

In silico differential display of defense-related expressed sequence tags from sugarcane tissues infected with diazotrophic endophytes

Marcio R. Lambais

Abstract

The expression patterns of 277 sugarcane expressed sequence tags (EST)-contigs encoding putative defense-related (DR) proteins were evaluated using the Sugarcane EST database. The DR proteins evaluated included chitinases, β -1,3-glucanases, phenylalanine ammonia-lyases, chalcone synthases, chalcone isomerases, isoflavone reductases, hydroxyproline-rich glycoproteins, proline-rich glycoproteins, peroxidases, catalases, superoxide dismutases, WRKY-like transcription factors and proteins involved in cell death control. Putative sugarcane WRKY proteins were compared and their phylogenetic relationships determined. A hierarchical clustering approach was used to identify DR ESTs with similar expression profiles in representative cDNA libraries. To identify DR ESTs differentially expressed in sugarcane tissues infected with *Gluconacetobacter diazotrophicus* or *Herbaspirillum rubrisubalbicans*, 179 putative DR EST-contigs expressed in non-infected tissues (leaves and roots) and/or infected tissues were selected and arrayed by similarity of their expression profiles. Changes in the expression levels of 124 putative DR EST-contigs, expressed in non-infected tissues, were evaluated in infected tissues. Approximately 42% of these EST-contigs showed no expression in infected tissues, whereas 15% and 3% showed more than 2-fold suppression in tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*, respectively. Approximately 14 and 8% of the DR EST-contigs evaluated showed more than 2-fold induction in tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*, respectively. The differential expression of clusters of DR genes may be important in the establishment of a compatible interaction between sugarcane and diazotrophic endophytes. It is suggested that the hierarchical clustering approach can be used on a genome-wide scale to identify genes likely involved in controlling plant-microorganism interactions.

INTRODUCTION

Plants can interact specifically with microorganisms forming mutualistic or pathogenic associations. During the development of such associations, the plant defense system is strictly regulated, determining whether the interaction will be successful. It has been proposed that during plant-microorganism interactions the plant defense system is regulated by complex signaling cross-talking and transduction pathways (Eugem *et al.*, 1999; McDowell and Dangl, 2000). In arbuscular mycorrhizae, as well as in legume-*Rhizobium* mutualistic symbioses, plant DR genes are transiently induced during the early stages of the interactions and suppressed later (Ruiz-Lozano *et al.*, 1999; Lambais, 2000). Attenuated suppression, and/or localized induction of DR genes has been observed in mycorrhizal roots grown at conditions unfavorable to the development of the symbioses, e.g. high phosphate concentrations in the soil (Lambais and Mehdy, 1998; Lambais, 2000). Activation of the plant defense system has also been observed in *Medicago sativa* upon infection with an exopolysaccharide (EPS I)-deficient mutant of *Sinorhizobium*, suggesting that exopolysaccharides may act as suppressors of the plant defense system (Niehaus *et al.*, 1993; Gonzalez *et al.*, 1996).

In plant-pathogen interactions, the activation of the plant defense system is dependent on the recognition of avirulence determinants by plant receptors. Upon recognition, specific signal transduction pathways are triggered, resulting in the production of reactive oxygen species, accumulation of pathogenesis-related proteins and phytoalexins, and localized cell death (McDowell and Dangl, 2000). Signal molecules generated at the infection site may induce defense responses systemically. Additionally, it has been shown that the recognition of several microbial signals is mediated by transmembrane receptor-like kinases and induces a cascade of phosphorylation/dephosphorylation events, resulting ultimately in the activation/repression of genes encoding several DNA-binding proteins. These transcription factors recognize specific DNA sequences in the promoter region of DR genes and regulate their expression (McDowell and Dangl, 2000).

Recently, a new class of zinc finger transcription factors containing the consensus sequence WRKYGQK has been identified (Eugem *et al.*, 2000). These proteins bind to W-boxes (TTGAC) of several genes encoding pathogenesis-related proteins, activating their transcription. It has been suggested that WRKY-like proteins may be involved in wounding as well as elicitor activation of plant

DR genes (Eulgem *et al.*, 1999; Yang *et al.*, 1999; Dellagi *et al.*, 2000; Hara *et al.*, 2000). WRKY proteins may contain one or two copies of the conserved WRKYGQK and zinc finger (C-X₄-C-X₂₂₋₂₃-H-X₁-H) domains. WRKY proteins may also have protein kinase C phosphorylation and nuclear localization sites (Eulgem *et al.* 2000; Dellagi *et al.*, 2000). The high number of WRKY-like proteins in plants, up to 100 in *Arabidopsis thaliana*, suggests that they might be involved in the regulation of several plant-specific biochemical pathways. The high divergence in protein sequences, despite the conserved domains, suggests that they might have different functions (Eulgem *et al.*, 2000).

It has been shown that a potato WRKY-like gene is induced by the treatment of potato leaves with *Erwinia carotovora* culture filtrate and pectate lyases (Dellagi *et al.*, 2000). This gene is also highly induced in a compatible interaction with *Phytophthora infestans*, and weakly induced in an incompatible interaction. Additionally, the potato WRKY-like gene is co-regulated with a class I endochitinase (Dellagi *et al.*, 2000). The tobacco TDBA12 WRKY-like gene is induced during the hypersensitive response to tobacco mosaic virus (Yang *et al.*, 1999), whereas the WIZZ (wound induced leucine zipper zinc finger) WRKY gene from tobacco is highly induced in response to wounding (Hara *et al.*, 2000). Using cDNA microarrays containing 25–30% of the *A. thaliana* genome for the determination of gene expression patterns under conditions that induce or repress systemic acquired resistance (SAR), it has been observed that the promoter regions of a co-regulated set of genes encoding pathogenesis-related proteins and proteins likely involved in SAR and disease resistance (PR-1 regulon) are enriched in WRKY-binding sites, suggesting that WRKY proteins may be essential for the transcriptional control of DR genes (Maleck *et al.*, 2000). In sugarcane, experimental data on the expression of WRKY-like genes is lacking.

