



**UNIVERSIDADE DE RIBEIRÃO PRETO-UNAERP PROGRAMA DE
PÓS-GRADUAÇÃO EM BIOTECNOLOGIA**

**Identificação de proteínas presentes no soro do látex da *Himatanthus
sucuuba* (Apocynaceae)**

**RIBEIRÃO PRETO - SP
2017**

JANETH SILVA PINHEIRO MARCIANO

**Identificação de proteínas presentes no soro do látex da *Himatanthus sucuba*
(Apocynaceae)**

Tese submetida à Coordenação do Programa de Pós-Graduação em Biotecnologia da Universidade de Ribeirão Preto, como requisito para obtenção do grau de Doutor em Biotecnologia.

Orientadora: Profa. Dra. Sônia Marli Zingaretti

**RIBEIRÃO PRETO - SP
2017**

Ficha catalográfica preparada pelo Centro de
Processamento Técnico da Biblioteca Central da UNAERP

- Universidade de Ribeirão Preto -

M319f Marciano, Janeth Silva Pinheiro, 1963-
Identificação de proteínas presentes no soro do látex da
Heliconia lutea (Apocynaceae) / Janeth Silva Pinheiro
Marciano. - - Ribeirão Preto, 2017.
132 f.: il. color.

Orientadora: Prof.^a Dr.^a Sônia Marli Zingaretti.

Tese (doutorado) - Universidade de Ribeirão Preto, UNAERP,
Biotecnologia. Ribeirão Preto, 2017.

1. Biotecnologia. 2. Eletroforese unidimensional. 3. Espectrometria
de massa. 4. Proteínas laticíferas. I. Título.

CDD 660

JANETH SILVA PINHEIRO MARCIANO

IDENTIFICAÇÃO DE PROTEÍNAS PRESENTES NO LÁTEX DA *HIMANTHUS*
SUCUUBA (APOCYNACEAE)

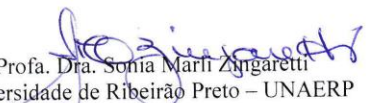
Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biotecnologia da Universidade de Ribeirão Preto, para obtenção do título de Doutor em Biotecnologia.

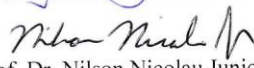
Área de Concentração: Biotecnologia Aplicada à Saúde

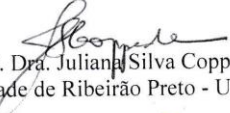
Data da defesa: 22 de MAIO de 2017

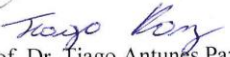
Resultado: Aprovada

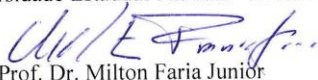
BANCA EXAMINADORA


Prof. Dra. Sonia Marli Zingaretti
Universidade de Ribeirão Preto – UNAERP


Prof. Dr. Nilson Nicolau Junior
Universidade Federal de Uberlândia - UFU


Prof. Dra. Juliana Silva Coppede
Universidade de Ribeirão Preto - UNAERP


Prof. Dr. Tiago Antunes Paz
Universidade Estadual Paulista - UNESP


Prof. Dr. Milton Faria Junior
Universidade de Ribeirão Preto - UNAERP

RIBEIRÃO PRETO
2017

Ao meu marido Ivan Marciano, por apoiar as minhas decisões, compreender a minha ausência, pelo seu companheirismo e amor, DEDICO.

A minha irmã Jacinete Pinheiro (*In Memoriam*), amor eterno, DEDICO.

AGRADECIMENTOS

A Deus que nos direciona em todos os momentos de nossas vidas.

À orientadora Profa. Dra. Sônia Marli Zingarett que me acolheu como orientanda, pela sua experiência científica e pelos seus ensinamentos.

Aos funcionários dos Laboratórios da Biotecnologia da UNAERP especialmente a Simone Cristina Zanpolo Torres.

A todos os Professores do Programa de Doutorado, pelos conhecimentos a nós transmitidos. Aos meus colegas do laboratório Matheus, Joice, Luana e Renan Silva pela amizade, apoio e cooperação.

Ao meu pai Juarez da Silva Pinheiro (*In Memoriam*) e a minha mãe Neide da Silva Pinheiro (*In Memoriam*) que muito contribuíram para formação dos alicerces de minha educação.

Aos meus irmãos que de alguma forma, sempre me ajudaram.

Ao Laboratório de Química de Proteínas da USP-Ribeirão Preto.

Ao Departamento de Química da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto – USP. Especificamente no Laboratório de Espectrometria de Massas Aplicada a Química de Produtos Naturais.

Ao professor Leo Kei Iwai do Instituto Butantan.

MUITO OBRIGADO!

O tempo pode ser cultivado como um amigo e um colaborador, pois ele é a inspiração de tudo que existe. O tempo permite que as coisas aconteçam; ele é o fluxo dos eventos, o desdobrar da experiência. O tempo nos dá as oportunidades preciosas de viver, de nos desenvolver, crescer e também de apreciar nossa natureza interior. Mesmo sabendo que o nosso tempo por fim se esgotará, que a vida terminará e que as nossas oportunidades se acabarão, ainda assim terá sido o tempo que permitiu o desenrolar de nossas vidas.

(Tarthang Tulku)

RESUMO

Himatanthus sucuuba é uma planta laticífera que pertence à família Apocynaceae, nativa da região Amazônica. É usada tradicionalmente na medicina popular como: vermífugo, tratamento de gastrite, artrite, laxativo, antianêmico, antifúngico e cicatrizante. Pouco se conhece sobre o proteoma do látex de *H. sucuuba* e o objetivo desse trabalho foi a identificação de proteínas presentes no soro do látex, através das técnicas de cromatografia, eletroforese bidimensional (2D), eletroforese (1D) SDS-PAGE e Espectrometria de massas, utilizando proteínas totais extraídas do soro do látex de *H. sucuuba* (HsLP). Os resultados foram analisados pelo MASCOT (integrado ao banco de dados NCBI nr/ Swiss-Prot), Blast2GO, em conjunto com a base de dados do KEGG. Foram identificadas no látex de *H. sucuuba* 1.471 proteínas das quais 587 proteínas apresentando score igual ou superior a 20. Dentre essas, 587 previamente identificadas, foram encontradas: 267 proteínas preditas (45%), 158 hipotéticas (27%), 128 caracterizadas (parciais ou inteiras) (22%), 24 putativas (4%) e 10 proteínas não definidas (2%), com massa molecular variando entre 13,33 e 164,65 KDa. Dentre as proteínas encontradas 32,64% são básicas ($pI > 8$), 24,82% são neutras ($6 \geq pI \geq 8$) e 60 proteínas que correspondem a mais de 40% do total são caracterizadas como ácidas ($pI < 6$). As enzimas mais frequentes encontradas pertencem às classes transferases, seguida das hidrolases, oxidoreductases, liases e ligases. Não foram identificadas enzimas do tipo isomerases. As enzimas identificadas no soro do látex de *H. sucuuba* resultaram na predição de 30 vias metabólicas. Entre elas, foi identificada uma enzima *NADH: ubiquinone reductase*, que compõe um dos principais complexos proteicos da fosforilação oxidativa nas mitocôndrias. Dentre outras enzimas, algumas da via de biossíntese de alcalóides isoquinolínicos foram identificadas. As enzimas *tetrahydroberberine oxidase*; *(S)-THB oxidase* e *(RS)-norcoclaurine 6-O-methyltransferase* catalisam reações que produzem produtos de interesse farmacológico, como os alcalóides berberine e palmatine, que são tóxicos para insetos e vertebrados e inibem a multiplicação de bactérias, fungos e vírus. A berberine tem várias atividades farmacológicas, incluindo propriedades antimicrobianas, de redução da glicose e do colesterol, antitumorais e imunomoduladoras. Além disso, identificou-se no látex de *H. sucuuba* o composto lupeol que apresenta atividade anti-inflamatória. A enzima óxido nítrico sintase que catalisa a produção de óxido nítrico uma importante molécula de sinalização celular, que desempenha funções em uma variedade de processos celulares como vasodilatação, neurotransmissão, agregação plaquetária, e está envolvido na angiogênese. The results obtained in this study provided important information about the proteins identified in the serum of *H. sucuuba* latex that make possible future research.

Palavras-chave: Eletroforese unidimensional, Espectrometria de massas, *Himatanthus sucuuba*, Proteínas laticíferas.

ABSTRACT

Himatanthus sucuuba is a latexiferous plant that belongs to the family Apocynaceae, native to the Amazon region. It is traditionally used in popular medicine as: vermifuge, in the fight against gastritis, arthritis, laxative, antianemic, antifungal and cicatrizante. Little is known about the latex protease of *H. sucuuba* and the objective of this work was to identify proteins present in the serum of the latex through the techniques of chromatography, two-dimensional electrophoresis (2D), electrophoresis (1D) SDS-PAGE and spectroscopy mass. For the proteomic analysis, total proteins extracted from the serum of *H. sucuuba* latex (HsLP) were used by electrophoresis (1D) SDS-PAGE and mass spectrometry. The result obtained from mass spectrometry was analyzed by MASCOT (integrated with the NCBI Inr / Swiss-Prot database). In the latex of *H. sucuuba*, 1.471 proteins were identified, of which 587 proteins had a score equal to or higher than 20. Of these, 587 previously identified, 267 predicted proteins (45%), 158 hypothetical proteins (27%), 128 (22%), 24 putative (4%) and 10 undefined proteins (2%), with molecular mass ranging from 13.33 to 164,65 KDa. Among the proteins found 32,64% are basic ($pI > 8$), 24,82% are neutral ($6 \leq pI \leq 8$) and 60 proteins, corresponding to more than 40% of the total are characterized as acidic ($pI < 6$). The most frequent enzymes that were found belong to the class transferases, followed by hydrolases, oxidoreductases, lyases and ligases. No isomerase-like enzymes have been identified. The enzymes identified in the serum of *H. sucuuba* latex by Blast2G0, together with the KEGG database, resulted in the prediction of a total of 30 metabolic pathways. In the metabolic pathways, an NADH: ubiquinone reductase enzyme, which makes up one of the major protein complexes of oxidative phosphorylation in the mitochondria, has been identified. Oxidative phosphorylation is the most important part of the metabolic pathways, which produces adenosine triphosphate (ATP) using energy, and It also produces reactive oxygen species, such as hydrogen peroxide and superoxide. Among these, the biosynthesis pathway of isoquinolinic alkaloids was an important pathway described. The enzymes identified in this pathway (tetrahydroberberine oxidase; (S) - THB oxidase and (RS) -norcoclaurine 6-O-methyltransferase) catalyze reactions and produce products of pharmacological interest, such as the alkaloids berberine and palmatine, are toxic to insects and vertebrates And inhibit the multiplication of bacteria, fungi and viruses. Berberine has several pharmacological activities, including antimicrobial, glucose and cholesterol lowering, antitumor and immunomodulatory properties. In addition, we identified in the latex of *H. sucuuba* the compound lupeol that shows anti-inflammatory activity. The enzyme nitric oxide synthase that catalyzes the production of nitric oxide is an important cell signaling molecule, which functions in a variety of cellular processes such as vasodilation, neurotransmission, platelet aggregation, and is involved in angiogenesis. The results obtained in this study provided important information about the proteins identified in the serum of *H. sucuuba* latex that allows for future research.

Key words: *Himatanthus sucuuba*, Latex proteins, Mass spectrometry, One-dimensional electrophoresis.

LISTAS DE FIGURAS

Figura 1	Anatomia de células laticíferas em oito espécies de plantas.	18
Figura 2	Atividade angiogênica na MCA.	19
Figura 3	Esquema simplificado da matéria prima para a produção de gel cicatrizante	20
Figura 4	REGEDERM® produto que possui atividade angiogênica e celeradora do processo cicatricial e é indicado para o tratamento e cicatrização de úlceras cutâneas de diversas etiologias.	20
Figura 5	Árvore <i>Himatanthus sucuuba</i> . No detalhe o tronco e casca rugosa da planta.	24
Figura 6	Partes da planta <i>H. sucuuba</i> (A fruto - B; semente- C; folha e flor - D)	24
Figura 7	Perfil eletroforético de amostra do soro do látex de <i>H. sucuuba</i> (15 µg) em PAGE-SDS	39
Figura 8	Cromatograma obtido da separação das proteínas do soro do látex de <i>H. sucuuba</i> , utilizando a coluna de troca catiônica HiTrap SP XL.	40
Figura 9	Cromatograma obtido da separação das proteínas do soro do látex de <i>H. sucuuba</i> utilizando a coluna de troca aniônica HiTrap QXL.	40
Figura 10	Fotos dos géis de segunda dimensão com amostras resultantes dos três métodos de extração testados. A- Extração de sulfato de amônio; B- Extração com TCA/Acetona e C - Extração com TCA/Acetona.	42
Figura 11	Perfil eletroforético em PAGE-SDS 12 % corado por Coomassie G-250, das Proteínas totais extraídas do soro do látex de <i>H. sucuuba</i> .	43
Figura 12	Proteínas potencialmente identificadas no látex de <i>Himatanthus sucuuba</i>	44
Figura 13	Classificação das proteínas caracterizadas de acordo com o organismo identificado.	55
Figura 14	Gráfico de dispersão para a relação de Score e pI das proteínas obtidas pela espectrometria de massas e identificadas pelo programa Mascot.	56
Figura 15	Anotação funcional para Processo Biológico das proteínas identificadas em <i>H. sucuuba</i> .	58
Figura 16	Anotação funcional para função molecular das proteínas identificadas em <i>H. sucuuba</i> .	59
Figura 17	Anotação funcional para componente celular das proteínas identificadas em <i>H. sucuuba</i> .	60
Figura 18	Classificação geral das enzimas encontradas no látex de <i>H. sucuuba</i> . As enzimas foram categorizadas utilizando-se o programa Blast2GO, de acordo com o banco de dados <i>Kyoto Encyclopedia of genes and Genomes pathway database</i> disponíveis em KEGG.	61
Figura 19	Esquema metabólico da reação de fosforilação oxidativa evidenciando a participação da enzima <i>NADH: ubiquinone reductase</i> , uma proteína identificada no látex de <i>H. sucuuba</i> . A caixa colorida na cor rosa, indica o número EC 1.6.5.3, correspondente da enzima identificada.	71

Figura 20	Resumo de algumas das vias metabólicas e enzimas preditas pelo Blast2GO, no soro do látex de <i>H. succuba</i> .	72
Figura 21	Reações catalisadas pela enzima EC: 1.3.3.8	74
Figura 22	Via metabólica dos alcalóides isoquinolínicos identificada em <i>H. succuba</i> . Enzima destacada em verde é a <i>(RS)-norcoclaurine 6-O-methyltransferase</i> (EC:2.1.1.128) em azul <i>(S)-tetrahydroberberine:oxygen oxidoreductase</i> (EC: 1.3.3.8), com seus respectivos números de EC.	75
Figura 23	Espectro de ESI-MS de ions totais da fração aquosa (modo positivo).	76
Figura 24	Espectro de ESI-MS/MS do ion m/z 409,4 (modo positivo).	77
Figura 25	Estrutura química do lupeol.	77
Figura 26	Espectro de ESI-MS de ions totais da fração acetato (modo positivo).	79
Figura 27	Espectro de ESI-MS de ions totais do látex bruto (modo positivo).	79
Figura 28	Espectro de ESI-MS/MS do ion m/z 409,4 (fração acetato; modo positivo).	80
Figura 29	Espectro de ESI-MS/MS do ion m/z 409,4 (látex bruto; modo positivo).	80

LISTA DE TABELAS

Tabela 1	Proteínas de látex identificadas em diferentes plantas por Espectrometria de massas .	23
Tabela 2	Lista de peptídeos e proteínas identificadas em látex de <i>Himatanthus sucuuba</i> por ESI-QUAD-TOF.	45
Tabela 3	Subclasses das enzimas oxiredutases identificadas no soro do látex de <i>H. sucuuba</i> . Categorizadas de acordo com os seus números de EC (<i>Enzyme comission</i>), com base nos dados obtidos no <i>Encyclopedia of genes and Genomes pathway database</i> (KEGG) pelo software Blast2GO.	62
Tabela 4	Subclasses das enzimas transferases identificadas no soro do látex de <i>H. sucuuba</i> . Categorizadas de acordo com os seus números de EC (<i>Enzyme comission</i>), com base nos dados obtidos no <i>Encyclopedia of genes and Genomes pathway database</i> (KEGG) pelo software Blast2GO.	63
Tabela 5	Subclasses das enzimas hidrolases identificadas no soro do látex de <i>H. sucuuba</i> . Categorizadas de acordo com os seus números de EC (<i>Enzyme comission</i>), com base nos dados obtidos no <i>Encyclopedia of genes and Genomes pathway database</i> (KEGG) pelo software Blast2GO.	64
Tabela 6	Subclasses das enzimas liases identificadas no soro do látex de <i>H. sucuuba</i> . Categorizadas de acordo com os seus números de EC (<i>Enzyme comission</i>), com base nos dados obtidos no <i>Encyclopedia of genes and Genomes pathway database</i> (KEGG) pelo software Blast2GO.	64
Tabela 7	Vias metabólicas das enzimas encontradas no soro do látex de <i>H. sucuuba</i> , preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (<i>Enzymes Commission</i>).	66

LISTA DE ABREVIATURAS E SIGLAS

1D	unidimensional
2D	bidimensional
ATP	trifosfato de adenosina
B2G	Blast2GO
CoA	coenzima A
DTT	ditiotreitól
EC	Enzyme Commission
EDTA	ácido etilenodiamino tetracético (dissódico)
EM	Espectrometria de massas
MS	Mass spectrometry
ESI	ionização por <i>electrospray</i> .
GO	Gene Ontology
IUBMB	International Union of Biochemistry and Molecular Biology
KEGG	Kyoto Encyclopedia of Genes and Genomes
<i>m/z</i>	Relação massa-carga
mL	mililitros
MM	massa molecular
NCBI	National Center for Biotechnology Information non-redundant protein database
PAGE	eletroforese em gel de poliacrilamida (<i>polyacrylamide gel electrophoresis</i>).
ROS	espécies reativas de oxigênio
SDS	dodecil sulfato de sódio (<i>sodium dodecyl sulfate</i>)
TCA	ácido tricarbóxico
TOF	tempo de vôo (<i>time-of-flight</i>)
PAGE	Eletroforese em gel de poliacrilamida
$\mu\text{g/mL}$	Microgramas/Mililitros
$\mu\text{g}/\mu\text{L}$	Microgramas/microlitros
MCA	Membrana cório-alantóidea
HsLP	Proteínas do Látex de <i>Himatanthus sucuba</i>

SUMÁRIO

1	INTRODUÇÃO	16
2	REVISÃO DA LITERATURA	17
2.1	Plantas laticíferas	17
2.2	Proteínas identificadas em diferentes espécies de plantas	22
2.3	Espécie <i>Himatanthus sucuuba</i>	23
2.4	Espectrometria de massas de proteínas	26
3	JUSTIFICATIVA	28
4	HIPÓTESE	29
5	OBJETIVOS	30
5.1	Objetivo geral	30
5.2	Objetivos específicos	30
6	MATERIAL E MÉTODOS	31
6.1	A coleta de material biológico	31
6.2	Fracionamento e diálise	31
6.3	Extração de proteínas do látex	31
6.4	Quantificação de proteínas totais	33
6.5	Eletroforese gel de poliacrilamida SDS-PAGE	33
6.6	Cromatografia de Troca Iônica	33
6.7	Eletroforese Bidimensional	33
6.7.1	Focalização Isoelétrica	34
6.7.2	Eletroforese Bidimensional ou Segunda Dimensão	35
6.8	Espectrometria de massas	35
6.8.1	Eletroforese em gel de poliacrilamida 12% –SDS-PAGE	36
6.8.1.1	Redução e alquilação	36
6.8.1.2	Digestão das proteínas	37
6.9	Identificação metabólitos secundário presentes no látex da espécie <i>H. sucuuba</i> (Apocynaceae) por Espectrometria de massas	37
7	RESULTADOS E DISCUSSÃO	39
7.1	Quantificação de proteínas	39
7.2	Eletroforese –SDS-PAGE	39

7.3	Cromatografia de troca iônica	39
7.4	Eletroforese Bidimensional	41
7.5	Identificação de proteínas por análise MS/MS	42
7.6	Perfil de proteínas	55
7.7	Pontos isoeletricos	56
7.8	Anotação funcional das proteínas identificadas pelo programa Blast2GO	57
7.9	Classificação das enzimas	61
7.9.1	Oxirredutases	62
7.9.2	Transferases	62
7.9.3	Hidrolases	63
7.9.4	Liases	64
7.9.5	Ligases	65
7.10	Vias metabólicas	65
7.10.1	Funções de algumas das proteínas e enzimas identificadas no soro do látex de <i>Himatanthus sucuuba</i>	71
7.11	Identificação de metabolitos secundário	76
8	CONCLUSÃO	81
	REFERÊNCIAS BIBLIOGRÁFICAS	82
	ANEXO	93

1.INTRODUÇÃO

O Brasil possui uma flora rica em plantas utilizadas pela população com fins terapêuticos, uma das mais antigas formas da prática medicinal da humanidade. Em alguns países do Oriente, a utilização de plantas medicinais é a principal fonte da maioria dos tratamentos na forma de chás, emplastos, pomadas e tinturas. As plantas têm sido também empregadas na indústria farmacêutica como matéria prima para a extração de princípios ativos (compostos químicos) que são utilizados no desenvolvimento de vários medicamentos como: quinina, extraído da casca da *Cinchona L.*, Taxol (paclitaxel isolado de *Taxus brevifolia*), a codeína e a morfina (isoladas do látex da *Papaver somniferum L.*), os alcaloides emetina e cefaelina (ambas isoladas de *Cephalis ipecacuanha*), entre outros. Nas últimas décadas várias pesquisas têm sido feitas buscando a comprovação das ações farmacológicas de diversas plantas de uso frequente pelas populações locais.

As plantas laticíferas, particularmente, são utilizadas na medicina popular em função de suas propriedades curativas e analgésicas, no uso tópico em locais de ferimentos ou por ingestão do látex.

Aproximadamente 20.000 espécies de plantas que compreendem aproximadamente 10% de todas as angiospermas identificadas contém látex. O látex apresenta-se como um líquido leitoso composto por uma mistura complexa de moléculas, na maioria das plantas o látex é branco, podendo ser encontrado também nas tonalidades marrom, laranja ou vermelho. Desempenha funções que vão desde barreira física a toxicidade e defesa da planta contra patógenos. No momento em que a planta sofre injúria mecânica nas folhas, caules e frutos pela ação do homem ou por insetos, essas plantas secretam imediatamente o látex. A exsudação do látex prossegue por alguns minutos até a formação de um coágulo entorno da injúria na planta que pode aprisionar pequenos insetos e apresentar efeito tóxico a outros animais. A coagulação do látex é importante para a defesa da planta contra um possível invasor.

A aplicação tópica de plantas medicinais (látex, folhas, raízes, caule) tem sido utilizada, na medicina popular, na cicatrização de feridas, no entanto, o mecanismo exato do processo de cicatrização das feridas não é plenamente esclarecido. Coagulação e fibrinólise são dois eventos proteolíticos importantes na cicatrização de feridas, essa atividade proteolítica encontra-se presente no látex de muitas plantas medicinalmente importantes, como: *Synadenium grantii*, *Calotropis gigantea*, *Euphorbia cf.*, *Tabernaemontana divaricata* e outras. O interesse medicinal e farmacêutico em plantas baseia-se na presença de compostos importantes como:

alcalóides, flavonóides, ácidos fenólicos entre outros que ocorrem na seiva leitosa (látex) de algumas plantas.

Atualmente, o látex tem sido bastante estudado e vem ganhando destaque no estudo de biomoléculas na defesa contra invasores nas plantas e atividade ant inflamatória e cicatrizante. Graças aos avanços tecnológicos nas áreas de química, bioquímica, biotecnologia e farmacologia, a extração, purificação, testes terapêuticos e a elucidação das estruturas moleculares complexas de constituintes naturais tem contribuído para obtenção de novos bioativos de origem vegetal.

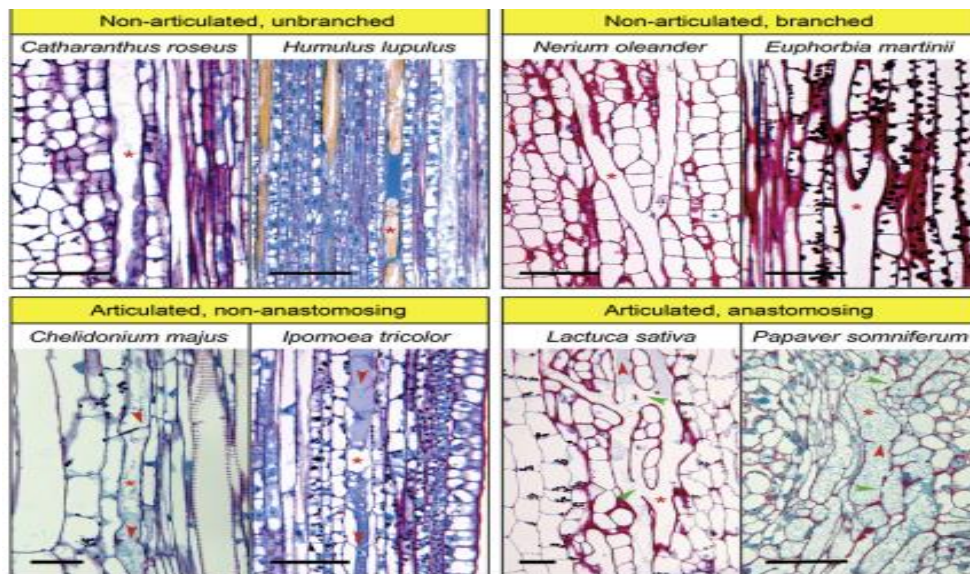
2. Revisão de Literatura

2.1 Plantas laticíferas

O látex é uma complexa emulsão composta por proteínas, alcalóides, açúcares, amido, óleos, taninos, resina e borracha, que muitas vezes acaba por coagular quando exposto ao ar (HAGEL et al., 2008). Estão presentes nos vasos laticíferos que são células secretórias alongadas do sistema laticífero (MAHLBERG, 1993; METCALFE, 1967). Estas células podem ser classificadas, de acordo com seus padrões de desenvolvimento e características morfológicas, como células únicas não-articulados ou uma série de células articuladas (Figura 1), contendo um leite branco, amarelo ou alaranjado, denominado de látex (PICKARD, 2008; MAHLBERG, 1993; HAGEL et al., 2008). Estudos fisiológicos mostraram que o látex está associado a metabólitos vitais na defesa das plantas contra insetos (KEKWICK, 2001).

A elucidação dos componentes ativos presentes nas plantas, bem como seus mecanismos de ação, apresenta-se como um dos maiores desafios para a química farmacêutica, bioquímica e farmacologia (GEBHARDT, 2000).

Figura 1. Anatomia de células laticíferas em oito espécies de plantas.

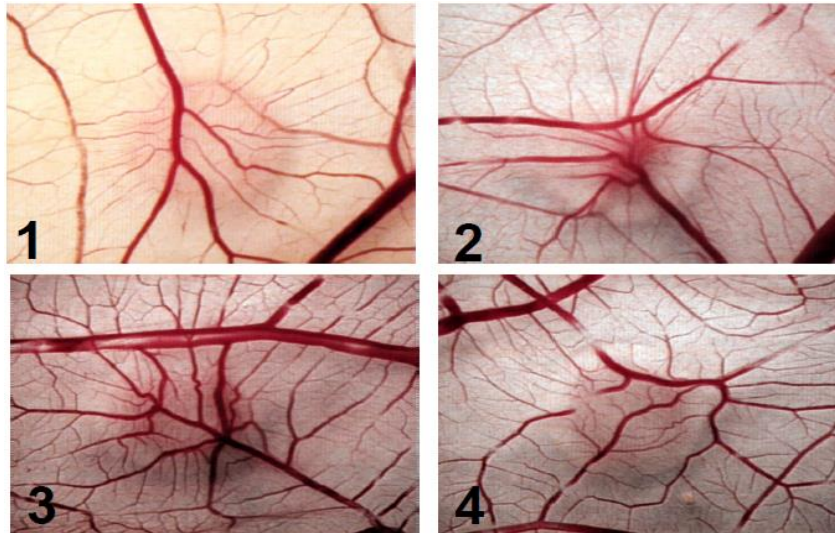


FONTE: HAGEL et al., 2008.

As plantas que mais produzem látex pertencem às famílias Euphorbiaceae e Apocynaceae (LYNN; CLEVETTE-RADFORD, 1987). A seringueira, *Hevea brasiliensis* (Euphorbiaceae), nativa da Região Amazônica, é a planta laticífera mais conhecida.

Mendonça et al. (2010), utilizaram uma fração sérica (FrHB1 0,3 mg; 1,5 mg de proteína) do soro do látex de *Hevea brasiliensis* em ensaios com membrana corioalantóica (MCA) de embrião de pinto (*Gallus gallus domesticus*) que mostrou ter efeito angiogênico e foi eficaz no aumento da permeabilidade vascular. Em outro teste com úlceras dérmicas, a fração do soro quando injetada acelerou significativamente a cicatrização das feridas, então foi sugerido que o látex natural de *H. brasiliensis* possui atividades de cicatrização de feridas e atividade angiogênica pela formação de novos vasos a partir de vasos preexistentes segundo (FOLKMAN; SHING, 1992; RAMJIWAN et al., 2017) (Figura 2). Em outro experimento utilizou-se uma fração proteica do soro do látex *H. brasiliensis* (FrHB) na confecção de uma biomembrana que foi usada como parte no tratamento para cicatrização de feridas no pé em pacientes diabéticos (ANDRADE et al., 2012), onde é citado amplamente na literatura devido os seus efeitos de cicatrização.

Figura 2. Atividade angiogênica na MCA. 1: Controle Negativo (água destilada); 2: FrHB 1 (0,3 mg de proteína); 3: FrHB 2 (1,5 mg de proteína); 4: FrHB 3 (1,5 mg de proteína). Em cada ensaio, 3,0 µL de uma solução (2,5 mg/mL) foram aplicados, utilizando-se de discos de papel de filtro (0,5 cm de diâmetro) como suporte.



FONTE: Mendonça et al., (2010).

As pesquisas com o látex da planta *Hevea brasiliensis* (seringueira), iniciadas pelo Professor Dr. Joaquim Coutinho Netto em 1993, já demonstravam inúmeras aplicabilidades na área médica, por sua ação indutora de angiogênese, semelhante ao fator de crescimento vascular endotelial, amplamente comprovada em animais de laboratório em 1997 (PELENOVA BIOTECNOLOGIA, 2017).

A empresa Pelenova Biotecnologia utiliza uma fração denominada proteína bioativa F1 do soro do látex da *H. brasiliensis* na regeneração tecidual, na formulação de cosméticos com propriedade anti-envelhecimento, géis (Figura 3), cremes, loções, na confecção de uma biomembrana (Biocure®) e REGEDERM® que possui atividade angiogênica e aceleradora do processo cicatricial e é indicado para o tratamento e cicatrização de feridas crônicas de diversas etiologias (Figura 4).

As feridas crônicas são as de origem venosa (por insuficiência de circulação das veias), as úlceras de pressão (feridas na região dorso-lateral dos quadris, lombo-sacral, dos glúteos e calcanhares) e as úlceras do pé diabético. Constituem-se em problema sério de saúde pública mesmo nos países desenvolvidos e acompanham-se com frequência de número importante de doenças e complicações (PELENOVA BIOTECNOLOGIA, 2017).

Figura 3. Esquema simplificado da matéria prima para a produção de gel cicatrizante.



FONTE: Pelenova biotecnologia, 2017.

Figura 4. REGEDERM® produto que possui atividade angiogênica e aceleradora do processo cicatricial, indicado para o tratamento e cicatrização de úlceras cutâneas de diversas etiologias.



FONTE: Pelenova Biotecnologia, 2017.

No látex de *Synadenium grantii* há uma grande quantidade de proteases, enzimas hidrolíticas que desempenham importante função na fisiologia das plantas e apresentam diversas ações farmacológicas, interferindo na hemostasia. Também foi relatado nesta espécie o isolamento e caracterização de uma protease (glicoproteína) que possui atividade fibrinolítica, com massa molecular de 34 KDa (RAJESH et al., 2006). Freitas et al., (2011) identificaram uma cisteína proteinase denominada osmotin presente no látex de *Calotropis procera*, possivelmente envolvida com o mecanismo de defesa da planta, contra fitopatógenos.

Em outro estudo as frações extraídas do látex seco de *Calotropis procera*, tanto a fração aquosa (AqDL) como a de extrato metanólico (MeDL) possuem propriedade anti-inflamatória e foram caracterizadas pelas suas propriedades bioquímicas, comparadas com a sua eficácia na

melhoria da febre em ratos e o seu mecanismo de ação foi elucidado. A administração oral das frações AqDL e MeDL produziu uma redução na temperatura corporal em ratos onde a febre foi induzida por levedura e o seu efeito foi comparável ao do medicamento padrão paracetamol. Estudos mostraram que tanto uma fração quanto a outra, são eficazes como potencial terapêutico no tratamento de várias condições febris (KUMAR et al., 2017; 1994).

Também foram isoladas do látex de algumas plantas proteínas com possível atividade cicatrizante entre os quais podemos destacar: proteinases no látex de *Carica candamasensis* utilizado em traumas dermatológicos e em aplicações clínicas (LEMOS et al., 2011; MELO et al., 2008); uma heme-peroxidase com propriedades anti-inflamatória e cicatrizante, identificada em *Artocarpus lakoocha*, que faz dela uma candidata com potencial em aplicações na indústria e biotecnologia (SONKAR et al., 2015), uma glicoproteína do látex de *Calotropis gigantea* com atividade fibrinogênica e de coagulação. Segundo Rajesh et al., (2005) o extrato de látex bruto contém muitas proteínas de natureza altamente básica e exibem uma forte atividade proteolítica. O látex de várias espécies de plantas tem demonstrado estar envolvido na homeostase (BOLAY, 1979; GUNTER et al., 2002), cicatrização de feridas e analgésica (THANKAMMA, 2003; DEWAN et al., 2000). *Jatropha curcas* Linn. (Euphorbiaceae), uma planta medicinal comumente cultivada nos trópicos, é tradicionalmente usada como hemostática (OSONIYI; FUNMI ONAJOBI, 2003).

A protease colagenolítica, extraída do látex de figo (*Ficus carica* var. Brown Turkey), é uma enzima com potencial para uso em várias aplicações bioquímicas, biotecnológicas e medicinais. Tem sido utilizada para o isolamento e cultivo de células de mamífero em cultura, tratamento de queimaduras, úlceras, feridas, elimina o tecido da cicatriz, também desempenha um papel importante no transplante de órgãos específicos, no tratamento da doença de Peyronie e vários tipos de fibrose destrutiva, tais como cirrose no fígado (RASKOVIK et al., 2014).

Ainda, segundo Alam et al., (2011), o látex extraído de diferentes espécies como de *Datura stramonium* L. (Solanaceae), *Euphorbia antiquorum* L. (Euphorbiaceae), *Euphorbia pilosa* (Euphorbiaceae), *Pelilanthus tithmaloids* (Euphorbiaceae), *Achyranthes aspera* L. (Amaranthaceae), *Alstonia scholaris* R.Br. (Apocinaceae), *Calotropis procera* Br (Asclepidaceae) apresenta propriedades curativas em feridas e queimaduras. O látex de *Ficus carica* é uma fonte rica de agentes antioxidantes naturais com características estabelecidas de atividade anti-carcinogênica que podem ser usadas em e Indústrias biofarmacêuticas (HASHEM et al., 2017). Em *Asclepias curassica* L., *Calotropis gigantea* R. Br., *Pergularia*

extensa R. Br. e *Cynanchum puciflorum* R. Br. que pertencem à família Asclepiadaceae foi identificada uma cisteína protease que exibe ação de trombina e plasmina validando o uso popular do látex dessas plantas para estancar sangramento e cicatrizar feridas (SHIVAPRASAD et al., 2009). *Euphorbia* cf. possui uma cisteína protease (EuP-82) com potencial para o uso no tratamento medicinal para trombose, essa enzima tem atividade fibrinogênica (SIRITAPETAWEI et al., 2015; SIRITAPETAWEI et al., 2015; SIRITAPETAWEI et al., 2017).

No látex da planta *Tabernaemontana divaricata* (L.) R. Br. foi identificada uma protease com potencial hemostático para uso tópico no tratamento de feridas (SINGH et al., 2015), enquanto no látex de Opium poppy (*Papaver somniferum*), o interesse medicinal baseia-se na síntese de compostos de interesse farmacêutico, referidos como alcalóides, com propriedades analgésica (morfina), antitussígeno (codeína) e relaxantes musculares (papaverina) (DECKER et al., 2000).

As proteínas solúveis presentes no látex de *Calotropis procera* (Apocynaceae), muito utilizado na medicina popular em infecções dermatológicas foram utilizadas para produção de uma membrana à base de álcool polivinílico quando testada em feridas induzidas na região dorsal de ratos observou-se que as proteínas do látex de *C. procera* atuaram significativamente no processo de cicatrização, comprovando a eficiência da membrana (FIGUEREDO et al., 2012). Segundo Singh et al., (2015), proteases do látex *Tabernaemontana divaricata* (L.) e *Artocarpus altilis* também possuem potencial hemostático.

2.2 Proteínas identificadas em diferentes espécies de plantas

Nos bancos de dados, verificou se que entre as proteínas provenientes de látex estão incluídas 186 proteínas de *H. brasiliensis* (com 172 proteínas funcionalmente anotadas), como a ‘REF, putative’ (gi/223548388), ‘SRPP’ (gi | 37622210) e ‘Cyclophilin A’ (gi | 167599641) (DAI et al, 2013). Em *Papaver somniferum* foram identificadas 75 proteínas, com 23 caracterizadas, como as proteínas ‘polyphenol oxidase’ (*Malus domestica*), ‘glutamine synthetase’ (*Glycine max*) e ‘malate-dehydrogenase’ (*Zea mays*) (DECKER et al., 2000). Em outras espécies, foram identificadas, 587 e 302 proteínas em *Lactuca sativa* (CHO et al., 2009; CHO et al., 2010), 21 e 334 em *Chelidonium majus* (NAWROT et al, 2007; NAWROT et al,

2016; 2017), 33 proteínas em *Thevetia peruviana* (FREITAS et al., 2016) e 161 proteínas em *Euphorbia tirucalli* (KITAJIMA et al., 2016).

No látex de *Carica papaya* L. (mamão) foram identificadas 160 proteínas (RODRIGUES et al., 2012), das quais algumas também foram identificadas no látex de *H. sucuuba*, como por exemplo, as proteínas ‘UDP-glucose pyrophosphorylase’ e ‘UBQ8, protein binding’ encontradas no mamão e as proteínas ‘UDP-glycosyltransferase’ e ‘Ubiquitin-like protein 1, partial’. Na tabela 1 estão apresentadas as proteínas identificadas em plantas laticíferas.

Tabela 1. Proteínas de látex identificadas em diferentes plantas por Espectrometria de massas.

Plantas (espécie)	Nº Proteínas	Técnicas de identificação de proteínas	Referências
<i>Thevetia peruviana</i>	33	2D-MS/MS	Freitas et al., 2016
<i>Euphorbia tirucalli</i>	161	LC/Orbitrap MS	Kitajima et al., 2016
<i>Carica papaya</i> L.	160	1D – LC/MS	Rodrigues et al., 2012
<i>Lactuca sativa</i>	587	LC-MS/MS	Cho et al., 2009
<i>Lactuca sativa</i>	302	LC-MS/MS	Cho et al., 2010
<i>Papaver somniferum</i>	75	2D-MALDI-MS	Decker et al., 2000
<i>Chelidonium majus</i>	334	1D-LC/MS	Nawrot et al., 2016
<i>Chelidonium majus</i>	21	2D-LC/MS	Nawrot et al., 2007
<i>Hevea brasiliensis</i>	186	1D-LC/MS	Dai et al., 2013

2.3 Espécie *Himatanthus sucuuba*

O gênero *Himatanthus*, da família Apocynaceae, é composto por 14 espécies (PLUMEL, 1991), dentre elas, *H. sucuuba* (Spruce) Woodson, popularmente conhecida como sucuuba, janaguba ou sucuba, é uma árvore de grande porte e nativa da região Amazônica, que fornece madeira para a construção civil (CÔRREA, 1984). É uma espécie latente, de tronco ereto e casca rugosa, detalhes da árvore estão apresentados na Figura 5.

Figura 5. Árvore *Himatanthus sucuuba*. No detalhe o tronco e casca rugosa da planta.



FONTE: Autor, 2017.

Possui folhas glabras, coriáceas e de margens inteiras. As inflorescências estão dispostas em cimeiras terminais com poucas flores, de cor branca (Figura 6 D) e os frutos apresentam forma de duplo folículo contendo sementes aladas (Figura 6 A e B). As folhas de *H. sucuuba* são simples e alternas, curtamente pecioladas, com aproximadamente 20 cm de comprimento e 7 cm de largura (Figura 6 C). São simétricas, de forma obovado-lanceolada, com ápice agudo a acuminado. A face adaxial apresenta coloração verde escura, enquanto a abaxial é verde amarelada (LARROSA; DUARTE, 2005).

Figura 6. Partes da planta *H. sucuuba* (A fruto - B; semente- C; folha e flor - D).



FONTE: Autor, 2017.

O caule, nas proximidades do ápice, em estrutura secundária incipiente, apresenta a epiderme como sistema de revestimento. Esta é composta por uma única camada de células, revestida por uma cutícula espessada e estriada. Os laticíferos são ramificados, comparativamente maiores do que as células que os ladeiam e possuem citoplasma denso (LARROSA; DUARTE, 2005).

As características anatômicas de ocorrência universal na família são floema interno e laticíferos, encontrados no córtex, no floema e no parênquima medular em *H. sucuuba*. Considerados estruturas especializadas, os laticíferos produzem látex, que é uma suspensão ou emulsão de pequenas partículas. Consistem de uma única célula longa ou de uma série de células (FAHN, 1990), sua parede celular pode ser irregularmente espessada, em razão da plasticidade da mesma (MAHLBERG, 1993). Apresenta-se inteiramente primária, contendo celulose, grande quantidade de substâncias pécticas e hemicelulose (MURUGAN; INAMDAR, 1987; FAHN, 1990).

A importância do gênero *Himatanthus* na medicina popular é considerável, visto que existem relatos de uso em diferentes comunidades locais no Brasil, algumas preparações a partir da casca e do látex para diversos tratamentos são comercializadas (FERREIRA et al., 2009).

O látex dessa planta apresenta alguns compostos, como iridóides (plumieride, isoplumieride) (BARRETO et al., 2007), terpenos (acetato de lupeol, α -amirina) (de MIRANDA et al., 2000), miscelaneos (cis-poliisoprene, arabinose, glicose, xilose) (SILVA et al., 2003) e compostos fenólicos (catecol, quercitrin) que apresentam propriedades biológicas interessantes como anti-inflamatória, anti-tumoral e analgésica (de MIRANDA et al., 2000). Segundo Suffredidi e Daly (2004), o látex tem sido utilizado para tratamentos de tumores, úlceras asma e tuberculose.

Este gênero é amplamente utilizado em distintos tratamentos fitoterapêuticos nos países da América do Sul (SUFFREDINI; DALY, 2004), como por exemplo, as folhas são utilizadas como antifúngico, antianêmico, vermífugo e no tratamento de gastrites e artrites (FERNANDES et al., 2000; DI STASI; HIRUMA-LIMA, 2002). A infusão feita a partir da casca do caule também tem sido usada para tratamento de furúnculos, edemas, artrites e ainda como vermífugo e laxativo (FERNANDES et al., 2000), efeito cicatrizante (VILLEGAS et al., 1997) e uma baixa toxicidade reprodutiva e teratogênica em raízes indicando que seu consumo é seguro na espécie humana no tratamento de gastrites e hemorroidas (GUERRA; PETERS, 1991).

Nos látex, de modo geral, são encontrados também terpenos, carboidratos, lipídios, aminoácidos e proteínas (MORCELLE et al., 2004). Recentemente muitas proteínas têm sido identificadas em diferentes espécies vegetais, e diferentes partes da planta como folhas, caules, raízes e frutos.

2.4 Espectrometria de massas de proteínas

As proteínas são moléculas com as mais distintas funções celulares, sendo capazes de orquestrar expressão de genes, catalisar reações metabólicas, além de compor a parte estrutural da célula (ARN, 2014). Proteoma designa o conjunto de proteínas que estão sendo expressas por uma célula, tecido ou organismo em um determinado momento. De forma distinta, a análise proteômica consiste no estudo do proteoma utilizando técnicas de separação e identificação, tais como eletroforese, cromatografia, espectrometria de massas e bioinformática (HEIN et al., 2013).

A análise proteômica apresenta maior número de variáveis em relação a análise genômica e do transcriptoma devido à grande diversidade química das proteínas e interconectividade destas em complexos e redes de sinalização, as quais podem variar muito de acordo com o tempo e espaço (ALTELAAR; MUNOZ; HECK, 2013). O estudo das proteínas requer ferramentas analíticas com alta seletividade, resolução e sensibilidade (POMASTOWSKI; BUSZEWSKI, 2014).

A espectrometria de massas é uma técnica que data do início do século passado (THOMPSON, 1913) e se baseia na formação de íons na fase gasosa (carregados positiva ou negativamente) que podem ser detectados baseado na sua razão massa/carga (m/z).

A habilidade da espectrometria de massas em analisar proteínas e extratos biológicos advém dos grandes avanços obtidos com o desenvolvimento das técnicas de ionização branda, como a técnica de ionização electrospray (ESI-MS) e a ionização/dessorção a laser assistida por matriz (MALDI), que são técnicas capazes de transformar macromoléculas em íons (TANAKA et al., 1988; FENN et al., 1989).

A técnica de ionização por electrospray foi desenvolvida para o uso em sistemas biológicos por Fenn (FENN et al., 1989). As amostras são dissolvidas em um tampão ou solvente que são bombeadas a um fluxo de microlitros por minuto através de uma agulha hipodérmica que está em uma alta voltagem para dispersar eletrostaticamente, ou electrospray,

gotas de tamanho micrométricos, que são rapidamente evaporadas e transmitem a sua carga para o analito. Duas teorias tentam explicar este fenômeno: “Charged-residue model” e “Ion evaporation model” (GASKELL, 1997). Este processo de ionização ocorre em condições ambientes e, portanto, é muito suave, ou seja, não ocorre a fragmentação dos íons do analito na fase gasosa (MANN; HENDRICKSON; PANDEY, 2001).

3. JUSTIFICATIVA

A busca por novos ativos isolados de plantas medicinais é constante, principalmente em plantas já utilizadas como medicamento pela população. A exemplo do que acontece com outras plantas laticíferas, como *Hevea brasiliensis* e *Calotropis procera*, cujo látex tem sido utilizado na produção de biomembranas e no tratamento de afecções da pele, a identificação das proteínas presentes no soro do látex de *H. succuba* pode fornecer informações importantes e possibilitar o desenvolvimento de novos fármacos.

4. HIPÓTESE

Tendo em vista a ampla utilização do látex de *Himatanthus sucuuba* pela população para o tratamento de várias doenças, nossa hipótese é que exista no látex de *H. sucuuba* proteínas com atividade farmacológica.

5. OBJETIVOS

5.1 Objetivo geral

O principal objetivo do presente trabalho foi a identificação de proteínas solúveis presentes no látex da *H. succuba* através das técnicas de cromatografia de troca iônica, eletroforese unidimensional, bidimensional e espectrometria de massas.

