



**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA**

RACHEL ALVES MAIA

**GENES DUPLICADOS DA OXIDASE TERMINAL DO PLASTÍDIO (PTOX) EM
SOJA [*Glycine max* (L.) Merrill]: CARACTERIZAÇÃO, EXPRESSÃO E
AFINIDADE DAS ISOFORMAS PROTEICAS PELO LIGANTE**

FORTALEZA

2018

RACHEL ALVES MAIA

GENES DUPLICADOS DA OXIDASE TERMINAL DO PLASTÍDIO (PTOX) EM SOJA
[*Glycine max* (L.) Merrill]: CARACTERIZAÇÃO, EXPRESSÃO E
AFINIDADE DAS ISOFORMAS PROTEICAS PELO LIGANTE

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Bioquímica. Área de concentração: Bioquímica Vegetal.

Orientador: Prof. Dr. José Hélio Costa.

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M188g Maia, Rachel Alves.

Genes duplicados da oxidase terminal do plastídio (PTOX) em soja [*Glycine max* (L.) Merrill] :
Caracterização, expressão e afinidade das isoformas proteicas pelo ligante / Rachel Alves Maia. – 2018.
98 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação
em Bioquímica, Fortaleza, 2018.

Orientação: Prof. Dr. José Hélio Costa.

1. Commodity. 2. Cloroplasto. 3. Fotossíntese. 4. RNA-seq. 5. Docking molecular. I. Título.

CDD 572

RACHEL ALVES MAIA

GENES DUPLICADOS DA OXIDASE TERMINAL DO PLASTÍDIO (PTOX) EM SOJA
[*Glycine max* (L.) Merrill]: CARACTERIZAÇÃO, EXPRESSÃO E
AFINIDADE DAS ISOFORMAS PROTEICAS PELO LIGANTE

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Bioquímica. Área de concentração: Bioquímica Vegetal.

Aprovada em: 19/02/2018.

BANCA EXAMINADORA

Prof. Dr. José Hélio Costa (Orientador)
Universidade Federal do Ceará (UFC)

Prof^a. Dr^a. Birgit Arnholdt-Schmitt
Universidade de Évora (UE)

Prof^a. Dr^a. Kátia Daniella da Cruz Saraiva
Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB)

A Deus.

Aos meus pais, Antonio e Regina.

AGRADECIMENTOS

A Deus, que me dá toda força e a Nossa Senhora, pela intercessão.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo fomento à pesquisa.

À Universidade Federal do Ceará e ao Programa de Pós-graduação em Bioquímica, pelo apoio institucional.

Ao Prof. Dr. José Hélio Costa, pela orientação e ensinamentos.

À Prof^{ra}. Dr^a. Birgit Arnholdt-Schmitt, por aceitar participar da minha banca examinadora e pelas valiosas contribuições.

À Prof^{ra}. Dr^a. Kátia Daniella da Cruz Saraiva, por aceitar participar da minha banca examinadora e pela ajuda imprescindível ao longo da execução desse trabalho. Obrigada pela paciência em ensinar e por sempre se mostrar disponível para tirar minhas dúvidas.

Aos professores Bruno Lopes e Cléverson Diniz, por prontamente se disponibilizarem a participar como suplentes da minha banca.

Aos colegas do Laboratório de Genômica Funcional e Bioinformática, de forma especial ao André, à Thais e ao Clesivan, por acompanharem mais de perto o desenvolvimento desse trabalho e sempre estarem disponíveis seja para esclarecer dúvidas, discutir ideias, corrigir textos, e principalmente pela amizade de vocês. À Karine, por sempre se mostrar acessível e me ajudar com muita boa vontade no *docking* molecular. À Beatriz, por todo apoio e por ajudar a revisar o texto. Aos demais colegas Fernanda, Mathias, Moaciria, Lyndefania, Sarah e Rodrigo, pelo apoio, pela convivência e momentos de descontração.

À minha família, por ser meu porto seguro em todos os momentos. Especialmente aos meus pais, Regina e Antonio, à minha irmã Jacqueline e ao meu querido sobrinho Pedro Paulo, que me traz tanta alegria. Obrigada por me compreenderem e pela paciência de vocês.

A todos os amigos que me deram apoio com orações ou palavras de ânimo, dizendo que “vai dar certo”. Obrigada pela amizade e por tornarem tudo mais leve.

Agradeço a todos que, de forma direta ou indireta, contribuíram para a realização desse trabalho.

“Descobri que há uma harmonia maravilhosa nas verdades complementares da fé e da ciência. O Deus da Bíblia é também o Deus do genoma. Deus pode ser encontrado na catedral e no laboratório. Investigando a criação incrível e majestosa de Deus, a ciência pode na verdade ser uma forma de louvor”.

Francis Collins

RESUMO

A soja [*Glycine max* (L.) Merrill] ocupa lugar de destaque na economia mundial, contudo, sua produtividade tem sido afetada por estresses ambientais e uma das formas de se contornar o problema é desenvolvendo marcadores moleculares voltados ao melhoramento genético. A oxidase terminal do plastídio (PTOX) é encontrada nos cloroplastos das plantas superiores e catalisa a oxidação do plastoquinol e a redução do oxigênio molecular a água. Diversos estudos mostram que a PTOX responde a diferentes estresses, portanto, avançar no entendimento de como ocorre a regulação da expressão da enzima na planta é necessário para desenvolver um potencial marcador molecular. Estudos prévios revelaram a presença de dois genes codificando a PTOX no genoma da soja, entretanto, a funcionalidade de ambos não foi comprovada experimentalmente, tampouco se tem dados sobre o papel desses parálogos na planta. O objetivo deste trabalho foi caracterizar e analisar o perfil de expressão dos dois genes *PTOX* em *G. max*, além de investigar a afinidade das respectivas isoformas proteicas pelo ligante. Inicialmente, as sequências que codificam para a PTOX foram identificadas e anotadas em *G. max* e em outras espécies leguminosas. A expressão dos genes foi avaliada em diversos órgãos de *G. max* durante o desenvolvimento e sob condições de estresse por meio de PCR em tempo real e da análise de dados transcriptômicos do GenBank. A modelagem das estruturas tridimensionais das duas isoformas proteicas foi seguida do ensaio de docagem molecular entre as isoformas e o seu ligante. Os resultados mostram que os genes *PTOX 1* e *PTOX 2* são funcionais e expressos em praticamente todas as partes da planta avaliadas e apresentam alta similaridade entre si. Contudo, durante o desenvolvimento de cotilédones, folhas e raízes o perfil de expressão dos genes é distinto, com *PTOX 1* prevalecendo nas partes mais jovens e *PTOX 2* tendo sua expressão aumentada à medida em que estes órgãos envelhecem. Já em condições estressantes, *PTOX 1*, de forma geral, é regulado positivamente enquanto os níveis de *PTOX 2* sofrem regulação negativa ou se mantêm estáveis. Entretanto, quando um estresse muito severo é aplicado, a expressão de *PTOX 2* aumenta consideravelmente. A afinidade pelo ligante também diferiu entre as isoformas, sendo maior para *PTOX 1*. É possível concluir que a duplicação dos genes *PTOX* no genoma da soja conferiu ganho de função para a planta. Também sugere-se que o gene *PTOX 1* esteja associado à eficiência fotossintética e que o gene *PTOX 2* esteja relacionado com a senescência de folhas e cotilédones.

Palavras-chave: Soja. Duplicação gênica. PTOX. Expressão gênica. RNA-seq. Docking.

ABSTRACT

Soybean [*Glycine max* (L.) Merrill] occupies a prominent place in global economy, however, its productivity can be affected by environmental stresses and one way to solve the problem is developing molecular markers for use in breeding programs. Terminal oxidase of plastid (PTOX) is an enzyme found in chloroplasts of higher plants that catalyzes the oxidation of plastoquinol and the reduction of molecular oxygen to water. Several studies link PTOX to different types of stresses responses, therefore, to advance in understanding how protein expression regulation occurs in plant is necessary to the development of a potential molecular marker. Two genes encoding PTOX have been identified in soybean genome; however, their existence has not been proven experimentally, either there is not any data about the effects of these paralogs on the plant. The objective of the present study was to characterize and to analyze the expression profile of the two *PTOX* genes in *G. max*, and to investigate the affinity of the respective protein isoforms for the ligand. Initially, sequences encoding PTOX were identified and annotated in *G. max* and other legume species. Expression of *PTOX* genes was evaluated in several organs of *G. max* during development and under stress conditions by real-time PCR and by analyses of transcriptomic data available in GenBank. The modeling of the three-dimensional structures of the two protein isoforms was followed by molecular docking assay between the isoforms and their ligand. The results reveal that *PTOX 1* and *PTOX 2* genes are functional and expressed in practically all evaluated parts of the plant and have high similarity to each other. However, during the development of cotyledons, leaves and roots the expression profile of the genes is distinct. *PTOX 1* prevails in younger organs whilst *PTOX 2* increases expression as these organs age. Already under stressful conditions, *PTOX 1* is generally positively regulated while *PTOX 2* levels are down-regulated or remain stable. However, when a very severe stress is applied, the expression of *PTOX 2* increases considerably. The affinity for the ligand also differed between the isoforms, being greater for PTOX 1. It is possible to conclude that *PTOX* genes duplication in soybean genome resulted in function gain to the plant. It is also suggested that *PTOX 1* gene is associated with photosynthetic efficiency whereas *PTOX 2* is related to leaf and cotyledon senescence.

Keywords: Soybean. Gene duplication. PTOX. Gene expression. RNA-seq. *Docking*.

LISTA DE ILUSTRAÇÕES

Figura 1 – Plântula de soja com indicação das principais partes e classificação taxonômica	15
Figura 2 – Mapa de produção agrícola da soja entre as regiões brasileiras referente à safra de 2016/2017	17
Figura 3 – Reações de transferência de elétrons durante a fotossíntese oxigênica	19
Figura 4 – Participação da oxidase terminal do plastídio (PTOX) na biossíntese de carotenoides	22
Gráfico 1 – Número de artigos científicos relacionados a genômica e transcriptômica da soja de 2007 a 2017 publicados na base de dados do PubMed	28
Figura 5 – Estrutura da plastoquinona	45
Figura 6 – Estrutura dos genes da oxidase terminal do plastídio (<i>PTOX</i>) de <i>Glycine max</i> representada pela ferramenta <i>Gene Structure Display Server 2.0</i>	46
Figura 7 – Alinhamento das sequências das oxidases terminais do plastídio com as oxidases alternativas de <i>Glycine max</i> feito pelo programa Clustal Omega	47
Figura 8 – Árvore filogenética dos genes da oxidase terminal do plastídio (PTOX) em leguminosas feita pelo programa MEGA 7.0	48
Figura 9 – Análise da integridade do RNA total extraído de diversas partes de <i>Glycine max</i> (A) em diferentes estádios de desenvolvimento e (B) tratados com polietilenoglicol (PEG) ou ácido salicílico (AS) em gel de agarose 1,5 %. Destaque para as bandas correspondentes às subunidades 18S e 28S do RNA ribossomal	52
Figura 10 – Eletroforese em gel de agarose 2 % dos <i>amplicons</i> do gene <i>EF1α 1a</i> para avaliação da qualidade do cDNA produzido por PCR semiquantitativa	52
Gráfico 2 – Expressão relativa dos genes <i>PTOX</i> de <i>Glycine max</i> em diferentes partes da planta durante o desenvolvimento avaliada por reação em cadeia da polimerase quantitativa em tempo real (RT-qPCR)	55
Gráfico 3 – Expressão relativa dos genes <i>PTOX</i> de <i>Glycine max</i> em folhas e raízes tratadas com polietilenoglicol (PEG) 100 g.L ⁻¹ ou ácido salicílico (AS) 0,5 mM em diferentes tempos de indução avaliada por reação em cadeia da polimerase quantitativa em tempo real (RT-qPCR)	56
Gráfico 4 – Expressão baseada em dados de RNA-seq dos genes <i>PTOX</i> em diferentes compartimentos das sementes de <i>Glycine max</i> nos estádios (A) globular, (B) cotiledonar e (C) de maturação precoce	58
Gráfico 5 – Expressão baseada em dados de RNA-seq dos genes <i>PTOX</i> de <i>Glycine max</i> durante o desenvolvimento e senescência de cotilédones e folhas trifolioladas ..	59
Gráfico 6 – Expressão baseada em dados de RNA-seq dos genes <i>PTOX</i> em raízes de <i>Glycine max</i> submetida a (A) desidratação e (B) estresse salino em diferentes tempos de indução	61

Gráfico 7 – Expressão baseada em dados de RNA-seq dos genes <i>PTOX</i> em folhas trifolioladas de dois cultivares de <i>Glycine max</i> com susceptibilidade e tolerância à seca submetidos a estresse por desidratação em diferentes tempos de indução	62
Gráfico 8 – Expressão baseada em dados de RNA-seq dos genes <i>PTOX</i> em zonas de abscisão foliar (LAZ) e de não abscisão (NAZ) de <i>Glycine max</i> expostos a alta concentração de etileno em diferentes tempos de indução	63
Gráfico 9 – Elementos <i>cis</i> regulatórios presentes nos promotores dos genes <i>PTOX 1</i> e <i>PTOX 2</i> de <i>Glycine max</i>	64
Figura 11 – Modelos tridimensionais preditos pelo servidor Robetta das proteínas PTOX 1 (A) e PTOX 2 (B) de <i>Glycine max</i> utilizados para o <i>docking</i> molecular	65
Figura 12 – Predição da estrutura secundária de PTOX 1 (A) e PTOX 2 (B) de <i>Glycine max</i> feita pela ferramenta PSIPRED	66
Figura 13 – Predição de regiões transmembranares em (A) PTOX 1 e (B) PTOX 2 de <i>Glycine max</i> feita pela ferramenta MEMSAT-SVM do servidor PSIPRED	67
Figura 14 – Diagrama de interação 2D proteína-ligante referente ao complexo PTOX1-PQH ₂ (plastoquinol)	68
Figura 15 – Diagrama de interação 2D proteína-ligante referente ao complexo PTOX2-PQH ₂ (plastoquinol)	69

LISTA DE TABELAS

Tabela 1	– Estádios de desenvolvimento da soja propostos por Ferh e Caviness (1977)	15
Tabela 2	– Órgãos coletados e idade das plantas usadas para os ensaios durante o desenvolvimento de <i>Glycine max</i>	33
Tabela 3	– Reagentes e volumes utilizados na reação de síntese de DNA complementar	35
Tabela 4	– Oligonucleotídeos iniciadores utilizados nas análises de expressão por Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR) de <i>Glycine max</i>	37
Tabela 5	– Experimentos de RNA-seq utilizados para analisar a expressão dos genes <i>PTOX</i> em <i>Glycine max</i>	40
Tabela 6	– Valores médios das concentrações e relações de densidade óptica (DO) dos RNAs extraídos de vários órgãos de <i>Glycine max</i> (cultivar BRS Pala) durante o desenvolvimento e tratados com polietilenoglicol (PEG) ou ácido salicílico (AS)	50
Tabela 7	– Temperaturas de anelamento ótimas e eficiência de amplificação dos <i>primers</i> dos genes <i>PTOX</i> de <i>Glycine max</i> e dos genes de referência utilizados no ensaio de PCR quantitativa em tempo real (RT-qPCR)	53

LISTA DE ABREVIATURAS E SIGLAS

AS	Ácido salicílico
EROs	Espécies Reativas de Oxigênio
NDH	Desidrogenase do NAD(P)H
PEG	Polietilenoglicol
PQ	Plastoquinona oxidada
PQH ₂	Plastoquinona reduzida (plastoquinol)
PSI	Fotossistema I
PSII	Fotossistema II
PTOX	Oxidase Terminal do Plastídio
RNA -seq	Sequenciamento de RNA
RT-qPCR	Reação em cadeia da polimerase quantitativa em tempo real

SUMÁRIO

1	INTRODUÇÃO	14
1.1	A soja [<i>Glycine max</i> (L.) Merrill]	14
1.2	A oxidase terminal do plastídio (PTOX)	18
1.2.1	<i>História evolutiva</i>	19
1.2.2	<i>Estrutura e sequências requeridas para a atividade</i>	20
1.2.3	<i>Papel fisiológico da PTOX</i>	21
1.2.4	<i>PTOX nas respostas a estresses</i>	23
1.3	Avanços na genômica, estudos de expressão gênica e estudos genômicos com a soja	25
2	OBJETIVOS	29
2.1	Objetivo geral	29
2.2	Objetivos específicos	29
3	METODOLOGIA	30
3.1	Caracterização dos genes <i>PTOX</i> em <i>Glycine max</i> e análise filogenética da proteína em leguminosas	30
3.2	Expressão gênica da <i>PTOX</i> em <i>Glycine max</i>	31
3.2.1	<i>Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR)</i>	31
3.2.1.1	<i>Material vegetal e condições de crescimento</i>	32
3.2.1.2	<i>Extração de RNA total</i>	33
3.2.1.3	<i>Quantificação, pureza e integridade do RNA total</i>	34
3.2.1.4	<i>Síntese de DNA complementar (cDNA)</i>	34
3.2.1.5	<i>Oligonucleotídeos iniciadores (primers) dos genes <i>PTOX</i> e dos genes normalizadores</i>	35
3.2.1.6	<i>3.2.1.6 Diluição, gradiente de temperatura, especificidade e curva de eficiência dos primers</i>	36
3.2.1.7	<i>PCR semi-quantitativa</i>	38
3.2.1.8	<i>Ensaio de PCR em tempo real (RT-qPCR)</i>	38
3.2.2	<i>Expressão in silico via RNA-seq</i>	39
3.2.2.1	<i>Obtenção das sondas</i>	39
3.2.2.2	<i>Ferramentas computacionais para cálculo da expressão diferencial</i>	42

3.2.3	<i>Análise da região promotora</i>	42
3.3	Análise estrutural das proteínas PTOX de <i>Glycine max</i>	43
3.3.1	<i>Predição e validação da estrutura tridimensional</i>	43
3.3.2	<i>Predição de estruturas secundárias</i>	44
3.3.3	<i>Docagem molecular das proteínas PTOX de <i>Glycine max</i></i>	45
4	RESULTADOS	46
4.1	Caracterização dos genes <i>PTOX</i> em <i>Glycine max</i> e distribuição da proteína em leguminosas	46
4.2	Expressão gênica da PTOX em <i>Glycine max</i>	49
4.2.1	<i>Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR)</i>	49
4.2.1.1	<i>Qualidade do RNA</i>	49
4.2.1.2	<i>Qualidade do DNA complementar obtido</i>	49
4.2.1.3	<i>Temperatura de anelamento, eficiência e especificidade dos primers</i>	53
4.2.1.4	<i>Expressão dos genes PTOX em <i>Glycine max</i> avaliada por RT-qPCR</i>	54
4.2.2	<i>Expressão in silico via RNA-seq</i>	57
4.2.3	<i>Análise da região promotora</i>	63
4.3	Modelagem da estrutura tridimensional das proteínas PTOX de <i>Glycine max</i> e interação com o plastoquinol	64
5	DISCUSSÃO	70
5.1	Os dois genes <i>PTOX</i> da soja são funcionais e expressos na maioria das partes da planta	70
5.2	O gene <i>PTOX 2</i> e sua relação com a senescência	72
5.3	<i>PTOX 1</i> e <i>PTOX 2</i> têm respostas distintas de acordo com a severidade do estresse	73
5.4	Afinidade das isoformas pelo plastoquinol e correlação com a função dos genes	75
6	CONCLUSÃO	77
	REFERÊNCIAS	78
	ANEXOS	90

1 INTRODUÇÃO

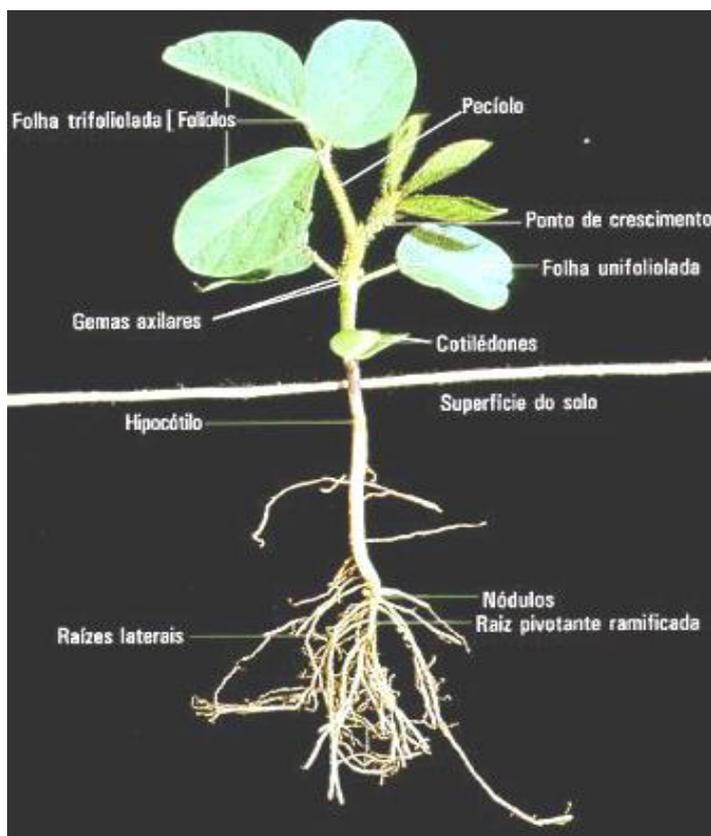
1.1 A soja [*Glycine max* (L.) Merrill]

A soja [*Glycine max* (L.) Merrill] é uma leguminosa amplamente cultivada em muitas regiões do mundo. Acredita-se que a espécie foi domesticada a partir da soja selvagem (*Glycine soja* Sieb. & Zucc) no leste da Ásia há cerca de 6000 a 9000 anos, apesar de ser relatado que a domesticação ocorreu de um complexo *G. soja/G. max* que divergiu de um ancestral comum às duas espécies (KIM *et al.*, 2012; KIM *et al.*, 2010; SEDIVY; WU; HANZAWA, 2017). O cultivo comercial da soja permaneceu por muito tempo restrito à China, chegando ao Ocidente na segunda década do século XX, de forma que desde meados da década de 1960, a produção da leguminosa vem crescendo exponencialmente (DALL'AGNOL, 2017).

A soja pertence à classe Magnoliopsida (das dicotiledôneas), ordem Fabales e família Fabaceae. É uma planta herbácea, de porte médio em torno de 1,0 metro e caule bastante ramificado. Durante seu ciclo de vida são identificados três tipos de folhas, as cotiledonares, as unifolioladas e as trifolioladas. Estas últimas, também chamadas de folhas verdadeiras, são alternadas, compostas de três folíolos ovalados ou lanceolados e geralmente amarelam e caem à medida que os frutos amadurecem. As flores podem ser brancas, amarelas ou violetas e os frutos são vagens achatadas de coloração cinza, amarela, palha ou preta. As sementes são arredondadas, achatadas ou alongadas, com coloração e tamanho variáveis e o sistema radicular é constituído de uma raiz principal bem desenvolvida com um grande número de raízes secundárias (MISSÃO, 2006; RITCHIE *et al.* 1985). A classificação taxonômica e as principais partes da planta são ilustradas na **Figura 1**.

O ciclo reprodutivo da soja tem duração variada de acordo com o cultivar, podendo ser precoce ou tardio (EMBRAPA SOJA, 2018). A metodologia para descrição dos estádios de desenvolvimento da soja proposta por Ferh e Caviness (1977) faz a divisão em estádios vegetativos (V) e reprodutivos (R) (**Tabela 1**) e é a mais utilizada mundialmente por ser precisa, universal e apresentar terminologia única (FARIAS; NEUMAIER; NEPOMUCENO, 2009).

Figura 1 – Plântula de soja com indicação das principais partes e classificação taxonômica.



Reino: Plantae
Divisão: Magnoliophyta
Classe: Magnoliopsida
Ordem: Fabales
Família: Fabaceae
Subfamília: Papilionoideae
Tribo: Pheseolae
Gênero: *Glycine*
Espécie: *Glycine max* (L.) Merrill

Fonte: National Center for Biotechnology Information (NCBI)/ Taxonomy. Disponível em: ncbi.nlm.nih.gov/taxonomy/. Foto: Ritchie *et al.* (1985) (traduzido).

Tabela 1 – Estádios de desenvolvimento da soja propostos por Ferh e Caviness (1977).

Estádios vegetativos	Estádios reprodutivos
VE – emergência	R1- início do florescimento
VC – cotilédones	R2 – pleno florescimento
V1 – primeiro nó	R3 – início da formação das vagens
V2 – segundo nó	R4 – plena formação das vagens
V3 – terceiro nó	R5 – início do enchimento das sementes
*	R6 – pleno enchimento das sementes
*	R7 – início da maturação
Vn – enésimo nó	R8 – maturação plena

Fonte: Ritchie *et al.* (1985) (traduzido). Observações: VE - cotilédones acima da superfície do solo; VC – cotilédones completamente abertos; V1 – corresponde às folhas unifolioladas completamente desenvolvidas (bordas não se tocam); V2 – corresponde à primeira folha trifoliolada completamente desenvolvida e assim sucessivamente até o Vn – enésima folha trifoliolada completamente desenvolvida. Nó cotiledonar não é considerado.

Os altos teores proteicos (38-45 %) e de óleo (aproximadamente 20 %) presentes nos grãos da soja, além da capacidade que a planta tem de fixar o nitrogênio atmosférico justificam o consolidado interesse na produção da leguminosa, que hoje é considerada uma das espécies cultivadas mais importantes do mundo (TYCZEWSKA *et al.*, 2016; DALL'AGNOL, 2017). O aproveitamento da soja se concentra principalmente na alimentação animal, se constituindo um alimento rico em proteínas, e na produção do óleo vegetal, o segundo óleo mais consumido mundialmente (TYCZEWSKA *et al.*, 2016).

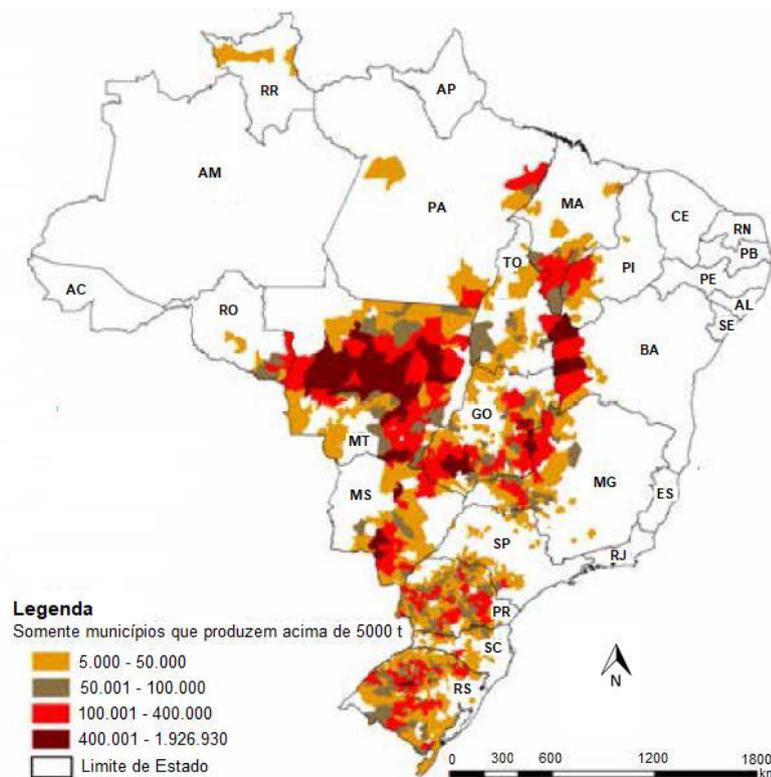
Além destas aplicações, a soja também é largamente utilizada na fabricação de produtos alimentícios como leite e farinha de soja, tofu e aditivos alimentares, bem como na obtenção de produtos industriais como cosméticos, plásticos e tintas. Ainda, a produção de biodiesel utiliza como matéria-prima o óleo extraído das suas sementes (TYCZEWSKA *et al.*, 2016). Há também relatos da presença de uma variedade de compostos bioativos na soja, tais como saponinas, inibidores de protease, ácido fítico e isoflavonas, os quais podem ser aplicados na área da saúde humana (BARRETT, 2006). Devido ao seu valor nutricional e econômico, a soja tem beneficiado produtores, indústrias e consumidores (TYCZEWSKA *et al.*, 2016).

Os Estados Unidos lideram a produção mundial da soja, seguidos do Brasil, Argentina e China. No último levantamento para a safra mundial da soja (2017/2018) feito pelo Departamento de Agricultura dos Estados Unidos (USDA), a previsão para a produção mundial é 348,6 milhões de toneladas do grão. O consumo global, que cresce há 10 anos de forma ininterrupta, está previsto para alcançar neste ano a faixa dos 345 milhões de toneladas (USDA, 2018; FIESP, 2018).

A soja é a principal *commodity* do Brasil, que lidera as exportações mundiais do grão (USDA, 2018). Entre os principais produtores nacionais estão Mato Grosso, Paraná e Rio Grande do Sul. No Nordeste, o principal produtor é a Bahia. (Companhia Nacional de Abastecimento - Conab, 2018). A **Figura 2** mostra a produção da soja em toneladas por safra entre os municípios brasileiros.

No levantamento mais recente das safras brasileiras da soja feito pela Conab, são observadas ampliações nas áreas plantadas de praticamente todos os estados detentores de plantações da leguminosa (CONAB, 2018), indicando que há um constante esforço para aumentar a produtividade da soja no Brasil, o que também reflete a demanda crescente global pelo grão.

Figura 2 – Mapa de produção agrícola da soja entre as regiões brasileiras referente à safra de 2016/2017.



Fonte: Conab (2018)

O evidente avanço na produção e comercialização da soja no cenário mundial requer que os cultivos alcancem o máximo de produtividade possível, entretanto, as condições climáticas são um ponto limitante neste processo. O crescimento, a produtividade e a qualidade da semente da soja são afetadas diretamente por uma gama de estresses bióticos e abióticos como ataque de patógenos, doenças virais, seca, acúmulo de sais nos solos, altas temperaturas, excesso de luminosidade, dentre outros. A baixa disponibilidade hídrica, contudo, tem sido elencada como o principal fator abiótico que interfere na produtividade agrícola, podendo reduzir o rendimento da produção da soja em torno de 40 %. Vale ressaltar que a soja é composta por aproximadamente 90 % de água e necessita, portanto, de suprimento hídrico suficiente para se desenvolver adequadamente, principalmente nos períodos de germinação-emergência e de floração-enchimento de grãos (FARIAS; NEUMAIER; NEPOMUCENO, 2009; MANAVALAN *et al.*, 2009; TRAN; MOCHIDA, 2010; MORANDO *et al.*, 2014).

Uma alternativa biotecnológica para contornar esse problema, portanto, é o desenvolvimento de marcadores moleculares para serem usados em programas de melhoramento genético vegetal (SAIBI *et al.*, 2013).

1.2 A oxidase terminal do plastídio (PTOX)

A oxidase terminal do plastídio (PTOX), também conhecida como oxidase terminal do plastoquinol ou ainda como IMMUTANS, produto do gene *IMMUTANS* de *Arabidopsis thaliana*, de onde foi isolada pela primeira vez, é uma enzima codificada no núcleo da célula, sendo encontrada somente em organismos capazes de realizar fotossíntese oxigênica (CAROL; KUNTZ, 2001; MCDONALD *et al.*, 2011).

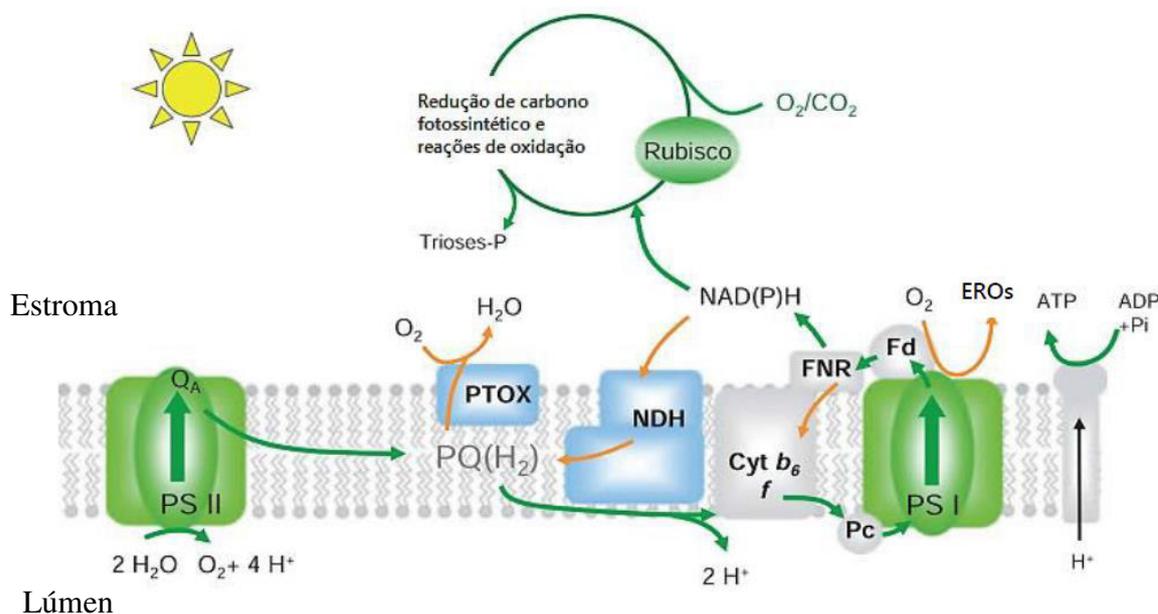
A reação catalisada pela PTOX (**Figura 3**) é a oxidação do plastoquinol (PQH₂) e a concomitante redução do oxigênio molecular a água (KUNTZ, 2004). O plastoquinol é uma molécula formada a partir da redução da plastoquinona pelo fotossistema II (PSII) durante o transporte de elétrons fotossintético, mas que, em determinadas condições, também pode ser gerado através da oxidação do NAD(P)H ou da ferredoxina (RUMEAU; PELTIER; CURNAC, 2007). A alta concentração de PQH₂ é reconhecida pela PTOX, que realiza sua reoxidação (KUNTZ, 2004;). Durante a reação, há a participação de dois íons ferros (Fe), cofatores da enzima, os quais doam elétrons para o oxigênio, sendo necessários dois PQH₂ para reduzir uma molécula de O₂ a duas moléculas de H₂O (MCDONALD *et al.*, 2011; NAWROCKI, 2015).

