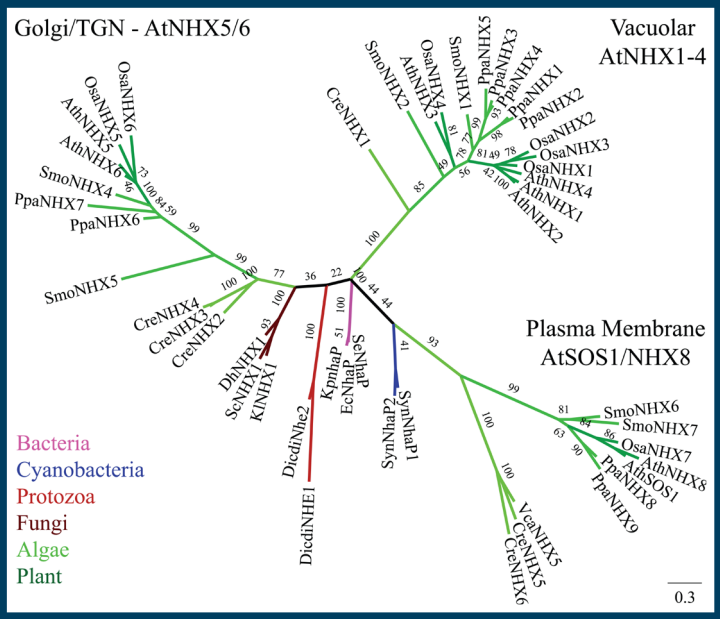
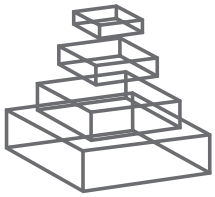
frontiers
RESEARCH TOPICS



EVOLUTION OF TRANSPORTERS IN PLANTS

Topic Editors
Heven Sze, Angus S. Murphy and Markus Geisler

frontiers in PLANT SCIENCE



frontiers

ISSN 1664-8714
ISBN 978-2-88919-228-1
DOI 10.3389/978-2-88919-228-1

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

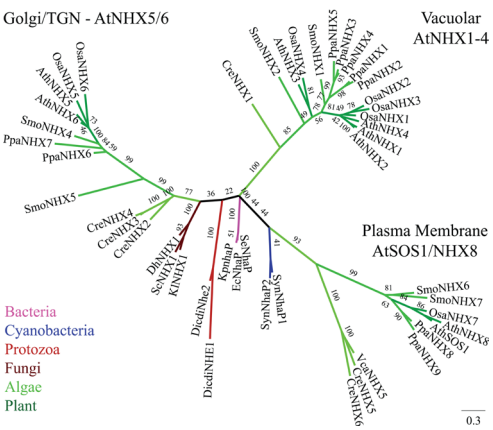
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

EVOLUTION OF TRANSPORTERS IN PLANTS

Topic Editors:
Heven Sze, University of Maryland, USA
Angus S. Murphy, University of Maryland, USA
Markus Geisler, University of Fribourg, Switzerland



In the past decade, many plant genomes have been completely sequenced ranging from unicellular alga to trees. This rich resource of information raises questions like: How did specific transporters evolve as early plants adapted to dry land? How did the evolution of transporters in monocot plants differ from that in dicots? What are the functional orthologs in food and energy crops of transporters characterized in model plants? How do we name the new genes/proteins? Phylogenetic analyses of transport proteins will shed light on these questions and potentially reveal novel insights for future studies to understand plant nutrition, stress tolerance, biomass production, signaling and development.

Table of Contents

frontiers in
PLANT SCIENCE

EDITORIAL
published: 08 January 2014
doi: 10.3389/fpls.2013.00547

Linking the evolution of plant transporters to their functions

Heven Sze[1]*, Markus Geisler[2] and Angus S. Murphy[3]

[1] *Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, USA*
[2] *Department of Biology, Plant Biology, University of Fribourg, Fribourg, Switzerland*
[3] *Plant Science and Landscape Architecture, University of Maryland, College Park, MD, USA*
**Correspondence: hsze@umd.edu*

Edited by:
Wendy A. Peer, University of Maryland, USA

Keywords: phylogeny of membrane transporters, early land plants, charophyte, molecular evolution, vascular development, angiosperms, pollen, moss

In the past decade, an increasing number of plant genomes ranging from unicellular alga to trees have been completely sequenced. As the transport of water, nutrients, hormones, and metabolites in aquatic plants could differ from that of non-vascular, vascular or flowering plants, this rich resource has the potential to answer many of our most urgent questions: How did specific transporters evolve as early plants adapted to dry land? Does the evolution of transporters in monocot plants differ from that in dicots? What are the orthologs in food and energy crops of critical transporters characterized in model plants? Can we infer functions or membrane localization from phylogeny? Phylogenetic analyses of transport proteins will shed light on these questions and potentially reveal novel insights for future studies to understand the contribution of transporters in plant nutrition, stress tolerance, biomass production, as well as in signaling and development.

We invited the plant biology community to participate in this effort as ∼5% of Arabidopsis thaliana (and probably other plant) genome encodes proteins with predicted transport roles. The *Arabidopsis thaliana* genome with 27 thousand protein-coding genes has over a 1000 genes classified as transporters, whereas the genomes of other flowering plants can be 2–7 times larger.

In this topic issue, several authors have examined the evolutionary history and diversification of a range of transporters and these initial results are yielding surprising insights. Some transporter families can be traced to the simplest green plants, including unicellular *Chlamydomonas reinhardtii* (Chlorophyta). Thus H^+-coupled antiporters like NHX (Chanroj et al., 2012), CAX, MHX, CCX (Emery et al., 2012), Sultr (Takahashi et al., 2012) and primary H^+ and Ca^{2+} pumps, such as AHA, ACA, ECA (Pedersen et al., 2012) are highly conserved in all green plants. In some cases, the gene family has diversified, for example, from 2 AHA genes in early land plant *Physcomitrella patens* to 11 genes in the flowering plant *A. thaliana*, indicating a need to regulate the proton motive force in specialized organs or cell types. In contrast, homologs of other transporter families in flowering plants were traced to a Charophyte, but not found in Chlorophyta. This history suggests that CHX co-transporters of K^+ (Chanroj et al.), LHT amino acid transporters (Tegeder and Ward, 2012) and SUT sucrose transporter (Reinders et al., 2012) had origins in an aquatic Charophyte, and further supports Charaphytes as basal to land plants. The increase in sibling genes within each family through gene duplication and diversification, can be attributed in part to innovations in vascular tissue development (LHT in phloem) and in sexual reproduction (LHT in pollen) on land. Strikingly, CHX genes have diversified in pollen especially in dicots, suggesting roles in the delivery of sperms in pollen tubes to distant ovules. Curiously, the purine uptake permease, PUP family exists in a vascular early land plant, *Selaginella moellendorfii*, but not in a non-vascular moss (*Physcomitrella patens*) (Jelesko, 2012). It is possible that PUP genes have a role in the transport of secondary metabolites for defense and/or reproductive success in vascular plants. Early land plants, such as moss and Selaginella, possess MIP homologs similar to PIP and TIP water channels (Anderberg et al., 2012), though it is not yet clear whether aquatic Charophytes need such proteins to transport water. Chlorophytes possess MIP-like homologs though their substrate(s) may be small uncharged molecules other than water.

A striking feature is that three branches of the AtNHX family correspond to $Na^+(K^+)/H^+$ antiporters associated with three distinct subcellular locations (plasma membrane, trans Golgi network and vacuolar membrane) in Arabidopsis (Chanroj et al., 2012). Though not all subclades of families are separated according to location, it does illustrate that finding orthologs from another species could aid in predicting function as well as membrane association (e.g., Reinders et al., 2012; Takahashi et al., 2012).

Facchinelli and Weber (2011) reviewed the evolutionary history of select metabolite transporters localized at the inner envelope of plastids. Phylogenetic studies are showing that *Arabidopsis thaliana* contributes 58% of the metabolite transporters vs. 12% from the endosymbiont cyanobacterium. The study illustrates that connecting plastid metabolism with that of the host cell is driving the evolution of metabolite transporters at the inner envelope. Authors also examine differences in phosphate transporters of photosynthetically-active plastids vs. heterotrophic plastids in non-green tissues. ATP/ADP exchangers localized to the inner envelope membrane of plastids are thought to provide ATP for reactions in heterotrophic plastids, or photosynthetic plastids during the night. However recent evidence for ATP/ADP antiporters on thylakoid membranes of Arabidopsis was surprising. Spetea et al. (2012) have traced the history of *thylakoid* ATP/ADP antiporters (TAAC) in plants and phylogenetic evidence suggests they arose once before the divergence of Chlorophyta and Streptophyta. These transporters are thought to supply ATP into the lumen for the biosynthesis and turnover of photosynthetic complexes.

Haferkamp and Schmitz-Esser (2012) re-analyzed the evolution of the Arabidopsis mitochondrial carrier family (MCF) catalyzing the transport of various substrates in mitochondria and other organelles. They found that various MCFs from Arabidopsis,

yeast and human built independent clades, allowing a prediction of their biochemical function. However, some MCFs "refuse" to cluster with previously described members of subclades; their characterization presents therefore a special challenge but also a unique opportunity. Related to this work, Linka and Esser (2012) summarized the current knowledge on peroxisomal transport proteins, mainly members of the ABC transporter family known as COMATOSE (CTS) or ABCD1 shown to transport substrates for β-oxidation, and three members of the MCFs required for the import of ATP and NAD. Based on analyses of 22 fully sequenced plant genomes, evolution of these four peroxisomal transporters was an early event as they were already present in the plant genome before the cyanobacterial endosymbiont engulfment.

Carraro et al. (2012) provided a comprehensive analysis of the three major families of auxin transporters, PIN, AUX/LAX, and ABCB, from the model woody species *Populus*. They found that the array of more than 40 putative auxin transporters is reflected by a pre-existing diversity of *ABCB* genes and an expansion of *PIN* and *AUX/LAX* genes caused by genomic and segmental duplications. Interestingly, they provide further evidence that diversification of *PIN* and *AUX/LAX* genes (*Populus* owns exactly twice as many members of these families as *Arabidopsis* but a comparable number of *ABCB* genes) occurred after the origin of land plants, while the ABCB gene family diversified much earlier and long before the diversification of land plants.

Forestan et al. (2012) reanalyzed the *PIN* family in maize and confirmed that the monocot PIN family is wider and more diverged than the dicot one, though this cannot be explained simply by a gene duplication event. Interestingly, beside an increased number of Arabidopsis orthologs, maize contains three monocot-specific *PINs* (*PIN9* and *PIN10a-b*) that based on their gene expression, protein localization and auxin accumulation pattern can be partially associated with the differentiation and development of monocot-specific organs and tissues.

As more crystal structures of prokaryote and eukaryote transport proteins become available, it is now possible to examine the evolution and mechanism of plant transporters using homology modeling. Bailly et al. (2012) conducted an evolutionary analysis of ABCB substrate specificity by homology modeling of plant and mammalian ABCB1 proteins with strikingly dissimilar substrate specificities. Interestingly, they identified kingdom-specific candidate substrate binding regions within the translocation chamber formed by transmembrane domains of the ABCBs, suggesting an early evolutionary separation of plant and mammalian substrate specificity. Obviously an experimental validation of these models has highest priority. Pantoja (2012) reviewed recent mutagenesis studies to identify residues that form the ammonium transport pore and other important amino acids of AMT/MEP transporters. Zelman et al. (2012) looked at the functional domains of plant cyclic nucleotide-gated channels (CNGC) using homology modeling to get insights on the gating, the selectivity filter, and the heteromeric structure of plant CNGCs. Nussaume et al. (2011) reviewed the biological functions of plasma membrane-localized H^+-coupled phosphate symporters of the PHT1 family, as well as their transcriptional and post-translational regulation.

Significantly, articles in this special issue demonstrate the conservation and diversification of transporters as plants evolved. The emerging patterns will allow us to predict with relative confidence the molecular and biological roles of uncharacterized genes from early land plants to energy crops, and to build a systems biology understanding of plant development, growth and survival.

REFERENCES

Anderberg, H. I., Kjellbom, P., and Johanson, U. (2012). Annotation of *selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants. *Front. Plant Sci.* 3:33. doi: 10.3389/fpls.2012.00033

Bailly, A., Yang, H., Martinoia, E., Geisler, M., and Murphy, A. S. (2012). Plant lessons: exploring ABCB functionality through structural modeling. *Front. Plant Sci.* 2:108. doi: 10.3389/fpls.2011.00108

Carraro, N., Tisdale-Orr, T. E., Clouse, R. M., Knöller, A. S., and Spicer, R. (2012). Diversification and Expression of the PIN, AUX/LAX, and ABCB families of putative auxin transporters in populus. *Front. Plant Sci.* 3:17. doi: 10.3389/fpls.2012.00017

Chanroj, S., Wang, G., Venema, K., Zhang, M. W., Delwiche, C. F., and Sze, H. (2012). Conserved and diversified gene families of monovalent cation/H± antiporters from algae to flowering plants. *Front. Plant Sci.* 3:25. doi: 10.3389/fpls.2012.00025

Emery, L., Whelan, S., Hirschi, K. D., and Pittman, J. K. (2012). Protein phylogenetic analysis of Ca2+/cation antiporters and insights into their evolution in plants. *Front. Plant Sci.* 3:1. doi: 10.3389/fpls.2012.00001

Facchinelli, F., and Weber, A. P. M. (2011). The metabolite transporters of the plastid envelope: an update. *Front. Plant Sci.* 2:50. doi: 10.3389/fpls.2011.00050

Forestan, C., Farinati, S., and Varotto, S. (2012). The maize PIN gene family of auxin transporters. *Front. Plant Sci.* 3:16. doi: 10.3389/fpls.2012.00016

Haferkamp, I., and Schmitz-Esser, S. (2012). The plant mitochondrial carrier family: functional and evolutionary aspects. *Front. Plant Sci.* 3:2. doi: 10.3389/fpls.2012.00002

Jelesko, J. G. (2012). An expanding role for purine uptake permease-like transporters in plant secondary metabolism. *Front. Plant Sci.* 3:78. doi: 10.3389/fpls.2012.00078

Linka, N., and Esser, C. (2012). Transport proteins regulate the flux of metabolites and cofactors across the membrane of plant peroxisomes. *Front. Plant Sci.* 3:3. doi: 10.3389/fpls.2012.00003

Nussaume, L., Kanno, S., Javot, H., Marin, E., Pochon, N., Ayadi, A., et al. (2011). Phosphate import in plants: focus on the PHT1 transporters. *Front. Plant Sci.* 3:83. doi: 10.3389/fpls.2011.00083

Pantoja, O. (2012). High affinity ammonium transporters: molecular mechanism of action. *Front. Plant Sci.* 3:34. doi: 10.3389/fpls.2012.00034

Pedersen, C. N. S., Axelsen, K. B., Harper, J. F., and Palmgren, M. G. (2012). Evolution of plant P-type ATPases. *Front. Plant Sci.* 3:31. doi: 10.3389/fpls.2012.00031

Reinders, A., Sivitz, A. B., and Ward, J. M. (2012). Evolution of plant sucrose uptake transporters (SUTs). *Front. Plant Sci.* 3:22. doi: 10.3389/fpls.2012.00022

Spetea, C., Pfeil, B. E., and Schoefs, B. (2012). Phylogenetic analysis of the thylakoid ATP/ADP carrier reveals new insights into its function restricted to green plants. *Front. Plant Sci.* 2:110. doi: 10.3389/fpls.2011.00110

Takahashi, H., Buchner, P., Yoshimoto, N., Hawkesford, M. J., and Han Shiu, S. (2012). Evolutionary relationships and functional diversity of plant sulfate transporters. *Front. Plant Sci.* 2:119. doi: 10.3389/fpls.2011.00119

Tegeder, M., and Ward, J. M. (2012). Molecular evolution of plant AAP and LHT amino acid transporters. *Front. Plant Sci.* 3:21. doi: 10.3389/fpls.2012.00021

Zelman, A. K., Dawe, A., Gehring, C., and Berkowitz, G. A. (2012). Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels. *Front. Plant Sci.* 3:95. doi: 10.3389/fpls.2012.00095

Received: 15 November 2013; accepted: 15 December 2013; published online: 08 January 2014.

Citation: Sze H, Geisler M and Murphy AS (2014) Linking the evolution of plant transporters to their functions. Front. Plant Sci. ***4****:547. doi: 10.3389/fpls.2013.00547*

This article was submitted to Plant Traffic and Transport, a section of the journal Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 21 February 2012
doi: 10.3389/fpls.2012.00031

Evolution of plant P-type ATPases

Christian N. S. Pedersen[1,2], Kristian B. Axelsen[3,4], Jeffrey F. Harper[5] and Michael G. Palmgren[3]*

[1] *Center for Membrane Pumps in Cells and Disease – PUMPKIN, Danish National Research Foundation, Aarhus University, Aarhus, Denmark*
[2] *Bioinformatics Research Centre (BiRC), Faculty of Science and Technology, Aarhus University, Aarhus, Denmark*
[3] *Department of Plant Biology and Biotechnology, Faculty of Life Sciences, Center for Membrane Pumps in Cells and Disease – PUMPKIN, Danish National Research Foundation, University of Copenhagen, Frederiksberg C, Denmark*
[4] *Swiss-Prot Group, Swiss Institute of Bioinformatics, Geneva, Switzerland*
[5] *Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV, USA*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Jin Chen, Michigan State University, USA
Manuel González-Guerrero, Universidad Politécnica de Madrid, Spain

****Correspondence:***
Michael G. Palmgren, Department of Plant Biology and Biotechnology, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark.
e-mail: palmgren@life.ku.dk

Five organisms having completely sequenced genomes and belonging to all major branches of green plants (Viridiplantae) were analyzed with respect to their content of P-type ATPases encoding genes. These were the chlorophytes *Ostreococcus tauri* and *Chlamydomonas reinhardtii*, and the streptophytes *Physcomitrella patens* (a non-vascular moss), *Selaginella moellendorffii* (a primitive vascular plant), and *Arabidopsis thaliana* (a model flowering plant). Each organism contained sequences for all five subfamilies of P-type ATPases. Whereas Na^+ and H^+ pumps seem to mutually exclude each other in flowering plants and animals, they co-exist in chlorophytes, which show representatives for two kinds of Na^+ pumps (P2C and P2D ATPases) as well as a primitive H^+-ATPase. Both Na^+ and H^+ pumps also co-exist in the moss *P. patens*, which has a P2D Na^+-ATPase. In contrast to the primitive H^+-ATPases in chlorophytes and *P. patens*, the H^+-ATPases from vascular plants all have a large C-terminal regulatory domain as well as a conserved Arg in transmembrane segment 5 that is predicted to function as part of a backflow protection mechanism. Together these features are predicted to enable H^+ pumps in vascular plants to create large electrochemical gradients that can be modulated in response to diverse physiological cues. The complete inventory of P-type ATPases in the major branches of Viridiplantae is an important starting point for elucidating the evolution in plants of these important pumps.

Keywords: evolution, P-type ATPases, plants, salt tolerance, Na^+ pumps

INTRODUCTION

P-type ATPases are primary transporters energized by hydrolysis of ATP with a wide range of specificities for small cations and apparently also phospholipids (Møller et al., 1996; Palmgren and Harper, 1999). P-type ATPases are characterized by forming a phosphorylated intermediate (hence the name P-type), by being inhibited by vanadate, and by having a number of sequence motifs in common (Serrano, 1989; Axelsen and Palmgren, 1998). Plant P-type ATPases are characterized structurally by having a single catalytic subunit, 8–12 transmembrane segments, N and C termini exposed to the cytoplasm, and a large central cytoplasmic domain including the phosphorylation and ATP binding sites.

P-type ATPases constitute a large and indispensable family in most organisms. The P-type ATPase family can be divided into five major evolutionarily related subfamilies, P1–P5, which group in a phylogenetic tree according to the ions they transport (Axelsen and Palmgren, 1998). The P-type ATPases are involved in a wide range of fundamental cellular processes such as the efflux or organismal redistribution of micronutrients (P1B Zn^{2+}- and Cu^{2+}-ATPases), cellular signaling and Ca^{2+} compartmentalization (P2A and P2B Ca^{2+}-ATPases), energizing the electrochemical gradient used as the driving force for the secondary transporters (P3A H^+-ATPases in plants and fungi and P2C Na^+/K^+-ATPases in animals), and being involved in membrane vesicle budding (P4 ATPases). The function of P5 ATPases is not known but they have been implicated in vesicle budding from the endoplasmic reticulum (Poulsen et al., 2008a).

The bioenergetic systems of flowering plants and animals use different coupling ions (Skulachev, 1988; Rodríguez-Navarro, 2000). In animals, the potential energy that can be harvested to drive, e.g., nutrient transport across the plasma membrane derives from a Na^+ gradient across this membrane. A Na^+ pump in the plasma membrane is a very efficient system for extrusion of toxic Na^+ (Gonzalez, 2011; Whittamore, 2012). A plasma membrane Na^+ pump in the plasma membrane is also important for the extrusion of toxic Na^+. Fish and invertebrates living in the salty oceans are dependent on such a pump for survival. For animals living in marine environments a Na^+/K^+ pumps (P2C ATPases) remain the sole system for formation of electrochemical ion gradients gradients in the plasma membrane (Morth et al., 2011).

In contrast, in plant and fungal cells, it is an electrochemical gradient of H^+ that energizes the plasma membrane (Morth et al., 2011). In cells of flowering plants, the plasma membrane completely lacks Na^+ pumps and depends entirely on plasma membrane H^+ pumps for establishing a steep electrochemical ion gradient across the plasma membrane. In those flowering plants in which salt tolerance mechanisms have been investigated, strategies other than primary active extrusion of Na^+ (i.e., efflux through a Na^+ pump) appear to have evolved (Tester and Davenport, 2003; Flowers and Colmer, 2008). Unfortunately, most flowering plants

remain very sensitive to saline environments and salinization of soils due to extensive irrigation is becoming an increasing problem for productivity of agriculture, especially in arid regions of the world (Tester and Davenport, 2003; Yamaguchi and Blumwald, 2005).

The apparent absence of Na^+ pumps in flowering plants raises several questions. In ancestral plants containing both Na^+ and H^+ pumps, did both pumps help energize the plasma membrane, and did those membranes function with co-transport systems that could utilize both Na^+ and H^+ gradients? During the evolution of flowering plants, was there a physiological reason why plasma membranes might not function well with both Na^+ and H^+ pumps? Did the flowering plant lineage evolve from an ancestor living exclusively in a fresh water environment, or were Na^+ pumps gradually lost at later points in evolution? Is it possible that Na^+ pumps are still present in some flowering plants that have not yet been studied?

With the completion of genome sequences from all major branches of Viridiplantae, it is now possible to study the evolution of primary transport capabilities throughout the green plant lineage and to compare these transporters with those of protists, animals, and fungi. Our analysis of five selected plant genomes suggest that Na^+ pumps coexisted with primitive plasma membrane H^+ pumps in both early aquatic and terrestrial plants. However, at some point in the evolution of vascular plants, Na^+ pumps appear to have been lost. The evolution of multicellular land plants also correlates with an expansion in number and potential regulatory features for H^+ pumps, Ca^{2+} pumps, heavy metal pumps, and lipid flippases. In contrast, the P5 subfamily did not show an equivalent change, with only one or two isoforms present in all five plant lineages. This study provides an evolutionary framework for considering how P-type ATPases contribute to the biology of all Viridiplantae, from single celled algae to multicellular flowering plants.

MATERIALS AND METHODS

The P-type ATPase sequences were identified by searching for sequences in UniProtKB (The UniProt Consortium, 2012) and in Phytozome[1] from the relevant genomes that matched the PFAM profile PF00122 and in UniProtKB also the PROSITE pattern PS00154 (DKTG[T,S][L,I,V,M][T,I]; Axelsen and Palmgren, 1998). The PROSITE pattern is unique for P-type ATPases while the PFAM profile is more inclusive and also matches sequences that cannot function as active P-type ATPases as they lack the phosphorylated Asp residue that is present in the PROSITE pattern. Since the sequences in Phytozome do not include information about matches to PROSITE patterns, only PFAM was used to select sequences from this resource. The identified sequences were aligned and duplicates were removed. All genomes except *Arabidopsis* are still draft versions in the databases, so often the predicted P-type ATPases do not represent complete proteins. Furthermore, the versions identified in UniProtKB and in Phytozome were not 100% identical. When possible, the UniProtKB sequence was chosen.

[1]http://www.phytozome.net/

The resulting dataset of 150 sequences were aligned using Muscle (Edgar, 2004). The resulting alignment of the 150 full-length sequences was used to construct a phylogenetic tree using the Neighbor Joining method as implemented in QuickTree (Howe et al., 2002). The standard parameters of Muscle and QuickTree were used. For visualization of the constructed tree we used Dendroscope (Huson et al., 2007). A phylogenetic tree for each subgroup (P1–P5) of the 150 sequences was constructed and visualized similarly. For each subgroup, the corresponding set of sequences was selected from the dataset of 150 sequences. The selected sequences and a few additional outlier sequences were aligned using Muscle and a phylogenetic tree was constructed using QuickTree. The trees for the individual subfamilies are all rooted with the sequence of the *E. coli* P1A ATPase KdpB (P03960) as outgroup. The 150 sequences as well as sequences for the outliers employed are available as Fasta files in Supplementary Information.

RESULTS

In this work we identified genes encoding P-type ATPases in the sequenced genomes from five representatives of the green plant lineage (Viridiplantae), which previously diverged from opisthokonts (animals, fungi, and Choanozoa; Yoon et al., 2004; **Figure 1**). The Chlorophytae (green algae, including *Chlamydomonas* and *Ostreococcus*) diverged from the Streptophytae (land plants and their close relatives; **Figure 1**) over a billion years ago. In this work, the genomes of two green algae were analyzed, namely those of *Ostreococcus tauri* (Derelle et al., 2006) and *Chlamydomonas reinhardtii* (Merchant et al., 2007), and three land plants, the moss *Physcomitrella patens* (Rensing et al., 2008), the primitive vascular plant *Selaginella moellendorffii*[2] and the flowering plant *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000). The latter has previously been investigated for its content of P-type ATPases (Axelsen and Palmgren, 2001).

Ostreococcus tauri is an extremely small (0.8 μm wide) unicellular green alga, which belongs to the Prasinophyceae, one of the

[2]http://genome.jgipsf.org/Selmo1/Selmo1.home.html

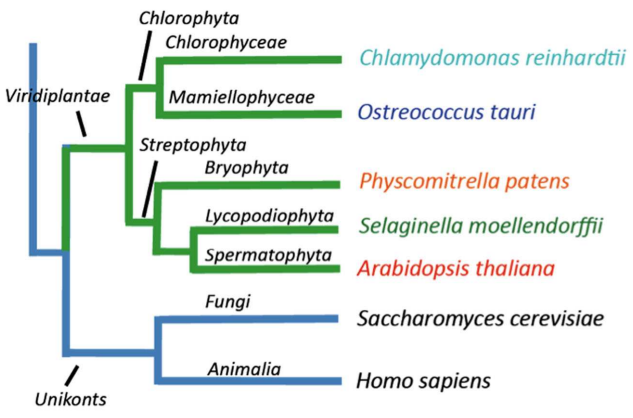


FIGURE 1 | Overview of the evolutionary relationship between the green plant genomes analyzed in this study.

most ancient groups within the lineage of green plants (Courties et al., 1998). This organism is a naked, non-flagellated cell possessing a single mitochondrion and a single chloroplast, and a common member of global oceanic picoplankton populations. *C. reinhardtii* is a much bigger (∼10 μm), likewise unicellular green alga. It lives in terrestrial soils and has multiple mitochondria, two anterior flagella for motility and mating, and a chloroplast (Rochaix, 1995). *P. patens* is a moss, a representative for primitive land plants without a vascular system. *S. moellendorffii* is a member of an ancient vascular plant lineage that first appears in the fossil record about 400 million years ago (Banks, 2009). *A. thaliana* was the first angiosperm to have its genome sequenced, and is a model plant for understanding the molecular biology of flowering plants.

Figure 2 shows the phylogenetic relationship of all the 150 sequences listed in **Table 2**, which were evaluated in this study. They are distributed in all five major families of P-type ATPases with some noteworthy comparisons discussed below.

P1B ATPases: HEAVY METAL PUMPS

Heavy metal pumps (P1B ATPases) are found in all life forms including bacteria. In eukaryotes, these pumps are typically encoded by multigene families (Axelsen and Palmgren, 1998). The model dicotyledonous plant *A. thaliana* contains eight P1B ATPases, which can be divided into three groups according to conserved sequence motifs and their putative substrate specificity (Axelsen and Palmgren, 2001; Williams and Mills, 2005; Argüello et al., 2007). AtHMA5 to 8 belong to group P1B-1 and are predicted to transport Cu/Ag, while AtHMA2 to 4 belong to group P1B-2 predicted to transport Zn/Cd, and AtHMA1 belongs to group P1B-4 with a predicted broad substrate specificity (Zn/Cu/Co/Cd/Pb/Ca).

P1B ATPases were found in all organisms investigated in this study (**Figure 3**). The genome of *C. reinhardtii* encodes five P1B ATPases, four of which (CrHMA2-5) belong to the P1B-1 cluster. A representative protein in *A. thaliana* from this cluster is AtHMA8/PAA2. This protein localizes to the thylakoid membrane of chloroplasts and is required for Cu delivery during the biogenesis of plastocyanin (Abdel-Ghany et al., 2005).

The only heavy metal ATPase in *C. reinhardtii* that is not a P1B-1 pump, is CrHMA1, which belongs to the P1B-4 cluster in which AtHMA1 is also found. AtHMA1 is localized to the chloroplast inner envelope membrane (Seigneurin-Berny et al., 2006; Kim et al., 2009). This ATPase has been reported to function as a transporter for Cu and Ca in addition to Zn/Cd/Co (Seigneurin-Berny et al., 2006; Moreno et al., 2008; Kim et al., 2009). All five organisms investigated in this study had one or two representatives of P1B-4 ATPases.

Zn transporting P1B ATPases are common in prokaryotes but have not been identified in fungi and animals. In *O. tauri*, which has one of the smallest known eukaryotic genomes, there are five P1B ATPases (∼30% of its P-type ATPase genes), which

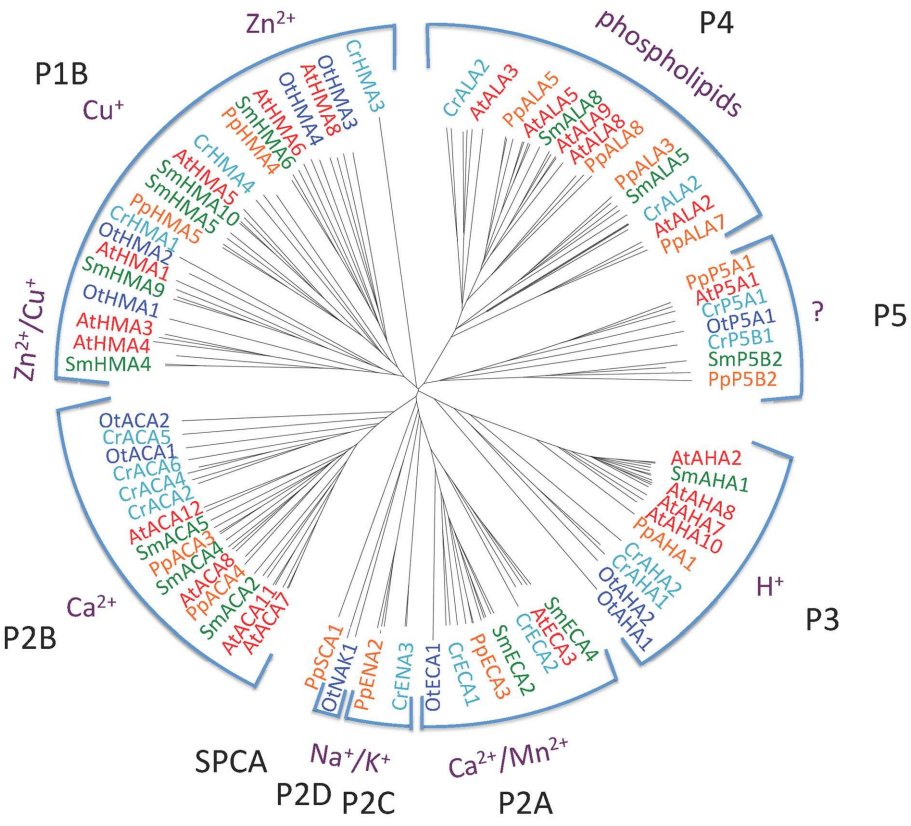
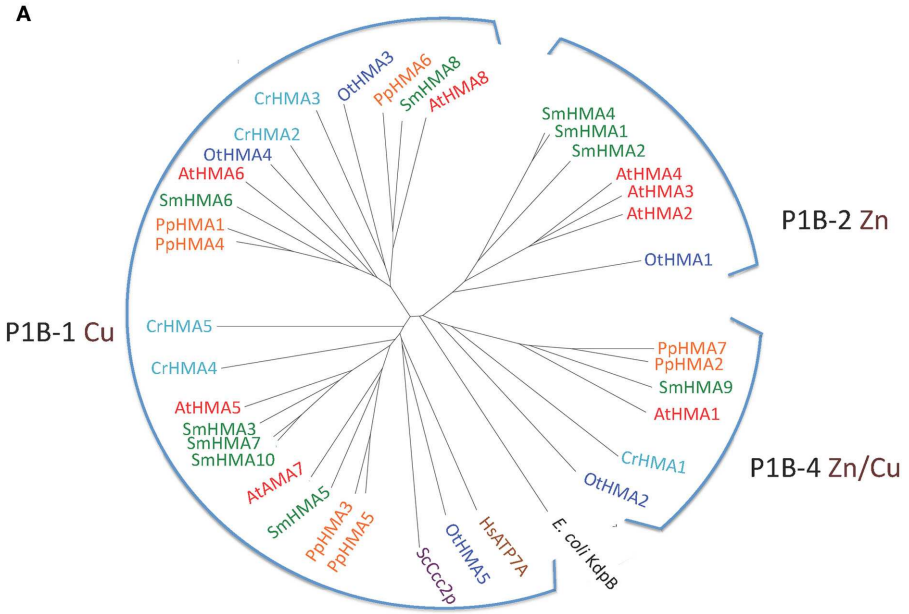


FIGURE 2 | Phylogenetic tree of P-type ATPases analyzed in this study. Not all branches are labeled. Accession numbers for sequences are given in **Table 2**.

A



B

	Organism				
Class	*O. tauri*	*C. reinhardtii*	*P. patens*	*S. moellend.*	*A. thaliana*
P1B-1	3	4	5	6	4
P1B-2	1	0	0	3	3
P1B-4	1	1	2	1	1

FIGURE 3 | (A) Phylogenetic tree of P1B ATPases (heavy metal pumps) from Viridiplantae. Accession numbers for sequences are given in **Table 2**. For comparison, the following outliers were included: a bacterial P1A ATPase *E. coli* KdpB (P03960), and HsATP7A (Q04656), and ScCcc2p (P38995), Cu^+ pumps from *H. sapiens* and *S. cerevisiae*, respectively. **(B)** Numbers of P1B ATPases by subgroups (Argüello et al., 2007) in the five plant genomes analyzed.

distribute among all three clusters of P1B ATPases, including a single P1B-2 pump. In our analysis, P1B-2 ATPases were found in *O. tauri* (OtHMA1), *S. moellendorffii* (SmHMA1, SmHMA2, and SmHMA4), and *A. thaliana* (AtHMA2, AtHMA3, and AtHMA4; **Figures 3** and **4**). P1B-2 ATPases could not be identified in *C. reinhardtii* and *P. patens*. Among the *A. thaliana* P1B-2 ATPases, AtHMA2, and AtHMA4 are localized to the plasma membrane. They show redundant function in cellular export of Zn and Cd into plant vascular tissues where they facilitate xylem loading and transport to the shoot (Hussain et al., 2004; Verret et al., 2004; Mills et al., 2003; Wong and Cobbett, 2009).

As P1B-2 Zn ATPases are absent from fungi and animals, but common in bacteria, it seems likely that in plants these pumps have evolved from chloroplastic pumps. With the advent of vascular plants, it is likely that a subset of P1B-2 ATPases was targeted to the plasma membrane and acquired a new role in redistribution of Zn within the plant body.

P2A ATPases: ER-TYPE Ca^{2+}-ATPASES

P2 ATPases form a large subfamily further divided into at least four clusters of pumps, two (P2A and P2B) having specificity for Ca^{2+} and two (P2C and P2D) for Na^+ as the transported ligand (**Figure 5**).

P2A ATPases were identified in all plants studied here (**Figure 5**). A single pump from each organism (OtECA2, CrECA4, PpECA4, SmECA4, and AtECA3) form a distinct subset of ER-type Ca^{2+}-ATPases (ECAs) closely related to the animal sarco-endoplasmic reticulum Ca^{2+} pump SERCA1 (**Figure 5**). In the flowering plant *A. thaliana*, AtECA3 is a Golgi-localized pump that can transport Ca^{2+} and Mn^{2+} (Li et al., 2008; Mills et al., 2008).

In a comparison among many eukaryotes, a subset of P2A Ca^{2+}-ATPases form a distinct cluster and have been named secretory pathway Ca^{2+}-ATPases (SPCAs; Wuytack et al., 2003). These pumps are identified in fungal and animal cells and are localized to the Golgi apparatus or other membranes of the secretory pathway. As evident from **Figure 6**, representatives of these pumps from fungi (*S. cerevisiae* and *S. pombe*) are characterized by having lost Ca^{2+} binding site 1.

Only a single likely SPCA protein was found in our analysis, namely PpSCA1. This protein clusters with other SPCAs and does not contain conserved residues expected for Ca^{2+} binding site 1 (**Figure 6A**). Among reference plant and animal genomes surveyed, PpSCA1 showed the greatest identity (30%) to the secretory pathway Ca^{2+}-ATPase Pmr1p from *S. cerevisiae* (Rudolph et al., 1989; Antebi and Fink, 1992).

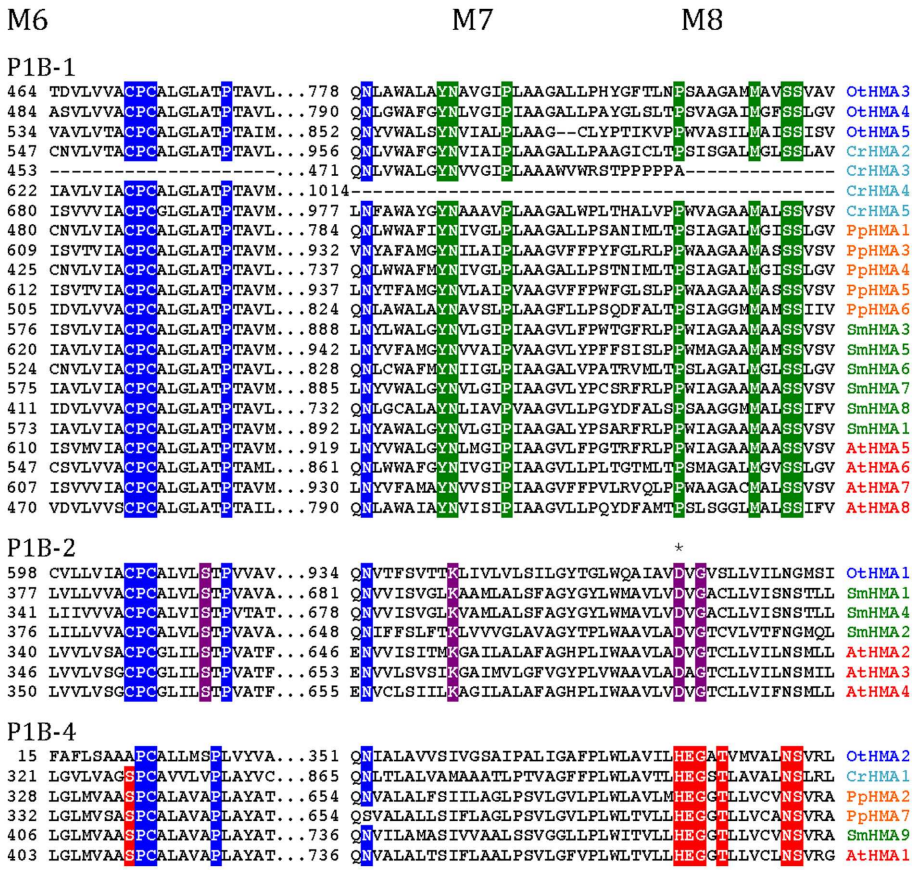


FIGURE 4 | Alignment of three transmembrane segments from Viridiplantae P1B ATPases analyzed in this study. Only the predicted transmembrane segments M6, M7, and M8 are shown. P1B-1: Cu^+ transporting ATPases; P1B-2: Zn^{2+} transporting ATPase: P1B-4: Mixed specificity heavy metal pumps. The *asterisk* indicates an Asp residue (D) conserved in P1B-2 ATPases, which could be important for coordination of Zn^{2+} (Dutta et al., 2006).

P2B ATPases: AUTOINHIBITED Ca^{2+}-ATPases

P2B ATPases are Ca^{2+} pumps that are activated by binding of calmodulin to autoinhibitory terminal domains. A marked difference between higher plant and animal P2B ATPases is that their calmodulin-binding domains (CMBDs) are situated in the N- and C-terminal domains, respectively (Sze et al., 2000). Autoinhibited Ca^{2+}-ATPases (ACAs) from flowering plants have been shown to be activated by Ca^{2+} in the presence of calmodulin and are characterized by an N-terminally situated CMBD (Malmström et al., 1997; Harper et al., 1998; Curran et al., 2000; Hwang et al., 2000). The CMBD overlaps partially with an autoinhibitory pump sequence (Bækgaard et al., 2006) and it has been proposed that calmodulin, by binding to the CMBD, neutralizes the constraint set by the autoinhibitory sequence on the pump molecule (Bækgaard et al., 2006). Animal P2B ATPases are likewise activated by calmodulin, but in these the CMBD is situated in an extended C-terminal domain (James et al., 1988).

Putative CMBDs were identified in the N-terminal domains of most P2B ATPases of *P. patens*, *S. moellendorffii*, and *A. thaliana* (**Figure 7**). The proposed CMBDs in AtACA12 and AtACA13 are weakly defined. Importantly, an N-terminal CMBD could not be identified in P2B ATPases from the chlorophytes *O. tauri* and *C. reinhardtii*.

As animal P2B ATPases are equipped with a C-terminally located CMBD, we analyzed all of the plant P2B ATPases for a putative regulatory domain in this location. CMBDs have very little similarity between calmodulin-binding proteins and can be difficult or impossible to predict with certainty although as a rule there is alternation between bulky aromatic and positively charged residues. No C-terminal extensions could be identified in Streptophyte P2B ATPases. In contrast, a conserved sequence with weak resemblance of a CMBD was identified in the C-terminal domain of P2B ATPases from the chlorophytes *O. tauri* and *C. reinhardtii* (**Figure 7**), which are missing a similar sequence from their N-terminal domain. The calmodulin-binding capacity of these putative CMBDs remains to be tested, but it is an attractive working hypothesis that swapping of the CMBD from the C-terminal to the N-terminal domain occurred at the split between Chlorophytae and Streptophytae.

P2C ATPases: Na^+/K^+-ATPases

The Na^+/K^+-ATPase of animal cells was the first P-type ATPase to be discovered (Skou, 1957), but all subsequent attempts to identify

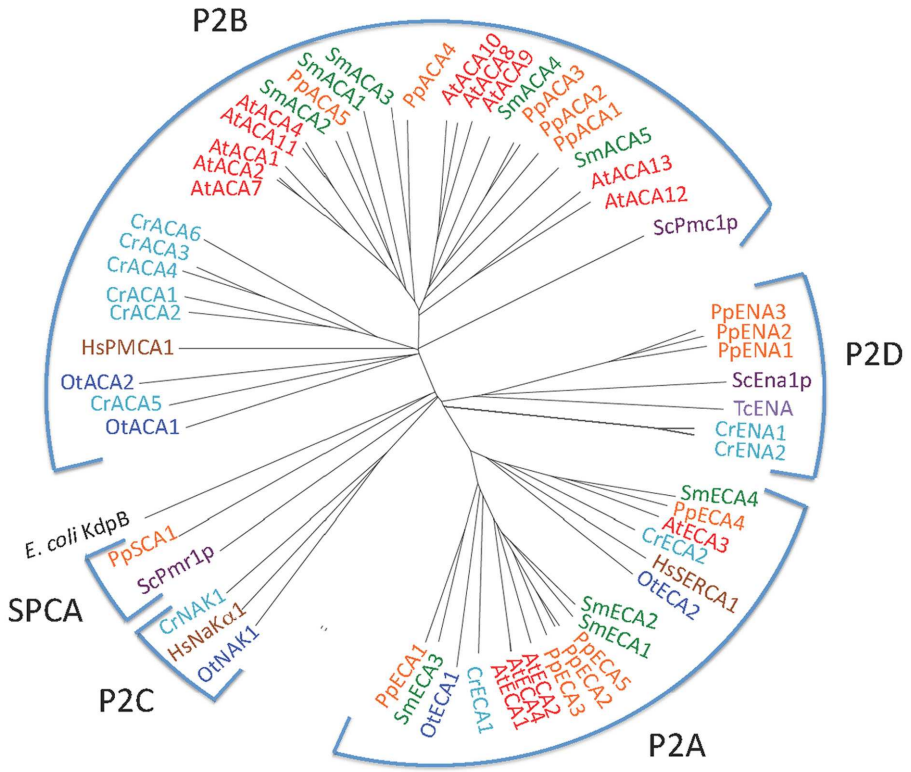


FIGURE 5 | Phylogenetic tree of P2 ATPases (Ca^{2+} and Na^+ pumps). Accession numbers for sequences are given in **Table 2**. For comparison, the following outliers were included: a bacterial P1A ATPase *E. coli* KdpB (P03960), the *S. cerevisiae* pumps ScPmr1p (P13586), ScPmc1p (P38929), and ScEna1p (P13857), the *H. sapiens* pumps HsPMCA1 (P20200), HsSERCA1 (O14983), and HsNaKα1 (P98194), and the *Trypanosoma cruzi* pump TcENA (Q76DT8).

a Na^+ pump in vascular plants failed. Vascular plants tend to be very sensitive to elevated Na^+ in the soil, which is probably due to the lack of an effective Na^+ extrusion system such as a Na^+/K^+-ATPase. In animal cells, the plasma membrane is energized by the Na^+/K^+-ATPase whereas in a typical plant, the plasma membrane H^+-ATPases (P3A ATPases) carry out this function. It has therefore been hypothesized that P2C Na^+/K^+-ATPases were lost at a branch point in the Streptophyta plant lineage, presumably in an organism that evolved in a fresh water environment and utilized a plasma membrane H^+-ATPase to energize its plasma membrane (Palmgren, 2001).

Biochemical evidence has pointed to the presence an electrogenic, vanadate-sensitive, ouabain-resistant Na^+-ATPase in the plasma membrane of marine chlorophytes (Popova et al., 1999, 2005; Gimmler, 2000) and chlorophyte expressed sequence tags with similarity to Na^+/K^+-ATPase have been identified (Barrero-Gil et al., 2005). Both chlorophytes analyzed in this work contain sequences with strong similarity to an animal Na^+/K^+-ATPase (**Figure 8**). No Na^+/K^+-ATPase has been crystallized in a form with bound Na^+, but a homology model has been built (Morth et al., 2011) based on the structure of a Na^+/K^+-ATPase with bound Rb^+ (as a substitute for K^+) and the structure of SERCA1 with bound Ca^{2+}. According to this model, residues in several transmembrane segments of Na^+/K^+-ATPase contribute with binding ligands to Na^+ (**Figure 8**). All these residues are conserved in OtNAK1 and CrNAK1, strongly suggesting that these ATPases operate as Na^+ pumps. *O. tauri* is a chlorophyte that lives in oceans, where a Na^+/K^+-ATPase is of obvious benefit for extrusion of Na^+ leaking in from sea water. However, it is peculiar that *C. reinhardtii*, a green alga of terrestrial soils, is also equipped with such a pump. This suggests that the presence of Na^+/K^+ pumps in green algae is a primitive character that was lost with the emergence of Streptophytae. This hypothesis is supported by the widespread presence of P2C pumps in other eukaryotes and archaea (Sáez et al., 2009).

As mentioned above, the substrate specificity of PpSCA1 is uncertain, with features that are both consistent and contrary to speculations on the transport of either Na^+ or Ca^{2+}. This pump does have some similarity with chlorophyte Na^+/K^+-ATPases, but ligands contributing to Na^+ Site 3 are not present and similarity to sites contributing to sites 1 and 2 are not absolute (**Figure 6A**). Biochemical and genetic experiments are needed to address the question of substrate specificity for PpSCA1.

P2D ATPases: Na^+ OR K^+ PUMPS OF MOSSES AND FUNGI

P2D ATPases form a unique group of pumps so far only found in mosses, fungi, and protozoa and confer Na^+ tolerance to organisms in which they are expressed (Rodríguez-Navarro and Benito, 2010). The moss *P. patens* encodes three P2D ATPases: PpENA1, PpENA2, and PpENA3. Related pumps are present in liverworts (Marchantiophyta), such as *Marchantia polymorpha* and *Riccia fluitans*, which are primitive non-vascular land plants related to

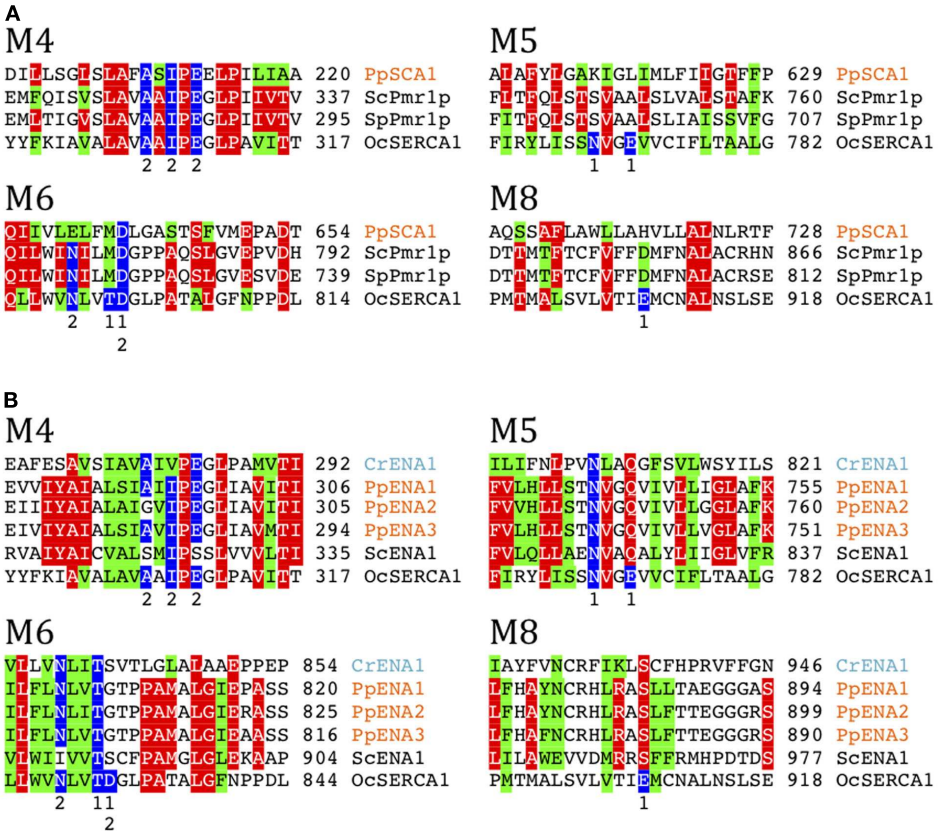


FIGURE 6 | Alignment of transmembrane segments showing differences in a potential cation binding site for selected Ca^{2+} and Na^{+} pumps. Secretory pathway Ca^{2+}-ATPases (SPCA pumps) and ENA Na^{+} pumps (P2D ATPases) are missing Ca^{2+} binding site 1 present in P2A pumps typified by rabbit OcSERCA1 (P04191). In P2A pumps a conserved Asp in M6 contributes to coordination of both Ca^{2+} ions. **(A)** Secretory pathway pumps (SPCAs) have a conserved Asp in M6. Fungal examples shown are *S. cerevisiae* Pmr1p (P13586) and *S. pombe* Pmr1p (O59868). **(B)** P2D ATPases (ENA pumps) do not have a conserved Asp in M6 like all other P2 ATPases. Residues contributing with Ca^{2+} coordinating oxygen molecules in rabbit SERCA1 (Toyoshima et al., 2000) are marked in blue. Only sequences including the predicted transmembrane segments M4, M5, M6, and M8 are shown.

mosses. A close fungal homolog to PpENA1 is from *Neurospora crassa* (Q9UUX8; 43% identity). When protein databases were searched for similar proteins outside plants and fungi, hits were only found in protist sequences.

In P2A ATPases, two Ca^{2+} sites are present: Site 1 and Site 2 (Toyoshima et al., 2000). In SPCAs (**Figure 6A**) and P2D ATPases (**Figure 6B**), Site 1 is missing. Thus, two negatively charged Glu residues in M6 and M8, which in SERCA1 contribute to Ca^{2+} coordination in Site 1, are absent in both secretory pathway pumps and P2D pumps. Notably, a negatively charged Asp in M6, which is conserved in all other P2-type ATPases, is replaced by a neutral residue in P2D pumps (Rodríguez-Navarro and Benito, 2010; **Figure 6B**). In the available structures of P2A, P2C, and P3A ATPases (Morth et al., 2007; Olesen et al., 2007; Pedersen et al., 2007), this Asp contributes to coordination of all transported cations including Ca^{2+}, Na^{+}, K^{+}, and H^{+}. The absence of the Asp in M6, in addition to other negatively charged amino acid residues in the membrane domain, therefore appears to be a hallmark of P2D ATPases. When the sequences of P-type ATPases retrieved in this study were analyzed in detail it appeared that two pumps from *C. reinhardtii*, here named CrENA1 and CrENA2, are likely to represent chlorophyte P2D ATPases as they lack a negatively charged Asp in M6 (**Figure 6B**). Although these branch out close to P2D pumps (**Figure 5**) they have less than 33% identity to these or any other plant pumps.

P3A ATPases: AUTOINHIBITED H^{+}-ATPases

We identified P3A ATPases in all genomes of Viridiplantae analyzed in this work (**Figure 9**). P3A ATPases energize the plasma membrane of plants and fungi by establishing a large proton gradient and membrane potential (negative on the inside) across the plasma membrane (Palmgren, 2001). The potential energy stored in this gradient serves as a proton motive force that drives a large number of transport processes carried out by secondary active transporters and channel proteins. The plasma membrane proton pump isoform 2 from *A. thaliana* (autoinhibited H^{+}-ATPase 2, AHA2) is expressed throughout the plant and, together with the closely related isoform AHA1, is essential for plant growth (Palmgren, 2001; Haruta et al., 2010). In this respect, they serve as functional analogs to the Na^{+}/K^{+}-ATPases of animal cells (Morth et al., 2011).

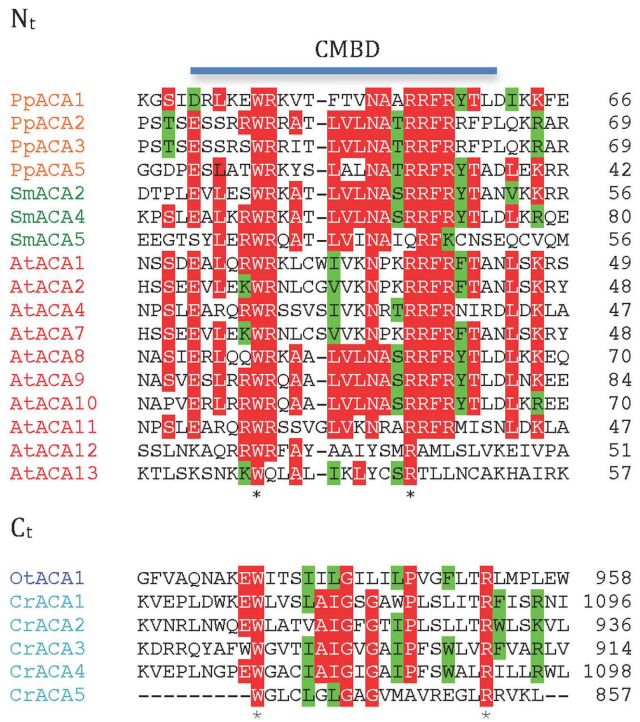


FIGURE 7 | Alignment of putative calmodulin-binding sites in the N- and C-terminal regions of ACA pumps. Nt, N-terminal domain; Ct, C-terminal domain.

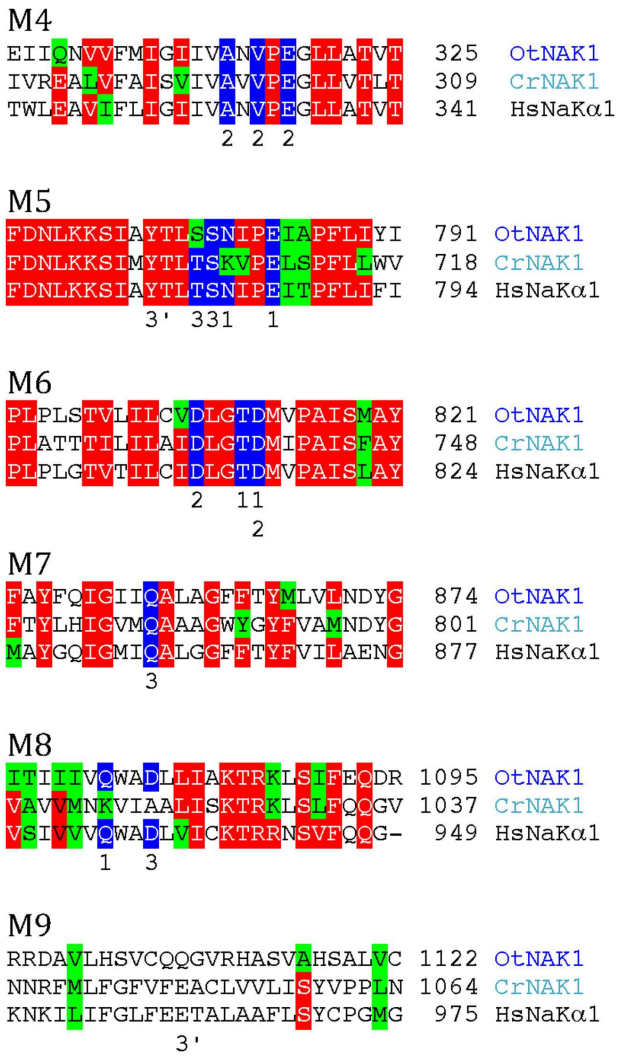


FIGURE 8 | Alignment of predicted transmembrane segments of putative chlorophyte Na^+/K^+-ATPases (OtNAK1 and CrNAK1) with similar regions in the human Na^+/K^+-ATPase α1 subunit (ATP1A1; P05023). Residues that in Na^+/K^+-ATPase are likely to contribute with oxygen atoms for coordination of Na^+ (Morth et al., 2007) are marked by blue. Numbers above these residues refer to which of the three Na^+ sites (1–3) they contribute to. Residues contributing to an alternative Na^+ site (site 3b) are marked 3′.

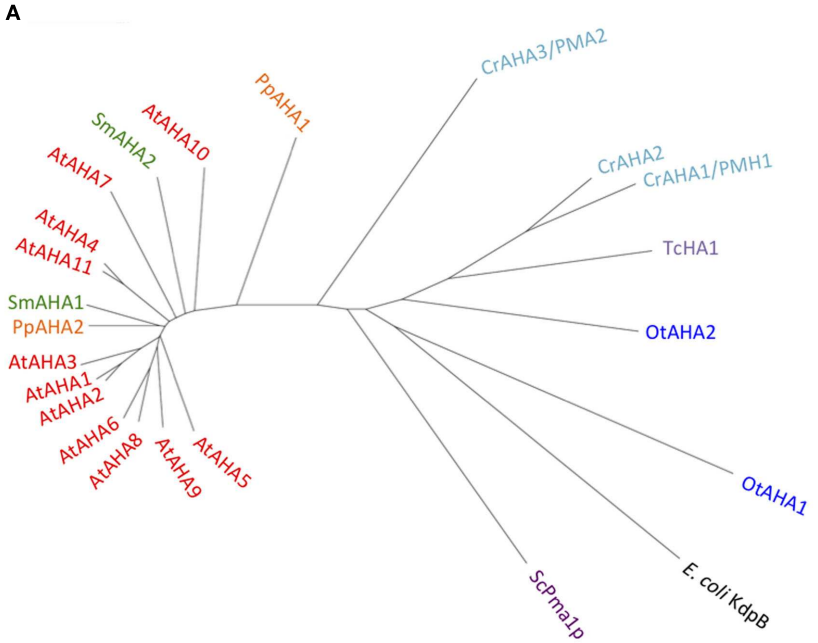
In *Chlamydomonas*, two plasma membrane H^+-ATPases have been described in the literature (Campbell et al., 2001). Closely related homologs to OtAHA2 are found in other green algae but not in streptophytes. Outside this group, OtAHA2 has highest similarity (42–44% identity) to plasma membrane H^+-ATPases of protists that have been characterized biochemically as P-type H^+-ATPases (Luo et al., 2002, 2006; **Figure 9**). OtAHA2 has lesser but marked similarity to AtAHA2 (38% identity). The most divergent P3A ATPase analyzed in this study is OtAHA1.

All P3A plasma membrane H^+-ATPases have conserved residues that have been implicated as being important for H^+ transport (Pedersen et al., 2007; **Figure 10**). These include the H^+ acceptor/donor Asp684 (AtAHA2 numbering) in M6 and the proposed gate-keeper residue Asn106 in M2 (Pedersen et al., 2007; Buch-Pedersen et al., 2009). Arg655, which in AtAHA2 has been proposed to prevent backflow of H^+ through the pump, a feature likely to be essential when electrochemical gradients get steep, is strictly conserved in all streptophyte pumps. Interestingly, this residue is absent in chlorophyte P3A ATPases except for CrAHA3 (**Figure 10**). CrAHA3 is the chlorophyte pump that shows the highest similarity to a streptophyte H^+ pump (**Figure 9**).

Notably, typical protists and chlorophytes are characterized by having both Na^+ and H^+ pumps. Protists are equipped with P2D Na^+ pumps and chlorophytes with P2C Na^+/K^+ pumps (**Table 1**; **Figures 8** and **9**). As plasma membrane H^+-ATPases in these organisms lack the residue corresponding to Arg655 (to block H^+ backflow), this would suggest that in organisms with co-expression of electrogenic H^+ and Na^+ pumps, it is the role of Na^+ ATPases to generate a plasma membrane electrochemical gradient, which in turn can be used as an energy source to drive a variety of cellular processes, such as secondary active transport. P3A ATPases in protists and chlorophytes might therefore have other roles than establishing electrochemical gradients, e.g., controlling intracellular pH.

Angiosperm plasma membrane H^+-ATPases are regulated by an extended C-terminal domain that functions as a pump auto inhibitor (Palmgren et al., 1991). All residues in this domain of ∼100 residues have been mutagenized and two clusters of

A



B

	Organism				
	O. tauri	*C. reinhardtii*	*P. patens*	*S. moellend.*	*A. thaliana*
Na^+/K^+ pumps present	+	+	-	-	-
ENA $Na^+(K^+)$ pumps present	-	+	+	-	-
A "gate keeper" (Arg) in M6 of H^+ pumps	-	(+)	+	+	+
Region II and 14-3-3 binding site in C-terminal R-domain of H^+ pumps	-	-	-	+	+

FIGURE 9 | Phylogenetic tree of P3 ATPases (H^+-pumps) (A) and table summarizing conserved features (B). Accession numbers for sequences are given in **Table 2**. For comparison, the following outliers were included: a bacterial P1A ATPase *E. coli* KdpB (P03960), the *S. cerevisiae* pump ScPma1p (P05030), and the *Trypanosoma cruzi* pump TcHA1 (Q8T7V7).

autoinhibitory sequences have been identified, Region I and Region II (Axelsen et al., 1999). Further, in the extreme C-terminal end, a 14-3-3 binding site has been identified. 14-3-3 binding results in pump activation, but in order for 14-3-3 binding to occur, the penultimate residue (a Thr or Ser) first has to become phosphorylated (Fuglsang et al., 1999; Svennelid et al., 1999; Maudoux et al., 2000).

When C-terminal sequences of putative P3A ATPases were analyzed, the complete set of regulatory sequences (Region I–II and the 14-3-3 binding site) could be identified in all AHAs of *S. moellendorffii* and *A. thaliana* (**Figure 11**). In *S. moellendorffii* the shorter C-terminal regions of CrAHA2 and PpAHA1, a stretch of residues with weak but notable similarity to Region I could be identified (**Figure 11**). In these pumps, sequences with similarity to Region II and the 14-3-3 binding site could not be observed. In Chlorophyte P3A ATPases resembling protist plasma membrane H^+-ATPases (OtAHA1, CrAHA1, and CrAHA3) no sequences with similarity to any of these regions could be identified. The presence of a putative Region I in the C termini of CrAHA2 and PpAHA1 suggests that the basic regulatory apparatus of the higher plant C-terminus could have been present in the first green plants, but that more complex features (e.g., a 14-3-3 binding site) evolved latter in the evolution of vascular plants. In support of an early origin of Region I, the *S. cerevisiae* plasma membrane H^+-ATPase Pma1p has a short autoinhibitory sequence in its C-terminal domain with weak similarity to Region I, which appears to be involved in regulation of pump efficiency (Portillo et al., 1989; Venema and Palmgren, 1995).

P4 ATPases: PUTATIVE LIPID PUMPS

P4 ATPases in plants have been implicated in flipping phospholipids across biological membranes (Poulsen et al., 2008a). They are completely absent from eubacteria and archaebacteria, whereas in eukaryotes they are typically encoded for by multigene families (Axelsen and Palmgren, 1998).

We identified P4 ATPases in all of the Viridiplantae investigated in this study (**Figure 12**). The chlorophyte *O. tauri* was the only organism with a single P4 ATPase. This solitary P4 ATPase groups in the phylogenetic tree in the same larger branch as AtALA3. AtALA3 activity is connected with transport of phosphatidylethanolamine, phosphatidylserine, and phosphatidylcholine in *A. thaliana* (Poulsen et al., 2008b). Further,

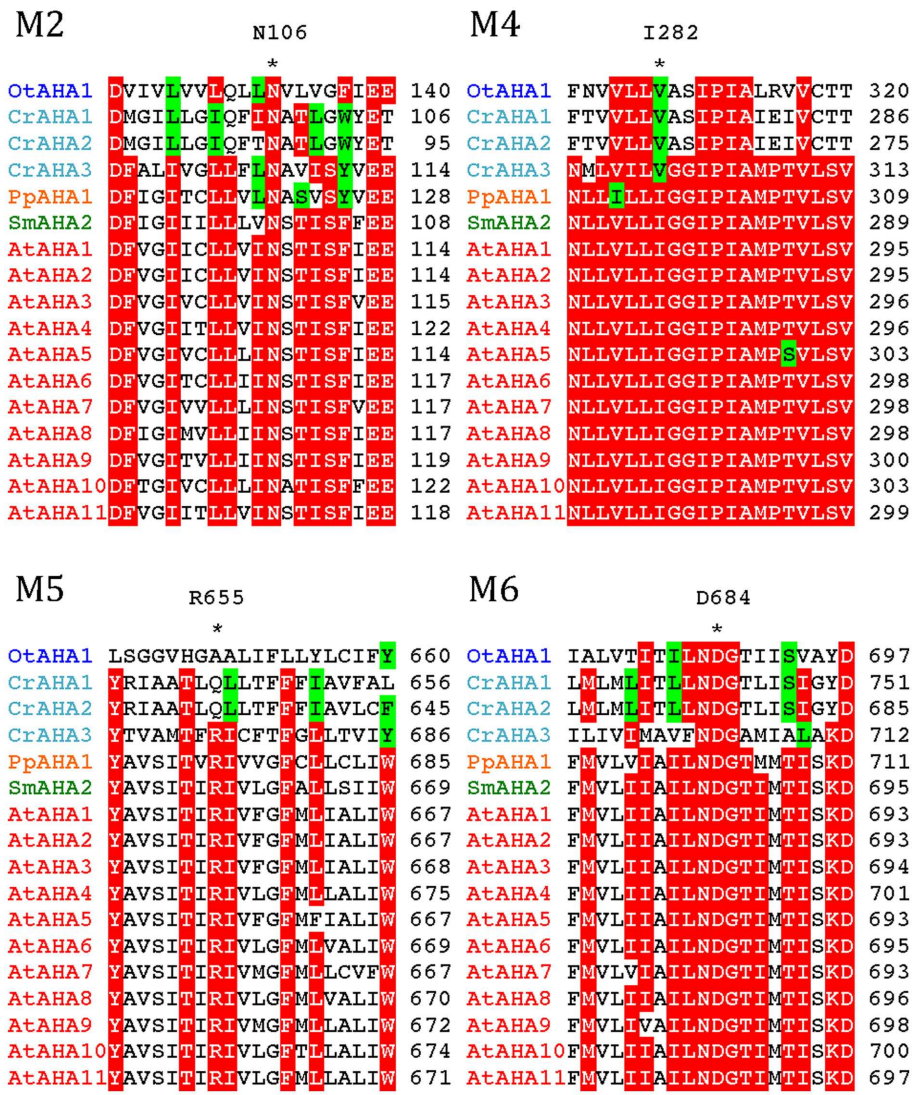


FIGURE 10 | Alignment of predicted transmembrane segments in P3A H⁺-ATPases. Asterisks mark residues of potential importance for proton coordination and pumping based on evidence from mutagenesis and analysis of a crystal structure for AtAHA2 (Pedersen et al., 2007). The number of the amino acid residue in AtAHA2 is indicated above each asterisk. R655 in M5, which seems important for controlling backflow of H^+ at high electrochemical gradients (Pedersen et al., 2007), is lacking in some chlorophyte H^+-ATPases.

Table 1 | Overview of number of pumps in different P-type ATPase subfamilies in Viridiplantae.

Class	Organism					Total
	O. tauri	*C. reinhardtii*	*P. patens*	*S. moellendorffii*	*A. thaliana*	
P1B	5	5	7	10	8	35
P2A	2	2	5	4	4	17
P2B	2	6	6	5	10	29
P2C	1	1	0	0	0	2
P2D	0	2	3	0	0	5
P3A	2	3	2	2	11	20
P4	1	3	8	8	12	32
P5A	1	1	1	1	1	5
P5B	0	1	2	2	0	5
Total	14	24	34	32	46	150

Table 2 | List of P-type ATPases analyzed in this study.

Class	Gene	aa	Accession #	Locus/ORF
OSTREOCOCCUS TAURI				
P1B	OtHMA1	1052	Q01EZ3	Ot02g01930
P1B	OtHMA2	681	Q00V55[1]	Ot15g01530
P1B	OtHMA3	1359	Q014R9	Ot07g03370
P1B	OtHMA4	861	Q00YQ6	Ot11g02480
P1B	OtHMA5	925	Q018N8	Ot05g03820
P2A	OtECA1	561	Q01FX0[1]	Ot01g04600
P2A	OtECA2	1013	Q01C29	Ot03g04740
P2B	OtACA1	1062	Q011R1	Ot09g02100
P2B	OtACA2	780	Q00U35[1]	Ot16g02150
P2C	OtNAK1	1172	Q01GZ3	Ot01g00900
P3A	OtAHA1	1043	Q00RY3[2]	Ot20g00330
P3A	OtAHA2	879	Q017J6[3]	Ot06g01560
P4	OtALA1	1258	Q016N2	Ot06g03680
P5A	OtP5A1	1398	Q00 × 11	Ot13g01720
CHLAMYDOMONAS REINHARDTII				
P1B	CrHMA1	1189	A8J6D5	CHLREDRAFT_195998
P1B	CrHMA2/CTP2	1086	A8IC93	CHLREDRAFT_205938
P1B	CrHMA3/CTP3	504	A8ICL1[4]	CHLREDRAFT_165985
P1B	CrHMA4/CTP1	1097	A8JBB5	CHLREDRAFT_206047
P1B	CrHMA5	1041	A8J829	CHLREDRAFT_195962
P2A	CrECA1	1069	A8I542	CHLREDRAFT_77047
P2A	CrECA2	845	g762350[5,6]	Cre22.g762350
P2B	CrACA1	1179	A8IZL7	CHLREDRAFT_196702
P2B	CrACA2	1009	A8IJV9	CHLREDRAFT_128099
P2B	CrACA3/FAB39	930	A8J0V2[1]	CHLREDRAFT_118223
P2B	CrACA4	1191	A8IS11	CHLREDRAFT_189266
P2B	CrACA5	873	A8JEM4[7]	CHLREDRAFT_179136
P2B	CrACA6	1434	A8HM60/g036350[8]	CHLREDRAFT_116583/Cre01.g036350
P2C	CrNAK1	1112	A8HX15	CHLREDRAFT_187139
P2D	CrENA1	1004	A8J4M6	CHLREDRAFT_104709
P2D	CrENA2	1001	A8J4M4	CHLREDRAFT_130931
P3A	CrAHA1/CrPMH1	1081	A8IFK0/Q9FNS3	CHLREDRAFT_54949
P3A	CrAHA2	802	A8IFH0[1]	CHLREDRAFT_38137
P3A	CrAHA3/CrPMA2	1081	Q93Z22	CHLREDRAFT_182602
P3A	–	1628	g316950[9]	Cre07.g316950
P4	CrALA1	1183	A8IVJ3	CHLREDRAFT_172401
P4	CrALA2	1300	A8IVJ6	CHLREDRAFT_190292
P4	CrALA3	1243	A8J8G9[6]	CHLREDRAFT_193025
P5A	CrP5A1	1168	A8JI26[1]	CHLREDRAFT_60708
P5B	CrP5B1	1308	A8I9J2[1]	CHLREDRAFT_186680
PHYSCOMITRELLA PATENS				
P1B	PpHMA1	902	A9RNK6	PHYPADRAFT_117222
P1B	PpHMA2	743	A9S8B9[1]	PHYPADRAFT_125638
P1B	PpHMA3	1004	A9T8Q3	PHYPADRAFT_192723
P1B	PpHMA4	841	A9SUQ2	PHYPADRAFT_215914
P1B	PpHMA5	1009	A9SME3	PHYPADRAFT_81365
P1B	PpHMA6	893	A9TB46	PHYPADRAFT_142921
P1B	PpHMA7	743	A9TQB5	PHYPADRAFT_148958
P2A	PpECA1	1060	A9RZK8	PHYPADRAFT_179791
P2A	PpECA2	1039	A9SH40	PHYPADRAFT_184915
P2A	PpECA3	1055	A9SHQ4	PHYPADRAFT_212461

(Continued)

Table 2 | Continued

Class	Gene	aa	Accession #	Locus/ORF
P2A	PpECA4	1000	A9TIL4	PHYPADRAFT_222630
P2A	PpECA5	1037	A9TK26	PHYPADRAFT_223041
P2A	PpSCA1	822	A9T9F0	PHYPADRAFT_220109
P2B	PpACA1/PCA1	1098	Q70TF0	PHYPADRAFT_202276
P2B	PpACA2/PCA2	1105	Q70TF1	PHYPADRAFT_224496
P2B	PpACA3	1074	A9SLT6	PHYPADRAFT_230135
P2B	PpACA4	948	A9RXA7[10]	PHYPADRAFT_121055
P2B	PpACA5	1035	A9RZJ8	PHYPADRAFT_121834
P2D	PpENA1	967	Q7XB51	
P2D	PpENA2	1058	Q7XB50	
P2D	PpENA3	963	C1L359	PHYPADRAFT_112089
P3A	PpAHA1	936	A9U0N9	PHYPADRAFT_153928
P3A	PpAHA2	465	A9TU44[1]	PHYPADRAFT_198308
P4	PpALA1	1219	A9RVW0	PHYPADRAFT_205967
P4	PpALA2	1062	A9S030	PHYPADRAFT_121975
P4	PpALA3	1229	A9S076	PHYPADRAFT_122321
P4	PpALA4	1194	A9SKC3	PHYPADRAFT_165384
P4	PpALA5	1251	A9SY94	PHYPADRAFT_189702
P4	PpALA6	1262	A9T6U6	PHYPADRAFT_88857
P4	PpALA7	1104	A9T776	PHYPADRAFT_168461
P4	PpALA8	1151	A9TDQ8	PHYPADRAFT_221270
P5A	PpP5A1	1178	A9RG70	PHYPADRAFT_113743
P5B	PpP5B1	435	A9TN06[1]	PHYPADRAFT_147950
P5B	PpP5B2	1290	A9SGC2	PHYPADRAFT_229474
SELAGINELLA MOELLENDORFFII[11]				
P1B	SmHMA1	819	D8RBL1	SELMODRAFT_89397
P1B	SmHMA2	696	D8QRZ3[1]	SELMODRAFT_60690
P1B	SmHMA3	952	D8SD62	SELMODRAFT_114297
P1B	SmHMA4	831	D8SJM4	SELMODRAFT_118425
P1B	SmHMA5	1018	D8SPX5	SELMODRAFT_122320
P1B	SmHMA6	910	D8TEP8	SELMODRAFT_138129
P1B	SmHMA7	953	D8RFP0	SELMODRAFT_92276
P1B	SmHMA8	790	D8QYH6[1]	SELMODRAFT_60775
P1B	SmHMA9	817	D8R2L0	SELMODRAFT_167322
P1B	SmHMA10	960	D8R2W8	SELMODRAFT_84115
P2A	SmECA1	1047	D8RSK1	SELMODRAFT_267711
P2A	SmECA2	1042	D8RUH8	SELMODRAFT_102055
P2A	SmECA3	1045	D8SA88	SELMODRAFT_112465
P2A	SmECA4	1009	D8SUV3	SELMODRAFT_158488
P2B	SmACA1	1014	D8QTC2	SELMODRAFT_437746
P2B	SmACA2	1030	D8QTC3	SELMODRAFT_266601
P2B	SmACA3	1068	D8T1F8	SELMODRAFT_129812
P2B	SmACA4	1105	D8S012	SELMODRAFT_451372
P2B	SmACA5	1069	D8T4Q6	SELMODRAFT_451597
P3A	SmAHA1	875	D8T8I0	SELMODRAFT_430150
P3A	SmAHA2	952	D8QT85[1]	SELMODRAFT_76771
P4	SmALA1	1207	D8QMQ3	SELMODRAFT_164122
P4	SmALA2	1157	D8R8G6	SELMODRAFT_86830
P4	SmALA3	1109	D8RG22	SELMODRAFT_410847
P4	SmALA4	1009	D8RKR6	SELMODRAFT_95836
P4	SmALA5	1208	D8S239	SELMODRAFT_107016
P4	SmALA6	1221	D8SBS1	SELMODRAFT_113552

(Continued)

Table 2 | Continued

Class	Gene	aa	Accession #	Locus/ORF
P4	SmALA7	1153	D8SGU6	SELMODRAFT_116847
P4	SmALA8	1184	D8TF22	SELMODRAFT_138337
P5A	SmP5A1	1109	D8TBK6	SELMODRAFT_187116
P5B	SmP5B1	1290	D8SFE9	SELMODRAFT_445079
P5B	SmP5B2	1246	D8R5P0	SELMODRAFT_439526
ARABIDOPSIS THALIANA				
P1B	AtHMA1	819	Q9M3H5	At4g37270
P1B	AtHMA2	951	Q9SZW4	At4g30110
P1B	AtHMA3	760	P0CW78	At4g30120
P1B	AtHMA4	1172	O64474	At2g19110
P1B	AtHMA5	995	Q9SH30	At1g63440
P1B	AtHMA6/PAA1	949	Q9SZC9	At4g33520
P1B	AtHMA7/RAN1	1001	Q9S7J8	At5g44790
P1B	AtHMA8/PAA2	856	Q9C594	At5g21930
P2A	AtECA1	1061	P92939	At1g07810
P2A	AtECA2	1054	O23087	At4g00900
P2A	AtECA3	998	Q9SY55	At1g10130
P2A	AtECA4	1061	Q9XES1	At1g07670
P2B	AtACA1	1020	Q37145	At1g27770
P2B	AtACA2	1014	O81108	At4g37640
P2B	AtACA4	1030	O22218	At2g41560
P2B	AtACA7	1015	O64806	At2g22950
P2B	AtACA8	1074	Q9LF79	At5g57110
P2B	AtACA9	1086	Q9LU41	At3g21180
P2B	AtACA10	1069	Q9SZR1	At4g29900
P2B	AtACA11	1025	Q9M2L4	At3g57330
P2B	AtACA12	1033	Q9LY77	At3g63380
P2B	AtACA13	1017	Q9LIK7	At3g22910
P2B	–	1049	F4KHQ2[12]	At5g53010
P3A	AtAHA1	949	P20649	At2g18960
P3A	AtAHA2	948	P19456	At4g30190
P3A	AtAHA3	949	P20431	At5g57350
P3A	AtAHA4	960	Q9SU58	At3g47950
P3A	AtAHA5	949	Q9SJB3	At2g24520
P3A	AtAHA6	949	Q9SH76	At2g07560
P3A	AtAHA7	961	Q9LY32	At3g60330
P3A	AtAHA8	948	Q9M2A0	At3g42640
P3A	AtAHA9	954	Q42556	At1g80660
P3A	AtAHA10	947	Q43128	At1g17260
P3A	AtAHA11	956	Q9LV11	At5g62670
P3A	–	813	Q9T0E0[13]	At4g11730
P4	AtALA1	1158	P98204	At5g04930
P4	AtALA2	1107	P98205	At5g44240
P4	AtALA3	1213	Q9XIE6	At1g59820
P4	AtALA4	1216	Q9LNQ4	At1g17500
P4	AtALA5	1228	Q9SGG3	At1g72700
P4	AtALA6	1240	Q9SLK6	At1g54280
P4	AtALA7	1247	Q9LVK9	At3g13900
P4	AtALA8	1189	Q9LK90	At3g27870
P4	AtALA9	1200	Q9SX33	At1g68710
P4	AtALA10	1202	Q9LI83	At3g25610
P4	AtALA11	1203	Q9SAF5	At1g13210

(Continued)

Table 2 | Continued

Class	Gene	aa	Accession #	Locus/ORF
P4	AtALA12	1184	P57792	At1g26130
P5A	AtP5A1	1179	Q9LT02	At5g23630

[1] *Fragment.*

[2] *A possible chimera. The last 200 amino acid residues do not match AtAHA4. Similarity ends at the position equivalent to position 862 in AtAHA4.*

[3] *Possibly a fragment. Maybe 80 amino acid residues missing from C-terminus (when compared to AtAHA2). The similarity extends to position 864 in AtAHA2.*

[4] *A possible fragment or pseudogene. When compared to AtHMA8/PAA1, positions 523–827 are missing (this includes the motif DKTGT). The last 60 amino acid residues are also missing.*

[5] *200 amino acid residues from the C-terminal are missing.*

[6] *No corresponding UniProtKB entry.*

[7] *When compared to AtACA9, amino acid residues 1–170 are missing.*

[8] *Fragment with insert. Similar to AtACA10 (107–529 + 576–1046; the positions on C. reinhardtii 3–462 + 856–1347). The UniProtKB entry only covers part of the protein taken from Phytozome.*

[9] *Possibly a pseudogene or chimera. Partly covered by A8I6H0 (460–576) and A8I6H2 (1521–1622). Many regions of low complexity. Similar to AtAHA10 (58–209 + 584–974, AtAHA10 numbering).*

[10] *First 130 aa missing compared to AtACA8.*

[11] *Near identical copies of most S. moellendorffii genes are present in databases indicating co-sequencing of two related cultivars.*

[12] *This protein is not likely to function as an ATPase as several key features are missing including core sequence 2 (TGES– > TASD), core seq. 4 (PEGL– > PVGL, core seq. 5 (the entire intron including DKTGTLT is missing), core sequence 6 (TGDN– > TDND), and core sequence 7 (VVAVTGDGTNDAPAL– > IVAATGMGIHDPKTL). cDNA: AY078942 (613–1049, but different splice variant), CB264713 (772–980), this splice variant. Probably similar to PMAX.*

[13] *Probable pseudogene (Axelsen and Palmgren, 2001).*

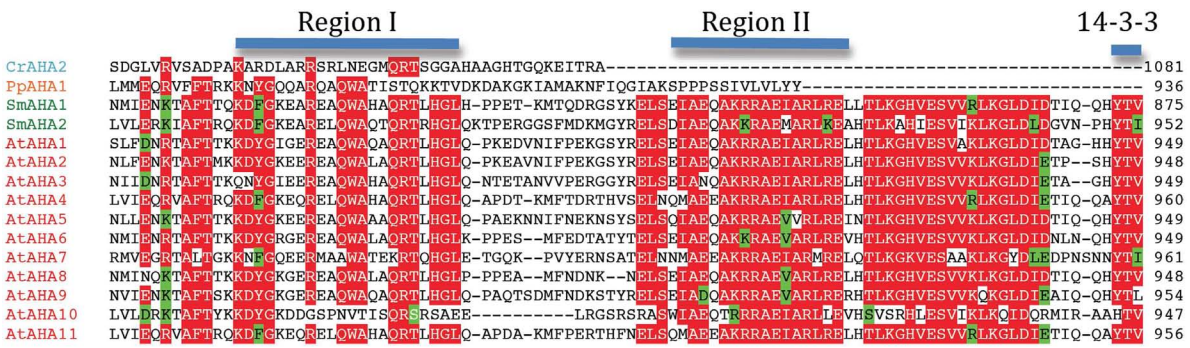


FIGURE 11 | Alignment of putative autoinhibitory regions in the C-terminal region of P3A H^+-ATPases.

it is a resident of the trans Golgi of root tip cells where it is connected to the generation of secretory vesicles leaving the Golgi apparatus (Poulsen et al., 2008b).

