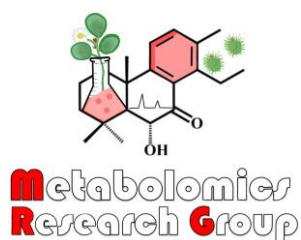




**UNIVERSIDADE DO ESTADO DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA APLICADA**



JONHENES SILVA DE OLIVEIRA

**AQUISIÇÃO DE TOLERÂNCIA À ESTRESSES ABIÓTICOS:
CONTRIBUIÇÃO DOS GENES DA FAMÍLIA TERPENO SINTASE EM
MAMONA (*Ricinus communis. L*)**

**Salvador
2022**

JONHENES SILVA DE OLIVEIRA

**AQUISIÇÃO DE TOLERÂNCIA À ESTRESSES ABIÓTICOS:
CONTRIBUIÇÃO DOS GENES DA FAMÍLIA TERPENO SINTASE EM
MAMONA (*Ricinus communis. L*)**

Dissertação apresentada ao Programa de Pós-Graduação em Química Aplicada, da Universidade do Estado da Bahia para obtenção do título de mestre em Química Aplicada sob orientação do Prof. Dr. Paulo Roberto Ribeiro de Jesus.

Salvador
2022

FICHA CATALOGRÁFICA

Biblioteca Professor **Edivaldo Machado Boaventura - UNEB – Campus I**

O48a Oliveira, Jonhenes Silva de

Aquisição de tolerância à estresses abióticos: contribuição dos genes da família terpeno sintase em mamona / Jonhenes Silva de Oliveira. – Salvador, 2022.

124f. : il.

Orientador: Paulo Roberto Ribeiro de Jesus.

Dissertação (Mestrado Acadêmico). Universidade do Estado da Bahia. Departamento de Ciências Exatas e da Terra. Programa de Pós-

UNIVERSIDADE DO ESTADO DA BAHIA
Autorização Decreto nº 9237/86. DOU 18/07/96. Reconhecimento: Portaria 909/95, DOU 01/08-95

DEPARTAMENTO DE CIÊNCIAS EXATAS E DA TERRA
CAMPUS I - SALVADOR




FOLHA DE APROVAÇÃO
**AQUISIÇÃO DE TOLERÂNCIA À ESTRESSES ABIÓTICOS:
CONTRIBUIÇÃO DOS GENES DA FAMÍLIA TERPENO SINTASE EM
MAMONA (*Ricinus communis*. L)**

JONHENES SILVA DE OLIVEIRA

Dissertação apresentada ao Programa de Pós-Graduação em Química Aplicada- PGQA, em 28 de julho de 2022, como requisito parcial para obtenção do grau de Mestre em Química Aplicada pela Universidade do Estado da Bahia, conforme avaliação da Banca Examinadora:

Professor Dr. PAULO ROBERTO RIBEIRO DE
JESUS
Universidade do Estado da Bahia - UNEB
Orientador(a) Presidente

Professor(a) Dr.(a) MARTA BRUNO LOUREIRO
Universidade Federal da Bahia – UFBA
Examinador(a) Externo

Documento assinado digitalmente
 MARCOS VINICIUS SILVA DE ANDRADE
Data: 28/07/2022 13:30:03 -0300
Verifique em <https://verificador.ti.br>

Professor Dr. MARCOS VINICIUS SILVA DE
ANDRADE
Universidade Federal da Bahia - UFBA
Examinador(a) Externo

JONHENES SILVA DE OLIVEIRA

**AQUISIÇÃO DE TOLERÂNCIA À ESTRESSES ABIÓTICOS:
CONTRIBUIÇÃO DOS GENES DA FAMÍLIA TERPENO SINTASE EM
MAMONA (*Ricinus communis. L*)**

BANCA EXAMINADORA:

Prof. Dr. Paulo Roberto Ribeiro de Jesus
Doutor em Bioquímica e Biologia Molecular de Plantas
Wageningen University (Holanda)
Universidade Federal da Bahia (UFBA, Brasil)

Prof. Dra. Marta Bruno Loureiro - Titular - ICS-UFBA

Pro. Dr. Marcos Vinicius Silva de Andrade

AGRADECIMENTOS

Primeiramente agradeço ao Deus vivo e presente em minha vida. Aos seus anjos bons, às pessoas de bom coração que cruzaram meu caminho, me fortalecendo e ensinando lições valiosas.

Ao meu orientador e amigo, Prof. Dr. Paulo Roberto Ribeiro de Jesus ao qual de muitas formas me ensinou o valor do esforço duro, mesmo entre lágrimas e sorrisos calorosos. Agradeço seu carinho, amparo e orientação.

Aos professores e professoras do Programa de Pós Graduação em Química Aplicada (PGQA – UNEB), em especial a Prof. Dr Lourdes Cardoso pelo carinho e ensinamentos valiosos.

Ao Programa de Pós-Graduação em Química Aplicada (PGQA) pelo suporte e apoio financeiro e ao Departamento de Ciências Exatas e da Terra I, da Universidade do Estado da Bahia pela infraestrutura para a realização do trabalho; Ao LBBB/ICS - UFBA, na presença da Prof^a. Dr^a. Luzimar pelo suporte instrumental e apoio dos colegas.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida;

A todos que contribuíram direta ou indiretamente para a realização desse trabalho.

Àqueles que sempre acreditaram em mim, àqueles que me emprestam os ombros, os braços, os abraços e não poucas vezes o coração.

LISTA DE FIGURAS

Figura 1: Espécie <i>Ricinus communis</i> com destaque em suas variadas anatomias morfológicas e aspectos frutíferos	16
Figura 2: Diversidade de cores e tamanhos das sementes da <i>Ricinus communis</i>	17
Figura 3: Estrutura cristalina da glicoproteína ricina. Destaque em verde para a subunidade catalítica RTA, e em laranja, a subunidade RTB	18
Figura 4: Visão geral da biossíntese dos terpenos em plantas.....	25
Figura 5: Comprimento de (a) hipocótilos e (b) raízes após quatro dias sob temperaturas de 20, 30 e 35 ° C.	47
Figura 6: Massa seca em gramas de (a) raízes (b) hipocótilos, (c) endospermas e (d) cotilédones, após quatro dias sob as temperaturas de 20 °C, 30 °C e 35 °C.	47
Figura 7: (a) Razão entre massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) Razão entre massa seca da raiz / comprimento da raiz (g /mm) e (c) razão entre a massa seca das raízes/ massa seca dos hipocótilos, após quatro dias sob temperaturas 20 ° C, 30 ° C e 35 ° C.	48
Figura 8: Comprimento (mm) dos (a) hipocótilos e (b) raízes, após quatro dias sob concentrações de estresse salino de 0 mM, 25, 50 e 100 mM de NaCl.....	49
Figura 9: Massa seca em gramas das (a) raízes e (b) hipocótilos, (c) endospermas e (d) cotilédones, após quatro dias sob estresse salino nas concentrações de 0 mM, 25, 50, e 100 mM de NaCl.	51
Figura 10: (a) Razão massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) Razão massa seca da raiz / comprimento da raiz (g / mm) após quatro dias de estresse salino sob concentrações de 0 mM, 25, 50 e 100 mM de de NaCl.	52
Figura 11: Comprimento de (a) hipocótilos e (b) raízes após quatro dias sob estresse osmótico sob os potenciais de 0, -0,2, -0,6 e -1,0 Mpa.	53
Figura 12: (a) Razão da massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) razão da massa seca da raiz / comprimento da raiz (g / mm), após quatro dias sob os potenciais -0,2; -0,6 e -1 Mpa.....	54
Figura 13: Massa seca em gramas das (a) raízes e (b) hipocótilos, (c) endospermas e (d) cotilédones após quatro dias sob estresse osmótico nos potenciais de -0,2; -0,6 e -1 MPa	55
Figura 14: Exemplos de terpenos e derivados comumente isolados de plantas, bactérias, organismos marinhos dentre outros.....	65

Figura 15: Rotas bioquímicas com destaque na localização celular onde ocorre a formação dos precursores IPP e DMAPP - blocos básicos de construção de terpenos em plantas.....	67
Figura 16: Relações filogenéticas dos genes TPS em <i>Ricinus communis</i> , <i>Populus Trichocarpa</i> e <i>Arabidopsis thaliana</i>	77
Figura 17: Porcentagem da distribuição das RcTPS em <i>Ricinus communis</i> de acordo com a predição da localização subcelular.	79
Figura 18: Motifs conservados e estrutura dos genes RcTPS a) Árvore filogenética NJ e distribuição de motifs conservados das 46 proteínas RcTPS. b) Disposição de exons e íntrons dos genes RcTPS	79
Figura 19: Análise da região promotora dos genes RcTPS em <i>Ricinus communis</i>	85

LISTA DE TABELAS

Tabela 1: Metabólitos já identificados em <i>Ricinus communis</i> e suas respectivas massas de alta resolução com adutos de Hidrogênio $[M + H]^+$ e Sódio $[M + Na]^+$	27
Tabela 2: Número de espécies TPS em <i>R. communis</i> L. e em outras espécies já relatadas na literatura.....	74

LISTA DE ABREVIATURAS E SIGLAS

UV	Ultra violeta
RNA	Ácido ribonucleico
DNA	Ácido desoxirribonucleico
%	Porcentagem
EBDA	Empresa Baiana de Desenvolvimento Agrícola
EMBRAPA	Empresa Brasileira de Pesquisas Agropecuária
LBBB	Laboratório de Bioquímica, Biotecnologia e Bioprodutos
MDA	Malondialdeído
NaCl	Cloreto de sódio
O ₂	Oxigênio molecular
OH ⁻	Hidroxila
PEG-6000	Polietilenioglicol
RAS	Regras de análise de sementes
SOD	Superóxido dismutase
UFBA	Universidade Federal da Bahia
EMPARN	Empresa de Pesquisa Agropecuária do Rio Grande do Norte
MVA	Ácido Mevalônico
MEP	2-C-Metil-D-eritritol 4-fostato
IPP	Difosfato de isopentenila
DMAPP	Difosfato de dimetilalila
FPP	farnesila

GPP	difosfato de geranila
GGPP	difosfato de geranylgeranila
AtTPS	Terpeno sintase em <i>Arabidopsis Thaliana</i>
RcTPS	Terpeno sintase em <i>Ricinus communis</i>
Na ⁺	Íon Sódio
Cl ⁻	Íon Cloreto
K ⁺	Íon Potássio
Ca ²⁺	Íon Cálcio
HSPs	Proteína de choque térmico

APRESENTAÇÃO

Das dificuldades que permeiam o estudo dos terpenos, destaca-se a complexidade dos genomas envolvidos e as restrições quanto ao estudo das vias biossintéticas para a maioria destes compostos. Portanto, este trabalho gera subsídios teóricos e metodológicos para a obtenção de linhagens de *Ricinus communis*. L. (mamona) com elevada tolerância à estresses abióticos que possibilitem o desenvolvimento de cultivares a agricultura familiar, mediante caracterização bioquímica e molecular do comportamento de genes da família terpeno sintase. Portanto, ambiciona-se que pequenos agricultores sejam favorecidos com este estudo, contribuindo na construção de políticas públicas ambientais com a finalidade de desenvolver cultivares da *R. communis* mais tolerantes a estresses abióticos, considerando as alterações climáticas e os prejuízos econômicos decorrentes. Este trabalho reúne instrumentos multidisciplinares de base científica, com a qual contribui para o fortalecimento dos estudos em metabolômica, bioquímica e biologia molecular de plantas aplicadas ao melhoramento genético na Bahia e no Nordeste.

OBJETIVOS:

Objetivo Geral:

Fornecer subsídios teóricos para a obtenção de linhagens de *Ricinus communis* com elevada tolerância à estresses abióticos que possam ser utilizadas em programas de melhoramento genético e para o desenvolvimento de cultivares que possam ser plantadas pelo pequeno agricultor familiar mediante a caracterização bioquímica e molecular de genes da família terpeno sintase.

Objetivos específicos:

- Selecionar e caracterizar dos genes da família terpeno sintase em *Ricinus communis* mediante ferramentas da bioinformática;
- Determinar condições estressoras e avaliar parâmetros morfológicos de sementes de *Ricinus communis* sob estresse induzido

RESUMO

Das dificuldades que permeiam o estudo dos terpenos, destaca-se a complexidade que envolvem os genomas das espécies envolvidas e as restrições quanto ao estudo das vias biossintéticas para a maioria destes compostos. Este trabalho objetivou fornecer subsídios teóricos e metodológicos para a obtenção de linhagens de *Ricinus communis* com elevada tolerância à estresses abióticos que possam ser utilizadas em programas de melhoramento genético e para o desenvolvimento de cultivares que possam ser plantadas pelo pequeno agricultor rural, mediante caracterização bioquímica e molecular do comportamento de genes da família terpeno sintase. Os resultados obtidos neste trabalho identificaram 46 genes responsáveis pela produção de enzimas altamente especializadas e conhecidas como Terpeno sintase. A localização subcelular, distribuição de éxons, íntrons, regiões regulatórias transcricionais bem como elementos cis promotores são relatados. A designação de subfamílias (TPS-a – TPS-g) são apresentadas de acordo com a classificação de *motifs* conservados encontrados nas sequências e estrutura dos genes. Foram encontrados em todos os genes domínios C-terminal e N-terminal no Pfan característicos das enzimas terpenos sintase (PF01397 e PF03936). Além disso, foram encontrados nos genes padrões de *motifs* conservados, DDXXD, DXDD, NSE/DTE e RRX₈W. Sequências de TPS foram alinhadas frente outras espécies modelos e foi comprovado a presença do *motif* canônico DDXXD, essencial para a regulação enzimática que catalisa a primeira ionização em IPP e DMAPP. Para analisar os efeitos da temperatura (20, 30 e 35 °C), concentração salina (25 mM; 50 mM; e 100 mM de NaCl) e déficit hídrico (-0,2 MPa; -0,6 MPa e -1 MPa) em plântulas de *Ricinus communis* do genótipo Paraguaçu, foi medido o comprimento de raízes e hipocótilos e a massa seca das raízes, hipocótilos, endospermas e cotilédones após 72 horas de germinação e submetidas aos estresses. forma geral comprometeu a alocação de biomassa nas raízes, hipocótilos, endosperma e cotilédones. Os comprimentos das raízes e dos hipocótilos evidenciaram um desvio da normalidade quando expostas ao estresse térmico em 20°C e não houveram diferenças significativas no comprimento das raízes e hipocótilos expostos à 35°C. Com estes achados, podemos observar possíveis translocações de biomassa, transferências de recursos que darão suporte a planta para atravessar os períodos de estresse.

ABSTRACT

Of the difficulties that permeate the study of terpenes, the complexity involved in the genomes of the species involved and the restrictions on the study of biosynthetic pathways for most of these compounds stand out. This work aimed to provide theoretical and methodological subsidies to obtain *Ricinus communis* strains with high tolerance to abiotic stresses that can be used in genetic improvement programs and for family farming, through biochemical and molecular characterization of the behavior of the family's genes. terpene synthase. The results obtained in this work identified 46 genes responsible for the production of highly specialized enzymes known as Terpene synthase. .Subcellular localization, distribution of exons, introns, transcriptional regulatory regions as well as cis promoter elements are reported. The designation of subfamilies (TPS-a – TPS-g) are presented according to the classification of conserved motifs found in the sequences and structure of the genes. C-terminal and N-terminal Pfam domains characteristic of terpene synthase enzymes (PF01397 and PF03936) were found in all genes. In addition, DDXXD, DXDD, NSE/DTE and RRX₈W were found in the standard genes of conserved motifs. TPS sequences were aligned against other model species and the presence of the canonical motif DDXXD, essential for the enzymatic regulation that catalyzes the first ionization in IPP and DMAPP, was confirmed. To analyze the effects of temperature (20, 30 and 35 °C), salt concentration (25 mM; 50 mM; and 100 mM NaCl) and water deficit (-0.2 MPa; -0.6 MPa and - 1 MPa) in seedlings of *Ricinus communis* of the Paraguaçu genotype, the length of roots and hypocotyls and the dry mass of roots, hypocotyls, endosperms and cotyledons were measured after 72 hours of germination and subjected to stress. generally compromised the allocation of biomass in roots, hypocotyls, endosperm and cotyledons. The lengths of roots and hypocotyls showed a deviation from normality when exposed to heat stress at 20°C and there were no significant differences in the length of roots and hypocotyls exposed to 35°C. With these findings, we can observe possible biomass translocations, transfers of resources that will support the plant to go through periods of stress.

SUMÁRIO

CAPÍTULO I	15
1.1 <i>Ricinus communis</i> e sua relevância na ciência brasileira	16
1.2 Plantas sob estresses ambientais	20
1.3 Terpenos: Ocorrência, classificação e biossíntese	22
1.4 Metabolômica no estudo de sistemas biológicos	26
1.4 Bioinformática como ferramenta no estudo de famílias gênicas em plantas	33
CAPÍTULO II	40
EFEITO DOS ESTRESSES ABIÓTICOS HIDRICO, SALINO E OSMÓTICO NO ESTABELECIMENTO DE PLÂNTULAS JOVENS DE <i>RICINUS COMMUNIS</i>	40
1 Introdução.....	41
2.0 Materiais e métodos.....	44
2.1 Tratamentos pré-germinativos em sementes de <i>Ricinus communis</i>	44
2.2 Germinação das sementes sob estresse térmico	Erro! Indicador não definido.
2.3 Germinação das sementes sob estresse salino.....	45
2.4 Germinação das sementes sob estresse hídrico.....	46
3. Resultados e discussões	46
4 Considerações finais	55
5 Referências	57
CAPÍTULO III	61
1 Introdução.....	64
2 Materiais e Métodos	71
2.1 Coleta de dados, classificação quanto a subfamílias e análise filogenética.....	71

2.2	Localizações subcelulares, estrutura das proteínas e análise da região promotora das RcTPS	72
3	Resultados e discussões	72
3.1	Identificação e análise filogenética da família TPS em <i>Ricinus communis</i> (RcTPS).	72
3.2	Análise da estrutura, localização subcelular e região promotora dos genes RcTPS .	78
4	Conclusões.....	86
5	Referências	87

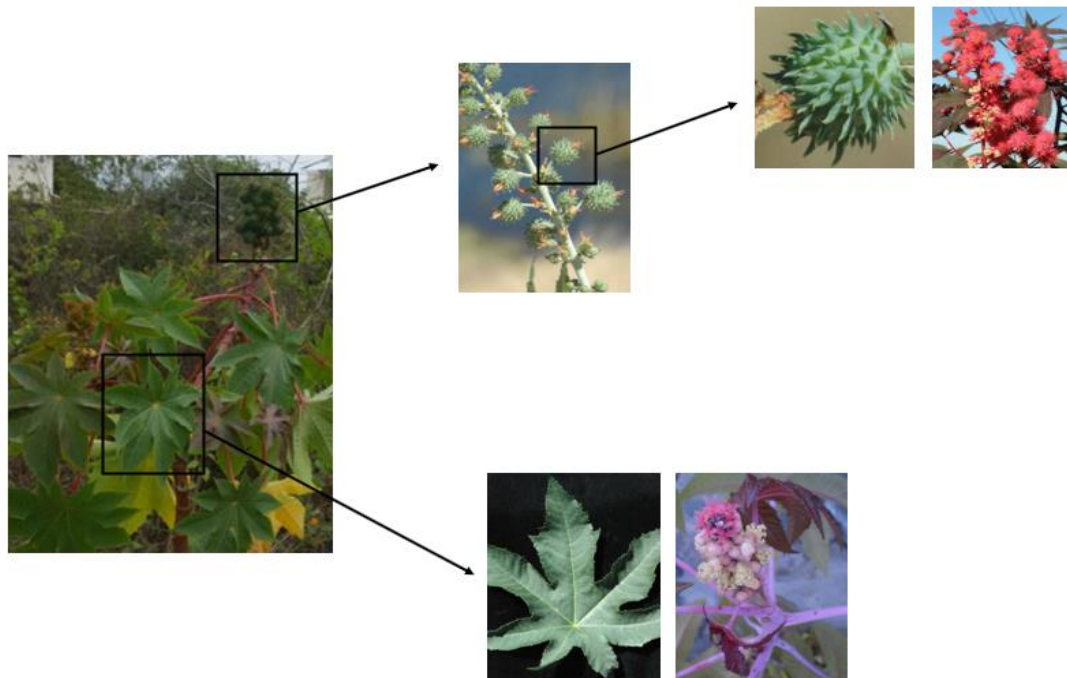
CAPÍTULO I
INTRODUÇÃO GERAL



1.1 *Ricinus communis* e sua relevância na ciência brasileira

A espécie *Ricinus communis*, pertence à família das Euphorbiaceae, do gênero *Ricinus*, também conhecida internacionalmente como Castor Bean. Sua alta diversidade morfológica se dá devido a elementos climáticos e geológicos dos lugares em que habita, gerando diferenças nos hábitos de crescimento, cor das folhas, caules, espinhos ou a ausência deles, ciclo anual e conteúdo lipídico (NÓBREGA, 2008). A maioria da comunidade científica concorda que sua origem se deu no continente Africano, possivelmente na Etiópia (MOSHKIN, 1986; NAIK, 2018). Esta espécie é moderadamente adaptável às condições de calor, salinidade e altas incidências de raios UV. Possui ótimo desenvolvimento nas zonas temperadas de 20 °C à 35 °C, exceto em solos pobres e argilosos. Por isso, pode ser cultivada em zonas tropicais e subtropicais do planeta, podendo alcançar tamanhos que variam de 90 cm a 10 metros de altura. Suas folhas são grandes, com variadas tonalidades verdejantes e avermelhadas que podem medir de 15 a 30 cm de largura com 5 a 11 lóbulos levemente serradas (**Figura 1**).

Figura 1: Espécie *Ricinus communis* com destaque em suas variadas morfologias foliares e aspectos dos frutos



Fonte: Adaptado de KEW SCIENCES, 2020

No Brasil, é popularmente conhecida como mamoeiro, carrapateira, rícino ou simplesmente mamona. Para além de sua popularidade, a *Ricinus communis* é uma oleaginosa

que tem se destacado nos últimos anos por sua importância econômica e qualificação científica. O Brasil é o 3º maior produtor mundial de mamona, ficando atrás apenas da China e da Índia. A soma da produção (biomassa e óleos) nesses países corresponde a 94% da produção mundial. Seus principais compradores de seus insumos são Estados Unidos e União Europeia (SOUZA, 2012; TELES, 2013).

As sementes da *Ricinus communis* (**Figura 2**), também possuem diferentes tonalidades e níveis elevados de óleos, em média cerca de 46 a 55% de óleo por peso (OGUNNIYI, 2006; PATEL et al., 2016). Possuem rica diversidade de compostos químicos presentes no óleo. A principal estrutura química é o ácido ricinoleico (R, Z-12-hidroxiocetadeca-9-enoico) cerca de 90,8% de sua constituição. Seguido de 4,6% de ácido oleico (ácido octadec-9-enóico), 3,5% de ácido linoleico (ácido cis, cis-9,12-octadecadienoico) e cerca de 1% de ácido palmítico (ácido hexadecanóico). As sementes também possuem ricinina e ricina, encontrada em diferentes concentrações em função dos diferentes genótipos

Figura 2: Diversidade de cores e tamanhos das sementes da *Ricinus communis*

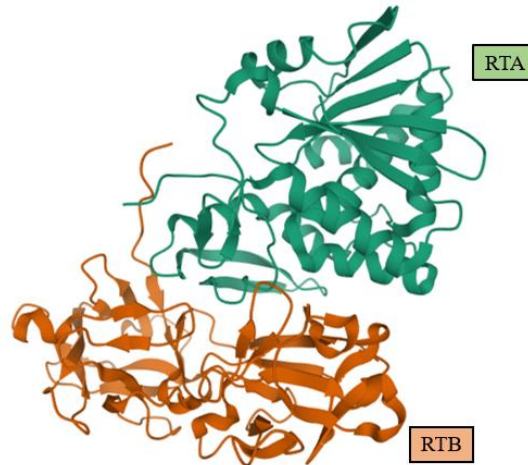


Fonte: Adaptado de WORBS et al., 2011

A ricinina é uma glicoproteína extremamente tóxica em seres humanos e animais, que possui o mecanismo responsável por impedir a atividade ribossomal, inibindo a tradução por remover uma adenina específica do RNA 28S, assim bloqueando a síntese de proteínas. Sua estrutura consiste em duas porções (RTA e RTB) unidas por interação dissulfeto (**Figura 3**). A subunidade RTB se liga aos diferentes resíduos de oligossacarídeos na superfície celular, como a galactose, glicolipídios e outras glicoproteínas que são abundantes em mamíferos, permitindo a entrada da subunidade A, a qual é responsável pelo dano celular (LESHIN et al., 2010; WORBS et al., 2011). Apesar disso, os produtos derivados das sementes de *Ricinus communis* são purificados e separados por diversos métodos, removendo a ricina e

reaproveitando-se todos os seus insumos na indústria cosmética, alimentícia, bem como na produção de corantes, fungicidas, tintas e colas, dentre outros (SEVERINO et al., 2012; MOHAMMAD et al., 2014; WORBS et al. 2015; KALB, et al., 2015; AHMAD et al., 2020).

Figura 3: Estrutura cristalina da glicoproteína ricina. Destaque em verde para a subunidade catalítica RTA, e em laranja, a subunidade RTB



Fonte: Adaptado de <https://www.rcsb.org/3d-view/3RTI>

A produção de mamona ganhou notoriedade no Brasil com o Programa Nacional de Produção e Uso do Biodiesel (PNPB) em 2004, que além de estimular a produção de biodiesel, buscava fomentar a participação ativa de agricultores familiares (CASTRO, 2011). Para isso, fora lançado o Decreto N° 5.297, de 6 de dezembro de 2004 que oficializava o Selo Combustível Social a ser concedido a produtores de biodiesel que promoviam a inclusão social da agricultura familiar e a Lei n° 11.097, de 13 de janeiro de 2005 que regulamenta a inclusão do biodiesel na matriz energética brasileira (BRASIL, 2020; PIRES e LOURENÇO, 2015).

Apesar deste avanço, as políticas públicas brasileiras voltadas para o uso de biodiesel advindos da mamona encontraram limitações significativas. Percebeu-se que grande parte do óleo extraído não era utilizado para produção de biodiesel, mas era vendido para as indústrias químicas por preços mais elevados. Estudos posteriores mostraram que devido à alta viscosidade, não era vantajoso empregá-lo como biocombustível, mas vendê-lo as indústrias químicas. O óleo de rícino é muito utilizado como precursor na produção de produtos manufaturados de alto valor econômico (KULKARNI, M. G.; SAWANT, 2003; DA SILVA e

BATALHA, 2010). Sua produção global se destaca frente a outras oleaginosas, por possuir a única fonte natural comercial de ácido graxo hidroxilado bem como por seu alto teor lipídico. Em média, sua produção global gira em torno de 1250 a 1500 litros por hectare e a Índia lidera esse cenário com produção média de 10^6 toneladas de sementes de mamona e em média $3,5$ a $5,0 \times 10^5$ toneladas de óleo. Para o biênio 2019/2020, indicaram que a Índia produziria 2052 kg / ha, com área total de plantio de 992.000 hectares. Representando um aumento de 29% em relação ao biênio anterior (SEVERINO *et al.*, 2012; KAUR e BHASKAR, 2020; BUSINESS STANDARD, 2020). De acordo com dados disponíveis na Companhia Nacional de Abastecimento (CONAB), a estimativa da produção da *Ricinus communis* para o biênio 2019/2020 é de 725 kg / h, alcançando em 2020/2021 52,7 mil hectares de área semeada, indicando um aumento de 15,85% em comparação com a temporada anterior. A região Nordeste é a principal responsável pela produção de mamona no Brasil, possuindo 43,5 mil hectares de plantações. O estado da Bahia é o grande protagonista, com 43,2 mil hectares e 34,3 mil toneladas produzidas em 2020 cerca de 75% da produção nacional. As estimativas para a produção em 2020 foram de 30,9 mil toneladas, um aumento de 10,3 % em relação ao ano anterior (CONAB, 2020).

No Brasil existem 23 cultivares registradas da *Ricinus communis*. Algumas foram produzidas por cruzamentos assistidos de modo a apresentar melhor adaptação ao clima. Dentre essas, destacam-se os genótipos do Instituto Agrônomo de Campinas (IAC Guarani) que apresentam morfologia rústica e elevada produtividade, o genótipo BRS Energia que apresenta porte baixo, ideal para pequenas plantações, os genótipos BRS Nordestina e BRS Paraguaçu, ambas com sistema radicular mais desenvolvido o que as tornam resistentes a estiagem característica da região Nordeste (CONAB, 2020; BRASIL, 2020; PIRES e LOURENÇO, 2015)

Os últimos três genótipos citados, foram desenvolvidos pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) em colaboração com a Empresa Baiana de Desenvolvimento Agrícola (EBDA) e a Empresa de Pesquisa Agropecuária do Rio Grande do Norte (EMPARN), com a finalidade de ajudar pequenos agricultores rurais em regiões semiáridas do Brasil a expandirem seus negócios e melhorarem sua qualidade de vida (CASTRO, 2011).

Neste cenário promissor, a ciência faz contribuições relevantes em todos os campos desde a produção até a colheita, investigando técnicas modernas que auxiliam todos os processos de produção do óleo e outros insumos da espécie. Contribuições de várias áreas do

conhecimento, da bioquímica à fisiologia vegetal, da genômica à ciências da computação, todas trabalhando juntas desvendando os mecanismos genômicos, transcricômicos, proteômicos e metabolômicos para melhor safra, mesmo em condições adversas (WANG et al., 2018; RIBEIRO et al., 2018; HAN, B. et al., 2020; YU; LI; LIU, 2020).

Em tempos recentes, algoritmos e cálculos de modelagem preditiva foram desenvolvidos para otimizar todos os processos de produção de óleo da mamona. O conhecimento de redes neurais acopladas a experimentos genéticos é útil em prever modelos estatísticos que organizam múltiplas variáveis, como composição do solo, carga nutricional, níveis pluviais, concentração salina, incidências de raios UV dentre outras, com a finalidade de obter as melhores condições para o plantio (HAJAR e VAHABZADEH, 2014; WANG et al., 2019) O Brasil e o mundo direcionam suas pesquisas mais recentes nesta direção, todas elas com importantes contribuições ao conhecimento sobre a *Ricinus communis*

1.2 Plantas sob estresses ambientais

Ao longo do processo evolutivo, os organismos desenvolveram diversos mecanismos biológicos que possibilitaram seu desenvolvimento e reprodução (sexuada ou assexuada), possibilitando-os gerar populações com características genéticas que conferem resistência apesar das adversidades impostas pelo ambiente em que vivem. Foi a capacidade de adaptar-se frente as condições adversas, a principal responsável por gerar a explosão de espécies diversificadas de microrganismos, plantas e animais que vivem hoje (CHRISTENHUSZ, JAMES, BYNG, 2016).

Plantas possuem considerável plasticidade frente a condições estressoras. A alta incidência de espécies demonstra sua habilidade de adaptar-se em condições impostas pelo ambiente. Um relatório produzido pelo Royal Botanic Gardens em 2016, estima que existam cerca de mais de 390 mil espécies de plantas já descritas na literatura científica. Destas, aproximadamente 308.312 são plantas vasculares com 295.383 plantas com flores. Angiospermas (monocotiledôneas: 74.273; eudicotiledôneas: 210.008). Para o reino protista, como as algas com cerca de 44.000, plantas hepáticas com 9 mil, hornworts com 225, musgos 12.700, licófitas 1.290, samambaias 10.560 e gimnospermas 1.079 mil espécies. Cerca de 2.000 mil novas espécies são descritas todos os anos (CHRISTENHUSZ, JAMES, BYNG, 2016).

No Brasil, de acordo com o último relatório Flora do Brasil (2020), existem cerca de 49.993 espécies de plantas brasileiras destacando-se as nativas, as cultivadas e naturalizadas no país. Destas, 4.993 são plantas, 35.553 são angiospermas, 1.610 são classificadas como Briófitas, 114 são gimnospermas e 1403 pertencem as Samambaias e Licófitas (Flora do Brasil, 2020). Esse número expressivo de espécies registradas no Brasil o coloca no pódio em primeiro lugar dentre os países com maior diversidade em espécies de plantas do mundo ao lado da China, Indonésia, México e África do Sul (FLORA BRASIL, 2020; ZAPPI et al., 2015; FIORAVANTI, 2016).

Todas as plantas vivem em ambientes sob constantes mudanças. Muitas vezes, estas não são favoráveis ao seu desenvolvimento, pois desequilibram funções vitais ao estabelecimento. Fatores estressores podem ser classificados como bióticos e abióticos. Causas tais como parasitas, infecção por patógenos, herbivoria, concorrentes naturais, dentre outros, são classificados como causas de estresses bióticos. Por outro lado, fatores tais com frio ou geada, calor, salinidade, crise hídrica, toxinas exógenas, contaminação por metais e poluentes, deficiência de nutrientes, irradiação ionizante, elevação da concentração de CO₂ na atmosfera, dentre outros tipos são classificados como estresses abióticos. (ZHU et al., 2016; FEI et al., 2017; ABHINANDAN et al., 2018).

Estresses abióticos tais como seca ou restrição hídrica, aumento da salinidade do solo e oscilações críticas na temperatura, afetam o desenvolvimento de culturas saudáveis bem como sua distribuição geográfica, o que se configura grave risco à segurança alimentar em muitos países, em especial aqueles em desenvolvimento. Tais fatores, tem sido ainda mais acentuados devido às mudanças climáticas observadas nas últimas décadas, o que provocam drásticas mudanças ecológicas com consequências à economia nacional e global (ALBERTO et al., 2013; BECKLIN et al., 2016)

Estudos demonstram que estresses térmicos, salinos e hídricos, individuais ou combinados são os principais responsáveis por alterações morfológicas e fisiológicas significativas em plantas, incluindo reduções do crescimento da área foliar, da taxa fotossintética, alterações na produção de clorofila e baixos rendimentos de insumos agrícolas. Naturalmente, ao serem expostas a tais estresses, as plantas sofrem alterações no padrão de expressão gênica e assinatura metabólica, comprometendo todas as etapas do seu metabolismo em todo o ciclo de vida (AINSWORTH e ROGERS, 2007; TAIZ e ZEIGER, 2017; BOWLING et al., 2020).