Em relação à localização subcelular, PTOX está presente nos cloroplastos, possuindo uma sequência de endereçamento na extremidade amino-terminal responsável pela importação da proteína até o plastídio (CAROL *et al.*, 1999;). Dentro dos cloroplastos, PTOX encontra-se ligada aos tilacoides, provavelmente como uma proteína de membrana voltada para o lado estromal (LENNON; PROMMEENATE; NIXON, 2003). Entretanto, evidências apontam que a ligação da PTOX 2 de *Chlamydomonas reinhardtii* à membrana é um processo dinâmico dependente do pH (KRIEGER-LISZKAY; FEILKE, 2016).

A primeira evidência para a existência da PTOX ocorreu em 1982, sendo comprovada na década seguinte através da caracterização de mutantes de *Arabidopsis* que não continham o gene *IMMUTANS*. A inativação do gene na planta resultou no aparecimento de fenótipos variegados, ou seja, apresentando folhas com setores verdes e brancos ou totalmente brancas, especialmente em condições de alta luminosidade (CAROL *et al.*, 1999; WU *et al.*, 1999).

Após essa descoberta, diversos trabalharam se empenharam em elucidar o papel fisiológico da PTOX e o seu envolvimento em processos metabólicos celulares, bem como a sua participação na resposta a estresses ambientais em plantas (CAROL; KUNTZ, 2001; KUNTZ, 2004; SUN; WEN, 2011).

Figura 3 – Reações de transferência de elétrons durante a fotossíntese oxigênica.



Fonte: Rumeau, Peltier e Cournac (2007). A oxidase terminal do plastídio (PTOX) está presente na membrana tilacoide e catalisa a oxidação do plastoquinol (PQH_2) e a redução do oxigênio molecular a água. Legenda: PSI – fotossistema I; PSII – fotossistema II; Cyt b_6f – complexo do citocromo b_6f ; FNR, ferredoxina NADP oxidorreductase; NDH - desidrogenase do NAD(P)H; Fd – ferredoxina; Pc – plastocianina; EROs – espécies reativas de oxigênio.

1.2.1 História evolutiva

A PTOX é homóloga da oxidase alternativa (AOX) de plantas, uma enzima resistente ao cianeto que atua transferindo elétrons do ubiquinol para o oxigênio molecular na via alternativa da respiração mitocondrial (SUN; WEN, 2011). Ambas são proteínas não-heme que apresentam um centro binuclear de ferro (Fe-Fe) ligado por quatro resíduos de glutamato (E) e dois de histidina (H), o que as torna pertencentes ao grupo das proteínas diferido carboxilato (BERTHOLD; STENMARK, 2003; NAWROCKI *et al.*, 2015). Além destas semelhanças, PTOX também apresenta sensibilidade aos mesmos inibidores da AOX, n-octil galato, n-propil galato e ácido salicilhidroxâmico (SHAM) (SUN; WEN, 2011; KUNTZ, 2004).

As duas proteínas são encontradas tanto em procariotos quanto em eucariotos e compartilham o mesmo ancestral, uma proteína diferido carboxilato primitiva de origem procariótica (NAWROCKI *et al.*, 2015). Um estudo feito por McDonald & Vanlerberghe (2006) aponta que PTOX e AOX divergiram antes dos eventos endossimbióticos primários que originaram os cloroplastos e as mitocôndrias. Concordando com esse fato, a análise filogenética entre as duas enzimas demonstrara que elas encontram-se em clados distintos,

estando mais relacionadas com as suas cópias presentes em procariotos e eucariotos do que entre si (MCDONALD; VANLERBERGHE, 2006; CAROL; KUNTZ, 2001).

A presença de mais de uma cópia do gene *PTOX* é relatada em vários organismos de diferentes grupos taxonômicos, como algas, diatomáceas e representantes dos filos Cryptophyta e Haptophyta (NAWROCKI *et al.*, 2015; MCDONALD *et al.*, 2011). A microalga verde *Chlamydomonas reinhardtii*, por exemplo, apresenta dois genes codificando a PTOX (FEILKE *et al.*, 2016; PAN *et al.*, 2018) e o musgo do gênero *Physcomitrella*, três (NAWROCKI *et al.*, 2015). De forma curiosa, as múltiplas cópias da proteína em um único organismo, na maioria das vezes, são possivelmente resultado da transferência horizontal de genes, uma vez que as cópias individuais da PTOX em certas espécies estão pouco relacionadas. Entretanto, algumas espécies não obedecem a esta regra, e apresentam cópias da PTOX mais relacionadas, como acontece, por exemplo, com o milho (SOUSA, 2008), eucalipto e algas vermelhas, o que pode ser consequência de eventos de duplicação gênica. Em vista disso, pode-se afirmar que a história evolutiva da PTOX é complexa e, considerando que a duplicação do gene é observada em algumas espécies, são levantados questionamentos acerca da importância funcional da enzima para esses organismos (NAWROCKI *et al.*, 2015).

1.2.2 Estrutura e sequências requeridas para a atividade

Proteínas difero carboxilato são compostas por um núcleo de quatro α -hélices agrupadas contendo o sítio de ligação ao ferro (quatro resíduos de glutamato e dois de histidina, como mencionado anteriormente) e, especialmente no caso da PTOX e da AOX, duas hélices adicionais estão presentes e são responsáveis pela ligação dessas proteínas à membrana (BERTHOLD; STENMARK, 2003; SHIBA *et al.*, 2013). Importante ressaltar que o sítio de ligação ao ferro da PTOX está voltado para o lado estromal (MCDONALD *et al.*, 2011).

Algumas sequências são relatadas por exercerem uma função relevante na atividade catalítica da PTOX da planta modelo *Arabidopsis thaliana*. O motivo de ligação ao ferro, já citado anteriormente, é um deles (FU; PARK, RODERMEL, 2005). Josse e colaboradores (2003) propuseram um sítio de ligação à quinona para PTOX e AOX, que consiste numa sequência de oito a nove resíduos de aminoácidos variáveis contendo uma histidina altamente conservada na posição central, sendo esta, portanto, presente em todas as sequências dessas enzimas. Outras sequências podem estar relacionadas não só com a

atividade catalítica, mas também com a estabilidade da enzima. É o que foi comprovado para a região que corresponde ao éxon 8 da PTOX (que não está presente em AOX) através de estudos de mutagênese *in vitro* e *in planta* (FU; PARK, RODERMEL, 2005). Um estudo semelhante com 14 sítios conservados entre PTOX e AOX identificou cinco resíduos de aminoácidos que desempenham papel importante na atividade dessas enzimas, dois dos quais são indispensáveis (FU; ALURU, RODERMEL, 2009).

PTOX ainda não foi cristalizada e, portanto, não possui estrutura resolvida. Entretanto, por ser homóloga à AOX, a sua estrutura tridimensional pode ser predita baseando-se no cristal da oxidase alternativa de *Trypanosoma brucei*, que encontra-se depositado no Banco de Dados de Proteínas (*Protein Data Bank*, PDB). Lima (2017) fez a predição da PTOX de diferentes ecotipos *A. thaliana* contendo polimorfismos de um único nucleotídeo (SNPs) com base nessa estrutura da AOX disponível no PDB e realizou ensaio de docagem molecular da enzima com o seu ligante, o plastoquinol. Abordagens desse tipo envolvendo a PTOX são importantes para avançar na compreensão da sua estrutura e função, entretanto, praticamente não são encontradas na literatura.

1.2.3 Papel fisiológico da PTOX

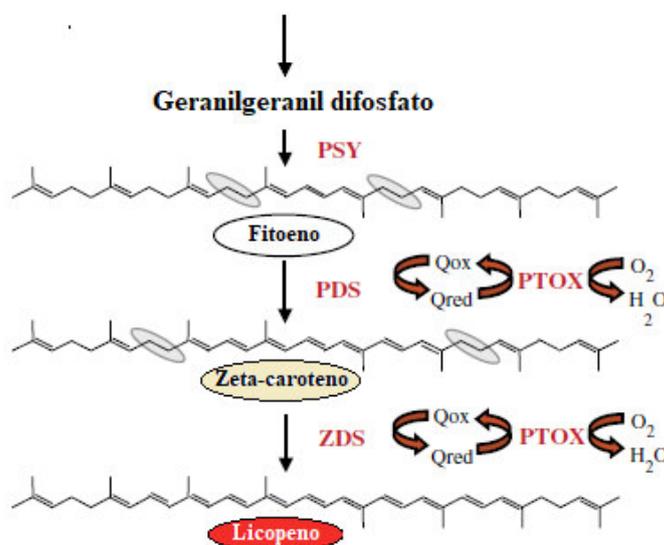
O fenótipo variegado, ou seja, com folhas apresentando setores verdes e brancos, resultante da inativação do gene *IMMUTANS* em *Arabidopsis* (como visto no tópico 1.2) ocorreu em plantas expostas a condições de luminosidade mediana e baixa, ao passo que as plantas que apresentaram somente folhas brancas foram submetidas a alta luminosidade. Nos setores brancos foi detectado acúmulo de fitoeno, precursor incolor da biossíntese de carotenoides. Além disso, também foi constatado que a descoloração destes setores era irreversível nos estádios iniciais do desenvolvimento da folha (CAROL *et al.*, 1999; WU *et al.*, 1999; KUNTZ, 2004; CAROL; KUNTZ, 2001). Em tomate (*Solanum lycopersicum*) com o gene *PTOX* (*ghost*) inativado, as folhas também exibiram manchas brancas, além das flores e frutos terem apresentado despigmentação (SHAHBAZI *et al.*, 2007).

Tais constatações levaram a algumas hipóteses referentes ao papel fisiológico da PTOX (*IMMUTANS*). Uma delas é de que um nível mínimo de carotenoides, moléculas conhecidas por exibirem atividade antioxidante, é necessário nos plastídios a fim de permitir a adequada proteção aos estresses foto-oxidativos. Nesse sentido, a PTOX seria essencial para alcançar esse nível mínimo, mas somente nos estádios iniciais do

desenvolvimento dos cloroplastos, uma vez que substitutos da enzima devem atuar nas fases pós-diferenciação (KUNTZ, 2004). A enzima estaria, portanto, envolvida na biogênese dos cloroplastos (JOHNSON; STEPIEN, 2016).

Como visto anteriormente, a ausência da PTOX em *Arabidopsis* também levou ao bloqueio da biossíntese de carotenoides, uma vez que o precursor dessa via foi acumulado nas folhas desprovidas de cor. Deste modo, tem sido proposto que a PTOX atua como cofator na biossíntese de carotenoides, fornecendo plastoquinona oxidada, produto da sua reação, para a dessaturase do fitoeno (PDS) e a dessaturase do zeta-caroteno (ZDS), enzimas dessa via (**Figura 4**). O estudo feito por Steccanella, Hansson & Jensen (2015) também apontou o envolvimento da PTOX na biossíntese de clorofila, uma vez que identificou que a plastoquinona oxidada é capaz de atuar como precursor da enzima ciclase redutase, que participa da biossíntese do pigmento.

Figura 4 – Participação da oxidase terminal do plastídio (PTOX) na biossíntese de carotenoides.



Fonte: Simkin *et al.* (2008). Legenda: PSY – sintase do fitoeno; PDS – dessaturase do fitoeno; ZDS – dessaturase do zeta-caroteno; Qox – plastoquinona oxidada; Qred – plastoquinona reduzida. São apresentadas somente as reações da via em que há participação da PTOX.

Outra função atribuída à PTOX é a sua participação na cloro-respiração, uma via que tem sido sugerida por atuar como uma válvula de segurança, uma vez que previne a super-redução do *pool* de plastoquinona (PQ) e protege, desta maneira, o PSII de danos foto-oxidativos causados pelo excesso de luz (MCDONALD *et al.*, 2011; NIYOGI, 2000). A cloro-respiração consiste na redução não-fotoquímica da PQ oxidada a PQH₂ pela

desidrogenase do NAD(P)H (NDH) e na sua posterior oxidação pela PTOX (CAROL; KUNTZ, 2001; JOHNSON; STEPIEN, 2016; SUN; WEN, 2011).

Sua atuação é especialmente requerida no escuro, desviando elétrons que não podem ser transferidos pelos fotossistemas devido à falta de luz e gerando a força próton-motriz necessária para a produção de ATP pela ATP-sintase. Deste modo, a cloro-respiração é capaz de evitar a ativação desnecessária dos mecanismos regulatórios do aparato fotossintético (NAWROCKI *et al.*, 2015; JOHNSON; STEPIEN, 2016). Entretanto, a via cloro-respiratória também pode ocorrer na presença de luz, como resposta ao excesso de elétrons produzido quando a planta é exposta a algum tipo de estresse ambiental. Neste sentido, o desvio de elétrons para evitar a super-redução da PQ evita a formação de espécies reativas de oxigênio (EROs) que possam vir a ser geradas e causar danos à célula (MCDONALD *et al.*, 2011; NAWROCKI *et al.*, 2015; JOHNSON; STEPIEN, 2016).

1.2.4 PTOX nas respostas a estresses

Conforme exposto no tópico anterior, a PTOX desempenha um papel importante na manutenção do estado redox da plastoquinona por meio da cloro-respiração, contribuindo, dessa forma, para a homeostase celular (KUNTZ, 2004; JOHNSON; STEPIEN, 2016). Uma vez envolvida em diversas funções fisiológicas (além da via cloro-respiratória), é de se esperar que a abundância da proteína e/ou dos transcritos seja alterada quando submetida a condições estressantes. De fato, uma variedade de trabalhos avalia a expressão da PTOX frente a diferentes estresses (NAWROCKI *et al.*, 2015; SUN; WEN, 2011).

Tallón & Quiles (2007) encontram que os níveis de PTOX em folhas jovens e maduras de *Avena sativa* aumentaram em resposta à exposição a alta intensidade luminosa e temperatura moderada. Paralelamente, os níveis de NDH, enzima que participa da via da cloro-respiração juntamente com a PTOX, também aumentaram. Um resultado semelhante foi encontrado para *Spathiphyllum wallisii*, uma planta aclimatada à sombra na qual tanto os níveis de PTOX quanto de NDH foram elevados em estresse induzido por alta temperatura e luminosidade. Em contrapartida, nas plantas aclimatadas ao sol (*Chrysanthemum morifolium*), este mesmo tratamento não alterou de forma significativa os níveis das duas proteínas, sugerindo que a via cloro-respiratória é mais importante para plantas aclimatadas à sombra (GAMBOA; MUÑOZ; QUILES, 2009).

Quando plantas de *Rosa meilandina* foram submetidas a alta temperatura e luminosidade, os níveis de PTOX e NDH também não sofreram alteração, somente aumentando quando o estresse foi combinado com déficit hídrico (PAREDES; QUILES, 2013). Uma resposta similar foi encontrada para plantas de *Hibiscus rosa-sinensis* incubadas sob alta intensidade luminosa e submetidas a seca (MUÑOZ; QUILES, 2013). A abundância relativa dos transcritos da PTOX também foi aumentada em folhas de *Coffea arabica* submetidas a estresse hídrico por 3 a 6 semanas (SIMKIN *et al.*, 2008).

Em plantas de arroz (*Oryza sativa*), o aumento no nível de transcritos da PTOX foi induzido por tratamento com cloreto de sódio (NaCl) e ácido abscísico (ABA). Os autores ainda identificaram que o gene PTOX (*OsIM*) sofre um mecanismo de *splicing* alternativo, sendo possível identificar a presença de dois transcritos, *OsIM1* e *OsIM2*, cuja abundância é diferencialmente regulada por estresse salino em cultivares sensíveis e tolerantes (KONG *et al.*, 2003).

Outros estudos evidenciando a regulação positiva nos níveis da PTOX são relatados por Díaz *et al.* (2007), para plantas de *Brassica* sob condições de alta temperatura e luminosidade e por Busch, Huner e Ensminger (2008) e Savitch *et al.* (2010), para plantas do gênero *Pinus* submetidas a estresse induzido por frio. Ainda, Sun *et al.* (2014) encontraram que o estresse hídrico aumenta a abundância de transcritos totais da PTOX em *G. max*, enquanto Hossain *et al.* (2017) identificaram aumento de expressão dos dois transcritos da PTOX na beterraba (*Beta vulgaris*) submetida a estresse salino.

Sabe-se que o transporte de elétrons fotossintético é flexível, ou seja, responde de forma dinâmica às diferentes condições ambientais as quais uma planta pode ser exposta, tais como seca, altas temperaturas, alta intensidade luminosa, salinidade elevada, dentre outras. Essa flexibilidade atua a fim de evitar danos foto-oxidativos no aparato fotossintético, que, por sua vez, são causados pela formação de espécies reativas de oxigênio (EROs) resultante da carga excessiva de elétrons gerada por esses estresses (MCDONALD *et al.*, 2011). Exemplos de vias que atuam desviando esse excesso de elétrons são a clororespiração, já mencionada anteriormente, e o transporte de elétrons cíclico envolvendo o fotossistema I (RUMEAU, 2007).

Analisando o papel bioquímico da PTOX, ou seja, sua função de equilibrar o estado redox da PQ, bem como as respostas da enzima frente a estresses, é possível observar o seu envolvimento no desvio de elétrons para evitar pressão oxidativa. Tais atributos da PTOX levaram a tentativas de superexpressar a proteína com o intuito de se obter plantas tolerantes a estresses. Contudo, essas tentativas falharam e, ainda de forma

mais surpreendente, foi detectado que a PTOX pode promover o estresse oxidativo (levar à formação de EROs) quando presente em altas concentrações (ROSSO *et al.*, 2006; HEYNO *et al.*, 2009).

Uma possível explicação para esse acontecimento é que a superexpressão não deveria ser induzida somente para PTOX, mas também para a desidrogenase do NAD(P)H (JOHNSON; STEPIEN, 2016). De fato, em vários casos de indução de estresse, o aumento da expressão da oxidase é acompanhado do aumento nos níveis de NDH, as duas enzimas que participam da cloro-respiração. Além disso, a produção de EROs pela PTOX (formação de intermediários de H₂O₂ durante a catálise) relatada por alguns estudos poderia ser uma sinalização para a ativação do metabolismo anti-oxidante (KRIEGER-LISZKAY; FEILKE, 2016).

Apesar das demonstrações do duplo papel exercido pela PTOX (função protetora e função pro-oxidante), continuam existindo boas evidências para o papel da PTOX como dissipadora de elétrons em algumas espécies e em determinadas condições. Nesse sentido, sugere-se que a PTOX é sim uma forte candidata a marcador molecular para o desenvolvimento de plantas tolerantes a estresse, entretanto, o entendimento incompleto acerca da sua função no metabolismo das espécies estudadas ainda é uma barreira para se alcançar esse objetivo (KRIEGER-LISZKAY; FEILKE, 2016; JOHNSON; STEPIEN, 2016).

1.3 Avanços na genômica, estudos de expressão gênica e estudos genômicos com a soja

O acesso a dados de sequências genômicas teve início a partir da década de 1970 com o desenvolvimento das técnicas de sequenciamento enzimático (SANGER *et al.*, 1977) e químico (MAXAM; GILBERT, 1977) da molécula de DNA. Os avanços no campo da genômica desde aquela época têm revolucionado o entendimento acerca da organização e funcionamento metabólicos dos organismos. De forma especial, isso foi viabilizado através do sequenciamento de genomas inteiros, como por exemplo o de *Arabidopsis thaliana* no ano de 2000, uma espécie de planta pertencente à família Brassicaceae e utilizada como modelo em diversos estudos.

O desenvolvimento da genômica logo levou à compreensão de que era necessário avançar também na análise dos produtos gênicos. Neste contexto, surgiram as ciências ômicas, como por exemplo a transcriptômica, que estuda o conjunto de RNAs (transcritos) presentes na célula em um determinado momento, também chamada de

genômica funcional; a proteômica, que refere-se à análise sistemática das proteínas e a metabolômica, que investiga o conjunto de metabólitos presentes na célula (BINNECK, 2004; SOUZA; RHODEN; PAMPHILE, 2014).

Um progresso ainda maior no campo da genômica foi alcançado após o surgimento das tecnologias de sequenciamento de segunda geração (NGS), que permitiram acesso a uma imensa quantidade de dados em um curto período de tempo. Importante ressaltar que todo esse avanço tem acontecido de forma concomitante ao desenvolvimento da bioinformática, de forma que um dos maiores desafios da era pós-genômica é o processamento e análise dessa quantidade massiva de dados (OZSOLAK; MILOS, 2010; GOODWIN *et al.*, 2016).

A análise da expressão gênica, ou seja, do transcriptoma de um organismo, tecido ou célula permite identificar que genes estão sendo expressos em determinadas condições e/ou em diferentes estádios de desenvolvimento da espécie estudada, fazer comparações do perfil de expressão entre momentos distintos, dentre outras abordagens (MCGETTIGAN, 2013). Dentre as metodologias utilizadas para esse tipo de análise, pode-se citar as técnicas que baseiam-se na hibridização dos transcritos com sequências complementares, como Northern-blot, mais clássica, e os microarranjos, técnica pela qual foi sequenciado o transcriptoma de *Saccharomyces cerevisiae* (LASHKARI *et al.*, 1997) e cuja aplicação nos anos que seguintes passou por uma expansão surpreendente. Entretanto, algumas limitações da técnica de microarranjos são o fato de que a análise dos transcritos se restringe apenas a genes conhecidos, não sendo possível identificar novos genes, além da grande quantidade de RNA requerida para as análises (OZSOLAK; MILOS, 2010; VERK *et al.*, 2013).

As metodologias para análise de expressão gênica evoluíram juntamente com as tecnologias de sequenciamento. Em 2008 foi publicado o primeiro estudo envolvendo as tecnologias NGS no sequenciamento de transcriptoma (NAGALAKSHMI *et al.*, 2008) e os microarranjos começaram a ser substituídos. Era o início da expansão da tecnologia denominada RNA-seq, que permite o sequenciamento de RNA em larga escala e consequentemente, novos avanços em relação à caracterização e quantificação de transcriptomas (OZSOLAK; MILOS, 2010). Essa técnica baseia-se na produção de dezenas de milhares de fragmentos (*reads*), os quais são alinhados a um genoma ou transcriptoma de referência, que, por sua vez, permitem a obtenção da sequência dos transcritos e o nível de expressão de cada gene (MARTIN *et al.*, 2013).

Para análise de expressão gênica utilizando RNA-seq não é necessário conhecimento prévio do transcriptoma, bem como é possível avaliar genes para os quais não existem sondas. Ainda, a técnica permite identificar diferentes isoformas, variações de *splicing* e os polimorfismos de um único nucleotídeo (SNPs), além de ser rápida e estar apresentando custo cada vez mais acessível. Além disso, a análise por RNA-seq dispensa etapas de clonagem e, portanto, requer pequenas quantidades de amostras. Desse modo, vem se estabelecendo como o principal método para análise em larga escala de transcriptomas (OZSOLAK; MILOS, 2010; MCGETTIGAN, 2013; MUTZ *et al.*, 2013; CONESA *et al.*, 2016). Entre as ferramentas mais utilizadas para análise de expressão diferencial a partir de dados de RNA-seq, pode-se destacar TopHat e Cufflinks (TRAPNELL *et al.*, 2012).

Outra técnica também bastante utilizada para a análise de expressão gênica é a reação em cadeia da polimerase quantitativa em tempo real (RT-qPCR), um método altamente eficiente para a detecção de RNAm em tecidos, visto que possui alta sensibilidade e especificidade. A técnica permite monitorar ao longo da reação a quantidade de produto (*amplicon*) formado a cada ciclo de amplificação e pode quantificar a expressão de forma absoluta, na qual se utiliza uma curva padrão, ou relativa, fazendo uso de genes normalizadores conhecidos por sua expressão estável em diversas condições e que funcionam como referência (ROCHA *et al.*, 2015).

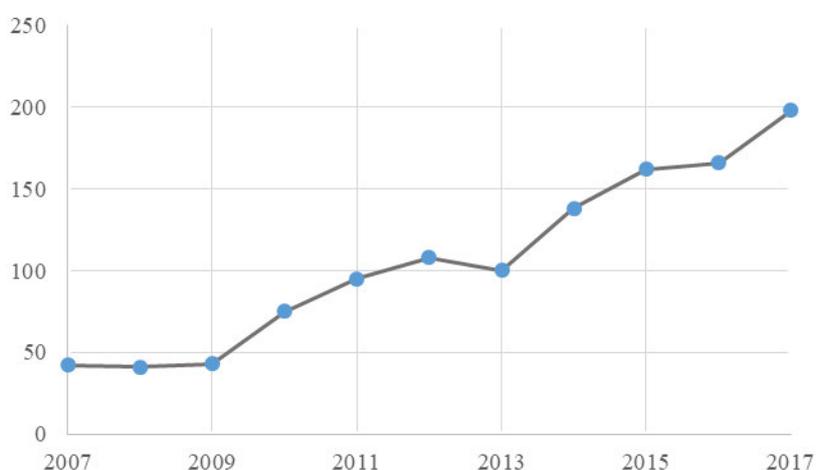
A demanda crescente pela produção da soja aliada à necessidade de desenvolver cultivares resistentes às condições ambientais adversas impulsionou os estudos moleculares envolvendo a leguminosa. O avanço da genômica contribuiu de forma expressiva neste processo, especialmente quando o genoma da leguminosa foi sequenciado e montado em cromossomos, tornando-se disponível nos bancos de dados em 2010 (SCHMUTZ, 2010).

A utilização da tecnologia de RNA-seq aplicada à soja também tem ganhado espaço, de forma que, nos últimos anos, uma grande quantidade de trabalhos analisando o transcriptoma da leguminosa nas mais diversas condições (durante o desenvolvimento ou sob condições de estresse) estão disponíveis no banco de dados SRA (*Sequence Read Archive*) do GenBank. Esses estudos vêm possibilitando importantes avanços na compreensão das respostas moleculares e dos mecanismos regulatórios envolvidos no metabolismo da soja e dão impulso, desse modo, para o desenvolvimento de estratégias de melhoramento genético (TRAN; MOCHIDA, 2010). Alguns exemplos desses estudos

podem ser encontrados nos trabalhos de Shamimuzzaman e Vodkin (2013), Leisner, Ming e Ainsworth (2014), Belamkar *et al.* (2014), Danzer *et al.* (2015), Prince *et al.* (2015) e Shin *et al.* (2015).

Uma busca simples feita na base de dados de artigos científicos do PubMed restrita às palavras “soybean” e “genome” ou “transcriptome” no título ou no resumo revelou que, em cinco anos (2013 a 2017), o número de trabalhos publicados anualmente praticamente dobrou, evidenciando o crescente interesse na área (**Gráfico 1**).

Gráfico 1 – Número de artigos científicos relacionados a genômica e transcriptômica da soja de 2007 a 2017 publicados na base de dados do PubMed.



Fonte: Autoria própria (2018).

Recentemente, Sun *et al.* (2015) identificaram a provável existência de dois genes codificando a PTOX em *G. max*. Como já evidenciado, essa enzima pode se constituir um potencial marcador molecular em programas de melhoramento genético, portanto, avançar na compreensão da função e dos mecanismos regulatórios envolvidos na expressão gênica dos dois parálogos da PTOX na soja é bastante pertinente para possíveis aplicações dessa natureza.

2 OBJETIVOS

2.1 Objetivo geral

Diante do exposto, o objetivo deste trabalho foi caracterizar e analisar o perfil de expressão de dois genes parálogos da oxidase terminal do plastídio (PTOX) em *Glycine max* (L.) Merrill e investigar a afinidade das respectivas isoformas proteicas pelo ligante.

2.2 Objetivos específicos

- ✓ Caracterizar genes da PTOX em *Glycine max* e outras plantas leguminosas usando genomas disponíveis em bancos de dados públicos;
- ✓ Fazer análise filogenética das proteínas PTOX em leguminosas;
- ✓ Analisar o perfil de expressão dos genes *PTOX* identificados em *Glycine max* durante o desenvolvimento e em condições de estresse através da Reação em Cadeia da Polimerase quantitativa em tempo real (RT-qPCR) e a partir de dados transcriptômicos (RNA-seq) disponíveis;
- ✓ Analisar a região promotora dos genes *PTOX* identificados em *Glycine max*;
- ✓ Predizer a estrutura tridimensional e secundária das isoformas da PTOX de *Glycine max*;
- ✓ Realizar ensaio de docagem molecular entre as isoformas proteicas da PTOX de *Glycine max* e o seu ligante plastoquinol.

3 METODOLOGIA

3.1 Caracterização dos genes *PTOX* em *Glycine max* e análise filogenética da proteína em leguminosas

Sun e colaboradores (2015) identificaram, através da ferramenta BLAST (*Basic Local Alignment Search Tool*, disponível em blast.ncbi.nlm.nih.gov/Blast.cgi) (ALTSCHUL *et al.*, 1997) do banco de dados GenBank (NCBI, *National Center for Biotechnology Information*, disponível em ncbi.nlm.nih.gov) dois possíveis genes parálogos da proteína PTOX em *Glycine max*, identificados pelos códigos Glyma0901130 e Glyma15g11950. Este mesmo grupo de pesquisadores também caracterizou a estrutura dos genes e a sequência teórica das proteínas PTOX da soja, dando fortes indícios para a funcionalidade desses genes no genoma da leguminosa; entretanto, não houve comprovação experimental para a existência dos parálogos. De forma semelhante, o grupo de pesquisa do Laboratório de Genômica Funcional e Bioinformática da Universidade Federal do Ceará também verificou a existência de dois possíveis genes para a oxidase terminal do plastídio em *G. max*. A fim de avançar na compreensão dessa descoberta, complementou-se a caracterização dos genes da PTOX da soja feita por Sun *et al.* (2015) e analisou-se filogeneticamente a distribuição da proteína entre outras espécies de leguminosas, pertencentes à ordem Fabales.

Para tanto, os valores teóricos de ponto isoelétrico (pI) e da massa molecular das sequências das proteínas PTOX da soja foram preditas com o auxílio da ferramenta *Compute pI/Mw* (GASTEIGER, 2005) e a predição do peptídeo trânsito foi feita pela ferramenta ChloroP 1.1 Server (EMANUELSSON, NIELSEN, e VON HEIJNE, 1999), todas disponíveis no servidor *online* ExPASy (expasy.org/resources). Uma série de outras caracterizações como número de íntrons e éxons de cada gene, localização no genoma, tamanho da sequência de aminoácidos das proteínas, dentre outras, também foi feita por Sun *et al.* (2015). A estrutura dos genes foi representada pelo servidor *Gene Structure Display Server 2.0* (HU *et al.*, 2015) e o alinhamento entre as sequências das proteínas PTOX deduzidas neste trabalho, como descrito a seguir, e das oxidases alternativas (AOX) da soja, classificadas por Costa *et al.* (2014), foi feito com o auxílio do programa Clustal Omega (ebi.ac.uk/Tools/msa/clustalo/) (SIEVERS *et al.*, 2014).

Em seguida, as sequências da PTOX de *Glycine max*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Lotus japonicus*, *Lupinus angustifolius*,

Medicago truncatula, *Phaseolus vulgaris*, *Vigna angularis* e *Vigna radiata* foram obtidas dos genomas sequenciados dessas espécies depositados no banco de dados GenBank através de alinhamento pela ferramenta BLAST usando como “isca” uma sequência de aminoácidos da PTOX do gênero *Vigna*. A anotação dos genes e a dedução dos DNAs complementares (cDNAs), ou seja, a determinação de éxons, íntrons e janelas abertas de leitura (ORFs, *open reading frame*) foram feitas manualmente, de acordo com a estratégia proposta por Costa *et al.* (2014), através da comparação das sequências identificadas com os cDNAs da PTOX do GenBank usando BLASTn. Os cDNAs deduzidos (a partir de DNA genômico) identificados nesse estudo foram traduzidos em sequências de aminoácidos usando a ferramenta *TranslateTool*, disponível no servidor ExPASy. Todas as proteínas deduzidas foram validadas alinhando suas sequências contra proteínas PTOX disponíveis no GenBank usando a ferramenta BLASTp.

A análise filogenética foi conduzida a partir das sequências de aminoácidos das proteínas PTOX encontradas em *G. max* e nas outras espécies de leguminosas citadas anteriormente. As sequências de *Arabidopsis thaliana* (Brassicales), *Oriza sativa* (Poales) e *Volvox carteri* (Chlamydomonadales) foram utilizadas como grupo externo. A árvore filogenética foi construída através do programa MEGA 7.0 (KUMAR *et al.*, 2016), utilizando o método de agrupamento *Neighbor-Joining* (SAITOU e NEI, 1987). Os seguintes parâmetros foram estabelecidos: método *Bootstrap* com 1000 replicatas, modelo de substituição “número de diferenças” e “completa deleção” na opção *Gaps/ tratamentos de dados perdidos*. As sequências de aminoácidos usadas nesta análise estão disponíveis no **Anexo A**.

