQTDGKGYARGHLIVNYQFVNCGD

RcPR-4a.2

>29598.m000458

MGTRVTVCIVFLAFLIARATCQSGSNVRATYHFYNPAQINWDLRAASAFCATWDADKPLAWRQKYGWTAF
CHPVGQQAACGKCLRVNTGTGAQVTVRVVDQCSNGGLDLDEGVFRQTDGRGIAQGHILVSYQFVNCGD

RcPR-4b

>29598.m000459

MGSVVLMVLLCSLGVAFGEQCGRQAGGKLCNNLCCSQYWCSTDDYCSPSKNCQSNCKGGGGGESAS
NVRATYHFYNPQQHGWDLNAVSAYCSWDANKPYSWRSKYGTAFCCGAPGPRQTSCGKCLRVTNTRTGA
ATTVRIVDQCSNGGLDLVNVFRQLDTDGIGYQQGHILVNYQFVNCGDSFNPLFSIMDDQ

RcPR-5a

>30170.m013991

MHYKLIKISILAIIVCGTNLYFCVKFAGNFSNATVFTLQNHCSYMIWPGTSLGNGAGTLGDGGFALPPGSS
IQLQAPPTWSGRFWARTGCTFDDLNGKCDTGDCGALKCTGGGAPPVSLAEFTIAPSSNEKDFYDVS LV
DGYNVGLGVKALGSGDCQYAGCVADLNGNCPPELVRTDSVSGSVVACKSACAAAFNAPEFCCTGDHSTPQ
TCSPTQYSALFKNACPSAYSAYYDDASSTCTCSGSDYMITFCPTGSG

RcPR-5b

>30152.m002402

MANWQFLTLVLISFLGSEAAIFTLQNNCRSTIWPGLTSGKGPLLINGGTRLKPGRTMRINVPKWSGRF
WARTGCKFDQTGKGCATGDCGGLKHCAGSVPVTLAEFTLDDLNFYDVS LVDGYNIPISYIPYGGSG
DCKKVECFSNLNRCPKDLQVINNGKIVACKSACTAYNTPKYCCSGAYGSPQTCCKPTGFSRIFKAACPTY
YSAYYDDATSTFTCKDANYRIKFC

RcPR-5c

>29669.m000825

MALWSFFTLCIILSGVEGRVITLENKCKDTIWPGLIEEGSTYPQELEGLVLRPGQAFNVTVPQHWTT
SVWGGNGCSYNSVGYGTCVTGGCDRQCTGERIPSSPVTLVVLYLDSQRDVYGVDSGGFNIPISVSPYG
KGSAECKAASCPTDVNRICPGVLQMKSNYHVACRSDDCLLRDRPERCCSNSSTAITSNSSSHSGTCMW
TDYLMALRTACPTATLNGFDWTNMFECGPNYLI SFC

RcPR-5d

>30116.m000371

MAQNPSLLLLLALFFIFFTITSGQNASSTKFTLLNCKETIWPGIITKGNDSGKGFPLKPGQTNLHYAP
 IGWSGRIWARTGCNFDKNGTGTCTGSCGAYLNCSPSSPPNTIAEFTLGDNDFFDI~~SLVNGFNLPVVIS~~
 PLNSKGNCS~~IAGCDGDLRQNCPEL~~SVRS~~DKVIACRS~~ACDVFN~~TDAY~~CCRGAYS~~DPEACAPS~~NY~~SRI~~FK
 QVCPASS~~SYAFDRDATS~~IITCS~~ASDYVVAFCASRNQ~~TFC~~SYHDNEVQCNTS~~AND~~SKALLPHGWWILMVTL~~
 PLAFILQTKI

RcPR-5e

>29669.m000834

MELCSHFILL~~TLISLLATGTDARVITLVNNC~~NTTIWPGFQVVMASQRSEFIEDGFELRPGQ~~SANLTAPAV~~
 WSLYIWGRSGCSFN~~SLGYGCVTGDCGG~~TLRCPLGRFPNLPITMVTLSLNNNLFDTYGVYLGAGFNIPVS
 ISPYGPG~~SDQCQESSCL~~SDVNSICPAELQVRVNNMIVACKSACLA~~FNRPEFCCLNTNGTNT~~ACKP~~TNYTT~~
 MFEGVCPAASADGSSQSIFKCSGSQFFVSFC

RcPR-6

>30099.m001658

MGSLKISQVLV~~SLLMGLALVVHPCHAQNS~~QQDY~~LDAHNAARAQVGVANIVWDNTVATYAQNYANSRIGDC~~
 NLVHSSGSYGENLAKGSSSSLTGTA~~AVNLWVAEKPYNY~~TSNSCTGGQQCLHY~~TQVVWSKSVRLG~~CARVQ
 CTNGWWFVTCNYDPPGNYIGQKPY

RcPR10-STH21

>29726.m004110

MGVSTFTQEFTSPVAPSRMFKALILDSNILIPKLLPQFIKSVDVIQGGDGAGTIEQVNFTEASDFKYIKN
 RIEELDKDNLVCKYTLIEGGPLGDKLDSIAYEVKFEAASDGGCICKITSNYFTLGKFTIKEEDIEDGK GK
 ALGIYRVVEAYLLENPQVYA

Tabela Suplementar 5. Elementos *cis* presentes na região promotora dos genes das Proteínas Relacionados a Patogêneses em *Ricinus communis* L.

GENE	NOME	FUNÇÃO	POSIÇÃO	NÚMERO DE ELEMENTOS	ELEMENTOS TOTAIS
RcPR-1a	AE-box	Light responsiveness	63+	1	
RcPR-1a	Box 4	Light responsiveness	1152-, 698+	2	
RcPR-1a	G-Box	Light responsiveness	1341-	1	9
RcPR-1a	G-box	Light responsiveness	776+, 1070-	2	
RcPR-1a	MRE	Light responsiveness	630+, 723+	2	
RcPR-1a	TCT-motivo	Light responsiveness	99-	1	
RcPR-1a	MYB	Cell cycle control in eukaryotes	396-, 681-	2	4
RcPR-1a	Myb-binding site	Cell cycle and cell proliferation response	396-, 681-	2	
RcPR-1a	MYC	Drought response	309+, 525-	2	3
RcPR-1a	as-1	Drought response	1125+	1	
RcPR-1a	Myb	Water response	351+, 1323-, 727+	3	3
RcPR-1a	ABRE	ABA response	1070-, 1342+, 777+	3	
RcPR-1a	ABRE3a	ABA response	1070-	1	5
RcPR-1a	ABRE4	ABA response	1070+	1	

RcPR-1a	GARE-motivo	gibberellin-responsive element	395+, 680+	2	2
RcPR-1a	CAT-box	Meristem specific activation	781-	1	4
RcPR-1a	ERE	Ethylene-responsive element	53+, 216-, 55-, 245+	4	4
RcPR-1a	TCA	Nutrient starvation stress response	867+, 1402+	2	2
RcPR-1a	TATA	Core promoter / enhancer element	1057+	1	1
RcPR-1a	TATA-box	Core promoter element around -30 of transcription start	27-, 166+, 214+, 232+, 234-, 240+, 491-, 492+, 495+, 496-, 497+, 639+, 836-, 875+, 876-, 877-, 892-, 893-, 993-, 994-, 1045+, 1046-, 1047-, 1056+, 1057-, 1093+, 1094-, 1095-, 1157-, 1164+, 1165-, 1166-, 1199-, 1200-, 1201-, 1250+, 1413-	37	37
RcPR-1a	CAAT-box	Common cis-acting element in promoter and enhancer regions	10+, 26+, 150-, 255+, 310-, 414-, 449-, 467-, 482-, 539-, 562-, 675+, 718+, 740-, 786+, 794+, 795+, 855+, 897+, 908-, 1013+, 1023+, 1024+, 1080+, 1131-, 1144-, 1181-, 1234+, 1255-, 1276-, 1298+, 1347+	32	32
RcPR-1a	CGTCA-motivo	MeJA-responsiveness	1125-	1	2
RcPR-1a	TGACG-motivo	MeJA-responsiveness	1125+	1	
RcPR-1a	Unnamed_2		630+	1	1
RcPR-1a	Unnamed_4		314-, 1464+, 1436+	3	3
RcPR-1b.1	AAGAA-motivo	Light responsiveness	1442-	1	14
RcPR-1b.1	Box4	Light responsiveness	50+, 1461-, 1002-	3	

RcPR-1b.1	G-Box	Light responsiveness	404+	1	
RcPR-1b.1	G-box	Light responsiveness	403+, 1143+, 424-, 1145-, 404+, 426+	6	
RcPR-1b.1	GATA-motivo	Light responsiveness	575-	1	
RcPR-1b.1	GT1-motivo	Light responsiveness	1288+	1	
RcPR-1b.1	chs-Unit 1 m1	Light responsiveness	989+	1	
RcPR-1b.1	MYB	Cell cycle control in eukaryotes	766+	1	2
RcPR-1b.1	Myb-binding site	Cell cycle and cell proliferation response	766+	1	
RcPR-1b.1	MYC	Drought response	81-, 1029-, 727-, 414-, 877-	5	
RcPR-1b.1	as-1	Drought response	959-	1	7
RcPR-1b.1	MBS	Drought-inducibility	1433+	1	
RcPR-1b.1	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1318+, 656+	2	2
RcPR-1b.1	MBSI	Flavonoid biosynthetic genes regulation	1281+	1	1
RcPR-1b.1	Myb	Water response	1063+, 1433+	2	2
RcPR-1b.1	ABRE	ABA response	1145-, 402+, 405+, 404+, 427+	5	
RcPR-1b.1	ABRE3a	ABA response	426+, 1145-	2	9
RcPR-1b.1	ABRE4	ABA response	426-, 1145+	2	

RcPR-1b.1	GARE-motivo	Gibberellin-responsive element	766-	1	1
RcPR-1b.1	ARE	Anaerobic induction	185-, 935-, 632+, 1437-	4	4
RcPR-1b.1	O2-site	Zein metabolism regulation	388+	1	1
RcPR-1b.1	TC-rich repeats	Defense and stress responsiveness	1439+	1	1
RcPR-1b.1	WRE3	Wounding and pathogenic response	996+	1	2
RcPR-1b.1	box S	Wounding and pathogenic response	732-	1	
RcPR-1b.1	TGACG-motivo	MeJA-responsiveness	959-	1	2
RcPR-1b.1	CGTCA-motivo	MeJA-responsiveness	959+	1	
RcPR-1b.1	AC-I	Tissue specific / preferential expressed	356+, 990+	2	2
RcPR-1b.1	TATA-box	Core promoter element around -30 of transcription start	17+, 1339-, 1055-, 1479-, 19+, 1477-, 1337+, 1481+, 18-, 1474-, 1056-, 1480-, 968-, 1478-, 1338-, 1482-	16	16
RcPR-1b.1	CAAT-box	Common cis-acting element in promoter and enhancer regions	322+, 324-, 81+, 483+, 484+, 506+, 521-, 550+, 573+, 627-, 672+, 675-, 727+, 786-, 876+, 877+, 879-, 902+, 1029+, 1031-, 1037-, 1059-, 1099-, 1132+, 1246+, 1249-, 1455+, 1473+	28	28
RcPR-1b.1	AT~TATA-box	Core promoter / enhancer element	1480-, 1478-	2	2
RcPR-1b.1	Unnamed_1	60K protein binding site	406+, 1144-, 428+	3	3
RcPR-1b.1	Unnamed_10		424-, 1143+	2	2
RcPR-1b.1	Unnamed_12		424-, 1143+	2	2

RcPR-1b.1	Unnamed_14		424-, 1143+	2	2
RcPR-1b.1	Unnamed_4		411+, 858-, 642-, 1104-, 458+, 745+	6	6
RcPR-1b.1	Unnamed_8		424-, 1143+	2	2
RcPR-1b.2	AAGAA-motivo	Light response	919+, 1036+	2	
RcPR-1b.2	AE-box	Light response	159-	1	5
RcPR-1b.2	F-box	Light response	1162+	1	
RcPR-1b.2	TCCC-motivo	Light response	1468+	1	
RcPR-1b.2	MYB	Cell cycle control in eukaryotes	7+, 1258+, 118+, 15-, 555+	5	5
RcPR-1b.2	MYB-like sequence		555+	1	1
RcPR-1b.2	MYB recognition site	Drought response	1176-	1	
RcPR-1b.2	MYC	Drought response	1234-	1	4
RcPR-1b.2	MBS	Drought-inducibility	122+, 697+	2	
RcPR-1b.2	Myb	Water response	122+, 697+	2	2
RcPR-1b.2	TGA-element	Auxin-responsive element	543-	2	2
RcPR-1b.2	TCA-element	Salicylic acid responsiveness	982+	1	1
RcPR-1b.2	ERE	Ethylene-responsive element	854+	1	1
RcPR-1b.2	STRE	Nutrient starvation stress response	442+, 1401-	2	2
RcPR-1b.2	WRE3	Wounding and pathogenic response	982+	1	1
RcPR-1b.2	GT1-motivo	Low-temperature responsiveness	333-	1	2

RcPR-1b.2	LTR	Low-temperature responsiveness	547-	1	
RcPR-1b.2	TATA-box	Core promoter element around -30 of transcription start	194+, 809-, 192-, 810-, 823-, 824-, 825-, 826-, 850-, 851-, 862-, 863-, 864-, 865-, 873-, 874-, 875-, 876-, 906-, 907-, 909-, 934-, 989-, 1029+, 1030-, 1031-, 1205-, 1403+, 1405-, 1411-	30	30
RcPR-1b.2	CCAAT-box		1176+	1	1
RcPR-1b.2	AT~TATA-box	Core promoter / enhancer element	907-	1	1
RcPR-1b.2	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)	1309-	1	1
RcPR-1b.2	CAAT-box	Common cis-acting element in promoter and enhancer regions	150-, 170-, 129-, 207-, 302+, 303+, 464-, 513+, 514+, 590-, 636-, 660-, 675-, 723-, 727+, 751-, 795-, 817+, 829-, 892+, 893+, 953+, 955-, 973+, 999-, 1017+, 1046-, 1076+, 1146-, 1150-, 1198+, 1223-, 1234+, 1295+, 1296+, 1320+, 1321+, 1449+	38	38
RcPR-1b.2	CAT-box	MYBHv1 binding site	62-	1	1
RcPR-1b.2	TATA	Core promoter / enhancer element	861-, 872-, 822-	3	3
RcPR-1b.2	Unnamed_2		1482+	1	1
RcPR-1b.2	Unnamed_4		44-, 565+, 314-, 1469+, 143+, 576-, 416-, 1486-, 99+, 573-, 375-, 1057+, 291-, 437-	14	14
RcPR-1c	AAGAA-motivo	Light response	666-	1	
RcPR-1c	AE-box	Light response	4-	1	11
RcPR-1c	Box 4	Light response	95+, 982-, 882-, 1478-	4	

RcPR-1c	Box II	Light response	1189+	1	
RcPR-1c	G-box	Light response	1190-	1	
RcPR-1c	GATA-motivo	Light response	840-	1	
RcPR-1c	GT1-motivo	Light response	599-	1	
RcPR-1c	I-box	Light response	838-	1	
RcPR-1c	MYC	Drought response	39+, 795-	2	
RcPR-1c	MBS	Drought-inducibility	173+	1	3
RcPR-1c	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1217+	1	1
RcPR-1c	Myb	Water response	173+, 294-	2	2
RcPR-1c	ABRE	ABA response	1190-	1	
RcPR-1c	ABRE3a	ABA response	1190-	1	3
RcPR-1c	ABRE4	ABA response	1190+	1	
RcPR-1c	ERE	Ethylene-responsive element	264+	1	1
RcPR-1c	ARE	Anaerobic induction	158-, 507-	2	2
RcPR-1c	W box	Wounding and pathogenic response	1281+	1	1
RcPR-1c	WRE3	Wounding and pathogenic response	1290-	1	1

RcPR-1c	CTAG-motivo	Determinant for repressor binding	747+	1	1
RcPR-1c	circadian	Circadian control	274+	1	1
RcPR-1c	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)	371+	1	1
RcPR-1c	TCA	Nutrient starvation stress response	1205-	1	1
RcPR-1c	AT~TATA-box	Core promoter / enhancer element	368+, 546+, 542+, 544+	4	4
RcPR-1c	TATA-box	Core promoter element around -30 of transcription start	45-, 545+, 544+, 546+, 46+, 547+, 543+, 548+, 53+, 551+, 66-, 552-, 68+, 553+, 367+, 558+, 368+, 568-, 370+, 611-, 424-, 612-, 430+, 613-, 440+, 614+, 441-, 1017-, 442+, 1019-, 487-, 1045-, 488-, 1239-, 489-, 1240-, 490+, 540+, 541-, 542+	40	40
RcPR-1c	CAAT-box	Common cis-acting element in promoter and enhancer regions	40-, 185+, 189+, 209-, 244+, 289+, 338-, 365+, 379+, 411-, 474-, 499-, 505-, 658-, 770+, 771+, 795+, 916+, 1014+, 1053+, 1054+, 1082-, 1104-, 1150-, 1159-, 1234-, 1277+, 1280-, 1335-, 1339+, 1346-, 1418+, 1446-, 1476+	34	34
RcPR-1c	Unnamed_1	60K protein binding site	998-	1	1
RcPR-1c	Unnamed_10		1188+	1	1
RcPR-1c	Unnamed_12		1188+	1	1
RcPR-1c	Unnamed_14		1188+	1	1
RcPR-1c	Unnamed_4		36+, 1294-, 710+, 283-, 964+	5	5
RcPR-1c	Unnamed_8		1188+	1	1