As culturas localizadas nas regiões áridas e semiáridas do mundo, como a *Ricinus communis* bem como cultivares alimentícia como soja, arroz e milho são as mais afetadas com o agravamento da crise climática. Estes fatores combinados contribuem para o aprofundamento da desigualdade social no campo, especialmente para microempreendedores rurais (LAL; ALAVALAPATI; MERCER, 2011; MASSETTI et al., 2013).

1.3 Terpenos: Ocorrência, classificação e biossíntese

Os terpenos compreendem a maior classe de produtos naturais já descritos na literatura científica, mais de 80 mil dessas substâncias já foram reportadas (CUI et al., 2015; LEAVELL et al., 2016; CHRISTIANSON, 2018). Seus derivados, os terpenóides, também chamados de isoprenóides, estão presentes em quase todos os organismos vivos incluindo plantas (AUBOURG; LECHARNY; BOHLMANN, 2002; THOLL, 2015; KARUNANITHI; ZERBE, 2019) bactérias (YAMAD; CANE; IKEDA, 2012), fungos (SCHMIDT-DANNERT, 2015) arqueias (MATSUMI et al., 2011; HENRY et al., 2018), amebas (CHEN et al 2019), insetos (ZHAO et al. 2018; LANCASTER et al., 2018; BONCAN et al., 2020) e organismos marinhos (GROSS; KÖNIG, 2006; KWON et al., 2020).

As plantas são de longe as campeãs no que tange a diversidade de terpenos produzidos. Estes exercem papéis regulatórios essenciais na germinação, crescimento e desenvolvimento das plantas saudáveis, como as giberelinas, brassinosteróides, carotenoides e clorofilas. A grande maioria destas espécies terpênicas são classificadas como metabólitos secundários, auxiliando a homeostase celular em processos biológicos vitais como sinalização hormonal, respostas à estresses bióticos e abióticos, comunicação e defesa (SINGH; SHARMA, 2015; BOCAN et al., 2020; IOVINELLA et al., 2020).

Os terpenos são convencionalmente classificados de acordo a quantidade de carbonos que possuem em seu esqueleto básico. Dessa forma, compostos que possuem 5 carbonos (C₅) são denominados hemiterpenos, monoterpenos (C₁₀), sesquiterpenos (C₁₅), diterpenos (C₂₀), triterpenos (C₃₀), tetraterpenos (C₄₀) e mais que que isso são classificados como politerpenos (C_{5_n}). Podem assumir formas cíclicas e acíclicas, anéis condensados, formas bicíclicas, grupos carbonilados e insaturações (CHRISTIANSON, 2006; NIEUWENHUIZEN et al., 2013).

A grande maioria dos terpenos são biossintetizadas no citosol e nos plastídios, através da ação das enzimas terpeno sintases (TPS). Atuam também na rota de produção dos terpenos as

enzimas prenilttransferases (PTs) e outras enzimas coadjuvantes pertencentes ao citocromo P450, são responsáveis por reações de ionização, alongamento da cadeia e reações do tipo redox (LIANG; KO; WANG, 2002; KOPCSAYOVÁ.; VRANOVÁ, 2019). As enzimas da família TPS já foram largamente caracterizadas em diversos organismos, como por exemplo em *Arabidopsis thaliana* (AUBOURG et al., 2002), *Populus trichocarpa* (IRMISCH, et al., 2014), *Solanum lycopersicum* (ZHOU E PICHERSKY, 2020) *Oryza sativa* (CHEN et al 2018), *Sorghun bicolor* (ZHUANG et al 2013), *Vitis vinifera* (MATARESE, F. et al., 2014; SMIT et al., 2020) e largamente em bactérias como em 30 espécies diferentes de *Streptomyces* e em quase todos as espécies de *Actinomycetales* (YAMADA et al., 2015; REDDY et al., 2020).

Em especial, as TPS regulam toda a produção de terpenos e derivados, conduzindo a rearranjos estereoespecíficos de carbocátions, rearranjos de metila e hidretos, mudanças conformacionais que levam a formação desta rica classe de metabólitos secundários. A biossíntese se inicia através da ionização do grupo fosfato e posterior condensação dos seus principais isômeros de 5 carbonos: Difosfato de dimetilalila (DMAPP) e o difosfato de isopentenila (IPP). Intermediários reacionais como o difosfato de farnesila (FPP), difosfato de geranila (GPP) e o difosfato de geranilgeranila (GGPP) são formados via condensação, tornando-se novos precursores de mono, sesqui e diterpenos, respectivamente. As plantas utilizam duas vias metabólicas distintas na construção desses dois isômeros C5. A via do mevalonato (MVA) que ocorre no citosol e tem Acetil-CoA como precursor e a via do 2- C - metil-D-eritritol-4-fosfato (MEP) que ocorre nos plastídios, o qual deriva do piruvato e do gliceraldeído-3-fosfato. As duas vias metabólicas também são observadas em fungos, arqueais e em algumas bactérias. Porém, a maioria das bactérias sintetizam unidades C5 via MEP (CHANG; CHENG; WANG, 2021) (**Figura 4**).

Estudos demonstram que a presença da via MVA em arqueias sugerem que esta rota biossintética também fazia parte do arsenal metabólico do último ancestral comum e que fora transmitido a todos seus descendentes. Por outro lado, estudos indicam que a via do MEP foi herdada por transferência horizontal de diferentes genes de origem bacteriana (LANGE et al., 2000; KARUNANITHI; ZERBE, 2019). Estudos de coexpressão genômica em *Arabidopsis thaliana* demonstraram interações significativas entre clusters de genes das vias MVA e MEP, o que sugerem que em plantas, estas vias buscam compensar o gasto energético empregado na produção simultânea de IPP e DMAPP, com vantagens que permitem maior controle de terpenos especializados em compartimentos celulares específicos bem como regulação e sinalização hormonal. Apesar de não conclusivos, estes estudos sugerem uma comunicação

entre as vias, com possível retenção e produção de terpenos em épocas distintas do desenvolvimento da planta (RODRÍGUEZ-CONCEPCIÓN; BORONAT, 2015; KARUNANITHI; ZERBE, 2019; GUO et al., 2020).

1.4 Metabolômica no estudo de sistemas biológicos

Nas últimas décadas, o estudo integrado da biologia de sistemas dinâmicos constitui grande desafio ao entendimento geral e específico de sistemas biológicos integrados, altamente complexos e especializados do ponto de vista bioquímico. As ciências “ômicas”, surgem em 1990 com a integração das várias áreas da ciência da vida como genômica, proteômica, transcriptômica, metabolômica e suas bases conceituais, no esforço de entender mecanismos genômicos em humanos. Desde então, seus estudos têm sido largamente aplicados em quase todas as espécies de organismos, sem muitas adaptações experimentais e de forma transversal no estudo da biologia de sistemas (MARTÍNEZ-ESPINOSA, 2017; ANTONELLI et al., 2019; CRANDALL et al., 2020).

A metabolômica contribui amplamente para o desenvolvimento da compreensão dos componentes químicos presentes nos organismos alvo, o que inclui um estudo do perfil químico, identificação e quantificação dos metabólitos primários e/ou secundários de baixo peso molecular, o que permite avaliar com maior compreensão as assinaturas metabólicas em qualquer faixa do desenvolvimento de qualquer organismo, mesmo quando submetidas a condições específicas, como estresses, por exemplo. Pode-se analisar tecidos, órgãos ou até células específicas. Desta forma, consegue-se avaliar diferentes processos biológicos e suas possíveis alterações e/ou disfunções desde a expressão gênica, síntese proteica até biossíntese de metabólitos que regulam processos bioquímicos tão essenciais à manutenção da vida. Os metabólitos consistem no produto final da atividade bioquímica, de modo que compreender sua formação e fatores biossintéticos interligados é de suma importância na caracterização bioquímica de determinado organismo (ARBONA et al., 2013; SINGH, 2020).

Como suporte instrumental neste estudo integrado, técnicas avançadas de separação analíticas, métodos da biologia molecular e análises de bioinformática são amplamente utilizadas para avaliar a presença ou ausência de componentes químicos, proteínas, expressão de genes, objetivando entender a regulação bioquímica dos organismos em estudo e seus seus desdobramentos fenotípicos. Entre elas, destacam-se a Eletroforese, Sequenciamento, Quantitative Reverse Transcription PCR (RT-qPCR), Espectrometria de Massas com variados detectores bem como a High Performance Liquid Chromatography (HPLC). A partir destes estudos avançados, consegue-se traçar um perfil metabolômico singular para cada organismo e compreender mecanismos de regulação bioquímica. Na **tabela 1**, listam-se os Metabólitos que

já

foram identificados em *Ricinus communis* bem como suas respectivas massas em alta resolução.

Tabela 1: Metabólitos já identificados em *Ricinus communis* e suas respectivas massas de alta resolução com adutos de Hidrogênio [M + H]⁺ e Sódio [M + Na]⁺

Number	Name	Formula	High resolution mass	[M + H] ⁺	[M + Na] ⁺	References
1	ricinine	C ₈ H ₈ N ₂ O ₂	164.0586	165.0664	187.0483	RIBEIRO et al., 2016
2	3-carboxy-4-methoxy-N-methyl-2-pyridone	C ₈ H ₉ NO ₄	183.0532	184.0610	206.0430	RIBEIRO et al., 2016
3	n-demethylricinine	C ₇ H ₆ N ₂ O ₂	150.0429	151.0508	173.0327	RIBEIRO et al., 2016
4	methyl 5-(3-cyano-1-ethyl-2-oxo-1,2-dihydropyridine-4-il) pentanoate	C ₁₃ H ₁₆ N ₂ O ₃	248.1161	249.1239	271.1059	RIBEIRO et al., 2016
5	1-methyl-4-(4-metilpentiloksi) pyridine-2 (1H)-on	C ₁₂ H ₁₉ NO ₂	209.1416	210.1494	232.1313	RIBEIRO et al., 2016
6	quercetin	C ₁₅ H ₁₀ O ₇	302.0427	303.0505	325.0324	RIBEIRO et al., 2016
7	quercetin-3-O-β-D-galactoside	C ₂₁ H ₂₀ O ₁₂	464.0955	465.1033	487.0853	RIBEIRO et al., 2016
8	quercetin-3-O-β-D-glucopyranoside	C ₂₁ H ₁₉ O ₁₂	463.0877	464.0955	486.0774	RIBEIRO et al., 2016
9	quercetin-3-O-β-D-rutinoside	C ₂₇ H ₃₀ O ₁₆	610.1534	611.1612	633.1432	RIBEIRO et al., 2016
10	quercetin-3-O-β-D-xylopyranoside	C ₂₀ H ₁₈ O ₁₁	434.0849	435.0927	457.0747	RIBEIRO et al., 2016
11	kaempferol-3-O-β-D-glucopyranoside	C ₂₁ H ₂₀ O ₁₁	448.1006	449.1084	471.0903	RIBEIRO et al., 2016
12	kaempferol-3-O-β-D-rutinoside	C ₂₇ H ₃₀ O ₁₅	594.1585	595.1663	617.1482	RIBEIRO et al., 2016
13	kaempferol-3-O-β-dxylopyranoside	C ₂₀ H ₁₈ O ₁₁	434.0849	435.0927	457.0747	RIBEIRO et al., 2016
14	catechin	C ₁₅ H ₁₄ O ₆	290.0790	291.0869	313.0688	RIBEIRO et al., 2016
15	epicatechin	C ₁₅ H ₁₄ O ₆	290.0790	291.0869	313.0688	RIBEIRO et al., 2016
16	luteolin	C ₁₅ H ₁₀ O ₆	286.0477	287.0556	309.0375	RIBEIRO et al., 2016
17	vitexin	C ₂₁ H ₂₀ O ₁₀	432.1056	433.1135	455.0954	RIBEIRO et al., 2016
18	ellagic acid	C ₁₄ H ₆ O ₈	302.0063	303.0141	324.9960	RIBEIRO et al., 2016
19	gallic acid	C ₇ H ₆ O ₅	170.0215	171.0294	193.0113	RIBEIRO et al., 2016
20	gentisic acid	C ₇ H ₆ O ₄	154.0266	155.0344	177.0164	RIBEIRO et al., 2016

21	vanillic acid	C ₈ H ₈ O ₄	168.0423	169.0501	191.0320	RIBEIRO et al., 2016
22	3,4-dimethoxy-6,8-dihydroxycoumarin	C ₁₁ H ₁₀ O ₆	238.0477	239.0556	261.0375	RIBEIRO et al., 2016
23	isofraxidine	C ₁₁ H ₁₀ O ₅	222.0528	223.0607	245.0426	RIBEIRO et al., 2016
24	scopoletin	C ₁₀ H ₈ O ₄	192.0423	193.0501	215.0320	RIBEIRO et al., 2016
25	α-tocopherol	C ₂₉ H ₅₀ O ₂	430.3811	431.3889	453.3709	RIBEIRO et al., 2016
26	β-tocopherol	C ₂₈ H ₄₈ O ₂	416.3654	417.3733	439.3552	RIBEIRO et al., 2016
27	δ-tocopherol	C ₂₇ H ₄₆ O ₂	402.3498	403.3576	425.3396	RIBEIRO et al., 2016
28	γ-tocopherol	C ₂₈ H ₄₈ O ₂	416.3654	417.3733	439.3552	RIBEIRO et al., 2016
29	(2R,4aR,8aR)- 3,4,4,8 atetrahydro-4ahydroxy- 2,6,7,8 atetramethyl-2-(4,8,12-trimethyltridecyl)-2H chromene- 5,8-dione	C ₂₉ H ₅₀ O ₄	462.3709	463.3787	485.3607	RIBEIRO et al., 2016
30	α-thujone	C ₁₀ H ₁₆ O	152.1201	153.1279	175.1099	RIBEIRO et al., 2016
31	1,8-cineole	C ₁₀ H ₁₈ O	154.1358	155.1436	177.1255	RIBEIRO et al., 2016
32	α -pinene	C ₁₀ H ₁₆	136.1252	137.1330	159.1150	RIBEIRO et al., 2016
33	camphor	C ₁₀ H ₁₆ O	152.1201	153.1279	175.1099	RIBEIRO et al., 2016
34	camphene	C ₁₀ H ₁₆	136.1252	137.1330	159.1150	RIBEIRO et al., 2016
35	ficusic acid	C ₁₁ H ₁₆ O ₃	196.1099	197.1178	219.0997	RIBEIRO et al., 2016
36	phytol	C ₂₀ H ₄₀ O	296.3079	297.3157	319.2977	RIBEIRO et al., 2016
37	callyspinol	C ₂₀ H ₃₈ O	294.2923	295.3001	317.2820	RIBEIRO et al., 2016
38	beyerene	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
39	Cembrene	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
40	kaurene	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
41	sandaracopimaradiene	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
42	trachylobane	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
43	casbene	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	RIBEIRO et al., 2016
44	(3E,7Z,11E)-19-hydroxycasba-3,7,11-trien-5-one	C ₂₀ H ₃₀ O ₂	302.2246	303.2324	325.2143	RIBEIRO et al., 2016
45	6-α-hydroxy-10-β-methoxy-7α,8α-epoxy-5-oxocasbane-20,10-olide	C ₂₁ H ₂₈ O ₆	376.1886	377.1964	399.1784	RIBEIRO et al., 2016
46	campesterol	C ₂₈ H ₄₈ O	400.3705	401.3783	423.3603	RIBEIRO et al., 2016
47	β-sitosterol	C ₂₉ H ₅₀ O	414.3862	415.3940	437.3759	RIBEIRO et al., 2016

48	stigmasterol	C ₂₉ H ₄₈ O	412.3705	413.3783	435.3603	RIBEIRO et al., 2016
49	3-O-β-D-glycosylsitosterol	C ₃₅ H ₆₀ O ₆	576.4390	577.4468	599.4288	RIBEIRO et al., 2016
50	7-oxo-β-sitosterol	C ₂₉ H ₄₈ O ₂	428.3654	429.3733	451.3552	RIBEIRO et al., 2016
51	stigmasterol arachidate	C ₄₉ H ₈₆ O ₂	706.6628	707.6706	729.6526	RIBEIRO et al., 2016
52	stigmasterol oleate	C ₄₇ H ₈₀ O ₃	692.6107	693.6186	715.6005	RIBEIRO et al., 2016
53	stigmasterol stearate	C ₄₇ H ₈₂ O ₂	678.6315	679.6393	701.6213	RIBEIRO et al., 2016
54	stigmast4-en-3-one	C ₂₉ H ₄₈ O	412.3705	413.3783	435.3603	RIBEIRO et al., 2016
55	stigmast-4-en-6-β-ol-3-one	C ₂₉ H ₄₈ O ₂	428.3654	429.3733	451.3552	RIBEIRO et al., 2016
56	Stigmast-4-en-3,6-dione	C ₂₉ H ₄₆ O ₂	426.3498	427.3576	449.3396	RIBEIRO et al., 2016
57	3-O-β-D-glycosylstigmasterol	C ₃₅ H ₅₈ O ₆	574.4233	575.4312	597.4131	RIBEIRO et al., 2016
58	erandone	C ₃₀ H ₄₆ O ₂	438.3498	439.3576	461.3396	RIBEIRO et al., 2016
59	lupeol	C ₃₀ H ₅₀ O	426.3862	427.3940	449.3759	RIBEIRO et al., 2016
60	30-nor-lupan-3β-ol-20-one	C ₂₉ H ₄₈ O ₂	428.3654	429.3733	451.3552	RIBEIRO et al., 2016
61	lup-20(29)-en-3β,15α-diol	C ₃₀ H ₅₀ O ₂	442.3811	443.3889	465.3709	RIBEIRO et al., 2016
62	acetylaleuritic acid	C ₃₂ H ₅₀ O ₄	498.3709	499.3787	521.3607	RIBEIRO et al., 2016
63	lup-20(29)-en-15α-ol-3-one	C ₃₀ H ₄₈ O ₂	440.3654	441.3733	463.3552	RIBEIRO et al., 2016
64	3-O-[β-D- glucuronopyranosyl-(1→3)-α-L-rhamnopyranosyl-(1→2)β-D-glucopyranosyl]-4α,20α-di(hydroxymethyl)olean-12-en-28-oic acid	C ₄₈ H ₇₆ O ₂₀	972.4930	973.5008	995.4828	RIBEIRO et al., 2016
65	ricinoleic acid	C ₁₈ H ₃₄ O ₃	298.2508	299.2586	321.2406	RIBEIRO et al., 2016
66	linoleic acid	C ₁₈ H ₃₂ O ₂	280.2402	281.2481	303.2300	RIBEIRO et al., 2016
67	linolenic acid	C ₁₈ H ₃₀ O ₂	278.2246	279.2324	301.2143	RIBEIRO et al., 2016
68	palmitic acid	C ₁₆ H ₃₂ O ₂	256.2402	257.2481	279.2300	RIBEIRO et al., 2016
69	stearic acid	C ₁₈ H ₃₆ O ₂	284.27153	285.2794	307.2613	RIBEIRO et al., 2016
70	methyl ricinoleate	C ₁₉ H ₃₆ O ₃	312.2664	313.2743	335.2562	RIBEIRO et al., 2016
71	methyl linoleate	C ₁₉ H ₃₄ O ₂	294.2559	295.2637	317.2457	RIBEIRO et al., 2016
72	aleuritic acid	C ₁₆ H ₃₂ O ₅	304.2250	305.2328	327.2147	RIBEIRO et al., 2016
73	gondoic acid	C ₂₀ H ₃₈ O ₂	310.2872	311.2950	333.2770	RIBEIRO et al., 2016
74	oleic acid	C ₁₈ H ₃₄ O ₂	282.2559	283.2637	305.2457	RIBEIRO et al., 2016

75	1-palmitic acid glycerol Ester	C ₁₇ H ₃₄ O ₄	302.2457	303.2535	325.2355	RIBEIRO et al., 2016
76	1-palmitic acidglycerol-2,3-dimethyl ketal ester	C ₂₀ H ₃₈ O ₄	342.2770	343.2848	365.2668	RIBEIRO et al., 2016
77	n-heptacosanyl oleate	C ₄₅ H ₈₈ O ₂	660.6784	661.6863	683.6682	RIBEIRO et al., 2016
78	triricinolein	C ₅₇ H ₁₀₄ O ₉	932.7680	933.7759	955.7578	RIBEIRO et al., 2016
79	arachidoyl arabinoside	C ₂₆ H ₄₈ O ₆	456.3451	457.3529	479.3349	RIBEIRO et al., 2016
80	n-butylricicommunioate	C ₁₂ H ₁₄ O ₅	238.0841	239.0920	261.0739	RIBEIRO et al., 2016
81	Ethyl brevifolincarboxylate	C ₁₅ H ₁₂ O ₈	320.0532	321.0610	343.0430	RIBEIRO et al., 2016
82	n-hexatriacont-14-ene	C ₃₆ H ₇₂	504.5634	505.5712	527.5532	RIBEIRO et al., 2016
83	methyl communisoate	C ₁₀ H ₁₀ O ₅	210.0528	211.0607	233.0426	RIBEIRO et al., 2016
84	triterpene glicosilate	C ₃₆ H ₅₆ O ₉	632.3924	633.4003	655.3822	SANTOS et al., 2018
85	lathyrane	C ₂₀ H ₂₈ O ₃	316.2038	317.2117	339.1936	SANTOS et al., 2018
86	tigliane	C ₂₀ H ₃₄	274.2661	275.2739	297.2558	SANTOS et al., 2018
87	ingenane	C ₂₀ H ₃₄	274.2661	275.2739	297.2558	SANTOS et al., 2018
88	jatrophone 4	C ₃₉ H ₅₂ O ₁₄	744.3357	745.3435	767.3255	KING et al. 2014
89	jatrophone 5	C ₄₁ H ₄₉ NO ₁₄	779.3153	780.3231	802.3051	KING et al. 2014
90	jatrophone 2	C ₃₉ H ₅₀ O ₁₅	758.3150	759.3228	781.3047	KING et al. 2014
91	ingenol mebutate	C ₂₅ H ₃₄ O ₆	430.2355	431.2434	453.2253	KING et al. 2014
92	resiniferatoxin	C ₃₇ H ₄₀ O ₉	628.2672	629.2751	651.2570	KING et al. 2014
93	β-caryophyllene	C ₁₅ H ₂₄	204.1878	205.1956	227.1776	DARMANIN et al. 2009
94	ent-Trachylobane	C ₂₀ H ₃₂	272.2504	273.2582	295.2402	JACKSON et al. 2014
95	bufotenine O-glucoside	C ₈ H ₈ N ₂ O ₂	164.0586	165.0664	187.0483	JACKSON et al. 2014
96	5α,6β-dihydroxysitosterol	C ₂₉ H ₅₂ O ₃	448.3916	449.3995	471.3814	JACKSON et al. 2014
97	all-trans-retinyl linolate	C ₃₈ H ₅₈ O ₂	546.4437	547.4515	569.4335	JACKSON et al. 2014
98	1-O-α-D-glucopyranosyl-1,2-eicosandiol	C ₂₆ H ₅₂ O ₇	476.3713	477.3791	499.3611	JACKSON et al. 2014
99	phylloquinone	C ₃₁ H ₄₆ O ₂	450.3498	451.3576	473.3396	JACKSON et al. 2014
100	momordol	C ₂₆ H ₄₈ O ₅	440.3502	441.3580	463.3399	JACKSON et al. 2014
101	6-hydroxymethylumazine	C ₇ H ₆ N ₄ O ₃	194.0440	195.0518	217.0338	SANTOS et al., 2018
102	8-methylcaffeine	C ₉ H ₁₂ N ₄ O ₂	208.0960	209.1039	231.0858	SANTOS et al., 2018

103	sodium oleate	$C_{18}H_{33}NaO_2$	304.2378	305.2457	327.2276	SANTOS et al., 2018
104	pubescenol	$C_{32}H_{40}O_{10}$	584.2622	585.2700	607.2519	SANTOS et al., 2018
105	phytenic acid	$C_{20}H_{38}O_2$	310.2872	311.2950	333.2770	SANTOS et al., 2018
106	22:3-Glc-campesterol	$C_{56}H_{94}O_7$	878.7000	879.7078	901.6897	SANTOS et al., 2018
107	behenic acid	$C_{22}H_{44}O_2$	340.3341	341.3420	363.3239	SANTOS et al., 2018
108	caprylic acid	$C_8H_{16}O_2$	144.1150	145.1229	167.1048	SANTOS et al., 2018
109	heptadecanoic acid	$C_{17}H_{34}O_2$	270.2559	271.2637	293.2457	SANTOS et al., 2018
110	methyl icosanoate	$C_{21}H_{42}O_2$	326.3185	327.3263	349.3083	SANTOS et al., 2018
111	methyl palmitate	$C_{17}H_{34}O_2$	270.2559	271.2637	293.2457	SANTOS et al., 2018
112	methyl paleate	$C_{17}H_{32}O_2$	268.2402	269.2481	291.2300	SANTOS et al., 2018
113	methyl stearate	$C_{19}H_{38}O_2$	298.2872	299.2950	321.2770	SANTOS et al., 2018
114	myristic acid	$C_{14}H_{28}O_2$	228.2089	229.2168	251.1987	SANTOS et al., 2018
115	4-hydroxycinnamic acid	$C_9H_8O_3$	164.0473	165.0552	187.0371	SANTOS et al., 2018
116	2-ketobutyric acid	$C_4H_6O_3$	102.0317	103.0395	125.0215	SANTOS et al., 2018
117	shikimic acid	$C_7H_{10}O_5$	174.0528	175.0607	197.0426	SANTOS et al., 2018
118	succinic acid	$C_4H_6O_4$	118.0266	119.0344	141.0164	SANTOS et al., 2018
119	allothreonine	$C_4H_9NO_3$	119.0582	120.0661	142.0480	SANTOS et al., 2018
120	arabitol/xylitol	$C_5H_{12}O_5$	152.0685	153.0763	175.0582	SANTOS et al., 2018
121	canavanine	$C_5H_{12}N_4O_3$	176.0909	177.0988	199.0807	SANTOS et al., 2018
122	fructose (sugarketose)	$C_6H_{12}O_6$	180.0634	181.0712	203.0532	SANTOS et al., 2018
123	mannitol (sugaralcohol)	$C_6H_{14}O_6$	182.0790	183.0869	205.0688	SANTOS et al., 2018
124	sedoheptulose	$C_7H_{14}O_7$	210.0740	211.0818	233.0637	SANTOS et al., 2018
125	sorbose (sugar ketose)	$C_6H_{12}O_6$	180.0634	181.0712	203.0532	SANTOS et al., 2018
126	squalene	$C_{30}H_{50}$	410.3913	411.3991	433.3810	SANTOS et al., 2018
127	threitol	$C_4H_{10}O_4$	122.0579	123.0657	145.0477	SANTOS et al., 2018
128	threose	$C_4H_8O_4$	120.0423	121.0501	143.0320	SANTOS et al., 2018
129	(-) - α -copaene	$C_{15}H_{24}$	204.1878	205.1956	227.1776	XIE et al., 2012
130	(+) - δ -cadinene	$C_{15}H_{24}$	204.1878	205.1956	227.1776	XIE et al., 2012
131	(E, E)- α -farnesene	$C_{15}H_{24}$	204.1878	205.1956	227.1776	XIE et al., 2012

132	30-Norlupan-3beta-ol-20-one	C ₂₉ H ₄₈ O ₂	428.3654	429.3733	451.3552	LI et al., 2014
133	lup-20(29)-en-3beta,15alpha-diol	C ₃₀ H ₅₀ O ₂	442.3811	443.3889	465.3709	LI, et al., 2014
134	acetylaleuritolic acid	C ₃₂ H ₅₀ O ₄	498.3709	499.3787	521.3607	LI, et al., 2014
135	stigmast-4-en-3-one	C ₂₉ H ₄₈ O	412.3705	413.3783	435.3603	LI, et al., 2014
136	stigmast4-en-3,6-dione	C ₂₉ H ₄₆ O ₂	426.3498	427.3576	449.3396	LI, et al., 2014
137	(3E.7Z.11E)-19-hydroxycasba-3,7,11-trien-5-one	C ₂₀ H ₃₀ O ₂	302.2246	303.2324	325.2143	TAN et al., 2009 al
138	6α-hydroxy-10β-methoxy-7α.8α-epoxy-5-oxocasbane-20,10-olide	C ₂₁ H ₂₈ O ₆	376.1886	377.1964	399.1784	TAN et al., 2009 al
139	15a-hydroxylup-20(29)-en-3-one	C ₃₀ H ₄₈ O ₂	440.3654	441.3733	463.3552	LIM, 2012
140	Docosane	C ₂₂ H ₄₆	310,3600	311,3678	333,3497	ABD-ALLA, et al., 2015
141	Ipsdienol	C ₁₀ H ₁₆ O	152.1201	153.1279	175.1099	ABD-ALLA, et al., 2015
142	9,10-Dibromo anthracene	C ₁₄ H ₈ Br ₂	333,8993	334.9071	356,8890	ABD-ALLA, et al., 2015
143	Germacra-1(10),5-dien-4-ol	C ₁₅ H ₂₆ O	222,1984	223.2062	245,1881	ABD-ALLA, et al., 2015
144	Longifolene	C ₁₅ H ₂₄	204.1878	205.1956	227.1776	ABD-ALLA, et al., 2015
145	2,6-bis-(1,1-Dimethyl-propyl)-4-methyl-phenol (DPMP)	C ₁₇ H ₂₈ O	248.2140	249.2218	271.2038	ABD-ALLA, et al., 2015
146	1-Isoquinoline carbonitrile	C ₁₀ H ₆ N ₂	1.540.530	155.0609	177.0428	ABD-ALLA, et al., 2015
147	2-Bromo-1,4-dimethoxybenzene	C ₈ H ₉ BrO ₂	2.159.785	216.9864	238.9683	ABD-ALLA, et al., 2015

1.4 Bioinformática como ferramenta no estudo de famílias gênicas em plantas

Com o desenvolvimento de novas tecnologias de sequenciamento genômico em diversas plataformas online, tornou-se disponível um imenso arsenal de dados genômicos que são utilizados para elucidar padrões transcriptômicos e metabolômicos de vários organismos. Com este conjunto de dados, pesquisadores no mundo inteiro se debruçam afim de elucidar padrões bioquímicos que regulam a homeostase nos organismos bem como quaisquer disfunções observadas como em patologias ou em exposição à fatores de estresses bióticos ou abióticos (ZHOU E PICHERSKY, 2020; MARTIN et al., 2010; IRMISCH et al., 2014).

A família terpeno sintase tem sido vastamente caracterizada em várias espécies de plantas, bem como organismos marinhos e microrganismos terrestres. Já foram utilizados diversos métodos em bioinformática desde a década de 90, com o objetivo de compreender rotas biossintética na produção de metabólitos, em especial em terpenos (AUBOURG et al., 2002). Em *Arabidopsis thaliana* uma família de 40 genes de terpenos sintases (AtTPS) foi parcialmente elucidadas por análises genômicas associadas a caracterizações bioquímicas posteriores. Os resultados sugeriram que a grande família de genes AtTPS agrupam-se em cinco subfamílias filogenéticas da superfamília TPS em espécies vegetais. Surpreendentemente, trinta AtTPS se assemelham muito, em todos os aspectos da arquitetura dos AtTPS, relação de sequência e posição filogenética, os genes para monoterpene sintases de plantas, sesquiterpenos sintases ou diterpenos sintases do metabolismo secundário da espécie (AUBOURG et al., 2002). Mais tarde, estruturas tridimensionais das enzimas TPS foram caracterizadas lançando luz sobre a funcionalidade destas enzimas na biossíntese de diversos metabólitos terpênicos.

De forma similar, várias espécies foram utilizadas no estudo biológico sistêmico dos terpenos desde os níveis da tradução e transcrição até a biossíntese de metabólitos, bem como suas variações frente a estresses bióticos e abióticos e interações ecológicas – este é o caso do *Solanum lycopersicum*, *Nicotiana benthamiana*, *Vitis vinífera*, *Populus trichocarpa*, *Oryza sativa* e *Sorghum bicolor*, dentre outras (ZHOU E PICHERSKY, 2020; MARTIN et al., 2010; IRMISCH et al., 2014; DANNER et al., 2012; CHEN et al., 2020; HAY et al., 2018). O estudo com estas espécies serviu de base teóricas e metodológicas para a investigação que compõe este estudo preliminar.

REFERÊNCIAS

ABHINANDAN, K. et al. Abiotic stress signaling in wheat—an inclusive overview of hormonal interactions during abiotic stress responses in wheat. **Frontiers in plant science**, v. 9, p. 734, 2018.

AHMAD, M. H. et al. Thermal Process of Castor and Plant Based Oil. **Indonesian Journal of Chemistry**, v. 20, n. 1, p. 237-247.

AINSWORTH, E. A.; ROGERS, A. The response of photosynthesis and stomatal conductance to rising [CO₂]: mechanisms and environmental interactions. **Plant, cell & environment**, v. 30, n. 3, p. 258-270, 2007.

ALBERTO, F. J. et al. Potential for evolutionary responses to climate change—evidence from tree populations. **Global Change Biology**, v. 19, n. 6, p. 1645-1661, 2013.

ANTONELLI, et al., Integrating imaging and omics data: A review. **Biomedical Signal Processing and Control**, v. 52, p. 264-280, 2019

AUBOURG, S.; LECHARNY, A.; BOHLMANN, J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of Arabidopsis thaliana. **Molecular Genetics and Genomics**, v. 267, n. 6, p. 730-745, 2002.

AUBOURG, S. et al. Structure, organization and putative function of the genes identified within a 23.9-kb fragment from Arabidopsis thaliana chromosome IV. **Gene**, v. 199, n. 1-2, p. 241-253, 1997.