3.2 Expressão gênica da PTOX em *Glycine max*

O perfil de expressão dos genes da PTOX (denominados neste trabalho como *PTOX 1* e *PTOX 2*) foi avaliado em diferentes órgãos durante o desenvolvimento da soja e sob condições de estresse, tanto através de PCR quantitativa em tempo real (RT-qPCR) quanto a partir de dados transcriptômicos disponíveis em bancos de RNA-seq.

3.2.1 Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR)

As condições experimentais e todo o procedimento adotado para a realização do ensaio de PCR quantitativa em tempo real (RT- qPCR) são descritos a seguir.

3.2.1.1 Material vegetal e condições de crescimento

As sementes do cultivar BRS Pala de soja [*Glycine max* (L.) Merrill] utilizadas foram cedidas pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa Trigo, Passo Fundo, Rio Grande do Sul) e encaminhadas ao Laboratório de Genômica Funcional e Bioinformática da Universidade Federal do Ceará (UFC), onde foram transferidas para frascos hermeticamente fechados e armazenadas em geladeira a 10 °C.

As sementes sem sinais de danos mecânicos foram selecionadas e esterilizadas em solução de hipoclorito de sódio (NaClO) 1% durante cinco minutos, a fim de eliminar possíveis patógenos em seu tegumento. Em seguida, foram deixadas em água destilada durante duas horas e, posteriormente, lavadas exaustivamente em água corrente e em água destilada.

O passo seguinte consistiu na germinação das sementes esterilizadas em papel de filtro embebido com a solução de Hoagland (HOAGLAND e ARNON, 1950) 25% em casa de vegetação (temperatura de 30 ± 1 °C durante o dia e 25 ± 1 °C durante a noite, fotoperíodo de 12 horas, umidade de 60 ± 8 %). Seis dias após a semeadura, as mudas foram transferidas para um sistema hidropônico em vasos de 5L contendo a solução nutritiva de Hoagland concentrada a 100%, sendo renovada semanalmente. As plantas foram divididas em dois grupos: o grupo 1, contendo mudas mantidas nas condições normais de crescimento anteriormente descritas, a fim de avaliar os diferentes estádios de desenvolvimento da soja; e o grupo 2, também contendo mudas mantidas nas condições já descritas, mas tratadas com polietilenoglicol (PEG), a fim de simular um estresse osmótico, ou com o fitormônio ácido salicílico (AS).

Para o grupo 1, as coletas dos órgãos para extração de RNA total foram feitas com 10, 22 e 45 dias após a semeadura (DAS) e 9 e 18 dias após a floração (DAF), que correspondem a 54 e 63 DAS, respectivamente. As partes coletadas foram raízes, folhas unifolioladas e trifolioladas, cotilédones, hipocótilos, epicótilos, flores e vagens. Além destes, também foram coletadas sementes no início da germinação (semente seca e semente embebida por 24 horas na solução de Hoagland 25 %), conforme apresentado na **Tabela 2**. No grupo 2, as plantas foram tratadas com polietilenoglicol (PEG) 100 g.L⁻¹ adicionado à solução nutritiva ou com ácido salicílico (AS) 0,5 mM borrifado sobre as folhas. Para as duas condições, a indução ocorreu 10 dias após a semeadura. Após a indução, folhas trifolioladas e raízes de todos os tratamentos (incluindo o controle) foram coletadas após 0, 6, 12 e 24 horas após a aplicação dos tratamentos.

Cada amostra foi constituída pelo material coletado de pelo menos três plantas ou sementes. Após a coleta dos órgãos dos dois grupos, as amostras foram imediatamente congeladas em nitrogênio líquido e armazenadas a -80 °C até o momento da extração de RNA.

Tabela 2 – Órgãos coletados e idade das plantas usadas para os ensaios durante o desenvolvimento de *Glycine max*.

Parte extraída	Idade da planta
Semente seca	-
Semente embebida na solução nutritiva de Hoagland	24 h após a embebição
Folha unifoliolada e cotilédones	10 DAS e 22 DAS
Folha trifoliolada, raiz, epicótilo e hipocótilo	10, 22 e 45 DAS
Flores	45 DAS
Vagens	9 e 18 DAF

Fonte: Autoria própria. Legenda: DAS – dias após a semeadura; DAF – dias após a floração.

3.2.1.2 Extração de RNA total

Aproximadamente 200 mg de material vegetal de cada uma das amostras foi macerado em nitrogênio líquido com auxílio de almofariz e pistilo. O *RNeasy Plant mini Kit* (Qiagen) foi utilizado para a extração de RNA total do material macerado, de acordo com as instruções do fabricante.

Foram adicionados às amostras 700 µL do tampão RLT e 10 µL de β-mercaptoetanol, homogeneizando-as em agitador vortex. A solução obtida foi transferida para a coluna lilás do kit (QIAshredder spin column, acoplada a um microtubo coletor de 2 mL) e centrifugada a 25 °C por dois minutos a 14.000 rpm. O eluído da coluna foi recuperado e transferido para um novo microtubo, onde foram adicionados 400 µL de etanol 95 %, sendo a solução homogeneizada por inversão.

Em seguida, a solução foi transferida para a coluna rósea do kit (RNeasy mini spin column, também acoplada a um microtubo coletor de 2 mL) e centrifugada a 25 °C por 15 segundos a 10.000 rpm. O eluído foi descartado e adicionados à coluna 350 µL do tampão RW1, seguido de centrifugação a 25 °C por 15 segundos a 10.000 rpm. Logo após, foram adicionados 80 µL de desoxirribonucleases (DNase) à coluna, a qual foi deixada em temperatura ambiente por 15 minutos. Após este período, adicionou-se novamente 350 µL

do tampão RW1 e fez-se nova centrifugação a 25 °C por 15 segundos a 10.000 rpm. Em seguida, a coluna foi acoplada a um novo microtubo coletor e adicionados 500 µL do tampão RPE à solução, que foi submetida a centrifugação a 25 °C por dois minutos a 10.000 rpm, a fim de eliminar possíveis resíduos de etanol que poderiam interferir na análise.

O eluído e o tubo coletor foram descartados e a coluna foi acoplada a um novo tubo coletor de 1,5 mL, onde foram adicionados 40 µL de água livre de RNases, submetendo a nova centrifugação por 1 minuto a 10.000 rpm para eluição do RNA. O RNA recuperado foi armazenado a -20 °C para as etapas posteriores.

3.2.1.3 Quantificação, pureza e integridade do RNA total

A concentração e pureza das amostras foram determinadas em espectrofotômetro Nanodrop 2000 (Thermo Scientific) em comprimentos de onda de 260 nm. A quantificação foi feita em triplicatas, sendo feita uma média das três leituras. Para estimar a pureza do RNA extraído, a razão entre as densidades ópticas (DO) de 260/280 nm foi utilizada para identificar contaminações por proteínas (razão deve estar entre 1,8 e 2,0) e a razão de 260/230 nm para identificar contaminações por polissacarídeos (deve estar entre 2,0 e 2,2).

A integridade do RNA total e a contaminação com DNA genômico foram verificadas por eletroforese em gel de agarose a 1,5 % a partir de 0,5 µg de RNA. O gel foi corado com brometo de etídio 0,5 µg/mL durante 10 minutos. Para a visualização das bandas correspondentes ao RNA ribossomal (18S e 28S) no gel, foi utilizado um transluminador de luz ultravioleta (UV), sendo o RNA fotodocumentado pelo sistema Mini BIS Pro (Bio-Imaging Systems) com o auxílio do software GEICapture™.

3.2.1.4 Síntese de DNA complementar (cDNA)

O RNA quantificado e submetido às análises de pureza e integridade anteriormente descritas foi, em seguida, utilizado para a síntese do DNA complementar (cDNA) com o kit da IMpromII™ Transcriptase Reverse (Promega, Estados Unidos), de acordo com as instruções do fabricante.

Para a reação, foram adicionados aproximadamente 1 µg de RNA total de cada amostra, oligodT₂₄, cloreto de magnésio (MgCl₂) 25 mM, dNTPs 10 mM, tampão de

reação 5X e água livre de RNases (Qiagen), totalizando um volume de 19 μL . A **Tabela 3** apresenta os reagentes com os respectivos volumes utilizados para esta reação.

A solução obtida foi submetida a 65 °C por 5 minutos para desnaturação e, em seguida, transferida para o gelo. Logo após, 1 μL da enzima transcriptase reversa foi adicionado em cada tubo de reação, completando o volume para 20 μL .

A reação de síntese da fita complementar ocorreu em termociclador (Techne TC-512) a 42 °C, durante 1 hora, seguida de um passo de desnaturação enzimática a 75 °C por 15 minutos. O cDNA obtido foi armazenado a -20 °C para os ensaios posteriores de qPCR, antes das quais foi diluído para uma concentração final de 25 ng/ μL , considerando-se que a concentração inicial era em torno de 1000 ng/ μL .

Tabela 3 – Reagentes e volumes utilizados na reação de síntese de DNA complementar.

Reagentes	Volumes
RNA total	Variável em função da concentração após a quantificação
Tampão de reação 5X	4 μL
MgCl ₂ (25 mM)	2,4 μL
dNTPs (10 mM)	1,0 μL
OligodT ₂₄ (20 $\mu\text{mol}/\mu\text{L}$)	3,0 μL
H ₂ O livre de RNases	Variável em função do volume de RNA total
Transcriptase reversa (160 $\mu\text{g}/\mu\text{L}$)	1 μL

Fonte: Autoria própria.

3.2.1.5 Oligonucleotídeos iniciadores (*primers*) dos genes PTOX e dos genes normalizadores

Os *primers* (senso e antisenso – *forward* e *reverse*) utilizados para as reações de qPCR foram desenhados por meio do programa Perl Primer V1.1.20 (MARSHALL, 2004) a partir da região 5' não traduzida (5' UTR) de cada gene da PTOX, uma vez que essas regiões são menos conservadas em comparação às regiões traduzidas, que apresentam alta identidade entre suas sequências nucleotídicas. Na **Tabela 4** são apresentadas as sequências dos *primers* dos dois genes da PTOX, identificados em *G. max*, e dos genes de referência *EF1 β* , *SKIP 16*, *MTP*, *UKN1*, *EF1 α 1a1*, *EF1 α 1a2*, *EF1 α 2a*,

EF1α 2b e *EF1α 3*, reportados por Hu *et al.* (2009) e Saraiva *et al.* (2014) como estáveis. Tais genes de referência foram usados como normalizadores para o cálculo da expressão relativa dos genes *PTOX*. O alinhamento entre os genes *PTOX* de *G. max* identificando a região dos *primers* e os genes normalizadores utilizados para cada parte da planta são encontrados nos **Anexos B e C**, respectivamente.

3.2.1.6 Diluição, gradiente de temperatura e curva de eficiência dos primers

Os *primers* adquiridos da empresa IDT[®] (*Integrated DNA Technologies*) foram diluídos para uma concentração de aproximadamente 20 μM e posteriormente quantificados em triplicata em espectrofotômetro Nanodrop 2000. Da média das concentrações obtidas, foi calculada a concentração de cada *primer*, os quais foram diluídos para uma concentração de 6 μM, a fim de se obter uma concentração final de 300 nM, a qual foi utilizada nas reações de qPCR.

A temperatura de anelamento ótima para cada par de *primer* foi determinada através de um gradiente de temperatura usando um *pool* das amostras de cDNA. O procedimento foi realizado em ensaio de PCR em tempo real no termociclador Mastercycler ep Realplex (Eppendorf[®]). As temperaturas usadas foram 47,2; 50,5; 52,7; 55; 57,5; 59,8; 63,6 e 65 °C. As temperaturas de menor C_t (*threshold cycle*) e com as curvas mais sigmóides foram selecionadas.

Para mensurar a eficiência dos *primers*, foram feitas diluições seriadas do *pool* de cDNA (1:1, 1:10, 1:100, 1:1000, 1:10000). A curva também foi feita através de qPCR. A eficiência foi calculada usando a equação $Ef = 10^{(-1/\text{slope})} - 1$.

Tabela 4 – Oligonucleotídeos iniciadores utilizados nas análises de expressão por Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR) de *Glycine max*.

Gene	Produto gênico	Sequência do primer (5'-3')	Tamanho do amplicon (pb)
<i>PTOX 1</i>	Oxidase terminal do plastídio 1	Fwd 5' GCCGTAATGCCGCTATG 3' Rev 5' AGAGGGAACCTTGTGTGTGT 3'	153
<i>PTOX 2</i>	Oxidase terminal do plastídio 2	Fwd 5' ACACACAAAAGTTCCCTCCTC 3' Rev 5' CCTGAAAAGTTGTGGAGCGG 3'	129
<i>EF1α 1a1</i>	Fator de alongação 1 α	Fwd 5' GATTCATGTAGCCGTAGCC 3' Rev 5' ATTTAAGACATCCCTCCTCAG 3'	182
<i>EF1α 1a2</i>	Fator de alongação 1 α	Fwd 5' ATCATCGTGGTTACTCCTTAT 3' Rev 5' TCAGACTCTTACCATCA 3'	199
<i>EF1α 2a</i>	Fator de alongação 1 α	Fwd 5' GGATGTCGTTTCTTATGGT 3' Rev 5' CAAACACAACACATTA AACAG 3'	189
<i>EF1α 2b</i>	Fator de alongação 1 α	Fwd 5' TTTCTGTACTCTTGTGTCTTCT 3' Rev 5' ACAAACACAACACATTA AACAC 3'	159
<i>EF1α 3</i>	Fator de alongação 1 α	Fwd 5' GTGCTGGATGGTTTATCAAGG 3' Rev 5' AAAGTTGCGAGATGACGGAG 3'	140
<i>EF1β</i>	Fator de alongação 1 β	Fwd 5' CCACTGCTGAAGAAGATGATGATG 3' Rev 5' AAGGACAGAAGACTGCCACTC3'	134
<i>MTP</i>	Metaloprotease	Fwd 5' CGCTCCAAGTGCTCCTCATTAG 3' Rev 5' TGAAGTAACCGACGCCAACG 3'	71
<i>SKIP 16</i>	Proteína de interação 16	Fwd GAGCCCAAGACATTGCGAGAG Rev CGGAAGCGGAAGA ACTGAACC	60
<i>UKNI</i>	Proteína hipotética	Fwd 5' TGGTGCTGCCGTATTTACTG 3' Rev 5' GGTGGAAGGA ACTGCTAACAAATC 3'	74

Fonte: Autoria própria. Legenda: pb – pares de base

3.2.1.7 PCR semi-quantitativa

A reação de PCR semi-quantitativa foi realizada para avaliar a qualidade do cDNA obtido (como descrito no tópico 3.2.1.4). Utilizou-se para tal procedimento um par de *primers* referente ao gene *EF1α 1a*. Para cada reação foi utilizado 1 μL de cDNA, tampão de reação 5X, 1 μL de dNTPs (5 mM), 1 μL de cada iniciador específico e 0,1 μL da enzima Go Taq DNA polimerase. O volume total foi completado com água livre de RNases, totalizando 25 μL. O resultado da amplificação foi visualizado em gel de agarose a 2 %.

3.2.1.8 Ensaio de PCR em tempo real (RT-qPCR)

Cada reação foi realizada em um volume final de 20 μL, que continha 4 μL de cDNA (100 ng), 10 μL do Power SYBR Green PCR Master Mix (Applied Biosystems, Warrington, UK), 1 μL de cada iniciador específico (300 nM cada) e 4 μL de H₂O ultrapura. Todas as reações foram realizadas em triplicata e um controle negativo, o qual não continha cDNA, foi realizado em paralelo. As reações de amplificação foram realizadas através de 40 ciclos térmicos, compostos por 15 s a 95 °C, seguido por um período de 15 s a uma temperatura de anelamento específica para cada iniciador e, por fim, 20 s a 60 °C. A desnaturação inicial foi realizada a 95 °C por 10 min. Para verificar a especificidade dos *primers*, também foi realizada a curva de dissociação (*melting*), que consistiu de 15 segundos de incubação a 95 °C, 15 segundos de incubação a 60 °C e uma subida de 95 °C.

O monitoramento em tempo real da PCR foi realizado em um termociclador RealPlex 4S (Eppendorf®) por meio da detecção dos níveis de fluorescência. As análises dos dados foram realizadas pelo software Realplex. Todas as reações, tanto dos genes alvo quanto dos genes de referência, foram realizadas em triplicatas.

Os níveis de expressão gênica foram determinados utilizando a média dos valores C_t (*Cycle threshold*). Os valores brutos de C_t foram importados para o programa qbase^{plus} (versão 2.4) e convertidos em dados de expressão relativa através do método do fator de normalização (HELLEMANS *et al.*, 2007). Os dados de expressão relativa do experimento que avaliou a expressão da PTOX durante o desenvolvimento foram expressos como média ± desvio padrão e os dados de expressão referentes aos tratamentos com PEG e AS foram convertidos em valores de \log_2 fold-change (\log_2FC), utilizando a seguinte equação: $\log_2(\text{valor de expressão relativa da amostra tratada}) - \log_2(\text{valor de expressão relativa da amostra controle})$.

3.2.2 Expressão *in silico* via RNA-seq

A análise de expressão diferencial também foi avaliada em dados de RNA-seq disponíveis no banco SRA (*Sequence Read Archive*) do NCBI, que contém, dentre outros, transcriptomas de *Glycine max* submetidos a diferentes estresses, bem como em diferentes partes da planta e estádios do desenvolvimento. A pesquisa pelos “bioprojetos” foi restringida à “*Glycine max*” e os experimentos selecionados foram baixados no formato “.fastq”. A **Tabela 5** apresenta todos os bioprojetos utilizados neste trabalho e detalha cultivares, tratamentos, número de replicatas, parte da planta avaliada e informações gerais de cada experimento.

3.2.2.1 Obtenção das sondas

As sequências gênicas (sondas) utilizadas como referência para serem mapeadas contra os diferentes experimentos do SRA foram selecionadas das regiões 3’ e 5’ não traduzidas do DNA complementar (cDNA) deduzido dos genes *PTOX 1* e *PTOX 2* da soja. Para desenhar as sondas, foi observado se não haviam regiões idênticas com mais de 49 pares de bases (pb) consecutivos entre as sequências dos genes, uma vez que os experimentos de SRA são compostos de dezenas de milhões de pequenas sequências (*reads*) contendo entre 50 e 150 pb. As sequências das sondas para os dois genes podem ser encontradas no **Anexo B**.

Tabela 5 – Experimentos de RNA-seq utilizados para analisar a expressão dos genes *PTOX* em *Glycine max*.

Identificador do Bioprojeto	Cultivar	Parte da planta	Tratamento/ Detalhes do experimento	Número de replicatas	Referência
PRJNA246315	Williams 82	Semente (estádio globular)	Desenvolvimento/ Oito compartimentos do estágio globular da semente foram analisados	Quatro	Não encontrado
PRJNA246783	Williams 82	Semente (estádio cotiledonar)	Desenvolvimento/ Dez compartimentos do estágio cotiledonar da semente foram analisados	Três ou quatro	Não encontrado
PRJNA197251	Williams 82	Semente (maturação precoce)	Desenvolvimento/ Dezesete compartimentos do estágio de maturação precoce da semente foram analisados	Três ou quatro	Danzer <i>et al.</i> (2015)
PRJNA262564	Williams 82	Cotilédones (3 estádios) e folhas trifolioladas (5 estádios)	Desenvolvimento e senescência/ Cotilédones coletados aos 4, 15 e 27 dias após a semeadura e folhas coletadas a partir do estágio V3 (1º estágio) e após 21, 35, 49 e 56 dias.	Três	Brown & Hudson (2015)

Continua

Tabela 5 (Continuação)– Experimentos de RNA-seq utilizados para analisar a expressão dos genes *PTOX* em *Glycine max*.

Identificador do Bioprojeto	Cultivar	Parte da planta	Tratamento/ Detalhes do experimento	Número de replicatas	Referência
PRJNA246058	Williams 82	Raízes	Seca e estresse salino/ Amostras coletadas nos tempos 0 (controle), 1, 6 e 12 horas sob desidratação ou tratadas com NaCl 100 mM	Três	Belamkar <i>et al.</i> (2014)
PRJNA259941	Benning (sensível a seca) e PI416937 (tolerante a seca)	Folhas trifolioladas	Seca/ Plantas no estágio R2 foram expostas ao ar. Folhas foram coletadas nos tempos 0, 6, 12 e 24 horas	Três	Shin <i>et al.</i> (2015)
PRJNA266538	Williams 82	Zona de abscisão foliar (LAZ) e pecíolo (NAZ)	Etileno/ Explantes provenientes de plantas com 21 dias de vida foram expostos a 25 µL/L de etileno de 0 a 72 horas	Três	Kim <i>et al.</i> (2015)

Fonte: Autoria própria.

3.2.2.2 Ferramentas computacionais para cálculo da expressão diferencial

O mapeamento (alinhamento) entre as sondas e os experimentos do banco de SRA foi realizado pelas ferramentas Bowtie2 v2.3.2 (<http://bowtie-bio.sourceforge.net/index.shtml>) (LANGMEAD E SALZBERG, 2013) e TopHat2 v2.0.13 (<http://tophat.cbcb.umd.edu/>) (KIM *et al.*, 2013). Somente as *reads* mapeadas foram fornecidas ao pacote Cufflinks (<http://cufflinks.cbcb.umd.edu/>) (TRAPNELL *et al.*, 2012) para a montagem do transcriptoma de cada biblioteca (condição/tratamento). A ferramenta Cuffmerge (incluída no pacote Cufflinks) é a responsável por essa montagem, na qual é gerado um arquivo “.gtf” que contém informações coordenadas do mapeamento das *reads* por transcrito e tratamento. Finalmente, através da ferramenta Cuffdiff (também incluída no pacote Cufflinks), o número de *reads* mapeadas por transcrito foi usado para estimar os genes *PTOX* diferencialmente expressos. Todas as etapas descritas foram operadas por linhas de comando através do terminal do sistema operacional LINUX (Ubuntu 16.04).

Os dados de expressão em FPKM (*Fragments Per Kilobase transcript per Million of mapped reads*) dos experimentos que avaliaram a expressão dos genes *PTOX* durante o desenvolvimento (sementes, cotilédones e folhas) foram expressos como média \pm desvio padrão das replicatas. Já os valores médios de FPKM dos tratamentos nos experimentos que avaliaram diferentes condições de estresse foram convertidos em \log_2 fold-change (\log_2FC) pela equação: $\log(\text{tratado};2) - \log(\text{controle};2)$. A análise de variância (ANOVA) foi realizada através do software GraphPad Prism 5.02, seguida do teste de Bonferroni para comparação entre médias e com nível de significância de 5 % ($p < 0,05$).

3.2.3 Análise da região promotora

A presença de elementos *cis* regulatórios foi analisada na região correspondente aos 1000 pares de bases que antecedem o sítio de início da tradução dos dois genes *PTOX* da soja. Para esta análise, as sequências foram submetidas aos servidores PLACE (sogo.dna.affrc.go.jp/PLACE) (HIGO *et al.* 1999), PlantPAN (plantpan.mbc.nctu.edu.tw) (CHOW *et al.* 2015), AtPAN (atpan.itps.ncku.edu.tw/) (CHEN *et al.* 2012) e PlantCARE (bioinformatics.psb.ugent.be/webtools/plantcare/html/) (LESCOT, 2002).

3.3 Análise estrutural das proteínas PTOX de *Glycine max*

A fim de avançar na compreensão da função das duas proteínas PTOX em soja, foi feita análise por meio da modelagem das estruturas tridimensionais e ensaio de docagem molecular com o ligante (substrato) específico das isoformas. Em todas estas análises foram utilizadas as proteínas maduras, ou seja, sem os 36 primeiros aminoácidos da sequência, a contar da extremidade amino-terminal e referente ao peptídeo trânsito, uma vez que essas sequências são clivadas quando a proteína é importada para o cloroplasto (BRUCE, 2000).

3.3.1 Predição e validação da estrutura tridimensional

A predição da estrutura tridimensional das proteínas PTOX da soja foi realizada pelo programa Modeller 9.18 (WEBB e SALI, 2014), que utiliza predição por modelagem comparativa ou homologia e pelo servidor Robetta (<http://robeta.bakerlab.org>), que fornece a predição tanto pelo método de primeiros princípios (*ab initio*) quanto por modelagem comparativa (KIM *et al.*, 2004).

Para a predição usando o Modeller, foi utilizado como molde a estrutura cristalizada da oxidase alternativa (AOX) de *Trypanosoma brucei*, que está depositada no Banco de Dados de Proteínas (PDB – *Protein Data Bank*) sob o identificador 3VV9 (DOI 10.2210/pdb3vv9/pdb) e apresenta 27 % de identidade (alinhamento feito pela ferramenta BLASTp) com as proteínas PTOX da soja. Para a execução da modelagem, seguiu-se um roteiro com algumas linhas de comando utilizadas na interface do programa, bem como foram utilizados dois arquivos, um no formato “.py” e outro no formato “.ali”, este último contendo o alinhamento da sequência alvo com a sequência do molde. Foram geradas 200 replicatas (modelos) para cada uma das proteínas, sendo aquela com menor função objetiva, ou seja, a mais otimizada geometricamente, a selecionada para as análises seguintes. Já para a predição por meio do Robetta, as sequências primárias das proteínas foram submetidas ao servidor, que retornou após alguns dias cinco modelos preditos, dos quais apenas o que apresentou menor valor de desvio médio dos átomos (RMSD, *Root Mean Square Deviation*) foi selecionado. Os íons ferro (Fe) foram adicionados aos modelos tomando como base as coordenadas do Fe do modelo 3VV9, uma vez que PTOX e AOX pertencem à classe de proteínas Dox (proteínas difero carboxilato quinol oxidases).

Para a validação das estruturas tridimensionais geradas foram usados os servidores SAVES (*Structure Analysis and Verification Server*, disponível em

services.mbi.ucla.edu/SAVE) (HOOFT, 1996) e MolProbity (molprobity.biochem.duke.edu/) (CHEN, 2010), que fornecem os valores de *Z-score* e o gráfico de Ramachandran, sendo estes os critérios escolhidos para a seleção dos modelos que participariam da análise de docagem molecular. O *Z-score* deve estar o mais próximo possível de zero e representa o desvio dos volumes atômicos da estrutura a ser validada em relação a valores padrão, ou seja, de estruturas altamente resolvidas e refinadas depositadas no PDB (PONTIUS, RICHELLE e WODAK, 1996) e o gráfico de Ramachandran fornece a porcentagem de resíduos de aminoácidos em posições favoráveis na estrutura (RAMACHANDRAN, RAMAKRISHNAN e SASISEKHARAN, 1963).

3.3.2 *Predição de estruturas secundárias*

A fim de dar maior confiabilidade às predições de estrutura tridimensional, as sequências proteicas das duas proteínas PTOX da soja foram submetidas ao servidor PSIPRED *Protein Sequence Analysis* (bioinf.cs.ucl.ac.uk/psipred/), sendo selecionadas as ferramentas PSIPRED v3.3, que prediz as estruturas secundárias da sequência, DISOPRED3, que prediz regiões de desordem intrínseca na estrutura e MEMSAT-SVM, que faz a predição de hélices transmembranares. (BUCHAN, 2013; JONES, 1999). Para esta última análise, também foi usada a ferramenta TMPred (HOFMANN e STOFFEL, 1993), disponível no portal ExpASy.

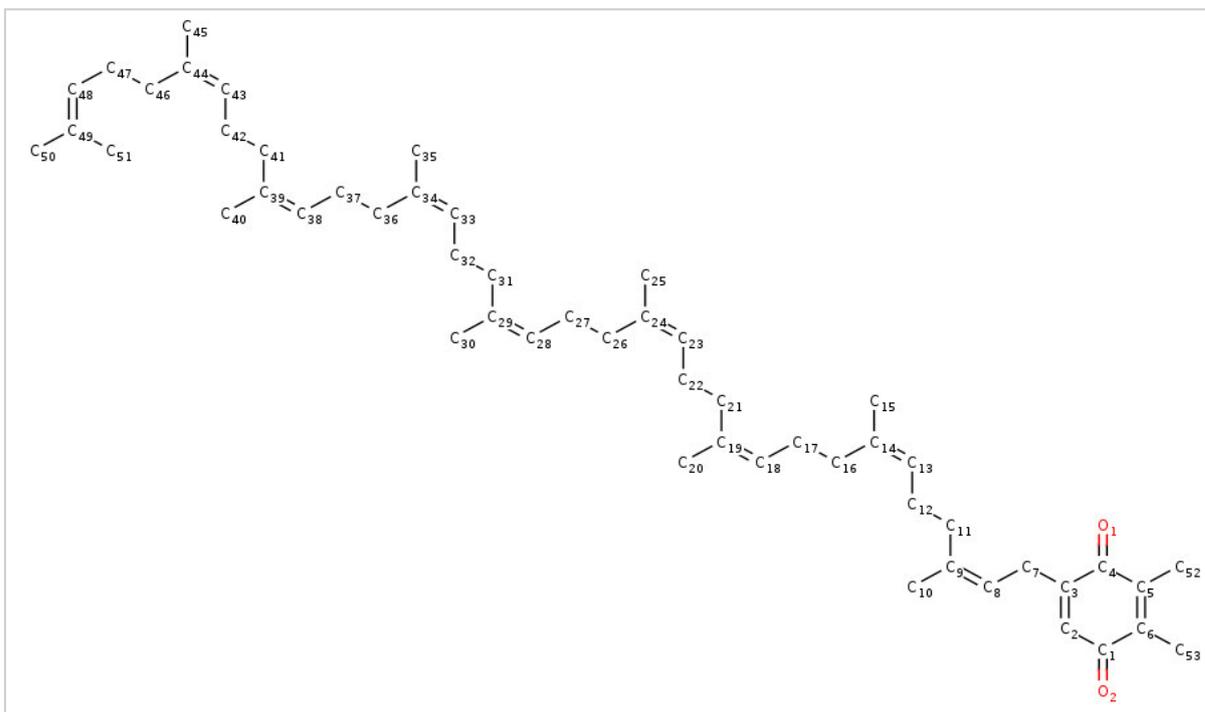
3.3.3 *Docagem molecular das proteínas PTOX de Glycine max*

A análise de *docking* ou atracamento molecular foi conduzida com as estruturas tridimensionais da PTOX geradas e o seu substrato (ligante), o plastoquinol (PQH₂). O ligante na sua forma não protonada (plastoquinona) foi obtido do PDB (identificador PL9) no formato “.sdf”, sendo em seguida convertido à sua forma reduzida pela adição de dois prótons através do software Avogrado (HANWELL, 2012), que também realizou a conversão para o formato “.pdb”. A **Figura 5** representa a estrutura da plastoquinona.

A interação entre as proteínas (PTOX 1 e PTOX 2) e o seu ligante foi analisada pelo software AutodockVina (TROTT e OLSON, 2010), que identifica os melhores modos de ligação da molécula do ligante no sítio de ligação da proteína. O programa requer, além das estruturas da proteína e do ligante no formato “.pdb”, que o usuário forneça o sítio de ligação. Para este trabalho, as coordenadas do sítio de ligação (centro de reação) fornecidas foram as

do carbono alfa do aminoácido histidina na posição 95, que está presente no motivo MSVLHMYES, sequência conservada nas duas proteínas e que é essencial para a sua atividade (JOSSE *et al.*, 2003). Também foram ajustados os seguintes parâmetros: tamanho da caixa, que corresponde à área na qual o ligante poderá rotacionar, para 50 angstroms (Å); “*exhaustiveness*”, que consiste no número de corridas realizadas pelo programa, para 150 e nove modos de ligação. Foram realizadas duas docagens, uma para cada proteína PTOX. Os complexos receptor-ligante ideais foram escolhidos com base na menor energia de interação (kcal.mol^{-1}), no menor valor de RMSD obtido e na menor distância entre o átomo de ferro e o hidrogênio do ligante envolvido na transferência de elétrons, que foi medida pela ferramenta “Measurement” no visualizador PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). Cada complexo ideal foi submetido à construção do diagrama de interação com o ligante pelo programa Maestro 11.1 (SCHRÖDINGER RELEASE 2017-2) considerando a distância padrão de 4 Å.

Figura 5 – Estrutura da plastoquinona.



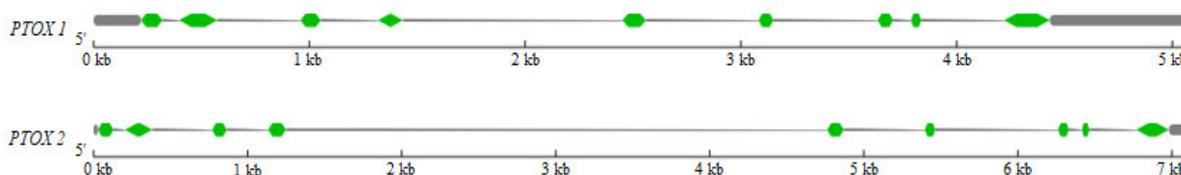
Fonte: *Protein Data Bank* (PDB). Fórmula molecular: $\text{C}_{53}\text{H}_{80}\text{O}_2$. Dois prótons foram adicionados aos átomos de oxigênio (em vermelho) pelo programa Avogadro para obtenção do plastoquinol, ligante da PTOX utilizado no *docking* molecular.

