Plant Antimicrobial Peptides: An Overview of SuperSAGE Transcriptional Profile and a Functional Review

E.A. Kido1, 2*, V. Pandolfi1, L.M. Houllou-Kido1, P.P. Andrade1, F.C. Marcelino1, A.L. Nepomuceno1, R.V. Abdelnoor3, W.L. Burnquist4 and A.M. Benko-Isepnon1

1Department of Genetics, Federal University of Pernambuco, Recife, PE, Brazil; 2Northeastern Center of Strategic Technologies -- CETENE, Recife, PE, Brazil; 3Embrapa Soybean, Londrina, PR, Brazil; 4Center of Sugarcane Technology -- CTC, Piracicaba, SP, Brazil

Abstract: Defensin, thionin and lipid transfer protein (LTP) gene families, which antimicrobial activity has an attractive use in protein engineering and transgenic production of agronomically important plants, have been here functionally reviewed. Also, a transcriptional overview of plant SuperSAGE libraries and analysis of 26 bp tags possibly annotated for those families are presented. Tags differentially expressed (p < 0.05) or constitutively transcribed were identified from leaves or roots from SuperSAGE libraries from important Brazilian plant crops [cowpea (Vigna unguiculata (L.) Walp.), soybean (Glycine max (L.) Merr.) and modern sugarcane hybrids (Saccharum spp.)] submitted to abiotic [salt (100 mM NaCl) or drought] or biotic stresses [fungus inoculation (Phakopsora pachyrhizi; Asiatic Soybean Rust phytopathogen)]. The diverse transcriptional patterns observed, probably related to the variable range of targets and functions involved, could be the first step to unravel the antimicrobial peptide world and the plant stress response relationship. Moreover, SuperSAGE opens the opportunity to find some SNPs or even rare transcripts that could be important on plant stress resistance mechanisms. Putative defensin or LTPs revealed by SuperSAGE following a specific plant treatment or physiologic condition could be useful for future use in genetic improvement of plants.

Keywords: PR proteins, defensin, thionin, lipid transfer protein, cowpea, sugarcane, soybean, stress.

INTRODUCTION

The present optimized plant systemic defense was produced by the combined evolution of different stress response mechanisms, which allowed these systems to interact with each other. The local defense activation initiates at pathogen invasion sites and normally produces a whole-plant resistance response, called systemic acquired resistance (SAR). During SAR expression, plants have an effective defense response against a wide range of pathogens and abiotic stresses [1, 2]. A group of plant-coded proteins, induced in response to different stresses, have an important role in plant defense against pathogenic agents and, in general, allows a better plant adaptation to stressful environment conditions [2]. The efficiency of such plant stress responses depends on well-timed stress recognition events. Some proteins have been associated with plant inducible defense reactions to various pathogenic microorganisms [3]. In this way, each plant cell acts as a recognition site against pathogen-derived metabolites, known as elicitors. A wide variety of natural and synthetic elicitors have been characterized, such as: glycoproteins, polysaturated fatty acids, fragments of chitin and 1,3-glucans, and others [4]. The antimicrobial peptides (AMPs) play a role in the "innate immunity" system of plants and seem to be a part of a developmentally regulated, pre-existing defense barrier [5]. These inducible defense-related proteins accumulate in plant tissue after a pathogen infection [6, 7]. The AMPs are widespread in nature and are synthesized in organisms from both Plant and the Animal kingdoms [8]. In plants, most of these proteins, named "pathogenesis-related (PR)" proteins, have been classified into 17 classes (numbered in the order in which they were discovered: from PR-1 to PR-17) [7, 9, 10]. Among these PR gene families, PR-12 type includes defensins [11], PR-13 type contains thionins [12, 13] and PR-14 type encompasses lipid transfer proteins - LTPs [14]. These PR protein classes embrace the main groups of antimicrobial peptides found in plants [15], identified during the last decade [10]. Some members of these families are recognized for their fundamental role in plant defense [6, 16, 17], exhibiting a vast anti-microbial activity. The PR protein main target is the microorganism plasma membrane, leading to fungal and bacterial pathogen plasmolysis, which in turn inhibits pathogen growth and development [18]. Some PR-12, PR-13 and PR-14 genes, producing many important plant AMPs, have been reported to be either constitutively expressed or induced by biotic and/or abiotic stresses in many plant species [13, 18-22].

The present review will focus on defensins, thionins and lipid transfer proteins AMP gene families, which antimicrobial activity has an attractive use in protein engineering and transgenic production of agronomically important plants. We will also present an overview of an on-going SuperSAGE analysis looking for AMP candidates following a specific plant treatment or physiologic condition for future use in Brazilian programs of plant genetic improvement.

*Address correspondence to this author at: Av. Prof. Morais Rego, 9/4, Cidade Universitária, Recife, PE, 50670-901, Brazil; Tel: +55 81 2126 8520; Fax: +55 81 2126 8521; E-mail: ederson.kido@gmail.com

© 2010 Bentham Science Publishers Ltd.

1389-2037/10 $55.00 + 0.00
SMALL ANTIMICROBIAL PEPTIDES IN PLANT DEFENSE

Plant Defensins – PR-12

Plant defensins are small (45–54 amino acids) highly basic cysteine-rich peptides that form one of the most prevalent families of AMP found in several types of organisms, including vertebrates, invertebrates, and plants [16, 20, 23-26]. Originally isolated from cereals, plant defensins have been first called γ-thionins, because of their size and cysteine content similar to thionins [20]. However, subsequent structural analysis allowed classifying γ-thionins as plant defensins [11]. Plant defensins have a characteristic three-dimensional folding pattern that is stabilized by eight disulfide-linked cysteines and this pattern is currently used to discriminate between plant defensins, thionins and LTPs [23, 27].