P5 ATPases: PUMPS WITH NO ASSIGNED SPECIFICITY

P5 ATPases constitute the least characterized group of P-type pumps and their transported ligand – if any – has not been identified. These pumps are absent from prokaryotes and are confined to eukaryotes where they reside in internal membrane systems (Møller et al., 2008). Based on sequence analysis, they are divided into two groups, P5A and P5B, each of which is predicted to transport different substrates based on differences in their transmembrane segments (**Figure 14**; Sørensen et al., 2010).

P5A ATPases have been found in all eukaryotic genomes analyzed so far (Møller et al., 2008) and were identified in all Viridiplantae analyzed in this study (**Figure 13**). Only a single P5A ATPase could be identified in each organism.

P5B ATPase sequences have so far been identified in the genomes of all eukaryotes examined, except for two plant lineages. Their widespread distribution supports a model in which they arose at an early point in the evolution of eukaryotes (Sørensen et al., 2010). When the Viridiplantae genomes analyzed here were searched for P5B sequences, we could identify P5B ATPases in *C. reinhardtii*, *P. patens*, and *S. moellendorffii*, but not in *O. tauri* and *A. thaliana* (**Figure 13**). This suggests that loss of P5B ATPases occurred at least twice in the evolution of Viridiplantae.

DISCUSSION

EVOLUTION OF PLANT P-TYPE ATPases

This survey of P-type ATPases provides evidence that members of the green plant lineage require at least five different types of P-type ATPases. Each of the five reference genomes analyzed showed at least one representative from each of the five subgroups of P-type ATPases. However, within the five different subgroups there is

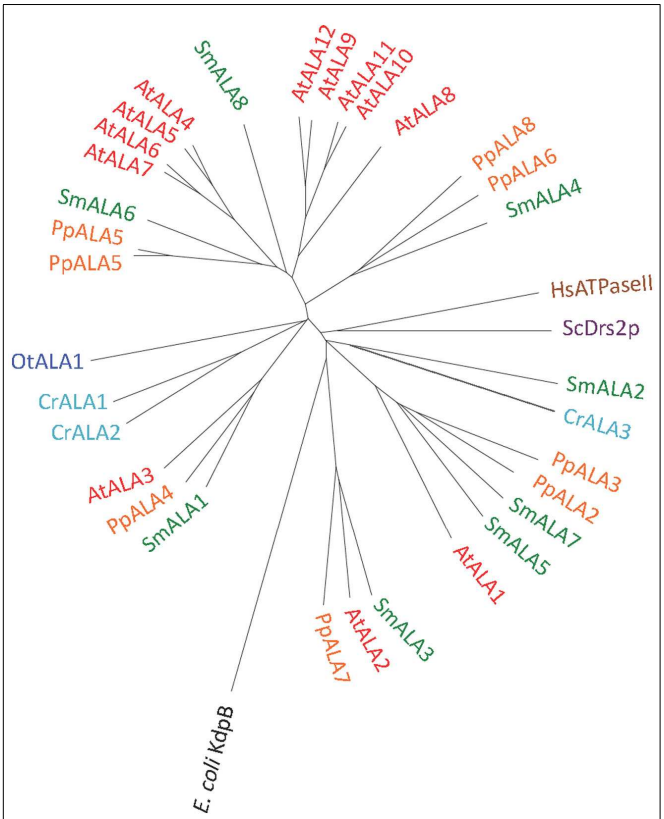


FIGURE 12 | Phylogenetic tree of P4 ATPases (lipid flippases). Accession numbers for sequences are given in **Table 2**. For comparison, the following outliers were included: A bacterial P1A ATPase *E. coli* KdpB (P03960), the *S. cerevisiae* pump ScDrs2p (P39524), and the *H. sapiens* pump HsATPaseII (ATP8A1; Q9Y2Q0).

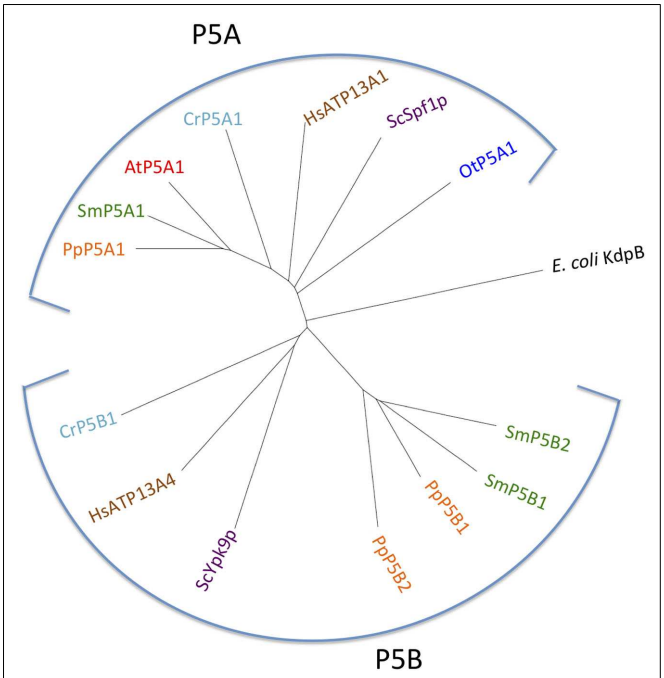


FIGURE 13 | Phylogenetic tree of P5 ATPases (having unknown transport activity). Accession numbers for sequences are given in **Table 2**. For comparison, the following outliers were included: a bacterial P1A ATPase *E. coli* KdpB (P03960), the *S. cerevisiae* pumps ScSpf1p (P39986) and ScYpk9 (Q12697), and the *H. sapiens* pumps HsATP13A1 (Q6NT90) and HsATP13A4 (Q4VNC1).

still evidence for considerable evolution of biochemical functions (**Figure 15**). For example, within the P2-type pumps, there are at least four subdivisions that delineate two subfamilies each of Na^+ and Ca^{2+} pumps (**Figure 5**).

In *Arabidopsis* there are 46 pumps, compared to 14 in *O. tauri*, 22 in *C. reinhardtii*, 33 in *P. patens*, and 32 in *S. moellendorffii*. This is consistent with a speculation that multicellular organisms require more pumps, presumably to provide specialized functions associated with more complex developmental programs or more variable environments. While loss of function phenotypes have been established for at least one member of each pump type in *Arabidopsis*, most of the 46 pumps remain uncharacterized at the genetic level.

The evolutionary diversity of P-type ATPases raises many questions to be explored. For example, why does *O. tauri* require only one lipid flippase (P4-type pump), whereas rice and *Arabidopsis* have 10 and 12 members, respectively (Baxter et al., 2003). What are the cellular functions of the lipid flippases? Are their biochemical functions limited to flipping lipids, or do they also help insert or remove lipids from membranes? At a structural level, how have the P4-pumps evolved from an ancestor that recognized simple cation substrates into having a dynamic interaction with lipids?

The relative expansion of the proton pump family (P3A) in flowering plants is also noteworthy, with 11 and 10 members in *Arabidopsis* and rice, respectively (Baxter et al., 2003), and only two in *P. patens* and *S. moellendorffii*. Do physiological complexities of flowering plants necessitate different isoform specific features?

In contrast, P5 pumps have only one or two representatives in all five reference organisms. The biochemical functions of these pumps are not clear, but their general importance to eukaryotes is supported by the presence of at least one representative in reference genomes from yeast to man.

The observation that all five reference genomes have two types of Ca^{2+} pumps, P2A and P2B suggests that both subgroups have conserved functions. A role in Ca^{2+} signaling has been proposed for P2B pumps, based on the presence of a regulatory domain that provides for activation of the pump by Ca^{2+}/calmodulin. Genetic evidence supports a model in which regulation of pump activity could modulate the magnitude or duration of a Ca^{2+} signal (Qudeimat et al., 2008; Boursiac et al., 2010; Zhu et al., 2010; Spalding and Harper, 2011). For the P2A pumps, genetic evidence suggests that they function in the homeostasis of both Ca^{2+} and Mn^{2+} (Wu et al., 2002; Li et al., 2008; Mills et al., 2008). Delineating signaling and nutritional functions of these pumps is an important challenge for the future.

THE EVOLUTION OF PLASMA MEMBRANE H^+ AND Na^+ PUMPS

The evolution of Na^+ pumps (type 2C and 2D) is of interest from several perspectives. First, the relatively close relationship between

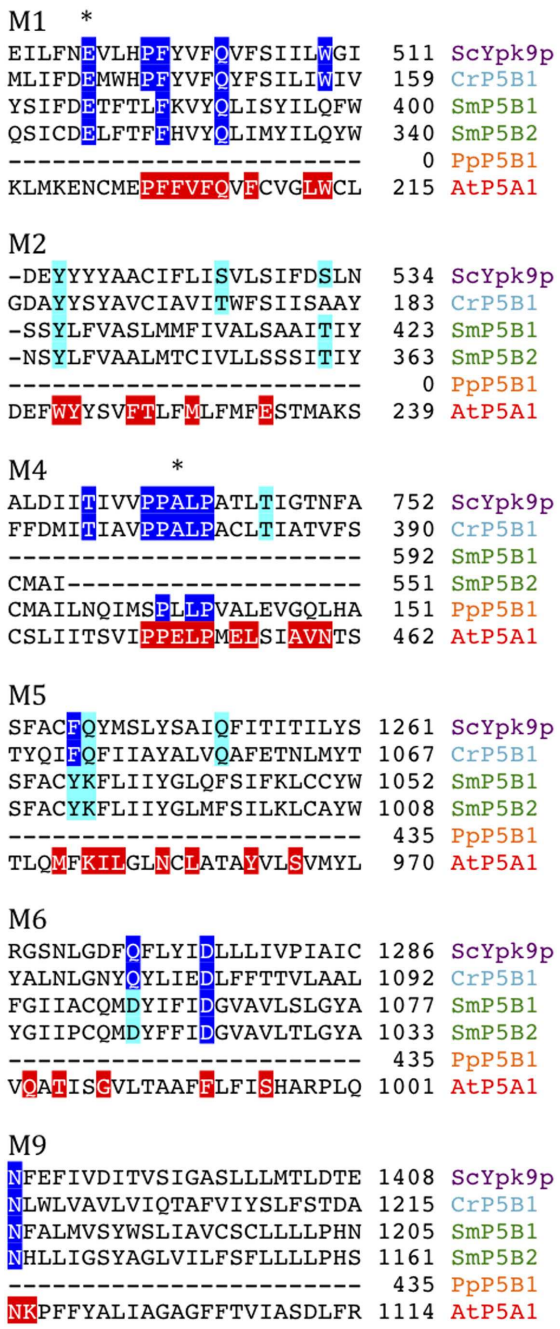


FIGURE 14 | Alignment of predicted transmembrane segments of putative P5B ATPases aligned with similar regions in the *S. cerevisiae* P5B ATPase ScYpk9p (Q12697). Residues conserved in all P5B ATPases (according to Sørensen et al., 2010) are marked in blue. Those that are highly conserved are marked in cyan. The P5A ATPase AtP5A/AtMIA is shown with residues conserved in P5A ATPases highlighted in red (Sørensen et al., 2010). Asterisks mark residues that are likely to play a role in ligand coordination.

Na^+ and Ca^{2+} pumps raises an interesting evolutionary question of which came first? Since both Na^+ and Ca^{2+} can be toxic within the cytosol, did both types of pumps arise early in evolution as a way to efflux toxic ions? Second, it appears that ancestral plants had both P2C and P2D ATPases. Both groups of pumps have remained in Chlorophyceae, represented here by the terrestrial green algae *C. reinhardtii*, but appear to have been lost in vascular plants (**Figure 15**).

An interesting question is why do the chlorophytes examined here have both H^+ and Na^+ pumps? For organisms such a *C. reinhardtii* that can live in fresh water, what is the evolutionary pressure for the retention of a Na^+ pump? For organisms that live in saline environments, such as *O. tauri*, why retain a plasma membrane H^+ pump? It seems reasonable to assume that Na^+ pumps alone could control cytoplasmic Na^+ levels and energize the plasma membrane for signaling and co-transport systems, as they do in typical animal cells. However, it is possible that the H^+ pumps actually evolved in marine organisms not to energize the plasma membrane, but rather to control cytoplasmic pH.

Regardless of their evolutionary origins, the observation that both H^+ and Na^+ pumps co-exist in the plasma membrane of well studied chlorophytes raises a question about which ion is used to drive secondary active transport systems? Are there different sets of H^+ or Na^+ specific co-transporters, or can the cotransporters be driven by either H^+ or Na^+ gradients? In vascular plants, Na^+ pumps appear to have been lost, leaving only plasma membrane H^+-ATPases to drive secondary transport systems. As a consequence it is thought that most plasma membrane cotransporters in vascular plants are H^+ coupled.

In the *Arabidopsis* pump AtAHA2, a conserved Arg in transmembrane segment M5 (Arg655) appears to be an important part of the H^+ pumping apparatus (Pedersen et al., 2007). This residue has been proposed to serve as a built-in cation that allows for rapid transition of the pump from the E1P to the E2P conformational state and as a gate-keeper that prevents H^+ from flowing backward ("backflow protection"; Pedersen et al., 2007; Buch-Pedersen et al., 2009). Further, in AHA2, a complex C-terminal regulatory domain with two autoinhibitory regions and a binding site for activating 14-3-3 protein is present (Axelsen et al., 1999). Backflow protection and multiple regulatory features are characteristic of all well studied H^+ pumps in flowering plants. The observation of similar features associated with the *S. moellendorffi* pumps (**Figure 11**), suggests that H^+ pumps with "advanced features" may be universal to all vascular plants.

In contrast, the plasma membrane H^+ pumps in chlorophytes appear to be more primitive in at least two aspects (**Figure 9B**). First, both P3A pumps of *O. taurii* lack the conserved Arg of M5 and two out of the three P3A pumps in *C. reinhardtii* also lack this residue (**Figure 10**). Second, a complex C-terminal regulatory domain is absent from the chlorophyte pumps.

The presence of "more advanced" H^+ pumps in vascular plants would suggest that these organisms have the capacity to utilize H^+ pumps to energize their plasma membranes with large electrochemical gradients. In flowering plants, H^+ pumps can create steep H^+ gradients and membrane potentials that can exceed 200 mV (negative on the inside; Hirsch et al., 1998). In comparison, Na^+/K^+ pumps in animal cells typically only produce membrane potentials around 60 mV. It is tempting to hypothesize that vascular plants have evolved to rely on these "more advanced" pumps to create very large electrochemical gradients for special purposes, such as signaling or nutrient transport. However,

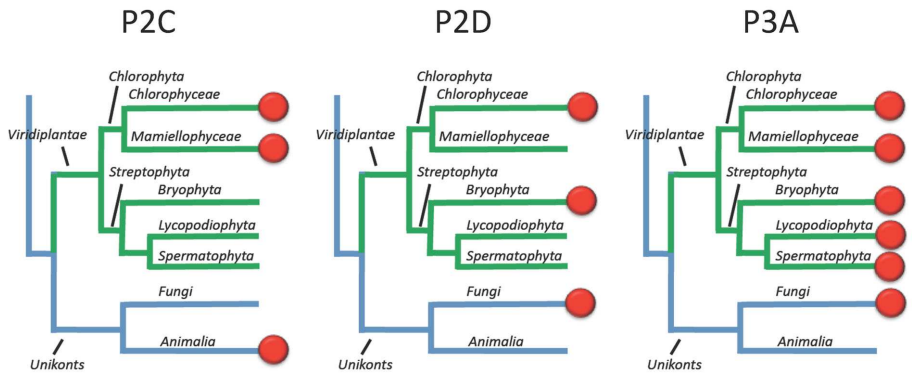


FIGURE 15 | Overview of the evolution of Na^+ and H^+ transporting P-type ATPases in Viridiplantae. Most likely, the ancestor of green plants had two types of Na^+ pumps (P2C and P2D) in addition to a plasma membrane H^+ pump (P3A). In present day plants, the terrestrial green algae *C. reinhardtii* still has all three types of pumps whereas Na^+ pumps have been lost completely in vascular plants (here represented by Lycopodiophyta and Spermatophyta).

during evolution, large H^+-based electrochemical gradients may have been problematic in the presence of a Na^+ pump. For example, a large membrane potential or pH gradient may have resulted in ancestral Na^+ pumps mis-functioning as H^+ or Na^+ leaks. As a result, Na^+ pumps may have been lost as vascular plants evolved to rely on their more advanced H^+ pumps.

FRONTIERS IN P-TYPE ATPase RESEARCH

The frontiers of P-type ATPase research can be divided into three areas. The first is to understand cellular and organismal functions of each pump. The second is to match biological functions with a structural understanding of how pumps transport specific substrates, and how their activities are modulated by signaling systems. The third is to explore ideas about how P-type ATPases can be altered or used to improve crop plants. For example, can Na^+ pumps from chlorophytes be moved into crop plants to provide improved Na^+ tolerance? Or would these pumps require re-engineering to prevent them from becoming H^+ or Na^+ leaks in cell types with large membrane potentials or H^+ gradients. As these frontiers are considered, it is worth remembering that P-type ATPases have already proven themselves as flexible substrates for evolution, producing pumps with a wide diversity of functions, from pumping protons to flipping lipids.

ACKNOWLEDGMENTS

Funding of Kristian B. Axelsen by The Swiss Federal Government through the Federal Office of Education and Science; and by the European Union (SLING: Serving Life-science Information for the Next Generation [226073] is gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Physiology/10.3389/fpls.2012.00031/abstract

REFERENCES

Abdel-Ghany, S. E., Muller-Moule, P., Niyogi, K. K., Pilon, M., and Shikanai, T. (2005). Two P-type ATPases are required for copper delivery in *Arabidopsis thaliana* chloroplasts. *Plant Cell* 17, 1233–1251.

Antebi, A., and Fink, G. R. (1992). The yeast Ca^{2+}-ATPase homologue, PMR1, is required for normal Golgi function and localizes in a novel Golgi-like distribution. *Mol. Biol. Cell* 3, 633–654.

Argüello, J. M., Eren, E., and González-Guerrero, M. (2007). The structure and function of heavy metal transport P_{1B}-ATPases. *Biometals* 20, 233–248.

Axelsen, K. B., and Palmgren, M. G. (1998). Evolution of substrate specificities in the P-Type ATPase superfamily. *J. Mol. Evol.* 46, 84–101.

Axelsen, K. B., and Palmgren, M. G. (2001). Inventory of the superfamily of P-type ion pumps in *Arabidopsis*. *Plant Physiol.* 126, 696–706.

Axelsen, K. B., Venema, K., Jahn, T., Baunsgaard, L., and Palmgren, M. G. (1999). Molecular dissection of the C-terminal regulatory domain of the plant plasma membrane H^+-ATPase AHA2: mapping of residues that when altered give rise to an activated enzyme. *Biochemistry* 38, 7227–7234.

Bækgaard, L., Luoni, L., De Michelis, M. I., and Palmgren, M. G. (2006). The plant plasma membrane Ca^{2+} pump ACA8 contains overlapping as well as physically separated autoinhibitory and calmodulin-binding domains. *J. Biol. Chem.* 281, 1058–1065.

Banks, J. A. (2009). *Selaginella* and 400 million years of separation. *Annu. Rev. Plant Biol.* 60, 223–238.

Barrero-Gil, J., Garciadeblás, B., and Benito, B. (2005). Sodium, potassium-ATPases in algae and oomycetes. *J. Bioenerg. Biomembr.* 37, 269–278.

Baxter, I., Tchieu, J., Sussman, M. R., Boutry, M., Palmgren, M. G., Gribskov, M., Harper, J. F., and Axelsen, K. B. (2003). Genomic comparison of P-type ATPase ion pumps in *Arabidopsis* and rice. *Plant Physiol.* 132, 618–628.

Boursiac, Y., Lee, S. M., Romanowsky, S., Blank, R., Sladek, C., Chung, W. S., and Harper, J. F. (2010). Disruption of the vacuolar calcium-ATPases in *Arabidopsis* results in the activation of a salicylic acid-dependent programmed cell death pathway. *Plant Physiol.* 154, 1158–1171.

Buch-Pedersen, M. J., Pedersen, B. P., Veierskov, B., Nissen, P., and Palmgren, M. G. (2009). Protons and how they are transported by proton pumps. *Pflugers Arch.* 457, 573–579.

Campbell, A. M., Coble, A. J., Cohen, L. D., Ch'ng, T. H., Russo, K. M., Long, E. M., and Armburst, E. V. (2001). Identification and DNA sequence of a new H^+-ATPase in the unicellular green alga *Chlamydomonas reinhardtii* (Chlorophyceae) *J. Phycol.* 37, 536–542.

Courties, C., Perasso, R., Chrétiennot-Dinet, M.-J., Gouy, M., Guillou, L., and Troussellier, M. (1998). Phylogenetic analysis and genome size of *Ostreococcus tauri* (Chlorophyta, Prasinophyceae). *J. Phycol.* 34, 844–849.

Curran, A. C., Hwang, I., Corbin, J., Martinez, S., Rayle, D., Sze, H., and Harper, J. F. (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.

Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A. Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié,

S., Cooke, R., Saeys, Y., Wuyts, J., Jabbari, K, Bowler, C., Panaud, O., Piégu, B., Ball, S. G., Ral, J. P., Bouget, F. Y., Piganeau, G., De Baets, B., Picard, A., Delseny, M., Demaille, J., Van de Peer, Y., and Moreau, H. (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11647–11652.

Dutta, S. J., Liu, J., Hou, Z., and Mitra, B. (2006). Conserved aspartic acid 714 in transmembrane segment 8 of the ZntA subgroup of P_{1B}-type ATPases is a metal-binding residue. *Biochemistry* 45, 5923–5931.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Flowers, T. J., and Colmer, T. D. (2008). Salinity tolerance in halophytes. *New Phytol.* 179, 945–963.

Fuglsang, A. T., Visconti, S., Drumm, K., Jahn, T., Stensballe, A., Mattei, B., Jensen, O. N., Aducci, P., and Palmgren, M. G. (1999). Binding of 14-3-3 protein to the plasma membrane H^+-ATPase AHA2 involves the three C-terminal residues Tyr^{946}-Thr-Val and requires phosphorylation of Thr947. *J. Biol. Chem.* 274, 36774–36780.

Gimmler, H. (2000). Primary sodium plasma membrane ATPases in salt-tolerant algae: facts and fictions. *J. Exp. Bot.* 51, 1171–1178.

Gonzalez, R. J. (2011). The physiology of hyper-salinity tolerance in teleost fish: a review. *J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.* (in press).

Harper, J. F., Hong, B., Hwang, I., Guo, H. Q., Stoddard, R., Huang, J. F., Palmgren, M. G., and 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.

Haruta, M., Burch, H. L., Nelson, R. B., Barrett-Wilt, G., Kline, K. G., Mohsin, S. B., Young, J. C., Otegui, M. S., and Sussman, M. R. (2010). Molecular characterization of mutant *Arabidopsis* plants with reduced plasma membrane proton pump activity. *J. Biol. Chem.* 285, 17918–17929.

Hirsch, R. E., Lewis, B. D., Spalding, E. P., and Sussman, M. R. (1998). A role for the AKT1 potassium channel in plant nutrition. *Science* 280, 918–921.

Howe, K., Bateman, A., and Durbin, R. (2002). QuickTree: building huge neighbour-joining trees of protein sequences. *Bioinformatics* 18, 1546–1547.

Huson, D. H., Richter, D. C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8, 460. doi:10.1186/1471-2105-8-460

Hussain, D., Haydon, M. J., Wang, Y., Wong, E., Sherson, S. M., Young, J., Camakaris, J., Harper, J. F., and Cobbett, C. S. (2004). P-type ATPase heavy metal transporters with roles in essential zinc homeostasis in *Arabidopsis*. *Plant Cell* 16, 1327–1339.

Hwang, I., Harper, J. F., Liang, F., and Sze, H. (2000). Calmodulin activation of an endoplasmic reticulum-located calcium pump involves an interaction with the N-terminal autoinhibitory domain. *Plant Physiol.* 122, 157–168.

James, P., Maeda, M., Fischer, R., Verma, A., Krebs, J., Penniston, J., and Carafoli, E. (1988). Identification and primary structure of a calmodulin binding domain of the Ca^{2+} pump of human erythrocytes. *J. Biol. Chem.* 263, 2905–2910.

Kim, Y. Y., Choi, H., Segami, S., Cho, H. T., Martinoia, E., Maeshima, M., and Lee, Y. (2009). AtHMA1 contributes to the detoxification of excess Zn(II) in *Arabidopsis*. *Plant J.* 58, 737–753.

Li, X., Chanroj, S., Wu, Z., Romanowsky, S. M., Harper, J. F., and Sze, H. (2008). A distinct endosomal Ca^{2+}/Mn^{2+} pump affects root growth through the secretory process. *Plant Physiol.* 147, 1675–1689.

Luo, S., Fang, J., and Docampo, R. (2006). Molecular characterization of *Trypanosoma brucei* P-type H^+-ATPases. *J. Biol. Chem.* 281, 21963–21973.

Luo, S., Scott, D. A., and Docampo, R. (2002). Trypanosoma cruzi H^+-ATPase 1 (TcHA1) and 2 (TcHA2) genes complement yeast mutants defective in H^+ pumps and encode plasma membrane P-type H^+-ATPases with different enzymatic properties. *J. Biol. Chem.* 277, 44497–44506.

Malmström, S., Askerlund, P., and Palmgren, M. G. (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.

Maudoux, O., Batoko, H., Oecking, C., Gevaert, K., Vandekerckhove, J., Boutry, M., and Morsomme, P. (2000). A plant plasma membrane H^+-ATPase expressed in yeast is activated by phosphorylation at its penultimate residue and binding of 14-3-3 regulatory proteins in the absence of fusicoccin. *J. Biol. Chem.* 275, 17762–17770.

Merchant, S. S., Prochnik, S. E., and Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Maréchal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riaño-Pachón, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245–250.

Mills, R. F., Doherty, M. L., López-Marqués, R. L., Weimar, T., Dupree, P., Palmgren, M. G., Pittman, J. K., and Williams, L. E. (2008). ECA3, a Golgi-localized P_{2A}-type ATPase, plays a crucial role in manganese nutrition in *Arabidopsis*. *Plant Physiol.* 146, 116–128.

Mills, R. F., Krijger, G. C., Baccarini, P. J., Hall, J. L., and Williams, L. E. (2003). Functional expression of AtHMA4, a P_{1B}-type ATPase of the Zn/Co/Cd/Pb subclass. *Plant J.* 35, 164–176.

Møller, A. B., Asp, T., Holm, P. B., and Palmgren, M. G. (2008). Phylogenetic analysis of P_5 P-type ATPases, a eukaryotic lineage of secretory pathway pumps. *Mol. Phylogenet. Evol.* 46, 619–634.

Møller, J. V., Juul, B., and le Maire, M. (1996). Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. Biophys. Acta* 1286, 1–51.

Moreno, I., Norambuena, L., Maturana, D., Toro, M., Vergara, C., Orellana, A., Zurita-Silva, A., and Ordenes, V. R. (2008). AtHMA1 is a thapsigargin-sensitive Ca^{2+}/heavy metal pump. *J. Biol. Chem.* 283, 9633–9641.

Morth, J. P., Pedersen, B. P., Buch-Pedersen, M. J., Andersen, J. P., Vilsen, B., Palmgren, M. G., and Nissen, P. (2011). A structural overview of the plasma membrane Na^+,K^+-ATPase and H^+-ATPase ion pumps. *Nat. Rev. Mol. Cell Biol.* 12, 60–70.

Morth, J. P., Pedersen, B. P., Toustrup-Jensen, M. S., Sørensen, T. L., Petersen, J., Andersen, J. P., Vilsen, B., and Nissen, P. (2007). Crystal structure of the sodium-potassium pump. *Nature* 450, 1043–1049.

Olesen, C., Picard, M., Winther, A. M., Gyrup, C., Morth, J. P., Oxvig, C., Møller, J. V., and Nissen, P. (2007). The structural basis of calcium transport by the calcium pump. *Nature* 450, 1036–1042.

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

Palmgren, M. G., and Harper, J. F. (1999). Pumping with plant P-type ATPases. *J. Exp. Bot.* 50, 883–893.

Palmgren, M. G., Sommarin, M., Serrano, R., and Larsson, C. (1991). Identification of an autoinhibitory domain in the C-terminal region of the plant plasma membrane H^+-ATPase. *J. Biol. Chem.* 266, 20470–20475.

Pedersen, B. P., Buch-Pedersen, M. J., Morth, J. P., Palmgren, M. G., and Nissen, P. (2007). Crystal structure of the plasma membrane proton pump. *Nature* 450, 1111–1114.

Popova, L., Balnokin, Y., Dietz, K. J., and Gimmler, H. (1999). Characterization of phosphorylated intermediates synthetized during the catalytic cycle of the sodium adenosine triphosphatase in the plasma membrane of the marine unicellular alga *Tetraselmis (Platymonas)*

viridis. *J. Plant Physiol.* 155, 302–309.

Popova, L. G., Shumkova, G. A., Andreev, I. M., and Balnokin, Y. V. (2005). Functional identification of electrogenic Na^+-translocating ATPase in the plasma membrane of the halotolerant microalga *Dunaliella maritima*. *FEBS Lett.* 579, 5002–5006.

Portillo, F., de Larrinoa, I. F., and Serrano, R. (1989). Deletion analysis of yeast plasma membrane H^+-ATPase and identification of a regulatory domain at the carboxyl-terminus. *FEBS Lett.* 247, 381–385.

Poulsen, L. R., López-Marqués, R. L., and Palmgren, M. G. (2008a). Flippases: still more questions than answers. *Cell. Mol. Life Sci.* 65,3119–3125.

Poulsen, L. R., López-Marqués, R. L., McDowell, S. C., Okkeri, J., Licht, D., Schulz, A., Pomorski, T., Harper, J. F., and Palmgren, M. G. (2008b). The Arabidopsis P4-ATPase ALA3 localizes to the Golgi and requires a β-subunit to function in lipid translocation and secretory vesicle formation. *Plant Cell* 20, 658–676.

Qudeimat, E., Faltusz, A. M., Wheeler, G., Lang, D., Brownlee, C., Reski, R., and Frank, W. (2008). A P_{IIB}-type Ca^{2+}-ATPase is essential for stress adaptation in *Physcomitrella patens*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19555–19560.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van de Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrano, R. S., and Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69.

Rochaix, J. D. (1995). *Chlamydomonas reinhardtii* as the photosynthetic yeast. *Annu. Rev. Genet.* 29, 209–230.

Rodríguez-Navarro, A., and Benito, B. (2010). Sodium or potassium efflux ATPase – a fungal, bryophyte, and protozoal ATPase. *Biochim. Biophys. Acta* 1798, 1841–1853.

Rodríguez-Navarro, A. (2000). Potassium transport in fungi and plants. *Biochim. Biophys. Acta* 1469, 1–30.

Rudolph, H. K., Antebi, A., Fink, G. R., Buckley, C. M., Dorman, T. E., LeVitre, J., Davidow, L. S., Mao, J. I., and Moir, D. T. (1989). The yeast secretory pathway is perturbed by mutations in PMR1, a member of a Ca^{2+} ATPase family. *Cell* 58, 133–145.

Sáez, A. G., Lozano, E., and Zaldívar-Riverón, A. (2009). Evolutionary history of Na,K-ATPases and their osmoregulatory role. *Genetica* 136, 479–490.

Seigneurin-Berny, D., Gravot, A., Auroy, P., Mazard, C., Kraut, A., Finazzi, G., Grunwald, D., Rappaport, F., Vavasseur, A., Joyard, J., Richaud, P., and Rolland, N. (2006). HMA1, a new Cu-ATPase of the chloroplast envelope, is essential for growth under adverse light conditions. *J. Biol. Chem.* 281, 2882–2892.

Serrano, R. (1989). Structure and function of proton translocating ATPase in plasma membranes of plants and fungi. *Biochim. Biophys. Acta* 947, 1–28.

Skou, J. C. (1957). The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochim. Biophys. Acta* 23, 394–401.

Skulachev, V. P. (1988). *Membrane Bioenergetics*. Berlin: Springer-Verlag.

Sørensen, D. M., Buch-Pedersen, M. J., and Palmgren, M. G. (2010). Structural divergence between the two subgroups of P5 ATPases. *Biochim. Biophys. Acta* 1797, 846–855.

Spalding, E. P., and Harper, J. F. (2011). The ins and outs of cellular Ca^{2+} transport. *Curr. Opin. Plant Biol.* 14, 715–720.

Svennelid, F., Olsson, A., Piotrowski, M., Rosenquist, M., Ottman, C., Larsson, C., Oecking, C., and Sommarin, M. (1999). Phosphorylation of Thr-948 at the C terminus of the plasma membrane H^+-ATPase creates a binding site for the regulatory 14-3-3 protein. *Plant Cell* 11, 2379–2392.

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

Tester, M., and Davenport, R. (2003). Na^+ tolerance and Na^+ transport in higher plants. *Ann. Bot.* 91, 503–527.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

The UniProt Consortium. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 40, D71–D75.

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

Venema, K., and Palmgren, M. G. (1995) Metabolic modulation of transport coupling ratio in yeast plasma membrane H^+-ATPase. *J. Biol. Chem.* 270, 19659–19667.

Verret, F., Gravot, A., Auroy, P., Leonhardt, N., David, P., Nussaume, L., Vavasseur, A., and Richaud, P. (2004). Overexpression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance. *FEBS Lett.* 576, 306–312.

Whittamore, J. M. (2012). Osmoregulation and epithelial water transport: lessons from the intestine of marine teleost fish. *J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.* 182, 1–39.

Williams, L. E., and Mills, R. F. (2005). P_{1B}-ATPases – an ancient family of transition metal pumps with diverse functions in plants. *Trends Plant Sci.* 10, 491–502.

Wong, C. K., and Cobbett, C. S. (2009). HMA P-type ATPases are the major mechanism for root-to-shoot Cd translocation in *Arabidopsis thaliana*. *New Phytol.* 181, 71–78.

Wu, Z., Liang, F., Hong, B., Young, J. C., Sussman, M. R., Harper, J. F., and 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.

Wuytack, F., Raeymaekers, L., and Missiaen, L. (2003). PMR1/SPCA Ca^{2+} pumps and the role of the Golgi apparatus as a Ca^{2+} store. *Pflugers Arch.* 446, 148–153.

Yamaguchi, T., and Blumwald, E. (2005). Developing salt-tolerant crop plants: challenges and opportunities. *Trends Plant Sci.* 10, 615–620.

Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21, 809–818.

Zhu, X., Caplan, J., Mamillapalli, P., Czymmek, K., and Dinesh-Kumar, S. P. (2010). Function of endoplasmic reticulum calcium ATPase in innate immunity-mediated programmed cell death. *EMBO J.* 29, 1007–1018.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 November 2011; accepted: 28 January 2012; published online: 21 February 2012.

Citation: Pedersen CNS, Axelsen KB, Harper JF and Palmgren MG (2012) Evolution of plant P-type ATPases. Front. Plant Sci. ***3****:31. doi: 10.3389/fpls.2012.00031*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 14 February 2012
doi: 10.3389/fpls.2012.00025

Conserved and diversified gene families of monovalent cation/H^+ antiporters from algae to flowering plants

Salil Chanroj[1], Guoying Wang[1], Kees Venema[2], Muren Warren Zhang[1], Charles F. Delwiche[1] and Heven Sze[1]*

[1] *Department of Cell Biology and Molecular Genetics, Maryland Agricultural Experiment Station, University of Maryland, College Park, MD, USA*
[2] *Departmento de Bioquímica, Biología Celular y Molecular de Plantas, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Granada, Spain*

Edited by:
Markus Geisler, University of Fribourg, Switzerland

Reviewed by:
Eduardo Blumwald, University of California Davis, USA
Wei-Hua Tang, Chinese Academy of Sciences, China
Enrico Martinoia, Universitaet Zuerich, Switzerland

***Correspondence:**
Heven Sze, Department of Cell Biology and Molecular Genetics, Maryland Agricultural Experiment Station, University of Maryland, Bioscience Research Building # 413, College Park, MD 20742, USA.
e-mail: hsze@umd.edu

All organisms have evolved strategies to regulate ion and pH homeostasis in response to developmental and environmental cues. One strategy is mediated by monovalent cation–proton antiporters (CPA) that are classified in two superfamilies. Many CPA1 genes from bacteria, fungi, metazoa, and plants have been functionally characterized; though roles of plant CPA2 genes encoding K^+-efflux antiporter (KEA) and cation/H^+ exchanger (CHX) families are largely unknown. Phylogenetic analysis showed that three clades of the CPA1 Na^+–H^+ exchanger (NHX) family have been conserved from single-celled algae to *Arabidopsis*. These are (i) plasma membrane-bound SOS1/AtNHX7 that share ancestry with prokaryote NhaP, (ii) endosomal AtNHX5/6 that is part of the eukaryote Intracellular-NHE clade, and (iii) a vacuolar NHX clade (AtNHX1–4) specific to plants. Early diversification of *KEA* genes possibly from an ancestral cyanobacterium gene is suggested by three types seen in all plants. Intriguingly, *CHX* genes diversified from three to four members in one subclade of early land plants to 28 genes in eight subclades of *Arabidopsis*. Homologs from *Spirogyra* or *Physcomitrella* share high similarity with AtCHX20, suggesting that guard cell-specific AtCHX20 and its closest relatives are founders of the family, and pollen-expressed *CHX* genes appeared later in monocots and early eudicots. AtCHX proteins mediate K^+ transport and pH homeostasis, and have been localized to intracellular and plasma membrane. Thus *KEA* genes are conserved from green algae to angiosperms, and their presence in red algae and secondary endosymbionts suggest a role in plastids. In contrast, AtNHX1–4 subtype evolved in plant cells to handle ion homeostasis of vacuoles. The great diversity of *CHX* genes in land plants compared to metazoa, fungi, or algae would imply a significant role of ion and pH homeostasis at dynamic endomembranes in the vegetative and reproductive success of flowering plants.

Keywords: cation homeostasis, pH homeostasis, dynamic endomembrane, secretory system, protein, cargo sorting

INTRODUCTION

Cells have evolved mechanisms to regulate ion and pH homeostasis in order to respond to developmental cues and to adapt to a constantly changing environment. In prokaryotes, this feat is accomplished with a diverse array of transporters at the plasma membrane. In eukaryotes, cells have increased in size and contain more intracellular membranes, including organelles and various compartments of the secretory and endocytic pathways. Consequently, the number of ion transporters localized to endomembranes has increased considerably. As plant ancestors initially living in an aquatic or marine environment colonized land, we wondered how transport systems required for growth, reproduction, and adaptation evolved in their structure and function. The genomes of diverse plants have been sequenced in recent years offering the first opportunity to understand how plant transporters evolved over one billion years. Here we focus on the monovalent cation/proton antiporter (CPA) superfamily whose members are thought to regulate cation and pH homeostasis (Brett et al., 2005a) by exchanging Na^+, Li^+, or K^+ for H^+. Prokaryote and eukaryote CPA members fall into two major families designated as CPA1 (2.A.36) and CPA2 (2.A.37; Saier, 2000; Brett et al., 2005a). In plants, CPA1 includes the well-studied Na^+–H^+ exchanger (NHX) family, and the relatively obscure CPA2 family that includes K^+ efflux Antiporter (KEA) and Cation–H^+ exchanger (CHX) families. The purpose of this study is to understand the evolutionary relationship of NHX, KEA, and CHX homologs from algae to flowering plants, to use comparative biology to infer functions of uncharacterized plant genes (Chang et al., 2004), and hopefully reveal new model systems to determine functions of KEA and CHX genes which remain largely uncharacterized.

MATERIALS AND METHODS

ORGANISMS

Plant genomes that have been completely sequenced and available in Phytozome were used. The list of plant species and their abbreviated names in parenthesis include *Manihot esculenta (Mes)*, *Ricinus communis (Rco)*, *Populus trichocarpa (Ptr)*, *Medicago truncatula*

(Mtr), Glycine max (Gma), Cucumis sativus (Csa), Arabidopsis thaliana (Ath), Arabidopsis lyrata (Aly), Carica papaya (Cpa), Eucalyptus grandis (Egr), Vitis vinifera (Vvi), Mimulus guttatus (Mgu), Sorghum bicolor (Sbi), Zea mays (Zma), Oryza sativa (Osa), Brachypodium distachyon (Bdi), Selaginella moellendorffii (Smo or club moss), Physcomitrella patens (Ppa or moss), Volvox carteri (Vca), and *Chlamydomonas reinhartdii (Cre).*

COLLECTING SEQUENCES FROM PLANTS

Deduced protein sequences from diverse plants were initially collected from Phytozome[1] by Blast analyses using NHX1, NHX5, NHX7, KEA1, KEA3, KEA4, or CHX17 from *Arabidopsis thaliana.* The sequences of each plant species were then compared with *Arabidopsis* protein by alignment with Clustal W or MUSCLE (Edgar, 2004) and analyzed with T-coffee (Notredame et al., 2000) or MEGA5 (Tamura et al., 2011). Abbreviated names were given to each gene according to the nomenclature used for *A. thaliana* (Maser et al., 2001). Blast analyses with *Arabidopsis* genes also recovered predicted proteins from an EST sequencing project of Charophyceae (Timme et al., 2012); though many transcripts were truncated and were not used in phylogenetic analysis. Other sequences from bacteria, cyanobacteria, fungi, protist, or metazoa were obtained by BLAST from JGI[2] or NCBI[3].

ALIGNMENT AND CONSTRUCTING TREES

RAxML

Sequences were assembled using Clustal X2 program (Larkin et al., 2007) to an original fasta file (fa.), and then aligned by Muscle software (Edgar, 2004). Alignment was exported and saved as fasta (afa), phylip (phy), or MSF format. RAxML was used for phylogenetic analysis and MSF was read by Genedoc (Nicholas et al., 1997[4]) for editing. After alignment with Clustal X2 or MUSCLE, sequences that included the Pfam domain (00999) Na/H exchanger domain were retained for further analysis. Short sequences or those lacking >30% of the transmembrane domain of ∼400 residues were dropped, though they resemble members of the family. When sequences outside the conserved domain diverged significantly, the N-terminus and C-tail were trimmed to maximize alignment (e.g., **Figure 1**). We used maximum-likelihood and tested phylogeny with a bootstrap of 100–500, JTT model (Jones et al., 1992), to construct trees with either

[1]Phytozome.net

[2]http://www.genome.jp/sit-bin/search_sequence?prog = blast&db = gvi&seqid = aa
[3]http://www.ncbi.nlm.nih.gov/protein
[4]http://www.nrbsc.org/gfx/genedoc/

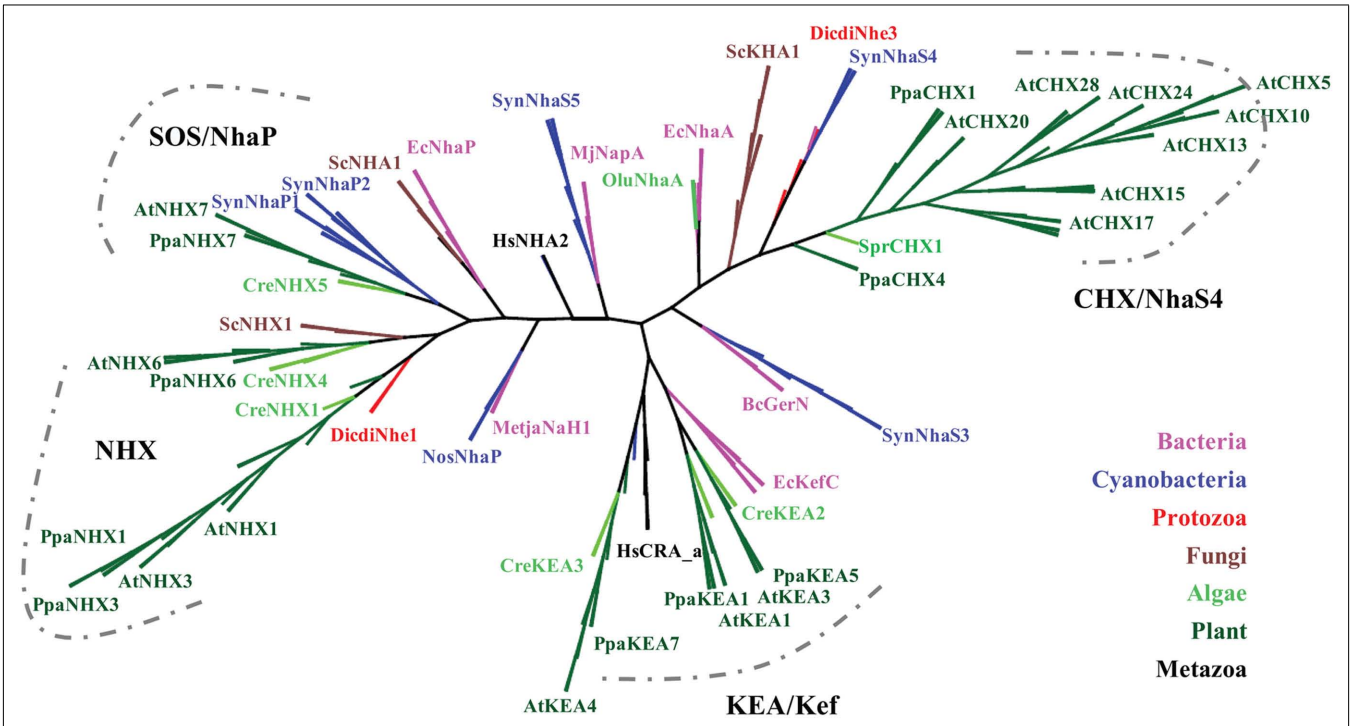


FIGURE 1 | Protein phylogeny of cation–proton antiporters shows the evolutionary history of CPA1 and CPA2 families in plants. The conserved Pfam 00999 domains from diverse organisms were aligned by MUSCLE and the evolutionary history of the sequences were determined by RAxML, maximum likelihood using the Jones–Taylor–Thornton (JTT) model with a bootstrap of 500. (Final ML optimization likelihood = −104241.644676). Both N and C ends were trimmed, so aligned TM domain for proteins was ∼400 residues, corresponding to AtCHX17 residue 33–426. Organisms are color-coded as follows: bacteria (magenta), Cyanobacteria (dark-blue), protist: *Dictyostelium* (red), fungi (brown), algae (green), and plants (dark green). NhaP and NhaA genes from ancestral bacteria likely gave rise to eukaryote CPA1 (NhaP and NHX) and CPA2 (KEA and CHX) families, respectively. Abbreviated protein names and organisms are defined in Table S1 in Supplementary Material. Unrooted rectangular tree is shown in Figure S1 in Supplementary Material.

RAxML (Stamatakis, 2006) or MEGA5 (Tamura et al., 2011). The bestTree file generated from the program gives the evolutionary value whereas the bipartition file informs the statistical value.

RESULTS

We first examined the evolutionary origin of plant cation–proton antiporters (CPA1 and CPA2) using *Arabidopsis* and rice as representatives of flowering plants. All these sequences had a conserved Na/H exchanger domain (Pfam00999) of about 400 residues. **Figure 1** shows three distinct branches which are labeled as CPA1 (NHX and NHAP), CPA2 with KEA, and with CHX families.

MONOVALENT CATION PROTON ANTIPORTER-1

Several prokaryotic and archaea NhaP genes are located at the base of the CPA1 branch supporting the idea that an ancestral NhaP gene gave rise to eukaryote CPA1 genes (**Figures 1** and **2**). This family of transporters is predicted to have 10–12 membrane-spanning domains, and bacterial NhaP mediates electroneutral Na^+/H^+ exchange (see **Table 1**). CPA1-type transporters are found in all kingdoms, including archaea, bacteria, fungi, plants, and metazoa (Brett et al., 2005a). Our analysis shows that ancestral NhaP evolved and diverged to give two distinct types: NHX and NHAP (**Figure 2**).

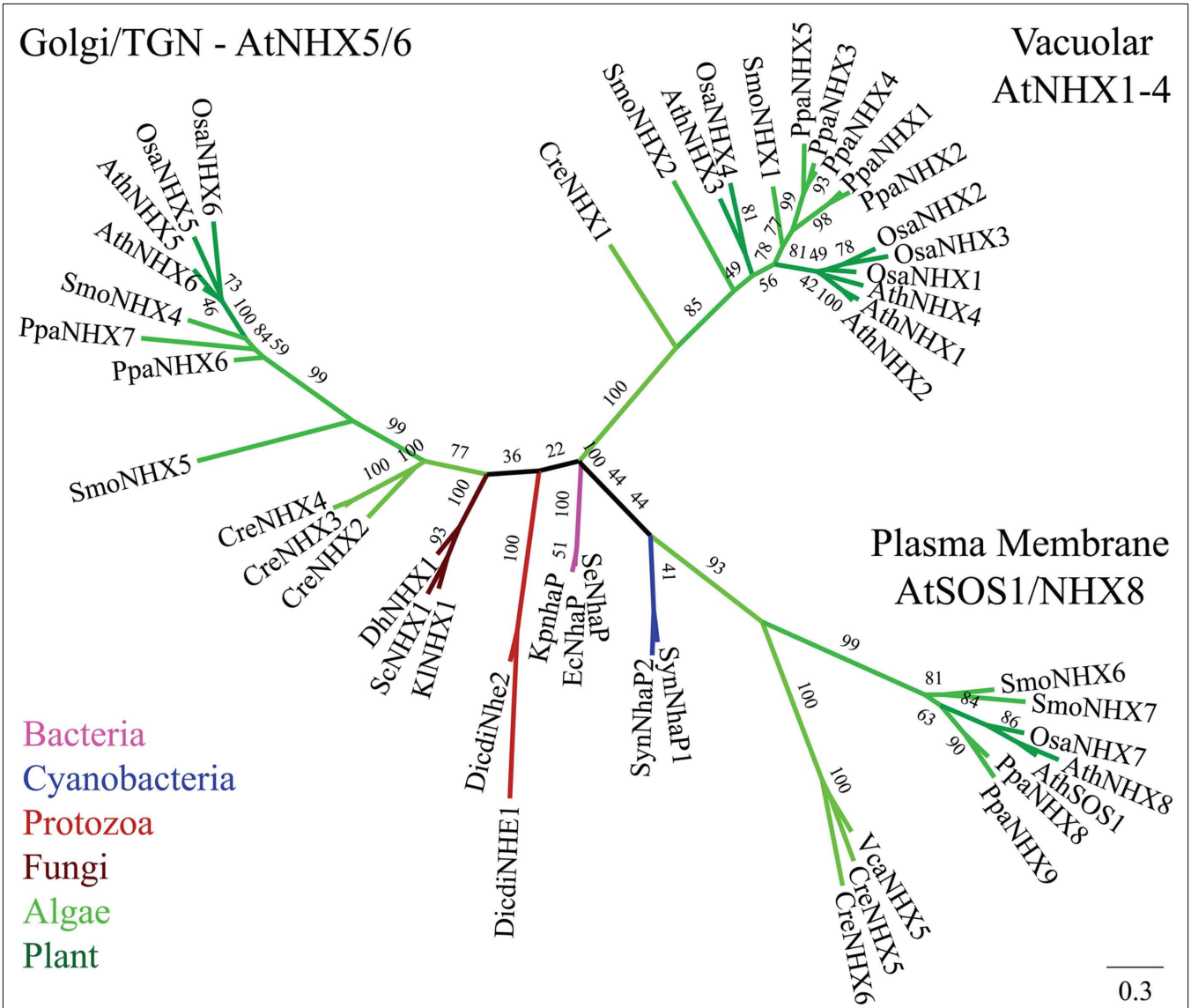


FIGURE 2 | Diversification of CPA1 genes into three clades preceded the evolution of early land plants. The C-tails were trimmed to maximize alignment, so AtNHX1 sequence included residues 1–448. Tree was generated with maximum likelihood of RAxML using the Jones–Taylor–Thornton (JTT) model and tested by bootstrap of 500. (Final ML optimization likelihood = −30244.238457). Single-celled alga (Cre), moss (Ppa), and club moss (Smo) NHX members were present in each of three clades, NhaP/SOS1, endosomal AtNHX5/6, and vacuolar AtNHX1–4. Organisms and proteins are defined in Table S2 in Supplementary Material, and color-coded as bacteria (magenta), Cyanobacteria (dark-blue), fungi (brown), protist (red), green algae (green), early land plants, and angiosperms (dark green). The "vacuolar" AtNHX1–4 clade is specific to plants.

Table 1 | Functions of prokaryote cation/proton antiporters (CPA).

Prokaryote isoform	Transport mode	Membrane location (aa)	Main functions	Plant homolog (CPA)	Reference
EcNhaA	(Na^+, Li^+)/$2H^+$ electrogenic	Plasma membrane (388)	Regulation of cellular acidity and salinity. Activated by external pH above 7	None (CPA2)	Arkin et al. (2007), Hunte et al. (2005)
SynNhaS1, S2 (SynNhaP, P2)	(Na^+, Li^+)/H^+ electroneutral	Plasma membrane (527, 540)	Salt tolerance	SOS1/AtNHX7/8 (CPA1)	Hamada et al. (2001), Inaba et al. (2001)
MjNhaP	(Na^+, Li^+)/H^+	Plasma membrane (426)	Cytoplasmic pH control using Na^+ gradient. Active between pH 6 and 7	SOS1/AtNHX7/8	Hellmer et al. (2002), Vinothkumar et al. (2005), Goswami et al. (2011)
SynNhaS3	(Na^+, Li^+)/H^+	Thylakoid membrane (461)	Salt tolerance. Ion homeostasis K^+, Na^+, H^+ in cytoplasm, and thylakoid lumen	KEA (CPA2)	Tsunekawa et al. (2009)
SynNhaS4	(K^+/H^+) suggested	? (410)	Growth in K^+ depleted conditions	CHX (CPA2)	Inaba et al. (2001)
SynNhaS5	?	? (698)	?	? (CPA2)	
EhNapA	(Na^+, Li^+)/H^+	Plasma membrane (383)	Salt tolerance at neutral pH	(CPA2)	Waser et al. (1992), Strausak et al. (1993)
BcGerN/GerT	(Na^+, Li^+)/(H^+–K^+)	Plasma membrane (387)	Spore germination/Spore outgrowth in alkaline or saline conditions	KEA (CPA2)	Thackray et al. (2001), Southworth et al. (2001), Senior and Moir (2008)
EcKefC	K^+ uniport or K^+/H^+ exchange (Rb^+, Li^+, Na^+)	Plasma membrane (620)	Cytoplasmic acidification coupled to K^+ efflux, survival exposure to electrophiles	KEA (CPA2)	Fujisawa et al. (2007), Ferguson et al. (2000)

"aa" Refers to the number of residues in the protein. Ec, Escherichia coli; Eh, Enterococcus hirae, Bc, Bacillus cereus, Syn, Synechocystis sp. PCC 6803; Mj, Methanococcus jannaschii. (See Table S1 in Supplementary Material for protein accession no.)

NHAP/SOS1 clade

Phylogenetic analysis showed that plant members of Na^+/H^+ antiporter (NHAP) clade arose from ancestral NhaP1 and NhaP2 genes of a cyanobacterium (**Figures 1** and **2**). The best characterized plant NHAP is SOS1/AtNHX7 which mediates electroneutral Na^+/H^+ exchange (Qiu et al., 2003). Although AtSOS1 and AtNHX8 were annotated previously as NHX, our analysis places this group within NHAP clade consistent with results of Brett et al. (2005a). AtSOS1 and its homolog, AtNHX8, are localized to the plasma membrane (Qiu et al., 2003; An et al., 2007). Intriguingly, plant NHAP, such as SOS1/NHX7, has a particularly long C-terminal tail compared to other plant NHX. Recent studies show that this long C-tail interacts with other regulatory proteins (Katiyar-Agarwal et al., 2006; Quintero et al., 2011).

Arabidopsis NHAP or AtSOS1 has been extensively studied since the discovery that *sos1* mutant is salt-overly sensitive (Wu et al., 1996; Quintero et al., 2002; Zhou et al., 2006; Fraile-Escanciano et al., 2010; see **Table 2**). SOS1 was shown to be associated with salt stress response (Shi et al., 2000, 2002) by mediating extrusion of Na^+ (Qiu et al., 2003; Olias et al., 2009). It was shown that SOS1 had a physiological role in protecting plasma membrane K^+ transporters from Na^+ inhibition (Qi and Spalding, 2004). Furthermore, interaction of AtSOS1 with radical-induced cell death (RCD1) prevented RCD1, which is important for oxidative stress response, from nuclear localization (Katiyar-Agarwal et al., 2006). This relationship suggested a role of SOS1 in signal transduction. Analysis of *sos1* mutant revealed the consequence of Na^+ accumulation which affected membrane trafficking, vacuolar biogenesis, and pH homeostasis in roots (Oh et al., 2010). Interestingly, cation specificity for AtSOS1 and AtNHX8 is limited to Na^+ and Li^+ (Qiu et al., 2003; An et al., 2007) consistent with the substrate specificity of prokaryotic NhaP (**Table 1**). Results indicate that the major role of plant SOS1 transporters, like other NHAP members, is to extrude Na^+ and confer tolerance to salt stress.

Two NHX clades

The diversification of NHAP to eukaryote intracellular-NHE (IC-NHE) and PM-NHE occurred early in evolution most likely when primitive eukaryote cells appeared (Brett et al., 2005a; **Figure 1**). Plant NHX members (excluding SOS1) belong to the IC-NHE family. *NHX* genes are present in single-celled algae like *Chlamydomonas*, early land plants, like *Physcomitrella patens*, as well as *Arabidopsis thaliana* (**Figure 2**). Furthermore, plant NHX has diversified into two clades that cluster either with *Arabidopsis* NHX5/6 or with AtNHX1–4. Consistent with the phylogenetic analysis, NHX proteins have been localized to intracellular membranes. In flowering plants, AtNHX5/6 is localized to Golgi or TGN, and AtNHX1 is at the vacuolar membrane (see **Table 2** for references). The functions and localization of plant NHX and NhaP/SOS1 proteins are summarized in **Table 2**. These findings support the idea that NHX genes evolved as endomembranes appeared in early eukaryote and plant cells.

Initial studies showed that vacuolar NHX1 mediated electroneutral Na^+/H^+ exchange (Darley et al., 2000), though later studies broadened cation specificity to K^+ in plants (**Table 2**). Vacuolar NHX was associated with salt stress responses and pH regulation (reviewed by Rodriguez-Rosales et al., 2009). InNHX1/2 of Japanese morning glory was shown to alkalinize vacuolar pH

Table 2 | Functions of CPA1 members (NHXs and SOS1) from *Arabidopsis* and other plants.

Plant isoform	Expression	Membrane location (aa)	Main functions	Reference
AtNHX1	Ubiquitous: epidermis of siliques, inflorescence, stems, leaves, petals induced by salt, osmolarity, ABA	Tonoplast (538)	$Na^+(K^+)/H^+$ antiport salt tolerance K^+ sequestered in vacuole vacuolar pH alkalinization cell expansion, interact with calmodulin (AtCAM15) regulating Na^+/K^+ selectivity	Apse et al. (1999, 2003), Gaxiola et al. (1999), Quintero et al. (2000), Shi et al. (2002), Venema et al. (2002), Yokoi et al. (2002), Yamaguchi et al. (2003, 2005), Leidi et al. (2010), Bassil et al. (2011b)
AtNHX2	Ubiquitous, induced by salt, osmolyte, ABA	Tonoplast (546)	$Na^+(K^+)/H^+$ antiport, K^+ sequestered in vacuole, vacuolar pH alkalinization, cell expansion, salt tolerance	Yokoi et al. (2002), Bassil et al. (2011b)
AtNHX3	Roots, germinating seeds, flowers, and siliques	Tonoplast (529)	K^+/H^+ Low K^+ tolerance	Yokoi et al. (2002), Liu et al. (2010)
AtNHX4	Ubiquitous: stem, induced by high Li^+ or K^+, ABA	Tonoplast (503)	Increase sensitivity to salt	Yokoi et al. (2002), Li et al. (2009)
AtNHX5/ LeNHX2	Ubiquitous, induced by salt	Golgi/TGN (521)	Vesicular trafficking, salt tolerance	Yokoi et al. (2002), Venema et al. (2003), Bassil et al. (2011a)
AtNHX6	Ubiquitous	Golgi/TGN (535)	Vesicular trafficking, salt tolerance	Yokoi et al. (2002), Bassil et al. (2011a)
AtSOS1 (AtNHX7)	Root-, stem- and leaf-xylem parenchyma, and epidermal cells of the root tip, induced by NaCl	Plasma membrane (1146)	Na^+/H^+, salt tolerance, Na^+ transport from root to shoot, protect K^+ permeability, regulated by protein kinase SOS2 and Ca^{2+} sensor SOS3 at C terminus, interact with RCD1 regulating oxidative stress response	Shi et al. (2000, 2002), Qiu et al. (2002), Quintero et al. (2002, 2011), Qi and Spalding (2004), Katiyar-Agarwal et al. (2006)
AtNHX8	Ubiquitous	Plasma membrane (756)	Li^+ tolerance	An et al. (2007)

"aa" Refers to number of residues in protein. At, Arabidopsis thaliana; Le, tomato.

causing purple buds to turn into blue flowers (Yamaguchi et al., 2001; Ohnishi et al., 2005). The increase in vacuolar pH was accompanied by K^+ accumulation indicating that NHX1 mediated K^+/H^+ exchange in *Ipomea tricolor* (Yoshida et al., 2009). Cation specificity was directly shown using proteoliposome where AtNHX1 catalyzed Na^+/H^+ and K^+/H^+ exchange with similar affinity (Venema et al., 2002). Moreover, silencing *LeNHX2* resulted in tomato plants that accumulated less K^+ and were susceptible to salt stress (Rodriguez-Rosales et al., 2008). In contrast, over-expression of AtNHX1 in tomato resulted in enhanced K^+ accumulation and resistance to salt (Leidi et al., 2010). Finally, *nhx1/nhx2* double mutants displayed reduced K^+ concentration and pH in vacuoles, supporting a role of AtNHX1 and AtNHX2 in mediating H^+ efflux coupled to K^+ uptake. Thus regulating intravacuolar (K^+) and pH are essential to cell expansion and flower development (Bassil et al., 2011b). There is increasing acceptance that plant NHX has a major role in pH and K^+ homeostasis in plant cells where salt tolerance is conferred by enrichment of intracellular K^+ pool and sequestration of Na^+ in vacuole.

Interestingly, endosomal NHXs, such as LeNHX2, AtNHX5, and AtNHX6 were able to confer salt tolerance (Yokoi et al., 2002; Rodriguez-Rosales et al., 2008). Double mutants of *nhx5/nhx6* were hypersensitive to salt (Bassil et al., 2011a). In previous studies of yeast, evidence emerged that ScNHX1 affected endosomal pH and K^+ homeostasis that influenced protein sorting and membrane trafficking (Bowers et al., 2000; Ali et al., 2004; Brett et al., 2005b; Wagner et al., 2006). A recent study with plants suggested that regulation of endosomal pH and K^+ homeostasis by AtNHX5 and AtNHX6 had a role in sorting of vacuolar protein and cellular stress responses (Bassil et al., 2011a), consistent with the phylogenetic analysis showing ScNHX1 and AtNHX5/6 are in the same clade (**Figures 1** and **2**).

Curiously, budding yeast (and probably most fungi) has only one endosomal NHX1 localized to endosome/prevacuolar compartment (PVC), and does not have genes related to "plant vacuolar-type" AtNHX1–4. However, *Dictyostelium* Nhe1 and Nhe2 are at the base of the plant vacuolar NHX clade (**Figure 2**), suggesting plant vacuolar NHX arose from a gene in an ancestral protozoa possibly with contractile vacuoles for osmoregulation (Allen and Naitoh, 2002; Gerisch et al., 2002; Uchikawa et al., 2011). The plant-specific AtNHX1–4 type has apparently undergone gene duplications (**Figure 2**) underscoring the significance and diversity of vacuole functions in multicellular land plants.

MONOVALENT CATION PROTON ANTIPORTER-2

Members of CPA2 superfamily are found in bacteria, fungi, and plants; however, they are rare in metazoa (Brett et al., 2005a). They are predicted to have 8–14 membrane-spanning domains with a Pfam00999 domain for $Na^+(K^+)/H^+$ exchange. Interestingly,

substrate and transport mode of CPA2 from bacteria vary among family members. The general transport mode of CPA2 is $M^+(in) + nH^+(out) \rightleftharpoons M^+(out) + (in)$ for the carrier-mediated mode, whereas some members may catalyze $M^+(in) \rightleftharpoons M^+(out)$ in the channel-mediated mode (Saier, 2000). The CPA2 superfamily consists of three families named here as NHA/NhaA, KEA/Kef, and CHX/NhaS4 based on phylogenetic analyses and prokaryote ancestry.

NHA clade

A eukaryote NHA clade was previously classified within CPA2 (Brett et al., 2005a), though they seem related to the NHAP clade in CPA1 (**Figure 1**). Eukaryote NHA, like NhaP, is localized at the plasma membrane (Kinclova et al., 2001). Functional studies of yeast show that plasma membrane ScNHA1 mediates both Na^+/H^+ and K^+/H^+ exchange (Banuelos et al., 1998; Sychrova et al., 1999; Ohgaki et al., 2005). Human NHA2 is localized to the PM and its expression in yeast conferred tolerance to Na^+ and Li^+, but not to K^+ (Xiang et al., 2007), suggesting it mediates Na^+ transport similar to NhaP or NHX of CPA1. Our phylogenetic analysis suggests that eukaryote NHA is likely evolved from an *NhaP* gene of an ancestral bacterium (**Figure 1**). Unlike NhaP, *E. coli* NhaA is active at pH 8 and inactive at pH 6 (Taglicht et al., 1991; see **Table 1**). EcNhaA mediated exchange of $2H^+$ for $1Na^+$ or $1Li^+$ (Padan, 2008). Intriguingly, only tiny marine green alga, such as *Ostreococcus lucimarinus* and *Micromonas pusilla* (Chlorophyta) have a protein homologous to bacterial NhaA (**Figure 1**; Table S1 in Supplementary Material). NHA members are thus found in bacteria, animals, and fungi (Brett et al., 2005a), but not in early land plants or flowering plants (**Figure 1**).

KEA family

Genes encoding putative KEA from higher plants were first classified in *Arabidopsis thaliana* (Maser et al., 2001); thus AtKEA1–AtKEA6 members are used as a reference. KEA-like transporters from bacteria have been extensively studied, and are characterized by an N-terminal Na_H exchanger domain (Pfam PF00999), and a C-terminal KTN NAD(H)-binding domain, also called TrkA_N domain (Pfam PF02254; **Figure 4**). For example, Kch, TrkA, Ybal, and KefB or KefC are K^+ channels or transporters with a C-terminal KTN domain (Choe, 2002). The minimal functional unit of the KTN domain is a dimeric molecule connected by a flexible hinge. Hinge motion can be physically coupled to transmembrane loops to control K^+ flux, providing a mechanism of gating control (Choe, 2002). The domain contains the typical Rossman fold GXGXXG. . .D Glycine motif involved in NAD binding, indicating metabolic control of transporter function (Jiang et al., 2001; Roosild et al., 2002).

Divergence of two KEA clades. To understand the evolutionary origin of higher plant KEA, we searched and analyzed homologs in prokaryotes, protists, metazoa, and plants. AtKEA genes diverged into two main branches with AtKEA1–3 in clade I and AtKEA4–6 in clade II. Clade I consists of the only eukaryote-encoded proteins with a complete C-terminal KTN domain (**Figures 3** and **4**), and are closely related to EcKefB and EcKefC-like proteins from bacteria, including cyanobacterium *Gloeobacter violaceus* (**Figure 3**).

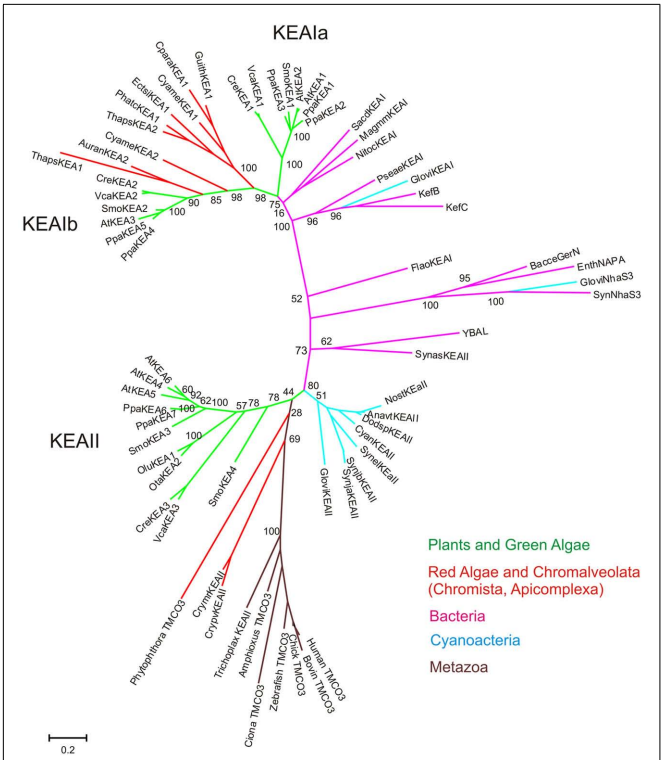


FIGURE 3 | Phylogenetic tree shows plant KEAs originated from ancestral genes in Cyanobacteria. The conserved Pfam 00999 domains in KEA sequences from diverse organisms were aligned by MUSCLE and the evolutionary history of the genes were determined by maximum likelihood using the Jones–Taylor–Thornton (JTT) model with a Bootstrap of 100. The percentage of trees in which the associated sequences clustered together is shown next to the branches. The analysis involved 66 amino acid sequences. All positions containing gaps and missing data were eliminated, resulting in a total of 307 positions in the final dataset. Branch colors refer to bacteria (pink), Cyanobacteria (blue), red alga (red), metazoa (brown), and plants (green). Organisms and accession numbers of protein sequences are listed in Table S3 in Supplementary Material.

Red algae as well as secondary photosynthetic organisms of the Kingdom Chromista contain sequences from clade I exclusively (**Figure 3**; Table S3 in Supplementary Material). In contrast, proteins of clade II lack a KTN domain (**Figures 3**, **4**, and **6**; Table S3 in Supplementary Material). Phylogenetic analysis suggests AtKEA4–6 arose from an ancestral cyanobacterium judging by the presence of related proteins in unicellular and filamentous cyanobacteria (**Figure 3**). Clade II genes are also present in parasitic organisms of the phylum Apicomplexa. These organisms contain a high number of plant and bacterial-like proteins (Abrahamsen et al., 2004) possibly due to apicoplast (plastid) resulting from secondary endosymbiont of an ancestral green or red alga (Keeling, 2004). A clade II protein is found in the oomycete, *Phytophthora infestans*, possibly also related to the presence of a plastid in a common photosynthetic ancestor of the chromista (Tyler et al., 2006). Oddly, AtKEA4–6 like proteins are also found in *Trichoplax adhaerens* (a Placozoa), *Ciona intestinalis*, a basal Chordata, and in higher Chordata, such as mammals, amphibians, and fish (**Figure 3**, Table S3 in Supplementary Material). These AtKEA4–6 like proteins contain

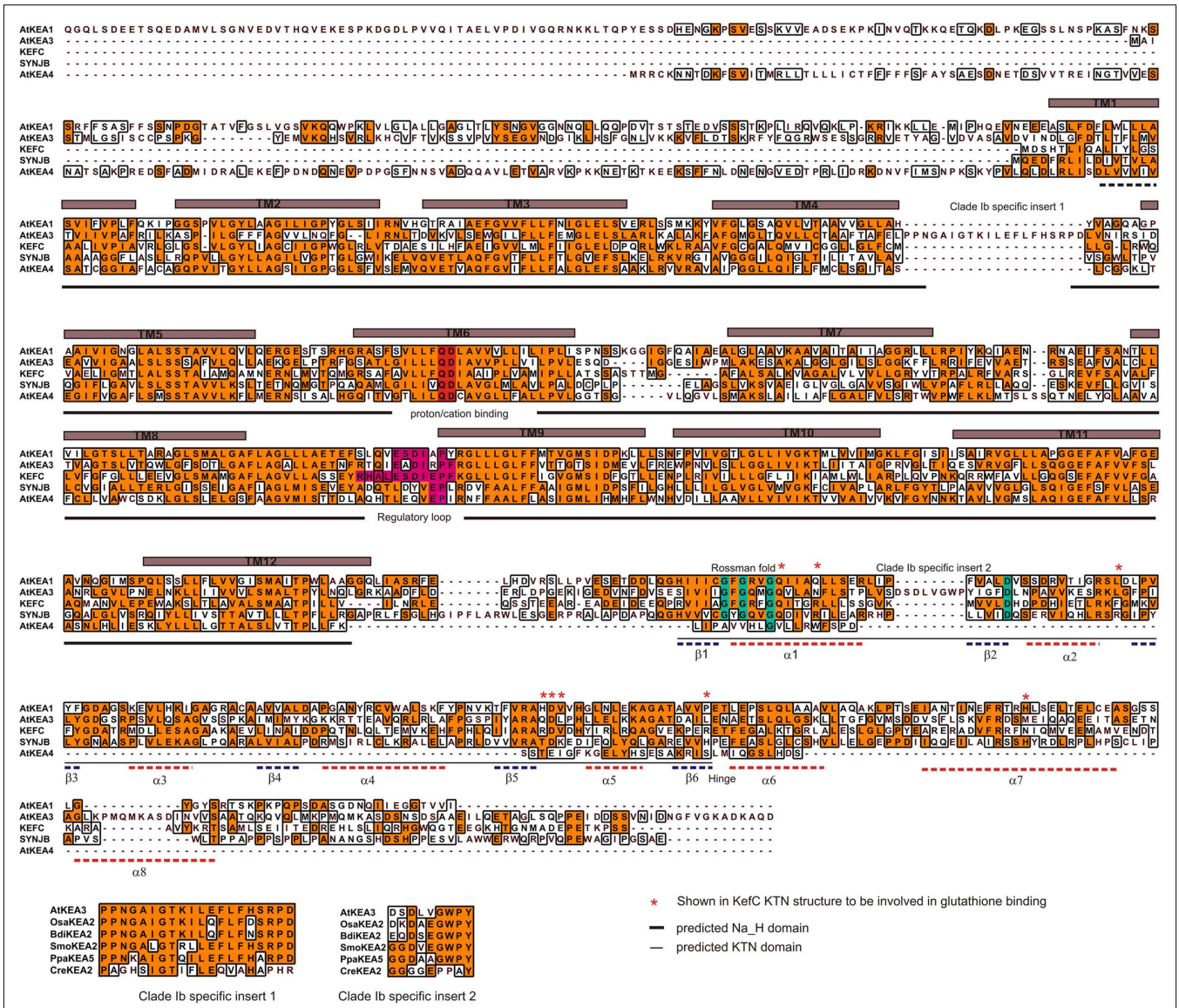


FIGURE 4 | Alignment of AtKEA sequences with their bacterial homologs – AtKEA sequences were aligned using MUSCLE (Edgar, 2004) with their closest prokaryotic homologs: *E. coli* KefC and Cyanobacterial sequence SynjbKEAII (Q2JI45 Table S3 in Supplementary Material). The soluble N-terminal domain of AtKEA1 was removed. Conserved residues are shown in orange. Approximate positions of possible transmembrane helices are shown above the alignment. Secondary structure of the KTN domain is indicated as well as residues that were shown to be involved in Glutathione-binding in the KefC KTN structure. A short cytoplasmic regulatory loop in the E. coli KefC sequence as well as the conserved residues in the *Arabidopsis* sequences are shown in magenta. The residues corresponding to the proton binding residues in the structure of the bacterial NhaA sequence are shown in red. The Rossman fold glycine motif is shown in green. Below the alignment the two specific clade Ib inserts in different plants are shown.

a trans membrane coiled coil protein 3 (TMCO3) and also lack the KTN domain.

These results would suggest that KEA genes were acquired by primary endosymbiosis of a cyanobacterium, and by secondary endosymbiosis in the Archaeplastida ancestor of red and green algae (Adl et al., 2005; Bowman et al., 2007). Furthermore, it appears the divergence of clade I and clade II KEAs occurred early as prokaryotes diversified perhaps before eukaryote cells evolved.

Intriguingly, KEA genes from clade I are found in a red alga (*C merolae*) and in chromista (**Figure 3**). Chromista possibly became photosynthetic by endosymbiosis with a unicellular red alga, as genes related to KEA sequences in *C. merolae* are also present in these species. In chromista of the cryptomad subgroup, a KEA gene constitutes one of three transport genes that are maintained in the highly reduced nucleomorph genome (Douglas et al., 2001). Nucleomorphs are the remnant nuclei of algal endosymbionts that were engulfed by non-photosynthetic eukaryotes (Tanifuji et al., 2011). This observation suggests that KEAs play important roles in these organisms, such as ion transport across one of four chloroplast membranes. Its role is apparently not in photosynthesis, as KEA-like gene is conserved in the

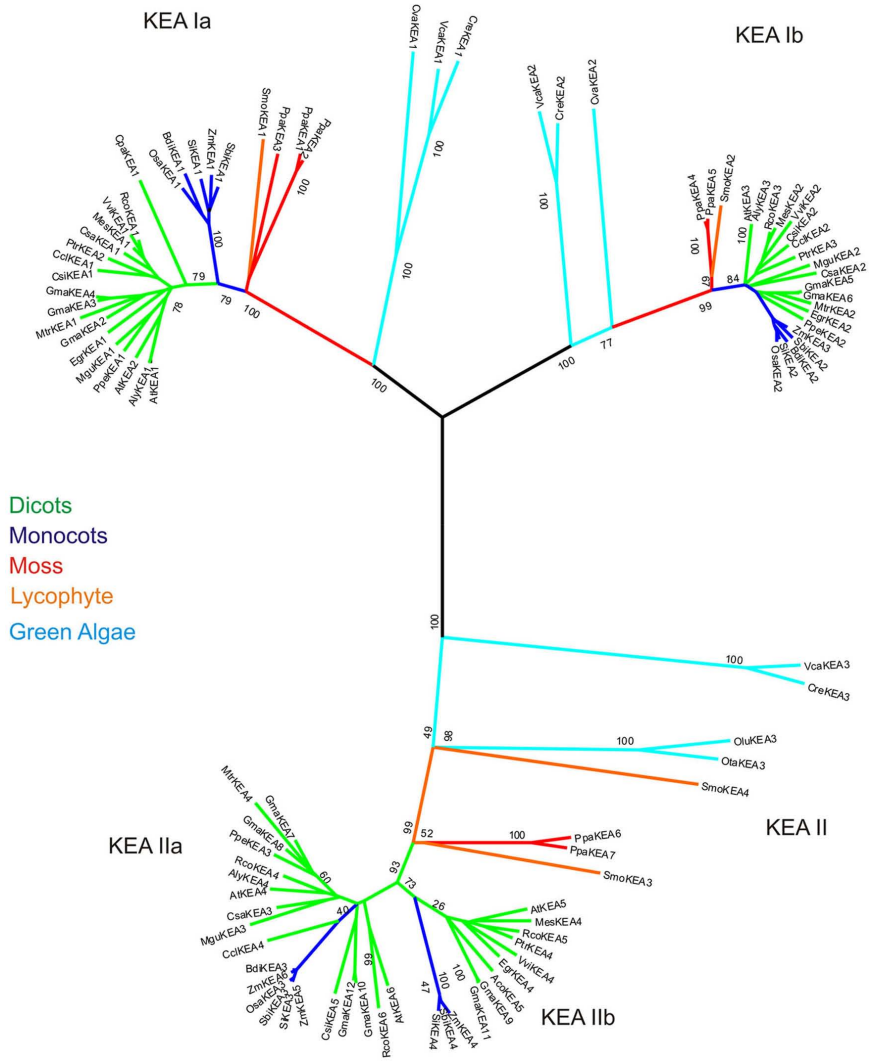


FIGURE 5 | K⁺-efflux antiporter diversified into three types in early photosynthetic eukaryotes. The analysis involved 97 amino acid sequences. The tree includes all proteins found in *A. thaliana, G. max, O. sativa, Z. mays, S. moellendorffii, P. patens, C. reinhardtii,* and *V. carteri,* as well as one representative sequence for each group from the remaining plant species listed in Table S3 in Supplementary Material. All positions containing gaps and missing data were eliminated. There were a total of 152 (335 excluding incomplete sequences GmaKEA11 and RcoKEA6) positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011). Branch colors refer to dicots (green), monocot (dark-blue), early land plants (orange–red), and green algae (light blue). Species and accession numbers of protein sequences are listed in Table S3 in Supplementary Material.

nucleomorph of a non-photosynthetic *Cryptomonas paramecium*, but not in nucleomorphs of green algal origin (Tanifuji et al., 2011).

Evolution of KEA in plants. Protein phylogenetic analyses showed that all plants, from unicellular algae to flowering plants, had three types of KEAs classified as clades Ia, Ib, and II (**Figure 5**). Green algae, like *C. reinhartdii* and *V. carteri*, contain only one gene in each group. In other plants, diversification occurred in clades KEA-Ia, KEA-II, or both.

Clade Ia KEAs possess a long N-terminal domain. Phylogenetic analysis showed that plant KEAs in clade I diverged into two types that differ in their protein organization. Within clade Ia, all plant sequences from moss to monocot and dicots have an long hydrophilic N-terminal domain of ∼570–770 amino acids (**Figure 6**) that is often missed in preliminary gene annotations. This domain is typically rich in negatively charged glutamate residues, and predicted coiled coil structures (Figure S4 in Supplementary Material). The role of this unusual extra sequence is unknown (Rose et al., 2004). The N-terminal domain of green algae Cre and Cva share some similarity (46%) and are slightly shorter than that of higher plants (**Figure 6**). The absence of this region in prokaryotes would indicate that the N-terminal domain is a recent acquisition that gained increasing relevance in land plants.

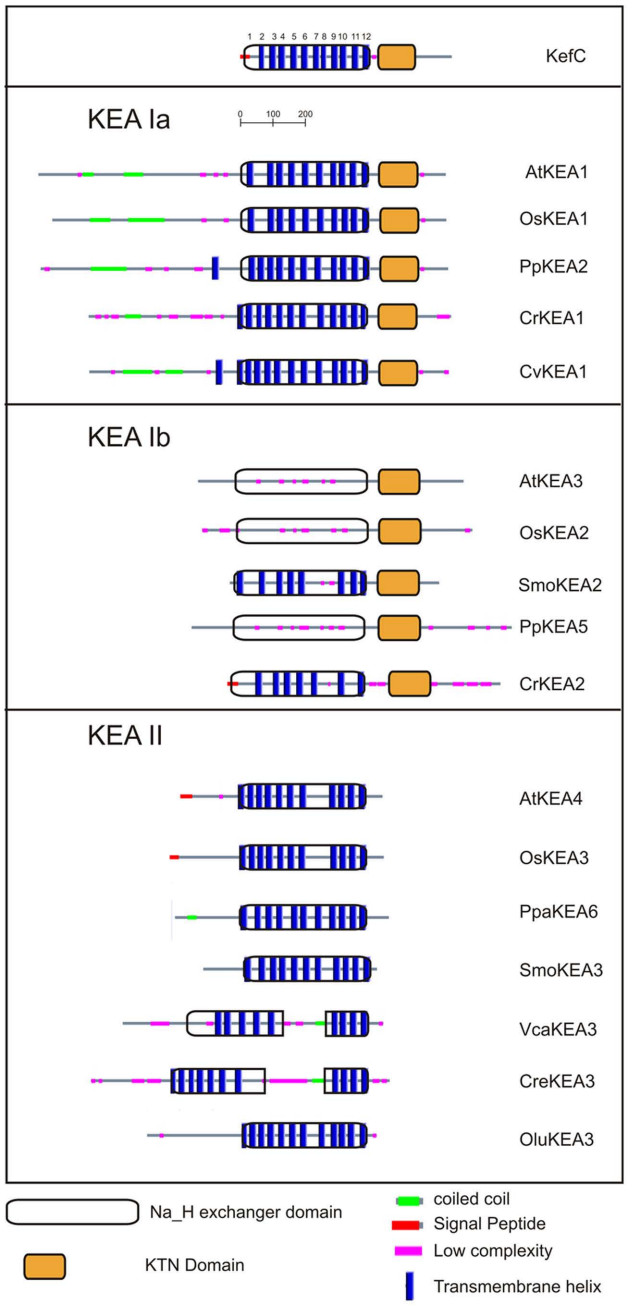


FIGURE 6 | Distinct protein domains of three KEA clades. The relative length and position of the KTN and Na_H exchanger domain in the three KEA clades is shown graphically using information obtained with the Simple Modular Architecture Research Tool (Smart, http://smart.embl-heidelberg.de/). All plant clade I proteins contain a KTN domain. In the clade II group this domain is lost, making it the shortest KEA proteins. The clade Ia proteins all have acquired a long N-terminal domain, typically containing predicted coiled coil structures. The C-terminus of the clade Ib proteins is slightly longer as compared to the other sequences. Some green algal clade II proteins have a split Na_H domain, which might be a result of erroneous gene predictions. The transmembrane helix predictions are very variable, and especially unsuccessful for the clade Ib group of proteins. The transmembrane helices shown in **Figure 2** are numbered from 1 to 12 in KefC.

