Genomic Comparison of P-Type ATPase Ion Pumps in Arabidopsis and Rice[1]

Ivan Baxter, Jason Tchieu, Michael R. Sussman, Marc Boutry, Michael G. Palmgren, Michael Gribskov, Jeffrey F. Harper*, and Kristian B. Axelsen

Department of Cell Biology, Plant Division, The Scripps Research Institute, La Jolla, California 92037 (I.B., J.F.H.); Department of Biology, University of California, San Diego, La Jolla, California 92093 (J.T., M.G.); Biotechnology Center, University of Wisconsin, Madison, Wisconsin 53706 (M.R.S.); Unite de Biochemie Physiologique, Institut des Sciences de la Vie, Universite Catholique de Louvain, Louvain-La-Neuve, Belgium (M.B.); Plant Physiology and Anatomy Laboratory, Department of Plant Biology, The Royal Veterinary and Agricultural University, Copenhagen, Denmark (M.G.P., K.B.A.); and SWISS-PROT group, Swiss Institute of Bioinformatics, Geneva, Switzerland (K.B.A.)

Members of the P-type ATPase ion pump superfamily are found in all three branches of life. Forty-six P-type ATPase genes were identified in Arabidopsis, the largest number yet identified in any organism. The recent completion of two draft sequences of the rice (*Oryza sativa*) genome allows for comparison of the full complement of P-type ATPases in two different plant species. Here, we identify a similar number (43) in rice, despite the rice genome being more than three times the size of Arabidopsis. The similarly large families suggest that both dicots and monocots have evolved with a large preexisting repertoire of P-type ATPases. Both Arabidopsis and rice have representative members in all five major subfamilies of P-type ATPases: heavy-metal ATPases (P_{1B}), Ca^{2+}-ATPases (endoplasmic reticulum-type Ca^{2+}-ATPase and autoinhibited Ca^{2+}-ATPase, P_{2A} and P_{2B}), H^+-ATPases (autoinhibited H^+-ATPase, P_{3A}), putative aminophospholipid ATPases (ALA, P_4), and a branch with unknown specificity (P_5). The close pairing of similar isoforms in rice and Arabidopsis suggests potential orthologous relationships for all 43 rice P-type ATPases. A phylogenetic comparison of protein sequences and intron positions indicates that the common angiosperm ancestor had at least 23 P-type ATPases. Although little is known about unique and common features of related pumps, clear differences between some members of the calcium pumps indicate that evolutionarily conserved clusters may distinguish pumps with either different subcellular locations or biochemical functions.

Ion pumps belonging to the P-type ATPases superfamily share a common enzymatic mechanism in which ATP hydrolysis is used to transport ions across a membrane. The name "P-type" comes from a phosphorylated intermediate that is characteristic of the enzyme's catalytic cycle (Pedersen and Carafoli, 1987). Two structures of 3.1 and 2.6 Å resolution of a calcium pump (sarcoplasmic/endoplasmic reticulum [ER] calcium ATPase [SERCA]) have greatly increased the understanding of the mechanism of ion transport by P-type ATPases (Toyoshima et al., 2000; Toyoshima and Nomura, 2002). These structures provide the foundation for understanding the mechanism of all P-type ATPases. For example, these structures have been used to model P-type proton (Bukrinsky et al., 2001; Kühlbrandt et al., 2002; Radresa et al., 2002) and sodium (Ogawa and Toyoshima, 2002) pumps.

P-type ATPases are found in all three branches of life and are used to translocate a diverse set of ions, including H^+, Na^+/K^+, H^+/K^+, Ca^{2+}, heavy metals, and possibly lipids (Axelsen and Palmgren, 1998). The superfamily has been divided into five major branches (with 10 sub-branches) that reveal a correlation between sequence identity and ion specificity. For example, all the heavy-metal pumps from bacteria, plants, and humans share significant sequence similarities and cluster together as the P_{1B} subfamily (Palmgren and Axelsen, 1998). Thus, the ion specificities characteristic of the different branches of P-type ATPases appear to have evolved very early.

The most P-type ATPases yet identified in a single organism have been found in the completed genome sequence of Arabidopsis. Following the 45 pumps identified in the initial genome release (Axelsen and Palmgren, 2001), an additional P-type ATPase was identified in further releases (P-type ATPase database; http://biobase.dk/~axe/Patbase.html) giving a total of 46 P-type ATPases in this model plant. In comparison, Brewer's yeast (*Saccharomyces cerevisiae*) has only 16 (Catty et al., 1997), fission yeast (*Schizosaccharomyces pombe*) has 14, *Caenhorabditis elegans* has

[1] Work is supported in part by the Department of Energy (grant no. DOE DE–FG03–94ER20152 to J.F.H.), Syngenta (grant to J.F.H.), by the National Science Foundation (grant no. DBI 0077378 to J.F.H. and M.G.), and by Human Frontiers Science Program (grant no. RG0268 to J.F.H., M.R.S., M.B., and M.G.P.).

* Corresponding author; e-mail harper@scripps.edu; fax 858–784–9840.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.103.021923.

21, fruitfly (*Drosophila melanogaster*) has 14, *Synechocystis* sp. PCC6803 has 9, and human (*Homo sapiens*) has 36 (P-type ATPase database; http://biobase.dk/~axe/Patbase.html).

The apparent expansion in numbers of P-type ATPases in Arabidopsis has at least three potential explanations. First, multiple isoforms may have evolved different modes of regulation and activity to provide specific cell types with unique biochemical properties. Second, multiple isoforms may have evolved under the control of different tissue-specific promoters to provide the appropriate levels of expression in specific tissues. Third, some of the multiple isoforms may represent functionally redundant duplications.

The recent completion of two shotgun sequences of the rice (*Oryza sativa*) genome (Goff et al., 2002; Yu et al., 2002) provides the first opportunity to compare the full complement of P-type ATPases from a monocot with that of a dicot, Arabidopsis. Here, we report the identification of 43 members of the P-type ATPase superfamily in rice. The observation of similarly large families in both rice and Arabidopsis suggests that both dicots and monocots have evolved with a large preexisting repertoire of P-type ATPases. Although the unique and redundant functions of these multiple pumps are largely unknown, comparison of the full complement of genes from two different plants allows for the grouping of similar pumps. Our genomic comparison of sequence identities and intron positions suggests that a common angiosperm ancestor harbored a core set of at least 23 P-type ATPases, providing the archetypal representatives for 23 clusters of P-type ATPases currently present in rice and Arabidopsis.

RESULTS AND DISCUSSION

Forty-Three Rice P-Type ATPase Genes

Our analysis identified 43 members of the P-type ATPase family in rice (Fig. 1; Supplemental Table A; supplementary data can be viewed at www.plantphysiol.org). This is three fewer than found in Arabidopsis. To ensure that all possible P-type ATPase genes were identified, multiple independent sequences were searched in several different ways. We searched both the Beijing Genomics Institute (Yu et al., 2002) and Syngenta (Goff et al., 2002) rice genome sequences, in addition to the partial sequences deposited at The Institute for Genome Research (by December 16, 2002). We searched each data set for either (a) sequences with overall sequence identity to known plant P-type ATPases or (b) sequences that matched one of five DKTGTXX motifs (where XX is LT, IT, VT, MT, or II), a sequence that is found in every known P-type ATPase (Axelsen and Palmgren, 1998). Because one additional Arabidopsis isoform was identified in subsequent releases of the "complete" Arabidopsis genome, it is possible that additional rice isoforms may be identified in future versions of the rice genome sequence.

The Expansion of the Rice Genome Did Not Increase the Number of P-Type ATPases

The rice genome is approximately 3.5 times the size of Arabidopsis (430 Mb compared with 120 Mb; Goff et al., 2002; Yu et al., 2002). This size difference has been attributed to (a) a 1.25- to 2.5-fold increase in the number of genes (32,000–55,000 in rice versus 26,000 in Arabidopsis), (b) an increase in intergenic distances, and (c) an increase in intron size and number. Of these variables, the only difference we confirmed for the rice P-type ATPase genes was a 2-fold increase in the average size of introns (356 to 164 bp).