The sugarcane defense responses to diazotrophic endophytes are largely unknown. *Herbaspirillum seropedicae* induces a hypersensitive-like response in the mottled stripe disease resistant sugarcane cultivar SP70-1143, as well as in the mottled stripe disease susceptible cultivar B-4362, and disease symptoms do not develop (Olivares *et al.*, 1997). In contrast, sugarcane B-4362 plants inoculated with *Herbaspirillum rubrisubalbicans* show disease symptoms on the leaves. Accumulation of polysaccharides and tannins in the parenchyma cells around the metaxylem of sugarcane clone Ja60-5 inoculated with *Gluconacetobacter diazotrophicus* has also been observed, suggesting that the plant defense system is activated during the interaction with the bacterium (Dong *et al.*, 1997). In sugarcane cultivar SP70-1143, no hypersensitive response has been observed after inoculation of leaves and stems with *G. diazo-*

trophicus, although an extracellular matrix accumulates around bacterial cells in the protoxylem and xylem parenchyma (James *et al.*, 2001).

In this work I used a hierarchical clustering approach coupled to color reordered data matrices to identify co-regulated sugarcane EST-contigs encoding putative DR genes, and DR EST-contigs differentially expressed in tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*, using the available sugarcane EST database (SUCEST).

MATERIAL AND METHODS

Data

The original data used in this study are available from the Sugarcane EST database (SUCEST; <http://sucest.lida.ic.unicamp.br/en/>). Construction of the cDNA libraries, sequencing methodology and procedure for EST clustering into contigs, using the Contig Assembly Program (CAP3) (Huang and Madar, 1999) are described elsewhere (Vettore *et al.*, 2001; Telles and Silva, 2001).

Expression of selected putative plant DR EST-contigs in sugarcane tissues

A total of 277 DR EST-contigs were selected for detailed expression analysis in cDNA libraries with more than 900 useful reads. The EST-contigs encode putative pathogenesis-related proteins (chitinases and β -1,3-glucanases), enzymes involved in the biosynthesis of isoflavonoid phytoalexins (phenylalanine ammonia-lyases, chalcone synthases, chalcone isomerases and isoflavone reductases), cell-wall proteins (hydroxyproline-rich and proline-rich glycoproteins), antioxidant enzymes (peroxidases, catalases and superoxide dismutases), WRKY-like transcription factors, and proteins likely involved in cell death control (*e.g.* nitric oxide synthase, gp91, LSD1 and NDR1/HIN1 homologues). EST-contigs were selected based on amino acid sequence homology to sequences in public databases, using the Basic Alignment Search Tool (BLAST) program (Altschul *et al.*, 1990) with a cut-off value of $e \leq 10^{-20}$ and a BLOSUM62 matrix. For each EST-contig, the frequency of reads in the selected libraries were computed and normalized for the total number of useful reads from each library. The EST-contigs and libraries were grouped by hierarchical clustering, using the Cluster and Tree View programs (Eisen *et al.*, 1998). Hierarchical clustering was performed using an un-centered correlation matrix and the average-linkage method (Eisen *et al.*, 1998). The data matrix was reordered, according to similarities in the pattern of gene expression and displayed as color arrays of EST-contigs, using a color scale representing the number of reads from a specific library in each EST-contig.

Analysis of WRKY-like EST-contigs from sugarcane

EST-contigs encoding homologues of WRKY-like proteins in public databases were selected for further analyses using the BLAST program with a cut-off value of $e \leq 10^{-20}$ and a BLOSUM62 matrix (Altschul *et al.*, 1990). The putative amino acid sequence of the protein was determined based on the consensus nucleotide sequence of the EST-contig using the Sequence Utilities program (<http://searchlauncher.bcm.tmc.edu>) (Smith *et al.*, 1996). The amino acid sequences of WRKY-like proteins from sugarcane, as well as *Arabidopsis thaliana*, *Avena sativa*, *Ipomoea batatas*, *Nicotiana tabacum*, *Oryza sativa*, *Petroselinum crispum*, *Pimpinella brachycarpa* and *Solanum tuberosum*, were aligned using the AlignX program (Vector NTI, Informax, Inc.) and default parameters. The phylogenetic relationships among WRKY-like proteins were determined by the Neighbor Joining method, using the AlignX and default parameters.

Expression of putative plant DR EST-contigs in sugarcane tissues infected with diazotrophic endophytes

From the data set used for the analyses of DR gene expression in different sugarcane tissues, 179 EST-contigs expressed in *in vitro* grown tissues infected with *Glucanacetobacter diazotrophicus* (library AD1) or *Herbaspirillum rubrisubalbicans* (library HR1), or non-infected tissues (libraries RT2 from roots and LV1 from leaves) were selected for further analyses. The EST-contigs were reorganized using the hierarchical clustering approach described above and the relative abundance of 124 putative DR EST-contigs expressed in non-infected tissues compared to infected tissues was calculated. The resulting matrix was log-transformed and reordered using the hierarchical clustering methodology as described. A new reordering was performed after exclusion of the EST-contigs not expressed in infected tissues (libraries AD1 and HR1). Suppression or induction of gene expression, compared to the non-infected control, is represented in green or red, respectively; no expression is represented in blue.

RESULTS AND DISCUSSION

Expression of selected putative plant DR EST-contigs in sugarcane tissues

Analysis of gene expression *in silico* can be performed based on the frequency of sequence tags in cDNA libraries, allowing comparisons of the expression profiles of specific genes in plant tissues (Ewing *et al.*, 1999). The methodology used in this study allowed the reordering of large data sets based on the similarity of expression profiles and the clustering by EST-contigs and/or cDNA libraries.