5.2 Objetivos específicos

- Obtenção do látex de *H. succuba*;
- Obter o soro do látex de *H. succuba*;
- Extrair proteínas solúveis presentes no soro do látex e identifica-las através das técnicas de cromatografia de troca iônica, eletroforese bidimensional e unidimensional;
- Identificação das proteínas por espectrometria de massas;
- Análise das sequências de aminoácidos por comparação em banco de dados de proteínas para a identificação;
- Identificação das vias metabólicas e enzimas atuantes.

6. MATERIAL E MÉTODOS

6.1 A coleta de material biológico

O látex foi extraído de árvores nativas de *H. succuba*, localizadas no município de Mazagão Velho, Amapá, Brasil. A exsicata de *H. succuba* (Marciano 001 HAMAB 018780), encontra-se depositada no Instituto de Pesquisas Científicas e Tecnológicas do Amapá (IEPA). A coleta foi realizada no período da manhã, através da técnica do gotejamento. Mediante o corte do caule da planta, foram coletados 420 ml de látex que, após a coleta, foi passado por uma peneira de malha fina para retirar qualquer resíduo de casca que pudesse interferir na qualidade do látex. Em seguida, o látex foi armazenado em tubos de polipropileno à 4°C para posterior utilização.

6.2 Fracionamentos e diálise

O látex foi diluído com água destilada (1/1), e centrifugado a 38.900 x g por 1 hora à 10 °C (Himac CR22G/High-Speed Refrigerated Centrifuge), para separação do soro (fase superior). Após a centrifugação, o sobrenadante foi colocado em membrana de diálise de 25 mm, com capacidade para retenção de moléculas de massa molecular superior a 12.000 KDa (Sigma, MWCO 12.000 Da). A seguir foi dialisado em água, por 72 horas, em câmara fria a 8°C, com trocas de água a cada 6 horas. Após a diálise, as amostras foram novamente centrifugadas e o sobrenadante liofilizado e guardado para posterior uso. A fração líquida (soro do látex) obtida após a centrifugação, foi alvo de todas as análises realizadas nesse trabalho.

6.3 Extração das proteínas do látex

Para a extração das proteínas solúveis presentes no látex de *H. succuba* foram avaliadas três metodologias: 1 - método de precipitação de proteínas com sulfato de amônio modificado por Rajesh (2005); 2 - método de extração de proteínas por ácido tricloroacético/acetona proposto por Li Dejun et al., (2009) e 3 - ácido acético como proposto por Posch et al., (1997), descritas a seguir.

A- Precipitação de proteínas com sulfato de amônio

Uma alíquota de 190 ml do soro foi submetida a precipitação com sulfato de amônio 80%. Ao soro foi adicionado o mesmo volume (1/1) de 10mM tampão fosfato pH 7,0 (1M NaH_2PO_4 /1M Na_2HPO_4) e mantido “overnight” a -4 °C. O sobrenadante foi decantado e centrifugado a 12000 x g por 20 minutos à 4 °C. Após a centrifugação, o sobrenadante foi submetido a precipitação das proteínas em 80% de sulfato de amônia. Centrifugado novamente a 12000 x g por 20 minutos a 4 °C e o pellet obtido foi dissolvido em 10 mM de tampão fosfato e dialisado no mesmo tampão para remoção do sulfato de amônia. Após a diálise a amostra foi liofilizada e ressuspensa em 600 µL água (4,5 ug/ul). Uma alíquota de 25 µL dessa amostra de proteínas (113,5 µg) foi diluída em 225 µL de tampão de solubilização (TS) composto por ureia 7M, tiureia 2M, CHAPS 4 % (w/v) para posterior utilização.

B- Precipitação de proteínas ácido tricloroacético/acetona

Em 35 ml de soro de látex de *H. sucuuba*, foi adicionado o mesmo volume de acetona gelada (-20 °C) contendo 20 % (v/v) TCA e 4 mmol/L de ditioneitol (DTT). A mistura foi homogeneizada e precipitada por 12 horas à - 20 °C. Após a precipitação, foi centrifugada à 18000 x g por 30 minutos à 4 °C. O sobrenadante foi descartado e o pellet foi lavado (3 vezes), com 10 ml de acetona gelada (- 20 °C) contendo 2 mmol/L de DTT. Posteriormente foi centrifugada por 15 minutos à 4 °C e 18000 x g. O pellet foi ressuspensa em 500 µL de água ultrapura (MilliQ). Após a quantificação, uma fração contendo 200 µg de proteína, foi solubilizada em 225µL de tampão de solubilização (TS) composto por ureia 7M, tiureia 2 M, CHAPS 4 % (v/v) e reservada para posterior utilização.

C. Precipitação de proteína do látex extraídas com ácido acético

Em 260 mL de soro do látex foram adicionados 10 % de ácido acético (v/v) e a mistura foi mantida em repouso por 30 minutos e centrifugada por 1 hora (18.000 x g) à 10 °C. Após a centrifugação o sobrenadante foi dialisado com água destilada por 72 horas e liofilizado. A amostra liofilizada foi diluída em água e as proteínas quantificadas (3,5 ug/uL). Foi utilizado o Clean-Up kit (GE Healthcare, USA) seguindo as instruções do fabricante para limpeza da proteína. Uma alíquota de 25µL dessa amostra contendo 200 µg de proteínas foi solubilizada em 225 µL de tampão de solubilização (TS) composto por ureia 7M, tiureia 2 M, CHAPS 4 % (v/v).

6.4 Quantificação das proteínas

A quantificação de proteínas totais, presentes no soro, foi realizada pelo método Lowry, modificado por Hartree, 1972. Foram utilizados 10 μL das amostras após a liofilização e solubilização. A concentração foi determinada a partir da curva padrão estabelecida, utilizando soro albumina bovina (BSA), (INLAB lote, 839250).

6.5 Eletroforese em gel de poliacrilamida

Após a quantificação, 15 μg de proteínas foram desnaturadas à 95 °C, e analisadas por eletroforese em gel de poliacrilamida 12 %, na presença de dodecil sulfato de sódio (SDS-PAGE), utilizando-se o padrão de massa molecular de 10 a 225 KDa (The Thermo Scientific Pierce Prestained Protein). O gel foi corado com Coomassie R-250 para visualização das proteínas.

6.6 Cromatografia de troca iônica

A cromatografia foi utilizada para a purificação das proteínas, presentes no soro do látex. Iniciou-se utilizando as colunas HiTrap SPXL (catiônica) e HiTrap QXL (aniônica), no equipamento AKTA Purifier (GE HEALTHCARE, Björkgatan, Suécia) seguindo as instruções do fabricante. Para tanto foram utilizados 100 μL (6 mg/mL) de amostra na coluna: HiTrap QXL (aniônica) e HiTrap SPXL (catiônica) respectivamente, ambas com fluxo de 1ml/minuto. A coluna aniônica foi equilibrada com tampão contendo 20 mM Tris-HCl, pH 8,0 e a amostra eluída com gradiente de 0-100% de tampão de eluição (20 mM Tris-HCl + 1M NaCl, pH 8,0), enquanto que para a coluna catiônica utilizou-se o tampão de equilíbrio (20 mM $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, pH 6,8) e a amostra foi eluída com gradiente de 0-100% de tampão de eluição (20 mM $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ + 1M NaCl, pH 6,8).

6.7. Eletroforese bidimensional

Para o processo de eletroforese bidimensional, o primeiro passo foi a reidratação das fitas IPG (com pH entre 3-10 de 13 centímetros). A solução de reidratação das IPGs (Immobiline DryStrip gel) foi preparada da seguinte forma: Ureia 7 M 10,5 g, Thiureia 2 M 3,8 g, CHAPS 2 % (v/v) 0,5 g, IPG Buffer 0,5 (v/v), 1 % de Bromophenol blue, ditiotreitol (7 mg

em 2,5 mL). Em um tubo eppendorf, adicionou-se primeiro 25 μ L de amostra (240 μ g), 1,25 μ L de IPG Buffer, 2 μ L de azul de bromofenol e 225 μ L de tampão de reidratação, homogeneizou-se e foi deixado por 30 minutos descansando em temperatura ambiente e centrifugado (Cenrtifuge 5424 R Eppendorf) por 1 min. O tampão de reidratação foi colocado na canaleta do IPG-box, evitando-se a formação de bolhas. Colocou-se a fita com gel em contato com a solução dentro da canaleta e deixou-se hidratando em temperatura ambiente durante 16 horas.

6.7.1 Focalização Isoelétrica (IEF)

Através do processo de focalização isoeletrica (IEF) as proteínas foram separadas de acordo com o seu ponto isoeletrico (pI) (primeira dimensão), usando o sistema Ettan IPG phor e as fitas IPGs 3 de 13 cm (GE Healthcare).

Após a hidratação, a fita foi colocada no Ettan IPGphor 3 (GE Healthcare). Foram aplicados 2 mL de óleo mineral (PlusOne DrysTtip Cover Fluid) para focalização isoeletrica realizada nas seguintes condições: 100 V por 4 horas, 500 V por 2 horas, 500 V por 2 horas, 1000 V por 2 horas, 1000 V por 2 horas, 5000 V por 2 horas, 5000 V por 2 horas, 8000 V por 2 horas, 8000 V por 2 horas até atingir 20000 V/h. Após a focalização isoeletrica, as fitas foram retiradas do aparelho Ettan IPGphor 3, lavadas em água destilada para tirar o excesso de óleo mineral, armazenadas em tubo de ensaio longo (com tampa de rosca) à -80 °C, até a sua utilização.

Imediatamente antes da segunda dimensão, a fita foi mantida em solução de equilíbrio (75 mM Tris-HCL pH 8,8 (10 mL), ureia 6 M 72,1g, glicerol (30 %) 69 ml/84,2g, duoadecil sulfato de sódio 4 g (SDS) Tracting dye (azul de bromofenol) acrescidas de ditiotreitól e 2 % (v/v) iodoacetamida. Esse procedimento acontece em duas etapas denominados primeira e segunda lavagens.

Na primeira lavagem são utilizados 10 mL da solução, acrescida de ditiotreitól 100 mg, por 15 minutos com agitação moderada. O ditiotreitól é utilizado como um agente redutor, quebrando as pontes de sulfeto das proteínas. Na segunda lavagem utiliza-se novamente 10 mL da solução de equilíbrio, acrescida de iodoacetamida (250 mg/10 mL) que é responsável pela alquilação.

6.7.2 Eletroforese bidimensional ou Segunda dimensão

As proteínas, após a focalização, foram submetidas a eletroforese vertical utilizando o sistema SE 600 Ruby (GE HealthCare) e gel bis-acrilamida de 12,5 % preparado segundo as recomendações do fabricante. Após a eletroforese, o gel foi submetido por 30 minutos a uma solução fixadora (40% v/v de etanol, 10 % de ácido acético), posteriormente, foi lavado duas vezes com água deionizada por 10 minutos e mantidos por 12 horas em solução de coloração Comassie Brilliant Blue G-250 incluir aqui a concentração do CBB. Depois de corado, o gel foi lavado 200 mL de ácido acético 1 % (v/v) até a eliminação excedente completa do corante (GE HealthCare, 2015).

6.8 Espectrometria de massas

A análise das proteínas totais por espectrometria de massas foi realizada no Laboratório de Química de Proteínas da USP.

Para a realização da espectrometria de massa foi utilizado o equipamento UPLC-nanoAcquity acoplado ao ESI-QUAD-TOF (Ultima, Waters, UK). Todos os dados gerados pelo espectrômetro de massas MS foram analisados através da ferramenta de busca MASCOT (<http://www.matrixscience.com>) e foi utilizado como banco de dados para as buscas o NCBIInr, Swiss-Prot.

Para as buscas no MASCOT (www.matrixscience.com) é necessário configurar alguns parâmetros, portanto, foram utilizados os seguintes parâmetros: i) tipo de pesquisa: MS/MS Ion; ii) enzima: tripsina; iii) modificação fixa: carbamidometilação das cisteínas; iv) modificação variável: oxidação da metionina; v) taxonomia: viridiplantae (Green plants); vi) valores de massa: monoisotopic; vii) massa de proteína: sem restrições; viii) tolerância de fragmentos de massa: ± 1.2 Da; ix) tolerância de massa de peptídeo: ± 0.6 Da; x) número de clivagem: no máximo 1; xi) tipo de instrumento: ESI-QUAD-TOF; e os valores de “ion score” considerados significantes foram os valores pré-estabelecidos pelo Mascot. A seleção das proteínas seguiu a recomendação de Yang et al. (2017) onde somente peptídeos com score igual ou superior a 20 apresentam ótima identidade, ou seja, selecionar sequências com score ≥ 20 reduz a probabilidade de identificação de falsos peptídeos. O programa MASCOT compara os dados obtidos pela espectrometria de massas com os dados de proteínas existentes em um determinado banco de dados. Para as anotações funcionais foi utilizado o software free

Blast2GO PRO Trial (v. 4.1), (CONESA et al., 2008). As sequências de proteínas identificadas pelo programa Mascot (v.2.4.1) de acordo com seu código de identificação, foram pesquisadas no UniProt (onde os arquivos com as sequências de proteínas foram salvas como .fasta) E anotadas de acordo com o Gene Ontology (GO), nos termos de *biological processes, molecular functions e cellular componentes*.

6.8.1 Eletroforese em gel de poliacrilamida 12% (SDS-PAGE)

A amostra contendo 100 µg de proteínas totais, foi submetida à eletroforese em gel de poliacrilamida 12 % contendo SDS, de acordo com Laemmli (1970). A amostra foi preparada em tampão redutor e desnaturante (12 mM Tris-HCL pH 6,8, 0,02 % (m/v) SDS, 1 % (v/v) glicerol, 20 mM de β-mercaptoetanol, DTT 0,2 M (1 %) e 0,004 % (m/v) de azul de bromofenol e foi submetida a 100 °C por 10 minutos em termociclador (PTC-100™ Programmable Thermal Controller MJ Research, Inc). A corrida eletroforética foi realizada em 20 mA por 3 horas e 30 minutos e o gel foi corado com Comassie Brilliant Blue G-250.

Após a lavagem do gel com água deionizada, foi submetido por 30 minutos a uma solução fixadora (40 % v/v de Etanol, 10 % de Ácido Acético), lavado duas vezes com água destilada deionizada por 10 minutos e mantido por 12 horas em solução de coloração (Comassie Brilliant Blue G-250). Depois de corado, o gel foi lavado com 1 % (v/v) de ácido acético até a eliminação excedente completa do corante. As bandas visualizadas no gel foram recortadas com um bisturi. Cada fragmento obtido foi colocado em tubo eppendorf (2 mL). Esses fragmentos foram lavados em água deionizada, descorados várias vezes na solução contendo 50 % v/v de acetonitrila (ACN) e 25 mM de bicarbonato de amônio (AMBIC), até a total remoção do corante (a cada lavagem adicionou-se 100 µL de solução e foi agitado por 2 horas). Posteriormente foram desidratados duas vezes em 100 % ACN por 10 min. A ACN foi removida através de pipeta e o resíduo remanescente do gel removido através da evaporação em Speed vac. SPD1010 and SPD2010 Integrated SpeedVac™ Systems (Thermo Scientific™)

6.8.1.1 Redução e alquilação

Os fragmentos do gel foram reidratados e reduzidos (esse procedimento desfaz as pontes dissulfetos, ajudando a eliminar a estrutura tridimensional dos polipeptídeos) em ditioneitol (20 mM DTT/50 mM AMBIC) a 60 °C por 40 minutos e alquilados (com iodoacetamida 55 mM IA /50 mM AMBIC) no escuro por 30 minutos. A iodoacetamida foi removida e os fragmentos

foram lavados em 25 mM de bicarbonato de amônio (o suficiente para cobrir a banda seca). Posteriormente foram desidratados em 100 % de acetonitrila. A acetonitrila foi descartada e o resíduo remanescente no gel foi evaporado em temperatura ambiente.

6.8.1.2 Digestão das proteínas

Os fragmentos do gel foram reidratados em 15 μ L de solução contendo 150 ng de tripsina em 25 mM de bicarbonato de amônio e incubados em banho seco a 37 °C por 24 horas. A ação da tripsina foi interrompida pela adição de 2 μ L de ácido fórmico puro.

Os peptídeos obtidos após a tripsinização foram eluídos em 50 μ L de uma solução aquosa de ácido fórmico 5 % (v/v) e metanol 60 % (v/v), centrifugados (Cenrtifuge 5424 R Eppendorf) por 1 min. Recolheu-se dessa solução 25 μ L que foram colocados em tubos cônicos para a espectrometria de massas. O sequenciamento dos peptídeos foi feito em espectrômetro de massas ESI-Q-TOF.

6.9. Identificação de metabólitos secundário presentes no látex da espécie *Himatanthus sucuuba* (Apocynaceae) por Espectrometria de massas

O látex de *H. sucuuba* foi fracionado por solvente de acordo com a sua polaridade. Utilizou se 20 ml de látex com pH 5 e diluído em 40 ml de água milli-Q (1:2) e a solução foi acidificada com ácido clorídrico até o pH 2. Em 60 ml de solução foi adicionado 100 mL de acetato de etila P.A (MERCK), agitou se algumas vezes e foi retirada a água, esse procedimento foi realizado três vezes e a cada retirada da água foram adicionados 100 ml de acetato de etila. Até ocorrer a total desidratação do látex, onde houve a separação de duas fases: fase aquosa e fase acetato (partição líquido/líquido). A fase aquosa foi liofilizada.

Para análise no espectrômetro de massas foram utilizadas quatro frações da amostra de latex: fase aquosa, acetato, soro liofilizado e soro bruto do látex de *H. sucuuba*. Nas frações aquosa e acetato foram eluídas com água e metanol e nas amostras de soro liofilizado e soro bruto foi adicionado metanol. Para à análise foi utilizado o espectrômetro de massas quadrupolo tandem Xevo™ TQ-S acoplado ao sistema UPLC ACQUITY H-Class, com ionização por electrospray e analisador quadrupolo (ESI-QUAD) para a identificação de compostos orgânicos. As análises foram realizadas no Departamento de Química da Faculdade de

Filosofia, Ciências e Letras de Ribeirão Preto – USP. Especificamente no Laboratório de Espectrometria de Massas Aplicada a Química de Produtos Naturais.

7. RESULTADOS e DISCUSSÃO

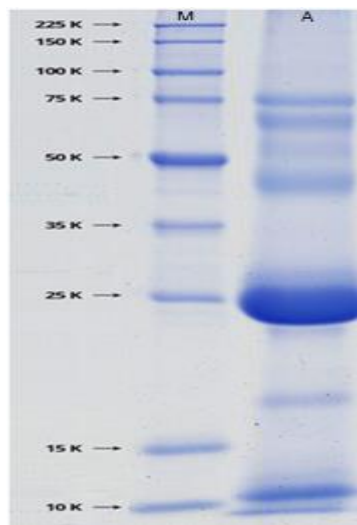
7.1 Quantificação de proteínas

A fração líquida (soro do látex) obtida após a centrifugação, foi alvo de todas as análises realizadas nesse trabalho. A concentração das proteínas solúveis presentes no soro foi de 8,4 mg/mL.

7.2 Eletroforese SDS-PAGE

Uma alíquota desta amostra de proteínas (15 µg), foi submetida a eletroforese em gel de poliacrilamida e está apresentada na Figura 7, onde se constata a presença de proteínas de diferentes pesos moleculares.

Figura 7. Perfil eletroforético de amostra do soro do látex de *H. sucuuba* (15 µg) em PAGE-SDS 12 % corado por Coomassie R-250. M: Padrão de peso molecular (Broad range Protein molecular weight markers, PROMEGA). A: amostra do soro com proteínas totais.



7.3 Cromatografia de troca iônica

Na cromatografia de troca catiônica, utilizando a coluna HiTrap SP XL, são separadas as proteínas de carga positiva. Embora tenham sido coletadas 26 amostras de 1 mL, como pode ser observado na Figura 08, não se verificou a presença de proteínas nas frações coletadas, apenas no primeiro tubo há um pico evidenciando que as proteínas não se ligaram a resina presente na coluna e portando é possível concluir que, não há proteínas catiônicas (+) nesta amostra de soro do látex da *Himatanthus sucuuba*. Quando se utilizou a coluna HiTrap QXL, que retêm as proteínas negativas da amostra, foi possível verificar que houve a separação de

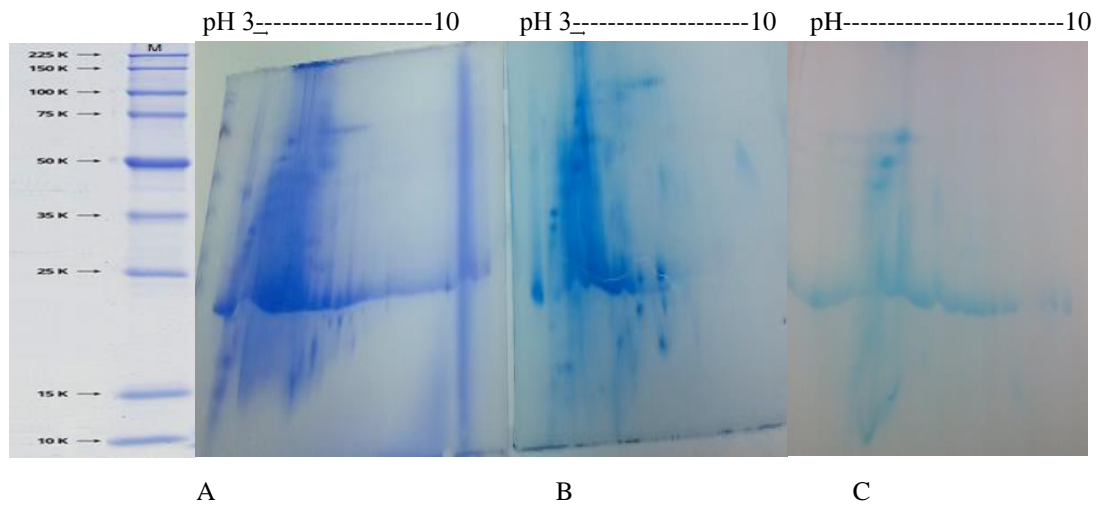
7.4 Eletroforese Bidimensional

A extração de proteínas do látex de plantas é frequentemente realizada por precipitação com Sulfato de amônio, TCA/acetona e ácido acético, porque são compatíveis com eletroforese 2-D e Espectrometria de massas. Mesmo utilizando esses métodos não se detectou a perfeita separação das proteínas nos géis de eletroforese 2D para viabilizar a análise por espectrometria de massas (Figura 10). Na figura 10A, está apresentado o gel da amostra de proteínas extraída com TCA/acetona, onde se verifica um arraste menor que o gel da figura 10B resultante do processo de extração com sulfato de amônia. Na figura 10C do gel contendo proteínas extraídas com ácido acético, o arraste é menos visível mas em compensação aparecem alguns spots. Possivelmente, por conter componentes não proteicos que interferem na focalização isoelétrica, o que inviabilizou o prosseguimento da análise por eletroforese bi-dimensional.

Dos três métodos utilizados na extração de proteínas totais do soro, o método TCA / acetona mostrou o maior rendimento protéico (14.76 mg de proteínas por ml de soro), seguido do método Sulfato de amônio (4,54 mg de proteínas por ml de soro) e método de extração de Acido acético foi o que teve o menor rendimento (3,5 mg de proteína por ml de soro). o rendimentos das proteínas variaram de acordo com diferentes métodos.

Segundo Duan et al. (2006), o látex é rico em compostos, tais como sal, sais minerais, lipídios, hidratos de carbono, e em particular os sistemas de membranas complexas, que podem ser co-extraídos com proteínas e interferir neste tipo de análise. Não existe nenhum método eficaz especificamente concebido para o isolamento de proteínas a partir de diferentes frações de látex de borracha. Assim optou-se pela separação de proteínas por eletroforese unidimensional (1D) e identificação por espectrometria de massas.

Figura 10. Fotos dos géis de segunda dimensão com amostras resultantes dos três métodos de extração testados. A- Extração de sulfato de amônio; B- Extração com TCA/Acetona e C - Extração com TCA/Acetona. A Focalização foi realizada em gradiente de pH 3-10. Coloração com comassie brilhante azul G250. (M: Padrão de peso molecular 10 a 225 KDa (Amersham™ ECL™ Rainbow™ Marker- Ful Rang

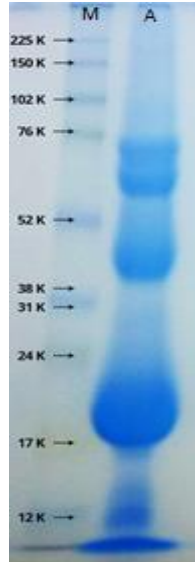


7.5 Identificação de proteínas por análise MS/MS

Em função da impossibilidade de separação e identificação das proteínas do soro utilizando eletroforese 2D, optou-se pela análise por espectrometria de massas a partir das bandas observadas na eletroforese de uma dimensão. Assim, 100 µg de proteínas foram separadas por eletroforese unidimensional em gel de acrilamida a 12 % (1D), na presença de dodecil sulfato de sódio (SDS-PAGE), Figura 11.

Após a eletroforese o gel foi corado com Comassie G250, fotografado e as bandas recortadas para serem tripsinizadas e submetidas a separação por MS/MS.

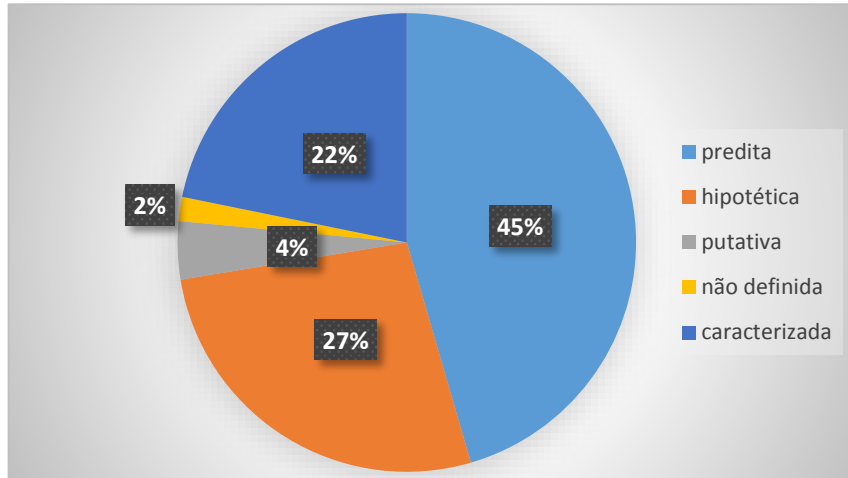
Figura 11. Perfil eletroforético em PAGE-SDS 12 % corado por Coomassie G-250, das Proteínas totais extraídas do soro do látex de *H. sucuuba*. M: Padrão de peso molecular 12 a 225 KDa (Amersham™ ECL™ Raibow™ Marker). A: amostra do soro com proteínas totais.



Os espectros de massas obtidos foram submetidos à análise pelo MASCOT um programa de análise de proteínas integrado aos bancos de dados NCBI - National Center for Biotechnology Information non-redundant e Swiss prot. Foram caracterizados com o número de acesso (ID), massa molecular (KDa), pI, número de aminoácidos, Score da proteína (parâmetro que indica o nível de confiabilidade do resultado obtido na comparação com as proteínas dos bancos de dados - significância do sequenciamento) e percentagem de cobertura do peptídeo em relação a sequência de aminoácidos identificada. Todas as proteínas identificadas no soro de *H. sucuuba* tem similaridade com as sequências de outras plantas.

A utilização da técnica de espectrometria de massas resultou na identificação de 1471 proteínas, das quais 587 proteínas apresentaram score igual ou superior a 20. Dentre essas 587 previamente identificadas, foram encontradas 267 proteínas preditas o que equivale a 45 %, 158 hipotéticas (27 %), 128 anteriormente caracterizadas (22 %), 24 putativas (4 %) e 10 proteínas não definidas (2 %). Os dados estão apresentados na Figura 12.

Figura 12. Proteínas potencialmente identificadas no látex de *Himatanthus sucuuba*.



Com base na repetibilidade de algumas proteínas, 136 foram selecionadas entre elas, caracterizadas e outras preditas estão apresentadas na Tabela 2. Nessa tabela, estão apresentados massa molecular da proteína, o ponto isoelétrico estimado, o número de aminoácidos da proteína, o valor do score, a sequência dos peptídeos obtidos, o código de identificação (ID) NCBI, o nome da proteína e o organismo de origem desta proteína. A partir do ID de cada proteína foi possível identificar e obter a sequência FASTA das proteínas, que estão apresentadas nos anexos.

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

	Massa molecular (KDa)	Número de aminoácidos	pI	Score da proteína	Sequências de aminoácidos dos peptídeos identificados	ID	Descrição da proteína
1	59,8	534	7,68	179	K.YGLAADNVIDAR.I R.ESMGEDLFWAIR.G	gi 743891207	tetrahydrocannabinolic acid synthase-like [<i>Populus euphratica</i>]
2	60.1	535	6,54	175	K.YGLAADNVIDAR.I K.ESMGEDLFWAIR.G	gi 496405183	FAD – dependet oxidoreductase [<i>Papaver somniferum</i>]
3	47.7	418	5,65	175	K.YGLAADNVIDAR.L R.KSMGEDLFWAIR.G	gi 30315245	nectarin 5, partial [<i>Nicotiana langsdorffii</i> x <i>Nicotiana sanderae</i>]
4	60.4	535	9,17	175	K.YGLAADNVLDAR.I R.KSMGEDLFWAIR.G	gi 747046464	tetrahydrocannabinolic acid synthase-like [<i>Sesamum indicum</i>]
5	61.4	543	5,42	175	K.YGLAADNVIDAR.L R.QSMGEDLFWAIR.G	gi 848927876	cannabidiolic acid synthase-like [<i>Erythranthe guttata</i>]
6	62.5	550	8,89	173	K.YGIAADNVLDAR.V R.ESMGEDLFWAIR.G	gi 848927873	tetrahydrocannabinolic acid synthase-like [<i>Erythranthe guttata</i>]
7	59.9	534	8,07	162	K.YGLAADNVIDAR.I	gi 1009123039	tetrahydrocannabinolic acid synthase-like [<i>Ziziphus jujuba</i>]
8	60.1	536	9,28	113	R.SMGEDLFWAIR.G K.EYFPELGLVR.E	gi 1040819157	tetrahydrocannabinolic acid synthase-like [<i>Daucus carota</i> subsp. <i>sativus</i>]
9	37.0	327	9,00	111	K.FGLSADNVLDAR.V R.ESMGEDLFWAIR.G	gi 970039800	tetrahydrocannabinolic acid synthase-like [<i>Solanum pennellii</i>]
10	60.8	538	6,62	89	K.YGLAADNIVDAR.L	gi 460393282	cannabidiolic acid synthase-like [<i>Solanum lycopersicum</i>]
11	62,6	544	8,88	88	R.KSMGEDLFWALR.G	A6P6V9	Cannabidinolic acid synthase [<i>Cannabis sativa</i>]
12	62,2	545	8,77	88	R.KSMGEDLFFWAI.R.G	Q33DQ2	Inactive tetrahydrocannabinolic acid synthase [<i>Cannabis sativa</i>]
13	59.1	532	9,19	80	R.TNMGEDLFWAIR.G	gi 729294682	reticuline oxidase-like protein [<i>Tarenaya hassleriana</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

14	62.0	566	6,41	79	R.ESMGEDIFWAI.R.G	gi 695072053	reticuline oxidase-like [<i>Musa acuminata</i> subsp. <i>malaccensis</i>]
15	61.6	548	6,36	79	R.ESMGEDLFWAL.R.G	gi 1040881446	cannabidiolic acid synthase-like 1 [<i>Daucus carota</i> subsp. <i>sativus</i>]
16	61.4	542	8,13	78	R.KSMGEDLFWAL.R.G	gi 20563190	Carbohydrate oxidase [<i>Helianthus annuus</i>]
17	64.7	575	9,13	78	R.KSMGEDIFWAI.R.G	gi 590614781	FAD-binding Berberine family protein [<i>Theobroma cacao</i>]
18	60.4	539	8,92	75	K.YGLAGDQVIDAR.L	gi 1040922945	flavin-dependent oxidoreductase FOX2-like [<i>Daucus carota</i> subsp. <i>sativus</i>]
19	39,1	352	6,2	57	K.YIAVGNEVDPVK.F	gi 565404247	glucan endo-1,3-beta-glucosidase A-like [<i>Solanum tuberosum</i>]
20	79,9	708	8,8	50	K.SVLAVLISK	gi 475663145	Hyaluronan synthase 2 [<i>Fusarium oxysporum</i> f.sp. <i>cubense</i> race 4]
21	86,5	810	6,41	46	K.EFPADVILGDGR.V	gi 320117871	Subtilisin-like protease 2 [<i>Phaseolus vulgaris</i>]
22	138,4	1.202	6,26	39	K.KVRSILSK.	gi 304325323	Rp1-like protein, partial [<i>Zea mays</i> subsp. <i>Parviglumis</i>]
23	39,7	350	5,32	37	K.YGLAAPAKFLVK.G	gi 745698892	Norcochlorine 6-O-methyltransferase [<i>Sinopodo phylum hexandrum</i>]
24	96,5	868	6,0	36	K.SFMSILSK.K	gi 635148931	Phototropin [<i>Desmidium aptogonum</i>]
25	112,4	1037	5,98	35	K.TGTLTTNQMSVIQ.V	P54209	Cation-transporting ATPase CA1 [<i>Dunaliella bioculata</i>]
26	25,6	221	10,9	35	K.YYMDGQTILGR	gi 573951433	Serine/arginine-rich SC35-like splicing factor SCL33 [<i>Oriza brachyantha</i>]
27	133,8	1.227	5,99	35	K.QNEDLNSNLLVEIK.S	gi 15228737	Carbohydrate-binding-like-fold-containing protein [<i>Arabidopsis thaliana</i>]
28	32,5	293	5,58	34	K.AMSLRAPVSTFK.V	Q9SU86	Cytidine deaminase 6 [<i>Arabidopsis thaliana</i>]
29	54,4	520	8,34	34	K.LNLVRRRCR.S	gi 560939381	Serine hydroxymethyltransferase [<i>Camellia sinensis</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

30	43,1	393	5,67	34	R.RALVELTTR.V	gi 527184620	Geranyl pyrophosphate synthase [<i>Genlisea aurea</i>]
31	128,2	1.170	5,71	33	K.LAKPDGAASVTISPR.V	gi 674863539	BnaCO4g54120D [<i>Brassica napus</i>]
32	121,3	1054	6,69	33	K.EHLSVFK.V	gi 901822427	Zinc finger CCCH domain- containing protein 4 [<i>Zostera marina</i>]
33	57,6	516	6,29	33	R.EHAEVFK.	gi 1024048764	14-3-3-like protein D-like [<i>Dorceras higrometricum</i>]
34	75	669	5,53	33	K.MRYVFK.N	gi 703096601	Translation factor GUF-1like protein [<i>Morus notabilis</i>]
35	19,5	170	8,31	32	MGVYVFK.R	gi 21666506	ADP-glucose pyrophosphorylase larg subnit, partial [<i>Metroxylon sagu</i>]
36	132,7	1.189	7,0	31	R.GGEAPTGLVVDAFADKV	gi 612397112	DNA repair and recombination protein RAD26 [<i>Bathyococcus prasinus</i>]
37	85,5	791	9,39	31	R.SALMTTAWMTNDKK.K	Q9FK76	Subtilisin-like protease SBT5.6 [<i>Arabidopsis thaliana</i>]
38	74,9	671	6,5	31	K.AVIVPHSMGVLYFLHFMK.W	Q9FNA9	Phospholipid:diacylglycerol acyltransferase 1 [<i>Arabidopsis thaliana</i>]
39	30,9	288	9,81	30	R.FVSVEWMKRL	gi 145343464	MC family transporter: uncoupling protein [<i>Ostreococcus lucimarinus CCE9901</i>]
40	60,9	512	9,46	30	K.AYSTSGGLDRKR.I	gi 1043378536	Maturase K (Chloroplast) [<i>Primula sinensis</i>]
41	36,4	319	5,81	29	K.ILKQIGGGFPIVVSSIAK.E	gi 922333312	Trehalose-6-phosphate phosphatase [<i>Medicago truncatula</i>]
42	20,9	185	6,06	29	K.DAKRLLSK	gi 590566736	Pectinesterase inhibitor-like protein [<i>Theobroma cacao</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

43	44,7	408	9,57	29	K.QKLDSSSTGGDHRR.E	gi 976923108	Aspartate/ glutamate/ uridylate kinase [<i>Cynara cardunculus</i> var. <i>scolymus</i>]
44	55,7	537	9,30	29	R.VTASLVPR.Q	gi 552846846	Expressed protein [<i>Chlorella variabilis</i>]
45	61,3	545	7,22	28	R.VMELLICKMIK.G	gi 590576851	Alpha/beta-hydrolases superfamily protein [<i>Theobroma cacao</i>]
46	84,5	754	8,23	28	K.NLHLHPVLMTWGYFPESKSDFKGK	gi 922363196	Long-chain-alcohol oxidase FAO2-like protein [<i>Medicago truncatula</i>]
47	78,6	691	8,96	28	R.WAIVMSAGIH	gi 976899861	Proteins kinase, ATP binding-site-containing protein [<i>Cynara cardunculus</i> var. <i>scolymus</i>]
48	24,9	220	5,23	27	MAFKLYGLPMSTNTTRAMICLHEK	gi 734433024	Glutathione S-transferase F13 [<i>Glycine soja</i>]
49	16,7	16	3,66	27	NFPLDLAAIEAPSTNV	gi 16305027	PsbA protein, partial (chloroplast) [<i>Artemisia rupestris</i>]
50	47,7	401	5,88	27	K.DYYPTR	gi 1012353235	KDEL motif-containing protein 1 [<i>Cajanus cajan</i>]
51	106,1	944	9,28	27	K.IYYPTR.H	gi 674947203	BnaAO6g30730D [<i>Brassica napus</i>]
52	37,9	347	10,1	26	GMAMLLTDGEVGALIKVSAAVVWVAMSYARLA AAR	gi 413957119	Acyltransferase [<i>Zea mays</i>]
53	35,6	326	5,74	26	K.SSSSSAPGSLNFDLR.I	gi 735880	Gearanylgeranyl pyrophosphate synthase-related protein [<i>Arabidopsis thaliana</i>]
54	82,6	717	6,3	26	K.QADAAWEILHR.S	gi 976910661	Alfa-N-acetylglucosaminidase [<i>Cynara cardunculus</i> var. <i>scolymus</i>]
55	54,4	490	8,88	26	R.LGVAVVEVGFATPLLTSR	gi 760449485	Serine palmitoyltransferase 2 [<i>Auxenochlorella protothecoides</i>]
56	133,2	1160	5,70	26	K.SLDRLNLSGCSR.L	gi 10177584	Disease resistance protein-like [<i>Arabidopsis thaliana</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

	Massa molecular (KDa)	Número de aminoácidos	pI	Score da proteína	Sequências de aminoácidos dos peptídeos identificados	ID	Descrição da proteína
57	23,0	212	5,51	26	K.NLQVLHSLRLGK	gi 1035946765	26S proteasome non-ATPase regulatory subunit 9 [<i>Ananas comosus</i>]
58	37,7	399	8,16	26	R.QGMSNLVDAINK.L	gi 1024000460	Methyl-CpG-binding protein 2-like [<i>Doroceras hygrometricum</i>]
59	27,9	239	10,24	26	K.DRGANTYNSSDNLGYIR.G	gi 575396380	Ribosomal protein S7 (mitochondrion) [<i>Dicranum scoparium</i>]
60	69,8	627	6,64	26	R.AAEEETMMYDVNVVQR.I	gi 703110509	BTB/POZ domain containing protein NPY2 [<i>Morus notabilis</i>]
61	29,4	249	6,30	26	R.EAAQDDFFWKCLCAR.R	gi 1024041777	F-box protein [<i>Doroceras hygrometricum</i>]
62	27,1	27	4,65	25	FVMYER.N	gi 634575543	Psba, partial (plastid) [<i>Eugenia bacopari</i>]
63	61,5	550	9,0	25	MTRGAAAAALLALALVAVAR	gi 301030585	L-ascorbate oxidase-like protein [<i>Cynodon dactylon</i>]
66	29,4	259	7,51	25	R.SAPHDMGMPYMEKKPLTEE R.A	gi 164605537	CMO545.270.nc [<i>Lotus japonicus</i>]
67	20,9	178	9,72	25	MSNHLKNVATVGLR	gi 971517319	Mitochondrial ribosomal protein L13 precursor [<i>Klebsormidium flaccidum</i>]
68	49,9	442	9,56	25	K.STAQRPSEEDSLKR.E	gi 971519052	Peptidase M15B and M15C [<i>Klebsormidium flaccidum</i>]
69	54,4	465	8,4	25	K.SLMILSMLIHR.S	gi 357498019	Eukaryotic translation initiation factor 3C [<i>Medicago truncatula</i>]
70	46,8	446	5,6	25	R.AGVEMALIDAIANSIR.I	gi 54290836	Cis,cis-muconate cycloisomerase-like [<i>Oryza sativa</i> Japonica Group]
71	51,9	454	6,20	25	AFLTEERSAVPDLETHVGGGMKSV ER.N	gi 566188069	TRNA isopentenyl transferase family protein [<i>Populus trichocarpa</i>]
72	69,1	610	8,35	24	K.EITVTITVVGASGDLAK.K	gi 901793024	Glucose-6-phosphate dehydrogenase [<i>Zostera marina</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

73	13,33	116	9,23	24	R.DCQQFRAIHR.A	gi 8347414	Itegrase/partial [<i>Oriza rhizomatis</i>]
74	76,2	660	8,90	24	R.ISMIKSMATATIATSADFDGVQVERT.K	gi 590647573	Ribonuclease H protein [<i>Theobroma cacao</i>]
75	22	205	6,0	24	K.HTILYSKYAGNDFKGS DGTNYIALR	gi 590613556	Chaperonin 20-like protein [<i>Theobroma cacao</i>]
76	59,7	534	6,2	24	K.TKPTADQDEDGIGTR	gi 297837155	F23N19.7 [<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>]
77	46,2	402	5,25	24	R.GKPHLMKNIHR.	gi 480328109	Heat stress transcription factor A-4a-like protein [<i>Caragana korschinskii</i>]
78	51,7	456	8,13	23	K.GDRILMMSFGAGFK	gi 703099618	3-ketoacyl-CoA synthase 21 [<i>Morus notabilis</i>]
79	55,1	498	6,24	23	R.AGSMDNGDGI AVGWLRHPIFXDNER.R	gi 33327957	Photosystem II CP47 protein, partial (Chloroplast) [<i>Euptelia polyandra</i>]
80	53,76	478	5,58	23	R.VGDWYEASTTNKR.K	gi 18394553	Protein maintenance of meristem [<i>Arabidopsis thaliana</i>]
81	102,2	914	6,43	23	R.QWAARMEAAGM ERVGLSYSGAMEAR.K	gi 77548634	GRAS family transcription factor containing expressed [<i>Oryza sativa Japonica Group</i>]
82	32,8	293	6,18	23	K.VEDEVSA DYRITGVPADGRCLFR.A	gi 1012326608	OTU domain-containing 6B [<i>Cajanus cajan</i>]
83	26,2	235	8,7	23	SPPSAGAGSRAAGVGEEQEML	gi 474429569	ABC transporter I family member 1 [<i>Cajanus cajan</i>]
84	164,6	1446	8,24	23	K.IAVVGRTGSGKSTLVQALFR	gi 922337059	Multidrug resistance protein ABC transport family protein [<i>Medicago truncatula</i>]
85	31,9	303	8,7	23	R.SRPVSDTMAALMAK	gi 413920278	Benzoxazinless1 [<i>Zea mays</i>]
86	35,4	312	9,30	23	K.TGLTDAIQTGI AKLNK	gi 987169608	Acetyl-CoA carboxylase carboxyltransferase beta subunit (chloroplast) [<i>Tetraplodon fuegianus</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

87	74,8	664	5,5	23	R.MVQLGKSLDPSRLWDAASGWIDPQDR	gi 545360397	Glycoside hydrolase [<i>Coccomyxa subelli psoidean C.169</i>]
88	74,4	678	9,7	23	R.SASPPHVTEPR	gi 971520561	Cyclic nucleotid-binding domain containing protein [<i>Klebsormidium flaccidum</i>]
89	124,2	1220	4,7	23	R.FTDCEFSSQGS LGQALVTKSALIVAGDR.S	gi 926785032	Outer membrane adhesion like protein [<i>Monoraphidium neglectum</i>]
90	77,5	684	5,36	23	K.LAVVLGLINTSPGQPLQVIK.N	gi 922353943	Pentatricopeptide (PPR) repeat protein [<i>Medicago truncatula</i>]
91	52,77	476	5,51	23	R.GTTLFDLTVADLK.T	gi 545364742	Argininosuccinate lyase [<i>Coccomyxa subellipsoidea C-169</i>]
92	34,7	315	8,55	22	R.GDPDPTMDPALDTKLVK	gi 357508881	Peroxidase Family protein [<i>Medicago truncatula</i>]
93	87	777	8,17	22	K.KDGLIAMIGDGINDAPALVK.K	gi 734325138	Cadmium/zinc-transporting ATPase 3 [<i>Glycine soja</i>]
94	61,7	515	9,88	22	R.ALAFPLXELHR.E	gi 294485969	Maturase K (Chloroplast) [<i>Gavilea litoralis</i>]
95	52	467	7,47	22	K.GSDWIVEMK.K	gi 545372182	NADH-ubiquinone oxidoreductase [<i>Coccomyxa subellipsoidea C-169</i>]
96	91,3	833	7,54	22	K.CDSPGVDR LGGLPDDVLGR.I	gi 475503908	Medium-chain-fatty-acid-CoA ligase [<i>Aegilops taishii</i>]
97	41,5	388	5,2	22	R.TPVATAVLLDR.R	gi 475614932	Peroxidase 70 [<i>Aegilops taishii</i>]
98	87,2	767	5,72	22	R.KPSMTRMLPEQAVGPPFFYFQNVAR.A	gi 475545788	DNA (cytosine -5) –methyltransferase DRMZ [<i>Aegilops taishii</i>]
99	45,8	403	4,8	22	K.FDMSASDKIAFG EK.M	gi 145354669	TRP-containing protein [<i>Ostreococcus lucimarinus CCE9901</i>]
100	127,7	1,087	5,86	22	R.AKEDILLFFR.	gi 728848320	Ubiquitin Carboxyl-terminal hydrolase 13-like protein [<i>Gossypium arboreun</i>]
101	65,2	571	9,29	22	R.AGAFYK.I	gi 475581696	Disease resistance protein RGA2 [<i>Aegilops tauschii</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

102	73,0	643	5,91	22	K.EKAMEELLVQK.L	gi 728829371	FAST kinase domain-containing 3 [<i>Gossypium arboreum</i>]
103	34,1	316	9,12	22	R.AHGERMAEMR.	gi 760440573	Nitric oxide synthase-interacting protein like protein [<i>Auxenochlorella protothecoids</i>]
104	62,3	547	8,08	22	R.RTGVLAPTSVVVKQMTLAR.	gi 357502015	Inactive poly [ADP-ribose] polymerase [<i>Medicago truncatula</i>]
105	96,4	835	7,46	22	R.GPSVLPKSLRVLK.W	gi 1012354712	TMV resistance protein N [<i>Cajanus cajan</i>]
106	58,0	531	9,07	22	K.IESNTSTNNINRPPAAK	gi 703134384	DNA polymerase delta subunit 3 [<i>Morus notabilis</i>]
107	36,6	347	9,05	22	R.SRPVSDTMAALMAK.G	gi 162461896	Indole-3-glycerol phosphate lyase, chloroplastic [<i>Zea mays</i>]
108	55,3	510	5,12	22	R.DLHRLVDDLLMEAAAK	gi 187373014	UDP-glycosyltransferase UGT705A4 [<i>Avena strigosa</i>]
109	16,84	141	8,72	22	K.SLMFTHDLTSVLFNSRTPPIPLWPR.V	gi 339787214	S-RNase, partial [<i>Coffea eugenioids</i>]
110	37,75	315	8,55	22	R.GKPDPTMDPALDTKLVK.L	gi 357508881	Peroxidase family protein [<i>Medicago truncatula</i>]
111	129,6	1.173	5,53	21	R.ILGFAWESGK.I	gi 15237426	LRR receptor-like serine/threonine-protein kinase FLS2 [<i>Arabidopsis thaliana</i>]
112	49,75	446	9,6	21	K.APPPVVVMPPAR.G	gi 674944880	BnaC06g37090D [<i>Brassica napus</i>]
113	61,3	532	4,93	21	MALFSSEDDASENR.I	gi 612385172	APG AC_XENLA cysteine protease APG4C [<i>Bathy cocais prasinos</i>]
114	62,35	588	5,06	21	K.ALLQDIAIMTGSEFQAK.	gi 901822509	60 KDa chaperonin2 [<i>Zostera marina</i>]
115	31,7	272	8,71	21	K.CGLIQGIEIRPK.	gi 15226939	F-box protein PP2-B10 [<i>Arabidopsis thaliana</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF. (Continua)

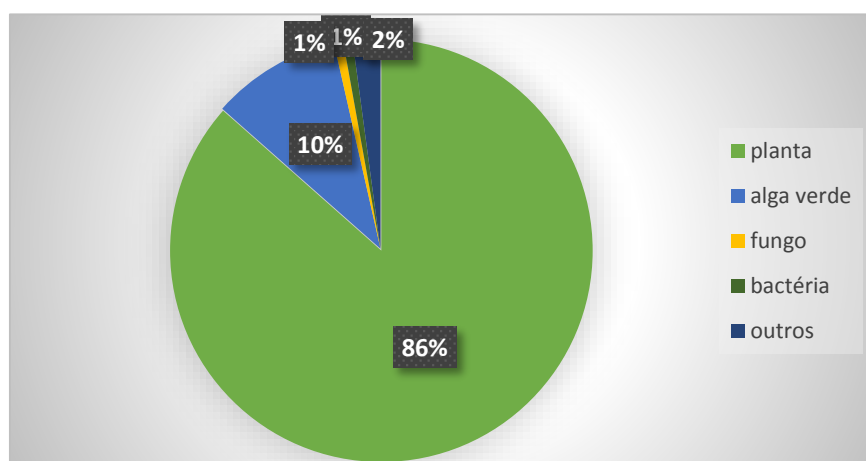
116	29,1	266	5,35	21	R.VCEASRIIGIDLNPNR.F	gi 2687444	Alcohol dehydro-genase, partial [<i>Gossypium robinsonii</i>]
117	100,4	939	4,7	21	K.IAGVNKQADKLTGGLDGIYK	gi 302837698	Ferroportin [<i>Volvox carteri</i>]
118	27,49	259	4,68	21	R.ALSATSGNVHAAVER.L	gi 304273284	ubiquitin-like protein 1, partial [<i>Glandiolus Glandiflorus</i>]
119	43,1	371	9,05	21	K.GAFFELGATVCPIAIK.Y	gi 156628062	Phospholipid/glycerol acyltransferase [<i>Helianthus annuus</i>]
120	31,6	278	6,09	21	R.REDGWMEVEVGEFLSGELEG VVK	gi 1012363201	F-box protein PP2-B15 [<i>Cajanus cajan</i>]
121	47,8	424	9,47	21	MERIVGGSYKLGR	gi 922330299	Casein kinase I-like protein [<i>Medicago truncatula</i>]
122	158	1470	5,81	21	R.FACETATSTER.R	gi 760450235	Ferredoxin dependent glutamate synthase, chloroplast [<i>Auxenochlorella protothecoides</i>]
123	19,3	176	5,8	21	K.FMMKDFTLRLHDTK.K	gi 729258063	DNA mismatch repair family, protein, partial [<i>Salix arbutifolia</i>]
124	13,8	121	3,87	21	K.SELHIILSNFGLI.R	gi 728807591	NADH-quinone oxidoreductase subunit D [<i>Gossypium arboreum</i>]
126	45,2	410	6,1	21	R.LGANIGSXQGPTGLGK.Y	gi 365823900	Photosystem II CP43, (Chloroplast) [<i>Zostera angustifolia</i>]
127	59	500	9,52	21	K.SIFASKGTFLINK.W	gi 909720495	Maturase K, partial, (Chloroplast) [<i>Brunonia australis</i>]
128	45,62	398	6,79	21	MRQFGPDFGGVTK.K	gi 255670779	Os02g0261400 [<i>Oriza sativa Japonica Group</i>]
129	18,71	165	9,8	21	R.KVDMLEGVSVXXPSEK.V	gi 817490197	Ribossomal protein L9, partial [<i>Sonneratis caseolis</i>]
130	56,8	492	9,03	21	K.VFMFGNPSVIVTSPEACR.R	gi 333394171	ent-kaurenoic acid oxidase [<i>Castanea mollissima</i>]
131	57,32	490	8,39	21	R.CFTKVTGLKFLTR.L	gi 590647948	Aminopeptidase isoform 1 [<i>Theobroma cacao</i>]

Tabela 2. Lista de peptídeos e proteínas identificadas em látex de *Himatanthus sucuuba* por ESI-QUAD-TOF.

