UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCAS BIOLÓGICAS

DEFENSINAS VEGETAIS: ROTINA DE IDENTIFICAÇÃO E ANÁLISE *IN SILICO*

LUÍS CARLOS BELARMINO DA SILVA

RECIFE
2010

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCAS BIOLÓGICAS

DEFENSINAS VEGETAIS: ROTINA DE IDENTIFICAÇÃO E ANÁLISE IN SILICO

LUÍS CARLOS BELARMINO DA SILVA

Dissertação apresentada ao programa de Pós-Graduação em Ciências Biológicas, Área de concentração em Biologia Molecular e Celular.

Orientadora: Dra. Ana Maria Benko-Iseppon

RECIFE
2010

Silva, Luís Carlos Belarmino da
Defensinas vegetais : rotina de identificação e análise in silico / Luís Carlos Belarmino da Silva. – Recife: O Autor, 2010.

167 folhas : il., fig., tab.
Orientadora: Ana Maria Benko-Iseppon.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCB. Ciências Biológicas, 2010.
Inclui bibliografia, anexos e apêndices.

1. Botânica 2. Genética vegetal 3. Melhoramento genético em plantas 4. Bioinformática 5. I. Título.

581.38 CDD (22.ed.) UFPE/CCB-2010-122

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
Programa de Pós-Graduação em Ciências Biológicas

PARECER DA BANCA EXAMINADORA DA DEFESA DE DISSERTAÇÃO DE MESTRADO DE
Luís Carlos Belarmino da Silva

DEFENSINAS VEGETAIS: ROTINA DE IDENTIFICAÇÃO E ANÁLISE *IN SILICO*

ÁREA DE CONCENTRAÇÃO: **BIOTECNOLOGIA**

A banca examinadora, composta pelos professores abaixo, sob a presidência do primeiro, considera o aluno **Luís Carlos Belarmino da Silva** como **aprovado.**

Recife, 22 de Janeiro de 2010.

Orientadora: Prof. Dra. Ana Maria Benko Iseppon
Departamento de Genética (UFPE)

Titular: Prof. Dr. Ederson Akio Kido
Departamento de Genética (UFPE)

Titular: Prof. Dr. Roberto D. Lins
Departamento de Bioquímica (UFPE)

A minha mãe, Terezinha Soares da Silva,
ao meu pai, José Belarmino da Silva e
a Daniel S. Barbosa

AGRADECIMENTOS

Agradeço ao professor Sergio Crovella por nos haver apresentado as defensinas e por sua presteza em ajudar na organização do material para publicação.

Sou imensamente grato à coordenação do programa de pós-graduação em Ciências Biológicas, que sempre me recebeu com entusiasmo e atenção, viabilizando a realização deste mestrado. A Adenilda, por organizar e solucionar todas as necessidades para a conclusão das atividades. Ao MCT/RENORBIO e a FACEPE agradeço pelo auxilio financeiro, através da concessão de bolsa, sem a qual não poderia realizar este trabalho.

O maior agradecimento é para minha mãe, que em toda a minha vida me incentivou nos estudos e trabalho, mostrando-me como deve ser alguém com caráter e ensinando-me a reconhecer o amor divino onde quer que eu vá. Agradeço também ao meu pai e meus irmãos, que sempre estiveram do meu lado, nos bons e maus momentos, incentivando constantemente meus trabalhos e estudos. Sinto-me particularmente grato a Daniel Sandro Barbosa, pelo companheirismo, dedicação, alegria e principalmente pelas mensagens de amor.

Meus agradecimentos são extensivos aos professores Laurent E. Dardene e Maria Tereza dos Santos Correia, pelo apoio no início do curso. Da mesma forma a Priscila Capriles, pela contribuição na elaboração deste trabalho; mesmo a distância suas ideias e construtivas discussões online vieram colaborar para o desenvolvimento de meu trabalho.

Agradeço ainda aos colegas de trabalho do Laboratório de Genética e Biotecnologia Vegetal, especialmente a Nina Mota, pela experiência compartilhada; com quem muito aprendi e vivenciei situações ímpares. A Alberto Vinicius Onofre e Lidiane Amorim, meus mais sinceros e nobres agradecimentos, pois desde os primeiros momentos mostraram-se voltado ao incentivo e ajuda, não somente na realização do curso e do trabalho ora apresentado como também dividindo seus conhecimentos, agregando valor científico e didático às disciplinas que frequentei. Do mesmo modo, agradeço à grande profissional e amiga, a minha orientadora, Ana Maria Benko-Iseppon, com quem dividi ideias e conhecimentos sobre bioinformática.

Luís Carlos Belarmino

RESUMO

A compreensão dos vários mecanismos de defesa vegetal tornou-se uma área central nas pesquisas mundiais. Em plantas, uma primeira linha de defesa contra invasores inclui uma variedade de peptídeos antimicrobianos, tais como defensinas - uma classe de pequenos peptídeos básicos, ricos em cisteínas, distribuídos por todos os reinos. A crescente disponibilidade de genomas completos recentemente sequenciados, grandes quantidades de sequências expressas (ESTs – *Expressed Sequence Tags*) e o desenvolvimento de tecnologias computacionais ofereceram vários recursos e algoritmos que possibilitaram a aplicação de abordagens baseadas em bioinformática para identificar potenciais peptídeos antimicrobianos. Assim, o presente trabalho visou ao desenvolvimento de uma rotina computacional para identificar e caracterizar prováveis defensinas nos genomas vegetais atualmente sequenciados. Tal rotina foi desenvolvida com base em programas de computadores gratuitos disponíveis pela internet, envolvendo a integração de dados provindos de sequências, reconhecimento de padrões e motivos conservados em proteínas, perfil diferencial de expressão, análise evolutiva e modelagem comparativa. Um exemplo da aplicabilidade dessa rotina foi demonstrado usando o genoma expresso da cana-de-açúcar, evidenciando a presença de 17 sequências de prováveis defensinas, evolutivamente relacionadas com peptídeos com atividade antifúngica descritos em outras espécies. Em cana-de-açúcar parecem existir seis grupos de defensinas envolvidas na defesa do organismo, cujos membros, apesar de bastante similares, apresentam um padrão de expressão tecidual específico. A resolução da estrutura tridimensional através de modelagem comparativa mostrou peptídeos globulares anfifilícos altamente compactos, estabilizados por quatro pontes de cisteínas. A rotina estabelecida se mostrou eficaz e potencialmente aplicável tanto às defensinas como a qualquer classe de peptídeos antimicrobianos. O presente trabalho incluiu também revisões sobre o estado de arte atual das pesquisas com defensinas vegetais, sua identificação e os principais bancos de dados que as compõem, revelando que a eficiência da estratégia utilizada em cana-de-açúcar, especialmente se integrada a dados oriundos de diversas fontes. Informações preliminares como as presentemente apresentadas são fundamentais para a escolha de moléculas com potencial antimicrobiano para o desenvolvimento de produtos farmacológicos.

Palavras-chave: desenvolvimento de fármaco, perfil diferencial de expressão, análise filogenética, estresse biótico, estresse abiótico, melhoramento vegetal.

ABSTRACT

Sugarcane is one of the most important tropical and subtropical crops. The culture has faced many losses due to pathogen attacks. Understanding the various mechanisms of plant defense has become a central area in worldwide research. In plants, a first line of defense against invaders includes a variety of antimicrobial peptides such as defensins - a class of small, basic, cystein-rich peptides, distributed throughout the kingdoms. The recent increasing availability of whole sequenced genomes, large amounts of expressed sequence tags (ESTs) and the computational development have offered several resources and algorithms that made possible the application of bioinformatics-based approaches to identify potential new antimicrobial peptides. Thus the present work was aimed at developing a computational routine to identify and characterize putative defensins in actually sequenced plant genomes. This routine was developed based on computer programs freely available on the internet, involving the integration of data coming from sequences, pattern recognition and conserved motifs in proteins, differential expression profile, evolutionary analysis and comparative modeling. An example of the applicability of this routine was demonstrated using the expressed genome of sugarcane, indicating the presence of 17 defensin-likesequences, evolutionarily related to peptides with antifungal activity described in other species. Sugarcane seems to have six groups of defensins involved in organism defense, whose members, although very similar, show a pattern of tissue-specific expression. The resolution of the three-dimensional structure by comparative modeling showed highly compact globular amphiphilic peptides, stabilized by four cysteine bridges. The routine established was effective and potentially applicable to both defensins as any class of antimicrobial peptides. The present work included also reviews regarding the actual state-of-art of defensin research in plants, its identification and main data-banks including such molecules revealing the applicability of the strategy used in sugarcane, especially if integrated with data from other sources. The here presented preliminary information are essential for the selection of molecules with antimicrobial potential for the development of pharmacological products.

Keyword: Drug discovery, differential expression profile, phylogenetic analysis, biotic stress, abiotic stress, plant breeding.

LISTA DE FIGURAS

Revisão bibliográfica

Figura 1: **Representação esquemática da história evolutiva proposta para as defensinas**. A árvore genealógica com a presença do supergrupo Opistoconta, incluindo fungos e animais, é usada para indicar o surgimento dos tipos diferentes de defensinas durante a evolução. Vários eventos importantes como duplicação gênica, ganho e perda gênica em linhagens específicas e surgimento da atividade antifúngica são indicados em diferentes pontos no tempo. **Fonte**: Zhu, 2008.

Figura 2: Alinhamento das sequencias de aminoácidos das defensinas NaD1 de *Nicotiana alata*; PhD1, PhD2 e PPT de *Petunia hybrida*; FST de *N. tabacum*, TPP3 e TGAS118 de *Solanum lycopersicon*; RS-AFP2 de *Raphanus sativus*; alfAFP de *Medicago sativa* e γ1-P de *Triticum aestivum*. O peptídeo sinal foi omitido. Resíduos idênticos foram destacados em preto, enquanto substituições conservativas foram destacadas em cinza. A seta indica o sitio de clivagem do pró-dominio C-terminal. Linhas sólidas representam o padrão de conectividade dissulfídica. Uma ponte dissulfídica adicional em PhD1 e PhD2 é mostrada como linha pontilhada. **Fonte:** modificada de Lay *et al*. (2003).

Figura 3: Duas classes de defensinas vegetais. Todas as defensinas vegetais são produzidas com um peptideos sinal de endereçamento ao retículo endoplasmático. (A) A maioria das defensinas vegetais possui apenas o domínio maduro além do peptídeo sinal. (B) Algumas defensinas vegetais isoladas apresentam um pró-domínio adicional na extremidade C-terminal. **Fonte**: Lay e Anderson (2005).

Figura 4: Estrutura 3D de defensinas vegetais como tipificada por Rs-AFP1. (A) Vv-AMP1 obtido por modelagem comparativa. A α-hélice e a β-folha estão respectivamente representadas em vermelho e marrom. Varetas amarelas representam as pontes dissulfídicas. (B) Representação globular, evidenciando a distribuição polarizada das cargas na superfície do peptídeo. **Fonte:** Beer e Vivier (2008); Zhu (2007).

Figura 5: Representação esquemática da estrutura gênica de defensinas vegetais conforme observado em *Vitis vinifera*. (A) Sequência codificante de Vv-AMP1 com seus respectivos aminoácidos deduzidos. (B) Sequência genômica. Os blocos em amarelo representam a sequência codificante do peptídeo sinal, enquanto o bloco vermelho a sequência codificante do peptídeo maduro. O bloco em cinza indica a posição do íntron. Números correspondem ao tamanho em bp em cada seção. **Fonte:** Beer e Vivier (2008).

Figura 6: Organização genômica de DEFL relacionadas em *Arabidopsis thaliana*. Cada ponto representa um gene de DEFL, sendo os genes agrupados em espaço de 100.000 bp representados em pilhas verticais. Os maiores clusters estão enumerados, sendo as sequências coloridas para refletir relação de parentesco entre os subgrupos. **Fonte:** Silverstein *et al*. (2005).

Figura 7: Modelo proposto para o modo de ação juntando evidências de fontes variadas. (A) O alvo molecular de peptídeos antimicrobianos de organismos multicelulares nas membranas celulares e a base de sua especificidade. (B) Modelo de Shai-Matsuzaki-Hunag do mecanismo de ação de um peptídeo antimicrobiano. **a**, recobrimento do folheto externo da membrana formando um tapete de peptídeos. **b**, integração do peptídeo na membrana e afinamento do folheto externo. A área de superfície do folheto externo da membrana se expande em relação ao folheto interno, resultando em uma tensão dentro da bicamada lipídica (setas denteadas). **c**, transição de fase e formação poros transientes se formam nessa fase. **d**, transporte de lipídeos e peptídeos para o folheto interno. **e**, difusão de peptídeos que encontram alvos intracelulares. **f**, colapso da membrana em fragmentos e rompimento físico da membrana-alvo. Lipídeos acídicos ou negativamente carregados estão representados com extremidade amarela, enquanto lipídeos sem carga geral estão representados com extremidade em preto. **Fonte:** Zasloff (2002).

Manuscrito 1 – Padovan *et al.*, 2010

Figure 1. Nucleotide sequence (FJ94789 GenBank) with predicted genes/exons and putative amino acid sequences of PDEF_VIGUN. The figure was prepared using GENSCAN.

Figure 2. Vigna unguiculata defensin DNA sequence compared with that of the other plant EST obtained from databank (for a more detailed comparison with more plant species see attached materials figure X) Aligned nucleotide sequences include *Vigna unguiculata* PDEF_VIGUN, *Gycine max* protease inhibitor (**U12150 GenBank) and *Prunus persica* defensin** (**AY078426 GenBank**). *Vigna/Glycine*: score 91; *Vigna/Prunus* score 77.
* identifies conserved residues.

Figure 3. Plant defensin gene structure after heterologous comparison. The intron boundaries are conserved in all compared members from Fabaceae, with the exception of *A. thaliana*. Black boxes indicate exon while grey boxes the introns. The nucleotide length of the DNA stretch is reported below the boxes. CDS: translated coding sequence; Arath: *Arabidopsis thaliana*; Glyma: *Glycine max*; Medtr: *Medicago truncatula*; Lotja: *Lotus japonicus*; Vigun: *Vigna unguiculata*. Phavu: *Phaseolus vulgaris* (TC147). PTR: Protein added to the alignment; Vigra: *Vigna radiata* (CAA34760), Vigun2: *Vigna unguiculata* (P18646); Vigra2: *Vigna radiata* (BAB82453). Disulphuric bonds are highlighted: Cys3-Cys47; Cys14-Cys43; Cys20-Cys41; Cys24-Cys34).

Figure 4. Amino acid sequence and putative structure analysis A) Amino acid sequence alignment of PDEF_VIGUN with closely related sequences from other plants. Plants are identified by common names, but the UniProt accession codes are provided. The degree of identity with PDEF_VIGUN and net charge are also shown. Highly conserved residues are shaded grey, strictly conserved residues are shaded black. The presence of conserved positively charged positions, negatively charged positions and hydrophobic residues positions are indicated by +, -, and Θ respectively. The sequence for bell pepper (*C. annuum*) defensin, the closest peptide for which a structure (1GPT) has been determined, is shown separately below. Buried (b) and exposed (e) residues resulted from visual analysis of each residue in the barley structure, as did the topological diagram (**B**). **C)** Schematic cartoon of the putative PDEF_VIGUN structure, based on that of bell pepper defensin (1GPT). The putative position of highly conserved residues in this schematic representation is approximately indicated. The dotted lines indicate areas where cationic residues would be concentrated.

Figure 5. A) Plant defensin expression in *Vigna unguiculata* after wounding and CPSMV inoculation. BRIT63 is a unique CPSMV (cowpea severe mosaic virus) resistant genotype derived from a breeding cross (BR14-Mulato x IT85F-2687), while IT85F is the CPSMV-susceptible parental cultivar sample. Negative controls consisted in one BRIT63 and one IT85F sample both not harvested with CPSMV. Threshold Cycle (Ct) values inversely correlate with the quantity of mRNA expression, β-defensin expression, after inoculation, appears to be fast and strongly during the first 30 min after inoculation (BRIT63-0,5h) and response increase for the following 30 min (BRIT63-1h). **B)** In the table, results for *Vigna* RT-qPCR for defensin transcript detection, after CPSMV inoculation are reported.

Manuscrito 3 – Belarmino *et al.*, 2010

Figure 1. Schematic representation of the routine application. Annotation is accomplished through a central strategy plus differential expression profiling, phylogeny analyses and comparative modeling. 20x corresponds to the number of interactions with the psiBLAST tool.

Figure 2. Comparison of amino acid sequences of sugarcane defensin-like. Evolutionary relationship is depicted left. Plant defensin hallmarks are highlighted below in the primary sequence. Linked bars are disulfide bonds: Cys1-Cys8; Cys2-Cys5; Cys3-Cys6; Cys4-Cys7. The expression profile is represented right: LR=Leaf roll; LB=Lateral Bark; RT=Root apex; FL=base of inflorescence; SD=Developing seed; CL=Calli.Dark red and light red correspond to higher and lower expression.

Figure 3. Evolutionary reconstruction with Maximum Parsimony inference. SODEF01-17 (Sugarcane candidate defensin 1 to 17); ZMDEF1-4 (*Zea mays* defensin 1 to 4); VVDEF (*Vitis vinifera* defensin); 1GPS (γ-1-P thionin: *Triticum aestivum*); 1GPT (γ-1-P thionin: *Hordeum vulgare*); 1JKZ (Defensin: *Pisum sativum*); 2GL1 (Defensin: *Vigna radiata*); 1TI5 (Defensin: *Vigna radiata*); 1MR4 (Defensin: *Nicotiana alata*); 1N4N (Defensin: *Petunia* x *hybrida*). Square places the NMR structures used in the comparative modeling step. Numbers at the base of each clade correspond to bootstrap means at 1000 replications.

Figure 4. Structural alignment of the comparative modeled sequences of Sugarcane candidate defensins. (a) SODEF01 in yellow and SODEF04 in blue. (b) SODEF05 in green. (c) SODEF06 in green. (d) SODEF08 in yellow, SODEF10 in blue and SODEF11 in green. (e) SODEF13 in yellow and SODEF16 in blue. (f) SODEF17 in green, N-terminus in red arrow and C-terminus in blue arrow. The first 6 and final 14 amino acids of SODEF17 are not presented because these regions could not be modeled by comparative modeling. All disulphide bonds are highlighted in the figures above, with special attention to the mutation Cys1/SER1 that produces an unbound Cys8 (in grey) in the SODEF10 protein, presented in (d). All structural alignments and figures were made with the VMD program.

Manuscrito 4 – Benko-Iseppon *et al.*, 2009

Figure 1. Main mechanisms of pathogen recognition and defense in plants. Pathogenic organisms (mainly virus, bacteria and fungi) secrete *avr* (avirulence) gene products that may be compatible with *R* gene products secreted by the plants. Compatible interactions lead to the activation of signal cascades inducing systemic resistance factors (as ethylene and jasmonic acid) and acquired resistance represented by 17 *PR* gene categories. From these, three categories (*PR-12*, *13* and *14*) include small cisteine-rich anti-microbial (AMP) peptides.

LISTA DE TABELAS

Revisão bibliográfica

Tabela 1: Levantamento das defensinas vegetais em relação à classe, família, espécie e tecido de obtenção bem como informação sobre o peptídeo e as atividades biológicas conhecidas, acompanhado do número de aquisição no GenBank ou do número do cluster sugerido (TC) no TIGR Gene índices.

Manuscrito 2 – Belarmino e Benko-Iseppon, 2010

Table 1: Main web based repositories including plant antimicrobial peptides.

Manuscrito 4 – Benko-Iseppon *et al.*, 2010

Table 1: Small, Cys-rich antimicrobial peptides from the plant kingdom, including their classification and structural features. PDB refers to the Protein Data Base, DB to number of Disulfide Bridges.

Lista de abreviaturas

AMP	-	*Antimicrobial Peptides*
AMSDb	-	*Antimicrobial Database*
ANTIMIC	-	*The data base of ANTIMICrobial sequences*
APD2	-	*Antimicrobrial Peptides Database 2*
CCP	-	*Cys-cluster Proteins*
CONAB	-	*Companhia Nacional de Abastecimento*
CRP	-	*Cysteine-rich Proteine*
CyBASE	-	*The Cyclic Protein Database*
DDBJ	-	*DNA Data Bank of Japan*
DEFL	-	*Defensin-like Sequences*
DGED	-	*cDNA Differential Gene Expression Displayer*
EMBL	-	*European Molecular Biology Laboratory*
EST	-	*Expressed Sequences Tags*
GO	-	*Gene Ontology*
HMM	-	*Hidden Markov Model*
MNR	-	*Magnetic Nuclear Resonance*
MtGI	-	*Medicago Truncatula Gene Index*
NCR	-	*Nodule Specific Cys-rich Proteins*
NordEST	-	*Projeto Nordeste de sequenciamento de ESTs do Feijão-caupi*
PDB	-	*Protein Data Bank*
PDBe	-	*Protein Data Bank in Europe*
PDBj	-	*Protein Data Bank in Japan*
PhytAMP	-	*Plant Antimicrobial Peptides Database*
PRP	-	*Proline-rich Proteins*
RALF	-	*Rapid Alkanization Factor*
RCSB	-	*Research Collaboratory of Structural Bioinformatics*
RENORBIO	-	*Rede Nordeste de Biotecnologia*
RMSD	-	*Root Mean Square Distance*
SAGE	-	*Serial Analysis of Gene Expression*
SUCEST	-	*Sugarcane EST Project*
TrEMBL	-	*Translated European Molecular Biology Laboratory Protein database*
UniProt	-	*Universal Protein Resource*
3D	-	*Three Dimensional*

SUMÁRIO

1. Introdução 12
2. Revisão bibliográfica 13
    2.1. Defensinas 13
        2.1.1. Origem, diversidade, classificação e estrutura 14
        2.1.2. Distribuição, organização gênica e genômica 18
        2.1.3. Perfil de expressão 20
        2.1.4. Patógenos-alvo 21
        2.1.5. Modos de ação antimicrobiana 22
        2.1.6. Outras atividades biológicas conhecidas 24
        2.1.7. Transgenia e expressão heteróloga 26
        2.1.8. Perspectivas e demandas biotecnológicas 27
    2.2. Bioinformática 29
        2.2.1. Bancos de dados 29
        2.2.2. Análise de sequências 31
        2.2.3. Modelagem comparativa das defensinas 34
        2.2.4. Análise evolutiva 38
        2.2.5. Perfil de expressão digital 40
    2.3. Cana-de-açúcar 42
        2.3.1. Características genômicas da Cana-de-açúcar 43
        2.3.2. Importância econômica da Cana-de-açúcar 43
        2.3.3. O banco de dados SUCEST 44
    2.4. Feijão-caupi 44
        2.4.1. Características genômicas do feijão-caupi 45
        2.4.2. Importância econômica do feijão-caupi 46
        2.4.3. O banco de dados NordEST 47
3. Objetivos 47
    3.1. Objetivos específicos 48
4. Perspectivas 48

REFERÊNCIAS BIBLIOGRÁFICAS 50

APÊNDICES 75
    Apêndice 1 - Padovan *et al.*, 2010 75
    Apêndice 2 - Belarmino e Benko-Iseppon, 2010 85
    Apêndice 3 - Belarmino *et al.*, 2010 98
    Apêndice 4 - Benko-Iseppon *et al.*, 2010 119

ANEXOS 143
    Anexo A – Tabela 143
    Anexo B – Instruções para submissão de trabalhos na revista Prot. Pept. Lett 149
    Anexo C - Instruções para submissão de trabalhos na revista CPPS 160

1. Introdução

As plantas são constantemente atacadas por microorganismos presentes no meio durante seu ciclo de vida. Apesar disso, o desenvolvimento de doenças infecciosas é um acontecimento excepcional, pois ao longo da evolução os vegetais desenvolveram, como todos os organismos multicelulares, mecanismos de defesa contra microorganismos patogênicos. A produção de peptídeos antimicrobianos (AMP – *Antimicrobial peptides*) é uma estratégia amplamente empregada por muitos organismos como uma primeira linha de defesa. AMPs constituem um componente principal da imunidade inata dos animais, tratando-se de um mecanismo de defesa ancestral presente numa diversidade de organismos, incluindo microorganismos [Asaduzzaman e Sonomoto, 2009], artrópodes [Bulet and Stoecklin, 2005], fungos [Marx, 2004], animais [Bulet *et al*., 2004] e plantas [Manners, 2007].

Várias famílias de AMPs em vegetais foram definidas baseadas no número e arranjo de resíduos de cisteína presentes, bem como no padrão de pontes dissulfídicas estabelecidas, dentre estas uma família ampla, coletivamente chamada de defensinas vegetais, muito comum na natureza e com uma surpreendente conservação de estrutura e função. Algumas defensinas conhecidas atualmente são consideradas tão poderosas quanto a penicilina, assim como a vancomicina, acreditando-se que novas defensinas sejam ainda mais potentes, agindo mais especificamente contra certos microorganismos, incluindo estirpes resistentes aos antibióticos convencionais atualmente em uso, portanto tendo implicações promissoras no desenvolvimento de fármacos para o tratamento de muitas infecções letais [Mygind *et al*., 2005].

A crescente disponibilidade recente de genomas completamente sequenciados, bem como grandes quantidades de sequências expressas (ESTs – *Expressed Sequence Tags*) e o desenvolvimento de ferramentas computacionais oferecem vários recursos e algoritmos que possibilitam a aplicação de abordagens baseadas em bioinformática para identificar e desenvolver novos peptídeos antimicrobianos [Belarmino e Benko-Iseppon, 2010; Belarmino *et al*., 2010], capacitando o uso de técnicas computacionais para tornar conhecidos tais genes em táxons específicos ou relacionados a determinados processos em certos tecidos vegetais, como por exemplo, o desenvolvimento de resistência contra doenças [Menossi *et al*., 2008]. Um dos aspectos mais importantes no uso de ESTs está na possibilidade de acessar a informação genética de espécies com genomas complexos que oferecem grande dificuldade de acesso usando a genética convencional. Este é o caso da cana-de-açúcar, uma cultura de

grande importância agrícola e energética cultivada nos trópicos e sub-trópicos [D'Hont e Glaszman, 2001; Grivet e Arruda, 2001; D'Hont, 2005]. Por outro lado, o feijão-caupi apresenta um genoma pequeno, não obstante, pouco conhecido em relação a outras leguminosas consideradas como sistemas-modelo que foram intensamente estudadas através de ferramentas moleculares [Gepts *et al.*, 2005; Sato *et al.*, 2007; Timko *et al.*, 2008].

Nesse contexto, ferramentas de bioinformática desenvolvidas para prever padrões de sequências biológicas podem facilitar a criação de uma rotina para identificar novas defensinas em bancos de dados de ESTs vegetais para testes posteriores de sua atividade antimicrobiana. Técnicas computacionais recentes aplicadas ao campo da biologia estrutural vêm auxiliar essa tarefa, oferecendo dados estruturais comparativos que adicionam valor à atribuição de função. Outras fontes, como o perfil digital de expressão diferencial aliado à reconstrução da história filogenética dessas moléculas oferecem ainda mais informações que auxiliarão na escolha de candidatos para simulações computacionais e testes experimentais em programas de melhoramento vegetal e projetos para o descobrimento de novos fármacos.

Assim, o presente trabalho visou ao desenvolvimento de uma rotina computacional para identificar e caracterizar sequências codificantes para defensinas no genoma expresso da cana-de-açúcar e do feijão-caupi, traçando seu perfil de expressão e inferindo sobre suas relações evolutivas, bem como modelando comparativamente suas estruturas protéicas. Além disso, no âmbito da pesquisa, foi realizada uma revisão sobre bancos de dados e ferramentas para identificação de peptídeos antimicrobianos vegetais.

2. Revisão bibliográfica

2.1. Defensinas

As defensinas são peptídeos de defesa, estrutural e funcionalmente relacionados, provavelmente presentes em todo o espectro de vida. Esses peptídeos já foram caracterizados em diversos organismos eucariotos, incluindo cnidários [Sunagawa *et al.*, 2009], nematóides [Zhang *et al.*, 2000], moluscos [Charlet *et al.*, 1996], crustáceos [Saito *et al.*, 1995], aracnídeos [Ceraul *et al.*, 2007], insetos [Bulet P, Stocklin R, 2005], peixes ósseos [Zou *et al.*, 2007], aves [Dijk *et al.*, 2008], mamíferos [Yang *et al.*, 2007], fungos [Mygind *et al.*, 2005] e plantas [Lay e Anderson, 2005].

Análises filogenéticas sugerem que as defensinas compartilham um ancestral comum [Charlet *et al*., 1996; Dijk *et al*., 2008]. Recentemente, análises *in silico,* seguidas de caracterização funcional demonstraram a presença de peptídeos semelhantes às defensinas com atividade anti-Plasmodium na bactéria *Anaeromyxobacter dehalogenans*, indicando uma origem procariótica para as defensinas eucarióticas e sugerindo que essa estratégia de defesa ancestral, com a produção de peptídeos antimicrobianos, foi transferida às linhagens eucarióticas em algum ponto durante a evolução [Zhu, 2007; Gao *et al*., 2009]. Em concordância com essas evidências, Zhu [2008] identificou *in silico* 25 novos peptídeos antimicrobianos no fungo basal *Rhyzopus oryzae*, incluindo defensinas anteriormente reconhecidas como pertencentes a classes de defensinas exclusivas de invertebrados, de insetos e comum a plantas e insetos (**Figura 1**).

2.1.1. Origem, diversidade, classificação e estrutura

Como indicado na **Figura 1**, as defensinas de eucariotos provavelmente surgiram de um ancestral procarioto e evoluíram por eventos que incluem o ganho e perda de genes em linhagens específicas. As primeiras defensinas vegetais foram purificadas a partir de sementes de trigo (*Triticum aestivum* L.), cevada (*Hordeum vulgare* L.) e urtiga (*Urtica dioica* L.), sendo inicialmente classificadas entre as tioninas, principalmente devido a similaridades no peso molecular, tamanho e constituição da sequência de aminoácido, bem como no número de pontes dissulfídicas entre resíduos de cisteínas [Broekaert *et al*., 1989; Colilla *et al*., 1990; Mendez *et al*., 1990]. Entretanto, diferenças estruturais demonstraram que as defensinas vegetais constituem uma família à parte da família das tioninas [Bruix *et al*., 1995]. O termo defensina vegetal foi então sugerido por Terras *et al* [1995] devido à relação estrutural e funcional observada entre os peptídeos vegetais e as defensinas de mamíferos e insetos [Terras *et al*., 1995; Wong e Ng, 2007].

Estudos recentes demonstram a presença de peptídeos relacionados às defensinas vegetais em moluscos e fungos, indicando que as defensinas de plantas sejam mais intimamente relacionadas com as de moluscos e artrópodes do que com aquelas dos demais taxa [Tincu e Taylor, 2004; Mygind *et al*., 2005]. Essa assertiva se baseia nos seguintes fatos surpreendentes: (i) plantas, insetos e escorpiões apresentam defensinas que, apesar de terem grandes diferenças em peso molecular e sequência de aminoácidos, mantêm um enovelamento similar, centrado num motivo αβ estabilizado por pontes de cisteínas [Fehlbaum *et al*., 1994;

Cornet *et al*., 1995; Landon *et al*., 1996]; (ii) a estrutura primária de uma defensina da libélula *Aeschna cyanea*, um inseto basal da ordem Odonata, apresenta maior similaridade com defensinas de moluscos e escorpiões – intimamente relacionadas com defensinas vegetais – do que com defensinas de insetos [Bulet *et al*., 1999] e (iii) a observação de defensinas de plantas/insetos conservadas em três reinos de eucariotos - plantas, fungos e invertebrados [Zhu, 2008].

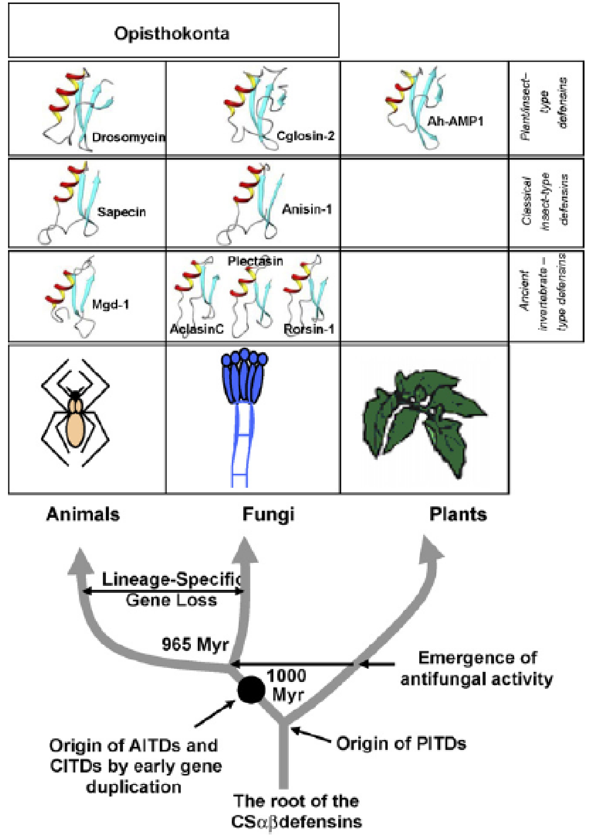


Figura 1: Representação esquemática da história evolutiva proposta para as defensinas. A árvore genealógica com a presença do supergrupo Opistoconta, incluindo fungos e animais, é usada para indicar o surgimento dos tipos diferentes de defensinas durante a evolução. Vários eventos importantes como duplicação gênica, ganho e perda gênica em linhagens específicas e surgimento da atividade antifúngica são indicados em diferentes pontos no tempo. **Fonte**: Zhu, 2008.

Vários estudos demonstraram um alto polimorfismo intraespecífico entre genes dessa família multigênica [Gu *et al*., 1992; Fedorova *et al*., 2002; Thomma *et al*., 2002; Mergaert *et al*, 2003; Hanks *et al*., 2005; Silverstein *et al*., 2005]. Entretanto, as defensinas vegetais exibem uma clara conservação de sequência (**Figura 2**), embora esta seja relativamente limitada ao arranjo de oito resíduos de cisteína e uma glicina entre o quarto e o quinto resíduo de cisteínas [Lay e Anderson, 2005].

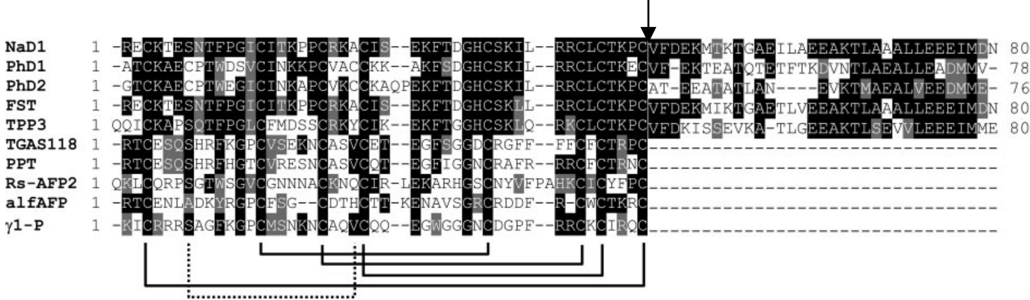


Figura 2: Alinhamento das sequências de aminoácidos das defensinas NaD1 de *Nicotiana alata*; PhD1, PhD2 e PPT de *Petunia hybrida*; FST de *N. tabacum*, TPP3 e TGAS118 de *Solanum lycopersicon*; RS-AFP2 de *Raphanus sativus*; alfAFP de *Medicago sativa* e γ1-P de *Triticum aestivum*. O peptídeos sinal foram omitidos. Resíduos idênticos foram destacados em preto, enquanto substituições conservativas foram destacadas em cinza. A seta indica o sítio de clivagem do pró-domínio C-terminal. Linhas sólidas representam o padrão de conectividade dissulfídica. Uma ponte dissulfídica adicional em PhD1 e PhD2 é mostrada como linha pontilhada. **Fonte:** Lay *et al*. (2003a), com modificações.

Conforme o efeito sobre o fungo *Fusarium culmorum*, as defensinas de plantas foram classificadas em quatro grupos. O grupo I, ou morfogênico, representado pela defensina RsAFP2 do rabanete (*Raphanus sativus* L.), a qual inibe o crescimento de *F. culmorum*, causando alterações morfológicas em sua membrana, com consequente redução do alongamento e ramificações das hifas. O grupo II, ou não morfogênico, representado pelas defensinas DmAMP1 isoladas de dália (*Dahlia merckii* Lehm.) e AhAMP1 isolada da castanha-da-índia (*Aesculus hippocastanum* L.), apresenta atividade antifúngica contra *F. culmorum* sem no entanto alterar sua morfologia. O grupo III é composto por defensinas que não apresentam atividade antifúngica, como por exemplo, as defensinas isoladas de sorgo (*Sorghum bicolor* (L.) Moench). O grupo IV inclui peptídeos com atividades antifúngica e antibacteriana, tal como a defensina isolada da semente da planta tropical de uso medicinal maravilha (*Mirabilis jalapa* L.) [Osborn *et al*., 1995; De Samblanx *et al*., 1997; Segura *et al*., 1998; Almeida *et al*., 2001].

Em uma segunda classificação, de acordo com o precursor codificante, as defensinas vegetais podem ser divididas em dois grupos (**Figura 3**). No maior deles, o mRNA maduro codificante para defensinas dá origem a duas partes distintas. A primeira parte, localizada na extremidade amino-terminal, é um peptídeo-sinal que direciona o peptídeo para o espaço extracelular. Os peptídeos-sinal de defensinas vegetais são frequentemente ácidos, excetuando-se alguns poucos exemplos. Semelhante região ácida foi relatada em defensinas de mamíferos, onde além de direcionar o peptídeo para as vias secretórias, houve ação

mitigante da atividade biológica do peptídeo maduro até que esta fosse necessária [Michaelson *et al.*, 1992]. A segunda parte do mRNA codifica o peptídeo maduro.

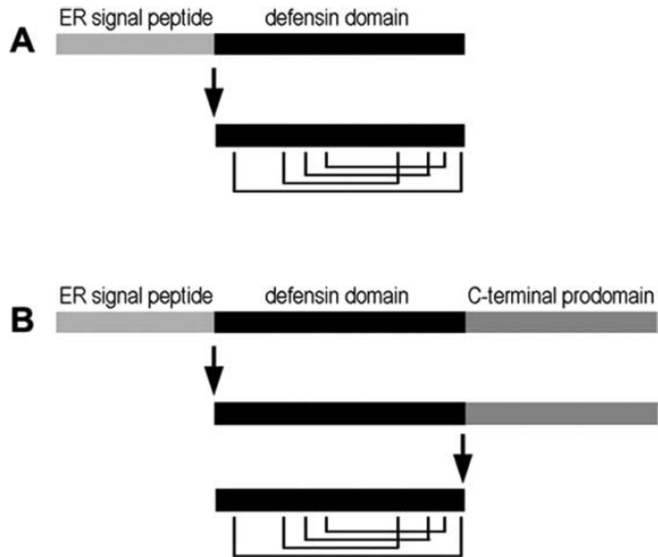


Figura 3: Duas classes de defensinas vegetais. Todas as defensinas vegetais são produzidas com um peptideos sinal de endereçamento ao retículo endoplasmático. (A) A maioria das defensinas vegetais possui apenas o domínio maduro além do peptídeo sinal. (B) Algumas defensinas vegetais isoladas apresentam um pró-domínio adicional na extremidade C-terminal. **Fonte**: Lay e Anderson (2005).

O segundo grupo de defensinas vegetais inclui peptídeos encontrados predominantemente em flores, codificados por precursores maiores, contendo, além das partes anteriormente mencionadas, um pró-domínio ácido e hidrofóbico na região carboxi-terminal, que contrabalanceia a carga positiva do domínio maduro. Uma região semelhante pode ser encontrada em defensinas de mamíferos e insetos, onde servem como sinal para o tráfico subcelular e processamento proteolítico pós-traducional, sendo também observada a ação mitigante [Gu *et al.*, 1992; Liu e Ganz, 1995; Milligan e Gasser, 1995; Brandstater *et al.*, 1996; Aluru *et al.*, 1999; Lay *et al.*, 2003a; Satchell *et al.*, 2003].

O processamento do peptídeo gera uma pequena molécula com peso molecular entre cinco e sete kDa, composta por 45 a 55 aminoácidos (**Figura 2**). Na estrutura primária do domínio maduro encontram-se oito resíduos de cisteína envolvidos em quatro pontes dissulfídicas, responsáveis pela estabilização da estrutura tridimensional, formando o motivo CSαβ presente em peptídeos revestidos de atividade antimicrobiana. A quarta ponte dissulfídica aproxima as extremidades amino e carboxi-terminais, criando um peptídeo pseudo-cíclico [Lay e Anderson, 2005]. A estrutura tridimensional de RsAFP1, determinada por ressonância magnética nuclear, tipifica o enovelamento global das defensinas vegetais (**Figura 4**). O enovelamento caracteriza uma estrutura globular compacta constituída de uma α-hélice e uma β-folha composta de três fitas antiparalelas, em uma configuração βαββ, que são estabilizadas

pelas pontes dissulfídicas intramoleculares, resultando numa estrutura muito compacta [Cornet *et al*., 1995; Fant *et al*., 1998; Jansen *et al*., 2003; Lay *et al*., 2003b].

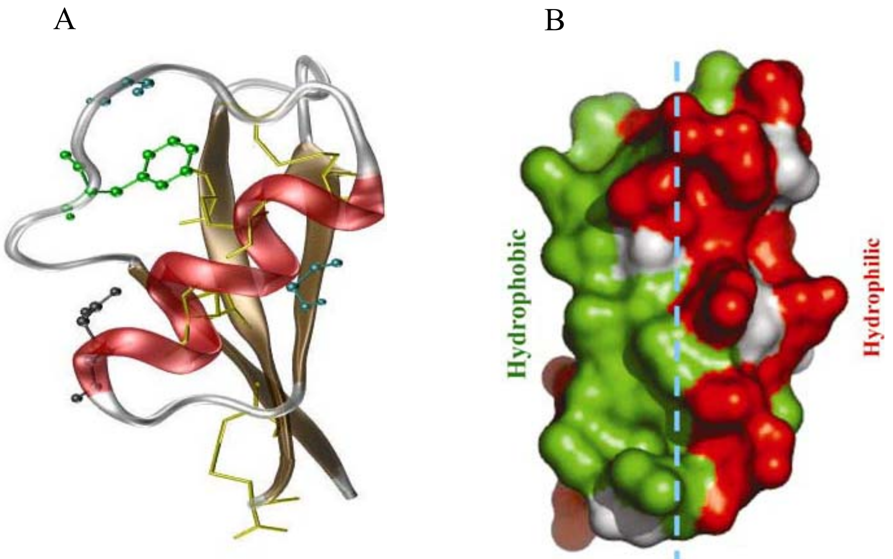


Figura 4: Estrutura 3D de defensinas vegetais como tipificada por RsAFP1. (A) Vv-AMP1 obtido por modelagem comparativa. A α-hélice e a β-folha estão respectivamente representadas em vermelho e marrom. Varetas amarelas representam as pontes dissulfídicas. (B) Representação globular, evidenciando a distribuição polarizada das cargas na superfície do peptídeo. **Fonte:** Beer e Vivier (2008); Zhu (2007).

2.1.2. Distribuição, organização gênica e genômica

Ao longo das duas últimas décadas várias defensinas foram isoladas em diferentes espécies vegetais. Atualmente, acredita-se que as defensinas estejam presentes em todas as espécies de plantas. De fato, peptídeos similares, representantes da família defensina, foram sucessivamente isolados e descritos em várias espécies vegetais (veja **Tabela 1** – Anexo A).

De acordo com análises disponíveis, a maioria dos genomas vegetais contém entre 15 e 50 genes para defensinas [Silverstein *et al*., 2005; Belarmino *et al*., 2010]. Em geral, o gene da defensina é composto por dois exons e um íntron de tamanho variável, inserido no meio da sequência que codifica o peptídeo sinal (**Figura 5**). O primeiro éxon codifica quase integralmente o peptídeo sinal, enquanto o segundo codifica o domínio funcional das defensinas [Terras *et al*., 1995; Doughty *et al*., 1998; Manners *et al*., 1998; Beer e Vivier, 2008; Pelegrini *et al*., 2008; Padovan *et al*., 2009]. Baseado em análises *in silico* no genoma de *Arabidopsis thaliana* (L.) Heynh., mais de 300 sequências de prováveis defensinas, nomeadas de DEFL (*defensin-like sequences*), foram descritas. Essas regiões gênicas apresentaram um padrão de distribuição genômica semelhante ao dos genes de resistência da

família NBS/LRR, organizado em 46 clusters espalhados nos cinco cromossomos de *A. thaliana* [Silverstein *et al*., 2005] (**figura 6**). Tal análise demonstrou que as defensinas formam uma grande família gênica e sugerem um modelo de evolução similar ao da família NBS/LRR, através de eventos sucessivos de duplicações em tandem e segmentares, seguidas de seleção positiva de genes ou de clusters gênicos.

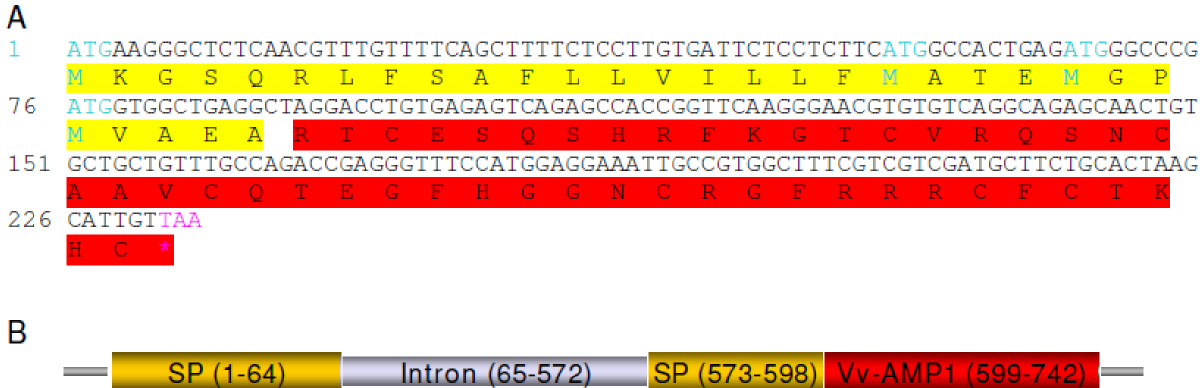


Figura 5: Representação esquemática da estrutura gênica de defensinas vegetais conforme observado em *Vitis vinifera*. (A) Sequência codificante de Vv-AMP1 com seus respectivos aminoácidos deduzidos. (B) Sequência genômica. Os blocos em amarelo representam a sequência codificante do peptídeo sinal, enquanto o bloco vermelho a sequência codificante do peptídeo maduro. O bloco em cinza indica a posição do íntron. Números correspondem ao tamanho em bp em cada seção. **Fonte:** Beer e Vivier (2008).

A organização genômica das defensinas ainda não está bem definida por duas razões: esses peptídeos ocorrem em formas numerosas, com uma origem diversa e variável e estão continuamente evoluindo; pouco se sabe sobre a organização genômica das defensinas em outras espécies de plantas. Entretanto, observou-se uma distribuição similar para algumas proteínas de nódulos ricas em cisteína específicas (NCRs - *Nodule specific Cys-Rich proteins*) , entre eles algumas defensinas, em uma análise preliminar de BACs do genoma de *Medicago truncatula*, sugerindo que esse padrão de organização possa ser uma tendência nos genomas vegetais [Alunni *et al*., 2007].

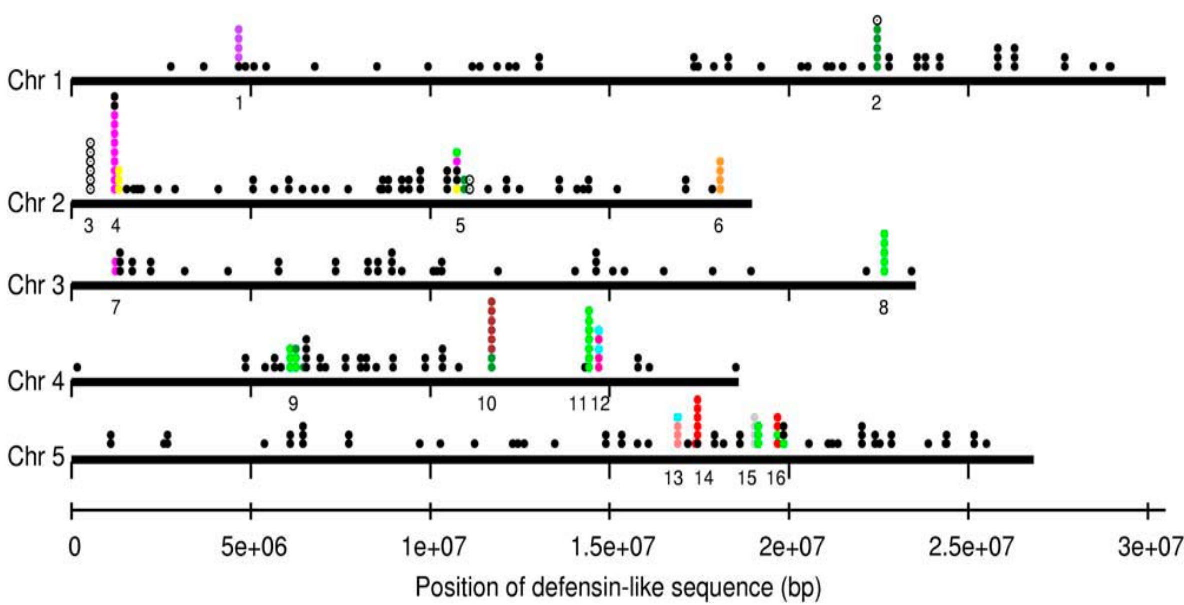


Figura 6: Organização genômica de DEFL relacionadas em *Arabidopsis thaliana*. Cada ponto representa um gene de DEFL, sendo os genes agrupados em espaço de 100.000 bp representados em pilhas verticais. Os maiores clusters estão enumerados, sendo as sequências coloridas para refletir relação de parentesco entre os subgrupos. **Fonte:** Silverstein *et al*. (2005).

2.1.3. Perfil de expressão

Os genes de defensinas vegetais apresentam um padrão de expressão muito complexo, com um amplo espectro de indução. Estudos de expressão com várias defensinas de *A. thaliana* revelaram diferentes padrões de expressão órgão-especifica, que provavelmente refletem funções distintas. Algumas defensinas são expressas constitutivamente, enquanto outras têm sua expressão aumentada em folhas após a infecção por patógenos. Em condições fisiológicas padrão, os genes para as defensinas foram encontrados expressos em tecidos diferentes, sendo diferencialmente regulados durante o desenvolvimento da planta [Epple *et al*., 1997; Thomma e Broekaert, 1998; Thomma *et al*., 1998].

A expressão de algumas defensinas pode ser induzida por estresses bióticos e abióticos, tais como frio [Koike *et al*., 2002; Carvalho *et al*., 2006], seca e salinidade [Yamada *et al*., 1997; Komori *et al*., 1997; Maitra e Cushman, 1998; Koike *et al*., 2002; Gaudet *et al*., 2003; Do *et al*., 2004]. Em geral, os genes de defensinas respondem e são induzidos pelo ácido jasmônico, ácido salicílico, ácido abscísico, etileno, benzotiadiazol, peróxido de hidrogênio, infecções fúngicas e ferimentos. Dependendo do modelo vegetal adotado no estudo, a resposta dos genes a esses estímulos pode ser muito diferente [Penninckx *et al*., 1996; De Samblanx *et al*., 1997; Terras *et al*., 1998]. Como resultado, a maioria dos tecidos vegetais

expressa constitutivamente dois ou mais genes de defensinas, sugerindo que defensinas individuais são expressas sob circunstâncias específicas ou em sítios específicos.

2.1.4. Patógenos-alvo

A relação com atividade antimicrobiana das defensinas vegetais foi descrita concomitantemente ao seu descobrimento no início da década de 1990. Essa atividade é observada principalmente contra fungos, mas os efeitos inibitórios contra bactérias gram-positivas e gram-negativas também podem ser observados, especialmente contra bactérias gram-positivas, embora a propriedade antibacteriana seja menos acentuada do que a antifúngica [Terras *et al.*, 1992; Moreno *et al.*, 1994; Zhang e Lewis, 1997; Segura *et al.*, 1998;]. A atividade inibitória contra várias bactérias de ambos os grupos já foram testadas, apresentando concentrações inibitórias tão baixas quanto 20 µM do peptídeo [Terras *et al.*, 1992; Terras *et al.*, 1993; Osborn *et al.*, 1995; Segura *et al.*, 1998; Koike *et al.*, 2002; Fujimura *et al.*, 2004; Chen *et al.*, 2005; Wong *et al.*, 2005a; Wong *et al.*, 2005b; Franco *et al.*, 2006; Wong *et al.*, 2006; Huang *et al.*, 2008; Van der Weerden *et al.*, 2008].