ARBONA, V. et al. Metabolomics as a tool to investigate abiotic stress tolerance in plants. **International journal of molecular sciences**, v. 14, n. 3, p. 4885–4911, 2013.

BRASIL, 2020. Ministério da Agricultura, Pecuária e Abastecimento. Programa Nacional de Produção e Uso do Biodiesel (PNPB). Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/agricultura-familiar/biodiesel/programa-nacional-de-producao-e-uso-do-biodiesel-pnpb>. Acessado em 12/08/2020.

_____. Decreto nº 5.297 de 6 de dezembro de 2004. Presidência da República, Casa Civil. Disponível em <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Decreto/D5297.htm> Acesso em 12/08/2020.

_____. Lei nº 11.097, de 13 de janeiro de 2005. Presidência da República, Casa Civil. Disponível em <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Lei/L11097.htm> Acesso em 12/08/2020.

BECKLIN, Katie M. et al. Examining plant physiological responses to climate change through an evolutionary lens. **Plant physiology**, v. 172, n. 2, p. 635-649, 2016.

BONCAN, D. A. T. et al. Terpenes and Terpenoids in Plants: Interactions with Environment and Insects. **International Journal of Molecular Sciences**, v. 21, n. 19, p. 7382, 2020.

BOWLING, L. C. et al. Agricultural impacts of climate change in Indiana and potential adaptations. **Climatic Change**, p. 1-23, 2020.

BUSINESS STANDARD. Castor seed production estimated to grow by 88% in 2019-20: Agriwatch. 2020. Disponível em <https://www.business-standard.com/article/economy-policy/castor-seed-production-estimated-to-grow-by-88-in-2019-20-agriwatch-120022200768_1.html>. Acesso em 15 de ago. de 2020

DANNER, Holger et al. Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of *Populus trichocarpa*. **Phytochemistry**, v. 72, n. 9, p. 897-908, 2011.

CASTRO, C. O programa nacional de produção e uso do biodiesel (PNPB) e a produção de matéria-prima de óleo vegetal no Norte e no Nordeste. Rio de Janeiro: IPEA. 2011.

COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB). Acompanhamento da safra brasileira - Cana-de-Açúcar. p. 76, 2018

CHANG, H.-Y.; CHENG, T.-H.; WANG, A. H.-J. Structure, catalysis, and inhibition mechanism of prenyltransferase. **IUBMB life**, v. 73, n. 1, p. 40–63, 2021.

CHRISTIANSON, D. W. Correction to structural and chemical biology of terpenoid cyclases. **Chemical reviews**, v. 118, n. 24, p. 11795-11795, 2018.

CHRISTENHUSZ, Maarten JM; BYNG, James W. The number of known plants species in the world and its annual increase. **Phytotaxa**, v. 261, n. 3, p. 201-217, 2016.

CHEN, X. et al. A terpene synthase-cytochrome P450 cluster in *Dictyostelium discoideum* produces a novel trisnorsesquiterpene. **Elife**, v. 8, p. e44352, 2019.

CHEN, H. et al. Combinatorial evolution of a terpene synthase gene cluster explains terpene variations in *Oryza*. **Plant Physiology**, v. 182, n. 1, p. 480-492, 2020.

CUI, G. et al. Functional divergence of diterpene syntheses in the medicinal plant *Salvia miltiorrhiza*. **Plant physiology**, v. 169, n. 3, p. 1607-1618, 2015.

DA SILVA C., A; BATALHA, M. O. Biodiesel production from castor oil in Brazil: A difficult reality. **Energy Policy**, v. 38, n. 8, p. 4031-4039, 2010.

FEI, S. et al. Divergence of species responses to climate change. **Science advances**, v. 3, n. 5, p. e1603055, 2017.

FIORAVANTI, C.. A maior diversidade de plantas do mundo. Pesquisa Fapesp, n. 241, p. 42-47, 2016.

Flora do Brasil 2020. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/>>. Acesso em: 08 abr. 2021

HAJAR, M.; VAHABZADEH, F. Artificial neural network modeling of biolubricant production using Novozym 435 and castor oil substrate. **Industrial Crops and Products**, v. 52, p. 430-438, 2014.

HAN, B. et al. 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, n. 2, p. 562, 2020.

HAY, Rebecca. Investigation of the role of gene clusters in terpene biosynthesis in *Sorghum bicolor*. 2018.

HENRY, L. K. et al. Contribution of isopentenyl phosphate to plant terpenoid metabolism. **Nature plants**, v. 4, n. 9, p. 721-729, 2018.

IOVINELLA, I. et al. Semiochemicals for intraspecific communication of the fig weevil *Aclees sp. cf. foveatus* (Coleoptera: Curculionidae): a first survey. **Scientific reports**, v. 10, n. 1, p. 1-8, 2020.

GROSS, H.; KÖNIG, G. M. Terpenoids from marine organisms: unique structures and their pharmacological potential. **Phytochemistry Reviews**, v. 5, n. 1, p. 115-141, 2006.

GUO, D. et al. Transcriptome profiling of spike provides expression features of genes related to terpene biosynthesis in lavender. **Scientific reports**, v. 10, n. 1, p. 6933, 2020.

KALB, S. R. et al. Recommended mass spectrometry-based strategies to identify ricin-containing samples. **Toxins**, v. 7, n. 12, p. 4881-4894, 2015.

KARUNANITHI, P. S.; ZERBE, P. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. **Frontiers in plant science**, v. 10, p. 1166, 2019.

KAUR, R.; BHASKAR, T. Potential of castor plant (*Ricinus communis*) for production of biofuels, chemicals, and value-added products. In: **Waste Biorefinery**. Elsevier, 2020. p. 269-310.

KEW SCIENCES. Plants of the World *Online*. 2020. Disponível em: <<http://www.plantsoftheworldonline.org/taxon/urn:lsid:ipni.org:names:355498-1#image-gallery>> Acesso em 15 de ago. de 2020.

KOPCSAYOVÁ, D.; VRANOVÁ, E. Functional Gene Network of Prenyltransferases in *Arabidopsis thaliana*. **Molecules** (Basel, Switzerland), v. 24, n. 24, p. 4556, 2019.

KULKARNI, M. G.; SAWANT, S. B. Some physical properties of castor oil esters and hydrogenated castor oil esters. **European journal of lipid science and technology**, v. 105, n. 5, p. 214-218, 2003.

KWON, O. et al. Cytotoxic scalarane sesterterpenes from the sponge *Hyrtios erectus*. **Marine drugs**, v. 18, n. 5, p. 253, 2020.

LANCASTER, J. et al. De novo formation of an aggregation pheromone precursor by an isoprenyl diphosphate synthase-related terpene synthase in the harlequin bug. **Proceedings of the National Academy of Sciences**, v. 115, n. 37, p. E8634-E8641, 2018.

LAL, P.; ALAVALAPATI, J.; MERCER, D. E. Socioeconomic impacts of climate change on rural communities in the United States. In: Alig, Ralph J., tech. coord. 2011. Effects of climate change on natural resources and communities: a compendium of briefing papers. Gen. Tech. Rep. PNWGTR-837. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 73-118., p. 73-118, 2011.

LEAVELL et al. Developing fermentative terpenoid production for commercial usage. **Current opinion in biotechnology**, v. 37, p. 114-119, 2016.

LESHIN, J. et al. Characterization of ricin toxin family members from *Ricinus communis*. **Toxicon**, v. 55, n. 2-3, p. 658-661, 2010.

LIANG, P. H.; KO, T.-P.; WANG, A. H.-J. Structure, mechanism and function of prenyltransferases: Structure, mechanism and function of prenyltransferases. **European journal of biochemistry**, v. 269, n. 14, p. 3339–3354, 2002.

MARTÍNEZ-ESPINOSA, R. M. Impact of the “omics sciences” in medicine: New Era for integrative medicine. **Journal of Clinical Microbiology and Biochemical Technology**, v. 3, n. 1, p. 009–013, 2017.

MARTIN, Diane M. et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. **BMC plant biology**, v. 10, n. 1, p. 1-22, 2010.

MASSETTI, E. et al. The impact of climate change on the brazilian agriculture: A ricardian study at microregion level. The Impact of Climate Change on the Brazilian Agriculture: A Ricardian Study at Microregion Level (December 2013). **CMCC Research Paper**, n. 200, 2013.

MATSUMI, R. et al. Isoprenoid biosynthesis in Archaea–biochemical and evolutionary implications. **Research in microbiology**, v. 162, n. 1, p. 39-52, 2011.

MOHAMMAD, T. H. et al. Bio-oil from pyrolysis of castor shell. **Int J Basic Appl Sci**, v. 14, p. 1-5, 2014.

MOSHKIN, V. A. History and origin of castor. Castor. **Oxonian Press Pvt. Ltd.**, New Delhi, p. 6-10, 1986.

NAIK, B. Botanical Descriptions of Castor Bean. In: **The Castor Bean Genome**. Springer, Cham, 2018. p. 1-14.

NÓBREGA, M.B. de M. Avaliação de genótipos de mamona (*Ricinus communis*) em cruzamentos dialélicos parciais. 2008. 77f. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, SP, 2008.

RIBEIRO, P. R. et al. Transcriptome profiling of *Ricinus communis* provides new insights underlying the mechanisms towards thermotolerance during seed imbibition and germination. **Industrial Crops and Products**, v. 126, p. 380-393, 2018.

RODRÍGUEZ-CONCEPCIÓN, M.; BORONAT, A. Breaking new ground in the regulation of the early steps of plant isoprenoid biosynthesis. **Current opinion in plant biology**, v. 25, p. 17–22, 2015.

PATEL, V. R. et al. Castor oil: properties, uses, and optimization of processing parameters in commercial production. *Lipid insights*, v. 9, p. LPI. S40233, 2016.

PIRES, V. B. L.; LOURENÇO, L. C. B. Biodiesel e inclusão social no Nordeste. **Revista de Política Agrícola**, v. 24, n. 3, p. 32-42, 2015.

OGUNNIYI, D. S. Castor oil: A vital industrial raw material. *Bioresource Technology*, v. 97, n. 9, p. 1086–1091, 2006. PENFIELD

SCHMIDT-DANNERT, C. Biosynthesis of terpenoid natural products in fungi. **Biotechnology of Isoprenoids**, p. 19-61, 2014.

SEVERINO, L. S. et al. A review on the challenges for increased production of castor. **Agronomy journal**, v. 104, n. 4, p. 853-880, 2012.

SINGH, B.; SHARMA, R. A. Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. 3 **Biotech**, v. 5, n. 2, p. 129-151, 2015.

SINGH, A. Tools for metabolomics. **Nature methods**, v. 17, n. 1, p. 24, 2020.

SOUZA, F. V. Expressão de genes em resposta a estresse por restrição hídrica em sementes de *Ricinus communis* (Euphorbiaceae). 2012. Dissertação. (Mestrado em Biotecnologia) Universidade Federal da Bahia, Salvador, Bahia, 2013.

TAIZ, L. et al. Fisiologia e desenvolvimento vegetal. **Artmed Editora**, 2017.

TELES, C. A. S. Aspectos fisiológicos e eventos do ciclo celular em sementes de *Ricinus communis* sob restrição hídrica. 2013. 92 f. Dissertação (Mestrado em Biotecnologia) Universidade Federal da Bahia, Salvador, Bahia, 2013.

THOLL, D. Biosynthesis and biological functions of terpenoids in plants. **Biotechnology of isoprenoids**, p. 63-106, 2015.

WANG, Y. et al. Salt-adaptive strategies in oil seed crop *Ricinus communis* early seedlings (cotyledon vs. true leaf) revealed from proteomics analysis. **Ecotoxicology and Environmental Safety**, v. 171, p. 12-25, 2019.

WANG, Yue et al. Gene structure, expression pattern and interaction of Nuclear Factor-Y family in castor bean (*Ricinus communis*). **Planta**, v. 247, n. 3, p. 559-572, 2018.

WORBS, S. et al. *Ricinus communis* intoxications in human and veterinary medicine—a summary of real cases. **Toxins**, v. 3, n. 10, p. 1332-1372, 2011.

WORBS, S. et al. Characterization of ricin and *R. communis* agglutinin reference materials. **Toxins**, v. 7, n. 12, p. 4906-4934, 2015.

YAMADA, Yuuki; CANE, David E.; IKEDA, Haruo. Diversity and analysis of bacterial terpene synthases. **Methods in enzymology**, v. 515, p. 123-162, 2012

YU, A.; LI, F.; LIU, A. Comparative proteomic and transcriptomic analyses provide new insight into the formation of seed size in castor bean. **BMC plant biology**, v. 20, n. 1, p. 48, 2020.

ZAPPI, D. C. et al. Growing knowledge: an overview of seed plant diversity in Brazil. *Rodriguésia*, v. 66, n. 4, p. 1085-1113, 2015.

ZHAO, R. et al. Volatile terpenes and terpenoids from workers and queens of *Monomorium chinense* (Hymenoptera: Formicidae). **Molecules**, v. 23, n. 11, p. 2838, 2018.

ZHU, J. Abiotic stress signaling and responses in plants. **Cell**, v. 167, n. 2, p. 313-324, 2016.

CAPÍTULO II

EFEITO DOS ESTRESSES ABIÓTICOS HIDRICO, SALINO E OSMÓTICO NO ESTABELECIMENTO DE PLÂNTULAS JOVENS DE *Ricinus Communis L.*



1 INTRODUÇÃO

Ricinus communis (Euphorbiaceae), conhecida popularmente no Brasil como mamona ou carrapateira, rícino, e internacionalmente como Castor Bean ou Castor Oil plant é uma cultura vegetal oleaginosa não comestível com destacada importância econômica. Outras como a mandioca (*Manihot esculenta* Crantz) a seringueira (*Hevea brasiliensis*), o pinhão-manso (*Jatopha curcas* L.), a soja (*Glycine max* L.) e a palma de óleo (*Elaeis guineenses*), a mamona se destaca em pesquisas na produção de derivados graxos para insumos industriais do ramo farmacêutico e cosmético, bem como biodiesel (FATHIMA, et al., 2022; MARIOD e TAHIR, 2022; ABHILASH et al., 2011; DIENG, et al., 2014; KANAI, et al., 2019; PINZI e DORADO, 2011).

De importância econômica para os pequenos agricultores do semiárido do nordeste brasileiro, a mamona é suscetível a uma variedade de fatores estressantes tanto de natureza biótica como aqueles causados por organismos vivos, como microrganismos, insetos e plantas invasoras, quanto abióticos tais como seca ou crise hídrica, aumento da salinidade no solo, aumento de incidência de UV, pobreza nutricional e temperaturas elevadas (HAN et al., 2020; LIU et al., 2019) Apesar de ser comprovadamente conhecida como sua alta tolerância sob tais condições adversas, a exposição a condições estressantes que podem agir isoladamente ou em conjunto contribui para a baixa produção das safras (DE BRITO et al., 2016; GOMES NETO et al., 2018; MUNNS et al., 2019).

O estresse hídrico é considerado uma das principais causas de desordens na homeostase celular em plantas, pois compromete diretamente na concentração de sais solúveis disponíveis no metabolismo celular, afetando por sua vez sua germinação de sementes e o estabelecimento de plântulas em níveis morfológicos, fisiológicos e metabólicos. Processos vitais como fotossíntese e regulação estomática são inteiramente ou parcialmente desregulados, afetando a concentração de CO₂ e O₂ no meio intra e extra celular. Tais alterações afetam o equilíbrio nutricional e hormonal do organismo provocando uma série de intoxicações por espécies químicas reativas de oxigênio (EROS), que comprometerão o processo de divisão celular e crescimento do organismo. Nos casos mais extremos, estes efeitos em cadeia ocasionam à morte do organismo (WANG et al., 2019; LIU et al., 2019; GOMES NETO et al., 2020)

Estudos já demonstraram que o estresse salino está entre as formas mais graves de estresses sofridos por organismos vegetais por causar lesão iônica – uma das principais lesões que afeta a membrana celular, ocasionando desordem no equilíbrio iônico celular entre os íons Na^+ , Cl^- , K^+ e Ca^{2+} , que ao se encontrarem dissociados em meio celular apresentam elevada toxicidade. Os tecidos radiculares das plantas são os primeiros órgãos a entrarem em contato com os nutrientes e assim selecionam os íons nutrientes, também protegem o organismo vegetal contra os íons tóxicos. O excesso desses íons nos tecidos celulares causa danos a membrana celular, e por consequência o comprometimento do desenvolvimento e morte do organismo (BITA et al., 2013; WANG et al., 2019; LIU et al., 2019; GOMES NETO et al., 2020)

Da mesma maneira, alterações significativas na homeostase celular são observadas a nível morfofisiológico quando as plantas são submetidas a elevadas temperaturas. Naturalmente, isto depende da suscetibilidade de cada espécie, porém de modo geral elevadas temperaturas afetam todos os estágios vegetativos e reprodutivos da planta. Pode-se notar lesões nas superfícies das folhas, caules, raízes e frutos, o que leva a diminuição da reprodutividade da espécie. São observadas também alterações significativas na estrutura celular, o que incluem organelas, citoesqueletos e funções membranares, além de alterações comprometedoras na síntese de proteínas, incluindo a tradução das proteínas de choque térmico (HSPs), produção de fitohormônios, como o ácido abscísico (ABA) e outras moléculas termoprotetoras, que visam o fechamento de estômatos e ativação de fatores de transcrição (FT) que auxiliam na adaptação da espécie as condições de estresse. As estruturas foliares e hipocotilares são atrofiadas, comprometendo a assimilação de líquido das partes aéreas e por consequência a distribuição total de biomassa e fotossíntese. (KUMAR, 2016; GOMES NETO et al., 2020) Todos os fatores de estresses ativam as espécies reativas de oxigênio (EROs), o que incluem radicais livres do tipo superóxidos como Hidroxil e Peróxidos de Hidrogênio que são produtos de sucessivas reações de redução proveniente do metabolismo aeróbico das plantas e que atuam como sinalizadores celulares para liberação de fitohormônios – mediadores de respostas celulares, como auxinas, ácido abscísico (ABA), ácido jasmônico (AJ) e etileno que também auxiliam na adaptação fisiológica da planta as condições estressoras (LIU et al., 2019; GOMES NETO et al., 2020; HAN et al., 2020).

Desta forma, diversas pesquisas são realizadas com o objetivo de contribuir para o entendimento dos mecanismos morfofisiológicos, bioquímicos e genéticos da *Ricinus communis* sob tais estresses a fim de propor estratégias para contornar os danos causados por

fatores de estresse. Tais danos podem ser visivelmente percebidos na germinação e no estabelecimento de plântulas saudáveis, velocidade e uniformidade de germinação bem como vigor ao serem submetidas a fatores com níveis estressores mínimos (SANTOS et al., 2013; RIBEIRO, 2015; SÁ, 2016; GOMES NETO et al., 2018; ARAUJO, 2019).

Estudos demonstram que os fatores de estresse já mencionados anteriormente afetam consideravelmente a distribuição de biomassa em todos os órgãos da planta. Em *Solanum lycopersicum*, por exemplo, o estudo realizado por Kim e colaboradores concluiu que a elevação da temperatura e exposição da espécie a incidências altas de luminosidades isoladamente ou combinados, levaram a diminuição da taxa fotossintética, diminuição da taxa de transpiração, diminuição total de biomassa dos frutos, menor espessura foliar. Houve também um aumento da biomassa das estruturas reprodutivas em detrimento da biomassa dos tecidos estruturais. Naturalmente estes danos são reflexos de desordens metabólicas da planta afetando a produtividade da cultura (KIM et al., 2019).

Em *Arabidopsis thaliana* o crescimento das plantas foi totalmente comprometido quando esta foi exposta a temperaturas mais altas. As plantas demonstraram redução da área foliar, redução de teor de água da parte aérea acompanhado do aumento de biomassa das raízes e comprimento das mesmas, o que possivelmente pode revelar um esforço para captação de água do solo. Alterações também foram observadas a nível metabólico e transcriptômico, como alta modulação gênica para terpenos devido a exposição a fatores de estresse e menor teor de água encontrado nos tecidos e baixa expressão de enzimas aquaporinas (CARRERA et al., 2018; JANSSON et al., 2018). Fenômenos similares ocorreu com plântulas de *Zea mays* (milho) cultivadas a 37°C e *Triticum aestivum* (Trigo), submetidas a estresse osmótico e salino, com pequenas diferenças na resposta metabólica dependente do estresse submetido (TRACHSEL et al., 2010; GUPTA et al., 2013; RIBEIRO et al., 2015; LIU et al., 2022)

Este estudo destina-se a avaliar o comprimento de raízes e hipocótilos e a distribuição de biomassa através da análise de massa seca em hipocótilos, endosperma, raízes e cotilédones de *Ricinus communis* submetidas a estresses abióticos do tipo térmico, salino e osmótico. Estes resultados nos permitirão correlacionar dados metabólicos e transcriptômicos aos dados morfológicos coletados com o objetivo de entender melhor os mecanismos de defesa da espécie estudada quando submetida a estresses abióticos desta natureza. Tais achados contribuirão também para o desenvolvimento de técnicas da

engenharia genética visando adaptação de plântulas a condições ambientais adversas em áreas semiáridas no mundo.

2.0 MATERIAIS E MÉTODOS

2.1 Tratamentos pré-germinativos em sementes de *Ricinus communis*

As sementes de *Ricinus communis* da cultivar Paraguaçu, foram gentilmente cedidas pelo Laboratório de Bioquímica, Biotecnologia e Bioprodutos (LBBB) no Instituto de Ciências da Saúde (ICS) da Universidade Federal da Bahia (UFBA). As mesmas foram compradas na Associação Olho D'água em Cafarnaum – Ba, em julho de 2019. Cafarnaum situa-se a 430 km de Salvador, a uma latitude de 11° 41' 37" sul e a uma longitude 41° 28' 06" oeste, com altitude de 770 metros no nordeste da Bahia, possuindo clima seco e árido. Foram cuidadosamente transportadas em saco de linho até o LBBB onde foram armazenadas sob temperatura controlada à 25 C°. Os ensaios experimentais foram executados no Laboratório de Bioquímica, Biotecnologia e Bioprodutos (LBBB) no Instituto de Ciências da Saúde (ICS) da Universidade Federal da Bahia (UFBA) nos períodos compreendidos entre outubro a dezembro de 2020.

O delineamento experimental foi inteiramente casualizado possuindo grupo controle. Todas as sementes utilizadas nos experimentos foram submetidas e desinfestação em solução de hipoclorito de sódio (0,25 %) e Tween-20 (1 gota para cada 100mL), indicado como a desinfestação mais eficaz por TELES (2013). As sementes foram postas em um béquer sob agitação constante durante 10 minutos e em seguida, foram enxaguadas 5 vezes com água destilada. Foram secas em papeis previamente autoclavados, após foram submetidas a escarificação mecânica e tiveram a carúncula retirada. Como substrato, foi utilizado papel de germinação (28 cm x 38 cm) umedecidos com água destilada na proporção de 3 vezes o peso do papel de germinação seco, seguindo as normas estabelecidas no Regras para Análise de Sementes (RAS) (BRASIL, 2009).

As sementes foram então dispostas no substrato de papel, com duas folhas na base e uma folha recobrando as sementes enfileiradas. Em seguida, os rolos foram postos para germinar envolvidas em sacos plásticos (evitando o ressecamento do substrato) em

germinadores do tipo B.O.D. (ELETROLAB, Mod. EL202/3). A temperatura foi ajustada para 30 °C, na ausência de luz e foram mantidas nesta condição por 72 horas. Foram utilizadas 3 repetições de 15 sementes cada. Sementes não germinadas ou deterioradas foram excluídas. Após a protrusão radicular, as plântulas foram expostas aos ensaios de estresse.

Todos os dados obtidos foram submetidos à Análise de Variância (ANOVA) com 5% de probabilidade, e as médias foram comparadas com o Teste de Tukey, utilizando o software de análise estatística SISVAR[®], desenvolvido pela Universidade Federal de Lavras (Ferreira, 2000).

2.2 Efeitos do estresse térmico sobre o desenvolvimento de plântulas

Após 72 horas sob germinação em água e com protrusão de radícula, as sementes de *Ricinus communis* foram transferidas para o estresse térmico sob as temperaturas de 20, 30 e 35 °C nas germinadoras tipo B.O.D. por um período de 96 horas. A temperatura de 30 °C foi escolhida como grupo controle, pois é a temperatura considerada ótima para o desenvolvimento da espécie. Os rolos de substrato foram irrigados com água destilada a cada 24 horas durante todo o experimento. Foram utilizados 5 ml para as temperaturas de 20 e 30 °C e 10 ml para a temperatura de 35 °C, para análise apenas do efeito da alta temperatura.

2.3 Efeitos do estresse salino sobre o desenvolvimento de plântulas

Após 72 horas sob germinação em água e com protrusão de radícula, as sementes de mamona foram transferidas para o estresse salino sob as concentrações de 0,0 mM (grupo controle); 25 mM; 50 mM e 100 mM de Cloreto de sódio (NaCl), sendo tais condições adotadas com base nas condições experimentais encontradas na literatura (WANG; VINOCUR; ALTMAN, 2003; WANG, et al. 2019; BRITO, et al. 2020). Para manter a concentração de sal constante durante todo o ensaio de estresse, a cada 48 horas foi realizada a troca completa do papel de germinação, utilizando as mesmas concentrações aplicadas inicialmente.

2.4 Efeitos do estresse osmótico sobre o desenvolvimento de plântulas

Para o ensaio de estresse osmótico foi utilizado o Polietilenoglicol 6000 (PEG 600; CAS: 25322-68-3) da fabricante Dinâmica, sendo as soluções preparadas seguindo as orientações de VILLELA et al.,1991). Após 72 horas sob germinação em água e com protrusão de radícula, as sementes de mamona foram expostas a condições de 0,0 Mpa (grupo controle); -0,2 MPa; -0,6 MPa e -1 MPa. Estes potenciais são amplamente descritos na literatura para ensaios de estresse osmótico ou de osmocondicionamento em sementes de várias espécies, incluindo *Ricinus communis* (FANTI; PEREZ; ANDRADE, 2003; CHEEMA, N. et al., 2010; OLIVEIRA, 2010; VASCONCELOS, et al. 2018, NETO et al., 2020).

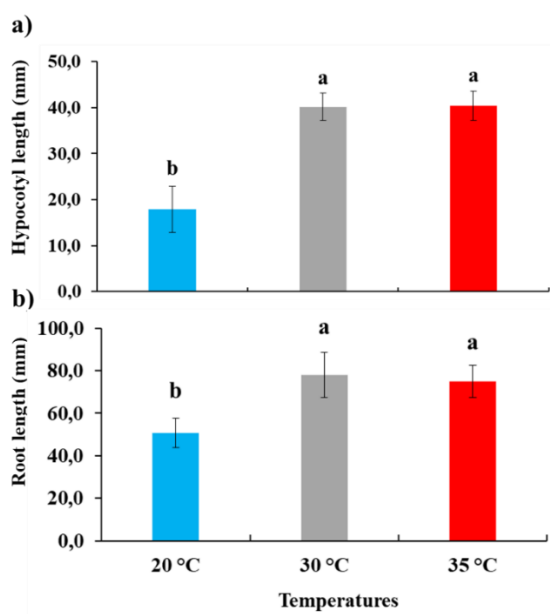
Para manter o potencial de PEG 6000 constantes durante todo o ensaio de estresse osmótico, a cada 48 horas foi realizada a troca completa do papel de germinação, utilizando os mesmos potenciais aplicados inicialmente no preparo do substrato.

3. RESULTADOS E DISCUSSÕES

Para analisar os efeitos da temperatura (20, 30 e 35 °C), concentração salina (25 mM; 50 mM; e 100 mM de NaCl) e déficit hídrico (-0,2 MPa; -0,6 MPa e -1 Mpa) em plântulas de *Ricinus communis* do cultivar Paraguaçu, foi medido o comprimento de raízes e hipocótilos e a massa seca das raízes, hipocótilos, endospermas e cotilédones após a exposição 72 horas de germinação e submetidas aos estresses térmico, salino e osmótico. Não foram observadas diferenças no comprimento das raízes e dos hipocótilos das plântulas submetidas a temperatura de 35°C em relação ao controle (30°C), porém uma diferença significativa foi observada no comprimento das raízes (cerca de 30 mm) e nos hipocótilos (cerca de 20) mm das plântulas submetidos à 20°C (**Figura 5**). As plântulas submetidas à temperatura de 20°C e 35°C não obtiveram desenvolvimento adequado das raízes, isso foi evidenciado pela perda de massa seca das raízes em comparação com o controle (30°) (**Figura 5**). Houve diferenças significativas na temperatura de 35 °C e 20°C. A massa seca do endosperma foi afetada quando submetidas à 20 e 35°C. Observou-se um acentuado aumento de biomassa dos cotilédones em 20°C e uma discreta acumulação de biomassa dos cotilédones em 35°C, quando comparadas com o controle (**Figura 6**). Um padrão de acumulação de biomassa dos

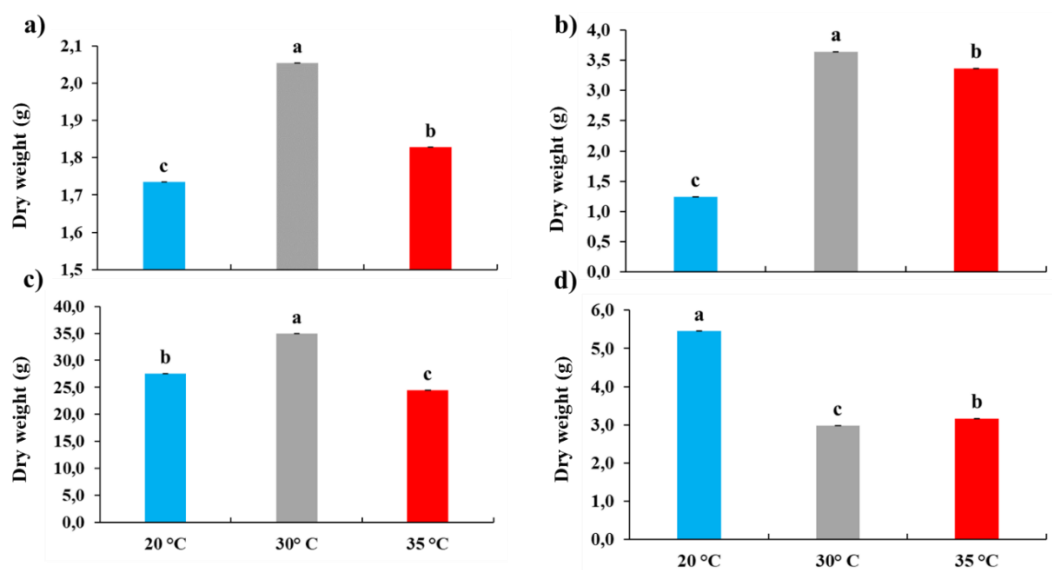
cotilédones diferente do genótipo Paraguaçu foi relatado para os genótipos IAC80 e MPA11 da mamona, onde relatou-se que a massa seca dos cotilédones aumentou com o aumento da temperatura para ambos os genótipos (RIBEIRO et al., 2015).

Figura 5: Comprimento de (a) hipocótilos e (b) raízes após quatro dias sob temperaturas de 20, 30 e 35 ° C. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras por HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

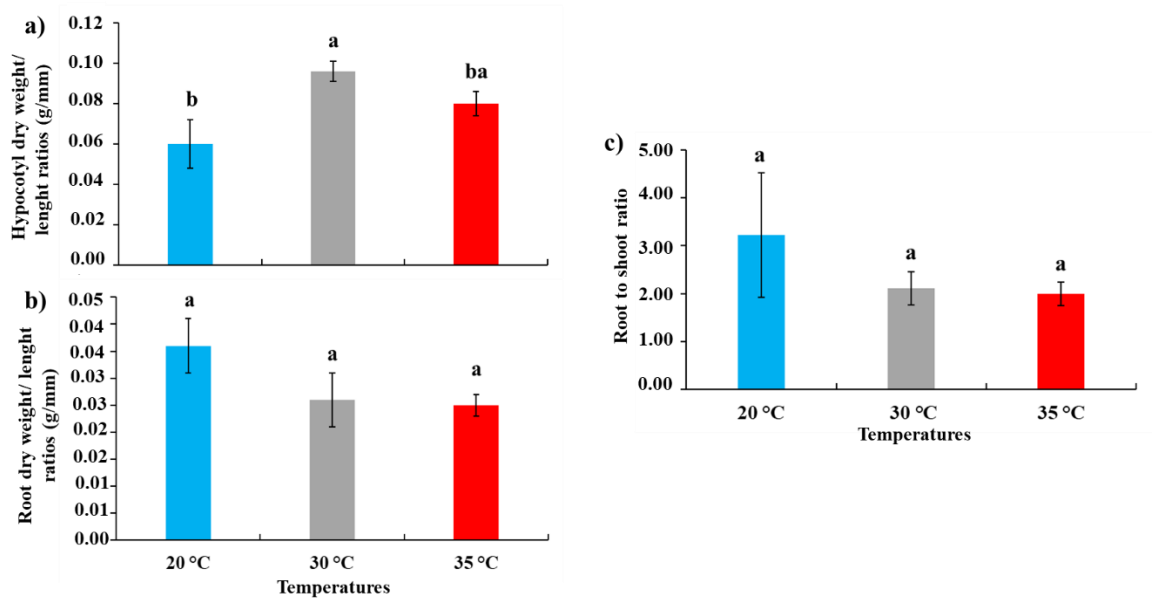
Figura 6: Massa seca em gramas de (a) raízes (b) hipocótilos, (c) endospermas e (d) cotilédones, após quatro dias sob as temperaturas de 20 °C, 30 °C e 35 °C. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras por HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

As raízes, folhas e caules das plantas possuem certa medida de independência funcional. Raízes, por exemplo, dependem exclusivamente das condições do solo, como disponibilidade de água e nutrientes. A parte aérea, depende em grande parte das condições acima do solo (níveis de CO₂, UV, etc). Para obter um panorama geral da distribuição de biomassa das raízes e hipocótilos em cada nível de estresse térmico, aplicamos a razão entre a massa seca das raízes pelo comprimento das raízes, também massa seca do hipocótilo pelo comprimento do hipocótilo, e a razão massa seca da raiz pela massa seca do hipocótilo (KRISHNAMURTHY et al., 1998; RIBEIRO et al; 2015; GOMES et al., 2018). Para os hipocótilos e raízes, não foi observada diferenças significativas entre os níveis de estresse térmico em comparação com o grupo controle (**Figura 7**). Estes resultados evidenciam um padrão diferente dos relatados na literatura acerca dos genótipos IAC80 e MPA11 da mamona. Para os genótipos MPA11, a razão aumentou com o aumento da temperatura e para o genótipo IAC80, relatou-se uma diminuição da razão (RIBEIRO et al., 2015). Dessa forma, conclui-se que possivelmente o genótipo Paraguaçu da *Ricinus communis* apresenta maior adaptabilidade quando expostas a níveis de estresse térmico de 20 e 35°C em comparação ao grupo controle em 30°C, considerando os critérios morfológicos descritos anteriormente.

