Genetics and Molecular Biology, 30, 3 (suppl), 888-905 (2007)

www.sbg.org.br

Research Article

Identifying water stress-response mechanisms in citrus by *in silico* transcriptome analysis

Gisele A.M. Torres, Marcos A. Gimenes, Vicente E. de Rosa Jr. and Vera Quecini

Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, Campinas, SP, Brazil.

Abstract

Water deficit is one of the most critical environmental stresses to which plants are submitted during their life cycle. The evolutionary and economic performance of the plant is affected directly by reducing its survival in the natural environment and its productivity in agriculture. Plants respond to water stress with biochemical and physiological modifications that may be involved in tolerance or adaptation mechanisms. A great number of genes have been identified as transcriptionally regulated for water deficit. EST sequencing projects provide a significant contribution to the discovery of expressed genes. The identification and determination of gene expression patterns is important not only to understand the molecular bases of plant responses but also to improve water stress tolerance. In our citrus transcriptome survey we have attempted to identify homologs to genes known to be induced and regulated under water stress conditions. We have identified 89 transcripts whose deduced amino acid sequences share similarities with proteins involved in uptake and transport of water and ion, 34 similar to components of the osmolyte metabolism, 67 involved in processes of membranes and proteins protection and 115 homologs of reactive oxygen species scavenger. Many drought-inducible genes identified are known to be regulated by development, salt, osmotic and low temperature. Their possible roles in specific or general mechanisms of water stress citrus responses are discussed.

Key words: environmental stress, CitEST, data mining, tolerance mechanisms, water deficit.

Received: July 24, 2006; Accepted: February 8, 2007.

Introduction

Water stress is one of the most critical environmental stresses that plants are exposed to. It affects both the evolutionary and the economic performance of a plant by directly reducing its survival in the natural environment and its productivity in agriculture. The International Water Management Institute estimates that by the year 2025, one third of the world population will inhabit regions of severe water scarcity. Moreover, the use of fresh water for irrigation competes with its use for human consumption, and the increased yields obtained with the employment of irrigation water have been predicted to be unsustainable (IWMI, 2005). The disruption of the plant water status and its detrimental effects on the plant performance are common features of several abiotic stresses, thus, likely to be present in the vast majority of agricultural zones.

In perennial species, seasonal variation in environmental conditions may influence water relationships. In *Citrus* trees, the large canopy and low hydraulic conductivities of the stem and the root contribute to high water deficit (Moreshet *et al.*, 1990). Moreover, at midday, a transient water deficit is a characteristic of citrus (Cohen *et al.*, 1997) and correlated to reduced photosynthesis rates (Brakke and Allen, 1995). Although temporary water deficit periods occur in many of cultivated regions, irrigation is rarely employed for citrus species in Brazil (Ortolani *et al.*, 1991).

The environmental amounts of water available to plants may decrease due to drought, altered ion content and water uptake caused by salinity or cellular dehydration induced by the formation of extracellular ice during freezing stress (Verslues *et al.*, 2006). Therefore, dehydration, osmotic stress, salinity and, to some extent, cold stresses have been treated as a group of factors imposing alterations in the plant water status (Vinocur and Altman, 2005; Verslues *et al.*, 2006). Water stress and dehydration are general terms to indicate a mild form of water deficit, a physiological condition where the water status of the plant body undergoes minor changes (Bray, 1997). The most severe form of water stress is desiccation, which consists in the loss of most protoplasmic "free" or "bulk" water, forcing the plant cells to survive exclusively with "bound water", that is, the water associated to the cell matrix (Ramanjulu and Bartels,

Send correspondence to Gisele A.M. Torres. Present address: Embrapa Trigo, BR 285, km 294, Caixa Postal 451, 99001-970, Passo Fundo, RS, Brazil. E-mail: gtorres@cnpt.embrapa.br.

2002). As a consequence of water stress, citrus canopy growth decreases whereas the root system appears to sustain less damage (Lloyd and Howie, 1989). Water stress may reduce leaf expansion and even lead to premature abscission or senescence of adult leaves (Munns, 2002).

Throughout evolution the sessile nature of plants has led to the development of adaptive strategies to cope with environmental stresses. An essential feature of these adaptive strategies is that they must be elicited in response to an external stimulus, providing utilization of resources when needed and preventing energy waste in the absence of the stress. Plant cells have evolved mechanisms to perceive distinct environmental signals, to integrate them and to modulate the expression of the required genes to respond accordingly. Plants are able to cope with water deficit via two general mechanisms: *i)* stress avoidance - by producing seeds before the establishment of the drought conditions or by developing specific morphological adaptations, such as leaf surfaces less prone to transpiration losses, reduced leaf area, sunken stomata, increased root length and density (Ramanjulu and Bartels, 2002) and, *ii)* stress tolerance - which consists of coordinated physiological and biochemical alterations at cellular and molecular levels, such as the accumulation of late embryogenesis-abundant (LEA) proteins associated with activity of the antioxidant system of the cells.

The molecular bases of water stress tolerance remains unknown. Candidate genes induced by water-deficit stress in plants which are relatively sensitive to cellular dehydration have been identified and characterized, mainly in the model plant *Arabidopsis thaliana* (Vinocur and Altman, 2005; Verslues *et al.*, 2006). The investigated plant systems have been shown to have common molecular and physiological components in a wide range of tolerance levels, indicating a major role for spatial and temporal gene expression regulation in water stress resistance (Ramanjulu and Bartels, 2002; Taji *et al.*, 2004).

The primary site for the detection of water stress in plants remains unknown (Ramanjulu and Bartels, 2002). In yeast and bacterial cells, well-characterized osmosensors are responsible for detection of osmotic stress. Osmosensors are members of the broad class of two-component systems, consisting of a histidine kinase sensor and an intracellular response-regulator which is responsible for relaying the phosphorylation signal to the next component of the pathway, leading to transcriptional regulation of gene expression (Wurgler-Murphy and Saito, 1997). A homolog to the yeast osmosensor SLN1 has been recently characterized in *Arabidopsis*, *AtHK1* (Urao *et al.*, 1999). Microarray whole-genome expression profiles in *A. thaliana* have shown water stress-induced alteration in the transcription of several sensor- and response regulator-like genes as early as 15 min after the onset of water deprivation stress (Seki *et al.*, 2002). The transcriptional regulation of two component system-like genes is similarly affected under osmotic and salt stress; however, the functional significance of these findings remains unknown (Seki *et al.*, 2002).

In eukaryotic cells, biotic and abiotic stresses trigger the production of reactive oxygen species (ROS) which causes the oxidation of cellular components and ultimately, cell death. The activation of oxidative stress signaling is evolutionarily conserved from yeast to mammals and involves mitogen-activated protein kinase (MAPK) cascades. Similarly, in plants, the production of ROS is induced by environmental, mechanical and biological stress conditions (Inzé and van Montagu, 1995). The observation of changes in the protein phosphorylation status of plants submitted to water deficit conditions indicates the involvement of reversible phosphate relay in the regulation of drought stress signal transduction (Pastori and Foyer, 2002). Several MAPKs are components of dehydration- and abscisic acid-induced signal transduction in plants (Mikolajczyk *et al.*, 2000). Thus, ROS-triggered signal transduction via a MAPK-based cascades induces the expression of detoxification and stress protection genes, such as heat shock proteins (HSP), glutathione-S-transferases (GST), peroxidases, superoxide-dismutases and pathogenesis-related (PR) proteins, protecting the plant from damage (Kovtun *et al.*, 2000). Consistently, the constitutive expression of a heterologous MAPK kinase kinase (MAPKKK) from tobacco induced both drought (Shou *et al.*, 2004a) and freezing (Shou *et al.*, 2004b) tolerance in transgenic maize.

Recent microarray studies in *Arabidopsis* indicate that a set of genes is involved in promoting tolerance, whereas another set responds to water deficit stress (Seki *et al.*, 2002). The genes responding to dehydration can be categorized into two distinct classes: *i)* early-response genes - in seconds or minutes, and *ii)* late-response genes - over hours, days or even weeks (Ramanjulu and Bartels, 2002). This temporal separation demonstrates distinct roles in the stress response; the early genes could provide initial protection and amplification of the signal transduction pathway while the late ones could be involved in adaptation to the stress condition. On the other hand, the manipulation of genes involved in protection and maintenance of cell components structure and cellular functions has been the major target of attempts to produce plants showing enhanced stress tolerance.

During late stages of embryogenesis in dico- and monocotyledonous plants, the cells undergo a severe dehydration process and simultaneously accumulate LEA proteins. Other stress situations, such as low temperatures, increased salinity and exogenous ABA treatment also induce LEA proteins which indicate their involvement in the general cellular protective system against water loss (Cuming, 1999). Molecularly, they are characterized by a biased amino acid composition: high hydrophilicity and high solubility in water. LEA proteins can be divided into five

groups according to their predicted biochemical properties and motif similarity (Ingram and Bartels, 1996; Cuming, 1999). LEA-like proteins were the most abundant transcripts found in the dehydration transcriptome of the bryophyte *Tortula ruralis*, demonstrating the importance of this class of proteins in the adaptive acquirement of tolerance to dehydration, as well as in the cellular rehydration repair response (Oliver *et al.*, 2004). Moreover, drought and salinity tolerance have been increased in rice (Xu *et al.*, 1996) and wheat (Sivamani *et al.*, 2000) with the introduction and expression of a heterologous LEA protein HVA1 from barley.

Aquaporins are members of a family of water channel proteins involved in the facilitation of its transport along transmembrane water potential gradients, thus regulating the hydraulic conductivity of membranes and water permeability (Maurel and Chrispeels, 2001). Several aquaporin-coding genes are upregulated by dehydration in *Arabidopsis* (Yamaguchi-Shinozaki *et al.*, 1992), tomato (Fray *et al.*, 1994) and *C. plantagineum* through the ABA-dependent and independent pathways (Mariaux *et al.*, 1998). However, the transcriptional regulation of aquaporin genes is complex with several hierarchic levels of control, and is responsive to both water deficit and numerous environmental and physiological factors (Maurel and Chrispeels, 2001; Tournaire-Roux *et al.*, 2003; Jang *et al.*, 2004).

The integrity of photosynthetic structures, especially of membrane-associated proteins, after a period of water stress is a crucial mechanism in desiccation tolerant plants (Godde, 1999; Bartels and Salamini, 2001). In *C. plantagineum*, three genes that are highly induced upon the onset of water deficit encode chloroplast-localized stress proteins (DSP). Two of them, DSP22 and DSP34, are thylakoid-associated and one, DSP21, is localized in the stroma (Schneider *et al.*, 1993; Alamillo and Bartels, 2001). Similarly stress-induced chloroplast protection proteins were found in potato (Pruvot *et al.*, 1996; Rey *et al.*, 1998) and in whole-genome expression profile in *T. ruralis* submitted to rehydration (Oliver *et al.*, 2004).

A common strategy for protection against water deficit in many organisms is the accumulation of compatible solutes or osmolytes. Osmolytes are only synthesized in response to osmotic stress and are biochemically inert in the cell, exclusively helping to maintain the osmotic balance necessary for growth and cellular metabolism under dehydration (Bray *et al.*, 2000). Besides their role in osmotic adjustment, osmolytes might also be involved in other protective mechanisms, such as ROS scavenging (Hong *et al.*, 2000).

The increased synthesis of osmolytes induced under water-stress conditions is caused by modulation of the expression and activity of key regulatory enzymes in their biosynthetic pathways (Ramanjulu and Bartels, 2002). In plants, sugars, polyols, proline, quaternary ammonium compounds and tertiary sulfonium compounds are often found to function as osmolytes. In citrus, the osmotic adjustment under salt stress is mostly dependent upon accumulation of proline and inorganic ions (Arbona *et al.*, 2005).

The accumulation of soluble sugars is a common feature of the desiccation process, in both desiccation-tolerant and desiccation-susceptible plants. Sugars have a role in osmotic adjustment, but also have indirect protective effects, such as protein stabilization (Carpenter *et al.*, 1990). In the desiccation-tolerant plant *C. plantagineum*, dehydration induces the conversion of 2-octulose, an eight-carbon sugar, to sucrose (Bianchi *et al.*, 1991). This conversion correlates to increases in the gene expression for sucrose synthase (*SUS*) and sucrose phosphate synthase (*SPS*) (Ingram *et al.*, 1997; Kleines *et al.*, 1999), which are considered key enzymes of sucrose synthesis/metabolism. Under conditions of dehydration/osmotic stress, the expression of genes coding for *SUS* isoforms is upregulated in several plants (Pelah *et al.*, 1997; Déjardin *et al.*, 1999). Similarly, antisense expression of the *SPS* coding sequence in potato plants completely suppressed the water stress-induced stimulation of sucrose synthesis (Geigenberger *et al.*, 1999). Thus, SUS and SPS in plants are crucial steps in the acclimation process of dehydration. Highly soluble sugars, such as the polyfructose molecules fructans, are involved in plant and bacterial adaptation to osmotic stress. Transgenic tobacco and sugar beet plants, overexpressing the gene *SacB* that codes for a levan sucrase from *Bacillus subtillis*, accumulated higher levels of fructans and performed better than the untransformed controls under water deficit conditions (Pilon-Smits *et al.*, 1995; Pilon-Smits *et al.*, 1999). Similar results were obtained with transgenic tobacco plants overexpressing a gene encoding the trehalose synthase subunit (*TPS1*) of the yeast trehalose synthase enzyme (Holmstrom *et al.*, 1996) and bacterial trehalose-6-phosphate synthase and trehalose-6-phosphate-phosphatase genes (Pilon-Smits *et al.*, 1998). The function of trehalose, a non-reducing disaccharide of glucose, in desiccation is hypothesized to involve the stabilization of membrane proteins and lipids and its use as a reserve metabolite. The accumulation of the methylated sugar alcohol, D-ononitol, in transgenic tobacco plants overexpressing the *IMT1* gene from *Mesembryanthemum crystallinum*, has led to increased salt and drought tolerance (Sheveleva *et al.*, 1997).

The regulation of the levels of proline in plants under water stress conditions is simultaneously controlled by upregulation of the P-5-C synthase (*P-5-CS*) gene and downregulation of the proline dehydrogenase gene (*ProDH*) (Yoshiba *et al.*, 1997). Furthermore, a proline and glycine betaine transporter (*LeProT1*) has been shown to be induced in tomato plants submitted to scarce water conditions (Schwacke *et al.*, 1999). Despite the complexity of the control of proline levels, transgenic tobacco and rice plants with higher proline levels due to the overexpression of *P-5-CS* gene, had higher biomass production under water

stress conditions (Kavi Kishor *et al.*, 1995; Zhu *et al.*, 1998).

Water shortage, like other biotic and abiotic stresses, causes the accumulation of enzymes responsible for the oxidative cellular defense system, such as superoxide dismutase, ascorbate, peroxidases, catalases, glutathione-S-transferases and glutathione peroxidases (Kovtun *et al.*, 2000). Other proteins involved in the repair of damaged cellular components have also been shown to be induced under water deficit conditions (Seki *et al.*, 2001; Oliver *et al.*, 2004).

The effect of individual genes involved in water stress tolerance is minimal. Molecular switches and regulatory genes have been proposed to be a better means to increase plant tolerance to water restrictive conditions (Ramanjulu and Bartels, 2002). Thus, one of the most successful strategies for plant modification for enhanced drought tolerance is based on the manipulation of genes coding for transcription factors and/or signaling partners that directly protect plant cells against water deficit. Furthermore, the majority of transcripts identified in association to plant stress responses are regulated by this condition instead of being *de novo* synthesized. For citrus plants submitted to water deficit during cyclic periods, as for other perennial species, it is important to explore constitutive gene expression that could be induced under constraint conditions.

Thellungiella halophila, an extremophyle plant displaying tolerance to high salinity, low humidity and freezing, exhibits higher pre-stress concentrations of several compounds that have been shown to have protective functions in osmotic imbalance, a common component between water- and salt-stress (Hasegawa *et al.*, 2000). The regulated expression of constitutive genes under stress conditions indicates further posttranscriptional regulation and may represent an early protection of the plant against water constraint. Drought is often interconnected to various environmental stresses that may induce similar cellular damage. As a consequence, similar cell signaling pathways are activated and oxidative stress is frequently induced causing protein denaturation (Smirnoff, 1998; Shinozaki and Yamaguchi-Shinozaki, 2000).

The plant organ submitted to stress conditions is another relevant aspect. Considered the hidden part of the plant, the root system is primordial to citrus physiology under water constraints, not only due to the provision of water and mineral nutrients but also as a storage organ. Roots accumulate carbohydrates in the winter and play a critical role in exporting them to developing fruitlets during early stages of fruit set. However, gene expression modulations and changes in physiological parameters, related to cell water status, are slighter in roots than in shoots (Schena *et al.*, 1995; Seki *et al.*, 2002; Torres *et al.*, 2006).

