Research Article

Iron homeostasis related genes in rice

Jeferson Gross[1,3], Ricardo José Stein[2], Arthur Germano Fett-Neto[1,2] and Janette Palma Fett[1,2]

[1]*Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia, Porto Alegre, RS, Brazil.*
[2]*Universidade Federal do Rio Grande do Sul, Departamento de Botânica, Porto Alegre, RS, Brazil.*
[3]*Botanical Institute of Ludwig-Maximilians-Universität, München, Germany.*

Abstract

Iron is essential for plants. However, excess iron is toxic, leading to oxidative stress and decreased productivity. Therefore, plants must use finely tuned mechanisms to keep iron homeostasis in each of their organs, tissues, cells and organelles. A few of the genes involved in iron homeostasis in plants have been identified recently, and we used some of their protein sequences as queries to look for corresponding genes in the rice (*Oryza sativa*) genome. We have assigned possible functions to thirty-nine new rice genes. Together with four previously reported sequences, we analyzed a total of forty-three genes belonging to five known protein families: eighteen YS (Yellow Stripe), two FRO (Fe^{3+}-chelate reductase oxidase), thirteen ZIP (Zinc regulated transporter / Iron regulated transporter Protein), eight NRAMP (Natural Resistance - Associated Macrophage Protein), and two Ferritin proteins. The possible cellular localization and number of potential transmembrane domains were evaluated, and phylogenetic analysis performed for each gene family. Annotation of genomic sequences was performed. The presence and number of homologues in each gene family in rice and *Arabidopsis* is discussed in light of the established iron acquisition strategies used by each one of these two plants.

Introduction

Iron is a key micronutrient for plants, taking part in redox centers of proteins essential for photosynthesis and respiration. Iron deficiency causes a metabolic imbalance deleterious to plant development (Briat, 1995). This problem is especially visible in calcareous soils, where the high pH dislocates the chemical balance towards the formation of iron insoluble complexes, leaving the cation sparsely available for plant uptake (Guerinot and Yi, 1994; Briat *et al.*, 1995). However, in anaerobic conditions and acid pH, as found in waterlogged soils, the ferrous form is stabilized and readily taken up by plants. When in excess, iron induces the production of hydroxyl radicals, the most potent oxidizing agents known. These radicals cause multiple damage to cellular structures, eventually leading to death (Guerinot and Yi, 1994; Briat *et al.*, 1995). Therefore plants must balance iron concentration in a homeostatic way, providing the necessary amounts of the micronutrient and preventing the condition of internal cation excess (Briat and Lobréaux, 1997, Grusak *et al.*, 1999).

The biotechnological manipulation of plant crops in a way to increase their capacity to adequate the iron content may result in yield improvement (Takahashi *et al.*, 2001; Guerinot, 2001). Engineering plant crops for efficient iron nutrition is also justified because about 60% to 80% of the world population suffers in some manner from organic deficiencies of this micronutrient (World Health Organization, 2002). The increase in iron content in plant staple foods can serve as a vector for iron ingestion in target populations, minimizing nutritional problems (Vansuyt *et al.*, 2000; Murrey-Kolb *et al.*, 2002). In this regard, a better understanding of iron homeostasis, involving knowledge of the basic physiological processes of iron absorption, distribution and storage in plants, can serve as starting point for the biotechnological manipulation of crops (Grusak *et al.*, 2002; Grotz and Guerinot, 2002). Plant iron homeostasis has been often analyzed by physiological and molecular genetics approaches. More recently, genomics also appears as an alternative for the understanding of this aspect of plant nutrition.

Traditional physiological studies gave special emphasis to the mechanisms for iron uptake from the soil. These have been classified in two distinct strategies (Marschner and Römheld, 1994). Strategy I is shared be-

Send correspondence to Janette Palma Fett. Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia, Caixa Postal 15005, 91501-970 Porto Alegre, RS, Brazil. E-mail: jpfett@dna .cbiot.ufrgs.br.

tween dicotyledoneous and non graminaceous mono-cotyledoneous species. Since iron in its ferric form is more soluble at low pH, these plants acidify the rhizosphere by means of a H^+-ATPase pump and release reductants/chelators such as organic acids and phenolics (Römheld and Marschner, 1983; Römheld *et al.*, 1984). The ferric ion is then chelated by the released compounds, as well as by soil bacteriosiderophores (Römheld and Marschner, 1983; Bar-Ness *et al.*, 1992). Subsequently, chelated Fe^{3+} is reduced by a low iron-inducible enzymatic activity present in the plasma membrane of root epidermal cells (Moog and Brüggermann, 1994) and Fe^{+2} is then absorbed by a specific transporter (Fox *et al.*, 1996). A different strategy to overcome the problem of low iron availability in the soil evolved in grasses (such as rice). The so called strategy II consists of releasing phytosiderophores, non-proteinogenic amino acids (structural derivatives of mugineic acid) that present mineral chelating properties (Sugiura and Nomoto, 1984; Römheld and Marschner, 1986). Depending on the mineral needs of the plant, these compounds are released to the rhizosphere, where they chelate the sparsely soluble Fe^{3+}, making it more available for uptake (von Wirén *et al.*, 1995). The Fe^{+3}-phytosiderophore complex is readily absorbed by the plant, without need of extracellular reduction (Römheld and Marschner, 1986; von Wirén *et al.*, 1995).

With this physiological model in mind, a new approach based on molecular genetics in the 1990s allowed the identification of molecular components related to the two strategies. The gene encoding the Fe^{+2} transporter related to strategy I, Iron Regulated Transporter 1 (*IRT*1), was cloned in *Arabidopsis thaliana* by functional complementation of *Saccharomyces cerevisiae* mutants defective in iron acquisition (Eide *et al.*, 1996). Other *IRT*1 homologues were soon identified in yeast showing capacity for zinc transport (Zinc Regulated Transporter, *ZRT*) (Zhao and Eide, 1996a,b), which led to the ZIP (Zrt/Irt-related Proteins) denomination of this new protein family (Grotz *et al.*, 1998). Some homologues of ZIP/IRT have already been characterized in *Arabidopsis*, tomato and soybean (Grotz *et al.*, 1998; Eckhardt *et al.*, 2001; Vert *et al.*, 2001; Moreau *et al.*, 2002). The ferric-chelate reductase component related to strategy I was identified in *Arabidopsis* (*AtFRO*2) (Robinson *et al.*, 1999) and later in pea (*PsFRO*1) (Waters *et al.*, 2002). The *FRO* genes belong to a super family of flavocytochromes located in the cellular membrane that transfer electrons from intracellular donors such as NADPH to extracellular acceptors such as iron or molecular oxygen. Transposon-tagging from a maize mutant defective in iron uptake allowed the cloning of *Yellow Stripe* 1 (*YS1*), the only molecular component related to the transport of the Fe^{3+}-phytosiderophore complex characterized to date (Curie *et al.*, 2001).

Additional components related to iron homeostasis have been characterized. Although not related to mineral uptake, these components are probably involved in iron intracellular targeting and storage. Members of the widely distributed NRAMP (Natural Resistance-Associated Macrophage Protein) family of cation transporters have also been characterized in *Arabidopsis* and rice (Belouchi *et al.*, 1997; Thomine *et al.*, 2000). Experimental evidence for a role of AtNRAMP1, 3 and 4 and OsNRAMP1 in iron homeostasis already exists. AtNRAMP1 may be related to iron subcellular transport and its targeting to storage compartments such as vacuoles or plastids (Curie *et al.*, 2000). Ferritin has a well established role in iron storage and buffering of the mineral availability in the cytoplasm of mammalian cells. In plants, Ferritin is localized in the plastid strome, where it performs a similar function (Ragland *et al.*, 1990). Ferritin cDNAs have been isolated from pea, soybean, maize and *Arabidopsis* (Briat and Lobréaux, 1997).

Currently, plant genomics allows the fast identification of molecular components related to mineral homeostasis. This is possible through database searches for sequences homologous to proteins already characterized. With this approach, recent reports describe the identification of six genes related to the *NRAMP* family and fifteen genes related to the *ZIP* family in *Arabidopsis* (Mäser *et al.*, 2001), as well as eight ORFs with homology to maize *YS1* (Curie *et al.*, 2001).

Once identified, these new components can be molecularly characterized and used for more detailed studies. The analysis of the *Arabidopsis* EST database and the availability of its complete genome allowed the identification of three new patterns of ferritin, in addition to the already described *AtFer1*. These new forms provide the first genome-wide study of the expression of ferritin genes in plants (Petit *et al.*, 2001b).

Together with *Arabidopsis*, rice stands out as a model for plant genomics (Shimamoto and Kyozuka, 2001). The International Rice Genome Sequencing Project (IRGSP), with roots in the Japanese Rice Genome Project, proposes the sequencing of the whole 430 Mb of the rice genome of the japonica variety, cultivar Nipponbare, based on a clone-by-clone approach, aiming at a high accuracy of 10^{-4} (Eckardt, 2000; Delseny *et al.*, 2001). In 2001, the IRGSP was highly propelled by the incorporation of the Monsanto rice genome draft and, at the present time, the IRGSP has submitted sequences to GenBank that comprise over 95% of the genome (http://www.tigr.org/tdb/e2k1/osa1/BACmapping/description.shtml). Sequencing is still in progress and the announcement of the draft conclusion is expected very soon (Burr, 2002). In parallel, the Beijing Genomics Institute (BGI) and Syngenta have separately published drafts of the rice genome based on a whole shotgun sequencing methodology. The Chinese team obtained a coverage of 92% of the genome of the indica variety, cultivar 93-11, with a prediction of 46,022 to 56,615 genes (Yu *et al.*, 2002). The Syngenta draft comprises 93% of the japonica variety genome, with an estimate of 32,000 to 50,000 genes (Goff *et al.*, 2002).

Using the indica variety genome, the IRGSP data, and the expressed sequence tags (ESTs) deposited in public databanks, we performed an analysis of the rice genome searching for sequences related to the YS, ZIP, FRO, NRAMP and Ferritin protein families. The purpose of this study was to contribute to the understanding of iron homeostasis dynamics in rice (focusing on gene families involved in uptake, intracellular targeting and storage), and to compare the molecular picture of iron nutrition in this model grass species to *Arabidopsis thaliana*.

Materials and Methods

Reference proteins of well established molecular function, representing each of the protein families investigated, were chosen as query sequences for searches in the rice (*Oryza sativa*) genome databases. These reference proteins were ZmYS1 (AAG17016), AtFRO2 (CAA70770), PsFRO1 (AAK95654), AtZIP1 (AAC24197), AtZIP2 (AAC24198), AtIRT1 (AAB01678), AtNRAMP1 (AAF36535), AtNRAMP2 (AAD41078) and ZmFer1 (CAA58146). Searches were made using the TBLASTN tool (Altschul *et al.*, 1997) against the GenBank databases Highthroughput Genomic Sequences (HTG) and Non Redundant (NR), with search specifications for *Oryza sativa* [Organism]. The BLAST server used was that of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/). The indica variety genome was screened in a specific database offered in the Genomic BLAST pages of the same server (http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/riceWGS.html). As selection criteria of BLAST hits for genomic sequences, a cutoff e-value of e-10 was previously set.

The genomic sequences found were used to predict putative genes contained within them. Whenever possible, genes were predicted on the basis of sequences generated by the IRGSP, since these sequences present a higher degree of accuracy. To that end, a mixed procedure was adopted combining *ab initio* gene prediction algorithms of genomic sequence alignments with similar sequences from expressed genes (ESTs and cDNAs). The prediction algorithms were GenScan (Burge and Karlin, 1997; http://genes.mit.edu/GENSCAN.html), GenomeScan (Burge and Karlin, 1997; http://genes.mit.edu/ genomescan.html), FGENESH (Salamov and Solovyev, 2000; http://www.softberry.com/berry.phtml?topic= gfind), GeneMark.hmm (Borodovsky and Lukashin, unpublished; http://opal.biology.gatech.edu/GeneMark/ eukhmm.cgi) and GrailEXP (Xu and Uberbacher, 1997; http://compbio.ornl.gov/grailexp/).

The similar expressed sequences were found by BLAST searches against EST and NR databases of GenBank, using the genomic sequence as query. The genomic sequences, predicted sequences and expressed gene sequences were then compared by multiple alignments using DIALIGN 2.1 (Morgenstern, 1999; http://bib

iserv.techfak.uni-bielefeld.de/cgi-bin/dialign_submit), that allowed the identification of exon and intron positions on the genomic sequences.

Each new predicted CDS served as a query sequence for new BLAST searches, leading to the identification of the largest possible number of related sequences.

The predicted CDS were translated into amino acids and compared to the reference sequences for the searches and to homologous proteins from *Arabidopsis*. The algorithm of choice for the multiple alignments of protein sequences was ClustalW 1.8 (Thompson *et al.*, 1994), available through the BCM Search Launcher server (http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html). The multiple alignments were edited with the help of GENEDOC (Free Software Foundation, Inc). All of the proteins with greater than 30% identity, with at least one of the reference proteins used in the searches, were regarded as functionally similar (homologous) to the reference proteins, receiving the same name (Orengo *et al.*, 1994; Rost and Valencia, 1996; Rost, 1999; Todd *et al.*, 2001). Those sequences that did not conform to this criterion were discarded. Only in the case of OsFRO2 the identity degree of 26% was accepted due to the conserved functional domains between this protein and the reference proteins.

The selected proteins were evaluated for the presence of transmembrane domains, using the SOSUI algorithm (Hirokawa *et al.*, 1998; http://sosui.proteome.bio.tuat.ac.jp/sosuimenu0E.html). In some cases, the proteins were analyzed in the protein secondary structure databases MetaFam (Silverstein *et al.*, 2001) and Conserved Domain Database (CDD) (Marchler-Bauer *et al.*, 2002) (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Protein alignments obtained with Clustal W 1.8 (Thompson *et al.*, 1994) were used as starting points for phylogenetic analysis, which were based on the Parsimony method with the PAUP* software (version 4.0b 10 (PPC/activec; Swofford, 2002). In all cases, 1000 Bootstrap replications tested the tree topology obtained.

Genes were positioned in the chromosomes in agreement with the IRGSP genetic map, available at the site of the Institute for Genomic Reseach, TIGR. Genetic markers were associated to the IRGSP clones using a genetic marker miner located at the TIGR genetic map browser (http://www.tigr.org/tdb/e2k1/osa1/BACmapping/description.shtml) and also from the TIGR Rice BLAST server (http://tigrblast.tigr.org/euk-blast/index.cgi?project=osa1) with searches against the genetic marker database.

