Genetics and Molecular Biology, 24 (1-4), 207-220 (2001)

Identifying sugarcane expressed sequences associated with nutrient transporters and peptide metal chelators

Antonio Figueira, Ederson Akio Kido and Raul Santin Almeida*

Abstract

Plant nutrient uptake is an active process, requiring energy to accumulate essential elements at higher levels in plant tissues than in the soil solution, while the presence of toxic metals or excess of nutrients requires mechanisms to modulate the accumulation of ions. Genes encoding ion transporters isolated from plants and yeast were used to identify sugarcane putative homologues in the sugarcane expressed sequence tag (SUCEST) database. Five cluster consensi with sequence homology to plant high-affinity phosphate transporter genes were identified. One cluster consensus allowed the prediction of a full-length protein containing 541 amino acids, with 81% amino acid identity to the *Nicotiana tabacum NtPT1* gene, consisting of 12 membrane-spanning domains divided by a large hydrophilic charged region. Putative homologues to *Arabidopsis thaliana* micronutrient transporter genes were also detected in some of the SUCEST libraries. Iron uptake in grasses involves the release of the phytosiderophore mugeneic acid (MA) which chelate Fe^{3+} which is then absorbed by a specific transporter. Sugarcane expressed sequence tag (EST) homologous to genes coding for three enzymes of the mugeneic acid biosynthetic pathway [nicotianamine synthase; nicotianamine transferase; and putative mugeneic acid synthetase (*ids3*)] and a putative Fe^{3+}-phytosiderophore transporter were detected. Seven sugarcane sequence clusters were identified with strong homology to members of the ZIP gene family (*ZIP1*, *ZIP3*, *ZIP4*, *IRT1* and *ZNT1*), while four clusters homologous to *ZIP2* and three to *ZAT* were found. Homologues to members of another gene family, *Nramp*, which code for broad-specificity transition metal transporters were also detected with constitutive expression. Partial transcripts homologous to genes encoding γ-glutamylcysteine synthetase, glutathione synthetase, and phytochelatin synthase (responsible for biosynthesis of the metal chelator phytochelatin) and all four types of the major plant metal-chelator peptide metallothionein (MT) were identified: Type I MT being the most abundant (>1% of seed-library reads), followed by Type II which had a similar pattern of expression as that described for *Arabidopsis* MT. Identifying and understanding the expression of genes associated with nutrient uptake and metal tolerance could lead to the development of more nutrient-efficient sugarcane cultivars, or might allow the use of sugarcane as a hyper-accumulator plant for the restoration of contaminated areas in phytoremediation programs.

INTRODUCTION

The potential of plant growth and development is limited by the ability of plants to efficiently absorb available nutrients from the soil. Worldwide, a significant portion of arable land presents some fertility constraints, either limiting concentrations of essential nutrients or toxicity, and crops frequently have to contend with limiting levels of essential nutrients in the rhizosphere, leading farmers to use fertilizers for maximum yield. Plants are the major source of micronutrients in the human diet, and mineral deficiencies (*e.g.* iron-deficiency) are important causes of major human nutritional disorders (Guerinot and Salt, 2001) while an excess of some minerals (*e.g.* zinc and copper) can be detrimental to human health (Kochian, 2000).

Nutrient uptake by plants is an active process, requiring energy to accumulate essential elements at levels in plant tissues above concentrations found in the soil solution (Fox and Guerinot, 1998) while, conversely, elevated concentrations of essential nutrients or the presence of toxic elements require active efflux systems and/or detoxification mechanisms to minimize the accumulation of ions. The transport of ions across the plasma membrane is based on an active efflux of protons which results in a pH gradient and/or a membrane potential which drives the movement of nutrients via carriers or channels (Hirsch and Sussman, 1999). Plants have evolved systems to maintain a controlled intracellular ion homeostasis.

Phosphorous is one of the most limiting nutrients for plant growth, and the low solubility of most inorganic mineral phosphates and the high sorption capacity of soil make phosphorous the macronutrient least available to roots (Kochian, 2000). Available soil phosphate is often found at μM levels, while plant tissue concentrations are in mM ranges (Raghothama, 1999). The mechanism of phosphate uptake involves an energy-dependent proton/phosphate symport process, driven by a proton gradient generated by a plasma membrane H^+-ATPase (Raghothama, 1999). Studies on the kinetics of phosphate absorption have shown the existence of a dual mechanism, with high-affinity transporters operating at low phosphate concentrations (μM)

Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Av. Centenário 303, C.P. 96, 13400-970 Piracicaba, São Paulo, Brazil.
**Send correspondence to Antonio Figueira. E-mail: figueira@cena.usp.br.*

and low-affinity transporters at high phosphate concentrations (mM) [Muchhal *et al.* 1996; Raghothama, 1999]. High-affinity phosphate transporter genes have been identified in *Arabidopsis* (Muchhal *et al.* 1996), tomato (Liu *et al.*, 1998), potato (Leggewie *et al.*, 1997) and other higher plants (Raghothama, 1999) based on functional complementation of yeast mutants defective in high-affinity transporters.

Some metal cations (*e.g.* such as zinc, copper, iron and manganese) are essential micronutrients for plants because they are required co-factors in many enzymatic systems. Micronutrient cations are highly reactive, potentially promoting the formation of active oxygen forms and uptake and transport must be maintained within limits to avoid oxidative cellular damage. Acquisition of micronutrients by plants is complicated by the limited availability of metal cations in the soil, where they have low solubility and tend to form metallo-organic complexes (Kochian, 2000).

Iron is predominantly found in soils in its oxidized ferric (Fe^{+3}) form, which is extremely insoluble at neutral or basic pH. Reduced ferrous iron (Fe^{+2}) is considerably more soluble but is readily oxidized to the ferric form under aerobic conditions. Plants use two strategies to acquire iron (Hirsch and Sussman, 1999). One iron-acquisition strategy (strategy I), used by most dicotyledonous and non-graminеous monocotyledonous plants, depends on the solubilization of ferric iron by the release of protons and organic acids, which acidify the rhizosphere, followed by the action of ferric-chelate reductase, and the reduced ferrous form is then taken into the plant by a specific transporter protein. The ferric-chelate reductase gene (*FRO2*), expressed when iron levels are low, has been cloned from *Arabidopsis* by Robinson *et al.* (1999). Based on complementation of yeast double mutants for low-affinity and high-affinity iron transporter (*fet3 fet4*), a plant iron-regulated transporter cDNA from *Arabidopsis* (*IRT1*) was identified (Eide *et al.*, 1996). *IRT1* is expressed in roots, induced by Fe deprivation and co-regulated with *FRO2* (Hirsch and Sussman, 1999).

An alternative iron-acquisition strategy (strategy II), adopted by grasses, depends on the release of a small peptide-like iron chelator, mugeneic acid (MA, also called phytosiderophore) to chelate ferric iron. The MA-Fe^{+3}-complex is then absorbed by a specific transporter protein into the plant (Mori, 1999). In barley, mugeneic acid biosynthesis is induced by iron-deficiency, starting with the formation of nicotianamine from 3 molecules of S-adenosyl methionine by nicotianamine synthase (NAS; EC 2.5.1.43). Nicotianamine then receives an amino group from nicotianamine aminotransferase (NAAT) and is reduced to deoxymugeneic acid and hydroxylated to mugeneic acid by a putative mugeneic acid synthetase (*ids3*). Genes encoding NAS (*nas*) and NAAT (*naat-A* and *naat-B*) have been cloned using partial peptide sequences from barley roots grown under iron-deficient conditions, while putative MA synthetase (*ids3*) was identified by differential hybridization of cDNAs from barley roots (Herbik *et al.*, 1999; Higuchi *et al.*, 1999; Mori, 1999; Takahashi *et al.*, 1999). A putative Fe^{+3}-phytosiderophore transporter has been cloned from a maize *yellow stripe 1* (*ys1*) mutant (Curie *et al.*, 2001).

The isolation of *IRT1* gene from *Arabidopsis thaliana* has allowed the identification of a new family of micronutrient transporters (Fox and Guerinot, 1998; Kochian, 2000). Sequence similarities between *IRT1* and yeast open reading frames of unknown function led to the identification of the *ZRT1* and *ZRT2*, which encode high-affinity and low-affinity zinc transporters, and this gene family was named ZIP for ZRT IRT-related Proteins (Grotz *et al.*, 1998). IRT1 was later demonstrated to have a broad specificity and was able to transport manganese and zinc (Korshunova *et al.*, 1999). Functional complementation of *zrt1 zrt2* mutant yeast by the expression of *Arabidopsis* cDNA sequences allowed the cloning of the *ZIP1*, *ZIP2* and *ZIP3*, which restored zinc-limited growth of the *zrt1/zrt2* yeast. Another member, *ZIP4*, was identified in *Arabidopsis* by sequence homology in GenBank, although, it was not able to complement the *zrt1 zrt2* yeast mutant. At least 24 members of the ZIP family (including 11 members from plants) have been identified and categorized into two subfamilies (Guerinot and Eide, 1999). Pence *et al*. (2000) cloned the *ZNT1* gene (homologous to the *ZIP4* gene) from the zinc-hyperaccumulating plant *Thlaspi caerulescens* using functional complementation in yeast and was able to show that ZNT1 was able to mediate both high-affinity zinc and low-affinity cadmium uptake.

