



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

ANDRÉ LUIZ MAIA ROQUE

**CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE EXPRESSÃO DOS
GENES DO CICLO ASCORBATO-GLUTATIONA DE SOJA (*Glycine max*)**

FORTALEZA

2018

ANDRÉ LUIZ MAIA ROQUE

CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE EXPRESSÃO DOS
GENES DO CICLO ASCORBATO-GLUTATIONA DE SOJA (*Glycine max*)

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: Ciências Biológicas II.

Orientador: Prof. Dr. José Hélio Costa.

FORTALEZA

2018

ANDRÉ LUIZ MAIA ROQUE

CARACTERIZAÇÃO, ANÁLISE FILOGENÉTICA E PERFIL DE EXPRESSÃO DOS
GENES DO CICLO ASCORBATO-GLUTATIONA DE SOJA (*Glycine max*)

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: Ciências Biológicas II.

Aprovada em: 16/02/2018.

BANCA EXAMINADORA

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

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

Prof. Dr. Daniel Ferreira Feijó
Universidade Federal do Rio de Janeiro (UFRJ)

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)

R69c Roque, André Luiz Maia.

Caracterização, Análise Filogenética e Perfil de Expressão dos Genes do Ciclo Ascorbato-Glutationa de Soja (*Glycine max*) / André Luiz Maia Roque. – 2018.

117 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. Ciclo Ascorbato-Glutationa. 2. Análise filogenética. 3. Expressão gênica. 4. In silico. 5. *Glycine max*. I. Título.

CDD 572

AGRADECIMENTOS

À **Deus**, pois sem sua proteção nada seria possível;

Aos meus pais **Antônio Roque e Lúcia**, e ao meu irmão **Matheus**, pelo apoio em todos os momentos da minha vida, em nome deles agradeço a todos da minha família;

Ao **Prof. Dr. José Hélio Costa**, pelo apoio e ensinamentos transmitidos;

À **Prof^a. Dr. Kátia Daniella da Cruz Saraiva**, pela amizade, ensinamentos e participação em minha banca;

Ao **Prof. Dr. Daniel Ferreira Feijó**, pela participação em minha banca;

Aos colegas **Clesivan, Diêgo, Rachel, Thais e Vinícius** pela amizade, diversão e compartilhamento de conhecimentos ao longo do meu mestrado;

A todos os colegas do laboratório de Genômica Funcional e Bioinformática (**Anthony Netto, Beatriz, Fernanda, Karine, Lyndefânia, Moaciria, Mathias, Sara e Rodrigo**), pelas contribuições e convívio agradável ao longo do mestrado;

Aos meus amigos de infância, **Carlos, Henrique, Gêssica e Washington** pelos conselhos e diversões ao longo de 20 anos de amizade;

A todos os meus colegas de departamento;

À minha companheira **Scarlety Soares**, pela amizade, paciência, companheirismo e por fazer eu enxergar a vida muito além das dificuldades;

À Universidade Federal do Ceará (**UFC**), por me proporcionar um ambiente enriquecedor de conhecimentos;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**), pelo apoio financeiro com a manutenção da bolsa de auxílio.

“Há uma força motriz mais poderosa que o vapor, a eletricidade e a energia atômica: a vontade.”

Albert Einstein

“Viver é arriscar tudo. Caso contrário você é apenas um pedaço inerte de moléculas montadas aleatoriamente à deriva onde o universo te sopra”.

Rick e Morty.

RESUMO

Em condições normais, o peróxido de hidrogênio (H_2O_2) desempenha um papel fundamental como mensageiro secundário em resposta à diferentes processos celulares. No entanto, os estresses ambientais geralmente levam a superprodução de H_2O_2 causando danos oxidativos na célula vegetal. Consequentemente, tais efeitos deletérios comumente resultam em baixa produtividade ou até mesmo na morte da planta. Na tentativa de mitigar os efeitos deletérios do H_2O_2 , as plantas possuem um sistema de defesa enzimático, representado pela ascorbato peroxidase (APX; E.C. 1.11.1.11), monodehidroascorbato redutase (MDHAR; E.C. 1.6.5.4), dehidroascorbato redutase (DHAR; E.C. 1.8.5.1) e glutathiona redutase (GR; E.C. 1.6.4.2), coletivamente conhecido por ciclo ascorbato-glutathiona. O objetivo do presente trabalho foi caracterizar, analisar a filogenia e o perfil de expressão das famílias multigênicas que codificam APX, MDHAR, DHAR e GR em plantas de soja (*Glycine max*) submetidas às condições de estresse e durante o desenvolvimento. Inicialmente, realizou-se a anotação dos genes que codificam para o ciclo ascorbato-glutathiona através de buscas no genoma de soja, de outras nove espécies leguminosas, *Arabidopsis thaliana* (referência) e da alga *Volvox carteri* (grupo externo) depositados no *GenBank*. Em seguida, as sequências de proteínas deduzidas foram utilizadas para análise filogenética através do programa MEGA 7. Enquanto, análises de expressão *in silico* dos genes que codificam para o ciclo ascorbato-glutathiona da soja foi conduzida usando dados de transcriptomas depositados no *GenBank*. Os resultados mostraram que o ciclo ascorbato-glutathiona é codificado por 24 genes em ambas *G. max* e *G. soja*, enquanto um máximo de 15 genes foram encontrados nas demais espécies leguminosas. A análise filogenética revelou que as proteínas APX, MDHAR, DHAR e GR são formados por cinco, três, dois e dois clados, respectivamente. Em seguida, as proteínas de cada clado foram nomeadas de acordo com a classificação previamente estabelecida para *A. thaliana*. Para *G. max*, as proteínas foram nomeadas de Gm-APX-1a, 1b, 2, 3a, 3b, 4a, 4b, 5a1, 5a2, 5b1 e 5b2; Gm-MDHAR-1, 2a, 2b, 3a e 3b; Gm-DHAR-1a, 1b, 2a e 2b; Gm-GR-1a, 1b, 2a e 2b. A análise *in silico* de expressão dos genes em *G. max* submetida à estresse e sob condições de desenvolvimento revelou que todos os genes são funcionais (exceto *Gm-APX-5a2*), contudo, apresentam perfis de expressão característicos do tecido, estresse e estágio de desenvolvimento. O número de transcritos de *Gm-APX-3a*, *Gm-GR-1a*, *Gm-MDHAR-2b* e *Gm-MDHAR-3b*, que codificam proteínas cloroplastidiais, citosólicas, peroxissomais e citosólicas, respectivamente, aumentou em condições de elevada salinidade, seca, alta concentração de etileno e ataque do fungo *Fusarium oxysporum*. Além disso, as regiões

promotoras dos genes *Gm-APX-3a* e *Gm-MDHAR-3b* apresentaram elementos *cis* MYB responsivos ao déficit hídrico enquanto que a região promotora do gene *Gm-MDHAR-2b* possui uma grande quantidade de elementos *cis* GT1 responsivos à salinidade. Os resultados sugerem que esses genes possuem um importante papel na proteção antioxidante em soja submetida à estresses. Enquanto que os diferentes perfis de expressão para desenvolvimento de cotilédones e sementes sugerem um importante envolvimento dos genes que codificam proteínas citosólicas (*Gm-APX-4a*, *Gm-APX-4b*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-GR-1a* e *Gm-GR-1b*) e peroxissomais (*Gm-APX-5b1*, *Gm-MDHAR-2a* e *Gm-MDHAR-2b*), respectivamente, as zonas de desenvolvimento da raiz apresentaram genes que codificam proteínas citosólicas (*Gm-APX-4b*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-DHAR-1a*, *Gm-GR-1a* e *Gm-GR-1b*), cloroplastidiais (*Gm-APX-2*) e peroxissomais (*Gm-APX-5a1*, *Gm-APX-5b2*) expressos positivamente. Os achados no presente trabalho revelam a diversidade de genes e perfis transcricionais das famílias APX, MDHAR, DHAR e GR em soja, sugerindo uma complementação da defesa antioxidante em diferentes compartimentos celulares durante o desenvolvimento e estresse.

Palavras-chave: Ciclo ascorbato-glutationa. Análise filogenética. Expressão gênica. *In silico*. *Glycine max*.

ABSTRACT

Under normal conditions, hydrogen peroxide (H₂O₂) plays a key role as a secondary messenger in response to different cellular processes. However, environmental stresses often lead to H₂O₂ overproduction causing oxidative damage in plants. Consequently, such deleterious effects often result in low yield or even plant death. In order to mitigate the harmful effects of H₂O₂, plants have a set of enzymes belonging to the ascorbate-glutathione cycle named ascorbate peroxidase (APX; EC 1.11.1.11), monodehydroascorbate reductase (MDHAR; EC 1.6.5.4), dehydroascorbate reductase (DHAR; EC 1.8.5.1), and glutathione reductase (GR; EC 1.6.4.2). The aim of this study was to characterize, analyze the phylogeny and expression profile of multigene families encoding APX, MDHAR, DHAR, and GR in soybean (*Glycine max*) during stress conditions and development. Initially, the genes coding for the ascorbate-glutathione cycle were searched through of soybean genome, nine other leguminous species, *Arabidopsis thaliana* (reference) and the *Volvox carteri* algae (outgroup) deposited in *GenBank*. Next, the deduced protein sequences were used for phylogenetic analysis through the MEGA 7 program. *In silico* expression analyses of the genes encoding the ascorbate-glutathione cycle of soybeans were conducted using transcriptome data deposited in *GenBank*. The results showed that the ascorbate-glutathione cycle is encoded by 24 genes in both *G. max* and *G. soja*, while a maximum of 15 genes were found in the other leguminous species. Phylogenetic analysis revealed that the APX, MDHAR, DHAR, and GR proteins are formed by five, three, two and two clades, respectively. Then, proteins of each clade were classified according to those previously established for *Arabidopsis thaliana*. For *G. max*, proteins were named Gm-APX-1a, 1b, 2, 3a, 3b, 4a, 4b, 5a1, 5a2, 5b1, and 5b2; Gm-MDHAR-1, 2a, 2b, 3a, and 3b; Gm-DHAR-1a, 1b, 2a, and 2b; Gm-GR-1a, 1b, 2a, and 2b. *In silico* expression analyses of *G. max* genes submitted to stress and development conditions revealed that all genes are functional (except for *Gm-APX-5a2*), however, they exhibit characteristic expression profiles of tissue, stress and stage of development. The number of *Gm-APX-3a*, *Gm-GR-1a*, *Gm-MDHAR-2b* and *Gm-MDHAR-3b* transcripts, encoding chloroplast, cytosol, peroxisomes and cytosol proteins, respectively, increased under conditions of high salinity, dry, high ethylene concentration, and attack of the *Fusarium oxysporum* fungus. In addition, promoter regions of *Gm-APX-3a* and *Gm-MDHAR-3b* genes presented cis MYB elements responsive to water deficit while promoter region of the *Gm-MDHAR-2b* gene has a large amount of GT1 elements responsive to salinity. The results suggest these genes play an important role in antioxidant protection in soybeans submitted to

stress. While different expression profiles for cotyledon and seed development suggest an important involvement *Gm-APX-4a*, *Gm-APX-4b*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-GR-1a* and, *Gm-GR-1b* and *Gm-APX-5b1*, *Gm-MDHAR-2a*, and *Gm-MDHAR-2b* genes (encoding cytosolic and peroxisomal proteins, respectively), root developmental zones showed genes that encoded cytosolic (*Gm-APX-4b*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-DHAR-1a*, *Gm-GR-1a*, and *Gm-GR-1b*), chloroplastidial (*Gm-APX-2*), and peroxisomal proteins (*Gm-APX-5a1* and *Gm-APX-5b2*) positively expressed. The present study reveal the diversity of genes and transcriptional profiles of APX, MDHAR, DHAR and GR families in soybean and suggest the complementation of antioxidant defense in different cellular compartments during development and stress.

Keywords: Ascorbate-glutathione cycle. Phylogenetic analyses. Gene expression. *In silico*. *Glycine max*.

LISTA DE FIGURAS

- Figura 1 – Produção mundial de grãos de soja (em milhões de toneladas) na safra 2016/2017 e estimativa para 2017/2018.....20
- Figura 2 – Representação esquemática do ciclo do ascorbato-glutationa.....26
- Figura 3 – Árvores filogenéticas das proteínas APX (A), MDHAR (B), DHAR (C) e GR (D) de diferentes espécies leguminosas (*Glycine max*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Glycine soja*, *Medicago truncatula*, *Phaseolus vulgaris*, *Vigna angularis* e *Vigna radiata*), bem como *A. thaliana* e *Volvox carteri*.....40
- Figura 4 – Configurações éxon/íntron dos genes APX (A), MDHAR (B), DHAR (C) e GR (D) em *G. max*.....48
- Figura 5 – Número de elementos *cis* regulatórios ABRE (ACGTG), GT1 (GGTTAA/GAAAAA), MYB (CTAACCA/CNGTTR/WAACCA/TAACCTG), LTRE (CCGAAA), BBF1 (ACTTTA) e DRE (GTCGAC) na região promotora dos genes do ciclo ASC-GSH em soja (*G. max*).....53
- Figura 6 – Expressão dos genes do ciclo ASC-GSH em raízes e folhas de soja (*G. max*) submetidas à estresse salino ou seca.....55
- Figura 7 – Expressão dos genes do ciclo ASC-GSH na zona de abscisão foliar e pecíolos de soja (*G. max*) submetidos à etileno.....57
- Figura 8 – Expressão dos genes APX (A), MDHAR (B), DHAR (C) e GR (D) de *G. max* em 72 h após a infecção com isolados patogênicos e não patogênicos de *Fusarium oxysporum*.....58
- Figura 9 – Expressão dos genes do ciclo ASC-GSH durante o desenvolvimento de cotilédones (A) e estádios de maturação da semente (B) de soja (*G. max*).....60
- Figura 10 – Expressão dos genes APX (A), MDHAR (B), DHAR (C) e GR (D) de *G. max* em zonas meristemática, de alongamento e de diferenciação de raízes de soja.....61

LISTA DE TABELAS

Tabela 1 – Estádios vegetativos e reprodutivos da soja.....	18
Tabela 2 – Produção nacional da soja dividida por regiões.....	21
Tabela 3 – Experimentos de RNA-seq utilizados para obter o perfil de expressão dos genes <i>APX</i> , <i>MDHAR</i> , <i>DHAR</i> e <i>GR</i> em <i>Glycine max</i>	36
Tabela 4 – Número de genes da <i>APX</i> , <i>MDHAR</i> , <i>DHAR</i> e <i>GR</i> para espécies leguminosas.....	38
Tabela 5 – Identidade entre as sequências de aminoácidos de <i>APX</i> em <i>G. max</i> utilizando alinhamento pelo BLASTp (%).....	45
Tabela 6 – Identidade entre as sequências de aminoácidos de <i>MDHAR</i> em <i>G. max</i> utilizando alinhamento pelo BLASTp (%).....	46
Tabela 7 – Identidade entre as sequências de aminoácidos de <i>DHAR</i> em <i>G. max</i> utilizando alinhamento pelo BLASTp (%).....	46
Tabela 8 – Identidade entre as sequências de aminoácidos de <i>GR</i> em <i>G. max</i> utilizando alinhamento pelo BLASTp (%).....	46
Tabela 9 – Predição da localização subcelular para as proteínas <i>APX</i> , <i>MDHAR</i> , <i>DHAR</i> e <i>GR</i> em <i>G. max</i>	51

LISTA DE ABREVIATURAS E SIGLAS

APX	Ascorbato peroxidase
ASC	Ascorbato reduzido
BLAST	<i>Basic Local Alignment Search Tool</i>
CAT	Catalase
cDNA	DNA Complementar
CICLO ASC-GSH	Ciclo Ascorbato-Glutationa
CIT	Citosol
CLO	Cloroplasto
CONAB	Companhia Nacional de Abastecimento
DHAR	Dehidroascorbato redutase
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EROs	Espécies Reativas de Oxigênio
FPKM	<i>Fragments per Kilobase Million</i>
FO	<i>Fusarium oxysporum</i>
GPX	Guaiacol Peroxidase
GR	Glutationa redutase
GSH	Glutationa reduzida
GSSG	Glutationa oxidada
IPCC	<i>Intergovernmental Panel on Climate Change</i>
MDHAR	Monodehidroascorbato redutase
MIT	Mitocôndria
NCBI	<i>National Center for Biotechnology Information</i>
NGS	<i>Next Generation Sequencing</i>
PEG	Polietilenoglicol
PER	Peroxisomo
SOD	Superóxido Dismutase
SRA	<i>Sequence Read Archive</i>
TIL	Tilacóide
USDA	<i>United States Department of Agriculture</i>
UTR	<i>Untranslated Region</i>

SUMÁRIO

1	INTRODUÇÃO	13
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	<i>Glycine max</i> (L.) MERR.	16
2.1.1	<i>Origem e aspectos gerais da espécie</i>	16
2.1.2	<i>A cultura de soja (Glycine max)</i>	19
2.2	Estresses ambientais em plantas	21
2.2.1	<i>Aspectos gerais dos principais estresses nas culturas</i>	22
2.3	O estresse oxidativo e as plantas	24
2.3.1	<i>Espécies reativas de oxigênio (ROS)</i>	24
2.3.2	<i>Ciclo ascorbato-glutationa</i>	26
2.3.2.1	<i>Características gerais das enzimas do ciclo ASC-GSH e perfil de expressão gênica</i>	27
2.4	Análise transcriptômica e estudos <i>in silico</i> de caracterização e expressão gênica	30
3	OBJETIVOS	32
3.1	Objetivo Geral	32
3.2	Objetivos Específicos	32
4	MATERIAL E MÉTODOS	33
4.1	Identificação e caracterização dos genes da APX, MDHAR, DHAR e GR em <i>Glycine max</i> e outras espécies leguminosas	33
4.2	Alinhamento das sequências e análise filogenética	33
4.3	Configuração éxon/intron dos genes da APX, MDHAR, DHAR e GR em <i>Arabidopsis thaliana</i> e <i>Glycine max</i>	34
4.4	Predição da localização subcelular e análise de domínios das enzimas do ciclo ascorbato-glutationa em <i>Glycine max</i>	34
4.5	Análise das regiões promotoras	34
4.6	Análises de expressão <i>in silico</i> utilizando dados de RNA-seq para <i>Glycine max</i>	35
4.7	Análise estatística	38
5	RESULTADOS	38
5.1	Identificação e análise filogenética dos genes APX, MDHAR, DHAR e GR	38

	em <i>Glycine max</i> e outras espécies leguminosas.....	
5.2	Identidade das proteínas e conservação do domínio proteico em parálogos das famílias APX, MDHAR, DHAR e GR em <i>Glycine max</i>.....	43
5.3	Caracterização estrutural dos membros gênicos da APX, MDHAR, DHAR e GR em <i>Glycine max</i>.....	47
5.4	Localização subcelular das proteínas APX, MDHAR, DHAR e GR em <i>Glycine max</i>.....	50
5.5	Possíveis elementos <i>cis</i> responsivos à estresses nas regiões promotoras nos genes APX, MDHAR, DHAR e GR.....	51
5.6	Expressão dos genes APX, MDHAR, DHAR e GR em soja submetida às condições de estresses.....	54
5.6.1	<i>Expressão dos genes APX, MDHAR, DHAR e GR em soja submetida à seca e salinidade.....</i>	54
5.6.2	<i>Expressão dos genes APX, MDHAR, DHAR e GR em soja submetida à alta concentração de etileno.....</i>	56
5.6.3	<i>Expressão dos genes APX, MDHAR, DHAR e GR em raízes de soja infectada por Fusarium oxysporum.....</i>	57
5.7	<i>Expressão dos genes APX, MDHAR, DHAR e GR em condições de desenvolvimento da soja.....</i>	58
6	DISCUSSÃO.....	62
7	CONCLUSÃO.....	68
8	REFERÊNCIAS.....	69
	ANEXO A - SEQUÊNCIAS DAS PROTEÍNAS DEDUZIDAS DOS DIFERENTES MEMBROS GÊNICOS DA APX, MDHAR, DHAR E GR UTILIZADAS PARA A FORMAÇÃO DA ÁRVORE FILOGENÉTICA....	80
	APÊNDICE A – SEQUÊNCIAS DOS DOMÍNIOS DAS PROTEÍNAS OBTIDAS PELA PLATAFORMA PFAM.....	113

1 INTRODUÇÃO

A soja (*Glycine max*) é uma das espécies leguminosas mais cultivadas no mundo, pois fornece proteínas e óleo vegetal utilizados na alimentação humana, ração animal e produção de biodiesel (RODRIGUES *et al.*, 2012; PRINCE *et al.*, 2015). Os dados da safra 2016/2017 revelaram uma produção de aproximadamente 350 milhões de toneladas do grão em uma área de cultivo em torno de 121 milhões de hectares (USDA, 2017). Entretanto, o aumento da produtividade com o objetivo de atender à crescente demanda global tem sido afetado por vários estresses ambientais tais como seca, altas temperaturas, salinidade do solo, altas concentrações de ozônio troposférico e patógenos (LANUBILE *et al.*, 2015; RODRIGUES *et al.*, 2015; SHIN *et al.*, 2015).

Em resposta a esses estresses, as plantas desencadeiam uma cascata de reações induzidas pelo aumento da produção intracelular de espécies reativas de oxigênio (EROs) (RODRIGUES *et al.*, 2008; GILL & TUTEJA, 2010). As EROs são subprodutos de processos respiratórios e fotossintéticos que incluem radicais livres, moléculas e íons derivados do oxigênio molecular representados pelo ânion superóxido ($O_2^{\cdot-}$), radical hidroxila ($\cdot OH$), oxigênio singlete (1O_2) e peróxido de hidrogênio (H_2O_2) (SCANDALIOS *et al.*, 1997; MHAMDI *et al.*, 2010; SHARMA *et al.*, 2012). A ocorrência de estresses severos induz um desbalanço na relação EROs/compostos antioxidantes, podendo gerar danos significativos na estrutura celular (peroxidação de lipídeos, desregulação transcricional, carbonilação de enzimas e degradação da parede celular). Conseqüentemente, tais efeitos deletérios comumente resultam em baixas produtividades ou até mesmo na morte do organismo (PANCHUK *et al.*, 2002; MOLLER *et al.*, 2007; DAS, 2014).

O peróxido de hidrogênio desloca-se em diferentes compartimentos celulares, pois possui o maior tempo de meia vida entre todas as EROs. Devido a sua estabilidade, ele atua como sinalizador e iniciador na resposta adaptativa das plantas às adversas condições ambientais. Entretanto, em altas concentrações, H_2O_2 leva à morte programada/apoptose da célula (BHATTACHARJEE *et al.*, 2012; NOCTOR *et al.*, 2014; HOSSAIN *et al.*, 2015). Portanto, na tentativa de neutralizar ou mitigar os efeitos danosos causados pelo excesso de H_2O_2 , as plantas regulam seus sistemas de defesa enzimático e não enzimático. O sistema de defesa enzimático inclui o ciclo ascorbato-glutationa (ASC-GSH) que é formado por um grupo de quatro enzimas que atua na tentativa de manter a homeostase redox/celular através de um sistema altamente coordenado dependente de ascorbato e glutatona como fonte de equivalentes redutores (FOYER *et al.*, 2001; GILL & TUTEJA *et al.*, 2010; SINGH *et al.*,

2015).

As enzimas do ciclo ascorbato-glutationa são denominadas de ascorbato peroxidase (APX), monodehidroascorbato redutase (MDHAR), dehidroascorbato redutase (DHAR) e glutatona redutase (GR) (CHEW *et al.*, 2003; LATOWSKI *et al.*, 2010). A enzima APX usa ascorbato como doador de elétron para a redução de H₂O₂ gerando água e monodehidroascorbato (FRUGOLI *et al.*, 1996; SOFO *et al.*, 2015). O monodehidroascorbato formado pode ser regenerado a ascorbato pela ação da enzima MDHAR ou espontaneamente convertido à dehidroascorbato. A enzima DHAR reduz dehidroascorbato à ascorbato usando glutatona como agente redutor. Finalmente, a GR regenera a glutatona mantendo seu *pool* reduzido na célula e o fluxo antioxidante via as isoformas da APX ativas (NOCTOR *et al.*, 2011; CRUZ-RUS *et al.*, 2011).

As enzimas do ciclo ascorbato-glutationa são codificadas por pequenas famílias multigênicas. Por exemplo, para APX, cinco genes foram identificados em pepino (*Cucumis sativus* L.) (OZYIGT *et al.*, 2016) enquanto oito foram encontrados para ambos o milho (*Zea mays* L.) e arroz (*Oryza sativa* L.) (TEIXEIRA *et al.*, 2006; OZYIGT *et al.*, 2016). Por outro lado, a GR é codificada por somente dois genes em *Arabidopsis thaliana* e feijão caupi (*Vigna unguiculata*) (CONTOUR-ANSEL *et al.*, 2006; GARNIK *et al.*, 2006). Duas isoformas da DHAR foram identificadas em espinafre (*Spinacia oleracea*) (SHIMAOKA, YOKOTA & MIYAKE, 2000) e em eucalipto (*Eucalyptus* spp.) (TEIXEIRA *et al.*, 2005). Em musgo (*Physcomitrella patens*) e *A. thaliana* foram encontradas três e cinco isoformas da MDHAR, respectivamente (LUNDE *et al.*, 2006).

Cada família multigênica pode exibir diferentes padrões de expressão sob as mais variadas condições de estresses ambientais e desenvolvimento. Em arroz, o gene *APX8* foi altamente expresso após 1 h de estresse com NaCl, enquanto os demais genes da APX não sofreram nenhum aumento significativo das suas expressões (HONG *et al.*, 2007). Sob seca, apenas o gene *GR* que codifica uma proteína citosólica foi regulado positivamente em dois cultivares do feijão-caupi (CONTOUR-ANSEL *et al.*, 2006). A regulação positiva do gene *DHAR1* do tomate foi observada durante as 4 h iniciais sob estresse salino com posterior decréscimo da sua expressão ao longo de 48 h (KABIR & WANG, 2010). Em ervilha (*Pisum sativum* L.), o gene *MDHAR1* foi regulado positivamente sob baixas temperaturas e rompimento mecânico (LETERRIER *et al.*, 2005). YIN e colaboradores (2014) relacionaram a redução dos níveis de expressão dos genes da *CAT1*, *APX1* e *MDHAR1* aos efeitos deletérios da acumulação de EROs provocados pelo envelhecimento artificial de embriões de arroz.

Como observado, a maioria dos estudos focam somente na expressão de uma família multigênica. Entretanto, a regulação do ciclo da ascorbato-glutationa pode envolver a expressão ou co-expressão coordenada de membros específicos que codificam as isoformas da APX, MDHAR, DHAR e GR. Nesse contexto, a exploração de bancos de dados genômicos e transcriptômicos usando ferramentas de bioinformática foi fundamental para a caracterização e entendimento de perfis de expressão dos genes de interesse nesse estudo.

Dessa maneira, as tecnologias de sequenciamento da nova geração (do inglês, *NGS*) têm permitido o sequenciamento e a montagem de mais de 230 genomas de plantas, disponibilizando um número cada vez maior de projetos transcriptômicos disponíveis para várias espécies vegetais, entre elas a soja. Tais informações estão disponíveis gratuitamente no Centro Nacional de Informação Biotecnológica (do inglês, *NCBI*) (<https://www.ncbi.nlm.nih.gov/>).

Considerando a disponibilidade de uma grande quantidade de dados genômicos e transcriptômicos para *Glycine max* e a escassez de estudos integrando a caracterização e expressão dos genes do ciclo ascorbato-glutationa, o presente estudo, inicialmente, buscou a anotação dos genes da APX, MDHAR, DHAR e GR de espécies da ordem Fabales. Posteriormente, as análises de expressão *in silico* foram realizadas em soja sob estresses salino, seca e patógenos e diferentes tecidos em desenvolvimento (cotilédones, sementes e raízes). Levando em conta esta abordagem integrativa, os resultados poderão contribuir para: 1) avançar no entendimento de como se comportam os genes do ciclo da ascorbato-glutationa sob condições estressantes; 2) o estabelecimento de uma relação entre distribuição filogenética e papéis fisiológicos desses genes e 3) a identificação de genes promissores para o desenvolvimento de cultivares melhores adaptados.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 *Glycine max* (L.) MERR

2.1.1 *Origem e aspectos gerais da espécie*

A soja é uma espécie leguminosa, nativa do leste da Ásia e teve como seu progenitor silvestre a *Glycine soja*, sendo domesticada primeiramente na China por volta de 1100 anos a.C (LI *et al.*, 2010). Provavelmente, a soja foi introduzida nas Américas no ano de 1765, sendo utilizada como forragem (FEDERIZZI, 2005). No Brasil, a soja proveniente dos Estados Unidos foi introduzida na Bahia em 1882 e cultivada pela primeira vez em 1892 no Estado de São Paulo. Entretanto, foi a região sul do país, em especial, Rio Grande do Sul, que obteve um maior destaque na produção de soja. Nesse estado, a cultura da soja apresentou melhor adaptação às condições edafoclimáticas, resultando na primeira exportação brasileira no ano de 1949. Nos anos 70 e 80, a expansão dessa cultura foi significativa na região centro-oeste do país (CÂMARA, 2015).

A soja pertence ao reino Plantae, divisão Magnoliophyta, classe Magnoliopsida, ordem Fabales, família Fabaceae, subfamília Faboideae, tribo Phaseoleae, gênero *Glycine*, espécie *Glycine max* e forma cultivada *Glycine max* (L.) Merr (KANCHANA *et al.*, 2015). É uma planta de cultivo anual, herbácea, autógama e de porte predominantemente ereto, na qual sua altura pode variar de 30 a 200 cm (MÜLLER, 1981).

O sistema radicular é constituído de uma raiz axial principal e raízes secundárias que apresentam nódulos que são produtos da relação simbiótica com bactérias do gênero *Bradyrhizobium* (fixadoras de nitrogênio) (FERH & CAVINESS, 1977; MASCARENHAS *et al.*, 2005)

O caule ereto e ramificado desenvolve-se a partir do eixo embrionário após o início da germinação. Quanto ao seu hábito de crescimento, os cultivares de soja podem ser de crescimento determinado (caule terminado por racemos florais) e indeterminado (caule não terminado por racemos florais), que varia de acordo com as características do ápice do caule principal (MUNDSTOCK & THOMAS, 2005).

As folhas são divididas em três tipos: cotiledonar, unifoliolada ou trifoliolada; todas com formato, tamanho e posicionamento diferentes. Possui em torno de duas a trinta e cinco flores por racemo e é considerada uma planta de dias curtos por precisar de um mínimo de horas de escuro para a indução floral (MUNDSTOCK & THOMAS, 2005; ROCHA, 2009).

O ciclo de desenvolvimento da soja varia de 75 a 200 dias (MÜLLER, 1981). Segundo Ferh e Caviness (1977), os estádios de desenvolvimento dessa leguminosa são divididos em vegetativo (V) e reprodutivo (R) (Tabela 1). O estágio vegetativo corresponde ao período compreendido da emergência da plântula até a abertura da primeira flor. As subdivisões da fase vegetativa são designadas como V1, V2, V3, até Vn, menos os dois primeiros estádios que são designados VE (emergência da plântula) e VC (estádio de cotilédone). O último estágio vegetativo (Vn) representa o número do último nó vegetativo e pode variar de acordo com as diferenças ambientais na qual o cultivar foi submetido. Já a fase reprodutiva está relacionada com o período que compreende a floração até o fim do ciclo da cultura. No total são oito estádios que podem ser divididos em quatro fases, florescimento (R1 e R2), desenvolvimento da vagem (R3 e R4), desenvolvimento da semente (R5 e R6) e maturação da planta (R7 e R8) (RITCHIE, 1985) (Tabela 1).

Tabela 1: Estádios vegetativos e reprodutivos da soja

Estádio	Designação	Característica
VE	Emergência	Fixação da raiz primária no solo junto com a elongação do hipocótilo estabelece uma alavanca que ergue o cotilédone à superfície do solo.
VC	Cotilédone	Expansão e desdobramento das folhas unifolioladas.
V1	Primeiro nó	Folhas unifolioladas completamente desenvolvidas.
V2	Segundo nó	Primeira folha trifoliolada completamente desenvolvida.
V3	Terceiro nó	Segunda folha trifoliolada completamente desenvolvida.
V4	Quarto nó	Terceira folha trifoliolada completamente desenvolvida.
V5	Quinto nó	Quarta folha trifoliolada completamente desenvolvida.
V6	Sexto nó	Quinta folha trifoliolada completamente desenvolvida.
V...
Vn	Enésimo nó	Anti-enésima folha trifoliolada completamente desenvolvida.
R1	Início do florescimento	Uma flor aberta em qualquer nó do caule principal.
R2	Pleno florescimento	Uma flor aberta em um dos dois nós superiores na haste principal com uma folha totalmente desenvolvida.
R3	Início da formação das vagens	Vagem de 5 mm de comprimento em um dos quatro nós superiores na haste principal com uma folha totalmente desenvolvida.
R4	Plena formação das vagens	Vagem de 2 cm de comprimento em um dos quatro nós superiores na haste principal com uma folha completamente desenvolvida.
R5	Início do enchimento das sementes	Grão de 3 mm de comprimento em uma vagem em um dos quatro nós superiores na haste principal com uma folha totalmente desenvolvida.
R6	Pleno enchimento das vagens	Vagem contendo grãos verdes que preenchem a sua cavidade em um dos quatro nós superiores na haste principal com uma folha completamente desenvolvida.
R7	Início da maturação	Uma vagem normal na haste principal que atingiu sua maturação.
R8	Maturação plena	95% das vagens que atingiram maturação.

Fonte: Adaptado de Ferh e Caviness (1977)

A caracterização dos estádios de desenvolvimento da soja é essencial para a descrição dos vários períodos que a lavoura atravessa durante o ciclo da cultura (FERH & CAVINESS, 1977). Essa linguagem unificada auxilia na comunicação dos diversos públicos para a realização do manejo mais adequado e, conseqüentemente, obter uma maior produtividade, tendo em vista o posto da soja como a leguminosa mais consumida no mundo (EMBRAPA, 2017).

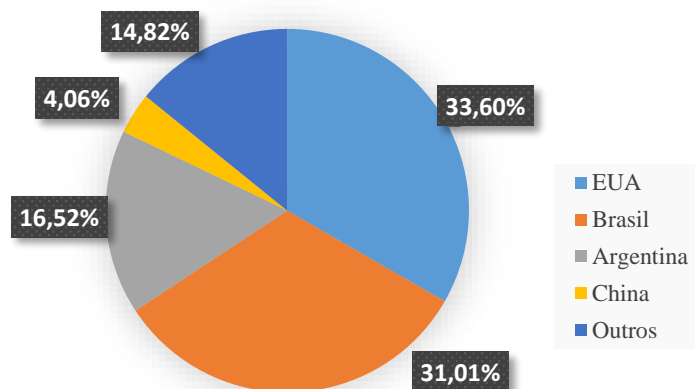
2.1.2 A cultura de soja (*Glycine max*)

Inicialmente, a soja foi utilizada para fins de adubação verde e, mais tarde, como fonte de óleo comestível, ração animal, matéria prima na indústria de alimentos, cosméticos, medicamentos e tintas. Dessa forma, ganhou um importante destaque na economia, sendo produzida e processada nos principais polos agroindustriais e químicos espalhados pelo mundo (FEDERIZZI, 2005; KOBERG, ABU-MUCH & GEDANKEN 2011; PRINCE *et al.*, 2015).

Em escala mundial, os principais produtores de soja do mundo são EUA, Brasil, Argentina e China (CONAB, 2017) (Figura 1). Dados da safra 2016/2017 revelaram que a produção mundial de soja chegou a mais de 351 milhões de toneladas em uma área plantada de aproximadamente 121 milhões de hectares (USDA, 2017). Só no Brasil, segundo maior exportador atrás somente dos EUA, a produção da soja atingiu 114 milhões de toneladas. Essa produção mantém a soja como o principal produto agrícola brasileiro com participação de 47,5% na safra nacional em 2017. Os principais estados brasileiros produtores do grão são Mato Grosso, Paraná e Rio Grande do Sul (CONAB, 2017).

Figura 1: Produção mundial de grãos de soja (em milhões de toneladas) na safra 2016/2017 e estimativa para 2017/2018

País	Safra	
	2016/2017	Estimativa 2017/2018
EUA	117,21	115,94
Brasil	114,00	107,00
Argentina	57,80	57,00
China	12,90	14,00
Outros	49,87	51,15
Total	351,78	345,09



Fonte: USDA/julho 2017

Inicialmente, a região sul do país era a produtora dominante de soja porque as variedades provenientes dos EUA eram muito sensíveis às mudanças de latitude e da época de plantio que ocasionam variações na duração do dia. Condições como a reação ao fotoperíodo e respostas à temperatura do ar provocavam um pequeno crescimento e rápido florescimento quando a cultura era implantada em outras regiões, resultando em baixo rendimento de grãos (MUNDSTOCK & THOMAS, 2005). Entretanto, após a introdução da característica período juvenil longo, a soja pôde se expandir para as regiões mais ao norte do Brasil (KIIHL & GARCIA, 1989). Durante o período juvenil a planta da soja não floresce mesmo em condições de fotoperíodo, permitindo que a soja cresça e desenvolva alta biomassa aparente, resultando em altos rendimentos (FEDERIZZI, 2005). Atualmente, o estado do Mato Grosso é o maior produtor de soja nacional (CONAB 2017) (Tabela 2).

Tabela 2: Produção nacional da soja dividida por regiões

	Safra 15/16	Safra 16/17
Norte	3.818,9	5,444,0
RR	79,2	99,0
RO	765,0	927,1
PA	1,288,0	1,595,8
TO	1.686,7	2.822,1
Nordeste	5.107,1	9.669,4
MA	1.250,2	2.498,0
PI	645,8	2.048,1
BA	3.211,1	5.123,3
Centro-oeste	43.752,6	50.149,9
MT	26.030,7	30.513,5
MS	7.241,4	8.575,8
GO	10.249,5	10.819,1
DF	231,0	241,5
Sudeste	7.574,9	8.126,6
MG	4.731,1	5.046,8
SP	2.843,8	3.079,8
Sul	35.181,1	40.540,3
PR	16.844,5	19.533,8
SC	2.135,2	2.292,6
RS	16.201,4	18.713,9
Brasil	95.434,0	113,930,2

Fonte: Adaptado de CONAB 2017

Apesar das técnicas de melhoramento genético que permitiram a expansão da soja, estima-se para 2017/2018 uma pequena queda da produção do grão (Figura 1). Essa queda deve-se principalmente às altas temperaturas e poucas chuvas nos principais estados produtores dos Estados Unidos e Brasil (USDA, 2017, CONAB, 2017). Outros estresses abióticos (e.g.: elevadas concentrações de ozônio, salinidade do solo, desregulação da luminosidade) e bióticos (e.g.: pragas e ataque de herbívoros) também ocasionam a redução significativa do rendimento das lavouras e restringem os locais, as épocas do ano e o solo onde a espécie pode ser cultivada (FARIAS, NEPOMUCENO & NEUMAIER 2007; DESHMUK *et al.*, 2014; LEISNER *et al.*, 2014; DAS *et al.*; 2016).

2.2 Estresses ambientais em plantas

As plantas, como organismos sésseis, vivem sob constantes mudanças ambientais que, na maioria das vezes, são desfavoráveis para o seu crescimento e desenvolvimento. Essas

condições adversas incluem os estresses abióticos representados por seca, calor, frio, deficiência de nutrientes, excesso de sais e metais no solo, elevada concentração de ozônio, dentre outros (LEISNER *et al.*, 2014; ZHU, 2016). Dentre esses estresses, seca, alto teor salino e elevadas temperaturas recebem atenção especial por serem mais frequentes em diversas regiões, principalmente em regiões áridas e semi-áridas, tais como o Nordeste brasileiro. (DE MENEZES & MORAIS, 2002; AHMAD *et al.*, 2012).

2.2.1 Aspectos gerais dos principais estresses nas culturas

A água é o principal constituinte dos vegetais, pois além de totalizar cerca de 90% da matéria verde, participa da maioria dos processos bioquímicos necessários para o crescimento e desenvolvimento da planta (TAIZ & ZEIGER, 2013). Entretanto, quando a absorção de água pelo sistema radicular não consegue atender às demandas fisiológicas da planta, diz-se que a mesma está sob estresse hídrico (FAN *et al.*, 2006).

A indisponibilidade hídrica é considerada como a condição ambiental de maior efeito na produtividade agrícola (PIMENTEL, 2004). Os efeitos da seca são bastante variáveis, pois dependem da sua intensidade e da velocidade de imposição do estresse além do estágio de desenvolvimento da planta na qual a seca ocorre (KRAMER & BOYER, 1995; PIMENTEL, 2004). A seca interfere na redução da fixação do carbono pelo aparato fotossintético devido ao fechamento dos estômatos, que afeta o acúmulo de fotoassimilados e, por consequência, influencia na absorção de nutrientes. Dessa forma, há um decréscimo da área foliar, aceleração da senescência e abscisão das folhas (TAIZ & ZEIGER, 2013). Além disso, a seca pode levar à danos nas membranas celulares e geração de espécies reativas de oxigênio (EROs) resultando na redução da biomassa ou até mesmo na morte da planta. Como resultado, há uma significativa diminuição do rendimento do produto agrícola (PRABA *et al.*, 2009; DAS *et al.*, 2016; FAHAD *et al.*, 2017).

O baixo índice pluviométrico associado às altas temperaturas e baixa qualidade da água utilizada na irrigação pode ocasionar o processo de salinização do solo e, conseqüentemente, influenciar no crescimento e desenvolvimento de culturas agrícolas (BABU & DEVARAJ, 2008; NETO *et al.*, 2012). A magnitude dos danos está relacionada com o tempo, concentração, tolerância da cultura e volume de água transpirado (SOUSA *et al.*, 2011). O efeito mais facilmente observado da salinidade sobre as plantas é a redução do crescimento em razão de desequilíbrios nutricionais que afetam o rendimento e a qualidade do produto agrícola (FERREIRA *et al.*, 2001).

O efeito potencial da salinidade em inibir processos de fotossíntese e transpiração estão relacionados com o grau de fechamento dos estômatos ou com o acúmulo excessivo de íons sódio e cloreto. O excesso desses íons causa um desbalanço na homeostase iônica nos cloroplastos, influenciando nos processos bioquímicos e fisiológicos da planta, bem como na estrutura dessa organela (ASHRAF & HARRIS, 2013; TAIZ & ZEIGER, 2013). Entre os danos fisiológicos também pode-se destacar a redução da área foliar, inibição no crescimento do sistema radicular e redução do número de folhas. Com relação aos danos bioquímicos, merece destaque o desbalanço no estado redox das células, através da produção excessiva de EROS (MUNNS, 2002; SHRIVASTAVA & KUMAR, 2015; TARI *et al.*, 2013).

Geralmente, em analogia ao estresse hídrico e salino, a temperatura elevada é um fator importante que afeta severamente o desenvolvimento das plantas. O grau de susceptibilidade da planta ao estresse térmico depende do seu estágio de desenvolvimento, entretanto altas temperaturas podem afetar todos os estádios de desenvolvimento da planta. Além disso, os efeitos do estresse variam entre espécies e genótipos (BARNABÁS, JÄGER & FEHÉR, 2008; SAKATA & HIGASHITANI, 2008).

As temperaturas elevadas causam severos danos fisiológicos às plantas, como a abscisão foliar e senescência, inibição do crescimento das raízes e danos aos frutos, levando à diminuição da produtividade da planta (VOLLENWEIDER & GÜNTHARDT-GOERG, 2005). A nível bioquímico, o estresse térmico causa danos em proteínas pertencentes ao fotossistema II e à fixação de carbono, reduzindo a eficiência do transporte de elétrons e, como resultado, proporcionando o acúmulo de espécies reativas de oxigênio (BITA & GERATS, 2013; FAHAD *et al.*, 2017).

Prevê-se que o problema da alta temperatura tende a se agravar, principalmente nas regiões áridas e semiáridas, haja vista que a temperatura global está aumentando entre 0,15 e 0,6 °C por década (BITA & GERATS, 2013; IPCC, 2013). Como resultado, o futuro da produção agrícola e, portanto, a segurança alimentar da população humana enfrentará desafios adicionais, tais como subsídio para políticas públicas e estímulo às pesquisas sobre culturas mais resistentes, e novas variedades, que muitas vezes demandam de décadas de estudos (BITA & GERATS, 2013).

Outro estresse que vem se intensificando nos últimos anos é a concentração elevada de ozônio troposférico (a nível do solo). Apesar dos seres vivos precisarem do ozônio na estratosfera para a proteção contra os raios ultravioletas, o ozônio troposférico é altamente tóxico e influencia diretamente no desenvolvimento das plantas através da geração excessiva de espécies reativas de oxigênio, aumento da senescência foliar, redução da assimilação de

carbono e até mesmo a morte celular programada (TAUSZ, GRULKE & WIESER, 2007; LUDWIKOW & SANDOWSKI, 2008).

Enquanto os estresses abióticos são causados por alterações de fatores edafoclimáticos, os estresses bióticos são danos causados às plantas por meio de organismos vivos, tais como fungos, bactérias e insetos. As plantas são hospedeiras de patógenos e pragas desde muito antes da sua domesticação. Um patógeno vegetal é qualquer organismo que utilize a planta hospedeira como substrato para completar uma parte ou a totalidade do seu ciclo de vida. Uma praga de planta, por outro lado, é um inseto herbívoro, nemátodo, mamífero ou pássaro que come tecido vegetativo, frutas e sementes (BUCHANAN, GRUISSEM & JONES, 2015). Os patógenos bem-sucedidos devem entrar na planta hospedeira, crescer longamente para obter nutrientes, suprimir as defesas e depois reproduzir para continuar seu ciclo de vida (BUCHANAN, GRUISSEM & JONES, 2015). Os patógenos são responsáveis por até 15% das perdas na produção mundial de alimentos e essa epidemia de doenças ou insetos deve continuar a causar perdas da produção de alimentos ao longo dos próximos anos (ONAGA & WYDRA, 2017).

Como visto, cada estresse pode induzir respostas fisiológicas, bioquímicas e moleculares específicas em plantas. Entretanto, uma resposta comum às plantas imposta por qualquer estresse é o desencadeamento de cascatas de reações induzidas por meio do aumento da produção de espécies reativas de oxigênio (EROs) que, em pequenas quantidades, podem induzir genes de defesa e respostas adaptativas. Porém, em altas concentrações, podem levar a oxidação de componentes celulares importantes como membranas celulares, lipídeos e até mesmo o material genético da célula. Essa oxidação leva à instabilidade das membranas, mutações prejudiciais e paralisação de processos metabólicos essenciais à sobrevivência por toxicidade, podendo levar a morte celular (GILL & TUTEJA, 2010; KARUPPANAPADIAN *et al.*, 2011; SHARMA *et al.*, 2012; PANDEY *et al.*, 2015).

2.3 O estresse oxidativo e as plantas

2.3.1 Espécies reativas de oxigênio (EROs)

Há aproximadamente 2,7 bilhões de anos o oxigênio molecular foi introduzido no ambiente terrestre por meio de organismos fotossintetizantes e, conseqüentemente, as espécies reativas de oxigênio (EROs) se tornaram compostos inevitáveis à vida aeróbica. Estima-se que até 1% do O₂ consumido por plantas é direcionado para a formação desses compostos

reativos (HALLIWELL, 2006; GILL & TUTEJA, 2010; MHAMDI *et al.*, 2010).