In the context of a possible 2.5-fold increase in the number of rice genes, it is noteworthy that the number of P-type ATPases did not show a parallel expansion. This same observation has been noted for 15 other gene families (Banuelos et al., 2002; Goff et al., 2002; Baumberger et al., 2003; Jasinski et al., 2003), raising a question about the function of the additional 6,000 to 29,000 rice-specific genes. At present, the expectation is that much of the increased number of genes is due to "over-predictions," because many of the novel rice genes are not supported by expressed sequence tag (EST) data and do not show similarity to known proteins in any organisms, including corn (*Zea mays*; Yu et al., 2002). Because of uncertainties in the accurate prediction of rice gene models, we did not attempt to evaluate the intragenic distances surrounding the rice P-type ATPases.

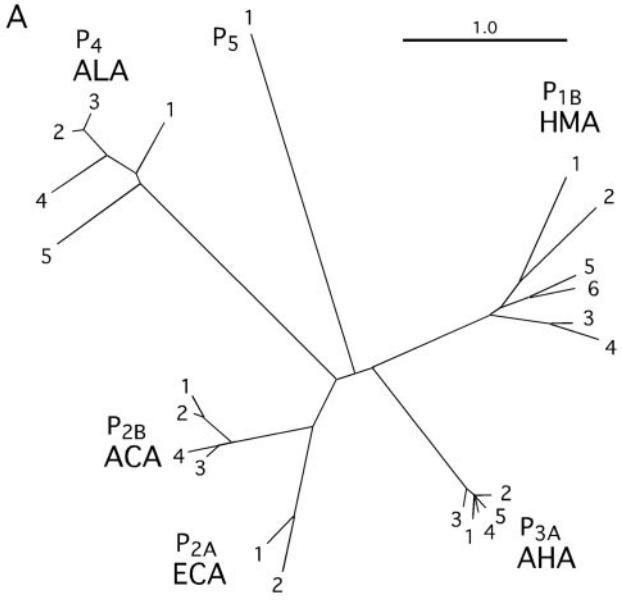
In contrast to a general expectation of more introns in rice genes compared with Arabidopsis (Yu et al., 2002), we observed the same average number (13) and highly similar positions within potential orthologs (Figs. 2–7). Thus, the approximate number and organization of P-type ATPases appear to have been maintained during the evolution of rice and Arabidopsis.

Twenty-Three Clusters of the P-Type ATPases in the Monocot-Dicot Ancestor

Of the 10 subfamilies of P-type ATPases identified to date, six are present in both rice and Arabidopsis: P_{1B} (heavy-metal ATPases [HMA]), P_{2A} (endoplasmic reticulum [ER]-type calcium ATPase [ECA]), P_{2B} (autoinhibited calcium ATPase [ACA]), P_{3A} (autoinhibited H^+ ATPase [AHA]), P_4 (putative aminophospholipid ATPase [ALA]), and P_5 (P-type ATPase, type 5, P5; Fig. 1). Absent from both organisms are the P_{1A} (K^+), P_{2C} (Na^+/K^+), P_{2D} (Na^+ or Ca^{2+}), and P_{3B} (Mg^{2+}) subfamilies.

We further divided the six subfamilies of plant P-type ATPases into 23 clusters, based on protein sequence similiarities, with members of each cluster present in each organism (Figs. 2–7). For ECAs (P_{2A})

A



B

Branch	Topology	Substrate	Cluster	# of members *O.s.*	# of members *A.t.*	Subcellular localization
P_{1B} HMA		Zn/Co/ Cd/Pb?	1	1	1	?
		Zn/Co/ Cd/Pb?	2	2	3	PM
		Cu/Ag?	3	2	1	?
			4	1	1	?
			5	1	1	?
			6	1	1	?
P_{2A} ECA		Ca/Mn	1	2	3	ER
			2	1	1	ER?
P_{2B} ACA		Ca	1	3	3	ER?
			2	4	2	Vac
			3	1	2	?
			4	3	3	PM
P_{3A} AHA		H	1	3	2	PM
			2	2	4	PM
			3	1	1	PM?
			4	3	3	PM?
			5	1	1	PM?
P_4 ALA		PL? Zn/Co?	1	3	1	?
			2	3	5	?
			3	1	4	?
			4	1	1	?
			5	2	1	?
P_5 P5		Ca ?	1	1	1	?
		total	23	43	46	

Figure 1. A, Phylogenetic tree showing 23 clusters of Arabidopsis and rice P-type ATPases. The tree was constructed by aligning full-length sequences with ClustalW, and then the Phylip programs protdist and fitch were used to create the tree. The major branches of the phylogenetic tree are named according to (Axelsen and Palmgren, 1998). B, Overview of plant P-type ATPases showing topology differences and ion specificities. Branches are designated from P_{1B} to P_5, with protein names underneath. The putative transported ions (substrates) are indicated. Boxes, Transmembrane segments; black circles, regulatory regions containing autoinhibitory sequences; white circles, HMA domains; black squares, CC dipeptide domains; white squares, poly-His domains. Cluster numbers correspond to the numbers found in the phylogenetic tree in A. For each cluster, the number of sequences in rice and Arabidopsis is indicated. If a subcellular location is known for a member of that group, it is noted in the right column. PM, Plasma membrane; Vac, vacuole.

and AHAs (P_{3A}), a similar delineation of clusters has previously been proposed (Pittman et al., 1999; Arango et al., 2003). For genes in all 23 clusters, an inspection of intron positions provides strong corroboration of clustering patterns (see Figs. 2–7). Within a cluster, there is a striking conservation of intron numbers and positions; within each cluster, an average of 82% of all intron positions are conserved. In contrast, intron patterns are highly dissimilar between clusters. Together this suggests that (a) the angiosperm ancestor had at least 23 P-type ATPases, (b) these 23 P-type ATPases had been present for a considerable time period before the split, and (c) the gene structure has been relatively stable since the split from the common angiosperm ancestor.

In total, we identified only three putative ATPase-like proteins that have features suggesting a dysfunctional ATPase or a pseudogene. Two were previously reported for Arabidopsis (At5g53010; Fig. 4; and At4g11730; Fig. 5) with similarity to ACA and AHA pumps, respectively (Axelsen and Palmgren, 2001), whereas a third was found here in rice (PlantsT no. 64568; Fig. 6) with similarity to ALA pumps. Although these putative proteins have regions of strong identity to P-type ATPases, they lack one or more regions considered essential for activity. Because none of these ATPase-like genes has a corresponding homolog in the other species, they either represent psuedogenes or encode unique, species-specific ATPase-like proteins.

Three Rice Genes Have Arabidopsis-Like GC Profiles

It has been observed that most monocot genes (with experimentally confirmed gene models) have a negative gradient of GC content starting with the start codon (Wong et al., 2002). This phenomenon has been observed in multiple monocots and is absent in all dicots studied. The change in GC content is mainly a function of the wobble position of monocot codons. As a result, the 5′ end of monocot genes can have up to 25% higher GC content than the 3′ end (Yu et al., 2002). These gradients have not been reported in Arabidopsis, which has an average GC content of 43% throughout the genic regions (Yu et al., 2002).