The distribution of Zn into intracellular compartments depends on another family of metal transporters, the Cation Diffusion Facilitator (CDF) family, which has been identified in various eukaryotes, including yeasts and mammals (Guerinot and Eide, 1999). A homologue to mammalian CDF genes was identified in *Arabidopsis* (*ZAT*) and it might be associated with vacuolar sequestration of Zn (van der Zaal *et al.*, 1999). The *Nramp* gene family is another class of integral membrane proteins possibly involved with broad specificity metal transport, including zinc (Guerinot and Eide, 1999), and at least five members of this family were identified in *Arabidopsis* (Thomine *et al.*, 2000). *AtNramp1*, *AtNramp3* and *AtNramp4* were able to complement a yeast mutants defective for manganese and iron transport, and they enhanced cadmium sensitivity when expressed in yeast and plants, suggesting that members of the *AtNramp* family function as broad specificity metal transporters.

A general mechanism for metal detoxification in plants includes chelation of metal by a ligand, often followed by compartmentalization of the metal-ligand complex (Cobbett, 2000). Plants contain two major metal peptide chelators, phytochelatin (PC) and metallothionein (MT), that utilize coordination of metals by cysteine resi-

dues (Robinson, 1993; Rauser, 1999). Phytochelatins are enzymatically synthesized polypeptides consisting of γ-glutamylcysteine repetitions (*i.e.* $(\gamma\text{-Glu-Cys})_n$-Gly, where $n = 2$ to 11) with a terminal glycine (Cobbett, 2000). Phytochelatins appear to be ubiquitous in plants and have been detected in some microorganisms. It is well established that glutathione (GSH, γ-Glu-Cys-Gly) is the substrate for phytochelatin synthesis. The transpeptidation of the γ-Glu-Cys moiety of a glutathione donor onto a glutathione receptor to form a n+1 oligomer is catalyzed by γ-Glu-Cys dipeptidyl transpeptidase (phytochelatin synthase, EC 2.3.2.15) (Cobbett, 2000). Phytochelatin synthase genes have been characterized in *Arabidopsis* (Vatamaniuk *et al.*, 1999; Ha *et al.*, 1999), wheat (Clemens *et al.*, 1999) and *Schizosaccharomyces pombe* (Clemens *et al.*, 1999; Ha *et al.*, 1999). Synthesis of glutathione depends on γ-glutamylcysteine synthetase (GCS; EC 6.3.2.2) and glutathione synthetase (GS; EC 6.3.2.3), and genes encoding these enzymes have been isolated from tomato and *Arabidopsis* (Goldsbrough, 1998).

Metallothioneins, small (4 to 8 kD) cysteine-rich proteins translated from mRNAs, are required for metal tolerance in fungi and animals but their role in plants is still unclear. There are reports that metallothioneins are important for copper tolerance in *Arabidopsis* and they may be associated with metal homeostasis in plants by regulating the intracellular concentration of ions (Robinson *et al.*, 1993; Zhou and Goldsbrough, 1995; Murphy and Taiz, 1995). Expression of metallothionein genes can be affected by treatment with metals (especially copper) and a number of stress-related stimuli (Giordani *et al.*, 2000).

Identifying and understanding the expression of genes responsible for, or associated with, nutrient uptake and movement could lead to the development of more nutrient-efficient sugarcane cultivars, reducing the application of fertilizers to sugarcane crops and consequently the environmental impact of fertilizer production and application. Such cultivars may also make it possible to use marginal soils with low levels of fertility for sugarcane production. A potential application of increased metal tolerance would be the use of sugarcane for the restoration of areas contaminated by metals, a strategy known as phytoremediation. The objectives of the work presented in this paper were to identify sugarcane expressed sequences homologous to characterized genes associated with plant nutrient uptake and transport and with peptide metal ligands, such as metallothioneins and enzymes of phytochelatin biosynthetic pathway.

MATERIAL AND METHODS

Searches were conducted on the sugarcane expressed sequence tag (SUCEST) database (http://sucest.lad.dcc.unicamp.br/en/Services/services.html), using keywords and the basic local alignment search tool (BLAST) in the tblastx mode and clusters derived from the database with the fragment assembly program Phrap (Telles and Silva, this issue). The expectation value (E-value) represents the number of distinct alignments with a better score than that expected to occur purely by chance in a database search. Identified clusters with a low E-value were individually re-analyzed using the BLAST program http://www.ncbi.nlm.nih.gov/BLAST/ to confirm their identities. The cluster nucleotide consensus sequences were conceptually translated at the Open Reading Frame Finder site at http://www.ncbi.nlm.nih.gov/gorf/gorf.html. The protein sequences at the correct reading frame were aligned with homologous proteins from other organisms using the ClustalW program (version 1.8) at http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html.

RESULTS AND DISCUSSION

Phosphate transporters

High affinity phosphate transporters are highly conserved in plants and the availability of several isolated genes encoding these transporters served as the basis for searching the SUCEST database for sugarcane homologues. We identified five clusters derived from 48 reads with E-values $< 10^{-34}$ (Table 1), cluster SCCCLR1001G10.g having the best homology with all the plant high-affinity phosphate transporters. An alignment of predicted amino acid sequences of the 5 clusters indicated that all clusters shared homology with cluster SCCCLR1001G10.g, except for cluster SCEPAM2011H02.g. The sugarcane high-affinity phosphate transporter identified here had an amino acid identity with other phosphate transporters which ranged from 73% with *Mendicago* MtP1 to 81% with tobacco NtPT1 (Table 2). Amino acid identity of the phosphate transporters from *Arabidopsis thaliana*, *Solanum tuberosum* (potato), *Lycopersicon esculentum* (tomato), *Nicotiana tabacum* (tobacco), *Catharanthus roseus*, *Mendicago truncatula* and *Triticum aestivum* (wheat) are all above 75%, except for *Arabidopsis AtPT6*, which is apparently more distantly related to the other phosphate transporters (Raghothama, 2000). The sugarcane cluster consensus SCCCLR1001G10.g allowed the prediction of the full-length protein, containing 546 amino acids (Figure 1), within the range of 518 to 587 amino acids observed for fungal and plant phosphate transporters. The predicted protein from the sugarcane cluster contained all the features of typical fungal and plant phosphate transporters, *i.e.* it was a membrane-associated protein with 12 membrane-spanning domains in two groups of six separated by a large hydrophilic charged region, common features of sugar, ion, antibiotic and amino acid co-transporters (Chrispeels *et al.*, 1999). The sugarcane phosphate transporter contained potential phosphorylation and N-glycosilation sites, as observed in fungal and plant phosphate transporters.

Table I - Predicted proteins derived from sugarcane expressed sequence tag (EST) clusters with homology to various high affinity phosphate transporters.

Sugarcane		NtPT1[a]	TaPT1[b]	CrTP1[c]	LePT1[d]	AtPT2[e]
EST Clusters	Number of reads	E-values				
SCCCLR1001G10.g	27	0.0	10^{-179}	0.0	0.0	0.0
SCEPAM2011H02.g	14	10^{-80}	10^{-65}	10^{-80}	10^{-80}	10^{-83}
SCVPHR1092B07.g	3	10^{-86}	10^{-50}	10^{-84}	10^{-84}	10^{-88}
SCSBST3095B06.g	1	10^{-68}	10^{-53}	10^{-67}	10^{-68}	10^{-67}
SCEZLR1052H04.g	3	10^{-49}	10^{-34}	10^{-47}	10^{-49}	10^{-50}

Table II - Amino acid identity (%) between the sugarcane expressed sequence tag (SUCEST) cluster SCCLR1001G10.g (Sc) and high affinity phosphate transporters from other plants.