132	73,3	632	6,2	21	K.ELSKGFQWQGRIADHEK	gi 157783541	Dehydration-response family protein S51 [<i>Brassica rapa</i>]
133	56,9	515	7,04	21	K.NFTAASMKSVQILR.K	gi 357443115	DUF 4283 domain protein [<i>Medicago truncatula</i>]
134	115,6	970	9,79	21	R.LRMEREEAMR.K	gi 922400551	TPX2-like protein [<i>Medicago truncatula</i>]
135	54,8	482	5,85	21	K.RIMIGDEAEEMRV	gi 1039217160	UGT73AL1 [<i>Punica granatum</i>]
136	29,1	251	4,37	20	K.QAFYR.	gi 124361102	Terpenoid synthase [<i>Medicago truncatula</i>]

Verificou-se que 86% das proteínas caracterizadas são sequências similares a proteínas encontradas em plantas (monocotiledôneas e/ou dicotiledôneas), 10% das sequências similares a proteínas identificadas em diferentes espécies de alga verde como *Auxenochlorella protothecoides* e *Dunaliella bioculata*), Figura 13.

Figura 13. Classificação das proteínas caracterizadas de acordo com o organismo identificado.



7.6 Perfil de proteínas

Das 1471 proteínas presentes no látex de *H. sucuuba*, apenas 128 proteínas estão caracterizadas, isso pode ser em função da especificidade da amostra (soro do látex) e ao número reduzido de proteínas provenientes de látex disponíveis nos bancos de dados, provavelmente pela falta de informação acerca do genoma dessa planta.

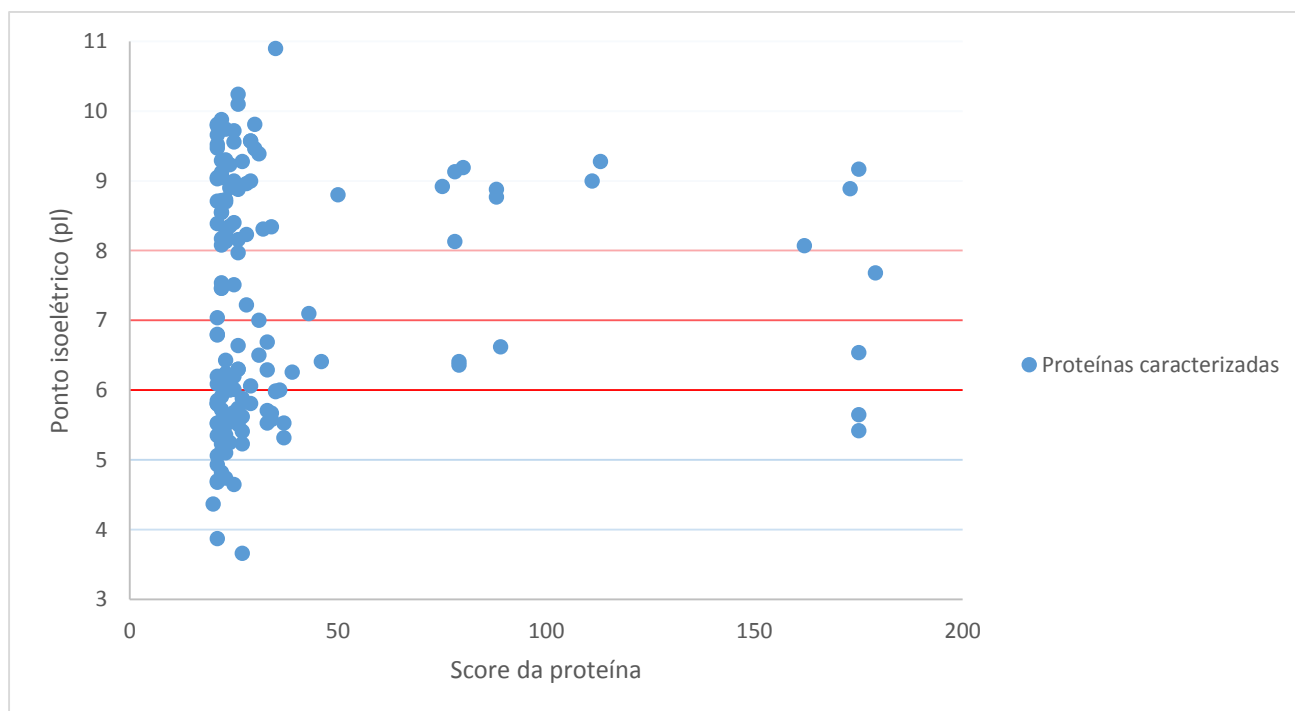
O látex em geral apresenta um conjunto de proteínas bastante diversificadas e por consequência, diferem em relação ao seu peso molecular. No látex de *H. sucuuba* as proteínas identificadas apresentaram peso molecular entre 13,8 e 164,6 KDa, similarmente em relação as proteínas encontradas em *H. brasiliensis*, que apresentou proteínas com peso molecular variando de 3,9 a 194,2 KDa (DAI et al., 2013). Em outro estudo com *H. brasiliensis*, identificou-se proteínas com peso molecular entre 12 e 44 KDa (WANG et al., 2010), diferentemente do observado para outras espécies laticíferas, com proteínas entre 20 e 100 KDa para *Thevetia peruviana* (FREITAS et al. 2016), entre 10,1 e 68,4 KDa para

Chelidonium majus (NAWROT et al., 2007), de 12,7 a 71,1 KDa para *Papaver somniferum* (DECKER et al, 2000) e de 4,7 a 367 KDa para as proteínas de *Lactuca sativa* (CHO et al., 2009).

7.7 Pontos isoelétricos

A diversidade de proteínas se reflete também em relação ao ponto isoelétrico das mesmas. Em *H. sucuuba* 32,64 % das encontradas se caracterizam como básicas ($pI > 8$), 24,82 % são neutras ($6 \geq pI \geq 8$) e 60 proteínas, o que corresponde a mais de 40% do total são caracterizadas como ácidas ($pI < 6$) como pode ser observado na Figura 14.

Figura 14. Gráfico de dispersão para a relação de Score e pI das proteínas obtidas pela espectrometria de massas e identificadas pelo programa Mascot.



Verificou-se uma quantidade superior de proteínas ácidas no látex de *H. sucuuba*, diferentemente do observado em *Lactuca sativa*, com proteínas apresentando ponto isoelétrico variando entre 5,8 e 13, sendo o número de proteínas básicas ($pI > 8$) 72 % do total (CHO et al., 2009). Já no estudo de DECKER et al. (2000), trabalhando com *Papaver somniferum*, somente 5 proteínas das 98 identificadas são básicas (com pI entre 8-9).

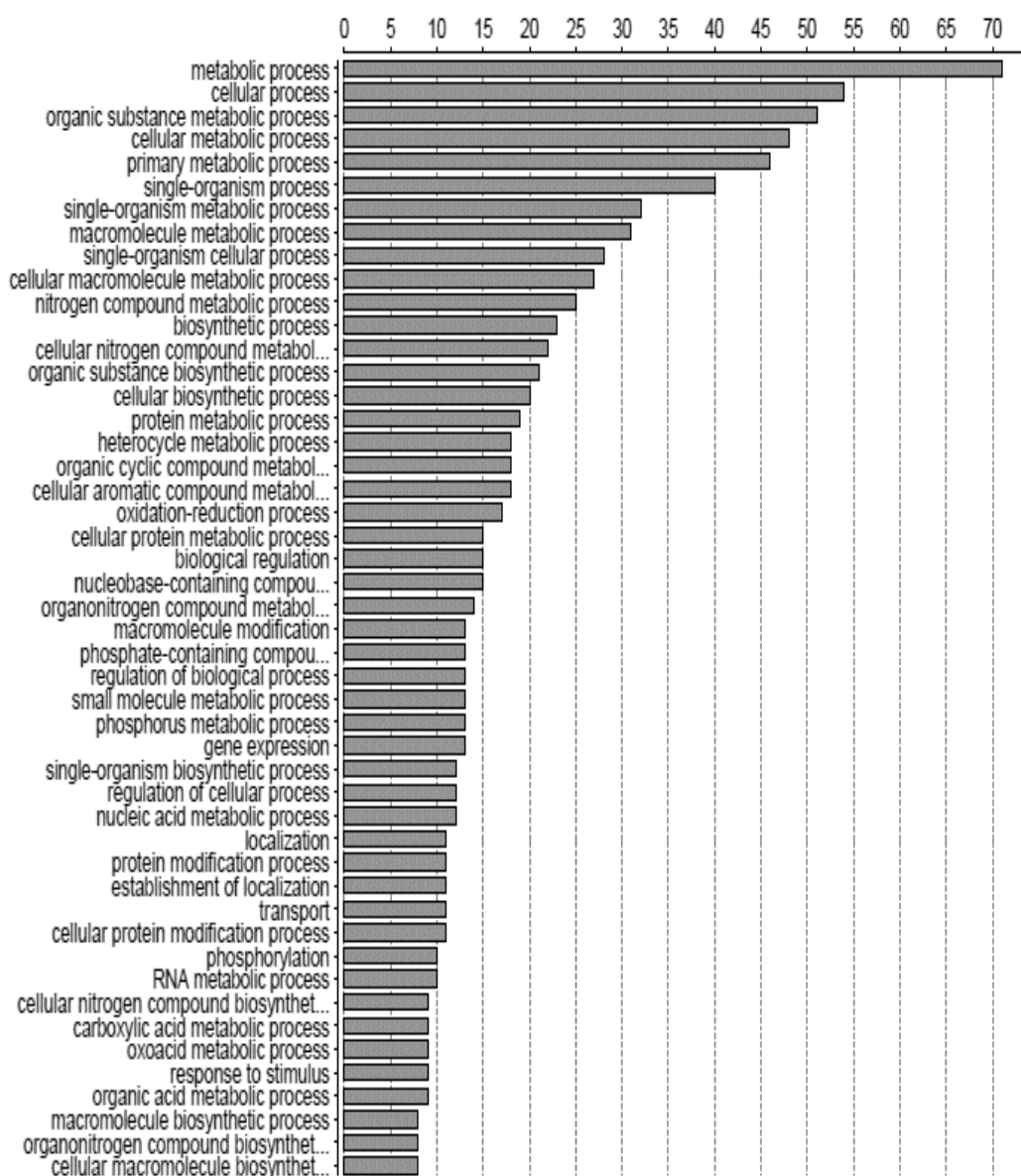
7.8 Anotação funcional das proteínas identificadas pelo programa Blast2Go.

Segundo (HOU, 2017), a anotação funcional das proteínas é vital para pesquisas biológicas, clínicas e outras aplicações devido as funções importantes que as proteínas desempenham em vários processos biológicos. Embora as funções de algumas proteínas tenham sido anotadas através de experiências biológicas, ainda há muitas proteínas cujas funções ainda não foram anotadas devido às limitações dos métodos existentes e ao alto custo.

A anotação funcional das proteínas de *H. succuba* foi realizada com o programa Blast2Go. A análise pelo programa Blast2Go consiste na anotação das sequências (associá-las às ontologias criadas pelo Gene Ontology Consortium). Essas ontologias visam padronizar a representação dos genes e de seus produtos para todos os sistemas biológicos, subdividindo-os nas categorias de *Biological process* (refere-se à atividade biológica que o gene ou produto contribui), *Molecular function* (atividade bioquímica do gene ou seu produto) e *Cellular component* (local na célula onde o gene ou seu produto é ativo) (dos SANTOS, 2013).

Das 136 sequências de proteínas (fasta), submetidas ao Blast2GO, 101 foram identificadas funcionalmente. As anotações funcionais das proteínas do soro de *H. succuba* foram realizadas da seguinte forma: 1- File load sequence; 2- Run Blast; 3- Run Mapping; 4- Run annotation; 5-Validate Annotation; 6-Run InterProScan; 7-Merge InterProScan Gos to Annotation; 8-Run GO-Slim; 9- Run GO-EnzymeCode Mapping e 10- Load Pathway Maps from KEGG. Com relação ao processo biológico em que as proteínas contribuem, mais de 70% atuam em processos metabólicos, e mais de 50% dos produtos identificados participam de processos celulares de modo geral. Outros processos em que há contribuição por parte das proteínas identificadas em *H. succuba* são processos biossintéticos, de oxidorredução, regulação biológica, expressão gênica, localização e transporte, entre outros (Figura 15).

Figura 15. Anotação funcional para Processo Biológico das proteínas identificadas em *H. succuba*.

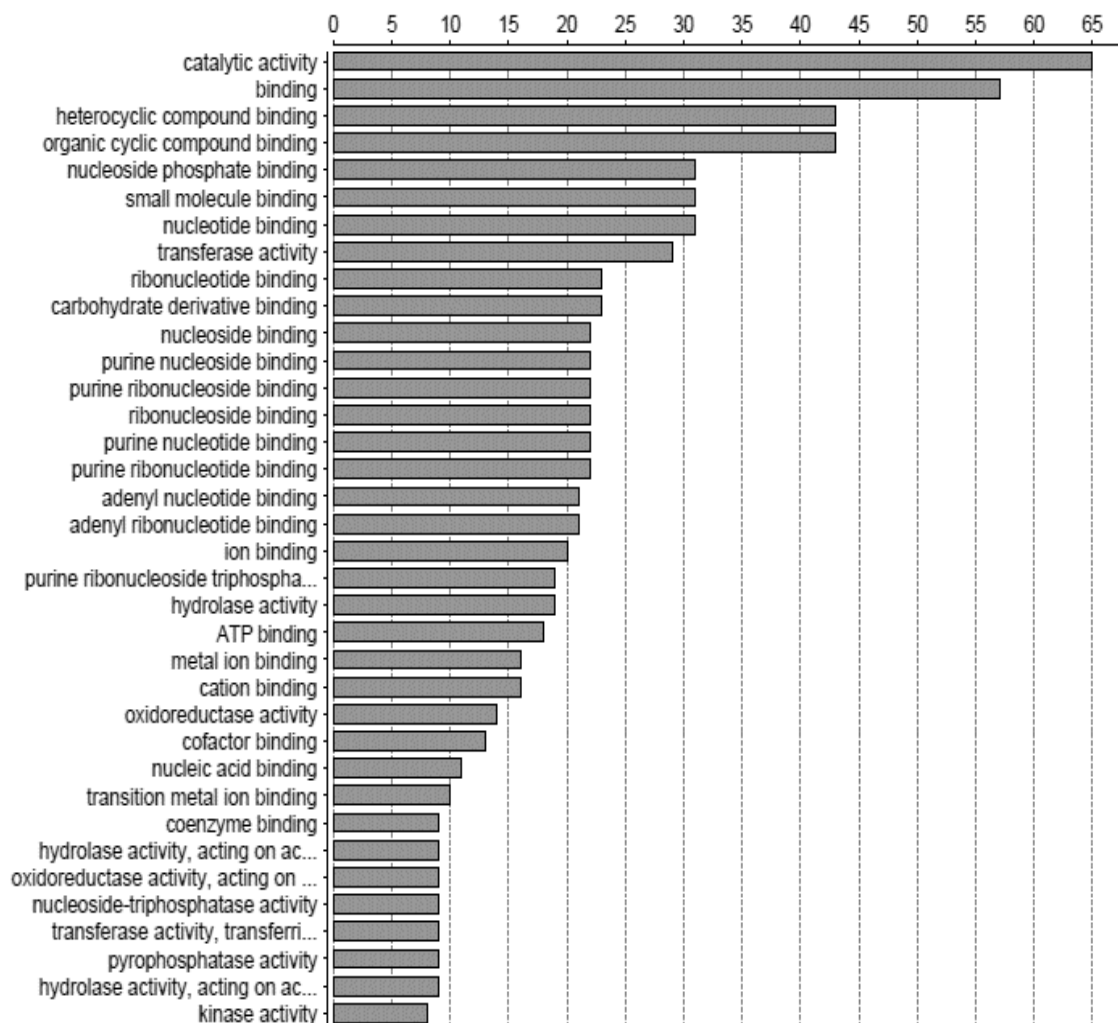


No trabalho realizado por Wahler e colaboradores (2012), no látex de *Taraxacum brevicorniculatum*, as proteínas identificadas estão associadas a funcionalidades como processos celulares básicos (31,4 %), transporte (10,4 %), processo metabólico de carboidrato (5,5 %) entre outros. Para o proteoma do látex de *Lactuca sativa*, mais da metade das proteínas identificadas estão agrupadas em cinco classes funcionais de proteínas: relacionadas ao metabolismo (20 %), resgate de células e defesa (12 %), função de ligação (binding) (10 %), transporte celular (8 %) (CHO et al., 2009). No látex de *H.*

brasiliensis, os processos relatados a partir da identificação das proteínas incluem: metabolismo de carboidrato, lipídeo, processo metabólico de oxidação-redução entre outros (HAVANAPAN et al., 2016).

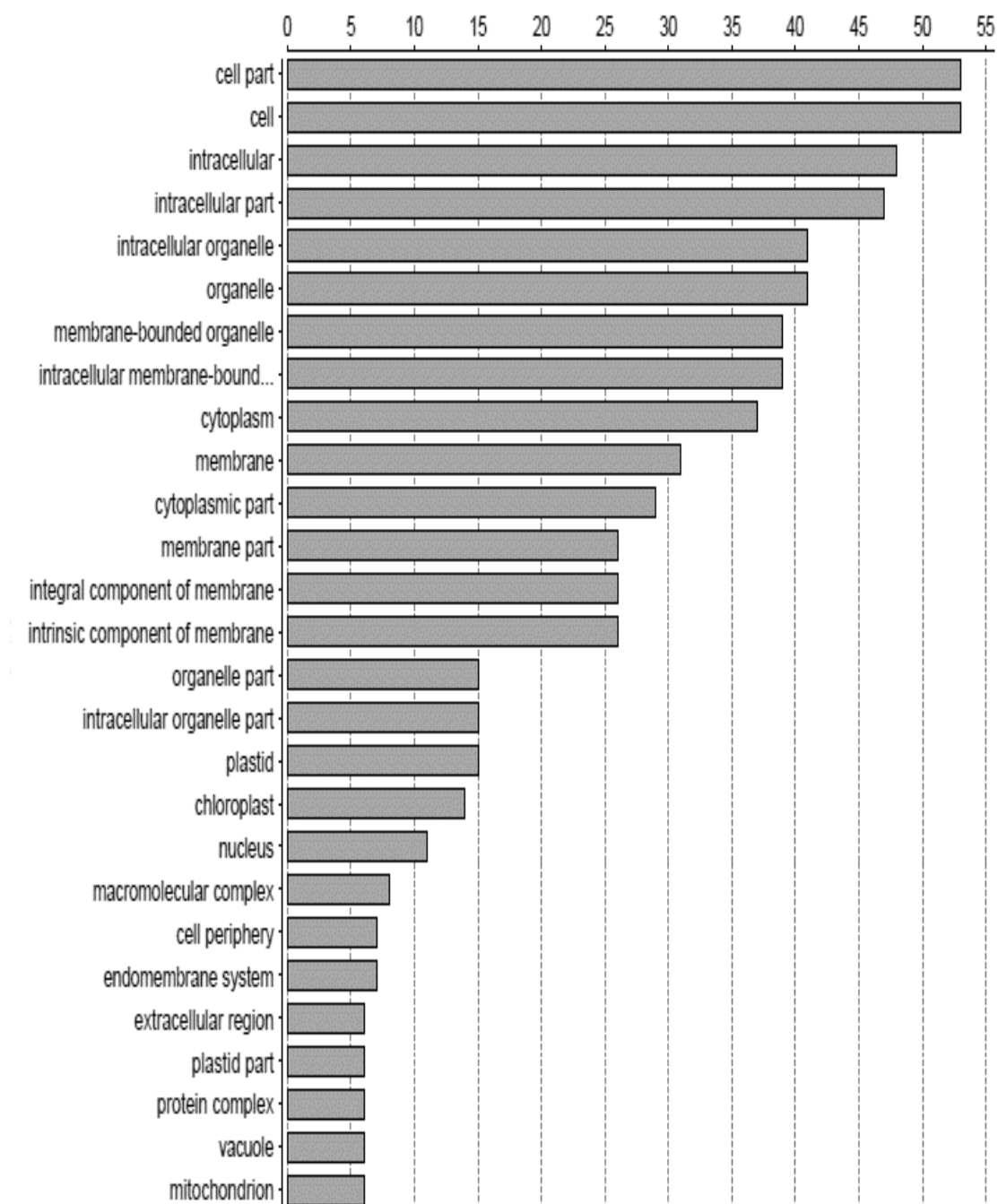
Em relação a anotação realizada para função molecular das proteínas, verificou-se que mais de 50% estão relacionadas a função de atividade catalítica e/ou ligação (binding). Outras funções identificadas são de ligação de pequenas moléculas, atividade transferase, atividade hidrolase, ligação ATP, atividade oxidoreductase, entre outras (Figura 16).

Figura 16. Anotação funcional para função molecular das proteínas identificadas em *H. sucuuba*.



Para os componentes celulares identificados, as proteínas de modo geral são ativas em parte da célula ou na célula como um todo (mais de 50 %), ou localizadas em outras regiões, sendo ativas na região intracelular, membrana, citoplasma, plastídio, cloroplasto, núcleo, vacúolo, mitocôndria, entre outras (Figura 17).

Figura 17. Anotação funcional para componente celular das proteínas identificadas em *H. sucuba*.

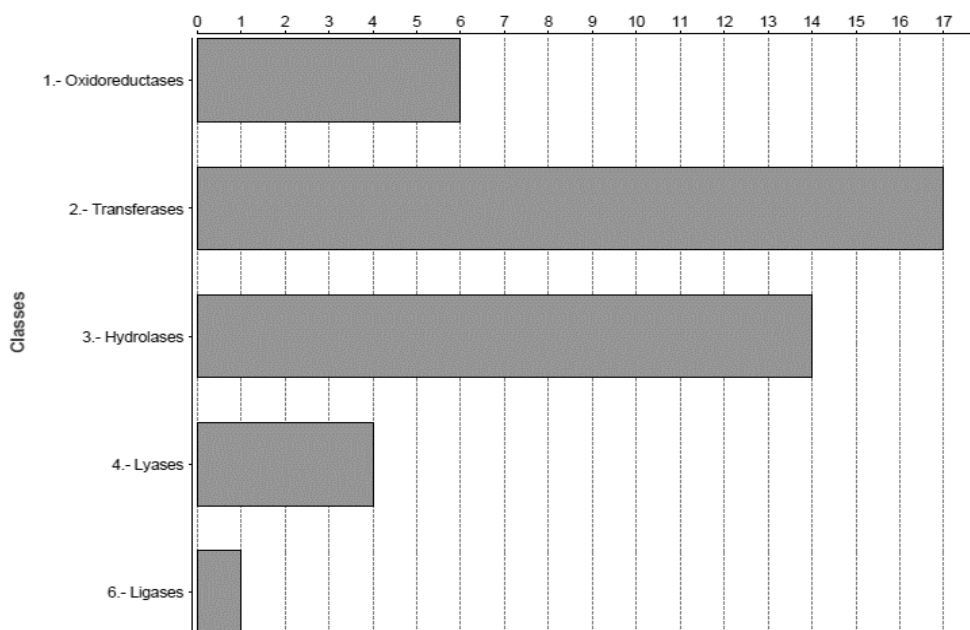


7.9 Classificação das enzimas realizadas pelo Blast2GO

A International Union of Biochemistry and Molecular Biology (IUBMB), classificou as enzimas conforme o tipo de reação que as enzimas catalisam. Cada enzima descrita recebe um número de classificação, conhecida por EC (*Enzyme Commission*), composto por 4 dígitos (EC: x.x.x.x), separados por pontos referente a: 1- classe; 2- subclasse; 3- grupos químicos específicos que participam de cada reação e 4- nome indicado a referida enzima. E são classificadas em seis grandes classes: 1 EC– Oxidoreductases; 2 EC – Transferases; 3 EC– Hidrolases; 4 EC – Liases; 5 EC – Isomerases; 6 EC – Ligases (CORNISH-BOWDEN, 2014; CUESTA et al., 2015; LEHNINGER et al., 2011).

Foram analisadas pelo programa Blast2GO 136 proteínas de *H. sucuuba* das quais 42 proteínas possuem alguma atividade enzimática (Figura 18). As enzimas mais frequentes foram transferases, seguida das hidrolases, oxidoreductases, liases e ligases. Não foram identificadas enzimas do tipo isomerases.

Figura 18. Classificação geral das enzimas encontradas no látex de *H. sucuuba*. As enzimas foram categorizadas utilizando-se o programa Blast2GO, de acordo com o banco de dados *Kyoto Encyclopedia of genes and Genomes pathway database* disponíveis em KEGG.



7.9.1 Oxidorredutases

As enzimas responsáveis pela transferência de elétrons para um aceptor, catalisando as reações de oxirredução são denominadas EC1. O grupo de enzimas oxidorredutases dividi-se em 24 subcategorias: EC: 1,1 à EC 1.23 e EC 1.97, incluindo outros tipos de oxidorredutases (Tabela 3).

Tabela 3. Subclasses das enzimas oxidorredutases identificadas no soro do látex de *H. succuba*. Categorizadas de acordo com os seus números de EC (*Enzyme commission*), com base nos dados obtidos no *Kyoto Encyclopedia of genes and Genomes pathway database* (KEGG) pelo software Blast2GO.

<i>Número EC (Enzyme Commission) - Subclasse</i>	<i>Nome das enzimas</i>	<i>Nº de enzimas</i>
<i>EC 1.1.1.1 - Atuam no grupo CH-OH de doadores</i>	<i>alcohol dehydrogenase</i>	1
<i>EC 1.1.1.49 - Atuam no grupo CH-OH de doadores</i>	<i>(glucose-6-phosphate dehydrogenase (NADP⁺))</i>	1
<i>EC 1.1.3.20 - Atuam no grupo CH-OH de doadores</i>	<i>long-chain-alcohol oxidase</i>	1
<i>EC 1.3.3.8 - Atuam no grupo CH-CH dos doadores</i>	<i>tetrahydroberberine oxidase</i>	1
<i>EC 1.4.1.2- Atuam no grupo CH-CH de doadores</i>	<i>glutamate dehydrogenase</i>	1
<i>EC 1.6.5.3- Atuam no NADH ou NADPH</i>	<i>NADH:ubiquinone reductase (H⁺-translocating)</i>	1

7.9.2 Transferases

São enzimas que catalisam, a transferência de grupos entre moléculas doadoras e aceptoras, classificadas como EC 2, sendo subdivididas em 10 subcategorias de EC 2.1 à EC2.10 (Tabela 4).

Tabela 4. Subclasses das enzimas transferases identificadas no soro do látex de *H. sucuuba*. Categorizadas de acordo com os seus números de EC (*Enzyme comission*), com base nos dados obtidos no *Kyoto Encyclopedia of genes and Genomes pathway database* (KEGG) pelo software Blast2GO.

<i>Número EC (Enzyme Comission) - Subclasse</i>	<i>Nome das enzimas</i>	<i>Nº de enzimas</i>
<i>EC 2.1.1.1- Transferem grupos de um carbono</i>	nicotinamide <i>N</i> -methyltransferase	1
<i>EC 2.1.1.128- Transferem grupos de um carbono</i>	(<i>RS</i>)-norcochlorine 6- <i>O</i> -methyltransferase	1
<i>EC 2.1.1.150- Transferem grupos de um carbono</i>	isoflavone 7- <i>O</i> -methyltransferase	1
<i>EC 2.1.2.1- Transferem grupos de um carbono</i>	glycine hydroxymethyltransferase	1
<i>EC 2.3.1.1- Aciltransferases</i>	amino-acid <i>N</i> -acetyltransferase	1
<i>EC 2.3.1.15- Aciltransferases</i>	glycerol-3-phosphate 1- <i>O</i> -acyltransferase	1
<i>EC 2.3.1.43- Aciltransferases</i>	phosphatidylcholine—sterol <i>O</i> -acyltransferase	1
<i>EC 2.4.1.1- Glicosiltransferases</i>	glycogen phosphorylase	1
<i>EC 2.4.2.30- Glicosiltransferases</i>	<i>NAD</i> ⁺ ADP-ribosyltransferase	1
<i>EC 2.5.1.18- Transferem grupos alquil ou aril, excluindo grupos metil</i>	glutathione transferase	1
<i>EC 2.5.1.27- Transferem grupos alquil ou aril, excluindo grupos metil</i>	adenylate dimethylallyltransferase	1
<i>EC 2.5.1.29- Transferem grupos alquil ou aril, excluindo grupos metil</i>	geranylgeranyl diphosphate synthase	1
<i>EC 2.5.1.75- Transferem grupos alquil ou aril, excluindo grupos metil</i>	tRNA dimethylallyltransferase	1
<i>EC 2.7.3.1- Transferem grupos que contenham fosforo</i>	guanidinoacetate kinase	1
<i>EC 2.7.11- Transferem grupos que contenham fosforo</i>	non-specific serine/threonine protein kinase	1
<i>EC 2.7.10.1-Transferem grupos que contenham fosforo</i>	receptor protein-tyrosine kinase	1
<i>EC 2.7.7.27- Transferem grupos que contenham fosforo</i>	glucose-1-phosphate denyltransferase	1

7.9.3 Hidrolases

São enzimas que catalisam reações de hidrólise, com transferência de grupos funcionais para a água, classificadas como EC 3. Estas enzimas são subdivididas em 13 subcategorias (EC 3.1 à EC 3.13) (Tabela 5).

Tabela 5. Subclasses das enzimas hidrolases identificadas no soro do látex de *H. sucuuba*. Categorizadas de acordo com os seus números de EC (*Enzyme comission*), com base nos dados obtidos no *Kyoto Encyclopedia of genes and Genomes pathway database* (KEGG) pelo software Blast2GO.

<i>Número EC (Enzyme Comission) - Subclasse</i>	<i>Nome das enzimas</i>	<i>Nº de enzimas</i>
<i>EC 3.1.3.12- Atuam em ligações éster</i>	<i>trehalose-phosphatase</i>	1
<i>EC 3.1.3.16- Atuam em ligações éster</i>	<i>protein-serine/threonine phosphatase</i>	1
<i>EC 3.1.3.26.4- Atuam em ligações éster</i>	<i>4-phytase</i>	1
<i>EC 3.4.11- Atuam em ligações peptídicas</i>	<i>leucyl aminopeptidase</i>	2
<i>EC 3.4.21- Atuam em ligações peptídicas</i>	<i>Serine endopeptidases</i>	3
<i>EC 3.4.22- Atuam em ligações peptídicas</i>	<i>Cysteine endopeptidases</i>	1
<i>EC 3.5.4.5- Atuam em ligações nitrogênio-carbono (ligações não peptídicas)</i>	<i>cytidine deaminase</i>	1
<i>EC 3.6.1.3- Atuam em anidridos ácidos</i>	<i>adenosinetriphosphatase</i>	1
<i>EC 3.6.3.3- Atuam em anidridos ácidos</i>	<i>Cd²⁺-exporting ATPase</i>	1
<i>EC 3.6.4.3- Atuam em anidridos ácidos</i>	<i>microtubule-severing ATPase</i>	1
<i>EC 3.6.3.8- Atuam em anidridos ácidos</i>	<i>Ca²⁺-transporting ATPase</i>	1

7.9.4 Liases

São enzimas que catalisam a clivagem de ligações, através hidrólise ou oxidação (EC 4) e são compostas por oito subcategorias de EC 4.1 à EC 4.7 e EC 4.99 que são outros tipos de liases (Tabela 6).

Tabela 6. Subclasses das enzimas liases identificadas no soro do látex de *H. succuba*. Categorizadas de acordo com os seus números de EC (*Enzyme comission*), com base nos dados obtidos no *Kyoto Encyclopedia of genes and Genomes pathway database* (KEGG) pelo software Blast2GO.

<i>Número EC (Enzyme Comission) – Subclasse</i>	<i>Nome das enzimas</i>	<i>Nº de enzimas</i>
<i>EC 4.1.2.8- Liasas carbono-carbono</i>	<i>indole-3-glycerol-phosphate lyase</i>	1
<i>EC 4.2.1.20- Liasas carbono-oxigênio</i>	<i>tryptophan synthase</i>	1
<i>EC 4.2.3.8- Liasas carbono-oxigênio</i>	<i>casbene synthase</i>	1
<i>EC 4.3.2.1- Liasas carbono-nitrogênio</i>	<i>Arginino succinate lyase</i>	1

7.9.5 Ligases

São enzimas que catalisam a formação de ligações químicas por reação de condensação acopladas à hidrólise de ATP (EC 6). Estas enzimas são subdivididas em seis subclasses de EC 6.1 à EC 6.6. Identificamos uma enzima Ligase, a EC 6.4.1.2 (*acetyl-CoA carboxilase*), que forma ligações carbono-carbono.

7.10 Vias metabólicas

As enzimas identificadas no soro do látex de *H. succuba* pelo Blast2GO, em conjunto com a base de dados do KEGG, resultaram na predição no total de 30 vias metabólicas, incluindo informações acerca das enzimas (proteínas) que participam de cada via metabólica e que foram identificadas no látex de *H. succuba*, bem como, o substrato e produto da reação catalisada por cada enzima (Tabela 7).

Tabela 7. Vias metabólicas das enzimas encontradas no soro do látex de *H. succuba*, preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (*Enzymes Commission*). (Continua)

Vias metabólicas	Número EC (<i>Enzyme Commission</i>)	Nome das enzimas	Substrato	Produto
Glicólise/gliconeogênese	1.1.1.1	<i>Alcohol dehydrogenase</i>	(1) um álcool primário + NAD ⁺ (2) um álcool secundário + NAD ⁺	(1) um aldeído + NADH + H ⁺ (2) uma cetona + NADH + H ⁺
Degradação de naftaleno	1.1.1.1	<i>Álcohol dehydrogenase</i>	(1) um álcool primário + NAD ⁺ (2) um álcool secundário + NAD ⁺	(1) um aldeído + NADH + H ⁺ (2) uma cetona + NADH + H ⁺
Fosforilação oxidativa	1.6.5.3	<i>NADH:ubiquinone reductase (H⁺-translocating)</i>	NADH + ubiquinone + 5 H ⁺ _{in}	NAD ⁺ + ubiquinol + 4 H ⁺ _{out}
Metabolismo do piruvato	6.4.1.2	<i>acetyl-CoA carboxylase</i>	ATP + acetil-CoA + HCO ₃ ⁻	ADP + fosfato + malonil-CoA
Metabolismo da tirosina	1.1.1.1	<i>Álcohol dehydrogenase</i>	(1) um álcool primário + NAD ⁺ (2) um álcool secundário + NAD ⁺	(1) um aldeído + NADH + H ⁺ (2) uma cetona + NADH + H ⁺
Metabolismo de glioxilato e dicarboxilato	2.1.2.1	<i>Glycine methyltransferase hydrox</i>	5,10-methylenetetrahydrofolate + glycine + H ₂ O	tetrahydrofolate + L-serine
Metabolismo do ácido cianoamino	2.1.2.1	<i>Glycine methyltransferase hydrox</i>	5,10-methylenetetrahydrofolate + glycine + H ₂ O	tetrahydrofolate + L-serine
Reservatório de carbono por folato	2.1.2.1	<i>Glycine methyltransferase hydrox</i>	5,10-methylenetetrahydrofolate + glycine + H ₂ O	tetrahydrofolate + L-serine

Tabela 7 – Vias metabólicas das enzimas encontradas no soro do látex de *H. sucuuba*, preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (*Enzymes Commission*). (Continua)

Metabolismo da glicina, serina e Treonina	4.2.1.20;	tryptophan synthase	L-serine + 1-C-(indol-3-yl)glycerol 3-phosphate	L-tryptophan + D-glyceraldehyde 3-phosphate + H ₂ O
	2.1.2.1;	<i>glycine hydrox methyltransferase</i>	5,10-methylenetetrahydrofolate + glycine + H ₂ O	tetrahydrofolate + L-serine
	1.1.1.1	<i>Álcohol dehydrogenase</i>	(1) um álcool primário + NAD ⁺ (2) um álcool secundário + NAD ⁺	(1) um aldeído + NADH + H ⁺ (2) uma cetona + NADH + H ⁺
Biossíntese de arginina	4.3.2.1	<i>argininosuccinate lyase</i>	2-(N ^o -L-arginino)succinate	fumarate + L-arginine
Biossíntese de fenilalanina, tirosina e triptofano	4.2.1.20	<i>tryptophan synthase</i>	L-serine + 1-C-(indol-3-yl)glycerol 3-phosphate	L-tryptophan + D-glyceraldehyde 3-phosphate + H ₂ O
Metabolismo do glicerol lipideo	2.3.1.15;	glycerol-3-phosphate 1-O-acyltransferase	acyl-CoA + <i>sn</i> -glycerol 3-phosphate	CoA + 1-acyl- <i>sn</i> -glycerol 3-phosphate
	2.3.1.158	phospholipid:diacylglycerol acyltransferase	phospholipid + 1,2-diacyl- <i>sn</i> -glycerol	lysophospholipid + triacylglycerol
Metabolismo do glicerolfosfolipideo	2.3.1.15;	glycerol-3-phosphate 1-O-acyltransferase	acyl-CoA + <i>sn</i> -glycerol 3-phosphate	CoA + 1-acyl- <i>sn</i> -glycerol 3-phosphate
	2.3.1.43	phosphatidylcholine—sterol O-acyltransferase	phosphatidylcholine + a sterol	1-acylglycerophosphocholine + a sterol ester

Tabela 7 – Vias metabólicas das enzimas encontradas no soro do látex de *H. sucuuba*, preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (*Enzymes Commission*). (Continua)

Biossíntese de alcaloides isoquinolínicos	1.3.3.8; 2.1.1.128	<i>tetrahydroberberine oxidase</i> <i>(RS)-norcoclaurine methyltransferase</i>	<i>6-O-</i> <i>S</i> -adenosyl-L-methionine + <i>(RS)-</i> norcoclaurine	berberine + 2 H ₂ O ₂ <i>S</i> -adenosyl-L-homocysteine + <i>(RS)-</i> coclaurine
Metabolismo da pirimidina	3.5.3.5	<i>formimidoylaspartate deiminase.</i>	<i>N</i> -formimidoyl-L-aspartate + H ₂ O	<i>N</i> -formyl-L-aspartate + NH ₃
Biossíntese do benzoxazinoide	4.1.2.8	<i>indole-3-glycerol-phosphate lyase</i>	<i>(1S,2R)</i> -1- <i>C</i> -(indol-3-yl)glycerol 3-phosphate	indole + D-glyceraldehyde 3-phosphate
Biossíntese de dipernoide	4.2.3.8	<i>casbene synthase</i>	geranylgeranyl diphosphate	casbene + diphosphate
Metabolismo de alanina, aspartato e Glutamato	4.3.2.1	<i>argininosuccinate lyase</i>	2-(<i>N</i> ^o -L-arginino)succinate	fumarate + L-arginine
Via da pentose fosfato	1.1.1.49	<i>glucose-6-phosphate dehydrogenase (NADP⁺)</i>	D-glucose 6-phosphate + NADP ⁺	6-phospho-D-glucono-1,5-lactone + NADPH + H ⁺
Biossíntese de ácidos graxos	6.4.1.2	<i>acetyl-CoA carboxylase</i>	ATP + acetyl-CoA + HCO ₃ ⁻	ADP + phosphate + malonyl-CoA
Metabolismo da glutatona	1.1.1.49; 2.5.1.18	<i>glucose-6-phosphate dehydrogenase (NADP⁺)</i> <i>glutathione transferase</i>	D-glucose 6-phosphate + NADP ⁺ RX + glutathione	6-phospho-D-glucono-1,5-lactone + NADPH + H ⁺ HX + R-S-glutathione
Metabolismo do propanoato	6.4.1.2	<i>acetyl-CoA carboxylase</i>	ATP + acetyl-CoA + HCO ₃ ⁻	ADP + phosphate + malonyl-CoA
Biossíntese de tetraciclina	6.4.1.2	<i>acetyl-CoA carboxylase</i>	ATP + acetyl-CoA + HCO ₃ ⁻	ADP + phosphate + malonyl-CoA
Metabolismo da purina	3.6.1.3	<i>adenosinetriphosphatase</i>	ATP + H ₂ O	ADP + phosphate
Metabolismo de drogas e outras enzimas	3.5.3.5	<i>cytidine deaminase</i>	cytidine + H ₂ O (2) 2'-deoxycytidine + H ₂ O	uridine + NH ₃ 2'-deoxyuridine + NH ₃

Tabela 7 – Vias metabólicas das enzimas encontradas no soro do látex de *H. sucuuba*, preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (*Enzymes Commission*).