RcPR-1d	Box 4	Light response	237+	1	
RcPR-1d	G-Box	Light response	836-	1	
RcPR-1d	G-box	Light response	833-, 835-, 834-, 836-	4	
RcPR-1d	GT1-motivo	Light response	383+	1	11
RcPR-1d	MRE	Light response	1073-, 1175-	2	
RcPR-1d	ATC-motivo	Light response	887-	1	
RcPR-1d	TCCC-motivo	Light response	1374+	1	
RcPR-1d	MYB	Cell cycle control in eukaryotes	82+	1	2
RcPR-1d	Myb-binding site	Cell cycle and cell proliferation response	82+	1	
RcPR-1d	MYB recognition site	Drought response	327-	1	3
RcPR-1d	MYC	Drought response	1009-, 1458-	2	
RcPR-1d	Myb	Water response	85-, 424+	2	2
RcPR-1d	ABRE	ABA response	834-, 837+, 836-	3	3
RcPR-1d	P-box	Gibberellin-responsive element	378-	1	1
RcPR-1d	ARE	Anaerobic induction	6-, 428-	2	2
RcPR-1d	W box	Wounding and pathogenic response	1194-	1	1
RcPR-1d	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)	52+	1	1
RcPR-1d	AT~TATA-box	Core promoter / enhancer element	390-, 1430-, 392+	3	6

RcPR-1d	TATA	Core promoter / enhancer element	311+, 1243+, 573+	3	
RcPR-1d	WUN-motivo	Wound-responsive element	805+	1	1
RcPR-1d	CAAT-box	Common cis-acting element in promoter and enhancer regions	60+, 103-, 115-, 146-, 211-, 248-, 360+, 397-, 434+, 534+, 585-, 606+, 612+, 638-, 727-, 748+, 754+, 766-, 821-, 825-, 852-, 948-, 965-, 1009+, 1011-, 1054-, 1135+, 1136+, 1197+, 1419-, 1452+, 1458+	32	32
RcPR-1d	CCAAT-box	MYBHv1 binding site	327+	1	1
RcPR-1d	Unnamed_1	60K protein binding site	835-	1	1
RcPR-1d	Unnamed_2		1172-	1	1
RcPR-1d	Unnamed_4		33-, 1375+, 511-, 452-, 983-	5	5
RcPR-1Be	AAGAA-motivo	Light responsiveness	671+, 1197-, 788+, 1201-	4	
RcPR-1Be	ACE	Light responsiveness	1259-	1	
RcPR-1Be	Box 4	Light responsiveness	52+, 216+, 97+	3	
RcPR-1Be	G-box	Light responsiveness	603+, 1001-	2	14
RcPR-1Be	I-box	Light responsiveness	7-, 1356+	2	
RcPR-1Be	MRE	Light responsiveness	779-	1	
RcPR-1Be	TCT-motivo	Light responsiveness	303+	1	
RcPR-1Be	MYB	Cell cycle control in eukaryotes	868+	1	1
RcPR-1Be	MYC	Drought response	22+, 1375+, 279+	3	
RcPR-1Be	MBS	Drought-inducibility	486+	1	4

RcPR-1Be	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	980+	1	1
RcPR-1Be	Myb	Water response	486+, 1097+, 1083-	3	3
RcPR-1Be	ABRE	ABA response	600-, 604+, 603+	3	
RcPR-1Be	ABRE3a	ABA response	603+	1	6
RcPR-1Be	ABRE4	ABA response	603-	1	
RcPR-1Be	AT~ABRE	ABA response	603+	1	
RcPR-1Be	CAT-box	Meristem expression	1002+	1	1
RcPR-1Be	ERE	Ethylene-responsive element	469+, 547-	2	2
RcPR-1Be	ARE	Anaerobic induction	1491+	1	1
RcPR-1Be	O2-site	Zein metabolism regulation	640-	1	1
RcPR-1Be	W box	Wounding and pathogenic response	1137+, 1363-	2	2
RcPR-1Be	circadian	Circadian control	637+, 1019-	2	2
RcPR-1Be	AT-rich sequence	element for maximal elicitor-mediated activation (2copies)	580+	1	1
RcPR-1Be	AT~TATA-box	Core promoter / enhancer element	399+, 415+, 407+, 541+, 405+, 419+, 411+, 1143-, 401+, 417+, 409+, 440+, 403+, 413	14	14

RcPR-1Be	CAAT-box	Common cis-acting element in promoter and enhancer regions	7-, 23-, 42+, 45-, 101-, 141-, 198+, 199+, 206-, 213+, 240-, 280-, 327-, 460+, 553-, 724-, 805+, 854-, 934+, 962+, 998-, 1110-, 1135-, 1159-, 1181+, 1222+, 1249-, 1295-, 1303-, 1330-, 1366+	31	31
RcPR-1Be	MYB-like sequence		868+	1	1
RcPR-1Be	TATA	Core promoter / enhancer element	578+	1	1
RcPR-1Be	TATA-box	Core promoter element around -30 of transcription start	66+, 178+, 180+, 398+, 399+, 400+, 401+, 402+, 403+, 404+, 405+, 406+, 407+, 408+, 409+, 410+, 411+, 412+, 413+, 414+, 415+, 416+, 417+, 418+, 419+, 420+, 421+, 438-, 440+, 442+, 541+, 542+, 543+, 545+, 578+, 622+, 624+, 650+, 851-, 885-, 886-, 887-, 936-, 937+, 939-, 940-, 941-, 942-, 943-, 984-, 985-, 994-, 1037-, 1047+, 1048-, 1143-, 1144-, 1145-, 1184-, 1185-, 1213+, 1450+, 1452-	63	63
RcPR-1Be	Unnamed_4		373-, 753+, 662+, 840+	4	4
RcPR-1Be	Unnamed_6		936-	1	1
RcPRB1-3f	AAGAA-motivo	Light response	1011+	1	
RcPRB1-3f	AE-box	Light response	958-,	1	
RcPRB1-3f	AT1-motivo	Light response	120+, 450+, 214-	3	
RcPRB1-3f	ATCT-motivo	Light response	329-	1	13
RcPRB1-3f	Box 4	Light response	837-, 1173-, 1090-, 1457-	4	
RcPRB1-3f	Box II	Light response	1138+	1	
RcPRB1-3f	GATA-motivo	Light response	328+	1	

RcPRB1-3f	TCT-motivo	Light response	366-	1	
RcPRB1-3f	MYB	Cell cycle control in eukaryotes	547-, 1153+	2	3
RcPRB1-3f	dOCT	Cell cycle and cell proliferation response	1239-,	1	
RcPRB1-3f	MYC	Drought response	726-, 1379+, 1253+	3	4
RcPRB1-3f	as-1	Drought response	276-,	1	
RcPRB1-3f	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	204-	1	1
RcPRB1-3f	TCA-element	Salicylic acid responsiveness	178-	1	1
RcPRB1-3f	CAT-box	Meristem expression	1245-	1	1
RcPRB1-3f	ERE	Ethylene-responsive element	70-, 947+, 461+, 160+, 872+	5	5
RcPRB1-3f	O2-site	Zein metabolism regulation	1039-	1	1
RcPRB1-3f	TC-rich repeats	Defense and stress responsiveness	366-, 1329+	2	2
RcPRB1-3f	WRE3	Wounding and pathogenic response	1346+	1	1
RcPRB1-3f	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)	694-	1	1

RcPRB1-3f	CGTCA-motivo	MeJA-responsiveness	276+	1	2
RcPRB1-3f	TGACG-motivo	MeJA-responsiveness	276-	1	
RcPRB1-3f	WUN-motivo	Wound-responsive element	1013-, 1095+,	2	2
RcPRB1-3f	AT~TATA-box	Core promoter / enhancer element	380+, 981-, 406+, 1442-	4	4
RcPRB1-3f	CAAT-box	Common cis-acting element in promoter and enhancer regions	62+, 63+, 114-, 226-, 246-, 258+, 336-, 413+, 428-, 543-, 603+, 726+, 810-, 823+, 846-, 918-, 940-, 969+, 1108-, 1112-, 1156+, 1157+, 1171+, 1254-, 1352+, 1428-, 1465+, 1473+	28	28
RcPRB1-3f	MYB-like sequence		547-	1	1
RcPRB1-3f	TATA	Core promoter / enhancer element	888+	1	1
RcPRB1-3f	TATA-box	Core promoter element around -30 of transcription start	379+, 765-, 764-, 766-, 380+, 767-, 722+, 782+, 382+, 783-, 404-, 784-, 406+, 825+, 407+, 826-, 408+, 827-, 432+, 869-, 444+, 882-, 451+, 887+, 452-, 888-, 453+, 937-, 463-, 981-, 468+, 982-, 469+, 983-, 526+, 1156+, 527-, 1442-, 528+, 1444-, 633+, 634-, 635+, 702+, 718-, 721-	46	46
RcPRB1-3f	Unnamed_4		156-, 1357+, 186-	3	3
RcPR-1g	AAGAA-motivo	Light response	714-	1	
RcPR-1g	AE-box	Light response	1048+	1	
RcPR-1g	Box 4	Light response	374+, 1174-, 1124-	3	11
RcPR-1g	G-Box	Light response	513+	1	
RcPR-1g	G-box	Light response	508-, 510-	2	
RcPR-1g	GATA-motivo	Light response	159-	1	

RcPR-1g	TCT-motivo	Light response	146+, 1084-	2	
RcPR-1g	MYC	Drought response	474-	1	1
RcPR-1g	ABRE	ABA response	513-	1	
RcPR-1g	P-box	Gibberellin-responsive element	644-	1	
RcPR-1g	TCA-element	Salicylic acid responsiveness	458-	1	
RcPR-1g	ERE	Ethylene-responsive element	740+, 1367-, 1162+	3	
RcPR-1g	AuxRR-core	Auxin responsiveness	882-	1	
RcPR-1g	ARE	Anaerobic induction	1027+, 1360+	2	2
RcPR-1g	STRE	Nutrient starvation stress response	1435-	1	
RcPR-1g	HD-Zip 1	Palisade mesophyll cells	748-	1	
RcPR-1g	WUN-motivo	Wound-responsive element	1011-	1	
RcPR-1g	CAAT-box	Common cis-acting element in promoter and enhancer regions	176+, 177+, 38-, 199+, 211-, 298+, 299+, 310+, 431-, 450-, 545+, 564+, 616+, 664+, 676+, 677+, 748+, 753-, 832-, 850+, 864-, 911+, 917+, 982-, 1042-, 1092-, 1156+, 1184+, 1270-, 1282+, 1285-, 1301+, 1363+, 1364+, 1407-	35	
RcPR-1g	AT~TATA-box	Core promoter / enhancer element	322+, 818-, 407+, 1420-, 320-, 820-, 495+, 59+, 388+	9	

RcPR-1g	TATA-box	Core promoter element around -30 of transcription start	59+, 388+, 234-, 318-, 232-, 405+, 313+, 389+, 61+, 406-, 306+, 390+, 233-, 407+, 314+, 408+, 60+, 409+, 235+, 418-, 231-, 494-, 311-, 495+, 496+, 497+, 386+, 320-, 552+, 387-, 553+, 319-, 691-, 321-, 817-, 322+, 818-, 324+, 819-, 820-, 821+, 822-, 903-, 927-, 988+, 989-, 990-, 1033-, 1035-, 1118-, 1119-, 1120-, 1121-, 1138-, 1139-, 1140-, 1141-, 1158-, 1159-, 1188-, 1420-, 1421+, 1422-	63	
RcPR-1g	Unnamed_1	60K protein binding site	8+, 512-	2	
RcPR-1g	Unnamed_4		16-, 240+, 130+, 266+	4	
RcPR-1g	Unnamed_6		228-	1	
RcPR-1h	AAGAA-motivo	Light response	1440+	1	
RcPR-1h	ACE	Light response	599+	1	
RcPR-1h	AT1-motivo	Light response	527-, 910+	2	
RcPR-1h	Box 4	Light response	844-, 81+, 198+, 1366-	4	
RcPR-1h	GATA-motivo	Light response	308+	1	14
RcPR-1h	GT1-motivo	Light response	396+, 521-	2	
RcPR-1h	LAMP-element	Light response	113-	1	
RcPR-1h	TCCC-motivo	Light response	1195+	1	
RcPR-1h	TCT-motivo	Light response	1481+	1	
RcPR-1h	MYB	Cell cycle control in eukaryotes	504+, 522+	2	2
RcPR-1h	MYC	Drought response	12+, 1359+, 1289-, 97+, 1352-	5	5
RcPR-1h	Myb	Water response	4+, 634-, 65-, 1125+	4	4

RcPR-1h	ABRE	ABA response	1425+	1	1
RcPR-1h	TCA-element	Salicylic acid responsiveness	1072+	1	1
RcPR-1h	ERE	Ethylene-responsive element	550+, 1111-, 1109+, 1460+	4	4
RcPR-1h	ARE	Anaerobic induction	25+, 1136-, 630+	3	3
RcPR-1h	O2-site	Zein metabolism regulation	176+	1	1
RcPR-1h	STRE	Nutrient starvation stress response	311+	1	1
RcPR-1h	TC-rich repeats	Defense and stress responsiveness	877+	1	1
RcPR-1h	W box	Wounding and pathogenic response	1068+, 1217+	2	2
RcPR-1h	WUN-motivo	Wound-responsive element	588+	1	1
RcPR-1h	AT~TATA-box	Core promoter / enhancer element	159+, 835-, 223+, 1388-, 221+, 837-, 225+	7	7
RcPR-1h	CAAT-box	Common cis-acting element in promoter and enhancer regions	87-, 364-, 195+, 407+, 99-, 232+, 196+, 408+, 97+, 434+, 139+, 440+, 441+, 443-, 469+, 499+, 508+, 525+, 526+, 595-, 663-, 709+, 760+, 812-, 958-, 1016+, 1023+, 1026-, 1057-, 1067-, 1086-, 1098-, 1134-, 1142+, 1155+, 1158-, 1216-, 1304+, 1351+, 1352+, 1354-, 1360-	42	42
RcPR-1h	MYB-like sequence		522+	1	1
RcPR-1h	TATA	Core promoter / enhancer element	692-	1	1

RcPR-1h	Unnamed_1	60K protein binding site	555+	1	1
RcPR-1h	Unnamed_4		1049+, 1200+, 1196+	3	3
RcPR-1i	3-AF1 binding site	Light responsive	300+	1	
RcPR-1i	ACE	Light responsive	480+	1	
RcPR-1i	AE-box	Light responsive	1298+	1	10
RcPR-1i	Box 4	Light responsive	939-, 1104-, 1023-	3	
RcPR-1i	GATA-motivo	Light responsive	1385+	1	
RcPR-1i	TCT-motivo	Light responsive	299-, 654+, 437+	3	
RcPR-1i	CGTCA-motivo	MeJA-responsiveness	1447+	1	2
RcPR-1i	TGACG-motivo	MeJA-responsiveness	1447-	1	
RcPR-1i	MYB	Cell cycle control in eukaryotes	164-	1	2
RcPR-1i	Myb-binding site	Cell cycle and cell proliferation response	164-	1	
RcPR-1i	MYC	Drought response	709+	1	
RcPR-1i	as-1	Drought response	1447-	1	3
RcPR-1i	MBS	Drought-inducibility	330+	1	
RcPR-1i	Myb	Water response	1173-, 330+, 757+	3	3
RcPR-1i	ERE	Ethylene-responsive element	732+, 885-	2	2
RcPR-1i	ARE	Anaerobic induction	1137-, 1456+	2	2
RcPR-1i	STRE	Nutrient starvation stress response	1420-	1	1

RcPR-1i	TC-rich repeats	Defense and stress responsiveness	181-	1	1
RcPR-1i	LTR	Low-temperature responsiveness	110+	1	1
RcPR-1i	WUN-motivo	Wound-responsive element	1266+	1	1
RcPR-1i	AP-1	Cadmium (Cd) response	178+	1	1
RcPR-1i	AT~TATA-box	Core promoter / enhancer element	802-	1	2
RcPR-1i	TATA	Core promoter / enhancer element	816-	1	
RcPR-1i	CAAT-box	Common cis-acting element in promoter and enhancer regions	57+, 84-, 189-, 313+, 314+, 512-, 520-, 628-, 699+, 709+, 711-, 718-, 763-, 780-, 834-, 919-, 972-, 1036-, 1087+, 1112+, 1239-, 1306-, 1362-, 1441+, 1450+	25	25
RcPR-1i	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	300-, 652+	2	2
RcPR-1i	TATA-box	Core promoter element around -30 of transcription start	209+, 210+, 265+, 360-, 453-, 455+, 497-, 498-, 499-, 500-, 501+, 565-, 567+, 634+, 668+, 767-, 800+, 801-, 802-, 803+, 804-, 810-, 811-, 812-, 813-, 817-, 818-, 819-, 820-, 861+, 862-, 863-, 868-, 869-, 870-, 871-, 872-, 883+, 945-, 946-, 947-, 948-, 967+, 968-, 992-, 1007+, 1008-, 1009-, 1027-, 1028-, 1029-, 1030-, 1041+, 1042-, 1061-, 1062-, 1064+,	67	67