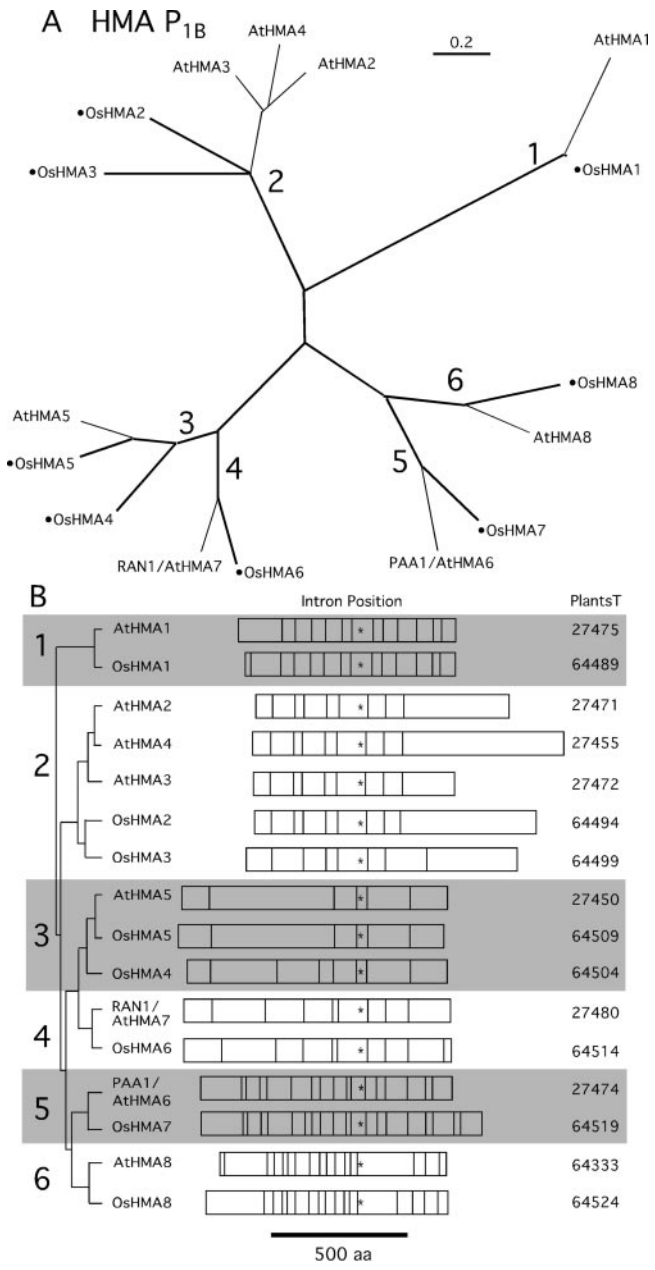


Figure 2. A, Phylogenetic tree of the P_{1B} subfamily revealing six clusters. The tree was constructed from full-length protein sequences. Numbers denote clusters. Branches leading to the rice isoforms are thicker and the rice isforms are preceded by a bullet. The bootstrap values for OsHMA2 and OsHMA3 having a separate branch were 51/100 so they were collapsed to the nearest node. All other bootstrap values are greater than 98/100. B, Distributions of introns and exons in the P_{1B} subfamily arranged in clusters. The phylogenetic tree used to define the clusters is taken from A but is shown in a different view. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star. Phylogenetic tree branch lengths are not to scale but are included to show groupings.

The majority of the rice P-type ATPases (33 of 43 genes) have 5′ GC gradients (Fig. 8). Of the Arabidopsis P-type ATPases, only *AtHMA5*, whose gene model has yet to be confirmed by a cDNA, has a small gradient. Interestingly, three genes, *OsAHA9*, *OsALA9*, and *OsALA10*, have flat GC profiles similar to Arabidopsis genes. Although the potential functions of these gradients in transcription and translation remain unknown, it is clear that since the monocot/dicot divergence, either the monocots have evolved these gradients or the dicots have lost them. Within the context of the rice P-type ATPases, the three genes with "flat-GC profiles" present an interesting opportunity to study this phenomenon from the perspective of regulatory function and genome evolution.

HMA, P_{1B}

There are eight P_{1B} ATPases in rice just like in Arabidopsis. The P_{1B} ATPases are divided into six

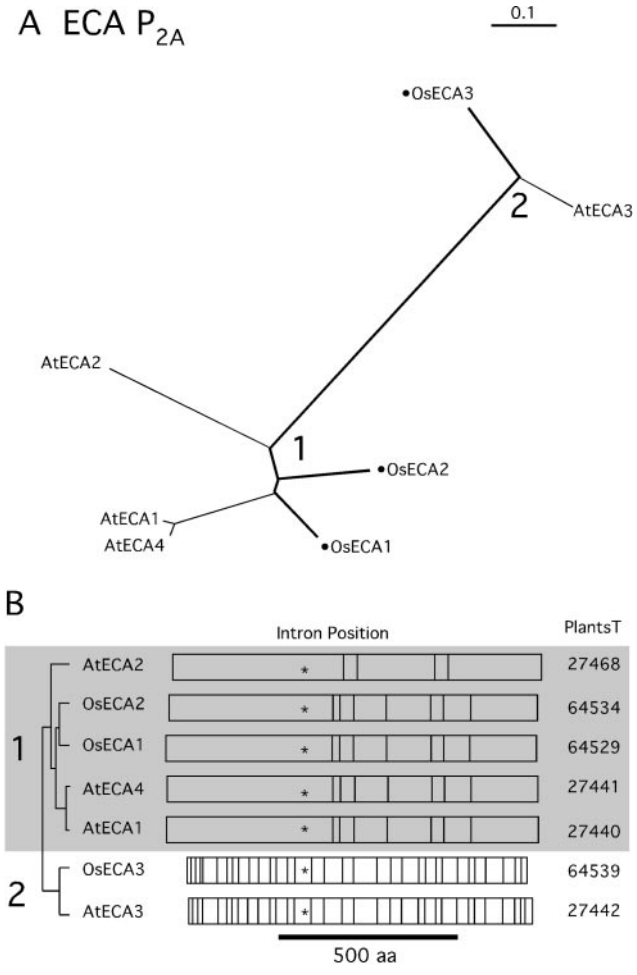


Figure 3. A, Phylogenetic tree of the P_{2A} subfamily revealing two clusters. The tree was constructed from full-length protein sequences. Numbers denote clusters. Branches leading to the rice isoforms are thicker, and the rice isforms are preceded by a bullet. All bootstrap values are greater than 90/100. B, Distributions of introns and exons in the P_{2A} subfamily arranged in clusters. The phylogenetic tree used to define the clusters is taken from A but is shown in a different view. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star. Phylogenetic tree branch lengths are not to scale, but are included to show groupings.

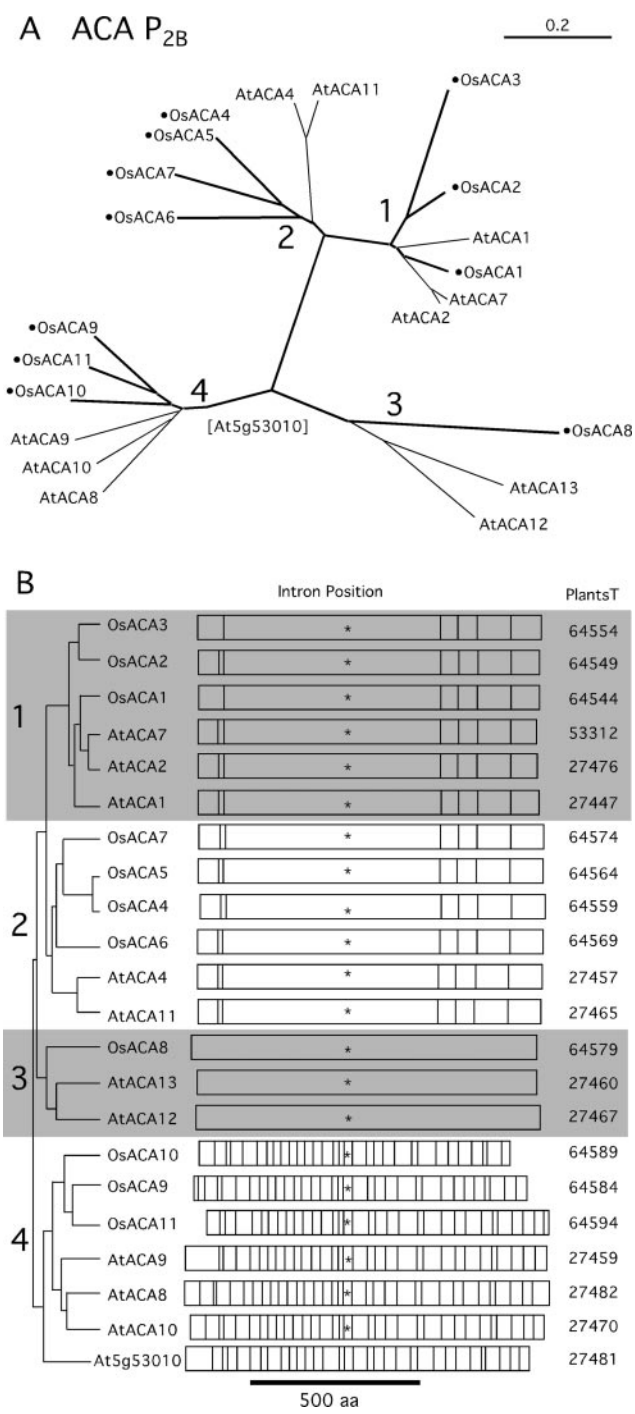


Figure 4. A, Phylogenetic tree of the P_{2B} subfamily revealing four clusters. The tree was constructed from full-length protein sequences. Numbers denote clusters. Branches leading to the rice isoforms are thicker, and the rice isoforms are preceded by a bullet. All bootstrap values are greater than 71/100. An ACA-like protein is noted in brackets. B, Distributions of introns and exons in the P_{2B} subfamily arranged in clusters. The phylogenetic tree used to define the clusters is taken from A but is shown in a different view. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star. Phylogenetic tree branch lengths are not to scale but are included to show groupings. An ACA-like protein is listed by its Arabidopsis Genome Institute (AGI) number.