The main characteristic of this PR-family member is its antimicrobial activity against a large number of phytopathogenic species, including fungi and bacteria [24, 28, 29]. Recently, plant defensins have been described in diverse plant families, such as Brassicaceae, Convolvulaceae, Fabaceae, Poaceae, Rosaceae and Solanaceae [30]. A defensin purified from Phaseolus vulgaris (PvD1) revealed high similarity to other Fabaceae species, such as V. unguiculata (93%), Cicer aritinum (95%) and Pachyrhizus erosus (87%), indicating a close relationship among defensins of this family. As many other defensins, this peptide showed a growth inhibition effect over a large variety of microorganism, such as, Candida albicans, C. parapsilosis, C. tropicalis, C. guilliermondii, Kluyveromyces marxianus, Saccharomyces cerevisiae, Fusarium oxysporum, F. solani, F. lateritium and Rizoctonia solani [31].

PR-12 peptides have also been isolated from a variety of tissues as leaves, tubers, flowers, pods and seeds [32] and their distribution inside the plant seems to be consistent with their putative defense response. In unstressed organs, AMPs are usually most abundant in the outer cell layer lining the organ, which is consistent with a role for the antimicrobial peptides in constitutive host defense against microbial invaders attacking from the outside [16]. These evidences indicate that a single defensin can play an important role in different plant stress responses and that defensins can be expressed in many plant tissues [30].

In addition to induction by pathogens, many defensin genes have been shown to be responsive to different signalizing molecules, such as ethylene [33], jasmonate [34] and salicylic acid [35], as well as abiotic stresses: cold [36], drought [35, 37], wounding, [35, 38], and to high NaCl levels [35]. Defensin genes have also been associated to heavy metal tolerance as in the case of ZnCl2 stress [39].

A clear differential expression level of defensins in several tissues was demonstrated by Bahramnej et al. analyzing eight Nicotiana defensin genes in different tissues in response to biotic and abiotic stresses by relative RT-qPCR. For example, NbDef 2.1 was the most expressed gene in all tissues tested (leaves, stem, roots, flower and developing seeds), while NbDef 2.1 and NbDef 2.2 were expressed only in flowers. Their results indicated that a small fraction of these genes was affected by a particular pathogen as well as by abiotic stress, e.g. NbDef 2.2 was strongly induced both by wounding and ethylene stresses [38].

In Arabidopsis thaliana, 15 defensin genes were reported, but only five had their expression levels investigated [24]. Additionally, Silverstein et al. showed that the A. thaliana genome contains 317 defensin-like sequences (DEFs) that have some homology to defensin sequences, but their specific expression profiles still need a careful analysis [40].

The correlation between defensin gene family and plant stress response mechanisms can also be reached through the observation of gene expression levels after a pathogen attack. For example, the inoculation of Pisum sativum (pea) with F. solani results in an increase of defensin expression [41]. Many studies in transgenic plants also indicated that plant defensins confer effective protection against pathogen attacks: Wang et al. observed that the expression of P. sativum defensin in Brassica napus increased Lepidospheara maculans resistance [42]. Similarly, B. oleracea defensin expressed in Oryza sativa increased resistance to Xanthomonas oryzae [43]. In the same way, the expression of a Raphanus sativus defensin in Nicotiana tabacum [11] and Lycopersicon esculentum [44] was responsible for an increased resistance against Alternaria longipes and A. solani, respectively.

Gao et al. identified a new plant defensin (alfalfa alfAPF gene) that inhibited Verticillium dahliae growth in vitro. This gene was introduced into potato plants and the transgenic plants displayed substantially increased levels of disease resistance against V. dahliae in greenhouse and field tests. A higher defensin content was observed in transgenic plant root and leaf tissues [45]. A wasabi defensin gene derived from a dicotyledonous plant (Wasabia japonica) was transferred by Agrobacterium-mediated transformation to monocotyledonous plants (O. sativa. cv. Sasanishiki), inducing increased expression in leaves, as evidenced by western blot analysis [46]; the transgenic rice exhibited enhanced resistance against the blast fungus (Magnaporthe grisea). The same defensin gene and the same transformation method were used to produce a transgenic Phalaenopsis orchid, which exhibited enhanced resistance against soft rot disease (Erwinia carotovora) [47]. Hanks et al. observed differentially expression of defensin genes after mycorrhizal root infection in Medicago truncatula. Mycorrhizal fungus Glomus versiforme induced defensin in roots two weeks post-inoculation. Another defensin gene was constitutively expressed in mock-inoculated as well as G. versiforme-infected roots [32].

Plant Thionins – PR-13

The thionin family includes peptides of 5 KDa, rich in basic and sulfur-containing residues (arginine, lysine and cysteine). The name thionin is used for two distinct but well characterized groups of plant peptides, α/β-thionins and γ-thionins, the latter being presently recognized as defensins. Traditionally, α/β-thionins were subdivided into five different classes (I, II, III, IV, and V) [48]. Detailed information on the structural characteristics of thionins can be found in Stec [49]. Thionins differ from defensin and LTP by their predominant location in the intracellular space [15, 16]. As other AMP described in this review, thionins also present a
toxic effects against bacteria, fungi and yeast such as *Pseudomonas, Xanthomonas, Erwinia, Corynebacterium, Thielaviopsis paradoxa* and *Drechslera teres* [12], although the direct mechanism of action is still controversial. They are also expressed in different tissues in distinct plant species, including monocots and eudicots [reviewed by 15 and 49]. Proteins similar to thionins were isolated from leaves and endosperms in barley, oat and rye. Evaluation of the genera *Triticum* and *Aegilops* brought evidences of thionins in the endosperm. In *Vicia faba* three proteins named viscothionins were found with structural similarities to thionins/purothionins. Proteins with similar properties have been isolated from other *Vascaea* family species, such as phorothionins A and B, denclatoxin B and ligatoxins A and B.