1075+, 1076-, 1077-, 1180-, 1181-,
1261+, 1262-, 1263-, 1461+, 1462-

RcPR-1i	Unnamed_4		203-, 1227-, 473-	3	3
RcPR-1i	Unnamed_6		948+	1	1
RcPR-4a.1	AAGAA-motivo	Light response	780-	1	
RcPR-4a.1	ATC-motivo	Light response	1378+	1	
RcPR-4a.1	Box 4	Light response	121+,1122-,988-,1218-,973-,1000-	6	9
RcPR-4a.1	chs-Unit 1 m1	Light response	909+	1	
RcPR-4a.1	MYB	Cell cycle control in eukaryotes	245+	1	1
RcPR-4a.1	MYC	Drought response	750-,145+	2	2
RcPR-4a.1	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1368-	1	1
RcPR-4a.1	ERE	Ethylene-responsive element	195+	1	1
RcPR-4a.1	ARE	Anaerobic induction	1489-	1	1

RcPR-4a.1	TCA	Nutrient starvation stress response	653+	1	1
RcPR-4a.1	box S	Wounding and pathogenic response	820+	1	1
RcPR-4a.1	GCN4_motivo	Endosperm expression	1174-	1	1
RcPR-4a.1	WUN-motivo	Wound-responsive element	74+	1	1
RcPR-4a.1	CAAT-box	Common cis-acting element in promoter and enhancer regions	44+,45+,144+,145+,147-,183+,239+,240+,277-,319-,392-,457+,497+,503+,508+,695+,745+,750+,800-,845-,963-,1063-,1144+,1279-,1290+,1291+,1314-	27	27
RcPR-4a.1	MYB-like sequence		245+	1	1
RcPR-4a.1	AT~TATA-box	Core promoter / enhancer element	1070-,1155-,1149-,1403-,1134-,1153-,1151-,309+,1147-	9	9
RcPR-4a.1	TATA-box	Core promoter element around -30 of transcription start	229+,270+,269+,309+,242+,310+,243+,311+,462+,536+,949-,1050-,1069-,1070-,1072-,1116-,1117-,1118-,1119-,1134-,1135-,1136-,1146-,1147-,1148-,1149-,1150-,1151-,1152-,1153-,1154-,1155-,1156+,1157-,1213+,1214-,1215-,1397-,1398-,1401+,1402-,1403-,1404+,1405-	44	44
RcPR-4a.1	Unnamed_4		223-,1473-,1227-,617-,1288+	5	5
RcPR-4a.2	4cl-CMA1b	Light responsive	971+	1	
RcPR-4a.2	Box 4	Light responsive	121+, 987-, 629+, 1463-, 164+, 1152-, 668+	7	
RcPR-4a.2	GATA-motivo	Light responsive	1071-	1	11
RcPR-4a.2	F-box	Light response	183+	1	
RcPR-4a.2	GT1-motivo	Light responsive	1037+	1	

RcPR-4a.2	MYB	Cell cycle control in eukaryotes	1036-	1	1
RcPR-4a.2	MYC	Drought response	145+, 1202-, 638-, 212+, 1185+	5	6
RcPR-4a.2	MBS	Drought-inducibility	934+	1	
RcPR-4a.2	Myb	Water response	934+, 1362+	2	2
RcPR-4a.2	TGA-element	Auxin-responsive element	1232-, 375-	2	2
RcPR-4a.2	CAT-box	Meristem expression	1403+, 1415-	2	2
RcPR-4a.2	ERE	Ethylene-responsive element	253+, 1268+, 405-, 255-, 893+	5	5
RcPR-4a.2	STRE	Nutrient starvation stress response	1059+	1	1
RcPR-4a.2	CTAG-motivo	Determinant for repressor binding	544-	1	1
RcPR-4a.2	MYB-like sequence		1036-	1	1
RcPR-4a.2	AT~TATA-box	Core promoter / enhancer element	242-, 244+, 818-	3	3
RcPR-4a.2	CAAT-box	Common cis-acting element in promoter and enhancer regions	144+, 321-, 171+, 495+, 145+, 484+, 213-, 529+, 85-, 382-, 147-, 531-, 411-, 723+, 734-, 757-, 812+, 868+, 912-, 974+, 975+, 1033-, 1107-, 1122+, 1130+, 1141+, 1142+, 1150+, 1202+, 1250+, 1256+, 1257+, 1258+, 1331+, 1332+, 1450+, 1467-	37	37
RcPR-4a.2	TATA-box	Core promoter element around -30 of transcription start	13+, 1004-, 1002-, 1299-, 240-, 1006+, 929-, 241-, 242-, 243-, 244+, 246+, 261+, 262+, 449+, 647+, 714+, 715+, 769-, 770-, 771-, 772-, 817-, 818-, 819-, 820-, 895-, 901-, 928-	29	29
RcPR-4a.2	Unnamed_4		517+, 1295+	2	2

RcPR-4a.2	Unnamed_6		237-	1	1
RcPR-4b	AAGAA-motivo	Light response	250+, 900-, 635+	3	
RcPR-4b	AE-box	Light response	328+	1	
RcPR-4b	ATC-motivo	Light response	1282+	1	
RcPR-4b	ATCT-motivo	Light response	484+, 910+	2	
RcPR-4b	Box 4	Light response	354+, 1050-, 365+, 358+, 728+	5	
RcPR-4b	Box II	Light response	17+	1	18
RcPR-4b	G-Box	Light response	839+	1	
RcPR-4b	LAMP-element	Light response	280-	1	
RcPR-4b	Sp1	Light response	1260+	1	
RcPR-4b	TCT-motivo	Light response	1194-	1	
RcPR-4b	chs-CMA2a	Light response	1405-	1	
RcPR-4b	CCGTCC motivo	Cell cycle and cell proliferation response	1135-	1	2
RcPR-4b	MYB	Cell cycle control in eukaryotes	797+	1	
RcPR-4b	DRE core	Drought response	1098+	1	
RcPR-4b	MYB recognition site	Drought response	1438-	1	8
RcPR-4b	MYC	Drought response	50-, 1181-, 761+, 52+, 778-	5	
RcPR-4b	as-1	Drought response	1119-	1	
RcPR-4b	Myb	Water response	949+	1	1
RcPR-4b	ABRE	ABA response	839-	1	1
RcPR-4b	TGA-element	Auxin-responsive element	241-	1	1

RcPR-4b	CAT-box	Meristem expression	990+, 1002-	2	3
RcPR-4b	CCGTCC-box	Meristem expression	1135-	1	
RcPR-4b	CGTCA-motivo	MeJA-responsiveness	1119+	1	2
RcPR-4b	TGACG-motivo	MeJA-responsiveness	1119-	1	
RcPR-4b	ARE	Anaerobic induction	375-, 620+	2	2
RcPR-4b	STRE	Nutrient starvation stress response	1258+	1	1
RcPR-4b	TC-rich repeats	Defense and stress responsiveness	287-	1	1
RcPR-4b	W box	Wounding and pathogenic response	250-, 1164-, 1158-	3	3
RcPR-4b	WRE3	Wounding and pathogenic response	211+	1	1
RcPR-4b	GCN4_motivo	Endosperm expression	93+, 460+	2	2
RcPR-4b	A-box	Cis-acting regulatory element	1135-	1	1
RcPR-4b	CAAT-box	Common cis-acting element in promoter and enhancer regions	46-, 98+, 107+, 112-, 152-, 260-, 278-, 310+, 369-, 552+, 602-, 623+, 624+, 656+, 682+, 706+, 718+, 778+, 824+, 830+, 831+, 832+, 908+, 918+, 919+, 960-, 1037+, 1047+, 1054-, 1181+, 1228-, 1314+, 1373-, 1386+	34	34
RcPR-4b	CCAAT-box	MYBHv1 binding site	1438+	1	1
RcPR-4b	HD-Zip 1	Palisade mesophyll cells	107+	1	1
RcPR-4b	MYB-like sequence		797+	1	1

RcPR-4b	TATA-box	Core promoter element around -30 of transcription start	82-, 83+, 168+, 301+, 302+, 378-, 379-, 380+, 429-, 430-, 431-, 432+, 437+, 517-, 518-, 519-, 520+, 531-, 532-, 533-, 534+, 589+, 590-, 591+, 615+, 616-, 617+, 794-, 795-, 886-, 1224-, 1225-, 1329-	33	33
RcPR-4b	TATA	Core promoter / enhancer element	530-	1	1
RcPR-4b	Unnamed_4		156-, 1400-, 513+, 1442-, 493+, 1277+	6	6
RcPR-4b	Unnamed_1	60K protein binding site	838-	1	1
<hr/>					
RcPR-5a	ARE	Light responsiveness	305-, 1105+, 634+	3	
RcPR-5a	ATCT-motivo	Light responsiveness	674+	1	
RcPR-5a	Box 4	Light responsiveness	131+, 513+, 391+, 1350-	4	
RcPR-5a	CCAAT-box	Light responsiveness	417-	1	22
RcPR-5a	G-Box	Light responsiveness	201+, 313+, 266+, 809-,	4	
RcPR-5a	G-box	Light responsiveness	201+, 411-, 266+, 226-, 313+	5	
RcPR-5a	GATA-motivo	Light responsiveness	696+, 1377-	2	
RcPR-5a	LAMP-element	Light responsiveness	1185+	1	
RcPR-5a	MYB	Cell cycle control in eukaryotes	1164+	1	1
RcPR-5a	MYB recognition site	Drought response	417+	1	3
RcPR-5a	MYC	Drought response	593-, 923	2	

RcPR-5a	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1258-	1	1
RcPR-5a	Myb	Water response	214+, 1334-	2	2
RcPR-5a	TATC-box	Gibberellin-responsiveness	95+	1	1
RcPR-5a	TCA-element	Salicylic acid responsiveness	716-	1	1
RcPR-5a	LTR	Low-temperature responsiveness	1470-	1	1
RcPR-5a	STRE	Nutrient starvation stress response	185+, 293+	2	2
RcPR-5a	AT~TATA-box	Core promoter / enhancer element	1355-	1	1
RcPR-5a	TATA-box	Core promoter element around -30 of transcription start	385-, 899-, 387+, 1135-, 7+, 1114+, 845-, 1357-, 286-, 1003-, 621-, 1355-, 386-, 1115-, 898-, 1363-, 8+, 900-, 620-, 1354-, 384-, 1002+, 622+, 1356+	24	24
RcPR-5a	CAAT-box	Common cis-acting element in promoter and enhancer regions	45+, 50+, 54+, 88+, 104-, 151-, 364-, 445+, 481-, 493-, 498-, 602-, 651+, 673+, 737+, 738+, 747+, 795+, 913-, 923+, 1012-, 1069+, 1077-, 1201+, 1203-, 1405-, 1430-, 1484-	28	28
RcPR-5a	CAG-motivo	MYBHv1 binding site	64+	1	1

RcPR-5a	Unnamed_1	60K protein binding site	265-	1	1
RcPR-5a	Unnamed_4		162-, 561-, 402+, 1152+, 193-, 724-, 505-, 1211+, 188-, 598-, 240-, 859-	12	12
RcPR-5b	AAGAA-motivo	Light responsiveness	1454+, 452+	2	
RcPR-5b	Box 4	Light responsiveness	53+, 811-, 175+, 819-, 47+, 815-, 676+	7	11
RcPR-5b	GATA-motivo	Light responsiveness	1199+	1	
RcPR-5b	GT1-motivo	Light responsiveness	1171-	1	
RcPR-5b	CCGTCC motivo	Cell cycle and cell proliferation response	121+	1	3
RcPR-5b	MYB	Cell cycle control in eukaryotes	1222-, 1172+	2	
RcPR-5b	MYC	Drought response	63-, 1176+	2	3
RcPR-5b	as-1	Drought response	1238-	1	
RcPR-5b	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1376-	1	1
RcPR-5b	Myb	Water response	78+	1	1
RcPR-5b	CCGTCC-box	Meristem specific activation	121+	1	1
RcPR-5b	ERE	Ethylene-responsive element	365-, 997-	2	2
RcPR-5b	ARE	Anaerobic induction	243+, 489-	2	2

RcPR-5b	W box	Wounding and pathogenic response	323+	1	1
RcPR-5b	CGTCA-motivo	MeJA-responsiveness	1238+	1	2
RcPR-5b	TGACG-motivo	MeJA-responsiveness	1238-	1	
RcPR-5b	MYB-like sequence		1172+, 1222-	2	2
RcPR-5b	A-box	Cis-acting regulatory element	121+	1	1
RcPR-5b	AT~TATA-box	Core promoter / enhancer element	28+, 1006-, 1004-, 1302-, 774-, 1008-, 840-, 1362-, 92+, 776-	10	12
RcPR-5b	TATA	Core promoter / enhancer element	702+, 559+	2	
RcPR-5b	TATA-box	Core promoter element around -30 of transcription start	28+, 92+, 72+, 311-, 29+, 310-, 91+, 27+, 64-, 30+, 312-, 73+, 93+, 313+, 94+, 345+, 346+, 362+, 363+, 392+, 436+, 542-, 544-, 545-, 546+, 556-, 557-, 558-, 559+, 632+, 683+, 701+, 702+, 773-, 774-, 775-, 776-, 777+, 778-, 839-, 840-, 841-, 842-, 878-, 879-, 883+, 884-, 885-, 925-, 926-, 927-, 928-, 929-, 1001-, 1003-, 1004-, 1005-, 1006-, 1007-, 1008-, 1009-, 1010-, 1302-, 1303+, 1304-, 1360-, 1362-, 1363+, 1364-	69	69
RcPR-5b	CAAT-box	Common cis-acting element in promoter and enhancer regions	113+, 874-, 495+, 1401+, 155+, 1177-, 758+, 1489-, 150+, 1018-, 591+, 1426+, 246+, 1313-, 765+, 1408+, 63+, 942+, 299+, 209-, 1435+, 723+, 151+, 247+, 1258+, 1061-, 1339+	27	27
RcPR-5b	Unnamed_4		162-, 561-, 402+, 1152+, 193-, 724-, 505-, 1211+, 188-, 598-, 240-, 859-	12	12
RcPR-5c	AAGAA-motivo	Light response	580+	1	12

RcPR-5c	Box 4	Light response	298+, 335+, 166+, 912-, 197+, 446+, 306+, 1401-	8	
RcPR-5c	G-box	Light response	1454-	1	
RcPR-5c	I-box	Light response	1225-	1	
RcPR-5c	TCT-motivo	Light response	1385-	1	
RcPR-5c	MYB	Cell cycle control in eukaryotes	454+	1	1
RcPR-5c	as-1	Drought response	736-	1	1
RcPR-5c	ABRE	ABA response	1455+	1	1
RcPR-5c	P-box	Gibberellin-responsive element	1359+	1	1
RcPR-5c	TCA-element	Salicylic acid responsiveness	1017+	1	1
RcPR-5c	ARE	Anaerobic induction	1009+	1	1
RcPR-5c	STRE	Nutrient starvation stress response	513+	1	1
RcPR-5c	W box	Wounding and pathogenic response	1334-	1	1
RcPR-5c	CGTCA-motivo	MeJA-responsiveness	736+	1	2
RcPR-5c	TGACG-motivo	MeJA-responsiveness	736-	1	
RcPR-5c	WUN-motivo	Wound-responsive element	272-	1	1
RcPR-5c	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)	364+	1	1
RcPR-5c	AT~TATA-box	Core promoter / enhancer element	267+, 980-, 227+	3	3

RcPR-5c	CAAT-box	Common cis-acting element in promoter and enhancer regions	85+, 180-, 195+, 262-, 333+, 418+, 435+, 463-, 543+, 641-, 669+, 727+, 787+, 815+, 989-, 1001+, 1013+, 1153+, 1161+, 1200-, 1305+, 1331-, 1396-, 1412+	24	24
RcPR-5c	MYB-like sequence		454+	1	1
RcPR-5c	TATA-box	Core promoter element around -30 of transcription start	99+, 982-, 980-, 1344-, 226-, 1139-, 979-, 1137+, 98+, 1185+, 227+, 228+, 229+, 243+, 266+, 267+, 268+, 269+, 294+, 323+, 324+, 339-, 340-, 341-, 342+, 378-, 379-, 380+, 450+, 451-, 452+, 459+, 460+, 527+, 528-, 529+, 614+, 711+, 712+	39	39
RcPR-5c	Unnamed_4		94-, 1311+, 1087-	3	3
RcPR-5d	Box 4	Light responsiveness	776-	1	
RcPR-5d	G-box	Light responsiveness	1397-	1	3
RcPR-5d	LAMP-element	Light responsiveness	173+	1	
RcPR-5d	MYB	Cell cycle control in eukaryotes	577+, 1323-	2	2
RcPR-5d	MYC	Drought response	936+, 467+	2	3
RcPR-5d	MBS	Drought-inducibility	386+	1	
RcPR-5d	Myb	Water response	190+, 386+, 276+	3	3
RcPR-5d	MYB-like sequence		577+	1	1
RcPR-5d	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	1171-	1	1