Thus, in order to cope with water deficit, the plant induces modifications of its physiological state and metabolic pathways using two major categories of responses: *i)* uptake and transport of water and ions, and *ii)* protection of membranes and proteins. These genes are known to be constitutively expressed and could actively participate in improving plant tolerance subject to water stress. The aim of this study was to survey citrus EST databases to identify components presenting similarity to genes functionally related to the fore mentioned classes of water stress responses that could represent interesting candidates for transgenic analyses.

Material and Methods

Database searches and alignments

Homologs of functionally characterized genes involved in dehydration responses were identified in BLAST searches (Altschul *et al.*, 1997) against EST contig sequences from the citrus index databases at CitEST. These consisted of approximately 176,200 ESTs obtained from the sequencing of 53 citrus libraries. Data validation was performed by tBLASTx and tBLASTn searches with BLOSUM80 scoring matrix of the retrieved citrus sequence against the databases at NCBI built inside the CitEST project. The resulting alignments were filtered by a threshold e-value of $1e^{-15}$ for the hits and further analyzed. Validated sequences were translated and protein (deduced amino acid) alignments were performed using ClustalX (Thompson *et al.*, 1997). When necessary, alignments were manually adjusted using Lasergene MegAlign (DNASTAR, Madison, WI, USA).

Motif analysis and *in silico* characterization

The identified citrus homologs were further investigated for the presence and sequence conservation of recognizable functional domains: described in several protein analysis and gene function databases (European Bioinformatics Institute - European Molecular Biology Laboratory - EMBL-EBI, **Ex**pert **P**rotein **A**nalysis **Sy**stem - ExPaSy of the Swiss Institute of Bioinformatics - SIB, and Protein Families - Pfam).

Phylogenetic analysis

The putative functionality of the citrus genes in comparison to their homologs from model systems was assessed by genetic distance and phylogenetic studies. Phylogenetic analyses were performed using distance and parsimony methods in the software PAUP* 4.0b10, using the software default parameters. Resampling bootstrap trees containing 1000 random samples were constructed using PSIGNFIT software and ClustalX (Thompson *et al.*, 1997).

In silico gene expression analysis

Qualitative gene expression profiling was performed by *in silico* analyses of the citrus EST database through the generation of a relational matrix between the number of ESTs corresponding to a determined gene in a given library

and normalizing the result to the number of reads of the library. Gene expression patterns of EST contigs and libraries were determined by hierarchical clustering, based on Spearman Rank correlation matrix, using Cluster and Tree View software packages (Eisen *et al.*, 1998) and cluster results were shown as their average expression pattern. The expression profile matrix was ordered accordingly and displayed in grayscale.

Results and Discussion

We have performed extensive BLAST and key word searches of the citrus transcriptome to identify homologs of the genes involved in responsive mechanisms to water deficit in citrus. We have searched for transcripts whose deduced amino acid sequences share similarity to proteins involved in uptake and transport of water and ions, osmolyte metabolism, processes of membrane and protein protection and reactive oxygen species scavenging. In CitEST databases, 305 assembled sequences and EST singlets sharing significant sequence identity with functionally characterized proteins were identified and analyzed (Table 1).

Ion transporters

We have identified 63 sequences in the citrus transcriptome showing significant deduced amino acid homology to functionally characterized ion transport-associated proteins (Table S1): 10 EST contigs and eight singlets are similar to ATPases involved in ion transport and H^+-exchange; whereas 24 EST contigs and 21 singlets show anion- and cation-binding and transporting functional motifs.

The families of Calcium (ACA), H^+ (AHA) and H^+/Na^+ exchanger (NHX) ATPases are approximately equally represented in citrus analyzed databases, including in libraries derived from non-stressed and non-infected tissues. The family of vacuolar A type H^+-ATPases (VHA-A) is characterized by lower expression levels in several model species. However, we were able to identify two ESTs from *C. sinensis* libraries that are highly similar to members of the family. Several ESTs containing motifs responsible for K^+ and Ca^{+2} transport were identified in citrus species. Interestingly, a gene encoding a putative Cu^{+2}-transporter was found in the citrus genome database (Table S1). The elevated sequence similarity of the deduced amino acid sequence from the citrus EST, with previously characterized haloacid dehalogenase-like hydrolases, indicate a role in heavy metal detoxification (Himelblau and Amasino, 2000).

Ion transporters are associated with water loss-derived secondary stresses (Verslues *et al.*, 2006) and several developmental processes including embryogenesis and fruit development involve cellular dehydration (Ramanjulu and Bartels, 2002). Interestingly, in citrus libraries derived from whole plants submitted to water stress, no significant increase in the frequency of ion transporter-like reads was observed. However, libraries obtained from the initial three stages of fruit development presented a high frequency of anion and cation transporter-related ESTs (Figure 1). Reads showing sequence similarity to monovalent cation transporters, such as K^+-transporting proteins and H^+-ATPases,

Table 1 - Citrus transcripts identified by tBLASTn searches of CitEST databases whose deduced amino acid sequences show similarity to drought-responsive mechanism components.

Functional categories	CitEST transcripts	
		Total
Ion transporters		63
Ion-transporting ATPases	18	
Ion transporters and transporter-associated proteins	45	
Major intrinsic proteins (MIP)		26
Plasma membrane intrinsic protein (PIP)	15	
Tonoplast intrinsic proteins (TIP)	3	
NOD26-like proteins (NIP)	5	
Small basic intrinsic protein (SIP)	3	
Osmolyte biosynthesis		34
Glycine betaine	3	
Sugars	12	
Mannose	2	
Proline	6	
Polyamines	5	
Trehalose	6	
Heat shock proteins		47
HSP60 family	1	
HSP70 family	19	
HSP90 family	9	
HSP100 family	8	
sHSP family	10	
LEA proteins		18
COR19 family	6	
CsDHN	1	
LEA group 1 protein	2	
LEA14_GOSHI	3	
LEA group 5 protein	1	
PgEMB8	4	
PsLEAm	1	
SP1 protein		2
ROS-scavenging enzymes		115
SOD-like	23	
APX /PRPX	46	
MDAR	6	
DHAR	10	
GR	4	
GST	5	
AOX	4	
CAT	15	
Thioredoxin	2	
Total		305

are highly induced in the first and second stages of fruit development in *C. sinensis* and *C. reticulata*. ESTs related to divalent cation transporters were more frequent in libraries derived from later stages of fruit development in *C. reticulata*; whereas, in *C. sinensis*, they were more frequent in the first and third stage, and were virtually absent from second stage-derived libraries (Figure 1). These observations suggest that in citrus, developmentally induced dehydration is responsible for the most significant changes in ion transporter gene expression. This indicates that these proteins may have secondary roles in environmentally induced-water loss protection.

Water channels

Water channel proteins are responsible for the transmembrane flow of water across the lipid bilayer plasma membrane, so they are involved in several aspects of plant-water relations. The proteins have a conserved pore-forming structure throughout evolution, consisting of six membrane spanning helices and two loops containing asparagines-proline-alanine (NPA) motifs (Figure 2A) and are members of large membrane intrinsic protein (MIP) families in plants (Chaumont *et al.*, 1998; Johanson *et al.*, 2001). In citrus, we have identified 18 EST contigs and eight EST singlets sharing deduced amino acid sequence homology with functionally characterized MIP family members (Table S2). Amino acid sequence alignment shows significant sequence conservation at the six membrane spanning helices and at the two NPA motifs between the citrus and *Arabidopsis* proteins (Figure 2B).

From the total 26 MIP-like proteins, 15 were more related to plasma membrane-associated water channels (PIP), three to tonoplast channels (TIP), five to nodulin-like proteins (NIP) and, three to the newly described family of small basic intrinsic proteins (SIP). Phylogenetic analysis indicates that the citrus MIP family has perhaps undergone an intraspecific duplication process, due to the presence of divergent citrus-exclusive PIP branch (Figure 2C). In *Arabidopsis*, the majority of the PIPs and some TIPs are constitutively highly expressed, whereas NIPs are categorized as rare transcripts (Alexandersson *et al.*, 2005). In citrus, a relatively high frequency of NIP and NIP-like transcripts were identified (Table S2), suggesting that in spite of the high level of sequence conservation between *Arabidopsis* and citrus MIP proteins, they may be submitted to distinct expression regulation. Dehydration transcriptionally and translationally down-regulates *PIP* genes in *Arabidopsis* (Alexandersson *et al.*, 2005); however, we were unable to identify changes in the frequency of MIP-like reads in citrus libraries submitted to environmental and developmental processes that trigger water loss. Thus at this point, the role of water channel proteins in desiccation protection in citrus remains unclear.

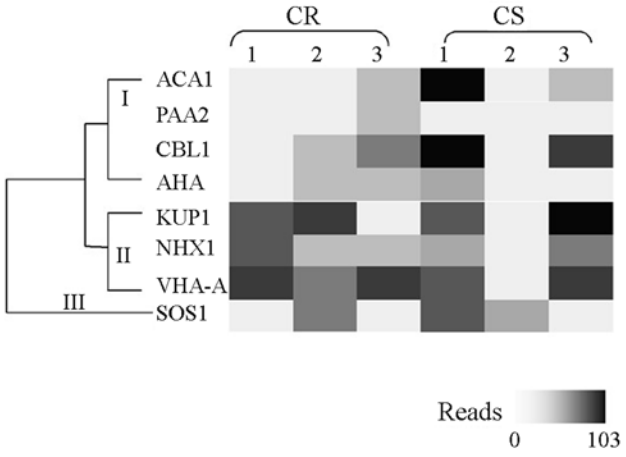

Figure 1 - Expression profile of ion transporter-like transcripts in fruit development libraries stage 1, 2 and 3 from *Citrus sinensis* (CS) and *Citrus reticulata* (CR). Data represents the normalized relative number of reads from a specific library showing sequence similarity to the ion transporters.

Osmolytes

Analyses of EST databases from citrus species revealed the presence of several transcripts which showed sequence similarity to genes encoding enzymes involved in the production of the osmolytes and osmoprotectants most commonly found in model plant species in response to stress (Table S3). These observations suggest the presence of extensive conservation in osmolyte and osmoprotectant metabolism between citrus and model plants. However, we were unable to identify genes coding for components involved in the metabolism of rare plant osmolytes or compounds associated to fungi- and bacteria osmoregulators such as those associated to production of DMSP, choline-O-sulfate or D-ononitol (Table S3).

Drought-induced osmotic stress causes detrimental changes in cellular components, which can be prevented by a wide range of metabolites, including amino acids (*e.g.* proline), quaternary and other amines (*e.g.* glycine-betaine and polyamines) and a variety of sugars and sugar alcohols (*e.g.* mannitol and trehalose). These metabolic changes were not evident in citrus transcriptome analyses, since the frequency of reads showing sequence similarity to transcripts involved in osmoprotection responses remained unaffected in libraries derived from drought-stricken tissues (data not shown). Thus, extensive conservation in osmolyte metabolism is observed in citrus, although its role in water stress protection remains to be established.

Several components of the betaine biosynthetic pathway were identified in citrus transcriptome, including the enzymes choline monooxygenase (CMO) and betaine aldehyde dehydrogenase (BADH) (Table S3). Citrus species CMO and BADH are highly similar to the enzymes from other organisms, including at catalytic sites (Table S3). This provides evidence for the production of glycine betaine rather than PRO- or Ala-betaines in citrus. Two singlet reads showed extensive similarity to CMO, whereas three BADH-like transcripts were found. The low fre-

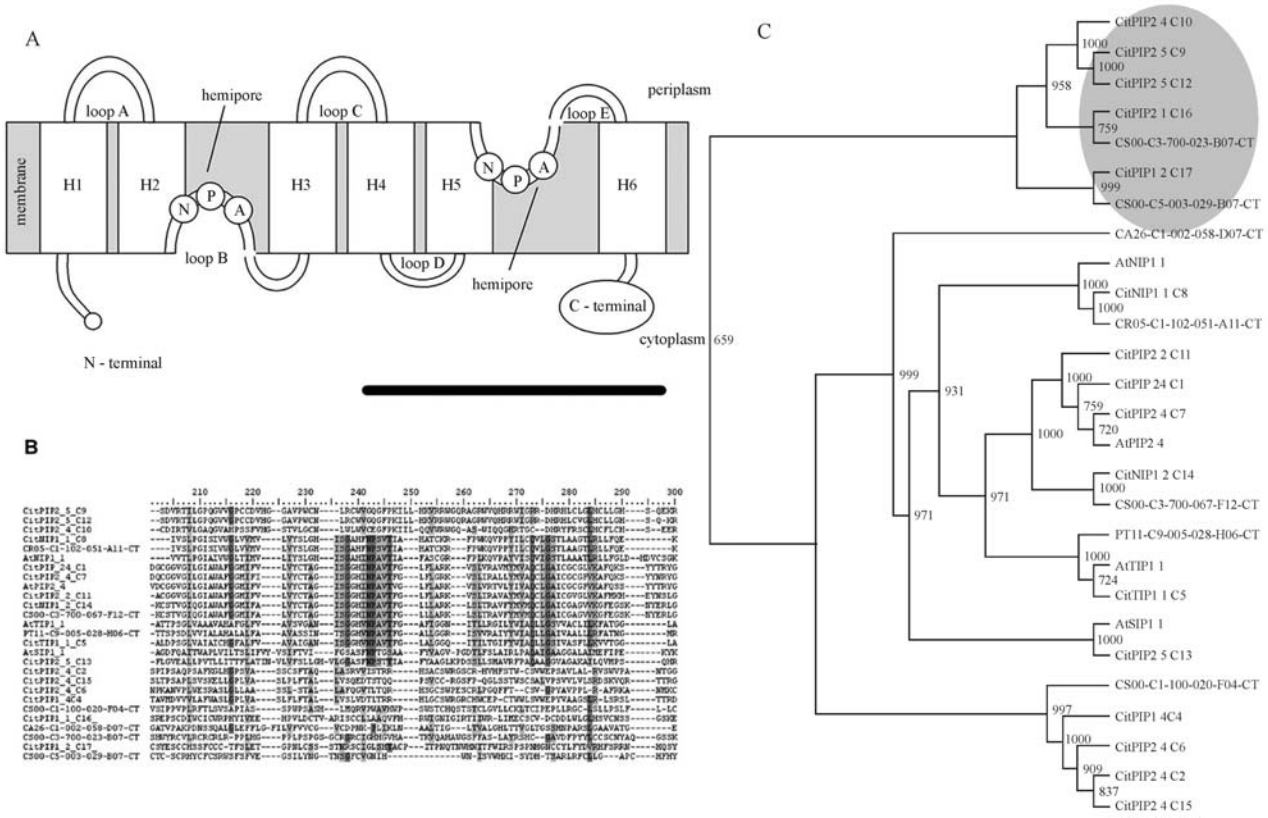

Figure 2 - Citrus membrane intrinsic proteins (MIP). (A) Schematic representation of the domain organization of plant MIP family. (B) Sequence alignment of loop C, H3, H4, loop D and H5 domains from *Arabidopsis* MIP proteins and their citrus counterparts. (C) Phylogenetic analysis of deduced amino acid sequence of citrus MIPs and prototypical representatives of *A. thaliana* PIP, TIP, NIP and SIP families. The shaded circle represents a citrus exclusive clade of divergent PIP-like aquaporins.

quency of betaine metabolism-related transcripts suggests that the pathway is expressed at low levels and/or that these genes are specifically upregulated by biotic and abiotic stresses, although such an induction was absent from drought-derived tissues. Alternatively, the low frequency of betaine biosynthesis transcripts could correspond to low levels of the compound *in planta*. Proline is biosynthetically derived from the amino acid L-glutamate and its direct precursor is the imino acid (S)-Δ1-pyrroline-5-carboxylate (P5C). Transcripts showing sequence homology to genes encoding enzymes responsible for proline biosynthesis and degradation were identified in citrus species transcriptome, including homologs to P5C synthase (P5CS), P5C reductase (P5CR) and pyrroline-5-carboxylate reductase (PROD) (Table S3). We have also identified transcripts showing similarity to the coding sequence of enzymes involved in sugar biosynthesis (Table S3). In citrus, the later category comprises trehalose synthesis enzymes (T6PP and T6PS) and mannose specific enzyme M6PR, along with sucrose metabolism components (enzymes SUSY and SPS) and one invertase (fructofuranosidase) that is likely to be involved in fructan metabolism (Table S3). Similarly to that observed for the betaine pathway, a role for transcripts showing sequence similarity to components of proline and osmolyte sugar metabolism in drought tolerance remains to be established.

Metabolic engineering of abiotic stress tolerance employs two general strategies; the first one aiming to increase the production of specific desired compounds or reduction in the levels of unwanted toxic compounds (Capell and Christou, 2004). However, modulation of a single enzymatic step is usually regulated by cellular systems that tend to restore homeostasis, thus limiting the potential of this approach (Vinocur and Altman, 2005). Alternatively, targeting multiple steps of the same pathway has been proposed as a means to control metabolic fluxes in a more predictable manner (Konstantinova *et al.*, 2002). The lack of correlation between the frequency of transcripts showing sequence conservation to members of osmolyte and osmoprotectant metabolism and drought-induced responses in citrus suggests that these protective mechanisms remain unsaturated under water shortage. This indicates potential for metabolic engineering.