A atividade antifúngica é a melhor caracterizada. Vários estudos mostraram um efeito inibitório muito potente (concentrações inibitórias do peptídeo tão baixas quanto 1 $\mu g.mL^{-1}$) contra várias espécies de fungos, incluindo muitos patógenos vegetais como *Alternaria brassicola*, *A. solani*, *Botrytis cinerea*, *Cladosporium colocasiae*, *C. sphaerospermum*, *Colletotrichum lindemuthianum*, *Diploidia maydis*, *Fusarium culmorum*, *F. decemcellulare*, *F. graminearum*, *F. oxysporum*, *F. verticillioides*, *Mycosphaerella arachidicola*, *M. fijinesis*, *Nectria haematococca*, *Penicillium digitatum*, *P. expansum*, *Pericularia oryzae*, *Phaeoisariopsis personata*, *Physalospora piricola*, *Rhizoctonia solani*, *Septoria tritici*, *Trichoderma viride*, *Verticilium albo-atrum*, *V. dahliae*, bem como contra o patógeno humano *Candida albicans* e os oomicetos *Phytophthora infestans* e *P. parasitica* [Terras *et al.*, 1992; Terras *et al.*, 1993; Osborn *et al.*, 1995; Park *et al.*, 2002; Ye e Ng, 2002; Wisniewski *et al.*, 2003; Chen *et al.*, 2005; Wong *et al.*, 2005a; Wong *et al.*, 2005b; Anaya-Lopez *et al.*, 2006; Olli e Kirti, 2006; Wang e Ng, 2006; Solis *et al.*, 2007; Finkina *et al.*, 2008; Odintsova *et al.*, 2008].

Alguns poucos exemplos relatam a presença de atividade inibitória contra a transcriptase reversa do HIV-1, muito embora os testes realizados tenham sido apenas *in vitro* [Ye e Ng, 2001; Ye e Ng, 2002; Wong *et al.*, 2005a; Wong *et al.*, 2005b].

2.1.5. Modos de ação antimicrobiana

Justamente por ser mais bem documentado, se conhece muito mais sobre o modo de ação antifúngico do que sobre o modo de ação contra os demais microorganismos, sobretudo se o peptídeo apresenta efeito antimicrobiano *in vivo*, como é o caso da propriedade antiviral contra HIV-1. O processo exato de inibição do crescimento fúngico ainda não é conhecido, porém os estudos disponíveis indicam que esse processo se iniciaria na membrana celular através da interação da defensina com glicoesfingolipídeos da membrana do fungo [Thevissen *et al*., 2003; Thevissen *et al*., 2005]. Glicoesfingolipídeos têm sido relacionados com moléculas eliciadoras do mecanismo de defesa de arroz, um papel consistente com seu envolvimento em vários processos celulares como a transdução de sinais, transporte de proteínas para a membrana [Bagnat *et al*., 2000], morte celular programada [Malisan *et al*., 1999] e adesão de patógenos à membrana do hospedeiro [Ghannoum *et al*., 1987; Jimenez-Lucho *et al*., 1990; Koga *et al*., 1998; Umemura *et al*., 2000].

Um modelo desenvolvido unindo vários resultados experimentais sugere que essa interação promova a formação de um tapete de defensinas na membrana externa do invasor, provocando a alteração da estrutura da membrana, com posterior formação de poros transientes e inserção na membrana, causando sua permeabilização que resulta no aumento do influxo de cálcio e do efluxo de potássio (**Figura 7**). Essa interação mostrou-se necessária, mas insuficiente para a inibição do crescimento do fungo *Pichia pastoris*, sugerindo o envolvimento de outros alvos para a atividade antifúngica [Thevissen *et al*., 2003]. Assim, de acordo com o modelo sugerido, pode haver uma ruptura violenta da membrana associada à interação das defensinas com componentes da membrana do fungo ou pode ocorrer a difusão de algumas defensinas que alcançam o meio intracelular e interagem com alvos internos, como a proteína ciclina F, interrompendo o término do ciclo celular [Zasloff, 2002; Lobo *et al*., 2007].

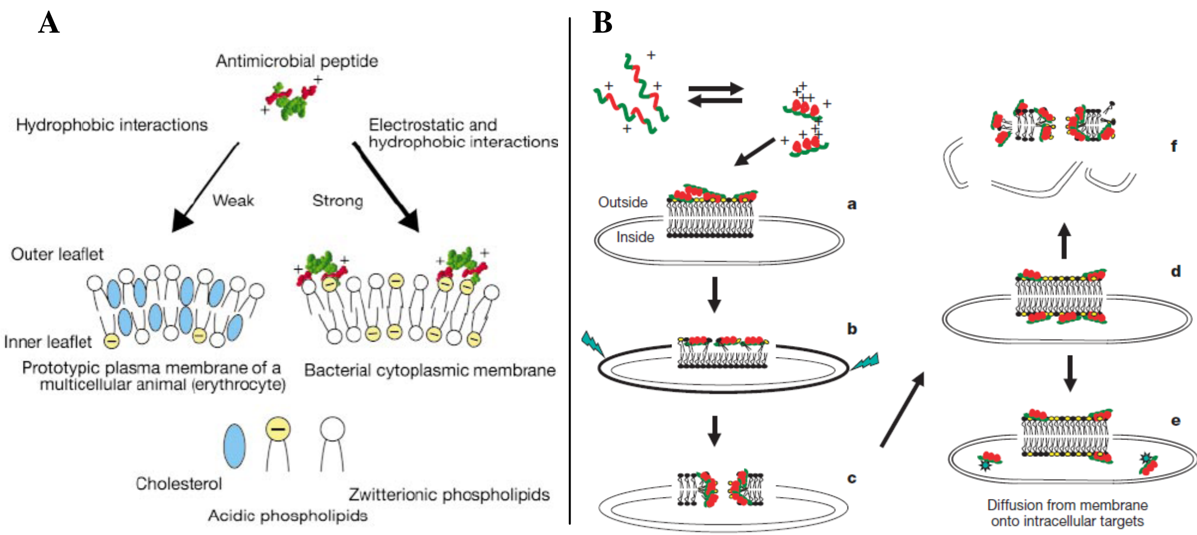


Figura 7: Modelo proposto para o modo de ação juntando evidências de fontes variadas. (A) O alvo molecular de peptídeos antimicrobianos de organismos multicelulares nas membranas celulares e a base de sua especificidade. (B) Modelo de Shai-Matsuzaki-Huang do mecanismo de ação de um peptídeo antimicrobiano. (**a**). recobrimento do folheto externo da membrana formando um tapete de peptídeos. (**b**). integração do peptídeo na membrana e afinamento do folheto externo. A área de superfície do folheto externo da membrana se expande em relação ao folheto interno, resultando em uma tensão dentro da bicamada lipídica (setas denteadas). (**c**). transição de fase e formação poros transientes se formam nessa fase. (**d**). transporte de lipídeos e peptídeos para o folheto interno. (**e**). difusão de peptídeos que encontram alvos intracelulares.(**f**). colapso da membrana em fragmentos e rompimento físico da membrana-alvo. Lipídeos acídicos ou negativamente carregados estão representados com extremidade amarela, enquanto lipídeos sem carga geral estão representados com extremidade em preto. **Fonte:** Zasloff (2002).

Consistente com o envolvimento de glicoesfingolipídeos na transdução de sinais, é possível que após a interação com essas moléculas outras defensinas permaneçam fora da célula e desencadeiem uma cascata de sinais intracelulares, através da produção de espécies de oxigênio reativo, induzindo à morte celular do fungo [Thevissen *et al.*, 1996; Thevissen *et al.*, 1999; Aerts *et al.*, 2007; Ramamoorthy *et al.*, 2007]. Esses resultados intensificam a idéia de que as defensinas devem atuar por vários e diferentes mecanismos, demonstrando o quão sofisticado e complicado é o mecanismo de defesa contra microorganismos.

O mecanismo de ação antibacteriano ainda não está bem compreendido, pois há poucos estudos sobre essa atividade das defensinas vegetais. Entretanto, algumas inferências podem ser feitas a partir das defensinas de animais que em geral possuem uma atividade antibacteriana mais acentuada (**Figura 7A**). Dado o alto potencial transmembrana de células bacterianas e a natureza de carga negativa da membrana devido à presença de ácido teóico, lipossacarídeos e fosfolipídios, as defensinas podem se ancorar à membrana das bactérias

através destes elementos, exercendo então seu efeito tóxico [Ganz e Lehrer, 1994; De Samblanx *et al*., 1997; Ganz, 2004; Brogden, 2005; Monk e Harding, 2005; Papo e Shai, 2005; Buscaglia *et al*., 2006; Van Djik *et al*., 2008]. O quadro atual mostra que ainda há muitos questionamentos com respeito ao modo de ação antibacteriano que a comunidade científica deverá investigar nos próximos anos. Entre eles, questiona-se qual seria o receptor das defensinas nas membranas bacterianas, quais aminoácidos interagiriam com esses elementos e quais seriam os efeitos dessa interação na inibição do crescimento bacteriano.

2.1.6. Outras atividades biológicas conhecidas

Além da atividade antimicrobiana, considera-se que as defensinas apresentem uma extensa gama de atividades biológicas *in vitro*. Surpreendentemente, defensinas individuais apresentam uma ou mais dessas atividades, mas não todas. A primeira atividade biológica descrita para as defensinas vegetais foi sua habilidade de inibir a tradução protéica em mamíferos e procariotos, porém sem qualquer efeito na tradução em plantas [Mendez *et al*., 1990; Mendez *et al*., 1996]. Entretanto, algumas defensinas de *Vigna radiata* (L.) R. Wilczek inibiram a tradução em sistemas vegetais, mas esses são os únicos relatos conhecidos no momento [Chen *et al*., 2002; Chen *et al*., 2004; Chen *et al*., 2005]. Pouco se sabe sobre o mecanismo de inibição da tradução usado pelas defensinas vegetais; porém, dados recentes indicam que esse processo pode ocorrer em etapas diferentes da tradução, visto que defensinas não possuem a habilidade de se ligar a moléculas de ácidos nucléicos [García-Olmedo *et al*., 1983; Mendez *et al*., 1990; Mendez *et al*., 1996].

Duas outras atividades conhecidas atualmente para as defensinas as inserem na classe de inibidores enzimáticos. No início da década de 1990 as atividades inibidoras de α-amilase e de proteinases apresentadas por algumas defensinas foram reconhecidas [Mendez *et al*., 1990; Bloch e Richardson, 1991; Wijaya *et al*., 2000; Wong *et al*., 2005b; Wong *et al*., 2006; Molosov e Valeuva, 2008]. Estudos subsequentes mostraram que a atividade inibidora da α-amilase está presente em algumas defensinas e ausentes em outras. Esses estudos destacaram que diferenças estéricas na estrutura tridimensional das defensinas de diferentes origens definem a atividade inibitória de α-amilase [Chen *et al*., 2002; Chen *et al*., 2004; Chen *et al*., 2005; Liu *et al*., 2006; Lin *et al*., 2007; Pelegrini *et al*., 2008]. Com relação ao mecanismo de inibição de tripsina especula-se que após ligar-se à tripsina, resíduos específicos da defensina interagem com o bolso S1 da proteinase, causando sua inibição [Melo *et al*., 2002].

Interessantemente, em oposição à sua ação inibitória de enzimas, algumas defensinas possuem ação enzimática dependente da glutationa envolvida no estado redox do ácido ascórbico, com prováveis implicações no modo de resposta vegetal a espécies de oxigênio reativo [Chen e Gallie, 2006; Huang *et al*., 2008].

Algumas defensinas surpreendentemente apresentaram atividade mediadora de tolerância ao zinco. Pouco se sabe sobre essa atividade até o momento, porém estudos de expressão de cDNA de *Arabidopsis halleri* (L.) O'Kane & Al-Shehbaz em *Sacharomyces cerevisiae*, posteriormente incubadas em meio contendo concentrações tóxicas de zinco, possibilitaram a identificação de quatro defensinas envolvidas na tolerância ao metal. Em seguida, esses cDNAs foram funcionalmente expressos em *A. thaliana* e *A. halleri,* promovendo tolerância ao zinco, porém o mecanismo responsável pela tolerância ainda não foi determinado [Pilon-Smits, 2005; Mirouze *et al*., 2006].

Outra propriedade conhecida para as defensinas vegetais surgiu de estudos com as defensinas γ1-zeationina e γ2-zeationina de milho. Tais defensinas apresentaram atividade inibidora de canais iônicos, bloqueando o fluxo de sódio respectivamente em concentrações de 62 μM e 33 μM. A defensina MsDEF1 de *Medicago sativa* L. bloqueou quase totalmente o fluxo de cálcio através do canal do tipo L; porém, essa atividade inibidora de canais iônicos não foi observada em outras defensinas. Nada se conhece sobre o modo de ação da atividade inibidora de canais iônicos, sendo proposto que essa atividade provavelmente seja devida à similaridade estrutural da defensina MsDEF1 com a proteína KP4 bloqueadora de canais de cálcio dependente de voltagem [Kushmerick *et al*., 1998; Spelbrink *et al*., 2004].

Finalmente, dada as propriedades das defensinas até agora descritas, vários estudos visaram ao teste de sua atividade biológica sobre células de mamíferos, através da determinação da viabilidade de células endoteliais de cordão umbilical, fibroblastos de músculos e da pele, bem como da lise de eritrócitos. Nenhuma das defensinas testadas reduziu a viabilidade das células mencionadas, como também a hemólise não foi observada mesmo em concentrações tão altas quanto 500 $\mu g.mL^{-1}$ [Terras *et al*., 1992]. Entretanto, estudos posteriores demonstraram que nem todas as defensinas vegetais são desprovidas de atividade biológica sobre células de mamíferos. A atividade mitogênica sobre esplenócitos de ratos foi relatada para algumas defensinas vegetais [Ye e Ng, 2001; Ye e Ng, 2002; Wong *et al*., 2006].

Adicionalmente, o efeito inibitório sobre certos tipos de células cancerígenas de humanos foram descritas para várias defensinas vegetais. A defensina de *Phaseolus vulgaris,* vulgarinina, diminuiu a proliferação das linhagens de células de leucemia L1210 e M1

respectivamente em cerca de 35% e 80%. A proliferação de células da linhagem de câncer de mama MCF-7 foi inibida em 80% por esse peptídeo. Outras linhagens de câncer, como a HeLa e a Bel-7402 foram inibidas em cerca de 80% em concentrações inibitórias médias de 43 μM e 28 μM, respectivamente [Wong *et al*., 2005a; Wong *et al*., 2005b; Anaya-Lopez *et al*., 2006; Wang *et al*., 2008]. O mecanismo inibitório da proliferação cancerígena ainda não foi elucidado, porém células cancerígenas apresentam propriedades que se assemelham mais a células de microorganismos do que àquelas dos próprios mamíferos [Papo e Shai, 2005], levando à suposição de que o mecanismo de inibição da proliferação de células cancerígenas deve se assemelhar ao postulado para a inibição do crescimento bacteriano.

2.1.7. Transgenia e expressão heteróloga

As propriedades descritas para as defensinas classificam-nas como bons candidatos para o desenvolvimento de transgênicos de plantas agronomicamente importantes, auxiliando no combate contra patógenos e pragas. Atualmente, várias plantas cultivadas foram transformadas com um gene de defensina vegetal, cuja expressão constitutiva representou um ganho de resistência em todos os estudos realizados. Por exemplo, plantas transgênicas de arroz, expressando constitutivamente a defensina de *Wasabia japonica* (Miq) Matsum, mostraram resistência aumentada contra o fungo *Magnoporthe grisea* [Kanzaki *et al*., 2002]. Plantas de mamão papaia (*Carica papaya* L.) transformadas com a defensina DmAMP1 tornaram-se mais resistentes a *Phytophthora palmivora* [Wong *et al*., 2006]. Um leve aumento de resistência contra a doença da canela preta, causada pelo fungo *Leptosphaeria maculans*, foi observado em plantas transgênicas de canola expressando uma defensina extraída de *Pisum sativum* [Wang *et al*., 1999].

Um estudo recente com plantas de tabaco e de amendoim transformadas com uma defensina de mostarda demonstrou que a transformação com um único gene de defensina pode conferir resistência a vários patógenos [Anuradha *et al*., 2008]. Além do mais, foi demonstrada a viabilidade do desenvolvimento de plantas transgênicas usando uma poli-proteína artificial composta de duas defensinas de origens diferentes, cujos transgênicos mostraram uma maior concentração das defensinas produzidas como poli-proteínas do que os transgênicos transformados com apenas uma delas, indicando um modo de aumentar os níveis de expressão de pequenas proteínas como também o potencial de obtenção de cultivares concomitantemente resistentes a vários patógenos [François *et al*., 2002].

Talvez o melhor exemplo do potencial das defensinas no desenvolvimento de plantas transgênicas de importância agrícola vem do estudo de Gao e colaboradores [2000] demonstrando que a expressão constitutiva da defensina alfAFP de alfafa (*M. sativa*) em batata (*Solanum tuberosum* L.) promoveu uma resistência robusta contra *Verticillium dahliae*, não apenas em casa de vegetação, como também em testes de campo durante vários anos e em diferentes localidades geográficas.

Além da transformação de sistemas vegetais, várias defensinas foram heterologamente expressas em bactérias, leveduras e fungos filamentosos. Entre as bactérias, *Escherichia coli* foi frequentemente usada para expressar defensinas. Esse sistema foi usado para expressar defensinas de *A. halleri*, *Tephrosia villosa* (L.) Pers. e da batata doce (*Ipomoea batatas* (L.) Lam.) [Huang *et al*., 2008; Vijyan *et al*., 2008; Marquès et al, 2009; Kovalskaya e Hammond, 2009]. Entre as leveduras, *S. cerevisiae* e *P. pastoris* têm sido geralmente escolhidas como os melhores sistemas para expressar defensinas [Cabral *et al*., 2003; Kant *et al*., 2009], porém apenas este último sistema tem sido usado para obter alta produção de defensinas. Sistemas de expressão baseados em fungos filamentosos, como *Aspergillus* e *Fusarium*, foram usados para a produção de algumas defensinas, tendo se mostrado superiores a sistemas baseados em leveduras, fornecendo níveis de expressão de dezenas de g/L nos casos mais bem sucedidos [Yoder e Lehmbeck, 2004; Mygind *et al*., 2005].

2.1.8. Perspectivas e demandas biotecnológicas

Dadas as propriedades das defensinas vegetais, fica claro o seu potencial biotecnológico para a descoberta de novas drogas e para a construção de plantas transgênicas com maior resistência às pragas e patógenos. Outros peptídeos têm sido utilizados como uma fonte particular de resistência, tais como transferidores de lipídeos, inibidores de enzimas digestivas, bem como genes *R* de resistência [Franco *et al*., 2002; Carvalho *et al*., 2006; Langen *et al*., 2006; Murad *et al*., 2007]. No entanto, as defensinas vegetais têm ocupado o 1º posto numa corrida mundial para obtenção de culturas agrícolas de alta produtividade, com resistência contra estresses bióticos e abióticos. Tal posição parece estar relacionada com a multifuncionalidade das defensinas, podendo agir contra bactérias, fungos, insetos e estresses abióticos [Pelegrini e Franco, 2005].

Defensinas vegetais são promissores agentes terapêuticos em humanos e outros animais, pois em contraste com outros peptídeos antimicrobianos, o seu modo de ação envolve a

ligação altamente específica às membranas fúngicas, explicando sua baixa citotoxicidade. A descoberta de que algumas defensinas vegetais são capazes de ligar-se a ciclinas abre um novo caminho para o desenvolvimento de estratégias terapêuticas para o tratamento de alguns tipos de câncer em humanos, onde a expressão de alguns tipos de ciclinas se apresenta aumentada [Kong *et al*., 2000; Yasuda *et al*., 2002; Lobo *et al*., 2007]. De fato, como discorrido anteriormente, o efeito inibitório da proliferação de alguns tipos de câncer já foi observado para algumas defensinas, embora o modo de ação pelo qual essas defensinas exercem seus efeitos ainda não tenha sido elucidado. As pesquisas em vegetais têm tradicionalmente contribuído para o desenvolvimento de novas drogas através da descoberta de compostos bioativos que são usados para tratar infecções e outras doenças, tais como a droga antiinflamatória aspirina e o agente anticancerígeno taxol [Van Baarlen *et al*., 2007]. As defensinas vegetais possivelmente representarão a próxima contribuição das ciências vegetais para as ciências médicas.

A produção de defensinas em plantas transgênicas também poderia ser utilizada com objetivos diferentes, além da obtenção de uma fonte de resistência. Defensinas vegetais têm sido indicadas como uma possível nova droga para o controle de infecções humanas, especialmente em relação a estirpes de bactérias resistentes. Nesse sentido, plantas transgênicas poderiam atuar como biofábricas, produzindo defensinas em larga escala para usos farmacêuticos. Várias técnicas foram desenvolvidas para esse propósito, visto que a produção industrial de defensinas frequentemente apresenta alguns problemas tais como formação de altos níveis de estruturas intermediárias sem o correto enovelamento em sistemas de expressão procarióticos como *E. coli* e mesmo eucariótico como *P. pastoris*, conexões imprecisas na formação das pontes dissulfídicas, resquícios frequentes de sequências de purificação que interferem na atividade do peptídeo, potencial recalcitrante para sistemas de expressão baseados em microorganismos como bactérias e fungos [Sels *et al*., 2007]. Afortunadamente esses problemas podem ser contornados utilizando cepas de bactérias adaptadas, linhagens de fungos resistentes pela atividade antifúngica das defensinas, através de protocolos de desnaturação e renaturação dos intermediários obtidos, ou mesmo obtenção *in planta* [Sels *et al*., 2007; Thevissen *et al*., 2007; Marquès *et al*., 2009].

Bioinformática

A vasta quantidade de dados biológicos diversos gerados pelo recente avanço biotecnológico levou ao desenvolvimento e evolução da bioinformática. Esse campo relativamente novo tem facilitado tanto a análise de dados genômicos e pós-genômicos, permitindo também a integração da informação provinda de várias fontes relacionadas, como a transcriptômica, a proteômica, a metabolômica e a fenômica. Tal integração tem permitido a identificação de genes e produtos gênicos, podendo elucidar relações funcionais entre genótipos e o fenótipo observado, desse modo permitindo uma ampla análise sistêmica partindo do genoma ao fenoma. Devido ao crescente valor e amplo alcance da biotecnologia, a bioinformática ocupa um papel muito importante na integração de vários dados gerados pela expansão das chamadas tecnologias "omicas" [Edwards e Batley, 2004].

2.2.1. Bancos de dados

O rápido crescimento da informação sobre sequências de nucleotídeos tornou necessário o desenvolvimento de bancos de dados específicos para armazená-las e distribuí-las. O maior deles surgiu em 1986 a partir da cooperação entre o GenBank e o EMBL (*European Molecular Biology Laboratory*), sendo em 1987 iniciada a participação do DDBJ (*DNA Data Bank of Japan*). Esse meta banco de dados é mundialmente considerado como sendo o repositório padrão para sequências de DNA [Edwards e Batley, 2004].

Alguns outros bancos de dados lidam com sequências protéicas. Ademais, um grande número de bancos de dados mais especializados está disponível como, por exemplo, bancos de dados de estruturas protéicas, de identificação de proteínas, de características especiais de genes e/ou proteínas, bem como de organismos específicos. Entre os bancos de proteínas mais expressivos estão o Entrez proteína [Wheeler *et al*., 2005; www.ncbi.nlm.nih.gov/protein/] e o UniProt [Bairoch *et al*., 2005]. Entre os bancos de dados especializados, se destacam os dedicados a organismos-modelo como *Homo sapiens* [Wain *et al*., 2002], *Mus musculus* [Eppig *et al*., 2005], *A. thaliana* [Swarbreck *et al*., 2008] e *O. sativa* [Ouyang *et al*., 2007]. Entre os bancos de identificação de proteínas, o banco de ontologias gênicas (GO – *Gene Ontology*) provê uma árvore de vocabulários controlados descrevendo a função molecular, o papel biológico e a localização celular de produtos gênicos [Camon *et al*., 2004]. Representações de interações protéicas podem ser acessadas a partir do banco IntAct que

trabalha usando anotações GO das proteínas para garantir a consistência da informação [Hermjakob *et al*., 2004]. Experimentos sobre expressão de proteínas podem ser acessados através do banco SWISS-2DPAGE, que disponibiliza experimentos de eletroforese em gel de poliacrilamida em duas dimensões, bem como de dodecil sulfato de sódio [Hermjakob *et al*., 2004].

Dados de estrutura protéica podem ser obtidos no banco de dados protéicos (PDB – *Protein Data Bank*), cuja colaboração entre o grupo de pesquisa colaborativa para bioinformática estrutural (RCSB – *Research Collaboratory for Structural Bioinformatics*), o banco de dados protéicos da Europa (PDBe – *Protein Data Bank in Europe*) e o banco de dados protéicos do Japão (PDBJ – *Protein Data Bank of Japan*) provêem dados de Raios-X e ressonância magnética nuclear para mais de 32.000 estruturas protéicas, ácidos nucléicos e carboidratos [Berman *et al*., 2000].

Muitos outros bancos de dados com informações mais específicas e direcionadas estão disponíveis pela internet para acesso público. O valor dessas fontes pode ser inferido à medida que uma rede interconectada de bases relacionadas possa ser estabelecida (Mesiti *et al*., 2009). De fato, muitos desses bancos mantêm referências cruzadas com outros bancos de dados, oferecendo informações básicas para estratégias mais sofisticadas de integração de dados. Alem disso, há uma tendência para o desenvolvimento de ferramentas capazes de encontrar a relação entre vários níveis de informação biológica através da mineração da literatura disponível (Krallinger *et al*., 2008).

Dado a importância dos AMPs, além de conhecer suas sequências de aminoácido, é muito importante entender a sua estrutura, topologia e função, os quais em conjunto permitem um melhor entendimento de sua ação efetiva contra patógenos. Dados relativos a várias classes de peptídeos antimicrobianos, incluindo as defensinas, podem ser acessados em diferentes repositórios na web, resultantes de vários esforços para coletar, processar e armazenar tais sequências. Cada um dos bancos de dados referidos foi desenvolvido com um propósito específico e relativo a um determinado conjunto de dados, agrupando moléculas de peptídeos antimicrobianos de uma diversidade de organismos, incluindo procariotos e eucariotos, além de algumas ferramentas para a sua avaliação comparativa (Brahmachary *et al*., 2004; Wang e Wang, 2004; Antcheva *et al*., 2006; Fjell *et al*., 2007; Seebah *et al*., 2007; Wang *et al*., 2008a; Wang *et al*., 2008b; Hammami *et al*., 2009).

Por exemplo, o banco de dados AMSDb (*Antimicrobial database*) é o banco de dados mais antigo sobre peptídeos antimicrobianos disponível na internet, contendo vários AMPs

vegetais. Este banco de dados permite buscas através de palavra-chave bem como por algumas características tais como organismo de origem, perfil de expressão e atividade, entre outras [Antcheva *et al*., 2006]. O banco de dados ANTIMIC (*data base of ANTIMICrobial sequences*) inclui 1.700 possíveis peptídeos antimicrobianos conhecidos, também integrando ferramentas para facilitar uma eficiente extração de dados e análise em nível molecular bem como pesquisar novos AMPs [Brahmachary *et al*., 2004]. Semelhantemente, Seebah *et al*. [2007] disponibilizaram um banco de dados de sequências e fontes de informação curados especificamente em relação às defensinas.

Outro recurso, o CyBASE (*The cyclic protein database*), foi desenvolvido por Wang *et al*. [2008a] como uma base de informações sobre AMPs com arcabouço circular, incluindo ferramentas de busca e interfaces de exposição de estrutura de sequência e função. Adicionalmente, Fjell *et al*. [2007] desenvolveram o banco de dados AMPer que usa o modelo estatístico de cadeias ocultas de Markov (HMM – *Hidden Markov Model*) aplicado a sequências publicamente disponíveis de AMPs. Este banco de dados permite o reconhecimento de classes de AMPs individuais, tais como defensinas, catelecidinas e cecropinas com uma precisão ≥ 99 %, constituindo uma excelente ferramenta de descoberta para a identificação de membros específicos da superfamília de AMPs.

Wang e Wang [2004] desenvolveram um banco de dados dedicado a AMPs de todas as formas de vida, desde bactérias a plantas e animais, inclusive humanos. Recentemente, este banco de dados foi atualizado sendo designado como APD2 (*Antimicrobial Peptides Database 2*) [Wang *et al*., 2008b], um recurso que provê dados estatísticos para peptídeos presentes em bancos de dados, bem como ferramentas para avaliação da relação estrutura/função relativo a AMPs.

Por fim, o banco de dados PhytAMP (*Plant Antimicrobial Peptides database*) é o único exclusivamente dedicado a plantas, apresentando informação valiosa relativa a 271 peptídeos antimicrobianos vegetais, incluindo informação taxonômica e microbiológica, além da disponibilização de ferramentas para inferências relativas à estrutura de sequências e à função [Hammami *et al*., 2009].

2.2.2. Análise de sequências

A atenção da comunidade científica agora está voltada para a anotação genômica, ou seja, o processo de adicionar análises e interpretações necessárias para extrair significado biológico

de sequências, inserindo-as no contexto de nosso entendimento sobre processos biológicos. Essa é uma tarefa que frequentemente é realizada em múltiplos passos, que podem ocorrer em três níveis: nucleotídico, protéico e de processos [Stein, 2001].

A primeira coisa a se fazer com uma sequência de nucleotídeos em mãos é identificar sua localização no genoma, o que juntamente com a descrição da estrutura gênica fornece um modo para conectar informações de várias fontes de pesquisas pré- e pós-genômicas. Várias ferramentas de alinhamentos entre sequências estão disponíveis, a exemplo do tradicional algoritmo BLAST [Altschul *et al*., 1990]. Tais ferramentas podem ser utilizadas para buscar similaridades e informações de outras espécies que podem ser extrapoladas para a sequência em mãos. Alguns desses algoritmos oferecem a possibilidade de passar de um nível a outro da informação, como é o caso do algoritmo BLASTx que pode retornar boas evidências de que uma sequência codifica uma proteína, acrescentando informação ao nível protéico, bem como dar suporte à suposição de que uma dada sequência pertence a um gene [Altschul *et al*., 1990].

Após buscar respostas sobre a localização de uma dada sequência, a anotação prossegue buscando informações sobre o que é codificado pela sequência. A comparação entre proteínas de espécies diferentes é uma fonte inestimável para a anotação funcional. Uma estratégia de anotação protéica buscará similaridades utilizando principalmente as ferramentas BLASTp e PSI-BLAST contra bancos de proteínas [Altschul e Koopin, 1998]. Uma estratégia complementar pode ser implementada buscando domínios funcionais em bancos de dados como o PFAM [Bateman *et al*., 2000] ou bancos locais usando o software HMMER [Durbin *et al*., 1998].

A última e mais desafiante parte da anotação está relacionada aos processos biológicos. Um marco no direcionamento deste nível foi o lançamento do consórcio GO para criar vocabulários descrevendo a princípio a função de genes eucariotos identificados em *Sacharomyces*, drosófila e rato [Ashburner *et al*., 2000]. Recentemente, outros organismos, como *A. thaliana*, *Zea mays* e *O. sativa* [Swarbreck *et al*., 2008] foram adicionados ao consórcio GO. A anotação em nível de processo se estende além de trabalhos puramente computacionais. Várias técnicas laboratoriais experimentais são utilizadas nesse nível para prover pistas vitais sobre o papel que genes e proteínas desempenham em um processo biológico. Em resumo, nesta fase a anotação de sequência começa a fundir-se com a tradicional pesquisa de bancada [Stein, 2001].

Até recentemente, a maioria dos AMPs foi determinada através de inferências proteômicas, após purificação a partir de extratos vegetais com posterior clonagem e avaliação dos seus respectivos genes [Odintsova e Egorov, 2007]. Entretanto, a disponibilidade crescente de genomas vegetais completamente sequenciados, bem como de ESTs, está mudando este cenário, especialmente considerando os recursos computacionais para comparação de genes e o reconhecimento de domínios específicos e assinaturas [Silverstein *et al*., 2007]. A quantidade de ESTs vegetais disponível permite o uso de métodos computacionais para a identificação de genes novos em taxa específicos ou em associação com determinados tecidos ou processos em plantas [Menossi *et al*., 2008].

Considerando a interação planta-micróbio em leguminosas, Fedorova *et al*. [2002], usando o formalismo Booleano, identificaram 340 prováveis genes específicos de nódulos radiculares. Surpreendentemente, 114 dos genes identificados codificavam pequenas Proteínas com agrupamentos de cisteínas (CCP - *Cys-Cluster Proteins*) com peptídeos-sinal cliváveis. Adicionalmente Mergaert *et al*. [2003] identificaram um número considerável de CCPs, mais de 300 NCRs no banco de dados MtGI (*Medicago truncatula Gene Índice*) usando a ferramenta BLAST e 19 sequências-sonda de famílias de plantas relacionadas.

Semelhantemente, Graham *et al*. [2004] usaram a ferramenta BLAST para comparar o conjunto de genes únicos de legumes com conjuntos de genes únicos de espécies não leguminosas e com sequências genômicas de *O. sativa* de *A. thaliana* e com o banco de dados não redundante de EST do GenBank. Posteriormente, os autores agruparam os prováveis genes específicos de legumes em suas famílias respectivas e avaliaram os prováveis motivos protéicos para cada familiar, ao compará-la com um banco de dados protéico, permitindo predições em relação a suas funções. Este procedimento permitiu a identificação de proteínas ricas em prolina (PRP – *Proline-Rich Proteins*) e em cisteína (CRP – *Cystein-Rich Protein*). Dentre estas, 300 CRP com similaridade com defensinas conhecidas foram expressas exclusivamente em nódulos radiculares e sementes, algumas delas relativas a quadros de leitura aberta desconhecidos no genoma de *A. thaliana* [Graham *et al*., 2004].

Após a identificação de defensinas nódulo-específicas por Graham *et al*. [2004], Silverstein *et al*. [2005, 2007] observaram que os genomas vegetais apresentam uma abundância inesperada de putativos AMPs ricos em cisteína com um potencial para adaptação funcional surpreendente. Usando uma estratégia baseada em HMM e buscas BLASTs, Silverstein *et al*. [2007] identificaram classes de CRPs contendo um peptídeo-sinal semelhante aos conhecidos em defensinas, thioninas, proteínas transferidoras de lipídios e

fatores de Alcanização rápida (RALF – *Rapid ALkanization Factor*), bem como genes hipotéticos codificantes de CRPs com assinaturas de cisteína não conhecidos em proteínas purificadas previamente. A estratégia usada por estes autores permitiu a identificação de 12. 824 sequências de CRPs distintas em 33 espécies de plantas, dando suporte a evidências anteriores da grande diversidade de CRPs no reino vegetal.

2.2.3. Modelagem comparativa das defensinas

A compreensão do mecanismo de função de uma proteína geralmente requer o conhecimento de sua estrutura tridimensional [Blundell *et al*., 1978; Weber, 1990], que em ultima instância é determinada por sua sequência de aminoácidos [Anfinsen, 1973]. Atualmente, existem cerca de dois milhões de sequências de proteínas no Swissprot e no TrEMBL [http://us.expasy.org/sprot/], entre as quais cerca de 50.000 proteínas tiveram suas estruturas resolvidos experimentalmente por métodos, tais como cristalografia de raios-X e espectroscopia de RMN [Johnson *et al*., 1994; http://www.rcsb.org/pdb/]. Esta enorme diferença entre o número de sequências disponíveis e as estruturas de proteínas experimentalmente obtidas poderia ser resolvida por métodos computacionais, como o método de modelagem por homologia, o qual obtém a estrutura tridimensional de uma dada sequência protéica baseada principalmente em sua similaridade de sequência com uma ou mais proteínas de estruturas conhecidas [Rost *et al*., 1996; Kolinski *et al*., 1999].

A predição de estrutura é um problema extremamente importante, simples de definir, mas difícil de resolver. Embora os métodos *ab initio* tenham alcançado progressos notáveis nos últimos anos, sua aplicação confiável ainda é pouco observada [Venclovas *et al*., 2003]. O gargalo é imposto principalmente pela imprecisão do campo de força e pela enorme e impraticável amostragem de conformações possíveis. A diferença de energia livre entre o enovelamento nativo e estados intermediários é de apenas alguns $kcal.mol^{-1}$, equivalente a várias interações atômicas de van der Waals, representando um desafio assustador para as funções de energia existentes. Além disso, a imprecisão do campo de força torna muito mais difícil a amostragem de conformações possíveis, uma vez que poderia orientar a minimização de energia para a conformação de um mínimo de energia global que seria totalmente diferente da verdadeira conformação nativa [Xiang *et al*., 1995; Wang *et al*., 1997].

Ao contrário, a modelagem por homologia tem assumido um papel cada vez mais importante na predição da estrutura de proteínas nos últimos anos, principalmente com o

advento de iniciativas de genômica estrutural em todo o mundo. Isso ocorre porque muitas sequências de proteínas são evolutivamente relacionadas e, portanto, podem ser classificadas em diferentes famílias. Proteínas das mesmas famílias frequentemente têm semelhanças perceptíveis e, portanto, compartilham de arquitetura tridimensional semelhante, permitindo uma descrição estrutural de todas as proteínas de uma família, mesmo quando apenas a estrutura de um único membro é conhecida. Esta relação evolutiva fornece os fundamentos para a genômica estrutural, um esforço sistemático de larga escala para a caracterização estrutural de todas as proteínas, onde uma proteína representante de cada família é escolhida para ser resolvida experimentalmente com o resto confiantemente previsto por um método de modelagem por homologia [Goldsmith-Fischman e Honing, 2003].

De fato, se assumimos que a estrutura da proteína é de um mínimo de energia global, a modelagem por homologia é simplesmente uma questão de vasculhar o espaço de conformações possíveis perturbando minimamente as soluções existentes, ou seja, as estruturas experimentalmente resolvidas. A vantagem óbvia é que a técnica de modelagem por homologia minimiza a exigência rigorosa do campo de força e a enorme procura de conformação, pois dispensa o cálculo de um campo de forças físico-químicas e o substitui, em grande parte, por uma medida de identidade entre sequências [Sanchez e Sali, 1997]. Dada uma sequência de proteínas, a modelagem por homologia consiste geralmente das seguintes etapas: (i) identificar o homólogo com estrutura conhecida a partir do Banco de dados protéico PDB; ii) alinhar a sequência alvo com o modelo de estrutura; (iii) construir o modelo com base no alinhamento; (iv) avaliar e aperfeiçoar o modelo [Sanchez e Sali, 1997; Fiser e Sali, 2003].

Quando a identidade de sequência é superior a 40%, o alinhamento ocorre sem maiores problemas, não há muitas lacunas, e 90% dos átomos da cadeia principal podem ser modelados com um erro calculado pela distância média quadrática (RMSD - *Root Mean Square Distance*) de cerca de 1 Å [Sanchez e Sali, 1997]. Quando a identidade de sequência é de cerca de 30-40%, obter o alinhamento correto torna-se difícil, sendo as inserções e deleções frequentes, onde 80% de átomos da cadeia principal podem ser previsto a uma RMSD de 3,5 Å. Quando a similaridade da sequência é inferior a 30%, o principal problema torna-se a identificação das estruturas homólogas e o alinhamento torna-se muito mais difícil [Harrison *et al.*, 1995; Mosimann *et al.*, 1995; Sauder *et al.*, 2000; Yang e Honing, 2000].

A predição da estrutura de proteínas foi um sonho para a comunidade científica durante décadas, não apenas para os químicos computacionais, como também para os físicos,

matemáticos e cientistas da computação. Recentemente essa tarefa tem obtido grande êxito com a explosão de sequências e informações estruturais, bem como com os avanços computacionais em diversas áreas, incluindo a análise de sequências e uma melhor compreensão dos determinantes energético da estabilidade protéica, possibilitando o desenvolvimento de uma série de novos métodos computacionais destinados a detectar novas relações entre sequência, estrutura e função. O progresso em modelagem *ab initio* possibilitará o aperfeiçoamento de modelos obtidos por homologia com maior precisão, servindo de base para uma análise mais detalhada das relações estrutura/função em comparação com os métodos disponíveis no passado e fornecendo ferramentas poderosas para a análise de dados experimentais e para a concepção de novos experimentos [Fiser e Sali, 2003].

Com relação às defensinas, até o momento oito estruturas tridimensionais (3D) de defensinas vegetais foram estudadas em detalhe por meio de cristalografia de raios-X e NMR, incluindo as defensinas g1-H de *Hordeum vulgare* e g1-P de *Triticum aestivum* [Bruix *et al.*, 1993], Rs-AFP1 de *R. sativus* [Fant *et al.*, 1998], Ah-AMP1 de *Aesculus hippocastanum* [Fant *et al.*, 1999], PsD1 de *Pisum sativum* [Almeida *et al.*, 2002], PhD1 de *Petunia hybrida* [Jansen et al., 2003], NaD1 de *Nicotiana alata* [Lay *et al.*, 2003] e VrD1 de *V. radiata* [Liu *et al.*, 2006], melhorando consideravelmente o conhecimento sobre a relação estrutura/função desses peptídeos. A predição da estrutura de defensinas vegetais através de modelagem comparativa foi utilizada para caracterizar a estrutura 3D de outras defensinas, tais como Vv-AMP1 de *Vitis vinifera* [Beer e Vivier, 2008], Cp-thioninI e Cp-thioninII de *Vigna unguiculata* [Melo *et al.*, 2002; Franco *et al.*, 2006], VrD1 de *V. radiata* [Shiau *et al.*, 2006] e Gbd de *Ginkgo biloba* [Shen *et al.*, 2005].

Todas as defensinas vegetais estudadas compartilham uma estrutura comum, também presente em defensinas de insetos e em algumas toxinas de escorpião [Cornet *et al.*, 1995; Krezel *et al.*, 1995], que embora relacionadas e compartilharem um cerne estrutural, apresentam grande variação funcional, levantando a questão de como elas atuam na defesa do organismo, dado que todas elas têm a mesma estrutura mas mostram diferentes funções. Conforme discorrido, a análise estrutural comparativa pode auxiliar no entendimento de sua estrutura 3D e elucidar o mecanismo por trás de suas funções protéicas, como foi demonstrado para a defensina VrD1 [Liu *et al.*, 2006; Shiau *et al.*, 2006].

Nesse sentido, a estrutura 3D da defensina Cp-thioninII com atividade antibacteriana foi modelada com base na estrutura resolvida do inibidor de α-amilase VrD1, que apresentaram

73% de identidade entre si [Franco *et al.*, 2006]. Os autores também investigaram a formação de multimeros de Cp-thioninII, visto que anteriormente foi demonstrado a formação de multimeros da defensina Cp-thioninI, bem como de dímeros de defensinas de *Pachyrrhizus erosus* [Melo *et al.*, 2002; Song *et al.*, 2005]. Dímeros de defensinas de mamíferos foram frequentemente relacionados à capacidade de inibir o crescimento bacteriano [Schibli *et al.*, 2002]. Mas essa relação mecanística parece não ocorrer em defensinas vegetais com atividade antibacteriana dado que a Cp-thioninII é observada apenas como monômeros, porém com atividade efetiva contra bactérias [Franco *et al.*, 2006].

A avaliação da RMSD da estrutura de várias defensinas em relação à estrutura de Cp-thioninII demonstrou que todas elas são basicamente idênticas, indicando uma baixa variação ao nível estrutural terciário [Murad *et al.*, 2007]. Um exame profundo das estruturas obtidas experimentalmente e por modelagem computacional logo denota a presença de duas regiões variáveis localizadas nos loops 1 e 2, provavelmente de grande importância para algumas das funções específicas das defensinas, tais como antifúngica e antibacteriana [Fant *et al.*, 1998; Almeida *et al.*, 2002; Franco *et al.*, 2006], bloqueador de canais iônicos [Shiau *et al.*, 2006] e inibidor de α-amilase e proteinase [Melo *et al.*, 2002; Liu *et al.*, 2006; Pelegrini *et al.*, 2008].

Por outro lado, a carga geral da superfície poderia auxiliar no entendimento e distinção funcional das defensinas. Entretanto, mais uma vez os estudos estruturais comparativos demonstram que é impossível delinear um padrão estrutura/função, uma vez que quase todas as defensinas examinadas apresentam as mesmas características em relação as cargas de superfície [Murad *et al.*, 2007].

A modelagem comparativa tem auxiliado também no descobrimento de outros processos nos quais as defensinas tomam parte, bem como no esclarecimento de questões evolutivas entre defensinas vegetais e de outros reinos [Doughty *et al.*, 1998; Zhu, 2007; Zhu, 2008; Gao *et al.*, 2009]. Por exemplo, a estrutura da proteína PCP-A1, envolvida no sistema de autoincompatibilidade de espécies da família Brassicaceae, foi obtida por modelagem comparativa, observando-se ser estruturalmente similar às defensinas γ1-P, γ1-H e Rs-AFP1, porém todas as outras sequências consenso observadas entre as defensinas vegetais com atividade antimicrobiana não estão presentes [Doughty *et al.*, 1998].

A modelagem de algumas famílias de defensinas fúngicas denotou a presença de defensinas conservadas em plantas, fungos e animais, bem como defensinas ancestrais e clássicas presentes em animais invertebrados e fungos, sendo possível propor uma história evolutiva para as defensinas [Zhu, 2008]. Análises de bioinformática estrutural das sequencias

de microorganismos possibilitaram a identificação de um peptídeo bacteriano AdDLP com similaridade estrutural com defensinas eucarióticas através de modelagem comparativa com a defensina fúngica Plectasin pertencente à classe de defensinas ancestrais de invertebrados [Zhu, 2007; Zhu, 2008]. Porém, predição *ab initio* da estrutura 3D de AdDLP demonstra uma estrutura mais similar às defensinas vegetais [Gao *et al*., 2009].

2.2.4. Análise evolutiva

Devido à riqueza de dados depositados nos bancos de dados e por meio de poderosos algoritmos de alinhamento de sequências e de reconstrução das relações evolutivas entre elas, torna-se cada vez mais interessante prever funções de proteínas em uma abordagem comparativa. Esta abordagem baseia-se essencialmente na observação de que proteínas homólogas retêm os aspectos da sua função durante longos períodos evolutivos [Pellegrini *et al*., 1999]. Isto permite uma predição de função com base na homologia, ou seja, a transferência de conhecimentos sobre a função entre proteínas homólogas. No entanto, apesar de homologia com frequência implicar uma preservação dos aspectos mecânicos funcionais, ela fornece poucas informações sobre o contexto funcional, isto é, sobre os processos nos quais uma proteína está envolvida, os mecanismos moleculares ou a natureza da associação funcional [Huynen *et al*., 2000]. Por exemplo, enzimas homólogas podem catalisar reações semelhantes, mas os substratos e produtos envolvidos na reação podem participar de vias diferentes. As conexões previstas por métodos baseados no contexto, como perfis filogenéticos, não são muito específicas e geralmente não fornecem informações sobre o papel exato da proteína em um processo, mas frequentemente são decisivas para orientar uma análise mais aprofundada. Por exemplo, a abordagem comparativa pode melhorar a qualidade dos dados considerando a conservação evolutiva da co-expressão gênica [Stuart *et al*., 2003; Van Noort *et al*., 2003; Bergmann *et al*., 2004; Snel *et al*., 2004].

Uma série de previsões sobre vários processos celulares têm sido feitas através de perfis filogenéticos e confirmados experimentalmente [Makarova *et al*., 2003; Sato *et al*., 2004]. Por exemplo, perfis filogenéticos identificaram enzimas da via MPE/DOXP, que nos cloroplastos das plantas, Apicomplexa, cianobactérias e inúmeras outras bactérias produzem os blocos constituintes de isoprenóides. Cunningham *et al*. [2000] determinaram os perfis de ocorrência das primeiras cinco enzimas conhecidas na via MPE/DOXP e encontraram duas outras proteínas que co-ocorreriam com a via, LytB e GcpE. Embora a participação da proteína LytB

na via tenha sido apoiada por evidências experimentais, eventualmente outras experiências genéticas foram necessárias para determinar as posições exatas de LytB e GcpE na via [Altincicek *et al.* 2001a, Altincicek *et al.* 2001b]. Geralmente, os exemplos mostram que o poder da caracterização filogenética deriva mais de uma abordagem em larga escala, que permite usá-la como um método exploratório e relativamente fácil de implementar. Após ter reduzido o espaço de busca a um número razoável de candidatos, linhas de evidências adicionais são necessárias. Estas podem incluir os resultados de outros métodos em um contexto genômico, como conservação de ordem genética, de experimentos em larga escala publicados ou a partir de estudos em pequena escala. Neste processo de descoberta de bons candidatos para novas experiências, todas estas diferentes fontes devem ser consideradas e a análise filogenética é uma delas [Dandekar *et al.*, 1998].

Semelhante a relacionar genes entre si através de sua co-ocorrência nos genomas, relacionar genes com fenótipos implica em associá-los a um processo biológico. No entanto, isso não requer conhecimento prévio sobre o envolvimento de outros genes no processo. Perfis fenótipo/genótipo ilustram o princípio da caracterização filogenética a partir de uma nova perspectiva, relacionando o ambiente de um organismo com sua evolução molecular [Jim *et al.*, 2004]. O ambiente e sua mudança são importantes fatores que influenciam a evolução, pois induzem adaptações mais ou menos específicas no organismo [Copley 2000; Pretzer *et al.*, 2005; Ternes *et al.*, 2006]. Isto sugere que a diversidade, a especificidade e a estrutura de correlação de fatores ambientais ativos em um ramo da filogenia irão determinar o valor das perdas e ganhos coordenados de genes neste ramo para traçar seu perfil filogenético e que, ao aumentar a resolução da árvore filogenética também é aumentada a resolução em termos de fatores seletivos. Esta pode ser uma razão para a melhora observada nos resultados de perfis filogenéticos tanto pelo aumento do número de espécies como por uma escolha equilibrada de espécies [Sun *et al.*, 2005].

A análise filogenética é um método versátil para a predição de interações funcionais, mas a sua cobertura é limitada aos sistemas celulares que evoluíram de forma modular [Campillos *et al.*, 2006]. Entretanto, basicamente todo método experimental e computacional apresenta limitações à medida que os programas consideram apenas um ou alguns aspectos da associação funcional. Inúmeros exemplos mostram que a evidência de evolução para uma interação funcional com base em perfis filogenéticos, bem como em outros métodos no contexto genômico, frequentemente não é forte nem específica o suficiente para assegurar a função exata de uma proteína. Por conseguinte, em praticamente todos os exemplos,

evidências experimentais adicionais foram utilizados para reforçar as conclusões obtidas a partir de perfis filogenéticos e para compreender a função da proteína em um nível muito mais detalhado [Altincicek *et al*. 2001a, Altincicek *et al*. 2001b; Makarova *et al*., 2003; Sato *et al*., 2004]. Portanto, a necessidade de combinar várias linhas de evidência se aplica aos perfis filogenéticos usados como um método que prediz em larga escala interações funcionais para um grande número de pares de genes. Assim, a análise filogenética tem sido integrada com outras fontes de dados, a fim de obter um retrato confiante da função específica de uma proteína [Dandekar *et al*., 1998; Snel *et al*., 2002; Spirin *et al*., 2006].

2.2.5. Perfil de expressão digital

Investigações com o objetivo de decifrar os eventos moleculares envolvido no início e progressão de processos biológicos como, por exemplo, a resistência a doenças em plantas ou o desenvolvimento de doenças em humanos, buscam principalmente caracterizar biomoléculas cuja expressão diferencial contribui para alterações na função celular, levando a um ajustamento ou a um estado patológico. Focalizando nos mecanismos de resistência, os cientistas pretendem identificar eventos críticos que podem ser alvo de melhoramento ou transfecção para culturas agrícolas importantes. Em humanos, as pesquisas objetivam identificar eventos moleculares críticos que podem ser alvo de novas estratégias terapêuticas. Características complexas, como resistência e doença, normalmente são poligênicas e estudos de uma única molécula não promoverão a percepção de uma resposta celular orquestrada conforme ela evolui em um tecido afetado. Visualizando a resposta completa, o investigador começa a entender a inter-relação complexa entre biomoléculas, que contribuem para mudanças no fenótipo da célula, conduzindo à resistência ou à doença. Assim, um ponto de partida fundamental nos estudos sobre os mecanismos de resistência/doença é decidir como identificar as biomoléculas associadas [Murray *et al*., 2007].

Técnicas capazes de quantificar a expressão gênica promovem o desenvolvimento de nossa compreensão sobre a distribuição e regulação de produtos gênicos em tipos celulares normais e anormais. Tais técnicas incluem uma variedade de microarranjos e análise serial da expressão gênica (SAGE – *Serial Analysis of Gene Expression*), das quais todas têm a capacidade de rápida e eficazmente examinar a expressão dos transcritos de todo um genoma [Velculescu *et al*., 1995; Lohar *et al*., 2006; Molina *et al*., 2008]. Adicionalmente, é possível explorar metodologias computacionais que traçam um perfil de expressão de todos os genes

de um modo quantitativo e direto e não apenas de genes conhecidos em chips de DNA. A disponibilidade de vastas quantidades de dados de sequência juntamente com os avanços em biologia computacional provém uma estrutura ideal para a análise *in silico* da expressão gênica. Todos estes avanços são dirigidos por estratégias computacionais em consonância com a disponibilidade de dados, claramente objetivando identificar padrões biologicamente pertinentes nos dados [Audic e Claverie, 1997].