Crambins, with high hydrophobicity and zero net charge, are remarkable among thionins because of their apparent lack of toxic activity [15, 48]. The *Pyrularia* thionin exhibits some features similar to crambin (a block of eight consecutively amino acid residues) and cereal thionins (four disulfide bridges). However, the *Pyrularia* thionin is most similar to the viscothionin thionin group. Altogether, these results indicated how diverse the thionin gene family can be [15, 48].

Several in vivo functions have been proposed for thionins based mainly on in vitro observations, such as regulatory proteins, thiol secondary messengers in the redox regulation of enzymes, as storage proteins, especially as source of sulfur and as generalized defense effectors [15].

Similarly to defensins, there are also several reports on increased resistance to fungal or bacterial pathogen attacks from transgenic plants expressing a thionin gene. For example, high-level expression of a hordothionin gene from barley in transgenic tobacco conferred resistance to *Pseudomonas syringae* [50]. A gene coding a thionin (Th2.1) when overexpressed in transgenic Arabidopsis plants, conferred resistance to *F. oxysporum* f. sp. *matthiolae*. These data clearly indicated that the overexpression of a single thionin unbalances the plant-pathogen relationship, giving more defense artifacts to plants [13]. This Th2.1 thionin gene was also efficient in transgenic tomato against the fungus *F. oxysporum* f. sp. *lycopersici* and against the bacterium *Ralstonia solanacearum* [51]. In addition, enhanced resistance to bacterial diseases was described in transgenic rice plants overexpressing an oat thionin gene [52].

**Plant Lipid Transfer Proteins – PR-14**

Plant Lipid Transfer Protein - LTP is another protein family involved in plant stress response [53-57]. The ability to transfer phospholipids between a donor and an acceptor membrane is the main property of LTPs found in plants. However, other lipids can also be transported and when the activity is not specific, the peptides are called non-specific lipid transfer protein [58]. The first plant LTP was discovered in spinach, in 1984 [58]. Since then, many other LTPs were identified in different plant species and were subsequently divided into two families of small peptides (LTP1 and LTP2). A concise review on these families can be found in Carvalho et al. [59] Although LTPs sequences of different plant species revealed significant divergence, a domain consisting of eight cysteine residues according to the pattern 2/3-C-8-C-12/15-CC-19-C-1-C-21/23-C-13-C-4/8 with four disulfide bonds, was found to be conserved [58-60], which promote the peptide tertiary structure stabilization. The LTP1 family members, with 90 – 95 amino acid residues and molecular masses around 9 kDa, seem to have a broader host pathogen interaction. The LTP2 family members, formed by peptides with molecular masses of approximately 7 kDa and around 70 amino acids, present a different pattern of disulfide bridges. Both families have a signal-peptide, 21 to 27 amino acids long in LTP1 family and 27 to 35 in the LTP2 family [59, 61], helping the targeting of these peptides to the apoplast through the cell secretory pathway [59]. The extracellular location of LTP1 at the cell wall prompted the suggestion of some possible functional activities for this protein family in defense against pathogens [62], cuticle synthesis [62, 63], pollen adherence [59] and in long-distance systemic signaling [57].

Wang et al. reported a non-specific lipid transfer peptide (nsLTP) isolated from mung bean (*Phaseolus mungo*) seeds which exhibited an antimicrobial activity against *F. solani*, *F. oxysporum*, *Pythium aphanidermatum*, *Sclerotium rolfsii* and *Staphylococcus aureus* [64].

Carvalho et al. investigated the VULTP (*V. unguiculata* lipid transfer protein) gene expression profile from cowpea (*V. unguiculata*) in different tissues of adult plants under normal physiological conditions and under biotic and abiotic stresses (by fungal infection and cold) in seedlings tissues. According to this study, LTP transcripts did not accumulate in adult tissues and VULTP gene expression seemed to be restricted to developing seeds and seedling tissues [65].

Based on the results indicating LTPs involvement in plant resistance responses [53-55, 58, 59,61], some efforts were recently attempted to increase plant pathogen defenses by plant transformation with LTP genes [56-58] and significant results were achieved in different plant species such as carrot [66], rice [67] and *Arabidopsis* [56]. In all cases, the transgenic plants presented increased pathogen resistance. In tobacco plants transformed with a lipid transfer protein from barley (LTP2), *P. syringae pv. tabaci* growth was retarded in comparison with the non-transformed control plants. The percentage of infection points evolving to necrotic lesions was reduced to 38%. The average size of these lesions was also reduced to 61 - 81% [66, 68].

However, LTP rank among the most common plant allergens. Their use in the generation of new transgenic plants resistant to pathogens or abiotic stresses request, therefore, a precautionary evaluation [69]. Nevertheless, a deeper understanding on LTP regulatory pathways and their mechanism of action in plants may unravel important clues on similar mechanism in animals and on host-parasite relationships in vertebrates.