Using hierarchical clustering and reordered data matrices with color scales to visualize the expression levels of DR EST-contigs it was possible to distinguish clusters of ESTs with similar patterns of expression in the different sugarcane cDNA libraries (Figure 1), suggesting that they might be co-regulated *in vivo*. Figure 1 also shows a zoomed image of four clusters of EST-contigs with highly correlated expression patterns (correlation coefficients varying from 0.65 to 0.80). Cluster I, II, III and IV are composed of EST-contigs preferentially expressed in leaves (library LV1), tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans* (libraries AD1 and HR1), and roots (library RT2), respectively. As a control for tissue specific expression, a sugarcane EST-contig encoding the ribulose biphosphate carboxylase/oxygenase (RUBISCO) small subunit (RBS1) was used. The RBS1 EST-contig was shown to be expressed only in cDNA libraries synthesized from plant material containing photosynthetic tissues. A similar approach has been used for clustering rice ESTs based on their expression profiles (Ewing *et al.*, 1999). The data showed that genes with similar functions or cDNA libraries expected to have similar patterns of gene expression cluster together. In this study, sugarcane cDNA libraries expected to have similar expression profiles also clustered together (data not shown), indicating that hierarchical clustering is a powerful methodology to group genes with similar expression patterns and/or tissues with similar gene expression profiles.

Based on *in silico* expression profiles it is possible to distinguish genes encoding differentially regulated isoforms of a specific protein family and sets of potentially co-regulated genes (regulons). Regulons normally share promoter elements, which interact with specific transcription factors. In this study, different EST-contigs encoding WRKY-like transcription factors were shown to be co-regulated with several putative DR EST-contigs. In cluster IV (Figure 1), there is a single WRKY-like EST-contig associated with the expression of β -1,3-glucanases (GLUC), chalcone synthases (CHS), isoflavone reductase (IFR), chitinase (CHT) and peroxidase (POX), whereas in cluster III there are two WRKY-like EST-contigs co-regulated with EST-contigs encoding putative phenylalanine ammonia-lyase (PAL), CHS, IFR, catalase (CAT), GLUC, POX and a gp91-like protein. The regulons depicted in clusters I and II (Figure 1) were not associated with EST-contigs encoding WRKY-like proteins, suggesting that sugarcane DR regulons are activated by different transcription factors.

Analysis of WRKY-like EST-contigs from sugarcane

The large number of EST-contigs homologous to WRKY genes (at least 26) in sugarcane, suggests, as in *A. thaliana*, that these proteins might take part in complexes regulatory pathways (Eugem *et al.*, 2000). Preliminary analyses of the phylogenetic relationships between