The EST searches were done both from reference proteins using TBLASTN and from predicted gene sequences, using BLASTN against the EST databases with search specifications for *Oryza sativa* [Organism]. The e-value cutoff was 0.01. The ESTs showing less than 5% of mismatches in an overlap extension of at least 100 nucleotides of BLAST alignment with a predicted CDS from the

genomic sequence were considered related to a predicted gene in the chromosome.

ESTs D22266 and AU064475 were sequenced in an automatic sequencer MegaBACE DNA Analysis System (Amersham Biosciences) with the M13 -20 and reverse primers.

Sequences and annotations of the new putative genes were submitted to GenBank. Accession numbers of new ferritins are AF519570 (*OsFer1*) and AF519571 (*OsFer2*). Other nucleotide sequence data reported are available in the Third Party Annotation Section of the DDBJ/EMBL/GenBank databases and their accession numbers are shown in the respective tables.

Results and Discussion

Genomic overview

With the availability of the BGI and the IRGSP data in GenBank, it was possible to construct an overview of iron homeostasis in rice. As a starting point, the protein families Yellow Stripe, FRO, ZIP, NRAMP and Ferritin, which have positive molecular implications on iron uptake, intracellular targeting and storage in other plants, were cho-

sen for analysis. Taking specific members of these families as query sequences, searches for homologous sequences in the non-redundant (NR), expressed sequence tag (EST) and high throughput genomic sequences (HTG) in GenBank current databases were carried out. Thirty-nine new putative genes were predicted. Together with four previously reported sequences, we analyzed in this study a total of forty-three genes possibly involved with iron homeostasis in rice: eighteen genes related to the Yellow Stripe Fe^{3+}-phytosiderophore transporter, two genes homologous to the *FRO* family, thirteen sequences related to the *ZIP* family (including two *IRT* genes), eight *NRAMP* putative genes and two ferritin genes. Since the PAC and BAC clones of the IRGSP are linked to genetic markers, it was possible to estimate the relative position of the sequences in the genetic map (Figure 1). The location of five putative new genes still remains unknown, which were predicted from the indica data, whose scaffolds were not anchored to genetic maps.

The yellow stripe family

A strong indication that the maize *YS1* gene encodes the Fe^{+3}-phytosiderophore transporter related to strategy II
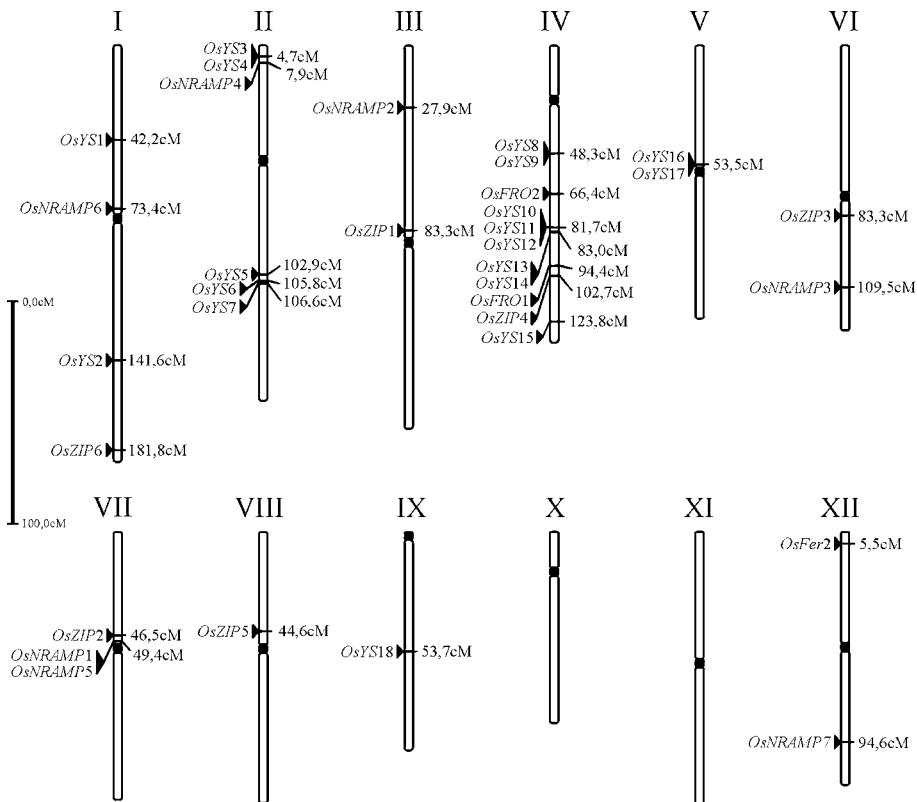
Figure 1 - Chromosome locations of the putative genes related to the iron homeostasis families in rice. All positions in centMorgans (cM) were estimated in accordance to genetic markers assigned to the BAC/PAC clones (as shown by the TIGR genetic map browser: http://www.tigr.org/tdb/e2k1/osa1/BACmapping/description.shtml) related to the predicted sequences. Roman numbers indicate each of the twelve rice chromosomes depicted in white with cetromeres shown in black. The proposed gene nomenclature is as follows: *OsYS, Oryza sativa Yellow Stripe-like*; *OsFRO, Oryza sativa Ferric-chelate reductase-like*; *OsZIP, Oryza sativa Zip-like*; *OsNRAMP, Oryza sativa Nramp-like*; *OsFer, Oryza sativa Ferritin-like*. It was not possible to assign a position for five putative genes (*OsIRT1, OsIRT2, OsZIP9, OsZIP10, OsZIP11*), since these genes were predicted from the BGI scaffolds not anchored to any rice chromosome map.

found in grass species came from the demonstration of its transport specificity: heterologous expression of a *Ys*1 clone in a *Saccharomyces cerevisiae* mutant lineage defective in iron acquisition resulted in functional complementation when Fe^{3+} was supplied chelated with deoxymugineic acid, but not when Fe-citrate was the mineral source. Expression of *Yellow Stripe*1 (*YS1*) in maize roots is increased under mineral deficiency (Curie *et al.*, 2001).

In addition to the eight members of the Yellow Stripe family found in *Arabidopsis*, there is also a vast number of homologues found in public DNA sequence databanks, related to expressed genes in other plants. In a broader sense, the YS1 protein also belongs to a superfamily of oligopeptide transporters, called OPT (Oligopeptide Transporter) (Yen *et al.*, 2001). The OPT family is particularly well characterized in yeast (Hauser *et al.*, 2001) and *Arabidopsis* (Koh *et al.*, 2002), in which its members function in the acquisition of oligopeptides (tetra and pentapeptides) from the environment. A wider role for this protein family in the transport of nicotianamine (NA) in plants has also been suggested (Curie *et al.*, 2001).

Through searches for homologous sequences related to the maize transporter YS1, we found several orthologs in the rice genome, with at least forty six related ESTs. All of these correspond to eighteen genomic sequences related to the IRGSP BAC/PAC clones. Only *OsYS1* had already been described and had CDS (coding DNA sequence) prediction submitted to the PLN subdivision of GenBank (for plant, fungal and algal sequences), which required the coding regions to be established for the other sequences. They were predicted from the respective IRGSP clones. The patterns of exons are shown in Table 1, with the numerals corresponding to the positions of the coding regions relative to the genomic clones. We propose a nomenclature for the maize *YS1* orthologs in rice as "*Oryza sativa Yellow Stripe-like*" (*OsYS*), numbered 1 to 18 as the order of appearance in the genome (Figure 1, Table 1). In four cases there are *OsYS* genes positioned in tandem in the same genomic locus, evidence of recent gene duplications leading to the expansion of this family. The CDS length of the new sequences is around 2000 nucleotides, which is similar to the 2049 found in maize *YS1* (Curie *et al.*, 2001). Most of the sequences have a structure of six to eight exons, which is in good agreement with the seven exons found in maize *YS1*. There are related ESTs for ten out of eighteen *YS1* rice orthologs, indicating that these are functional genes. For the putative CDS without related ESTs, further studies will be required to attest their functionality.

The seventeen newly identified genes and the previously reported *OsYS1* were translated into amino acids and, together with maize YS1, analyzed in the secondary protein database MetaFam. In all cases, hits occurred with the profile of the OPT family. For maize YS1 protein, the obtained e-value related to the OPT family profile was 2.7e-98, con-

firming that the Yellow Stripe family belongs to the OPT super family of oligopeptide transporters (Yen *et al.*, 2001). Similar e-values were found for the rice sequences. The similarity with the OPT family also seems to be functional, since maize YS1 transports compounds derived from mugineic acid, a non-protein amino acid, structurally similar to small peptides. The eighteen putative OsYS proteins, maize YS1 and the eight orthologs found in *Arabidopsis* were compared at protein level based on multiple alignments. All the putative OsYS proteins showed more than 30% identity with maize YS1. The most similar proteins were OsYS7 with 75% identity and OsYS6, 13 and 14 with more than 60% identity with the maize transporter. The rice YS members also share a high level of similarity with the AtYS-like proteins, with more than 40% identity in most cases. The presence of transmembrane domains was also investigated for the eighteen OsYS proteins. The result was somewhat heterogeneous with the majority of proteins presenting 12 to 14 transmembrane domains (Table 1), close to the 12 observed in the maize YS1 (Curie *et al.*, 2001).

A better visualization of the relationships between YS protein sequences was achieved by a phylogenetic analysis. Besides the YS members of rice, *Arabidopsis* and the maize YS1, the OPT family members from *Arabidopsis* (AtOpt1 and AtOpt5) were also included in the analysis (Koh *et al.*, 2002), as well as two sequences from the Fungi kingdom, *Candida albicans* CaOpt1 (Lubkowitz *et al.*, 1997) and *S. cerevisiae* ScOpt1 (Hauser *et al.*, 2000). The tree shown in Figure 2 was obtained after 6603 steps of a heuristic search using branch-swapping options. The Consistency Index (CI) was 0,6441. The YS homologues were clustered separately from the representatives of the OPT family. OsYS16 and OsYS17 figured at the base of the YS branch, followed by OsYS1, showing that these proteins were more divergent from maize YS1 and the other YS members were more closely clustered. These results indicated that these new putative genes found in rice are homologous to the maize YS and probably have a similar transporter function.

The ferric reductase oxidase (FRO) family

In strategy I plants, the ability to change the redox state of Fe^{3+} in the soil was early recognized as an obligatory step for iron uptake (Chaney *et al.*, 1972). The ferric-chelate reducing activity of yeast cells provided the first insight into the molecular basis of this physiological function. The *FRE1* gene was isolated based on complementation of the *frd*-1 yeast mutant, defective in the ferric-chelate reductase activity (Dancis *et al.*, 1992).

By means of PCR screenings of cDNA and genomic libraries using degenerate primers specially designed for the conserved cofactor binding sites from yeast and human sequences, it was possible to isolate the *AtFRO1*, *AtFRO2* and *AtFROHC* genes from *Arabidopsis* (Robinson *et al.*, 1997; Robinson *et al.*, 1999) and *PsFRO1* from pea (*Pisum Sativum*) (Waters *et al.*, 2002). The genomic location of

Table 1 - New sequences found in the rice genome homologous to the maize Fe^{3+}-phytosiderophore transporter YS1. The new genes predicted from IRGSP clones are shown in the first column according to the proposed nomenclature. The exon positions shown in the sixth column correspond to the related IRGSP clone order. The accession numbers in brackets correspond to protein sequences.

Sequence	IRGSP BAC/PAC accession number	BGI scaffold accession number	Chromosome position	CDS length	Predicted exons positions relative to the IRGSP genomic clones	Gene accession number	ESTs accession number	Predicted number of transmembrane domains	Identities with protein ZmYS1
OsYs1	AP002855	AAAA01004433	Chrm. 1 42.2 cM	2127	(82624-82968); (83124-83284); (83439-83538); (83617-83938); (84598-85796)	OSJNBa0086P08.15 (BAB17205)	AU108928	16	37%
OsYs2	AP003734	AAAA01000694	Chrm. 1 141.6 cM	2037	(121967-121632); (121517-121357); (121255-121156); (121070-120941); (120792-120597); (120489-119376)	BK000548 (DAA00248)	no one	13	37%
OsYs3	AP004121	AAAA01006534	Chrm. 2 4.7 cM	2085	(84196-85729); (86323-86873)	BK000597 (DAA01398)	BE039869	12	41%
OsYs4	AP004121	AAAA01007950	Chrm. 2 4.7 cM	2052	(90808-90473); (90384-90130); (90031-89898); (89814-88488)	BK000598 (DAA01399)	AU182396; AU183814; C98607; C98536; BE230224; C98608; AU183815; C72032; C72649; C98535; AU165407; C98740;AU182428; C98741; BE230485	14	34%
OsYs5	AP004192	AAAA01002189	Chrm. 2 102.9 cM	2184	(74992-75480); (76821-76980); (77068-77168); (77267-77397); (77630-77822); (77938-79050)	BK000599 (DAA013400)	C91598; D24925; AU064354; AU165061; AU064343; AU064290; AU031949; D21962	12	47%
OsYs6	AP004868	AAAA01008088	Chrm. 2 105.8 cM	2031	(56302-56685); (56793-56966); (57780-57870); (57950 -58089); (58193 -58386); (58478 -58592); (58671-59603)	BK000600 (DAA013401)	D42870; BE228381; BE230357	14	68%
OsYs7	AP004175	AAAA01002699	Chrm. 2 106.6 cM	2019	(21016-20639); (19919-19750); (18774-18675); (18572-18453); (18342-18148); (18061-17949); (17823-16890)	BK000601 (DAA013402)	AU031880; AU082716; AU184448; AU032408; D24664; D24682;	10	75%
OsYs8	AL731600	AAAA01003711	Chrm. 4 48.3 cM	2037	(13665-14036); (14721-14891); (15030-15128); (15801-15935); (16094-16285); (16705-16837); (16925-17859)	BK000602 (DAA013403)	AU032502; BI118727; BI809647; BM419938; BI810349	12	49%
OsYs9	AL731600	AAAA01003711	Chrm. 4 48.3 cM	1959	(19731-19742); (20165-20441); (20690-20862); (21090-21186); (21848-21984); (22538-22731); (23007-23135); (23333-24272)	BK000603 (DAA013404)	no one	15	44%
OsYs10	AL663010	AAAA01008180	Chrm. 4 81.7 cM	2139	(2996-2550); (2078-1915); (1825-1729); (1638-1508); (1420-1225); (1139-36)	BK000604 (DAA013405)	no one	14	45%
OsYs11	AL663010	AAAA01008180	Chrm. 4 81.7 cM	2154	(8963-8499); (7160-7000); (6923-6824); (6713-6583); (6231-6038); (5937-4835)	BK000605 (DAA013406)	AU068199; AU068198; AU068880	14	49%
OsYs12	AL663010	AAAA01001289	Chrm. 4 81.7 cM	2067	(16227-16294); (16403-16703); (17836-17997); (18090-18185); (18297-18431); (18538-18731); (18803-19913)	BK000606 (DAA013407)	AU181651; BM421051; BI812310	13	47%
OsYs13	AL606695	AAAA01006916	Chrm. 4 83.0 cM	2055	(74403-73982); (73885-73716); (72612-72509); (72421-72293); (72178-71985); (71903-71792); (71715-70792)	BK000607 (DAA013408)	no one	10	63%
OsYs14	AL606695	AAAA01002849	Chrm. 4 83.0 cM	2148	(103442-1029\36); (102843-102675); (102275-102175); (102053-101922); (101797-101607); (101468-101350); (101259-100331)	BK000608 (DAA013409)	AU082725; AU082724	12	67%
OsYs15	AL606457	AAAA01001708	Chrm. 4 123.8 cM	2223	(68683-69204); (70036-70195); (70316-70416); (70523-70654); (70901-71089); (71161-72279)	BK000609 (DAA013410)	no one	14	45%
OsYs16	AC104274	AAAA01005390	Chrm. 5 53.5 cM	2061	(106079-105742); (103614-103446); (103328-103224); (103127-102996); (102388-102200); (101276-100149)	BK000610 (DAA013411)	no one	15	32%
OsYs17	AC121362	AAAA01005390	Chrm. 5 53.5 cM	2064	(5728-5395); (5112-4946); (3123-3025); (2929-2800); (2245-2055); (1473-331)	BK000611 (DAA013412)	no one	15	31%
OsYs18	AP004556	AAAA01007771 AAAA01015634	Chrm. 8 53.7 cM	2097	(145831-146208); (150603-150866); (151077-151205); (151381-152706)	BK000612 (DAA013413)	no one	13	38%