	Sc[a]	NtPT1[b]	TaPT1[c*]	CrPT1[d]	LePT1[e]	StPT1[f]	AtPT2[g]	Os[h*]	AtPT4[i]	AtPT1[j]	LePT2[k]	MtPT1[l]
Sc	100											
NtPT1	81	100										
TaPT1	81	80.3	100									
CrPT1	80	88	78	100								
LePT1	80	89	78	84	100							
StPT1	80	94	80	86	77	100						
AtPT2	77	83	75	82	80	83	100					
Os*	76	71	68	72	70	70	67	100				
AtPT1	75	80	74	79	77	80	78	70	100			
AtPT4	75	80	73	79	79	82	79	68	100	94		
LePT2	74	81	71	78	79	81	77	70	79	78	100	
MtPT1	73	77	72	77	76	78	77	69	80	79	81	100

[a]SUCEST cluster SCCCLR1001G10.g; [b]*Nicotiana tabacum* (GenBank AB20061); [c]*Triticum aestivum* (AAD26146); [d]*Catharanthus roseus* CrPT1 (BAA20522); [e]*Lycopersicum esculentum* (AAB82146); [f]*Solanum tuberosum* (CAA67395); [g]*Arabidopsis thaliana* (AAB17265), [h]*A. thaliana* (AAB88291) [i]*A. thaliana*(AAB69122); [j]*Oryza sativa* (AAF40188) [k]*L. esculentum* (AAB82147); [l]*Medicago truncatula*, (AAB81346). *partial sequence.

The SCCCLR1001G10.g cluster was assembled from 27 reads, 15 of which were derived from root and root-shoot transition zone libraries. Genes encoding high-affinity phosphate transporter are preferentially expressed in roots, the transporters being rapidly induced under phosphate deprivation (Raghothama, 2000). However, high-affinity phosphate transporters are also expressed in other plant parts (*e.g.* leaves, stems, tubers and flowers) probably associated with the transport of phosphorous within the plant (Raghothama, 1999).

A low-affinity phosphate transporter gene, *Pht2;1*, was cloned and characterized from *Arabidopsis* (Daram *et al.*, 1999). *Pht2;1* codes for a phosphate:proton symporter, with a 12-membrane spanning region, and is preferentially expressed in leaves of *Arabidopsis*. Homologues to *AtPht2;1* have been identified in rice and other plants but no *AtPht2;1* homologues was found in sugarcane.

Iron acquisition

According to Kochian (2000) strategy II is adopted for iron acquisition by most species of grass, and sequences of barley genes encoding NAS (*nas1* to *nas6*; Higuchi *et al.*, 1999; Herbik *et al.*, 1999), NAAT (*naat-A* and *naat-B*; Takahashi *et al.*, 1999) and the putative mugeneic acid synthetase gene *ids3* (Nakanishi *et al.*, 2000) were searched in the SUCEST database. Homologues to a putative Fe^{+3}-phytosiderophore transporter which has been cloned from maize by Curie *et al.* (2001) was also searched. The *Arabidopsis FRO2* and *IRT1* genes which are involved in strategy I iron acquisition in plants were also used in our searches.

We identified four NAS clusters (from 9 reads) with E-values < 10^{-24} (Table 3) and 10 NAAT clusters (from 28 reads) with E-values < 10^{-37}. Three clusters of sugarcane NAS-like sequences presented the best alignment with *nashor2* gene from barley (GenBank AF136942) characterized by Herbik *et al.* (1999), whereas only one sugarcane cluster, SCSBFL4061E06.g, showed some alignment with the *nas4* described by Higuchi *et al.* (1999) [data not shown]. The sugarcane clusters with the best E-values for NAAT-like sequences exhibited good alignment with both *naat-A* and *naat-B* (data not shown). For the putative mugeneic acid synthetase *ids3*, 14 clusters were identified with

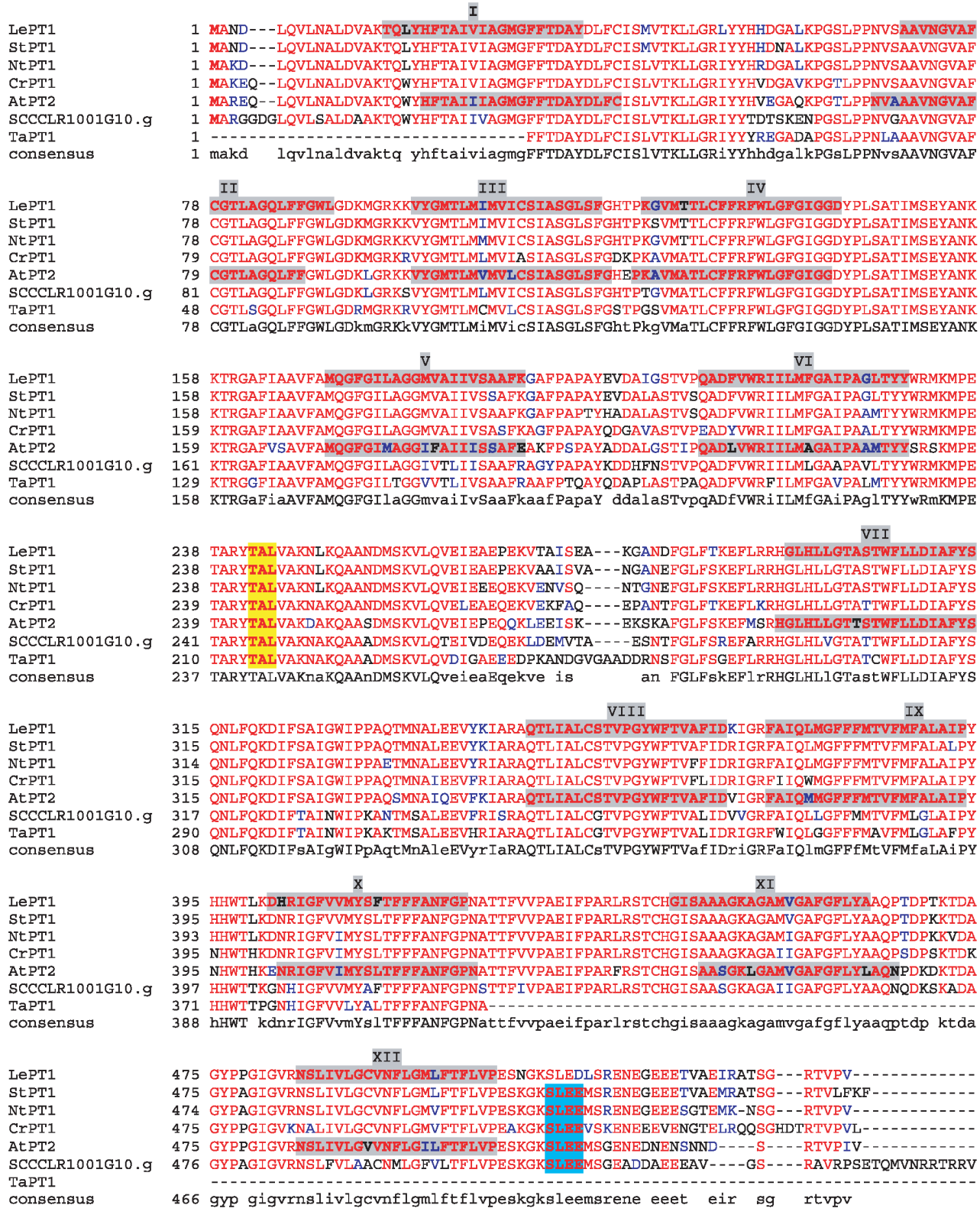

Figure 1 - Alignment of the deduced amino acid sequence of sugarcane cluster SCCCLR1001.G10.g with that of *Lycopersicum esculentum* (LePT1), *Solanum tuberosum* (StPT1), *Nicotiana tabacum* (NtPT1), *Catharanthus roseus* (CrPT1), *Arabidopsis thaliana* (AtPT2) and the partial amino acid sequence of *Triticum aestivum* (TaPT1). Identical amino acids are indicated in red and conserved substitution in blue. The 12 membrane-spanning domains are shaded gray and identified by roman numerals (I-XII). Sequences highlighted in yellow are consensus sites for phosphorylation by casein kinase II, and sequences highlighted in blue are consensus sites for phosphorylation by protein kinase C.

Table III - Sugarcane EST clusters homologues to two known enzymes involved in Strategy II iron acquisition.

Enzymes (number of amino acids)	Clusters	Number of reads	E-values	Amino acid identity	Amino acid position
Nicotianamine synthase (317-336)	SCJFRZ1006A09.g	6	10^{-111}	173/220	15 to 234
	SCJLRZ1020E11.b	1	10^{-91}	168/209	75 to 283
	SCSBFL4061E06.g	1	10^{-79}	145/178	105 to 283
	SCJFAD1C12D02.b	1	10^{-24}	50/63	75 to 137
Nicotianamine amino-transferase (461)	SCSBSB1051H05.g	2	10^{-147}	254/308	140 to 448
	SCBGLR1100G12.g	2	10^{-138}	241/286	162 to 448
	SCJLHR1029E03.g	4	10^{-92}	152/265	190 to 454
	SCCCAM1C01C10.g	4	10^{-73}	124/211	243 to 453
	SCJLRT1018B07.g	3	10^{-71}	133/234	19 to 254
	SCEZLR1009D01.g	4	10^{-58}	116/218	30 to 243
	SCSGFL5C05G01.g	2	10^{-44}	98/144	25 to 177
	SCVPRT2073H05.g	3	10^{-43}	93/144	23 to 165
	SCRFAD1023E01.g	2	10^{-39}	85/134	23 to 160
	SCRLFL3008C04.g	2	10^{-37}	64/126	259 to 384

E-value < 10^{-35} (data not shown). One cluster, SCQGHR1010F12.g, (from 10 reads) contained most of the putative protein with 69% identity with *ids3,* containing 237 out of 342 amino acids and having an E-value of 10^{-141} (Figure 2). Nineteen clusters containing 84 reads showed homology with the maize putative Fe^{+3}-phytosiderophore transporter gene *ys1* with an E-value < 10^{-22}.