Uma área chave desta pesquisa envolve a definição da população de genes diferencialmente expressos em um tecido doente ou em modelos do processo de doença. O conhecimento da identidade de tais transcritos fornece um ponto de partida útil na busca por eventos moleculares críticos que contribuem para a resistência ou doença. Para o biólogo moderno, há numerosas estratégias computacionais que podem ser empregadas para analisar a expressão de genes. EST e bibliotecas de SAGE são uma fonte ideal para traçar um perfil de expressão, pois em principio tanto a frequência do clone de EST quanto da etiqueta de SAGE são proporcionais ao nível de expressão do gene correspondente em um determinado tecido, permitindo comparações estatisticamente importantes do nível de expressão entre duas populações de cDNA [Adams *et al*., 1991; Adams *et al*., 1993; Lal *et al*., 1999; Schmitt *et al*., 1999].

Para explorar esta grande quantidade de informação foram desenvolvidos algoritmos computacionais para a descoberta de ambos os genes novos [Vasmatzis *et al*., 1998] e genes com distribuição limitada a um tecido e/ou de expressão específica sob dado estresse [Brinkmann *et al*., 1998]. Por exemplo, SAGEmap é uma ferramenta usuário-amigável mantida pelo NCBI especialmente projetada para interpretar dados de SAGE. Embora sua interface web tenha sido descontinuada em.25 de setembro de 2007, é possível acessar o mapeamento para algumas espécies através do servidor ftp do NCBI disponível em *ftp://ftp.ncbi.nlm.nih.gov/pub/sage/mappings* [Vasmatzis *et al*., 1998]. cDNAxProfiler é outra ferramenta dedicada a humanos e ratos, que compara expressão de genes entre duas amostras de bibliotecas, onde cada amostra pode ser uma única biblioteca ou um agrupamento de várias bibliotecas [http://cgap.nci.nih.gov/Tissues/xProfiler]. O expositor digital da expressão gênica (DGED – *Digital Gene Expression Displayer*) é uma ferramenta para a comparação da expressão gênica entre duas amostras de bibliotecas, podendo ser usado para comparar bibliotecas de cDNA ou bibliotecas de SAGE [http://cgap.nci.nih.gov/Tissues/GXS].

Todos os eventos biológicos na célula são governados principalmente pelas mudanças na expressão de genes fundamentais. A habilidade de uma célula de regular a expressão gênica

ao longo do tempo dirige toda função e atividade biológica. Portanto, muito interesse é destinado a delinear um perfil de expressão gênico para identificar os genes-chave e grupos de genes-chave cuja expressão é alterada em certos estados. O desenvolvimento de novas metodologias experimentais e analíticas tem redefinido de várias maneiras as ferramentas que os biólogos possuem para acessar e analisar tal nível de informação molecular. Comparando o perfil de expressão gênica sob condições diferentes, podem ser identificados genes individuais ou grupos de genes que desempenham um papel fundamental em cascatas de sinalizações ou em determinado processo celular ou ainda na etiologia de uma doença. Ademais, o perfil de expressão também é importante para compreender a função gênica e identificar alvos terapêuticos.

2.3. Cana-de-açúcar

A cana-de-açúcar (*Saccharum* spp.) pertence à família Poaceae, sendo mais cultivadas nos trópicos e sub-trópicos devido ao seu colmo peculiar com acumulação bastante alta de sacarose. Esta família é uma das mais importantes para a economia mundial, oferecendo várias fontes de forragem para o gado e grãos para a alimentação humana. A família Poaceae também tem contribuído com vários organismos-modelo para o entendimento da genética de plantas superiores, a exemplo do milho, arroz, sorgo e trigo, cujos genomas foram intensamente investigados por meio de múltiplas técnicas genéticas e moleculares [http://www.gramene.org/]. Dentro deste importante grupo vegetal, a cana-de-açúcar se destaca pelo seu uso na obtenção de açúcar e combustível, sendo o Brasil o maior produtor mundial destes derivados, exercendo uma forte influência sobre a variação dos preços no mercado internacional de açúcar e de energia limpa e renovável [CONAB, 2009; http://www.unica.com.br/].

Até o século XIX, a maioria das variedades comerciais provinha de cruzamentos entre *S. officinarum*, *S. barberi* e *S. sinensis*, sendo que os primeiros híbridos interespecíficos artificiais foram produzidos na Índia a partir especialmente de *S. officinarum*, *S. spontaneum* e *S. barberi.* As variedades comerciais atuais resultam do cruzamento dessas espécies, acrescida da contribuição do genoma de *S. robustum* [Daniels e Roach, 1987]. No Brasil, o desenvolvimento de variedades comerciais visou ao aumento da produtividade, através da obtenção de resistência a pragas e doenças, bem como da adaptação aos climas e solos

regionais brasileiros, além do melhoramento destinado a técnicas de corte e manejo [Galvão *et al*., 2005].

2.3.1. Características genômicas da cana-de-açúcar

Os cultivares modernos de cana-de-açúcar são híbridos derivados do cruzamento de *S. officinarum* e *S. spontaneum*, que normalmente apresentam o número cromossômico 2n = 80 e 2n = 40 a 2n = 128, respectivamente. Tratam-se de espécies poliplóides, cujos genomas são respectivamente múltiplos do número básico X = 10 e X = 8 cromossomos [Grivet e Arruda, 2001]. O número cromossômico pode variar entre os cultivares modernos. Devido a diferenças estruturais entre cromossomos das duas espécies, os híbridos possuem proporções diferentes de cromossomos, conjunto cromossômico variável e eventos complexos de recombinação. Os híbridos são altamente poliplóides e aneuplóides, contendo em média 2n = 100-120 cromossomos com um tamanho nas células somáticas estimado em 10.000 Mbp [D'Hont, 2005]. O tamanho do genoma básico de cana-de-açúcar varia entre 760 Mbp e 960 Mbp, o que é duas vezes o tamanho do genoma do arroz (389 Mbp) e semelhante ao tamanho do genoma do sorgo (760 Mbp) [D'Hont e Glaszman, 2001].

Devido a essa complexidade, o genoma da cana-de-açúcar recebeu no passado relativamente pouca atenção e investimentos por parte dos cientistas vegetais no desenvolvimento de ferramentas biotecnológicas e genéticas para essa cultura. Contudo, abordagens de citogenética molecular e o uso de marcadores moleculares levaram a avanços significativos no estabelecimento da origem e da estrutura genômica da cana-de-açúcar [Grivet e Arruda, 2001].

2.3.2. Importância econômica da cana-de-açúcar

Produzida em mais de 70 países, a cana-de-açúcar desempenha um papel importante na economia nacional desde a sua introdução, sendo a principal fonte de renda para o Brasil, especialmente em relação às regiões Nordeste e Centro-sul. Os números parciais obtidos para a safra 2009 sinalizam um aumento nas divisas geradas em comparação aos números obtidos no mesmo período em 2008, uma tendência que vem sendo observada nos últimos anos [CONAB, 2009; http://www.unica.com.br/]. As estimativas para a safra 2010 indicam que esta expansão se manterá, sendo um reflexo da expansão nas áreas de cultivo e do

melhoramento das tecnologias aplicadas à cana-de-açúcar, decorrente da demanda mundial por açúcar e combustível. Porém, houve uma mudança no destino do caldo da cana que foi direcionado para a produção de açúcar em detrimento da produção de álcool, refletindo os preços atuais do açúcar no mercado mundial. A produção de açúcar cresceu 16% esse ano, enquanto a produção de álcool cresceu 4,79 %. Contudo, estatísticas provenientes da indústria automobilística nacional apontam um mercado consumidor de álcool hidratado em grande expansão, sendo os carros "*Flex-Fuel*" responsáveis por 92% das vendas de automóveis, com uma frota atual de 9, 4 milhões veículos [CONAB, 2010; http://www.unica.com.br/].

2.3.3. O banco de dados SUCEST

Nos últimos anos, grandes coleções de EST se tornaram disponíveis para explorar o grande genoma alopoliplóide da cana-de-açúcar, renovando consequentemente o interesse dos pesquisadores na genética desse vegetal [Butterfield *et al*., 2001; Grivet e Arruda, 2001; Ming *et al*., 2006], provendo um meio para a descoberta e identificação de novos genes, o que é essencial para programas de melhoramento, tanto para o desenvolvimento de plantas transgênicas como para o melhoramento assistido por marcadores. Somente o projeto de EST da cana-de-açúcar (SUCEST – *Sugarcane EST Project*) gerou 237.954 ESTs que foram agrupadas em 43.141 prováveis transcritos únicos de cana-de-açúcar, originados dos principais órgãos que foram coletados em vários estágios do desenvolvimento [Vettore *et al*., 2001; http://sucest.lad.ic.unicamp.br/en/]. Esse banco de EST representa um dos mais ricos em genes, apresentando uma variedade de tecidos que foram amostrados a partir de cultivares diferentes, capacitando o acesso à informação transcricional de genes expressos em vários sistemas biológicos, enquanto a maioria dos projetos de ESTs vegetais enfatizou situações contrastantes e amostrou a planta inteira ou determinados tecidos sob condições de estresse. Desse modo, o banco de dados SUCEST é um excelente modelo para a análise da expressão diferencial de genes.

2.4. Feijão-Caupi

O feijão-caupi (*Vigna unguiculata* (L.) Walp.) pertence à família Fabaceae, sendo, semelhantemente à cana-de-açucar, mais cultivado nos trópicos e sub-trópicos principalmente devido ao seu conteúdo nutricional, adaptabilidade e versatilidade [Ehlers e Hall, 1997]. Esta

família é uma das mais importantes para a economia mundial, oferecendo juntamente com os cereais (Poaceae) várias fontes de forragem para o gado e grãos para a alimentação humana. A família Fabaceae também tem contribuído com vários organismos-modelo para o entendimento da genética de processos vegetais como a fixação de nitrogênio, a interação planta/microorganismo, bem como do processo de adaptação a estresses abióticos, a exemplo da soja (*Glycine max*) e da erva-médica (*Medicago truncatula*) [Gepts *et al.*, 2005; Sato *et al.*, 2007]. Dentro deste importante grupo vegetal, o feijão-caupi se destaca por sua rusticidade além de sua proximidade evolutiva com outras leguminosas economicamente importantes tais como a própria soja e o feijão comum (*Phaseolus vulgaris*), servindo como uma espécie-modelo de adaptação de plantas domesticadas a estresses bióticos e abióticos [Gepts *et al.*, 2005].

Acredita-se que as principais variedades de feijão-caupi foram introduzidas pelos colonizadores e escravos que entravam no Brasil principalmente pelos portos de Recife, Salvador, Rio de janeiro e São Luiz, trazendo consigo muito de sua cultura bem como sementes de plantas usadas em seus pratos favoritos ou em rituais [Curtin, 1969 apud Simon *et al.*, 2007]. O desenvolvimento de variedades comerciais visou principalmente ao aumento da produtividade com cultivares que apresentassem estabilidade produtiva em ambientes variados, através da obtenção de resistência a nematóides, pragas e doenças, bem como da adaptação a fatores abióticos, além do melhoramento da arquitetura da planta para facilitar o plantio e manejo [EMBRAPA; Singh *et al.*, 2002].

2.4.1. Caracteristicas genômicas do feijão-caupi

A origem do feijão-caupi é controversa, havendo grande divergência entre os pesquisadores não somente quanto a sua origem como também quanto ao inicio de sua domesticação. Algumas evidências apontam a África como provável centro de origem e domesticação, uma vez que se observa um maior número de espécies do gênero, bem como de espécies endêmicas nesse continente. Contudo, outras evidências sugerem que a planta pode ter se originado na América do Sul ou na Ásia [Freire Filho, 1988; Magloire, 2005; Simon *et al.*, 2007].

Vigna unguiculata é uma espécie diplóide com número cromossômico 2n=22, apresentando genoma com um tamanho nas células somáticas estimado em 450-500 Mbp, o que é aproximadamente similar ao genoma do feijão comum e um terço do genoma do

amendoim (*Arachis hypogaea* L.), representando um dos menores genomas entre as leguminosas [Pignone *et al*., 1990; Benko-Iseppon, 2001].

Apesar dessa simplicidade e da grande importância e de sua importância econômica e social em paises em desenvolvimento, o genoma do feijão-caupi recebeu relativamente pouca atenção e investimentos em relação a pesquisas, permanecendo até certo ponto uma cultura pouco explorada [Nelson *et al*., 2004]. Porém, a transposição de ferramentas biotecnológicas e genéticas tradicionalmente utilizadas para a pesquisa em sistemas-modelo levaram a avanços significativos no estabelecimento da origem e da estrutura gênica e genômica do feijão caupi nos últimos cinco anos [Timko *et al*., 2008]. Atualmente existem 189.139 sequências de EST em bancos de dados públicos [http://www.ncbi.nlm.nih.gov] e aproximadamente 35.000 novas ESTs estão sendo produzidas pelo projeto NordEST da REde NORdestina de BIOtecnologia (RENORBIO) [Benko-Iseppon, Comunicação pessoal].

2.4.2. Importância econômica do feijão-caupi

O feijão-caupi, feijão-de-corda ou feijão-macassar (*Vigna unguiculata*) representa uma das principais culturas alimentares com atributos desejáveis da população de baixa renda do nordeste brasileiro e oeste africano [Onwuluri e Obu, 2002], constituindo uma excelente fonte de proteínas (23-25% em média), apresentando todos os aminoácidos essenciais, exceto pela deficiência de sulfurados, carboidratos (62%, em média), vitaminas e minerais, além de possuir grande quantidade de fibras dietéticas, baixa quantidade de gordura (teor de óleo de 2%, em média) e não conter colesterol e apresentar baixa atividade inibitória de tripisina [Frota *et al*., 2008]. Além de seu grande potencial alimentar, a cultura do feijão-caupi apresenta ciclo curto, baixa exigência hídrica e rusticidade para se desenvolver e manter a produtividade em face de estresses abióticos como seca, altas temperaturas, solos de baixa fertilidade, tendo a habilidade de fixar nitrogênio por meio da simbiose com bactérias do gênero *Rhizobium*, propriedades que tornam seu plantio uma cultura interessante para a subsistência em áreas semi-áridas do mundo, onde pode inclusive ser associado a rotações de culturas de cereais e tubérculos, uma vez que tem o potencial de aumentar a fertilidade do solo [Ellis *et al*., 1994; Sanginga *et al*., 2003; Hall, 2004].

Pelo seu valor nutritivo, o feijão-caupi é cultivado principalmente para a produção de grãos, secos ou verdes, visando o consumo humano in natura, na forma de conserva ou desidratado. Além disso, o caupi também é utilizado como forragem verde, feno, ensilagem,

farinha para alimentação animal e, ainda, como adubação verde e proteção do solo [EMBRAPA]. A área ocupada com feijão-caupi, no mundo, está em torno de 12,5 milhões de ha, com 8 milhões (64% da área mundial) na parte oeste e central da África. A outra parte da área está localizada na América do Sul, América Central e Ásia, com pequenas áreas espalhadas pelo sudoeste da Europa, sudoeste dos Estados Unidos e da Oceania. Entre todos os países, os principais produtores mundiais são Nigéria, Niger e Brasil [Ehlers e Hall, 1997; http://faostat.fao.org], sendo o nordeste brasileiro responsável por grande parte da produção nacional, com uma produtividade de 303,5 kg/ha [EMBRAPA meio-norte].

2.4.3. O banco NordEST

Nos últimos anos, o uso de marcadores moleculares foram paulatinamente empregados na caracterização genômica do feijão-caupi e na determinação de sua origem [Simon *et al*., 2007; Wang *et al*., 2008; Muchero *et al*., 2009]. Coleções de EST para explorar o genoma diplóide do feijão-caupi se tornaram disponíveis, impulsionando o interesse dos pesquisadores na genética desse vegetal [http://www.ncbi.nlm.nih.gov], provendo um meio para a descoberta e identificação de novos genes e do modo como o feijão-caupi responde e se adapta as principais situações ambientais que limitam seu crescimento e produtividade. Somente o projeto de EST do feijão-caupi no âmbito da RENORBIO gerou 500.000 ESTs, incluindo 12 bibliotecas de EST sob condições de estresse biótico (vírus do mosaico severo e potyvirus) e abiótico (salinidade), além de cinco bibliotecas de SuperSAGE para o vírus do mosaico severo e duas bibliotecas de LongSAGE para o potyvirus. Esse banco de dados representa um dos mais elaborados para a espécie, representando tecidos que foram amostrados a partir de diferentes condições, capacitando o acesso à informação transcricional de genes expressos em vários momentos durante a resposta do vegetal frente aos desafios experimentalmente aplicados [Benko-Isepon *et al*., 2009].

3. Objetivos

Visto que os estudos indicam a ubiquidade de defensinas no reino vegetal, pretende-se desenvolver uma rotina computacional para identificar e caracterizar defensinas em bancos de ESTs vegetais utilizando técnicas computacionais aplicadas à genômica, transcriptômica, proteômica, bioinformática estrutural e filogenética.

3.1. Objetivos específicos

- Catalogar sequências aminoacídicas de defensinas vegetais em bancos de dados biológicos públicos;
- Avaliar a estrutura gênica, identificando regiões codificantes, domínios e motivos conservados em defensinas vegetais;
- Buscar sequências similares em um banco de ESTs vegetais;
- Realizar estudo filogenético;
- Caracterizar as estruturas primária, secundária e terciária das defensinas;
- Delinear o perfil de expressão *in silico*;
- Aplicar a rotina no transcriptoma da cana-de-açúcar.

4. Perspectivas

Os estudos recentes indicam que as defensinas provavelmente estão presentes em todos os genomas vegetais, possuindo uma incrível capacidade multifuncional provavelmente decorrente de sua plasticidade primária com estrita conservação terciária. Essas características possivelmente levaram as defensinas a desempenhar um papel em variados e importantes processos biológicos, destacando-se entre eles a defesa vegetal. Não obstante, dada suas características, as defensinas são raramente identificadas em sua totalidade nos genomas, principalmente os genomas vegetais que aparentemente carregam uma grande quantidade de peptídeos semelhantes às defensinas com um grande potencial de adaptação funcional. Este trabalho baseou-se no conhecimento obtido sobre o nível primário, secundário e terciário das defensinas vegetais, bem com em outras fontes de informações como perfil diferencial de expressão e história filogenética para desenvolver uma rotina prática de mineração de defensinas nos genomas vegetais. Inicialmente, a rotina se mostrou eficiente e aplicável, identificando 17 defensinas no transcriptoma da cana-de-áçúcar. Com estes peptídeos foi possível obter a estrutura tridimensional das defensinas de cana-de-açúcar por modelagem comparativa, evidenciando um enovelamento altamente conservado, porém com algumas diferenças observadas nas regiões de *turn* e *loop*, indicando que múltiplas classes de defensinas estão potencialmente envolvidas na defesa deste organismo e provavelmente participam em diferentes processos moleculares. Ainda com relação às defensinas identificadas, estas podem ser usadas em manipulações genéticas, podendo se transformar em

ferramentas na produção de novos compostos ativos contra patógenos e pestes da cultura da cana-de-açúcar, bem como na transformação genética de plantas resistentes. Alternativamente, a estrutura desses peptídeos pode ser explorada para estudo ou mesmo para o melhoramento da eficácia de sua ação antimicrobiana através de ferramentas biotecnológicas, podendo ser utilizada para o desenho racional de fármacos. Consequentemente, a rotina é aplicável e transponível para todo e qualquer genoma vegetal, representando uma primeira linha de evidência para a escolha de peptídeos para estudos mais detalhados e desenvolvimento de produtos agrobiotecnológicos e farmacêuticos de origem vegetal. Com relação a amplificação de defensinas, foi possível desenvolver um primer heterólogo capaz de amplificar eficazmente um região gênica, cuja análise da esturua e padrão de processamento se mostrou em concordância com o verificado para outros vegetais. A partir do estudo realizado pretende-se extender a pesquisa para outros genomas vegetais e adicionalmente amplificar membros dessa família gênica em plantas nativas da região nordeste que apresenta uma grande diversidade de plantas de uso medicinal.

Referências bibliográficas

ADAMS M D; KELLEY J M; GOCAYNE J D; DUBNICK M; POLYMEROPOULOS M H; XIAO H; MERRIL C R; WU A; OLDE B; MORENO R F. Complementary DNA sequencing: expressed sequence tags and human genome project. **Science.** v. 252, p. 1651-1656, 1991.

ADAMS M D; KERLAVAGE A R; FIELDS C; VENTER J C. **3**,400 new expressed sequence tags identify diversity of transcripts in human brain. **Nat. Genet.** v. 4, p. 256-267, 1993.

AERTS A M; FRANÇOIS I E J A; MEERT E M K; LI Q; CAMMUE B P A; THEVISSEN K. The antifungal activity of RsAFP2, a plant defensin from *Raphanus sativus*, involves the induction of reactive oxygen species in *Candida albicans*. **J. Mol. Microbiol. Biotechnol.** v. 13, p. 243–247, 2007.

ALMEIDA M S; CABRAL K M S; MEDEIROS L N; VALENTE A P; ALMEIDA F C L; KURTENBACH E. cDNA cloning and heterologous expression of functional Cys-rich antifungal protein Psd1 in the yeast *Pichia pastoris*. **Arch. Biochem. Biophys.** v. 395, p. 199-207, 2001.

ALMEIDA M S; CABRAL K M S; ZINGALI R B; KURTENBACH E; Characterization of two novel defensin peptides from pea (*Pisum sativum*) seeds. **Arch. Biochem. Biophys.** v. 378, n. 2, p. 278-286, 2000.

ALMEIDA M S; CABRAL K M S; KURTENBACH E; ALMEIDA F C L; VALENTE A P. Solution structure of *Pisium sativum* defensin 1 by resolution NMR: plant defensins, identical backbone with different mechanisms of action. **J. Mol. Biol.** v. 315, p.749-757, 2002.

ALTINCICEK B; KOLHAS A; EBERL M; WIESNER J; SANDERBRAND S; HINTZ M; BECK E; JOMAA H. LytB, a novel gene of the 2-C-methyl-D-erythritol 4-phosphate pathway of isoprenoid biosynthesis in *Escherichia coli*. **FEBS Lett.** v. 499, p. 37–40, 2001a.

ALTINCICEK B; KOLHAS A; EBERL M; WIESNER J; SANDERBRAND S; HINTZ M; BECK E; JOMAA H. GcpE is involved in the 2-C-methyl-D-erythritol 4-phosphate pathway of isoprenoid biosynthesis in *Escherichia coli*. **J. Bacteriol.** v. 183, p. 2411–2416, 2001b.

ALTSCHUL S F; KOOPIN E V. Iterated profile searches with PSI-BLAST - a tool for discovery in protein databases. **Trends Biochem. Sci.** v. 23, p. 444–447, 1998.

ALTSCHUL S F; GISH W; MILLER W; MYERS E W; LIPMAN D J. Basic local alignment search tool. **J. Mol. Biol.** v. 215, p. 403–410, 1990.

ALUNNI B; KEVEI Z; REDONDO-NIETO M; KONDOROSI A; MERGAERT P; KONDOROSI E. Genomic Organization and Evolutionary Insights on *GRP* and *NCR* Genes, Two Large Nodule-Specific Gene Families in *Medicago truncatula.* **Mol. Plant Microbe Interact.** v. 20, n. 9, p. 1138-1148, 2007.

ALURU M; CURRY J; O'CONNELL M A. Nucleotide sequence of a defensin or γ-thionin-like gene (accession no. AF128239) from *habanera chili*. **Plant Physiol.** v.120, p.633, 1999.

ANAYA-LOPES J L; LÓPES-MEZA J E; BAIZABAL-AGUIRRE V M; CANO-CAMACHO H; OCHOA-ZARZOSA A. Fungicidal and cytotoxic activity of a *Capsium chinese* defensin expressed by endothelial cells. **Biotechnol Lett.** v. 28, p. 1101-1108, 2006.

ANFINSEN C B; Principles that govern the folding of protein chains. **Science**. v. 181, p. 223-230, 1973.

ANTCHEVA N; ZELEZETSKY L; TOSSI A. Cationic antimicrobial peptides- The defensins. **In the handbook of biologically active peptides**; A. J. Kastin, Ed.; Elsevier Science B. V: Amsterdam, p. 55-56, 2006.

ANURADHA T S; DIVYA K; JAMI S K; KIRTI P B. Transgenic tobacco and peanut plants expressing a mustard defensin show resistance to fungal pathogens. **Plant Cell Rep.** v. 27, p. 1777–1786, 2008.

ASADUZZAMAM S M; SONOMOTO K. Lantibiotics: diverse activities and unique modes of action. **J. Biosci. Bioeng.** v. 107, n. 5, p. 475-487, 2009.

ASHBURNER M; BALL C A; BLAKE J A; BOTSTEIN D; BUTLER H; CHERRY J M; DAVIS A P; DOLINSKI K; DWIGHT S S; EPPIG J T; HARRYS M A; HILL D P; ISSEL-TARVER L; KASARSKIS A; LEWIS S; MATESE J C; RICHARDSON J E; RINGWALD M; RUBIN G M; SHERLOCK. Gene ontology: tool for the unification of biology. **Nat. Genet.** v. 25, p. 25–29, 2000.

AUDIC S; CLAVERIE J M. The significance of digital gene expression profiles. **Genome Res.** v. **7**, p. 986-995, 1997.

BAIROCH A; APWEILER R; WU C H; BARKER W C; BOECKMANN B; FERRO S; GASTEIGER E; HUANG H; LOPEZ R; MAGRANE M; MARTIN M J; NATALE D A; O'DONOVAN C; REDASCHI N; YEH –S L. The universal protein resource (UniProt). **Nucleic Acids Res.** 33, D154–D159, 2005.

BAGNAT M; KERANEN S; SHEVCHENKO A; SIMONS K. Lipid rafts function in biosynthetic delivery of proteins to the cell surface in yeast. **Proc. Natl. Acad. Sci**. v. 97, p. 3254-3259, 2000.

BATEMAN A; COIN L; DURBIN R; FINN R D; HOLLICH V; GRIFFITHS-JONES S; KHANNA A; MARSHALL M; MOXON S; SONNHAMMER E L L; STUDHOLME D J; EDDY C Y S R. The Pfam protein families' database. **Nucleic Acids Res.** v. 28, p. 263–266, 2000.

BEER A; VIVIER M A. Vv-AMP1, a ripening induced peptide from *Vitis vinifera* shows strong antifungal activity. **BMC Plant Biol.** v.8, p. 75, 2008.

BELARMINO L C; CAPRILES P V S Z; CROVELLA S; DARDENE L E; BENKO-ISEPPON A M. EST-based search of plant defendinss – an example using the sugarcane, a large and complex genome. **Curr. Protein Pept. Sci.** 2010, *In Press.*

BELARMINO L C; BENKO-ISEPPON A M. Databank based mining on the track of antimicrobial weapons in plant genomes. **Curr. Protein Pept. Sci.** 2010, *In Press.*

BENKO-ISEPPON A M. Estudos moleculares e citogenéticos no caupi e em espécies relacionadas: Avanços e perspectivas. In: Avanços tecnológicos no Feijão-caupi. **Anais da V Reunião nacional de pesquisa de caupi**, EMBRAPA Meio Norte, Terezina, Piauí, p. 327-332, documentos 56, 2001.

BERMAN H M; WESTBROOK J; FENG Z; GILLILAND G; BHAT T N; WEISSIG H; SHINDYALOV I N; BOURNE P E. The protein data bank. **Nucleic Acids Res.** v. 28, p. 235–242, 2000.

BERGMANN S; IHMELS J; BARKAI N. Similarities and differences in genome-wide expression data of six organisms. **PLoS Biol**. v. 2, E9, 2004.

BLOCH C J; RICHARDSON M. A new family of small (5 kDa) pro tein inhibitors of insect a-amylases from seeds of *sorghum* (*Sorghum bicolor* (L) Moench) have sequence homologies with wheat g-purothionins. **FEBS Lett**. v. 279, n. 1, p. 101–104, 1991.

BLUNDELL T L; BEDARKAR S; RINDERKNECH E; HUMBLE R E. Insulin-like growth factor: a model for tertiary structure accounting for immunoreactivity and receptor binding. **Proc. Natl. Acad. Sci.** v. 75, p. 180-184, 1978.

BRAHMACHARY M; KRISHNAN S P T; KOH J L Y; KHAN A M; SEAH S H; TAN T W; BRUSIC V; BAJIC V B. ANTIMIC: a database of antimicrobial sequences. **Nucleic Acids Res.** v. 32, p. 586-589, 2004

BRANDSTADTER J; ROSSBACH C; THRES K. Expression of genes for a defensin and a proteinase inhibitor in specific areas of the shoot apex and the developing flower in tomato. **Mol. Gen. Genet**. v. 252, p.146-154, 1996.

BRINKMANN U; VASMATZIS G; LEE B; YERUSHALMI N; ESSAND M; PASTAN I. PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in normal andneoplastic prostate, testis, and uterus. **Proc. Natl. Acad. Sci.** v. 95, p. 10757-10762, 1998.

BROEKAERT W; VAN PARIJS J; LEYNS F; JOOS H; PEUMANS W A. A chitin-binding lectin from stinging nettle rhizomes with antifungal properties. **Science**. v. 245, p. 1100-1102, 1989.

BROGDEN K A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? **Nat. Rev. Microbiol.** v. 3, p. 238–250, 2005.

BRUIX M; JIMENEZ M A; SANTORO J; GONZALEZ C; COLILLA F J; MENDEZ E; RICO M. Solution structure of γ1-H and γ1-P thionins from barley and wheat endosperm determined by H-NMR: a structural motif common to toxic arthropod proteins. **Biocheimstry**. v. 32, n. 2, p. 715-724, 1993.

BRUIX M; GONZALEZ C; SANTORO J; SORIANO F; ROCHER A, MENDEZ E. H-NMR studies on the structure of a new thionin from barley endosperm. **Biopolymers**. v. 36, p. 751–763, 1995.

BULET P; HETRU C; DIMARCQ J-L; HOFMANN D. Antimicrobial peptides in insects; structure and function. **Dev. Comp. Immunol.** v. 23, p. 329–344, 1999.

BULET P; STOECKLIN R; MENIN L. Anti-microbial peptides: from invertebrates to vertebrates. **Immunological Reviews**. v.198, p. 169–184, 2004.

BULET P; STOECKLIN R. Insect Antimicrobial Peptides: Structures, Properties and Gene Regulation. **Protein Pept. Letter.** v. 12, p. 3-11, 2005 .

BUSCAGLIA C A; CAMPO V A; FRASCH A C C; NOIA J M D. *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. **Nat. Rev. Microbiol.** v. 4, p. 229–236, 2006.

BUTTERFIELD M K; D'HONT A; BERDING N. In Proceedings of the South African Sugar Technologists' Association (SASTA '01), Durban, South Africa, July-August. **The sugarcane genome: a synthesis of current understanding, and lessons for breeding and biotechnology**. 75, p. 1–5, 2001.

CABRAL K M; ALMEIDA M S; VALENTE A P; ALMEIDA F C; KURTENBACH E. Production of the active antifungal *Pisum sativum* defensin 1 (Psd1) in *Pichia pastoris*: overcoming the inefficiency of the STE13 protease. **Protein Expr. Purif.** v. 31, n.1, p. 115-122, 2003.

CAMPILLOS M; VON MERING C; JENSEN L J; BORK P. Identification and analysis of evolutionarily cohesive functional modules in protein networks. **Genome Res.** v. 16, p. 374-382, 2006.

CAMOM E; MAGRANE M; BARRELL D; LEE V; DIMMER E; MASLEN J; BINNS D; HARTE N; LOPEZ R; APWEILER R. The Gene Ontology annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. **Nucleic Acids Res.** v. 32, p. 262–266, 2004.

CGAP (Cancer Genome Anatomy Project). cDNA xProfiler. Disponível em: <**http://cgap.nci.nih.gov/Tissues/xProfiler**> Acesso em: 13 nov. 2009.

CGAP (Cancer Genome Anataomy Project). cDNA Digital Gene Expression Displayer (DGED). Disponível em: <**http://cgap.nci.nih.gov/Tissues/GXS>** Acesso em:14 nov. 2009.

CARVALHO A O; FILHO G A S; FERREIRA B S; BRANCO A T; OKOROKOVA-FAÇANHA, GOMES V M. Cloning and characterization of a cDNA encoding a cowpea seed defensin and analysis of its expression. **Protein Pept. Lett.** v. 13, p. 1029–1036, 2006.

CARVALHO A O; MACHADO O L T; DA CUNHA M; SANTOS I S; GOMES V M. Antimicrobial peptides and immunolocalization of a LTP in vigna unguiculata seeds. **Plant Physiol. Biochem.** v. 39, p. 137- 146, 2001.

CECILIANI F; BORTOLOTTI F; MENEGATTI E; RONCHI S; ASCENZI P; PALMIERI S. Purification, inhibitory properties, amino acid sequence and identification of the reactive site of a new serine proteinase inhibitor from oil-rape (*Brassica napus*) seed. **FEBS Lett.** v. 342, n. 2, p. 221-224, 1994.

CERAUL S M; DREHER-LESNICK S M; GILLESPIE J J; RAHAMAN M S; AZAD A F. New Tick Defensin Isoform and Antimicrobial Gene Expression in Response to *Rickettsia montanensis* Challenge. **Infection and Immunity**. v. 75, n. 4, p. 1973-1983, 2007.

CHARLET M; CHERNYSH S, PHILIPPE H; HETRU C; HOFFMANN J A, BULET P. Isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusk, *Mytilus Edulis*. **J. Bio. Chem.** v. 271, n. 36, p. 21808–21813, 1996.

CHEN G-H; HSU M-P; TAN C-H; SUNG H-Y; KUO C G; FAN M-J, *et al*. Cloning and characterization of a plant defensin VaD1 from azuki bean. **J. Agric. Food Chem.** v. 8, p. 982–988, 2005.

CHEN K-C; LIN C-Y; Kuan C-C; SUNG H-Y; CHEN C-S. A novel defensin encoded by a mungbean cDNA exhibits insecticidal activity against bruchid. **J. Agric. Food Chem.** v. 50, p. 7258–7263, 2002a.

CHEN K-C; LIN C-Y; CHUNG M-C; KUAN C-C; SUNG H-Y; TSOU S C S; et al. Cloning and characterization of a cDNA encoding an antimicrobial protein from mung bean seeds. **Bot. Bull. Acad. Sin.** v. 43, p. 251-259, 2002b.

CHEN Jr J; CHEN G-H; HSU H-C; LI S-S; CHEN C-S. Cloning and functional expression of a mungbean defensin VrD1 in *Pichia pastoris*. **J. Agric. Food Chem.** v. 52, p. 2256–2261, 2004.

CHEN Z; GALLIE DR. Dehydroascorbate reductase affects leaf growth, development and function. **Plant Physiol.** v. 142, p. 775–787, 2006

CHIANG C C; HADWIGER L A. The *Fusarium solani*-induced expression of a pea gene family encoding high cysteine content proteins. **Mol. Plant Microbe Interact**. v. 4, n. 4, p. 324-331, 1991.

CHOI Y; CHOI Y D; LEE J S. Nucleotide sequence of a cDNA encoding a low molecular weight sulfur-rich protein in soybean seeds. **Plant Physiol.** v. 101, n. 2, p. 699-700, 1993.

COLILLA F J; MENDEZ R A; MENDEZ E. γ-purothionins: amino acid sequence of two polypeptides of a new family of thionins from wheat endosperm. **FEBS Lett.** v. 270, n.1, p. 191-194, 1990.

CONAB (Companhia Nacional de Abastecimento). **Acompanhamento da safra Brasileira de cana-de-açúcar 2009/2010**: terceiro levantamento. Brasília.16p. 2009.

COPLEY S D. Evolution of a metabolic pathway for degradation of a toxic xenobiotic: the patchwork approach. **Trends Biochem. Sci.** v. 25, p. 261–265, 2000.

CORNET B; BONMATIN J-M; HETRU C, HOFFMANN J A; PTAK M; VOVELLE F. Refined three-dimensional solution structure of insect defensin A. **Structure**. v. 3, n.5, p. 435-448, 1995.

CUNNINGHAM Jr F X; LAFOND T P; GANTT E. Evidence of a role for LytB in the nonmevalonate pathway of isoprenoid biosynthesis. **J. Bacteriol.** v. 182, p. 5841–5848, 2000.

DA-HUI L; GUI-LIANG J; YING-TAO Z; TIE-MIN A. Bacterial expression of a *Trichosanthes lirilowii* defensin (TDEF1) and its antifungal activity on *Fusarium oxysporum*. **Appl. Microbiol. Biotechnol.** v. 74, p.146–151, 2007.

DANDEKAR T; SNEL B; HUYNEN M; BORK P. Conservation of gene order: a fingerprint of proteins that physically interact. **Trends Biochem. Sci.** v. 23, p. 324-328, 1998.

DANIEL S J; ROACH B T. Taxonomy and Evolution. In: HEINZ, D J (Ed.Téc.). **Sugarcane improvement through breeding**. Amsterdam: Elsevier. p. 7-84, 1987.

DE SAMBLANX G W; GODERIS I J; THEVISSEN K; RAEMAEKERS R; FANT F; BORREMANS F; ACLAND D P; OSBORN R W; PATEL S; BROEKAERT W F. Mutational Analysis of a Plant Defensin from Radish (*Raphanus sativus* L.) Reveals Two Adjacent Sites Important for Antifungal Activity. **J. biol. chem.** v. 272, n. 2, p. 1171–1179, jan 1997.

D'HONT A. Unraveling the genome structure of polyploids using FISH and GISH examples of sugarcane and banana. **Cytogen. Gen. Res.** v. 109, n. 1-3, p. 27-33, 2005.

D'HONT A; GLASZMAN J C. Sugarcane genome analysis with molecular markers, a first decade of research. **Proceedings of the International Society for Sugar Cane Technology.** v. 24, p. 556-559, 2001.

DIJK A V; VELDHUIZEN E J A; HAAGSMAN H P. Avian defensins. **Veterinary Immunology and Immunopathology**. v. 124, p. 1–18, 2008.

DO H M; LEE S C; JUNG H W; SOHN K H; HWANG B K. Differential expression and in situ localization of a pepper defensin (CADEF1) gene in response to pathogen infection, abiotic elicitors and environmental stresses in Capsicum annuum. **Plant Sci.** v. 166, p.1297–1305, 2004.

DOMON C; EVRARD J L; HERDENBERGER F; PILLAY DT; STEINMETZ A. Nucleotide sequence of two anther-specific cDNAs from sunflower (*Helianthus annuus* L.). **Plant Mol. Biol.** v. 15, n. 4, p. 643-646, 1990.

DOUGHTY J; DIXON S; HISCOCK S J; WILLIS A C; PARKIN I A P; DICKNSONA H G. PCP-A1, a Defensin-like Brassica Pollen Coat Protein That Binds the *S* Locus Glycoprotein, Is the Product of Gametophytic Gene Expression. **Plant Cell**. v. 10, p. 1333–1347, 1998.

DURBIN R; EDDY S R; KROGH A; MITCHISON G. Biological sequence analysis: probabilistic models of proteins and nucleic acids. **Cambridge University Press Cambridge**, UK, 1998.

EDWARDS D; BATLEY J. Plant bioinformatics: from genome to Phenome. **Trends Biotechnol.** v. 22, n. 5, p. 232-237, 2004.

EHLERS J D, HALL A E. Cowpea (*Vigna unguiculata* (L.) Walp.). **Field Crops Res**. v 53, p. 187-204, 1997.

ELLIS R H, LAWER R J, SUMMERFIELD R J, ROBERTS E H, CHAY P M, BROUWER J B, ROSE J L, YEATES S J. Towards the reliable prediction on time to flowering in six annual crops. III. Cowpea (*Vigna unguiculata*). **Experimental Agriculture**, v.30, p.17-29, 1994.

EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária). Disponível em: <**http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Feijao/FeijaoCaupi/importancia.htm#topo**>. Acesso em: 17 nov de 2009.

EPPIG J T; BULT C J; KADIN J A; RICHARDSON J E; BLAKE J A. The Mouse Genome Database Group. The Mouse Genome Database (MGD): from genes to mice - a community resource for mouse biology. **Nucleic Acids Res.**v. 33, D471–D475, 2005.

EPPLE P; APEL K; BOHLMANN H. ESTs reveal a multigene family for plant defensins in *Arabidopsis thaliana*. **FEBS Lett.** v. 400, p. 168-172, 1997.

FANT F; VRANKEN W F; MARTINS J C; BORREMANS F A M. Determination of the three-dimensional solution structure of *Raphanus sativus* antifungal protein 1 by 1HNMR. **J. Mol. Biol**. v. 279, p. 257-270, 1998.

FANT F; VRANKEN W F; BORREMANS F A M. The three-dimensional solution structure of *Aesculus hippocastanum* antimimicrobial protein 1 determined by 1H nuclear magnetic resonance. **Proteins**. v. 37, p. 388–403, 1999.

FAO (Food and Agriculture Organization of the United Nations). Disponível em: <**http://faostat.fao.org/default.aspx**>. Acesso em: 13 nov de 2009.

FEDOROVA M; VAN DE MORTEL J; MATSUMOTO P A; CHO J; TOWN C D; VANDENBOSCH K A; GRANTT J S; VANCE C P. Genome-wide identificationof nodule-specific transcripts in the model legume Medicago truncatula. **Plant Physiol.** v. 130, p. 519 e 537, 2002.

FEHLBAUM P; BULET P; MICHAUT L; LAGUEUX M; BROEKAERT W F; HUTRU C; HOFFMAN J A. Septic injury of rosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. **J. Biol. Chem.** v. 269, n. 52, p. 33159-33163, 1994.

FINKINA E I; SHARAMOVA E I; TAGAEV A A; OVCHINNIKOVA T V. A novel defensin from the lentil Lens culinaris seeds. **Biochem. Biophys. Res. Commun.** v. 371, p. 860–865, 2008.

FISER A; SALI A. **In Protein Structure** (Chasman, D., Ed.).p. 167-206. 2003. Marcel Dekker, Inc. New York.

FJELL C D; HANCOCK R E W; CHERKASOV A. AMPer: a database and an automated discovery tool for antimicrobial peptides. **Bioinformatics**. v. 23, p. 1148- 1155, 2007.

FRANCO O L; MURAD A M; LEITE J R; MENDES P A M; PRATES M V; BLOCH Jr C. Identification of a cowpea g-thionin with bactericidal activity. **FEBS J.** v. 273, p. 3489–3497, 2006.

FRANCO O L, RIGDEN D J, MELO F R, GROSSI DE SÁ M F, Plant α-amylase inhibitors and their interaction with insect α-amylases structure, function and potential for crop protection. **European Journal of Biochemistry**. v.269, p. 397- 412, 2002.

FRANCOIS I E J A; DE BOLLE M F C; GODERIS I J W ; WOURTORS P F J; VERHAERT P D; PROOST P, SCHAAPER W M M, CAMMUE B P A; and Willem F. Broekaert. Transgenic Expression in Arabidopsis of a Polyprotein Construct Leading to Production of Two Different Antimicrobial Proteins1. **Plant Physiol.** v. 128, p. 1346–1358, 2002

FREIRE FILHO F R. Cowpea taxonomy and introduction to Brazil. In WATT E E, ARAÚJO J P P. Cowpea research in Brazil. **IITA, EMBRAPA**, Brasilia, p. 3-10, 1988.

FROKJAER S; OTZEN D E. Protein drug stability: a formulation challenge. **Nat. Rev. Drug Discov.** v. 4, p. 289-306, 2005.

FROTA K M G, SOARES R A M, ARÊAS JOSÉ A G. Composição química do feijão-caupi (*Vigna unguiculata* (L.) Walp.), cultivar BRS-Milênio. **Ciência e Tecnologia de Alimentos**. v. 28, n. 2, p. 470-476, 2008.

FUJIMURA M; IDEGUCHI M; MINAMI Y; WATANABE K; TADERA K. Purification, characterization and sequencing of novel antimicrobial peptides Tu-AMP1 and Tu-AMP2 from bulbs of tulip (*Tulipa gesneriana* L.). **Biosci. Biotechnol. Biochem.** v. 68, n 3, p. 571–577, 2004.

GALVÃO L S; FORMAGGIO A R; TISOT D A. Discriminação de variedades de cana-de-açúcar com dados hiperespectrais do sensor Hyperion/EO-1. **Rev. Bras. Cartogr.** v. 57, n. 1, p.7-14, 2005.

GANZ T. Defensins: antimicrobial peptides of vertebrates. **CR Biol.** v. 327, p. 539–549, 2004.

GANZ T; LEHRER R I. Defensins. **Curr. Opin. Immunol.** v. 6, p. 584–589, 1994.

GAO B; RODRIGUEZ M Del M; LANZ-MENDOZA H; ZHU S. AdDLP, a bacterial defensin-like peptide, exhibits anti-Plasmodium activity. **Biochem. Biophys. Res. Commun.** v. 387, p. 393–398, 2009.

GAO A G; HAKIMI S M; MITTANCK C A; WU Y; WOERNER B M; STARK D M; et al. Fungal pathogen protection in potato by expression of a plant defensin peptide. **Nat. Biotechnol.** v. 18, p. 1307–10, 2000.

GARCÍA-OLMEDO F; CARBONERO P; HERNANDEZ-LUCAS C; PAZ ARES J; PONZ F; VICENTE O; et al. Inhibition of eukaryotic cell-free protein synthesis by thionins from wheat endosperm. **Biochim. Biophys. Acta.** v. 740, p. 52–56, 1983.

GAUDET D A; LAROCHE A; FRICK M; HUEL R; PUCHALSKI B. Cold induced expression of plant defensin and lipid transfer protein transcripts in winter wheat. **Plant Physiol.** v. 117, p. 195-205, 2003.

GEPTS P, BEAVIS W D, BRUMMER E C, SHOEMAKER R C, STALKER H T, WEEDEN N F, YOUNG N D. Legumes as a Model Plant Family. Genomics for Food and Feed Report of the Cross-Legume Advances through Genomics Conference. **Plant Physiol**. V 137, p. 1228-1235, 2005.

GHANNOUM M; ABU EL-TEEN A K; RADWAN S S. Blocking adherence of *Candida albicans* to buccal epithelial cells by yeast glycolipids, yeast wall lipids and lipids from epithelial cells. **Mykosen**. v. 30, p. 371-378, 1987.

GOLDSMITH-FISCHMAN S; HONING B. Structural genomics: Computational methods for structure analysis. **Protein Sci.** v. 12, n. 9, p. 1813-1821, 2003.

GRAHAM M A; SILVERSTEIN K A; CANNON S B; VANDENBOSCH K A. Computational Identification and Characterization of Novel Genes from Legumes. **Plant. Physiol.** v. 135, p. 1179-1197, 2004.

GRAMENE. Disponível em: <**http://www.gramene.org/**> Acesso em: 11 nov. 2009.

GRIVET L; ARRUDA P. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Curr. Opin. Plant Biol.** v. 5, n. 2, p. 122-127, 2001.

GU Q; KAWATA E E; MORSE M J; WU H M; CHEUNG A Y. A flower-specific cDNA encoding a novel thionin in *tobacco*. **Mol. Gen. Genet.** v. 234, p. 89-96, 1992.

HALL A E. Breeding for adaptation to drought and heat in cowpea. **Eur J Agron**. v. 21, p.447–454, 2004.

HAMMAMI R; HAMIDA J B; VERGOTEN G; FLIS I. PhytaAMP: a database dedicated to antimicrobial plant peptides. **Nucleic Acids Res.** v. 37, D963-D968, 2009.

HANKS J N; SNYDER A K; GRAHAM M A; SHAH R P; BLAYLOCK L A; HARRISON M J; SHAH D M. Defensin gene family in Medicago truncatula: structure, expression and induction by signal molecules. **Plant Mol. Biol.** v. 58, p.385 e 399, 2005.

HARRISON R W; CHATTERJEE D; WEBER I.T. Analysis of six protein structures predicted by comparative modeling techniques. **Proteins**. v. 23, n. 4, p. 463-671, 1995.

HERMJAKOB H; MOTECCHI-PALAZZI L; LEWINGTON C; MUDALI S; KERRIER S; ORCHARD S; VINGRON M; ROECHERT B; ROEPSTORFF P; VALENCIA A; et al. IntAct: an open source molecular interaction database. **Nucleic Acids Res.** v. 32, D452–D455, 2004.

HIMLY M; JAHN-SCHMID B; DEDIC A; KELEMEN P; WOPFNER N; ALTMANN F; VAN REE R; BRIZA P; RICHTER K; EBNER C; FERREIRA F. The major allergen of mugwort pollen, is a modular glycoprotein with a defensina-like and a hydroxyproline-rich doimain. **FASEB J.** v. 17, n. 1, p. 106-108, 2003.

HU X; BIDNEY D L; YALPANI N; DUVICK J P; CRASTA O; FOLKERTS O; LU G. Over expression of a gene encoding hydrogen peroxide-generating oxalate oxidase evokes defense responses in sunflower. **Plant Physiol.** v. 133, n. 1, p. 170-181, 2003

HUANG G-J; LAI H-C; CHANG Y-S; SHEU M-J; LU T-L; HUANG S-S, et al. Antimicrobial, dehydroascorbate reductase, and monodehydroascorbate reductase activities of defensin from sweet potato [*Ipomoea batatas* (L.) Lam. 'tainong57'] storage roots. **J. Agric. Food Chem.** v. 56, p. 2989–2995, 2008.

HUYNEN M; SNEL B; LATHE W; BORK P. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. **Genome Res.** v. 10, p. 1204–1210, 2000.

ISHIBASHI N; YAMAUCHI D; MINAMIKAWA T. Stored mRNA in cotyledons of *Vigna unguiculata* seeds: nucleotide sequence of cloned cDNA for a stored mRNA and induction of its synthesis by precocious germination. **Plant Mol. Biol.** v. 15, n. 1, p. 59-64, 1990.

JANSSEN B J C; SCHIRRA H J; LAY F T; ANDERSON M A; CRAIK D J. Structure of Petunia hybrida defensin 1, a novel plant defensin with five disulfide bonds. **Biochemistry**. v. 42, p. 8214-8222, 2003.

JIM K; PARMAR K; SINGH M; TAVAZOIE S. A crossgenomic approach for systematic mapping of phenotypic traits to genes. **Genome Res.** v. 14, p. 109-115, 2004.

JIMENEZ-LUCHO V; GINSBURG V; KRIVAN H C. *Cryptpcoccus neoformans*, *Candida albicans*, and other fungi bind specifically to the glycosphingolipid lactosylceramide (Galβ1-4Glcβ1-1Cer), a possible adhesion receptor for yeasts. **Infection Immunology**. v. 58, p. 2085-2090, 1990**.**

JOHNSON M S; SRINIVASAN N; SOWDHAMINI R; BLUNDELL T L. Knowledge-based protein modellingCrit. **Rev. Biochem. Mol. Biol.** v. 29, n. 1, p. 1-68, 1994**.**

KANZAKI H; NIRASAWA S; SAIOTH H; ITO M; NISHIHARA M; TERAUCHI R; et al. Overexpression of the wasabi defensin gene confers enhanced resistance to blast fungus (*Magnaporthe grisea*) in transgenic rice. **Theor. Appl. Genet.** v. 105, p. 809–814, 2002.

KANT P; LIU W Z; PAULS K P. PDC1, a corn defensin peptide expressed in *Escherichia coli* and *Pichia pastoris* inhibits growth of *Fusarium graminearum*. **Peptides**. v. 30, n. 9, p.1593- 1599, 2009.

KOGA J; YAMAUCHI T; SHIMURA M; OGAWA N; OSHIMA K; UMEMURA K; KIKUCHI M; OGASAWARA N. Cerebrosides A and C, phingolipid elicitors of hypersensitive cell death and phytoalexin accumulation in rice plants. **J. Biol. Chem.** v. 273, p.31985-31991, 1998.

KOIKE M; OKAMOTO T; TSDUA S; IMAI R. A novel plant defensin-like gene of winter wheat is specifically induced during cold acclimation. **Biochem. Biophys. Res. Commun.** v. 298, p. 46–53, 2002.

KOLINSKI A; ROTKIEWICZ P; ILKOWSKI B; SKOLNICK J. A method for the improvement of threading-based protein models. **Proteins**. v. 37, n. 4, p. 592-610, 1999.

KOMORI T; YAMADA S; IMASEKI H A cDNA clone for γ-thionin from *Nicotiana paniculata* (accession no. AB005250; PGR97–132). **Plant Physiol.** v. 1, n. 15, p. 314, 1997.

KONG M; BARNES E A; OLLENDORFF V; DONOGHUE D J. Cyclin F regulates the nuclear localization of cyclin B1 through a cyclin–cyclin interaction. **EMBO J.** v. 19, p. 1378–1388, 2000.