clusters, based on sequence alignments and intron positions (Fig. 2). Seven Arabidopsis P_{1B} ATPases were previously identified (Axelsen and Palmgren, 2001), but the completion of the Arabidopsis sequencing revealed an eighth gene, *AtHMA8*, which is most identical (43%) to *PAA1/AtHMA6*. It was previously observed that Arabidopsis underwent an expansion in numbers of P_{1B} ATPases compared with other eukaryotic organisms, which only have one or two (Axelsen and Palmgren, 2001). The parallel discovery of eight P_{1B} ATPases in rice suggests that this expansion was important for the evolution of angiosperms (for a recent review of heavy-metal transport in plants, see Williams et al., 2000).

Clusters 1 and 2 are possible $Zn^{2+}/Co^{2+}/Cd^{2+}/Pb^{2+}$ ATPases, whereas isoforms in clusters 3 to 6 are believed to be Cu^{+}/Ag^{+} ATPases (Axelsen and Palmgren, 2001). Although pumps within clusters share at least 50% identity, the identity between the farthest clusters is less than 27%, giving the HMA (P_{1B}) subfamily the distinction of the most diverse subfamily of plant pumps.

Several heavy-metal-binding motifs have been identified in the plant HMAs, including heavy-metal-associated domains, CC dipeptide motifs, and poly-His motifs (Fig. 1). One key difference between clusters is the kind, numbers, and location of heavy-metal-binding motifs. For example, pumps from clusters 1 and 2 lack the N-terminal heavy-metal-associated domains found in pumps from clusters 3 to 6. Instead, they have a poly-His motif (cluster 1) or a CC dipeptide (cluster2) in the N-terminal domain.

HMA Regulation

No regulatory mechanisms have been elucidated for the plant HMAs. However, the presence of long C- and N-terminal cytoplasmic tails containing heavy-metal-binding motifs point to these regions as potentially significant for ion sensing or regulation.

HMA Phenotypes

There are two examples of experimental evidence that support the predicted substrate specificity of plant HMAs: RAN1/AtHMA7 (cluster 4) and PAA1/AtHMA6 (cluster 5), both of which are putative Cu^{+}/Ag^{+}-pumps. First, experiments by Hirayama et al. (1999) looking for mutants that show ethylene phenotypes in response to trans-cyclooctene, an ethylene antagonist, have implicated RAN1/AtHMA7 in ethylene-signaling pathways. They showed that the "ethylene-related" phenotype could be rescued by excess Cu and that the plant pump could complement a mutation in yeast of *ccc2*Δ, a yeast copper ATPase. Second, experiments by Shikanai, Pilon, and Niyogi have shown PAA1/HMA6 mutants have restricted electron transport between the PS and PQ1 proteins of the photosynthetic apparatus, less holo-

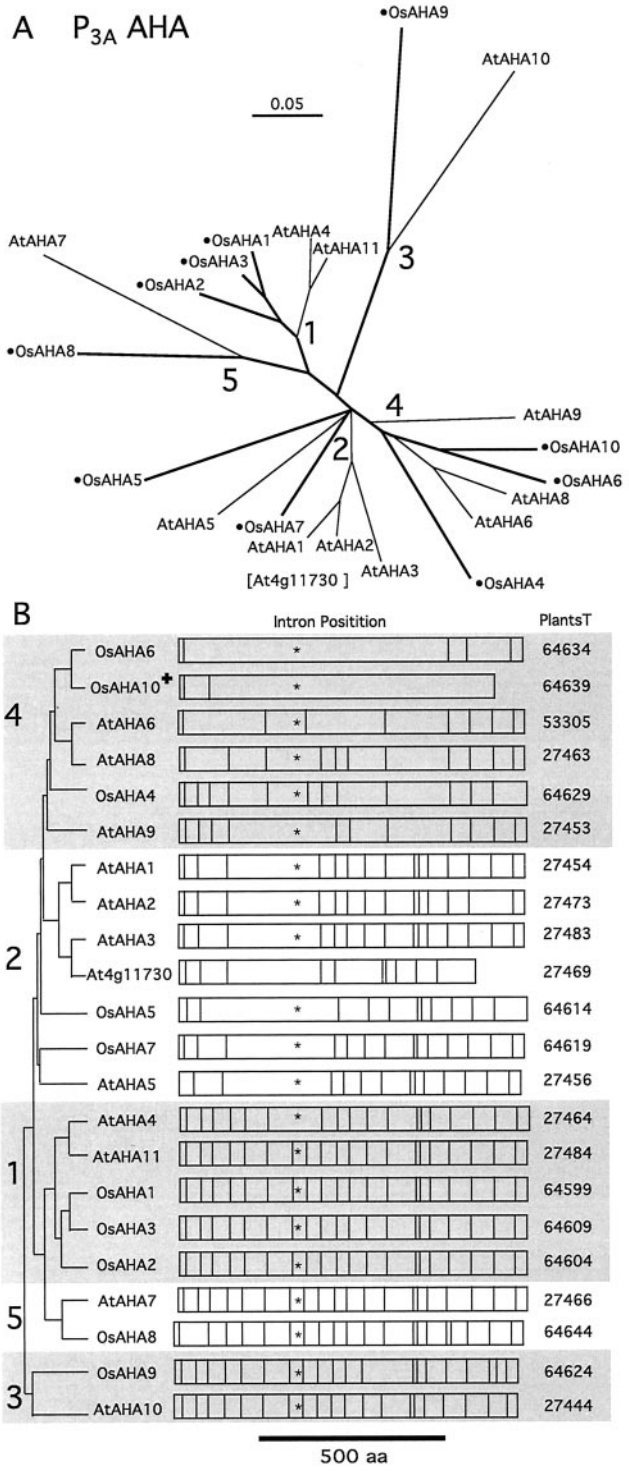


Figure 5. A, Phylogenetic tree of the P_{3A} subfamily revealing five clusters. The tree was constructed from full-length protein sequences. Numbers denote clusters. Branches leading to the rice isoforms are thicker, and the rice isoforms are preceded by a bullet. The branching of OsAHA5 and 7 and AtAHA7 is ambiguous (bootstrap values 24/100), so the branches were collapsed to a common node. All other values are greater than 73/100. AHA-like protein is noted in brackets. B, Distributions of introns and exons in the P_{3A} subfamily arranged in clusters. The phylogenetic tree used to define the clusters is taken from A but is shown in a different view. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star. The position is marked on the scale below. Phylogenetic tree branch lengths are not to scale but are included to show groupings. An AHA-like protein is listed by its AGI number. OsAHA10$^+$ is missing a C-terminal autoinhibitor domain.

plastocyanin in the thylakoid lumen, and less holo-Cu/ZnSOD in the chloroplast stroma (T. Shikanai, M. Pilon, and K. Niyogi, personal communication). Cu^{2+} levels in chloroplasts isolated from PAA1 mutant plants are lower, whereas the levels in whole leaves are the same as in wild-type. The electron transport defect is rescued by Cu^{2+} feeding. Nevertheless, a direct biochemical test of Cu^{2+} transport by RAN1/AtHMA7 and PAA1/AtHMA6 has not been reported.