In contrast, proteins in clade Ib, such as AtKEA3, do not harbor the N-terminal extension, instead they have gained specific insertions between predicted membrane helices 4 and 5 and in the Rossman fold glycine motif of the KTN domain (**Figures 4** and **6**). Although all hydrophobic regions are conserved in the clade Ib proteins, transmembrane helices are poorly identified by transmembrane prediction programs (**Figure 6**). In general, only one gene of this type is present per plant, except in *G. max* and *P. patens*, where two genes are present (**Table 4**; Table S3 in Supplementary Material; **Figure 5**).

Recent diversification of clade II KEAs. Genes encoding AtKEA4–6 have a hydrophilic N-terminal domain of ∼157–235 residues that is predicted to include a signal peptide (**Figures 4** and **6**). Green algal sequences found in *V. carteri* and *C. reinhartdii* have an anomalous Na_H exchanger domain that is interrupted by a similar sequence in both species (**Figure 5**). Interestingly, clade KEA-II has diversified further in angiosperms where AtKEA5 is clearly separate from AtKEA4 and AtKEA6 indicating this occurred recently after the evolution of early land plants.

Function of KEA. Nearly nothing is known about plant KEA function, so bacterial homologs could provide insights. KefB and KefC in *E. coli* are glutathione regulated K^+ efflux transporters, that are important for bacterial survival during exposure to toxic metabolites (Booth, 2003). Electrophiles react with glutathione to form glutathione adducts, thereby releasing KefC and KefB inhibition by GSH, and eliciting K^+ efflux and H^+ uptake. Cytoplasmic acidification is suggested to prevent DNA damage (Ferguson et al., 2000; Miller et al., 2000). Reduced GSH stabilizes the association of the two KTN domains, whilst the larger adducts disrupt this interaction (Roosild et al., 2010). Based on mutagenesis studies, a short central hydrophilic regulatory loop was proposed to interact with the KTN domain. This regulatory loop is partially conserved in the plant KEA1–3 sequences, but the residues forming the glutathione-binding pocket are less conserved in the plant sequences (**Figure 2**), indicating that plant KEA1–3 are regulated in a different way (Roosild et al., 2010). Although originally proposed to function as K^+ channels, recent studies indicate KefC has K^+/H^+ antiport activity (Fujisawa et al., 2007).

Recently, AtKEA1 and AtKEA3 peptides were detected by mass spectroscopy in *Arabidopsis* chloroplast preparations (Zybailov et al., 2008). Moreover, localization to thylakoid membrane of SynNha3 (Tsunekawa et al., 2009), a *Synechocystis* protein closely related to KEA-I and KEA-II proteins (**Figure 1**), would suggest a role of plant KEA in chloroplast ion homeostasis. So far only one cyanobacterium, *Gloeobacter violaceus*, has a gene encoding a KEA-I homolog (**Figure 3**). Perhaps ancestral primary plastids originated just after the separation of primitive cyanobacterium, like *Gloeobacter*, from present-day lineages (Criscuolo and Gribaldo, 2011). In *Gloeobacter*, thylakoids are absent and photosystems reside at the plasma membrane (Vothknecht and Westhoff, 2001). We hypothesize that the long N-terminal sequence of AtKEA1/2 has a role in chloroplast-specific processes in plant (Vothknecht and Westhoff, 2001). These and other possibilities need to be tested.

Evolutionary origin of CHX clade in plants

The CHX clade is populated by genes from higher plants, but none from metazoa so far. Nearly all deduced proteins from plants have ∼800 residues (Table S4 in Supplementary Material) that consist of a hydrophobic Na^+/H^+ exchanger domain of ∼400 residues at the amino terminus, and a carboxyl hydrophilic domain of 300–400 residues. The long hydrophilic C-tail does not contain any known conserved motifs, so its role is a complete mystery. NhaS4 from the cyanobacterium *Synechocystis* PCC 6803 encodes a much shorter protein of 410 residues, though the predicted protein from *Synechococcus elongatus* PCC 6301 has 715 residues.

Based on analyses using the TM domain alone, we show that the present-day NhaS4 from cyanobacteria, shares the highest similarity with AtCHX16–CHX20 (**Figure 1**; Sze et al., 2004). NhaS4 gene is found in single-celled *Synechocystis* PCC 6803 (Inaba et al., 2001) as well as filamentous *Anabaena variabilis* (Ava1632). Many non-cyanobacteria do not have genes that cluster with the CHX/NhaS4 clade. One exception is MyxxaNhaS4 (Table S1 in Supplementary Material) from the gliding bacterium *Myxococcus xanthus* (a member of Myxococcales) which form fruiting body in response to nutrient starvation (Mauriello et al., 2010). Intriguingly, the mechanisms of motility and establishment of cell polarity in *M. xanthus* and eukaryotes share several similarities (Leonardy et al., 2010). Although genes from metazoa are absent from this clade, *Dictyostelium discoideum* (a protist and slime mold) has two members, Nhe3 and Nhe4, of unknown function that cluster with NhaS4 (**Figure 1**). The function and membrane location of cyanobacterial NhaS4 are still unclear. Interestingly, all fungi examined have a single gene *KHA1* that encodes a homolog of *Synechocystis* NhaS4 as seen in **Figure 1**. Fungi include unicellular *Saccharomyces cerevisiae*, and filamentous fungi such as *Neurospora crassa*, *Aspergillus oryzae*, and *Ustilago maydis* (**Figure 1**; Table S1 in Supplementary Material). Mutants of *ScKHA1* have served as a useful heterologous system to test functions of *Arabidopsis* CHX16–20 (Chanroj et al., 2011; **Table 3**).

Comparative analyses of full-length protein sequences revealed additional insights on the origin of plant CHX members. One CHX homolog was found in *Spirogyra pratensis*, a Zygnematales, though none was detected in *Chlamydomonas reinhartdii*, *Volvox carteri* (**Figures 1** and **7**), or red algae (not shown). Early land plants, like *Physcomitrella* and *Selaginella*, had four and three genes, respectively encoding CHX homologs (**Figure 1**). Interestingly, all these proteins cluster in one subclade closest to *Arabidopsis* CHX20 (**Figure 7**) and CHX16–19. Results suggest that AtCHX16–CHX20 genes arose early in CHX gene family diversification. In *Arabidopsis*, these genes are expressed in leaves and roots, especially in dermal tissues, such as root epidermis (Cellier et al., 2004; Chanroj, 2011), root cap, and guard cells (Padmanaban et al., 2007; see **Table 3**).

Multiple subclades within the CHX family. Analysis of CHX homologs from 14 higher plants unveiled the diversification of this family in flowering plants (**Figure 7**). In general, CHX genes diverged less in monocots than in dicots. First, the average number of CHX genes (16) per monocot genome is less than that seen in many dicots (>26; **Table 4**). Second, **Figure 7** shows that CHX members in *Arabidopsis* and other dicots can be separated into eight subclades, and five of these are conserved in monocots (**Table 5**). For example, orthologs of AtCHX20 and AtCHX16–CHX19 were evident in rice, corn, and sorghum, as shown by the blue branches in **Figure 7** (Table S4 in Supplementary Material). Furthermore, monocot orthologs of AtCHX15, AtCHX1/2, and AtCHX28 were suggested by their close association with these subclades. Many of these genes in *Arabidopsis* are preferentially expressed in pollen (Sze et al., 2004; **Table 5**). Oddly, homologs of AtCHX21/23 were not obvious in monocot

Table 3 | Functions of characterized CHX genes (CPA2) in *Arabidopsis thaliana*.

Plant isoform	Tissue expression	Membrane localization (aa)	Main functions and phenotype	Reference
AtCHX13	Pollen, induced by K^+ starvation in root tip and elongation zone	Plasma membrane in plant and yeast cells (832)	K^+ uptake in yeast and plant cells, restore growth of K^+-uptake deficient yeast, *chx13* mutant plants are sensitive to low K^+, essential for K^+ acquisition	Zhao et al. (2008), Sze et al. (2004)
AtCHX17	Pollen (microspore), root epidermis, root cortex, leaf, induced by salt stress, K^+ starvation, acidic pH, ABA	PVC in plant, PM in plant, endomembrane in yeast (820)	K^+ transport and pH homeostasis, *chx17* mutants have less K^+ in roots when K^+ starved, restore growth of *kha1* yeast at alkaline pH, membrane trafficking: reduce CPY secretion in KTA40–2 yeast; confer HygB tolerance in *kha1* mutant	Cellier et al. (2004), Sze et al. (2004), Maresova and Sychrova (2006), Chanroj et al. (2011), Chanroj (2011)
AtCHX20	Guard cells, root tip/cap	Reticulate endomembrane (842)	K^+ transport and pH homeostasis, guard cell movement and osmoregulation, *chx20* mutants show reduced stomatal aperture, pH homeostasis in yeast	Padmanaban et al. (2007), Chanroj et al. (2011)
AtCHX23	Pollen grain, pollen tube	Reticulate endomembrane: ER in pollen tube (867)	K^+ transport (E coli), *chx21chx23* mutant pollen tube failed to target the ovule, signaling: unresponsive to female cues	Sze et al. (2004), Lu et al. (2011)

(aa) refers to total residues in protein.

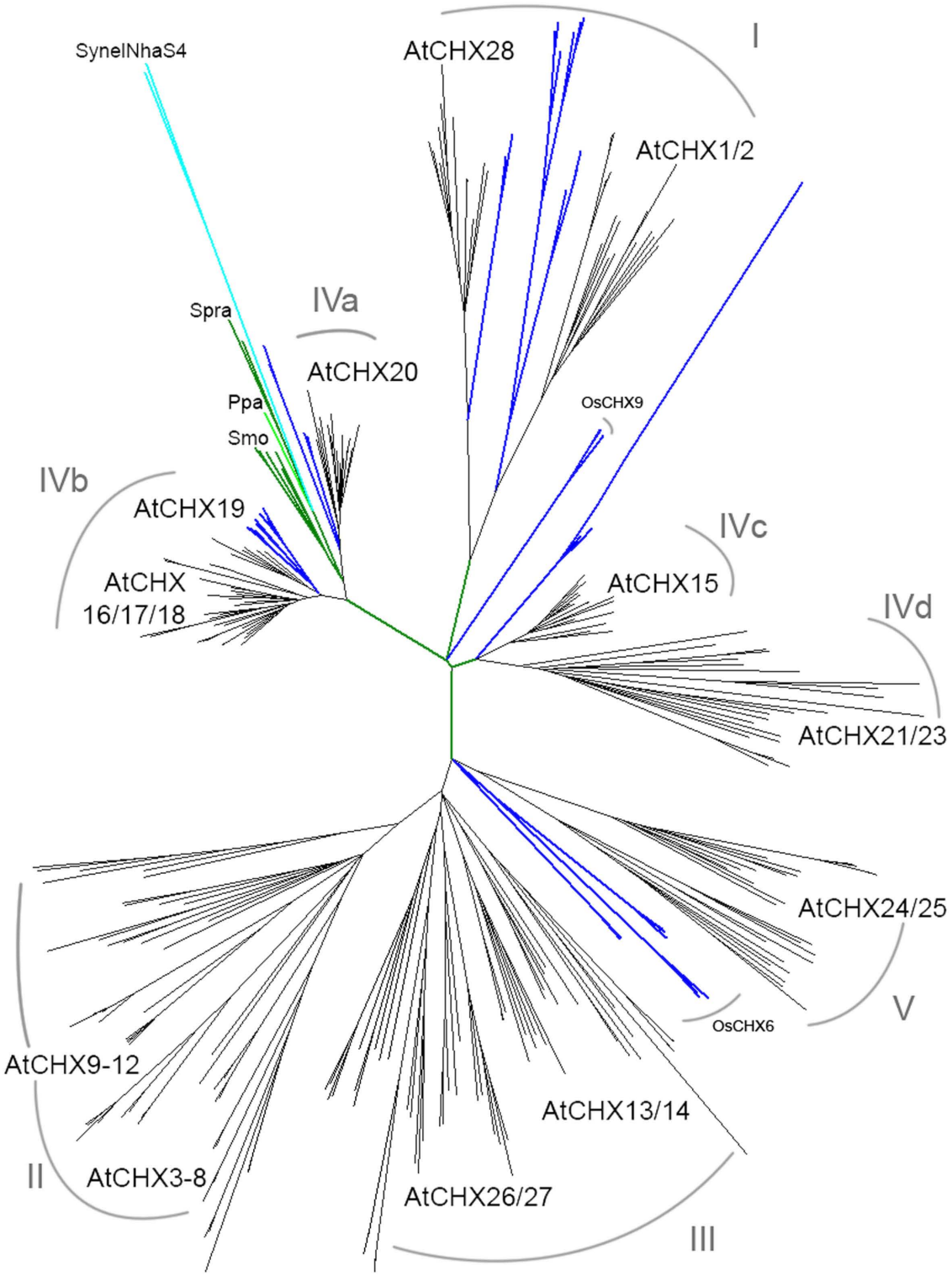


FIGURE 7 | CHX homologs diversified from moss to flowering plants. Full-length CHX proteins from 15 species were aligned and then analyzed by maximum likelihood. *Spirogyra* (Spra), moss (Ppa), and club moss (Smo; green) CHXs clustered with AtCHX20 (subclade IVa). Both monocot (blue) and dicot (black) plants had orthologs of AtCHX20, AtCHX16–19 (IVb), AtCHX28 (I), AtCHX1/2 (I), and AtCHX15 (IVc). Corn, sorghum, or rice had two to four genes encoding OsCHX6 homologs in a monocot cluster near AtCHX24/25. Most dicots had additional CHX homologs that clustered with AtCHX24/25 (V), 26/27, and 13/14 (III). AtCHX3–12 proteins (subclade II) are specific to *A. thaliana* and *A. lyrata*. A cluster of CHX resulting from multiple gene duplications is specific to *Medicago truncatula* (Mtr). See summary in **Table 5**. Species, protein accession numbers and protein properties are described in Table S4 in Supplementary Material, and unrooted rectangular tree is in Figure S4 in Supplementary Material.

genomes. However, rice had four other genes (OsCHX5–8) that clustered near AtCHX24/25, though none were closely related to 16 additional *Arabidopsis* gene products, CHX3–12, 13/14, or 26/27. Three to four CHX genes from other monocot plants including *Brachypodium*, a model of an energy grass, sorghum, and corn also clustered with rice OsCHX5–8, suggesting these represent a monocot subclade of CHX homologs (**Figure 7**).

In contrast, some dicot plants shared additional genes that are potential orthologs of AtCHX24/25, 13/14, and 26/27. Other dicot CHXs from poplar, *Citrus clementina*, *Mimulus,* and *Manihot esculenta* (cassava) formed clusters that were separate from AtCHX13/14 or 24/25, 26/27 subclades, suggesting these genes had diverged later in speciation. Notably, AtCHX3–8 and AtCHX9–12 proteins were in general specific to *Arabidopsis* genus, indicating diversification of CHX genes during *Arabidopsis* speciation. *Medicago truncatula* also had a cluster of 12 genes that could be a result of additional duplications during speciation.

Functions of plant CHX proteins. Functional studies of plant CHX proteins are just beginning and so far limited to *Arabidopsis thaliana*. According to results from the founding members of the family, AtCHX16–AtCHX20, these proteins are implicated in K^+ and pH homeostasis of dynamic endomembranes. First, expression of AtCHX17 or its homologs CHX16–CHX20, rescued alkaline-sensitive growth phenotype of the host yeast strain (Chanroj et al., 2011). Second, CHX20 caused acidification and alkalization of the cytosol and vacuole of yeast, respectively; although CHX17 did not alter cytoplasmic or vacuolar pH (Chanroj et al., 2011). Third, AtCHX17 restored growth of bacteria defective in several K^+ uptake systems, and tracer studies demonstrated that CHX17, CHX20, and CHX23 mediated monovalent cation transport with a preference for K^+ over Na^+ (**Table 3**). Thus CHXs, including plasma membrane-localized CHX13, transport K^+ (Zhao et al., 2008), though their ability to restore yeast growth depends on different external pH's. Interestingly, these findings of plant CHX are consistent with mutant studies that suggest *Synechocystis* NhaS4 probably transported K^+, but not Na^+ (Inaba et al., 2001).

The association of AtCHX proteins with endomembranes and their role in pH and cation homeostasis would suggest they perform important roles in membrane trafficking events that affect protein and cargo sorting. Two lines of evidence support this idea. First, AtCHX17, AtCHX18, or AtCHX19, but not CHX20, conferred tolerance to hygromycin B in yeast lacking multiple cation-handling mechanisms (Chanroj et al., 2011). Sensitivity to aminoglycosides in yeast has been related to altered intracellular pH homeostasis and membrane trafficking. Second, Chanroj et al. (2011) showed that yeast mutants expressing CHX17 secreted less vacuolar-destined carboxypeptidase Y (CPY) to the external medium similar to cells expressing NHX1. Thus cation and pH homeostasis in endosomes, such as PVC, appeared to influence sorting and trafficking of proteins in the endomembrane system. Based on these findings and analysis of *Atchx20* mutants (Padmanaban et al., 2007), we propose that CHX20 affects endomembrane dynamics and osmoregulation needed for guard cell swelling and stomatal opening.

The multiplicity of CHX genes in *Arabidopsis* pollen and possibly other plants is surprising and not understood. The first study on pollen CHX proteins showed that AtCHX21 and AtCHX23 are functionally related and that they are involved in either the reception or the transduction of female signals that target pollen tube to the ovule (Lu et al., 2011). Pollen carrying double mutant *chx21chx23* developed into mature grains, extended a tube into the transmitting tract, however the tube failed to find the egg (Lu et al., 2011). Pollen tube elongates by tip growth and failure to shift the axis of polarity in the double mutant suggests that localized cation and pH homeostasis by CHX has a role in the establishment, or maintenance of polarity or both. Homologs of *Arabidopsis* CHX21 and CHX23 are present in dicot, but not in monocot, plants (**Figure 7**; **Table 5**; Table S4 in Supplementary Material). Whether they influence pollen tube guidance of dicots in general or have other roles in reproductive fitness will need to be investigated.

DISCUSSION

ORIGIN OF PLANT KEA AND CHX GENES FROM CYANOBACTERIA

Our phylogenetic analyses confirm that plant KEA and CHX genes of the CPA2 superfamily most likely originated from ancestral NhaA/NapA genes of prokaryotes (Chang et al., 2004; Brett et al., 2005a) of cyanobacterial origin. Although many bacteria, such as *E. coli* and *K. pneumonia* possess both NhaP and NhaA genes, CPA2 genes related to KEA/KefB/C are detected in a subset of bacteria, including Cyanobacteria *Gloeobacter* and *Anabaena variabilis* (**Figure 3**). Oddly, KEA or Kef-like genes have not been found in any fungi yet. These results suggest that plant CPA2 genes were derived from a cyanobacterium after it formed an endosymbiotic relationship with a primitive non-photosynthetic eukaryote cell. *Gloeobacter violaceus* may be a descendant of that primitive cyanobacterium (Criscuolo and Gribaldo, 2011), as it has genes encoding proteins homologous to NhaP, NhaS3, NhaS4, and KefB (KEA1/2). Similarly, NhaS4 from predominantly Cyanobacteria is highly homologous to plant CHX (**Figure 1**). However, NhaS4 also shares high similarity with fungi KHA1 and Nhe3/4 from a protist *Dictyostelium discoideum.*

Our analysis suggests that KEA transporters in eukaryotes serve functions preferentially in plant cells, and possibly in chloroplasts or plastids. First, we find that all plants sequenced to date possess three subtypes of KEA (**Table 4**). Green algae, *Chlamydomonas* and *Volvox*, have one KEA in each subtype, while higher plant KEAs diversified in subtype KEA1/2 or KEA4–6 (**Figure 3**). Second, KEAs are found mainly in plastid-containing organisms, including red algae and secondary endosymbiont like Chromista, but are not detected in fungi. Third, peptide sequences of KEA-Ia subtype were identified in proteomic analysis of isolated *Arabidopsis* plastids (Zybailov et al., 2008). Fourth, a functional study showed that *Synechocystis* NhaS3 was localized to thylakoid membranes and that it conferred tolerance to high Na and low K^+ when expressed in a salt-sensitive *E. coli* (Inaba et al., 2001; Tsunekawa et al., 2009). As NhaS3 is related to KefB and KEAs, plant KEA-I subtype may participate in monovalent cation and pH homeostasis of thylakoid membranes in chloroplasts. Plant KEA-Ia proteins are distinguished by a long N-terminal domain. The roles of this unusual domain and of KEA subtypes in plants need to be determined.

Table 4 | Overview of genes encoding cation–proton antiporters (CPA) genes in plants.

Classification	Species	CPA1			CPA2			
		NHX subtypes			KEA subtypes			CHX
				NHAP	KEA-Ia	KEA-Ib	KEA-II	
		NHX1-4	NHX5/6	NHX7/8	KEA1/2	KEA3	KEA4/5/6	CHX
EurosidI	*M. esculenta*				1	1	4	28
	R. communis	4	2	1	2	1	3	26
	P. trichocarpa	2	4	2	2	1	4	30
	M. truncatula	5	1	?	1	1	2	45
	G. max	7	3	1	4	2	6	46
	C. sativus				1	1	3	19
EurosidII	*A. thaliana*	4	2	2	2	1	3	28
	A. lyrata	4	2	2	2	1	4	28
	C. papaya				1	1	2	28
	E. grandis	6	1	2	1	1	2	19
Eudicot	*V. vinifera*				1	1	2	15
	M. guttatus				1	1	3	17
Monocots	*S. bicolor*	4	2	1	1	1	2	16
	Z. mays	6	2	1	2	1	3	15
	O. sativa	4	2	1	1	1	2	17
	B. distachyon	4	2	2	1	1	2	15
Lycophyte	*S. moellendorffii*	2	2	2	1	1	2	3
Bryophyte	*P. patens*	5	2	2	3	2	2	4
Zygnematales	*S. pratensis**	1	1	1	1	1	1	1
Volvocales	*V. carteri*	1–2	1	2	1	1	1	0
Volvocales	*C. reinhartdii*	1	3	1	1	1	1	0
Fungi	*S. cerevisiae*	0	1	1	0	0	0	1 KHA1
Cyanobacteria	*Synechocystis*	0	0	NhaP1/2	(NhaS3/5)			NhaS4
Bacteria	*E. coli*	0	0	1 NhaP	2 KefB/C		1 YBal	0

*Gene numbers are taken from the data in PHYTOZOME (http://www.phytozome.net/) from genomes of Manihot esculenta, Ricinus communis, Populus trichocarpa, Medicago truncatula, Glycine max, Cucumis sativus, Arabidopsis thaliana, Arabidopsis lyrata, Carica papaya, Eucalyptus grandis, Vitis vinifera, Mimulus guttatus, Sorghum bicolor, Zea mays, Oryza sativa, Brachypodium distachyon, Selaginella moellendorffii, Physcomitrella patens, Volvox carteri, and Chlamydomonas reinhartdii. *Data for Spirogyra pratensis are unpublished results of C. Delwiche. Gene numbers from Saccharomyces cerevisiae, Escherichia coli K12, and Synechocystis PCC 6803 are provided for comparison. Blanks indicate "not analyzed."*

VACUOLAR NHX1–4 CLADE IS SPECIFIC TO PLANTS

The finding that single-celled algae and early land plants have homologs of plasma membrane NhaP/SOS1, endosomal AtNHX5/6, and vacuolar AtNHX1–4 suggest these genes were present in ancestral algae before plants colonized land (**Table 4**). Furthermore, their roles are conserved and fundamental to plant cells, as the three NHX subtypes have persisted from single-celled algae to flowering plants. Endosomal AtNHX5/6 is part of the IC-NHE clade found in all eukaryote, including protist, fungi, and metazoa (Brett et al., 2005a). It is noteworthy that yeast NHX1 is localized to PVC and shares more similarity to endosomal NHX5/6 than to plant vacuolar NHX1. Importantly, our analysis extends a previous study (Brett et al., 2005a) that a clade of "vacuolar" AtNHX1–4 is specific to plants as seen from *Chlamydomonas* and moss to rice and *Arabidopsis* (**Figure 2**). Importantly, gene diversification of the "plant vacuolar NHX" in monocots and dicots (**Table 4**) underscores the diverse and critical roles vacuoles play in plant life (Martinoia et al., 2007; Muntz, 2007).

DIVERSIFICATION OF CHX GENE FAMILY DURING PLANT EVOLUTION

The multiplicity of the CHX gene family as plants colonized terrestrial habitats is striking. Phylogenetic analysis demonstrated this gene family in plants was founded by members similar to AtCHX20 and AtCHX16–19. First, cyanobacterium SynNhaS4 shares highest similarity with AtCHX17–20. Second, one CHX from a Charophyte, *Spirogyra*, is homologous to AtCHX20. Third, CHX genes from *Selaginella* and *Physcomitrella* encode proteins that are most homologous to AtCHX20 and AtCHX16–19, and fourth all flowering plants sequenced so far have CHX genes orthologs to guard cell-specific AtCHX20 and AtCHX16–19 (**Figure 7**). Thus a CHX member in *Spirogyra*, an alga that is closely related to land plants (Karol et al., 2001), is conserved in early land plants. In *Arabidopsis*, AtCHX20 is preferentially expressed in guard cells (Padmanaban

Table 5 | CHX homologs from algae to angiosperms cluster into distinct subclades.

Subclade	IVa	IVb	IVc	IVd	I	I-m	V	III	II
A. thaliana	CHX20	CHX16-19	CHX15	CHX21/23	CHX1/2/28	OsCHX9	CHX24/25	CHX13/14, 26/27	CHX3-12
Expression in At only	Guard	Root/leaf	Pol/Veg	Pollen	Pollen		Pollen	Pollen	Pollen
ALGAE									
Cre	0								
Spra	1								
EARLY LAND PLANT									
Ppa moss	4								
Smo	3								
MONOCOT									
Osa	1	3	2		4	(2)	(4)		
Bdi	2	3	2		4	(1)	(3)		
Zma	1	3	3		4	(1)	(3)		
Sbi	2	4	2		4	(2)	(2)		
DICOT									
Aco	1	2	3					11	
Mgu	1	3	3	1	2		2	3	1
Vvi	2	4	1	1	4		1	2	
Egr	2	1	1	1	2		4	3	3
Ccl	2	4	1	3	3		6	5	4
Ath	1	4	1	2	3		2	4	11
Aly	1	4	1	2	3		2	4	11
Csa	1	4	1	1	2		1	5	3
Mtr	3	5	2	5	2		2	6	12
Ptr	1	4	1	4	4		2	5	7
Mes	n.d.	3	2	3	4		3	5	2

*Number of genes affiliated with each subclade is shown. The AtCHX family is grouped into five subclades, I–V (Sze et al., 2004), and subclade IV splits in four groups, IVa–IVd, according to phylogenetic analysis (**Figure 7**) and functional studies. Expression refers only to CHX genes of Arabidopsis thaliana. The founding members are likely genes in subclade IVa. CHX gene family diversified in flowering plants, especially in dicots. Genes in two clusters specific to monocots are in parenthesis. See text for full names of plant species. Aco, Aquilegia coelulea; Ccl, Citrus clementina.*

et al., 2007), and CHX16–19 are expressed in various root cells, leaf hydathodes, as well as at the micropylar end of developing seeds (Chanroj, 2011). It would be interesting to determine if one or more of moss CHX genes are expressed in cells forming leaf pores and/or stomata which first appeared in mosses as they developed specialized epidermal cells to prevent desiccation (Freeman, 2008; Peterson et al., 2010).

Analysis of higher plants showed several monocot CHXs are orthologous to AtCHX15, AtCHX1/2, and AtCHX28. Thus these genes were present before the diversification of dicots and monocots. As these *Arabidopsis* genes are expressed in pollen (Sze et al., 2004[5]), we suggest CHX genes evolved as plants developed pollen grains to protect and transfer sperms for successful reproduction on land. Several rice CHX genes are expressed in developing anthers when pollen is at the bicellular and tricellular stages (Fujita et al., 2010[6]), supporting this idea. In general, monocots have less CHX genes than dicots encoding proteins homologous to pollen-expressed AtCHXs (**Table 4**). Thus CHX genes further diversified as dicot plants evolved as seen in AtCHX3–12 (Sze et al., 2004) from *Arabidopsis thaliana* or *A. lyrata*. Several homologs exist in *Medicago truncatula* in a sister clade to pollen-expressed *Arabidopsis* CHX. Appearance of pollen is an important innovation for successful reproduction on land (Freeman, 2008). Whether CHX homologs are related to development of monocolpate versus tricolpate pollen seen in monocots and dicots, respectively, remains to be seen.

STRUCTURAL AND FUNCTIONAL INNOVATIONS FOR SURVIVAL ON LAND

Based on functional studies of CHXs and phylogenetic results, we propose a simple model. The emergence of CHX genes in early land plants and their multiplicity in flowering plants, but not in higher metazoa, suggest CHX function is tied to land plant-specific cellular processes. Their roles in guard cell movement (Padmanaban et al., 2007) and pollen tube guidance (Lu et al., 2011) would support this idea. Based on these two studies, it is obvious that endomembrane-associated NHX members cannot substitute for CHX functions (Chanroj et al., 2011), even though vacuolar and endosomal NHX homologs are expressed in most cell types. Studies showed that AtCHX17 and AtCHX20 are associated with dynamic endomembranes of the endocytic and exocytic pathways suggesting roles in

[5]www.genevestigator.com/
[6]http://www.plexdb.org/

osmoregulation, remodeling of membranes as well as in secretion. Land plants secrete a variety of compounds for growth, development, and defense, including signaling molecules, soluble proteins, membrane complexes, extracellular and wall components. We propose that CHX transporters modulate the cation and pH environment of diverse intracellular compartments for protein activities within the lumen and/or for directed vesicle trafficking to deliver cargo in flowering plants to designated locations.

SUMMARY

Based on the prevalence of CPA1 genes in all organisms, we suggest that plant NHX gene family evolved to serve fundamental needs for cation and pH homeostasis in all eukaryote cells. NhaP/SOS1 regulates cation/H^+ fluxes at the plasma membrane as in prokaryotes, and endosomal NHX appeared later to modulate homeostasis and support functions of multiple intracellular compartments in primitive eukaryote cells. Further diversification of plant-specific vacuolar AtNHX1–4 type genes occurred in algal ancestors and in multicellular plants as vacuolar functions increased. In contrast, plant CPA2 genes likely originated from an ancestral cyanobacterium, where NhaS3- or NhaS5-like genes were transferred to the nuclear genome of ancestral plant cells. Three subtypes of KEA proteins are conserved from algae to flowering plants. Although their functions are still unknown, the evolutionary history would implicate a role of some KEA in ion homeostasis of plastids. A NhaS4-like gene of an ancient cyanobacterium is a likely progenitor of the CHX gene family which has multiplied and diversified specifically in flowering plants. Their functions in guard cell movement and pollen tube guidance determined so far would indicate they evolved to support vegetative and reproductive innovations required by plants to adapt and survive on land.

ACKNOWLEDGMENTS

The sequence data encoding CHX homologs from many plants were produced by the US Department of Energy Joint Genome Institute (www.Phytozome.net). This work was support in part by National Science Foundation Grant IBN0209788 and US Department of Energy Grant BES DEFG0207ER15883 to Heven Sze, grant BIO2008-01691 from Spanish Plan Nacional I + D + I to Kees Venema, and a Royal Thai Government Graduate Fellowship to Salil Chanroj. Work of CFD was supported by NSF grants #MCB-0523719 and DEB1036506.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Physiology/10.3389/fpls.2012.00025/abstract

Table S1 | Cation–proton antiporters 1 and CPA2 genes and proteins from prokaryotes, algae, protist, metazoa, and plants used in Figure 1.

Table S2 | Gene descriptions of selected CPA1 proteins used in Figure 2.

Table S3 | Gene descriptions of selected KEA/Kef proteins in Figure 3.

Table S4 | Description of CHX genes and proteins from Figure 7.

Table S5 | Xls. Gene descriptions of selected KEA/Kef proteins in Figure 3 with hyperlinks.

Figure S1 | Phylogenetic tree showing CPA1 and CPA2 proteins used in Figure 1.

Figure S2 | Tree showing all CPA1 and NHX proteins used in analysis of Figure 2.

Figure S3 | Tree of KEA/Kef proteins used in Figure 3.

Figure S4 | Alignment of the N-terminal domain of KEA clade Ia sequences. The N-terminal domain of the clade Ia protein sequences found in the different plant taxa were aligned by MUSCLE (Edgar, 2004). Conserved residues are shown in orange. Regions of predicted coiled coil structures are underlined in blue. The database references for the used species are listed in Table S3.

Figure S5 | Tree showing all CHX proteins used in analysis of Figure 7.

REFERENCES

Abrahamsen, M. S., Templeton, T. J., Enomoto, S., Abrahante, J. E., Zhu, G., Lancto, C. A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G. A., Xu, P., Bankier, A. T., Dear, P. H., Konfortov, B. A., Spriggs, H. F., Iyer, L., Anantharaman, V., Aravind, L., and Kapur, V. (2004). Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* 304, 441–445.

Adl, S. M., Simpson, A. G., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., Brugerolle, G., Fensome, R. A., Fredericq, S., James, T. Y., Karpov, S., Kugrens, P., Krug, J., Lane, C. E., Lewis, L. A., Lodge, J., Lynn, D. H., Mann, D. G., McCourt, R. M., Mendoza, L., Moestrup, O., Mozley-Standridge, S. E., Nerad, T. A., Shearer, C. A., Smirnov, A. V., Spiegel, F. W., and Taylor, M. F. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 52, 399–451.

Ali, R., Brett, C. L., Mukherjee, S., and Rao, R. (2004). Inhibition of sodium/proton exchange by a Rab-GTPase-activating protein regulates endosomal traffic in yeast. *J. Biol. Chem.* 279, 4498–4506.

Allen, R. D., and Naitoh, Y. (2002). Osmoregulation and contractile vacuoles of protozoa. *Int. Rev. Cytol.* 215, 351–394.

An, R., Chen, Q. J., Chai, M. F., Lu, P. L., Su, Z., Qin, Z. X., Chen, J., and Wang, X. C. (2007). AtNHX8, a member of the monovalent cation: proton antiporter-1 family in *Arabidopsis thaliana*, encodes a putative Li/H antiporter. *Plant J.* 49, 718–728.

Apse, M. P., Aharon, G. S., Snedden, W. A., and Blumwald, E. (1999). Salt tolerance conferred by overexpression of a vacuolar Na^+/H^+ antiport in *Arabidopsis*. *Science* 285, 1256–1258.

Apse, M. P., Sottosanto, J. B., and Blumwald, E. (2003). Vacuolar cation/H^+ exchange, ion homeostasis, and leaf development are altered in a T-DNA insertional mutant of AtNHX1, the *Arabidopsis* vacuolar Na^+/H^+ antiporter. *Plant J.* 36, 229–239.

Arkin, I. T., Xu, H., Jensen, M. O., Arbely, E., Bennett, E. R., Bowers, K. J., Chow, E., Dror, R. O., Eastwood, M. P., Flitman-Tene, R., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Shan, Y., and Shaw, D. E. (2007). Mechanism of Na^+/H^+ antiporting. *Science* 317, 799–803.

Banuelos, M. A., Sychrova, H., Bleykasten-Grosshans, C., Souciet, J. L., and Potier, S. (1998). The Nha1 antiporter of *Saccharomyces cerevisiae* mediates sodium and potassium efflux. *Microbiology* 144(Pt 10), 2749–2758.

Bassil, E., Ohto, M. A., Esumi, T., Tajima, H., Zhu, Z., Cagnac, O., Belmonte, M., Peleg, Z., Yamaguchi, T., and Blumwald, E. (2011a). The *Arabidopsis* intracellular Na^+/H^+ antiporters NHX5 and NHX6 are endosome associated and necessary for plant growth and development. *Plant Cell* 23, 224–239.

Bassil, E., Tajima, H., Liang, Y. C., Ohto, M. A., Ushijima, K., Nakano, R., Esumi, T., Coku, A., Belmonte, M., and Blumwald, E. (2011b). The *Arabidopsis* Na^+/H^+ antiporters NHX1 and NHX2 control vacuolar pH and K^+ homeostasis to regulate growth, flower development, and reproduction. *Plant Cell* 23, 3482–3497.

Booth, I. R. (2003). Bacterial ion channels. *Genet. Eng. (N.Y.)* 25, 91–111.

Bowers, K., Levi, B. P., Patel, F. I., and Stevens, T. H. (2000). The sodium/proton exchanger Nhx1p is required for endosomal protein trafficking in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 11, 4277–4294.

Bowman, J. L., Floyd, S. K., and Sakakibara, K. (2007). Green genes-comparative genomics of the green branch of life. *Cell* 129, 229–234.

Brett, C. L., Donowitz, M., and Rao, R. (2005a). Evolutionary origins of eukaryotic sodium/proton exchangers. *Am. J. Physiol. Cell Physiol.* 288, C223–C239.

Brett, C. L., Tukaye, D. N., Mukherjee, S., and Rao, R. (2005b). The yeast endosomal Na^+K^+/H^+ exchanger Nhx1 regulates cellular pH to control vesicle trafficking. *Mol. Biol. Cell* 16, 1396–1405.

Cellier, F., Conejero, G., Ricaud, L., Luu, D. T., Lepetit, M., Gosti, F., and Casse, F. (2004). Characterization of AtCHX17, a member of the cation/H exchangers, CHX family, from *Arabidopsis thaliana* suggests a role in K homeostasis. *Plant J.* 39, 834–846.

Chang, A. B., Lin, R., Keith Studley, W., Tran, C. V., Saier, M. H. Jr. (2004). Phylogeny as a guide to structure and function of membrane transport proteins. *Mol. Membr. Biol.* 21, 171–181.

Chanroj, S. (2011). *Plant-Specific K^+ Transporters with Distinct Properties and Their Emerging Roles in Endomembrane Trafficking.* Ph.D. Dissertation, University of Maryland, College Park.

Chanroj, S., Lu, Y., Padmanaban, S., Nanatani, K., Uozumi, N., Rao, R., and Sze, H. (2011). Plant-specific cation/H^+ exchanger 17 and its homologs are endomembrane K^+ transporters with roles in protein sorting. *J. Biol. Chem.* 286, 33931–33941.

Choe, S. (2002). Potassium channel structures. *Nat. Rev. Neurosci.* 3, 115–121.

Criscuolo, A., and Gribaldo, S. (2011). Large-scale phylogenomic analyses indicate a deep origin of primary plastids within Cyanobacteria. *Mol. Biol. Evol.* 28, 3019–3032.

Darley, C. P., van Wuytswinkel, O. C., van der Woude, K., Mager, W. H., and de Boer, A. H. (2000). *Arabidopsis thaliana* and *Saccharomyces cerevisiae* NHX1 genes encode amiloride sensitive electroneutral Na^+/H^+ exchangers. *Biochem. J.* 351, 241–249.

Douglas, S., Zauner, S., Fraunholz, M., Beaton, M., Penny, S., Deng, L. T., Wu, X., Reith, M., Cavalier-Smith, T., and Maier, U. G. (2001). The highly reduced genome of an enslaved algal nucleus. *Nature* 410, 1091–1096.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Ferguson, G. P., Battista, J. R., Lee, A. T., and Booth, I. R. (2000). Protection of the DNA during the exposure of *Escherichia coli* cells to a toxic metabolite: the role of the KefB and KefC potassium channels. *Mol. Microbiol.* 35, 113–122.

Fraile-Escanciano, A., Kamisugi, Y., Cuming, A. C., Rodriguez-Navarro, A., and Benito, B. (2010). The SOS1 transporter of *Physcomitrella patens* mediates sodium efflux in planta. *New Phytol.* 188, 750–761.

Freeman, S (2008). "Green plants," in *Biological Sciences*, 3rd Edn. San Francisco, CA: Pearson Education, 626–663.

Fujisawa, M., Ito, M., and Krulwich, T. A. (2007). Three two-component transporters with channel-like properties have monovalent cation/proton antiport activity. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13289–13294.

Fujita, M., Horiuchi, Y., Ueda, Y., Mizuta, Y., Kubo, T., Yano, K., Yamaki, S., Tsuda, K., Nagata, T., Niihama, M., Kato, H., Kikuchi, S., Hamada, K., Mochizuki, T., Ishimizu, T., Iwai, H., Tsutsumi, N., and Kurata, N. (2010). Rice expression atlas in reproductive development. *Plant Cell Physiol.* 51, 2060–2081.

Gaxiola, R. A., Rao, R., Sherman, A., Grisafi, P., Alper, S. L., and Fink, G. R. (1999). The *Arabidopsis thaliana* proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1480–1485.

Gerisch, G., Heuser, J., and Clarke, M. (2002). Tubular-vesicular transformation in the contractile vacuole system of *Dictyostelium*. *Cell Biol. Int.* 26, 845–852.

Goswami, P., Paulino, C., Hizlan, D., Vonck, J., Yildiz, O., and Kuhlbrandt, W. (2011). Structure of the archaeal Na^+/H^+ antiporter NhaP1 and functional role of transmembrane helix 1. *EMBO J.* 30, 439–449.

Hamada, A., Hibino, T., Nakamura, T., and Takabe, T. (2001). Na^+/H^+ antiporter from *Synechocystis* species PCC 6803, homologous to SOS1, contains an aspartic residue and long C-terminal tail important for the carrier activity. *Plant Physiol.* 125, 437–446.

Hellmer, J., Patzold, R., and Zeilinger, C. (2002). Identification of a pH regulated Na(+)/H(+) antiporter of *Methanococcus jannaschii*. *FEBS Lett.* 527, 245–249.

Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E., and Michel, H. (2005). Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* 435, 1197–1202.

Inaba, M., Sakamoto, A., and Murata, N. (2001). Functional expression in *Escherichia coli* of low-affinity and high-affinity Na(+) (Li(+))/H(+) antiporters of *Synechocystis*. *J. Bacteriol.* 183, 1376–1384.

Jiang, Y., Pico, A., Cadene, M., Chait, B. T., and MacKinnon, R. (2001). Structure of the RCK domain from the *E. coli* K^+ channel and demonstration of its presence in the human BK channel. *Neuron* 29, 593–601.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282.

Karol, K. G., McCourt, R. M., Cimino, M. T., and Delwiche, C. F. (2001). The closest living relatives of land plants. *Science* 294, 2351–2353.

Katiyar-Agarwal, S., Zhu, J., Kim, K., Agarwal, M., Fu, X., Huang, A., and Zhu, J. K. (2006). The plasma membrane Na^+/H^+ antiporter SOS1 interacts with RCD1 and functions in oxidative stress tolerance in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18816–18821.

Keeling, P. J. (2004). Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* 91, 1481–1493.

Kinclova, O., Ramos, J., Potier, S., and Sychrova, H. (2001). Functional study of the *Saccharomyces cerevisiae* Nha1p C-terminus. *Mol. Microbiol.* 40, 656–668.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.

Leidi, E. O., Barragan, V., Rubio, L., El-Hamdaoui, A., Ruiz, M. T., Cubero, B., Fernandez, J. A., Bressan, R. A., Hasegawa, P. M., Quintero, F. J., and Pardo, J. M. (2010). The AtNHX1 exchanger mediates potassium compartmentation in vacuoles of transgenic tomato. *Plant J.* 61, 495–506.

Leonardy, S., Miertzschke, M., Bulyha, I., Sperling, E., Wittinghofer, A., and Søgaard-Andersen, L. (2010). Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J.* 29, 2276–2289.

Li, H. T., Liu, H., Gao, X. S., and Zhang, H. X. (2009). Knock-out of *Arabidopsis* AtNHX4 gene enhances tolerance to salt stress. *Biochem. Biophys. Res. Commun.* 382, 637–641.

Liu, H., Tang, R. J., Zhang, Y., Wang, C. T., Lv, Q. D., Gao, X. S., Li, W. B., and Zhang, H. X. (2010). AtNHX3 is a vacuolar K^+/H^+ antiporter required for low-potassium tolerance in *Arabidopsis thaliana*. *Plant Cell Environ.* 33, 1989–1999.

Lu, Y., Chanroj, S., Zulkifli, L., Johnson, M. A., Uozumi, N., Cheung, A., and Sze, H. (2011). Pollen tubes lacking a pair of K^+ transporters fail to target ovules in *Arabidopsis*. *Plant Cell* 23, 81–93.

Maresova, L., and Sychrova, H. (2006). *Arabidopsis thaliana* CHX17 gene complements the kha1 deletion phenotypes in *Saccharomyces cerevisiae*. *Yeast* 23, 1167–1171.

Martinoia, E., Maeshima, M., and Neuhaus, H. E. (2007). Vacuolar transporters and their essential role in plant metabolism. *J. Exp. Bot.* 58, 83–102.

Maser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J., Sanders, D., Harper, J. F., Tchieu, J., Gribskov, M., Persans, M. W., Salt, D. E., Kim, S. A., and Guerinot, M. L. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* 126, 1646–1667.

Mauriello, E. M., Mignot, T., Yang, Z., and Zusman, D. R. (2010). Gliding motility revisited: how do the myxobacteria move without flagella? *Microbiol. Mol. Biol. Rev.* 74, 229–249.

Miller, S., Ness, L. S., Wood, C. M., Fox, B. C., and Booth, I. R. (2000). Identification of an ancillary protein, YabF, required for activity of the KefC glutathione-gated potassium efflux system in *Escherichia coli*. *J. Bacteriol.* 182, 6536–6540.

Muntz, K. (2007). Protein dynamics and proteolysis in plant vacuoles. *J. Exp. Bot.* 58, 2391–2407.

Nicholas, K. B., Nicholas, H. B., and Deerfield, D. W. (1997). GeneDoc: analysis and visualization of genetic variation. *EMBNEW NEWS* 4, 14.

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-coffee: a novel method for fast and accurate

multiple sequence alignment. *J. Mol. Biol.* 302, 205–217.

Oh, D. H., Lee, S. Y., Bressan, R. A., Yun, D. J., and Bohnert, H. J. (2010). Intracellular consequences of SOS1 deficiency during salt stress. *J. Exp. Bot.* 61, 1205–1213.

Ohgaki, R., Nakamura, N., Mitsui, K., and Kanazawa, H. (2005). Characterization of the ion transport activity of the budding yeast Na^+/H^+ antiporter, Nha1p, using isolated secretory vesicles. *Biochim. Biophys. Acta* 1712, 185–196.

Ohnishi, M., Fukada-Tanaka, S., Hoshino, A., Takada, J., Inagaki, Y., and Iida, S. (2005). Characterization of a novel Na^+/H^+ antiporter gene InNHX2 and comparison of InNHX2 with InNHX1, which is responsible for blue flower coloration by increasing the vacuolar pH in the Japanese morning glory. *Plant Cell Physiol.* 46, 259–267.

Olias, R., Eljakaoui, Z., Li, J., De Morales, P. A., Marin-Manzano, M. C., Pardo, J. M., and Belver, A. (2009). The plasma membrane Na^+/H^+ antiporter SOS1 is essential for salt tolerance in tomato and affects the partitioning of Na^+ between plant organs. *Plant Cell Environ.* 32, 904–916.

Padan, E. (2008). The enlightening encounter between structure and function in the NhaA Na^+-H^+ antiporter. *Trends Biochem. Sci.* 33, 435–443.

Padmanaban, S., Chanroj, S., Kwak, J. M., Li, X., Ward, J. M., and Sze, H. (2007). Participation of endomembrane cation/H^+ exchanger AtCHX20 in osmoregulation of guard cells. *Plant Physiol.* 144, 82–93.

Peterson, K. M., Rychel, A. L., and Torii, K. U. (2010). Out of the mouths of plants: the molecular basis of the evolution and diversity of stomatal development. *Plant Cell* 22, 296–306.

Qi, Z., and Spalding, E. P. (2004). Protection of plasma membrane K^+ transport by the salt overly sensitive1 Na^+-H^+ antiporter during salinity stress. *Plant Physiol.* 136, 2548–2555.

Qiu, Q. S., Barkla, B. J., Vera-Estrella, R., Zhu, J. K., and Schumaker, K. S. (2003). Na^+/H^+ exchange activity in the plasma membrane of *Arabidopsis. Plant Physiol.* 132, 1041–1052.

Qiu, Q. S., Guo, Y., Dietrich, M. A., Schumaker, K. S., and Zhu, J. K. (2002). Regulation of SOS1, a plasma membrane Na^+/H^+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8436–8441.

Quintero, F. J., Blatt, M. R., and Pardo, J. M. (2000). Functional conservation between yeast and plant endosomal Na(+)/H(+) antiporters. *FEBS Lett.* 471, 224–228.

Quintero, F. J., Martinez-Atienza, J., Villalta, I., Jiang, X., Kim, W. Y., Ali, Z., Fujii, H., Mendoza, I., Yun, D. J., Zhu, J. K., and Pardo, J. M. (2011). Activation of the plasma membrane Na/H antiporter Salt-Overly-Sensitive 1 (SOS1) by phosphorylation of an auto-inhibitory C-terminal domain. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2611–2616.

Quintero, F. J., Ohta, M., Shi, H. Z., Zhu, J. K., and Pardo, J. M. (2002). Reconstitution in yeast of the *Arabidopsis* SOS signaling pathway for Na^+ homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9061–9066.

Rodriguez-Rosales, M. P., Galvez, F. J., Huertas, R., Aranda, M. N., Baghour, M., Cagnac, O., and Venema, K. (2009). Plant NHX cation/proton antiporters. *Plant Signal. Behav.* 4, 265–276.

Rodriguez-Rosales, M. P., Jiang, X., Galvez, F. J., Aranda, M. N., Cubero, B., and Venema, K. (2008). Overexpression of the tomato K^+/H^+ antiporter LeNHX2 confers salt tolerance by improving potassium compartmentalization. *New Phytol.* 179, 366–377.

Roosild, T. P., Castronovo, S., Healy, J., Miller, S., Pliotas, C., Rasmussen, T., Bartlett, W., Conway, S. J., and Booth, I. R. (2010). Mechanism of ligand-gated potassium efflux in bacterial pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19784–19789.

Roosild, T. P., Miller, S., Booth, I. R., and Choe, S. (2002). A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell* 109, 781–791.

Rose, A., Manikantan, S., Schraegle, S. J., Maloy, M. A., Stahlberg, E. A., and Meier, I. (2004). Genome-wide identification of *Arabidopsis* coiled-coil proteins and establishment of the ARABI-COIL database. *Plant Physiol.* 134, 927–939.

Saier, M. H. Jr. (2000). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol. Mol. Biol. Rev.* 64, 354–411.

Senior, A., and Moir, A. (2008). The *Bacillus cereus* GerN and GerT protein homologs have distinct roles in spore germination and outgrowth, respectively. *J. Bacteriol.* 190, 6148–6152.

Shi, H., Ishitani, M., Kim, C., and Zhu, J. K. (2000). The *Arabidopsis thaliana* salt tolerance gene SOS1 encodes a putative Na^+/H^+ antiporter. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6896–6901.

Shi, H., Quintero, F. J., Pardo, J. M., and Zhu, J. K. (2002). The putative plasma membrane Na(+)/H(+) antiporter SOS1 controls long-distance Na(+) transport in plants. *Plant Cell* 14, 465–477.

Southworth, T. W., Guffanti, A. A., Moir, A., and Krulwich, T. A. (2001). GerN, an endospore germination protein of *Bacillus cereus*, is an Na(+)/H(+)-K(+) antiporter. *J. Bacteriol.* 183, 5896–5903.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Strausak, D., Waser, M., and Solioz, M. (1993). Functional expression of the *Enterococcus hirae* NaH-antiporter in *Escherichia coli. J. Biol. Chem.* 268, 26334–26337.

Sychrova, H., Ramirez, J., and Pena, A. (1999). Involvement of Nha1 antiporter in regulation of intracellular pH in *Saccharomyces cerevisiae. FEMS Microbiol. Lett.* 171, 167–172.

Sze, H., Padmanaban, S., Cellier, F., Honys, D., Cheng, N. H., Bock, K. W., Conejero, G., Li, X., Twell, D., Ward, J. M., and Hirschi, K. D. (2004). Expression patterns of a novel AtCHX gene family highlight potential roles in osmotic adjustment and K^+ homeostasis in pollen development. *Plant Physiol.* 136, 2532–2547.

Taglicht, D., Padan, E., and Schuldiner, S. (1991). Overproduction and purification of a functional Na^+/H^+ antiporter coded by nhaA (ant) from *Escherichia coli. J. Biol. Chem.* 266, 11289–11294.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

Tanifuji, G., Onodera, N. T., Wheeler, T. J., Dlutek, M., Donaher, N., and Archibald, J. M. (2011). Complete nucleomorph genome sequence of the nonphotosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. *Genome Biol. Evol.* 3, 44–54.

Thackray, P. D., Behravan, J., Southworth, T. W., and Moir, A. (2001). GerN, an antiporter homologue important in germination of *Bacillus cereus* endospores. *J. Bacteriol.* 183, 476–482.

Timme, R. E., Bachvaroff, T. R., and Delwiche, C. F. (2012). Broad phylogenomic sampling and the sister lineage of land plants. *PLoS ONE* 7, e29696.

Tsunekawa, K., Shijuku, T., Hayashimoto, M., Kojima, Y., Onai, K., Morishita, M., Ishiura, M., Kuroda, T., Nakamura, T., Kobayashi, H., Sato, M., Toyooka, K., Matsuoka, K., Omata, T., and Uozumi, N. (2009). Identification and characterization of the Na^+/H^+ antiporter Nhas3 from the thylakoid membrane of *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 284, 16513–16521.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M. K., McDonald, W. H., Medina, M., Meijer, H. J., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. (2006). Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266.

Uchikawa, T., Yamamoto, A., and Inouye, K. (2011). Origin and function of the stalk-cell vacuole in *Dictyostelium. Dev. Biol.* 352, 48–57.

Venema, K., Belver, A., Marin-Manzano, M. C., Rodriguez-Rosales, M. P., and Donaire, J. P. (2003). A novel intracellular K^+/H^+ antiporter related to Na^+/H^+ antiporters is important for K^+ ion homeostasis in plants. *J. Biol. Chem.* 278, 22453–22459.

Venema, K., Quintero, F. J., Pardo, J. M., and Donaire, J. P. (2002). The arabidopsis Na^+/H^+ exchanger AtNHX1 catalyzes low affinity Na^+ and K^+ transport in reconstituted liposomes. *J. Biol. Chem.* 277, 2413–2418.

Vinothkumar, K. R., Smits, S. H., and Kuhlbrandt, W. (2005).

pH-induced structural change in a sodium/proton antiporter from *Methanococcus jannaschii*. *EMBO J.* 24, 2720–2729.

Vothknecht, U. C., and Westhoff, P. (2001). Biogenesis and origin of thylakoid membranes. *Biochim. Biophys. Acta* 1541, 91–101.

Wagner, M. C., Molnar, E. E., Molitoris, B. A., and Goebl, M. G. (2006). Loss of the homotypic fusion and vacuole protein sorting or golgi-associated retrograde protein vesicle tethering complexes results in gentamicin sensitivity in the yeast *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother.* 50, 587–595.

Waser, M., Hess-Bienz, D., Davies, K., and Solioz, M. (1992). Cloning and disruption of a putative NaH-antiporter gene of *Enterococcus hirae*. *J. Biol. Chem.* 267, 5396–5400.

Wu, S. J., Ding, L., and Zhu, J. K. (1996). SOS1, a genetic locus essential for salt tolerance and potassium acquisition. *Plant Cell* 8, 617–627.

Xiang, M., Feng, M., Muend, S., and Rao, R. (2007). A human Na^+/H^+ antiporter sharing evolutionary origins with bacterial NhaA may be a candidate gene for essential hypertension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18677–18681.

Yamaguchi, T., Aharon, G. S., Sottosanto, J. B., and Blumwald, E. (2005). Vacuolar Na^+/H^+ antiporter cation selectivity is regulated by calmodulin from within the vacuole in a Ca^{2+}- and pH-dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16107–16112.

Yamaguchi, T., Apse, M. P., Shi, H. Z., and Blumwald, E. (2003). Topological analysis of a plant vacuolar Na^+/H^+ antiporter reveals a luminal C terminus that regulates antiporter cation selectivity. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12510–12515.

Yamaguchi, T., Fukada-Tanaka, S., Inagaki, Y., Saito, N., Yonekura-Sakakibara, K., Tanaka, Y., Kusumi, T., and Iida, S. (2001). Genes encoding the vacuolar Na^+/H^+ exchanger and flower coloration. *Plant Cell Physiol.* 42, 451–461.

Yokoi, S., Quintero, F. J., Cubero, B., Ruiz, M. T., Bressan, R. A., Hasegawa, P. M., and Pardo, J. M. (2002). Differential expression and function of *Arabidopsis thaliana* NHX Na^+/H^+ antiporters in the salt stress response. *Plant J.* 30, 529–539.

Yoshida, K., Miki, N., Momonoi, K., Kawachi, M., Katou, K., Okazaki, Y., Uozumi, N., Maeshima, M., and Kondo, T. (2009). Synchrony between flower opening and petal-color change from red to blue in morning glory, *Ipomoea tricolor* cv. Heavenly Blue. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 85, 187–197.

Zhao, J., Cheng, N. H., Motes, C. M., Blancaflor, E. B., Moore, M., Gonzales, N., Padmanaban, S., Sze, H., Ward, J. M., and Hirschi, K. D. (2008). AtCHX13 is a plasma membrane K^+ transporter. *Plant Physiol.* 148, 796–807.

Zhou, G. A., Jiang, Y., Yang, Q., Wang, J. F., Huang, J., and Zhang, H. S. (2006). Isolation and characterization of a new Na^+/H^+ antiporter gene OsNHA1 from rice (*Oryza sativa* L.). *DNA Seq.* 17, 24–30.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and van Wijk, K. J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3, e1994. doi:10.1371/journal.pone.0001994

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 October 2011; accepted: 21 January 2012; published online: 14 February 2012.

Citation: Chanroj S, Wang G, Venema K, Zhang MW, Delwiche CF and Sze H (2012) Conserved and diversified gene families of monovalent cation/H^+ antiporters from algae to flowering plants. Front. Plant Sci. ***3****:25. doi: 10.3389/fpls.2012.00025*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 January 2012
doi: 10.3389/fpls.2012.00001

Protein phylogenetic analysis of Ca^{2+}/cation antiporters and insights into their evolution in plants

Laura Emery*[1], *Simon Whelan*[1], *Kendal D. Hirschi*[2] *and Jon K. Pittman*[1]

[1] *Faculty of Life Sciences, University of Manchester, Manchester, UK*
[2] *Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX, USA*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Sakiko Okumoto, Virginia Tech, USA
Guillaume Pilot, Virginia Tech, USA

***Correspondence:**
Jon K. Pittman, Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK.
e-mail: jon.pittman@manchester.ac.uk

Cation transport is a critical process in all organisms and is essential for mineral nutrition, ion stress tolerance, and signal transduction. Transporters that are members of the Ca^{2+}/cation antiporter (CaCA) superfamily are involved in the transport of Ca^{2+} and/or other cations using the counter exchange of another ion such as H^+ or Na^+. The CaCA superfamily has been previously divided into five transporter families: the YRBG, Na^+/Ca^{2+} exchanger (NCX), Na^+/Ca^{2+}, K^+ exchanger (NCKX), H^+/cation exchanger (CAX), and cation/Ca^{2+} exchanger (CCX) families, which include the well-characterized NCX and CAX transporters. To examine the evolution of CaCA transporters within higher plants and the green plant lineage, CaCA genes were identified from the genomes of sequenced flowering plants, a bryophyte, lycophyte, and freshwater and marine algae, and compared with those from non-plant species. We found evidence of the expansion and increased diversity of flowering plant genes within the CAX and CCX families. Genes related to the NCX family are present in land plant though they encode distinct MHX homologs which probably have an altered transport function. In contrast, the NCX and NCKX genes which are absent in land plants have been retained in many species of algae, especially the marine algae, indicating that these organisms may share "animal-like" characteristics of Ca^{2+} homeostasis and signaling. A group of genes encoding novel CAX-like proteins containing an EF-hand domain were identified from plants and selected algae but appeared to be lacking in any other species. Lack of functional data for most of the CaCA proteins make it impossible to reliably predict substrate specificity and function for many of the groups or individual proteins. The abundance and diversity of CaCA genes throughout all branches of life indicates the importance of this class of cation transporter, and that many transporters with novel functions are waiting to be discovered.

Keywords: calcium transport, cation transport, evolution, H^+/Ca^{2+} exchanger, Na^+/Ca^{2+} exchanger, phylogeny, CaCA

INTRODUCTION

The importance of transporters as "gatekeepers" of the cell is exemplified in recent genomic studies. For example, the central role of transporters in controlling cell growth was indicated in a study showing that increased yeast growth correlates with increased cell surface area, and thus increased nutrient uptake (Groeneveld et al., 2009). Indeed, of the 1312 distinct reaction steps in a reconstruction of yeast metabolism, 401 involve transport (Herrgard et al., 2008). Likewise, the analysis of fully sequenced photosynthetic eukaryotes finds that transporters consistently make up a significant proportion (∼5%) of the genome (Mäser et al., 2001; Merchant et al., 2007; Rensing et al., 2008). Calcium (Ca^{2+}) is a critical element in all organisms; it is an essential nutrient which also has a conserved signaling role (Hirschi, 2004; Case et al., 2007). Transporters that mediate the movement of this ion are therefore particularly important. These Ca^{2+} transporters impact upon cell division, growth, development, and adaptation to environmental conditions. Three classes of membrane transporters help mediate Ca^{2+} flux across a membrane and regulate cytosolic Ca^{2+} levels: Ca^{2+}-permeable channels, Ca^{2+}-ATPases, and Ca^{2+}/cation antiporters (CaCAs). Biochemical identification and functional analysis of CaCAs in an array of organisms have helped conceptualize the functional properties and important physiological roles of these exchangers in cellular ion homeostasis (Lytton, 2007; Manohar et al., 2011).

The majority of the ion-coupled Ca^{2+} exchange activity within a cell is encoded by members of the Ca^{2+}/cation antiporter (CaCA) gene superfamily which are present widely in archaea, bacteria, fungi, plants, and animals (Saier et al., 1999; Cai and Lytton, 2004a; **Table 1**). These proteins serve as essential components in Ca^{2+} cycling systems. CaCA proteins promote Ca^{2+} efflux across membranes, normally against its concentration gradient, by using a counter-electrochemical gradient of other ions such as H^+, Na^+, or K^+ to energize the process. Animal proteins principally use Na^+ gradients as the driving force while plant and bacterial exchangers exclusively utilize H^+. Based on functional and phylogenetic analysis, the CaCA superfamily is composed of at least five families (Cai and Lytton, 2004a): the Na^+/Ca^{2+} exchanger (NCX) family, the Na^+/Ca^{2+}, K^+ exchanger (NCKX) family and the CCX family, the H^+/cation exchanger (CAX) family, and the YRBG family

Table 1 | Definition of the five major gene families that make up the CaCA superfamily.

CaCA gene family	Described major function	Species domain distribution	Key selected members
YRBG	Putative Na^+/Ca^{2+} exchanger	Bacteria, archaea	*EcyrbG* (bacteria)
NCX	Na^+/Ca^{2+} exchanger	Eukaryotes	*NCX1 – 3* (mammalian), *AtMHX* (plant)
NCKX	K^+-dependent Na^+/Ca^{2+} exchanger	Eukaryotes – excluding land plants	*NCKX1 – 5* (mammalian)
CCX	Cation/Ca^{2+} exchanger	Eukaryotes	*NCKX6* (mammalian), *AtCCX3* (plant)
CAX	Ca^{2+}/H^+ exchange	Bacteria, eukaryotes – excluding mammals, insects, nematodes	*ScVCX1* (yeast), *AtCAX1* (plant), *EcchaA* (bacteria)

Table 2 | Summary of the major properties and functions of selected CaCA genes.

Gene name (species)	CaCA family	Transport function	Sub-cellular location	Main function	Reference
yrbG (*E. coli*)	YRBG	Putative Na^+/Ca^{2+} exchange – unconfirmed	Inner membrane	Unknown; not involved in regulating cytosolic Ca^{2+}; not essential	Naseem et al. (2008)
NCX1 (human, mouse)	NCX	Na^+/Ca^{2+} exchange; stoichiometry: $3Na^+:1Ca^{2+}$	Plasma membrane	Ca^{2+} efflux from cell in heart, brain, kidney	Blaustein and Lederer (1999)
AtMHX (*Arabidopsis*)	NCX (MHX)	Mg^{2+}/H^+ and Zn^{2+}/H^+ exchange; electrogenic	Vacuole	Vacuolar Mg^{2+} sequestration; Mg^{2+} homeostasis	Shaul et al. (1999)
NCKX1 (bovine, human)	NCKX	K^+-dependent Na^+/Ca^{2+} exchange; stoichiometry: $4Na^+:1Ca^{2+}$, $1K^+$	Plasma membrane	Ca^{2+} efflux in rod photoreceptor retina cells to allow light-triggered signal generation	Reilander et al. (1992), Schnetkamp (1995)
NCKX6/NCLX (human, mouse)	CCX	Na^+/Ca^{2+} and Li^+/Ca^{2+} exchange, K^+-independent?	Mitochondria	Mitochondrial Ca^{2+} efflux by Na^+/Ca^{2+} exchange	Cai and Lytton (2004b), Palty et al. (2004, 2010)
AtCCX3 (*Arabidopsis*)	CCX	K^+/H^+, Na^+/H^+, and Mn^{2+}/H^+ exchange	Endo-membrane	Vesicle ion uptake; K^+ homeostasis in flower tissues	Morris et al. (2008)
chaA (*E. coli*)	CAX (CHAA)	Ca^{2+}/H^+, Na^+/H^+, and K^+/H^+ exchange	Inner membrane	Outward and inward cation flux; ion homeostasis	Ivey et al. (1993), Ohyama et al. (1994), Radchenko et al. (2006)
VCX1 (yeast)	CAX	Ca^{2+}/H^+ exchange; stoichiometry: $2H^+:1Ca^{2+}$	Vacuole	Vacuolar Ca^{2+} loading; regulates cytosolic Ca^{2+}	Cunningham and Fink (1994), Denis and Cyert (2002)
AtCAX1 (*Arabidopsis*)	CAX	Ca^{2+}/H^+ exchange; stoichiometry: $3H^+:1Ca^{2+}$ (based on red beet studies); can also mediate Mn^{2+}/H^+ and Cd^{2+}/H^+ exchange	Vacuole	Vacuolar Ca^{2+} loading; provides Ca^{2+} tolerance; determines cellular Ca^{2+} concentration; involved in response to stress	Blackford et al. (1990), Hirschi et al. (1996), Cheng et al. (2003), Conn et al. (2011)
CrCAX1 (green microalgae)	CAX	Ca^{2+}/H^+ and Na^+/H^+ exchange	Vacuole (predicted)	Provides vacuolar Ca^{2+} sequestration and Ca^{2+} tolerance when expressed in yeast	Pittman et al. (2009)
DrCAX (zebra fish)	CAX	Ca^{2+}/H^+ exchange	Endo-membrane	Neural crest development	Manohar et al. (2010)

named after the *Escherichia coli* gene *yrbG* (**Table 1**). Detailed characterization of selected isoforms from each of these families has begun to shed light on their functions (summarized in **Table 2**).

The YRBG family has been the least characterized. Sequence similarity suggests *E. coli* YrbG may function as a NCX but the transporter is not essential for bacterial growth and it is unknown whether YrbG is a Ca^{2+} transporter (Cai and Lytton, 2004a; Naseem et al., 2008). In *E. coli*, pH and monovalent cations can regulate cytosolic free Ca^{2+} through Ca^{2+} influx and efflux, but direct measurement of cytosolic Ca^{2+} levels using the Ca^{2+} reporter aequorin suggests that YrbG does not play a role in regulating cytosolic Ca^{2+} (Naseem et al., 2008). Orthologs of *yrbG* are present in other prokaryotes but do not appear to be present in eukaryotes (Cai and Lytton, 2004a).

Na^+/Ca^{2+} exchangers, present at the plasma membrane of most animal cells, are a fast and high-capacity Ca^{2+} transport system that are important regulators of cellular Ca^{2+} homeostasis by causing efflux of Ca^{2+} from the cell. Detailed functional and

molecular studies have revealed two distinct families of Na^+/Ca^{2+} exchange proteins (Lytton, 2007). The first of these, the NCX transporters are composed of three distinct types in mammals: NCX1, NCX2, and NCX3. All three exchangers share about 70% overall amino acid identity and have identical transport function. Invertebrates have a single NCX gene, whereas vertebrate species have multiple NCX genes as a result of at least two duplication events (On et al., 2008). The presence of NCX genes in all animal species' genomes allowed the construction of a phylogenetic tree that correlated to animal evolution and revealed that the origin of NCX duplication was initiated at the emergence of vertebrate organisms (On et al., 2008). A plant Mg^{2+}/H^+ exchanger termed *Arabidopsis thaliana* Mg^{2+}/H^+ exchanger (AtMHX), is also a member of the NCX family. This transporter concentrates Mg^{2+} into vacuoles and has no apparent Na^+/Ca^{2+} exchange activity (Shaul et al., 1999). Another putative Mg^{2+}/H^+ exchanger (XNTA) has been discovered in *Paramecium*, but although it appears to be a member of the CaCA superfamily, it is distinct from AtMHX and does not appear to fall within the NCX clade (Haynes et al., 2002).

The second Na^+/Ca^{2+}-transporting exchanger family, the NCKX, operate in cellular Ca^{2+} efflux by extruding one Ca^{2+} ion and one K^+ ion in exchange for four Na^+ ions (Lytton, 2007). Like the NCX transporters, the NCKX play an important role in Ca^{2+} homeostasis, notably in retinal and neuronal cells, and are important in various roles including photoreceptor function and skin pigmentation (**Table 2**). Also like NCX transporters, NCKX are reversible and can import Ca^{2+} into the cell upon reversal of the transmembrane Na^+ gradient. NCKX proteins differ from NCX proteins in their requirement for K^+, their Ca^{2+} kinetics, and sequence divergence. To date there has been no evidence of NCKX activity or NCKX-like genes present in any plants.

Initial analyses of the CaCA superfamily did not identify the CCX family and this family has only recently been created and considered a stand-alone group. A novel gene from mammals, originally thought to be a NCKX gene, named *NCKX6*, has significant sequence variation from other known *NCX* and *NCKX* genes (Cai and Lytton, 2004b). Subsequent phylogenetic analysis has grouped *NCKX6* separately from the NCX and NCKX family in the CCX family, which includes members from vertebrates, invertebrates, fungi, protozoa, and plants (Cai and Lytton, 2004a). NCKX6 (also named NCLX) appears to be a bona fide NCX but it is slightly contentious as to whether it is K^+-dependent or not; other studies have suggested that this transporter is K^+-independent and can also transport Ca^{2+} by Li^+ exchange (Palty et al., 2004; **Table 2**). Several transporters in plants, some formerly referred to as CAX, have been reassigned to the CCX family (Shigaki et al., 2006). For example, *Arabidopsis* CCX3 (formerly CAX9) is a H^+-dependent cation transporter but with no apparent Ca^{2+} transport properties (Morris et al., 2008).

Members of the CAX family have been identified from many species including plants, algae, some invertebrate animals, fungi, protozoa, and bacteria, but are absent from the genomes of mammals, insects, and nematodes (Shigaki et al., 2006; Manohar et al., 2011). CAXs have been classified into three major categories based on phylogenetic analysis: Type 1, 2, and 3 (Shigaki et al., 2006). Several Type 1 and 3 CAXs have been characterized across species, particularly in plants and yeast, and have been shown to be important for Ca^{2+}/H^+ exchange, predominantly at the vacuolar membrane (Manohar et al., 2011; Pittman, 2011). In particular, many plant and fungal CAXs are important for providing tolerance to excess Ca^{2+}, for regulating Ca^{2+} homeostasis and nutrition, and the modulation of environmental or developmental events via Ca^{2+} signal generation (McAinsh and Pittman, 2009; Cunningham, 2011). However, to date only two Type 2 CAXs, the *S. cerevisiae* Vnx1 and zebrafish DrCax1 have been characterized (Cagnac et al., 2007, 2010; Manohar et al., 2010).

The Type 1 plant CAXs have been further divided into two phylogenetically distinct groups (Type 1A and Type 1B; Shigaki et al., 2006). This distinction suggests functional differentiation between the two groups; however, no clear-cut functional differences have so far been identified. *Arabidopsis* and rice (*Oryza sativa*) each have six CAX ORFs equally divided between the two phylogenetic groups. One suggestion has been that there are differences in ion substrate selectivity between members of the two groups; for example, Type 1B proteins such as AtCAX2 and AtCAX5 have been shown to transport multiple ions including Mn^{2+} and Cd^{2+} in addition to Ca^{2+} (Hirschi et al., 2000; Shigaki et al., 2003; Edmond et al., 2009) while Type 1A members such as AtCAX1 and AtCAX3 are important for Ca^{2+} homeostasis and were originally thought to be specific for Ca^{2+} (Shigaki et al., 2003; Conn et al., 2011). However, Type 1A CAXs can mediate the transport of multiple ions *in planta* and when heterologously expressed in yeast (Kamiya et al., 2005; Shigaki et al., 2005, 2010; Koren'kov et al., 2007; Mei et al., 2009). Thus plant CAXs have a broad cation specificity and hence the CAX family has more recently been referred to as H^+/cation exchangers rather than just H^+/Ca^{2+} exchangers. The Type 1B sub-group may contain more diverse CAX proteins. Two non-canonical Type 1B CAXs with an additional transmembrane (TM) insertion between TM4 and TM5 were recently reported from plants in the Asteraceae family (Jain et al., 2009), although the relevance of this alteration is unknown. Recently, a CAX gene from the chlorophyte microalgae *Chlamydomonas reinhardtii* was identified and suggested to group within the Type 1C sub-group alongside protozoan CAX genes (Pittman et al., 2009). Like the higher plant CAXs, CrCAX1 appears to have broad substrate specificity, and in addition to Ca^{2+} can also transport the monovalent ion Na^+.

All CaCA proteins share a similar topological model: on average 10 TM helices are predicted to form a Ca^{2+} transduction pathway which includes two conserved α-repeats regions within TM helices 2–3 and 7–8. These two α-repeats are important for ion selectivity, binding, and transport in diverse CaCA members (Kamiya and Maeshima, 2004; Ottolia et al., 2005; Shigaki et al., 2005) and are predicted from topology models to be orientated on opposite sides of the membrane (**Figure 1**), and in the folded structure of the protein face toward each other to form an ion conductance channel (Nicoll et al., 2007). This topology is exemplified by NCX (**Figure 1**). The sizes of individual members of the CaCA superfamily vary from ∼300 to 1000 amino acids, and although the core topological structure is conserved there are some exceptions; for example, the CCX proteins have additional TM domains within the C-terminal half of the protein (Cai and Lytton, 2004b). The clusters of 5 + 5 TM helices within both halves of a typical CaCA are separated by a large cytosolic loop (**Figure 1**) that regulates

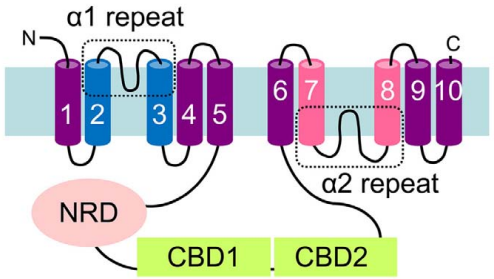


FIGURE 1 | Topological model of the Na^+/Ca^{2+} exchanger (NCX) protein. Barrel structures denote the transmembrane spanning domains identified by hydropathy analysis using TMHMM version 2. The conserved α1- and α2-repeat regions which span TM 2 and 3, and TM 7 and 8 are highlighted. The Ca^{2+}-binding domains (CBD) and the predicted Na^+ regulatory domain (NRD) present on the large central cytosolic loop are highlighted.

exchange activity in some members. It is likely that both halves of the protein formed following an ancient duplication event. While split halves of NCX or CAX proteins are not able to function alone (Ottolia et al., 2001; Zhao et al., 2009), it is not inconceivable that an ancestral form was useful as part of a related protein domain if this event occurred very early in the evolution of the gene family.

In most genomes putative membrane transporters abound with few transporters representing unique sequences in the genome (Bock et al., 2006). Despite fundamental differences in organismal complexity, an emerging theme is that diverse species have similar protein-coding potentials and alterations in CaCA transporters among taxa may be associated with modifications in function. The recent sequencing of genomes from diverse photosynthetic eukaryotes, including various higher plants, bryophytes, lycophytes, unicellular and multi-cellular green, red, and brown alga, and diatoms, act as powerful genomic resources for continued comparative analyses of CaCA proteins. Evolutional relationships among the characterized and uncharacterized CaCAs will lead to a better understanding of their potential roles. In this study we have performed a multi-genome identification of CaCA genes and a phylogenetic analysis that have provided insights into the evolution of CaCA transporters within higher plants and the green plant lineage.

MATERIALS AND METHODS

DATA SOURCES

Predicted protein-coding genes from complete and near complete genome sequences were downloaded for 74 species in May 2011 from a variety of databases (Table S1 in Supplementary Material). Species were selected for their broad taxonomic range in order to capture the full diversity of CaCA sequences, and where possible, photosynthetic organisms were included. The species examined included 12 land plants (including one lycophyte and one bryophyte), and 12 algae of various phyla and class (Table S1 in Supplementary Material), four fungi, eight other unicellular eukaryotes (including a variety of protists and oomycetes), seven animals (including two mammals), 21 bacteria (of which 10 were members of the cyanobacteria), and 10 archaea.

CaCA IDENTIFICATION

Putative CaCA proteins were identified using the basic local alignment search tool (BLAST; Altschul et al., 1997). A local database was constructed from the predicted protein sequences for the 74 species examined. The database was queried with 147 previously identified CaCA protein-coding sequences (Cai and Lytton, 2004a) including all known CaCA proteins from *A. thaliana*. The initial results contained a variety of putative ion transporter sequences from the CaCA superfamily and from other families, in particular those with known functions in Na^+/H^+, arginine, and citrate transport, which were discarded. To obtain sequences specific to the CaCA transporter family, only those matches with an *e*-value of <1, that spanned a region of >100 bp in length, with no less than 15% sequence identity, and where the predicted coding sequence corresponding to the match was >280 bp in length were considered for further analysis. The remaining 1328 candidate CaCA homologs were then screened by hydropathy analysis using TMHMM v2.0[1] (Krogh et al., 2001) to remove all non-TM proteins and to remove those whose hydropathy profile was clearly dissimilar to a selection of representative CaCA genes from the NCX, NKCX, CCX, CAX, and YRBG families from human, *A. thaliana*, *O. sativa*, *S. cerevisiae*, *E. coli*, and *Synechocystis* sp. Further screening of candidate sequences was performed by aligning them with a small selection of well-characterized representative CaCA proteins, and visual inspection was used to remove those that were clearly not members of the CaCA family. Finally 308 putative CaCA homologs were retained for phylogenetic analysis, plus 34 novel CaCA-like proteins (named EF–CAX) that were shown by Pfam analysis[2] (Finn et al., 2010) to contain putative Ca^{2+}-binding EF-hand domains (Pfam ID: PF00036) and so-called "Na_Ca_ex" (consensus CaCA) domains (Pfam ID: PF01699; Table S2 in Supplementary Material).

PHYLOGENETIC ANALYSES

Phylogenetic analysis of full-length CaCA proteins was not suitable because their diversity makes sequence alignment difficult. Several steps were taken in order to obtain reasonable alignments. Firstly, poorly conserved and extended N- and C-terminal regions of the predicted proteins were removed. TMHMM was then used to identify and extract putative TM domains for each protein. Furthermore, the greatest sequence distance between two TM domains was recorded and inferred to be the length of the central loop region. Visual inspection of preliminary alignments suggested that the number of TM domains can be variable between species, which may affect the accuracy of sequence alignment. To account for these differences we used MAFFT v6.0 (Katoh et al., 2002) under the E-INS-i option, which is specifically designed to cope with homologous regions separated by unalignable regions. Other approaches were also investigated for aligning these regions, and all yielded comparable results. The final phylogeny was estimated using maximum likelihood under the WAG + F model of amino acid substitution with Γ-distributed rates across sites, as implemented in RAxML v7.1 (Stamatakis, 2006). Other substitution models were examined and found to yield very similar or

[1] www.cbc.dtu.dk/services/TMHMM
[2] http://pfam.sanger.ac.uk

substantially worse log-likelihood values. Confidence in the tree was assessed using the fast bootstrap approach (Stamatakis et al., 2008). The tree was viewed using the FigTree program[3]. The resulting phylogeny appeared to be broadly consistent with previous findings, with five major clades corresponding to the CAX, CCX, NCX, NCKX, and YRBG genes. Thus gene products were classified according to their position in the phylogeny. For the phylogenetic analysis of the EF–CAX, due to poor sequence similarity between the N-terminal half of these proteins with N-terminal sequence of the other CaCA proteins, the conserved α2-repeat region sequence within the C-terminal half of CaCA and EF–CAX proteins was used instead. This sequence was identified using the consensus α2-repeat region sequence as determined previously for CaCA proteins (Cai and Lytton, 2004a), and extracted from each CaCA protein. The α2-repeat sequences were realigned to confirm that the correct consensus sequence was obtained and were then used to generate trees using maximum likelihood, as described above.

AMINO ACID SEQUENCE COMPARISONS

For the visual analysis and comparison of CaCA protein sequences, multiple sequence alignments were performed using ClustalW2[4] (Larkin et al., 2007) and shaded using the BOXSHADE v.3.2 viewer and manually annotated.

RESULTS AND DISCUSSION

IDENTIFICATION OF GENES ENCODING CaCA HOMOLOGS IN PHOTOSYNTHETIC EUKARYOTES

To examine the evolution and potential diversification of the CaCA superfamily in photosynthetic eukaryotes and throughout the land plant lineage, we screened the sequenced genomes of 10 flowering plant species. These included six dicots: *A. thaliana*, a tree species – poplar (*Populus trichocarpa*), grape (*Vitis vinifera*); and three legumes: soybean (*Glycine max*), *Medicago truncatula*, and *Lotus japonica*, and four monocots: rice (*O. sativa*), corn (*Zea mays*), sorghum (*Sorghum bicolor*), and *Brachypodium distachyon*. The genomes of two non-flowering land plants were available for analysis; the bryophyte moss *Physcomitrella patens*, and the lycophyte spikemoss *Selaginella moellendorffii*. A range of diverse algal species were also chosen. These comprised four freshwater algae: the model unicellular chlorophyte alga *C. reinhardtii*, the unicellular chlorophyte *Coccomyxa* sp. C-169 (formerly named *Chlorella* sp. C-169), and the multi-cellular chlorophyte *Volvox carteri*; and a selection of marine algae: the prasinophytes *Ostreococcus tauri* and *Micromonas pusilla*, the rhodophyte *Cyanidioschyzon merolae*, the haptophytes *Emiliania huxleyi*, the brown alga *Ectocarpus siliculosus*, the pelagophyte *Aureococcus anophagefferens*, and photosynthetic members of the stramenopiles, including the diatoms *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, and *Fragilariopsis cylindrus*. In addition, two non-photosynthetic stramenopiles, the oomycetes *Phytophthora sojae*, and *Pythium ultimum* were examined. A selection of mammalian, invertebrate, fungal, protozoan, bacterial, cyanobacterial, and archaeal species with completed genome sequences were chosen that provide a cross section of life for comparison with the plant and algal species (Table S1 in Supplementary Material). Using a dataset of previously identified CaCA query sequences, the genomes of all these species were screened by BLAST analysis and the results were further screened by hydrophobicity analysis, sequence alignment, and individual protein BLAST to validate each of the outputs and generate a final list of putative CaCA proteins (see Materials and Methods). In 13 of the species screened (including five bacteria and seven archaea) no obvious CaCA genes were identified (Table S1 in Supplementary Material). We were also unable to identify a clear-cut CaCA gene from the eukaryotic protozoan parasite *Giardia lamblia*. In contrast, the remaining 18 prokaryote species and all other eukaryotes examined including all of the plant and algal species possessed CaCA gene members (**Figure 2**). In all 342 CaCA genes were identified from 74 species (including the 13 species with no CaCA genes), giving an average of 4.6 CaCA genes per species. From the 12 land plants, 159 genes were determined (13.25 genes per species) and from the 12 algal species, 70 genes were identified (5.8 genes per species). The search method used multiple queries for each genome examined with the robust search method employed suggesting confidence that the search returned all expected genes and that CaCA genes are truly lacking in 13 of the species.