Heat-shock proteins

Following heat stress, the amount of cellular proteins is diminished. However, some proteins are accumulated under those conditions and are thus called ‘heat shock pro-

teins' (HSP). The HSPs are ubiquitous and present high sequence conservation. They are present in all cellular compartments and their classification follows their kDa molecular mass: HSP60, HSP70, HSP90, HSP100 and the small HSP, which range from 15 to 30 kDa (Vierling, 1991).

Distinct proteins belonging to HSP classes have been related to plant water stress responses. In this study, we have identified 47 sequences presenting similarity to heat shock proteins in the citrus transcriptome (Table S4): one contig reveals homology to a chaperonin from HSP60 class; 18 contigs and one singlet show high similarity to HSP70; six contigs and three singlets, to HSP90; five contigs and three singlets to HSP100; and finally, 10 contigs are similar to small HSP (sHSP). Most of the sequences are derived from HSP70 class (38%) followed by ESTs encoding proteins from HSP90 class (34%) (Figure 3).

HSP70 are essential in helping to prevent the aggregation and assisting in the folding of proteins under normal and stress conditions (Sung *et al.*, 2001a). They also play a regulatory role in stress-associated gene expression (Lee and Schöffl, 1996). Plant HSP70 genes are encoded by a highly conserved multigene family and are localized in several cellular compartments (Sung *et al.*, 2001b; Wang *et al.*, 2004). HSP70s are known to be differentially regulated in response to developmental stages and to a wide range of stresses. Several studies suggest the association of HSP70 to other stress-related responses of plants. These proteins could coordinate to prevent cellular damage and to re-establish cellular homeostasis. In citrus species, they could represent interesting targets for biotechnological manipulations aimed at improving plant tolerance to water deficit.

HSP90 family members have been isolated from animals and plants. They encode structurally related proteins ranging from 80 to 90 kDa. HSP90 is one of the major species of molecular chaperones that requires ATP for its functions (Wang *et al.*, 2004). Three putative ATP-binding motifs are highly conserved among members of *hsp90* gene family. *Hsp90* genes are developmentally regulated in plants (Koning *et al.*, 1992; Marrs *et al.*, 1993; Krishna and Gloor, 2001) and their expression is induced upon stress conditions in both prokaryotes and eukaryotes. HSP90 proteins are distinct from many other molecular chaperones due to their interaction with steroid hormone receptors and signaling kinases (Young *et al.*, 2001). Thus, it is not surprising that they were found in high proportion in our transcriptome survey. Furthermore, Hsp90 acts as part of a multichaperone system together with Hsp70 and cooperates with several co-chaperones (Wang *et al.*, 2004).

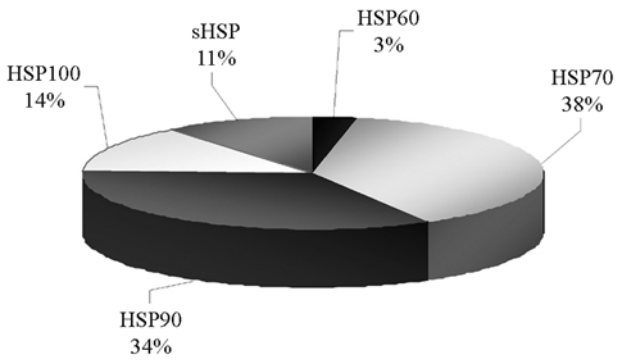

Figure 3 - Relative abundance of HSP families involved in water stress response in the citrus EST database.

Comparing the data from control and water-stressed libraries of *Citrus limonia* roots, we have found alterations in the expression levels of five sequences related to HSP. A putative cytosolic HSP70 (AT3G12580) and a chaperonin containing TCP-1 domain belonging to HSP60 class present high levels of expression that diminish in response to water stress. The other three sequences (a HSC70-1 homolog; a sequence similar to an HSP81-2 and another encoding to a putative HSP100) were not detected in control plants and are induced in stressed roots following water deficit.

The heat shock cognate proteins (HSC70) are expressed in normal growth conditions and may or may not be induced in response to stress (DeRocher and Vierling, 1995; Miernyk, 1997; Sung *et al.*, 2001a). In *Arabidopsis thaliana* seedlings and bean plants, the expression of homologs of *hsc70* genes was induced after dehydration (Kiyosue *et al.*, 1994; Torres *et al.*, 2006). Interestingly, the transcript encoding an HSC70 was induced during the light period in untreated roots and in a manner probably independent of ABA biosynthesis (Torres *et al.*, 2006). Thus, HSC70 homologs appear as promising candidates for functional analyses. The early expression changes verified in other plant species may be an indication of a mechanism capable of responding rapidly to water constraint.

We could also identify a contig encoding a chaperonin with very high similarity to a protein isolated from *Bruguiera sexangula*. It has a chaperone activity *in vitro* and is capable of improving salt stress tolerance when expressed in *E. coli* (Yamada *et al.*, 2002). BsCCTα has three highly conserved domains: an equatorial domain that contains the ATP-binding site; an apical domain that binds to peptides; and an intermediate domain. By expression studies in *E. coli*, the authors verified an increased salt tolerance and identified a region of 218 amino acids as responsible and sufficient to improve stress tolerance. This region is equally conserved in the citrus contig (Table S4), which also presents a decrease in mRNA expression under water deficit conditions. A detailed protein expression study would provide further information about its role in water stress. At this point, assuming a high turnover rate, there are indications that these transcripts play an important role in cellular protection.

HSP81-2 is a member of HSP90 family that is expressed abundantly in root apical meristem of *Arabidopsis* and is induced by NaCl (Yabe *et al.*, 1994). The *hsp90* gene

family was characterized in *A. thaliana* plants and calli in response to heat and heavy metals (Milioni and Hatzopoulos, 1997). Although a similar induction profile for all six genes in response to stress conditions was observed, the results were specifically dependent on the kinetic experiment analysis. Furthermore, a combination of heat and drought could induce a more rapid change in *hsp90* genes expression. In *Citrus*, it would be interesting to analyze the profile of transcript levels of *hsp90* mRNAs in response to the kinetics of water stress in different organs.

Genes coding for HSP100 have been isolated from various plant species and can be also termed 'Clp' proteins due to their sequence similarity to *E. coli* ClpA (Gottesman *et al.*, 1990). They are found in all cellular compartments and are hypothesized to participate in proteolysis regulation, protein translocation and acquired thermotolerance. Thermotolerance in yeast was associated to higher stability of membrane proteins which may be the result of the protective role of HSPs (Swan, 1997). Under water stress conditions, a similar function could also be required. In *Phaseolus lunatus* leaves, a ClpB homolog has been identified as mediating the response of the chloroplast to heat stress. Surprisingly the citrus ESTs were present in libraries derived from roots. The discrimination between mitochondrial and plastidial transit peptides remains uncertain and some proteins have been demonstrated to be dually targeted (Peeters and Small, 2001; Zhang and Glaser, 2002). Based on mitochondrial proteome analyses, Heazlewood *et al.* (2004) demonstrated that only one half of the proteins were correctly predicted. Thus, these data require future investigations to clarify the specific role of ClpB in roots and the subcellular location of this protein.

Comparing data from healthy leaves of six citrus species (Figure 4), we have observed a great number of reads corresponding to HSP81-2, one of the sequences already described as differentially expressed in response to water deficit. In *Citrus aurantium*, we were unable to detect BiP homologs. This protein belongs to HSP70 family and is targeted to endoplasmic reticulum. Its overproduction in transgenic tobacco plants enhanced their tolerance to water deficit (Alvim *et al.*, 2001). BiP plays a role in translocation/retranslocation, folding and assembly of ER proteins and its expression could be regulated by various environmental conditions (Noh *et al.*, 2003). Exclusively in *C. aurantifolia* and *C. aurantium*, we have observed high levels of expression of small heat shock proteins (HSP18.2 and HSP17.9, respectively). The sHSP were directly implicated in improving plant tolerance to water stress (Sun *et al.*, 2002). This kind of differential regulation in both species could be an interesting tool for future studies on the mechanisms of drought tolerance. Finally, in *C. aurantifolia* and *Poncirus trifoliata* we have found a remarkably distinct pattern of expression of mtHSP70. Bean mtHSP70 homologs were found in the outer mitochondrial membrane facing the cytosol and equally in the mitochon-

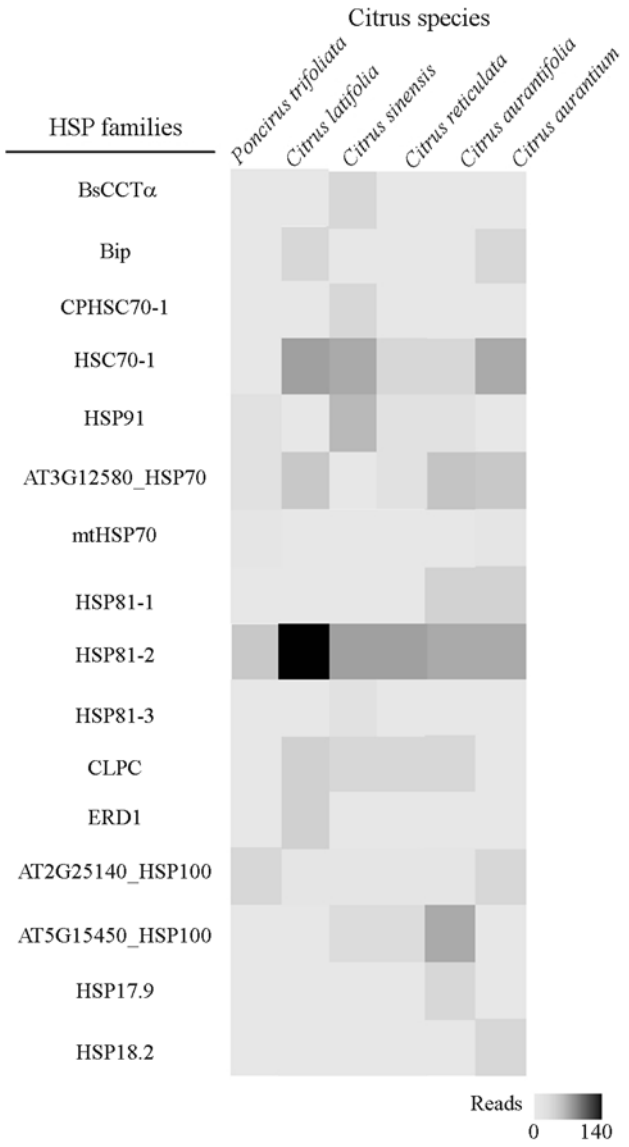

Figure 4 - Expression profile of heat shock proteins-like transcripts in healthy leaves from *Citrus aurantifolia*, *Citrus aurantium*, *Citrus latifolia*, *Citrus sinensis*, *Citrus reticulata* and *Poncirus trifoliata*. Data represents the normalized relative number of reads from a specific library showing sequence similarity to the ion transporters.

drial matrix, playing a role in protein translocation (Vidal *et al.*, 1993). Mitochondria perform a variety of functions in eukaryotic cells, notably responding to cellular signals as oxidative stress. The precise and eventually specific roles of this protein in *C. aurantifolia* and *Poncirus trifoliata* remain to be elucidated.

LEA proteins

LEA proteins were initially described as being present during the late period of seed development that is accompanied by dehydration. They are characterized by low molecular weight ranging mainly from 10 to 30 kDa and above 30 kDa (Hong-Bo *et al.*, 2005) and most of them constitute a more widespread group called "hydrophilins"

(Garay-Arroyo *et al.*, 2000). First studied in developing cotton seeds (Dure and Croudh, 1981), LEA proteins have been detected in several plant species (Close, 1996; Han and Kermode, 1996; Chen *et al.*, 2003). The proteins are regulated not only during seed development but also upon most diverse environmental conditions. However, the knowledge of the biochemical functions of LEA proteins is still incomplete (Bartels and Salamini, 2001) but some of them certainly contribute to improving plant drought tolerance (Babu *et al.*, 2004; Hara *et al.*, 2004).

In citrus, we have identified nine EST contigs and nine EST singlets sharing deduced amino acid sequence homology to LEA protein coding genes (Table S5). From the total 18 LEA-like proteins, six were more related to COR19, one to CsDHN, two to LEA group 1, three to LEA group 4, four to PgEMB8, one to LEA group 5 and one to the recently identified PsLEAm.

The most frequent sequences correspond to homologs of LEA 14-A coding genes followed by *LEA*s from group 1. These sequences were mainly found in libraries obtained from fruits in different developmental stages in both *C. sinensis* and *C. reticulata* species (Figure 5). In *C. sinensis*, a defined pattern of changes in LEA14-A and COR19 family genes expression was not observed. In contrast, in *C. reticulata*, we have observed an increase in LEA-14A expression during the fruit maturation process. This fact indicates that these genes are down regulated throughout fruit maturation stages, which are known to be accompanied by dehydration.

Moreover, in citrus seed-derived libraries, we have observed generally high levels of expression of LEA-coding genes. These genes could exert specific roles related to seed development and the imposition of a desiccation step. A novel mitochondrial LEA protein has been identified in pea seeds (Grelet *et al.*, 2005). The corresponding mRNAs are responsive to maturation of seeds and to water deficit. Conversely, they are down-regulated during seed germination. We identified transcripts sharing sequence similarity to this putative mitochondrial LEA exclusively in *Poncirus trifoliata* seeds. Therefore, it would be interesting to analyze its profile of expression under different conditions.

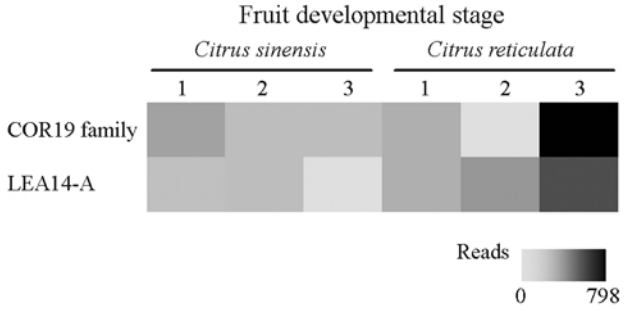

Figure 5 - Expression profile of late embryogenesis proteins-like transcripts in fruit development libraries stage 1, 2 and 3 from *Citrus sinensis* and *Citrus reticulata*. Data represents the normalized relative number of reads from a specific library showing sequence similarity to the ion transporters.

On the other hand, *Cor19* genes and the *CsDHN* gene were absent from seed libraries. *Cor19* and dehydrin genes in general may be involved in response to several environmental stresses and to development stages. The kinetics of the response and their regulation varies according to the species and to the stress applied. Recently, Hara *et al.* (2004) demonstrated that CuCOR19 acts as radical scavenger and may reduce the oxidative damage induced by water deficit in *Citrus unshiu*. Generally, as dehydrin genes are expressed during water stress responses in plants, it is speculated that they protect plants from damage caused by cell desiccation. Therefore their specific role in citrus drought tolerance remains to be clarified.

Soluble protein 1

Soluble protein 1 (SP1) has been described as a member of a novel class of plant stress-response proteins. Initially isolated from an expression library of water-stressed aspen plants, SP1 homologs are widespread over a range of different organisms (Wang *et al.*, 2002). In citrus transcriptome survey, we have identified one EST contig and one singlet with high similarity to *Populus tremula*, *Arabidopsis thaliana* and *Oryza sativa* SP1-coding ESTs (Table S6). Citrus sequences have conserved Phe residues at four positions and the consensus motif "K-F-WG-D" located in the middle portion of the sequences. Aspen SP1 transcripts are expressed under non-stressed conditions and are also induced upon water and hypo-osmotic stress. These transcripts are also maintained at high levels during stress recovery (Wang *et al.*, 2002). The identified sequences of SP1 homologs in citrus were present in different parts of the plant, such as leaves, bark and fruits. Interestingly, SP1-like mRNAs were present in *Poncirus trifoliata* plants infected with *Citrus* tristeza virus. Nevertheless, SP1 is stress-related and has high thermostability as small heat shock proteins (sHSP), it does not present amino acid sequence nor function similarity in stress protection (Dgany *et al.*, 2004). Therefore, Wang *et al.* (2006) have described aspen SP1 as a remarkably resistant protein. It is boiling-stable and resistant to proteases, organic solvents and high levels of ionic detergent. However, at this point its function and involvement in repair of cellular damage remains to be elucidated.