(Continua)

Biossíntese de antiobióticos	4.2.1.20;	<i>tryptophan synthase</i>	L-serine + 1-C-(indol-3-yl)glycerol 3-phosphate	L-tryptophan + D-glyceraldehyde 3-phosphate + H ₂ O
	2.5.1.1;	<i>dimethylallyltranstransferase</i>	dimethylallyl diphosphate + isopentenyl diphosphate	diphosphate + geranyl diphosphate
	6.4.1.2;	<i>acetyl-CoA carboxylase</i>	ATP + acetyl-CoA + HCO ₃ ⁻	ADP + phosphate + malonyl-CoA
	2.1.2.1;	<i>glycine hydroxymethyltransferase</i>	5,10-methylenetetrahydrofolate + glycine + H ₂ O	tetrahydrofolate + L-serine
	1.1.1.1;	<i>alcool dehidrogenase</i>	(1) um álcool primário + NAD ⁺ (2) um álcool secundário alcool + NAD ⁺	um aldeido + NADH + H ⁺ uma cetona + NADH + H ⁺
	2.5.1.29;	<i>geranylgeranyl diphosphate synthase</i>	(2E,6E)-farnesyl diphosphate + isopentenyl diphosphate	diphosphate + geranylgeranyl diphosphate
	1.1.1.49;	<i>glucose-6-phosphate dehidrogenase (NADP⁺)</i>	D-glucose 6-phosphate + NADP ⁺	6-phospho-D-glucono-1,5-lactone + NADPH + H ⁺
	4.3.2.1	<i>argininosuccinate lyase</i>	2-(N ^o -L-arginino)succinate	fumarate + L-arginine

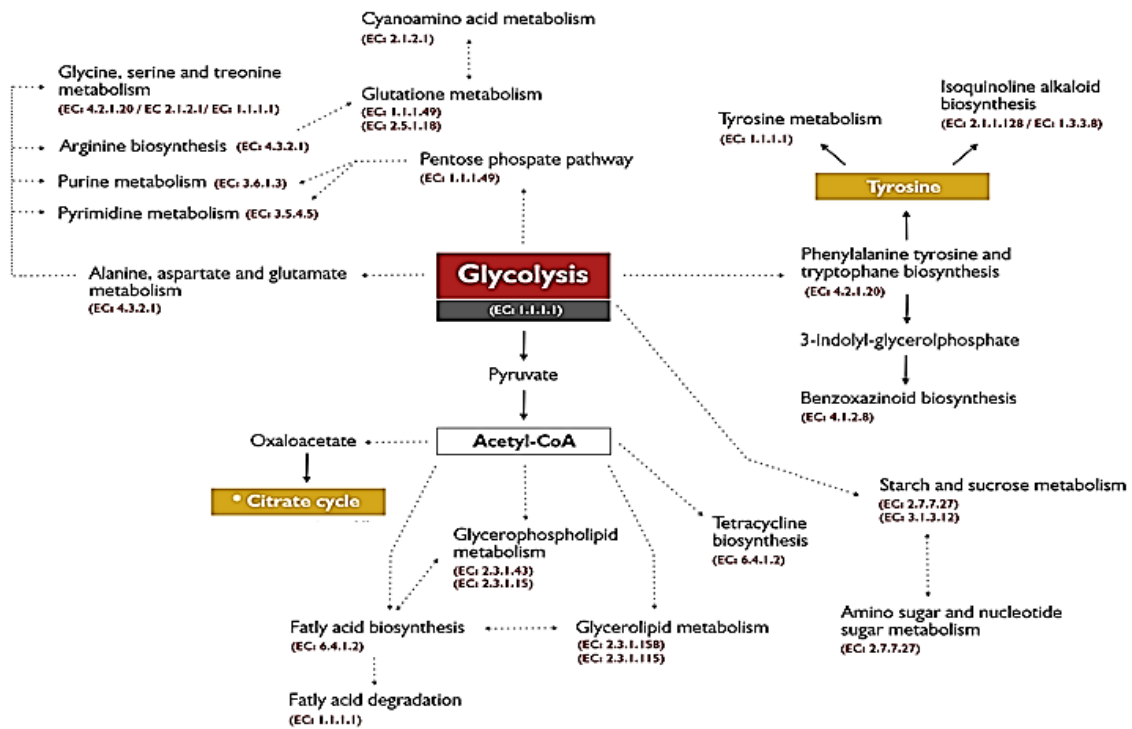
Tabela 7 – Vias metabólicas das enzimas encontradas no soro do látex de *H. sucuuba*, preditas pelo Blast2GO em conjunto com a base de dados KEGG pathway. Com seus respectivos números e nomes EC (*Enzymes Commission*).

Metabolismo de amido e sacarose	3.1.3.12;	trehalose-phosphatase	α,α -trehalose 6-phosphate + H ₂ O	α,α -trehalose + phosphate
	2.7.7.27	glucose-1-phosphate adenylyltransferase	ATP + α -D-glucose 1-phosphate	diphosphate + ADP-glucose
Biossíntese da zeatina	2.5.1.27;	<i>Adenylate dimethylallyltransferase (AMP-dependent).</i>	dimethylallyl diphosphate + AMP	diphosphate + N ⁶ -(dimethylallyl)adenosine 5'-phosphate
	2.5.1.75	<i>tRNA dimethylallyltransferase</i>	dimethylallyl diphosphate + AMP	diphosphate + N ⁶ -dimethylallyladenine ³⁷ in tRNA
Degradação de ácido graxo	1.1.1.1	<i>Álcool dehydrogenase</i>	(1) um álcool primário + NAD ⁺	(1)um aldeído + NADH + H ⁺
			(2) um álcool secundário + NAD ⁺	(2) uma cetona + NADH + H ⁺
Metabolismo de açúcares, aminoácidos e nucleotídeos	2.7.7.27	glucose-1-phosphate adenylyltransferase	ATP + α -D-glucose 1-phosphate	diphosphate + ADP-glucose

celular, que desempenha funções em uma variedade de processos celulares como vasodilatação através de nervos nitrígenos periféricos, neurotransmissão, agregação plaquetária, secreção de insulina. Agindo como molécula de defesa contra bactérias intercelulares e células tumorais. E está envolvido na angiogênese, inflamação e remodelação de tecidos (FORSTERMANN; SESSA, 2012; HERCULANO et. 2011).

Na Figura 20 apresenta-se um esquema com as principais enzimas e respectivas vias metabólicas identificadas a partir da análise (proteoma) do látex de *H. sucuuba*.

Figura 20. Resumo de algumas das vias metabólicas e enzimas previstas pelo Blast2GO, no soro do látex de *H. sucuuba*.



FONTE: Elaborada pelo autor.

Verificou-se ainda a presença da enzima *alcohol dehidrogenase* (EC:1.1.1.1) que atua como uma oxidoreductase na via metabólica da glicólise, que consiste do processo de conversão de glicose em piruvato e trata-se de um caminho central que produz importantes metabólitos precursores e também foi identificada em látex de *Lactuca sativa* (CHO et al., 2009). A *alcohol*

dehidrogenase é uma enzima encontrada em animais, leveduras e plantas. Em plantas esta enzima atua na via de produção de etanol em resposta a condições de hipoxia e anoxia. Principal adaptação ao stress anaeróbico em plantas (KIMMERER; STRIGER, 1998). A desidrogenase alcoólica (ADH) é uma das enzimas da via de fermentação etanólica que é responsável pela redução do acetil aldeído, que é tóxico para os tecidos vegetais. Resultando na regeneração contínua de NAD no citoplasma (CHUNG; FERL, 1999). Assim, a indução de ADH pode aumentar a sobrevivência de plantas em condições inundadas (JOHNSON et al., 1994).

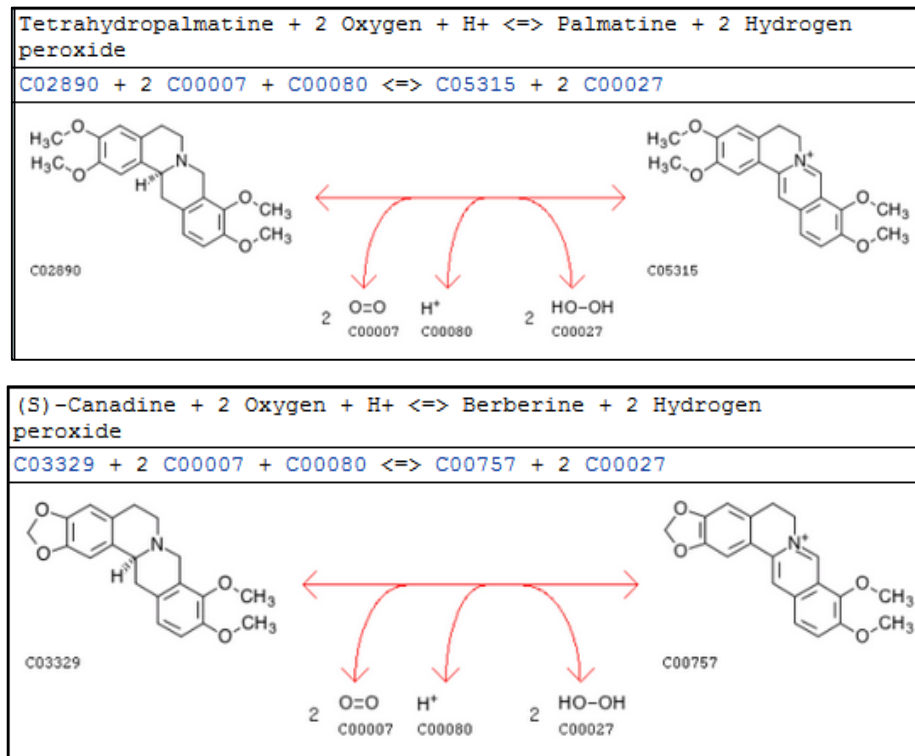
O *acetil-CoA*, produzido pela descarboxilação oxidativa do piruvato, é uma Co-enzima que tem uma forte contribuição no armazenamento de lipídeos e no metabolismo energético global. Catalisa a formação de *malonil-CoA*, um substrato essencial para a síntese de ácidos graxos (ABU-ELHEIGA et al., 2001; ABU-ELHEIGA et al., 2003). A produção de ácidos graxos ocorre nos plastídeos que é catalisada por duas enzimas, *acetil-CoA Carboxilase (ACCase)* e *sintase de ácido graxo* nas células de plantas, esta organela não existe em animais (TRONCOSO-PONCE et al., 2016; SASAKI et al., 1995; SASAKI et al., 2004). Outras vias que estão relacionadas ao *acetil-CoA* e as respectivas enzimas identificadas no látex, são as vias de metabolismo de glicerolfosfolipídio (enzimas *acyl-CoA:sn-glycerol-3-phosphate 1-O-acyltransferase*; *phosphatidylcholine:sterol O-acyltransferase*), metabolismo de glicerolipídio (*acyl-CoA:sn-glycerol-3-phosphate 1-O-acyltransferase*; *phospholipid:1,2-diacyl-sn-glycerol O-acyltransferase*) e biossíntese de tetraciclina (*acetyl-CoA:carbon-dioxide ligase (ADP-forming)*).

Algumas enzimas estão presentes no látex de *H. sucuuba* e participam de diferentes processos metabólicos, como metabolismo de alanina, aspartato e glutamato (enzima identificada nessa via foi a *2-Nomega-L-arginino succinate arginine-lyase (fumarate-forming)*, via da pentose fosfato (com a enzima *D-glucose-6-phosphate:NADP+ 1-oxidoreductase*), metabolismo de amido e açúcar (*trehalose-6-phosphate phosphohydrolase*; *glucose-1-phosphate adenylyltransferase*), e biossíntese de fenilalanina, tirosina e triptofano [*L-serine hydro-lyase [adding 1-C-(indol-3-yl) glycerol 3-phosphate, L-tryptophan and D-glyceraldehyde-3-phosphate-forming]*].

Foram também identificadas duas enzimas que participam da via metabólica de biossíntese de alcaloides isoquinolínicos. A enzima (*S*)-*tetrahydroberberine:oxygen oxidoreductase* (EC: 1.3.3.8) classificada como uma oxidoreductase, atua em três momentos da via e catalisa a produção de compostos de interesse farmacológico como a *Palmatine* e

Berberine (Figura 21) e a enzima *(RS)-norcoclaurine 6-O-methyltransferase* (EC:2.2.1.128) que atua em duas reações, catalisando a produção de (S)-Coclaurine e (S)-6-O-Methylnorlaudanosoline.

Figura 21. Reações catalisadas pela enzima EC: 1.3.3.8.



FONTE: KEGG, 2017

As enzimas estão ressaltadas no esquema da via dos alcalóides de isoquinolínicos apresentado na Figura 22.

As enzimas identificadas participam de reações que apresentam como produto, alcalóides com potencial atividade e interesse farmacológico como os agentes berberina e palmatina, que apresentam variadas ações farmacológicas e têm sido amplamente utilizados na medicina popular. A biossíntese de alcalóides de isoquinolínicos prossegue a via de descarboxilação de tirosina para produzir dopamina, que em conjunto com 4-hidroxifenilacetaldeído, um aldeído derivado de tirosina, é convertido em reticulina, um precursor importante de vários alcalóides de benzilisoquinolina (KEGG,2017).

pode ser importante na manutenção dos estados celulares normais, pois participam de processos que envolvem adesão celular, como a diferenciação, reparo e cicatrização (VALKO et al., 2007; SILVA e JASIULIONIS, 2014).

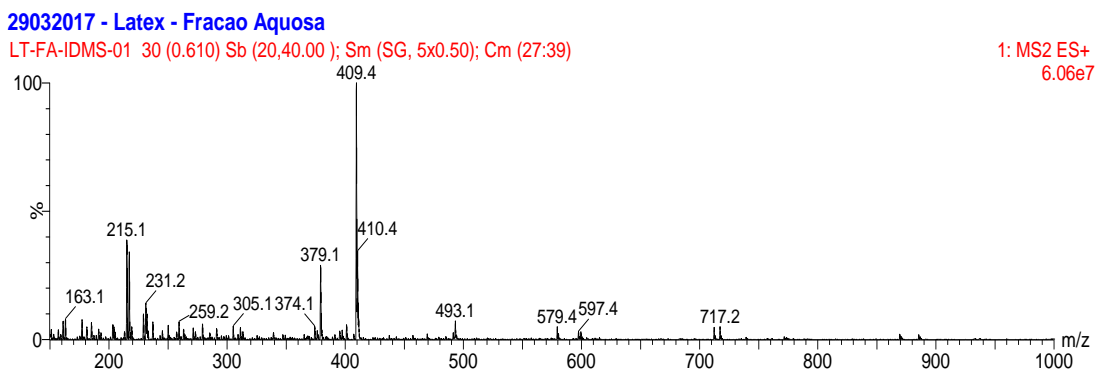
Os alcalóides berberina e palmatina são tóxicos para insetos e vertebrados e inibem a multiplicação de bactérias, fungos e vírus. Em consequência, estas atividades bioquímicas podem mediar a defesa química contra microorganismos, vírus e herbívoros nas plantas que produzem estes alcalóides (SEUNG et al., 2002).

Outras atividades farmacológicas da berberina incluem propriedades antimicrobianas; redução da glicose e do colesterol; antitumorais e imunomoduladoras. Estudos sugerem que a berberina pode ser benéfica para o tratamento da doença de Alzheimer (DA) limitando a patogênese de placas amilóides extracelulares e emaranhados neurofibrilares intracelulares. Evidências crescentes indicam que a berberina exerce um papel protetor na aterosclerose relacionada com as propriedades hipolipemiantes e hipoglicemiantes, implicando que a berberina tem o potencial de inibir estes fatores de risco para DA. A base farmacológica através da qual a berberina pode retardar o estresse oxidativo e a neuroinflamação para exibir seu papel protetor na DA, constituindo-se numa potencial abordagem terapêutica, foi explicada numa revisão recente publicada por Zhiyou et al. (2016).

7.11 Identificação de metabólitos secundário

No espectro de ESI-MS de íons totais da fração aquosa, modo positivo, foi possível identificar um íon m/z 409,4 $[M+H-18]^+$, apresentado na Figura 23.

Figura 23. Espectro de ESI-MS de ions totais da fração aquosa (modo positivo).



Analisando o fracionamento deste ion, cujo espectro de ESI-MS/MS do íon m/z 409,4 está apresentado na Figura 24, observou-se que o referido ions apresentava o padrão de fragmentação do lupeol, Figura 26, uma vez que mostrou um fragmento característico da molécula lupeol de m/z 137.

Figura 24. Espectro de ESI-MS/MS do ion m/z 409,4 (modo positivo).

29032017 - Latex - Fracao Aquosa - MSMS m/z 409
LT-FA-MSMS-03 13 (0.419) Sm (SG, 5x0.50); Cm (11:14)

2: Daughters of 409ES+
2.50e6

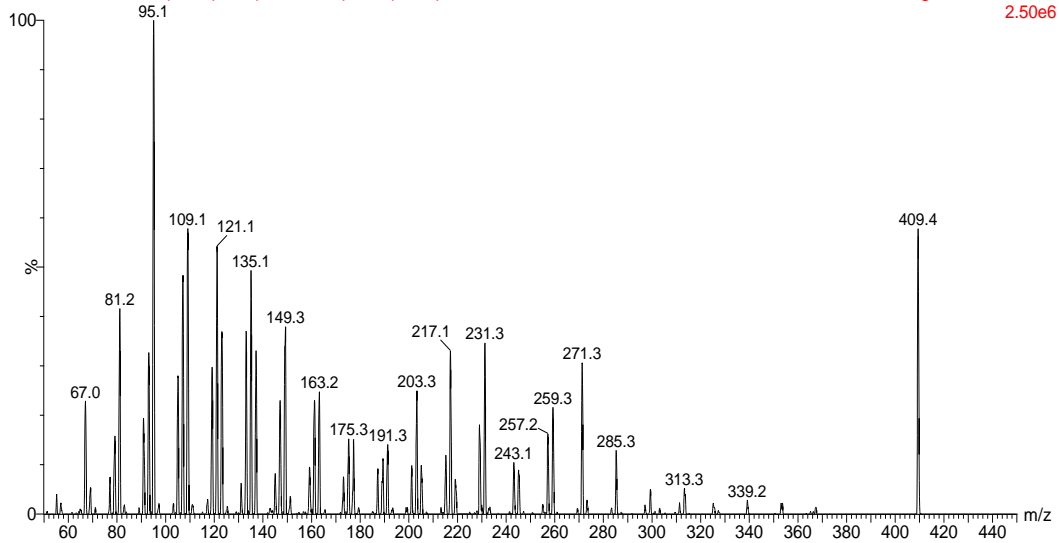
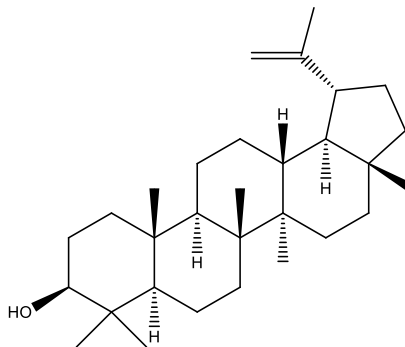


Figura 25. Estrutura química do lupeol.



FONTE: Mo et. al., 2013.

Nos espectros de ESI-MS de íons totais da fração acetato bem como do látex bruto (Figura 26 e 27) também mostram o mesmo íon m/z 409,4 Os espectros de ESI-MS/MS do íon

m/z 409,4 da fração acetato e do látex bruto apresentaram o mesmo perfil de fragmentação (Figura 28 e 29). Entre as proteínas identificadas na análise proteômica esta a enzima geranyl pirofosfato sintase (EC: 2.5.1.29) que atua na biossíntese de lupeol, um triterpeno da classe dos terpenóides. A presença do terpenoide no soro detectado pela espectroscopia de massas ESI-MS/MS confirma a presença do lupeol no látex de *Himatanthus sucuuba* e corrobora com os dados descritos na literatura (SALEEM, 2009; MO et al; 2013).

O efeito anti-inflamatório do látex de *Himatanthus sucuuba* é provavelmente oriundo de lupeol cujo efeito foi verificado na inibição de edema e constrictões abdominais (SALEEM, 2009; SINGH et al., 2017). Vários estudos relatam a atividade anti-inflamatória desse composto (GEETHA et al., 1999; DE MIRANDA et al., 2000; VASCONCELOS et al., 2008; NGUEMFO et al., 2009), o lupeol apresenta-se como anti-microbiano, anti-protozoário, anti-angiogénico (SIDDIQUE et al., 2011). Atividades farmacológicas benéficas incluindo os antioxidantes, anti-hiperglicémicos, anti-dislipémicos e anti-mutagénicos, anti-diabético, anti-asmático, anti-artrítica, cardioprotectora, hepatoprotectora, nefroprotectora, neuroprotectora também atribuídas a esse composto (TSAI et al., 2016).

O lupeol pode ser utilizado como um potente anti-oxidante contra a genotoxicidade induzida por pesticidas em trabalhadores agrícolas. Aumentando os níveis de enzimas anti-oxidantes, superóxido dismutase (SOD) e catalase e diminuição da geração de ROS (espécies reativas de oxigênio) e LPO (peroxidação de lipídeos) intracelular (SRIVASTAVA et al., 2016).

Figura 26. Espectro de ESI-MS de ions totais da fração acetato (modo positivo).

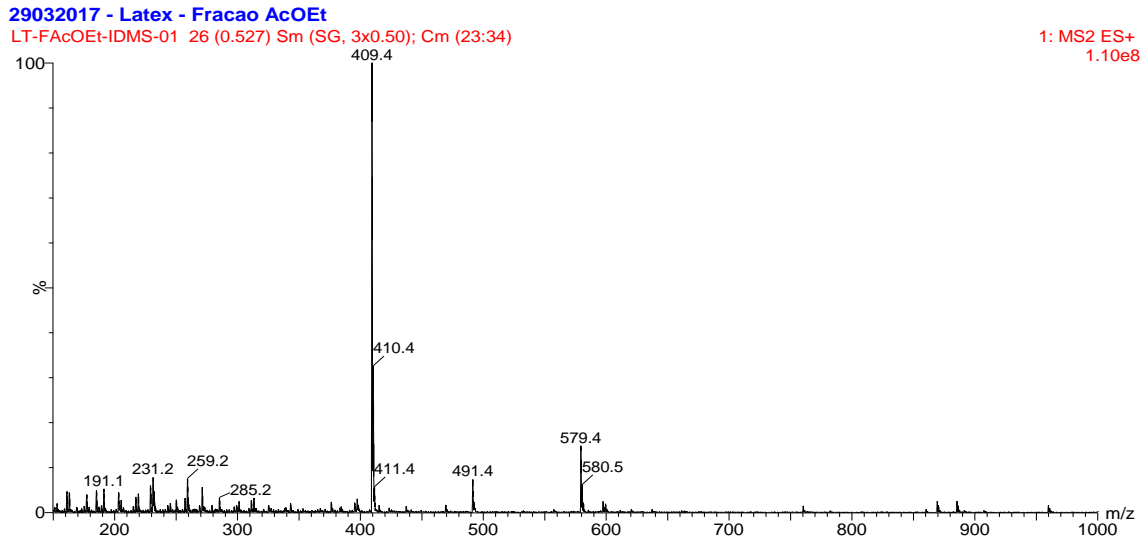


Figura 27. Espectro de ESI-MS de ions totais do látex bruto (modo positivo).

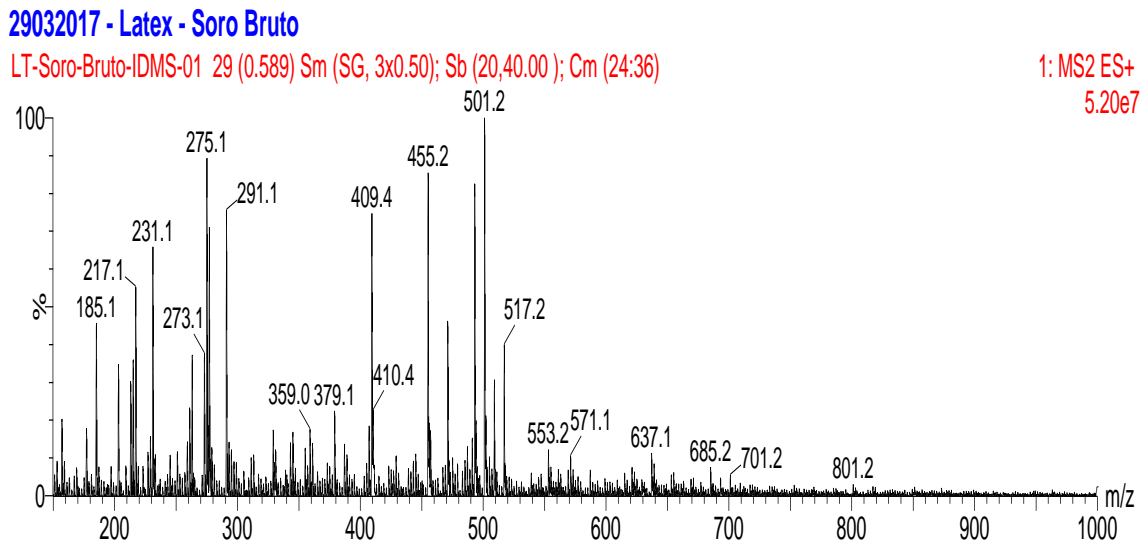


Figura 28. Espectro de ESI-MS/MS do ion m/z 409,4 (fração acetato; modo positivo).

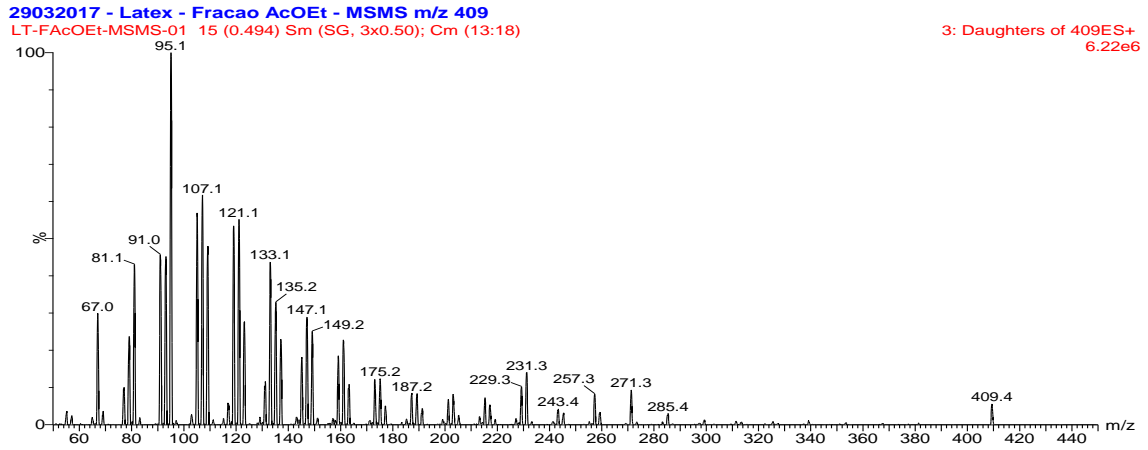
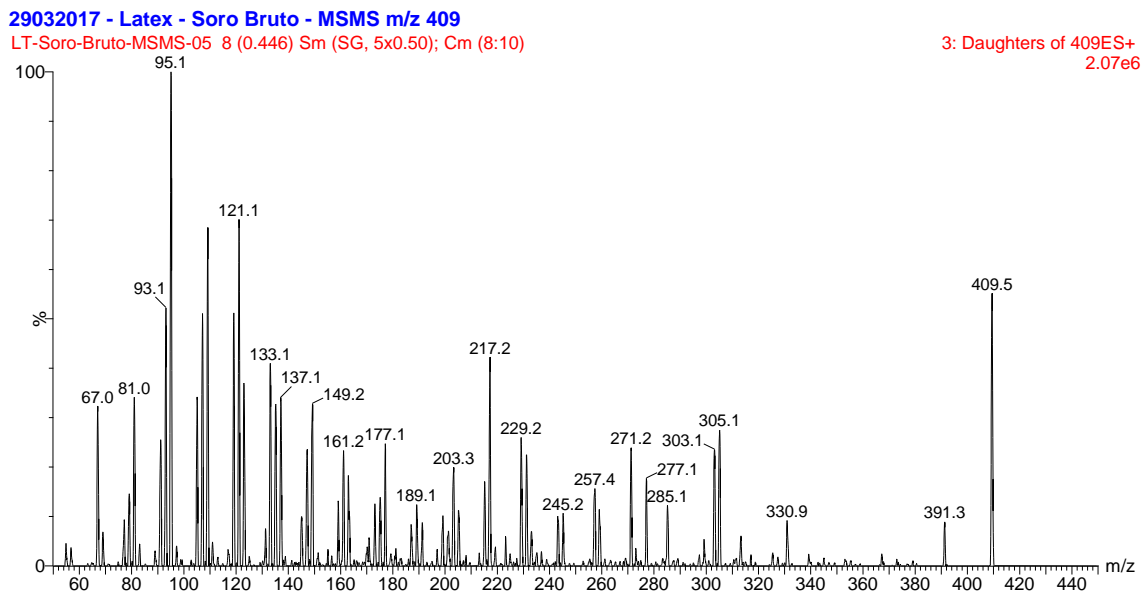


Figura 29. Espectro de ESI-MS/MS do ion m/z 409,4 (látex bruto; modo positivo).



CONCLUSÕES

- i) A utilização da técnica de espectrometria de massas resultou na identificação de 1471 proteínas, das quais 587 proteínas apresentaram score igual ou superior a 20.
- ii) Dentre essas 587 proteínas, foram encontradas 267 proteínas preditas o que equivale a 45 %, 158 hipotéticas (27 %), 128 anteriormente caracterizadas (22 %), 24 putativas (4 %) e 10 proteínas não definidas (2 %).
- iii) Das proteínas encontradas 40% apresentaram ponto isoelétrico menor que 6, 32,64 % apresentaram ponto isoelétrico maior que 8.
- iv) Entre essas proteínas caracterizadas determinou-se que mais de 70 % atuam em processos metabólicos e 50% em processos celulares variados, identificou-se que grande parte contribui de modo geral em processos metabólicos e celulares, com atividades catalíticas e de ligação. São encontradas em distintos locais subcelulares, exibindo atividades enzimáticas como transferases, hidrolases, oxirredutases e outras funções.
- v) As proteínas identificadas foram alocadas em 30 distintas vias metabólicas, e dentre essas, a via de biossíntese de alcalóides isoquinolínicos. As enzimas identificadas nessa via (*tetrahydroberberine oxidase*; (*S*)-*THB oxidase* e (*RS*)-*norcoclaurine 6-O-methyltransferase*) catalisam reações e produzem produtos de interesse farmacológico, como os alcalóides berberine e palmatine.
- vi) Este é o primeiro trabalho que relata a identificação e predição de proteínas presentes no látex de *H. sucuuba* e os resultados obtidos neste estudo forneceram informações importantes que possibilitam futuras pesquisas.

REFERÊNCIAS BIBLIOGRÁFICAS

ABU-ELHEIGA L; MATZUK M.M; ABO-HASHEMA K.A; WAKIL S.J. Continuous Fatty Acid Oxidation and Reduced Fat Storage in Mice Lacking Acetyl-CoA Carboxylase 2 *Science*, n. 291, p. 2613–2616, 2001.

ABU-ELHEIGA L; OH W; KORDARIP; WAKIL S. J. Acetyl-CoA carboxylase 2 mutant mice are protected against obesity and diabetes induced by high-fat-high-carbohydrat diets. *Proceeding of the National Academy of Sciences of the United State of America*, n.100, p. 10207-10212, 2003.

ANDRADE, T.; G. CAETANO; D.MASSON; C. LANDIN; J. COUTINHO-NETO, M. FOSS; M. FAAD. Protein from *Hevea brasiliensis* látex rubber tree enhances wound healing in diabetic rats. *Endocrine Abstrates*, v. 29, p. 542, 2012.

ALAM, G., MANJUL PRATAP SINGH, ANITA SINGH. Wound healing potential of some medicinal plants. *International Journal of Pharmaceutical Sciences Review and Research*, v. 9, n. 026, August 2011.

ALTELAAR, A. F.; MUNOZ, J.; HECK, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nature Reviews Genetics, London*, v. 14, n. 1, p. 35-48, 2013.

ARN, P. H. Phenylketonuria (PKU). In: Aminoff, M. J. e Daroff, R. B. *Encyclopedia of the Neurological Sciences (Second Edition)*. Oxford: Academic Press, p.887-889, 2014.

BARRETO, ALAÍDE DE SÁ; ANA CLAUDIA F. AMARAL. ÁCIDO 15-DESMETILISOPLUMIERÍDEO, UM NOVO IRIDÓIDE ISOLADO DAS CASCAS DE *Plumeria rubra* E DO LÁTEX DE *Himatanthus sucuuba* *Quimica Nova*, v. 30, n. 5, p. 1133-1135, 2007.

BHATTACHARJEE, S. Sites of generation and physicochemical basis of formation of reactive oxygen species in plant cell. In: GUPTA, S.D. Reactive oxygen species and antioxidants in higher plants. *Enfi eld: Science Publishers*, p.1-30, 2010.

BIRDSALL, TC; KELLY GS. Berberine: Therapeutic potential of an alkaloid found in several medicinal plants. *Alternative Medicine Review*, n. 2, p. 94–103, 1997.

BOLAY, E. Feigenund wurgefeigen. *Pharmazie in Unserer Zeit*, n. 4, p. 97–112, 1979.

CHO, KYONG WON, XIONG-YAN CHEN, NAZIM MOHAMAD UDDIN, YEONGGIL RIM, UYEON MOON, JIN-HEE JUNG, CHUNLIN SHI, HYOSUBCHU, SUWHA KIM, SEON-WON KIM, ZEE-YONG PARK, JAE-YEAN KIM. Comprehensive proteome analysis of lettuce latex using multidimensional protein-identification technology. *Phytochemistry*, n.70, p. 570-578, 2009.

CHO, WON KONG; XIONG-YAN CHEN; EONGGIL RIM; HYOSUB CHU; JO Y; SUWHA KIM; ZEE-YONG; JAE-YEAN KIM. Extended latex proteome analysis deciphers additional roles of the lettuce laticifer. **Plant Biotechnology Reports**, n. 4, p. 311-319, 2010.

CHO, WONG KYONG; YEONHWA JO; HYOSUB CHU; SANG-HO PARK; KOOK-HYYUNG KIM. Integration of latex protein sequence data provides comprehensive functional overview of latex proteins. **Molecular Biology Reports**, v. 41, n. 3 p. 1469-1481, 2014.

CHUNG, H. J., AND R. J. FERL. Arabidopsis alcohol dehydrogenase expression in both shoots and roots is conditioned by root growth environment. **Plant Physiology**, n.121, p. 429–436, 1999.

CKLESS, K; SCHLOTTFELDT, JL; PASQUAL, M; MOYNA P; HENRIQUES JA; WAJNER M. Inhibition of in-vitro lymphocyte transformation by the isoquinoline alkaloid berberine. **Journal of Pharmacy and Pharmacology**, v. 47, n. 12, p.1029-31, 1995.

CONESA, A; GÖTZ S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. **International Journal of Plant Genomics**, p. 1–12, 2008.

CORREA, M. P. Dicionário das plantas úteis do Brasil e das exóticas cultivadas no Rio de Janeiro: Imprensa Nacional, v.6, 1984.

CORNISH-BOWDEN, A. Current IUBMB recommendations on enzyme nomenclature and kinetics. **Perspectives in Science**, v. 1, p. 74-87, 2014.

CUESTRA, SERGIO MARTINEZ; SYED ASAD RAHMAN; NICOLAS FURNHAM AND JANET M. THORTON. Biophysical Perspective. **Biophysical Journal**. v.109, p. 1082-1086, 2015.

DA SILVA, CAMILA TAINAH; JASIULIONIS, MIRIAM GALVONAS. Relação entre estresse oxidativo, alterações epigenéticas e câncer. *Ciencia e Cultura*, v. 66, n.1, São Paulo, 2014.

DAI, LONGJUM; GUIUAN KANG; YU LI; ZHIYI NIE; CUIFANG DUAN; RIZHONG ZENG. In-deph proteome analysis of the rubber particle of *Hevea brasiliensis* (para rubber tree). **Plant Molecular Biology**, n. 82, p. 155-168, 2013.

DECKER, GABRIELE; GERHARD WANNER; MEINHARD H. ZENK; FRIEDRICH LOTTSPREICH. Characterization of proteins in latex of the opium poppy (*Papaver somniferum*) using two-dimensional gel electrophoresis and micro sequencing, n. 21, p. 3500-3516, **Electrophoresis**, 2000.

DE MIRANDA, A.L.P; SILVA, J.R.A; REZENDE, C.M.; PINTO, A.C.; PINHEIRO, M.L.B; CORDEIRO, M.C; TAMBORINI, E; PARRINI, J.S; Anti-inflammatory and analgesic activities of the látex containing triterpenes from *Himatanthus sucuuba*. **Planta Medica**, n. 66, p. 284-286, 2000.

DEWAN, S., SANGRAULA, H., & KUMAR, V. L. Preliminary studies on the analgesic activity of latex of *Calotropis procera*. **Journal of Ethnopharmacology**, n. 73, p.307–311, 2000.

DI STASI, L.C.; HIRUMA-LIMA, C.A. **Plantas medicinais na Amazônia e na Mata Atlântica**. 2. ed. São Paulo: UNESP, p. 604, 2002.

DOS SANTOS, A. A. **Exploração de uma biblioteca genômica de *Passiflora edulis* f. *flaviacarpa* por sequenciamento de BAC-ends**. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2013.

DUAN, C. F.; NIE, Z. Y.; ZENG, R. Z. Establishment of 2DE system and primary analyses on the membrane proteins of rubber particles in *Hevea brasiliensis* by MALDITOF. **Chinese Journal of Tropical Crops**, n. 27, p. 22–29, 2006.

FAHN, A. Plant anatomy 4^a ed. Oxford: **Pergamon Press**, 1990.

FERREIRA, J. L. PINTO; ANA C. F. AMARAL; RENATA B. ARAÚJO; JUNIOR R. CARVALHO; CAROLYN E. B. PROENÇA; SANDRA A. P. M. FRAGA; JEFFERSON ROCHA A. SILVA. Pharmacognostical Comparison of Three Species of *Himatanthus*. **International Journal of Botany**, v. 5, n. 2, p. 171-175, 2009.

FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M. Electrospray ionization for mass spectrometry of large biomolecules. **Science, New York**, v. 246, n. 4926, p. 64-71, Oct 1989.

FERNANDES, M.Z.L.C.M; FERNANDES, R.M. SOUSA, M.C.B.B; LOPES, J.B. Determinação da toxicidade aguda da *Himatanthus sucuuba* (Spruce). Woodson (Apocynaceae) em camundongos. **Rev Bras Farm**, n. 81, p. 98-100, 2000.

FIGUEREDO, I. S. T.; OLIVEIRA, R. S. B.; FREITAS, L. B. N.; PINHEIRO, R. S. P.; ARAGÃO, K. S.; GONZAGA, M. L. C.; RICARDO, N. M. P. S.; BRITO, G.A.C.; RAMOS, M. V.; ALENCAR, N. M. Wound healing modulation by a biomembrane of laticifers proteins from *Calotropis procera* (AIT.). R. B. **Plant Med**, 78, 2012.

FORSTERMANN, ULRICH FO; SESSA, WILLIAM C. Basic science for the clinician Nitric oxide synthases: regulation and function. **European Heart Journal**, v. 33, p. 829–837, 2012.

FOLKMAN, J.; Y. SHING. Angiogenesis. **Journal of Biological Chemistry**, v.267, n.16, p.10931-10934, 1992.

FREITAS, C. T.; WALLACE T. da CRUZ; MARIA Z. SILVA; ILKA M. VASCONCELOS; FREDERICO B. M. MORENO; RENATO A. OLIVEIRA; ANA C. O. MONTEIRO-MOREIRA; LUCIANA M. R. ALENCAR; JEANLEX S. SOUSA; BRUNO A. M. ROCHA; MARCIO V. RAMOS. Proteomic analysis and purification of an unusual germin-like protein with proteolytic activity in the latex of *Thevetia peruviana*. **Plant**, 2016.

FREITAS, C. D. T.; LOPES, J. L. S.; BELTRAMINI, L. M. ; OLIVEIRA, R. S. B.; OLIVEIRA, J. T. A. Osmotin from *Calotropis procera* latex: New insights into structure and antifungal properties. **Bochimica et Biophysica Acta (BBA)- Biomembranes**, v. 1808, n. 10, p. 2501-2507, 2011.

FREITAS, C. D. T.; OLIVEIRA, J. S.; MIRANDA, M. R. A.; MACEDO, N. M. R.; SALES, M. P.; VILLAS-BOAS L. A.; RAMOS, M. V. Enzymatic activities and protein profile of latex from *Calotropis procera*. **Plant Physiology and Biochemistry**, v. 45, p. 781-789, 2007.

GASKELL, S. J. Electrospray: Principles and Practice. **Journal of Mass Spectrometry, Chichester**, v. 32, n. 7, p. 677-688, Jul 1997.

GEBHARDT, R. In vitro screening of plant extracts and phytopharmaceuticals: Novel approaches for the elucidation of active compounds and their mechanisms. **Planta medica**, v. 66, n. 2, p. 99-105, 2000.

GEETHA, T; VARALAKSHMI P. Anticomplement activity of triterpenes from *Crataeva nurvala* stem bark in adjuvant arthritis in rats. **General Pharmacology**, v. 32, n. 4, p. 495-497, 1999.

GUERRA, MO; PETERS VM. Screening for reproductive toxicity in rats for a decoction of *Himatanthus sucuuba* stem bark. **Journal Ethnopharmacology**, n. 34, p. 195-199, 1991.

GUNTER, R., HANS, P.S., FRIEDRICH, D., PETER, L. Activation and inactivation of human factor X by proteases derived from *Ficus carica*. **British Journal Haematology**, n. 119, p. 1042-1051, 2002.

HAGEL, JILLIAN M.; EDWARD C.; YEUNG; PETER J. FACCHINI. Got milk? The secret life of laticifers. **Trends in Plant Science**, v. 13, n. 12, p. 631-639, 2008.

HASHEM, SEYYED ABBAS; MADANI, SEYYED ABDOLLAH; FAZLI, MEHRAN ABEDIANKENARI, SAIED. *Ficus Carica* Latex; a Review with Focus on Cancer. **Annals of Advanced Sciences**, v. 1, n. 1, 2017.

HARTREE, E. F. Determination of protein: A modification of the lowry method that gives a linear photometric response. **Analytical Biochemistry**, v. 48, n. 2, p. 422-427, 1972.

HAVANAPAN, PHATTARA-ORN; APICHAJ BOURCHOOKARN; ALBERT KETTERMAN; CHARTCHAI KRITTANAI. Comparative proteome analysis of rubber latex serum from pathogenic fungi tolerant and susceptible rubber tree (*Hevea brasiliensis*). **Journal of Proteome**, n. 131, p. 82-92, 2016.

HEIN, M. Y. et al. Chapter 1 - Proteomic Analysis of Cellular Systems. In: Walhout, A. J. M., Vidal, M. Handbook of Systems Biology. San Diego: **Academic Press**, p. 3-25, 2013.

HERCULANO, RONDINELLI DONIZETTI; TZU, LEE CHEN; SILVA, CECILIA PEREIRA; BRUNELLO, CARLOS ALBERTO; QUEIROZ, ÁLVARO ANTÔNIO ALENCAR DE QUEIROZ; KINOSHITA, ANGELA; GRAEFF, CARLOS FREDERICO DE OLIVEIRA GRAEFF. Nitric Oxide Release Using Natural Rubber Latex as Matrix. **Materials Research**, v. 14, n. 3, p. 355-359, 2011.

HOU, JINGYU. New Approaches of Protein Function Prediction from Protein Interaction Networks. **Academic Press**, p. 124, 2017.

JOHNSON, J. R., B. G. COBB, AND M. C. DREW. 1994. Hypoxic induction of anoxia tolerance in roots of *Adh1* null *Zea mays*. **Plant Physiology**, n. 105, p. 61–67, 1994.

KANEDA, Y; TORII M; TANAKA T; AIKAWA, M. In vitro effects of berberine sulphate on the growth and structure of *Entamoeba histolytica*, *Giardia lamblia* and *Trichomonas vaginalis*. **Annals Tropical Medicine and Parasitology**, 85, n. 4, p. 417-25, 1991.

KARUPPANAPANDIAN, T; JUN-CHEOL MOON; CHANGSOO KIM; KUMARIAH, MANOHARAN; WOOK, KIM. Reactive oxygen species in plants: their generation, signal transduction, and scavenging mechanisms. **Australian Journal of Crop Science**, v.5, n. 6, p.709-725, 2011.

KEKWICK, R. G. O. Látex and Laticifers. **Encyclopedia of Life Sciences**, p.1-6, 2001.

KINNULA, V.L; CRAPO, J. D; KO, RAIVIO. Generation and disposal of reactive oxygen metabolites in the lung. **Laboratory Investagation**, v.73, n.1, p.3-19, 1995.

KIMMERER. THOMAS W; MARY A. STRINGER. Alcohol Dehydrogenase and Ethanol in the Stems of Trees EVIDENCE FOR ANAEROBIC METABOLISM IN THE VASCULAR CAMBIUM. **Plant Physiology**, n. 87, p. 693-697, 1988.

KITAJIMA, SAKIHITO; KENJI MIURA; WATARU AOKI; KATSUYUKI T. AMATO; TOKI TAIRA; RYUTA MURAKAMI; SHUNSUKE ABURAYA. Transcriptome and proteome analyses provide insight into laticifers defense of *Euphorbia tirucalli* against pests. **Plant Physiology and Biochemistry**. n. 108, p. 434-446, 2016.

KUMAR, V. L.; BASU, N. Anti-inflammatory activity of the latex of *Calotropis procera*. **Journal of Ethnopharmacology**, v.44, p.123–125, 1994.

KUMAR VIJAY L.; B. GURUPRASAD; SYED MERAJ A. FATMI 1 & PRIYANKA CHAUDHARY; NYLANE MARIA NUNES ALENCAR; JOSÉ VITOR MOREIRA LIMA-FILHO; MÁRCIO VIANA RAMOS. In Vivo Efficacy of Latex from *Calotropis procera* in Ameliorating Fever—Biochemical Characteristics and Plausible Mechanism **Applied Biochemistry Biotechnology**, 2017.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, n.224, p. 680-685, 1970.

LARROSA, C. R. R.; M. R. DUART. Contribuição ao estudo anatômico do caule de *Himatanthus sucuuba* (Spruce ex Müll. Arg.) Woodson, Apocynaceae. **Revista Brasileira de Farmacognosia**, 15 (2), p. 110-114, 2005.

LEHNINGER, DAVID NELSON; MICHAEL M. COX. **Princípios de Bioquímica de Leninger**. 5ª edição, Ed. Sarvier, p. 183-185, 2011.

LEMONS, O.L.; FERREIRA, L. A.M.; CARDOSO, V. N.; CASSALI, G.D.; SALAS, C.E.; LOPES, M.T.P. Skin-healing activity and toxicological evaluation of a proteinase fraction *Carica candamarcensis*. **European Journal Dermatology**, v. 21, n. 5, p.722-730, 2011.

LI DEJUN, YAN JIE, DENG ZHI, CHEN CHUNLIU, HE PENG, WU MIN, CHEN SHOUCAI. Evaluation of Three Methods For Protein Extraction Suitable for 2-DE of *Hevea brasiliensis* C-serum. **Chinese Agricultural Science Bulletin**, v. 26, n. 16, p. 273-279, 2009.

LYNN, K. R.; CLEVETTE-RADFORD, N. A. Acid phosphatase from lattices of Euphorbiaceae. **Phytochemistry**, v. 26, p. 655-657, 1987.

MAHLBERG, P. G. Laticifers: An historical perspective. **Botanical Review**, n. 59, p. 1-23, 1993.

MANN, M.; HENDRICKSON, R. C.; PANDEY, A. Analysis of proteins and proteomes by mass spectrometry. **Annual Review of Biochemistry**, Palo Alto, v. 70, n 1, p. 437-473, Jul 2001.

METCALF, C.R. Distribution of latex in the plant kingdom. **Economic Botany**, v. 21, n. 2, p. 115-127, 1967.

MELLO, V. J.; GOMES, M. T. R.; LEMOS, F. O.; DELFINO, J. L.; ANDRADE, S. P.; LOPES, M. T. P.; SALAS, C. E. The gastric ulcer protective and healing role of cysteine proteinases from *Carica candamarcensis*. **Phytomedicine**, n. 15, p. 237-244, 2008.

MENDONÇA, RICARDO JOSÉ; VANESSA BEATRIZ MAURICIO; LARISSA DE BERTOLLI TEIXEIRA; JOÃO JOSÉ LACHAT AND JOAQUIM COUTINHO-NETO. Increase vascular permeability, angiogenesis and wound healing induced by the sérum of natural latex of the rubber tree *Hevea brasiliensis*. **Phytotherapy Research**, v. 24, p. 764-768, 2010.