The classification of the genes into one of the five CaCA families was determined by phylogenetic analysis of the full superfamily. Initial amino acid sequence alignments using full-length sequence generated poor alignments and resulted in a low confidence tree. To improve the sequence alignments and generate the tree, the highly variable N- and C-terminal tail sequences and the hydrophilic loop sequences were therefore removed. The predicted size of the central loop region, the longest hydrophilic loop region of the CaCA protein, situated between TM5 and 6 of the consensus NCX (**Figure 1**), NCKX, CCX, and YRBG proteins, and between TM6 and 7 of the consensus CAX proteins, varied largely between the CaCA sequences, ranging from 19 amino acids in one of the CAX proteins to 602 amino acids in one of the CCX proteins. The NCX, NCKX, and CCX proteins all had on average a longer central loop region than the CAX and YRBG proteins; however, within all of these families there was significant variation in central loop length and the predicted number of TM domains (Table S3 in Supplementary Material). In particular, there was large variation within the NCX, NCKX, CCX, and CAX families. The significant variation within the CAX family was caused in part by the topological variation and variation in protein size between the Type 1 and Type 2 CAX protein sub-types (Shigaki et al., 2006). There was also large variation in the size of the N-terminal regions before the first TM span. Many of the animal NCX, NCKX, and CCX proteins are predicted to possess an N-terminal signal peptide (Lytton, 2007), while many of the plant CAX proteins possess a long N-terminal tail that has been shown to play a post-translational regulatory role (Pittman et al., 2002; Pittman, 2011), but which is absent in the fungal and bacterial CAX proteins.

The resulting tree indicates the presence of the five main CaCA families identified in previous analysis (Cai and Lytton, 2004a) although the bootstrap values at the base of the NCKX, NCX, and CCX branches were low (**Figure 3**). Likewise, some of the exact details for this tree and the individual family trees should be viewed with caution because long branches could lead to phylogenetic

[3]http://tree.bio.ed.ac.uk/software/figtree
[4]http://www.ebi.ac.uk/Tools/msa/clustalw2

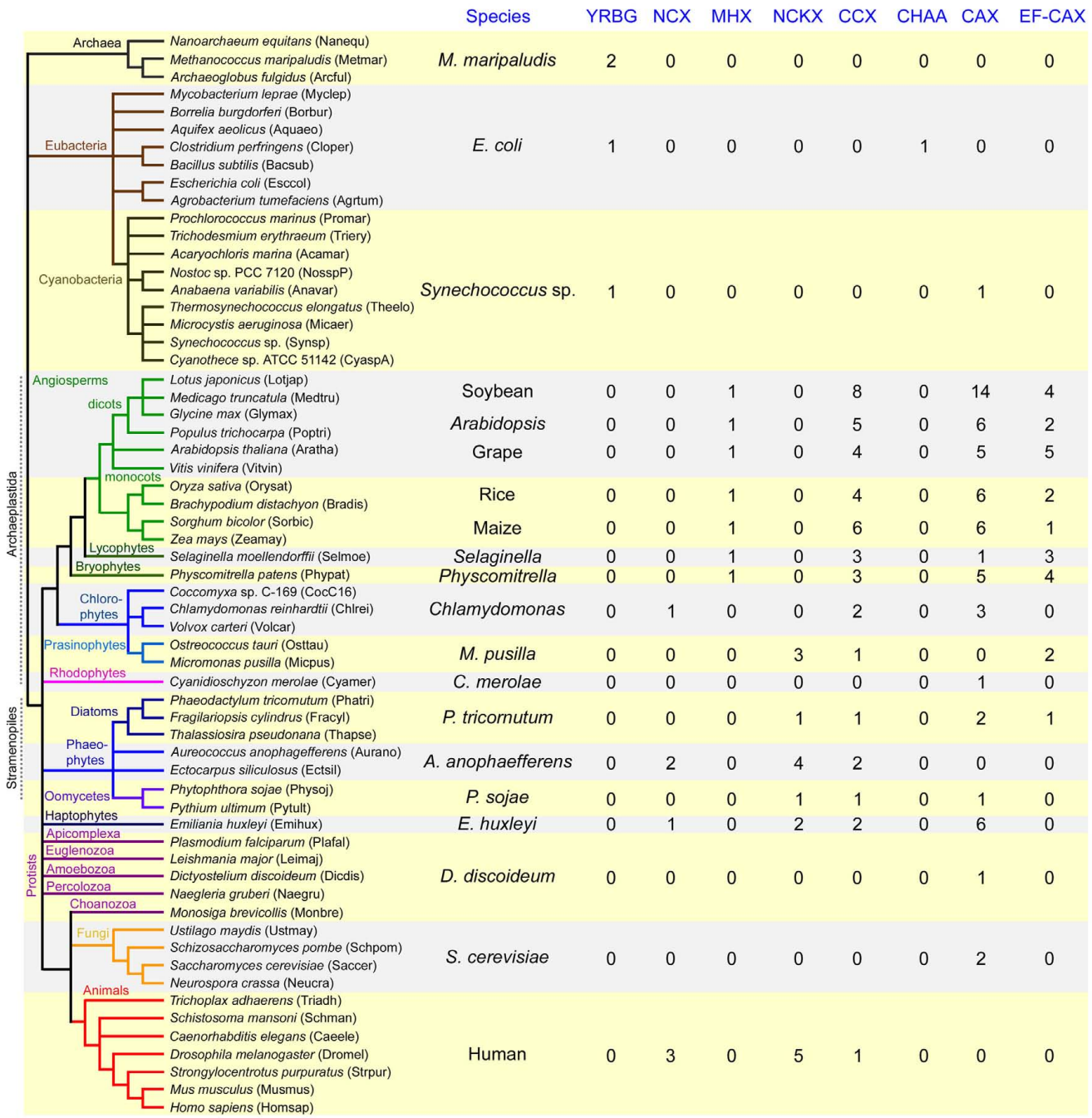


FIGURE 2 | Variation in the distribution of genes within each of the major CaCA gene families (YRBG, NCX, NCKX, CCX, CHAA, CAX) and sub-families (MHX, EF–CAX) across eukaryotic and prokaryotic organisms. Gene numbers determined by phylogenetic analysis from representative species for each of the classes of organisms examined are shown. The species from which CaCA genes were identified are listed (left hand side) with the species name abbreviation in parentheses. The evolutionary relationships of the species is shown, generated using iTOL (http://itol.embl.de) and based on the NCBI taxonomy tree. For the full gene distribution for each of the species analyzed see Table S1 in Supplementary Material.

artifacts such as long branch attraction. The YRBG family contains only prokaryotic genes, including *E. coli yrbG* and genes from other bacteria (including cyanobacteria), but also archaea (**Figure 3**). None of the prokaryotic species analyzed here contained genes that fell within the NCX, NCKX, and CCX families. These families possess higher plant (in the case of NCX and CCX), algal, and animal genes, with some fungal-derived genes in the CCX family. The algal genes are widely dispersed amongst these three

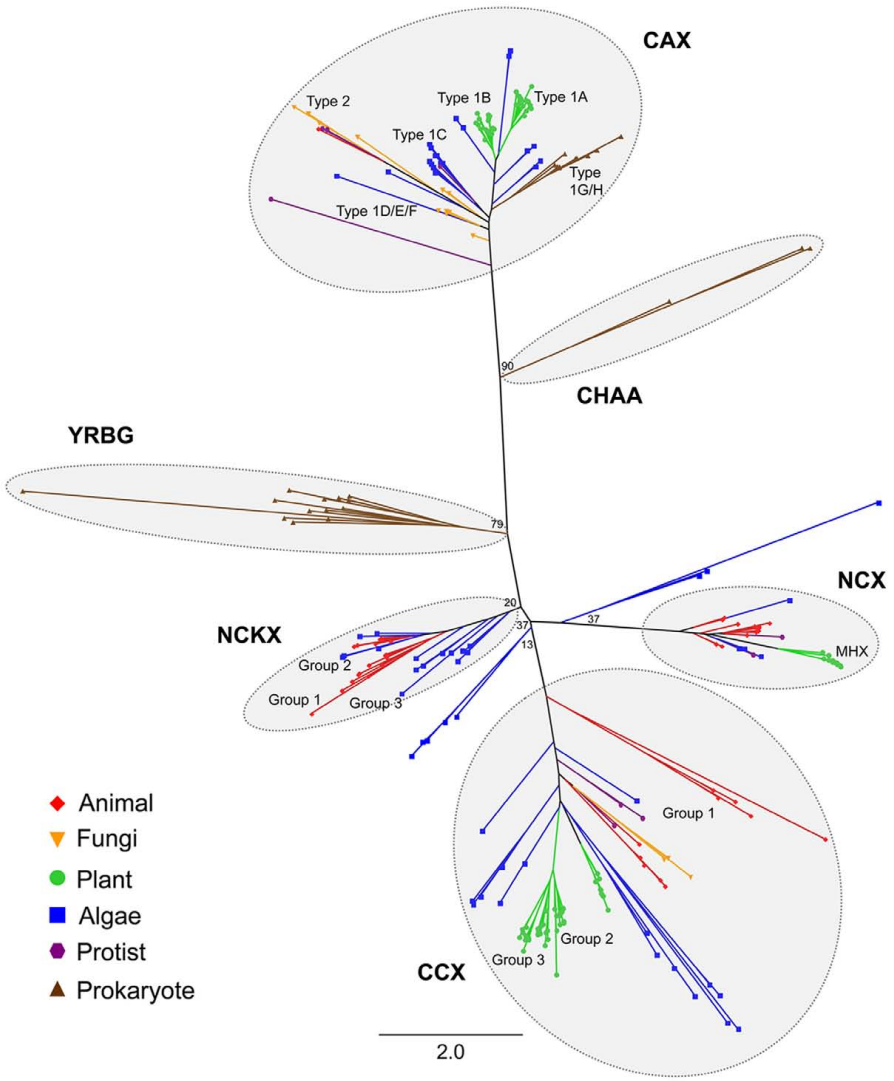


FIGURE 3 | A phylogenetic tree of the CaCA superfamily. The tree was constructed using the maximum likelihood method and derived from alignments of conserved hydrophobic region sequences identified from the genomes of selected animal, fungal, land plant, algae, protist, bacterial, and archaebacterial species, listed in **Figure 2**. Six major groups are highlighted. The sub-groups determined for the CAX, CCX, NCX, and NCKX families as shown in **Figures 4**, **6**, **7**, and **9**, are also given. Line colors and symbols denote the species class. Bootstrap values are indicated at the nodes of major branches. The branch length scale bar indicates the evolutionary distance of two amino acid substitutions per site.

families and it is mostly the algal gene product sequences that yield the longest branch lengths within the tree indicating significant sequence variation. Furthermore, algal genes were also identified which were basal to the NCX and NCKX groups and could not be distinguished between the two families based on phylogeny (**Figure 3**). The largest proportion of CaCA genes identified was from the CAX family (120 of the 342 genes, which is 1.6 CAX genes per species; Table S1 in Supplementary Material). These genes were from all the classes of organisms examined including some invertebrate animals, as shown previously (Shigaki et al., 2006), but there were no mammalian or insect CAX genes. The CAX family included some bacterial and cyanobacterial genes but none from any of the archaeal species examined. As with the CCX family, the size and branching of the CAX family indicates significant sequence variation and potential diversification. The *E. coli* protein ChaA, originally identified as a H^+/Ca^{2+} exchanger (Ivey et al., 1993) and subsequently suggested to function as a H^+/Na^+ exchanger (Ohyama et al., 1994), was not identified as a CaCA homolog by the BLAST search performed by Cai and Lytton (2004a). These authors commented that the ChaA protein had limited sequence similarity with other CaCA proteins. More recent phylogenetic analysis of the CAX family suggested that *chaA* falls within the Type 3C branch of the CAX tree (Shigaki et al., 2006). The BLAST search in this study did identify *E. coli chaA* and our global CaCA phylogenetic analysis indicates that *chaA* and other related genes from *Agrobacterium tumefaciens* and *Mycobacterium*

leprae are distinct from the other CAX genes and form a separate group which we have named here CHAA (**Figure 3**). Like the YRBG group, the CHAA group is specifically prokaryotic.

NCKX AND "ANIMAL-TYPE" NCX GENES ARE ABSENT IN LAND PLANTS BUT PRESENT IN ALGAE

Our analysis of the selected genomes indicated that NCX and NCKX genes appear to be ubiquitous in vertebrate and invertebrate species, as previously determined (Cai and Lytton, 2004a; On et al., 2008). In contrast, no genes from any of the higher plants, bryophyte, or lycophyte species were identified that group within the NCKX family (**Figure 3**). Likewise, no genes from any of the land plants were found to cluster closely with the NCX genes from animals (**Figure 4**), although more distantly related NCX-like genes were present (see below). Higher plants probably lack sufficient transmembrane Na^+ gradients to energize the counter transport of Ca^{2+} or other cations and hence have retained Ca^{2+}/H^+ exchangers instead of Na^+/Ca^{2+} exchange activity (Hirschi, 2001). NCX genes were however, detected in algae including the chlorophyte *Chlamydomonas*, the pelagophyte *A. anophagefferens*, and the haptophyte *E. huxleyi* (**Figure 4**). These algal NCX homologs have high sequence similarity with mammalian NCX proteins including conservation of the key residues in the α1- and α2-repeat regions (**Figure 5**) that are critical for Na^+/Ca^{2+} exchange activity (Lytton, 2007). Furthermore, the genes from both of the marine algae, which are more distantly related from the land plants and chlorophyte algae, cluster more closely with animal NCX genes (**Figure 4**), and we would predict them to have a similar function.

The NCKX tree also contains genes from various algal species (**Figure 6**). The NCKX tree has three main groups with Group 1 being animal-specific, containing invertebrate, insect, and mammalian genes, including human NCKX3, NCKX4, and NCKX5. Group 2 has animal genes including human NCKX1 and NCKX2. Algal NCKX genes from the two Stramenopile species *A. anophagefferens* and *P. tricornutum*, and the prasinophyte *Micromonas* cluster close to Group 2 but the lower bootstrap values and longer branch lengths of these genes do not allow us to confidently assign

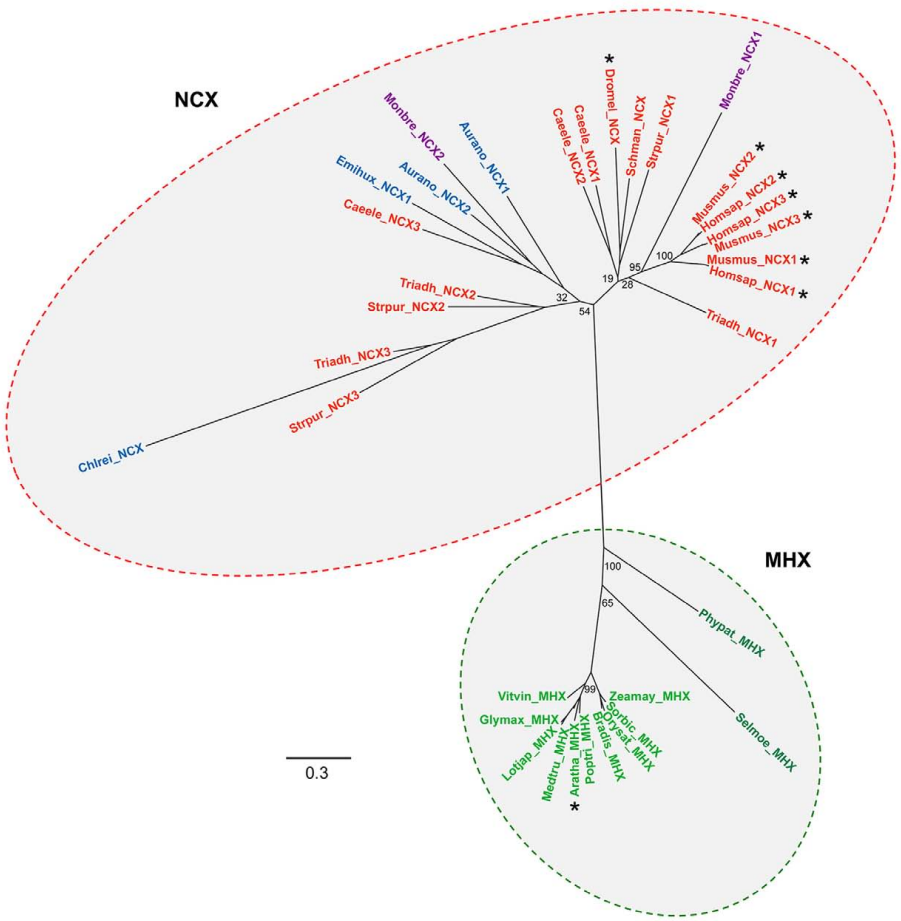


FIGURE 4 | Na^+/Ca^{2+} exchanger family phylogenetic tree. The tree was constructed using genes identified in the NCX group of the full CaCA tree (**Figure 3**). A separate NCX sub-group named MHX is highlighted which includes all the plant genes including *Arabidopsis* MHX. Branch labels include the species name abbreviation defined in **Figure 2** and are colored according to the species class using the same color code as in **Figure 3**. Bootstrap values are indicated at the nodes of major branches. The branch length scale bar indicates the evolutionary distance of 0.3 amino acid substitutions per site. Genes that are highlighted with an asterisk encode transporters for which the ion specificity has been determined.

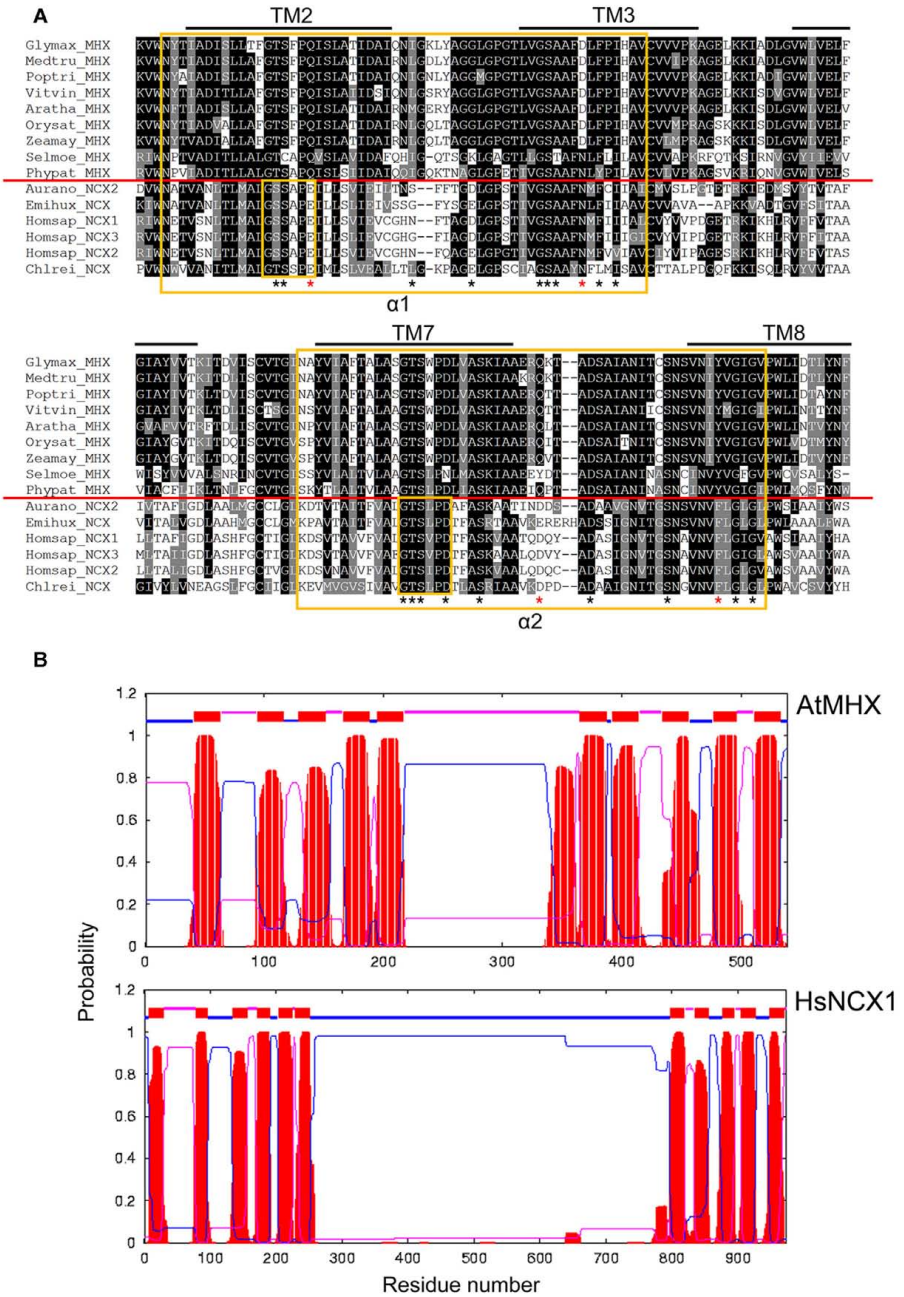


FIGURE 5 | Sequence and structural variation of NCX and MHX homologs. (A) Multiple amino acid sequence alignments of the α1- and α2-repeat region sequence from selected plant MHX, human NCX, and algal NCX proteins. Alignments were performed using ClustalW. Amino acids that are identical or similar are shaded black or gray, respectively. Predicted hydrophobic regions and putative transmembrane spans are over-lined. The α-repeat regions and signature residues are boxed in yellow. The red line separates the MHX and NCX sequences. Asterisks indicate residues shown to be important for Na^+/Ca^{2+} exchange activity in NCX1, with those in red being not conserved in the MHX sequences. **(B)** Topology models of AtMHX and HsNCX1 generated by TMHMM. Red areas indicate predicted TM spans.

them to Group 2 rather than Group 3. Group 3 is an algal NCKX group, including another *Micromonas* gene and two genes from the related *O. tauri* picoalgae, two genes from *E. huxleyi*, and two genes from the brown alga *E. siliculosus*, plus genes from the two oomycete species, that are related to the Stramenopile algae (**Figure 6**). It is notable that all of the algal NCKX genes are from marine species and from non of the freshwater species, suggesting a potential link between the high Na^+ and K^+ concentration in seawater and the presence of a putative K^+-regulated Na^+-coupled Ca^{2+} transporter. Indeed all of the marine algae examined here, with the exception of the red algae *C. merolae*, either possessed an NCKX gene or a gene that was basal to the

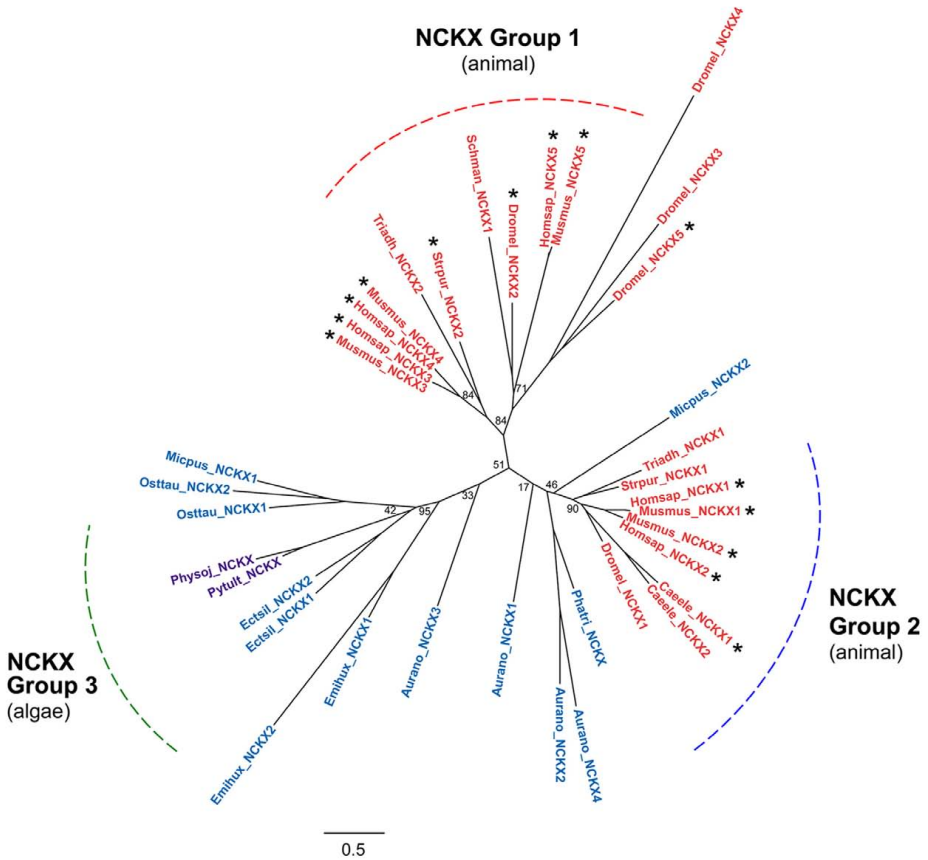


FIGURE 6 | $Na^+/Ca^{2+}/K^+$ exchanger family phylogenetic tree. The tree was constructed using genes identified in the NCKX group of the full CaCA tree (**Figure 3**). Three NCKX sub-groups are highlighted. Branch labels include the species name abbreviation defined in **Figure 2** and are colored according to the species class using the same color code as in **Figure 3**. Bootstrap values are indicated at the nodes of major branches. The branch length scale bar indicates the evolutionary distance of 0.5 amino acid substitutions per site. Genes that are highlighted with an asterisk encode transporters for which the ion specificity has been determined.

NCKX group (**Figure 2**), suggesting that these transporters may be critical for life in a marine environment. The apparent absence of NCX and NCKX genes from all land plants but their presence in various classes of algae, including the more closely related chlorophytes and prasinophytes, may suggest that they have been lost in all plants. However, it is also possible that these genes were also absent from the common ancestor species in the green eukaryote lineage and that the algae lineage subsequently acquired these genes through horizontal transfer.

Algal NCX and NCKX proteins have not yet been cloned and characterized, further work is needed to determine if these transporters encode Na^+/Ca^{2+} and K^+-dependent Na^+/Ca^{2+} exchange activities. However, plasma membrane Na^+/Ca^{2+} exchange activity has been detected in some algae, such as the marine chlorophyte *Dunaliella*, and in the freshwater chlorophyte *Chlamydomonas* when assayed in the presence of external Na^+, with a predicted ratio of 3 Na^+:1 Ca^{2+} (Karimova et al., 2000), equivalent to the stoichiometry of animal NCX. While the presence of a high external Na^+ concentration in marine algae and therefore a steep Na^+ gradient across the plasma membrane might be expected to make the presence of a Na^+-coupled ion transporter energetically favorable, it is perhaps somewhat surprising that a freshwater alga like *Chlamydomonas* possesses a putative NCX in addition to CAX genes. However, *Chlamydomonas* is also a soil-living microorganism and may encounter saline conditions in some environments. NCX and NCKX genes did not appear to be ubiquitous in all classes of algae, particularly other freshwater algae. They could not be identified in the freshwater chlorophytes *V. carteri* and *Coccomyxa* sp., or in the genome of the red alga *C. merolae*, although future analysis of these genomes will be needed to confirm that these gene losses and not due to poor genome sequence quality.

ADAPTIVE EVOLUTION OF MHX GENES IN LAND PLANTS

While no obvious NCX genes were detected across the plant genes analyzed, distantly related NCX-like genes were detected in all plant species, and formed a tight monophyletic group which includes the *A. thaliana AtMHX* gene, hence we named this group MHX (**Figure 4**). *AtMHX* was previously found to be closely related to mammalian NCXs but was shown to function as a H^+-coupled Mg^{2+} and Zn^{2+} exchanger and not a NCX (Shaul et al., 1999). A single *AtMHX* homolog was identified from the

closely related zinc-hyperaccumulator species *Arabidopsis halleri* (Elbaz et al., 2006) but the extent of MHX genes within plants and other species has not been previously examined. The putative *Paramecium* Mg^{2+} exchanger *XNTA* appears to be a CaCA homolog but has significant divergence from the NCX gene family and therefore is not regarded as a MHX homolog (Haynes et al., 2002). An NCX-like gene was identified in each of the flowering plant genomes examined. Each of these gene sequences had high identity to *AtMHX* and formed a clear MHX sub-group within the NCX family (**Figure 4**). None of the plants examined had more than one MHX gene in contrast to the variation in numbers of CAX and CCX genes between plant species. Single MHX genes were also present in the lower land plants *Physcomitrella* and *Selaginella* but MHX genes do not appear to be present in algae as none of the NCX or NCX-related genes from any of the algal species grouped with the plant MHX genes (**Figures 3** and **4**). Furthermore, MHX-like genes were not present in any non-plant species. The birth of MHX genes appears to be an example of adaptive evolution which has taken place most probably in the common ancestor of land plants.

There is 20–22% amino acid sequence identity (31–33% similarity) between the plant MHX sequences and human NCX1, and both sets of proteins share a number of conserved residues particularly within the hydrophobic regions and in the α1- and α2-repeat regions, yet there are clear distinctions (**Figure 5A**). Many of the residues that have been shown to critical for Na^+/Ca^{2+} exchange activity in NCX1 (Nicoll et al., 1996; Iwamoto et al., 2000; Philipson and Nicoll, 2000; Ottolia et al., 2005) are conserved in the MHX sequences but there are some specific differences. Key residues Glu-113 and Asp-130 in the α1-repeat region of NCX1, and Asp-825 and Phe-844 in the NCX1 α2-repeat region, are substituted in the equivalent positions in each of the plant MHX homologs (**Figure 5A**). For example, Glu-113 is replaced by a Gln residue in each MHX protein. Such a G113Q substitution was shown to abolish Na^+-dependent Ca^{2+} transport activity of NCX1 (Nicoll et al., 1996). Mutations of Asp-130 and Asp-825 of NCX1 were found to substantially reduce the affinity of the exchanger for Ca^{2+} (Iwamoto et al., 2000). Asn-143 of NCX1 is also a critical residue and is substituted by Asp in the higher plant MHX proteins but not the lower plant MHXs (**Figure 5A**). Interestingly, the N143D mutation of NCX1 showed wild type activity (Ottolia et al., 2005). These amino acid changes in the α1-repeat region of AtMHX compared to NCX1 therefore appear to explain the lack of Na^+/Ca^{2+} exchange activity by AtMHX and probably by the other plant MHX transporters. The MHX genes also encode proteins with a much shorter central loop than the NCX genes (**Figure 5B**; Table S3 in Supplementary Material) and lack the Ca^{2+} binding domain sequence motifs present on the NCX proteins in this region (Lytton, 2007), further explaining the loss of Ca^{2+} activity and regulation by the AtMHX transporter. The residues in the α1-repeat region are thought to be involved in ion selectivity (Shigekawa et al., 2002; Ottolia et al., 2005) and therefore we might presume that the residues within the MHX α1-repeat region encode the Mg^{2+} and Zn^{2+}/H^+ exchange activity observed for AtMHX. Whether each of the plant MHX transporters is involved in Mg^{2+} transport or in the transport of other cations remains to be seen.

THE DIVERSIFICATION OF CCX GENES IN LAND PLANTS

To date only two CCX members have been functionally characterized in depth, mammalian NCKX6/NCLX, which has been shown to have $Na^+(Li^+)/Ca^{2+}$ exchange activity (Cai and Lytton, 2004b; Palty et al., 2004) and CCX3 from *Arabidopsis* which appears to function as a H^+/K^+ exchanger and which can also transport Na^+ and Mn^{2+} but not Ca^{2+} (Morris et al., 2008). Thus, there is clearly variation in function within the CCX transporter family as indicated by the CCX tree (**Figure 7**). Apart from the duplication of insect and nematode CCX genes (four to five genes), single genes are present in mammal and fungi species, and one or two CCX genes were present in each algal species examined, except in *Coccomyxa* sp. and *C. merolae* where CCX genes could not be identified. In contrast, there has been a significant diversification in CCX genes within land plants, with the bryophyte and lycophyte species each possessing three genes, and individual flowering plant species possessing three to eight CCX genes (**Figure 2**), suggesting that these are important genes within plants. There is significant variation between these plant CCX genes which are found within each of the three main groups of this tree (**Figure 7**). There are two plant-specific groups; Group 3 contains solely angiosperm genes and includes *Arabidopsis CCX1* and *CCX2*, while Group 2, which includes *Arabidopsis CCX3* and *CCX4*, also contains CCX genes from *Physcomitrella* and *Selaginella*. Each of the 10 flowering plant species analyzed has at least one member in each of these two groups. There is evidence of recent gene duplication within both groups, particularly within Group 3 where there are mostly two genes for each dicot species and three genes for each monocot species, while within Group 2 three of the dicot species (*Arabidopsis*, soybean, and poplar), plus *Physcomitrella* and *Selaginella*, have two CCX genes. Group 1 contains all the algal and non-plant genes, including *NCKX6/NCLX*, single CCX genes from *Physcomitrella* and *Selaginella*, and single CCX genes each from seven of the 10 higher plant species, including *CCX5* from *Arabidopsis*, but CCX5-like genes from the remaining three monocot species (*Brachypodium*, rice, and sorghum) appear to have been lost (**Figure 7**). There is considerable variation among the algal CCX genes which are widely dispersed throughout Group 1, as indicated by their long tree branch lengths, but as yet nothing is known about these genes or their function to explain this variation.

Further biochemical analysis of the plant CCX proteins is required before we can discern whether the CCX members from each group differ in substrate specificity. However, it is tempting to speculate that members of Group 2 may share AtCCX3 H^+/K^+ exchange activity and that the plant members of Group 1 like AtCCX5 may be more likely to have Ca^{2+} transport activity as they are more closely grouped to NCKX6/NCLX. A closer look at the amino acid sequence of selected plant, algal, and human CCX sequences from each of the three groups shows that they all have strong sequence conservation within the α1- and α2-repeat regions, including 100% identity in the α1 signature motif GNGAPD and the α2 signature motif G(N/D)SxGD (**Figure 8**). Yet there are still some obvious sequence differences in both α-repeat regions between the Group 1 and the Group 2/3 CCX proteins, in particular between the Ca^{2+} transporting NCKX6/NCLX and the non-Ca^{2+} transporting AtCCX3, that might determine cation specificity.

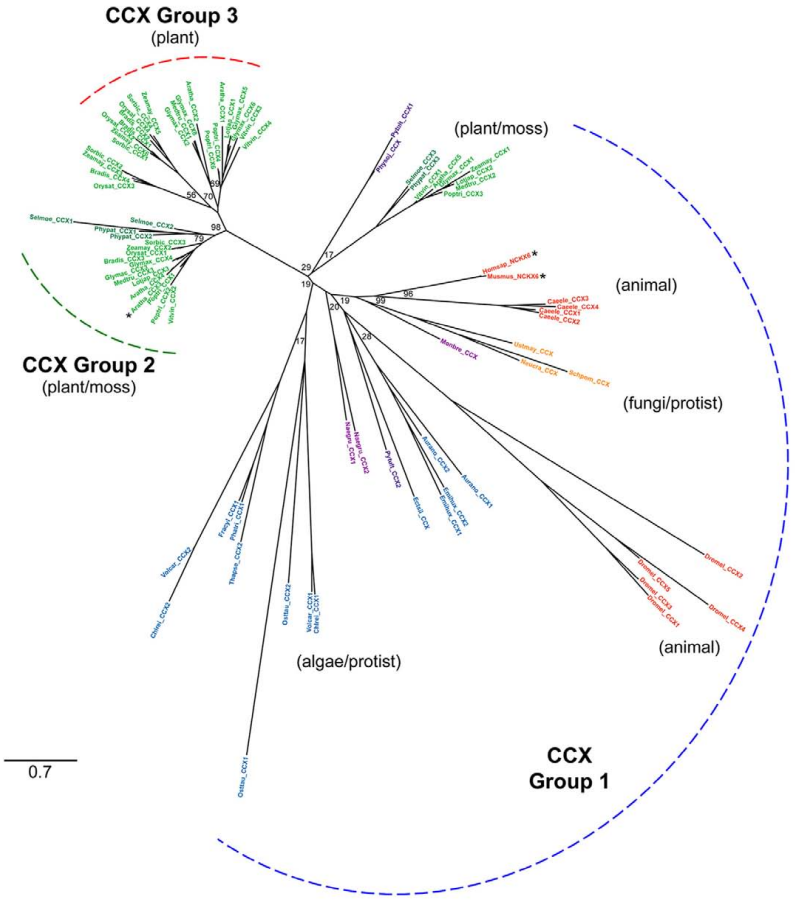


FIGURE 7 | Cation/Ca^{2+} exchanger family phylogenetic tree. The tree was constructed using genes identified in the CCX group of the full CaCA tree (**Figure 3**). Three CCX sub-groups are highlighted. Branch labels include the species name abbreviation defined in **Figure 2** and are colored according to the species class using the same color code as in **Figure 3**. Bootstrap values are indicated at the nodes of major branches. The branch length scale bar indicates the evolutionary distance of 0.7 amino acid substitutions per site. Genes that are highlighted with an asterisk encode transporters for which the ion specificity has been determined.

EXPANSION OF THE CAX GENE FAMILY THROUGHOUT THE PLANT KINGDOM

Previous analysis has examined the diversification and evolution of CAX genes and identified distinct phylogenetic groups that encompass genes from bacterial, fungal, plant, and invertebrate species (Shigaki et al., 2006). Here we have screened through a wider range of flowering plant genomes and a wide selection of algal species, from which CAX genes have not been examined in detail. Shigaki et al. (2006) identified a number of distinct Type 1 phylogenetic groups, many of which were comprised of fungal (Type 1D, 1E, and 1F) and bacterial/cyanobacterial (Type 1G and 1H) CAX genes. As this study has focused predominantly on plant and algae species, only a small number of selected fungal and bacterial species were examined and therefore the CAX tree generated here lacks distinction between the fungal and bacterial Type 1 groups (**Figure 9A**). The *Leishmania* parasitic protist gene stands out with an unusually long branch that may be reflective of the effect of the parasitic lifestyle on molecular evolution. The Type 2 group which contains distinct CAX genes including yeast *VNX1* and invertebrate CAX genes was also clearly observed in this tree, but this group contains no plant or algal genes. A number of CAX genes were identified from various algal species, although three of the species, *O. tauri*, *M. pusilla*, and *A. anophagefferens* appeared to lack CAX homologs (**Figure 2**). A previously cloned CAX gene from *Chlamydomonas* was shown to cluster within the Type 1C group which includes CAX genes from parasitic protists such as *Plasmodium falciparum* (Pittman et al., 2009). Genes from other freshwater chlorophyte algae including a second gene from *Chlamydomonas*, likewise fall within this group, plus genes from some of the marine algae and the oomycetes. However, five CAX genes from the haptophyte alga *E. huxleyi*, two stramenopile algae genes, and two chlorophyte algae genes from *Chlamydomonas* and *Volvox*, cluster more closely to the land plant CAX genes and might be considered as Type 1B-like genes. It might therefore be expected that these algal CAX homologs share some functional characteristics of the plant Type 1B CAX transporters. Despite the sequence divergence between *Chlamydomonas* CAX1 in the Type 1C group and *Arabidopsis* CAX1 in the Type 1A group, both of these transporters have very similar function; when expressed in a yeast heterologous expression

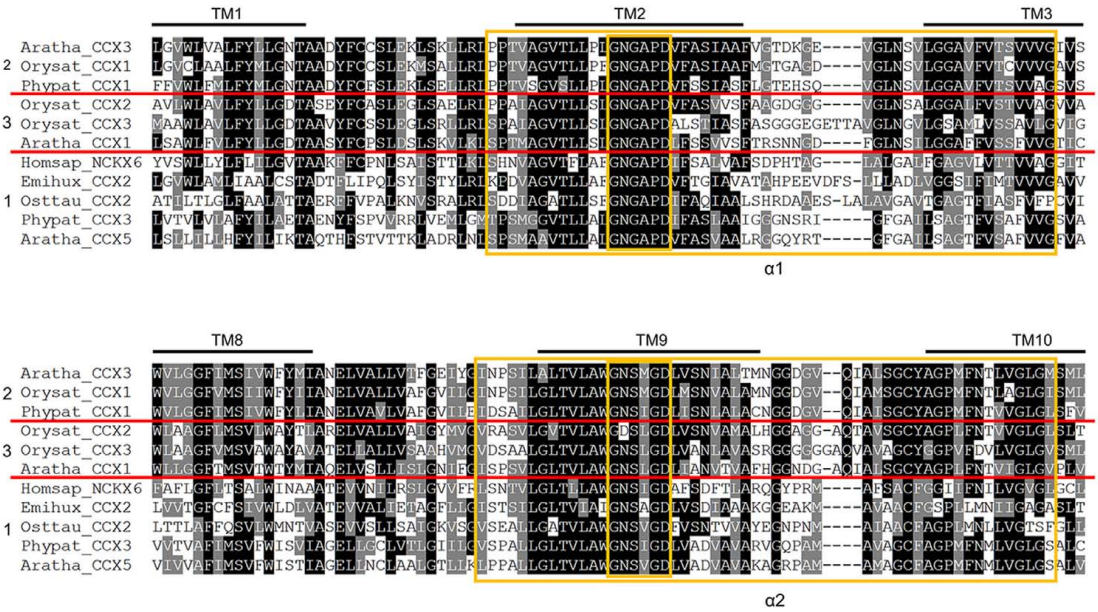


FIGURE 8 | Sequence variation of CCX homologs. Multiple amino acid sequence alignments of the α1- and α2-repeat region sequence from selected plant, moss, algal, and human CCX proteins. Alignments were performed using ClustalW. Amino acids that are identical or similar are shaded black or gray, respectively. Predicted transmembrane spans are over-lined. The α-repeat regions and signature residues are boxed in yellow. The red lines separate the Group 1, 2, and 3 CCX sequences.

system both function as vacuolar-localized H^+/Ca^{2+} exchangers with equivalent regulatory characteristics (Pittman et al., 2009). The main functional difference appears to be that CrCAX1 can also transport Na^+. To date all of the biochemically characterized CAX transporters show Ca^{2+} transport activity, including those in the Type 2 group (Manohar et al., 2010), therefore it is not clear why there is so much sequence diversity within the CAX family compared to the NCX or NCKX families which yield much tighter families (**Figure 3**).

The previously analyzed flowering plant CAX genes were found to cluster into either the Type 1A or Type 1B groups, and this was consistently seen here with all the flowering plant genomes examined. Within the Type 1A group, clear distinction between the genes from monocot and dicot plants was observed (**Figure 9B**) but this was not as obvious within the Type 1B group which showed greater overall divergence and also included the CAX genes from *Physcomitrella* and *Selaginella* (**Figure 9C**). This analysis indicates that there has been a diversification of CAX genes within land plants. Within the Chlorophyta, the species possess one to three CAX genes; in contrast, in the land plants each species has on average five to six CAX genes (**Figure 2**). This expansion may have occurred prior to the evolution of the flowering plants as the bryophyte *Physcomitrella* has five CAX genes. The most parsimonious explanation is that there was likely to have been at least one duplication event in the flowering plant ancestor. However, the evolution of CAX genes in plants seems to have been a dynamic process with many apparent instances of gains and losses, for example, the lycophyte *Selaginella* appears to have just a single CAX gene. With two exceptions, all the monocots and dicots analyzed here have between four and six CAX genes, generally equally divided between the Type 1A and Type 1B groups. 14 CAX genes were identified from soybean, also spread between both groups, but we could only identify two CAX genes from *L. japonica*, both within the Type 1A group. The genes in most of these species await future studies to see whether any novel functional traits will be uncovered or whether the range of characteristics determined for the *Arabidopsis* CAX proteins is conserved for the other flowering plants.

IDENTIFICATION OF GENES ENCODING CAX-LIKE PROTEINS CONTAINING EF-HAND DOMAINS IN SOME PLANTS AND ALGAE

In addition to the CaCA genes described, the BLAST search identified a number of genes which were clearly distinct from the others. Further analysis indicated that they were most closely related to the plant CAX genes. Amino acid sequence alignment showed that these gene products have higher similarity with selected plant CAX sequences in the C-terminal half of the proteins but the sequences aligned poorly against the N-terminal half of the CAX proteins (**Figure 10**) despite having an overall CAX-like topology including 6 + 5 TM domains separated by a central loop, although this loop is longer than in most CAX proteins (**Figure 11A**). Pfam analysis (Finn et al., 2010) was performed on these proteins and confirmed that they contain a "Na_Ca_ex" domain (Pfam ID: PF01699) in the C-terminal half of the protein. This domain denotes the core consensus 5 TM hydrophobic region present in NCXs and in all other CaCA proteins[5] including CAX proteins (**Figure 11A**). A closer look at the sequence of these proteins found that there was a high number of similar residues shared with the α2-repeat region of the CAX sequences including the presence of residues conserved throughout CaCA proteins (Cai and

[5]http://pfam.sanger.ac.uk/family/Na_Ca_ex

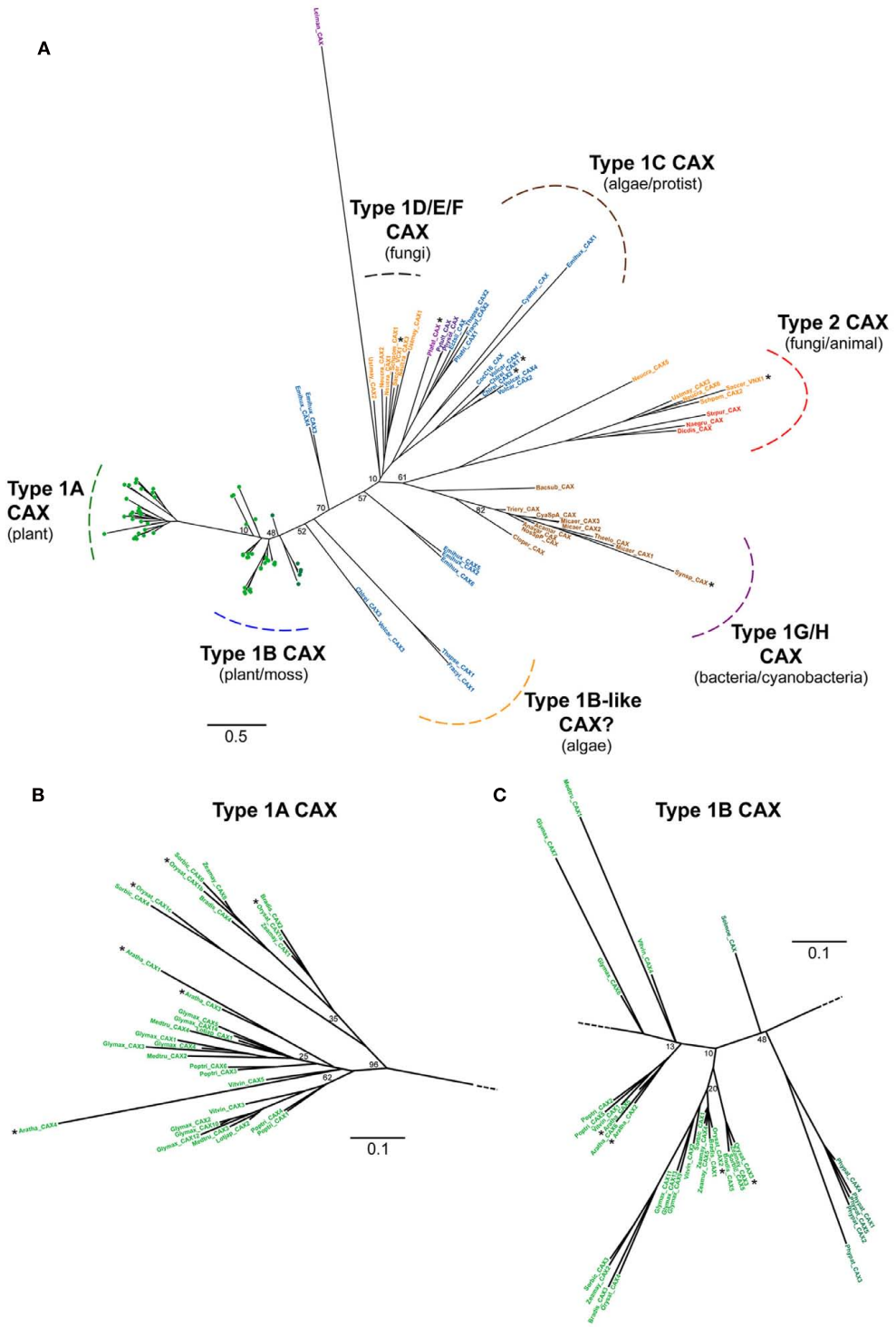


FIGURE 9 | CAX family phylogenetic tree. (A) The tree was constructed using genes identified in the CAX group of the full CaCA tree (**Figure 3**). The Type 1A-H and Type 2 CAX sub-groups are highlighted and determined using the nomenclature described previously (Shigaki et al., 2006). Branch labels include the species name abbreviation defined in **Figure 2** and are colored according to the species class using the same color code as in **Figure 3**. The individual gene labels for the Type 1A and 1B groups are replaced by circles and the gene name labels are given in the enlarged tree sections in **(B,C)**. Bootstrap values are indicated at the nodes of major branches. The branch length scale bar indicates the evolutionary distance of 0.5 amino acid substitutions per site [0.1 amino acid substitutions per site in **(B,C)**]. Genes that are highlighted with an asterisk encode transporters for which the ion specificity has been determined.

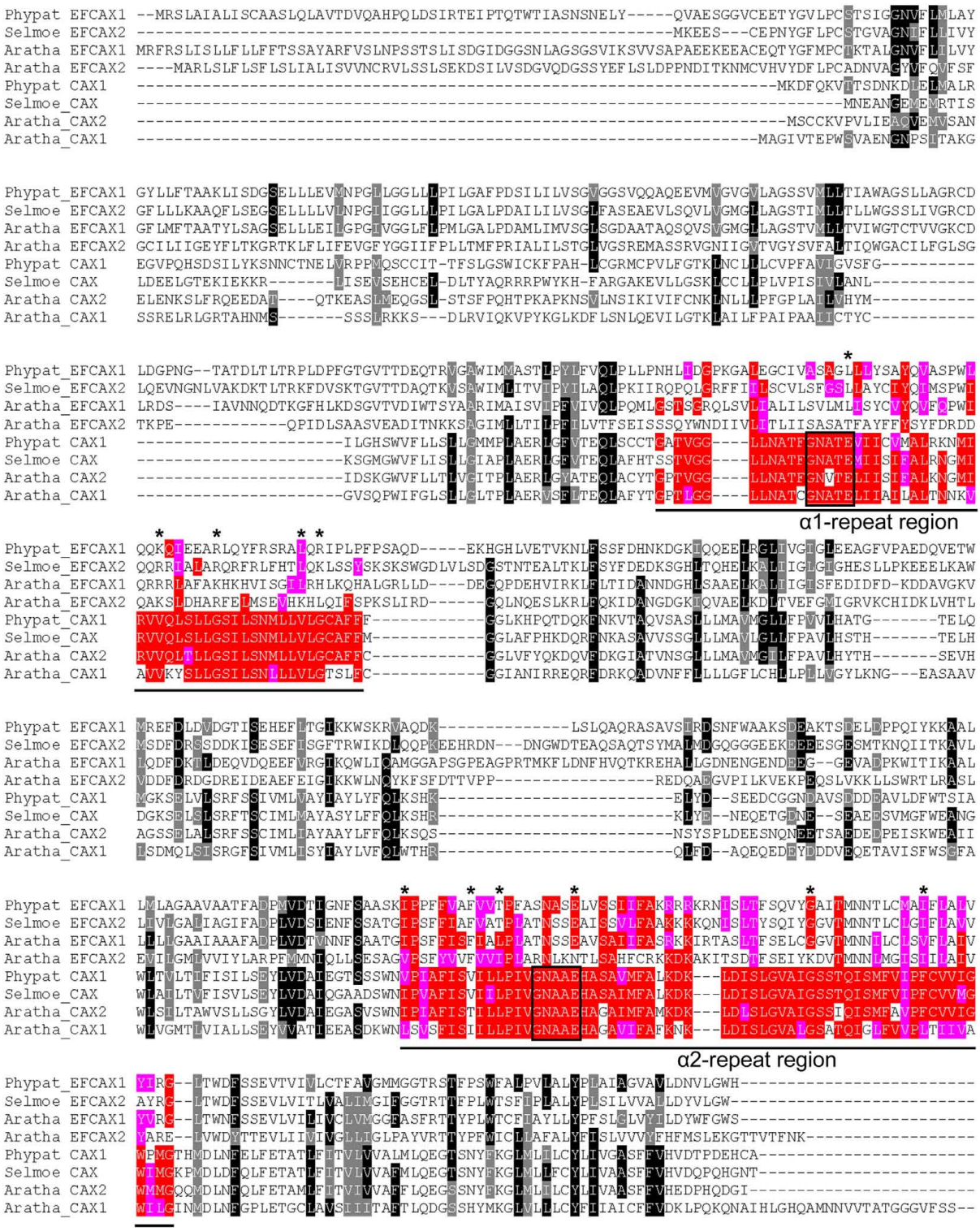


FIGURE 10 | Multiple sequence alignment of CAX and EF–CAX sequences. Full-length amino acid sequences of selected *Arabidopsis*, *Physcomitrella*, and *Selaginella* CAX and EF–CAX proteins were aligned. The alignment was performed using ClustalW. Amino acids that are identical or similar are shaded black/red or gray/pink, respectively. The α-repeat regions and signature residues are highlighted. Amino acids that are conserved in >90% of CaCA proteins (Cai and Lytton, 2004a) are highlighted (asterisk) within the α1- and α2-repeat regions.

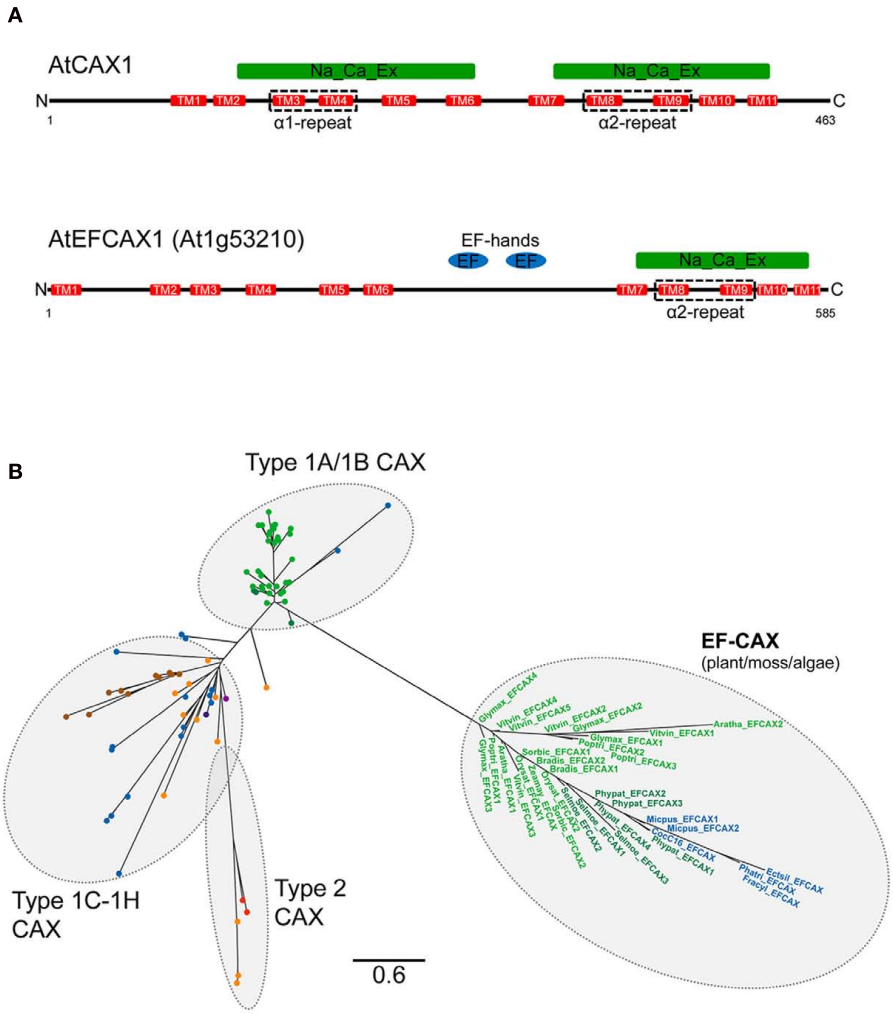


FIGURE 11 | Structural and phylogenetic comparisons of CAX and EF–CAX proteins. (A) Schematic representation of the topology structure of a typical CAX protein (*Arabidopsis* CAX1) and a representative EF–CAX protein (*Arabidopsis* EFCAX1). The black line denotes the protein from N-terminus to C-terminus, red bars denote the TM domains as predicted by TMHMM, and the numbers indicate the first and last amino acid residues. The locations of the α1- and α2-repeat regions are highlighted. The identified Pfam domains are indicated above the protein schematic. **(B)** CAX gene family phylogenetic tree including the EF–CAX genes. The tree was constructed using the maximum likelihood method and derived from alignments of α2-repeat region sequences extracted from genes identified in the CAX group of the full CaCA tree (**Figure 3**) and the EF–CAX genes as determined by Pfam analysis. The CAX sub-groups are highlighted in addition to the EF–CAX sub-group. Branch labels include the species name abbreviation defined in **Figure 2** or are replaced by a circle and are colored according to the species class using the same color code as in **Figure 3**. The branch length scale bar indicates the evolutionary distance of 0.6 amino acid substitutions per site.

Lytton, 2004a) but poor sequence similarity with the α2-repeat region (**Figure 10**). In addition the CAX-like proteins possess one or more EF-hand-type Ca^{2+}-binding domains (Pfam ID: PF00036) within the long central loop (**Figure 11A**). One of these genes from *Arabidopsis* (At1g53210) has been identified from a genome search (Shigaki et al., 2006), but as yet they have not been cloned or characterized. From all of the genomes analyzed here, 34 of these CAX/EF-hand (EF–CAX) genes were identified, predominantly from land plants (*Arabidopsis*, poplar, grape, soybean, *Brachypodium*, maize, rice, and sorghum) but also from *Physcomitrella*, *Selaginella*, and from some diverse algae species: the chlorophyte *Coccomyxa* sp., the prasinophyte *M. pusilla*, and three of the stramenopile algae (*E. siliculosus*, *F. cylindrus*, and *P. tricornutum*). Interestingly, *Selaginella* has more EF–CAX genes (three genes) than CAX genes (one gene), while *M. pusilla*, which has two EF–CAX genes, has no identifiable CAX genes. These genes were not identified from any of the bacterial, fungal, or animal genomes, but the presence of these genes in these select algae species suggests that they either evolved once in an ancestral organism from an ancestral CAX gene, or on at least two occasions, such as within the Stramenopiles and within the Archaeplastida. Using the α2-repeat region sequence alone to generate a phylogenetic tree, we confirmed that the EF–CAXs are more closely related to the CAX family members (**Figure 11B**) than to any of the other

CaCA families (data not shown). There are some long branch lengths within the EF–CAX clade indicating a degree of divergence between these genes, although the divergence between plant and algae EF–CAX genes appears to be less than between plant and algae CAX genes. Furthermore, the analysis demonstrated that the EF–CAX genes are more distant from the plant Type 1A and Type 1B CAX genes than the plant Type 1A/1B CAX genes are from the non-plant CAX genes (**Figure 11B**).

The functions of the EF–CAX proteins are unknown and future work will be need to determine if these proteins function as a transporter or merely as a Ca^{2+} binding protein. It is unclear how effective a CaCA transporter would be with a single α-repeat region. Experiments with animal NCX proteins indicates that both α-repeat regions are required for transport function and it has been suggested that together these regions interact within the tertiary structure to form an ion conductance pore (Nicoll et al., 2007). For the *Arabidopsis* CAX proteins, it has been shown that separated N- or C-terminal halves of CAX1 or a mixture of N-terminal CAX1 with N-terminal CAX3 lack transport activity (Zhao et al., 2009). In contrast, a mutated NCKX6/NCLX protein containing either a single α1- or α2-repeat region appears to function; however, oligomerization between two copies of these α1- or α2-repeat proteins thus yielding a protein with two α-repeat regions is required (Palty et al., 2006). If the EF-CAX proteins were to function as a transporter, these proteins may need to form oligomers to allow transport activity to occur. Alternatively, one possibility might be that the N-terminal half of the EF-CAX protein contains a region which can function as an α-repeat region but that the sequence has changed significantly from the consensus α-repeat sequence and is no longer recognizable. Clearly future functional analysis is needed to understand these novel proteins.

CONCLUSION

Ca^{2+} homeostasis is critical for all organisms and therefore a suite of Ca^{2+} regulators including transporters must be present. The CaCA superfamily represents an ancient class of genes which have members present in most species. A major function of these transporters is in Ca^{2+} transport as is evident from the functional analysis of CaCA gene products from different classes of organisms and from each of the different CaCA families: for example, *E. coli* yrbG, *Drosophila* NCX, *C. elegans* NCKX1, human NCKX6, and rice CAX1a all appear to mediate Ca^{2+} transport activity (Schwarz and Benzer, 1997; Szerencsei et al., 2000; Cai and Lytton, 2004b; Kamiya and Maeshima, 2004; Naseem et al., 2008), yet the analysis here shows that they have minimal sequence conservation. It is therefore difficult to predict the function and ion specificity of these many putative transporters, yet some of them are certainly likely to function in the transport of ions other than Ca^{2+}. This divergence in function is particularly apparent in higher plants, such that the *Arabidopsis* NCX and CCX proteins AtMHX and AtCCX3 possess Mg^{2+} transport and K^+ transport activity, respectively, rather than Ca^{2+} transport (Shaul et al., 1999; Morris et al., 2008). Overall, the CaCA complement and function in plants is different than in most animals. The plants possess a wide diversity of CAX genes, some or all of which encode proteins that will have Ca^{2+} transport and probably other cation transport function, a diversity of CCX genes with mostly unknown function, and a single MHX genes. This diversification appears to mirror the evolution of plants on land, as shown by the analysis of *Physcomitrella* and *Selaginella* in this study, which also multiple CAX and CCX genes plus single MHX genes. In contrast, many of the algal species have few or no CAX and CCX genes, which are divergent from those of the land plants, and animal-like NCX and NCKX genes. Recent analysis of Ca^{2+} channels in algal species has likewise found that many of the animal Ca^{2+} channel genes that are lost in land plants are present in algae (Wheeler and Brownlee, 2008; Verret et al., 2010), indicating that many algal species possess animal-like systems for Ca^{2+} signaling and suggesting that there are significant differences in Ca^{2+} signaling and homeostasis mechanisms between algae and land plants. The motile lifestyle of most algae species compared to the sessile land plants plus the significant differences in environmental conditions between the sets of organisms may in part explain the presence of very different Ca^{2+} regulatory mechanisms. Algae may therefore represent an excellent model for studying the evolution of Ca^{2+} signaling and aspects of animal-like signaling.

ACKNOWLEDGMENTS

We thank Toshiro Shigaki for comments on the manuscript and Mousheng Wu for help preparing **Figure 1**.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Physiology/10.3389/fpls.2012.00001/abstract

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Blackford, S., Rea, P. A., and Sanders, D. (1990). Voltage sensitivity of H^+/Ca^{2+} antiport in higher-plant tonoplast suggests a role in vacuolar calcium accumulation. *J. Biol. Chem.* 265, 9617–9620.

Blaustein, M. P., and Lederer, W. J. (1999). Sodium/calcium exchange: its physiological implications. *Physiol. Rev.* 79, 763–854.

Bock, K. W., Honys, D., Ward, J. M., Padmanaban, S., Nawrocki, E. P., Hirschi, K. D., Twell, D., and Sze, H. (2006). Integrating membrane transport with male gametophyte development and function through transcriptomics. *Plant Physiol.* 140, 1151–1168.

Cagnac, O., Aranda-Sicilia, M. N., Leterrier, M., Rodriguez-Rosales, M.-P., and Venema, K. (2010). Vacuolar cation/H^+ antiporters of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285, 33914–33922.

Cagnac, O., Leterrier, M., Yeager, M., and Blumwald, E. (2007). Identification and characterization of Vnx1p, a novel type of vacuolar monovalent cation/H^+ antiporter of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 282, 24284–24293.

Cai, X. J., and Lytton, J. (2004a). The cation/Ca^{2+} exchanger superfamily: phylogenetic analysis and structural implications. *Mol. Biol. Evol.* 21, 1692–1703.

Cai, X. J., and Lytton, J. (2004b). Molecular cloning of a sixth member of the K^+-dependent Na^+/Ca^{2+} exchanger gene family, NCKX6. *J. Biol. Chem.* 279, 5867–5876.

Case, R. M., Eisner, D., Gurney, A., Jones, O., Muallem, S., and Verkhratsky, A. (2007). Evolution of calcium homeostasis: from birth of the first cell to an omnipresent signalling system. *Cell Calcium* 42, 345–350.

Cheng, N. H., Pittman, J. K., Barkla, B. J., Shigaki, T., and Hirschi, K. D. (2003). The *Arabidopsis* cax1 mutant exhibits impaired ion homeostasis, development, and hormonal responses and reveals interplay among vacuolar transporters. *Plant Cell* 15, 347–364.

Conn, S. J., Gilliham, M., Athman, A., Schreiber, A. W., Baumann, U., Moller, I., Cheng, N. H., Stancombe, M. A., Hirschi, K. D., Webb, A. A. R., Burton, R., Kaiser, B. N., Tyerman, S. D., and Leigh, R. A. (2011). Cell-specific vacuolar calcium storage mediated by CAX1 regulates apoplastic calcium concentration, gas exchange, and plant productivity in *Arabidopsis*. *Plant Cell* 23, 240–257.

Cunningham, K. W. (2011). Acidic calcium stores of *Saccharomyces cerevisiae*. *Cell Calcium* 50, 129–138.

Cunningham, K. W., and Fink, G. R. (1994). Calcineurin-dependent growth-control in *Saccharomyces cerevisiae* mutants lacking Pmc1, a homolog of plasma-membrane Ca^{2+} ATPases. *J. Cell Biol.* 124, 351–363.

Denis, V., and Cyert, M. S. (2002). Internal Ca^{2+} release in yeast is triggered by hypertonic shock and mediated by a TRP channel homologue. *J. Cell Biol.* 156, 29–34.

Edmond, C., Shigaki, T., Ewert, S., Nelson, M., Connorton, J., Chalova, V., Noordally, Z., and Pittman, J. K. (2009). Comparative analysis of CAX2-like cation transporters indicates functional and regulatory diversity. *Biochem. J.* 418, 145–154.

Elbaz, B., Shoshani-Knaani, N., David-Assael, O., Mizrachy-Dagri, T., Mizrahi, K., Saul, H., Brook, E., Berezin, I., and Shaul, O. (2006). High expression in leaves of the zinc hyperaccumulator *Arabidopsis halleri* of AhMHX, a homolog of an *Arabidopsis thaliana* vacuolar metal/proton exchanger. *Plant Cell Environ.* 29, 1179–1190.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., and Bateman, A. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222.

Groeneveld, P., Stouthamer, A. H., and Westerhoff, H. V. (2009). Super life – how and why 'cell selection' leads to the fastest-growing eukaryote. *FEBS J.* 276, 254–270.

Haynes, W. J., Kung, C., Saimi, Y., and Preston, R. R. (2002). An exchanger-like protein underlies the large Mg^{2+} current in Paramecium. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15717–15722.

Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., Bluthgen, N., Borger, S., Costenoble, R., Heinemann, M., Hucka, M., Le Novere, N., Li, P., Liebermeister, W., Mo, M. L., Oliveira, A. P., Petranovic, D., Pettifer, S., Simeonidis, E., Smallbone, K., Spasic, I., Weichart, D., Brent, R., Broomhead, D. S., Westerhoff, H. V., Kirdar, B., Penttila, M., Klipp, E., Palsson, B. O., Sauer, U., Oliver, S. G., Mendes, P., Nielsen, J., and Kell, D. B. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* 26, 1155–1160.

Hirschi, K. (2001). Vacuolar H^+/Ca^{2+} transport: who's directing the traffic? *Trends Plant Sci.* 6, 100–104.

Hirschi, K. D. (2004). The calcium conundrum. Both versatile nutrient and specific signal. *Plant Physiol.* 136, 2438–2442.

Hirschi, K. D., Korenkov, V. D., Wilganowski, N. L., and Wagner, G. J. (2000). Expression of *Arabidopsis* CAX2 in tobacco. Altered metal accumulation and increased manganese tolerance. *Plant Physiol.* 124, 125–133.

Hirschi, K. D., Zhen, R. G., Cunningham, K. W., Rea, P. A., and Fink, G. R. (1996). CAX1, an H^+/Ca^{2+} antiporter from *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8782–8786.

Ivey, D. M., Guffanti, A. A., Zemsky, J., Pinner, E., Karpel, R., Padan, E., Schuldiner, S., and Krulwich, T. A. (1993). Cloning and characterization of a putative Ca^{2+}/H^+ antiporter gene from *Escherichia coli* upon functional complementation of Na^+/H^+ antiporter-deficient strains by the overexpressed gene. *J. Biol. Chem.* 268, 11296–11303.

Iwamoto, T., Uehara, A., Imanaga, I., and Shigekawa, M. (2000). The Na^+/Ca^{2+} exchanger NCX1 has oppositely oriented reentrant loop domains that contain conserved aspartic acids whose mutation alters its apparent Ca^{2+} affinity. *J. Biol. Chem.* 275, 38571–38580.

Jain, N., Nadgauda, R., and Shigaki, T. (2009). Mining cation (CAX) transporter diversity for nutrition-enhanced crops and phytoremediation. *Int. J. Integr. Biol.* 7, 22–25.

Kamiya, T., Akahori, T., and Maeshima, M. (2005). Expression profile of the genes for rice cation/H^+ exchanger family and functional analysis in yeast. *Plant Cell Physiol.* 46, 1735–1740.

Kamiya, T., and Maeshima, M. (2004). Residues in internal repeats of the rice cation/H^+ exchanger are involved in the transport and selection of cations. *J. Biol. Chem.* 279, 812–819.

Karimova, F. G., Kortchouganova, E. E., Tarchevsky, I. A., and Iagoucheva, M. R. (2000). The oppositely directed Ca^{2+} and Na^+ transmembrane transport in algal cells. *Protoplasma* 213, 93–98.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.

Koren'kov, V., Park, S., Cheng, N. H., Sreevidya, C., Lachmansingh, J., Morris, J., Hirschi, K., and Wagner, G. J. (2007). Enhanced Cd^{2+}-selective root-tonoplast-transport in tobaccos expressing *Arabidopsis* cation exchangers. *Planta* 225, 403–411.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.

Lytton, J. (2007). Na^+/Ca^{2+} exchangers: three mammalian gene families control Ca^{2+} transport. *Biochem. J.* 406, 365–382.

Manohar, M., Mei, H., Franklin, A. J., Sweet, E. M., Shigaki, T., Riley, B. B., MacDiarmid, C. W., and Hirschi, K. (2010). Zebrafish (*Danio rerio*) endomembrane antiporter similar to a yeast cation/H^+ transporter is required for neural crest development. *Biochemistry* 49, 6557–6566.

Manohar, M., Shigaki, T., and Hirschi, K. D. (2011). Plant cation/H^+ exchangers (CAXs): biological functions and genetic manipulations. *Plant Biol.* 13, 561–569.

Mäser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J. M., Sanders, D., Harper, J. F., Tchieu, J., Gribskov, M., Persans, M. W., Salt, D. E., Kim, S. A., and Guerinot, M. L. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* 126, 1646–1667.

McAinsh, M. R., and Pittman, J. K. (2009). Shaping the calcium signature. *New Phytol.* 181, 275–294.

Mei, H., Cheng, N. H., Zhao, J., Park, P., Escareno, R. A., Pittman, J. K., and Hirschi, K. D. (2009). Root development under metal stress in *Arabidopsis thaliana* requires the H^+/cation antiporter CAX4. *New Phytol.* 183, 95–105.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marechal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q. H., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Grigoriev, I. V., Rokhsar, D. S., Grossman, A. R., Annotation, C., and Team, J. G. I. A. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245–251.

Morris, J., Tian, H., Park, S., Sreevidya, C. S., Ward, J. M., and Hirschi, K. D. (2008). AtCCX3 is an *Arabidopsis* endomembrane H^+-dependent K^+ transporter. *Plant Physiol.* 148, 1474–1486.

Naseem, R., Holland, I. B., Jacq, A., Wann, K. T., and Campbell, A. K. (2008). pH and monovalent cations regulate cytosolic free Ca^{2+} in *E. coli*. *Biochim. Biophys. Acta* 1778, 1415–1422.

Nicoll, D. A., Hryshko, L. V., Matsuoka, S., Frank, J. S., and Philipson, K. D. (1996). Mutation of amino acid residues in the putative transmembrane segments of the cardiac sarcolemmal Na^+-Ca^{2+} exchanger. *J. Biol. Chem.* 271, 13385–13391.

Nicoll, D. A., Ren, X. Y., Ottolia, M., Phillips, M., Paredes, A. R., Abramson, J., and Philipson, K. D. (2007). What we know about the structure of NCX1 and how it relates to its function. *Ann. N. Y. Acad. Sci.* 1099, 1–6.

Ohyama, T., Igarashi, K., and Kobayashi, H. (1994). Physiological role of the chaA gene in sodium and calcium circulations at a high pH in *Escherichia coli*. *J. Bacteriol.* 176, 4311–4315.

On, C., Marshall, C. R., Chen, N., Moyes, C. D., and Tibbits, G. F. (2008). Gene structure evolution of the Na^+/Ca^{2+} exchanger (NCX) family. *BMC Evol. Biol.* 8, 127. doi:10.1186/1471-2148-8-127

Ottolia, M., John, S., Qiu, Z. Y., and Philipson, K. D. (2001). Split Na^+-Ca^{2+} exchangers – implications for function and expression. *J. Biol. Chem.* 276, 19603–19609.

Ottolia, M., Nicoll, D. A., and Philipson, K. D. (2005). Mutational analysis of the alpha-1 repeat of the cardiac Na^+-Ca^{2+} exchanger. *J. Biol. Chem.* 280, 1061–1069.

Palty, R., Hershfinkel, M., Yagev, O., Saar, D., Barkalifa, R., Khananshvili, D.,

Peretz, A., Grossman, Y., and Sekler, I. (2006). Single R-domain constructs of the Na+/Ca2+ exchanger, NCLX, oligomerize to form a functional exchanger. *Biochemistry* 45, 11856–11866.

Palty, R., Ohana, E., Hershfinkel, M., Volokita, M., Elgazar, V., Beharier, O., Silverman, W. F., Argaman, M., and Sekler, I. (2004). Lithium-calcium exchange is mediated by a distinct potassium-independent sodium-calcium exchanger. *J. Biol. Chem.* 279, 25234–25240.

Palty, R., Silverman, W. F., Hershfinkel, M., Caporale, T., Sensi, S. L., Parnis, J., Nolte, C., Fishman, D., Shoshan-Barmatz, V., Herrmann, S., Khananshvili, D., and Sekler, I. (2010). NCLX is an essential component of mitochondrial Na^+/Ca^{2+} exchange. *Proc. Natl. Acad. Sci. U.S.A.* 107, 436–441.

Philipson, K. D., and Nicoll, D. A. (2000). Sodium-calcium exchange: a molecular perspective. *Annu. Rev. Physiol.* 62, 111–133.

Pittman, J. K. (2011). Vacuolar Ca^{2+} uptake. *Cell Calcium* 50, 139–146.

Pittman, J. K., Edmond, C., Sunderland, P. A., and Bray, C. M. (2009). A cation-regulated and proton gradient-dependent cation transporter from *Chlamydomonas reinhardtii* has a role in calcium and sodium homeostasis. *J. Biol. Chem.* 284, 525–533.

Pittman, J. K., Shigaki, T., Cheng, N. H., and Hirschi, K. D. (2002). Mechanism of N-terminal autoinhibition in the Arabidopsis Ca^{2+}/H^+ antiporter CAX1. *J. Biol. Chem.* 277, 26452–26459.

Radchenko, M. V., Tanaka, K., Waditee, R., Oshimi, S., Matsuzaki, Y., Fukuhara, M., Kobayashi, H., Takabe, T., and Nakamura, T. (2006). Potassium/proton antiport system of *Escherichia coli*. *J. Biol. Chem.* 281, 19822–19829.

Reilander, H., Achilles, A., Friedel, U., Maul, G., Lottspeich, F., and Cook, N. J. (1992). Primary structure and functional expression of the Na/Ca,K-exchanger from bovine rod photoreceptors. *EMBO J.* 11, 1689–1695.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van de Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L. X., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrano, R. S., and Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69.

Saier, M. H., Eng, B. H., Fard, S., Garg, J., Haggerty, D. A., Hutchinson, W. J., Jack, D. L., Lai, E. C., Liu, H. J., Nusinew, D. P., Omar, A. M., Pao, S. S., Paulsen, I. T., Quan, J. A., Sliwinski, M., Tseng, T. T., Wachi, S., and Young, G. B. (1999). Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochim. Biophys. Acta* 1422, 1–56.

Schnetkamp, P. P. M. (1995). Calcium homeostasis in vertebrate retinal rod outer segments. *Cell Calcium* 18, 322–330.

Schwarz, E. M., and Benzer, S. (1997). Calx, a Na-Ca exchanger gene of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10249–10254.

Shaul, O., Hilgemann, D. W., de-Almeida-Engler, J., Van Montagu, M., Inze, D., and Galili, G. (1999). Cloning and characterization of a novel Mg^{2+}/H^+ exchanger. *EMBO J.* 18, 3973–3980.

Shigaki, T., Barkla, B. J., Miranda-Vergara, M. C., Zhao, J., Pantoja, O., and Hirschi, K. D. (2005). Identification of a crucial histidine involved in metal transport activity in the *Arabidopsis* cation/H^+ exchanger CAX1. *J. Biol. Chem.* 280, 30136–30142.

Shigaki, T., Mei, H., Marshall, J., Li, X., Manohar, M., and Hirschi, K. D. (2010). The expression of the open reading frame of *Arabidopsis* CAX1, but not its cDNA, confers metal tolerance in yeast. *Plant Biol.* 12, 935–939.

Shigaki, T., Pittman, J. K., and Hirschi, K. D. (2003). Manganese specificity determinants in the *Arabidopsis* metal/H^+ antiporter CAX2. *J. Biol. Chem.* 278, 6610–6617.

Shigaki, T., Rees, I., Nakhleh, L., and Hirschi, K. D. (2006). Identification of three distinct phylogenetic groups of CAX cation/proton antiporters. *J. Mol. Evol.* 63, 815–825.

Shigekawa, M., Iwamoto, T., Uehara, A., and Kita, S. (2002). Probing ion binding sites in the Na^+/Ca^{2+} exchanger. *Ann. N. Y. Acad. Sci.* 976, 19–30.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.* 57, 758–771.

Szerencsei, R. T., Tucker, J. E., Cooper, C. B., Winkfein, R. J., Farrell, P. J., Iatrou, K., and Schnetkamp, P. P. M. (2000). Minimal domain requirement for cation transport by the potassium-dependent Na/Ca-K exchanger – comparison with an NCKX paralog from *Caenorhabditis elegans*. *J. Biol. Chem.* 275, 669–676.

Verret, F., Wheeler, G., Taylor, A. R., Farnham, G., and Brownlee, C. (2010). Calcium channels in photosynthetic eukaryotes: implications for evolution of calcium-based signalling. *New Phytol.* 187, 23–43.

Wheeler, G. L., and Brownlee, C. (2008). Ca^{2+} signalling in plants and green algae – changing channels. *Trends Plant Sci.* 13, 506–514.

Zhao, J., Connorton, J. M., Guo, Y. Q., Li, X. K., Shigaki, T., Hirschi, K. D., and Pittman, J. K. (2009). Functional studies of split *Arabidopsis* Ca^{2+}/H^+ exchangers. *J. Biol. Chem.* 284, 34075–34083.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2011; accepted: 01 January 2012; published online: 13 January 2012.

*Citation: Emery L, Whelan S, Hirschi KD and Pittman JK (2012) Protein phylogenetic analysis of Ca^{2+}/cation antiporters and insights into their evolution in plants. Front. Plant Sci. **3**:1. doi: 10.3389/fpls.2012.00001*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 19 January 2012
doi: 10.3389/fpls.2011.00119

Evolutionary relationships and functional diversity of plant sulfate transporters

Hideki Takahashi[1], Peter Buchner[2], Naoko Yoshimoto[3], Malcolm J. Hawkesford[2] and Shin-Han Shiu[4]**

[1] Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, USA
[2] Plant Science Department, Rothamsted Research, Harpenden, UK
[3] Graduate School of Pharmaceutical Sciences, Chiba University, Chiba, Japan
[4] Department of Plant Biology, Michigan State University, East Lansing, MI, USA

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Li-Qing Chen, Carnegie Institution for Science, USA
Joseph M. Jez, Washington University in St. Louis, USA

***Correspondence:**
Hideki Takahashi, Department of Biochemistry and Molecular Biology, Michigan State University, 209 Biochemistry Building, East Lansing, MI 48824, USA.
e-mail: htakaha@msu.edu;
Shin-Han Shiu, Department of Plant Biology, Michigan State University, S308 Plant Biology Building, East Lansing, MI 48824, USA.
e-mail: shius@msu.edu

Sulfate is an essential nutrient cycled in nature. Ion transporters that specifically facilitate the transport of sulfate across the membranes are found ubiquitously in living organisms. The phylogenetic analysis of known sulfate transporters and their homologous proteins from eukaryotic organisms indicate two evolutionarily distinct groups of sulfate transport systems. One major group named Tribe 1 represents yeast and fungal SUL, plant SULTR, and animal SLC26 families. The evolutionary origin of SULTR family members in land plants and green algae is suggested to be common with yeast and fungal SUL and animal anion exchangers (SLC26). The lineage of plant SULTR family is expanded into four subfamilies (SULTR1–SULTR4) in land plant species. By contrast, the putative SULTR homologs from Chlorophyte green algae are in two separate lineages; one with the subfamily of plant tonoplast-localized sulfate transporters (SULTR4), and the other diverged before the appearance of lineages for SUL, SULTR, and SLC26. There also was a group of yet undefined members of putative sulfate transporters in yeast and fungi divergent from these major lineages in Tribe 1. The other distinct group is Tribe 2, primarily composed of animal sodium-dependent sulfate/carboxylate transporters (SLC13) and plant tonoplast-localized dicarboxylate transporters (TDT). The putative sulfur-sensing protein (SAC1) and SAC1-like transporters (SLT) of Chlorophyte green algae, bryophyte, and lycophyte show low degrees of sequence similarities with SLC13 and TDT. However, the phylogenetic relationship between SAC1/SLT and the other two families, SLC13 and TDT in Tribe 2, is not clearly supported. In addition, the SAC1/SLT family is absent in the angiosperm species analyzed. The present study suggests distinct evolutionary trajectories of sulfate transport systems for land plants and green algae.

Keywords: evolution, plant, sulfate, transporter

INTRODUCTION

Sulfate is an essential nutrient and the initial substrate for biosynthesis of sulfur-containing metabolites in plants, algae, and microorganisms (Leustek et al., 2000; Saito, 2004; Takahashi et al., 2011). The organic sulfur metabolites synthesized in these autotrophic organisms are the sulfur nutritional resource for animals. However, animals are not devoid of sulfate transport proteins as they play significant roles in reabsorbing sulfate in renal systems to maintain ion homeostasis (Markovich and Murer, 2004; Mount and Romero, 2004). A sulfate transporter is also known to be essential for cartilage formation as it may contribute to supplying sulfate for synthesis of sulfated proteoglycans (Hästbacka et al., 1994). Sulfate transport proteins are found therefore across diverse organisms, although they may facilitate transport of sulfate for different purposes. Apart from the ubiquitous presence among organisms, the expansion of the family members in multicellular organisms is most likely an evolutionary development to provide sulfate transporters which are specifically functional in different organs or tissues. The expansion led to distinct spatial distribution and organization of biochemically diversified forms of sulfate transporters, which is necessary for coordinating the overall transport of sulfate within complex biological systems. In addition, the ionic environmental factors are highly variable that may have contributed to develop distinct types of transport systems and regulatory mechanisms.