Reactive oxygen species-scavenging enzymes

Aerobic organisms utilize oxygen as electron receptors during respiration. Under optimal conditions, there is the reduction of O_2 to H_2O, following the reception of four electrons. However, when O_2 receives one, two or three electrons, reactive oxygen species (ROS) are formed (Levine, 1999). ROS have been demonstrated to have signaling function in several environmental responses and developmental processes, including biotic and abiotic stress

responses, allelopathic plant-plant interactions, cell division and elongation, and programmed cell death (Apel and Hirt, 2004; Foyer and Noctor, 2005). Moreover, normal cell metabolism constantly generates ROS; thus, their basal levels are tightly controlled. In *Arabidopsis thaliana*, the ROS gene network comprises at least 152 genes, such as the scavenging enzymes (superoxide dismutases - SODs, ascorbate peroxidases - APXs, catalases - CATs, glutathione peroxidases - GPX, and peroxiredoxins - PRPX) and enzymes involved in ascorbate-glutathione cycle (monodehydroascorbate reductase - MDAR, dehydroascorbate reductase - DHAR and glutathione reductase - GR) (Mittler *et al.*, 2004).

Under water stress conditions, the plants may activate the antioxidant-defense system to control ROS overproduction (Bartels, 2001). Nonetheless, this activation is dependent on the plant species, the developmental stage, the time and intensity of stress conditions, as on the radicals formed and the cellular compartment localization (Levine, 1999; Bowler and Fluhr, 2000).

In citrus transcriptome analysis, we have identified 115 transcripts that share sequence conservation to *Arabidopsis* ROS metabolism (Table S7 to S12): 23 SOD-like (Table S7), 46 plant APX/PRPX (Table S8), six MDAR, 10 DHAR, four GR, five GST (Table S9), four AOX (Table S10), 15 CAT (Table S11) and two transcripts sharing sequence similarity to potato thioredoxin (Table S12).

In citrus transcriptome, copper/zinc SOD (CDS) transcripts were the most abundant subfamily (48%) of this class of ROS scavenging enzymes, whereas iron (FDS) and manganese (MDS) were less frequent (30% and 22%, respectively) (Table S7). The deduced amino acid sequence of citrus species SOD transcripts is highly similar to the *Arabidopsis* proteins, especially for FDS- and MDS-like sequences (Figure 6). Eight citrus transcripts showed extensive conservation of the deduced amino acid sequence to *Arabidopsis* APX proteins (Table S8). Interestingly, the sequence conservation between citrus and *Arabidopsis At*GPX family is less significant, although a higher number of homologous transcripts was identified, representing 57% of all citrus peroxidase-like mRNAs (Table S8).

A high ratio of reduced peroxidized ascorbic acid and glutathione is believed to be essential for the proper scavenging of ROS in cells. It is maintained by glutathione reductase (GR), monodehydroascorbate reductase (MDAR) and dehydroascorbate reductase (DHAR) using NADPH as reducing power (Mittler, 2002). APX catalyses the reduction of H_2O_2 with simultaneous oxidation of ascorbate generating monodehydroascorbate (MDHA) (Yoon *et al.*, 2004) or dehydroascorbate (DHA). MDAR and DHAR have auxiliary functions in the maintenance of proper ascorbate concentration in cells by reducing the MDHA and DHA radical directly to ascorbate (Mittler *et al.*, 2004).

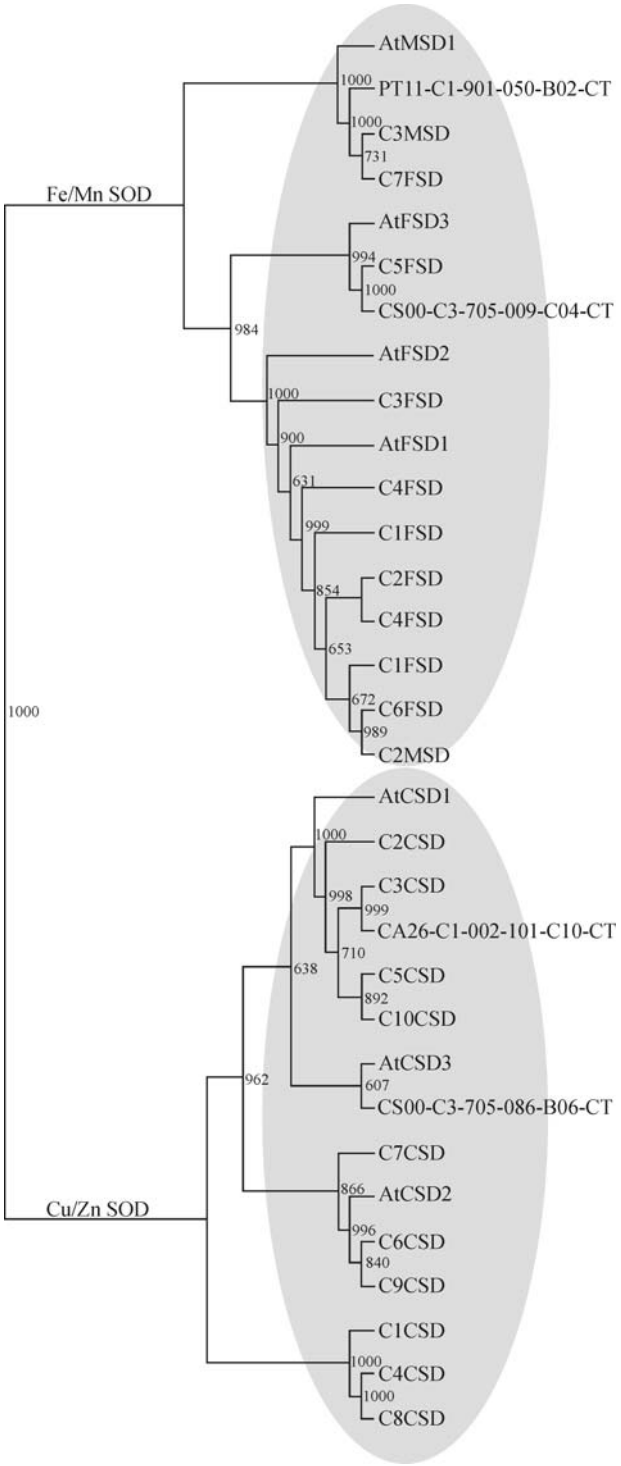

Figure 6 - Phylogenetic analysis of citrus superoxide dismutase (SOD) homologs and the *Arabidopsis thaliana* prototypical SOD proteins. Neighbor-joining trees for citrus deduced amino acid and *Arabidopsis* full length sequences aligned with ClustalX are shown. Bootstrap values are indicated above each branch. At, *Arabidopsis thaliana*; C Number, contig number; CSD, copper/zinc superoxide dismutase; FDS, iron superoxide dismutase; MDS, manganese superoxide dismutase.

MDAR reduces the MDHA to ascorbate at the expense of a NAD(P)H and DHAR reduces DHA to ascorbate using glutathione (GSH) as a reducing agent, resulting in GSSG

that is reduced to GSH by glutathione reductase (GR) using NADPH (Allen, 1995). Glutathione is a ubiquitous tripeptide that is synthesized via two ATP-dependent reactions. In plants, most of glutathione exists in the reduced form (GSH) and the content of the oxidized form (GSSG) normally does not exceed 10% of the total GSH content. Six transcripts showed extensive conservation of the deduced amino acid sequence to *Pisum sativum* MDAR, 10 to *Arabidopsis* DHAR, and four to *Arabidopsis* GR (Table S9).

Glutathione S-transferases (GST) are a family of multifunctional, dimeric enzymes that catalyze the nucleophilic attack of glutathione on lipophilic compounds with electrophilic centers. In plants, it has been reported mainly associated with herbicide detoxification, although they could respond to diverse environmental stresses, notably dehydration (Marrs, 1996). In citrus EST database, we were able to identify four EST contigs and one singlet coding for GST homologs (Table S9).

Alternative oxidase (AOX) mediates the conversion of O_2 to H_2O in a one-step reaction coupled with ATP production. This enzyme is thought to function as a homodimeric protein and is encoded by multigene families in several plant species (Whelan *et al.*, 1995; Finnegan *et al.*, 1997; Ito *et al.*, 1997; Saisho *et al.*, 1997). Expression of *AOX* genes is developmentally and environmentally regulated (Vanlerberghe and McIntosh, 1997). In our citrus EST search, we have found three EST contigs and one singlet presenting similarity to *Arabidopsis AOX* transcripts (Table S10). One contig and one singlet related to *AOX1A* and two contigs related to *AOX2*. The expression of *AOX* homologs was not affected in water stress-derived libraries. In tobacco- and *Arabidopsis*- transgenic plants, AOX exerts a role in diminishing the oxidative stress, evidenced by changes at transcriptional levels (Maxwell *et al.*, 1999; Saisho *et al.*, 2001). Thus, in order to fully characterize citrus AOX, more studies are necessary.

Catalase multigene family in *A. thaliana* consists of three genes (*CAT1*, *CAT2* and *CAT3*) encoding individual subunits, which associate to form at least six isozymes that are readily resolved by non-denaturing gel electrophoresis (McClung, 1997). Catalase is a tetramer and catalase activity gels reveal that three isozymes are detectable throughout the *Arabidopsis* life cycle (Salomé and McClung, 2002). CAT2 and CAT3 are clock-regulated and CAT1 is not (Michael and McClung, 2002).

All three mRNAs are detectable in freshly imbibed seeds, although the pattern of mRNA relative abundance varies among the three genes during early germination (McClung, 1997). *Arabidopsis* transgenic plants with high levels of tolerance to chilling and oxidative stresses presented an induction of transcript levels and the activity of CAT1. *CAT1* appears to be an important responsive gene to oxidative stress and to be induced in water deficit-tolerant plants (Hsieh *et al.*, 2002a; Hsieh *et al.*, 2002b). In our citrus transcriptome survey, we have identified one EST contig that is highly related to *CAT1*, 10 EST contigs and two singlets similar to *CAT2* and finally, one EST contig homologous to *CAT3* (Table S11). Although a great number of EST coding for catalase in citrus ESTs libraries is available, we were unable to observe a bias in these transcript levels in response to stresses or developmental conditions. As observed in *Arabidopsis* (Salomé and McClung, 2002), *CAT2* transcripts were the most abundant in citrus libraries.

Thioredoxins are small proteins containing a Cys-Gly-Pro-Cys active site domain that is able to reduce disulfide bridges on target proteins (Eklund *et al.*, 1991). They exert a general role in enzyme activity regulation via thiol redox control. CDSP32 is a recently identified new thioredoxin highly induced under drought and oxidative stress conditions (Rey *et al.*, 1998; Broin *et al.*, 2000). Lines lacking CDSP32 are more susceptible to photoxidative treatments (Broin *et al.*, 2002). CDSP32 is a critical component in the defense system against lipid peroxidation in the photosynthetic apparatus (Broin and Rey, 2003). In the present citrus data mining effort, we have found two EST contigs coding for homologs of CDSP32 (Table S12). Reads similar to *CDSP32* were absent from citrus water-stressed libraries. However, a slight down-regulation of these transcripts was observed in *C. reticulata*, *C. sinensis* and *Poncirus trifoliata* that are infected with CVC or CTV. Its role in a general plant defense mechanism prompts a detailed investigation.

Recently, ROS signals have been demonstrated to possess a certain degree of specificity and selectivity, which allows them to act efficiently in a variety of developmental processes and environmental responses (Gadjev *et al.*, 2006). The chemical nature of ROS and/or their subcellular site of production could be critical for the specificity and selectivity of these signals. During abiotic stress conditions, ROS accumulation has been hypothesized to consist in toxic stress by-products and signal transduction molecules responsible for the activation of defense mechanisms (Mittler, 2002). Dehydration-induced transcription of ROS scavenging proteins has been demonstrated in the desiccation-tolerant plant *Craterostigma plantagineum* and in the model system *A. thaliana* (Mittler, 2002) and is thought to be mediated by abscisic acid signaling (Zhang *et al.*, 2006). ABA and ROS treatment induce the expression of antioxidant genes and all activities of the antioxidant enzymes catalase, ascorbate peroxidase, glutathione reductase and superoxide dismutase via a mitogen-activated protein kinase (MAPK)-mediated pathway (Zhang *et al.*, 2006).

Our survey of citrus transcriptome has identified cDNAs showing sequence similarity to ROS scavenging enzymes in the majority of the libraries available in the CitEST database, including those from distinct developmental stages, pathogen-attacked tissues and water stress.

In *C. sinensis* and *C. reticulata*, peroxidase-like reads are identified by BLAST searches and are equally frequent in libraries obtained from fruit developmental stages, pathogen-attacked and water stressed tissue. Interestingly, in *P. trifoliata,* the frequency of peroxidase homologs appears to be biased; glutathione (ATGPX) and L-ascorbate (APX) homologs were identified in fruit development libraries, whereas peroxiredoxin (PRPX) and APX-like sequences were found in libraries obtained from pathogen-attacked tissues (Figure 7). Water stress appears to induce the transcription of homologs of all three peroxidase families in *P. trifoliata*, with a smaller prevalence of APX-like mRNAs (Figure 7). Thus, at this point a consistent association between a specific sub-set of oxidative stress metabolism enzymes and water stress-induced ROS scavenging in citrus remains to be established, although *APX*-like genes appear to be suitable candidates for responding to environmentally caused water loss.

Concluding Remarks

Water stress is arguably the most serious constraint to agriculture (Araus *et al.*, 2002). However, stress tolerance by genetic modification is difficult to achieve due to the involvement of complex traits in plants. The manipulation of certain classes of proteins, especially those having a direct protective role, could reveal good candidates to improve plant stress tolerance (Wang *et al.*, 2003).

Various genes are known to be stress-induced (Seki *et al.*, 2002). Expression profiling has become an important tool to investigate how an organism responds to environmental changes and, subsequently how these transcriptional changes may thus define both tolerant and sensitive responses (Hazen *et al.*, 2003). The difference between tolerance and susceptibility may arise from the regulation of a basic set of genes (Xiong and Zhu, 2002; Taji *et al.*, 2004; Kant *et al.*, 2006). In the *Citrus* transcriptome survey, we have identified 305 genes implicated in different functions known to be intrinsically related to plant water stress tolerance (Figure 8).

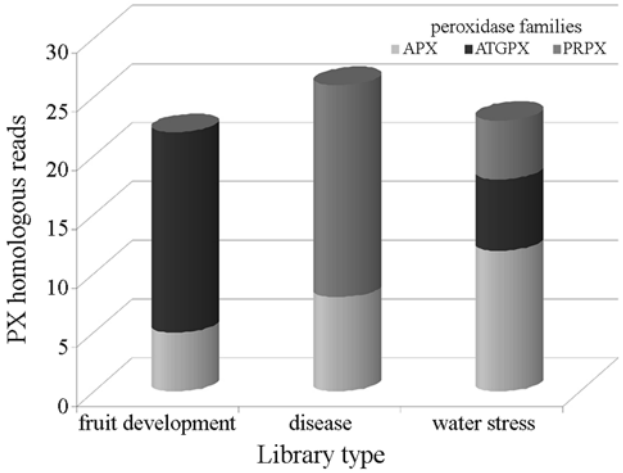

Figure 7 - Expression of peroxidase (PX) families in *Poncirus trifoliata* tissues from different developmental stages and stress conditions. The y-axis represents the number of reads showing sequence similarity to PX family members in BLAST searches. Citrus libraries were grouped according to the treatment: fruit development (700 series), disease (100, 200 and 300 series) and water stress (500 series). APX, L-ascorbate peroxidase; ATGPX, *Arabidopsis thaliana* glutathione peroxidase; PRPX, peroxiredoxin peroxidase.

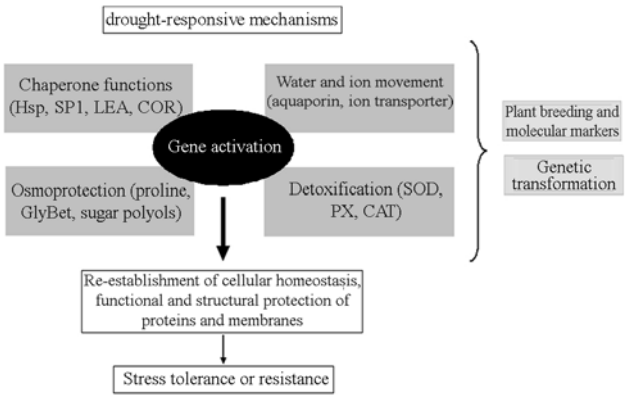

Figure 8 - Drought responsive mechanisms in plants leading to acquired stress tolerance. Stress-responsive mechanisms reestablish homeostasis, protect and repair damaged proteins and membranes. Genetic engineering and conventional plant breeding combined with the use of molecular markers and quantitative trait loci (QTLs) provide invaluable tools to achieve acquired tolerance to abiotic stress. Abbreviations: CAT, catalase; CDPK, calcium-dependent protein kinase; COR, cold-responsive protein; Hsp, heat shock protein; LEA, late embryogenesis abundant; PX, peroxidase; ROS, reactive oxygen species; SOD, superoxide dismutase; SP1, stable protein 1.