KOVALSKAYA N; HAMMOND R W. Expression and functional characterization of the plant antimicrobial snakin-1 and defensin recombinant proteins. **Protein Expr. Purif.** v. 63, n. 1, p.12-17, 2009.

KRAGH K M; NIELSEN J E; NIELSEN K K; DREBOLDT S; MIKKELSEN J D. Characterization and localization of a new antifungal cysteine-rich proteins form *Beta vulgaris*. **Mol. Plant Microb. Interact.** v.8, n. 5, p. 424–434, 1995.

KRALLINGER M; VALENCIA A; HIRSCHMAN L. Linking genes to literature: text mining, information extraction and retrieval appliacations for biology. **Gen. Biology**. v. 9, s. 2, 2008.

KREZEL A M; KASIBHATLA C; HIDALGO P; MACKINNON R; WAGNER G. Solution structure of the potassium channel inhibitor agitoxin 2: caliper for probing channel geometry. **Protein Science**. v. 4, p. 1478-1489, 1995.

KURUNANANDAA B; SINGH A; KAO T H. Characterization of a predominantly pistil-expressed gene encoding a gamma-thionin-like protein of *Petunia inflate*. **Plant Mol. Biol.** v. 26, n. 1, p. 459-464, 1994.

KUSABA M; DWYER K; HENDERSHOT J; VREBALOV J; NASRALLAH J B; NASRALLAH M E. Self-incompatibility in the genus Arabidopsis: characterization of the S locus in the outcrossing A. *lyrata* and its autogamous relative A. *thaliana*. **Plant Cell**. v. 13, n. 3, p. 627-643, 2001.

KUSHMERICK C; CASTRO M S; CRUZ J S; BLOCH Jr C; BEIRÃO P S L. Functional and structural features of g-zeathionins, a new class of sodium channel blockers. **FEBS Lett.** v. 440, p. 302–306, 1998.

LAL A; LASH A E; ALTSCHUL S F; VELCULESCU V; ZHANG L, MCLENDON R E; MARRA M A; PRANGE C; MORIN P J; POLYAK K; PAPADOPULOS N; VOGELSTEIN B; KINZLER K W; STRAUSBERG R L; RIGGINS G J. A public database for gene expression in human cancers. **Cancer Res.** v. 59, p. 5403-5407, 1999.

LANDON C; CORNET B; BONMATIN J-M; KOPEYAN C; ROCHAT H; VOVELLE F; PTAK M. H-NMR-derived secondary structure and the overall fold of the potent anti-

mammal and anti-insect toxin I11 from the scorpion *Leiurus quinquestriatus quinquestria.* **Eur. J. Biochem.** v. 236, p. 395-404, 1996.

LANGEN G; IMANI J; ALTINCICEK B; KIESERITZKY G; KOGEL K H; VILCINKAS A. Transgenic expression of gallerimycin, a novel antifungal insect defensin from the greater wax moth *Galleria mellonella*, confers resistance to pathogenic fungi in tobacco. **J. Biol. Chem.** v. 387, p. 549-557, 2006.

LAY F –M; DELONG C; MEI K; WIGNES T; FOBERT P R. Analysis of the DRR230 family of pea defensins: gene expression pattern and evidence of broad host-range antifungal activity. **Plant Science**. v. 163, n. 4, p. 855-864, 2002.

LAY F T; ANDERSON M A. Defensins – Components of the Innate Immune System in Plants. **Curr. Protein Pept. Sci.** v. 6, p. 85-101, 2005.

LAY F T; BRUGLIERA F; ANDERSON M A. Isolation and Properties of Floral Defensins from Ornamental Tobacco and Petunia. **Plant Physiol.** v. 131, p. 1283-1293, 2003a**.**

LAY F T; SCHIRRA H J; SCANLON M J; ANDERSON M A; CRAIK D J. J. The three-dimensional solution structure of NaD1, a new floral defensin from *Nicotiana alata* and its application to a homology model of the crop defense protein alfAFP. **Mol. Biol.** v. 325, p.175-188, 2003b.

LIN K F; LEE T R; TSAI P H; HSU M P; CHEN C ; LYU P C. Structure-based protein engineering for a-amylase inhibitory activity of plant defensin. **Proteins**. v. 68, p. 530–540, 2007.

LIU Y J; CHENG C S; LAI S M; HSU M P; CHEN C S; LYU P C. Solution structure of the plant defensin VrD1 from mung bean and its possible role in insecticidal activity against bruchids. **Proteins**. v. 63, p. 777–786, 2006.

LIU L; GANZ T. The pro region of human neutrophil defensin contains a motif that is essential for normal subcellular sorting. **Blood**. v. 85, p.1095-1103, 1995.

LOBO D S; PEREIRA I B; FRAGEL-MEDEIRA L; MEDEIROS L N; CABRAL L M; FARIA J; BELLIO M; CAMPOS R C; LINDEN R; KURTENBACH E. Antifungal *Pisum sativum* defensin1 interacts with *Neurospora crassa* cyclin F related to the cell cycle. **Biochemistry**. v. 46, p. 987–996, 2007.

LOHAR D P; SHAROPOVA N; ENDRE G; PEÑUELA S; SAMAZ D; TOWN C; SILVERSTEIN K A T; VANDENBOSCH K A. Transcript Analysis of Early Nodulation Events in Medicago truncatula1, 2[W]. **Plant Physiol.** v. 140, p. 221–234, 2006.

MAKAROVA K S; WOLF Y I; KOONIN E V. Potential genomic determinants of hyperthermophily. **Trends Genet.** v. 19, p. 172–176, 2003.

MAGLOIRE N. The genetic, morphological and physiological evaluation of African cowpea genotypes. 119 p. Thesis (Magister Scientiae Agriculturae) – University of the Free State, Bloemfontein, South Africa, 2005.

MAITRA N; CUSHMAN J C. Characterization of a drought-induced soybean cDNA encoding a plant defensin. **Plant Physiol.** v. 118, p. 1536, 1998.

MALISAN F; RIPPO M R; DEMARIA R; TESTI R. Lipid and glycolipid mediators in CD95-induced apoptotic signaling. **Results Probl. Cell Differ.** v. 23, p. 65-76, 1999.

MANNERS J M. Hidden weapons of microbial destruction in plant genomes. **Genome Biology**. v. 8, p. 225, 2007.

MANNERS J M; PENNINCKX I A M A; VERMAERE K; KANZAN K; BROWN R L; Morgan A. The promoter of the plant defensin gene PDF1.2 from Arabidopsis is systemically activated by fungal pathogens and responds to methyl jasmonate but not to salicylic acid. **Plant Mol. Biol.** v. 38, p.1071–1080, 1998.

MARQUÈS L M; OOMEN R J; AUMELAS A; LE JEAN M; BERTHOMIEU P. Production of an *Arabidopsis halleri* foliar defensin in *Escherichia coli*. **J. Appl. Microbiol.** v. 106, n. 5, p.1640-1648, 2009.

MARX F. Small, basic antifungal proteins secreted from filamentous ascomycetes: a comparative study regarding expression, structure, function and potential application. **Appl. Microbiol. Biotechnol.** v. 65, p. 133–142, 2004.

MELO F R; RIGDEN D J; FRANCO O L; MELLO L V; ARY M B; GROSSI DE SÁ M F; et al. Inhibition of trypsin by cowpea thionin: characterization, molecular modeling, and docking. **Proteins**. v. 48, n. 2, p. 311–319, 2002.

MENDEZ E; MORENO A; COLILLA F J; PALAEZ F; LIMAS G G; MENDEZ R; et al. Primary structure and inhibition of protein synthesis in eukaryotic cell-free system of a novel thionin, g-hordothionin, from barley endosperm. **Eur. J. Biochem.** v. 194, p. 533-539, 1990.

MENDEZ E; ROCHER A; CALERO M; GIRBES T; CITORES L; SORIANO F. Primary structure of v-hordothionin, a member of a novel family of thionis from barley endosperm, and its inhibition of protein synthesis in eukaryotic and prokaryotic cell-free systems. **Eur. J. Biochem.** v. 239, p. 67–73, 1996.

MENOSSI M; SILVA-FILHO M C; VICENTZ M,VAN-SLUYS M –A; SOUZA G M. Sugarcane Functional Genomics: Gene Discovery for Agronomic Trait Development. **Int. J. Plant Genomics**. v. 2008, p. 1-11, 2008.

MERGAERT P; NIKOVICS K; KELEMEN Z; MAUNOURY N; VAUBERT D; KONDOROSI A; KONDOROSI E. A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteinemotifs. **Plant Physiol.** v. 132, p.161 e 173, 2003.

MESITI M; JIMÉNEZ-RUIZ E; SANZ I; BERLANGA-LLAVORI R; PERLASCA P; VALENTINI G; MANSET. XML-based approaches for the integration of heterogeneous bio-molecular data. **BMC Bioinformatics**. v. 10, s. 12, 2009.

MICHAELSON D, RAYNER J, COUTO M, GANZ T. Cationic defensins arise from charge-neutralized propeptides: a mechanism for avoiding leukocyte autocytotoxicity?. **J. Leukoc. Biol.** v. 51, p. 634-639, 1992.

MILLIGAN S B; GASSER C S. Nature and regulation of pistil-expressed genes in tomato. **Plant Mol. Biol.** v. 28, p. 691-711, 1995.

MING D; HELLEKANT G. Brazzein, a new high potency thermostable sweet protein from *Pentadiplandra brazzeana* B. **FEBS Lett.** v. 355, p.106–108, 1994.

MING R; MOORE P H; WU K K; et al. Sugarcane improvement through breeding and biotechnology. **Plant Breeding Reviews**. v. 27, p. 15–118, 2006.

MIROUZE M; SELS J; RICHARD O; CZERNIC P; LOUBET S; JACQUIER A; et al. A putative novel role for plant defensins: a defensin from the zinc hyper-accumulating plant, Arabidopsis halleri, confers zinc tolerance. **Plant J.** v. 47, p. 329–342, 2006.

MYGIND H P; FICSHER R L; SCHNORR K M; HANSEN M T; SONKSEN C P; LUDVIGSEN S; RAVENTÓS D; BUSKOV S; CHRISTENSEN B; DE MARIA L; TABOUREAU O; YAVER D; JORGENSEN S G E; SORENSEN M V; CHRISTENSEN B E; KJÆRULFFI S; FRIMODT-MOLLER N; LEHRER R L; ZASLOFF M; KRISTENSEN H -H. Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. **Nature**. v. 437, n. 13, p. 975-980, 2005.

MOLINA C; ROTTER B; HORRES R; UDUPA S M; BESSER B; BELARMINO L; BAUM M; MATSUMURA H; TERAUCHI R; KAHLL G; WINTER P. SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. **BMC Genomics**. v. 9, p. 553, 2008.

MOLOSOF V V; VALEUVA T A. Proteinase inhibitors in plant biotechnology: a review. **Appl. Biochem. Microbiol.** v. 44, n. 3, p. 233–240, 2008.

MONK B C; HARDING D R K. Peptide motifs for cell-surface intervention. **Biodrugs**. v. 19, n. 4, p. 261–278, 2005.

MORENO M; SEGURA A; GARCÍA-OLMEDO F. Pseudothionin-Stl, a potato peptide active against potato pathogens. **Eur. J. Biochem.** v. 223, p. 135–139, 1994.

MOSIMANN S; MELESHKO R; JAMES M N. A critical assessment of comparative molecular modeling of tertiary structures of proteins. **Proteins**. v. 23, n. 3, p. 301-317, 1995.

MURAD A M; PELEGRINI P B; SIMONE M N; FRANCO O L. Novel Findings of Defensins and their Utilization in Construction of Transgenic Plants. **Transg. Plant J.** v. 1, n. 1, p. 39-48, 2007.

MURRAY D; DORAN P; MACMATHUNA P; MOSS A C. *In silico* gene expression analysis – an overview. **Mol. Cancer**. v. 6, p.50, 2007.

NCBI. Disponível em< http**://www.ncbi.nlm.nih.gov/protein**>. Acesso em: 12 nov 2009.

NELSON R J, NAYLOR R L, JAHN M M. The role of genomics research in improvement of "Orphan" crops. **Crop Science**. v 44, p. 1901-1904, 2004.

NEUMANN G M; CONDRON R; POLYA G M. Purification and mass spectrometry-based sequencing of yellow mustard (*Sinapis alba L.*) 6 kDa proteins. Identification as antifungal proteins. **Interact J. Pept. Protein Res.** v. 47, n. 6, p. 437-446, 1996.

NGAI P H; NG T B. *Coccinin*, an antifungal peptide with antiproliferative and HIV-1 reverse transcriptase inhibitory activities from large scarlet runner beans. **Peptides**. v. 25, n. 12, p. 2063-2068, 2004.

ODINTSOVA T I; ROGOZHIN E A; BARANOV Y; MUSOLYAMOV A K; NASSER Y, EGOROV T A;et al. Seed defensins of barnyard grass *Echinochloa crusgalli* (L.). **Beauv. Biochim.** v. 90, p.1667–1673, 2008.

ODINTSOVA T I; EGOROV T A. **In peptidomics: method and application**; SOLVIER M; SHAW C; ANDRÉN P EDS; WILEY J; SONS I N C. Inc: New Jersey, p. 99-117, 2007.

OKUDA S; TSUTSUIL H; SHIINAL K; SPRUNCK S;TAKEUCHIL H; YUIL R; KASAHARA R D; HAMAMURA Y;MIZUKAMI A;SUSAKI D; KAWANOL N; SAKAKIBARA T; NAMIKIL S; ITOH K; OTSUKA K; MATSUZAKI M; NOZAKI H; KUROIWA T; NAKANO A; KANAOKA M M; DRESSELHAUS T; SASAKIL N; HIGASHIYAMA T. Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. **Nature**. v. 458, p. 357-362, 2009.

OLLI S; KIRTI P B. Cloning, characterization and antifungal activity of defensina Tfgd1 from *Trigonella foenum-graecum* L. **J. Biochem. Mol. Biol.** v. 39, n. 3, p.278–283, 2006.

ONWULIRI A V, OBU A J. Lipids and other constituents of *Vigna unguiculata* and *Phaseolus vulgares* grown in northern Nigéria. **Food Chemistry**, Oxford, v. 78, n. 1, p. 1-7, 2002.

OSBORN R W; DE SAMBLANX G W; THEVISSEN K; GODERIS I; TORREKENS S; VAN LEUVEN F; ATTENBOROUGH S; REES S B; BROEKAERT W F. Isolation and characterization of plants defensins from seeds of *Asteraceae, Fabaceae, Hippocastanaceae* and *Saxifragaceae*. **FEBBS letters**. v. 368, p. 257-262, 1995.

OUYANG S; ZHU W; HAMILTON J; LIN H; CAMPBELL M; CHILDS KEVIN; THIBAUD-NIESSEN F; MALEK R L; LEE Y; ZHENG L; ORVIS J; HAAS B; WORTMAN J; BUELL C R. The TIGR Rice Genome Annotation Resource: improvements and new features. **Nucleic Acids Res.** v. 35, Database issue D883–D887, 2007.

PADOVAN L; SEGAT L; TOSSI A; CALSA Jr T; KIDO E A; BRANDÃO L; GUIMARÃES R L; PESTANA-CALSA M C; PANDOLFI V; BELARMINO L C; BENKO-ISEPPON A M; CROVELLA S. Characterization of a new defensin from cowpea (*Vigna unguiculata* (L.) Walp.). **Protein Pept. Lett.** 2010, *In press*.

PAPO N; SHAI Y. Host defense peptides as new weapons in cancer treatment. **CMLS Cell Mol. Life Sci.** v. 62, p. 784–790, 2005.

PARK H C; KANG Y H; CHUN H J; KOO J C; CHEONG Y H; KIM C Y; et al. Characterization of a stamen-specific cDNA encoding a novel plant defensin in Chinese cabbage. **Plant Mol. Biol.** v. 50, p. 59–69, 2002.

PARK Y S; JEON M H; LEE S H; MOON J S; CHA J S; KIM H Y; CHO T J. Activation of defense responses in Chinese cabbage by a no host pathogen, Pseudomonas syringae pv. Tomato. **J. Biochem. Mol. Biol.** v. 38, n. 6, p. 748-754, 2005.

PDB (Protein Data Bank). Disponível em: <http://www.rcsb.org/pdb/home/home.do>. Acesso em: 16 nov de 2009.

PELEGRINI P B; FRANCO O L. Plant γ-thionins: Novel insights on the mechanisms of a multi-functional class of defense proteins. **Int. J. Biochem. Cell Biol.** v. 37, p. 2239-2253, 2005.

PELEGRINI P B; LAY F T; MURAD A M; ANDERSON M A; FRANCO O L. Novel insights on the mechanism of action of α-amylase inhibitors from the plant defensin family. **Proteins**. v. 73, p. 719-729, 2008.

PELLEGRINI M; MARCOTTE E M; THOMPSON M J; EISENBERG D; YEATES T O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. **Proc. Natl. Acad. Sci.** v. 96, p. 4285–4288, 1999.

PENNINCKX I A M A; EGGERMONT K; TERRAS F R G; THOMMA B P H J; DE SAMBLANX G W. Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. **Plant Cell.** v. 8, p. 2309–2323, 1996.

PERVIEUX A I; BOURASSA M; LAURANSB F; HAMELINA R; SÉGUINA A. A spruce defensin showing strong antifungal activity and increased transcript accumulation after wounding and jasmonate treatments Isa. **Physiol. Mol. Plant Pathol.** v. 64, p. 331-341, 2004.

PIGNONE D, CIFARELLI S, PERRINO P. Chromosome identification in Vigna unguiculata (L.) Walp. In: NG N Q, MONTI L M, (Eds.), **Cowpea genetic resources**. IITA, Ibadan, Nigeria, 1990.

PILON-SMITS E. Phytoremediation. **Annu Rev. Plant Biol**. v. 56, p. 15–39, 2005.

PRETZER G; SNEL J; MOLENAAR D; WIERSMA A; BRON P A; LAMBERT J; DE VOS W M; VAN DER MEER R; SMITS M A; KLEEREBEZEM. Biodiversity-based identification and functional characterization of the mannose-specific adhesion of *Lactobacillus plantarum*. **J. Bacteriol**. v. 187, p. 6128–6136, 2005.

RAHMAN H; UCHIYAMA M; KUNO M; HIRASHIMA N; SUWABE K; TSUCHIYA T; KAGAYA Y; KOBAYASHI I; KAKEDA K; KOWYAMA Y. Expression of stigma- and anther-specific genes located in the S locus region of *Ipomoea trifida.* **Sex. Plant Reprod.**v. 20, n. 2, p. 73-85, 2007.

RAMAMOORTHY V; ZHAO X; SNYDER A K; XU J R; SHAH D M. Two mitogen-activated protein kinase signaling cascades mediate basal resistance to antifungal plant defensinsin Fusarium graminearum. **Cell Microbiol.** v. 9, p. 1491–1506, 2007.

ROST B; FARISELLI P; CASADIO R. Topology prediction for helical transmembrane proteins at 86% accuracy-Topology prediction at 86% accuracy. **Protein Sci.** v. 5, n. 8, p. 1704-1718, 1996.

SAITO T; KAWABATA S; SHIGENAGA T; TAKAYENOKI Y; CHO J; NAKAJIMA H. A novel big defensin identified in horseshoe crab hemocytes: isolation, amino acid sequence, and antibacterial activity. **J. Biochem.** v.117, p. 1131–117, 1995.

SAITOH H; KIBA A; NISHIAHARA M; YAMAMURA S; SUZUKI K; TERAUCHI R. Production of antimicrobial defensin in Nicotiana benthamiana with a potato virus X vector. **Mol. Plant Microbe Interact.** v. 14, n. 2, p. 111-115, 2001.

SANCHEZ R; SALI A. Comparative protein structure modeling as an optimization problem. **J. Mol. Struct.** (Theochem.) p. 398-399, p. 489-496, 1997.

SANGINGA N, et al. Sustainable resource management coupled to resilient germplasm to provide new intensive cereal–grain–legume–livestock systems in the dry savanna. **Agric Ecosys Environ**. v.100, p.305–314, 2003.

SATCHELL D P; SHEYNIS T; SHIRA FUJI Y; KOLUSHEVA S; QUELLETTE A J; JELINEK R. Interactions of mouse paneth cell alpha- defensins and alpha- defensins precursors with membranes. Prosegment inhibition of association with biomimetic membranes. **J. Biol. Chem.** v. 278, p. 13838-13846, 2003.

SATO S, NAKAMURA Y, ASAMIZU E, ISOBE S, TABATA S. Genome Sequencing and Genome Resources in Model Legumes. **Plant Physiol.** V 144, p. 588-593, 2007.

SATO T; IMANAKA H; RASHID N; FUKUI T; ATOMI H; IMANAKA T. Genetic evidence identifying the true gluconeogenic fructose-1,6-bisphosphatase in *Thermococcus kodakaraensis* and other Hyperthermophiles. **J. Bacteriol.** v. 186, p. 5799–5807, 2004.

SAUDER J M; ARTHUR J W; DUNBRACK R L Jr. Large-scale comparison of protein sequence alignment algorithms with structure alignments. **Proteins**. v. 40, n.1, p. 6-22, 2000.

SCHIBLI D J; HUNTER H N; ASEYEV V; STARNER T D; WIENCEK J M; MCCRAY P B; TACK B F; VOGEL H J. The solution structures of the human b-defensins lead to a better understanding of the potent bactericidal activity of HBD3 against Staphylococcus aureus. **J. Biol. Chem.** v. 277, p. 8279–8289, 2002.

SCHMITT A O; SPECHT T; BECKMANN G; DAHL E; PILARSKY C P; HINZMANN B; ROSENTHAL A. Exhaustive mining of EST libraries for genes differentially expressed in normal and tumour tissues**. Nucleic Acids Res.** v. 27, p. 4251-426, 1999.

SEEBAH S; SURESH A; ZHUO S; CHOONG Y H; CHUA H; CHUON D; BEUERMAN R; VERNA C. Defensins knowledgebase: a manually curated database and information source

focused on the defensins family of antimicrobial peptides. **Nucleic Acids Res.** v. 35, D265-268, 2007.

SEGURA A; MORENO M; MOLINA A; GARCÍA-OLMEDO F. Novel defensin subfamily from spinach (Spinacia oleracea). **FEBS Letter**. v. 435, p-159-162, 1998.

SELS J; DELAURÉ S L; AERTS A M; PROOST P; CAMMUE B P; DE BOLLE M F. Use of a PTGS-MAR expression system for efficient in planta production of bioactive *Arabidopsis thaliana* plant defensins. **Transgenic Res.** v. 16, n. 4, p. 531-538, 2007.

SENTHIL-KUMAR M; GOVIND G; KANG L; MYSORE K S; UDAYAKUMAR M. Functional characterization of *Nicotiana benthamiana* homologs of peanut water deficit-induced genes by virus-induced gene silencing. **Planta**. v. 225, n. 3, p. 523-539, 2007.

SHARMA P; LÖNNEBORG A. Isolation and characterization of a cDNA encoding a plant defensin-like protein from roots of Norway spruce. **Plant Mol. Biol.** v. 31, n. 3, p. 707-712, 1996.

SHEN G; PANG Y; WU W; MIAO Z; QIAN H; ZHAO L. Molecular cloning, characterization and expression of a novel jasmonate-dependent defensina gene from *Ginkgo biloba*. **J. Plant Physiol.** v.162, p.1160–1168, 2005.

SHIAU Y-S; HORNG S-B; CHEN C-S; HUANG P-T; LIN C; HSUEH Y-C; LOU K-L. Structural analysis of the unique insecticidal activityof novel mungbean defensin VrD1 reveals possibility of homoplasy evolution between plant defensinsand scorpion neurotoxins. **J. Mol. Recognit.** v.19, p.441-450, 2006.

SILVERSTEIN K A T; GRAHAM M A; PAAPE T D; VAN DEN BOSCH K A. Genome organization of more than 300 defensin-like genes in Arabidopsis. **Plant Physiol.** v. 138, p. 600 e 610, 2005.

SILVERSTEIN K A; MOSKAL W A Jr; WU H C; UNDERWOOD B A; GRAHAM M A; TOWN C D; VANDENBOSCH K A. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. **Plant J.** v. 51, p. 262-280, 2007.

SIMON M V, BENKO-ISEPPON A M, RESENDE L V, WINTER P, KAHL G. Genetic diversity and phylogenetic relationships among Vigna Savi germplasm revealed by DNA amplification fingerprinting. **Genome**. V 50, p. 538-547, 2007.

SINGH B B, EHLERS J D, SHARMA B, FREIRE FILHO F R. Recent progress in cowpea breeding. In: FATOKUN C A, TARAWALI S A, SINGH B B, KORMAWA P M, TAMO M. (Eds.), Challenges and opportunity for enhancing sustainable cowpea production. **Proceedings of the World Cowpea Conference III held at International Institute of Tropical Agriculure (IITA)**, Ibadan, Nigeria, 4–8 September 2000. IITA, Ibadan, Nigeria, 2002.

SNEL B; HUYNEN M A. Quantifying modularity in the evolution of biomolecular systems. **Genome Res.** v. 14, p. 391–397, 2004.

SOLIS J; MEDRANO G; GHISLAIN M. Inhibitory effect of a defensin gene from the Andean crop maca (*Lepidium meyenii*) against *Phytophthora infestans*. **J. Plant Physiol.** v. 164, p.1071–1082, 2007.

SONG X; ZHOU Z; WANG J; WU F; GONG W. Purification, characterization and preliminary crystallographic studies of a novel plant defensina from *Pachyrrhizus erosus* seeds. **Acta Crystallogr.** v. 60, p. 1121-1124, 2004.

SONG X; WANG J; WU F; LI X; TENG M; GONG W. cDna cloning, functional expression and antifungal activities of a dieric plant defensins SPE10 from *Pachyrrhizus erosus* seeds. **J. Plant Mol. Biol.** v. 57, n. 1, p. 13-20, 2005.

SPELBRINK R G; DILMAC N; ALLEN A; SMITH T J; SHAH D M; HOCKERMAN G H. Differential antifungal and calcium channel-blocking activity among structurally related plant defensins. **Plant Physiol.** v.135, p. 2055–2067, 2004.

SPIRIN V; GELFAND M S; MIRONOV A A; MIRNY L A. A metabolic network in the evolutionary context: musticale structure and modularity. **Proc. Natl. Acad.** v. 103, p. 8774-8779, 2006.

STANCHEV B S; DOUGHTY J; SCUTT C P; DICKINSON H; CROY R R. Cloning of PCP1, a member of a family of pollen coat protein (PCP) genes from *Brassica oleracea* encoding novel cysteine-rich proteins involved in pollen-stigma interactions. **Plant J.** v. 10, n. 2, p. 303-313, 1996.

STEIN L. Genome annotation: from sequence to biology. **Nat. Rev. Genet.** v. 2, p. 493–503, 2001.

STIEKEMA W J; HEIDEKAMP F; DIRKSE W G; VAN BECKUM J; DE HAAN P; TEN BOSCH; LOWERSE J D. Molecular cloning and analysis of four potato tuber mRNAs. **Plant Mol. Biol.** v. 11, p. 255-269, 1988.

STUART J M; SEGAL E; KOLLER D; KIM S K. A gene coexpression network for global discovery of conserved genetic modules. **Science**. v. 302, p. 249–255, 2003.

SUCEST(Sugarcane EST project). Disponível em:<**http://sucest.lad.ic.unicamp.br/en/**> Acesso em: 28 out.2009.

SUN J; XU J; LIU Z; LIU Q; ZHAO A; SHI T; LI Y. Refined phylogenetic profiles method for predicting protein–protein interactions. **Bioinformatics**. v. 21, p. 3409–3415, 2005.

SUNAGAWA S; DESALVO M K, VOOLSTRA C R, REYES-BERMUDEZ, MEDINA M. Identification and Gene Expression Analysis of a Taxonomically Restricted Cysteine-Rich Protein Family in Reef-Building Corals. **PLoS ONE**. v. 4, n.3: e4865. doi:10.1371/journal.pone.0004865, 2009.

SWARBRECK D; WILKS C; LAMESCH P; BERARDINI T Z; GARCIA-HERNADEZ M; FOERSTER H; LI D; MEYER T; MULLER R; PLOETZ L; RADENBAUGH A; SINGH S; SWING V; TISSIER C; ZHANG P; HUALA E. The Arabidopsis Information Resource

(TAIR): gene structure and function annotation. **Nucleic Acids Res.** v. 36, Database issue D1009–D10142008, 2008.

SWISS-PROT -Protein knowledgebase; TrEMBL -Computer-annotated supplement to Swiss-Prot. Disponível em: <http://www.expasy.org/sprot/> acesso em: 8 nov de 2009.

TERNES P; SERLING P; ALBRECHT S; FRANKE S; CREGG J M; WARNECKE D; HEINZ E. Identification of fungal sphingolipid C9-methyltransferases by phylogenetic profiling. **J. Biol. Chem.** v. 281, p. 5582–5592, 2006.

TERRAS F R G; EGGERMONT K; KOVALEVA V; RAIKHEL N V; OSBORN R W; KESTER A; REES S; TORREKEMD S;VAN LEUVEN F; VANDERLEYDEN J; CAMMUE B P A; BROEKART W F. Small Cysteine-Rich Antifungal Proteins from Radish: Their Role in Host Defense. **Plant Cell.** v. 7, p. 573-588, 1995.

TERRAS F R G; PENNINCKX I A M A; GODERIS I J; BROEKAERT W F. Evidence that the role of plant defensins in radish defense responses is independent of salicylic acid. **Planta**. v. 206, p. 117–24, 1998.

TERRAS F R G; SCHOOFS H M E; DE BOLLE M F C; VAN LEUVEN F; REES S B; VANDERLEYDEN J; et al. Analysis of two novel classes of plant antifungal proteins from radish (Raphanus sativus L.) seeds. **J. Biol. Chem.** v. 267, n. 22, p. 15301–15309, 1992.

TERRAS F R G; TORREKENS S; VAN LEUVEN F; OSBORN R W; VANDERLEYDEN J; CAMMUE B P A; et al. A new family of basic cysteine-rich plant antifungal proteins from *Brassicaceae* species. **FEBS**. v. 316, n. 3, p. 233–240, 1993.

TINCU J A; TAYLOR S W. Antimicrobial Peptides from Marine Invertebrates. **Antimicrobial agents and chemotherapy**. v. 48, n. 10, p. 3645-3654, 2004.

THEVISSEN K; FRANCOIS I E J A; TAKEMOTO J Y; FERKET K K A; MEERT E M K; CAMMUE B P A. DmAMP1, an antifungal plant defensin from dahlia (*Dahlia merckii*), interacts with sphingolipids from Saccharomyces cerevisiae. **FEMS Microbiol. Lett.** v. 226, p. 169–173, 2003.

THEVISSEN K; GHAZI A; DE SAMBLANX G W; BROWNLEE C; OSBORN R W;BROEKAERT W F. Fungal membrane responses induced by plant defensins and thionins. **J. Biol. Chem.** v. 271, p.15018–15025, 1996.

THEVISSEN K; IDKOWIAK-BALDYSB J; IMB Y –J; TAKEMOTOB J; FRANCOIS I E J A; KATHELINE K A; FERKETA; AERTA A M; MEERTA E M K; WINDERICKXC J; ROOSENC J; CAMMUE B PA. SKN1, a novel plant defensin-sensitivity gene in *Saccharomyces cerevisiae*, is implicated in sphingolipid biosynthesis. **FEBS Letters**. v. 579, p. 1973–1977, 2005.

THEVISSEN K; KRISTENSEN H-H; THOMMA B P H J; CAMUE B P A; FRANÇOISE I E J A. Therapeutic potencial of antifungal plant and insect defensins. **Drug Discov. Today.** v. 12, n. 21/22, p. 966-971, 2007.

THEVISSEN K; TERRAS F R; BROEKAERT W F.Permeabilization of fungal membranes by plant defensins inhibits fungal growth. **Appl. Environ. Microbiol.** v. 65, p. 5451–5458, 1999.

THOMMA B P H J; BROEKAERT A W F. Tissue-specific expression of plant defensin genes *PDF2.1* and *PDF2.2* in *Arabidopsis thaliana* .**Plant Physiol. Biochem.** v. 36, p. 533-537, 1998.

THOMMA B P H J; CAMMUE B P A; THEVISSEN K. Plant defensins. **Planta**. v. 216, p. 193 e 202, 2002.

THOMMA B P H J; EGGERMONT K; PENNINCKX I A M A; MAUCH-MANI B; VOGELSANG R; CAMMUE B P A; BROEKAERT W F. Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. **Proc. Natl. Acad. Sci.** v. 95, p. 15107-15111, 1998.

TIMKO M P, RUSHTON P J, LAUDEMAN T W, BOKOWIEC M T, CHIPUMURO E, CHEUNG F, TOWN C D, CHEN X. Sequencing and analysis of the gene-rich space of cowpea. **BMC Genomics**. v. 9, p.103, 2008.

TREGEAR J M; MORCILLO F; RICHAUD F; BERGER A; SINGH R; CHEAH S C. Characterization of a defensin gene expressed in oil palm inflorescences: induction during tissue culture and possible association with epigenetic somaclonal variation events. **J. Exp. Bot.** v. 53, n. 373, p.1387–1396, 2002.

UMEMURA K; OGAWA N; YAMAUCHI T; IWATA M; SHIMURA M; KOGA J. Cerebroside elicitors found in diverse phytopathogens activate defense responses in rice plants. **Plant Cell Physiol.** v. 61, p. 676-683, 2000.

UNICA (União da Indústria da Cana-de-Açúcar). Disponível em:<**http://www.unica.com.br/>** Acesso em: 03 nov. 2009.

URDANGARÍN M C; NORENO N S; BROEKAERT W F; DE LA CANAL L. A defensin gene expressed in sunflower inflorescence. **Plant Physiol. Biochem.** v. 38, n. 3, p. 253-258, 2000.

VAN BAARLEN P; VAN BELKUM A; SUMMERBELL R C; CROUS P W; THOMMA B P. Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? **FEMS Microbiol. Rev.** v. 31, p. 239–277, 2007.

VAN DEN HEUVEL K J; HULZINK J M; BARENDSE G W; WULLEMS G J. The expression of tgas118, encoding a defensin in *Lycopersicon esculentum*, is regulated by gibberellins. **J. Exp. Bot.** v. 52, n. 360, p. 1427-1436, 2001.

VAN DER WEERDEN N L; LAY F T; ANDERSON M A. The plant defensin, NaD1, enters the cytoplasm of *Fusarium oxysporum hyphae*. **J. Biol. Chem.** v. 283, n. 21, p. 14445–14452, 2008.

VAN DJIK A; VELDHUIZEN E J A; HAAGSMAN H P. Avian defensis. **Vet. Immunol Immunopathol**. v.124, p.1–18, 2008.

VAN NOORT V; SNEL B; HUYNEN M A. Predicting gene function by conserved co-expression. **Trends Genet.** v. 19, p. 238–242, 2003.

VANOOSTHUYSE V; MIEGE C; DUMAS C; COCK J M. Two large Arabidopsis thaliana gene families are homologous to the *Brassica* gene superfamily that encodes pollen coat proteins and the male component of the self-incompatibility response. **Plant Mol. Biol.** v. 46, n. 1, p. 17-34, 2001.

VASMATZIS G; ESSAND M; BRINKMANN U; LEE B; PASTAN I. Discovery of three genes specifically expressed in human prostate by expressed sequence tag database analysis**.** **Proc. Natl. Acad. Sci.** v. 96, p. 300-304, 1998.

VELCULESCU V E; ZHANG L; VOGELSTEIN B; ZINZLER K W. Serial analysis of gene expression. Science. v. 270, n. 5235, p. 484-487, 1995.

VENCLOVAS C; ZEMBLA A; FIDELIS K; MOULT J. Assessment of progress over the CASP experiments. **Proteins**. v. 53, Suppl. 6, p.585-595, 2003.

VETTORE A L; DA SILVA F R; KEMPER E L; ARRUDA P. The libraries that made SUCEST. **Gen. Mol. Biol.** v. 24, p. 1–7, 2001.

VIJYAN S; GURUPRASAD L; KIRTI P B. Prokaryotic expression of a constitutively expressed *Tephrosia villosa* defensin and its potent antifungal activity. **Appl Microbiol Biotechnol**. v. 80, n. 6, p.1023-1032, 2008.

WAIN H M; LUSH M; DUCLUZEAU F; POVEY S. Genew: the human nomenclature database. **Nucleic Acids Res.** v. 30, p. 169–171, 2002.

WANG C; KAAS Q; CHICHE L; CRAIK D L. CyBase: a database of cyclic protein sequences and structures with applications in protein discovery and engineering. **Nucleic Acids Res.** v. 36, D206-210, 2008a.

WANG C; WAN S; XIANG Z; SHI Y. Incorporating Hydration Force Determined by Boundary Element Method into Stochastic Dynamics. **J. Physiol. Chem.** v. 101, p. 230-235, 1997.

WANG H X; NG T B. An antifungal peptide from baby lima bean. **Appl. Microbiol. Biotechnol.** v. 73, p. 576–581, 2006.

WANG H X; NG T B. Isolation and characterization of an antifungal peptide with antiproliferative activity from seeds of Phaseolus vulgaris cv. ‘Spotted Bean’. **Appl. Microbiol. Biotechnol.** v. 74, p.125–130, 2007.

WANG G; LI X; WANG Z. APD2: the updated antimicrobial peptide database and its application in peptide design. **Nucleic Acids Res.** v. 37, D933-D937, 2008b.

WANG S; RAO P; Ye X. Isolation and biochemical characterization of a novel leguminous defense peptide with antifungal and antiproliferative potency. **Appl. Microbiol. Biotechnol.** doi:10.1007/s00253-008-1729-2, 2008.

WANG Y; NOWAK G; CULLEY D; HADWIGER L A; FRISTENSKY B. Constitutive expression of pea defense gene DRR206 confers resistance to blackleg (*Leptosphaeria maculans*) disease in transgenic canola (*Brassica napus*). **Mol. Plant Microb. Interact**. v. 12, p. 410–418, 1999.

WANG Z; WANG G. A.P.D: the Antimicrobial Peptide Database. **Nucleic Acids Res.** v. 32, D590-D592, 2004.

WEBER I T. Evaluation of homology modeling of HIV Protease. **Proteins**. v. 7, n. 2, p. 172-184, 1990.

WHEELER D L; BARRETT T; BENSON D A; BRYANT S H; CANESE K; CHURCH D M; DICUCCIO M; EDGAR R; FEDERHEN S; HELMBERG W, D.A; et al. Database resources of the National Centre for Biotechnology Information. **Nucleic Acids Res.** v. 33, D39–D45, 2005.

WHITTALL J B; MEDINA-MARINHO A; ZIMMER E A; HODGES S A. Generating single-copy nuclear gene data for a recent adaptive radiation. **Mol. Phylogenet. Evol.** v. 39, p. 124-134, 2006.

WIJAYA R; NEUMANN G M; CONDRON R; HUGHES A B; POLY G M. Defense proteins from seed of *Cassia fistula* include a lipid transfer protein homologue and a protease inhibitory plant defensin. **Plant Sci.** v. 159, p. 243–255, 2000.

WISNIEWSKI M E; BASSET C L; ARTLIP T S; WEBB R P; JANISIEWICZ W J; NORELLI J L. Characterization of a defensin in bark and fruit tissues of peach and antimicrobial activity of a recombinant defensin in the yeast, Pichia pastoris. **Physiol. Plant.** v. 119, p. 563–572, 2003.

WONG J H; NG T B. Sesquin, a potent defensin-like antimicrobial peptide from ground beans with inhibitory activities toward tumor cells and HIV-1 reverse transcriptase. **Peptides**. v. 26, p. 1120–1126, 2005a

WONG J H; NG T B. Vulgarinin, a broad-spectrum antifungal peptide from haricot beans (*Phaseolus vulgaris*). **IJBCB**. v. 37, p.1626–1632, 2005b.

WONG J H; XIA L; NG T B. A Review of Defensins of Diverse Origins. **Curr. Protein Pept. Sci.** v. 8, p. 446-459, 2007.

WONG J H; ZHANG X Q; WANG H X; NG T B. Amitogenic defensin from white cloud beans (*Phaseolus vulgaris*). **Peptides**. v. 27, p. 2075–2081, 2006.

XIANG Z; XU Y; SHI Y. Solving the finite-difference, nonlinear, Poisson-Boltzmann equation under a linear approach. **J. Comp. Chem.** v. 16, n. 2, p. 200-206, 1995.

XIAO H; JIANG N; SCHAFFNER E; STOCKIGER E J; VAN DER KNAAP E. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. **Science**. v. 319, n. 5869, p. 1527-1530, 2008.

YAMADA S; KOMORI T; IMASEKI H. cDNA cloning of γ-thionin from *Nicotiana excelsior* (accession no. AB005266; PGR97–131). **Plant Physiol.** v. 115, p. 314, 1997.

YANG A S; HONIG B. An integrated approach to the analysis and modeling of protein sequences and structures. I. Protein structural alignment and a quantitative measure for protein structural distance. **J. Mol. Biol.** v. 301, n. 3, p. 665-678, 2000.

YANG D;LIU Z –H, TEWARY P, CHEN Q, DE LA ROSA G; OPPENHEIM J J. Defensin Participation in Innate and Adaptive Immunit**y. Curr. Pharm. Des.** v. 13, p. 3131-3139, 2007.

YASUDA M; TAKESUE F; INUTSUKA S; HONDA M N T; KORENAGA D. Overexpression of cyclin B1 in gastric cancer and its clinicopathological significance: an immunohistologicalstudy. **J. Cancer Res. Clin. Oncol.** v. 128, p.412–6, 2002.

YE X Y; NG T B. A new antifungal peptide from rice beans. **J. Peptide Res.** v. 60, p. 81–87, 2002.

YE X Y; NG T B. Peptides from pinto bean and red bean with sequence homology to cowpea 10-kda protein precursor exhibit antifungal, mitogenic, and HIV-1 reverse transcriptase-inhibitory activities. **Biochem. Biophys. Res. Commun.** v. 285, p.424–429, 2001.

YODER W T; LEHMBECK J. Heterologous expression and protein secretion in filamentous fungi. **In Advances in Fungal Biotechnology for Industry, Agriculture, and Medicine** (Tkacz, J.S. and Lange, L., eds), p. 201–219, Kluwer Academic/Plenum Publishers, 2004.

YOKOYAMA S; KATO K; KOBA A; MINAMI Y; WATANABE K; YAGI F. Purification, characterization, and sequencing of antimicrobial peptides, Cy-AMP1, Cy-AMP2, and Cy-AMP3, from the Cycad (*Cycas revoluta*) seeds. **Peptides**. v. 29, p. 2110-2117, 2008.

ZASLOFF, M. Antimicrobial peptides of multicellular organisms. **Nature**. v. 415, p. 389-395, 2002.

ZÉLICOURT A; LETOUSEY P; THOIRON S; CAMPION C; SIMONEAU P; ELMORJANI K; et al. Ha-DEF1, a sunflower defensina, induces cell death in Orobanche parasitic plants. **Planta**. v. 226, p. 591-600, 2007.

ZHANG H; YOSHIDA S; AIZAWA T; MURAKAMI R; SUZUKI M; KOGANEZAWA N; MATSUURA A; MIYAZAWA M; KAWANO K; NITTA K; KATO Y. In Vitro Antimicrobial Properties of Recombinant ASABF, an Antimicrobial Peptide Isolated from the Nematode *Ascaris suum.* **Antimicrob. agents chemother.** v. 44, n. 10, p. 2701-2705, 2000.

ZHANG Y; LEWIS K. Fabatins: new antimicrobial plant peptides. **FEMS Microbiology Letters**. v. 149, n. 1, p. 59-64, 1997.

ZHOU Y; NORIOKA N; LI S; NORIOKA S. Nucleotide Sequence of a cDNA Encoding Plant Defensin-Like Protein(PDN1) in Pollen Grains of Japanese Pear (*Pyrus pyrifolia*). **J. Plant Physiol. Mol. Biol.** v. 28, n. 407, 2002.

ZHU S. Evidence for myxobacterial origin of eukaryotic defensins. **Immunogenetics**. v. 59, p. 949–954, 2007.

ZHU S. Discovery of six families of fungal defensin-like peptides provides insights into origin and evolution of the CSαβ defensins. **Molecular Immunology.** v. 45, p. 828-838, 2008.

ZOU J; MERCIER C; KOUSSOUNADIS A; SECOMBES C. Discovery of multiple beta-defensin like homologues in teleost fish. **Molecular Immunology**. v. 44, p. 638–647, 2007.

Apêndices

Manuscrito de artigo científico 1

Padovan *et al*., 2010

Caracterização of a new defensin from cowpea (Vigna unguiculata (L.) Walp.)

Protein and Peptide Letters, Aceito em Maio de 2009

Characterization of a new defensin from cowpea (*Vigna unguiculata* (L.) Walp.)

Lara Padovan*[1,2], Ludovica Segat[2], Alessandro Tossi[1], Tercilio Calsa Jr.[4], Ederson Akio Kido[4], Lucas Brandao[3], Rafael L. Guimarães[3], Valesca Pandolfi[4], Maria Clara Pestana-Calsa[4], Luís Carlos Belarmino[4], Ana Maria Benko-Iseppon[4], Sergio Crovella[2,4].

Affiliations

1. Department of Life Sciences, University of Trieste - Via Giorgieri, 1 - 34127 TRIESTE, Italy
2. Genetic Service, IRCCS Burlo Garofolo and Department of Developmental and Reproductive Sciences, University of Trieste -Via dell'Istria, 65/1 - 34137 TRIESTE, Italy
3. Laboratório de Imunopatologia Keizo Azami (LIKA) Federal University of Pernambuco Campus Universitário, Cidade Universitária CEP: 50670-901 Recife - PE – Brazil
4. Universidade Federal de Pernambuco, Centro de Ciências Biológicas, Departamento de Genética, Recife, PE – Brazil

Corresponding author:
Lara Padovan
Genetic Service, IRCCS Burlo Garofolo and Department of Developmental and Reproductive Sciences, University of Trieste -Via dell'Istria, 65/1 - 34137 TRIESTE, Italy
Tel. +39.040.3785422
Fax. +39.040.3785540
e-mail: padovan@burlo.trieste.it

Abstract

Using available Phaseoleae defensins in databases a putative defensin gene was isolated in cowpea (*Vigna unguiculata* (L.) Walp.) and cloned from genomic cowpea DNA. The putative mature defensin sequence displays the characteristic defensins residues arrangement, secondary and tertiary structures were predicted and splicing analysis was performed. Using RT-PCR, defensin expression and differences in response to biotic stimuli between infected and non infected plants were tested.

Key words: antimicrobial peptides, cowpea mosaic virus, evolution, Fabaceae, plant defensin, protein function.

GeneBank accession number for *Vigna unguiculata* defensin gene sequence: FJ94789.

Abbreviations:
EST = Expressed Sequences Tags
PDEF_VIGUN = Vigna unguiculata defensin
PCR = Polymerase Chain Reaction
RT-PCR = Reverse Transcription Polymerase Chain Reaction
CPSMV = Cowpea Severe Mosaic Virus

Introduction

Antimicrobial peptides are important components of the innate immune systems of all living organisms and are widely distributed throughout the plant kingdom [1, 2, 3]. In particular plant defensins play an important role in host defense against pathogens, having a direct antimicrobial effect against both fungi and bacteria [1]. Defensins are cysteine-rich peptides consisting of 45-54 amino acids and characterised by the presence of conserved intramolecular disulfide bonds [4, 5]. The three-dimensional structure of plant defensins presents a triple-stranded, antiparallel β-sheet platform overlaid by a single α-helix and is stabilised by four disulfide bonds. It closely resembles defensin structures observed in vertebrate and invertebrate animals as well as moulds [1, 6, 7].

The first plant defensin was isolated from wheat in 1990 [8] and since then the number of identified plant defensins number has rapidly increased [6]. In particular, defensins from several leguminous species were reported, such as *Phaseolus vulgaris* [9], *Pisum sativum* [5], *Vicia faba* [10] and *Clitoria ternatea* [11].

We report herein the characterisation of a putative defensin in *Vigna unguiculata* (L.) Walp. (cowpea), a leguminous plant grown extensively as a food and fodder crop in West Africa, north-eastern Brazil, part of the Middle East and southern regions of North America [12, 13]. In *V. unguiculata*, the structures and activities of two other plant defensins (formerly also known as thionins) named Cp thionins I II have been already described [14, 15]. We have identified another region in the cowpea genome codifying for a novel defensin, with a different primary structure to those previously reported, but sharing with them the distinctive features of plant defensins. This study confirms that multiple defensins act within the host defence pathways of important legume food crops, in this as well as in other cultivated plants species (such as sugarcane, maize and rice).

Materials and Methods

Computational search for new defensins: sequences from cowpea (*V. unguiculata* Taxonomy ID 3917 GenBank) and other Phaseoleae were obtained from the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/sites/entrez). BLAST searches were performed at the Gene Index Databases (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi). Cowpea sequence tags (EST) that could refer to a defensin were not found in Gene Index databases, so common bean (*Phaseolus vulgaris* that belongs to the same tribe as *V. unguiculata*) was chosen as reference and its ESTs were blasted against Phaseoleae sequences present in databases.

Computational searches showed interesting similarities between *P. vulgaris* TC147 EST (a cluster from Gene Index Databases) and sequences codifying for defensins that belong to other plants families, in particular sunflower (*Helianthus annuus*, AF141131 GenBank) and peach (*Prunus persica*, AY078426 GenBank) defensins. Thus, TC147 was chosen for primers design to perform a heterologous PCR on *V. unguiculata*.

Primer design: primers were selected by using “Primer 3” software (http://frodo.wi.mit.edu) according to default parameters and were: primer forward: 5’-TCCATGGCTCGCTCTGTGTCTT-3’ and reverse 5’-TGAAGTTTTAACAGTGTTTGGTGCACAAG-3'.

Genomic DNA and RNA extraction: Genomic DNA was extracted from young leaves of the Brazilian cultivar BR14-Mulato (kindly provided by EMBRAPA-CPAMN, Teresina, Piauì State), using a modified CTAB (cetyl-trimethyl-ammoniumbromide)-based protocol [16]. DNA concentration was determined electrophoretically using known amounts of phage Lambda/Hind III DNA marker as reference.

Total RNA was extracted from 30-days old plants of a unique CPSMV (Cowpea Severe Mosaic Virus) resistant genotype, named BRIT63, derived from a breeding cross (BR14-Mulato x IT85F-2687) grown in the greenhouse and either native or inoculated using a standard protocol [17]. Leaves were harvested 30, 60, 90 min and 16 h after mechanical wounding with Carborundum™ and virus inoculation . Negative controls consisted in BRIT63 and IT85F (the CPSMV-susceptible parental cultivar) samples, both neither infected nor mechanically injured. Harvested plant material was immediately frozen in liquid nitrogen and stored at -80ºC until RNA extraction [18].

PCR: Genomic DNA of *V. unguiculata* was amplified in a 25 µL volume containing 50 ng of DNA template; 6 mmol $MgCl_2$, 200 mmol dNTPs, 10 pmol of each primer, 1.0 U Taq DNA polymerase (AmpliTaq Gold – PE Biosystem) and 2.5 L buffer 10x (500 mM KCl, 100 mM Tris-HCl pH 8.3, 1.4 mM $MgCl_2$). Amplification was carried out in an automated thermal cycler (Applied Biosystem) according to the following program: an initial denaturation at 95°C for 10 min, after which 40 cycles of: denaturation (20 s at 95°C), primer annealing (30 s at 58°C) and primer extension (30 s at 72°C) were performed, followed by a final extension at 72°C for 7 min. PCR products were visualized on 2% agarose gel.

Sequences analysis: Sequences were analysed with SeqMan Genome Assembler (DNASTAR, Inc., Madison, WI, USA), multiple alignments of the nucleotide and amino acid sequences were performed using the

CLUSTALW software (http://www.ebi.ac.uk/Tools/clustalw). The candidate defensin nucleotide sequence (FJ94789 GenBank) was then analysed and translated into the corresponding amino acid sequence using the GENSCAN package (http://genes.mit.edu/GENSCAN.html) and Expasy software (http://www.expasy.org/cgi-bin/dna_aa). The obtained sequence was blasted against available plant defensins in databases using both the NCBI BLAST service (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and FASTA Nucleotide Similarity Search at EMBL-EBI (http://www.ebi.ac.uk/fasta33/nucleotide.html).
Disulfide-bonds formation was predicted using the DISULFIND program at the PredictProtein server http://www.predictprotein.org.
Structure prediction was carried out by identifying equivalent sequence positions in the closest plant defensin structure present in the PDB databank: 1GTP belonging to a barley defensins (http://www.rcsb.org/pdb/home/home.do).
Gene structure comparison: The predicted nucleotide sequence encoding for PDEF_VIGUN protein was aligned against the genome of *Arabidopsis thaliana* (TAIR; http://www.arabidopsis.org) and *Glycine max* (PHYTOZOME http://www.phytozome.net). Additionally, the predicted protein was blasted against BAC sequences from *Medicago truncatula* and *Lotus japonicus* (http://www.kazusa.or.jp). The genome location with best match was excised and compared to the genomic sequence of *V. unguiculata*. For evaluation of the splicing patterns, the exon/intron boundaries of *A. thaliana* and *G. max* were used, since these are well defined in the databanks. Nevertheless it was performed a Blastx search against the nr protein databank and megaBlast against EST database (dbEST and NordEST, a Brazilian cowpea EST database initiative) to confirm these boundaries and to identify the putative boundaries from *V. unguiculata*, *M. truncatula* and *L. japonicus*. Subsequently, the nucleotide sequences were aligned with ClustalW and manually adjusted to fit the ATG initiation codon of the ORFs, as well as the remaining coding positions in the first and second exons, as observed in the Blastx step. Signal peptide and the probable cleavage site was identified with SignalP [19].