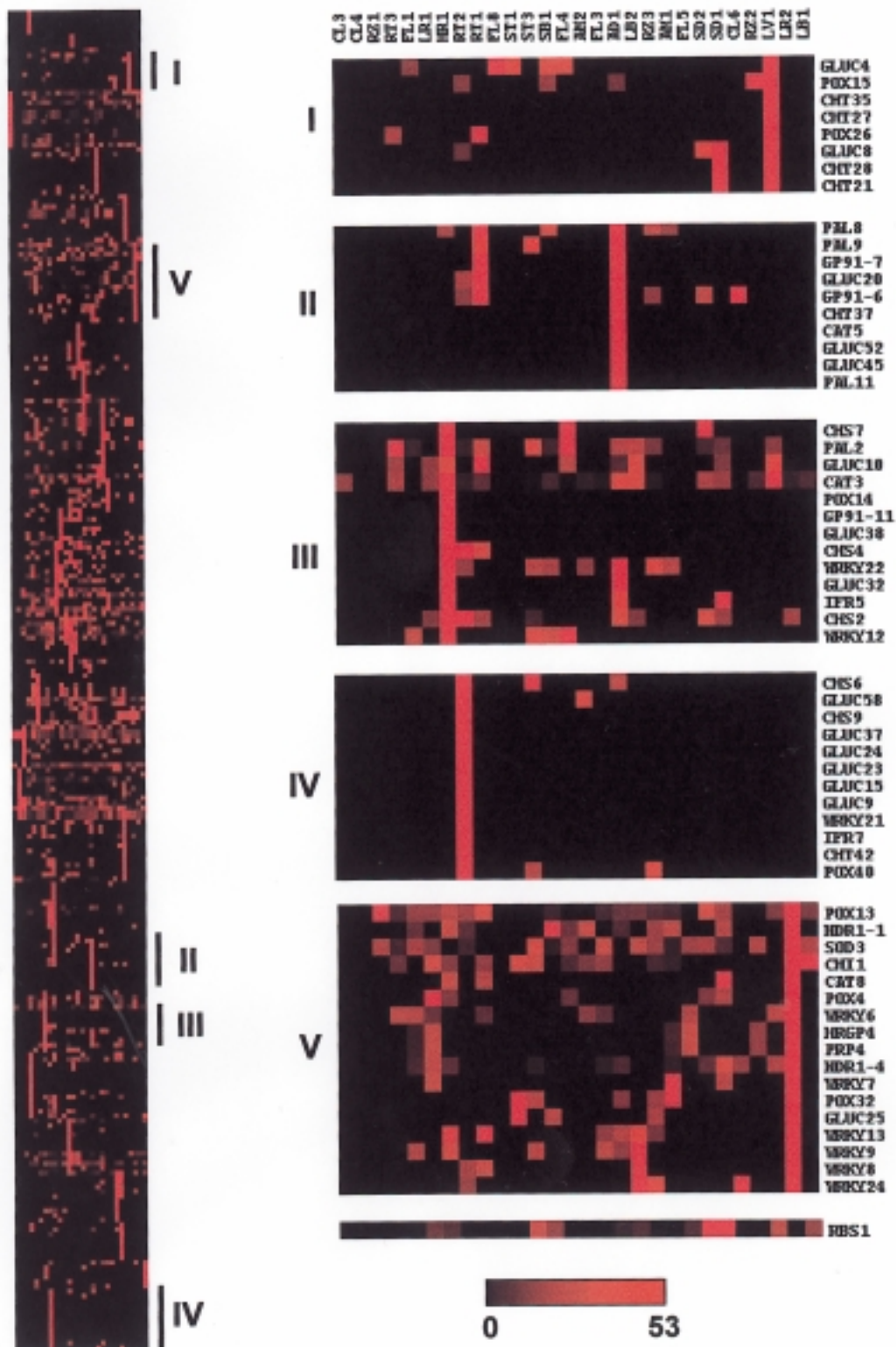


Figure 1 - Expression profiles of 277 putative defense-related sugarcane EST-contigs in selected cDNA libraries from the SUCEST database. I, II, III and IV represent clusters of EST-contigs preferentially expressed in leaves (LV1), tissues infected with *G. diazotrophicus* (AD1), tissues infected with *H. rubrisubalbicans* (HR1) and roots (RT2), respectively. V represents a regulon containing several EST-contigs encoding WRKY-like proteins. An EST-contig encoding the RUBISCO small subunit (RBS1) was used as control for the detection of tissue specific expression. Data represent the relative number of reads from a specific library in each EST-contig per 10,000 reads. Each EST-contig is represented by a single row, each library is represented by a single column.

EST-contigs encoding WRKY-like proteins in sugarcane (Figure 2) suggest that there are at least 12 sub-families of

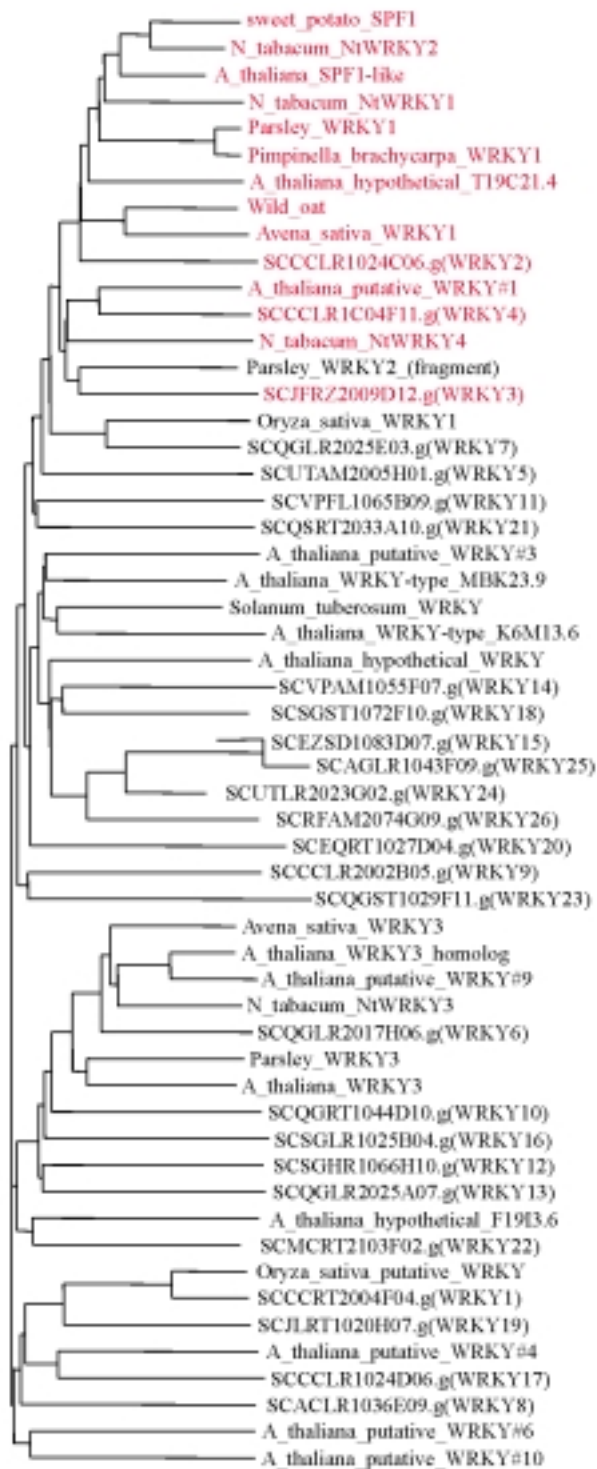


Figure 2 - Phylogenetic relationships between WRKY-proteins. Clustering was performed using the deduced amino acid sequence from the sugarcane EST-contig consensus nucleotide sequence and amino acid sequences in public databases. Sugarcane EST-contigs encoding putative WRKY-proteins were selected based on similarity to WRKY-like proteins from public databases (BLAST, $e \leq 10^{-20}$). The proteins highlighted in red contain two conserved WRKY-domains.

WRKY-like proteins in this species. The alignment of the amino acid sequences of different WRKY-like proteins showed the presence of very conserved WRKYGQK and zinc finger domains (Figures 3 and 4). Out of 26 sugarcane WRKY-like proteins analyzed, three (WRKY2, WRKY3 and WRKY4) presented two conserved WRKYGQK and zinc finger domains (Figure 4) and were phylogenetically related to other WRKY-proteins containing two WRKYGQK domains, forming a distinct sub-family (Figure 2). However, the number of sugarcane proteins with two WRKYGQK domains may be higher, since the sequences of several ESTs are only partial. Conserved protein kinase C phosphorylation sites were observed only in proteins with two conserved WRKYGQK domains (Figure 4), whereas no conserved nuclear localization sites were observed in sugarcane WRKY-like proteins. Proteins containing two WRKYGQK domains also have a putative tyrosine kinase C phosphorylation site in the vicinity of the N-terminal WRKYGQK domain (Figure 4).

Based on the data set analyzed, the sugarcane EST-contigs encoding WRKY-like proteins were distributed in 14 regulons ($r > 0.60$), most of which were associated with a single WRKY EST-contig, although five regulons were associated with two to six WRKY EST-contigs. Even though WRKY6, WRKY7, WRKY8, WRKY9, WRKY13 and WRKY24 showed similar expression patterns in sugarcane tissues (Figure 1), they belong to different sub-families (Figure 2). Similarly, WRKY proteins from the same sub-family (WRKY2, WRKY3 and WRKY4) were associated with different regulons. WRKY3 and WRKY11 were not significantly associated with the detected regulons (correlation coefficients of 0.58 and 0.13, respectively). These data suggest that in sugarcane, as in *A. thaliana*, WRKY-like proteins might be involved in the regulation of multiple biochemical pathways.