4 RESULTADOS

4.1 Caracterização dos genes *PTOX* em *Glycine max* e distribuição da proteína em leguminosas

A identificação e caracterização dos dois possíveis genes codificando a oxidase terminal do plastídio em soja feita por Sun *et al.* (2005) revelou que esses parálogos estão em cromossomos distintos, o 9 e o 15. A partir dessas informações, neste trabalho, o gene alocado no cromossomo 9 foi denominado *PTOX 1* e o localizado no cromossomo 15, *PTOX 2*. Ambos os genes possuem nove éxons e oito íntrons (**Figura 6**). O quarto e o sexto íntrons de *PTOX 2* apresentam 2442 e 313 pares de bases (pb) a mais que os correspondentes em *PTOX 1*, respectivamente. Já a sequência da região codante é altamente conservada entre os dois genes. Seu tamanho é de 999 pb, os quais codificam proteínas com 332 aminoácidos e que compartilham 97 % de identidade (alinhamento feito pela ferramenta BLAST). A proteína possui um peptídeo trânsito que corresponde aos 36 primeiros aminoácidos das sequências de *PTOX 1* e *PTOX 2* e que é responsável pelo seu transporte até o cloroplasto. A massa molecular e o ponto isoelétrico preditos para a proteína madura (sem o peptídeo trânsito) foram, respectivamente, 34,34 kDa e 5,62 para *PTOX 1* e 34,31 kDa e 5,77 para *PTOX 2*.

Figura 6 – Estrutura dos genes da oxidase terminal do plastídio (*PTOX*) de *Glycine max* representada pela ferramenta *Gene Structure Display Server 2.0*.



Fonte: Baseado em Sun *et al.* (2015). Os éxons são indicados por caixas verdes e os íntrons por linhas. Caixas em cinza representam as regiões não traduzidas 5' e 3'. Os números em kb (quilobases) indicam o comprimento das sequências gênicas.

O alinhamento entre as sequências da *PTOX* e da *AOX* da soja é exibido na **Figura 7**. Foram identificadas nove mutações na sequência de aminoácidos entre as sequências da *PTOX*, duas das quais na região do peptídeo trânsito (marcadas em rosa). Em amarelo, são identificados os resíduos de cisteína envolvidos na ligação dissulfeto que ocorre somente em *AOX*, indicando que *PTOX* é encontrada como um monômero (SUN *et al.*, 2015). Em verde, encontra-se a sequência consenso para proteínas diferro carboxilato **E-X_n**-

Figura 7 (Continuação) – Alinhamento das sequências das oxidases terminais do plastídio com as oxidases alternativas de *Glycine max* feito pelo programa Clustal Omega.

CLUSTAL O(1.2.4) multiple sequence alignment

```

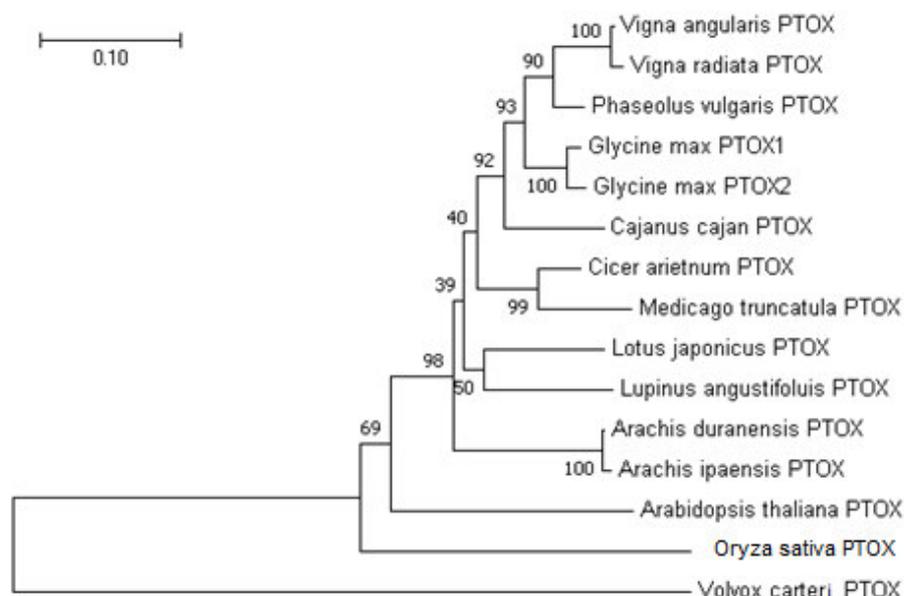
PTOX1      LRSPhSYAEDDDSSVCALEADCEGIVDCIKKSVHSNPAKVK      332
PTOX2      LRSPhSYAEDDDSSVCALEADCEGIVDCIKKSVHSNPAKVK      332
AOX1       AAPPIGYH-----                               321
AOX2a      APAPIGYH-----                               333
AOX2d      APAPIGYH-----                               326
           : * . *

```

Fonte: Baseado em Sun *et al.* (2015). Aminoácidos marcados em rosa representam as mutações entre PTOX 1 e PTOX 2. Em amarelo, os resíduos de cisteína responsáveis pela ligação dissulfeto que ocorre em AOX. Regiões azul claro indicam o sítio de ligação à quinona, com destaque para o resíduo de histidina (sublinhado) conservado em PTOX e AOX. Marcada em verde, a sequência consenso típica de proteínas diferido carboxilato. Em cinza, as regiões referentes ao éxon 8 da PTOX, não encontrada em AOX.

A busca nos bancos de dados do GenBank evidenciou que nenhuma outra espécie leguminosa analisada possui mais de um gene codificando a PTOX. A anotação dos genes indicou que o número de éxons (nove) e de íntrons (oito) do gene *PTOX* é o mesmo em todas as espécies da ordem Fabales, com exceção de *M. truncatula*, que apresentou oito éxons e sete íntrons. O alinhamento entre as proteínas deduzidas revelou identidade de 76 a 92 % entre as proteínas PTOX da soja e das demais leguminosas. A análise filogenética demonstrou, ainda, que é mais provável que tenha ocorrido um evento de duplicação em soja, o que possibilitou a presença das duas cópias do gene, ao invés da deleção de uma suposta segunda cópia do gene nas demais plantas desta ordem (**Figura 8**).

Figura 8 – Árvore filogenética dos genes da oxidase terminal do plastídio (PTOX) em leguminosas feita pelo programa MEGA 7.0.



Fonte: Autoria própria. Os genes *PTOX* de *Arabidopsis thaliana*, *Oryza sativa* e *Volvox carteri* foram utilizados como grupo externo.

4.2 Expressão gênica da PTOX em *Glycine max*

4.2.1 Reação em cadeia da DNA polimerase quantitativa em tempo real (RT-qPCR)

4.2.1.1 Qualidade do RNA

A quantificação do RNA total extraído de vários órgãos de *G. max* em diferentes estádios de desenvolvimento e do RNA total isolado de raízes e folhas tratadas com PEG ou AS mostrou que a relação de densidade óptica $OD_{260/280}$ (para averiguar contaminações com proteínas) obedeceu aos limites estabelecidos, variando de 2 a 2,15, sendo a qualidade do RNA considerada boa. Já a relação $OD_{260/230}$ (para averiguar contaminações com polissacarídeos) variou de 1,0 a 2,29, e, apesar de não ter obedecido de forma ideal aos limites (entre 2,0 e 2,2), a qualidade também foi considerada aceitável para uso. Todos estes valores, bem como as concentrações das amostras (ng/ μ L) são apresentados na **Tabela 6**.

Com relação à integridade do RNA total extraído, os perfis observados no gel de agarose revelam que a amostra não sofreu degradação, uma vez que é possível visualizar as duas bandas referentes às subunidades 18S e 28S do RNA ribossomal (**Figura 9**).

4.2.1.2 Qualidade do DNA complementar obtido

A qualidade do cDNA produzido via transcrição reversa foi verificada através de PCR semi - quantitativa, usando todas as amostras de cDNA produzido. A banda referente ao *amplicon* do gene EF1 α 1a, usado para as reações, foi evidenciada nas diferentes amostras analisadas, demonstrando boa qualidade do cDNA produzido (**Figura 10**).

Tabela 6 – Valores médios das concentrações e relações de densidade óptica (DO) dos RNAs extraídos de vários órgãos de *Glycine max* (cultivar BRS Pala) durante o desenvolvimento e tratados com polietilenoglicol (PEG) ou ácido salicílico (AS).

Órgão ou tratamento	Concentração de RNA (ng/μL)	DO _{260/280}	DO _{260/230}
Semente seca	592,4	2,11	2,3
Semente embebida	238,0	2,15	2,07
Cotilédones 10 DAS	841,1	2,14	2,31
Cotilédones 22 DAS	136,9	2,11	2,31
Folha unifoliolada 10 DAS	591,95	2,10	1,8
Folha unifoliolada 22 DAS	370,2	2,14	2,06
Folha trifoliolada 10 DAS	783,65	2,1	2,3
Folha trifoliolada 22 DAS	545,0	2,09	2,26
Folha trifoliolada 45 DAS	246,5	2,13	2,13
Raiz 10 DAS	788,4	2,12	1,88
Raiz 22 DAS	529,6	2,11	1,76
Raiz 45 DAS	330,7	2,12	1,45
Hipocótilo 10 DAS	260,5	2,11	1,9
Hipocótilo 22 DAS	159,9	2,14	1,54
Hipocótilo 45 DAS	409,4	2,13	1,96
Epicótilo 10 DAS	817,7	2,13	1,43
Epicótilo 22 DAS	382,3	2,13	1,0
Epicótilo 45 DAS	578,3	2,11	1,78
Vagens 9 DAF	1149,7	2,15	2,04
Vagens 18 DAF	758,0	2,12	2,21
Flores 45 DAS	420,4	2,11	1,53
Controle 0 h folha trifoliolada	313,8	2,13	1,49
Controle 6 h folha trifoliolada	523,0	2,11	2,0
Controle 12 h folha trifoliolada	361,0	2,0	1,5
Controle 24 h folha trifoliolada	462,3	2,15	2,1
Controle 0 h raiz	346,7	2,0	1,62

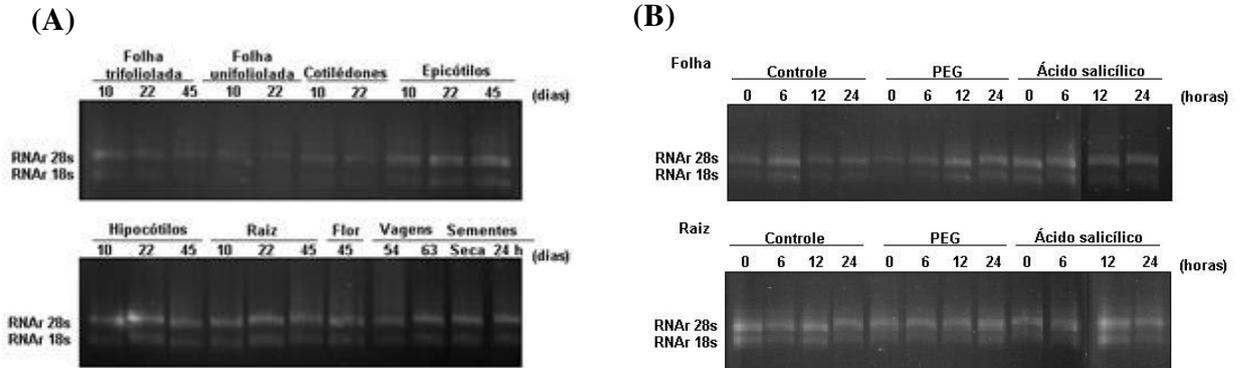
Continua

Tabela 6 (Continuação) – Valores médios das concentrações e relações de densidade óptica (DO) dos RNAs extraídos de vários órgãos de *Glycine max* (cultivar BRS Pala) durante o desenvolvimento e tratados com polietilenoglicol (PEG) ou ácido salicílico (AS).

Órgão ou tratamento	Concentração de RNA (ng/ μ L)	DO _{260/280}	DO _{260/230}
Controle 6 h raiz	587,0	2,09	2,21
Controle 12 h raiz	407,2	2,13	1,65
Controle 24 h raiz	592,7	2,04	1,42
PEG 0 h folha trifoliolada	426,2	2,1	1,51
PEG 6 h folha trifoliolada	569,1	2,12	2,29
PEG 12 h folha trifoliolada	589,1	2,09	2,0
PEG 24 h folha trifoliolada	551,9	2,13	1,80
PEG 0 h raiz	559,0	2,11	1,23
PEG 6 h raiz	369,6	2,12	1,75
PEG 12 h raiz	363,1	2,1	2,07
PEG 24 h raiz	248,2	2,11	2,05
AS 0 h folha trifoliolada	620,2	2,10	2,0
AS 6 h folha trifoliolada	371,9	2,10	1,7
AS 12 h folha trifoliolada	381,5	2,13	2,24
AS 24 h folha trifoliolada	341,26	2,12	1,39
AS 0 h raiz	449,7	2,12	1,5
AS 6 h raiz	610,2	2,08	2,17
AS 12 h raiz	682,26	2,11	2,16
AS 24 h raiz	417,2	2,12	1,7

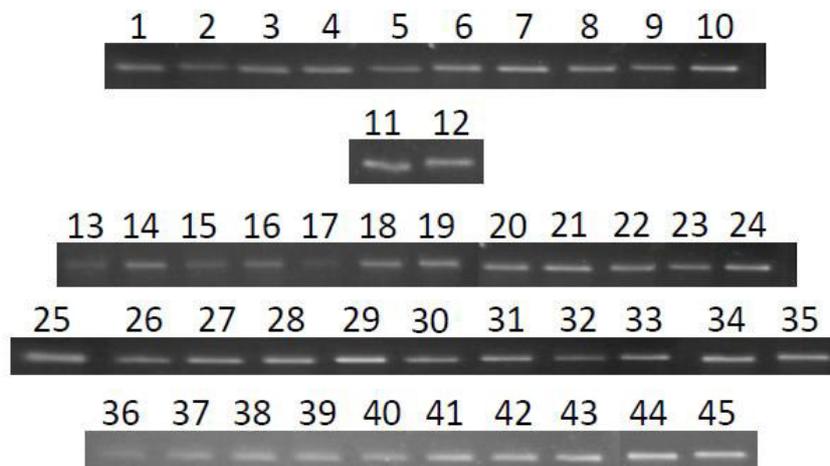
Fonte: Autoria própria. Legenda: DAS – dias após a semeadura; DAF: dias após a floração; DO – densidade óptica; PEG – polietilenoglicol; AS – ácido salicílico

Figura 9 – Análise da integridade do RNA total extraído de diversas partes de *Glycine max* (A) em diferentes estádios de desenvolvimento e (B) tratados com polietilenoglicol (PEG) ou ácido salicílico (AS) em gel de agarose 1,5 %. Destaque para as bandas correspondentes às subunidades 18S e 28S do RNA ribossomal.



Fonte: Autoria própria.

Figura 10 – Eletroforese em gel de agarose 2 % dos amplicons do gene *EF1a 1a* para avaliação da qualidade do cDNA produzido por PCR semiquantitativa.



Fonte: Autoria própria. Legenda: raias 1 a 12 – folhas de soja de plantas controle e tratadas com polietilenoglicol (PEG) ou ácido salicílico (AS). Raias 13 a 24 - raízes de soja de plantas controle e tratadas com PEG ou AS. Raias 25 a 45 – amostras de vários órgãos (semente seca, semente embebida, cotilédones, folhas unifolioladas e trifolioladas, raízes, hipocótilos, epicótilos, vagens e flores) da planta em diversos estádios do desenvolvimento, compreendendo estádios vegetativos e reprodutivos.

4.2.1.3 Temperatura de anelamento, eficiência e especificidade dos primers

A melhor temperatura de anelamento, escolhida a partir de um gradiente de temperatura, variou para cada par de *primer*. Os valores são apresentados na **Tabela 7**.

A eficiência de amplificação dos *primers* variou de 82 a 129 %, mostrando que todos os iniciadores estavam adequados para os ensaios de RT-qPCR. Os valores de eficiência dos *primers* também são mostrados na Tabela 5.

As curvas de dissociação (curvas de *melting*) analisadas durante o ensaio de RT-qPCR mostraram a presença de um único pico, evidenciando que um único produto foi amplificado e que não houve formação de dímeros de *primers* (dados não mostrados). A especificidade dos iniciadores também foi confirmada por eletroforese em gel de agarose 2 %, onde se observou a presença de um único fragmento com o tamanho esperado do *amplicon* para todos os genes testados (dados não mostrados).

Tabela 7 – Temperaturas de anelamento ótimas e eficiência de amplificação dos *primers* dos genes *PTOX* de *Glycine max* e dos genes de referência utilizados no ensaio de PCR quantitativa em tempo real (RT-qPCR).

Gene	Temperatura de anelamento dos <i>primers</i> (°C)	Eficiência de amplificação dos <i>primers</i> (%)
<i>PTOX 1</i>	57,5	94
<i>PTOX 2</i>	59,8	123
<i>EF1α 1a1</i>	59,8	113
<i>EF1α 1a2</i>	59,8	95
<i>EF1α 2a</i>	57,5	82
<i>EF1α 2b</i>	57,5	95
<i>EF1α 3</i>	63,6	129
<i>EF1β</i>	63,6	94
<i>MTP</i>	63,3	93
<i>SKIP 16</i>	50,7	102
<i>UKNI</i>	63,6	96

Fonte: Autoria própria.

4.2.1.4 Expressão dos genes *PTOX* em *Glycine max* avaliada por RT-qPCR

A análise de expressão por RT-qPCR revelou que os dois genes *PTOX* foram expressos em todos os órgãos examinados durante o desenvolvimento da soja (**Gráfico 2**), evidenciando tanto a presença quanto a funcionalidade dos parálogos da oxidase terminal do plastídio em *G. max*.

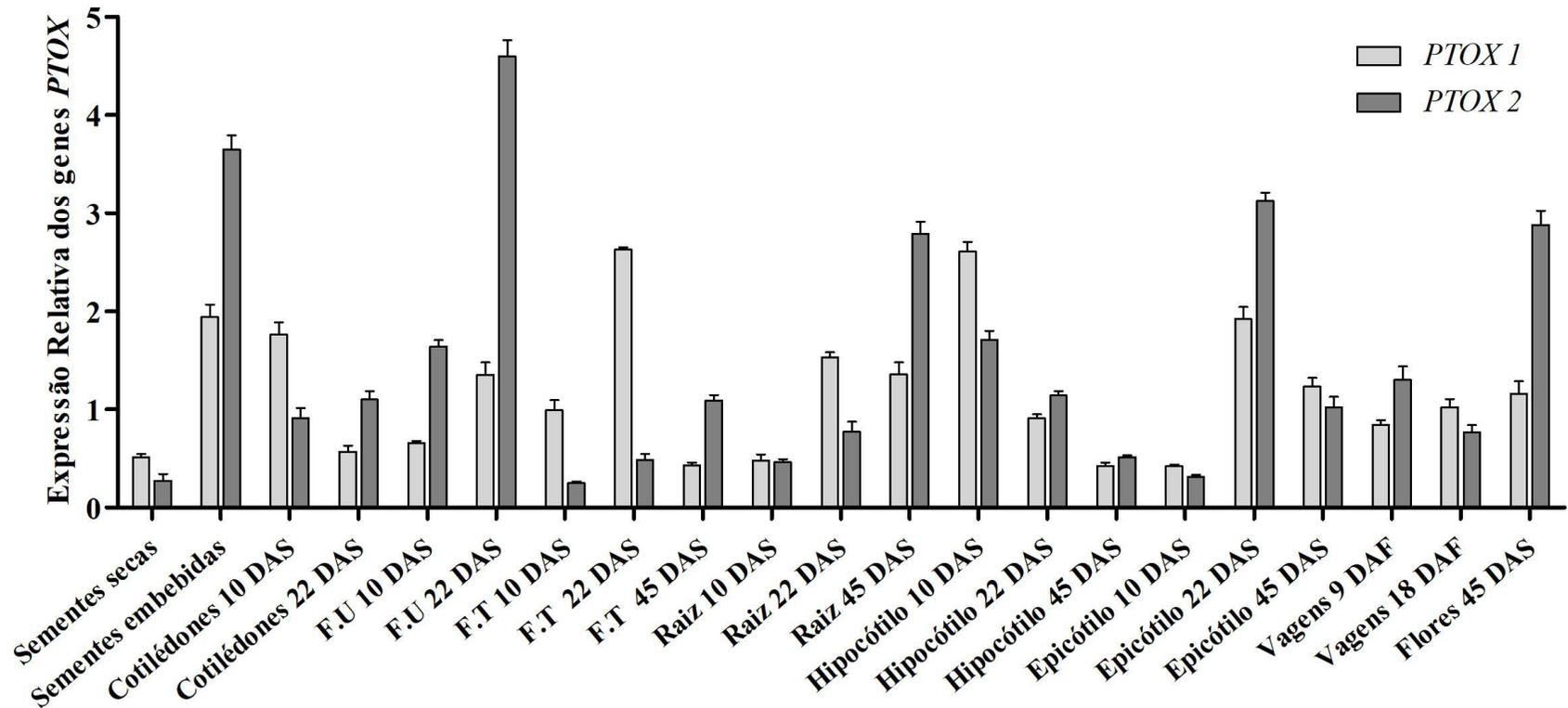
Em sementes secas, vagens, hipocótilo e epicótilo, o padrão de expressão entre os dois transcritos foi similar, com alguma variação no nível de expressão entre os genes. Já em flores e sementes embebidas o gene da *PTOX 2* apresentou maiores níveis de expressão. Um perfil de expressão interessante foi observado em cotilédones, folhas (unifolioladas e trifolioladas) e raízes (**Gráfico 2**).

Nas estruturas mais jovens (cotilédones aos 10 DAS, folhas trifolioladas e raízes aos 22 DAS), maior abundância dos transcritos de *PTOX 1* foi detectada. Entretanto, nos órgãos mais velhos (cotilédones aos 22 DAS, folhas unifolioladas aos 22 DAS, folhas trifolioladas e raízes aos 45 DAS), foi observado um aumento no número de transcritos para *PTOX 2* (**Gráfico 2**).

Já em condições estressantes, o perfil de expressão para *PTOX 1* e *PTOX 2* (**Gráfico 3**) apresentou comportamento distinto. Mesmo para o tratamento com polietilenoglicol (PEG) em folhas, onde ambos os genes foram regulados positivamente, *PTOX 1* mostrou maior expressão 24 horas após a indução, apresentando um aumento na expressão de 1,16 fold (aproximadamente duas vezes) em relação ao controle. Já em raízes, neste mesmo tratamento, os dois transcritos apresentaram um padrão oposto de expressão: *PTOX 1* foi regulado positivamente 6 horas após a indução e permaneceu estável nos demais tempos, enquanto *PTOX 2* foi regulado negativamente nos tempos 12 h e 24 h, mantendo-se estável nos tempos 0 h e 6h (**Gráfico 3**).

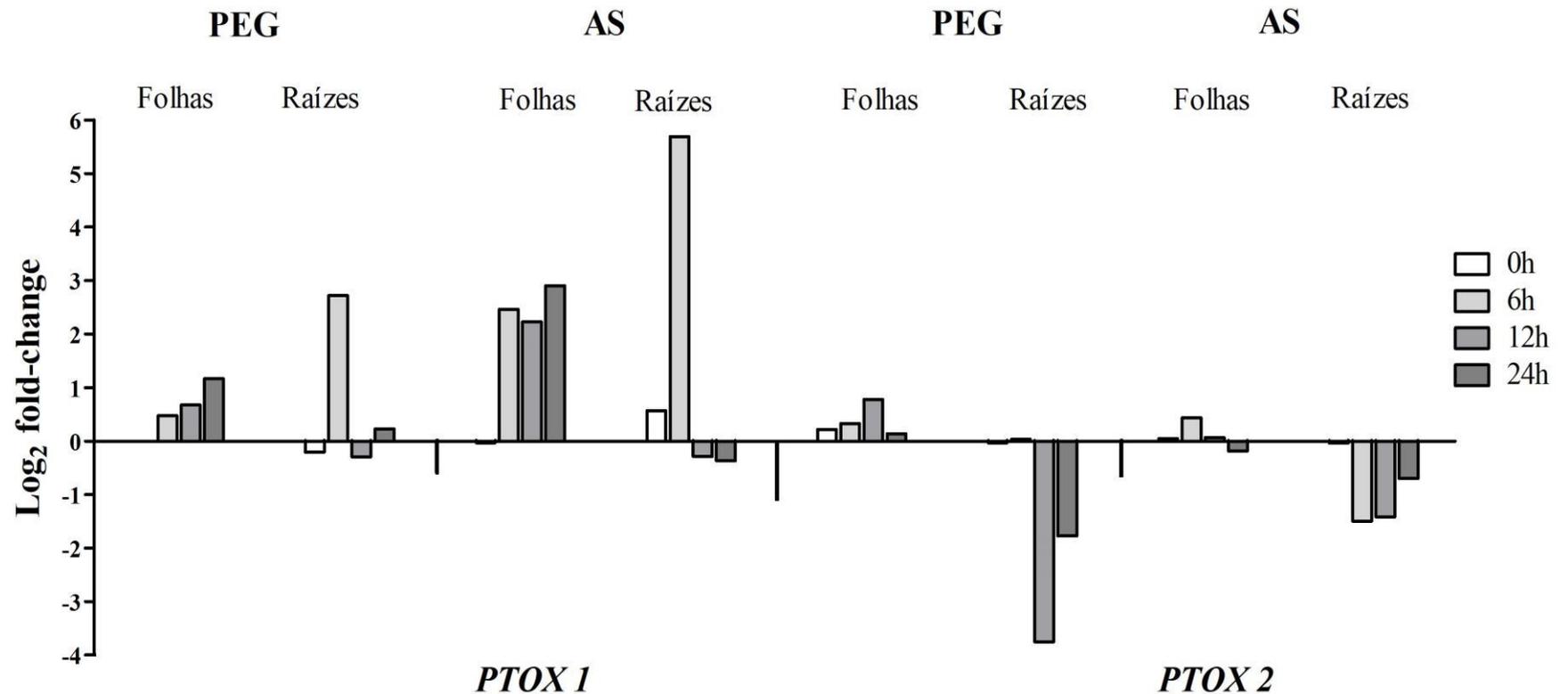
No tratamento com ácido salicílico (AS), o gene *PTOX 1* apresentou altos níveis de expressão, maior abundância de transcritos foi verificada nos tempos 6, 12 e 24 h em folhas e 6 h em raízes (5,6 fold). O gene *PTOX 2*, em contrapartida, não apresentou variação na expressão em folhas e foi regulado negativamente em raízes (**Gráfico 3**).

Gráfico 2 – Expressão relativa dos genes *PTOX* de *Glycine max* em diferentes partes da planta durante o desenvolvimento avaliada por reação em cadeia da polimerase quantitativa em tempo real (RT-qPCR).



Fonte: Autoria própria. Legenda: DAS – dias após a sementeira; DAF- dias após a floração; F.U: folhas unifolioladas; F.T: folhas trifolioladas. Os valores de expressão relativa são expressos como média \pm desvio padrão de três reações independentes a partir de um *pool* de RNA total. Os genes de referência usados para normalizar a expressão de *PTOX* em cada parte da planta são apresentados no Anexo C.

Gráfico 3 – Expressão relativa dos genes *PTOX* de *Glycine max* em folhas e raízes tratadas com polietilenoglicol (PEG) 100 g.L⁻¹ ou ácido salicílico (AS) 0,5 mM em diferentes tempos de indução avaliada por reação em cadeia da polimerase quantitativa em tempo real (RT-qPCR).



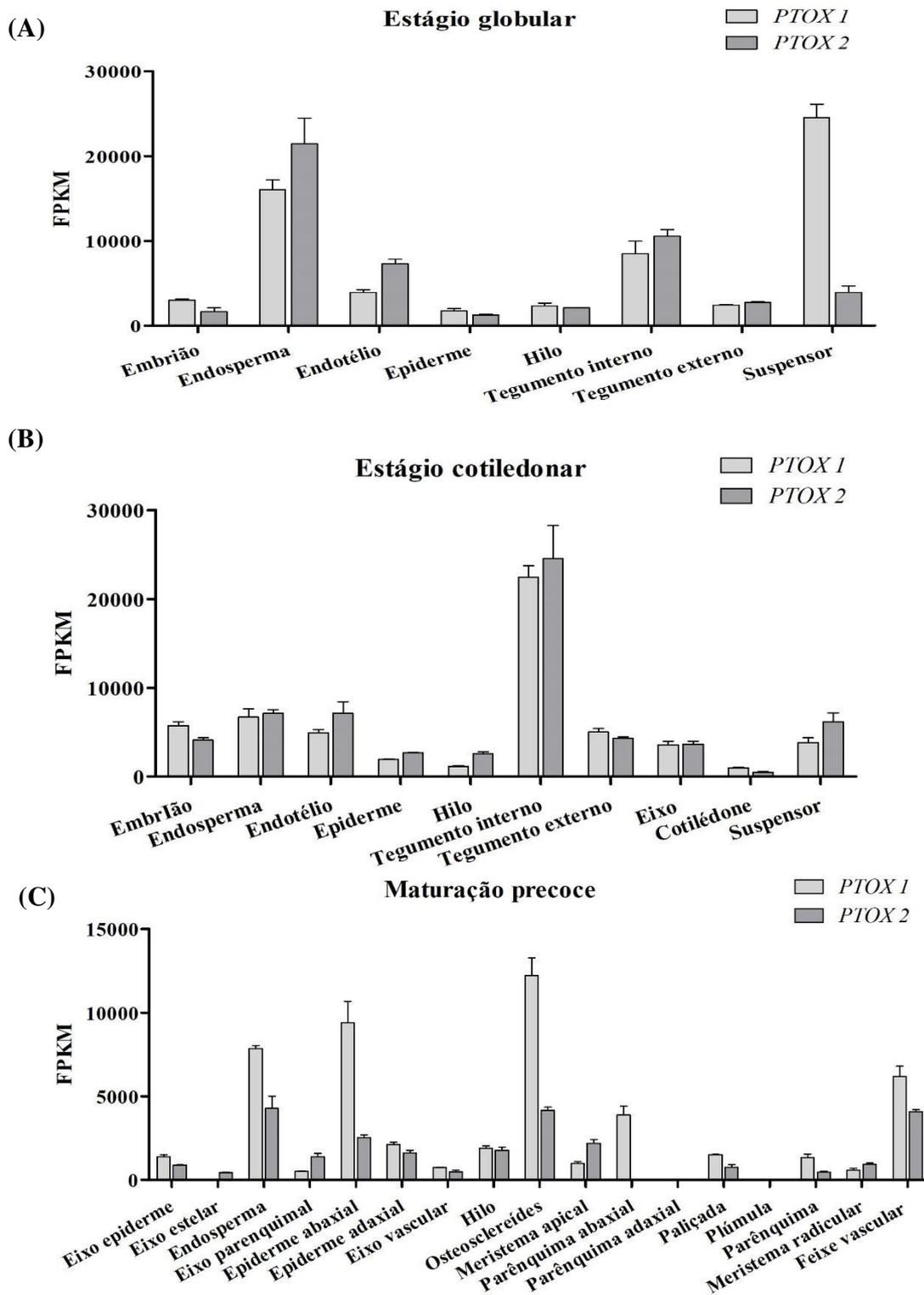
Fonte: Autoria própria. Valores de $\log_2 FC$ positivos indicam aumento da abundância do gene na condição estudada comparada ao controle, enquanto valores negativos indicam diminuição da abundância do gene na condição analisada comparada ao controle.

4.2.2 Expressão *in silico* via RNA-seq

A análise de expressão gênica da PTOX avaliada nos dados transcriptômicos (banco SRA) da soja disponíveis também revelou que os dois genes da oxidase terminal do plastídio são expressos e funcionais em diversas partes da planta e condições de tratamento em *G. max*. Esta constatação corrobora o achado pela expressão por RT-qPCR, conferindo robustez aos resultados e revela, ainda, que as sondas utilizadas no mapeamento foram específicas para cada um dos genes, sendo adequadas para a análise.

Foram avaliados experimentos que analisaram as sementes de soja nos estádios globular (PRJNA246315), cotiledonar (PRJNA246783) e maturação precoce (PRJNA197251). Observou-se que tanto *PTOX 1* quanto *PTOX 2* foram expressos em quase todos os compartimentos estudados (**Gráfico 4**). No estágio globular, a expressão dos dois genes foi maior no endosperma e no tegumento interno, enquanto a expressão de *PTOX 1* foi maior no suspensor. Já no estágio cotiledonar, praticamente não houve variação na expressão dos dois genes, sendo ambos mais expressos no tegumento interno. A semente no estágio de maturação precoce, por sua vez, apresentou alguns compartimentos em que não foi detectada expressão de um ou ambos os genes *PTOX*, tais como eixo estelar, parênquima adaxial e plúmula. Contudo, foi possível observar maior abundância de *PTOX 1* no endosperma, na epiderme abaxial e nas osteosclereídes, um tipo de célula que compõe o tegumento externo (BELTRATI; PAOLI, 2006).