RT-QPCR: Retro-transcription and quantitative amplification were performed by using the RT-PCR Core kit (Applied Biosystems, Foster City, CA) following the manufacturer's protocol. Real-time PCR was performed on the obtained cDNA by using the SYBR Green I chemistry and Corbett Rotor Gene 3000 real time PCR platform (Corbett Research). After an initial denaturation step at 95°C for 10 min, forty cycles were performed by two steps at 95°C for 16 s, and 60°C for 1 min. We used PCR primers designed in the exons flanking regions: primer forward 5'-CTGGTGGCCACTGAGATG-3' and reverse 5'-TCACACATGGTCCCTTGAAA-3' in order to avoid genomic DNA amplification. Plant defensin primers concentrations were optimized in order to determine the minimum concentration required to give the maximal signal (ΔRn), the lowest threshold cycle (Ct) and minimizing non-specific amplification.

The specificity of defensin amplification was verified by melting temperature assay (MTA) protocol including a slow cooling from 95°C to 60°C in 20 min at the end of PCR. The Corbett Rotor Gene 3000 MTA software allowed the identification of the point at which the re-association occurred (flexus point) as well as the melting temperature and the melting curve.

Results and Discussion

Nucleotide Sequence Analysis

By performing a heterologous PCR on *V. unguiculata*, using primers designed from *P. vulgaris* TC147 EST, we have amplified a region (so on named PDEF_VIGUN) that shows a 150 bp insertion in comparison with the reference sequence TC147. By using the GENSCAN program we described the putative organisation of the genomic region we amplified (Figure (**1**)): the sequence is expected to account for two exons (namely exon 1 and exon 2), separated by an intron (the additional 150 bp). The position of the intron, localized among the signal peptide, is consistent with other genomic defensins that have already been isolates both in *V. unguiculata* (VuD1) (Pelegrini 2008) and, among other plants, such as *P. inflata* (Karunanandaa 1994), *C. annuum* (Houlne, 1998) and *S. officinarum* (Padovan 2008).
PDEF_VIGUN nucleotide sequence was then blasted against EMBL and NCBI databases and a similarity higher than 98% emerged with some EST sequences from *V. unguiculata* and *P. vulgaris* common bean; moreover PDEF_VIGUN shows significant similarity with sequences that codify for defensins, thionins and protease inhibitors belonging to different plants families; this suggests to us that the amplified sequence codifies for a defensin and/or might act as a protease inhibitor. In fact the highest similarity among deposited nucleotide sequences emerged with a protease inhibitor from *G. max* (U12150 GenBank) (91% similarity); this plant shares with *V. unguiculata* not only the same family (Fabaceae) but also the same tribe "Phaseoleae". It is noteworthy that *G. max* is the only plant sequence belonging to Phaseoleae that shows high identity with *V. unguiculata* nucleotide sequence (Figure (**2**)). PDEF_VIGUN was compared with the other *Vigna* defensins available in the genomic databases and few homologies were found.

In fact VuD1 partial cds show only little similarity with PDEF_VIGUN nulceotide sequence; we then extended the analysis to the defensins belonging to other *Vigna* species such as *V.radiata* and *V. Nakashimae* and the highest similarity (score 55) was found between PDEF_VIGUN and: *V. radiata* "D1" (FJ591131 GenBAnk), *V. radiata Cys-rich protein* (VrCRP) (AF326687 GenBank), *V. radiata* "VrD1" (AY437639 GenBank) and *V. nakashimae* "Nak95" (AY856095 GenBank). However these similarities are lower than those found between *V. unguiculata* and plants belonging to other families, in particular *V. unguiculata* and *G. max* (score 91). This high variability among defensins in *Vigna* nucleotide sequences suggest how multiple peptides contribute to the innate immunity of this plant.

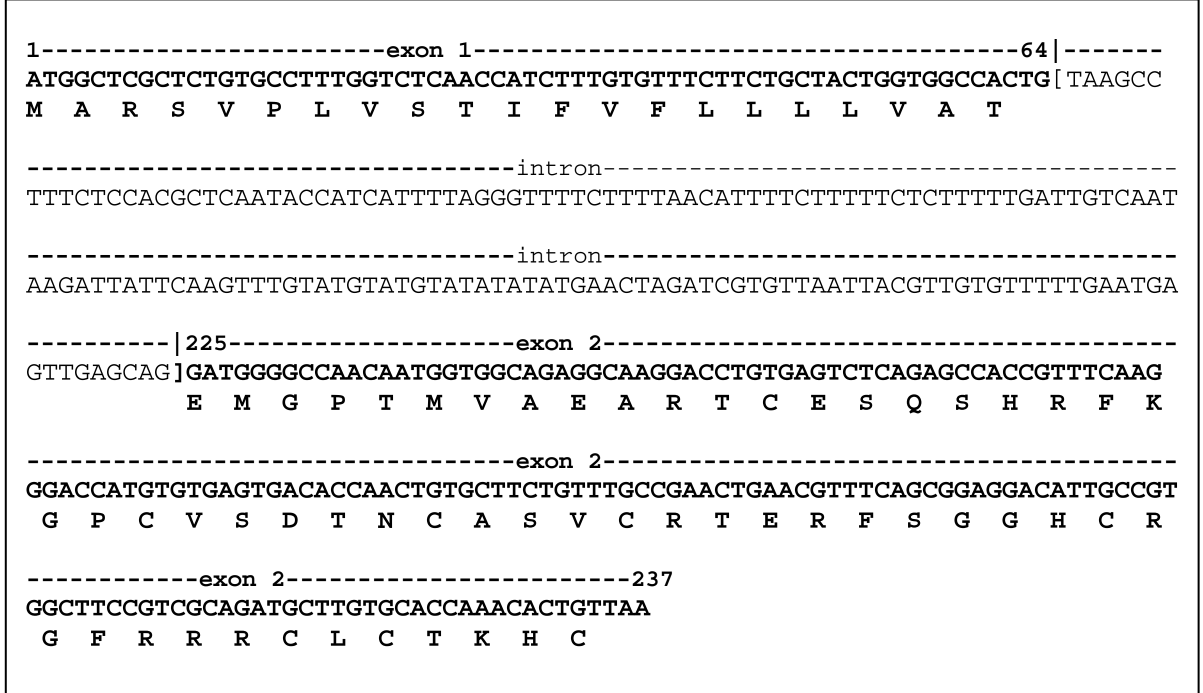


Figure 1. Nucleotide sequence (FJ94789 GenBank) with predicted genes/exons and putative amino acid sequences of PDEF_VIGUN. The figure was prepared using GENSCAN.

Among the other isolated defensins in plants, *Prunus persica* (AY078426 GenBank) defensin shows the highest similarity (77%) with the amplified sequence PDEF_VIGUN (Figure (**2**)), while for the other matched defensins from *Petunia inflata* (L27173 GenBank), *Castanea sativa* (AF417297 GenBank) gamma-thionin, *Nelumbo nucifera* (EF421192 GenBank) and *Capsicuum annum* (AF442388 GenBank) defensins and *Nicotiana tabacum* (AB034956 GenBank) thionin-like protein the similarity is lower than 75%. However, these significant alignments do not reflect the degree of identity of the putative mature peptides, which show a different order of identities that can reach up to 85% (see below).

Gene structure comparison

The availability of complete eukaryote genome sequences allows us to address fundamental evolutionary questions about the evolution of exon/intron gene structure. Thus, we compared the structural organization of PDEF_VIGUN with orthologs from other legumes as *G. max*, *M. truncatula*, *L. japonicus* and to the model dicot *A. thaliana*. Interesting features emerged from this analysis. First, with the exception of *A. thaliana*, the intron boundaries seem to be conserved in all compared members from Fabaceae, suggesting that also PDEF_VIGUN follows this trend showing a phase 1 intron rather than a phase 0 intron pointed by GENSCAN (Figure (**3**)). While the intron consensus sequences (A/C)AG|GT(A/G) at the donor splice site and CAG|G at the acceptor splice site are conserved, the intron length varies and the region at the intron boundaries are likely to be hotspots for indels. Curiously this frequently occurs in the most C-terminal portion of the signal peptide and could contribute to the diversity in the cleavage site thus allowing the formation of new defensins with diversified amino terminal portion.

```
Vigna           ATGGCTCGCTCTGTGCCTTTGGTCTCAACCATCTTTGTGTTTCTTCTGCTACTGGTGGCC 60
Glycine         ATGTCTCGCTCCGTGCCTTTGGTTTCAACCATTTGTGTCTTGCTTCTGCTTCTGGTGGCC 60
Prunus          ATGGAGCGCTCCATGCGTTTATTTTCAACTGCCTTCGTCTTCTTTCTGCTTCTGGCAGCT 60
                ***   *****  *** ***  * *****    *  ** **  ******* ****  **

Vigna           ACTG-GATG---GGGCCAACAATGGTGGCAGAGGCAAGGACCTGTGAGTCTCAGAGCCAC 117
Glycine         ACTGAGATGATGGGGCCAACAATGGTGGCAGAAGCAAGAACTTGTGAGTCTCAGAGCCAC 120
Prunus          GCTGGGATGATGATGGGGCCAATGGTTGCTGAGGCTAGGACCTGTGAGTCTCAGAGTAAT 120
                 *** ****     *    ******* ** ** ** ** ** **************  *

Vigna           CGTTTCAAGGGACCATGTGTGAGTGACACCAACTGTGCTTCTGTTTGCCGAACTGAACGT 177
Glycine         CGTTTCAAGGGGCCATGTTTGAGTGACACCAACTGTGGCTCTGTTTGCCGAACCGAACGT 180
Prunus          CGGTTCAAGGGAACTTGCGTGAGTACAAGCAACTGTGCATCTGTTTGCCAAACTGAGGGC 180
                ** ********  * **   *****   * ********  ********** *** **   *

Vigna           TTCAGCGGAGGACATTGCCGTGGCTTCCGTCGCAGATGCTTGTGCACCAAACACTGTTAA 237
Glycine         TTCACTGGAGGACACTGCCGTGGCTTCCGTCGCAGATGCTTCTGCACCAAACATTGTTAA 240
Prunus          TTCCCTGGTGGCCATTGTCGTGGCTTTCGCCGCAGATGCTTTTGCACTAAACATTGTTAA 240
                ***   ** ** ** ** ******** ** *********** ***** ***** ******
```

Figure 2 Vigna unguiculata defensin DNA sequence compared with that of the other plant EST obtained from databank (for a more detailed comparison with more plant species see attached materials figure X) Aligned nucleotide sequences include *Vigna unguiculata* PDEF_VIGUN, *Gycine max* protease inhibitor (**U12150 GenBank) and *Prunus persica* defensin** (**AY078426 GenBank**). *Vigna/Glycine*: score 91; *Vigna/Prunus* score 77.
* identifies conserved residues.

Amino acid sequence analysis

The deduced amino acid sequence of PDEF_VIGUN consists on 78 residues with a putative signal sequence of 31 amino acids at the N-terminus (Figure (**3**)); the putative mature peptide was identified by homology with the conserved cleavage site present in known defensin sequences and using SignalP software [19].
We then compared PDEF_VIGUN with other *V. unguiculata* defensins already described in the literature. Cp_thionin I and II [14, 15] show 68% and 37% similarity with the deduced PDEF_VIGUN, while for VuD1 (Pelegrini 2008) only 27% similarity was found. The presence of a new defensin in *V. unguiculata* confirms how multiple peptides contribute to the innate immunity of this plant, as it occurs in *Oryza sativa* and *Zea mays*.
Although in databases Fabaceae defensins are present, blasting the PDEF_VIGUN amino acid sequence against plant sequences returned homologies higher than 75% emerged only with defensins (formerly known as γ-thionins) isolated from plant species belonging to families other than Fabaceae (see Figure (**4A**)), for example *H. annuus* (sunflower, Asteraceae) γ-thionin precursor (P82659 GenBank), *P. persica* (peach, Rosaceae) (Q84UH1 UniProt), *Solanum pimpinellifolium* and *S. lycopersicum* (tomato, Solanaceae) defensin (e.g. Q9XG53 UniProt) and *Nicotiana tabacum* (tobacco, Solanaceae) thionin-like protein (Q9MB66 UniProt). It seems strange that except for the *G. max* (Q39807 UniProt) protease inhibitor with 91% similarity PDEF_VIGUN, other plants that share with *V. unguiculata* the same family and tribe, Fabaceae and Phaseoleae respectively, only show an identity lower than 75% with peptides identified as a defensin in *Phaseolus vulgaris* γ-thionin (kidney bean, A0JJX6 UniProt) and as kunits trypsin inhibitor protein in *P. coccineus* (scarlet runner bean, Q9FUP3 UniProt) and not in other bean species.
The similarity to peptides that belong to classes alternatively defined as defensins or protease inhibitors suggests that these plant peptides are likely characterized by a multifunctional activity and could act either in defence from microbes or from insects, and/or be involved in activation of other stress-responsive mechanisms.
Likewise, PDEF_VIGUN presents amino acids characteristic in plant defensins (see Figure (**3**)), namely a serine residue at position 7, an aromatic residue at position 10, two glycines at position 12 and 32, a glutamic acid at position 27 as well as the 8 cysteines residues that give the typical disulphuric bonds pattern [20]. Curiously, all the abovementioned defensins/protease inhibitor sequences show a conserved phenylalanine residue at position 42, whereas in PDEF_VIGUN this is replaced by leucine. It is notewothy that a large variation in plant defensins primary sequences is evident (apart from the cysteine residues) when comparing PDEF_VIGUN with other plant defensin sequences (Fig. **3**). Despite this high variability they all share the three dimensional structures as it is described below.

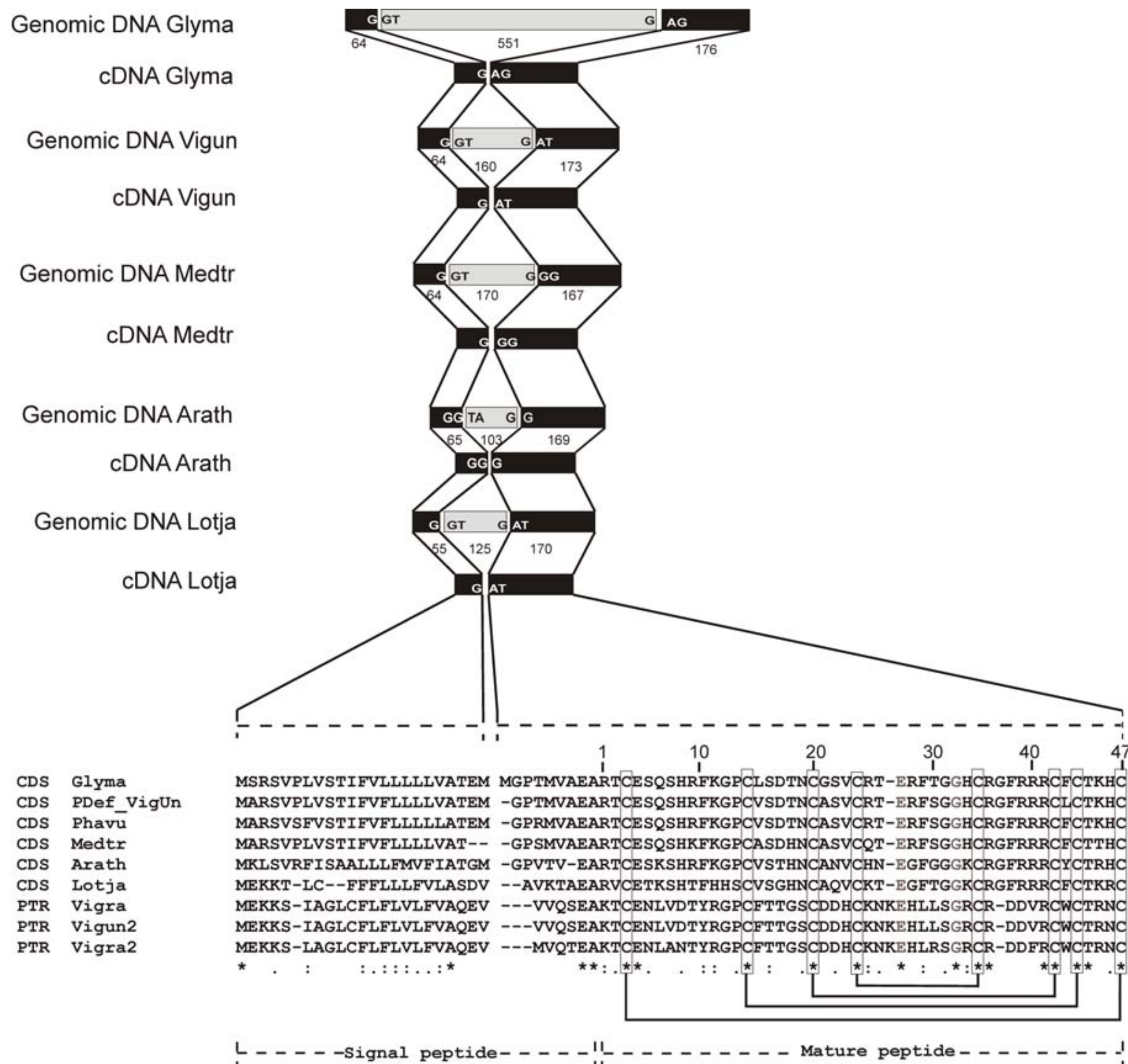


Figure 3. Plant defensin gene structure after heterologous comparison. The intron boundaries are conserved in all compared members from Fabaceae, with the exception of *A. thaliana*. Black boxes indicate exon while grey boxes the introns. The nucleotide length of the DNA stretch is reported below the boxes. CDS: translated coding sequence; Arath: *Arabidopsis thaliana*; Glyma: *Glycine max*; Medtr: *Medicago truncatula*; Lotja: *Lotus japonicus*; Vigun: *Vigna unguiculata*. Phavu: *Phaseolus vulgaris* (TC147). PTR: Protein added to the alignment; Vigra: *Vigna radiata* (CAA34760), Vigun2: *Vigna unguiculata* (P18646); Vigra2: *Vigna radiata* (BAB82453). Disulphuric bonds are highlighted: Cys3-Cys47; Cys14-Cys43; Cys20-Cys41; Cys24-Cys34).

Protein structure prediction

Blasting the PDEF_VIGUN sequence against the PDB database reveals the sequence of *Hordeum vulgare* defensins (gamma-hordothionin) as having the closest identity (53%, see Figure (**4A**)). The sequence identity, as well as other primary structural features, suggests that its structure (1GTP) is representative of the PDEF_VIGUN tertiary structure. The peptide's topology, as well as a schematic representation of the structure are shown in Figure (**4B**) and (**4C**). An analysis of this structure and the predicted PDEF_VIGUN structure (derived by a direct residue substitution), is revealing. Conservation of residues Ser_7 and Glu_{27} are explained by the fact that although polar, they are both buried and form a network of H-bonds that likely has an essential role in stabilising the defensin scaffold. The conserved aromatic residues at positions 10 and 29 and the conserved valine at position 23 contribute to the core of the molecule. Gly_{12} seems at a key position in a loop connecting the first strand with the helix (see Figure (**4C**)) and its main chain amide and carboxyl form a network of H-bonds with other conserved residues, Phe_{10} and Arg_{40}. Gly_{32} is in a highly hindered position where strand 2 passes directly over the helix. Apart from this, other conserved residues in PDEF-VIGUN may have important structural roles. Asn_{19}, for example, is involved in a network of H-bonds that may indicate it acts as a helix-stabilising cap, also involving Ser_{16}.

Leu$_{42}$, that replaces Phe$_{42}$ present in all the other closely related sequences, would not seem to be placed in the molecules core, but rather project from the -sheet platform (on the rear of Figure (**4C**)).

PDEF_VIGUN is quite cationic, with a winged arrangement of the basic residues that are clustered at the base and at the top of the structure, as indicated in Figure (**4C**), mostly in the loop regions connecting the -strands and -helix.

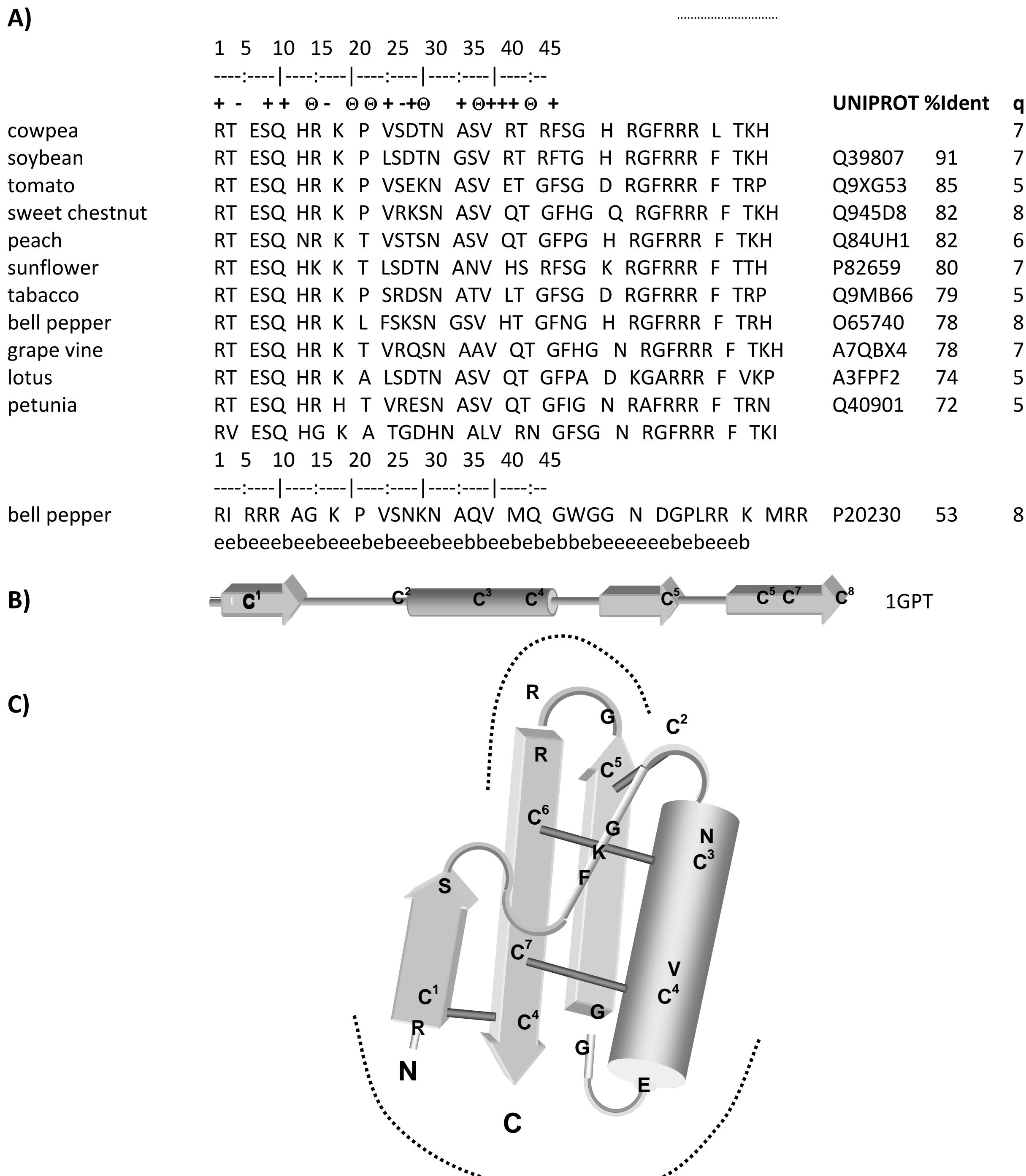
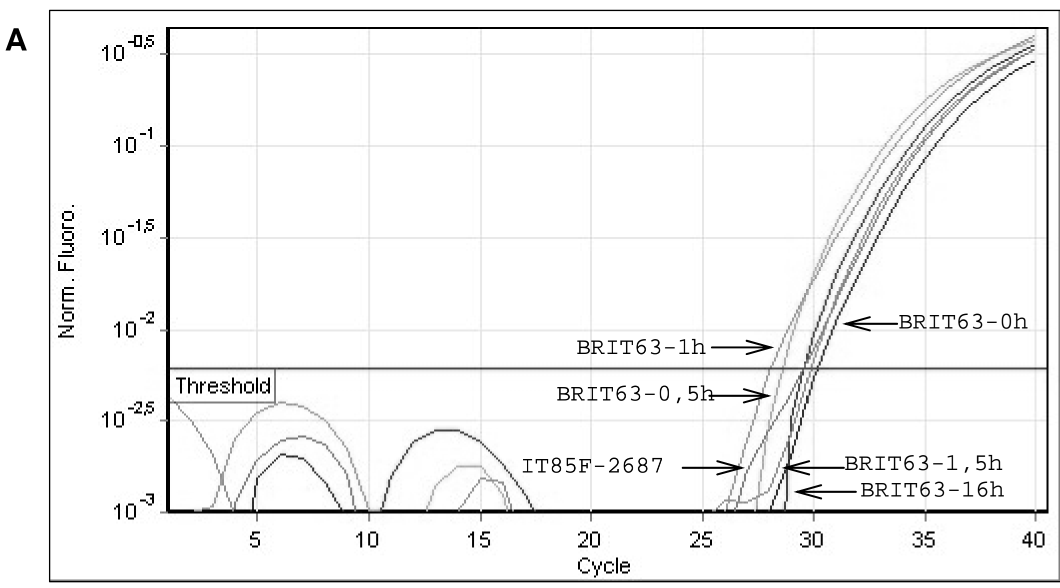


Figure 4. Amino acid sequence and putative structure analysis A) Amino acid sequence alignment of PDEF_VIGUN with closely related sequences from other plants. Plants are identified by common names, but the UniProt accession codes are provided. The degree of identity with PDEF_VIGUN and net charge are also shown. Highly conserved residues are shaded grey, strictly conserved residues are shaded black. The presence of conserved positively charged positions, negatively charged positions and hydrophobic residues positions are indicated by +, -, and Θ respectively. The sequence for bell pepper (*C. annuum*) defensin, the closest peptide for which a structure (1GPT) has been determined, is shown separately below. Buried (b) and exposed (e) residues resulted from visual analysis of each residue in the barley structure, as did the topological diagram (**B**). **C)** Schematic cartoon of the putative PDEF_VIGUN structure, based on that of bell pepper defensin (1GPT). The putative position of highly conserved residues in this schematic representation is approximately indicated. The dotted lines indicate areas where cationic residues would be concentrated.

RT-QPCR

PDEF_VIGUN expression and its response to biotic stimuli were tested by RT_QPCR. RNA samples were extracted from 30-days old plants BRIT63 (a unique CPSMV resistant genotype) inoculated with CPSMV, while BRIT63 and the CPSMV-susceptible parental cultivar (IT85F non infected nor mechanically injured) were used as negative controls. RT-qPCR data obtained after CPSMV inoculation suggest that defensin gene expression is responsive to biotic stimuli and wounding stress (Figure (**5**)). Although the average Ct value did not present marked changes, the exponential nature of the amplification may indicate a significant variation in transcription intensity. In time-course samples, defensin expression was observed to occur rapidly and to be strongly induced by CPSMV infection, already during the first 30 min after inoculation. An increasing responsiveness is observed for a further 30 min before expression starts to decrease from around 1.5 hours after inoculation, to a level which is maintained for the next several hours.

A



B

Sample	Harvest time after virus inoculation (h)	Average Ct	Expression [(2^-Ct)*(10^9)]
IT85F-2687	0*	29.63	1.204
BRIT63	0*	30.17	0.828
	0.5	28.61	2.441
	1.0	28.07	3.549
	1.5	29.90	0.998
	16.0	29.56	1.263

Figure 5. A) Plant defensin expression in *Vigna unguiculata* after wounding and CPSMV inoculation. BRIT63 is a unique CPSMV (cowpea severe mosaic virus) resistant genotype derived from a breeding cross (BR14-Mulato x IT85F-2687), while IT85F is the CPSMV-susceptible parental cultivar sample. Negative controls consisted in one BRIT63 and one IT85F sample both not harvested with CPSMV. Threshold Cycle (Ct) values inversely correlate with the quantity of mRNA expression, β-defensin expression, after inoculation, appears to be fast and strongly during the first 30 min after inoculation (BRIT63-0,5h) and response increase for the following 30 min (BRIT63-1h). **B)** In the table, results for *Vigna* RT-qPCR for defensin transcript detection, after CPSMV inoculation are reported.

A similar pattern was observed regarding the transcription profile of a pepper defensin (*CADEF1*) with increased expression already 30 min after wounding stress with needles while the same gene was strongly expressed 2 h after *X. campestris* pv. *vesicatoria* infection in an incompatible interaction [21]. Such a picture points to a regulation model for this *V. unguiculata* defensin gene expression that might be considered as early responsive to wounding and biotic (viral) stress. In addition, PDEF_VIGUN may perform direct roles as virus antagonist molecules, since some plant defensins (γ-thionins) have presented a very large spectrum of action against diverse plant parasitic bacterial genera through

impairing cell growth [15]. Noteworthy, legume defensins have already been described as harbouring effective antiproliferative and anti-HIV-1 reverse transcriptase activities in mammal pathosystems [22]. Indeed, PDEF_VIGUN may be part of a plant defense mechanisms acting as one of the initial responses, necessary to activate, or allow time for activation of secondary and/or late defense-related metabolic and physiological responses more directly associated to resistant phenotype.

In conclusion we have isolated from *V. unguiculata* a sequence with characteristics that suggest it corresponds to a defensin-like gene. Amino acid sequences analyses indicate that the deduced PDEF_VIGUN sequence is homologous to already known defensin proteins and thus might have similar function to these other defensins. Moreover, preliminary functional results demonstrate the presence of mRNA for this cowpea defensin in the plant tissue and also an increase in expression in response to biotic stimuli such as CPSMV infection. This suggests that a lack of defensins expression could make the plant more susceptible to the invasion by pathogens.

ACKNOWLEDGEMENTS

This work was partially supported by RC03/04 from IRCCS Burlo Garofolo. L.P. is recipient of a fellowship from FVG2008. None of the authors has any potential financial conflict of interest related to this manuscript.

References

[1] Broekaert, W.F.; Terras, F.R.; Cammue, B.P.; Osborn, R.W. Plant defensins: novel antimicrobial peptides as components of the host defense system. *Plant Physiol*., **1995**, *108*(4),1353-1358.

[2] Thomma, B.P.; Cammue, B.P.; Thevissen, K. Plant defensins. *Planta*, **2002**, *216*(2), 193-202.

[3] Castro, M.S.; Fontes, W. Plant defense and antimicrobial peptides. *Protein Pept Lett*., **2005**, *12*(1), 13-18.

[4] Bloch, C. Jr; Richardson, M. A new family of small (5 kDa) protein inhibitors of insect alpha-amylases from seeds or sorghum (*Sorghum bicolor* (L) Moench) have sequence homologies with wheat γ-purothionins. *FEBS Lett*., **1991**, *279*(1), 101-104.

[5] Almeida, M.S.; Cabral, K.M.; Zingali, R.B.; Kurtenbach, E. Characterization of two novel defense peptides from pea (*Pisum sativum*) seeds. *Arch Biochem Biophys*, **2000**, *378*(2), 278-86.

[6] Antcheva, N.; Zelezetsky, I.; Tossi, A. Cationic Antimicrobial *Peptides*—The Defensins. In *The Handbook of Biologically Active Peptides*; A.J. Kastin, Ed.; Elsevier Science B. V: Amsterdam, **2006**; pp. 55-66.

[7] Liu, Y.J.; Cheng, C.S.; Lai, S.M.; Hsu, M.P.; Chen, C.S.; Lyu, P.C. Solution structure of the plant defensin VrD1 from mung bean and its possible role in insecticidal activity against bruchids. *Proteins*, **2006**, *63*(4), 777-786.

[8] Mendez, E.; Moreno, A.; Colilla, F.; Pelaez, F.; Limas, G.G.; Mendez, R. Soriano F, Salinas M, de Haro C. Primary structure and inhibition of protein synthesis in eukaryotic cell-free system of a novel thionin, γ -hordothionin, from barley endosperm. *Eur J Biochem*, **1990**, *94*, 533-539.

[9] Games, P.D.; Dos Santos, I.S.; Mello, E.O.; Diz, M.S.; Carvalho, A.O.; de Souza-Filho, G.A.; Da Cunha, M.; Vasconcelos, I.M.; Ferreira, B. dos S.; Gomes, V.M. Isolation, characterization and cloning of a cDNA encoding a new antifungal defensin from *Phaseolus vulgaris* L. seeds. *Peptides*, **2008**, *29*(12), 2090-2100.

[10] Zhang, Y.; Lewis, K. Fabatins: new antimicrobial plant peptides. *FEMS Microbiol Lett*, **1997**, *149*, 59-64.

[11] Dimarcq, J.L.; Zachary, D.; Hoffmann, J.A.; Hoffmann, D.; Reichhart, J.M. Insect immunity: expression of the two major inducible antibacterial peptides, defensin and diptericin, in *Phormia terranovae*. *EMBO J*, **1990**, *9*, 2507-2515.

[12] Ehlers, J.D.; Hall, A.E. Cowpea (*Vigna unguiculata* L. Walp.). *Field Crops Research*, **1997**, *53*(1–3),187-204.

[13] Azevedo, H.; Houllou-Kido, L.; Benko-Iseppon, A.M. Análise do Potencial Regenerativo in vitro de Diferentes Cultivares de Feijão-Caupi. *Revista Brasileira de Biociências*, **2007** , *5*(2), 528-530.

[14] Melo, F.R.; Rigden, D.J.; Franco, O.L.; Mello, L.V.; Ary, M.B.; Grossi de Sá, M.F.; Bloch Jr, C. Inhibition of trypsin by cowpea thionin: characterization, molecular modeling, and docking**.** *Proteins*, **2002**, *48*(2), 311-319.

[15] Franco, O.L.; Murad, A.M.; Leite, J.R.; Mendes, P.A.; Prates, M.V.; Bloch Jr, C. Identification of a cowpea gamma-thionin with bactericidal activity. *FEBS J*., **2006**, *273*(15), 3489-3497.

[16] Weising, K. *DNA fingerprinting in plants: principles, methods, and applications*, 2nd ed.; CRC Press: Boca Raton, FL, **2005**.

[17] Pio-Ribeiro, G.; Paz, C.D.; Andrade, G.P.; Assis Filho, F.M. Efeito de quatro isolados do vírus do mosaico severo, CPSMV, na produção de quatro cultivares de caupi, em condições de casa de vegetação. *Caderno Ômega*, **2001**, *12*, 44-45.

[18] Chang, S.; Puryear, J.; Cairney, J. A simple and efficient method for isolating RNA from pine trees. *PMB Reporter*, **1993**, *11*, 113-116.

[19] Bendtsen, J.D.; Nielsen, H.; von Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol*., **2004**, *340*(4), 783-795.

[20] Janssen, B.J.; Schirra, H.J.; Lay, F.T.; Anderson, M.A.; Craik, D.J. Structure of *Petunia hybrida* defensin 1, a novel plant defensin with five disulfide bonds. *Biochemistry*, **2003**, *42*(27), 8214-8222.

[21] Do, H.M; Lee, S.C.; Jung, H.W.; Sohn, K.H.; Hwang, B.K. Differential expression and in situ localization of a pepper defensin (CADEF1) gene in response to pathogen infection, abiotic elicitors and environmental stresses in *Capsicum annuum*. Plant Science, **2004**, *166(5)*, 1297-1305.

[22] Ngai, P.H.; Ng, T.B. Phaseococcin, an antifungal protein with antiproliferative and anti-HIV-1 reverse transcriptase activities from small scarlet runner beans. *Biochem Cell Biol* **2005** *83*(2), 212-220.

Manuscrito de artigo científico 2

Belarmino e Benko-Iseppon, 2010

Databank based mining on the track of antimicrobial weapons in plant genomes

Current Protein and Peptide Science, aceito em Outubro de 2009

Databank based mining on the track of antimicrobial weapons in plant genomes

L. C. Belarmino[1] and A. M. Benko-Iseppon[1]*

[1]Federal University of Pernambuco, Genetics Dept., Recife, Brazil.

*Corresponding Author

Ana Maria Benko Iseppon

Universidade Federal de Pernambuco, Genetics Dept. (UFPE/CCB/Genetica)

Av. Prof. Moraes Rego, 1235, CEP 50.670-420, Recife, PE, Brazil.

Phone: 55-81-2126-8620 and 2126-7816

Fax: 55-81-2126-8522

E-mail: ana.benko.iseppon@pq.cnpq.br

Abstract

The expressive amount of nucleotide sequences from diverse plant species in databanks enables the use of computational approaches to discovery still unidentified genes and to infer about their function, structure and role in some biological processes. Of special interest are the antimicrobial peptides (AMP), whose functionalities have a very important role in defense against microbial infection in multicellular eukaryotes, being considered less susceptible to bacterial resistance than traditional antibiotics, with potential to develop a new class of therapeutic agents. Recent computational developments have provided various algorithms and resources to profit from the overwhelming information in data banks for biomining such peptides. This review focuses on the computational and bioinformatic approaches so far used for the identification of antimicrobial peptides in plant systems, highlighting alternative means of mining the entire plant peptide space that has recently become available.

Keywords: antimicrobial peptides, bioinformatics, computational modeling, databases.

Introduction

Despite of the intense exposure of plants to microorganisms present in the environment, the occurrence of infectious processes is quite uncommon, especially in native plants, due to the development of defense mechanisms during plant-microorganism co-evolution. The production of antimicrobial peptides (AMP) is a common strategy applied by many organism classes as a first defense response. AMP constitutes a main component of animal innate immunity, being an ancestral defense mechanism in all living groups including prokaryotes [1], arthropods [2], fungi [3], animals [4] and plants [5]. In such groups AMP may be expressed constitutively or in response to wounding and microbial invasion, composing pre- or post-established defense barriers [1-5].

Recent bioinformatic analysis in plants revealed an unexpected abundance of coding genes for such peptides, representing a dynamic and diversified defense arsenal against pathogenic invaders [6-10]. The present review summarizes the computational methods adopted for the identification of antimicrobial peptides discussing their diversity considering available information in data banks and also their potential for future biotechnological uses.

Biomining of Antimicrobial peptides

Until recently, the majority of antimicrobial peptides have been determined by proteomic inferences, after purification from plant extracts with posterior cloning and evaluation of their respective genes [11]. Meanwhile, the growing availability of plant complete sequenced genomes, as well as expressed sequence tags (ESTs), is changing this scenario, especially considering the computational resources for gene comparison and recognition of specific domains and signatures. The amount of plant ESTs available allows the use of computational methods for the identification of new genes in specific taxa or in association with given tissues or processes in plants.

Due to its importance, the plant-microbe interaction is among the most studied processes, with legumes figuring as model systems for genomic evaluations in order to understand the symbiotic associations within this group, but also for understanding the mechanisms of innate defense in

higher plants. Considering these goals, Fedorova *et al.* [7], using the Boolean formalism, identified 340 genes putatively specific for legume root nodules. Surprisingly 114 of the identified genes coded for small Cys Cluster Proteins (CCP) bearing cleavable signal peptides. Additionally Mergaert *et al*. [8] identified a considerable number of CCPs (more than 300 Nodule specific Cys-rich Proteins - NCR) by searching the MtGI (*Medicago truncatula* Gene Index) database using BLAST and 19 seed sequences from related plant families.

Similarly, Graham *et al*. [9] used BLAST tools to compare unigene sets from legumes to non-legumes unigene sets and to genomic sequences from rice (*Oryza sativa*), *Arabidopsis thaliana* and the non-redundant EST database from GenBank. Afterwards the authors grouped the putative legume specific genes in their respective families and evaluated the putative protein motifs for each family, when comparing with a protein database, allowing predictions regarding their function. This procedure allowed for the identification of proline- and cysteine-rich proteins, known respectively as Pro-rich Proteins (PRP) and Cys-rich Proteins (CRP). From these, 300 CRP were exclusively expressed in root nodules and seeds, with similarity to known defensins, some of them regarding unknown open reading frames in the *A. thaliana* genome. Despite having other goals, the described approach revealed the high number and diversity that apparently is characteristic of leguminous AMPs, especially regarding root nodules. This feature suggests the development of secondary defense systems for this group by recruiting genes from the primary defense system to avoid the invasion by other microorganisms present in the soil while the symbiotic specific association with *Rhizobium* bacteria is promoted [7-9].

After the identification of the nodule-specific defensins by Graham *et al*. [9], Silverstein *et al*. [10,11] observed that plant genomes present an unexpected abundance of putative cysteine-rich AMPs with an astonishing potential for functional adaptation. Using an approach based on Hidden Markov Models (HMM) and BLAST searches, Silverstein *et al.* [11] identified CRP classes bearing a signal peptide as known for defensins, thionins, lipid transfer proteins and Rapid Alkanization Factors (RALF), as well as hypothetical coding CRP genes with cysteine signatures not previously

identified in purified proteins. The strategy used by these authors allowed the identification of 12,824 distinct CRP sequences in 33 plant species, supporting previous evidences regarding their massive diversity in the plant kingdom.

These initial studies were valuable for the demonstration of the abundance and diversity of antimicrobial peptides in plants. However, the strategies adopted by the authors do not favor the identification of new peptides in which the genealogical relationships are not evident by the given sequence signatures.

Web resources and repositories

Given the importance of antimicrobial peptides, beyond knowing their amino acid sequences, it is very important to understand their structure, topology and function that together allow a better understanding of their effective action against pathogens, including some bacteria resistant to conventional antibiotics, important for the development of new antimicrobial drugs.

Actually, data regarding AMPs may be accessed in different web repositories (Table 1). These databases are the result of many efforts to collect, process and store such sequences together with some tools for their comparative evaluation. Each of the referred databases was developed with a specific purpose and regarding a given data set, grouping AMP molecules from a diversity of organisms, including pro- and eukaryotes.

For example, the AMSDb database is the oldest database online and contains several plant AMPs. This database allows search by keyword as well by some features like organism of source, expression profile and activity, among others [12]. The ANTIMIC database includes 1,700 known putative antimicrobial peptides, also integrating tools to facilitate an efficient data extraction and analysis at molecular level as well as to search new AMPs [13]. Similarly, Seebah *et al*. [14] provided a curated sequence database and information source regarding the defensin AMPs.

Another resource, the CyBASE, was developed by Wang *et al*. [15] as a base for AMP bearing circularized scaffolds, including search tools and interfaces for exposition of sequence structure and

function. Additionally, Fjell *et al*. [16] developed the AMPer database using HMM applied to public available AMP sequences. This data bank allows the recognition of individual AMP classes, as defensins, catelecidins and cecropins with a precision of ≥ 99 %, constituting an excellent discovery tool for the identification of specific AMP members.

Wang *et al*. [17] developed a database dedicated to AMP from all life forms, from bacteria to plants and animals, including humans. Recently this database was updated being designated as APD2 (Antimicrobial Peptides Database) [18], a resource that provides statistical data for peptides present in data banks, as well as tools for evaluation of the structure/function relationship regarding AMP.

Finally the PhytAMP database is the only one exclusively dedicated to plants [19], presenting valuable information regarding plant antimicrobial peptides and including taxonomic and microbiological information, besides the availability of tools for inferences regarding sequence structure and function. This database includes 271 sequences, a number that is vastly underestimated and that reflects the lack of extensive evaluations in plants, not a deficiency in the amount of available plant AMPs.

Table 1: Main web based repositories including plant antimicrobial peptides.

Database	Description	Web site	Reference
AMSDb	Animal and plant oriented antimicrobial database containing sequence and annotation information concerning accepted structure-based classification.	http://www.bbcm.units.it/~tossi/amsdb.html	Antcheva *et al.* [12]
AMPer	A database of Hidden Markov Model designed for various AMP classes.	http://marray.cmdr.ubc.ca/cgi-bin/amp.pl	Fjell *et al.* [16]
ANTIMIC	A database of antimicrobial sequences with facility tools for data extraction.	http://research.i2r. a-star.edu.sg/Templar/DB/ANTIMIC/ (Discotinued)	Brahmachary *et al.* [13]
APD	A database of antimicrobial peptide of diverse origin.	http://aps.unmc.edu/AP/main.html	Wang and Wang [17]; Wang *et al.* [18]
CyBase	A database of cyclic protein sequences and structures, with applications in protein discovery and engineering.	http://research1t.imb.uq.edu.au/cybase	Wang *et al.* [15]
Defensins knowledgebase	A manually curated database and information source on the defensin family of antimicrobial peptides.	http://defensins.bii.a-star.edu.sg/	Seebah *et al.* [14]
PhytAMP	A database dedicated to antimicrobial peptides in plants.	http://phytamp.pfba-lab-tun.org	Hammami *et al.* [19]

Computational modeling of antimicrobial peptides

The growing interest on AMP molecules is justified by their important role in the defense mechanisms of all living organisms. To date, about 2,000 AMP sequences have been identified in the main specific databases (ANTIMIC, APD2 and PhytAMP), from which 271 regard plant AMPs [13,18,19]. Among peptide analysis available in data banks, Nagarajan *et al.* [20] identified an indexing method based on Fourier transformation, using biologically distinctive features of known antimicrobial peptides. The developed algorithm was used to randomly scan 10,000 proteins of small size (16 amino-acids), obtaining three positive results with high probability of antimicrobial activity as predicted by the APD database.

Observing the modular pattern frequently seen in natural AMPs Loose *et al*. [21] proposed the existence of an AMP specific pattern or "language" that putatively could shelter given "grammatical rules" traceable in their nucleotides or amino-acid sequences. In biochemical terms, the sentences would be analog to peptides while the words in a sentence would be analog to the amino acids of a given peptide. The authors found the searched group of grammatical rules (or patterns) to describe AMPs using the Teiresias pattern discovery tool, identifying 684 grammatical patterns in 526 well characterized eukaryotic AMP sequences available in the databases. This approach allowed the development of 700 grammatical models with 20 amino acids in size (considered to be the medium size of AMP in data banks), named unnatural AMPs, comparing them with the data bank of natural AMP sequences. From this set, all 20-mers that had six or more amino-acids in a row in common with a naturally occurring AMPs were selected and compared to the remaining sequences on the basis of similarity, allowing the selection of 42 sequences to synthesize and assay for antimicrobial activity. The tested peptides were able to inhibit bacterial growth with a minimum concentration of 256 $\mu g.mL^{-1}$, confirming the successful application of this approach for the recognition and development of small artificial AMPs, without tertiary and quaternary complex structures.

In a later work based on the APD database, Lata *et al*. [22] analyzed 476 antibacterial peptides, observing that in AMPs some amino-acids were preferred to the detriment of others. The group used Artificial Neural Network (ANN), Quantitative Matrix (QM) and Support Vector Machine (SVM) algorithms to predict AMPs with high accuracy, allowing the development of models based only on 15 residues of the N- and C-terminal sites, considering that bactericidal AMPs normally bear only 20-30 amino-acid residues. By using the described approach, they obtained impressive performance for the three modules based on QM, ANN and SVM, with the SVM approach presenting the best results. However, the results are not representative for plants, since the majority of the available antibacterial AMP sequences are from non plant sources.

Conclusions

The recognition and characterization of plant AMPs is still in its infancy. Besides plant specific approaches, the evaluation of public databanks have the potential to reveal an amazing diversity and abundance of such sequences in plants. Considering the up to date status, one may recommend that data mining groups should focus on procedures to aid the identification of coding genes for small proteins.

Bioinformatic strategies have been very effective in the recognition of well known AMPs with given features derived from their common ancestry. On the other hand, unknown, not previously described peptides remain undetected by the available search methods. An alternative to solve this problem would be the unraveling of given features of biologically distinct plant antimicrobial peptides, developing algorithms able to recognize these features. Meanwhile, an effort to collect, process, store and validate AMP specific information, carried out by our group* [23] shall help the construction of a large panel of features able to allow the recognition of still non described AMP specific patterns.

Acknowledgements

The authors thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) and FACEPE (Fundação de Amparo à Pesquisa do Estado de Pernambuco, Brazil) for the concession of fellowships. None of the authors has any potential financial conflict of interest related to this manuscript.

Footnote: *Padovan, L.; Segat, L.; Tossi A.; Calsa Jr, T.; Ederson, A.K.; Brandao, L.; Guimarães, R.L.; Pandolfi, V.; Pestana-Calsa, M.C.; Belarmino, L.C.; Benko-Iseppon, A.M.; Crovella, S. Characterization of a new defensin from cowpea (*Vigna unguiculata* (L.) Walp.) *Protein Pept. Lett. accepted*

References

[1] Asaduzzaman, S.M.; Sonomoto, K. Antibiotics: Diverse activities and unique modes of action. *J. Biosci. Bioeng*., **2009**, *17*, 475-487.

[2] Bulet, P.; Stöcklin, R. Insect antimicrobial peptide: Structures, properties end gene regulation. *Protein Pept. Lett*., **2005**, *12*, 3-11.

[3] Marx, F. Small basic antifungal proteins secreted from filamentous ascomycetes: a comparative study regard expression, structure, function and potential application. *Appl. Microbiol. Biotechn*., **2004**, *65*, 133-142.

[4] Bulet, P.; Stoecklin, R.; Menin, L. Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev*., **2004**, *198*, 169-184.

[5] Manners, J.M. Hidden weapons of microbial destruction in plant genomes. *Genome Biol*., **2007**, *8*, 225.

[6] Silverstein, K.A.; Moskal, W.A. Jr; Wu, H.C.; Underwood, B.A.; Graham, M.A.; Town, C.D.; VandenBosch, KA. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. *Plant J*., **2007**, *51*, 262-280.

[7] Fedorova, M.; Van de Mortel, J.; Matsuomoto, P.A.; Cho, J.; Town, C.D.; Vandenbosch, K.A.; Gantt, J.S.; Vance, C.P. Genome-Wide Identification of Nodule-Specific Transcripts in the Model Legume *Medicago truncatula*. *Plant Physiol*., **2002**, *130*, 519-537.

[8] Mergaert, P.; Nikovics, K.; Kelemen, Z.; Maunoury, N.; Vaubert. D.; Kondorosi, A.; Kondorosi, E. A Novel Family in *Medicago truncatula* Consisting of More Than 300 Nodule-Specific Genes Coding for Small, Secreted Polypeptides with Conserved Cysteine Motifs. *Plant Physiol*., **2003**, *132*, 161–173.

[9] Graham, MA; Silverstein, KA; Cannon, SB; VandenBosch, KA. Computational Identification and Characterization of Novel Genes from Legumes. *Plant Physiol*., **2004**, *135*, 1179-1197.

[10] Silverstein, K.A.; Graham, M.A.; Paape, T.D.; VandenBosch, K.A. Genome organization of more than 300 defensin-like genes in *Arabidopsis*. *Plant Physiol*., **2005**, *138*, 600-610.

[11] Odintsova, T.I.; Egorov, T.A. In *Peptidomics: method and application*; Soloviev, M.; Shaw, C.; Andrén, P., Eds.; John Wiley and Sons, Inc: New Jersey, **2007,** pp. 99-117.

[12] Antcheva, N.; Zelezetsky, I. and Tossi, A. *Cationic Antimicrobial Peptides — The Defensins* in The Handbook of Biologically Active Peptides, A.J. Kastin Ed.; Elsevier, Burlington, **2006**, MA. Chapter 11, pp. 55-66.

[13] Brahmachary, M; Krishnan, SPT; Koh, JLY; Khan, AM; Seah, SH; Tan, TW; Brusic, V; Bajic, VB. ANTIMIC: a database of antimicrobial sequences. *Nucl. Acids Res*., **2004**, *32*, 586-589.

[14] Seebah S.; Suresh, A.; Zhuo, S.; Choong, Y.H.; Chua, H.; Chuon, D.; Beuerman, R.; Verma, C. Defensins knowledgebase: a manually curated database and information source focused on the defensins family of antimicrobial peptides. *Nucl. Acids Res*., **2007**, *35*, D265-8.

[15] Wang, C.K.; Kaas, Q.; Chiche, L.; Craik D.J. CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucl. Acids Res*., **2008**, *36*, D206-10.

[16] Fjell, C.D.; Hancock, R.E.W.; Cherkasov, A. AMPer: a database and an automated discovery tool for antimicrobial peptides. *Bioinformatics*, **2007**, *23*, 1148-1155.

[17] Wang, Z.; Wang, G. A.P.D: the Antimicrobial Peptide Database. *Nucl. Acids Res*., **2004**, *32*, D590-D592.