**SUPERSAGE: DIGITAL GENE EXPRESSION**

The identification of new AMPs groups and families is hindered by the small size of the proteins, their high sequence diversity and the lack of a large set of conveniently annotated sequences. However, numerous plant genome and transcriptome sequencing projects for different proposals have been launched, providing a singular opportunity for assignment and eventual classification of these important antimicrobial peptides. SuperSAGE [70], the optimization of
conventional SAGE - Serial Analysis of Gene Expression [71], is the best technology presently available for transcriptional profiling [72], in digital sequence analysis assays [73]. New sequencing technologies [the Genome Analyzer (Solexa/Illumina), 454 (Roche), and ABI-SOLiD (Applied Biosystems)] can readily supply a large number of sequenced tags, which can be counted and compared to database sequences. The tag size in this case (26 bp) is longer than the original SAGE tag (14 pb) or even than LongSAGE tag (18 - 20 bp). This features provides SuperSAGE technology some advantages, such as better tag-to-gene annotation, simultaneous analysis of two interacting eukaryotic organisms (e.g. a host and a pathogen), full-length cDNAs amplification using tags as primers, potential use of tags via RNA interference (RNAi) in gene function studies, or even tags directly spotted onto DNA chips for high-throughput applications [74].

By applying SuperSAGE to _M. grisea_ (blast)-infected rice leaves, gene expression profiles of both the rice host and blast fungus were simultaneously monitored [70, 75]. The same group developed a SuperSAGE array combining the quantitative expression analysis with high-throughput microarray technology for 1,000 tags in rice. The results showed to be very reproducible. The same methodology was also applied with success to _Nicotiana benthamiana_, an organism for which sufficient genome sequence information is not available [74]. Trying to understand the transcriptome changes during the early resistance response in pepper plants (_Capsicum chinense_) against a tobanovirus (_Pepper mild mottle virus_ - PMMoV), SuperSAGE was applied and some genes identified to be up- or downregulated [76]. SuperSAGE was also applied to develop the most comprehensive analysis of the chickpea drought-response transcriptome available to date [77]. The authors demonstrated that – inter alias – signal transduction, transcription regulation, osmolyte accumulation, and ROS scavenging undergo strong transcriptional remodelling in chickpea roots already 6 h after drought stress. Some characterized transcript isoforms could be potential targets for genetic breeding in a non-model crop like chickpea. Mitsuya _et al_. applied SuperSAGE analysis to understand the polyamine spermine (Spm) functions as a signaling molecule to evoke defense reactions in avirulent pathogen-attacked tobacco plants. The identified Spm-responsive genes support the existence of a polyamine spermine signaling pathway in _A. thaliana_. Close to 90% of the Spm-responsive genes were also involved in cucumber mosaic virus (CMV)-elicited hypersensitive response. Spm modulated the expression of genes of redox components, and genes involved in protein folding and secretion, protein degradation and defense [78]. Genome-wide gene expression profiling like those described above could be the basis for further identification of new PR peptides and contribute to unravel the still largely unknown biological role of those peptides [70].

**SUPERSAGE: THE BRAZILIAN EXPERIENCE LOOKING FOR IN SILICO PLANTS AMP TAGS**

In the search for possible AMPs transcripts we looked for SuperSAGE tags (26 bp) from different tissues (leaves or roots) and plants (cowpea (*V. unguiculata* (L.) Walp.), soybean (*Glycine max* (L.) Merr.) and modern hybrids sugarcane (*Saccharum* spp.) submitted or not (controls) to different stress conditions (salt (cowpea: 100 mM NaCl; bulk of three time points - 30, 60 and 120 min); drought (sugarcane: 24 h after water suppression on soil (40 kg.m⁻²¹); bulks of tolerant or sensible genotypes; soybean: root dehydration (bulk of five time points (last time point: 150 min) after plant suspension of hydroponic solution); soybean: *Phakopsora pachyrhizi* inoculation (Asiatic Soybean Rust pathogen): bulk of three time points after inoculation (12, 24, 48 h)). The unique and total tags available by experiment are showed in Table 1. These libraries were generated by the UFPE (Federal University of Pernambuco, Brazil) group in collaboration with other Brazilian partners (soybean: Embrapa Soybean, Londrina, PR; sugarcane: Center of Sugarcane Technology - CTC, Piracicaba, SP; cowpea: Northeast Biotecnology Network - RENORBIO) and GenXPro GmbH (Frankfurt, Germany).

Initially, defenses (PR-12), thionins (PR-13) and LTPs (PR-14) sequences were searched by keywords in nucleotide and EST NCBI databases [79]. These searches were restricted to Viridiplantae and all results were downloaded (FASTA format), except for chromosome data. The number of sequence retrieved from the database are displayed in Table 2. The FASTA files were used locally in BlastN analysis [80] with the unique SuperSAGE tags extracted from the libraries. The number of tags with good alignments (score ≥ 50, maximum of one mismatch) by PR class, after exclusion of all unknown hits, is showed in Table 3. All libraries were normalized (10⁶ tags, library¹) and the Fold Change (FC) of one specific tag was calculated considering the relation between the tag frequencies in a stressed library in relation to the unstressed control. The differentially expressed tags (p < 0.05) were identified using the software DiscoverySpace 4.0 [81]. Some results are summarized below.

**Plant Defensin Tags**

Three tags annotated as defensins were identified in root cowpea libraries under salt stress (Table 4). One of them (CATGTGTAGTGTGACACCAACTGTC) displayed 100% identity with part of a *V. unguiculata* defensin precursor sequence (gi|225548305). The _in silico_ translation of this sequence using the NCBI ORF Finder tool [82] allowed to identify the tag as part of the second exon, in a region around the second and third cysteine residues (Fig. 1A). The sequence had all the characteristics of a mature plant defensin peptide, including the eight conserved cysteine residues (3-C-10-1-C-5-C-3-C-9-C-8-C-1-C-3-C) [83] supposed to form the same bonds displayed in other plant defensins, as predicted by on line DISULFIND tool [84]. Other conserved residues typical of almost all plants defensins, such as a serine at position 7, an aromatic residue at position 10, two glycines at position 12 and 32 and a glutamic acid at position 27 [20, 85] are also showed. The same tag was observed in all libraries, its frequency ranging from 73 to 465 considering 10³ tags library¹. Moreover, it was upregulated (p < 0.05) after two hours of salt exposition, for both genotypes (salt-tolerant and salt-susceptible), and the FC increased ca. three times, considering the salt tolerant genotype under stress compared to the same unstressed genotype, and five times, considering the salt susceptible genotype in relation to the unstressed control (Table 4). The same tag and upregulated
Table 1. Unique and Total Tags (26bp) Extracted in Plant SuperSAGE Experiments