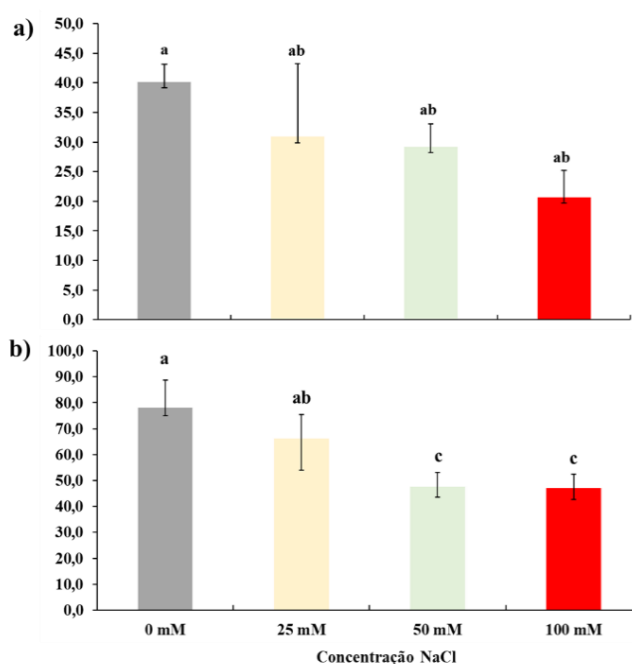
Figura 7: (a) Razão entre massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) Razão entre massa seca da raiz / comprimento da raiz (g / mm) e (c) razão entre a massa seca das raízes/ massa seca dos hipocótilos, após quatro dias sob temperaturas 20 ° C, 30 ° C e 35 ° C. Média e erros padrão de quatro repetições biológicas contendo 20 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey (p < 0,05).



Fonte: Dados obtidos do autor

Para avaliar os efeitos do estresse salino, as plantas foram submetidas às concentrações de 25, 50 e 100 mM de Cloreto de sódio (NaCl), após 72 horas de germinação em água. (**Figura 8**) Os resultados obtidos apontaram que houveram diferenças significativas no comprimento dos hipocótilos e no comprimento das raízes em comparação ao grupo controle (0 mM de NaCl). Estudos demonstram que a espécie apresenta certa resistência a estresse salino, possuindo até bons níveis de germinação, porém é na fase de crescimento e estabelecimento que as plantas são gravemente afetadas pelo estresse salino (PINHEIRO et al., 2008; GOMES et al., 2018).

Figura 8: Comprimento (mm) dos (a) hipocótilos e (b) raízes, após quatro dias sob concentrações de estresse salino de 0 mM, 25, 50 e 100 mM de NaCl. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

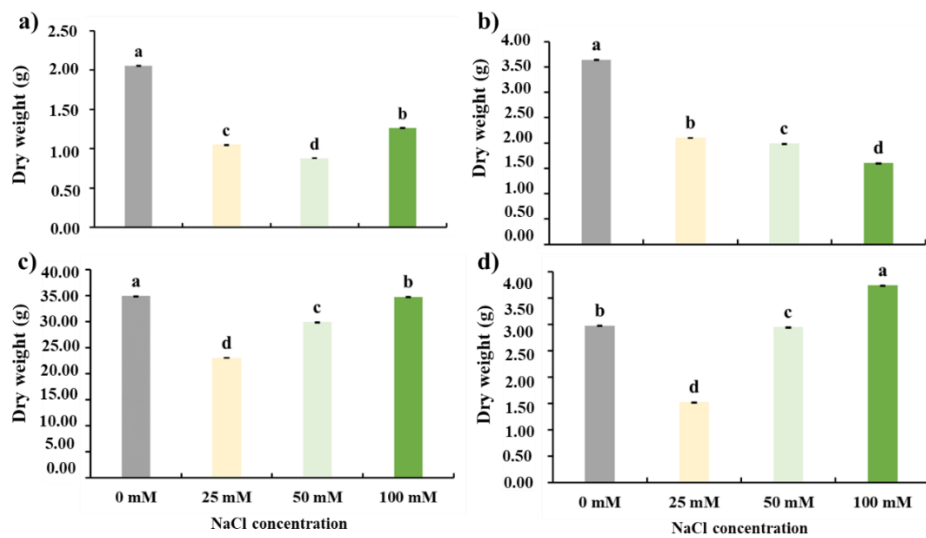
Resultados semelhantes também foram obtidos com os genótipos de mamona AIC 226 sob a estresses simulados com Cloreto de Potássio (KCl), apesar de não afetarem a germinação, a uniformidade e velocidade de germinação foram severamente afetados. O mesmo ocorre para *Jatropha curcas* L. e *Ricinus communis* (ANDRÉO-SOUZA et al., 2010; GOMES et al., 2018).

A massa seca das raízes demonstrou uma drástica diminuição nas concentrações de 25 mM e 50 mM, com moderada ascendência na concentração de 100 mM de NaCl, porém com diferença de 0,79 g, em relação ao grupo controle. Isto indica que no nível mais alto de estresse salino aplicado neste ensaio, a planta conseguiu desenvolver estratégias para preservar sua biomassa (CHEN et al., 2010; SANTOS et al., 2014). Os hipocótilos diminuíram sua biomassa à medida que se aumentava a concentração do sal, em relação ao controle. Os endospermas apresentaram padrões distintos de alocação de biomassa, sendo a concentração de sal mais estressora a de 25 mM, porém com maior alocação de biomassa em 50 mM e 100 mM, o que possivelmente indica uma adaptação mais rápida a concentrações mais elevadas do sal, em relação ao controle.

Os cotilédones também apresentaram padrões distintos de alocação de biomassa, considerando a menor alocação em 25 mM e o maior acúmulo de biomassa em 100 mM, curiosamente tendo nesse último nível de estresse salino uma biomassa maior que a do grupo

controle. Tanto para os endospermas como para cotilédones, observamos uma alocação de biomassa maior no nível mais alto de estresse salino (**Figura 9**). O mesmo fenômeno já foi relatado na literatura, onde o teor de umidade, massa seca e conteúdo relativo de água nos tecidos se mantiveram iguais ao controle até uma concentração de 150 mM de NaCl, indicando um certo ajustamento osmótico o que proporciona um acúmulo de soluto osmoprotetores da célula (CHEN et al., 2010; SANTOS et al., 2014).

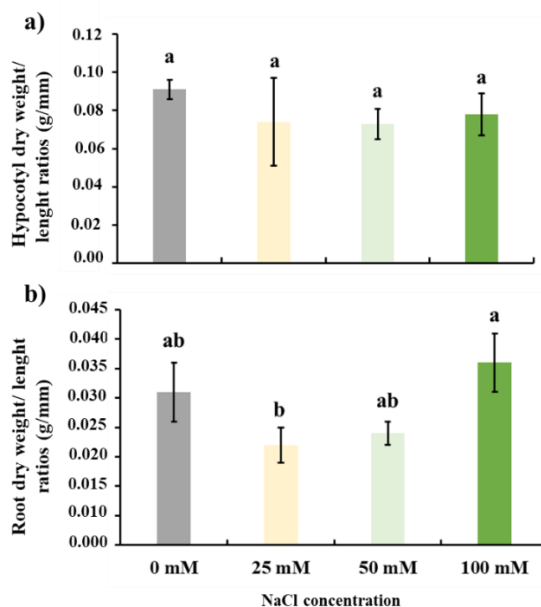
Figura 9: Massa seca em gramas das (a) raízes e (b) hipocótilos, (c) endospermas e (d) cotilédones, após quatro dias sob estresse salino nas concentrações de 0 mM, 25, 50, e 100 mM de NaCl. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

A razão massa seca da raiz pelo comprimento da raiz revelou não haver diferenças entre os níveis de estresse salino aplicados. Isto também se refletiu na razão massa seca dos hipocótilos pelo comprimento dos hipocótilos (**Figura 10**).

Figura 10: (a) Razão massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) Razão massa seca da raiz / comprimento da raiz (g / mm) após quatro dias de estresse salino sob concentrações de 0 mM, 25, 50 e 100 mM de de NaCl. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



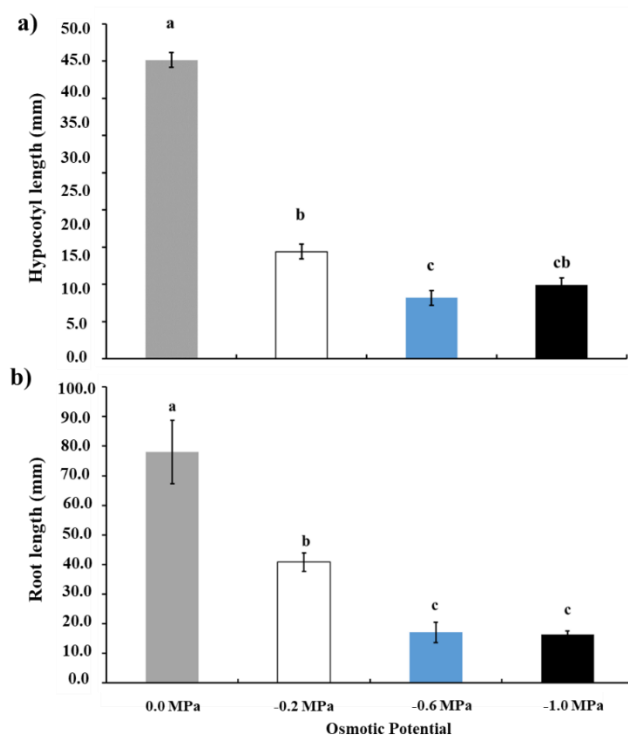
Fonte: Dados obtidos do autor

Os ensaios de estresse osmótico ou restrição hídrica foram simulados utilizando Polietilenoglicol (PEG 6000) sob os potenciais osmóticos de 0,0 Mpa; -0,2 MPa; -0,6 MPa e -1 Mpa. A restrição hídrica atua de forma mais danosa nos organismos, reduzindo drasticamente a velocidade dos processos biológicos e desregulando funções celulares importantes para o desenvolvimento das plantas. Os efeitos de uma restrição hídrica são relatados vastamente na literatura como menor desenvolvimento foliar, menor acúmulo de biomassa em todos os tecidos, desregulação fotossintética e respiração. Todas as enzimas que participam das reações bioquímicas, como as hidrolíticas ou digestivas são estabilizadas por moléculas de água, além disso, o transporte de íons e toda a maquinaria celular funcionam dependente de um equilíbrio sensível de vários íons em solução como os íons cálcio e potássio. Ao aumentar a concentração desses íons dispersos no ambiente intra e extracelular através da restrição hídrica (redução do solvente), instala-se uma série de reações em cadeia que afetam especialmente a regulação metabólica do organismo, por consequência o desenvolvimento saudável da espécie (SARA, 2016; ZHU, 2016).

Os resultados obtidos neste ensaio revelaram que em as plântulas submetidas ao osmocondicionamento foram severamente afetadas, o que comprometeu o desenvolvimento saudável das raízes e hipocótilos expressos através do comprimento destes. Observa-se uma

drástica redução do comprimento das estruturas submetidos a restrição hídrica em potenciais osmóticos de -0,2 Mpa e -0,6 e -1,0 Mpa, não possuindo diferenças significativas entre estas últimas (**Figura 11**). Resultados semelhantes também obtidos em mamona conforme relata estudos anteriores (MORAES et al., 2015; ARAUJO et al., 2016; GOMES et al., 2018).

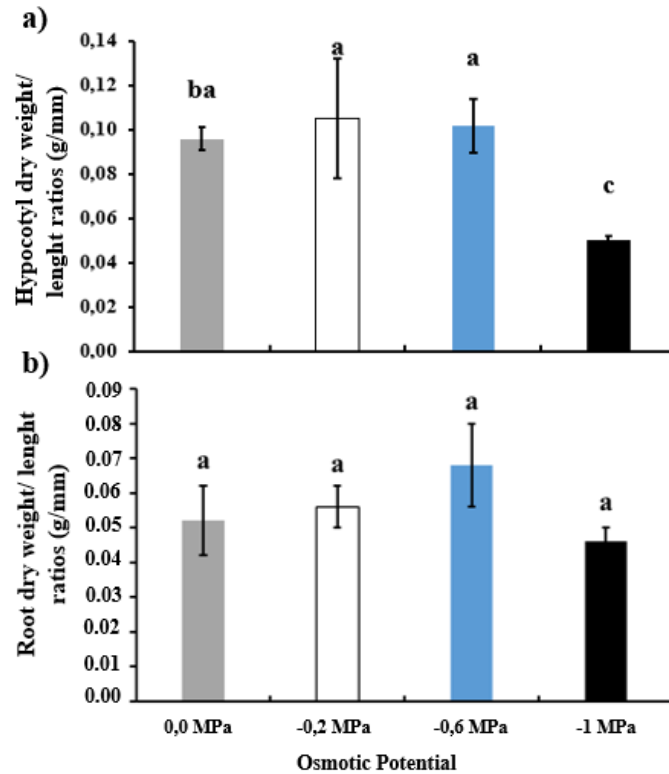
Figura 11: Comprimento de (a) hipocótilos e (b) raízes após quatro dias sob estresse osmótico sob os potenciais de 0, -0,2, -0,6 e -1,0 Mpa. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

A razão da massa da raiz pelo comprimento da raiz e massa do hipocótilo pelo comprimento do hipocótilo mostra que não há diferenças significativas entre as plântulas submetidas aos potenciais utilizados, exceto para o hipocótilo das plântulas submetidas ao estresse em -1 Mpa, no nível máximo de restrição hídrica, com comprometimento de todos do desenvolvimento da planta (**Figura 12**). Estudos anteriores indicaram que a restrição hídrica tem vários efeitos adversos sobre as plântulas, particularmente sob a biomassa, área foliar, taxa fotossintética líquida e diminuição do teor de clorofila. São relatadas também um aumento das atividades de enzimas antioxidantes e aumento da condutividade elétrica relativa. Alterações em todos os níveis bioquímicos também são relatados vastamente na literatura (SOUZA, 2013; ARAUJO et al., 2016; GOMES et al., 2018, YI et al., 2020)).

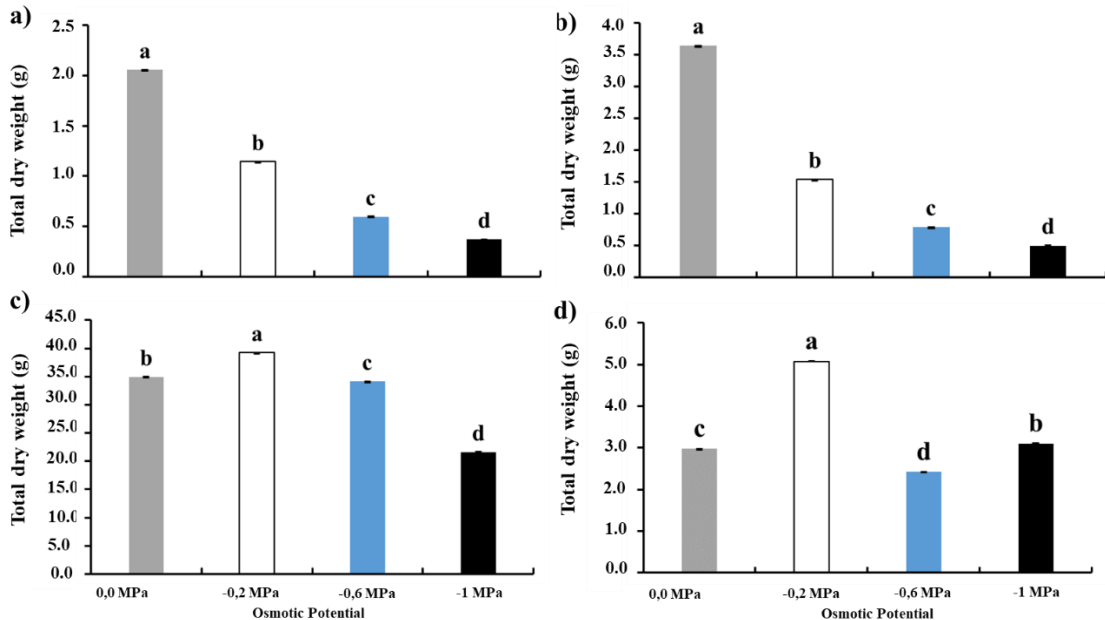
Figura 12: (a) Razão da massa seca do hipocótilo / comprimento do hipocótilo (g / mm), (b) razão da massa seca da raiz / comprimento da raiz (g / mm), após quatro dias sob os potenciais -0,2; -0,6 e -1 Mpa. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

A alocação de biomassa nas raízes e hipocótilos diminuiu à medida que diminuía o potencial osmótico, ou seja, à medida que se agravava o estresse hídrico. Um padrão distinto foi observado nos endospermas e cotilédones. A um potencial de -0,2 observou-se um acúmulo maior de biomassa em comparação com o do grupo controle, porém as plântulas expostas a níveis mais alto de estresse hídrico nos potenciais de -0,6 e -1, sofreram grande perda de biomassa. Nos cotilédones, o mesmo fenômeno foi encontrado, com acentuado aumento da biomassa no potencial -0,2 com perda de biomassa nas plântulas expostas ao potencial -0,6 e com uma alocação maior de biomassa no potencial -1 (**Figura 13**)

Figura 13: Massa seca em gramas das (a) raízes e (b) hipocótilos, (c) endospermas e (d) cotilédones após quatro dias sob estresse osmótico nos potenciais de -0,2; -0,6 e -1 MPa. Média e erro padrão de três repetições biológicas contendo 15 sementes cada. Letras diferentes acima das barras indicam diferenças significativas entre as amostras pelo HSD de Tukey ($p < 0,05$).



Fonte: Dados obtidos do autor

Isso possivelmente nos sugere uma translocação de recursos nutricionais para o cotilédone em um esforço fisiológico para auxiliar a planta a suportar a crise hídrica, ao menos por um dado período. Estas variações de biomassa e translocações para melhor adaptação as condições adversas são comuns e muito relatadas na literatura (ARAÚJO et al., 2016; GOMES et al., 2018, RIBEIRO et al., 2015)

4 CONSIDERAÇÕES FINAIS

Neste estudo observou-se que a exposição das plantas aos ensaios de estresse térmico (20, 30 e 35 °C), salino (25, 50 e 100 mM de Cloreto de sódio) e osmótico (-0,2; -0,6 e -1 Mpa em PEG 6000) forma geral comprometeu a alocação de biomassa nas raízes, hipocótilos, endosperma e cotilédones. Algumas exceções ocorreram, como no caso do acúmulo de biomassa dos endospermas e cotilédones quando expostos ao nível mais alto de estresse térmico e osmótico. Os comprimentos das raízes e dos hipocótilos evidenciaram um desvio da normalidade quando expostas ao estresse térmico em 20°C e não houveram diferenças significativas no comprimento das raízes e hipocótilos expostos à 35°C. Com estes achados,

podemos observar possíveis translocações de biomassa, transferências de recursos que darão suporte a planta para atravessar os períodos de estresse. Esses dados fisiológicos fornecem informações sobre o desenvolvimento da espécie *Ricinus communis* expostas a fatores estressores e fornecem subsídios teóricos no estudo da tolerância e adaptação à temperatura durante o estabelecimento. Além disso, as mudanças fisiológicas observadas em resposta a exposição a estresse térmico, salino e osmótico fornecem pistas para a adaptação destas a condições ambientais flutuantes, o que seria muito útil no desenvolvimento de estratégias para pesquisa de melhoramento de culturas de *R. communis*

REFERÊNCIAS

- ABHILASH, P. C. et al. Revisited *Jatropha curcas* as an oil plant of multiple benefits: critical research needs and prospects for the future. **Environmental Science and Pollution Research**, v. 18, n. 1, p. 127-131, 2011.
- ANDRÉO-SOUZA, Yara et al. Effect of salinity on physic nut (*Jatropha curcas* L.) seed germination and seedling initial growth. **Revista Brasileira de Sementes**, v. 32, p. 83-92, 2010.
- BAKOUMÉ, C.; WICKNESWARI, R.; SIJU, S.; RAJANAIDU, N.; KUSHAIRI, A.; BILLOTTE, N. Genetic diversity of the world's largest oil palm (*Elaeis guineensis* Jacq.) field genebank accessions using microsatellite markers. **Genetic Resources and Crop Evolution**, 62, n. 3, p. 349-360, 2014.
- BITA, Craita E.; GERATS, Tom. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. **Frontiers in plant science**, v. 4, p. 273, 2013.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Regras para análise de sementes. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Defesa Agropecuária. Brasília, DF: MAPA/ACS, 2009. p. 346.
- BRITO, et al. Overexpression of *Ricinus communis* .malate sintase enhances seed tolerance to abiotic stress during germination. **Industrial Crops and Products**, v. 145, p. 112110, 2020.
- CARRERA, Dániel Á. et al. Comparative proteomic analysis of plant acclimation to six different long-term environmental changes. **Plant and Cell Physiology**, v. 59, n. 3, p. 510-526, 2018.
- CHEEMA, N. et al. Influence of temperature and osmotic stress on germination induction of different castor bean cultivars. **Pakistan J. Bot**, v. 42, p. 4035-4041, 2010.
- DE BRITO, C. D. et al. Osmoconditioning prevents the onset of microtubular cytoskeleton and activation of cell cycle and is detrimental for germination of *Jatropha curcas* L. seeds. **Plant Biology**, v. 18, n. 6, p. 1053–1057, 2016

DIENG, Amadou et al. Effects of *Jatopha curcas* L. plantation on soil bacterial and fungal communities. **Soil Biology and Biochemistry**, v. 72, p. 105-115, 2014.

FANTI, S.; PEREZ, S. ANDRADE, C.; Efeito do estresse hídrico e envelhecimento precoce na viabilidade de sementes osmocondicionadas de paineira (*Chorisia speciosa*). **Pesquisa Agropecuária Brasileira**, v. 38, n. 4, p. 537-543, 2003.

FATHIMA, Anwar Aliya et al. Cassava (*Manihot esculenta*) dual use for food and bioenergy: A review. **Food and Energy Security**, p. e380, 2022.

FERREIRA, D. F. Sistema SISVAR para análises estatísticas: manual de orientação. Lavras: Universidade Federal de Lavras/Departamento de Ciências Exatas, 2000, 66p.

YI, Lita et al. Mechanisms of drought response in *Populus*. Southern Forests: **Journal of Forest Science**, v. 82, n. 4, p. 359-366, 2020.

JANSSON, Christer et al. Climate-smart crops with enhanced photosynthesis. **Journal of experimental botany**, v. 69, n. 16, p. 3801-3809, 2018.

GOMES NETO, V. et al. Characterization of the superoxide dismutase gene family in seeds of two *Ricinus communis* genotypes submitted to germination under water restriction conditions. **Environmental and Experimental Botany**, v. 155, p. 453–463, 1 nov. 2018.

GUPTA, N. K. et al. Effect of short-term heat stress on growth, physiology and antioxidative defence system in wheat seedlings. **Acta Physiologiae Plantarum**, v. 35, n. 6, p. 1837-1842, 2013.

KANAI, Masatake et al. Soybean (*Glycine max* L.) triacylglycerol lipase GmSDP1 regulates the quality and quantity of seed oil. **Scientific reports**, v. 9, n. 1, p. 1-10, 2019.

KIM, Hye-Ji; LIN, Meng-Yang; MITCHELL, Cary A. Light spectral and thermal properties govern biomass allocation in tomato through morphological and physiological changes. **Environmental and experimental botany**, v. 157, p. 228-240, 2019.

KRISHNAMURTHY, L. et al. Length to weight ratio of chickpea roots under progressively receding soil moisture conditions in a Vertisol. **Field Crops Research**, v. 58, n. 3, p. 177-185, 1998.

- LIU, Hao et al. Genome-wide analysis and functional characterization of CHYR gene family associated with abiotic stress tolerance in bread wheat (*Triticum aestivum* L.). **BMC Plant Biology**, v. 22, n. 1, p. 1-14, 2022.
- MARIOD, Abdalbasit Adam; TAHIR, Haroon Elrasheid. Biological activities of rubber (*Hevea brasiliensis*) oil. In: *Multiple Biological Activities of Unconventional Seed Oils*. **Academic Press**, 2022. p. 253-256.
- MORAES, Patrícia Favoretto et al. Expressão gênica diferencial em genótipos de mamona (*Ricinus communis*) submetidos a déficit hídrico induzido por PEG. **Bragantia**, v. 74, p. 25-32, 2015.
- MUNNS, R. Comparative physiology of salt and water stress. **Plant, Cell and Environment**, v. 25, n. 2, p. 239–250, fev. 2002.
- NETO et al. Sequence analysis of *Ricinus communis* small heat-shock protein (sHSP) subfamily and its role in abiotic stress responses. **Industrial Crops and Products**, v. 152, p. 112541, 2020.
- OLIVEIRA, A. Envelhecimento artificial, osmocondicionamento e estresse osmótico em sementes de sorgo: parâmetros fisiológicos, bioquímicos e Citoplasmaquímicos. 2010. Tese (Doutorado em Agronomia) Universidade Federal do Ceará, Fortaleza, 2010.
- PINHEIRO, H. A. et al. Leaf gas exchange, chloroplastic pigments and dry matter accumulation in castor bean (*Ricinus communis* L) seedlings subjected to salt stress conditions. **Industrial Crops and Products**, v. 27, n. 3, p. 385–392, 2008.
- PINZI, S.; DORADO, M. P. Vegetable-based feedstocks for biofuels production. In: **Handbook of Biofuels Production**. Woodhead Publishing, 2011. p. 61-94.
- RIBEIRO, Paulo R. et al. Effect of temperature on biomass allocation in seedlings of two contrasting genotypes of the oilseed crop *Ricinus communis*. **Journal of plant physiology**, v. 185, p. 31-39, 2015.
- RIBEIRO, P. R.; DE CASTRO, R. D.; FERNANDEZ, L. G. Chemical constituents of the oilseed crop *Ricinus communis* and their pharmacological activities: A review. **Industrial Crops and Products**, v. 91, p. 358–376, 2016.

SOUZA, Felipe Vieira de. Expressão de genes em resposta a estresse por restrição hídrica em sementes de *Ricinus communis*(Euphorbiaceae). 2013.

TELES, C. A. S. Aspectos fisiológicos e eventos do ciclo celular em sementes de *Ricinus communis* .sob restrição hídrica. 2013. 92 f. Dissertação (Mestrado em Biotecnologia) Universidade Federal da Bahia, Salvador, Bahia, 2013.

TRACHSEL, S. et al. Effect of high temperatures, drought and aluminum toxicity on root growth of tropical maize (*Zea mays* L.) seedlings. **Maydica**, v. 55, n. 3/4, p. 249-260, 2010.

VASCONCELOS, P.; et al., New insights into the mechanism underlying *Ricinus communis* .tolerance to drought stress during germination. **Industrial Crops and Products**, v. 103, p. 99-106, 2017.

VILLELA, F., et al. Tabela de Potencial Osmótico em Função da Concentração de Polietileno Glicol 6.000 e da Temperatura. Área de Informação da Sede-Artigo em periódico indexado (ALICE), 1991.

WANG, W.; VINOCUR, B.; ALTMAN, A. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. **Planta**, v. 218, n. 1, p. 1-14, 2003.

WANG, Yingnan et al. Physiological adaptive strategies of oil seed crop *Ricinus communis* early seedlings (cotyledon vs. true leaf) under salt and alkali stresses: From the growth, photosynthesis and chlorophyll fluorescence. **Frontiers in Plant Science**, v. 9, p. 1939, 2019.

WANG, Y. *et al.* Salt-adaptive strategies in oil seed crop *Ricinus communis*early seedlings (cotyledon vs. true leaf) revealed from proteomics analysis. **Ecotoxicology and Environmental Safety**, v. 171, p. 12-25, 2019.

ZHU, Jian-Kang. Abiotic stress signaling and responses in plants. **Cell**, v. 167, n. 2, p. 313-324, 2016.

CAPÍTULO III
CONTRIBUIÇÕES DA BIOINFORMÁTICA NO ESTUDO DAS TERPENO SINTASE
EM MAMONA (*Ricinus communis*L.)



RESUMO

Terpeno sintase (TPS) são responsáveis por uma vasta diversidade de produtos terpênicos em plantas, bactérias, fungos e organismos marinhos. As TPS são responsáveis pela produção de metabólitos altamente especializados que auxiliam na resposta contra estresses bióticos e abióticos, herbívoros, patógenos, sinalização ecológica bem como produção de hormônios reguladores da homeostase. Com o sequenciamento crescente de genoma de plantas, uma variedade de estudos com foco na caracterização das TPS e seus genes tem obtido destaque científico. Os mecanismos gênicos e bioquímicos tem sido costumeiramente elucidado e previsto por ferramentas computacionais com o auxílio da biologia molecular e outras ciências complementares. Neste contexto, uma extensa caracterização estrutural e filogenética das TPS é apresentada na oleaginosa *Ricinus communis*.L, utilizando ferramentas da bioinformática. Foram encontrados através do BLASp 46 genes TPS canônicos na espécie. A localização subcelular, distribuição de éxons, íntrons, regiões regulatórias transcricionais bem como elementos cis promotores são relatados. A designação de subfamílias (TPS-a – TPS-g) são apresentadas de acordo com a classificação de *motifs* conservados encontrados nas sequências e estrutura dos genes. Foram encontrados em todos os genes domínios C-terminal e N-terminal no Pfam característicos das enzimas terpenos sintase (PF01397 e PF03936). Além disso, foram encontrados nos genes padrões de *motifs* conservados, DDXXD, DXDD, NSE/DTE e RRX₈W. Sequências de TPS foram alinhadas frente outras espécies modelos e foi comprovado a presença do motif canônico DDXXD, essencial para a regulação enzimática que catalisa a primeira ionização em IPP e DMAPP. A identificação e caracterização estrutural e filogenética dos genes RcTPS fornecem uma rica compreensão da família TPS em *Ricinus communis* além de oportunizar futuras validação funcionais desta família gênica.

ABSTRACT

Terpene synthase (TPS) are responsible for a wide diversity of terpenic products in plants, bacteria, fungi, and marine organisms. TPS are responsible for the production of highly specialized metabolites that assist in the response against biotic and abiotic stresses, herbivores, pathogens, ecological signaling as well as the production of hormones that regulate homeostasis. With the growing sequencing of plant genomes, a variety of studies focusing on the characterization of TPS and their genes have gained scientific prominence. Genetic and biochemical mechanisms have been usually elucidated and predicted by computational tools with the help of molecular biology and other complementary sciences. In this context, an extensive structural and phylogenetic characterization of the TPS is presented in the oilseed *Ricinus communis*L, using bioinformatics tools. We found 46 canonical TPS genes in the species through BLASp. Subcellular location, exon distribution, introns, transcriptional regulatory regions as well as cis promoting elements are reported. The designation of subfamilies (TPS-a - TPS-g) are presented according to the classification of conserved *motifs* found in the sequences and structure of the genes. C-terminal and N-terminal domains in Pfan, characteristic of the terpene synthase enzymes (PF01397 and PF03936), were found in all genes. In addition, they were found in the conserved motif pattern genes, DDXXD, DXDD, NSE / DTE, and RRX8W. TPS sequences were aligned with other model species and the presence of the canonical motif DDXXD, essential for the enzymatic regulation that catalyzes the first ionization in IPP and DMAPP, has been proven. The identification and structural and phylogenetic characterization of the RcTPS genes provide a rich understanding of the TPS family in *Ricinus communis* .in addition to providing opportunities for future functional validation of this vast gene family.