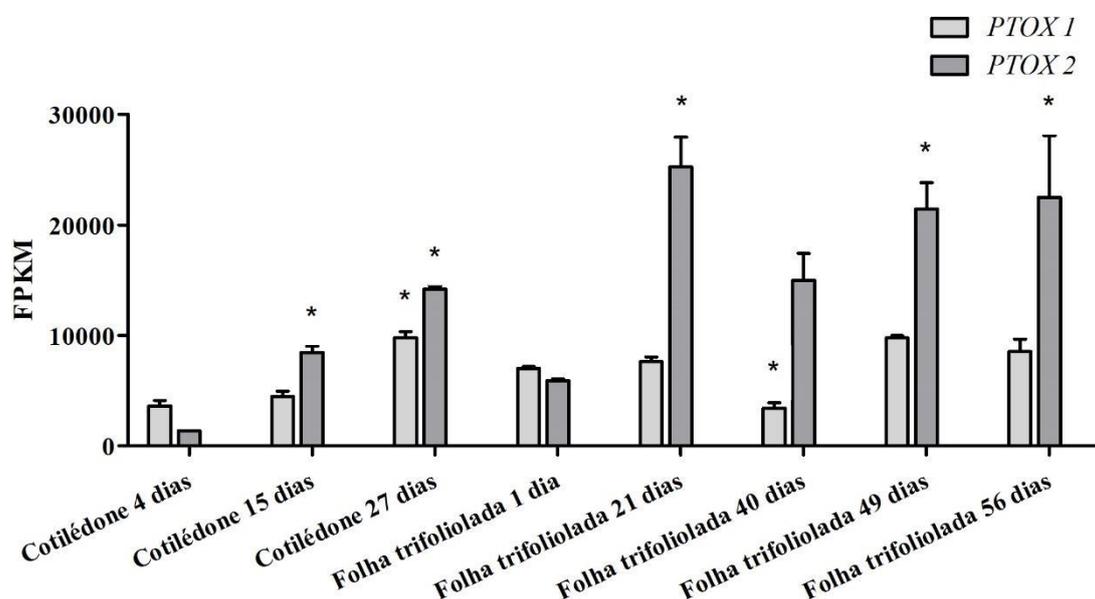
Gráfico 4 – Expressão baseada em dados de RNA-seq dos genes *PTOX 1* e *PTOX 2* em diferentes compartimentos das sementes de *Glycine max* nos estádios (A) globular, (B) cotiledonar e (C) de maturação precoce.



Fonte: Autoria própria. Valores expressos em FPKM (*Fragments Per Kilobase transcript per Million of mapped reads*) como média \pm desvio padrão das replicatas. Foram analisados oito (A), dez (B) e dezessete (C) compartimentos da semente em diferentes estádios do desenvolvimento.

Também foi analisada expressão dos genes *PTOX* em cotilédones e folhas trifolioladas em diferentes estádios de desenvolvimento, desde os órgãos recém formados até os senescentes (PRJNA262564). Além da expressão dos genes *PTOX* em todos os estádios e órgãos, também foi observado que o gene *PTOX 2* aumenta sua expressão à medida que cotilédones e folhas vão envelhecendo (**Gráfico 5**). Em cotilédones com 27 dias (estádio 3), foi observado um aumento de aproximadamente seis vezes na expressão de *PTOX 2* em relação aos cotilédones com 4 dias (estádio 1). Já em folhas trifolioladas, houve um aumento na abundância de transcritos de *PTOX 2* de quase quatro vezes em folhas com 56 dias (estádio 5) em relação às folhas mais jovens (estádio 1). A expressão de *PTOX 1* também aumentou com o envelhecimento dos cotilédones, mas de forma menos pronunciada; já em folhas, tendeu a se manter constante durante o desenvolvimento. Comportamento semelhante foi identificado para *PTOX 2* nos dados de RT-qPCR, onde houve maior expressão do gene em órgãos mais velhos (cotilédones e folhas unifolioladas com 22 dias e folhas trifolioladas com 45 dias) em comparação aos respectivos órgãos mais jovens.

Gráfico 5 – Expressão baseada em dados de RNA-seq dos genes *PTOX* de *Glycine max* durante o desenvolvimento e senescência de cotilédones e folhas trifolioladas.



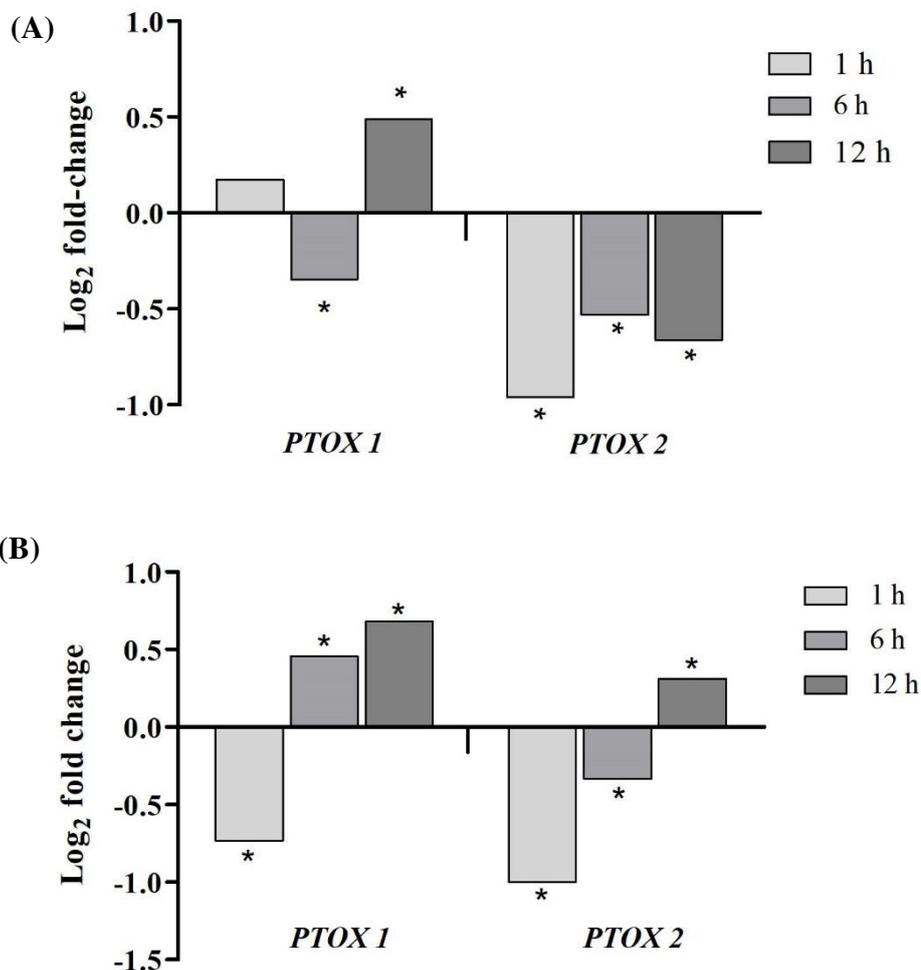
Fonte: Autoria própria. Valores expressos em FPKM (*Fragments Per Kilobase transcript per Million of mapped reads*) como média \pm desvio padrão das replicatas. Foram considerados como controle os estádios mais novos para os dois genes (cotilédones com 4 dias e folhas trifolioladas com 1 dia). A análise estatística (*one-way* ANOVA seguida do teste de Bonferroni) foi aplicada a cada gene separadamente. Diferenças significativas dos controles são indicadas por asterisco (*) com $p < 0,05$.

Também foram analisados experimentos com diferentes componentes da planta de *G. max* expostos a condições de estresse, tais como desidratação e salinidade (PRJNA246058), seca (PRJNA259941) e alta concentração de etileno (PRJNA266538). Foi possível observar uma resposta distinta para cada um dos genes nessas condições de estresse.

A exposição das raízes da soja à desidratação levou *PTOX 1* a sofrer regulação positiva no tempo 12 horas. Em contrapartida, *PTOX 2* sofreu regulação negativa de forma mais marcante após 1 hora de tratamento, tendo sua expressão diminuída em torno de duas vezes em relação ao controle, mantendo-se esse perfil de expressão nos demais tempos (**Gráfico 6A**). No experimento de PCR em tempo real, os transcritos de *PTOX 2* em raízes também foram regulados negativamente quando a planta foi exposta a estresse osmótico (tratamento com polietilenoglicol), o que dá consistência aos resultados.

Já quando a planta foi exposta a estresse salino (100 mM de NaCl) (**Gráfico 6B**), os dois transcritos da proteína PTOX sofreram uma diminuição na expressão 1 hora após o tratamento. Apenas *PTOX 1* foi regulado positivamente nos tempos seguintes (6 e 12 horas), enquanto *PTOX 2* manteve regulação negativa no tempo 6 h, passando a ser mais expresso no tempo 12 horas.

Gráfico 6 – Expressão baseada em dados de RNA-seq dos genes *PTOX* em raízes de *Glycine max* submetida a (A) desidratação e (B) estresse salino em diferentes tempos de indução.

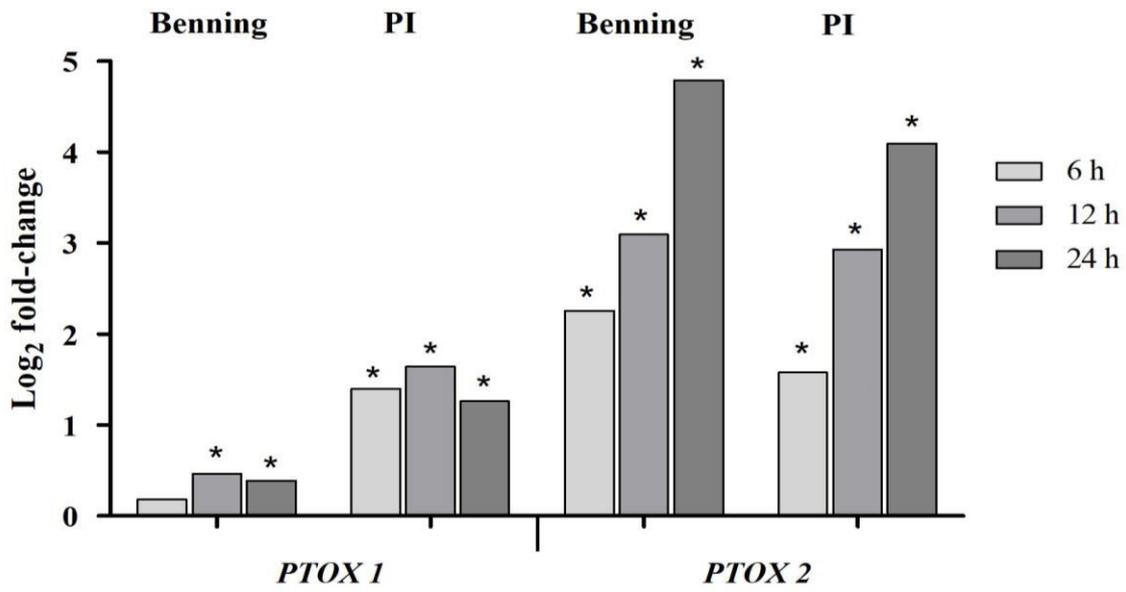


Fonte: Autoria própria. (A) Para o tratamento de desidratação, plantas no estágio V1 foram removidas do papel de germinação e deixadas no ar sob condições de limitação de água. (B) O tratamento de estresse salino foi aplicado transferindo as plantas para solução com 100 mM de NaCl. Valores de $\log_2 FC$ positivos indicam aumento da abundância do gene na condição estudada comparada ao controle, enquanto valores negativos indicam diminuição da abundância do gene na condição analisada comparada ao controle. A análise estatística (*one-way* ANOVA seguida do teste de Bonferroni) foi aplicada a cada gene separadamente. Diferenças significativas dos controles são indicadas por asterisco (*) com $p < 0,05$.

A expressão gênica da *PTOX* também foi avaliada em folhas de soja submetidas a condições drásticas de seca, sendo as plantas expostas ao ar por 6, 12 e 24 horas (PRJNA259941). Neste experimento, foram analisadas dois cultivares: Benning, que é sensível à seca e PI416937, tolerante à desidratação. Observou-se que apenas os transcritos de *PTOX 2* sofreram aumentos significativos na sua abundância, cerca de quatro, seis e dez vezes para os tempos 6, 12 e 24 horas, respectivamente, para este mesmo cultivar (Gráfico 7). Já para o cultivar tolerante à seca, ambos os transcritos *PTOX* foram diferencialmente expressos em todos os tempos testados. *PTOX 1* teve sua expressão aumentada mais de duas vezes em todos os tempos em relação ao controle e *PTOX 2* sofreu aumentos graduais de expressão ao

longo do tempo: 1,57 fold no tempo 1 hora, 2,92 fold no tempo 6 horas e 4,09 fold 24 horas após o tratamento.

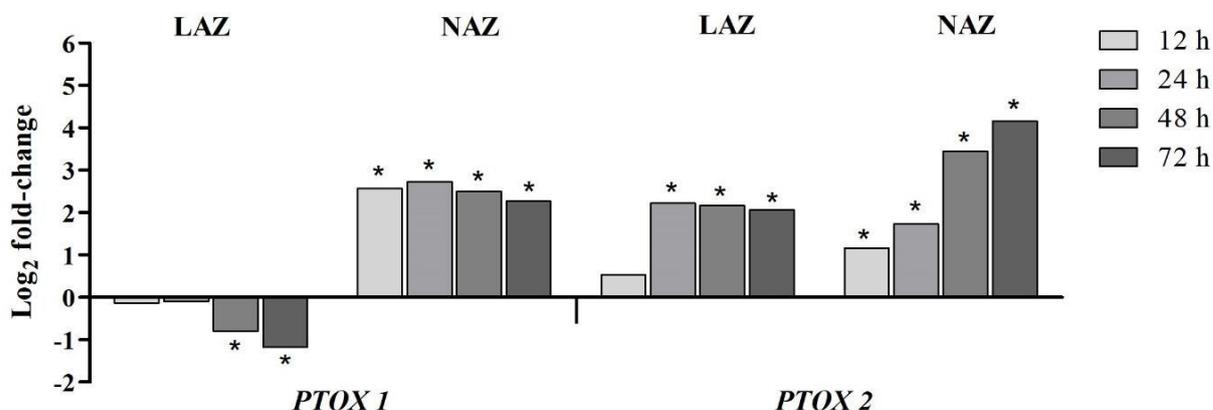
Gráfico 7 – Expressão baseada em dados de RNA-seq dos genes *PTOX* em folhas trifolioladas de dois cultivares de *Glycine max* com susceptibilidade e tolerância à seca submetidos a estresse por desidratação em diferentes tempos de indução.



Fonte: Autoria própria. Cultivar Benning é susceptível à seca e PI 416937, tolerante. Para o tratamento de desidratação, as plantas no estágio R2 foram removidas dos vasos e expostas ao ar por 0 (controle), 6, 12 e 24 horas. Valores de $\log_2 FC$ positivos indicam aumento da abundância do gene na condição estudada comparada ao controle, enquanto valores negativos indicam diminuição da abundância do gene na condição analisada comparada ao controle. A análise estatística (*one-way* ANOVA seguida do teste de Bonferroni) foi aplicada a cada gene separadamente. Diferenças significativas dos controles são indicadas por asterisco (*) com $p < 0,05$.

O experimento com explantes de folhas (zona de abscisão foliar) e pecíolo (zona de não abscisão foliar) do cultivar Williams 82 de soja expostos a 25 $\mu\text{L/L}$ do hormônio vegetal etileno (PRJNA266538) revelou perfil de expressão de *PTOX 1* e *PTOX 2* oposto na zona de abscisão foliar (**Gráfico 8**). Enquanto *PTOX 1* manteve níveis basais nos primeiros dois tempos testados, sendo regulada negativamente após 48 e 72 horas do tratamento, *PTOX 2* sofreu aumento de expressão de cerca de quatro vezes em relação ao controle, nos tempos de 24 h em diante. Já em pecíolos, ambos os transcritos foram regulados positivamente e diferencialmente expressos para todos os tempos testados; contudo, *PTOX 1* manteve os níveis de expressão constantes ao longo do tempo, enquanto *PTOX 2* aumentou gradualmente sua expressão ao longo do tempo, atingindo o máximo de 4,15 fold no tempo de 72 horas.

Gráfico 8 – Expressão baseada em dados de RNA-seq dos genes *PTOX* em zonas de abscisão foliar (LAZ) e de não abscisão (NAZ) de *Glycine max* expostos a alta concentração de etileno em diferentes tempos de indução.



Fonte: Autoria própria. Legenda: LAZ - zona de abscisão foliar (junção da folha com o pecíolo); NAZ - zona de não abscisão (parte intermediária do pecíolo). O tratamento consistiu na exposição de explantes excisados da planta 25 μ L/L de etileno por 0 (controle), 12, 24, 48 e 72 horas. Valores de $\log_2 FC$ positivos indicam aumento da abundância do gene na condição estudada comparada ao controle, enquanto valores negativos indicam diminuição da abundância do gene na condição analisada comparada ao controle. A análise estatística (*one-way* ANOVA seguida do teste de Bonferroni) foi aplicada a cada gene separadamente. Diferenças significativas dos controles são indicadas por asterisco (*) com $p < 0,05$.

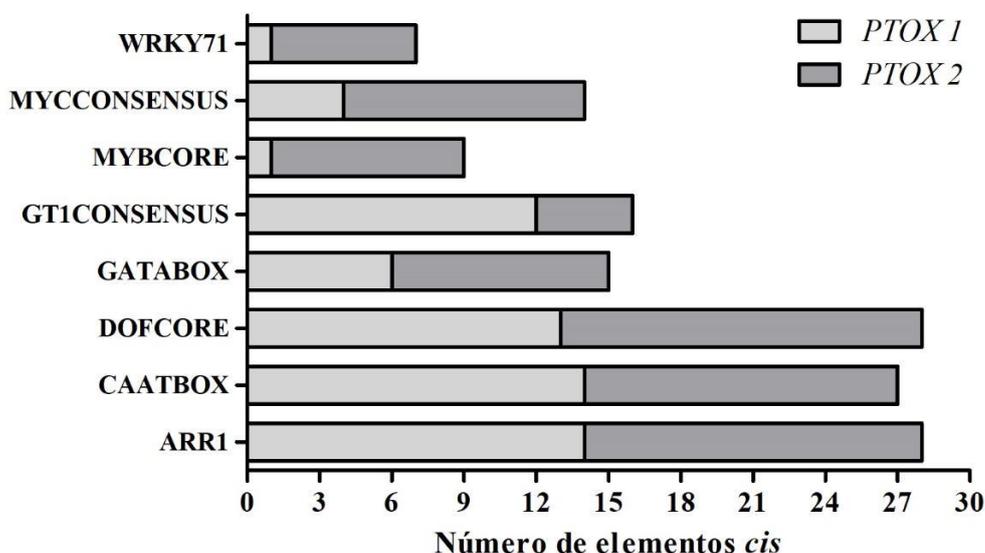
4.2.3 Análise da região promotora

A análise dos promotores dos genes *PTOX* de *G. max* revelou a presença de alguns elementos *cis* regulatórios comuns aos dois genes e de outros característicos de cada gene (Gráfico 9). No primeiro caso, destacam-se os sítios: ARR1 [(A/T/C/G)GATT], aos quais se ligam ativadores transcricionais (SAKAI; AOYAMA; OKA, 2008); CAATBOX1 (CAAT), um elemento *cis* comum em promotores e regiões potenciadoras (SHIRSAT *et al.*, 1989); DOFCORE (AAAG), um sítio requerido para a ligação de proteínas Dof, que desempenham diversos papéis na célula (YANAGISAWA; SCHMIDT, 1999) e GATABOX (GATA), uma região à qual se liga o fator de transcrição ASF-, requerido para expressão induzida pela luz (TEAKLE *et al.*, 2002).

Já o elemento GT1CONSENSUS [G(A/G)(A/T)AA(A/T)], um sítio também presente em genes regulados pela luz (VILLAIN; MACHE; ZHOU, 1996), foi mais encontrado no promotor de *PTOX 1*. Os elementos mais presentes em *PTOX 2*, por sua vez, foram MYCCONSENSUS [CA(A/T/C/G)(A/T/C/G)TG] e MYBCORE [C(A/T/C/G)GTT(A/G)], que são encontrados em promotores do gene *rd22* responsivo à desidratação e em vários outros genes em *Arabidopsis* (ZHU, 2002). Os elementos MYB também são, juntamente com outras famílias de fatores de transcrição como WRKY, relatados

como pertencentes às principais famílias relacionadas com a senescência (BALAZADEH; RIAÑO-PACHÓN; MUELLER-ROEBER, 2008; BROWN; HUDSON, 2015). De fato, também foram detectados no promotor de *PTOX 2* vários sítios de ligação WRKY71 (TGAC), um elemento que, em *Oryza sativa*, apresentou-se responsivo à senescência das folhas (LIU *et al.*, 2010).

Gráfico 9 – Elementos *cis* regulatórios presentes nos promotores dos genes *PTOX 1* e *PTOX 2* de *Glycine max*.



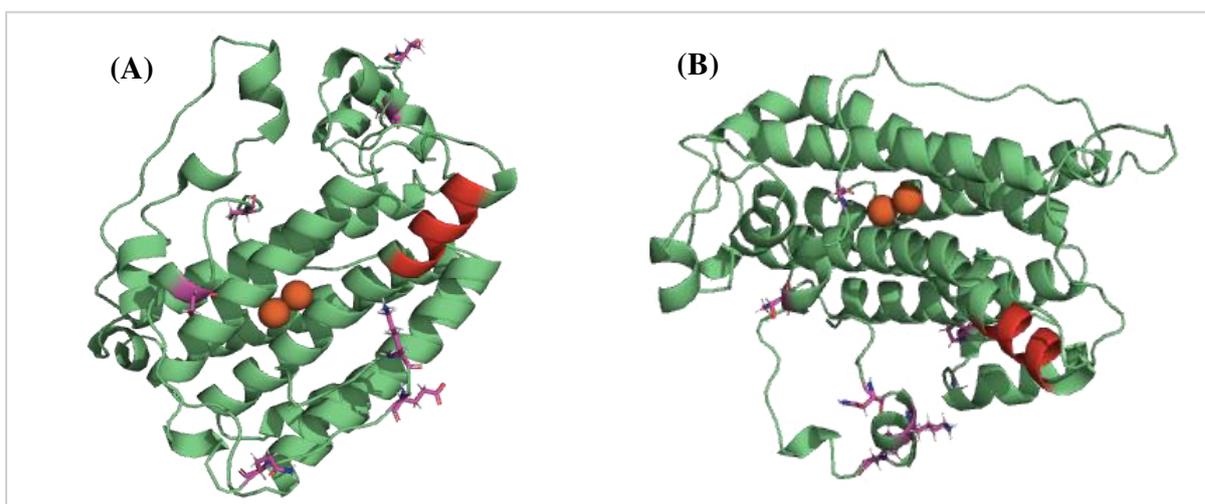
Fonte: Autoria própria.

4.3 Modelagem da estrutura tridimensional das proteínas PTOX de *Glycine max* e interação com o plastoquinol

A estrutura tridimensional das duas proteínas PTOX da soja foi predita por dois servidores, Modeller e Robetta. Dos duzentos modelos gerados pelo Modeller para cada uma das proteínas, os que apresentaram menor função objetiva, ou seja, os mais otimizados geometricamente, foram os escolhidos (modelo 102 para ambas). Já dentre os cinco modelos fornecidos pelo Robetta, foram selecionados aqueles com menor valor de RMSD (modelo 1 para ambas). Para as duas proteínas, as estruturas dadas pelo Modeller abrangeram os aminoácidos nas posições 74 a 243; já para o Robetta, as duas estruturas abrangeram todos os aminoácidos da sequência. A validação das estruturas pelos servidores SAVES e MolProbitry forneceram o *Z-score* e o gráfico de Ramachandran. O *Z-score* para os modelos do Modeller foi -0,365 para as duas estruturas, já os valores para os modelos do Robetta foram 0,113 para

PTOX 1 e 0,157 para PTOX 2. A porcentagem de resíduos de aminoácidos em posições favoráveis dada pelo gráfico de Ramachandran foi de 94 % para PTOX 1 e PTOX 2 para os modelos gerados pelo Modeller e de 96,3 % e 94,9 % para PTOX 1 e PTOX 2, respectivamente, para os modelos gerados pelo Robetta (**Anexo D**). Tendo em vista os resultados obtidos na validação da estrutura tridimensional (*Z-score* mais próximo de zero e maior porcentagem de aminoácidos em posições favoráveis), os modelos para as duas proteínas PTOX da soja gerados pelo servidor Robetta foram os escolhidos para dar prosseguimento à análise de docagem molecular (**Figura 11**).

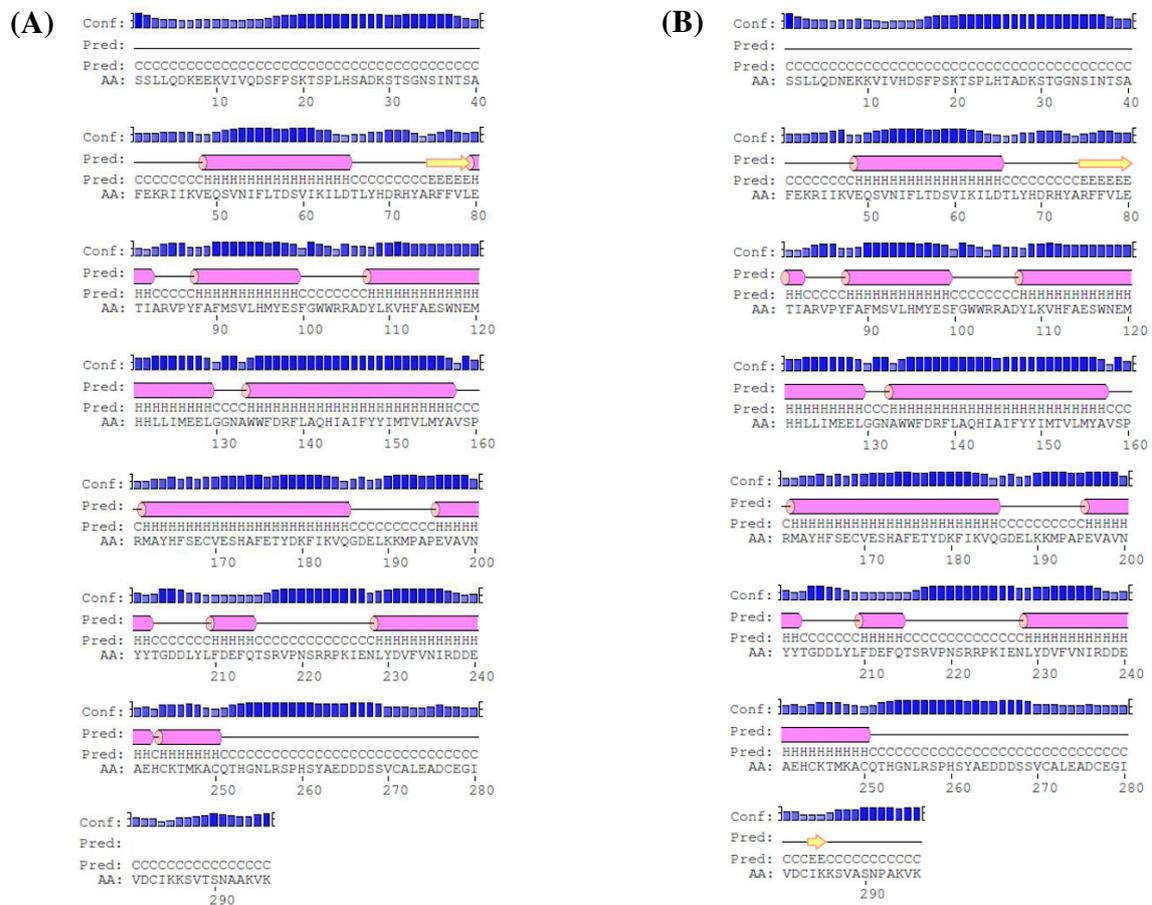
Figura 11 – Modelos tridimensionais preditos pelo servidor Robetta das proteínas PTOX 1 (A) e PTOX 2 (B) de *Glycine max* utilizados para o *docking* molecular.



Fonte: Autoria própria. A visualização das estruturas foi feita no programa PyMOL v. 2.0.1. As esferas laranjas representam os íons ferro necessários para a atividade da enzima. As regiões em vermelho correspondem ao centro de reação (MSVLHMYES) e em rosa (bastões), os aminoácidos alterados entre as duas estruturas da PTOX. As regiões desordenadas são representadas por linhas.

A predição da estrutura secundária das proteínas feita pela ferramenta PSIPRED v3.3 (**Figura 12**) corroborou com a localização de α -hélices na visualização das estruturas tridimensionais, apesar de ter feito a previsão equivocada de uma folha β em uma curta região nas duas proteínas; entretanto, o próprio programa avaliou esta predição (folha β) como não muito confiável. Para as duas sequências proteicas, as estruturas tridimensionais apresentaram seis cadeias principais de α -hélice, o que foi condizente com a estrutura modelada da PTOX 2 de *Chamydomonas reinhardtii* apresentada por Nawrocki *et al.* (2015).

Figura 12 – Predição da estrutura secundária de PTOX 1 (A) e PTOX 2 (B) de *Glycine max* feita pela ferramenta PSIPRED.



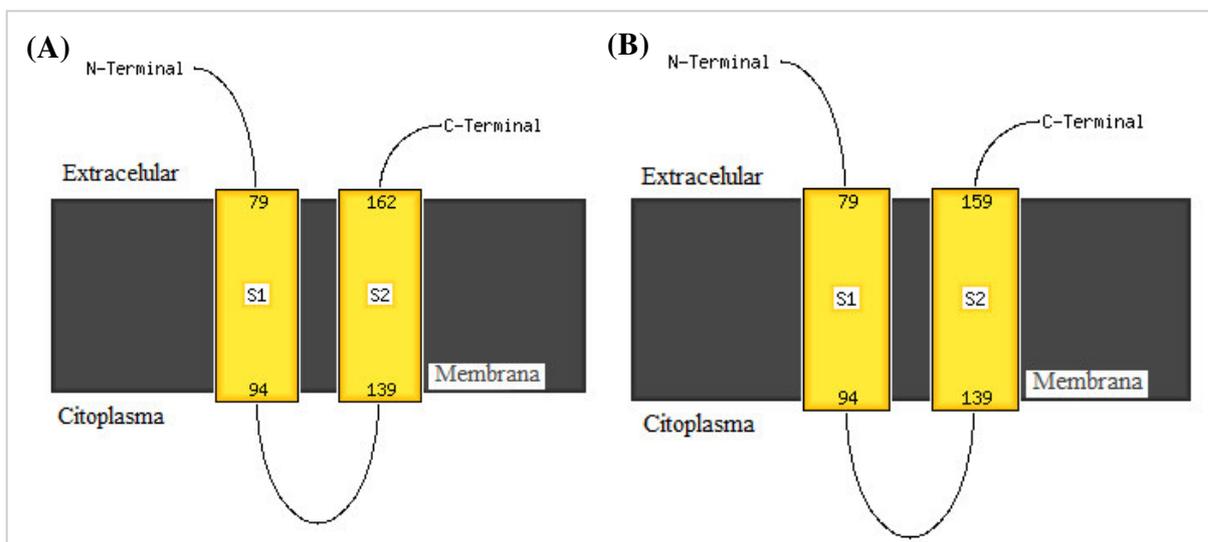
Fonte: PSIPRED. Legenda: caixas rosa – α -hélice; setas amarelas – folhas β ; Linhas – regiões de desordem; Conf – confiança da predição; Pred – estrutura secundária predita; AA – sequência alvo. Para as duas proteínas, foram retiradas a sequência do peptídeo trânsito (primeiros 36 aminoácidos de cada sequência).

A predição de estruturas desordenadas feita pelo DISOPRED3 revelou apenas regiões nas extremidades N- e C-terminais das proteínas: aminoácidos 1 - 9 e 287 - 296 para PTOX 1 e aminoácidos 1, 2, 5 e 289 - 296 para PTOX 2, enquanto na estrutura tridimensional é possível observar um maior número de regiões de desordem (regiões representadas por linhas na Figura 11).

A ferramenta MEMSAT-SVM propôs que duas α -hélices seriam transmembranares para as duas proteínas, sendo em PTOX 1 as regiões compreendidas pelos aminoácidos 79 a 94 e 139 a 162 e em PTOX 2 as regiões abrangidas pelos aminoácidos 79 a 94 e 139 a 159 (Figura 13). A ferramenta TMPred também previu duas α -hélices transmembranares para as duas proteínas em regiões muito próximas, quando não as mesmas, daquelas fornecidas pelo MEMSAT-SVM: aminoácidos 76 a 103 e 139 a 159 para ambas as sequências. Nawrocki *et al.* (2015) e Castro *et al.* (2017) também encontraram duas hélices

transmembranares para as estruturas previstas da PTOX de *C. reinhardtii* e AOX de *Citrus clementina*.