RcPR-5d	ABRE	ABA response	1397-	1	
RcPR-5d	ABRE3a	ABA response	1397-	1	3
RcPR-5d	ABRE4	ABA response	1397+	1	
RcPR-5d	P-box	Gibberellin-responsive element	1351+	1	2
RcPR-5d	TATC-box	Gibberellin-responsive element	1421+	1	
RcPR-5d	TCA-element	Salicylic acid responsiveness	566-	1	1
RcPR-5d	CAT-box	Meristem expression	625+	1	1
RcPR-5d	ARE	Anaerobic induction	813-	1	1
RcPR-5d	W box	Wounding and pathogenic response	464-, 860-	2	2
RcPR-5d	CAAT-box	Common cis-acting element in promoter and enhancer regions	17-, 44+, 58+, 162-, 269+, 298-, 360-, 405-, 423+, 467+, 469-, 483+, 495+, 526-, 532-, 593+, 633+, 719-, 750+, 810-, 907+, 937-, 1125+, 1126+, 1160+, 1207-, 1248-, 1275+, 1299-, 1321-, 1335+	31	31
RcPR-5d	TATA	Core promoter / enhancer element	827+, 1436+	2	4
RcPR-5d	AT~TATA-box	Core promoter / enhancer element	272+, 1026-	2	
RcPR-5d	TATA-box	Core promoter element around -30 of transcription start	61+, 133+, 134+, 158-, 159+, 170-, 247-, 248-, 249-, 250+, 271+, 272+, 273+, 274+, 305-, 307+, 349-, 400-, 401-, 402+, 424-, 454+, 455-, 456+, 684+, 685-, 686+, 784-, 786-, 797-, 820-, 821-, 822-, 827-, 871-, 920-, 921-, 922-, 923-, 947-, 1025-, 1026-, 1028-, 1200-, 1435+, 1436-, 1475-, 1477-	48	48

RcPR-5d	Unnamed_1	60K protein binding site	825-	1	
					2
RcPR-5d	Unnamed_1	60K protein binding site	1275+	1	
RcPR-5d	Unnamed_4		69-, 1307-, 320-, 1387+, 225-, 664+	6	6
RcPR-5e	ACE	Light responsiveness	837+	1	
					2
RcPR-5e	Box 4	Light responsiveness	919-	1	
RcPR-5e	CCGTCC motivo	Cell cycle and cell proliferation response	181+	1	
RcPR-5e	Myb-binding site	Cell cycle and cell proliferation response	620+	1	8
RcPR-5e	MYB	Cell cycle control in higher eukaryotes	249-, 1310+, 620+, 1314+, 334-, 1306+	6	
RcPR-5e	MYC	Drought response	406+, 1318-, 832-	3	
					5
RcPR-5e	MBS	Drought-inducibility	139-, 349+	2	
RcPR-5e	Myb	Water response	139-, 1055-, 349+	3	3
RcPR-5e	TATC-box	Gibberellin-responsiveness	6+	1	1
RcPR-5e	CAT-box	Meristem expression	948+	1	
					2
RcPR-5e	CCGTCC-box	Meristem expression	181+	1	
RcPR-5e	ARE	Anaerobic induction	533+	1	1
RcPR-5e	TC-rich repeats	Defense and stress responsiveness	688-	1	1
RcPR-5e	A-box	Cis-acting regulatory element	181+	1	1

RcPR-5e	AT~TATA-box	Core promoter / enhancer element	1426-, 1442-, 1432-, 1448-, 1428-, 1444-, 1436-, 1452-, 1150-, 1440-, 1434-, 1450-, 1430-, 1446-, 1438-, 1454-	16	16
RcPR-5e	CAAT-box	Common cis-acting element in promoter and enhancer regions	27+, 55-, 65-, 79-, 84+, 130+, 213-, 236+, 331-, 407-, 480+, 486+, 677-, 751-, 783+, 832+, 935+, 1043+, 1177+, 1188+, 1197+, 1264-, 1336+, 1356-, 1481+	25	25
RcPR-5e	CARE	GA response	75-	1	1
RcPR-5e	MYB-like sequence		249-, 1306+, 334-	3	3
RcPR-5e	Myc	Myc is a renowned transcription factor controlling disparate aspects of cell physiology that, together, allow efficient proliferation of somatic cells.	818-	1	1
RcPR-5e	TATA-box	Core promoter element around -30 of transcription start	29+, 30+, 319+, 320+, 337-, 338+, 355-, 356-, 357-, 358+, 372-, 416+, 417-, 418+, 558+, 764+, 765-, 766-, 930-, 1021+, 1022-, 1023-, 1071-, 1096+, 1097-, 1098-, 1104+, 1150-, 1151-, 1152-, 1351+, 1352-, 1353-, 1381-, 1419-, 1420-, 1421-, 1424+, 1425-, 1426-, 1427-, 1428-, 1429-, 1430-, 1431-, 1432-, 1433-, 1434-, 1435-, 1436-, 1437-, 1438-, 1439-, 1440-, 1441-, 1442-, 1443-, 1444-, 1445-, 1446-, 1447-, 1448-, 1449-, 1450-, 1451-, 1452-, 1453-, 1454-, 1455-, 1456-	70	70
RcPR-5e	Unnamed_16		176+	1	1
RcPR-5e	Unnamed_4		587+, 1036+, 891+	3	3

RcPR-6	AE-box	Light response	582+	1	
RcPR-6	Box 4	Light response	4+, 990-, 154+	3	
RcPR-6	I-box	Light response	1450+	1	
RcPR-6	MRE	Light response	882-	1	10
RcPR-6	Pc-CMA2a	Light response	614+	1	
RcPR-6	TCCC-motivo	Light response	1337-	1	
RcPR-6	TCT-motivo	Light response	523-, 390+	2	
RcPR-6	MYB	Cell cycle control in eukaryotes	827-, 815-	2	3
RcPR-6	Myb-binding site	Cell cycle and cell proliferation response	815-	1	
RcPR-6	Myb	Water response	1261-	1	1
RcPR-6	ERE	Ethylene-responsive element	458+, 974+, 503+, 465+, 505-	5	5
RcPR-6	ARE	Anaerobic induction	28-, 1012-	2	2
RcPR-6	TC-rich repeats	Defense and stress responsiveness	343+, 386+	2	2
RcPR-6	LTR	Low-temperature responsiveness	1151+	1	1
RcPR-6	AT~TATA-box	Core promoter / enhancer element	1395-, 1407-, 1401-, 1413-, 1397-, 1409-, 1403-, 1415-, 733+, 1405-, 1399-, 1411-	12	12
RcPR-6	CAAT-box	Common cis-acting element in promoter and enhancer regions	74+, 537+, 411-, 618+, 302+, 617+, 536+, 703-, 129+, 578+, 373-, 762-, 587+, 836+, 850-, 872-, 913+, 927+, 943-, 961-, 988+, 1000-, 1010-, 1140+, 1164+, 1165+, 1213-, 1252+, 1361-	29	29
RcPR-6	MYB-like sequence		827-	1	1

RcPR-6	TATA	Core promoter / enhancer element	81-, 396+, 170- 82-, 83-, 84-, 85+, 163-, 164-, 165-, 166+, 171-, 172-, 173-, 174+, 256+, 257+, 396+, 439+, 440-, 441+, 447-, 497+, 498+, 732+, 733+, 734+, 735+, 777-, 778-, 779-, 780-, 781-, 839-, 840-, 841-, 842-, 867-, 994-, 995-, 996-, 997-, 1041-, 1042-, 1043-, 1095- , 1096-, 1394-, 1395-, 1396-, 1397-, 1398-, 1399-, 1400-, 1401-, 1402-, 1403-, 1404-, 1405-, 1406-, 1407-, 1408-, 1409-, 1410-, 1411+, 1412-, 1413-, 1414-, 1415-, 1416-, 1417-, 1428+, 1430-	3	3
RcPR-6	TATA-box	Core promoter element around -30 of transcription start	81-, 396+, 170- 82-, 83-, 84-, 85+, 163-, 164-, 165-, 166+, 171-, 172-, 173-, 174+, 256+, 257+, 396+, 439+, 440-, 441+, 447-, 497+, 498+, 732+, 733+, 734+, 735+, 777-, 778-, 779-, 780-, 781-, 839-, 840-, 841-, 842-, 867-, 994-, 995-, 996-, 997-, 1041-, 1042-, 1043-, 1095- , 1096-, 1394-, 1395-, 1396-, 1397-, 1398-, 1399-, 1400-, 1401-, 1402-, 1403-, 1404-, 1405-, 1406-, 1407-, 1408-, 1409-, 1410-, 1411+, 1412-, 1413-, 1414-, 1415-, 1416-, 1417-, 1428+, 1430-	70	70
RcPR-6	Unnamed_1	60K protein binding site	854+	1	1
RcPR-6	Unnamed_4		1339-, 1102+, 1426+	3	3
RcPR10-STH21	ATCT-motivo	Light responsiveness	1418+	1	
RcPR10-STH21	Box 4	Light responsiveness	699+	1	4
RcPR10-STH21	G-box	Light responsiveness	592+	1	
RcPR10-STH21	GT1-motivo	Light responsiveness	714+	1	
RcPR10-STH21	ABRE	ABA response	593+	1	
RcPR10-STH21	ABRE3a	ABA response	592+	1	3
RcPR10-STH21	ABRE4	ABA response	592-	1	
RcPR10-STH21	ARE	Anaerobic induction	562+, 38-	2	2
RcPR10-STH21	O2-site	Zein metabolism regulation	520+	1	1

RcPR10-STH21	TC-rich repeats	Defense and stress responsiveness	650+	1	1
RcPR10-STH21	circadian	Circadian control	668+	1	1
RcPR10-STH21	AT~TATA-box	Core promoter / enhancer element	692+, 694+, 543+	3	3
RcPR10-STH21	TATA-box	Core promoter element around -30 of transcription start	52-, 600+, 338+, 694+, 95+, 692+, 544+, 696+, 53+, 674+, 543+, 695+, 319+, 693+, 545+, 673+, 51+, 691+, 339-, 94-, 340+	21	21
RcPR10-STH21	CAAT-box	Common cis-acting element in promoter and enhancer regions	228+, 706-, 375+, 1422+, 281+, 787+, 472-, 1445-, 159+, 770-, 406-, 1421+, 372-, 584+	14	14
RcPR10-STH21	Unnamed_2		713-, 529+	2	2
RcPR10-STH21	Unnamed_4		509+, 802+, 609+, 575+, 749-	5	5

Figura suplementar 6. Alinhamento múltiplo entre Proteínas Relacionadas a Patogênese em *Ricinus communis* (*RcPRs*), *Nicotiana tabacum* (*NtPRs*) e *Arabidopsis thaliana* (*AtPRs*).

PR-1 <i>R. communis</i>	-----MSNGLFFAEGRFQCNC	17
PR-1 <i>R. communis</i>	ML--SCGALGTFLLF--LLLASASEVSS-----GR-----SSLI RTQ	33
PR-1 <i>N. tabacum</i>	MKPFLLFLPLFLSKITSIHTL-QIPSKSLPLSLDQTTLNKNLNDNTIYKISKQLCWC	59
PR-1 <i>R. communis</i>	MKPHSFLLLIFLITIFIT--NNL-LVTSQ--SPSPAETPLNKRNDTIYKSKQLCWC	56
PR-1 <i>R. communis</i>	MRTYCYCVVAISIS--LFLTTR-ATQSSA-----AAESAIA-----HQHAY	40
PR-1 <i>A. thaliana</i>	---MFHKVAIETLVLLLINYLQIDVSS---AQY---S-----QYPQS	35
PR-1 <i>R. communis</i>	MYL-----NDFFA-ATVMVGTVLVHI-----SLA	23
PR-1B <i>N. tabacum</i>	MGFFLFSQMPSSFLLV-STLLFLIISHS-----SHA	30
PR-1 <i>A. thaliana</i>	MNFTGYSR---FLI-VFVALVGVLP-----SKA	26
PR-1 <i>A. thaliana</i>	MKVTSYSR---ILI-ILAAALVGVLP-----LKA	26
PR-1 <i>A. thaliana</i>	MKIFNSSQ---NLF-LAITFFLVLIVH-----LKA	26
PR-1 <i>N. tabacum</i>	--MGFLT---TIV-ACFITFAILIHS-----SKA	23
PR-1 <i>R. communis</i>	MAMPKVSL---PLF-YFMIMSLALLFS-----SNA	26
PR-1 <i>R. communis</i>	--MLSLKI---SAA-LSFLISLSLIS-----SHA	24
PR-1 <i>R. communis</i>	--MKQFDP---SLA-LVCIISIALIFS-----SHA	24
PR-1 <i>R. communis</i>	--MGFNK---FIV-ALCLMTLITPV-----SLA	23
PR-1 <i>R. communis</i>	--MGFTN---LVI-ALCLVASALSPT-----CLA	23
PR-1 <i>R. communis</i>	QSLFSSLSLWLFNRDITFNTCTPNWGFSSWINAVEPAFPYSGSQLVFSLMGAKDWESR	77
PR-1 <i>R. communis</i>	RETIKQYKPHNRERAKL-----GLRPLKWSNKLA-----SFASSWAHQ	72
PR-1 <i>N. tabacum</i>	LGEAIQFLFHNLVVRATK-----MELPLTWDSNLE-----KYAKWASS	98
PR-1 <i>R. communis</i>	IGEAQLFLYAHNLVVRASK-----WELPLTWDSQLE-----RYARWAGT	95
PR-1 <i>R. communis</i>	LSTANQFLGPHNAARAAL-----RMPPLIWDTRLA-----NYAQWYANQ	79
PR-1 <i>A. thaliana</i>	HEYPSYLRPHNAARAIV-----KVKPLRWDGIA-----TVAQDYANH	74
PR-1 <i>R. communis</i>	QNSPQDYVNAHNTVRAEI-----GVGPITWNTVVA-----AYAQTAYS	62
PR-1B <i>N. tabacum</i>	QNSQDYVLDHNTARADV-----GVEPLTWDNVA-----AYAQNYVSQ	69
PR-1 <i>A. thaliana</i>	QDSQDYLRVHNQARGAV-----GVGPMQWDERVA-----AYARSYAEQ	65
PR-1 <i>A. thaliana</i>	QDSQDYVNAHNAARSQI-----GVGPMQWDEGLA-----AYARNYANQ	65
PR-1 <i>A. thaliana</i>	QDSQDYLVAVHNRARAIV-----GVGPLRWDEKVA-----AYARNYANQ	65
PR-1 <i>N. tabacum</i>	QNSPQDYLRPHNAARRQV-----GVGPMTWDRNLA-----AYAQNYANQ	62
PR-1 <i>R. communis</i>	QDTNQDYINSHNDARVAVGA-----GLGNTWNETVA-----DYARDYANQ	67
PR-1 <i>R. communis</i>	QNLQDYVNAHNAARAIV-----GVGPMTWNTVA-----AYAQNYANQ	63
PR-1 <i>R. communis</i>	QDTQDYVNAHNSARGDV-----GLGPLTWDDKVA-----SYAQHMANQ	63
PR-1 <i>R. communis</i>	QNSQDYVAAHNEVRAEL-----GLGPVWNEKLA-----LYARKYIQT	62
PR-1 <i>R. communis</i>	RNSRGDYLDHNAVRAEV-----GVDPLVWNTKLA-----DYAKASP--	60
	: * : * :	
PR-1 <i>R. communis</i>	R-----KSPVNVVWSPMDWVRPPYNYFR-----NFETIKESDSTIQ	112
PR-1 <i>R. communis</i>	R-QGDCALLHSNS---NYGENLFWGSGKDWKPGDAVAWAEEKCYNHNTNTCTKNC--	125
PR-1 <i>N. tabacum</i>	R-KEDCRLMHSFPESDFKLGENIYWGSGTWTPTDAVNAWADEKQYNYASNSCVEGQ--	155
PR-1 <i>R. communis</i>	R-KQDCQLEHSFPEDGFKLGENIYWGSGTAWTPRDAVSAWASEEKYYTYATNSCEEQ--	152
PR-1 <i>R. communis</i>	R-RFDCDLRHSNG---PYGENIYWGSGTGWTPAQAVTAWVSEKWNYYNSNSCYGHQ--	132
PR-1 <i>A. thaliana</i>	LASGPCSLHSSG---PYGENLAFSGSDM-SAAQAVAMWVHEKSYDYFYSNSCHGP---	126
PR-1 <i>R. communis</i>	R-IESCEFEHSYG---PYGENIAEGYGNL-NGDVAVNMWSEKPNYDYGNSCKGGE-D	115
PR-1B <i>N. tabacum</i>	L-AADCNLVHSHG---QYGENLAQSGSDMFAAKAVEMWDEKQYDHSNTCAAGQ---	122
PR-1 <i>A. thaliana</i>	L-RGNCRLLHSGG---PYGENLAWGSGDL-SGVSAVINMWSEKANYNYAANTCNG---	115
PR-1 <i>A. thaliana</i>	L-KGDCRLVHVRG---PYGENLAKSGGDL-SGVAAVNLWNEKANYNYDTNTCNG---	115
PR-1 <i>A. thaliana</i>	R-KGDCAMKHSYG---SYGENIAWSSGSM-TGVAADVMDWDEKQYDYSNTCAAWDK--	117
PR-1 <i>N. tabacum</i>	R-IGDCGMIHSHG---PYGENLAAAFPOL-NAAGAVKMWDEKRFYDYNNSNSCVGG---	113
PR-1 <i>R. communis</i>	R-IADCNLVHSDG---PYGENLAWGSGDL-SGLDAVRMWDEKAFYDYNSTCTGGQ--	119
PR-1 <i>R. communis</i>	R-INDCRLVHSGG---RYGENIAWSSGDL-SGTAAVKLWDEKAFYNYNSNTCAAGQ--	115
PR-1 <i>R. communis</i>	H-KSDCSLVHSEG---PYGENLAWSSGDL-SGTDVAKMWIDEKPYDYNNSNSCASGQ--	115
PR-1 <i>R. communis</i>	K-VETCILEHSNG---PYGENLAKSGGEGFSGVDAVKLWDEKPNYDYNLSNSCAGG---	114
PR-1 <i>R. communis</i>	-----QPNFNAKDAVKIWADEKFFYDRKNSCKGG---	90
PR-1 <i>R. communis</i>	LVGS--WVKWRMALEEISELA--AT-----QLPFI SFISRTGNQVVDGLAK	156
PR-1 <i>R. communis</i>	DCLHYTQVWVRQSLKVGCAVICTSG-DTLITCNYDPHGVIGERPF-----	171
PR-1 <i>N. tabacum</i>	LCGHYTQVWVWSTRRVRGCAVICTSG-DVFMTCNYFPNGYIGEKPYPRB-N-tabacum	212
PR-1 <i>R. communis</i>	MCGHYTQVWVWTRRIGCARVVCDDG-DVFMTCNYDPPGNYIGEKPY-----	198
PR-1 <i>R. communis</i>	ECGHYTQVWVRKTRRIGCAKVTCSDDLGVFMTCNYDPPGNYIGERPY-----	179
PR-1 <i>A. thaliana</i>	ACGHYTQVWVRGSRARLGCCKAKCNG-ASIVVCNYDPAGNYIGARPY-----	172
PR-1 <i>R. communis</i>	ECLHYTQVWVRNSVHLGCGRAKCKTG-WMFVTCNYHPVGNIEGQRPF-----	161
PR-1B <i>N. tabacum</i>	VCGHYTQVWVRNSVRVGCARVQCNG-GYVVCNYDPPGNYIGQSPY-----	168
PR-1 <i>A. thaliana</i>	VCGHYTQVWVRKSVRLGCAKVRCCNG-GTIISCNYPGRNYVNEKPY-----	161
PR-1 <i>A. thaliana</i>	VCGHYTQVWVRNSVRLGCAKVRCCNG-GTIISCNYPGNYANQKPY-----	161
PR-1 <i>A. thaliana</i>	QCGHYTQVWVRNSERLGCARVRCNG-QTFITCNYDPPGNYVGVWEPY-----	163
PR-1 <i>N. tabacum</i>	VCGHYTQVWVRNSVRLGCAKVRCCNG-WFFITCNYDPPGNIGQRPFGDLEEQPFDSKLE	172
PR-1 <i>R. communis</i>	QCGHYTQVWVRDSISLGCARVTCNGLGLTITCNYDPPGNYVIGQRPY-----	166
PR-1 <i>R. communis</i>	QCGHYTQVWVRNSVRLGCAKVRCCNG-GTFITCNYDPPGNYVFNQRPY-----	161
PR-1 <i>R. communis</i>	ICGHYTQVWVRNSVRLGCAKVRCCNG-GTFICNYDPPGNYIGQRPY-----	162
PR-1 <i>R. communis</i>	MCGHYTQVWVRDTSQVGCARVCKKNG-WTYISCNYPGNYIGERPF-----	160
PR-1 <i>R. communis</i>	ECGHYTQVWVRDTSQVGCARVCKKNG-HTFISCNYPGNYIGQSPY-----	136
	: * : * :	