As EROs são um grupo de radicais livres, moléculas reativas e íons derivados do O₂ representados principalmente pelo o ânion superóxido (O^{2•-}), radical hidroxila (•OH), oxigênio singlete (¹O₂), e peróxido de hidrogênio (H₂O₂) (ASADA, 2006; SHARMA *et al.*, 2012). Entre as EROs, o H₂O₂ recebe atenção especial por possuir um tempo de meia vida relativamente longo (1 ms) e cujo tamanho pequeno o torna capaz de difundir-se por meio de aquaporinas nas membranas e migrar em compartimentos diferentes (NEILL, DESKIAN & HANCOCK, 2002; CHOUDHURY *et al.*, 2017). Além disso, o H₂O₂ participa da formação do •OH, o oxidante mais reativo dentre todas as EROs (KARUPPANAPADIAN *et al.*, 2011).

Nas plantas, os principais sítios de geração de EROs são apoplastos, cloroplastos, peroxissomos e mitocôndrias (ASSADA, 1999; MITTLER *et al.*, 2004; GILL & TUTEJA, 2010; CHOUDHURY *et al.*, 2017). Nos cloroplastos, a formação de EROs está relacionada com eventos da fotossíntese. Outra fonte produtora de EROs é a fotorrespiração nos peroxissomos. Nas mitocôndrias, cerca 1 a 5 % do O₂ consumido em condições de escuro é convertido em EROs (GILL & TUTEJA, 2010).

Em baixas concentrações, as EROs podem funcionar como mensageiros secundários sinalizando uma cascata de mecanismos que mediam várias respostas em plantas, incluindo o fechamento e abertura dos estômatos, morte celular programada e aquisição de tolerância aos estresses bióticos e abióticos (SHARMA *et al.*, 2012). Entretanto, a produção excessiva de EROs provoca uma desregulação na relação EROs/compostos antioxidantes e, conseqüentemente, leva a planta ao estresse oxidativo (MITTLER *et al.*, 2004; CHOUDHURY *et al.*, 2017).

Devido a sua reatividade, as EROs reagem com biomoléculas celulares tais como lipídeos, proteínas e DNA. Os efeitos danosos a essas biomoléculas resultam em alterações da fluidez de membranas, transporte de íons, perda da atividade enzimática, inibição da síntese de proteínas, desregulação transcricional, dentre outros (GILL & TUTEJA, 2010; SHARMA *et al.*, 2012; SINGH *et al.*, 2016).

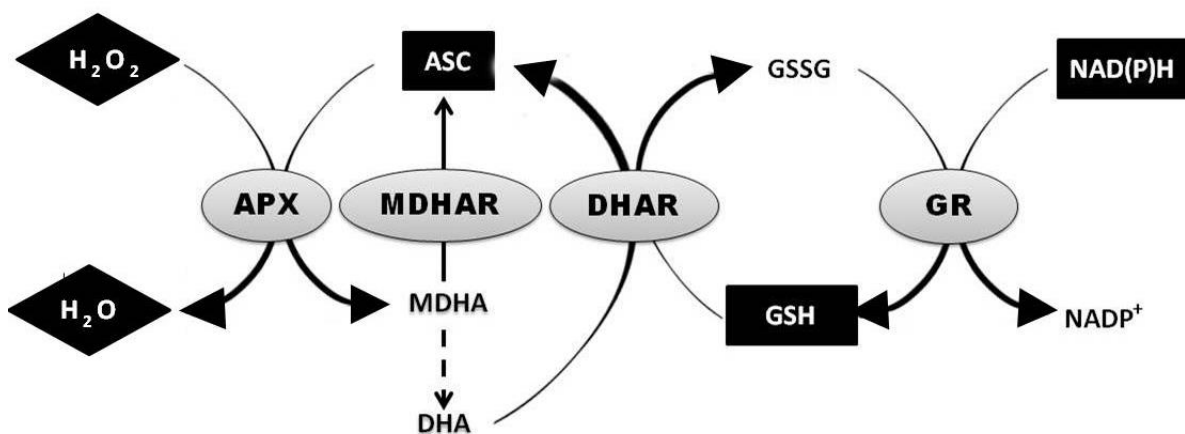
No entanto, na tentativa de neutralizar ou mitigar os efeitos negativos causados pelo excesso de EROs, as plantas regulam seus mecanismos de defesa não enzimático e enzimático. O mecanismo de defesa não enzimático inclui os principais tampões redox da célula, ascorbato e glutatona, além de compostos fenólicos, carotenoides e tocoferóis. Eles não somente protegem os componentes celulares de danos, mas também possuem um importante papel no crescimento e desenvolvimento por ajuste de processos como mitose, alongação celular, senescência e morte celular (DAS & ROYCHOUDHURY, 2014). Por sua vez, o

sistema antioxidante enzimático inclui várias enzimas tais como superóxido dismutase (SOD), catalase (CAT), guaiacol peroxidase (GPX) e enzimas do ciclo ascorbato-glutationa (Ciclo ASC-GSH) (MITTLER *et al.*, 2004; SHARMA *et al.*, 2012; PANDEY *et al.*, 2015).

2.3.2 Ciclo ascorbato-glutationa

O ciclo ascorbato-glutationa é uma via metabólica que promove a redução do peróxido de hidrogênio em água. O ciclo envolve compostos não enzimáticos como ascorbato, glutatona e NADPH e um conjunto de quatro enzimas denominadas de ascorbato peroxidase (APX), monodehidroascorbato redutase (MDHAR), dehidroascorbato redutase (DHAR) e glutatona redutase (GR). De maneira geral, o ciclo atua regenerando o ascorbato utilizado pela APX no processo de desintoxicação do H_2O_2 . O monodehidroascorbato formado a partir da catálise da APX pode ser regenerado a ascorbato pela ação da MDHAR ou espontaneamente convertido à dehidroascorbato. O DHAR reduz dehidroascorbato à ascorbato usando glutatona como agente redutor. Finalmente, a GR regenera a glutatona para manter o *pool* de glutatona reduzida utilizada tanto para o sistema redox quanto para outros sistemas (GILL & TUTEJA, 2010; FOYER & NOCTOR *et al.*, 2011; CRUZ-RUS *et al.*, 2011; VALERO *et al.*, 2016) (Figura 2).

Figura 2: Representação esquemática do ciclo ascorbato-glutationa. A linha pontilhada representa uma conversão espontânea. ASC, ascorbato reduzido; APX, ascorbato peroxidase; MDHA, monodehidroascorbato; MDHAR, monodehidroascorbato redutase; DHA, dehidroascorbato; DHAR, dehidroascorbato redutase; GSH, glutatona reduzida; GR, glutatona redutase; GSSG, glutatona oxidada.



Fonte: Adaptado de Pandey *et al.*, 2015

Estudos mostram que essas enzimas estão presentes em pelo menos quatro

compartimentos celulares: cloroplasto, citosol, mitocôndrias e peroxissomos (GILL & TUTEJA, 2010). Além disso, cada enzima é codificada por uma família multigênica que varia em quantidade ou não de uma espécie para outra e a expressão de seus diferentes membros gênicos está sob uma complexa regulação. Entretanto, os mecanismos de regulação desses membros ainda não são completamente compreendidos (PANDEY *et al.*, 2015).

2.3.2.1 Características gerais das enzimas do ciclo ASC-GSH e perfil de expressão gênica

A enzima ascorbato peroxidase (E.C.1.1.11.1) é um membro da Classe I da superfamília das hemes peroxidases sendo um componente central do ciclo ASC-GSH (DAS & ROYCHOUDHURY, 2014; SHARMA *et al.*, 2012). Uma vez que as isoformas APX apresentam alta afinidade por H₂O₂, sugere-se que as APXs estão envolvidas na modulação fina dos níveis de H₂O₂ (RIBEIRO *et al.*; 2017). Entretanto, deve-se considerar ainda a ação da APX nos cloroplastos, uma vez que a CAT (outra enzima que desintoxica o H₂O₂) não está presente nesse compartimento (PANDEY *et al.*, 2015).

Em relação as isoformas, oito enzimas APX com propriedades químicas específicas foram encontradas em *Arabidopsis* e arroz (TEIXEIRA *et al.*, 2006). Entretanto, esse número varia entre algumas espécies, por exemplo, cinco e sete APXs foram identificadas em espinafre (*Spinacia oleracea*) (CAVERZAN *et al.*, 2012) e tomate (*Solanum lycopersicum*) (NAJAMI *et al.*, 2008) respectivamente, enquanto que em *Eucalyptus grandis* foram encontrados seis isoformas (TEIXEIRA *et al.*, 2005).

Estudos relatam diferentes perfis de expressão para os genes APX sob condições de estresse. Por exemplo, em tabaco, um gene APX que codifica uma isoforma citosólica (e não a cloroplastidial) foi regulada positivamente após exposição à luz por 6h (PIGNOCCHI *et al.*, 2003). O mRNA de uma APX que codifica uma isoforma citosólica apresentou alta regulação durante seca em nódulos de alfafa (*Medicago sativa*) (NAYA *et al.*, 2007). Ainda em relação à seca, a expressão da APX1 (isoforma citosólica) aumentou significativamente em dois genótipos de trigo, entretanto, outra APX citosólica (APX2) foi regulada positivamente somente no genótipo tolerante à seca (SEČENJI *et al.*, 2010). Em arroz submetido a estresse salino, a expressão de ambas APX2 (isoforma citosólica) e APX7 (isoforma cloroplastidial) foram aumentadas enquanto que o acúmulo dos transcritos da APX8 (isoforma cloroplastidial) foi fortemente suprimido, entretanto, em altas concentrações de ABA, a APX8 foi regulada positivamente (TEIXEIRA *et al.*, 2006). O nível de transcritos da APX1 que codifica uma

isoforma citosólica da cevada (*Hordeum vulgare*) aumentou significativamente sob estresse salino (SHI *et al.*, 2001; DRABOWSKA *et al.*, 2007). Sob altas temperaturas, uma APX citosólica foi altamente induzida em folhas de batata doce (*Ipomoea batatas*) submetidas a altas temperaturas (PARK *et al.*, 2004; CAVERZAN *et al.*, 2012).

A monodehidroascorbato redutase (MDHAR; E.C. 1.6.5.4) é uma enzima flavoproteína (FAD, flavina adenina dinucleotídeo). Além da função de regenerar o MDHA, a MDHAR encontrada no cloroplasto também participa do processo de mediação da fotoredução do oxigênio (O_2) à $O_2^{\bullet-}$ quando o substrato MDHA está ausente (SHARMA *et al.*, 2012).

A MDHAR é codificada por uma pequena família multigênica. Em estudos com musgo (*Physcomitrella patens*) (LUNDE *et al.*, 2006) e tomate (*Solanum lycopersicum*) (STEVENS *et al.*, 2007), por exemplo, foram identificados três membros gênicos da MDHAR. Esses membros, por sua vez, codificam proteínas que podem estar localizadas em diferentes compartimentos celulares. Uma MDHAR citosólica foi purificada do pepino (*Cucumis sativus*) (HOSSAIN & ASADA, 1985), da soja (*Glycine max*) (DALTON *et al.*, 1993) e de tubérculos de batata (*Solanum tuberosum*) (SHARMA *et al.*, 2012). Já uma MDHAR cloroplastidial foi purificada do espinafre (*Spinacia oleracea*) (SANO *et al.*, 2005).

Estudos de expressão e superexpressão dos membros da MDHAR foram realizados em condições de estresse. Em um deles, o gene *MDHAR1* foi regulado positivamente em ervilha submetida ao frio, danos mecânicos e 2 a 4 dias de seca (LETERRIER *et al.*, 2005). Já as plantas de tabaco superexpressando o gene *MDHAR1* de *Arabidopsis thaliana* foram mais tolerantes ao estresse sob ozônio, sal e polietileno glicol (PEG) em termo de significativo aumento da taxa fotossintética (ELTAYEB *et al.*, 2007). A superexpressão de um membro que codifica uma proteína citosólica da MDHAR de acerola (*Malpighia glabra*) em plantas de tabaco resultou no aumento dos níveis de ascorbato e aumentou a tolerância da planta ao estresse salino (ELTELIB, FUJIKAWA & ESAKA, 2012).

O radical MDHA produzido na reação catalisada pela APX tem uma pequena meia vida e, se não rapidamente reduzido à ascorbato pela ação da MDHAR, converte-se em dehidroascorbato (DHA). O DHA pode ser hidrolisado irreversivelmente à ácido 2,3-dicetogulônico ou reciclado à ascorbato por meio da redutase da dehidroascorbato redutase (DHAR) (GILL & TUTEJA, 2010; SHARMA *et al.*, 2012). A DHAR (EC 1.8.5.1) catalisa a redução da DHA à ascorbato usando glutatona (GSH) como substrato redutor e, portanto, possui um importante papel na manutenção do ascorbato em sua forma reduzida (KABIR & WANG, 2010; DAS *et al.*, 2014).

Entre dois e quatro membros da *DHAR* foram encontradas em algumas espécies vegetais, tais como arroz, batata, *Arabidopsis*, choupo (*Populus trichocarpa*) e *Solanum patens* (YOSHIDA *et al.*, 2006; QIN, SHI & YU, 2011; LIU *et al.*, 2013). Dois diferentes membros foram identificados em folhas de espinafre, sendo um deles direcionado aos cloroplastos (SHIMAOKA, YOKOTA & MIYAKE, 2000). Teixeira e colaboradores (2005) identificaram dois genes para *DHAR* que codificam enzimas citosólicas e cloroplastidiais em eucalipto (*Eucalyptus* spp.).

Em relação ao perfil de expressão, uma *DHAR* que codifica uma proteína citosólica teve sua expressão aumentada em *Lotus japonicus* sob estresse salino, tornando essa espécie mais tolerante a esse estresse em relação às outras leguminosas (RUBIO *et al.*, 2009). Em *Pennisetum glaucum*, os níveis de transcritos da *DHAR* que também codifica uma proteína citosólica aumentaram significativamente em resposta ao aumento de temperatura (PANDEY *et al.*, 2014). A superexpressão de outra *DHAR* que codifica uma proteína citosólica em folhas e tubérculos de batata aumentaram o conteúdo de ascorbato nesses tecidos (QUIN, SHI & YU, 2011).

Por fim, é necessário que a glutathiona, oxidada (GSSG) pela DHAR, volte à forma reduzida (GSH) para a manutenção da alta taxa celular de GSH/GSSG. A redutase da glutathiona (GR; E.C. 1.6.4.2), uma enzima dependente de NAD(P)H, promove essa redução por meio de dois passos: (1) a fração da flavina é reduzida por NADPH, na qual a flavina é oxidada e uma ponte dissulfeto redox é reduzida para a produção do ânion tiolato e uma cisteína e (2) a redução de GSSG via reação de dupla troca com o tioldissulfeto (SHARMA *et al.*, 2012; TRIVERDI *et al.*, 2013). A GR é predominantemente encontrada em cloroplastos (nomeada de GR1 em *Arabidopsis*), mas algumas isoformas também foram encontradas em citosol (nomeada de GR2 em *Arabidopsis*). Entretanto, os transcritos que codificam as isoformas cloroplastidial e citosólica podem ser direcionados às mitocôndrias e peroxissomos, respectivamente (YOUSUF *et al.*, 2012; TRIVERDI *et al.*, 2013).

Os papéis desempenhados por ambas GR1 e GR2 em seus diferentes compartimentos foram investigados no passado e ainda continuam sendo alvo de estudos nos dias atuais (CREISSEN *et al.*, 1995; KATAYA & REUMANN, 2010; DING *et al.*, 2016). Um estudo sugere que sob excesso de luz, a GR2 de *Arabidopsis* possui um importante papel na manutenção da função de aceptor do PSII e no reparo de danos causados por excesso de luminosidade ao PSII, pois previne a acumulação de H₂O₂ (DING *et al.*, 2016). Elevados níveis da GR cloroplastidial em plantas transgênicas de tabaco aumentaram a tolerância do aparato fotossintético ao estresse causado por ozônio (BROADBENT *et al.*, 1995). Sob seca,

dois cultivares do feijão caupi apresentaram altos níveis de transcritos para a *GR* que codifica uma proteína citosólica, na qual observou-se uma maior tolerância de ambos os cultivares a esse estresse (CONTOUR-ANSEL *et al.*, 2006).

Como observado, a complexa regulação dos membros gênicos do ciclo ascorbato-glutationa é alvo de muito estudos. Entretanto, os mecanismos de regulação e ação conjunta das isoformas sob condições de estresse e desenvolvimento ainda estão longe de ser totalmente compreendidos (PANDEY *et al.*, 2015). Dessa forma, abordagens globais tais como a transcriptômica vêm sendo amplamente utilizadas para avançar na compreensão desses mecanismos regulatórios, proporcionando informações mais aprofundadas sobre a tolerância ao estresse que podem resultar em estratégias para o melhoramento de plantas.

2.4 Análise transcriptômica e estudos *in silico* de caracterização e expressão gênica

A regulação da expressão gênica é de fundamental importância para o desenvolvimento e diferenciação de células. Essa regulação se caracteriza pela abundância dos transcritos sob diferentes condições e o seu conhecimento pode promover a descoberta de novas funções genéticas e caminhos regulatórios (BUNNIK & ROCH, 2013). Nesse sentido, a análise transcriptômica ganha um importante destaque, pois permite identificar e analisar o perfil de expressão de todos os genes de uma célula ou população de células expressos em condições de desenvolvimento ou estresse, fornecendo uma visão temporal e espacial de integração do todo (MCGETTIGAN, 2013).

A primeira abordagem transcriptômica foi fornecida pela técnica de microarranjos (SCHENA *et al.*, 1995). No entanto, esta tecnologia se baseava na hibridização, tornando a análise transcriptômica restrita somente aos genes conhecidos. Além disso, a técnica de microarranjos não fornece informação da sequência e requer grande quantidade de RNA (VAN VERK *et al.*, 2013).

Entretanto, em 2008, foi publicado o primeiro estudo aplicando as tecnologias NGS no sequenciamento de transcriptoma, técnica essa denominada RNA-seq (sequenciamento de RNA) (NAGALAKSHMI *et al.*, 2008). A técnica de RNA-seq é capaz de identificar novos transcritos que eram previamente desconhecidos, diferentes isoformas, variações de *splicing* e o polimorfismo em um único nucleotídeo (SNP), promovendo uma melhor cobertura da expressão genica (DESHMUKH, 2014). Além disso, a técnica vem se tornando cada vez mais rápida e com um custo mais acessível (MCGETTIGAN, 2013).

As vantagens da técnica de RNA-seq proporcionaram o sequenciamento e montagem

do genoma bem como estudos transcriptômicos de um grande número de espécies vegetais que são depositados em bancos de dados de acesso gratuito pelo site do Centro Nacional de Informação Biotecnológica (do inglês, *National Center for Biotechnology Information*, NCBI) (<https://www.ncbi.nlm.nih.gov/>).

Diante da sua importância, a soja foi a primeira espécie leguminosa que teve o sequenciamento completo do genoma. O sequenciamento mostrou que o genoma da soja (*Glycine max* var. Williams 82) é um dos maiores entre os organismos vegetais já sequenciados, pois possui 1,1 Gb distribuídos em 20 cromossomos. Mais ainda, foram preditos 46.430 genes que codificam proteínas, 70% a mais do que em *Arabidopsis*. O número maior de genes registrado para *G. max* é devido a eventos de duplicações no genoma que ocorreram em aproximadamente 59 e 13 milhões de anos atrás que resultou em cópias múltiplas em cerca de 75% dos genes (SCHMUTZ *et al.*, 2010).

Atualmente, aproximadamente 200 estudos transcriptômicos estão disponíveis no NCBI relacionados à soja. Entretanto, ainda são poucos os trabalhos que abordam estudos *in silico* de expressão e caracterização de genes em soja e outras leguminosas. Dessa forma, a exploração desses dados *in silico* para estudos de expressão gênica pode ser utilizada na tentativa de descobrir genes responsivos às adversas condições ambientais e estádios de desenvolvimento em soja visando cultivares melhores adaptados às limitações do ambiente.

3 OBJETIVOS

3.1 Objetivo geral

Caracterizar, analisar filogeneticamente e avaliar o perfil de expressão dos genes das famílias multigênicas APX, MDHAR, DHAR e GR em soja (*Glycine max*) submetida a diferentes condições de estresse e em diferentes tecidos em desenvolvimento.

3.2 Objetivos específicos

- Identificar todos os membros gênicos das famílias APX, MDHAR, DHAR e GR de soja (*Glycine max*) bem como em outras espécies leguminosas como *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Glycine soja*, *Medicago truncatula*, *Phaseolus vulgaris*, *Vigna angularis* e *Vigna radiata*;
- Analisar filogeneticamente as famílias multigênicas APX, MDHAR, DHAR e GR nas espécies leguminosas citadas;
- Analisar o perfil de expressão dos genes do ciclo ascorbato-glutationa em soja submetida a estresses e em desenvolvimento;
- Identificar genes alvos para o desenvolvimento de cultivares melhores adaptados às condições de estresses.

4 MATERIAL E MÉTODOS

4.1 Identificação e caracterização dos genes da APX, MDHAR, DHAR e GR em *Glycine max* e outras espécies leguminosas

A identificação e a anotação dos genes do ciclo ascorbato-glutationa foram realizadas por meio de buscas através da ferramenta BLAST (ALTSCHUL *et al.*, 1997) nos genomas de *Glycine max*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Glycine soja*, *Medicago truncatula*, *Phaseolus vulgaris*, *Vigna angularis* e *Vigna radiata* disponíveis em bancos de dados do *GenBank* (NCBI, *National Center for Biotechnology Information*) (<http://www.ncbi.nlm.nih.gov>). As proteínas de *Arabidopsis thaliana* foram utilizadas como sequências iscas para localizar genes homólogos nas espécies anteriormente citadas.

A anotação do gene e dedução do cDNA (determinação de éxons/íntrons e quadro de leitura aberta) foi realizado manualmente contra os cDNAs da APX, MDHAR, DHAR e GR disponíveis no *GenBank* (NCBI) através de pesquisas por meio da ferramenta BLASTn. O cDNA deduzido foi traduzido em sequências de aminoácidos com o auxílio da ferramenta online ExPASy (<http://web.expasy.org/translate>). A verificação da homologia e da extensão completa das sequências de proteínas deduzidas foi confirmada através de homólogas depositadas no *GenBank* (não redundante) usando a ferramenta BLASTp. A estratégia de anotação (dedução do cDNA e proteína) empregada foi a mesma realizada por Costa e colaboradores 2014.

4.2 Alinhamento das sequências e análise filogenética

As análises filogenéticas das sequências de proteínas deduzidas foram conduzidas para cada família multigênica anotada em *Glycine max* e em outras nove espécies leguminosas. As sequências homólogas de *Arabidopsis thaliana* (Brassicales) e *Volvox carteri* (alga verde) foram incluídas como grupo externo. Todas as sequências foram alinhadas através do programa CLUSTAL W e a árvore filogenética foi construída com o programa MEGA 7.0 (KUMAR *et al.*, 2016) usando o método de agrupamento *Neighbor-joining* (SAITOU & NEI, 1987) com valores de *bootstrap* gerados a partir de 1000 replicatas (TAMURA *et al.*, 2011). Todas as sequências de proteínas utilizadas para análise encontram-se disponíveis no ANEXO A.

4.3 Configuração éxon/íntron dos genes da APX, MDHAR, DHAR e GR em *Arabidopsis thaliana* e *Glycine max*

A disposição dos éxons e íntrons de cada gene da APX, MDHAR, DHAR e GR deduzido para soja e *Arabidopsis* foi realizada com o uso da ferramenta *online Gene Structure Display Server* (GSDS 2.0) (HU *et al.*, 2015) através do alinhamento das sequências do cDNA com suas correspondentes sequências genômicas. Em seguida, as estruturas dos genes de *A. thaliana* foram utilizadas como referências para a identificação de inserção de íntrons em ortólogos de *G. max*.

4.4 Predição da localização subcelular e análise de domínios das enzimas do ciclo ascorbato-glutationa em *Glycine max*

A localização subcelular das proteínas deduzidas foi determinada por meio de três passos: (1) acesso a três servidores online distintos: *TargetP 1.1* (<http://www.cbs.dtu.dk/services/TargetP/>), *CELLO V.2.5* (<http://cello.life.nctu.edu.tw/>) e *WoLF PSORT* (<http://www.genscript.com/wolf-psort.html>). (EMANUELSSON *et al.*, 2000; YU, LIN & HWANG, 2004 ; YU *et al.*, 2006); (2) cada sequência de proteína foi lançada na ferramenta BLASTp (NCBI) para comparação com aquelas ortólogas encontradas no banco de dados (*non-redundant protein sequence*) que apresentavam identificação da localização subcelular e (3) por meio de artigos científicos com descrição da localização subcelular de cada proteína da APX, MDHAR, DHAR e GR em *Arabidopsis thaliana*, ortólogas àquelas de *Glycine max*.

As análises dos domínios conservados entre isoformas foram realizadas pela busca do modelo *Hidden Markov* (HMM) através do servidor *online* PFAM (<http://pfam.xfam.org/search/sequence>) (FINN *et al.*, 2016).

4.5 Análise das regiões promotoras

A presença de possíveis elementos *cis* regulatórios responsivos à estresse [ABRE (ACGTG),GT1 (GGTTAA/GAAAAA), MYB (CTAACCA/CNGTTR/WAACCA/TAAC TG), LTRE (CCGAAA), BBF1 (ACTTTA) e DRE (GTCGAC)] na região *upstream* de 1000 pares de base do códon de início da tradução foi analisada por meio de dados disponíveis nos seguintes servidores online: New Place (<https://sogo.dna.affrc.go.jp>) (HIGO *et al.*, 1999), PlantPAN (<http://plantpan.mbc.nctu.edu.tw/>) (CHANG *et al.*, 2008), PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/p>) (LESCOT *et al.*, 2002), e AtPAN

(<http://atpan.itps.ncku.edu.tw/>) (CHEN *et al.*, 2012).

4.6 Análises de expressão *in silico* utilizando dados de RNA-seq para *Glycine max*

As análises de expressão dos genes do ciclo ascorbato-glutationa foram geradas a partir da exploração de bioprojetos disponíveis no banco de dados SRA (*Sequence Read Archive*) do *GenBank*. No geral, foram selecionados experimentos com triplicatas biológicas cobrindo o transcriptoma de diferentes cultivares de soja durante estágios de desenvolvimento (embrião, raiz, cotilédone e semente) e estresses (seca, etileno, NaCl, patógeno). As condições experimentais de cada bioprojeto explorado estão detalhados na Tabela 3.

As sondas representativas da região UTR 5' e 3' dos cDNAs de cada membro gênico da APX, MDHAR, DHAR e GR foram desenhadas a partir do cDNA deduzido a fim de evitar o alinhamento múltiplo e, conseqüentemente, fornecer um perfil de expressão preciso dos genes parálogos que apresentam alta identidade. Posteriormente, as sondas das regiões UTR de 5' e 3' foram definidas como sequências de referência para mapear cada transcriptoma (cada tratamento com dezenas de milhões de *reads* entre 50 a 150 pares de base) usando os programas Bowtie2 v2.3.2 (LANGMEAD & SALZBERG, 2012) e TopHat2 v2.0.13 (<http://tophat.cbcb.umd.edu/>) (KIM *et al.*, 2013). Somente os *reads* mapeados foram fornecidos ao pacote Cufflinks (<http://cufflinks.cbcb.umd.edu/>) (TRAPNELL & SALZBERG, 2009) para a contagem do número de *reads* mapeados por transcrito e inferir o perfil de expressão. Somente os genes que apresentaram $\log_2 \textit{fold-change} \leq -1$ ou $\geq +1$, com significância estatística (p-value <0.05), foram considerados como diferencialmente expressos.

Tabela 3: Experimentos de RNA-seq utilizados para obter o perfil de expressão dos genes *APX*, *MDHAR*, *DHAR* e *GR* em *Glycine max*

Acesso do Bioprojeto	Cultivar	Tecido	Estágio da planta	Tratamentos	Detalhamento experimental	Replicatas	Referência
PRJNA246058	Williams 82	Raízes	V1	Controle, sal e desidratação	Raízes foram coletadas em 0, 1, 6 e 12h de tratamento com sal (100 mM NaCl) ou desidratação (raízes expostas ao ar).	Três	BELAMKAR <i>et al.</i> , 2014
PRJNA259941	Benning e PI416937	Folhas	R2	Controle e desidratação	Folhas expostas ao ar foram coletadas depois de 0, 6, 12 e 24h em ambos os cultivares Benning (sensível à seca) e PI416937 (tolerante a sal)	Três	SHIN <i>et al.</i> , 2015
PRJNA324522	Williams 82	Folhas	V4	Controle e seca	Estresse foi imposto por interrupção da irrigação por sete dias.	Três	CHEN <i>et al.</i> , 2016
PRJNA266538	Williams 82	Zona de abscisão foliar e pecíolos	V1	Controle e etileno	Explantos com plantas provenientes de 21 dias foram expostos a 25 μ L/L de etileno de 0 a 72 h.	Três	KIM <i>et al.</i> , 2015
PRJNA278121	Forrest	Raízes	-	Controle, FO36 e FO40	Raízes foram infectadas com <i>Fusarium oxysporum</i> FO36 (não patogênico) e FO40 (patogênico). Em seguida, as raízes foram coletadas em 72 horas após a inoculação (hpi)	Três	LANUBILE <i>et al.</i> , 2015

Tabela 3: Continuação

Acesso do Bioprojeto	Cultivar	Tecido	Estágio da planta	Tratamentos	Detalhamento experimental	Replicatas	Referências
PRJNA262564	Williams 82	Cotilédones		Desenvolvimento dos cotilédones	Cotilédones foram coletados em 4, 15 e 27 dias após a semeadura.	Três	BROWN & HUDSON, 2015
PRJNA388955	Williams 82	Embrião	-	Quatro estágios de desenvolvimento do embrião	Embriões coletados no estágio cotiledonar, início, metade e final da maturação	Três	-
PRJNA271595	Williams 82	Raízes	-	Zona meristemática (ZM, divisão celular e formação da coifa da raiz), zona de alongação (ZE; expansão celular) e zona de diferenciação (ZD; maturação celular)	O segmento ZM representa a porção terminal da raiz na qual cortou-se na posição em que o comprimento das células começou a exceder sua largura. O segmento ZD inclui o primeiro cabelo da raiz iniciado até o ponto em que os cabelos da raiz atingem o comprimento total. O segmento ZE representa a parte da raiz entre o ZM e ZD.	Três	HUANG & SCHIEFELBEIN, 2015

Fonte: Elaborado pelo autor

4.7 Análise estatística

O *software* GraphPad Prism 5.0 foi utilizado para realizar a análise estatística dos dados oriundos de experimentos de RNA-seq. Os resultados foram expressos com a média +/- erro padrão das três replicatas biológicas. As diferenças em variáveis foram analisadas por *one-way ANOVA*, seguido pelo teste de Bonferroni. As diferenças foram consideradas significantes em $p < 0,05$.

5 RESULTADOS

5.1 Identificação e análise filogenética dos genes *APX*, *MDHAR*, *DHAR* e *GR* em *Glycine max* e outras espécies leguminosas

Um total de 24 genes (11 *APX*, 5 *MDHAR*, 4 *DHAR* e 4 *GR*) que codificam as enzimas do ciclo ASC-GSH foram identificados em ambos os genomas de *Glycine max* e *Glycine soja* (Tabela 4), espécies pertencentes à tribo Phaseoleae. Em outras plantas da tribo Phaseoleae tais como *Cajanus cajan*, *Phaseolus vulgaris*, *Vigna radiata* e *Vigna angularis*, um menor número de genes, entre 14 e 15, foram identificados. Um padrão similar também foi encontrado em espécies das tribos Cicereae, Trifolieae and Dalbergieae, nas quais foram identificados um máximo de 15 genes (Tabela 4).

Tabela 4: Número de genes da *APX*, *MDHAR*, *DHAR* e *GR* para espécies leguminosas

Espécies leguminosas	Tribo	Número de genes				Total
		<i>APX</i>	<i>MDHAR</i>	<i>DHAR</i>	<i>GR</i>	
<i>A. duranensis</i>	Dalbergieae	7	3	2	3	15
<i>A. Ipaensis</i>		6	3	2	3	14
<i>M. truncatula</i>	Trifolieae	5	3	2	2	12
<i>C. arietinum</i>	Cicereae	6	3	2	2	13
<i>C. Cajan</i>		8	3	2	2	15
<i>G. max</i>		11	5	4	4	24
<i>G. soja</i>	Phaseoleae	11	5	4	4	24
<i>P. vulgaris</i>		7	3	2	2	14
<i>V. angularis</i>		7	3	2	2	14
<i>V. radiata</i>		7	3	2	2	14

Fonte: Elaborado pelo autor

Todas as proteínas identificadas em espécies leguminosas foram classificadas em relação às proteínas homólogas de *A. thaliana* utilizando a *V. carteri* (alga verde) como grupo externo (Figura 3). As análises filogenéticas para APX (Figura 3A), MDHAR (Figura 3B), DHAR (Figura 3C) e GR (Figura 3D) revelaram a necessidade para reclassificar algumas proteínas de *Arabidopsis* em relação às suas classificações do *GenBank*. As proteínas renomeadas para *G. max* estão destacadas nas respectivas figuras. Ambos os genomas de *G. max* e *G. soja* comumente revelaram genes que codificam proteínas altamente similares (e.g. APX-1a e 1b; 5a1 e 5a2; 5b1 e 5b2; MDHAR-2a e 2b; DHAR-1a e 1b; GR-1a e 1b) (Figura 3).

Oito proteínas APX de *A. thaliana* foram distribuídas em cinco clados e denominadas como At-APX1 (clado 1), At-APX2 (clado 2), At-APX3a-b (clado 3), At-APX4a-b (clado 4) e At-APX5a-b (clado 5) (Figura 3A). Em acordo com essa distribuição, as proteínas APX nas espécies leguminosas foram classificadas revelando que a família multigênica da APX em *G. max* e *G. soja* é constituída por APX-1a, 1b, 2, 3a, 3b, 4a, 4b, 5a1, 5a2, 5b1 e 5b2 (Figura 3A). Em geral, a maior quantidade de proteínas APX encontradas em ambas *G. max* e *G. soja* em relação às outras leguminosas foi principalmente devido às duplicações encontradas em APX-1, 3, 4 e 5a. Essas duplicações também foram encontradas em APX-3, 4 e 5 de *A. thaliana*. Interessantemente, a APX-5a foi detectada somente em *A. thaliana* e espécies da tribo Phaseoleae, indicando que esse gene foi perdido em outras espécies leguminosas estudadas aqui (Figura 3A).

Com relação à MDHAR, as cinco proteínas anotadas para *A. thaliana* foram distribuídas em três clados e classificadas como At-MDHAR-1 (clado 1), At-MDHAR-2 (clado 2), e At-MDHAR-3a-b1-b2 (clado 3) (Figura 3B). Após essa classificação, as respectivas proteínas ortólogas MDHAR em *G. max* foram nomeadas como Gm-MDHAR-1, 2a, 2b, 3a e 3b (Figura 3B). As *G. max* e *G. soja* apresentaram a maior quantidade de proteínas MDHAR em relação às outras leguminosas devido às duplicações encontradas em MDHAR-2 e 3. Além disso, cópias similares em MDHAR-3 também foram encontradas para *A. thaliana* (Figura 3B).

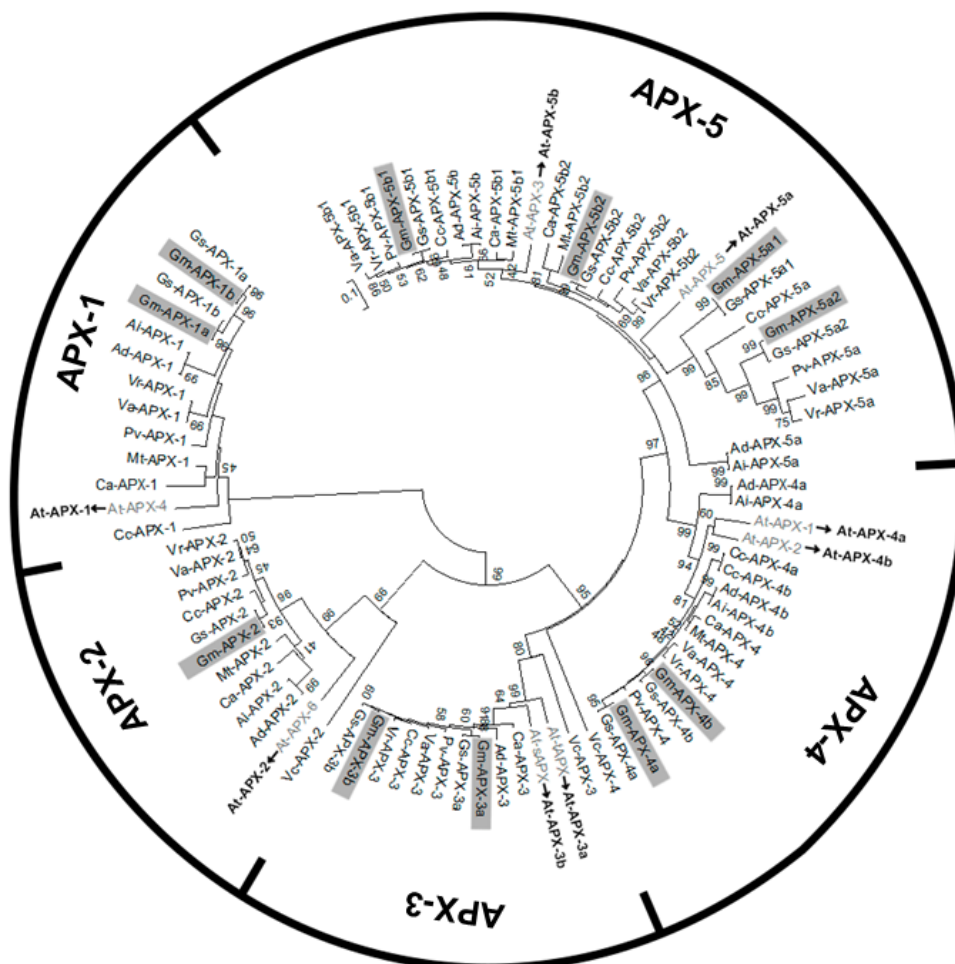
Para DHAR, a análise filogenética revelou dois clados distintos fortemente apoiados pelos altos valores de *bootstrap* (Figura 3C). Três proteínas DHAR de *A. thaliana* foram distribuídas em clado 1 (At-DHAR-1a-b) e clado 2 (At-DHAR-2). Como resultado, proteínas DHAR ortólogas em *G. max* e *G. soja* foram classificadas como DHAR-1a, 1b, 2a e 2b (Figura 3C). Duplicações em *G. max* e *G. soja* foram encontradas em DHAR-1 e 2, entretanto

em *Arabidopsis*, duplicações de proteínas DHAR foram identificadas apenas em DHAR-2 (Figura 3C).

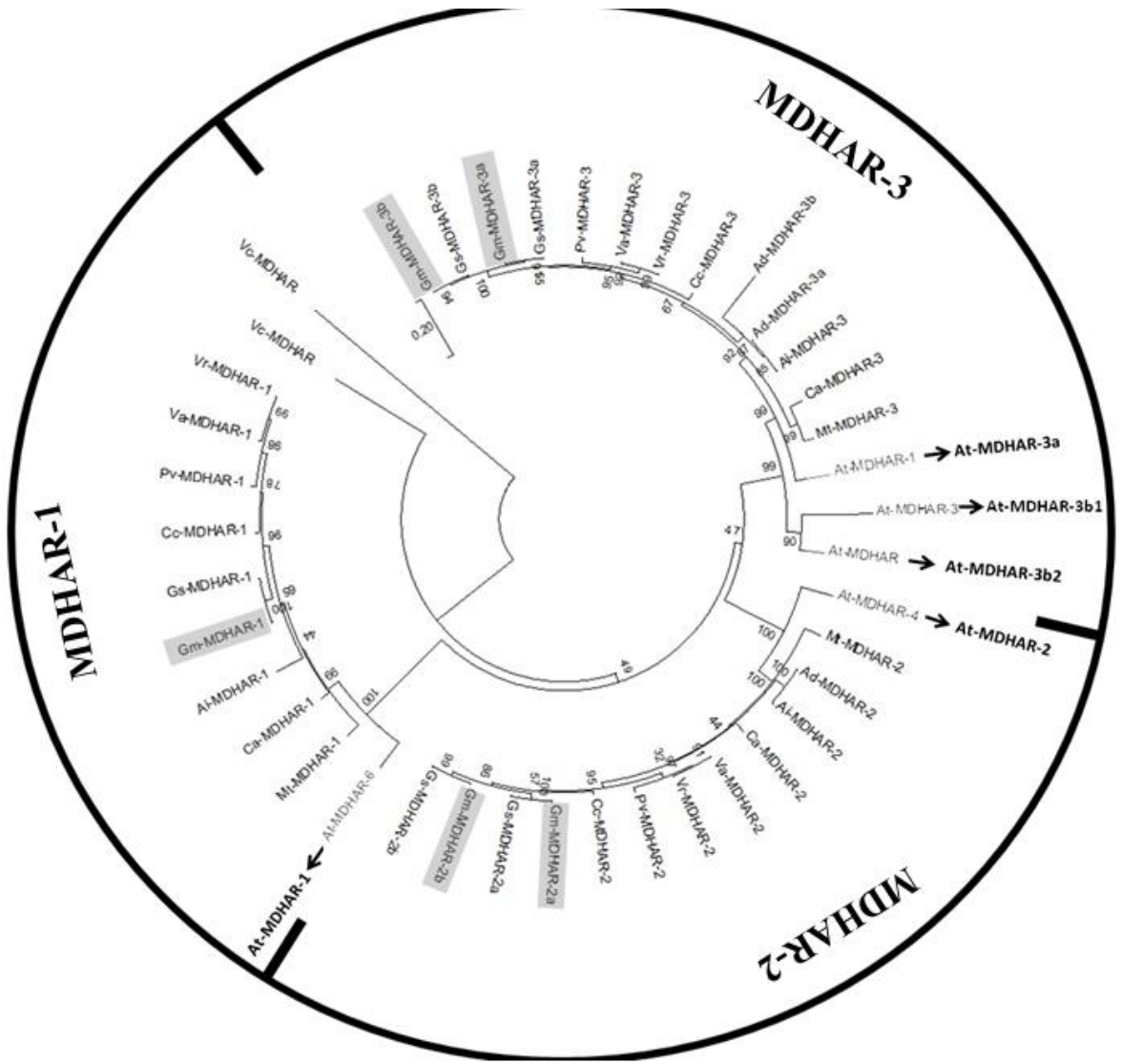
Finalmente, para GR, as proteínas de *A. thaliana* também se separaram em dois clados distintos, nomeados como At-GR-1 (clado 1) e At-GR-2 (clado 2). Então, de acordo com esta distribuição filogenética, as quatro proteínas GR em *G. max* foram classificadas como Gm-GR-1a, 1b, 2a e 2b (Figura 3D). Cópias similares encontradas em GR-1 e 2 para ambas *G. max* e *G. soja* revelam a maior quantidade das proteínas GR para essas espécies em relação às outras leguminosas e *A. thaliana* (Figura 3D).

Figura 3: Árvores filogenéticas das proteínas APX (A), MDHAR (B), DHAR (C) e GR (D) de diferentes espécies leguminosas (*Glycine max*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum*, *Glycine soja*, *Medicago truncatula*, *Phaseolus vulgaris*, *Vigna angularis* e *Vigna radiata*), bem como *A. thaliana* e *Volvox carteri*. As caixas cinzas representam as proteínas classificadas para *G. max*. Em *Arabidopsis*, as setas partem da classificação encontrada no *GenBank* para a nova classificação (em negrito). A árvore foi obtida pelo método de agrupamento *Neighbour-joining* (SAITOU & NEI, 1987) com valores de *bootstrap* de 1000 replicatas usando o programa MEGA 7 (KUMAR *et al.*, 2016).