Apart from the complex regulatory network between drought and other environmental stresses, a tight correlation amongst distinct classes of proteins has been demonstrated. For example, the overexpression of *BiP* (a scarce transcript in citrus libraries) was apparently associated to a decreased responsiveness of antioxidative enzymes under water deficit, notably in relation to SOD, identified in a large number of *CitEST* libraries (Alvim *et al.*, 2001). New LEA proteins hypothesized to function as radical scavengers under oxidative stress conditions imposed by the dehydration have recently been described (Hara *et al.*, 2004). In contrast, mitochondria-localized LEAs could exert a complementary role to osmolytes, which mainly participate in cytosol stabilization (Grelet *et al.*, 2005). Another aspect that should be considered is the capacity of the plant to recover from an imposed constraint. The regulation of certain transcripts could be an indication of this capacity, as it is the case for aquaporins in salt-stressed rice roots (Kawasaki *et al.*, 2001).

Finally, it has been demonstrated that the capacity to cope with adverse conditions is strongly dependent on a stress-anticipatory preparedness in tolerant species (Gong *et al.*, 2005). Therefore, the approach adopted in this study will bring about trends to be pursued in the functional characterization and transgenic analyses of the identified components.

References

Alamillo JM and Bartels D (2001) Effects of desiccation on photosynthesis pigments and the ELIP-like dsp 22 protein complexes in the resurrection plant *Craterostigma plantagineum*. Plant Sci 160:1161-1170.

Alexandersson E, Fraysse L, Sjovall-Larsen S, Gustavsson S, Fellert M, Karlsson M, Johanson U and Kjellbom P (2005) Whole gene family expression and drought stress regulation of aquaporins. Plant Mol Biol 59:469-484.

Allen R (1995) Dissection of oxidative stress tolerance using transgenic plants. Plant Physiol 107:1049-1054.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.

Alvim FC, Carolino SMB, Cascardo JCM, Nunes CC, Martinez CA, Otoni WC and Fontes EPB (2001) Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress. Plant Physiol 126:1042-1054.

Apel K and Hirt H (2004) Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373-399.

Araus JL, Slafer GA, Reynolds MP and Royo C (2002) Plant breeding and drought in C3 cereals: What should we breed for? Ann Bot (Lond) 89:925-940.

Arbona V, Marco AJ, Iglesias DJ, López-Climent MF, Talon M and Gómez-Cadenas A (2005) Carbohydrate depletion in roots and leaves of salt-stressed potted *Citrus clementina* L. Plant Growth Regul 46:153-160.

Babu RC, Zhang J, Blum A, Ho T-HD, Wu R and Nguyen HT (2004) *HVA1*, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. Plant Sci 166:855-862.

Bartels D (2001) Targeting detoxification pathways: An efficient approach to obtain plants with multiple stress tolerance? Trends Plant Sci 6:284-286.

Bartels D and Salamini F (2001) Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*. A contribution to the study of drought tolerance at the molecular level. Plant Physiol 127:1346-1353.

Bianchi G, Gamba A, Murelli C, Salamini F and Bartels D (1991) Novel carbohydrate metabolism in the resurrection plant *Craterostigma plantagineum*. Plant J 1:355-359.

Bowler C and Fluhr R (2000) The role of calcium and activated oxygens as signals for controlling cross-tolerance. Trends Plant Sci 5:241-246.

Brakke M and Allen LH (1995) Gas exchange of citrus seedlings at different temperatures, vapor-pressure deficits, and soil water content. J Am Soc Hort Sci 120:497-504.

Bray EA (1997) Plant responses to water deficit. Trends Plant Sci 2:48-54.

Bray EA, Bailey-Serres J and Weretilnyk E (2000) Responses to abiotic stresses. In: Buchnau BB, Gruissem W and Jones RL (eds) Biochemistry and Molecular Biology of Plants. American Society of Plant Physiologists, Rockville, pp 1158-1203.

Broin M and Rey P (2003) Potato plants lacking the CDSP32 plastidic thioredoxin exhibit overoxidation of the BAS1 2-cysteine proxiredoxin and increased lipid peroxidation in thylakoids under photooxidative stress. Plant Physiol 132:1335-1343.

Broin M, Cuiné S, Eymery F and Rey P (2002) The plastidic 2-cysteine peroxiredoxin is a target for a thioredoxin involved in the protection of the photosynthetic apparatus against oxidative damage. Plant Cell 14:1417-1432.

Broin M, Cuiné S, Peltier G and Rey P (2000) Involvement of CDSP32, a drought-induced thioredoxin, in the response to oxidative stress in potato plants. FEBS Lett 467:245-248.

Capell T and Christou P (2004) Progress in plant metabolic engineering. Curr Opin Biotechnol 15:148-154.

Carpenter JF, Crowe JH and Arakawa T (1990) Comparison of solute-induced protein stabilization in aqueous solution and in the frozen and dried states. J Dairy Sci 73:3627-3636.

Chaumont F, Barrieu F, Herman EM and Chrispeels MJ (1998) Characterization of a maize tonoplast aquaporin expressed in zones of cell division and elongation. Plant Physiol 117:1143-1152.

Chen Y, Qu YQ and Jia X (2003) The characters and gene expression of rice seed proteins. Hereditas 25:367-372.

Close TJ (1996) Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins. Plant Physiol 97:795-803.

Cohen S, Moreshet S, Le Guillou L, Simon JC and Cohen M (1997) Response of citrus trees to modified radiation regime in semi-arid conditions. J Exp Bot 48:35-44.

Cuming A (1999) LEA proteins. In: Shewry PR and Casey R (eds) Seed Proteins. Kluwer Academic Publishers, Dordrecht, pp 753-780.

Déjardin A, Sokolov LN and Kleczkowski LA (1999) Sugar/osmoticum levels modulate differential abscisic acid-independent expression of two stress-responsive sucrose synthase genes in *Arabidopsis*. Biochem J 344:503-509.

DeRocher A and Vierling E (1995) Cytoplasmic HSP70 homologues of pea: Differential expression in vegetative and embryonic organs. Plant Mol Biol 27:441-456.

Dgany O, Gonzalez A, Sofer O, Wang W, Zolotnitsky G, Wolf A, Shoham Y, Altman A, Wolf SG, Shoseyov O, *et al.* (2004) The structural basis of the thermostability of SP1, a novel plant (*Populus tremula*) boiling stable protein. J Biol Chem 279:51516-51523.

Dure L III and Croudh M (1981) Developmental biochemistry of cotton seed embryogenesis, and termination: Changing messenger ribonucleic and populations as shown by *in vitro* and *in vivo* protein synthesis. Biochem 20:4162-4168.

Eisen MB, Spellman PT, Brown PO and Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863-14868.

Eklund H, Gleason FK and Holmgren A (1991) Structural and functional relations among thioredoxins of different species. Proteins Struct Funct Genet 11:13-28.

Finnegan PM, Whelan J, Millar AH, Zhang Q, Smith MK, Wiskich JT and Day DA (1997) Differential expression of the multigene family encoding the soybean mitochondrial alternative oxidase. Plant Physiol 114:455-466.

Foyer CH and Noctor G (2005) Oxidant and antioxidant signalling in plants: A re-evaluation of the concept of oxidative stress in a physiological context. Plant Cell Environ 28:1056-1071.

Fray RG, Wallace A, Grierson D and Lycett GW (1994) Nucleotide sequence and expression of a ripening and water stress-related cDNA from tomato with homology to the MIP class of membrane channel proteins. Plant Mol Biol 24:539-543.

Gadjev I, Vanderauwera S, Gechev TS, Laloi C, Minkov IN, Shulaev V, Apel K, Inzé D, Mittler R and Van Breusegem F (2006) Transcriptomic footprints disclose specificity of reactive oxygen species signaling in Arabidopsis. Plant Physiol 141:436-445.

Garay-Arroyo A, Colmenero-Flores JM, Garciarrubio A and Covarrubias AA (2000) Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. J Biol Chem 275:5668-5674.

Geigenberger P, Reimholz R, Deiting U, Sonnewald U and Stitt M (1999) Decreased expression of sucrose phosphate synthase strongly inhibits the water stress-induced synthesis of sucrose in growing potato tubers. Plant J 25:119-129.

Godde D (1999) Adaptations of the photosynthetic apparatus to stress conditions. In: Lerner HR (ed) Plant Responses to Environmental Stresses. From Phytohormones to Genome Reorganization. Marcel Dekker, New York, pp 449-474.

Gong Q, Li P, Ma S, Rupassara SI and Bohnert HJ (2005) Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. Plant J 44:826-839.

Gottesman S, Squires C, Pichersky E, Carrington M, Hobbs M, Mattick JS, Dalrymple B, Kuramitsu H, Shiroza T, Foster T, *et al.* (1990) Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. Proc Natl Acad Sci USA 87:3513-3517.

Grelet J, Benamar A, Teyssier E, Avelange-Macherel MH, Grunwald D and Macherel D (2005) Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. Plant Physiol 137:157-167.

Han B and Kermode AR (1996) Dehydrin-like proteins in castor bean seeds and seedlings are differentially produced in response to ABA and water-deficit-related stresses. J Exp Bot 47:933-939.

Hara M, Fujinaga M and Kuboi T (2004) Radical scavenging activity and oxidative modification of citrus dehydrin. Plant Physiol Biochem 42:657-662.

Hasegawa PM, Bressan RA, Zhu JK and Bohnert HJ (2000) Plant cellular and molecular responses to high salinity. Annu Rev Plant Physiol Plant Mol Biol 51:463-499.

Hazen SP, Wu Y and Kreps JA (2003) Gene expression profiling of plant responses to abiotic stress. Funct Integr Genom 3:105-111.

Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J and Millar AH (2004) Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. Plant Cell 16:241-256.

Himelblau E and Amasino RM (2000) Delivering copper within plant cells. Curr Opin Plant Biol 3:205-210.

Holmstrom K, Mantyla E, Welin B, Mandal A, Palva ET, Tunnela OE and Londesborough J (1996) Drought tolerance in tobacco. Nature 25:683-684.

Hong Z, Lakkineni K, Zhang Z and Verma DP (2000) Removal of feedback inhibition of delta(1)-pyrroline-5-carboxylate synthetase results in increased proline accumulation and protection of plants from osmotic stress. Plant Physiol 122:1129-1136.

Hong-Bo S, Zong-Suo L and Ming-An S (2005) LEA proteins in higher plants: Structure, function, gene expression and regulation. Colloids and Surf B Biointerfaces 45:131-135.

Hsieh TH, Lee JT, Charng YY and Chan MT (2002a) Tomato plants ectopically expressing Arabidopsis CBF1 show enhanced resistance to water deficit stress. Plant Physiol 130:618-626.

Hsieh TH, Lee JT, Yang PT, Chiu LH, Charng YY, Wang YC and Chan MT (2002b) Heterology expression of the Arabidopsis *C-repeat/dehydration response element binding factor 1* gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. Plant Physiol 129:1086-1094.

Ingram J and Bartels D (1996) The molecular basis of dehydration tolerance in plants. Annu Rev Plant Physiol Plant Mol Biol 47:377-403.

Ingram J, Chandler JW, Gallagher L, Salamini F and Bartels D (1997) Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to sugar interconversions associated with dehydration in the resurrection plant *Craterostigma plantagineum* Hochst. Plant Physiol 115:113-121.

Inzé D and van Montagu M (1995) Oxidative stress in plants. Curr Opin Biotechnol 6:153-158.

Ito Y, Saisho D, Nakazono M, Tsutsumi N and Hirai A (1997) Transcript levels of tandem-arranged alternative oxidase genes in rice are increased by low temperature. Gene 203:121-129.

Jang JY, Kim DG, Kim YO, Kim JS and Kang H (2004) An expression analysis of a gene family encoding plasma membrane aquaporins in response to abiotic stresses in *Arabidopsis thaliana*. Plant Mol Biol 54:713-725.

Johanson U, Karlsson M, Johansson I, Gustavsson S, Sjovall S, Fraysse L, Weig AR and Kjellbom P (2001) The complete set of genes encoding major intrinsic proteins in Arabidopsis provides a framework for a new nomenclature for major intrinsic proteins in plants. Plant Physiol 126:1358-1369.

Kant S, Kant P, Raveh E and Barak S (2006) Evidence that differential gene expression between the halophyte, *Thellungiella halophila* and *Arabidopsis thaliana* is responsible for higher levels of the compatible osmolyte proline and tight control of Na^+ uptake in *T. halophila*. Plant Cell Environ 29:1220-1234.

Kavi Kishor PB, Hong Z, Miao G-H, Hu C-AA and Verma DPS (1995) Over-expression of [delta]-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. Plant Physiol 108:1387-1394.

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D and Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell 13:889-905.

Kiyosue T, Yamaguchi-Shinozaki K and Shinozaki K (1994) Cloning of cDNAs for genes that are early-responsive to dehydration stress (ERDs) in *Arabidopsis thaliana* L.: Identification of three ERDs as HSP cognate genes. Plant Mol Biol 25:791-798.

Kleines M, Elster RC, Rodrigo MJ, Blervacq A-S, Salamini F and Bartels D (1999) Isolation and expression analysis of two stress-responsive sucrose synthase genes from the *Craterostigma plantagineum* (Hochst). Planta 25:13-24.

Koning AJ, Rose R and Comai L (1992) Developmental expression of tomato heat-shock cognate protein 80. Plant Physiol 100:801-811.

Konstantinova T, Parvanova D, Atanassov A and Djilianov D (2002) Freezing tolerant tobacco, transformed to accumulate osmoprotectants. Plant Sci 163:157-164.

Kovtun Y, Chiu WL, Tena G and Sheen J (2000) Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. Proc Natl Acad Sci USA 97:2940-2945.

Krishna P and Gloor G (2001) The Hsp90 family of proteins in *Arabidopsis thaliana*. Cell Stress Chaperones 6:238-246.

Lee JH and Schöffl F (1996) An *Hsp70* antisense gene affects the expression of HSP70/HSC70, the regulation of HSF and the acquisition of thermotolerance in transgenic *Arabidopsis thaliana*. Mol Gen Genet 252:11-19.

Levine A (1999) Oxidative stress as a regulator of environmental responses in plants. In: Lerner HR (ed) Plant Responses to Environmental Stresses. From Phytohormones to Genome Reorganization. Marcel Dekker, New York, pp 247-264.

Lloyd J and Howie H (1989) Salinity, stomatal responses and whole-tree hydraulic conductivity of orchard Washington navel orange, *Citrus sinensis* (L.) Osbeck. Aust J Plant Physiol 16:169-179.

Mariaux JB, Bockel C, Salamini F and Bartels D (1998) Desiccation- and abscisic acid-responsive genes encoding major intrinsic proteins (MIPs) from the resurrection plant *Craterostigma plantagineum*. Plant Mol Biol 38:1089-1099.

Marrs KA (1996) The functions and regulation of glutathione S-transferases in plants. Annu Rev Plant Physiol Plant Mol Biol 47:127-158.

Marrs KA, Casey ES, Capitant SA, Bouchard RA, Dietrich PS, Mettler IJ and Sinibaldi RM (1993) Characterization of two maize HSP90 heat shock protein genes: Expression during heat shock, embryogenesis, and pollen development. Devel Genet 15:27-41.

Maurel C and Chrispeels MJ (2001) Aquaporins: A molecular entry into plant water relations. Plant Physiol 125:135-138.

Maxwell DP, Wang Y and McIntosh L (1999) The alternative oxidase lowers mitochondrial reactive oxygen production in plant cells. Proc Natl Acad Sci USA 96:8271-8276.

McClung CR (1997) Regulation of catalases in *Arabidopsis*. Free Radic Biol Med 23:489-496.

Michael TP and McClung CR (2002) Phase-specific circadian clock regulatory elements in *Arabidopsis*. Poster abstracts of the 13th International Conference on Arabidopsis Research. Abstract 9-16.

Miernyk JA (1997) The 70 kDa stress-related proteins as molecular chaperones. Trends Plant Sci 2:180-187.

Mikolajczyk M, Awotunde OS, Muszynska G, Klessig DF and Dobrowolska G (2000) Osmotic stress induces rapid activation of a salicylic acid-induced protein kinase and a homolog of protein kinase ASK1 in tobacco cells. Plant Cell 12:165-178.

Milioni D and Hatzopoulos P (1997) Genomic organization of *hsp90* gene family in Arabidopsis. Plant Mol Biol 35:955-961.

Mittler R (2002) Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci 7:405-410.

Mittler R, Vanderauwera S, Gollery M and Van Breusegem F (2004) The reactive oxygen gene network in plants. Trends Plant Sci 9:490-498.

Moreshet S, Cohen Y, Green GC and Fuchs M (1990) The partitioning of hydraulic conductances within mature orange trees. J Exp Bot 41:833-839.

Munns R (2002) Comparative physiology of salt and water stress. Plant Cell Environ 25:239-250.