PR-STH21 <i>N. tabacum</i>	MGVTTYTHEASTTVAPTRLFKALVLDADNLIPKLMQVVKNIETVEGDGGVGSIKKMNFB	60
PR-10STH21 <i>R. communis</i>	MGVSTFTQEFSTSPVAPSRMFKALILDSNILIPKLLPQFIKSDVIQGGDGAGTIEQVNF	60
	:*:*:* :: **:*:*:**:**:**: **:**:*.....*..*:*:**:**:	
PR-STH21 <i>N. tabacum</i>	EGGPIKYLKHKLHVIDDKNLVTKYSLIEGDVLGDKLESITYDVKFETSAKGGCICKTSTE	120
PR-10STH21 <i>R. communis</i>	EASDFKYIKNRIEELDKDNLVCKYTLIEGGPLGDKLDSIAYEVKFEAASDGGCICKITSN	120
	*.. **:*:**:.. *..** **:*:**, *****:**:*:**:**:**: ***** ::	
PR-STH21 <i>N. tabacum</i>	YHTKGDYVFKEEEHNEGKEKAMELFKWEDYLLANSTVYA 160	
PR-10STH21 <i>R. communis</i>	YFTLGKFTIKEEDIEDGKGKALGIYRVVEAYLLENPQVYA 160	
	. *..:**:**: **:** **:* **:* **:* **:* **:* **:*	

Aminoácidos idênticos são indicados por um asterisco (*), substituições conservativas por dois pontos (:) e substituições semiconservativas por um ponto (.) .

Figura Suplementar 7. Motivos conservados. MEME Suits.

	Motif	Sequência Consenso
1	4.6e-123	
2	1.1e-121	
3	3.2e-081	
4	3.9e-076	
5	7.4e-050	
6	1.4e-042	
7	1.7e-038	
8	5.1e-031	
9	1.3e-019	
10	1.0e-018	
11	9.0e-016	
12	4.6e-011	
13	1.4e-008	
14	9.9e-006	
15	1.5e-004	
16	2.7e-004	

ARTIGO 2**Resolução da estrutura proteica das Proteínas Relacionadas à Patogênese (PRs) via modelagem por homologia em *Ricinus communis* L.**

¹Artigo a ser ajustado para posterior submissão ao Comitê Editorial do periódico científico *Plus One*, em versão na língua inglesa.

Resolução da estrutura proteica das Proteínas Relacionadas à Patogênese (PRs) via modelagem por homologia em *Ricinus communis* L.

RESUMO: A mamona (*Ricinus communis* L.) é um importante oleaginosa onde mamonocultura é prejudicada devido ao ataque de moléstias e pragas. Portanto, nosso trabalho teve por objetivo realizar a construção, de forma computacional, de prováveis modelos de cinco Proteínas Relacionadas à Patogênese (Proteínas PR), nomeadas neste trabalho como RcPR-1c; RcPR-4b; RcPR-5a; RcPR-6 e RcPR10-STH21. A anotação existente no PHYTOZOME V 12.1 do genoma de *R. communis* permitiu a identificação de 20 genes e suas respectivas proteínas classificados em cinco diferentes famílias: PR-1; PR-4; PR-5; PR-6 e PR-10 STH21. A partir disso, foi escolhida uma proteína PR de cada família gênica que apresentou boa identidade com outras PR já depositadas no *Protein Data Bank* – PDB. Os modelos tridimensionais previstos foram obtidos usando o servidor SWISS-MODEL e o programa *Swiss-Pdb Viewer* v.3.7. A visualização das estruturas sobrepostas das proteínas RcPRs com as depositadas no PDB, foi realizada no programa PyMOL v. 2.3.3. A qualidade estereoquímica das proteínas RcPRs foi calculada pelo *Procheck* versão 3.4 e no servidor *Protein Structure Analysis - ProSa*, pelos índices: *escore-Z*; *ERRAT* e *VERIFY3D*. O grau de identidade variou de 49,37 a 76,23%, com uma semelhança de 44 a 56% e uma cobertura de 60 a 99%. A qualidade estereoquímica de acordo com o Diagrama de *Ramachandran* mais de 90% dos resíduos de aminoácidos das cinco proteínas PR encontram-se nas regiões favorecidas. Os demais índices de avaliação da qualidade dos modelos tridimensionais – 3D, que levam em consideração a interação entre as moléculas gerados de forma *in silico*, permite afirmar que são dignos de confiança. Estes resultados fornecem informações importantes para futuros estudos de interação molecular das proteínas com possíveis fármacos-receptor/sítio ativo. Pois possibilitam o melhor direcionamento nos processos de acoplamento molecular com receptores alvo de forma a auxiliar no controle de a moléstias como o mofo-cinzento na cultura da mamona.

Palavras-Chaves: Análise *in silico*, Modelo tridimensional, Qualidade estereoquímica

Resolution of the protein structure of Pathogenesis-Related Proteins (PRs) via homology modeling in *Ricinus communis* L.

ABSTRACT: Castor bean (*Ricinus communis* L.) is an important oilseed where mammonoculture is harmed due to the attack of diseases and pests. Therefore, our work aimed to perform the computational construction of probable models of five Pathogenesis-Related Proteins (PR Proteins), named in this work as RcPR-1c; RcPR-4b; RcPR-5a; RcPR-6 and RcPR10-STH21. The existing annotation in PHYTOZOME V 12.1 of the genome of *R. communis* allowed the identification of 20 genes and their respective proteins classified in five different families: PR-1; PR-4; PR-5; PR-6 and PR-10 STH21. From this, a PR protein was chosen from each gene family that showed good identity with other PRs already deposited in the Protein Data Bank – PDB. The predicted three-dimensional models were obtained using the SWISS-MODEL server and the Swiss-Pdb Viewer v.3.7 program. The visualization of the overlapping structures of the RcPRs proteins with those deposited in the PDB was performed using the PyMOL v. 2.3.3. The stereochemical quality of the RcPRs proteins was calculated by Procheck version 3.4 and in the Protein Structure Analysis - ProSa server, using the indices: Z score; ERRAT and VERIFY3D. The degree of identity ranged from 49.37 to 76.23%, with a similarity of 44 to 56% and a coverage of 60 to 99%. Stereochemical quality according to the Ramachandran Diagram more than 90% of the amino acid residues of the five PR proteins are found in the favored regions. The other indices for evaluating the quality of the three-dimensional models – 3D, which take into account the interaction between the molecules generated in silico, allow us to say that they are trustworthy. These results provide important information for future studies of the molecular interaction of proteins with possible drugs-receptor/active site. Because they enable better targeting of molecular coupling processes with target receptors in order to help control diseases such as gray mold in castor bean culture.

Keywords: Three-dimensional model, Steriochemical quality, in silico analysis

INTRODUÇÃO

A imunidade das plantas é atribuída a diferentes camadas de mecanismos de defesa que englobam a Resposta de Hipersensibilidade - RH que leva à produção de Espécies Reativas de Oxigênio - ERO e síntese de Proteínas Relacionadas à Patogênese - Proteínas PR que funcionam como elementos antimicrobianos. Essas proteínas PR atuam na Resistência Adquirida Sistêmica - RAS, aumentando a morte celular programada que reduz o crescimento de patógenos, reduzindo assim a taxa de infecção (CHASSOT et al., 2007; VAN BAARLEN et al., 2007).

Estas proteínas respondem a diferentes tipos de infecções causadas por (bactérias, fungos, vírus e nematóides), além de vários insetos. No total, 19 famílias de proteínas PR foram identificadas em plantas mono e dicotiledôneas. Sabe-se que as proteínas PR foram identificadas primeiramente no genoma de *Nicotiana tabaccum* (tabaco) infectadas pelo vírus do mosaico do tabaco. Desde então, essa família de proteínas tem sido exaustivamente estudada no Reino *Plantae*, em diferentes culturas, à exemplo da *Manihot esculenta* Crantz (mandioca); *Solanum tuberosum* (batata); *Allium sativum* L. (alho) e no *Glycine max* (soja) (IRIGOYEN et al., 2020; ARORA et al., 2020; ZAYNAB et al., 2021; ANISIMOVA et al., 2021; ALMEIDA-SILVA et al., 2022).

Em *Ricinus communis* L. (Mamona), esta classe de proteína será caracterizada pela primeira vez. Apesar de ser uma cultura de grande importância econômica, tendo o seu óleo como matéria prima para a produção de mais de 400 coprodutos, os estudos com as proteínas PR ainda são ínfimos. Assim, esforços tem se mantido através do Programa de Melhoramento Genético que visa investir em pesquisas que agregam o uso da Bioinformática, como ferramenta útil no auxílio do melhoramento genético da espécie.

As linhagens que têm sido conservadas e estudadas no Núcleo de Biotecnologia Melhoramento Genético – NBIO da Universidade Federal do Recôncavo da Bahia – UFRB tem sido caracterizada de forma morfoagronômicas e molecular. Os estudos *in silico* auxiliam positivamente no melhoramento por ser um tipo de experimentação que utiliza a simulação computacional com rapidez na execução e de baixo custo computacional, devolvendo resultados confiáveis e dignos de aceitação. Contudo, este tipo de abordagem não visa

substituir os experimentos *in vivo*, tradicionais, e sim ser uma ferramenta a mais que se deve ser explorada, principalmente pelo baixo a zero custo financeiro envolvido.

Assim, nosso trabalho tem por objetivo realizar a modelagem por homologia de cinco proteínas-PR, de diferentes famílias gênicas (PR-1, PR-4, PR-5, RcPR-6 e RcPR10-STH21), anteriormente caracterizadas em nível de gene e proteína, identificadas no genoma de *Ricinus communis* L. e por meio disso, descrever de forma detalhada as características de cada proteína.

MATERIAL E MÉTODOS

Identificação das proteínas-PR no genoma de *Ricinus communis* L.

A anotação existente no Banco de Dados PHYTOZOME V 12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>) do genoma de *Ricinus communis* permitiu a identificação de 20 genes e suas respectivas proteínas classificados em cinco diferentes famílias: PR-1; PR-4; PR-5; PR-6 e PR-10 STH21.

Modelagem molecular

Para escolha da proteína a ser modelada, as 20 sequências de aminoácidos foram alinhadas individualmente com o Banco de Dados de Proteínas – PDB (do inglês *Protein Data Bank*), usando a Ferramenta Básica de Pesquisa em Alinhamento Local de Proteínas – BlastP (<https://blast.ncbi.nlm.nih.gov/>).

Dentre as 20 RcPRs caracterizadas neste trabalho, foi escolhida uma RcPR de cada família identificadas neste trabalho (RcPR-1c; RcPR-4b; RcPR-5a; RcPR-6 e RcPR10-STH21). A escolha foi baseada no maior grau de homologia com proteínas já depositadas no *Protein Data Bank* - PDB. A RcPR-1c com 63,70% de identidade com a Proteína Relacionada a Patogênese P14A, depositada no PDB sob o código 1CFE; a RcPR-4b com 76,23% de identidade com a Proteína de Domínio *Barwin* em sementes de cevada depositada no PDB sob o código 1BW3, a RcPR-5a com 53,67% de identidade com a Proteína semelhante à Taumatina, isolada de cerejas maduras, depositada no PDB sob o código 2AHN, a RcPR-6 com 67,41% também com o código 1CFE e a RcPR10-

STH21 que apresentou 49,37% de identidade com a PR-10 de alérgeno de pólen de bétula Bet v 1, depositada no PDB sob o código 6ALK.

Os modelos tridimensionais previstos foram obtidos usando o servidor SWISS-MODEL (<https://swissmodel.expasy.org>) e o programa Swiss-PdbViewer v.3.7 (GUEx et al., 2009). A visualização das estruturas sobrepostas das proteínas RcPRs com as depositadas no PDB, foi realizada no programa PyMOL v. 2.3.3 (DELANO, 2002). A qualidade estereoquímica das proteínas RcPRs foi calculada pelo Procheck versão 3.4 (LASKOWSKI et al., 1993) e no servidor de Análise de Estrutura de Proteínas – ProSa (do inglês *Protein Structure Analysis*), pelos índices: escore-Z (SIPPL, 1993); ERRAT e VERIFY3D, disponíveis no (<https://prosa.services.came.sbg.ac.at/prosa.php>).

Predição de Sítios Ativos

No servidor *Protein Allosteric and Regulatory Sites* (<http://bioinf.uab.cat/cgi-bin/pars-cgi/mainUploadPdb.pl?jobClass=pars>), foi previsto o sítio das cinco proteínas RcPR modeladas (RcPR-1c; RcPR-4b; RcPR-5a; RcPR-6 e RcPR10-STH21). Inicialmente, foram preditas cavidades com possíveis resíduos de aminoácidos que potencialmente compõem os sítios de ligação. A partir disso, foi considerada a cavidade de cada RcPR que apresentou valores acima de 50% de conservação da estrutura proteica, pois quanto maior este valor, maior são as chances desta cavidade apresentar os resíduos que compõe o potencial sítios ativo.

RESULTADOS

Modelagem Molecular das RcPRs

No servidor *SWISS-MODEL* (<https://swissmodel.expasy.org>), os cinco sequências-alvos foram alinhados com outras sequências proteicas contidas no PDB, através da ferramenta *BLAST*, dessa forma, foram obtidos os modelos tridimensionais das cinco RcPRs (Figura 1).

O grau de identidade entre as sequências-alvo e modelo foi de 49,37 a 76,23%, com uma semelhança de 44 a 56% e uma cobertura de 60 a 99%. Os resultados destes parâmetros permitiram escolher os modelos mais adequados

para a determinação das estruturas proteicas tridimensionais - 3D das cinco RcPRs (Tabela 1).

Tabela 1. Índice de Qualidade de homologia das Proteínas com estruturas tridimensionais depositadas no PDB com as Proteínas RcPR.

PDB	Proteínas	RcPRs	Identidade (%)	Semelhança (%)	Cobertura (%)
1CFE	Proteína Relacionada a Patogênese P14A	RcPR-1c	63,7	52	84
1BW3	Proteína de Domínio <i>Barwin</i> em sementes de cevada	RcPR-4b	76,23	56	60
2AHN	Proteína semelhante à Taumatina, de cerejas maduras	RcPR-5a	53,67	47	85
1CFE	Proteína Relacionada a Patogênese P14A	RcPR-6	67,41	53	82
6ALK	PR-10 de alérgeno de pólen de bétula Bet v 1	RcPR10-STH21	49,37	44	99

A análise de validação via diagrama de Ramachandran (Tabela 1 e Figura Suplementar 1) do modelo das cinco RcPRs mostrou que mais de 94% dos resíduos de aminoácidos estavam nas regiões mais favorecidas e nas regiões permitidas adicionais, que são consideradas aceitáveis, com exceção de 2% dos resíduos, que se encontraram em regiões com provável incompatibilidade estereoquímica.