HMA Subcellular Localization

Very little is known about the subcellular locations for members of clusters 1, 3, 4, and 6. However, in cluster 2, AtHMA2 has been localized to the plasma membrane by green fluorescent protein fusions and immunodetection in membrane fractionation with an anti-HMA2 antiserum (Y. Wang and J.F. Harper, unpublished data). In cluster 5, PAA1/AtHMA6 has been localized to the chloroplast by fusing an N-terminal peptide to a passenger protein and showing chloroplast import (T. Shikanai, M. Pilon, and K. Niyogi, unpublished data). The PAA1/AtHMA6 phenotypes suggest that it is localized to the chloroplast envelope.

ECA, P_{2A}

The P_{2A} ATPases include the well-characterized mammalian SERCA Ca^{2+}-ATPases (Carafoli and Brini, 2000; Geisler et al., 2000a; Sze et al., 2000). Three members of this subfamily are found in the rice genome, compared with four in Arabidopsis. The P_{2A} ATPases are divided into two clusters, based on sequence alignments and intron positions (Fig. 3). Within cluster 1, *AtECA2* may be special, because it has lost three of seven introns and has diverged from the other pumps in its cluster (Fig. 3). Experiments in yeast suggest that ECAs can transport Ca^{2+}, Mn^{2+}, and Zn^{2+} (Wu et al., 2002).

The previously reported gene model for *OsECA1* (Chen et al., 1997) has been changed here as a result of comparisons with other homologs. The gene *OsECA1* was first identified as a gibberillic acid-induced gene in the aleurone layer (Chen et al., 1997). The new version (PlantsT no. 64529) has been corrected to match the updated rice genomic sequence as well as its most similar Arabidopsis homolog, *AtECA3*.

ECA Regulation

In mammals, the P_{2A} pumps are regulated by phospholambans. Neither phospholambans nor other reg-

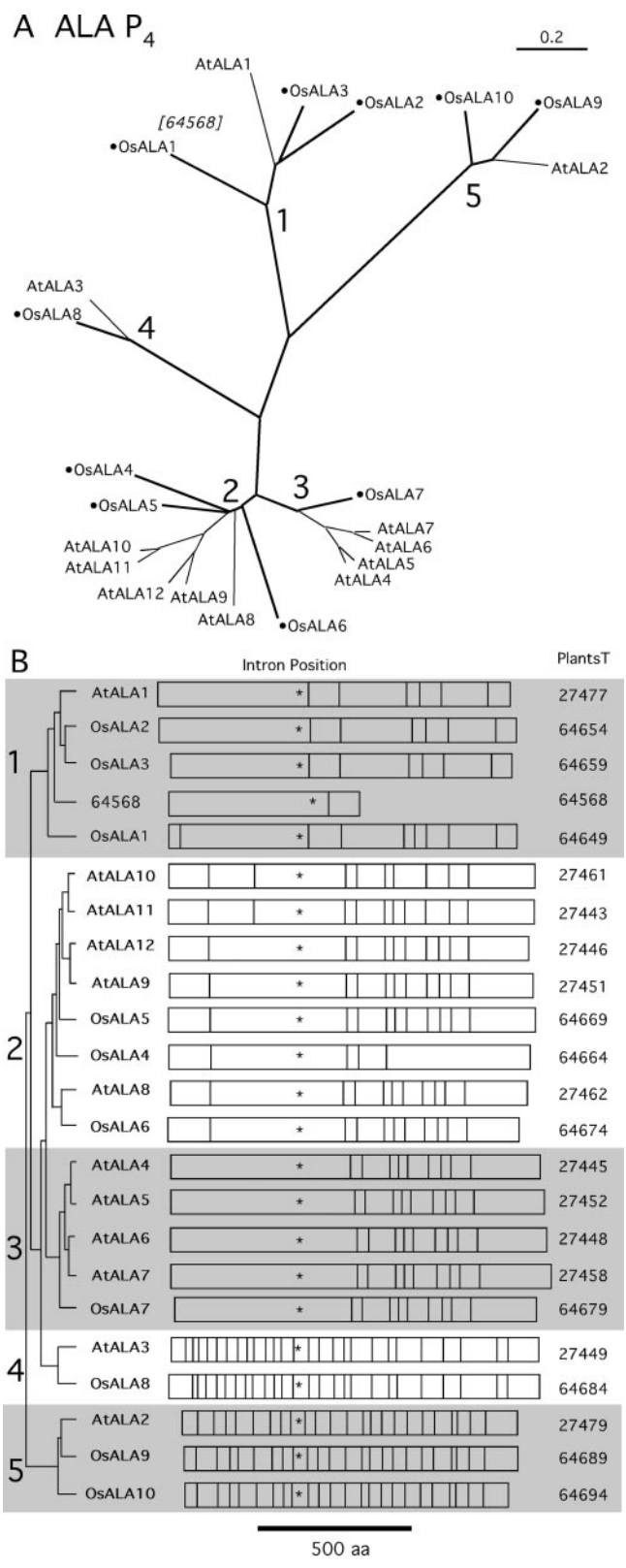


Figure 6. A, Phylogenetic tree of the P_4 subfamily revealing five clusters. The tree was constructed from full-length protein sequences. Numbers denote clusters. Branches leading to the rice isoforms are thicker, and the rice isoforms are preceded by a bullet. ALA-like gene is in brackets near the appropriate cluster. The bootstrap values for OsALA4 and OsALA5 having a separate branch were 43/100 so they were collapsed to the nearest node. The rest of the bootstrap values were greater than 76/100. B, Distributions of introns and exons in the P_4 subfamily arranged in clusters. The phylogenetic tree used to define the clusters is taken from A but is shown in a different view. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star. Phylogenetic tree branch lengths are not to scale but are included to show groupings. An ALA-like gene is noted by its PlantsT number.

ulators of P_{2A} proteins have been identified in plants, and the phospholamban-binding site on human SERCA2 is not conserved in the plant ECA proteins.

ECA Phenotypes

Plants with a transfer DNA disruption of AtECA1 grow poorly on low Ca^{2+} or high Mn^{2+} concentrations in the media and have reduced ability to grow root hairs when grown on high Mn^{2+} (Wu et al., 2002).

ECA Subcellular Localization

The only localization experiments reported for an Arabidopsis ECA confirmed an ER location for *AtECA1* (Hong et al., 1999). This is consistent with a similar ER location for the closest related animal homologs (SERCAs). Nevertheless, one report of a tomato (*Lycopersicon esculentum*) ECA isoform located at the plasma membrane (Ferrol and Bennett, 1996) suggests that the location of every ECA isoform needs to be experimentally verified in an isoform-specific fashion.

ACA, P_{2B}

The plant P_{2B} ATPases are related to the mammalian plasma membrane calmodulin-stimulated ATPases, but differ in having an N-terminal rather than a C-terminal autoinhibitor (Carafoli and Brini, 2000; Geisler et al., 2000; Sze et al., 2000). There are 10 P_{2B} ATPases in Arabidopsis and 11 in rice. The P_{2B} ATPases are divided into four clusters, based on sequence alignments and intron positions (Fig. 4). Cluster 3 is special because it is the only cluster of P-type ATPases that harbors intron-less genes. In contrast, genes in cluster 4 have as many as 33 introns.

ACA Regulation

In plants, the four best characterized P_{2B} ATPases have been shown to have an N-terminal, calmodulin-regulated autoinhibitory domain (i.e. AtACA2, AtACA4, AtACA8, and cauliflower [*Brassica oleracea*] BCA1; Malmström et al., 1997; Harper et al., 1998; Bonza et al., 2000; Geisler et al., 2000b). When Ca^{2+}/calmodulin is bound, the ATPases are stimulated (for recent reviews, see Geisler et al., 2000a; Sze et al.,

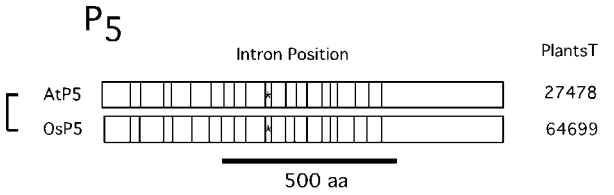


Figure 7. Distributions of introns and exons in the P_5 subfamily forming a single cluster. The genes are aligned around the phosphorylation site, [DKTGT], marked by a star.