Noh SJ, Kwon CS, Oh DH, Moon JS and Chung WI (2003) Expression of an evolutionarily distinct novel *BiP* gene during the unfolded protein response in *Arabidopsis thaliana*. Gene 311:81-91.

Oliver MJ, Dowd SE, Zaragoza J, Mauget AS and Payton PR (2004) The rehydration transcriptome of the desiccation-tolerant bryophyte *Tortula ruralis*: Transcript classification and analysis. BMC Genomics 5:89-135.

Ortolani AA, Pedro Júnior MJ and Alfonsi RR (1991) Agroclimatologia e o cultivo de citros. In: Rodrigues O, Viégas F, Pompeu Júnior J and Amaro AA (eds) Citricultura Brasileira. 2nd edition. Fundação Cargill, Campinas, pp 153-195.

Pastori GM and Foyer CH (2002) Common components, networks, and pathways of cross-tolerance to stress: The central role of "redox" and abscisic acid-mediated controls. Plant Physiol 129:460-468.

Peeters N and Small I (2001) Dual targeting to mitochondria and chloroplasts. Biochim Biophys Acta 1541:54-63.

Pelah D, Wang W, Altman A, Shoseyov O and Bartels D (1997) Differential accumulation of water-stress related proteins, sucrose synthase and soluble sugars in *Populus* species that differ in their water stress response. Physiol Plant 25:153-159.

Pilon-Smits EAH, Ebskamp MJM, Paul MJ, Jeuken MJW, Weisbeek PJ and Smeekens SCM (1995) Improved performance of transgenic fructan-accumulating tobacco under drought stress. Plant Physiol 107:125-130.

Pilon-Smits EAH, Terry N, Sears T, Kim H, Zayed AM, Hwang SB, van Dun K, Voogd E, Verwoerd TC, Krutwagen WHH, *et al.* (1998) Trehalose-producing transgenic tobacco plants show improved growth performance under drought stress. J Plant Physiol 152:525-532.

Pilon-Smits EAH, Terry N, Sears T and van Dun K (1999) Enhanced drought resistance in fructan-producing sugar beet. Plant Physiol Biochem 37:313-317.

Pruvot G, Cuine S, Peltier G and Rey P (1996) Characterization of a novel drought-induced 34-kDa protein located in the thylakoids of *Solanum tuberosum* L. plants. Planta 198:471-479.

Ramanjulu S and Bartels D (2002) Drought- and desiccation-induced modulation of gene expression in plants. Plant Cell Environ 25:141-151.

Rey P, Pruvot G, Becuwe N, Eymery F, Rumeau D and Peltier G (1998) A novel thioredoxin-like protein located in the chloroplast is induced by water deficit in *Solanum tuberosum* L. plants. Plant J 13:97-107.

Saisho D, Nakazono M, Lee KH, Tsutsumi N, Akita S and Hirai A (2001) The gene for alternative oxidase-2 (AOX2) from *Arabidopsis thaliana* consists of five exons unlike other AOX genes and is transcribed at an early stage during germination. Genes Genet Syst 76:89-97.

Saisho D, Nambara E, Naito S, Tsutsumi N, Hirai A and Nakazono M (1997) Characterization of the gene family for alter-

native oxidase from *Arabidopsis thaliana*. Plant Mol Biol 35:585-596.

Salomé PA and McClung CR (2002) Characterization of the *Arabidopsis* catalase family. Poster abstracts of the 13th International Conference on *Arabidopsis* Research. Abstract 7-49.

Schena M, Shalon D, Davis RW and Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467-470.

Schneider K, Wells B, Schmelzer E, Salamini F and Bartels D (1993) Desiccation leads to the rapid accumulation of both cytosolic and chloroplastic proteins in the resurrection plant *Craterostigma plantagineum* Hochst. Planta 189:120-131.

Schwacke R, Grallath S, Breitkreuz KE, Stransky E, Stransky H, Frommer WB and Rentsch D (1999) *Le*ProT1, a transporter for proline, glycine betaine and γ-amino butyric acid in tomato pollen. Plant Cell 25:377-391.

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y and Shinozaki K (2001) Monitoring the expression pattern of *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13:61-72.

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, *et al.* (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31:279-292.

Sheveleva E, Chmara W, Bohnert HJ and Jensen RG (1997) Increased salt and drought tolerance by d-ononitol production in transgenic *Nicotiana tabacum* L. Plant Physiol 25:1211-1219.

Shinozaki K and Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: Differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol 3:217-223.

Shou H, Bordallo P and Wang K (2004a) Expression of the *Nicotiana* protein kinase (NPK1) enhanced drought tolerance in transgenic maize. J Exp Bot 55:1013-1019.

Shou H, Bordallo P, Fan JB, Yeakley JM, Bibikova M, Sheen J and Wang K (2004b) Expression of an active tobacco mitogen-activated protein kinase kinase kinase enhances freezing tolerance in transgenic maize. Proc Natl Acad Sci USA 101:3298-3303.

Sivamani E, Bahieldin A, Wraith JM, Al-Niemi T, Dyer WE, Ho TD and Qu R (2000) Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. Plant Sci 155:1-9.

Smirnoff N (1998) Plant resistance to environmental stress. Curr Opin Biotechnol 9:214-219.

Sun W, Montagu MV and Verbruggen N (2002) Small heat shock proteins and stress tolerance in plants. Biochim Biophys Acta 1577:1-9.

Sung DY, Kaplan F and Guy CL (2001a) Plant Hsp70 molecular chaperones: Protein structure, gene family, expression and function. Physiol Plant 113:443-451.

Sung DY, Vierling E and Guy CL (2001b) Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. Plant Physiol 126:789-800.

Swan TM (1997) Membrane fatty acid composition and membrane fluidity as parameters of stress tolerance in yeast. Can J Microbiol 43:70-77.

Taji T, Seki M, Satou M, Sakurai T, Kobayashi M, Ishiyama K, Narusaka Y, Narusaka M, Zhu JK and Shinozaki K (2004) Comparative genomics in salt tolerance between Arabidopsis and Arabidopsis-related halophyte salt cress using Arabidopsis microarray. Plant Physiol 135:1697-1709.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The CLUSTALX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882.

Torres GAM, Pflieger S, Corre-Menguy F, Mazubert C, Hartmann C and Lelandais-Brière C (2006) Identification of novel drought-related mRNAs in common bean roots by differential display RT-PCR. Plant Sci 171:300-307.

Tournaire-Roux C, Sutka M, Javot H, Gout E, Gerbeau P, Luu DT, Bligny R and Maurel C (2003) Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins. Nature 425:393-397.

Urao T, Yakubov B, Satoh R, Yamaguchi-Shinozaki K, Seki M, Hirayama T and Shinozaki K (1999) A transmembrane hybrid-type histidine kinase in *Arabidopsis* functions as an osmosensor. Plant Cell 11:1743-1754.

Vanlerberghe GC and McIntosh L (1997) Alternative oxidase: From gene to function. Annu Rev Plant Physiol Plant Mol Biol 48:703-734.

Verslues PE, Agarwal M, Katiyar-Agarwal S, Zhu J and Zhu JK (2006) Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. Plant J 45:523-539.

Vidal V, Ranty B, Dillenschneider M, Charpenteau M and Ranjeva R (1993) Molecular characterization of a 70 kDa heat-shock protein of bean mitochondria. Plant J 3:143-150.

Vierling E (1991) The roles of heat shock proteins in plants. Annu Rev Plant Physiol Plant Mol Biol 42:579-620.

Vinocur B and Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: Achievements and limitations. Curr Opin Biotechnol 16:123-132.

Wang W, Vinocur B and Altman A (2003) Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. Planta 218:1-14.

Wang W, Vinocur B, Shoseyov O and Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9:244-252.

Wang WX, Dgany O, Wolf SG, Levy I, Algom R, Pouny Y, Wolf A, Marton I, Altman A and Shoseyov O (2006) Aspen SP1, an exceptional thermal, protease and detergent resistant self-assembled nano-particle. Biotechnol Bioeng 95:161-168.

Wang WX, Pelah D, Alergand T, Shoseyov O and Altman A (2002) Characterization of SP1, a stress-responsive, boiling-soluble, homo-oligomeric protein from Aspen. Plant Physiol 130:865-875.

Whelan J, Smith MK, Meijer M, Yu JW, Badger MR, Price GD and Day DA (1995) Cloning of an additional cDNA for alternative oxidase in tobacco. Plant Physiol 107:1469-1470.

Wurgler-Murphy SM and Saito H (1997) Two-component signal transducers and MAPK cascades. Trends Biochem Sci 22:172-176.

Xiong L and Zhu J-K (2002) Molecular and genetic aspects of plant responses to osmotic stress. Plant Cell Environ 25:131-139.

Xu D, Duan X, Wang B, Ho T-HD and Ray W (1996) Expression of a late embryogenesis abundant protein gene *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol 110:249-257.

Yabe N, Takahashi T and Komeda Y (1994) Analysis of tissue-specific expression of *Arabidopsis thaliana* HSP90-family gene *HSP81*. Plant Cell Physiol 35:1207-1219.

Yamada A, Sekiguchi M, Mimura T and Ozeki Y (2002) The role of plant CCTalpha in salt- and osmotic-stress tolerance. Plant Cell Physiol 43:1043-1048.

Yamaguchi-Shinozaki K, Koizumi M, Urao S and Shinozaki K (1992) Molecular cloning and characterization of 9 cDNAs for genes that are responsive to desiccation in *Arabidopsis thaliana*: Sequence analysis of one clone that encodes a putative transmembrane channel protein. Plant Cell Physiol 33:217-224.

Yoon H-S, Lee H, Lee I-A, Kim K-Y and Jo K (2004) Molecular cloning of the monodehydroascorbate reductase gene from *Brassica campestris* and analysis of its mRNA level in response to oxidative stress. Biochim Biophys Acta 1658:181-186.

Yoshiba Y, Kiyosue T, Nakashima K, Yamaguchi-Shinozaki K and Shinozaki K (1997) Regulation of levels of proline as an osmolyte in plants under water stress. Plant Cell Physiol 38:1095-1102.

Young JC, Moarefi I and Hartl FU (2001) Hsp90: A specialized but essential protein folding tool. J Cell Biol 154:267-273.

Zhang A, Jiang M, Zhang J, Tan M and Hu X (2006) Mitogen-activated protein kinase is involved in abscisic acid-induced antioxidant defense and acts downstream of reactive oxygen species production in leaves of maize plants. Plant Physiol 141:475-487.

Zhang X-P and Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends Plant Sci 7:14-21.

Zhu BC, Su J, Chan MC, Verma DPS, Fan Y-L and Wu R (1998) Over-expression of a [delta]-pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water-stress and salt-stress in transgenic rice. Plant Sci 25:41-48.

Internet Resources

Citrus Biotechnology Laboratory, http://citest.centrodecitricultura.br (September 13, 2006).

Cluster v.2.11 Software, http://rana.lbl.gov/EisenSoftware.htm.

DNASTAR Lasergene Software, http://www.dnastar.com/web/index.php.

European Bioinformatics Institute-European Molecular Biology Laboratory (EMBL-EBI), www.ebi.ac.uk/interpro/ (September 4, 2006).

Expert Protein Analysis System (ExPaSy), http://www.expasy.org/prosite/ and http://www.us.expasy.org/sprot/ (October 5, 2006).

International Water Management Institute (IWMI), http://www.iwmi.cgiar.org/pubs/AReps/2004_2005/home.htm (October 27, 2006).

National Center for Biotechnology Information (NCBI) http://www.ncbi.nlm.nih.gov/BLAST/ (October 27, 2006).

PAUP* 4.0b10 Software, http://paup.csit.fsu.edu/.

Protein Families (Pfam), http://www.sanger.ac.uk/Software/Pfam/ (October 15, 2006).

PSIGNFIT Software, http://www.bootstrap-software.org/.

Tree View v.1.6 Software, http://rana.lbl.gov/EisenSoftware.htm.

Supplementary Online Material

Table S1 - Citrus ESTs with homology to functionally characterized genes involved in ion transporters from *Arabidopsis thaliana* and other model species.

Table S2 -Citrus ESTs with homology to the major intrinsic protein (MIP) gene family from *Arabidopsis thaliana* and other model species.

Table S3 -Citrus ESTs with homology to genes involved in osmolyte biosynthesis.

Table S4 -Citrus ESTs with homology to the heat shock proteins (HSP) gene family from *Arabidopsis thaliana* and other model species.

Table S5 - Citrus ESTs with homology to the late embryogenesis abundant proteins (LEA) gene family from *Arabidopsis thaliana* and other model species.

Table S6 -Citrus ESTs with homology to the SP1-related coding genes from *Arabidopsis thaliana*.

Table S7 - Citrus ESTs with homology to superoxide dismutase protein families from *Arabidopsis thaliana*.

Table S8 - Citrus ESTs with homology to peroxidase protein families from *Arabidopsis thaliana*.

Table S9 - Citrus ESTs with homology to ascorbate-glutathione cycle-related protein families from *Arabidopsis thaliana*.

Table S10 - Citrus ESTs with homology to alternative oxidase protein families from *Arabidopsis thaliana*.

Table S11 - Citrus ESTs with homology to catalase protein families from *Arabidopsis thaliana*.

Table S12 - Citrus ESTs with homology to thioredoxin protein from *Solanum tuberosum*.

Supplementary References

Are provided with the electronic version of this article.

This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Marco Aurélio Takita

Table S1 – Citrus ESTs with homology to functionally characterized genes involved in ion transporters from *Arabidopsis thaliana* and other model species.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
Ion-transporting ATPases						
ACA1 and *ACA* family from other species	AT1G27770	C6-CS/CL (2)	21.2	1e-149	calcium-transporting ATPase 1, plasma membrane-type	Reddy and Reddy, 2002
		C16-CS/CR (4)	19.6	1e-137		
		C31-CS (3)	82.3	1e144		
		CS00-C3-704-035-H04-CT	26.8	1e-121		
		CS00-C3-702-063-G02-CT	20.4	1e-117		
*AHA1*and *AHA* family from other species	AT2G18960	C17-CR (3)	13.4	1e-91	plasma membrane-type, proton-exporting ATPase, E1-E2 type; cation transporter	Houlne and Boutry, 1994
		C20-CS (2)	16.4	1e-144		
		CS00-C3-703-054-E07-CT	82.4	1e-133		
		CR05-C3-702-002-D12CT	80.2	1e-105		
NHX1 and *H^+/ATPase* from	AT5G27150	C8-CS/CR (5)	23.7	3e-96	Na^+/H^+-antiporter, monovalent cation:proton	Maser *et al.*, 2001
		C28-PT (3)	24.3	1e-128		

other species		C33-LT (2)	24.6	5e-98	antiporter (CPA1) family	
		CR05-C3-700-023-F11-UV	21.9	1e-67	member	
NHX2	AT3G05030	C32-CG (2)	23.0	1e-29	Na^+/H^+antiporter family	Maser *et al.*, 2001
NHX6	AAM08407	CS00-C3-704-005-F12-CT	76.2	3e-85	Na^+/H^+antiporter family	Yokoi *et al.*, 2002
		CR05-C3-702-002-D12-CT	84.9	2e-97		
VHA-A	AT1G78900	C22-CS (13)	81.6	1e-180	vacuolar H^+-ATPase	Kluge *et al.*, 2003
		CS00-C1-451-005-H09-CT	18.2	2e-58	catalytic subunit, H^+-transporting two-sector ATPase, α/β subunit	
Ion transporters and transporter-associated proteins						
CBL1 and *CBL* family from other species	AT4G17615	C3-CS/PT (2)	54.9	1e-68	calcineurin B-like protein, calcium binding, calcium ion binding, N-terminal protein myristoylation	Guo *et al.*, 2002
		C5-CS/CR (2)	27.9	1e-51		
		C7-CS (2)	58.8	1e-72		
		C23-CG (4)	19.1	2e-74		
		PT11-C1-900-068-H12-CT	40.7	1e-102		
		CG32-C1-003-082-D10-CT	31.1	1e-101		

		CG32-C1-003-094-E09-CT	16.4	8e-60		
		PT11-C9-005-022-G08-CT	19.3	4e-42		
		CL06-C4-500-037-E11-CT	15.6	3e-35		
CBL2	AT5G55990	C25-PT (2)	50.9	1e-77	calcineurin B-like protein	Guo *et al.*, 2002
		C29-PT (3)	41.4	5e-41		
CBL6	AT4G16350	CS13-C1-001-016-H11-CT	57.2	2e-88	calcineurin B-like protein	Guo *et al.*, 2002
		CS00-C3-705-076-F08-CT	51.5	1e-78		
CBL9	AT5G47100	C1-CS (5)	19.4	1e-82	calcineurin B-like protein	Guo *et al.*, 2002
FLA10	AT3G46550	C13-CR/LT (2)	27.3	1e-70	Fasciclin-like arabino-galactan protein 10 precursor, KUP/HAK/KT transporter family	Very and Sentenac, 2002
HAK5	AT4G13420	C10-CS (2)	30.4	4e-97	KUP/HAK/KT transporter family member, similar to HAK2 (*Hordeum vulgare*)	Rus *et al.*, 2004
KT2	AT2G40540	C18-CR (3)	20.9	2e-97	KUP/HAK/KT transporter	Rus *et al.*, 2004