Ademais, outros indicadores gerais de qualidade da estrutura da proteína foram utilizados. Os valores do escore-Z, abaixo de -5,00, mostram que as cinco proteínas RcPRs modeladas apresentam uma excelente qualidade do modelo geral e se encontram dentro da faixa de pontuações normalmente encontradas para proteínas nativas de tamanho semelhante (Tabela 2 e Figura Suplementar 2). Já os gráficos de energia, mostram a qualidade do modelo local, plotando valores de energia em função da posição da sequência dos aminoácidos, de forma geral, os valores negativos correspondem também a uma boa qualidade de estrutura (Tabela 2 e Figura Suplementar 2).

Além desses índices, pela análise do gráfico ERRAT, mais de 74% dos resíduos encontram-se dentro da região aceitável ($50 <$) (Tabela 2 e Figura Suplementar 3). E pela análise do Verify3D, observa-se que mais de 82% dos resíduos apresentam compatibilidade de modelo atômico (3D) com sua própria sequência de aminoácido (1D), o que indica um ambiente aceitável da cadeia lateral, com $3D-1D \geq 0,2$ (Tabela 2 e Figura Suplementar 4).

Tabela 2. Análise da Qualidade Estereoquímica das cinco proteínas RcPRs modeladas por homologia. Regiões A, B, L mais favorecidas; Regiões a, b, l, p favorecidas.

RcPRs	Diagrama de Ramachandran (%)				Escore-Z	ERRAT (%)	Verify3D (%)	Valores de Energia
	Regiões A, B, L	Regiões a, b, l, p	Somatório das Regiões	Regiões Não Favorecidas				
RcPR-1c	79,5	15,4	94,9	3,4	-5,57	85,16	94,89	0 a -3
RcPR-4b	79,6	17,5	97,1	2,9	-5,3	78,22	99,18	0 a -2
RcPR-5a	86,9	13,1	100	0,0	-6,75	88,83	97,27	0 a -3
RcPR-6	80,8	15,8	96,6	2,5	-5,49	74,42	99,28	0 a -3
RcPR10-STH21	85,4	11,7	97,1	1,5	-6,14	78	82,28	1 a -2

A estrutura tridimensional das cinco proteínas RcPRs, ainda mostrou a presença das β -folhas e das α -hélices (Figura 2). Além disso, permitiu identificar os resíduos de cisteínas que formam as pontes dissulfeto (Figura3).

Figura 1. Modelagem por Homologia das cinco RcPRs. **A)** Estruturas sobrepostas da RcPR-1c de cor azul e da 1CFE de cor rosa. **B)** Estruturas sobrepostas da RcPR-4b de cor dourada e da 1BW3 de cor azul. **C)** Estruturas sobrepostas da RcPR-5a de cor amarela e a 2AHN de cor dourada. **D)** Estruturas sobrepostas da RcPR-6 de cor rosa e a 1CFE de cor verde. **E)** Estruturas sobrepostas da RcPR10-STH21 de cor verde e a 6ALK de cor dourada. Estas imagens foram geradas no programa no PyMOL.

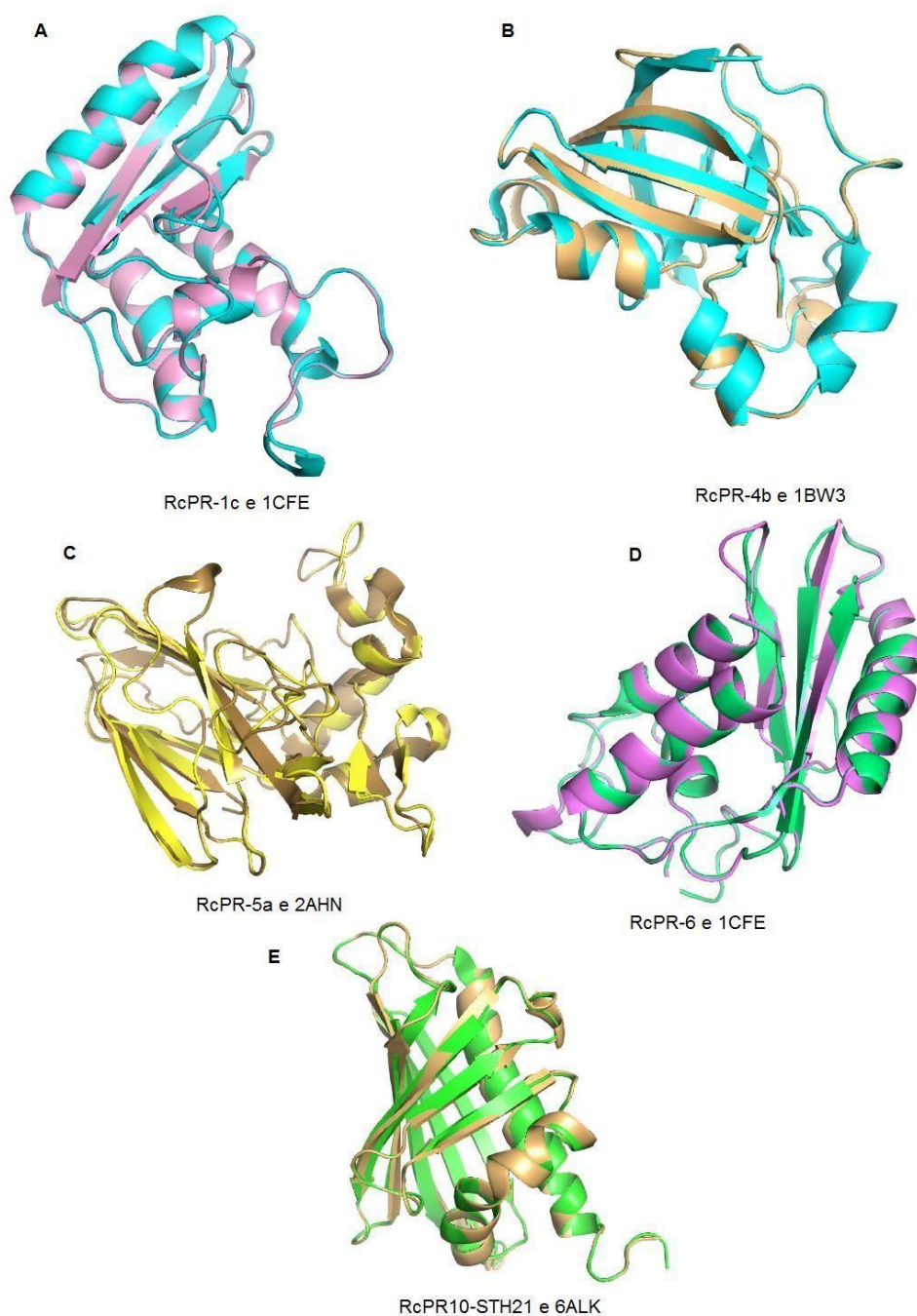


Figura 2. Estrutura 3D das cinco RcPRs. **A)** Estrutura da RcPR-1c com três β -folhas e cinco α -hélices. **B)** Estrutura da RcPR-4b com quatro β -folhas e α -hélices. **C)** Estrutura da RcPR-5a com 12 β -folhas e cinco α -hélices. **D)** Estrutura RcPR-6 com três β -folhas e cinco α -hélices. **E)** Estrutura da RcPR10-STH21 com sete β -folhas e três α -hélices. Todas estas estruturas foram visualizadas no programa PyMol.

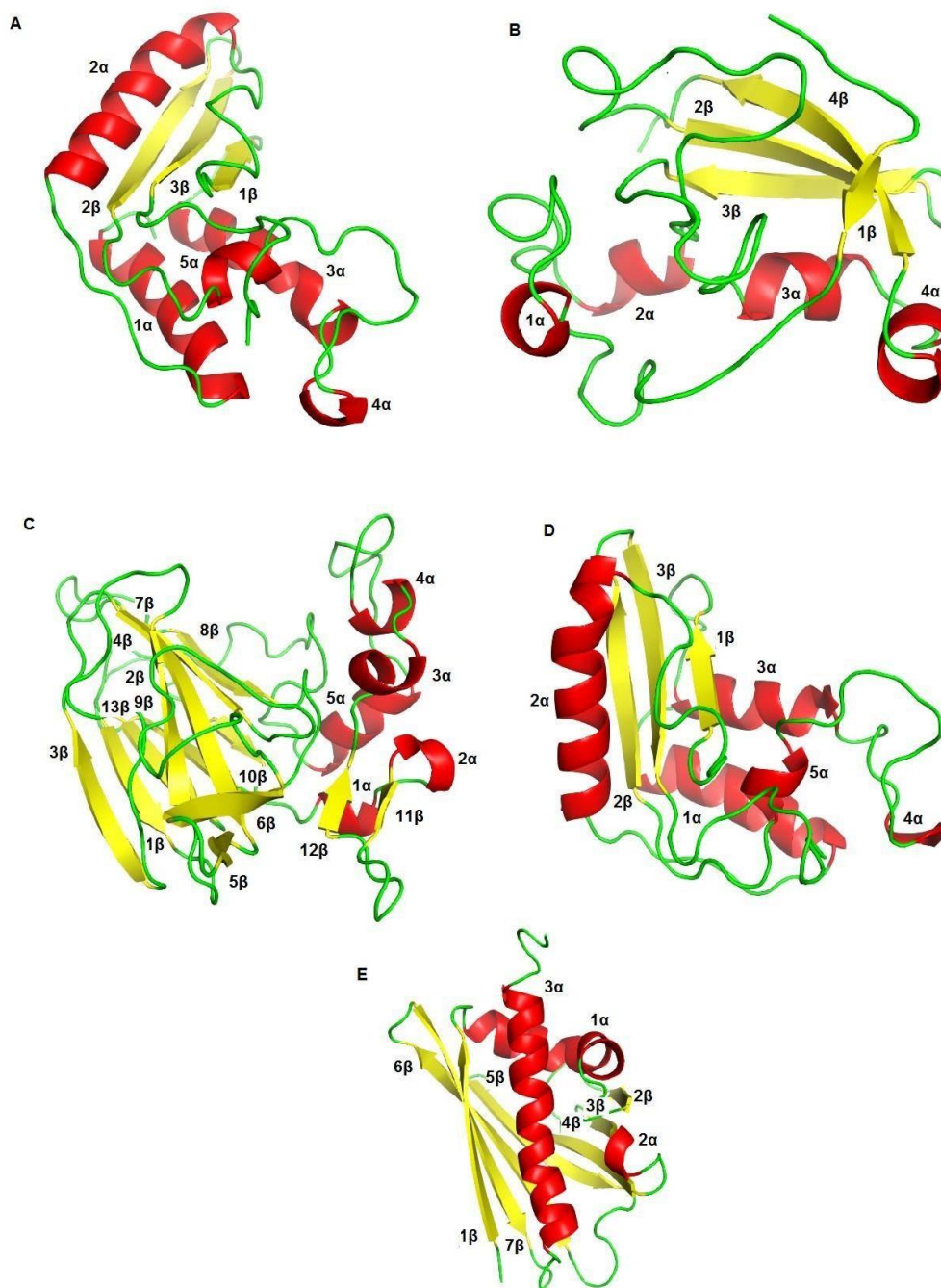
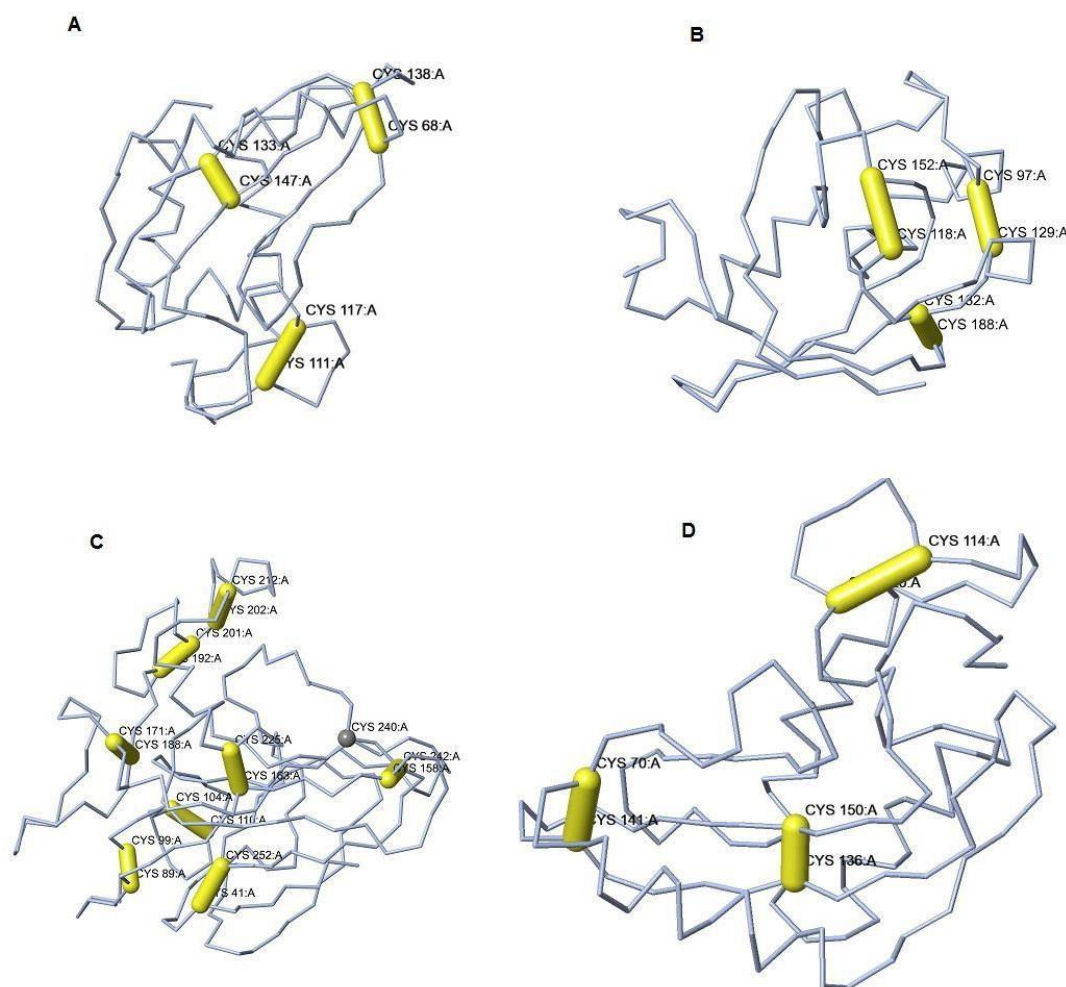


Figura 3. A estrutura 3D de quatro das cinco proteínas RcPRs modeladas que apresentaram resíduos de cisteínas, formando as pontes dissulfeto de cor amarela. **A, B e D)** Estrutura da RcPR-1c; RcPR-4b e RcPR-6 com seis resíduos de cisteína, formando três pontes dissulfeto, cada. **C)** Estrutura da RcPR-5a com 17 cisteínas formando oito pontes dissulfeto. Estas visualizações foram realizadas no servidor Jmol (firstglance.jmol.org).