Expression of putative plant DR EST-contigs in sugarcane tissues infected with diazotrophic endophytes

In order to identify DR ESTs differentially expressed in sugarcane tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*, 181 putative DR EST-contigs expressed in non-infected tissues (libraries LV1 and RT2) and/or infected tissues (libraries AD1 and HR1) were selected and arrayed by similarity of expression patterns. Three clusters showing correlation coefficients between 0.80 and 0.94, comprising EST-contigs expressed preferentially in non-infected tissues, tissues infected with *G. diazotrophicus* and tissues infected with *H. rubrisubalbicans* could be distinguished (Figure 5A).

Since the sugarcane cultivar inoculated with the diazotrophic endophytes was not the same one used for the preparation of the control libraries (LV1 + RT2), only DR EST-contigs expressed in non-infected tissues (approximately 69% of the EST-contigs expressed in non-infected

and/or infected tissues) were used to evaluate relative changes in expression levels due to the interaction with the diazotrophic endophytes, as compared to the control (Figure 5B). Approximately 57% of the selected EST-contigs showed no expression in infected tissues (Figure 5B; blue cells). These EST-contigs were removed from the data matrix and a new hierarchical clustering was performed (Figure 5C). This procedure was adopted to ensure that only genes expressed in both sugarcane varieties were evaluated.

Suppression in gene expression higher than 2-fold in tissues infected with *G. diazotrophicus* was observed for approximately 28% of the selected putative DR EST-contigs (Figure 5C). Approximately 50% of the putative DR EST-contigs suppressed in *G. diazotrophicus*-infected tissues showed no expression in *H. rubrisubalbicans*-infected tissues. Additionally, in *H. rubrisubalbicans*-infected tissues only 7% of the EST-contigs selected showed suppression higher than 2-fold. Approximately 25 and 15% of the DR EST-contigs selected showed induction higher than 2-fold in tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*, respectively. These data indicate that the responses of sugarcane to *G. diazotrophicus* and *H. rubrisubalbicans* are distinct, suggesting that the control of the host infection and/or colonization processes depends on the endophyte.

It has been proposed that nitric oxide, as well as salicylic acid, may play an essential role in regulating the accumulation of hydrogen peroxide during resistance responses to pathogens by inhibiting catalase and ascorbate peroxidase activities (Clark *et al.*, 2000). Nitric oxide may also induce accumulation of salicylic acid and mediate cell death activation in plants (Delledonne *et al.*, 1998; Durner *et al.*, 1998). Suppression of the accumulation of EST-contigs encoding a putative nitric oxide synthase (NOS1) and a LSD1 homologue (LSD1-1) in sugarcane tissues infected with *H. rubrisubalbicans* suggests that the defense-responses are localized and no systemic induction of defense genes occurs in this particular interaction. In contrast, the induction of NOS1 and LSD1-1 in *G. diazotrophicus*-infected tissues suggests that a systemic defense response may occur. *G. diazotrophicus*-infected tissues also showed attenuated induction (2.5-fold) of an EST-contig encoding a catalase isoform (CAT3), as compared to *H. rubrisubalbicans*-infected tissues (5-fold). Additionally, the expression of three gp91-homologues in *G. diazotrophicus*-infected tissues was 2 to 4-fold higher than in the non-infected control, whereas in *H. rubrisubalbicans*-infected tissues these EST-contigs were not detected. Based on the differential induction of the EST-contig encoding CAT3 in *G. diazotrophicus* and *H. rubrisubalbicans* infected tissues, it is likely that these endophytes avoid the defense responses controlling the accumulation of H₂O₂ in the plant tissues. Induction of catalases in conditions that

favor microbial colonization of plant tissues has also been observed in mycorrhizal symbioses (Lambais, 2000).

In contrast, the 3 to 5-fold induction of EST-contigs encoding a peroxidase (POX9) and a phenylalanine ammonia-lyase (PAL2) in infected tissues, compared to the non-infected control, might be necessary to restrain the colonization process, through the strengthening of plant cell walls. The induction of other putative DR genes suggests that part of the plant defense system is kept active during the interactions with *G. diazotrophicus* and *H. rubrisubalbicans*. A partially activated defense system would be important to rapidly eliminate bacteria with low nitrogen fixation efficiency.

Even though the hierarchical clustering of EST-contigs is a valuable approach to identify genes possibly involved in the regulation of sugarcane-diazotrophic endophyte interactions, the actual role of each gene or set of genes should be evaluated *in vivo*.

Using a hierarchical clustering approach associated with reordered expression data matrices displayed as color arrays, it was possible to identify specific isoforms of the sugarcane WRKY-like transcription factor associated with DR regulons. It was also possible to identify DR genes likely involved in controlling sugarcane interactions with diazotrophic endophytes. Such an approach for analyzing EST data sets may be useful for the selection of potentially important genes involved in development and/or responses to environmental stresses, as well as to microorganisms.