		C19-CA (3)	19.8	1e-113	family member	
KUP1/HAK1/KT	AT2G30070	C12-PT (2)	25.1	3e-56	KUP/HAK/KT transporter	Rus *et al.*, 2004
and K^+		C21-CS (4)	50.9	1e-164	family	
transporters from		C30-PT (2)	19.9	3e-83		
other species		CS00-C1-101-025-B09-CT	31.0	1e-83		
		CR05-C1-100-068-D06CT	23.0	5e-64		
		CA26-C1-002-088-H06CT	22.2	9e-64		
		CG32-C1-003-054-E05CT	15.3	7e-63		
		CS00-C3-703-059-F03-CT	11.5	6e-37		
		CS00-C3-704-029-B10-CT	26.4	4e-42		
		CR05-C3-701-065-D07CT	19.1	1e-35		
		CS12-C1-001-029-E04-CT	10.7	2e-22		
KUP3	AT3G02050	C2-CS (2)	29.5	3e-21	KUP/HAK/KT transporter	Rus *et al.*, 2004
		C4-CR (4)	71.4	4e-74	family	
		C15-CR/PT (5)	23.9	3e-65		
		C24-CS/CR (2)	83.1	1e-169		

KUP7	AT5G09400	C14-CS/CR (2)	54.9	3e-63	KUP/HAK/KT transporter family	Rus *et al.*, 2004
KUP10	AT1G31120	C27-PT (3)	26.2	1e-151	KUP/HAK/KT transporter family	Rus *et al.*, 2004
PAA2	AT5G21930	C34-LT/CR (2)	21.6	6e-85	P-Type ATPase, haloacid dehalogenase-like hydrolase, E1-E2 type, heavy metal transport/detoxification protein , copper-translocating	Shikanai *et al.*, 2003
SOS2/CIPK24	AT5G35410	PT11-C1-900-068-H12-CT	61.1	4e-96	calcineurin B-like protein, calcium binding, calcium ion binding, N-terminal protein myristoylation	Guo *et al.*, 2002
		CR05-C1-100-039-F07-CT	58.6	7e-93		
		CS00-C1-102-017-H10-CT	57.5	2e-90		
SOS3	AT5G24270	C9-CS (5)	77.7	1e-109	calcineurin B-like protein	Guo *et al.*, 2002

		CG32-C1-003-082-D10CT	65.9	2e-55		
		CS13-C1-001-016-H11-CT	62.1	4e53		
SOS4	AT5G37850	C11-CS/CR (4)	79.6	1e-112	calcineurin B-like protein	Guo *et al.*, 2002
		C26-CR (4)	58.1	1e-94		
SOS5/FLA8	AT3G46550	CR05-C1-100-074-A04-CT	51.3	4e-51	calcineurin B-like protein	Guo *et al.*, 2002

[a]Gene name abbreviations: *ACA*: autoinhibited Ca^{2+} -ATPase, *AHA*: plasma membrane H^+-ATPase, *CBL*: calcineurin B-like, CIPK: serine-threonine protein kinases interact with CBL proteins, *FLA*: fasciclin-like arabino-galactan protein, *HAK*: high-affinity K^+ transporter ,*KT*: potassium transporter , *KUP*: K^+ uptake protein, *NHX*: Na^+/H^+ exchanger, *PAA*: P-type ATPase, *SOS*: salt overly sensitive, *VHA-A*: vacuolar-type H^+-ATPase subunit A.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S2 –Citrus ESTs with homology to the major intrinsic protein (MIP) gene family from *Arabidopsis thaliana* and other model species.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
Plasma Membrane Intrinsic Protein (PIP) family						
PIP1;1	AT3G61430	C16-CS/LT (3)	21.2	1e-130	major intrinsic protein family, two tandem repeats containing three membrane-spanning domains and a pore-forming loop (signature motif Asn-Pro-Ala/Thr -NPA), isoform 1, substrate inespecificity	Heymann and Engel, 1999
		CA26-C1-002-058-D07-CT	24.9	6e-31		
PIP1;2	AT2G45960	C17-CS/PT (2)	17.7	1e-123	major intrinsic protein family, isoform 1, substrate inespecificity	Heymann and Engel, 1999
		CS00-C3-700-067-F12-CT	78.0	1e-123		
PIP1;4	AT4G00430	C4-CS (2)	45.4	1e-105	major intrinsic protein family, isoform 1, substrate	Heymann and Engel, 1999

					inespecificity	
PIP1;5	AT4G23400	C2-CS (22)	81.4	1e-109	major intrinsic protein family, isoform 1, substrate inespecificity	Heymann and Engel, 1999
PIP2;4	AT5G60660	C1-CS/CR (3)	60.5	1e-117	major intrinsic protein family, isoform 2, water channel activity (substrate specificity for water)	Heymann and Engel, 1999
		C3-CS (23)	58.1	1e-110		
		C6-CS (5)	39.4	1e-121		
		C10-CS (9)	40.2	1e-110		
		C18-CS (2)	21.8	1e-54		
PIP2;5	AT2G16850	C9-CS (18)	82.2	1e-132	major intrinsic protein family, isoform 2, water channel activity	Heymann and Engel, 1999
		C13-CR (3)	19.0	1e-108		
		CR05-C1-102-051-A11-CT	20.5	2e-20		
PIP2;7	AT3G54820	C11-CS (10)	74.7	1e-144	major intrinsic protein family, isoform 2, water channel activity	Heymann and Engel, 1999
Tonoplast Intrinsic Proteins (TIP) family						

TIP1;1	AT2G36830	C5-CS (8)	78.9	1e-119	tonoplast intrinsic proteins, isoform α (expression in seeds), unspecific substrate: water, amino acids and/or peptides	Chaumont *et al.*, 1998; Jauh *et al.*, 1998
TIP2;1	AT3G26520	C7-CS (6)	19.3	1e-98	tonoplast intrinsic proteins, isoform α (expressed in the roots), unspecific substrate: water, amino acids and/or peptides	Chaumont *et al.*, 1998; Jauh *et al.*, 1998
		PT11-C9-005-028-H06-CT	62.1	2e-91		
NOD26-like proteins (NIP) family						
NIP1;2	AT4G18910	C14-CS (20)	20.1	1e-77	similar to nodulin-26, a major component of the peribacteroid membrane induced during nodulation in legume roots after *Rhizobium* infection, glycerol as substrate	Baiges *et al.*, 2002
		CR05-C1-102-066-G09-CT	49.7	7e-61		
		CS00-C5-003-029-B07-CT	17.2	2e-50		

NLM2	AT4G18910	C8-CS (5)	54.5	5e-71	NOD-26 like major intrinsic protein	Baiges *et al.*, 2002
NIP4;2	AT5G37820	C12-CS (6)	17.8	2e-38	NOD-26 like major intrinsic protein	Baiges *et al.*, 2002
Small Basic Intrinsic Protein (SIP) family						
SIP1;1	AT3G04090	C15-PT (2)	10.5	1e-55	small basic intrinsic proteins	Johanson *et al.*, 2001;
		CS00-C1-100-020-F04-CT	15.6	3e-66	sub-family, small proteins,	Ishikawa *et al.*, 2005
		CS00-C3-700-023-B07-CT	18.1	3e-64	similar to TIPs, basic like PIPs and many of the NLMs	

[a]Gene name abbreviations: *PIP*: plasma membrane intrinsic protein, *TIP*: tonoplast intrinsic protein, *NIP:* nodulin26-like protein, *NLM*: nodulin26-likemajor intrinsic protein, *SIP*: small basic intrinsic protein.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S3 –Citrus ESTs with homology to genes involved in osmolyte biosynthesis.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
BADH	AT1G74920	C1-CS (2)	82	e-126	oxidation of betaine aldehyde,	Wood *et al.*, 1996
		CS00-C1-102-011-B04-CT	75	2e-97	the second step of gly betaine	
					biosynthesis	
CMO	AT4G29890	C1-CR (2)	68	3e-53	catalyzes the primary,	Rathinasabapathi *et al.*,
					regulatory step of GB	1997
					biosynthesis by oxidizing	
					choline to betaine aldehyde	
INVA	AT1G12240	C2-CA/CR/CS (4)	87	0.0	beta-fructofuranosidase	Van den Ende *et al.*, 2002
		C3-CA/CL/CS/PT (5)	90	0.0	(invertase), which accumulate	
		C1-CR/CS (5)	99	e-169	as soluble polypeptides in the	
					vacuole; degradation of sucrose	
					metabolism	
NADP-	AT2G21250	C1-CA/CG/CR/CS (17)	86	e-159	mannose biosynthesis	Everard *et al.*, 1997

M6PR		CS00-C3-701-029-D03-CT	61	1e-90		
P5CR	AT5G14800	C1-CR/CS (6)	78	e-117	second reaction in proline biosynthesis	Kiyosue *et al.*, 1996
P5CS	AT2G39800	C2-CA/CG/CL/CR/CS/PT(54)	82	0.0	first reaction in proline biosynthesis	Kiyosue *et al.*, 1996
		C1-CS (3)	88	e-123		
		CS00-C3-702-024-F04-CT	82	1e-58		
PROD	AT3G30775	C2-CR (3)	39	7e-46	first enzyme of proline catabolism	Kiyosue *et al.*, 1996
		C1-CS (2)	44	4e-51		
SAMDC	AT5G15950	C1-CR/CS/PT (18)	93	e-140	biosynthesis of. the polyamines such as spermine and spermidine	Franceschetti *et al.*, 2001
		C4-CG/CS/PT (23)	97	0.0		
		C3-CS (4)	60	4e-98		
		C2-CS (4)	75	e-117		
		CS00-C1-650-029-H09-CT	56	7e-59		
SPS	AT5G20280	C1-CR/CS/PT (8)	98	0.0	catalyses of the last regulated step in sucrose synthesis	Heim *et al.*, 1996
		C5-CS/PT (6)	68	e-119		
		C3-CR/CS (2)	84	e-162		

		C4-CS/PT (4)	68	e-161		
		CR05-C3-700-084-A11-CT	95	4e-59		
SUSY	AT3G43190	C4-CG/CL/CR/CS/LT/PT(158)			catalyses the cleavage of	Komatsu *et al.*, 2002
		C3-CA/CR/CS/PT (32)	99	0.0	sucrose (reversible) in the	
		C1-CS (5)	99	0.0	presence of UDP into UDP-	
		CR05-C1-100-059-D04-CT	67	e-161	glucose and fructose	
			82	e-113		
T6PS	AT4G17770	C2-CA/CR/CS/LT (10)	82	0.0	trehalose biosynthesis	Vogel *et al.*, 2001
		C3-CA/CR/CS/PT (16)	78	0.0		
		C4-CS (4)	83	0.0		
		PT11-C2-301-032-B06-CT	85	3e-73		
		PT11-C1-900-081-D08-CT	62	2e-79		
		CA26-C1-002-098-D02-CT	90	5e-61		

[a]Gene name abbreviations: *BADH*: betaine-aldehyde dehydrogenase; *CMO*: choline monooxygenase; *INVA*: acid invertase; *NADP-M6PR*: NADPH-dependent mannose 6-phosphate reductase; *P5CR*: Pyrroline-5-carboxylate reductase; *P5CS*: Delta 1-pyrroline-5-carboxylate synthetase; *PROD*: proline dehydrogenase; *SAMDC*: S-adenosylmethionine decarboxylase; *SPS*: sucrose-phosphate synthase; *SUSY*: sucrose synthase; *T6PS*: trehalose phosphatase.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S4 –Citrus ESTs with homology to the heat shock proteins (HSP) gene family from *Arabidopsis thaliana* and other model species.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
HSP60 family	BSCTP-1	C1-CL/CR/CS/PT (18) *BsCCTα*[d]	93.0	0	Encodes a putative cytoplasmic chaperonin (CCT1), T-complex protein 1 alpha subunit (TCP-1-alpha)	Yamada *et al.*, 2002
HSP70 family	AT5G28540	C16-CA/CR/CS/LT/PT (19) *BiP-1*	91.9	0	Response to heat, response to virus, protein folding, ATP binding function	Sorin *et al.*, 2006
		PT11-C9-005-035-C11-CT(*BiP-1*)	85.4	e-116		
		C3-CA/CG/CR/CS/PT (29) *CPHSC70-1*	86.3	0		
		C18-CR/CS (3) *CPHSC70-1*	88.0	e-171		
		C11-CR/CS (2) *CPHSC70-1*	65.1	e-100		
		C23-CA/CG/CL/CR/CS/LT/PT (42) *HSC70-1*	94.0	0		
		C21-CA/CG/CL/CS/PT (25) *HSC70-1*	92.3	0		
		C2-CA/CS (2) *HSC70-1*	92.7	e-120		

		C22-CG/CR/CS/PT (11) *HSP91*	84.8	0		
		C17-CR/CS (10) *HSP91*	84.9	e-117		
		C8-CS (2) AT1G16030	88.5	0		
		C7-CG/CS/PT (12) AT3G12580	93.0	0		
		C5-CS/PT (5) AT3G12580	82.5	0		
		C19-CA/CG/CR/CS/PT (16) AT3G12580	93.5	e-153		
		C4-CL/CR/CS/LT (13) AT3G12580	89.5	1e-84		
		C15-CA/CR/CS/LT (15) AT3G12580	89.8	7e-98		
		C10-CS (6) AT3G12580	90.5	e-126		
		C20-PT (4) AT3G12580	90.6	6e-97		
		C9-CA/CR/CS/PT (14) AT4G37910	44.5	3e-53		
HSP90 family	AT5G52640	C3-CA/CG/CS (8) *HSP81-1*	92.6	0	response to heat, response to arsenic, induced by IAA and NaCl	Yabe *et al.*, 1994; Milioni and Hatzopoulos, 1997
		CR05-C3-700-048-F08-CT(*HSP81-1*)	76.8	e-117		
		C5-				

		CA/CG/CL/CR/CS/LT/PT (92) *HSP81-2*	93.5	0		
		C9-PT (6) *HSP81-2*	92.7	0		
		C7-CA/CR/CS/LT (51) *HSP81-2*	99.7	e-168		
		C6-CA/CR/CS/LT/PT (43) *HSP81-2*	86.3	e-133		
		C1-CS (2) *HSP81-2*	66.4	5e-84		
		PT11-C1-901-009-C09-CT(*HSP81-2*)	94.3	3e-52		
		CS00-C1-100-114-E05(*HSP81-3*)	63.0	1e-43		
HSP100 family	AT5G50920	C1-CR/CS/PT (9) *ATHSP101*	88.0	0	members of the AAA^+ family of ATPases that mediate ATP-dependent protein unfolding reactions, may have ATP-dependent peptidase activity	Zheng *et al.*, 2002; Tran *et al.,* 2004
		C4-CR/CS/LT/PT (29) *CLPC*	94.0	0		
		C5-CR/CS/PT (4) *CLPC*	74.0	e-132		
		C3-CR/CS/LT (4) *ERD1*	77.0	e-155		
		CG32-C1-003-008-A07-CT(*ERD1*)	79.3	e-116		
		CA26-C1-002-054-G04-CT(AT2G25140)	63.0	1e-68		
		PT11-C1-900-095-B06-				

		CT(AT2G25140)	80.2	4e-99		
		C2-CG/CL/CR/CS/PT (34) AT5G15450	82.0	0		
sHSP family	P05478	C2-CS (3) *HSP17.4*	77.2	3e-69	17.4-22 kDa class I or class II heat shock protein, response to heat, response to oxidative stress	Kim *et al.*, 2005
		C3-CS (9) *HSP17.9-D*	81.9	2e-71		
		C4-CG/CR/CS (13) *HSP17.9-D*	81.9	2e-71		
		C6-PT (5) *HSP17.9-D*	81.9	1e-71		
		C9-CS/PT (7) *HSP18.2*	77.0	7e-64		
		C8-PT (5) *HSP18.2*	77.9	1e-68		
		C1-CG/CR/CS/PT (7) *HSP18.2*	77.3	4e-68		
		C10-PT (8) *HSP18.5-C*	75.3	1e-66		
		C5-CS (4) *HSP18.5-C*	74.7	3e-67		
		C7-PT (2) *HSP22*	65.7	1e-78		

[a]Gene name abbreviations: *BsCCTα*: *Bruguiera sexangula* chaperonin containing TCP-1 (TCP-1, t-complex peptide-1), *BiP-1*: binding protein 1, *CPHSC70-1*: chloroplast heat shock cognate protein 70-1, *HSC70-1*: heat shock cognate protein 70-1, *HSP91*: heat shock protein 91, *ATHSP101*: *Arabidopsis thaliana* heat shock protein 101, *CLPC*: Clp-C/HSP100 molecular chaperone, *ERD1*: early responsive to dehydration 1, *HSP22*: heat shock protein 22 from *Petunia* x *hybrida*.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads), *Locus name*.