Figure 2 - Phylogenetic analysis of the Yellow Stripe family using PAUP (Swofford, 2002) after alignment of the sequences with Clustal W (Thompson *et al.*, 1994). The resulting tree shows the relationships between the new 18 members predicted from the rice genome (OsYS) and the eight representatives found in *Arabidopsis* (AtYS) and the *Zea mays* YS1 (ZmYS1). Opt members from *Arabidopsis* (AtOpt), *Candida albicans* (CaOpt1) and *S. cerevisiae* (ScOpt1) were joined together as an outgroup. Accession numbers of nucleotide sequences and corresponding predicted OsYS proteins are shown in Table 1. Accession numbers to the other proteins are as follows: ZmYS1: AAG17016; AtYS1: NP_567694; AtYS2: NP_197826; AtYS3: NP_200167; AtYS4: NP_198916; AtYS5: NP_564584; AtYS6: NP_566806; AtYS7: NP_176750; AtYS8: NP_564525; AtOpt1: BAB08658; AtOpt5: CAB43855; CaOpt1: AAB69628; ScOpt1: NP_012323. Values associated with tree branches indicate the number of times (in percent) that each branch topology was found during bootstrap analysis, with 1000 replications.

AtFRO2 proved to be allelic to the ferric-chelate reductase impaired mutation, that could be restored with the ectopic expression of *AtFRO2* (Robinson *et al.*, 1999). PsFRO1 also demonstrated ferric ion-chelating activity in assays using *PsFRO1*-expressing yeast cells (Waters *et al.*, 2002). The expression patterns of *AtFRO2*, *AtFROHC* and *PsFRO1* in roots were increased under limiting iron availability. The analysis of predicted proteins from these new genes puts them in the same superfamily of flavocytochromes, to which belong the yeast FRE1 protein and a subunit of the respiratory burst oxidase (gp91-phox) from human neutrophyls (Robinson *et al.*, 1997; Robinson *et al.*,

1999; Waters *et al.*, 2002). Additional homologues have been detected in *Arabidopsis* and rice which could presumably act as burst oxidases generating reactive oxygen species as a defense against pathogens (Groom *et al.*, 1996; Torres *et al.*, 1998). A common feature of this superfamily is the high similarity around the FAD, NADPH and heme binding sites, suggesting catalytic similarity between FRO and mammalian proteins. The main difference is that in the FRO members Fe^{3+} acts as the final acceptor for the transferred electron instead of molecular oxygen, as in the case of burst oxidases (Robinson *et al.*, 1999).

We have searched for homologous genes in the genomic and expressed rice sequences using the PsFRO1 and AtFRO2 proteins as query sequences. Only two genomic sequence patterns with significant BLAST hits were found in the IRGSP clones and BGI scaffolds. The predicted CDS showed related proteins with lengths and structures similar to the query sequences and identities higher than 25% with the reference proteins. The new CDS were named *Oryza sativa FRO1* (*OsFRO1*) and *OsFRO2* (Table 2). *OsFRO*2 is a functional gene since there are four identical rice ESTs found in the expressed gene databases. There is only one wheat cDNA with high similarity for *OsFRO*1 (529 identical nucleotides of 538 as shown by BLAST alignment), derived from a substracted cDNA library of cold -stressed seedlings. Excluding this cDNA and a few other ESTs from wheat, it is interesting to note that most expressed genes similar to *OsFRO1* found in databases are from non-graminaceous plants.

The derived OsFRO proteins were analyzed for putative transmembrane domains. OsFRO1 contains 10 transmembrane spanning regions whereas OsFRO2 has 11. These proteins were also searched against the Conserved Domain Database (CDD), presenting significant hits only with the profile of Ferric reductase transmembrane family, to which gp91-phox and FRE1 from yeast and the *Arabidopsis* and pea FRO proteins belong. In order to examine with more attention the similarities of the FRO family proteins, we performed a multiple alignment of the predicted OsFRO polypeptides with the previously described FRO members from *Arabidopsis* and pea, also including AtRbohF, a respiratory burst oxidase member from *Arabidopsis* (Figure 3). The OsFRO proteins have all the expected sites for heme and dinucleotide cofactor binding, suggesting a similar functional activity for these new enzymes. OsFRO1 has greater than 39% identity with the reference proteins. OsFRO2 showed only 26% identity with AtFRO2 and PsFRO1, within a range (20-30% identity) for which it is difficult to infer homology, the so-called "twilight zone" (Rost and Valencia, 1996). However, considering the highly conserved amino acids around the catalytic sites (Figure 3), it is reasonable to regard OsFRO2 as functionally related to other FRO proteins.

A dendrogram was constructed based on a multiple alignment including the previously compared sequences,

Table 2 - Orthologs found in the rice genome relative to the ferric-chelate reductases of *Arabidopsis* (AtFRO2) and pea (*Pisum sativum*, PsFRO1). Column 2 shows the genomic clone accession number from IRGSP where the ortholog was found based on a BLAST search. The third column corresponds to the relative BGI scaffold accession number. All the predictions (shown in column 6) were made relative to the IRGSP clones since their sequences are produced more accurately. Accession numbers in brackets correspond to protein sequences.

Sequence	IRGSP BAC/PAC accession number	BGI scaffold accession number	Chromosome position	CDS length	Predicted exon positions relative to the IRGSP genomic clones	Gene accession number	EST accession number	Identities with AtFRO2	Identities with PsFRO1
OsFRO1	AL606624	AAAA01018388	Chr. 4 94.4cM	2001	(64494-64342); (64188- 64097); (63977- 63783); (63662-63428); (63344-63017); (62939-61942)	BK 000589 (DAA 01390)	-	39%	40%
OsFRO2	AL731593	AAAA01001281 AAAA01077767	Chr. 4 66.4cM	2262	(90658-90430); (90317- 90221); (89835-89638); (89192-88958); (90658-90430); (90317- 90221); (89835-89638); (89192-88958); (88656-88389); (88301-88031); (87930-87675); (87591-87135); (87003- 86813)	BK 000590 (DAA 01391)	BI813613; BI813084; BI805110; AA753007	26%	26%

six *Arabidopsis* FRO relatives found in the GenBank non redundant database and another burst oxidase sequence from rice (Figure 4). The resulting tree was obtained after 3720 steps of a branch and bound search. The Consistency Index was 0,8798. The tree topology clearly puts the burst oxidases in a separate clade from the FRO sequences. There is also a subdivision in the FRO family in two separate groups. OsFRO1 was clustered in the group comprising the AtFRO1, 2 and C and the PsFRO1, indicating that this rice protein is closely related to the previously described members of the FRO family. OsFRO2 was grouped together with three *Arabidopsis* proteins that present a lower degree of similarity with AtFRO2 and PsFRO1 (nearly 25% identity and 45% similarity) but are clearly distinct from the burst oxidases members.

The ZIP family

Members of the ZIP family are amongst the most studied metal transporters in plants (Guerinot, 2000; Mäser *et al.*, 2001). *Arabidopsis IRT1* is the best characterized so far. Functional complementation assays of yeast mutants clearly established a preference for the transport of Fe^{2+} instead of Fe^{3+} (Eide *et al.*, 1996). Moreover, *IRT1* is expressed in roots under iron starvation and the corresponding gene product only accumulates under the same circumstances, even when the gene is overexpressed by means of a 35S promoter construct. This pattern suggests control at the protein stability level (Connolly *et al.*, 2002). Reporter gene expression and *in situ* hybridization of the IRT1 protein showed its localization at the plasma membrane of root epidermal cells. These results, combined with the fact that *Arabidopsis IRT*1 knock out mutants are lethal, are strong evidence that this protein is the main component responsible for iron uptake from soil in limiting conditions of this mineral (Vert *et al.*, 2002). *IRT*2 from *Arabidopsis* is also expressed during iron starvation in root epidermal cells (Vert *et al.*, 2001).

There is molecular evidence for a function in cation transport for *Arabidopsis ZIP1*, *ZIP2* and *ZIP3*, as they could restore zinc uptake in the yeast double mutant *zrt1* and *zrt*2. Northern blot experiments showed that *ZIP1* and *ZIP3* are expressed in roots under zinc deficiency, while *ZIP4* is expressed both in roots and shoots under the same condition. Therefore, these ZIP members have been implicated in Zn^{2+} uptake from the soil (Grotz *et al.*, 1998). A feature observed for some members of this family is a broader affinity for cation transport. The *Arabidopsis* ZIP members are currently being investigated by reverse genetic methods, which will provide further information about the molecular functions of these genes (Mäser *et al.*, 2001).

The IRT1, ZIP1 and ZIP2 *Arabidopsis* proteins were used as query sequences to search for ZIP family homologues in rice. Several sequences were found, corresponding to thirteen putative predicted genes (Table 3). The
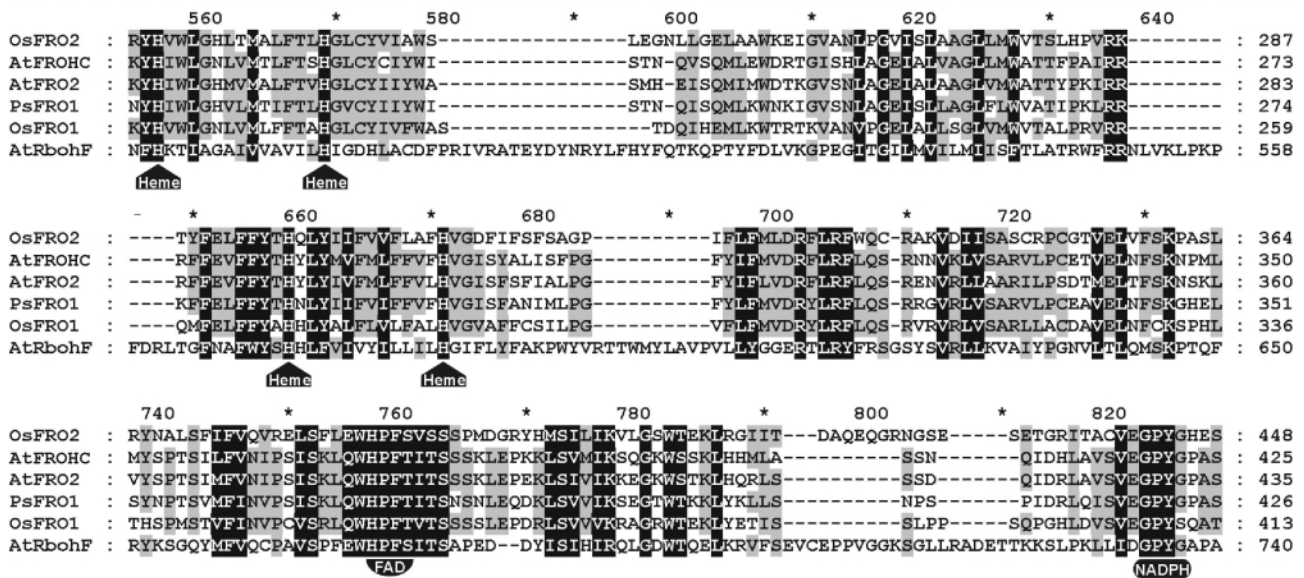
Figure 3 - Conservation of amino acids around the prosthetic binding sites in the FRO family. The multiple alignment was performed using ClustalW algorithm. Only the central parts of the proteins are shown. The four heme-coordinating histidine (H) residues and the dinucleotide binding motifs (FAD, HPFT and NAPH, GPYG) are indicated under the block of alignments in the respective positions. Accession numbers of nucleotide sequences and corresponding predicted OsFRO proteins are shown in Table 2. Accession numbers to the other proteins are as follows: AtFrohC: AAB72168; AtFRO2: CAA70770; PsFRO1: AAK95654; AtRbohF (*Arabidopsis thaliana* Respiratory burst oxidase F): BAA28953. Black color indicates regions in the alignment with 100% similarity and gray color indicates regions with at least 60% similarity.

previously described gene *OsIRT1* finds matching sequences in the BGI scaffolds AAAA01002856 and AAAA01010833. A gene product presenting 77% identity with OsIRT1 and 49% identity with AtIRT1 was predicted based on the BGI AAAA01010833 scaffold. Because of its high similarity with *OsIRT1*, this gene was named *OsIRT2*. The other eleven predicted genes, due to their greater similarity to *AtZIP1* and *AtZIP2*, were named OsZIP1 to *OsZIP11*. Most CDS were predicted from clones sequenced by the IRGSP, whereas *OsZIP9* to *OsZIP11* and *OsIRT1* and OsIRT2 were only found in BGI scaffolds. CDS *OsZIP9*, *OsZIP10* and *OsZIP11* are incomplete, since the corresponding scaffolds are of short extension. Sixteen ESTs corresponding to six of these predicted CDS confirm that at least these six genes are functional (Table 3). The fact that *OsIRT1* and *OsZIP9* to *OsZIP11* were exclusively predicted from genomic sequences corresponding to BGI scaffolds suggests that these could be genes restricted to the indica variety.