Previous studies have indicated that sulfate transport proteins can be classified to four different types according to the mechanisms mediating transport of sulfate across the membranes (**Figure 1**). The influx of sulfate can be coupled with co-transport of positively charged counter ions such as proton (H^+) and sodium (Na^+). For these mechanisms, the concentration gradients of counter ions serve as driving force for the influx of sulfate across the membranes. The proton gradient is suggested to be the driving force for sulfate uptake systems in yeast and plants (Roomans et al., 1979; Lass and Ullrich-Eberius, 1984; Hawkesford et al., 1993). SUL1 and SUL2 in yeast (Smith et al., 1995a; Cherest et al., 1997) and SULTR family members in plants (Smith et al., 1995b, 1997; Buchner et al., 2004; Takahashi, 2010) are the suggested components of the proton/sulfate co-transport systems. In contrast, animals may use different mechanisms. The sulfate transport activities of SLC13 family proteins are known to be dependent on sodium (Markovich and Murer, 2004; Pajor, 2006). An

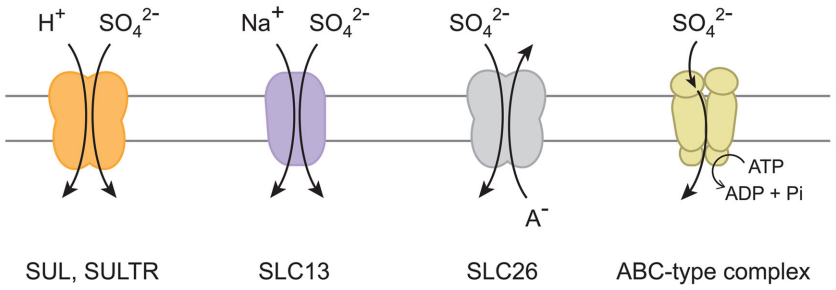


FIGURE 1 | Sulfate transport mechanisms. Proton/sulfate co-transporter, sodium/sulfate co-transporter, sulfate/anion (A^-) exchanger, and ABC-type sulfate transporter complex are illustrated. The names of sulfate transport proteins are indicated below the suggested mechanisms.

alternative mechanism is the anion exchange systems facilitating the counter transport of sulfate and other negatively charged ions, such as chloride (Cl^-), iodide (I^-), and bicarbonate (HCO_3^-). The SLC26 family proteins facilitate sulfate/anion exchanges in animals (Mount and Romero, 2004). Transport of sulfate can be also driven by an ATP-binding cassette (ABC) transporter complex in bacteria and algal chloroplasts (Sirko et al., 1990; Laudenbach and Grossman, 1991; Lindberg and Melis, 2008). These mechanisms are suggested to have evolved in various ancestral species depending on the ionic environments where those transporters were to be operated.

The various sulfate transport systems are intimately linked with the subsequent metabolism of their transported molecule, sulfate. Once sulfate is delivered to the cell, it serves as a substrate for the sulfur assimilatory enzyme, ATP sulfurylase, both in the cytoplasm and the plastids in plants (Takahashi et al., 2011). The metabolic flux of ATP sulfurylase and subsequent reduction steps in plastids defines the primary requirement of sulfate in metabolism (Vauclare et al., 2002; Kopriva, 2006). Before entering the steps of metabolic conversion, sulfate in the cytoplasm can be sequestered to vacuoles (Buchner et al., 2004; Takahashi, 2010). The export of sulfate to the extra-cellular space would be another factor affecting the rate of sulfate uptake across the plasma membrane. Unknown passive transport systems are suggested for those mechanisms as the membrane potentials are positive at extra-cellular and vacuolar lumen sides (Buchner et al., 2004; Takahashi, 2010). In addition, a steep upward concentration gradient of sulfate may be generated across the plasma membrane under sulfate-starved conditions; active transport systems are necessary to drive the influx of sulfate efficiently under such circumstances. The systems should contain selective mechanisms either coupled with transport of counter ions, or energized by ATP, as catalyzed by ABC transporters (**Figure 1**).

The recent expansion of genome sequencing information has enabled the identification of a number of sulfate transporters and homologous proteins from higher plants (Takahashi, 2010). This study focuses on the molecular evolution of the families of sulfate transporters in the green lineage (i.e., land plants and green algae). Phylogenetic analysis was conducted using a diverse set of relevant protein sequences from yeast, fungi, algae, bryophyte, lycophyte, seed plants, and animals to reinterpret their biochemical diversification with respect to the evolution of eukaryotic organisms and to assess the lineage-specific expansion of the family members. The present study aims to provide information of family classifications of sulfate transporters and related proteins based on their phylogenetic relationships and evolution.

FAMILY CLASSIFICATIONS

The protein sequences of sulfate transporters were identified from the following organisms: *Saccharomyces cerevisiae*, *Aspergillus niger*, *Aspergillus nidulans*, *Chlamydomonas reinhardtii*, *Volvox carteri*, *Physcomitrella patens*, *Selaginella moellendorffii*, *Arabidopsis thaliana*, *Glycine max*, *Populus trichocarpa*, *Oryza sativa*, *Brachypodium distachyon*, *Sorghum bicolor*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Danio rerio*, and *Homo sapiens*. Metazoan, plant, budding yeast, and *Aspergillus* protein sequences were obtained from Ensembl (release 62[1]), Phytozome (ver. 5[2]), SGD (Feb 0, 2011[3]), and BROAD[4], respectively. To identify sulfate transporter, a three step analysis pipeline was used. First, annotated protein sequences from the 17 representative species were searched against known sulfate transporter from plants and human with a low Expect value threshold of one to include as many candidates as possible. Second, presence of transmembrane regions in these candidate sequences was identified with TMHMM (Krogh et al., 2001). Two types of candidates were analyzed further: (1) the sequence has an Expect value $>$ 1e–5 but has $\geq$7 transmembrane regions and (2) the sequence has an Expect value $\leq$ 1e–5 and has $\geq$1 transmembrane regions. The first criterion is to ensure that divergent transporters are captured. The second is to include partial sequences of true sulfate transporters given many genomes analyzed were not heavily annotated.

In the final step, candidates passing the Expect value and transmembrane region criteria were aligned with annotated sulfate transporters for phylogenetic reconstruction. Due to the sheer number of sequences analyzed, the phylogenetic analysis was done in three iterations starting with computationally straightforward neighbor-joining algorithm with bootstrap as implemented in MEGA (Tamura et al., 2011). After dividing candidates into "tribes" based on neighbor-joining trees, bootstrapped maximum likelihood (ML) trees were generated with RAxML (Stamatakis, 2006). This program has advantages for computation of large

[1]http://www.ensembl.org
[2]http://www.phytozome.net/
[3]http://yeastgenome.org/
[4]http://www.broadinstitute.org

phylogenetic trees as in this study. We chose these methods considering accuracy and computational performance. Based on ML tree topology, candidates were subdivided into families and sequences were excluded if they do not resides in the same well supported (>50%) clades with known sulfate transporters. Finally, bootstrapped ML trees were generated for each family.

The members of chloroplast-localized sulfate transporter from *Chlamydomonas* (Melis and Chen, 2005; Lindberg and Melis, 2008) correspond to bacterial ABC-type sulfate transporter complex, which is composed of a sulfate binding protein, periplasmic membrane-bound proteins, and an ABC protein that hydrolyzes ATP and provides the energy for transport (Sirko et al., 1990; Laudenbach and Grossman, 1991; **Figure 1**). These components are present only in bacteria and algae but not in any land plant species analyzed (Takahashi, 2010; Takahashi et al., 2011). Horizontal gene transfer of ABC-type sulfate transporter complex may have occurred from bacteria to algae, but it is possible that the complex has been lost in land plants when their ancestor diverged from green algae approximately one billion years ago. Based on phylogenetic analysis, there is no evidence suggesting that the proteins for ABC-type sulfate transporter complex are homologous to other groups of sulfate transporters focused in this study. Although both the eukaryotic sulfate transporters and the ABC-type complexes may transport sulfate, they are structurally and mechanistically different (**Figure 1**). Plant SULTR, metazoan SLC26, and yeast SUL proteins are predicted to contain 10–14 hydrophobic transmembrane regions (Smith et al., 1995a,b; Cherest et al., 1997; Hawkesford, 2003; Mount and Romero, 2004). SLC13 is predicted to have 8–13 transmembrane regions (Markovich and Murer, 2004; Pajor, 2006). By contrast, the bacterial/algal ABC-type complexes are composed of multiple subunit proteins sharing individual roles in facilitating transport of sulfate across the membranes (Sirko et al., 1990; Laudenbach and Grossman, 1991; Melis and Chen, 2005; Lindberg and Melis, 2008). It is apparent that ABC-type sulfate transporter complex radiated in prokaryotes and algae, and their evolutionary trajectories were completely different from those for the eukaryotic-type sulfate transporters focused in this article.

Two major groups of eukaryotic sulfate transporters were identified and designated Tribe 1 and Tribe 2 (**Figures 2–4**). Tribe 1 is composed of three major lineages, *Family P*, *Family A1*, and *Family A2*, respectively (**Figure 2**). The plant SULTR family members are found exclusively in *Family P* (**Figures 2** and **3**). The animal SLC26 family members are found in *Family A1* and *A2*. The yeast SUL1/SUL2 and their fungal homologs are found in *Family A1*, although a few additional homologs including yeast YPR003C and YGR125W exist in groups that diverged earlier than the emergence of plant and animal lineages. The members of algal SULTR also split into two groups; one present in *Family P* and the other in clades regarded as out-groups. In Tribe 1, the evolutionary origin of plant SULTR family may be tracked back to the fungal–animal–plant common ancestor based on the relationships of plant SULTR (Takahashi, 2010) to yeast and fungal SUL (Smith et al., 1995a; Cherest et al., 1997) and to animal sulfate/anion exchangers (SLC26; Mount and Romero, 2004; **Figure 2**). The *Family P* lineage appears to be associated with the *Family A1* lineage. Although there still remains ambiguity regarding the exact origin of *Family P*, the results may well suggest that the ancestral forms of the existing yeast SUL1/SUL2 were the founders of the three major lineages, *Family P*, *A1*, and *A2* in Tribe 1 (**Figure 2**). In contrast, the two other yeast SUL homologs, YPR003C and YGR125W, are suggested to have diverged prior to the emergence of those major lineages (**Figure 2**). Tribe 2 is also composed of three distinct lineages (**Figure 4**). These lineages represent the families of animal sodium-dependent sulfate/carboxylate transporters (SLC13; Markovich and Murer, 2004; Pajor, 2006), plant tonoplast-localized dicarboxylate transporters (TDT; Emmerlich et al., 2003), and algal putative sulfur-sensing proteins (SAC1; Davies et al., 1996), and SAC1-like transporters (SLT; Pootakham et al., 2010), respectively. The phylogenetic relationships of the family and subfamily members in these two tribes will be discussed in the following sections.

PLANT SULFATE TRANSPORTERS (SULTR)

The plant specific lineage of Tribe 1 is composed of four distinct subfamilies of SULTR-type sulfate transporters, SULTR1, SULTR2, SULTR3, and tonoplast-localized sulfate transporters (SULTR4; **Figure 2**, *Family P*; **Figure 3**). These four subfamilies correspond with the nomenclatures of sulfate transporters identified from *Arabidopsis* and other vascular plant species (Hawkesford, 2003; Buchner et al., 2004; Takahashi, 2010). *Family P* was fairly specific to land plant species, although part of algal SULTR members (Pootakham et al., 2010) was present in an out-group diverged from SULTR4 subfamily (**Figure 2**, *node 2*). The first gene duplication event in *Family P* was the split of the land plant SULTR1/2/3 and SULTR4 forms (**Figure 2**, *node 1*). Two distinct types of SULTR were subsequently generated. The lineage of SULTR4 (Kataoka et al., 2004a) was diverged by subsequent gene duplication (**Figure 2**, *node 2*). The vacuole localization of the ancestral form of SULTR4 was probably defined after this gene duplication event, because a plasma membrane localized sulfate transporter from *Chlamydomonas* (SULTR2; Pootakham et al., 2010) is found in the clade diverged from this lineage. The branching of SULTR4 subfamily members in bryophyte and lycophyte is somewhat irregular as *Selaginella* SULTR4 appears to have diverged earlier than *Physcomitrella* SULTR4. This may be due to the ambiguity of branch position of *Physcomitrella* SULTR4 with relatively low bootstrap support.

SULTR1/2/3 subsequently split into SULTR1/2 and SULTR3 clades (**Figure 2**, *node 3*; **Figure 3**). SULTR1 and SULTR2 in *Arabidopsis* are functional sulfate transporters that can restore the sulfate uptake activity of the yeast *sul1 sul2* mutant (Takahashi et al., 1997, 2000; Shibagaki et al., 2002; Yoshimoto et al., 2002, 2007). The phylogenetic relationships between SULTR1 and SULTR2 support their functional similarities as being sulfate transporters. Within the clade of SULTR1/2, there is a sister group of sulfate transporters for *Physcomitrella* and *Selaginella*. The ancestor of this group likely emerged prior to the division of the SULTR1 and SULTR2 subfamilies. Their phylogenetic relationships with SULTR1/2 lineage may well suggest that these bryophyte and lycophyte SULTR homologs would have sulfate transport activities, although their functional identities are not confirmed. With respect to the substrate specificities of the SULTR1 and SULTR2 subfamilies, the results suggest that the differences of their kinetic properties are evolutionarily derived. Consistent with previous findings, SULTR1 and SULTR2 can be defined as subfamilies of

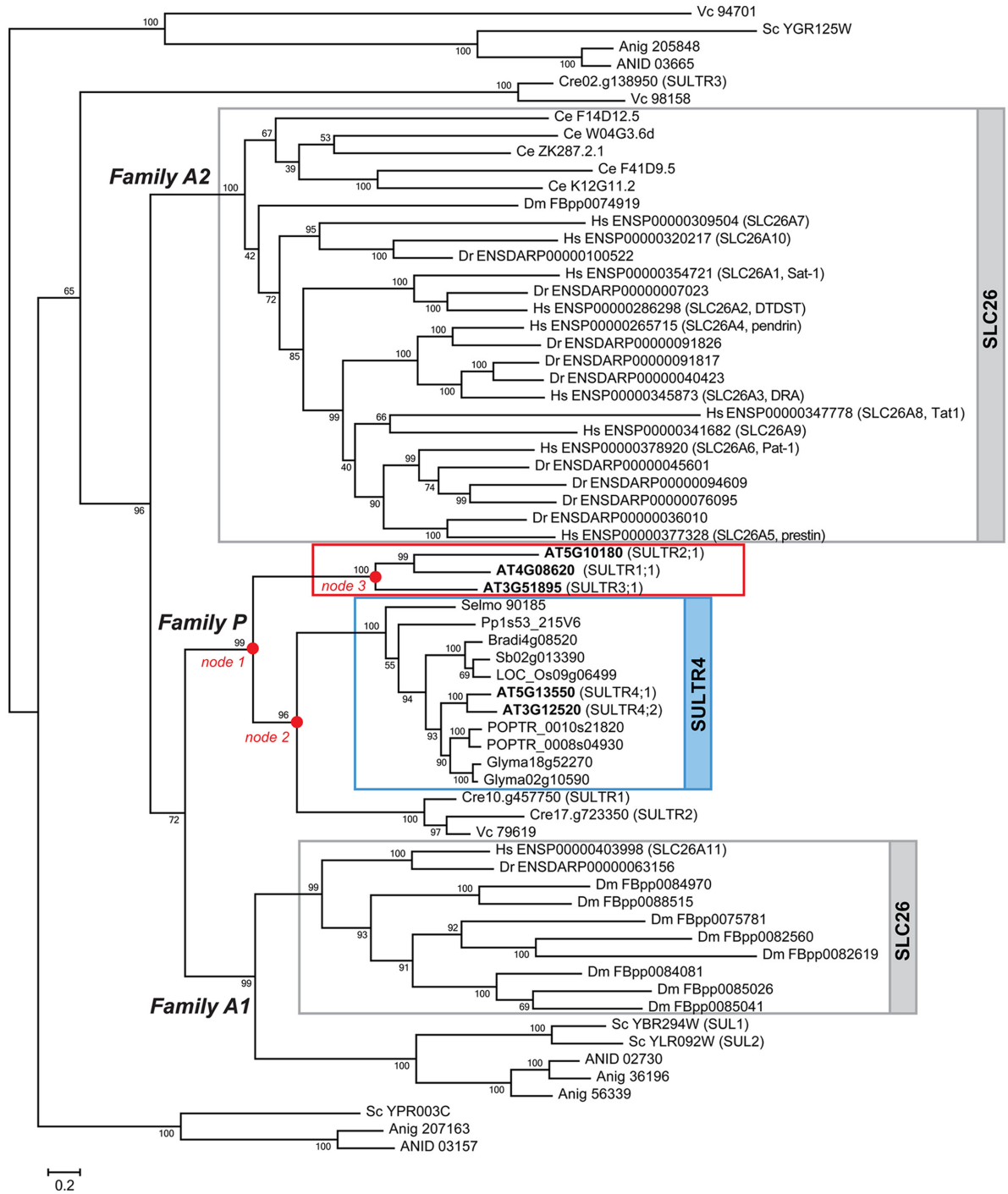


FIGURE 2 | Phylogenetic relationships of SUL, SULTR, and SLC26 in Tribe 1. The lineage that splits to SULTR1, SULTR2, and SULTR3 subfamilies (**Figure 3**) are boxed in red. SULTR1;1, SULTR2;1, and SULTR3;1 from *Arabidopsis* were selected as representatives of these subfamilies to construct the phylogenetic tree. The SULTR4 subfamily in plants and SLC26 in animals are highlighted in blue and gray, respectively. The nodes 1–3 where the plant SULTR lineage splits into subfamilies are indicated by red dots. The bootstrap values are shown at the nodes. The locus numbers or gene IDs are indicated according to Ensembl (http://www.ensembl.org), Phytozome (http://www.phytozome.net/), SGD (http://yeastgenome.org/), JGI (http://www.jgi.doe.gov/), and BROAD (http://www.broadinstitute.org). Prefix abbreviations of locus numbers or gene IDs indicate genus and species names: AT, *Arabidopsis thaliana*; ANID, *Aspergillus nidulans*; Anig, *Aspergillus niger*; Bradi, *Brachypodium distachyon*; Ce, *Caenorhabditis elegans*; Cre, *Chlamydomonas reinhardtti*; Dr, *Danio rerio*; Dm, *Drosophila melanogaster*; Glyma, *Glycine max*; Hs, *Homo sapiens*; Os, *Oryza sativa*; Pp, *Physcomitrella patens*; POPTR, *Populus trichocarpa*; Selmo, *Selaginella moellendorffii*; Sc, *Saccharomyces cerevisiae*; Sb, *Sorghum bicolor*; Vc, *Volvox carteri*. The locus numbers of *Arabidopsis* SULTR are highlighted in bold letters. The names of *Arabidopsis* SULTR, *Chlamydomonas* SULTR, human SLC26A, and yeast SUL family members are shown in the parentheses next to the locus numbers.

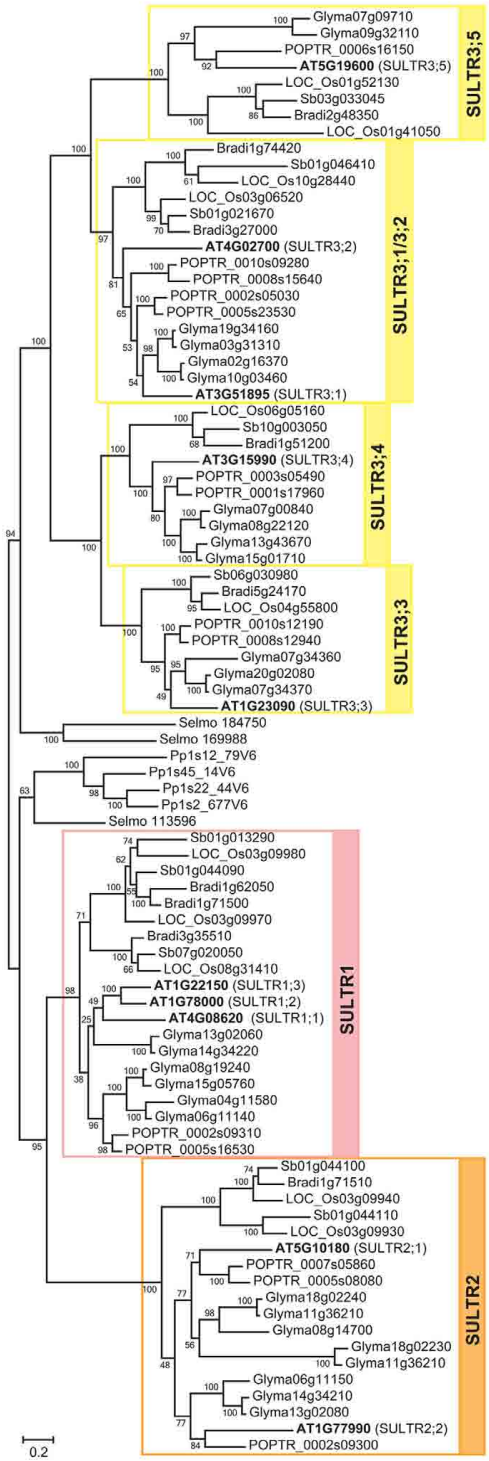


FIGURE 3 | Expansion of SULTR1–SULTR3 subfamilies in plants. The phylogenetic tree indicates expansion of the SULTR1/2/3 lineage (**Figure 2**) in seed plants. The SULTR1, SULTR2, and SULTR3 subfamilies are highlighted in pink, orange, and yellow, respectively. The bootstrap values are shown at the nodes. The locus numbers or gene IDs are described the same as in **Figure 2**. The names of *Arabidopsis* sulfate transporters are shown in the parentheses next to the locus numbers.

high- and low-affinity sulfate transporters, respectively (Takahashi et al., 2000; Yoshimoto et al., 2002). Within each subfamily, SULTR1 first splits to dicotyledonous and monocotyledonous groups and subsequently duplicates to have the subfamily members in different flowering plant lineages. SULTR2 also splits to dicotyledonous and monocotyledonous groups, followed by specialization of SULTR2;1 and SULTR2;2. Supported by the experimental evidence for the kinetic properties of *Arabidopsis* SULTR2;1 and SULTR2;2 (Takahashi et al., 2000), the divergence of these two forms at least in dicots may indicate the difference in their affinities to sulfate.

SULTR3 is principally composed of subfamily members from angiosperms. Before the expansion in the angiosperms, the ancestral lineage of SULTR3 gave rise to *Selaginella* SULTR (**Figure 3**). Since there is no *Physcomitrella* SULTR in this subfamily, the lineage of SULTR3 appears to be specific to vascular plant species. The SULTR3 subfamily subsequently divided to four classes. They were designated SULTR3;5, SULTR3;1/3;2, SULTR3;3, and SULTR3;4 according to the names of the subfamily members from *Arabidopsis* (Takahashi, 2010). As with the SULTR1 subfamily, SULTR3;3, SULTR3;4, and SULTR3;5 first split to dicotyledonous and monocotyledonous groups and then diverge to have the subfamily members in individual plant species. With respect to the expansion of the SULTR3;1/3;2 subfamily, SULTR3;1 and SULTR3;2 are founded after the division of dicotyledonous and monocotyledonous plants as described for the evolution of SULTR2;1/2;2. The SULTR3 family members in *Arabidopsis* are suggested to be involved in internal transport of sulfate in vasculature and developing seeds (Kataoka et al., 2004b; Zuber et al., 2010). In addition, a SULTR3;5 homolog in *Lotus japonicus* appears to mediate intracellular transport of sulfate to symbiosomes (Krusell et al., 2005). In spite of these indications from the physiological characterizations of plant mutant lines, the exact biochemical features of SULTR3 subfamily members are yet unverified. At this point, their sulfate uptake activities are suggested to be very low or barely detectable (Kataoka et al., 2004b) except for the case in *L. japonicus* SST1 (Krusell et al., 2005). In contrast to the SULTR1/2 subfamilies, the lack of biochemical information hampers us to interpret the evolutionary diversification of the individual classes of SULTR3 subfamily based on their molecular functions.

Some earlier studies have annotated an additional family of putative sulfate transporter (group 5) based on its partial sequence similarities with plant SULTR (Hawkesford, 2003; Buchner et al., 2004). However, the transmembrane structured proteins in group 5 contain no sulfate transporter motif (Leves et al., 2008) and STAS domain (Aravind and Koonin, 2000) which is the typical signature for SUL, SULTR, and SLC26 proteins. Later identification of its role as a molybdate transporter (MOT) may explain its considerable divergence from sulfate transporters (Tejada-Jiménez et al., 2007; Tomatsu et al., 2007; Baxter et al., 2008; Gasber et al., 2011). Consistent with functional studies, we did not find clear support for an evolutionary relationship between MOT and families of sulfate transporters in both prokaryotes and eukaryotes.

SUL AND SLC26

The yeast and fungal SUL and animal SLC26 sulfate/anion exchangers are present in two distinct lineages of Tribe 1 (**Figure 2**,

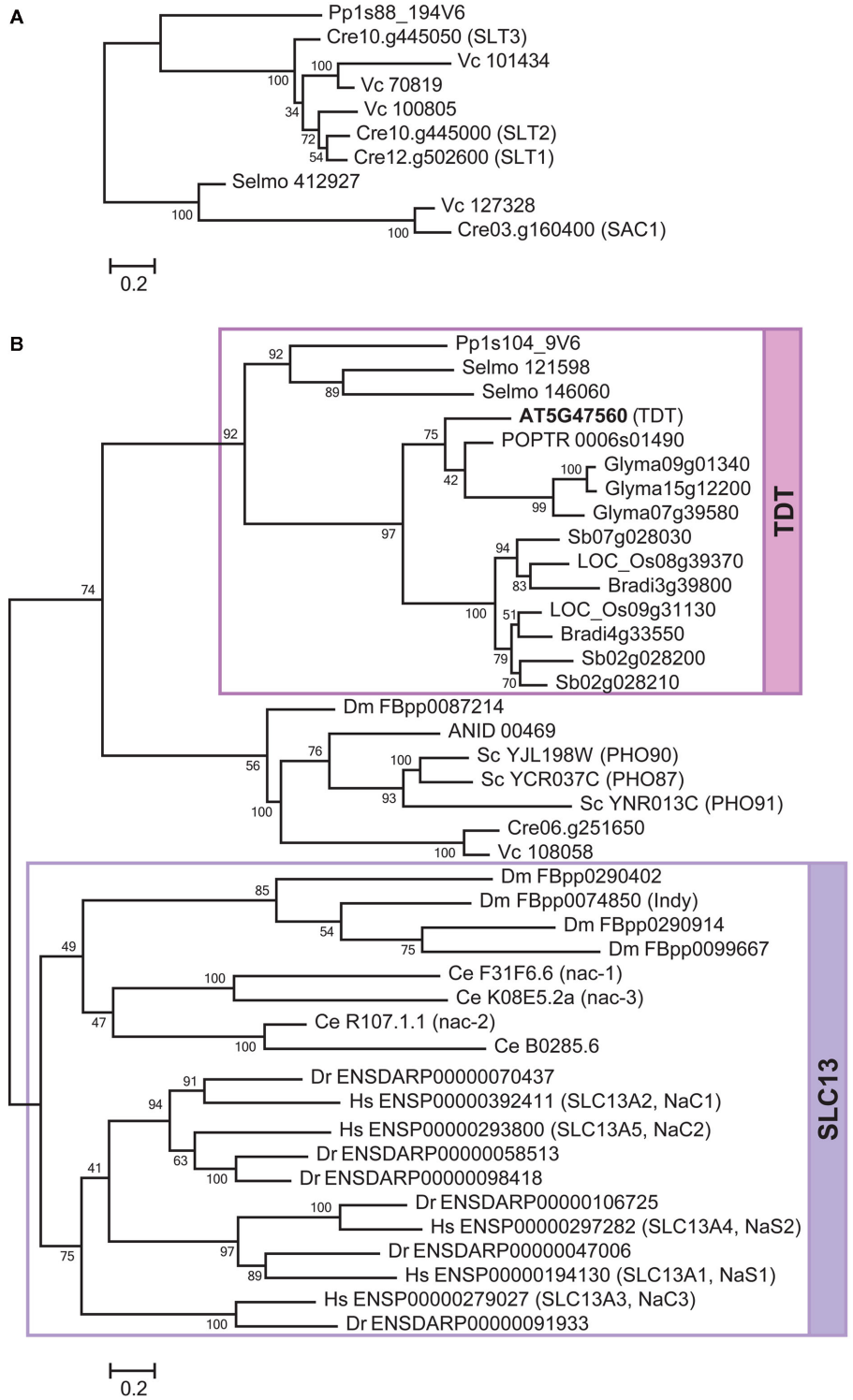


FIGURE 4 | Phylogenetic relationships of SAC1/SLT, SLT13, and TDT in Tribe 2. The phylogenetic trees of SAC1/SLT **(A)**, and SLC 13 and TDT **(B)**, were constructed separately. The lineages for SLC13 and TDT are highlighted in violet and magenta, respectively. The bootstrap values are shown at the nodes. The locus numbers or gene IDs are described the same as in **Figure 2**. The locus number of *Arabidopsis* TDT is highlighted in bold letters. The names of *Chlamydomonas* SAC1 and SLT, *Arabidopsis* TDT, and animal SLC13 family members are shown in the parentheses next to the locus numbers.

Family A1, and *Family A2*). *Family A1* contains SUL1/SUL2 sulfate transporters from yeast (Smith et al., 1995a; Cherest et al., 1997) and their closest homologs in *Aspergillus*. An animal-specific lineage is present in this family. It was composed of SLC26A11 from human and zebra fish, and eight homologous proteins from *Drosophila*. It is likely that the majority of the existing SLC26 members in *Drosophila* occurred through gene duplications in this animal-specific lineage in *Family A1*. Only one SLC26 member from *Drosophila* is found in the other animal clade (*Family A2*). By contrast, SLC26A11 is the only member from human and zebra fish present in *Family A1*, while the rest of the members (SLC26A1–A10) are found in *Family A2*. *Family A2* is a lineage of SLC26 from *C. elegans*, human, and zebra fish (**Figure 2**). This animal-specific lineage eventually expands to human and zebra fish SLC26A1–A10 members. The ancestor of this group appears to have first duplicated to form a subfamily of *C. elegans* SLC26, and subsequently expanded to the vertebrate SLC26A1–A10 members (Mount and Romero, 2004).

The high bootstrap value provides supports that the existing members of sulfate transporters in *Family P*, *A1*, and *A2* are originated from the same ancestral form. The results from phylogenetic analysis further suggested additional members of putative sulfate transporters which could have diverged prior to the emergence of these three major lineages (**Figure 2**). The SULTR homologs from *Chlamydomonas* (SULTR3; Pootakham et al., 2010) and *Volvox* (Vc 98158) in the out-group may derive from earlier evolutionary events. Yeast YPR003C and YGR125W and their homologs in *Aspergillus* form additional clades, suggesting their ancestral forms may have first diverged from the major lineages of existing sulfate transporters. The yeast YPR003C and YGR125W are putative sulfate transporters yet to be characterized. They may contribute to the residual sulfate transport activities in yeast *sul1* and *sul1 sul2* mutants (Smith et al., 1995a; Cherest et al., 1997), although the biochemical function of these putative sulfate transporters awaits further investigation. It is notable that a SULTR homolog from *Volvox* is present in the clade of YGR125W (**Figure 2**; Vc 94701). The phylogenetic relationship between algal SULTR and putative sulfate transporters from yeast and fungi suggests that they have shared a common ancestor prior to the divergence of the green algal and fungal lineages.

PUTATIVE SULFUR-SENSING PROTEIN (SAC1) AND SAC1-LIKE TRANSPORTERS (SLT)

Among the three lineages in Tribe 2, the SAC1/SLT is relatively independent of the other two lineages, SLC13 and TDT (**Figure 4**). Considering the overall similarities of protein sequences among these family members, SAC1/SLT (Davies et al., 1996; Pootakham et al., 2010), SLC13 (Markovich and Murer, 2004; Pajor, 2006), and TDT (Emmerlich et al., 2003) may have originated from a common ancestor. However, the phylogenetic relationships of SAC1/SLT with SLC13 and TDT were not clearly supported according to the ML tree (**Figure 4**). SAC1/SLT is a family specific to Chlorophyte algae *Chlamydomonas* and *Volvox*, a bryophyte *Physcomitrella* and a lycophyte *Selaginella* (**Figure 4A**). No seed plant homologs have been identified. Within this family, a SAC1 homolog is identified from *Selaginella*, suggesting that SAC1/SLT is present at least in the ancestor of vascular plants but has been lost when seed plant ancestors diverged from non-seed plants approximately 400 million years ago.

The phylogenetic tree provides further support that SLT1–SLT3 (Pootakham et al., 2010) are well conserved in Chlorophyte green algae as they are found in *Chlamydomonas* and *Volvox* (**Figure 4A**). The closest family member from *Physcomitrella* is suggested to have diverged from this algal SLT group, although the phylogenetic relationship is not strongly supported. The branch organization apparently suggests that *Physcomitrella* and *Selaginella* may have lost SAC1 and SLT, respectively. Since the position of a *Physcomitrella* homolog is not supported (**Figure 4A**), it may be also associated with the SAC1 clade. Accordingly, the clade of SLT1–SLT3 will become distinguishable as an algal specific lineage. Although the biological functions of SAC1/SLT from *Physcomitrella* and *Selaginella* are yet to be verified, this alternative interpretation may simply explain the divergence of algal SLT sulfate transporters from SAC1. It is hypothesized that *Chlamydomonas* or green algae in general may have the flexibility to utilize proton/sulfate transporter (SULTR) or sodium/sulfate transporter (SLT) depending on the environmental conditions such as pH and sodium concentrations which they need to acclimate (Pootakham et al., 2010). The sodium-dependency of sulfate transport activity of SLT needs to be verified to support this model.

SLC13 AND TDT

The lineage of SLC13 is composed of animal SLC13 (Markovich and Murer, 2004; Pajor, 2006; **Figure 4B**). No plant or algal proteins are associated with this family. Based on the phylogeny, the expansion pattern of this SLC13 family can be interpreted in a straightforward way. The SLC13 ancestor first diverged to form two lineages; one specific to vertebrates and the other that further duplicates to form clades specific to *Drosophila* and *C. elegans* (**Figure 4B**). A *Drosophila* lifespan determinant protein, Indy, is known to function as a sodium-independent electroneutral citrate transporter (Rogina et al., 2000; Inoue et al., 2002). By contrast, NaC family members of *C. elegans* co-transport sodium and dicarboxylate (Fei et al., 2003). These lines of evidence suggest that sodium-dependency is not always conserved among SLC13. Among the five SLC13 proteins in human, two are sodium/sulfate transporters (NaS1 and NaS2) and the rest three are sodium/carboxylate transporters (NaC1, NaC2, and NaC3; Markovich and Murer, 2004; Pajor, 2006). The phylogenetic tree indicates that both NaS1/NaS2 and NaC1/NaC2 originally come from the ancestor of NaC3. It is suggested that the substrate specificities of NaS1 and NaS2 for sulfate may have been developed later when they diverged from the lineage leading to NaC1 and NaC2.

The TDT family is specific to land plants (**Figure 4B**). The tonoplast-localized dicarboxylate transporter from *Arabidopsis* provides the biochemical evidence for this family (Emmerlich et al., 2003). The *Arabidopsis* TDT is capable of transporting malate and fumarate to vacuoles (Emmerlich et al., 2003; Hurth et al., 2005). There is no direct experimental evidence showing a sulfate transport activity, although a certain degree of sequence similarity with SLC13 has been detected. In addition, it is reported that sodium does not stimulate the dicarboxylate transport activity of TDT (Emmerlich et al., 2003), suggesting it is not functionally equivalent to SLC13. The phylogenetic tree indicates that

the ancestral protein of TDT was first split to give a clade specific to seed plants and a separate clade for *Selaginella* and *Physcomitrella* (**Figure 4B**). The family members in dicotyledonous and monocotyledonous plant species are suggested to have expanded through subsequent gene duplications. There appears to be two separate forms for monocotyledonous TDT, although their functional differences are not known.

Intriguingly, a group of phosphate transporters (PHO) that may share a common ancestry with TDT was identified. This group is composed of low-affinity PHO from yeast (Wykoff and O'Shea, 2001; Hürlimann et al., 2007). Yeast PHO87, PHO90, and PHO91, and homologs from *Aspergillus*, *Chlamydomonas*, *Volvox*, and *Drosophila* are present in this clade. It may be hypothesized that the substrate specificity could have been low during the ancient period and was defined when the ancestor split to TDT and PHO. The low substrate affinity of the existing PHO87 and PHO90 (Wykoff and O'Shea, 2001) may be considered as a remnant of the ancestral trait.

CONCLUSION

Sulfate transporters are essential biological components. The occurrence of types of sulfate transporters may vary depending on genotype, the environment, as well as location at the organ or subcellular compartment level. A number of studies have described the biochemical and physiological functions of sulfate transporters from various organisms. The present study is intended to provide a framework to reinterpret the biological information in the context of their evolutionary relationships. Using protein sequences of sulfate transporters from a diverse set of organisms, distinct evolutionary origins of sulfate transporters were identified. The results suggest that they subsequently underwent gene duplications and eventually expanded to have multiple subfamily members playing potentially specialized roles in sulfate transport processes.

The phylogenetic analysis indicates the evolutionary trajectories of two distinct families of sulfate transporters in green algae and land plant species: (i) Chlorophyte green algae contain both SULTR and SAC1/SLT family members; (ii) SULTR family in chlorophytes is associated with the lineage of plant SULTR4 subfamily but also contains divergent members likely originating from earlier evolutionary events; (iii) Angiosperms has multiple SULTR1–SULTR4 subfamily members but are devoid of SAC1/SLT; (iv) *Selaginella* and *Physcomitrella* are at intermediate positions between algae and angiosperms, as they seem to have partially developed SULTR subfamilies, and have SAC1/SLT homologs as well that are absent in angiosperms. These lines of evidence suggest that plants and algae individually may have developed mechanisms for sulfate transport and sensing to adapt to their natural environments. The upstream sulfur-sensing system may be different between plants and algae, as suggested by the absence of SAC1 in plants. In addition, their sulfate transport systems appear to have different biochemical characteristics and physiological roles.

ACKNOWLEDGMENTS

Shin-Han Shiu is partially supported by NSF grants IOS-1126998 and MCB-0929100. Rothamsted Research receives grant aided support from the Biotechnology and Biological Sciences Research Council (BBSRC), UK.

REFERENCES

Aravind, L., and Koonin, E. V. (2000). The STAS domain – a link between anion transporters and antisigma-factor antagonists. *Curr. Biol.* 10, R53–R55.

Baxter, I., Muthukumar, B., Park, H. C., Buchner, P., Lahner, B., Danku, J., Zhao, K., Lee, J., Hawkesford, M. J., Guerinot, M. L., and Salt, D. E. (2008). Variation in molybdenum content across broadly distributed populations of *Arabidopsis thaliana* is controlled by a mitochondrial molybdenum transporter (MOT1). *PLoS Genet.* 4, e1000004. doi: 10.1371/journal.pgen.1000004

Buchner, P., Takahashi, H., and Hawkesford, M. J. (2004). Plant sulphate transporters: co-ordination of uptake, intracellular and long-distance transport. *J. Exp. Bot.* 55, 1765–1773.

Cherest, H., Davidian, J. C., Thomas, D., Benes, V., Ansorge, W., and Surdin-Kerjan, Y. (1997). Molecular characterization of two high affinity sulfate transporters in *Saccharomyces cerevisiae*. *Genetics* 145, 627–635.

Davies, J. P., Yildiz, F. H., and Grossman, A. (1996). Sac1, a putative regulator that is critical for survival of *Chlamydomonas reinhardtii* during sulfur deprivation. *EMBO J.* 15, 2150–2159.

Emmerlich, V., Linka, N., Reinhold, T., Hurth, M. A., Traub, M., Martinoia, E., and Neuhaus, H. E. (2003). The plant homolog to the human sodium/dicarboxylic cotransporter is the vacuolar malate carrier. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11122–11126.

Fei, Y. J., Inoue, K., and Ganapathy, V. (2003). Structural and functional characteristics of two sodium-coupled dicarboxylate transporters (ceNaDC1 and ceNaDC2) from *Caenorhabditis elegans* and their relevance to life span. *J. Biol. Chem.* 278, 6136–6144.

Gasber, A., Klaumann, S., Trentmann, O., Trampczynska, A., Clemens, S., Schneider, S., Sauer, N., Feifer, I., Bittner, F., Mendel, R. R., and Neuhaus, H. E. (2011). Identification of an *Arabidopsis* solute carrier critical for intracellular transport and inter-organ allocation of molybdate. *Plant Biol.* 13, 710–718.

Hästbacka, J., de la Chapelle, A., Mahtani, M. M., Clines, G., Reeve-Daly, M. P., Daly, M., Hamilton, B. A., Kusumi, K., Trivedi, B., Weaver, A., Coloma, A., Lovett, M., Buckler, A., Kaitila, I., and Lander, E. S. (1994). The diastrophic dysplasia gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell* 78, 1073–1087.

Hawkesford, M. J. (2003). Transporter gene families in plants: the sulfate transporter gene family – redundancy or specialization? *Physiol. Plant* 117, 155–163.

Hawkesford, M. J., Davidian, J. C., and Grignon, C. (1993). Sulphate/proton cotransport in plasma-membrane vescicles isolated from roots of *Brassica napus* L.: increased transport in membranes isolated from sulphur-starved plants. *Planta* 190, 297–304.

Hürlimann, H. C., Stadler-Waibel, M., Werner, T. P., and Freimoser, F. M. (2007). Pho91 is a vacuolar phosphate transporter that regulates phosphate and polyphosphate metabolism in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 18, 4438–4445.

Hurth, M. A., Suh, S. J., Kretzschmar, T., Geis, T., Bregante, M., Gambale, F., Martinoia, E., and Neuhaus, H. E. (2005). Impaired pH homeostasis in *Arabidopsis* lacking the vacuolar dicarboxylate transporter and analysis of carboxylic acid transport across the tonoplast. *Plant Physiol.* 137, 901–910.

Inoue, K., Fei, Y. J., Huang, W., Zhuang, L., Chen, Z., and Ganapathy, V. (2002). Functional identity of *Drosophila melanogaster* Indy as a cation-independent, electroneutral transporter for tricarboxylic acid-cycle intermediates. *Biochem. J.* 367, 313–319.

Kataoka, T., Watanabe-Takahashi, A., Hayashi, N., Ohnishi, M., Mimura, T., Buchner, P., Hawkesford, M. J., Yamaya, T., and Takahashi, H. (2004a). Vacuolar sulfate transporters are essential determinants controlling internal distribution of sulfate in *Arabidopsis*. *Plant Cell* 16, 2693–2704.

Kataoka, T., Hayashi, N., Yamaya, T., and Takahashi, H. (2004b). Root-to-shoot transport of sulfate in *Arabidopsis*: evidence for the role of SULTR3;5 as a component of low-affinity sulfate transport system in the root vasculature. *Plant Physiol.* 136, 4198–4204.

Kopriva, S. (2006). Regulation of sulfate assimilation in *Arabidopsis* and beyond. *Ann. Bot.* 97, 479–495.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.

Krusell, L., Krause, K., Ott, T., Desbrosses, G., Krämer, U., Sato, S., Nakamura, Y., Tabata, S., James, E. K., Sandal, N., Stougaard, J., Kawaguchi, M., Miyamoto, A., Suganuma, N., and Udvardi, M. K. (2005). The sulfate transporter SST1 is crucial for symbiotic nitrogen fixation in *Lotus japonicus* root nodules. *Plant Cell* 17, 1625–1636.

Lass, B., and Ullrich-Eberius, C. L. (1984). Evidence for proton/sulfate cotransport and its kinetics in *Lemna gibba* G1. *Planta* 161, 53–60.

Laudenbach, D. E., and Grossman, A. R. (1991). Characterization and mutagenesis of sulfur-regulated genes in a cyanobacterium: evidence for function in sulfate transport. *J. Bacteriol.* 173, 2739–2750.

Leustek, T., Martin, M. N., Bick, J. A., and Davies, J. P. (2000). Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51, 141–165.

Leves, F. P., Tierney, M. L., and Howitt, S. M. (2008). Polar residues in a conserved motif spanning helices 1 and 2 are functionally important in the SulP transporter family. *Int. J. Biochem. Cell Biol.* 40, 2596–2605.

Lindberg, P., and Melis, A. (2008). The chloroplast sulfate transport system in the green alga *Chlamydomonas reinhardtii*. *Planta* 228, 951–961.

Markovich, D., and Murer, H. (2004). The SLC13 gene family of sodium sulphate/carboxylate cotransporters. *Pflugers Arch.* 447, 594–602.

Melis, A., and Chen, H. C. (2005). Chloroplast sulfate transport in green algae – genes, proteins and effects. *Photosyn. Res.* 86, 299–307.

Mount, D. B., and Romero, M. F. (2004). The SLC26 gene family of multifunctional anion exchangers. *Pflugers Arch.* 447, 710–721.

Pajor, A. M. (2006). Molecular properties of the SLC13 family of dicarboxylate and sulfate transporters. *Pflugers Arch.* 451, 597–605.

Pootakham, W., Gonzalez-Ballester, D., and Grossman, A. R. (2010). Identification and regulation of plasma membrane sulfate transporters in *Chlamydomonas*. *Plant Physiol.* 153, 1653–1668.

Rogina, B., Reenan, R. A., Nilsen, S. P., and Helfand, S. L. (2000). Extended life-span conferred by cotransporter gene mutations in *Drosophila*. *Science* 290, 2137–2140.

Roomans, G. M., Kuypers, G. A. J., Theuvenet, A. P. R., and Borst-Pauwels, G. W. F. H. (1979). Kinetics of sulfate uptake by yeast. *Biochim. Biophys. Acta* 551, 197–206.

Saito, K. (2004). Sulfur assimilatory metabolism. The long and smelling road. *Plant Physiol.* 136, 2443–2450.

Shibagaki, N., Rose, A., Mcdermott, J. P., Fujiwara, T., Hayashi, H., Yoneyama, T., and Davies, J. P. (2002). Selenate-resistant mutants of *Arabidopsis thaliana* identify SULTR1;2, a sulfate transporter required for efficient transport of sulfate into roots. *Plant J.* 29, 475–486.

Sirko, A., Hryniewicz, M., Hulanicka, D., and Böck, A. (1990). Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequence and expression of the cysTWAM gene cluster. *J. Bacteriol.* 172, 3351–3357.

Smith, F. W., Hawkesford, M. J., Ealing, P. M., Clarkson, D. T., Vanden Berg, P. J., Belcher, A. R., and Warrilow, A. G. S. (1997). Regulation of expression of a cDNA from barley roots encoding a high affinity sulfate transporter. *Plant J.* 12, 875–884.

Smith, F. W., Hawkesford, M. J., Prosser, I. M., and Clarkson, D. T. (1995a). Isolation of a cDNA from *Saccharomyces cerevisiae* that encodes a high affinity sulphate transporter at the plasma membrane. *Mol. Gen. Genet.* 247, 709–715.

Smith, F. W., Ealing, P. M., Hawkesford, M. J., and Clarkson, D. T. (1995b). Plant members of a family of sulfate transporters reveal functional subtypes. *Proc. Natl. Acad. Sci. U.S.A.* 92, 9373–9377.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Takahashi, H. (2010). Regulation of sulfate transport and assimilation in plants. *Int. Rev. Cell Mol. Biol.* 281, 129–159.

Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu. Rev. Plant Biol.* 62, 157–184.

Takahashi, H., Watanabe-Takahashi, A., Smith, F. W., Blake-Kalff, M., Hawkesford, M. J., and Saito, K. (2000). The role of three functional sulfate transporters involved in uptake and translocation of sulfate in *Arabidopsis thaliana*. *Plant J.* 23, 171–182.

Takahashi, H., Yamazaki, M., Sasakura, N., Watanabe, A., Leustek, T., de Almeida Engler, J., Engler, G., Van Montagu, M., and Saito, K. (1997). Regulation of sulfur assimilation in higher plants: a sulfate transporter induced in sulfate starved roots plays a central role in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 11102–11107.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

Tejada-Jiménez, M., Llamas, A., Sanz-Luque, E., Galván, A., and Fernández, E. (2007). A high-affinity molybdate transporter in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20126–20130.

Tomatsu, H., Takano, J., Takahashi, H., Watanabe-Takahashi, A., Shibagaki, N., and Fujiwara, T. (2007). An *Arabidopsis* thaliana high-affinity molybdate transporter required for efficient uptake of molybdate from soil. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18807–18812.

Vauclare, P., Kopriva, S., Fell, D., Suter, M., Sticher, L., von Ballmoos, P., Krähenbühl, U., den Camp, R. O., and Brunold, C. (2002). Flux control of sulphate assimilation in *Arabidopsis thaliana*: adenosine 5′-phosphosulphate reductase is more susceptible than ATP sulphurylase to negative control by thiols. *Plant J.* 31, 729–740.

Wykoff, D. D., and O'Shea, E. K. (2001). Phosphate transport and sensing in *Saccharomyces cerevisiae*. *Genetics* 159, 1491–1499.

Yoshimoto, N., Inoue, E., Watanabe-Takahashi, A., Saito, K., and Takahashi, H. (2007). Posttranscriptional regulation of high-affinity sulfate transporters in *Arabidopsis* by sulfur nutrition. *Plant Physiol.* 145, 378–388.

Yoshimoto, N., Takahashi, H., Smith, F. W., Yamaya, T., and Saito, K. (2002). Two distinct high-affinity sulfate transporters with different inducibilities mediate uptake of sulfate in *Arabidopsis* roots. *Plant J.* 29, 465–473.

Zuber, H., Davidian, J. C., Aubert, G., Aimé, D., Belghazi, M., Lugan, R., Heintz, D., Wirtz, M., Hell, R., Thompson, R., and Gallardo, K. (2010). The seed composition of Arabidopsis mutants for the group 3 sulfate transporters indicates a role in sulfate translocation within developing seeds. *Plant Physiol.* 154, 913–926.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2011; paper pending published: 10 November 2011; accepted: 31 December 2011; published online: 19 January 2012.

Citation: Takahashi H, Buchner P, Yoshimoto N, Hawkesford MJ and Shiu S-H (2012) Evolutionary relationships and functional diversity of plant sulfate transporters. Front. Plant Sci. ***2****:119. doi: 10.3389/fpls.2011.00119*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 15 February 2012
doi: 10.3389/fpls.2012.00022

Evolution of plant sucrose uptake transporters

Anke Reinders[1], Alicia B. Sivitz[2] and John M. Ward[1]*

[1] *Department of Plant Biology, University of Minnesota, St. Paul, MN, USA*
[2] *Department of Biological Sciences, Dartmouth College, Hanover, NH, USA*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Christopher Peter Grof, University of Newcastle, Australia
Uener Kolukisaoglu, University of Tuebingen, Germany

***Correspondence:**
John M. Ward, Department of Plant Biology, University of Minnesota, 250 Biological Sciences Center, 1445 Gortner Avenue, St. Paul, MN 55108, USA.
e-mail: jward@umn.edu

In angiosperms, sucrose uptake transporters (SUTs) have important functions especially in vascular tissue. Here we explore the evolutionary origins of SUTs by analysis of angiosperm SUTs and homologous transporters in a vascular early land plant, *Selaginella moellendorffii*, and a non-vascular plant, the bryophyte *Physcomitrella patens*, the charophyte algae *Chlorokybus atmosphyticus*, several red algae and fission yeast, *Schizosaccharomyces pombe*. Plant SUTs cluster into three types by phylogenetic analysis. Previous studies using angiosperms had shown that types I and II are localized to the plasma membrane while type III SUTs are associated with vacuolar membrane. SUT homologs were not found in the chlorophyte algae *Chlamydomonas reinhardtii* and *Volvox carterii*. However, the characean algae *Chlorokybus atmosphyticus* contains a SUT homolog (CaSUT1) and phylogenetic analysis indicated that it is basal to all other streptophyte SUTs analyzed. SUTs are present in both red algae and *S. pombe* but they are less related to plant SUTs than CaSUT1. Both *Selaginella* and *Physcomitrella* encode type II and III SUTs suggesting that both plasma membrane and vacuolar sucrose transporter activities were present in early land plants. It is likely that SUT transporters are important for scavenging sucrose from the environment and intracellular compartments in charophyte and non-vascular plants. Type I SUTs were only found in eudicots and we conclude that they evolved from type III SUTs, possibly through loss of a vacuolar targeting sequence. Eudicots utilize type I SUTs for phloem (vascular tissue) loading while monocots use type II SUTs for phloem loading. We show that HvSUT1 from barley, a type II SUT, reverted the growth defect of the Arabidopsis *atsuc2* (type I) mutant. This indicates that type I and II SUTs evolved similar (and interchangeable) phloem loading transporter capabilities independently.

Keywords: sucrose transporter, SUT, phylogeny, evolution

INTRODUCTION

In angiosperms, H^+-coupled sucrose-uptake transporters (SUTs) are involved in the long-distance transport of sucrose. They function to load sucrose into the phloem (vascular tissue) and in uptake of sucrose into sink tissues such as seeds and flowers. The physiological functions of SUTs have been reviewed recently (Braun and Slewinski, 2009; Kuhn and Grof, 2010; Ayre, 2011). In this paper we focus on the phylogenetic relationship between SUTs in photosynthetic organisms from algae to angiosperms. SUTs are members of the glycoside–pentoside–hexuronide (GPH): cation symporter family which is distantly related to the major facilitator superfamily (Chang et al., 2004). Transporters homologous to SUTs are found in bacteria, fungi, and animals. SUT function in angiosperms predominates in the phloem, and yet SUTs clearly existed prior to evolution of phloem tissue. So it is interesting to investigate SUT sequences in more simple non-vascular land plants and algae to understand the origins of angiosperm SUTs. Analysis of the structure/function of more divergent homologs may also help us understand the SUT transport mechanism.

The SUT homolog SpSUT1 from *Schizosaccharomyces pombe* is a proton-coupled α-glucoside symporter that has a higher affinity for maltose than sucrose (Reinders and Ward, 2001). SUT homologs in animals, including humans, are associated with melanosomes and mutations in the respective genes generally cause hypopigmentation. MATP (Harada et al., 2001) in humans is encoded by *AIM1* or SLC45A2 and mutations result in oculocutaneous albinism type 4 (OCA4; Inagaki et al., 2006). In horse, *MATP* is associated with cream coat color (Mariat et al., 2003). Similarly, in mouse a SUT homolog is encoded by the *underwhite* (*uw*) gene (Newton et al., 2001; Costin et al., 2003), mutations in the AIM1 gene in medaka fish reduce melanin content (Fukamachi et al., 2001) and in birds, plumage color is controlled by alleles of the gene encoding MATP (Gunnarsson et al., 2007). The only animal SUT homolog for which transport activity has been reported is SCRT from *Drosophila*, SCRT is able to transport sucrose and it is localized to subcellular vesicles that resemble melanosomes (Meyer et al., 2011).

In plants, the first SUT was cloned using yeast functional expression (Riesmeier et al., 1992). SUTs are encoded by small gene families in all flowering plants and phylogenetic analysis shows the presence of three groups of SUTs called type I, II, and III (Aoki et al., 2003). Interestingly, type I SUTs are only found in eudicot species. Type I SUTs are necessary for essential functions in eudicots such as phloem loading (Riesmeier et al., 1994; Gottwald et al., 2000) and normal pollen function (Sivitz et al., 2008). All land plant species contain type II and III SUTs. Monocot species utilize type II SUTs for phloem loading (Slewinski et al., 2009). This indicates that evolution of type I SUTs coincided with

monocot and eudicot divergence. Type III SUTs were first cloned from *Arabidopsis*, potato and tomato and characterized as H^+-coupled symporters (Weise et al., 2000). Type III SUTs are localized at the vacuolar membrane (Endler et al., 2006; Reinders et al., 2008) and function in sucrose-uptake into the cytoplasm (Reinders et al., 2008; Schulz et al., 2011).

Advances in genome sequencing allow us for the first time to investigate the origins of angiosperm SUTs. Complete genome sequence is available for representative bryophyte (*Physcomitrella patens*), lycophyte (*Selaginella moellendorffii*), and chlorophytes (*Chlamydomonas reinhardtii* and *Volvox carterii*). In addition, partial sequence is available for the red algae *Galdieria sulphuraria* and *Cyanidioschyzon merolae* and EST sequence is available for several charophyte algae (Timme and Delwiche, 2010). The main questions that we can address by phylogenetic analysis are whether type I SUTs were derived from type II or type III SUTs and whether both type II and III SUTs were represented in the earliest land plants and algae.

MATERIALS AND METHODS

SUT PROTEIN SEQUENCES

All SUT protein sequences were obtained from the following species in which genome sequence is available: the eudicot *Arabidopsis thaliana*, the monocot rice (*Oryza sativa*), the lycophyte *Selaginella moellendorffii*, and the bryophyte *Physcomitrella patens* using BLAST searches on the Phytozome website[1]. The same database was searched for SUT protein sequences from the chlorophytes *Chlamydomonas reinhardtii* and *Volvox carterii*. Dr. Charles F. Delwiche and Mr. James Thierer, University of Maryland, provided support by searching their algal sequence database (Timme and Delwiche, 2010) for SUT homologs in the charophytes *Chlorokybus atmosphyticus*, *Klebsormidium flaccidum*, *Spirogyra pratensis*, *Coleochaete* sp., *Chaetosphaeridium globosum*, *Penium marinum*, and *Nitella hyalina*. In addition, the genome sequence of the red algae *Galdieria sulphuraria*[2] (Barbier et al., 2005) and *Cyanidioschyzon merolae*[3] were searched for the presence of SUTs. Sequences of *Galdieria sulphuraria* SUTs were provided by Dr. Andreas P. M. Weber, University of Düsseldorf.

PHYLOGENETIC ANALYSIS

Multiple protein sequence alignments were generated with Clustal X (Larkin et al., 2007). The variable length N- and C-terminal regions of the alignment were removed. Percent protein sequence identity is presented, based on the trimmed alignment, as average for each cluster (±SD). Sequences with greater than 90% overall sequence identity were not included in the phylogenetic analysis. Phylogenetic analysis was performed through the iPlant Collaborative website[4]. Maximum likelihood analysis was done using PhyML 3.0 with 100 bootstrap replicates (Guindon and Gascuel, 2003; Guindon et al., 2010). Trees were visualized using the FigTree program[5].

[1]http://phytozome.net
[2]http://genomics.msu.edu/cgi-bin/galdieria/blast.cgi
[3]http://merolae.biol.s.u-tokyo.ac.jp/
[4]http://www.iplantcollaborative.org/
[5]http://tree.bio.ed.ac.uk/software/figtree/

COMPLEMENTATION OF THE *ARABIDOPSIS* atsuc2-1 MUTANT

Constructs for plant transformation contained the AtSUC2 (At1g22710) promoter, coding region of either AtSUC2 or HvSUT1 (CAJ20123.1) cDNAs and the AtSUC2 3′UTR. The AtSUC2 promoter (2 kb) was amplified using the primers 5′ggggac aactttgtatagaaaagttgtaccagatttcggtaaatt and 5′ggggactgctttttttgtaca aacttgaagaaagtaagaaaaaaaagaaatt and cloned into the pDONR P4-P1R vector (Invitrogen) using BP clonase II. The AtSUC2 ORF was amplified using 5′caccggtttgtcaaatatggtcagccatcc and 5′atgaaatcccatagtagctttgaag. The HvSUT1 ORF was amplified using 5′caccggtttgtcaaatatggcgcgcggcggcgg and 5′tcagtgaccgccgccg ctgac. The two ORFs were cloned into pENTR/D/TOPO (Invitrogen). The AtSUC2 3′UTR (500 bp) was amplified using 5′gggg acagctttcttgtacaaagtggattgaattttagcagtggt and 5′ggggacaactttgtataa taaagttgaattaactaaaatagataa and cloned into pDONR P2R–P3 (Invitrogen). Constructs were assembled into the pB7m34GW binary vector (Karimi et al., 2005) by directional multi-fragment recombination cloning using LR Clonase Plus (Invitrogen). *Agrobacterium tumefaciens* strain C58C1 containing these plasmids was used to transform heterozygous *atsuc2-1 Arabidopsis* (WS ecotype) plants (Gottwald et al., 2000) by the floral dipping method (Clough and Bent, 1998). Basta-resistant transformed plants were selected on soil. Homozygous *atsuc2-1* mutants were identified by PCR.

RESULTS

Phylogenetic analysis shows that angiosperm SUTs form three main groups (**Figure 1**). Here, we follow the nomenclature of Aoki et al. (2003) and name these groups type I, II, and III. All SUTs encoded by the eudicot *Arabidopsis thaliana* (seven sequences), the monocot *Oryza sativa* (five), the basal non-vascular moss *Physcomitrella patens* (four), the vascular non-seed spikemoss *Selaginella moellendorffii* (five), and the yeast *Schizosaccharomyces pombe* (one) were included as representatives of those groups where full genome sequence is available. In addition, SUTs that have been functionally characterized were included. In **Figure 1** the commonly used abbreviated names for the transporter genes are listed. In **Table 1** the protein accession number, gene name, protein length, and species are presented and sorted by phylogenetic group. Sequences from single-celled red algae *Cyanidioschyzon merolae* and *Galdieria sulphuraria* that are homologous to SUTs but did not cluster with SUTs encoded by land plants are present in a separate group in **Figure 1** and **Table 1**. Additionally, the fungal sequence SpSUT1 from *Schizosaccharomyces pombe* is homologous (Reinders and Ward, 2001) but did not cluster with plant SUTs. The genome sequence from two chlorophyte green algae, *Chlamydomonas reinhardtii* and *Volvox carteri* is available, however no SUT sequences were identified in these chlorophytes. A single charophyte algal sequence from *Chlorokybus atmosphyticus* (CaSUT1) is present just basal to the plant SUT sequences but does not cluster with type I, II, or II (marked with an asterisk in **Figure 1**).

TYPE I SUTs

Type I SUTs were only found in eudicots. The 26 type I SUTs analyzed here cluster into a single group with an average of 69% (±5%) identity. The lack of type I SUTs in non-vascular land plants

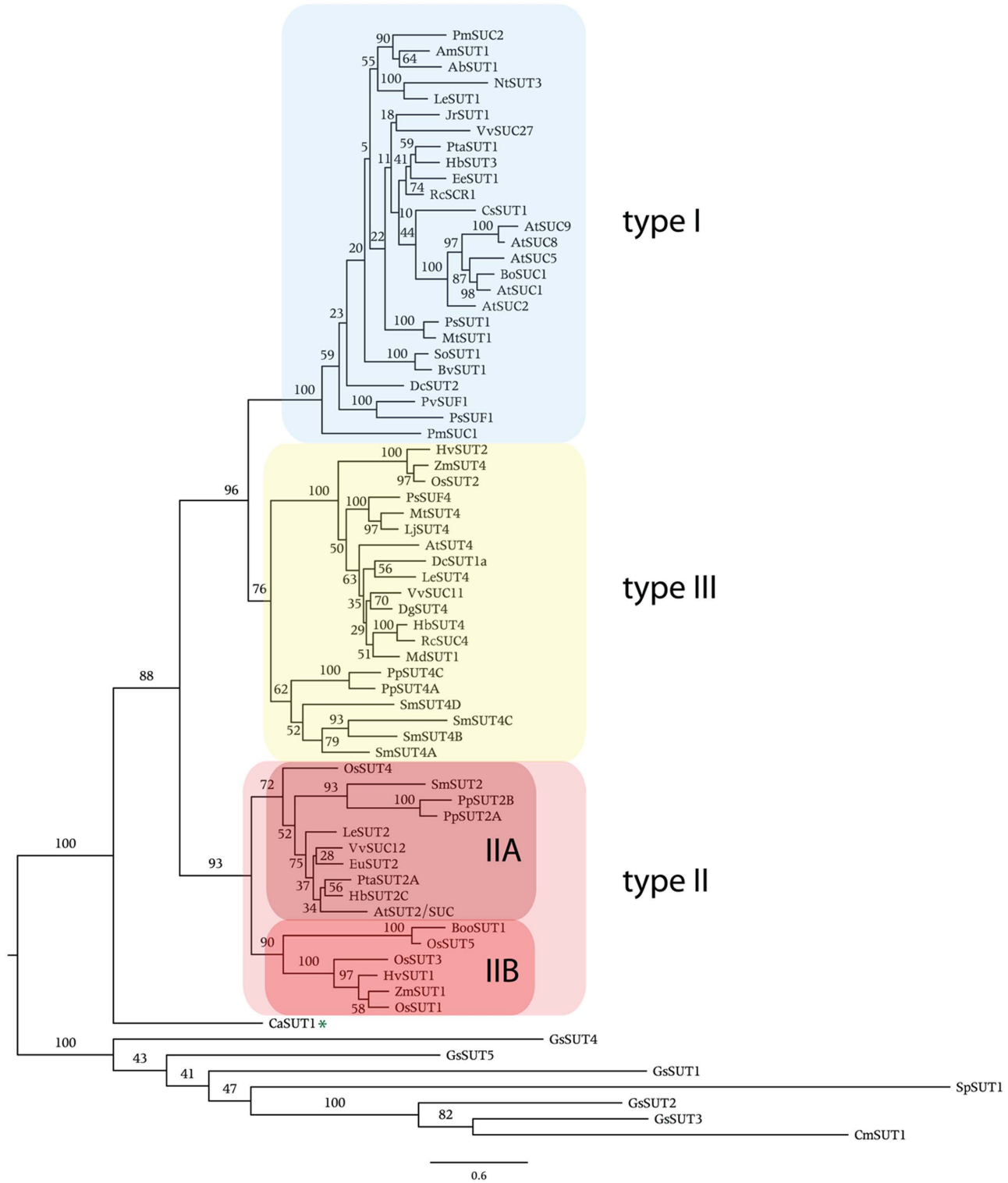


FIGURE 1 | Phylogenetic analysis of plant sucrose transporters and homologs. Protein alignment was done using Clustal X. Sequences with greater than 90% identity were not used in construction of the tree (they are shown in **Table 1**). The variable length N- and C-terminal regions were trimmed from the alignment. The maximum likelihood tree was generated using PhyML 3.0. Numbers indicate percent of 100 bootstrap analyses. Asterisk indicates the single charophyte SUT sequence, CaSUT1, from *Chlorokybus atmosphyticus*.

Table 1 | Sucrose transporter homologs.

Type	Organism	Common name	Gene	Prot ID	Length (aa)	Reference
I	*Alonsoa meridionalis*		AmSUT1	AAF04295	502	Knop et al. (2001)
I	*Arabidopsis thaliana*	Thale cress	AtSUC1 (At1g71880)	CAA53147	513	Sauer and Stolz (1994)
I	*Arabidopsis thaliana*	Thale cress	AtSUC2 (At1g22710)	CAA53150	512	Sauer and Stolz (1994)
I	*Arabidopsis thaliana*	Thale cress	AtSUC5 (At1g71890)	AAG52226	512	Theologis et al. (2000)
I	*Arabidopsis thaliana*	Thale cress	AtSUC8 (At2g14670)	AAC69375	492	Lin et al. (1999)
I	*Arabidopsis thaliana*	Thale cress	AtSUC9 (At5g06170)	BAB09682	491	Tabata et al. (2000)
I	*Asarina barclaiana* (*Maurandya barclaiana*)	Twining snapdragon	AsSUT1	AAF04294	510	Knop et al. (2001)
I	*Beta vulgaris*	Sugar beet	BvSUT1	CAA58730	523	Vaughn et al. (2002)
I	*Brassica oleracea*	Broccoli	BoSUC1	AAL58071	513	Gapper et al. (2005)
I	*Citrus sinensis*	Sweet orange	CsSUT1	AAM29150	528	Li et al. (2003)
I	*Daucus carota*	Carrot	DcSUT2	CAA76369	515	Shakya and Sturm (1998)
I	*Euphorbia esula*	Leafy spurge	EeSUT1	AAF65765	530	
I	*Hevea brasiliensis*	Para rubber tree	HbSUT3/ HbSUT1A	ABK60190	535	Tang et al. (2010)
I	*Juglans regia*	English walnut	JrSUT1	AAU11810	516	Decourteix et al. (2006)
I	*Solanum lycopersicum* (*Lycopersicon esculentum*)	Tomato	LeSUT1	CAA57726	512	Barker et al. (2000)
I	*Medicago truncatula*	Barrel medic	MtSUT1	TC175182, TC184317*	525	http://compbio.dfci.harvard.edu/tgi/
I	*Nicotiana tabacum*	Common tobacco	NtSUT3	AAD34610	521	Lemoine et al. (1999)
I	*Phaseolus vulgaris*	Common bean	PvSUF1	ABB30165	509	Zhou et al. (2007)
I	*Pisum sativum*	Pea	PsSUT1	AAD41024	524	Tegeder et al. (1999)
I	*Pisum sativum*	Pea	PsSUF1	ABB30163	511	Zhou et al. (2007)
I	*Plantago major*	Common plantain	PmSUC1	CAA59113	503	Gahrtz et al. (1996)
I	*Plantago major*	Common plantain	PmSUC2	CAA53390	510	Gahrtz et al. (1996)
I	*Populus trichocarpa*	Black poplar	PtaSUT1/ PtSUT1.2	18221401†	535	Tuskan et al. (2006)
I	*Ricinus communis*	Castor bean	RcSCR1	CAA83436	533	Weig and Komor (1996)
I	*Spinacia oleracea*	Spinach	SoSUT1	CAA47604	526	Riesmeier et al. (1992)
I	*Vitis vinifera*	Grape	VvSUC27	AAF08331	505	Davies et al. (1999)
IIA	*Arabidopsis thaliana*	Thale cress	AtSUT2/AtSUC3 (At2g02860)	CAB92307	595	Meyer et al. (2000), Schulze et al. (2000)
IIA	*Eucommia ulmoides*	Gutta-percha tree	EuSUT2	AAX49396	604	Pang et al. (2008)
IIA	*Hevea brasiliensis*	Para rubber tree	HbSUT2C/ HbSUT2A	CAM33449	539	Dusotoit-Coucaud et al. (2009)
IIA	*Oryza sativa japonica*	Rice	OsSUT4 (Os02g58080)	BAC67164	595	Aoki et al. (2003)
IIA	*Physcomitrella patens*		PpSUT2A	18051919†	635	Rensing et al. (2008)
IIA	*Physcomitrella patens*		PpSUT2B	18064412†	557	Rensing et al. (2008)
IIA	*Populus trichocarpa*	Black poplar	PtaSUT2A	18241865†	602	Tuskan et al. (2006)
IIA	*Selaginella moellendorffii*		SmSUT2	15412113†	521	Banks et al. (2011)
IIA	*Solanum lycopersicum* (*Lycopersicon esculentum*)	Tomato	LeSUT2	AAG12987	605	Barker et al. (2000)
IIA	*Vitis vinifera*	Grape	VvSUC12	AAF08330	612	Davies et al. (1999)

(Continued)

Table 1 | Continued

Type	Organism	Common name	Gene	Prot ID	Length (aa)	Reference
IIB	*Bambusa oldhamii* (*Dendrocalamopsis oldhamii*)	Bamboo	BooSUT1	AAY43226	525	
IIB	*Hordeum vulgare*	Barley	HvSUT1	CAB75882 CAJ20123	523	Weschke et al. (2000), Sivitz et al. (2005)
IIB	*Oryza sativa japonica*	Rice	OsSUT1 (Os03g07480)	BAA24071	537	Hirose et al. (1997)
IIB	*Oryza sativa japonica*	Rice	OsSUT3 (Os10g26740)	BAB68368	506	Aoki et al. (2003)
IIB	*Oryza sativa japonica*	Rice	OsSUT5 (Os02g36700)	BAC67165	535	Aoki et al. (2003)
IIB	*Saccharum hybrid cultivar*	Sugarcane	ShSUT1[#]	AAV41028	517	Rae et al. (2005)
IIB	*Zea mays*	Corn	ZmSUT1	BAA83501	521	Aoki et al. (1999)
III	*Arabidopsis thaliana*	Thale cress	AtSUT4 (At1g09960)	AAL59915	510	Weise et al. (2000)
III	*Datisca glomerata*	Durango root	DgSUT4	CAG70682	498	Schubert et al. (2010)
III	*Daucus carota*	Carrot	DcSUT1a	CAA76367	501	Shakya and Sturm (1998)
III	*Hevea brasiliensis*	Para rubber tree	HbSUT4/ HbSUT4A	ABK60191	498	Tang et al. (2010)
III	*Hordeum vulgare*	Barley	HvSUT2	CAB75881	506	Weschke et al. (2000)
III	*Lotus japonicus*		LjSUT4	CAD61275	511	Flemetakis et al. (2003)
III	*Malus x domestica*	Apple	MdSUT1	AAR17700	499	Fan et al. (2009)
III	*Medicago truncatula*	Barrel medic	MtSUT4	17466537[†]	504	
III	*Oryza sativa japonica*	Rice	OsSUT2 (Os12g44380)	BAC67163	501	Aoki et al. (2003)
III	*Physcomitrella patens*		PpSUT4A	18040351[†]	532	Rensing et al. (2008)
III	*Physcomitrella patens*		PpSUT4B[#]	18037160[†]	500	Rensing et al. (2008)
III	*Physcomitrella patens*		PpSUT4C	18053343[†]	524	Rensing et al. (2008)
III	*Pisum sativum*	Pea	PsSUF4	ABB30162	507	Zhou et al. (2007)
III	*Ricinus communis*	Castor bean	RcSUC4	AAU21439	509	
III	*Selaginella moellendorffii*		SmSUT4A	15419655[†]	514	Banks et al. (2011)
III	*Selaginella moellendorffii*		SmSUT4B	15407332[†]	492	Banks et al. (2011)
III	*Selaginella moellendorffii*		SmSUT4C	15417411[†]	493	Banks et al. (2011)
III	*Selaginella moellendorffii*		SmSUT4D	15402611[†]	531	Banks et al. (2011)
III	*Solanum lycopersicum* (*Lycopersicon esculentum*)	Tomato	LeSUT4	AAG09270	501	Weise et al. (2000)
III	*Vitis vinifera*	Grape	VvSUC11	AAF08329	501	Davies et al. (1999)
III	*Zea mays*	Corn	ZmSUT4	AAT35810	501	
	Chlorokybus atmosphyticus	Soil alga	CaSUT1			
	Cyanidioschyzon merolae		CmSUT1	CMO328C[‡]	502	Matsuzaki et al. (2004)
	Galdieria sulphuraria		GsSUT1	Gs18190[§]	471	Weber et al. (2004), Barbier et al. (2005)
	Galdieria sulphuraria		GsSUT2	Gs34550[§]	546	Weber et al. (2004), Barbier et al. (2005)
	Galdieria sulphuraria		GsSUT3	Gs56570[§]	430	Weber et al. (2004), Barbier et al. (2005)
	Galdieria sulphuraria		GsSUT4	Gs29860[§]	526	Weber et al. (2004), Barbier et al. (2005)
	Galdieria sulphuraria		GsSUT5	Gs08920[§]	638	Weber et al. (2004), Barbier et al. (2005)
	Schizosaccharomyces pombe	Fission yeast	SpSUT1	NP594387	553	Reinders and Ward (2001)

**sequence from DFCI (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=medicago).*

[†] sequence from Phytozome v7.0 (http://www.phytozome.net/).

[#] not included in the phylogenetic analysis (>90% identical to another SUT).

[‡] sequence from Cyanidioschyzon merolae genome project (http://merolae.biol.s.u-tokyo.ac.jp/).

[§] sequence from Galdieria sulphuraria genome project (http://genomics.msu.edu/galdieria/index.html).

(*Physcomitrella*), vascular non-seed land plants (*Selaginella*), and the monocot branch of angiosperms indicates that development of type I SUTs occurred after divergence of monocots and eudicots, around 150 MYR ago (Laroche et al., 1995). Genome sequence representing early diverging eudicots such as *Papaver* sp. (poppy) and *Ranunculus* sp. (buttercup; Soltis et al., 2003) would be useful to more clearly determine the origins of type I SUTs. It is interesting that type I SUT genes were amplified in *Arabidopsis* and acquired specialized functions. *Arabidopsis thaliana* has five genes in this group (**Figure 1**; **Table 1**) and an additional two that have been identified as pseudogenes (Sauer et al., 2004) that were not included in the analysis.

In *Arabidopsis thaliana*, type I SUTs display specialization in both expression and transport function. AtSUC2 is necessary for loading sucrose into the phloem (Gottwald et al., 2000). It has a K_m (affinity) for sucrose of 1.4 mM (Chandran et al., 2003) and a wide substrate specificity for α and β glucosides that is shared with other type I SUTs (**Figure 2**; Chandran et al., 2003). AtSUC1 transport activity is very similar to AtSUC2 but its expression pattern is quite different. AtSUC1 is expressed in trichomes, pollen and roots (Sivitz et al., 2007). AtSUC1 is necessary for normal pollen function (Sivitz et al., 2008). Expression of AtSUC1 in the phloem, under control of the AtSUC2 promoter, has been shown to revert the growth defects of *atsuc2* mutants (Wippel and Sauer, 2011). There are also examples of type I SUTs with modified transport activity. AtSUC9 has a much higher affinity for sucrose compared to other type I SUTs (66 μM; Sivitz et al., 2007) while the substrate specificity is typical of other type I SUTs (**Figure 2**; Sivitz et al., 2007).

TYPE II SUTs

Type II SUT sequences were identified in eudicots, monocots, non-vascular land plants (*Physcomitrella*), and vascular non-seed land plants (*Selaginella*). A total of 16 SUT sequences clustered in the type II group with an average of 62% (±9%) identity. The type II group was divided into two subgroups IIA and IIB. These two subgroups were identified previously (Braun and Slewinski, 2009). There is also a structural difference between type IIA and IIB SUTs. Type IIA proteins have a longer central cytoplasmic loop compared to type IIB SUTs. This is reflected in the average length of proteins in type IIA of 587 amino acids (aa) compared to 523 aa in type IIB (**Table 1**). Each angiosperm genome appears to have one gene in the IIA subgroup. Sequences from *Physcomitrella* (two) and *Selaginella* (one) are also included in the IIA subgroup. PpSUT2A and B from *Physcomitrella* and SmSUT2 contain longer central loops with conserved sequence characteristic of angiosperm type IIA transporters. Overall, this indicates that a type IIA transporter with a longer central loop was an ancestral form of the type II SUTs found in angiosperms.

The type IIB subgroup is monocot specific, rice encodes three type IIB transporters. This group contains the monocot phloem loading SUTs. ZmSUT1 has been shown to be expressed in vascular tissue and to function in phloem loading (Slewinski et al., 2009). Similar to the amplification of type I SUTs in *Arabidopsis*, type IIB SUTs appear to have been amplified in rice. Transport activities of OsSUT1 and OsSUT5 were analyzed by expression in oocytes and electrophysiology. OsSUT5 was found to have a higher affinity for sucrose (2.3 mM) compared to OsSUT1 (7.5 mM) and the activity of OsSUT5 was found to be less pH dependent (Sun et al., 2010).

It is interesting to note that monocots and eudicots utilize different SUTs to load sucrose into the phloem. Differences in substrate specificity between type I SUTs such as AtSUC2 that transport sucrose into the phloem in eudicots and type II SUTs such as HvSUT1 that performs the same function in monocots have been identified (Chandran et al., 2003; Sivitz et al., 2005, 2007; Reinders et al., 2006, 2008; Sun et al., 2008). **Figure 2** shows a summary of substrate specificity results for five sucrose transporters. AtSUC2 and AtSUC9 are both type I sucrose transporters and although AtSUC9 has approximately a 20-fold lower $K_{0.5}$ for sucrose (Sivitz et al., 2007) compared to AtSUC2, they have almost

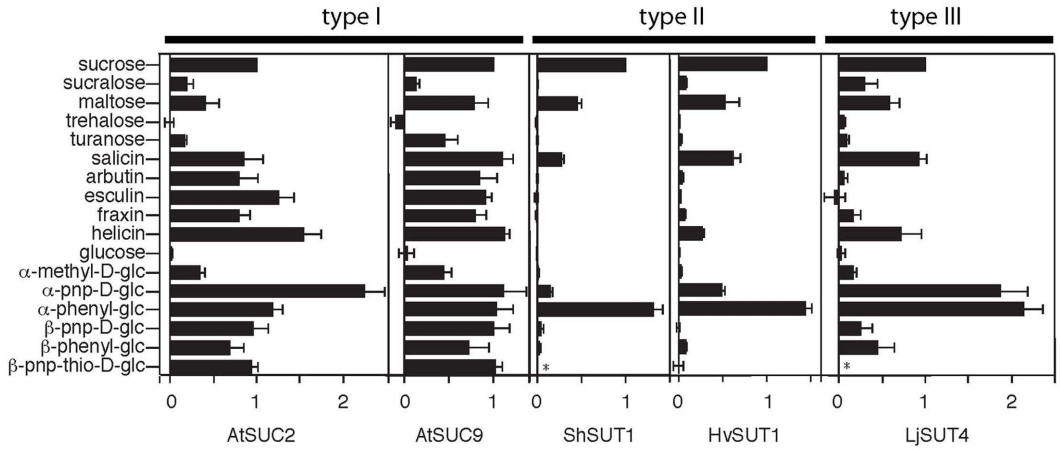


FIGURE 2 | Substrate specificities of type I (AtSUC2, AtSUC9), type II (ShSUT1, HvSUT1), and type III (LjSUT4) plant sucrose transporters. Transport activity was assayed by expression in *Xenopus* oocytes and two-electrode voltage clamping. Oocytes were bathed in sodium Ringer solution containing substrates at concentrations between 0.5 and 25 mM (depending on the transporter affinity and substrate solubility). All currents were normalized to sucrose-dependent currents and are presented as mean ± SE with at least three oocytes per mean. ∗Indicates substrate not tested. Modified with permission from Chandran et al. (2003), Sivitz et al. (2005, 2007), Reinders et al. (2006, 2008).

identical substrate specificities. These type I SUTs transport the plant β-glucosides salicin, arbutin, esculin, fraxin, and helicin. Notably, arbutin, esculin, and fraxin are not transported by the type II transporters ShSUT1 and HvSUT1 (**Figure 2**). Synthetic β phenyl glucosides are also transported by type I and not by type II SUTs (**Figure 2**).

The differences in substrate specificity between type I and type II SUTs might suggest that the specificity of phloem loading in eudicots is different from that in monocots. It is possible that type I SUTs load other glucosides, in addition to sucrose, into the phloem. To begin to address this question we used either AtSUC2 or HvSUT1 to complement the *Arabidopsis atsuc2-1* mutant (Gottwald et al., 2000). The homozygous *atsuc2-1* mutant has greatly reduced growth and accumulates starch in source leaves due to its reduced ability for phloem loading (**Figure 3A**). By comparison, growth of the *atsuc2-1* heterozygous plants is indistinguishable from wild-type (**Figures 3A,B**). As expected, the *atsuc2-1* mutant growth phenotype was complemented by expression of the AtSUC2 gene. Expression of the HvSUT1 coding region driven by the AtSUC2 promoter also resulted in growth that was indistinguishable from wild-type (**Figure 3B**). The type II SUT HvSUT1 appears to revert the growth reduction caused by the loss of AtSUC2 in *Arabidopsis*. This indicates that differences in substrate specificity between type I and II SUTs might not reflect a significant difference in physiological function, although this result is preliminary. Further work is necessary to determine if HvSUT1 fully complements under different growth and stress conditions.

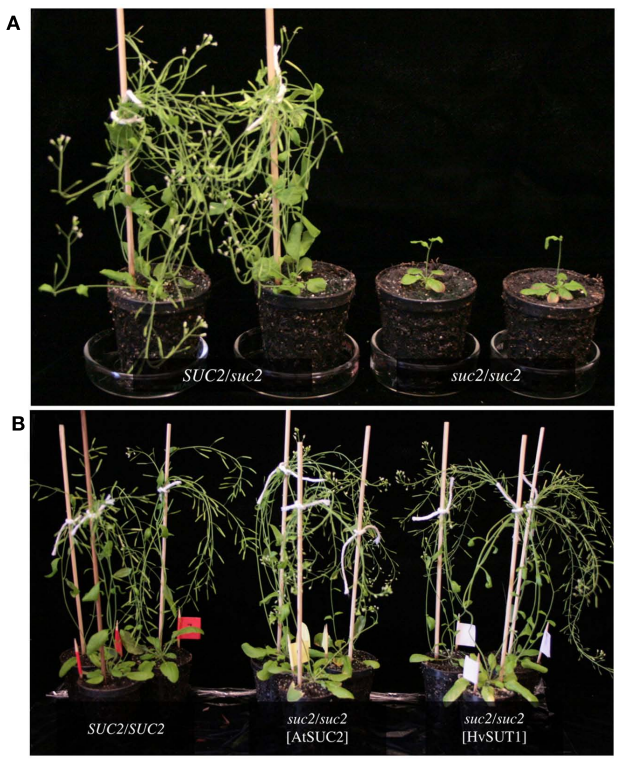


FIGURE 3 | Complementation of *suc2-1* sucrose transporter mutant. **(A)** *Arabidopsis* plants heterozygous for the *suc2-1* insertion left, plants homozygous for the *suc2-1* insertion right. **(B)** Both *AtSUC2* and *HvSUT1* complemented the growth defect of the homozygous *suc2-1* mutant. The WS wild-type (left) is shown for comparison. All plants shown in **(A)** and **(B)** are 8 weeks old.

Finally, the grouping of moss type II SUTs can give us a few more clues about the evolution and function of these ancestral type SUTs. The type II moss and spikemoss sequences cluster with type IIA and contain longer central loops. Both *Physcomitrella* and *Selaginella* lack type I and type IIB SUTs. If early vascular plants such as *Selaginella* have SUTs that function in phloem loading, those transporters are likely to be type IIA such as SmSUT2 and are different from those used by monocots and eudicots. Also, type IIA SUTs in angiosperms do not compensate for loss of the main phloem loading SUT as evidenced by mutant phenotypes of *atsuc2* (Gottwald et al., 2000) and *zmsut1* (Slewinski et al., 2009) mutants.