MO, SHUNYAN; LINLIN DONG; W. JEFFREY HURST; RICHARD B. VAN BREEMEN. Quantitative analysis of phytosterols in edible oils using APCI liquid chromatography-tandem mass spectrometry. **Lipids**, v. 48, n. 9, p. 949–956, 2013.

MORCELLE, S.R.; CAFFINI, N.O.; PRIOLO, C.N. Proteolytic properties of *Funastrum clausum* latex. **Fitoterapia** v.75, p. 480-493, 2004.

MURUGAN, V; INANDAR JA. Organographic distribution , structure and ontogeny of laticifers in *Plumeria alba* Linn. **Proceeding: Plant Sciencis**, v. 97, p. 25-31, 1987.

NAWROT, ROBERT; ANDRZEJ KALINOWSKI; ANNA GOZDZICKA-JOZEFIAK. Proteomic analysis of *Chelidonium majus* Milk sap using two-dimensional gel electrophoresis and tandem mass spectrometry. **Phytochemistry**, v. 68, n. 12, p 1612-1622, 2007.

NAWROT, ROBERT; RICO LIPPMANN; ANDREA MATROS; OSKAR MUSIDLAK; GRZEGORZ NOWICKI; HANS-PETER MOCK. Proteomic comparison of *Chelidonium majus* L. latex in different phases of plant development. **Plant Physiology and Biochemistry**, n. 112, p. 312-325, 2017.

NAWROT, ROBERT; JAKUB BARYLSKI; RICO LIPPMANN; LOTHAR ALTSCHMIED; HANS-PETER MOCK. Combination of transcriptomic and proteomic approaches helps to unravel the protein composition of *Chelidonium majus* L. milky sap. **Planta**, v. 244 n. 5, p.1055–1064, 2016.

NGUEMFO, E. L.; DIMO, T; DONGMO, A. B.; AZEBAZE, A. G.; ALAOUI, K.; ASONGALEM, A. E. Anti-oxidative and anti-inflammatory activities of some isolated constituents from the stem bark of *Allanblackia monticola* Staner L.C (Guttiferae). **Inflammopharmacology**, v. 17, n. 1, p. 37–41, 2009.

OSONIYI, O.; ONAJOBI, F. Coagulant and anticoagulant activities in *Jatropha curcas* latex. **Journal Ethnopharmacology**, v. 89, n. 1, p. 101-105, 2003.

PELENOVA BIOTECNLOGIA, 2017. Látex de seringueira. Disponível em: <<http://www.pelenova.com.br>>. Acesso em 14 de jan. 2017.

PICKARD, W. F. Laticifer and secretory ducts: two other tube systems in plants. **New Phytologist**, v. 177, nn. 4, p. 877-888, 2008.

PLUMEL, M.M. Le genre *Himatanthus* (Apocynaceae) révision taxonomique. **Bradea** 5: p. 1-118, 1991.

POMASTOWSKI, P.; BUSZEWSKI, B. Two-dimensional gel electrophoresis in the light of new developments. **Trends in Analytical Chemistry, Amsterdam**, v. 53, p. 167-177, Jan 2014.

POSCH, ANTON; ZHIPING, CHEN; COLIN WHEELER; MIKE J. DUNN; MONIKA RAULF-HEIMSOTH; XAVIER BAUR. Characterization and identification of látex allergens by two-dimensional electrophoresis and protein micro sequencing. **J. Allergy Clin Immunol**, v. 99, n. 3, p. 385-395, 1997.

RACKOVA, L.; MAJEKOVA, M.; KOST'ALOVA, D.; STEFEK, M. Antiradical and antioxidante activities of alkaloids isolated from *Mahonia aquifolium*. Structural aspects. **Bioorganic Medicinal Chemistry**, v.12, n. 17, p. 4709-4715, 2004.

RAJESH, R. A.; NATARUJU, C. D. R. GOWDA. B. M.; FREY, F. J.; FREY, B. S. V. Purification and characterization of a 34-KDa, heat stable glycoprotein from *Synadenium grantii* latex: action on human fibrinogen and fibrin clot. **Biochimie**, v. 88, n. 10, p. 1313-1322, 2006.

RAJESH, R.A.; C.D. RAGHAVENDRA GOWDA; A. NATARAJU; B.L. DHANANJAYA; K. KEMPARAJU; B.S. VISHWANATH. Protocoagulant activity of *Calotropis gigantean* latex associated with fibrin (ogen) olytic activity. **Toxicon**, v. 46, p. 84-92, 2005.

RAMJAWAN, RAKESH R.; GRIFFIOEN, ARJAN W.; DUDA, DAN G. Anti-angiogenesis for cancer revisited: Is there a role for combinations with immunotherapy? **Angiogenesis**, v.20, n. 2, p.185–204, 2017.

RASKOVIC, BRANKICA; OLGA BOZOVIC; RADIVOJJE PRODANOVIC; VESNA NIKETIC and NATALIJA POLOVIC. Identification, purification and Characterization of a novel collagenolytic serine protease from fig (*Ficus carica* var. Brown Turkey) latex. **Journal of Bioscience and Bioengineering**, v. 118, n. 6, p. 622-627, 2014.

RODRIGUES SILAS P.; JOSE A AVENTURA; CLEMENTE AGUILAR; ERNESTO S. NAKAYASU; HYUNG WON CHOI; TIAGO J. P. SOBREIRA; LILIAN NOHARA; LUCIANA S. WERMELINGER; IGOR C. ALMEIDA; RUSSOLINA B. ZINGALI; PATRICIA M. B. FERNANDES. Label-free quantitative proteomics reveals differentially regulated proteins in latex of sticky diseased *Carica papaya* L. plants. **Journal of Proteomics**, n. 75, p. 3191-3198, 2012.

SALEEM, MOHAMMAD. Lupeol, A Novel Anti-inflammatory and Anti-cancer Dietary Triterpene. **Cancer Lett**, v. 28, n. 285(2), p. 109–115, 2009.

SASAKI, YUKIKO; KONISHI, TOMOKAZU; NAGANO, YUKIO. The Compartmentation of Acetyl-Coenzyme A Carboxylase in Plants. **Plant Physiology**, n. 108, p. 445-449, 1995.

SASAKI, YUKIKO; NAGANO, YUKIO. Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. **Bioscience Biotechnology Biochemistry**, v. 68, n. 6, p. 1175-84, 2004.

SETHURAJ, M. R. Natural Rubber, Biology, Cultivation and Technology. Elsevier Science. **Netherlands**, 1992.

SEUNG JUNE OH; JEONGYOON KANG; JEONG YUN JEONG; KYUNG HOON LEE; SUNG JOON KIM; YOUNG SHIN CHUNG; EUN KYUNG HONG; KWANG MYUNG KIM, Pharmacological Effects of Berberine and Palmatine on the Prostatic and Urethral Smooth Muscle of the Rabbit. Fundamental Science for Neurourology. **Journal of the Korean Continence Society**, v.6, n. 2, p. 62-71, 2002.

SIDDIQUE, HIFZUR RAHMAN; MOHAMMAD, SALEEM. Beneficial health effects of lupeol triterpene: A review of preclinical studies. **Life Sciences**, v. 88, n. 7–8, p. 285–293, Fevereiro, 2011.

SILVA, JRA; AMARAL, A. C. F; SIANI AC; REZENDE CM; FELCMAN J; PINTO AC. Contribution to the study of *Himathanthus sucuuba*: latex macromolecules, microelements and carbohydrates. **Acta Amazonica**, n. 33, p. 105–10, 2003.

SINGH, PAYAL; DEEPIKA, ARORA; YOGESHWER, SHUKLA. Enhanced chemoprevention by the combined treatment of pterostilbene and lupeol in B[a]P-induced mouse skin tumorigenesis. **Food and Chemical Toxicology**, v. 99, p. 182-189, 2017.

SINGH, MAHESHWARI KUMARI; R. USHA; K. R. HITHA SHREE. O. BINDHU. Hemostatic potential of latex protease from *Tabernaemontana divaricata* (L.) R. Br. ex. Roem. and Schult. And *Artocarpus altilis* (Parkinson ex. F. A. Zorn) Forberg. **Journal Thrombosis and Thrombolysis**, v. 39, n. 1, p. 43-49, 2015.

SIRITAPETAWE, JARUWAN; PUNCHAPAT SOJIKUL; SOMPONG KLAYNONGSRUANG. Biochemical characterization of a new glycosylated protease from *Euphorbia cf. lactea* latex. **Plant Physiology and Biochemistry**, n. 92, p. 30-38, 2015.

SIRITAPETAWE, J; LIMPHIRAT, W.; KANTACHOT, C.; KONGMARK, C. The effects of metal ions in *Euphorbia cf. lactea* latex on the fibrinolytic activity of a plant protease. **Applied Biochemistry and Biotechnology**, v. 175, n.1, p. 232-242, 2015.

SIRITAPETAWE, J; TALABNIN C; VANICHTANANKUL J; SONGSIRIRITTHIGUL C; THUMANU K; CHEN C. J; KOMANASIN N. Characterization of the binding of a glycosylated serine protease from *Euphorbia cf. lactea* latex to human fibrinogen. **Biotechnology Applied Biochemistry**. 2017.

SHIVAPRASAD, H. V.; M. RIAZ; R. VENKATESH KUMAR; K. K. DHARMAPPA; SHAISTA TARANNUM; J. M. SIDDESHA; R. RAESH; B. S. VISHWANATH. Cysteine proteases from the Asclepiadaceae plants latex exhibited thrombin and plasmin like activities. **Journal Thrombosis Thrombolysis**, n. 28, p. 304-308, 2009.

SHIRWAIKAR, A; ARAJENDRAN, K, PUNITHA IS. In vitro antioxidant studies on the benzyl tetra isoquinoline alkaloid berberine. **Biological and Pharmacological Bulletin**, v. 29, n.9, p. 1906-1910, 2006.

SRIVASTAVA, AMIT KUMAR; SANJAY, MISHRA; WAHID ALI; YOGESHWER, SHUKLA. Protective effects of lupeol against mancozeb-induced genotoxicity in cultured human lymphocytes. **Phytomedicine**, v. 23, n. 7, p. 714-724, 2016.

SONKAR, KIRTI SHILA; MANENDRA PACHAURI; AMIT KUMAR; ANKITA SHURLA; MONIKA PATEL. Heme-peroxidase from medicinal plant *Artocarpus lakoocha*: Purification, characterization and wound healing studies. **Biocatalysis and Agricultural Biotechnology**, v. 4, p. 180-190, 2015.

SUFFREDINI IB, DALY DC. O Rio Negro como Cenário na Busca de Novos Medicamentos. **In Florestas do Rio Negro**, p. 257-281, 2004.

TANAKA, K.; WAKI, H.; IDO, Y.; AKITA, S.; YOSHIDA, Y.; YOSHIDA, T.; MATSUO, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. **Rapid Communications in Mass Spectrometry**, London, v. 2, n. 8, p. 151-153, Aug 1988.

THANKAMMA, L. Hevea latex as wound healer and pain killer. **Current Science**, v. 84, n. 8, p. 971-972, 2003.

THOMPSON, J.J.; Rays of Positive Electricity and their Application to Chemical Analysis. London: Longmans, **Green and Co. Ltd.**, 1913.

TRONCOSO-PONCE, MANUEL ADRIAN; KRISZTINA NIKOVICS; CHLOE MARCHIVE; LOIC LEPINIEC; SEBASTIEN BAUD. New insights on the organization and regulation of the fatty acid biosynthetic network in the model higher plant *Arabidopsis thaliana* **Biochimie**, n. 120, p. 3-8, 2016.

TSAI, FS; LIN, LW; WU, CR. Lupeol and Its Role in Chronic Diseases. **Advances in Experimental Medicine Biology**, v. 929, p.145-175, 2016.

VALKO, M; LEIBFRITZ, D; MONCO, L J; CRONIN, M. T; MAZUR M, TELSER J. Free radicals and antioxidants in normal physiological functions and human disease. **International Journal of Biochemistry & Cell Biology**, v.39, p.44-84, 2007.

VASCONCELOS, JF; TEIXEIRA MM; BARBOSA-FILHO JM; LÚCIO AS; ALMEIDA JR; DE QUEIROZ LP. The triterpenoid lupeol attenuates allergic airway inflammation in a murine model. **International Immunopharmacology**, n. 8, p. 1216–1221, 2008.

VILLEGAS, I. F.; FERNANDEZ, I. D.; MALDONADO, R. TORRES, A. ZAVALETA, A. J. Evaluation of the wound healing activity of selected traditional medicinal plants from Perú. **Journal Ethnopharmacol**, n. 55, p. 193-200, 1997.

WAHLER DANIELA; THOMAS COLGY; NATALIE A. KOWALSKI; ANNAER HARZEN; SANDRAY. WOTZKA; ANDREA HILLEBRAND; RAINER FISCHER; JOHANNES, JURGEN SCHMIDT; CHRISTIAN SCHULZE GRONOVER and DIRK PRÜFER. Proteomic analysis of latex from the rubber-producing plant *Taraxacum brevicorniculatum*. **Proteomic**, n. 12, p 901-905, 2012.

WANG, XUCHU; MINJING, SHI; XIULI, LU; RUIFENG ,MA; CHENGGONG, WU; ANPING GUO; MING PENG; WEIMIN TIAN. A method for protein extraction from different subcellular fractions of laticifer latex in *Hevea brasiliensis* compatible with 2-DE and MS. **Proteome Science**, v. 8, n. 35, 2010.

YANG, SONG; MENG-MENG DING; FANG CHEN; YING XU. Proteomic analysis of latex from *Jatropha curcas* L. stems and comparison of two classic proteomic sample isolation methods: The phenol extraction and the TCA/acetone extraction. **Electronic Journal of Biotechnology**, n. 27, p. 14–24, 2017.

ZHIYOU CAI; CHUANLING WANG; WENMING YANG. Role of berberine in Alzheimer's disease **Neuropsychiatr Dis Treat**, n. 12, p. 2509–2520, 2016.

ANEXOS

ANEXO 01

Sequências FASTA obtidas a partir dos peptídeos das proteínas por MS/MS

>gi|743891207|ref|XP_011039282.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Populus euphratica*]

MERSSFSRLRLLLLLILLVSLARSENSSSAHKMFLQCFSSHMQHKSYSSEVILTKNSSAYSSVLQSSFRNFR
 FLNTSTLKPQFIITPFNEFEIQAAIVCAKKYDIQIRVRSGGHDYEGLSFLSYQEFVLIDLAELSFISVDI
 ENETAWVGAGASTGELYRYIAEKSEVHAFAPAGSCPTVGVGGHFSGGGFGTIFRKYGLAADNVIDARIVDA
 NGRILDR~~ESMGEDLFWAIR~~GGGAASFGVVF~~SWKVR~~LVSVPPTVTVFNIEKTLQQGATNLLHKWQNI~~GDKL~~
 HEDLFLHATI~~AVATSSPHG~~NKTIQVSFASLFLGRAGELLPMQDSFPELGLMSENCSEMSWIQSVLYFGG
 FSPSDSLDVLLSRTAQFKGFFK~~GKSDYVKEPI~~SETGLEGLYKRLLEEEASTLILTPYGGRM~~SGISD~~SETP
 FPHRSGNIFEIQYIITWDVEEETEKNLKW~~MRKLYAY~~MAPYVSDSPRAAYLN~~YRDLDLGR~~NNYGN~~T~~SFAKA
 SVWGLKYFKNNFKRLARVKTATDPSNFFRNEQSIPVLQ~~RKR~~NLK

>AGL44335.1 FAD-dependent oxidoreductase [*Papaver somniferum*]

MNMIRSTQSSSSLILLYAFLLLSISLVTSSSEYEDDYNFLQCLSQHSDPSILTYTSKDSNFSSVLFSTI
 QSLRFYSPAIRKPRVIVT~~PLKESHVQASVIC~~SKRHGFQIRVRSGGHDYEGLSYVSDVPFVVVDLSN~~LRSI~~
 KIDVENSTAWVESGAI~~TGELYRYIAEK~~SRNLGFP~~SVCPTVGV~~GGSFQGGGYGNMLRKYGLAADNVIDAR
 IVDAQGRVLDK~~ESMGEDLFWAIR~~GGGMSFGIV~~SWKIRL~~VYPTNVTVFTINKNLDQGATKLVHRWQ~~EVA~~
 SELPHEL~~FVRV~~SISV~~VNSTTKEG~~QKTI~~LASFPS~~LYLGN~~TENLLAV~~MKERFPELGL~~ETKDC~~AEMT~~WIQ~~SQL
 YINGFPVNGSLDILL~~SRNQVKRYAKIK~~SDYVKEPI~~PETGLEGLW~~KKILEEK~~S~~VARMN~~FSPNGGR~~MAEISE
 CEIPFPHRQGNLYSIQYVVGWEGAGSEAAEPHIR~~WMREL~~HEYMPYV~~SISPREAY~~LN~~YRDID~~VGQSINGT
 ATYLEGMVWGSKYFKNNYERLVQVKS~~KVDPEN~~FFRNEQSIPAVSY

>gi|30315245|gb|AAP30841.1| nectarin 5, partial [*Nicotiana langsdorffii* x
Nicotiana sanderae]

YEGLSYVSEDPFVLIDLVGHRNITINLDDKTAWVETG~~STIGELYKISK~~SKTLGFPAGLCPTVGVGGHI
 SGGGTGVM~~L~~RKYGLAADNVIDARLMDANGRILDR~~KSMGEDLFWAIR~~GGGGNTFGLVLAWKIKLVDVPEKV
 IVFTIDKTLEQNATKLVHKWQYVSSKLHQDLYIRIF~~IHKDEQNI~~FLASFVSI~~FLGDIDR~~LLIMQENFPE
 LGLVRENCIEMSWIESTLYFAGFPRGESLDVLR~~SRGLP~~PTLYSEAKADYVQKPI~~S~~VQQLEGIWDFNAGE
 AKFEQMI~~FTPYGGRM~~DEISEYELP~~PHRPGN~~LYEIQYLMFWDEEGVEEAERHMR~~WMRRL~~YAHMEPLVSTS
 PRAAYIN~~YRDLDIGV~~NNKKGNTSYAQAKVWG~~IKYFKNN~~FDRLVQVKT~~KVDPSN~~VFRNEQSIPPLVEQE

>gi|747046464|ref|XP_011099776.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Sesamum indicum*]

MKTRTVVHMPSLVLSLCLVFLACTVFAHQAPHKFLHCVQSHNSEISISQVIYSKNNSSYLSNLLFSIQNI
 RFISPSMPKPAFIVIPFTESQIQAVVLCASKASGLQVRVRCGGHDYEGLSYSSHVPFVIIDLRNLSSIRID
 AKKRVAWVEGGALLGNLSYQIAKRSGNLAYPNGLCPTVGVGGHFGSGGGYGTLLRKYGLAADNVLDARIIIN
 ANGEILDRKSMGEDLFWAIRGGGAASFGVVTAWKIRLVKVPDIVTVFTVNRRTLEQNATDLVHRWQYVADK
 FDRDLFVRILVTRVNSSHHEKKTTVQAAFNSIFLGLKIDRLLPLMQESFPELGLVQEDCKEKRWIEAVLYF
 SDLPDGSTLDDLNRNPNPKTYKAKSDFVQVPIPKHGLKGLWKFFDEDEAADAQLIFAPYGGRMSEIPE
 SQIPFPHRAGNIYEIQHLAYWDNERNKQAERYINWLRRLYEYLTVPVSKNPRAAYLNRYRDLDLGVNNDGN
 TSYKQASVWGTQYFKNNFKRLVEVKSKVDPNSNFFRNEQSIPILST

>gi|848927876|ref|XP_012827618.1| PREDICTED: cannabidiolic acid synthase-like [*Erythranthe guttata*]

MKTPSISTLSFILFVIFSCSCASVSADGHEDFLECLSEEFHNYPSISNVITYTPINNSSSYSSILRFSIQN
 LRFTSESTPKPLVIITPEHESQIPPIIHCAKDSEVQIRTRSGGHDYEGLSYVSPVPFIIDDLINFSEITV
 DAVNKTAWVESGSTIGSLYYRIA EKSPVLGFPAGVCPTVGVGGHFGSGGGYGNLLRKYGLAADNVLDARLI
 DVNGRILDRQSMGEDLFWAIRGGGGASFGVILAWKVQLVDVPERVTVFTVQKTLEQNASRLIHRWQYVAS
 KFDEDLFIRII IARANSSSSQTGGNNVTLNASFNSIFLGGVDRLLPMMQESFPELGLVREDCTEMSWIES
 ILYFDGFPIENREVLNRTQRSVRYFKAKSDYVQTPIPEFALEGIWRLFFEPEADEAVVIFSPYGGRMDE
 VPESIPFPHRAGNLYKIQHLVYWEKEEDQNSGRYLSWIRRLYGYVAPYVSRFPRAAYINRYRDLDIGMNR
 NEGRTSYAEANVWGMKYFKGNFERLVRVKTMVDPRNFFRNEQSIPSLREEIDD

>gi|848927873|ref|XP_012827617.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Erythranthe guttata*]

MGNFAVVISILFILSISCLISCDETKYQTFIQICISDHSSGQNKILQRIHTPNSPAYVSLLESAAQNPRWT
 NSTAQRPFLLIAAPHREAEIRAVILCSKMFHQIRVKSGGHDYEGLSFRSETPLSFIMVDLSNLNRTKIDL
 EEEIAWIQTGVKLGQLYYEIANKSSTLAFPGGLYPTVGVGGHISGGGLGTLMRKYGIAADNVLDARVMDA
 NGVTLDRESMGEDLFWAIRGGGGASFCVILAWKLLVVRVPEKLTVFTVRRKFEPENLHLLQKWQNTAHKV
 SKDLFIRILLQTAPKKDTSLEKLVAVGYNGLFLGPADEFVSYMETIFPEFDLKIEDCFSAPAGNYSCSDR
 PCIKKECFQVPWIKSVLYFAGKKLDDPPEILLQQRVNKHSYSKGTSDFLKSPIPDEGWKMIHKMFNTDDH
 QMMIMDPLGGKIDEIWEDEIPFPHRKGNLFNQYLYKWEVNTKNESKHIKWMRNHLHKKMKPYVAQSPRT
 AYINYKDLDLGRNDENCSYSRAKIWGEKYFKGNFKRLARVKGKVDPDNFFRNEQSIPVLL

>gi|1009123039|ref|XP_015878332.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Ziziphus jujuba*]

MFCSSSLAFLFMVLLQSEISLAISTDSTLDKFITCAQVHSSVTIPISKTIFNPKNSSYTSVLQNPVNL
 RFIEPSVRKPEFIFFTPLHDAHVQAAVICSKQLGIQIRIRSGGHDEYEGVSYTSVFETRIVVDIGNLRSII
 VDIEKSSAWVQAGATNGEVYRIAEEKSIHAYPAGACTSLGIGGHITGGGYGSLHRK**YGLAADNVIDARI**
 VDANGRILDREAMGEDLFWAIRGGGGASFGIILWWKIKLVPVPETVTVFTVTKTLEEGATKLLYRWQOVA
 DKVDEDLLIRVLISPSNVSSSTRTISTTYDGFFLGNSDRLLQVIQESFPELGVTRKDCVELDWIGSVMYLS
 NNPSGTSVEFLLQRKSVFRTYFKAKSDFVRDPIPETALEGIWKLLLEEDSPVLLSPYGGKMSKISESET
 PFPHRSGTIFIFEYMTVWTDGEGKTEKHKGWIRKLYGYLTPYVSKFPREAYVNYRDFDLGVNKNSTSYI
 EASVWGKKYFKNNFDRDLVKVKTEVDPDNFFRHEQSIPPLPVWKN

>gi|1040819157|ref|XP_017239542.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Daucus carota subsp. sativus*]

MATSSILCSFILLIVFSSFSLQTSASFSRQTSGDSAQRVQCLTRYAKNSESITQVVFTPANASYNPILO
 LNLQNLRFNTSGTRKPLAIVTPVEDTQVQAVIYCARKNMNVTRGGGHDFEGVSYTADVPFVLLDMINY
 NRVNIDLKTSTAWVQSGISLGEFYKISQTSVDLAFPAGLVSTVGLTGLLGGGGYGMKRYALAADNTL
 DARIVDYNGKILDRK**SMGEDLFWAIR**GGDPASFCVILELKLQVLPVVKLVTYFAVQRTLEQNGSALFQKW
 QSTAVNVFPRDLVRRVVVDITITSNSSAREDKKTVRVIFQSLYLGKIDTLLPIMK**EYFPELGLVREDCVET**
 SWIKTAPMFSFFPVGTDPKILLNKTA PTRNPVKIKSSFTTQPISLEGLNGIWDLWLKQPVQTTLIQYTPF
 GGMNEFAESALPFPHRPGVLYMINIAVTLNQNAEATLQWINDLFKYYPYVTKNPRTSYVNYRDADLGT
 GSKTYQEASIWGRKYYKNNFDRDLVKIKSVVDPQNFNHNKQSIPLLM

>gi|970039800|ref|XP_015081242.1| PREDICTED: tetrahydrocannabinolic acid synthase-like [*Solanum pennellii*]

MEKYFQFLPFLQKYQKTYTFSKITNIFYLLQYGQKNPRWLNSSSSHPIFIITPINESGIGAIIFCSQKFG
 IQIRVNSGGHDYEGLSFRFEKGERFVIVDLIKLDEIYTDKNEQTAWIQGTIGQFYIEIAKKGENLAFP
 GGLYPSVSGGKISGGGIGTLMRK**FGLSADNVLDAR**VMNVNGEILDR**ESMGEDLFWAIR**GGGGSSFCVIL
 AWKVKLVRVPLKLTVFTVRKHLHGETINLLQRWQNVSHELSDLFIRVLIQNMGIGSKKIVQVSFQGLFL
 GRMNELLPMLNQTFPEFGLVTKDCFQDPVVNCAIELPCIKKECFEVP

>gi|460393282|ref|XP_004242239.1| PREDICTED: cannabidiolic acid synthase-like [*Solanum lycopersicum*]

MKISWFSFIFVLLVLSASWSTLADNHEEFIQCLSHNNQTSSNIYTPNNSSFQSILQFSIQNLRFNTSET
 PKPLVIVTPVSESEVQRVILCAKKTGMHVRVRGAGHDYEGLSYVSEVPFVIVDLINLRTINVDVNDKSAW
 VEAGSTIGELYRIAIEKSKTLGFPAGVCPTVGVGGHFGGGYGVMLRKYGLAADNIVDARLIDANGRILD
 RASMGEDLFWAIRGGGNSFGLVLAWKVKLVDPVEIVTVFTLDKTLQATKLVHKWQYVAPRFHQDLFI
 RILVSRNLSSNQDDNNQQTIVASFNSIFLGGIDRLLPIMQENFPELGLRRADCIEMSWIESILYFAGFP
 TDGPLDVLLSRVQLSTRYFKAKSDYVYQPIPEGGLDGIWRFFFEDEAQSSQVILSPYGGRMDEISPSAIP
 FPHRAGNLYKIQHLVYWDEEGEEVAERHISWIRRLYSYMAPFVSKSPRAAYINYRDLDIGVNNIKGYTSY
 VQAKVWGIKYFKNNFDRLVHVTKTKVDPSNFFRNEQSIPSLTQWKNKGE

>A6P6V9.1 RecName: Full=Cannabidiolic acid synthase; AltName: Full=CBDA synthase; Flags: Precursor

MKCSTFSFWFVCKIIFFFFSFNITSIANPRENFKCFSQYIPNNATNLKLVYTQNNPLYMSVLNSTIHN
 LRFTSDTTPKPLVIVTPSHVSHIQGTILCSKKVGLQIRTRSGGHDSEGMSYISQVPFVIVDLRNMRSIKI
 DVHSQTAWVEAGATLGEVYYWVNEKNENLSLAAGYCPTVCAGGHFGGGYGPLMRNYGLAADNII DAHLV
 NVHGKVLDRKSMGEDLFWALRGGGAESFGIIVAWKIRLVAVPKSTMFSVKKIMEIHELKLVNWKWQNIAY
 KYDKDLLMTHFITRNI TDNQGNKTAIHTYFSSVFLGGVDSLVDLMNKSFPPELGIKKTDRCQLSWIDTI
 IFYSGVVNYD TDNFNKEILLDRSAGQNGAFKIKLDYVKKPI PESVVFVQILEKLYEEDIGAGMYALYPYGG
 IMDEISESAIPFPHRAGILYELWYICSWEKQEDNEKHLNWI RNINFMTPYVSKNPRLAYLNYRDLDIGI
 NDPKNPNNYTQARIWGEKYFGKNFDRLVKVKTLVDPNNFFRNEQSIPPLPRHRH

>Q33DQ2.1 RecName: Full=Inactive tetrahydrocannabinolic acid synthase; Flags: Precursor

MNCSAFSEFWFVCKIIFFFLSEFNISIANPQENFLKCFSEYIPNNPANPKFIYTQHDQLYMSVLNSTIQN
 LRFTSDTTPKPLVIVTPSNVSHIQASILCSKKVGLQIRTRSGGHDAGLSYISQVPFAIVDLRNMHTVKV
 DIHSQTAWVEAGATLGEVYYWINEMNENFSFPGGYCPTVGVGGHFGGGY GALMRNYGLAADNII DAHLV
 NVDGKVLDRKSMGEDLFWAIRGGGGENFGIIAAWKIKLVVPSKATIFSVKKNMEIHGLVKLFNKWQONIA
 YKYDKDLMLTTHFRTRNITDNHGKNTTVHGYFSSIFLGGVDSLVDLMNKSFPPELGIKKTDCKELSWIDT
 TIFYSGVVNYNTANFKKEILLDRSAGKKTAFS IKLDYVKKLIPETAMVKILEKLYEEEVGVMYVLYPYG
 GIMDEISESAIPFPHRAGIMYELWYTATWEKQEDNEKHINWVRSVYNFTTPYVSNPRLAYLNYRDLDLG
 KTNPEPNNY TQARIWGEKYFGKNFNRLVKVKTKADPNNFFRNEQSIPPLPPHHH

13>gi|729294682|ref|XP_010537313.1| PREDICTED: reticuline oxidase-like protein [*Tarenaya hassleriana*]

MAFDTMKGFI PGFVVCLLVLASDASAVKHDVEGFLGCLPRQASPSSPISGAVYTPKNSSFDSVFSYTKN
 KRFLNPRFTKPVAIVTAQHVSHPATVVC AKHHGLQIRVRSGGHDYEGLSYTSVVPFVILDMFNLRIDV
 DLYSETAWVQAGATMGELYTKIADASKLRAF PAGVCPTLGAGGHISGGYGNLIRKYGISVDHVVDALLV
 DVNGRVLNR TNMGEDLFWAIRGSGGASFGVILAWKIKLVTVPETVTVFRVNKTLEQGATDVLYKWQLVAS
 KLPEELFLRAMPEVSRGRDGNRTIVVLFIAQFLGPAENLLAI VKNLPELGLKRQDCIEMSWLNSTVFWA
 DFPAGTPTTVLLDRPTSPGISFKSKSDYVKRPIKPEGMEKMWKSMLKFNSLWMQWNPYGGVMDRIPETST

14>gi|695072053|ref|XP_009383129.1| PREDICTED: reticuline oxidase-like
[*Musa acuminata subsp. malaccensis*]

MAESYRHHVILLLLSCCCFFVVSFASYPVAPSGFDGVRKLVLCCLASAGVDNYTVPTADPLISDTSLYYFLN
FSIQNLRFAFRPGLARPAAIIVLPGNRTHLRSTVLCRAAGFGVIRIRSGGHSYEGLSYSEDSGVAAPFVVV
DLMRLNRVRVDPESTRATWVESGATLGETYHAI AASSDSLAFSAGSCPTVGS GGHIAGGGFGLLSRKYGLA
ADNVVDAVLIDADGRVMDRESMGEDIFWAIRGGGGGGWGAVYAWRIQLLP IPARVTGFIVNRPGSTRLVA
ELVHKWQLVAPSLPDEFYLSAFVAGLPELDRVAMSATFKGFYLGPKSEAVSIMARSPFELGLVDTDCHE
TSWIESVLFSSGLPNGSTVSDLKDRILRGKKSFKAKSDFVRIPIKSNLTRAFLDLSQEPKAYLIMDPYG
GAMARIRSDHIPFPHRSGNHLYIQYMI EW TSEEEASSEQHLEWIRGYYNHMGD YVSKGPRAAYVNYLDLDL
LGTNRWTVGMGDLNDR TADARSWGEKYFLANYDRLVRAKTA VDPYDVFNNAQSIPPNSSSEQKDMISRHIA
TEVHIS

15>gi|1040881446|ref|XP_017250456.1| PREDICTED: cannabidiolic acid
synthase-like 1 [*Daucus carota subsp. sativus*]

MKAARAPTS LTHFSLVFLSSFHLLFSFSAAEHYPGQAFIQCLIFQSNDSIANVVYSPNNASYEQIYEFSL
RNPRLNNSNRLQPQVIVTPVSESQVQAVRCAQQNHLRIRTRSGGHDFEGLSYSSTYDIPFVLLDMINLR
NVSVNATARTATIGAGATLGEVYWIYRASGTLA FPAGVWSTVGATGLICGGGYGVLRRKYGFAADNVLN
VRIVDVYGNILDRKSMGEDLFWALRGGGCSSFGVLSWELNLVEVPETVAIFNISRTLEQNATQMILPFQ
TVARNLPVEVDLRVLMSTTISNTSVRPDNKTVRLSFTSTYLGPAD ELYKIIDKNLPEIGLLRSDIQDVSW
IQAIMQGSFFPLFDEAPT PEDLLDR TFLADIPFKGKSD FVREPISEQINGLWDKLEVG PGETTLVFTP
YGGVLDEYLES AIPF PNRNGTLYMVYMRVLWVGNTTQKLEWIRSLYGYLRPYVSRNPRRAYNYNDLDL G
MNNARGPINYITARRWGRSYFNNNFMR LVGVKTRVDPLNFFRHEQSIPPFSLAMVSDM

16>gi|20563190|gb|AAM27915.1|AF364866_1 carbohydrate oxidase [*Helianthus
annuus*]

MNNSRSVFLLVLALSFCVVSFGALSSIFDVTSTSEDFITCLQSNNSNNVTTISQLVFTPANTS YIPIWQAAA
DPIRFNKSYIPKPSVIVTPTDETQIQ TALLCAKKHGYEFRI RDGGHDFEGNSYTANAPFVMLDLVNMRAI
EINVENRTALVQGGALLGELYTISQKTD TLYFPAGIWAGVGVSGFLSGGGYGNLLRKYGLGADNVLDIR
FMDVNGNILD RKSMGEDLFWALRGGGASSFGIVLQWKLNLV PPERVTLFSVSYTLEQGATDIFHKYQYV
LPKFDRDLLIRVQLNTEYIGNTTQKTVRILFHGIYQGNIDTLLPLLNQSFPELNV TREVCQEV RMVQTTL
EFGGFNISTPTSVLANRSAIPKLSFKGKSDYVRTPIPRSGLRKLRKMFENDNSQTLFMYTFGGKMEEYS
DTAIPYPHRAGVLYQVFKRVDFVDQPSDKTLISLRRLAWLRSFDKTL EPYVTSNPREAYMNYNDLDLGF D
SAAAYEEASEWGERYWKRENFKKLIRIKAKVDPENFFRHPQSI PVFSRPLSDM

17>gi|590614781|ref|XP_007023039.1| FAD-binding Berberine family protein
[*Theobroma cacao*]

MTCIVFKRFLFGTTHSSLNLFKFSYKFLHLLAFGIIQKTMGSLRPAAVVSLLSVLLLSISLQGTSDSAQE
TFLQCLLDNSHPSYPISEAIFTPQNPSYSSVLQSSIRNLRNFNETFTPKPFLILTAKHESHIQAAIVCARK
DNIQMKIRSGGHDEGLSYVATVPPFFVLDMFNLRSIDVDVANETAWVQTGATLGEVYRRISEKSKTHGFP
AGVCPTVGVGGHFGGGGYGNMMRKYGLSVDNIVDAYFIDVNGRLHDKSMGEDIFWAIRGGGAASFGVVL
AYKIKLVRVPETVTVFRVEKTLEENATDIVDQWQHVADKLPEDLFVRLVLDVNVSSRNTGEKTVRAAFIS
LFLGDSERLLSIMNERFPALGLTQSDCIETSWVQSVLFWTNIPIETETAILLDRTPSSLVFLKRKSDYVK
KPIPKAGLEWLWKRMIELQVPQLLFNPYGGRMAEIPSTATPFPHRAGNLWKIQYVTNWNNEAGTEAADHYI
GLTRKLHGYMTRFVSKNPREAFLNYRDIDLGVNHNGRQSYMGRVYGIKYFKGNFNRLVQIKTRVDPGNF
FRNEQSIPTLPYKGN

18>gi|1040922945|ref|XP_017223696.1| PREDICTED: flavin-dependent
oxidoreductase FOX2-like [*Daucus carota subsp. sativus*]

MKLPTSFLPLLLLLSTFSFSSSATKTDDFLNCLAKSSDSTTISKLVYTPANSTFDAALTYSSINNLRFAQA
STPKPLVIVTPTTESQIQNVIYCTKKTGLEMRIRSGGHSFEGFSYVSSLQFIVLDRNINKVTPDMSTAT
AWVDSGVTNGELYYYISKATSAYGFPSGLWSNVGVGGILSGGGYGLRKKYGLAGDQVIDARLIDANGRI
LTRKTMGENLFWAIRGGGGSGFVVVSWRVNLVPPPIVTVFRVFRVLEEDMTNIFYKWQSVAPVLPKEL
DIRCNGQVILSNSSTRSDKKTQMNFESLYLGPASEVLAIMGERLPELGLVREDLFEVSYIQAMVFFSQF
PIQAPPEILLDKTILPRPAFKGRSDFFKKPMPIEGLLGLWDYMFQLPDNQAFQYTPYGGRMNEISATAL
PFPYRAGYLYMNFFAVTCPTREKCAEETARMDWVRTVDKYLTPYVTSNPRSAYVNYVNVWVGQNNPTGS
TSYAQASQWGKRYFGVNFDKLVMIKSLADPFNFRHEQSIPVFSLWSDM

>XP_006367558.1 PREDICTED: glucan endo-1,3-beta-glucosidase A-like [*Solanum
tuberosum*]

MSFLSFV GASFLLVGLLIQMTGAQSIGVCYKGIANNLPSDQDVKNLYNSNGIKKMRIYYPDTNVFNALKG
SDIEIILDVNPQDLEALANPSSANGWVQDNIISNFPDVKFKYIAVGNEVDPVKFKYIAVGNEIDPDTNTG
QYTQYVGPAMENVYNALTSAGLQDQINVSTATYLGLLTNTYPPSDSIFREEYKSFINPIIEFLARNNLPL
LANIYPYFGHIDNTNDVPLSYALFNDQGTNSAGYQNLFDALDLSMYFATEKLGQNIIEIVSESGWPSEG
HPAITLENAQTYTNLINHVKGGVGTPKKPGRAIETYLFAMFDENQKDGQSSEQHFLFYDPQRAKYQLN
FN

>ANN02876.1 UGT73AL1 [*Punica granatum*]

MANEGETNDTKSTTSKLIHIFFPFSPKSGHTIPMIDMAKLFSSRGVKSTLISTPHNEPSFLRSIERTQKLG
FDIGVVTVKLPLEKVLPEHCQILDSTNSPDMINKYWMAIRMLDKQLEKLIKEHCACVITDTFLPWTVD
VAAKFGIPRLAFHGTSHFAMCALECTRLYKPHLNISSDSEPFVIPNFPGEITMTRAQLPDFIKEDTEFSK
LYIEMMESELRSYGVIMNSFKELEPVYTDHYRDVLGRRSWRVGPVLLCNQDTEKLRGTRATIGEDSCL
KWLDSREAGSVVYICFGSRTDFSASQLHEIAEGLEASRQPFVWVVKKDESVEEGKEEWLPQGYEERMQK
GLIIRGWAPQVLILHKKAVSGFVTHCGWNSTLEGITAGVRMVTWPVAAEQFCNEKLITEVLRIGVPVGAK
QWMVKVGDVSESEKVEKAVKRIMIGDEAEEMRVRAIELSKIAKKAVEEGSSHSDLTALLES

>EMT60941.1 Hyaluronan synthase 2 [*Fusarium oxysporum* f. sp. cubense race 4]

MSGKQATVEVSAGVESSSASSSDSIKASRFDSASVFDTPVRKPLRINPFSAAINIIIAFGLIAALITGYII
GLRVLDLGITSVGLYGTVLITEYLIQLTCAIFNRFIDIRIVKNVRKKTANAQAAGDCEKSASNPADNLL
NPDGDVSMMAVVGREDEQAWRQCLRSQTQTALPKCVVGVVDGNDEPDLSMAQAFVDEFQDQKATLIHLP
ILLSELHRRTYFDHVPADTRFFLVKFWHYLVGRYRPGHEAALATAREVVVQQVLEWEEKWNISSLKAVCF
AOPHGHKRTAMFTAFAMSLYAFRTRDAIFTTDSDTLVQSNGLDEMLTLRSPQIGGVTADVKIWNRAES
LLARMCAARYWFAFNIERACQSLWRCVCGCLSGPMSMYRASDLETILGPWNLQTFGGKETTFGDDRHLTNQ
LLAHGLRTRYTHRTWCDESEPTSFVRWVAQQTRWSKSFRESFWFPASFAYQSPWMLVEMTVQALYPFIL
IATVLHFLFDTSDDSSPWRPIIWLVTMFGVALFKSVLAVLISKDPWLLLFSAYGFYFFGLLPSKIYALIT
MNQTGWGTSARSSSERKRGQSFLQRSFHICHLVIWYTAVFVGIGFFVWRVFGNPLYILIGAVALIPIAFL
YWQGSFSLGGLFRNKVASETSTPIQVDDEMLPQKPEKALSISSKRSSNTLVKKTITIVSVKDVTDEVEK
PVNSASAA

>ADW11233.1 subtilisin-like protease 2 [*Phaseolus vulgaris*]

MLNVGTLQWSTLQRRIPVPIRIRPCVADSMKAPKASSLTLISSPILTFVYSLVPDLSHPPSDAPRTYIVH
VAQSQKPRFLTHHNWYTSILHLPSSHPATLLYTTRAAAGFSVRITPSQLSHLRRHPAVLAVEPEPGPPH
PPPTHTPRFLGLAESFGLWPNSDYADDVIVGVLDGTGIWPELRSFSDNLSVPSTWKGSCVSRDFPAS
SCNRKIIGAKAFYKGYEAYLDGPIDESAESKSPRDETEGHGHTSSTAAGGVSNASLFHYAQGEARGMAT
KARIAAYKICWKYGCDFSDILAAMDEAVADGVHVISLSVSGSSGYAPQYFRDSIALGAFGAARHNVLVSCS
AGNSGPGPFTAVNIAPWILTVGASTIDREFPADVILGDGRVFGGVSLYYGESLPDFQLRLVYAKDCGNRY
CYLGSLEASKVQVKIVVCDRGGNARVEKGSVAVKLAGAGGLGVIMANTAESGEELLADAHLLAATMVGQIA
GDEIKKYIRLSQYPTATIEFKGTVIGGSPSAPQVASFSSRGNHNTSEILKPDVIAPGVNLAGWTGRVG
PTDLIDIPRRVEFNIISGTSMSCPHASGIAALLRKAYPEWSPAAIKSALMTTAYNVDNSGGNIKDLGTGK
ESNPFTHGAGHVDPNRALNPGLVYDSINDYLAFLCSIGYDANQIAVFTREPAAANPCEGKVGRTGRLAS
PGDLNYPFSFVELGRGSDLVKYKRVVTNVGSSVDAVYTVKVNAPPGVDVTVAPNTLVFSGENKTQAFEVA
FSRVTPATSDSFGSIEWTDGSHVVRSPIAVRWSGDSSSSL

>ADM25048.1 Rpl-like protein, partial [*Zea mays* subsp. *parviglumis*]

AQKSPHRGKLESWLRRLKEAFYDAEDLLDEHEYNVLKAkakSGKGPLLREDESSSTATVMKPFHSAMNR
ARNLLPGNRRLISKMNELKAILTEAKQLRDLGLPHGNTVEWPAAAPTSVPTTSLPTSKVFRDRDRDR
IVKFPLGKTTTAEASSTKYSGLAIVGLGGMGKSTLAQYVYNDKRIEFCFVDMWICISRKLDVHRHTREI
IESAKKGECPRVDNLDTLQCKLRDILQESQKFLVLDDVWFEEKSHNETEWELFLAPLVSKQSGSKVLVTS
QSGTLPAAICCEQEHVHLENMDDTEFLALFKHHAFSGAEIKDQLLRTKLEDTAEEIAKRLGQCPLAAKV
LGSRLCRKKDIAEWKAALKLGDLSDPFTSLLWSYEKLDPRLQRCFLYCSLLPKGHGYRPEELVHLWVAEG
FVGSCNLSRRTLEEVGMDYFNDMVSVSFFQLVSQMYCDSYYVMHDILHDFAESLSREDCFRLEDDNVTEI
PCTVRHLSVHVQSMQKHKQIICKLYHLRTIICIDPLMDGPSDIFDGMLRNQRKLRVLSLSFYSSSKLPES
IGELKHLRYLNLIRTLVSELPTSLCTLYHLQLLWLNHMVENLPDKLCNLRKLRHLGAYVNDFAIEKPICQ
ILNIGKLTSLQHIYVFSVQKQGYELRQLKDLNELGGSLLKVNLENVIGKDEAVESKLYLKSRLKELAFE
WSSENGMDAMDILEGLRPPPQLSKLRIKGYRSDTYPGWLLERSYFENLESFELSNCSSLEGLPPDTELLR
NCSRLRINFVPNLKELSNLPAGLTDLSIDWCPLLMFITNNELGQHDRENI I IKADELASKLTMWEVDS
GK**KVRSILSK**DYSSLKQLMTLMMDDDISKHLQIIESGLEEREDKVMKENI I KAWLFCHEQRIRFIYGR
MEMSLVLPGLYKLSLSSCIITDEALAICLSGLTSLRTELELYNMTLTTLTLPSEAFQOMTKLKCFAISGC
WCLKSLGGLHAAPSLSALDCWDCPSLELARGAELMPLNLASYLDIQCILAADSFTNYVPDLKQLTIINC
RCSPSLSIGHLTSLESLSQLIGLDPDLYFVEGLSSHLKRLKLGDVANLTAKCFSQFRVMESLTVSSSVLLN
QMLMAEGFMVPPNLEFLYCKEPSILFEAPANLSSVKCLNFSLCETESLPRNLKSLSSLESLEIGFCPNIA

SLPDLPSSLERI

>AJD20222.1 norcoclaurine 6-O-methyltransferase [Sinopodophyllum hexandrum]
 MEAQKENISSQAKLWNFIYGFAESLVLKCAVELDFANIIHNHGKPMTLSELASQLPVMQPVNTNSLYRVM
 RYLVHINIFTKTLDENDGETKYGLAAPAKFLVKGWDNCMVGSILGITDKVFMEPWYYIKDELAPGTGTAF
 ELALGKDIWEYMGENPEKNKLFNAAMACDTSMIMSALISECKDKFNGIRTLVDVGGGTGTAARNIARAFP
 NIKCTVYDLPHVIADSPVYPEINRVSGDMFKCIPNADAILMKCILHDWEDKECIEILKRCKEAVPVDGGK
 VIIIDVVLGDGESEHPYTKVRLNSDLDMMLNTEGKERTKEGWKKLFKAAGYRDYNTQISALQSVIEAFPY