The ESTs related to the *OsIRT*2 gene were generated from cDNAs derived from both the japonica variety, cultivar Nipponbare, and the indica variety, cultivar Pokkali. The same was observed for the other genes with ESTs that correspond to cDNAs derived from different varieties (Table 3). Therefore, the greater number of genomic sequences related to the indica variety may reflect the higher genomic coverage level of the BGI libraries.

The predicted gene sequences were translated into amino acids and compared both by multiple alignment and phylogenetic analysis to the *Arabidopsis* ZIP proteins and yeast ZRT1 and ZRT2. Four trees were obtained after 4058 steps of a heuristic search with branch and bound options. A strict consensual method generated the tree depicted in Figure 5. The Consistency Index was 0.6353. The tree topology was tested with 1000 Bootstrap replications. The proteins resulting from the new rice sequences showed an extension from 350 to 390 amino acid residues, with a similarity among themselves ranging from 33 to 77% and more than 35% identity to at least one of the three reference proteins used in the searches (Table 3). All these proteins are grouped in the phylogenetic tree together with the *Arabidopsis* ZIP polypeptides in a superclade that separated them from the yeast ZRT proteins. OsZIP1 and OsZIP6 are more distantly related to the other ZIP proteins, being grouped with AtZIP2 and AtZIP11, closer to the root of the ZIP family.

Transmembrane domains were predicted for the OsIRT1, OsIRT2 and OsZIP1 to OsZIP8 proteins, leaving out the incomplete polypeptides OsZIP9, 10 and 11. Seven to nine transmembrane domains were found (Table 3), in good agreement with the eight regions observed for the *Arabidopsis* ZIP family members (Mäser *et al.*, 2001). Also in line with the *Arabidopsis* ZIP proteins, a central portion with a low degree of conservation, the so called variable region (Guerinot, 2000), was also present in the rice ZIP proteins. This region contained, in the majority of examined sequences, several histidine residues, presumably metal binding sites. Three other histidine residues, located in the II, IV and V transmembrane domains, are conserved in all rice ZIP sequences. Point mutations in these residues of the
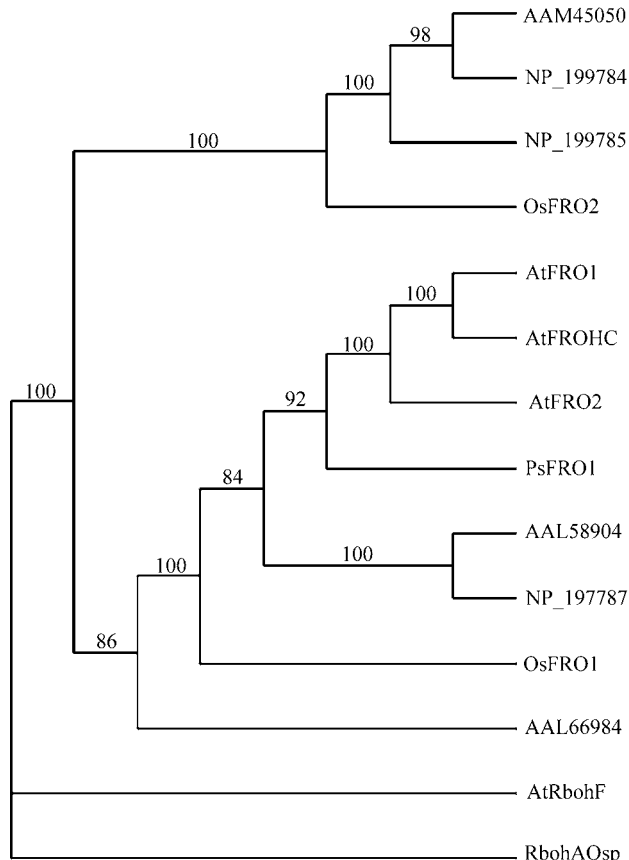
Figure 4 - Phylogenic analysis of the FRO protein family. Accession numbers of nucleotide sequences and corresponding predicted OsFRO proteins are shown in Table 2. Accession numbers to the other proteins are as follows: AtFRO1: CAA70769; AtFRO2: CAA70770; AtFrohC: AAB72168; PsFRO1: AAK95654; AtRbohF (*Arabidopsis thaliana* Respiratory burst oxidase F): BAA28953; RbohAOsp (*Oryza sativa* respiratory burst oxidase A): AAB87790. The six additional *Arabidopsis* paralogs are discriminated in the dendrogram with the respective accession numbers.

Figure 5 - Phylogenetic analysis of the ZIP protein family from *Oryza sativa* and *Arabidopsis* with Zinc regulated transporters ZRT1 and two from *S. cerevisiae* as outgroup. Accession numbers of nucleotide sequences and corresponding predicted OsZIP proteins are shown in Table 3. Accession numbers to the other proteins are as follows: AtIRT1: AAB01678; AtIRT2: NP_193703; AtIRT3: NP_564766; AtZIP1: AAC24197; AtZIP2: AAC24198; AtZIP3: AAC24199; AtZIP4: NP_172566; AtZIP5: NP_172022; AtZIP6: NP_180569; AtZIP7: AAL38434; AtZIP8: AAL83293; AtZIP9: NP_195028; AtZIP10: NP_174411; AtZIP11: NP_564703; AtZIP12: NP_201022; ScZRT1: P32804; ScZRT2: NP_013231.

Arabidopsis IRT1 abolished the protein capacity to transport iron and other minerals (Rogers *et al.*, 2000). The conservation of these same residues in the rice ZIP proteins suggests an important catalytic function kept unaltered in the large gene family.

The NRAMP family

The first *Nramp* (Natural Resistance-Associated Macrophage Protein) gene to be identified was the mammalian *Nramp1*, which encodes a macrophage membrane protein believed to control divalent cation concentrations within the phagosome, and, by this means, to regulate replication of engulfed bacteria (Williams *et al.*, 2000). The NRAMP protein family, with members also in *Drosophila melanogaster*, *Saccharomyces cerevisiae* and rice (OsNRAMP1), was defined by a highly conserved hydrophobic core with 10 transmembrane domains, several invariant charged residues and other features typical of transporters or channels (Cellier *et al.*, 1995).
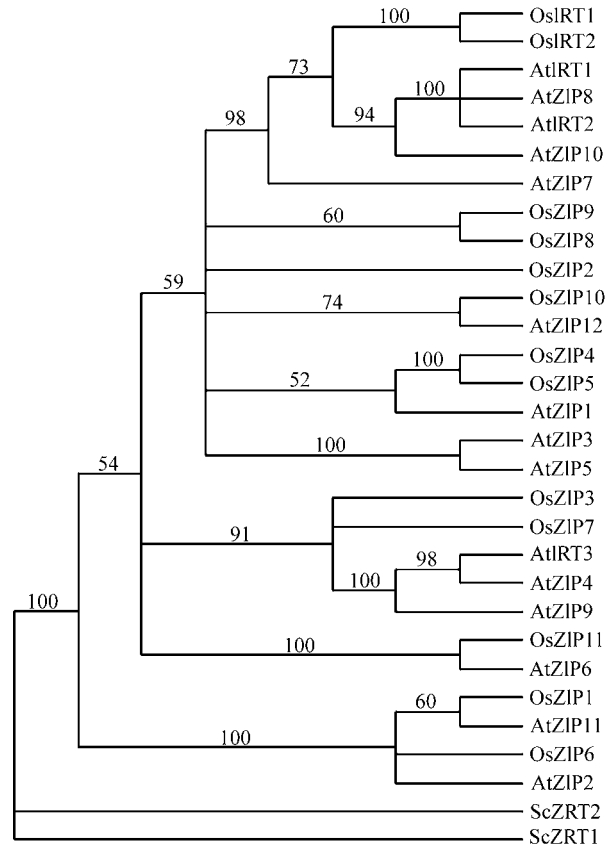
Two groups presented evidence for the involvement of plant NRAMP proteins in iron transport. Curie *et al.* (2000) showed that both rice and *Arabidopsis* NRAMP1 were able to complement the *fet3fet4* yeast mutant, defective in low- and high-affinity iron transport, with the *Arabidopsis* protein being more efficient than the rice one. They also showed that *AtNramp1* transcript accumulated in response to iron deficiency in roots. Furthermore, overexpression of AtNRAMP1 in *Arabidopsis* increased its resistance to toxic iron concentrations. Thomine at al. showed that *AtNramp3* and *AtNramp4* were more efficient than *AtNramp1* in complementing the iron uptake yeast mutant *fet3fet4*. Iron starvation resulted in up-regulation of expression of the *Arabidopsis Nramp* genes. Disruption of *AtNramp3* led to increased cadmium resistance and overexpression of the same gene led to cadmium hypersensitivity. It was suggested that the AtNRAMP3 protein may be involved in iron uptake in roots, with its contribution rel-

Table 3 - Paralogs found in the rice genome relative to the ZIP family. When possible the new genes were predicted from IRGSP clones, since their sequences are produced more accurately. For many genes there is more than one corresponding genomic sequence. The accession number related to the genomic sequence used to predict the exon positions shown in the sixth column is underlined. *OsZip5* was predicted from 3 different genomic sequences since these clones were incomplete or presented bad quality at the gene location. Sequences *OsZip9*, 10 and 11 are incomplete, thus the identities shown in columns 10, 11 and 12 are relative to alignments with reference sequences restricted only to the known extension of the incomplete proteins. Accession numbers in brackets correspond to protein sequences.

Sequence	IRGSP BAC/PAC accession number	BGI scaffold accession number	Chromosome position	CDS length	Predicted exons positions relative to genomic clones	Number of transmembrane domains	Gene accession number	ESTs accession number	Identi-ties with AtIRT1	Identi-ties with AtZIP1	Identi-ties with AtZIP2
OsIrt1	-	AAAA01002856 AAAA01010833	?	1125	AAAA01002856: (10673-11338) AAAA01010833: (822-888)	8	AB070226 (BAB85123)	BE040809[a]	48%	39%	15%
OsIrt2	-	AAAA01010833	?	1113	AAAA01010833: (9174-8518); (8437-7982)	7	BK000613 (DAA01414)	D49213[b] AU174128[b] BE040823[a]	49%	36%	13%
OsZip1	AC092262	AAAA01020606 AAAA01003957	Chr.3 83,3cM	1077	AC092262: (40976-40485); (39838-39779); (39662-39534); (38390-37995)	9	BK000614 (DAA01415)	BQ907959 BQ909168 D47931[b] AU032700[b]	14%	17%	46%
OsZip2	AP004297 AP005246	AAAA01036704 AAAA01011601	Chr.7 46,5cM	1173	AP004297 : (51814-51305); (51229-51100); (50423-50272); (49198-48893)	8	BK000615 (DAA01416)	-	35%	41%	14%
OsZip3	AP005456 AP005458	AAAA01026116 AAAA01002944	Chr.6 76,5cM	1083	AP005456: (116028-116098); (116231-116286); (116393-116892); (117414-117563); (117667-117972)	8	BK000616 (DAA01417)	-	35%	35%	14%
OsZip4	AL731610	AAAA01002111	Chr. 4 102,7cM	1095	AL731610: (63256-63894); (64227-64376); (64884-65189)	8	BK000617 (DAA01418)	AU030756[b] AU030757[b]	35%	48%	14%
OsZip5	AP003875 AP003919	AAAA01083113 AAAA01009964	Chr.8 44,6cM	1191	Exon 1: (AP003875: 125127-125576 + AAAA01083113: 371-249); AP003919: Exon 2: (48203-48367); Exon 3:(48894-49043); (49878-50183)	9	BK000618 (DAA01419)	-	34%	46%	14%
OsZip6	AP003277	AAAA01061663 AAAA01028642 AAAA01050973	Chr.1 181.8cM	1059	AP003277: (93501-92838); (92271-91877)	9	13810566 (BAB63683)	D23954[b] AU032374[b] AU164595[b] AU032375[b]	17%	16%	53%
OsZip7	AC136519	AAAA01010830	Chr. 5 51,8 cM	1107	AC136519: (104427-103782); (103686-103531); (102290-101986)	7	BK000619 (DAA01420)	-	34%	38%	13%
OsZip8	AC135423	AAAA01012059	Chr. 5 95,8 cM	1062	AC135423: (136570-137173); (137321-137473); (138389-138693)	7	BK000620 (DAA01421)	AA751631[c] AU162281[b]	37%	42%	15%
OsZip9	-	AAAA01063001	?	N-term. region: 633	N-term. Region: AAAA01063001: (161-793)	-	BK000621 (DAA01422)	-	36%	38%	9%
OsZip10	-	AAAA01018830	?	C-term. region: 309	C-term. Region: AAAA01018830: (2544-2852)	-	BK000622 (DAA01423)	-	46%	56%	21%
OsZip11	-	AAAA01016699	?	C-term. region: 508	C-term. region: AAAA01016699: (19-211); (1547-1861)	-	BK000623 (DAA01424)	-	47%	50%	18%

[a]EST derived from cDNA library provenient from *Oryza sativa* indica variety Pokkali.
[b]EST derived from cDNA library provenient from *Oryza sativa* japonica variety cultivar Nipponbare.
[c]EST derived from cDNA library provenient from *Oryza sativa* indica variety cultivar Milyang23.

ative to IRT1 being dependent on pH (Thomine *et al.*, 2000).

The sequences of the *Arabidopsis* NRAMP1 and NRAMP2 proteins were used as queries to search for NRAMP orthologs in rice. Besides the three *NRAMP* genes previously known in rice (Belouchi *et al.*, 1997), five new members of this family were predicted based on BGI and IRSGP genomic sequences. These new putative genes were called *OsNRAMP4* to *OsNRAMP8* (Table 4). The previously deposited (U60767) CDS Os*NRAMP3* was incomplete, so we used the predicted complete CDS, based on the IRGSP genomic sequence AP004989, in our analyses. *OsNRAMP7* was annotated on the basis of two genomic clones, since the AL772426 sequence has a gap in the region corresponding to the 5'end of the gene. The complete CDS was obtained using scaffold AAAA01007693, that provided the missing region. The *OsNRAMP* sequences were approximately 1500 bp long. The corresponding genes had 12 exons in *OsNRAMP6* and 13 exons in *OsNRAMP1*, *3*, *4* and *5*. In contrast, *OsNRAMP8* had only 3 exons and *OsNRAMP2* and *7* had 4 exons. There were nineteen ESTs related to *OsNRAMP1*, *2*, *5*, *6* and *7*, indicating that these are expressed genes. For the other predicted genes, molecular evidence needs to be acquired in order to fully establish their functionality. The two *NRAMP* members with the largest number of related ESTs, *OsNRAMP1* and *OsNRAMP5*, were found in the same chromosomal locus, suggesting a recent event of gene duplication.