In terms of expression, the sequences homologous to NAS were found at low levels in roots, flower, apical meristems and *in vitro* plantlets infected with *Gluconacetobacter diazotroficans*, while NAAT-like sequences were found primarily in roots and flowers but were also present in most of the tissues evaluated. The sugarcane homologue to the putative mugeneic acid synthetase *ids3*

```
ids3              1 MENILHATPAHVSLPESFVFASDKVP--PATKAVVSLPIIDLSCGRDEVRRSILEAGKEL
SCQGHR1010F12.g   1 MEKMLHAAPVHATLPECFIFPADKLPQAKATSATVSLPIIDLSLGRDEVRRAILDAGKEI
consensus         1 MEkmLHAaPvHatLPEcFiFpaDKlPqakATsAtVSLPIIDLSlGRDEVRRaILdAGKEi

ids3             59 GFFQVVNHGVSKQVMRDMEGMCEQFFHLPAADKASLYSEERHKPNRLFSGATYDTGGEKY
SCQGHR1010F12.g  61 GFFQVVNHGVSLEAMRDMEAVCQEFFALPAEDKAGLYSEDTGKTTRIYSSTMFDTGAEKY
consensus        61 GFFQVVNHGVSleaMRDMEavCqeFFaLPAeDKAgLYSEdtgKttRiySstmfDTGaEKY

ids3             119 WRDCLRLACPFP-VDDSINEWPDTPKGLRDVIEKFTSQTRDVGKELLRLLCEGMGIQADY
SCQGHR1010F12.g  121 WRDCLRLACSFPDVGDSPKNWPDKPARLREVVERFTVQTRGLGMQILRLLCEGLGLRPDY
consensus        121 WRDCLRLACsFPdVgDSpknWPDkParLReVvErFTvQTRglGmqiLRLLCEGlGlrpDY

ids3             178 FEGDLSGGNVILNINHYPSCPNPDKALGQPPHCDRNLITLLLPGAVNGLEVSYKGDWIKV
SCQGHR1010F12.g  181 LEGDISGGDVVLHVNHYPPCPDPTATLGLPPHCDRNLLTLLLPSMVRGLEVAYSGDWIKV
consensus        181 lEGDiSGGdVvLhvNHYPpCPdPtatLGlPPHCDRNLlTLLLPsmVrGLEVaYsGDWIKV

ids3             238 DPAPNAFVVNFGQQLEVVTNGLLKSIEHRAMTNSALARTSVATFIMPTQECLIGPAKEFL
SCQGHR1010F12.g  241 DPVPGAFVVNFGCQLEVVTNGVLKSIEHRVMTNLGVARTTVATFIMPTTDCLIGPAAEFL
consensus        241 DPvPgAFVVNFGcQLEVVTNGvLKSIEHRvMTNlgvARTtVATFIMPTtdCLIGPAaEFL

ids3             298 SKENPPCYRTTMFRDFMRIYNVVKLGSSLNLTTNLKNVQKEI
SCQGHR1010F12.g  301 SDDNPPAYRTLTFGEFKRIYSVVKLGSSLNLTTNLKDVQKEL
consensus        301 SddNPPaYRTltFgeFkRIYsVVKLGSSLNLTTNLKdVQKEl
```

Figure 2 - Alignment of the predicted amino acid sequences of one sugarcane clusters consensus and the sequence predicted for the barley *ids3* gene, a putative mugeneic acid synthetase.

was expressed mostly in seeds and roots. Interestingly, most of these barley genes were not expressed constitutively, but were inducible under conditions of iron-deprivation. Higuchi *et al.* (1999) found that when barley was grown in the presence of sufficient iron *nas1* transcripts were not detectable but were highly induced in roots grown under conditions of iron-deficiency. Based on northern analysis, Takahashi *et al.* (1999) found that *naat-B* was expressed constitutively in barley roots and that both *naat-A* and *naat-B* were induced by iron deprivation, although both transcripts were absent in leaves. Nakanishi *et al.* (2000) observed that *ids3* was only expressed in the roots of iron-deficient barley and rye, and was not detected in the roots of eight other species. Curie *et al.*, (2001) stated that maize *ys1* is constitutively expressed in roots and shoots but higher levels are induced under conditions of iron deprivation. The fact that constitutive homologues of the above genes in sugarcane were detected even at low levels of expression and under iron-sufficient growth conditions, suggests that sugarcane may have a distinct control of expression of these genes. Sequences with homology to the putative *ys1* gene were primarily observed in the sugarcane root libraries, but they were also identified in most of the other sugarcane tissues.

Sequence homologues to *FRO2* were not identified in the SUCEST database, supporting the belief that sugarcane utilizes Strategy II for iron acquisition. Sequences with homology to *IRT1* are discussed below.

Zinc transporters

Genes from the ZIP family of micronutrient transporters (ZRT-IRT-related proteins) were used to BLAST search the SUCEST database. Searches were also performed for sequences related to the *Arabidopsis ZAT* (van der Zaal *et al.*, 1999) and *Nramp* gene family (Thomine *et al.*, 2000).

Seven sugarcane sequence clusters which had high homology (E-value < 10^{-20}) to various members of the ZIP gene family were identified (Table 4), although because of the high homology between the members of the ZIP family (*ZIP1*, *ZIP3*, *ZIP4*, *ZNT1* and *IRT1*) it was difficult to classify the sugarcane clusters according to their homology to a specific member, except for *ZIP2*. Table 4 shows that sugarcane clusters SCJLLR1103B12.g and SCEZRZ 1017H11.g shared high homology with all members of the ZIP family. Several sugarcane clusters (*e.g.* SCQSST1036E04.g and SCEZRZ1015A04.g) showed higher homology to *ZIP1*, *ZIP3* and *IRT1* and less homology to *ZIP4* and *ZNT1*, while clusters SCRULB2064B09.b and SCJFST1014H01.g were homologous only to *ZIP4* and *ZNT1*. We found these ZIP-like sequences expressed in various sugarcane tissues, with the exception of root tissue. Grotz *et al.* (1998) stated that *ZIP1* and *ZIP3* were barely detectable in roots and shoots of plants growing under zinc-sufficient conditions but were induced at higher levels in zinc-deprived roots, whereas *ZIP4* was induced in zinc-deprived roots and shoots.

When the *ZIP2* sequence (which had less homology to the other members of the ZIP family) was used in the search four clusters were identified from 13 reads, the *ZIP2*-like sequences being detected in various tissues, including roots. None of the sugarcane ZIP-like sequences were complete nor did they include the amino terminal of the predicted proteins.

Three sugarcane clusters (preferentially expressed in shoots) homologous to the *Arabidopsis ZAT* gene were identified at E-values < 10^{-13}, including one cluster (SCVPLR1049G12.g) that contained the amino terminal of the ZAT protein.

Table 5 shows the ten clusters which were identified as being homologous to members of the *Nramp* family. In general, the E-values for each sugarcane cluster were similar for *AtNramp2*, *AtNramp3* and *AtNramp4*, but the inverse of that for *AtNramp1* (*i.e.* when the E-value was high for *AtNramp1* it was low for the other three and *vice versa*). Sugarcane clusters SCCCST1008B09.g SCCCRT 2002F07.g, SCSFST3080A03.g and SCCCAM2001H06.g gave the highest homologies to *AtNramp1*. Cluster SCEZLB1008B05.g was homologous only to *AtNramp1*

Table IV - Sugarcane expressed sequence tag (EST) clusters with homology to members of the *ZIP* micronutrient transporter gene family.

Sugarcane		ZIP1	ZIP3	ZIP4	ZNT1	IRT1
EST clusters	Number of reads	E-values				
SCJLLR1103B12.g	6	10^{-40}	10^{-35}	10^{-50}	10^{-50}	10^{-43}
SCEZRZ1017H11.g	4	10^{-34}	10^{-33}	10^{-59}	10^{-59}	10^{-45}
SCQSST1036E04.g	3	10^{-22}	10^{-24}	nd	nd	nd
SCEZRZ1015A04.g	1	10^{-22}	10^{-23}	nd	nd	10^{-23}
SCVPCL6063B11.g	1	nd[a]	10^{-24}	10^{-29}	10^{-30}	10^{-30}
SCRULB2064B09.b	4	nd	nd	10^{-20}	10^{-20}	nd
SCJFST1014H01.g	5	nd	nd	10^{-20}	10^{-20}	nd

[a]not detected.