2000). However, the extrapolation that all members of this subfamily are regulated in a similar way was brought into question by an in silico analysis of cluster 3 (*OsACA3*, *AtACA12*, and *13*). When cluster 3 pumps were aligned with other ACAs, they showed a marked divergence in the putative calmodulin-binding region (Supplemental Data D). In addition, they showed a substitution of Asp to Asn in the otherwise conserved motif SDY[R/K/Q]Q juxtaposed to transmembrane 2 on the cytosolic surface. A substitution of Asp to Asn in AtACA2 has been shown to render this isoform calmodulin independent (i.e. constitutively active; Curran et al., 2000). Thus, isoform-specific differences in regulation are expected among the ACAs.

ACA Phenotypes

Although no studies of plant lines harboring P_{2B} genes have been published, a disruption of every calcium pump has been obtained (J.F. Harper and M.G. Palmgren, unpublished data) from large insertional mutant collections (Young et al., 2001; Sessions et al., 2002). Thus, we anticipate great advances in our understanding of the role of the ACAs in plants in the near future.

ACA Subcellular Localization

In contrast to animals, where the analogous P_{2B} pumps are located only in the plasma membrane, plants appear to have ACA pumps in multiple subcellular locations. AtACA2, a member of cluster 1, has been localized to the ER (Harper et al., 1998). AtACA4, a member of cluster 2, has been localized to the vacuole (Geisler et al., 2000b). AtACA8, a member of cluster 4, has been shown to reside in the plasma membrane (Bonza et al., 2000). A representative of cluster 3 has not yet been characterized.

AHA, P_{3A}

The P_{3A} H^+-ATPases are not found in animal systems. Arabidopsis has 11 while rice has 10. Of the six subfamilies of P-type ATPases, the P_{3A} branch shows the least divergence. The P_{3A} ATPases are divided into five clusters, based on sequence alignments and intron positions (Fig. 5). Although the bootstrap values that delineate clusters 2 and 4 are quite low, this branch is further supported by a similar clustering analysis that included additional tobacco (*Nicotiana tabacum*) genes (Arango et al., 2003).

AHA Regulation

P_{3A} ATPases have been subjected to substantial biochemical characterization, because they energize the plasma membrane and are likely to be under tight regulation to be able to quickly respond to the cells' needs (Palmgren, 2001). Differences in biochemical activity between isoforms when expressed in Brewer's yeast have been noted for ATPases from Arabidopsis and tobacco. The C terminus has been shown to harbor an autoinhibitory region as well as other

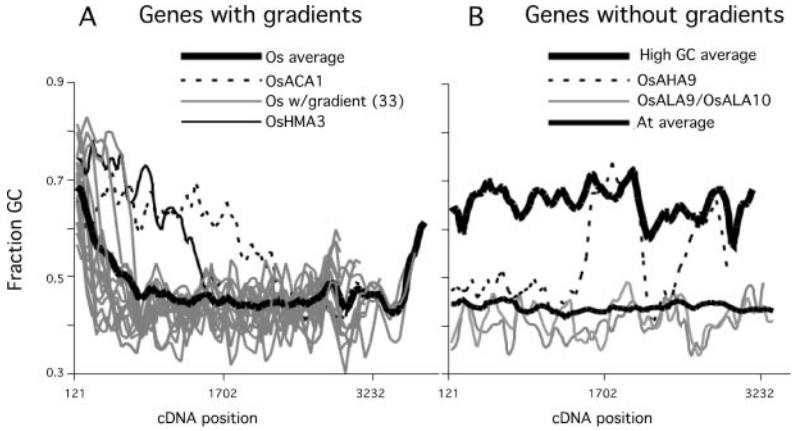


Figure 8. GC gradient analysis reveals three Arabidopsis-like profiles in rice. GC content of DNA sequences was calculated using an in-house program called GeneGC (http://plantst.sdsc.edu/plantst/html/geneGC.shtml). The GC content was calculated in all of the full-length P-type ATPases in rice and Arabidopsis using a 121-bp sliding window that moved in steps of 51 bp. The GC content of the Arabidopsis sequences varies between 40% and 50%. The average of all the Arabidopsis genes is shown in B for comparison. All of the rice genes with 5′ gradients are shown in A. The genes in B do not have 5′ gradients and were classified as having high GC content or Arabidopsis-like "normal" GC content. Only the averages of the genes with high GC content are shown. Genes with unusual GC patterns have different markers.

regulatory sites. The C-terminal regions (after the final transmembrane helix) are the most divergent regions of the proteins (36% identical versus 73% for the protein core). The C termini bind 14-3-3 proteins as a part of the activation mechanism, a binding that depends upon phosphorylation of the conserved penultimate Thr (for recent reviews, see Palmgren, 2001; Arango et al., 2003). Interestingly, OsAHA10 is predicted to be missing the entire C-terminal domain, which would make it the only constitutively active proton pump found in plants (Arango et al., 2003). Although there are ESTs indicating that *OsAHA10* is expressed, a sequence confirmation of a "truncated end" has not been established.

AHA Phenotypes

Genetic disruption and RNA interference have been carried out on the members of cluster 1 in both Arabidopsis and tobacco. In Arabidopsis, nondominant, double disruptions of *AtAHA4* and *AtAHA11* (the only Arabidopsis genes in the cluster) have no apparent effect on the health of the plant (I. Baxter and J.F. Harper, unpublished data). RNA interference of the three tobacco genes from this cluster also had no observable effect on plant development (F. Gévaudant, J. Kanczewska, and M. Boutry, unpublished data). Unexpectedly, a transfer DNA insertion in AtAHA4, which segregates as a semidominant mutation, confers salt sensitivity (Vitart et al., 2001).

Genetic disruption of several isoforms from cluster 2 shows deleterious effects on the health of the plants. RNA interference of PMA4 in tobacco causes defects in development, sugar transport, and male sterility (Zhao et al., 2000). The double knockout of *AtAHA1* and *2* is embryonic lethal, and *AtAHA3* is essential for male gametogenesis (M.R. Sussman, unpublished data). At present, all biochemical, genetic disruption, and RNAi studies of proton pumps in plants have been performed with members of clusters 1 and 2. Evidence for similar or unique functions of pumps from cluster 3, 4, or 5 still needs to be explored.

AHA Subcellular Localization and Tissue Specificity

The plasma membrane proton pump has long been considered a marker of the plasma membrane, based on immunocytology and membrane fractionation (DeWitt et al., 1996). Nevertheless, a PM targeting destination has been confirmed for only two of the 11 Arabidopsis isoforms using an isoform-specific epitope tagging strategy. Thus, the potential of an endomembrane targeting location for one of the uncharacterized isoforms remains a formal possibility.

The favored rationale to explain the presence of so many isoforms is the need to express different amounts of pump in different tissues. Tissue-specific expression patterns have been observed for different isoforms in multiple plant species, including Arabidopsis and tobacco (Arango et al., 2003). For example, *AtAHA10* has been localized to the developing seed (Harper et al., 1994). *AtAHA4* is expressed in the root endodermis, in flowers, and in maturing siliques (Vitart et al., 2001). AHA2 is expressed predominantly but not exclusively in the root epidermis (J.F. Harper, unpublished data). *AtAHA9* is expressed in anthers (Houlne and Boutry, 1994). ESTs of *OsAHA10* have been found in the rice panicle at flowering stage. However, potential orthologous isoforms from tobacco and Arabidopsis do not always show the same tissue-specific pattern of expression, suggesting that the distinction between clusters is not related to tissue specificity (Arango et al., 2003).

Putative ALA, P_4

This large subfamily of divergent ATPases, which is only found in eukaryotes, includes 12 and 10 genes in Arabidopsis and rice, respectively. The P_4 ATPases are divided into five clusters, based on sequence alignments and intron positions (Fig. 6). The previously reported gene model for *AtALA3* has been changed based on the similarity with its ortholog, *OsALA8*. The donor site of the second-to-last exon has been moved four nucleotides upstream, resulting in the last exon being read in another frame and making the protein 90 amino acids longer. The P_4 ATPases have been implicated in aminophospholipid flipping. However, it is not clear if the flipase activity is the primary function of P_4 ATPases or an indirect effect of its activity.