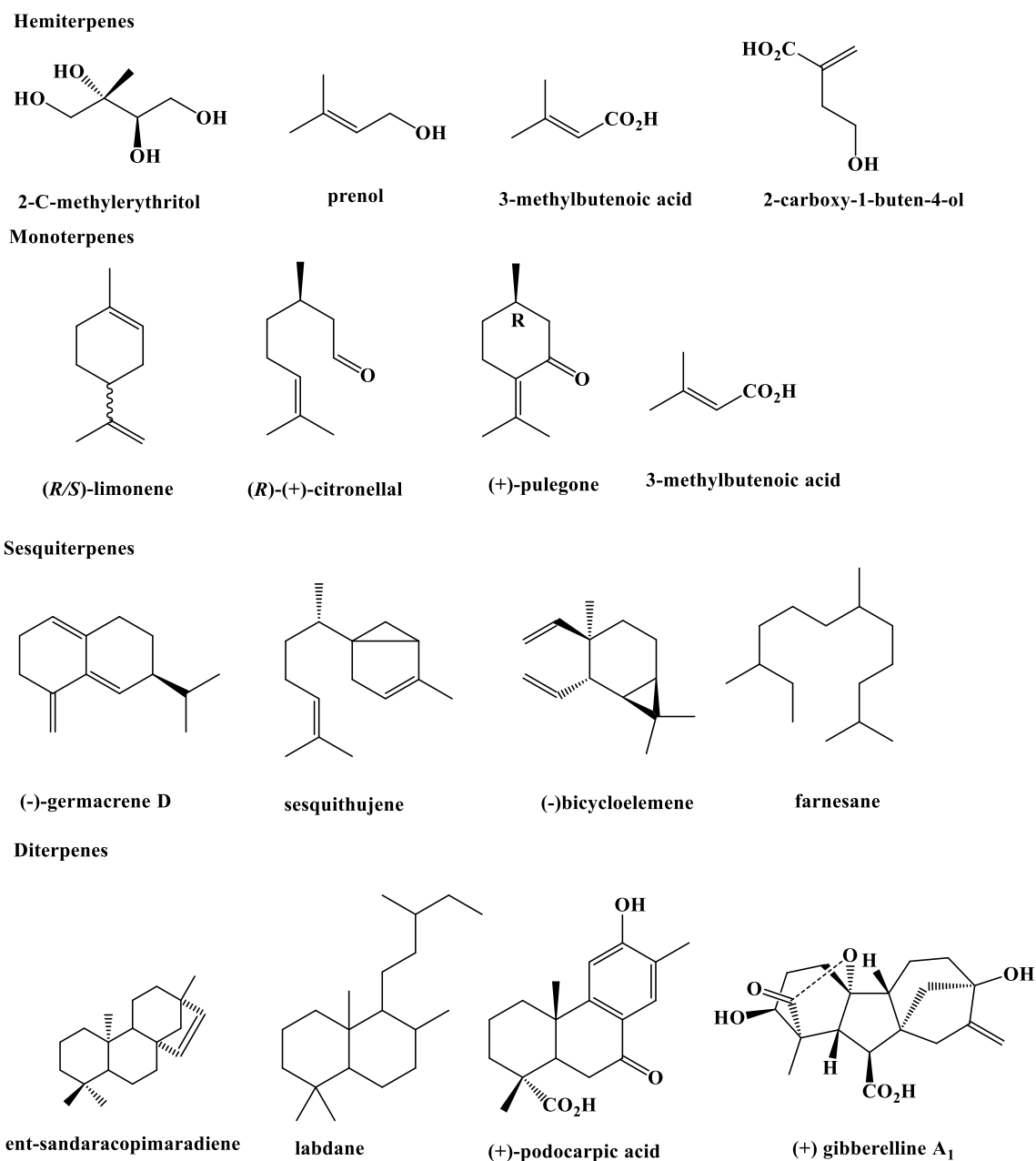
1 Introdução

As plantas apresentam a incrível capacidade de sintetizar diversos metabólitos secundários altamente especializados (Chen et al., 2011; Karunanithi e Zerbe, 2019; Luck et al., 2020). Entre esta rica produção natural, encontramos a maior classe de produtos naturais já registrada na literatura, os terpenos e seus derivados, os terpenóides. Sua heterogeneidade e elegância química se tornam quase onipresentes em todas as espécies de plantas (Hillwig et al., 2011; Nagel et al., 2014). Mais de 60 mil compostos terpênicos já foram identificados em plantas, organismos marinhos, insetos, bactérias e fungos (Gross e König, 2006; Yamada et al., 2012; Soltani et al., 2016; Lancaster et al., 2018; Karunanithi e Zerbe, 2019). Os terpenos são constituídos de um esqueleto básico de hidrocarbonetos, que unidos formam blocos de 5 carbonos (hemiterpenos), 10 carbonos (monoterpenos), 15 carbonos (sesquiterpenos), 20 carbonos (diterpenos), 30 carbonos (triterpenos), 40 carbonos (tetraterpenos) ou mais, C_{5n} (politerpenos). Estas espécies químicas podem assumir formas cíclicas, acíclicas além de apresentarem anéis fundidos. Podem ainda conter uma ou mais instaurações além de possuírem grupos carbonilados e possibilidade de estereoquímica (**Figura 14**) (Bohlmann et al., 1998; Christianson, 2006; Nieuwenhuizen et al., 2013).

Alguns terpenos e derivados possuem funções fisiológicas primárias, sendo designados como metabólitos primários, como a giberelina e o ácido abscísico. Estes auxiliam na formação de pigmentos como carotenoides, clorofilas e plastoquinonas na fotossíntese, outros atuam como hormônios vegetais, assessores na cadeia transportadora de elétrons e carboidratos além de estabilizadores de membranas (Tholl et al., 2011; Xie et al., 2012). Todavia, suas formas mais diversificadas foram encontradas em plantas como metabólitos secundários, atuando especialmente como agentes de comunicação com o ambiente, defesa a herbivoria e patógenos. Muitos destes compostos apresentam baixos pesos moleculares, portanto são voláteis sendo expelidos através das flores, frutos e tecidos (Dhar et al., 2013; Abbas et al., 2017). Terpenos e derivados também apresentam alto valor econômico, sendo largamente utilizados na produção de cosméticos, fragrâncias, óleos essenciais e especialmente medicamentos (Paduch et al., 2014; Bergman et al., 2019; Mahizan et al., 2019). Destes últimos, destacam-se o TAXOL^(R) que apresenta potente ação antineoplásica e a

arteminina, que apresenta excelentes atividades antimaláricas (Nagegowda e Gupta, 2020; Zhou and Pichersky, 2020b).

Figura 14: Exemplos de terpenos e derivados comumente isolados de plantas, bactérias, organismos marinhos dentre outros.



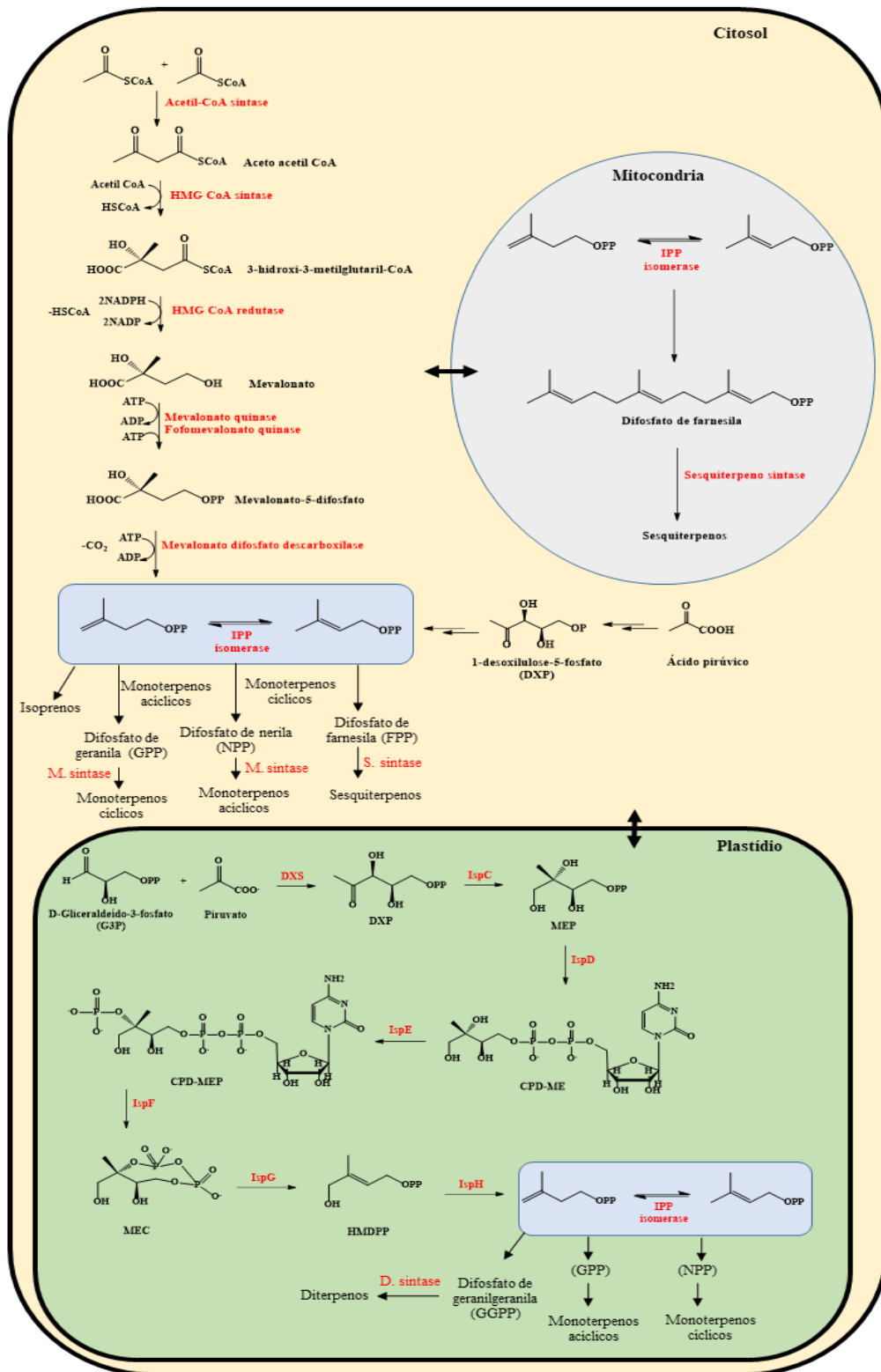
Fonte: ChemDraw, 2021

Todas as classes de terpenos, são biossintetizadas por enzimas TPS a partir de dois blocos fundamentais de unidades C5: Difosfato de isopentenila (IPP) e o Difosfato de

dimetilalila (DMAPP), ambos isômeros. Estes dois precursores atuam como os principais atores na biossíntese de todas as demais classes de terpenos e seus derivados, a partir de sucessivas reações de condensação, ciclização, óxido-redução, metilação, rearranjos, reações de eliminação, adição e substituição, bem como reações de natureza ácido-base. Estes precursores básicos são produzidos por duas vias distintas, a via do ácido mevalônico ou mevalonato (MEV) presente no citoplasma das plantas, animais e fungos e a via do D-eritritol-4 fosfato ou fosfato de metileritritol (MEP), operando nos plastídeos vegetais (Trapp e Croteau, 2001; Aubourg et al., 2002; Oldfield e Lin, 2012; Kumar et al., 2018; Alicandri, et al., 2020). Um grupo de enzimas especializadas em condensação, denominadas preniltransferases atuam na condensação inicial desses precursores (IPP e DMAPP) produzindo intermediários com diversos comprimentos de cadeias carbônicas. Os produtos da condensação são variados e nomeados de acordo com as unidades de C5 presentes na estrutura: Disfosfato de geranila (GPP; C10), difosfato de farnesila (FPP; C15), difosfato de geranilgeranila (GGPP; C20) e escaleno (C30). Os produtos GPP, FPP e GGPP sofrem sucessivas reações com outras unidades de IPP, resultando em sesquiterpenos, diterpenos, triterpenos e tetraterpenos (C40) (**Figura 15**).

O conjunto de enzimas terpeno sintase (TPS) atuam mediando essas transformações por estabilizar o carbocátion reativo no sítio ativo enzimático, o que posteriormente favorece a conversão em diversos compostos derivados. Para a formação dos hemiterpenos (C5), atuam as hemiterpenos sintase. Para os monoterpenos (C10), as monoterpenos sintase. Para os sesquiterpenos (C15), as sesquiterpenos sintase e para os diterpenos (C20), atuam as diterpenos sintase. Todas estas apresentam um sofisticado grau de especificidade, formando produtos especializados e adequados às necessidades fisiológicas do organismo frente às adversidades naturais durante o processo evolutivo (Bohlmann et al., 1998; Trapp e Croteau, 2001; Aubourg et al., 2002; Kumar et al., 2018).

Figura 15: Rotas bioquímicas com destaque na localização celular onde ocorre a formação dos precursores IPP e DMAPP - blocos básicos de construção de terpenos em plantas.



Fonte: Autor, adaptado de BONCAN et al., 2020.

Enzimas são representadas em vermelho. DXPS, 1-deoxy-d-xylulose-5-phosphate sintase; DXPRI DXP-reductoisomerase; MEPCT, MEP citidiltransferase; CDP-MEK, CDP-ME quinase; MECDPS, MECDP-sintase;

HMBPPS, (*E*) - 4-hidroxi-3-metilbut-2-enil-difosfato-sintase; HMBPPR, HMBPP redutase; IPP; isopentenil difosfato sintase (Karunanithi e Zerbe, 2019).

Apesar dos genes da família terpeno sintase (TPS) possuírem uma única origem evolutiva, eles podem ser subdivididos em duas categorias básicas – quanto a sua classe (I, II e III) e subfamília (TPS-a, TPS-b, TPS-c, TPS-d, TPS-e/f, TPS-g e TPS-h) (Chen et al., 2011; Jia, 2016). De acordo com o número, posição dos íntrons e comprimento de éxons identifica-se a classe dos genes TPS. Genes com 12 – 14 íntrons são designados a classe I, genes com 9 íntrons, classe II e genes com 6 íntrons pertencem à classe III (Bohlmann et al., 1998; Chen et al., 2011). *Motifs* conservados também são utilizados para designar as classes aos genes TPS. Nos genes de classe I são encontrados os padrões DDXXD ou NSE/DTE, geralmente encontrados em C-terminal funcionais (domínio α). Os genes de classe II, são encontrados os *motifs* DXDD relacionado ao N-terminal (domínio β) junto ao EDXXD (domínio γ) com elevada acidez o que auxilia na atividade catalítica das enzimas de classe II. Ainda são encontrados os *motifs* RRX₈W bem conservado localizado próximo ao peptídeo de transito N-terminal (Kumar et al., 2018).

As subfamílias são comumente agrupadas de acordo com a similaridade, número de éxons, classificação taxonômica e atividade enzimática. (Bohlmann et al., 1998; Nieuwenhuizen et al., 2013; Alquézar et al., 2017; Zhou et al. 2020). A subfamília TPS-a é subdividida em dois grupos que abarcam algumas sesquiterpenos sintase. As TPS-a-1 são distribuídas em plantas eudicotiledôneas e TPS-a-2 em monocotiledôneas. As TPS-b, englobam as monoterpenos sintase e isopreno sintase e estão amplamente distribuídas ente as angiospermas, seu papel influencia na síntese de monoterpenos cíclicos. As TPS-c possuem domínios codificadora característicos de difosfato de copalila (CPP) e *ent*-caureno (KS) bem como outras possíveis mono, sesqui e diterpenos. As TPS-c ocorre com menor distribuição em plantas terrestres, porém com incidência em todas elas. Enzimas TPS-c estão relacionadas a síntese de metabólitos primários, como a giberelina e o ácido adscísico. Enquanto as TPS-a, TPS-b e TPS-c se relacionam a síntese de metabólitos secundários como o citronelal e pulegona.

As TPS-d, possui ampla ocorrência em gimnospermas, e é subdividida em 3 categorias. TPS-d-1 que reúne especialmente as mototerpenos sintase e algumas sesquiterpenos sintase e TPS-d-2 que reúne apenas sesquiterpenos sintase. As TPS-d-3 que

além das sesquiterpenos sintase, reúne especialmente as diterpenos sintase. Em plantas vasculares, encontra-se alta incidência das TPS-e/f. Estas eram classificadas como dois cladogramas distintos, contudo foram convencionalmente fundidas devido as numerosas similaridades estruturais e funcionais entre si. As TPS-g, presentes especialmente nas angiospermas, agrupam as monoterpenos sintase, sesquiterpenos sintase e diterpenos sintase. Estudos recentes indicaram 18 membros TPS em *Selaginella moellendorffii* que apresentaram diterpenos sintase teoricamente bifuncionais que não pertencem às subfamílias TPS-c e TPS-e/f, formando um novo clado de TPS nesta espécie, a subfamília TPS-h (Chen et al., 2011; Li et al., 2012; Christianson, 2006).

Vários estudos demonstram a caracterização estrutural dos genes e proteínas TPS bem como seus respectivos produtos em diversos organismos vivos. Em plantas, as TPS já foram amplamente caracterizadas em *Arabidopsis thaliana* (Aubourg et al., 2002; Michael et al., 2020), *Solanum lycopersicum* (Spyropoulou et al., 2014; Zhou e Pichersky, 2020) *Oryza sativa* (Chen et al 2018; Kiryu et al 2018), *Sorghum bicolor* (Zhuang et al 2013), *Vitis vinifera* (Matarese, F. et al., 2014; Smit et al., 2020) e outras espécies, o que possibilitou descrever a família TPS em plantas como família de tamanho médio, com número de genes variando de 1 em *Physcomitrella patens* e 113 em *Eucalyptus grandis* (Butler et al., 2018).

Métodos modernos utilizam técnicas da bioinformática e biologia molecular com amparo das ciências ôhmicas no intuito de compreender melhor as características evolutivas e bioquímicas desta grande família gênica, além de propor alternativas para a obtenção dos seus produtos visando maior qualidade de vida humana em sociedade. Xie e colaboradores trabalharam na caracterização funcional de quatro sesquiterpenos sintase (nomeados de RcSeTPS1, RcSeTPS5, RcSeTPS7 e RcSeTPS10) oriundos da espécie *Ricinus communis* (Xie et al., 2012). Jackson e colaboradores utilizaram técnicas da biologia sintética aplicada na caracterização bioquímica de quatro enzimas KS (*ent*- caurene sintase) encontradas em *Ricinus communis*L (Jackson et al., 2014). Contudo, esta espécie carece de estudos com um panorama geral acerca da estrutura, biossíntese e atribuições funcionais das TPS. Este artigo, reflete o esforço de caracterizar *in silico* os genes da família terpeno sintase em *Ricinus communis*L., utilizando métodos modernos em bioinformática que contribuem para a riqueza do conhecimento da espécie.

2 Materiais e Métodos

2.1 Coleta de dados, classificação quanto a subfamílias e análise filogenética

Inicialmente, a região conservada de uma terpeno sintase em *Ricinus communis*, retirada do NCBI (<https://www.ncbi.nlm.nih.gov/>), foi usada como referência na consulta ao genoma da mesma espécie no banco de dados Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) com valor de corte e-value > e- 04. Com base na homologia, as sequências putativas de proteínas terpeno sintase em *Ricinus communis* (RcTPS) foram submetidas ao banco de dados Pfam (<http://pfam.xfam.org/search/sequence>) para a análise de domínios característicos das terpeno sintase (PF01397 e PF03936). Foi realizado a coleta das sequências genômicas, CDS e de aminoácidos, com seus respectivos códigos de identificação (ID).

Para estudar as relações filogenéticas, uma árvore filogenética não enraizada em seu centro foi construída usando o repertório completo de genes TPS identificados neste trabalho com genes da família TPS disponíveis na literatura das espécies *Arabidopsis thaliana* (Aubourg et al., 2002) e *Populus trichocarpa* (Irmisch, et al., 2014). Esta análise foi realizada utilizando as sequências de aminoácidos através do software MEGA 7.0 (<https://www.megasoftware.net/>) com alinhamento MUSCLE e *gaps* deletados. A topologia foi calculada através do método *Neighbor-Joining* com distâncias calculadas por p-distance. O *bootstrapping* foi ajustado para 1000 réplicas. As subfamílias de TPS (a-g) foram designadas de acordo a homologia das sequências das espécies comparativas neste estudo designadas pelo software Terzyme (<http://www.nipgr.ac.in/terzyme.html>) que utiliza fundamentos da *Machine Learning* através do algoritmo Modelo de Markov Ocultos de Perfil e Precisão Preditiva, com confiabilidade superior a 86% (Priya et al., 2016, 2018; Boncan et al., 2020). Foram retiradas da árvore filogenética as sequências RcTPS15, RcTPS27, RcTPS28, RcTPS29, RcTPS32, RcTPS35, RcTPS39, RcTPS43 e RcTPS45, por serem curtas demais a fim de serem comparadas com as demais sequências na árvore ou por se tratarem de sequências incompletas no banco de dados.

2.2 Localizações subcelulares, estrutura das proteínas e análise da região promotora das RcTPS

As localizações subcelulares das proteínas RcTPS foram previstas utilizando os softwares CELLO v.2.5 (<http://cello.life.nctu.edu.tw/>). Os respectivos pesos moleculares, ponto isoelétrico (pI), Instability index, Aliphatic index e GRAVY das proteínas RcTPS foram mensurados através do ProtParam (<https://web.expasy.org/protparam/>). Para analisar quais *motifs* são conservados, foi utilizado o software The MEME Suite 5.3.0 (<http://meme-suite.org/tools/meme>). O número máximo de *motifs* identificados foi ajustado para 10. Para a análise estrutural de íntrons, exons, upstream e downstream foi utilizado o software Gene Structure Display Server 2.0 (<http://gsds.gao-lab.org/>). Duplicações em tandem, foram reveladas pelo software PTGBase (<http://ocri-genomics.org/PTGBase/>) (Gomes et al., 2018, 2020). Para o alinhamento foi utilizado o software Uniprot (<https://www.uniprot.org/align/>), com método Clustal Omega de alinhamento. Uma sequência representativa das espécies *Arabidopsis thaliana*, *Solanum lycopersicum*, *Oryza sativa*, *Selaginella moellendorffii*, *Populus trichocarpa* e *Ricinus communis* foram utilizadas observando vestígios de similaridade entre aminoácidos e *motifs* conservados. Para a análise *in silico* dos elementos promotores regulatórios foi utilizado o banco de dados PlantCare (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Para isso, analisamos 1 kb upstream do códon de iniciação ATG. Os elementos cis foram classificados quanto a presença de regiões que regulam ou expressam ácido abscísico, auxina, giberelina, metil jasmonato (MeJa), respostas à luminosidade, MYB, MYC, indução anaeróbica, resposta de defesa ou estresses, expressão induzida por crise hídrica, baixas temperaturas, ácido salicílico, expressão de meristema, AT-rich, expressão de zeína, controle circadiano e respostas a feridas/herbivoria.

3 Resultados e discussões

3.1 Identificação e análise filogenética da família TPS em *Ricinus communis* (RcTPS).

As investigações acerca do quantitativo dos genes da família terpeno sintase em *Ricinus communis* utilizando as anotações disponíveis no banco de dados NCBI, revelaram

51 genes putativos para terpeno sintase em *Ricinus communis*. As pesquisas de similaridade frente ao proteoma da espécie no Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) e no Pfam (<http://pfam.xfam.org/search/sequence>), revelaram que apenas 46 genes possuíam domínios de terpeno sintase registrados no Pfam (Figura Suplementar S1). Portanto, estes foram identificados neste estudo como genes da família terpeno sintase em *Ricinus communis*, nomeados neste estudo como RcTPS1 – RcTPS46 (**Tabela suplementar S2**).

As relações observadas na árvore filogenética construída pelo MEGA 7.026 (<https://www.megasoftware.net/>) com outras angiospermas eudicotiledôneas, a saber *Populus trichocarpa* e *Arabidopsis thaliana*, mostraram que a família RcTPS (46 genes) é quantitativamente maior em *R. communis* L. (**Figura 16**). Possuem homologia com 32 genes de *Arabidopsis thaliana* (Aubourg *et al.*, 2002; Zhou and Pichersky, 2020a) e 38 genes de *Populus trichocarpa* (Irmisch *et al.*, 2014). A análise filogenética ainda revelou que as classificações das subfamílias (TPSa - TPSg) das RcTPS tendem ao agrupamento com as mesmas subfamílias de TPS das espécies TPS em *Populus trichocarpa* e *Arabidopsis thaliana*, evidenciado por vários ramos de uma única cor, todos agrupados, com algumas exceções (Potri.002G052100.1, Potri.011G031800.1 e AT3G29110.1). As subfamílias TPS-d e TPS-h não foram incluídas nesta análise, pois são subfamílias específicas para gimnospermas e *Selaginella moellendorffii* respectivamente. Em alguns casos, os genes RcTPS mostraram maior identidade com os genes ortólogos de *Populus trichocarpa* como nas sequências RcTPS44 – Potri.004G38400.1; RcTPS38 – Potri.005G210300.1 e RcTPS5 – Potri004G037900.1. Para a *Arabidopsis thaliana* a maior identidade com genes ortólogos se deu entre as sequências RcTPS16 – AT4G02780.

O número de genes RcTPS dispostos na árvore filogenética variou de 1 à 13 por subfamília (RcTPSa - RcTPSg), sendo 13 genes pertencentes a subfamília RcTPS-a, 10 genes pertencentes a subfamília RcTPS-b, apenas 1 gene pertencente a subfamília RcTPS-c, 3 genes pertencentes a subfamília RcTPS-e/f, e 4 genes pertencentes a subfamília RcTPS-g. Nove genes foram retirados da análise filogenética, pois suas sequências possivelmente são curtas demais para serem comparadas as outras nesta análise, ou por estarem incompletas no banco de dados. Destas, 6 genes pertencem a subfamília TPS-b (RcTPS15, RcTPS27, RcTPS28, RcTPS29, RcTPS32, RcTPS45) e 3 genes pertencem a subfamília TPS-a (RcTPS35,

RcTPS39, RcTPS43). Conforme dados disponíveis na literatura, a subfamília TPS-a é a mais numerosa dentre as terpeno sintases em pelo menos 7 espécies destacadas (**Tabela 2**).

Tabela 2: Número de espécies TPS em *R. communis* L. e em outras espécies já relatadas na literatura

Species	TPS Subfamily							Total	References
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>ef</i>	<i>g</i>	<i>h</i>		
<i>Ricinus communis</i>	16	16	1	0	9	4	0	46	This research
<i>Arabidopsis thaliana</i>	22	6	1	0	2	1	0	32	Aubourg et al. (2002)
<i>Solanum lycopersicum</i>	15	9	2	0	6	2	0	34	Zhou and Pichersky (2020)
<i>Oryza sativa</i>	18	0	3	0	9	2	0	32	Chen et al. (2014)
<i>Sorghum bicolor</i>	15	2	1	0	3	3	0	24	Paterson et al. (2009)
<i>Vitis vinifera</i>	30	19	2	0	1	17	0	69	Martin et al. (2010)
<i>Populus trichocarpa</i>	16	14	2	0	3	3	0	38	Irmisch et al. (2014)

De acordo a árvore filogenética (**Figura 3**), o PtTPS20 (Potri.001G415100) é o ortólogo mais próximo ao RcTPS13 (29842.m003696). Apesar de apresentarem alta similaridade – o que possivelmente sugere que esses genes conduzem a enzimas com funcionalidades semelhantes, pertencem a subfamílias diferentes, levando a dois produtos putativos diferentes – ambos sesquiterpenos (C15). O gene PtTPS20 já foi bem caracterizado, e catalisa a formação do sesquiterpeno germacreno D – um isômero da família dos germacrenos que apresenta atividades antimicrobianas e inseticidas. Já o RcTPS13 lhe é previsto a formação do nerolidol – um sesquiterpeno natural encontrado nos óleos essenciais de muitos tipos de plantas. Tanto o germacreno D quanto o nerolidol são moléculas participantes em processos sinalizadores na defesa induzida de plantas mediada pela atração de inimigos de herbívoros. O mesmo ocorre com os pares de genes PtTPS18 (Potri.015G032100) e RcTPS14 (29842.m003692), na formação dos compostos (-) germacreno D e alfa-farneseno, respectivamente. Ocorre também nos genes PtTPS36 (Potri.004G037900) e RcTPS5 (29910.m000944), na formação dos compostos (3*E*, 6*E*) - alfa farneseno e o diterpeno casbeno; nos genes PtTPS35 (Potri.005G210300) e RcTPS38 (29842.m003698), na formação dos produtos putativos *Ent*- kaureno e nerolidiol e nos genes PtTPS 16 (Potri.004G030400) e RcTPS44 (30156.m001715), na formação dos produtos putativos nerolidol e (*E*) – beta ocineno (Irmisch et al., 2014; Danner et al., 2011; Boncan et al., 2020).

Os genes PrTPS31 (Potri.019G045400) e RcTPS2 (29840.m000603) pertencem a subfamília TPS-a, responsável pela formação de vários monoterpenos incluindo mirceno,

limoneno, terpinoleno e linalol. A alta similaridade do gene RcTPS2 com o PtTPS31, possivelmente indica que estes genes possuem funções semelhantes, dando origem aos mesmos produtos. Isto ocorre também nos genes PtTPS7 (Potri.001G308200) e RcTPS1 (29771.m000073), na formação dos compostos putativos (-) alfa- terpineol e 1, 8 cineol, respectivamente – mediante ionização por hidrólise do grupo difosfato do GPP.

Porém, apesar de possuírem alta similaridade e mesma subfamília, os pares de genes PtTPS21 (Potri.011G142800) e RcTPS17 (28166.m001038), participam em processos de formação de dois compostos putativos diferentes – germacreno D e (+) -delta-cadineno, respectivamente. O mesmo ocorre com os pares de genes PtTPS18 (Potri.015G032100) e RcTPS14 (29842.m003692), na formação dos compostos (-) germacreno D e alfa-farneseno, respectivamente. O mesmo foi observado nos genes PtTPS1 (Potri.019G023000.1) e RcTPS 8 (30156.m001703), na formação dos compostos putativos (-) alfa- terpineol e beta-ocimeno, respectivamente, e nos genes PtTPS12 (Potri.017G041700) e RcTPS25 (30156.m001700) na formação do Isopreno e dos compostos putativos (+/-) - alfa-terpineol, respectivamente (Wiberley et al., 2008; Irmisch et al., 2014; Smit et al., 2020).

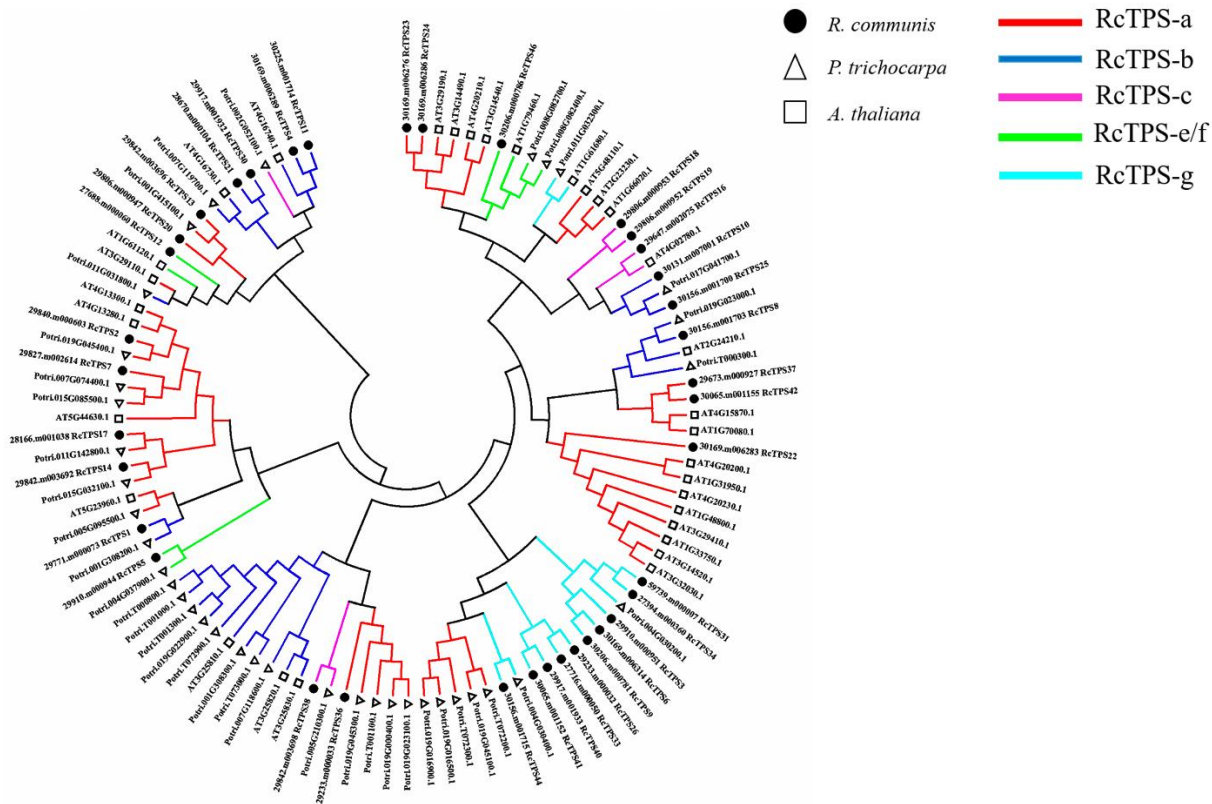
O gene RcTPS16 (29647.m002075) apresentou maior similaridade com o ortólogos de *Arabidopsis thaliana* – AtTPS31 (AT4G02780) ao qual catalisa a conversão de difosfato de geranilgeranila (GGPP) em difosfato de copalila (CPP) na biossíntese de giberelina. O fato de ambos pertencerem a mesma subfamília (TPS-c) e participam no processo de formação do intermediário Ent-kaurene e seus precursores catiônicos – catalisada pela enzima bifuncional *Ent*-kaurene sintetase A (5.5.1.13) mediante sucessivos rearranjos de carbocátions, demonstra que possivelmente os genes RcTPS16 e AtTPS31 possuem a mesma função. Estudos indicam que a liberação de *Ent*-kaurene em resposta a estresses abióticos em *Arabidopsis thaliana* auxiliam no processo de homeostase celular na regulação dos níveis de giberelina bioativos da espécie, além de auxiliar no crescimento e comunicação interespecies (Fleet et al., 2003; Otsuka et al., 2004; Jackson et al., 2014; Jia et al., 2017).

Estes achados indicam que a família de genes terpeno sintase encontra-se distribuída funcionalmente na espécie *Ricinus communis assim* como ocorre também nas espécies comparativas usadas neste estudo e outras disponíveis na literatura (Pazouki and Niinemets,

2016; Alquézar et al., 2017; Kumar, Y. et al., 2018; Martín-Sánchez et al., 2019; Zhou and Pichersky, 2020a; Yu et al., 2020) (**Figura 16**). Geralmente, a espécie *Ricinus communis* apresenta mais proteínas TPS do outras espécies de plantas. Para *A. thaliana* 32 TPS putativos já foram identificados. Para a espécie *Populus trichocarpa* já foram relatadas 38 TPS. Para as espécies *S. lycopersicum* e *S. bicolor* foram relatadas 34 e 24 TPS, respectivamente. O que possivelmente garantiu a estas espécies diferentes estratégias combinadas para a adaptação, especialmente a ambientes estressores (Tholl; Lee, 2011; Kumar et al., 2018; Zhou and Pichersky, 2020).