Table V - Sugarcane expressed sequence tag (EST) clusters with homology to members of the *AtNramp* family of metal transporters.

Sugarcane		*AtNramp1*	*AtNramp2*	*AtNramp3*	*AtNramp4*
EST Clusters	Number of reads	E-values			
SCCCST1008B09.g	8	10^{-77}	10^{-44}	10^{-46}	10^{-46}
SCCCRT2002F07.g	2	10^{-61}	10^{-36}	10^{-36}	10^{-36}
SCSFST3080A03.g	1	10^{-56}	10^{-24}	10^{-22}	10^{-21}
SCCCAM2001H06.g	1	10^{-50}	10^{-39}	10^{-39}	10^{-40}
SCBFRZ2017F03.g	7	10^{-40}	10^{-106}	10^{-103}	10^{-98}
SCEZLB1008B05.g	1	10^{-38}	nd	nd	nd
SCJLFL1049E03.g	3	10^{-34}	10^{-102}	10^{-100}	10^{-100}
SCVPRT2079F02.g	7	10^{-27}	10^{-56}	10^{-60}	10^{-59}
SCQGAM2029C12.g	1	nd[a]	10^{-37}	10^{-35}	10^{-36}
SCBFST3135C04.g	2	nd	10^{-24}	10^{-27}	10^{-27}

[a]not detected.

while clusters SCQGAM2029C12.g and SCBFST 3135C04.g had no homology with *AtNramp1*. Clusters SCBFRZ2017F03.g, SCJLFL1049E03.g and SCVPR T2079F02.g were most homologous to *AtNramp2, AtNramp3* and *AtNramp4.* Thomine *et al.*, (2000) has shown that the predicted proteins encoded by *AtNramp2*, *AtNramp3* and *AtNramp4* are between 65 to 75% identical, whereas *AtNramp1* is more distantly related (<37% identity).

We detected *Nramp*-like sequences most frequently in the root library and, less frequently, in the flower library, but we detected them in various other tissue libraries as well. Thomine *et al.* (2000) proposed that *Nramp* genes may play a role in constitutive metal transport, and found that in *Arabidopsis AtNramp1* was preferentially expressed in the roots of plants grown *in vitro* but was also detectable in shoots, while *AtNramp3* and *AtNramp4* were equally expressed in roots and shoots. These authors also found that expression of these three genes was induced in roots by metal starvation, with only *AtNramp4* being induced in shoots as well. Rice *Nramp* homologue *OsNramp2* was expressed only in leaves, whereas *OsNramp3* was expressed predominantly in roots, although it was also expressed in shoots.

Phytochelatins

The SUCEST database was searched using keywords for the 3 enzymes involved in phytochelatin biosynthesis (Table 6). Genes encoding γ-glutamylcysteine synthetase (GCS), glutathione synthetase (GS) and phytochelatin synthase (PCS) have been isolated from *Arabidopsis*, *e.g. gsh1*, (May and Leaver, 1995), *gsh2* (Wang and Oliver, 1996), *cad1* (Ha *et al.*, 1999) and *AtPCS1* (Vatamaniuk *et al.*, 1999).

Table VI - Sugarcane expressed sequence tag (EST) clusters homologues to the three enzymes of the phytochelatin biosynthetic pathway.

Enzymes (number of amino acids)	Sugarcane EST clusters	Number of reads	E-values	Amino acid identity	Amino acid position
γ-Glutamylcysteine synthetase (508-523)	SCVPRZ2041G04.g	7	10^{-109}	181/222	302 to 523
	SCJLRT1018B02.g	2	10^{-71}	112/143	366 to 508
	SCQGLR2025D03.g	3	10^{-25}	55/77	432 to 508
Glutathione synthetase (478-552)	SCRLLR1059A07.g	2	10^{-81}	86/123	407 to 529
	SCQSRT1034D10.g	3	10^{-76}	140/186	344 to 529
	SCBGLR1096A08.g	2	10^{-68}	93/129	169 to 552
	SCEZHR1055B12.g	1	10^{-70}	65/111	181 to 291
	SCAGFL3029G12.g	2	10^{-34}	66/98	155 to 448
	SCJLLB2077F06.g	1	10^{-31}	68/97	220 to 316
	SCBGLB2075H01.g	1	10^{-31}	52/84	181 to 264
PC synthase (485-500)	SCJFST1012F10.g	4	10^{-94}	176/239	137 to 390
	SCBFFL1142D06.g	2	10^{-92}	166/238	204 to 441
	SCJFST1047E04.g	2	10^{-82}	142/179	75 to 611

Three sugarcane clusters (E-values < 10^{-25}) were identified with partial homology to γ-glutamylcysteine synthetase (Table 6). These three clusters, derived from 12 reads, were expressed principally in root tissues (5 reads) but transcripts were also detected in apical meristems and flowers (2 reads each), leaves, the shoot-root transition zone and leaf-roll. No transcript contained the amino terminal of the protein.

We also identified seven clusters (E-values < 10^{-31}) derived from 12 reads that showed partial homology to glutathione synthetase, which also was chiefly expressed in roots (4 reads), leaf-roll (3 reads) and lateral buds and flowers (2 reads each). May and Leaver (1994) have shown that genes for these enzymes are constitutively expressed in *Arabidopsis*, while Xiang and Oliver (1998) have reported that such genes are inducible by specific metals (*e.g.* Cd and Cu) and by jasmonic acid.

Table 6 shows that the search of the SUCEST database came up with 3 phytochelatin synthase clusters (E-value < 10^{-82}) derived from 8 reads. These clusters had high similarity with wheat phytochelatin synthase (Clemens *et al.*, 1999), with an amino acid identity of 73%, but none of the clusters had a complete transcript. The reads were found in libraries prepared from stem bark, stem internode, root, flower and calli. Phytochelatin synthase is constitutively expressed in cell cultures and plant roots and stems and its activity is regulated by the presence of metal ions (Cobbett, 2000). Transcription of the *Arabidopsis* gene *CAD1* (*AtPCS1*) is not affected by heavy metal stress (Ha *et al.*, 1999; Vatamaniuk *et al.*, 1999), whereas the wheat gene (*TaPCS1*) is activated when exposed to cadmium (Clemens *et al.*, 1999). From our work it appears that the sugarcane tissues in which γ-glutamylcysteine synthetase, glutathione synthetase and phytochelatin synthase are expressed, and the levels of expression of these genes, are consistent with the data for other systems where expression is constitutive and ubiquitous.

Metallothioneins

Plants contain genes that are structurally and functionally homologous to animal and fungal metallothionein genes (Robinson *et al.*, 1993; Rauser, 1999), but according to Zhou and Goldsbrough, (1995) the homology between animal and fungal metallothioneins is confined to the cysteine residues, with low homology at nucleotide level. The first plant metallothionein protein was isolated from wheat embryos (E_c) as a zinc-binding protein with cysteine residues arranged in three domains (Lane *et al.*, 1987). Since this first isolation, at least 58 genes encoding cysteine-rich metallothionein proteins have been identified from a range of plants and tissues (Rauser, 1999), although most of the predicted proteins have not been isolated and sequenced, except for *Arabidopsis* metallothionein (Murphy *et al.*, 1997) and even their metal-binding capacity has not been demonstrated.

Most plant metallothioneins contain two cysteine-rich domains separated by a large cysteine-free central domain, and are classified into 4 types according to the distribution of the cysteine residues in the two domains (Rauser, 1999). Type I metallothioneins have 6 pairs of cysteine residues all present in a Cys-Xaa-Cys (Cys = cysteine residue and Xaa = amino acids other than cysteine) arrangement distributed equally between the two domains (Figure 3). Type II metallothioneins contain 4 pairs of cysteine residues in the amino terminal region, starting with a Cys-Cys configuration and ending the first domain a Cys-Xaa-Xaa-Cys configuration (Figure 3), the remaining 3 pairs of cysteines at the carboxy terminal being arranged as in the Type I metallothioneins (Robinson *et al.*, 1993). Type III MTs have cysteine mainly at the amino terminal end (Murphy *et al.*, 1997). Type IV metallothioneins are related to the wheat E_c genes expressed during the development of wheat embryos.