>AHZ63910.1 phototropin [Desmidium aptogonum]
 MGPPPEGSSLVKGTTHDKVAGGGSVPTARRYSLGLSPEDDPRRSSNSQAGKVLGSKSELRDALTAQQSFC
 MVDATKPDLPVLFASEGFYQLTGYSALETIGKNPRFLQGQETDRAEVAKLKQAIQAGESWCGRLLNYKKD
 GTPFWNLLTVTPVKDDGGKVVKFIGMLVEVTKYTEGAKDKETRPNALPVSLIKYDARQKEEAESSASELL
 EEAQRHPLLDMSGGKDKGEGGMDKMLQPKVEEEGGGSAEKKPERRKSFMSILSKKDPKSTSQGAPQG
 GEAQSGAAVAEEEDGDVDRKNRKGMDLATTLERIQKNFVITDPRLPDNPPIIFASDDFLELTEYSREEIIG
 RNCRFLQGPDTPNPKTVQKIREAINNQEDITVQLLNYTKSGKPFWNLFHLQAVKDNKGMQLQYFIGVQLDAS
 QYLDPNIQGLEDRFAQEGERKIVETANNIDGAVRELADPGADAKDLWSIHSVPAVVKPHKRQDPAWQAIL
 DVVAKDGRGLGKHFRPIKPLGAGDTGSVHLVELRDTGRLFAMKAMDKEVMINRNKVRHACTERDILGRLD
 HPFLPTLYASFQATATHVCLITEFCSSGELYGVLERQKGRFPENVAKVFAAEVLLALEFLHTQGVVYRDL
 KPENVLLMESGHAMLSDFDLSFLSSSTPKLELSPDDAKKKPKKLLKKKGPLSDAERAQWKAELNAVVPML
 VAEPSTSSNSFVGTTEEYIAPEIINGTGHSAPVDWWAFGIFLHEMLYGKTPFRGRNRQRTFTNVLMKELTF
 PDTVPVSDAKALIRLLLLERDPEKRLGSKKGAAEIRAHPPFLDIDWALIRHKPPPTPSIPMKLITTEADS
 ARQSMVAEEELDWEHEARPSESLDYG

>P54209.1 RecName: Full=Cation-transporting ATPase CA1
 MVSHASSGRPSSRDGMVYLGGLGMQDAYSSSEVQEVAAFYHVDLDRGLSDRDVQQARIKYGRNQMEAEQST
 PLWKLILKQFDDLLVKILLGAAIVDFIIAISEGESIQSGLIEPMVILLILVANATVGVVTERNAEKAIEQ
 LKSYEADDATVLRNGQLQLIPADIVPGDIVEAVGNKVPADTRVSHIYTTSLKIDQSLLTGESQAVEKH
 TEVVHNEQAVYQDKLNMLFSGTLVVAGRARGIVVGTGSNTAIGKIRDAMGVEEDVVTPLKAKLDEFGALL
 SKVIAGICVLVWVVNINRFNDPALGGWFQGAIHVFYKIAVALAVAAIPEGLPAVVTTCLALGTRKMARHNA
 IVRTLPSVETLGCTTVICSDKTGTLTTNQMSVIKVAAVQSSSSQLAEFDVTGTTFSPEGMVLGPGGVVLR
 QPADTPCLAHAAQCAALCNDSSQVFVAQKTGTLQRIGESTEIALRVFAEKIGLPSSIRPDRPISRSQFGTN
 NFWQEDVERLALLEFSRDRKMSVVLKGSQRQHNIIWSKGAPEFVLRKCSHVLANNEGAVPLTDNMRQAI
 LSDMQAFGSRQALRCLALAFKSVPTTTTKLDYSDESGLTFIGLLGMHDPPRPECRSALSTCHNAGIKVIM
 VTGDNKGTAEAVARQVGALSPSTALAGSDDENLGISYTGREFEEMGALGQAAATRNLVLSRVEPMHKL
 RLVELLKAQGHVVAMTGDGVNDAPALLRADIGIAMSGTAVAKHAADMVLGDDNFATIVFAVAEGRVIFN
 NTKQFIRYMISSNIGEVAIFLAALLGLPEVLTVPVQLLWVNLVTDGLPATALGFNRADKMMARGPRRVD
 DPIVNGWLFRLYLIIGMYVGIVTVYGFIIWYISFPEGGNMTWSQLTHFQACASQPGGAKDCEVFHSHKPT
 TISMSVLVVVEMFNALNLSLSEDSSLLRIPPWDNKWLVGAIATSMALHFGILYTGASAMFGVTGLSFAEWT
 MVIKLSAPVILVDEIMKAWSRRRQRHPASSRGGPVSLMEIQVPLTSSSRDEAALKL

>XP_006658000.1 PREDICTED: serine/arginine-rich SC35-like splicing factor SCL33 [*Oryza brachyantha*]

MGRGYSYSPSPPPRSYRRRASSPIPRGRYGGGRDLPTSLLVNLRDRCPEDIRRFPGQFGRLLKDVYIP
RDYYSGEPRGFVQYDPDDAADAKY~~YMDGQTI~~LGREIAVVF~~AEENRKKPAEMRARDRISGRGRSYDQ~~
RYSRSPRYSPPPRGYSPPRRGRSPYRSPYSRSPSPRYARRMRERSYSPVDSRSRSRSPIDEGYGG
TRRERSLSVSG

>NP_191795.1 Carbohydrate-binding-like fold [*Arabidopsis thaliana*]

MAASRKICHSLIVFLIAISTVYGVSAADSIKCGGFVEASSSLVRSRKGSDGKLD~~FSHITVELQTV~~DGLVK
DSTQCAPNGYYFIPVYDKGSFILKINGPDGWSWNPDKVTVVDDSSC~~NNDDINFHFTGFTLSGKVLGAV~~
GGESCLIKNGGPADVNVELLSSD~~GSEDPVASVLTSSDGSYLFKNIIPGTYNIRASHPELQVEVRGSTEVE~~
LGFANGMVDDIFFVLGYDLKGSVVAQGNPILGVHIYLHSD~~DDVSMVDCPQGS~~GDAAGERKSLCHAVS~~DAEG~~
IFSFKSI~~PCGKYELVPHYK~~GENTVFDVSPVMPVSVEHQHVTVPQKFQVTGFSIGGRVVDGNSV~~GEVVK~~
ILVDGSLRSVTDKEGYKLDQPAKL~~VVTQPLLVNFLRLLES~~RVKNMARFLKVTSNQYTI~~DAVKEHYKFDK~~
LKKFMVLPNMASLPDINAVSYDICGVVRMFGSRHKAKVALTHGPTNVK~~PQMKLTD~~ETGAFCFEVP~~PGEYR~~
LSALAAATPKGASELLFLPAYVDVAVKSP~~LLNIEFSQARVNVHGSVTC~~KEKCGPSVSVVLVGAAGDR~~DKKT~~
VVL~~TDESSQFL~~FS~~DILPGKYR~~VEVKSISPEAA~~SDSDWCWDRSSIDVNVGTEDIK~~GIEFVQKGYWINIIS
THEVDARIAHPDGSP~~TSLKIKKGSQKICIES~~PGGHELQ~~LS~~SDSCMSFGS~~NSIKIDVSNPQPIHLKAEK~~YLL
KGLIN~~VSSSTIESELQENFIVDIQDKKGNVINTIAAKLASD~~GS~~GVY~~EY~~TWASL~~GEKISFV~~PQDS~~RGNV
EKKMLFYPKEIHAVVSKDGCQASVSPFTGRLGLYIQGSVSPPLPGVNIKIFA~~AKDSLIS~~LKKGEIAIET
STLSAGSFVAGPLYDDIPYATEASKPGYHIKRLGPY~~SFSCQKLGQISVRVNSK~~DNAETSIP~~PLLS~~LS~~SGD~~
HGYR~~NNSISGAGGLFV~~FD~~SLFPGNFYLRPLLKEYSFKP~~STLAIELNSGESSEAVFEATRVAYSAMGRVAL
LSGQPQEGVAIEARSDSKGYEETSDINGNYRLRGLHPDTAYVIKVS~~KKIGSANNQIERAS~~PESVSLQI
GYEDINGLDFLVFEQPETTILTCHVEGKQ~~NEDLNSNLLVEIK~~SAIDKSKIEN~~VFPLPLSNFFQVKGLPKG~~
KHLVQLKSSRPLISHKVESEIEVDFETNAQIHIGPLRYSIVADHQSQEVT~~PAAILPLVIGVSAIALFLS~~
IPRLKDIYQATVGISSPGFTTS~~AKREPRKAVARKKTF~~

>Q9SU86.1 RecName: Full=Cytidine deaminase 6

MKFVYTPSEAAEEGVRG~~PSDLPKLIDKAMSLARAPVSTFK~~VGAVGLTSSGEVFLGVNVEFPNLPLHHTIH
AEQFLVTNLALNSMKKLTHIAVSVTGTIFGAPCGHCRQFYQEMRNAPEIEILIKRPKDGIDEFMSL~~KSLM~~
PERFGPDSILPEDASLLEQRD~~NSLVLSDPEEICS~~DPEDCSHTKCRALAAANKSYAPYSKCP~~SGVALICG~~
GEVYKGWYIESVAYNPSLGPVEAALVDFVARGGGKEFNEITEVVLVEMKDVKVSQEATARTFLDKIAPK
DFKVLHCYKTKN

>AHB32114.1 serine hydroxymethyltransferase [*Camellia sinensis*]

MGSIQQPIWTKGSTFSLKSGFNGFPHQVK~~LNLVKRCR~~SSQLEGNLVTGKSPSSVSIPATKIGNGSCFI
DHGLSEADPEVRTIIDKEKQ~~RQLRSLELIA~~SENFTSRVMEAVGSCLTNKYSEGLPGKRYGGNEYIDEL
ETLCQERALAAFHLDGKKWGVNVQPLSGSPANFEVYTALLNPHDRIMGLDLPHGGHLSHGFMTPKRRVSG
TSIYFESMPYRLDESTGIVDYDMLEKTATLFRPKLIAGASAYPRDFYPRMRKIADAVGAF~~LMMDMAHI~~
SGLVAASV~~VADPF~~EYCDVVT~~TTTHKSLRGR~~PGMIFFKRETVHGV~~DLESAINNAV~~FPLQGGPHNHTIGG
LAVCLNHAQSPEFRAYQNEVVS~~NCRALASRLIE~~LG~~YKLV~~SGSDNHLVLV~~DLRPLGLD~~GARVEK~~VLDIAS~~
ITLNKNSVPGDKSALVPGGIRIGSPAMTTRGFKEKEFIVTADFIHEGVQISLEAKRSVPGSKLQDFMKFV
GSPDFPLMHRVSDLRRRVEELSTQFPMPGL

>EPS58436.1 geranyl pyrophosphate synthase [Genlisea aurea]
MHTGTLPLENAAHPYESSQRVSGRRVLFWSVSNVSVNGQQPQFQSNLSVEDHLDPFSLVADELSSLGN
RLRSMVVAEVPKLASAAEYFFKMGVEGKRFRPTVILLMATALDSSIGRQDALVGDFAELRARQQCIAEI
TEMIHVASLLHDDVLDADTRRGIGSLNFVMGNKLAVLAGDFLLSRACVALASLKNTEVVSLLAKVVEHL
VTGETMQMTSTADKRCSEYYMEKTFYKTASLISNSCKAIALLAGQTAEVSTLAYEYGNLGLAYQLIDD
VLDFTGTSASLGKGSLSDIRHGIITAPILFAMEEFPQLRVVDRGFENSNDVLDLAEYLKGSHGIERRE
LAVKHASLASAAISSLPKNDEEDVMRSR**RALVELTTR**VITRTRK

>CDY68509.1 BnaC04g54120D [Brassica napus]
MGNDLIVLAAGVEDGEIAAASGNASIEVRQSTVADGGDVDVGGVTGGGRGGGFDGNSRVWMTMRDLMTK
YPEYRGYANGLSNFAWAQAVQNKPLSEGLGKEYETREGGGDKIVIEDSDDEKEEGEELEEGEIDLDSTR
DDEMETESLVVLTSADELEDDRVRKERELETKVKLIRDVLESTSLVQAQIAFEGVCSRLLGALESREL
VSDNDDFPKRDITMQLSFASLQITINSVFTSMNNMSKELNKDTMSRLVSLVNDHCSRFLSSNQRIEIEAMN
QNLRRSAISLSAGASSEENVNRMTQTSNGDLFPKLNLSLEGTRRGAFYARSRLPLLDLHMDHDADSLPSP
TRETTPSLPVNGRHMMARPGFFPGKEGQTSEVAKVHHHPYESEALKAVSNYQKKFGVNSLFKTDDLPSPTP
SGEPNDGNGGTGGEVSSSVVASKKPGTLMTYGQDVPLPSTFSSRSMPVVSSTVPPHPLSIYGMSTPAGAT
QTVLASDQTVKPSAKSRDPRLR**LAKPDGAASVTISPR**VVPSAELVNQRKQKATSELFIDGPTWKRKSDN
DAQKATNIGGWLEDTESSGHPKLESKPRLIEAGVTSMKTSVMPTNAVTVTPKVTTATSTEALSSLFKDFA
ENPTMIMNILKMGQKQTVPEKAPQKPMDDPRRAAQLPGSSSVPPVAPPVSI PASNALPANFPQPGAPKDE
SGSIRMKPRDPRRILLGSTLQRTDSVAEKQSKLNDSSTLKKGKTEVLETPSQLVPRQSI SLNGTSNMRVSG
EPVRGKTPDFTKNLKNVADMSVLSQQVGNPLATTHAADLKTDKDQEEASVSAASVTAAGPTRSMNSWG
DVEHLFEGYDDKQRVAIQKERTRRLEEQKMFSGSKKLSLVLDLHLLNSAKFHEVETAHEAMLRKKEEQ
DRDKPYRHLFRFPHMGMWTKLRPGIWNFLEKASKLYELHLYTMGNKLYATEMAKLLDPEGVLFNGRVISK
GDDGDFLDGDERVPSKDLGVMGMESSVVIIDDSVRVWPHNKMNLIVVERYTYFPCRRQFGLLGPSSL
EVDFDERPEEGTLASTLAVIERIHKNFHSHTSLDEADVARNILASEQRKILAGCRIVFSRIFFVGEANPQL
HPLWQTAEQFGAVCTTQADEHITHVVTNSLGTDKVNWALSRGKFVVHPGW

>KMZ74127.1 Zinc finger CCCH domain-containing protein 4 [Zostera marina]
MSSRAQSLEPTCPIGRDQSMGTNEATRDPQSSSLSSPGNASVPPLSQGTSSDARSSFPARHTLPVFA
RIVDKIRENRVTLVIGDTGCGKSSQVPQFLLEEDIQPIILCTQPRRFVAVAIARMVAEARNCEVGGVGYH
IGHSNVSSSSSKIIFKTAGVLLQMRDRGLAALKYKVIILDEVHERSVESDLVLACVKQFMMKNKGLRVV
LMSATADISRYKDYFKDLGRSERVEVLAIPYVPQQVIFQRKVLYLEQISELVLDLKETTDKICITNCKDLG
DAGLKNEVHTLIHKLILHIHNNESDIEKSILVFLPTYLLERQWILLK**EHLVSVFK**VHILHRSIDTESALS
AMQIWKSHRKVILATNIAESSVTIPGVAFVIDSCRSLQVFDQNRNIDTPKLVVWVSKSQAEQRKGRTRT
CDGHIYRLVPGQFYSSFDDEFPAILRSLRQHVLVCCADSKAINDPKVLLQKLMPPYPVVVENAINQ
LLNIKALNKHPHTHKGSYEPTFYGRLLDSLPLTFDPSVLLLKFGDIGLLHEGIIISILMDIQPFPIIQPF
KLYLHEQYVANFFEDDKTIGKKENILMGNLSSYFQWQCFKDKLRMDRLKQLNGAEPKSSLSLIEHELEK
EWCYLNHLVPTTLNNSIDIYEDALCAIHRFRPMFLQADSLPSYDPYDFKHTCLIQSESHSIDSVNENP
ESTDRSKCTSLPFVTSNDFQVSYISEEFVKLIKEIRGHNTDEKYEDRKRIGTSGCAAATKICKYFIRGL
CNKNDQCIFSHSFQAKKPPCKFFFTLQGCYGNSCIFSHDNQLQNSSFMNAFLCSPEDSNERQTYDIFLH
LLPAASEGCILILNDKHLQITSILAKYIDPHKLIVSIPDQPSSEVSSSLIGLTKVSDTSMPSWLNLTETK
EDCTPWAQVRSILWFSDFDVDEAEKHHNLLKFFQCLAIKFLAGILPPHSVIITMNNLRFQSIEVEKLAR
ECFFFLTQSFLFDASTFGKISHDCAVKPMQASSTISYIFEIHLPPAHTMEVKDYSSSSNNNHIGLMYNT
EWQF

>KZV49446.1 14-3-3-like protein D-like [Dorcoceras hygrometricum]
 MFELPVDGLADVSEMPKDKIFDAKSIVSMTEEPVILSGLKSQMKIHYRLLCDIMAKSISVKAGSFNALT
 EKFSLMTAVVCRVKMNWGRVLFVILKMKMTSGTKQAKGFAIQISLLLANVPNLELGELSEFPSSKILTEK
 TVHRYIAVIDKSGAQDADAAPKVTKAPKKRTTAVPADVPVVKRKRRTTKKKDSSSKYNLELMAVAQEAVPI
 QMIEGSTAPAVDDTDDQFAAVDAFPADRPAAEIESVERVDEQLAAEPAVHVPAKEPAVEVSRVNVDDPET
 VIRQVLDQLDFVTEDEPVVAKVDRATLSEQFNESQSGISSRLHKIEQGLRDSLR**EHAEVFK**NLNFQGARQEG
 RTIDDVQTLCFNEFCKNILAQNASNFTGLADVRKEVQEVNAKVDIIASRVNEVRKNVEETKEALSHQLLE
 FQSQAQENQNILHAQLSELVNYINRGSADKKGESSRGPQQPPNVQNLESGQSSISLEETAERIREADRRQ
 AEMERERERQRRIIRRLSGSSKRRRGY

>XP_010095888.1 Translation factor GUF1-like protein [Morus notabilis]
 MAANYMTSSSSSSQALLLSTCHHHHHSRTTTLSSSLSSFLPSISKALQFSHSSSYSLYSSSRFSISCO
 AAASGTAVSDDLVAQVGRDRLKVPISNIKNFCIIAHIDHGKSTLADKLLQMTGTVQKREMKDQFLDNM
 DLERERGITIKLQAAR**MRYVFK**NEPYCLNLIDTPGHVDFSYEVSRLVACEGALLVVDASQGVQAOTLAN
 VYLALENNLEIIPVLNKIDLPGAEPDRVIREIEEIVGLDCSNAIRCSAKEGIGISEILDAIVERIPPPQD
 TADKPLRALIFDSYDYPYRGVIVYFRVVDGKIKKGDRIYFMASDKDYFADEIGVLSPSQLQVDELYAGEV
 GYLSASIRSVADARVGDITITHYVRAEDSLPGYEEATPMVFCGLFPVDADQFPELRDALEKLQLNDAALK
 FEPETSSAMGFGFRGFLGLLHMEIVQERLEREYNLSLITTPSVVYRVNVCVNGDTECSNPSSLPEPGK
 RTSIEEPFVKIEMLMPKDYIGPLMELAQDRRGDFKEMKFITENRASITYELPLAEMVGDFDQLKSRSGK
 YASMEYTFVGYKESDLIKLDIQINGDRVEPLATIVHKDKACIGAKVIASEALS AIRKDV LAKCYGGDITR
 KKKLLRKQAEKGKRMKAIGKVDVPQEAFAVVLKLEKEVL

>AAM73733.1 ADP-glucose pyrophosphorylase large subunit, partial
 [Metroxylon sagu]
MGVYVFKRDVLLKLLRWYPKSNDFGSEILPSAVKERNVQAYLFNDYWEDIGTIKSFDDANLALTDQLPK
 FQFYDPRTPFFTSRPLFPPTKIKKCRILDAIISHGCFLERCVQHSIVGVRSLDYGADLKDMMMGADS
 YETEAETASLLAEGKVPVIGVGQNTKISNCI

>XP_007514075.1 DNA repair and recombination protein RAD26 [Bathycoccus
 prasinus]
 MRNEEGDTDALSNLVLGPLSVVARDERHVQRDVLRLDAIVDDDRGGGGGEEETKMETMVLRSERAVEKL
 ERSLEFIDREIEAVEAATKDLGVEEAKKKEEKTPLRDVKLRKQVAAKRVAGLREKRSKVEKELDAALEVS
 EVERRAFEREKKKKKDASDADAKKTGGVPPEEQRRRPKVVVQDDDFDAELDAVQKKT'TNNLLGGGSNG
 ETERERLIRIGAMTPFDRLDGFDKARTDEAGKKLKEKAALLQSAKSKLKTIDLKDPKQMEKMHSAIGE
 AISRRVKPTKNGDSAAKKLALKRKQWKEQSEQQQNKKNNGASARKRRSSFQAYSSDEEELDGDADEDG
 DDVIEAEEDVEFEGGLSVDGDRFAKLLPHQKTAVKWLWELHCQRAGGIIGDEMGLGKTVQVA AFLGALS
 SNLYQASVVVCPATMLRQWRRELKIWAPELKPVVLHDSAITQDALKVANGNRKNAMKNAIRNATRDPKGL
 VITTYECLRGMREDLLTVRWGYAVLDEGHKIRNPEADITVVSRLRRTVHRIIMTGAPVQNRSELWSLID
 FVYPGKLGTLPVFQAQFAVPIQIGGYVNASDQAATTAYRCAVALKDLISPYLLRRLKQDLIDINLPDKTEQ
 VLFPCPMTENQRDAYKGFSSREVEDIIDGRREALGGIDVLRKIVNHPDLLERNRAGDANYGDPVRSGLK
 QVALKILSMWKSQGHRCLVFSQTQQMLDILEQAVANEGYTYRRMDGTPVAHRMGLVDSFNDAGNVGEEG
 VAAEDMQEPVVFVLLTTKVGGLGINLTGANRVLLFPDWNPTDAQARERAWRIGQTKAVTIYRLITGT
 IEEKVYHRQIYKEFLTGKVLKDPKQRRFFKARDMMDLFAYDDPEEKQRGGGVAGSAAMGGGAANETAELF
 AEVEGEIILAADCKDEDEESLITVEGDESLEEGETTTANNGTIVEGVQRVETNRLNVNNKDDNGKGDAAIL
 KSLFDGEGGLHSAMCHDKILSAADSRRRAKIAFADRIARQAAEAVKRSRGRGEMNGHSNTRVQGGQQQQH
 INATTTSTTTIRTATNNTGREGSNNAGSRRLFSRIQQRREEDAAI IATNQANANQDEEARFAQTLL
 KDIIQFLKSR**GGEAPTGLVVDAFADK**VTAERRVIFRNLLKQCARLERNPTTNDGNKGFSAWVLKSEYDT

>Q9FK76.1 RecName: Full=Subtilisin-like protease SBT5.6; AltName: Full=Subtilase subfamily 5 member 6; Short=AtSBT5.6; Flags: Precursor
 MKKLTSLFPLLFLLIPLLASCAEEKQVYIVYFGEHKGDKAFHEIEEHHHSYLQSVKESEEDARASLLYSYK
 HSINGFAAELTPDQASKLEKLAEVVSVFKSHPRKYEAHTRSWEFVGLLEEEETDSVPRRKNDAADDRFRV
 GRNFLKKAHGDGIIVGVLDSDGVWPESKSFNDKGMGPVPSWKGICQTGVAFNSSHCNRKIIGARYYVKG
 YERYYGAFNATANKDFLSPRDPDGHGSHASTAVGRRVLGASALGGFAKGSASGGAPLARLAIYKACWAK
 PNAEKVEGNICLEEDMLAAIDDAIADGVHVISISIGTTEPFPTQDGIAMGALHAVKRNIVVAASAGNSG
 PKPGTLSNLAPWIIITVGASTLDRAVFGGLVLGNGYTIKTDSITAFKMDKFAPLVYASNVVPGIALNETS
 QCLPNSLKPPELVSGKVVLCRLGAGSRIGKMEVVRAGGAGMILGNIAANGNEVPSDSHFVPTAGVTPTVV
 DKILEYIKTDKNPKAFIKPGKTVYKYQAAPSMTEGSSRGNVVDPNILKPDITAPGLYILAAWGSADSPS
 KMSVDQRVAGYNIYSGTSMSCPHVAGAIALLKAIHPKWSSAAIR **SALMTTAWMTNDKK**KPIQDTTGLPAN
 PFALGSGHFRPTKAADPGLVYDASYRAYLLYGCSVNITNIDPTFKCPSKIPPGYNHNYPISIAVNLKKT
 TVKRTVTNVGTGNTSTYLFVSKPPSGISVKAIPNILSFNRIGQKQRFKIVIKPLKNQVMNATEKGQYQF
 GWFSWTDKVVHVRSPIAVSLA

>Q9FNA9.1 RecName: Full=Phospholipid:diacylglycerol acyltransferase 1; Short=AtPDAT
 MPLIHRKKPTEKPTPPSEEVVHDEDSQKKPHESSKSHHKKSNNGGKWSCIDSCCWFIGVCVCTWWFLLF
 LYNAMPASFPQYVTERITGPLPDPGPKLKEGLKAKHPVFIPIGIVTGGLELWEGKQCADGLFRKRLWG
 GTFGEVYKRPLCWVEHMSLDNETGLDPAGIRVAVSGLVAADYFAPGYFVWAVLIANLAHIGYEEKNMYM
 AAYDWRLSFQNTTEVRDQTLSRMKSNIELMVSTNGGK **AVIVPHSMGVLYFLHFMK**WVEAPAPLGGGGPD
 WCAKYIKAVMNIIGGPFVLPKAVAGLFSAEAKDVAVARAIAPGFLDIDIFRLQTLQHVMMRTRTWDSTMS
 MLPKGGDTIWGGLDWSPEKGHGCCGKQKNNETCGEAGENGVSCKSPVNYGRMISFGKEVAEAPSEINN
 IDFRGAVKQGISPNHTCRDVWTEYHDMGIAGIKAI AEYKYVTAGEAIDLLHYVAPKMMARGAAHFSYGIA
 DDLDDTKYQDPKYWSNPLETKLPNAPEMEIYSLYGVGIPTERAYVYKLNQSPDSCIPFQIFTSAAHEEDED
 SCLKAGVYNVDGDETVPVLSAGYMCACAWRGKTRFNPSGIKTYIREYNHSPPANLLEGRGTQSGAHVDIM
 GNFALIEDIMRVAAGNGSDIGHDQVHSGIFEWSERIDLKL

>XP_001416343.1 MC family transporter: uncoupling protein [Ostreococcus lucimarinus CCE9901]
 MASGAVTHPIDLVKVRMQLRGEVDKAAAAASSRASTRAPGMVSTFAHVLRVEGALGLYKGLTASLMRQAS
 FIGTKFGAYDALKAAALRSEGDEKLPFWKMTMCGIGAGAIGAAVGNPADLAMVRMQADGRLPVELRRNYRN
 GADALMRVAREEGVGALWRGCAPTVMRAMIVTASQMAVYDQAKHYIVEHTSLNDGLLAQTGASFGAGVVA
 ALTSNPIDLAKSRLMSMKADEHGKMPYSGTLDCIAKTVRREGVFAVYKGLVPTTARQVPLNMVR **FVSVEW**
MKRLLEPL

>YP_009266863.1 maturase K (chloroplast) [Primula sinensis]
 MEEFKRYLELDRSQQHYFLYPLIFQEIYALAHNHSLTRSSLELEKAGYDNNFSLLTAKRLITHLITQM
 DRQNHSFLTNDNQNPFGLHNTTLYSQIILEGFVVVEIPFSIRVISSLEGKEIVKSHNTNLSIHSIFPF
 LEDKFSHLNYVLDLLIPHSIHLEILVQTLRYWVKDASSLHLLRAFLHXYHNHNWNNLITIKSSFSFSKR
 NKRFLFFLYNFHVYYESIFAFLRNQSSHLQSKFYRSLLDRIYFYEKRDHFVEVFTKYFQAVLCSFKDSF
 MHYVRYQKGKALLASKGTFLINKWNYLVNFWQCYFDMWSQPGRIHINKLSNHSLLGLYLSSVGLNSSM
 VRNQMLENSFLIASASKKFDTTVPIIPLIGSLKAKFCNLLGHPISKPVWADLSDSDIIDFRGRIYRNIF
 HYYSGSSKMSLYRIKIYIRLSCARTLARKHKSTVRAFLKRIGSEFLEEFFMEEELIFSLTFPK **AYSTSG**
GLDRKRIWYLDIFCINDLANHS

>XP_013444647.1 trehalose-6-phosphate phosphatase [*Medicago truncatula*]
 MKSLLPASLNGDNNEESILSSYNWLENHPSALENFEKVMNIAREKKIVVFLDYDGTLSQIVDDDPKAYM
 TDAMRAAVREVASYFPTAIISGRSRNKVYDFVKLNIIYAGSHGMDISTSLGSSKYHDKNHQTKGVDEKG
 NEVVLYHPAEFFLPTIQEIIKILKDNIRVINGSTIEDNTFCFTVHYRRVKNREDVEVLKEIVESIMKDYP
 DFLISGGKEIMEIRPNVNWNGDALMYFLDTLGYNTFDDVLPYIYGDDRTDEDAFKILKQIGGGFPIVVS
 SIAKETNASYSLRDPADVKTFLTHLAKWKKNLIHKTQQR

>XP_007010319.1 Pectinesterase inhibitor-like protein [*Theobroma cacao*]
 MASLSRSSLLVALSFAIFFINPSFAKPRPNVTDAEIITICSKTPAPSFCLKVLSNETLHANQTSLHGLAK
 ISIELALASADETQVEISPLIKQAENYTVREGYTLCSQNYQEAVASLKDAKRLLSKHDYRGVVRVQALAAAL
 EEAEACEHDLRIPAFNPSPLHDKNEEFKHYCNI IWAITNRLVDYY

>KVI07455.1 Aspartate/glutamate/uridylylate kinase [*Cynara cardunculus* var.
scolymus]
 MASSTNVESPPIGEESNNVNNSETPATAAAVNSPATTAAVNSPTTAAVNSPGTAEDVTKTSEITPNGR
 RSKKQKLDSSSTGGDHRREEREWSDTAIAILLDSYTEKFMALNRSNLRGKDWEVVAELVAESGDKQPRKSI
 EQCKNKIDNLKKRYKLETQRMENGINESSWVWFKKMDIVFGSLMGPKSGAGTGATSDDDKSVGASSPHRL
 RRSARLAPTSAPVRSNFKTTTNIKWKRVVFKVSGTALAGSGQSIDPKLCRLQKKWQQRPGMVWMSMMATVM
 NSVLLQSALEKLGITRVQSAFLMPEIAEPYNRLRAMRHLDKGRVVI FVNADAVIKGTNVNGICDKHKVP
 LDQISFRDAVSRDCGSMDLMAIQFCEENSIPVVI FNMLERGNVAKALTGGQVGTLDQ

>XP_005851718.1 expressed protein [*Chlorella variabilis*]
 MAMANGLLANDVLLWKAIGAAVWAAAIAVAASGALGLMLSPATLFS PARLVGGSFSLSAWLSATALVLAQ
 APAFAGAAAAALRACEPRPAHLHRLHWP RCAPASMLIGKLAARLGSVADAASTAAFFATHLSAALFLSV
 FAAAMGNAMGGARSTLQYSLWLATAYLLHWVHCSQDVLAFPVSVQRHRYFRMKQRLPRAAVQAAQLAAAAF
 SCAVTTSLLRSNVLGEASPGASGAPLTLGGGTAALLAGALCTFCWLMSSAALEVVFTERLRPDDYSDRDV
 LRAMVACLAKRGGMLMQGLALHDASLLAGDVGRAALRRADMFADESGDRWKPVAVACIAELDAVSAAVGA
 ALQQQKPAAGGAAMP SAATQSHKWNGLGSSKQQLDALLAVRCGYPRAAQAAQALAGFACASLKEDRFGV
 LQLTQFGLGDVLLCLL GALGATQQLMRVTASLVPRQLSLGPWRGNGDAAAWAGCYSSPSVDAAYALQDT
 LTVALYRTSTTFGAGLGKVLADCTGKPAYGSASEAAAALLQQFQRGQA

>XP_007013072.1 Alpha/beta-Hydrolases superfamily protein [*Theobroma cacao*]
 MGDLOTSEGMNETPLSVYHGVSNNICHPLDFLVFPSLMDDIFTRLLAWTAYMNRFFINFVENLVLQDVYR
 CSEDSVAITSGQNPRSSSSQSICKSAREGCSSSSSTCTENSYS AICGRSSELREELLANTDSRNTITYHS
 KRGAPFIFQGLMFPLFGIRLAWSLALASWRCALYHLRSTQAQVYSIKSRMQRTL RGSSDDIGWLQRNPGM
 APVENGTARFLELLQAIRNGEHTLPNSFVYLLVPGLLSNHGPLYFVATKRFFSKMGLACHI AKIPSEASV
 AQNAWELKRYIEELYWGSGRVMLLGHSGKGVDSAAALS IYWSELEGKVAGLALVQSPYGGTPLASDILR
 EGQIADKETRRVMELLICKMIKGDIRALEDLTYEKRKEFIMKHKLPEGVPLISFHSEARVAPGFLATMTH
 IAHAELPWLPLPKFGSAEFDVSGRLGHQVPIVPIPI SAAMAACALHLLLLRYGEKSDGLVTCRDAEVPGSV
 VRPDQKLDHAWMVYSSWKKNPNEPDACEMCEALLTLLVELGERKQEEMEKFSSS

>XP_003607656.2 long-chain-alcohol oxidase FAO2-like protein [Medicago truncatula]

MRRECHPLLSGMRENKRYKHGFSVAEMESLTSICEVMFPSLPMDDALTKDDESSKDVQSFFNISASKYPI
 PDEVAEMFEKRAVIEAVILIRVILWLLATRLGTLICGLLCLSKKWPFINNFSSLSLDKREKIVQRGLKH
 KFLTPFRLAFVYIKILCLFVFFSWVDENGDNPAAWKAIGYEVSTADEKMINDTKKRPLEKGIIEIMHEHDT
 TLQQSLSNKGLNVTLDKNNILKIKCDAVVVSGCGGGVAASVLSKAGYKVVVLEKGNFYVPKDYSSLEG
 PSMDQQYENGGILASIDSRILILAGSTVGGGSAVNWSACIKTPEKVLKEWSEKHKLSLFDSEIYLSAMET
 VCERIGVNECTQEGFQNVLRKGCQNLGLKVDYVPRNSPGNHYCGSCGYGCPKGEKQGTQATWLVDVAVD
 KGAVIITGCKAERFLFENNYRNANTRKNKKCLGLVAKTLNSGVTMQLQFEAKVTISAGGALLTPPLMISS
 GLTKNKNIGRNLHLHPVLMTWGYFPESKSDFKGKANEGGIITSVHRVPSSSNDDSIDSTRAIETPLLGP
 SSFSALYPWESGLDYKQRMNLNYPRTAHFITIIIRDKAGQVTTERRISYKLNIDKENMRAGIQQALRILI
 AAGAVEVGTTHRSQQRIKCNENTSEKEIEEFIDSVYPMEGALWPGENWNLYTSAHQMGSCRMGVNEKEGA
 VDENGESWEAEGLFVCDASVLPVAVGVNPMITIQSTAFICISNRIVDFLRKRQEP

>KVH88216.1 Protein kinase, ATP binding site-containing protein [Cynara cardunculus var. scolymus]

MPQLRSGVRRGRPSKRPIAAERTEPEGVVEAVRKTDKRGRKVNNTKQVNRVEQKNGRGRKKAVQEPVVV
 SDEDSEEKNAVRTTPEEEEEKERKPELAASAAEIKDKEKEEEEEAGEKKMDDNDSVAPSGDKGLGADEGST
 APLPERVQVGGSPSYKIEKKLKGKGGFGQVYVGRINAPVHERTGSGAVEVALKFEHRSSKGCNYGPPYE
 WQVYNVLGGSHGVPRVHYKGRQGDYYIMVMDMLGPSLWDVWNNNSHTMSIEMVACIAIEAISILEKMHSR
 GYVHGDVVKPENFLLGSPGTSDEKKLFLVDLGLATRWRSASGLHVEYDQRPDVFRGTVRYASVHAHLGRT
 GSRRDDLESLAYTLVFLLRGRLPWQGFQGENKGLVCKKKMATSPETLCCFCPAPFRHFVEHVNLKFDDE
 EPNYAKYISLFDGIVGNPDIRPINTDGAQKLMYQVGHKRGRLMMEEDYEQPKKKVRMGMPATQWISVY
 NARRPMKQRYHYNVADARLPQHIEKGNEDGLFISCVASCSNLWALIMDAGTGFTSQVYQLSPMFLHKEWI
 MEQWEKNYYSIAIAGANNSSLVVMSKGTQYIQQSYKVSSEFPFKWINKKWREGFHVMTAMATAGSRWAIIV
 MSRGIHRRWDAGYRITSTAATWDQAAFVLRKVGKESVHCFCVLWSNCFMSSRSSSSFSFLDC

>KHN46581.1 Glutathione S-transferase F13 [Glycine soja]

MAFKLYGLPMSTNTTRAMICLHEKEVDFELVPVNVFAAEHKQPPFLSKNPFGLIPLLEDGDLTLFESRAI
 TAYVAEKFKETGADLIRHKDAKEAALVKVWTEVESHYEPAVSPPIIYEYFVAPFQKPEPKSVIDTNVEK
 LKKVLDVYEAKLSSTKYLADDFYSLADLSNVSETHYLMQTPCASTVNERPHVKAWWEDISSRPAFTKVVG
 GMTFGQNQEE

>CAC04434.2 psbA protein, partial (chloroplast) [Artemisia rupestris]

NFPLDLAAIEAPSTNV

>KYP64423.1 KDEL motif-containing protein 1 [Cajanus cajan]

MTLTCPKDYYPTRFEQYQDSSTESTCPEYFKWIHEDLKPWKRTGITREMVERGQNVSHFRLVLIQKAYA
 EKYAYSQYTRDVFTIWGILQLLRLYPGNIPDLELLFETGDRTVVEKQNFQESPPPIIFHYCGQKNAYDIV
 FPDWSFWGWAELAIRPWEALLQSIEEGNKKIKWKDRLPYAFWKGNTMVSYKRYDLTKCNASDQHHSYAH
 YPLSWDKIEIRGFKNKLEHQCVHRYKIYVEGVAVSVSEKYIILACDSMTLFIPIFYDFTRSLVPRKHY
 WPISNTNQSMCNDIKYAVDWGNANPDKAEAIKAGTSFIEENLKMKFVYDYMFLHLLNEYARLLRFEPTIP
 AGAVEICSENFACPLNGIWREYMVESMVKSPSDTPPCTMASPYGENEGEEK

>CDX86356.1 BnaA06g30730D [Brassica napus]

MEGNSNWKPNQQGGDSLASNANDWRSQLAPDMRKKVILAIVEKLLIYYPTRHPNAIKNTAFSFEKGIYA
AAKDKDDYMRTIKGNIMNFRKRLQSSNVQSGSSVNGTNPAPAAQALNQGQSTPTSQQWLHQNNNNIQSN
LNILDSFKSSSRFSKTDTPWKGTSSSSAATEFQLLPESNGSTALERGSTTFIHAAAASEQKRQEREQLIS
HLMNDKDTQQNHLTPQQNNGEKQAAFRASSSQNNIASFQERPLQNNSIQORLYSHKVGQSQTMIQQQYQ
PQHTMQQAQRNLIQQPLDDTQRFQASGSLLTQQNQPYQLQRTSPANTFSIDYTNTFFFLASQDSTGQTV
NASGGGDWQEETYQKIKALKEKYILVVGALYQKLSNKLREIDAHPPQKIQHGHMEKLRASKATLKLVLVLF
LNVSRNAITESHREKFNIYEEQLLRFVKHNQTVTRRPMQQQQQQQQVHLPPSQTHQTALQSQSGHQVFH
VPQSSALSNTTSHAMPSSQTRPKMEPKEETNIMTLPASNPQPSMFQQKQFHLLSMQQRQQQPQKNH
QQLOMPKNEMNDVRMSQRVNNKAGLRQQNISPNQRHLAKPLASPQLVDQQILPTTFNKNGTSSQSGGSPF
VAPSSNLGDPENPISVESPSHDYQLQPAAQEHPPEPNAERPIDRLIKAFQSSSPESLAQSNEMSSVIS
LTDRLAGCVQSIGGSRARVPQDLSETRRLRQGETNPNTNKRFRSITTPIDITSETERYKQFSSLESE
VDSTASSGSKANKIEAGLALLQEIMEVNRRLVETMVSICSEVDGPSEVTTGTIVMCSYAPVALCDTFQAL
YKSGHVSQIQPLRLLVPENYPHSPILNIENIPFDSSVNKHEDLSARTRSFRGLSMKEFSEPMSLTEIAQAW
DACARATMAEYAERHGGGTFSSKHGHWEPLRAS

>AFW89768.1 acyltransferase [Zea mays]

MGMAMVLTGDGEV GALIKVSAAVVWAMS YARLAAARLRPGALRLLALLPVVALLCAIPFAFSTSTFRGTSG
FFLAWLGSFKLLLLGAGIGPLDPSLRSLSHFVCSATLPVKLRRQSKEKSQAPARGPARILLSGAVIPGVIY
AYQFKSSMGRYQLLALYSVHIYFSLDLLLATVHTVIHDLGMEPEPQVDRPYLASSLRDFWGRWNLMVP
SILRPSVFRPVRARLNAAGVHATFLVSGLMHELMFYIIMRSAPSGEVTAFFLLHGACAAAEGWWASHAG
WWRPPRPAAVPLTLAFVAGTGFWLFPLPAMVKGLDEMVLRECQGMVVLMEQAARRLAGATDLVSSMT

>AAA81879.1 geranylgeranyl pyrophosphate synthase-related protein
[Arabidopsis thaliana]

MLFSGSAIPLSSFCSLPEKPHTLPMKLSPAIRSSSSSAPGSLNFDLRTYWTTLITEINQKLDEAIPVKH
PAGIYEAMRYSVLPQGPKRAPPVMCAACELEFGDRLAAPTACALEMVHAASLIHDDLPCMDDDPVRRG
KPSNHTVYVYSGMAILAGDALFPLAFQHIVSHTPPDLVPRATILRLITEIARTVVGSTGMAAGQYVDLEGGP
FPLSFVQEKKFGAMGECSAVCGLLGGATEDELQSLRRYGRAVGMLYQVDDITEDKKKSYDGGAEKGM
EMAEELKEKAKKELQVFDNKYGGGDTLVPLYTFVDYAAHRHFLLPL

>KVH97338.1 Alpha-N-acetylglucosaminidase [Cynara cardunculus var.
scolymus]

MFNSKLSFLILLLLLPLSSSSSSPEAIQGLVERLDTQRPSFQLQESAALRRLLPTHLSSEFKILSK
DACGGYSCFWITNYENSSNKSAEITVQGTAVEIASGLHWYLYWCGAHVSWDKTGGIQUIASIPPTGNLP
LVKKDGIMIQRPVQWNYQNVVTSNSYVWWDWERWEKEIDWMALQGINLPLAFTGQETIWQKVFMDFNIS
AQDLNFFGGPAFLAWARMGNLHAWGGPLSQNWLDQQLALQKQILSRMIELGMPVLPFSFGNVPAALRE
IFPSANITRLGDWNTVDGNPRWCCTYLLDPSDPLFIQIGEAFIKRQIKEYGDVTDIYNCDTFNENSPPTS
DPAYISSLGSVYKAMSKADKDAVWLMQGWLFYSDDSSFWKPPQMKALLHSVPFGKMIVLDLFDVVKPWE
SSSQFYGTPTYVWMLHNFNGNIEMYGVLDALASGPIDARISENSTMVGVGMCMEGIEHNPVVEYELPEMA
FRKDKVRVEEWLRVYSRRRYGKSVKQADAWEILHRSIYNCSDDGIADHNTDYIVKFPDWDPSLNTYSTFS
KQNRTPSFITTHRNRFILTETQSTLPQPHLWYSTRDAINALKLFIDAGQDLAGSLTYRTQITMWYDITK
TNQSQLHDYANKFWGSLLDVYLLPRASMYFNRLSESLRENTFRKIVEWRTEWISYSNKWQQDTKLYLVKA
QGDALAI STSLFQKYFG

>XP_011401811.1 Serine palmitoyltransferase 2 [Auxenochlorella protothecoides]

MGTGTLAPSLKQALFAYTSFTLLYIFGQFRDFFRKRWSKASLKGYPAPIRQDYEDFYTRRMYRIHDCFNRPITGPPSSWMDVLERTPVNRQEPLTTTGKIKRCLNLGSYNLGFASSDPYCTPRVLDTLSQHGWSLCSNRADAGTTPVHEELERELASFLGKEAALTCGMGFATNSAFLPVLATPGTLVLSDALNHSSIVAGVRAAKGTVRVFDHNDPEHLERLLRTALVEGQPRIHRPWRKVIIVVEGIYSMEGEVVDLAGFVAVAKKYKAYIWLDEAHSIGAMGATGRGVCEHLGVSSADIDVLMGTFSKSFACGGYVAANKDVIDYFKAQCPAHLIYATAMTPPAVQMVLSALRLITGADGSDRGAHKLSQLRTNANYFRSGLLRGLGFNVLDGWDSPVMPVMIFQPAKLSATSREMLRLGVAVVEVGFATPLLTSMRVCISASHTRQDLDYALQLFDYVAERCDIRYGGDTARALANTGPKIQAV

>BAB10815.1 disease resistance protein-like [Arabidopsis thaliana]

MASSSSSSSRNWSYHVFPSPFSGEDVRNTFLSHFLKELDRKLIISFKDNEIERSQSLDPELKHGIRNSRIAVVVFSTKYASSSWCLNELLEIVKCKKEFGQLVIPIFYNLDPShVRKQTGDFGKIFEKTCRNKTVDKIRWKEALTDVANILGYHIVTWDNEASMIEEIANLILGKMNISPSNDFEDLVGIEDHITKMSSLLHLESEEVRMVGIWGPSGIGKTTIARALFSRLSCQFQSSVFDKVFISKSMEVYSGANLVDYMKLHLQRAFLAEIFDKKDIKIHVGAMEKMKHRKALIVIDDLDQDVLADALADQTQWFGSGSRIIVVTENKHFRLANRIDHIYKVCPLSNALALEMFCRSAFKKNSPDDFLELSSEVALRAGNPLGLNVLGSNLRGINKGYWIDMLPRLQGLDGIKIKTLRVSVDGLNLRKDEAIFRHIACIFNGEKVSDIKLLANSNLDVNIKLNLDVDRSLICERFNTLEMHSLLOELGKEIVRTQSNQPGEREFLVDLKDICDVLEHNTGTTKVLGITLIDETDELHIHESFKGMHNLFLKIYTKKLDQKKKVRWHLPERFDYLP SRLRLRFRDRYPSKCLPSNFHPENLVKLMQSQSKLEKLWDGVHSLAGLRNMDLRGSRNLKEIPDLSMATNLETCLKSSCSLVELPSSIQYLNKLNLDMSYCDHLETIPSGVNLKSLDRNLNSGCSR LKSFLDIPTNISWLDIGQTADIPSNLRLQNLDELILCERVQLRTPMLTMLSPTLTLRLTFSNNPSPFVEVPSSIQNLYQLEHLEIMNCRNLVTLPTGINLDSLISLDSLHCSQLKTFPDI STNISDLNLSYTAIEEVPLSIEKLSLLCYLDMNGCSNLLCVSPNISKLKHLERADFSDCVELTEASWNGSSSEMVKLLPADNFSTVKLNFINCFKLDLTALIQNTFFMQLIITGEEVPSYFTHRTSGDSISLPHISVCQSFFSFRGCTVIDVDVSFSTISVSFDIEVCCRFDIDRFGNHFDSTDFPGYFITTKLGGHLVVFDCYFPFNEEFTTFLDGOFNVDHVDIQFRLTNDNSQLKLGCGILLSEDPVSLDNRPCSPNIPGVCEDSALERRSFRTKMRMITEVASTVVRSDDEEARNGDIRESREYGKALFLSAKDFVSGNDTEE