The predicted CDS for the eight *OsNRAMP* genes were translated into amino acids and analyzed in multiple alignment comparatively to the *Arabidopsis* NRAMP. All gene products showed over 35% identity with the reference proteins (Table 4). OsNRAMP2, 7 and 8 had higher similarity with AtNRAMP2, displaying 60 to 66% identity with this protein. The other polypeptides were more closely related to AtNRAMP1, reaching up to 68% identity in the case of OsNRAMP3. Also, the OsNRAMP products were highly conserved when compared among themselves, showing up to 72% identity and 81% similarity, as in the case of OsNRAMP7 and OsNRAMP8. All of the OsNRAMP proteins were highly conserved in the region referring to the Consensus Transport Motif (CTM), located between two transmembrane regions in the carboxy-terminal region of the proteins. This motif was first identified in bacteria and is present in all *Arabidopsis* NRAMP proteins (Curie *et al.*, 2000). The transmembrane domain prediction showed that most of the proteins had 10 of these domains, two of them had 11 transmembrane domains and one had only 8 (Table 4). The phylogenetic analysis of protein sequences of the rice NRAMP, *Arabidopsis* NRAMP and SMF proteins of *Saccharomyces cerevisiae* is shown in Figure 6. The depicted tree was obtained after 2765 steps of a heuristic search with branch-swapping options. The Consistency Index was 0.8119. The resulting topology clearly placed the NRAMP family in a different branch from the yeast SMF proteins. The NRAMP family was arranged in 2 separated clades. One grouped the proteins similar to AtNRAMP1 and the other involved those related to AtNRAMP2. In fact, the previously known *Arabidopsis* and rice NRAMP proteins seem to belong to two sub-classes, one including AtNRMP1 and 6 and OsNRAMP1 and 3 and the other including AtNRAMP2 to 5 and OsNRAMP2 (Curie *et al.*, 2000; Maser *et al.*, 2001). In the present study, a correlation was apparent between the number of exons of *OsNRAMP* genes and the observed branching of the NRAMP protein phylogenetic tree. The proteins with less exons present in their genomic sequences, OsNRAMP2 (4 exons) and OsNRAMP7 and 8 (3 exons each), were grouped in the same branch with AtNRAMP2. The NRAMP proteins with corresponding genomic sequences having a larger number of exons, OsNRAMP6 (12 exons) and OsNRAMP1, 3, 4 and 5 (13 exons each) were in a different branch, clustered with AtNRAMP1.

The ferritin family

Ferritins are iron storage proteins with highly conserved structure among plants, animals and bacteria. The protein has 24 subunits organized in a spherical shell, and can store up to 4500 iron atoms in its central cavity (Briat and Lobréaux, 1997). In mammals, two different types of ferritin subunits are assigned different functions: the H subunit contains a ferroxidase site responsible for rapid Fe^{2+} oxidation, leading to rapid iron uptake to inside the protein, and the L subunit lacks this site but is rich in E residues facing the central cavity of the protein, allowing better nucleation of Fe^{3+} for long-term storage (Petit *et al.*, 2001b and references therein). Plant ferritins combine both functions in their subunits, which share higher homology with animal ferritins than with bacterioferritins, although they are localized in the plastids, a typical prokaryotic environment (Briat and Lobréaux, 1998).

Plant ferritin genes do not contain IREs (Iron Regulatory Elements), conserved regulatory sequences responsible for the translational derepression of animal ferritin mRNA in response to iron overload (Petit *et al.*, 2001b and references therein). Instead, in the promoters of maize and *Arabidopsis* ferritin genes, the presence of Iron-Dependent Regulatory Sequences (IDRS), responsible for repression of transcription under low iron supply, was recently shown (Petit *et al.*, 2001a). In maize, regulation of *ZmFer1* by iron is ABA-independent and requires the IDRS, while induction of its expression by H_2O_2 does not require a functional IDRS. One form of maize ferritin (*ZmFer2*) is expressed in shoots in response to ABA treatment, as well as one *Arabidopsis* ferritin (*AtFer2*), which is expressed in seeds (Fobis-Loisy *et al.*, 1995; Petit *et al.*, 2001b). *AtFer1*, *3* and *4* are expressed in response to iron overload, although with different kinetics in the latter. *AtFer1* and, to a lesser extent, *AtFer3*, are also expressed in response to H_2O_2. Both *AtFer3* and *AtFer4* promoters have partly conserved IDRS

Table 4 - New sequences found in the rice genome homologous to the NRAMP family. The genes *OsNramp1* to *OsNramp6* were predicted from the IRGSP sequences. *OsNramp8* derived from a BGI scaffold while *OsNramp7* were predicted merging one IRGSP sequence with one BGI scaffold. In all cases the accession number related to the genomic sequence used to predict the exon positions shown in the sixth column is underlined. Accession numbers in brackets correspond to protein sequences.

| Sequence	IRGSP BAC/PAC accession number	BGI Scaffold accession number	Chromosome position	CDS length	Predicted exons positions relative to genomic clones	Number of transmembrane domains	Gene accession number	ESTs accession number	Identities with AtNRAMP1	Identities with AtNRAMP2
OsNramp1	AP003755	AAAA01008708	Chr. 7 49,4 cM	1554	AP003755: (108266-108185); (107140-107060); (106990-106951); (106819-106740); (106658-106556); (106315-106200); (106018-105841); (105623-105553); (105465-105294); (105160-105054); (104847-104718); (104471-104375); (104259-103960)	10	L41217 S81897 (AAB62273)	C25866 D15268 BE040339 C98020 BI805108 AU082598 BQ907728 BQ907313 BQ907757	53%	31%
OsNramp2	AC105928	AAAA01015560	Chr. 3 27,9 cM	1395	AC105928: (77591-77988); (79380-79516); (79612-80267); (80766-80970)	10	L81152 (AAB61961)	AU182598 BI810889	34%	60%
OsNramp3	AP004989	AAAA01000179	Chr. 6 109,5 cM	1653	AP004989: (148972-149084); (149455-149533); (149616-149656); (149816-149894); (150000-150104); (150381-150490); (150768-150949); (151034-151103); (151345-151518); (151842-151946); (152488-152619); (152761-152856); (152942-153292)	11	BK 000591 (DAA01392) U60767 (AAC49720)	-	68%	32%
OsNramp4	AP004077 AP004150	AAAA01001581	Chr. 2 7,9 cM	1638	AP004077: (51839-51705); (51608-51533); (51418-51378); (50481-50403); (50302-50198); (50097-49985); (49886-49707); (49601-49533); (49443-49270); (49102-48998); (48811-48680); (48568-48473); (48339-48007)	10	BK 000592 (DAA01393)	-	46%	30%
OsNramp5	AP004176	AAAA01001334	Chr. 7 49,4 cM	1617	AP004176: (66470-66531); (66798-66876); (67005-67045); (67344-67422); (67505-67609); (68609-68718); (68943-69124); (69224-69293); (69584-69757); (70039-70143); (71130-71261); (71857-71952); (72962-73294)	10	BK 000593 (DAA01394)	BE607374 AA750229 AU070750 AU091988	54%	32%
OsNramp6	AP003144	AAAA01005115	Chr. 1 73,4 cM	1527	AP003144: (114206-114012); (112489-112411); (112154-112114); (112007-111929); (111849-111745); (110798-110689); (108271-108090); (107770-107701); (107261-107091); (106811-106709); (106081-105988); (104193-103900)	8	BK 000594 (DAA01395)	BI813536 AA749826	36%	25%
OsNramp7	AL772426 AL772418	AAAA01007693	Chr. 12 94,6 cM	1626	Exon 1: (AAAA01007693: 6133-5948 + AL772426: 41630-41188) others: AL772426: (39927-39791); (39703-39054); (38536-38327)	11	BK 000595 (DAA01396)	BM419904 D25033	32%	66%
OsNramp8	AC133860	AAAA01009439	Chr. 3 89,2 cM	1608	AC133860: (123279-122640); (121691-121555); (121461-120800); (120188-120021)	10	BK 000596 (DAA01397)	-	31%	62%
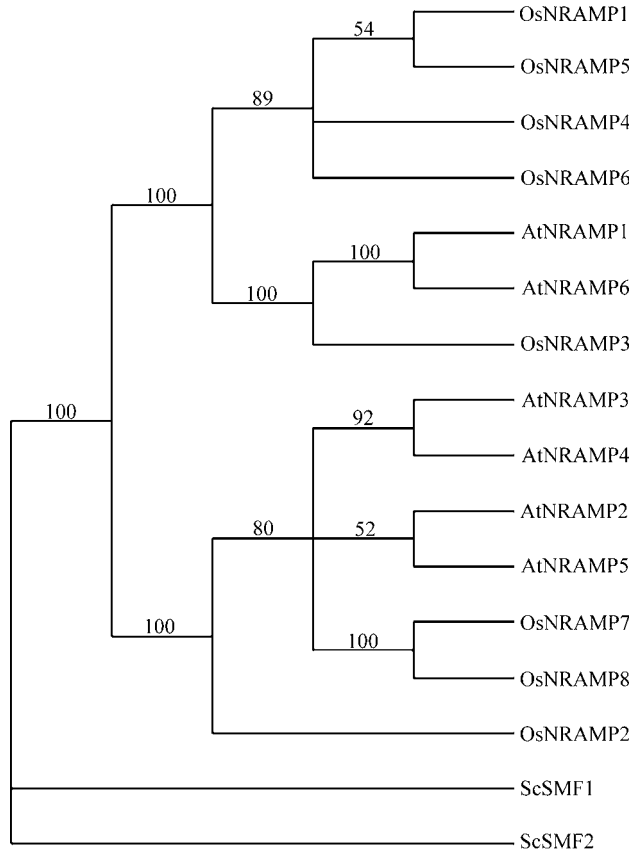
Figure 6 - Phylogenetic analysis of the NRAMP protein family. The new OsNRAMP proteins were analyzed together with the NRAMP proteins from *Arabidopsis* and the *S. cerevisiae* SMF1 and SMF2 polypeptides. Accession numbers of nucleotide sequences and corresponding predicted OsNRAMP proteins are shown in Table 4. Accession numbers to the other proteins are: AtNRAMP1: AAF36535; AtNRAMP2: AAD41078; AtNRAMP3: NP_179896; AtNRAMP4: AAF13279; AtNRAMP5: NP_193614; AtNRAMP6: NP_173048; ScSMF1: AAB48984; ScSMF2: B45154.

sequences, but it is not known if they are functional (Petit *et al.*, 2001b).

Post-transcriptional regulation of plant ferritin genes has also been shown, with accumulation of ferritin mRNA but not protein in the maize *ys1* mutant (with low iron content in leaves) treated with iron (Fobis-Loisy *et al.*, 1996). Expression of the soybean ferritin gene in rice seeds resulted in up to a threefold increase in iron content (Goto *et al.*, 1999).

The search for sequences homologous to the maize ferritin (ZmFER1) yielded several related sequences in rice. These include 20 ESTs, 2 sequenced cDNAs, 2 genomic sequences from the IRGSP and 4 BGI scaffolds (Table 5). A comparative analysis of ESTs showed two distinct patterns of expressed genes, which differ in the 5'region by point variations and two small deletions. In order to better analyze these differences, we sequenced two cDNAs from the japonica variety, corresponding to ESTs D22266 and AU064475, representing the two observed patterns. The full length cDNA information confirmed the

Table 5 - Sequences found in the rice genome homologous to the maize Fer1 protein. The two Ferritin sequences found in rice were obtained from full-length cDNA sequencing. The corresponding genomic sequences are in the second and third columns. The exon positions for the OsFer1 were predicted from two genomic clones (sixth column) since a quimeric sequence, mixing OsFer1 and OsFer2, was observed in the BGI scaffold with the accession number AAAA01002317. Since the two OsFer genes are very similar, it was impossible for many ESTs (shown in the tenth column) to determine if they are related to OsFer1 or OsFer2. Those which are unequivocally related to either of the two paralogs are shown in the ninth column.

Sequence	IRGSP BAC/PAC accession number	BGI scaffold accession number	Chromosome position	CDS length	Predicted exons potions relative to genomic clones	Protein accession number	Gene accession number	Unequivocally corresponding ESTs accession number	Ambiguously corresponding ESTs accession number	Identities with ZmFer1
OsFer1	BX000497	AAAA01002317 AAAA01012161	Chrm. 11 1,7 cM	756	BX000497 (13971-13685); (13285-13202); (13121-13061); (12957-12871); (12309-12272); (12146-12081); (11967-11905); (11798-11756)	AAM74942	AF519570 AX031243	D22266	D17774 D42549 D49189 C96748 AA751899 AU057833 C99470 AU082276 AA754602 AU198187	82%
OsFer2	BX000494	AAAA01008574 AAAA01002317	Chr. 12 5.5 cM	768	BX000494: (55680-55382); (54675-54592); (54511-54451); (54348-54262); (53509-53445); (53311-53250); (53136-53070); (52963-52921)	AAM74943	AF519571 AF370029	AA752950 AU064475 AU069989	AU182829 AU198187 AU198188 BI810307 BI810170 BI812873	81%

two distinct gene patterns, which were named *OsFER1* and *OsFER2*. The new genes are very similar, with 96% identity. In addition to 15 SNPs along the complete cDNA, *OsFER1* has two deletions of 12 bp not seen in *OsFER2*. One of the deletions is located in the 5'untranslated region and the other immediately downstream from the start codon. Due to the high similarity between the two sequences, very few ESTs found in the data banks could be unequivocally related to one of the gene types (Table 5, column 9). *OsFER2* is related to the BX000494 genomic sequence of the IRGSP and to two BGI scaffolds. The 5'region of OsFER1, containing the two deletions characteristic of this sequence, is found in the final part of the AAAA01012161 scaffold, whereas the rest of the gene is found in scaffold AAAA01002317. Eight exons were found in the genomic sequences *OsFER*1 and *OsFER*2, which is the number found in all plant ferritins described so far.

The CDS predicted from scaffold AAAA01002317 has similarities with portions of both genes that have been identified based on expressed sequences. There is identity with *OsFER*2 in the 19513-19409 region of the referred scaffold, as seen from the SNPs and insertions found in the *OsFER*2 sequence relative to *OsFER*1. From the 19410 position of scaffold AAAA01002317 onwards, all SNPs correspond to those observed in *OsFER*1. This is possibly due to a chimera in the assembly of the scaffold, which joined sequences derived from two very similar loci, one encoding *OsFER*2, followed by another coding for *OsFER*1. Errors in the computational assembly of sequences involving similar regions in the genome have been pointed out as the main disadvantage of the total shotgun sequencing methodology (Waterston *et al.*, 2002), employed by BGI.