Based on a keyword-search for metallothionein, 84 SUCEST clusters were recovered with E-values < 10^{-5}. This high E-value is due to the small size of plant metallothioneins, which typically contain from 45 to 80 amino acids (Robinson *et al.*, 1993). All 84 clusters were examined, but only those with E-values < 7.10^{-7} were analyzed, resulting in a total of 69 clusters. Additionally, four other clusters were identified by the BLAST program using nucleotide sequences from various Type III metallothioneins from *Arabidopsis thaliana* GenBank Z26008; *Carica papaya* Y08322; *Malus domestica* U61974 (Murphy *et al.* 1997) and type IV metallothioneins (E_c metallothioneins) from the monocotyledonous plants *Triticum aestivum* (GenBank X68289 and X68288) *Oryza sativa* (AAG13588) *Zea mays* U10696 (White and Rivin, 1995) and the dicotyledonous plant *Arabidopsis thaliana* (Z27049 and 473284) (Murphy *et al.*, 1997). From this total of 73 sugarcane clusters we obtained 60 which were conceptually translated and contained the full length protein, and 55 clusters (from 849 reads) aligned with metallothioneins from other species when we classified these metallothioneins into the four plant metallothionein types (Figure 3).

Figure 3 shows a representative of each variant of the 27 clusters that we classified as Type I metallothioneins, 17 of which had identical sequences to cluster SCQSHR1020G10.g. Only 4 clusters out of the 27 differed by one amino acid, 5 by 2 amino acids and the single-read cluster SCCCLB1025A03.g by only 4 residues from the consensus. Most of the substitutions occurred in the central cysteine-free spacer domain, and all the cysteine residues were conserved. These variants may arise from inaccurate sequencing and/or cluster assembly, or they may represent gene or allelic variants. The original reads need to be re-sequenced to confirm the variation. We tried to correlate the pattern of expression of the variants in the different li-

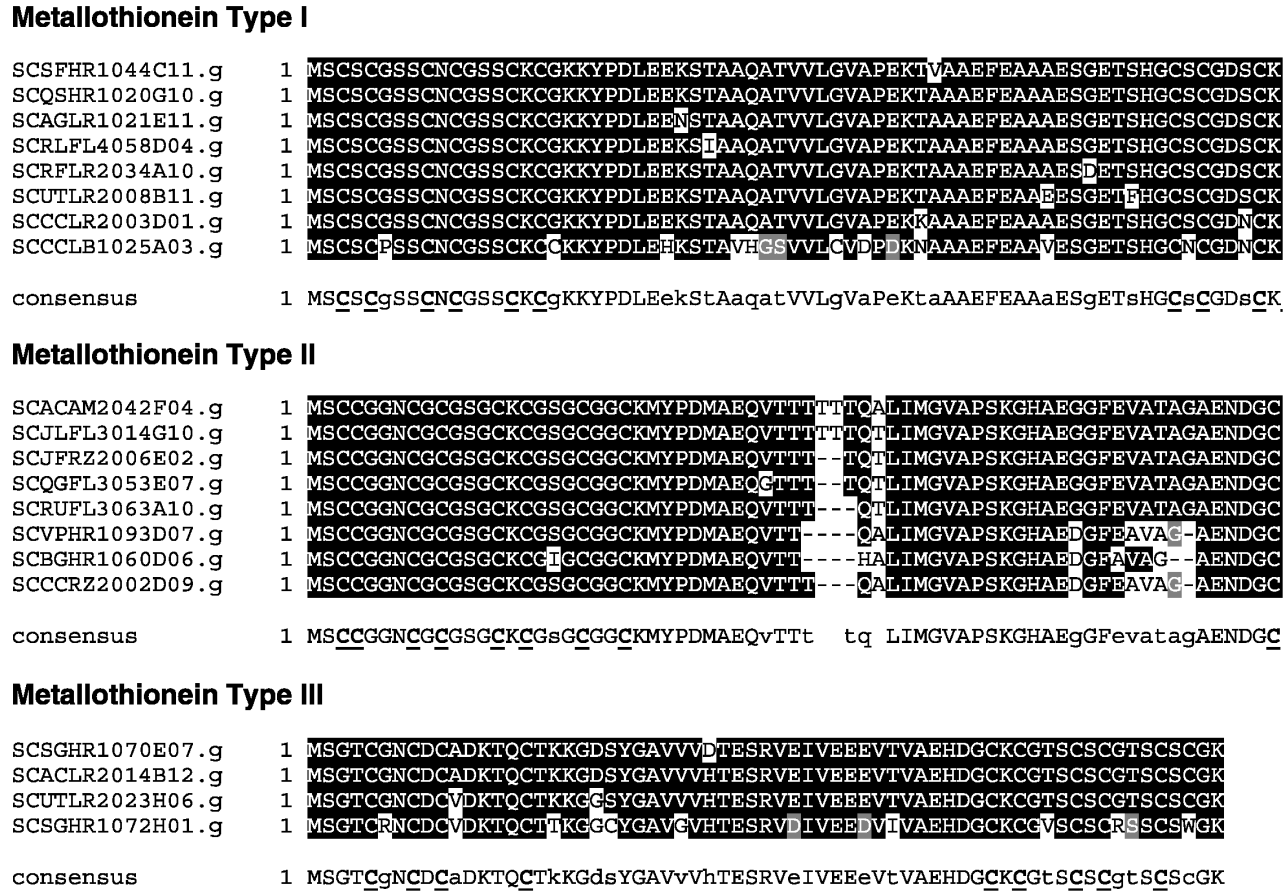

Figure 3 - Predicted amino acid sequences for the three Types of metallothionein-like proteins identified in sugarcane tissues, with consensus. Cystein residues are in bold and underlined in consensus sequences.

braries with the sequence variation but this failed (data not shown). In general, most plant species contain gene families encoding a single metallothionein class (Murphy *et al.*, 1997)

We found 21 clusters related to Type II metallothionein proteins, of which 8 were protein sequence variants with minor amino acid changes, each represented by 1 to 4 clusters (Figure 3). We attempted to further classify these clusters based on *Arabidopsis MT2* gene sequences (*MT2a* and *MT2b*) but were not able to. We were able to classify five clusters as Type III metallothioneins, with clusters SCSGHR1070E07.g and SCCCAD1004C07.g having identical protein sequences (Figure 3). These sequences showed some homology with metallothionein Type III proteins from *Arabidopsis*, kiwi and papaya (Murphy *et al.*, 1997). Only cluster SCSFSD1066A01.g, with an incomplete sequence and derived from only 4 reads expressed in seeds, could be classified as a Type IV metallothionein with homology to the wheat E_c metallothionein.

Of the 291.689 ESTs in the SUCEST database, a total of 849 reads (0.29%) encoding metallothionein-like proteins were identified. These reads were present in most of the tissue libraries, with the level of expression varying from 0.005% in the FL1 flower cDNA library (1 in 18523) to 1.5% (156 in 10336) in the SD2 seed library, indicating that metallothionein genes are highly expressed in various tissues. It is well established that at least some metallothionein genes are expressed at relatively high levels in terms of mRNA abundance, and appear frequently in EST collections (Goldsbrough, 1998). We found that the overall average of metallothionein-like transcripts (0.29%) was similar to value obtained for the normalized library NR2, 0.26% (1 in 384).

Type I metallothionein transcripts were most abundant with 501 reads, followed by Type II metallothionein transcripts with 324 reads. Surprisingly, Type I metallothioneins were very abundant in seed libraries with 119 reads in the seed library SD1 and 146 reads in the SD2 (Table 7). Type IV metallothioneins, including wheat E_c, are usually expressed in seeds (Robinson *et al.*, 1993) and, indeed, we found Type IV metallothionein transcripts in only in seed tissues, but they were present only at low levels (4 reads, 0.019%). White and Rivin, (1995) observed low transcript levels of seed-specific E_c-like maize metallothionein in immature embryos, with transcript levels only reaching peak levels during the mid-maturation phase, although Robinson *et al.* (1993) found that wheat E_c tran-

Table 7 - Summary of the number of reads and percentage of each specific metallothionein-like protein type sequenced in each library, with total number of sequenced reads per type of libraries.

	Shoot-root zone	Roots	Lateral bud	Stem bark	Stem internodes	Leaf	Leaf roll	Apical	Flowers	Seeds	Callus	AD[a]	HR[b]
Total reads	24096	31487	18047	16318	20762	6432	18141	28128	64095	21406	11872	18137	12000
Type I	17 (0.070%)	63 (0.20%)	22 (0.122%)	0 (0.000%)	7 (0.034%)	13 (0.202%)	43 (0.237%)	4 (0.014%)	5 (0.001%)	265 (1.238%)	6 (0.050%)	15 (0.083%)	40 (0.333%)
Type II	53 (0.220%)	5 (0.016%)	16 (0.089%)	4 (0.024%)	42 (0.202%)	9 (0.140%)	6 (0.033%)	49 (0.174%)	87 (0.136%)	12 (0.056%)	17 (0.143%)	5 (0.028%)	19 (0.158%)
Type III	2 (0.001%)	0 (0.000%)	0 (0.000%)	1 (0.001%)	2 (0.001%)	2 (0.031%)	7 (0.038%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	2 (0.011%)	8 (0.066%)
Type IV	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	0 (0.000%)	4 (0.019%)	0 (0.000%)	0 (0.000%)	0 (0.000%)

[a]library made from *in vitro* plantlets without developed leaves and roots infected with *Herbaspirillum rubri* ssp. *albicans*.
[b]library made from *in vitro* plantlets without developed leaves and roots infected with *Gluconsugarcane diazotroficans*.

scripts accumulated at the early stages of embryogenesis. The SUCEST cDNA seed libraries SD1 and SD2 were prepared from equal amounts of mRNA extracted from seeds at three stages of development up to seed maturation, so it is possible that the peak level of sugarcane Type IV metallothionein transcripts might have been missed. Using *in situ* hybridization, Garcia-Hernandez *et al.* (1998) detected a strong signal for *Arabidopsis MT1a* RNA, but not for *MT2a* RNA, in tissues that function in the transport of nutrients into the placenta and funiculus *of* developing *Arabidopsis* seeds. Table 7 shows that a few Type II metallothionein transcripts in tissues from sugarcane seeds were detected but no Type III metallothionein transcript.