ALA Phenotypes

The subfamily is relatively uncharacterized; only one isoform in plants, AtALA1, has been studied (Gomes et al., 2000). Antisense suppression of AtALA1 demonstrated a role in cold tolerance of Arabidopsis plants (Gomes et al., 2000). Work on other P_4 ATPases showed that the yeast DRS2 protein is involved in formation of clathrin-coated vesicles (Gall et al., 2002), whereas human ATP8B1 is involved in transport of conjugated bile acids (Bull et al., 1998).

Unknown (P5, P_5)

There is a single member of the P_5 ATPases in each organism; the proteins are 77% identical (Fig. 7). The P_5 ATPases are the least studied of the P-type ATPase subfamilies. The only protein to have been studied in some detail is SPF1/COD1, one of two members in Brewer's yeast (Cronin et al., 2002; Vashist et al., 2002). The studies have given no clear idea of ion specificity, but they do suggest a functional connection to lipid biogenesis and pathogen response.

CONCLUSIONS

The current draft sequences of the rice genome have enabled the first comprehensive comparison of P-type ATPase genes in two plant species. Even though the respective pumps and their genes are remarkably similar, the analysis of the rice sequences has done more than confirm the previous findings in Arabidopsis. The genomic comparison has increased our knowledge in five ways. (a) Because the majority of the 89 P-type ATPase genes in these two organisms have not been cloned as full-length cDNAs, we have been able to correct and verify predicted gene models, such as the change described here for AtALA3 based on similarity to OsALA8. (b) Comparison of protein sequences and intron organization from the two plant species has confirmed that at least 23 ATPases evolved very early in evolution before the divergence of dicots and monocots. (c) The observation that no cluster was lost in Arabidopsis or rice suggests that all 23 archetypal clusters play an important role for plant survival at the evolutionary scale. The observation that knockout or RNAi of some clusters (as discussed above) has little impact might simply indicate that these clusters are not essential in tested growth conditions but that they presumably confer some advantage to the plant in nature. (d) Of the three ATPase-like proteins in Arabidopsis or rice, none are conserved in both organisms, suggesting that they represent pseudogenes or have species-specific functions. (e) Within some clusters, the number of genes varies between rice and Arabidopsis. The greatest expansion was in ALA cluster 3, with one isoform in rice and four in Arabidopsis. This example of gene duplication or deletion may reflect gene specialization and species-specific differences in transport activities. The most challenging task ahead is to understand the conserved and species-specific functions of all 89 different rice and Arabidopsis P-type ATPases.

MATERIALS AND METHODS

Identification

To identify all P-type ATPases in the rice genome, we searched three different databases: (a) the Rice Annotated Protein Database at The Institute for Genome Research (including all sequences predicted from the International Rice Genome Sequencing Project), (b) the Syngenta Rice Genomic Sequence of *Lotus* subsp. *japonica* (Goff et al., 2002; http://www.tmri.org), and (c) the Bejing Genomics Institute 93–11 cultivar genomic sequences of *Lotus* subsp. *Indica* (downloaded May 27, 2002; http://btn.genomics.org.cn/rice). Both the Syngenta and the Bejing Genomics Institute sequences were reported to cover more than 90% of known rice genes or Unigene cluster at the time of analysis (Goff et al., 2002; Yu et al., 2002). If the contigs from the three sequencing projects are evenly distributed over the genome, we searched more than 99% of the genome for P-type ATPases.

Contigs that contained P-type ATPase genes were identified by TBlastN (Altschul et al., 1990) searches using either the 46 Arabidopsis P-type ATPases or the six currently known variations of the DKTGTXX motif that are conserved in all known P-type ATPases (http://biobase.dk/~axe/Patbase.html). Forty-eight occurrences of the DKTGTXX motifs were found in the Syngenta genomic sequence and 56 in the Chinese genomic sequence. A nonredundant set of 43 rice P-type ATPases was identified after making gene model predictions.

Editing

Gene models (open reading frames) were predicted by submitting 8,000 bp of genomic sequence encompassing each potential gene to Genemark (http://opal.biology.gatech.edu/GeneMark/; M. Borodovsky and A. Lukashin, unpublished data). In some cases, longer sequences were examined to ensure the inclusion of the 5′ or 3′ end of the genes. The predicted coding sequences were aligned with the closest Arabidopsis P-type ATPase homolog using ClustalW (Thompson et al., 1994). The candidate rice genes that either did not have any recognizable open reading frames or coded for proteins other than P-type ATPases were eliminated. Sequences that appeared to be bacterial in nature, including several putative Mg^{2+} ATPases (P_{3B}), were also eliminated from the analyses. The sequences in question are being stored in a separate database on the PlantsT Web site (http://plantst.sdsc.edu). Possible pseudogenes were removed from analyses where appropriate.

The remaining 43 sequences were edited by hand, using the closest Arabidopsis homolog as a template, to improve intron/exon predictions. Although Genemark was useful as a starting point for creating gene models, this program correctly predicted only one of the rice genes, which is in agreement with the predicted accuracy of gene model prediction programs (Pertea and Salzberg, 2002). Although most gene models were based entirely on the Syngenta *japonica* sequence, in a few cases we combined sequencing data from both the Syngenta and Chinese sequences as well as from EST data to make complete predictions. The final gene models are deposited in GenBank (http://www.ncbi.nlm.nih.gov) and PlantsT (http://plantst.sdsc.edu) databases. Updated annotations to our predicted genes can be submitted to the PlantsT Web site.

Analysis

Phylogenetic analyses were performed using the Phylip package (Phylogeny Inference Package, v3.2, Department of Genetics, University of Washington, Seattle). Initial alignments were done with ClustalW using all of the rice and Arabidopsis P-type ATPases. The segments conserved among all P-type ATPases (Axelsen and Palmgren, 1998) were extracted using the initial alignments with the Arabidopsis isoforms (http://biobase.dk/~axe/Patbase.html). Alignments using the conserved sequences produced the same phylogenetic trees as alignments using full-length sequences, so we used full-length sequences for all analysis. After alignment with ClustalW, bootstrapping was performed using the Phylip programs seqboot, protdist, and consense. Subsequently, the proteins were divided up into subfamily and alignments, and bootstrapping analysis was done on each subfamily individually (subfamily alignments are included as Supplemental Data B–G). Similarity was calculated using the protdist program from Phylip v3.6a3 (Department of Genetics, University of Washington, Seattle). Introns, exons, and gene lengths were calculated by comparing the cDNA sequence to the genomic sequence with the Sim4 program (Florea et al., 1998). A program was written to calculate the GC content along the length of a gene (GeneGC, http://plantst.sdsc.edu/plantst/html/geneGC.shtml). The GC content across the genes was calculated using a 129-bp sliding window that moved in 51-bp steps.

Received February 9, 2003; returned for revision February 25, 2003; accepted February 25, 2003.

LITERATURE CITED

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Arango M, Gévaudant F, Oufattole M, Boutry M (2003) The plasma membrane proton pump ATPase: the significance of gene subfamilies. Planta **216:** 355–365

Axelsen KB, Palmgren MG (1998) Evolution of substrate specificities in the P-type ATPase superfamily. J Mol Evol **46:** 84–101

Axelsen KB, Palmgren MG (2001) Inventory of the superfamily of P-type ion pumps in Arabidopsis. Plant Physiol **126:** 696–706

Banuelos MA, Garciadeblas B, Cubero B, Rodriguez-Navarro A (2002) Inventory and functional characterization of the HAK potassium transporters of rice. Plant Physiol **130:** 784–795

Baumberger N, Doesseger B, Guyot R, Diet A, Parsons RL, Clark MA, Simmons MP, Bedinger P, Goff SA, Ringli C et al. (2003) Whole-genome comparison of leucine-rich repeat extensins in Arabidopsis and rice: A conserved family of cell wall proteins form a vegetative and a reproductive clade. Plant Physiol **131:** 1313–1326

Bonza MC, Morandini P, Luoni L, Geisler M, Palmgren MG, De Michelis MI (2000) At-ACA8 encodes a plasma membrane-localized calcium-ATPase of Arabidopsis with a calmodulin-binding domain at the N terminus. Plant Physiol **123:** 1495–1506