<table>
<thead>
<tr>
<th>Species / Tissue</th>
<th>Stress</th>
<th>Phenotypes</th>
<th>Number of Libraries</th>
<th>Unique Tags</th>
<th>Total Number of Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigna unguiculata / root</td>
<td>Salt (100 mM NaCl)</td>
<td>Tolerant Sensible</td>
<td>04</td>
<td>133,964</td>
<td>4,726,458</td>
</tr>
<tr>
<td>Glycine max / root</td>
<td>Drought (0 – 150 min)</td>
<td>Tolerant Sensible</td>
<td>04</td>
<td>449,964</td>
<td>2,540,271</td>
</tr>
<tr>
<td>Glycine max / leaves</td>
<td>Phakopsora pachyrhizi</td>
<td>Tolerant</td>
<td>02</td>
<td>422,945</td>
<td>1,961,313</td>
</tr>
<tr>
<td>Saccharum spp. / root</td>
<td>Drought (0 – 24h)</td>
<td>Tolerant Sensible (bulks)</td>
<td>04</td>
<td>288,404</td>
<td>8,787,313</td>
</tr>
</tbody>
</table>

Table 2. NCBI Search Results (by Keywords) in Nucleotide and EST Databases Considering the Keyword in the Annotation/Description of Sequence (A) or Unknown/Others Descriptions (N)

<table>
<thead>
<tr>
<th>Keyword</th>
<th>EST</th>
<th>Nucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A / N</td>
<td>Total</td>
</tr>
<tr>
<td>defensin</td>
<td>292/274</td>
<td>566</td>
</tr>
<tr>
<td>thionin*</td>
<td>45/37</td>
<td>70</td>
</tr>
<tr>
<td>lipid transfer protein</td>
<td>2,040/396</td>
<td>2,436</td>
</tr>
</tbody>
</table>

* including word-derivatives (like purothionin), crambin and phoratoxin.

Table 3. Total of Tags Aligning (BlastN; score ≥ 50) to Defensins, Thionins* and Lipid Transfer Proteins (LTP) in SuperSAGE Experiments (Species / Tissue and Stress)

<table>
<thead>
<tr>
<th>Species / tissue</th>
<th>Stress</th>
<th>Defensins</th>
<th>Thionins*</th>
<th>LTPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigna unguiculata / root</td>
<td>Salt (100 mM NaCl)</td>
<td>3</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Glycine max / root</td>
<td>Drought (0 – 150 min)</td>
<td>5</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Glycine max / leaves</td>
<td>Phakopsora pachyrhizi</td>
<td>0</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Saccharum spp. / root</td>
<td>Drought (0 – 24h)</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

* including word-derivatives (like purothionin), crambin and phoratoxin.

Table 4. SuperSAGE Tags, Their Possible Defensin Annotation (GI Number), Frequencies in Normalized Libraries (10⁴ Tags) and Tags Expression (E; p = 0.05) in Stressed (S) Versus Control Libraries (C)

<table>
<thead>
<tr>
<th>Tag (26 bp) (gi; number, description)</th>
<th>Crop / Stress</th>
<th>S¹</th>
<th>C²</th>
<th>E</th>
<th>S³</th>
<th>C⁴</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATGACGACGACGACCTGCGGCAACG (gi170522416; defensin precursor mRNA)</td>
<td>Sugarcane / drought</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>CN</td>
</tr>
<tr>
<td>CATGAAATTATTCCTTGGTGTTGC (gi24205345; defensin precursor mRNA)</td>
<td>Soybean / drought</td>
<td>0</td>
<td>17</td>
<td>DR</td>
<td>4</td>
<td>13</td>
<td>CN</td>
</tr>
<tr>
<td>CATGCTTCTGCACCAACATTTGTTAAG (gi226772163; defensin protein 1 mRNA)</td>
<td>Soybean / drought</td>
<td>0</td>
<td>2</td>
<td>CN</td>
<td>0</td>
<td>1</td>
<td>CN</td>
</tr>
<tr>
<td>CATGTTTGAGTGACACAACTGTTGGC (gi533691; protease inhibitor mRNA)</td>
<td>Soybean / drought</td>
<td>81</td>
<td>101</td>
<td>CN</td>
<td>45</td>
<td>108</td>
<td>DR</td>
</tr>
<tr>
<td>CATGTTGAGTGACACAACTGTCTGCTG (gi225254835; defensin precursor gene)</td>
<td>Cowpea / salt</td>
<td>465</td>
<td>153</td>
<td>UR</td>
<td>386</td>
<td>73</td>
<td>UR</td>
</tr>
<tr>
<td>CATGTTGAGTGACACAACTGTTGGTG (gi225254835; defensin precursor gene)</td>
<td>Cowpea / salt</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>2</td>
<td>CN</td>
</tr>
<tr>
<td>CATGCTTCTGCTTTGGTCT (gi225253895; defensin precursor gene)</td>
<td>Cowpea / salt</td>
<td>1</td>
<td>2</td>
<td>CN</td>
<td>4</td>
<td>0</td>
<td>UR</td>
</tr>
</tbody>
</table>

* Libraries with bulk of genotypes; salt: 100 mM NaCl; S¹: tolerant genotype(s) under stress; C²: tolerant genotype(s) without stress; S³: susceptible genotype(s) under stress; C⁴: susceptible genotype(s) without stress; UR: upregulated; CN: constitutive; DR: downregulated; base substitution (bold).
expression pattern were observed in cowpea leaves after mechanical injury with abrasive material (data not shown). Another tag was only counted twice in one of the control libraries and showed a single mismatch in relation to the previous tag (Table 4). We cannot discard the possibility of a sequence error rather than a SNP. Anyway, the ORF analysis revealed that this nucleotide substitution was synonymous. A third tag showed a mismatch just over the first nucleotide and the annotation could be wrong. The recognition of the initial CATG in a similar sequence is an important step in the tag annotation.