TYPE III SUTs

The first type III SUTs were isolated from *Arabidopsis*, tomato, potato, and barley and named AtSUT4, LeSUT4, StSUT4, and HvSUT2, respectively (Weise et al., 2000; Weschke et al., 2000). AtSUT4 from *Arabidopsis* and HvSUT2 from barley (Endler et al., 2006), LjSUT4 from *Lotus japonicus* (Reinders et al., 2008), and OsSUT2 from rice (Eom et al., 2011) were demonstrated to localize to the vacuole membrane. Twenty type III SUT sequences were included in this study (**Figure 1**; **Table 1**) and these have an average of 65% (±8%) identity. Each angiosperm genome appears to contain a single type III SUT gene. Both *Selaginella* and *Physcomitrella* contain multiple type III SUT genes. No type III SUT homologs have been identified in green algae.

Transport activity has been characterized in detail for type III SUT LjSUT4 (Reinders et al., 2008). The substrate specificity of LjSUT4 is intermediate between type I and II SUTs (**Figure 2**). Like other type III SUTs (Weise et al., 2000; Weschke et al., 2000) LjSUT4 functions as a H^+-coupled sucrose-uptake transporter. This indicates that its physiological function in the vacuolar membrane is sucrose-uptake into the cytoplasm from the vacuolar lumen. This activity for AtSUT4 has been demonstrated in *Arabidopsis* vacuoles (Schulz et al., 2011).

SUTS IN *CHLOROKYBUS ATMOSPHYTICUS, GALDIERIA SULPHURARIA, CYANIDIOSCHYZON MEROLAE*, AND *SCHIZOSACCHAROMYCES POMBE*

No SUT sequences were found in chlorophytes *Chlamydomonas reinhardtii* and *Volvox carteri*. Charophyte green algae are considered to represent ancestors of land plants. A single SUT sequence was found in the charophyte *Chlorokybus atmosphyticus* (CaSUT1). It did not cluster with type I, II, or III SUTs from land plants but appears to be basal to these clades (**Figure 1**). Since a complete genome sequence of a charophyte is not yet available it remains to be determined whether additional SUTs are present in charophyte genomes. The central loop of CaSUT1 is not extended as in type IIA SUTs. Also, the N-terminal sequence for CaSUT1 is not available so we could not determine if the putative vacuole targeting sequence is present (see Discussion).

Galdieria sulphuraria and *Cyanidioschyzon merolae* are closely related, unicellular red microalgae. While *G. sulphuraria* can grow

on 27 different sugars and sugar alcohols (Gross and Schnarrenberger, 1995), *C. merolae* can not grow heterotrophically (Matsuzaki et al., 2004). Five SUT homologs were identified in *G. sulphuraria* (GsSUT1-5) and one, CmSUT1, was identified in the *C. merolae* genome (**Figure 1**; **Table 1**). This is consistent with the larger number of genes encoding transporters and enzymes involved in carbohydrate metabolism identified in *G. sulphuraria* compared to *C. merolae* (Barbier et al., 2005).

DISCUSSION

THE ORIGIN OF PLANT SUTs IN CHAROPHYTE ALGAE

SUTs function as H^+-coupled cellular sucrose uptake transporters. In angiosperms, type I and II SUTs are localized to the plasma membrane while type III SUTs are localized to the vacuole membrane. They are important for the long-distance transport of sucrose in apoplastic phloem loaders (requiring transmembrane transport). Another important function for SUTs in angiosperms is in sucrose-uptake into sinks that are symplastically isolated such as seeds and pollen. The availability of bryophyte (non-vascular), lycophyte (early vascular), and algal genome sequences allows us to begin to analyze the origins of SUTs in land plants. The presence of CaSUT1 in the charophyte alga *Chlorokybus atmosphyticus* as well as the absence of SUTs in chlorophyte algae (*Chlamydomonas reinhardtii* and *Volvox carterii*) is consistent with the hypothesis that charophyte algae are ancestral to land plants (McCourt et al., 2004).

The physiological function of SUT homologs in *Chlorokybus*, which exists as small clusters of cells and in the unicellular red algae *Galdieria* and *Cyanidioschyzon* is currently unknown but will depend on their membrane localization. They are likely to function as H^+-coupled symporters for glucoside uptake into the cytosol whether they are localized to the plasma membrane or an internal membrane. Interestingly, *Cyanidioschyzon* lacks a central vacuole (Barbier et al., 2005), so it is more likely that CmSUT1 is a plasma membrane transporter. Bryophytes lack true vascular tissue yet *Physcomitrella* contains both type IIA and type III SUTs. In angiosperms, type IIA SUTs are localized to the plasma membrane (Barker et al., 2000; Meyer et al., 2000) while type III SUTs are vacuolar (Endler et al., 2006; Reinders et al., 2008). Therefore, it is likely that *Physcomitrella* contains both plasma membrane and vacuolar SUTs but this will need to be determined experimentally. Long-distance transport of photosynthate in mosses involves leptoid cells and the mechanism appears to be symplasmic, involving plasmodesmata not transmembrane transport (Raven, 2003). Therefore, if SUTs are localized to the plasma membrane in bryophytes their function is not in phloem loading but may be involved in recovery of sucrose that is released to the apoplast. Although leptoid cells evolved independently of phloem, many groups of angiosperms that utilize a similar passive mechanism for phloem loading (Rennie and Turgeon, 2009) also encode SUTs. The function of type III SUTs in bryophytes is likely to be the same as in angiosperms. Sucrose is transiently stored in the vacuole in angiosperms and type III SUTs function in the vacuole membrane to return sucrose from the vacuole lumen to the cytoplasm (Reinders et al., 2008; Schulz et al., 2011). The more recent development of type I SUTs in eudicots and type IIB SUTs in monocots is likely to be linked to the evolution of active phloem loading requiring energy and transmembrane transport.

PUTATIVE VACUOLAR TARGETING MOTIF IN TYPE III SUTs

Recently, a dileucine-like motif (LXXXLL) in the N-terminal cytoplasmic domain of the *Arabidopsis* monosaccharide transporter ESL1was shown to be necessary for localization of the transporter to the vacuole membrane (Yamada et al., 2010). Dileucine-like motifs are recognized by a clathrin-associated, heterotetrameric adaptor protein (AP-3) complex and function in sorting of vacuole membrane proteins in yeast (Vowels and Payne, 1998). Similar dileucine motifs contain an acidic residue spaced several residues prior to the leucine pair with a consensus of DXXLL or [DE]XXXL[LI] (Braulke and Bonifacino, 2009). The AP-3 complex has been shown to be necessary for normal vacuole function in *Arabidopsis* (Zwiewka et al., 2011). An LXXLL motif is found in the cytoplasmic N-terminus of type III SUTs (**Figure 4**) but is lacking in type I and II SUTs. All of the angiosperm type III SUTs contain a perfect LXXLL motif with the exception of AtSUT4 that has the sequence KRVLL (**Figure 4**). AtSUT4 has been demonstrated to localize to the vacuole membrane (Endler et al., 2006) so it is likely that the first leucine of the motif is not strictly required. Recently, localization of AtSUT4 to the vacuole membrane in *Arabidopsis* was shown to be dependent on AP-3 (Wolfenstetter et al., 2012). None of the *Physcomitrella* or *Selaginella* type III SUTs contain a

Type	Name	Sequence
I	AtSUC1	28SPLRKIISVASIAAGV43
	PsSUT1	33SPLRKIMVVASIAAGV48
II	AtSUT2/SUC3	58CSLVTLVLSCTVAAGV73
	OsSUT1	47ISLGRLILSGMVAGGV62
III	SmSUT4A	15VP**L**RS**L**ARVACVAAGV30
	SmSUT4B	15VP**L**KA**L**ARVASVAAGV30
	SmSUT4C	22VP**L**RG**L**ARVASVALGV37
	PpSUT4A	12VPIRA**L**IQVASVAAGV27
	PpSUT4C	12VPIRA**L**IQVASVAAGV27
	SmSUT4D	26IRQRQ**L**FRVSSVAAGI41
	DcSUT1a	25VS**L**RL**LL**RVASVACGI40
	LeSUT4	24VP**L**RL**LL**RVASVAGGI39
	DgSUT4	21VS**L**RK**LL**RVSSVACGI36
	MdSUT1	21VP**L**RQ**LL**RVASVACGI36
	VvSUC11	25VP**L**RR**LL**RVASVACGI40
	RcSUC4	32VS**L**RK**LL**RVTSIAGGI47
	HbSUT4	21VP**L**RQ**LL**RVTSVAGGI36
	AtSUT4	38VSKRV**LL**RVASVACGI53
	LjSUT4	36VP**L**RQ**LL**RVASVASGI51
	MtSUT4	33TP**L**RQ**LL**RVASVASGI48
	PsSUF4	32VP**L**TK**LL**RVASVAGGI47
	OsSUT2	22VP**L**RK**LL**RAASVACGV37
	ZmSUT4	17VP**L**RK**LL**RAASVACGV32
	HvSUT2	25VP**L**RS**LL**RAASVACGV40
	Motif	**..LXXLL.....**

FIGURE 4 | Putative dileucine-like vacuolar targeting sequence in type III SUTs. A part of the multiple protein alignment of sucrose transporters is shown. All type III SUTs and selected type I and II SUTs are shown for comparison. Numbers indicate the amino acid positions for each protein. Amino acid positions that conform to the dileucine-like motif LXXLL are shown in bold.

complete LXXLL motif and it is unknown whether they localize to the vacuole membrane.

THE ORIGIN OF TYPE I SUTs

Type I SUTs are localized to the plasma membrane in eudicots. Based on phylogeny (**Figure 1**) and substrate specificity (**Figure 2**) they are more similar to type III SUTs than to type II SUTs. Since type III SUTs are present in bryophytes and lycophytes, we suggest that type I SUTs are derived from vacuolar-type III SUTs. This would likely involve mutation of the vacuolar targeting information resulting in localization to the plasma membrane, the default targeting pathway for membrane proteins in plants. We hypothesize that the LXXLL motif found in type III SUTs serves as the vacuolar targeting domain but this needs to be tested directly.

CONCLUSION

Angiosperm SUTs clustered into three groups, type I, II, and III. Type I SUTs, only found in eudicots appear to have evolved from vacuolar-type III SUTs which were found in all land plants from bryophytes to angiosperms. Type II SUTs were divided into an ancestral form, type IIA, that exist in all land plants and have an extended central loop. Type IIB SUTs only exist in monocots and include the phloem loading transporters in those species. Here we identify an algal SUT (CaSUT1) from the charophyte *Chlorokybus atmosphyticus*. Based on phylogenetic analysis, CaSUT1 appears basal to three types of land plant SUTs and this is consistent with the hypothesis that charophytes are ancestral to land plants.

ACKNOWLEDGMENTS

We thank Dr. Charles F. Delwiche and Mr. James Thierer, Cell Biology and Molecular Genetics, University of Maryland for providing sequences from charophyte algae prior to publication. We thank Dr. Andreas Weber, University of Düsseldorf for providing sequences of SUT homologs from *Galdieria sulphuraria*. This work was supported by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences of the U.S. Department of Energy grant DE-FG02-10ER15886 (John M. Ward).

REFERENCES

Aoki, N., Hirose, T., Scofield, G. N., Whitfeld, P. R., and Furbank, R. T. (2003). The sucrose transporter gene family in rice. *Plant Cell Physiol.* 44, 223–232.

Aoki, N., Hirose, T., Takahashi, S., Ono, K., Ishimaru, K., and Ohsugi, R. (1999). Molecular cloning and expression analysis of a gene for a sucrose transporter in maize (*Zea mays* L.). *Plant Cell Physiol.* 40, 1072–1078.

Ayre, B. G. (2011). Membrane-transport systems for sucrose in relation to whole-plant carbon partitioning. *Mol. Plant* 4, 377–394.

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., Depamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., Debarry, J., Dreyer, I., Elias, M., Engstrom, E. M., Estelle, M., Feng, L., Finet, C., Floyd, S. K., Frommer, W. B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K. G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T. P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T., Mueller-Roeber, B., Nelson, D. R., Obara, M., Oguri, Y., Olmstead, R. G., Onodera, N., Petersen, B. L., Pils, B., Prigge, M., Rensing, S. A., Riano-Pachon, D. M., Roberts, A. W., Sato, Y., Scheller, H. V., Schulz, B., Schulz, C., Shakirov, E. V., Shibagaki, N., Shinohara, N., Shippen, D. E., Sorensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theissen, G., Ulvskov, P., Wakazuki, S., Weng, J. K., Willats, W. W., Wipf, D., Wolf, P. G., Yang, L., Zimmer, A. D., Zhu, Q., Mitros, T., Hellsten, U., Loque, D., Otillar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D., and Grigoriev, I. V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963.

Barbier, G., Oesterhelt, C., Larson, M. D., Halgren, R. G., Wilkerson, C., Garavito, R. M., Benning, C., and Weber, A. P. (2005). Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol.* 137, 460–474.

Barker, L., Kuhn, C., Weise, A., Schulz, A., Gebhardt, C., Hirner, B., Hellmann, H., Schulze, W., Ward, J. M., and Frommer, W. B. (2000). SUT2, a putative sucrose sensor in sieve elements. *Plant Cell* 12, 1153–1164.

Braulke, T., and Bonifacino, J. S. (2009). Sorting of lysosomal proteins. *Biochim. Biophys. Acta* 1793, 605–614.

Braun, D. M., and Slewinski, T. L. (2009). Genetic control of carbon partitioning in grasses: roles of sucrose transporters and tie-dyed loci in phloem loading. *Plant Physiol.* 149, 71–81.

Chandran, D., Reinders, A., and Ward, J. M. (2003). Substrate specificity of the *Arabidopsis thaliana* sucrose transporter AtSUC2. *J. Biol. Chem.* 278, 44320–44325.

Chang, A. B., Lin, R., Keith Studley, W., Tran, C. V., and Saier, M. H. Jr. (2004). Phylogeny as a guide to structure and function of membrane transport proteins. *Mol. Membr. Biol.* 21, 171–181.

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743.

Costin, G. E., Valencia, J. C., Vieira, W. D., Lamoreux, M. L., and Hearing, V. J. (2003). Tyrosinase processing and intracellular trafficking is disrupted in mouse primary melanocytes carrying the underwhite (uw) mutation. A model for oculocutaneous albinism (OCA) type 4. *J. Cell Sci.* 116, 3203–3212.

Davies, C., Wolf, T., and Robinson, S. P. (1999). Three putative sucrose transporters are differentially expressed in grapevine tissues. *Plant Sci.* 147, 93–100.

Decourteix, M., Alves, G., Brunel, N., Ameglio, T., Guillio, A., Lemoine, R., Petel, G., and Sakr, S. (2006). JrSUT1, a putative xylem sucrose transporter, could mediate sucrose influx into xylem parenchyma cells and be up-regulated by freeze-thaw cycles over the autumn-winter period in walnut tree (*Juglans regia* L.). *Plant Cell Environ.* 29, 36–47.

Dusotoit-Coucaud, A., Brunel, N., Kongsawadworakul, P., Viboonjun, U., Lacointe, A., Julien, J. L., Chrestin, H., and Sakr, S. (2009). Sucrose importation into laticifers of *Hevea brasiliensis*, in relation to ethylene stimulation of latex production. *Ann. Bot.* 104, 635–647.

Endler, A., Meyer, S., Schelbert, S., Schneider, T., Weschke, W., Peters, S. W., Keller, F., Baginsky, S., Martinoia, E., and Schmidt, U. G. (2006). Identification of a vacuolar sucrose transporter in barley and *Arabidopsis* mesophyll cells by a tonoplast proteomic approach. *Plant Physiol.* 141, 196–207.

Eom, J. S., Cho, J. I., Reinders, A., Lee, S.W., Yoo, Y., Tuan, P.Q., Choi, S. B., Bang, G., Park, Y. I., Cho, M. H., Bhoo, S. H., An, G., Hahn, T. R., Ward, J. M., and Jeon, J. S. (2011). Impaired function of the tonoplast-localized sucrose transporter in rice, OsSUT2, limits the transport of vacuolar reserve sucrose and affects plant growth. *Plant Physiol.* 157, 109–119.

Fan, R. C., Peng, C. C., Xu, Y. H., Wang, X. F., Li, Y., Shang, Y., Du, S. Y., Zhao, R., Zhang, X. Y., Zhang, L. Y., and Zhang, D. P. (2009). Apple sucrose transporter SUT1 and sorbitol transporter SOT6 interact with cytochrome b5 to regulate their affinity for substrate sugars. *Plant Physiol.* 150, 1880–1901.

Flemetakis, E., Dimou, M., Cotzur, D., Efrose, R. C., Aivalakis, G., Colebatch, G., Udvardi, M., and Katinakis, P. (2003). A sucrose transporter, LjSUT4, is up-regulated during *Lotus japonicus* nodule

development. *J. Exp. Bot.* 54, 1789–1791.

Fukamachi, S., Shimada, A., and Shima, A. (2001). Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. *Nat. Genet.* 28, 381–385.

Gahrtz, M., Schmelzer, E., Stolz, J., and Sauer, N. (1996). Expression of the PmSUC1 sucrose carrier gene from *Plantago major* L. is induced during seed development. *Plant J.* 9, 93–100.

Gapper, N. E., Coupe, S. A., McKenzie, M. J., Sinclair, B. K., Lill, R. E., and Jameson, P. E. (2005). Regulation of harvest-induced senescence in broccoli (*Brassica oleracea* var. italica) by cytokinin, ethylene, and sucrose. *J. Plant Growth Regul.* 24, 153–165.

Gottwald, J. R., Krysan, P. J., Young, J. C., Evert, R. F., and Sussman, M. R. (2000). Genetic evidence for the in planta role of phloem-specific plasma membrane sucrose transporters. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13979–13984.

Gross, W., and Schnarrenberger, C. (1995). Heterotrophic growth of 2 strains of the acido-thermophilic red alga *Galdieria sulphuraria. Plant Cell Physiol.* 36, 633–638.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.

Gunnarsson, U., Hellstrom, A. R., Tixier-Boichard, M., Minvielle, F., Bed'hom, B., Ito, S., Jensen, P., Rattink, A., Vereijken, A., and Andersson, L. (2007). Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. *Genetics* 175, 867–877.

Harada, M., Li, Y. F., El-Gamil, M., Rosenberg, S. A., and Robbins, P. F. (2001). Use of an in vitro immunoselected tumor line to identify shared melanoma antigens recognized by HLA-A*0201-restricted T cells. *Cancer Res.* 61, 1089–1094.

Hirose, T., Imaizumi, N., Scofield, G. N., Furbank, R. T., and Ohsugi, R. (1997). cDNA cloning and tissue specific expression of a gene for sucrose transporter from rice (*Oryza sativa* L.). *Plant Cell Physiol.* 38, 1389–1396.

Inagaki, K., Suzuki, T., Ito, S., Suzuki, N., Adachi, K., Okuyama, T., Nakata, Y., Shimizu, H., Matsuura, H., Oono, T., Iwamatsu, H., Kono, M., and Tomita, Y. (2006). Oculocutaneous albinism type 4: six novel mutations in the membrane-associated transporter protein gene and their phenotypes. *Pigment Cell Res.* 19, 451–453.

Karimi, M., De Meyer, B., and Hilson, P. (2005). Modular cloning in plant cells. *Trends Plant Sci.* 10, 103–105.

Knop, C., Voitsekhovskaja, O., and Lohaus, G. (2001). Sucrose transporters in two members of the Scrophulariaceae with different types of transport sugar. *Planta* 213, 80–91.

Kuhn, C., and Grof, C. P. (2010). Sucrose transporters of higher plants. *Curr. Opin. Plant Biol.* 13, 288–298.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.

Laroche, J., Li, P., and Bousquet, J. (1995). Mitochondrial DNA and monocot-dicot divergence time. *Mol. Biol. Evol.* 12, 1151–1156.

Lemoine, R., Burkle, L., Barker, L., Sakr, S., Kuhn, C., Regnacq, M., Gaillard, C., Delrot, S., and Frommer, W. B. (1999). Identification of a pollen-specific sucrose transporter-like protein NtSUT3 from tobacco. *FEBS Lett.* 454, 325–330.

Li, C. Y., Weiss, D., and Goldschmidt, E. E. (2003). Effects of carbohydrate starvation on gene expression in citrus root. *Planta* 217, 11–20.

Lin, X., Kaul, S., Rounsley, S., Shea, T. P., Benito, M. I., Town, C. D., Fujii, C. Y., Mason, T., Bowman, C. L., Barnstead, M., Feldblyum, T. V., Buell, C. R., Ketchum, K. A., Lee, J., Ronning, C. M., Koo, H. L., Moffat, K. S., Cronin, L. A., Shen, M., Pai, G., Van Aken, S., Umayam, L., Tallon, L. J., Gill, J. E., Adams, M. D., Carrera, A. J., Creasy, T. H., Goodman, H. M., Somerville, C. R., Copenhaver, G. P., Preuss, D., Nierman, W. C., White, O., Eisen, J. A., Salzberg, S. L., Fraser, C. M., and Venter, J. C. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana. Nature* 402, 761–768.

Mariat, D., Taourit, S., and Guerin, G. (2003). A mutation in the MATP gene causes the cream coat colour in the horse. *Genet. Sel. Evol.* 35, 119–133.

Matsuzaki, M., Misumi, O., Shin, I. T., Maruyama, S., Takahara, M., Miyagishima, S. Y., Mori, T., Nishida, K., Yagisawa, F., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara, Y., and Kuroiwa, T. (2004). Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428, 653–657.

McCourt, R. M., Delwiche, C. F., and Karol, K. G. (2004). Charophyte algae and land plant origins. *Trends Ecol. Evol. (Amst.)* 19, 661–666.

Meyer, H., Vitavska, O., and Wieczorek, H. (2011). Identification of an animal sucrose transporter. *J. Cell Sci.* 124, 1984–1991.

Meyer, S., Melzer, M., Truernit, E., Hummer, C., Besenbeck, R., Stadler, R., and Sauer, N. (2000). AtSUC3, a gene encoding a new *Arabidopsis* sucrose transporter, is expressed in cells adjacent to the vascular tissue and in a carpel cell layer. *Plant J.* 24, 869–882.

Newton, J. M., Cohen-Barak, O., Hagiwara, N., Gardner, J. M., Davisson, M. T., King, R. A., and Brilliant, M. H. (2001). Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4. *Am. J. Hum. Genet.* 69, 981–988.

Pang, Y., Zhang, J., Cao, J., Yin, S. Y., He, X. Q., and Cui, K. M. (2008). Phloem transdifferentiation from immature xylem cells during bark regeneration after girdling in *Eucommia ulmoides* Oliv. *J. Exp. Bot.* 59, 1341–1351.

Rae, A. L., Perroux, J. M., and Grof, C. P. (2005). Sucrose partitioning between vascular bundles and storage parenchyma in the sugarcane stem: a potential role for the ShSUT1 sucrose transporter. *Planta* 220, 817–825.

Raven, J. A. (2003). Long-distance transport in non-vascular plants. *Plant Cell Environ.* 26, 73–85.

Reinders, A., Sivitz, A. B., Hsi, A., Grof, C. P., Perroux, J. M., and Ward, J. M. (2006). Sugarcane ShSUT1: analysis of sucrose transport activity and inhibition by sucralose. *Plant Cell Environ.* 29, 1871–1880.

Reinders, A., Sivitz, A. B., Starker, C. G., Gantt, J. S., and Ward, J. M. (2008). Functional analysis of LjSUT4, a vacuolar sucrose transporter from *Lotus japonicus. Plant Mol. Biol.* 68, 289–299.

Reinders, A., and Ward, J. M. (2001). Functional characterization of the alpha-glucoside transporter Sut1p from *Schizosaccharomyces pombe*, the first fungal homologue of plant sucrose transporters. *Mol. Microbiol.* 39, 445–454.

Rennie, E. A., and Turgeon, R. (2009). A comprehensive picture of phloem loading strategies. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14162–14167.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin, I. T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van De Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrano, R. S., and Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69.

Riesmeier, J. W., Willmitzer, L., and Frommer, W. B. (1992). Isolation and characterization of a sucrose carrier cDNA from spinach by functional expression in yeast. *EMBO J.* 11, 4705–4713.

Riesmeier, J. W., Willmitzer, L., and Frommer, W. B. (1994). Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. *EMBO J.* 13, 1–7.

Sauer, N., Ludwig, A., Knoblauch, A., Rothe, P., Gahrtz, M., and Klebl, F. (2004). AtSUC8 and AtSUC9 encode functional sucrose transporters, but the closely related AtSUC6 and AtSUC7 genes encode aberrant proteins in different *Arabidopsis* ecotypes. *Plant J.* 40, 120–130.

Sauer, N., and Stolz, J. (1994). SUC1 and SUC2: two sucrose transporters from *Arabidopsis thaliana*; expression and characterization in baker's yeast and identification of the histidine-tagged protein. *Plant J.* 6, 67–77.

Schubert, M., Melnikova, A. N., Mesecke, N., Zubkova, E. K., Fortte, R., Batashev, D. R., Barth, I., Sauer,

N., Gamalei, Y. V., Mamushina, N. S., Tietze, L. F., Voitsekhovskaja, O. V., and Pawlowski, K. (2010). Two novel disaccharides, rutinose and methylrutinose, are involved in carbon metabolism in *Datisca glomerata*. *Planta* 231, 507–521.

Schulz, A., Beyhl, D., Marten, I., Wormit, A., Neuhaus, E., Poschet, G., Buttner, M., Schneider, S., Sauer, N., and Hedrich, R. (2011). Proton-driven sucrose symport and antiport are provided by the vacuolar transporters SUC4 and TMT1/2. *Plant J.* 68, 129–136.

Schulze, W., Weise, A., Frommer, W. B., and Ward, J. M. (2000). Function of the cytosolic N-terminus of sucrose transporter AtSUT2 in substrate affinity. *FEBS Lett.* 485, 189–194.

Shakya, R., and Sturm, A. (1998). Characterization of source- and sink-specific sucrose/H^+ symporters from carrot. *Plant Physiol.* 118, 1473–1480.

Sivitz, A. B., Reinders, A., Johnson, M. E., Krentz, A. D., Grof, C. P., Perroux, J. M., and Ward, J. M. (2007). *Arabidopsis* sucrose transporter AtSUC9. High-affinity transport activity, intragenic control of expression, and early flowering mutant phenotype. *Plant Physiol.* 143, 188–198.

Sivitz, A. B., Reinders, A., and Ward, J. M. (2005). Analysis of the transport activity of barley sucrose transporter HvSUT1. *Plant Cell Physiol.* 46, 1666–1673.

Sivitz, A. B., Reinders, A., and Ward, J. M. (2008). *Arabidopsis* sucrose transporter AtSUC1 is important for pollen germination and sucrose-induced anthocyanin accumulation. *Plant Physiol.* 147, 92–100.

Slewinski, T. L., Meeley, R., and Braun, D. M. (2009). Sucrose transporter 1 functions in phloem loading in maize leaves. *J. Exp. Bot.* 60, 881–892.

Soltis, D. E., Senters, A. E., Zanis, M. J., Kim, S., Thompson, J. D., Soltis, P. S., Ronse De Craene, L. P., Endress, P. K., and Farris, J. S. (2003). Gunnerales are sister to other core eudicots: implications for the evolution of pentamery. *Am. J. Bot.* 90, 461–470.

Sun, A. J., Xu, H. L., Gong, W. K., Zhai, H. L., Meng, K., Wang, Y. Q., Wei, X. L., Xiao, G. F., and Zhu, Z. (2008). Cloning and expression analysis of rice sucrose transporter genes OsSUT2M and OsSUT5Z. *J. Integr. Plant Biol.* 50, 62–75.

Sun, Y., Reinders, A., Lafleur, K. R., Mori, T., and Ward, J. M. (2010). Transport activity of rice sucrose transporters OsSUT1 and OsSUT5. *Plant Cell Physiol.* 51, 114–122.

Tabata, S., Kaneko, T., Nakamura, Y., Kotani, H., Kato, T., Asamizu, E., Miyajima, N., Sasamoto, S., Kimura, T., Hosouchi, T., Kawashima, K., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A., Nakayama, S., Nakazaki, N., Naruo, K., Okumura, S., Shinpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., Sato, S., De La Bastide, M., Huang, E., Spiegel, L., Gnoj, L., O'Shaughnessy, A., Preston, R., Habermann, K., Murray, J., Johnson, D., Rohlfing, T., Nelson, J., Stoneking, T., Pepin, K., Spieth, J., Sekhon, M., Armstrong, J., Becker, M., Belter, E., Cordum, H., Cordes, M., Courtney, L., Courtney, W., Dante, M., Du, H., Edwards, J., Fryman, J., Haakensen, B., Lamar, E., Latreille, P., Leonard, S., Meyer, R., Mulvaney, E., Ozersky, P., Riley, A., Strowmatt, C., Wagner-McPherson, C., Wollam, A., Yoakum, M., Bell, M., Dedhia, N., Parnell, L., Shah, R., Rodriguez, M., See, L. H., Vil, D., Baker, J., Kirchoff, K., Toth, K., King, L., Bahret, A., Miller, B., Marra, M., Martienssen, R., McCombie, W. R., Wilson, R. K., Murphy, G., Bancroft, I., Volckaert, G., Wambutt, R., Dusterhoft, A., Stiekema, W., Pohl, T., Entian, K. D., Terryn, N., Hartley, N., Bent, E., Johnson, S., Langham, S. A., McCullagh, B., Robben, J., Grymonprez, B., Zimmermann, W., Ramsperger, U., Wedler, H., Balke, K., Wedler, E., Peters, S., van Staveren, M., Dirkse, W., Mooijman, P., Lankhorst, R. K., Weitzenegger, T., Bothe, G., Rose, M., Hauf, J., Berneiser, S., Hempel, S., Feldpausch, M., Lamberth, S., Villarroel, R., Gielen, J., Ardiles, W., Bents, O., Lemcke, K., Kolesov, G., Mayer, K., Rudd, S., Schoof, H., Schueller, C., Zaccaria, P., Mewes, H. W., Bevan, M., Fransz, P., Kazusa DNA Research Institute, Cold Spring Harbor and Washington University in St Louis Sequencing Consortium, and European Union *Arabidopsis* Genome Sequencing Consortium. (2000). Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. *Nature* 408, 823–826.

Tang, C., Huang, D., Yang, J., Liu, S., Sakr, S., Li, H., Zhou, Y., and Qin, Y. (2010). The sucrose transporter HbSUT3 plays an active role in sucrose loading to laticifer and rubber productivity in exploited trees of *Hevea brasiliensis* (para rubber tree). *Plant Cell Environ.* 33, 1708–1720.

Tegeder, M., Wang, X. D., Frommer, W. B., Offler, C. E., and Patrick, J. W. (1999). Sucrose transport into developing seeds of *Pisum sativum* L. *Plant J.* 18, 151–161.

Theologis, A., Ecker, J. R., Palm, C. J., Federspiel, N. A., Kaul, S., White, O., Alonso, J., Altafi, H., Araujo, R., Bowman, C. L., Brooks, S. Y., Buehler, E., Chan, A., Chao, Q., Chen, H., Cheuk, R. F., Chin, C. W., Chung, M. K., Conn, L., Conway, A. B., Conway, A. R., Creasy, T. H., Dewar, K., Dunn, P., Etgu, P., Feldblyum, T. V., Feng, J., Fong, B., Fujii, C. Y., Gill, J. E., Goldsmith, A. D., Haas, B., Hansen, N. F., Hughes, B., Huizar, L., Hunter, J. L., Jenkins, J., Johnson-Hopson, C., Khan, S., Khaykin, E., Kim, C. J., Koo, H. L., Kremenetskaia, I., Kurtz, D. B., Kwan, A., Lam, B., Langin-Hooper, S., Lee, A., Lee, J. M., Lenz, C. A., Li, J. H., Li, Y., Lin, X., Liu, S. X., Liu, Z. A., Luros, J. S., Maiti, R., Marziali, A., Militscher, J., Miranda, M., Nguyen, M., Nierman, W. C., Osborne, B. I., Pai, G., Peterson, J., Pham, P. K., Rizzo, M., Rooney, T., Rowley, D., Sakano, H., Salzberg, S. L., Schwartz, J. R., Shinn, P., Southwick, A. M., Sun, H., Tallon, L. J., Tambunga, G., Toriumi, M. J., Town, C. D., Utterback, T., Van Aken, S., Vaysberg, M., Vysotskaia, V. S., Walker, M., Wu, D., Yu, G., Fraser, C. M., Venter, J. C., and Davis, R. W. (2000). Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. *Nature* 408, 816–820.

Timme, R. E., and Delwiche, C. F. (2010). Uncovering the evolutionary origin of plant molecular processes: comparison of *Coleochaete* (Coleochaetales) and *Spirogyra* (Zygnematales) transcriptomes. *BMC Plant Biol.* 10, 96. doi:10.1186/1471-2229-10-96

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G. L., Cooper, D., Coutinho, P. M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J. C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D. R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C. J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., and Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604.

Vaughn, M. W., Harrington, G. N., and Bush, D. R. (2002). Sucrose-mediated transcriptional regulation of sucrose symporter activity in the phloem. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10876–10880.

Vowels, J. J., and Payne, G. S. (1998). A dileucine-like sorting signal directs transport into an AP-3-dependent, clathrin-independent pathway to the yeast vacuole. *EMBO J.* 17, 2482–2493.

Weber, A. P., Oesterhelt, C., Gross, W., Brautigam, A., Imboden, L. A., Krassovskaya, I., Linka, N., Truchina, J., Schneidereit, J., Voll, H., Voll, L. M., Zimmermann, M., Jamai, A., Riekhof, W. R., Yu, B., Garavito, R. M., and Benning, C. (2004). EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. *Plant Mol. Biol.* 55, 17–32.

Weig, A., and Komor, E. (1996). An active sucrose carrier (Scr1) that is predominantly expressed in the seedling of *Ricinus communis* L. *J. Plant Physiol.* 147, 685–690.

Weise, A., Barker, L., Kuhn, C., Lalonde, S., Buschmann, H., Frommer, W. B., and Ward, J. M. (2000). A new subfamily of sucrose transporters, SUT4, with low affinity/high capacity localized in enucleate sieve elements of plants. *Plant Cell* 12, 1345–1355.

Weschke, W., Panitz, R., Sauer, N., Wang, Q., Neubohn, B., Weber, H., and Wobus, U. (2000). Sucrose transport into barley seeds: molecular characterization of two transporters and implications for seed development and starch accumulation. *Plant J.* 21, 455–467.

Wippel, K., and Sauer, N. (2011). *Arabidopsis* SUC1 loads the phloem in suc2 mutants when expressed from the SUC2 promoter. *J. Exp. Bot.* 63, 669–679.

Wolfenstetter, S., Wirsching, P., Dotzauer, D., Schneider, S., and Sauer, N. (2012). Routes to the tonoplast: the sorting of tonoplast transporters in *Arabidopsis* mesophyll protoplasts. *Plant Cell.* [Epub ahead of print].

Yamada, K., Osakabe, Y., Mizoi, J., Nakashima, K., Fujita, Y., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2010). Functional analysis of an *Arabidopsis thaliana* abiotic stress-inducible facilitated diffusion transporter for monosaccharides. *J. Biol. Chem.* 285, 1138–1146.

Zhou, Y., Qu, H., Dibley, K. E., Offler, C. E., and Patrick, J. W. (2007). A suite of sucrose transporters expressed in coats of developing legume seeds includes novel pH-independent facilitators. *Plant J.* 49, 750–764.

Zwiewka, M., Feraru, E., Moller, B., Hwang, I., Feraru, M. I., Kleine-Vehn, J., Weijers, D., and Friml, J. (2011). The AP-3 adaptor complex is required for vacuolar function in *Arabidopsis*. *Cell Res.* 21, 1711–1722.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 December 2011; accepted: 20 January 2012; published online: 15 February 2012.

Citation: Reinders A, Sivitz AB and Ward JM (2012) Evolution of plant sucrose uptake transporters. Front. Plant Sci. ***3****:22. doi: 10.3389/fpls.2012.00022*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 February 2012
doi: 10.3389/fpls.2012.00021

Molecular evolution of plant AAP and LHT amino acid transporters

Mechthild Tegeder*[1] *and* ***John M. Ward*[2]**

[1] *School of Biological Sciences, Washington State University, Pullman, WA, USA*
[2] *Department of Plant Biology, University of Minnesota, St. Paul, MN, USA*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Daniel R. Bush, Colorado State University, USA
Guillaume Pilot, Virginia Tech, USA
Torgny Näsholm, Swedish University of Agricultural Sciences, Sweden

***Correspondence:**
Mechthild Tegeder, School of Biological Sciences, Washington State University, Abelson Hall 401, Pullman, WA 99164-4236, USA.
e-mail: tegeder@wsu.edu

Nitrogen is an essential mineral nutrient and it is often transported within living organisms in its reduced form, as amino acids. Transport of amino acids across cellular membranes requires proteins, and here we report the phylogenetic analysis across taxa of two amino acid transporter families, the amino acid permeases (AAPs) and the lysine–histidine-like transporters (LHTs). We found that the two transporter families form two distinct groups in plants supporting the concept that both are essential. AAP transporters seem to be restricted to land plants. They were found in *Selaginella moellendorffii* and *Physcomitrella patens* but not in Chlorophyte, Charophyte, or Rhodophyte algae. AAPs were strongly represented in vascular plants, consistent with their major function in phloem (vascular tissue) loading of amino acids for sink nitrogen supply. LHTs on the other hand appeared prior to land plants. LHTs were not found in chlorophyte algae *Chlamydomonas reinhardtii* and *Volvox carterii*. However, the characean alga *Klebsormidium flaccidum* encodes KfLHT13 and phylogenetic analysis indicates that it is basal to land plant LHTs. This is consistent with the hypothesis that characean algae are ancestral to land plants. LHTs were also found in both *S. moellendorffii* and *P. patens* as well as in monocots and eudicots. To date, AAPs and LHTs have mainly been characterized in *Arabidopsis* (eudicots) and these studies provide clues to the functions of the newly identified homologs.

Keywords: amino acid, nitrogen, transporter, AAP, LHT, membrane, evolution

INTRODUCTION

Nitrogen is a critical mineral nutrient in all living organisms since it is required for synthesis of a large number of compounds including hormones, nucleotides, and amino acids. As the basic building blocks of proteins, amino acids are needed for metabolism, cellular structure, growth, and development. Amino acid uptake into cells and cellular compartments depends on membrane-integral transporter proteins, and amino acid transporters have been identified in many organisms including bacteria, fungi, animals, and plants (Chang et al., 2004; Boudko, 2010). In plants, amino acid transporters are found in two families within the amino acid–polyamine–choline (APC) transporter superfamily, the amino acid/auxin permease (AAAP), and the APC family. The AAAP family includes transporters from plants, animals, and fungi (Chang et al., 2004), and in plants contains the amino acid permeases (AAPs), lysine–histidine-like transporters (LHTs), proline transporters (ProTs), γ-aminobutyric acid transporters (GATs), ANT1-like aromatic, and neutral amino acid transporters and auxin transporters (AUXs; Wipf et al., 2002; Rentsch et al., 2007). Cationic amino acid transporters (CATs) belong to the APC family and are present in both animals and plants[1])

This study addresses the phylogeny of the plant AAP and LHT transporters. These have been characterized in angiosperms (flowering plants), and specifically in eudicots, and detailed overviews on their substrate specificity, localization, and biological functions have recently been presented (see Fischer et al., 2002; Lee and Tegeder, 2004; Rentsch et al., 2007; Tegeder and Rentsch, 2010; Tegeder et al., 2011). In *Arabidopsis* the AAP family consists of eight members (AtAAP1–8) that generally transport neutral and acidic amino acids with moderate affinity, with the exception of AtAAP3 and AtAAP5 that also transport basic amino acids (Fischer et al., 1995, 2002; Rentsch et al., 2007; Svennerstam et al., 2008). All *Arabidopsis* AtAAPs analyzed to date have been localized to the plasma membrane and they function as H^+-coupled amino acid uptake systems (see Tegeder and Rentsch, 2010). AAPs have been suggested to be involved in a number of physiological processes in plants including amino acid uptake from the soil (Hirner et al., 2006; Lee et al., 2007; Svennerstam et al., 2008), phloem loading or xylem–phloem transfer (Schulze et al., 1999; Okumoto et al., 2002; Koch et al., 2003; Tegeder et al., 2007; Tan et al., 2008; Hunt et al., 2010; Zhang et al., 2010; see also Tegeder and Rentsch, 2010), and seed loading (Schmidt et al., 2007; Tegeder et al., 2007; Tan et al., 2008; Sanders et al., 2009).

Much less is known about the LHTs, a family of 10 members (AtLHT1–10) in *Arabidopsis*. AtLHT1 was originally described as a lysine and histidine selective transporter (Chen and Bush, 1997), but other studies with AtLHT1 and AtLHT2 suggests that LHTs preferentially transport neutral and acidic amino acids with high affinity (Lee and Tegeder, 2004; Hirner et al., 2006; Svennerstam et al., 2007, 2008). Like the AAPs, AtLHTs are localized to the plasma membrane and transport a broad spectrum of amino acids

[1]http://www.tcdb.org/superfamily.php

from the cell wall space into the cell (Hirner et al., 2006; Foster et al., 2008). Based on promoter–*GUS* studies, LHTs have been suggested to be involved in import of organic nitrogen into root and mesophyll cells (Hirner et al., 2006), as well as into pollen and other cells of reproductive floral tissue (Lee and Tegeder, 2004; Foster et al., 2008).

AAPs and LHTs have not yet been described in any organisms other than angiosperms. With the recent progress in genome sequencing we are however now in the excellent position to determine whether AAP and LHT amino acid transporters are present in ancestors of seed plants and to examine the phylogenetic relationship of AAP and LHT proteins. Three major clades form the large monophyletic plant kingdom. These include the green plants (Viridiplantae), Rhodophytes (red algae), and Glaucophytes (freshwater microscopic algae; **Figure 1**; Anderberg et al., 2011). The green plants are grouped into the Chlorophytes that contain algae such as *Chlamydomonas reinhardtii*, and the Streptophytes with algal species (Charophytes) and land plants (Finet et al., 2010; Banks et al., 2011). The land plants are divided in non-vascular plants (Bryophytes; i.e., liverworts, mosses, and hornworts) and vascular plants that split into Lycophytes (non-seed plants) and Euphyllophytes. The Lycophytes contain clubmosses, quilworts, and spikemosses while Euphyllophytes consist of ferns (Monilophytes) and seed-bearing plants (Spermatophytes), which are often grouped into angiosperms (flowering plants) and gymnosperms (i.e., cycads, *Ginkgo*, conifers, and gnetophytes). However, the evolutionary relationships of Spermatophytes are not clearly resolved (Magallon and Sanderson, 2002; Mathews, 2009).

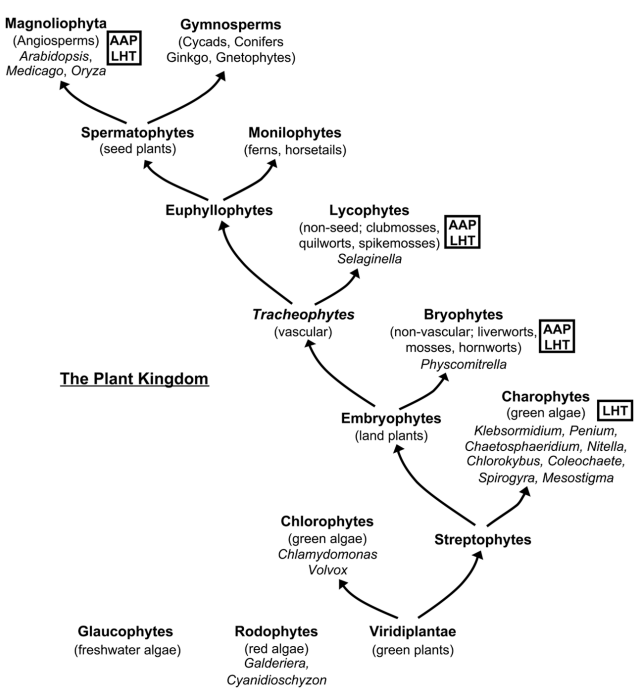


FIGURE 1 | Simplified overview of the plant kingdom and analyzed genera. Taxa containing AAPs and/or LHTs are marked. Note, that genomes of gymnosperms and monilophytes have not been used for the analysis.

Here, sequences from red algae (*Galdieria sulfuraria* and *Cyanidioschyzon merolae*), green algae (Chlorophytes: *Chlamydomonas reinhardtii* and *Volvox carterii*; Charophytes: *Penium marinum*, *Spirogyra praetensis*, *Coleochaete* sp., *Chaetosphaeridium globosum*, *Mesostigma viride*, *Nitella hyalina*, *Klebsormidium flaccidum*, *Chlorokybus atmosphyticus*), and basal non-vascular (*Physcomitrella patens*), non-seed vascular (*Selaginella moellendorffii*), and vascular land plants (eudicots: *Arabidopsis thaliana*, *Medicago sativa*; monocots: *Oryza sativa*) were analyzed for AAP and LHT proteins (**Figure 1**). Phylogenetic reconstruction was performed to determine diversification of the AAP and LHT amino acid transporters as well as their lineage association.

RESULTS AND DISCUSSION

AAP AND LHT TRANSPORTERS FORM TWO DISTINCT GROUPS

Database searches for AAP and LHT proteins in red algae, green algae, basal non-vascular and vascular land plants, and seed plants resulted in 44 AAP and 39 LHT protein sequences (**Figure 2**; **Tables 1** and **2**). Predicted protein sequences for the AAPs averaged 478 ± 14 amino acids (mean ± SD). The length of LHT sequences was similar (463 ± 31 amino acids). The LHT sequence from *Klebsormidium flaccidum* (KfLHT13) is an incomplete cDNA and contains the C-terminal 388 amino acids. A maximum-likelihood

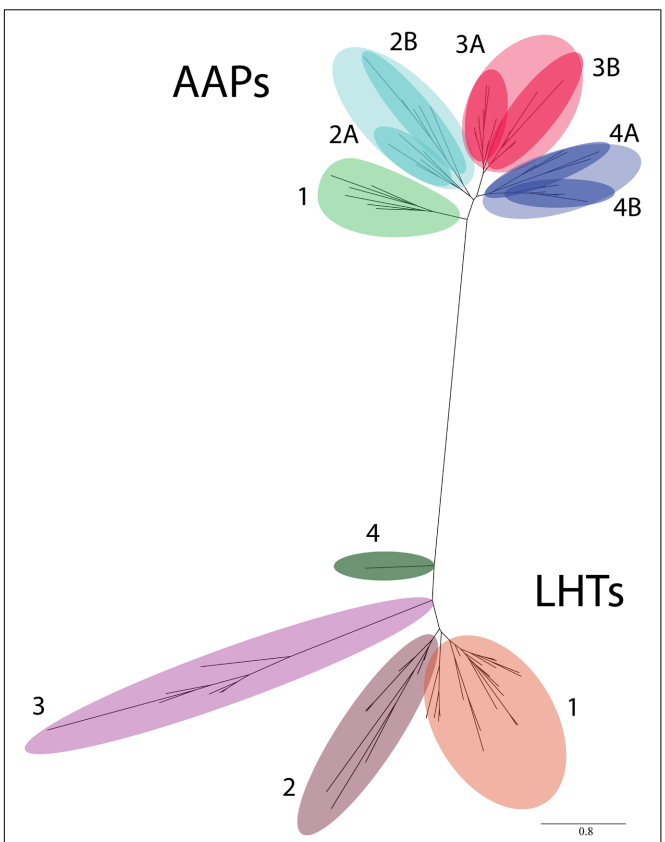


FIGURE 2 | Phylogenetic tree of AAP and LHT proteins found across taxa. A total of 44 AAP and 39 LHT proteins were identified. AAPs and LHTs form two distinct groups, each sub-divided in clusters or subclusters. The clusters (numbers) and subclusters (letters) were labeled and color-coded.

Table 1 | Amino acid permeases protein sequences sorted by subcluster.

Abbreviation	Gene	Organism	Protein size (aa)
SUBCLUSTER 1			
AtAAP7	At5g23810	*Arabidopsis thaliana*	467
MtAAP7A	Medtr3g080570	*Medicago truncatula*	460
MtAAP7B	Medtr5g104490	*Medicago truncatula*	462
OsAAP7A	Os04g39489	*Oryza sativa*	466
OsAAP7B	Os04g56470	*Oryza sativa*	469
OsAAP7C	Os02g49060	*Oryza sativa*	469
SUBCLUSTER 2A			
PpAAP9A	Pp1s107_87V6[1]	*Physcomitrella patens*	515
PpAAP9B	Pp1s387_41V6[1]	*Physcomitrella patens*	500
SmAAP9A	442676[1]	*Selaginella moellendorffii*	503
SmAAP9B	166966[1]	*Selaginella moellendorffii*	467
SmAAP9C	90661[1]	*Selaginella moellendorffii*	479
SUBCLUSTER 2B			
OsAAP10A	Os06g12350	*Oryza sativa*	507
OsAAP10B	Os12g09300	*Oryza sativa*	468
OsAAP10C	Os01g65660	*Oryza sativa*	465
OsAAP10D	Os01g65670	*Oryza sativa*	466
SmAAP10	442677[1]	*Selaginella moellendorffii*	495
SUBCLUSTER 3A			
AtAAP2	At5g09220	*Arabidopsis thaliana*	493
AtAAP3	At1g77380	*Arabidopsis thaliana*	476
AtAAP4	At5g63850	*Arabidopsis thaliana*	466
AtAAP5	At1g44100	*Arabidopsis thaliana*	480
MtAAP2A	Medtr4g143430	*Medicago truncatula*	475
MtAAP2B	Medtr5g017170	*Medicago truncatula*	465
MtAAP2C	Medtr3g142750	*Medicago truncatula*	466
MtAAP2D	Medtr3g142780	*Medicago truncatula*	477
MtAAP2E	Medtr3g142720	*Medicago truncatula*	465
OsAAP3	Os02g01210	*Oryza sativa*	518
SUBCLUSTER 3B			
OsAAP11A	Os12g08090	*Oryza sativa*	475
OsAAP11B	Os12g08130	*Oryza sativa*	475
OsAAP11C	Os11g09020	*Oryza sativa*	476
OsAAP11D	Os12g09320	*Oryza sativa*	468
OsAAP11E	Os01g66010	*Oryza sativa*	488
OsAAP11F	Os05g34980	*Oryza sativa*	496
OsAAP11G	Os04g41350	*Oryza sativa*	471
SUBCLUSTER 4A			
OsAAP12A	Os06g36180	*Oryza sativa*	487
OsAAP12B	Os06g36210	*Oryza sativa*	474
OsAAP12C	Os06g12330	*Oryza sativa*	484
MtAAP12A	Medtr1g008290	*Medicago truncatula*	457
MtAAP12B	Medtr1g008320	*Medicago truncatula*	473
SUBCLUSTER 4B			
AtAAP1	At1g58360	*Arabidopsis thaliana*	485
AtAAP6	At5g49630	*Arabidopsis thaliana*	481
AtAAP8	At1g10010	*Arabidopsis thaliana*	475
MtAAP6A	Medtr1g008410	*Medicago truncatula*	481
MtAAP6B	Medtr3g127950	*Medicago truncatula*	491
OsAAP6	Os07g04180	*Oryza sativa*	487

[1] *Phytozome gene identifier.*

Table 2 | Lysine–histidine-like transporters protein sequences sorted by subcluster.

Abbreviation	Gene	Organism	Protein size (aa)[1]
SUBCLUSTER 1			
AtLHT1	At5G40780	*Arabidopsis thaliana*	446
AtLHT2	At1G24400	*Arabidopsis thaliana*	441
AtLHT3	At1G61270	*Arabidopsis thaliana*	451
AtLHT5	At1G67640	*Arabidopsis thaliana*	441
AtLHT6	At3G01760	*Arabidopsis thaliana*	455
AtLHT8	At1G71680	*Arabidopsis thaliana*	448
AtLHT9	At1G25530	*Arabidopsis thaliana*	440
AtLHT10	At1G48640	*Arabidopsis thaliana*	453
MtLHT1A	Medtr2g122930	*Medicago truncatula*	453
MtLHT1B	Medtr6g025000	*Medicago truncatula*	484
MtLHT2A	AC233656_24.1	*Medicago truncatula*	471
MtLHT2B	Medtr3g103290	*Medicago truncatula*	436
MtLHT3	Medtr8g109640	*Medicago truncatula*	425
MtLHT8	Medtr3g013200	*Medicago truncatula*	469
MtLHT9A	Medtr1g117410	*Medicago truncatula*	437
MtLHT9B	Medtr1g117800	*Medicago truncatula*	437
MtLHT9C	Medtr1g117420	*Medicago truncatula*	437
MtLHT9D	Medtr1g117790	*Medicago truncatula*	437
OsLHT1	Os08g03350	*Oryza sativa*	447
OsLHT2	Os12g14100	*Oryza sativa*	446
OsLHT8	Os05g14820	*Oryza sativa*	456
OsLHT9	Os04g38860	*Oryza sativa*	444
PpLHT11A	Pp1s79_71V6.1[2]	*Physcomitrella patens*	480
PpLHT11B	Pp1s105_62V6.1[2]	*Physcomitrella patens*	465
PpLHT11C	Pp1s5_176V6.1[2]	*Physcomitrella patens*	453
SmLHT11A	270979[2]	*Selaginella moellendorffii*	473
SmLHT11B	127260[2]	*Selaginella moellendorffii*	430
SUBCLUSTER 2			
PpLHT11A	Pp1s79_71V6.1[2]	*Physcomitrella patens*	480
PpLHT11B	Pp1s105_62V6.1[2]	*Physcomitrella patens*	465
PpLHT11C	Pp1s5_176V6.1[2]	*Physcomitrella patens*	453
SmLHT11A	270979[2]	*Selaginella moellendorffii*	473
SmLHT11B	127260[2]	*Selaginella moellendorffii*	430
SmLHT11C	75458[2]	*Selaginella moellendorffii*	427
SmLHT11D	17345[2]4	*Selaginella moellendorffii*	468
SmLHT11E	12727[2]0	*Selaginella moellendorffii*	450
SUBCLUSTER 3			
AtLHT4	At1G47670	*Arabidopsis thaliana*	519
AtLHT7	At4G35180	*Arabidopsis thaliana*	478
MtLHT4A	Medtr2g014200	*Medicago truncatula*	520
MtLHT4B	Medtr2g013940	*Medicago truncatula*	520
MtLHT7	Medtr5g023220	*Medicago truncatula*	534
OsLHT4A	Os12g30040	*Oryza sativa*	508
OsLHT7	Os04g47420	*Oryza sativa*	512
PpLHT4	Pp1s77_57V6.1[2]	*Physcomitrella patens*	559
SUBCLUSTER 4			
KfLHT13	kfla_Contig1880	*Klebsormidium flaccidum*	(388)

[1] *Partial sequences are listed in parentheses.*

[2] *Phytozome gene identifier.*

tree was constructed using PhyML 3.0 (Guindon et al., 2010) based on the alignment of full-length AAP and LHT sequences and the truncated KfLHT13 (**Figure 2**). In addition, trees were made using alignments in which the variable-length N- and C-terminal

regions of the alignment were removed (data not shown). These trees did not differ from those based on full-length AAPs and LHTs, and KfLHT13 (**Figures 2–4**). Both AAPs and LHTs were found in eudicots, monocots, *Selaginella* and *Physcomitrella*, but AAPs and LHTs form two distinct groups supporting functional differences between the two transporter families in the analyzed organisms (see also **Figure 1**). The absence of AAP or LHT genes in Chlorophytes is consistent with the hypothesis that chlorophyte algae are not ancestors of land plants (Turmel et al., 1999; Karol et al., 2001; Kapraun, 2007).

AAPs EVOLVED AT THE SAME TIME AS LAND PLANTS

When searching the databases, AAP proteins were found in non-vascular land plants (*Physcomitrella patens*; 2 proteins), non-seed vascular plants (*Selaginella moellendorffii*, 4 proteins), and seed plants including *Arabidopsis thaliana* (8 proteins), *Oryza sativa* (19 proteins), and *Medicago truncatula* (11 proteins). No AAPs were found in algal sequences of Rhodophytes (*Galdieria* and *Cyanidioschyzon*), Chlorophytes (*Chlamydomonas* and *Volvox*), or Charophytes (*Penium*, *Spirogyra*, *Coleochaete*, *Chaetosphaeridium*, *Nitella*, *Klebsormidium*, and *Chlorokybus*). However, the Charophyte search was based on EST sequences, and until the whole genome sequences are available we cannot rule out that AAPs are present in Charophytes. The identified AAP proteins are grouped into four main clusters (1, 2, 3, and 4), with cluster 2–4 being subdivided into two subclusters A and B (**Table 1**; **Figure 3**). Cluster 1, 3, and 4 only contain AAPs of seed plants while cluster 2 contains non-vascular and non-seed vascular plant, and angiosperm proteins.

Cluster 1 contains proteins from monocots and eudicots that are related to *Arabidopsis* AtAAP7. It holds AtAAP7 and two *Medicago* proteins (MtAAP7A and 7B) consistent with a genome duplication in legumes relative to *Arabidopsis* (Cannon et al., 2006). In addition, it includes three rice AAPs (OsAAP7A–7C) that likely represent an amplification of AAP7 genes in monocots. While the specific function of AtAAP7 and related proteins is still unknown, the phylogenetic analysis supports that they are important for seed plants since they are maintained in both monocot and eudicot lineages.

Cluster 2 contains AAP proteins from non-vascular and non-seed plants, and monocots, but lacks eudicot proteins. Subcluster 2A includes proteins only from the moss *Physcomitrella patens* (PpAAP9A and 9B) and spikemoss *Selaginella moellendorffii* (SmAAP9A–9C), but no proteins from seed plants, suggesting differences in amino acid transporter function between early and higher land plants. Differences in function might be based on (i) differences in phloem loading or source-sink transport between Spermatophytes that have complex leaf venation and the Bryophytes and Lycophytes with no vasculature or microphylls with only a single vascular strands (Reinhart and Thomas, 1981; Aldous, 2002; Beerling, 2005), or on (ii) differences in reproduction (flower versus spores and spore-bearing structures; Prigge and Bezanilla, 2010). For example, specific transporters might be needed for uptake of amino acids into moss sporophytes (Caussin et al., 1983).

In contrast, subcluster 2B contains four proteins from rice (OsAAP10A–10D) and one from *S. moellendorffii*. It is interesting

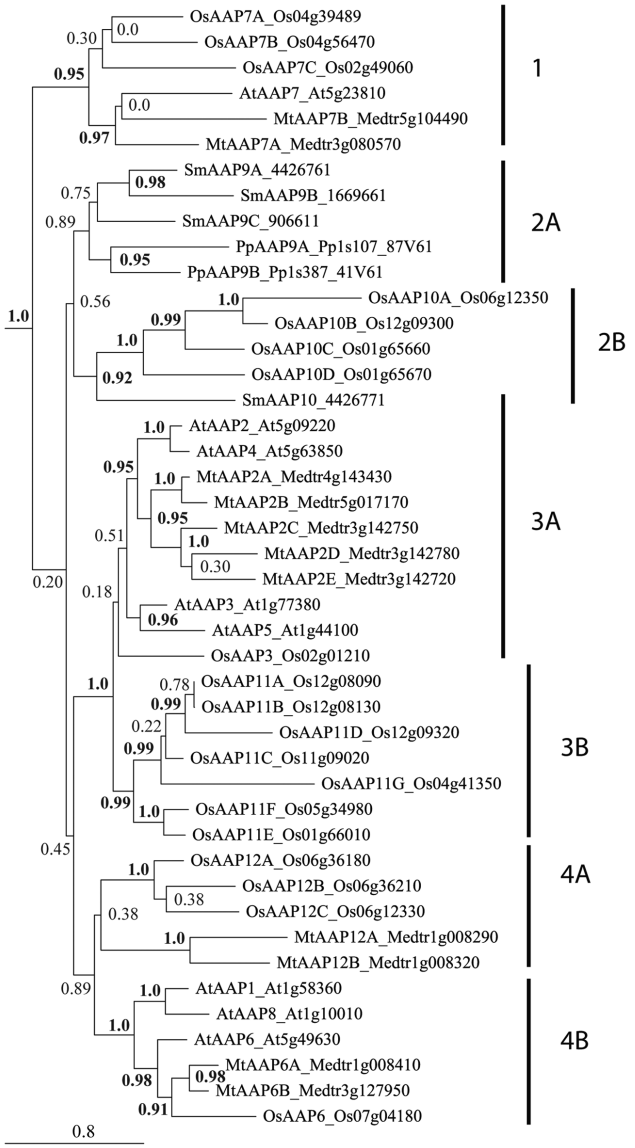


FIGURE 3 | Phylogenetic analysis of AAP proteins found in *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Medicago truncatula* (Mt), *Selaginella moellendorffii* (Sm), and *Physcomitrella patens* (Pp). Multiple protein sequence alignment was done using CLUSTAL X (Thompson et al., 2007). A maximum-likelihood tree was constructed using PhyML 3.0 (Guindon et al., 2010). Numbers at the nodes indicate SH-like branch support (Shimodaira and Hasegawa, 1999; Buckley et al., 2001). Values above 0.9 (in bold) show significant phylogenetic support. Accession numbers of sequences (see **Table 1**) are provided after the species transporter names. The clusters (numbers) and subclusters (letters) are labeled.

that cluster 2 lacks eudicot representation, which indicates that AAPs of this type were lost from eudicot genomes after divergence from monocots. This also suggests that AAPs in cluster 2 represent the earliest AAP sequences. No information is available concerning the function of AAPs in cluster 2. However, the presence of monocot AAPs indicates that an essential amino acid transport activity, required for non-vascular and non-seed vascular plants,

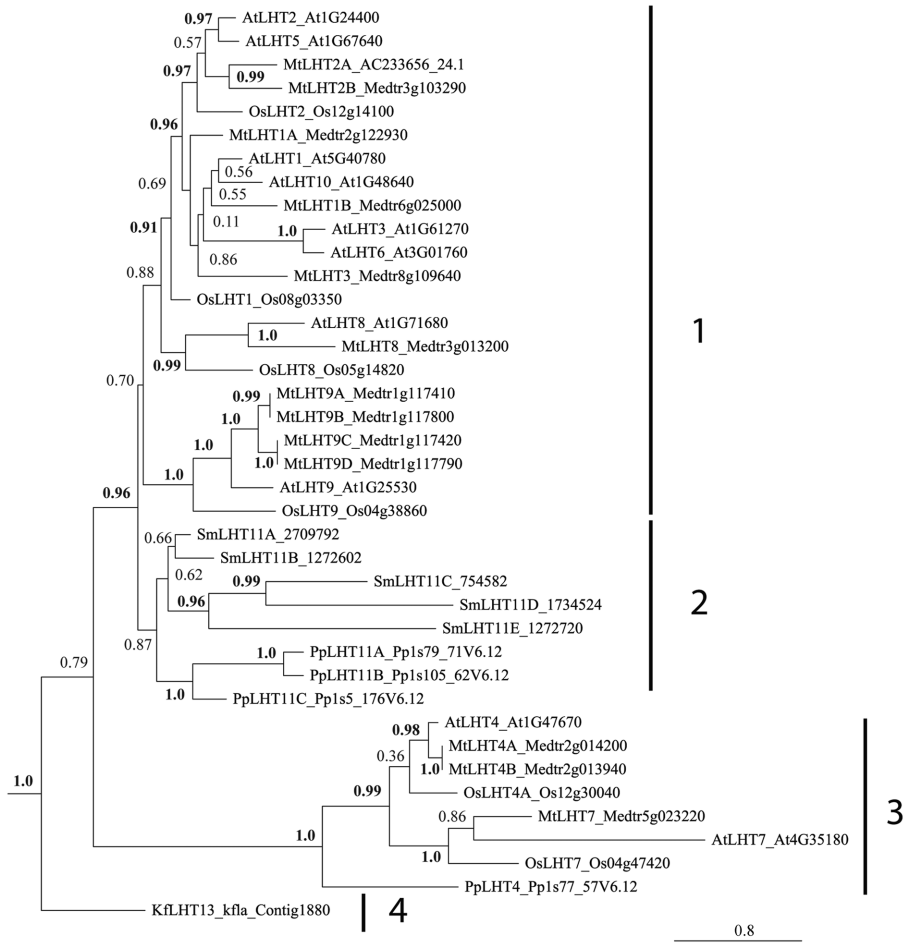


FIGURE 4 | Phylogenetic analysis of LHT proteins found in *Arabidopsis thaliana* (At), *Oryza sativa* (Os), *Medicago truncatula* (Mt), *Selaginella moellendorffii* (Sm), *Physcomitrella patens* (Pp), and *Klebsormidium flaccidum* (Kf). Multiple protein sequence alignment was done using CLUSTAL X (Thompson et al., 2007). A maximum-likelihood tree was constructed using PhyML 3.0 (Guindon et al., 2010). Numbers at the nodes indicate SH-like branch support (Shimodaira and Hasegawa, 1999; Buckley et al., 2001). Values above 0.9 (in bold) show significant phylogenetic support. Accession numbers of sequences (see **Table 2**) are provided after the species transporter names. The clusters are numbered.

was maintained in monocots and likely replaced in eudicots by other AAPs or amino acid transporters in other families.

Cluster 3 contains only monocot and eudicot sequences. Subcluster 3A includes *Arabidopsis* AtAAP3 and AtAAP5 and one related rice protein (OsAAP3), as well as *Arabidopsis* AtAAP2 and AtAAP4 and five legume/*Medicago* AAPs (MtAAP2A–2E). All of the *Arabidopsis* AtAAPs in subcluster 3A appear to be involved in loading of amino acids into the phloem. With the evolution of vascular plants, two vascular tissues were established, the xylem and the phloem. While the xylem functions in water and nutrient transport from the root to the shoot, phloem is important for long-distance transport of nutrients from source (e.g., mature leaves or roots) to sinks such as developing roots, flowers and seeds. In most herbaceous plants such as *Arabidopsis*, rice, or *Medicago*, phloem loading follows the apoplasmic route, where nutrients are loaded from the apoplast into the sieve element-companion cell complex of the collection phloem (Rennie and Turgeon, 2009). In addition, xylem to phloem amino acid transfer might occur along the transport pathway from source to sink (Pate et al., 1975, 1977). These loading steps into the collection or transport phloem require the activity of plasma membrane transporters.

In *Arabidopsis*, AtAAP3 function in the phloem seems to be restricted to the root (Okumoto et al., 2004), while AtAAP5 probably functions in import of amino acids into the companion cells (Brady et al., 2007; Zhang et al., 2008) of different organs including roots and leaves (Fischer et al., 1995; Cartwright et al., 2009; see also Tegeder and Rentsch, 2010). Preliminary results from the Tegeder lab indicate that AtAAP4 also plays a role in leaf phloem loading (Garneau and Tegeder, unpublished). AtAAP2 on the other hand is essential for phloem loading along the transport path (Hirner et al., 1998; Zhang et al., 2010). One rice and five *Medicago* proteins are related to the *Arabidopsis* phloem loaders and we cautiously speculate that they are involved in amino acid import into the sieve element/companion cells complex in legumes and monocot species. This prediction however requires proof through cellular and subcellular localization studies, and functional analysis in plants in

future. Nevertheless, it receives some support from the fact that other legume AAPs, specifically *Phaseolus vulgaris PvAAP1* and *Pisum sativum PsAAP1*, have also been localized to the phloem and group within cluster 3A (Tegeder et al., 2007; Tan et al., 2008).

Cluster 3B only contains seven rice proteins (OAAP11A–11G); eudicots are not represented. It is possible that this large group of monocot AAPs all function in phloem loading of amino acids in different tissues considering the similarity of subcluster 3B to 3A and the presence of only one rice AAP (and multiple eudicot sequences) in 3A. Note that the placement of OsAAP3 in cluster 3A only has weak phylogenetic support. Further, the lack of *Selaginella* and *Physcomitrella* sequences in cluster 3B suggests that AAPs developed independently in monocots, rather than the alternative, that AAPs of this group were lost in eudicots. None of the currently known eudicot amino acid transporters including AAPs from *Arabidopsis* (**Figure 3**), tomato (LeAAPs), potato (StAAPs), pea (PsAAPs), faba bean (VfAAPs), canola (BnAAPs), and *Ricinus* (RcAAPs) falls into cluster 3B (see Tan et al., 2008), providing further support for this hypothesis. Future research needs to determine if the monocot AAP proteins of cluster 3B differ in function from eudicot AAP proteins.

Amino acid permeases in cluster 4 are also divided into two subclusters. Cluster 4A holds three rice (OsAAP12A–12C) and two *Medicago* AAPs (MtAAP12A and 12B) with unknown function. Subcluster 4B contains *Arabidopsis* AtAAP6 related proteins and is branched into a group with AtAAP6 and two (duplicated) *Medicago* proteins, a related single rice protein (OsAAP6), and a group harboring *Arabidopsis* amino acid transporters AtAAP1 and AtAAP8. AtAAP6 is localized to the leaf xylem parenchyma (Okumoto et al., 2002). Although not directly involved in phloem loading, it is predicted to be important for xylem to phloem transfer of amino acids in *Arabidopsis* (Okumoto et al., 2002; Hunt et al., 2010), and the AAP6 relatives in legumes and monocots might have similar functions. It is interesting that one group within cluster 4B only contains two *Arabidopsis* transporters. It seems that *Arabidopsis* has gained two extra copies of AtAAP6: AtAAP1 and AtAAP8. Both AtAAP1 and AtAAP8 proteins are involved in seed loading, rather than phloem loading of amino acids (Schmidt et al., 2007; Sanders et al., 2009), supporting that they are AAP6 paralogs. It is tempting to hypothesize that monocot AAPs in group 4A function in seed loading as nothing is known to date concerning the function of OsAAP12A, B, or C.

Nevertheless, it is important to point out that *AAP* expression is generally not phloem or seed specific (see Ortiz-Lopez et al., 2000; Rentsch et al., 2007; Tegeder and Rentsch, 2010). For example, *Arabidopsis AAP*s (i.e., *AtAAP1* and *AtAAP5*) are also expressed in root epidermal and cortex cells suggesting that they fulfill additional functions in plants including amino acid uptake from the soil (Lee et al., 2007; Svennerstam et al., 2008; Cartwright et al., 2009). Further, the function of AAPs and other amino acid transporters including LHTs seems to be influenced by the physiology of the plant as nitrogen starvation and nitrate re-feeding affects their expression patterns (Liu and Bush, 2006).

In angiosperm, analysis of the relatedness of the rice and legume AAPs with *Arabidopsis* proteins might help with prediction of their function. For example a placement of rice and/or *Medicago* proteins with *Arabidopsis* AAP2, 3, 4, 5, and 6 might suggest a function in phloem loading. Interestingly, rice lacks close relatives of AAP2 and AAP4, and *Medicago* has no AAP3 and AAP5 phloem loaders. Some of the duplicated *Arabidopsis* AAPs may be functionally redundant. At least for AtAAP3 this appears to be the case, since mutant analysis did not result in a functional phenotype (Okumoto et al., 2004). On the other hand, some of the evolved AAPs in legumes (*Medicago*) and monocots seem not to be present in *Arabidopsis*/non-legume dicots (see **Figure 3**, subcluster 2B, 3B, and 4A) further supporting differences in AAP function among angiosperms. This is also in agreement with the large variation in the number of AAP proteins between *Arabidopsis* (8 proteins), *Medicago* (11 proteins), and rice (19 proteins). For example monocots might require additional or different amino acid transporters than eudicots due to differences in morphology and physiology between these distinct groups of seed plants. In legumes, additional AAP proteins might be needed for amino acid transport processes related to N_2 fixation and nodule function.

Taken together, AAPs are mainly found in euphyllophytes, including monocots and eudicots/legumes, which is in agreement with the main functions of AAPs in phloem and seed loading in support of amino acid translocation from source to sink (seeds). Non-vascular and non-seed vascular plants only have AAPs that are more closely related to AAP7, a transporter that remains to be characterized and might differ in function from the other AAPs.

LHTs EVOLVED PRIOR TO LAND PLANTS

Phylogenetic analysis revealed that LHTs are present in Charophytes (*Klebsormidium flaccidum*, 1 protein), non-vascular land plants (*P. patens*, 4 proteins), non-seed vascular plants (*S. moellendorffii*, 5 proteins), and seed plants (*A. thaliana*, 10 proteins; *O. sativa*, 6 proteins; *M. truncatula*, 13 proteins), demonstrating that LHTs evolved before the occurrence of early land plants (**Table 2**). The identified LHT proteins group into four clusters (**Table 2**; **Figure 4**). Cluster 1 includes LHT proteins of euphyllophytes, cluster 2 has five *Selaginella* and three *Physcomitrella* LHT proteins, cluster 3 contains euphyllophyte sequences and one *Physcomitrella* LHT and cluster 4 only contains one *Klebsormidium* protein.

Cluster 1 contains 8 of the 10 *Arabidopsis* AtLHTs, including AtLHT1, 2, 3, 5, 6, 8, 9, and 10. It appears that a duplication event has occurred in *Arabidopsis* and placement of the LHT proteins suggests that AtLHT2 and AtLHT5, AtLHT1 and AtLHT10, and AtLHT3 and AtLHT6, respectively are the result of such duplication. Relatives of AtLHTs are present in *Medicago* and 10 of the 13 MtLHTs are present in cluster 1. The presence of only four rice LHT sequences in group 1 indicates that LHTs were not as extensively duplicated in monocots as in eudicots. Recent studies have shown that *Arabidopsis* LHT transporters of cluster 1 including AtLHT1, 2, 4, 5, and 6 are expressed in male and female floral tissue, such as anther tissue, tapetum, mature pollen, pollen tubes, and pistil transmitting tissue (Hirner et al., 2006; Foster et al., 2008; see also Tegeder and Rentsch, 2010), and it was suggested that they might be essential for successful sexual plant reproduction. This is also in agreement with the observation that LHTs of cluster 1 are only present in flowering plant species. However, experimental proof for LHT function in reproduction is still missing, and AtLHTs of cluster 1 seem to have additional functions in plants as they are expressed in other organs besides flowers (Hirner

et al., 2006; see Arabidopsis eFP Browser, Winter et al., 2007). As for example recently demonstrated for AtLHT1, the transporter is important for amino acid uptake into root and mesophyll cells (Hirner et al., 2006; Svennerstam et al., 2008). LHT8 and LHT9 proteins of cluster 1 form separate subgroups. One group includes *Arabidopsis* AtLHT8 and an ortholog each in rice and *Medicago*, and the second group contains *Arabidopsis* AtLHT9, rice OsLHT9, and four closely related *Medicago* MtLHT9 transporters. Localization of these transporters has not been resolved and similar to most other LHTs, their physiological functions remain to be elucidated.

Cluster 2 includes three *P. patens* (PpLHT11A–C) and five *S. moellendorffii* LHT11 proteins (SmLHT11A–E) but none from Spermatophytes. *LHT11* genes may have evolved independently in *Physcomitrella* and *Selaginella* suggesting that in early land plants these LHTs serve functions in cellular amino acid transport processes that are not required in higher plants (see above). However, as in seed plants, Bryophytes and Lycophytes seem to need both LHTs and AAPs for growth and development (**Figures 1–4**).

Cluster 3 contains LHT4 and LHT7 proteins from angiosperms and *Physcomitrella* PpLHT4. Recent expression studies suggest that *Arabidopsis* AtLHT4 and AtLHT7 might be involved in reproduction, specifically in anther and pollen development (Bock et al., 2006; Foster et al., 2008). However, at least AtLHT4 has most certainly additional functions since it is also expressed in root and stem (Winter et al., 2007). This might explain its phylogenetic divergence from other LHTs and its placement with PpLHT4. In early land plants, transporters may be critical for amino acid movement over relatively short distances. As plants colonized dry land, translocation of amino acids from source to sink cells occurred probably by cell to cell transport (symplasmic) and between cells (apoplasmic), especially in non-vascular mosses (Trachtenberg and Zamski, 1978; Reinhart and Thomas, 1981). Uptake of the apoplasmic amino acids required membrane proteins including H^+-coupled, high affinity LHT symporters, as indicated by the phylogenetic analysis.

Cluster 4 only contains a LHT protein from green algae *Klebsormidium* called KfLHT13, suggesting its evolutionary divergence from LHTs of land plants and differences in function. While we are not aware of amino acid transport studies in Charophytes, research with Chlorophytes such a *Chlamydomonas* and *Chlorella* spp. and marine microalgae demonstrate that in algae different transport systems are present (Kirk and Kirk, 1978a; Cho et al., 1981; Cho and Komor, 1985; Shehawy and Kleiner, 2001; Kato et al., 2006; see also Flynn and Butler, 1986 and references within). Although an LHT transporter was only found in *Klebsormidium*, the screened charophyte sequences were obtained from EST projects and we predict that LHTs are also present in other charophytes besides *Klebsormidium*. Placement of KfLHT13 further supports that LHTs have evolved before land plants and that their function is important to green algae as well.

Gene function of lysine–histidine-like transporters was likely important in ancestors of plants, as a gene is detected in a charophyte, contributing to its high affinity and substrate selectivity for neutral and acid amino acids. Localization and expression studies of *Arabidopsis AtLHTs* suggest that, in addition to other functions, LHTs have a major role in sexual plant reproduction in seed plants. This also indicates a difference in LHT function between angiosperms and non-seed/non-land plants. While LHT functions still need to be demonstrated *in planta*, this is in agreement with the phylogenetic analysis showing a grouping of angiosperm LHTs while LHT proteins from non-vascular and non-seed plants, and green algae are present in separate groups.

Based on the phylogeny, genes in the LHT family of land plants likely arose from an ancestral gene similar to the charophyte LHT. The ancestral gene diversified as plants colonized dry land, as seen by the presence of multiple LHT in moss, a non-vascular plant. In contrast, no algal genes encoding AAP transporters were detected in our analysis of the Charophytes, but it may be too early to conclude whether genes belonging to the AAP family are present, as the charophyte genome has not been sequenced completely. Algae generally acquire amino acids from the environment for growth, and some variation with respect of the kind and amount of amino acids that are taken up has been observed between and within species (Cho et al., 1981; Cho and Komor, 1983, 1985; Flynn and Butler, 1986; Kato et al., 2006). In addition, leakage of amino acids from the cells into the apoplast might occur and requires transporters for retrieval. Physiological studies have demonstrated that in algae active transport systems with varying specificity and affinities (high and low) are present (Kirk and Kirk, 1978a,b; Sauer et al., 1983; Cho and Komor, 1985; Shehawy and Kleiner, 2001; Kato et al., 2006; see also Flynn and Butler, 1986), which might point to the presence of both, LHT and AAP transporters, or additional amino acid transporters in Charophyte algae.

CONCLUSION

Recent functional studies support that AAP and LHT proteins have essential roles in transport of a broad range of amino acids in eudicots (see Tegeder and Rentsch, 2010). Here, phylogenetic analysis supports that AAPs, which generally present moderate and low affinity systems for neutral and acidic amino acids, are important to land plants with a main function in phloem loading and that they are not required in red algae or green algae (Charophytes or Chlorophytes). In contrast, LHTs are found in green algae, non-seed plants and angiosperms suggesting the need for high affinity amino acid transporters across the different organisms.

Both AAPs and LHTs were found in all land plants analyzed consistent with essential and distinct functions for both transporter families. To date, information on the role of AAP and LHT transporters is almost exclusively based on studies in *Arabidopsis* and in some cases in legumes, and suggests differences between AAPs and LHTs in substrate selectivity, transport affinity and cellular function (see Rentsch et al., 2007; Tegeder and Rentsch, 2010; see also above). However, phylogenetic analyses indicates that function of some AAP and LHT transporters diverged in monocots, non-seed vascular plants, non-vascular plants and green algae, and future studies need to address the role of the amino acid transporters across land plants and in green algae.

MATERIALS AND METHODS

IDENTIFICATION AND ANNOTATION OF AAP AND LHT PROTEINS

Genome sequences are available for *Arabidopsis thaliana*, rice (*Oryza sativa*), *Medicago truncatula*, *Selaginella moellendorffii*, *Physcomitrella patens*, *Chlamydomonas reinhardtii*, and *Volvox carterii*. AAP and LHT sequences were selected from rice, *M.*

truncatula, *S. moellendorffii*, and *P. patens* predicted protein sequences using BLAST searches with known *Arabidopsis* AAP and LHT transporters (see **Tables 1** and **2**) on the Phytozome website[2]. The same database was searched for AAP and LHT protein sequences from the Chlorophytes *C. reinhardtii* and *V. carterii*. Dr. Charles F. Delwiche and Dr. James Thierer, University of Maryland provided support by screening their EST (Expressed Sequence Tag) databases for AAP and LHT relatives in charophytes, specifically in *Penium marinum*, *Spirogyra praetensis*, *Coleochaete* sp., and *Chaetosphaeridium globosum*, *Mesostigma viride*, *Nitella hyalina*, *Klebsormidium flaccidum*, *Chlorokybus atmosphyticus*[3]. In addition, the genome (protein) sequences of the red algae *Galdieria sulfuraria* available through http://genomics.msu.edu/cgi-bin/galdieria/blast.cgi (Barbier et al., 2005) and *Cyanidioschyzon merolae* at http://merolae.biol.s.u-tokyo.ac.jp/and were searched for the presence of AAPs and LHTs.

SEQUENCE ALIGNMENTS AND PHYLOGENETIC ANALYSIS

Multiple protein sequence alignments were generated with Clustal X (Thompson et al., 2007) and, for comparison, with MUSCLE (Edgar, 2004). Phylogenetic analysis was performed through the iPlant Collaborative website[4]. Maximum-likelihood analysis was done using PhyML 3.0 (Guindon and Gascuel, 2003; Guindon et al., 2010) and statistical analysis of phylogenetic trees was performed using a Shimodaira–Hasegawa-like test (SH-like test; Shimodaira and Hasegawa, 1999; Buckley et al., 2001). The values for SH-like branch support are presented at the nodes on the trees. Values above 0.9 show significant phylogenetic support. Trees were visualized using the FigTree program[5]

NAMING OF AAP AND LHT TRANSPORTERS

The identified AAP and LHT sequences were named based on clustering with *Arabidopsis* AtAAP and AtLHT protein sequences. In *Arabidopsis* 8 AAP (AtAAP1–8) and 10 LHT (AtLHT1–10) transporters have been previously identified (see Rentsch et al., 2007). Phylogenetic grouping was used to name the transporters from other species. In cases where more than one *Arabidopsis* relative was found from a given species, letter labeling was chosen in addition. For example, three rice relatives of AtAAP7 were named OsAAP7A, 7B, and 7C. Transporters that did not group with *Arabidopsis* proteins were given numbers not found for the *Arabidopsis* transporters such as AAP9 or LHT11.

ACKNOWLEDGMENTS

We thank Dr. Charles Delwiche and Mr. James Thierer, Cell Biology and Molecular Genetics, University of Maryland for providing sequences from Charophyte algae prior to publication. Mechthild Tegeder appreciates the financial support by the National Science Foundation Grant IOS 1021286 and the Agricultural and Food Research Initiative Competitive Grant no. 2010-65115-20382 from the USDA National Institute of Food and Agriculture.

[2]http://phytozome.net
[3]http://www.clfs.umd.edu/labs/delwiche/Charophyte.html
[4]http://www.iplantcollaborative.org/
[5]http://tree.bio.ed.ac.uk/software/figtree/.

REFERENCES

Aldous, A. R. (2002). Nitrogen translocation in *Sphagnum* mosses: effects of atmospheric nitrogen deposition. *New Phytol.* 156, 241–253.

Anderberg, H. I., Danielson, J. Å. H., and Johanson, U. (2011). Algal MIPs, high diversity and conserved motifs. *BMC Evol. Biol.* 11, 110. doi:10.1186/1471-2148-11–110

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., DeBarry, J., Dreyer, I., Elias, M., Engstrom, E. M., Estelle, M., Feng, L., Finet, C., Floyd, S. K., Frommer, W. B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K. G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T. P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T., Mueller-Roeber, B., Nelson, D. R., Obara, M., Oguri, Y., Olmstead, R. G., Onodera, N., Petersen, B. L., Pils, B., Prigge, M., Rensing, S. A., Riaño-Pachón, D. M., Roberts, A. W., Sato, Y., Scheller, H. V., Schulz, B., Schulz, C., Shakirov, E. V., Shibagaki, N., Shinohara, N., Shippen, D. E., Sørensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theissen, G., Ulvskov, P., Wakazuki, S., Weng, J. K., Willats, W. W., Wipf, D., Wolf, P. G., Yang, L., Zimmer, A. D., Zhu, Q., Mitros, T., Hellsten, U., Loqué, D., Otillar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D., and Grigoriev, I. V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963.

Barbier, G., Oesterhelt, C., Larson, M. D., Halgren, R. G., Wilkerson, C., Garavito, R. M., Benning, C., and Weber, A. P. (2005). Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol.* 137, 460–474.

Beerling, D. J. (2005). Leaf evolution: gases, genes and geochemistry. *Ann. Bot.* 96, 345–352.

Bock, K. W., Honys, D., Ward, J. M., Padmanaban, S., Nawrocki, E. P., Hirschi, K. D., Twell, D., and Sze, H. (2006). Integrating membrane transport with male gametophyte development and function through transcriptomics. *Plant Physiol.* 140, 1151–1168.