[18] Wang, G.; Li, X.; Wang, Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucl. Acids Res*., **2008**, *37*, D933-D937.

[19] Hammami, R.; Hamida, J.B.; Vergoten, G.; Fliss, I. PhytAMP: a database dedicated to antimicrobial plant peptides. *Nucl. Acids Res*., **2009**, *37*, D963-D968.

[20] Nagarajan, V.; kaushik, N.; Murali, B.; Zhang, C.; Lakhera, S.; Elasri, M.O.; Deng, Y. A Fourier transformation based method to mine peptide space for antimicrobial activity. *BMC Bioinformatics*, **2006**, *7*, S2.

[21] Loose, C.; Jensen, K.; Rigoutsos, I.; Stephanopoulos G. A linguistic model for the rational design of antimicrobial peptides. *Nature*, **2006**, *443*, 867-869.

[22] Lata, S.; Sharma, B.K.; Raghava, G.P.S. Analysis and prediction of antibacterial peptides. *BMC Bioinformatics*, **2007**, *8*, 263

[23] Padovan, L.; Segat, L.; Tossi, A.; Antcheva, N.; Benko-Iseppon, A.M.; Ederson, A.K.; Brandao, L.; Calsa, T.Jr.; Crovella, S. A plant-defensin from sugarcane (*Saccharum* spp.). *Protein Pept. Lett*., **2009**, *16*, 430-436.

Manuscrito de artigo científico 3

Belarmino *et al.*, 2010

EST-Database search of plant defensins – an example using sugarcane, a large and complex genome

Current Protein and Pepite Science, aceito em Novembro de 2009

EST-Database search of plant defensins – an example using sugarcane, a large and complex genome

L.C. Belarmino[1], P.V.S.Z. Capriles[2], S. Crovella[1], L.E. Dardene[2], A.M. Benko-Iseppon[1]*

[1]Federal University of Pernambuco, Genetics Dept., Recife, Brazil.

[2]Grupo de Modelagem Molecular de Sistemas Biológicos, Laboratório Nacional de Computação Científica, Petrópolis, Rio de Janeiro, Brazil.

*Corresponding Author

Ana Maria Benko Iseppon

Universidade Federal de Pernambuco, Genetics Dept. (UFPE/CCB/Genética)

Av. Prof. Moraes Rego, 1235, CEP 50.670-420, Recife, PE, Brazil.

Phone: 55-81-2126-8620 and 2126-7816

Fax: 55-81-2126-8522

E-mail: ana.benko.iseppon@pq.cnpq.br

Abstract

EST (Expressed Sequence Tags) databases are increasing in number and size, especially regarding cultivated plants. Sugarcane is one of the most important tropical and subtropical crops, presenting a complex polyploid genome of hybrid origin, bearing a challenge for the understanding of genetic processes in higher plants. In the present work a general search was carried out on the largest Sugarcane EST Database (SUCEST) that includes 237,954 ESTs aiming to identify defensin antimicrobial peptides – a class of small, basic, cysteine-rich peptides distributed throughout the kingdoms. Using a computational approach 17 new defensin isoforms could be identified. Main steps for the search, characterization and evaluation of the defensin expression profile are presented. Prevalent expression tissues were leaf roll, lateral bark, root apex, base of inflorescence, developing seed, and calli. Bioinformatics and phylogenetic analysis of the primary structure of the sugarcane defensin candidates as well as the 3D structures obtained by comparative modeling support their role as antimicrobials.

Keywords: data mining, EST-database, comparative modeling, defense system, crop evolution, drug discovery.

Plant EST databases

Expressed sequences as ESTs and SuperSage tags, represent stretches of the expressed genome portions, allowing the potential identification of genes encoding proteins, natural antisense transcripts [1-6], miRNA, transacting siRNA precursors [7, 8, 9], and noncoding RNA [10]. EST projects to acquire information about the transcriptome have been carried out for hundreds of organisms including plants and animals. Currently, more than 63 million ESTs are available through the dbEST entry of GenBank (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_sumary.html). However, only 32% of dbEST release corresponds to plant sequences as of November 03, 2009. The large volume of short lengths and the lack of functional annotation are common obstacles to the use of EST sequence for gene modeling and structure identification. Despite of their fragmentary and inaccurate nature, ESTs were proved to be an effective technique for the discovery of new genes, particularly those involved in given biological processes like plant disease [11].

Several groups succeeded to assemble and annotate ESTs and cDNAs from plants, in order to reduce the redundancy and to potentiate sequence information regarding the transcripts. Each of these groups used a specific approach to select the plant species and the types of sequences to be clustered and included in the database. The first effort to offer the scientific community with such a database was the Unigene project from NCBI (*http://www.ncbi.nlm.nih.gov/sites/entrez?db =unigene*) that nowadays contains unique transcription loci for 36 plant species [12]. More recently, two databases exclusively dedicated to plant EST assembly have been launched: the PlantGDB and the plant Transcript Assembly (TA). The PlantGDB provides annotated transcript assembly for 202 plant species including more than 10,000 published ESTs available in GenBank and Uniprot (*http://www.plantgdb.org/*). Every four months, the PUT (PlantGDB-generated Unique Transcript) data are refreshed, aligned to Uniprot entries and the best match along with the associated Gene Ontology (GO) term are stored [13]. Finally, the Institute for genomic research generates plant Transcript Assembly (TA) for all species that have more than 1,000 publicly available ESTs and cDNA sequences in GenBank (*http://plantta.tigr.org*). The PlantTA database is the most

comprehensive assembly database known to date, including 252 plant species, being updated every three months [14].

Sugarcane's genome

A useful aspect of ESTs is in accessing genetic information of species with complex genomes, whose access is difficult using conventional genetics. This is the case of sugarcane, an important crop that is cultivated in the tropics and subtropics due to its unique stalk with high sucrose accumulation. Modern sugarcane cultivars are hybrids derived from the crossing of *Saccharum officinarum*, usually having 2n = 80 chromosomes and *Saccharum spontaneum*, 2n = 40 - 128 chromosomes. Both species are polyploid with genomes that are multiple of 2n = 10 and 2n = 8 chromosomes, respectively [15].

The chromosome number can vary among commercial cultivars. Due to structural differences between chromosomes of the two species, the hybrids possess different proportions of chromosomes, varying chromosome sets and complex recombination events. The hybrids are highly polyploid, aneuploid and on average contain 100–120 chromosomes with an estimated somatic cell size of 10,000 Mbp [16]. The basic genome size ranges from 760 to 926 Mbp, which is twice the size of the rice genome (389 Mbp) and similar to sorghum's genome (760 Mbp) [17]. Because of this complexity, the sugarcane genome has received in the past relatively little attention and investments from plant scientists in the development of biotechnology and genetic tools for this crop. Nonetheless, molecular cytogenetic approaches and the use of molecular markers have led to significant advances in establishing the origin and genomic structure of sugarcane [15].

The SUCEST databank

Accordingly to the aforementioned, large EST collections have become available to explore the large allopolyploid sugarcane genome and consequently renewed the interest in sugarcane genetics [15, 18, 19], unrevealing the way for the discovery and identification of new genes that is essential

for breeding programs, either for transgenic plant development or for marker-assisted breeding. The Sugarcane EST project (SUCEST – *http://sucest.lad.ic.unicamp.br/en/*) generated 237,954 ESTs, which were assembled into 43,141 putative unique sugarcane transcripts from all major organs, harvested at several developmental stages [20]. To our knowledge this is one of the most gene rich databases representing a variety of tissues that were sampled from different cultivars, enabling transcript information accession of genes expressed in many biological sources, while the majority of plant EST projects have emphasized contrasting situations and have sampled either whole plant or certain tissues under stress conditions. The SUCEST databank is thus an excellent model for analyzing differential gene expression.

Plant defensins – old actors in the new era of biotechnology

All life forms have to defend themselves against microbial invaders - bacteria, fungi and viruses - and to do this they produce antimicrobial peptides such as defensins. Some defensin known to date possess as much power as penicillin as well as vancomycin [21]. It is believed that new defensins are even more potent and target certain microorganims more specifically than these antibiotics, including strains that are now resistant to conventional antibiotics. The discovery of novel candidates among such peptides has implications for the development of defensins as a treatment against many deadly infections, and may commence a new era of antibiotic discovery and development, a field that has not changed much since the discovery of penicillin by Alexander Fleming in 1929.

The finding that defensins are present in all plant species so far investigated and the existence of about 60.000 species of higher plants in Brazil [22] unravel a new horizon for exploring novel antimicrobial peptides. This is important either because increasing bacterial resistance to conventional antibiotics threatens the future of many antibiotics in current use or the recurrent application of pesticides in agriculture represents risks for environment and human health.

EST-based discovery of plant defensin – Bioinformatic steps and analyses

Small genes are generally difficult to predict by *ab initio* approaches. Moreover, plant sequence annotation programs have adopted cut off values that leave undetected potential small genes. Aside of being codified by small genomic regions, defensins are highly divergent in primary structures making them difficult to find by simple blast searches based on sequence similarity. Despite this diversity in amino acid sequences, bioinformatic analyses showed that defensins share common 3D structures, due to their stabilization by key amino acids with constrained mutation rates [23]. The observation of such key sites has led to the design of sequence motifs believed to correlate with the antimicrobial function of this peptide family, and thus useful for structure-based identification of such a divergent gene family. The development of bioinformatic tools for predicting patterns in biological sequences has now enabled the establishment of a routine to mine in plant EST databases for new defensins to posterior testing of their antimicrobial activity. This task encounters the recent needs of clean and efficient alternatives for pharmaceutical and agrobiotechnological applications.

Using PCR amplification, a putative sugarcane defensin (*SODEF*, *Saccharum officinarum* defensin - acc: ACB20518) has been previously reported by our laboratories [24]. The sequence amplified carried two exons and one intron that putatively encode a 30-amino acid long signal peptide coupled with a sequence displaying the eight-cysteine residue conservation, characteristic of mature defensins. Starting from this sequence as template for motif-based search, 17 defensin-like sequences were identified in several tissues of sugarcane. Fig. (**1**) shows the basic components of the routine used to identify defensin-like sequences in EST databases. Functional assignment is accomplished through an elaborate annotation procedure that consists of a central strategy plus expression profiling, phylogeny analyses and comparative modeling.

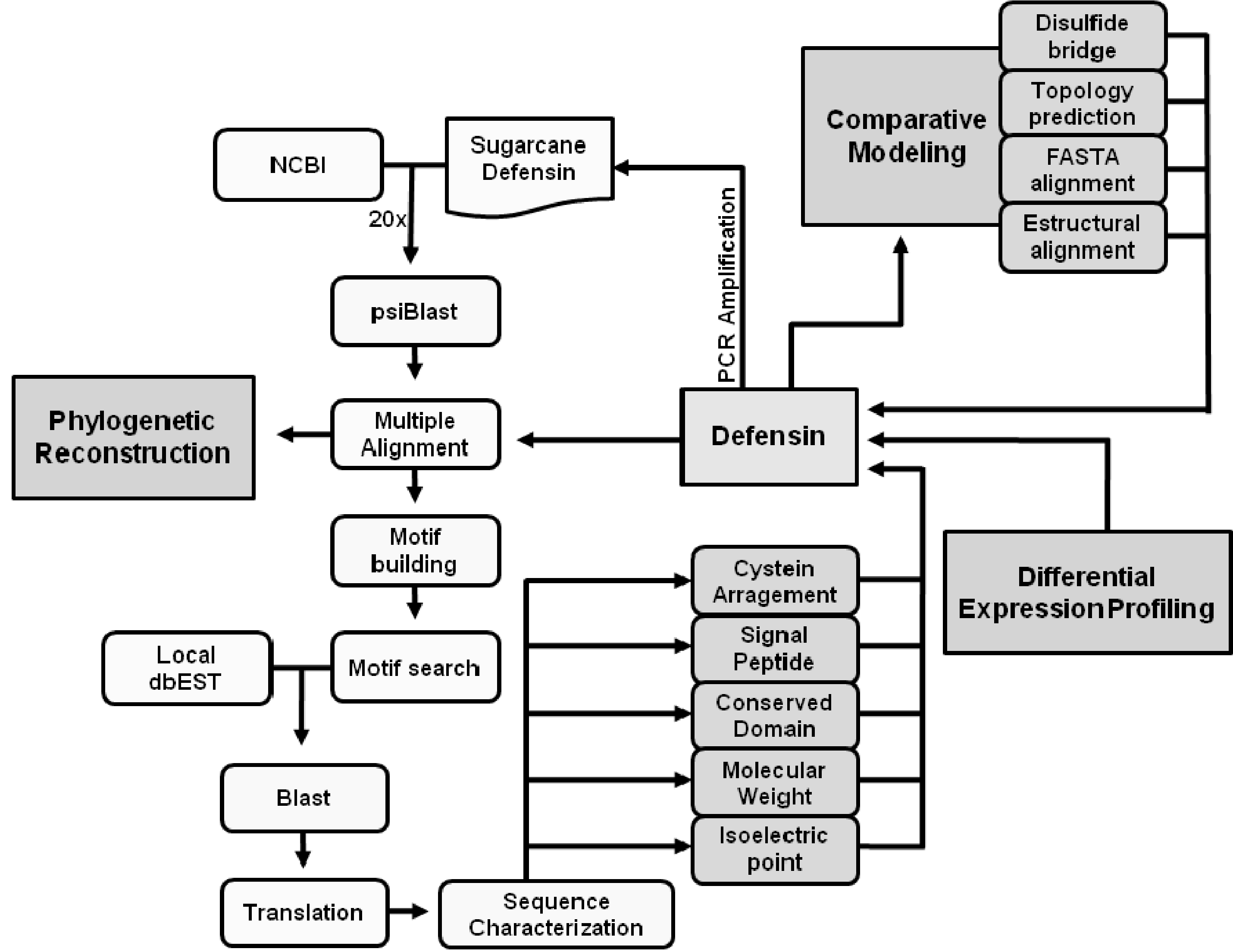


Figure 1 . Schematic representation of the routine application. Annotation is accomplished through a central strategy plus differential expression profiling, phylogeny analyses and comparative modeling. 20x corresponds to the number of interactions with the psiBLAST tool.

Motif Searching

The pro-peptide of SODEF was used as seed sequence in a psiBlast (Position specific iterative BLAST) search with 20 interactions on the NCBI non-redundant protein database, allowing for the identification of 229 distant related sequences for motif building [25]. The retrieved sequences were aligned to each other with ClustalW [26], without the signal peptide because they were too divergent to be used for motif design. Using the resulting alignment, a motif was built in the hmmbuild program, based on the statistical model hidden Markov model (HMM), and subsequently used with the program hmmsearch to screen the SUCEST database translated in all Six Frames Translation of Sequence [27]. This search retrieved 22 candidate sequences, five of which (~22%) were considered false positive because they lacked some key amino acids.

Sequence Annotation

The remaining candidate defensins obtained (*SODEF01* to *SODEF17*) were scrutinized to work out traces normally present in plant defensins. Further, the NCBI non-redundant protein and nucleotide databases were searched for similar sequences using Blastx and Blastn [28]. Based on the frame indicated in this step, the candidate nucleotide sequences were translated to amino acids using ORFfinder. These proceedings showed that the sugarcane sequences have a significant similarity with other plant defensins available in these databases.

Then, the isoelectric point and the molecular weight of the putative proteins were calculated with the program Jvirgel [29]. The proteins were predicted by Jvirgel to be highly basic and small with molecular weight ranging from 5.41 kDa to 8.6 kDa. The signal peptide was identified in the polypeptides with the program Philius that also predicted the overall topology of the proteins (*comparative modeling section*) [30]. All translated sequences presented the signal peptide, but its length varied among sequences Fig. (**2**). Subsequently, the translated proteins were submitted to rpsBlast (reverse specific position BLAST) to evaluate the conserved γ-thionin domain. All putative sugarcane defensins but the SODEF10 sequence presented the characteristic γ-thionin domain for

plant defensins. Moreover, the γ-core motif (common to all antimicrobial peptides) could be identified in all candidate sequences using multiple alignment as was also the Csαβ motif Fig. (**2**).

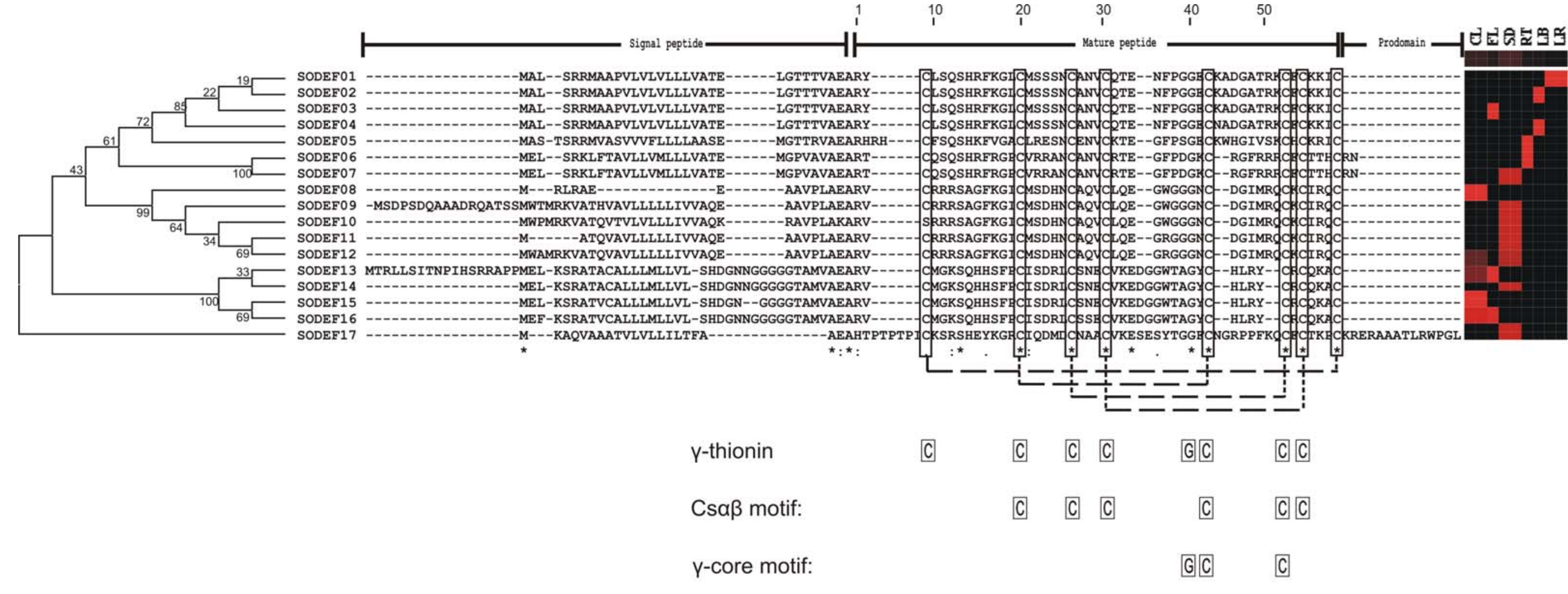


Figure 2. Comparison of amino acid sequences of sugarcane defensin-like. Evolutionary relationship is depicted left. Plant defensin hallmarks are highlighted below in the primary sequence. Linked bars are disulfide bonds: Cys1-Cys8; Cys2-Cys5; Cys3-Cys6; Cys4-Cys7. The expression profile is represented right: LR=Leaf roll; LB=Lateral Bark; RT=Root apex; FL=base of inflorescence; SD=Developing seed; CL=Calli.Dark red and light red correspond to higher and lower expression.

Differential Expression Profiling

The expression profile of the sugarcane defensin-like sequences was delineated with a hierarchical clustering approach applied to normalized data and observed in the TreeView program [31]. The expressed sequences displayed tissue-specific expression patterns, mostly in reproductive tissues Fig. (**2**).

Phylogeny Analysis

The evolutionary history of the defensin candidates was reconstructed with the program MEGA4 at the nucleotide level as well as at the protein level [32]. The dendrogram generated at protein level showed at least five clades including sugarcane defensins with high degree of identity among the member of each group Fig. (**2**). This grouping could also be observed at nucleotide level but with higher diversity at the UTR regions among sequences. Additionally, a phylogenetic inference was carried out with related amino acid sequences, including the sequence of the 3D templates, in order to deduce function relatedness clustering - the signal peptide was again removed in this step due to its exclusion of the mature peptide. The generated dendrogram with related proteins from diverse organisms split the candidates defensin up into two major groups but six internal groups were still delineated Fig. (**3**).

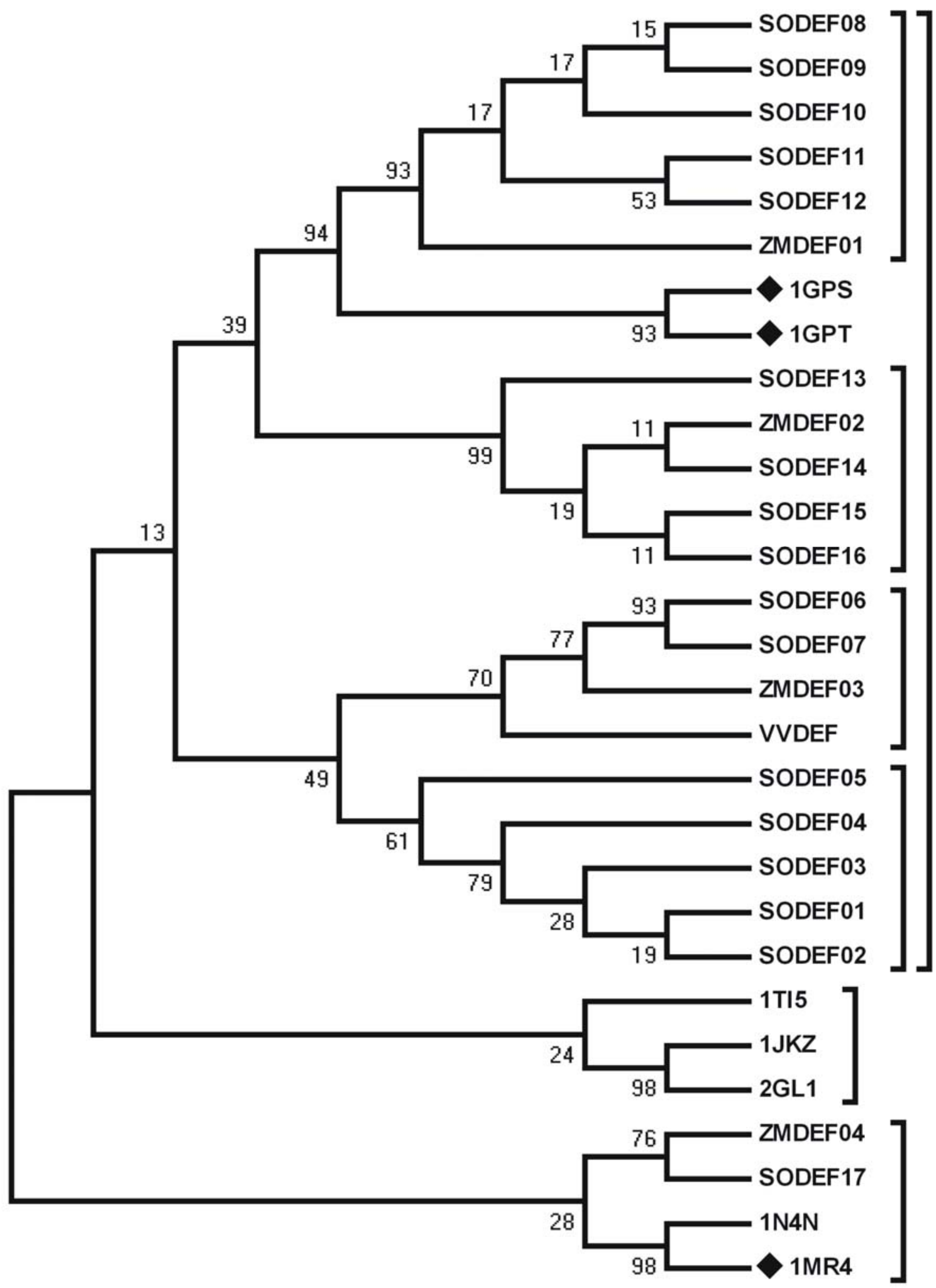


Figure 3 - Evolutionary reconstruction with Maximum Parsimony inference. SODEF01-17 (Sugarcane candidate defensin 1 to 17); ZMDEF1-4 (*Zea mays* defensin 1 to 4); VVDEF (*Vitis vinifera* defensin); 1GPS (γ-1-P thionin: *Triticum aestivum*); 1GPT (γ-1-P thionin: *Hordeum vulgare*); 1JKZ (Defensin: *Pisum sativum*); 2GL1 (Defensin: *Vigna radiata*); 1TI5 (Defensin: *Vigna radiata*); 1MR4 (Defensin: *Nicotiana alata*); 1N4N (Defensin: *Petunia* x *hybrida*). Square places the NMR structures used in the comparative modeling step. Numbers at the base of each clade correspond to bootstrap means at 1000 replications.

Comparative Modeling

The disulfide bonds that are known to stabilize the 3D structure of plant defensins were predicted with disulfind [33]. The typical four disulfide bonds were predicted to be formed among the eight canonical cysteines in all probable sugarcane defensins, except in the SODEF10 where the first cysteine involved in the formation of the fourth disulfide bridge was replaced by a serine Fig. (**2**). To detect best PDB templates and construct the structural models of the probable defensins, we submitted their amino acid sequences to the MHOLline workflow (*www.mholline.lncc.br*). The structures used as templates were defensins from *Triticum turgidum* (PDB1GPS), *Hordeum vulgare* (PDB1GPT) and *Nicotiana tabacum* (PDB1MR4) with growth inhibition activity toward fungal pathogens. Using MHOLline workflow the best templates found were PDB1GPS for sequences SODEF01-04 and SODEF06-07, PDB1GPT for sequences SODEF05 and SODEF08-12, and PDB1GPS+PDB1MR4 for SODEF17. The sequences SODEF05, SODEF13-16 and SODEF17 were not automatically modeled by MHOLline, thus they had their 3D structures constructed manually, via Modeller program [34, 35], using the best templates suggested by the MHOLline analysis, visualized and aligned by VMD [36].

The quality of 3D protein models were analyzed using Procheck [37], Molprobity [38] and QMEAN [39] programs. The modeled structures comprised a cysteine-stabilized αβ motif made up of one α-helix and a triple-stranded β-sheet, organized in βαββ architecture Fig. (**4**). The analysis performed by Procheck, Molprobity and QMEAN programs indicated that the obtained structural models have a good level of quality. The multiple sequence alignment Fig. (**2**) done with ClustalW and the structural alignments performed by VMD program Fig. (**4**), supported the presence of six distinct groups of probable defensins, with internal identities ranging from 95% to 100%. Fig. (**4**) shows the first 6 and final 14 amino acids of SODEF17 sequence that could not be modeled comparatively, due to template absence, but the final 14 residues probably constituted a pro-domain already observed in some floral defensins and thus is likely to be removed from the mature peptide.

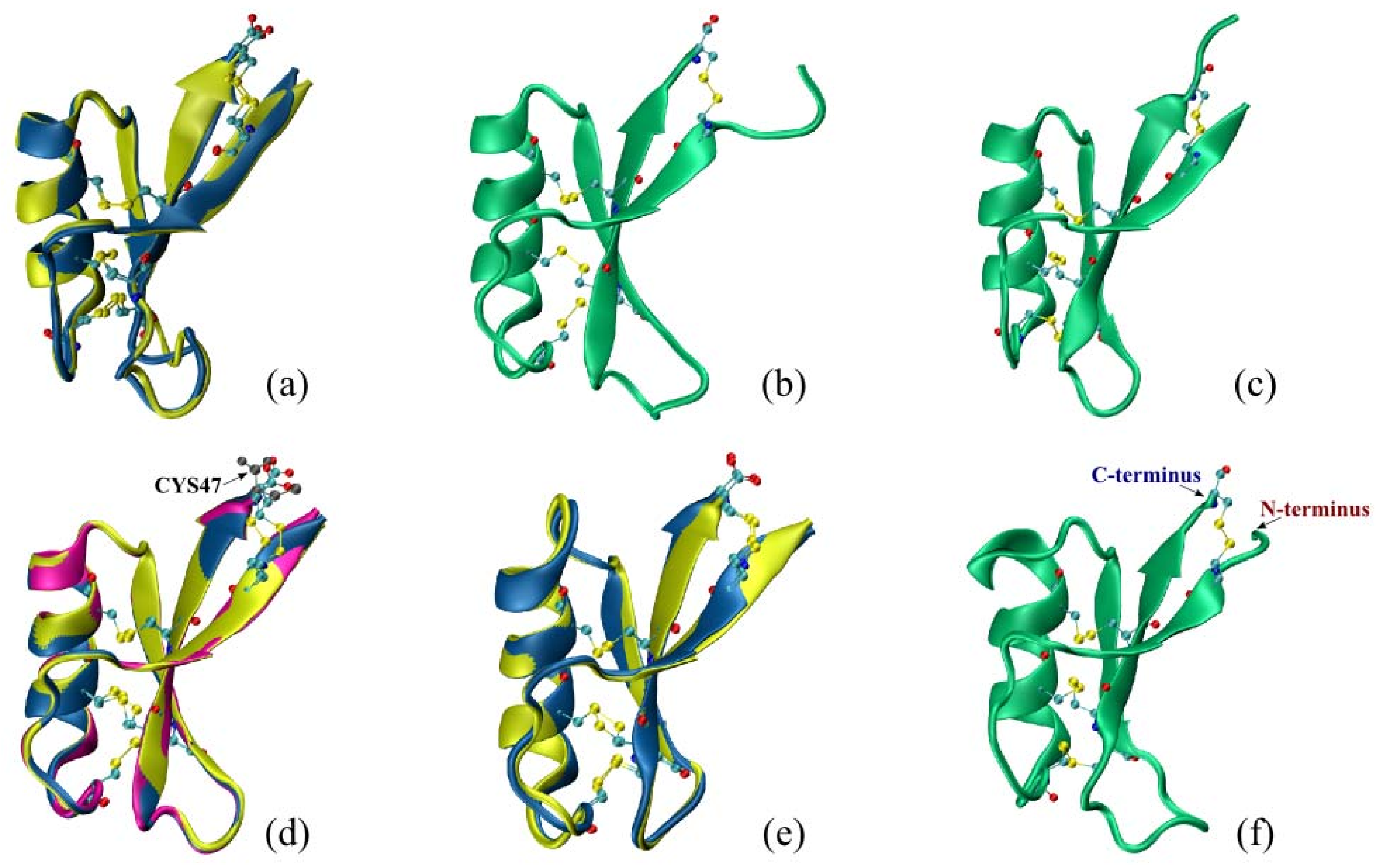


Figure 4. Structural alignment of the comparative modeled sequences of Sugarcane candidate defensins. (a) SODEF01 in yellow and SODEF04 in blue. (b) SODEF05 in green. (c) SODEF06 in green. (d) SODEF08 in yellow, SODEF10 in blue and SODEF11 in green. (e) SODEF13 in yellow and SODEF16 in blue. (f) SODEF17 in green, N-terminus in red arrow and C-terminus in blue arrow. The first 6 and final 14 amino acids of SODEF17 are not presented because these regions could not be modeled by comparative modeling. All disulphide bonds are highlighted in the figures above, with special attention to the mutation Cys1/SER1 that produces an unbound Cys8 (in grey) in the SODEF10 protein, presented in (d). All structural alignments and figures were made with the VMD program.

Defensins: a fairy tale of controlling infection disease in the genomics era.

Using PCR amplification, a putative sugarcane defensin has been previously reported by our laboratories [24]. Using this sequence as a start point for motif-based search, 17 defensin-like sequences were identified in several tissues of sugarcane.

The here described defensin-like sequences were grouped phylogenetically in six clades indicating that diversified defensin classes are constitutively expressed and may be responsible for the resistance against pathogens in sugarcane, as observed for other organisms Fig. (**2**). Considering the high level of amino acid identity (great than 97% identity) among identified sequences, it is to suppose that they are separated by recent divergence events, as it is the case in the related species that compose the hybrid and allopolyploid genome of sugarcane. Indeed, members of the same clade have minor amino acid sequence divergence, the major diversity among close related sequences being only in the UTR regions. Despite the conservation at amino acid level, the members seemed to be tissue and developmentally regulated and this could be correlated with the sequence diversity in the UTR regions, known to present regulatory elements. Another important observation is the specific expression of defensins in seeds and other reproductive tissues. We have noted expression in the root apex, base of inflorescence, lateral bud, developing seeds and meristematic tissues (sugarcane internodes and calli) Fig. (**2**). It has been proposed that this expression could be implied in the prevention from pathogen attack of nutrient-rich resources [40].

The observation of a major clade containing only dicot defensins plus a sugarcane candidate suggest that some defensins in sugarcane were already present before the divergence of the major plant lineages while others may have arisen after this event. In fact, the minor groups observed in the dendrogram also give evidence of conserved sequences across diverse plant taxa as well as taxon specific genes Fig. (**3**). The sugarcane probable defensins had their structures obtained by comparative modeling and showed distinctive structure although all the sequences have exhibited the linear hallmarks of plant defensins Fig. (**4**). The difference in the structure resided primarily in the loop and turn regions that show important solvent exposed sites, responsible for antifungal and

bactericidal activity. Thus, it is reasonable to believe that these sugarcane defensin candidates have diverse activities or even target different microorganisms.

Conclusions

All putative members of sugarcane SODEF family held the features observed in plant defensins, though some of them present extra features that may be reflected in the individual structure. The expression pattern observed, the evolutionary relationship to plant defensins already described for other crops, and the common 3D-structural similarity together suggest an antimicrobial activity for the here presented putative defensin-like sequences from sugarcane, revealing that these sequences may exert their activity in diverse tissues by different ways and toward distinct or overlapping groups of pathogens. The here used routine for "mining" defensins in sugarcane proved to be efficient, with the advantage that tools and softwares used were publicly available in the web. Those interested in "mining" defensins have notwithstanding to handle with a plethora of programs. Therefore, this study sets the basis for the development of a tool to mine such small peptides as well as opens the way for further computational and experimental investigation of their antimicrobial activity and potential protective role in this important crop.

Acknowledgements

The authors thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) and FACEPE (Fundação de Amparo à Pesquisa do Estado de Pernambuco, Brazil) for the concession of fellowships. None of the authors has any potential financial conflict of interest related to this manuscript.

References

[1] Wang, X.-J.; Gaasterland, T.; Chua, N.-H. Genome-wide prediction and identification of cis-natural antisense transcripts in *Arabidopsis thaliana*. *Genome Biol.,* **2005**, *6*, 30.

[2] Wang, H.; Chua, N.-H. and Wang, X.-J. Prediction of transantisense transcripts in *Arabidopsis thaliana*. *Genome Biol.*, **2006**, *7*, 92**.**

[3] Vazquez, F. *Arabidopsis* endogenous small RNAs: highways and byways. *Trends in Plant Sci.,* **2006**, *11*, 460–468.

[4] Ma, J.; Morrow, D.J.; Fernandes, J.; Walbot, V. Comparative profiling of the sense and antisense transcriptome of maize lines. *Genome Biol.*, **2006**, *7*, 22.

[5] Molina, C.; Rotter, B.; Horres, R.; Udupa, S.M.; Besser, B.; Bellarmino, L.; Baum, M.; Matsumura, H.; Terauchi, R.; Kahl, G.; Winter, P. SuperSage: the drought stress-responsive transcriptome of chickpea roots. *BMC Genomics*, **2008**, *9*, 553.

[6] Kido, E.A.; Pandolfi, V.; Houlou-kido, L.M.; Andrade, P.P.; Marcelino, F.C.; Nepomuceno, A.L.; Abdelnoor, R.V.; Burnquist, W.L.; Benko-Iseppon, A.M. Plant antimicrobial peptide: An overview of SuperSage transcriptional profile and a functional review. *Curr. Prot. Pept.* In press. (this volume).

[7] Zhang, B.H.; Pan, X.P.; Wang, Q.L.; Cobb, G.P.; Anderson, T.A. Identification and characterization of new plant microRNAs using EST analysis. *Cell Res.,* **2005**, *15*, 336–360.

[8] Zhang, B.H.; Pan, X.P.; Anderson, T.A. Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett.*, **2006**, *580*, 3753–3762.

[9] Allen, E.; Xie, Z.; Gustafson, A.M. and Carrington, J.C. microRNA-directed phasing during Trans-acting siRNA biogenesis in plants. *Cell*, **2005,** *121*, 207–221.

[10] Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Genet.*, **2006**, *15*, 17–29.

[11] Menossi, M.; Silva-Filho, M.C.; Vincentz, M.; Van-Sluys, M.-A.; Souza, G.M. Sugarcane Functional Genomics: Gene Discovery for Agronomic Trait Development. *Int. J. Plant Genom.*, **2008**, ID 458732.

[12] Wheeler, D.L.; Barret, T.; Benson, D.A.; Bryant, S.H.; Canese, K.; Chetvernin, V.; Church, D.M.; DiCuccio, M.; Edgar, R.; Federhen, S.; Feolo, M.; Geer, L.Y.; Helmberg, W.; Kapustin, Y.; Khovaykho, O.; Landsman, D.; Lipman, D.J.; Madden, T.L.; Maglott, D.R.; Miller, V.; Ostell, J.; Pruitt, K.D.; Schuler, G.D.; Shumway, M.; Sequeira, E.; Sherry, S.T.; Sirotkin, K.; Souvorov, A.; Starchenko, G.; Tatusov, R.L.; Tatusova, T.A.; Wagner, L.; Yaschenko, E. Database resources of the National Center for Biotechnology Information. *Nucl. Acids Res.*, **2008**, *36*, D13–D21.

[13] Duvick, J.; Fu, A.; Muppirala, U.; Sabharwal, M.; Wilkerson, M.D.; Lawrence, C.J.; Lushbough, C.; Brendel, V. PlantGDB: a resource for comparative plant genomics. *Nucl. Acids Res.*, **2008**, *36*, D959–D965.

[14] Childs, K.L.; Hamilton, J.P.; Zhu, W.; Ly, E.; Cheung, F.; Wu, H.; Rabinovicz, P.D.; Town, C.D.; Buell, C.R.; Chan, A.P. The TIGR Plant Transcript Assemblies database. *Nucl. Acids Res.*, **2006**, *35*, D846–D851.

[15] Grivet, L.; Arruda, P. Sugarcane genomics: depicting the complex genome of an important tropical crop. *Curr. Opin. Plant Biol.*, **2001**, *5*, 122–127.

[16] D'Hont, A. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogen. Genome Res.*, **2005**, *109*, 27–33.

[17] D'Hont, A.; Glaszman, J.C. Sugarcane genome analysis with molecular markers, a first decade of research. *Proc. Int. Soc. Sugar Cane Techn.*, **2001**, *24*, 556–559.

[18] Butterfield, M.K.; D'Hont, A.; Berding, N. In *The sugarcane genome: a synthesis of current understanding, and lessons for breeding and biotechnology*. Proceedings of the South African Sugar Technologists' *Association (SASTA '01)*, Durban, South Africa, **2001**, *75*, 1–5.

[19] Ming, R.; Moore, P.H.; Wu, K.K.; D'Hont, A.; Glaszmann, J.C.; Tew, T.L.; Mirkov, T.E.; da Silva, J.; Jifon, J.; Rai, M.; Schnell, R.J.; Brumbley, S.M.; Lakshmanan, P.; Comstock, J.C.; Paterson, A.H. Sugarcane improvement through breeding and biotechnology. *Plant Breed. Rev.*, **2006**, *27*, 15–118.

[20] Vettore, A.L.; Da Silva, F.R.; Kemper, E.L.; Arruda, P. The libraries that made SUCEST. *Gen. Mol. Biol.*, **2001**, *24*, 1–7.

[21] Mygind, P.H.; Fischer, R.L.; Schnorr, K.M.; Hansen, M.T.; Sönksen, C.P.; Ludvigsen, S.; Raventós, D.; Buskov, S.; Christensen, B.; De Maria, L.; Taboureau, O.; Yaver, D.; Elvig-Jørgensen, S.G.; Sörensen, M.V.; Christensen, B.E.; Kjaerulff, S.; Frimodt-Moller, N.; Lehrer, R.I.; Zasloff, M.; Kristensen, H.H. Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. *Nature*, **2005**, *437*, 975–980.

[22] Giulietti, A.M.; Harley, R.M.; Queiroz, L.P.; Wanderley, M.G.L.; Berg, C.V.D. Biodiversidade e conservação das plantas no Brasil. *Megadiversidade*, **2005**, *1*, 52–61.

[23] Yount, N.Y.; Yeaman, M.R. Multidimensional signatures in antimicrobial peptides. *PNAS*, **2004**, *101*, 7363–7368.

[24] Padovan, L.; Segat, L.; Tossi, A.; Antcheva, N.; Benko-Iseppon, A.M.; Ederson, A.K.; Brandao, L.; Calsa, T.J.; Crovella, S. A plant defensin from sugarcane (*Saccharum* spp). *Protein Peptide Letters,* **2009**, *16*, 430–436.

[25] Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.,* **1997**, *25*, 3389–3402.

[26] Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.,* **1994**, *22*, 4673–4680.

[27] Durbin, R.; Eddy, S.R.; Krogh, A.; Mitchison, G. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press Cambridge, UK, **1998**.

[28] Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Alignment Search Tool. *J. Mol. Biol.*, **1990**, *215*, 403–410.

[29] Hiller, K.; Schobert, M.; Hundertmark, C.; Jahn, D.; Munch, R. JVirGel: Calculation of Virtual Two-Dimensional Protein Gels. *Nucl. Acids Res.,* **2003**, *31*, 3862–3865.

[30] Reynolds, S.M.; Käll, L.; Riffle, M.E.; Bilmes, J.A.; Noble, W.S. Transmembrane Topology and Signal Peptide Prediction Using Dynamic Bayesian Networks. *PLoS Comput. Biol.*, **2008**, *4*, e1000213.

[31] Eisen, M.B.; Spellman, P.T.; Brown, P.O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *PNAS*, **1998**, *95*, 14863–14868.

[32] Tamura, K.; Dudley, J.; Nei, M. and Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution, **2007**, 24, 1596–1599.

[33] Ceroni, A.; Passerini, A.; Vullo, A.; Frasconi, P. Disulfind: a disulfide bonding state and cysteine connectivity prediction server. *Nucl. Acids Res*., **2006**, *34*, W177–W181.

[34] Braun, W.; Go, N. Calculation of protein conformations by proton–proton distance constraints. A new efficient algorithm. *J. Mol. Biol.,* **1985**, *186***,** 611–626.

[35] Eswar, N.; Eramian, D.; Webb, B.; Shen, M.Y.; Sali, A. Protein Structure Modeling with Modeller. *Meth. Mol Biol*., **2008**, *426*, 145–159.

[36] Hsin, J.; Arkhipov, A.; Yin, Y.; Stone, J.E.; Schulten, K. Using VMD: an introductory tutorial. *Curr. Protoc. Bioinf.,* **2008**, *24*, 5.7.

[37] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.,* **1993**, *26*, 283–291.

[38] Lovell, S.C.; Davis, I.W.; Arendall III W.B.; de Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. *Prot.: Struc. Funct. Genet.,* **2003,** *50*, 437–450.

[39] Benkert, P.; Künzli, M.; Schwede, T. QMEAN Server for Protein Model Quality Estimation. *Nucl. Acids Res.,* **2009**, *37*, W510–W514.

[40] Silverstein, K.A.T.; Graham, M.A.; Paape, T.D.; VandenBosch, K.A. Genome Organization of More Than 300 Defensin-Like Genes in *Arabidopsis*. *Plant Physiology,* **2005**, *138*, 600–610.

Manuscrito de artigo científico 4

Benko-Iseppon *et al.*, 2010

Plant Antimicrobial Peptides

Current Protein and Peptide Science, aceito em Outubro de 2009

Plant Antimicrobial Peptides

Ana Maria Benko-Iseppon[1*], Suely Lins Galdino[1], Tercilio Calsa Jr.[1], Ederson Akio Kido[1], Alessandro Tossi[2], L. C. Belarmino[1], Sergio Crovella[1,3]

[1]Federal University of Pernambuco, Center of Biological Sciences, Genetics and Antibiotics Depts. - Av. Prof. Moraes Rego, 1235, Recife, Brazil.

[2]University of Trieste, Department of Life Sciences - Via Giorgieri, 1 – 34127, Trieste, Italy.

[3]University of Trieste, Genetic Service, IRCCS Burlo Garofolo and Department of Developmental and Reproductive Sciences -Via dell'Istria, 65/1 - 34137 Trieste, Italy.

*Corresponding Author

Ana Maria Benko Iseppon

Universidade Federal de Pernambuco, Genetics Dept. (UFPE/CCB/Genetica)

Av. Prof. Moraes Rego, 1235, CEP 50.670-420, Recife, PE, Brazil.

Phone: 55-81-2126-8620 and 2126-7816

Fax: 55-81-2126-8522

E-mail: ana.benko.iseppon@pq.cnpq.br

Abstract

Mechanisms related to biotic interactions, such as pathogen attack, herbivory and symbiosis are important challenges to higher plants and have been widely studied especially for breeding purposes. The present review focuses on a special category of defense molecules, the plant antimicrobial peptides, providing an overview of their main molecular features and structures.

Key words: pathogenesis related (PR) protein, defensin, thionin, lipid transfer protein (LTP), hevein-like protein, knotin-type protein, snakin, cyclotide.

Introduction

In plants and animals, the surface barriers are the first obstacle to pathogen penetration. In animals, cutaneous tissues, mucous membranes and respiratory tract lining fluid represent the first non-specific protection against the infections. Analogously, the external natural barriers of the plants include leaf epicuticular layers, suberized, cutinized and lignified epidermal tissues. Plants are known to present both constitutive and inducible resistance mechanisms against pathogen attack. Defensive weapons include morphological barriers, secondary metabolites (e.g. phytoalexins), and antimicrobial proteins that in combination impair pathogen invasion. Countless expression assays have shown, in the past decade that a single plant/pathogen interaction (either compatible or incompatible) is able to recruit or silence hundreds of genes, many of them already known, while others which remain to be described.

Immediately after perception of a pathogen by the products of specific *R* (Resistance) genes, a signal cascade is released inducing secondary non-specific (systemic) defense mechanisms. The first specific recognition mechanism relies on transmembrane pattern recognition receptors that respond to slowly evolving microbial- or pathogen-related molecular patterns (PRAMPs) [1].

The second acts largely inside the cell, using the polymorphic NBS-LRR (Nucleotide Biding Site-Leucine Rich Repeat) protein products encoded by most *R* genes [2, 3, 4]. For each existing *R* gene in the host plant there will be a corresponding *avr* (avirulence) gene in the pathogen. The plant will be resistant and invasion by the pathogen will be impaired when both *R* and *avr* genes are present and compatible, leading to a hypersensitive response (HR) that leads to local cell death, preventing the spreading of the pathogen throughout the plant [5, 6]. Besides this local reaction, the HR activates a signal cascade that induces systemic defense to a wide spectrum of pathogens. This corroborates lines of evidence recognized by researchers more than 100 years ago that like animals plant may be "immunized" against a given pathogen once exposed to another pathogen or derived elicitor molecules [7].

After the contact with the pathogen and activation of the specific defense system, a series of secondary responses is initiated, including reinforcement of cell walls, the production of plant antibiotics (phytoalexins), and the synthesis of defense-related factors, so called PR (Pathogenesis Related) defense proteins [8, 9, 10].

The present mini-review aims to provide a brief introduction to the main classes of plant AMP peptides, their features, structures and sources of isolation.

Antimicrobial Peptides and Pathogenesis Related (PR) Defense Proteins

PR defense mechanisms react against many classes of pathogenic agents, including oomycetes, fungi, bacteria, viruses, or even insects [10]. Resistance and susceptibility does not depend only upon the 'quality' of the activated defense genes, in many cases associated with differences in the timing and magnitude of their expression, but also on the contemporary expression of different sets of genes [11].

Pathogenesis-related proteins have been classified into 17 families (see revision in [10, 12]); with some seeming to occur more specifically in particular plant species. Most PR proteins are induced through the action of the signaling compounds salicylic acid, jasmonic acid, or ethylene, and possess antimicrobial activities *in vitro* through lytic activities on cell walls, contact toxicity, and perhaps an involvement in defensive signaling [10].

Among the different PR groups, three (namely PR-12, 13 and 14) consist of proteins with a molecular size below 10 kDa, which are encoded by a single gene, have overall net positive charge, tolerance to acid and organic solvents, thermostability, and a wide range of biological activities, with emphasis on specific antimicrobial activities [10, 12], characteristics which are in common with antimicrobial peptides (AMPs) [13].

In mammals, including humans, AMPs are particularly abundant in leukocytes and epithelial cells, where they are constitutively expressed and also induced by infections [14]. In plants AMPs are secreted in most, if not all plant species, and are found constitutively in storage organs (i.e. seeds) and peripheral cell layers of generative tissues (reproductive organs, fruits and flowers), besides

being induced in vegetative tissues, following infection or wounding [15]. Figure 1 brings a general overview of the main mechanisms of plant defense signaling, with emphasis on *PR* genes and AMPs.

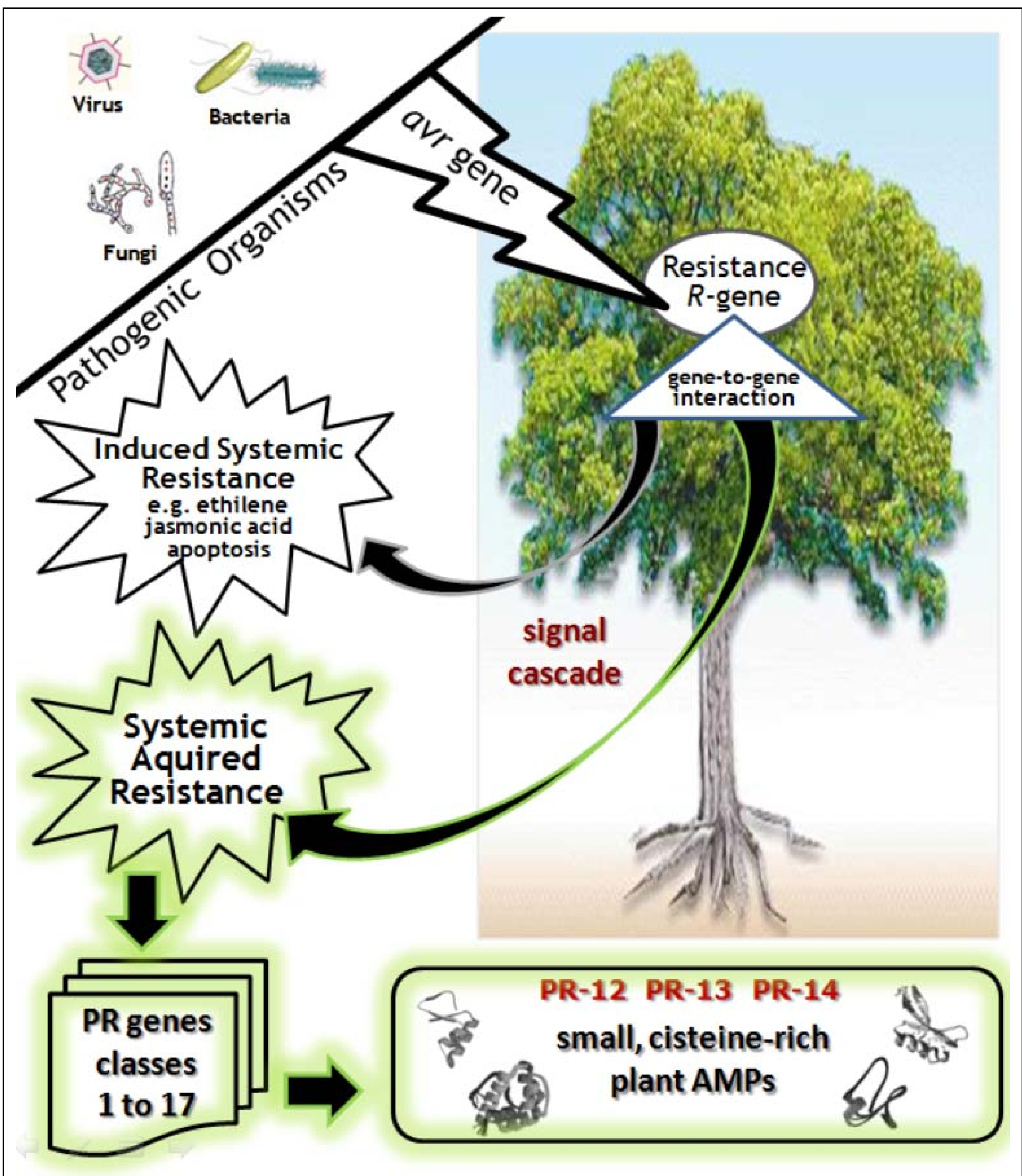


Figure 1. Main mechanisms of pathogen recognition and defense in plants. Pathogenic organisms (mainly virus, bacteria and fungi) secrete *avr* (avirulence) gene products that may be compatible with *R* gene products secreted by the plants. Compatible interactions lead to the activation of signal cascades inducing systemic resistance factors (as ethylene and jasmonic acid) and acquired resistance represented by 17 *PR* gene categories. From these, three categories (*PR-12*, *13* and *14*) include small cisteine-rich anti-microbial (AMP) peptides.