[c]Identity percentage at the amino acid level.

Table S5 – Citrus ESTs with homology to the late embryogenesis abundant proteins (LEA) gene family from *Arabidopsis thaliana* and other model species.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST[b]**	**%[c]**	**e value**	**biological process**	**References**
COR19 family	BAA74736	C9-PT (5) COR11	100.0	9e-59	domain of dehydrin with a K-segment similar to that of gymnosperms and in having a serine cluster (S-segment) at an unusual position at the C-terminus, LEA group 2, response to cold, drought and flooding, may act as a radical scavenging protein	Cai *et al.*, 1995; Hara *et al.*, 2004; Sanchez-Ballesta *et al.*, 2004
		C3-CS (7) COR15	94.9	3e-74		
		CR05-C3-700-068-A07-CT(COR15)	58.8	2e-23		
		CS00-C3-700-045-G09-CT(CuCOR19)	66.7	2e-18		
		CS00-C3-704-017-A06-CT(CuCOR19)	64.3	1e-25		
		CR05-C1-100-002-D10-CT(CuCOR19)	63.1	2e-26		
CsDHN	AAN78125	PT11-C1-901-093-B12-CT	57.1	8e-38	common angiosperm-type dehydrin domain K-segment	Porat *et al.*, 2004

					consensus sequence, completely different genes from the previously defined group of citrus dehydrins, their expression in the fruit peel tissue is down-regulated by many environmental stresses, such as wounding, UV irradiation, water stress, and exposure to ethylene and low oxygen concentrations	
LEA group 1 protein	AT5G06760	C5-CS (3) PT11-C9-005-028-G06-CT	61.0 56.0	3e-47 1e-22	LEA group 1 domain-containing protein, involved in embryonic development, molecular function unknown, expressed during dry seed stage	Tai *et al.*, 2005

LEA14_GOSHI	LEA14-A	C2-CG/CR/CS (32) C14-PT (3) C15-PT (2)	76.8 74.8 74.8	2e-63 3e-62 3e-62	Belongs to the LEA type 2 family (group 4), induced by water stress in leaves	Galau *et al.*, 1993
LEA group 5 protein	AT3G22490	PT11-C9-005-012-F05-CT	58.8	1e-65	domain of seed maturation protein, accumulated specifically in mature seeds, might be involved in maturation and desiccation tolerance of seeds	Yang *et al.*, 1996
PgEMB8	EMB8	C4-CR/CS (5) C11-PT (7) CR05-C1-102-039-G10-CT CS00-C1-650-026-B10-CT	65.0 62.1 49.6 75.0	e-119 e-108 3e-38 4e-24	Domains of alpha/beta hydrolase fold and esterase/lipase/thioesterase, expressed during embryonic development	Dong and Dunstan, 1999
PsLEAm	CAF32327	C1-PT (6)	38.0	2e-35	LEA domain (group 3), located in mitochondrial, not expressed	Grelet *et al.*, 2005

							in vegetative tissues, induced by ABA application and water stress, could participate in the stabilization of mitochondrial matrix proteins in the dry state and hence contribute to desiccation tolerance of the seed

[a]Gene name abbreviations: *Cor*: cold-responsive; *CsDHN*: *Citrus sinensis* dehydrin; *LEA*: late embryogenesis abundant; *LEA14_GOSHI*: late embryogenesis abundant protein-14 from *Gossypium hirsutum*; *PgEMB8*: *Picea glauca* embryogenesis-associated protein; *PsLEAm*: *Pisum sativum* mitochondrial late embryogenesis abundant protein.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads), Gene name.

[c]Identity percentage at the amino acid level.

Table S6 –Citrus ESTs with homology to the SP1-related coding genes from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
SP1	AT3G17210	C1-CA/CR/CS (4)	75.5	2e-43	stable protein 1-related, similar	Wang *et al.*, 2002; Dgany *et*
		PT11-C1-901-096-C07-CT	76.1	7e-43	to stable protein 1	*al.*, 2004; Lytle *et al.*, 2004
					(GI:13445204) from *Populus*	
					tremula, has a ferredoxin-like	
					fold, there are strong	
					interactions between each two	
					molecules creating a	
					stable dimer	

[a]Gene name abbreviations: *SP1*: stable protein 1.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S7 – Citrus ESTs with homology to superoxide dismutase protein families from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
CSD family	AT1G08830	C3-LT/CS/CR (17) CSD1	83.6	2e-71	copper/zinc superoxide dismutase,	Van Camp *et*
		C5-CA/CR (3) CSD1	83.5	8e-72	cytoplasm, response to oxidative	*al.*, 1990;
		C2-CS (3) CSD1	82.9	4e-71	stress, removal of superoxide	Drazkiewicz *et*
		C9-CG (3) CSD2	72.7	2e-83	radicals	*al.*, 2004
		C6-CR/PT (5) CSD2	70.8	2e-83		
		C7-PT (3) CSD1	55.9	3e-82		
		C1-CS/CR (50) CSD1	40.1	7e-64		
		C8-PT (5) CSD1	40.1	6e-67		
		C4-CS/CR (2) CSD1	39.5	7e-36		
		CA26-C1-002-101-C10-CT (CSD1)	82.9	1e-75		
		CS00-C3-705-086-B06-CT (CSD3)	55.5	3e-69		

FSD family	AT4G25100	C6-CS/CR/PT/LT (37) FSD1	73.6	1e-104	iron superoxide dismutase,	Van Camp *et*
		C2-CS/PT (10) FSD1	73.1	1e-103	chloroplast, mitochondrion,	*al.*, 1990;
		C1-CS (12) FSD1	73.1	1e-104	removal of superoxide radicals	McKersie *et*
		C4-CR (3) FSD1	62.3	3e-86		*al.*, 2000
		C3-CR (2) FSD1	56.6	6e-73		
		C5-CS/CR (3) FSD3	44.9	3e-64		
		CS00-C3-705-009-C04-CT (FSD1)	42.2	7e-50		
MSD family	AT3G10920	C3-CS/CR/PT/LT (38) MSD1	74.0	1e-100	manganese superoxide dismutase,	Van Camp *et*
		C1-CS (7) MSD1	31.6	1e-103	mitochondrion, removal of	*al.*, 1990;
		C2-CR (4) MSD1	31.6	1e-104	superoxide radicals	Gadjev *et al.*,
		C4-PT (6) MSD1	31.6	1e-104		2006
		PT11-C1-901-050-B02-CT (MSD1)	40.8	1e-45		

[a]Gene name abbreviations: *CSD*: cooper/Zinc superoxide dismutase, *FSD:* iron superoxide dismutase, *MSD*: manganese superoxide dismutase.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads), Gene name.

[c]Identity percentage at the amino acid level.

Table S8 – Citrus ESTs with homology to peroxidase protein families from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
APX family	AT1G07890	C2-CS/CR (11) APX2	84.9	1e-126	cytosolic ascorbate peroxidase,	Santos *et al.*,
		C6-PT (15) APX1	83.6	1e-123	scavenges hydrogen peroxide in the	1996; Milla *et*
		C3-CS/CR/CA/LT (51)			cytosol and chloroplasts, induction	*al.*, 2003
		APX1	82.4	1e-122	of heat shock proteins	
		C5-CS/CR/PT (29) APX3	76.3	1e-126		
		C4-CS/CR (2) APX1	58.4	8e-84		
		C1-CS/PT (3) APX2	45.8	1e-146		
		CS00-C1-100-011-F12-CT				
		(APX1)	69.6	1e-97		
		CS00-C3-702-011-B08-CT				
		(APX2)	65.7	5e-93		
ATGPX	AT2G25080	C2-CS/CR (10) ATGPX2	74.6	7e-71	phospholipid hydroperoxide	Bartling *et al.*,
family		C8-CS/CR (2) ATGPX2	74.6	2e-70	glutathione peroxidase, chloroplast,	1993; Bianchi

		C10-PT (24) ATGPX2	67.5	2e-77	response to oxidative stress	*et al.*, 2002
		C6-CS/CR (10) ATGPX2	66.9	5e-78		
		C1-CS/CR/CG (22) ATGPX1				
		C9-CS/PT (4) ATGPX2	66.5	6e-88		
		C7-CS/CR/PT (8) ATGPX2	65.1	6e-86		
		C11-LT/CG (3) ATGPX1	61.5	5e-69		
		C4-CS/CR (3) ATGPX2	58.1	6e-83		
		C3-CS (3) ATGPX2	57.4	2e-69		
		C5-CR (2) ATGPX2	39.6	3e-31		
			33.7	3e-21		
PER/PRXR	AT1G48130	C16-CS/CR (100)	81.2	1e-158	haem peroxidase, plant peroxidase	Jespersen *et*
family		C7-CS (11)	64.1	1e-121	domain, endomembrane system,	*al.*, 1997
		C4-CS/CR/PT (6)	49.2	7e-88	response to oxidative stress	
		C3-CS (3)	44.0	1e-114		
		C14-PT (3)	40.0	3e-85		
		C10-CS/PT (6)	38.3	1e-115		

C8-CG/CS (6)	38.0	4e-99
C9-CR/PT (2)	37.0	3e-88
C12-CR/CA (2)	36.6	2e-98
C6-CS (6)	36.5	3e-76
C13-PT (2)	36.2	6e-83
C5-CS (7)	34.7	2e-94
C2-CS/CR (9)	34.4	1e-96
C1-CS/PT (2)	34.0	2e-99
C11-CR (6)	33.4	1e-102
C15CS/CR (4)	32.4	1e-90
CS00-C1-101-096-D04-EU	63.4	1e-101
PT11-C9-005-047-B10-CT	43.2	7e-64
CS00-C3-702-059-G03-CT	35.9	5e-56
PT11-C9-005-046-D11-CT	34.9	2e-83
CS00-C1-100-114-D06-CT	34.1	1e-70
CS00-C3-703-065-B12-CT	33.9	2e-69

CS00-C3-701-050-F04-CT	33.2	9e-57
PT11-C1-900-034-E10-CT	33.2	8e-67
CA26-C1-002-065-E02-CT	33.1	4e-58
PT11-C1-900-037-E12-CT	32.3	3e-55
CS00-C3-700-054-G03-CT	31.9	5e-56

[a]Gene name abbreviations: *APX*: L-ascorbate peroxidase, *ATGPX*: glutathione peroxidase, *PER*: 1-cysteine peroxiredoxin, *PRPX:* peroxiredoxin, *sAPX*: stromal L-ascorbate peroxidase, *tAPX*: thylakoid L-ascorbate peroxidase.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads), Gene name.

[c]Identity percentage at the amino acid level.

Table S9 – Citrus ESTs with homology to ascorbate-glutathione cycle-related protein families from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
MDAR	AY6626551	C1-CR (7)	82.0	1e-177	Pyridine nucleotide-disulphide	Leterrier *et al.*, 2005
	(*Pisum*	C1-CS (33)	80.0	0.0	oxidoreductase, FAD-	
	sativum)	C1-PT (8)	77.0	0.0	dependent pyridine nucleotide-	
		CA26-C1-002-010-D04-CT	66.0	3e-69	disulphide, reduces	
		CG32-C1-003-015-H05-CT	65.0	7e-74	monodehydroascorbate to	
		C1-LT (2)	72.0	1e-152	ascorbate	
DHAR	AAL71855	C1-CR (2)	70.0	6e-88	Glutathione S-transferase, C-	Chen *et al.*, 2003
		C1-CS (7)	77.0	3e-95	terminal and Glutathione S-	
		C2-CS (23)	77.0	2e-93	transferase, N-terminal, reduces	
		C3-CS (2)	77.0	1e-72	dehydroascorbate to ascorbate	
		C1-PT (9)	75.0	1e-91		
		C1-CA (4)	75.0	1e-92		
		CS00-C3-700-034-B08-CT	60.0	1e-53		

		CA26-C1-002-080-D03-CT	57.0	1e-46		
		C1-CG (6)	82.0	1e-86		
		C1-LT (2)	74.0	3e-91		
GR	AT3G24170	CR05-C3-700-020-G10-			electron transport, glutathione	Cho and Seo, 2005; Henmi
		EU(*GR*)	83.4	e-108	metabolism, located in	*et al.*, 2005
		C1-CR/CS/LT/PT (19)			cytoplasm or chloroplast, has	
		ATGR1	81.2	0	disulfide oxidoreductase	
		C4-CR/CS (5) *ATGR1*	78.4	e-113	activity, glutathione-disulfide	
		CS00-C3-700-035-A12-			reductase activity,	
		CT(*ATGR1*)	76.2	e-101	oxidoreductase activity, FAD	
					binding	
ATGSTU	AT1G17180	CS13-C1-001-003-C09-			Encodes glutathione transferase	Wagner *et al.*, 2002
		CT(*ATGSTU25*)	62.6	1e-62	belonging to the tau class of	
		C2-CS (2) *ATGSTU27*	65.1	8e-61	GSTs. Naming convention	
		C3-CR/CS/PT (7) *ATGSTU27*			according to Wagner *et al.*	
		C5-CR/CS/LT (4)	62.1	1e-72	(2002). Involved in toxin	

ATGSTU27			catabolism, located in
C6-PT (2) *ATGSTU27*	63.7	2e-62	cytoplasm, glutathione
	60.2	3e-71	transferase activity

[a]Gene name abbreviations: *DHAR*: dehydroascorbate reductase, *MDAR:* monodehydroascorbate reductase. *GR*: glutathione-reductase; *ATGSTU*: *Arabidopsis thaliana* glutathione transferase belonging to the tau class of GSTs.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads), Gene name.

[c]Identity percentage at the amino acid level.

Table S10 – Citrus ESTs with homology to alternative oxidase protein families from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
AOX1A	AT3G22370	C2-PT (2)	92.3	2e-98	mitochondrion, response to	Umbach *et al.*, 2005
		CS00-C5-003-045-D07-CT	89.4	3e-63	cold, alternative oxidase activity, cellular respiration	
AOX2	AT5G64210	C1-CR/CS/PT (8)	66.0	e-129	mitochondrial envelope,	Saisho *et al.*, 2001
		C3-PT (2)	56.0	5e-72	electron transport, alternative oxidase activity	

[a]Gene name abbreviations: *AOX*: alternative oxidase.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S11 – Citrus ESTs with homology to catalase protein families from *Arabidopsis thaliana*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**	**%**[b]	**e value**	**biological process**	**References**
CAT1	AT1G20630	C2-LT (2)	87.8	e-120	catalyzes the reduction of hydrogen peroxide using heme group as cofactor, protects cells from toxicity by H_2O_2	Hsieh *et al.*, 2002
CAT2	AT4G35090	C1-CA (3)	75.1	e-162	Encodes a peroxisomal catalase, highly expressed in bolts and leaves. mRNA expression patterns show circadian regulation with mRNA levels being high in the subjective early morning	Orendi *et al.*, 2001
		C2-CA (3)	92.4	0		
		C1-CG (8)	89.5	0		
		C2-CG (6)	76.6	3e-86		
		C1-CR (32)	78.7	e-179		
		C2-CR (57)	89.2	0		
		C1-CS (28)	78.9	0		
		C2-CS (79)	89.4	0		
		C1-LT (4)	73.4	e-173		

		C1-PT (34)	89.4	0		
		C2-PT (2)	82.3	4e-84		
		CS00-C1-102-062-F11-CT	98.1	1e-55		
		LT33-C1-003-004-D02-CT	77.9	7e-51		
CAT3	AT1G20620	C3-PT (2)	75.7	5e-81	Involved in hydrogen peroxide catabolism, catalyzes the breakdown of hydrogen peroxide (H_2O_2) into water and oxygen, located in peroxisome, has catalase activity, expressed during senescence, expressed in inflorescence and leaf	Park *et al.*, 1998

[a]Gene name abbreviations: *CAT*: catalase.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.

Table S12 – Citrus ESTs with homology to thioredoxin protein from *Solanum tuberosum*.

Arabidopsis thaliana		***CitEST***			**Protein motifs and**	
Name[a]	**Gene**	**EST**[b]	**%**[c]	**e value**	**biological process**	**References**
CDSP32	CAA71103	C1-CR/CS/LT (28)	72.8	e-117	function as a physiological electron donor to the BAS1 peroxiredoxin, has a role in plastid defense against oxidative damage	Broin *et al.*, 2002
		C2-PT (4)	69.3	e-103		

[a]Gene name abbreviations: *CDSP32*: chloroplastic drought-induced stress protein of 32 kDa from *Solanum tuberosum*.

[b]C: contig, CA: *Citrus aurantium*, CG: *Citrus aurantifolia*, CR: *Citrus reticulata*, CS: *Citrus sinensis,* LT: *Citrus latifolia,* PT: *Poncirus trifoliata*, (number of reads).

[c]Identity percentage at the amino acid level.