Bukrinsky JT, Buch-Pedersen MJ, Larsen S, Palmgren MG (2001) A putative proton binding site of plasma membrane H^+-ATPase identified through homology modelling. FEBS Lett **494:** 6–10

Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, Lomri N, Berger R, Scharschmidt BF et al. (1998) A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. Nat Genet **18:** 219–224

Carafoli E, Brini M (2000) Calcium pumps: structural basis for and mechanism of calcium transmembrane transport. Curr Opin Chem Biol **4:** 152–161

Catty P, de Kerchove d'Exaerde A, Goffeau A (1997) The complete inventory of the yeast *Saccharomyces cerevisiae* P-type transport ATPases. FEBS Lett **409:** 325–332

Chen X, Chang M, Wang B, Wu B (1997) Cloning of a Ca^{2+}-ATPase gene and the role of cytosolic Ca^{2+} in the gibberellin-dependent signaling pathway in aleurone cells. Plant J **11:** 363–371

Cronin SR, Rao R, Hampton RY (2002) Cod1p/Spf1p is a P-type ATPase involved in ER function and Ca^{2+} homeostasis. J Cell Biol **157:** 1017–1028

Curran AC, Hwang I, Corbin J, Martinez S, Rayle D, Sze H, Harper JF (2000) Autoinhibition of a calmodulin-dependent calcium pump involves a structure in the stalk that connects the transmembrane domain to the ATPase catalytic domain. J Biol Chem **275:** 30301–30308

DeWitt ND, Hong B, Sussman MR, Harper JF (1996) Targeting of two Arabidopsis H^+-ATPase isoforms to the plasma membrane. Plant Physiol **112:** 833–844

Ferrol N, Bennett AB (1996) A single gene may encode differentially localized Ca^{2+}-ATPases in tomato. Plant Cell **8:** 1159–1169

Florea L, Hartzell G, Zhang Z, Rubin GM, Miller W (1998) A computer program for aligning a cDNA sequence with a genomic DNA sequence. Genome Res **8:** 967–974

Gall WE, Geething NC, Hua Z, Ingram MF, Liu K, Chen SI, Graham TR (2002) Drs2p-dependent formation of exocytic clathrin-coated vesicles *in vivo*. Curr Biol **12:** 1623–1627

Geisler M, Axelsen KB, Harper JF, Palmgren MG (2000a) Molecular aspects of higher plant P-type Ca^{2+}-ATPases. Biochim Biophys Acta **1465:** 52–78

Geisler M, Frangne N, Gomes E, Martinoia E, Palmgren MG (2000b) The ACA4 gene of Arabidopsis encodes a vacuolar membrane calcium pump that improves salt tolerance in yeast. Plant Physiol **124:** 1814–1827

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science **296:** 92–100

Gomes E, Jakobsen MK, Axelsen KB, Geisler M, Palmgren MG (2000) Chilling tolerance in Arabidopsis involves ALA1, a member of a new family of putative aminophospholipid translocases. Plant Cell **12:** 2441–2454

Harper JF, Hong B, Hwang I, Guo HQ, Stoddard R, Huang JF, Palmgren MG, Sze H (1998) A novel calmodulin-regulated Ca^{2+}-ATPase (ACA2) from Arabidopsis with an N-terminal autoinhibitory domain. J Biol Chem **273:** 1099–1106

Harper JF, Manney L, Sussman MR (1994) The plasma membrane H^+-ATPase gene family in Arabidopsis: genomic sequence of AHA10 which is expressed primarily in developing seeds. Mol Gen Genet **244:** 572–587

Hirayama T, Kieber JJ, Hirayama N, Kogan M, Guzman P, Nourizadeh S, Alonso JM, Dailey WP, Dancis A, Ecker JR (1999) RESPONSIVE-TO-ANTAGONIST1, a Menkes/Wilson disease-related copper transport, is required for ethylene signaling in Arabidopsis. Cell **97:** 383–393

Hong B, Ichida A, Wang Y, Gens JS, Pickard BG, Harper JF (1999) Identification of a calmodulin-regulated Ca^{2+}-ATPase in the endoplasmic reticulum. Plant Physiol **119:** 1165–1176

Houlne G, Boutry M (1994) Identification of an *Arabidopsis thaliana* gene encoding a plasma membrane H^+-ATPase whose expression is restricted to anther tissue. Plant J **5:** 311–317

Jasinski M, Ducos E, Martinoia E, Boutry M (2003) The ATP-binding cassette transporters: structure, function, and gene family comparison between rice and Arabidopsis. Plant Physiol **131:** 1169–1177

Kühlbrandt W, Zeelen J, Dietrich J (2002) Structure, mechanism, and regulation of the *Neurospora* plasma membrane H^+-ATPase. Science **297:** 1692–1696

Malmström S, Askerlund P, Palmgren M (1997) A calmodulin-stimulated Ca^{2+}-ATPase from plant vacuolar membranes with a putative regulatory domain at its N-terminus. FEBS Lett **400:** 324–328

Ogawa H, Toyoshima C (2002) Homology modeling of the cation binding sites of Na^+K^+-ATPase. Proc Natl Acad Sci USA **99:** 15977–15982

Palmgren MG (2001) Plant plasma membrane H^+-ATPases: powerhouses for nutrient uptake. Annu Rev Plant Physiol Plant Mol Biol **52:** 817–845

Palmgren MG, Axelsen KB (1998) Evolution of P-type ATPases. Biochim Biophys Acta **1365:** 37–45

Pedersen P, Carafoli E (1987) Ion motive ATPases. I. Ubiquity, properties, and significance to cell function. Trends Biochem Sci **4:** 146–150

Pertea M, Salzberg SL (2002) Computational gene finding in plants. Plant Mol Biol **48:** 39–48

Pittman JK, Mills RF, O'Connor CD, Williams LE (1999) Two additional type IIA Ca^{2+}-ATPases are expressed in *Arabidopsis thaliana*: evidence that type IIA sub-groups exist. Gene **236:** 137–147

Radresa O, Ogata K, Wodak S, Ruysschaert JM, Goormaghtigh E (2002) Modeling the three-dimensional structure of H^+-ATPase of *Neurospora crassa*. Eur J Biochem **269:** 5246–5258

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C et al. (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell **14:** 2985–2994

Sze H, Liang F, Hwang I, Curran AC, Harper JF (2000) Diversity and regulation of plant Ca^{2+} pumps: insights from expression in yeast. Annu Rev Plant Physiol Plant Mol Biol **51:** 433–462

Thompson JD, Higgins DG, 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

Toyoshima C, Nakasako M, Nomura H, Ogawa H (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. Nature **405:** 647–655

Toyoshima C, Nomura H (2002) Structural changes in the calcium pump accompanying the dissociation of calcium. Nature **418:** 605–611

Vashist S, Frank CG, Jakob CA, Ng DT (2002) Two distinctly localized P-type ATPases collaborate to maintain organelle homeostasis required for glycoprotein processing and quality control. Mol Biol Cell **13:** 3955–3966

Vitart V, Baxter I, Doerner P, Harper JF (2001) Evidence for a role in growth and salt resistance of a plasma membrane H^+-ATPase in the root endodermis. Plant J **27:** 191–201

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

Wong GK-S, Wang J, Tao L, Tan J, Zhang J, Passey DA, Yu J (2002) Compositional gradients in *Gramineae* genes. Genome Res **12:** 851–856

Wu Z, Liang F, Hong B, Young JC, Sussman MR, Harper JF, Sze H (2002) An endoplasmic reticulum-bound Ca^{2+}/Mn^{2+} pump, ECA1, supports plant growth and confers tolerance to Mn^{2+} stress. Plant Physiol **130:** 128–137

Young JC, Krysan PJ, Sussman MR (2001) Efficient screening of Arabidopsis T-DNA insertion lines using degenerate primers. Plant Physiol **125:** 513–518

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science **296:** 79–92

Zhao R, Dielen V, Kinet JM, Boutry M (2000) Cosuppression of a plasma membrane H^+-ATPase isoform impairs sucrose translocation, stomatal opening, plant growth, and male fertility. Plant Cell **12:** 535–546