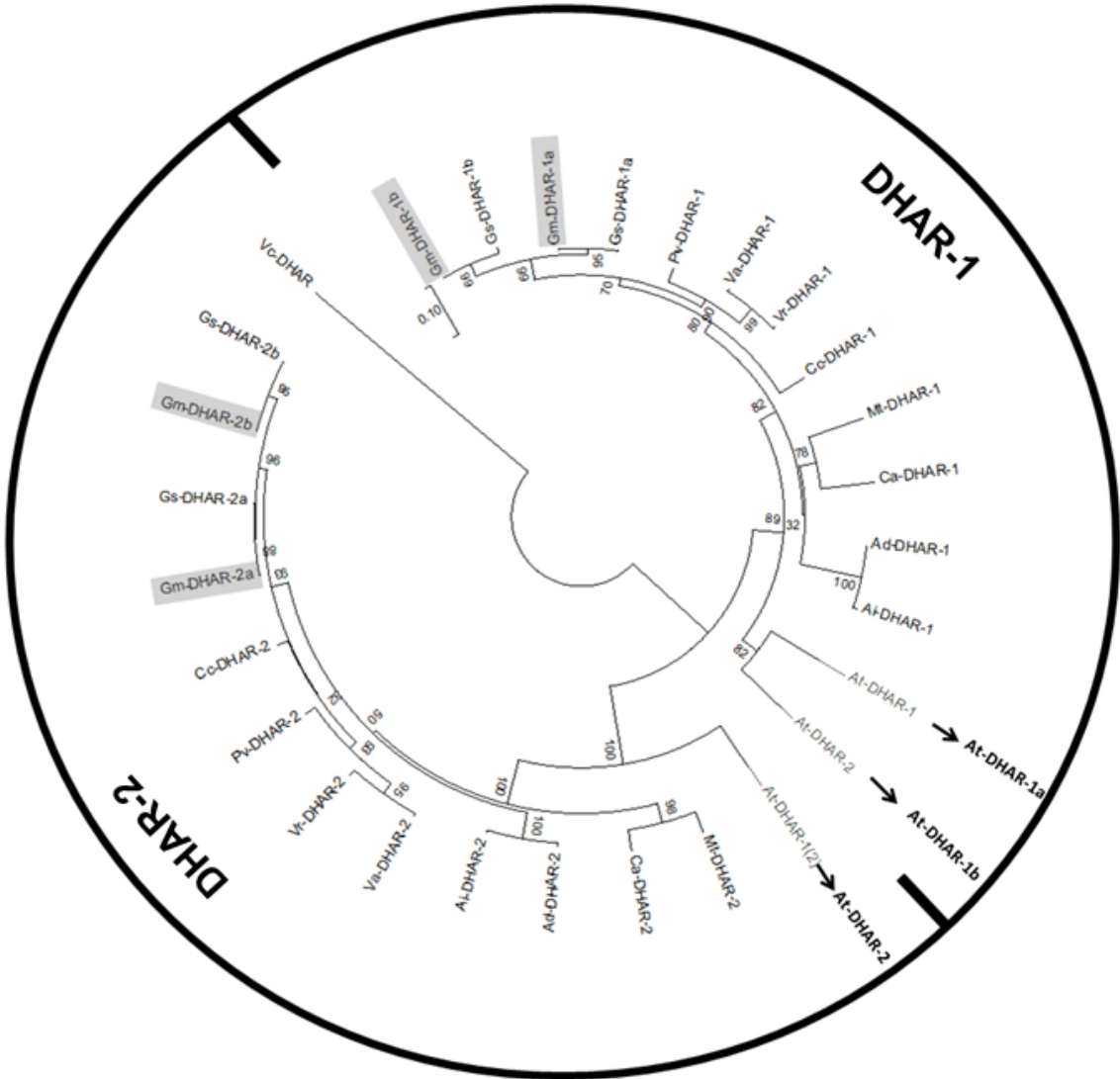
A



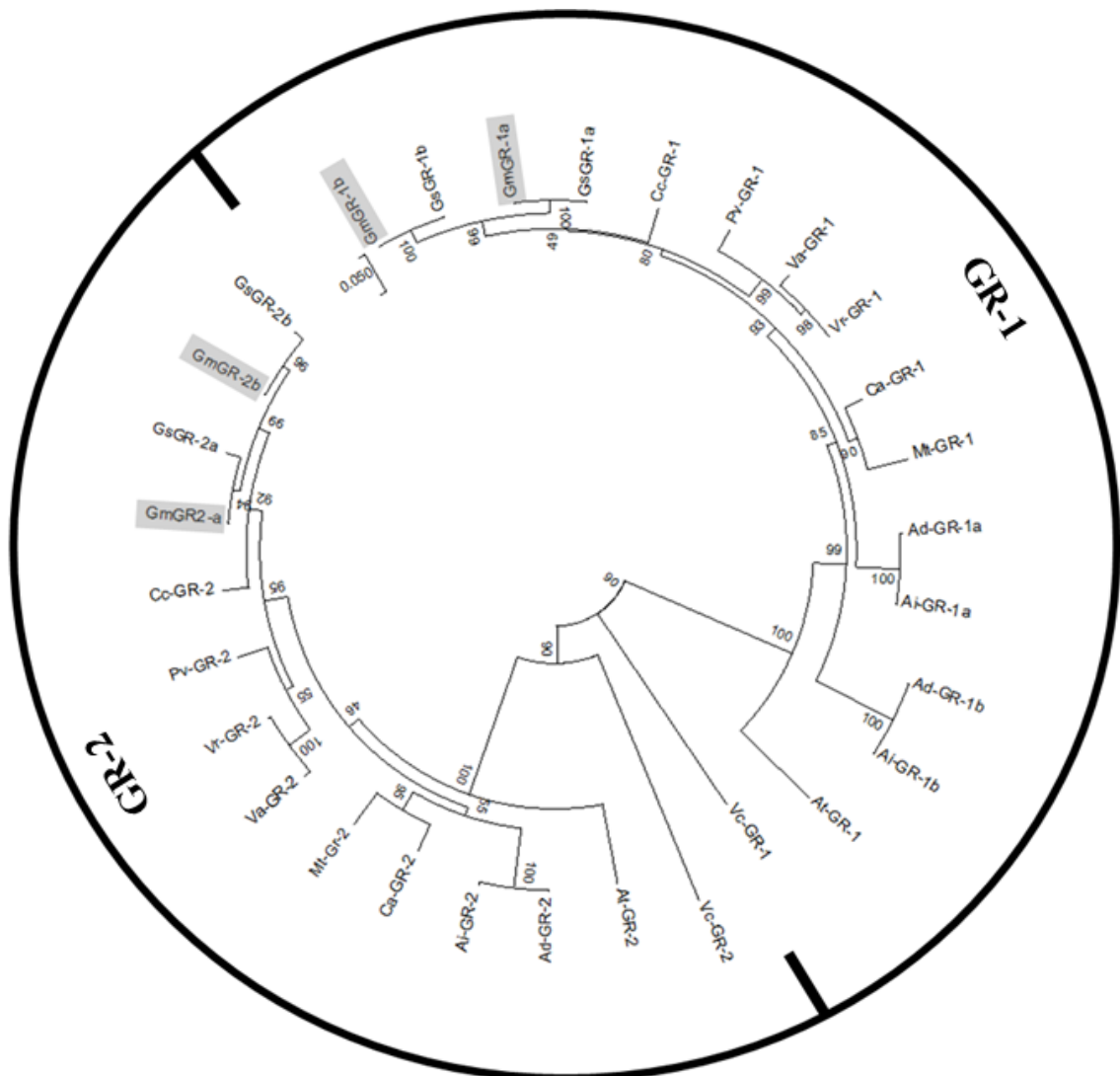
B



C



D



Fonte: Elaborado pelo autor

5.2 Identidade das proteínas e conservação do domínio proteico em parálogos das famílias APX, MDHAR, DHAR e GR de *Glycine max*

As análises das sequências das enzimas APX, MDHAR, DHAR e GR revelaram características específicas para genes parálogos em cada família, tais como altas e baixas variações de identidade entre as sequências de aminoácidos e conservados domínios proteicos.

No geral, a maioria das cópias similares dentro de cada clado observadas anteriormente para *G. max* apresentaram elevados valores de identidade e, à medida que se comparavam parálogos mais distantes evolutivamente, a identidade diminuía consideravelmente (Tabela 5, 6, 7 e 8).

Para APX, a maioria das cópias similares apresentaram 95% de identidade entre si (e.g. Gm-APX-1a e 1b; Gm-APX-3a e 3b; Gm-APX-4a e 4b), entretanto, outros parálogos como Gm-APX-3a e Gm-APX-1a apresentaram baixa identidade (até 27 %) (Tabela 5). Em relação a MDHAR, as identidades das proteínas variaram de 99% (Gm-MDHAR-3a e 3b) a 44% (Gm-MDHAR-1 e Gm-MDHAR-2a) (Tabela 6). A identidade das proteínas DHAR variou de 96% (Gm-DHAR-1a e 1b) a 61% (Gm-DHAR-1b e 2b) (Tabela 7). Para GR, a identidade das proteínas variou de 97% (Gm-GR-1a e 1b; Gm-GR-2a e 2b) a 53 % (e.g. Gm-GR-2b e Gm-GR-1b) (Tabela 8).

Buscas pela plataforma *online* Pfam foram realizadas para identificar domínios conservados e seus tamanhos em cada sequência de aminoácidos das enzimas APX, MDHAR, DHAR e GR. Para todas as famílias multigênicas, isoformas altamente similares, exceto Gm-APX-5a1 e Gm-APX-5a2, apresentaram a mesma quantidade de aminoácidos para o domínio (Anexo B).

Em relação à APX, todas as proteínas apresentaram um único domínio peroxidase (PF00141) com variação de 175 (Gm-APX-1a e 1b) a 232 aminoácidos (Gm-APX-3a e 3b). Já as proteínas MDHAR apresentaram o domínio oxidoreductase de piridina nucleotídeo dissulfeto (PF07992) com 315 (Gm-MDHAR-1, 2a e 2b) ou 316 aminoácidos (Gm-MDHAR-3a e 3b). Dois domínios [Glutathione S-transferase, N-terminal (PF13417) (69 aa) e C-terminal (PF13410) (63 aa)] foram identificados para Gm-DHAR-1a e 1b, enquanto que Gm-DHAR-2a e 2b apresentaram somente o domínio Glutathione S-transferase, N-terminal (PF13417) que contém 70 aa. As proteínas GR também apresentaram dois domínios: oxidoreductase da piridina nucleotídeo-dissulfeto (PF07992) e domínio de dimerização (PF02852) (Anexo B).

Tabela 5: Identidade entre as sequências de aminoácidos de APX em *G. max* utilizando alinhamento pelo BLASTp (%)

	APX-1a	APX-1b	APX-2	APX-3a	APX-3b	APX-4a	APX-4b	APX-5a1	APX-5a2	APX-5b1	APX-5b2
APX-1a	-	95	33	27	27	32	29	42	43	32	33
APX-1b	95	-	34	28	28	33	28	41	43	33	32
APX-2	33	33	-	35	35	38	38	35	28	38	38
APX-3a	27	27	35	-	95	44	46	41	37	50	47
APX-3b	27	27	35	94	-	45	46	41	37	51	46
APX-4a	32	33	38	44	45	-	95	49	41	64	63
APX-4b	28	28	38	46	46	95	-	51	43	64	64
APX-5a1	45	43	35	41	41	49	51	-	58	60	59
APX-5a2	37	37	28	37	37	41	43	58	-	50	51
APX-5b1	32	33	38	50	51	64	64	54	46	-	73
APX-5b2	32	31	38	46	46	63	64	59	51	73	-

Fonte: Elaborado pelo autor

Tabela 6: Identidade entre as sequências de MDHAR em *G. max* utilizando o alinhamento pelo BLASTp (%)

	MDHAR-1	MDHAR-2a	MDHAR-2b	MDHAR-3a	MDHAR-3b
MDHAR-1	-	44	45	46	46
MDHAR-2a	44	-	95	55	55
MDHAR-2b	45	95	-	55	55
MDHAR-3a	46	55	55	-	99
MDHAR-3b	46	55	55	99	-

Fonte: Elaborado pelo autor

Tabela 7: Identidade entre as sequências de DHAR em *G. max* utilizando o alinhamento pelo BLASTp (%)

	DHAR-1a	DHAR-1b	DHAR-2a	DHAR-2b
DHAR-1a	-	96	63	63
DHAR-1b	96	-	62	61
DHAR-2a	63	62	-	95
DHAR-2b	63	61	95	-

Fonte: Elaborado pelo autor

Tabela 8: Identidade entre as sequências de GR em *G. max* utilizando o alinhamento pelo BLASTp (%)

	GR-1a	GR-1b	GR-2a	GR-2b
GR-1a	-	97	53	53
GR-1b	97	-	53	53
GR-2a	53	52	-	97
GR-2b	53	53	97	-

Fonte: Elaborado pelo autor

5.3 Caracterização estrutural dos membros gênicos da APX, MDHAR, DHAR e GR em *Glycine max*

O quadro de leitura aberta (ORF, *open reading frame*), bem como a distribuição éxon/íntron dos genes de *G. max* foram determinados e comparados com as respectivas estruturas de genes ortólogos de Arabidopsis (Figuras 4A, B, C, D). No geral, os genes de *G. max* são maiores em tamanho que aqueles de Arabidopsis, principalmente devido ao maior tamanho dos íntrons. Entretanto, em alguns casos, a extensão do tamanho dos genes de *G. max* ocorreu por meio de eventos de ganho de íntrons.

Para APX, o tamanho do ORF dos genes em Arabidopsis variou de 753 pb (*At-APX-4a*) a 1.281 pb (*At-APX-3a*). Em *G. max*, *Gm-APX-4a* (ortólogo à *At-APX-4a*) apresentou o menor tamanho do ORF (753 pb) enquanto que o ORF de *Gm-APX-3b* foi a maior (1.161 pb) (Dados não mostrados). Com relação à estrutura éxon/íntron, inserções de íntron em *Gm-APX-1a-b*, *Gm-APX-3b* e *Gm-APX-4a-b* foram identificadas quando comparado aos genes de Arabidopsis (Figura 4A). Além disso, *Gm-APX-5a2* apresentou um evento de deleção que resultou em uma menor quantidade de éxons em comparação ao parólogo *Gm-APX-5a1* e ao ortólogo *At-APX-5a* (Figura 4A).

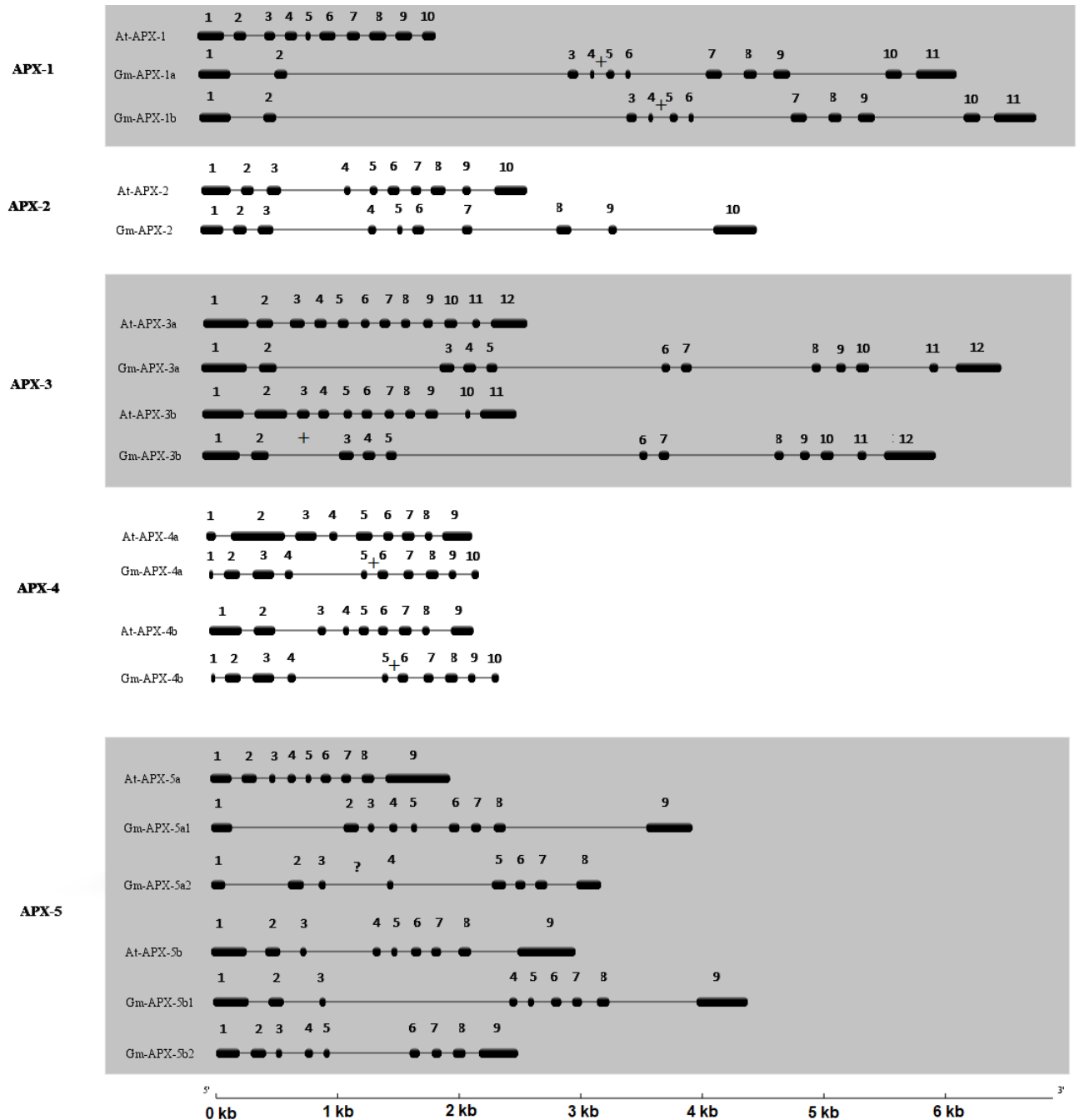
O tamanho do ORF dos membros gênicos de MDHAR em Arabidopsis variou de 1.326 pb (*At-MDHAR-3b1*) à 1.467 pb (*At-MDHAR-2*), já em *G. max*, a variação foi de 1.302 pb (*Gm-MDHAR-3a* e *3b*) à 1.467 pb (*Gm-MDHAR-1*). Em adição, inserções de íntrons foram encontradas em *Gm-MDHAR-1* e *Gm-MDHAR-3b* quando comparados aos respectivos ortólogos em Arabidopsis (Figura 4B) e em *At-MDHAR-3b2* em relação ao seu parólogo *At-MDHAR-3b1* (Figura 4B).

Em relação à DHAR, o tamanho do ORF em genes de Arabidopsis variou de 642 pb (*At-DHAR-1a* e *1b*) a 777 pb (*At-DHAR-2*). Por outro lado, os genes *Gm-DHAR-1a* e *1b* também possuem o ORF com 642 pb, enquanto que o ORF de *Gm-DHAR-2a* tem 786 pb (Dados não mostrados). A inserção de três íntrons em *Gm-DHAR-1a* e *1b* foram identificadas em comparação aos ortólogos *At-DHAR-1a* e *1b*, respectivamente (Figura 4C).

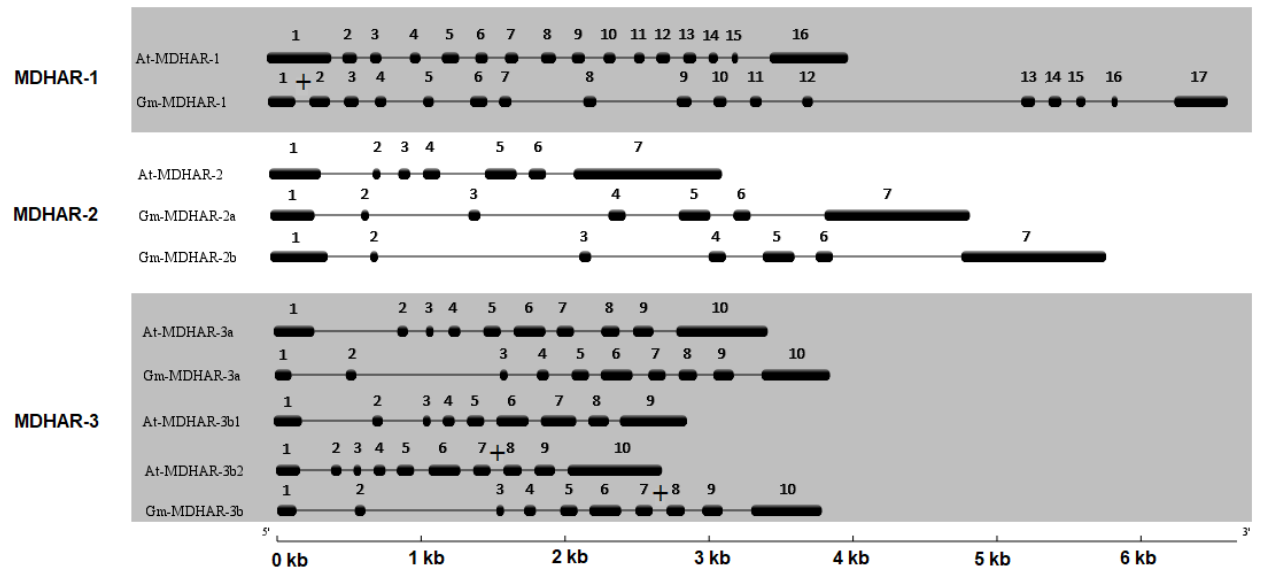
Finalmente, a análise do ORF em genes GR de Arabidopsis revelaram que *At-GR-1* e *At-GR-2* apresentam 1.500 e 1.698 pb, respectivamente. Em *G. max*, o tamanho do ORF variou de 1.441 pb (*Gm-GR-1b*) a 1.638 pb (*Gm-GR-2a* e *2b*) (Dados não mostrados). Interessantemente, nenhum evento de inserção de íntrons foi identificado em genes GR de *G. max* em relação àqueles de Arabidopsis (Figura 4D).

Figura 4: Configurações éxon/íntron dos genes *APX* (A), *MDHAR* (B), *DHAR* (C) e *GR* (D) em *G. max*. Os éxons são representados por caixas pretas e os íntrons por linhas finas. O símbolo + representa a inserção de íntron.

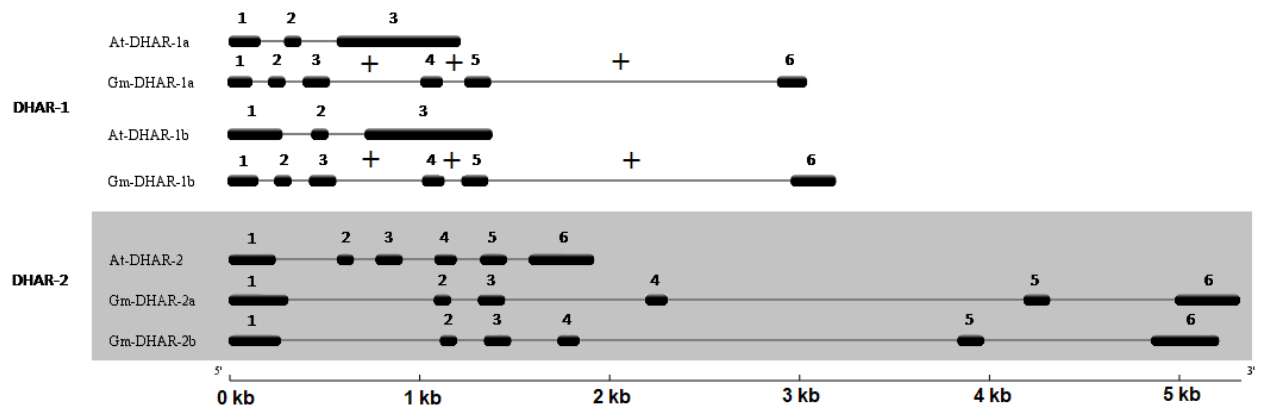
A



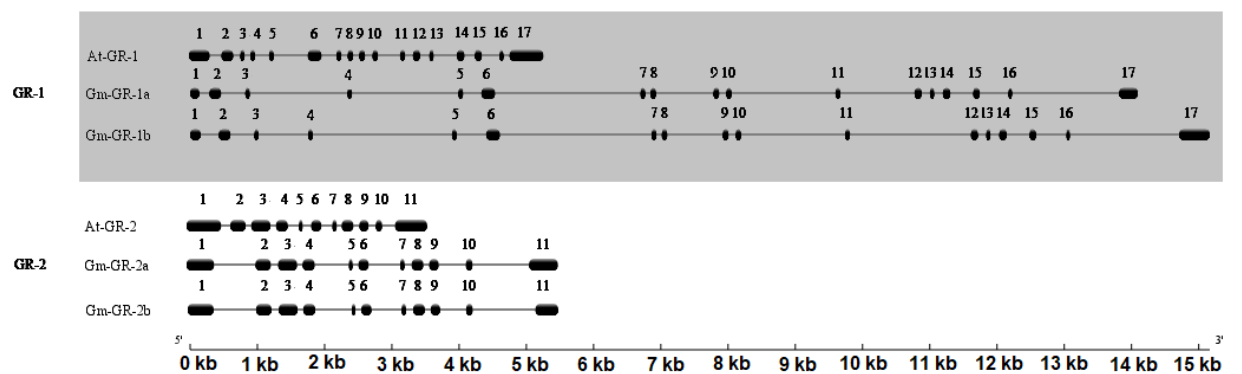
B



C



D



Fonte: Elaborado pelo autor

5.4 Localização subcelular das proteínas APX, MDHAR, DHAR e GR em *Glycine max*

A predição da localização subcelular das proteínas do ciclo ascorbato-glutationa em *G. max* foi realizada primeiramente a partir de três servidores *onlines* (Tabela 9). No geral, as proteínas se encontram no citosol e em cloroplastos. Entretanto, a busca por proteínas homólogas no *GenBank*, bem como por publicações científicas que predizem a localização de ortólogos em *Arabidopsis* (organismo modelo) revelam outras localizações subcelulares para algumas isoformas (peroxissomos, mitocôndrias e tilacóides) (Tabela 9).

Em relação à predição da localização para APX, todas as proteínas pertencentes à Gm-APX-1, APX-2 e APX-3 são cloroplastidiais, porém, enquanto que as proteínas Gm-APX-1a e 1b são endereçadas às membranas tilacoidais, as Gm-APX-2, Gm-APX-3a e 3b são estromais (Tabela 9). Além disso, as proteínas Gm-APX-4 são citosólicas enquanto que as Gm-APX-5 são peroxissomais (Tabela 9).

Para MDHAR, a proteína Gm-MDHAR-1 pertence ao cloroplasto, Gm-MDHAR-2a e 2b ao peroxissomo e Gm-MDHAR-3a e 3b ao citosol (Tabela 9). Já para as proteínas DHAR, Gm-DHAR-1a e 1b e Gm-DHAR-2a e 2b são endereçadas ao citosol e cloroplasto, respectivamente (Tabela 9). Por outro lado, as proteínas GR podem ser endereçadas a dois locais distintos: Gm-GR-1a e 1b ao citosol ou peroxissomo e Gm-GR-2a e 2b ao cloroplasto ou mitocôndria (Tabela 9).

Tabela 9: Predição da localização subcelular para as proteínas APX, MDHAR, DHAR e GR em *G. max*. Clo, cloroplasto; Cit, citosol; Mit, mitocôndrias; M.P. Membrana plasmática; Per, peroxissomos; Til, tilacóides.

Predição da localização subcelular para as proteínas de <i>G.max</i>						
Família	Proteínas de <i>G. max</i>	TargetP	CELLO	WoLF PSORT	NCBI	Artigos publicados (<i>A. thaliana</i>)
APX	APX-1 ^a	-	Clo	Clo	Til	Clo (Wang, Hecker, Hauser, 2014)
	APX-1b	Clo	Clo	Clo	Til	Clo (Wang, Hecker, Hauser, 2014)
	APX-2	Clo	Clo	Clo	Clo	-
	APX-3 ^a	Clo	Clo	Clo	Clo	Clo (Jespersen et al., 1997)
	APX-3b	Clo	Clo	Clo	Clo	Clo (Jespersen et al., 1997)
	APX-4 ^a	Outro	Cit	Cit	Cit	Cit (Kubo et al., 1993)
	APX-4b	Outro	Cit	Cit	Cit	Cit (Kubo et al., 1993)
	APX-5a1	Outro	Nuc	Gol	Per	Per (Zhang et al., 1997; Narendra et al., 2006)
	APX-5a2	Outro	Cito	Cito	Per	Per (Zhang et al., 1997; Narendra et al., 2006)
	APX-5b1	Outro	Cito	Cito	Per	Per (Zhang et al., 1997; Narendra et al., 2006)
	APX-5b2	Outro	Cito	Cito	Per	Per (Zhang et al., 1997; Narendra et al., 2006)
MDHAR	MDHAR-1	Clo	Clo	Mit	Clo/Mit	Chlo (Obara, Sumi, Fukuda, 2002)
	MDHAR-2a	-	M. P.	Cit	Cit/Per	Pero (Lisenbee, Lingard & Trelease, 2005)
	MDHAR-2b	-	Clo	Clo	Clo	Pero (Lisenbee, Lingard & Trelease, 2005)
	MDHAR-3a	Outro	Cit	Cit	Cit	Cit (Lisenbee, Lingard & Trelease, 2005)
	MDHAR-3b	Outro	Cit	Cit	Cit	Cit (Lisenbee, Lingard & Trelease, 2005)
DHAR	DHAR-1a	Outro	Cit	Cit	Cit	Cit (Dixon, Davis & Edwards, 2002)
	DHAR-1b	Outro	Cit	Cit	Cit	Cit (Dixon, Davis & Edwards, 2002)
	DHAR-2a	Clo	Clo	Clo	Clo	Clo (Dixon, Davis & Edwards, 2002)
	DHAR-2b	Clo	Clo	Clo	Clo	Clo (Dixon, Davis & Edwards, 2002)
GR	GR-1 ^a	Outro	Cit	Cit	Cit	Cit/Per (Kataya & Reumann, 2010)
	GR-1b	Outro	Cit	Cit	Cit	Cit/Per (Kataya & Reumann, 2010)
	GR-2 ^a	Clo	Clo	Clo	Clo	Clo/Mit (Creissen et al., 1995; Ding et al., 2016)
	GR-2b	Clo	Clo	Clo	Clo	Clo/Mit (Creissen et al., 1995; Ding et al., 2016)

Fonte: Elaborado pelo autor

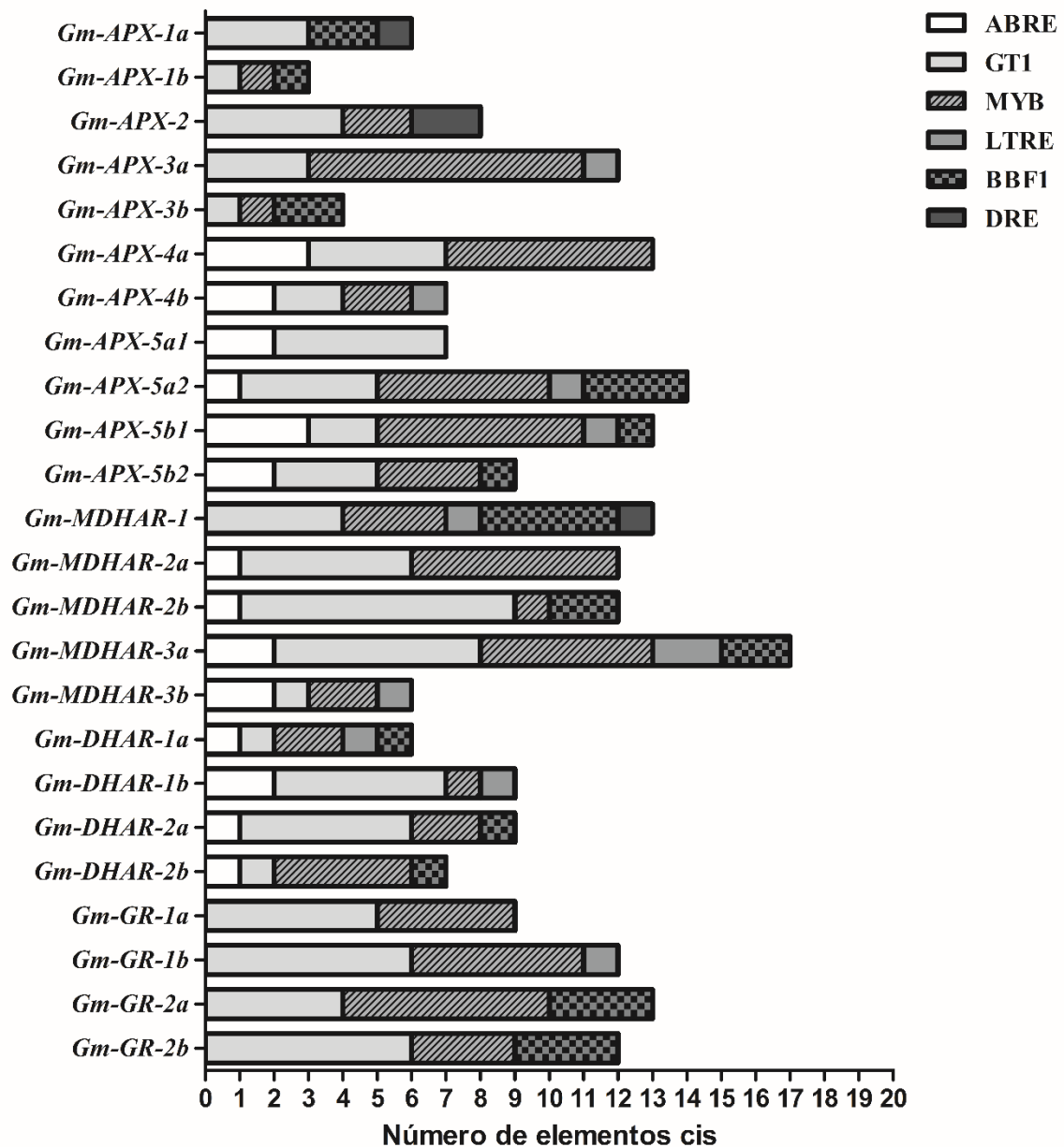
5.5 Possíveis elementos *cis* responsivos à estresses nas regiões promotoras nos genes *APX*, *MDHAR*, *DHAR* e *GR*

As análises das regiões promotoras (1000 pb *upstream* do códon de iniciação) dos

genes *APX*, *MDHAR*, *DHAR* e *GR* em *G. max* revelaram a presença de elementos *cis* regulatórios associados ao estresse, divididos em seis grupos: ABRE (ACGTG), GT1 (GGTTAA/GAAAAA), MYB (CTAACCA/CNGTTR/WAACCA/TA ACTG), LTRE (CCGAAA), BBF1 (ACTTTA) e DRE (GTCGAC) (Figura 5).

Os elementos *cis* GT1 e MYB foram detectados em todos os genes, exceto em *Gm-APX-1a* e *Gm-APX-5a1* (Figura 5). Em adição, os genes *Gm-MDHAR-2b* e *Gm-APX-3a* apresentaram a maior quantidade (8) de elementos *cis* GT1 e MYB, respectivamente. Em 14 genes foram identificados os elementos *cis* ABRE e BBF1. O número de elementos *cis* ABRE variou de um (*Gm-APX-5a2*, *Gm-MDHAR-2a*, *Gm-MDHAR-2b*, *Gm-DHAR-1a*, *Gm-DHAR-2a* e *Gm-DHAR-2b*) à quatro (*Gm-APX-4a* e *Gm-APX-5b1*). Da mesma forma, o número de elementos *cis* BBF1 variou de um (*Gm-APX-1b*, *Gm-APX-5b1*, *Gm-APX-5b2*, *Gm-DHAR-1a*, *Gm-DHAR-2a* e *Gm-DHAR-2b*) à quatro (*Gm-MDHAR-1*). Dez genes apresentaram somente um (*Gm-APX-3a*, *Gm-APX-4b*, *Gm-APX-5a2*, *Gm-APX-5b1*, *Gm-MDHAR-1*, *Gm-MDHAR-3b*, *Gm-DHAR-1a*, *Gm-DHAR-1b* e *Gm-GR-1b*) ou dois (*Gm-MDHAR-3a*) elementos *cis* LTRE, enquanto que apenas três genes (*Gm-APX-1a*, *APX-2*, e *Gm-MDHAR-1*) continham o elemento *cis* DRE (Figura 5).

Figura 5: Número de elementos *cis* regulatórios ABRE (ACGTG), GT1 (GGTTAA/GAAAAA), MYB (CTAACCA/CNGTTR/WAACCA/TAACCTG), LTRE (CCGAAA), BBF1 (ACTTTA) e DRE (GTCGAC) na região promotora dos genes do ciclo ASC-GSH em soja (*G. max*)



Fonte: Elaborado pelo autor

5.6 Expressão dos genes *APX*, *MDHAR*, *DHAR* e *GR* em soja submetida à condições de estresses

A expressão dos genes do ciclo ASC-GSH foi avaliada em soja submetida a condições de estresses, como salinidade, seca, etileno e ataque de patógenos usando dados de RNA-seq disponíveis do NCBI. Os resultados mostraram que todos os genes foram expressos, com exceção do *Gm-APX-5a2*. Em adição, o perfil de expressão desses genes revelou padrões específicos que variaram em relação ao tecido e/ou tratamento.

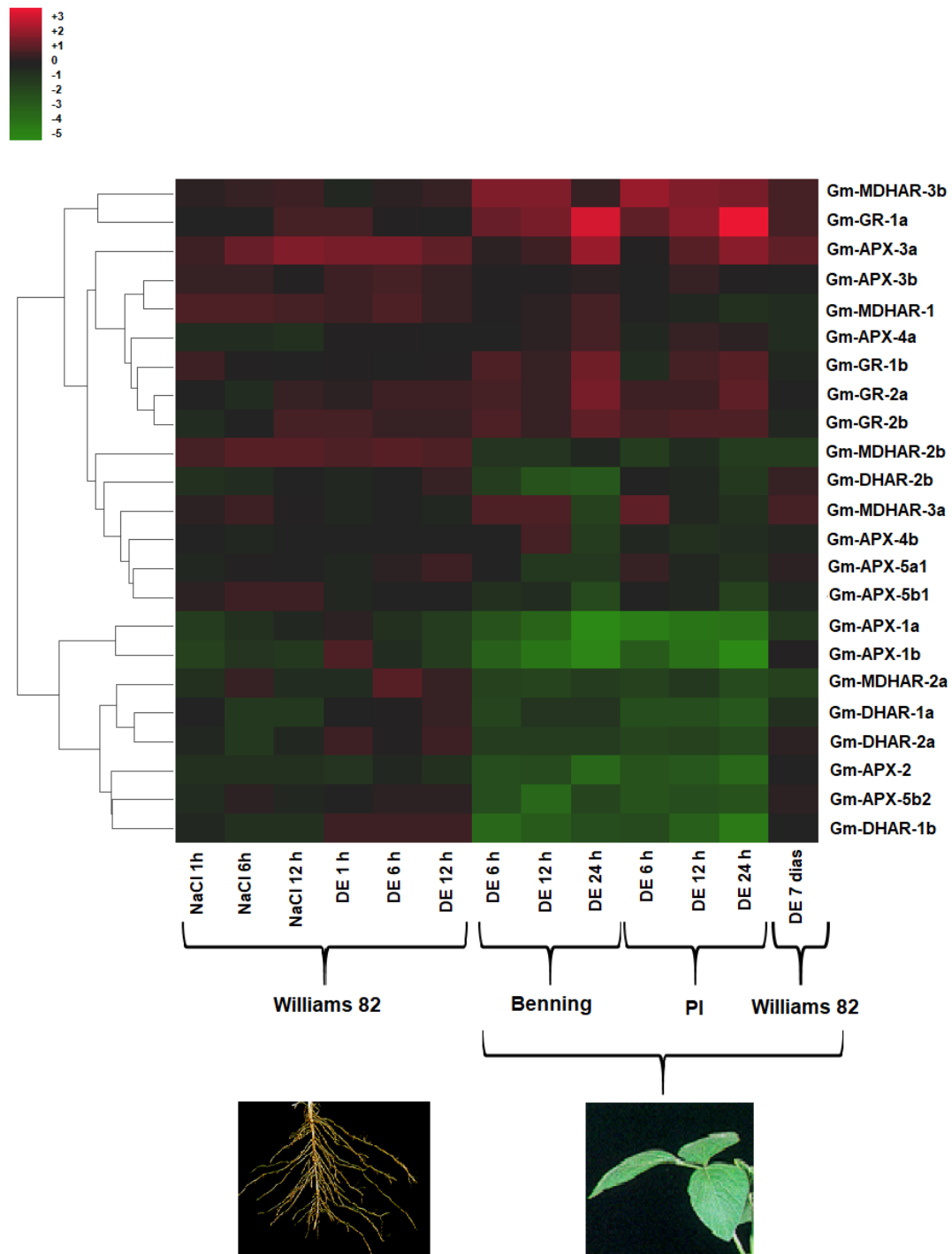
5.6.1 Expressão dos genes *APX*, *MDHAR*, *DHAR* e *GR* em soja submetida à seca e salinidade

Três experimentos de RNA-seq em soja submetida à diferentes condições de desidratação foram investigados. No primeiro experimento, as raízes do cultivar Williams 82 foram submetidas a (1) 100 mM de NaCl ou (2) expostas ao ar por 1, 6 e 12 horas (Figura 6). O gene *Gm-APX-3a* foi regulado positivamente em todos os tempos (exceto 1 h NaCl) para ambos os tratamentos. Em adição, *Gm-MDHAR-2b* foi regulado positivamente em 12 h de NaCl e 6 h de desidratação. Por outro lado, *Gm-APX-1a* e *1b* foram regulados negativamente em ambos os tratamentos de NaCl (1 e 12 h) e desidratação (12 h). Os genes *Gm-DHAR-1a* e *1b* tiveram sua expressão diminuída somente em NaCl (6 e 12 h). Com relação à *GR*, todos os genes não apresentaram expressão diferencial em ambos os tratamentos (Figura 6).

No segundo experimento, as folhas de ambos os cultivares PI (tolerante à seca) e Benning (sensível à seca) apresentaram o gene *Gm-APX-3a* regulado positivamente em 24 h e a expressão de *Gm-MDHAR-3b* e *Gm-GR-1a* aumentada em praticamente todos os tempos do estresse hídrico (Figura 6). O cultivar Benning também apresentou os demais genes GR genes altamente regulados em 24 h. Em contraste, *Gm-APX-1a* e *1b*, *Gm-MDHAR-2a* e todos os genes DHAR foram regulados negativamente em ambos os cultivares durante 24 h de tratamento (Figura 6). Interessantemente, o cultivar sensível induziu e reprimiu uma maior quantidade de genes quando comparado ao tolerante (Figura 6).

O terceiro ensaio avaliou o efeito da seca por sete dias em folhas do cultivar Williams 82 e revelou que somente o gene *Gm-APX-3a* aumentou diferencialmente seu número de transcritos (Figura 6). Em acordo com o segundo experimento, *Gm-APX-1a*, *Gm-MDHAR-2a* e *2b* foram regulados negativamente (Figura 6).

Figura 6: Expressão dos genes do ciclo ASC-GSH em raízes e folhas de soja (*G. max*) submetidas à estresse salino ou seca. A escala de cor representa os valores de Log_2 *fold-change* transformados a partir do método de normalização FPKM (*Fragments Per Kilobase Million*)

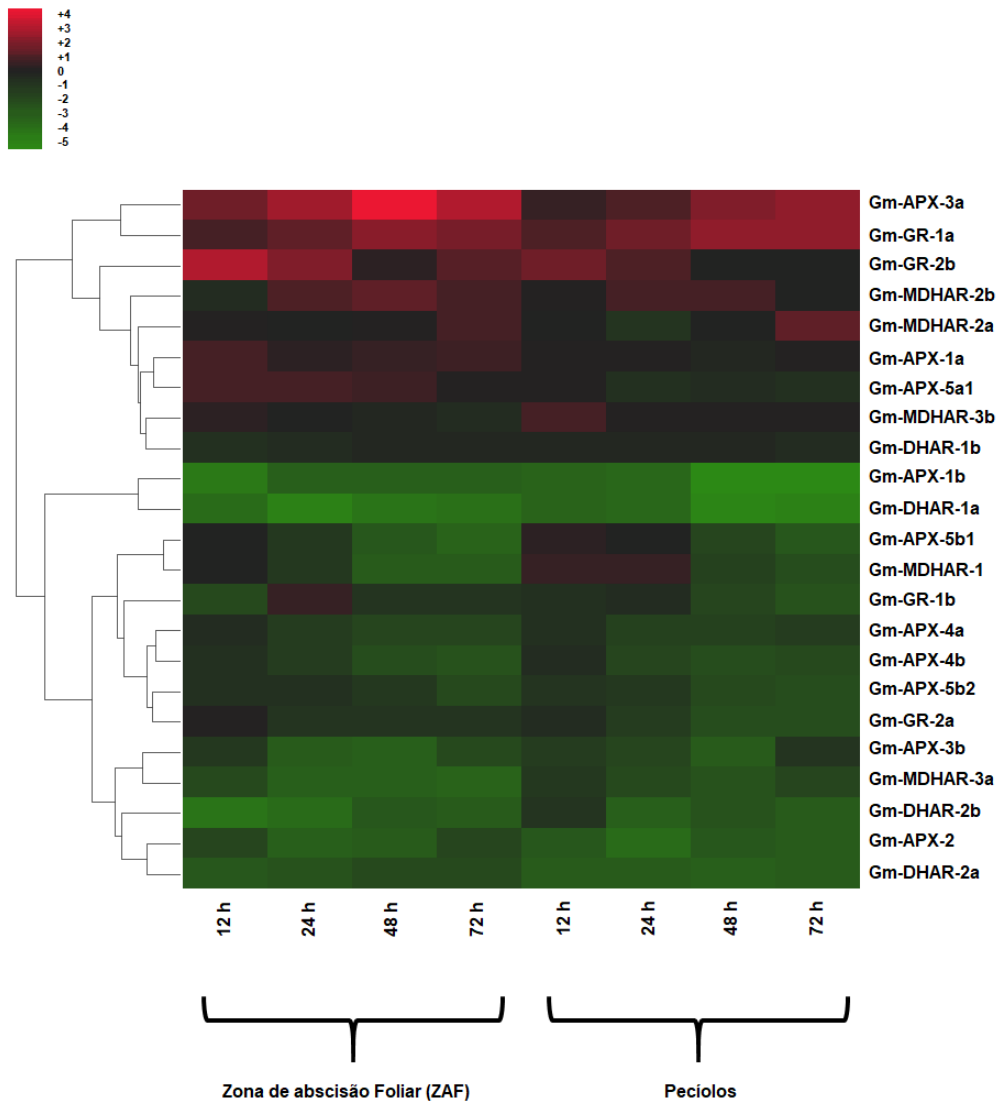


Fonte: *Heatmap* elaborado pelo autor. Imagens da raiz e folhas de soja adaptadas de EMBRAPA soja.

5.6.2 Expressão dos genes APX, MDHAR, DHAR e GR em soja submetida à alta concentração de etileno

Os efeitos do etileno (25 $\mu\text{L/L}$) na expressão dos genes do ciclo ASC-GSH foram avaliados em pecíolos e zona de abscisão foliar (ZAF) nos tempos de 0, 12, 24, 48 e 72 h. A análise revelou que os genes *Gm-APX-3a* e *Gm-GR-1a* foram regulados positivamente em praticamente todos os tempos em ambos os tecidos (Figura 7). Em adição, os níveis de transcritos para *Gm-MDHAR-2b* aumentaram em ZAF em 48 h e *Gm-MDHAR-2a* em pecíolos em 72 h. Entretanto, vários genes (*Gm-APX-1b*, 2, 3b, 5b1, 5b2, *Gm-MDHAR-3a*, *Gm-DHAR-1b*, 2a e 2b) reduziram a sua expressão em ambos os tecidos (Figura 7). O decréscimo dos níveis de transcritos para *Gm-GR-1b* e 2a em pecíolos também foi detectado após 24 h de exposição ao etileno (Figura 7).