>OAY71028.1 26S proteasome non-ATPase regulatory subunit 9 [Ananas comosus]

MVGANVKAETLSLMERRAAMEADMNAIDALSRPGPGISGNLDSQGFPRPDIDIPTVRAQRHRLAELRNDHKDITDKIDK NLQVLHSLRLGK NEPAIPGDAGASASHGVSSQYSPMEEDPIVRIPIFAIIDEIVDDSPA AEDGLQLGDEIVKFGSVEIGDGLQAKLVSEAQSNQGRTVPLIIIRHGSVMNLNVI PRQWRGRLLGCHFRIL

>KZV29384.1 methyl-CpG-binding protein 2-like [Dorcoceras hygrometricum]

MDLTVSGQEGKDTLGRDYRKGNWTIQETMVLIEAKKMDDERMKRSCDQRAKQAE LRWKWVEDYCWRNGCFRSQNCNDKWDNLMRDFKKVREYERRVVDAGVRGEGGADKSYWKIEKNERRENGLP SHMLIQIYEALVDVVEKKDQTMVASALAAAPASGGAASSNVDPNTGSGKSSLMALGQPSILPLPVQQQAPIMLLPLPPPAEPIPIFPFSQPLPTVDS EDTSECSSEPAKRRKKEGGGGG EAGTSGGPLGEIGNAISKSASIMAKAIQSREEADGRRHREV VGLHERRLQIEESKAEIKR QGMSNLVD AINK LADSIQALAANNKNQESFK

>AHG59274.1 ribosomal protein S7 (mitochondrion) [Dicranum scoparium]

MNLFVKLSNFSFVLGSSLDWFHQSKLSEKAGMKKNENWPFSGKNLFFFSESFARRLLHCRYLCYALPGHAPS RPK DRGANTYNSSDNLGYIR GLHSKQKQLIKKLVHICMINGKKTRSRAIVYKTFHCLAQHGDILRLLVNAIENVKPVCEVKKVRI SGTTLVPSIIATNRQETLAIRWML EAAAKRRMNKKSMSLDQCLADEILDAS

RKMGIARKKRDDLHKLAQANRSFSHYRWW

>XP_010099608.1 BTB/POZ domain-containing protein NPY2 [Morus notabilis]
 MKFMKLGSKPDSFQTDGKNVRYVASELASDITVIVGDVKFYLHKFPLLSKSAHLQKLVSTATDEHSDGVY
 ISDIPGGPEAFEICVKFCYGMTVTLNAHNVVVVRCAA EYLG MHETIEKGNLIYKIDVFLSSSIFRSWKDS
 IIVLETTNSLTSMPELKLVGRCVESIATKACVDVSKVDWSYSYRKKLPEENGNNDPNWNGVVKRSVPK
 DWVVEDLCELEIDLYKIVLATIKSKAIVSNEVIGEALKAYAYRRLPGFSKGMICQGDMMVKHQTTVDIVW
 LLPAEKGSVSCSFLLRLLKAANLVGSCDVVKEELVKRIGQQLEEASVNDILIR **AAEEETMMYDVNVVQRI**
 VEVFLRQDLNTEIESLEDDDDQLQKMRPGILSDASKLMVAKLIDGYLAEIAKDPNPLPSKFVDLAEMVSGI
 SRPAHDGIYRAIDMYLKEHPGISKSERKRICKLMDCKLSVDACMHAVQNERLPLRVVQVLFQVRAA
 ATSGSSTPDLPKSIKDLNASHGSSRSATNTNEEDWDAVATAEELRALRGELASLRLSNGVGGSERNGGD
 GAKSGVDKAAISKMRGLLKS KKI FTKLWSSKGGQGENSGSDSSESLGSLNPEEAKSTPSRNRHSVS

>KZV45624.1 F-box protein [Dorcoceras hygrometricum]
 MPFEEILKVVFPLLEGKDLVSCMLVCRQWREAAQDDFFWKCLCARRWPSICKKSSPPTLTYHKLKFNFYR
 RRPPIRKILPPRLSFSLEFYIDIWTGERLILSEVIPGVPVLHKGMGTLPPGICDMLRFHLEGPEYKMTFSV
 QPRFNIPFDQTVSVSVLVGRRDHKKVACVINKSIFDYIDRTAHRALAFDYLD FSPAHPFVSGIRAWFSL
 FMDHGDEGIVDVFGEI LDFCDVADSE DQVLWLLDMLDWK

>AHZ60327.1 PsbA, partial (plastid) [Eugenia bacopari]
FVMXERNAHNFPLDLAAVEAPSTNG

>ADK47981.1 L-ascorbate oxidase-like protein [Cynodon dactylon]
MTRGAAAAALLALALVAVARAEDPYHFFEWKVTYGTKNIMGTPQKVILINDMFPGPPTINCTSNNNIVINV
 FNMLDQPLLETFWHGIQQRKNSWQDGMPTMCPKPGTNTFYHWQPKDQIGSFFYFPSIAMQRSAGGYGLI
 SVHSRDLIPVLFADAPDDFPVLVGDWYTKDHTVLAKHLDAGKGIGRPAGLIINGKNDKDAASAPMYNFEEA
 GKTYRFRVCNVGIKASLNVRVPGHNLKLVEMEGSHTVQNMYSLDVHVGQCLSFLVTADQKPADYFLVVS
 TRFIKEVSTITALIRYKGSSTPPSPKLPEGPSGWAWSINQWRSFRWNLTASAARNPQGSYHYGQINITR
 TIKLSTSRGKVDGKERYALNGVSHVDAETPLKLA EYFNATDGVFQYNLISDVPPKAGTPIKLAPNVLSAE
 FRTFIEVVFNPEKSIDSFHIDGYAFFTAGMGPWTWSPQSRKTYNLLD TVSRHTIQVYPRSWTAVMLTFD
 NAGMWNVRSNLWERQYLGEQMYISVISPARSLRDEYNMPETSLRCGKVVG LPPMPPSYLPA

>BAF98603.1 CM0545.270.nc [Lotus japonicus]
 MEHTEAGTVQAVL FQEEAVETIKGTQKMEPSEAGTVQMSDNVLRKLLRGP RYDPPADCGWETCYNCGE
 EGHATVKCAA AKELKKPCYLCGSLMHQAKRCKKEIQCYVCKSFGHLCCANTTGSTPIEISCYKCGQTGHT
 GLARCSHVQGNFFTENAGVMVKEEVRKRGH TSNTE SPTFQKENG YMGDR **SAPHDMGMPYMEKKPLTEE**
RAITTQQPSKHRGGCRE DQAMLFITGSQHLGIGESEK LQQVVA FQNSGL

>GAQ80887.1 mitochondrial ribosomal protein L13 precursor [Klebsormidium
 flaccidum]
MSNHLKNVATVGLRWRLVDAKGEVLGRLASQVSMILMGKDKPTYTPHWDEGDVVVVKNARHVELTGKKVK
 EKVKYKWTGYVGLKERTVEEQFEREPT E VLRKAVERMLPKNRLRDDRMRKLRIFPDEEHTYEGLELWKF
 EMPPRNKREL RPPEARLLRKQQAEGEQKLHNGASVEPS

>GAQ79155.1 Peptidase M15B and M15C [Klebsormidium flaccidum]
 MATTTASPASRSLQRLFCKSSPSQAPSKLLLQTTPLKSNLFLDLLSPRSLPRGGLKWLHSSKLLRNRSR
 QFSTQDICFSGYQRAFFQQLRCTRPTVRTAAFSADRSNPLPDAGAYRK**STAQRPSEEDSLKR**EIDGERTGA
 EWRRDSVERSAGPVLDPDTREFLSFERNVKAIDPESIQQAVRIRPEDRLRKQKQKAAQPPWVAAAQWVLAG
 LALCAFSWRVAKLGGGDIHEPPIATVELDSRPQKGLWAEQLKKAQAEDGFREQLLGHFRTEEAPKDDLK
 LLRDGTVKLRSAAYSSFFAMQQAAWKEGVQLLPISGFRSIEDQKDVFFGIKAERNQSAKERAKVSAPPGY
 SEHHTGYALDIGDAKAANTDLEFTFDQTSAFEWLQRNGAKFHFEMSFPDRNPYGVAYEPWHWRYVGNVHS
 LQMFHGHERVNSFPASDATART

>XP_003619298.1 eukaryotic translation initiation factor 3c [Medicago truncatula]
 MDNNDWGRLQESFDKINKQLEKIRRVSEKIPKLYIRTLVVPKDFMAVSRDKDVKKMIPKLKNNNKQYED
 LINK**SLMILSMLIHR**SHGKLFNKKFKWLVAVRGRKKTERFEQVDHILPTNQNWLKTPAQELQILFSVVS
 QFDVNSSLIGGHMPINVWKKCVHNMLVILDILVQYPNINVDSEEPYESETKKGADYNFFKSLQCIDPHT
 CEYIERLQDEPFMGDFEGSLKVALMRVELIYYKPQEVYDAMTTLVEPEPLNIFLLESVYLISVMLLEVFN
 IAANVHDVKRKIISKNFSSRLLEISDKKHSTVLPKMLRIMSWLSQCFLSMETSTRLDKTKKEALIRTYLIT
 FSSSYESECSSVMNYEHARWDQPSGCIVFRNVEPSMVQALAFELTEKLSILAKSSERATEAWLGSVGI
 ALPLLQMVGVKTCRGSLLLILINILFLSLFFSNLLCTPFTICLF

>BAD61475.1 cis,cis-muconate cycloisomerase -like [Oryza sativa Japonica Group]
 MDSAISTSSSFSPLRLSQNHGSIMSGSTRAQVLPVGRQQQPPAGGATRLRAVSPSPSPTPPAPQPAETF
 GFDALKEAFSVDVVAEAEARPLNVPLAAPTIASSRLDAVSNVAVRVELRSGAVGWGEAPVLPVTAEDQP
 GALAAATRACGALAGAPAAPLGAVLQDVASALPGHDFASAR**AGVEMALIDAIANSIR**IPLWRFLGGASDS
 VTTDITIPVTPNEAAQLAAKYRGQGFQTLKLVGKNLNSDIEVLKAIIRLAHPDCSFILDANEGYTANQA
 IEALDRLNEMGVTPVLFEPVHRDDWEGLRDVSIVAKEKYRVAAVADESCRSLDDAQKIMDGNLAHVINI
 KLAKLGLGALVEIDAARKARIALMIGGMVETRIAMGFAGHLAAGLGCFSFVDLDTPLLLSEDPVFGGYE
 VSGPVYKFTNARGHGGFLHLDNNGLK

>XP_002313657.2 TRNA ISOPENTENYLTRANSFERASE family protein [Populus trichocarpa]
 MGPTGSGKSKLAIDLAAHFVEIINADSMQVYRGLDVLTKVPISDQEGVPHHLLGTLNPNVEFTAKDFR
 DSAIPLINEILSRNCLPVVGGTNYIYIQAIVSPFLDDTTNDLDESLLNHPSGDEQTDHATDSGRESFNH
 SYDYLRELDPVAANRLHPNNHRKINQYLNLYARSGILPSKLYQGKAAENWGCMDNYRFHCCFCVDADIP
 VLDRYVEQRVDSMIDAGLLGEVCEVYNADYTRGLRQAIGVREFDNFLRVYMSDEKGDHSMGSLFLQSK
 NEDVKLLKDNMREILHSSDDNQLKILLAEADKVKANTRRLRVQKRRLTRLQTFFGWNIHYVDATEFIS
 CKTDELWAGQVSSAVNVIR**AFLTEERSAVPDLETHVGGGMKSVER**NLWTQYICKACGNRVLRGAHEWEQ
 HKQGRGHRKRISRLRKSQGHSSYSLVEQEVISNSS

>KMZ57095.1 glucose-6-phosphate dehydrogenase [Zostera marina]
 MAFSASRCLFSISPFGFSTSLSSSPSSSSSSSLFPCNLYSGSVQTRGISLDQRRRCQWSTFVCKQMEAVK
 PSSGAKEVVLDKESSLMGRSVNGALPTTSSLSKDNENGLLIGEDK**EITVTITVVGASGDLAK**KKIFPA
 LFALYYEDCFPKHFTIFGYARSNMTDAELRIMVGKTLTCRIDKSENCQKMKEFLSRCFYHSGQYDSEDN
 FSKLDEKLKEHEAGRISNRLFYLSIPPNIFFIDVVKCASTSASSTSGWTRVIVEKPFGRDLESSALTKEG
 KQYLDDEDQIFRIDHYLGKELVENLSVLRFSNLVFEPLWSRQYIRNVQLIFSEDFGTEGRGGYFDSYGIIR
 DIMQNHLQLILALFAMETPVSLDAEDIRNEKVKVLRSMRPLEIENVVVGQYKGHVKGGSYPAYTDDNTV
 PKGSLTPTFAAAALFIDNARWDGVPFLMKAGKALHRSKRAEIRVQFRHVPGNLYKNTFSVDIDRATNELVI

RVQPDEAIYLKINNKVPGLGMRLDRSKLNLHYAAKYSKEIPDAYERLLLDAIEGERRLFIRSDELDAAWK
LFTPLLKELEDRKIHPELYPYGSRGPVGAHYLAAKYNVRWGDINAENYKT

>AAF74625.1 integrase, partial [Oryza rhizomatis]

KGCRDCQQFRAIHAPASVMHPIIKPWPFRGWGIDMIGQINPPSSKGHKYILVATDYFTKWVEAVPFKKV
DSKDAIQFVKEHIIYRFGLPQTITTDQGSIFASDEFVRFADSPASF

>XP_007031938.1 Ribonuclease H protein [Theobroma cacao]

MHHLPRVQSDHRPLLVVLDDHQNGQSPSVCHFQSAWLTHEDFGNFVQRRLLARLGGVEKALEHRVARRK
NKLKIIRLKDEQGNWCDDQSTLKLQAVAFFQKLYTKDNGTLSSYSIRGTFLLTSDKDKLRLTQLVESKEV
YDALFEMKPLKTPGLDGLPTLFFQSQWAIVGQSLVKNVSNIMEGGDFGDNICSSSLIVLIHKVFNPETISQ
FRPIILLPVAFKRDKWLSDKTLADITCRVANPALDKVVVREFLNPNGHWDYDKLSYCLPNEVVLQVVQTM
PPTVIIAQDMPYWGESASGQFTVASVYDYLRLQLSSPAKARPSGIWQGAWKWQGSQVRVTFRLFQCLHGRL
TNRERLHRQLTTDSLCPQCRMEDETVTHVLRDCMVATSLWVKIIPQHEQNDFFTFPLREWLVSNLQKQQL
ILGNPWSVVVFLACWCLWKWRNGVVFYAAFNPTRKRISMIKSMATATIATSADFDGVQVERRKKEEVLIG
WRTPQVGWVCLNTDEAYKRSIEEASTGGVIRNAEGDWQAEFLAKLGKCSAYRAELGWVLHGLRLAWDSGF
KKVQVQVDNKMVPAVSTNKLI PGANTDLIRA IKDVLQKEWEVSMHTYCEGNMVTDYLASAFVLEKSY
IVLEQAPT GARKLLMYDMLGVCLPRMIPIQ

>XP_007031938.1 Ribonuclease H protein [Theobroma cacao]

MHHLPRVQSDHRPLLVVLDDHQNGQSPSVCHFQSAWLTHEDFGNFVQRRLLARLGGVEKALEHRVARRK
NKLKIIRLKDEQGNWCDDQSTLKLQAVAFFQKLYTKDNGTLSSYSIRGTFLLTSDKDKLRLTQLVESKEV
YDALFEMKPLKTPGLDGLPTLFFQSQWAIVGQSLVKNVSNIMEGGDFGDNICSSSLIVLIHKVFNPETISQ
FRPIILLPVAFKRDKWLSDKTLADITCRVANPALDKVVVREFLNPNGHWDYDKLSYCLPNEVVLQVVQTM
PPTVIIAQDMPYWGESASGQFTVASVYDYLRLQLSSPAKARPSGIWQGAWKWQGSQVRVTFRLFQCLHGRL
TNRERLHRQLTTDSLCPQCRMEDETVTHVLRDCMVATSLWVKIIPQHEQNDFFTFPLREWLVSNLQKQQL
ILGNPWSVVVFLACWCLWKWRNGVVFYAAFNPTRKRISMIKSMATATIATSADFDGVQVERRKKEEVLIG
WRTPQVGWVCLNTDEAYKRSIEEASTGGVIRNAEGDWQAEFLAKLGKCSAYRAELGWVLHGLRLAWDSGF
KKVQVQVDNKMVPAVSTNKLI PGANTDLIRA IKDVLQKEWEVSMHTYCEGNMVTDYLASAFVLEKSY
IVLEQAPT GARKLLMYDMLGVCLPRMIPIQ

>XP_007022697.1 Chaperonin 20-like protein [Theobroma cacao]

MPFRFPCLVVKQGGGIVSAFPRKASSLSTDRLLFLCQPPFLRVHLLSQPTVLPVPLSSSPLSAGRLPLSQRW
VFPFLSTAFTRGSAPLSQPLSPVFPVSLSAVHSRVAVKLVASSFEVVRSMVVAEAEKTAGRLLLTEAS
KEKPSIGSVIAVGPGLTDEEGNKKPLSVAPRHTILYSKYAGNDFKGS DGTNYIALRASDVMVAVLS

>XP_002886459.1 F23N19.7 [Arabidopsis lyrata subsp. lyrata]

MAKSCYFRPALLLLLLLVRAESRQGFEPKILLPTEKTKPTADQDEDGIGTRWAVLVAGSSGYGNRHOA
DVCHAYQILRKGGGLKEENIVVLMYDDIANHPLNPRPGTIINHDPGDDVYAGVPKVLHNNYSDSDSRDICY
GKLNLMCGPLIGIAPRFFIATSYPFLIVCSDYTGNSNVTAANFYAVLLGDQKAVKGGSGKVIASKPNDDHI
FVYYADHGGPGVGLMPNTPHIYATDFIETLKKKHASGTYKEMVIYVEACESGSI FEGIMPKDLNIYVTTA
SNAQESSYGTYCPGMNPSPPSEYITCLGDLYSVAWMEDETHNLKKETIKQQYQTVKMRTSNYNTYSGGS
HVMEYGNNSIKSEKLYFYQGFDPATVNLPLKLPVNSQVGVVNQRDADLLFLWHMYRASEDGSRRKDDTL
KELTETTRHRKHLASVELIGTILFGPAMNVLNSVREPGLPLVDDWECLKSMVRAFETHCGSLTQYGMKH
MRAFANVCNNGVSKELMEEASAAACGGYNEARYTLHPSVLGYSA

>AGJ83756.1 heat stress transcription factor A-4a-like protein [Caragana korshinskii]

MDEAQGSSSSSLPPFLAKTYEMVDDPSSDPIVSWSVTNKSFIVWNPPEFARVLLPRFFKHNNFSSFIRQLN
 TYGFRKVDPEQWEFANDDFIRGKPHLMKNIHRKRPVHSHSLQNLQSQVPLSESERQSLNDEVEKQDKE
 RLLMELKRYQQEQWQTYEIQIHCSKDRLEKLEQKQKQMVSSVSQVLQKPVIALNLLPLTETMDRKRRLPRS
 GYFNNEANTEDAAETSQMFPRENAEGTSVLTSSMERLDQLESSMVFWETVAHEVGDKYVHIHSNMDLDES
 TSCADSLSISCVQLDVEVRPKSPGIDMNSEPAAAVVEPIASKEQPVGITTAATGVNDVFWEQFLTEDPS
 ALAEQEVQSERKDYNGKKNENPNDLGRFVWNRNANNLPEQMGHVSQAekt

>XP_010096695.1 3-ketoacyl-CoA synthase 21 [Morus notabilis]

MELLMAISLFLVISYAFHLTKSLLQKQHQC CYMLAYECYKPLEETKLSTDACAKIVLRNKNISVDELRF
 LLNAMVSSGIGDESYGPKNVLEGREESPTLTD AISEMDEVI FTTLDSLFAKTGVSPAQIDILVVNSLFS
 PAPSLSRIVNRYKMRENVM SYNLSGMGCSASII GIDLQHLFKTHENSYAIVVSTESFGLHWYCGKEKS
 MMLSNCLEFRSGGCSMLFTNKNELKKKAILK LKCLVRTHLGADDEAYECCIQLEDEKGEHGFRLTKKLTKA
 AAKALRLNLRVLPKILPVREILRYWIVNLIRNTN NKGDPQLTSSGANNLDFKTGAEHFCIHPGGRAVID
 GVGMSLGLSEYDLEPSRMALHRFGNTSAGGF WYVLGYMEAKRRLKKGDRILMMSFGAGFKC NNCLWEVMR
 DLDDDEDGASSVRQHFGGRVLRVRFVGHGKKEAQERR

>AAQ09385.1 photosystem II cp47 protein, partial (chloroplast) [Euptelea polyandra]

VVLNDPGRLLSVHIMHTALVXGWAGSMALYELAVFDPSPVLXPMWRQGMFVI PFMTRLGITNSWGGWSI
 TGGTITNPGIWSYEGVAGAHIVFSGLCFLAAIWHWVYWDLEIFCDERTGKPSLDLPKIFGIHLFLSGVAC
 FGFAGFHVTVGLYGPGIWVSDPYGLTGKVQSVNPAWGVGFDPFVPGGIASHHIAAGTLGILAGLHLSVR
 PPQRLYKGLRMGNIE TVLSSSIAAVFFAAFFVAGTMWYGSATTPIELFGPTRYQWDQGYFQQE IYRRVGA
 GLAENLSLSEAWSKIPEKLAFYDYIGNNPAKGG LFRAGSMDNGDGIAGVGLRHPIFXDNERREL FVRRMP
 TFFETFPPVVLVDGDGIVRADVPFRRAESKYSVEQVGVTVFEFYGGELNGVSYSDPATVKKYARRAQLGEIF
 ELDRATLKS DGVFRSSPRGWFTFGHASFALLFFF GHIWHGARTLFRDVFVXGIDPDLDAQVEFGAFQKLG D
 PTTRRQVV

>NP_564042.1 Aminotransferase-like, plant mobile domain family protein [Arabidopsis thaliana]

MPVLYEQDKHVSSA IITGQERGVLRQCQERTSLLHHWKLTKEQIALVEKAGFGWFRLVGSISLNNSLISAL
 VERWRRETNTFFHFCGEMTITLDEVSLILGLAVD GKPVVGVKEKDEDPSQVCLRLLGKLPKGELSGNRVT
 AKWLKESFAECPKGATMKEIEYHTRAYLIYIVG STIFATTDPSKISVDYLILFEDFEKAGEYAWGAAALA
 FLYRQIGNASQRSQSIIGGCLTLLQCWSYFHLNIDRPKRTRQFPLALLWKGRQQSRSKNDLFKYRKALD
 DLDPNSVSWCPFEGDLDIVPQSFKDNLL LGRSRTKLIGPKVVEWHFPDRCKMQFGLCQVIPGEVPPRKNE
 KNHDEDLLEDMNTADEEWMRRRENIVENGGNG DESEYMQWFNSITVPKLHRDTSLEADIMNVQAAILQF
 DEVASTLSLEDLHPEEREAEIEEAVMSMSNALR VGDWYEASTTNKRKRREEQQQTDWSE

>ABA91431.1 GRAS family transcription factor containing protein, expressed [Oryza sativa Japonica Group]

MMHGLWVQDQGVVDHLAQLVPLLHECASHVTEGSFEKADFSFKKIRMLTIADGPLQRLSTIIIVDSL AHR L
 LSSIQGLPGALIDPSDYFEKSTLRAARHNFFKLN PYLSTGFVTINWAIMEAMEDEKVDLQVVHIVDLSCS
 AAHPWQWPKLDDDFHGRPGGAPELYLTVLHDDNDFLADMQSLLS KKAESLGVSFHFISVIGRLETLD FSN
 LRSTFQIKFGVAVAISCALQMHRLLLVDNLSSTSIAQLQKMANFTQPKQMASSVCS PASTLNYLQTPSP
 RTPKLLARLLSAIRALKPNIMLIMEQDADHNTLLFRDRFNEVLNYAALFDCFHAVAAANPGR TDERLRV
 DRMILREEIKNILVCEGVHRHERHERLDQWAMHMEESGFHNVQLSFS AIREAYVWQLKVQADNLR LCCTD

RGMFQDDMLSSATSSPASSVYSPSPSPSNGSWVQELSHDQQSVRLIGLLYQCAAQVSAAGSFDLANLCL
 ITQLASLDAPHALQRLAAVFADALARKLLNLILGLSRALLSSANSADAHLVPVARRHMFVLPFLKLAYL
 TTNHAILLEAMEGERFVHVDFSGPAANPVQWIALFHAFRGRREGPPHLRITAVHDSKEFLANMAAVLSKE
 AEAFDIAFQFNAVEAKLDEMDFDALRHDLGVRSGEALAVSVVLQLHRLAVDDGRRHAAAGCLTPVQIIA
 RSSPRSFGELELRELNTRLQLSPDASVSSLSPHSPAAATAAHPPTSTPKLGSFSLAVRSLSPKIMVMT
 QEANHNNGGAFQERFDEALNYASLFDCLQRSAAAAAERARVERVLLGEEIRGVVACEGAERVERHERARQ
 WAARMEAGMERVGLSYSGAMEARKLLQSCGWAGPYEVRHDAGGHGFFFCWHRKPLYAVTAWRPAASRRG
 HTRS

>KYP38396.1 OTU domain-containing protein 6B [Cajanus cajan]
 MLGVLCATRPKPWLLSLVHASLPRLAAASLSLSLSLSLSASPPRRHHSTACKLLAHAGGAASIWHAIRPR
 GAHGFR LAVHDPKGEKGSWNVAWDARPARWLHRSASAWLLFGVCACLAPPACADADADAFSPDETCGPRLP
 LDGKVEDEVSADYRITGVPADGRCLFRATAHGAACLRNGEKAPDENRQRELADELRAKVVDELLKRREETE
 WFIEGDFDITYVKRIQQPYVWGGEPELLMASHVLKTPISVFMRTGSLDLVNI AKYGEEYRNQKEISIDVL
 FHGYGHYDILETL

>EMS67907.1 ABC transporter I family member 1 [Triticum urartu]
 MALNLPISLKSPPSAGAGSRAAGVGEEQEMRKELELLTKPRKGEKGSFLVKGLPGSTSGLSKKGQINAVP
 FHEAILWNGHDVTS PGVFQQYKQLQNLNWSLKDVAKEKLTVLENVQWFELLEKGHGRSAPAIELMGLGRML
 NDKARMLSMGQRKRLQLARLLAIDRPIWLLDEPSVALDSEGVKLEYYIAEHRKKGIVIVATHLPIEIE
 DSMNLRLPQRELYELEHSADQMRLQ

>XP_013446519.1 multidrug resistance protein ABC transporter family protein
 [Medicago truncatula]
 MAILDVLLGTINVTFFYVILIWVLFDSLRSQSTRNNLQHFKHTPTIFSYITVFFNAVISLLNIAFVFDYD
 TRGIIGFNYSFGLTWVLTATMVSFYSMKKTLENKRFPFVLILWVFFVTFVHIISLSLKLKVNKSKSINLW
 ILLLEKNTVETVSLPMLLMCFNAFPNVCVREQSEIEERLLQKEFESSTFEDEEAFKAGVWSKLTFRWL
 NPIFEMGRIQKLEHVNVPSVPPSETAASASSMLEESIRKQKLECGSLSKAIVDSVWKSALNAVLAGVNT
 IAAYIGPLLI SNFVNFLSNDDNSNIKYGLILAFIFFLAKTVESLSQRQWYFGAQRIGIQVRAALMALVY
 SKSLMIKCGGPTHGKIINLINVDVERIGDFCWVYHGVWLLPVQIILALVILYINLGCPTSAALAVTILV
 MVCNTPPLANMQEGLHASKIMEAKDSRIKMTSETMKNIRILKLSWESTFLQKLLQLRDTEKKWLHKYLYLC
 SAVATLFWASPTLVSVFTFGACILVKTELTAATVLSALATFRILQEPIYNLPELISMITQTKVSVDRIOE
 FIKEEDQNQFMNRHASKTSTIAIEIKPGEYAWENDQFLKKPTIHIAEKLMIKKGQKAVVCGPVGSGKSS
 LLCSMLGEISLVSGAATKVYGRSYPQSPWIQSGTIRENIFLGKQMNKDFYENVVDGCALLQDINLWSD
 GDLTMVEERGINLSGGQKRIQLARAVYNDSDIYFLDDPFSAVDAHTGSHMFKECLMKLLYDKTVVYATH
 QLEFLEAADLILVMKDGKIVESGRYRDLIACPHSEFVQMAAHEETVCQIPCRKDDSVCCRPCQKNPTEI
 AEENIQEIMLDWKRTRREEAMTGRVKWSVYSTFVTLAYRGALVPIILLCQILFQVMQMGSNYWMWATEK
 KGRVDNVQLMGVFALLSGSSIFILGRTVLMATVSVETAQRLFHGMIA SVFRAPVSFFDTPSSRILSRS
 STDQSTVDTDIPYRLAGLVFALIQLLSIIVLMSQAQWQVILLFFVVLALS VWYQAYYITTARELARMVGI
 RKAPILHFFSESIAGAATIRCFKQEKIFLTKVMVLIDDYSRVAFHNYATMEWLSVRINFLNLFVYFVLV
 ILVTLPRSAINPSLAGLVATYGLNLNVLQAWIWNLCNVENKMISVERILQFSNIPSEAPLI IQDCRPEP
 EWPKEGKIEFLNLHIQYDPSGPMVLKGVTCVFPQKQKIAVVGRGTSGSKSTLVQALFRVVEPLEGCILIDG
 VNISKIGLQDLRSKLGII PQDPTLFLGTVRTNLDPLEQHTDQDLWEVLRKCHLAEIVQQDPRLLDAPVAE
 NGENWSVGQRQLVCLARLLLKRRKILVLDEATASIDTATDNLIQKTIREETSGCTVITVAHRIPTVIDND
 LVLVLNEGTTAEYDQPSQLLQANSSSFSKLVSEFLRRSSQSNCKKR

>AFW60210.1 benzoxazinless1 [Zea mays]
 MAFAPKTSSSSSLSALQAAQSPPLLLRRMSSTATPRRRYDAAVVVTTTTARAAAAAVTVPAAPPQAPA
 PAVVPPKQAAAPAERRSRPVSDTMAALMAK GKTAFIPIYITAGDPDLATTAELRLLDGGCADVIELGVPC
 SDPYIDGPIIQASVARALASGTTMDAVLEMLREVTPELSCPVVLLSYKPIMSRSLAEMKEAGVHGLIVP
 DLPYVAAHSLWSEAKNNNLELVLLTTPAIPEDRMKEITKASEGFVYLVSVNGVTGPRANVNPRVESLIQE
 VKKVTNKPVAVGFGISKPEHVKQ

>AMD61995.1 acetyl-CoA carboxylase carboxyltransferase beta subunit
 (chloroplast) [Tetraplodon fuegianus]
 MSMLNWFEDKRRFRGGLIGASIEKATKGYILNERKRLKVNNTNGLWTRRDNCENILYVKFLKQNKSIICEC
 GYHLQISSSTERIELLVDRGTWQPMDEDMVARDVLKFSDEDSYKSRVIFYQKR TGLTDAIQTGIAKLNKLN
 IALGVMDFQFMGSMGSMVVGEEKITRLIEYATEKSMPLIIVCSSGGARMQEGTSLMQMAKISSVLQIHQA
 KKKLLYIAILTYPTTGGVTASFGLGDI IIAEPKAYIAFAGKRVIEQTLRQKIPYGFQVAESLFDHGLLD
 LIVPRNLLKGVLGKIFELYGLAAYKERNNYFF

>XP_005645574.1 glycoside hydrolase [Coccomyxa subellipsoidea C-169]
 MLVALVPCLFILAAANQVAPAPAPSPSDPTKDCNPAVQAVSGVCKPFRRLTPWSCMVPDPKNSLPEYPRP
 QLMRPDYLSLNGVWEWERADSWQSGPVPVNRTEGSI VVPYPLEGDLSGVL DHNKTGVFRMWYRRSFQVA
 KDWALGHNTNVLHFGAIDWEAEVWVNGERLQHQGGYDKFYDITEALKKGNGTHELIVRVYDPTETA
 HIPTGKQRHYVDRDIFYKGS SGIWQTVWLEPVPAAYVTRLDMPDIDSSSLNVTVLGSKAAEGLTVQID
 MYIPQNGTAVVSAQGIVGRPLSVTIPNQRLWAGVADPFLYDVRVRI LQTPSTSSFRKAMPVNLVSAH
 DEVMSYIGLRKWSLDKAGDIRELRVYLNNKPFVFFGPLDQGYHPDGIYTAATDEALAWDLEAIKALGMNM
 VRKHIKVESDRWYHADRLGLMVWQDAVSMFWEKPYTEGEQFRTAGVKAQFELEMRRMIEEHVSSPSVGI
 YTVFNEGWGQFDTQRMVQLGKSLDPSRLWDAASGWIDPQDR TAFGDFNNKGPWYHYKGYVGLRDDHNY
 PDAKASGATSTRANVAGEYGGGLGLFTTGHWTWVDSRDKDRFIGGLHHVRDLVKSSSGLSGAVYTEITDVED
 EVNGLYTYDRQVFKMPDLEAVQKEIKDLLSTDVN

>GAQ77714.1 Cyclic nucleotide-binding domain containing protein
 [Klebsormidium flaccidum]
 MSAFWRKFRRTAVRFLNKAKGPAEKSADQASERELVACQVAILQKSPEARTDEDLAVIDDILKSVKFTAEL
 PLRVRLEICKNCTYKQGAAGQELFKQDGSKEFFIILSGIVTCSAINTSTGQASVVARLTIHQSFGERGL
 QSKNATRQETVTLESRCDFLVISRDNFDVINQHLQSNESRKKVDFLRNVGVFAHLPQSSQLMIATAMAAR
 TCPKNSVIKRQGGDSTEIYFLVTGTCRVIRDVTFSSRESIQLSLVLVDAIDSASDAAAGRRRPGDAAA
 PDTILGLPLAPTSAAGTFALPASLAVPLTRTEGQRSSAEAAAGSEGGQESGTPSTAREAEAEEDDKEGAL
 IGVDKRADDPISAETVESRKSRSVFSQLPSSTPPERGRSGTSSQSRKSSDPVGRTPDSAPSTSGRERTDF
 PARAYAGAGENVSPKRTSSGPPSFRAPPEPFFPAGINPSHGERIYPWERSQKWPKTTPRSR SASPPHVTE
 SPRARRARASADPYSSKRLTVTQMLRAMPKRKEKSRRLFLELGQLNPPAYFGEAGLLRNEARWTTIVAAT
 NVELYVMNKLDQVRVGRDIIKQMEQRIPEYPTDDAIRKQFLRNVLWSNYKAALVSEVTAHRSGNGKPTS
 AVPLPSANFRQPDGQTVRLRLLERKTNASAPNVKLLKGAVMPIPTLD

>XP_013900990.1 outer membrane adhesin like proteiin [Monoraphidium
 neglectum]
 MQMHQHASLAVLLLVLAAAMVSSSTGAAPVSLVVTNDQDTGPGSLRDALSTVQPAATTIITFSSSLAGSTID
 LTASGSLVLPALVGNASLTVQAPAVAGSPVLIVQPLVAGGNPETALGSKTATSATNPVVEFINFQFAE
 MTFAYVDGELRFTDCEFFSSQGSGLQALVTKSALIVAGDR SKLTLTRVRMTGFRFTITGNPPGPAANGAAI
 RADPSISAAAYPLRPTGVVATDLRVISCEASNGAIFVQDTPFTCTRCTFLSNVGTNGGGLYATGTGAVAL
 TDCQFSANIAQTAGGMHVINGPTVVVRRVVFIGNQATFTGGGLHYSSSTSEAVLIYESSFVGNALVSSA

AARGGGAYIASGNSLPTPPYATIQIHSTTFYNNELSGNPNPGEESGLYLQLLQGPPANASIIISCILWRDG
 FNAYPVYYTSPFSLANAAMPFYNNDSGAQAAGVVGFGNMDANPLLHPVSTFWIIPFPSSPVIDAGANPN
 GDTADVRGLARVVGAAADIGAAENQAPVAVDGALTVNQGTVVTAQAPGVLGAYDPDGDAPIRAVLVSGP
 TNGFLTLSSSDGSFQYTPSVNYVGPDTIAFRASEGPRLSAVHQFIITLVRVNHAPVAGNVAYSVKQDFTLN
 QPAPGLLASSSDPDFGDRLSVAAVIPPANGALVPQADGSFRYTPNHGFSGSDSFSFQVMDGAGATSTAQG
 TITVVSNQPPAIATNLNYTGKESTALVVDSAAGVLSTASDPDGDVPLSALLAPGGSPTHGALTLSANGSFT
 YSPTPGYSGPDVFKFTVVDTRNGTVTRQANINVVANLPPVATDLSYTAQSVRLSVNASVGLLSTASDPD
 GDLPLVASVVAGSGPAHGTLNLQPDGSFTYAPTGYSGADAFNFTVTDHRNGTVSRQATITVDANLPPVA
 TDLSYTAQSTNLSVNASAGVLSTASDPDSDLPLVASVVGSGPAHGTLTLQPDGSFTYTPSPGYSGADA
 FNFTVTDHRNGTVSHQANITIVANLPPVATDLSYTAQSVALS VNASVGLLSTASDPDGDPLVASVVAR
 SGPAHGTLNLQADGSFMYAPTGYSGADAFNFTVTDHRNGSVSRQATITVVVNLPPVASDLSFTAKQSVT
 LSVASAGLLSTASDPDGDPLPLVASVVAGSGPAHGTLTLQPDGSFTYAPTLGYSGADAFNFTVTDHRNGS
 VSRHATITVGEMLNAKRWTLKSNNMYVCLYT

>XP_013452787.1 pentatricopeptide (PPR) repeat protein [Medicago truncatula]

MSSLVHMYLKCNRIDDAQKLFDMCDRDVIVWSAMIAGYSRLGNVDRAKEVFCEMRKEGVEPNLVSWMGM
 IAGFGNVGLYDEAVRLFHEMVEGFLPDGSTVSCVLPVGNLEEDVLMGKQVHGVIKLGLESKYVVSAL
 LDYGRGCAPEMSRVFDEIDQTEIGSLNAFLTGLSRNGLVDTALDVFKKFKAGELELNVTWTSIIASC
 VQNGKDMEALELFRDMQADGVEPNAVITPSLIPACGNISALHTGKEIHCFSLRKGI FDDVYVGSALIDMY
 ANCGRIRLSQNCFDEMPFRNLVSWNSIMSGYAMHGKAKETIEMFHMLQSGQKPDSTFTSVLSACTQNG
 LTEEGWHYFNSMSKEYDVKPKMEHYACMVTLLSRVGLKEEAYSIIKEMPFEPDACVWGALLSSCRVHNL
 SLGEIAAEKLFVLEPDNPGNYILLNIIYASKGMWDEENRVRDMMKSKGLQKNPGCSWIEIGHRVHTLVSG
 DKSHPMKEILEKSEKLSIEIKESGCLPMTKSVLQDVEEQDKEQILCGHSEK **LAVVLGLINTSPGQPLQV**
IKNLRICDDCHAVIKVISRLEGREIFVRDTRNFHFKKEGVCSCADFCIYMPPTVIPGGDLVKVRRTFCI
 ISNSTNVAEVFGRKFAMTTSITCKPSVHLSTETDYKEVGAESSEGGKEDIDNY

>XP_005647413.1 argininosuccinate lyase [Coccomyxa subellipsoidea C-169]

MASAQAVQEAAITGEKKKLWGGRTGATDPLMEKFNESLPFDKRMWREDIRGSQAYAKALARAGVLTAAE
 AEQIVEGLGKVELEWEAGTFKIVSGDEDIHTANERRLTELIGSVGGKLHTGRSRNDQVATDTRLWLALYAL
 QDTRGYLADLIKTAIDRAELDQVIMPGFTHLQPAQAVRWSHWLGHAAAWQRDDMLRDLMPRVATLPL
 GSGALAGNPFVDRQFIARELGFVGGVCPNSMDAVSDRDYVLDLTLYFVAVHMAHLRWAEDLIYSSGLF
 KFVQCSDAYATGSSILMPQKKNPDALELIRGKAGRCQGNLAGLLAVMKGAPTYYNKDFQEAQWLMFDSVDT
 ANDCIRIATGVLSTIRINPDRMLAGLSPDMLATDLAEYLVRKGVPPFRETHHISGAAVKMAEDR **GTTLFDL**
TVADLKTIHPLFEDDVLEVWDYNRSAEMRDTEGGASKRSVLEQAQKLKRYLGAEGL

>XP_003624729.1 peroxidase family protein [Medicago truncatula]

MKQISIIIFFFILPLAFADLELGFYASSCRKAESIVKQVVQKRFNRDKSITAALLRMHFHDCFVRGCDAS
 LLIDSTKNNISEKDTGANDSVRGYDLIDDVKEAIEAACPSTVSCADIVALATRDAVALSGGPKYNIPTGR
 RDGLIANRDDVDLPGPNIPIGALSQFFAAKGITTEEMVTLLGAHTVGVVAHCGFFASRLSSVRGKPDPTMD
 PALDTKLVLKCKSNSDGAFLDQNTSFTVDNEFYKQILLKRGIMQIDQQLALDKSTSTFVSNFASNGDKF
 VKSFATAMIKMGKVGVLVGNENGEIRKNCRVFNKRN

>AKF41605.1 ribosomal protein L9, partial [Sonneratia caseolaris]

PEGVSIKVNKAVIEVEGPRGKLTRNFKHLNLDLQLIKDEDTGKRKLKIEAWFGSRKTSAAIRTALSHVGN
 LITGVTKGYRYXMRVYAHFPINASITNGNNSIEIRNFLGKEKRV **KVDMLEGVSVXPSEK**VKDELIVDGN
 DIELVSRSCALINQKCHVKNKDIRK

>KHN05237.1 Cadmium/zinc-transporting ATPase 3 [Glycine soja]
 MSSLTSMTPQKAVIAETGERVDVNDVKINTILAVKAGDAIPLDGIVVEGKCEVDEKMLTGESLPVIKELD
 SVLWAGTINCKNYCVGKRHMVARMSKHVEEASSRKSQTQRFIDNFAQSTIFLQVTVGLISAGIAVVPAAAL
 KVHDIKPWFHLVIVVLLIACPCALILSTPIAIFCALTCAAISGLLLKGGDYIETLSGIKTVAFDKTGTIT
 RGEFTVTFDFSVVDDISIKTLLYWSSIQSKSSHMAAALVEYGMSNSIKPIPKNVENFENLPGEGVLGTI
 DGKDICIGNRRIVFSLVDTCRAGALEAIEEPKLLGVRSVMLTGDRSQDAIASPSEKAVI IENFK **KDGLIA**
MIGDGIN DAPALVKKALSLSGEVTKINVTQLPKTNKLTVPKSVNQVDGRRPLFRPKQFVTLTSMDFL
 NEMEREKPIGFTVEQLRIATDNYSLLGSGGSGAVYKGSFSDGTSIAVKVLRGSSEKRI IEQFMAEVATIG
 KVHHFNLVRLHGFCFESHFRGLVYEMANDTLEKYLFCCKSMFLSFEKRHEIAVGTPRGIAYLHEECQQR
 IYYDIKPGNILLDNWFCPKVADFGGLAKLCNRDNAHITLTRGTPGFAAPELWMPNFPVTHKYDVYSFGMLL
 FEIIGRRRNHNINLPESQVWFPMWVWKRFDAEQVRDLITACGIEDQNCIEAERFVRVALSCVQYRLESRP
 IMSVVVKMLGGFIEVPKPMNPFPHLVDWSSPHTLVQASQINADKSIKFDSSVMLTKSVHVITTPILTKEYE
 IYLKKFQ

>ADE87856.1 maturase K (chloroplast) [Gavilea littoralis]
 MYINGRITKIFRLKKTDFRQQNFLYPLLLQESXSYLAHDHSFNFSFIFYEPVEILGSDKKSSLVLVKRLIT
 RMYHQKSLIYSVNYSNQNGFWGHKNSFSSQILSEGFGIILEIPFSSRLVSSFEEKKRPKYQNLRSIHSIF
 PFLEEKFSHLNSVSDLLIPHPHLEILVQILKCWIKDVPSLHLLRLIFYEYHNLKNFITSNKFIVFLKR
 KKRFFWFLHNSYVYECEYLFLEFLRKQSSYLRSSTFGVFLERTHFYKIEYLLGVYFNSFQRIWFLKDP
 INYVRYQKAILASKGTLILMKKWKYHLVNFQFYFHFWSQPYRIHIKQLPNSSFSFLGYFLSVQKNTLV
 VRNQMLENSFLINTLTKKLDTIAPVISLIGSLSKAQFCTISGHPISKPIWTDLSDSDILDRFCRICRNL
 RYHSGSSKKQVLYRIKIYILRLSCARTLARKHKSTVRTFMRLGSGFLEEFFMEEEQALSFI FLQKR **ALAF**
PLXELHRERIWYLDIIRMNDLVAHS

>XP_005650585.1 NADH-ubiquinone oxidoreductase [Coccomyxa subellipsoidea C-169]
 MLVGRSRAAAARELQKSLRSFSAQASPPPTPPPSEKTTYGGLRDEDRIFTNLYRQGDPFKIGAMARGDW
 YRTKDLVEK **GSDWIVGEMK**KSGLRGRGGAGFPSPGLKWSFMPKVS DGRPSYLVNADESEPGTCKDREIMR
 HDPHKLIEGCLIAGVGMRARAGYIYIRGEYYNERLALERALSEAYAKGFLGKNACGSGYDFDLNVAYGAG
 AYICGEETALIESLEGGKQKPRPKPPFPANVGLYGCPTTVTNVE TVAVAPTILRRGPEWFASFGRKNNAG
 TKLFCISGHVNTPTVEEELSIPLELIERHAGGVRGGWDLAIIPGGSSVPCLPKRICDDVLMDFDAL
 KEAQSGLGTAIVVMDKSTDIIDAIARLSYFYKHESCGQCTPCREGAPWLYDIMTRMKKG DARLEEIDML
 WEVTQKQIEGHTICALGDAAAAPVQGLIRHFRPEMEERIGRFCEQIAA