In addition to their strong expression in seed cDNA libraries, Type I metallothionein transcripts were also highly expressed in root libraries (Table 7) and also found at high level of expression in the HR1 library prepared from *in vitro* plantlets infected with *Herbaspirillum rubri* ssp. *albicans* and at a significant level of expression in the leaf-roll, and leaf and stem internodes libraries (Table 7). Zhou and Goldsbrough (1995) found that *Arabidopsis MT1a* was expressed at high levels in roots from both young and mature plants, and at a lower level in leaves and was poorly expressed in inflorescence and siliques and concluded that, in general, expression of most Type I metallothionein sequences tends to be more abundant in roots than in leaves. Garcia-Hernandez *et al.* (1998) concluded that localization of the of *Arabidopsis MT1a* in vascular bundles, especially the phloem or in the placenta and funiculus tissues involved in transport of nutrients to the seed, suggest that *MT1a* may play a role in metal ion transport and/or vascular development.

Type II metallothionein transcripts were more abundant in libraries made from the shoot-root transition zone, stem internodes, apical meristem and flower tissues (Table 7). In contrast to Type I metallothionein transcripts, the pattern of the accumulation of Type II transcripts was less frequent in seed, root, and leaf-roll libraries. Garcia-Hernandez *et al* (1998) observed the same trend for *Arabidopsis MT1* and *MT2*, and concluded that these metallothioneins play distinct roles in metal homeostasis. It is interesting to note that in our research the libraries that exhibited most accumulation of Type II metallothionein transcripts (shoot-root transition zone, stem internodes, apical meristem and flowers) contained a significant portion of sugarcane stems. Zhou and Goldsbrough (1995) showed that *Arabidopsis MT2* was constitutively expressed in the leaves, inflorescences and roots of mature plants, but was also detected in roots from young plants and siliques.

Table 7 shows that only 24 reads of Type III metallothionein transcripts were detected from libraries prepared from *in vitro* plantlets infected with *H. rubri* ssp. *albicans* (HR1) and from leaf roll. In *Arabidopsis*, Type III metallothionein transcripts are expressed at higher levels in leaves than in roots (Murphy *et al.*, 1997), and have been detected

in fruits and developing embryos (Garcia-Hernandez *et al.*, 1998).

In *Arabidopsis*, metallothionein genes of various types have been shown to be part of a small gene family consisting of two or more genes, each gene having distinct temporal and spatial expression patterns (Murphy *et al.*, 1997). We found that the level of expression of metallothionein transcripts detected in sugarcane under standard growth conditions indicates that these genes are constitutively expressed at high levels, particularly the gene for Type I metallothionein, although it will be necessary to conduct northern hybridization analyses to confirm the specificity of metallothionein gene expression in sugarcane tissues. The function of metallothioneins is still unknown, but the evidence indicates that these proteins may play an important role in metal homeostasis, probably associated with metal transport within the plant.

CONCLUSIONS

The expression and function of the sugarcane gene homologues reported in this paper need to be confirmed by northern expression analysis, micro-arrays or macro-arrays and appropriate induction conditions, or by the functional complementation of yeast cells or model plants such as *Arabidopsis* or rice. Over-expressing or co-suppressing these sequences in transgenic sugarcane or rice should allow the evaluation of expected phenotypes under suitable conditions and enable us to validate the identity and expression patterns of these genes. The first obvious candidates for such treatment are the high affinity phosphate transporter genes and sequences associated with metal chelating peptides such as metallothioneins and phytochelatins.

RESUMO

A absorção de nutrientes pelas plantas é um processo ativo, requerendo energia para o acúmulo de nutrientes essenciais em níveis mais elevados nos tecidos vegetais do que na solução do solo, enquanto que a presença de metais tóxicos ou excesso de nutrientes requererem mecanismos para modular o acúmulo de íons. Genes que codificam transportadores de íons, isolados de plantas e de leveduras, foram usados para identificar homólogos presumíveis no banco de dados de seqüências expressas de cana-de-açúcar (SUCEST). Cinco consensos de grupos de seqüências com homologia a genes de transportadores de fosfato de alta afinidade foram identificados. O consenso de um dos grupos permitiu a predição da proteína completa, com 541 amino ácidos e 81% de identidade com o gene *NtPT1* de *Nicotiana tabacum*, consistindo de 12 domínios transmembrana divididos por uma grande região hidrofílica. Homólogos presumíveis a genes transportadores de micronutrientes de *Arabidopsis thaliana* também foram detectados em algumas bibliotecas do SUCEST. A absorção de ferro em gramíneas envolve a liberação de um composto fito-sideróforo, ácido mugenêico (MA), que complexa com Fe^{3+}, sendo então absorvido por um transportador específico. Seqüências expressas (EST, *expressed sequence tag*) de cana-de-açúcar homólogas aos genes que codificam as três enzimas da via de biossíntese do ácido mugenêico (nicotianamina sintase; nicotianamina transferase; e a sintetase presumível do ácido mugenêico *ids3*), além de um transportador presumível de Fe^{3+}-fito-sideróforo foram também detectados. Sete grupos de seqüências de cana-de-açúcar foram identificados com grande homologia com os membros da família de genes ZIP (*ZIP1*, *ZIP3*, *ZIP4*, *IRT1* e *ZNT1*), enquanto que quatro grupos apresentaram homologia com *ZIP2* e três com *ZAT*. Seqüências homólogas aos membros de uma outra família de genes, *Nramp*, que codificam transportadores de metais de ampla espectro, foram também detectados com expressão constitutiva. Transcritos parciais homólogos aos genes que codificam γ-glutamilcisteína sintetase, glutationa sintetase e fitoquelatina sintase (responsáveis pela biossíntese da proteína quelante de metais, fitoquelatina) e todos os quatro tipos do outro principal peptídeo quelante de metais em plantas, metalotioneína (MT), foram identificados: MT do tipo I sendo a mais abundante (> 1% das seqüências na biblioteca de sementes), seguido pela MT do tipo II, com padrão de expressão similar àquele descrito para MT de *Arabidopsis*. A identificação e a compreensão da expressão de genes associados com a absorção de nutrientes e tolerância a metais poderiam possibilitar o desenvolvimento de variedades de cana-de-açúcar mais eficientes nutricionalmente, ou permitiriam o uso da cana-de-açúcar como planta hiper-acumuladora para a restauração de área contaminadas em programas de fitorremediação.

ACKNOWLEDGMENTS

This work is supported by FAPESP (99/0248-7). The authors thank the Brazilian agencies CNPq (A.F.), CAPES (E.A.K.) and FAPESP (R.S.A.) for fellowships.

REFERENCES

Chrispeels, M.J., Crawford, N.M. and **Schroeder, J.I.** (1999). Proteins for transport of water and mineral nutrients across the membranes of plant cells. *Plant Cell 11*: 661-676.

Clemens, S., Kim E.J., Neuman D. and **Schroeder J.I.** (1999). Tolerance to toxic metals by a gene family of phytochelatin synthases from plants and yeast. *EMBO J. 18*: 3325-3333.

Cobbett, C.S. (2000). Phytochelatins and their roles in heavy metal detoxification. *Plant Physiol. 123*: 825-832.

Curie, C., Panaviene, Z., Loulergue, C., Dellaporta, S.L., Briat, J.F. and **Waler, E.L.** (2001). Maize *yellow stripe 1* encodes a membrane protein directly involved in Fe (III) uptake. *Nature 409*: 346-349.

Daram, P., Brunner S., Rausch C., Steiner C., Amrhein N. and **Bucher, M.** (1999). *Pht2;1* encodes a low-affinity phosphate transporter from *Arabidopsis*. *Plant Cell 11*: 2153-2166.