Figura 7: Expressão dos genes do ciclo ASC-GSH na zona de abscisão foliar e pecíolos de soja (*G. max*) submetidos à etileno. A escala de cor representa os valores de $\text{Log}_2 \text{fold-change}$ transformados a partir do método de normalização FPKM (*Fragments Per Kilobase Million*)



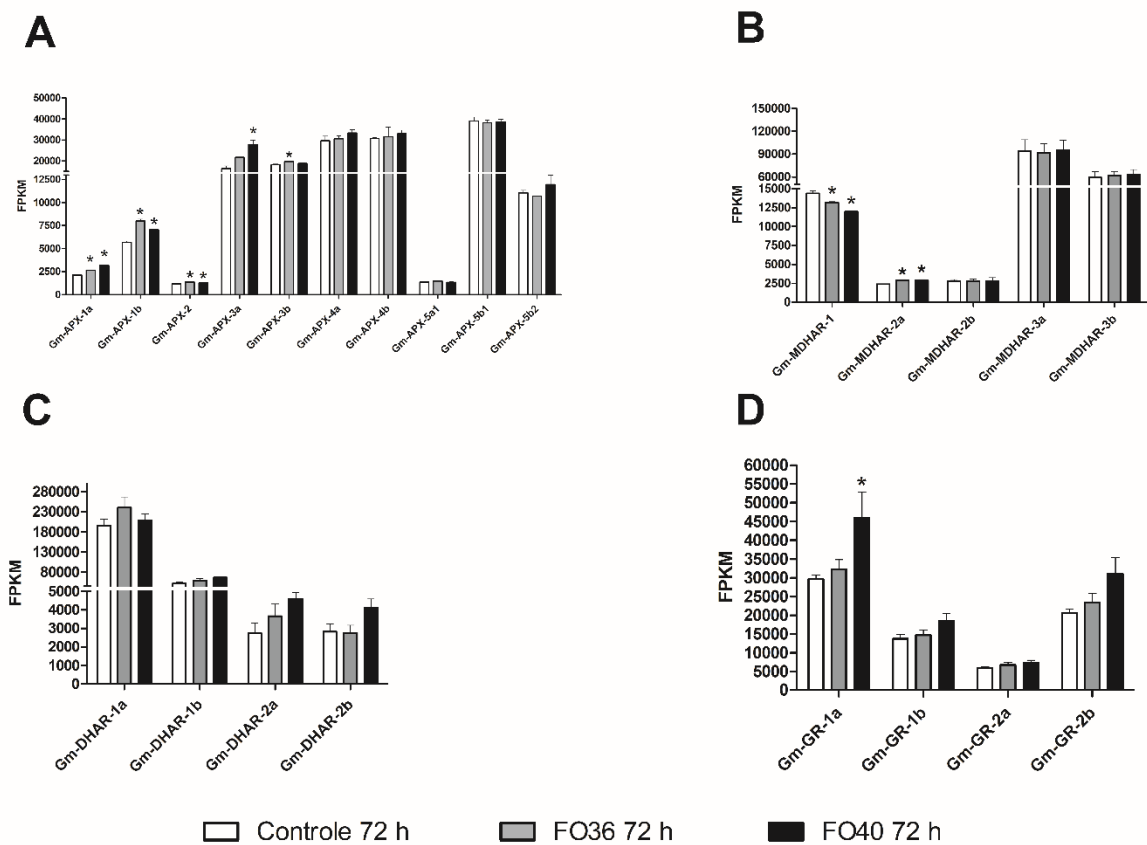
Fonte: Elaborado pelo autor

5.6.3 Expressão dos genes APX, MDHAR, DHAR e GR em raízes de soja infectada por *Fusarium oxysporum*

Os efeitos da infecção de raízes de soja pelo fungo *Fusarium oxysporum* foram investigados no tempo de 72 h utilizando isolados não patogênicos (FO36) e patogênicos (FO40). Os resultados mostram que apesar de todos os genes do ciclo ASC-GSH serem induzidos (exceto *Gm-APX-5a2*), nenhum deles foi diferencialmente expressos ($\text{log}_2 \text{fold-}$

$change \leq -1$ ou $\geq +1$) (Figura 8). Porém, os genes *Gm-APX-3a*, *Gm-MDHAR-2a* e *Gm-GR-1a* mostraram significativo aumento nos níveis de transcritos em 72 horas após a infecção com isolados patogênicos de *Fusarium oxysporum* (Figuras 8A e 8D).

Figura 8: Expressão dos genes *APX* (A), *MDHAR* (B), *DHAR* (C) e *GR* (D) de *G. max* em 72 h após a infecção com isolados patogênicos e não patogênicos de *Fusarium oxysporum*. Os bancos de dados de RNA-seq foram utilizados com o método FPKM (*Fragments Per Kilobase Million*). A análise estatística (*one-way* ANOVA seguido pelo teste de Bonferroni) foi aplicada para cada gene separadamente. As diferenças significativas a partir do controle são indicadas por * ($p < 0,05$).



Fonte: Elaborado pelo autor

5.7 Expressão dos genes *APX*, *MDHAR*, *DHAR* e *GR* em tecidos durante o desenvolvimento da soja

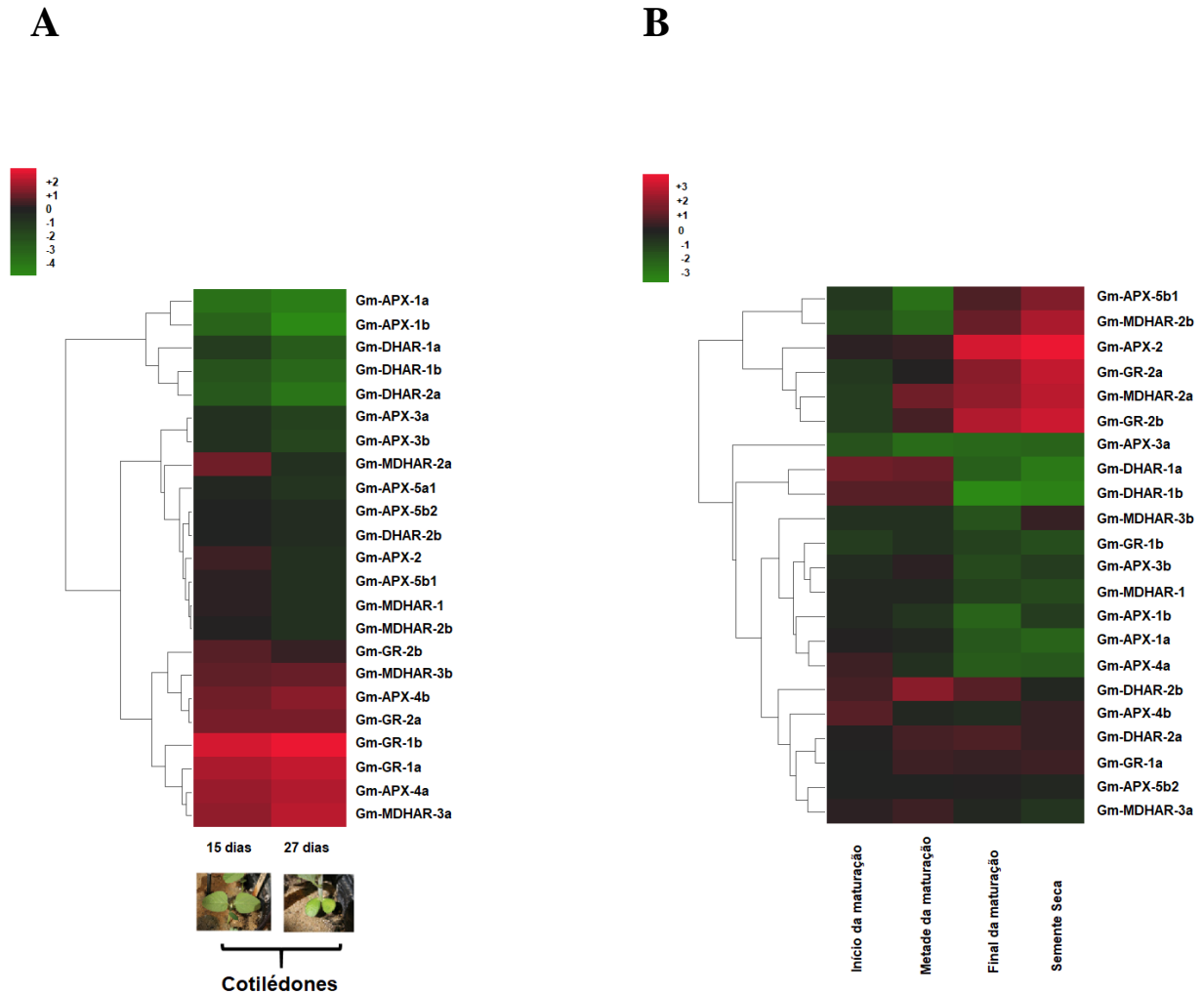
A expressão dos genes do ciclo ASC-GSH foi também investigada em cotilédones, sementes e raízes de *G. max* durante o desenvolvimento (Figuras 9A, 9B e 10). A análise de

expressão mostrou que todos os genes são funcionais nas condições analisadas, exceto o gene *Gm-APX-5a2*, com perfil de expressão característico do tecido e do estágio de desenvolvimento.

A análise de expressão de cotilédones foi realizada comparando o estágio de 15 e 27 dias em relação ao quarto dia de desenvolvimento desse órgão (Figura 9A). O estágio de desenvolvimento de cotilédones em 27 dias foi definido como o estágio de senescência. Oito genes foram regulados positivamente de 4 a 15 dias (*Gm-MDHAR-2a*, *Gm-MDHAR-3b*, *Gm-APX-4b*, *Gm-GR-2a*, *Gm-GR-1b*, *Gm-GR-1a*, *Gm-APX-4a* e *Gm-MDHAR-3a*). Em 27 dias, todos os genes acima mencionados continuaram com sua expressão aumentada quando comparado a 4 dias (exceto *Gm-MDHAR-2a*) (Figura 9A).

A expressão dos genes *APX*, *MDHAR*, *DHAR* e *GR* também foi investigada em estádios de maturação da semente representados pelo estágio inicial, intermediário, final e a dessecação em comparação com o estágio cotiledonar (Figura 9B). A análise revelou que os genes *Gm-APX-5b1*, *Gm-MDHAR-2b*, *Gm-APX-2*, *Gm-GR-2a*, *Gm-MDHAR-2a* e *Gm-GR-2b* apresentaram expressão positiva a partir da etapa final de maturação da semente (Figura 9B). Em contrapartida, a etapa final de maturação da semente também apresentou uma repressão do número de transcritos para *Gm-APX-1a*, *Gm-APX-1b*, *Gm-APX-3a*, *Gm-APX-3b*, *Gm-APX-4a*, *Gm-MDHAR-1*, *Gm-MDHAR-3b*, *Gm-DHAR-1a*, *Gm-DHAR-1b* e *Gm-GR-1b* (Figura 9B).

Figura 9: Expressão dos genes do ciclo ASC-GSH durante o desenvolvimento de cotilédones (A) e estádios de maturação da semente (B) de soja (*Glycine max*). A escala de cor representa os valores transformados do $\text{Log}_2 \text{fold-change}$ a partir do método FPKM (*Fragments Per Kilobase Million*) normalizado (p-value <0,05).

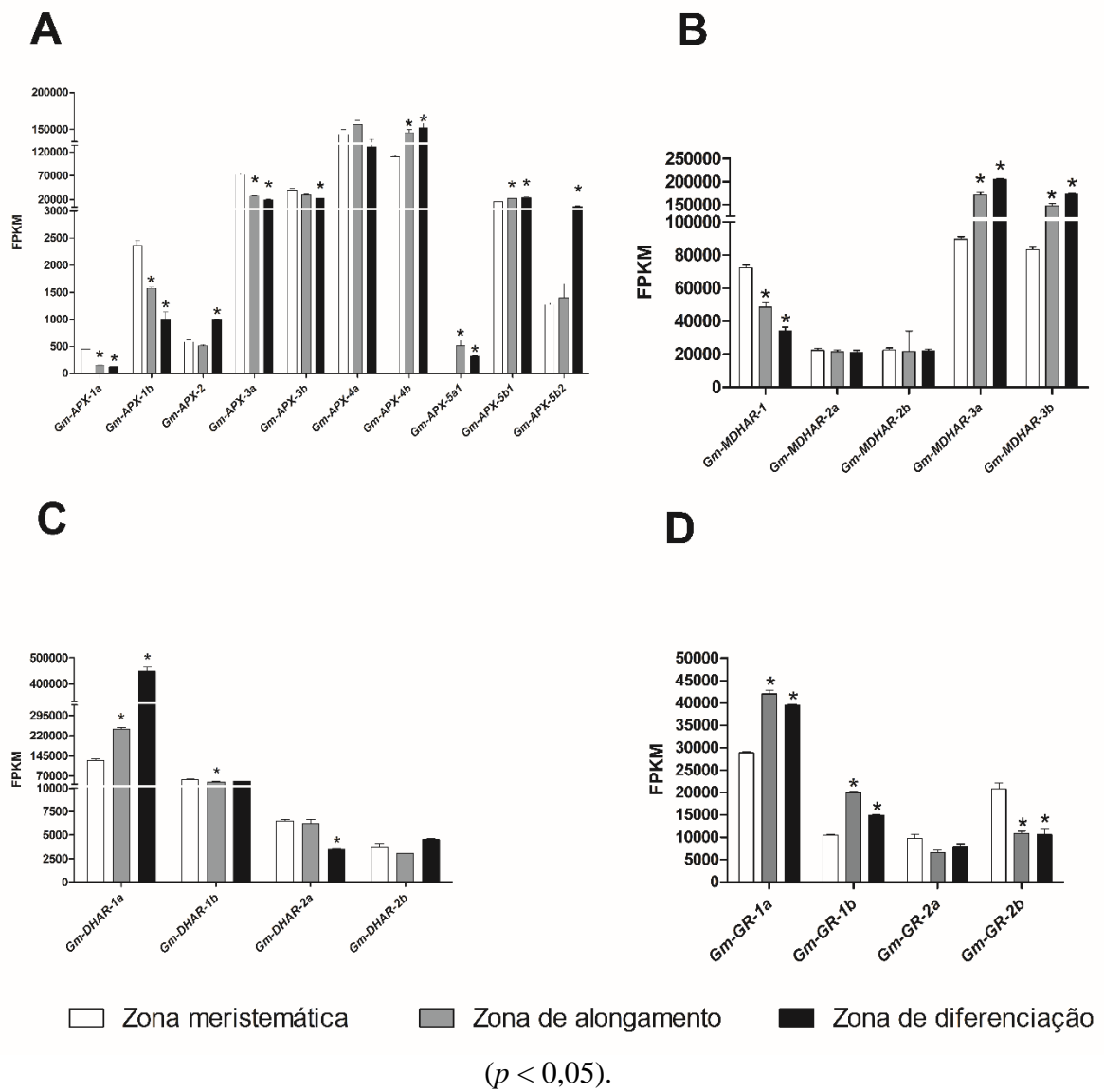


Fonte: *Heatmaps* elaborados pelo autor. Imagens dos estádios do cotilédone adaptados de BROWN & HUDSON, 2015.

Finalmente, a expressão dos genes do ciclo ASC-GSH foi investigada em três zonas de desenvolvimento da raiz de soja: zona meristemática, zona de alongamento e zona de diferenciação (Figura 10A, B, C e D). A análise mostrou que, quando comparadas as zonas de alongamento e diferenciação à zona meristemática, os genes *Gm-APX-2*, *Gm-APX-4b*, *Gm-APX-5a1*, *Gm-APX-5b2*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-DHAR-1a*, *Gm-GR-1a* e *Gm-GR-1b* tiveram um significativo aumento no número de transcritos enquanto os transcritos de

Gm-APX-1a, *Gm-APX-1b*, *Gm-APX-3a*, *Gm-APX-3b*, *Gm-MDHAR-1*, *Gm-DHAR-2a* e *Gm-GR-2b* foram reprimidos (Figura 10).

Figura 10: Expressão dos genes *APX* (A), *MDHAR* (B), *DHAR* (C) e *GR* (D) em zonas meristemática, de alongamento e de diferenciação de raízes de soja. Os bancos de dados de RNA-seq foram utilizados com o método FPKM (*Fragments Per Kilobase Million*). A análise estatística (one-way ANOVA seguido pelo teste de Bonferroni) foi aplicada para cada gene separadamente. As diferenças significativas a partir da zona meristemática são indicadas por *



Fonte: Elaborado pelo autor

6 DISCUSSÃO

No presente estudo, os genes do ciclo ASC-GSH da soja e de outras espécies leguminosas foram identificados, caracterizados e analisados sua filogenia utilizando a espécie *Arabidopsis* como referência. Posteriormente, análises de expressão *in silico* traçaram perfis de expressão dos membros gênicos em *Glycine max* sob diferentes condições de estresse e em tecidos durante o desenvolvimento.

Um total de 24 genes (11 *APX*, 5 *MDHAR*, 4 *DHAR* e 4 *GR*) que codificam as enzimas do ciclo ASC-GSH foram identificados em ambos os genomas de *G. max* e *G. soja*, enquanto para as demais espécies leguminosas foram identificados um máximo de 15 genes (Tabela 4). O número de genes dessas famílias multigênicas em outras espécies também é variado. Por exemplo, foram encontradas oito isoformas da *APX* em *Arabidopsis* e arroz (TEIXEIRA *et al.*, 2005) enquanto que em espinafre foram identificados cinco isoformas (ISHIKAWA *et al.*, 1995, 1996, 1998; CAVERZAN *et al.*, 2012). Um total de três isoformas *MDHAR* foram identificadas em tomate (STEVENS *et al.*, 2007). Duas diferentes isoformas da *DHAR* foram relatadas em espinafre (SHIMAOKA, YOKOTA & MIYAKE, 2000) e eucalipto (TEIXEIRA *et al.*, 2005). Em feijão caupi, duas isoformas da *GR* foram clonadas e identificadas (CONTOUR-ANSEL *et al.*, 2006).

O número de membros gênicos de cada espécie está relacionado ao processo de acúmulo de genes duplicados e a divergência evolutiva entre suas cópias como uma das possíveis respostas de adaptação durante a evolução (LESPINET *et al.*, 2002; LYNCH, 2002). Além disso, a quantidade de genes do ciclo ASC-GSH em cada espécie pode também estar associada aos diferentes sítios de produção de H₂O₂ na célula (cloroplastos, mitocôndrias e peroxissomos) (PANDEY *et al.*, 2015).

Em *G. max*, a maior quantidade de genes identificados está relacionada com cópias similares provenientes de eventos de duplicação recente do genoma que cobriu aproximadamente 75% dos genes dessa espécie (SCHMUTZ *et al.* 2010). É possível que a maioria das espécies de plantas tenham sofrido processos de duplicações do genoma inteiro, porém, a perda de genes redundantes após milhares de anos de evolução torna menos óbvia as marcas deixadas por tais eventos (JAILLON *et al.*, 2007). Em soja, o gene *Gm-APX-5a2* (provável pseudogene) não foi expresso em nenhum dos experimentos analisados, demonstrando que devido a sua redundância com *Gm-APX-5a1*, o primeiro deixou de ser funcional. Nas demais duplicações, verificou-se uma maior divergência na região não traduzida (UTR, *untranslated region*) (Dados não mostrados), indicando que as outras

duplicações estão passando por modificações gradativas de modo a divergir ao longo do tempo. Nesse sentido, acredita-se que a provável duplicação desses genes teve um importante papel na evolução do organismo, pois divergências em regiões regulatórias tendem à mudanças no padrão de expressão (XU *et al.*, 2012).

Em *G. max*, as proteínas dentro de cada família multigênica possuem variados valores de identidade. Por exemplo, enquanto Gm-APX-1a e Gm-APX-3a apresentaram somente 27 % de identidade, Gm-MDHAR-3a e Gm-MDHAR-3b são 99% idênticas (Tabelas 5 e 6). Essas variações de identidade também foram identificadas em outras espécies, tais como *Arabidopsis*, arroz, sorgo e tomate (TOGNOLLI *et al.*, 2002; KABIR & WANG *et al.*, 2010; OZYIGIT *et al.*, 2016). Aqui, a maior divergência entre as proteínas parálogas contribuiu para a formação de clados diferentes que permitiram a predição e caracterização de ortólogos dentro de cada família multigênica. Dessa forma, as proteínas APX, MDHAR, DHAR e GR de leguminosas foram divididas em cinco, três, dois e dois clados, respectivamente. Todas as espécies (com exceção da *V. carteri*) apresentaram proteínas representativas de cada clado formado (Figura 3A-D), entretanto, o grupo APX-5a foi detectado somente em *A. thaliana* e espécies da tribo Phaseoleae, indicando que esse gene tenha sido perdido nas outras espécies leguminosas aqui estudadas por meio de eventos de deleção durante a evolução.

Com relação à estrutura éxon/intron, a comparação entre os genes de *G. max* e seus respectivos ortólogos em *Arabidopsis* (espécie modelo) revelou uma alta conservação baseada nas quantidades similares de éxons e íntrons (Figura 4). Entretanto, a extensão e, em alguns casos, maior número de íntrons nos genes de soja proporcionaram um maior tamanho dos mesmos quando relacionados aos seus ortólogos de *Arabidopsis*. A variação nas estruturas éxon-intron em genes parálogos tão bem como entre ortólogos pode ser originada a partir do mecanismo de éxon/intron perda/ganho, que contribuiu para a divergência estrutural. Entretanto, na prática, nem sempre é fácil determinar se um éxon é perdido ou o intron é ganho por parálogos/ortólogos a menos que o estado ancestral seja conhecido (XU *et al.*, 2012). Essas divergências estruturais podem resultar na tradução de proteínas com diferentes tamanhos e organização dos domínios. Em *G. max*, os domínios proteicos de parálogos apresentaram as mesmas funções, indicando forte conservação da atividade dessas enzimas ao longo da evolução (ALBERTS *et al.*, 2010). Em adição, as análises da estrutura gênica das sequências de *G. max* em relação às de *Arabidopsis* tão bem como os domínios condizentes com a literatura contribuíram para conferir uma maior consistência e confiabilidade nos resultados obtidos através da anotação gênica aqui realizada.

Neste trabalho, foi observado que os membros gênicos das famílias APX, MDHAR,

DHAR e GR são expressos em diferentes tecidos e cultivares de soja em condições de estresses, exceto *Gm-APX-5a2*. Para fins didáticos, ficou estabelecido que entre parênteses está descrita a localização subcelular da proteína referente ao transcrito mencionado.

Sob condições de desidratação, em ambas, raízes e folhas de soja, o gene *Gm-APX-3a* (cloroplastos) foi regulado positivamente (Figura 6). Entretanto, o número de transcritos para o gene *Gm-MDHAR-2b* (peroxissomos) aumentou em raízes enquanto o gene *Gm-MDHAR-3b* (citosol) foi expresso positivamente em folhas (Figura 6). Apenas em folhas, o gene *Gm-GR-1a* (citosol) foi diferencialmente expresso em todos os tempos. Além disso, as regiões promotoras dos genes *Gm-APX-3a* e *Gm-MDHAR-3b* apresentaram elementos *cis* MYB responsivos ao déficit hídrico enquanto que a região promotora do gene *Gm-MDHAR-2b* possui uma grande quantidade de elementos *cis* GT1 responsivos à salinidade (Figura 5). Dessa forma, o perfil de expressão e a presença de elementos *cis* responsivos à salinidade e seca em regiões promotoras sugerem um importante papel dos genes acima mencionados em resposta à essas condições de estresse.

De maneira similar à soja, raízes de arroz submetidas à estresse salino apresentaram um membro da APX que codifica uma isoforma cloroplastidial (*OsAPX8*) regulada positivamente com o aumento dos níveis de ABA e obteve um significativo aumento da sua atividade enzimática (HONG *et al.*, 2007). Em ervilha (*Pisum sativum* cv. Puget) submetida à altas concentrações de NaCl houve maior indução de uma APX que codifica uma isoforma cloroplastidial (GOMEZ *et al.*, 2014). Sečenji e colaboradores (2010) relataram que duas MDHAR que codificam proteínas citosólicas aumentaram significativamente seus níveis de transcritos em um cultivar de trigo sensível à seca sob condições limitadas de água, no entanto, não observaram qualquer regulação de transcritos para os MDHARs que codificam proteínas cloroplastidiais ou peroxissomais. Há poucos relatos na literatura de expressão de uma MDHAR peroxissomal em plantas submetida à estresse hídrico ou salino. Porém, sob danos mecânicos e frio, uma MDHAR que codifica uma proteína peroxissomal em ervilha (*Pisum sativum*) foi expressa positivamente. A presença do gene *Gm-GR-1a* citosólico pode estar relacionado principalmente ao mecanismo antioxidante protetor ativo que atua no citosol (FOYER & NOCTOR 2003).

Em contrapartida, o gene que codifica uma proteína tilacoidal *Gm-APX-1b* apresentou uma supressão do número de transcritos em praticamente todos os tempos de estresse salino e seca impostos à soja. O *Gm-APX-1b* também apresentou poucos elementos *cis* responsivos a estresse dentre todas as regiões promotoras avaliadas. Alguns estudos sugerem que uma maior sensibilidade dos tilacóides é evidenciada pela prematura supressão das APX dessa organela,

podendo ser um fator limitante no sistema antioxidante durante o estresse oxidativo em cloroplastos (TEIXEIRA *et al.*, 2006; HONG *et al.*, 2007; GOMEZ *et al.*, 2014).

A expressão dos genes do ciclo ASC-GSH também foi investigada sob concentração elevada de etileno em zona de abscisão foliar e pecíolos de soja (Figura 7). Interessantemente, o número de transcritos de *Gm-APX-3a*, *Gm-MDHAR-2b* e *Gm-GR-1a* também aumentaram significativamente. Em adição, o gene *Gm-GR-2b* (cloroplastos) também foi regulado positivamente (Figura 7). A indução desses genes pode estar associada ao dano mecânico ocasionado pela remoção dos pecíolos das plantas e/ou aplicação do etileno. O gás etileno é produzido em vários tecidos vegetais e está relacionada com várias respostas da planta, dentre elas a abscisão foliar e senescência (TAIZ & ZEIGER, 2013). Nesse sentido, uma possível aceleração da senescência ocasionada pela aplicação de etileno em zonas específicas de abscisão foliar pode ter estimulado a produção de H₂O₂ que induziu os genes mencionados acima. Entretanto, estudos serão necessários para melhor esclarecer o envolvimento desses genes em resposta ao etileno.

A análise da expressão dos genes do ciclo ASC-GSH em soja submetida ao ataque do patógeno *Fusarium Oxysporum* revelou um aumento significativo nos níveis de transcritos de *Gm-APX-3a*, *Gm-MDHAR-2a* (peroxissomal) e *Gm-GR-1a* (Figura 8). Entretanto, nenhum gene foi diferencialmente expresso. O acúmulo de H₂O₂ pode ser diretamente tóxico ao patógeno, contribuindo para a morte da célula hospedeira ou atuar diretamente na morte do patógeno (TAIZ & ZEIGER, 2013). Dessa forma, apenas um pequeno aumento do *pool* de proteínas constitutivas em cada compartimento celular pode ser o suficiente para sustentar a proteção antioxidante (NOGUEIRA, SASSAK & MAIA, 2011). Entretanto, estudos identificaram alguns genes do ciclo responsivos à interação planta-patógeno. Por exemplo, em arroz submetido ao ataque do fungo *Magnaporthe oryzae*, um gene *APX* que codifica uma proteína citosólica foi expresso positivamente (WANG *et al.*, 2015). Genótipos resistentes de milho à fungos do gênero *Fusarium* apresentaram altos níveis de expressão das enzimas do ciclo ASC-GSH em relação ao genótipo susceptível (LANUBILE *et al.*, 2012). Esses resultados indicam que, apesar das plantas adquirirem uma sofisticada estratégia defensiva contra patógenos ao longo da evolução, cada espécie vegetal (e sua interação com diferentes patógenos) pode induzir de forma específica seu sistema antioxidante enzimático (CARVALHO, 2012).

Em relação aos genes *DHAR* em soja, observou-se uma repressão no número de transcritos em folhas submetidas à seca, em pecíolos e zona de abscisão foliar submetidas à etileno e raízes sob salinidade enquanto que em raízes sob desidratação e fungo não

apresentaram diferenças significativas. Além disso, a presença de elementos cis responsivos a estresses nas regiões promotoras dos genes *DHAR* são bem inferiores aos genes *MDHAR* (Figura 5). Esses resultados sugerem que os genes *MDHAR*, em especial *Gm-MDHAR-2b* e *Gm-MDHAR-3b*, desempenham um papel fundamental na manutenção da homeostase redox celular na restauração do ascorbato oxidado pela ação das isoformas da APX em plantas de soja quando comparados aos genes *DHAR*.

Há inúmeras evidências que provam que os genes do ciclo ASC-GSH possuem um papel protetivo sob várias condições de estresse em plantas, incluindo salinidade, seca e patógenos. Porém, a regulação dos genes *APX*, *MDHAR*, *DHAR* e *GR* sob estresse é complexa e varia de um organismo para o outro. Nesse sentido, os genes *Gm-APX-3a*, *Gm-GR-1a*, *Gm-MDHAR-2b* e *Gm-MDHAR-3b* foram os mais responsivos aos estresses analisados. No entanto, outros estudos envolvendo diferentes intensidades de estresse em tecidos de *G. max* e outras leguminosas poderão ser promissores para elucidar o papel de cada membro das famílias APX, MDHAR, DHAR e GR. De início, os resultados aqui encontrados apontam esses genes como candidatos para a aplicação de ferramentas biotecnológicas visando o desenvolvimento de cultivares melhor adaptados às condições de estresse.

A análise de expressão durante o desenvolvimento da soja detectou a presença de todos os genes nas diferentes partes estudadas da planta, com exceção *Gm-APX-5a2* (Figuras 9 e 10). Em cotilédones senescentes, os genes *Gm-APX-4a*, *Gm-MDHAR-3a*, *Gm-GR-1a* e *Gm-GR-1b* (citosol) aumentaram em até quatro vezes seu número de transcritos (Figura 9A). Além desses, os genes *Gm-MDHAR-2a* (peroxissomos), *Gm-MDHAR-3b* (citosol), *Gm-APX-4b* (citosol), *Gm-GR-2a* (cloroplastos), em menor escala, também foram regulados positivamente (Figura 9A). Esse achado sugere que a função desses genes possa estar relacionada com a senescência dos cotilédones. Dentre os vários estímulos responsáveis pelo controle da senescência, a produção de H₂O₂ em excesso participa deste processo, e a modulação do sistema antioxidante é provavelmente responsável por esse início (RIBEIRO *et al.*, 2017). Ribeiro e colaboradores (2017) observaram que o envelhecimento das folhas de arroz naturalmente acumula EROs.

Assim como em soja, a senescência de folhas de *A. thaliana* mostrou que o gene *At-GR2* (ortólogo à *Gm-GR-2a*) foi expresso positivamente, sugerindo que *At-GR2* poderia estar envolvido fortemente nesse processo (DING *et al.*, 2016). Panchuk e colaboradores (2005) verificaram que os níveis de transcritos das isoformas da APX não aumentaram em senescência de folhas de *A. thaliana*, porém, assim como neste trabalho, a expressão de uma APX que codifica uma proteína tilacoidal foi altamente reprimida (PANCHUK *et al.*, 2005).

Em estádios de maturação da semente, os genes *Gm-APX-5b1*, *Gm-MDHAR-2a* e *Gm-MDHAR-2b* (peroxissomos) e os genes *Gm-APX-2*, *Gm-GR-2a* e *Gm-GR-2b* (cloroplastos) apresentaram regulação positiva a partir da etapa final de maturação da semente (Figura 9B). Essa regulação pode estar relacionada à produção de H₂O₂, que por sua vez, está presente em todas as fases do ciclo de vida da semente. O acúmulo de H₂O₂ é importante para quebrar a dormência das sementes e estimular a germinação e proteção contra agentes patogênicos. No entanto, a produção excessiva de EROs pode ser prejudicial (KUMAR *et al.*, 2015).

Muitos estudos revelaram um aumento da atividade das enzimas APX, MDHAR, DHAR e GR no processo de maturação de sementes de plantas (TOMMASI, PACIOLLA & ARRIGONI, 1999; BAILLY *et al.*, 2001; GARNCZARSKA, BEDNARSKI & JANCELEWICZ, 2008). Entretanto, poucos relacionaram o processo de maturação com os níveis de transcritos que codificam essas enzimas. Por exemplo, Chen e colaboradores (2014) relacionaram a maior abundância de transcritos de uma APX citosólica (*APX6*) ao papel desse gene na proteção da maturação da semente em *Arabidopsis*.

Em zonas de desenvolvimento da raiz, a expressão dos genes *Gm-APX-4b*, *Gm-MDHAR-3a*, *Gm-MDHAR-3b*, *Gm-GR-1a*, *Gm-DHAR-1a* e *Gm-GR-1b* (citosol), *Gm-APX-5a1* e *Gm-APX-5b2* (peroxissomos) e *Gm-APX-2* (cloroplastos) aumentaram significativamente em relação à zona meristemática (Figura 10). Dunand, Crèvecoeur & Penel (2006) sugeriram que a zona de diferenciação apresenta uma maior produção de H₂O₂ que parece estar envolvido na restrição do crescimento e formação de pelos da raiz em *Arabidopsis* e relacionaram com a presença de peroxidases. Em um outro estudo, a superexpressão de um gene que codifica uma peroxidase em *Arabidopsis* proporcionou um maior crescimento da raiz (PASSARDI *et al.*, 2006).

Entretanto, não foram encontrados estudos relacionando a expressão dos genes do ciclo ASC-GSH com zonas de desenvolvimento de raízes de plantas. Dessa forma, estudos ainda deverão ser feitos para investigar se há uma relação entre a expressão desses genes e seu papel na zona de diferenciação que está diretamente relacionada ao processo de maturação celular em raízes.

Os resultados de expressão para desenvolvimento encontrados aqui revelam que membros das famílias APX, MDHAR, DHAR e GR em diferentes localizações subcelulares estão relacionados às necessidades específicas dos órgãos/tecidos da planta de soja durante seu ciclo de vida. Entretanto, estudos mais detalhados serão necessários para melhor elucidar os papéis desempenhados pelo H₂O₂ e, conseqüentemente, a indução dos genes do ciclo ASC-GSH em soja.

7 CONCLUSÕES

O ciclo ascorbato-glutationa é representado pelas famílias multigênicas da APX, MDHAR, DHAR e GR. Em *G. max* foram identificados um total de 24 genes, classificados como *Gm-APX-1a, 1b, 2, 3a, 3b, 4a, 4b, 5a1, 5a2, 5b1 e 5b2*; *Gm-MDHAR-1, 2a, 2b, 3a e 3b*; *Gm-DHAR-1a, 1b, 2a e 2b*; *Gm-GR-1a, 1b, 2a e 2b*. O maior número de genes em *G. max* e *G. soja*, comparado às outras espécies leguminosas e *A. thaliana* sugere a ocorrência de eventos de duplicação recentes no genoma da soja.

Em relação ao perfil de expressão, todos os genes em *G. max* são funcionais, exceto *Gm-APX-5a2*. Além disso, a predominância do aumento de transcritos de *Gm-APX-3a, Gm-GR-1a, Gm-MDHAR-2b e Gm-MDHAR-3b* que codificam proteínas cloroplastidiais, citosólicas, peroxissomais e citosólicas, respectivamente, em soja submetida à condições de estresse sugere esses genes como potenciais candidatos para o programa de melhoramento genético.

Em desenvolvimento, os diferentes perfis de expressão observados para cotilédones, semente e raiz sugerem que ocorra uma expressão gênica espacial/temporal entre os membros do ciclo ASC-GSH. Geralmente, os membros gênicos regulados positivamente são:

- Senescência de cotilédones: *Gm-APX-4a, Gm-APX-4b, Gm-MDHAR-3a, Gm-MDHAR-3b, Gm-GR-1a e Gm-GR-1b* (codificam proteínas citosólicas);
- Semente seca: *Gm-APX-5b1, Gm-MDHAR-2a e Gm-MDHAR-2b* (codificam proteínas peroxissomais); *Gm-APX-2 e Gm-GR-2a* (codificam proteínas cloroplastidiais);
- Zona de diferenciação da raiz: *Gm-APX-2, Gm-APX-4b, Gm-APX-5a1, Gm-APX-5b2, Gm-MDHAR-3a, Gm-MDHAR-3b, Gm-DHAR-1a, Gm-GR-1a e Gm-GR-1b* (codificam proteínas cloroplastidiais, peroxissomais e citosólicas).

REFERÊNCIAS

- ADAMS, M. D., CELNIKER, S. E., HOLT, R. A., EVANS, C. A., GOCAYNE, J. D., AMANATIDES, P. G., ... GEORGE, R. A. The genome sequence of *Drosophila melanogaster*. **Science**, v. 287, p. 2185-2195, 2000.
- ALBERTS, B., JOHNSON, A., LEWIS, J., MORGAN, D., RAFF, M., ROBERTS, K., HUNT, T. (2010). *Biologia molecular da célula*. Artmed Editora.
- AHMAD, S. **Oxidative stress and antioxidant defenses in biology**. Springer Science & Business Media, 2012.
- 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 research**, v. 25, p. 3389-3402, 1997.
- ASADA, K. Production and scavenging of reactive oxygen species in chloroplasts and their functions. **Plant physiology**, v. 141, n. 2, p. 391-396, 2006.
- ASHRAF, M. H. P. J. C., HARRIS, P. J. C. Photosynthesis under stressful environments: an overview. **Photosynthetica**, v. 51, n. 2, p. 163-190, 2013.
- BARNABÁS, B., JÄGER, K., FEHÉR, A. The effect of drought and heat stress on reproductive processes in cereals. **Plant, cell & environment**, v. 31, p. 11-38, 2008.
- BITA, C. E., GERATS, T. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. **Frontiers in plant science**, v. 4, 2013.
- BELAMKAR, V., WEEKS, N. T., BHARTI, A. K., FARMER, A. D., GRAHAM, M. A., CANNON, S. 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**, v. 15, p. 950, 2014.
- BROADBENT, P., CREISSEN, G. P., KULAR, B., WELLBURN, A. R., MULLINEAUX, P. M. Oxidative stress responses in transgenic tobacco containing altered levels of glutathione reductase activity. **The Plant Journal**, v. 8, p. 247-255, 1995.
- BROWN, A. V., HUDSON, K. A. Developmental profiling of gene expression in soybean trifoliolate leaves and cotyledons. **BMC plant biology**, v. 15, p. 169, 2015.
- BUCHANAN, B. B., GRUISSEM W., JONES R. L. **Biochemistry and molecular biology of plants**. John Wiley & Sons, 2015.
- BUNNIK, E. M., LE ROCH, K. G. An Introduction to Functional Genomics and Systems Biology. **Advances in wound care**, v. 2, p. 490-498, 2013.
- CÂMARA, G.M.S. Introdução ao agronegócio da soja. In: CÂMARA, G. M.S. (Ed.) **Soja: tecnologia da produção**. Piracicaba, p. 1-30, 2015.

CAVERZAN, A., PASSAIA, G., ROSA, S. B., RIBEIRO, C. W., LAZZAROTTO, F., MARGIS-PINHEIRO, M. Plant responses to stresses: role of ascorbate peroxidase in the antioxidant protection. **Genetics and molecular biology**, v. 35, p. 1011-1019, 2012.

CHANG, W. C., LEE, T. Y., HUANG, H. D., HUANG, H. Y., PAN, R. L. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. **BMC genomics**, v. 9, p. 561, 2008.

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

CHEN, W., YAO, Q., PATIL, G. B., AGARWAL, G., DESHMUKH, R. K., LIN, L., WANG, B., WANG, Y., PRINCE, S. J., SONG, L., XU, D., AN Y. C., VALLIYODAN, B., VARSHNEY, R. K., NGUYEN, H. T. Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. **Frontiers in plant science**, v. 7, p. 1044, 2016.

CHOUDHURY, F. K., RIVERO, R. M., BLUMWALD, E., MITTLER, R. Reactive oxygen species, abiotic stress and stress combination. **The Plant Journal**, v. 90, p. 856-867, 2017

COMPANHIA NACIONAL DE ABASTECIMENTO. **Perspectivas para a Agropecuária**. Brasília: CONAB, 2017, V.5, p. 1-111. Disponível em <<http://www.conab.gov.br>>. Acesso em: 25 dez. 2017.

CONTOUR-ANSEL, D., TORRES-FRANKLIN, M. L., CRUZ DE CARVALHO, M. H., D'ARCY-LAMETA, A., ZUILY-FODIL, Y. Glutathione reductase in leaves of cowpea: cloning of two cDNAs, expression and enzymatic activity under progressive drought stress, desiccation and abscisic acid treatment. **Annals of botany**, v. 98, p. 1279-1287, 2006.

COSTA, J. H., MCDONALD, A. E., ARNHOLDT-SCHMITT, B., DE MELO, D. F. A classification scheme for alternative oxidases reveals the taxonomic distribution and evolutionary history of the enzyme in angiosperms. **Mitochondrion**, p. 172-183, 2014.

CREISSEN, G., REYNOLDS, H., XUE, Y., MULLINEAUX, P. Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. **The Plant Journal**, v. 8, p. 167-175, 1995.

CRUZ-RUS, E., AMAYA, I., SÁNCHEZ-SEVILLA, J. F., BOTELLA, M. A., VALPUESTA, V. Regulation of L-ascorbic acid content in strawberry fruits. **Journal of Experimental Botany**, v. 62, p. 4191-4201, 2011.

DAS, K., ROYCHOUDHURY, A. Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. **Frontiers in Environmental Science**, v. 2, p. 53, 2014.

DĄBROWSKA, G., KATA, A., GOC, A., SZECHYŃSKA-HEBDA, M., SKRZYPEK, E. Characteristics of the plant ascorbate peroxidase family. **Acta Biol Cracow Ser Bot**, v. 49, p. 7-17, 2007.

- DALTON, D. A., BAIRD, L. M., LANGEBERG, L., TAUGHER, C. Y., ANYAN, W. R., VANCE, C. P., SARATH, G. Subcellular localization of oxygen defense enzymes in soybean (*Glycine max* [L.] Merr.) root nodules. **Plant Physiology**, v. 102, p. 481-489, 1993.
- DE AZEVEDO NETO, A. D., PRISCO, J. T., ENÉAS-FILHO, J., DE ABREU, C. E. B., GOMES-FILHO, E. Effect of salt stress on antioxidative enzymes and lipid peroxidation in leaves and roots of salt-tolerant and salt-sensitive maize genotypes. **Environmental and Experimental Botany**, v. 56, p. 87-94, 2006.
- DE MENEZES, E. O., & MORAIS, J. M. L. **A seca no Nordeste: desafios e soluções**. Atual Editores, 2002.
- DESHMUKH, R., SONAH, H., PATIL, G., CHEN, W., PRINCE, S., MUTAVA, R., VUONG T., VALLIYODAN B., NGUYEN H. T. Integrating omic approaches for abiotic stress tolerance in soybean. **Frontiers in Plant science**, v. 5, 2014.
- DING, S., JIANG, R., LU, Q., WEN, X., LU, C. Glutathione reductase 2 maintains the function of photosystem II in Arabidopsis under excess light. **Biochimica et Biophysica Acta (BBA)-Bioenergetics**, v. 1857, p. 665-677, 2016.
- DUNAND, C., CREVECOEUR, M., PENEL, C. Distribution of superoxide and hydrogen peroxide in. **New Phytologist**, 2006.
- ELTAYEB, A. E., KAWANO, N., BADAWI, G. H., KAMINAKA, H., SANEKATA, T., SHIBAHARA, T., INANAGA, S., Tanaka, K. Overexpression of monodehydroascorbate reductase in transgenic tobacco confers enhanced tolerance to ozone, salt and polyethylene glycol stresses. **Planta**, v. 225, p. 1255-1264, 2007.
- ELTELIB, H. A., FUJIKAWA, Y., ESAKA, M. Overexpression of the acerola (*Malpighia glabra*) monodehydroascorbate reductase gene in transgenic tobacco plants results in increased ascorbate levels and enhanced tolerance to salt stress. **South African journal of botany**, v. 78, p. 295-301, 2012.
- EMANUELSSON, O., NIELSEN, H., BRUNAK, S., VON HEIJNE, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. **Journal of molecular biology**, v. 300, p. 1005-1016, 2000.
- EMBRAPA SOJA. Disponível em <<https://www.embrapa.br/soja/cultivos/soja1>>. Acesso em: 30/12/2017.
- FAHAD, S., BAJWA, A. A., NAZIR, U., ANJUM, S. A., FAROOQ, A., ZOHAIB, A., SADIA, S., NASIM W., ADKINS S., SAUD, S., IHSAN, M. Z., ALHARBY H., WU, C. WANG, D., HUANG J. Crop production under drought and heat stress: plant responses and management options. **Frontiers in plant science**, v. 8, p. 1147, 2017.
- FAN, L., LINKER, R., GEPSTEIN, S., TANIMOTO, E., YAMAMOTO, R., NEUMANN, P. M. Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics. **Plant physiology**, v. 140, p. 603-612, 2006.

FARIAS, J. R. B., NEPOMUCENO, A. L., & NEUMAIER, N. (2007). Ecofisiologia da soja. **Embrapa Soja-Circular Técnica (INFOTECA-E)**, 2007.

FEDERIZZI L.C. A SOJA COMO FATOR DE COMPETITIVIDADE NO MERCOSUL: HISTÓRICO, PRODUÇÃO E PERSPECTIVAS FUTURAS. In: III Encontro CEPAN: Vantagens Competitivas dos Agronegócios no Mercosul, Porto Alegre, **CD dos Anais**, Centro de estudos e Pesquisas em Agronegócio – CEPAN/UFRGS, 2005.

FEHR, W. R.; CAVINESS, C. E. Stages of soybean development. Ames: Iowa State University/ Cooperative Extension Service 1977.

FERREIRA, R. G., TÁVORA, F. J. A. F., HERNANDEZ, F. F. F. Distribuição da matéria seca e composição química das raízes, caule e folhas de goiabeira submetida a estresse salino. **Pesquisa Agropecuária Brasileira**, v. 36, p. 79-88, 2001.

FINN, R.D., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., MISTRY, J., MITCHELL, A. L., POTTER, S. C., PUNTA, M., QURESHI, M., SANGRADOR-VEGAS, A., SALAZAR, G. A., TATE, J., BATEMAN A. The Pfam protein families database: towards a more sustainable future. *Nucleic acids research*, v. 44, p. D279-D285, 2015.

FOYER, C. H., NOCTOR, G. Ascorbate and glutathione: the heart of the redox hub. **Plant physiology**, v. 155, p. 2-18, 2011.

FOYER, C. H., NOCTOR, G. Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. **Physiologia plantarum**, v. 119, p. 355-364, 2003.

GARNCZARSKA, M., BEDNARSKI, W., & JANCELEWICZ, M. Ability of lupine seeds to germinate and to tolerate desiccation as related to changes in free radical level and antioxidants in freshly harvested seeds. **Plant Physiology and Biochemistry**, v. 47, p. 56-62, 2009.

GILL, S. S., TUTEJA, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant physiology and biochemistry**, v. 48, p. 909-930, 2010.

GOMEZ, J. M., JIMENEZ, A., OLMOS, E., SEVILLA, F. Location and effects of long - term NaCl stress on superoxide dismutase and ascorbate peroxidase isoenzymes of pea (*Pisum sativum* cv. Puget) chloroplasts. **Journal of Experimental Botany**, v. 55, p. 119-130, 2004.

HALLIWELL, B. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. **Plant physiology**, v. 141, p. 312-322, 2006.

HIGO, K., UGAWA, Y., IWAMOTO, M., KORENAGA, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. **Nucleic acids research**, v. 27, p. 297-300, 1999.

HONG, C. Y., HSU, Y. T., TSAI, Y. C., KAO, C. H. Expression of ASCORBATE PEROXIDASE 8 in roots of rice (*Oryza sativa* L.) seedlings in response to NaCl. **Journal of experimental botany**, v. 58, p. 3273-3283, 2007.