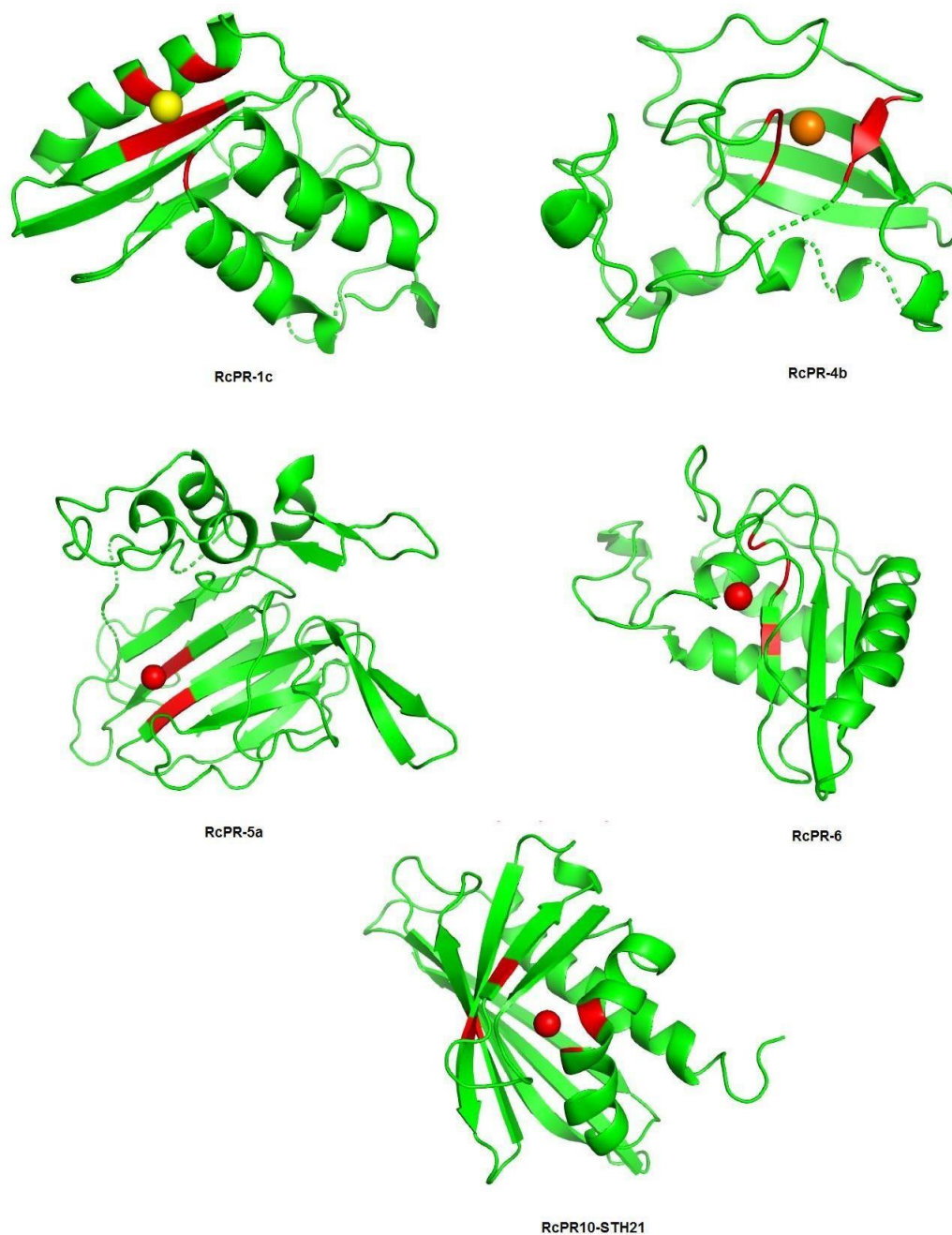


Predição dos sítios ativos

Foi previsto um sítio ativo para cada uma proteína RcPR que teve a resolução da sua estrutura tridimensional. Os resíduos de aminoácidos que compõem o sítio ativo da proteína RcPR-1c são a ALA 54, GLN 58, CYS 133, ALA 134 e a LYS 135. Na proteína RcPR-4b são a VAL 74, ARG 75, ALA 76, GLY 155 e 156; e a LEU 157. Na proteína RcPR-5a são a SER 80, GLY 81 e o GLU 122. Na proteína RcPR-6 são a TYR 79, GLU 81 e a LEU 83. Na proteína

RcPR10-STH21 são a ILE 24, SER 27, GLU 55, ILE 72 e a TYR 84 (Figura 4).

Figura 4. Predição do sítio ativo das cinco proteínas RcPR. Cada RcPR apresenta sua cavidade composta de resíduos de aminoácidos que fazem parte dos sítios ativos preditos de cor vermelha.



DISCUSSÃO

A modelagem molecular permitiu confirmar que as proteínas modeladas, RcPR-1c, RcPR-4b, RcPR-5a, RcPR-6 e RcPR10-STH21 são altamente homólogas com as proteínas com estruturas já depositadas no PDB. Estas cinco RcPRs modeladas, compartilham o mesmo domínio estrutural dos seus respectivos modelos. Com isto, acredita-se que ambas apresentam funções semelhantes no contra-ataque a agentes fitopatogênicos. Além delas desencadearem RH e RAS em plantas contra infecções por agentes fitopatogênicos como já descrito outrora.

A presença das β -folhas e das α -hélices nas cinco RcPRs são de suma importância, pois contribuem para estabilidade da molécula, sendo as β -folhas formadas por uma estrutura rígida e as α -hélices conformações de interação entre o O da carboxila (C=O) de um resíduo de aminoácido com o grupo amino (-NH) de um outro resíduo que está a quatro posições depois na cadeia polipeptídica por meio das ligações de hidrogênio (H). Este padrão de ligação leva a estrutura helicoidal, isto é, volta entre as hélices, onde o agrupamento *R* de cada aminoácido se fixa na parte externa da alfa hélice, onde ficam livres para interagir (NELSON, 2014).

Em geral, a função de uma proteína é especificada por sua estrutura tridimensional (ALBERTS, 2002), assim, acredita-se que a resolução da estrutura tridimensional das cinco RcPRs, permitirá uma exploração mais completa da função específica das PRs no sistema imune de plantas de *Ricinus communis* L., bem como, no auxílio de estudos *in vitro* e *in vivo* para obtenção do conhecimento mais completo sobre esta família.

Ademais, a predição dos resíduos de aminoácidos conservados que compõem possivelmente os sítios ativos de cada uma das cinco RcPRs modeladas, permite sugerir possíveis acoplamentos moleculares com moléculas que possam atuar como um recurso de reconhecimento de fitopatógenos, e a possível ligação sugestiva de uma RcPR com ligantes, pode levar a inibição do crescimento de fitopatógenos, causando toxicidade, por exemplo, a componentes da parede celular ou membrana plasmática de fitopatógenos oportunistas, resultando na morte de suas células (NUSRAT et al., 2017; TELLIS et al., 2017).

Com base nisso, especificamente, espera-se um possível acoplamento entre a RcPR-4b e prováveis ligantes de reconhecimento ao *Amphobotrys ricini* B., causador do mofo-cinzento em *R. communis*, considerada umas das doenças de maior importância econômica na espécie (YAMUNA et al., 2015), que tenha como alvo, por exemplo, componentes da matriz extracelular do fitopatógeno, que podem ocasionar a morte celular desse fungo, contribuindo desta forma para que o *A. ricini* não tenha êxito no estabelecimento da relação parasitária com a espécie.

Perspectivas Futuras

Neste trabalho, a caracterização realizada dos genes e sequências proteicas das proteínas-PR na espécie de *Ricinus communis* L., que participam de respostas ambientais, de desenvolvimento na planta e no combate a fitopatógenos, agora pode ser examinado em estudos experimentais sobre diversas abordagens biotecnológicas ou por meio de marcadores de resistência.

Embora *in silico*, os dados apresentados neste trabalho representam um ponto de partida para elucidar as complexas respostas das plantas em *R. communis*, a estresses tanto de origem biótica quanto abiótica. As cinco famílias das proteínas-PR aqui analisadas podem ser úteis para verificar mecanismos de imunidade inata na presente espécie, a fim de investigar se há resistência induzida/não hospedeira ou hospedeira a fitopatógenos de importância econômica na cultura que ocorre com a participação destas proteínas.

Atualmente, o gerenciamento de doenças causadas por fitopatógenos, principalmente os fúngicos, em *R. communis* é alcançado principalmente por aplicações de produtos químicos com elevados custos anualmente, alguns dos quais provavelmente serão proibidos a sua aplicação em um futuro próximo. Além disso, devido à rápida evolução da população de fitopatógenos, cepas resistentes podem ser selecionadas também (GOODWIN et al., 1996).

Com o constante crescimento populacional, o efeito das mudanças climáticas, o movimento de material contaminado, a prevalência de monoculturas e o aumento da área total cultivada, a evolução dos fitopatógenos é cada vez mais observada entre as espécies, dessa maneira, variedades e/ou cultivares comerciais tornam-se cada vez mais susceptíveis a doenças.

O presente estudo contribuirá mais especificadamente nas pesquisas que serão desenvolvidas, *in vivo*, pelo grupo de pesquisa Núcleo de Melhoramento Genético e Biotecnológico - NBIO, que visa identificar moléculas que fazem parte do mecanismo de defesa em genótipos descritos como susceptíveis ao fungo *Amphobotrys ricini* B., causador da doença mofo-cinzeno. Dentre as moléculas, espera-se identificar níveis de expressão gênica de precursores das proteínas-PR. Esta possível identificação abrirá novas possibilidades para o desenvolvimento de genótipos resistentes ao *A. ricini*.

Em geral, estas pesquisas proporcionarão ao Melhoramento Genético da Mamona avanços significativos nos trabalhos que envolvem resistência não só ao mofo-cinzeno como também a outras doenças de importância econômica na espécie.

Entende-se que este conhecimento abrangente, tanto da variação proteica quanto da sua caracterização estrutural, agrega informações valiosas que podem melhorar as estratégias de manipulação dessas moléculas, permitindo compreender o papel e sistema regulador molecular no mecanismo imune de *R. communis*. Além de gerar subsídios para o avanço no estudo sobre os padrões de expressão destas famílias gênicas em diferentes tecidos na espécie e assim, investigar melhor as funções biológicas dessa classe de proteínas e até sugerir o envolvimento com outras moléculas também relacionadas aos processos de resposta ao reconhecimento de infecção por fitopatógenos.

CONCLUSÃO

As diferentes abordagens de verificação da estrutura modelada por homologia permitiram a construção de cinco modelos tridimensionais confiáveis das proteínas RcPR-1c, RcPR-4b, RcPR-5a, RcPR-6 e RcPR10-STH21 que auxiliará na exploração mais completa da função específica das proteínas-PR no sistema imune de plantas de *Ricinus communis* L.

Estes dados contribuirão para o desenvolvimento de novas estratégias que visam aumentar a resistência na espécie de *R. communis* contra agressores.

AGRADECIMENTOS

Agradecemos a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudo que permitiu o desenvolvimento desta pesquisa.

REFERÊNCIAS

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**, 4th edition. Garland Science, New York 2002.

ALMEIDA-SILVA, F.; VENANCIO, T. M. Pathogenesis-related protein 1 (PR-1) genes in soybean: Genome-wide identification, structural analysis and expression profiling under multiple biotic and abiotic stresses. **Gene**, v. 809, p. 1-9, 2022.

ANISIMOVA, O. K.; SHCHENNIKOVA, A. V.; KOCHIEVA, E. Z.; FILYUSHIN, M. A. Pathogenesis-Related Genes of PR1, PR2, PR4, and PR5 Families Are Involved in the Response to Fusarium Infection in Garlic (*Allium sativum* L.). **International Journal of Molecular Sciences**, v. 22, p. 1-20, 2021.

ARORA, R.; KUMAR, A.; SINGH, I. K.; SINGH, A. Pathogenesis Related Proteins: a defensin for plants but an allergen for humans. **International Journal of Biological Macromolecules**. v. 157, p. 659-672, 2020.

CHASSOT, C.; NAWRATH, C.; MÉTRAUX, J-P. Cuticular defects lead to full immunity to a major plant pathogen. **The Plant Journal**, v. 49, p. 972-980, 2007.
NELSON, D. L.; COX, M. M. Princípios de bioquímica de Lehninger. 6. ed. Porto Alegre: Artmed, 2014.

NUSRAT, Y.; SALEEM, M.; NAS, M.; GUL, R.; REHMAN, H. M. Molecular Characterization, Structural Modeling, and Evaluation of Antimicrobial Activity of Basrai Thaumatin-Like Protein against Fungal Infection. **BioMed Research International**, v. 2017, p. 1-10, 2017.

TELLIS, M.; MATHUR, M.; GURJAR, G.; KADDOO, N.; GUPTA, V. Identification and functionality prediction of Pathogenesis-Related Proteins 1 from legume family. **Proteins**, v. 85, p. 2066-2080, 2017.

VAN BAARLEN, P.; VAN BELKUM, A.; SUMMERBELL, R. C.; CROUS, P. W.; THOMMA, B. P. H. J. Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? **FEMS Microbiology Reviews**, V. 31, P. 239-277, 2007.

YAMUNA, C.; VARMA, P. K.; PRASAD, R. D.; LAKSHMI, K. V. Morphological and molecular characterization of anamorph associated with gray mold of castor (*Ricinus communis* L.). **Journal of Oilseeds Research**, v. 32, p. 63-67, 2015.

ZAYNAB, M.; PENG, J.; SHARIF, Y.; AL-YAHYAI, R.; JAMIL, A.; HUSSAIN, A.;

KHAN, K. A.; ALOTAIBI, S. S.; LI, S. Expression profiling of pathogenesis-related Protein-1 (PR-1) genes from *Solanum tuberosum* reveals its critical role in *Phytophthora infestans* infection. **Microbial Pathogenesis**, v. 161, p. 1-11, 2021.

ANEXOS

Tabela Suplementar 1: Sequências de aminoácidos das Proteínas Relacionadas a Patogênese em *Ricinus communis* L. Peptídeo sinal: sublinhado; Domínio Funcional Conservado: hachurada em cinza, Sítios de Fosforilação (Serina, Treonina e Tirosina) em azul e os Sítios de N-glicosilação (Asparagina-x-Serina/Treonina, em que x pode ser qualquer aminoácido, exceto a prolina) em vermelho.

RcPR-1c

>30099.m001652

MLSLKISAAALSFLISLSLILSSHAQNLPQDYINAHNAARAQVGVGPMTWDNTVAAYAQN~~YANQRINDCRL~~
 VHSSGRYGENIAWSSGDLSGTA~~AVKLWVDEKAFYNYNSNTCAAGQQCGHYTQVVWRNSVRLGCAKVRCRN~~
 GGTFITCNYDPPGNFVNQRPY

RcPR-4b

>29598.m000459

MGSVVLMVLLCSLGVAFGEQCGRQAGGKLCFNLLCCSYGWCGSTDDYCSPSKNCQSNCKGGGGESS
 NVRATYHFYNPQQHGWDLNAVSAYCSTWDANKPYSWRSKYGWTAFCGPAGPRGQTSCGKCLRVTNTRTGA
 ATTVRIVDQCSNGGLDLDVNVFRQLDTDGIGYQQGHLIVNYQFVNCGDSFNPLFSIMDDQ

RcPR-5a

>30170.m013991

MHYKLIKILAIIVCGTNLYFCVKFAGNFSNATVFTLQNHCSYMIWPGTLSGNGAGTLDGGFALPPGSS
 IQLQAPPTWSGRFWARTGCTFFDDLGNKCDTGDCCGALKCTGGGAPPVSLAEFTIAPSSNEKDFYDVSLV
 DGINVGLGKALGSSGDCQYAGCVADLNGNCPPELVRTDSVSSVVACKSACAAAFNAPEFCCTGDHSTPQ
TCSPTQYSALFKNACPSAYSYAYDDASSTCTCSGSDYMITFCPTGSG

RcPR-6

>30099.m001658

MGSLKISQVLVSLLMGLALVVHPCHAQNSQODYLDAHNAARAQVGVANIVWDNTVATYAQNYANSRIGDC
 NLVHSSGSYGENLAKGSSSSLTGTAAVNLWVAEKPYNYTSNSCTGGQQCLHYTQVVWSKSVRLGCARVQ
 CTNGWWFVTCNYDPPGNYIGQKPY

RcPR10-STH21

>29726.m004110

MGVSTFTQEFTSPVAPSRMFKALILDSNLIIPKLLPQFIKSVDVIQGGDAGTIEQVNFTEASDFKYIKN
 RIEELDKDNLVCKYTLIEGGPLGDKLDSIAYEVKFEAASDGGCICKITSNYFTLGKFTIKEEDIEDGKG
 ALGIYRVVEAYLLENPQVYA

Figura Suplementar 1. Gráfico de Ramachandran. Verificador de qualidade estereoquímica da estrutura proteica.

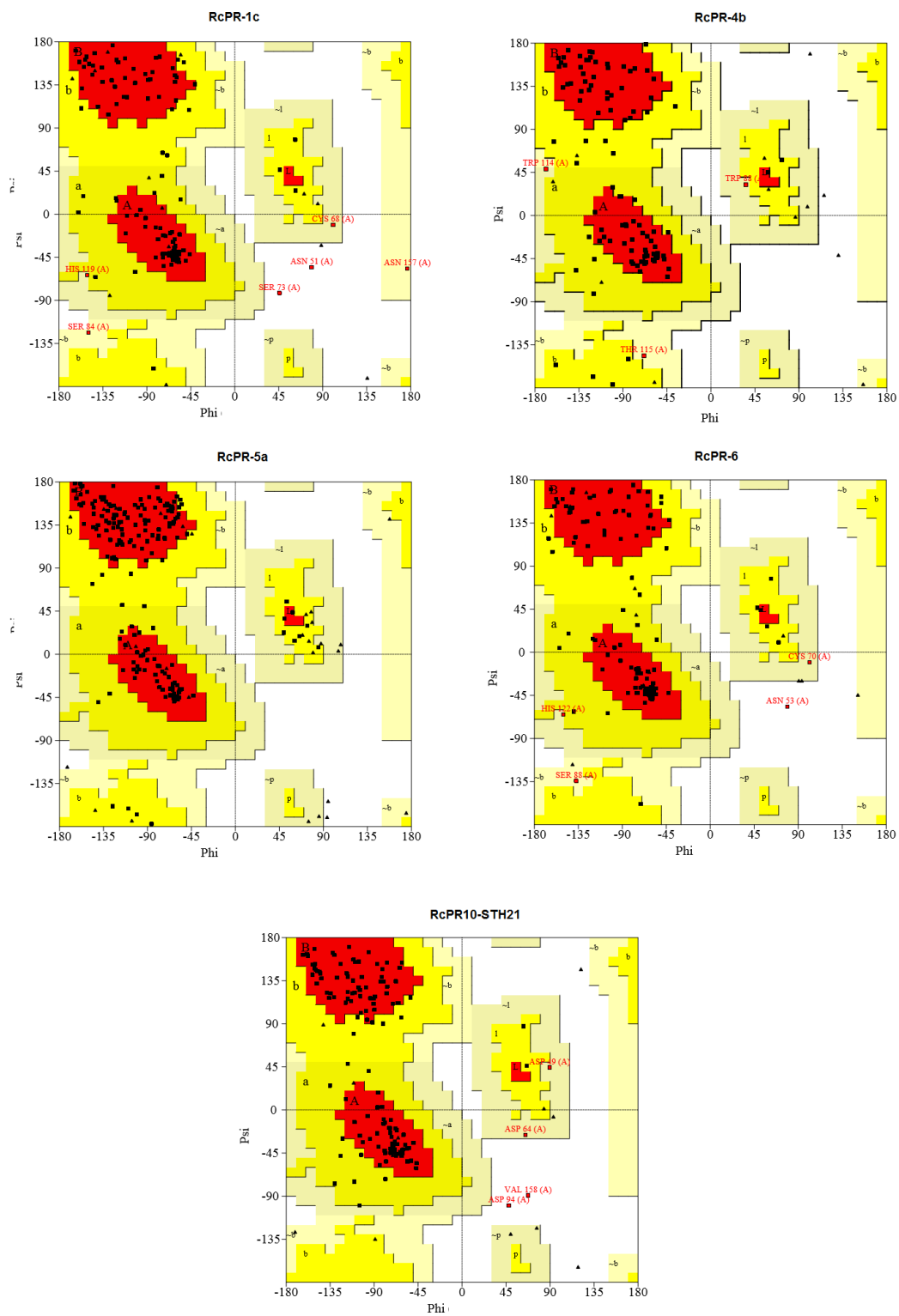


Figura Suplementar 2. Fator de qualidade *Score-Z*. Verificador de qualidade estereoquímica da estrutura proteica.