Plant defensins may be highly expressed or upregulated in plants challenged by various abiotic stresses, including salinity, but their potential role in abiotic stress tolerance is presently unknown. Some authors [39] suggested defensins could play a role in metal tolerance (more specifically through a Zn tolerance, as demonstrated in Arabidopsis halleri), more than in osmotic stress tolerance, although the induction of plant defensin gene expression by salt stress has been reported. For instance, the defensin CADEF1 gene in pepper plants was differentially activated in response to high salinity, drought and wounding stresses, as well as after Xanthomonas campestris pv. vesicatoria infection [35]. Moreover, PDF1.2, a defensin gene, is highly and constitutively expressed in the halophyte salt cress Thellungiella halophila [86].

Surprisingly, the BlastN analysis of soybean leaves tags after P. pachyrhizi inoculation returned no sequences showing an acceptable similarity to defensin-like genes. Only a single hit was found similar to a protein of unknown function but bearing a γ-thionin domain. On the other hand, three tags (two barely expressed) matched with protease inhibitor mRNA (gi|533691) from our thionin database (see discussion below), and its corresponding protein (gi|533692) showed eight conserved cysteine residues (3-C-3-9-C-6-C-1-C-3-C). Plant defensins with enzyme inhibitory activity against insect protease inhibitors or α-amylase, as well their use in transgenic plants are reviewed by [29]. The most expressed defensin tag (CATGT TTGAGTGCAC CAACTTGGGC) was observed in the mock library (48 versus 32 tags in the normalized infected library). The results suggest that this defensin does not play a major role in this plant-pathogen interaction, although the analyzed genotypes (plant and pathogen) may have been inadequate or the defensin may not be targeted to this pathogen. Indeed, the expression of a sugarcane defensin-like gene (gi|170522417) in a modern sugarcane hybrid, as observed by RTqPCR, was not induced by infection of the plant with Fusarium and Trichoderma, when compared to the non-infected control [85]. Another possible explanation for the apparent unresponsiveness of plant defensin genes to infection is the time points of sample collection used in our experiments. RTqPCR and microarray results with canola (Brassica napus L.) inoculated with the fungal pathogen Sclerotinia sclerotiorum showed enhanced expression of the PDF1.1 defensin gene (almost 6-fold comparing with actin gene) starting only 48 h after inoculation and sharply decreasing after 72 h [87].

On the other hand, five tags were identified in soybean roots submitted to drought stress. Four of them showed a perfect alignment (100% identity) with a defensin precursor gene (gi|24205345; similar to Uniprot TR: O65740). Other five tags showed excellent alignments (score ≥ 50) with sequences from our thionin database, but all the corresponding NCBI sequences were actually defensin-like proteins, in spite of having other putative functions annotated (see discussion below). The relative higher number of defensin tags compared to those described before could be a consequence of the experimental design. The drought stress was obtained with plant dehydration in air (until 150 minutes). This methodology was used before [88] and the gene expression results were comparable with those observed with plants in a pot-based system and greenhouse conditions. Besides these five tags, no other was similar to defensin genes. However, two particular tags, (CATGAAATTTATCTCT GTTG GTTGC and ATGTTTGAATGACACCCACTGTC GC), observed in both genotypes (tolerant and susceptible to water deficit) had their frequencies decreased at a significant level (5%) after stress exposition (Table 4). The first tag matches over an untranslated region (3'-UTR), and not in the ORF, as observed to the second tag, matching over the coding region to the second and third cysteines, as also observed for the tags displayed in the Fig. (1). The results suggest that both genes are downregulated, although in different genotypes, as shown in Table 4.

In relation to the drought sugarcane libraries, three tags aligned with annotated defensin genes. One of them (CATGAGCA GCAGCACTGCACCCAACG) had a perfect match with Saccharum officinarum defensin precursor (PDEF) mRNA (gi|170522416), but was very seldom transcribed and only in the bulk of drought susceptible genotypes (Table 4). This could be a consequence of the experimental design, aimed primarily at the early sugarcane response to drought stress; accordingly, RNA samples were collected after 24 h of irrigation suppression (bulk of tolerant genotypes), after three months in greenhouse condition, in pots of 40 kg soil. Similarly to the cowpea defensin precursor protein this tag was also mapped in the ORF region (second exon) (Fig. 1B). This sugarcane defensin-like gene, as indicated by preliminary RTqPCR functional results, showed a lack of expression in the Brazilian sugarcane hybrid specifically infected with Fusarium and Trichoderma, in relation to the non-infected control [85].