- HOSSAIN, M. A., ASADA, K. Monodehydroascorbate reductase from cucumber is a flavin adenine dinucleotide enzyme. **Journal of Biological Chemistry**, v. 260, p. 12920-12926, 1985.
- HU, B., JIN, J., GUO, A. Y., ZHANG, H., LUO, J., GAO, G. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics**, v. 31, p. 1296-1297, 2014.
- HUANG, L., SCHIEFELBEIN, J. Conserved gene expression programs in developing roots from diverse plants. **The Plant Cell**, v. 27, p. 2119-2132, 2015.
- ISHIKAWA, T., SAKAI, K., TAKEDA, T., SHIGEOKA, S. Cloning and expression of cDNA encoding a new type of ascorbate peroxidase from spinach. **FEBS letters**, v. 367, p. 28-32, 1995.
- ISHIKAWA, T., SAKAI, K., YOSHIMURA, K., TAKEDA, T., SHIGEOKA, S. cDNAs encoding spinach stromal and thylakoid-bound ascorbate peroxidase, differing in the presence or absence of their 3'-coding regions. **FEBS letters**, v. 384, p. 289-293, 1996.
- ISHIKAWA, T., YOSHIMURA, K., SAKAI, K., TAMOI, M., TAKEDA, T., SHIGEOKA, S. Molecular characterization and physiological role of a glyoxysome-bound ascorbate peroxidase from spinach. **Plant and cell physiology**, v. 39, n. 1, p. 23-34, 1998.
- JAILLON, O. *et al.* The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. **Nature**, v. 449, p. 463, 2007.
- KABIR, M. H., WANG, M-H. Expression pattern of two dehydroascorbate reductase genes from tomato (*Solanum lycopersicum* L.) in response to stress. **Journal of the Korean Society for Applied Biological Chemistry**, v. 53, p. 668-676, 2010.
- KANCHANA, P.; SANTHA, M. L.; RAJA, K. D. A review on *Glycine max* (L.) Merr (soybean). **WJPPS**, v. 5, p. 356-371, 2015.
- KARUPPANAPANDIAN, T., MOON, J. C., KIM, C., MANOHARAN, K., KIM, W. Reactive oxygen species in plants: their generation, signal transduction, and scavenging mechanisms. **Australian Journal of Crop Science**, v. 5, p. 709, 2011.
- KATAYA, A. RA; REUMANN, S. Arabidopsis glutathione reductase 1 is dually targeted to peroxisomes and the cytosol. **Plant signaling & behavior**, v. 5, p. 171-175, 2010.
- KIM, D., PERTEA, G., TRAPNELL, C., PIMENTEL, H., KELLEY, R., SALZBERG, S. L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome biology**, v. 14, p. R36, 2013.
- KIIHL, R. A. S.; GARCIA, A. The use of long juvenile trait in breeding soybean cultivars. In: **World Soybean Research Conference**. 1989.
- KIM, J., SUNDARESAN, S., PHILOSOPH-HADAS, S., YANG, R., MEIR, S., TUCKER, M. 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**, v. 6, p. 1109, 2015.

- KOBERG, M.; ABU-MUCH, R.; GEDANKEN, A. Optimization of bio-diesel production from soybean and wastes of cooked oil: Combining dielectric microwave irradiation and a SrO catalyst. **Bioresource technology**, v. 102, p. 1073-1078, 2011.
- KRAMER, P J., BOYER, J. S. **Water relations of plants and soils**. Academic press, 1995.
- KUMAR, S., STECHER, G., TAMURA, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular biology and evolution**, v. 33, p. 1870-1874, 2016.
- LANUBILE, A., MUPPIRALA, U. K., SEVERIN, A. J., MAROCCO, A., MUNKVOLD, G. P. Transcriptome profiling of soybean (*Glycine max*) roots challenged with pathogenic and non-pathogenic isolates of *Fusarium oxysporum*. **BMC genomics**, v. 16, p. 1089, 2015.
- LEISNER, C. P., MING, R., AINSWORTH, E. A. Distinct transcriptional profiles of ozone stress in soybean (*Glycine max*) flowers and pods. **BMC plant biology**, v. 14, p. 335, 2014.
- LETERRIER, M., CORPAS, F. J., BARROSO, J. B., SANDALIO, L. M., LUIS, A. Peroxisomal monodehydroascorbate reductase. Genomic clone characterization and functional analysis under environmental stress conditions. **Plant Physiology**, v. 138, n. 4, p. 2111-2123, 2005.
- LESCOT, M., DÉHAIS, P., THIJS, G., MARCHAL, K., MOREAU, Y., VAN DE PEER, Y., ROUZÉ, P., ROMBAUTS, S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic acids research**, v. 30, p. 325-327, 2002.
- LANGMEAD, B., SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, v. 9, p. 357, 2012.
- LI YH.; LI W.; ZHANG C, YANG L.; CHANG RZ.; GAUT B.S. QIU LJ. Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for single sequence repeat and single-nucleotide polymorphism loci. **New Phytologist**, v. 188, p. 242-253, 2010.
- LIU, Y. J., HAN, X. M., REN, L. L., YANG, H. L., ZENG, Q. Y. Functional divergence of the glutathione S-transferase supergene family in *Physcomitrella patens* reveals complex patterns of large gene family evolution in land plants. **Plant physiology**, v. 161, p. 773-786, 2013.
- LUDWIKOW, A., SADOWSKI, J. Gene networks in plant ozone stress response and tolerance. **Journal of integrative plant biology**, v. 50, p. 1256-1267, 2008.
- LUNDE, C., BAUMANN, U., SHIRLEY, N. J., DREW, D. P., FINCHER, G. B. Gene structure and expression pattern analysis of three monodehydroascorbate reductase (MdhAR) genes in *Physcomitrella patens*: implications for the evolution of the MDHAR family in plants. **Plant molecular biology**, v. 60, p. 259-275, 2006.
- MASCARENHAS, H.A.A.; TANAKA R. T.; WUTKE, E. B.; KIKUTI, H. Nitrogênio: a soja aduba a lavoura de cana. **O Agrônomo**. Campinas, p-12-13, 2005.

- MCGETTIGAN, P. A. Transcriptomics in the RNA-seq era. **Current opinion in chemical biology**, v. 17, p. 4-11, 2013.
- MITTLER, R., VANDERAUWERA, S., GOLLERY, M., VAN BREUSEGEM, F. Reactive oxygen gene network of plants. **Trends in plant science**, v. 9, p. 490-498, 2004.
- MÜLLER, L.; MIYASAKA, S.; MEDINA, J. C. Morfologia, anatomia e desenvolvimento. **A soja no Brasil**. Campinas: ITAL, 1981.
- MUNDSTOCK, C. M.; THOMAS, A. L. Soja: fatores que afetam o crescimento e o rendimento de grãos. **Porto Alegre: Universidade Federal do Rio Grande do Sul**, 2005.
- MUNNS, R. Comparative physiology of salt and water stress. **Plant, cell & environment**, v. 25, p. 239-250, 2002.
- MUTZ, K. O., HEILKENBRINKER, A., LÖNNE, M., WALTER, J. G., STAHL, F. Transcriptome analysis using next-generation sequencing. **Current opinion in biotechnology**, v. 24, p. 22-30, 2013.
- 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.
- NAGESH BABU, R.; DEVARAJ, V. R. High temperature and salt stress response in French bean (*Phaseolus vulgaris*). **Aust J Crop Sci**, v. 2, p. 40-48, 2008.
- NAJAMI, N., JANDA, T., BARRIAH, W., KAYAM, G., TAL, M., GUY, M., VOLOKITA, M. Ascorbate peroxidase gene family in tomato: its identification and characterization. **Molecular Genetics and Genomics**, v. 279, p. 171-182, 2008.
- NAYA, L., LADRERA, R., RAMOS, J., GONZÁLEZ, E. M., ARRESE-IGOR, C., MINCHIN, F. R., BECANA, M. The response of carbon metabolism and antioxidant defenses of alfalfa nodules to drought stress and to the subsequent recovery of plants. **Plant Physiology**, v. 144, p. 1104-1114, 2007
- NEILL, S., DESIKAN, R., HANCOCK, J. Hydrogen peroxide signalling. **Current opinion in plant biology**, v. 5, p. 388-395, 2002.
- NOGUEIRA, F. T. S., SASSAKI, F. T., MAIA, I. G. *Arabidopsis thaliana* uncoupling proteins (AtUCPs): insights into gene expression during development and stress response and epigenetic regulation. **Journal of bioenergetics and biomembranes**, v. 43, p. 71-79, 2011.
- ONAGA, G., WYDRA, K. Advances in plant tolerance to abiotic stresses. In: **Plant Genomics**. InTech, 2016.
- OZYIGIT, I. I., FILIZ, E., VATANSEVER, R., KURTOGLU, K. Y., KOC, I., ÖZTÜRK, M. X., ANJUM, N. A. Identification and comparative analysis of H₂O₂-scavenging enzymes (ascorbate peroxidase and glutathione peroxidase) in selected plants employing bioinformatics approaches. **Frontiers in plant science**, v. 7, p. 301, 2016.

- PANCHUK, I. I., ZENTGRAF, U., & VOLKOV, R. A. Expression of the *Apx* gene family during leaf senescence of *Arabidopsis thaliana*. **Planta**, v. 222, p. 926-932, 2005.
- PANDEY, P., ACHARY, V. M. M., KALASAMUDRAMU, V., MAHANTY, S., REDDY, G. M., REDDY, M. K. Molecular and biochemical characterization of dehydroascorbate reductase from a stress adapted C4 plant, pearl millet [*Pennisetum glaucum* (L.) R. Br]. **Plant cell reports**, v. 33, p. 435-445, 2014
- PANDEY, P., SINGH, J., ACHARY, V., REDDY, M. K. Redox homeostasis via gene families of ascorbate-glutathione pathway. **Frontiers in Environmental Science**, v. 3, p. 25, 2015.
- PASSARDI, F., TOGNOLLI, M., DE Meyer, M., PENEL, C., DUNAND, C. Two cell wall associated peroxidases from *Arabidopsis* influence root elongation. **Planta**, v. 223, n. 5, p. 965-974, 2006.
- PARK, S. Y., RYU, S. H., JANG, I. C., KWON, S. Y., KIM, J. G., KWAK, S. S. Molecular cloning of a cytosolic ascorbate peroxidase cDNA from cell cultures of sweetpotato and its expression in response to stress. **Molecular Genetics and Genomics**, v. 271, p. 339-346, 2004.
- PIGNOCCHI, C., FLETCHER, J. M., WILKINSON, J. E., BARNES, J. D., FOYER, C. H. The function of ascorbate oxidase in tobacco. **Plant Physiology**, v. 132, p. 1631-1641, 2003.
- PIMENTEL, C. A relação da planta com a água. **Seropédica: Edur**, 2004.
- PRABA, M. L., CAIRNS, J. E., BABU, R. C., LAFITTE, H. R. Identification of physiological traits underlying cultivar differences in drought tolerance in rice and wheat. **Journal of Agronomy and Crop Science**, v. 195, p. 30-46, 2009.
- PRINCE, S. J., JOSHI, T., MUTAVA, R. N., SYED, N., VITOR, M. D. S. J., PATIL, G., SONG, L., WANG J. J., LIN, L., CHEN, W., SHANNON, J. G., VALLIYODAN B., XU, D., NGUYEN H. Y. Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting. **Plant Science**, v. 240, p. 65-78, 2015.
- QIN, A., SHI, Q., YU, X. Ascorbic acid contents in transgenic potato plants overexpressing two dehydroascorbate reductase genes. **Molecular biology reports**, v. 38, p. 1557-1566, 2011.
- RIBEIRO, C. W., KORBES, A. P., GARIGHAN, J. A., JARDIM-MESSEDER, D., CARVALHO, F. E., SOUSA, R. H. V., CAVERZAN, A., TEIXEIRA F. K., SILVEIRA J. A. G., MARGIS-PINHEIRO, M. Rice peroxisomal ascorbate peroxidase knockdown affects ROS signaling and triggers early leaf senescence. **Plant Science**, v. 263, p. 55-65, 2017.
- RITCHIE, S. W., HANWAY, J. J., THOMPSON, H. E. **How a soybean plant develops**. Iowa State University of Science and Technology, Cooperative Extension Service, 1985.
- ROCHA, R. S. **Avaliação de variedades e linhagens de soja em condições de baixa latitude. 2009, 59p.** 2009. Dissertação (Mestrado) - Universidade Federal do Piauí, Piauí.
- RUBIO, M. C., BUSTOS - SANMAMED, P., CLEMENTE, M. R., BECANA, M. Effects of

salt stress on the expression of antioxidant genes and proteins in the model legume *Lotus japonicus*. **New Phytologist**, v. 181, p. 851-859, 2009.

SAITOU, N., NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular biology and evolution**, v. 4, p. 406-425, 1987.

SAKATA, T., HIGASHITANI, A. Male sterility accompanied with abnormal anther development in plants—genes and environmental stresses with special reference to high temperature injury. **International Journal of Plant Developmental Biology**, v. 2, p. 42-51, 2008.

SANO, S., TAO, S., ENDO, Y., INABA, T., HOSSAIN, M. A., MIYAKE, C., MATSUO, M., AOKI, H., ASADA, K., Saito, K. Purification and cDNA cloning of chloroplastic monodehydroascorbate reductase from spinach. **Bioscience, biotechnology, and biochemistry**, v. 69, p. 762-772, 2005.

SCHENA, M., SHALON, D., DAVIS, R. W., BROWN, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science**, v. 270, n. 5235, p. 467-470, 1995.

SCHMUTZ, J. *et al.* Genome sequence of the palaeopolyploid soybean. **nature**, v. 463, p. 178, 2010.

SEČENJI, M., LENDVAI, Á., MISKOLCZI, P., KOCSY, G., GALLÉ, Á., SZŰCS, A., HOFFMANN, B., SÁRVÁRI, P., SCHWEIZER, P., STEIN, N., DUDITS, D., GYORGYEY, J. Differences in root functions during long - term drought adaptation: comparison of active gene sets of two wheat genotypes. **Plant Biology**, v. 12, p. 871-882, 2010.

SHARMA, P., JHA, A. B., DUBEY, R. S., PESSARAKLI, M. Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. **Journal of botany**, v. 2012, 2012.

SHI, W. M., MURAMOTO, Y., UEDA, A., TAKABE, T. Cloning of peroxisomal ascorbate peroxidase gene from barley and enhanced thermotolerance by overexpressing in *Arabidopsis thaliana*. **Gene**, v. 273, p. 23-27, 2001.

SHIMAOKA, T., YOKOTA, A., MIYAKE, C. Purification and characterization of chloroplast dehydroascorbate reductase from spinach leaves. **Plant and Cell Physiology**, v. 41, p. 1110-1118, 2000.

SHIN, J. H., VAUGHN, J. N., ABDEL-HALEEM, H., CHAVARRO, C., ABERNATHY, B., DO KIM, K., JACKSON, S. A., LI, Z. Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. **BMC plant biology**, v. 15, p. 26, 2015.

SHRIVASTAVA, P., KUMAR, R. Soil salinity: a serious environmental issue and plant growth promoting bacteria as one of the tools for its alleviation. **Saudi journal of biological sciences**, v. 22, p. 123-131, 2015.

SINGH, R., SINGH, S., PARIHAR, P., MISHRA, R. K., TRIPATHI, D. K., SINGH, V. P.,

CHAUHAN, D. K., PRASAD, S. M. Reactive Oxygen Species (ROS): beneficial companions of plants' developmental processes. **Frontiers in plant science**, v. 7, 2016.

SOUSA, A. D., BEZERRA, M. A., FARIAS, F. C. Germinação e desenvolvimento inicial de clones de cajueiro comum sob irrigação com água salina. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 15, p. 390-394, 2011.

STEVENS, R., BURET, M., DUFFÉ, P., GARCHERY, C., BALDET, P., ROTHAN, C., CAUSSE, M. Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. **Plant Physiology**, v. 143, p. 1943-1953, 2007.

SYTYKIEWICZ, H. Expression patterns of genes involved in ascorbate-glutathione cycle in aphid-infested maize (*Zea mays* L.) seedlings. **International journal of molecular sciences**, v. 17, p. 268, 2016

TAIZ, L.; ZEIGER, E. Fisiologia Vegetal. 5th edn (Porto Alegre: Artmed). 2013.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M., KUMAR, S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Molecular biology and evolution**, v. 28, p. 2731-2739, 2011.

TARI, I., LASKAY, G., TAKÁCS, Z., POÓR, P. Response of sorghum to abiotic stresses: a review. **Journal of Agronomy and Crop Science**, v. 199, n. 4, p. 264-274, 2013

TAUSZ, M., GRULKE, N. E., WIESER, G. Defense and avoidance of ozone under global change. **Environmental Pollution**, v. 147, p. 525-531, 2007.

TEIXEIRA, F. K., MENEZES-BENAVENTE, L., GALVÃO, V. C., & MARGIS-PINHEIRO, M. Multigene families encode the major enzymes of antioxidant metabolism in *Eucalyptus grandis* L. **Genetics and Molecular Biology**, v. 28, p. 529-538, 2005.

TEIXEIRA, F. K., MENEZES-BENAVENTE, L., GALVÃO, V. C., MARGIS, R., & MARGIS-PINHEIRO, M. Rice ascorbate peroxidase gene family encodes functionally diverse isoforms localized in different subcellular compartments. **Planta**, v. 224, p. 300, 2006.

THE C. ELEGANS SEQUENCING CONSORTIUM. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. **Science**, p. 2012-2018, 1998.

TOGNOLLI, M., PENEL, C., GREPPIN, H., SIMON, P. Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. **Gene**, v. 288, p. 129-138, 2002.

TOMMASI, F., PACIOLLA, C., ARRIGONI, O. The ascorbate system in recalcitrant and orthodox seeds. **Physiologia Plantarum**, v. 105, p. 193-198, 1999.

TRAPNELL, C., SALZBERG, S. L. How to map billions of short reads onto genomes. **Nature biotechnology**, v. 27, p. 455, 2009.

TRIVEDI, D. K., GILL, S. S., YADAV, S., TUTEJA, N. Genome-wide analysis of glutathione

reductase (GR) genes from rice and Arabidopsis. **Plant signaling & behavior**, v. 8, p. e23021, 2013.

UNITED STATES DEPARTMENT OF AGRICULTURE. **World agricultural production**. Foreign Agricultural Service/USDA, 2018.

VALERO, E., MACIÀ, H., ILDEFONSO, M., HERNÁNDEZ, J. A., GONZÁLEZ-SÁNCHEZ, M. I., GARCÍA-CARMONA, F. Modeling the ascorbate-glutathione cycle in chloroplasts under light/dark conditions. **BMC systems biology**, v. 10, p. 11, 2015.

VAN VERK, M. C., HICKMAN, R., PIETERSE, C. M., VAN WEES, S. C. RNA-Seq: revelation of the messengers. **Trends in plant science**, v. 18, p. 175-179, 2013.

VOLLENWEIDER, P., GÜNTHARDT-GOERG, M. S. Diagnosis of abiotic and biotic stress factors using the visible symptoms in foliage. **Environmental Pollution**, v. 137, p. 455-465, 2005.

WANG, Y., WU, J., CHOI, Y. W., JUN, T. H., KWON, S. W., CHOI, I. S., KIM, Y. C., GUPTA, R., KIM, S. T. Expression Analysis of *Oryza sativa* Ascorbate Peroxidase 1 (*OsAPx1*) in Response to Different Phytohormones and Pathogens. **Journal of Life Science**, v. 25, p. 1091-1097, 2015.

XU, G., GUO, C., SHAN, H., KONG, H. Divergence of duplicate genes in exon–intron structure. **Proceedings of the National Academy of Sciences**, v. 109, p. 1187-1192, 2012.

YOSHIDA, S., TAMAOKI, M., SHIKANO, T., NAKAJIMA, N., OGAWA, D., IOKI, M., AONO, M., KUBO, A., KAMADA, H., INOUE, Y., SAJI, H. Cytosolic dehydroascorbate reductase is important for ozone tolerance in *Arabidopsis thaliana*. **Plant and cell physiology**, v. 47, p. 304-308, 2006.

YOUSUF, P. Y., HAKEEM, K. U. R., CHANDNA, R., AHMAD, P. Role of glutathione reductase in plant abiotic stress. In: **Abiotic Stress Responses in Plants**. Springer New York, 2012. p. 149-158.

YU, C. S., CHEN, Y. C., LU, C. H., HWANG, J. K. Prediction of protein subcellular localization. **Proteins: Structure, Function, and Bioinformatics**, v. 64, p. 643-651, 2006.

YU, CS., LIN, CJ., HWANG, JK., Predicting subcellular localization of proteins for Gram - negative bacteria by support vector machines based on n - peptide compositions. **Protein science**, v. 13, p. 1402-1406, 2004.

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

ANEXO A: SEQUÊNCIAS DAS PROTEÍNAS DEDUZIDAS DOS DIFERENTES MEMBROS GÊNICOS DA APX, MDHAR, DHAR E GR UTILIZADAS PARA A FORMAÇÃO DA ÁRVORE FILOGENÉTICA.

>Ad-APX-4a

MGKSYPTVSEDEYKAIQKARRKLRGLIFEKQCAPLMLRLAWHSAGTFDVKTGTGGPF
GTMKHEAEQKHEANRGLDIAVNLLLEPLKQHFPNITYADFYQLAGVVAVEVTGGPEVP
FHPGREDKADPPPEGRLPDATKGNHDLRELQVQTMGLTDQDIVALSGAHTLGRCHKD
RSGFEGPWTSNPLVFDNSYFTELVGGEREGLLQLPTDKALLEDPVFRPLVEKYASDED
AFFADYTAHLKLSLSEGFADA

>Ad-APX-5b

MALPVVVDAYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDVNTKTGG
ANGSIRNEEYSHGANGLKKAIDFCEEVKAKYSKISYADLYQLAGVVAVEVTGGPT
VEFVPGRKDSKVSTKEGRLPDAKLGVSHLRSIFYRMGLSDKDIVALSGAHTLGRAHQ
ERSGFEGPWTDPLKFDNSYFVELLKGESAGLLQLPTDKALLEDPEFRRYVELYAKDE
DAFFRDYAESHKKLSELGFATSSKPIPKDGTILAQSAVGVVVTAAVVILGYLYEVHKR
GK

>Ad-APX-4b

MGKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLMLRLAWHSAGTFDVATKSGGP
FGTIKHPSELAHGANAAGLDIAVRLLEPIKEQFPTLSYADFYQLAGVVAVEITGGPEIPFH
PGREDKPEPPPEGRLPDATKGSDDLDRDVFVKAMGLSDQDIVALSGGHTLGAHKERS
GFEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDPVFRPLVEKYAADEDAFF
ADYAEHLKLSLSEGFADA

>Ad-APX-5a

MAKEGVAVDFEYAKQIDKARRELRALISSRNCAPLMLRLAWHDAGTYDAKTRTGGA
NGSIRNQELNHQANRGLTAVQFCEEVKAKCPRVSYADLYQLAGVIAVEVTGGPTIR
FVPGRKDSLESPPPEGRLPDATKDAQHLRDVFYRMGLTDKDIVALSGGHTLGAHKERS
SGFEGQWTENHLKFDNSYFVELMDSKDSFKLPTDKVLVEDPEFRKYVVRYKNDEDA
FFADYAESHKKLSELGFNPNRVSDSRIVRGVIGFLVASTVIVMGYLIEINRKS

>Ad-APX-3

MAERVFLTHPSLTLPSQPFPTRMASTATAAIAASRIIPSTATALSSSASRSFFSSSSTSRSSL
KCLHLHSSSRISHLFVNQRRAEVRVSSGDYATVSSPKSLASDPDQLKSAREDIKELLRT
KFCHPILIRLGWHDAGTYNKNIEEWPQRGGANGLRFEVELKHAANAGLVNALKLIQ
PIKDKYARVYADLFQLAGATAVEEAGGPKIPMKYGRVDVSSPEQCPEEGRLPDAGPP
SPADHLRQVFYRMGLNDKEIVALSGAHTLGRSRPERSGWGKPKETKYTKDGPAPGG
QSWTVQWLKFDNSYFKYIKEKKDEDLLVLPDVAFLFEDPSFKVYAEKYAEDQVAFFN
DYAEAHAKLSNLGAKFDPPEGIVLDPSPKQAQAEKFEAAKYSTGKD

>Ad-APX-2

MSSTLTNHHPLLCSISTASASAQQHKPLTNFPAKSQRSTFSTVNFLPNHPPPPSSGPNL
VVGRRGLVSIIVTLPLVPLTQIMGCLQAYAAQPSTKEYLLIKEEVRKVLKSKGAAGVL
RLVFHDAGTFDMDDKTGGMNGSIVYELERPENSGLKKSIVLQKAKIQIDAIQPVSW
GDMIAVAGAEAIELCGGPSIQVSLGRLDSQGPDPPEGKLPESLNASSLKKCFQRKGFST
QELVALSGAHTLGSKGFGSPTSFDNAYYKILLEKPWANPGGMSSMIGLPSDHALVEDD

ECLRWIKKYAENENIFFEDFKNAYVKLVNSGVRWNNL

>Ad-APX-1

MGVSFLSTLPSLLPLAHDSSPTTTTTIRHHPRLVTCGYNKIKASYHVIDEGHFSLKRRD
ILKSLGMTIGLELIGSSGSLVGTANAADLIQRRQRSEFQSKIKGTLFTAIGNPDVPSLL
TLALNDAVTYDKATKSGGPNGSIRFSSEISRPENTRLSGALSLIEEAKKEIDSYSKAGPI
SYADLIHYAAQSAVKSTFIAAAIRKCGGNEEKGNLLYTAYGSNGQWGLFDKQFGRED
AQEPDPEGRVPLWEKASVQEMKDKFSAVGLGPRQLAVLSAFLGPDQAATEAKLASDK
DVAPWVDKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYAPPKIDITKLKL

>Ai-APX-5b

MALPVVVDAEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDVNTKTGG
ANGSIRNEEYSHGANGLKKAIDFCEEVKAKYSKISYADLYQLAGVVAVEVTGGPT
VEFVPGRKDSKVSTKEGRLPDAKLGVSHLRSIFYRMGLSDKDIVALSGAHTLGRAHQ
ERSGFEGPWTDPLKFDNSYFVELLKGESAGLLQLPTDKALLEDPEFRRYVELYAKDE
DAFFRDYAESHKKLSELGFATSSKPIFPKDGTLAQS AVGVVVTAAVVILGYLYEVHKR
GK

>Ai-APX-4b

MGKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLMLRLAWHSAGTFDVATKSGGP
FGTIKHPSELAHGANAGLDIAVRLLLEPIKEQFPTLSYADFYQLAGVVAVEITGGPEIPFH
PGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDQDIVALSGGHTLGAHKERS
GFEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDPVFRPLVEKYAADEDAFF
ADYAE AHLK LSELGFADA

>Ai-APX-4a

MGKSYPTVSEDEYKAIEKARRKLRGLIFEKQCAPLMLRLAWHSAGTFDVKTGTGGPF
GTMKHEAEQKHEANRGLDIAVNLEPLKQHFPNITYADFYQLAGVVAVEITGGPEVPF
HPGREDKADPPPEGRLPDATKGN DHLREL FVQTMGLTDQDIVALSGAHTLGRCHKDR
SGFEGPWTSNPLVFDNSYFTELVGGERDGLLQLPTDKALLEDPVFRPLVEKYAKDEDA
FFADY TQAHLK LSELGFADA

>Ai-APX-5a

MAKEGVAVDFEYAKQIDKARRELRALISSRNCAPLMLRLAWHDAGTYDAKTRTGGA
NGSIRNQELNHQANRGLETAVQFCEEVKAKCPRVSYADLYQLAGVIAVEVTGGPTIR
FVPGRKDSLESPPPEGRLPDATKDAQH LRDV FYRMGLTDKDIVALSGGHTL GKAKKER
SGFEGQW TENHLKFDNSYFVELMDSKDSFKLPTDKVLVEDPEFRKYVVRYKNDEDA
FFADYAVSHKKLSELGFNPNRVSDSRIVHGVIGFLVASTVIVMGYLIEINRKS

>Ai-APX-2

MSSTLTNHHPLLCSISTASASAQQHKPLTNFPAKSQRPTFSTVNFHPNHPPPSSSGPNL
VVGRRGLVSIATLPCLVPLTQIMGCLQAPTKEYLLIKEEVRKVLSKGKAAGVLRLVF
HDAGTFDMDDKTGGMNGSIVYELERPENSGLKKSIKVLQKAKIQIDAIQVSWGDMI
AVAGAEAIELCGGPSIQVSLGRLDLGPDPGKLP EESLSASSLKKCFQRKGFSTQELV
ALSGAHTLGSKGFSGPTSFDNAYYKILLEKPWANPGGMSSMIGLPSDHALVEDNECLR
WIKKYAENENIFFEDFKNAYVKLVNSGVRWNSL

>Ai-APX-1

MGVSFLSLFLLFFHLLMIHHLPPPPSDTTLVTCGYNKIKASYHVIDEGHFSLKRRDIL
KSLGMTIGLELIGSSGSLVGTANAADLIQRRQRSEFQCKIKGTLFTAIGNPDVPSLLT

LALNDAV TYDKATKSGGPN GSIRFSSEISR PENRR LSGALSLIEEAKKEIDSY SKAGPIS
 YADLIHYAAQS AVKSTFIAAAIRKCGGNEEKGNLLYTAYG SNGQWGLFDKQFGREDA
 QEPDPEGRVPLWEKASVQEMKDKFS AVGLGPRQLAVLSAFLGPDQAATEAKLASDKD
 VAPWVDKYQRSRET VVSQTDYEVDLITFTKLSTLGQQINYEAYTYAPPKIDITKLKL

>Cc-APX-4a

MGKSYPTVSADYQKAVDKAKKRLRGLIAEKRCAPLMLRIAWHSAGTYDVKTKSGGP
 FGTIKHPAELAHGANGLDIAVRLLEPVKA EFPILSYADFYQLAGVVAVEVTGGPEIPF
 HPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDQDIVALSGGHTLGSCHKER
 SGFEGPWTTNPLIFDNSYFKELLSGEKEGLIQLPTDKSLLSDPVFRPLVEKYAADEDAF
 FADYAESHKLKSELGFAEA

>Cc-APX-5a

MALVDEEYLEEIEKARRQLRLLITINKCAPLMLRLAWNDACTYDAQTRTGGANASIR
 TPQELKHQANIGLERAVRLCEEVKAKLKKVSYADLYQLAGVVAIELTGGPTINFYPGR
 QDSDECPPEGRHLSGEEDAKSLRNTFSRMGLRDDKHIVALCGGVRTISYTKGPYQW
 TKDPLKFDNSYFVLPMDALVKDEKFRHFVELYAKDEDTFKDYAMSHEKLSSELGCN
 LKKPNQPKGSFEKLNQPKGLIGLVLVVTVILGYLMKKRKKNLKVEEDLDTNL

>Cc-APX-5b2

MVKATVGGEYLKEFEKARRDLRALISNRNCAPLMLRLAWHDAGTYDAKTKTGPN
 GSIRTEEELNHAANKGLRTAVEFC EEVKAKHPKVSYADLYQLAGVVAVEVTGGPTIDF
 LPGRKDSLESPAEGRLPDAKQGASHLREIFYRMGLGDRDIVALSGGHTLGKAHKDRS
 DFHGQWTKDPLKFDNSYFVELLKGDSKDLLKLPTDKALVEDLKFRKYVELYAKDED
 AFFADYASSHKKLSSELGFNVKDRSMLVKG AIGIAIALTAVILGYLRELNINKTS

>Cc-APX-5b1

MALPVVVD AEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGPN
 NGSIRNEEEYSHGANSGLKKAIDFCEEVKAKHPKITYADLYQLAGVVAVEVTGGPTID
 FVPGRRDSKISPKEGRLPDAKQGVPHLRDIFYRMGLNDQDIVALSGGHTLGRAHPERS
 GFEGPWTE DPLKFDNSYFVELLKEDYAMLLKLPTDKALLEDPNFRRFVELYAKDEDA
 FFRDYAESHKKLSSELGFVPSSRATCVKDG TILAQGAVGVVVTAAVVILSYLYEVRKRG
 K

>Cc-APX-4b

MGKSYPTVSADYQKAVDKAKKRLRGLIAEKRCAPLMLRIAWHSAGTYDVKTKGGXP
 AELAHGANGLDIAVRLLEPVKA EFPILSYADFYQLAGVVAVEVTGGPEIPFHPGRED
 KPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDQDIVALSGGHTLGSCHKERSGFEGP
 WTTNPLIFDNSYFKELLSGEKEGLIQLPTDKSLLSDPVFRPLVEKYAADEDAFFADYAE
 SHLKLSELGGKGGKRRKGEKGREEGCNRRR RRRSDGGRDRQRRHRRRRRATGLRK
 GEEKKRRKRKREEEEEEEKKEKKKSNTIGGMMGGPTRRRRRSGVMREEEKNKRKR
 RRRRRNVPEKKEKVEKNERNTYLLIYFINY

>Cc-APX-3

MAERVLLTPPSLPLPSQHSPTTMALLGGAASARVIPSASLSASSRSFFSFSSSSSLQCLRS
 SPRISHLFLNQRRAEVRVSSGGYGTVSAPKSFASDPDQLKSAREDIKELLSKFCHPILI
 RLGWH DAGTYNKNIEEWPQRGGANGLRFEIELKHAANAGLVNALKLLQPIKDKYS
 GVTYADLFQLAGATAVEEAGGPQIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLR
 QVFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGK PETKYTKDGP GAPGGQSWTAQ
 WLKFDNSYFKDIKEKKDEDLLVLPTDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAH
 AKLSNLGAKFDPPEVCSF

>Cc-APX-2

MRLGVSKDVVRCSIGTTPGESKTHFPAKLQPSSFSTVKFRTDNASSDLAVSSSSRRGLI
 CIATLPCLLPLIGSLPANAMQPGAKEYLLIKEEVRKVLKAKAAGVLRVLFHDAGTFDI
 DDNTGGMNGSIVYELERPENAGLKKSIAKACFVLQKAKTQIDAIQPVSWADMIAVAGA
 EAVEVCGGPPIQVSLGRIDTLVPDPEQRLPEESLGASDLKCKCFQRKGFSTQELVALSGA
 HTLGSKGFGSPISFDNSYYKVLLEKPWTSPGGMSSMVGLPSDHALVEDNECLRWIKK
 YADSENLFEDFKNAYVKLVNSGVRWNGV

>Cc-APX-1

MSNSTDLPWLSFSILTAQVRRTLFTAIGNPDLTPSILKLALNDALTYEKATKSGGPNG
 SIRFSSEIGRPENTGLSAAVSLLEEAKKEIDSYAKGGPISYADLIQYAAQSAVKSTFLAA
 AIRKCGGNEEKGNILYNAYGSNGQWGLFDKQFGRADAQDPDPEGRVPIWEKASVQE
 MKDKFSAVGFGRQVAVLSAFLGPDQSATETLLASDPDVAPWVQKYQRSRETVSQTD
 YEVDLITTFTKLVFPENCSTQAYSLNRQPKFLIKPMRTSPNVWEFYL

>Ca-APX-4

MGKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLILRLAWHSAGTFDVKTGTGGPF
 GTIKHQAEALAHGANSGLDIAVRLLEPLKEQFPIISYADFYQLGGVVAVEVTGGPEVPFH
 PGREDKPEPPPEGRLPDATKGSDDLDRDVFVKAMGLSDQDIVALSGGHTIGAAHKERS
 GFEGPWTSNPLIFDNSYFTELLNGEKEGLLKLPSDKALLSDPVFRPLVEKYAADEDAFF
 ADYAEHLKLSELGFAEA

>Ca-APX-5b2

MAQDAEYVKEIQEARAQLASLITNKKCAPMLRLAWHDAGTYDAKTRTGGPNGSIR
 NQQELNHDANKGLKTAVELCGEVKVKHPKISYADLYQLAGVVAVEVTGGPTIHFLPG
 RKDSLESPQEGRLPDAKKGASHLRDIFYRMGLNDRDIVALSGGHTLGKAHKDRSDFE
 GQWTRDPNKFDNSYFVELLNSKSKELLKLPTDKALIEDSVFRKYVELYAKDEDAFFR
 DYAISHKKLSELGFINPNCNYRSKLAKGVIGFAIASTVVILGYLLELNKKIY

>Ca-APX-5b1

MALPVVVDAEYLKEIDKARRDLRALIANRNCAPIMLRLAWHDAGTYDAKTKTGGPN
 ASIRNEEEFSDANNGLKKAIDFCEEVKAKHPKISYADLYQLAGVVAVEVTGGPTVDF
 TPGRRDSKVSTREGRLPDAKQGVPHLRDIFYRMGLTDKDIVALSGAHTLGRAHPERS
 GFDGPWTEPLKFDNSYFVGLLKEDATGLLKLPTDMALLDDPEFLKYVELYAKDEDA
 FFRDYAESHKKLSELGFVPSSKATSPKDGTLAQS AVGA AVTAAVVILGYLYEVHKRG
 K

>Ca-APX-2

MMCSSVSKHVVC SIGTASGVKSQACSFRSDEHGSSSV PSSRRGVLCIATLPCLLPLT
 HIFASLPANAISVTPATNEYLRKQEVKVLKSKGAAGVLRVLFHDAGTFQIDDTTGG
 MNGSIVYELERPENAGLKKSIVLQKAKTQIDAIHPVSWADIISVAGAEAVEVCGGPAIQ
 VRLGRQDSLGPDSEGKLPEETLDASGLKCKCFQKKGSTQELVALSGAHTLGSKGFSP
 TSFDNSYYKVLLEKPWISSGGMPSMIGLPSDHALVEDDECLRWIKNYADNENMFFED
 FRNAYVKLVNSGVRWNSL

>Ca-APX-3

MAERVLLTPSLPLPTTRPSTTTMATHISGTAAARMIPSA TRATVSLSSSSRSFFSFSSSSSV
 SSLQCLRSSPRISHLFLDQRRADV RVSSGRYGTASVTKALASDPDQLKSAREDIKELLR
 TKYCHPLMIRLGWHDAGTYNKNIEEWPQRGGANGLRFEAELKHGANAGLVNALK

LLQPIKDKYSGVTYADLFQLAGATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDA
 GPPSPADHLRQVFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGGPAP
 GGQSWTAQWLKFDNSYFKDIKEKKDEDLLVLPTDAALFEDPSFKVYAEKYAEDQEAF
 FKDYAEAHAKLSNLGSKFDPPEGIVIDGSPNAEKFVAAKYSSGKE

>Ca-APX-1

MGVSFLSISTLPSLLPPLVAHNPKSTLTQYQPRSVaicCTKINTDVSDEEEQFHWKRRDI
 LKcIGVTIGLELITNSGPLVGMANAADLIERRQRSEFQSQIKGTLYTAIKGNPDIVPSLL
 TLALNDALTYDKATKTGGPNgsIRFSSEINRPENKGLSAALNFIEEAKKEIDSYSKGGPI
 SYSDLIQYAAQSATKATFLASAIRKCGGNEEKGNLLYTAYGSNGQWGLFDRQFGRAD
 TQEPDPEGRIWdkASVQEMKDKFSAVGLGPRQYAEVRRYTLFKNSVV

>Gm-APX-4b

MgKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLMLRLAWHSAGTYDVSSKTGGP
 FGTIKHPSELAHGANNGLDIAVRLLEPLKAefPILSYADfyQLAGVVAVEVTGGPEVFPF
 HPGREDKPEPPPEGRLPDATKGSdHLRDVFGKAMGLSDRDiVALSGGHTIGAAHKER
 SGFEGPWTSNPLIFDnsYfKELLSGEKEGLLQLPSDKALLSDPVFRPLVEKYASDEDaf
 FADYAEAHQKLSELGFAEA

>Gm-APX-4a

MgKSYPTVSADYQKAVEKAKKKLRGFIAEKRCAPLMLRLAWHSAGTFDKGKTGGP
 FGTIKHPAELAHSANGLDIAVRLLEPLKAefPILSYADfyQLAGVVAVEVTGGPEVFPF
 HPGREDKPEPPPEGRLPDATKGSdHLRDVFGKAMGLTDQDiVALSGGHTIGAAHKER
 SGFEGPWTSNPLIFDnsYfTELLSGEKEGLLQLPSDKALLSDPVFRPLVDKYAADEDaf
 FADYAEAHQKLSELGFADA

>Gm-APX-5b1

MALPVVVDaEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGGP
 NGSIRNEEEYSHGANGLKKAIDFCQEVKAKYPKITYADLFQLAGVVAVEVTGGPTID
 FVPGRRDSKvSPNEGRLPDAKKGVPHLRDIFyRMGLTDRDiVALSGGHTLGRAHPERS
 GFDGPWTEdPLKFDNSYfVELLKEDSAGLLKLPTDKALLEDaEFRCYVELYAKDEDA
 FFRDYAESHKKLSSELGFVPSSKPIsIKEGTILAQGAVGVVVATAVVILSYLYEVRKRgK

>Gm-APX-5b2

MAKPRVDAEYLKEIEKARRDLRALISNRNCAPLMLRLAWHDAGTYDAKTNTGGPNg
 SIRNRQELNHAANKGLEtALAFCEEVKAkHPKISYADLYQLAGVVAVEVTGGPTINFV
 PGRKDSLESPAEGRLPDAKQGASHLRDIFyRMGLGDkDiVALSGGHTLgKAHKDRSD
 FHGQWTKDPLKFDNSYfVELLRGESKDLLKLPTDKALVEDPNFRKYVELYAKDEDAf
 FSDYATSHKKLSSELGFIFKNHRSILAKGVIGIAIVLTAVILGYLRELNKkTN

>Gm-APX-5a1

MAEPVVDDEYLKEIDKARRELRAFITSNQCAPLMLRLAWNDAAATYDARNRAGGPNG
 SIRTDKELKHEANEGLLKATQLCEHVKAkLKKVSYADLYQLAGVVAIEVSGGPTIEFL
 PGRKDSMESSAEGLLPDVKQGASiIRNIFSRMGISDDKHIVALCGGLTWGETLKDRSDS
 KGQWPkdPLKFDNSYyKkILSKDLSSRLPIEDALLTDQSFRRHVEEYSKDENSFFKEY
 AMSHKKLSSELGYNLKKQnQSKVPDEEQNQPkgPYEKLNQHRGLIGIGIVSVVTVILG
 YLLKRkKNQLKK

>Gm-APX-3b

MAERVSLPLPSQSQLSPTTTMALLGAASSSARIIPSVSLSSSSRSFFSLSSSSSKLQCLRS
 SPRISHLFLNQRRAEVRVSSGGYGTVSAPKSVASDPDQLKSAREDIKELLNSKFCHPILI

RLGWH DAGTYNKNIEEWPQRGGANGLRFEIELKHGANAGLVNALKLLQPIKDKYS
 GVTYADLFQLAGATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLR
 QV FYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGP GAPGGQSWTVQ
 WLKFDNSYFKDIKEKKDEDLLVLP TDAALFEDPSFKVYAEKYAEDQE AFFKDYAEAH
 AKLSNLGAKFDPPEGIVIDDSPNAGAEKFVAAKYSTGKE

>Gm-APX-3a

MAERVLLLTPSQSHHSPTTMAVLGGAASARIIPSVSLSTSSRSFFSLSSSSKLCRSPR
 ISHLFLNQQRAEVRVSSGGYGTVSAPKSVASDPDQLKSAREDIKELLNSKFCHPILIRL
 GWH DAGTYNKNIEEWPLRGGANGLRFEVELKHGANAGLLNALKLLQPIKDKYSGV
 TYADLFQLASATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLRQV
 FYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGP GAPGGQSWTVQWL
 KFDNSYFKDIKEKRDEDLLVLP TDAALFEDPSFKVYAEKYAEDQE AFFKDYAEAHAK
 LSNLGAKFDPLEGIVIDDSPNAGGEKFVAAKYSTGKE

>Gm-APX-2

MRLGVS VSVRCSIGTASGNTQTDYPAKFQRSSFSTANFLPDNASSDLAVSISSSSSRRG
 LIRIATLPCLLPLIGCLPANAMQPGSKEYLLIKEEVRKVL SKGKAAGVLR LRVFHDAGTF
 DIDDSTGGMNGSIVYELERPENAGLKKSVKVLQKAKTQIDAIQPVSWADMIAVAGAE
 AVEVCGPPIQVSLGRLDTLVPDPEGRLPEESLNASGLKKCFQSKGFSTQELVALSGAH
 TIGSKGFGSPISFDNSYYKVLLEKPWTSSGGMP SMIGLPSDHALVEDDECLRWIKKYA
 DSENLFFEDFKNAYVKLVNSGVRNSL

>Gm-APX-5a2

MVEPIVVDEEYRKEIEMARRELGVFITNNKCAPLMLQFAW NDAATYDAKSRRGGPN
 GSIRLRIGQELKHEANKGLEKAVQYCEIVKTKLKLKVS YANLYQDSNESPRTEGRFI
 DGEEDARNLRKIFSRMGLSDEQDIVALCGGHTLIRTMYPKVS MGETHKDRSKFEEGK
 STNKPLKFDNSYFKELLIK DASFSRLPMDYALVEDPKFH HYVERYDEEIFFKEYAISHK
 KLSELGFNLNLDQPKGPYAKLNQHKGLIGIGIASVVVTVILGYLLKRKKNQLKN

>Gm-APX-1a

MGV SFLSTCPSLLPLVAYDTPPSARYQPRSVSICCNKTISDVSDGEPFHWRRRDV LKCA
 GATVGLELIGSSGSLVEMANAADLIQRRQRSEFQSKIKETLFTA IKGNDLIPSLTLAL
 NDALTYDKATKSGGPN GSIRFSSEIGRPENTGLS AALNLEEAKKEIDSDSKGGSISYA
 DLIHIAAQSAAKATFLAAAIRKCGGKEEKGKTLYNAYGSNGQWGLFDRQFGRADAQ
 DPDPEGRVPLWEKASVQEMKDKFVAVGFGPRQLAVLSAFLGPDQ NATETLLASDPDV
 APWVQKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYPPPKIDITKLL

>Gm-APX-1b

MGV SFLSTCPSLLPLVAYDTPPSTRYQPRSVSICCNKTISNVSD EEPFHWRRREILK CAG
 ATVGLELIGSSGSLVGMANAADLIQRRQRSEFQSKVKETLFTA IKGNDLIPSLTLAL
 NDALTYDKATKSGGPN GSIRFSSEIGRPENTGLS AAMNLEEAKKEIDSYSKGGPISYA
 DLIHFAAQSAAKATFLAAAIRKCGGNEEKGKTLYNAYGSNGQWGLFDRQFGRADTQ
 DPDPEGRVPIWEKASVQEMKDKFVAVGFGPRQLAVLSAFLGPDQ TATETLLATDPDVA
 PWVQKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYPPPKIDITKLL

>Gs-APX-4b

MGKSYPTVSADYQKAVEKAKKKLIGFIAEKRCAPLMLRLAWHSAGTYDVSSKTGGP
 FGTIKHPSELAHGANNGLDIAVR LLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPF
 HPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDR DIVALSGGHTIGA AAKER

SGFEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDPVFRPLVEKYASDEDADFADYAEAHQKLSSELGFAEA

>Gs-APX-4a

MGKSYPTVSADYQKAVEKAKKKLRGIIAEKRCAPLMLRLAWHSAGTFDKGKTGGP
FGTIKHPAELAHSANGLDIAVRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVFP
HPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLTDQDIVALSGGHTIGAAHKER
SGFEGPWTSNPLIFDNSYFTELLSGEKEGLLQLPSDKALLSDPVFRPLVDKYAADEDADFADYAEAHQKLSSELGFADA

>Gs-APX-5b1

MALPVVVDAEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGGP
NGSIRNEEEYSHGANGLKKAIDFCQEVKAKYPKITYADLFQLAGVVAVEVTGGPTID
FVPGRRDSKVSPNEGRLPDAKKGVPHLRDIFYRMGLTDRDIVALSGGHTLGRAHPERS
GFDGPWTE DPLKFDNSYFVELLKEDSAGLLKLPTDKALLED AEFRCYVELYAKDEDA
FFRDYAESHKKLSELGFVPSKPISIKEGTLAQQAVGVVVATAVVILSYLYEVRKRK