Boudko, D. Y. (2010). "Molecular ontology of amino acid transport," in *Epithelial Transport Physiology*, ed. G. A. Gerencser (New York, NY: Springer), 379–472.

Brady, S. M., Orlando, D. A., Lee, J. Y., Wang, J. Y., Koch, J., Dinneny, J. R., Mace, D., Ohler, U., and Benfey, P. N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806.

Buckley, T. R., Simon, C., Shimodaira, H., and Chambers, G. K. (2001). Evaluating hypotheses on the origin and evolution of the New Zealand alpine cicadas (*Maoricicada*) using multiple-comparison tests of tree topology. *Mol. Biol. Evol.* 18, 223–234.

Cannon, S. B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., Wang, X., Mudge, J., Vasdewani, J., Schiex, T., Spannagl, M., Monaghan, E., Nicholson, C., Humphray, S. J., Schoof, H., Mayer, K. F., Rogers, J., Quétier, F., Oldroyd, G. E., Debellé, F., Cook, D. R., Retzel, E. F., Roe, B. A., Town, C. D., Tabata, S., Van de Peer, Y., and Young, N. D. (2006). Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14959–14964.

Cartwright, D. A., Brady, S. M., Orlando, D. A., Strumfels, B., and Benfey, P. N. (2009). Reconstructing spatiotemporal gene expression data from partial observations. *Bioinformatics* 25, 2581–2587.

Caussin, C., Fleurat-Lessard, P., and Bonnemain, J. L. (1983). Absorption of some amino acids by sporophytes isolated from *Polytrichum formosum* and ultrastructural characteristics of the haustorium transfer cells. *Ann. Bot.* 51, 167–173.

Chang, A. B., Lin, R., Studley, W. K., Tran, C. V., and Saier, M. H. Jr. (2004). Phylogeny as a guide to structure and function of membrane transport proteins. *Mol. Membr. Biol.* 21, 171–181.

Chen, L., and Bush, D. R. (1997). LHT1, a lysine- and histidine-specific amino acid transporter in *Arabidopsis*. *Plant Physiol.* 115, 1127–1134.

Cho, B. H., and Komor E. (1983). Mechanism of proline uptake by *Chlorella vulgaris*. *Biochim. Biophys. Acta* 735, 361–366.

Cho, B. H., and Komor, E. (1985). The amino acid transport systems of the autotrophically grown green alga *Chlorella*. *Biochim. Biophys. Acta* 821, 384–392.

Cho, B. H., Sauer, N., Komor, E., and Tanner, W. (1981). Glucose induces two amino acid transport systems in *Chlorella*. *Proc. Natl. Acad. Sci. U.S.A.* 78, 3591–3594.

Edgar, R. C. (2004). MUSCLE multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Finet, C., Timme, R. E., Delwiche, C. F., and Marlétaz, F. (2010). Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Curr. Biol.* 20, 2217–2222.

Fischer, W. N., Kwart, M., Hummel, S., and Frommer, W. B. (1995). Substrate specificity and expression profile of amino acid transporters (AAPs) in *Arabidopsis*. *J. Biol. Chem.* 270, 16315–16320.

Fischer, W. N., Loo, D. D. F., Koch, W., Ludewig, U., Boorer, K. J., Tegeder, M., Rentsch, D., Wright, E. M., and Frommer, W. B. (2002). Low and high affinity amino acid H^+-cotransporters for cellular import of neutral and charged amino acids. *Plant J.* 29, 717–731.

Flynn, K. J., and Butler, I. (1986). Nitrogen sources for the growth of marine microalgae: role of dissolved free amino acids. *Mar. Ecol. Prog. Ser.* 34, 281–304.

Foster, J., Lee, Y. H., and Tegeder, M. (2008). Distinct expression of members of the LHT amino acid transporter family in flowers indicates specific roles in plant reproduction. *Sex. Plant Reprod.* 21, 143–152.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.

Hirner, A., Ladwig, F., Stransky, H., Okumoto, S., Keinath, M., Harms, A., Frommer, W. B., and Koch, W. (2006). *Arabidopsis* LHT1 is a high-affinity transporter for cellular amino acid uptake in both root epidermis and leaf mesophyll. *Plant Cell* 18, 1931–1946.

Hirner, B., Fischer, W. N., Rentsch, D., Kwart, M., and Frommer, W. B. (1998). Developmental control of H^+/amino acid permease gene expression during seed development of *Arabidopsis*. *Plant J.* 14, 535–544.

Hunt, E., Gattolin, S., Newbury, H. J., Bale, J. S., Tseng, H. M., Barrett, D. A., and Pritchard, J. (2010). A mutation in amino acid permease AAP6 reduces the amino acid content of the *Arabidopsis* sieve elements but leaves aphid herbivores unaffected. *J. Exp. Bot.* 61, 55–64.

Kapraun, D. F. (2007). Nuclear DNA content estimates in green algal lineages: Chlorophyta and Streptophyta. *Ann. Bot.* 99, 677–701.

Karol, K. G., McCourt, R. M., Cimino, M. T., and Delwiche, C. F. (2001). The closest living relatives of land plants. *Science* 294, 2351–2353.

Kato, Y., Ueno, S., and Imamura, N. (2006). Studies on the nitrogen utilization of endosymbiotic algae isolated from Japanese *Paramecium bursaria*. *Plant Sci.* 170, 481–486.

Kirk, D. L., and Kirk, M. M. (1978a). Carrier-mediated uptake of arginine and urea by *Chlamydomonas reinhardtii*. *Plant Physiol.* 61, 556–560.

Kirk, M. M., and Kirk, D. L. (1978b). Carrier-mediated uptake of arginine and urea by *Volvox carteri f. nagariensis*. *Plant Physiol.* 61, 549–555

Koch, W., Kwart, M., Laubner, M., Heineke, D., Stransky, H., Frommer, W. B., and Tegeder, M. (2003). Reduced amino acid content in transgenic potato tubers due to antisense inhibition of the leaf H^+/amino acid symporter StAAP1. *Plant J.* 33, 211–220.

Lee, Y. H., Foster, J., Chen, J., Voll, L. M., Weber, A. P. M., and Tegeder, M. (2007). AAP1 transports uncharged amino acids into roots of *Arabidopsis*. *Plant J.* 50, 305–319.

Lee, Y. H., and Tegeder, M. (2004). Selective expression of a novel high-affinity transport system for acidic and neutral amino acids in the tapetum cells of *Arabidopsis* flowers. *Plant J.* 40, 60–74.

Liu, X., and Bush, D. R. (2006). Expression and transcriptional regulation of amino acid transporters in plants. *Amino Acids* 30, 113-–120.

Magallon, S., and Sanderson, M. J. (2002). Relationships among seed plants inferred from highly conserved genes: sorting conflicting phylogenetic signals among ancient lineage. *Am. J. Bot.* 89, 1991–2006.

Mathews, S. (2009). Phylogenetic relationships among seed plants: persistent questions and the limits of DNA sequence data. *Am. J. Bot.* 96, 228–236.

Okumoto, S., Koch, W., Tegeder, M., Fischer, W. N., Biehl, A., Leister, D., Stierhof, Y. D., and Frommer, W. B. (2004). Root phloem-specific expression of the plasma membrane amino acid proton co-transporter AAP3. *J. Exp. Bot.* 55, 2155–2168.

Okumoto, S., Schmidt, R., Tegeder, M., Fischer, W. N., Rentsch, D., Frommer, W. B., and Koch, W. (2002). High affinity amino acid transporters specifically expressed in xylem parenchyma and developing seeds of *Arabidopsis*. *J. Biol. Chem.* 277, 45338–45346.

Ortiz-Lopez, A., Chang, H. C., and Bush, D. R. (2000). Amino acid transporters in plants. *Biochim. Biophys. Acta* 1465, 275–280.

Pate, J. S., Sharkey, P. J., and Atkins, C. A. (1977). Nutrition of a developing legume fruit: functional economy in terms of carbon, nitrogen, water. *Plant Physiol.* 59, 506–510.

Pate, J. S., Sharkey, P. J., and Lewis, O. A. M. (1975). Xylem to phloem transfer of solutes in fruiting shoots of legumes, studied by a phloem bleeding technique. *Planta* 122, 11–26.

Prigge, M. J., and Bezanilla, M. (2010). Evolutionary crossroads in developmental biology: *Physcomitrella patens*. *Development* 137, 3535–3543.

Reinhart, D. A., and Thomas, R. J. (1981). Sucrose uptake and transport in conducting cells of *Polytrichum commune*. *Bryologist* 84, 59–64.

Rennie, E. A., and Turgeon, R. (2009). A comprehensive picture of phloem loading strategies. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14162–14167

Rentsch, D., Schmidt, S., and Tegeder, M. (2007). Transporters for uptake and allocation of organic nitrogen compounds in plants. *FEBS Lett.* 581, 2281–2289.

Sanders, A., Collier, R., Trethewy, A., Gould, G., Sieker, R., and Tegeder, M. (2009). AAP1 regulates import of amino acids into developing *Arabidopsis* embryos. *Plant J.* 59, 540–552.

Sauer, N., Komor, E., and Tanner, W. (1983). Regulation and characterization of two inducible amino-acid transport systems in *Chlorella vulgaris*. *Planta* 159, 404–410.

Schmidt, R., Stransky, H., and Koch, W. (2007). The amino acid permease AAP8 is important for early seed development in *Arabidopsis thaliana*. *Planta* 226, 805–813.

Schulze, W., Frommer, W. B., and Ward, J. M. (1999). Transporters for ammonium, amino acids and peptides are expressed in pitchers of the carnivorous plant *Nepenthes*. *Plant J.* 17, 637–646.

Shehawy, R. M., and Kleiner, D. (2001). "Nitrogen limitation," in *Algal Adaptation to Environmental Stresses-Physiological, Biochemical and Molecular Mechanisms*, eds L. C. Rai and J. P. Gaur (Berlin: Springer), 45–64.

Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.

Svennerstam, H., Ganeteg, U., Bellini, C., and Näsholm, T. (2007). Comprehensive screening of *Arabidopsis* mutants suggests the lysine histidine transporter 1 to be involved in plant uptake of amino acids. *Plant Physiol.* 143, 1853–1860.

Svennerstam, H., Ganeteg, U., and Näsholm, T. (2008). Root uptake of cationic amino acids by *Arabidopsis* depends on functional expression of amino acid permease. *New Phytol.* 180, 620–630.

Tan, Q. M., Grennan, A. K., Pelissier, H. C., Rentsch, D., and Tegeder, M. (2008). Characterization and expression of French bean amino acid transporter *PvAAP1*. *Plant Sci.* 174, 348–356.

Tegeder, M., and Rentsch, D. (2010). Uptake and partitioning of amino acids and peptides. *Mol. Plant* 3, 997–1011.

Tegeder, M., Rentsch, D., and Patrick, J. W. (2011). "Organic carbon and nitrogen transporters," in *Plant Plasma Membrane: Plant Cell Monographs*, eds A. Murphy, W. Peer, and B. Schulz (Berlin: Springer), 331–352.

Tegeder, M., Tan, Q., Grennan, A. K., and Patrick, J. W. (2007). Amino acid transporter expression and localisation studies in pea (*Pisum sativum*). *Funct. Plant Biol.* 34, 1019–1028.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (2007). The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.

Trachtenberg, S., and Zamski, E. (1978). Conduction of ionic solutes and assimilates in the leptom of *Polytrichum juniperinum* Willd. *J. Exp. Bot.* 29, 719–727.

Turmel, M., Lemieux, C., Burger, G., Lang, B. F., Otis, C., Plante, I., and Gray, M. W. (1999). The complete mitochondrial DNA sequences of *Nephoroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within green algae. *Plant Cell* 11, 1717–1730.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G., and Provart, N. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2, e718. doi:10.1371/journal.pone.0000718

Wipf, D., Ludewig, U., Tegeder, M., Rentsch, D., Koch, W., and Frommer, W. B. (2002). Conservation of amino acid transporters in fungi, plants and animals. *Trends Biochem. Sci.* 27, 139–147.

Zhang, C. Q., Barthelson, R. A., Lambert, G. M., and Galbraith, D. W. (2008). Global characterization of cell-specific gene expression through fluorescence-activated sorting of nuclei. *Plant Physiol.* 147, 300–340.

Zhang, L., Tan, Q., Lee, R., Trethewy, A., Lee, Y.-H., and Tegeder, M. (2010). Altered xylem-phloem transfer of amino acids affects metabolism and leads to increased seed yield and oil content in *Arabidopsis*. *Plant Cell* 22, 3603–3620.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 October 2011; accepted: 19 January 2012; published online: 13 February 2012.

Citation: Tegeder M and Ward JM (2012) Molecular evolution of plant AAP and LHT amino acid transporters. Front. Plant Sci. ***3****:21. doi: 10.3389/fpls.2012.00021*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 10 May 2012
doi: 10.3389/fpls.2012.00078

An expanding role for purine uptake permease-like transporters in plant secondary metabolism

John G. Jelesko*

Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg, VA, USA

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Roland Krause, Max Planck Institute of Molecular Cell Biology and Genetics, Germany
Marcelo Desimone, Instituto Multidisciplinario de Biología Vegetal, Argentina

***Correspondence:**
John G. Jelesko, Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, 548 Latham Hall, AgQuad Drive, Blacksburg, VA 24061-0390, USA.
e-mail: jelesko@vt.edu

For the past decade, our understanding of the plant purine uptake permease (PUP) transporter family was primarily oriented on purine nucleobase substrates and their tissue-specific expression patterns in *Arabidopsis*. However, a tobacco PUP-like homolog demonstrating nicotine uptake permease activity was recently shown to affect both nicotine metabolism and root cell growth. These new findings expand the physiological role for PUP-like transporters to include plant secondary metabolism. Molecular evolution analyses of PUP-like transporters indicate they are distinct group within an ancient super family of drug and metabolite transporters (DMTs). The PUP-like family originated during terrestrial plant evolution sometime between the bryophytes and the lycophytes. A phylogenetic analysis indicates that the PUP-like transporters were likely derived from a pre-existing nucleotide-sugar transporter family within the DMT super family. Within the lycophyte *Selaginella*, there are three paralogous groups of PUP-like transporters. One of the three PUP-like paralogous groups showed an extensive pattern of gene duplication and diversification within the angiosperm lineage, whereas the more ancestral PUP-like paralogous groups did not. Biochemical characterization of four closely related PUP-like paralogs together with model-based phylogenetic analyses indicate both subfunctionalization and neofunctionalization during the molecular evolution of angiosperm PUP-like transporters. These findings suggest that members of the PUP-like family of DMT transporters are likely involved in diverse primary and secondary plant metabolic pathways.

Keywords: adenine, alkaloid, evolution, nicotine, transport

INTRODUCTION

The purpose of this mini review is to integrate recent molecular and phylogenetic findings about a tobacco nicotine uptake permease (NUP) with prior knowledge about the closely related plant-specific purine uptake permease (PUP) transporter family. The founding member of the PUP-like transporter family was *Arabidopsis thaliana* PUP1 (AtPUP1) that was identified in a cDNA complementation screen of a yeast adenine uptake mutant (Gillissen et al., 2000). Yeast expressing the *AtPUP1* cDNA show high affinity adenine uptake activity that is abrogated by protonophore-inhibitors, indicating a proton-facilitated substrate symport mechanism. This adenine uptake activity is also efficiently competed by structurally-related purine ring-containing cytokinins and caffeine. Nicotine also reduces AtPUP1-mediated adenine uptake rates, though considerably less efficiently than the above purine compounds.

In *Arabidopsis*, the *PUP*-like genes comprise a moderate size gene family. There are 21 *PUP*-like genes in *Arabidopsis*, including one pseudogene. PUP-like transporters typically have 10 predicted transmembrane spanning domains. To date, only three *Arabidopsis PUP*-like genes have been the focus of detailed studies into expression profiles and substrate specificity (Gillissen et al., 2000; Burkle et al., 2003). This limited characterization of the *Arabidopsis* PUP-like transporter family nevertheless demonstrates evolutionary patterns of both subfunctionalization and neofunctionalization. In a parsimony phylogenetic analysis comprised of only *Arabidopsis* PUP-like proteins, AtPUP1 and AtPUP2 showed closest sequence identity (64%) and phylogenetic association to each other, than to any other AtPUP transporter. Both AtPUP1 and AtPUP2 transport adenine and cytokinins, albeit with different kinetics. *AtPUP1* gene expression localizes to leaf hydathode tissue (specifically epithem cells) and the stigmatic surface, whereas *AtPUP2* localizes to vascular tissue, specifically to the phloem. Therefore, AtPUP1 and AtPUP2 have diversified in terms of tissue specificity, but not in terms of ligand specificity. Thus, AtPUP1 and AtPUP2 display a pattern of subfunctionalization (Force et al., 1999). In contrast, AtPUP3 is more distantly related to AtPUP1 (29% protein identity) and *AtPUP3* gene expression is restricted to pollen. Heterologously expressed recombinant AtPUP3 does not transport adenine during the same conditions as AtPUP1 and AtPUP2. AtPUP3 is presumed to transport another, as yet to be determined, substrate into pollen cells. Thus, in terms of both cell type-specific expression and substrate recognition, AtPUP3 exhibits patterns of neofunctionalization relative to AtPUP1 and AtPUP2. AtPUP1 is presumed to be involved in purine and/or cytokinin uptake during leaf guttation and at the stigma, whereas AtPUP2 is postulated to be involved with uptake of these substrates into the phloem. These predicted physiological roles await confirmation using plants with altered *AtPUP1* or *AtPUP2* expression levels. Given these patterns of both subfunctionalization and neofunctionalization, together with the relatively low degree of sequenced

identity between these *Arabidopsis* transporters, ranging from 16% to 64% identity, it is likely that other PUP-like transporters recognize different substrates involved in either primary or secondary metabolism.

A PUP-LIKE TRANSPORTER NUP1 AFFECTS PYRIDINE ALKALOID METABOLISM AND PHYSIOLOGY

The physiological impact of PUP-like transporters was recently expanded to include plant secondary metabolism. Transcriptional profiling of a tobacco mutant affecting total alkaloid accumulation levels and nicotine biosynthetic gene expression levels identified transcripts in roots that are coordinately regulated with several nicotine biosynthetic genes (Kidd et al., 2006). One such cDNA fragment encodes a predicted peptide with 56% identity and 67% similarity to AtPUP1. The corresponding full-length *Nicotiana tabacum* cDNA encodes a protein with nicotine uptake permease activity, called NUP1 (Hildreth et al., 2011). In contrast to AtPUP1, NUP1-mediated nicotine uptake activity is not efficiently competed by either purines or cytokinins, neither it is inhibited by closely-related pyridine alkaloids nor less-related tropane alkaloids. Therefore, NUP1 shows a high degree of substrate specificity for nicotine and is not an ortholog of AtPUP1. A NUP1–GFP fusion localizes primarily to the tobacco plasma membrane, suggesting NUP1 transports apoplastic nicotine into the cytoplasm. Steady state *NUP1* mRNA levels accumulate to highest levels at root tips, where nicotine biosynthesis is localized (Dawson, 1942a,b; Baldwin, 1988).

The association of NUP1 with nicotine metabolism was confirmed using tobacco lines with reduced steady state *NUP1* mRNA levels. Transgenic *NUP1-RNAi* tobacco plants have reduced foliar nicotine levels (Hildreth et al., 2011). However, the lower foliar nicotine accumulation levels are not due to altered nicotine transport from roots to shoots *per se*, because *NUP1-RNAi* plants are not compromised in their ability to proportionally transport exogenous nicotine fed to roots into leaves. Rather, less foliar nicotine is correlated with lower nicotine levels in the roots. During non-induced conditions *NUP1-RNAi* hairy root lines have lower nicotine levels in the hairy roots. Thus, reduced *NUP1* transcript levels correlate with significant reductions in nicotine accumulation levels in leaves, roots, and hairy roots. During non-inducing culture conditions the *NUP1-RNAi* hairy root cultures showed significantly more nicotine accumulation in the spent culture media, compared to wild type. It was not determined whether the increased nicotine in the culture media of *NUP1-RNAi* lines was due to nicotine released from the open xylem vessels at the ends of the hairy roots, nicotine release from the root epidermis, or diminished nicotine re-uptake from the rhizosphere. While there were consistent patterns of decreased nicotine accumulation levels in *NUP1-RNAi* roots, during several conditions examined there was an overall poor correlation of reduced nicotine levels with corresponding reductions the steady state transcript levels of several essential nicotine biosynthetic genes. Thus, while *NUP1* expression levels clearly affect overall nicotine metabolism, the molecular mechanism by which NUP1 affects nicotine accumulation levels is currently poorly understood.

Nicotine biosynthesis is one of a few plant specialized metabolic pathways that are integrated with ongoing root growth. On one hand, root growth is essential for net nicotine accumulation levels (Solt, 1957; Baldwin, 1988). On the other hand, tobacco roots are not immune to the toxic effects of either exogenous nicotine treatment (Baldwin and Callahan, 1993; Baldwin and Ohnmeiss, 1994; Shoji et al., 2009) or the stimulation of endogenous nicotine biosynthesis (Hildreth et al., 2011). Both conditions result in reduced wild type root growth. This indicates there are complex interactions between root growth and nicotine metabolism/accumulation that could have homeostatic properties. Another interesting phenotype of *NUP1*-RNAi lines is they exhibit increased seedling root elongation rates during both control conditions and several treatments that increase overall nicotine accumulation levels. This latter result highlights the current lack of knowledge in the field about the molecular mechanisms of how nicotine accumulation levels are coordinated with root growth. Therefore, it is noteworthy that NUP1 is the first molecular component demonstrated to affect both root growth and nicotine metabolism.

NUP1 may also have implications in nicotine physiology far beyond the root tip. Whole plant nicotine levels are held at allometric set points relative to total nicotine-free plant biomass (Ohnmeiss and Baldwin, 1994; Baldwin and Karb, 1995). How allometric control of whole plant nicotine levels is enforced, when the site of nicotine biosynthesis is restricted to growing root tips yet most nicotine is unequally distributed throughout the plant, is currently an enigma. Clearly, simple models of metabolic control based upon homeostatic apoplastic nicotine levels cannot account for the observed whole plant nicotine allometry. Perhaps allometric control of nicotine biosynthesis is enforced by the integration of nicotine biosynthesis with root tip growth. An appealing aspect of this model for whole plant nicotine allometric control is that growing root tips are a classic sink tissue whose growth is a function of shoot primary productivity that in turn is scaled to shoot biomass. If this is the case, NUP1 has several properties that make it a good candidate for participation in whole plant nicotine allometric set points. NUP1 presumably transports apoplastic nicotine into cells near the root tip, reduced *NUP1* expression levels results in reduced root nicotine accumulation levels, and concomitant increased root elongation rates. The remainder of this article will focus on the molecular evolution of the PUP-like transporter family.

PUP-LIKE TRANSPORTERS ORIGINATED EARLY DURING TERRESTRIAL PLANT EVOLUTION, POSSIBLY FROM A DMT UDP-SUGAR TRANSPORTER FAMILY

Two previous molecular evolution studies indicated that PUP-like homologs were restricted to plant taxa (Gillissen et al., 2000; Hildreth et al., 2011). Another study, based upon reciprocal BLASTP searches, assigned the PUP transporters to an ancient drug and metabolite transporter (DMT) super family (Jack et al., 2001). The plant PUP-like transporters comprise a distinct family within the DMT super family named the plant organocation permease (POP) family and were given the transporter classification (TC) code designation 2.A.7.14. A NUP1 BLAST query on the NCBI Conserved Domain Database (CDD) identified a conserved pfam 03151 domain (BLASTP expectation value = 7.1 e^{-28}). On CDD the pfam 03151 domain is associated with the EamA

super family, citing Jack et al. (2001) as the primary reference for the EamA super family, despite the said article does not use the EamA nomenclature. This highlights the considerable diversity in nomenclature associated with PUP-like transporters, both in the literature and promulgated in electronic databases. Henceforth, this mini review will use both the International Union of Biochemistry and Molecular Biology-approved TC coding (Jack et al., 2001; Saier Jr. et al., 2009) and the PUP terminology to discuss the molecular evolution of the PUP/2.A.7.14 family of transporters.

Published PUP and NUP similarity searches and phylogenies are consistent with the PUP/2.A.7.14 family originating in terrestrial plants after the bryophytes (Gillissen et al., 2000; Hildreth et al., 2011). However, it is improbable that a progenitor PUP/2.A.7.14 transporter formed entirely *de novo* in plants. Rather, it is more likely that it was derived from a different DMT family pre-existing in the plantae. Jack et al. (2001) used the program GAP (Devereux et al., 1984) to assign binary affinity of PUP/2.A.7.14 family to the DMT 2.A.7.3 family of prokaryotic DMT transporters. This binary association barely met the minimum GAP score SD value cutoff of 9, and this purported association has not been independently corroborated by phylogenetic analyses. Since a BLASTP e-value cutoff of e^{-4} identified only plant PUP/2.A.7.14 homologs (Hildreth et al., 2011), a new PUP/2.A.7.14 phylogenetic analysis was performed that used the transitive nature of protein homology sensu stricto to identify distantly related DMT homologs. Briefly, three PUP-like homologs representing three well-supported clades in *Selaginella* were used as queries in the SSEARCH program on a local NCBI RefSeq database. Proteins returned by at least two searches established homology, regardless of expectation value. A total of 185 proteins meeting this criterion were aligned and subjected to four independent Bayesian runs. The resulting joint Bayesian phylogeny is displayed in **Figure 1**. Again, no PUP/2.A.7.14 homologs were identified in bryophyte or algal taxa, further supporting the hypothesis that the PUP/2.A.7.14 family originated during terrestrial plant evolution. Plant PUP/2.A.7.14 transporters formed a well-supported monophyletic group (posterior probability equal to 1.0). Because this phylogeny contained eukaryotic taxa other than plants, it effectively rooted the monophyletic PUP/2.A.7.14 clade and established PUP/2.A.7.14 paralogous relationships within the plantae. The most basal plant PUP/2.A.7.14 clade included both *Selaginella* and angiosperm homologs and may represent the earliest PUP/2.A.7.14 orthologous group. The large PUP/2.A.7.14 clade was rooted within a tritomy of other eukaryotic DMT transporters, which precludes establishing the relative branch order. Nevertheless, 78% of these eukaryotic DMT transporters were most similar to 2.A.7.12 family members that are presumed to transport nucleotide–sugar conjugates. The next more basal eukaryotic clade was likewise comprised of 2.A.7.12 family members and it also defined the bacteria – eukaryotic split with strong Bayesian support. Eighty-six percent of the prokaryotic DMT transporters belonged to the 2.A.7.3 family. This is the same family identified in the binary association analysis reported in the initial characterization of the DMT super family (Jack et al., 2001). Given the abundance of 2.A.7.12 family members in both the basal and sister groups to the monophyletic plant PUP/2.A.7.14 transporters, it is likely that the PUP/2.A.7.14 family arose from a pre-existing 2.A.7.12 family of nucleotide-sugar transporters sometime after the bryophytes, and was well established by the lycophytes.

There is asymmetry in the pattern of PUP/2.A.7.14 transporter evolution within the plantae. Of the three paralogous PUP/2.A.7.14 groups in *Selaginella*: one group is restricted only to *Selaginella*, one group has just one or two homologs in diverse taxa, whereas the third paralogous group comprises a sister group to an extensive pattern of gene duplication during angiosperm evolution. The widespread diversification of the PUP/2.A.7.14 gene family in the angiosperms likely promoted increased complexity in both plant development and plant chemical diversity. It follows that the phylogenetically more derived homologs may be involved with correspondingly more derived plant traits/characters. Consistent with this idea, the phylogenetically more derived AtPUP1, AtPUP2, APUP3 transporters show a pattern of subfunctionalization with respect to their restricted cell-type expression patterns in *Arabidopsis*, whereas AtPUP3 also shows neofunctionalization in substrate specificity. Similarly, tobacco NUP1 shows neofunctionalization with respect to both tissue localization and substrate specificity. These evolutionary patterns suggests that PUP/2.A.7.14 transporters were important for the evolution of highly specialized plant tissue/cell types. Assuming the patterns of subfunctionalization and neofunctionalization observed with these four characterized PUP/2.A.7.14 members are typical for this transporter family, then the other phylogenetically highly-derived PUP/2.A.7.14 family members shown in **Figure 1** are expected to show considerable diversity in both tissue-specific expression and/or ligand specificity/selectivity.

The discovery that NUP1 affects pyridine alkaloid metabolism has implications for other plant alkaloid biosynthetic pathways. Both tropane and monoterpenoid indole alkaloid (MIA) biosynthetic pathways are cell non-autonomous processes, in which specific biosynthetic enzymes are expressed in non-adjacent specialized cell types (reviewed in Ziegler and Facchini, 2008). This means that alkaloid intermediates must be transported both out of and into dispersed cells. Despite the recognized importance of these inferred transport processes during tropane and MIA biosynthesis, specific transporters mediating these processes are a conspicuous gap in current knowledge. While it is currently uncertain if nicotine biosynthesis is a cell autonomous process, NUP1 nevertheless provides an important precedent that a member of the PUP/2.A.7.14 transporter family can impact the metabolism of at least one plant alkaloid. This may have important implications for the biosynthesis other plant alkaloids for which intercellular transport of alkaloid intermediates is believed to be an essential process.

OUTLOOK FOR PUP-LIKE TRANSPORTER RESEARCH

The discovery that NUP1 is involved in secondary metabolism significantly broadens the scope of the under-investigated PUP/2.A.7.14 family of DMT transporters. The molecular

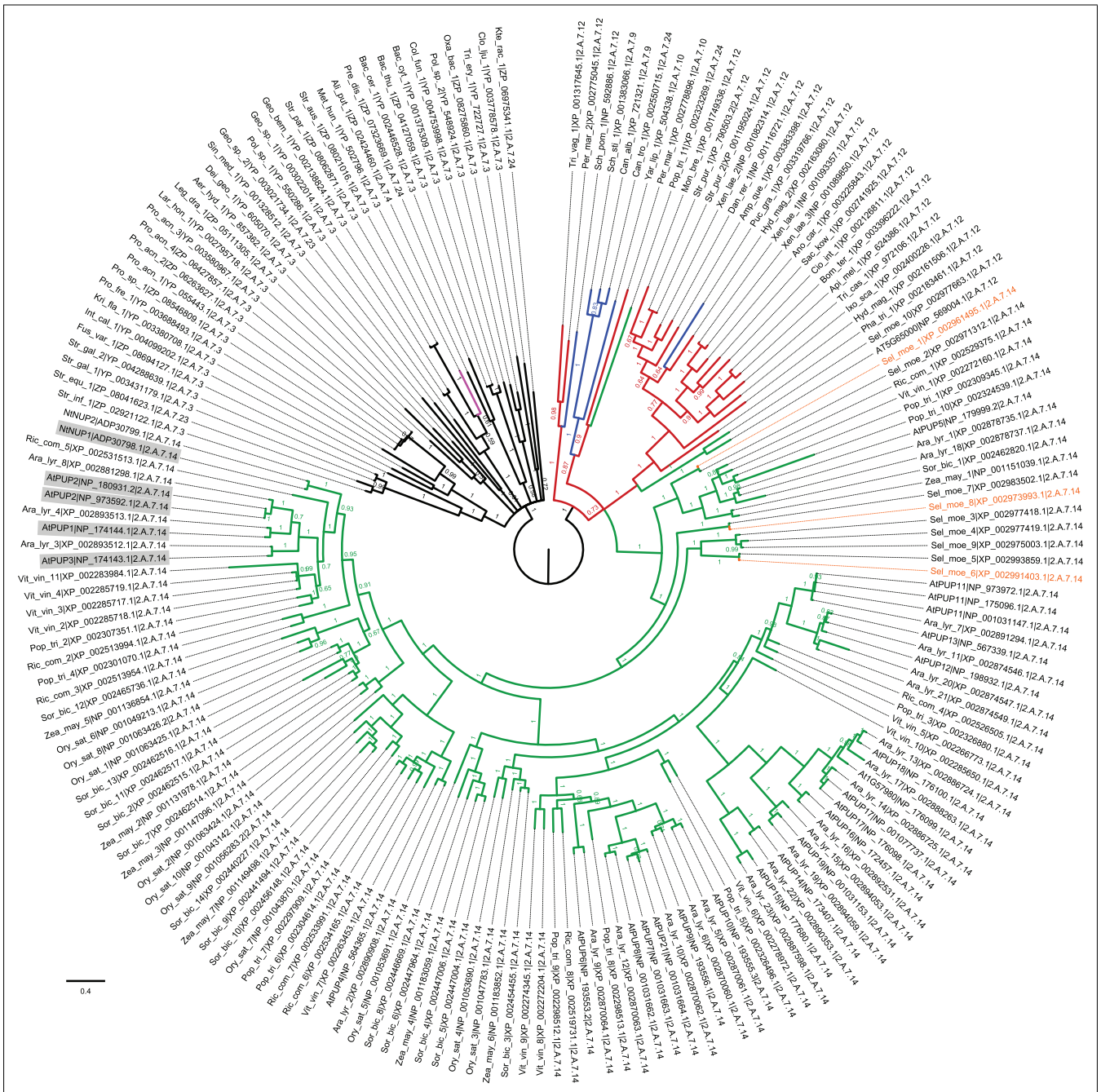


FIGURE 1 | Plant PUP-like transporters are closely related to the DMT 2.A.7.12 family of nucleotide-sugar transporters. Bayesian phylogram of DMT homologs related to the PUP/2.A.7.14 family. Taxonomic coding was as follows: black, bacteria; pink, archaea; red, diverse eukaryota; green, plantae; and blue, fungi. Orange indicates three *Selaginella* homolog queries used in SSEARCH. *Arabidopsis* and tobacco PUP-like transporters were labeled by their published protein names, whereas other PUP-like homologs were identified by a three letter abbreviated Genus_species_serial number|GenBank accession|TC code. TC codes were assigned to each transporter based upon the TC code of the optimal BLASTP homolog identified at http://www.tcdb.org. Biochemically characterized PUP-like transporters were highlighted in gray. The tree was rooted on the node separating prokaryotes from eukaryotes. *Methods*: Homology was defined by three *Selaginella* query sequences (orange) using SSEARCH (Pearson, 1996) with the BLOSUM 50 matrix on a local copy of the NCBI RefSeq database. Proteins identified in any two of the three queries defined homology sensu stricto. These were combined into a non-redundant dataset that was aligned using MAFFT (Katoh et al., 2005) and analyzed using MrBayes ver. 3.1.2 (Ronquist and Huelsenbeck, 2003) implementing a LG + G + F model (determined by ProtTest; Darriba et al., 2011) with four independent runs each with eight heated chains (temp 0.1) for 10 million generations, and a 2.5 million generation burn-in. Convergence diagnostics were evaluated using Tracer (Rambaut and Drummond, 2003). Skyline plots of all parameters showed stable distributions of posterior parameter estimates. The effective sample sizes (ESS) of the LnL values ranged from 837 to 948, and the ESS values for the TL parameter ranged from 450 to 626. Therefore, the four runs did not show evidence of non-convergence, so a joint analysis was performed. The resulting consensus tree was displayed and annotated using Figtree (Rambaut, 2009).

evolution of the PUP/2.A.7.14 family in angiosperms combined with the narrow substrate specificity/selectivity profiles of the few biochemically characterized homologs support the hypothesis that other PUP-like transporters are likely involved in other highly specialized metabolic pathways. Characterization of additional PUP/2.A.7.14 homologs will be required to determine if this hypothesis is well-supported or not. Toward that end, there are many PUP/2.A.7.14 transporters that await investigation.

ACKNOWLEDGMENTS

Alex Weisberg provided scripting to reformat tree taxa labels and Sarah Kashanian assigned TC codes to tree taxa labels. Dr. Angus Murphy provided helpful comments during the early drafting of the manuscript. I am also grateful for the helpful suggestions of the reviewers. This work was supported by a USDA award administered through the Virginia Agriculture and Experiment Station Biodesign and Bioprocessing Research Center to John G. Jelesko.

REFERENCES

Baldwin, I. T. (1988). Damage-induced alkaloids in tobacco: pot-bound plants are not inducible. *J. Chem. Ecol.* 14, 1113–1120.

Baldwin, I. T., and Callahan, P. (1993). Autotoxicity and chemical defense: nicotine accumulation and carbon gain in solanaceous plants. *Oecologia* 94, 534–541.

Baldwin, I. T., and Karb, M. J. (1995). Plasticity in allocation of nicotine to reproductive parts in *Nicotiana attenuata*. *J. Chem. Ecol.* 21, 897–909.

Baldwin, I. T., and Ohnmeiss, T. E. (1994). Swords into plowshares? *Nicotiana sylvestris* does not use nicotine as a nitrogen source under nitrogen-limited growth. *Oecologia* 98, 385–392.

Burkle, L., Cedzich, A., Dopke, C., Stransky, H., Okumoto, S., Gillissen, B., Kuhn, C., and Frommer, W. B. (2003). Transport of cytokinins mediated by purine transporters of the PUP family expressed in phloem, hydathodes, and pollen of *Arabidopsis*. *Plant J.* 34, 13–26.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164–1165.

Dawson, R. F. (1942a). Accumulation of nicotine in reciprocal grafts of tomato and tobacco. *Am. J. Bot.* 29, 66–71.

Dawson, R. F. (1942b). Nicotine synthesis in excised tobacco roots. *Am. J. Bot.* 29, 813–815.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12, 387–395.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545.

Gillissen, B., Burkle, L., Andre, B., Kuhn, C., Rentsch, D., Brandl, B., and Frommer, W. B. (2000). A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in *Arabidopsis*. *Plant Cell* 12, 291–300.

Hildreth, S. B., Gehman, E. A., Yang, H., Lu, R.-H., Ritesh, K. C., Harich, K. C., Yu, S., Lin, J., Sandoe, J. L., Okumoto, S., Murphy, A. S., and Jelesko, J. G. (2011). Tobacco nicotine uptake permease (NUP1) affects alkaloid metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18179–18184.

Jack, D. L., Yang, N. M., and H. Saier, M. (2001). The drug/metabolite transporter superfamily. *Eur. J. Biochem.* 268, 3620–3639.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 511–518.

Kidd, S. K., Melillo, A. A., Lu, R. H., Reed, D. G., Kuno, N., Uchida, K., Furuya, M., and Jelesko, J. G. (2006). The *A* and *B* loci in tobacco regulate a network of stress response genes, few of which are associated with nicotine biosynthesis. *Plant Mol. Biol.* 60, 699–716.

Ohnmeiss, T. E., and Baldwin, I. T. (1994). The allometry of nitrogen allocation to growth and an inducible defense under nitrogen-limited growth. *Ecology* 75, 995–1002.

Pearson, W. R. (1996). Effective protein sequence comparison. *Methods Enzymol.* 266, 227–258.

Rambaut, A. (2009). *FigTree*, ver. 1.2.2. Available at: http://tree.bio.ed.ac.uk/software/figtree

Rambaut, A., and Drummond, A. J. (2003). *Tracer*, ver. 1.5. Available at: http://tree.bio.ed.ac.uk/software/tracer

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.

Saier, M. H. Jr., Yen, M. R., Noto, K., Tamang, D. G., and Elkan, C. (2009). The Transporter Classification Database: recent advances. *Nucleic Acids Res.* 37, D274–D278.

Shoji, T., Inai, K., Yazaki, Y., Sato, Y., Takase, H., Shitan, N., Yazaki, K., Goto, Y., Toyooka, K., Matsuoka, K., and Hashimoto, T. (2009). Multidrug and toxic compound extrusion-type transporters implicated in vacuolar sequestration of nicotine in tobacco roots. *Plant Physiol.* 149, 708–718.

Solt, M. L. (1957). Nicotine production and growth of excised tobacco root cultures. *Plant Physiol.* 32, 480–484.

Ziegler, J., and Facchini, P. J. (2008). Alkaloid biosynthesis: metabolism and trafficking. *Annu. Rev. Plant Biol.* 59, 735–769.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 December 2011; accepted: 09 April 2012; published online: 10 May 2012.

*Citation: Jelesko JG (2012) An expanding role for purine uptake permease-like transporters in plant secondary metabolism. Front. Plant Sci. **3**:78. doi: 10.3389/fpls.2012.00078*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 08 February 2012
doi: 10.3389/fpls.2012.00016

The maize *PIN* gene family of auxin transporters

Cristian Forestan, Silvia Farinati and Serena Varotto*

Department of Agronomy, Food, Natural Resources, Animal and Environment, University of Padova, Legnaro, Italy

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
June M. Kwak, University of Maryland, USA
Christian Luschnig, University of Natural Resources and Life Sciences, Austria

****Correspondence:***
Serena Varotto, Department of Agronomy, Food, Natural Resources, Animal and Environment DAFNAE, University of Padova, Viale dell'Università 16, 35020 Legnaro, Padova, Italy.
e-mail: serena.varotto@unipd.it

Auxin is a key regulator of plant development and its differential distribution in plant tissues, established by a polar cell to cell transport, can trigger a wide range of developmental processes. A few members of the two families of auxin efflux transport proteins, PIN-formed (PIN) and P-glycoprotein (ABCB/PGP), have so far been characterized in maize. Nine new *Zea mays* auxin efflux carriers *PIN* family members and two maize *PIN-like* genes have now been identified. Four members of *PIN1* (named *ZmPIN1a–d*) cluster, one gene homologous to *AtPIN2* (*ZmPIN2*), three orthologs of *PIN5* (*ZmPIN5a–c*), one gene paired with *AtPIN8* (*ZmPIN8*), and three monocot-specific *PINs* (*ZmPIN9*, *ZmPIN10a*, and *ZmPIN10b*) were cloned and the phylogenetic relationships between early-land plants, monocots, and eudicots PIN proteins investigated, including the new maize PIN proteins. Tissue-specific expression patterns of the 12 maize *PIN* genes, 2 *PIN-like* genes and *ZmABCB1*, an ABCB auxin efflux carrier, were analyzed together with protein localization and auxin accumulation patterns in normal conditions and in response to drug applications. *ZmPIN* gene transcripts have overlapping expression domains in the root apex, during male and female inflorescence differentiation and kernel development. However, some *PIN* family members have specific tissue localization: *ZmPIN1d* transcript marks the L1 layer of the shoot apical meristem and inflorescence meristem during the flowering transition and the monocot-specific *ZmPIN9* is expressed in the root endodermis and pericycle. The phylogenetic and gene structure analyses together with the expression pattern of the *ZmPIN* gene family indicate that subfunctionalization of some maize PINs can be associated to the differentiation and development of monocot-specific organs and tissues and might have occurred after the divergence between dicots and monocots.

Keywords: PIN-formed, auxin efflux carriers, kernel development, inflorescences, monocots, polar auxin transport, *Zea mays*

INTRODUCTION

Auxin is a key regulator of plant development: its differential distribution in plant tissues, established by a polar transport (polar auxin transport, PAT) in response to internal and external stimuli can trigger a wide range of developmental processes, such as organ initiation at the shoot apex (Reinhardt et al., 2000, 2003; Benkova et al., 2003; Heisler et al., 2005), leaf venation (Scarpella et al., 2006, 2010), apical dominance (Leyser, 2005), tropisms (Blakeslee et al., 2004; Kimura and Kagawa, 2006; Palme et al., 2006), embryo axis formation (Friml et al., 2003; Weijers et al., 2005; De Smet and Jurgens, 2007), root architecture (Blilou et al., 2005; Dubrovsky et al., 2008), and fruit ripening (Ellis et al., 2005). Both metabolic changes and transport of auxin have been demonstrated to play a major role in plant organogenesis and differentiation (Petrasek and Friml, 2009). Auxin is synthesized in many plant tissues throughout several different pathways (Normanly, 2010; Zhao, 2010) and subjected to long-range transport via the vascular system by mass flow. Together with this long-range transport a cell to cell transport is present in the tissues, which covers both long and short distance and is mostly polar according to the chemiostatic theory (reviewed in Zazimalova et al., 2010). Several experimental evidences indicate that the control of auxin transport by carrier proteins is fundamental to patterning processes involving auxin (reviewed in Petrasek and Friml, 2009). Three main types of membrane auxin transporters have been characterized in vascular plants and are responsible for cell to cell auxin movement: the auxin permease 1 (AUX1)/LAX influx carriers (reviewed in Kerr and Bennett, 2007), the ATP-BINDING cassette subfamily B [ABCB; previously known as multidrug resistance (MDR)/Phosphoglycoprotein (PGP)] transporters (reviewed by Geisler and Murphy, 2006; Verrier et al., 2008; Yang and Murphy, 2009) and the PIN-formed (PIN) efflux carriers (reviewed by Tanaka et al., 2006; Vieten et al., 2007; Zazimalova et al., 2007; Krecek et al., 2009; Petrasek and Friml, 2009).

The two families of auxin efflux carriers are characterized by specific features. ABC is one of the largest and most ubiquitous of the transporter families and its ATP-driven transport is associated with the movements of a wide variety of small molecules, nutrients, and xenobiotics (Verrier et al., 2008). A small group of the ABCB subclass of ABC transporters, comprising ABCB1 and ABCB19 in *Arabidopsis thaliana* (Blakeslee et al., 2005), function in long-distance auxin transport and localized loading of auxin into the transport system across plant species, while they do not appear to be involved in setting vectorial auxin flows that control organogenesis (reviewed in Titapiwatanakun and Murphy, 2008). The best characterized ABCB auxin transporters in *A. thaliana*, ABCB1, ABCB4, and ABCB19, have a stable and non-polar plasma membrane localization in the cell. Particularly,

Arabidopsis ABCB1/PGP1 has been shown to function in exporting IAA from shoot and root meristematic cells into the long-distance polar auxin stream, and its homologs in maize *Brachytic2* (*BR2*) and sorghum *Dwarf3* (*DW3*) have also been characterized (Multani et al., 2003; Knoller et al., 2010; McLamore et al., 2010). Mutations in *BR2* and *DW3* genes determine several plant defects, such as altered seedling auxin transport, altered vasculature of the stalk, and reduced stalk height due to shortened lower internodes (Multani et al., 2003; Pilu et al., 2007), confirming the ABCB role in long-range auxin transport. BR2 is also expressed in epidermal and hypodermal tissues at the root apex, where it is involved on shootward auxin transport from the apex to the elongation zone (McLamore et al., 2010). Although some differences exist between *Arabidopsis abcb1* and maize *br2* phenotypes and these differences are likely to reflect differences in developmental patterns, a combination of phylogenetic, phenotypic, and physiological analyses recently showed that ABCB1/BR2 function is conserved between dicots and monocots (Knoller et al., 2010).

The plant specific *PIN* family of efflux carriers comprises integral membrane proteins (belonging to the "mem_trans" group of the PFAM database) and has been associated to PAT (Petrasek et al., 2006; Zazimalova et al., 2007; Krecek et al., 2009; Yang and Murphy, 2009). In *A. thaliana* there are eight *PIN* genes (At*PIN1–AtPIN8*) codifying for proteins that differ in the length of hydrophilic loop in the middle of their polypeptide chain (Krecek et al., 2009; Zazimalova et al., 2010; Peer et al., 2011). The long PIN proteins of *Arabidopsis* (PIN1–4 and 7) show mostly PM-localization and their polar localization determines the direction of auxin flux. PIN-mediated PAT regulates many differentiation and developmental processes that have been extensively characterized. Furthermore, long AtPINs do not reside statically in plasma membrane but constitutively cycle between the PM and endosomal compartments and their re-location can be triggered by environmental stimuli (Friml, 2010; Richter et al., 2010). Three PIN proteins of *A. thaliana*, namely PIN5, 6, and 8, have a shorter central hydrophilic domain and both PIN5 and PIN8, in which this domain is particularly short, have been shown to localize in endoplasmic reticulum, suggesting a possible role in regulating intracellular auxin homeostasis (Mravec et al., 2009; Wabnik et al., 2011). The classification of AtPIN6 in long or short PINs is rather more controversial since the hydrophilic loop is only partially reduced while the transmembrane regions show high sequence similarity with long PINs (Krecek et al., 2009; Mravec et al., 2009). In addition to the eight AtPIN proteins, the *Arabidopsis* genome encodes for seven further *PIN-like* genes. Phylogenetic analysis revealed that the two families form distinct clusters and the role of these PIN-like proteins has not yet been demonstrated (Paponov et al., 2005).

Many homologous *PIN* genes were cloned in monocot species and some members of the rice (*Oryza sativa*; Xu et al., 2005; Wang et al., 2009) and maize (*Zea mays*; Carraro et al., 2006; Forestan et al., 2010; Forestan and Varotto, in press) families have been characterized. Both specific features and homologies between monocots and *Arabidopsis* (a eudicot) PIN families have been shown. Monocot-specific features comprise both sequence clustering in phylogenetic analyses and expression pattern at transcript and protein level. In rice, the sequence analysis of the 12 *PIN* genes present in the genome showed that rice has four *PIN1* genes and one *OsPIN2*, while no OsPIN protein was grouped into the AtPIN3, AtPIN4, AtPIN7 cluster. Four *OsPIN* genes encode for rice PIN proteins with a short central hydrophilic domain: three *OsPIN5* and one *OsPIN8*. Furthermore, three OsPIN proteins appear monocot-specific: OsPIN9, OsPIN10a, and OsPIN10b. OsPIN9 has a central hydrophilic domain intermediate in length between long and short PINs of *Arabidopsis* and its expression analysis at transcription level suggests a possible function in adventitious root differentiation. OsPIN10a and OsPIN10b have a long central hydrophilic domain and based on their expression pattern they could be involved in rice tillering (Wang et al., 2009). Until now, three *PIN1* genes were identified in maize (*ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c*) and their transcript and protein localized in different tissues (Carraro et al., 2006; Forestan et al., 2010). Using an antibody raised against the AtPIN1 protein, ZmPIN1 proteins were shown to localize in both a polar and apolar way in the cell plasma membrane in some tissues or to accumulate in the endomembranes inside the cell in a specialized tissue of the endosperm (Forestan and Varotto, 2010; Forestan et al., 2010). Finally, 11 genes belonging to the *PIN* family have been identified in *Sorghum bicolor*, confirming that also in this monocot species at least three members of the family are grouped in the *AtPIN1* cluster and another three in *AtPIN5* cluster (Shen et al., 2010).

In this work we report the identification and characterization of nine new *Z. mays* PIN family members and two maize *PIN-like* genes. The phylogenetic and gene structure analyses together with the expression pattern of the members of *PIN* gene family indicate that subfunctionalization of some maize *PINs* can be associated to the differentiation and development of specific monocot organs and tissues and might have occurred after the divergence between dicots and monocots.

MATERIALS AND METHODS

MAIZE AUXIN EFFLUX CARRIERS IDENTIFICATION AND SEQUENCE ANALYSIS

The B73 maize genome sequence has recently been completed and published (Schnable et al., 2009), and annotated CDS and filtered protein have also become available (release 5b.60). Two genome browsers are available at MaizeGDB[1] and MaizeSequence[2] websites. Putative maize auxin efflux carriers were initially identified by Blast searches against the reference genome using *A. thaliana* and *O. sativa PIN* transcript and protein sequences as queries (see Table S1 in Supplementary Material, for accession numbers of all the sequences used in this study). Following this approach, we retrieved the sequences of nine new maize PIN-related genes with their relative annotations (see Table S1 in Supplementary Material, for accessions). Additional Blast searches were performed to identify genomic survey sequences (GSSs), transcripts, and transcript assemblies (TAs) from PlantGDB[3], from J. Craig Venter Institute[4], from NCBI[5], and from DFCI/TIGR database[6]. A large amount

[1]www.maizegdb.org
[2]www.maizesequence.org
[3]http://www.plantgdb.org/cgi-bin/blast/PlantGDBblast
[4]http://blast.jcvi.org/tgi_maize/index.cgi
[5]www.ncbi.nlm.nih.gov
[6]http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb = maize

of genomic and expressed sequences was retrieved and univocally assigned to each corresponding primary identified sequence by alignment and consensus generation using ClustalX (Thompson et al., 2002), Geneious, and Lasergene DNAStar tools. In addition, GeneSeqer software[7] was used to assign each TA to its corresponding genomic sequence and indirectly validate the splicing models of genome sequencing derived sequences. Redundant sequences were immediately discarded, while three PlantGDB GSSs without matching sequences in the genome database were further analyzed (AZM5_3988, AZM5_105354, and ZmGSStuc11-12-04.463.1). Two of these GSS (AZM5_105354 and AZM5_3988) resulted partial: the first corresponding to the 5′ end and the other to the 3′ end of a putative maize *PIN* gene and both showing high sequence similarity with a single *PIN* sequence of *S. bicolor* and, to a lesser extent, with *Arabidopsis PIN2*. Based on maize/sorghum alignment the two maize GSSs resulted as not overlapping and so the construction of a contig was not possible. However, using primer pairs designed on the putative 5′UTR and 3′UTR it was possible to amplify a single genomic and cDNA sequence from B73 and close the gap. This sequence and the third GSS (ZmGSStuc11-12-04.463.1) are not included in the B73 genome databases.

As described for the two partial GSS, several primers (Table S2 in Supplementary Material) were designed on the retrieved full-length sequences and were then used in PCR experiments using B73 inbreed genomic DNA and cDNA as template to directly amplify *ZmPIN* genes and further validate them. Cloning of the PCR amplification products was performed with the pCR® II-TOPO® TA-Cloning kit (Invitrogen) or with the pGEM®-T Easy Vector Systems (Promega). *E. coli* competent cells were transformed by heat-shock and selected on plates containing LB medium, plus ampicillin, and X-Gal (IPTG was added when necessary). Plasmids were then purified using the Plasmid Mini Kit (QiAgen), following the manufacturer's instructions. After sequencing on both strands, cloned sequences were edited, aligned against the reference genome, and translated to putative protein sequences using Geneious and Lasergene DNAStar softwares.

Despite many attempts, the *ZmPIN10b* full-length cDNA sequences resulted as being un-amplifiable from any of the tissues used in our study, while the genomic sequence was easily amplified and cloned. Curiously no ESTs or TAs corresponding to this gene were found in public databases, suggesting that this gene could be expressed at very low levels or in extremely specific and restricted domains/cell-types (see also the genome-wide atlas of gene expression data reported in Figure S5 in Supplementary Material; Sekhon et al., 2011). A similar situation concerns the *ZmPIN5b* locus. Annotation on genome browser shows a stop codon soon after the junction between the first and second exon, resulting in a truncated protein. Anyway the alignment between *ZmPIN5b* genomic sequence and sorghum ortholog CDS seems to indicate that the CDS predicted in the maize annotation could be wrong: a single base change in the first exon/intron results in the skipping of stop codon and in the "*in silico*" translation of a functional protein homolog to *PIN5* of different species. Unfortunately just the first exon was successfully amplified from cDNA, while several primers designed from the second to the fifth exon sequences did not work in PCR amplifications. As for *ZmPIN10b*, no EST corresponding to *ZmPIN5b* were retried from databases and no expression has been detected by a genome-wide transcriptome analysis (Figure S5 in Supplementary Material; Sekhon et al., 2011), making it impossible to verify the actual CDS sequence of this gene. Studies described below (phylogenetic and protein profiles analyses) were done using the full-length protein sequence.

PHYLOGENETIC ANALYSIS

The phylogenetic relationships between newly identified maize PIN proteins and PINs of eudicots and monocots were determined through a robust phylogenetic analysis which also allowed determination of the evolution of the PIN auxin efflux carriers family from early-land plants (mosses) to monocots and eudicots. PIN protein sequences of *A. thaliana* and *Medicago truncatula* (two herbaceous eudicots), *Populus trichocarpa* and *Vitis vinifera* (two arboreal eudicots), *Z. mays*, *O. sativa*, *S. bicolor*, *Brachypodium distachyon*, and *Setaria italica* (five monocots), and *Physcomitrella patens* and *Selaginella moellendorffii* (a bryophyte and lycophyte, respectively) were retrieved by family blast searches in the Phytozome v7.0 database[8]. All the accessions of the sequences employed in this study are reported in Table S1 in Supplementary Material. Given the large amount of sequences, a subset of "pocket" trees was generated excluding PIN proteins of *M. truncatula*, *V. vinifera*, *S. bicolor*, and *S. italica* (**Figure 1**; Figures S2 and S3 in Supplementary Material), while a poster tree with all the sequences is reported in Figure S1 in Supplementary Material.

Evolutionary models of PIN proteins were performed using three different approaches and the results compared. First PIN proteins were aligned with ClustalX 1.81 (Blosum Weight Matrix; Gap Opening Penalty: 5; Gap Extension Penalty: 0.20; Thompson et al., 2002), and a neighbor-joining phylogenetic tree was prepared with Phylip package (Felsenstein, 1989) using 100 bootstraps. In this case the yeast PIN-like protein AEL1 (NP_593365 from *Schizosaccharomyces pombe*) was included in the analysis as outgroup (Yang and Murphy, 2009; Zazimalova et al., 2010). The phylograms were drawn using Tree-View 1.6.6[9] and reported in **Figure 1** and Figure S1 in Supplementary Material. In addition, PIN protein sequences were aligned with MAFFT algorithm (Katoh et al., 2002, 2009) available at the EMBL-EBI website[10]. Phylogeny reported in Figure S3 in Supplementary Material was inferred with the Molecular Evolutionary Genetic Analysis 5 (MEGA5; Tamura et al., 2011) program using the Maximum Likelihood method (10,000 bootstraps, Jones–Taylor–Thornton model). Finally the same alignment was also used to construct a NeighborNet Network (Figure S2 in Supplementary Material) with SplitsTree4 software (Huson and Bryant, 2006).

Based on these phylogenetic analyses, newly identified *ZmPIN* genes are named according to the cluster of the *Arabidopsis* and rice *PIN* family to which they belong.

[7]http://www.plantgdb.org/cgi-bin/GeneSeqer/PlantGDBgs.cgi

[8]www.phytozome.net/

[9]http://darwin.zoology.gla.ac.uk/wrpage/treeviewx

[10]http://www.ebi.ac.uk/Tools/msa/mafft/

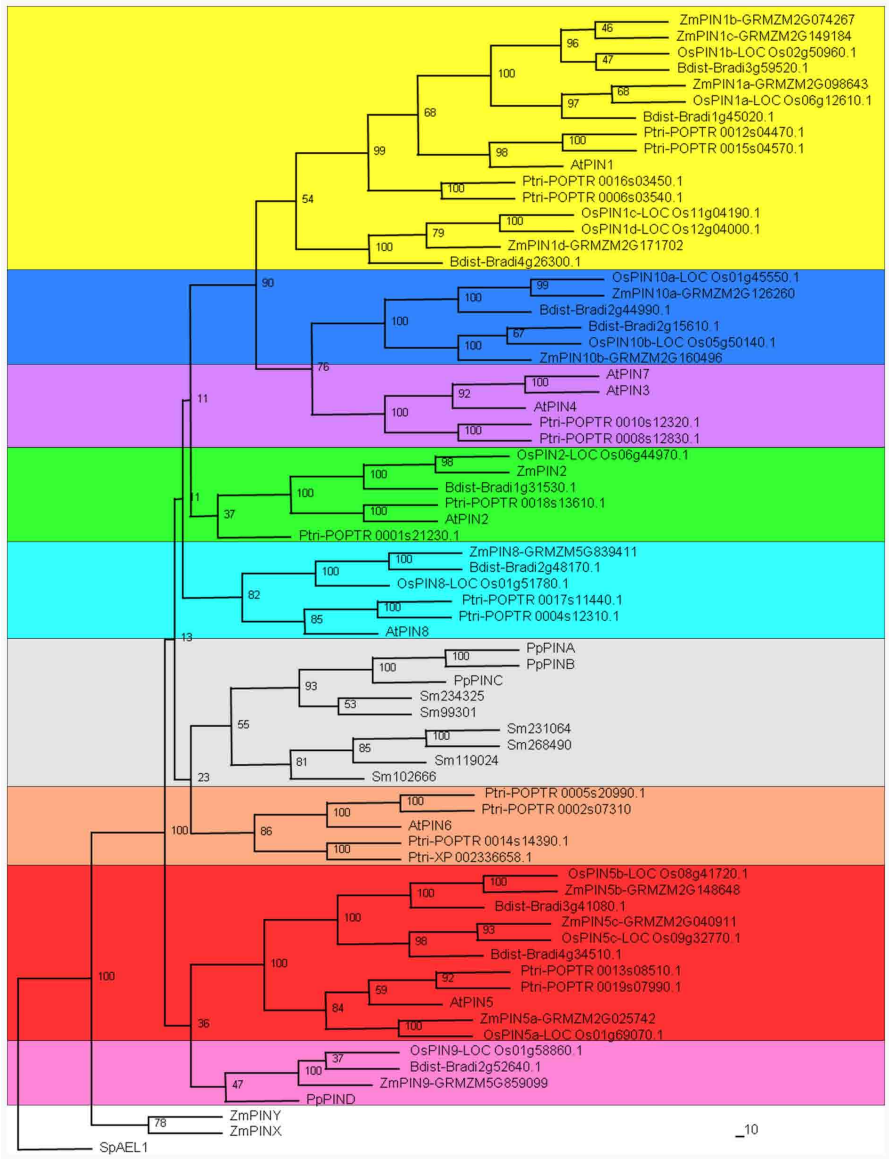


FIGURE 1 | Phylogenetic analysis of PIN auxin efflux carriers from *Zea mays, Arabidopsis thaliana, Oryza sativa, Brachypodium distachyon, Populus trichocarpa, Physcomitrella patens,* and *Selaginella moellendorffii*. The phylogenetic analysis between newly identified maize PIN proteins and PINs of eudicots and monocots, summarizes the evolutionary relationships among the 70 members of the PIN protein family. PIN protein sequences of *Z. mays* (*Zm*), *A. thaliana* (*At*), *P. trichocarpa* (*Ptri*), *O. sativa* (*Os*), *B. distachyon* (*Bdist*), *P. patens* (*Pp*), and *S. moellendorffii* (*Sm*) were retrieved by family blast searches in the Phytozome v7.0 database (www.phytozome.net/). Accession numbers are reported in Table S1 in Supplementary Material. PIN proteins were aligned with ClustalX 1.81 (Blosum Weight Matrix; Gap Opening Penalty: 5; Gap Extension Penalty: 0.20), and a neighbor-joining phylogenetic tree was prepared with Phylip package using 100 bootstraps. The yeast PIN-like protein AEL1 (NP_593365 from *Schizosaccharomyces pombe*) was included in the analysis as outgroup. Newly identified ZmPIN proteins are named according to the cluster of the *Arabidopsis* and rice PIN family to which they belong.

GENE STRUCTURES, MAP POSITIONS, DUPLICATION ANALYSIS, AND PROTEIN PROFILE ANALYSIS

Exon/intron structures of maize *PIN* genes were obtained from the MaizeGDB Genome browser[11], except for *ZmPIN2* and *ZmPINX* that are not present in the genome database and for which exon/intron profiles have been drawn with GSDS, Gene Structure Display Server (Guo et al., 2007)[12] based on the alignment between the previously amplified and sequenced genomic and cDNA sequences.

ZmPIN map positions were identified by searching the www.maizesequence.org database and are manually depicted in **Figure 2** and Figure S4 in Supplementary Material; the exact

[11] http://gbrowse.maizegdb.org/cgi-bin/gbrowse/maize_v2/

[12] http://gsds.cbi.pku.edu.cn/index.php

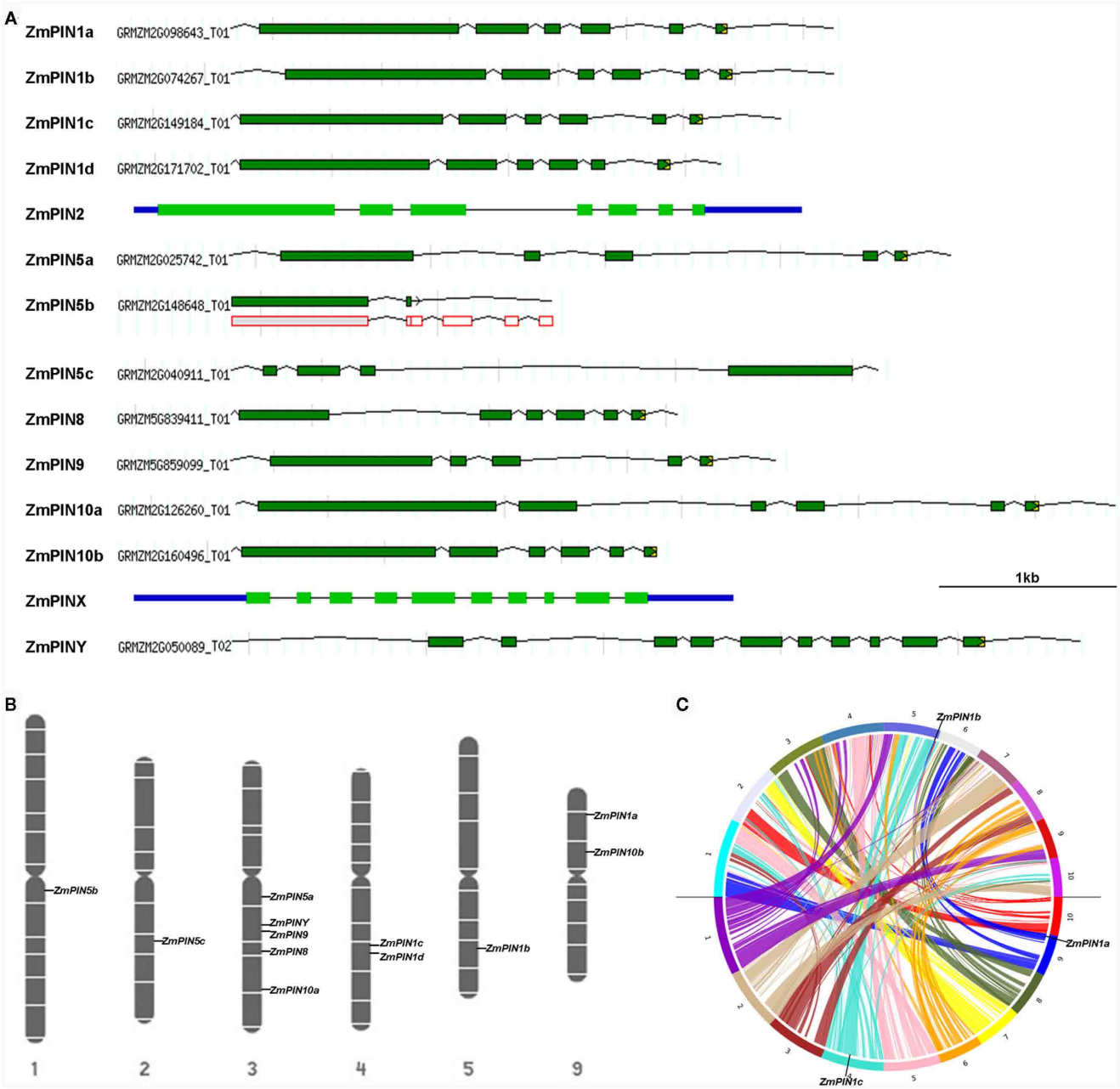


FIGURE 2 | *ZmPINs* gene structure, map position, and duplication analysis. (A) Exon/intron structures of maize PIN genes were obtained from the MaizeGDB Genome browser (http://gbrowse.maizegdb.org/cgi-bin/gbrowse/maize_v2/), except for *ZmPIN2* and *ZmPINX*, for which exon/intron profiles were drawn with GSDS, Gene Structure Display Server (Guo et al., 2007; http://gsds.cbi.pku.edu.cn/index.php) based on the alignment between the previously amplified and sequenced genomic and cDNA sequences. **(B)** Schematic representation of maize chromosomes showing the location of *ZmPIN* loci identified by searching the www.maizesequence.org database. The exact physical position is instead reported in Table S3 in Supplementary Material. **(C)** Circular representation of the 10 maize chromosomes showing duplicated blocks. *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* position in duplicated region has been displayed with Synteny Mapping and Analysis Program (SyMAP) v3.4 (Soderlund et al., 2011; http://www.symapdb.org/).

physical position is instead reported in Table S3 in Supplementary Material.

Duplication analysis of maize *PIN* genes and synteny between maize and rice chromosomes were identified and displayed by Synteny Mapping and Analysis Program (SyMAP) v3.4 (Soderlund et al., 2011)[13]

[13]http://www.symapdb.org/.

Finally, ZmPIN proteins transmembrane helices were predicted using TMHMM2 (Krogh et al., 2001)[14]: the red peaks in **Figure 3** show the predicted transmembrane domains of proteins.

PLANT MATERIAL

Zea mays B73 line was used for DNA and RNA extraction, for expression analysis, immunolocalization, and *in situ* hybridization. Plants were grown in the field or greenhouse, with supplemental lighting when necessary during winter/spring. Kernels were harvested and analyzed at different days after pollination (DAP) after manual self-pollination of plants in the field trials. Roots were collected 2–3 days after kernel germination in rolled adsorbent paper towels soaked with water in a growth chamber at 26˚C with a photoperiod of 16 h light and 8 h dark as described in Woll et al., 2005. Kernels were previously surface sterilized with 70% ethanol for 10 min and then rinsed three times with water.

The maize mutant line *br2* (Multani et al., 2003; Pilu et al., 2007) was kindly supplied by Dr. Roberto Pilu (Università di Milano) after introgression for six generations in B73 background. Homozygous *br2* mutant plants were grown as previously described. Furthermore, the reporter lines pZmPIN1a::ZmPIN1a:YFP and DR5::mRFP (Gallavotti et al., 2008)[15], kindly provided by Prof. Dave Jackson, were introgressed for three additional generations in B73 keeping the transgene in a hemizygous state.

NUCLEIC ACID EXTRACTION

Genomic DNA was extracted from maize leaves with the DNeasy Plant Mini Kit (QiAgen) according to the manufacturer's instructions. Total RNA was extracted from maize tissues (seedlings, roots, leaves, tassels, ears, and kernels) using to the RNeasy Plant Mini Kit (QiAgen) and subjected to on-column DNase treatment (QiAgen). Before retrotranscription, total RNA was quantified by spectrophotometer analysis.

EXPRESSION ANALYSES BY RT-PCR, REAL-TIME PCR, AND *IN SITU* HYBRIDIZATION ASSAYS

For the semi-quantitative RT-PCR expression analysis, *ZmPIN* and *ZmABCB1/BR2* gene transcriptions were investigated in vegetative and reproductive tissues and kernels. Vegetative tissues included 3 days-old seedlings, young leaves, the apex (from 0 to 1.5 cm) and the elongation/mature zone (from 1.5 to 3 cm) of primary roots, and the seventh and fifth node, while for the reproductive tissues ear and tassel at the stage of 1 and 3 cm were considered. Kernels collected at different DAP (0, 3, 5, 6, 7, 8, 9, 10, 12) were also taken into account for the *ZmPIN* expression analysis.

For the detailed quantitative *PINs* expression analysis in B73 line and *br2* mutant, young roots were collected from seedlings vertically grown in a growth chamber for 3 days as previously described. These roots were divided into specific segments, which were labeled from A to E indicating the segment from 0 to 0.2 cm (A), from 0.2 to 0.5 cm (B), from 0.5 to 1 cm (C), from 1 to 1.5 cm (D), and from 1.5 to 2 cm (E; **Figure 6A**). In these root segments the expression of *ZmPIN* genes and *ZmABCB1/BR2* was investigated by Real-Time PCR.

For both expression analyses, cDNA synthesis was performed with the SuperScript III reverse transcriptase kit (Invitrogen), according to the manufacturer's instructions. One microgram of total RNA was used as a template together with 1 μl oligo (dT) 12–18 (0.5 μg/μl – Invitrogen). In the semi-quantitative RT-PCR expression analysis, each cDNA was then diluted 1:10 and 1 μl was used for PCR amplification in a volume of 25 μl. The constitutively expressed *GAPC2* gene was used as housekeeping internal control of the cDNA/RNA quantity. Expression analyses were performed with at least two primer combinations for each gene. Primer sequences are reported in Table S2 in Supplementary Material. In order to obtain semi-quantitative results, the number of cycles of the PCR was adjusted for each gene to obtain barely visible bands in agarose gels (see Table S2 in Supplementary Material for the number of amplification cycles for each gene). Aliquots of the PCR were loaded on 1.2% agarose gels and stained with Sybr-Safe (Invitrogen). For all *PINs* the different primer combinations tested gave comparable amplification results. *ZmABCB1/BR2* amplification products (see Pilu et al., 2007 for primers used in RT-PCR amplification) were checked on 2% agarose gel. Images were captured with Gel Logic 100 Imaging System and Molecular Imaging software (Kodak).

Quantitative Real-Time PCR expression analysis was performed using an ABI 7500 Real-Time PCR System (Applied Biosystems) and the POWER SYBR® GREEN PCR Master Mix (Applied Biosystems), following the manufacturer's guidelines. Real-time conditions were: 2 min at 50˚C, 10 min at 95˚C, 40 cycles of: 15 s at 95˚C and 1 min at 60˚C. For each reaction, we observed product melting curves by heating from 60 to 95˚C at 0.2˚C/s. For all transcripts, this procedure allowed identification of a single product, which we confirmed by analysis on 2% agarose gel. Two different biological samples, obtained by different RNA preparations from separate root pools, were processed; three replicates were carried out for each primer combination and for each biological sample. The absolute quantification was carried out by the generation of standard curves for each gene. Expression levels of each auxin transporter gene were normalized to *GAPC2* transcript quantities. Primer sequences are reported in Table S2 in Supplementary Material.

In situ hybridization experiments were conducted as described by Varotto et al., 2003. Plant material (kernels, vegetative apexes, roots, and inflorescences) was fixed in 4% paraformaldehyde (Sigma) in 0.1 M phosphate buffer (pH 7.2) for 16 h at 4˚C, then dehydrated by ethanol series and embedded in Paraplast Plus (Sigma). Sections (7–10 μm) were cut using a Leica RM 2135 microtome (Leica, Germany) and collected on xylane-coated SuperFrost Plus slides (Menzel-Glazer). Slides were then deparaffinized and treated with 10 μg/ml proteinase K.

In vitro transcription of the DIG-UTP (Roche) labeled RNA sense and antisense probes was obtained using T7 and SP6 polymerases. More information on primer sequences used to amplify the probes for each *PIN* gene can be found in Table S2 in Supplementary Material. The hybridization was performed in a 50% formamide buffer at 48˚C overnight. DIG detection and

[14] http://www.cbs.dtu.dk/services/TMHMM-2.0/
[15] http://maize.jcvi.org/cgi-bin/maize/cellgenomics/geneDB_list.pl

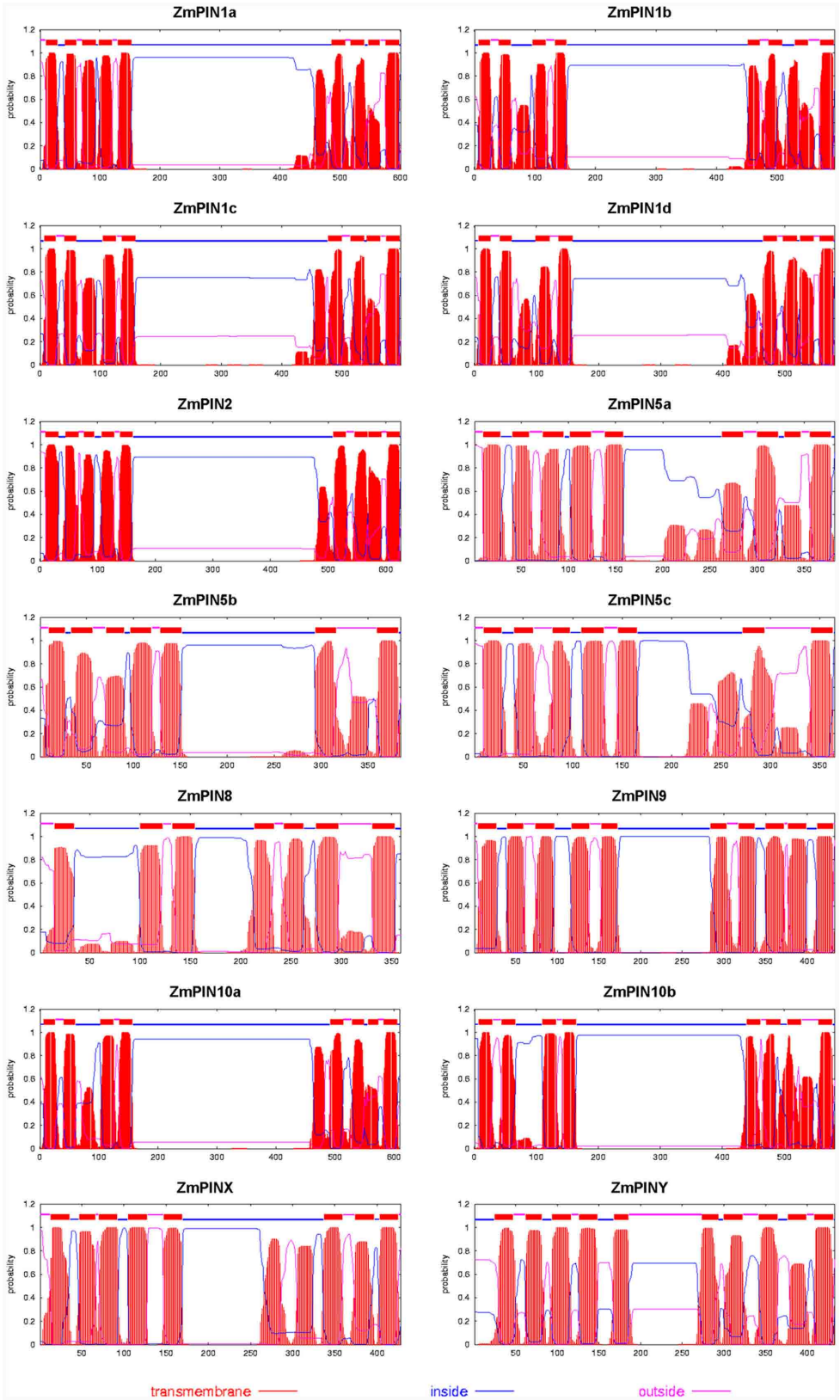


FIGURE 3 | The predicted protein profiles of the maize PIN transporters. The transmembrane helices of PIN proteins were predicted using TMHMM2 (Krogh et al., 2001; http://www.cbs.dtu.dk/services/TMHMM-2.0/): the red peaks show the predicted transmembrane domains of proteins.

signal visualization were done using Anti-Digoxigenin-AP antibody (Roche) and NBT plus BCIP (Roche), following the manufacturer's instructions. Slides were air-dried and mounted with DPX mounting medium (Fluka Biochemika). Negative controls obtained with sense riboprobes are reported in Figures S6A–D in Supplementary Material.

MOLECULAR ANALYSES UNDER NPA AND NAA TREATMENTS

Zea mays B73 line and *br2* mutant seeds were surface sterilized with 70% ethanol for 10 min and then rinsed three times with sterile water. Kernels were placed on soaked 3MM paper in horizontal position for 1 day. After imbibition, the seeds were transferred in rolled adsorbent paper towels in vertical position for successive treatments. Seedlings were treated with 50 μM NPA (1-naphthylphthalamic acid, a PAT inhibitor), 50 μM NAA (2-naphthalen-1-ylacetic acid, a synthetic auxin), and water as control growth condition for 2 days. Roots were collected and divided in two specific segments indicated with A (from 0 to 0.2 cm) and B (from 0.2 to 0.5 cm).

The expression profiles of *ZmPIN* genes and *ZmABCB1/BR2* in untreated, NPA- and NAA-treated roots were investigated by Real-Time PCR using the standard curve method, as reported above.

The reporter lines pZmPIN1a::ZmPIN1a:YFP and DR5::mRFP were treated with NPA and NAA at the same concentrations and observed at the confocal microscope.

IMMUNOLOCALIZATION ASSAYS

PIN1 immunostaining was performed on wax-embedded (9:1 – PEG400 distearate:1-hexadecanol 99%, Sigma) roots collected on 2–3 days-old plants vertically grown on rolled paper. Roots were fixed in 4% paraformaldehyde in PBS 1× solution for 1 h under vacuum conditions at room temperature, embedded, and cut (10–12 μm) with a Leica RM 2135 microtome (Leica, Germany). After dewaxing sections were blocked for 2 h in 1% BSA in 1× PBS buffer and stained overnight with the anti-AtPIN1 antibody (aP20 sc-27163, Santa Cruz Biotechnology) at a 1:250 dilution in PBS 1× and 1% BSA. The next day, the sections were washed twice for 10 min in PBS 1× and stained for 1 h with the secondary antibody (anti-goat IgG conjugated with Alexa568, Molecular Probes) diluted 1:400 in PBS 1×. Sections were washed twice for 10 min with PBS 1× and mounted with VECTASHIELD® with DAPI Mounting Medium (Vector Laboratories, USA). For the negative controls of immunostaining experiments hybridization was performed with the secondary antibody only.

A monoclonal anti-IAA antibody (Biofords – Agdia – PMD09346/0096) was used for IAA immunolocalization in maize roots. Tissues were prefixed in 3% (w/v) *N*-Ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDAC, Sigma) 0.1 M in PBS 1×, post-fixed in 4% paraformaldehyde in the same buffer and embedded in Paraplast Plus (Sigma). Sections (7–10 μm) were cut using a Leica RM 2135 microtome (Leica, Germany), collected on polysine slides (Menzel-Glazer) and processed as previously reported by Avsian-Kretchmer et al., 2002. The anti-IAA primary antibody was used at a concentration of 0.05 mg/ml. The secondary antibody, an anti-mouse IgG-AP (Santa Cruz Biotechnology), was used at a 1:250 dilution. Signal visualization was done using SIGMAFAST™Fast Red Tablets (Sigma) and when purple was observed, the colorimetric reaction was stopped in water. Negative controls of IAA immunostaining experiments were performed on samples no pre-fixed with EDAC or hybridizing slides with the IgG-AP secondary antibody only and are reported in Figures S6E–G in Supplementary Material.

MICROSCOPY AND REPORTER LINES OBSERVATIONS

pZmPIN1a::ZmPIN1a:YFP and DR5::mRFP reporter lines roots were hand sectioned, mounted with 50% glycerol in water and observed with a Leica TCS SP2 laser confocal microscope (Leica Microsystems, Heidelberg, Germany) using the follow settings: (i) ZmPIN1a–YFP excitation at 514 nm and signal collection between 530 and 580 nm; (ii) DR5 excitation at 543 nm and signal collection between 580 and 640 nm. Images were coded red for RFP and yellow for the YPF reporter line.

Immunolocalization and *in situ* hybridization images were taken with a Leica DM4000B Digital microscope, equipped with a Leica DC300F Camera, and Leica Image Manager 50 software (Leica Microsystems, England). Images were coded red for Alexa568 and blue for DAPI.

RESULTS

THE MAIZE PIN FAMILY OF AUXIN EFFLUX CARRIERS: IDENTIFICATION, PHYLOGENETIC ANALYSIS, GENE STRUCTURES, AND PROTEIN PROFILES

Members of the maize PIN family were initially identified by a comprehensive Blast search in public databases using *A. thaliana* and *O. sativa PIN* transcript and protein sequences as queries. MaizeGDB and MaizeSequence were used as primary databases to obtain the genomic sequences coming from the B73 Maize Genome Sequencing project, release 5b.60, and their relative annotations, while GSSs from PlantGDB and from J. Craig Venter Institute, transcript sequences from NCBI and TAs from DFCI/TIGR were used to initially confirm and integrate the previously identified sequences and splicing models. In addition to the three *ZmPIN1* genes previously characterized (Carraro et al., 2006; Forestan et al., 2010), nine full-length sequences were found in the genomic release 5b.60 and a tenth sequence was identified in a GSS from PlantGDB. Furthermore, two partial, not overlapping, GSSs were assembled together using a sorghum *PIN* sequence to fill the gap, for a total of 11 new putative auxin efflux carrier genes. A preliminary *in silico* characterization was performed to identify putative ORFs and amino acid sequences. All the full-length sequences were later amplified from genomic DNA and cDNA (except *ZmPIN5b* and *ZmPIN10b* for which the amplification of the cDNA full-length proved unsuccessful), to validate the sequences of the 11 maize *PINs* that were named according to the cluster of *Arabidopsis* and rice PIN family to which they belonged.

The phylogenetic relationships between early-land plants, monocots and eudicots PIN proteins were investigated including the new maize PIN proteins in a phylogenetic analysis (**Figure 1**; Figures S2 and S3 in Supplementary Material), together with PIN proteins of *A. thaliana*, *P. trichocarpa*, *O. sativa*, *B. distachyon*, *P. patens*, and *S. moellendorffii*. A wider evolutionary study was performed by also including PIN proteins of *M. truncatula*, *V. vinifera*,

S. bicolor, and *S. italica* (Figure S1 in Supplementary Material). PIN protein sequences of different species were retrieved from Phytozome v7.0 and analyzed with different sequence alignment and phylogenetic analysis tools, to develop a robust evolutionary model of transporters. Auxin efflux carrier family members were aligned with ClustalX 1.81 (Blosum Weight Matrix; Gap Opening Penalty: 5; Gap Extension Penalty: 0.20; Thompson et al., 2002) and a neighbor-joining phylogenetic tree was prepared with the Phylip package (Felsenstein, 1989) using 100 bootstraps (**Figure 1**; Figure S1 in Supplementary Material). In addition, PIN proteins were aligned with MAFFT algorithm (Katoh et al., 2002, 2009) and phylogeny inferred using MEGA5 program (Maximum Likelihood method; 10,000 bootstraps; Jones–Taylor–Thornton model; Tamura et al., 2011; Figure S3 in Supplementary Material) or constructing a NeighborNet Network with SplitsTree4 software (Huson and Bryant, 2006; Figure S2 in Supplementary Material). The three different approaches showed comparable results on the evolutionary history of auxin efflux carrier proteins in early-land plants, monocots and dicots, also suggesting the evolutionary relationships between the 11 newly identified maize proteins, *Arabidopsis*, and rice PIN proteins (**Figure 1**; Figures S1–S3 in Supplementary Material). Indeed, the four different phylogenetic trees show the same clustering of newly identified maize PIN proteins with *Arabidopsis* and rice orthologs, while differences in few nodes with low support can be observed for *V. vinifera* and *P. trichocarpa* proteins (**Figure 1**; Figures S1–S3 in Supplementary Material).