RESUMO

Os padrões de expressão de 277 contíguos de ESTs de cana-de-açúcar codificando proteínas possivelmente associadas ao sistema de defesa vegetal, *e.g.* quitinases, β -1,3-glucanases, fenilalanina amônia liases, chalcona sintases, chalcona isomerases, isoflavona redutases, glicoproteínas ricas em hidroxiprolina, glicoproteínas ricas em prolina, peroxidases, catalases, superóxido dismutases, fatores de transcrição similares a WRKY e proteínas envolvidas no controle de morte celular, foram avaliados utilizando o banco de dados do SUCEST. Proteínas WRKY putativas de cana-de-açúcar foram comparadas e suas relações filogenéticas determinadas. Análises de agrupamento hierárquico foram utilizadas para a identificação de ESTs relacionadas à defesa vegetal com padrões de expressão similares em bibliotecas de cDNA representativas. Visando identificar ESTs relacionadas à defesa vegetal com expressão diferencial em tecidos de cana-de-açúcar infectados com *Glucanacetobacter diazotrophicus* ou *Herbaspirillum rubrisubalbicans*, 179 contíguos de ESTs possivelmente associados à defesa expressos em tecidos não-infectados (folhas e raízes) e/ou tecidos infectados foram selecionados e agrupados por similaridade de seus perfis de expressão. Alterações nos níveis de expressão de 124 contíguos de ESTs relacionados à defesa expressos em tecidos não-infectados foram avaliadas em tecidos infectados. Aproxima-

damente 42% desses contíguos de ESTs não apresentaram expressão em tecidos infectados, enquanto 15% e 3% apresentaram supressão de mais de duas vezes em tecidos infectados com *G. diazotrophicus* ou *H. rubrisubalbicans*, res-

pectivamente. Aproximadamente 14 e 8% dos contíguos de ESTs avaliados apresentaram indução superior a duas vezes em tecidos infectados com *G. diazotrophicus* ou *H. rubrisubalbicans*, respectivamente. A expressão diferen-

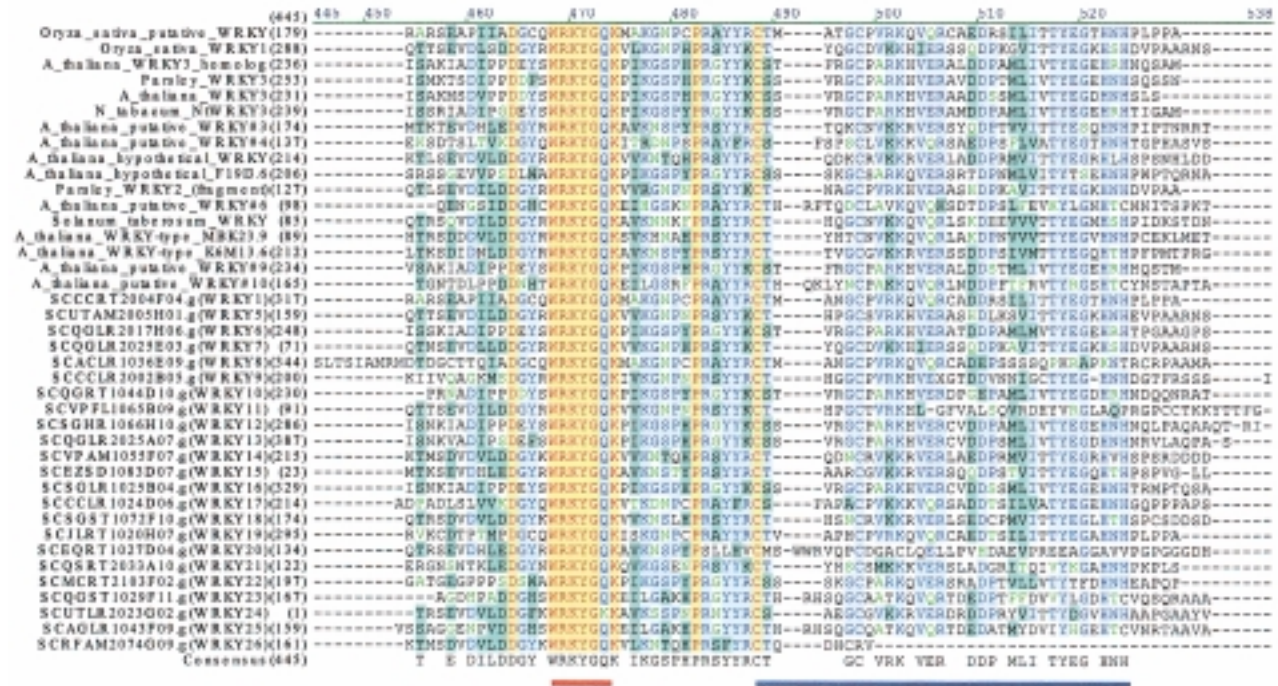


Figure 3 - Alignment of WRKY-like proteins containing one conserved WRKYGQK domain from different plant species. Amino acids highlighted in yellow are identical in all proteins. Blue represents consensus residues derived from a block of similar residues, green represents consensus residues derived from the occurrence of more that 50% of a single residue at a given position. The conserved consensus WRKYGQK domain is underlined in red and the conserved consensus zinc-finger domain is underlined in blue.