>Gs-APX-5b2

MAKPRVDAEYLKEIEKARRDLRALISNRNCAPLMLRLAWHDAGTYDAKTNTGGPNG
SIRNRQELNHAANKGLETALAFCEEVKAKHPKISYADLYQLAGVVAVEVTGGPTINFV
PGRKDSLESPAEGRLPDAKQGASHLRDIFYRMGLGDKDIVALSGGHTLTKAHKDRSD
FHGQWTKDPLKFDNSYFVELLRGESKDLLKLPTDKALVEDPNFRKYVELYAKDEDA
FSDYATSHKKLSELGFIFKNHRSILAKGVIGIAIVLTAVILGYLRELNKKTN

>Gs-APX-5a1

MAEPVVDDDEYLKEIDKARRELRAFITSNQCAPLMLRLAWNDAATYDARNRAGGPNG
SIRTDKELKHEANEGLLKATQLCEHVKAKLKKVSYADLYQLAGVVAIEVSGGPTIEFL
PGRKDSMESSAEGLLPDVKQGASIIRNIFSRMGISDDKHIVALCGGLTWGETLKDRSDS
KGQWPKDPLKFDNSYKILSKDLSSRLPIEDALLTDQSFRRHVEEYSKDENSFFKEY
AMSHKKLSELGYNLKKQNQSKVPDEEQNPKGPYEKLNQHRGLIGIGIVSVVTVILG
YLLKRKKNLKK

>Gs-APX-3a

MAERVLLLTPSQSHHSPTTMAVLGGAASARIIPSVSLSTSSRSFFSLSSSSKCLKLRSSPR
ISHLFLNQRAEVRVSSGGYGTVSAPKSVASDPDQLKSAREDIKELLNSKFCHPILIRL
GWH DAGTYNKNIEEWPLRGGANGLRFEVELKHGANAGLLNALKLLQPIKDKYSGV
TYADLFQLASATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLRQV
FYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGPAGGGQSWTVQWL
KFDNSYFKDIKEKRDEDLLVLP TDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAHAK
LSNLGAKFDPLEGIVIDDSPNAGGEKFVAAKYSIGKE

>Gs-APX-3b

MAERVSLPLPSQSLSPTTTMALLGAASSARIIPSVSLSSSSRSFFSLSSSSSKLQCLRS
SPRISHLFLNQRRAEVRVSSGGYGTVSAPKSVASDPDQLKSAREDIKELLNSKFCHPILI
RLGWHDAGTYNKNIEEWPLRGGANGLRFEIELKHGANAGLVNALKLLQPIKDKYS
GVTYADLFQLAGATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLR
QVFYHMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGPAGGGQSWTVQ
WLKFDNSYFKDIKEKKDEDLLVLP TDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAH
AKLSNLGAKFDPPPEGIVIDDSPNAGAEKFVAAKYSTGKE

>Gs-APX-2

MRLGVSVSVRCSIGTASGNTQTDYPAKFQRSSFSTANFLPDNASSDLAVSISSSSSRRG
 LIRIATLPCLLPLIGCLPANAMQPGSKEYLLIKEEVRKVLKSKGAAGVLRVLFHDAGTF
 DIDDSTGGMNGSIVYELERPENAGLKKSVKACSVSLFSHSPSITISITIWRIHPDSEVL
 QKAKTQIDAIQPVSWADMIAVAGAEAVEVCGGPPIQVSLGRLDTLVPDPEGRLPEESL
 NASGLKKCFQSKGFSTQELVALSGAHTIGSKGFGSPISFDNSYYKVLLEKPWTSSGGM
 PSMIGLPSDHALVEDDECLRWIKKYADSENLFEDFKNAYVKLVNSGVRRNSL

>Gs-APX-5a2

MVEPIVVDEEYRKEIEMARRELGVFITNNKCAPLMLQFAWDAATYDAKSRRGGPN
 GSIRLRIGQELKHEANKGLENVQYCEIVKTKLKLKVVSYADLYQDSNESPRTEGRFI
 DGEEDARNLRKIFSRMGLSDEQDIVALCGGHTLIRTMYPKVPMPGETHKDRSKFEEGK
 STNKPLKFDNSYFKELLSKDASFSRLPMDYALVEDPKFHHCVERYDKEIFFKEYAISHG
 KLSELGFNLNLDQPKGPYAKLNQHKGLIGIGIASVVVTVILGYLLKRKKNQLKN

>Gs-APX-1a

MGVSFLSTCPSLLPLVAYDTPPSTRYQPRSVSICCNKTISNVSDDEEPFHWRRREILKCAG
 ATVGLELIGSSGSLVGMANAADLIQRRQRSEFQSKVKETLFTAIGKNPDLIPSLTLAL
 NDALTYDKATKSGGPNGSIRFSSEIGRPENTGLSAAMNLEEAKKEIDSYSKGGPISYA
 DLIHFAAQSAAKATFLAAAIRKCGGNEEKGKTLYNAYGSNGQWGLFDRQFGRADTQ
 DPDPEGRVPIWEKASVQEMKDKFVAVGFGPRQLAVLSAFLGPDQONATETLLASDPDV
 APWVQKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYPPPKIDITKLKL

>Gs-APX-1b

MGVSFLSTCPSLLPLVAYDTPPSARYQPRSVSICCNKTISDVSDGEPFHWRRRDVLKCA
 GATVGLELIGSSGSLVEMANAADLIQRRQRSEFQSKIKETLFTAIGKNPDLIPSLTLAL
 NDALTYDKATKSGGPNGSIRFSSEIGRPENTGLSAALNLEEAKKEIDSDSKGGSISYA
 DLIHIAAQSAAKATFLAAAIRKCGGKEEKGKTLYNAYGSNGQWGLFDRQFGRADAQ
 DPDPEGRVPLWEKASVQEMKDKFVAVGFGPRQLAVLSAFLGPDQONATETLLASDPDV
 APWVQKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYPPPKIDITKLKL

>Mt-APX-4

MGKSYPTVSADYQKAVEKAKRKLRFIAEKKCAPLMLRLAWHSAGTFDSKTKTGGP
 FGTIKHQAELAHGANGLDIAVRLLEPLKEQFPIISYADFYQLAGVVAVEVTGGPEVPF
 HPGREDKPEPPPEGRLPDATKGSDDLVDVFGKAMGLSDQDIVALSGGHTIGAAHKER
 SGFEGPWTSNPLIFDNSYFTELLGGEKEGLLQLPSDKALLSDPVFRPLVEKYAADEDA
 FFADYAEAHQKLSLGFAEA

>Mt-APX-5b2

MAKYAVDAEYVKEIEEVRDLRSFITNKKCAPLMLRLAWHDAGTYDAKTRTGGPNG
 SIRNQELNHAANKGLKTAVELCEEVVKVHPKVSADLYQLAGVVAVEVTGGPTIQF
 VPGRKDSLESPQEGRLPDAKQGASHLREIFYRMGLTDKDIVALSGGHTLGKAHKDRS
 NFEQWTRDPLKFDNSYFVELMKLESNDLLKLPTDKALVDDPAFRKYVELYAKDEN
 AFFRDYATSHKKLSELGFNPSCSYRSQLAKAALGMVIASTVVVLGYLLELNKKIN

>Mt-APX-5b1

MAFPVVVDSEYLKEIEKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGGPN
 ASIRNEEEFSGANGLKKAIDFCEEVKAKHPKISYADLYQLAGVVAVEVTGGPTVNF
 VPGRRDSKVCTRDGRLPDAKQGVSHLRDIFYRMGLTDKDIVALSGAHTLGRAHPERS
 GFDGPWTEPLKFDNSYFQILLEEDSAALLKLPTDRALLDDPEFRRYVELYAKDEDAF

FRDYAESHKKLSELGFAPSSKATAPVDATVLAQSAFGVVAAA AVVILGYLYETHKRGK

>Mt-APX-2

MMCSRVSNRIVRCASGGENSYPAKLQPSCLSTVKFRSNKPNDHASSDVSSSSRRRAVIFS
SIATLPCLLPLTHIFGSLQANAMPPGTKEYLLIKEELRKVLTKGKAAGVLRLVFDHAGT
FEIDDNTGGMNGSIVYELERPENTGLKKSVKVLQKAKTQIDAIHPVSWADVIAVAGTE
AVEVCGGPTITVSLGRQDSPGPDPEGKLPETLDASGLKRCFHKKGFSTQELVALSGA
HTLGSKGFGSPTSFDNSYKVLLEKPWTPSGGMSTMIGLPSDHALVEDDECLRWIKK
YAENENMFFEDFKNVYVKLVNSGVKWNLSL

>Mt-APX-1

MGVSFLSISTLPSLHPLGGSHYPKTTSTRHQPRSVSICCAKIKTDVSDEEEQFFHWRRR
DVLKCIGLTIGLESISSSGSLLGTANAADLIERRQRSDFQSQIKGTLYTAIKANPDLTPSIL
TLALNDALTYDKATKTGGPNGSIRFSSEISRPENKGLS AALNFIEEAKKEIDSYSKGGPI
SYADLIQYAAQSATKAQFLASAIRKCGGNEEKGNLLYTAYGSNGQWGLFDRQFGRTD
TDEADPEGRIPLWEKASVQEMKDKFS AVGFGPRQLAVLSAFIGPDQDATETLLASDPD
VAPWVQKYQRSRET VSR TDYEVDLITFTKLSTLGQKINYEAYTYPRKKIDITKLKL

>Pv-APX-4

MGKSYPSVSADYQKAIEKAKRKL RGFIAEKRCAPLMLRLAWHSAGTFDVSTKTGGPF
GTIKHPSELAHGANNGLDI AVR LLEPLKAEFPILSYADFYQLAGVVGVEVTGGPEVPF
YPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDQDIVALSGGHTIGA A HKER
SGFEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDPVFRPLVDKYAADEDA
FFADYAVAHQKLSSELGFAEA

>Pv-APX-5b1

MALPVVDSEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGGP
NGSIRTEEEYSHGSNNGLKK AIDFCEEVKAKHPKITYADLYQLAGVVAVEVTGGPTVD
FVPGRKDSKVSPKEGRLPDAKLG VPHLRDIFHRMGLTDRDIVALSGGHTLGRAHPERS
GFDGPWTE DPLKFDNSYFVELLKGDYAGLLKLPTDKALLEDPEFRRYVELYAKDEDA
FFRDYAESHKKLSELGFVPSSKAISVKDGTVLAQGA VGVVVTA AVVILSYLYEVRKRK
K

>Pv-APX-5b2

MAKAIADAEYVKEIEKARRDLRALISSRNCAPLMLRLAWHDAGTYDVKTRTGGPNG
SIRNARELNHAANKGLETAVVLCEQVKAKHPNVS YADLYQLAGVVAVEITGGPTIDFV
PGRKDALESPAEGRLPDAKQGASHLREMFYRMGLGDKDIVALSGAHTLGKAHRDRS
NFEGQWTKDPLKFDNSYFVELLKGESKDLLKLPTDKALVEDPIFRNYVELYAKDEDV
FFSDYATSHKKLSELGFILKNHRSM LAKGVVGF GIALTAVILGYLHELNKRTN

>Pv-APX-3

MAERAFLTPSSLPLPSQRTPTTMALFSGAASARIIPSVLS SSSSR SFFSL SSSSPSLQCLRS
SPRISHLFLNQRRAEVRVSSGGYGTVSAPKSFASDPDQLKNAREDIKELLRSKFCHPILI
RLGWHDAGTYNKNIEEWPQRGGANGLRFEIELKHAANAGLVNALKLLQPIKEKYS
GVTYADLFQLAGATAVEEAGGPKLPMKYGRVDVSKPEQCPEEGRLPDAGPPSPADHL
RQVFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPE SKYTKDGP GAGPGQSWTVQ
WLKFDNSYFKDIKEKKDEDLVLPTDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAH
AKLSNLGAKFDPPEGIVIDDSPNPKAEKFVAAKYSTGKE

>Pv-APX-2

MRLGVSKHVVRCSIGTVSGDSKTHFPAKSQRSSFSTVKFRSDNGSSDLAVSSSSRRELI
 CIVTWPCLLPLVGCLPANALQAGAKEYLLIKEEVRKVLKSKGKAAGVLRVLFHDAGTF
 DIDDNTGGMNGSLVYELERPENAGLKRSVKVLQKAKTQIDAIQPVSWADMIAVAGAE
 AVEICGGPPIQVSLGRDLALVPDPEGRLPEESLNASALKKCFQSKGLSTQELVALSGAH
 TIGSKGFGSPNSFDNSYYKILLEKPWTSSGGMSSMIGLPSDHALVEDDECLRWIKKYA
 DSENLFFEDFKNAYVKLVNSGARWKNM

>Pv-APX-5a

MGEVDEEYLSEIEMARRQLRSFITNKNKCAPLMLQLAWNDAATGGPNASIRFITKHEAF
 KDLEKAVEYCEIVKAKLKLKKVSYADLYQLAGVVAIEVTQGPTIDFVPGRKDLNESPR
 VQIRFLNGEEDARSLRRKFSRMGLSEDKDIVVLCGGHTLTRSMYPKVPMPRETPEGET
 HKDSSKFDERKWTKDPLKFDNSYFKELLSKGASFSRLPIDNALVEDQIFRHYVERYAK
 DEDIFFKEYAISHKKLSELGFDPKKNLQDKGLCQKLNQHKGLVGIGMISVVLTAILGYL
 HKKKKNLQKN

>Pv-APX-1

MGVSFSTCPSLLSIAANDTPPSTRYQPRSVSICCNKTISDVAAEPPFHWRREVLKCV
 GATLGLEIFGSSGLVGMANAADLIQRRQRSEFQSQVKETLFNAIKGNPDIIPSLTLAI
 NDAVTYDKATKSGGPNGSIRFSSEISRPENAGLSAALNLLLEEAKKEIDSYSKGGPISYA
 DLINYAAQSAVKATFLAAAIKCGGNVEKGSILYNAYGSSGQWGLFDRQFGRADAQD
 PDPEGRVPVWEKASVQEMKDKFSALGFGPRQLAVLSAFLGPDQSATETLLASDPDVA
 PWVQKYQRSRETVSRDYEVDLITFTKLSTLGGQINYEAYTYPPPKIDITKLKL

>Va-APX-4

MGKSYPTVSADYQKA FEKAKKKLRGFIAEKRCAPLMLRLAWHSAGTFD VSTKTGGP
 FGTIKHS AELAHGANGLDIAVRLLEPIKA EFPILSYADFYQLAGVVAVEITGGPEVPFH
 PGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLSDQDIVALS GGHTIGA AHKERS
 GFEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDSVFRPLVEKYAADEDAFF
 ADYAVAHQKLSLGFAEA

>Va-APX-5b1

MALPVVVDSEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKSKTGGP
 NGSIRNEEEF SHGSNGLKKAIDFCEEVKEKHPKITYADLYQLAGVVAVEVTGGPTIDF
 VPGRKDSKVSPKEGRLPDAKQGVSHLRNIFYRMGLTDRDIVLSGGHTLGRAHPERS
 GFDGPWTE DPLKFDNSYFVELLKGEYIGLLKLPTDKALLEDPEFRRYVELYAKDEDAF
 FRDYAEAHKKLSELGFVPSSKAISIKDGTVLAQGA VGVVTA AVVILSYLYEVRKR GK

>Va-APX-5b2

MAKAIADA EYVKEIEKARRDLRTLIS SRNCAPLMLRLAWHDAGTYDVKTRTGGPNG
 SIRNARELNHAANKGLETAVLFCEDVKAKHPKVS YADLYQLAGVVAVEITGGPTIDFV
 PGRKDSLES PAEGRLPDAKQGASHIREIFYRMGLSDR DIVALS GGHTLGKAHRDRSDF
 EGQWTKDPFKFDNSYFVELLKGESKDLLKLPTDKALVEDPMFRKYVELYAKDEDAFF
 KDYATSHKKLSELGFS LKNHRSILAKGVIGFGIALTAVILGYLHELNKRTN

>Va-APX-3

MAERVFLTPSSPPLPSQHTPTAMALFSGAASARIIPSVLS SSSRSFFSIS SSSSSSLQCLRSS
 PRISHLFLNQRRVEVRVSSEGYGTVSAPKSFASDPDQLKGAREDIKELLRSKFCHPILIR
 LGWHDAGTYNKNIEEWPQRGGANGSLRFEIELKHGANAGLVNALKLLQPIKDKYSG
 VTYADLFLQLAGATAVEEAGGPKLPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLR
 QVFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGPGAPGGQSWTVQ

WLKFDNSYFKDIKEKKDEDLLVLPTDAALFEDPSFKVYAEKYAEDQEAFFKDYAEAH
AKLSNLGAKFDPPEGIVIDESPNAGAIEKFVAAKYSTGKVFIFAHIKSYYSFC

>Va-APX-2

MRLGVSKHVVRCSIGTVSGESKTHFPAKSQRSSFSAVKFRPGDGSSDLAVSSSSRRELI
CIVTWPCLLPLIGCLPANALQAGAKEYLLIKEEVKVLKSKGKAAGVLRLVFDAGTF
DINDNSGGMNGSIVYELERPENAGLKKSVKVLQKAKTQIDAIQPVSWADMIAVAGAE
AVEVCGGPPIQVSLGRDLALVPDPEGRLPEESLNASGLKKCFQSKGLSTQELVALSGA
HTIGSKGFGSPNSFDNSYKVLLEKPWTSSGGMSSMIGLPSDHALVEDDECLRWIKKY
ADSENLFEDFKNAYVKLVDSGVRWKNM

>Va-APX-5a

MDAVDEEYLKEIEIARRELRSFITNKNKCAPLIQLAWNDAATYDARSRTGGPNGSIRFIT
KHEAIKELQKAVQYCDIVKHRLKLLKVSADLYQLAGVVAIEVTQGPIIDFVPRKDS
NESPRVQAHFLNGEEDARSLRRKFSRMGLSEDEDIVVLCGGHALIRAMHPKVAMRET
PKGETHEDRSKFEERKWTKEPLKFDNSYFKELLSKGASFSRLPMDALVEDQRFYHY
VERYAKDEDIFFKEYAISHKKLSELGFDPQNLNKSGLYQKLNQHRGLVGIGIISVAVA
GILGYLHKKKKKNQLKD

>Va-APX-1

MGVSFLSTCPSLLSLVANDTPPSTRYHPRSVSICCNKTISDVVGEDPFHWRRREVLKCV
GATVGLEIIGSSGLVAMANAADLIERRQRSEFQSQVKGTLYNAIKGNPDIIPSLLKLAL
NDALTYEKATKSGPNGSIRFSSEISRPENTGLSPALNLEEAKKEIDSYSKGGPISYAD
LINYAAQSAVKATFLAAAIRKCGGNEEKGNILYNAYGSNGQWGLFDRQFGRDLAQDP
DPEGRVPIWEKASVQEMKDKFSAVGFGRQLAVLSAFLGPEQSATETLLASDPDVAPW
VQKYQRSRETVSQTDYEVDLITFTKLSTLGQQINYEAYTYPRAKIDITKLL

>Vr-APX-4

MGKSYPTVSADYQKAIEKAKRKLGRFIAEKRCAPLILRLAWHSAGTFDVKTKTGGPF
GTIKHPAELAHGANGLDIAVRLLEPIKAIEFPILSYADFYQLAGVVAVEITGGPEVPHF
GREDKPEPPPEGRLPDATKGSDDLDRDVFVKAMGLSDQDIVALSGGHTIGAAHKERSG
FEGPWTSNPLIFDNSYFKELLSGEKEGLLQLPSDKALLSDPVFRPLVEKYAADEDAFFA
DYAVAHQKLSLSELGFAEA

>Vr-APX-5b1

MALPVVDSEYLKEVDKARRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGGP
NGSIRNEEEYSHGSNGLKKAIDFCEEVKAKHPKITYADLYQLAGVVAVEITGGPTID
FVPRKDSKVSPEGRLPDAKQGVSHLRDIFYRMGLTDRDIVALSGGHTLGRAHPER
SGFDGPWTEPLKFDNSYFLELLKGEYIGLLKLPTDKALLEDPEFRRYVELYAKDEDA
FFKDYAEAHKKLSELGFVSSKAISIKDGTVLAQGAVGVVTAAVVILSYLYEVRKRG
K

>Vr-APX-5b2

MAKAIADAIEYVKEIEKARRDLRALISSRNCAPLMLRLAWHDAGTYDVKTRTGGPNG
SIRNARELNHAANKGLETAVLFCEDVKAKHPKVSADLYQLAGVVAVEITGGPTIDFV
PGRKDSLESPAEGRLPDAKQGASHLREIFYRMGLSDRDIVALSGGHTLKGAKHRDRSDF
EGQWTKDPFKFDNSYFVELLKGESKDLLKLPTDKALVEDPMFRKYVELYAKDEDAFF
KDYATSHKKLSELGFILKNHRSILATGVIGFGIALTAVILGYLHELKRTN

>Vr-APX-3

MAERVFLTASSRPLPSQQHTSTAMALFSGAASARIIPSVLSSTSTRSFFSLSSSSSLQCL
 RSSPRISHLFLNQRRAEVRFSSGGYGTVSAPKSFASDPDQLKSAREDIKELLRSKFCHPI
 LIRLGWHDAGTYNKNIEEWPQRGGANGLRFEIELKHGANAGLVNALKLLQPIKDKY
 SGVTYADLFLQLAGATAVEEAGGPKLPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADH
 LRQVFYRMGLNDKEIVVLSGAHTLGRSRPDRSGWGKPKETKYTKDGGAPGGQSWTV
 QWLKFDNSYFKDIKEKKDEDLLVLPDAAALFEDPSFKVYAKEYAEDQEAFFKDYAEA
 HAKLSNLGAKFDPPEGIVIDESPNAGAIEKFAAKYSTGKE

>Vr-APX-2

MRLGVNKHVVRCSIGTVSGESKTHFPAKSQRSSFSVAVKFRPGDGSSDLAVSSSSRRELI
 CIVTWPCLLPLIGCLPANALQAGAKEYLLIKEEVRKVLSKGKAAGVLRVLFHDAGTF
 DINDNSGGMNGSIVYELERPENAGLTKSVKVLQKAKTQIDAIQPVSWADMIAVAGAE
 AVEVCGPPIQVSLGRLDALVPDPEGRLPEESLNASGLKMCQSKGLSTQELVALSGA
 HTIGSKGFGSPNSFDNSYYKVLLEKPWTSSGGMSSMIGLPSDHALVEDDECLRWIKKY
 ADSENLFEDFKNAYVKLVNSGVRWKNM

>Vr-APX-5a

MGAVDDEDYLKEIEIARRELRFSFIANNKCAPLMLQLAWNDAATYDARSRTGGPNASIRF
 ITKHAAIKELEKAVQNCDIVKHRLKLRKVSADLYQLAGVVAIEVTQGPIIEFVPRKRD
 SNESPRVQANFLNGEEDARSLRRKFSRMGLSEDKDIVVLCGGHALIRTMHPKVAMRE
 TPKGETHEDRSRFEERKWTKEPLKFDNSYFKELLSKGASFSRLPMDALSALVEDQRFYH
 YVERYAKDEDIFFKEYAISHKKLSELGFDPKNLNKSGLYQKLNQHRGLVGIGIISVAV
 AGILVYLHKMKKNQLKD

>Vr-APX-1

MGVSFSLSTCPSLLSLVANDTPPSTRYHPRSVPICCNKTIIDVVGEEPFHWRRREVLCV
 GATVGLIIGSSGSLVGMANAADLIERRQRSEFQSQVKGTLYNAIKGNPDIIPSLLKLAL
 NDALTYEKATKSGGPNGSIRFSSEISRPENTGLSPALNLEEAKKEIDSYSKGGPISYAD
 LINYAAQSAVKATFLAAAIKCGGNEEKGNILYNAYGNSGQWGLFDRQFGRVDAQDP
 DPEGRVPIWEKASVQEMKDKFSAVGFGPRQLAVLSAFLGPDQSATETLLASDPDVAP
 WVQKYQRSRETVSRTDYEVDLITFTKLSTLGGQINYEAYTYPRAKIDITKLKL

>At-APX-4a

MTKNYPTVSEDYKKAVEKCRRLRGLIAEKNCAPIMVRLAWHSAGTFDCQSRTGGPF
 GTMRFDAEQAHGANSIHALRLLDPIREQFPTISFADFHQLAGVVAVEVTGGPDIPFH
 PGREDKPQPPEGRLPDATKGCDDLDRDVFQKQMGSLDKDIVALSGAHTLGRCHKDRS
 GFEGAWTSNPLIFDNSYFKELLSGEKEGLLQLVSDKALLDDPVFRPLVEKYAADEDAF
 FADYAEAHMKLSELGFADA

>At-APX-4b

MVKKSYPEVKEEYKAVQRCKRKLRLGLIAEKHCAPIVLRLAWHSAGTFDVKTGTGG
 PFGTIRHPQELAHDANGLDIAVRLLDPIKELFPILSYADFYQLAGVVAVEITGGPEIPF
 HPGRLDKVEPPPEGRLPQATKGVLDHLRDVFRMGLNDKDIVALSGGHTLGRCHKERS
 GFEGAWTPNPLIFDNSYFKELLSGEKEGLLQLPTDKALLDDPLFLPFVEKYAADEDAFF
 EDYTEAHLKLSELGFADKE

>At-APX-5b

MAPIVDAEYLKEITKARRELRSLIANKNCAPIMLRLAWHDAGTYDAQSKTGGPNGS
 IRNEEEHTHGANSGLKIALDLCEGVKAKHPKITYADLYQLAGVVAVEVTGGPDIVFVP
 GRKDSNVCPKEGRLPDAKQGFQHLRDVFRMGLSDKDIVALSGGHTLGRAHPERSG

FDGPWTQEPLKFDNSYFVELLKGESEGLLKLPTDKTLLEDPEFRRLVELYAKDEDAFF
RDYAESHKKLSELGFNPNSSAGKAVADSTILAQSAFGVAVAAAVVAFGYFYEIRKRMK

>At-APX-3b

MAERVSLLTNGTLLSPPPTTTTTTMSSSLRSTTAASLLLRSSSSSSRSTLTLSASSLSL呢
RSLVSSPRLSSSSLSQKKYRIASVNRSFNSTTAATKSSSSDPDQLKNAREDIKELLSTK
FCHPILVRLGWH DAGTYNKNIKEWPQRGGANGSLRFDIELKHAANAGLVNALNLIKD
IKEKYSGISYADLFQLASATAIEEAGGPKIPMKYGRVDASGPEDCPEEGRLPDAGPPSPA
THLREVFYRMGLDDKDIVALS GAHTLGRSRPERSGWGKPKETKYTKEGPGAPGGQSW
TPEWLKFDNSYFKEIKEKRDEDLLVLPTDAAIFEDSSFKVYAEKYAADQDAFFKDYAV
AHAKLSNLGAEFNPPEGIVI

>At-APX-5a

MAVNVD AEYLKEIEKTRRDLRALISSRNCAPIMLRLAWHDAGTYDAKKKTGGANGSI
RFKEELNRPHNKGLEKAVAFCEEVKAKHPRVSYADLYQLAGVVAVEVTGGPAIPFTPG
RKDADSADDGELPNPNEGASHLRTLFSRMGLLDRDIVALSGGHTLGRAHKERSDFEG
PWTQDPLKFDNSYFVELLKGETPGLLQLKTDKALLDDPKFHPFVKLYAKDEDMFFKA
YAISHKKLSELGFNPPRRIPSAVTQQTLGIAVAAAVVIFTICYEASRRGK

>At-APX-2

MTTTTASLVKTLFRCDSSFKFCKFESPAKTRLLSPATEKHVVRSSRAWRIRCLSD
DPGSSHV FVASRRKMVLLSTVQLLHMLPQNGNAAEIYPVMQNEIRKVVTKGKAA
GVLRLVFHDAGTFELDDHSGGINGSIAYELERPENIGLKKSLKVLAKAKVKVDEIQPV
SWADMISVAGSEAVSICGGPTIPVVLGRLDSAQPDPEGKLPETLSASGLKECFKRKGF
STQELVALSGAHTIGSKGFGDPTVFDNAYYKILLEKPWTSTSKMTSMVGLPSDHALVQ
DDECLRWVKRYAEDQDKFFEDFTNAYIKLVNSGAKWNML

>At-APX-3a

MSVLSAASHLLCSSTRVSLSPAVTSSSSSPVVALSSSTSPHSLGVSASSLFPHSSFVLQ
KKHPINGTSTRMISPKCAASDAAQLISAKEDIKVLRLTKFCHPILVRLGWH DAGTYNK
NIEEWPLRGGANGSLRFEAELKHAANAGLLNALKLIQPLKDKYPNISYADLFQLASAT
AIEEAGGPDIPMKYGRVDVVAPEQCPEEGRLPDAGPPSPADHLRDVFYRMGLDDKEIV
ALSGAHTLGRARPDRSGWGKPKETKYTKTGPEAGGQSWTVKWLKFDNSYFKDIKE
KRDDDLLVLPTDAAALFEDPSFKNYAEKYAEDVAFFKDYAEAHAKLSNLGAKFDPPE
GIVIENVPEKFVAAKYSTGKKELSDSMKKKIRAEYEAIGGSPDKPLPTNYFLNIIIAIGV
LVLLSTLFGGNNNSDFSGF

>At-APX-1

MGGVSFLSTVPSFTNTTNHQHLTTLSSSSHRSAVIRCSKIEPQVSGESLAFHRRDVLKL
AGTAVGMELIGNGFINNVD AKAADLNQRRQRSEFQSKIKILLSTTIKAKPELVPSLLK
LALNDAMTYDKATKSGGANGSIRFSSELSRAENGLSDGLSLIEEVKKEIDSISKGGPIS
YADIIQLAGQSAVKFTYLA SAIRKCGGNEEKGNLLYTAYGSAGQWGLFDRNFRSDA
TEADPEGRVPQWGKATVQEMKDKFIAVGLGPRQLAVMSAFLGPDQAATEQLLATDPQ
VAPWVQKYQRSRETVSQTDYEVDLITAF TKLSCLGQQINFEAYTYPVERINLSKLLK

>Vc-APX-4

MATSAPRMLRLGSLLARQAAPSSLLQQQFRTFSDNAAA EAKAAPEAVKADSGSKG
KIRGVQALIFGGLGAGMLYGYSLFTSQPPSGPLAGPPTLPTGTAPKTSTPGDYAAVRD
SISSLLEAADYDDGSYGPLLRLAWHASGTYDKSSCTGGSNGATMRFPPECEWAANR
GLAIARQLLEPVKAAHPWISYADLWTLAGVVAIEDMGGPSVAWRPGREDYSDGSKIV

PDGRLPNATLGAKHLRDIFHRMGFDDRDIVALS GAHTLGRCHPDRSGFSGPW TNAPT
TFSNLYFQELVNNKWRPKKWDGPLQYEDAKTGTLMMLPTDLALLSDRTFKKYVAQY
AKDEEAFFKDFAVAFGKLELGVPPFSAEKAAATS

>Vc-APX-3

MQTCFPKGTTKRAVFGSCRRRTPTVCVRAAVNVEQLKALRQELFSYINSRGCNPIIVR
LGWHDSGTYDKNIAEWPARGGATGSIRFKPEIDHAANKGLAIALGILGPMKKKFPEVS
YADLFQMASAVAVEAAGGPKIPMRYGRKDATSPEQCVPDGRLPAAHHPFADGSTSPA
EHLRRVFRMGLTDQEIVVLSGGHTLGRARPERSGFGADKTKYTDVGPSTSSASPSG
SPDRPVTPKPVGQLGTSWTANWLEFDNSYFTEVKAKRDADLLVLPDACLFEEDGFR
PYAEKYAADQEAFFADYALAAQQKLELGVVEWDEGAPVTI

>Vc-APX-2

MFGVLHRPPCPDTLASRTPGAKQRNPLLIRADVAHQARVPSTGHDRDQTKEQPALS
RQVLGLSAPLLGALIGGVLPISIGQLVATAPATAAAISVPVTPSREVKAAALNKALARAIP
KAKTAVALLRVFHDAGTFSFPPGNGGLNASIQYELDRPENAGLKRGWRIIEQVRDEL
GTPAEGVATDADLVALAGAYAVRLCGGPDIALALGRPVAASPDPPDRMPGENFTA
QLKANFAAKGLSVQEMVALS GAHTLGSKGFDPTRFDNEYRALLRRPWTPNDMS
ASMIGLPSDHVLPDDPECLPYIERAEDQDAFFADFAAAYVKLTS LGVPGWAV

>At-MDHAR-1

MALASTTLPTKSGLSLWCPSSPLARRFPARFSPIGSRIASRSLVTASFANENREFVIVG
GGNAAGYAARTFVENGMADGRLCIVTKEAYAPYERPALTKAYLFPPEKKPARLPGFH
TCVGGGGERQTPDWYKEKGEVIYEDPVAGADFEKQTLTTDAGKQLKYGSLIATGCT
ASRFPDKIGGHLPGVHYIREVADADSLIASLGKAKKIVIVGGGYIGMEVAAA VAWN
DTTIVFPEDQLLQRLFTP SLAQKYEELYRQNGVKFVKGASINNLEAGSDGRVSAV
KLAGSTIEADTVVIGIGAKPAIGPFETLAMNKSIGGIQVDGLFRTSTPGIFAIGDVA
AFPLKIYDRMTRVEHVDHARRSAQHCVKSLTAHTDTYDYLPHYFYSRVFEYEGSP
RKVWWQF FGDNVGETVEVGNFDPKIATFWIESGRLKGVLVESGSPEEFQLLPKLARS
QPLVDKAK LASASSVEEALeiaQAALQS

>At-MDHAR-2

MGRAFVYVILGGGVAAGYAALEFTRRGVSDGELCIISEEPVAPYERPALS KGFLLPEAP
ARLPSFHTCVGANDEKLTPKWKDHGIELVLGTRVKSVDVRRKTLLSSTGETISYKFL
IIATGARALKLEEFVGEVSDAENVCYLRDLADANRLATVIQSSNGNAV VIGGGYIGM
ECAASLVINKINVTMVFPEAHC MARLFTP KIASLYEDYRAKGVKFIKGTVLTSFEFDS
NKKVTAVNLKDGSHLPADLVVVGIGIRPNTSLFEGQLTIEKGGIKVNSRMQSSDSSVYA
IGDVATFPVKLFGEMRRLEHVD SARKSARHAVSAIMDPIKTGDFDYL PFFYSRVFAFS
WQFYGDPTGDVVHFGEYEDGKSFGAYWVKKGHLVGSFLEGGTKEEYETISKATQLK
PAVTIDLEELEREGLGFAHTVVSQQKVPEVKDIPSAEMVKQSASVVMIKKPLYVWHA
ATGVVVAASVAFAFWYGRRRRRW

>At-MDHAR-3a

MAEKSFKYIILGGGVSAGYAAKEFANQGVQPGELAVISKEAVAPYERPALS KGYLFPE
GAARLPGFHCCVGS GGKLLPESYKQKGIELILSTEIVKADLSAKSLVSATGDVFKYQ
TLIATGSTVLRRLTDFGVKGADSKNILYLREIDDADKLVEAIKAKKGGKAVVGGGYI
GLELSAVLRINNLDVTMVFPEPWCMRPLFTADIAAFYETYTYTNKGVKIIKGTVASGFT
AQPNGEVKEVQLKDGR TLEADIVIVGVGAKPLTSLFKGQVEEDKGGIKTDAFFKTSVP
DVYAVGDVATFPLKMYGDVRRVEHVDHSRKS AEQAVKAIKAAEGGA AVEEYDYLPH
FYSRSFDLSWQFYGDNVGDSV LFGDSNPSNPKPRFGAYWVQGGKVVGAFMEGGSG

DENKALAKVAKARPSAESLDELVKQGISFAAKI

>At-MDHAR-3b1

MAEEKSYKYVVIIGGGVAGGYAAREFSNQGLKPGELAIISKEPVPFFERPELTKVYIDLE
 VNPTLANIYVCAGTGEAKQYPNWKYKEKGIDLIVGTEIVKADLASKTLVSDDGKIYKY
 QTLIATGSTNIRLSEIGVQEADVKNIFYLREIEDSDELALAMELYVQRGKAVIIGGGFL
 GLEISSALRANNHEVTMVFPEPWLVHRFFTAETASFYESYANKGIKIKGTVATGFSTN
 SDGEVTEVKLEDGRTLEANIVVAGVGARPATSLFKGQLEEEKGGIKTDGFFKTSVPDV
 YALGDVATFPMKMYGGTRRVEHADNARKSAAQAVKAIKAGEEGKTIPDYDYLPHYFY
 SRFFKLSWEFYGENVGESVLFGDNDPKSPKPKFGTYWVKDGVVGVFLEGGTQEEH
 KAIKVARAQPSVESLDVLSEEGLSFATKIFYSTSL

>At-MDHAR-3b2

MAEEKSFKYVIVGGGVAAGYAAREFFNQGVKPGELAIISREQVPPYERPALSKGYIHL
 ENKATLPNFYVAAGIGGERQFPQWYKEKGIELILGTEIVKADLAAKTLVSGTGQVFKY
 QTLAATGSSVIRLSDFGVPGADAKNIFYLRELEADADYLAYAMETKEKGKAVVGGG
 YIGLELGAALKANNLDVTMVYPEPWCMPLRFTAGIASFYEGYYANKGINIVKGTVAS
 GFTTNSNGEVTEVKLKDGRVLEADIVVGVGGRPISLFDQVEEEKGGLKTDGFFKTS
 LPDVYAIGDVATFPMKLYNEMRRVEHVDHARKSAEQAVKAIKAAEEGNSIPEYDYL
 YFYSRAFDLSWQFYGDNVGESVLFGDNDPESPKPKFGSYWIKERKVVGAFLEGGSP
 ENNAIAKLARAQPSVESLEVLSKEGLSFATNI

>Ad-MDHAR-3a

MAKTFKYVILGGGVAAGYAAREFAKQGLKPGELAIISKEAVAPYERPALSKAYLFPESP
 ARLPGFHVVCVGGGERLLPEWYNEKGIELILSTEIVKADLAAKSLTSAGGETFSYQILL
 IATGSTVIRLSDFGVQGADAKNIFYLREVDADKLYEAIKAKKNGKAVVGGGYIGL
 ELSAVLRLNLDVTMVYPEPWCMPLRFTSGIAAFYEGYYAKKGVNIKGTVAVGFTA
 NSDGEVKEVKLKDGRVLEADIVVGVGGRPQTALFKGQVEEEKGGIKTDAFFKTSVP
 NVYAVGDVATFPLKLYNELRRVEHVDHSRKSAEQAAKAIKAAEEGKTVEEYDYLPHY
 YSRFELSWQFYGDNVGDTVLFGDSSPESSPKPKFGSYWIKDGVVGAFLSGTPEEN
 QAIKVARVQPPVQDQDQLASEGLSFASKI

>Ad-MDHAR-2

MGRAFVYVILGGGVAAGYAALFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
 PARLPSFHTCVGANEERLTPKWYKENGIELVLGTGVKSADVCRKTLTTTGETISYKIL
 IVATGARALKLEEFVSGSDAQNVCYLRNIADANRLVNAVQSSPGGNVAVVIGGGYIG
 MECAASLVINKINVTMVFPEAHCMARLFTPKIASYYEYYKSRGVNFVKGTVLSSFD
 FDSGKVTAVNLRDGRKISVDMVVVGGIRPNTSLFEGQLTLEKGGIKVNGKLQSSNS
 SVYAIGDVAAPVKAFFGETRRLEHVDSARQSARYVVAIMEPGKTGEFDYLPFFYSRV
 FTLSWQFYGDNVGEVVYFGDFSGSTFGAYWVNKGHLVGAFLGGTKEEYEAVAKAT
 RFRPAIEDLTELERQGLTFAVTVSQKQLSSPPQIQVTSSTDVLVLEKPLYAWHATAGVILA
 ASVAAFAYYYGKKRRRW

>Ad-MDHAR-3b

MVKTFFKYIILGGGVAAGYAAREFAKQGLKPGELAIISKEAVAPYERTVLSKSYLHPEFP
 SRLPTFYVYAGPRREILFPEWYSEKGIELILSTEIVKADLAAKSLTSAGGETFSYQILLIA
 TGSTVIRLSDLGVQGADAKNIFYFREIDDADKLYEAIKAKKNGKAVVGGGEHIGMEL
 SAMLRCNNLDVTMVYPEPWCMRSLFTSDIAAFNEGYAKKGVNIMKRTVVVGFTAN
 SDGEVKEVKLKDGRVLEADIVVVDGIRPQTALFKGQVEEEKGGIKTDGFFKTSVFN
 VYAVGDVATFPVKLYNELRRVEYAEHACQSARKAVEAIKAAEEGKTVEEYDYLPLVLY

SHQFGHAWEFYGDNVGDTVLFQDSSPESSIPKFGSYWIKDGKVVGAFWQGRIEEEERD
AMAYVARVRPSVQNVDELAREGFSFARKIENFGL

>Ai-MDHAR-3

MAKTFKYVILGGGVAAGYAAREFAKQGLKPGELAIISKEAVAPYERPALSKAYLFPESP
ARLPGFHVCVGGGERLLPEWYNEKGIELILSTEIVKADLAAKSLTSAGGETFSYQILL
IATGSTVIRLSDFGVQGADAKNIFYLREVDADKLYEAIKAKKNGKAVVVGGGYIGL
ELSAVLKLNLDVTMVYPEPWCMPLRFTSGIAAFYEGYYAKKGVNIVKGTAVGFTA
NSDGEVKEVKLKDGRVLEADIVVVGVGGRPQTTLFKGQVEEEKGGIKTDAFFKTSVP
NVYAVGDVATFPLKLYNELRRVEHVDHSRKSAEQAVKAIKAAEEGKTIEEYDYLPHYFY
SRSELSWQFYGDNVGDTVLFQDSSPESSKPKFGSYWIKDGKVVGAFLSFGTPEENQ
AIAKVARVQPPVQDQDLASEGLSFASKI

>Ai-MDHAR-2

MGRAFVYVILGGGVGAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
PARLPSFHTCVGANEERLTPKWKYKENGIELVLGTGVKSADVVRKTLTTLTGETISYKIL
IVATGARALKLEEFVSGSDAQNVCYLRNIADANRLVNAVQSSPGGNVAVVIGGGYIG
MECAASLVINKINVTMVFPEAHC MARLFTPKIASYEEYYKSRGVNFVKGTVLSSFD
FSDGKVTAVNLRDGRKISVDMVVVIGIRPNTGLFEGQLTLEKGGIKVNGKQLQSSNS
SVYAIGDVAAPVKAFFGETRRLEHVD SARQSARYVVAIMEPGKTEEFDYLPFFYSRV
FTLSWQFYGDNIGEVVYFGDFSGSTFGAYWVNGHVLVGAFLGEGTKEEYEA VAKATR
FRPAIEDLTELERQGLTFAVTVSQKQLSPPQIQVTSSTDLVLEKPLYAWHATAGVILAA
SVAAFAYYYGKKRRRW

>Ai-MDHAR-1

MPSSAVRRLAMAAASSLSLKQGLLLRPLRTNPLTRPRPSSTSLPSRSFRRCYAAFANE
NREYVIVGGGNAAGYAARTFVEHGMADGRLCIVTKEAYAPYERPALTKAYLFPDDKK
PARLPGFHTCVGGGERQTPEWYQEKGIEMLYQDPVKDIDIEKQTLITNSGKHLKYGS
LIATGSSASRFPEKIGGNLPGVHYIREVADADALISSLGKAKKVVVVGGGYIGMEIAA
ATVAWKLDTTIIFPEDHLLQRLFTPSLARRYEELYQKNGVKILKGASIKNLEAGSDGHV
AGVKLGDGSVVEADTVIVGIGAKPAVSPFERVGLNTEVGGQLVKPMNRLFTTKLLIS
CSPDVFLKTGQIYNRTARVEHVDHARRSAQHCVKSLLAQTHTYDYLPHYFYSRVFEY
EGSPRRVWWQFFGDNVGETVEIGNFDPKVATFWIESGKLGVLLES GSPEEFQLPKL
ARNQPVIDKAKLQSASSVEEALDIAWRSLQVEATV

>Cc-MDHAR-3

MAKTFKYIILGGGVSAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESP
ARLPGFHVCVGGGERLLPEWYQEKGIELILSTEIVKADLAAKSLISAAGETFSYQILII
ATGSTVIRLTDVFGVEGADAKKIFYLREVDADKLYEAIKAKKNGKAVVVGGGYIGLE
LSAVLRLNNIDVTMVYPEPWCMPLRFTSGIAAFYEGYYAKKGVNIKGTAVGFTANS
DGEVKEVKLKDGRVLEADIVVVGVGGRPQTALFKGQIEEEKGGIKTDSFFKTNLSDV
YAVGDVATFPLKLYGELRRVEHVDHSRKSAEQAVKAIKAAEDGKTVEEYDYLPHYFYS
RSFDLSWQFYGDNVGDTVLFQDNNPESSKPKFGTYWVKDGKVVGAFLSFGTPEENQ
AIAKVARVQPPAADIDQLAKEGLSFASKI

>Cc-MDHAR-2

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
AARLPSFHTCVGANEERLTPKWKYKEHGIELVLGTGVKSADVVRKTLTTLTATGETISYKF
LIVATGARALKLEEFVSGSDAENVCYLRDIADANRLVDVMQSSPGGNVAVVIGGGYV
GMECAASLVINKIHVTMVFPEEHCMARLFTSKIANFYEEYYKSRGVNFIKGTVLSSFD

FDSNGKVTAVNLRDGSTLSVDMVVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSN
 SSVYAIGDVAAFPVKAFGETRRLEHVDSARKSARHAVSAIMEPDKTGEFDYLPFFYSR
 VFTLSWQFYGDNVGEVVYYGDMSGSAFGAYWVSKGHLVGAFLEGGTKEEYEAIAK
 ATKLRPAIEDLTELERQGLKFAVTVSQKPVAPPVEVRASDLLLEKPLYAWHATAGVIIA
 ASIAAFAYFYGKKRRRW

>Cc-MDHAR-1

MPSPSVRRLTMTAISNSLSLKQGFSLARPRTPLTRTPSSASVKFTRCYAAAFANENREY
 VIVGGGNSAGYAARTFVEHGMADGRLCIVTKEPYAPYERPALTKAYLFPPDKKPARLP
 GFHTCVGSGGERQTPEWYKENGIEMLYEDPVKDIDIEKQTLTSSGKLLKYGSLIATG
 CTASRFPEKIGGNLPGVHYIREVADADALILSLEKAKKVVVVGGGYIGMEVAAAAGV
 WKLDTTIIFPEDHLLQRLFTPSSLARRYEELYQKNGVKILKGASIKNLEAGSNHVGAV
 KLDGDSIVEADTVIIGIGAKPAVSPFERVGLNTDVGGIQVDGLFRTSTPGIFAVGDVAAF
 PLKIYDRVARVEHVDHARRSAQHCVKALLSAQTHTYDYLPHFYSRVFEYEGSPRKVW
 WQFFGDNVGETVEIGNFDPKIANFWIESGKLGVLLES GSPEEFQLLPKLARSQPLIDK
 AKLQNATSVEEALEIARGSLQGEAAV