Plant Thionin Tags

Unfortunately, no reliable data resulted for this PR class. Besides the alignment results presented in Table 3, no thionin were confirmed after an accurate analysis. Two causes could be envisaged: a) all the six hits (score ≥ 50) were assigned to unknown function (like gi|255640773) or other functions [PR-6 type like Bowman-Birk Protease Inhibitor mRNA (gi|18747), protease inhibitor mRNA (gi|533691), low-molecular-weight cysteine-rich protein LCR70 precursor (gi|19568755), etc.]; after a careful evaluation, all of them proved to be γ-thionin (defensins) or at least presented this domain; b) the small number of thionin nucleotide sequences (most of them are deposited in databases as amino acid sequences) correctly annotated and available in the NCBI databases, c) the low number of reliable thionin bearing SuperSAGE tags in their ORFs or even
in the most variable 3'UTRs. The thionin keyword search retrieves, as expected, sequences like puro-thionin, hordo-thionin and hellethionin but γ-thionin, the first name of plant defense. No BlastN results also were found with crambin, phoratoxin or viscotoxin sequences.

Plant Lipid Transfer Protein Tags

Nine tags from cowpea salt stress libraries had excellent alignments with LTP genes (eight with 100 % identity) and some of them are shown in Table 5. The tag CATGTGTAATCCAGGGA (194 – 418 tags by normalized library). In both genotypes (tolerant and susceptible), the tag frequency increased after salt stress exposure, but only in the tolerant genotype it was upregulated (p = 1.02 E-12; Table 5). In the same way, two potential SNPs were also upregulated in the tolerant genotype after salt stress exposure in relation to the unstressed control. One of them was also upregulated in the salt-stress susceptible genotype when compared with the same genotype in the unstressed condition. However, the observed polymorphism needs to be further analyzed do discard an eventual sequence error. A pepper LTP from Capsicum annuum (CALTP II) was specifically induced by high salinity (100 mM NaCl), although other abiotic conditions were also studied [61].

Eighteen tags identified in the soybean leaves under P. pachyrhizi inoculation were associated by BlastN to LTPs or related genes. A specific tag (CATGTGTGATCCAGGGA ATTATATGT), well represented in both libraries, showed a 4-fold decrease in the infected library versus the unstressed control (Table 5), suggesting downregulation (p = 7.75 E-21) in the resistant genotype after infection. Although some LTPs could increase fungal resistance [67], the diverse LTP functions and specificity could modify the expression pattern.

Considering the soybean drought stress, 25 SuperSAGE tags could be annotated as LTPs or related proteins. Some of them are also displayed in Table 5. The most frequent LTP-related tag (CATGTGTAATCCAGGGA) seemed to be constitutively expressed [no significant difference at the studied level (p < 0.05)] in the root system of susceptible genotype under drought condition in relation to the same genotype under unstressed condition. The same tag seemed to be upregulated (p < 3.59 E-10) in the drought tolerant genotype, as compared to its unstressed control (Table 5).

Other differentially expressed tags are also displayed in Table 5, including identical tags identified in different species or tissues submitted to stress or unstressed condition. Different expression patterns were also observed with LTPs (CALTP I and CALTP III) genes from pepper (Capsicum annuum); both genes were expressed in different infected tissues and were also induced by drought, high salinity, low temperature and wounding stresses [61].

On the other hand, only four sugarcane tags were associated to LTPs. All of them were seldom transcribed (only 1-2 copies by normalized libraries). If we compare with those identified in legume libraries, the number is almost inexpressive. One explanation could be the fact that LTPs expression in root is quite rare and they are frequently detected only in young tissues [63]. Moreover, the root sampling in sugarcane does not favor the young meristemic cells, as root tips tend to break and remain in the soil. The hydroponic system used in the drought soybean experiment, on the contrary, enable a better sampling of young root tips. Another aspect related to the marked difference in LTP tag frequency in sugarcane could be the presence of LTP-like proteins as
Table 5. SuperSAGE Tags, their Possible Lipid Transfer Protein (LTP) Annotation, Alignments in ORF or UTR of the Protein Sequence (GI), Frequencies in Normalized Libraries (10^5 Tags) and Tags Expression (E; p = 0.05) in Stressed (S) Versus Control Libraries (C)

<table>
<thead>
<tr>
<th>SuperSAGE Tag (26 bp) (gi number; description; alignment)</th>
<th>Crop / Stress</th>
<th>GI</th>
<th>C^2</th>
<th>E</th>
<th>GI</th>
<th>C^2</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATGTCATTCCATTTCCGTTTCCGTTG (gi170743846; LTP; 3'UTR)</td>
<td>Cowpea / salt</td>
<td>335</td>
<td>194</td>
<td>UR</td>
<td>418</td>
<td>436</td>
<td>CN</td>
</tr>
<tr>
<td></td>
<td>Soybean / ASR</td>
<td>2</td>
<td>4</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>291</td>
<td>336</td>
<td>CN</td>
<td>456</td>
<td>234</td>
<td>UR</td>
</tr>
<tr>
<td>CATGGAATCCCTGTTAAGTGCAAGG (gi1420886; nsLTP; ORF)</td>
<td>Cowpea / salt</td>
<td>2</td>
<td>0</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / ASR</td>
<td>0</td>
<td>1</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>73</td>
<td>115</td>
<td>DR</td>
<td>32</td>
<td>185</td>
<td>DR</td>
</tr>
<tr>
<td>CATGGTGTCGTGCGCAGAGGACAGAGGA (gi10237890; nsLTP; 3'UTR)</td>
<td>Cowpea / salt</td>
<td>0</td>
<td>0</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>6</td>
<td>17</td>
<td>CN</td>
<td>18</td>
<td>12</td>
<td>CN</td>
</tr>
<tr>
<td>CATGCTCTGGCATAACTCCAAGAAC (gi10848236; nsLTP; ORF)</td>
<td>Soybean / ASR</td>
<td>1</td>
<td>0</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>20</td>
<td>36</td>
<td>CN</td>
<td>0</td>
<td>44</td>
<td>UR</td>
</tr>
<tr>
<td>CATGTTTGCTCTACCTCCTCTCAG (gi23729321; nsLTP; 3'UTR)</td>
<td>Soybean / ASR</td>
<td>2</td>
<td>0</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>47</td>
<td>66</td>
<td>CN</td>
<td>4</td>
<td>57</td>
<td>DR</td>
</tr>
<tr>
<td>CATGCCCTGTGTGTTACCTCAAAAC (gi4395670; LTP-like protein; ORF)</td>
<td>Soybean / ASR</td>
<td>2</td>
<td>0</td>
<td>CN</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>12</td>
<td>20</td>
<td>CN</td>
<td>4</td>
<td>47</td>
<td>DR</td>
</tr>
<tr>
<td>CATGTTTGATCCAGGAGATGATTAGA (gi19270333; LTP; 3'UTR)</td>
<td>Soybean / ASR</td>
<td>0</td>
<td>14</td>
<td>DR</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / ASR</td>
<td>715</td>
<td>2735</td>
<td>DR</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Soybean / drought</td>
<td>47</td>
<td>64</td>
<td>CN</td>
<td>25</td>
<td>61</td>
<td>DR</td>
</tr>
<tr>
<td>CATGTTGAATGTTATTTTCTCCTAA (gi110458369; LTP; 3'UTR)</td>
<td>Cowpea / salt</td>
<td>8</td>
<td>0</td>
<td>UR</td>
<td>1</td>
<td>2</td>
<td>CN</td>
</tr>
<tr>
<td></td>
<td>Cowpea / salt</td>
<td>50</td>
<td>0</td>
<td>UR</td>
<td>13</td>
<td>0</td>
<td>UR</td>
</tr>
</tbody>
</table>