The OsFER1 and OsFER2 proteins have 82 and 81% identity with the reference polypeptide ZmFER1, respectively, and share 96% identity. The protein alignment of OsFER1 and OsFER2 to six plant ferritins and to the structures of equine Heavy Chain Ferritin and human H ferritin (tridimensional structure already defined) is shown in Figure 7. The rice ferritins have more than 30% identity in relation to the mammalian ferritins and between 58 and 82% with other plant ferritins. A transit peptide and an extension peptide are only seen in plant ferritins. The transit peptide is involved in the subcellular localization of ferritin to the plastid in maize, soybean and pea (Briat and Lobréaux, 1998). The limits of the transit peptide indicated in Figure 7 are based on the extension observed in Ferritin 1 of maize (Fobis-Loisy *et al.*, 1995). The cleavage of the transit peptide generates a mature protein that has the so called extension peptide, which represents an extra portion not seen in the mammalian ferritins (Briat and Lobréaux, 1998). The latter are composed of a bundle of 5 alpha helices forming a subunit. Twenty-four of these subunits compose a sphere-like holoprotein, which has channels or specific pores formed by the junction of 3 to 4 subunits (Harrison *et al.*,

1998). The 3-fold channel (3FC) is formed by the junction of 3 subunits and is functionally related to iron incorporation. The amino acids related to the structure of this channel are hydrophilic and are identical in almost all ferritins aligned in Figure 7. The 4-fold channel (4FC) is formed by the junction of 4 different subunits. In mammalian ferritins, this channel is hydrophobic in nature and is presumably related to the release of iron from the holoprotein. The amino acids related to the 4FC and the remaining residues that compose the E helix are not much conserved between ferritins of mammals and plants, although in the latter the iron release mechanism is also related to the E-helix through its proteolytic cleavage (Masuda *et al.*, 2001). Amino acid residues related to the catalytic iron oxidation (FO) in H ferritin are conserved in plant ferritins, indicating a possible catalytic role for these residues also in plant ferritins (Briat and Lobréaux, 1998; Harrison *et al.*, 1998).

The same ClustalW alignment in Figure 7 was used as starting point for the phylogenetic analysis shown in Figure 8. The depicted tree was obtained after 542 steps of a branch and bound search. The resulting consistency index was 0.9059. The dendrogram separates the mammalian ferritins from their plant homologs. Within the plant ferritins, the four *Arabidopsis* sequences are placed in a separate clade from the grass ferritins. In a similar fashion, within the grass ferritins, there is a division separating the rice sequences from those of corn.

Analysis of the promoter regions of the newly described rice ferritin genes revealed the presence of the IDRS (Iron-Dependent Regulatory Sequence) in both of them. The sequence (CACGAGCCCGCCAC) is the same in *OsFer1* and *OsFer2*, and has one base difference in relation to both the *AtFer1* IDRS (C instead of G in the seventh position) and the *ZmFer1* IRDS (G instead of T in the tenth position) (Petit *et al.*, 2001a). Interestingly, these are the only two bases in which there are differences between the two previously reported IRDS, and the one found in rice shares one of the bases with maize and the other one with *Arabidopsis*, establishing a trend to a consensus sequence CACGAG(C or G) CC(G or T)CCAC. Experimental data is needed to characterize the functionality of this element and its responsiveness to iron levels in regulating expression of the ferritin genes in rice.

Multiplicity of transporters and members in the gene families

Mineral homeostasis in plants is just starting to be analyzed at the molecular level. A number of studies have shown that there are several types of transporters and homologs of each of them in *Arabidopsis* and other plants (Williams *et al.*, 2000; Maser *et al.*, 2001). The multiple DNA sequences related to iron homeostasis in rice found in this study is in agreement with the diversity observed in other species. What are the biological implications of different protein families involved in the transport of the same
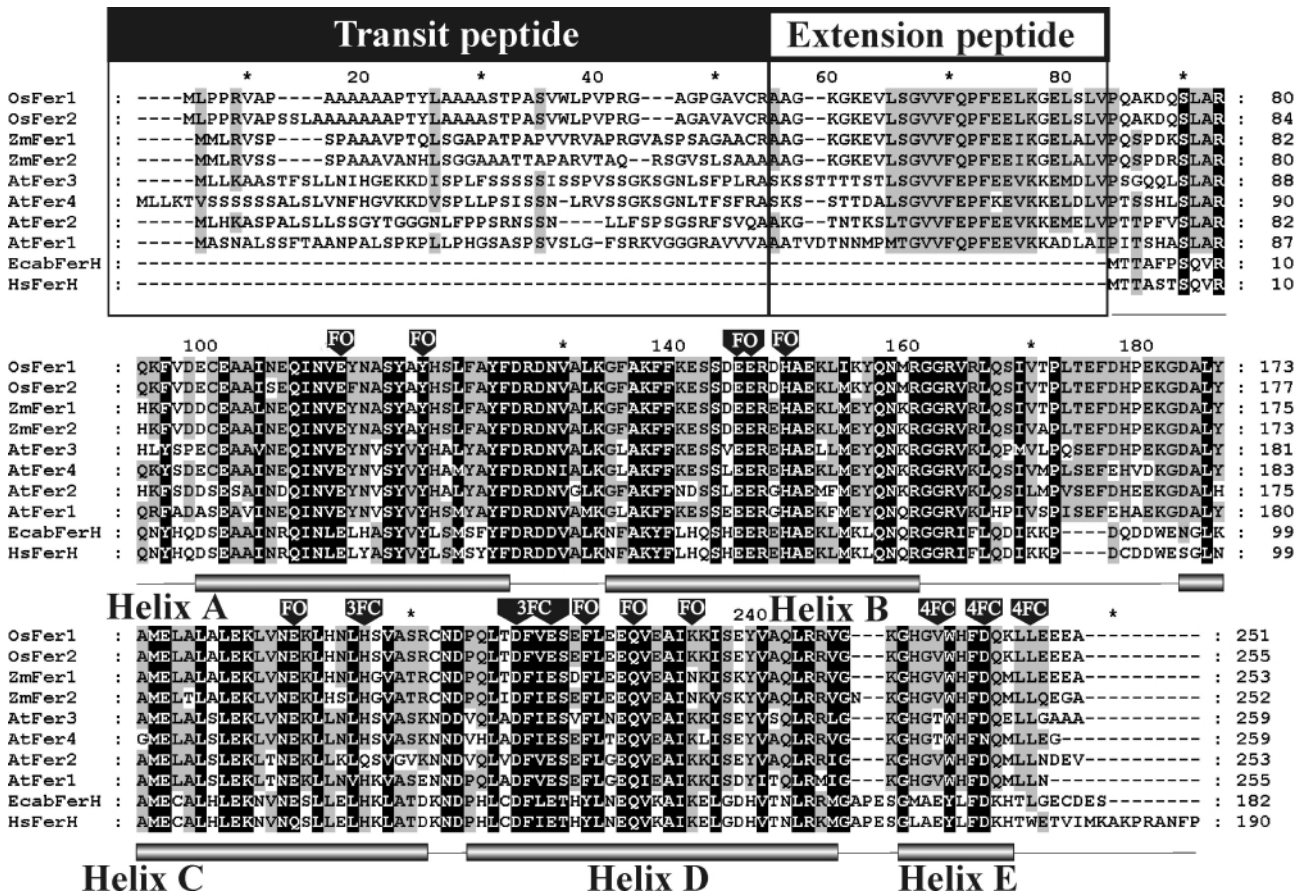
```
                    Transit peptide                        Extension peptide
                    *        20        *        40        *        60        *        80        *
OsFer1    : ----MLPPRVAP----AAAAAAPTYLAAAASTPASVWLPVPRG---AGPGAVCRAAG--KGKEVLSGVVFQPFEELKGELSLVPQAKDQSLAR :  80
OsFer2    : ----MLPPRVAPSSLAAAAAAAPTYLAAAASTPASVWLPVPRG---AGAVAVCRAAG--KGKEVLSGVVFQPFEELKGELSLVPQAKDQSLAR :  84
ZmFer1    : -----MMLRVSP----SPAAAVPTQLSGAPATPAPVVRVAPRGVASPSAGAACRAAG--KGKEVLSGVVFQPFEEIKGELALVPQSPDKSLAR :  82
ZmFer2    : -----MMLRVSS----SPAAAVANHLSGGAAATTAPARVTAQ--RSGVSLSAAAAAG--KGKEVLSGVVFQPFEEIKGELALVPQSPDRSLAR :  80
AtFer3    : -----MLLKAASTFSLLNIHGEKKDISPLFSSSSSISSPVSSGKSGNLSFPLRASKSSTTTTSTLSGVVFEPFEEVKKEMDLVPSGGQLSLAR :  88
AtFer4    : MLLKTVSSSSSSSALSLVNFHGVKKDVSPLLPSISSN-LRVSSGKSGNLTFSFRASKS--STTDALSGVVFEPFKEVKKELDLVPTSSHLSLAR :  90
AtFer2    : -----MLHKASPALSLLSSGYTGGGNLFPPSRNSSN----LLFSPSGSRFSVQAAKG--TNTKSLTGVVFEPFEEVKKEMELVPTTPFVSLAR :  82
AtFer1    : -----MASNALSSFTAANPALSPKPLLPHGSASPSVSLG-FSRKVGGGRAVVVAAATVDTNNMPMTGVVFQPFEEVKKADLAIPITSHASLAR :  87
EcabFerH  : -----------------------------------------------------------------------------MTTAFPSQVR :  10
HsFerH    : -----------------------------------------------------------------------------MTTASTSQVR :  10

                 100      FO        FO        *        140        FO FO        160        *        180
OsFer1    : QKFVDECEAAINEQINVEYNASYAYHSLFAYFDRDNVALKGFAKFFKESSDEERDHAEKLIKYQNMRGGRVRLQSIVTPLTEFDHPEKGDALY : 173
OsFer2    : QKFVDECEAAISEQINVEFNASYAYHSLFAYFDRDNVALKGFAKFFKESSDEERDHAEKLMKYQNMRGGRVRLQSIVTPLTEFDHPEKGDALY : 177
ZmFer1    : HKFVDDCEAALNEQINVEYNASYAYHSLFAYFDRDNVALKGFAKFFKESSDEEREHAEKLMEYQNKRGGRVRLQSIVTPLTEFDHPEKGDALY : 175
ZmFer2    : HKFVDDCEAALNEQINVEYNASYAYHSLFAYFDRDNVALKGFAKFFKESSDEEREHAEKLMEYQNKRGGRVRLQSIVADLTEFDHPEKGDALY : 173
AtFer3    : HLYSPECEAAVNEQINVEYNVSYVYHALYAYFDRDNVALKGLAKFFKESSVEEREHAELLMEYQNKRGGRVKLQPMVLPQSEFDHPEKGDALY : 181
AtFer4    : QKYSDECEAAINEQINVEYNVSYVYHAMYAYFDRDNIALKGLAKFFKESSLEEREHAEKLMEYQNKRGGRVKLQSIVMPLSEFEHVDKGDALY : 183
AtFer2    : HRFSDDSESAINDQINVEYNVSYVYHALYAYFDRDNVCLKGFAKFFNDSSLEERGHAEMFMEYQNKRGGRVKLQSILMPVSEFDHEEKGDALH : 175
AtFer1    : QRFADASEAVINEQINVEYNVSYVYHSMYAYFDRDNVAMKGLAKFFKESSEEEREHAEKFMSYQNQRGGRVKLHPIVSPISEFEHAEKGDALY : 180
EcabFerH  : QNYHQDSEAAINRQINLELHASYVYLSMSFYFDRDDVALKNFAKYFLHQSHEEREHAEKLMKLQNQRGGRIFLQDIKKP----DQDDWENGLK :  99
HsFerH    : QNYHQDSEAAINRQINLELYASYVYLSMSYYFDRDDVALKNFAKYFLHQSHEEREHAEKLMKLQNQRGGRIFLQDIKKP----DCDDWESGLN :  99

                    Helix A        FO   3FC    *        3FC FO FO    FO    240  Helix B    4FC 4FC 4FC        *
OsFer1    : AMELALALEKLVNEKLHNLHSVASRCNDPQLTDFVESEFLEEQVEAIKKISEYVAQLRRVG---KGHGVWHFDQKLLEEEA---------- : 251
OsFer2    : AMELALALEKLVNEKLHNLHSVASRCNDPQLTDFVESEFLEEQVEAIKKISEYVAQLRRVG---KGHGVWHFDQKLLEEEA---------- : 255
ZmFer1    : AMELALALEKLVNEKLHNLHGVATRCNDPQLTDFIESDFLEEQVEAINKISKYVAQLRRVG---KGHGVWHFDQMLLEEEA---------- : 253
ZmFer2    : AMELTLALEKLVNEKLHSLHGVATRCNDPQLIDFIESEFLEEQVEAINKYSKYVAQLRRVGN--KGHGVWHFDQMLLQEGA---------- : 252
AtFer3    : AMELALSLEKLVNEKLLNLHSVASKNDDVQLADFIESVFLNEQVEAIKKISEYVSQLRRLG---KGHGTWHFDQELLGAAA---------- : 259
AtFer4    : GMELALSLEKLVNEKLLNLHSVASKNNDVHLADFIESEFLTEQVEAIKLISEYVAQLRRVG---KGHGTWHFNQMLLEG------------ : 259
AtFer2    : AMELALSLEKLTNEKLLKLQSVGVKNNDVQLVDFVESEFLGEQVEAIKKISEYVAQLRRIG---KGHGVWHFDQMLLNDEV---------- : 253
AtFer1    : AMELALSLTEKLLNVHKVASENNDVQLVDFVESEFLGEQIEAIKKISDYITQLRRMG---------KGHGVWHFDQMLLN---------- : 255
EcabFerH  : AMECALHLEKNVESLLELHKLATDKNDPHLCDFLETHYLNEQVKAIKELGDHVTNLRRMGAPESGMAEYLFDKHTLGECDES-------- : 182
HsFerH    : AMECALHLEKNVNQSLLELHKLATDKNDPHLCDFIETHYLNEQVKAIKELGDHVTNLRKMGAPESGLAEYLFDKHTWETVIMKAKPRANFP : 190

                    Helix C                    Helix D                    Helix E
```

Figure 7 - ClustalW multiple alignment of the Ferritin proteins. The figure shows the structure of plant Ferritins of *Oryza sativa* (OsFer1: AAM74942; OsFer2: AAM74943), *Zea mays* (ZmFer1: CAA58146; ZmFer2: CAA58147), *Arabidopsis thaliana* (AtFer1: CAA63932; AtFer2: CAC85498; AtFer3: CAC85399; AtFer4: CAC85400) and compares them with the mammalian homologues of *Homo sapiens* heavy chain (HsFerH: CAA72086) and *Equus caballus* heavy chain (EcabFerH: AAM51631). A transit peptide and an extension peptide are delimited in the corresponding boxes. Alpha helixes are positioned below the alignment block according to the extensions from the HsFerH. Some functional aminoacids are pointed above the alignment block in the positions that they occur in the HsFerH. In this protein they are involved in the Ferroxidase catalytic center (FO), constitution of the 3-fold channel (3FC) and the 4-fold channel (4FC).

mineral and the occurrence of various paralogs within a family of transporters?