AMPs are polypeptides with less than 200 amino acids in length (very often less than 50), found in host defense settings, and exhibiting antimicrobial activity at physiologic ambient conditions and peptide concentrations [15]. Most of the AMPs are cationic at physiological pH due to an excess of positively charged residues such as arginine and lysine compared to negatively charged ones. The cationic nature of AMPs, associated with a tendency to adopt amphipathic structures (net separation of charged/polar and hydrophobic surfaces), facilitates their interaction and insertion into the anionic cell walls and phospholipid membranes of microorganisms [16].

In plants, besides acting locally at the sites of infection, they also accumulate in uninfected districts from the infection point, a phenomenon known as systemic acquired resistance (SAR) [17]. These proteins are collectively part of the innate immune system and provide a relatively rapid response of the hosts, with a comparatively lower energy cost, when compared to the adaptive immune system of higher vertebrates or to the production of secondary metabolites resulting from complex metabolic pathways.

Main AMP Categories

Many AMPs are rich in a given amino acid type. In plants, these peptides share some common characteristics: besides their small molecular weight (~10 kDa) and highly basic character, they frequently contain an even number of cysteine residues that stabilize the protein structure through formation of disulfide bridges [18].

There are several classes of such peptides (Table 1), including plant defensins, thionins, lipid transfer proteins, hevein-like proteins, and knotin-type proteins, besides the antimicrobial peptides from *Macadamia integrifolia* and *Impatiens balsamina*. Circular peptides known as cyclotides form a class of uncommon cysteine-rich peptides found in plants of the families Rubiaceae and Violaceae [19-22].

In general, while AMPs are expressed in several tissues, they are better represented in tissues that are in constant contact with the external environment, and thus continuously exposed to the

microbial biota. These peptides are generally produced as pre-proteins, but some are produced as larger precursors with a C-terminal pro-domain (additional to the N-terminal signal peptide as in animals). There is a considerable variability regarding amino acid sequences, but a peptide can be designated as a member of a given family according to a specific and conserved arrangement and connectivity of the cysteine residues, even in the absence of a high level of sequence homology (Table 1). Despite this amino acid variability, members of the same family show a folding pattern that is globally comparable (Table 1), adopting a tridimensional structure that involves the formation of secondary structural elements such as β-sheets and α-helices that are stabilized by the intramolecular disulfide bridges.

Despite the abundance of different families, and the presence of some species-specific AMPs, the main groups of plant AMPs are represented by defensins, thionins, lipid transfer proteins, cyclotides and snakins.

Defensins

The first antifungal plant defensin was identified in *Urtica dioica* L. roots and rhizomes [39], but other similar peptides have been successively isolated from several species and organs - for example defensins from radish leaves (*Raphanus sativus* L. [40]), pepper fruit (*Capsicum annum* L. [41]), mung bean seeds (*Vigna radiata* (L.) R. Wilcz. [42], *Petunia* x *hybrida* Vilm. Seeds [43], tobacco and *Petunia* flowers (*N. tabacum* [44]), bean seeds (*Phaseolus vulgaris* L. cv. 'white cloud', [45]); sugarcane (*Saccharum* spp. [46]) and cowpea leaves (*Vigna unguiculata* (L.) Walp.; Padovan *et al*., in press*).

Plant defensins are small (~5 kDa, 45 to 54 amino acids), basic, cysteine-rich peptides. Plant defensins identified so far have 8 cysteines that form four structure-stabilizing disulfide bridges. Their three-dimensional structure comprises a triple-stranded β-sheet with an aligned α-helix , and is stabilized by four disulfide bridges connecting these structural elements [47-49].

*Padovan, L.; Segat, L.; Tossi, A.; Calsa Jr, T.; Kido, E.A.; Brandão, L.; Guimarães, R.L.; Pestana-Calsa, M.C.; Pandolfi, V.; Belarmino, L.C.; Benko-Iseppon, A.M.; Crovella, S. Characterization of a new defensin from cowpea (*Vigna unguiculata* (L.) Walp.). *Protein Pept. Lett.*, **2009**, in press.

Table 1: Small, Cys-rich antimicrobial peptides from the plant kingdom, including their classification and structural features. PDB refers to the Protein Data Base, DB to number of Disulfide Bridges.

Family	Name	Protein Structure	PDB entry	DB	Cys Arrangement*	Reference
Plant Defensin	Rs-AFP2		1AYJ	4	3-C-10-C-5-C-3-C-9-C-8-C-1-C-3-C	[23]
α- and β-Thionin (8-Cystein-type)	Alpha-1-purothionin		1BHP	4	2-CC-7-C-3-C-8-C-3-C-1-C-8-C-6	[24]
Thionin (6-cystein-type)	Crambin		1AB1	3	2-CC-11-C-9-C-5-C-7-C-6	[25]
non-specific Lipid transfer Protein	Ace-AMP1		1T12	4	3-C-9-C-12-CC-18-C-1-C-23-C-15-C-4	[26]
Hevein-like protein	Ace-AMP2		1HEV	4	3-C-4-C-4-CC-5-C-6-C-2	[27]
Knotin	Mj-AMP1		1DKC	3	1-C-6-C-8-CC-3-C-10-C-3	[28]
Macadamia	MiAMP1		1C01	3	10-C-9-C-1-C-25-C-14-C-11-C	[29, 30]
Impatiens	ib-AMP1	-	-	-	5-CC-8-C-3-C	[31, 32]
Maize-AMP	MBP-1	-	-	-	6-C-3-C-13-C-3-C-4	[33]
Puroindolines	Puroindoline A	-	-	-	10-C-8-C-7-C-9-WR[W]$_2$K[W]$_2$K-C-6-C-9-CRC-35-C-5-C-7	[34, 35]
Snakin	SN1	-	-	-	[X]$_n$-C-3-C-2-RC-8-C-3-C-2-CC-2-C-1-CVP-1-G-2-GN-3-C-1-CY-10-KCP	[36, 37]
Cyclotide	Kalata B1		1BH4	3	1-C-3-C-4-C-4-C-1-C-4-C-6	[38]

*Cys arrangement in antimicrobial peptides is interpolated with the number of amino acids frequently observed between conserved Cys-residues.

The name "plant defensin" was introduced in 1995 by Terras and colleagues [17], who isolated two antifungal proteins (Rs-AFP1 and Rs-AFP2) from radish (*Raphanus sativus* L.) seeds. Plant defensins have a widespread distribution throughout the plant kingdom and are likely to be present in most, if not all, plants [50-52].

Most plant defensins have been isolated from seeds, where they are abundant, and many have been characterized at molecular, biochemical and structural levels [52]. Defensin expression has also been observed in other tissues including leaves [17, 53], pods [54], tubers [55], fruits [41], roots [56] and floral tissues [55].

Over the last two decades, numerous plant defensins have been purified with subsequent test of their biological activity [51, 57]. While several different methods have been reported for defensin purification, many of these rely on the intrinsic physio-biochemical properties of the peptides such as their small size, overall net positive charge, tolerance to acids and organic solvents, and their thermostability. A review of defensin purification and production methods is given in Padovan *et al.**.

Thionins

Thionins are a group of small antimicrobial of the PR-13 protein group present in higher plants. They are basic peptides, with low molecular weight (~5 kDa), rich in basic and sulfur-containing residues (arginine, lysine and cysteine). Thionins are predominantly located intracellularly, but are also found in the extracellular space, being reported in several plant tissues, such as seeds, stems and roots [58]. Different members present a high level of sequence and structure similarity, besides displaying toxic effects against bacteria, fungi, yeast, animal and plant cells [59].

Studies analyzing thionin sequences and three-dimensional structures in plants have demonstrated that they are able to act directly on the cellular membranes of pathogens. Additionally, a conserved DNA-binding motif was identified, suggesting that they might interact with DNA within the plant cells [60]. Furthermore, thionin expression in transgenic plants has been found to protect against

*Padovan, L.; Crovella, S.; Tossi, A.; Segat, L. Techniques for plant defensin production. *Curr. Prot. Pept. Sci.* (present volume).

pathogenic bacteria in many plant species, including *Arabidopsis thaliana* (L.) Heynh. [61], rice (*Oryza sativa* L. [62] and tobacco (*Nicotiana tabacum* L. [63]).

Lipid Transfer Proteins

Membrane biogenesis involves movement of lipids from their sites of synthesis, especially the endoplasmic reticulum, to other organelles like chloroplasts and mitochondria. The search for carrier proteins led to the discovery of lipid transfer proteins (LTPs) in various plant species. The development of adequate assays to determine lipid transfer activity constitutes one of the most important steps of LTPs characterization. These assays are based on the monitoring of marked lipids transferring from a donor membrane to an acceptor membrane. The donor membranes are generally liposomes, while the acceptors are cytoplasmic organelles, like chloroplasts and mitochondria, which can be easily separated from liposomes by centrifugation [59]. The first plant LTP was discovered 35 years ago in potato (*Solanum tuberosum* L. [64]) tubers being so denominated due to their *in vitro* ability to facilitate the transfer of phospholipids between a donor and an acceptor membrane [65].

Plant lipid transfer proteins (LTP) are small cationic peptides, subdivided into two families, both comprising conserved patterns of cysteine residues and a three-dimensional structure with an internal hydrophobic cavity that corresponds to the lipid binding site [66]. Members of the LTP1 family are approximately 10 kDa, have 90–95 amino acid residues (with 8 cysteine residues in conserved positions), and have isoeletric points (pI) between 9 and 10 [65, 66]. The LTP2 family includes smaller peptides (~7 kDa), possessing on average 70 amino acids, 4 conserved disulfide bridges, and high pI in common with the LTP1 family [67, 68, 69]. Both families present an N-terminal signal peptide of 21-27 amino acids for the LTP1 family, and 27- 35 amino acids, for the LTP2 family. This signal peptide is excised in the mature peptide and serves to target the LTPs to the cell secretory pathway where they are exported to the apoplast [65, 66].

LTP1 of various plants species are localized at the cell wall, as demonstrated in castor bean (*Ricinus communis* L. [70]), *A. thaliana* [71], broccoli leaves (*Brassica oleracea* L. var. *italica* [72]) and in cowpea seeds (*V. unguiculata* [73]). These findings were inconsistent with the hypothesis of a biological role for the LTPs in the plant cell cytoplasm. In addition to the given lipid transfer function, which is inconsistent with their cell wall location [74], other roles, with emphasis on plant defense against phytopathogen attack, have been described in various plants, including radish seeds (*Raphanus sativus* L.; [23]), maize leaves (*Zea mays* L. [75]), arabidopsis and spinach (*Spinacea oleracea* L.) leaves [76] , wheat seeds (*Triticum aestivum* L. [77], sunflower seeds (*Helianthus annuus* L. [78]), chili pepper seeds (*C. annuum* [79]) and from rape seeds (*Brassica campestris* L. [80]).

Cyclotides

Cyclotides are plant-derived proteins of 28-37 amino acids in size and characterized by the unique structural features of a cyclic peptide backbone and a knotted arrangement of three conserved disulphide bonds joined in a so called 'cystine knot' motif. Cyclotides are ultra-stable plant proteins because of their circular peptide backbone cross-linked by the cystine knot , being exceptionally resistant to thermal, chemical and enzymatic degradation [81]. They were found initially in large quantities in plants of the Violaceae and Rubiaceae families [81, 82]. Additional studies revealed that these peptides prevail in some taxa within Rubiaceae, being present also in the family Apocynaceae, revealing a wider distribution and abundance than initially supposed [83].

Cyclization is achieved by recruitment of plant proteolytic enzymes and operating them in 'reverse' to form a peptide bond between the N- and C-termini of a linear precursor. It has been suggested that circular proteins are more common in the plant kingdom than previously thought, while their exceptional stability has led to their application as protein-engineering templates in pharmaceutical applications [83, 84].

Snakins

Snakins are cysteine-rich peptides present in different plants species with a broad-spectrum of antimicrobial activities *in vitro.* Two described peptides represent this class: Snakin-1 (StSN1) [36] and Snakin-2 (StSN2) [37] both isolated from potato (*Solanum tuberosum* cv. Jaerla) tubers. They present a wide spectrum of antimicrobial activities – different from those of defensin peptides from the same tissues – against specific bacterial (*Ralstonia solanacearum* and *Erwinia chrysanthemi*) and fungal (*Botritys cinerea*) pathogens.

Snakin-1 is a ubiquitous 63-residue peptide with 12 cysteines, while Snakin-2 (St-SN2) is a paralogue of St-SN1 (also containing 12 cysteine residues) with a low sequence similarity (38%). Snakins are upregulated on wounding and fungal infection and downregulated by bacterial infections. Orthologues of snakins have been detected in other plants, including tomato [85], *A. thaliana* [86], *Gerbera* x *hybrida* Hort. [87], *Fragraria* x *annanasa* Duch., *Ricinus communis* and *Petunia* x *hybrida* [88]. Experimental assays demonstrated that StSN2 is upregulated by gibberellin and downregulated by abscisic acid, in contrast to St-SN1, which does not respond to gibberellin [37].

Final Considerations

Interestingly, plant AMPs display antimicrobial activity not only against plant pathogens but also against fungi, viruses, parasites, Gram-positive and Gram-negative bacteria that infect humans [89, 90, 91]. The principal mechanism of AMPs for rapid killing of microbial pathogens is attributed to perturbation of the microbial cell membrane [92, 93]. In addition to their antimicrobial role, AMPs also serve as important effector molecules during inflammation, immune activation, and wound healing [94-96]: for the above mentioned reasons, plant AMPs are being considered as amongst the most promising molecules for the development of new drugs with antibiotic and anti-inflammatory actions in the future.

Besides the medicinal potential of plant AMPs, they can be equally useful to improve agricultural plant yields. These small peptides present a wide-range of inhibitory activities on phyto-pathogenic microorganisms (mainly fungi), and have already been used to enhance crop resistance to pathogen attack through either genetic breeding or transgenic manipulation [17]. Additionally, plant AMPs may be helpful in increase abiotic stress tolerance, as has been demonstrated in transgenic plants using different peptide classes.

Thus, screening and isolation of plant peptides associated with defense against pathogens aim to provide new alternatives to reduce yield losses in agriculture, as well as to identify novel potentially useful antimicrobial molecules in medicine and industry. In spite of the large number of research results available regarding plant AMPs, there is however little information on such peptides derived from Brazilian or South American wild plant species [97]. Considering plant organs, AMPs have been extensively studied in roots, leaves and seeds, but they are also produced in flowers, which may harbor additional molecules with inhibitory action on pathogens. To date, defensins (γ-thionins), lipid transfer proteins, hevein-like and snakin-like peptides have been identified in flower tissues and their functions were associated not only with antifungal and antibacterial activities, but have also been considered as protease inhibitors with insecticidal, and allergenic properties, involved in pollen compatibility and floral reproductive differentiation processes [98]. Flower myrosinase-binding proteins, or MBPs – another plant cationic AMP class with two S-S bonds, are significantly more expressed in immature floral tissues, but detailed structural data is still missing and the mechanism of action is not fully described (probable ionophore-like function on microbial cell membranes [20].

Acknowledgements

This study has been partly supported by the R3A2 grant from FVG (2008) region (Italy). The authors also thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) for the concession of fellowships.

References Cited

[1] Zipfel, C.; Felix, G. Plants and animals: a different taste for microbes? *Curr. Opin. Plant. Biol.*, **2005**, *8*, 353-360.

[2] Dangl, J.L.; Jones, J.D. Plant pathogens and integrated defence responses to infection. *Nature*, **2001**, *411*, 826-33.

[3] Barbosa-da-Silva, A.; Wanderley-Nogueira A.C., Soares-Cavalcanti; N.M.; Benko-Iseppon, A.M. In silico survey of resistance (R) genes in *Eucalyptus* transcriptome. *Genet. Mol. Biol*., **2005**, *28*, 562-574.

[4] Wanderley-Nogueira, A.C.; Mota, N.; Lima-Morais, D.; Silva, L.C.B.; Silva A.B.; Benko-Iseppon, A.M. Abundance and Diversity of Resistance (R) Genes in the Sugarcane Transcriptome. *Genet. Mol. Res.*, **2007**, *6*, 866-889.

[5] Bonas, U.; Anckerveken, G.V. Gene-for-gene interactions: bacterial avirulence proteins specify plant disease resistance. *Curr. Opin. Plant Biol*., **1999**, *2*, 94-98.

[6] Staskawicz, B. First insights into the genes that control plant-bacterial interactions. *Mol. Plant Pathol*., **2009**, *10*, 719-720.

[7] Chester, K.S. The Problem of Acquired Physiological Immunity in Plants. *Quart. Rev. Phytopathol*., **1933,** *42*, 185-209.

[8] Slusarenko, A.J.; Fraser, R.S.S.; Van Loon, L.C. *Mechanisms of Resistance to Plant Diseases*. Kluwer, Dordrecht, **2000**.

[9] Salvaudon, L.; Heraldet V.; Shykoff, J.A. Parasite-host fitness trade-offs change with parasite identity: genotype-specific interactions in a plant-pathogen system. *Int. J. Org. Evol*., **2005**, *59*, 2518-2524.

[10] van Loon, L.C.; Rep, M.; Pieterse, C.M.J. Significance of Inducible Defense-related Proteins in Infected *Plants Annu. Rev. Phytopathol*., **2006**, *44*, 135-162.

[11] Tao, Y.; Xie, Z.; Chen, W.; Glazebrook, J.; Chang, H-S.; Han, B.; Zhua, T.; Zou, G.; Katagiri, F. Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *Plant Cell*, **2003**, *15*, 317-330.

[12] Sels, J.; Mathys, J.; De Coninck, B.M.; Cammue, B.P.; De Bolle, M.F. Plant pathogenesis-related (PR) proteins: A focus on PR peptides. *Plant Physiol. Biochem.,* **2008**, *46*, 941-50.

[13] Tossi, A.; Sandri, L. Molecular diversity in gene-encoded, cationic antimicrobial polypeptides. Curr. Pharm. Design, 2002, 8, 743-761.

[14] Klotman, M.E.; Chang, T.L. Defensins in innate antiviral immunity. *Nat. Rev. Immunol.*, **2006**, *6*, 447-456.

[15] Ganz, T. Defensins: antimicrobial peptides of vertebrates. *C.R. Biol.*, **2004**, *327*, 539-549

[16] Oren, Z.; Shai, Y. Mode of action of linear amphipathic alpha-helical antimicrobial peptides. *Biopolymers*, **1998**, *47*, 451-463.

[17] Terras, F. R.; Eggermont, K.; Kovaleva, V.; Raikhel, N.V.; Osborn, R.W.; Kester, A.; Rees, S.B.; Torrekens, S.; Van Leuven, F.; Vanderleyden, J.; Cammue, B.P.A.; Broekaert, W.F. Small cysteine-rich antifungal proteins from radish: their role in host defense. *Plant Cell*, **1995,** *7*, 573-88.

[18] Lay, F.T.; Anderson, M.A. Defensins – components of the innate immune system in plants. *Curr. Prot. Pept. Sci.*, **2005,** *6*, 85-101.

[19] Marcus, J.P.; Goulter, K.C.; Green, J.L.; Harrison, S.J.; Manners, J.M. Purification, characterization and cDNA cloning of an antimicrobial peptide from *Macadamia integrifolia*. *Eur. J. Biochem*., **1997**, *244*, 743-749.

[20] Tailor, R.H.; Acland, D.P.; Attenborough, S.; Cammue, B.P.; Evans, I.J.; Osborn, R.W.; Ray, J.A.; Rees, S.B.; Broekaert, W.F. A novel family of small cysteine-rich antimicrobial peptides from seed of *Impatiens balsamina* is derived from a single precursor protein. *J. Biol. Chem*.**,** **1997**, *272*, 24480-24487.

[21] Patel, S.U.; Osborn, R.; Rees, S.; Thornton, J.M. Structural studies of *Impatiens balsamina* antimicrobial protein (Ib-AMP1). *Biochemistry*, **1998**, *37*, 983-990.

[22] McManus, A.M.; Nielsen, K.J.; Marcus, J.P.; Harrison, S.J.; Green, J.L.; Manners, J.M.; Craik, D.J. MiAMP1, a novel protein from *Macadamia integrifolia* adopts a Greek key beta-barrel fold unique amongst plant antimicrobial proteins. *J. Mol. Biol*., **1999**, *293*, 629-638.

[23] Terras, F.R., Goderis, I.J., Van Leuven, F., Vanderleyden, J., Cammue, B.P.; Broekaert, W.F. *In vitro* antifungal activity of a radish (*Raphanus sativus* L.) seed protein homologous to nonspecific lipid transfer proteins. *Plant Physiol*., **1992**, *100*, 1055-1058.

[24] Ohtani, S.; Okada, T.; Yoshizumi, H.; Kagamiyama H. Complete primary structures of two subunits of purothionin A, a lethal protein for brewer's yeast from wheat flour. *J. Biochem*., **1977,** *82*, 753-767.

[25] Schrader-Fischer, G.; Apel, K. Organ-specific expression of highly divergent thionin variants that are distinct from the seed-specific crambin in the crucifer *Crambe abyssinica*. *Mol. Gen. Genet*., **1994,** *245*, 380-389.

[26] Cammue, B.P.; Thevissen, K.; Hendriks, M.; Eggermont, K.; Goderis, I.J.; Proost, P.; Van Damme, J.; Osborn, R.W.; Guerbette, F.; Kader, J.C.; Broekaert, W.F. A potent antimicrobial

protein from onion seeds showing sequence homology to plant lipid transfer proteins. *Plant Physiol*., **1995,** *109*, 445-455.

[27] Broekaert, W.F.; Mariën, W.; Terras, F.R.; De Bolle, M.F.; Proost, P.; Van Damme, J.; Dillen, L.; Claeys, M.; Rees, S.B.; Vanderleyden, J.; Cammue, B.P.A. Antimicrobial peptides from *Amaranthus caudatus* seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins. *Biochemistry*, **1992,** *31*, 4308-4314.

[28] Cammue, B.P.; De Bolle, M.F.; Terras, F.R.; Proost, P.; Van Damme, J.; Rees, S.B.; Vanderleyden, J.; Broekaert, W.F. Isolation and characterization of a novel class of plant antimicrobial peptides form *Mirabilis jalapa* L. seeds. *J. Biol. Chem*., **1992**, *267*, 2228-2233.

[29] Marcus, J.P.; Goulter, K.C.; Green, J.L.; Harrison, S.J.; Manners, J.M. Purification, characterization and cDNA cloning of an antimicrobial peptide from *Macadamia integrifolia*. *Eur. J. Biochem*., **1997**, *244*, 743-749.

[30] McManus, A.M.; Nielsen, K.J.; Marcus, J.P.; Harrison, S.J.; Green, J.L.; Manners, J.M.; Craik, D.J. MiAMP1, a novel protein from *Macadamia integrifolia* adopts a Greek key beta-barrel fold unique amongst plant antimicrobial proteins. *J. Mol. Biol*., **1999**, *293*, 629-638.

[31] Tailor, R.H.; Acland, D.P.; Attenborough, S.; Cammue, B.P.; Evans, I.J.; Osborn, R.W.; Ray, J.A.; Rees, S.B.; Broekaert, W.F. A novel family of small cysteine-rich antimicrobial peptides from seed of *Impatiens balsamina* is derived from a single precursor protein. *J. Biol. Chem***.,** **1997**, *272*, 24480-24487.

[32] Patel, S.U.; Osborn, R.; Rees, S.; Thornton, J.M. Structural studies of *Impatiens balsamina* antimicrobial protein (Ib-AMP1). *Biochemistry*, **1998**, *37*, 983-990.

[33] Duvick, J.P.; Rood, T.; Rao, A.G.; Marshak, D.R. Purification and characterization of a novel antimicrobial peptide from maize (*Zea mays* L.) kernels. *J. Biol. Chem*., **1992**, *267*, 18814-18820.

[34] Blochet, J.E.; Chevalier, C.; Forest, E.; Pebay-Peyroula, E.; Gautier, M.F.; Joudrier, P.; Pézolet, M.; Marion, D. Complete amino acid sequence of puroindoline, a new basic and cystine-rich protein with a unique tryptophan-rich domain, isolated from wheat endosperm by Triton X-114 phase partitioning. *FEBS Lett*., **1993,** *329*, 336-340.

[35] Gautier, M.F.; Aleman, M.E.; Guirao, A.; Marion, D.; Joudrier, P. *Triticum aestivum* puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant. Mol. Biol*., **1994,** *25*, 43-57.

[36] Segura, A.; Moreno, M.; Madueño, F.; Molina, A.; García-Olmedo, F. Snakin-1, a peptide from potato that is active against plant pathogens. *Mol. Plant-Microbe Interact*., **1999**, *12*, 16-23.

[37] Berrocal-Lobo, M.; Segura, A.; Moreno, M.; López, G.; García-Olmedo, F.; Molina, A. Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiol*., **2002,** *128*, 951-961.

[38] Jennings, C.; West, J.; Waine, C.; Craik, D.; Anderson, M. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. USA*, **2001,** *98*, 10614-10619.

[39] Broekaert, W.; Van Parijs, J.; Leyns, F.; Joos, H.; Peumans, W. A chitin-binding lectin from stinging nettle rhizomes with antifungal properties. *Science*, **1989**, *245*, 1100-1102.

[40] Terras, F.R.; Torrekens, S.; Van Leuven, F.; Osborn, R.W.; Vanderleyden, J.; Cammue, B.P.; Broekaert, W.F. A new family of basic cysteine-rich plant antifungal proteins from *Brassicaceae species*. *FEBS Lett*., **1993**, *316*, 233-240.

[41] Meyer, B., Houlne, G., Pozueta-Romero, J., Schantz, M.L. and Schantz, R. Fruit-specific expression of a defensin-type gene family in bell pepper. Upregulation during ripening and upon wounding. *Plant. Physiol.*, **1996**, *112*, 615-622.

[42] Chen, K.C.; Lin, C.Y.; Kuan, C.C.; Sung, H.Y.; Chen, C.S. A novel defensin encoded by a mung bean cDNA exhibits insecticidal activity against bruchid. *J. Agric. Food Chem.*, **2003**, *50*, 7258-7263.

[43] Janssen, B.J.; Schirra, H.J.; Lay, F.T.; Anderson, M.A.; Craik, D.J. Structure of *Petunia hybrida* defensin 1, a novel plant defensin with five disulfide bonds. *Biochem.*, **2003**, *42*, 8214-8222.

[44] Lay, F.T.; Brugliera, F.; Anderson, M.A. Isolation and properties of floral defensins from ornamental tobacco and petunia. *Plant Physiol.*, **2003**, *131*, 1283-1293.

[45] Wong J.H.; Zhang X.Q.; Wang H.X.; Ng T.B. A mitogenic defensin from white cloud beans (*Phaseolus vulgaris*). *Peptides*, **2006**. *27*, 2075-81.

[46] Padovan, L.; Segat, L.; Tossi, A.; Antcheva, N.; Benko-Iseppon, A.M.; Kido, E.A.; Brandao, L.; Calsa Jr., T.; Crovella, S. A plant-defensin from sugarcane (*Saccharum* spp.). *Protein Pept. Lett.*, **2009**, *16*, 430-436.

[47] Bloch, C., Jr.; Patel, S.U.; Baud, F.; Zvelebil, M.J.; Carr, M.D.; Sadler, P.J.; Thornton, J.M. 1H NMR structure of an antifungal gamma-thionin protein SIalpha1: similarity to scorpion toxins. *Proteins*, **1998**, *32*, 334-49.

[48] Fant, F.; Vranken, W.; Broekaert, W.; Borremans, F. Determination of the three-dimensional solution structure of *Raphanus sativus* antifungal protein 1 by 1H NMR. *J. Mol. Biol.*, **1998**, *279*, 257-70.

[49] Almeida, M.S.; Cabral, K.M.; Kurtenbach, E.; Almeida, F.C.; Valente, A.P. Solution structure of Pisum sativum defensin 1 by high resolution NMR: plant defensins, identical backbone with different mechanisms of action. *J. Mol. Biol.*, **2002**, *315*, 749-757.

[50] Broekaert, W.F.; Terras, F.R.; Cammue, B.P.; Osborn, R.W. Plant defensins: novel antimicrobial peptides as components of the host defense system. *Plant Physiol.*, **1995**, *108*, 1353-1358.

[51] Osborn, R.W.; De Samblanx, G.W.; Thevissen, K.; Goderis, I.; Torrekens, S.; Van Leuven, F.; Attenborough, S.; Rees, S.B.; Broekaert, W.F. Isolation and characterization of plant defensins from seeds of Asteraceae, Fabaceae, Hippocastanaceae and Saxifragaceae. *FEBS Lett.*, **1995**, *368*, 257-262.

[52] Broekaert, W.F.C., Cammue, B.P.A.; De Bolle, M.F.C.; Thevissen, K.; De Samblanx, G.W.; Osbron, R.W. Antimicrobial peptides in plants. *Crit. Rev. Plant. Sci.*, **1997**, *16*, 297-323.

[53] Segura, A., Moreno, M., Molina, A. and Garcia-Olmedo, F. Novel defensin subfamily from spinach (*Spinacia oleracea*). *FEBS Lett.*, **1998**, *435*, 159-162.

[54] Chiang, C.C.; Hadwiger, L.A. The *Fusarium solani*-induced expression of a pea gene family encoding high cysteine content proteins. *Mol. Plant. Microbe Interact.*, **1991**, *4*, 324-331.

[55] Moreno, M.; Segura, A.; Garcia-Olmedo, F. Pseudothionin-St1, a potato peptide active against potato pathogens. *Eur. J. Biochem.*, **1994**, *223*, 135-139.

[56] Sharma, P.; Lonneborg, A. Isolation and characterization of a cDNA encoding a plant defensin-like protein from roots of Norway spruce. *Plant. Mol. Biol.*, **1996**, *31*, 707-712.

[57] Terras, F.R., Schoofs, H.M., De Bolle, M.F., Van Leuven, F., Rees, S.B., Vanderleyden, J., Cammue, B.P.; Broekaert, W.F. Analysis of two novel classes of plant antifungal proteins from radish (*Raphanus sativus* L.) seeds. *J. Biol. Chem.*, **1992**, *267*, 15301-15309.

[58] Selitrennikkoff, C.P. Antifungal proteins. *Appl. Environ. Microbiol*., **2001**, *67*, 2883-2894.

[59] Castro, M.S.; Fontes, W. Plant defense and antimicrobial peptides. *Prot. Pept. Lett*., **2005**, *12*, 13-18.

[60] Diamond, G.; Zasloff, M.; Eck, H.; Brasseur, M.; Maloy, W.L.; Bevins, C.L. Tracheal antimicrobial peptide, a cysteine-rich peptide from mammalian tracheal mucosa: peptide isolation and cloning of a cDNA. *Proc. Natl. Acad. Sci. USA*, **1991**, *88*, 3952-6.

[61] Epple, P.; Apel, K.; Bohlmann, H. Overexpression of an endogenous thionin enhances resistance of *Arabidopsis* against *Fusarium oxysporum*. *Plant Cell*, **1997**, *9*, 509-520.

[62] Iwai, T.; Kaku, H.; Honkura, R.; Nakamura, S.; Ochiai, H.; Sasaki, T.; Ohashi, Y. Enhanced resistance to seed-transmitted bacterial diseases in transgenic rice plants overproducing an oat cell-wall-bound thionin. *Mol. Plant. Microb. Interact*., **2002**, *15*, 515-521.

[63] Choi, Y.; Choi, Y.D.; Lee, J.S. Antimicrobial activity of gamma-thionin-like soybean SE60 in *E. coli* and tobacco plants. *Biochem. Biophys. Res. Commun*., **2008**, *375*, 230-234.

[64] Kader, J-C. Proteins and the intracellular exchange of lipids: stimulation of phospholipid exchange between mitochondria and microssomal fractions by proteins isolated from potato tuber. *Biochim. Biophys. Acta*, **1975**, *380*, 31-44.

[65] Kader, J-C. Lipid-transfer proteins in plants. *Annu. Rev. Plant Physiol. Plant. Mol. Biol*., **1996**, *47*, 627-654.

[66] Carvalho, A.O.; Gomes, V.M. Role of plant lipid transfer proteins in plant cell physiology — A concise review. *Peptides,* **2007**, *28*, 1144-1153.

[67] Douliez, J-P.; Pato, C.; Rabesona, H.; Molle, D.; Marion, D. Disulfide bond assignment, lipid transfer activity and secondary structure of a 7-kDa plant lipid transfer protein, LTP2. *Eur. J. Biochem*., **2001**, *268*, 1400-1403.

[68] Liu, Y-J.; Samuel, D.; Lin, C-H.; Lyu, P-C. Purification and characterization of a novel 7-kDa non-specific lipid transfer protein-2 from rice (*Oryza sativa*). *Biochem. Biophys. Res. Commun.*, **2002**, *294*, 535-540.

[69] Castro, M.S.; Gerhardt, I.R.; Orru, S.; Pucci, P.; Bloch, C.J.R. Purification and characterization of a small (7.3 kDa) putative lipid transfer protein from maize seeds. *J. Chromat. B: Analyt. Techn. Biomed. Life Sci.*, **2003**, *794*, 109-114.

[70] Tsuboi, S.; Osafune, T.; Tsugeki, R.; Nishimura, M.; Yamada, M. Nonspecific lipid transfer protein in castor bean cotyledons cells: subcellular localization and a possible role in lipid metabolism. *J. Biochem*., **1992**, *111*, 500-508.

[71] Thoma, S.L.; Kaneko, Y.; Somerville, C. An *Arabidopsis* lipid transfer protein is a cell wall protein. *Plant J*., **1993**, *3*, 427-37.

[72] Pyee, J.; Yu, H.; Kolattukudy, P.E. Identification of a lipid transfer protein as the major protein in the surface wax of broccoli (*Brassica oleracea*) leaves. *Arch. Biochem. Biophys*., **1994**, *311*, 460-468.

[73] Carvalho, A.O.; Teodoro, C.E.S.; Da Cunha, M.; Okorokova-Façanha, A.L.; Okorokov, L.A.; Fernandes, K.V.S.; Gomes, V.M. Intracellular localization of a lipid transfer protein in *Vigna unguiculata* seeds. *Physiol. Plant*, **2004**, *122*, 328-336.

[74] Farrokhi, N.; Whitelegge, J.P.; Brusslan, J.A. Plant peptides and peptidomics. *Plant Biotechn. J.,* **2008**, *6*, 105-134.

[75] Molina, A.; Segura, A.; García-Olmedo, F. Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibitors of bacterial and fungal plant pathogens. *FEBS Lett*., **1993**, *316*, 119-122.

[76] Segura, A.; Moreno, M.; Garcia-Olmedo, F. Purification and antipathogenic activity of lipid transfer proteins (LTPs) from the leaves of *Arabidopsis* and spinach. *FEBS Lett*. **1993**, *332*, 243-246.

[77] Dubreil, L.; Gaborit, T.; Bouchet, B.; Gallant, D.J.; Broekaert, W.F.; Quillien, L.; Marion, D. Spatial and temporal distribution of the major isoforms of puroindolines (puroindoline-a and puroindoline-b) and non specific lipid transfer protein (ns-LTP1e1) of *Triticum aestivum* seeds. Relationships with their *in vitro* antifungal properties. *Plant Sci*., **1998**, *138*, 121-135.

[78] Gonorazky, A.G.; Regente, M.C.; de la Canal, L. Stress induction and antimicrobial properties of a lipid transfer protein in germinating sunflower seeds. *J. Plant Physiol*., **2005**, *162*, 618-624.

[79] Diz, M.S.; Carvalho, A.O.; Rodrigues, R.; Neves-Ferreira, A.G.; Da Cunha, M.; Alves, E.W.; Okorokova-Facanha, A.L.; Oliveira, M.A.; Perales, J.; Machado, O.L.; Gomes, V.M. Antimicrobial peptides from chili pepper seeds cause yeast plasma membrane permeabilization and inhibit the acidification of the medium by yeast cells. *Biochim. Biophys. Acta*, **2006**, *1760*, 1323-1332.

[80] Lin, P.; Xia, L.; Ng, T.B. First isolation of an antifungal lipid transfer peptide from seeds of a *Brassica* species. *Peptides*, **2007**, *28*, 1514-1519.

[81] Craik, D.J.; Daly, N.L.; Bond, T.; Waine, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.*, **1999**, *294*, 1327-1336.

[82] Simonsen, S.M.; Sando, L.; Ireland, D.C.; Colgrave, M.L.; Bharathi, R.; Goransson, U.; Craik, D.J. A continent of plant defense peptide diversity: Cyclotides in Australian *Hybanthus* (Violaceae). *Plant Cell*, **2005**, *17*, 3176-3189.

[83] Gruber, C.W.; Elliott, A.G.; Ireland, D.C.; Delprete, P.G.; Dessein, S.; Göransson, U.; Trabi, M.; Wang, C.K.; Kinghorn, A.B.; Robbrecht, E.; Craik, D.J. Distribution and Evolution of Circular Miniproteins in Flowering Plants. *Plant Cell*, **2008**, *20*, 2471-2483.

[84] Craik, D.J. Circling the enemy: cyclic proteins in plant defence. *Trends Plant Sci.*, **2009**, *14*, 328-335.

[85] Shi, L.; Gast, R.T.; Gopalraj, M.; Olszewski, N.E. Characterization of a shoot-specific, GA3- and ABA-regulated gene from tomato. *Plant J.*, **1992**, *2*, 153-159.

[86] Herzog, M.; Dorne, A.-M.; Grellet, F. GASA, a gibberellins regulated gene family from *Arabidopsis thaliana* related to the tomato GAST1 gene. *Plant Mol. Biol.*, **1995**, *27*, 743-752.

[87] Kotilainen, M.; Helariutta, Y.; Mehto, M.; Pollanen, E.; Albert, V.A.; Elomaa, P.; Teeri, T.H. *GEG* participates in the regulation of cell and organ shape during corolla and carpel development in *Gerbera hybrida*. *Plant Cell*, **1999**, *11*, 1093-1104.

[88] Ben-Nissan, G.; Weiss, D. The petunia homologue of tomato *gast1*: Transcript accumulation coincides with gibberellin-induced corolla cell elongation. *Plant Mol. Biol*., **1996**, *32*, 1067-1074.

[89] Hancock, R.E.; Lehrer, R. Cationic peptides: a new source of antibiotics. *Trends Biotechnol.*, **1998**, *16*, 82-88.

[90] Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature*, **2002**, 415, 389-95.

[91] Brogden, K.A.; Ackermann, M.; McCray, P.B.Jr.; Tack, B.F. Antimicrobial peptides in animals and their role in host defences. *Int. J. Antimicrob. Agents*, **2003**, *22*, 465-478.

[92] van't Hof, W.; Veerman, E.C.; Helmerhorst, E.J.; Amerongen, A.V. Antimicrobial peptides: properties and applicability. *Biol. Chem.*, **2001**, *382*, 597-619.

[93] Shai, Y. Mode of action of membrane active antimicrobial peptides. *Biopolymers*, **2002**, *66*, 236-248.

[94] Koczulla, A.R.; Bals, R. Antimicrobial peptides: current status and therapeutic potential. *Drugs*, **2003**, *63*, 389-406.

[95] Yang, D.; Biragyn, A.; Kwak, L.W.; Oppenheim, J.J. Mammalian defensins in immunity: more than just microbicidal. *Trends Immunol.*, **2002**, *23*, 291-296.

[96] Yang, D.; Biragyn, A.; Hoover, D.M.; Lubkowski, J.; Oppenheim, J.J. Multiple roles of antimicrobial defensins, cathelicidins, and eosinophil-derived neurotoxin in host defense. *Annu. Rev. Immunol.*, **2004**, *22*, 181-215.

[97] Pelegrini, P.B.; Farias, L.R.; Saúde, A.C.M.; Costa, F.T.; Bloch-Jr, C.; Silva, L.P.; Oliveira, A.S.; Gomes, C.E.M.; Sales, M.P.; Franco, O.L. A novel antimicrobial peptide from *Crotalaria pallida* seeds with activity against human and phytopathogens. *Curr. Microbiol.*, **2009**, *59*, 400-404.

[98] Tavares, L.S.; Santos, M.O.; Viccini, L.F.; Moreira, J.S.; Miller, R.N.G.; Franco, O.L. Biotechnological potential of antimicrobial peptides from flowers. *Peptides*, **2008**, *29*, 1842-1851.

Anexos

Anexo A

Tabela 1: Levantamento das defensinas vegetais em relação à classe, família, espécie e tecido de obtenção bem como informação sobre o peptídeo e as atividades biológicas conhecidas, acompanhado do número de aquisição no GenBank ou do número do cluster sugerido (TC) no TIGR Gene Indices.

Classe	Família	Espécie	Aquisição ou TC	Peptídeo sinal pI	Peptídeo maduro			Tecido	Atividade biológica	Referências
					pI	nR/kDa	nC/pD			
Monocotiledônea	Arecaceae	*Elaeis guineensis*	-	6.52	8.99	47/5.31	8/4	Inflorescência	VS	Tregear *et al.*, 2002
	Liliaceae	*Tulipa gesneriana*	-	nd	9.08	46/4.99	8/4	Bulbo	nd	Fujimura *et al.*, 2004
	Orchidaceae	*Gymnadenia conopsea*	EF051327	8.35	9.08	68/7.52	8/4	nd	nd	np
	Poaceae	*Echinochloa crusgalli*	-	nd	8.74	46/5.05	8/4	Semente	F	Odintsova *et al.*, 2008
		Hordeum vulgare	-	nd	9.77	47/5.25	8/4	Endosperma	IST	Mendez *et al.*, 1990
			-	nd	8.75	48/5.52	8/4	Endosperma	IST	np
		Sorghum bicolor	-	nd	8.51	47/5.38	8/4	Semente	A	Bloch e Richardson, 1991
			-	nd	8.75	48/5.31	8/4	Semente	A	np
			-	nd	8.97	48/5.39	8/4	Semente	A	np
		Triticum aestivum	-	8.50	8.51	49/5.53	8/4	Coroa	F, B$^-$	Koike *et al.*, 2002
Dicotiledônea	Amaranthaceae	*Spinacea oleracea*	-	nd	9.35	52/5.80	8/4	Folha	F, B$^+$, B$^+$	Segura *et al.*, 1998
		Beta vulgaris	P81493	nd	8.21	46/5.08	8/4	Folha, Flor	F	kragh *et al.*, 1995
			P82010	nd	8.51	46/5.18	8/4	Folha, Flor	F	np
	Asteraceae	*Artemísia vulgaris*	AAO24900	4.53	8.17	108/10.80	8/4	Pólen	nd	Himly *et al.*, 2003
		Dhalia merckii	-	nd	7.80	50/5.53	8/4	Semente	F	Olli e Kirti PB, 2006
		Helianthus annuus	-	5.90	9.14	47/5.35	8/4	Flor	F	Urdangarín *et al.*, 2000
			-	8.35	7.09	50/5.59	8/4	Flor	IPP	Zélicourt *et al.*, 2007
			AAM27914	8.35	8.19	80/8.93	8/4	Folha	F	HU *et al.*, 2003
			P22357	3.80	5.72	146/13.86	8/4	Antera	nd	Domon *et al.*, 1990
	Brassicaceae	*Arabidopsis halleri*	-	4.53	8.92	51/5.70	8/4	Gérmen	TMP	Mirouze *et al.*, 2006
			-	4.53	8.51	51/5.53	8/4	Gérmen	nd	np
			-	4.53	8.49	51/5.57	8/4	Gérmen	nd	np
			-	4.53	8.91	51/5.71	8/4	Gérmen	nd	np
		Arabidopsis lyrata	Q9AVE3	8.00	8.71	63/7.34	8/4	Antera	AI	Kusaba *et al.*, 2005
		Arabidopsis lyrata	Q9AVE2	5.59	8.69	55/6.47	8/4	Antera, Micrósporos	AI	np
		Arabidopsis thaliana	P30224	nd	nd	nd	nd	Semente	F	Terras *et al.*, 1993
		Brassica napus	AAB25088	nd	nd	nd	nd	Semente	F	np
			AAB25089	nd	nd	nd	nd	Semente	F	np
			P80301	nd	8.19	60/6.60	8/4	Semente	IP	Ceciliane *et al.*, 1994
			Q39313	4.53	8.51	50/5.49	8/4	nd	nd	np
		Brassica rapa	AAB25086	nd	nd	nd	nd	Semente	F	Terras *et al.*, 1993
			AAB25087	nd	nd	nd	nd	Semente	F, B$^+$	np
			AAN23105	8.35	8.51	51/5.73	8/4	Folha	nd	Park *et al.*, 2005
			AAQ92328	8.35	8.81	51/5.78	7/3	nd	nd	np
			BAH03381	4.53	8.51	50/5.49	8/4	Folha	nd	Chen *et al.*, 2008
		Brassica juncea	ABB59548	4.53	8.72	51/5.69	8/4	nd	nd	np
		Brassica oleracea	CAA06464	8.35	9.26	55/6.43	8/4	Antera, Pólen	AI	Doughty *et al.*, 1998
			CAA65766	8.35	9.05	58/6.47	8/4	Antera, Pólen	AI	Stanchev *et al.*, 1996
			CAA65767	8.35	9.05	56/6.20	8/4	Antera, Pólen	AI	np
			ABS11037	4.53	8.51	50/5.49	8/4	nd	nd	Np
			CAC37558	4.53	8.91	51/5.73	8/4	Plântula	nd	Np
			CAC19879	nd	8.71	64/6.92	8/4	Pólen	nd	Vanoosthuyse *et al.*, 2001
		Eutrema wasabia	BAB19054	4.53	8.50	51/5.71	8/4	Folha	F	Saito *et al.*, 2001
		Lepidium meyenii	AAV85992	8.35	8.73	51/5.77	8/4	Folha	F	Solis *et al.*, 2007
		Orychophragmus violaceus	ACS93640	4.53	8.91	51/5.69	8/4	nd	nd	Np
		Raphanus sativus	O24331	4.53	8.51	51/5.74	8/4	Folha	F	Terras *et al.*, 1992; Terras *et al.*, 1995
			O24332	4.53	8.51	50/5.49	8/4	Folha	nd	Np
			P30230	4.25	9.08	51/5.73	8/4	Semente	nd	Np
			P69241	4.53	8.72	51/5.69	8/4	Semente	nd	Np
		Sinapis Alba	P30231	nd	8.72	51/5.69	8/4	Semente	F	Neumann *et al.*, 1996; Terras *et al.*, 1993
			P30232	nd	8.92	51/5.72	8/4	Semente	F	Np
			Q10989	nd	9.27	52/5.85	8/4	Semente	F	Np
			AAY15221	4.53	8.51	51/5.61	8/4	nd	nd	Np

Tabela 1 (continuada)

Classe	Família	Espécie	Aquisição ou TC	Peptídeo sinal pI	Peptídeo maduro			Tecido	Atividade biológica	Referências
					pI	nR/kDa	nC/pD			
	Convolvulaceae	*Ipomoea batatas*	AAS65426	5.75	8.94	55/6.04	8/4	Túbero	DHAR, F, B⁺, IP	Huang *et al.*, 2008
		Ipomoea trifida	AAS79586	9.52	8.40	89/10.13	9/4	Pólen	AI	Rahman *et al.*, 2007
			AAS79587	5.90	8.46	50/5.50	8/4	Pólen	Nd	np
			BAF36317	nd	8.68	57/6.52	8/4	Pólen	Nd	np
			BAF36359	5.38	8.67	50/5.55	8/4	Pólen	Nd	np
			BAF52542	4.53	9.03	52/5.96	8/4	Pólen	Nd	np
			BAF52543	5.82	8.69	57/6.44	8/4	Pólen	Nd	np
			BAF52544	5.38	9.00	50/5.55	8/4	Pólen	Nd	np
			BAF52547	5.90	8.67	50/5.55	8/4	Pólen	Nd	np
	Curcubitaceae	*Trichosanthes kirilowii*	ABF74600	4.49	9.58	47/5.61	8/4	Folha	F	Da-Hui *et al.*, 2007
	Fabaceae	*Arachis diogoi*	AAP92330	4.78	8.53	47/5.44	8/4	nd	nd	np
			AAO72633	4.78	8.51	45/5.18	8/4	nd	nd	np
		Arachis hypogaea	ABC46711	4.53	6.89	47/5.40	8/4	nd	Nd	Senthil-Kumar *et al.*, 2007
		Cajanus cajan	AAP49847	4.78	8.51	45/5.18	8/4	nd	Nd	np
		Cassia fistula	-	nd	8.69	50/5.47	8/4	Semente	IP	Wijaya *et al.*, 2000
		Cicer arietinum	AAO38756	4.78	8.51	45/5.18	8/4	nd	Nd	np
			ABC59238	4.68	7.75	47/5.41	8/4	Semente	Nd	np
			ABC02867	4.41	7.75	47/5.41	8/4	Semente	Nd	np
		Clitoria ternatea	AAB34971	nd	8.51	49/5.61	8/4	Semente	F	Osborn *et al.*, 1995
		Glycine Max	AAC97524	8.00	8.97	58/6.56	8/4	Folha	Nd	np
			ACU20671	5.82	9.69	47/5.45	8/4	nd	Nd	np
			ACU15056	4.78	8.20	47/5.50	8/4	nd	Nd	np
			CAA79164	5.90	9.19	47/5.37	8/4	Semente	Nd	np
			ACU14316	8.34	9.37	47/5.35	8/4	nd	Nd	np
			Q07502	5.90	9.19	46/5.26	8/4	Semente	Nd	Choi *et al.*, 1993
		Lens culinaris	ABP04037	4.78	8.20	47/5.45	8/4	Semente	F	Finkina *et al.*, 2008
		Medicago sativa	AAG40321	4.78	8.51	45/5.19	8/4	Semente	F, BCI	Frokjaer e Otzen, 2005; Spelbrink, *et al.* 2004
			AAV85437	4.78	8.50	45/5.16	8/4	Semente, Folha, Flor	Nd	Hanks *et al.*, 2005
	Fabaceae	*Medicago sativa*	AAV85436	4.78	8.50	45/5.16	8/4	Semente, Folha, Flor	Nd	
			AAV85435	4.78	8.51	45/5.19	8/4	Semente, Folha, Flor	Nd	
			AAV85434	4.78	8.51	45/5.19	8/4	Semente, Folha, Flor	Nd	
			AAV85433	4.78	8.51	45/5.19	8/4	Semente, Folha, Flor	Nd	
			AAV85432	4.78	8.51	45/5.19	8/4	Semente, Folha, Flor	Nd	
			AAT66095	4.78	8.22	45/5.18	8/4	nd	Nd	np
			AAT66096	4.78	7.72	47/5.42	8/4	nd	Nd	np
			AAQ91290	4.78	8.21	45/5.15	8/4	Semente	F	Spelbrink *et al.*, 2004
			AL385826	nd	8.84	57/6.33	8/4	Raiz	Nd	Hanks *et al.*, 2005
			AJ499425	5.75	9.10	49/5.46	8/4	Raiz	Nd	
			BG452001	nd	8.75	52/5.82	8/4	Plântula	Nd	
			TC85327	5.75	8.97	47/5.32	8/4	Semente, Raiz, Nódulo, Folha, Flor	Nd	
			CA990416	8.50	8.76	57/6.14	8/4	Semente	Nd	
			TC89352	5.94	9.54	47/5.32	8/4	Vagem com semente	Nd	
			AJ498901	4.78	8.20	45/5.12	8/4	Vagem com semente	Nd	
			TC82368	4.78	6.42	46/5.20	8/4	Semente, Folha	Nd	
			TC87634	7.98	4.55	55/45.90	8/4	Nódulo	Nd	
			TC76542	6.70	8.21	55/5.73	8/4	Gérmen, Raiz, Plântula, Folha, Vagem com semente	Nd	
			TC90828	7.76	5.45	55/5.78	8/4	Raiz	Nd	
			AJ499268	5.40	8.20	53/5.84	8/4	Raiz	Nd	
			TC77480	5.75	8.21	50/5.54	8/4	Raiz	Nd	
			TC90864	5.75	8.73	51/5.62	8/4	Raiz	Nd	
			TC87273	4.53	9.46	107/11.91	16/nd	Raiz, Nódulo	Nd	
			TC84759	4.53	9.20	50/5.63	8/4	Raiz	Nd	
		Pachyrhizus erosus	AAT80338	nd	7.72	47/5.50	8/4	Semente	TMP	Song *et al.*, 2004
		Phaseolus coccineus	P84785	nd	nd	nd	nd	Semente	F, AC, ITR	Ngai e Ng, 2004
			AAG17880					Semente	Nd	Np
		Phaseolus limensis	-	nd	6.86	48/5.50	8/4	Semente	F, ITR	Wang e Ng, 2006
			-	nd	nd	nd	nd	Semente	F, B⁺, AC	Wang *et al.*, 2008
		Phaseolus vulgaris	CAL68581	7.98	8.97	47/5.17	8/4	Folha	Nd	np
			-	nd	nd	nd	nd	Semente	F, M, ITR	Ye e Ng, 2001
			-	nd	8.20	47/5.44	8/4	Semente	F	Gao *et al.*, 2000
			-	nd	nd	nd	nd	Semente	F, AC	Wang e Ng, 2007
			-	nd	7.72	47/5.47	8/4	Semente	F, B⁺, ITR, IST, AC, M	Wong *et al.*, 2006
			-	nd	nd	nd	nd	Semente	F, AC, IST, ITR	Wong e Ng, 2005b