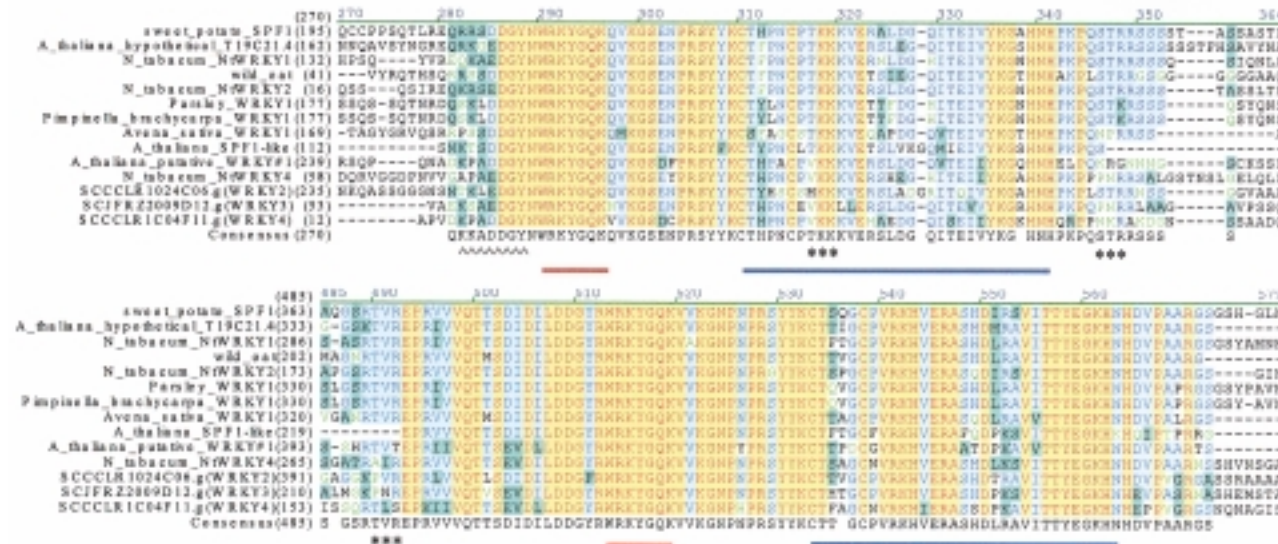


Figure 4 - Alignment of WRKY-like proteins containing two conserved WRKYGQK domains from different plant species. Amino acids highlighted in yellow are identical in all proteins. Blue represents consensus residues derived from a block of similar, green represents consensus residues derived from the occurrence of more that 50% of a single residue at a given position. The conserved consensus WRKYGQK domains are underlined in red. The conserved consensus zinc-finger domains are underlined in blue. Putative protein kinase C phosphorylation sites in the consensus sequence are indicated by ***. A putative tyrosine kinase phosphorylation site in the consensus sequence is indicated by ^^^^^.

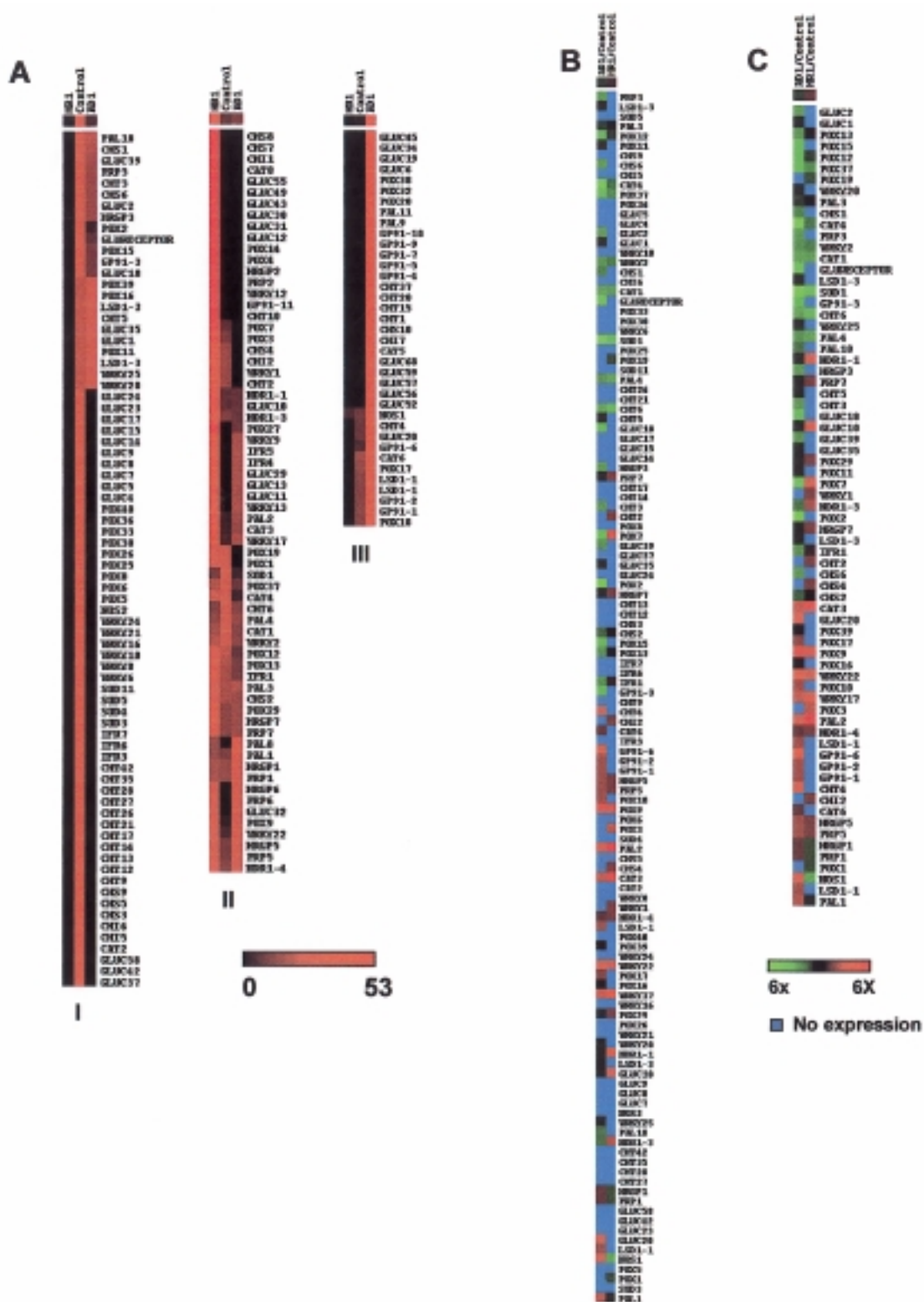


Figure 5 - Differential expression of putative sugarcane defense-related EST-contigs. **A.** Expression of defense-related EST-contigs in non-infected tissues, and/or tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*. Data represent the relative number of reads from a specific library in each EST-contig per 10,000 reads. Each EST-contig is represented by a single row; each library is represented by a single column. **B.** Differential expression of the putative defense-related EST-contigs expressed in non-infected tissues in tissues infected with *G. diazotrophicus* or *H. rubrisubalbicans*. **C.** Differential display of the putative defense-related EST-contigs expressed in tissues infected with either *G. diazotrophicus* or *H. rubrisubalbicans*. For **B.** and **C.**, data represent induction (red cells) or suppression (green cells) of expression in relation to the non-infected control, no expression is indicated in blue. cDNA libraries from *in vitro* grown roots (RT2) and leaves (LV1) were used as the non-infected control. Each EST-contig is represented by a single row; each library is represented by a single column.

cial de agrupamentos de genes relacionados à defesa vegetal pode ser importante para o estabelecimento de interações compatíveis entre plantas e diazotróficos endofíticos. Os resultados sugerem que o agrupamento hierárquico pode ser utilizado em escala genômica para a identificação de genes possivelmente envolvidos no controle de interações planta-microrganismos.