>Ca-MDHAR-3

MAQSFKYIIVGGGVSAGYAAREFVNQGVKPGELAIISKEAVAPYERPALSKAYLPESP
 ARLPGFHTCVGSGGERLLPEWYNEKGIQLLNTEIVKADLAGKLLTSKGENFNQYQT
 LVIATGSTVIRLTDVFGVEGADAKNIFYLREVDADKLYESIKAKKNGKAVVVGGGYIG
 LEISAVLKLNDVDVTMVYPEPWCMPLFTSEIAAFYEGYANKGVNIVKGTVAVGFT
 TNSDGEVKEVKLKDGRVLEADIVVVGVGGRPQISLFGQVEEEKGGIKTDPFFKTNV
 PDVYAVGDVATFPLKLYNEVRRVEHVDHARKSAEQAVKAIIAAETGKTIEEYDYLPHYF
 YSRFDLSWQFYGDNVGDTVLFGDNNPASSKPKFGTYWIKDGKLGAFLEGGTPDEN
 KAIKVARVLPAVEDVSQLAKEGLSFASKI

>Ca-MDHAR-1

MPSSVRRLTMATITNSLSLKHTLSFGFPQSKTLTRNRIPSSNSLAFRCYASANENREY
 VIVGGGNAAGYAARTFVEHGVADGRLCIVTKEAFAPYERPALTKAYLFPPDKKPARLP
 GFHTCVGSGGERQTPEWYKEKGIEMLYEDPVENVVDIEKQTLTSSGKLLKYGSLIAT
 GCTASRFPAKIGGDLGSHYIRDVADADALISSLGKAKKVVVVGGGYIGMEVAAAAGV
 AWNLDTTIIFPEDHLLQRLFTPSSLARRYEELYQKNGVKILKGASIKNLEASSNGNHVA
 AVKLDGDSIVEADTVIIGIGSKPAVGPFERVGLNTDVGGIKVDGQFRTSIPGVFAVGDA
 AFPLKIYNRIARVEHVDHARRSAQHCVKALLSAQNTYDYLPHYFYSRVFEYEGSLRK
 VWWQFFGDNVGETVEIGNFDPKIATFWIESGKLGVLLES GSPEEFQLLP ELARSQPPI
 DKAKLQNASSVEEALDIARGSLKAAV

>Ca-MDHAR-2

MGRAFVYVILGGGVSAGYAALEFVKRGVNLGELCIISDEPVAPYERPALSKGFLLEPA
 AARLPSFHTCVGANEEERLTPKWYKEHGVELVLGTEVKSADV KRKTLLTATGETISYKF
 LIVATGARALKLEEFVSGSDAENVCYLRNIADANRLVNAIQSSPGGNVAVVIGGGYVG
 MECAASLVINKINVTMVFPEAHC MARLFTP KIASY YEEYKSRGVNFVKGTVLSSFD
 FSDGKVTAVDLRDGSKLSVDMVVVVGIGIRPNTSLFEGQLTLEKGGIKVNGMLQSSNS
 SVYAIGDVAAFPVKAFATRLEHVDSARKSAKHAVSAIMEPDKTGEFDYLPFFYSRV
 FTLSWQFYGDNIGEVVYYGDLSDSSTFGAYWVSKGHLVGAFLEGGTKEEYEAIAKAT
 RVRPAIEDLTELERQGLGFAVTVSEKPVTLPPPIEVNTSGLLLEKPLYAWHAAAGVILA
 ASVATFAYFYGKRRRRW

>Gm-MDHAR-3b

MAKTFKYIILGGGVSAGYAAREFAKQGVPKPGELAIISKEAVAPYERPALSKAYLFPESP
 ARLPGFHVCVGSggerLLPEWYTEKGIELILSTEIVKVDLAAKSLISAGGETFSYQILIV
 ATGSTVIRLTDFGVEGADAKNIFYLREVDADKLYEAIKAKKNGKAVVVGgGYIGLE
 LSAVLKLNIDVTMVYPEPWCMPLRFTAGIAEFYEEYYKNKGVNIIKGTAVGFTSNS
 DGEVKEVKLKDGRVLEADIVVVGVGGRPQTALVKGQVEEEKGGIKTDAFFKTNLSD
 VYAVGDVATFPLKLYGELRRVEHVDHSRKSAEQAVKAIKAAEEGRTVEEYDYLPHYFY
 SRSFDLSWQFYGDNVGDTVLFGDNNPASSKPKFGTYWIKDGKVVGVFLEGGTPEEN
 QAIKVAKVQPPVADVDQLAKEGLSFASKI

>Gm-MDHAR-3a

MAKTFKYIILGGGVSAGYAAREFAKQGVPKPGELAIISKEAVAPYERPALSKAYLFPESP
 ARLPGFHVCVGSggerLLPEWYTEKGIELILSTEIVKVDLAAKSLISAGGETFSYQILIV
 ATGSTVIRLTDFGVEGADAKNIFYLREVDADKLYAAIKAKKNGKAVVVGgGYIGLE
 LSAVLKLNIDVTMVYPEPWCMPLRFTAGIAEFYEGYYANKGVNIIKGTAVGFTSNS
 DGEVKEVKLKDGRVLEADIVVVGVGGRPQTVLVKGQVEEEKGGIKTDAFFKTNLSD
 VYAVGDVATFPLKLYGELRRVEHVDHSRKSAEQAVKAIKAAEEGKTVEEYDYLPHYFY
 SRSFDLSWQFYGDNVGDTVLFGDNNPASSKPKFGTYWIKDGKVVGVFLEGGTPEEN
 QAIKVAKVQPPVADVNQLAKEGLSFASKI

>Gm-MDHAR-2a

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
 AARLPSFHTCVGANEERLTPKWKYKEHGIELVLGTGVKSTDVVRKTLTLLTTGETISYKI
 LIVATGARALKLEEFVSGSDAENVCYLRHIADANRLVDVMQSCPGGNAVVIgGGYI
 GMECAASLVINKINVTMVFPeeHcMARLFTSKIANYYEEYYKSRGVNFIKGTVLSSFD
 FDSNGKVTAVNLRDGTTLSDMVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSN
 SSVYAIGDVAAFPVKAFGETRRLEHVDSARKSAIHVVSSIMEPNKTGEFDYIPFFYSRV
 FTLSWQFYGDNVGEVVYYGDMSGSAFGAYWVSKGHLVGAFLEGGTKEEYEAIAKA
 TRLRPAIEDLTELERQGLSFAVTVSQKPVVPPPVEVRASDLLLEKPLYAWHATAGVIAA
 SIAAFAYFYGKKRRRW

>Gm-MDHAR-2b

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
 AARLPSFHTCVGANEERLTSKWKYKEHGIELVLGTGVKSADVVRKTLTLLTTGETISYKI
 LIVATGARALKLEEFVGTGSDAGNVCYLRDIADANRLVDVMQSCPGGNAVVIgGGYI
 GMECAASLVINKINVTMVFPeeHcMGRlFTSKIANYYEEYYKSRGVNFIKGTVLSSFD
 FDSNGKVTAVNLRDGSTLSVDMVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSN
 SSVYAIGDVAAFPVKAVGETRRLEHVDSARKSARHVVSaimePNKTREFDYLPFFYSR
 IFTLSWQFYGDNVGEVVYYGDMSGSTFGAYWVSKDHLVGAFLEGGTKEEYEAIAKA
 TRLRPVIEDLTELERQGLRFAVTVSEKPMASPIEVRASDLLLEKPLYAWHATAGVIAA
 SIAGFAYFYGKKRRHW

>Gm-MDHAR-1

MPSPSVRRLTMAAISNSLSLKHsFSLSRAPPSSSIKFKFRRCYAAFANENREYVIVGGG
 NAAGYAARTFVEHGMADGRLCIVTKEPYAPYERPALTKAYLFPDKKPARLPGFHTC
 VGSGGERQTPGWYKENGIEMLYEDPVKDIDIEKHTLTSSGKLLKYGSLVIATGCTAS
 RFPEKIGGNLPGVHYIRDVGDADALILSLEKAKKVVVVGgGYIGMEVAAAAGWKL
 DTTIIFPEDHLLQRLFTPSLARRYEELYQKNGVKILKGASIKNLEAGSNGHVAAVKLG
 GSLVEADTVIIGIGAKPAVTPFERVGLNTEVGGIQVDGQFRTSTPGIFAVGDVAAFPLKI
 YNRISRVEHVDHARRSAQHCVKALLSAQTHTYDYLPHYFYSRIFEYEGSPRKVWWQFF
 GDNVGETVEIGNFDPKISTFWIESGKLGVLLESgsPEEFQLLPKLARSQPLIDKAKLQ

NATSVEEALEIARESLOGEAAV

>Gs-MDHAR-3b

MAKTFKYIILGGGVSAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESP
ARLPGFHVCVSGGGERLLPEWYTEKGIELILSTEIVKVDLA AKSLISAGGETFSYQILIV
ATGSTVIRLTDGVEGADAKNIFYLREVDADKLYEAIKAKKNGKAVVVGGGYIGLE
LSAVLKLNNIDVTMVYEPWCMPRLFTAGIAEFYEEYYKNKGVNIIKGTAVGFTSNS
DGEVKEVKLKDGRVLEADIVVVGVGGRPQTALVKGQVEEEKGGIKTDAFFKTNLSD
VYAVGDVATFPLKLYGELRRVEHVDHSRKS AEQAVKAIKAAEEGRTVEEYDYLPHYFY
SRSFDLSWQFYGDNVGDTVLFGDNNPASSKPKFGTYWIKDGKVVGVFLEGGTPEEN
QAIKAVKVPVADVDQLAKEGLSFASKI

>Gs-MDHAR-3a

MAKTFKYIILGGGVSAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESP
ARLPGFHVCVSGGGERLLPEWYTEKGIELILSTEIVKVDLA AKSLISAGGETFSYQILIV
ATGSTVIRLTDGVEGADAKNIFYLREVDADKLYAAIKAKKNGKAVVVGGGYIGLE
LSAVLKLNNIDVTMVYEPWCMPRLFTAGIAEFYEGYYANKGVNIIKGTAVGFTSNS
DGEVKEVKLKDGRVLEADIVVVGVGGRPQTVLVKGQVEEEKGGIKTDAFFKTNLSD
VYAVGDVATFPLKLYGELRRVEHVDHSRKS AEQAVKAIKAAEEGKTVEEYDYLPHYFY
SRSFDLSWQFYGDNVGDTVLFGDNNPASSKPKFGTYWIKDGKVVGVFLEGGTPEEN
QAIKAVKVPVADVNQLAKEGLSFASKI

>Gs-MDHAR-2a

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
AARLPSFHTCVGANEERLTPKWYKEHGIELVLGTGVKSTDV KRKTLT TTTGETISYKI
LIVATGARALKLEEFVSGSDAENVCYLRHIADANRLVDVMQSCPGGNAV VIGGGYI
GMECAASLVINKINVTMVFP EEHC MARLFTSKIANYYEEYYKSRGVNFIKGTVLSSFD
FDSNGKVTAVNLRDGT TLSVDMVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSN
SSVYAIGDVA AFPVKAFGETRRLEHVDSARKSAIHVVSSIMEPNKTGEFDYIPFFYSRV
FTLSWQFYGDNVGEVVYYGDMSGSAFGAYWVSKGHLVGAFLEGGTKEEYEAIAKA
TRLRPAIEDLTELERQGLSFAVTVSQKPVVPPPVEVRASD L LLEKPLYAWHATAGVIAA
SIAAFAYFYGKKRRRW

>Gs-MDHAR-2b

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
AARLPSFHTCVGANEERLTPKWYKEHGIELVLGTGVKSADV KRKTLT TTTGETISYKI
LIVATGARALKLEEFVGTGSDAGNVCYLRDIADANRLVDVMQSCPGGNAV VIGGGYI
GMECAASLVINKINVTMVFP EEHC MARLFTSKIANYYEEYYKSRGVNFIKGTVLSSFD
FDSNGKVTAVNLRDGT T LSVDMVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSN
SSVYAIGDVA AFPVKAVGETRRLEHVDSARKSARHVVS AIMEPNKTREFDYLPFFYSR
IFTLSWQFYGDNVGEVVYYGDMSGSTFGAYWVSKDHLVGAFLEGGTKEEYEAIAKA
TRLRPVIEDLTELERQGLRFAVTVSEKPMASPPIEVRASD L LLEKPLYAWHATAGVIAA
SIAGFAYFYGKKRHRW

>Gs-MDHAR-1

MAAISNSLSLKHFSLSRAS P S S S S I K F K F R R C Y A A F A N E N R E Y V I V G G G N A A G Y A A R T
FVEHGMADGRLCIVTKEPYAPYERPALTKAYLFPDPKPARLPGFHTCVSGGGERQTP
GWYKENGIEMLYEDPVKDIDIEKHTLT T S S G K L L K Y G S L V I A T G C T A S R F P E K I G G N L P
GVHYIRDVG D A D A L I L S L E K A K K V V V V G G G Y I G M E V A A A V G W K L D T T I I F P E D H L L
QRLFTPSLARRYEELYQKNGVKILKGASIKNLEAGSNGHVA AVKLG D G S L V E A D T V I I

GIGAKPAVTPFERVGLNTEVGGIQVDGQFRTSTPGIFAVGDVAAFPLKIYNRISRVEHVD
 HARRSAQHCVKALLSAQTHTYDYLPYFYSRIFEYEGSPRKVWWQFFGDNVGETVEIG
 NFDPKISTFWIESGKLGVLLES GSPEEFQLLPKLARSQPLIDKAKLQNATSVEEALEIA
 RESLQGEAAV

>Mt-MDHAR-1

MAAITNSLSLKHTLSIGSPQTNLLTRTRPSSTLSFRTHRCYASYANENREYVIVGGGN
 AAGYAARTFVEHGVADGRLCIVTKEGVAPYERPALKAYLFPPDKKPARLPGFHTCVG
 SGGERQTPEWYKEKGIEMLYEDPVENIDTEKQTLTSSGKLLKYGSLIATGCTASRFP
 DKIGGNLPGVHYVRDVADADALVSSLEKTKKVVVVGGGYIGMEIAAAVAWKLDTT
 IIFPEDHLLQRLFTPSLARRYEELYQKNGVKILKGASIKSLEASSNGNGGVASVKLGDG
 SIVEADTVIIGIGAKPAVGPFERVGLNTDVGGIQVDGLFRTSIPGIFAVGDVAAFPLKIYS
 RISRVEHVDHARRSAQHCVKALLSAQTNTYDYLPYFYSRVFEYEGSPRKVWWQFFG
 DNVGETVEIGNFDPKIATFWIESGKLGVLVESGSPEVISTIFLV

>Mt-MDHAR-2

MGRAFYVYVILGGGVSAGYAALEFVKRGVSN GELCIISDES VAPYERPALS KGYLLPEA
 AARLPSFHTCV GANEERLTPK WYREHGVELVLGTGVKSADVKRKTLLTTTGETISYK
 VLIVATGARALKLEEFVNGSDAENVCYLRDIADANRLVSAIQSSPGGNVAVVGGGYI
 GMECAASLVINKISVTMVFPEEHCMARLLTPKIASYYEYYKSRGVNFVKGTVLSSFD
 FDDNGKVTAVNLKTGTKISVDMVVVGIGIRPNTGLFEGQLTLEKGGIKVNGMFQSSNS
 SVYAIGDVAAPVKA FGEMRRL EHVDSARKSAKHAVSAIMEPGKTGEFDYIPFFYSRV
 FTLSWQFFGDNVGEVVYYGEISAAGSTFGAYWVNKGHVVGAFLEGGTKEEYEAIAK
 VTRLRPAIEDLAELERQGLAFAVTVSQKPVASPPPIEVSSSTS VLLLERPLYAWHATAGV
 VLAASIAAFAYYYGKRRRRW

>Mt-MDHAR-3

MAEHSFKYIIVGGGVSAGYAAAREFVNQGVKPGELAIISKEAVAPYERPALS KAYLPES
 PARLPGFHTCV GSGGERLLPEWYSEKGVQLHLSTEIVKADLAAKSLTSAKGETFKYQ
 TLVIATGSTVIRLDFGVEGADAKNIFYLREVDDADKLYEAIKGGKNAKAVVGGGYI
 GLELSAVLRLNNIDVTMVYPEPWCMPLRFTAIEIAAFYEGYYANKGVTIKGTVATGFT
 SNSDGEVKEVKLKDGRVLEADIVVVGVGGRPQISLFGQVEEEKGGIKTDSFFKTNV
 SNVYAVGDVATFPLKLYNEVRRVEHVDHARKSAEQAAKAIIAEAGKTVEEYDYLPY
 FYSRFDLSWQFYGDNVGETVLFGDNNPASSKPNFGTYWIKDGKVVGFLEGGTPDE
 NKAIKVARALPAVEDVNQLAKEGISFASKI

>Pv-MDHAR-3

MAKTFKYIIVGGGVAAGYAAAREFVKQGVNPGELAIISKEAVAPYERPALS KAYLPESP
 ARLPGFHVCV GSGGERLLPEWYTEKGIELILSTEIVKADLAAKSLISAAGETFSYQTLV
 IATGSTVIRLDFGVEGADAKNIFYLREVDDADKLYEAIKAKKNGKAVVGGGYIGLE
 LSAVLKLNNDVTMVYPEPWCMPLRFTAIAAFYEGYYANKGVNIKGTAVGFTSNS
 DGEVNEVKLKDGRVLEADIVVVGVGGRPQTALFKGQVEEDKGGIKTDSFFKTNLSDV
 YAVGDVATFPLKLYGELRRVEHVDHSRKS AEQAVKAIKAAEDGKTVEEYDYLPYFYS
 RAFDLSWQFYGDNVGDTVLFGDNEPASP KPFYTYWIKDGKVVGVFLENGTPEENS
 AIAKVARVQPPVADVDQLAKEGLSFASKI

>Pv-MDHAR-1

MAAISNSLSLKHAFLSRPQRPFFTPAPSSSTSIKFKRCYAAFANENREYVIVGGGNSAG
 YAARTFVEHGLADGRLCIVTKEPYAPYERPALKAYLFPPDKKPARLPGFHTCV GSGG
 ERQTPEWYKENGIEMLYEDPVKDIHIENQTLTNSGKLLKYGSLIATGCTASRFPEKIG

GNLPGVHYIRNVADADALILSLEKAKKVVVVGGGYIGMEVAAAAGVWNLDTTIIFPE
 DKLLQRLFTPSLSRRYEELYQKNGVKILKGASIKKLEAGSNHVGAGVELGDGSIVEAD
 TVIIGIGAKPAVSPFESVGLNTDVGGIQVDGLFRTRIPGIFAVGDVAAPLKIYDRISRVE
 HVDHARRSAQHCVKALLSAQTHTYDYLPHYFYSRVFEYEGSPRKVWWQFFGDNVGE
 AVEIGNFDPKIATFWIESGKLGVLLES GSPEEFQLLPKLARSQPVIDKAKLQNATSVE
 EALEIARASLQGEAAV

>Pv-MDHAR-2

MGRAFVYVILGGGVAAGYAALEFVKKGVSYGELCIISDEPVAPYERPALS KGFLPEA
 AARLPSFHTCVGANEEERLTPKWYKEHGIELVLGTGVKSADV KRKTLTLLTTTGETISYKF
 LIVATGARALKLEEFVSGSDAENVCYLRDIADANRLVNVIQSCPEGNAV VIGGGYIG
 MECAASLVINKINVTMVFPEEHCMARLFTTKIANYYEYKYSRGVNFVKGTVMSSFD
 FDSNGKVTAVNLRD GSTLSVEMVVVIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSNS
 SVYAIGDVAAFPVKAFGETRRLEHVD SARKSAKHVVAAIM EPDKTGEFDYLPFFYSRI
 FTLSWQFYGDNVGEVIYYGDMMSGSAFGAYWVSKGHLVGAFLEGGNREEYEAIAKVT
 RLRPAIEDLTELERQGLGYAVTVSQKPVASPAVEVRASDLLLEKPLYAWHATAGVIIAA
 SIAAFAYFYGKKRRRW

>Va-MDHAR-1

MPSPSARRLTMAAISNSLSLKHGFSLARQPFLTRAPPSSTSIKFRRCYAAFANENREY
 VIVGGGNSAGYAARTFVEHGMADGRLCIVTKEPYAPYERPALTKAYLFPPDKKPARLP
 GFHTCVGSGGERQTPEWYKENGIEMLYEDPVKDIHIENQTLTNSGKLLKYGSLIATG
 CTASRFPEKNGGNLPGVHYIRNVADADALILSLEKSKKVVVVGGGYIGMEVAAAAGV
 WNLDTTIIFPEDQLLQRLFTPSLSRRYEELYQKNGVKVLKGASIKKLEAGSNHVGAGV
 KLDGDSMVEADTVIIGIGAKPAVSPFERVGLNTDVAGIQVDGLFRSIPGIFAVGDVAAP
 PLKIYDRISRVEHVDHARRSAQHCVKALLSAQTQTYDYLPHYFYSRIFEYEGSPRKVW
 WQFFGDNVGETVEVGNFDPKIATFWIGSGKLGVLLES GSPEEFQLLPKLARNQPLIDI
 AKLQNATTVEEALEIARASLQAAA

>Va-MDHAR-2

MGRAFVYVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALS KGFLPEA
 AARLPSFHTCVGANEEERLTPKWYKEHGIELVLGTGVKSADV KRKTLTLLTTTGETISYKF
 LIVATGARALKLEEFVSGSDAENVCYLRDIADANRLVNVIQSCPEGNAV VIGGGYIG
 MECAASLVINKINVTMVFPEEHCMARLFTTKIANYYEYKYSRGVNFVKGTVLSSFD
 FDSTGKVTAVNLRD GSTLSVDMVVVIGIRPNTGLFEGQLTLEKGGIKVNGMLQSSNS
 SVYAIGDVAAFPVKAFGETRRLEHVD SARKSAKHVVAAIM EPDKTGEFDYIPFFYSRIF
 TLSWQFYGDNIGEVVYFGDMMSGSAFGAYWVSKGQLVGAFLEGGNKEEYEAIAKVTR
 LRPAIEDLTELERQGLRYAVTVSQKPAVSPPEVEVRASDLLLEKPLYAWHATAGVIIAASI
 AAFAYFYGKKRRRW

>Va-MDHAR-3

MAKTFKYIIVGGGVAAGYAAAREFAKQGVKPGELAIISKEAVAPYERPALS KAYLPESP
 ARLPGFHVVCVSGGERLLPEWYTEKGIELILSTEIVKADLATKSLISASGETFSYQTLIIA
 TGSTVIRLSDFGVEGADAKNIFYLREIDDADKLYEAIKAKKNGKAVVVGGGYIGLELS
 AVLKLNNDVTMVYPEPWCMRPLFTADIAEFYEGYYANKGINIIKGTVAVGFTSNSDG
 EVKEVKLKDGRVLEADIVVVG VGGRPQTALFKGQVEEDKAGIKTDSFFKTNLSDVYA
 VGDVATFPLKLYGELRRVEHVDHSRKS AEQAVKAIIKAAEDGKTVEEYDYLPHYFYSRA
 FDLWQFYGDNVGDTVLFGDNKPTS AKPKFGSYWIKDGKVVGAFLEGGTPEENSAI
 AKVARAQPPVADIDQLAKEGLSFASKI

>Vr-MDHAR-3

MAKTFKYIIAGGGVAAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESP
 ARLPGFHVCVGGGERLLPEWYTEKGIELILSTEIVKADLATKSLISASGETFSYQTLIIA
 TGSTVIRLTDGVEGADAKNIFYLREIDDADKLYEAIKAKKNGKAVVVGGGYIGLELS
 AVLKLNNDVTMVYPEPWCMRLFTADIAEFYEGYYANKGINIIKGTAVVGFTSNSDG
 EVKEVKLKDGRVLEADIVVVGVGGRPQTALFKGQVEEKGGIKTDSFFKTNLSDVYAV
 GDVATFPLKLYGELRRVEHVDHSRKS AEQAVKAIKAAEDGKTVEEYDYLPYFYSRAF
 DLSWQFYGDNVGDTVLFQDNKPTS AKPKFGSYWIKDGKVVGAFLGGTPEENSAIA
 KVARAQPPVADIDQLAKEGLSFASKI

>Vr-MDHAR-2

MGRAFVYVILGGGVAAGYAALFVKKGVSHGELCIISDEPVAPYERPALSKGFLPEA
 AARLPSFHTCVGANEERLTPKWYREHGIELVLGTGVKSADVKRKTLTLLTTGETISYKF
 LIVATGARALKLEEFVSGSDAENVCYLRDIADANRLVNVIQSCPEGNNAVIGGGYIG
 MECAASLVINKMNVTMVFP EEHC MARLFTTKIANYYEEYYKSRGVNFVKGTVLSFF
 DFDSNGKVTAVNLRD GSTLSVDMVVVIGIRPNTGLFEGQLTLEKGGIKVNGMLQSS
 NSSVYAIGDVA AFPVKAFGETRRLEHVDSARKSAKHVVA AIMEPDKTGEFDYIPFFYS
 RIFTL SWQFYGDNVGEVVYFGDMSGSAFGAYWVSKGQLVGAFLEGGNKEEYEAIAK
 VTSLRPAIEDLTELERQGLRYAVTVSQKPAVSPPEVVRASDLLLEKPLYAWHATAGVIIA
 ASIAAFAYFYGKKRRRW

>Vr-MDHAR-1

MPSPSARRLTMAAISNSLSLKHGFSLARQP RPF LTPAPPYSTSIKFRRCYAAFANENREY
 VIVGGGNSAGYAARTFVEHGMADGRLCIVTKEPYAPYERPALTKAYLFPDCKPARLP
 GFHTCVGSGGERQTPEWYKENGIEMLYEDPVKDIHIENQTLTNSGKLLKYGSLIIATG
 CTASRFPEKIGGNLPGVHYIRNVEDADALILSLEKSKKVVVVGGGYIGMEVAAA AVG
 WNLDTTIIFPEDQLLQRLFTP SLRRYEELYQKNGVKILKGASIKKLEAGSNHGVAGV
 KLDG GSTVEADTVIIGIGAKPAVSPFERVGLNTDVGGIQVDGLFRTSIPGIFAVGDVAAF
 PLKIYDRISRVEHVDHARRSAQHCVKALLSAQTQTYDYLPYFYSRIFEYEGSPRKVW
 WQFFGDNVGETVEVGNFDPKIATFWIESGKLGVLLES GSPEEFQLLPKLARNQPLID
 KAKLQHATTVEEAL EIARASLQGEAAV

>Vc-MDHAR

MSSYKIIFLGGGNAAGYAARAFVENCLKAGELAIITEEPYVAYERPALSKGYLLGAAR
 LPGFHTCVGGGGERQAPEWYSEKGITYLTNSRVVKADLANKALTASGEVLSYEKLI
 GTGARPTRLTEFGVPGADLGGFLYLRDVKDGDVLA AVAATKEAGGKAVVIGGGYIG
 MEVAAGLSSSGLSVTMVFPEDRILSRLLPQLAAVYERLYDAKGIKMKVKGAKVTGFD
 GVDGK VSCVRFTDASGSQQSLDAGLVVGVGARNVELFQGGQLEIAAGGIKVDGQM
 ATSVDPVYAVGDVAAFPLTSVASGEVSYARQEHVTHCRLSAAQA AKAILGLSPPPYDY
 LPFFYSRVFALS WVFYGEAPADATAVHFGDMPEAKCFGCLWLGAGGKLVGAFLEGG
 ADDAAVLKAAVAGRLTIPTTEEGGLGAAAGSGAAVVAQLKAKL

>Vc-MDHAR

MLARVPFPRATLKRAARALTTVCNSSMAADPAKKRVLIVGGVAGGASCAARLRMD
 EQAEITIFERGPYVSFANCGLPYFIGDVIKEQSSLLVANA AKFQKWFNVTVKESTEVVA
 IDRAAKKITARDVNNGQESVHPYDYLVLSPGGHAI RPLPGADLPGIFVVKTI PDANAI
 KAWITEKGAKSAVVVGGGFIGLEVLENLVHRGLKCSLVEMLPQVMPPYDPEVVEPVH
 DRLRAAGVELHLGDGVSGFEAGPAGSAALLAVKTQSGKTHAADLVVLAIGVRPETTL
 ARQAGLELGSRRGIKVDERMRTSDSAIFAVGDVAEVKDFVTGEQTLIPLAGPANRQGR
 IVADVISGNEPSRFRGSQGTSVVGLFGLTLAATGASEKTLRRLGRTFRKVYLHINN HAG

YYPGAKQIDIKLLFDPTDGRVLGMQAVGEEGVEKRVDVVMAMAIQKGGTVFDLEEAEEL
 CYAPQYGSADKDPVNMVGMVAANVLRNQHPLTSWDELQALDADPGCVLVDVREPP
 ELDKVGGKVESAVNMPLSSLRQQLAKLQDKDKKYYVYCQVGLRGYIATRQFLQAGFD
 AVNVSGGYKSFQQQLQEQQRPAL

>Pv-DHAR-1

MAVEVAVKAAVGAAPTILGDCPFSQRVLLTLEEKKIPYKLHLIDLSNKPEWFLGPNPEG
 KVPVALFDGKWVSDSDVIVGILEEKYPEISLITPPEFATVGSKIFGSFVTFLLKSKDPNDG
 TEQALLAELSALDEHLKANGPYVAGEKVTAVDLSLAPKLYHLVVVLGHFKKWSVPES
 LAHVHNYTKLLFSRESFEKTKPPKEEYVIAGWAPKVNG

>Pv-DHAR-2

MSTARVQVSACALSTAVNHLRLRPNAVVSFTNHFRKSLRVVSMSSVPPSQPFEIAVK
 ASLTPNRLGDCPFCQRVLLTLEEKHLPYEPKLVDFTNPEWFRFTVNPDKVPVIKDFE
 KWVPDSDVITQTLEEKYPSPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQALLSEL
 SSFNNDYIKENGPFGINGNEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLSYTKAI
 FLRESFIKTSAQPDVIEGWRPKVEG

>Mt-DHAR-2

MSTVRIQVSACALSATVNHLRYRPNYAVSTSSFNHFSIKPLKVSMSAPPSEPLEVAV
 KASLTPNKIGDCPFSQRVLLTLEEKHLPYEPKLVDLRNKPEWFLEISPEGKVPVINFD
 GKWVADSDLITQTLEEKYPSPLVTPPEKATAGSKIFSTFIGFLKSKDPNDGTEQALLNE
 LSSFNNDYIKENGPFGINGKDISAADLSLGPPLYHLEIALGHYKKWTVPDSLTFLLKSYLK
 EIFSRESFINTRAQPEDVIEGWRPKVEG

>Mt-DHAR-1

MALEVAVKAAVGAAPTILGDCPFSQRVLLTLEEKKIPHNIHLINLTDKPQWFLEVNPEGK
 VPVVKFDGKWVSDSDVIVGILEDKYPEPSLVSPAQFSSVGSNIFASFSSFLKSKDSNDG
 TEQALLAELNALDEHLKANGPFVAGEKVTAVDLSLAPKLYHLVVALRHFKSWTIPESL
 AKVHNYIKLLFSRESFEKTKAAEEYIIAGWAPKVNA

>At-DHAR-1a

MALEICVKAAGAPDHLGDCPFSQRALLTLEEKSLTYKIHLINLSDKPQWFLDISPQGK
 VPVLKIDDKWVTDSVIVGILEEKYPDPPLKTPAEFASVGSNIFGTFTGFLKSKDSNDG
 SEHALLVELEALENHLKSHDGPFIAGERVSAVDLSLAPKLYHLQVALGHFKSWSVPES
 FPHVHNYMKTLFSLDSFEKTKTEEKYVISGWAPKVNP

>At-DHAR-2

MISLRFQPSTTAGVLSASVSRAGFIKRCGSTKPGRVGRFVTMATAASPLEICVKASITTP
 NKLGDPCFCQKVLLTMEKKNVPYDMKMVDLSNKPEWFLKISPEGKVPVVKFDEKW
 VPDSDVITQALEEKYPEPPLATPPEKASVGSKIFSTFVGFLLKSKDSGDGTEQVLLDELT
 TFNDYIKDNGPFINGEKISAADLSLAPKLYHMKIALGHYKNWSVPDSLFPVKSYPEN
 VFSRESFTNTRAETEDVIAGWRPKVMG

>At-DHAR-1b

MALDICVKVAVGAPDVLGDCPFSQRVLLTLEEKKLPYKTHLINVSDKPQWFLDISPEG
 KVPVVKLDGKWVADSDVIVGLLEEKYPEPSLKTPEFASVGSKIFGAFVTFLLKSKDAN
 DGSEKALVDELEALENHLKTHSGPFVAGEKITAVDLSLAPKLYHLEVALGHYKNWSV
 PESLTSVRNYAKALFSRESFENTKAKKEIVVAGWESKVNA

>Va-DHAR-2

MSTARVQVSACALSAAVNNLRLRPTAVVSFTNHFRKSLRVVSMSSVPPSQPFEIAVK
 ASVTPNRLGDCPFCQRVLLTLEEKHLPYEPKLVDFTNRPEWFLKANPDGKVPVIKDFD
 EKWVPDSDVITQTLEEKYPIPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQALLSE

LSSFNDYIKENGPYINGNEICAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKAYTK
AIFSRESFIKTSAPQPDVIEGWRPKVEG

>Va-DHAR-1

MALEVAVKAAAGAPDVLGDCPFCQRVLLTLEEKKIPYKLNLDLSNKPEWFLGVNPE
GKVPVALFDGKWVPDSDVIVGILQEKEYPEISLVTPPEFATVGSKIFGSFVSFLKSKDPND
GTEQALLAELSALDEHLKAHGPYIAGEKVTAVDLSLAPKLYHLVVVLGHFKKWSVPE
SLSHVHIYTKLLFSRESFEKTKPPKEEFVIAGWAPKVNA

>Vr-DHAR-2

MSTARVQVSACAFSAAVNNLRLRPTAVVSFTNHFRKKSLRVVSMSSVPPSQPFEIAVK
ASVTPNRLGDCPFCQRVLLTLEEKNLPEYKLVDFSNRPEWFLKANPDGKVPVVKF
DEKWVPDSDVITQTLEEKYPIPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQALLS
ELSSFNDYIKENGPYINGNEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKAYTK
AIFSRESFIKTSAPQPDVIEGWRPKVEG

>Vr-DHAR-1

MALEVAVKAAAGAPDVLGDCPFCQRVLLTLEEKKIPYKLHLIDLSNKPEWFLGVNPE
GKVPVALFDGKWVPDSDVIVGILQEKEYPEISLVTPPEYATVGSKIFGSFVSFLKSKDPN
DGTEQALLAELSALDEHLKAHGPYVAGEKVTAVDLSLAPKLYHLVVVLGHFKKWSV
PELSHVHIYTKLLFSRESFEKTKPPKEEFVIAGWAPKVNA

>Gm-DHAR-2a

MSAVRVQVSACAVSATVNHLRLRQNAVVSFRKKKPLTLRVVSMSSVPPSQPFEIAVKA
SVTPNRLGDCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDE
KWVPDSDVITQTLEEKYPSPLVTPPERATAGSKIFSTFIGFLKSKDPNDGTEQALLSEL
SSFNDYIKENGPYINGSEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKSYMKA
IFSRESFVKTSAPQPDVIEGWRPKVEG

>Gm-DHAR-2b

MSTVRVQVSACVFSATVNNLCLRQNAVVSFRKKKLLRLVSMSSVPPSQPFEIAVKASV
TPNRLGDCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDEK
WVPDSDIITQTLEEKYPSPLVTPPEKATAGSKIFSTFIGFLKSKDPNDGTEQALLSELSS
FSDYIKENGPYINGSEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKSYMKVIFS
RESFVKTSAPQPDVIEGWRPKVEG

>Gm-DHAR-1b

MALEVAVKAAVGAPNVLDGCPFSQRVLLTLEEKKIPYKLHLIDLSNKPEWFLGVNPEG
KVPVVLFDGKWVADSDVIVGILEEKYPEPSLVTPPEFASVGSKIFGSFVSFLKSKDTND
GTEQTLVAELSALDEHLKAHGPYIAGEKVTAVDLSLAPKLYHLVVALSHFKNWNIPES
LTHVHNYTKLLFSRESFEKTKPPKVEYVISGWAPKVNA

>Gm-DHAR-1a

MALEVAVKAAVGAPNVLDGCPFSQRVLLTLEEKKIPYKLHLIDLSNKPEWFLGVNPEG
KVPVVLFDGKWVADSDVIVGILEEKYPEPSLVTPPEFASVGSKIFGSFVSFLKSKDTND
GTEQALVAELSALDEHLKTHGPYIAGEKVTAVDLSLAPKLYHLVVALSHFKNWNIPES
LVHVHNYTKLLFSRESFEKTKPPKEEYVIAGWAPKVNA

>Gs-DHAR-2b

MSTVRVQVSACVFSATVNNLCLRQNAVVSFRKKKLLRLVSMSSVPPSQPFEIAVKASV

TTPNRLGDCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDEK
WVPDSDIITQTLEEKYPSPLLPPEKATAGSKIFSTFIGFLKSKDPNDGTEQALLSELSS
FSDYIKENGPFIGSEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKSYMKVIFS
RESFVKTSAQPQDVIEGWRPKVEG

>Gs-DHAR-2a

MSAVRVQVSACAVSATVNHLRLRQNAVVSFRKKKPLTLRVVSMSSVPPSQPFEIAVKA
SVTTPNRLGDCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDE
KWVPDSDVITQTLEEKYPSPLVTPPERATAGSKIFSTFIGFLKSKDPNDGTEQALLSEL
SSFNDYIKENGPFIGSEISAADLSLGPPLYHLEIALGHYKKWTVPDSLTSLKSYMKAI
FSRESFVKTSAQPQDVIEGWRPKVEG

>Gs-DHAR-1b

MALEVAVKAAVGAAPNVLGDCPFSQRVLLTLEEKIPYKLHLIDLSSKPEWFLGVNPEG
KVPLVLFDGKWVADSDVIVGILEEKYPEPSLVTPEFASVGSKIFGSFVSFLKSKDTND
GTEQTLVAELSALDEHLKAHGPYIAGEKVTAVDLSLAPKLYHLVVALSHFKNWNIPES
LTHVHNYTKLLFSRESFEKTKPPKVEYVISGWAPKVNA

>Gs-DHAR-1a

MALEVAVKAAVGAAPNVLGDCPFSQRVLLTLEEKIPYKLHLIDLSSKPEWFLGVNPEG
KVPVVLFDGKWVADSDVIVGILEEKYPEPSLVTPEFASVGSKIFGSFVSFLKSKDTND
GTEQALVAELSALDEHLKTHGPYIAGEKVTAVDLSLAPKLYHLVVALGHFKNWNIPES
LVHVHNYTKLLFSRESFEKTKPPKEEYVIAGWAPKVNA

>Ca-DHAR-1

MALEVAVKAAVGAAPNVLGDCPFSQRVLLTLEEKIQFNTHLIDVSNKPQWFLEVYPD
GKVPVLKIDDKWVPDSDVIIAILEDKYPQPPLATPPQFSSVGSKIFGSFVSFLKSKDAN
DGTEQALVAELNALEEHLKAHGPYVAGENITAVDLSLAPKLYHLVVALGHFKNWTISE
SLTHVHNYIKLLFSRESFEKTKAAEEYVIAGWAPKVNA

>Ca-DHAR-2

MSTVRIQVSACALSATVNHLRYPNYVVSTSHFNNHFSFKPLRVSMSSVPPSEPLEVV
VKASITTPNKIGDCPFSQRVLLTLEEKHLPYDPKLVNLANKPEWFLEISPEGKVPVIKFD
GKWVPDSDVITQTLEEKYPSPLITPEKATIGSKIFSTFIGFLKSKDSNDGTEQALLNE
LSSFNDYIKENGPFIGKDISAADLSLGPPLYHLEIALGHYKKWTVIPDSLTFKLSYIKDI
FSRESFINTRAQPQDVIEGWRPKVEG

>Ad-DHAR-1

MALEVAVKAAVGAAPNVLGDCPFCQRVLLTLEEKKVPYSTHLINFDEKPQWFLDVNPE
GKVPVIKFDDKWISDSVIVGILEEKYPQPSLVTPEASVGSKLFVAFVKFLKSKDPN
DGSEQALLAELKDLDEHLKNHGPFVAGEKVTAVDLSLAPKLYHLVITLEHFKNWTIPQ
DLAHVHNYTKLLFSLESFEKTKAAKEYVIAGWAPKVNA

>Ad-DHAR-2

MSTVRIQASACALYSTFNHLRHRPNSVVSLSRNAKNSSALKVLSMSSSPPSAPFEICVK
ASITTPNKLGDPCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKINAEGKVPVIKLDE
KWVSDSDIITQILEEKYPSPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQALLSELS
SFNDYIKENGPVNGKDISAADLSIGPKLYHLEIALGHYKKWTVIPDSLQSLKSYMKDI
FSRESFIKTRAQPQDVIEGWRPKVEG

>Ai-DHAR-1

MALEVAVKAAAGAPNVLGDCPFCQRVLLTLEEKKVPYSTHLINFDEKQPWFVDVNPE
GKVPVLKFDDKWISDSVIVGILEEKYPQPSLVTTPESASVGSKLFVAFVKFLKSKDPN
DGSEQALLAELKDLDEHLKNHGPFVAGEKVTAVDLSLAPKLYHLVITLHFKNWTIPQ
DLAHVHNYTKLLFSLESFEKTKAAKEYVIAGWAPKVNA

>Ai-DHAR-2

MSTVRIQASACALYSTFNHLRHRPNSVVSVCRNAKNSSALKVLSMSSPPSAPFEICV
KASITTPNKLGDPCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKINAEGKVPVIKLD
EKWVSDSDIITQILEEKYPSPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQALLSEL
SSFNDYKENGPLVNGKDISAADLSLGPPLYHLEIALGHYKKWKVPDSLQSLKSYMK
DIFSRESFIKTRAQPQDVIEGWRPKVEG

>Cc-DHAR-1

MALEVAVKAAVGAAPNILGDCPFSQRVLLTLEEKKVSYKHLHLIDLSNKPEWFLGVNPEG
KVPVVQFDGKWWADSDVIVGILEEKYPQPSLLTPPEFASVGSKIFGSFVSFLKSKDSND
GTEQALVSELSALDDHLKAHGPYVAGEKVTAVDLSLAPKLYHLVITLGHFKSWSVPES
LAHVHNYTKLLFSRESFEKTKAAKEYVIAGWAPKVNA

>Cc-DHAR-2

MSTVRVQVSACALTATVNHLRLRPNAVVSFPNNFALRKKPLRVLSMSSLPPSQPFEVA
VKASVTTPNRLGDPCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVI
KFDEKWWPDSVITQILEEKYPSPLVTPPEKATVGSKIFSTFIGFLKSKDPNDGTEQA
LTELSSFNDYIKENGPFGKEISAADLSLAPKLYHLEIALGHYKKWTVPDSLTYLKS
YTKAIFSRESFIKTSAPQDVIEGWRPKVEG

>Vc-DHAR

MLLPRVSPQRYLTRATTYKHLRSVSVAAMANTVIYVKGDPANKTLGDPCFCHRALLT
FERKKVPYTLDYIDFANKPSWLQDVS GGKVPVIKEDGQPYMPDSVIVVHLEEKYPE
PSMKSSVPPEIGAKLFFAFRGVLMGPPEELADKQALLISELKAMNDYLEAHQAEGPLF
GGHQINATDAAVAPKLYHAVVALKHFKGWELPPEFAAVRRYMAAIQQLPEWKKTDY
GEAMIIKGWERHMAHH

>At-GR-2

MASTPKLTSTISSSSPSLQFLCKKLPPIAHLPSSSSSSFLSLPKTLTSLYSLRPRIALLSNHR
YYHSRRFSVCASTDNGAESDRHYDFDLFTIGAGSGGVRAARFASFGASA AVCEL PFS
TISSDTAGGVGGTCVLRGCVPKLLVYASKYSHEFEDSHGFGWKYETEPSHDWTTLI
ANKNAELQRLTGIYKNILSKANVKLIEGRGKVIDPHTVDVDGKIYTTNRNIIAVGGRPF
IPDIPGKEFAIDSDAALDLPSPKPKIAIVGGGYIALEFAGIFNGLNCEVHV FIRQKKVLR
GFDEDVRDFVGEQMSLRGIEFHTEESPEAIKAGDGSFSLKTSKGTVEGF SHVMFATG
RKPNTKNLGLENVGVKMAKNGAIEVDEYSQTSVPSIWA VGDVTD RINLTPVALMEGG
ALAKTLFQNEPTKPDYRAVPCAVFSQPPIGTVGLTTEEQAIEQYGDVDVYTSNFRPLKA
TLGSLPDRVFMKLIVCANTNKVLGVHMCGEDSPEIIQGFVAVKAGLTKADFDATVG
VHPTAAEEFVTMRAPTRKFRKDSSEGKASPEAKTAAGV

>At-GR-1

MARKMLVDGEIDKVAADEANATHYDFDLFVIGAGSGGVRAARF SANHGAKV GICEL
PFHPISSEEIGGVGGTCVIRGCVPKKILVYGATYGGEELEDAKNYGWEINEKVDFTWKK
LLQKKTDEILRLNNIYKRLLANAAVKLYEGEGRVVGPNVEVEVRQIDGTKISY TAKHILI
ATGSRAQKPNIPGHELAI TSDEALSLEEFPKRAIVLGGGYIAVEFASIWRGMGATVDLF

FRKELPLRGFDDEMRAALVARNLEGRGVNLHPQTSLTQLTKTDQGIKVISSHGEEFVAD
 VVLFATGRSPNTKRLNLEAVGVVELDQAGAVKVDEYSRTNIPSIWAVGDATNRINLTPVA
 LMEATCFANTAFGGKPTKAEYSNVACAVFCIPPLAVVGLSEEEAVEQATGDILVFTSGF
 NPMKNTISGRQEKTLMKLVDEKSDKVIGASMC GPDAAEIMQGIAIALKCGATKAQF
 DSTVGIHPSSAEFVTMRSVTRRIAHPKPKTNL

>Ad-GR-1b

MESRKMLSEAETNKSADQVASFDFDLFVIGGGSGGVRAARFSANFGAKVGICELPYH
 PISSDKHGGFGGTGC VIRGCVPKKILVYGGTFGGDLEDARNFGWELSEKIDFNWKKLL
 QKKTDEISRLNGVYKRLSSAGVKT FEGQGKITGPNEVEVTQIDGTKLSY TAKHILIAT
 GSRAQFPNIPGQELGITSDEALSLEEIPKRVVHGAGYIAVEFASIWRGMGSEVHLVYR
 KELPLRGFDDEMRAVVAKNLEGRGIVLHPSTNLAQLIKTEDGIKVITDHGEELMADTV
 LFATGRAPNSKRLNLEAVGVQVDKIGAIKVDEYSRTNVPSIWAIGDVTNRMNLTPVAL
 MEGTYFANTVFGNRPSKPDYSNIPCAVFCNPPLSCVGLSEEQAIKQANGDVLVFTSSF
 NPMKNTISGRQEKTVMKLVDAATDKVLGASMC GPDAPEIMQGIAIALKCDATKAQ
 FDTTVGIHPSSAEFVTMRSVTRRIPVGTTPKTNL