No decorrer do processo evolutivo, eventos naturais excluíram e/ou adicionaram determinados genes TPS da *Ricinus communis* aos quais produziram uma vasta diversidade de terpenos especializados, tais como monoterpenos, sesquiterpenos, diterpenos (voláteis e/ou não voláteis) que contribuíram com sua adaptação aos climas mais áridos e quentes, considerável tolerância a herbivoria, atração de insetos predadores de herbívoros, atração de polinizadores e mecanismos físico-químicos que serviram de barreiras contra a alimentação e oviposição de insetos, bem como outros mecanismos regulatórios da sobrevivência e adaptação desta espécie (Bohlmann et al., 1998; Keeling and Bohlmann *et al.*, 2006; Chen *et al.*, 2011; Zerbe *et al.*, 2015; Karunanithi and Zerbe *et al.*, 2019) (**Figura 16**).

Figura 16: Relações filogenéticas dos genes TPS em *Ricinus communis*, *Populus Trichocarpa* (IRMISCH et al., 2014) e *Arabidopsis thaliana* (Aubourg et al. 2002; Zhou e Pichersky, 2020). A classificação das subfamílias (a, b, c, e/f, g) baseia-se na classificação de terpeno sintase do software Terzyme. Os círculos preenchidos representam genes TPS da espécie *Ricinus communis*. Os quadrados, genes TPS de *Arabidopsis thaliana* e os triângulos genes TPS de *Populus Trichocarpa*. As cores dos ramos representam as diferentes subfamílias: TPS-a (vermelho), TPS-b (azul marinho), TPS-c (rosa), TPS-e/f (verde claro) e TPS-g (azul claro).



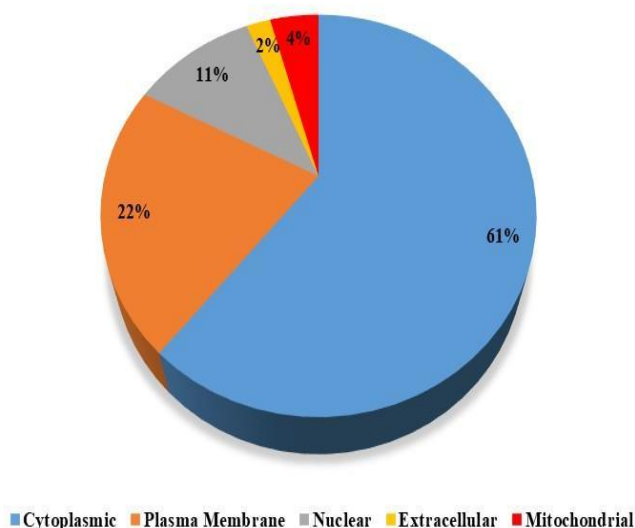
3.2 Análise da estrutura, localização subcelular e região promotora dos genes RcTPS

O peso molecular das RcTPS em *Ricinus communis* se diversificou consideravelmente entre 93570.55 Da (29806.m000947) de peso máximo à 10960.66 Da (29878.m000241) de peso mínimo. O comprimento de aminoácidos variou de 816 AA (29806.m000947) à 98 AA (29878.m000241), o número máximo semelhante a *Arabidopsis thaliana* (884 e 547 AA) (Aubourg et al., 2002). O ponto Isoelétrico (pI) também oscilou entre 9.32 (30206.m000785) à 4.60 (30156.m001715). Desta forma, o presente estudo demonstra que a maior parte das RcTPS possuem caráter moderadamente ácido, exceto as RcTPS37 (29673.m000927), RcTPS38 (29842.m003698), RcTPS40 (29917.m001933), RcTPS42 (30065.m001155), RcTPS43 (29878.m000241) e RcTPS45 (30206.m000785) com pI de 8.21, 8.49, 7.02, 7.64, 8.66 e 9.32 respectivamente. As localizações subcelulares das RcTPS foram previstas e os resultados indicaram que a grande parte das RcTPS possuem localização Citosólica, 28 no total, assim como em *Populus trichocarpa* (**Figura 17**). Dez genes possuem localização na Membrana Plasmática, quatro possuem localização no núcleo celular, dois possuem localização Mitocondrial e apenas um possui localização Extracelular (**Tabela suplementar S1 e S2**). Comparamos nossos resultados com as localizações subcelulares previstas de Terpeno sintase em *Arabidopsis thaliana* e constatamos similaridades no número de genes previstos localizados na mitocôndria e membrana plasmática – 10 e 2 genes, respectivamente (Tholl and Lee, 2011; Irmisch et al., 2014). Os ortólogos PtTPS12 (Potri.017G041700) e RcTPS 25 (30156.m001700); PtTPS7 (Potri.001G308200) e RcTPS1 (29771.m000073) foram previstos como participantes da mesma família e mesma localização subcelular, porém dão origem a intermediários distintos – Isoprenos (C5) e Monoterpenos (C10), respectivamente. O que possivelmente demonstra que nas duas espécies, estas enzimas atuam de diferentes formas, talvez consecutivas, no mesmo compartimento celular, especializando e diversificando a produção de terpenos que auxiliam no processo de adaptação em ambientes distintos (Irmisch et al., 2014; Jackson et al., 2014).

Outras espécies de angiospermas também possuem grande parte das TPS presentes no Citosol, por exemplo *Solanum lycopersicum* (Zhou and Pichersky, 2020), *Arabidopsis thaliana* (Aubourg et al. 2002) e *Glycine max* (LIU, J. et al., 2014). Isso é razoável, pois a biossíntese de terpenos ocorre especialmente no Citosol produzindo diversos sesquiterpenos e

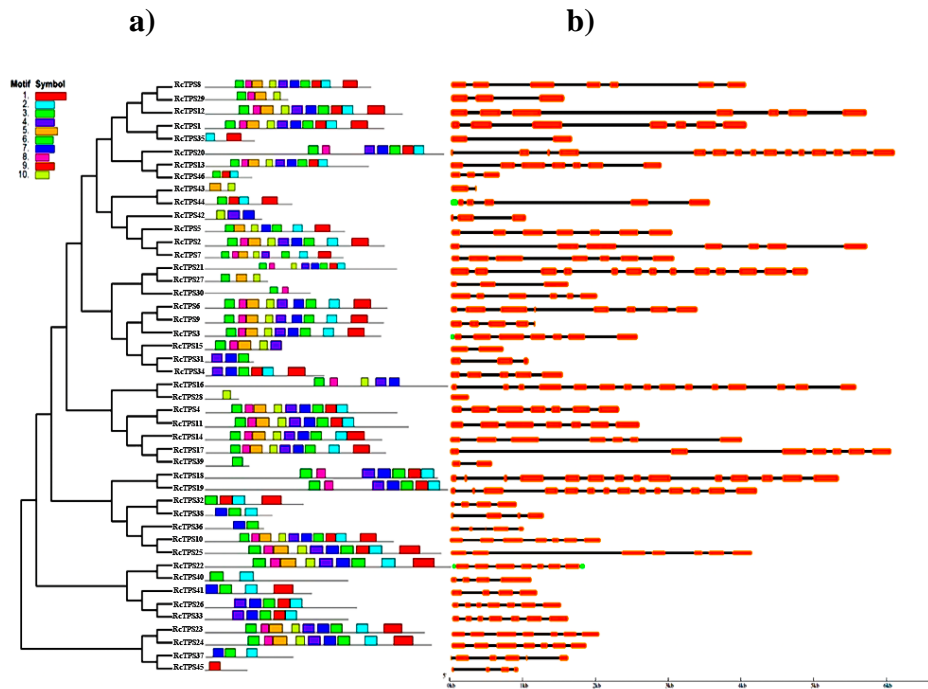
monoterpenos, além de ocorrer também nos plastídios formando derivados de monoterpenos e diterpenos (Chen et al., 2011; Pazouki e Niinemets, 2016; Alquézar, et al., 2017; Alicandri, et al., 2020).

Figura 17: Porcentagem da distribuição das RcTPS em *Ricinus communis* de acordo com a predição da localização subcelular.



A análise dos *motifs* conservados (**Figura 18a**) foi realizada através da *Multiple for Motif Elicitation de Motifs* (MEME) e foram encontrados 10 *motifs* conservados para as RcTPS, com comprimento variável entre 27 a 43 aminoácidos. Para todos foram encontrados domínios pertencentes a terpeno sintase (código Pfan: PF01397 e PF03936), exceto para o *motif* 10 (**Tabela suplementar S4**). Entretanto, todas as 46 sequências apresentadas neste estudo apresentam ao menos um domínio que as caracteriza como RcTPS (**Figura suplementar S1**).

Figura 18: Motifs conservados e estrutura dos genes RcTPS a) Árvore filogenética NJ e distribuição de motifs conservados das 46 proteínas RcTPS. Dez motifs previstos estão demonstrados em caixas de cores diferentes, conforme legenda ao lado. b) Disposição de exons e íntrons dos genes RcTPS. As caixas laranjas representam os exons, os seguimentos de linha preta representam íntrons e as caixas verdes representam as regiões de upstream ou downstream (5' and 3' UTR). Dez motifs foram identificados por meio do software MEME e destacados com cores diferentes.



As análises previstas elaboradas pelo *Gene Structure Display Server 2.0* (**Figura 18b**), revelaram adicionalmente quantos éxons, íntrons, *upstream* e *downstream* possuem as sequências dos genes RcTPS em *Ricinus communis* – Vinte genes RcTPS (43,5%) possuem 7 éxons e 6 íntrons, o que evidencia uma estrutura razoavelmente conservada. Vinte e um genes RcTPS (45,7%) possui menos que 5 éxons e 5 íntrons. Cinco genes RcTPS (10,9%) possuem mais de 13 éxons e 12 íntrons. Apenas 3 genes (RcTPS3, RcTPS22 e RcTPS44) (6,5%), mostraram regiões relacionadas a sítios reguladores de transcrição *upstream* e *downstream*. O gene RcTPS28 com localização prevista como nuclear, mostrou apenas um éxon – o menor previsto desta família de gênica.

Os genes que apresentaram maiores ocorrências de éxons localizaram-se no Citoplasma (RcTPS16 e RcTPS18) e na região próxima a Membrana Plasmática (RcTPS19, RcTPS20 e RcTPS21). A maioria dos genes RcTPS, possuem quantidades expressivas de íntrons, o que garante certo grau de proteção a eventos naturais causadores de mutações em regiões codificantes (Jo et al., 2015). Os genes localizados na região Mitocondrial (RcTPS29

e RcTPS42) possuem exatamente a mesma quantidade de éxons e íntrons, 3 éxons e 2 íntrons respectivamente. Os genes localizados na região Nuclear (RcTPS12, RcTPS6, RcTPS39 e RcTPS45), apresentam quantidades dessemelhantes de éxons e íntrons. Os comprimentos também diferiram entre os 4 genes citados. O gene RcTPS43, obteve previsão de localização Extracelular e, é o segundo menor gene com 0,5 kilobases.

Os eventos remotos de duplicação ou hexaploidização de genes ancestrais são os principais responsáveis pela grandiosa diversificação e especificidade de produtos naturais obtidos das plantas (Jaillon et al., 2007; Chan et al., 2010; Yang et al., 2017; Zeng et al., 2017). Estudos sugerem que eventos em duplicação em tandem é a principal causa de diversificação evolutiva das famílias gênicas em dicotiledôneas (Hudson et al., 2015; Herath e Verchot, 2021). A família das RcTPS contém 12 genes resultantes de eventos naturais de duplicação em tandem (**Tabela suplementar S3 e Figura suplementar S4**). A maioria destes localizados na região Citoplasmática, exceto RcTPS19 (29806.m000952) e RcTPS 46 (30206.m000786) que foram localizados na região da Membrana Plasmática e RcTPS28 (30206.m000791) localizado na região Nuclear.

Vários estudos apontam para um padrão de *motifs* conservados existentes na família terpeno sintase em plantas e em outros organismos. Geralmente, as TPS necessitam de cofatores metálicos divalentes, comumente Mg^{2+} , que se ligam covalentemente a um *motif* canônico DDXXD presente em quase todas as TPS (Chen et al., 2011; Christianson, 2006) Estudos demonstram que a eficiência da função catalítica das TPS está relacionada a estes cofatores que se ligam afim de auxiliar nas primeiras etapas de ionização do difosfato de isopentenila e difosfato de dimetilalila, o que corresponderá nas sucessivas etapas de formação dos terpenos e suas classes (Zhou e Peters, 2009). Outros padrões menos conservados e menos compreendidos também são encontrados na família TPS em plantas, como DDXX, DXDD, DXDD, NSE/DTE e RRX₈W (Kumar et al., 2018; Zhou and Pichersky, 2020b). Em todas as RcTPS em *Ricinus communis* foram encontradas pelo menos um dos *motifs* descritos anteriormente, exceto para RcTPS15, RcTPS28, RcTPS30, RcTPS35, RcTPS43 e RcTPS45, aos quais não foram identificados *motifs* característicos, possivelmente isto se dá devido a incompletude das sequências ou a existência de outros padrões de *motifs* ainda desconhecidos (**Tabela suplementar S6**). Trinta e seis sequências de RcTPS (78,26%) apresentaram o *motif* canônico DDXXD. Vinte e nove (63%) apresentaram *motif* conservado

padrão NSE/DTE. Sete (15,2%) apresentaram *motif* conservado padrão RRR₈W e apenas duas sequências (4,34%) apresentaram padrão de *motif* conservado DXDD. Em algumas sequências de RcTPS apresentaram mais de um padrão de *motifs* conservados: Cinco RcTPS apresentaram três padrões distintos de *motifs*. Vinte e cinco sequências RcTPS apresentaram dois padrões distintos de *motifs* conservados e dez sequências apresentaram apenas um padrão de *motif* conservado (**Tabela complementar S6**).

Elementos reguladores da região promotora foram analisados pelo *PlantCare database* com sequências de 1 kb de *upstream* do códon de iniciação ATG. Tais elementos atuam em processos reguladores de transcrição e estão relacionados a regulação gênica. Os elementos cis também podem se relacionados a estratégias de defesa da planta a estímulos estressores e foram classificados quanto a presença de regiões relacionadas ao ácido abscísico, auxina, giberelina, metil jasmonato (MeJa), respostas a luminosidade, MYB, MYC, indução anaeróbica, resposta de defesa ou estresses, expressão induzida por crise hídrica, baixas temperaturas, ácido salicílico, expressão de meristema, AT-rich, expressão de zeína, controle circadiano e respostas a ferimentos/herbivoria. Todos estes elementos desempenham papéis importantes nas respostas à estresses bióticos e abióticos. (Mosharaf et al., 2020).

Elementos cis responsivos à luz foram encontrados em todos os 45 genes RcTPS, exceto para o gene RcTPS31, ao qual não foi encontrado nenhum elemento promotor registrado no banco de dados Phytozome. Elementos cis relacionados a resposta a incidência de luz, também foram largamente encontrados em maior quantidade em *Arabidopsis thaliana* (Shariatipour e Heidari, 2018), *Citrus sinensis* L. (Mosharaf et al., 2020) e em *Sorghum bicolor* L. (Maheshwari et al., 2019). As respostas a indução anaeróbica foram encontradas em 18 genes RcTPS (RcTPS4, RcTPS10, RcTPS11, RcTPS15, RcTPS19, RcTPS20, RcTPS22, RcTPS26, RcTPS27, RcTPS30, RcTPS33, RcTPS34, RcTPS36, RcTPS37, RcTPS38, RcTPS45 e RcTPS46). Os elementos cis relacionados a resposta a baixas temperaturas foram encontrados em 5 genes (RcTPS13, RcTPS20, RcTPS32, RcTPS34 e RcTPS35). Foram encontrados elementos promotores de expressão por seca em 12 genes (RcTPS6, RcTPS8, RcTPS14, RcTPS18, RcTPS19, RcTPS20, RcTPS22, RcTPS26, RcTPS30, RcTPS34, RcTPS38, RcTPS43). Elementos relacionados ao controle circadiano foram encontrados em 7 genes (RcTPS1, RcTPS3, RcTPS28, RcTPS36, RcTPS40). Elementos relacionados a respostas por ácido abscísico foram encontrados em 19 genes

(RcTPS1, RcTPS2, RcTPS5, RcTPS6, RcTPS9, RcTPS12, RcTPS13, RcTPS14, RcTPS16, RcTPS17, RcTPS22, RcTPS23, RcTPS23, RcTPS24, RcTPS25, RcTPS26, RcTPS27, RcTPS28, RcTPS37, RcTPS39, RcTPS41). Elementos relacionados a auxina foram encontrados em 11 genes (RcTPS1, RcTPS4, RcTPS8, RcTPS11, RcTPS12, RcTPS22, RcTPS26, RcTPS28, RcTPS30, RcTPS33, RcTPS34).

Elementos para a expressão de giberelina foram encontrados em 11 genes (RcTPS2, RcTPS5, RcTPS16, RcTPS17, RcTPS18, RcTPS23, RcTPS28, RcTPS34, RcTPS35, RcTPS42, RcTPS44). Regiões promotoras relacionadas ao ácido metil jasmonato (MeJa) foram encontrados em 15 genes (RcTPS5, RcTPS6, RcTPS8, RcTPS16, RcTPS19, RcTPS20, RcTPS22, RcTPS26, RcTPS27, RcTPS28, RcTPS29, RcTPS38, RcTPS41, RcTPS42, RcTPS43). Aqueles relacionados ao ácido salicílico foram encontrados em 13 genes (RcTPS4, RcTPS5, RcTPS6, RcTPS8, RcTPS10, RcTPS11, RcTPS13, RcTPS15, RcTPS29, RcTPS30, RcTPS33, RcTPS34, RcTPS46). Os elementos promotores relacionados a expressão de MYB foram encontrados em 7 genes (RcTPS1, RcTPS3, RcTPS23, RcTPS24, RcTPS28, RcTPS38, RcTPS43). Não foram encontrados elementos promotores relacionados ao MYC. Estudos sugerem que os elementos MYB e MYC estão associados a maior tolerância de estresses abióticos e bióticos, como resistência a seca, frio, salinidade e restrição hídrica e exposição a patógenos em *Arabidopsis thaliana* e *Citrus sinensis* L. (Seo e Park, 2010; Guo et al., 2017; Mosharaf et al., 2020). Regiões promotoras relacionadas a defesa e estresse conjuntos foram encontrados em 10 genes (RcTPS9, RcTPS10, RcTPS13, RcTPS18, RcTPS19, RcTPS21, RcTPS26, RcTPS30, RcTPS34, RcTPS44). Para a expressão de meristema, foram encontrados elementos promotores em 13 genes (RcTPS5, RcTPS6, RcTPS7, RcTPS11, RcTPS17, RcTPS20, RcTPS21, RcTPS27, RcTPS28, RcTPS34, RcTPS36, RcTPS37, RcTPS41).

Elementos cis regulatórios relacionados a expressão de AT-rich foram localizados em 6 genes (RcTPS2, RcTPS6, RcTPS9, RcTPS17, RcTPS23, RcTPS32). Elementos promotores relacionados a proteína zeína foram encontrados em 8 genes (RcTPS11, RcTPS16, RcTPS18, RcTPS27, RcTPS28, RcTPS35, RcTPS42, RcTPS43) e em apenas 1 gene foi encontrado elementos promotores relacionados a respostas à ferimentos/herbivoria (RcTPS1). (**Figura 19 e Tabela suplementar S7**). A distribuição de elementos cis regulatórios distribuídos nos genes RcTPS apontam que esses genes estão intimamente relacionados nas respostas em defesa a variados tipos de estresses bióticos e abióticos.

O alinhamento múltiplo caracterizado pelo Uniprot (<https://www.uniprot.org/align/>), com método Clustal Omega, foi produzido utilizando as sequências de proteínas das espécies *Arabidopsis thaliana*, *Solanum lycopersicum*, *Oryza sativa*, *Selaginella moellendorffii*, *Populus trichocarpa* e *Ricinus communis*. Esta análise revelou a existência de padrões de similaridade. Em todas elas, o *motif* canônico DDXXD comum para todos os terpenos foi encontrado, até mesmo em *Selaginella moellendorffii*, um organismo modelo do rol das plantas vasculares considerada especial na genômica comparativa por seu registro fóssil datar de 400 milhões de anos. Portanto, estes achados corroboram para a significativa contribuição da família terpeno sintase no processo evolutivo de todas as plantas e organismos que as possuem (**Figura suplementar S7 e S8**).

4 Conclusões

No presente estudo, descobrimos e analisamos 46 genes da família terpeno sintase do genoma de *Ricinus communis* usando métodos de análises da bioinformática. O estudo fornece uma visão abrangente sobre as anotações da grande família de genes RcTPS (*Ricinus communis*L.) acerca de sua estrutura gênica, classificação de subfamílias, *motifs* conservados, localização subcelular e elementos cis promotores. Os resultados obtidos, indicam que a família de genes RcTPS é estruturalmente e funcionalmente agrupada em subfamílias (TPS-a – TPS-g) em todo o seu genoma, fato que se assemelha a outras espécies usadas como comparação neste trabalho. Os fatos apresentados indicam que a família terpeno sintase em *Ricinus communis* apresenta uma extraordinária diversidade estrutural e funcional, o que conduz a variadas entidades químicas produzidas a fim de garantir a estabilidade bioquímica frente a desafios do mundo natural. Estes dados podem ser úteis no planejamento de estratégias de engenharia genética no melhoramento destas espécies objetivando melhor adaptabilidade a estresses abióticos, resistência a herbivoria bem como biossíntese de compostos terpenícos de alta eficiência farmacêutica. Caracterizar os genes da família TPS em *Ricinus communis* por métodos *in silício*, contribui significativamente para o conhecimento desta espécie e posterior caracterização funcional.

5 REFERÊNCIAS

- ABBAS, F. et al. Volatile terpenoids: multiple functions, biosynthesis, modulation and manipulation by genetic engineering. **Planta**, v. 246, n. 5, p. 803-816, 2017.
- ALICANDRI, E. et al. On the Evolution and Functional Diversity of Terpene Sintases in the Pinus Species: A Review. **Journal of Molecular Evolution**, p. 1-31, 2020.
- ALQUÉZAR, B. et al. Genomic analysis of terpene sintase family and functional characterization of seven ynthasepene sintases from Citrus sinensis. **Frontiers in plant science**, v. 8, p. 1481, 2017.
- AUBOURG, S; LECHARNY, A.; BOHLMANN, J. Genomic analysis of the terpenoid sintase (AtTPS) gene family of Arabidopsis thaliana. **Molecular Genetics and Genomics**, v. 267, n. 6, p. 730-745, 2002.
- BERGMAN, M. E.; DAVIS, B.; P., M. A. Medically useful plant terpenoids: biosynthesis, occurrence, and mechanism of action. **Molecules**, v. 24, n. 21, p. 3961, 2019.
- BOHLMANN, J.; MEYER-GAUEN, G.; CROTEAU, R. Plant terpenoid sintases: molecular biology and phylogenetic analysis. **Proceedings of the National Academy of Sciences**, v. 95, n. 8, p. 4126-4133, 1998.
- BUTLER, Jakob B. et al. Annotation of the Corymbia terpene sintase gene family shows broad conservation but dynamic evolution of physical clusters relative to Eucalyptus. **Heredity**, v. 121, n. 1, p. 87-104, 2018.
- CHAN, A. P. et al. Draft genome sequence of the oilseed species Ricinus communis. **Nature biotechnology**, v. 28, n. 9, p. 951-956, 2010.
- CHEN, F. et al. The family of terpene sintases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. **The Plant Journal**, v. 66, n. 1, p. 212-229, 2011.
- CHEN, X. et al. The rice terpene sintase gene Os TPS 19 functions as an (S)-limonene sintase in planta, and its overexpression leads to enhanced resistance to the blast fungus Magnaporthe oryzae. **Plant biotechnology journal**, v. 16, n. 10, p. 1778-1787, 2018.

CHRISTIANSON, D. W. Structural biology and chemistry of the terpenoid cyclases. **Chemical reviews**, v. 106, n. 8, p. 3412-3442, 2006.

DHAR, M. K.; KOUL, A.; KAUL, S. Farnesyl pyrophosphate sintase: a key enzyme in isoprenoid biosynthetic pathway and potential molecular target for drug development. **New biotechnology**, v. 30, n. 2, p. 114-123, 2013.

FLEET, C. M. et al. Overexpression of AtCPS and AtKS in Arabidopsis confers increased ent-kaurene production but no increase in bioactive gibberellins. **Plant physiology**, v. 132, n. 2, p. 830-839, 2003.

GUO, H. et al. Expression of the MYB transcription factor gene Bpl MYB 46 affects abiotic stress tolerance and secondary cell wall deposition in *Betula platyphylla*. **Plant biotechnology journal**, v. 15, n. 1, p. 107-121, 2017.

GROSS, H. KÖNIG, G. M. Terpenoids from marine organisms: unique structures and their pharmacological potential. **Phytochemistry Reviews**, v. 5, n. 1, p. 115-141, 2006.

LANCASTER, J. et al. De novo formation of an aggregation pheromone precursor by an isoprenyl diphosphate sintase-related terpene sintase in the harlequin bug. **Proceedings of the National Academy of Sciences**, v. 115, n. 37, p. E8634-E8641, 2018.

LI, G. et al. Nonseed plant *Selaginella moellendorffii* has both seed plant and microbial types of terpene sintases. **Proceedings of the National Academy of Sciences**, v. 109, n. 36, p. 14711-14715, 2012.

LIU, J. et al. Genome-wide analysis of terpene sintases in soybean: functional characterization of GmTPS3. **Gene**, v. 544, n. 1, p. 83-92, 2014.

LUCK, K. et al. The reconstruction and biochemical characterization of ancestral genes furnish insights into the evolution of terpene sintase function in the Poaceae. **Plant Molecular Biology**, v. 104, n. 1, p. 203-215, 2020.

KARUNANITHI, P. S.; ZERBE, P. Terpene sintases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. **Frontiers in Plant Science**, v. 10, p. 1166, 2019.

KEELING, C. I.; BOHLMANN, J. Diterpeno resin acids in conifers. **Phytochemistry**, v. 67, n. 22, p. 2415-2423, 2006.

KIRYU, M. et al. Rice terpene sintase 18 (OsTPS18) encodes a ynthasepene sintase that produces an antibacterial LENGTH-nerolidol against a bacterial pathogen of rice. **Journal of General Plant Pathology**, v. 84, n. 3, p. 221-229, 2018.

KUMAR, Y. et al. Genome-wide detection of terpene sintase genes in holy basil (*Ocimum sanctum* L.). **PloS one**, v. 13, n. 11, p. e0207097, 2018.

HAY, R. Investigation of the role of gene clusters in terpene biosynthesis in *Sorghum bicolor*. 2018.

HERATH, V.; VERCHOT, J. Insight into the bZIP Gene Family in *Solanum tuberosum*: Genome and Transcriptome Analysis to Understand the Roles of Gene Diversification in Spatiotemporal Gene Expression and Function. **International Journal of Molecular Sciences**, v. 22, n. 1, p. 253, 2021.

HILLWIG, M. L. et al. Domain loss has independently occurred multiple times in plant terpene sintase evolution. **The Plant Journal**, v. 68, n. 6, p. 1051-1060, 2011.

HUDSON, C. J. et al. Genomic patterns of species diversity and divergence in *Eucalyptus*. **New Phytologist**, v. 206, n. 4, p. 1378-1390, 2015.

IRMISCH, S. et al. Terpene sintases and their contribution to herbivore-induced volatile emission in western balsam poplar (*Populus trichocarpa*). **BMC Plant Biology**, v. 14, n. 1, p. 1-16, 2014.

JACKSON, A. J. et al. Biochemical characterization of the castor bean ent-kaurene sintase (-like) family supports quantum chemical view of ynthase cyclization. **Phytochemistry**, v. 103, p. 13-21, 2014.

JAILLON, O. et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. **Nature**, v. 449, n. 7161, p. 463, 2007.

JIA, Q. Computational Identification of Terpene Sintase Genes and Their Evolutionary Analysis. 2016.

JIA, et al. A pair of residues that interactively affect diterpene synthase product outcome. **ACS chemical biology**, v. 12, n. 3, p. 862-867, 2017.

JO, Bong-Seok; CHOI, Sun Shim. Introns: the functional benefits of introns in genomes. **Genomics & informatics**, v. 13, n. 4, p. 112, 2015.

MAHIZAN, N. A. et al. Terpene derivatives as a potential agent against antimicrobial resistance (AMR) pathogens. **Molecules**, v. 24, n. 14, p. 2631, 2019.

MARTÍN-SÁNCHEZ, L. et al. Phylogenomic analyses and distribution of terpene sintases among *Streptomyces*. **Beilstein journal of organic chemistry**, v. 15, n. 1, p. 1181-1193, 2019.

MATARESE, F. et al. Expression of terpene sintase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. **Phytochemistry**, v. 105, p. 12-24, 2014.

MICHAEL, R. et al. Light-regulated expression of terpene sintase gene, AtTPS03, is controlled by the bZIP transcription factor, HY5, in *Arabidopsis thaliana*. **Biochemical and Biophysical Research Communications**, v. 529, n. 2, p. 437-443, 2020.

Maheshwari, P. et al. Genome-wide identification and expression profile analysis of Nuclearear factor Y family genes in *Sorghum bicolor* L. (Moench). *PLoS One* 14: 1–27 (2019).
<https://doi.org/10.1371/journal.pone.0222203>

MOSHARAF, P. et al. In silico identification and characterization of AGO, DCL and RDR gene families and their associated regulatory elements in sweet orange (*Citrus sinensis* L.). **PloS one**, v. 15, n. 12, p. e0228233, 2020.

NAGEL, R. et al. Overexpression of an isoprenyl diphosphate sintase in spruce leads to unexpected terpene diversion products that function in plant defense. **Plant Physiology**, v. 164, n. 2, p. 555-569, 2014

NAGEGOWDA, D. A.; GUPTA, P.. Advances in biosynthesis, regulation, and metabolic engineering of plant specialized terpenoids. **Plant Science**, v. 294, p. 110457, 2020.

NETO, V. G. et al. Sequence analysis of *Ricinus communis* small heat-shock protein (sHSP) subfamily and its role in abiotic stress responses. **Industrial Crops and Products**, v. 152, p. 112541, 2020.

NETO, V. G. et al. Characterization of the superoxide dismutase gene family in seeds of two *Ricinus communis* .genotypes submitted to germination under water restriction conditions. **Environmental and Experimental Botany**, v. 155, p. 453-463, 2018.

NIEUWENHUIZEN, N. J. et al. Functional genomics reveals that a compact terpene sintase gene family can account for terpene volatile production in apple. **Plant physiology**, v. 161, n. 2, p. 787-804, 2013.

OLDFIELD, E.; LIN, F.. Terpene biosynthesis: modularity rules. *Angewandte Chemie International Edition*, v. 51, n. 5, p. 1124-1137, 2012.

PAZOUKI, L.; NIINEMETS, Ü.. Multi-substrate terpene sintases: their occurrence and physiological significance. **Frontiers in Plant Science**, v. 7, p. 1019, 2016.

PRIYA, P.; KUMARI, S.; YADAV, G. Quantification of the plant terpenome: predicted versus actual emission potentials. **Indian Journal of Plant Physiology**, v. 21, n. 4, p. 569-575, 2016.

PRIYA, P. et al. Terzyme: a tool for identification and analysis of the plant terpenome. **Plant Methods**, v. 14, n. 1, p. 1-18, 2018.

SEO, P.; PARK, C.. MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in Arabidopsis. **New Phytologist**, v. 186, n. 2, p. 471-483, 2010.

SMIT, S. J.; VIVIER, M. A.; YOUNG, P. Comparative (within species) genomics of the *Vitis vinifera* L. terpene synthase family to explore the impact of genotypic variation using phased diploid genomes. **Frontiers in Genetics**, v. 11, p. 421, 2020.

SOLTANI, J. Secondary metabolite diversity of the genus *Aspergillus*: recent advances. **New and future developments in microbial biotechnology and bioengineering**, p. 275-292, 2016.

SPYROPOULOU, E. A.; HARING, M. A.; SCHUURINK, R. C. RNA sequencing on *Solanum lycopersicum* trichomes identifies transcription factors that activate terpene synthase promoters. **BMC genomics**, v. 15, n. 1, p. 1-16, 2014.

THOLL, D.; LEE, S. Terpene specialized metabolism in *Arabidopsis thaliana*. The Arabidopsis book, v. 9, p. e0143, 2011.

THOLL, D. et al. The biochemistry of homoterpenes—common constituents of floral and herbivore-induced plant volatile bouquets. **Phytochemistry**, v. 72, n. 13, p. 1635-1646, 2011.

TRAPP, S. C.; CROTEAU, R. B. Genomic organization of plant terpene synthases and molecular evolutionary implications. **Genetics**, v. 158, n. 2, p. 811-832, 2001.

WIBERLEY et al. Regulation of isoprene emission in *Populus trichocarpa* leaves subjected to changing growth temperature. **Plant, Cell & Environment**, v. 31, n. 2, p. 258-267, 2008.

XIE, X.; K., J.; KEASLING, J. D. Functional characterization of four synthase genes from *Ricinus communis* (castor bean). **Phytochemistry**, v. 78, p. 20-28, 2012.

YAMADA, Y.; CANE, D. E.; IKEDA, H. Diversity and analysis of bacterial terpene synthases. **Methods in enzymology**, v. 515, p. 123-162, 2012.

YANG, Jun et al. Haplotype-resolved sweet potato genome traces back its hexaploidization history. **Nature plants**, v. 3, n. 9, p. 696-703, 2017.

YU, Z. et al. Genome-wide identification and expression profile of TPS gene family in *Dendrobium officinale* and the role of DoTPS10 in linalool biosynthesis. **International journal of molecular sciences**, v. 21, n. 15, p. 5419, 2020.

Yu C., Lin C., Hwang J. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. **Protein Science** 2004, 13:1402-1406.