Tabela 1 (Continuada**)**

Classe	Família	Espécie	Aquisição ou TC	Peptídeo sinal pI	Peptídeo maduro			Tecido	Atividade biológica	Referências
					pI	nR/kDa	nC/pD			
		Pisum sativum	P81929	nd	7.73	46/5.20	8/4	Semente	F	Almeida *et al.*, 2000
			P81930	nd	8.52	47/5.40	8/4	Semente	F	
			ACI15746	4.78	7.73	46/5.20	8/4	nd	nd	np
			AAN60445	3.80	6.42	40/4.45	6/3	Vagem, Folha	F	Lay *et al.*, 2002
			Q01784	8.35	5.65	54/6.06	8/4	Vagem	nd	Chiang e Hadwiger, 1991
			AAA79118	8.35	5.65	54/6.06	8/4	Vagem	nd	
			CAA36473	8.35	5.65	54/6.06	8/4	Vagem	nd	
			Q01783	4.78	8.21	45/5.02	8/4	Vagem	nd	
			AAA79117	4.78	8.21	45/5.02	8/4	Vagem	nd	
			CAA36474	4.78	8.21	45/5.02	8/4	Semente	nd	
			AAG43285	5.01	8.22	46/5.09	8/4	nd	nd	np
		Tephrosia villosa	AAX86993	4.78	8.20	47/4.47	8/4	Folha	F, ICR	Vijyan *et al.*, 2008
		Trigonella foenum-graecum	AAP48592	4.78	8.53	47/5.44	8/4	Folha	F	Olli e Kirti, 2006
			AAO72632	4.78	8.53	47/5.44	8/4	nd	nd	np
		Vicia faba	P81456	nd	9.12	47/5.22	8/4	Semente	B^+, B^-,	Zhang e Lewis, 1997
			P81457	nd	9.12	47/5.20	8/4	Semente	B^+, B^-	
	Fabaceae	*Vicia faba*	ACI02057	5.94	9.12	47/5.24	8/4	Semente	nd	Np
			ACI02058	4.78	7.02	46/5.17	8/4	Semente	nd	Np
			ACI02059	4.78	8.21	46/5.43	8/4	Semente	nd	Np
			ACI02060	4.78	8.21	46/5.43	8/4	Semente	nd	Np
			ACI02061	4.78	7.02	46/5.17	8/4	Semente	nd	Np
			ACI02062	4.78	7.73	46/5.17	8/4	Semente	nd	Np
		Vigna angularis	-	nd	nd	nd	nd	Semente	F, M, ITR	Ye e Ng, 2001
			-	4.87	9.20	46/5.20	8/4	Semente	F, AIS, IST	Chen *et al.*, 2005
		Vigna nakashimae	AAX98673	4.87	8.20	47/5.35	8/4	nd	nd	Np
			AAX98672	4.87	9.06	46/5.12	8/4	nd	nd	Np
		Vigna radiata	AAG45227	4.87	9.06	46/5.12	8/4	Semente	F, AIS, IST	Chen *et al.*, 2004; Chen *et al.*, 2002a; Chen *et al.*, 2002b
			ACN22750	4.87	9.06	46/5.12	8/4	nd	nd	Np
			AAR08912	4.87	9.06	46/5.12	8/4	Semente	AIS	Chen *et al.*, 2004
			CAA34760	4.78	7.72	47/5.44	8/4	Cotilédone	nd	Ishibashi *et al.*, 1990
			BAB82453	4.78	8.51	47/5.50	8/4	Semente	nd	Np
		Vigna sesquipedalis	P84868	nd	nd	nd	nd	Semente	F, B^+, B^-, AC,ITR	Wong e NG, 2005a
		Vigna umbellata	-	nd	nd	nd/5.00	8/4	Semente	F, M, ITR, IST	Ye e NG, 2002
		Vigna unguiculata	ACJ06538	nd	7.72	47/5.41	8/4	Semente	F, A	Carvalho *et al.*, 2001; Carvalho *et al.*, 2006; Pelegrini *et al.*, 2008
			ACN93800	5.75	9.17	55/6.12	8/4	Raiz, Folha	nd	Padovan *et al.*, 2010
			P84920	nd	9.20	46/5.24	8/4	Semente, Folha	B+, B-	Franco *et al.*, 2006
			P83399	nd	8.97	47/5.17	8/4	Cotilédone	IP	Melo *et al.*, 2002
			P18646	4.78	7.72	47/5.44	8/4	Cotilédone		Ishibashi *et al.*, 1990
	Linderniaceae	*Torenia fournieri*	BAH29762	6.49	8.16	47/5.14	8/4	Óvulo	nd	Okuda *et al.*, 2009
	Nelumbonaceae	*Nelumbo nucifera*	ABN46979	5.99	8.75	47/5.11	8/4	nd		Np
	Oleaceae	*Olea europea*	ABS72000	nd	8.93	56/6.18	8/4	nd		Np
	Plantaginaceae	*Plantago major*	CAH58740	5.75	9.12	47/5.21	8/4	Tecido vascular	nd	Np
	Pentadiplandraceae	*Pentadiplnadra brazzeana*	P56552	nd	5.72	54/6.50	8/4	Fruto	Sabor adocicado	Ming e Hellekant, 1994
	Rosaceae	*Prunus pérsica*	AAL85480	5.90	9.17	47/5.23	8/4	Súber	F	Wisniewski *et al.*, 2003
		Pyrus pyrifolia	BAB64929	4.53	9.53	64/6.99	8/4	Pólen		Zhou *et al.*, 2002
			BAB64930	5.90	9.06	64/6.98	8/4	Pólen		
			BAB64931	3.90	9.22	57/6.34	8/4	Pólen		
	Ranunculaceae	*Aquilegia brevistyla*	ABA86588	4.53	8.92	49/5.47	8/4	nd	nd	Whittall *et al.*, 2006
		Aquilegia chrysantha	ABA86587	5.72	8.74	49/5.43	8/4	nd	nd	
		Aquilegia formosa	ABA81896	4.53	8.92	49/5.47	8/4	nd	nd	
			ABA86586	4.53	8.92	49/5.47	8/4	nd	nd	
		Aquilegia olympica	ABA86590	4.53	8.92	49/5.47	8/4	nd	nd	
		Aquilegia pyrenaica	ABA86589	4.53	8.92	49/5.47	8/4	nd	nd	
	Sapindaceae	*Aesculus hippocastanum*	AAB34970	nd	7.73	50/5.86	8/4	Semente	F	Osborn *et al.*, 1995
	Saxifragaceae	*Heuchera sanguinea*	AAB34974	nd	8.49	54/5.95	8/4	Semente	F	Osborn *et al.*, 1995
	Solanaceae	*Capsicum annuum*	AAL35366	6.51	9.37	47/5.38	8/4	Folha	nd	Do *et al.*, 2004
			CAA65045	4.37	8.52	48/5.19	8/4	Fruto	F	Meyer *et al.*, 1996
			Q43413	4.37	8.52	48/5.19	8/4	Fruto	F	Meyer *et al.*, 1996
			CAA65046	5.75	9.37	47/5.38	8/4	Fruto, Flor	F	Meyer *et al.*, 1996
			AAR90845	9.52	8.52	58/6.44	8/4	nd	nd	Np
			AAF16413	8.34	8.41	59/6.70	9/3	Fruto	F	Oh *et al.*, 1999

Tabela 1 (continuada)

Classe	Família	Espécie	Aquisição ou TC	Peptídeo sinal pI	Peptídeo maduro			Tecido	Atividade Biológica	Referências
					pI	nR/kDa	nC/pD			
	Solanaceae	*Capsicum annuum*	ABY66953	8.34	5.22	66/7.49	9/4	nd	nd	Np
			ACB30363	8.50	8.41	59/6.70	9/3	nd		Np
			AAF18936	8.34	8.41	59/6.70	9/3	Folha		Lee *et al.*, 2000
		Capsicum chinensis	ABW99097	nd	8.74	55/5.98	8/4	Folha		Np
			AAD21200	8.50	6.87	82/9.26	8/4	nd	nd	Np
		Capsicum frutescens	AAD21200	8.50	6.87	82/9.26	8/4	Folha	nd	Np
		Nicotiana alata	Q8GTM0	5.75	9.08	47/5.30	8/4	Botão floral	F	Lay *et al.*, 2003a e b
		Nicotiana attenuata	AAS13434	nd	9.41	52/5.93	8/4	Tricoma	nd	Lou Y, Baldwin IT, 2004
			AAS13436	5.75	6.46	81/9.06	8/4	Tricoma	nd	
		Nicotiana benthamiana	ABU62754	nd	4.96	53/5.87	nd/nd	Flor	nd	Np
		Nicotiana excelsior	BAA21113	nd	5.36	79/8.78	8/4	nd	nd	Np
			BAA21114	4.37	6.75	80/8.88	8/4	Folha	nd	Yamada *et al.*, 1997
		Nicotiana megalosiphon	ACR46857	4.14	8.51	45/5.13	8/4	nd	F	Portieles *et al.*, 2010.
		Nicotiana paniculata	O24115	5.75	6.44	81/8.92	8/4	Folha	nd	Yamada *et al.*, 1997
		Nicotiana tabacum	BAA95697	5.75	8.78	55/6.05	8/4	nd	nd	Np
			P32026	5.75	6.75	80/8.93	8/4	Flor	nd	Gu *et al.*, 1994
			ABU40984	5.75	6.75	80/8.90	8/4	Flor	nd	Np
			BAA06149	7.95	9.10	60/7.16	8/4	Cultura celular deficiente em fósforo	nd	Ezaki *et al.*, 1995
		Petunia hybrida	Q8H6Q1	4.37	8.90	47/5.21	10/5	Pétala	F	Lay *et al.*, 2003a
			Q8H6Q0	4.53	8.76	49/5.40	8/4	Pétala	F	
		Petunia integrifolia	Q40901	9.35	8.79	55/6.12	8/4	Pistilo	nd	Karunanandaa *et al.*, 1994
		Solanum chacoense	ACB32234	9.50	8.76	53/5.91	8/4	Flor	nd	np
		Solanum lycopersicum	AB036637	9.50	8.76	53/5.97	8/4	Fruto	Variação morfológica	Xiao *et al.*, 2008.
			AB036635	5.75	8.99	47/5.29	8/4	Fruto	Variação morfológica	
			AB036638	5.38	8.99	47/5.26	8/4	Fruto	Variação morfológica	
			AAA80496	4.53	7.68	80/9.00	8/4	Pistilo	nd	Milligan e Gasser, 1995
			CAB42006	9.50	8.76	53/5.97	8/4	Antera	nd	Van Den Heuvel *et al.*, 2001
		Solanum pimpinellifolium	AB036641	9.50	8.74	53/5.95	8/4	Fruto	Variação morfológica	Xiao *et al.*, 2008
			AB036643	9.50	8.95	58/6.54	8/4	Fruto	Variação morfológica	
			AB036640	5.75	8.99	47/5.29	8/4	Fruto	Variação morfológica	
			AB036642	5.38	8.99	47/5.26	8/4	Fruto	Variação morfológica	
		Solanum tuberosum	P20346	9.50	9.14	58/6.55	8/4	Tubérculo	nd	Stiekema *et al.*, 1988
			ACJ26760	9.50	8.76	58/6.55	8/4	nd	nd	np
			AAB31351	nd	nd	nd/5.00	nd/nd	Flor, Tubérculo, Caule, Folha	F, B+, B-,	Moreno *et al.*, 1994
	Vitaceae	*Vitis vinifera*	TC69032	5.90	9.37	47/5.35	8/4	Bago	F	Beer e Vivier, 2008
			CBI24167	nd	9.00	55/6.39	8/4	nd	nd	np
			CBI27688	4.78	8.94	47/5.22	8/4	nd	nd	np
			CBI27507	5.90	9.37	47/5.35	8/4	nd	nd	np
			XP_002272913	5.90	8.94	47/5.22	8/4	nd	nd	np
			XP_002263380	6.10	9.47	47/5.54	8/4	nd	nd	np
			XP_002281189	8.25	8.72	57/6.38	8/4	nd	nd	np
			XP_002274353	5.90	9.37	47/5.35	8/4	nd	nd	np
			CAN69975	4.53	8.94	47/5.22	8/4	nd	nd	np
			CAN64556	8.25	8.72	57/6.38	8/4	nd	nd	np
Coniferopsida	Pinaceae	*Picea abies*	AAN40688	4.68	9.22	50/5.44	8/4	Semente, Acícula	nd	np
			CAA62761	4.68	9.06	50/5.37	8/4	Raiz	nd	Sharma e Lönneborg, 1996
	Pinaceae	*Picea glauca*	AAR84643	4.68	9.06	50/5.37	8/4	Semente	F	Pervieux *et al.*, 2004
		Picea sitchensis	ABK25062	4.32	9.06	50/5.37	8/4	Líder	nd	np
			ABK23404	4.68	9.06	50/5.37	8/4	Líder, Súber	nd	np
Cycadopsida	Cycadaceae	*Cycas revoluta*	-	nd	8.47	44/4.58	8/nd	Semente	F, B+, B-	Yokoyama *et al.*, 2008

Tabela 1 (continuada)

Classe	Família	Espécie	Aquisição ou TC	Peptídeo sinal pI	Peptídeo maduro			Tecido	Atividade Biológica	Referências
					pI	nR/kDa	nC/pD			
Cycadopsida	Cycadaceae	*Cycas revoluta*	-	nd	8.47	44/4.57	8/nd	Semente	F, B+, B-	Yokoyama *et al.*, 2008
Ginkopsida	Ginkgoaceae	*Ginhkgo biloba*	AAU04859			50/5.68	8/4	Raiz, Folha,	nd	Shen *et al.*, 2005

Peso molécula (kDa) e ponto isoelétrico (pI) foram calculados com o programa Jvirgel. As pontes dissulfídicas foram previstas com o programa Disulfind. Nos casos em que as sequencias não foram depositadas em repositórios online tais informaçãoes foram obtidas diretamente do artigo, quando disponível. Atividades biológicas: (A) inibição de α-amilase; (AC) anticancerígena; (B^{+}) inibição de bactérias Gram-positivas; (B^{-}) inibição de bactérias Gram-negativas; (DHAR) dehidroascorbato redutase; (F) fungistática; (TMP) tolerância a metais pesados; (BCI) bloqueador de canais iônicos; (IP) inibição de protease; (ICR) inibição do crescimento radicular; (AIS) anti-insetos; (M) mitogênico; (IPP) inibição do crescimento radicular de plantas parasitas; (IST) inibição do sistema de tradução; (ITR) inibição da transcriptase reversa do HIV; (VS) variação somaclonal; (AI) auto-incopatibilidade; (nd) não determinado. (np) nenhuma publicação referente. (-) sequência não depositada em repositórios online. (nd) não determinado.

Anexo B – Instruções para submissão de trabalhos na revista Protein & Peptide Letters

ISSN: 0929-8665 - Volume 17, 12 Issues, 2010

Instructions for Authors

Regular manuscripts may be submitted directly to the Editorial Office.

ONLINE MANUSCRIPT SUBMISSION: An online submission and tracking service via Internet facilitates a speedy and cost-effective submission of manuscripts. The full manuscript has to be submitted online via Bentham's Content Management System (CMS) at http://www.bentham-editorial.org/ (View Instructions)

Manuscripts must be submitted by one of the authors of the manuscript, and should not be submitted by anyone on their behalf. The principal/corresponding author will be required to submit a Covering Letter along with the manuscript, on behalf of all the co-authors (if any). The author(s) will confirm that the manuscript (or any part of it) has not been published previously or is not under consideration for publication elsewhere. Furthermore, any illustration, structure or table that has been published elsewhere must be reported, and copyright permission for reproduction must be obtained.

For all online submissions, please provide soft copies of all the materials (main text in MS Word or Tex/LaTeX), figures / illustrations in TIFF, PDF or JPEG, and chemical structures drawn in ChemDraw (CDX) / ISISDraw (TGF) as separate files, while a PDF version of the entire manuscript must also be included, embedded with all the figures / illustrations / tables / chemical structures etc.

It is imperative that before submission, authors should carefully proofread the files for special characters, mathematical symbols, Greek letters, equations, tables, references and images, to ensure that they appear in proper format.

References, figures, tables, chemical structures etc. should be referred to in the text at the appropriate place where they have been first discussed. Figure legends/captions should also be provided.

A successful electronic submission of a manuscript will be followed by a system-generated acknowledgement to the principal/corresponding author. Any queries therein should be addressed to sarwat@benthamscience.org

MANUSCRIPTS PUBLISHED: The Journal accepts mini-reviews, letters, research papers and crystallization reports written in English. Single topic/thematic issues may also be considered for publication.

Single Topic Issues: These issues may either contain invited review articles or a mixture of research, letter and review articles. A Single Topic Special Editor will offer a short Perspective and co-ordinate the solicitation of manuscripts between 3-5 (for a mini-hot topic) to 6-10 (for full-length hot topic) from leading scientists. Authors interested in editing a single topic issue in an emerging topic of proteins and peptide sciences may submit their proposal to the Editor-in-Chief at ppl@benthamscience.org for his consideration. Visit http://www.bentham.org/ppl/Special-Issues.htm for viewing forthcoming issues.

MANUSCRIPT LENGTH:

Mini-Reviews: The maximum page length for a mini review article is twenty five journal pages. Each journal page is on average 900 words.

Letters Length: The maximum total page length for Letter type articles published in the journal is nine journal pages. Each journal page is on average 900 words.

Research Articles: The maximum total page length for Research articles is ten journal pages, and the maximum thirty pages. Each journal page is on average 900 words.

Crystallization Reports: Crystallization reports are no longer routinely published by the journal. Only reports on the first crystallization of proteins of substantial interest or reports on new crystallization techniques of wide application will be considered. Crystallization reports should include data describing the source of the protein, composition, pH and concentrations of precipitants and buffers etc., the method employed, crystal size and morphology, and time of growth etc. Unit cell dimensions, space-group assignments, resolution limit, specific volume and asymmetric unit contents must also be provided. The maximum page length for Crystallization Reports is ten journal pages. Each journal page is on average 900 words.

There is no restriction on the number of figures, tables or additional files e.g. video clips, animation and datasets, that can be included with each article online. Authors should include all relevant supporting data with each article (Refer to Supplementary Material section).

MANUSCRIPT PREPARATION:

The manuscript should be written in English in a clear, direct and active style. All pages must be numbered sequentially, facilitating in the reviewing and editing of the manuscript.

For further convenience, the customer support team available at Bentham Publishing Services (www.benthampublishingservices.com) can provide assistance to authors for the preparation of manuscripts.

MANUSCRIPT SECTIONS FOR PAPERS: Manuscripts for review articles submitted to the journal may be divided into the following sections:

Covering letter
Title
Title page
Abstract
Keywords
Text organization
List of abbreviations (if any)
Conflict of interest (if any)
Acknowledgements (if any)
References
Appendices
Figures/illustrations (if any)
Chemical structures (if any)
Tables and captions (if any)
Supportive/supplementary material (if any)

Covering Letter: It is mandatory that a signed covering letter also be submitted along with the manuscript by the author to whom correspondence is to be addressed, delineating the scope of the submitted article declaring the potential competing interests, acknowledging contributions from authors and funding agencies, and certifying that the paper is prepared according to the 'Instructions for Authors'. All inconsistencies in the text and in the reference section and any typographical errors must be carefully checked and corrected before the submission of the manuscript. The article contains no such material or information that may be unlawful, defamatory, fabricated, plagiarized, or which would, if published, in any way whatsoever, violate the terms and conditions as laid down in the copyright agreement. The authors acknowledge that the publishers have the legal right to take appropriate action against the authors for any such violation of the terms and conditions as laid down in the copy right agreement. DOWNLOAD COVERING LETTER.

Title: The title of the article should be precise and brief and must not be more than 120 characters. Authors should avoid the use of non-standard abbreviations. The title must be written in title case except for articles, conjunctions and prepositions.

Authors should also provide a short 'running title'. Title, running title, byline, correspondent, footnote and key words should be written as presented in original manuscripts.

Title Page: Title page should include paper title, author(s) full name and affiliation, corresponding author(s) names complete affiliation/address, along with phone, fax and email.

Abstract: The abstract should not exceed 250 words for review and research papers and should be limited to only 150 words for letters, summarizing the essential features of the article. The use of abbreviations should be reduced to a minimum and the references should not be cited in the abstract.

Key Words: Provide 6 to 8 key words in alphabetical order.

Text Organization: The main text should begin on a separate page and should be divided into separate sections. The manuscript should be divided into title page, abstract and the main text. The text may be subdivided further according to the areas to be discussed, which should be followed by the Acknowledgement (if any) and Reference sections. For Letter/Research Articles, the manuscript should begin with the title page and abstract followed by the main text, which must be structured into separate sections as Introduction, Materials and Methods, Results, Discussion, Conclusion, Acknowledgements and References. The review article should mention any previous important recent and old reviews in the field and contain a comprehensive discussion starting with the general background of the field. It should then go on to discuss the salient features of recent developments. The authors should avoid presenting material which has already been published in a previous review. The authors are advised to present and discuss their observations in brief. The manuscript style must be uniform throughout the text and 10 pt Times New Roman fonts should be used. The full term for an abbreviation should precede its first appearance in the text unless it is a standard unit of measurement. The reference numbers should be given in square brackets in the text. Italics should be used for Binomial names of organisms (Genus and Species), for emphasis and for unfamiliar words or phrases. Non-assimilated words from Latin or other languages should also be italicized e.g. per se, *et al*. etc.

Greek Symbols and Special Characters: Greek symbols and special characters often undergo formatting changes and get corrupted or lost during preparation of manuscript for publication. To ensure that all special characters used are embedded in the text, these special characters should be

inserted as a symbol but should not be a result of any format styling (Symbol font face) otherwise they will be lost during conversion to PDF/XML.2

Authors are encouraged to consult reporting guidelines. These guidelines provide a set of recommendations comprising a list of items relevant to their specific research design. Chemical equations, chemical names, mathematical usage, unit of measurements, chemical and physical quantity & units must conform to SI and Chemical Abstracts or IUPAC.

All kinds of measurements should be reported only in International System of Units (SI).

List of Abbreviations: If abbreviations are used in the text either they should be defined in the text where first used, or a list of abbreviations can be provided.

Conflict of Interest: Financial contributions to the work being reported should be clearly acknowledged, as should any potential conflict of interest.

Acknowlegements: Please acknowledge anyone (individual/company/institution) who has contributed to the study by making substantial contributions to conception, design, acquisition of data, or analysis and interpretation of data, or who was involved in drafting the manuscript or revising it critically for important intellectual content. Please list the source(s) of funding for the study, for each author, and for the manuscript preparation in the acknowledgements section.

This journal complies with the International Committee of Medical Journal Editors' Uniform Requirements for Manuscripts Submitted to Biomedical Journals http://www.icmje.org and the FDA's Good Reprint Practices for the Distribution of Medical Journal Articles and Medical or Scientific Reference Publications on Unapproved New Uses of Approved Drugs and Approved or Cleared Medical Devices http://www.fda.gov/oc/op/goodreprint.html

References: References must be listed in the ACS Style only. All references should be numbered sequentially [in square brackets] in the text and listed in the same numerical order in the reference section. The reference numbers must be finalized and the bibliography must be fully formatted before submission.

See below few examples of references listed in the ACS Style:

Journal Reference:
[1] Bard, M.; Woods, R.A.; Bartón, D.H.; Corrie, J.E.; Widdowson, D.A. Sterol mutants of Saccharomyces cerevisiae: chromatographic analyses. Lipids, 1977, 12(8), 645-654.

[2] Zhang, W.; Brombosz, S.M.; Mendoza, J.L.; Moore, J.S. A high-yield, one-step synthesis of o-phenylene ethynylene cyclic trimer via precipitation-driven alkyne metathesis. J. Org. Chem., 2005, 70, 10198-10201.

Book Reference:
[3] Crabtree, R.H. The Organometallic Chemistry of the Transition Metals, 3rd ed.; Wiley & Sons: New York, 2001.

Book Chapter Reference:
[4] Wheeler, D.M.S.; Wheeler, M.M. In: Studies in Natural Products Chemistry; Atta-ur-Rahman, Ed.; Elsevier Science B. V: Amsterdam, 1994; Vol. 14, pp. 3-46.

Conference Proceedings:
[5] Jakeman, D.L.; Withers, S.G.E. In: Carbohydrate Bioengineering: Interdisciplinary Approaches, Proceedings of the 4th Carbohydrate Bioengineering Meeting, Stockholm, Sweden, June 10-13, 2001; Teeri, T.T.; Svensson, B.; Gilbert, H.J.; Feizi, T., Eds.; Royal Society of Chemistry: Cambridge, UK, 2002; pp. 3-8.

URL(WebPage):
[6] National Library of Medicine. Specialized Information Services: Toxicology and Environmental Health. http://sis.nlm.nih.gov/Tox/ToxMain.html (Accessed May 23, 2004).

Patent:
[7] Hoch, J.A.; Huang, S. Screening methods for the identification of novel antibiotics. U.S. Patent 6,043,045, March 28, 2000.

Thesis:
[8] Mackel, H. Capturing the Spectra of Silicon Solar Cells. PhD Thesis, The Australian National University: Canberra, December 2004.

E-citations:
[9] Citations for articles/material published exclusively online or in open access (free-to-view) , must contain the exact Web addresses (URLs) at the end of the reference(s), except those posted on an author's Web site unless editorially essential, e.g. 'Reference: Available from: URL'.

Some important points to remember:

• All references must be complete and accurate.
• All authors must be cited and there should be no use of the phrase *et al*.
• Online citations should include the date of access.
• Journal abbreviations should follow the Index Medicus/MEDLINE.
• Take special care of the punctuation convention as described in the above-mentioned examples.
• Superscript in the in-text citations and reference section should be avoided.
• Abstracts, unpublished data and personal communications (which can only be included if prior permission has been obtained) should not be given in the reference section but they may be mentioned in the text and details provided as footnotes.
• The authors are encouraged to use a recent version of EndNote (version 5 and above) or Reference Manager (version 10) when formatting their reference list, as this allows references to be automatically extracted.

Appendices: In case there is a need to present lengthy, but essential methodological details, use appendices, which can be a part of the article. An appendix must not exceed three pages (Times New Roman, 12 point fonts, 900 max. words per page).The information should be provided in a condensed form, ruling out the need of full sentences. A single appendix should be titled APPENDIX, while more than one can be titled APPENDIX A, APPENDIX B, and so on.

Figures / Illustrations:

All authors must strictly follow the guidelines below for preparing illustrations for publication in Protein & Peptide Letters. If the figures are found to be sub-standard, then the manuscripts will be rejected/ and the authors offered the option of figure improvement professionally by Bentham Publishing Services. The costs for such improvement will be charged to the authors.

The authors should provide the illustrations as separate files, as well as embedded in the text file, numbered consecutively in the order of their appearance. Each figure should include a single illustration. Each figure should be closely cropped to minimize the amount of white space surrounding the illustration.

If a figure consists of separate parts, it is important that a single composite illustration file be submitted, containing all parts of the figure.

Photographs should be provided with a scale bar if appropriate, as well as high-resolution component files.

Scaling/Resolution

For Line Art image type, which is generally an image based on lines and text and does not contain tonal or shaded areas, the preferred file format is TIFF or EPS, with colour mode being Monochrome 1-bit or RGB, in a resolution of 900-1200 dpi.

For Halftone image type, which is generally a continuous tone photograph and contains no text, the preferred file format is TIFF, with colour mode being or RGB or Grayscale, in a resolution of 300 dpi.

For Combination image type, which is generally an image containing halftone in addition to text or line art elements, the preferred file format is TIFF, with colour mode being or RGB or Grayscale, in a resolution of 500-900 dpi.

Formats

For illustrations, the following file formats are acceptable:

- Illustrator
- EPS (preferred format for diagrams)
- PDF (also especially suitable for diagrams)
- PNG (preferred format for photos or images)
- Microsoft Word (version 5 and above; figures must be a single page)
- PowerPoint (figures must be a single page)
- TIFF
- JPEG (conversion should be done using the original file)
- BMP
- CDX (ChemDraw)
- TGF (ISISDraw)

Bentham Science does not process figures submitted in GIF format.

If the large size of TIFF or EPS figures acts as an obstacle to online submission, authors may find that conversion to JPEG format before submission results in significantly reduced file size and upload time, while retaining acceptable quality. JPEG is a 'lossy' format, however. In order to maintain acceptable image quality, it is recommended that JPEG files are saved at High or Maximum quality.

Files should not be compressed with tools such as Zipit or Stuffit prior to submission as these tools

will in any case produce negligible file-size savings for JPEGs and TIFFs, which are already compressed.

Please do not:

1. Supply embedded graphics in your word processor (spreadsheet, presentation) document;
2. Supply files that are optimized for screen use (like GIF, BMP, PICT, WPG); the resolution is too low;
3. Supply files that are too low in resolution;
4. Submit graphics that are disproportionately large for the content.

Image Conversion Tools:

There are many software packages, many of them freeware or shareware, capable of converting to and from different graphics formats, including PNG.

Good general tools for image conversion include GraphicConverter on the Macintosh, PaintShop Pro, for Windows, and ImageMagick, which is available on Macintosh, Windows and UNIX platforms.

Note that bitmap images (e.g. screenshots) should not be converted to EPS, since this will result in a much larger file size than the equivalent JPEG, TIFF, PNG or BMP, with no increase in quality. EPS should only be used for images produced by vector-drawing applications such as Adobe Illustrator or CorelDraw. Most vector-drawing applications can be saved in, or exported as, EPS format. In case the images have been originally prepared in an Office application, such as Word or PowerPoint, then the original Office files should be directly uploaded to the site, instead of being converted to JPEG or another format that may be of low quality.

Color Figures / Illustrations:

• The cost for the first published page of color figures is US$ 893; the second additional page will be for US$ 652 and each subsequent page for US$ 441.

• Color figures should be supplied in CMYK not RGB colors.

Chemical Structures:

Chemical structures MUST be prepared in ChemDraw (CDX) file and provided as separate file.

Structure Drawing Preferences:

[as according to the ACS style sheet]

Drawing Settings:

Chain angle 120°

Bond spacing 18% of width

Fixed length 14.4 pt (0.500cm, 0.2in)

Bold width 2.0 pt (0.071cm, 0.0278in)

Line width 0.6 pt (0.021cm, 0.0084in)

Margin width 1.6 pt (0.096cm)

Harsh spacing 2.5 pt (0.088cm, 0.0347in)

Text settings:

Font Times New Roman

Size 12pt

Under the Preference Choose:

Units points

Tolerances 3 pixels

Under Page Setup Use:

Paper US letter

Scale 100%

Tables:

- Data Tables should be submitted in Microsoft Word table format.
- Each table should include a title/caption being explanatory in itself with respect to the details discussed in the table. Detailed legends may then follow.
- Table number in bold font i.e. Table 1, should follow a title. The title should be in small case with the first letter in caps. A full stop should be placed at the end of the title.
- Tables should be embedded in the text exactly according to their appropriate placement in the submitted manuscript.
- Columns and rows of data should be made visibly distinct by ensuring that the borders of each cell are displayed as black lines.
- Tables should be numbered in Arabic numerals sequentially in order of their citation in the body of the text.
- If a reference is cited in both the table and text, please insert a lettered footnote in the table to refer to the numbered reference in the text.
- Tabular data provided as additional files can be submitted as an Excel spreadsheet.

Supportive/Supplementary Material: We do encourage to append supportive material, for example a PowerPoint file containing a talk about the study, a PowerPoint file containing additional screenshots, a Word, RTF, or PDF document showing the original instrument(s) used, a video, or the original data (SAS/SPSS files, Excel files, Access Db files etc.) provided it is inevitable or endorsed by the journal's Editor.

Published/reproduced material should not be included unless you have obtained written permission from the copyright holder, which must be forwarded to the Editorial Office in case of acceptance of

your article for publication.

Supportive/Supplementary material intended for publication must be numbered and referred to in the manuscript but should not be a part of the submitted paper. In-text citations as well as a section with the heading "Supportive/Supplementary Material" before the "References" section should be provided. Here, list all Supportive/Supplementary Material and include a brief caption line for each file describing its contents.

Any additional files will be linked into the final published article in the form supplied by the author, but will not be displayed within the paper. They will be made available in exactly the same form as originally provided only on our Web site. Please also make sure that each additional file is a single table, figure or movie (please do not upload linked worksheets or PDF files larger than one sheet). Supportive/ Supplementary material must be provided in a single zipped file not larger than 4 MB.

Authors must clearly indicate if these files are not for publication but meant for the reviewers'/editors' perusal only.

PERMISSION FOR REPRODUCTION: Published/reproduced material should not be included unless you have obtained written permission from the copyright holder, which should be forwarded to the Editorial Office in case of acceptance of your article for publication.

For obtaining permission for reproducing any material published in an article by Bentham Science Publishers, please fill in the request FORM and send to ppl@benthamscience.org for consideration.

AUTHORS AND INSTITUTIONAL AFFILIATIONS: The author will be required to provide their full names, the institutional affiliations and the location, with an asterisk in front of the name of the principal/corresponding author. The corresponding author(s) should be designated and their complete address, business telephone and fax numbers and e-mail address must be stated to receive correspondence and galley proofs.

PAGE CHARGES: No page charges will be levied to authors for the publication of their article.

LANGUAGE AND EDITING: Manuscripts submitted containing many English typographical errors will not be published. Authors from non-English language countries are advised to use the services of our professional language editing department prior to submitting their manuscript to the Journal. Please contact Bentham Publishing Services www.benthampublishingservices.com for a language editing quote at e-mail: info@benthampublishingservices.com stating the total number of words of the article to be edited.

PROOF CORRECTIONS: Authors will receive page proofs of their accepted paper before publications. To avoid delays in publication, proofs should be checked immediately for typographical errors and returned within 48 hours. Major changes are not acceptable at the proof stage. If unable to send corrections within 48 hours due to some reason, the author(s) must at least send an acknowledgement on receiving the galley proofs or the article will be published exactly as received and the publishers will not be responsible for any error occurring in the published manuscript in this regard.

The corresponding author will be solely responsible for ensuring that the revised version of the manuscript incorporating all the submitted corrections receives the approval of all the co-authors of the manuscript.

REPRINTS: Each first-named (corresponding) author will receive electronically a certain number of free reprints. Reprints may be ordered from the Publisher prior to publication of the article. First named authors may also order a personal print and online subscription of the journal at 50% off the normal subscription rate by contacting the subscription department at e-mail: subscriptions@bentham.org

OPEN ACCESS PLUS: Accepted articles can be published online for free open access for all to view. Open access publishing provides the maximum dissemination of the article to the largest audience. Authors must pay for this service. All corresponding authors will be asked to indicate whether or not they wish to pay to have their paper made freely available on publication. If authors do not select the Open Access option, then their article will be published with standard subscription-based access at no charge.

Bentham Science offers authors the choice of open access publication of their articles at a fee of US$ 2,800 per published article which allows indefinite free-to-view online publication with Bentham Science.

Alternatively, authors may choose to publish their article with Bentham Science at a reduced fee for a limited open access period. Bentham Science is the first and only publisher to offer authors the choice at a reduced open access fee to have their article published for a limited open access period. For open access publication for a period of either twelve months or two months (limited open access option) the per article fee is US $1,400 or US $700, respectively.

For more information please contact us at e-mail: openaccess@bentham.org

REVIEWING AND PROMPTNESS OF PUBLICATION: All papers submitted for publication will be immediately subjected to editorial scrutiny, usually in consultation with members of the Editorial Advisory Board. Every effort will be made to assess the papers quickly. The papers will be typeset and the proofs dispatched to the authors normally within 4 weeks of their acceptance.

COPYRIGHT: Authors who publish in Bentham Science print & online journals will transfer copyright to their work to Bentham Science Publishers. Submission of a manuscript to the respective journals implies that all authors have read and agreed to the content of the Covering Letter or the Terms and Conditions. It is a condition of publication that manuscripts submitted to this journal have not been published and will not be simultaneously submitted or published elsewhere. Plagiarism is strictly forbidden, and by submitting the article for publication the authors agree that the publishers have the legal right to take appropriate action against the authors, if plagiarism or fabricated information is discovered. By submitting a manuscript the authors agree that the copyright of their article is transferred to the publishers if and when the article is accepted for publication. Once submitted to the journal, the author will not withdraw their manuscript at any stage prior to publication.

E-Pub Ahead of Schedule: Bentham Science Publishers are pleased to offer electronic publication of accepted papers prior to scheduled publication. These peer-reviewed papers can be cited using the date of access and the unique DOI number. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue.

Disclaimer: Articles appearing in E-Pub Ahead-of-Schedule sections have been peer-reviewed and accepted for publication in this journal and posted online before scheduled publication. Articles appearing here may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Accordingly, Bentham Science Publishers, the editors and authors and their

respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained of articles in the E-Pub Ahead-of-Schedule.

Anexo C – Instruções para a submissão de trabalhos na revista Current Protein & Peptide Science

ISSN: 1389-2037 - Volume 10, 6 Issues, 2009

Instructions for Authors

Manuscripts should be submitted openly to the Editor-in-Chief without prior invitation at the following address:

Prof. Ben M. Dunn
Editor-in-Chief
Current Protein and Peptide Science
Department of Biochemistry and Molecular Biology
University of Florida, College of Medicine
P.O. Box 100245, Gainesville
Florida 32610-0245
USA
Fax: +1 352 846 0412
E-Mail: bdunn@ufl.edu

For full review articles, the manuscript, written in English, should not exceed 70 pages in length (ca. 30 printed pages) and for mini-reviews a maximum of 40 pages in length (ca. 15 printed pages). The manuscript, written in English, should be typed in double spacing. All pages should be numbered. Abbreviations should be defined the first time that they are used in the manuscript and a list of abbreviations used should be provided.

Manuscript Organisation: The manuscript should be divided as: Title page, abstract and the main text. The text may be subdivided further according to the areas to be discussed. This should be followed by the Acknowledgement (if any) and Reference sections. The first page should contain the title, the authors names (initials and surname only), with an asterisk in front of the name of the principal author.

COVERING LETTER: It is a mandatory requirement that a signed covering letter also be submitted along with the manuscript by the author to whom correspondence is to be addressed, delineating the scope of the submitted article declaring the potential competing interests, acknowledging contributions from authors and funding agencies, and certifying that the paper is prepared according to the 'Instructions for Authors'. All inconsistencies in the text and in the reference section, and any typographical errors must be carefully checked and corrected before the submission of the manuscript. The article contains no such material or information that may be unlawful, defamatory, fabricated, plagiarized, or which would, if published, in any way whatsoever, violate the terms and conditions as laid down in the agreement. The authors acknowledge that the publishers have the legal right to take appropriate action against the authors for any such violation of the terms and conditions as laid down in the agreement. Download the Covering letter

Conflict of Interest: Financial contributions to the work being reported should be clearly acknowledged, as should any potential conflict of interest.

ACKNOWLEDGEMENTS: Please acknowledge anyone (individual/company/institution) who has contributed to the study by making substantial contributions to conception, design,

acquisition of data, or analysis and interpretation of data, or who was involved in drafting the manuscript or revising it critically for important intellectual content. Please list the source(s) of funding for the study, for each author, and for the manuscript preparation in the acknowledgements section.

This journal complies with the International Committee of Medical Journal Editors' Uniform Requirements for Manuscripts Submitted to Biomedical Journals http://www.icmje.org and the FDA's Good Reprint Practices for the Distribution of Medical Journal Articles and Medical or Scientific Reference Publications on Unapproved New Uses of Approved Drugs and Approved or Cleared Medical Devices http://www.fda.gov/oc/op/goodreprint.html

Abstract: The abstract should not exceed 250 words and it should condense the essential features of the review article, with the focus on the major advances in the field.

Key Words: Authors must supply up to eight key words along with their manuscript.

Text: The main text should begin on a separate page and it may be subdivided into separate sections. The review article should mention any previous important reviews in the field and contain a comprehensive discussion starting with the general background of the field. It should then go on to discuss the salient features of recent developments. The authors should avoid presenting material which has already been published in a previous review. The reference numbers should be given in square brackets in the text. Acknowledgements should be kept to a minimum.

Tables:

Data Tables should be submitted in Microsoft Excel or in Microsoft Word table format.
Each table should include a title/caption explaining what the table shows. Detailed legends may then follow.
Tables should be given on separate pages with indications on the left hand margin to the text, as to the appropriate placement of the tables.
Tables should not contain vertical rules.
Columns and rows of data should be made visibly distinct by ensuring the borders of each cell display as black lines.
Tables should be numbered consecutively in order of their citation in the body of the text, with Arabic numerals.
If a reference is cited in both the table and text, insert a lettered footnote in the table to refer to the numbered reference in the text.
Tabular data provided as additional files can be submitted as an Excel spreadsheet.
Figures/Illustrations: All authors must strictly follow the guidelines below for preparing illustrations for publication in Current Protein & Peptide Science. If the figures are found to be sub-standard, then the manuscripts will be rejected/ and the authors offered the option of figure improvement professionally by Bentham Publishing Services. The costs for such improvement will be charged to the authors.

Editorial Requirements:

The quality of the illustrations printed in the journal largely depends on the quality of the figures/illustration provided by the author. The acceptance of a manuscript for publication is subject to the figures being up to the highest standards in terms of clarity and resolution. In

case the figures supplied are not up to the standards required, then the manuscripts will be rejected, and the authors given the option to either return the manuscripts within one week with appropriate drawings/figures or they may opt to have these professionally prepared through the paid professional service available from Bentham Publishing Services.

Submit the original artwork or a photographic print of the original for publication.
Do not submit photocopies. (Hardcopy submission requirement).
Illustrations should be provided as separate files, not embedded in the text. (Hardcopy submission requirement).
Remove all color from graphics except for those illustrations (Greyscale), which are intended to be printed in color.
Good quality of hardcopy originals are a requirement.
Figures should be referred to as "Fig. (1)", "Fig. (2)" etc. in the text with figure numbers in bold within round bracket.
Revised manuscripts should contain a complete set of artwork and photographs.
Technical Details:

Figures should be provided on high quality, white, smooth opaque paper. (Hardcopy requirement).
Format & Resolution:

The following file formats can be accepted. (Preference in order of appearance): PDF, TIFF, JPEG, GIF.
Illustrations must fit a single or double column format on the journal page, according to the following guideline below:
HEIGHT
WIDTH
SINGLE COLUMN
24 cm (max)
8.5 cm (max)
DOUBLE COLUMN
24 cm (max)
18 cm (max)

Whenever possible, submit graphics that do not have to be reduced to fit the standard figure size.
Use the best resolution available (hardcopy submission requirement). For online submission the maximum resolution is 300 dpi.
Avoid textures and shadings giving a three-dimensional effect to the illustration.
Symbols and lettering in the illustration should be of similar size.
Smaller lettering is used.
Font: Times New Roman or Helvetica is preferable. Font size: 10pt (maximum).
Line drawings should have clear and sharp lines and should be of uniform density. Moreover, lines should be continuous without any breaks.
Do not use lines thinner than 1 point.
Color Figures/Illustrations:

The cost for the first published page of color figures is US $850; the second additional page will be US $621; and each subsequent page, US $420.
Color figures should be supplied in CMYK not RGB colors.

Required resolution for online submission is 300 dpi. In case of hardcopy submission, please use the maximum resolution available. In case the figure(s) are of an inferior quality they will need to be improved either by the author or by the publisher on a charged basis.
Photographs: Submit high-contrast prints that are of single- or double-column width so that they will not have to be reduced when printed. Negatives are not acceptable. However, in case of hardcopy submission, color illustrations and plates can be published from 35 mm color slides.

Chemical Structures: Chemical structures must be prepared according to the guidelines below.

Structures should be prepared using a suitable drawing program e.g. ChemDraw and provided as separate file, submitted both on disk and in printed formats.

Numbers in bold should identify all chemical structure. If a name is also used to describe the compound, the associated number in bold should be in parenthesis, e.g. "compound (1)".

Structure Drawing Preferences:

[As according to the ACS style sheet]

Drawing Settings:

Chain angle 120°

Bond spacing 18% of width

Fixed length 14.4 pt (0.500cm, 0.2in)

Bold width 2.0 pt (0.071cm, 0.0278in)

Line width 0.6 pt (0.021cm, 0.0084in)

Margin width 1.6 pt (0.096cm)

Harsh spacing 2.5 pt (0.088cm, 0.0347in)

Text settings:

Font Times New Roman / Helvetica

Size 10pt

Under the Preference Choose:

Units points

Tolerances 3 pixels

Under Page Setup Use:

Paper US letter

Scale 100%

References: References must be listed in the ACS Style only. All references should be numbered sequentially [in square brackets] in the text and listed in the same numerical order in the reference section. The reference numbers must be finalized and the bibliography must be fully formatted before submission. Titles of journals are abbreviated according to Chemical Abstract Service Source Index.

See below few examples of references listed in the ACS Style:

Examples:

Journal Reference:
[1] Banner, D.W.; Harvary, P. Crystallographic analysis at 3.0-A resolution of the binding to human thrombin of four active site-directed inhibitors. J. Biol. Chem., 1991, 266, 20085-20093

Book Reference:
[2] Crabtree, R.H. The Organometallic Chemistry of the Transition Metals, 3rd ed.; Wiley & Sons: New York, 2001.

Book Chapter Reference:
[3] Wheeler, D.M.S.; Wheeler, M.M. In Studies in Natural Products Chemistry; Atta-ur-Rahman, Ed.; Elsevier Science B. V: Amsterdam, 1994; Vol. 14, pp. 3-46.

Conference Proceedings:
[4] Jakeman, D.L.; Withers, S.G. E. In Carbohydrate Bioengineering: Interdisciplinary Approaches, Proceedings of the 4th Carbohydrate Bioengineering Meeting, Stockholm, Sweden, June 10-13, 2001; Teeri, T.T.; Svensson, B.; Gilbert, H.J.; Feizi, T., Eds.; Royal Society of Chemistry: Cambridge, UK, 2002; pp. 3-8.

URL(Web Page):
[5] National Library of Medicine. Specialized Information Services: Toxicology and Environmental Health. http://sis.nlm.nih.gov/Tox/ToxMain.html (accessed May 23, 2004).

Patent:
[6] Hoch, J.A.; Huang, S. Screening methods for the identification of novel antibiotics. U.S. Patent 6,043,045, March 28, 2000.

Thesis:
[7] Kirby, C.W. Thesis, University of Waterloo, 2000.

Some important points to remember:

*All references must be complete and accurate.
*All authors must be cited and there should be no use of the phrase *et al* (the term "*et al.*" should be in italics).
*Online citations should include the date of access.

*Journal abbreviations should follow the Index Medicus/MEDLINE.
*Take special care of the punctuation convention as described in the above-mentioned examples.
*Avoid using superscript in the in-text citations and reference section.
*Abstracts, unpublished data and personal communications (which can only be included if prior permission has been obtained) should not be given in the reference section but they may be mentioned in the text and details provided as footnotes.
*The authors are encouraged to use a recent version of EndNote (version 5 and above) or Reference Manager (version 10) when formatting the reference list, as this allows references to be automatically extracted.

SUPPORTIVE/SUPPLEMENTARY MATERIAL: We do encourage to append supportive material, for example a PowerPoint file containing a talk about the study, a PowerPoint file containing additional screenshots, a Word, RTF, or PDF document showing the original instrument(s) used, a video, or the original data (SAS/SPSS files, Excel files, Access Db files etc.) provided it is inevitable or endorsed by the journal's Editor.

Published/reproduced material should not be included unless you have obtained written permission from the copyright holder, which must be forwarded to the Editorial Office in case of acceptance of your article for publication.

Supportive/Supplementary material intended for publication must be numbered and referred to in the manuscript but should not be a part of the submitted paper. In-text citations as well as a section with the heading "Supportive/Supplementary Material" before the "References" section should be provided. Here, list all Supportive/Supplementary Material and include a brief caption line for each file describing its contents.

Any additional files will be linked into the final published article in the form supplied by the author, but will not be displayed within the paper. They will be made available in exactly the same form as originally provided only on our Web site. Please also make sure that each additional file is a single table, figure or movie (please do not upload linked worksheets or PDF files larger than one sheet). Supportive/ Supplementary material must be provided in a single zipped file not larger than 4 MB.

Authors must clearly indicate if these files are not for publication but meant for the reviewers'/editors' perusal only.

Online Manuscript Submission: An online submission and tracking service via Internet facilitates a speedy and cost-effective submission of manuscripts.1 The full manuscript has to be submitted online via Bentham's Content Management System (CMS) at http://www.bentham-editorial.org/ View Instructions

Alternatively, you may also submit your full manuscript by e-mail to cpps@benthamscience.org

Manuscripts must be submitted by one of the authors of the manuscript, and should not be submitted by anyone on their behalf. The principal/corresponding author will be required to submit a Covering Letter along with the manuscript, on behalf of all the co-authors (if any). The author(s) will confirm that the manuscript (or any part of it) has not been published previously or is not under consideration for publication elsewhere. Furthermore, any

illustration, structure or table that has been published elsewhere must be reported, and copyright permission for reproduction must be obtained.

For all online submissions, please provide your complete manuscript in the form of a single zipped folder containing soft copies of all the materials (main text in MS Word or Tex/LaTeX), figures / illustrations in TIFF, PDF or JPEG, and chemical structures drawn in ChemDraw (CDX) / ISISDraw (TGF) as separate files, while a PDF version of the entire manuscript must also be included, embedded with all the figures / illustrations / tables / chemical structures etc.

A successful electronic submission of a manuscript will be followed by a system-generated acknowledgement to the principal/corresponding author within 72 hours of the dispatch of the manuscript. Any queries therein should be addressed to mahmood@bentham.org and copied to Jalil@bentham.org

Page Charges: No page charges will be levied to authors.

Language Editing: Manuscripts submitted containing many English typographical errors will not be published. Authors from non-English language countries are advised to use the services of our professional language editing department prior to submitting their manuscript to the Journal. Please contact Bentham Publishing Services www.benthampublishingservices.com for a language editing quote at e-mail: info@benthampublishingservices.com stating the total number of words of the article to be edited.

Proofs: Authors are sent page proofs. To avoid delays in publication, proofs should be checked immediately for typographical errors and returned within 48 hours. Authors will be charged for excessive changes made at the page proof stage.

Reprints: Each first-named (corresponding) author will receive electronically a certain number of free reprints. Reprints may be ordered from the Publisher prior to publication of the article. First named authors may also order a personal print and online subscription of the journal at 50% off the normal subscription rate by contacting the subscription department at e-mail: subscriptions@bentham.org

Open Access Plus: Accepted articles can be published online for free open access for all to view. Open access publishing provides the maximum dissemination of the article to the largest audience. Authors must pay for this service. All corresponding authors will be asked to indicate whether or not they wish to pay to have their paper made freely available on publication. If authors do not select the Open Access option, then their article will be published with standard subscription-based access at no charge.

Bentham Science offers authors the choice of open access publication of their articles at a fee of USD 2,800 per published article which allows indefinite free-to-view online publication with Bentham Science.

For more information please contact us at e-mail: openaccess@bentham.org

Reviewing and Promptness of Publication: All papers submitted for publication will be immediately subjected to editorial scrutiny, usually in consultation with members of the Editorial Advisory Board. Every effort will be made to assess the papers quickly. The papers

will be typeset and the proofs dispatched to the authors normally within 4 weeks of their acceptance.

Copyright: Authors who publish in Bentham print & online journals will transfer copyright to their work to Bentham Science Publishers. Submission of a manuscript to the respective journals implies that all authors have read and agreed to the content of the Covering Letter or the Terms and Conditions. It is a condition of publication that manuscripts submitted to this journal have not been published and will not be simultaneously submitted or published elsewhere. Plagiarism is strictly forbidden, and by submitting the article for publication the authors agree that the publishers have the legal right to take appropriate action against the authors, if plagiarism or fabricated information is discovered. By submitting a manuscript the authors agree that the copyright of their article is transferred to the publishers if and when the article is accepted for publication. Once submitted to the journal, the author will not withdraw their manuscript at any stage prior to publication.

E-Pub Ahead of Schedule: Bentham Science Publishers are pleased to offer electronic publication of accepted papers prior to scheduled publication. These peer-reviewed papers can be cited using the date of access and the unique DOI number. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue.

Disclaimer: Articles appearing in E-Pub Ahead-of-Schedule sections have been peer-reviewed and accepted for publication in this journal and posted online before scheduled publication. Articles appearing here may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Accordingly, Bentham Science Publishers, the editors and authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained of articles in the E-Pub Ahead-of-Schedule.