>Ad-GR-2

MAVATSLTVTTPKKLSSLSSPTLHSLFFSKALSFS PSSSSFLSLRKTLLLSPSHHTRTRRR
 SFTLSASANPHNYDFDLFTIGAGSGGVRASRFAANYGASVAICELPFSTISSDVTGGVG
 GTCVLRGCVPKKLLVYSSKYSHEFEESNGFGWKYESEPKHDWSTLIANKNAELQRLT
 GIYKNILKNSNVQLIEGRGKIVDPHTVDVDGKLYTARHILSVGGRRPFIPDIPGSEHAID
 SDAALDLPKPEKIAIVGGGYIALEFAGIFNGLKSDVHV FIRQKKVLRGFDEEIRDFVG
 EQMALRGIEFHTEESPQAIVKSADGSLSLKTNKGTVEGF SHIMFATGRKPNTKNIGLES
 VGVKIDKKGAIENVESQSSVPSIWAIGDVTDRINLTPVALMEGVALAKTLFLNEPTK
 PDYSYVPSAVFSQPPIGQVGLTEEQAVEQYGDVDIFTSNFRPLKATLSGLPDRTFMKLI
 VCAKTNKVLGLHMCGEDSPEITQGFAVAIKAGLTKGDFDATVGIHPTAAEFVTMRT
 TRKIRKTNPSEGRSECEVKAVAGT

>Ad-GR-1a

MARKMLIDGEPAAALEEEATDFDFDLFVIGAGSGGVRASRFSANFGAKVGICELPYHP
 ISSETIGGVGGTGC VIRGCVPKKILVYGASFGGDLEDARNYGWELSEKIDFNWKKLLQK
 KTDEINRLNGIYKRLSSNAGVKLFEGEGKVVG PHEVEVTQVDGTKLSYS AKHILIATG
 GRAHLDPDIPGKELGITSDEALSLEELPKRVVILGGGYIAVEFASIWQGMGSTVNLVFRK
 ELPLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTEDGIKAVTDHGEELMADVVL
 FATGRVPNTKRLNLEAVGVVELDSNGAIKVNEYSRTNIPSIWAVGDVTD RMNLTPVALM
 EGTCFAKTVFGAQESKPDYSNIPYAVFSIPPLSVVGLSEEQAI EQANGDVLVFTSTFNP
 MKNTISGRQEKTVMKLVDAETDKVLGASMC GPDAPEIMQGI AVALKCGATKAQFD
 STVGIHPSSAEFVTMRSVTRTVA AAKPRTSL

>Ai-GR-1b

MESRKMLSEAETNKSADQVASFDFDLFVIGGGSGGVRAARFSANFGAKVGICELPYH
 PISSDKHGGFGGTGC VIRGCVPKKILVYGGTFGGDLEDARNFGWELSEKIDFNWKKLL
 QKKTDEISRLNGVYKRLSSAGVKT FEGQGKITGPNEVEVTQIDGTKLSY TAKHILIAT
 GSRAQFPNIPGQELGITSDEALSLEEIPKRVVHGAGYIAVEFASIWRGMGSEVHLVYR
 KELPLRGFDDEMRAVVAKNLEGRGIVLHPSTNLAQLIKTEDGIKVITDHGEELMADTV
 LFATGRAPNSKRLNLEAVGVQVDKIGAIKVDEYSRTNVPSIWAIGDVTNRMNLTPVAL
 MEGTYFANTVFGNRPSKPDYSNIPCAVFCNPPLSCVGLSEEQAIKQANGDVLVFTSSF
 NPMKNTISGRQEKTVMKLVDAATDKVLGASMC GPDAPEIMQGIAIALKCGATKAQ
 FDTTVGIHPSSAEFVTMRSVTRRIPVGTTPKTNL

>Ai-GR-1a

MARKMLIDGEPAAALEEEATDFDFDLFVIGAGSGGVRASRFSANFGAKVGCICELPYHP
 ISSETIGGVGGTCVIRGCVPKKILVYGASFGGDLEDARNYGWELSEKIDFNWKKLLQK
 KTDEINRLNGIYKRLLSNAGVKLFEFEGEKVVGPEVEVTQLDGTKLSYSAKHILIATG
 GRAHLDPDIPGKELGITSDEALSLEELPKRVVILGGGYIAVEFASIWQGMGSTVNLVFRK
 ELPLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTEDGIKAVTDHGEELIADVLF
 ATGRVPNTKRLNLEAVGVELDSNGAIKVNEYSRTNIPSIWAVGDVTDNRMNLTVALME
 GTCFAKTVFGAQESKPDYSNIPYAVFSIPPLSVVGLSEEQAIEQANGDVLVFTSTFNPM
 KNTISGRQEKTVMKLVDAETDKVLGASMCGBPDAPEIMQGIKVAALKCGATKAQFDST
 VGIHPSSAEFVTMRSVTRTVAAAKPRTSL

>Ai-GR-2

MAVATSLTVTTPKKLSSLSLSSPTLHSLFFSKALSFSPPSSSFLSLRKTLLSPSHHTRTR
 RRSFTLSASANPHNYDFDLFTIGAGSGGVRASRFAANYGASVAICELPFSTISSDVTGG
 VGGTCVLRGCVPKKLLVYSSKYSHEFEESNGFGWKYESEPKHDWSTLIANKNAELQ
 RLTIYKNILKNSNVQLIEGRGKIVDPHTVDVDGKLYTARHILVSVGGRRPFIPDIPGSEH
 AIDSDAALDLPKPEKIAIVGGGYIALEFAGIFNGLKSDVHVFIQKQVLRGFDEEIRDF
 VGEQMALRGIEFHTEESPQAVVKSADGSLSLKTNKGTVEGFHSHIMFATGRKPNTKNIG
 LESVGVKIDKKGAIEVNEYSQSSVPSIWAVGDVTDNRINLTVALMEGVALAKTLFLNEP
 TKPEYSYVPSAVFSQPPIGQVGLTEEQAVEQYGDVDIFTSNFRPLKATLSGLPDRTFMK
 LIVCAKTNKVLGLHMCGEDSPEITQGFVAIAKAGLTKADFDATVGIHPTAAEEFVTMR
 TPTRKIRKTNPSEGRSESEVKAVAGT

>Cc-GR-1

MARKMLIDGEPELVQGGEEGTHFDFDLFIIGAGSGGVRARFSSNFGAKVGCICELPFH
 PISSETVGGVGGTCVIRGCVPKKILVYGASFGGDLEDARNYGWELNEKVDFNWKLL
 QKKTDEINRLNGVYKRLLSNAGVKLFEFEGEKIAGPNEVEVTQLDGTKLSYSAKHILIA
 TGGRAQRPNIPGQELGITSDEALSLEELPKRAVVLGGGYIAVEFASIWRGMGSTVDLVF
 RKELPLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKSEDGIKVITDHGEELIADV
 LFATGRAPNSKRLNLEAVGVELDKIGAIKVDEYSRTNIPSIWAVGDVTNRMNLTVAL
 ESSCLSKTIFGGQSIKPDYSNIPCAVFCIPPLSVVGLSEEQAIEQTKGDLLIFTSTFNPMK
 NSISGRQEKTIMKLVDAETDKVLGASMCGBPDAPEIIQGIKVAALKCGATKAQFDSTVGI
 HPSSAEFVTMRSVTRRVTGSGKPKTNL

>Cc-GR-2

MATSLSFSPSLNNTLFFAKTFPLSRPSFLPLSKSLLSLSTRRRTFVRAESQNGVDAAS
 ANYDFDLFTIGAGRGGVRSSRFAANYGASVAICELPFSTISSDTTGGVGGTCVIRGCV
 KLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTIYKNILNNSG
 VKLIEGHGKIIDPHTVDVNGKLYSAKHILVAVGGRRPFIPDIPGKEYAIDSDASLDLASKP
 EKIAIVGGGYIALEFAGIFSGLKSEVHVFIQKQVLRGFDEEIRDFVAEQMSLRGIEFHT
 EESPQAITKSADGSFSLKTNKGTVDGFHSHIMFATGRRPNTKNLGLSVGVKLAKDGAI
 EVDEYSQTSVPSIWAVGDVTNRINLTVALMEGGALVKTLFQDNPTKPDYRAVPSAVF
 SQPPIGQVGLTEEQAVQQYGDIDIYTANFRPLKATLSGLPDRTFMKLVCAKTNEVIGL
 HVCGEDAPEIVQGFVAIAKARLTKDDFDATVGIHPSAAEEFVTMRTPTRKIRKSESSEG
 KSGSEVKTAAGV

>Ca-GR-2

MATSLSLSCCSPTFTLTRNTLFFTKTVPFSPFSSNPSTNTLFSLSPPRRYFAIRADSQNA
 VDPARHYDFDLFTIGAGSGGVRASRFATTFGASAAICELPFSTISSDTTGGVGGTCVIR

GCVPKCLFVYSSKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTGIYKNIL
 KNSGVKCLIEGRGKIIDPHTVDVDGKLYSAKHILVSVGGRRPFIPDIPGKEYAIDSDAALD
 LPSKPEKIAIVGGGYIALEFAGIFNGLKSEVHVFIHQKKVLRGFDEEVRDFVAEHMSLR
 GIEFHTEESPVAITKAADGSLSLKTNKGTEEGFSHIMFATGRRPNTKNLGLSVGVKM
 ANDGSIEVDEYSQTSVPSIWAIGDVTNRINLTPVALMEGVALSKTLFQNEPTKPDYRAI
 PSAVFSQPPIGVVGLTEEQAVEQYGDVDIFTANFRPLKATLSGLPDRVLMKLVSAKTN
 KVLGLHMCDDAAEIMQGFVAIAKAGLTKADFDATVGIHPTAAEEFVTMRTPTRKIRK
 SHASEGKSDSEVKTVAR

>Ca-GR-1

MARKMLIDGEPDLRKGEEEGTNFDFDLFVIGAGSGGVRAARFSANFGAKVGCICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASYGGELEDARNYGWELSENVDNFNWKKL
 LQKKTDEINRLNGIYKRLLSNAGVKLFEGEGKIVGPNEVEVTQLDGTKLSYSAKHILI
 ATGSRAQRPNIPGQELGITSDEALSLEEFKRAVILGGGYIAVEFASIWRGMGSNVHLV
 FRKELPLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTDDGIKVIDDHGEELIADV
 VLFATGRAPNSKRLNLEDVGVELDKTGAIVVNEDSRTNIPSIWAVGDVTNRMNLTTPVA
 LMEASCFAQTVFGGQPSKPDYNDIPYAVFSIPPLSVVGLSEEQAIEQTNGDVLVFTSTF
 NPMKNTISGRQEKTVMKLVDAETDKVLGASMCQPDAPETIIQGIAIALKCGATKAQF
 DSTVGVHPSAAEEFVTMRSVTRRVTAATAKPKTNL

>Gm-GR2-a

MATSLSVSPSLSLNTLFFAKALPLSRPSLLSLPLPKSLLSLSTRRRTFAVRAESQNGADA
 VPAHYDFDLFTIGAGSGGVRAARFAANYGASVAICELPFSTISSETTGGVGGTCVIRGC
 VPKKLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTGIYKNILNN
 AGVKLIEGHGKIIDPHTVDVNGKLYSAKHILVTVGGRRPFIPDIPGNEYAIDSDAALDLP
 TKPEKIAIVGGGYIALEFAGIFNGLKSEVHVFIHQKKVLRGFDEEIRDFVSEQMSVRGIE
 FHTEESPQAITKSADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNLGLSVGVKIAKD
 GAIEVDEYSQTSVPSIWAIGDVTNRINLTPVALMEGGALVKTTLFQDNPTKPDYRAVPS
 AVFSQPPIGQVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKTNEV
 LGLHMCDDAPEIVQGFVAIAKARLTKADFDATVGIHPSAAEEFVTMRTPTRKIRKSE
 SSEGKSGSEAKTAAGV

>Gm-GR-1b

MARKMLIDGEAEIAQGGGEEVSNFDFDLFIIGAGSGGVRAARFSSNFGAKVGCICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASFGGDLQDARNYGWELSEKVDNFNWKRL
 LQKKTDEINRLNGLYKRMLSNWGVKLFEGEGKIVGPNEVEVTQLDGTKLSYSAKHIL
 IATGGRAQRPNIPGQELGITSDEALSLEDLPKRAVVLGGGYIAVEFASIWRGMGSTVDL
 VFRKELPLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTEDGIKVIDDHGEELIAD
 VVLFATGRAPNSKRLNLEAVGVELDNTGAIKVDEYSCTSIPSIWAVGDVTNRMNLTTPV
 ALMEASYFAKTVFGGQKIKPDYNNIPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSIF
 NPMKNTISGRQEKTVMKLVDAETDKVLGASMCQPDAPETIIQGIAIALQCGATKAQF
 DSTVGIHPSAAEEFVTMRSVTRRVAGTGSAKPKTNL

>Gm-GR-2b

MATSLSVSPSLSLNTLFIKALPLSRPSFSLPLPKSLLSLSTRRRTFIVRAESQNGADPV
 PAHYDFDLFTIGAGSGGVRAARFAANYGASVAICELPFSTISSETTGGVGGTCVIRGCV
 PKKLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSFIANKNAELQRLTGIYKNILNNA
 GVKLIEGHGKIIDPHTVDVNGKLYSAKHILVAVGGRRPFIPDIPGKEYAIDSDAALDLPTK
 PVKIAIVGGGYIALEFAGIFNGLKSEVHVFIHQKKVLRGFDEEIRDFVVEEQMSVRGIEF
 HTEESPQAITKSADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNLGLSVGVKLAKD

AIEVDEYSQTSVSSIWAVGDVTNRINLTPVALMEGGALVKTLFQDNPTKPDYRAVPSA
 VFSQPPIGQVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKTNEVL
 GLHMCGEDAPEIVQGFAVALKARLTKADFDATVGIHPSAAEEFVTMRTPTRKIRKSES
 SEGKSGSQAKAAAGV

>Gm-GR-1a

MARKMLIDGEAELAQGGGEEVSNFDFDLFIIGAGSGGVRAARFSSNFGAKVVICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASFGGDLEDARNYGWELSEKVDNFNWKKL
 LQKKTDEINRLNGLYKRMLSNWGVKLFEGEGKIVGPNEVEVTQLDGTKLSYTAKHIL
 IATGSRAQRPNIPGQELGITSDEALSLEDLPKHAVVLGGGYIAVEFASIWRGMGSTVDL
 VFRKELPLRGFDDEMRAAVARNLEGRGINLHPRTNLTQLIKTEDGIKVITDHGEELIAD
 AVLFATGRAPNSKRLNLETVGVELDNTGAIKVDEYSRTSIPSIWAVGDVTNRMNLTPV
 ALMEASYFAKTVFGGQTIKPDYSNIPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSTF
 NPMKNTISGRQEKTVMKLVVDAETDKVLGASMCGBPDAPEIMQGIAVALKCGATKAQ
 FDSTVGIHPSAAEEFVTMRSVTRRVAGTGSVKPKTNL

>Gs-GR-2a

MATSLSVSPSLSLNTLFFAKALPLSRPSLLSLPLPKSLLSLSTRRRTFAVRAESQNGADA
 VPAHYDFDLFTIGADSDGVRAARLATNYGASVAICEFPFSTISSETTGGVDGTCVIRGC
 VPKKLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTGIYKNILNN
 AGVKLIEGHGKIIDPHTVDVNGKLYSAKHILVTVGGRPFIPIPDIPGKEYAIDSDAALDLP
 TKPEKIAIVGGGYIALEFAGIFNGLKSEVHVVFIRQKKVLRGFDEEIRDFVSEQMSVRGIE
 FHTEESPQAITKSADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNLGLESVGVKIAKD
 GAIEVDEYSQTSVPSIWAVGDVTNRINLTPVALMEGGALVKTLFQDNPTKPDYRAVPS
 AVFSQPPIGQVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKTNEV
 LGLHMCGDDAPEIVQGFAVALKARLTKADFDATVGIHPSAAEEFVTMRTPTRKIRKSE
 SSEGKSGSEAKTAAGV

>Gs-GR-2b

MATSLSVSPSLSLNTLFIKALPLSRPSFSLPLPKSLLSLSTCHRFTTARFESQIDVDLVP
 AHYDFNLFTIDAGSGRVRAARFAANYGASVAICELPFSTISSETTGGVGGTCVIRGCV
 KLLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSFIANKNAELQRLTGIYKNILNNAG
 VKLIEGHGKIIDPHTVDVNGKLYSAKHILVAVGGRPFIPDIPGKEYAIDSDAALDLPTKP
 VKIAIVGGGYIALEFAGIFNGLKSEVHVVFIRQKKVLRGFDEEIRDFVEEQMSVRGIEFH
 TEESPQAITKSADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNIGLESVGVKLAKDGA
 EVDEYSQTSVSSIWAVGDVTNRINLTPVALMEGGALVKTLFQDNPTKPDYRAVPSAVF
 SQPPIGQVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKTNEVLG
 LHMCGEDAPEIVQGFAVALKARLTKADFDATVGIHPSAAEEFVTMRTPTRKIRKSESSE
 GKSGSQAKAAAGV

>Gs-GR-1b

MARKMLIDGEAIEAQQGGGEEVSNFDFDLFIIGAGSGGVRAARFSSNFGAKVVICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASFGGDLDARNYGWELSEKVDNFNWKRL
 LQKKTDEINRLNGLYKRMLSNWGVKLFEGEGKIVGPNEVEVTQLDGTKLSYSAKHIL
 IATGGRAQCPNIPGQELGITSDEALSLEDLPKRAVVLGGGYIAVEFASIWRGMGSTVDL
 VFRKELPLRGFDDEMRAVAVARNLEGRGINLHPRTNLTQLIKTEDGIKVITDHGEELIAD
 VVLFATGRAPNSKRLNLEAVGVELDNTGAIKVDEYSCTSIPSIWAVGDVTNRMNLTPV
 ALMEASYFAKTVFGGQKIKPDYNNIPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSIF
 NPMKNTISGRQEKTVMKLVVDAETDKVLGASMCGBPDAPEIIGQIAIALQCGATKAQF
 DSTVGIHPSAAEEFVTMRSVTRRVAGTGSAPKPTNS

>Gs-GR-1a

MARKMLIDGEAELAQGGGEEVSNFDFDLFIIGAGSGGVRAARFSSNFGAKVVICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASFGGDLEDARNYGWELSEKVDNFNWKLL
 LQKKTDEINRLNGLYKRMLSNWGVKLFEGEGKIVGPNEVEVTQLDGTKLSYTAKHIL
 IATGSRAQRPNIPGQELGITSDEALSLEDLPKHAVVLGGGYIAVEFASIWRGMGSTVDL
 VFRKELPLRGFDDEMRAAVARNLEGRGINLHPRTNLTQLIKTEDGIKVITDHGEELIAD
 AVLFATGRAPNSKRLNLETVGVELDNTGAIKVDEYSRTSIPSIWAVGDVTNRMNLTVP
 ALMEASYFAKTVFGGQTIKPDYSNIPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSTF
 NPMKNTISGRQEKTVMKLVDAETDKVLGASMCGPDAPEIMQGI AVALKCGATKAQ
 FDSTVGIHPSAAEEFVTMRSVTRRVAGTGSVKPKTNL

>Mt-GR-1

MARKMLKDGEPLDNKGAEETNDFDFDLFIIGAGSGGVRAARFSSNHGAKVAICELPF
 HPISSETIGGVGGTCVIRGCVPKKILVYGASYGGDLEDARNFGWELSENIDFNWKKLL
 QKKCLKNKDPSCLSNTDEINRLNGIYKRLSNAGVKLFEGEGKIAGPHEVEVTQLDG
 TKLSYSAXHILIATGSRAQRPNIPGQELGITSDEALSLEEFPKRVVILGGGYIAVEFASI
 RGMGANVNLVFRKELPLRGFDDEMRAVAVARNLEGRGINLHPRTNLTQLIKTEDGIKV
 TTDHGEELIADVLFATGRAPNSKRLNLEKAGVEVDKTGAIVVNEYSCNTNIPSIWAVG
 DVTNRMNLTVPALMEASHFANTVFGGKTQKPDYRDIPYAVFSIPPLSVVGLSEEEAVE
 QTNGDLLVFTSTFNPMKNTISGRQEKTVMKIVVDAQTDKVLGASMCGPDAPEIVQGI
 AIALKCGATKAQFDSTVGIHPSAAEEFVTMRSVTRRVGTGSAKPKTNL

>Mt-GR-2

MKQAAMATSLSLSCSSPLLTRNTLFFTKTFSFSRSFSLSPPLSTKTLISQSPRRRTFAVRA
 ESQNGVDPTGHYDFDLFTIGAGSGGVRAARFANFGANAAICELPFSTISSDTTGGVG
 GTCVIRGCVPKKLLVYASKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTG
 IYKNILKNSNVKLIENGRGKIVDPHTVDVDGKLYTAKHILVSVGGRRPFIPDIPGKEYAIDS
 DAALDLPSKPEKIAIVGGGYIALEFAGIFNGLKSDVHVFIRQKQVLRGFDEEIRDFVAE
 NVALRGIEFHTEESPVAITKAADGSLSLKTSKGTVDGFSHIMFATGRKPNTKNLGLS
 VGVKIAKDGSIEVDEYSQTSVPSIWAIGDATNRINLTVPALMEGVALTKTLFQNEPTKP
 DYRAVPAAVFSQPPIGVGFTEEQAVEQDGDVDIYTANFRPLKATLSGLPDRVFMKLV
 SAKTNEVLGLHMCDDAPEIVQGFAVAIKAGLTKADFDATVGIHPTAAEEFVTMRTP
 RKIRKSLDSQGKPDSEAKAVAGT

>Pv-GR-2

MATSFSPSLSLNTLSRNTLFFTKTIPFSRSSFLPRSLSKSLISISTRCTFTVRAQSQNG
 ADAVATHYDFDLFTIGAGSGGVRAARFAANNNGASVAICELPFSTVSSETTGGVGGTCV
 IRGCVPKKLLVYASKFAHEFEESNGFGWRYGSEPKHDWSSLIANKNAELQRLTGIYKN
 ILNAGVKLIEGHGKIIDSHTVDVNGKQYSAXHILVAVGGRRPFIPDIPGKEYAIDSDIAL
 DLPSKPGKIAIVGGGYIALEFAGIFNGLQSEVHVFIRQKKVLRGFDEEIRDFITEQMSLR
 GIEFHNEESPQAITKSADGTFLSKTNKGTVDGFSHIMFATGRRPNTKNLGLTAGVKL
 AKDGAIEVDEYSQTSVPSIWAIGDVTNRMNLTVPALMEGGALVKTLFQDNPTKPDYRA
 VPSAVFSQPPIGVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRAFMKLLVSAKT
 NQVVGLHMCGEDAPEIIQGFAIAIKAGLTKAEFDATVGVHPSAAEEFVTMRTPTRKIR
 KSQSSEGKSGSDVKAAGV

>Pv-GR-1

MARKMLIDGEAELVPGVEQGMNDFDFDLFIIGAGSGGVRAARFSSNYGAKVVICELPF
 HTISSETIGGVGGTCVIRGCVPKKILVYGASYGGELEDARNYGWELSEKVDNFNWKLL

LQKKTDEINRLNGIYKRLLSNAGAKLFEGEGKIVGPNEVEVTQLDGTKLSYSAXHILI
 ATGSRAQRPNIPGQELGITSDEALSIEDLPKRVLGGGYIAVEFASIWRGMGSEVHLVF
 RKELLRLRGFDDEMRTVVARNLEGRGINLHPRTNLTQLIRTENGIKVIDHGEELIADV
 LFATGRAPNSKRLNLEAVGVELDKAGAIKVDEYSRTSIPTIWAAGVDVTNRMNLTVAL
 MEASCFAKTVFNGQTSKPDYSNIACAVFSIPPLSVVGLSEEEAIEQTKGDLLIFTSSFP
 MKNTISGRQEKTVMKLIVEAETDKVLGASMCPPDAPEIIQGIKALCGATKAQFDST
 VGIHPSASEEFVTMRTVTRTVTGNGKPKTNL

>Va-GR-2

MATSLSFSPSLNNTLFFNKTLFPFRSLLRRPFSKSLIPIFTRRCTFAVRAQSQNGADAV
 AVHYDFDLFTIGAGSGGVRAARFAANNGASVAICELPFSTVASETTGGVGGTCVIRGC
 VPKKLLVYASKFSHEFEESHGFGWSYDSEPKHDWSSLIANKNAELQRLTGIYKNILNN
 AGVKLIEGHGKIIDAHTVDVNGKQYSAXHILVAAGGRPFIPDIPGKEYAIDSDIALDLPS
 QPGKIAIVGGGYIALEFAGIFNGLKSDVHLFIRKKVLRGFDEEIRDFIGEQLSRGIEIH
 PEESPQAITKSADGSFTLKTNGKTMNGFSHIMFATGRRPNTKNIGLESAGVKLAKDGA
 IEVDEYSQTSVPSIWAAGVDVTNRINLTVALMEAGALVKTLFQDSPTKPDYRAVPAAVF
 SQPPIGQVGLTEEQALQQYGDIDIFTADFRPLKSTLSGLPDRVFMKVVVSAKTNEVLG
 LHMCGEDAPEIIQGFAIAIKAGLTKADFDATVGVHPSAAEEFVTMRTPTRRIRKSQSS
 GKSGSVKAAAGV

>Va-GR-1

MARKMLIDGEAESVPGVEQGFHDFDLFIIGAGSGGVRAARFSSNYGAKVVICELPFN
 AISSETIGGVGGTCVIRGCVPKKILVYGASFGGELEDARNYGWELSEKVDNFNWKLL
 QKKTDEINRLNGIYKRLLSNAGVCLFEGEAKIVGPNEVEVTQLDGTKLSYSAXHILIA
 TGSRAQRPNIPGQELGITSDEALSIEDLPKRAVVLGGGYIAVEFASIWRGMGSEVDL
 RKELLRLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTENGIKVIDHGEELIADV
 LFATGRAPNSRRLNLEAVGVEVDKAGAIKVDEYSRTSIPSIWAAGVDVTNRMNLTVAL
 MEASCFAKTVFNGQTSKPDYSNIACAVFSIPPLSVVGLSEEQAIEQTKGDLLIFTSTFP
 MKNTISGRQEKTVMKLVDAETDKVLGASMCPPDAPEIIQVFL

>Vr-GR-2

MATSLSFSPSLNNTLFFNKTLFPFRCSFLRRPLSKSLIPIFTRRCTFAVRAQSQNGADAV
 AVHYDFDLFTIGAGSGGVRAARFAANNGASVAICELPFSTVASETTGGVGGTCVIRGC
 VPKKLLVYASKFSHEFEESHGFGWSYDSEPKHDWSSLIANKNAELQRLTGIYKNILNN
 AGVKLIEGHGKIIDAHTVDVNGKQYSAXHILVAAGGRPFIPDIPGKEYAIDSDIALDLPS
 QPGKIAIVGGGYIALEFAGIFNGLKSDVHLFIRQKVVLRGFDEEIRDFIGEQLSRGIEIH
 PEESPQAITKSADGSFTLKTNGKTMNGFSHIMFATGRRPNTKNIGLESAGVKLAKDGA
 IEVDEYSQTSVPSIWAAGVDVTNRINLTVALMEAGALVKTLFQDNPTKPDYRAIPSAV
 SQPPIGQVGLTEEQAVQQYGDIDIFTADFRPLKSTLSGLPDRVFMKVVVSAKTNKVLG
 LHMCGEDAPEIIQGFAIAIKAGLTKADFDATVGVHPSAAEEFVTMRTPTRRIRKSESSG
 GKSGSVKAAAGV

>Vr-GR-1

MARKMLIDGEAESVPGVEQGFHDFDLFIIGAGSGGVRAARFSSNYGAKVVICELPFH
 TISSETIGGVGGTCVIRGCVPKKILVYGASFGGELEDARNYGWELSEKVDNFNWKLL
 QKKTDEINRLNGIYKRLLSNAGVCLFEGEGKIVGPNEVEVTQLDGTKLSYSAXHILIA
 TGSRAQRPNIPGQELGITSDEALSIEDLPKRAVVLGGGYIAVEFASIWRGMGSEVDL
 RKELLRLRGFDDEMRAVVARNLEGRGINLHPRTNLTQLIKTENGIKVIDHGEELIADV
 LFATGRAPNSKRLNLEAVGVEVDKAGAIKVDEYSRTSIPSIWAAGVDVTNRMNLTVAL
 MEASCFAKTVFNGQTSKPDYSNIACAVFSIPPLSVVGLSEEQAIEQTKGDLLIFTSTFP

MKNTISGRQEKTVMKLVVDAETDKVLGASMCGPDAPEIIQGIAIALKCGATKAQFDS
TVGIHPSAAEEFVTMRTVTRRVTSVGGKPKTNL

>Vc-GR-2

MRNLVGTGFLTQVELPISAMNVAKLTSGLASKRCLRVVAPARFGRGLSTRVAASASPG
GPGGDFDFDLFTIGAGSGGVRGSRFAASYGAKVAVCELPFDYVSSETRGGVGGTCVL
RGCVPKCLMVYASEYADDFRHAAGFGWSVPTEGLQHSWTSFIEAKRKELQRLNNAY
KNTLGNKAVTLLEGRGRVVDPHTEVDGRRYSANKNLIIVGGKPSKLDIPGAELCITS
DEALELPECPRKVAVLGGGYIAVEFSGIFARMGAEVHTVYRQPLPLRGFDGEVRKFAA
EQYAAAGLVVHPELTPQSISRQLDGRSLVVARSDGSTVLQDLQVLMATGRVPKT
DGLGLAEAGVELGPKGQVLVDEYCRVAVPSIWAAGDVIDRIQLTPVALMEGMAVAKSV
ALNQLTQPDYWAVPSAVFSNPEIATVGYTEEAASARFGDLDIYASFRPMPRNTVSGSPL
RTFMKIVVDAASQKVVGMHVMVGAEAAEIMQGFVAVKVGVTKAQLDSVVGIHPSA
AEEFVTMRSVSRKVRREVPANPVAPATAAA

>Vc-GR-1

MATLDDQYDLVTIGAGSGGVRASRFAATLYGAKVAVCELPFGFISSESVGGAGGTCVI
RGCVPKCLLVYGAFAEEFTDARGFGWALGGGAAAADGATGNGPVHDWSSLIKLN
KEIQRLNTTYGNILKNSGVTLIEGRGRILDATTVEVAAPDGAVRHLRAKNILLATGGVA
TKIPIEGAEHAIMSDNALALEALPPGPIAVLGAGYIATEFAGIFRGTHHYPVHLMFRGD
KVLRGFDTECRDQVQTNLVARGVHVHPCQPTRIEKKGEHQLVLHYRDSRGAGEQQ
LEVAMVMMATGRKPRVEGLGLEAVGVALDSSGAVQVDEYSRTSVPGIWAAGDVTNRI
NLTPVALMEGMAFAKSAFGGELTKPDYRNVASAVFCQPPLATVGYTEEQAVREFSGDI
DVYVSRFRPMKYTISGRDEKTLMKLVVHAESDKVLGCHMVGPDPEIMQGLAVALK
CGATKAQFDATVGIHPTAAEEFVTMRTTRTRTPVPGTGASKL

APÊNDICE A - SEQUÊNCIAS DOS DOMÍNIOS DAS PROTEÍNAS OBTIDAS PELA PLATAFORMA PFAM.

Gene	Quantidade (aa)	Sequência do domínio	Pfam domínio
<i>Gm-APX-1a</i>	175	KETLFTAIGKGNPDLIPSLTLALNDALTYDKATKSGGPNGSIRFSSEIGRPENTGLSAALNLLLEEAKKEI DSDSKGGSISYADLIHIAAQAQSAKATFLAAAIRKCGGKEEKGKTLYNAYGSNGQWGLFDRQFGRAD AQDPDPEGRVPLWEKASVQEMKDKFVAVGFGPRQLAVLS	Peroxidase (PF00141)
<i>Gm-APX-1b</i>	175	KETLFTAIGKGNPDLIPSLTLALNDALTYDKATKSGGPNGSIRFSSEIGRPENTGLSAAMNLLLEEAKKE IDSYSKGGPISYADLIHFAAQAQSAKATFLAAAIRKCGGNEEKGKTLYNAYGSNGQWGLFDRQFGRA DTQDPDPEGRVPIWEKASVQEMKDKFVAVGFGPRQLAVLS	Peroxidase (PF00141)
<i>Gm-APX-2</i>	202	EEVRKVLKSKGAAGVLRVLFHDAGTFDIDDSTGGMNGSIVYELERPENAGLKKSVKVLQKAKTQI DAIQPVSWADMIAVAGAEAVEVCGGPPIQVSLGRLDTLVPDPEGRLPPEESLNASGLKKCFQSKGFST QELVALSGAHTIGSKGFGSPISFDNSYYKVLLEKPWTSSGGMPSMIGLPSDHALVEDDECLRWIKKY ADS	Peroxidase (PF00141)
<i>Gm-APX-3a</i>	232	REDIKELLNSKFCHPILIRLWHDAGTYNKNIEEWPLRGGANGSLRFEVELKHGANAGLLNALKLL QPIKDKYSGVTYADLFQLASATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLRQ VFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGPAPGGQSWTVQWLKFDNSYFK DIKEKRDEDLVLPTDAALFEDPSFKVYAEKYAED	Peroxidase (PF00141)
<i>Gm-APX-3b</i>	232	REDIKELLNSKFCHPILIRLWHDAGTYNKNIEEWQRGGANGSLRFEIELKHGANAGLVNALKLL QPIKDKYSGVTYADLFQLAGATAVEEAGGPKIPMKYGRVDVSGPEQCPEEGRLPDAGPPSPADHLRQ VFYRMGLNDKEIVALSGAHTLGRSRPDRSGWGKPKETKYTKDGPAPGGQSWTVQWLKFDN SYFKDIKEKKDEDLVLPTDAALFEDPSFKVYAEKYAED	Peroxidase (PF00141)
<i>Gm-APX-4a</i>	208	KKKLRGFIAEKRCAPMLRLAWHSAGTFDKGKTGGPFGTIKHPAELAHSANGLDIAVRLLLEPLK AEFPILSYADFYQLAGVVAVEVTGGPEVPPFHPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLT DQDIVALSGGHTIGAANKERSGFEGPWTSNPLIFDNSYFTELLSKEKEGLLQLPSDKALLSDPV FRPLVDKYAAD	Peroxidase (PF00141)
<i>Gm-APX-4b</i>	208	KKKLRGFIAEKRCAPMLRLAWHSAGTYDVSSKTGGPFGTIKHPSELAHGANNGLDIAVRLLLEPLK AEFPILSYADFYQLAGVVAVEVTGGPEVPPFHPGREDKPEPPPEGRLPDATKGS DHLRDVFGKAMGLS DRDIVALSGGHTIGAANKERSGFEGPWTSNPLIFDNSYFKELLSKEKEGLLQLPSDKALLSDPVFRPL VEKYASD	Peroxidase (PF00141)

APÊNDICE A – CONTINUAÇÃO...

Gene	Quantidade (aa)	Sequência do domínio	Pfam domínio
<i>Gm-APX-5a1</i>	205	RELRAFITSNQCAPLMLRLAWNDAATYDARNRAGGPNGSIRTDKELKHEANEGLLKATQLCEHV KAKLKKVSYADLYQLAGVVVAIEVSGGPTIEFLPGRKDSMESSAEGLLPDVKQGASIIRNIFSRMGIS DDKHIVALCGGLTWGETLKDRSDSKGQWPKDPLKFDNSYYKKILSKDLSSRLPIEDALLTDQSFRR HVEEYSKD	Peroxidase (PF00141)
<i>Gm-APX-5a2</i>	193	ITNNKCAPLMLQFAWNAATYDAKSRRGGPNSIRLRIGQELKHEANKGLEKAVQYCEIVKTKLK LKKVSYANLYQDSNESPRTEGRFIDGEEDARNLRKIFSRMGLSDEQDIVALCGGHTLIRTMYPKVS MGETHKDRSKFEEGKSTNKPLKFDNSYFKELLIKDASFSRLPMDYALVEDPKFHYYVERYDE	Peroxidase (PF00141)
<i>Gm-APX-5b1</i>	207	RRDLRALIANRNCAPLMLRLAWHDAGTYDAKTKTGPPNGSIRNEEEYSHGANGLKKAIDFCQE VKAKYPKITYADLFQLAGVVAVEVTGGPTIDFVPGRRDSKVSPNEGRLPDAKKGVPHLRDIFYRM GLTDRDIVALSGGHTLGRAHPERSGFDGPWTEPLKFDNSYFVELLKEDSAGLLKLPTDKALLED AEFRCYVELYAKD	Peroxidase (PF00141)
<i>Gm-APX-5b2</i>	207	RRDLRALISNRNCAPLMLRLAWHDAGTYDAKTNTGGPNSIRNRQELNHAANKGLETALAFCEE VKAKHPKISYADLYQLAGVVAVEVTGGPTINFVPGRKDSLESPAEGRLPDAKQGASHLRDIFYRM GLGDKDIVALSGGHTLGKAHKDRSDFHGQWTKDPLKFDNSYFVELLRGESKDLLKLPTDKALVE DPNFRKYVELYAKD	Peroxidase (PF00141)
<i>Gm-MDHAR-1</i>	315	EYVIVGGGNAAGYAARTFVEHGMADGRLCIVTKEPYAPYERPALKAYLFPPDKKPARLPGFHTC VGSGGERQTPGWYKENGIEMLYEDPVKDIDIEKHTLTSSGKLLKYGSLVIATGCTASRFPEKIGGN LPGVHYIRDVDADALILSLEKAKKVVVVGGGYIGMEVAAAAGWKLDTTIIFPEDHLLQRLFTP SLARRYEELYQKNGVKILKGASIKNLEAGSNGHVA AVKLDGSLVEADTVIIGIGAKPAVTPFERVG LNTEVGGIQVDGQFRTSTPGIFAVGDVAAFPLKIYNRISRVEHVDHARRSA	Oxidoreductase de piridina nucleotídeo dissulfeto (PF07992)
<i>Gm-MDHAR-2a</i>	315	YVILGGGVAAGYAALEFVKKGVSHGELCIISDEPVAPYERPALSKGFLLEAAARLPSFHTCVGAN EERLTPKWKYKEHGIELVLGTGVKSTVDKRTLLTTTGETISYKILIVATGARALKLEEFVSGSDAE NVCYLRIADANRLVDVMQSCPGGNNAVIGGGYIGMECAASLVINKINVTMVFPEEHCMARLFTS KIANYEYEEYKSRGVNFIKGTVLSSFDSDSNGKVTAVNLRDGTTLSDVMVVGIGIRPNTGLFEG QLTLEKGGIKVNGMLQSSNSSVYAIGDVAAFPVKAFGETRRLEHVDSARKSA	Oxidoreductase de piridina nucleotídeo dissulfeto (PF07992)

APÊNDICE A – CONTINUAÇÃO...

Gene	Quantidade (aa)	Sequência do domínio	Pfam domínio
<i>Gm-MDHAR-2b</i>	315	YVILGGGVAAGYAALFVKKGVSHGELCIISDEPVAPYERPALSKGFLLPEAAARLPSFHTCVGAN EERLTSKWYKEHGIELVLGTGVKSADVKRKTLLTTTGETISYKILIVATGARALKLEEFVGTGSDA GNVCYLRDIADANRLVDVMQSCPGGNVAVVIGGGYIGMECAASLVINKINVTMVFPPEEHCGRFLF TSKIANYYEEYKSRGVNFIKGTVLSSFDSDSNGKVTAVNLRDGSTLSVDMVVVVGIGIRPNTGLFE GQLTLEKGGIKVNGMLQSSNSSVYAIGDVAAPVAVGETRRLEHVDSARKSA	Oxidoreductase de piridina nucleotídeo dissulfeto (PF07992)
<i>Gm-MDHAR-3a</i>	316	KYIILGGGVSAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESPARLPGFHVCVSG GERLLPEWYTEKGIELILSTEIVKVDLAAKSLISAGGETFSYQILIVATGSTVIRLDFGVEGADAKNI FYLREVDDADKLYAAIKAKKNGKAVVGGGYIGLELSAVLKLNNIDVTMVYPEPWCMPLFTAG IAEFYEGYYANKGVNIIKGTAVGFTSNSDGEVKEVKLKDGRVLEADIVVVGVGGRPQTVLVKGQ VEEEKGGIKTDAFFKTNLSDVYAVGDVATFPLKLYGELRRVEHVDHSRKA	Oxidoreductase de piridina nucleotídeo dissulfeto (PF07992)
<i>Gm-MDHAR-3b</i>	316	KYIILGGGVSAGYAAREFAKQGVKPGELAIISKEAVAPYERPALSKAYLFPESPARLPGFHVCVSG GERLLPEWYTEKGIELILSTEIVKVDLAAKSLISAGGETFSYQILIVATGSTVIRLDFGVEGADAKNI FYLREVDDADKLYEAIKAKKNGKAVVGGGYIGLELSAVLKLNNIDVTMVYPEPWCMPLFTAG IAEFYEEYYKNKGVNIIKGTAVGFTSNSDGEVKEVKLKDGRVLEADIVVVGVGGRPQTALVKGQ VEEEKGGIKTDAFFKTNLSDVYAVGDVATFPLKLYGELRRVEHVDHSRKA	Oxidoreductase de piridina nucleotídeo dissulfeto (PF07992)
<i>Gm-DHAR-1a</i>	69	DCPFSQRVLLTLEEKKIPYKLHLIDLNSKPEWFLGVNPEGKVPVVLFDGKQWVADSDVIVGILEEKY PP	Glutathione S-transferase, N-terminal (PF13417)
	63	EQALVAELSALDEHLKTHGPIYAGEKVTAVDLSLAPKLYHLVVALGHFKNWNIPESLVHVHNY	Glutathione S-transferase, C-terminal (PF13410)
<i>Gm-DHAR-1b</i>	69	DCPFSQRVLLTLEEKKIPYKLHLIDLSSKPEWFLGVNPEGKVPVVLFDGKQWVADSDVIVGILEEKY PEP	Glutathione S-transferase, N-terminal (PF13417)
	63	EQTLVAELSALDEHLKAHGPIYAGEKVTAVDLSLAPKLYHLVVALSHFKNWNIPESLTHVHNY	Glutathione S-transferase, C-terminal (PF13410)
<i>Gm-DHAR-2a</i>	70	DCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDEKWVPDSVITQTLEEK YPSPP	Glutathione S-transferase, N-terminal
<i>Gm-DHAR-2b</i>	70	DCPFCQRVLLTLEEKHLPYDPKLVDLTNKPEWFLKVNPDGKVPVIKFDEKWVPDSIITQTLEEKY PSPP	Glutathione S-transferase, N-terminal

APÊNDICE A – CONTINUAÇÃO...

Gene	Quantidade (aa)	Sequência do domínio	Pfam domínio
<i>Gm-GR-1a</i>	327	DLFIIGAGSGGVRAARFSSNFGAKV GICELPFHPISSETIGGVGGTCVIRGCVPKKILVY GASFGGDLEDARNYGWELSEK VDFNWKLLQKKTDEINRLNGLYKRMLSNWGVKL FEGEGKIVGPNEVEVTQLDGTKLSY TAKHILIATGSRAQRPNIPGQELGITSDEALSLED LPKHAVVLGGGYIAVEFASIWRGMGSTVDLVFRKELPLRGFDDEMRAAVARNLEGRGI NLHPRTNLTQLIKTEDGIK VITDHGEELIADAVLFATGRAPNSKRLNLETVGVELDNTG AIKVDEYSRTSIPSIWAVGDVTNRMNLT PVALMEA	Oxidoreutase da piridina nucleotídeo-dissulfeto (PF07992)
	110	IPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSTFNPMKNTISGRQEKTVMKLVVDAET DKVLGASMCGPDAPEIMQGI AVALKCGATKAQFDSTVGIHPSAAEEFVTM	Piridina nucleotídeo-dissulfeto, domínio de dimerização (PF02852)
<i>Gm-GR-1b</i>	327	DLFIIGAGSGGVRAARFSSNFGAKV GICELPFHPISSETIGGVGGTCVIRGCVPKKILVY GASFGGDLQDARNYGWELSEK VDFNWKRLQKKTDEINRLNGLYKRMLSNWGVKL FEGEGKIVGPNEVEVTQLDGTKLSY SAKHILIATGGRAQRPNIPGQELGITSDEALSLE DLPKRAVVLGGGYIAVEFASIWRGMGSTVDLVFRKELPLRGFDDEMRAV VARNLEGR GINLHPRTNLTQLIKTEDGIK VITDHGEELIADVLFATGRAPNSKRLNLEAVGVELDN TGAIKVDEYSCTSIPSIWAVGDVTNRMNLT PVALMEA	Oxidoreutase da piridina nucleotídeo-dissulfeto (PF07992)
	110	IPYAVFSIPPLSVVGLSEEQAIEQTNGDLLVFTSIFNPMKNTISGRQEKTVMKLVVDAET DKVLGASMCGPDAPEIIQGI AIALQCGATKAQFDSTVGIHPSAAEEFVTM	Piridina nucleotídeo-dissulfeto, domínio de dimerização (PF02852)
<i>Gm-GR-2a</i>	323	DLFTIGAGSGGVRAARFAANYGASVAICELPFSTISSETTGGVGGTCVIRGCVPKKLLV YASKFSHEFEESNGFGWRYDSEPKHDWSSLIANKNAELQRLTG IYKNILNAGVKLIE GHGKIIDPHTVDVNGKLYSAKHILVTVGGRRPFIPDIPGNEYAIDSDAALDLPTKPEKIAI VGGGYIALEFAGIFNGLKSEVHV FIRQKKVLRGFDEEIRDFVSEQMSVRGIEFHTEESP QAITKSADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNLGLESVGVKIAKDGAIEVDE YSQTSVPSIWAVGDVTNRINLT PVALMEG	Oxidoreutase da piridina nucleotídeo-dissulfeto (PF07992)
	109	VPSAVFSQPPIGQVGLTEEQAVQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKT NEVLGLHMCGDDAPEIVQGFAVALKARLTKADFDATVGIHPSAAEEFVTM	Piridina nucleotídeo-dissulfeto, domínio de dimerização (PF02852)

APÊNDICE A – CONTINUAÇÃO...

Gene	Quantidade (aa)	Sequência do domínio	Pfam domínio
<i>Gm-GR-2b</i>	323	DLFTIGAGSGGVRAARFAANYGASVAICELPFSTISSETTGGVGGTCVIRGCVPKLLVY ASKFSHEFEESNGFGWRYDSEPKHDWSSFIANKNAELQRLTGIYKNILNNAGVKLIEGH GKIIDPHTVDVNGKLYSAKHILVAVGGRPFIPDIPGKEYAIDSDAALDLPTKPKVIAIVGG GYIALEFAGIFNGLKSEVHV FIRQKKVLRGFDEEIRDFVEEQMSVRGIEFHTEESPQAITK SADGSFSLKTNKGTVDGFSHIMFATGRRPNTQNLGLESVGVKLAKDGAIEVDEYSQTSV SSIWAVGDVTNRINLTPVALMEG	Oxidoreutase da piridina nucleotídeo-dissulfeto (PF07992)
	109	VPSAVFSQPPIGQVGLTEEQAVQQYGDIDIFTANFRPLKATLSGLPDRVFMKLVVCAKTN EVLGLHMCGEDAPEIVQGFAVALKARLTKADFDATVGIHPSAAEEFVTM	Piridina nucleotídeo-dissulfeto, domínio de dimerização (PF02852)

Fonte: Elaborado pelo autor