- Yu C., Chen Y., Lu C., Hwang J.: Prediction of protein subcellular localization. *Proteins: Structure, Function and Bioinformatics* 2006, 64:643-651.
- ZENG, Liping et al. Resolution of deep eudicot phylogeny and their temporal diversification using Nuclear genes from transcriptomic and genomic datasets. **New Phytologist**, v. 214, n. 3, p. 1338-1354, 2017.
- ZERBE, Philipp; BOHLMANN, Jörg. Plant ynthase sintases: exploring modularity and metabolic diversity for bioengineering. **Trends in biotechnology**, v. 33, n. 7, p. 419-428, 2015.
- ZHOU, F.; PICHERSKY, E.. The complete functional ynthasepeneon of the terpene sintase family in tomato. **New Phytologist**, v. 226, n. 5, p. 1341-1360, 2020.
- ZHOU, F.; PICHERSKY, E. More is better: the diversity of terpene metabolism in plants. *Current opinion in plant biology*, v. 55, p. 1-10, 2020.
- ZHOU, H. et al. Analysis of terpene sintase family genes in *Camellia sinensis* with an emphasis on abiotic stress conditions. **Scientific reports**, v. 10, n. 1, p. 1-13, 2020.
- ZHUANG, X. et al. Dynamic evolution of herbivore-induced ynthasepene biosynthesis in sorghum and related grass crops. **The Plant Journal**, v. 69, n. 1, p. 70-80, 2012.
- ZHOU, K.; PETERS, R. J. Investigating the conservation pattern of a putative second terpene sintase divalent metal binding motif in plants. **Phytochemistry**, v. 70, n. 3, p. 366-369, 2009.

ANEXOS



Tabela suplementar S1: Características gerais das proteínas terpeno sintase em *Ricinus communis*

Nº	Nome do gene	ID	Subfamília	L. celular	AA	P. molecular (Da)	p.I. teórico	I. index	A. index	GRAV Y	Ortólogos em Arabidopsis	Ortólogos em Populus
1	RcTPS1	29771.m000073	b	Citoplasma	551	64630.08	5.29	39.64	92.74	-0.320	-	Potri.001G308200
2	RcTPS2	29840.m000603	a	Citoplasma	550	63679.60	5.19	44.59	93.76	-0.285	-	Potri.019G045400
3	RcTPS3	29910.m000951	g	Citoplasma	547	63692.92	5.35	54.55	93.36	-0.259	-	Potri.004G030200
4	RcTPS4	30169.m006289	b	Citoplasma	563	64488.54	6.12	35.62	91.71	-0.268	AT4G16740	-
5	RcTPS5	29910.m000944	e/f	Citoplasma	476	55462.30	5.08	44.66	81.20	-0.326	-	Potri.004G037900
6	RcTPS6	30169.m006314	g	Citoplasma	552	64359.12	5.27	39.54	88.35	-0.349	-	Potri.004G030200
7	RcTSP7	29827.m002614	a	Citoplasma	551	63819.66	5.29	44.41	89.75	-0.302	-	Potri.015G085500/ Potri.007G074400
8	RcTPS8	30156.m001703	b	M. Plasmática	584	67733.64	5.82	43.42	94.83	-0.226	AT2G24210	Potri.019G023000
9	RcTPS9	30206.m000781	g	Citoplasma	555	64359.12	5.32	37.29	88.94	-0.344	-	Potri.004G030200
10	RcTPS10	30131.m007001	b	Citoplasma	599	69129.06	5.80	35.09	91.00	-0.294	-	Potri.017G041700
11	RcTPS11	30225.m001714	b	M. Plasmática	572	65426.40	5.64	50.93	94.27	-0.197	AT4G16740	-
12	RcTPS12	27688.m000060	e/f	Nuclear	587	67929.42	5.87	48.86	88.72	-0.334	AT1G61120	-
13	RcTPS13	29842.m003696	a	Citoplasma	572	65857.87	5.89	41.80	92.05	-0.309	-	Potri.001G415100
14	RcTPS14	29842.m003692	a	Citoplasma	521	60706.53	6.93	43.55	87.75	-0.417	-	Potri.015G032100
15	RcTPS15	30206.m000792	b	Citoplasma	222	26283.10	5.77	38.26	103.24	-0.345	*	*
16	RcTPS16	29647.m002075	c	Citoplasma	800	91352.61	5.60	38.33	85.32	-0.369	AT4G02780	-
17	RcTPS17	28166.m001038	a	Citoplasma	558	64391.82	5.37	40.06	94.19	-0.278	-	Potri.011G142800
18	RcTPS18	29806.m000953	c	Citoplasma	775	88179.10	5.23	40.07	96.00	-0.185	AT4G02780	-
19	RcTPS19	29806.m000952	c	M. Plasmática	778	88763.60	5.27	40.76	91.38	-0.233	AT4G02780	-
20	RcTPS20	29806.m000947	a	M. Plasmática	816	93570.55	6.12	41.60	93.08	-0.217	-	Potri.001G415100
21	RcTPS21	28670.m000104	b	M. Plasmática	808	92019.54	5.68	35.73	95.09	-0.151	AT4G16730	Potri.007G119700
22	RcTPS22	30169.m006283	a	Citoplasma	601	68965.60	5.35	41.00	85.51	-0.239	AT4G20200/AT1G31950	-
23	RcTPS23	30169.m006276	a	Citoplasma	614	70814.85	6.10	43.28	79.74	-0.343	AT3G29190/AT3G14490	-
24	RcTPS24	30169.m006286	a	Citoplasma	599	68775.44	5.44	42.01	84.97	-0.255	AT3G29190/AT3G14490	-
25	RcTPS25	30156.m001700	b	Citoplasma	596	68363.30	6.01	34.64	87.57	-0.303	-	Potri.017G041700
26	RcTPS26	29233.m000032	g	Citoplasma	408	478380.01	5.54	46.30	92.25	-0.337	-	Potri.004G030400
27	RcTPS27	30206.m000787	b	Citoplasma	226	26444.15	4.95	41.17	103.10	-0.260	*	*
28	RcTPS28	30206.m000791	b	Nuclear	99	11763.74	4.97	74.39	114.24	-0.097	*	*

29	RcTPS29	30156.m001716	b	Mitocôndria	275	31884.60	6.80	50.95	87.24	-0.493	*	*
30	RcTPS30	29917.m001932	b	M. Plasmática	411	46930.01	5.17	51.05	86.42	-0.160	AT4G16730	Potri.007G119700
31	RcTPS31	59739.m000007	g	Citoplasma	143	17319.10	5.06	41.46	101.68	-0.153	-	Potri.004G030400
32	RcTPS32	29771.m000074	b	Citoplasma	257	29568.90	4.95	36.96	90.00	-0.211	*	*
33	RcTPS33	27716.m000050	g	Citoplasma	408	47634.67	5.65	51.52	93.24	-0.369	-	Potri.004G030400
34	RcTPS34	27394.m000360	g	Citoplasma	351	40294.21	5.67	38.55	91.68	-0.166	-	Potri.004G030400
35	RcTPS35	28359.m000286	a	Citoplasma	178	20102.01	4.88	36.34	88.71	-0.126	*	*
36	RcTPS36	29233.m000033	a	Nuclear	161	18627.10	5.39	36.07	87.89	-0.289	-	Potri.019G045300
37	RcTPS37	29673.m000927	a	M. Plasmática	271	31556.32	8.21	37.99	80.96	-0.269	AT4G15870/ AT1G70080	-
38	RcTPS38	29842.m003698	c	M. Plasmática	180	20877.23	8.49	35.05	106.67	-0.063	-	Potri.005G210300
39	RcTPS39	29878.m000240	a	Nuclear	126	14429.30	6.96	40.71	85.87	-0.501	*	*
40	RcTPS40	29917.m001933	g	Citoplasma	335	38713.60	7.02	50.63	87.07	-0.375	-	Potri.004G030400
41	RcTPS41	30065.m001152	g	Citoplasma	282	32471.20	5.29	32.11	85.46	-0.177	-	Potri.004G030400
42	RcTPS42	30065.m001155	a	Mitocôndria	156	18672.60	7.64	34.57	90.00	-0.340	AT4G15870/ AT1G70080	-
43	RcTPS43	29878.m000241	a	Extracelular	98	10960.66	8.66	20.57	94.59	-0.148	*	*
44	RcTPS44	30156.m001715	g	M. Plasmática	288	32558.40	4.60	39.04	100.87	0.081	-	Potri.004G030400
45	RcTPS45	30206.m000785	b	Nuclear	123	14156.60	9.32	46.68	90.41	-0.258	*	*
46	RcTPS46	30206.m000786	e/f	M. Plasmática	192	22128.42	4.90	46.55	95.00	-0.029	AT1G79460	Potri.008G082700/ Potri.008G082400

- destacam a ausência de ortólogos para a espécie indicada

* indicam as sequências que foram retiradas da construção da árvore filogenética.

Gene ID: 29647.m002075_RcTPS16

Significant Pfam-A Matches

Show or hide all alignments.

Table with 14 columns: Family, Description, Entry type, Clan, Envelope (Start, End), Alignment (Start, End), HMM (From, To), HMM length, Bit score, E-value, Predicted active sites, Show/hide alignment. Contains two entries for Terpen synth and Terpen synth_C.

Gene ID: 28166.m001038_RcTPS17

Significant Pfam-A Matches

Show or hide all alignments.

Table with 14 columns: Family, Description, Entry type, Clan, Envelope (Start, End), Alignment (Start, End), HMM (From, To), HMM length, Bit score, E-value, Predicted active sites, Show/hide alignment. Contains two entries for Terpen synth_C and Terpen synth.

Gene ID: 29806.m000953_RcTPS18

Significant Pfam-A Matches

Show or hide all alignments.

Table with 14 columns: Family, Description, Entry type, Clan, Envelope (Start, End), Alignment (Start, End), HMM (From, To), HMM length, Bit score, E-value, Predicted active sites, Show/hide alignment. Contains two entries for Terpen synth_C and Terpen synth.

Gene ID: 29806.m000952_RcTPS19

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene_synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	455	719	455	718	1	266	267	209.4	6.1e-62	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															
	Terpene_synth	Terpene synthase, N-terminal domain	Domain	n/a	221	413	221	413	1	192	192	148.3	2.1e-43	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29806.m000947_RcTPS20

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene_synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	491	755	491	754	1	266	267	221.1	1.6e-65	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															
	Terpene_synth	Terpene synthase, N-terminal domain	Domain	n/a	257	449	257	448	1	191	192	139.1	1.4e-40	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 28670.m000104_RcTPS21

Gene ID: 28359.m000286_RcTPS35

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene_synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	1	123	1	123	145	267	267	165.0	2.1e-48	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29233.m000033_RcTPS36

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene_synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	52	158	56	156	14	114	267	96.5	1.6e-27	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29673.m000927_RcTPS37

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene_synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	4	227	9	221	23	240	267	253.1	2.9e-75	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29842.m003698_RcTPS38

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	1	93	3	88	13	98	267	86.9	1.4e-24	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	85	180	90	180	128	219	267	65.4	4.9e-18	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29878.m000240_RcTPS39

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth	Terpene synthase, N-terminal domain	Domain	n/a	32	126	32	126	1	109	192	87.0	1.3e-24	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29917.m001933_RcTPS40

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	1	198	1	197	63	266	267	153.7	5.8e-45	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 30065.m001152_RcTPS41

Significant Pfam-A Matches

Show or hide all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	1	225	1	225	38	267	267	280.9	9.2e-84	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 30065.m001155_RcTPS42

Significant Pfam-A Matches

Show or [hide](#) all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	69	156	69	156	1	88	267	135.2	2.6e-39	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 29878.m000241_RcTPS43

Significant Pfam-A Matches

Show or [hide](#) all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth	Terpene synthase, N-terminal domain	Domain	n/a	9	82	12	82	118	192	192	69.3	3.3e-19	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 30156.m001715_RcTPS44

Significant Pfam-A Matches

Show or [hide](#) all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	23	233	30	232	63	266	267	214.7	1.5e-63	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 30206.m000785_RcTPS45

Significant Pfam-A Matches

Show or [hide](#) all alignments.

	Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
					Start	End	Start	End	From	To					
	Terpene synth_C	Terpene synthase family, metal binding d...	Domain	CL0613	43	120	44	114	176	246	267	59.7	2.8e-16	n/a	Hide
#HMM															
#MATCH															
#PP															
#SEQ															

Gene ID: 30206.m000786_RcTPS46

Tabela Suplementar S2: Relação das localizações subcelulares das RcTPS e seus respectivos valores de confiabilidade apontados no CELLO v.2.5.

Nome do gene	ID	L. subcelular	Índice de confiança*
RcTPS1	29771.m000073	Citoplasma	2.733
RcTPS2	29840.m000603	Citoplasma	2.287
RcTPS3	29910.m000951	Citoplasma	3.494
RcTPS4	30169.m006289	Citoplasma	1.523
RcTPS5	29910.m000944	Citoplasma	3.704
RcTPS6	30169.m006314	Citoplasma	1.507
RcTSP7	29827.m002614	Citoplasma	2.310
RcTPS8	30156.m001703	M. Plasmática	1.554
RcTPS9	30206.m000781	Citoplasma	1.507
RcTPS10	30131.m007001	Citoplasma	1.365
RcTPS11	30225.m001714	M. Plasmática	1.579
RcTPS12	27688.m000060	Nuclear	1.347
RcTPS13	29842.m003696	Citoplasma	1.430
RcTPS14	29842.m003692	Citoplasma	2.646
RcTPS15	30206.m000792	Citoplasma	2.330
RcTPS16	29647.m002075	Citoplasma	2.832
RcTPS17	28166.m001038	Citoplasma	2.162
RcTPS18	29806.m000953	Citoplasma	1.398
RcTPS19	29806.m000952	M. Plasmática	1.274
RcTPS20	29806.m000947	M. Plasmática	1.765
RcTPS21	28670.m000104	M. Plasmática	1.937
RcTPS22	30169.m006283	Citoplasma	1.782
RcTPS23	30169.m006276	Citoplasma	1.509
RcTPS24	30169.m006286	Citoplasma	1.938
RcTPS25	30156.m001700	Citoplasma	1.611
RcTPS26	29233.m000032	Citoplasma	4.091
RcTPS27	30206.m000787	Citoplasma	1.399
RcTPS28	30206.m000791	Nuclear	1.496

RcTPS29	30156.m001716	Mitocôndria	1.890
RcTPS30	29917.m001932	M. Plasmática	1.755
RcTPS31	59739.m000007	Citoplasma	1.412
RcTPS32	29771.m000074	Citoplasma	2.196
RcTPS33	27716.m000050	Citoplasma	3.675
RcTPS34	27394.m000360	Citoplasma	1.472
RcTPS35	28359.m000286	Citoplasma	1.659
RcTPS36	29233.m000033	Nuclear	1.632
RcTPS37	29673.m000927	M. Plasmática	1.556
RcTPS38	29842.m003698	M. Plasmática	2.358
RcTPS39	29878.m000240	Nuclear	1.901
RcTPS40	29917.m001933	Citoplasma	1.845
RcTPS41	30065.m001152	Citoplasma	1.577
RcTPS42	30065.m001155	Mitocôndria	1.435
RcTPS43	29878.m000241	Extra celular	1.720
RcTPS44	30156.m001715	M. Plasmática	2.469
RcTPS45	30206.m000785	Nuclear	2.525
RcTPS46	30206.m000786	M. Plasmática	2.118

Figura suplementar S4: Duplicações em tandem das RcTPS

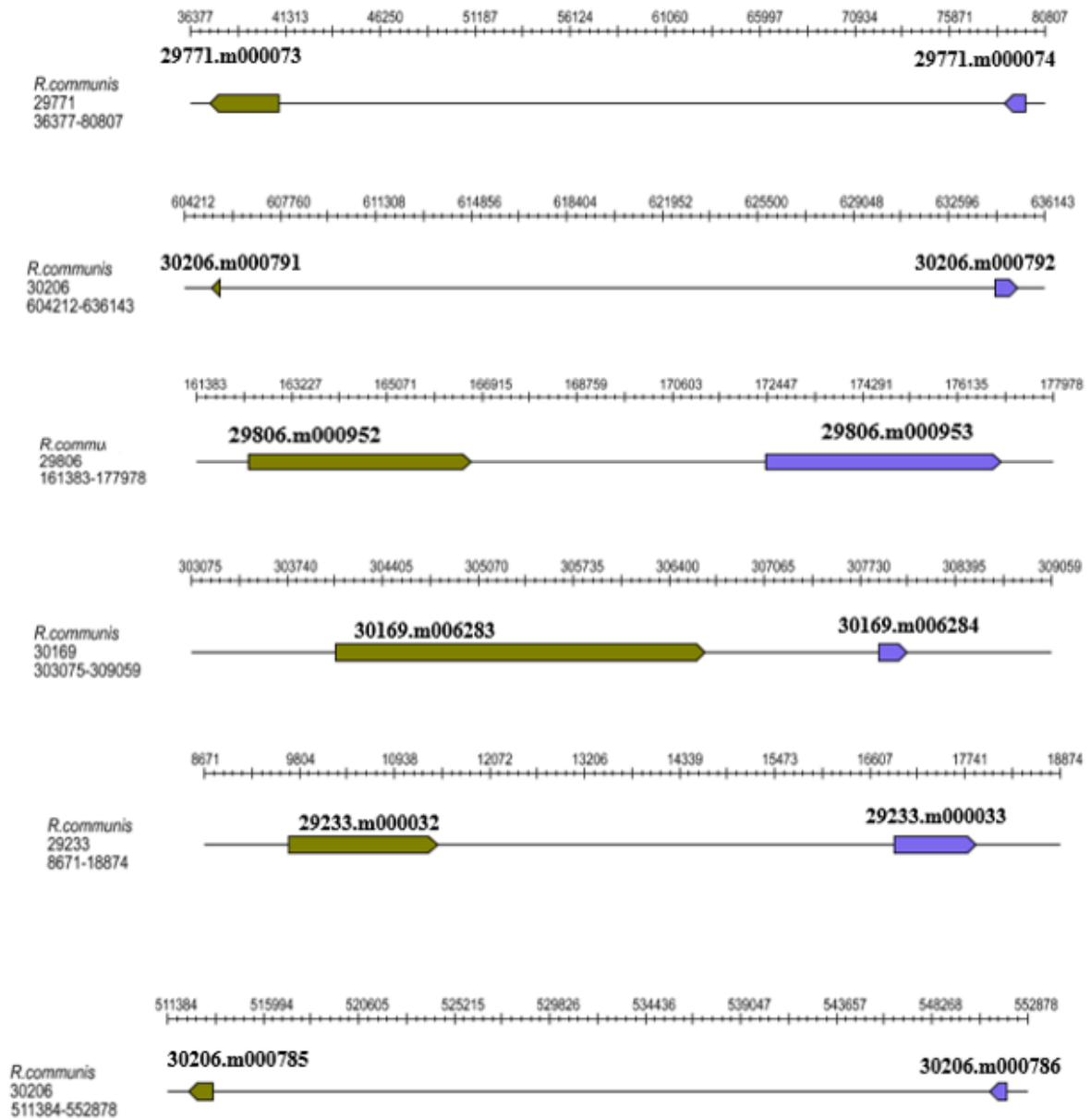


Tabela suplementar S3: Características gerais sobre os genes duplicados em tandem das RcTPS.

Query ID	Subject ID	% identity	A. Length	Q. start	Q. end	S.start	S. end	E-Value	Scores
29771.m000073	29771.m000074	87.6	250	302	551	1	250	8.00E ⁻¹³²	467
29771.m000074	29771.m000073	87.6	250	1	220	302	551	3.00E ⁻¹³²	467
30206.m000791	30206.m000792	82.29	96	4	99	110	205	6.00E ⁻³⁹	154
30206.m000792	30206.m000791	82.29	96	110	205	4	99	2.00E ⁻³⁸	154
29806.m000952	29806.m000953	72.69	791	1	778	1	775	0	1191
29806.m000953	29806.m000952	72.69	791	1	775	1	778	0	1191
30169.m006283	30169.m006284	98.44	64	538	601	1	64	2.00E ⁻³²	137
30169.m006284	30169.m006283	98.44	64	1	64	538	601	1.00E ⁻³³	137
29233.m000032	29233.m000033	77.32	97	101	197	60	156	1.00E ⁻³⁹	160
29233.m000033	29233.m000032	77.32	97	60	156	101	197	2.00E ⁻⁴⁰	160
30206.m000785	30206.m000786	68.18	110	16	95	74	183	5.00E ⁻³⁵	142
30206.m000786	30206.m000785	68.18	110	74	183	16	95	1.00E ⁻³⁴	142

Figura suplementar S5: Estrutura dos genes RcTPS

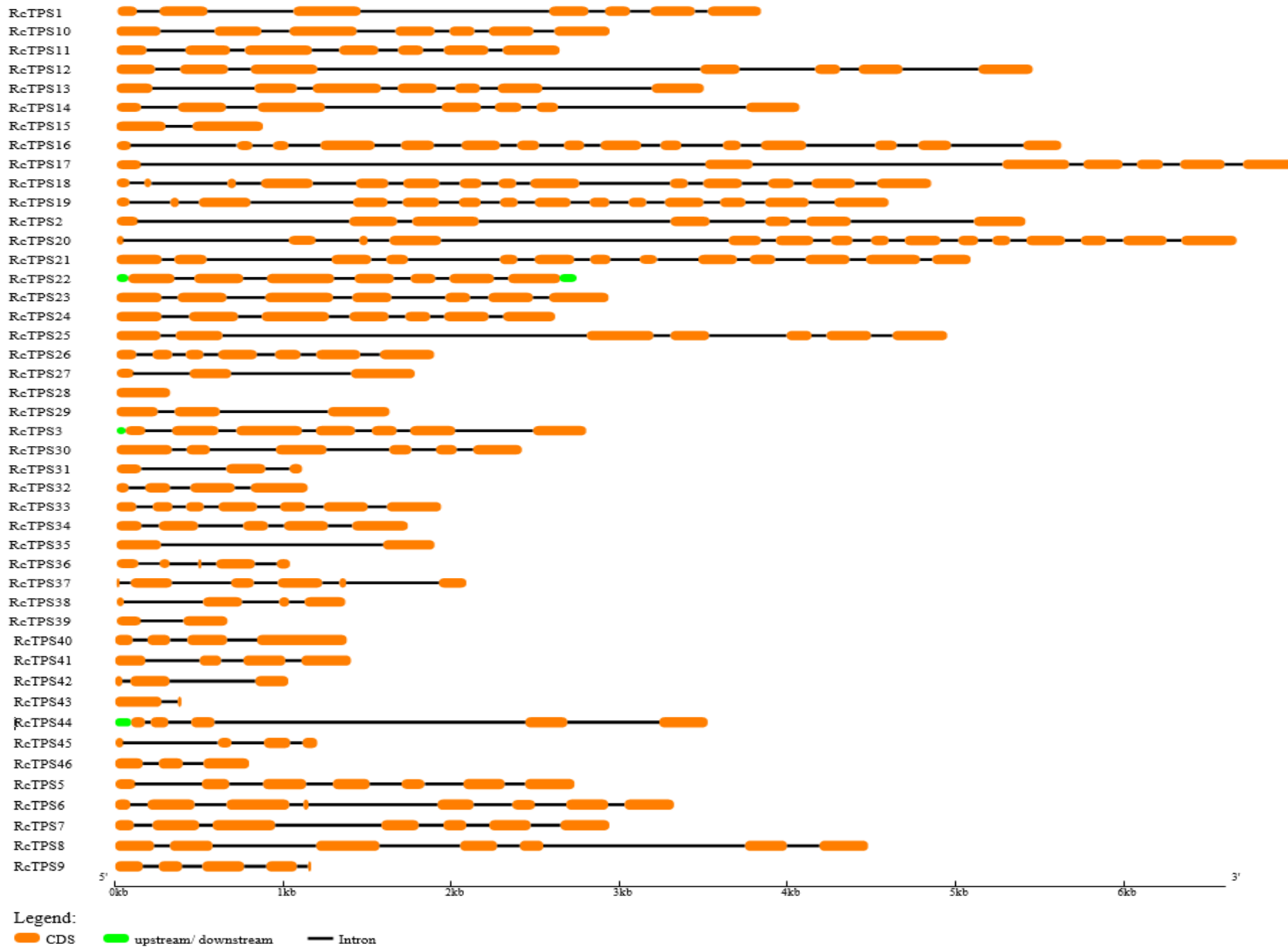


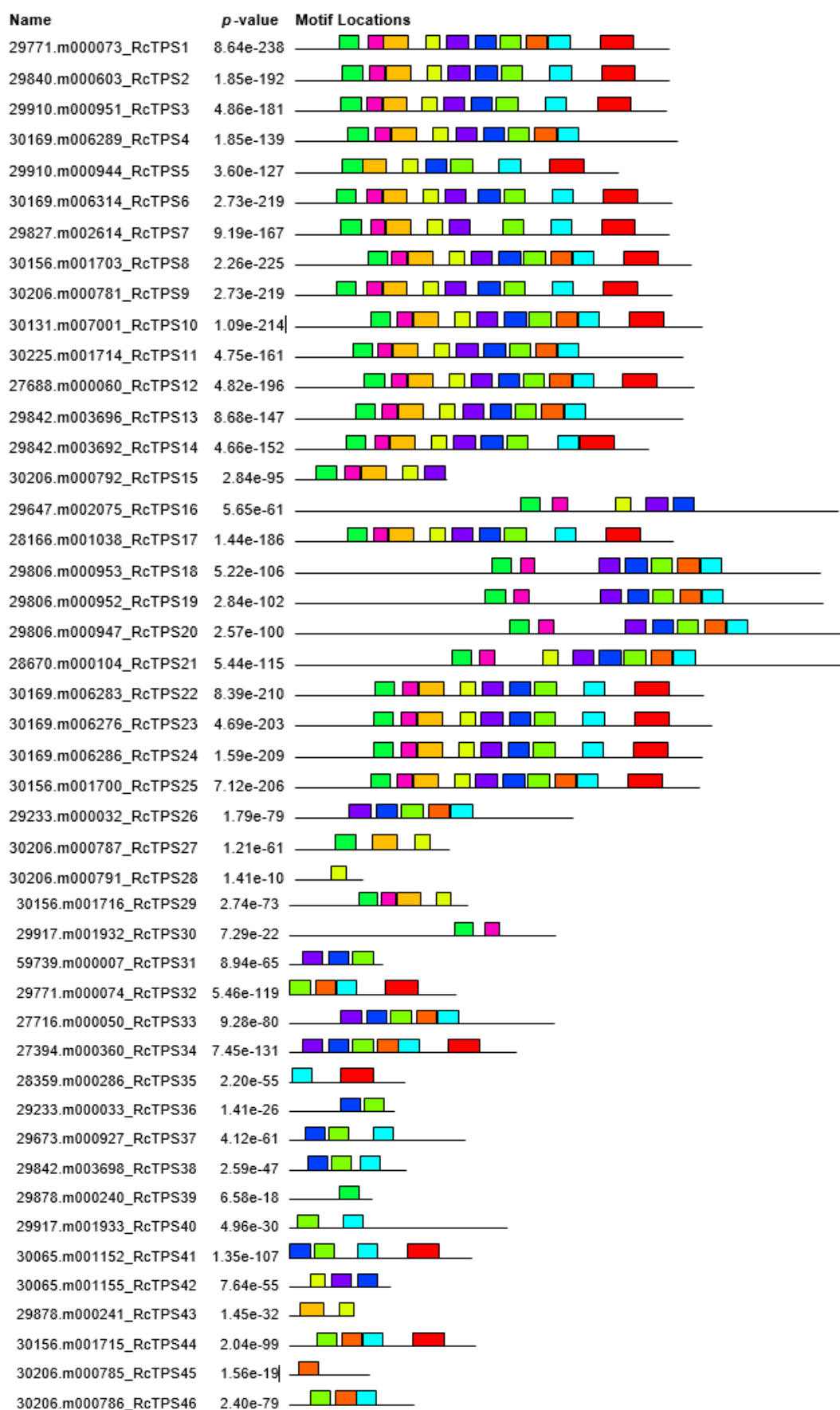
Figura suplementar S6: Motifs conservados dos genes RcTPS

Tabela Suplementar S4: *Motifs* conservados dos genes RcTPS que apresentaram domínios no Pfan

Nº Motif	Domínio registrado no Pfan	HMM
1	Sim	syereqkrqgdvassvecymkehgvsseeAveeikkliedawkd
2	Sim	lvkaylkeakWreekyvPsfeEylenslvss
3	Sim	IDDtyDvygtleEleklteaierWdesave
4	Sim	LeLAkldFnllqslhqkElkelsrWwk
5	Sim	FkdeegnfkelsedvkgllsLYeAshlstegeeiL
6	Sim	LelidtLqrLGisyhFekEIkeiLdeiyr
7	Sim	aRdrlvecyfwalgvvfepeysraRiilaKv
8	Sim	aRdrlvecyfwalgvvfepeysraRiilaKv
9	Sim	ykalntfneledelekeeksdrvskylkeaw
10	Não encontrado	-

Tabela suplementar S5: Caracterização das subfamílias (TPSa – TPSg) em *Ricinus communis* e seus possíveis produtos previstos apresentada pelo software Terzyme.

Gene	Subfamília	Score	E-Value	Nº de domínios	Atividade predita
RcTPS1	b	863.0	2.0E-262	1	Monoterpeno Sintase
RcTPS2	a	830.5	5.4e-252	1	Sesquiterpeno Sintase
RcTPS3	a	721.3	5.1e-219	1	Sesquiterpeno Sintase
RcTPS4	g	759.3	6.2e-231	1	Monoterpeno Sintase
RcTPS5	a	431.5	1.5e-131	3	Sesquiterpeno Sintase
RcTPS6	a	691.1	6.6e-210	1	Sesquiterpeno Sintase
RcTPS7	a	618.0	7.4e-188	1	Sesquiterpeno Sintase
RcTPS8	b	741.8	1.1e-225	1	Monoterpeno Sintase
RcTPS9	a	691.1	6.6e-210	1	Sesquiterpeno Sintase
RcTPS10	b	811.0	1.2e-246	1	Monoterpeno Sintase
RcTPS11	g	662.2	1.6e-201	1	Monoterpeno Sintase
RcTPS12	b	775.4	7.3e-236	1	Monoterpeno Sintase
RcTPS13	g	810.0	2.7e-246	1	Monoterpeno Sintase
RcTPS14	b	634.9	2.4e-193	2	Sesquiterpeno Sintase
RcTPS15	b	285.4	1.8e-87	1	Monoterpeno Sintase
RcTPS16	c	1225.3	0	1	Diterpeno Sintase
RcTPS17	a	783.6	7.7e-238	1	Sesquiterpeno Sintase
RcTPS18	e/f	928.9	5.5e-282	1	Diterpeno Sintase
RcTPS19	e/f	985.1	5.5e-299	1	Diterpeno Sintase
RcTPS20	e/f	912.5	5.2e-277	1	Diterpeno Sintase
RcTPS21	e/f	1111.5	0	1	Diterpeno Sintase
RcTPS22	a	644.1	1,00E-195	1	Sesquiterpeno Sintase
RcTPS23	a	679.2	2.6e-206	1	Sesquiterpeno Sintase
RcTPS24	a	646.2	2.4e-196	1	Sesquiterpeno Sintase
RcTPS25	b	750.6	2.3e-228	1	Monoterpeno Sintase
RcTPS26	e/f	445.2	1,00E-135	1	Diterpeno Sintase
RcTPS27	b	252.3	1.8e-77	1	Monoterpeno Sintase
RcTPS28	b	116.2	2.9e-36	1	Monoterpeno Sintase
RcTPS29	b	260.8	4.9e-80	1	Sesquiterpeno Sintase
RcTPS30	e/f	397.5	2.8e-121	1	Diterpeno Sintase
RcTPS31	b	235.9	1.6e-72	1	Monoterpeno Sintase

RcTPS32	b	416.2	4.3e-127	1	Monoterpeno Sintase
RcTPS33	e/f	458.8	8.4e-140	1	Diterpeno Sintase
RcTPS34	b	471.5	7.6e-144	1	Monoterpeno Sintase
RcTPS35	a	262.9	1.2e-80	1	Sesquiterpeno Sintase
RcTPS36	e/f	136.1	3.2e-42	1	Diterpeno Sintase
RcTPS37	a	240.7	6.1e-74	1	Sesquiterpeno Sintase
RcTPS38	g	268.8	2.1e-82	2	Monoterpeno Sintase
RcTPS39	a	103.5	1.6e-32	1	Sesquiterpeno Sintase
RcTPS40	e/f	322.2	1.7e-98	1	Diterpeno Sintase
RcTPS41	a	329.2	1.2e-100	1	Sesquiterpeno Sintase
RcTPS42	a	182.9	1.7e-56	1	Sesquiterpeno Sintase
RcTPS43	a	104.2	9.4e-33	1	Sesquiterpeno Sintase
RcTPS44	b	342.5	8.6e-105	1	Monoterpeno Sintase
RcTPS45	b	115.0	6.7e-36	2	Monoterpeno Sintase
RcTPS46	b	297.5	3.7e-91	1	Monoterpeno Sintase

Tabela suplementar S6: Relação das RcTPS e existência (sim) ou ausência (não) de padrões de *motifs* conservados rastreados no alinhamento.