>EMT04584.1 Medium-chain-fatty-acid--CoA ligase [Aegilops tauschii]
 MKKRLANGTIK **CDS PGVDRLGGLPDDVLGR**ILGFLPTPLAVRATQLSRRWRRLWPAHV FALNLSVQDCKN
 RGVGVRFPDL CARALARFP IFSIPSISLEFCTRQIGVGKAKAWYAEAMERAAGSVSVTVLRGAFPLALP
 RFTQAEALSLTLTHRIDLLEPAAGDDPLLQACSLTTEHLELDLVLDPDMLDNWLGPEERGS PAGEDLMR
 HVPPLPRVTVLSLKV RWGIGGDVGPCLASLLSRVPSVATHVGPAPYCLTVLGGAVVPRGECRWGRCVDE
 RSSGGQLGSLREIVVHGLRGT DGEESLVEVLLGTVPPIERISLRCSKLGPNPANSCPLTPLGFLELAA
 TAYRDCPSVIYHDTVYTSQTFRRCLRLASALASIGIAHRDVVSVLLPNVPAMYEMQFGVPMSGAILNNI
 NTRL DARTVSVLLHHS GSKLVLVDPASLQLE DALRLLPPEHSAPRVVADDPHEKELL PAPATALTYER
 LLEKGDPDFAWVRPANEWEPMILNYTSGTTSAPKGVLHCHRGIFLITLDSLVDWGMPTYPYTLWTLPMFH
 VIGWSFPWGMVVGTTNVCLRRVDAADVAAISRHGVT HLCGAPVVLNMLANAPEAVRKTLPGNVEILT
 GAPPPAAALHRTEAIGFKVRHGYGLCETAGLATSCVWNAGWDKLPATERARFKARQGVRTPATAEVDVVD
 EKSGRSMRRD GSTVGEVIRGGGGDPGDGWFYTG DVGVMHPDGYLEIRDRSKDVIISGGENISSVEVESV

LGELTEAELVAWSRERIPRYMVPKTVVFRDELPKTSTGKIQKYVLRNIAKEMGPTSRGVSSKM
 >EMT29259.1 Peroxidase 70 [Aegilops tauschii]
 MASDLSGVAQWRGLGVRVWMTMDSRRRTTSLGTVVASMAERPSKVDASVSDLRMDRWWKEVTAIAACACATC
 AAVANAQLSENYYGSSCPAALLTIRTPVATAVLLDRRMGASLLRLYFHDCFVQASPQDLPLLPWSSFNI
 GCDASVLLDDTPSFTGEKGAGPNAGSLRGFEVIDRIKLLLELICPGTVSCADILAVAAHDAVVHLGGPSW
 TVLLGRRDATASASLANSDLPGPNSNLNDLLAAFSSKGLSSTDMVALSGGHTIGRAQCQNYRNRIYADT
 DIDGTFAAASLRGDCPQGGNDGNLAPLDASSPDYFDNSYFSGLLSHRGLLHSDQALYDGGSTDGLVRSYA
 SNNDQFGSDFAAAMVKLGNLGVLTGAYGEIRVNCRAVN

>EMT10537.1 DNA (cytosine-5)-methyltransferase DRM2 [Aegilops tauschii]
 MADRDSDDNVKFDWESDGEAEPSSAPAFRNSDAPGPSTLDSNGQANEEAPSTALIEEYVAMGFPEIV
 VKGMKEIGNTQRHSDADALLELILTYQALGADDVAVGNCSTSGCAPQGVVEEEDDDDLDFENWDGDDDDVG
 GRGTNCDDPGDEDFLQEMSQKDKKINSLVDMGFPEDEANIAITRCGVDADLCVLVDSISASLVAGDFNSR
 NISDHQVMDFDSFGGRKKARLMEESSKRRMQYAQSGSSFAGSHDEPTRIPDPMVGFNLPSDRKPSMT
 RMLPEQAVGPPFFYFQNVARAPRGAWTTISKKFYDIQPEFVDSKYFCAASRESGYIHNLPINREALLPF
 PLKTVFDAFFHYKKWWPSWDPRRQLNCLQASVATAKLTQIQRTLARSGNPSVQKHVVECKTWDLVVVG
 KKNVAQLEPDEMESLLGFPRDHTRGVVKTERTYKSLANSFQVDTVAYHLSVLRDMFPNGVNVLSLSTGIGG
 GEVALHRLGIRMRTVVSVEEGETNRRIFEGWWDQTQTGKLVQIANLKSILTNERIASLVGRFGCFDLVIGG
 SQGQTYFWKNTSCLMSLSSCMIFKVWENKPIRLPGYQVSQVCKVQNLGSDFVIGILHGEEDEDERERKKRD
 KKERRRSGRDGRGSDDEEEERRRRKKKHGDRGDKERDSKERRSKEKEKSKRKKDKDAPSLKQDIKEISKDD
 YFAKNNEFATWLKEEKGYFSDLSESARDIFLKFVKQWNGKGLPSQYYVGITTGPRSAHNWNKKA

>XP_001421601.1 TRP-containing protein [Ostreococcus lucimarinus CCE9901]
 MSLPPELLDFCVCKGPGDTISEQQQLADAGVKRAVIHGDAALDTKNALAVTSLCKVRLKWKFDDEECWKQD
 EFVVYDGTVPSGVELAIMNLHEGESVAFNARGEAALTDPEPLPKADISQVFRSGQAEVLVDKLERKDKFD
 MSASDKIAFGEKMNEIGKKLFANGRLARAAEVVKRAAEVFAVLEPEDEDVNPNGMKNNACHEVARKVYL
 NMALVMYKLENFTECELYCCDVLDTRVDEPTYSQIKSKALFRRGCARLKLKRVSLGSRITADTAEADFFL
 ASELDSSIDVAAHFDACTELREQLLASGVEIPSDDVETPNVEDEMNHYPDIEHALRQIRRENLDGLAF
 EHPQKRAAILEERMVYRRLMDETPWSALRGTMFDADSEELQKTATTAREVDT

>KHG27763.1 Ubiquitin carboxyl-terminal hydrolase 13 -like protein
 [Gossypium arboreum]
 MTLMTPPPLDQEDDEMLVPHNDFVDGPPLEVAKEAASKVDAQAVDDPPSGRFTWTIENFSRLNNTKKLYS
 DIFFVGGYKWRILIFPKGNNVDHLSMYLDVADSATLPYGWSRYAQFSLAVCNQIHNKYTIRKDTQHGFNV
 RESDWGFTSFMPGLGELYDPSRGLVNDTCVVEADVAVRKYVADYWSHDSKKEGTGYVGLKNQGATCYMNSLL
 QTLYHIPYFRKAVYHMPTTENDMPTGSIPLALQSLFYKLYSDTSVATKELTKSFGWDTYDSFMQHDVQE
 LNRVLCEKLEDKMKGTVEGTIQQLFEGHMHNYIECINVDYKSTRKESFYDLQLDVKGCKDVYASFDKYV
 EVECLEGDNRYHAEQFGLQDARKGVLFIDFPPVLQLQLKRFYDFMRDTMVKINDRYEFPLQLDLREDG
 KYLSPAADRSVRNLYTLHSLVHSGGVHGGHYAYIRPTLSQWFKFDDERVTKEDVKRALEEQYGGEEE
 LPQTNPGFNNTPFKFTKYSNAYMLVYIRESKDKIICNVDEKDIAEHLRIRLKKQEKEQKRKEKAEAH
 LYTVIKVARNEDLVEQIGRDIYFDLVDHEKVRSFRIQKQMPFNVFKEEVAKEFGIPVQYQRFWLWAKRQN
 HTYRPNRPLTYQEEAQSVGQLREVSNKANNAELKLFLEVELGPDLPVPPPEKAKEDILLFFKLYDPFKE
 ELRYIGRMFVKGAGKPTIILARINKMAGFGPDEEIELYEIEKFEPNVMCEHIDKKLTFRTSQLEDGDIIC
 LQKYSEVASEQCRYPDVPSFLESKLNHYDDVVERVAHHLGLDDPSKIRLTSNHCYSQQPKPQPIKYRGVE
 HLSMDLIHYNQTSDILYEVLDIPLPELQGLKTLKVAFHATKDEVVIHTIRLPKQSTVGDVLDLTKTKV
 ELSHPGAELRLLLEVYHKIYKIFPLSEKIENINDQYWTLRAEEIPEEEKELGPHDRLIHVYHFMKDATQN
 QQVQNFGEPPFFLVIHETETLAEVKVRVQKRLQVPDEEFAKWKFAFLSLGRPEYLQSDVVSSTRFQRRDVY

GAWDQYLGLEHSDNAPKRSYAANQNRHTFEKPVKIYN

>EMT18915.1 Disease resistance protein RGA2 [Aegilops tauschii]

MEPLLSAIVGDLVSRALSMVTQRYLQSKGAEELKQLRLRAVLLRIDATVEEAEGRHITNQAMLQQLQMLR
 RGMRYGYYTLDTFRYRAHRDEAYGEASTRRSLALSRSFSPSSSRRTFFSVCDAPSRDLVLDAGSCSNLDM
 LSSLERMIGDMQEFVFMFLASYPRISRQPYSAFLYFHDQVMFGRQMEKETILSFLLCREAARNGSLGVLPII
 GPARVKGSTLVEHVCHDERVRRRFSLIVFFSGNDLNGGNLNLATLKGSGVIKHQNLASSPRGDSLAVIELAG
 DMDEETWRGLYTSAAANHLTPGSKIILTSRSEKIASLGTAQSLVLKFLSQEAYWYFFKTAAFGSTDPGEHP
 NLAALGMEIVVHMNRSFVAANTVASLLRTNLDTRFWRRVLRCLRDFASKHLSMFGEHPTNLLQKDQPVYM
 WMAKTSNVVVRDIYHEPSPQISEVPRITAQDVLGRVTSEGAFAVAWRSRIPPICYTYLVSFVVSRTD
 EHVVLACSMDLQLYLHKPSTELDITGLVELLTSQIDRTIFICKAGWQRKQVDPQHYRRAGAFYKISKRLCR
 FQQLLHCTVYM

>KHG08814.1 FAST kinase domain-containing 3 [Gossypium arboreum]

MEFLNPFPTQTYLKPLIFMPKQVHNLSLVKLRGGLEFPRRNFAVSGKNSSISTGNSVNDDEMEDWELEF
 VGEIDPFSYQAPKKRKKQEKSRVLNDEEGMDWCLKARKMALKSIESRGLAHKVEDLIKKKKKKKRLAKK
 DKISKEVEEIEEGFDQDIGNEGFDNQIDDVNSRLRDTVSSMGDGMFLEKKEKAMEELVQKLAQISGPS
 DRKKEVNLNKDIIQSQTAEIEILEITSEMIMAVGKGLSPSPLSPLNIATALHRIAKSMEKVSMPSTRRLAF
 TRQREMSMLIGIAMTALPECSAQVSNVAWALSIGGDMFLSEMDRVAEVALAKVTEFNSQNVANIAGA
 FATMRHSAPDLFEGGLAKRAAAI IHSFHEQELAQMLWAFASLYEPADTLLQAI DTVFNANQIKCLSSET
 VKDDEETGVENSRIKFGGIPDPPVLTLSRYQLGNIAWSYAVFGQLDRMFFSHVWKTLSNFEEQRISEQY
 RGDVMFASQVHLVNQCLKLEYSHLQLSIDGELESKIVRAGKTKRFNQKTTSSFOKEVGRLLVSTGLDWVR
 EHPTDCYTIDAVLVDQKVALEIDGPTHFSRNTGTPLGHTVLKSRHIAASGWRVLSYQEWEEHEGEVEQ
 LEYLRKILEDHLG

>XP_011397355.1 Nitric oxide synthase-interacting protein-like protein
 [Auxenochlorella protothecoides]

MGKHSKNAGVMGAESLTYHERKALGFQTVSERLKGDSIGNYYDCRLTLRPAVDPVCTPDGFLFSREAILE
 NLLAQKKANKRKLAAWKAQQAERQRKAEKAAIEAAVLAFLDRQNHAGASAGAAARLTASIAAEADALL
 ADKRVASGAVNIRAHGERMAEMRAFVLPSTPEAAANLEAPDGVTRCPASGKPLRLDLTSVALTPLPDG
 AAQEFMDPVSRLPFTNATRLMVIKPTGDVVSEETWRTCILPEGSYRGVPLGPEDAIELQRGGTGFAEHDK
 QAQAKKFFHLGPGSGLSDRRGQHGGPVSKFGLRIAN

>XP_003621296.1 inactive poly [ADP-ribose] polymerase RCD1 [Medicago
 truncatula]

MEENPSKALDGVSLNKKRKRDSQGVSRRTGVLAPTSVVVKQMTLAR DICKPDDSCPDISKSLVSYLNYK
 KSGIPKRLMFYKNGKWVDYPEDVVDLVKKDFKIKKAI VEVELNGQEVVLDLFLHMNHVDMKTCLQPIGWI
 DEEGNYFFPKVFGVSTEEPNNIREGEEERLNKKEPHEIKLHISNEINGADESKLRKYSRESNATKNVK
 AEGHAMTTKIGIQNVAIDINQEPDIDLNDYSESLYGKLDVNSTQKMFLKGMSSLGISESDIVGIYRSSGR
 SMQMRLQLFEKQADI IKGIRGVANIRYAWLACSKEEISTMMEYELSHYELSPSKCIYGPVHLAAITHPF
 VCSLSCDEEDNGIKHMFVLCRVIMGNMELLRPSKQLRPSDCEYDNGVDDIQCPKCYVWVNMNMNTHIYPE
 FVVVFSKSPFGFEENICSNESKKQGYRHSILDTVKATGVPANIRRTPGSASVPFVLLSSLRNMVPSNIMF
 LLDAQYALLMSNKISRGCVLKLRRLIVANNVLKSAVYNVQSKRPPIYGSNGSNVNQE

>KYP65899.1 TMV resistance protein N [Cajanus cajan]

MDDEELRKGEEITQALLKAIQESRIAIIFSPNYASSTFCLQELEDIMDCLNHKDKLVWPVFKVDPSPDV
 RYQKGSYAEALAKHESRISDKDKVNWRSILQAAANLSGWHFKDGCECEIIGKIIQEVSKRINRNTLHVT
 HYPVGLSERVQKVKSLLDLESNDEVHMGGIYGIGGIGKTTLACAVYNFIAHQFEGLCFLADIGKKIMEHE
 LVQLQEKLSDIVGKKDTKLGNIKQGIPIIKSRLSKKKVVLLIIDDVDKQLEVLGGLEWFGPSSRIII
 TSNKHLRLRVHGVERTYEMEGLNQEEALELFSWNAFKRKRVPNGYLDISKRVLQHSNGLPLSLEIIGSDL
 YGKTELEWKSALDTYERIPHEDIQQILRVSYDGLKEFEKEIFLDIACFFKGYRLNDVISILHSGRGSAPN
 YAIQVLVDKCLIKIDLDCVKMHDLIEDMGREIVRLESPPKPGERSRLWFFKDILRVFKQSKGSDKTEIIM
 LHLPKVKEVQWDGTALEKMNENKILVVENAHFSR**GPSVLPKSLRVLK**WTAYPESSLPANFDPKLVILDL
 PMSYFTFADQVIMKFKSLMEMNLSGCELLKMPDMSGAPNLKLDLQCCRNLVEIHSDGLLNRLEWLN
 HNCTSLTSFPEILGKMEKMRFLDLSDSGIRGLPFSIGNLVGLKYLYLDGCKWLELPESSIFTLQKRLEA
 KYCERLALVQNGEGQGEIMSSANSMSYINFECYHLTDQFLCPVLPWMRNVTILSLNYSNITILPSSI
 SVCLSLKALYLNCKRLREIRGLPPNFKILSAFNCLSLSSSESNEMLLNEVSYLCFEYCLILKVNA

>XP_010105627.1 DNA polymerase delta subunit 3 [Morus notabilis]

MAQIETLGILQDIESLVSDDLQVVSQKWLRSNYLLSSNSAKRLLQEFVEKHGNGFEVVYALSGWLKGDSP
 SYHIRLVTGTKLEEAKQEFDFGSCSVQVYSVQACIPKDPVAVLWNVEFVQAEELFKQSTIDNCLRDNRFC
 ISNSFIKRNVEGTPSSVSAAQFQSGVGVLSKSGLTSSQKIAAPQSLQKGEKYSSKVGFPSPNLVKDVKS
 ENTGLQDPVNQAPTNEKVPMSAANKKKVQNGKGSSTEGSLANFWSRASAKTKASCPAETSSGASAEAO
 ICAREAVNVEGASSDDDGQDNFKRASNGESTRKRVRVLYFSDDDEYENAVSLSSPEAGDPKPKGESINS
 LVPEKSNAKLDKREDDKSEVKEKVTIDKGSTKLMKNDASAIINDKITGTSPKEKALCSVQENDVNKRDKL
 TKTTPESPKRKRVLKTQIDERGREVTEVIWEGEDTEAKKATNGTEEKADSNTRKK**IESNTSTNNINRPPA**
AKKSPA VGNTAPSNPTGKAGNKKTGKNDKSKQGNILSFFKRV

>NP_001105219.1 indole-3-glycerol phosphate lyase, chloroplastic precursor
 [Zea mays]

MAFAPKTSSSSSLSALQAAQSPPLLRMSSTATPRRRYDAAVVVTTTTTARAAAAAVTVPAAPPQAPA
 PAPVPPKQAAAPAERR**SRPVSDTMAALMAK**GKTAFIPIYITAGDPLATTAEALRLLDGCGADVIELGVPC
 SDPYIDGPIIQASVARALASGTTMDAVLEMLREVTPELSCPVLLSYYKPIMSRSLAEMKEAGVHGLIVP
 DLPYVAASHLWSEAKNNNLELVLLTTPAIPEDRMKEITKASEGFVYLVSVNGVTGPRANVNPRVESLIQE
 VKKVTNKPVAVGFVGFISKPEHVQIAQWADGVIIGSAMVRQLGEAASPKQGLRRLEEYARGMKNALP

>ACD03241.1 UDP-glycosyltransferase UGT705A4 [Avena strigosa]

MASNDNVPTAVTSSINKKLRVLLIPIILATSHIGPFTTELAISLAATNDAVEATVAVTPANVSIVQSMLEHR
 GGHSVKVATYPFPAVDGLPEGVENFGSAATPEQSMCIMVATKSEALTRPVHETLIRSQSPDAVVTDMTFL
 WNSGIAAELGVPCVVFVSMGAFSMLAMHHLEDAGVDRDDQDDDDDDAVVEVPFGFPPIRIPRTELPGF
 LRRPDYSITNLFISLKAANCFGLAMNTSSELEKQYCELYTTPPEEGGGGLRRAYFLGPLALALPPPISS
 SSSSSDCCSIMAWLDSKPSRSVVYVSFGSMAHVKDVLDELALGLETSGISFLWVVRGREEWSPPKGWEA
 RVQDRGFIIRAWAPQISILGHHAAGAFVTQCGWNSVLETVAAPMLTWPLAFEQFITERLVTDLVIGIGV
 RLWPDGAGLRSESYQEHEVIPRQDVARALVEFMRPAAGGPSSIRDMARTKLMDSAKLHAAVAQGGSSHR
DLHRLVDDLLMEAAAKRPRT

>XP_013467412.1 TPX2-like protein [Medicago truncatula]

MEATRVPKANLYPYTTDYPVVLKSHQKPEPKHRTKPEPFQLESLLRHKEEMQREQEER**LRMEREAAQMRK**
 FKALPVKRGPNPSSREGPQTPYTNSRV

>AEK11755.1 S-RNase, partial [Coffea eugenioides]

PLKDQRSINAQDDYWFDFYFLLNPPSANGNWRIEQRFWAQEWSRHGTCSENVFNQQSYFNLAKSLMFTHDL
TSVLFNSRTPPIPLPWPRVSDVMTLISKSPRVRPELRCHYYGNNKILVEVALCYNFTGKQLIDCPRPGSVY
C

>XP_003624729.1 peroxidase family protein [Medicago truncatula]

MKQISIIIFFFILPLAFADLELGFYASSCRKAESIVKQVVQKRFNRDKSITAALLRMHFHDCFVRGCDAS
LLIDSTKNNISEKDTGANDSVRGYDLIDDVKEAIEAACPSTVSCADIVALATRDAVALSGGPKYNIPTGR
RDGLIANRDDVDLPGPNIPIGALSQFFAAKGITTEEMVTLPGAHTVGVAHCGFFASRLSSVRGKPDYTMDD
PALDTKLKLVKLCCKSNSDGAFLDQNTSFTVDNEFYKQILLKRGIMQIDQQLALDKSTSTFVSNFASNGDKF
VKSFATAMIKMGKVGVLVGNENGEIRKNCRVFNKRN

>NP_199445.1 Leucine-rich receptor-like protein kinase family protein
[Arabidopsis thaliana]

MKLLSKTFLILTLTFFFGIALAKQSFEP EIEALKSFKNGISNDPLGVLSDWTIIGSLRHNCNWTGITCDS
TGHVVSVSLEKQLEGVLSPAIANLTYLQVLDLTSNSFTGKIPAEIGKLT ELNQLILYLNYFSGSIPSGI
WELKNIFYLDRNNLLSGDVPEEICKTSSLVLIGFDYNNLTGKIPECLGDLVHLQMFVAAGNHLTGSIPV
SIGTLANLTDLDLGNQLTGKIPRDFGNLLNLQSLVLTENLLEGDIPAEIGNCSSLVQLELYDNQLTGKI
PAELGNLVQLQALRIYKNKLTSSIPSSLFRLTQLTHLGLSENHLVGPIS EEIFGFLESLEVLTLSNNFTG
EFFQSI TNLRNLTVLTVGFNNISGELPADLGLLTNLRNLSAHDNLLTGPIPSSISNCTGLKLLDLSHNQM
TGEIPRGFGRMNLTFISIGNHFTGEIPDDIFNCSNLETLSVADNNLTGTLKPLIGKQLKRLIQVSYNS
LTGPIPREIGNLKDNLILYLSNGFTGRIPREMSNLTLLQGLRMYSDNLEGP IPEEMFDMKLLSVLDLSN
NKFSGQIPALFSKLES LTYLSLQGNKFNGSIPASLKSLSLLNTFDISDNLLTGTIPGELLASLKNMQLYL
NFSNNLLTGTIPKELGKLEMVQEIDLNNLFSGSI PRSLQACKNVFTLDFSQNNLSGHI PDEVFQGMDMI
ISLNLRSNFSGEI PQSFGNMTHLVSLDLSNNLTGEI PESLANLSTLKHKLKLSNLLKGVHPESGVFKN
INASDLMGNTDLCGSKKPLKPTIKQKSSHFSKRTRVILII LGSAAALLLVLLLVLILTCKKKEKKIEN
SSESSLPDLD SALKLKRFEPELEQATDSFN SANIIGSSSLSTVYKGQLEDGTVI AVKVLNLKEFSAESD
KWFYTEAKTLSQLKHRNLVKILGFAWESGKTKALVLPFMENGNLEDTIHGSAAPIGSLLEKIDLCVHIAS
GIDYLSHSGYGFPIVHCDLKPANILLSDRVAHVSDFGTARILGFREDGSTTASTSAFEGTIGYLAPEFAY
MRKVTTKADVFSFGIIMMELMTKQRPTSLNDEDSQDMTLRQLVEKSIGNGRKG MVRVLDMELGDSIVSLK
QEEAIEDFLKLC LCTSSRPEDRPMNEIILTHLMKLRGKANSFREDRNE DREV

>CDX88265.1 BnaC06g37090D [Brassica napus]

MGNFCPNGSSSVGHEEIPVSTKPLPAVKHGQNPVNQKAPPPVVVMPPARGSNPKFPAAETRTVDNTPRQK
QQEKKPRSVDDPSREKQQEKKPRSVETPPSKPVEKQGVGKKA VPPSGKIVTPNLKMFTLTDLMTATKNFR
PESMIGEGGFQVFKGWVDEETLSPSRAGVGPIAVKKSNDPSAQGLHEWQARKFHHPNLVKLLGYCWEE
NQFLLVYEF LPKGSLENHLFSKGDGLTWDTRLKIAIEAAQGLTFLHNSEKSVIYRDFKASNILLDSNFNA
KLSDFGLAKHGPI NGYSHVTTRVMGTQGYAAPEYVATGHLYVLSDVYGFV LLELLTGLRALDPNRPSA
QQNLVEWAKPVLSQKKIKQLMDPRLEHKYPLLA VTKTAALILRCLEADPKNRPPMDDV LRELEIVRTIR
EQPKKEKRNRNNGHGS PHVRKTGRTR

>ABN09074.1 Terpenoid synthase [Medicago truncatula]

MILMMPYGTVEEELFTQAIQRWDFSLIQPLPECMKVVFNTIVELWDEIEITLVETGKSNLVLQYIKQA
FYKLAQSYLVETKWN EGIPTYDEYKANGLISSTIPLSIIISFVGLGEFSNEELLDWLSSEPTIVNAVSA
IGRLADDVSSHKFEQQRVHVAAA VECCMKQYNMSQEEAYKHINKDIEDFWMDINEEFLKLDYIPNPVLEC
ILNVARITEFTYENYEDKYTN GELLKDYVVALLDPISTGR

>XP_007508107.1 APG4C_XENLA Cysteine protease APG4C (ISS) [Bathycoccus prasinus]

MALFSSSEDDASENR TNPLWKVLQKSFVSYKKLQFSKRMAMKNSIFSRRIHDESNRQVVRNTVVMILGERM
TAFDRLISRVDAFFFFSSYVSEFPPIESKMFLETTNNTNTTTNTKETTTMKMKSIDFVANQVGGYTTDCGW
GCTLRSAQMLFGEALMRSARARRFRRLSLRRSEEDVSEEEEEEGEEFKRKKEIVDMFSDNDNDDENEDKKE
KKKKKTKNVFGLRRVYEGDEDENALCPGQWMPSEICKRYGKMMNRLDSFQNVRCILIGDGC GGVP EFY
PERVREEMKTHADKDVLLVPLRCGASDAINPEYVKS LQKFLSVRECVGIVGGKKTASYIYVGF TSGKKS
SDSYSGGEKEEEEEKEEEENEDEEEEEEEEEETRAIYLDPHVAKAYVSPRERSRDESTESAYYRSFFG
SASEHGILYTPFHALDPSLVVGFVGN DTYDEMNNASSSSSLDAFVDVLTNIERESGSTPLITVVTKKVL
SKKKMNSSMNSSSSSKRTENNKRGEEDEDDWEIIDGCEGEEF

>KMZ74171.1 60 kDa chaperonin 2 [Zostera marina]

MASANAISTAIPRSPRQPEVQRRIQAQRRVSYRPSMNGKSRFVVRSSAKEIAFDQISRSTMQNGIDKL
ADAVGVTLGPRGRNVVLEDFGNPKVVNDGVTIARAIELANPMENAGAALIREVASKTND SAGDGT TASV
LAREIIRLGLLNVTSGANPVSIKKGIDKTVIELINVLEGKSLPVKGRNDIKAIATISAGNDDQIGTMIAD
AIDKVGPDGVL SIESSSFFETTVDVEEGMEIDRGYISPQFVTNPEKLIVEFENARVLITDQKISAIKEII
PLLEKTSQLRAPLLIIAEDITGEALATLVVNKLRGIINVAAIKAPGFGERRK ALLQDIAIMTGSEFQAKD
LGLLIEDASVESLGIARKITISQSSTTIADAASKDEIQARISQIKNELAETDSVYDSEKLAERIAKLSG
GVAVIKVGAATETELED RKLRIEDAKNATFAAIEEGIVPGGAAVHLS TCVPAIKELFEDADERLGADI
IQKALVSPASLIAANAGVEGEVVVEKVRDSEWEIGYNAMTDKYENLVESGVIDPAKVTRCALQNAASVAG
MVLTTQAIVVEKPKPKVPVAAPPQGLTV

>NP_178341.1 phloem protein 2-B10 [Arabidopsis thaliana]

MGRKRRVKSESSPFDSPEDCISYIISFTNPRDACVAATVSKTFESTVKSDIIWEKFLPADYESLIPPSR
VFSSKKELYFSLCNDPVLFD DDKKSVWLEKASGKRCLMLSAMNLSIIWGDNPQYQWQWIPI PESRFEKVAK
LRDVCWFEIRGR TNTRVLSRTRYSAYIVFKGVDKCYGFQNV AIEAAVGVVQEPSRRLICFSEAIRRGR
RNVVKPKQREDGWMEIELGEFFNDGGIMDNDEIEMSALETKQLNRK CGLIIQGIEIRPAKIL

>AAC63569.1 alcohol dehydrogenase, partial [Gossypium robinsonii]

LSVEEIEVAPPQKNEVRVKIHF TSLCHTDVYFWEAKGQNPLFPRI LGHEAGGVVESVGEVMDVQPGDHV
LPIFTGECKECPHCLSEASN MCDLLRINTERGGMLHDGQTRFSKDGKPIYHFLGTSTFSEYTVVHVQIA
KINPEAPLDPKVCVLS CGMSTGFATVNVAKPKKGGSVAVFGLGAVGLAAAEGAR VCEASRIIGIDLNPNR
FEEAKKFGCTEFVNP KDHNKPVQEVIAEMTGGGVDCSIECTGSTQAMVSAFECVHD

>XP_002950408.1 ferroportin [Volvox carteri f. nagariensis]

MEPQVPVRAKVF LCASYALAAWARSWEFIVALVLI ELYPDSLLMVSAYGLLDN MARVLLGPAVGSYVDR
HERMPGAQAMLR LQNL CIGGSAALVLLWPRSAATEHKAVYWSLMWLLTVLGSASSAGSTGVSISVERE
AVKTL CGADGRAL AALNSVMRAIDL TALLCAPLAAGLLMTAAGPFTAVAAMAAYCGVAYVPEVLLLGA AF
RAAPVLGQPKVRAASDS AVSGGQAAGLGGADADTEERTGLLAGCADCAGEQAISR RDPHTSGAGGKAGPA
VTVAATGTDPEHLRSCFEDSTGGC ISSGGTRGEPQGGDRGSRR AEGSDDDDGDFGDRADSARLNQPLLL
DAAESLPTSVSATWEGDDRLVTGDRSAGRSQHAI SVKVGVMVAAGGGGELSNTSSVTALEMTSPTAA
TSCGGGLARLRVLLRAVRSQFGLYCD SWRVYLQQSVLLLCVALALLYMTVLSLGLFMTSFLKWSGLSEA
EVSGYRGIGALTGLAATAIFPPLSARAGLLFCAVAGVTYQLACLAAGVLPVVTPTMAGGDGQRPSVPQVR
ILVAGLVSSRTGLWLYDLAVTQLIQEEVRQDQLGSVYGVQSS LQASFEMMSFVAGLIGGCFFATMMSFR
VSTLLILAVATARCLAQY EYDYGYDYFDTEDDFYDEYGYGDEDESYYEYGYG YADDFAGYEDDYAE

EYDFFECCFFDTNGQPVLANGTASCDIK **IAGVVKQADKLTGGLDGIYK**LTSCYNGKPMYKRQSSPPGEDR
 VLWYSSTFGDWDVSKGSEPNAAEILMYGGEMEHASVPLFVGSWHLGGDLKSDTALGEDDYLPINNVNSCA
 DGTVYTAGLNSYKQVVGVPVLTDAEIEAKYSYIYEKYSKSDPSPTINFVLLVMSGLTIVLAIPIYFLLR
 KKGAKGGSISVSFAQMLTQSKKKSSGHIN

>ADM18306.1 ubiquitin-like protein 1, partial [Gladiolus grandiflorus]
 NDSAPNPFAALLGNQGTASREPLTNSSDPASGSPAPNTNPLPNPWGNTTAGAAQATNTRATPTTGTRAP
 GIAGLGGGLGPELERMVGQVDPALINQVMQNPAMTQMMQNLLSDPQYMNQMLSMS PQMRTMMENTQFRE
 MLQNPEFIRQLTSPDPTMQQLLSFQQSVISQLGRQQPSQEQNQTTGGGSGLES LMNMFGGLGAGGLGVPNPS
 KVPPEELYTTQLSQLREMGFFDTQENIR **ALSATSGNVHAAVER**LLGNLG

>ABU88984.1 phospholipid/glycerol acyltransferase [Helianthus annuus]
 MSKLTTSRSELDLDRPNIEDYLPDPSIQQPHTKLRLLRDLDISPTL TEAAGAI VDDSFTRCFKSNPPEPW
 NWNIIYLFPLWCLGVVVRYGILFPGRVLILTIGWIIIFLSCYIPVHVLLKGHDKLRKRLERALVELICSFV
 ASWTGVVKYHGRPCARPKQVFNHTSMIDFIVLEQMTAFVIMQKHPGWVGLLQSTILESVGCIFWNR
 SEAKDREIVARKLREHVEGTDNNPLLI FPEGTCVNNNYTVMFKK **GAFELGATVCPIAIK**YNKI FVDAFWN
 SRKQSFTTHLLQLMTSWAVVCDVWYLEPQNMKPGETAIEFAERVRSIISIRAGLKMVPWDGYLKYSRSP
 KHRESKQQSFAESVLRRLLEEK

>KYP74384.1 F-box protein PP2-B15 [Cajanus cajan]
 MESKTQQQLVSFDDDECPTKIEALAEDCVSKILSYTSPADACRF SMLSSSLRSAADSDLLWRSFLPSDHTH
 ILSRALSPLTLNSSSFKHLFHALSHPLLLDAGNISFKLDKSSAQKSYILSARELSITWSNDPLYWCWRPA
 PESRFKEVVLRVTVSWLEIQGKIETKILTQNTSYGAYLIMKISHREYGLDSVVCVSVQLGNKVQSGRVY
 LCKKLDEKKDEGIGIPRR **REDGWMEVEVGEFLS GE GEEVVK**ISLMEVGYRLKGGGLIVEGLEVRPKHHS

>XP_003630752.2 casein kinase I-like protein [Medicago truncatula]
MERIVGGSYKLGRKIGSGSFGEIYLATHIDT FEIVAVKIENSKTKHPQLLYEAKLYNILQGGSGIPGLKW
 FGIDREDNVLAIDLGLPSLEDL FVYCGRKFSLKTVLMLADQMITRIEYVHSGYLRHDIKPDNFMGLGR
 KANQVYIIDFGLAKRYRESTTNRHIPYRENKSLTGTARYASCNTHLGIEQSRDDLES LGYVLLYFLRGS
 LPWQGLRAATKKQKYEKICQKKVSTPIEVLCKSHPEFASYFHYCHSLTFDQRPDYGFLKRLFRDLFARE
 GHEFDFVFDWTILKYQQSQKNLVQPSLSPVPGASNRHAVAMDVDNHQGHGEERIRSGNATGSGVKVQFKS
 PAGRNLSNDNPRDKNIFGEANMASASYS PAGTSKRNPLN PALS AEP SNTGHGQASKIGPSSSWMSSLQRI
 SSSK

>XP_011402186.1 Ferredoxin-dependent glutamate synthase, chloroplastic
 [Auxenochlorella protothecoides]
 MPSPLOAVPVLGPGPAPDTEPYVGVAVDNE DLLREKDACGVGF IANLKAERTHDIVTKALMALGCMEHRG
 ACSSDNVSGDGAGLMLEIPWDL LAEEVPGTAQGSTGVGMVFLPDDEAAAGTARGIVA AVAEAE GFDLLGW
 RDVPVDVA AVGEVARATMPRIAQVVLRSRAGLAGDDLERELFIVRKLIERRSERELGPELAPGFYFCSLS
 CRTIVYKGM LNSGAVARFFRDLTDP RYVTCFAVYHRRYSTNTTPRWPLAQPMRLLGHNGEINTLQGNLW
 VASRQ GALRAEAWGGRERELLPLCDARES DSANLDHLAELLVRS GTDAREALMVLVPEAYRNHPDLAAEY
 AAVADFYRYEGLQEGWDGPALLVFTDGKKGARLDRNGLRPARFWRTADGVVYVASEVGLGDFVSSAP
 NIVAKGRLGPGQMLCADLETGEFQGHVDIARDVASRHPYADWLAAC THRLADIVPPRTVLEGPAMAAGDA
 LRLMAASGYGLEDTAMVIEGMAGSGAEPTYCMGDDVPLPALARS PHLLYDYFKQRF AQVTNPPIDPLREG
 LVMSLEMRLGRRGNLLSPGPDSYQQTSLVVE TAGCVSTHHLAALVGYGAHAVVPYLA FEAAARQWRAS PRT
 ENLVRAGKLPDVSMAAAQRNYKKALEKGLMKI LSKMGISLLSCYHGAQIF EAYGLGREVIDAA FVGTVSR

IGGLTLDDVQRETEAFWAKGFPEKKLTKLEDYGYIQARPKGEHHANNQAMSKLLHKAIGLGGHGPAGAGA
 YEAYLQHFRAPVTVLRDLELHTGAAAVDVEEVESVADIVARFCTGGMSLGAISRETHETIAIAMNRLG
 GKSNSGEGGEDPVRWLELDGVDKGGRSALLPHLKGLRAGDTATSRIKQVASGRFGVTPEFLVNADQGGQL
 PGKKVSPYIADLRRSKPGVPLISPPPHHDIYSIEDLAQLIYDLHQVNPAARVSVKLVAQAGIGTVASGVA
 KAGSDVIQVSGHDGGTGASPISSIKHAGGPMEMGLAEVHQTLTANGLRDRVVLRADGGMRTGRDVLVAAA
 LGGDEFGFGTVAMIATGCIMARVCHTNCPVGVASQREELRKRFPGTPEDLVNYFMFVAEEVRAGLASLG
 LRSLELVLVGRGDYLRQRTDVSALAKTSALDLSILTR**FAGETATSTER**RGAPPHDNGSEWDDVILADADEVQK
 AIDEEGTVRRSYDIANTDRSALGRLGGAIKRYGDDRFQGRIELDLRGSAGQSFACFITRGVHVTTLTGEA
 NDYVCKGMAGGEVAIRPPEESRFPAREASIVGNTCLYGATGGALFVNGRAGERFAVRNSRAAAVVEGVGD
 HACEYMTGGVVVVLGKSGRNVAAGMTGGLGYFYDPQGTFPDRVNAEIVTVQRVLTGTGAAQLKGLLERHV
 ELTGSAAGQELLGDWEAALPRFWQLRFQHLEIGAHDTDLLASPGPTGPVQYPLRAPSFQPLPGPFQGTRV

>AIZ05147.1 DNA mismatch repair family protein, partial [*Salix arbutifolia*]
 DSSQLCPPLEESKMFSKQLSANGNDSEETQTAEDSSPLMTVEVKSKTFQAGERSIHDIEEK**FMMKDFTL**
RLHDTKKTDSL TNSNSCKATKHLNAETDRNARCLSRGVERVKGDSNGPSGYFQSKLSNFVTVNKRKREDI
 TPQLSEAPVLRNQ TSECQLKISDINMHDAVTS LPFN

>KHF97672.1 NADH-quinone oxidoreductase subunit D [*Gossypium arboreum*]
 MLYRRVTRFVSSLVSDK**SELHIILSNFGLI**RKMALQIADFCPHGRVSYSCMTHGHVARSSVPWCGNEIEV
 NMLHTPHTQECDLAVLHKS SVYPTILPQSDTRACVAISKGTRAKHTGMWLAM

>AEX01583.1 photosystem II CP43, partial (plastid) [*Zostera angustifolia*]
 MKTLYSXRRXYPVXTLFNGXLAXAXRDQETTXFAWGXGXARXINLSGKLLGAHVAHAGLIVFWAGAMNLF
 EVAHFVPEKPMYEQGLILLPHLATLWGVGPGGEVIDTFPYFVSGVLHLISSAVLGFGGIYHALLGPETL
 EESFPFFGYVWVKDRNKMTTILGIHLILLGIGAFLLVFKALYFGGVYDTWAPGGGDVRKITNLTLSPSVIF
 GYLLKSPFGGEGWIVSVDDLEDIIGGHVWLGSICIFGGIWHILTKPFAWARRALVWSGEAYLSYXLXALA
 VFGFIACCFVWFNXTAYXSEFYGPTGPEASQAQAF TXXXRDXR**LGANIGSXQGPTGLGK**YLMRSPTGEVI
 FGGEXMRWFWDLRAPWLEPLRGPNGLDLSRLKKDIQPWQERRSAEYMTHAPLGSLSNSVGGV

>AKT94711.1 maturase K, partial (chloroplast) [*Brunonia australis*]
 MEEFQNHLELDRSQQHYFLYPLIFQEYIYVLTHNDGFNMLENAGYENKSRFLIVKR LISQMYQKNYLHMI
 LSSNASKQSLFWGYKRCYSQVMSKGF SIMEIPLFMRLISSLERKALVKSENLR SIQSIFSFLEDNFSY
 FNYVLDLLIPYPAHLQILVQALRYWIKDASSLHLVKLFLEYSNFNLSLIPSNKSKAPSSFLNQRLFFFLY
 NFHVCEYESGLIFLRKQSSYLRSTSFGTLLERIFFYKIDQLADVFRSFQATLWLFKEPFMHYVRYQ GK
SIFASKGTFLLINKWKYFYFIKFWQFYFVLSQPGTIYINQLSNHSLDLLGYCSIVRLKPLMVHSQM QENT
 FLIDNRINKFETLVPILPLIDLLAKLKVCAVGHPIGKASWADLSDSEIIDRFGRLYRNLSHFHSGCSKR
 NSLSRVKYILQLSCLRTLARKHKSTVRAFLKRFGESEFLEEFFGEEVLSLTFPRVSSISRKLSRKRIWYL
 DVXNDXANHD

>BAF08402.2 Os02g0261400 [*Oryza sativa Japonica Group*]
MRQFGPDFGGVTKKSFLHLK KIELVGLPELVEWGGDHCMFSKLLSIRCEDCPNLTVLLLPSFECSISD
 TKDINTIWFNLC SLKIRNCPRLSLPPLPHTSMLT CVTVKEDD TDLMYFDGKSLRLNRYGSALAFHNLNK
 VEDMEIVDMPLVSWTGLQKLN SPRSMQSMGLLSNLSSLTHLELVNCDNLRVDGFDPLTTCNLKEMAVYNS
 KNHHP SIAADLFSVAMMEV I PAGSFQQLEQLSVD S ISAVLVAPICNLLASTLCKMEFPYDMWME SF TET
 QEEALQLLTS LQCLGFYVCPRLQSLPEGLHRLSSLRELI IHKCPEIRALPKEGFPASLRVVFAYEGISVD
 LKDQLK LKASTPGLRDRCSRSIPLFIEISKGKTARGCSIVYRINSRK

>AEF32085.1 ent-kaurenoic acid oxidase [Castanea mollissima]
 MELGPICNVLICIFGVLVVLKVVKNANWWLYETQLGKQYSLPPGDLGWPFIGNMWSFLSAFKSKDPDS
 FVSSFVSRFGRTGIYK**VFMFGNPSVIVTSPEACR**RVLSDDDCFKPGWPKSTVALIGKKSFIGISFEEHKR
 LRRRLTAAPVNGYEALSMYTKYIEEIVTSSLDKWTMGEIEFLTELKLTFRIMYIFLSSESEPVMTALE
 REYTALNYGVRAMAINIPGFAYHKALKARKNLVATFQSIVDDRRNQKKVSVPTKKKDMMDALIDVKDENG
 RKLTDDEEIIDILVMYLNAGHESSGHTAMWATIFLNQHPEYLQKAKKEQEEIVRKRLPEQKGLTLKEIREM
 EYLSKVIDETLRVITFSLTVFREAKQDVNIAGYTIKPGWRVLVWFRSVHLDPEIYENPKFKPSRWDNFT
 PKAGAFLPFGAGTRLCPGNDLAKLEISIFLHHFLLNYQLERLNPGPSMYLPHTRPVDNCLARIKKCPSS
 SL

>XP_007032041.1 Aminopeptidase isoform 1 [Theobroma cacao]
 MGKKKKDI IKLERESVPIVKQKLITTLADLIEGKSERAEFLKLCQRYEYTI RAWYHLQFDELMQLFALF
 EPVHGAQKLEQQNLSPEEIDDLEQNFLSYLFEMMDKSNFKITDDEEIDVALSAEYRLNLPVIVNESKLDK
 RLFTRYFEKNPHKDLPLYFADKFIIFRRFGIDHMTAYFFKAKVNTIIIVRFWR**CFTKVTGLKFLTR**LFFGK
 KVKPKYKDSAEPTTELKIEGEQNELYVERIRIEKLFKFSISNFLRKITIQEPTFERIIVVYRRMPRKHETQR
 NIYVKHFKNIPMADLEIVLPEKKNPGLTPMDVWKFLVSAAGLVTVISSLSMPKPDIRVIFGILSAVVG
 CVKTYFSFQQNLVDYQNLITRCVYDKQLDSGRGTLHLHCDEVIQQEVKEVIVSFFVLMEQKGVSKGDLD
 LLCEQLIIIEKFGDCNFDVDDAIRKLEKLGIVSQDNIGTYTCVNLRRANEIIGTTTEEVLKAKGGGGDP

>ABV72578.1 dehydration-responsive family protein S51 [Brassica rapa]
 MNLSTRISRDKKSNNLYYITLVAVLCIGSYLLGVWQNTTVNPRAAFDTSTDAPPCEKFSKTTSTTDLDFN
 AHHNPHDPPPSAVTAVSFPSDAALSEHTPCEDAKRSLKFSRERLEYRQRHCPDREEALKCRIPAPYGYK
 TPFWRPESRDVAWFANVPHTELTVEKKNQNWRYENDRFWFPGGGTMFPRGADAYIDDIGRLIDLSDGSI
 RTAIDTSCGVASFGAYLLSRNITMSFAPRDTHEAQVQFALERGVPAMIGIMATIRLPYPSRAFDLAHCS
 RCLIPWGKNDGVYLMEVDRVLRPGGYWILSGPPINWQKRKKGWERTMDDLNEEQTQIEQVARSLCWKVV
 QRDDLAIWQKPFNHICHCKMRQVLKNPEFCRYDQDPDMAWYTKMDSCLTPLPEVDESEDLKTVAGGKVEK
 WPARLNAVPPRVNNGDLKEITPEAFLEDTELWKQVRSYYKKLDYQLGETGRYRNLLDMNAYLGGFAAALA
 DEPVWVMNVVPEAKHNTLGVIYERGLIGTYQNWCEAMSTYPRTYDFIHADSVFTLYQDKCEPEDILLEM
 DRVLRPGGGVIIRDDVDVLIKVK**ELSKGFQWQGRADHEK**GIPHERVKIYYAVKQYWTVPAPPEEDKNNTKA
 LS

>XP_003591835.1 DUF4283 domain protein [Medicago truncatula]
 MAATNPTTSGKSFAQALSGAVSCDSFLSKLPPKIVMGDSVRVKISQAVYESGLAACQCNLHGRLTLHKGD
 SPLTTQALKVKLNLCITIEEMKQVWAIGVVNLKPGFMHFSCWTKDFISKAQAQTHAQVWVRLMNLPHHEYW
 GKQTI FEIASGLGSPLTIDDATVNRREFGLFARVLVDVLANKMFDVIVEREHNLGHSIHTCSKMNHET
 PTNVTRKAHNVNTKAQSKASLNSKKDEETIVHKIAPSSNEAVGKKPAKTQIPSDNMEPSNSITAVEGGE
 TLCDNVTKELVKTMHVMSNQAEIHTTDKNLDDSPVFEVTESAKQASKESVCLEKVTSSDPCTLEKDSSQ
 LVVLPNTQVLES DHGRHPMPDPYVDTSLYHFPLVTPITTTDEVLPDKGKVIITGSSK**NFTAASMKS**
ILRKFWGDEEDTDPATDSTMDPDTNTEKLDTCETS IAGKYL VQHPHNSFVKPGRKASKRQVNHNGKTT
 DGVTSSSEHIQTRSKKGVIKSNPKYV