Eide, D., Broderius M., Fett J. and **Guerinot, M.L.** (1996). A novel iron-regulated metal transporter from plants identified

by functional expression in yeast. *Proc. Natl. Acad. Sci. (USA) 93*: 5624-5628.

Fox, T.C. and **Guerinot, M.L.** (1998). Molecular biology of cation transport in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol. 49*: 669-696.

Garcia-Hernandez, M., Murphy, A. and **Taiz, L.** (1998). Metallothioneins 1 and 2 have distinct but overlapping expression patterns in *Arabidopsis*. *Plant Physiol. 118*: 387-397.

Giordani, T., Natali, L., Maserti, B.E., Taddei, S. and **Cavallini, A.** (2000). Characterization and expression of DNA sequences encoding putative type-II metallothioneins in the seagrass *Posidonia oceanica*. *Plant Physiol. 123*: 1571-1581.

Goldsbrough, P.B. (1998). Metal Tolerance in Plants: the role of phytochelatins and metallothioneins. In: *Phytoremediation of contaminated soils soil and water* (Terr, N. and Banuelos, G.S., eds). CRC Press, Boca Raton, EUA, pp. 221-233.

Grotz, N, Fox, T., Conolly, E., Park, W., Guerinot, M.L. and **Eide, D.** (1998). Identification of a family of zinc transporter genes from *Arabidopsis* that respond to zinc deficiency. *Proc. Natl. Acad. Sci. (USA) 95*: 7220-7224.

Guerinot, M.L. and **Eide, D.** (1999). Zeroing in on zinc uptake in yeast and plants. *Current Opinion in Plant Biology 2*: 244-249.

Guerinot, M.L. and **Salt, D.E.** (2001). Fortified foods and phytoremediation. Two sides of the same coin. *Plant Physiol. 125*: 164-167.

Ha, S.B., Smith, A.P., Howden, R., Dietrich, W.M., Bugg, S., O'Connell, M.J., Goldsbrough, P.B. and **Cobbett, C.S.** (1999). Phytochelatin synthase genes from *Arabidopsis* and the yeast *Schizosaccharomyces pombe*. *Plant Cell 11*: 1153-1163.

Herbik, A., Koch, G., Mock, H.P., Dushkov, D., Czihal, A., Thielman, J., Stephan, U.W. and **Baulmlein, H.** (1999). Isolation, characterization and cDNA cloning of nicotianamine synthase from barley. *European J. Biochem. 265*: 231-239.

Higuchi, K., Suzuki, K., Nakanishi, H., Yamaguchi, H., Nishiwaza, N.K. and **Mori, S.** (1999). Cloning of nicotianamine synthase genes, novel genes involved in the biosynthesis of phytosiderophores. *Plant Physiol. 119*: 471-479.

Hirsch, R.E. and **Sussman, M.R.** (1999). Improving nutrient capture from soil by the genetic manipulation. *Trends in Biotechnology 17*: 356-360.

Kochian, L.V. (2000). Molecular physiology of mineral nutrient acquisition, transport and utilization. In: *Biochemistry and Molecular Biology of Plants*. (Buchan, B.B., Gruissen, W. and Jones, R.L., eds.). American Society of Plant Physiology, Rockville, EUA, pp. 1204-1249.

Korshunova, Y.O., Eide, D., Clark, W.G., Guerinot, M.L. and **Eide, D.** (1998). The IRT1 protein from *Arabidopsis thaliana* is a metal transporter with broad specificity. *Plant Mol. Biol. 40*: 37-44.

Lane, B., Kajioka, R. and **Kennedy, T.** 1987. The wheat germ E_c protein is a zinc-containing metallothionein. *Biochem. Cell. Biol. 65*: 1001-1005.

Leggewie, G., Willmitzer, L. and **Riesmeier, J.W.** (1997). Two cDNAs from potato are able to complement a phosphate uptake-deficient yeast mutant: identification of phosphate transporters from higher plants. *Plant Cell 9*: 381-392.

Liu, C., Muchhal, U.S., Uthappa, M., Kononowicz, A.K. and **Rhagothama, K.G.** (1998). Tomato phosphate transporter genes are differentially regulated in plant tissues by phosphorous. *Plant Physiol. 116*: 91-99.

May, J.M. and **Leaver, C.J.** (1995). *Arabidopsis thaliana* g-glutamylcysteine synthetase is structurally unrelated to mammalian, yeast and *Escherichia coli. Proc. Natl. Acad. Sci (USA) 91*: 10059-10063.

Mori, S. (1999). Iron acquisition by plants. *Current Opinion in Plant Biology 2*: 250-253.

Muchhal, U.S., Pardo, J.M. and **Raghothama, K.G.** (1996). Phosphate transporters from the higher plant *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci (USA) 93*: 10159-10523.

Murphy, A.S. and **Taiz, L.** (1995). Comparison of metallothionein gene expression and non-protein thiols in ten *Arabidopsis* ecotypes. Correlation with copper tolerance. *Plant Physiol. 109*: 1-10.

Murphy, A.S., Zhou, J., Goldsbrough, P.B. and **L., Taiz.** (1997). Purification and immunological identification of metallothioneins 1 and 2 from *Arabidopsis thaliana*. *Plant Physiol. 113*: 1293-1301.

Nakanishi, H., Yamaguchi, H., Sasakuma, T., Nishizawa, N.K. and **Mori, S.** (2000). Two dioxygenases genes, *Ids3* and *Ids2*, from *Hordeum vulgare* are involved in the biosynthesis of mugeneic acid family phytosiderophores. *Plant Mol. Biol. 44*: 199-207.

Pence, N.S, Larsen, P.B., Ebbs, S.D., Letham, D.L.D., Lasat, M.M., Gravin, D.F., Eide, D. and **Kochian, L.V.** (2000). The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. *Proc. Natl. Acad. Sci (USA) 97*: 4956-4960.

Raghothama, K.G. (1999). Phosphate acquisition**.** *Annual Review Plant Physiology Plant Molecular Biology 50*: 665-686.

Raghothama, K.G. (2000). Phosphate transport and signaling**.** *Current Opinion in Plant Biology 3*: 182-187.

Rauser, R.W. (1999). Structure and function of metal chelators produced by plants: the case for organic acids, amino acids, phytin and metallothioneins. *Cell Biochem. Biophys. 31*: 19-48.

Robinson, N.J., Tommey, A.T., Kuske, C. and **Jackson, P.J.** (1993). Plant metallothioneins. *Biochem. J. 295*: 1-10.

Robinson, N.J., Procter, C.M., Conolly, E.L. and **Guerinot, M..L.** (1999). A ferric chelate reductase for iron uptake from soils. *Nature 397*: 694-697.

Takahashi, M., Yamaguchi, H., Nakanishi, H., Shiori, T., Nishizawa, N.K. and **Mori, S.** (1999). Cloning two genes for nicotianamine aminotransferase, a critical enzyme in iron acquisition (Strategy II) in graminaceous plants. *Plant Physiol. 121*: 947-956.

Teles, G.P and **Silva, F.R.** (2001). Trimming and clustering sugarcane ESTS. *Genet and Mol. Biol. 24* (1-4): 17-23.

Thomine, S., Wang, R., Ward, J.W., Crawford, N.M. and **Schroeder, J.I.** (2000). Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to *Nramp* genes. *Proc. Natl. Acad. Sci. (USA) 97*: 4991-4996.

Van der Zaal, B.J., Neuteboom, L.W., Pinas, J.E., Chardonnens, A.N., Schat, H., Verkeij, J.A.C. and **Hoykass, P.J.J.** (1999). Overexpression of a novel *Arabidopsis* gene related to putative zinc-transporter genes from animals can lead to enhanced zinc resistance and accumulation. *Plant Physiol. 119*: 1047-1055.

Vatamaniuk, O.K., Mari, S., Lu, Y.P. and **Rea, P.A.** (1999). AtPCS1, a phytochelatin synthase from *Arabidopsis*: isolation and *in vitro* reconstitution. *Proc. Natl. Acad. Sci. (USA) 96*: 7110-7115.

Wang, C. and **Oliver, D.J.** (1996). Cloning of the cDNA and genomic clones for glutathione synthetase from *Arabidopsis thaliana* and complementation of a *gsh2* mutant from fission yeast. *Plant Mol. Biol. 31*: 1093-1104.

White, C.N. and **Rivin, C.J.** (1995). Characterization and expression of a cDNA encoding a seed-specific metallothionein in maize. *Plant Physiol. 108*: 831-832.

Xiang, C. and **Oliver, D.J.** (1998). Glutathione metabolic genes coordinately respond to heavy metals and jasmonic acid in *Arabidopsis*. *Plant Cell 10*: 1539-1550.

Zhou, J. and **Goldsbrough, P.B.** (1995). Structure, organization and expression of the metallothionein gene family in *Arabidopsis*. *Mol. and Gen. Genet. 248*: 318-328.