Initially, distinct iron transporter families were explained on the basis of the different Fe uptake strategies (Guerinot and Yi, 1994). In strategy I, found in dicots and non-graminaceous monocots, after the solubilization and reduction steps, Fe^{2+} would be taken up by a specific transporter, later on characterized as IRT1 (Eide *et al.*, 1996). In contrast, grasses would take up Fe^{3+} in the form of complexes with phytosiderophores by another class of transporters, later defined as Yellow Stripe1 (Curie *et al.*, 2001). With the complete genomic sequence of *Arabidopsis*, it was found that eight homologs of the grass YS transporter were present in this dicot. In the present study, it was also observed that there are IRT homologs in the rice genome. Therefore, the notion that a specific family of transporters is related to a certain physiological strategy for iron uptake, although real, seems not to be exclusive and must be revised.

The presence of YS family members in *Arabidopsis*, a species that does not produce phytosiderophores, has been explained on the basis of nicotianamine (NA) transport (Curie *et al.*, 2001). This compound is ubiquitous in plant tissues and is a structural analog of the derivatives of mugineic acid (MA), the phytosiderophores. It has been proposed that NA plays a fundamental role in iron homeostasis, as a transport form for minerals in the phloem (Stephan and Scholtz, 1993). The corn YS1 transporter is capable of transporting Fe^{2+}-nicotianamine complexes based on yeast functional complementation assays (Walker *et al.*, 2001). Immunodetection studies showed an increase of NA-chelated iron in tomato tissues exposed to excess concentration of the metal (Pich *et al.*, 2001). Therefore, the YS transporters in dicots could play a role in the transport of NA-complexed minerals, not being involved in uptake, whereas the ZIP family of transporters would be responsible for Fe^{2+} uptake in roots, with a possible additional function in micronutrient distribution throughout the plant.

What about the rationale for the presence of IRT orthologs in grasses such as rice? The mechanism based on

the release of phytosiderophores for Fe^{+3} chelation and uptake under mineral deficiency represents an evolutionary advantage relative to the mechanism of strategy I, since the latter requires an energy dependent reduction step prior to the uptake (Grusak *et al.*, 1999). Presumably, grasses acquired the strategy II components in addition to those of strategy I, which progressively assumed a more secondary role in mineral uptake and transport. The larger number of YS orthologs in rice (18) compared to ZIP orthologs (13) supports the greater importance of YS in mineral uptake and transport. The opposite scenario is seen in *Arabidopsis*, in which there are fifteen ZIP members and only eight YS orthologs, in agreement with a predominant role for the ZIP family in this species. The recently cloned *OsIRT1* in rice, expressed in roots under mineral deficiency and capable of rescuing iron deficient yeast mutants, has been proposed as an actively engaged iron uptake system, especially under limiting conditions (Bughio *et al.* 2002). The functional characterization of *OsIrt1* is evidence that rice also possesses an Fe^{2+} uptake system that could be involved in ferrous ion absorption under waterlogged conditions, in which this form of iron is prevalent. This result suggests that grasses may not be restricted to a single strategy for iron uptake, but can adjust the system according to environmental requirements (Bughio *et al.*, 2002).

The NRAMP family members coexist with YS and ZIP families in *Arabidopsis* and, as seen in this study, in rice. Except for a suggested role for AtNRAMP3 (Thomine *et al.*, 2000), the NRAMP proteins do not seem to be involved in iron uptake from the soil, but probably function in the subcellular transport of the mineral. Transgenic *Arabidopsis* lines overexpressing *NRAMP1* showed higher resistance to toxic concentrations of iron than wild type plants, proportionately to the transgene expression. Resistance could be caused by iron relocation, promoted by the NRAMP1 protein, from the cytosol to the vacuole or the plastids (Curie *et al.*, 2000).

Another reason for the presence of three different families of iron transporters in plants could be differential affinity for iron in different oxidation states. The NRAMP family is known as a transporter for divalent cations (Williams *et al.*, 2000). Similarly, IRT1 has affinity for Fe^{2+} (Eide *et al.*, 1996), whereas only the YS transporter would be capable of ferrying the cation in its ferric form (Curie *et al.*, 2001). The availability of various transporters may fulfill the needs for iron transport in different redox environments or under fluctuating redox conditions.

Different iron carrier molecules may also explain the diversity of iron transporter families in plants. The YS1 protein, for example, transports Fe^{3+} phytosiderophores, and possibly Fe^{+2} NA (Walker *et al.*, 2001), structurally similar to phytosiderophores. Another important mineral carrier in plants is citrate, mainly associated with iron transport in the xylem (Grusak *et al.*, 1999). Since members of the YS family were unable to transport citrate-complexed
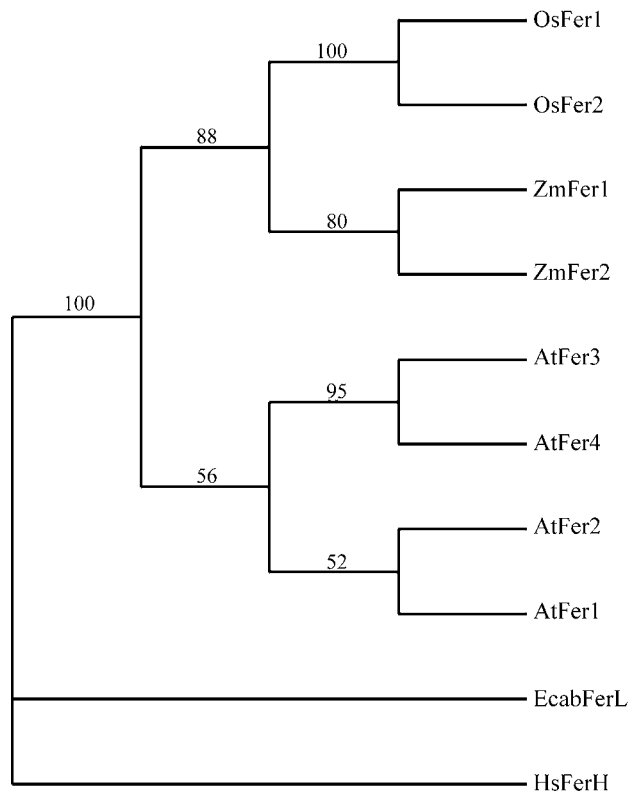

Figure 8 - Phylogenetic relationships of *Oryza sativa* Ferritins to plant and mammalian Ferritins. The proteins are from *Oryza sativa* (OsFer1: AAM74942; OsFer2: AAM74943), *Zea mays* (ZmFer1: CAA58146; ZmFer2: CAA58147), *Arabidopsis thaliana* (AtFer1: CAA63932; AtFer2: CAC85498; AtFer3: CAC85399; AtFer4: CAC85400) and the mammalian homologues *Homo sapiens* heavy chain (HsFerH: CAA25086) and *Equus caballus* heavy chain (EcabFerH: AAM51631).

minerals (Curie *et al.*, 2001), it may be possible that NRAMP or ZIP members could transport such complexes.

The variety of paralogs found within each family of transporters could be explained by at least three reasons: 1) paralogs with different affinities for different minerals or distinct affinities for the same mineral; 2) different paralogs operating in the various organs and tissues; 3) different paralogs for each subcellular compartment, such as vacuoles or plastids. Yeast complementation assays showed that plant mineral transporters are able to transport several minerals, often displaying preferences for the transport of one or more of these.

The IRT proteins are mostly involved in iron transport. Although with lower affinity, *Arabidopsis* IRT1 was also capable of transporting Zn^{2+}, Mn^{2+} and Cd^{2+}, as shown by radioisotope and functional complementation assays (Eide *et al.*, 1996; Korshunova *et al.*, 1999; Rogers *et al.*, 2000). Similarly, *Arabidopsis* IRT2 seems to be specific for iron and Zn^{2+} transport, not showing affinity for other cations (Vert *et al.*, 2001). Tomato IRT1 and 2 may also be capable of transporting zinc, copper and manganese (Eckhardt *et al.*, 2001). On the other hand, ZIP1 through 4 show specificity for zinc transport, but not for iron. Zinc

uptake by *ZIP2*-complemented yeast was strongly inhibited by Cd^{2+} and Cu^{2+} (Grotz *et al.*, 1998). It is quite possible that similar features are found in the rice ZIP family members. The plant NRAMP family members have been characterized for the transport of iron and cadmium. Expression of *AtNramp1, 3* and *4* resulted in increased cadmium sensitivity and cadmium content in yeast, and the same genes can also complement a yeast mutant defective in manganese uptake (Thomine *et al.*, 2000). Moreover, mammalian and yeast orthologs have a wide spectrum of mineral transport (Williams *et al.*, 2000). This may lead to the hypothesis that the eight rice NRAMP members could be partly explained by diverse affinities of the individual members for different minerals. The important role of NA in mineral homeostasis may be one of the reasons for the large number of YS family members in the rice genome. NA is well known for its capacity to chelate a variety of minerals (Benes *et al.*, 1983), being involved in the transport of copper in the phloem and possibly performing the same function for zinc, iron and manganese (Stephan and Scholtz *et al.*, 1993). In this regard, it is reasonable to suggest that different YS paralogs may coordinate the transport of various NA-metal complexes. The release of phytosiderophores is not only stimulated by iron deficiency, but by zinc as well (Tolay *et al.*, 2001), indicating that this mineral may also be transported in complexed form with phytosiderophores, perhaps by a YS paralog.

Different paralogs may also have different affinities for the same cation. In yeast, the ZIP orthologs constitute two systems with different affinities for zinc. ZRT1 has high affinity for the metal, being active when zinc is limiting in the cell, whereas ZRT2 shows low affinity, transporting the mineral when it is abundant in the environment (Guerinot, 2000). Yeast complementation assays indicated that *Arabidopsis* ZIP1, 2, and 3 have different zinc uptake activities, with Km values ranging from 2 µM in ZIP2 to 14 µM zinc in ZIP3 (Grotz *et al.*, 1998). Such different affinities for the mineral may also occur in the rice ZIP, NRAMP, and YS families, constituting a possible explanation for the diversity of paralogs observed in this species.

The mineral flow throughout the plant requires the transport across barriers that delineate tissue compartments of long distance transport, such as xylem and phloem, and cell compartments in other plant tissues. Therefore, different transporters could be associated with specific tissues or cell types. Northern blot experiments have provided strong evidence that paralog diversity is related to their expression in different organs and tissues. In rice, *NRAMP1* is expressed in roots, whereas *NRAMP2* is expressed in leaves and *NRAMP*3 in both organs (Belouchi *et al.*, 1997). In *Arabidopsis*, organ-specific expression is also observed. *NRAMP1* and 2 are preferentially expressed in roots, whereas *NRAMP3* and *4* are equally expressed in roots and shoots. Moreover, although *NRAMP3* expression occurs in shoots under iron availability, expression under iron defi-

ciency is only induced in roots for this paralog (Thomine *et al.*, 2000). *Arabidopsis IRT1* and *IRT2* are exclusively expressed in roots (Eide *et al.*, 1996; Vert *et al.*, 2001), the same happening with *ZIP1* and *ZIP3* of the same family. *ZIP4* is expressed under iron deficiency in shoots and roots (Grotz *et al.*, 1998). The EST patterns found in this study for the rice YS family are in good agreement with these data on organ-specific expression. The ESTs that make up the *OsYS4* pattern were obtained from cDNA libraries derived from flowering panicules, whereas ESTs related to the *OsYS6* paralog originated in cDNA libraries built from roots of seedlings.

Besides the need for organ or tissue specific transporters, it is possible that different isoforms of the iron reducing enzymes are also required. Under limited iron availability, AtFROHC and PsFRO1 were shown to be expressed in leaves, evidence that FRO members may act in a reducing step before iron moving through the xylem is absorbed by leaves.

Some paralogs observed in the various transporter families may also be related to the subcellular transport of iron and other minerals. The varying patterns of expression of different *Nramp* genes in plants are consistent with possible localization in distinct membranes in the plant cell, such as the plasma membrane, the tonoplast and the chloroplast envelope. AtNRAMP1 could be involved in the sequestration of free cytosolic iron into the plastids or the vacuole (Curie *et al.*, 2000). The subcellular localization of each one of the isoforms will help to clarify this point. A role in internal iron transport to a subcellular location has also been proposed for IRT2, since its constitutive expression could not suppress the defect in iron uptake from the *Arabidopsis Irt1* T-DNA mutant (Vert *et al.*, 2002). On the basis of a predicted transit peptide, it is also believed that *Arabidopsis ZIP4* may function as a zinc transporter to the plastid. The subcellular localization signal predictions conducted in the present study pointed toward some evidence of plastid localization of OsYS12 and OsFer2 encoded in chromosome 12, mitochondrial localization of OsZIP1 and OsNRAMP2, as well as involvement in secretory pathways for OsNRAMP5. These and other proteins may be responsible for coordinating influx and possible efflux of minerals in subcellular compartments.

Perspectives

The novel putative components of iron homeostasis in rice reported in this study open new perspectives in the study of iron and other mineral nutrients uptake and distribution in this important crop species. More refined studies, involving complementation analysis in yeast strains deficient in iron acquisition, gene expression analysis, immunodetection and protein level analyses, may define the individual role that each of these new transporters plays in mineral homeostasis, as well as provide the basis for efficient biotechnological manipulation of iron content in rice

plants. We are currently investigating the expression patterns of several of these new genes in rice plants exposed to varying levels of iron. We are especially interested in comparing expression between rice varieties considered resistant and susceptible to iron toxicity under waterlogged conditions.

Acknowledgments

References

Altschul SF, Madden TL, Schaffer 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.

Bar-Ness E, Hadar Y, Chen Y, Romheld V and Marschner H (1992) Short-term effects of rhizosphere microorganisms on Fe uptake from microbial siderophores by maize and oat. Plant Physiol 100:451-456.