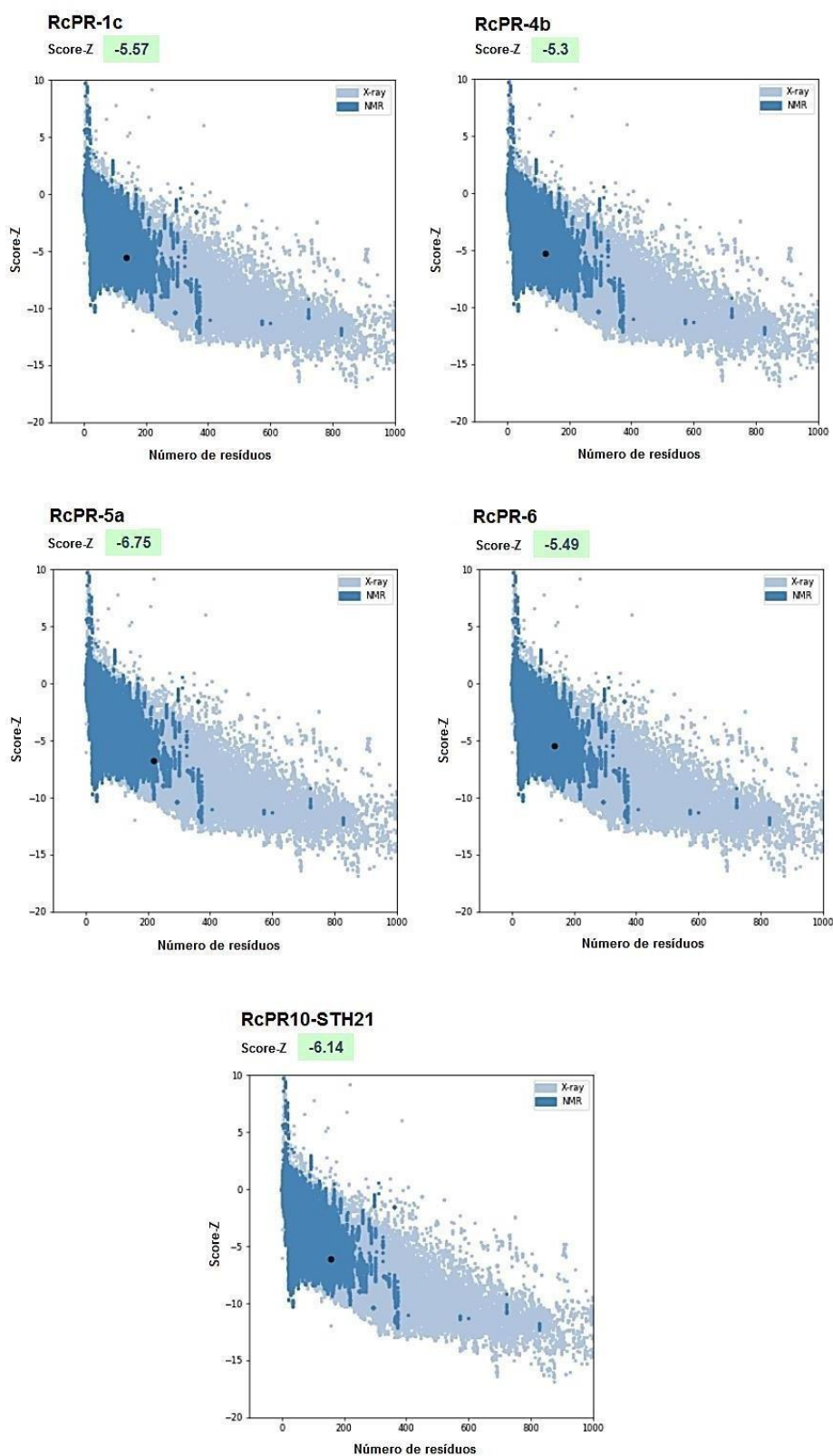


Figura Suplementar 3. Fator de qualidade ERRAT. Verificador de qualidade estereoquímica da estrutura proteica.

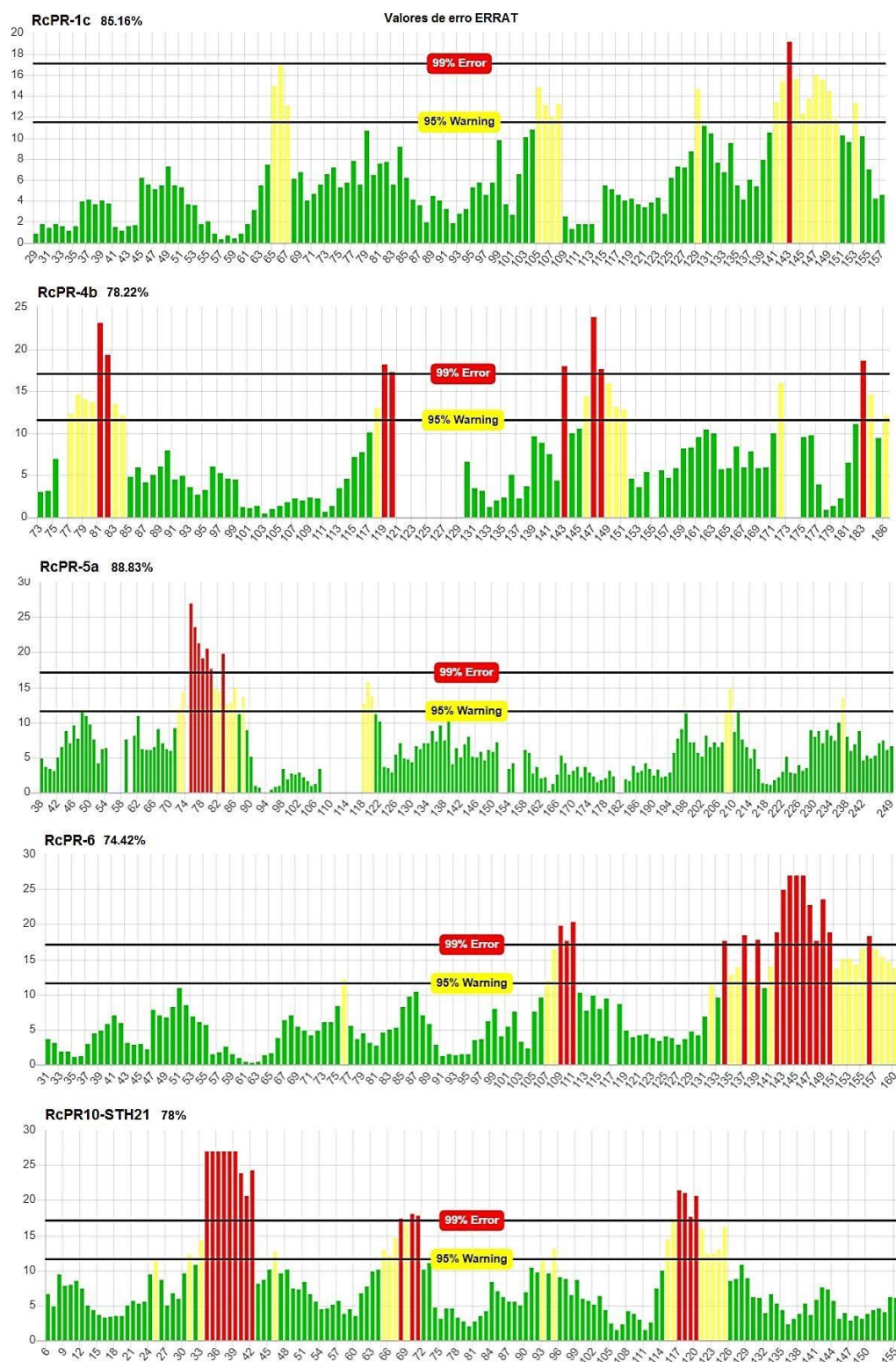


Figura Suplementar 4. Verify 3D. Verificador de qualidade estereoquímica da estrutura proteica.

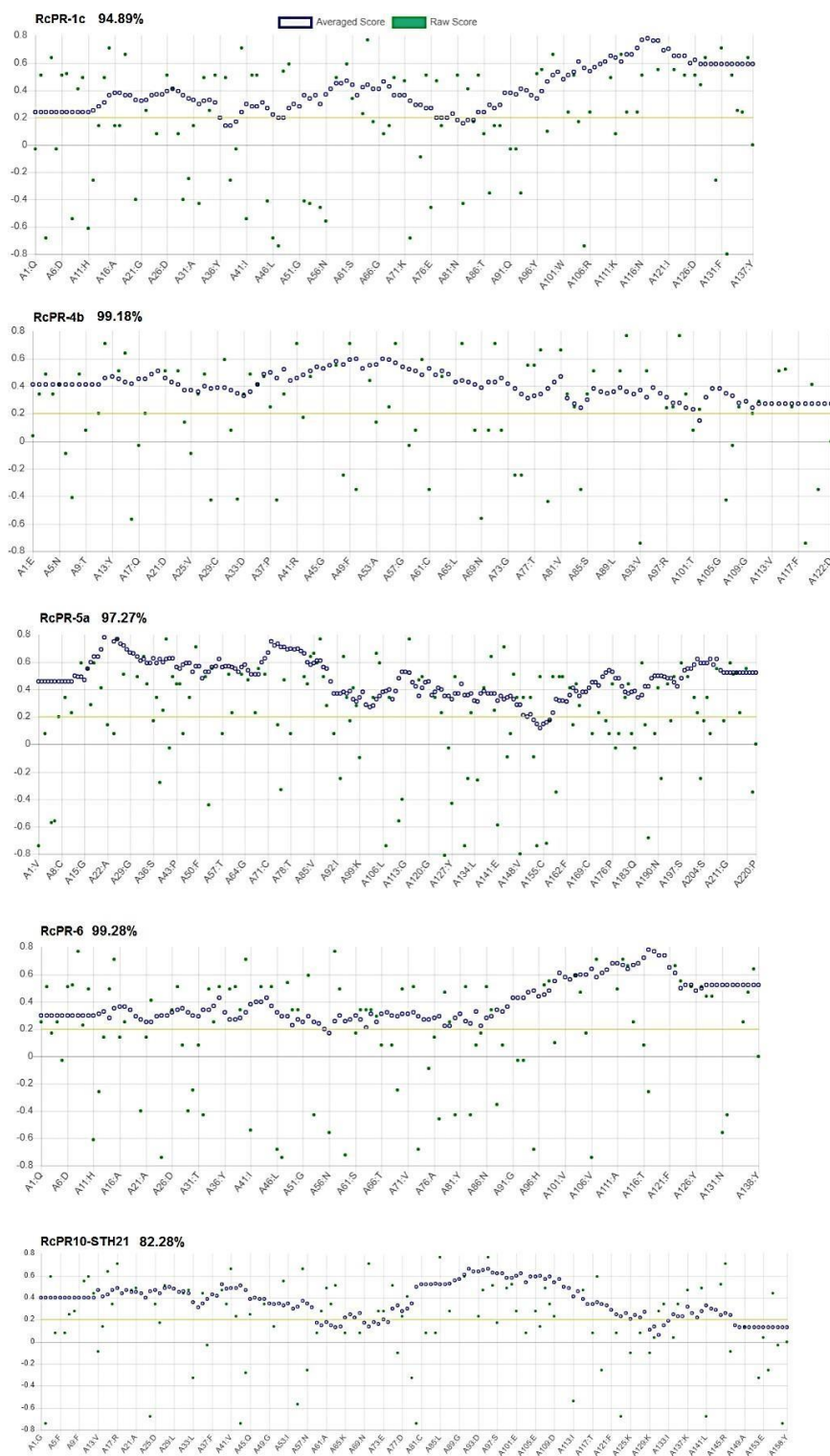
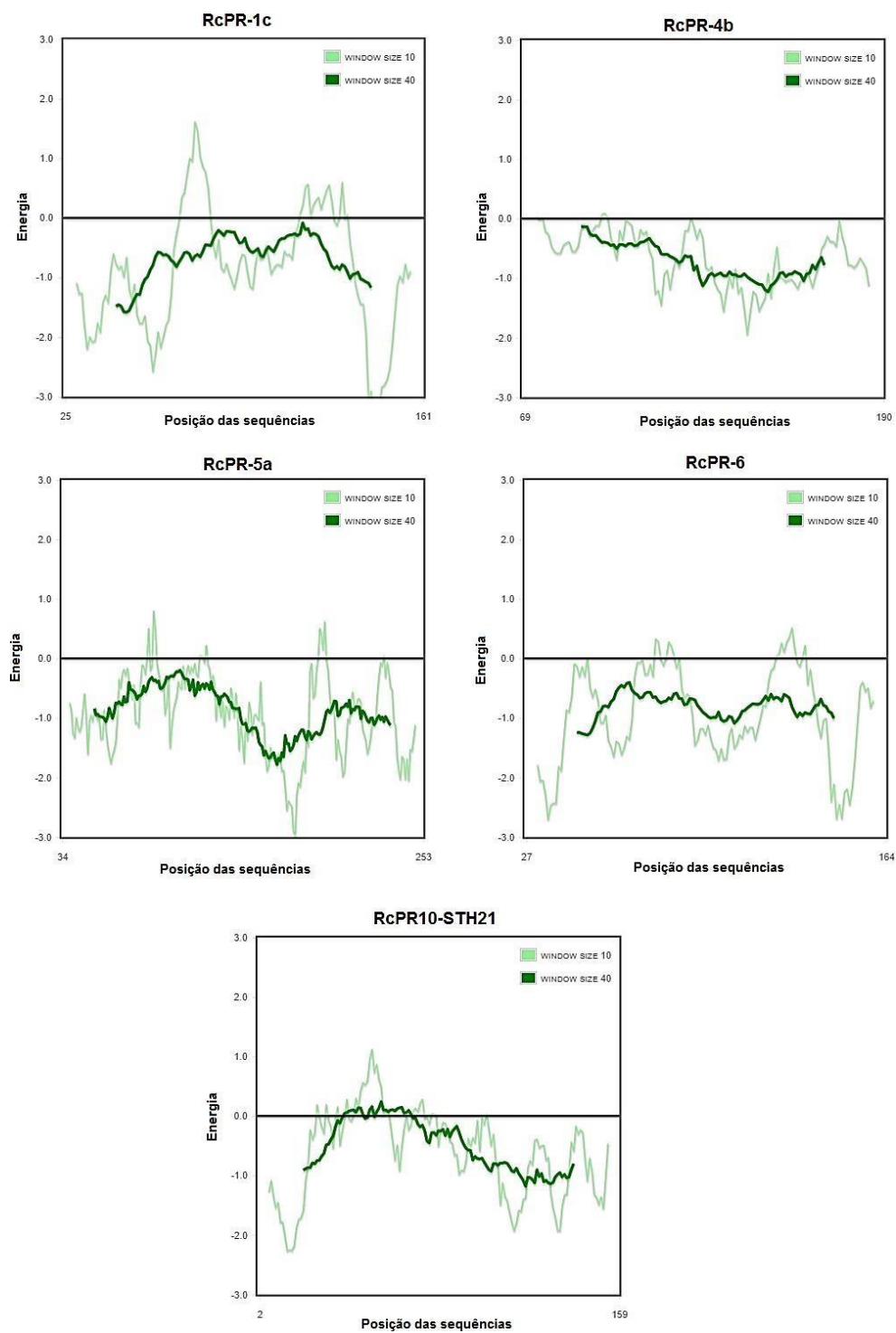


Figura Suplementar 5. Gráfico de energia.



CONSIDERAÇÕES FINAIS

Este trabalho apresentou resultados relevantes sobre a descrição das importantes famílias gênicas das RcPR, amplamente distribuídas no genoma de *R. communis*, atuantes em respostas a ataque de pragas e doenças, além de contribuir a respostas ligadas a outros tipos de estresses.

Este foi um trabalho pioneiro na cultura e contribuirá para um melhor entendimento do comportamento das Proteínas Relacionadas à Patogeneses no genoma da referida espécie e assim propor, no futuro, a identificação de moléculas que fazem parte do mecanismo de defesa em genótipos descritos como susceptíveis ao fungo *Amphobotrys ricini* B., causador da doença mofo-cinzento.