ACKNOWLEDGMENTS

Funding for this project was provided by the FAPESP (São Paulo, Brazil) as part of the Sugarcane Genome Project. The author is supported by the CNPq (Brasília, Brazil). Thanks to Jualiano C. Cury for preparing part of the data matrices used in this study.

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Clark, D., Durner, J., Navarre, D.A. and Klessig, D.F. (2000). Nitric oxide inhibition of tobacco catalase and ascorbate peroxidase. *Mol. Plant-Microbe Interact.* 13: 1380-1384.
- Dellagi, A., Heilbronn, J., Avrova, A.O., Montesano, M., Palva, E.T., Stewart, H.E., Toth, I.K., Cooke, D.E.L., Lyon, G.D. and Birch, P.R.J. (2000). A potato gene encoding a WRKY-like transcription factor is induced in interactions with *Erwinia carotovora* subsp. *atroseptica* and *Phytophthora infestans* and is co-regulated with class I endochitinase expression. *Mol. Plant-Microbe Interact.* 13: 1092-1101.
- Delledonne, M., Xia, Y., Dixon, R.A. and Lamb, C. (1998). Nitric oxide functions as a signal in plant disease resistance. *Nature* 394: 585-588.
- Dong, Z., McCully, M.E. and Canny, M.J. (1997). Does *Gluconacetobacter diazotrophicus* live and move in the xylem of sugarcane stems? Anatomical and physiological data. *Ann. Bot.* 80: 147-158.
- Durner, J., Wendehenne, D. and Klessig, D.F. (1998). Defense gene induction in tobacco by nitric oxide, cyclic GMP, and cyclic ADP-ribose. *Proc. Natl. Acad. Sci. USA* 95: 10328-10333.
- Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* 95: 14863-14868.
- Eulgem, T., Rushton, P.J., Robatzek, S. and Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5: 199-206.
- Eulgem, T., Rushton, P.J., Schmelzer, E., Hahlbrock, K. and Somssich, I.E. (1999). Early nuclear events in plant defence signaling: rapid activation by WRKY transcription factors. *EMBO J.* 18: 4689-4699.
- Ewing, R.M., Kahla, A.B., Poirot, O., Lopez, F., Audic, S. and Claverie, J.-M. (1999). Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* 9: 950-959.
- Gonzalez, J.E., Reuhs, B.L. and Walker, G.C. (1996). Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. *Proc. Natl. Acad. Sci. USA* 93: 8636-8641.
- Hara, K., Yagi, M., Kusano, T. and Sano, H. (2000). Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. *Mol. Gen. Genetics* 263: 30-37.
- Huang, X. and Madan, A. (1999). CAP3: A DNA Sequence Assembly Program. *Genome Res.* 9: 868-877.
- James, E.K., Olivares, F.L., de Oliveira, A.L.M., dos Reis, F.B., da Silva, L.G. and Reis, V.M. (2001). Further observations on the interaction between sugarcane and *Gluconacetobacter diazotrophicus* under laboratory and greenhouse conditions. *J. Exp. Bot.* 52: 747-760.
- Lambais, M.R. (2000). Regulation of plant DR genes in arbuscular mycorrhizae. In: *Current advances in mycorrhizae research* (Podila, G.K. and Douds Jr, D.D., eds.). APS Press. St. Paul, Minnesota. pp. 45-59.
- Lambais, M.R. and Mehdy, M.C. (1998). Spatial distribution of chitinases and β -1,3-glucanase transcripts in bean arbuscular mycorrhizal roots under low and high soil phosphate conditions. *New Phytol.* 140: 33-42.
- Maleck, K., Levine, A., Eulgem, T., Morgan, A., Schmid, J., Lawton, K.A., Dangl, J.L. and Dietrich, R.A. (2000). The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat. Genet.* 26: 403-410.
- McDowell, J.M. and Dangl, J.L. (2000). Signal transduction in the plant immune response. *Trends Biochem. Sci.* 25: 79-82.
- Niehaus, K., Kapp, D. and Puhler, A. (1993). Plant defense and delayed infection of alfalfa pseudonodules induced by an exopolysaccharide (EPS-I)-deficient *Rhizobium meliloti* mutant. *Planta* 190: 415-425.
- Olivares, F.L., James, E.K., Baldani, J.I. and Dobereiner, J. (1997). Infection of mottled stripe disease-susceptible and resistant sugarcane varieties by the endophytic diazotroph *Herbaspirillum*. *New Phytol.* 135: 723-737.
- Ruiz-Lozano, J.M., Roussel, H., Gianinazzi, S. and Gianinazzi-Pearson, V. (1999). Defense genes are differentially induced by a mycorrhizal fungus and *Rhizobium sp* in wild-type and symbiosis-defective pea genotypes. *Mol. Plant-Microbe Interact.* 12: 976-984.
- Smith, R.F., Wiese, B.A., Wojzynski, M.K., Davison, D.B. and Worley, K.C. (1996). BCM Search Launcher - an integrated interface to molecular biology data base search and analysis services available on the World Wide Web. *Genome Res.* 6: 454-462.
- Telles, G.P. and Silva, F.R. (2001). Trimming and clustering sugarcane ESTs. *Gen. Mol. Biol.* 24 (1-4): 17-23.
- Vettore, A.L., Silva, F. R., Kemper E.L. and Arruda, P. (2001). The libraries that made SUCEST. *Gen. Mol. Biol.* 24 (1-4): 1-7.
- Yang, P., Chen, C., Wang, Z., Fan, B. and Chen, Z. (1999). A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. *Plant J.* 18: 141-149.