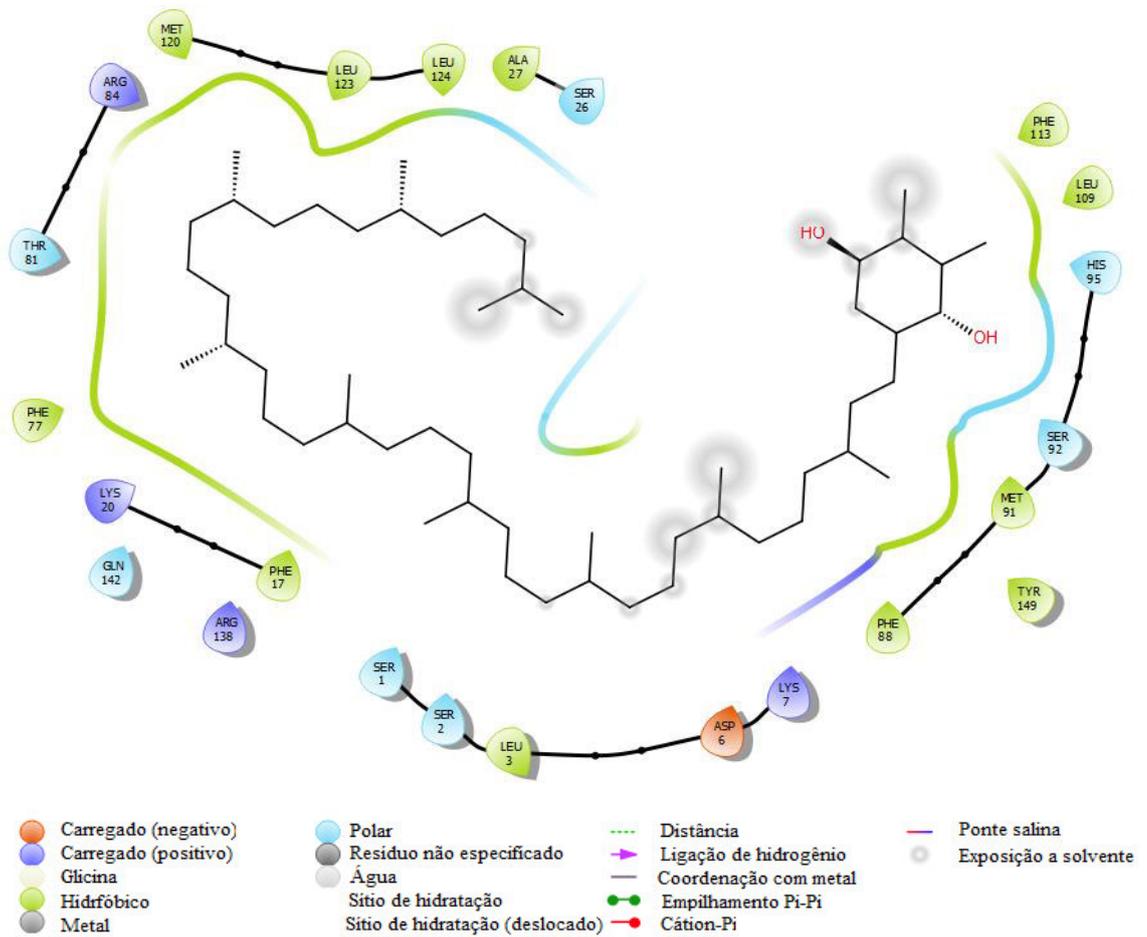
Figura 13 – Predição de regiões transmembranares em (A) PTOX 1 e (B) PTOX 2 de *Glycine max* feita pela ferramenta MEMSAT-SVM do servidor PSIPRED.



Fonte: MEMSAT-SVM. (A) perfil de regiões transmembranares para a sequência de aminoácidos de PTOX 1, proposto com duas α -hélices cruzando a membrana. (B) perfil de regiões transmembranares para a sequência de aminoácidos de PTOX 2, proposto com duas α -hélices cruzando a membrana. Para as duas proteínas, foram retiradas a sequência do peptídeo trânsito (primeiros 36 aminoácidos de cada sequência).

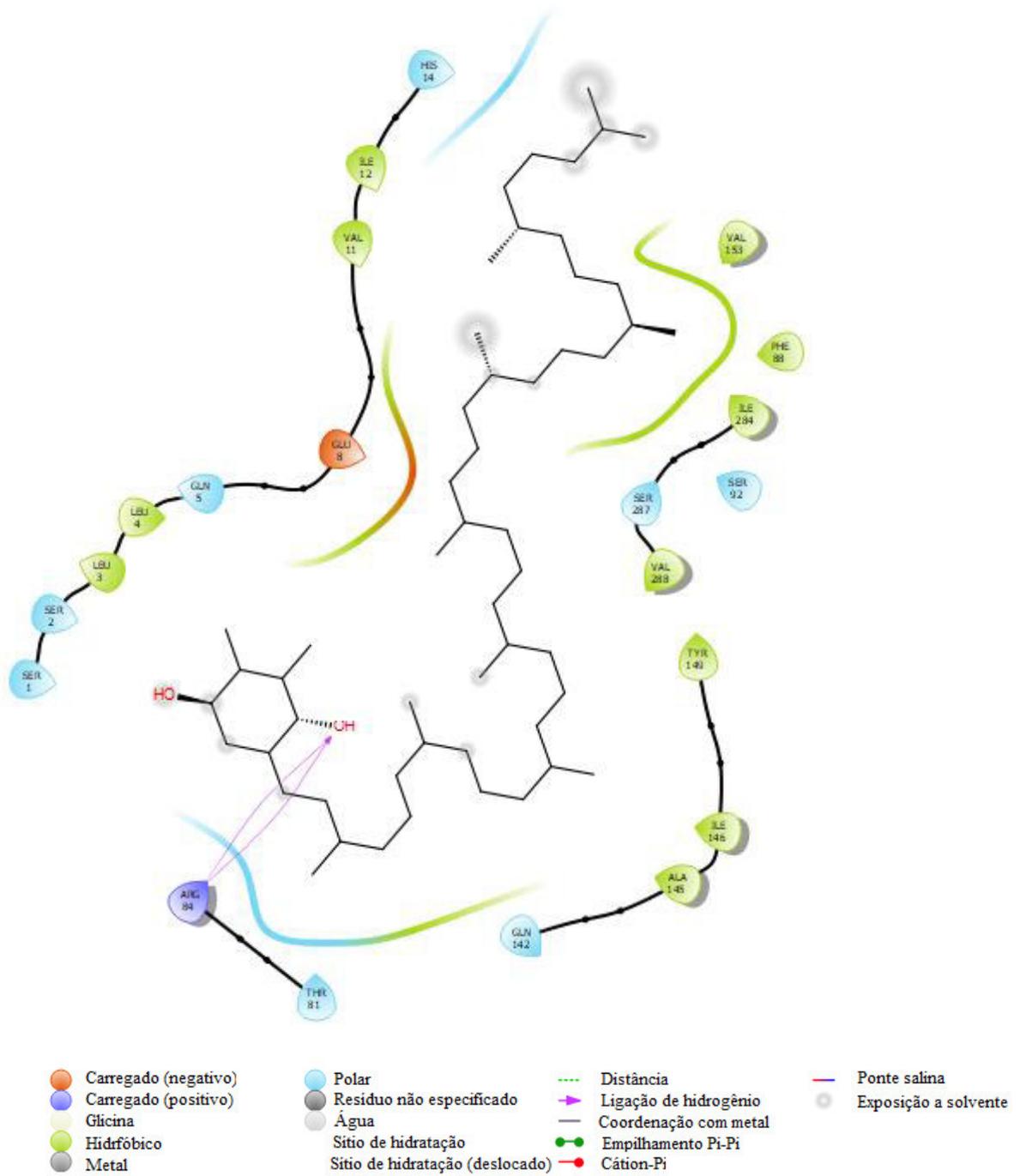
Os resultados do *docking* em energia de interação proteína-ligante foram -7,6 kcal/mol para o complexo formado por PTOX 1 e -6,5 para PTOX 2, ambos com RMSD de zero, revelando que a afinidade de PTOX 1 pelo substrato é maior que a de PTOX 2. Os diagramas de interação fornecidos pelo programa Maestro são apresentados na **Figuras 14 e 15**. O diagrama para o complexo da PTOX 1 revela a participação da histidina na posição 95 (H95), que é essencial para a atividade da enzima. Já no complexo formado pela PTOX 2 pode-se observar a formação de uma ligação de hidrogênio entre o aminoácido arginina (R84) e uma das hidroxilas do ligante. É possível observar, ainda, que a conformação da molécula de plastoquinol é diferente em cada um dos complexos formados e que dois resíduos de aminoácidos que estão presentes em PTOX 1 (K7 e S26), mas não em PTOX 2 interagem com o substrato. De forma análoga, o resíduo H14, somente presente em PTOX 2, interage com o ligante.

Figura 14 – Diagrama de interação 2D proteína-ligante referente ao complexo PTOX1-PQH₂ (plastoquinol).



Fonte: Diagrama gerado pelo programa Maestro.

Figura 15 – Diagrama de interação 2D proteína-ligante referente ao complexo PTOX2-PQH₂ (plastoquinol).



Fonte: Diagrama gerado pelo programa Maestro.

5 DISCUSSÃO

A ocorrência de dois genes (*PTOX1* e *PTOX 2*) codificando a oxidase terminal do plastídio em *Glycine max* foi anteriormente relatada por Sun *et al.* (2015), entretanto, não se tem informação sobre a funcionalidade de ambos os genes nas respostas fisiológicas da planta. O presente trabalho, portanto, não só corroborou o achado como também comprovou que ambos os genes são expressos, apresentando expressão diferencial em vários estádios de desenvolvimento da planta, bem como em resposta a diferentes condições de estresse. Além disso, investigou, por meio de ferramentas computacionais, a interação das duas isoformas da proteína com seu ligante.

5.1 Os dois genes *PTOX* da soja são funcionais e expressos na maioria das partes da planta

As sequências do DNA complementar e da proteína para os dois genes compartilham 97 % de identidade, sendo altamente conservadas. Isso pode ser constatado pelo alinhamento entre as sequências de aminoácidos das PTOXs e das oxidases alternativas (AOXs) da soja (Figura 7), visto que, por serem enzimas evolutivamente relacionadas, conservam todas as regiões necessárias para a atividade enzimática, como o centro de reação e os motivos de ligação aos íons ferro, característico de proteínas diferido carboxilato (BERTHOLD; STENMARK, 2003). Tal fato também atesta para a funcionalidade dos dois genes da PTOX em *G. max*.

A investigação dos genes da PTOX em outras espécies da ordem Fabales também revelou um elevado grau de conservação das sequências gênicas e proteicas entre espécies, sendo presente, em todas elas, o mesmo número de éxons e íntrons e os motivos requeridos para a atividade da proteína. Entretanto, em nenhuma outra leguminosa avaliada foram encontradas duas cópias da PTOX e a análise filogenética (Figura 8), demonstrou que a presença dos parálogos em *G. max* é consequência de um evento de duplicação recente, ao invés da deleção da segunda cópia em outras espécies leguminosas. Sabe-se que o genoma da soja é altamente duplicado, apresentando múltiplas cópias de aproximadamente 75 % dos genes (SCHMUTZ *et al.*, 2010). Saraiva *et al.* (2016) também identificaram o dobro do número de cópias do gene do fator de alongação 1 alfa (*EF1 α*) em soja, quando comparada com as espécies *Phaseolus vulgaris* e *Vigna radiata*, ambas pertencentes à sua mesma tribo (Phaseoleae). Entretanto, apesar da alta identidade entre os genes duplicados da PTOX na soja, eles encontram-se em cromossomos distintos (*PTOX 1* no cromossomo 9 e *PTOX 2* no

cromossomo 15), indicando que os eventos de duplicação foram seguidos de rearranjos gênicos. De fato, é sabido que as duplicações ocorridas no genoma de *G. max* foram seguidas de numerosos rearranjos cromossômicos (SCHMUTZ *et al.*, 2010).

A PTOX ocorre em procariotos e eucariotos, sendo que a maioria das cianobactérias e plantas apresentam apenas uma cópia do gene. Entretanto, são encontradas espécies que contêm duas cópias do gene entre diatomáceas, algas verdes, vermelhas e pardas e também entre representantes dos filos Cryptophyta e Haptophyta; já o gênero de musgo *Physcomitrella* contém três cópias (NAWROCKI *et al.*, 2015). Recentemente, Hossain *et al.* (2017) também identificaram a presença de dois genes da PTOX na beterraba (*Beta vulgaris*), uma dicotiledônea assim como a soja.

Uma análise filogenética utilizando mais de cem sequências de aminoácidos da PTOX feita por Nawrocki *et al.* (2015) revelou que as múltiplas cópias da proteína em um único organismo, na maioria dos casos, não são decorrentes de eventos de duplicação ocorridos nas espécies, mas possivelmente de transferência horizontal de genes, visto que as cópias individuais da PTOX em certas espécies estão pouco relacionadas. Esta mesma análise, contudo, também identificou algumas espécies em que as duas cópias da oxidase estavam mais relacionadas, como o milho e o eucalipto. A duplicação do gene em *Glycine max*, entretanto, foi relatada pelos autores como um falso positivo decorrente de artefatos durante a montagem do genoma. Entretanto, o presente trabalho, destaca a presença de dois genes parálogos da PTOX em soja, bem como suas funcionalidades, comprovada experimentalmente por RT-qPCR e dados transcriptômicos (RNA-seq).

Os eventos de duplicação de genes podem resultar em ganho de função, perda de função ou redundância (WANG *et al.*, 2011). Os resultados de expressão gênica obtidos neste trabalho sugerem que a duplicação do gene da PTOX em *G. max* conferiu um ganho de função à planta, tendo em vista que os genes parecem ter expressão preferencial em determinados estádios do ciclo celular, tecido, órgão ou condição de estresse.

Transcritos de *PTOX 1* e *PTOX 2* foram detectados em todos os órgãos e idades avaliados durante o desenvolvimento da planta (sementes, cotilédones, folhas, raízes, flores, vagens, epicótilo e hipocótilo), o que pode ser evidenciado tanto nos experimentos de RT-qPCR quanto nos dados transcriptômicos de RNA-seq. Ou seja, ambos os genes foram expressos mesmo quando a planta não sofreu com a aplicação de algum estresse, condição em que é sabido haver alteração na abundância dos transcritos da PTOX (NAWROCKI *et al.*, 2015). Esses resultados atestam para o caráter ubíquo da proteína, mesmo sendo esta relatada como um componente minoritário nas células e apresentando uma razão PTOX/PSII de 1:100

em espinafre (LENNON; PROMMEENATE; NIXON, 2003). Todavia, seus transcritos foram expressos em praticamente todos os tecidos e órgãos de *Arabidopsis thaliana* (ALURU *et al.*, 2001) e níveis da proteína foram detectados em folhas de *A. thaliana* e *Pisum sativum* durante todo o período de desenvolvimento avaliado (LENNON; PROMMEENATE; NIXON, 2003).

O estudo da expressão dos genes *PTOX* em diferentes estádios do desenvolvimento da semente (globular, cotiledonar e maturação precoce) avaliada em dados de RNA-seq (Gráfico 4) também revelou a presença dos transcritos em praticamente todos os compartimentos avaliados, mas em maior abundância no endosperma, suspensor, tegumento interno e células osteosclereides. Todos estes compartimentos estão, direta ou indiretamente, relacionados com a nutrição do embrião em desenvolvimento (GEHRING; CHOI; FISCHER, 2004; MOISE *et al.*, 2005; YEUNG; MEINKE, 1993).

Sabe-se que a fotossíntese é o processo básico que provê energia para a planta através da produção de ATP e NADPH que são utilizados nas reações de fixação do CO₂ para formação de carboidratos (NIYOGI; WOLOSUK; MALKIN, 2015). Apesar da *PTOX* não ser uma proteína essencial para que a fotossíntese ocorra, sua atividade está intimamente relacionada com esse processo, visto que é bastante relatada sua atuação como uma válvula alternativa a nível do PSII, evitando a redução excessiva do *pool* de plastoquinona, um dos carreadores da cadeia de transporte de elétrons fotossintética (CTE) (CAROL; KUNTZ, 2001; RUMEAU; PELTIER; COURNAC, 2007; NAWROCKI *et al.*, 2015). Além disso, também é relatado seu envolvimento na biossíntese de carotenoides e clorofila, dois pigmentos fotossintéticos (CAROL; KUNTZ, 2001; STECCANELLA; HANSSON; JENSEN, 2015). O perfil transcricional de *Arabidopsis* durante o desenvolvimento do embrião revelou um considerável aumento na expressão de genes que codificam componentes envolvidos na fotossíntese ainda no estágio globular da semente (SPENCER *et al.*, 2007). Também foi comprovado que embriões de *Brassica* foram capazes de fixar CO₂, contribuindo para o seu crescimento e aumento de biomassa (GOFFMAN *et al.*, 2005). Ademais, Borisjuk *et al.* (2005) demonstraram que uma alta taxa de transporte de elétrons fotossintético ocorre nos estádios iniciais da maturação em sementes de *Glycine max*.

5.2 O gene *PTOX 2* e sua relação com a senescência

A expressão dos genes *PTOX* durante o desenvolvimento e senescência de cotilédones e folhas de *G. max* apresentou um padrão interessante identificado tanto na expressão por PCR em tempo real quanto nos dados de RNA-seq. Verificou-se um aumento

na abundância de *PTOX 2* à medida que a planta envelhecia, como visto nos Gráficos 2 e 5. Cotilédones com 22 e 27 dias, ou seja, em processo de senescência, apresentaram níveis consideráveis de *PTOX 2* quando comparados aos seus estádios mais jovens (10 e 4 dias, respectivamente). Ocorrência semelhante também é observada no padrão de expressão de *PTOX 2* em folhas unifolioladas (Gráfico 2) e trifolioladas (Gráficos 2 e 5) com mais dias de vida.

Corroborando com esses resultados, foram encontrados na região promotora de *PTOX 2* vários sítios de ligação aos fatores de transcrição pertencentes às famílias MYB e WRKY (Gráfico 9), que estão entre os principais grupos relacionados com a senescência (BALAZADEH; RIAÑO-PACHÓN; MUELLER-ROEBER, 2008; BROWN; HUDSON, 2015). Uma das funções que se tem atribuído à *PTOX* é a sua participação na biossíntese de carotenoides, uma vez que a plastoquinona oxidada, produto da sua reação, é requerida por duas importantes enzimas dessa via, a dessaturase do fitoeno e a dessaturase do caroteno (JOHNSON; STEPIEN, 2016; CAROL; KUNTZ, 2001; MCDONALD *et al.*, 2011, NAWROCKI *et al.*, 2015). Também é relatado que em alguns casos a produção de carotenoides pode ser aumentada durante a senescência (THOMAS *et al.*, 2015).

5.3 *PTOX 1* e *PTOX 2* têm respostas distintas de acordo com a severidade do estresse

Muitos relatos na literatura também apontam para a *PTOX* como uma enzima responsiva a vários tipos de estresses, como seca, salinidade, altas temperaturas e alta luminosidade (SUN; WEN, 2011). De fato, a análise das regiões promotoras dos dois genes confirma que a *PTOX* está envolvida nestes processos, uma vez que há um considerável número de elementos *cis* responsivos à luz (GT1) e ao estresse hídrico (MYB) nos promotores de *PTOX 1* e *PTOX 2* (Gráfico 9). De acordo com os resultados obtidos, os genes *PTOX* da soja também responderam aos estresses impostos, havendo casos de aumento na expressão de até quase dez vezes em relação ao tratamento controle, como visto no Gráfico 7.

As condições estressantes estudadas neste trabalho foram estresse osmótico (PEG), desidratação ou seca, estresse salino e altas concentrações de fitormônios (ácido salicílico e etileno). O perfil de expressão de *PTOX 1* e *PTOX 2* apresentou um padrão distinto para cada tipo de tratamento, parecendo ser, na maioria das vezes, dependente do rigor do estresse aplicado. Foi possível observar que nas condições em que foram mantidos todos os quesitos básicos para a sobrevivência da planta, apesar das altas concentrações dos agentes estressantes, os níveis de *PTOX 1*, de forma geral, tenderam a aumentar e os de *PTOX*

2 a diminuir (Gráficos 3 e 6). Em contrapartida, quando um estresse muito severo foi aplicado, a ponto de ameaçar a vida da planta, os níveis de transcritos dos dois genes sofreram aumento, de maneira mais pronunciada para *PTOX 2* (Gráficos 7 e 8).

Nos estresses induzidos por PEG (RT-qPCR, Gráfico 3), desidratação e NaCl (RNA-seq, Gráfico 6), os genes *PTOX* apresentaram perfis de expressão inverso entre eles e semelhante entre tratamentos. *PTOX 1* foi regulado positivamente nas folhas das plantas tratadas com PEG e nas raízes das plantas submetidas a desidratação e NaCl no tempo 12 horas. Já no tempo 6 horas, este mesmo gene sofreu aumento de expressão em raízes para os tratamentos com PEG e NaCl. Com relação à expressão de *PTOX 2*, houve regulação negativa nas raízes das plantas tratadas com PEG e sob desidratação no tempo 12 h, ao passo que, no tempo 6 h, a diminuição da expressão do gene ocorreu nas raízes das plantas submetidas a desidratação e alta concentração de NaCl. Os estresses (PEG, desidratação e NaCl), aqui analisados, causam respostas fisiológicas similares nos momentos iniciais da exposição, que é a restrição da absorção de água pela planta (ZHU, 2001). Assim, o perfil de expressão inverso dos genes *PTOX 1* e *2* similar entre os tratamentos, obtidos com técnicas distintas, corroboram para a confiabilidade dos resultados obtidos.

A atividade da *PTOX* foi aumentada em folhas de *Hibiscus* mantidas em estresse hídrico e incubadas com alta incidência luminosa (MUÑOZ; QUILES, 2013). Resultado semelhante foi encontrado para *Rosa meillandina*, que teve a via da clororespiração (da qual a *PTOX* participa) estimulada por seca e sob altas temperaturas e luminosidade (PAREDES; QUILES 2013). O estresse hídrico também regulou positivamente o gene *PTOX* em *G. max* após 6 horas da aplicação do tratamento (SUN *et al.*, 2014); mesmo a espécie analisada neste caso sendo a soja, a expressão verificada foi em relação aos transcritos totais, não havendo distinção entre *PTOX 1* e *PTOX 2*. Já em *Beta vulgaris*, que apresenta duas cópias da *PTOX*, a expressão dos dois transcritos foi avaliada em folhas frente a estresse salino após 3 horas e 14 dias de tratamento. Ambos os genes foram regulados positivamente em relação ao controle, mas um deles alcançou maiores valores de expressão relativa (HOSSAIN *et al.*, 2017).

No tratamento com ácido salicílico (AS), o perfil de expressão de *PTOX 1* e *PTOX 2* foi similar ao observado para os tratamentos com PEG, desidratação e salinidade. Enquanto *PTOX 1* parece ser regulado positivamente, como observado claramente em folhas e raízes (6 h de tratamento) *PTOX 2* não sofreu alteração de expressão em folhas, sendo regulada negativamente em raízes (Gráfico 3). É relatado que o AS atua como um indutor do sistema de defesa vegetal contra o ataque de vírus (LEE *et al.*, 2008), portanto, o tratamento com esse fitormônio buscou simular um estresse biótico por meio da indução dessa resposta

de defesa da planta. Tang *et al.* (2017) relataram que tanto a aplicação isolada do fitormônio quanto combinada com estresse hídrico em plântulas de soja aumentam a expressão dos genes que codificam PTOX (apenas um) e AOX.

O tratamento de seca com um cultivar sensível (Benning) e um tolerante (PI) à desidratação revelou um padrão destoante da maior parte dos tratamentos analisados, porém interessante (Gráfico 7). Até então, o perfil de expressão dos dois genes frente a estresses mostrava-se antagônico: enquanto a expressão de *PTOX 1* subia em resposta a um tratamento, *PTOX 2* tendia a ser regulada negativamente. Neste experimento, todavia, o que se observou foi um aumento de expressão de ambos os genes, mas de forma muito mais marcante de *PTOX 2* e para o cultivar sensível. Ainda, os níveis de *PTOX 1* se mantiveram praticamente constantes durante todo o período do tratamento, ao passo que os níveis de *PTOX 2* aumentaram gradualmente à medida que o estresse prolongava-se. Comportamento semelhante pode ser visto nos pecíolos de *G. max* tratados com etileno (NAZ – Gráfico 8). Além deste padrão observado, os dois experimentos também têm em comum as condições drásticas a que as plantas foram expostas. No primeiro, as plantas inteiras foram deixadas no ar, em completa desidratação. No segundo, além do tratamento com etileno, hormônio conhecido por participar em respostas a estresses e promover a senescência (LEYSER; DAY, 2015), os segmentos tratados foram excisados da planta. Fica claro, portanto, o caráter adverso dessas condições de tratamento.

Quando a planta está passando pelo processo de senescência ou enfrenta uma condição de dano físico, além das respostas fisiológicas, uma intrincada rede de respostas moleculares é ativada, dentre as quais pode-se destacar a superexpressão de alguns genes em detrimento da supressão de outros; além disso, os fatores de transcrição exercem papel central nessas redes regulatórias (THOMAS *et al.*, 2015; BALAZADEH; RIAÑO-PACHÓN; MUELLER-ROEBER, 2008). Como já citado anteriormente, o promotor de *PTOX 2* contém regiões de ligação a fatores de transcrição envolvidos na senescência e em outros estresses. Neste sentido, o aumento expressivo dos transcritos de *PTOX 2* nas condições descritas há pouco deve, portanto, ser uma dessas respostas moleculares.

5.4 Afinidade das isoformas pelo plastoquinol e correlação com a função dos genes

A interação entre as duas isoformas da PTOX com o seu substrato revelou que PTOX 1 tem mais afinidade pelo ligante do que PTOX 2. Corroborando o resultado, um valor próximo ao encontrado neste trabalho para a energia de afinidade entre o complexo PTOX1-

PQH₂ (-7,6 kcal/mol) foi encontrado por Castro *et al.* (2017) em ensaio de *docking* molecular com uma das isoformas da AOX de *Citrus clementina* com seu substrato, ubiquinol (-7 kcal/mol). AOX e PTOX, por serem proteínas pertencentes à mesma classe, mantêm conservados na sua estrutura o centro de reação e os motivos de ligação ao ferro, que são essenciais para a atividade das enzimas.

Os dados obtidos neste trabalho entram em consonância e permitem supor que o gene *PTOX 1* de *Glycine max* parece ser transcrito preferencialmente, de forma geral, quando a célula enfrenta condições de estresse que não ameaçam a sobrevivência da planta. Neste sentido, a maior afinidade da enzima PTOX 1 pelo substrato pode explicar esta preferência que a célula tem pela expressão desse gene. Por outro lado, o gene *PTOX 2*, que possui dois íntrons maiores que *PTOX 1* (o que significa um maior gasto de energia por parte da célula para realizar a sua transcrição) e cuja proteína correspondente possui menor afinidade de ligação ao substrato, parece ser preferencialmente expresso quando em folhas e cotilédones se dá início ao processo de senescência, um evento altamente regulado e dependente de energia (THOMAS *et al.*, 2015) e quando as condições do meio são muito severas a ponto de serem capazes de ameaçar a sobrevivência da planta.

6 CONCLUSÃO

Os genes *PTOX 1* e *PTOX 2* são funcionais e expressos em praticamente todas as partes da planta de *Glycine max* analisadas. Os resultados também permitem inferir que as duas cópias em cromossomos distintos são resultado de eventos de duplicação e de rearranjos cromossômicos que ocorreram no genoma da soja e que resultou em ganho de função para a planta. Tanto as duas sequências dos cDNAs quanto as de proteínas compartilham alto grau de identidade, porém, o gene *PTOX 2* apresenta dois íntrons maiores que em *PTOX 1*. O perfil de expressão de cada gene é distinto durante o desenvolvimento de folhas, cotilédones e raízes e também é dependente do rigor das condições estressantes impostas. Além disso, a afinidade pelo ligante é maior para a isoforma da proteína codificada por *PTOX 1*. Desse modo, sugere-se que o gene *PTOX 1* em soja esteja associado com a eficiência fotossintética, prevalecendo, de forma geral, nas partes mais jovens e tendo a sua expressão aumentada preferencialmente quando o estresse aplicado não ocasiona danos severos à célula. Em contrapartida, sugere-se que o gene *PTOX 2*, por possuir na região promotora sítios de ligação a fatores de transcrição relacionados à senescência e requerer mais energia para a sua transcrição (tem íntrons maiores), além do seu produto proteico não ter tanta afinidade pelo ligante quanto *PTOX 1*, esteja mais relacionado a eventos específicos, como a senescência de folhas e cotilédones e com respostas a estresse muito severos, os quais ameaçam a manutenção da vida da planta.

REFERÊNCIAS

ALTSCHUL, S. F.; MADDEN, T.L., SCHÄFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res.**, [s.l.], v. 25, n. 17, p.3389-3402, set. 1997.

ALURU, M. R.; BAE, Hanhong; WU, Dongying; RODERMEL, Steven R. The Arabidopsis immutans Mutation Affects Plastid Differentiation and the Morphogenesis of White and Green Sectors in Variegated Plants. **Plant Physiology**, [s.l.], v. 127, n. 1, p.67-77, 1 set. 2001.

BALAZADEH, Salma; RIAÑO-PACHÓN, Diego Mauricio; MUELLER-ROEBER, Bern. Transcription factors regulating leaf senescence in Arabidopsis thaliana. **Plant Biology**, [s.l.], v. 10, p.63-75, set. 2008.

BARRETT, John Richard. The Science of Soy: What Do We Really Know? **Environmental Health Perspectives**, [s.l.], v. 114, n. 6, p. 352-358, 2006.

BELAMKAR, Vikas; WEEKS, Nathan T.; BHARTI, Arvind K.; FARMER, Andrew D.; GRAHAM, Michelle A; CANNON, Steven B. Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (*Glycine max*) during dehydration and salt stress. **Bmc Genomics**, [s.l.], v. 15, n. 1, p.950-975, 2014.

BELTRATI, C. M.; PAOLI, A. A. S. Semente. In: GLÓRIA, Beatriz Apezato da; GUERREIRO, Sandra Maria Carmello. **Anatomia Vegetal**. 2. ed. Viçosa: UFV, 2006. Cap. 15, p. 420

BERTHOLD, Deborah A.; STENMARK, Pål. MEMBRANE-BOUND IRON CARBOXYLATE PROTEINS. **Annual Review Of Plant Biology**, [s.l.], v. 54, n. 1, p.497-517, jun. 2003.

BINNECK, Eliseu. As ômicas: integrando a bioinformação. **Biotecnologia Ciência & Desenvolvimento**, [s.l.], v. 1, n. 32, p.28-37, jun. 2004.

BORISJUK, Ljudmilla; NGUYEN, Thuy Ha; NEUBERGER, Thomas; RUTTEN, T.; TSCHIRSCH, H.; CLAUS, B.; ... WOBUS, U. Gradients of lipid storage, photosynthesis and plastid differentiation in developing soybean seeds. **New Phytologist**, [s.l.], v. 167, n. 3, p.761-776, 7 jun. 2005.

BROWN, Anne; HUDSON, Karen. Developmental profiling of gene expression in soybean trifoliolate leaves and cotyledons. **Bmc Plant Biology**, [s.l.], v. 15, n. 1, p.1-14, 3 jul. 2015.

BRUCE, Barry D. Chloroplast transit peptides: structure, function and evolution. **Trends in Cell Biology**, v. 10, p. 440-447, 2000.

BUCHAN, Daniel W. A.; MINNECI, Federico; NUGENT, Thomas C.; BRYSON, Kevin.; JONES, David T. Scalable web services for the PSIPRED Protein Analysis Workbench. **Nucleic Acids Research**. V. 41, n. W1, p. 340-348, 2013.

BUSCH, F.; HUNER, N. P.a.; ENSMINGER, I. Increased Air Temperature during Simulated Autumn Conditions Impairs Photosynthetic Electron Transport between Photosystem II and Photosystem I. **Plant Physiology**, [s.l.], v. 147, n. 1, p.402-414, 21 mar. 2008.

CAROL, Pierre; STEVENSON, David; BISANZ, Cordelia; BREITENBACH, Jürgen; SANDMANN, Gerhard; MACHE, Regis; COUPLAND, George; KUNTZ, Marcel. Mutations in the Arabidopsis gene IMMUTANS cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. **The Plant Cell**, [s.l.], v.11, n. 1, p. 57-68, 1999.

CAROL, Pierre; KUNTZ, Marcel. A plastid terminal oxidase comes to light: implications for carotenoid biosynthesis and chlororespiration. **Trends In Plant Science**, [s.l.], v. 6, n. 1, p.31-36, jan. 2001.

CASTRO, Jacqueline Araújo; FERREIRA, Monique Drielle Gomes; SILVA, Raner José Santana; ANDRADE, Bruno Silva; MICHELI, Fabienne. Alternative oxidase (AOX) constitutes a small family of proteins in Citrus clementina and Citrus sinensis L. Osb. **Plos One**, [s.l.], v. 12, n. 5, p.1-17, 1 maio 2017.

CHEN, Vincent B.; ARENDALL, W. Bryan; HEADD, Jeffrey J.. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica Section D Biological Crystallography**, [s.l.], v. 66, n. 1, p.12-21, 21 dez. 2009.

CHEN, Y.A.; WEN, Y.C.; CHANG, W.C. "AtPAN: an integrated system for reconstructing transcriptional regulatory networks in Arabidopsis thaliana", **BMC Genomics** 13:85, 2012.

CHOW, Chi-Nga; ZHENG, Han-Qin; WU, Nai-Yun ; CHIEN, Chia-Hung; HUANG, Hsien-Da; Tzong-Yi LEE; CHIANG-HSIEH, Yi-Fan ; HOU, Ping-Fu ; YANG, Tien-Yi ; CHANG, Wen-Chi. PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. **Nucleic Acids Research**, n. 44, p. 1154-1160, 2015.

CONAB, Companhia Nacional de Abastecimento. In: Levantamentos de Safra. Disponível em <www.conab.gov.br>. Acesso em 22 jan. 2018.

CONESA, A.; MADRIGAL, P.; TARAZONA, S.; GOMEZ-CABRERO, D. CERVERA, A.; MCPHERSON, A.; SZCZESNIAK, M.W.; GAFFNEY, D.J.; ELLO, L.L.; ZHANG, X.; MORTAZAVI, A. A survey of best practices for RNA-seq analysis. **Genome Biol.**, 17:13, 2016.

COSTA, José Hélio; MCDONALD, Allison E.; ARNHOLDT-SCHMITT, Birgit; DE MELO, Dirce Fernandes. A classification scheme for alternative oxidases reveals the taxonomic distribution and evolutionary history of the enzyme in angiosperms. **Mitochondrion**, [s.l.], v. 19, p.172-183, nov. 2014.