3'UTR: 3'Untranslated Region; ORF: open read frame; salt: 100 mM NaCl; ASR: Asian Soybean Rust; S^1: tolerant genotype(s) under stress; S^2: tolerant genotype(s) without stress; S^3: susceptible genotype(s) under stress; C^1: susceptible genotype(s) without stress; UR: upregulated; CN: constitutive; DR: downregulated; base substitution (bold).

consequence of root nodules. Some of the antimicrobe-like polypeptides may not be acting as antimicrobial factors but as signals to mediate nodule formation and development [89]. Indeed, the different expression regulation patterns could be the result of a variable range of possible LTP targets and functions involved in different metabolic pathways.

Besides the SuperSAGE tags mentioned before, a total of 298 new tags from different libraries (cowpea/roots: 18; soybean/leaves: 122; soybean/roots: 135; sugarcane/roots: 23) showed excellent BlastN alignments with sequences present in the LTP FASTA file used here. The elevated tag frequencies, as well as their excellent matching to LTP gene sequences could be a first indication of the LTP relevance in stress response.

CONCLUDING REMARKS

The antimicrobial activity against a variable range of microorganisms highlights PR proteins as important candidates for the development and production of novel drugs, as well as for the improvement of resistance against phytopathogens in crops. This aspect could result in the future release of agronomically important crops resistant to various diseases of economic importance. SuperSAGE technology opens the opportunity to find new PR genes, possible SNPs or even rare transcripts that could be important on plant stress resistance mechanisms. The present results identified tags annotated as defensins, LTPs or relatives, some of them bearing a constitutive gene expression in different tissues or plant species, after biotic or abiotic stress or even unstressed condition, while others were differentially expressed after stress. These results are being now validated by RT-qPCR. Besides potential identification of specific genes, a SuperSAGE transcriptome profile could provide a first basis to further unravel the still largely unknown biological role of AMP peptides, particularly the signaling pathway and the mechanisms of gene regulation and its various aspects such as receptors, signal transducing cascades and molecular targets. Also, a great number of newly un-annotated PDF-like, thionin-like and LTP-like sequences have been found in many plant species and need to be carefully studied. Defensins or LTPs potentially identified by SuperSAGE following a
specific plant treatment or physiological condition could be useful for future use in programs of genetic improvement of plants.

ACKNOWLEDGEMENTS
The authors thank Prof. Dr. Günter Kahl (University of Frankfurt, Germany) and Dr. Peter Winter and Dr. Björn Rotter (GenXPtro GmbH) for scientific and technical advice. This work has been supported by the Banco do Nordeste do Brasil - BNBR / Science and Technology Support Foundation of Pernambuco State (FACEPE), The Brazilian Innovation Agency (FINEP 01.08.0023.00 and 01.06.0290.00), Brazilian National Council for Scientific and Technological Development (CNPq 552735/2007-8) and Brazilian Ministry of Science and Technology.

ABBREVIATIONS
alFP  =  Alkalalfi antifungal peptide defensin
AMPs  =  Antimicrobial peptides
ASR  =  Asiatic soybean rust
BlastN  =  Basic Local Alignment Search Tool
DEFLs  =  Defensin-like sequences
EST  =  Expressed Sequence Tag
HPLC  =  High-performance liquid chromatography
kDa  =  Kilodalton
LTPs  =  Lipid transfer proteins
NBD  =  Nicotiana benthamiana defensin
nsLTP  =  Non-specific lipid transfer peptide
ORF  =  Open reading frame
PR  =  Pathogenesis-related
PvD  =  Phaseolus vulgaris defensin
RNAs  =  RNA interference
RT-PCR  =  Reverse transcription-polymerase chain reaction
RT-qPCR  =  Real time quantitative polymerase chain reaction
SAGE  =  Serial Analysis of Gene Expression
SAR  =  Systemic acquired resistance
SNP  =  Single nucleotide polymorphism
Thi2.1  =  Thionin gene
UTR  =  Untranslated region
VULTP  =  Vigna unguiculata lipid transfer proteins

REFERENCES