Nome do Gene	ID	AA	DDXXD	DXDD	NSE/DTE	RRX8W	Score	E-Value	Putative function
RcTPS1	29771.m000073	551	Sim	Não	Sim	Sim	1115.5	0	(<i>R</i>)-limonene synthase, putative
RcTPS2	29840.m000603	550	Sim	Não	Sim	Não	1145.2	0	(+)-delta-cadinene synthase isozyme A, putative
RcTPS3	29910.m000951	547	Sim	Não	Sim	Não	1101.7	0	(+)-delta-cadinene synthase isozyme A, putative
RcTPS4	30169.m006289	563	Sim	Não	Sim	Não	1156.7	0	1,8-cineole synthase, chloroplast precursor, putative
RcTPS5	29910.m000944	476	Sim	Não	Sim	Não	989.2	0	Terpinolene synthase, chloroplast precursor, putative
RcTPS6	30169.m006314	552	Sim	Não	Sim	Não	1124.8	0	Casbene synthase, chloroplast precursor, putative
RcTSP7	29827.m002614	551	Sim	Não	Sim	Sim	1145.2	0	Casbene synthase, chloroplast precursor, putative
RcTPS8	30156.m001703	584	Sim	Não	Não	Sim	1166.8	0	(<i>R</i>)-limonene synthase, putative
RcTPS9	30206.m000781	555	Sim	Não	Não	Não	1154.8	0	(<i>R</i>)-limonene synthase, putative
RcTPS10	30131.m007001	599	Sim	Não	Não	Não	1250.7	0	(<i>R</i>)-limonene synthase, putative
RcTPS11	30225.m001714	572	Sim	Não	Não	Não	1124.0	0	1,8-cineole synthase, chloroplast precursor, putative
RcTPS12	27688.m000060	587	Sim	Não	Não	Sim	1226.9	0	(<i>R</i>)-limonene synthase, putative
RcTPS13	29842.m003696	572	Sim	Não	Não	Não	1165.2	0	1,8-cineole synthase, chloroplast precursor, putative
RcTPS14	29842.m003692	521	Sim	Não	Não	Não	1083.9	0	(<i>R</i>)-limonene synthase, putative
RcTPS15	30206.m000792	222	Não	Não	Não	Não	420.2	5.8e ⁻¹⁵⁰	(<i>R</i>)-limonene synthase, putative
RcTPS16	29647.m002075	800	Sim	Sim	Sim	Não	1670.6	0	Ent-kaurene synthase A, chloroplast precursor, putative
RcTPS17	28166.m001038	558	Sim	Sim	Sim	Não	1159.4	0	(+)-delta-cadinene synthase isozyme A, putative
RcTPS18	29806.m000953	775	Não	Não	Sim	Não	1579.7	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS19	29806.m000952	778	Sim	Não	Sim	Não	1616.7	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS20	29806.m000947	816	Sim	Não	Sim	Não	1649.8	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS21	28670.m000104	808	Sim	Não	Sim	Não	1650.2	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS22	30169.m006283	601	Sim	Não	Sim	Não	1141.7	0	Casbene synthase, chloroplast precursor
RcTPS23	30169.m006276	614	Sim	Não	Sim	Não	1287.7	0	Casbene synthase, chloroplast precursor, putative
RcTPS24	30169.m006286	599	Sim	Não	Não	Não	1167.1	0	Casbene synthase, chloroplast precursor, putative

RcTPS25	30156.m001700	596	Sim	Não	Sim	Sim	1246.9	0	(<i>R</i>)-limonene synthase, putative
RcTPS26	29233.m000032	408	Sim	Não	Sim	Não	808.9	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS27	30206.m000787	226	Não	Não	Sim	Sim	427.9	6.3e ⁻¹⁵³	(<i>R</i>)-limonene synthase, putative
RcTPS28	30206.m000791	99	Não	Não	Não	Não	201.1	1.5e ⁻⁶⁷	geranyl-diphosphate cyclase, putative
RcTPS29	30156.m001716	275	Não	Não	Sim	Sim	528.9	0	D-cadinene synthase, putative
RcTPS30	29917.m001932	411	Não	Não	Não	Não	860.5	0	Alpha-bisabolene synthase, putative
RcTPS31	59739.m000007	143	Sim	Não	Sim	Não	294.7	5.9e ⁻¹⁰³	(<i>R</i>)-limonene synthase, putative
RcTPS32	29771.m000074	257	Sim	Não	Sim	Não	533.1	0	(<i>R</i>)-limonene synthase
RcTPS33	27716.m000050	408	Sim	Não	Sim	Não	850.9	0	Ent-kaurene synthase B, chloroplast precursor, putative
RcTPS34	27394.m000360	351	Sim	Não	Sim	Não	734.9	0	(<i>R</i>)-limonene synthase, putative
RcTPS35	28359.m000286	178	Não	Não	Não	Não	374.0	3.6e ⁻¹³³	D-cadinene synthase, putative
RcTPS36	29233.m000033	161	Sim	Não	Não	Não	308.5	6.2e ⁻¹⁰⁸	Ent-kaurene synthase, putative
RcTPS37	29673.m000927	271	Não	Não	Sim	Não	571.6	0	casbene synthase, putative
RcTPS38	29842.m003698	180	Sim	Não	Sim	Não	369.8	1.8e ⁻¹³¹	(<i>R</i>)-limonene synthase, putative
RcTPS39	29878.m000240	126	Sim	Não	Não	Não	264.6	1.2e ⁻⁹¹	D-cadinene synthase, putative
RcTPS40	29917.m001933	335	Sim	Não	Sim	Não	699.5	0	ent-kaurene synthase, putative
RcTPS41	30065.m001152	282	Sim	Não	Sim	Não	588.6	0	Casbene synthase, chloroplast precursor, putative
RcTPS42	30065.m001155	156	Sim	Não	Não	Não	327.0	2.3e ⁻¹¹⁵	casbene synthase, putative
RcTPS43	29878.m000241	98	Não	Não	Não	Não	204.1	7.5e ⁻⁶⁹	D-cadinene synthase, putative
RcTPS44	30156.m001715	288	Sim	Não	Sim	Não	593.6	0	1,8-cineole synthase, chloroplast precursor, putative
RcTPS45	30206.m000785	123	Não	Não	Não	Não	258.1	2.9e ⁻⁸⁹	(<i>R</i>)-limonene synthase, putative
RcTPS46	30206.m000786	192	Sim	Não	Sim	Não	400.2	5.4e ⁻¹⁴³	(<i>R</i>)-limonene synthase, putative

Figura suplementar S7: Alinhamento múltiplo utilizando sequencias TPS das espécies *Ricinus communis*L (29771.m000073), *Arabidopsis thaliana* (AT3G25810.1), *Populus trichocarpa* (Potri.T000300), *Solanum lycopersicum* (Solyc01g105960), *Oryza sativa* (LOC_Os01g23530) e *Selaginella moellendorffii* (412139_SmTPS) com destaque nas similaridades entre as sequencias de aminoácidos.

29771.m000073	1	-----	0
AT3G25810.1	1	-----	0
Potri.T000300	1	-----	0
Solyc01g105960	1	-----	0
LOC_Os01g23530	1	-----	0
412139_SmTPS	1	MSHVLLISFQRIGSGSLQLFGGTNLKSNQAFNFRKVS SCASKPAEAEVKGWLMKLAGS	60
29771.m000073	1	-----	0
AT3G25810.1	1	-----	0
Potri.T000300	1	-----MAL	3
Solyc01g105960	1	-----	0
LOC_Os01g23530	1	-----	0
412139_SmTPS	61	KGKISPSSTYDTAWVARIPSDSDSSLPEFFEALEWIINSQLPDGSDDDCTKFLRENMIKL	120
29771.m000073	1	-----MDF	4
AT3G25810.1	1	-----M-ATLCI---GSAPI---YQNAICHH---FRL---QRPRRFISKAMKTIIP	38
Potri.T000300	4	THFA---SSF-CTLPLTASPSRAL--AGFVSSKN---Y-----RSLP-THVRCWA	44
Solyc01g105960	1	-----M-YKLEM-----TMSISKS-----NLI-SKLEVP-K	23
LOC_Os01g23530	1	-----MPADHMEKTVF	11
412139_SmTPS	121	NHDDGDFLSGFVFTFMMLEHAKRVGLKLPDYDTEFTLLEISTEKLAKIPLDLRHSTPT	180
29771.m000073	5	IN-SN-----ETIIRRSAN--YRPFWDVDFVQSLKSEFVGELN-IKRIDK	46
AT3G25810.1	39	DA-YP-----LDLRRSSGN--YQPSWDHSYLLSTENKYVNEKE-VITRHV	80
Potri.T000300	45	TE-AA-----GQIVRSAN--YQTSIWEYDFVQSLTSKYKGEFY-TARSEK	86
Solyc01g105960	24	SC-IS-----NVPIRSSGN--YQPSIWDYHQLSLKNHYSDEKFM-MRRRNE	65
LOC_Os01g23530	12	TV-VFARFAN-----SEGERKSTH--FHPSLWGDFFDNYEFPFAPQACMKGRARV	59
412139_SmTPS	181	TLIYSLGELGQLEIDWQKILKLSNGLSFLSFSSTACVIRKTKDK-----KSLQV	291
29771.m000073	47	EDNK---MMLN-K-TMAPSQEIEIDTQRLGAYHFGDEIKRIWKSITYNS---HRN-	97
AT3G25810.1	81	KKVK---KMLEEVEIKSRLEKIEIDDLQKLGVSYHFEQENNIITNFHLE---NGKN	134
Potri.T000300	87	NANIR---MMLA-N-ASKPLDQELIDALQRLGVSYHFEQEKSTKSLFYK---NHV-	137
Solyc01g105960	66	MEVK---IMLSDR-NMKLQEQEIEIDLQRLGVSYHFEQETYSIINNLSDK-----	114
LOC_Os01g23530	60	EEVR---TLKGL--KEVQIIEIDTMAQRLGDMYVNEEINELHFVYSS-----	107
412139_SmTPS	232	KAMKHQNYAVSCHYFIDLFESRIVVDTEERGLDLFRDEIKKAVLDNLYSFWTNEGIG	291
29771.m000073	98	DNTMKEDLHTIALQRLGQHGYNV-SQEIIDFDELGNFKECL---HEDIEGMSL	152
AT3G25810.1	135	IWKCKEEDLHATALEERLQHGFGV-SEDIIDVIEDKIESN--TF--KSDNITSEITE	189
Potri.T000300	138	ENTKIVHDLVATALEERLQHGKQV-PQEVNHFHDEQGNFRAWI---HEDLKGMIFF	192
Solyc01g105960	115	---GSKRDHLYAKALEERLQHGKQVVSQETGGFVNDITGFGEIH---HNEITKGMILE	169
LOC_Os01g23530	108	---DYNDKDLNLVSLRFLKQKGNV--SSGIELSFDKDNENF-----IVDTRSLNLI	157
412139_SmTPS	292	WGSTCIVNDIEDTAMARILRMHGYSV-SPDANQFWLPQDKFCCFVGLSHVSAHINI	350
29771.m000073	153	ASYLEEGENI--EVAREFAASCYKQV-NK---DQLSMIVSHSSEVFLHWRMFR	207
AT3G25810.1	190	ASYLESTKSDTKHVVIRPFATEQRNFVDE-SETYINIMREMAIHALEIYHWRMFR	248
Potri.T000300	193	ASYLEVLESESL--EDARDFITNLEKVKKCNF---SEYSKMVSHALELAWRMRLR	248
Solyc01g105960	170	ASYLEALIESEKEL--ELARNTEEHREYLDQNKMDVQGNVFLVHHALELHWRMRLR	228
LOC_Os01g23530	158	ASYLELVYGEKVI--GEAATFTSRISGVLESS---SISTEVSFAEATTFARARI	211
412139_SmTPS	351	ASQVDFPHEETIKR--TFKISHVYLVN-ESSHTSATKKNMSEVIVELANVHDCDLFR	408
29771.m000073	208	ETRWFDIMEKKQ-GM-----NPLLEAKLDFNIVQATYHEDLKYVTSWRN	255
AT3G25810.1	249	ETRWVDAIEKKH-DM-----NLFLEAFKIDFNIVQATAHQEDVYVYCWKE	296
Potri.T000300	249	LESNWFNVETKT-DM-----EPVLEAKLDFNIVQAIHQEDLRKSSRWKR	296
Solyc01g105960	229	LETWFWNYKKRQDKM-----IPFLEAKLDFNIVQAAHIEDLKYVVARWKE	277
LOC_Os01g23530	212	LEMRRNYEPINEIEA-TR-----NETLEAFKLNFNLLLYCEEKLNITLWKE	259
412139_SmTPS	409	LYNNAYKHSIDDFWVAKTYRPLVKNKVLLEAKRYAQCCSYQPSLKLILVKNHHS	468
29771.m000073	256	TGLGKLSFA---RDRVENFLWVSNVFPQGYFRSEKVNSEIIVDDIYVYGI	311
AT3G25810.1	297	TGLGSQLHFV---RDRVENFLWVSNVIEPQGYFRSVAIVAAIIVDDIYVYGI	352
Potri.T000300	297	TGLGKLDFA---RDRVENFLWVSNVIEPQGNCRMLKVNSEIIVDDIYVYGI	352
Solyc01g105960	278	TCLAENLFFA---RDRVENFLWVSNVIEPQGYFRSEKVNSEIIVDDIYVYGI	333
LOC_Os01g23530	260	LKVKNLSFS---RDRVENFLWVSNVIEPQGYFRSEIIVDDIYVYGI	315
412139_SmTPS	469	SHFEDPSTRFKANIMLFYIMVVICATFHEQEGPLTFEKKTCNTLEFELISAKS	528
29771.m000073	312	LDELQLEINAVERVNVANDQLPEYMKLCLLAIHNSINENAYDALRE---QGFIIPYL	367
AT3G25810.1	353	PELELITAMVQNDINRDELPEYMKLCLLTFINEINAGCDVLKC---KNIDIPYF	408
Potri.T000300	353	LDELLELTDAVVNDLNCNDRLPDYMKCLLAFNSVNEINAYDILKY---QGVDLPYL	408
Solyc01g105960	394	LDELQITHAERNSEIDEDRLPDNMKCYALDNFINOADDAAFE---QGFIPSPYK	389
LOC_Os01g23530	316	TESMLLAEANRWDESAAGLLPEYIRGFVAYLLATFDSFE---ELGPEK	363
412139_SmTPS	529	IRELDSQNMESWDTSFISQKE--YKIIQEEYNTLVMTKIASKINENLSLEFHKYL	586
29771.m000073	368	KKAWADCKSYLVEAKWYIYGTETLQYMDNAWESISAPVIVHAYFLEGSPTNEALK	427
AT3G25810.1	409	KKSWADCKSYLVEAKWYKGGYKSEVEYMDNAWESISAPTIMHFYCAFSQDISVQILE	468
Potri.T000300	409	KKAWADCKSYLLEAKWYFSGYETLQYMDNAWESISAPVIVHAYFYVSNPTTEASQ	468
Solyc01g105960	390	RNSWADCKSYLLEAKWYHSGYETSMYMDNAWESISAPVIVHAYFLVAVNPKLEASQ	449
LOC_Os01g23530	364	RYRLKRVQAYTKLWRDEDTKTELEHFVSNRSGGFTLAASVFGMDIATKDF	423
412139_SmTPS	587	LAIYTKIKSGIIEARKWIKQSGYGFYEMNAKVSRIATHVLMGILFCGDPLELLN	646
29771.m000073	428	SLK-EYFDIHWSSMILRFADLGGSSDELKRGDNPMSQCYIYET-GVSELKAREHQY	485
AT3G25810.1	469	SLVQQQQVVRCSATVLRANDLAISPELARGDVLKSVCCMNET-GVSEEARHVVQ	527
Potri.T000300	469	FMD-EYFDIHWSSMILRFADLGGSTDELKRGDISKSCCMHEA-GVSEKAREHNRN	526
Solyc01g105960	450	VLENNYHDIRCSALTLRANDEGSSDELKRGDVPFSCCMNET-QAEEERARQVRL	508
LOC_Os01g23530	424	EMLISYSPLEKTFIVLRSHVSNKSGTSGHYATIKSCCMNET-QAEEERARQVRL	482
412139_SmTPS	647	IYY--DSKMLILGSISSRCHNDQYKLEMKLGSAGVSCMKHHPGAREEDVAWVLS	704
29771.m000073	486	IGEWKKNKEREY---IDSPFSKIEVATNLARMAQCNVQHGDD---HGIEEG-ET	537
AT3G25810.1	528	MSHWDENNYEARTVARISSLSRGEVETAMNLARMSQCNVQHGDD---HGCPKAKI	583
Potri.T000300	527	LEENWKKINDY-QF---DNPCISQTFIGIAMNLARMAQCNVQHGDD---HGCVHLE-ET	577
Solyc01g105960	509	LISQWKKNEAHL---AADPFKIEVTCAMNLARMAQCNVQHGDD---HGGNNS-TT	560
LOC_Os01g23530	483	LEDQWKNVEHCTNP---IDDOPLVIEQTVVNFARTVITVTHGDA---FTSSHT--I	533
412139_SmTPS	705	LEKTKKQLEIYFTE---KDLFKNKRFSPD---MMETNKAQKDLFKNPNK-EL	754
29771.m000073	538	KDHVLSLLVKPIEM-----	551
AT3G25810.1	584	VDRVQLLVDPILD-----	598
Potri.T000300	578	KDRVKSLLIKPL-----	599
Solyc01g105960	561	KHVMALLFESVFLGHKHSAAEKEDHSMVNVYREKPMI	597
LOC_Os01g23530	534	KEMSSYVWPIQV-----	548
412139_SmTPS	755	QSMIEFPLET-----	765

Figura suplementar S8: Alinhamento múltiplo utilizando sequencias TPS das espécies *Ricinus communis*L (29771.m000073), *Arabidopsis thaliana* (AT3G25810.1), *Populus trichocarpa* (Potri.T000300), *Solanum lycopersicum* (Solyc01g105960), *Oryza sativa* (LOC_Os01g23530) e *Selaginella moellendorffii* (412139_SmTPS) com destaque no padrão de *motif* canônico conservado DDXD entre as espécies.

29771.m000073	1	-----	0
AT3G25810.1	1	-----	0
Potri.T000300	1	-----	0
Solyc01g105960	1	-----	0
LOC_Os01g23530	1	-----	0
412139_SmTPS	1	MSKVLISFQRIGSGSLQFGGTLNLSNAQPNRFRKVSVCASKPAEAEVKGWMLKLAGS	60
29771.m000073	1	-----	0
AT3G25810.1	1	-----	0
Potri.T000300	1	-----MAL	3
Solyc01g105960	1	-----	0
LOC_Os01g23530	1	-----	0
412139_SmTPS	61	KGKISPSYDAMVARIPSDSDSSLPEFFPEALEMIINSQLPDGSDDDCTKFLRENMIKL	120
29771.m000073	1	-----MIDP	4
AT3G25810.1	1	-----M-ATLCI---GSAPI--YQACIHN---FRL---QRFRRFISKAMTKTIP	38
Potri.T000300	4	THFA---SSF-CTLPTLASFSRAL--AGFVSSKN---Y-----RSLP-TKVRCMVA	44
Solyc01g105960	1	-----M-YKLEM-----TMSISKS-----NLI-SKLEVF-K	23
LOC_Os01g23530	1	-----MFADHMEKTPV	11
412139_SmTPS	121	NHDDGDFLSGFEVTFMMLHEAKRVGLKLPYDTEFTALLEISTEKLAKIFLDRLHSTPT	180
29771.m000073	5	IM-SN-----ETIIRRSAN--YRPFIDFDFVQSLKSEFVGEIN-IKRIDK	46
AT3G25810.1	39	DA-YP-----LDLRRRSN--YQPSWDSHLYLLSTENKYVNEKE-VITRHV	80
Potri.T000300	45	TE-AA-----GQIVRRSAN--YQTSIWEYDFVQSLTSKYVGEFY-TARSEK	86
Solyc01g105960	24	SC-IS-----NVPVRSN--YQPSIWDYHQLSKLKHYSDEKF-MRRARNE	59
LOC_Os01g23530	12	TU-VEARPPN-----SEGERKSTN--FHPSLWGFLLNVEPPTAFQACMKGRAEV	59
412139_SmTPS	181	TLLYSLEGLQGLEIDWQIKLQKSDKSLFSSFSSTACVYLKTKDK-----KSLQY	231
29771.m000073	47	LKEDVK---MMLN-K-TMAPSDQLELIDTLQRLGLAVHFGDEIKRIVKSIYNS---HRN-	97
AT3G25810.1	81	LKKRVK---KMLEEVEKSRLEKLELIDDLQRLGVSVHFQEINNILTNFHE---NGKN	134
Potri.T000300	87	LKANIR---MMLA-N-ASKPLDQLELIDALQRLGLSYHFVEIKSTLKSIFYK---NHV-	137
Solyc01g105960	66	LKMEVK---IMLSDR-NMKQLELEIDNQLRGLSYHFDEIYSILNLSDK-----	114
LOC_Os01g23530	60	LREEVR---TLKGL--KEVQKILDIIMAIQRLGLDMYVNEINELHVFVSS-----	107
412139_SmTPS	232	LQKAMKHQYAVSCHVPIDLFESLRVVDTIERLQIDLFFRDEIKAVLDNLYSFWTNEGIG	291
29771.m000073	98	-DNWIKEDLHTIALQFRLLRQHGYNISQEIFDFRDELGNFKECL---HEDIEGMLSL	152
AT3G25810.1	135	IWKCKEEDLHATALEFRLLRQHGFGV-SEDFPDIIDKIESN--TF--KSDNITSITTL	189
Potri.T000300	138	-ENTKTVHDLVATALEFRLLRQHGFKV-FQEVFNHFQDEQGNFRAM---HEDLKGMLFL	192
Solyc01g105960	119	---GSKRDHLVAKALEFRLLRQHGYNISQEIFDFRDELGNF---HVEDTKGMLYL	169
LOC_Os01g23530	108	---DYNKDLNLEVTALFRYLLRQHGYNM-SGDFPISFKDNEGF---IVDDTRSLNLL	157
412139_SmTPS	292	WGSTCIVNDIEDTAMAFRLRMHGYSV-SPDAFQFWLFGDKFCQFVSELSEVSAAMLN	350
29771.m000073	153	YEASVLEEGENIL-EVAREFAASCCLKYIQV-NK---DQLLSMIVSHSLEVPLHWRMRP	207
AT3G25810.1	190	YEASVSTKSDTKLHVIRPFAPTEQIRNFVDE-SEYINIMLREMAHALEIPYHWRMR	248
Potri.T000300	193	YEASVLEEGESIL-EDARDFTTKNLEKYVKKCNP---SEYLSKQVSHALELPLAWRMR	248
Solyc01g105960	170	YEASFLAIEGKEL-ELARNLTHELREYADQKNNDVQNLVELVHHALEPLHWRMR	228
LOC_Os01g23530	158	YNAANLRVYGEKVL-DEAATFTSRLEGVLESS-----DSILSTEVSFALEAPIFRARRI	211
412139_SmTPS	351	YRASQVDFNEETLTK-TFKYSHDVLNV-ESSHTSATKKNLMGEVITELANPFHDLCP	408
29771.m000073	208	LETRWFIDIEKKQ-GM-----NPLLELAKLDFNNVQATYHEDLKYVTSWWRN	255
AT3G25810.1	249	LETRWFIDAYEKKH-DM-----NLFLEAFKIDFNIVQTAHQEDVKVYSCWWR	296
Potri.T000300	249	LESNWFINVYETK-DM-----EPVLELAKLDFNNVQATYHEDLKHSSRWKR	296
Solyc01g105960	229	LETRWFIDIEKKQ-QDKM-----IPFLELATLDFNIVQAAHIEDLKVVARWKE	277
LOC_Os01g23530	212	VENRNYTPYETEATR-----NETILEFAKLNFLQLLICEELNKITLWKE	259
412139_SmTPS	409	IYNNAYIKHYSIDDFWVAKTIYRLPLVKNKVLLELANRYAQCCQSHPSELKILVKNWHS	468
29771.m000073	256	TGLGKLSFA---RDRLMENLWTVGNVFPQGFYFRIRITKVNLSLITV	311
AT3G25810.1	297	TGLGSLQHFV---RDRIVENYFWTVGVIPEQGFYIRRIVAIVAALITV	352
Potri.T000300	297	TGLGKLDFA---RDRLVENLWTVGVIPEQGFYFRIRITKVNLSLITV	352
Solyc01g105960	278	TCLAENLFFA---RDRLVENFWTVGVIPEQGFYFRIRITKVNLSLITV	333
LOC_Os01g23530	260	LKVKSLSFS---RDRIVENYFWNMGALVEPHYSHSRILITVAVFTMI	315
412139_SmTPS	469	SHFEDIPSTRFKANINMLFYIYVICATFHEQFGPLRTFFTKCCMTL	528
29771.m000073	312	LDELQFLTNAVERDWNAMDQLPEYMKLCLFALHNSINEMAYDALRE---QGFIIPYL	367
AT3G25810.1	353	PEELEFTAMVQNDINRLEDELPEYMKLCLFLTFNEINAMGCDVLKCC---KNIDVIPPY	408
Potri.T000300	353	LDELELFTDAVVRWDLNCDRLPDYMKLCLFLTFNSVNEMAYDILKY---QGVDFILPYL	408
Solyc01g105960	334	LDELELFTHAIERWSDIDELDRLPDNNKQCYALDNFINQLADDAFEE---QGFIIPYL	389
LOC_Os01g23530	316	TEESMLLAEAINRWDESAIGLLPEYIRGFYAYLLKTFDSFEE-----ELGPEK	363
412139_SmTPS	529	IRELDSLQNMIESWDTSFSLQKE--YKIIQEFYNTILVMTKIASKINENLSLEFTHKYL	586
29771.m000073	368	KKAWADLCKSYLVEAKWYIYGYFTLQEYMDNNAWISISAFVILVHAFVLEGSFITNEALK	427
AT3G25810.1	409	KKSWADLCKAYLVEAKWYKGYKFSVEEYMQNNAWISISAFVILVHAFVLEGSFITNEALK	468
Potri.T000300	409	KKAWADLCKSYLVEAKWYIYGYFTLQEYMDNNAWISISAFVILVHAFVLEGSFITNEALK	468
Solyc01g105960	390	RNSWADLCKSYLVEAKWYHYSQYFSMEYMDNNAWISISAFVILVHAFVLEGSFITNEALK	449
LOC_Os01g23530	364	RVRLKRLVQAYTKELWRDDEYTKTELEHFEVSMRSGGFTLAAASVFGMDDIATKDF	423
412139_SmTPS	587	LAIYTKLKSGLIIEARWKIQGYIFGFEYEMENAKYSIAIATHVLMGLFCGDLPLTEELLN	646
29771.m000073	428	SLK-EYPDIIQWSSMILRFADDLGTSSDELKRGDNPKSIQCVIYET-GVSELKAREHIQY	485
AT3G25810.1	469	SLVQQQDQVVRCSATVLRLANDLATSPDELARGDVLKSVQCYMHET-GVSEEARATHWQQ	527
Potri.T000300	469	FMD-EYPDIIQWSSMILRFADDLGTSSDELKRGDVKSIQCYMHET-GVSEEARATHWQQ	526
Solyc01g105960	450	YLENNYHDIIRCSALILRLANDLGTSSDELKRGDVPKSIQCYMHET-QASEEARQYIRL	508
LOC_Os01g23530	424	EWLSYPSLFTKFDIIVRLSNDIVSNKREQTGDHYASTIQCYMKEH-GTITHETVQRLRE	482
412139_SmTPS	647	IYI--DSKLLKLSIISRICNDIQYKIEMLKGSAGQVSCYMKDFHGATEEDAVAYLQS	704
29771.m000073	486	LIGETWKKINKEREY---IDSPFSKIFIEVATNLARMAQCMYQHGDS---HGIEDG-ET	537
AT3G25810.1	528	MISHTWDEMHYARTVARISLRSRGTVEAMNLRMSQCMYQHGDS---HGCPFKAKI	583
Potri.T000300	527	LIENTWKKINDI-QF---DNFCSQTFEGIAMNLRMAQCMYQHGDS---HGNGHL-ET	577
Solyc01g105960	509	LISQWKKLNEAHL---AASPFEKIFTCAMNLRMAQCMYQHGDS---HGNGHS-TI	560
LOC_Os01g23530	483	LIEDSWKDMVEHCINP---IDDPFLVQGVVNFARTVITWYTHGDA---FTSHT--I	533
412139_SmTPS	705	LLEKTKQLNEIYFTE---KDLFKNIKRFSFD---MMFTYKNAKQDLKFNPNK-EL	754
29771.m000073	538	KDHVLSLLVKPIPM-----	551
AT3G25810.1	584	VDRVQTLVVDPIPLD-----	598
Potri.T000300	578	KDRVKSLLIKPL-----	589
Solyc01g105960	561	KNHIMALLFESVPLGHKHSAAEKEDHSVNVYREKEMI	597
LOC_Os01g23530	534	KEMISSIYVPIQV*-----	548
412139_SmTPS	755	QSMIEFCLET*-----	765

DDXD

Tabela suplementar S7: Lista de elementos cis- reguladores associados aos genes RcTPS relatados neste estudo e seus respectivos quantitativos de elementos cis. Ácido abscísico (AA), auxina (AX), giberelina (GB), metil jasmonato (MeJa), respostas a luminosidade (L), MYB, MYC, indução anaeróbica (IA), resposta de defesa ou estresses (DE), expressão induzida por restrição hídrica(ES), baixas temperaturas (TB), ácido salicílico (AS), expressão de meristema (EM), AT-rich, expressão de zeína (EZ), controle circadiano (CC) e Respostas a feridas/herbivoria.

Gene name	ID gene	Subfamília	AA	Ax	GB	MeJa	Luz	MYB	MYC	IA	DE	ES	TB	AS	EM	AT rich	RZ	WR	CC
RcTPS1	29771.m000073	b	2	1	0	0	9	3	0	0	0	0	0	0	0	0	0	1	1
RcTPS2	29840.m000603	a	1	0	1	0	7	0	0	0	0	0	0	0	0	1	0	0	0
RcTPS3	29910.m000951	g	0	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0	2
RcTPS4	30169.m006289	b	0	1	0	0	7	0	0	1	0	0	0	1	0	0	0	0	0
RcTPS5	29910.m000944	e/f	3	0	1	1	2	0	0	0	0	0	0	1	1	0	0	0	0
RcTPS6	30169.m006314	g	4	0	0	2	6	0	0	0	0	1		1	1	1	0	0	0
RcTPS7	29827.m002614	a	0	0	0	0	6	0	0	0	0	0	0	0	2	0	0	0	0
RcTPS8	30156.m001703	b	0	1	0	2	7	0	0	0	0	1	0	2	0	0	0	0	0
RcTPS9	30206.m000781	g	1	0	0	0	3	0	0	0	1	0	0	0	0	1	0	0	0
RcTPS10	30131.m007001	b	0	0	0	0	5	0	0	1	2	0	0	2	0	0	0	0	0
RcTPS11	30225.m001714	b	0	1	0	0	8	0	0	2	0	0	0	1	1	0	1	0	0
RcTPS12	27688.m000060	e/f	3	1	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0
RcTPS13	29842.m003696	a	1	0	0	0	11	0	0	0	1	0	1	1	0	0	0	0	0
RcTPS14	29842.m003692	a	1	0	0	0	10	0	0	0	0	2	0	0	0	0	0	0	0
RcTPS15	30206.m000792	b	0	0	0	0	5	0	0	1	0	0	0	1	0	0	0	0	0
RcTPS16	29647.m002075	c	2	0	1	3	9	0	0	0	0	0	0	0	0	0	1	0	0
RcTPS17	28166.m001038	a	1	0	1	0	4	0	0	0	0	0	0	0	1	1	0	0	0
RcTPS18	29806.m000953	c	0	0	2	0	6	0	0	0	2	2	0	0	0	0	2	0	0
RcTPS19	29806.m000952	c	0	0	0	1	6	0	0	1	1	1	0	0	0	0	0	0	0
RcTPS20	29806.m000947	a	0	0	0	2	5	0	0	1	0	1	1	0	1	0	0	0	0
RcTPS21	28670.m000104	b	0	0	0	0	4	0	0	0	2	0	0	0	1	0	0	0	0
RcTPS22	30169.m006283	a	5	1	0	2	16	0	0	1	0	1	1	0	0	0	0	0	0
RcTPS23	30169.m006276	a	1	0	1	0	9	1	0	0	0	0	0	0	0	1	0	0	0

RcTPS24	30169.m006286	a	4	0	0	0	12	2	0	0	0	0	0	0	0	0	0	0	0
RcTPS25	30156.m001700	b	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
RcTPS26	29233.m000032	g	1	1	0	1	2	0	0	1	1	1	0	0	0	0	0	0	0
RcTPS27	30206.m000787	b	1	0	0	1	3	0	0	1	0	0	0	0	1	0	2	0	0
RcTPS28	30206.m000791	b	5	1	2	1	4	2	0	0	0	0	0	0	1	0	1	0	1
RcTPS29	30156.m001716	b	0	0	0	1	3	0	0	0	0	0	0	1	0	0	0	0	0
RcTPS30	29917.m001932	b	0	1	0	0	5	0	0	1	1	1	0	1	0	0	0	0	0
RcTPS31	59739.m000007	g	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RcTPS32	29771.m000074	b	0	0	0	0	5	0	0	0	0	0	0	0	0	1	0	0	0
ARcTPS33	27716.m000050	g	0	1	0	0	6	0	0	2	0	0	0	1	0	0	0	0	0
RcTPS34	27394.m000360	g	0	2	1	0	5	0	0	1	1	1	1	1	1	0	0	0	0
RcTPS35	28359.m000286	a	0	0	2	0	6	0	0	0	0	0	1	0	0	0	1	0	0
RcTPS36	29233.m000033	a	0	0	0	0	8	0	0	2	0	0	0	0	1	0	0	0	1
RcTPS37	29673.m000927	a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RcTPS38	29842.m003698	c	0	0	0	1	4	3	0	1	0	3	0	0	0	0	0	0	0
RcTPS39	29878.m000240	a	1	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0
RcTPS40	29917.m001933	g	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1
RcTPS41	30065.m001152	g	2	0	0	2	6	0	0	0	0	0	0	0	1	0	0	0	0
RcTPS42	30065.m001155	a	0	0	2	1	4	0	0	0	0	0	0	0	0	0	1	0	0
RcTPS43	29878.m000241	a	0	0	0	2	6	1	0	0	0	1	0	0	0	0	1	0	0
RcTPS44	30156.m001715	g	0	0	1	0	8	0	0	0	1	0	0	0	0	0	0	0	0
RcTPS45	30206.m000785	b	0	0	0	0	5	0	0	2	0	0	0	0	0	0	0	0	0
RcTPS46	30206.m000786	e/f	0	0	0	0	5	0	0	1	0	0	0	1	0	0	0	0	0