DALL'AGNOL, Amélio. **A saga da soja no Brasil e no Mundo**. 2017. Disponível em: <https://www.agrolink.com.br/colunistas/coluna/a-saga-da-soja-no-brasil-e-no-mundo_400724.html>. Acesso em: 29 jan. 2018.

DANZER, John; MELLOTT, Eric; Q BUI, Anthu; LE, Brandon H; MARTIN, Patrick; HASHIMOTO, Meryl; PEREZ-LESHER, Jeanett; CHEN, Min; PELLETIER, Julie M.; SOMERS, David A.; GOLDBERG, Robert B.; HARADA, John, J. Down-Regulating the Expression of 53 Soybean Transcription Factor Genes Uncovers a Role for SPEECHLESS in Initiating Stomatal Cell Lineages during Embryo Development. **Plant Physiology**, [s.l.], v. 168, n. 3, p.1025-1035, 11 maio 2015.

DÍAZ, Milagros; DE HARO, Virginia; MUÑOZ, Romualdo; QUILES, María José. Chlororespiration is involved in the adaptation of Brassica plants to heat and high light intensity. **Plant, Cell & Environment**, [s.l.], v. 30, n. 12, p.1578-1585, dez. 2007.

EMANUELSSON, Olof; NIELSEN, Henrik; VON HEIJNE, Gunnar. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. **Protein Sci.** v. 8, n. 5, p. 978-84, 1999.

EMBRAPA SOJA. Disponível em < <https://www.embrapa.br/soja/cultivos/soja1>>. Acesso em 19 jan. 2018.

FARIAS, Jose Renato Boucas; NEUMAIER, Norman; NEPOMUCENO, Alexandre Lima. Soja. In: MONTEIRO, José Eduardo B. A. **Agrometeorologia dos Cultivos: O fator meteorológico na produção agrícola**. 1. ed. Brasília: INMET, p.263-277, 2009.

FEILKE, Kathleen; STREB, Peter; CORNIC, Gabriel; PERREAU, FRANÇOIS; KRUK, Jerzy; KRIEGER-LISZKAY, Anja. Effect of Chlamydomonas plastid terminal oxidase 1 expressed in tobacco on photosynthetic electron transfer. **The Plant Journal**, [s.l.], v. 85, n. 2, p.219-228, jan. 2016.

FEHR, W. R.; CAVINESS, C. E. Stages of soybean development. **Special Report**. Ames, Yowa: Yowa State University of Science and Technology, Cooperative Extension Service, n. 80, 11 p., 1977.

FIESP, Federação das Indústrias do Estado de São Paulo. In: Informativo DEAGRO janeiro de 2018. Disponível em < <http://www.fiesp.com.br/>>. Acesso em: 22 jan. 2018.

FU, Aigen; ALURU, Maneesha; RODERMEL, Steven R.. Conserved Active Site Sequences in Arabidopsis Plastid Terminal Oxidase (PTOX). **Journal of Biological Chemistry**, [s.l.], v. 284, n. 34, p.22625-22632, 19 jun. 2009.

FU, Aigen; PARK, Sungsoon; RODERMEL, Steven. Sequences Required for the Activity of PTOX (IMMUTANS), a Plastid Terminal Oxidase. **Journal of Biological Chemistry**, [s.l.], v. 280, n. 52, p.42489-42496, 25 out. 2005.

GAMBOA, Joaquín; MUÑOZ, Romualdo; QUILES, María José. Effects of antimycin A and n-propyl gallate on photosynthesis in sun and shade plants. **Plant Science**, [s.l.], v. 177, n. 6, p.643-647, dez. 2009.

GASTEIGER, Elisabeth; HOOGLAND, Christine; GATTIKER, Alexandre; DUVAUD, S'everine Duvaud.; WILKINS, Marc R.; APPEL, Ron D.; BAIROCH, Amos. Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): **The Proteomics Protocols Handbook**, Humana Press, 2005.

GEHRING, Mary; CHOI, Yeonhee; FISCHER, Robert L. Imprinting and Seed Development. **The Plant Cell Online**, [s.l.], v. 16, n. 1, p.203-213, 12 mar. 2004.

GOFFMAN, Fernando. D.; ALONSO, Ana P.; SCHWENDER, Jörg.; SHACHAR-HILL, Yair; OHLROGGE, John B. Light Enables a Very High Efficiency of Carbon Storage in Developing Embryos of Rapeseed. **Plant Physiology**, [s.l.], v. 138, n. 4, p.2269-2279, 1 ago. 2005.

GOODWIN, S.; MC PHERSON, J. D.; MC COMBIE W. R. Coming of age: ten years of next-generation sequencing technologies. **Nat. Ver. Genet.**, v. 17, p. 333-351, 2016.

HANWELL, Marcus D.; CURTIS, Donald Ephraim; LONIE, David C.; VANDERMEERSCH, Tim.; ZUREK, Eva; HUTCHISON, Geoffrey R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. **Journal of Cheminformatics**, v. 4, n. 17, 2012.

HELLEMANS, Jan.; MORTIER, Geert; PAEPE, Anne De; SPELEMAN, Frank; VANDESOMPELE, Jo. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. **Genome Biol.**, 8:R19, 2007.

HEYNO, Eiri; GROSS, Christine M.; LAUREAU, Constance; CULCASI, Marcel; PIETRI, Sylvia; KRIEGER-LISZKAY, Anja. Plastid Alternative Oxidase (PTOX) Promotes Oxidative Stress When Overexpressed in Tobacco. **Journal of Biological Chemistry**, [s.l.], v. 284, n. 45, p.31174-31180, 9 set. 2009.

HIGO, Kenichi; UGAWA, Yoshihiro; IWAMOTO, Masao; KORENAGA, Tomoko. Plant cis-acting regulatory DNA elements (PLACE) database. **Nucleic Acids Res.**, v. 27, n. 1, p. 297-300, 1999.

HOAGLAND, Denis Robert; ARNON, Daniel Israel. The water culture method for growing plants without soil. **California Agriculture Experimental Station Circular**. 1950, 347 p.

HOFMANN, Kay; STOFFEL, Wilhelm. TMbase - A database of membrane spanning proteins segments. **Biol. Chem. Hoppe-Seyler**, V. 374, p. 166, 1993.

HOOFT, Rob W. W.; VRIEND, Gert; SANDER, Chris; ABOLA, Enrique E. Errors in protein structures. **Nature**. V. 381, p. 272, 1996.

HOSSAIN, M. Sazzad; ELSAYED, Abdelaleim Ismail; MOORE, Marten; DIETZ, Karl-Josef. Redox and Reactive Oxygen Species Network in Acclimation for Salinity Tolerance in Sugar Beet. **Journal Of Experimental Botany**, [s.l.], v. 68, n. 5, p.1283-1298, 15 fev. 2017.

HU, Ruibo; FAN, Chengming; LI, Hongyu; ZHANG, Qingzhu; FU, Yong-Fu. Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. **Bmc Molecular Biology**, [s.l.], v. 10, n. 1, p.93-102, 2009.

HU, Bo; JIN, Jinpu; GUO, An-Yuan, ZHANG, He; LUO, Jingchu; GAO, Ge. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics**, 31(8):1296-1297, 2015.

JONES, David T. Protein secondary structure prediction based on position-specific scoring matrices. **J. Mol. Biol.** V. 292, p. 195-202, 1999.

JOSSE, Eve-marie. ALCARAZ, Jean-pierre; LABOURE, Anne-marie; KUNTZ, Marcel. In vitro characterization of a plastid terminal oxidase (PTOX). **European Journal Of Biochemistry**, [s.l.], v. 270, n. 18, p.3787-3794, set. 2003.

JOHNSON, Giles N.; STEPIEN, Piotr. Plastid Terminal Oxidase as a Route to Improving Plant Stress Tolerance: Known Knowns and Known Unknowns. **Plant And Cell Physiology**, [s.l.], p.1-10, 2 mar. 2016.

KIM, David; CHIVIAN, Dylan; BAKER, David. Protein structure prediction and analysis using the Robetta server. **Nucleic Acids Research**. V. 32, p.526-531, 2004.

KIM, Daehwan; PERTEA, Geo; TRAPNELL, Cole; PIMENTEL, Harold; KELLEY, Ryan; SALZBERG, Steven L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome Biol.**, v. 14, n. 4, 2013.

KIM, Joonyup; SUNDARESAN, Srivignesh; PHILOSOPH-HADAS, Sonia; YANG, Ronghui; MEIR, Shimon; TUCKER, Mark L. Examination of the Abscission-Associated Transcriptomes for Soybean, Tomato, and Arabidopsis Highlights the Conserved Biosynthesis of an Extensible Extracellular Matrix and Boundary Layer. **Frontiers in Plant Science**, [s.l.], v. 6, p.1-15, 15 dez. 2015.

KIM, Moon Young; LEE, Sunghoon; VAN, Kyujung; KIM, Tae-Hyung; JEONG, Soon-Chun; CHOI, Ik-Young; KIM, Dae-Soo; LEE, Yong-Seok; PARK, Daeui; MA, Jianxin; KIM, Woo-Yeon; KIM, Byoung-Chul; PARK, Sungjin; LEE, Kyung-A; KIM, Dong Hyun; KIM, Kil Hyun; SHIN, Jin Hee; JANG, Young Eun; KIM, Kyung Do; LIU, Wei Xian; CHAISAN, Tanapon; KANG, Yang Jae; LEE, Yeong-Ho; KIM, Kook-Hyung; MOON, Jung-Kyung; SCHMUTZ, Jeremy; JACKSON, Scott A.; BHAK, Jong; LEE, Suk-Ha. Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. **Proceedings of the National Academy of Sciences**, [s.l.], v. 107, n. 51, p.22032-22037, 3 dez. 2010.

KIM, Moon Young; VAN, Kyujung; KANG, Yang Jae; KIM, Kil Hyun; LEE, Suk-Ha. Tracing soybean domestication history: From nucleotide to genome. **Breeding Science**, [s.l.], v. 61, n.5, p.445-452, 2012.

KONG, Jin; GONG, Ji-Ming; ZHANG, Zhi-Gang; ZHANG, Jin-Song; CHEN, Shou-Yi. A new AOX homologous gene OsIM1 from rice (*Oryza sativa* L.) with an alternative splicing mechanism under salt stress. **Tag Theoretical And Applied Genetics**, [s.l.], v. 107, n. 2, p.326-331, 1 jul. 2003.

KRIEGER-LISZKAY, Anja; FEILKE, Kathleen. The Dual Role of the Plastid Terminal Oxidase PTOX: Between a Protective and a Pro-oxidant Function. **Frontiers In Plant Science**, [s.l.], v. 6, p.1-2, 5 jan. 2016.

- KUMAR, Sudhir; STECHER, Glen; TAMURA, Koichiro. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution**, [s.l.], v. 33, n. 7, p.1870-1874, 22 mar. 2016.
- KUNTZ, Marcel. Plastid terminal oxidase and its biological significance. **Planta**, [s.l.], v. 218, n. 6, p.896-899, 1 abr. 2004.
- LANGMEAD, Ben; SALZBERG, Steven L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 4, n. 9, p. 357-359, 2013.
- LASHKARI, D. A.; DERISI, J. L.; MCCUSKER, J. H.; NAMATH, A. F.; GENTILE, C.; HWANG, S. Y.; BROWN, P. O.; DAVIS, R. W. Yeast microarrays for genome wide parallel genetic and gene expression analysis. **Proc. Natl. Acad. Sci. U S A**, v. 94, p. 13057-13062, 1997.
- LEE, Wing-sham; FU, Shih-feng; LI, Zheng; MURPHY, A. M.; DOBSON, E. A.; GARLAND, L.; ... & CARR, J. P. Salicylic acid treatment and expression of an RNA-dependent RNA polymerase 1 transgene inhibit lethal symptoms and meristem invasion during tobacco mosaic virus infection in *Nicotiana benthamiana*. **Bmc Plant Biology**, [s.l.], v. 16, n. 1, p.1-14, 13 jan. 2016.
- LEISNER, Courtney P; MING, Ray; A AINSWORTH, Elizabeth. Distinct transcriptional profiles of ozone stress in soybean (*Glycine max*) flowers and pods. **Bmc Plant Biology**, [s.l.], v. 14, n. 1, p.1-13, 28 nov. 2014.
- LENNON, Adrian M.; PROMMEENATE, Peerada; NIXON, Peter J. Location, expression and orientation of the putative chlororespiratory enzymes, Ndh and IMMUTANS, in higher-plant plastids. **Planta**, [s.l.], v. 218, n. 2, p.254-260, 1 dez. 2003.
- LESCOT, Magali; DÉHAIS, Patrice; THUIS, Gert; MARCHAL, Kathleen; MOREAU, Yves; VAN DE PEER, Yves; ROUZÉ, Pierre; ROMBAUTS, Stephane. PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic Acids Res.**, Database issue, v. 30, n. 1, p. 325-327, 2002.
- LEYSER, Ottoline; DAY, Stephen. Signal Transduction: Ethylene signal transduction. In: BUCHANAN, Bob B.; GRUISSEM, Wilhelm; JONES, Russell L. **Biochemistry & Molecular Biology of Plants**. 2. ed. Oxford: Wiley Blackwell, 2015. Cap. 18, p. 847.
- LIMA, Karine Thiers Leitão. **Polimorfismos do gene da plastoquinol oxidase (PTOX) em ecotipos de *Arabidopsis thaliana* (projeto 1001 genomas): correlações de expressão gênica e de variantes estruturais da proteína com a distribuição geográfica dos ecotipos**. 2017. 87 f. Dissertação (Mestrado) - Curso de Bioquímica, Universidade Federal do Ceará, Fortaleza, 2017.
- LIU, L.; ZHOU, Y.; SZCZERBA, M. W.; LIN, Y. Identification and Application of a Rice Senescence-Associated Promoter. **Plant Physiology**, [s.l.], v. 153, n. 3, p.1239-1249, 3 maio 2010.

- MANAVALAN, Lakshmi P.; GUTTIKONDA, Satish K.; TRAN, Lam-Son Phan. Physiological and Molecular Approaches to Improve Drought Resistance in Soybean . **Plant & Cell Physiology**, [S.l.], v. 50 , n. 7, p. 1260-1276, 2009.
- MARSHALL, Owen J. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. **Bioinformatics**, [s.l.], v. 20, n. 15, p.2471-2472, 8 abr. 2004.
- MARTIN, L. B. B.; FEI, Z.; GIOVANNONI, J. J.; ROSE, J. K. C. Catalyzing plant science research with RNA-seq. **Front. Plant Sci.**, 4:66, 2013.
- MAXAM, A. M.; GILBERT, W. A new method for sequencing DNA. **Proc. Natl. Acad. Sci. USA**, v. 74, p. 560-564, 1977.
- MCDONALD, Allison E.; VANLERBERGHE, Greg C. Origins, evolutionary history, and taxonomic distribution of alternative oxidase and plastoquinol terminal oxidase. **Comparative Biochemistry and Physiology Part D: Genomics and Proteomics**, [s.l.], v. 1, n. 3, p.357-364, set. 2006.
- MCDONALD, Allison E.; IVANOV, Alex G.; BODE, Rainer; MAXWELL, Denis P.; RODERMEL, Steven R.; HÜNER, Norman P.A. Flexibility in photosynthetic electron transport: The physiological role of plastoquinol terminal oxidase (PTOX). **Biochimica Et Biophysica Acta (bba) - Bioenergetics**, [s.l.], v. 1807, n. 8, p.954-967, ago. 2011.
- MCGETTIGAN, P. A. Transcriptomics in the RNA-seq era. **Curr. Opin. Chem. Biol.**, v.17, p. 4-11, 2013.
- MISSÃO, Mauricio Roberto. Soja: origem, classificação, utilização e uma visão abrangente do mercado. **Revista de Ciências Empresariais**, [s.l.], v. 3, n. 1, p.7-15, jan. 2006.
- MOISE, Jaimie A.; HAN, Shuyou; GUDYNAITE-SAVITCH, Loreta; JOHNSON L., Douglas A.; MIKI, Brian L. A. Seed coats: Structure, development, composition, and biotechnology. **In Vitro Cellular & Developmental Biology - Plant**, [s.l.], v. 41, n. 5, p.620-644, set. 2005.
- MORANDO, Rafaela; SILVA, Alessandro O. da; CARVALHO, Leidiane C.; Pinheiro, Mírian P. M. A. DÉFICIT HÍDRICO: EFEITO SOBRE A CULTURA DA SOJA. **Journal Of Agronomic Sciences**, Umuarama, v. 3, n. especial, p.114-129, 2014.
- MUÑOZ, Romualdo; QUILES, María. Water Deficit and Heat Affect the Tolerance to High Illumination in Hibiscus Plants. **International Journal Of Molecular Sciences**, [s.l.], v. 14, n. 3, p.5432-5444, 7 mar. 2013.
- MUTZ, K. O.; HEILKENBRINKER, A.; LÖNNE, M.; WALTER, J. G.; STAHL, F. Transcriptome analysis using next-generation sequencing. **Curr. Opin. Biotechnol.**, v. 24, p. 22-30, 2013.

NAWROCKI, Wojciech J.; TOURASSE, Nicolas J.; TALY, Antoine; RAPPAPORT, Fabrice; ANDR'EWOLLMAN; Francis. The Plastid Terminal Oxidase: Its Elusive Function Points to Multiple Contributions to Plastid Physiology. **Annual Review Of Plant Biology**, [s.l.], v. 66, n. 1, p.49-74, 29 abr. 2015.

NAGALAKSHMI, U.; WANG, Z.; WAERN, K.; SHOU, C.; RAHA, D.; GERSTEIN, M.; SNYDER, M. The transcriptional landscape of the yeast genome defined by RNA sequencing. **Science**, v. 320, p. 1344-1349, 2008.

NIYOGI, Krishna K. Safety valves for photosynthesis. **Current Opinion in Plant Biology**, [s.l.], v. 3, n. 6, p.455-460, dez. 2000.

NIYOGI, K. K.; WOLOSUIK, R. A.; MALKIN, R. Photosynthesis. In: BUCHANAN, B.B.; GRUISSEM, W.; JONES, R. L. **Biochemistry & Molecular Biology of Plants**. 2. ed. Oxford: Wiley Blackwell, 2015. Cap. 12. p. 508-511.

OZSOLAK, Fatih; MILOS, Patrice M.. RNA sequencing: advances, challenges and opportunities. **Nature Reviews Genetics**, [s.l.], v. 12, n. 2, p.87-98, 30 dez. 2010.

PAN, Chang-Gui; PENG, Feng-Jiao; SHIB, Wen-Jun; HU, Li-Xin; WEI, Xiao-Dong; YING, Guang-Guo. Triclosan-induced transcriptional and biochemical alterations in the freshwater green algae *Chlamydomonas reinhardtii*. **Ecotoxicology And Environmental Safety**, [s.l.], v. 148, p.393-401, fev. 2018.

PAREDES, Miriam; QUILES, María José. Stimulation of chlororespiration by drought under heat and high illumination in *Rosa meillandina*. **Journal Of Plant Physiology**, [s.l.], v. 170, n. 2, p.165-171, jan. 2013.

PONTIUS, Joan; RICHELLE, Jean; WODAK, Shoshana J. Deviations from Standard Atomic Volumes as a Quality Measure for Protein Crystal Structures. **Journal of Molecular Biology**, [s.l.], v. 264, n. 1, p.121-136, nov. 1996.

PRINCE, Silvas J.; JOSHI, Trupti; MUTAVA, Raymond N.; SYED, Naeem; MALDONADO, Joao Vitor Dos Santos; PATIL, Gunvant; SONG, Li; WANG, Jiao Jiao; LIN, Li; CHEN, Wei; SHANNON, Grover J.; VALLIYODAN, Babu; XU, Dong; NGUYEN, Henry T. Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting. **Plant Science**, [s.l.], v. 240, p.65-78, nov. 2015.

RAMACHANDRAN, G.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, [s.l.], v. 7, n. 1, p.95-99, jul. 1963.

RITCHIE, Steven W. How a Soybean Plant Develops. **Special Report**, Iowa, n. 53, p.1-20, jul. 1985.

ROCHA, Antônio J.; Monteiro-Júnior, José E.; Freire, José E.C., Sousa, Antônio J.S.; Fonteles, Cristiane S.R. Real Time PCR: the Use of Reference Genes and Essential Rules Required to Obtain Normalisation Data Reliable to Quantitative Gene Expression. **Journal of**

Molecular Biology Research, [s.l.], v. 5, n. 1, p.45-55, 20 jul. 2015.

ROSSO, Dominic.; IVANOV, Alexander. G.; FU, Aigen; GEISLER, Matt; STEWART, Gregory; KROL, Marianna; HURRY, Vaughan; RODERMEL, Steven R.; MAXWELL, Denis P.; HÜNER, Norman P.A. IMMUTANS Does Not Act as a Stress-Induced Safety Valve in the Protection of the Photosynthetic Apparatus of Arabidopsis during Steady-State Photosynthesis. *Plant Physiology*, [s.l.], v. 142, n. 2, p.574-585, 11 ago. 2006.

RUMEAU, Dominique; PELTIER, Gilles; COURNAC, Laurent. Chlororespiration and cyclic electron flow around PSI during photosynthesis and plant stress response. **Plant, Cell and Environment**, v. 30, n. 9, p. 1041-1051, 2007.

SAIBI, Walid; BRINI, Faical; HANIN, Moez; MASMOUDI, Khaled. Development of Energy Plants and their Potential to Withstand Various Extreme Environments. **Recent Patents on Dna & Gene Sequences**, [s.l.], v. 7, n. 1, p.13-24, 1 fev. 2013.

SAITOU, Naruya; NEI, Masatoshi. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology And Evolution**, [s.l.], v. 4, n. 4, p.406-425, jul. 1987.

SAKAI, Hiroe; AOYAMA, Takashi; OKA, Atsuhiko. Arabidopsis ARR1 and ARR2 response regulators operate as transcriptional activators. **The Plant Journal**, [s.l.], v. 24, n. 6, p.703-711, 18 jul. 2000.

SANGER, F. DNA sequencing with chain-terminating inhibitors. **Proc. Natl. Acad. Sci. USA**, v. 74, p. 5463-5467, 1977.

SARAIVA, Kátia Daniella da Cruz MELO, Dirce Fernandes de; MORAIS, Vanessa D.; VASCONCELOS, Ilka Maria; COSTA, José Hélio. Selection of suitable soybean EF1 α genes as internal controls for real-time PCR analyses of tissues during plant development and under stress conditions. **Plant Cell Reports**, [s.l.], v. 33, n. 9, p.1453-1465, 13 maio 2014.

SARAIVA, Kátia Daniella da Cruz; OLIVEIRA, Antonio Edson Rocha; DOS SANTOS, Clesivan Pereira; LIMA, Karine Thiers Leitão; DE SOUSA, Janaína Martins; MELO, Dirce Fernandes de; COSTA, José Hélio. Phylogenetic analysis and differential expression of EF1 α genes in soybean during development, stress and phytohormone treatments. **Molecular Genetics and Genomics**, [s.l.], v. 291, n. 4, p.1505-1522, 16 mar. 2016.

SCHMUTZ, Jeremy; CANNON, Steven B.; SCHLUETER, Jessica ... JACKSON, Scott, A. Genome sequence of the palaeopolyploid soybean. **Nature**, [s.l.], v. 463, n. 7278, p.178-183, jan. 2010.

SCHRÖDINGER RELEASE 2017-2: Maestro, Schrödinger, LLC, New York, NY, 2017.

SEDIVY, Eric J.; WU, Faqiang; HANZAWA, Yoshie. Soybean domestication: the origin, genetic architecture and molecular bases. **New Phytologist**, [s.l.], v. 214, n. 2, p.539-553, 30 jan. 2017.

SHAHBAZI, Maryam; GILBERT, Matthias; LABOURÉ, Anne-Marie; KUNTZ, Marcel. Dual Role of the Plastid Terminal Oxidase in Tomato. **Plant Physiology**, [s.l.], v. 145, n. 3, p.691-702, 7 set. 2007.

SHAMIMUZZAMAN, Md; VODKIN, Lila. Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-Seq. **Bmc Genomics**, [s.l.], v. 14, n. 1, p.477-494, 2013.

SHIBA, Tomoo; KIDOA, Yasutoshi; SAKAMOTOA, Kimitoshi; INAOKAA, Daniel Ken; TSUGEA, Chiaki; TATSUMIA, Ryoko; TAKAHASHIB, Gen; BALOGUNA, Emmanuel Oluwadare; NARAD, Takeshi; AOKID, Takashi; HONMAE, Teruki; SHIN, Jin Hee; VAUGHN, Justin N; ABDEL-HALEEM, Hussein; CHAVARRO, Carolina; ABERNATHY, Brian; KIM, Kyung Do; JACKSON, Scott A.; LI, Zenglu. Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. **Bmc Plant Biology**, [s.l.], v. 15, n. 1, p.26-39, 2015.

SHIN, Jin Hee; VAUGHN, Justin N.; ABDEL-HALEEM, Hussein, CHAVARRO, Carolina;

ABERNATHY, Brian; KIM, Kyung Do; JACKSON, Scott A.; LI, Zenglu . Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. **Bmc Plant Biology**, [s.l.], v. 15, n. 1, p.26-39, 2015.

SHIRSAT, Anil; WILFORD, Neville; CROY, Ronald; BOULTER, Donald. Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. **Mol Gen Genet.**, [s.l.], v. 215, n. 2, p.326-331, jan. 1989.

SIEVERS, Fabian; WILM, Andreas; DINEEN, David; GIBSON, Toby J.; KARPLUS, Kevin; LI, Weizhong; LOPEZ, Rodrigo; MCWILLIAM, Hamish; REMMERT, Michael; SO'DING, Johannes; THOMPSON, Julie D.; HIGGINS, Desmond G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, [s.l.], v. 7, n. 1, p.539-539, 16 abr. 2014.

SIMKIN, Andrew J.; MOREAU, Helene; KUNTZ, Marcel; PAGNY, Gaëlle; LIN, Chenwei; TANKSLEY, Steve; MCCARTHY, James. An investigation of carotenoid biosynthesis in *Coffea canephora* and *Coffea arabica*. **Journal of Plant Physiology**, [s.l.], v. 165, n. 10, p.1087-1106, jul. 2008.

SOUSA, Francisco Yuri Maia de. **Caracterização bioquímica e molecular da oxidase termial da plastoquinona (PTOX) em Zea mays L.** 2008. 73 f. Dissertação (Mestrado) - Curso de Bioquímica, Universidade Federal do Ceará, Fortaleza, 2008.

SOUZA, Luciani de Liz; RHODEN, Sandro Augusto; PAMPHILE, João Alencar. A importância das ômicas como ferramentas para o estudo da prospecção de microrganismos: perspectivas e desafios. **Uningá Review**, [s.l.], v. 18, n. 2, p.16-21, jun. 2014.

SPENCER, M. W.b.; CASSON, S. A.; LINDSEY, K.. Transcriptional Profiling of the Arabidopsis Embryo. **Plant Physiology**, [s.l.], v. 143, n. 2, p.924-940, 22 dez. 2006.

SUN, Xin; LEI, Tao; DU, Jun-Bo; YANG, Wen-YU. Identification and Characterization of Two Paralogous Plastid Terminal Oxidase Genes in Soybean. **International Journal of Agriculture and Biology**, [s.l.], v. 17, n. 06, p.1275-1278, 1 nov. 2015.

SUN, Xin; WEN, Tao. Physiological roles of plastid terminal oxidase in plant stress responses. **Journal Of Biosciences**, [s.l.], v. 36, n. 5, p.951-956, 11 nov. 2011.

SUN, Xin; YANG, C.Q; WEN, T.; ZENG, F.C.; WANG, Q.; YANG, W.Y.; LIN, H.H. Water Stress Enhances Expression of Genes Encoding Plastid Terminal Oxidase and Key Components of Chlororespiration and Alternative Respiration in Soybean Seedlings. **Zeitschrift Für Naturforschung C**, [s.l.], v. 69, n. 7-8, p.300-308, 1 ago. 2014.

STECCANELLA, V.; HANSSON, M.; JENSEN, P. E. Linking chlorophyll biosynthesis to a dynamic plastoquinone pool. **Plant Physiology and Biochemistry**, v. 97, p. 207–216, 2015.

TALLÓN, Carlos; QUILES, María José. Acclimation to heat and high light intensity during the development of oat leaves increases the NADH DH complex and PTOX levels in chloroplasts. **Plant Science**, [s.l.], v. 173, n. 4, p.438-445, out. 2007.

TANG, Yanping; SUN, Xin; WEN, Tao; LIU, Mingjie; YANG, Mingyan; CHEN, Xuefei. Implications of terminal oxidase function in regulation of salicylic acid on soybean seedling photosynthetic performance under water stress. *Plant Physiology And Biochemistry*, [s.l.], v. 112, p.19-28, mar. 2017.

TEAKLE, Graham R.; MANFIELD, Iain W.; GRAHAM, John F.; GILMARTIN, Philip M. Arabidopsis thaliana GATA factors: organisation, expression and DNA-binding characteristics. **Plant Mol Biol.**, [s.l.], v. 50, n. 1, p.43-57, set. 2002.

THE PYMOL MOLECULAR GRAPHICS SYSTEM, Version 2.0 **Schrödinger**, LLC.

THOMAS, Howard; OUGHAM, Helen; MUR, Luis; JANSSON, Stefan. Senescence and Cell Death. In: BUCHANAN, Bob B.; GRUISSEM, Wilhelm; JONES, Russell L. **Biochemistry & Molecular Biology of Plants**. 2. ed. Oxford: Wiley Blackwell, 2015. Cap. 20. p. 925-952.

TRAN, Lam-son Phan; MOCHIDA, Keiichi. Functional genomics of soybean for improvement of productivity in adverse conditions. **Functional & Integrative Genomics**, [s.l.], v. 10, n. 4, p.447-462, 27 jun. 2010.

TRAPNELL, Cole; ROBERTS, Adam; GOFF, Loyal; PERTEA, Geo; KIM, Daehwan; KELLEY, David R; PIMENTEL, Harold; SALZBERG, Steven L.; RINN, John L; PACHTER, Lior. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. **Nature Protocols**, [s.l.], v. 7, n. 3, p.562-578, 1 mar. 2012.

TROTT, Oleg; OLSON, Arthur J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. 455–461, 2010.

TYCZEWSKA, Agata; GRACZ, Joanna; KUCZYŃSKI, Jakub; TWARDOWSKI, Tomasz. Deciphering the soybean molecular stress response via high-throughput approaches. **Acta Biochimica Polonica**, [s.l.], v. 63, n. 4, p.631-643, 2016.

USDA, United States Department of Agriculture. In: Oil Crops Outlook: January 2018. Disponível em < <https://www.ers.usda.gov/topics/farm-economy/commodity-outlook/>>. Acesso em 19 jan. 2018.

VILLAIN, Patricia; MACHE, Régis; ZHOU, Dao-xiu. The Mechanism of GT Element-mediated Cell Type-specific Transcriptional Control. **Journal Of Biological Chemistry**, [s.l.], v. 271, n. 51, p.32593-32598, 20 dez. 1996.

VERK, M. C.V.; HICHMAN, R.; PIETERSE, C. M. J.; WEES, S. C. M. V. RNA-Seq: revelation of the messengers. **Trends Plant Sci.**, v. 18, p. 175-179, 2013.

WANG, Yupeng; WANG, Xiyin; TANG, Haibao; TAN, Xu; FICKLIN, Stephen P.; FELTUS, Alex; PATERSON, Andrew H. Modes of Gene Duplication Contribute Differently to Genetic Novelty and Redundancy, but Show Parallels across Divergent Angiosperms. **Plos One**, [s.l.], v. 6, n. 12, p.1-17, 2 dez. 2011.

WEBB, Ben; SALI, Andrej. Comparative Protein Structure Modeling Using Modeller. **Current Protocols in Bioinformatics**. John Wiley & Sons Inc., V. 5, n. 6, p. 1-32, 2014.

WU, Dongying; WRIGHT, David A.; WETZEL, Carolyn; VOYTAS, Daniel F.; Steven. The IMMUTANS variegation locus of Arabidopsis defines a mitochondrial alternative oxidase homolog that functions during early chloroplast biogenesis. **The Plant Cell Online**, [s.l.], v. 11, n. 1, p.43-56, 1 jan. 1999.

YANAGISAWA, Shuichi; SCHMIDT, Robert J.. Diversity and similarity among recognition sequences of Dof transcription factors. **The Plant Journal**, [s.l.], v. 17, n. 2, p.209-214, jan. 1999.

YEUNG, Edward C.; MEINKE, David W. Embryogenesis in Angiosperms: Development of the Suspensor. **The Plant Cell Online**, [s.l.], v. 5, n. 10, p.1371-1381, 1 out. 1993.

ZHU, Jian-Kang. Cell signaling under salt, water and cold stresses. **Current Opinion in Plant Biology**, [s.l.], v. 4, n. 5, p.401-406, out. 2001.

ZHU, Jian-Kang. Salt and drought stress signal transduction in plants. **Annual Review of Plant Biology**, [s.l.], v. 53, n. 1, p.247-273, jun. 2002.