Belouchi A, Kwan T and Gros P (1997) Cloning and characterization of the OsNramp family from *Oryza sativa*, a new family of membrane proteins possibly implicated in the transport of metal ions. Plant Mol Biol 33:1085-1092.

Benes I, Schreiber K, Ripperger H and Kircheiss A (1983) Metal complex formation by nicotianamine, a possible phytosiderophore. Experimia 39:261-262.

Briat JF, Fobis-Loisy I, Grignon N, Lobréaux S, Pascal N, Savino G, Thoiron S, von Wirén N and van Wuytswinkel O (1995) Cellular and molecular aspects of iron metabolism in plants. Biol Cell 84:69-81.

Briat JF and Lobréaux S (1997) Iron transport and storage in plants. Trends Plant Sci 2:187-193.

Briat JF and Lobréaux S (1998) Iron storage and ferritin in plants. In: Sigel A and Sigel H (eds) Metal Ions in Biological Systems: Iron Transport and Storage in Microorganisms, Plants and Animals, Marcel Dekker Inc, New York, pp 563-584.

Bughio N, Yamaguchi H, Nishizawa NK, Nakanishi H and Mori S (2002) Cloning an iron-regulated metal transporter from rice. J Exp Bot 53:1677-1682.

Burge C and Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78-94.

Burr B (2002) Mapping and Sequencing the Rice Genome. Plant Cell 14:521-523.

Cellier M, Privé G, Belouchi A, Kwan T, Rodriques V, Chia W and Gros P (1995) Nramp defines a family of membrane proteins. Proc Natl Acad Sci USA 92:10089-10093.

Curie C, Alonso JM, Le Jean M, Ecker JR and Briat JF (2000) Involvement of NRAMP1 from *Arabidopsis thaliana* in iron transport. Biochem J 347:749-755.

Curie C, Panaviene Z, Loulergue C, Dellaporta SL, Briat JF and Walker EL (2001) Maize yellow stripe1 encodes a membrane protein directly involved in Fe(III) uptake. Nature 409:346-349.

Chaney RL, Brown, JC and Tiffin LO (1972) Obligatory reduction of ferric chelates in iron uptake by soybean. Plant Physiol 50:208-213.

Connolly EL, Fett JP and Guerinot ML (2002) Expression of the IRT1 metal transporter is controlled by metals at the levels of transcript and protein accumulation. Plant Cell 14:1347-1357.

Dancis A, Roman DG, Anderson GJ, Hinnebusch AG and Klausner RD (1992) Ferric reductase of *Saccharomyces cerevisiae*: molecular characterization, role in iron uptake, and transcriptional control by iron. Proc Natl Acad Sci USA 89:3869-3873.

Delseny M, Salses J, Cooke R, Sallaud C, Regad F, Lagoda P, Guiderdoni E, Ventelon M, Brugidou C and Ghesquière A (2001) Rice genomics: present and future. Plant Physiol Biochem 39:323-334.

Eckardt NA (2000) Sequencing the rice genome. Plant Cell 12:2011-2017.

Eckhardt U, Marques AM and Buckhout TJ (2001) Two iron-regulated cation transporters from tomato complement metal uptake-deficient yeast mutants. Plant Mol Biol 45:437-448.

Eide D, Broderius M, Fett JP and Guerinot ML (1996) A novel iron-regulated metal transporter from plants identified by functional expression in yeast. Proc Natl Acad Sci USA 93:5624-5628.

Fobis-Loisy I, Loridon K, Lobreaux S, Lebrun M and Briat JF (1995) Structure and differential expression of two maize ferritin genes in response to iron and abscisic acid. Eur J Biochem 231:609-619.

Fobis-Loisy I, Aussel L and Briat J-F (1996) Post-transcriptional regulation of plant ferritin accumulation in response to iron as observed in the maize mutant *ys1*. FEBS Letters 397:149-154.

Fox TC, Shaff JE, Grusak MA, Norvell WA, Chen Y, Chaney RL and Kochian LV (1996) Direct measurement of ^{59}Fe-labeled Fe^{2+} influx in roots of pea using a chelator buffer system to control free Fe^{2+} in solution. Plant Physiol 111:93-100.

Goff SA, Ricke D, Lan T-H, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun W, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A and Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92-100.

Groom QJ, Torres MA, Fordham-Skelton AP, Hammond-Kosack KE, Robinson NJ and Jones JD (1996) rbohA, a rice homologue of the mammalian gp91phox respiratory burst oxidase gene. Plant J 10:515-522.

Grotz N, Fox T, Connolly E, Park W, Guerinot ML 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.

Grotz N and Guerinot ML (2002) Limiting nutrients: an old problem with new solutions? Curr Opin Plant Biol 5:158-163.

Grusak MA, Pearson JN and Marentes E (1999) The physiology of micronutrient homeostasis in field crops. Field Crops Res 60:41-56.

Grusak, MA (2002) Enhancing mineral content in plant food products. J Am Coll Nutr 21:178S-183S.

Guerinot ML and Yi Y (1994) Iron: nutritious, noxious, and not readily available. Plant Physiol 104:815-820.

Guerinot ML (2000) The ZIP family of metal transporters. Biochim Biophys Acta 1465(1-2):190-198.

Guerinot ML (2001) Improving rice yields-ironing out the details. Nat Biotechnol 19:417-418.

Harrison PM, Hempstead PD, Artymiuk PJ and Andrews SC (1998) Structure-function relationships in the ferritins. In: Sigel A and Sigel H (eds) Metal Ions in Biological Systems: Iron Transport and Storage in Microorganisms, Plants and Animals, Marcel Dekker Inc, New York, pp 435-477.

Hauser M, Donhardt AM, Barnes D, Naider F and Becker JM (2000) Enkephalins are transported by a novel eukaryotic peptide uptake system. J Biol Chem 275:3037-3041.

Hauser M, Narita V, Donhardt AM, Naider F and Becker JM (2001) Multiplicity and regulation of genes encoding peptide transporters in *Saccharomyces cerevisiae*. Mol Membr Biol 18:105-112.

Hirokawa T, Boon-Chieng S and Mitaku S (1998) SOSUI: Classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378.

Koh S, Wiles AM, Sharp JS, Naider FR, Becker JM and Stacey G (2002) An oligopeptide transporter gene family in *Arabidopsis*. Plant Physiol 128:21-29.

Korshunova YO, Eide D, Clark WG, Guerinot ML and Pakrasi HB (1999) The IRT1 protein from *Arabidopsis thaliana* is a metal transporter with a broad substrate range. Plant Mol Biol 40:37-44.

Lubkowitz MA, Hauser L, Breslav M, Naider F and Becker JM (1997) An oligopeptide transport gene from *Candida albicans*. Microbiology 143:387-396.

Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY and Bryant SH (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 30:281-283.

Marschner H and Römheld V (1994) Strategies of plants for acquisition of iron. Plant Soil 165:261-274.

Mäser P, Thomine S, Schroeder JI, Ward JM, Hirschi K, Sze H, Talke IN, Amtmann A, Maathuis FJ, Sanders D, Harper JF, Tchieu J, Gribskov M, Persans MW, Salt DE, Kim SA and Guerinot ML (2001) Phylogenetic relationships within cation transporter families of *Arabidopsis*. Plant Physiol 126:1646-1667.

Masuda T, Goto F and Yoshihara T (2001) A novel plant ferritin subunit from soybean that is related to a mechanism in iron release. J Biol Chem 276:19575-19579.

Moog PR and Brüggermann W (1994) Iron reductase systems on the plant plasma membrane- a review. Plant Soil 165:241-260.

Morgenstern B (1999) DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics 15:211-218.

Moreau S, Thomson RM, Kaiser BN, Trevaskis B, Guerinot ML, Udvardi MK, Puppo A and Day DA (2002) GmZIP1 encodes a symbiosis-specific zinc transporter in soybean. J Biol Chem 277:4738-4746.

Murray-Kolb LE, Takaiwa F, Goto F, Yoshihara T, Theil EC and Beard JL (2002) Transgenic rice is a source of iron for iron-depleted rats. J Nutr 132:957-960.

Orengo CA, Jones DT and Thornton JM (1994) Protein superfamilies and domain superfolds. Nature 372:631-634.

Petit J-M, van Wuytswinkel O, Briat J-F and Lobréaux S (2001a) Characterization of an iron-dependent regulatory sequence involved in the transcriptional control of *AtFer1* and *ZmFer1* plant ferritin genes by iron. J Biol Chem 276:5584-5590.

Petit J-M, Briat J-F and Lobreaux S (2001b) Structure and differential expression of the four members of the *Arabidopsis thaliana* ferritin gene family. Biochem J 359:575-582.

Pich A, Manteuffel R, Hillmer S, Scholz G and Schmidt W (2001) Fe homeostasis in plant cells: Does nicotianamine play multiple roles in the regulation of cytoplasmic Fe concentration? Planta 213:967-976.

Ragland M, Briat JF, Gagnon J, Laulhere JP, Massenet O and Theil EC (1990) Evidence for conservation of ferritin sequences among plants and animals and for a transit peptide in soybean. J Biol Chem 265:18339-18344.

Robinson NJ, Sadjuga and Groom QJ (1997) The froh gene family from *Arabidopsis thaliana*: Putative iron-chelate reductases. Plant and Soil 196:245-248.

Robinson NJ, Procter CM, Connolly EL and Guerinot ML (1999) A ferric-chelate reductase for iron uptake from soils. Nature 397:694-697.

Rogers EE, Eide DJ and Guerinot ML (2000) Altered selectivity in an *Arabidopsis* metal transporter. Proc Natl Acad Sci USA 97:12356-12360.

Römheld V and Marschner H (1983) Mechanism of iron uptake by peanut plants. I- Fe(III) reduction, chelate splitting, and release of phenolics. Plant Physiol 71:949-954.

Römheld V and Marschner H (1986) Evidence for a specific uptake system for iron phytosiderophores in roots of grasses. Plant Physiol 80:175-180.

Römheld V, Müller C and Marschner H (1984) Localization and capacity of proton pumps in roots of intact sunflower plants. Plant Physiol 76:603-606.

Rost B and Valencia A (1996) Pitfalls of protein sequence analysis. Curr Opin Biotechnol 7:457-461.

Rost B (1999) Twilight zone of protein sequence alignments. Protein Eng 12:85-94.

Salamov AA and Solovyev VV (2000) Ab initio gene finding in *Drosophila* genomic DNA. Genome Res 10:516-522.

Shimamoto K and Kyozuka J (2001) Rice as a model for comparative genomics of plants. Annu Rev Plant Physiol Plant Mol Biol 53:399-419.

Silverstein KAT, Shoop E, Johnson JE, Kilian A, Freeman JL, Kunau TM, Awad IA, Mayer M and Retzel EF (2001) The MetaFam Server: a comprehensive protein family resource. Nucleic Acids Res 29:49-51.

Stephan UW and Scholtz G (1993) Nicotianamine: mediator of transport of iron and heavy metals in the phloem? Physiol Plant 88:522-529.

Sugiura Y amd Nomoto K (1984) Phytosiderophores: structures and properties of mugineic acids and their metal complexes. Struct Bonding 58:107-135.

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, MA.

Takahashi M, Nakanishi H, Kawasaki S, Nishizawa NK and Mori S (2001) Enhanced tolerance of rice to low iron availability in alkaline soils using barley nicotianamine aminotransferase genes. Nat Biotechnol 19:466-469.

Thomine S, Wang R, Ward JM, Crawford NM and Schroeder JI (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.

Thompson JD, Higgins DG and Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680.

Todd AE, Orengo CA and Thornton JM (2001) Evolution of function in protein superfamilies, from a structural perspective. J Mol Biol 307:1113-1143.

Tolay I, Erenoglu B, Römheld V, Braun HJ and Cakmak I (2001) Phytosiderophore release in *Aegilops tauschii* and *Triticum* species under zinc and iron deficiencies. J Exp Bot 52:1093-1099.

Torres MA, Onouchi H, Hamada S, Machida C, Hammond-Kosack KE and Jones JD (1998) Six *Arabidopsis thaliana* homologues of the human respiratory burst oxidase (gp91phox). Plant J 14:365-370.

Vansuyt G, Mench M and Briat JF (2000) Soil-dependent variability of leaf iron accumulation in transgenic tobacco overexpressing ferritin. Plant Physiol Biochem 38:499-506.

Vert G, Briat JF and Curie C (2001) *Arabidopsis* IRT2 gene encodes a root-periphery iron transporter. Plant J 26:181-189.

Vert G, Grotz N, Dedaldechamp F, Gaymard F, Guerinot ML, Briat JF and Curie C (2002) IRT1, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. Plant Cell 14:1223-1233.

Walker EL, Burke T, Panaviene Z and Abbey PJ (2001) Possible roles for the Yellow Stripe-Like (YSL) family of proteins in the transport of metal-nicotianamine complexes. American Society of Plant Biologists Annual Meeting Abstracts. Mini-symposium 1 - Membrane Transport. Abstract 11004 (also available at http://www.rycomusa.com/aspp2001/public/M01/0986.html).

Waters BM, Blevins DG and Eide DJ (2002) Characterization of FRO1, a pea ferric-chelate reductase involved in root iron acquisition. Plant Physiol 129:85-94.

Waterston RH, Lander ES and Sulston JE (2002) On the sequencing of the human genome. Proc Natl Acad Sci USA 99:3712-3716.

Williams LE, Pittman JK and Hall JL (2000) Emerging mechanisms for heavy metal transport in plants. Biochim Biophys Acta 1465:104-126.

von Wirén N, Marschner H and Römheld V (1995) Uptake kinetics of iron-phytosiderophores in two maize genotypes differing in iron efficiency. Plant Physiol 93:611-616.

World Health Organization (2002) Nutrition. Reduction of iron deficiency anaemia. http://www.who.int/nut/ida.htm, 09/12/2002.

Xu Y and Uberbacher EC (1997) Automated gene identification in large-scale genomic sequences. J Comput Biol 4:325-338.

Yen M-R, Tseng YH and Saier Jr MH (2001) Maize Yellow Stripe1, an iron-phytosiderophore uptake transporter, is a member of the oligopeptide transporter (OPT) family. Microbiology 147:2881-2883.

Yu J, Hu S, Wang J, Wong G K-S, Li S, Liu B, Deng Y, Daí L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Jian Li, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L and Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79-92.

Zhao H and Eide D (1996a) The yeast ZRT1 gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation. Proc Natl Acad Sci USA 93:2454-2458.

Zhao H and Eide D (1996b) The ZRT2 gene encodes the low affinity zinc transporter in *Saccharomyces cerevisiae*. J Biol Chem 271:23203-23210.

Editor: Márcio de Castro Silva Filho