**ANEXO A - SEQUÊNCIAS DE AMINOÁCIDOS DAS PROTEÍNAS PTOX DA
ORDEM FABALES E DOS GRUPOS EXTERNOS USADAS NA CONSTRUÇÃO DA
ÁRVORE FILOGENÉTICA.**

>*Glycine max* PTOX1

MAAIFSSSLFVSSPLPKPLSSSYSRIHKAPQLFRARSSLLQDKEEKVIVQDSFSPKTSPLHSADKSTS
GNSINTSAFEKRI IKVEQSVNIFLTDSVIKILDITLYHDRHYARFFVLETIARVPYFAFMSVLHMYESF
GWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAVSPRMAYHFSE
CVESHAFETYDKFIKVQGDDELKMPAPEVAVNYYTGDDLYLFDEFQTSRVPNSRRPKIENLYDVFNIR
RDDEAEHCKTMKACQTHGNLRSPhSYAEDDDSSVCALEADCEGIVDCIKKSVTSNAKVK

>*Glycine max* PTOX2

MAAIFSSSLFASSPLPKPLSSSYSRIHTAPQLFRARSSLLQDNEKKVIVHDSFSPKTSPLHTADKSTG
GNSINTSAFEKRI IKVEQSVNIFLTDSVIKILDITLYHDRHYARFFVLETIARVPYFAFMSVLHMYESF
GWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAVSPRMAYHFSE
CVESHAFETYDKFIKVQGDDELKMPAPEVAVNYYTGDDLYLFDEFQTSRVPNSRRPKIENLYDVFNIR
RDDEAEHCKTMKACQTHGNLRSPhSYAEDDDSSVCALEADCEGIVDCIKKSVASNPAAKVK

>*Arachis duranensis* PTOX

MAATTLSSSLFATTTFKAPPPHSSYFSHNFPCTTRQSLPRVRATMLQDNEEKVKVKESLPSKISTFD
DAGKASIESGDSTSSSSSSSFEKFKVIVQDSVNIIFLTDSVIKILDITLYHDRDYARFFVLETIARVPYF
AFMSVLHMYESFGWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVFMY
AISPMMAYHFSECVESHAFETYDKFIKAQGEDLKKLPAPEVAVNYYTGDDLYLFDEFQTSRIPNTRRP
KIENLYDVFNIRRDDEAEHCKTMKACQTHGNLRSPhSDSADDDGDAALCNLEADCEGIVDCLKKSPTS
NPAKLLK

>*Arachis ipaensis* PTOX

MAATTLSSSLFATTTFKAPPPHSSSFSHNFPCTTRQSLPRVRATMLQDNEEKVKLKESSLPSKISTFD
DAGKASIESGDSTSSSSSSSFEKFKVIVQDSVNIIFLTDSVIKILDITLYHDRDYARFFVLETIARVPYF
AFMSVLHMYESFGWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVFMY
AISPMMAYHFSECVESHAFETYDKFIKAQGEDLKKLPAPEVAVNYYTGDDLYLFDEFQTSRIPNTRRP
KIENLYDVFNIRRDDEAEHCKTMKACQTHGNLRSPhSDSADNDGDAALCNLEADCEGIVDCLKKSPTS
NPAKLLK

>*Cajanus cajan* PTOX

MAAILSSSLLLPKHTRTSRPLFRARASLLQNNEEKLILPDSFSPKTSPLHTTTTKTTDADATKTS
AQKRI IKVEQSVNIFLTDSVIKILDALYHDRNYARFFVLETIARVPYFAFMSVLHMYESFGWWRRADY
LKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAI SPMMAYHFSECVERHAFE
TYDKFIKQGEELKMPAPEVAVDYTGDDLYLFDEFQTSRVPNSRRPKIENLYDVFNIRRDDEAEHC
KTMKACQTHGNLRSPhSYTADDDGSVCALEPGCEGIVDCITKSVTSSPAKVK

>*Cicer arietinum* PTOX

MAALFSSSLFPTTPLNKSQIHRPIHFPLSYRPPSLRICASLLQDKENKVIQDSFSPKTSPLDSSSTEN
SSSNNNNNNNNNNNSSSPSGWEKGVIVQDSVNIIFLTDSVIKILDALYHDRNYARFFVLETIARVPY
FAFMSVLHMYESFGWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYFMTALM
YTISPMMAYHFSECVESHAFETYDKFIKQGEELKMPAPEVAVNYYTGDDLYLFDEFQTSRVPNTRR
PKIENLYDVFMNIRRDDEAEHCKTMRACQTHGNLRSPhSYAEDDDDDSVCTPEAGCEGIVDCIKKSVTS
NPAKVK

>*Lotus japonicus* PTOX

MAAIFSSSIFSTSSLNKSPPSHYPHRHHNFSPLSLRSPLSRVRATVVKDTEGKVVVEKTFPSKIPLQT
GAEENVSDSDSSPPTAFELGVIVQDSINIFLTDSVIKILDALYDRNYARFFVLETIARVPYFAFMS
VLHMYESFGWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYFMTALMYTISP
RMAYHFSECVENHAFETYDKFIKQGEELKLPAPEVAVNYYTGDDLYLFDEFQTSRVPNTRRPKIEN
LYDVFNVRDDEAEHCKTMKACQTHGNLRSPhSYAEDDEDTVCPLEADCEGIVDCIKKSVTSSPAKVK

>*Lupinus angustifolius* PTOX

MAAILSSAVFATSPLKFTIPKHPSSLSPSSQLQFHPLRYNPLSSRSPLLRVRATVLKDEEEVITQES
FPSKTSPPFDVAKETNADSLNTPGAVEKWAIVQDSINIFLTDSVIKILDALYHDRNYARFFVLET
IARVPYFAFMSILHMYESFGWWRRADYLKVHFAESWNEMHLLIMEELGGNAWWFDRFLSQHVAVFYY
IMTVIMYAI SPMMAYHFSECVENHAFETYDKFIKAQGEELKLPAPEVAVNYYTGDDLYLFDEFQTSR
IPNTRRPKIENLYDVFNIRRDDEAEHCKTMKACQTHGNLRSPhSYTEDDDVNAVCDPEADCEGIVDCI

KKSVSSNPAKLK

>*Medicago truncatula* PTOX

MASTAMFSSSLFPITPLNKLSSSRNSLIFRPLSFRPPLFRIRSSLLQDKEDKVIQTQNTFPSKTSPLDS
VTENSTNDDDTSSSTSAWEKGVIKVEQSVNIFLTDSVIKILDALYRDRNYARFFVLETIARVPYFAFMS
ILHMYESFGWRRADYLVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYFMTALMYLISP
RMAYHFSECVESHAFETYDKFIKEQGEELKKMPAPEVAVNYYTGGDLYLDFEFQTSRVPNTRRPTIDN
LYDVFLNIRDDEAEHCKTMRACQTYGNLRS PHSYADAEDDDESVCTIEAGCEGIVDCIKKSVTSNPAK
VK

>*Phaseolus vulgaris* vPTOX

MAAILSSSLFATSSLPKPFYSRTHHTAPQFLRARASLLQDNEEKVIVQDSFSPKSSPLHTANGKSAAD
SINTSAFEKRIKVEQSVNIFLTDSVIKILDLYHDRDYARFYVLETIARVPYFAFMSVLHMYESFGW
WRRADYLVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAVSPRMAYHFSECV
ESHAFETYDKFIKAQGEELKKMPPEVAVNYYTGGDLYLDFEFQTSRVPNSRRPKIENLYDVFNIRD
DEAEHCKTMKACQTHGNLRS PHSYAEDDDVSVCALEADCEGIVDCIKKSVTPNPAKVK

>*Vigna angularis* PTOX

MAAILSSSLFASSCLPKPFSHPRTHHTASSLLKARASLLQDNEEKVIVQDSFSPKSSPLHTTNGKSSVD
SISTSAFEKGIKVEQSVNIFLTDSVIKILDLYHDRNYARFYVLETIARVPYFAFMSVLHMYESFGW
WRRADYLVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAISPRMAYHFSECV
ESHAFETYDKFIKVQGEELKKMPPEVAVKYTGGDLYLDFEFQTSRVPNSRRPKIDNLYDVFLNIRD
DEAEHCKTMKACQIHGNLRS PHSYAEDGDDVSACALEVDCEGIVDCIKKSVNPDPAKVK

>*Vigna radiata* PTOX

MAAILSSSLFASSCLPKPFSHPRTHHTASSLLKARASLLQDNEEKVIVQDSFSPKSSPLHTTNGKSSVD
SIGTSAFEKGIKVEQSVNIFLTDSVIKILDLYHDRHYARFYVLETIARVPYFAFMSVLHMYESFGW
WRRADYLVHFAESWNEMHLLIMEELGGNAWWFDRFLAQHIAIFYYIMTVLMYAISPRMAYHFSECV
ESHAFETYDKFIKVQGEELKKMPPEIAVKYTGGDLYLDFEFQTSRVPNSRRPKIDNLYDVFLNIRD
DEAEHCKTMKACQIHGNLRS PHSYAEDDDVVSACALEVDCEGIVDCLKKSVPNPDPAKVK

>*Arabidopsis thaliana* PTOX

MAAISGISSGTLMISRPLVTLRRSRAAVSYSSSHRLLHHLPLSSRRLLLRNNHRVQATILQDDEEKVV
VEESFKAETFTGKEPLEEPNMSSSSTSAFETWIIKLEQGVNVLFTDSVIKILDLYRDRTYARFFVLE
TIARVPYFAFMSVLHMYETFGWRRADYLVHFAESWNEMHLLIMEELGGNSWWFDRFLAQHIATFY
YFMTVFLYILSPRMAYHFSECVESHAYETYDKFLKASGEELKNMPAPDIAVKYTGGDLYLDFEFQTS
RTPNTRRPVIENLYDVFNIRDDEAEHCKTMRACQTLGCLRS PHSILEDDEEESGCVVPEEAHCEG
IVDCLKKSITS

>*Oryza sativa* PTOX

MAAVASASPLPAAAAPSTRCSPPPAFLPLRAHRPRVGTVATRRVFRAEAMRTQREKEQTEVAVEESFP
FRETAPPDEPLVTAEESWVVKLEQSVNIFLTVESVITILDGLYRDRNYARFFVLETIARVPYFAFISVL
HMYETFGWRRADYIKVHFAESWNEFHLLIMEELGGNSLWVDRFLARFAAFFYYFMTVAMYMVSPRM
AYHFSECVVERHAYSTYDKFIKLHEDELKKLPAPEAALNYYLNEDLYLDFEFQTARVPCSRPKIDNLY
DVFNIRDDEAEHCKTMKACQTHGNLRS PHSMQKCLETDTECVIPEDDCEGIVDCVKKSLVSKE

>*Volvox carteri* PTOX

MLPCPSLSRSGLHARRRTRGALRVHAVAAPIIPGKGECPLYRDRGTGKLI PAMCADYGFRSGAGRLYQE
SYGEVPKDVWQLAKDNRYRHELEQLRRAVRYPPSDVQRQSPHPVAKALHTANGVGAALASLDKALEEAR
VLPQLPPVRSALLETQEFKEIRARLDQLRLDADDVIAVERERATGGGDMESPLWVKAPFYALCWL
DIMYDNKPIEKFWVLETVARI PYFAYISILHLYESLGFWRAGAELRKIHFAEENEMHLLQIMESLGG
DRAWMDRFIAEHSVAFYYWVLIIFYLVSPRMAYNFMQRVELHAADTYTAFLQRNAAVLESIPPMVAL
QYYSEDLYLDFEFQTASRGAPRRRRCETLLDVFKNIRDDEMEHVKTMIACQNSTIAKDIAAASASS
SSSPSPSATELPAPATPPKRAAASTVVEE

PTOX1	GCATTTTGCTGAGAGCTGGAATGAAATGCATCATTGCTCATCATGGAAGAAGCTAGGGGG	865
PTOX2	GCATTTTGCTGAGAGCTGGAATGAAATGCATCATTGCTCATCATGGAAGAAGCTAGGGGG *****	759
PTOX1	GAATGCTTGGTGGTTTGACCGCTTCTCGCTCAACATATTGCAATATTTTATTATATTAT	925
PTOX2	GAATGCTTGGTGGTTTGATCGGTTCTCGCTCAACATATTGCAATATTTTATTATATTAT ***** ** *****	819
PTOX1	GACAGTTTTAATGTATGCAGTAAGCCCAAGAATGGCATATCACTTTTTCGGAATGTGTAGA	985
PTOX2	GACAGTTTTAATGTATGCAGTAAGCCCAAGAATGGCATATCACTTTTCTGAATGTGTAGA ***** *****	879
PTOX1	GAGTCATGCATTTGAAACTTATGACAAATTTATCAAGGTCCAAGGAGATGAATTGAAAAA	1045
PTOX2	GAGTCATGCATTTGAAACTTATGACAAATTTATCAAGGTCCAAGGAGATGAATTGAAAAA ***** *****	939
PTOX1	AATGCCAGTCCAGAGGTTGCTGTGAATTATTATACAGGAGATGACTTGTATTTATTTGA	1105
PTOX2	AATGCCAGTCCAGAGGTTGCTGTGAATTATTATACAGGAGATGACTTGTATTTATTTGA ***** *****	999
PTOX1	TGAATTTCAAACCTCCAGAGTTCCTAATTCAGAAGACCTAAGATAGAGAATCTGTACGA	1165
PTOX2	TGAATTTCAAACCTCCAGAGTTCCTAATTCAGAAGACCTAAGATAGAGAATCTGTACGA *****	1059
PTOX1	TGTATTTGTAAACATCAGAGATGATGAAGCTGAACATTGTAAGACAATGAAGGCTTGTC	1225
PTOX2	TGTATTTGTAAACATCAGAGATGATGAAGCTGAACATTGTAAGACAATGAAGGCTTGTC ***** ***** *****	1119
PTOX1	AACCCATGGGAACCTCCGGTCACCTCATTTCATATGCAGAGGATGATGACAGCTCAGTCTG	1285
PTOX2	AACCCATGGGAACCTCCGGTCACCTCATTTCATATGCAGAGGATGATGACAGCTCAGTCTG *****	1179
PTOX1	TGCTCTTGAAGCAGATTGTGAAGGAATAGTAGACTGTATAAAGAAATCTGTTACTTCTAA	1345
PTOX2	TGCTCTTGAAGCAGATTGTGAAGGAATAGTAGACTGTATAAAGAAATCTGTTACTTCTAA ***** *****	1239
PTOX1	TGCAGCTAAGGTGAAGTAAAGATAAGTGGCACAATACATAGTACATTCATTTAATAAGATA	1405
PTOX2	TGCAGCTAAGGTGAAGTAAAGATAAGTGGCACAATACATAGTACATTCATTTAATAAGATA * *****	1299
PTOX1	GGAGGCAGAGAATAATAATTGGAGGTAGTAAATAGTTGAAATTTGGATAAAGGTTGCTCTC	1465
PTOX2	GGAGGCAGAGAATAATAATTGGAGGTAGTAAATAGTTGAAATTTGGAAAAGGTTGCTCTC * ***** ***** ***** *****	1359
PTOX1	CGGTAAATTGTC-TTTATAGAAAGTAAAGGTTAA-----	1498
PTOX2	CGGTGTAATTGCTTTTATAGAAAGTAAAGGTTATTCACCTTAAATTTTCTAAGATCAA *** * ***** *	1419
PTOX1	-----ATATTCATACCACACTCT	1518
PTOX2	GATCATGAAATATCTTATAGTGACAACCTTCTGTAGGTTAAATGTTTCATACCACACTCT * * * *****	1479
PTOX1	GAGGATTATCAAAGAAAATAATATTATGGAGATTGTCTATAAATTTGCTCCTGCTCTTTT	1578
PTOX2	GAGGATTATCAAAGAAAATAATATTATGGAGATTGTCTATAAATTTGCTCCTGCTCTTTT ***** ***** ***** *****	1539
PTOX1	CCTTGTTTTTGGAATTGGAACTTTGTTCATTAGAGCTAGAGGGTCCATGTGTTGACTTT	1638
PTOX2	CCTTGTTTTTGGAATTGGAACTTTGTCTTTAGAGCTAGAGGGTGCATGCGTTGACTTT ***** ***** ***** ***** *****	1599
PTOX1	TAGAAACCATCCATCACTGAAATATTTAGATTGTGACAACATA	1680
PTOX2	-----	1599

**ANEXO C – GENES DE REFERÊNCIA PARA CADA PARTE DA PLANTA
UTILIZADOS PARA NORMALIZAÇÃO NO ENSAIO DE PCR EM TEMPO REAL**

Desenvolvimento

Flores e vagens: *EF1 β* e *EF1 α 1a1*

Sementes: *EF1 β* e *SKIP 16*

Cotilédones: *EF1 α 3* e *EF1 α 1b*

Folhas unifolioladas: *UKN1* e *SKIP 16*

Folhas trifolioladas: *UKN1* e *SKIP 16*

Raízes: *EF1 α 2a* e *UKN1*

Hipocótilo: *EF1 α 1a1* e *EF1 β*

Epicótilo: *EF1 α 1a2*, *EF1 α 2b*, *EF1 α 1a1*

Condições estressantes

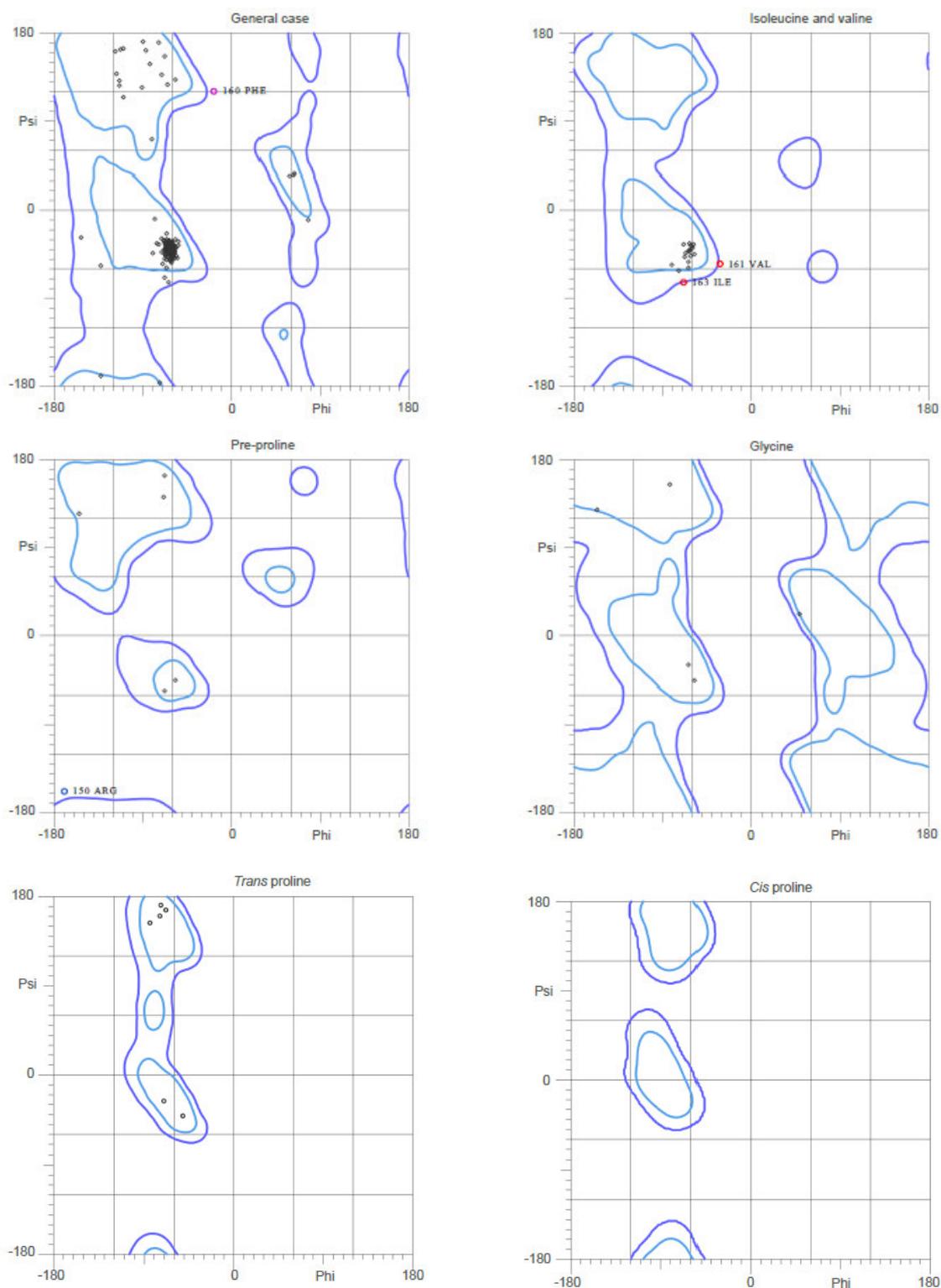
Folhas AS: *UKN1*, *SKIP 16*, *EF1 β* , *EF1 α 1a2* e *MTP*

Folhas PEG: *UKN1* e *EF1 β*

Raízes SA: *UKN1*, *EF1 α 1a1*, *EF1 β* e *SKIP 16*

Raízes PEG: *EF1 β* , *SKIP 16*, *UKN1*, *MTP* e *EF1 α 2b*

ANEXO D - GRÁFICOS DE RAMACHANDRAN PARA AS ESTRUTURAS TRIDIMENSIONAIS PREDITAS DAS ISOFORMAS DA PTOX EM *Glycine max.*
PTOX 1 – Predição Modeller

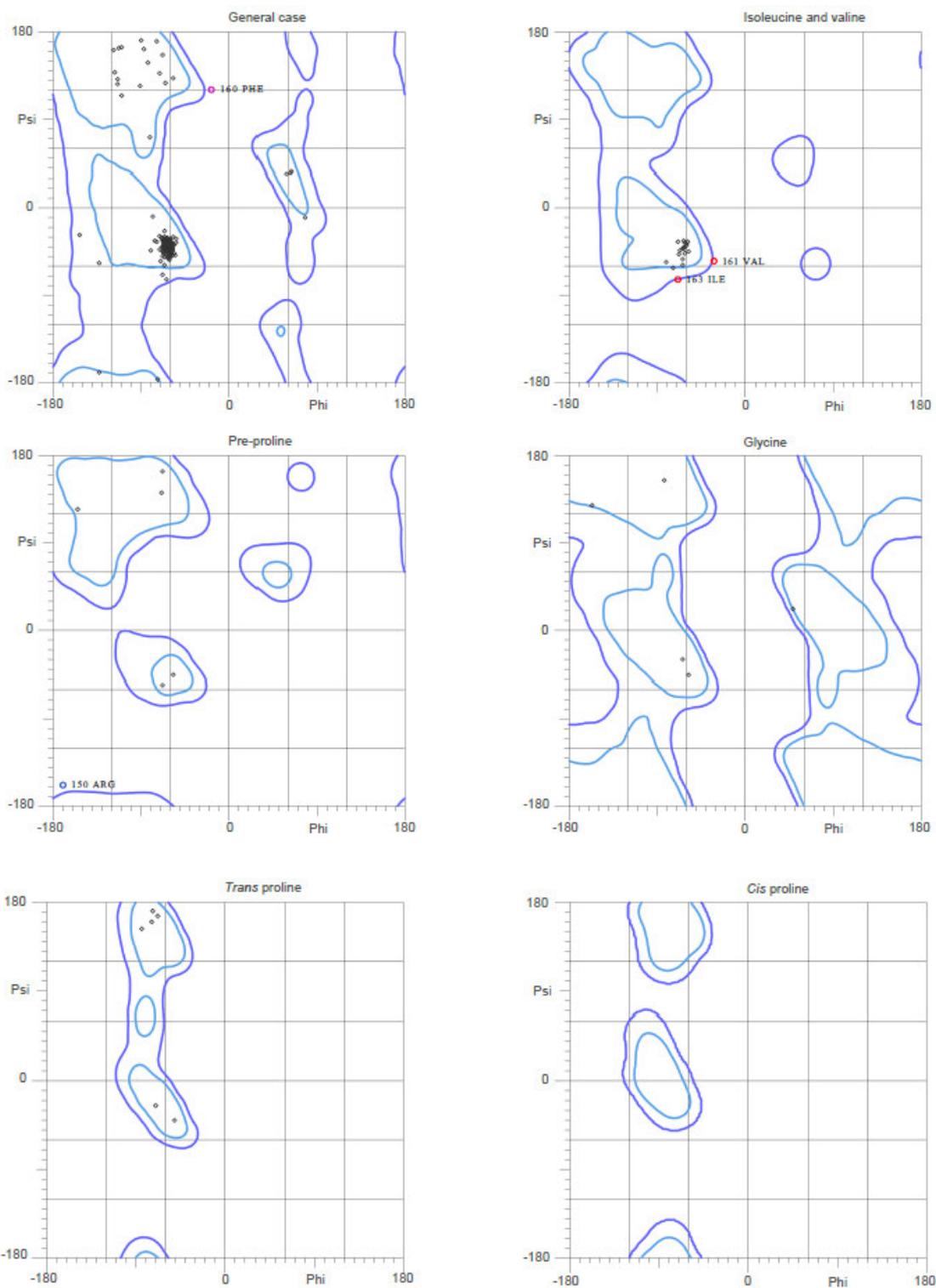


94.0% (158/168) of all residues were in favored (98%) regions.
 97.6% (164/168) of all residues were in allowed (>99.8%) regions.

There were 4 outliers (phi, psi):

- 150 ARG (-170.9, -160.4)
- 160 PHE (-18.5, 122.0)
- 161 VAL (-32.4, -56.0)
- 163 ILE (-69.0, -75.3)

PTOX 2 – Predição Modeller

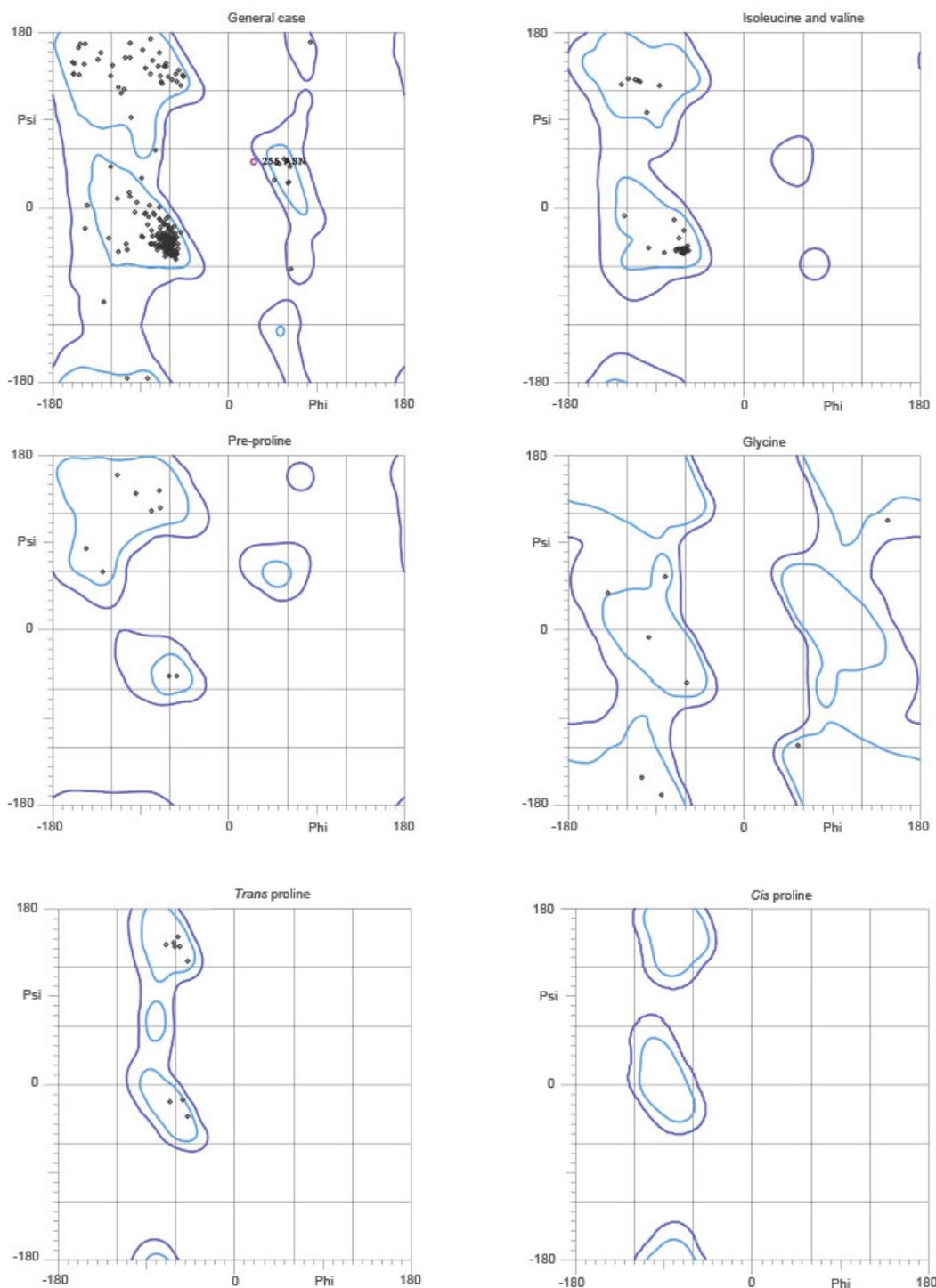


94.0% (158/168) of all residues were in favored (98%) regions.
 97.6% (164/168) of all residues were in allowed (>99.8%) regions.

There were 4 outliers (phi, psi):

150 ARG (-170.9, -160.4)
 160 PHE (-18.5, 122.0)
 161 VAL (-32.4, -56.0)
 163 ILE (-69.0, -75.3)

PTOX 1 – Predição Robetta



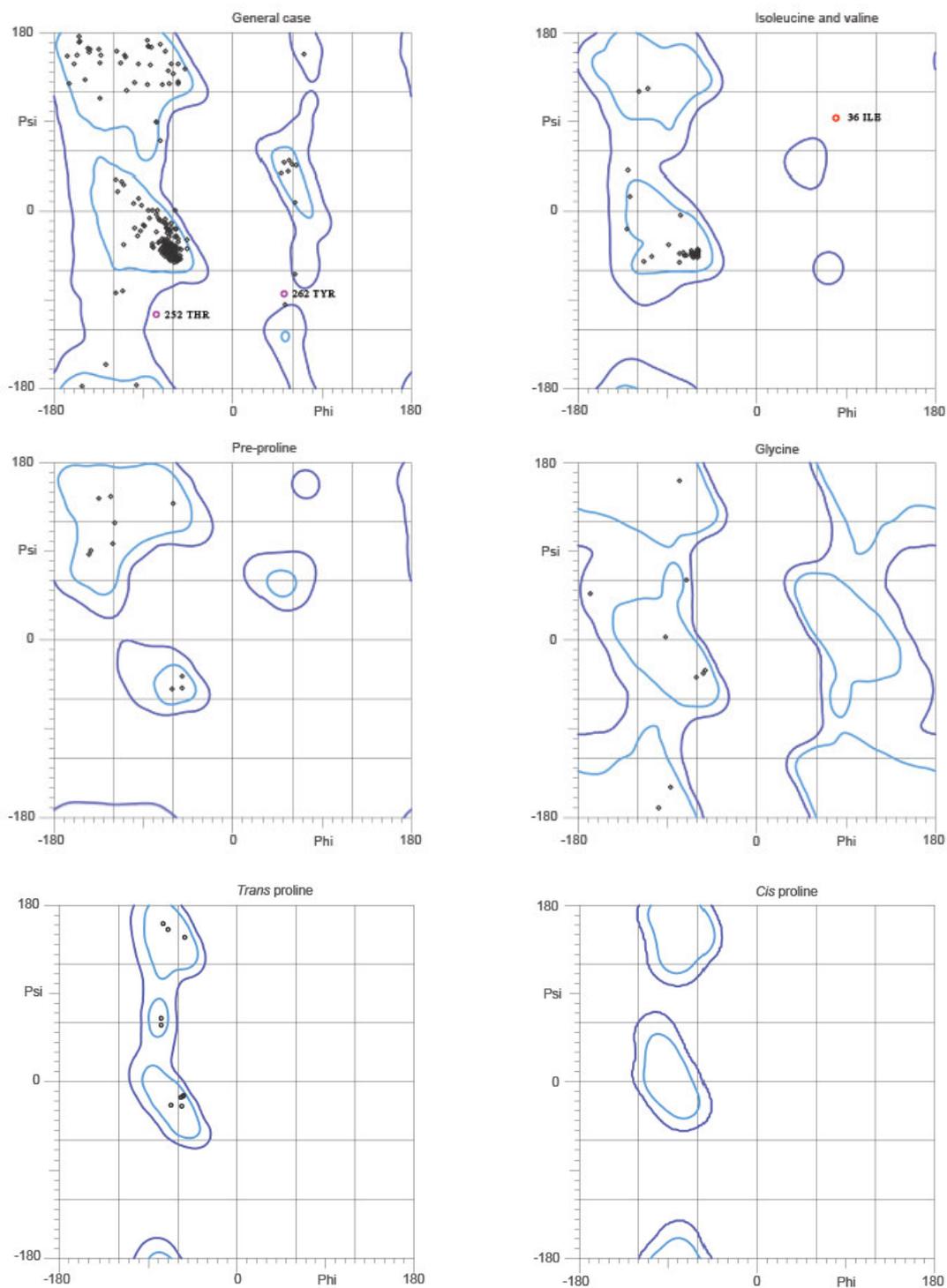
96.3% (283/294) of all residues were in favored (98%) regions.

99.7% (293/294) of all residues were in allowed (>99.8%) regions.

There were 1 outliers (phi, psi):

255 ASN (26.5, 47.2)

PTOX 2 – Predição Robetta



94.9% (279/294) of all residues were in favored (98%) regions.
 99.0% (291/294) of all residues were in allowed (>99.8%) regions.

There were 3 outliers (phi, psi):

36 ILE (80.3, 94.2)
 252 THR (-77.4, -106.9)
 262 TYR (52.6, -85.9)