In addition to the previously characterized *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c*, a fourth gene phylogenetically close to *AtPIN1* was identified. We named this gene, which clusterizes together with rice *OsPIN1c* and *OsPIN1d* genes, *ZmPIN1d*. The PlantGDB database search and the phylogenetic analysis also allowed identification of a putative *ZmPIN2* gene; indeed the two available partial GSSs codify for a protein closely related to AtPIN2 and OsPIN2.

The gene structure analysis showed that also the fourth *ZmPIN1* gene, *ZmPIN1d*, has six exons, whose distribution is similar to that of *AtPIN1* and the four *OsPIN1* genes (**Figure 2A**). Like the *Arabidopsis* and rice orthologs, *ZmPIN2* gene presents a complex exon/intron structure, with seven exons and six introns. We also identified four genes that encode for proteins lacking the central hydrophilic domain that normally characterizes long PIN proteins. This peculiar hydropathic profile is typical of the *Arabidopsis* PIN5 and PIN8 proteins and our phylogenetic analysis confirmed that three of the maize proteins clusterize with *Arabidopsis* and rice PIN5 (ZmPIN5a, ZmPIN5b, and ZmPIN5c), while the remaining one represents the maize *PIN8* ortholog (*ZmPIN8*). So, as recently reported in rice, three duplicated *PIN5* genes and a *PIN8* gene are also present in the maize genome. *ZmPIN5a* and *ZmPIN5b* have four introns, while *ZmPIN5c* has three with the first of about 2 kb in length (**Figure 2A**). As described in the Section "Material and Methods," genome annotation of *PIN5b* gene showed the presence of a stop codon soon after the first splicing site, resulting in a truncated protein (**Figure 2A**). Anyway, the alignment with the sorghum ortholog CDS showed that the splicing model annotated could be wrong and a shift of a single base pair resulted in a full-length PIN5 protein. Unfortunately RT-PCR using several primer combinations proved unsuccessful in the amplification of the complete CDS. Moreover no ESTs or TAs corresponding to *ZmPIN5b* were found on public databases and a recent Genome-Wide Expression Atlas revealed that *ZmPIN5b* is not expressed in the 60 distinct analyzed tissues representing 11 major organ systems of inbred line B73 (Figure S5 in Supplementary Material; see text footnote 1; Sekhon et al., 2011). *ZmPIN8* has a very similar exon/intron structure to that of *OsPIN8* (**Figure 2A**).

Finally, we identified three proteins that do not clusterize with any *Arabidopsis* PIN proteins, confirming, as suggested for rice, the existence of monocot-specific *PIN* genes. Based on the phylogenetic analysis, these maize genes were named *ZmPIN9*, *ZmPIN10a*, and *ZmPIN10b*. These monocot-specific PIN genes have a conserved structure when compared to their rice counterparts. Phylogenetic analyses also revealed that *ZmPINX* and *ZmPINY* proteins clusterize together with the *Arabidopsis* PIN-like proteins (Figure S1 in Supplementary Material).

Chromosome map positions of maize *PINs* have been identified by genome browser searches and are reported in **Figure 2B** and Figure S4 in Supplementary Material. Duplication analysis revealed that *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* are located in duplicated blocks on chromosomes 9, 5, and 6, respectively (**Figure 2C**; Figure S4 in Supplementary Material). Interestingly, both maize and rice *PIN5a*, *PIN8*, *PIN9*, and PIN10a map on highly syntenic regions of chromosomes 3 and 1, respectively (Figure S4 in Supplementary Material, Wang et al., 2009).

The predicted transmembrane architecture of the ZmPIN proteins showed the typical structure with a conserved N-terminal region of transmembrane motifs, a variable middle domain representing the cytoplasmic domain of transmembrane and a C-terminal domain of transmembrane segments (**Figure 3**). Interestingly, ZmPIN9 putative protein presents a hydropathic profile in between that of long PINs (AtPIN1 and AtPIN2, which are plasma membrane localized) and that of short PINs (the ER-localized AtPIN5 and AtPIN8).

EXPRESSION ANALYSIS OF MAIZE AUXIN TRANSPORTERS

The tissue-specific expression pattern of 12 *ZmPIN* genes (*ZmPIN1a*, *ZmPIN1b*, *ZmPIN1c*, *ZmPIN1d*, *ZmPIN2*, *ZmPIN5a*, *ZmPIN5b*, *ZmPIN5c*, *ZmPIN8*, *ZmPIN9*, *ZmPIN10a*, and *ZmPIN10b*), two *ZmPIN-like* genes (*ZmPINX* and *ZmPINY*) and *BR2*, a maize *MDR/PGP/ABCB* efflux transporter were analyzed by semi-quantitative RT-PCR and the results are reported in **Figure 4**. These results are in agreement with those obtained by the genome-wide transcriptome analysis previously mentioned and reported in Figure S5 in Supplementary Material (Sekhon et al., 2011).

ZmPIN1a, *ZmPIN1b*, and *ZmPIN1c* expression analysis confirmed that these genes are ubiquitously expressed but differentially modulated in maize vegetative and reproductive tissues and during kernel development (see also Carraro et al., 2006; Forestan et al., 2010). A different expression pattern was observed for *ZmPIN1d* that is specifically expressed in tassels, ears and in the fifth node of adult plants and not detectable in caryopses after pollination. *ZmPIN2* is expressed in root tip, male and female inflorescences and its expression is modulated during the early stages of kernel development. *ZmPIN5a* is mainly expressed in the elongation/mature zone of the primary root, in nodes and in the young seed till 10 DAP. *ZmPIN5b* is up-regulated in the

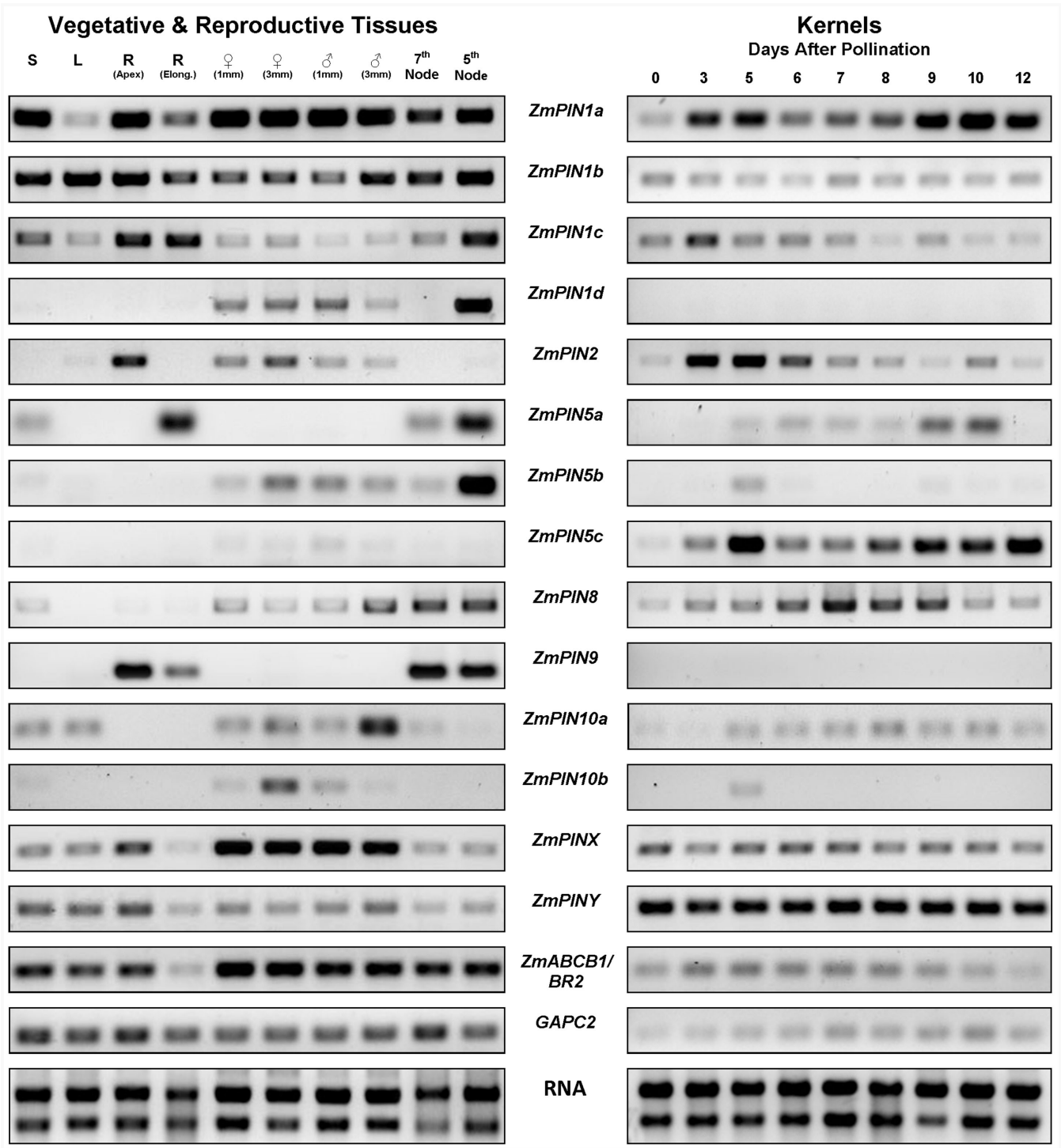


FIGURE 4 | Expression analysis of maize auxin transporters. Semi-quantitative RT-PCR analysis of 12 *ZmPIN* genes, two maize *PIN-like* genes and *BR2*, a maize MDR/PGP/ABCB efflux transporter in vegetative and reproductive tissues and in kernels from 0 to 12 days after pollination (DAP). The constitutively expressed *GAPC2* gene was used as housekeeping internal control of the RNA/cDNA quantity. To get semi-quantitative results, the number of cycles of the PCR was adjusted for each gene to obtain barely visible bands in agarose gels (see Table S2 in Supplementary Material for the number of amplification cycles for each gene). *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* are expressed ubiquitously and differentially up-regulated in the tissues. *ZmPIN1d* is expressed in tassels, ears, and in the 5th node of adult plants. *ZmPIN2* is expressed in root apex (R), in male and female inflorescences and during maize kernel development. *ZmPIN5a* is expressed in seedlings (S), in the root elongation/mature zone, in nodes and in the young seed till 10 DAP. *ZmPIN5b* is expressed in male and female inflorescences,

(Continued)

FIGURE 4 | Continued
in nodes and at a low level during kernel development. *ZmPIN5c* is highly expressed during kernel development. *ZmPIN8* is expressed in male and female inflorescences during kernel development. *ZmPIN9* is expressed in roots and nodes. *ZmPIN10a* is expressed in seedlings and leaf (L), in male and female inflorescences and at low level during kernel development. *ZmPIN10b* is expressed in ears and tassels. *ZmPINX* and *ZmPINY* are both expressed ubiquitously and differentially up-regulated in the tissues. *BR2* shows a constitutive expression but it appeared down-regulated in the root elongation zone in comparison with root apexes.

fifth node, and has a weak expression in male and female inflorescences and during kernel development (note that the amplified fragment corresponds to the first exon while no amplifications were obtained with downstream located primers). *ZmPIN5c* is highly expressed during kernel development from 3 to 12 DAP. *ZmPIN8* is expressed in all the tissues analyzed with the exception of the root and is detectable during kernel development, where it appeared up-regulated between 6 and 9 DAP. Of the monocot-specific genes, *ZmPIN9* is exclusively expressed in roots and nodes while *ZmPIN10a* is expressed during male and female inflorescence development and at very low level during kernel development. Finally a very weak expression of *ZmPIN10b* is detectable in young ears and tassels.

The expression analyses of the two maize *PIN-like* genes showed that they are ubiquitously expressed and differentially up-regulated in all the analyzed tissues. In detail, *ZmPINX* is up-regulated in root apex and male and female inflorescences, while *ZmPINY* is highly expressed during kernel development.

Finally, *BR2* transcript is detectable in all the tested tissues with a lower expression level in the root elongation zone compared to other vegetative and reproductive tissues. Moreover, the *BR2* transcript amount is detectable at a very low level during tested kernel development stages.

LOCALIZATION OF *PIN* TRANSPORTERS AND AUXIN MAXIMUM IN MAIZE PRIMARY ROOTS

The longitudinal structure of the maize root includes the root cap at the terminal end, the root apical meristem, and the elongation and root hair zones. In cross section, the root is organized in a series of concentric cell layers: the outermost epidermic layer and exodermis, the pluristratified cortex, the endodermis, and the pericycle that surrounds the central vascular cylinder (see Hochholdinger et al., 2004 for a review on root development in maize). Six *ZmPIN* genes are expressed in maize root: indeed, quantitative RT-PCR data showed that *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* and the monocot-specific *ZmPIN9* are expressed both in root apical and distal zones (**Figure 4**). In addition, *ZmPIN2* appeared up-regulated in root apex, while *ZmPIN5a* is up-regulated in the root elongation/mature zone (**Figure 4**). The expression pattern of *ZmPIN* genes expressed in the root apex was determined at tissue level using *in situ* hybridization (**Figures 5A–O**). In root apex longitudinal sections, comprising the root cap and meristematic area, a *ZmPIN1* probe, which recognized the three gene transcripts expressed in the root showed the hybridization signal in the root cap with a clear up-regulation in calyptrogen, in the root meristematic area, in the epidermis and in the differentiating vascular tissues of the central cylinder (data not shown). Specific probes for *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* were also transcribed *in vitro* DIG-labeled and hybridized on root apex longitudinal sections. The results showed that *ZmPIN1a* is the only *PIN1* transcript localized in the calyptrogen and in the meristem region of the root (**Figure 5A**), while *ZmPIN1b* is mainly expressed in epidermis, root cap, and vasculature (**Figures 5B–D**). Finally a specific *ZmPIN1c* probe showed that *ZmPIN1c* is localized in the epidermis and vasculature of the central cylinder (**Figures 5E–G**).

The localization pattern of the three *ZmPIN1* genes was also confirmed by immunolocalization experiments, using an anti-AtPIN1 antibody, which was shown to recognize the proteins of the three maize *PIN1* genes expressed in the root. PIN1 proteins are polarly localized in the epidermis, meristem region, and central cylinder, where they are inserted in the basal cell membrane (**Figure 5P**) while they are detectable in the cytoplasm of the calyptrogens and proliferating cap (**Figure 5Q**). At protein level, the localization pattern of ZmPIN1a protein was observed using a ZmPIN1a–YFP reporter line (Gallavotti et al., 2008). At the confocal microscope the YFP signal was evident in the basal cell membranes of the central cylinder cells and in the cytoplasm of root cap cells and meristematic region of the root apex (**Figures 5T–V** and **7H**).

ZmPIN2 transcripts were localized in root epidermis, exodermis, and multiple layers of cortex in cross section of the root apex (**Figures 5H–M**). Finally, *ZmPIN9* the monocot-specific PIN gene is mainly expressed in endodermis, pericycle, and in phloem of the central cylinder (**Figures 5N,O**).

To correlate the localization pattern of the PIN genes expressed in maize root with auxin distribution and maximum in the same tissues, observations at the confocal microscope of a DR5::mRFP reporter line (Gallavotti et al., 2008) and immunolocalizations with an anti-IAA antibody on root longitudinal sections were both performed. Auxin response was observed in the root cap and vasculature in the DR5-RFP reporter line fresh root dissected at the confocal microscope (**Figures 5W–Y** and **7A,B**). Auxin accumulation was evident in the root cap, in particular in the calyptrogen, in the epidermis at the level of root meristem and vasculature in longitudinal root sections labeled with the anti-IAA antibody (**Figures 5R–S**).

ZmPIN GENES EXPRESSION ANALYSIS IN ROOTS OF B73 LINE AND *br2* MUTANT

The *br2* maize mutant exhibits reduced shootward auxin transport at the root apex and reduced root gravitropic growth (McLamore et al., 2010). To investigate whether this alteration in auxin fluxes influenced the expression of *ZmPIN* transporters, the mRNA levels of *ZmPIN1a*, *ZmPIN1b*, *ZmPIN1c*, *ZmPIN2*, and *ZmPIN9* genes, the *PIN* members expressed in root apex, were analyzed by Real-Time PCR in different segments of B73 and *br2* young primary roots.

As shown in **Figure 6B**, in *br2* mutant the expression level of *ZmABCB1/BR2* transcript was drastically reduced in all root segments (from A to E) compared to the levels observed in B73 line.

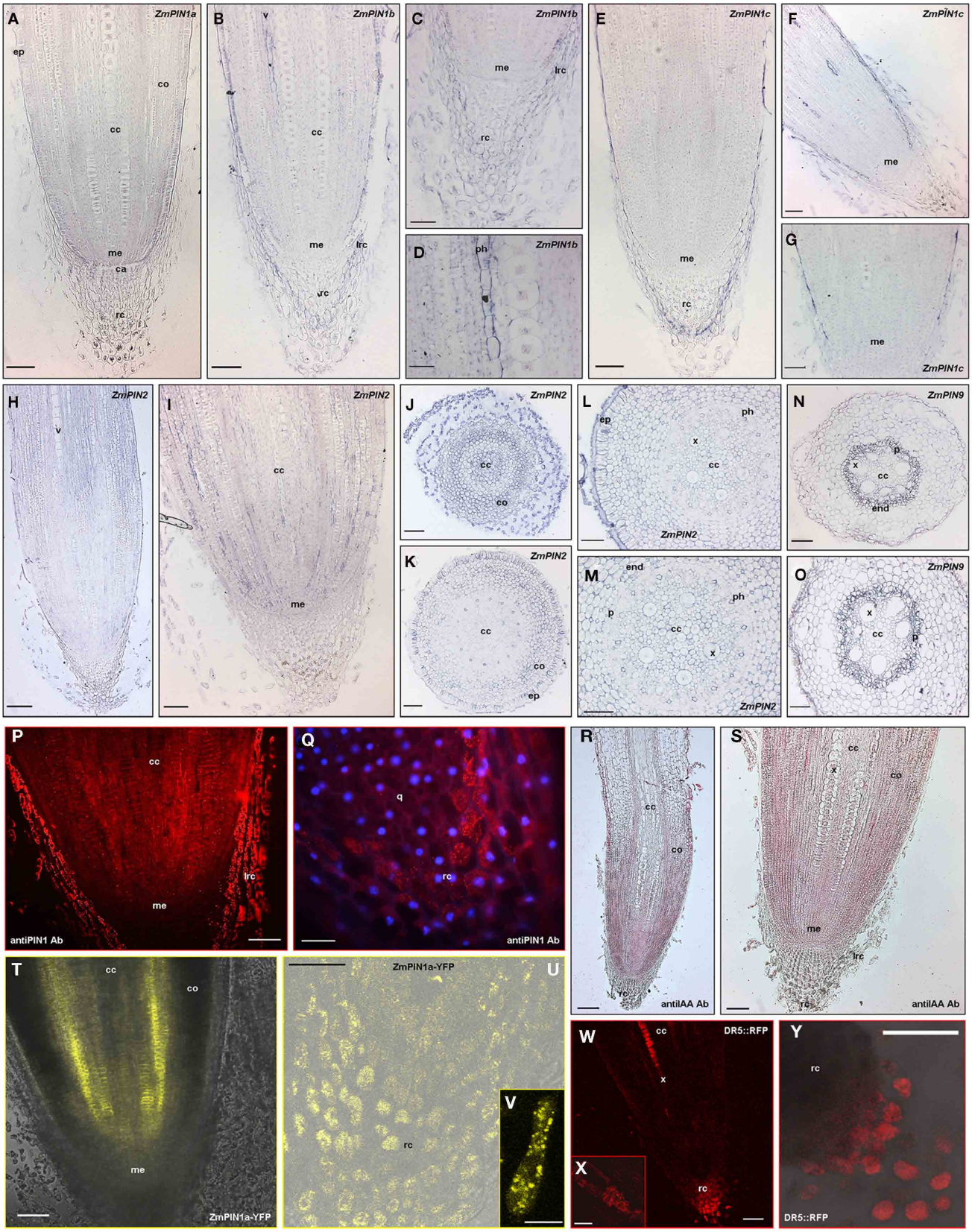


FIGURE 5 | Continued

FIGURE 5 | Continued
Localization of PIN transporters and auxin maximum in maize root. *In situ* hybridization on maize root sections with DIG-labeled antisense probes of maize *PIN* transporters expressed in the root **(A–P).** The hybridization signal is represented by the blue staining. Images were acquired with a Leica DC300F camera. ZmPIN1 protein immunolocalization **(Q,R)** and ZmPIN1a::YFP reporter line observation at the confocal microscope **(U,V)** in maize primary root can be correlated with auxin accumulation revealed by DR5 activity **(W,X)** or IAA immunolocalization textbf(S–T). **(A)** Longitudinal section of primary root hybridized with the *ZmPIN1a* antisense transcript; the hybridization signal is localized in the root cap, particularly in the calyptrogen and root meristematic region. **(B–D)** Longitudinal sections through a primary maize root hybridized with *ZmPIN1b* antisense transcript showing the hybridization signal in root cap, epidermis, endodermis, and vasculature of the central cylinder. In **(B)** the signal is visible in root cap, epidermis, and vasculature. **(C)** portrays a magnification of the root cap and **(D)** is a higher magnification of **(B)** showing the blue staining in the phloem cells. **(E–G)** Longitudinal sections of a maize primary root hybridized with a *ZmPIN1c* antisense probe: the signal is detectable in root cap, epidermis, and vasculature. **(H–M)** *In situ* hybridization on maize primary root sections with a *ZmPIN2* DIG-labeled antisense RNA probe. **(H,I)** are longitudinal sections of a primary root showing the hybridization signal of *ZmPIN2* in root cortex and central cylinder. **(J–M)** depict cross sections through a maize primary root hybridized with the *ZmPIN2* antisense probe. **(J)** The picture portrays a cross section of a root tip at the level of proximal root meristem in which the signal is visible in the proliferating cells of the root cap, cortex, and differentiating central cylinder. **(K–M)** Cross sections showing the concentric tissues composing the maize primary root: epidermis, exodermis, cortex, endodermis, xylem elements, phloem, and pericycle. In **(K)** the hybridization signal of *ZmPIN2* is detectable in exodermis, cortex, and phloem. **(L,M)** show the hybridization signal in the cortex and phloem. **(M)** is higher magnifications of **(L)**. **(N,O)** Cross sections of a maize primary root hybridized with *ZmPIN9* antisense riboprobe. The signal is localized in epidermis, exodermis, endodermis, and pericycle that surround the central cylinder. **(P,Q)** Anti-PIN1 immunolocalization on longitudinal sections root apexes. The anti-AtPIN1 primary antibody shows protein localization on basal cell membranes of root meristematic region **(P)** and in proliferating root cap **(P,Q)**. DAPI was used to counterstain the nuclei in **(Q)**. Both panels represent epifluorescence images acquired with a Leica DC300F camera. **(R,S)** Auxin immunolocalization in maize primary roots. In these longitudinal sections an anti-IAA monoclonal antibody shows a gradient of auxin toward the root tip, with the concentration maximum in the lateral root cap and in calyptrogens cells. **(T–V)** Confocal images of the pZmPIN1a::ZmPIN1a:YFP transgenic line root apex. ZmPIN1a–YFP is localized in proliferating root cap **(U)**, in calyptrogens **(U)**, in the quiescent center **(T)**, in the proximal root meristem **(T)** and in the basal membranes of central cylinder cells **(T)**. The single root cap cell in **(V)** shows the cytosolic localization of YFP signal in the root cap cells. **(W–Y)** Confocal images of a DR5::mRFP maize reporter line. Auxin accumulation is indirectly visualized in the tip of the root cap **(W,Y)** and in the vasculature of the central cylinder **(W)**. The inset **(X)** represents a single xilematic cell showing DR5 activity while **(Y)** depicts details of **(W)**. ca, calyptrogen; cc, central cylinder; co, cortex; ep, epidermis, end, endodermis; ex, exodermis; me, root meristem; p, pericycle; ph, phloem; q, quiescent center; rc, root cap; v, vasculature; x, xylem. Scale Bars: 200 μm in **(H,R)**; 150 μm in **(W)**; 100 μm in **(A–C,E–G,I–P,S,T,Y)**; 75 μm in **(U)**; 50 μm in **(D,Q,X)**; and 20 μm in **(V)**.

ZmPIN1a transcript showed a slight higher expression in the root segments A–C and lower levels in segments D–E of *br2* mutant when compared with the same segments of B73. On the contrary, the mRNA level of *ZmPIN1b* was higher in A–B–C root segments of *br2* mutant, while in the D–E segments its level was comparable to that observed in wt. Also *ZmPIN1c* showed a different expression pattern in *br2* compared to B73 line: in particular, an up-regulation was observed in segment A, while a down-regulation occurs in segment D.

The transcript level of *ZmPIN2* was slightly higher in all tested segment of *br2* mutant, in particular in the distal region of the root (segments D and E) where *PIN2* is not detectable in B73. Finally, *ZmPIN9* transcript was up-regulated in segments C–D of *br2* mutant.

Taken all together these results highlight that the expression profile of *PIN* transporters is altered in *br2* mutant primary roots.

GENE EXPRESSION ANALYSIS IN ROOTS OF B73 LINE AND *br2* MUTANT UNDER DRUG TREATMENTS

The expression profile of *ZmPIN* genes and *ZmABCB1/BR2* was investigated by Real-Time PCR in root apex of B73 and *br2* mutant line grown in water (control), in presence of exogenous auxin (NAA), and of a PAT inhibitor (NPA; **Figure 6C**). Due to the strong reduction of root elongation caused by NAA application, the analysis was done just on the first two root segments indicated with A (from 0 to 0.2 cm) and B (from 0.2 to 0.5 cm).

In both lines the addition of NAA and NPA affected the expression of *ZmPIN* genes and *ZmABCB1/BR2* in both root segments. In particular in B73, NPA strongly inhibited *ZmABCB1/BR2* expression in segment A, while NAA application doubled its expression levels in segment B. The same NAA effect was observed in *br2* mutant roots. Both NAA and NPA treatments increased *ZmPINa* and *ZmPIN1c* transcript levels in segment B in both genotypes; on the contrary NAA application strongly reduced *ZmPIN1c* levels in the segment A of *br2* root. Similarly, *ZmPIN1b* transcript levels were strongly reduced after NAA application in *br2* mutant roots (segments A and B), while NPA affected its expression exclusively in the B segment. NAA treatment increased *ZmPIN9* transcript levels in the segment B of mutant roots, while *ZmPIN2* expression was affected in a lower extent by the drug applications.

Experimental evidences showed that auxin transporter expression is altered by exogenous auxin or PAT inhibitor applications and these alterations are more dramatic in the *br2* mutant background.

DRUG TREATMENTS ALTER MORPHOLOGY, AUXIN DISTRIBUTION, AND ZmPIN1a–YFP LOCALIZATION IN PRIMARY ROOTS

The effects of NAA and NPA applications were investigated at cellular level by microscope observation of DR5::mRFP and ZmPIN1a–YFP reporter lines (**Figure 7**). NAA applications caused the main effects on root anatomy, auxin responses, and PIN1a localization (**Figures 7B,E,F,I,K**). Exogenous auxin application indeed resulted in a size reduction of both root cap and meristem (compare **Figures 7B,D**) and in the formation of a pluristratified epidermis (**Figures 7E,F**). After the treatment, DR5 activity is diffused in the lateral root cap cells (**Figure 7B**) and not limited to the root cap tip as in the control (**Figure 7D**). Auxin response was detectable in the vasculature that differentiates very close to the root meristem (**Figures 7B,E**) and also the pluristratified epidermis showed DR5 activity (**Figure 7F**). Although PIN1a–YFP always localized in the stele cells in both NAA-treated and control

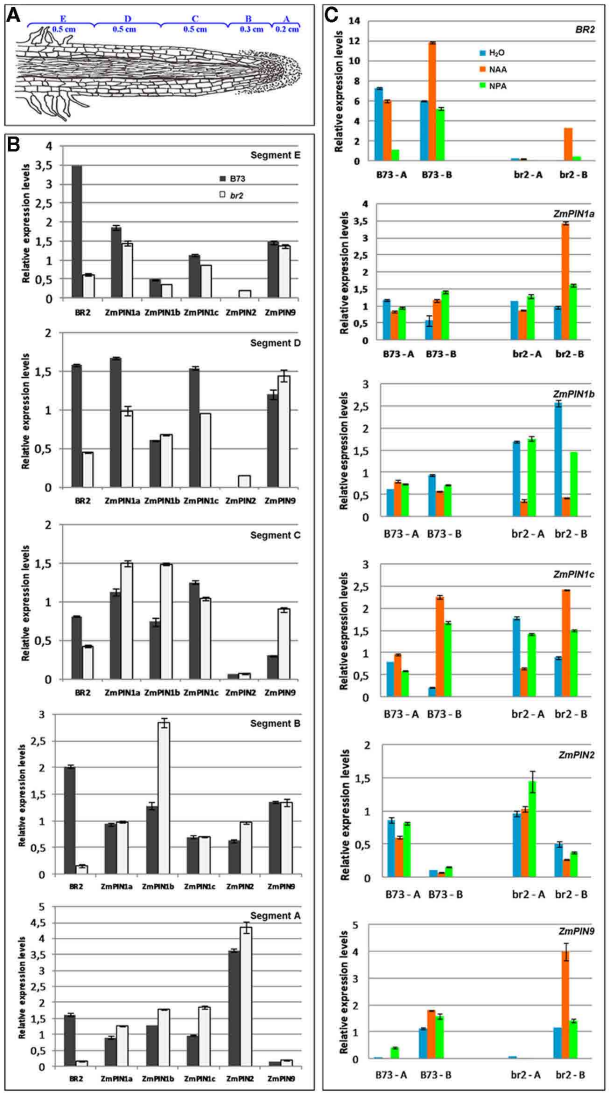


FIGURE 6 | *br2/abcb1* mutation and drug applications alter auxin transporter expression levels. Quantitative RT-PCR analysis of *ZmABCB1/BR2, ZmPIN1a, ZmPIN1b, ZmPIN1c, ZmPIN2,* and *ZmPIN9* was performed in root segments (from A–E) of B73 line and *br2* mutant under different treatments. The average value from three replicates and two biological samples was estimated by the generation of standard curves for each gene. Expression levels of each auxin transporter gene were normalized to *GAPC2* transcript quantities. Standard deviation bars are indicated in the diagram. **(A)** Schematic representation of a maize root indicating the different segments used in expression analysis. **(B)** Quantitative RT-PCR expression analysis of auxin transporters in B73 and *br2* mutant roots. The altered auxin shootward transport in the mutant root results in different expression levels of *PIN* mRNAs. **(C)** Quantitative RT-PCR expression analysis of auxin transporters after NAA and NPA application in B73 and *br2* mutant roots. Applications of 50 μM NAA and 50 μM NPA to young seedlings result in a tissue-specific modulation of PIN transcripts. The most dramatic effects are observable in the mutant roots.

roots, some observed differences in protein localization reflected the changes in root morphology, particularly at the level of the vasculature and meristem (**Figures 7H,I,K**).

Following NPA applications, auxin signal was detectable in root cap as in the control (data not shown) while a diffuse DR5 activity marked the whole central cylinder and the cortex cells (**Figures 7C,G**). This suggests that NPA caused a failure in creating the auxin maximum detectable in the vascular cell in the control roots (**Figure 7A**). Similarly, the ZmPIN1a–YFP reporter line showed a more diffuse localization of the PIN1a protein that marked also the cortex cells (**Figures 7J,L**).

ZmPIN LOCALIZATION PATTERN IN MAIZE INFLORESCENCES

In maize, at the transition from vegetative to reproductive development the shoot apical meristem (SAM) becomes an inflorescence meristem (IM) which starts to produce lateral meristems called branch meristems (BMs). BMs develop spikelet-pair meristems (SPMs), each bearing two spikelet meristems (SMs). SMs form two floral meristems (FMs), the upper FM and lower FM, which produce the floral organ of the tassel. In each flower of the tassel, after gynoecial development, the gynoecium degenerates, and an imperfect staminate flower results. Within a few weeks after the SAM transition, a lateral shoot meristem in the axil of a leaf becomes an ear IM and produces multiple rows of SPMs, which follow the same step as in the tassel until the lower floret aborts as do the stamens of the upper floret. Indeed, in the mature ear a pistillate flower is present per spikelet (see McSteen et al., 2000 for a review on maize inflorescence development). Since *ZmPIN1d* transcript is specifically up-regulated in both male and female inflorescences, *in situ* hybridizations were performed in the SAM, starting from the transition stage to SPMs and spikelet differentiation, in male inflorescences, and during female inflorescence development (**Figures 8A–H**). During flowering transition, which is characterized by the elongation of the SAM and the adoption of an inflorescence identity to form a tassel primordium, *ZmPIN1d* hybridization signal is detectable in the L1 layer of the IM apical meristem (**Figure 8A**); later on, *ZmPIN1d* is localized in the L1 layer of the developing tassel SPMs primordium (data not shown). In the ear, *ZmPIN1d* showed the same localization patterns as in the tassel, being localized in the L1 layer of the IM during the transition phase and maintaining its localization in the L1 cell layer during SPM differentiation (**Figures 8B,C**). During SMs differentiation and growth, *ZmPIN1d* is localized in the tips of the developing glumes (**Figure 8D**). Later on in ear development, during the pistillate flower differentiation, *ZmPIN1d* hybridization signal is detectable in the outer cell layers of the gynoecium, till the differentiation of the gynoecium ridge that will encircle the ovule region giving rise to the stylar canal (**Figures 8E–G**) and in the nucellus surrounding the mother cells of the megaspore (**Figure 8H**). The same localization pattern was observed in the pistillate flowers of female inflorescence sections hybridized with *ZmPIN1a* (**Figures 8I–K**), *ZmPIN1b* (**Figure 8L**), and *ZmPIN2* (**Figures 8M,N**) antisense DIG-labeled mRNA probes.

ZmPIN LOCALIZATION PATTERN IN MAIZE KERNEL

RT-PCR showed that *ZmPIN2*, *ZmPIN5c*, *ZmPIN8*, and *ZmPIN10a* are up-regulated during the early phases of kernel development, from 3 to 12 DAP. To precisely determine the localization of these genes at tissue level, *in situ* hybridization experiments were done on longitudinal sections of maize kernels from 5

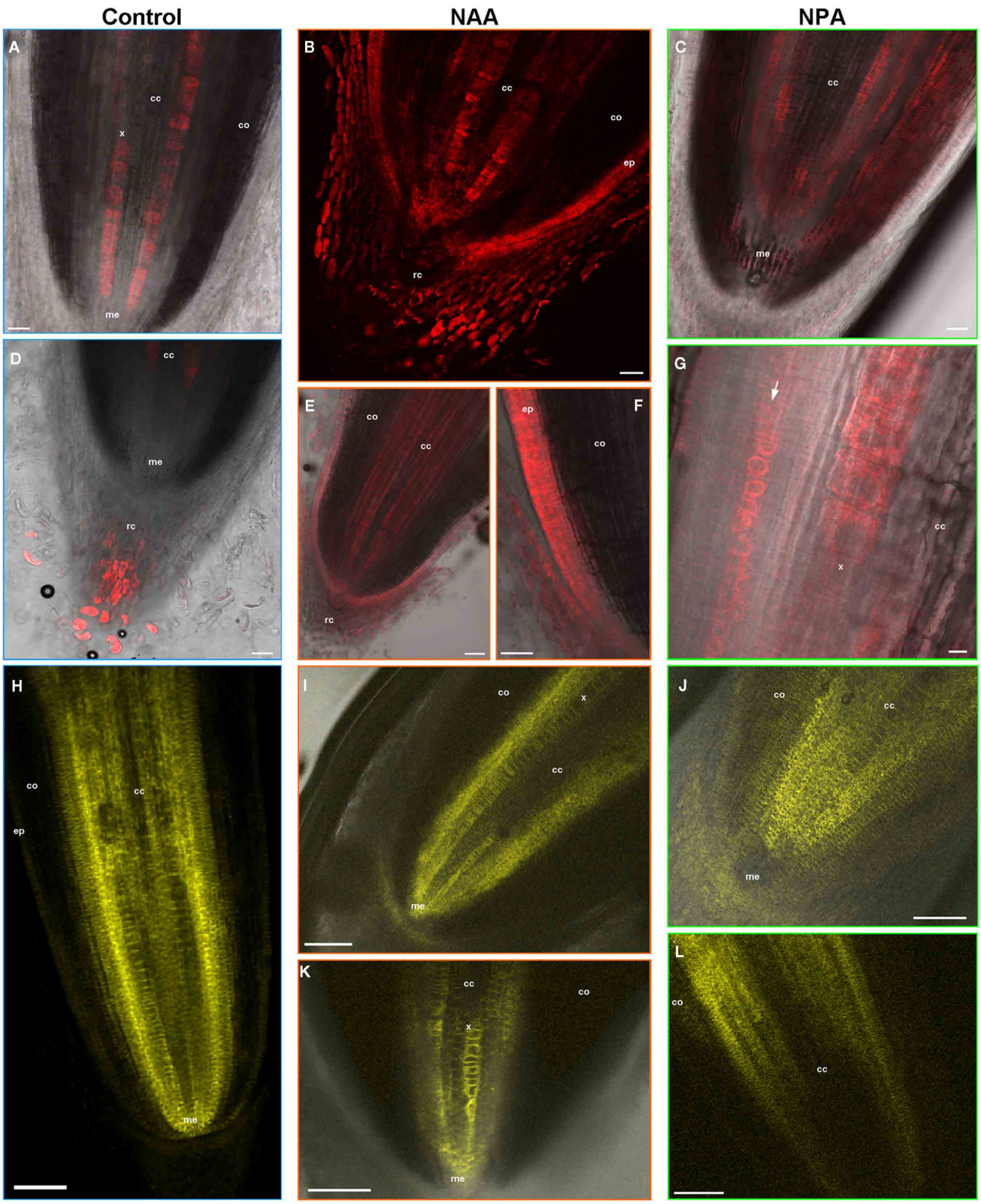


FIGURE 7 | Continued

FIGURE 7 | Continued
NAA and NPA applications alter root anatomy. auxin response and PIN1a–YFP localization. The effects of NAA and NPA treatments on root anatomy, auxin gradients, and PIN1a protein localization are visualized by DR5:mRFP **(A–G)** and ZmPIN1a::YFP reporter lines observation at the confocal microscope **(H–L). (A,D)** Confocal images of a DR5::mRFP reporter line root grown in water. DR5 activity is visualized in the tip of the root cap **(D)** and in the vasculature of the central cylinder **(A)**. **(B,E,F)** Exogenous auxin application results in root anatomic alteration **(B,F)** and in a diffused DR5 signal in the lateral root cap cells **(B)**, in the central cylinder **(B,E)** and in the pluristratified epidermis **(F)**. **(C,G)** Following NPA applications a diffuse DR5 activity marks the whole central cylinder and the cortex cells. In particular high DR5 signal is visible in the endodermis cells [arrow in **(G)**]. **(H)** Confocal image of the pZmPIN1a::ZmPIN1a:YFP transgenic line root apex. ZmPIN1a–YFP is localized in the quiescent center, in the proximal root meristem and in the basal membranes of central cylinder cells. **(I,K)** PIN1a–YFP localization alteration in NAA-treated primary roots reflects the modified root anatomy at the level of the meristem and vascular tissues. **(J,L)** After NPA treatment ZmPIN1a–YFP reporter line shows a more diffuse signal of the PIN1a protein that marks also the cortex cells. cc, central cylinder; co, cortex; ep, epidermis, me, root meristem; rc, root cap; x, xylem. Scale Bars: 150 μm in **(H)**; 100 μm in **(E,I–L)**; 50 μm in **(A–D and F)** and 10 μm in **(G)**.

to 12 DAP (**Figure 9**). During this early stage of maize kernel development the main cellular domains of the endosperm, the Basal Transfer Layer (BETL), the embryo surrounding region (ESR), the aleurone, and sub-aleurone differentiate, before the accumulation of reserves occurs. At 5 DAP *ZmPIN2* is expressed in the differentiating endosperm ESR and in the embryo proper (**Figure 9B**); at 12 DAP later on during kernel development when the main cellular domains of the endosperm are completely differentiated, *ZmPIN2* transcript is localized in the BETL, ESR, aleurone, and developing embryo at the transition stage (**Figures 9A,C,K**). *ZmPIN5c* transcripts are localized in ESR, BETL, and maternal chalazal cells of the kernel and in the embryo at both 5 and 12 DAP (**Figures 9D–G**); at 12 DAP its hybridization signal appears particularly evident in the outer epidermic layer of the embryo scutellum (**Figure 9F**), in the transition zone between BETL and aleurone (**Figure 9G**), while transcripts are not detectable in the aleurone layer (**Figure 9L**). *ZmPIN10a* transcript is also localized in both BETL and ESR endosperm cellular domains and embryo starting at 5 DAP (**Figures 9H–I**), and are also detectable in the chalazal and pedicel kernel tissues at 12 DAP (**Figure 9H**). Finally, *ZmPIN8* is expressed in BETL cells, maternal chalazal tissues (**Figure 9J**), and aleurone layer (**Figure 9M**).

DISCUSSION

This work reports the identification and characterization of the *PIN* auxin efflux carrier family members in *Z. mays*. In addition to the previously characterized *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* (Carraro et al., 2006; Forestan et al., 2010) we identified nine new members of the family, that is lastly made up by four genes homologous to *Arabidopsis PIN1*, three to *AtPIN5*, a *PIN2*, and a *PIN8* ortholog and by three monocot-specific *PINs* (*ZmPIN9*, *ZmPIN10a*, and *ZmPIN10b*). Anyway, more efforts are necessary to amplify and characterize the full-length CDSs of *ZmPIN5b* and *ZmPIN10b*, excluding (or confirming) the possibility they represent two pseudogenes.

Genes homologous to the *Arabidopsis PINs* are present in genomes throughout the plant kingdom, from the model moss *P. patens* to all vascular plants, and the relatively high amino acid identity between PIN proteins suggests that all the *PIN* genes diverged from a single ancestral sequence. At least four independent phylogenetic studies of *PIN* sequences from different plant species, including early-land plants, monocots, and dicots, revealed that the monocot *PIN* family is wider and more divergent than the dicots one, with two or three genes homologous to one single *Arabidopsis PIN* gene and also with the presence of at least one monocot-specific *PIN* gene (Paponov et al., 2005; Xu et al., 2005; Zazimalova et al., 2007; Krecek et al., 2009; Wang et al., 2009).

Our results strongly confirmed this hypothesis, showing a widening of PIN auxin efflux carriers family in maize, and more in general in monocots. By now it is largely accepted that maize genome undergoes a whole genome duplication (Helentjaris et al., 1988; Gaut and Doebley, 1997; Schnable et al., 2009) and several maize genes were found in duplicated chromosomic regions (Bomblies et al., 2003; Nardmann et al., 2004; Nardmann and Werr, 2006). Anyway, the widening of the monocots PIN family cannot be explained only by whole genome duplication, since in the diploid *B. distachyon* only two *PINs* (a *PIN1* and a *PIN5* ortholog) are missing compared to maize and rice. On the contrary, a whole genome duplication event could explain the widening of PIN members belonging to the *P. trichocarpa* family (Tuskan et al., 2006).

In detail, of the four *PIN1* maize orthologs, chromosomal duplication analysis revealed that *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* are located on duplicated regions on chromosomes 9, 5, and 4, respectively (**Figures 2B,C** and Figure S4 in Supplementary Material), and, together with *ZmPIN1d* share a common exon/intron structure (six exons and five introns, **Figure 2A**). These four genes encode for canonical long PIN proteins, with two conserved hydrophobic domains at the N- and C-termini and a long variable hydrophilic central region. Four *PIN1* genes have also been found in rice (Wang et al., 2009): subfunctionalization of monocots *PIN1* genes is confirmed by the expression analysis in different maize and rice plant tissues. *ZmPIN1a* and *ZmPIN1b* are ubiquitously expressed, *ZmPIN1c* is preferentially expressed in the post-embryonic root and stem while *ZmPIN1d* shows a high expression specificity, marking the transition from the vegetative to the reproductive development in SAM and IMs. Similarly in rice, *OsPIN1a* and *OsPIN1b* showed a constitutive expression, *OsPIN1c* is mainly expressed in root and stem while the expression pattern of *OsPIN1d* is still awaiting characterization (Wang et al., 2009).

A certain degree of expression specificity can be observed also analyzing maize *PIN1* gene localization in the post-embryonic root: *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* are expressed in root cap, epidermis, and central cylinder. However *ZmPIN1b* is expressed in the lateral root cap and in the epidermis, *ZmPIN1c* is mainly detectable in the root cap tip and in the differentiating vasculature and *ZmPIN1a* in the calyptrogens and in the root meristem. The specific *ZmPIN1a* transcript localization in root calyptrogen was confirmed at protein level by both immunolocalization experiments and observations of a ZmPIN1a:YFP reporter line. Both

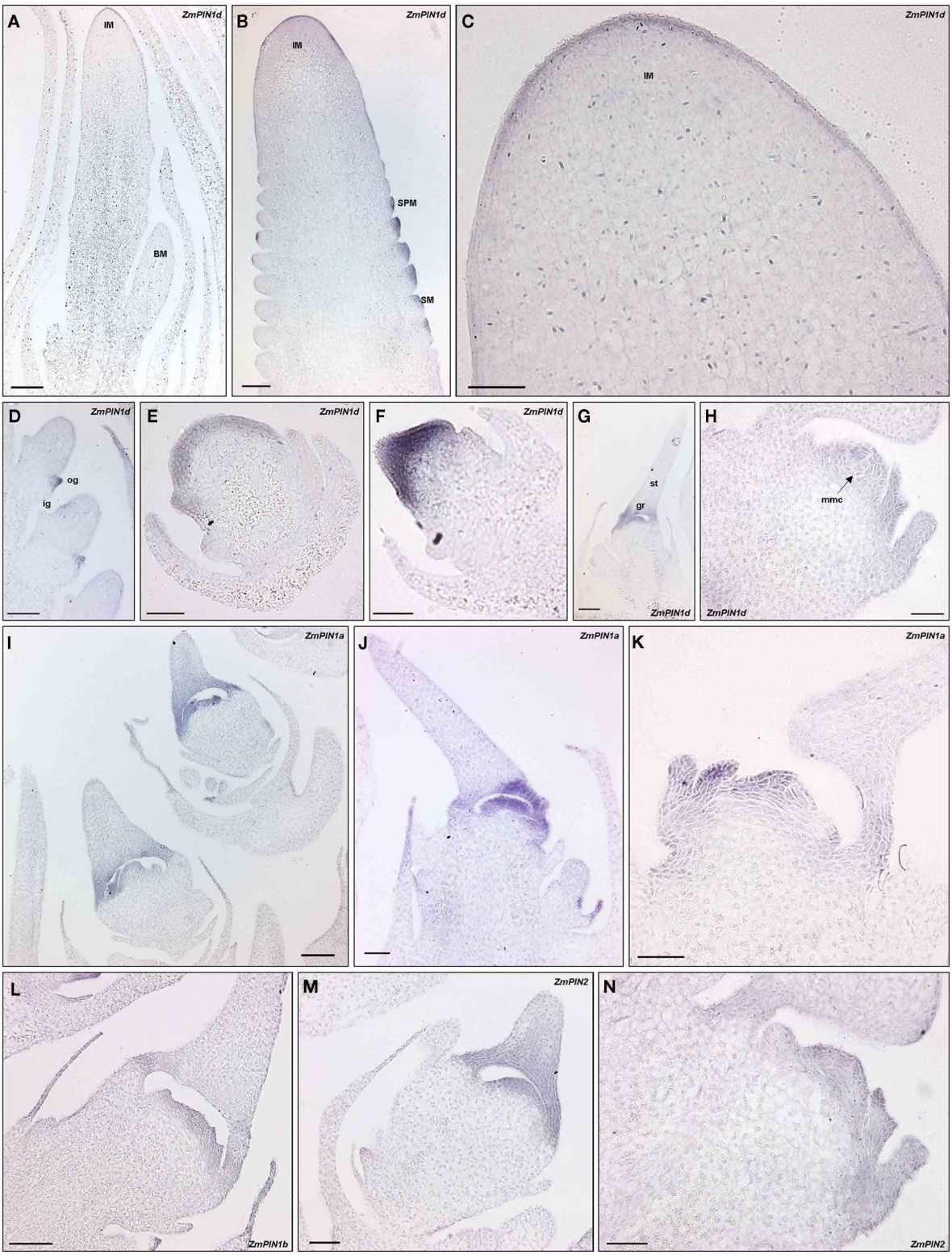


FIGURE 8 | Contined

FIGURE 8 | Continued
***PIN* transcripts localization pattern in maize inflorescences. Longitudinal sections through maize developing inflorescences hybridized with a *ZmPIN* DIG-labeled antisense riboprobes.** The hybridization signal is visualized by the blue staining. The images were acquired with a Leica DC300F camera. **(A–H)** Inflorescence longitudinal sections hybridized with *ZmPIN1d* antisense probe. **(A)** Tassel inflorescence meristem (IM) showing an up-regulation of *ZmPIN1d* expression at its L1 cell layer. **(B,C)** Ear IM forming its spikelet meristems (SMs): both IM L1 layer and SMs show *ZmPIN1* expression. **(C)** is a higher magnification of **(B)**. **(D)** shows a strong signal from the newly formed glume primordia of ear SMs. **(E–H)** Longitudinal sections of developing pistillate flowers. *ZmPIN1d* hybridization signal is detectable in the outer cell layers of the gynoecium till the differentiation of the gynoecium ridge that will encircle the ovule region giving rise to the stylar canal and in the nucellus surrounding the mother cells of the megaspore. (**I–N)** *ZmPIN1a* **(I–K)**, *ZmPIN1b* **(L)**, and *ZmPIN2* **(M–N)** expression domains during the development of pistillate flowers overlap that of *ZmPIN1d*. IM, inflorescence meristem; SM, spikelet meristem; SPM, spikelet-pair meristem; ig, inner glume; og, outer glume; st, stylar canal; gr, gynoecium ridge; mmc, megaspore mother cell. Scale Bars: 100 μm in **(A–B,D,G,I–J,L)**; 50 μm in **(C,E,F,H,K,M,N)**.

localization procedures showed that PIN1 proteins are localized in the proximal root meristem. The anti-AtPIN1 antibody also shows the polarization of ZmPIN1 proteins in the cortex and central cylinder cells, suggesting that the proteins are involved in auxin transport toward root apex, as modeled for *Arabidopsis* root. Interestingly, in calyptrogen ZmPIN1a is not inserted in cell plasma membrane but is accumulated in the cytoplasm. The same ZmPIN1 accumulation pattern was previously observed in the ESR during maize endosperm development (Forestan et al., 2010) and ZmPIN1a has been observed to accumulate in intracellular compartments also by the transformation of tobacco protoplast with a ZmPIN1a::GFP fusion construct (Forestan and Varotto, 2010). Interestingly, in maize endosperm ESR the accumulation of PIN1 proteins in the cell endomembranes is coupled with an auxin maximum in the same cells. Similarly in maize root, both anti-IAA immunolocalization and observation of a DR5-reporter line showed the presence of an auxin maximum in the root cap.

An auxin accumulation also characterizes the *Arabidopsis* root cap and columella cells (Sabatini et al., 1999): auxin is supplied to the root cap by the PIN1- and PIN4-dependent acropetal route and then laterally redistributed by PIN3- and PIN7-mediated efflux from columella cells (reviewed in Peer et al., 2011). The simultaneous and coordinated expression of *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c* in the maize root apex could suggest that these three proteins play the role of the *Arabidopsis* PIN3, PIN4, and PIN7 efflux carriers. Indeed, as in rice, no *ZmPIN* genes were grouped into the AtPIN3, AtPIN4, or AtPIN7 cluster (Wang et al., 2009). A previous phylogenetic analysis suggested a close evolutionary relationship between these three *Arabidopsis* transporters: AtPIN3 and AtPIN7 are located on two copies of a duplicated block of 285 genes on chromosome 1 and the rate of synonymous substitution analysis indicates that these blocks are the result of a recent duplication. Similarly, AtPIN3 and AtPIN4 share the same relative position in duplicated blocks of genes. However, these blocks are too small to be considered as statistically significant (Paponov et al., 2005). Our Phylip-phylogenetic study seems to suggest an evolutionary relationship between the AtPIN3/AtPIN4/AtPIN7 cluster and PIN10a/PIN10b monocots proteins that belong to a monophyletic group. However, this hypothesis is not confirmed by the evolutionary analysis based on Maximum Likelihood method, making any speculation weak. Nor can the expression analysis of the two maize *PIN10* genes support this evidence: *ZmPIN10a* is expressed in leaves and kernels and both *ZmPIN10a* and *ZmPIN10b* are expressed in male inflorescence at the limit of detection, as confirmed by a search in Genome-Wide Expression Atlas (Figure S5 in Supplementary Material; Sekhon et al., 2011), with an expression pattern totally different from *AtPIN3*, *AtPIN4*, and *AtPIN7*.

Although the diploid monocot *B. distachyon* has only two *PIN5* genes (the transporters lacking the central hydrophilic domain), rice, sorghum, and maize have three *PIN5* sequences in their genome. In maize there is no evidence of duplication events involving regions in which *PIN5* genes map; anyway both the divergent expression pattern of the three genes and the putative pseudogene nature of *ZmPIN5b* suggest that the tracks of any potential duplication event might have been lost during chromosomal rearrangements and the subfunctionalization of these sequences.

In addition to *PIN5s*, a fourth short PIN transporter has been identified as the *AtPIN8* ortholog and named *ZmPIN8*. All the short PIN transporters, the three PIN5 and PIN8 proteins, have a conserved hydrophobic profile with their orthologs in *Arabidopsis* that are localized in ER. However, this peculiar subcellular localization, which has been suggested as having an important implication from an evolutionary point of view (Wabnik et al., 2011), is still awaiting confirmation in monocots.

The monocot-specific ZmPIN9 and OsPIN9 proteins do not correspond to the traditional classification in long and short PINs, based on the length of hydrophilic domain. Their transcripts are specifically localized in the root and their proteins show an intermediate length of the hydrophilic domain. Furthermore, they grouped (together with sorghum and *Brachypodium* orthologs) in the same cluster with the *P. patens* PIND, whose intron sequences analysis suggested the possibility of its horizontal transfer from monocots (Krecek et al., 2009; Mravec et al., 2009). Unlike PpPINA, PpPINB, and PpPINC, the three closely related *Physcomitrella* PINs, which were localized in ER in transient transformation assays using tobacco protoplasts, PpPIND protein showed a non-specific cytosolic localization pattern (Mravec et al., 2009). All these observations make the investigation on subcellular localization of monocot PIN9 proteins particularly interesting.

Although *ZmPIN2* gene was expressed in vegetative and reproductive tissues and during kernel development, its locus was not identified in the maize genome 5b.60 release. *ZmPIN2* has a conserved intron/exon structure with *AtPIN2* but, similarly to its ortholog in rice (Wang et al., 2009) it does not have the same specific expression pattern at the root level. *ZmPINX* gene is a second missing gene in the maize genome database and together with *ZmPINY* represents the first *PIN-like* sequences characterized in maize so far. As for *Arabidopsis*, phylogenetic analysis revealed that *PIN* and *PIN-like* families form distinct clusters (Figure S1 in Supplementary Material; Paponov et al., 2005).

FIGURE 9 | Continued

FIGURE 9 | Continued
Localization pattern of *PINs* transporters in maize kernel. Longitudinal sections through maize kernels at different DAP, hybridized with antisense mRNA DIG-labeled probes corresponding to the *PIN* transporters expressed in maize caryopses. The hybridization signal is visualized by the blue staining. **(A–C)** *ZmPIN2* mRNA localization during kernel development. **(A)** In 12 DAP endosperm section *ZmPIN2* transcripts localize in the basal endosperm transfer layers facing the maternal placenta of the seed. **(B,C)** Depict longitudinal section through two maize kernels at 6 and 12 DAP, respectively: in both sections, the hybridization signal is detectable in the endosperm embryo surrounding region (ESR) and in the embryo where it marks the outer cell layers. **(D–G)** *ZmPIN5c* transcript *in situ* hybridization in embryos and endosperm. **(D,G)** Show longitudinal sections through a 12 DAP kernel; *ZmPIN5c* hybridization signal is localized in the completely differentiated transfer cells facing the placenta in **(D)** and also in the differentiating cells at the flanks of the kernel, where the aleurone starts in **(G)**. **(E,F)** Are longitudinal sections of maize kernel at 9 and 14 DAP, respectively, showing *ZmPIN5c* hybridization signal in the embryo. In transition stage embryos ZmPIN5c marks the embryo proper, while no signal is detectable in the suspensor **(E)**; later on development, in L1-stage embryo the transcripts localize in the scutellum and Root Apical Meristem **(F)**. **(H,I)** *ZmPIN10a* antisense hybridizations in longitudinal sections of kernels and embryos. In 12 DAP kernels the hybridization signal is visible in both the BETL layer and maternal tissues (placenta and pedicel) of the seed **(H)**, while at 5 DAP the ESR and the embryo proper are labeled **(I)**. **(J)** *ZmPIN8* mRNA localization in longitudinal section of a maize kernel at 12 DAP. The blue staining is detectable in the outer cell layers of the endosperm comprising the BETL and the aleurone. **(K–M)** *ZmPIN2*, *ZmPIN5c*, and *ZmPIN8* expression in the aleurone layer at 12 DAP. Both *ZmPIN2* and *ZmPIN8* are highly expressed in the single cell layer of the aleurone **(K,M)**, while no expression of *ZmPIN5c* is detectable in this domain **(L)**. BETL, basal endosperm transfer layer; ESR, embryo surrounding region; cha, maternal chalazal region; pe, pedicel; se, starchy endosperm; al, aleurone; ep, embryo proper; su, suspensor; scu, scutellum. Scale bars: 200 μm in **(J)**; 100 μm in **(A–I)**; 50 μm in **(K–M)**.

Together with *ZmPIN1a*, *ZmPIN1b*, *ZmPIN1c*, *ZmPIN2*, and *ZmPIN9*, *ZmABCB1/BR2* is expressed in maize root apexes, particularly in root epidermis as previously reported (Knoller et al., 2010; McLamore et al., 2010). In *br2* mutant root segments the expression levels of *ZmPIN* transcript are differentially altered, depending on the gene and segment. These observed fluctuations at transcript level can be correlated to the reduced auxin transport that characterizes *br2* root (McLamore et al., 2010). This statement is confirmed by the expression analysis of *ZmPIN* genes in wt and *br2* root treated with NAA and NPA. Indeed, also in B73 and *br2* treated-root the expression levels of *ZmPIN* transcript were differentially altered, depending on the gene and segment. However, observing the ZmPIN1a and DR5-reporter lines, it became evident that both NAA and NPA treatments had a strong effect not only on auxin distribution but also on root tissues differentiation and ZmPIN1a protein localization. All together these evidences indicate that the NAA and NPA treatment (and may be also the *br2* mutation), affect the root tissue differentiation indirectly changing the localization of the *PIN* transcripts and proteins, which appear to have a highly tissue-specific expression in maize root. Interestingly, the combination of *br2* mutation and NAA or NPA treatment has additional effect on the alteration of *ZmPIN* gene expression levels.

This high tissue-specific regulation of *PIN* genes is a further indication that, in spite of a partial overlapping in their transcript localization in the root tissues, maize *PIN* genes might have acquired subfunctionalization during evolution.

Unlike in *Arabidopsis*, where each *PIN* has a specific expression domain, expression analysis revealed that many *ZmPIN* genes are simultaneously expressed in the same organ or tissue. Indeed *ZmPIN1a*, *ZmPIN1b*, *ZmPIN1d*, and *ZmPIN2* mark the development of the gynoecium and the formation of the silk in the female inflorescence, *ZmPIN2*, *ZmPIN5c*, and *ZmPIN8* are expressed, together with *ZmPIN1a*, *ZmPIN1b*, and *ZmPIN1c*, in the basal endosperm transfer layer (BETL) cells. These overlapping expression domains in different tissues and at different developmental stages indicate a great degree of functional redundancy among the members of maize PIN family. To understand the biological meaning behind this redundancy it is necessary to further investigate the different endogenous and exogenous stimuli able to regulate *PIN* gene expression and determine auxin gradients in the different plant tissues.

Another open point concerns auxin distribution in maize root apex: anti-IAA immunolocalization experiments and confocal images of a DR5-reporter line showed contrasting auxin distribution and accumulation patterns. It is well known that these two experimental approaches visualize free-auxin distribution and activity of auxin responsive promoter, respectively (Schlicht et al., 2006). Our results suggest that these approaches might be used together to better analyze auxin distribution and accumulation pattern in plant tissues.

In maize root apex auxin distribution and accumulation patterns differ from that reported for *Arabidopsis*, pointing out the importance of the root cap tip and epidermis in the creation of auxin gradients. Further investigations are needed to elucidate whether these differences in the localization of auxin gradients reflect the different anatomical organization between monocot and dicot roots or vice-versa.

ACKNOWLEDGMENTS

The authors thank N. Carraro and L. Ceccato for helpful discussion during the preparation of the manuscript, Prof. D. Jackson for providing us with the reporter lines, Dr. R. Pilu for the *br2* mutant seeds and A. Garside for checking the English. Authors also thank the two reviewers for their useful comments and suggestions. The work represents the starting point of FIRB – Progetto Giovani of the Italian MIUR.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Traffic_and_Transport/10.3389/fpls.2012.00016/abstract

Figure S1 | Neighbor-joining phylogenetic analysis of PIN auxin efflux carriers from *Zea mays, Arabidopsis thaliana, Medicago truncatula, Oryza sativa, Brachypodium distachyon, Sorghum bicolor, Setaria italica, Populus trichocarpa, Vitis vinifera, Physcomitrella patens,* and *Selaginella moellendorffii.* The phylogenetic analysis between newly identified maize PIN proteins, PIN and PIN-like proteins of eudicots and monocots, summarizes the evolutionary relationships among the 120 members of the PIN protein family.

PIN protein sequences from two herbaceous eudicots (*A. thaliana*, *At*, and *M. truncatula*, *Medtr*), two arboreal eudicots (*P. trichocarpa*, *Ptri*, and *V. vinifera*, *Vvinifera*), four monocots (*O. sativa*, *Os*; *S. bicolor*, *Sb*; *B. distachyon*, *Bdist*, and *S. italica*, *Sitalica*), a bryophyte (*P. patens*, *Pp*), and a lycophyte (*S. moellendorffii*,) were retrieved by family blast searches in the Phytozome v7.0 database (www.phytozome.net/) and were aligned together with newly identified ZmPIN and PIN-like proteins of maize. Alignment were performed with ClustalX 1.81 (Blosum Weight Matrix; Gap Opening Penalty: 5; Gap Extension Penalty: 0.20), and a neighbor-joining phylogenetic tree was prepared with Phylip package using 100 bootstraps. The yeast PIN-like protein AEL1 (NP_593365 from *Schizosaccharomyces pombe*) was included in the analysis as outgroup. Newly identified ZmPIN proteins are named according to the cluster of the *Arabidopsis* and rice PIN family to which they belong. Accession numbers are reported in Table S1.

Figure S2 | Phylogenetic network of PIN auxin efflux carriers from *Zea mays, Arabidopsis thaliana, Oryza sativa, Brachypodium distachyon, Populus trichocarpa, Physcomitrella patens,* and *Selaginella moellendorffii.* PIN protein sequences of *Z. mays* (*Zm*), *A. thaliana* (*At*), *P. trichocarpa* (*Ptri*), *O. sativa* (*Os*), *B. distachyon* (*Bdist*), *P. patens* (*Pp*), and *S. moellendorffii* (*Sm*) were retrieved by family blast searches in the Phytozome v7.0 database (www.phytozome.net/) and aligned with MAFFT algorithm available at the EMBL-EBI website (http://www.ebi.ac.uk/Tools/msa/mafft/). A NeighborNet Network was constructed with SplitsTree4 software. In comparison with the classical phylogenetic tree, this network-based approach allows a richer visualization of the data, taking in account that many events such as hybridization, horizontal gene transfer, recombination, or gene duplication and loss might be involved in evolution of the PIN family across plant species.

Figure S3 | Maximum Likelihood phylogenetic analysis of PIN auxin efflux carriers from *Zea mays, Arabidopsis thaliana, Oryza sativa, Brachypodium distachyon, Populus trichocarpa, Physcomitrella patens,* and *Selaginella moellendorffii.* PIN protein sequences of *Z. mays* (*Zm*), *A. thaliana* (*At*), *P. trichocarpa* (*Ptri*), *O. sativa* (*Os*), *B. distachyon* (*Bdist*), *P. patens* (*Pp*), and *S. moellendorffii* (*Sm*) were retrieved by family blast searches in the Phytozome v7.0 database (www.phytozome.net/) and aligned with MAFFT algorithm available at the EMBL-EBI website (http://www.ebi.ac.uk/Tools/msa/mafft/). PIN proteins were aligned with MAFFT algorithm available at the EMBL-EBI website (http://www.ebi.ac.uk/Tools/msa/mafft/). Phylogeny was inferred with the Molecular Evolutionary Genetic Analysis 5 (MEGA5) program using the Maximum Likelihood method (10,000 bootstraps, Jones–Taylor–Thornton model).

Figure S4 | *ZmPINs* duplication analysis. The upper panel shows the duplicated regions between the 10 maize chromosomes. The map positions of the maize *PIN* genes have been manually reported in the scheme. In the bottom panel the duplications of *ZmPIN1a–ZmPIN1b* (on chromosomes 9 and 5, respectively) and *ZmPIN1b–ZmPIN1c* (on chromosomes 5 and 4, respectively) blocks is depicted. Finally, *ZmPIN5a*, *ZmPIN9*, *ZmPIN8*, and *ZmPIN10a* map on long arm of chromosome 3, in a region highly syntenic to the rice chromosome 1, where *OsPIN5a*, *OsPIN9*, *OsPIN8*, and *OsPIN10a* map too. Duplicated and syntenic blocks have displayed with Synteny Mapping and Analysis Program (SyMAP) v3.4 (Soderlund et al., 2011; http://www.symapdb.org/).

Figure S5 | Maize auxin transporter expression analysis from the Genome-wide atlas of transcription during maize development. *ZmPINs*, *ZmPIN-like* and *BR2/ZmABCB1* expression levels analyzed in 60 distinct tissues representing 11 major organ systems of inbred line B73 (Sekhon et al., 2011). Expression graphs have been download from www.maizegdb.org: the green bars indicate tissues in which the gene is expressed.

Figure S6 | *In situ* hybridization and anti-IAA negative controls. Specific sense RNA probes for each *PIN* genes were used as negative controls in every *in situ* hybridization experiment **(A–D)** while two different negative controls, no EDAC pre-fixation and without anti-IAA primary antibody were used. **(A)** Longitudinal section of a primary root hybridized with a *ZmPIN1a* sense probe. **(B)** Longitudinal section of a primary root apex hybridized with a *ZmPIN1b* sense probe. **(C)** Cross section of primary root hybridized with a *ZmPIN9* sense probe. **(D)** Longitudinal section of a female inflorescences showing the mother cell of the megaspore inside the ovule hybridized with a *ZmPIN1d* sense probe. **(E)** Anti-IAA negative control on longitudinal section of a primary root not pre-fixed with EDAC. **(F,G)** Anti-IAA negative control on longitudinal sections of a primary root hybridized with the secondary antibody only. cc, central cylinder; co, cortex; ep, epidermis, end, endodermis; me, root meristem; p, pericycle; ph, phloem; rc, root cap; x, xylem; st, stylar canal; mmc, megaspore mother cell. Scale Bars: 200 μm in **(A)**; 100 μm in **(B–G)**.

Table S1 | PIN sequences accession numbers. Accession numbers list, common name (where available) and respectively protein sequences of auxin efflux carriers considered in this work for phylogenetic analysis are reported for each species. While *Arabidopsis thaliana* and *Zea mays* PIN and PIN-like sequences were retrieved from TAIR (www.arabidopsis.org) and MaizeGDB (www.maizegdb.org) databases, respectively, PIN protein sequences of the other species were identified by a family blast search in the Phytozome v7.0 database (www.phytozome.net/).

Table S2 | Primer list. Primers used for cDNA and genomic full-length amplification, amplification of probes for *in situ* hybridization and expression analysis in vegetative and reproductive tissues (with relative number of circles considered in each PCR) are listed for each maize *PIN* and *PIN-like* gene.

Table S3 | Maize *PIN* and *PIN-like* map position. Chromosomal position (obtained from www.maize.sequence.org database) of the newly identified maize *PIN* and *PIN-like* sequences.

REFERENCES

Avsian-Kretchmer, O., Cheng, J. C., Chen, L., Moctezuma, E., and Sung, Z. R. (2002). Indole acetic acid distribution coincides with vascular differentiation pattern during *Arabidopsis* leaf ontogeny. *Plant Physiol.* 130, 199–209.

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., and Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* 115, 591–602.

Blakeslee, J. J., Bandyopadhyay, A., Peer, W. A., Makam, S. N., and Murphy, A. S. (2004). Relocalization of the PIN1 auxin efflux facilitator plays a role in phototropic responses. *Plant Physiol.* 134, 28–31.

Blakeslee, J. J., Peer, W. A., and Murphy, A. S. (2005). Auxin transport. *Curr. Opin. Plant Biol.* 8, 494–500.

Blilou, I., Xu, J., Wildwater, M., Willemsen, V., Paponov, I., Friml, J., Heidstra, R., Aida, M., Palme, K., and Scheres, B. (2005). The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* 433, 39–44.

Bomblies, K., Wang, R. L., Ambrose, B. A., Schmidt, R. J., Meeley, R. B., and Doebley, J. (2003). Duplicate FLORICAULA/LEAFY homologs zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. *Development* 130, 2385–2395.

Carraro, N., Forestan, C., Canova, S., Traas, J., and Varotto, S. (2006). ZmPIN1a and ZmPIN1b encode two novel putative candidates for polar auxin transport and plant architecture determination of maize. *Plant Physiol.* 142, 254–264.

De Smet, I., and Jurgens, G. (2007). Patterning the axis in plants – auxin in control. *Curr. Opin. Genet. Dev.* 17, 337–343.

Dubrovsky, J. G., Sauer, M., Napsucialy-Mendivil, S., Ivanchenko, M. G., Friml, J., Shishkova, S., Celenza, J., and Benkova, E. (2008). Auxin acts as a local morphogenetic trigger to specify lateral root founder cells. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8790–8794.

Ellis, C. M., Nagpal, P., Young, J. C., Hagen, G., Guilfoyle, T. J., and Reed, J. W. (2005). AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* 132, 4563–4574.

Felsenstein, J. (1989). PHYLIP – phylogeny inference package. *Cladistics* 5, 164–166.

Forestan, C., Meda, S., and Varotto, S. (2010). ZmPIN1-mediated auxin transport is related to cellular differentiation during maize embryogenesis and endosperm development. *Plant Physiol.* 152, 1373–1390.

Forestan, C., and Varotto, S. (2010). PIN1 auxin efflux carriers localization studies in *Zea mays*. *Plant Signal. Behav.* 5, 436–439.

Forestan, C., and Varotto, S. (inpress). The role of PIN auxin efflux carriers in polar auxin transport and accumulation and their effect on shaping maize development. *Mol. Plant.* doi:10.1093/mp/ssr103

Friml, J. (2010). Subcellular trafficking of PIN auxin efflux carriers in auxin transport. *Eur. J. Cell Biol.* 89, 231–235.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., and Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* 426, 147–153.

Gallavotti, A., Yang, Y., Schmidt, R. J., and Jackson, D. (2008). The Relationship between auxin transport and maize branching. *Plant Physiol.* 147, 1913–1923.

Gaut, B. S., and Doebley, J. F. (1997). DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6809–6814.

Geisler, M., and Murphy, A. S. (2006). The ABC of auxin transport: the role of p-glycoproteins in plant development. *FEBS Lett.* 580, 1094–1102.

Guo, A. Y., Zhu, Q. H., Chen, X., and Luo, J. C. (2007). GSDS: a gene structure display server. *Yi Chuan* 29, 1023–1026.

Heisler, M. G., Ohno, C., Das, P., Sieber, P., Reddy, G. V., Long, J. A., and Meyerowitz, E. M. (2005). Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the *Arabidopsis* inflorescence meristem. *Curr. Biol.* 15, 1899–1911.

Helentjaris, T., Weber, D., and Wright, S. (1988). Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. *Genetics* 118, 353–363.

Hochholdinger, F., Park, W. J., Sauer, M., and Woll, K. (2004). From weeds to crops: genetic analysis of root development in cereals. *Trends Plant Sci.* 9, 42–48.

Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267.

Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537, 39–64.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.

Kerr, I. D., and Bennett, M. J. (2007). New insight into the biochemical mechanisms regulating auxin transport in plants. *Biochem. J.* 401, 613–622.

Kimura, M., and Kagawa, T. (2006). Phototropin and light-signaling in phototropism. *Curr. Opin. Plant Biol.* 9, 503–508.

Knoller, A. S., Blakeslee, J. J., Richards, E. L., Peer, W. A., and Murphy, A. S. (2010). Brachytic2/ZmABCB1 functions in IAA export from intercalary meristems. *J. Exp. Bot.* 61, 3689–3696.

Krecek, P., Skupa, P., Libus, J., Naramoto, S., Tejos, R., Friml, J., and Zazimalova, E. (2009). The PIN-FORMED (PIN) protein family of auxin transporters. *Genome Biol.* 10, 249.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.

Leyser, O. (2005). The fall and rise of apical dominance. *Curr. Opin. Genet. Dev.* 15, 468–471.

McLamore, E. S., Diggs, A., Calvo Marzal, P., Shi, J., Blakeslee, J. J., Peer, W. A., Murphy, A. S., and Porterfield, D. M. (2010). Non-invasive quantification of endogenous root auxin transport using an integrated flux microsensor technique. *Plant J.* 63, 1004–1016.

McSteen, P., Laudencia-Chingcuanco, D., and Colasanti, J. (2000). A floret by any other name: control of meristem identity in maize. *Trends Plant Sci.* 5, 61–66.

Mravec, J., Skupa, P., Bailly, A., Hoyerova, K., Krecek, P., Bielach, A., Petrasek, J., Zhang, J., Gaykova, V., Stierhof, Y. D., Dobrev, P. I., Schwarzerova, K., Rolcik, J., Seifertova, D., Luschnig, C., Benkova, E., Zazimalova, E., Geisler, M., and Friml, J. (2009). Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. *Nature* 459, 1136–1140.

Multani, D. S., Briggs, S. P., Chamberlin, M. A., Blakeslee, J. J., Murphy, A. S., and Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science* 302, 81–84.

Nardmann, J., Ji, J., Werr, W., and Scanlon, M. J. (2004). The maize duplicate genes narrow sheath1 and narrow sheath2 encode a conserved homeobox gene function in a lateral domain of shoot apical meristems. *Development* 131, 2827–2839.

Nardmann, J., and Werr, W. (2006). The shoot stem cell niche in angiosperms: expression patterns of WUS orthologues in rice and maize imply major modifications in the course of mono- and dicot evolution. *Mol. Biol. Evol.* 23, 2492–2504.

Normanly, J. (2010). Approaching cellular and molecular resolution of auxin biosynthesis and metabolism. *Cold Spring Harb. Perspect. Biol.* 2, a001594.

Palme, K., Dovzhenko, A., and Ditengou, F. A. (2006). Auxin transport and gravitational research: perspectives. *Protoplasma* 229, 175–181.

Paponov, I. A., Teale, W. D., Trebar, M., Blilou, I., and Palme, K. (2005). The PIN auxin efflux facilitators: evolutionary and functional perspectives. *Trends Plant Sci.* 10, 170–177.

Peer, W. A., Blakeslee, J. J., Yang, H., and Murphy, A. S. (2011). Seven things we think we know about auxin transport. *Mol. Plant* 4, 487–504.

Petrasek, J., and Friml, J. (2009). Auxin transport routes in plant development. *Development* 136, 2675–2688.

Petrasek, J., Mravec, J., Bouchard, R., Blakeslee, J. J., Abas, M., Seifertova, D., Wisniewska, J., Tadele, Z., Kubes, M., Covanova, M., Dhonukshe, P., Skupa, P., Benkova, E., Perry, L., Krecek, P., Lee, O. R., Fink, G. R., Geisler, M., Murphy, A. S., Luschnig, C., Zazimalova, E., and Friml, J. (2006). PIN proteins perform a rate-limiting function in cellular auxin efflux. *Science* 312, 914–918.

Pilu, R., Cassani, E., Villa, D., Curiale, S., Panzeri, D., Badone, F., and Landoni, M. (2007). Isolation and characterization of a new mutant allele of brachytic 2 maize gene. *Mol. Breed.* 20, 83–91.

Reinhardt, D., Mandel, T., and Kuhlemeier, C. (2000). Auxin regulates the initiation and radial position of plant lateral organs. *Plant Cell* 12, 507–518.

Reinhardt, D., Pesce, E. R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., and Kuhlemeier, C. (2003). Regulation of phyllotaxis by polar auxin transport. *Nature* 426, 255–260.

Richter, S., Anders, N., Wolters, H., Beckmann, H., Thomann, A., Heinrich, R., Schrader, J., Singh, M. K., Geldner, N., Mayer, U., and Jurgens, G. (2010). Role of the GNOM gene in *Arabidopsis* apical-basal patterning–from mutant phenotype to cellular mechanism of protein action. *Eur. J. Cell Biol.* 89, 138–144.

Sabatini, S., Beis, D., Wolkenfelt, H., Murfett, J., Guilfoyle, T., Malamy, J., Benfey, P., Leyser, O., Bechtold, N., Weisbeek, P., and Scheres, B. (1999). An auxin-dependent distal organizer of pattern and polarity in the *Arabidopsis* root. *Cell* 99, 463–472.

Scarpella, E., Barkoulas, M., and Tsiantis, M. (2010). Control of leaf and vein development by auxin. *Cold Spring Harb. Perspect. Biol.* 2, a001511.

Scarpella, E., Marcos, D., Friml, J., and Berleth, T. (2006). Control of leaf vascular patterning by polar auxin transport. *Genes Dev.* 20, 1015–1027.

Schlicht, M., Strnad, M., Scanlon, M. J., Mancuso, S., Hochholdinger, F., Palme, K., Volkmann, D., Menzel, D., and Baluska, F. (2006). Auxin immunolocalization implicates vesicular neurotransmitter-like mode of polar auxin transport in root apices. *Plant Signal. Behav.* 1, 122–133.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., Belter, E., Du, F., Kim, K., Abbott, R. M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S. M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M. J., McMahan, L., Van Buren, P., Vaughn, M. W., Ying, K., Yeh, C. T., Emrich, S. J., Jia, Y., Kalyanaraman, A., Hsia, A. P., Barbazuk, W. B., Baucom, R. S., Brutnell, T. P., Carpita, N. C., Chaparro, C., Chia, J. M., Deragon, J. M., Estill, J. C., Fu, Y., Jeddeloh, J. A., Han, Y., Lee, H., Li, P., Lisch, D. R., Liu, S., Liu, Z., Nagel, D. H., McCann, M. C., SanMiguel, P., Myers, A. M., Nettleton, D., Nguyen, J., Penning, B. W., Ponnala, L., Schneider, K. L., Schwartz,

D. C., Sharma, A., Soderlund, C., Springer, N. M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T. K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J. L., Dawe, R. K., Jiang, J., Jiang, N., Presting, G. G., Wessler, S. R., Aluru, S., Martienssen, R. A., Clifton, S. W., McCombie, W. R., Wing, R. A., and Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115.

Sekhon, R. S., Lin, H., Childs, K. L., Hansey, C. N., Buell, C. R., de Leon, N., and Kaeppler, S. M. (2011). Genome-wide atlas of transcription during maize development. *Plant J.* 66, 553–563.

Shen, C., Bai, Y., Wang, S., Zhang, S., Wu, Y., Chen, M., Jiang, D., and Qi, Y. (2010). Expression profile of PIN, AUX/LAX and PGP auxin transporter gene families in *Sorghum bicolor* under phytohormone and abiotic stress. *FEBS J.* 277, 2954–2969.

Soderlund, C., Bomhoff, M., and Nelson, W. M. (2011). SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* 39, e68.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

Tanaka, H., Dhonukshe, P., Brewer, P. B., and Friml, J. (2006). Spatiotemporal asymmetric auxin distribution: a means to coordinate plant development. *Cell. Mol. Life Sci.* 63, 2738–2754.

Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics* 2, 2.3.

Titapiwatanakun, B., and Murphy, A. S. (2008). Post-transcriptional regulation of auxin transport proteins: cellular trafficking, protein phosphorylation, protein maturation, ubiquitination, and membrane composition. *J. Exp. Bot.* 60, 1093–1107.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G. L., Cooper, D., Coutinho, P. M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J. C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D. R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C. J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., and Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604.

Varotto, S., Locatelli, S., Canova, S., Pipal, A., Motto, M., and Rossi, V. (2003). Expression profile and cellular localization of maize Rpd3-type histone deacetylases during plant development. *Plant Physiol.* 133, 606–617.

Verrier, P. J., Bird, D., Burla, B., Dassa, E., Forestier, C., Geisler, M., Klein, M., Kolukisaoglu, U., Lee, Y., Martinoia, E., Murphy, A., Rea, P. A., Samuels, L., Schulz, B., Spalding, E. J., Yazaki, K., and Theodoulou, F. L. (2008). Plant ABC proteins – a unified nomenclature and updated inventory. *Trends Plant Sci.* 13, 151–159.

Vieten, A., Sauer, M., Brewer, P. B., and Friml, J. (2007). Molecular and cellular aspects of auxin-transport-mediated development. *Trends Plant Sci.* 12, 160–168.

Wabnik, K., Kleine-Vehn, J., Govaerts, W., and Friml, J. (2011). Prototype cell-to-cell auxin transport mechanism by intracellular auxin compartmentalization. *Trends Plant Sci.* 16, 468–475.

Wang, J., Hu, H., Wang, G., Li, J., Chen, J., and Wu, P. (2009). Expression of PIN genes in rice (*Oryza sativa* L.): tissue specificity and regulation by hormones. *Mol. Plant.* 2, 823–831.

Weijers, D., Sauer, M., Meurette, O., Friml, J., Ljung, K., Sandberg, G., Hooykaas, P., and Offringa, R. (2005). Maintenance of embryonic auxin distribution for apical-basal patterning by PIN-FORMED-dependent auxin transport in *Arabidopsis*. *Plant Cell* 17, 2517–2526.

Woll, K., Borsuk, L. A., Stransky, H., Nettleton, D., Schnable, P. S., and Hochholdinger, F. (2005). Isolation, characterization, and pericycle-specific transcriptome analyses of the novel maize lateral and seminal root initiation mutant rum1. *Plant Physiol.* 139, 1255–1267.

Xu, M., Zhu, L., Shou, H., and Wu, P. (2005). A PIN1 family gene, OsPIN1, involved in auxin-dependent adventitious root emergence and tillering in rice. *Plant Cell Physiol.* 46, 1674–1681.

Yang, H., and Murphy, A. S. (2009). Functional expression and characterization of *Arabidopsis* ABCB, AUX 1 and PIN auxin transporters in *Schizosaccharomyces pombe*. *Plant J.* 59, 179–191.

Zazimalova, E., Krecek, P., Skupa, P., Hoyerova, K., and Petrasek, J. (2007). Polar transport of the plant hormone auxin – the role of PIN-FORMED (PIN) proteins. *Cell. Mol. Life Sci.* 64, 1621–1637.

Zazimalova, E., Murphy, A. S., Yang, H., Hoyerova, K., and Hosek, P. (2010). Auxin transporters – why so many? *Cold Spring Harb. Perspect. Biol.* 2, a001552.

Zhao, Y. (2010). Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* 61, 49–64.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 October 2011; accepted: 17 January 2012; published online: 08 February 2012.

*Citation: Forestan C, Farinati S and Varotto S (2012) The maize PIN gene family of auxin transporters. Front. Plant Sci. **3**:16. doi: 10.3389/fpls.2012.00016*

This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 07 February 2012
doi: 10.3389/fpls.2012.00017

Diversification and expression of the PIN, AUX/LAX, and ABCB families of putative auxin transporters in *Populus*

Nicola Carraro[1], Tracy Eizabeth Tisdale-Orr[2], Ronald Matthew Clouse[3], Anne Sophie Knöller[4] and Rachel Spicer[5]*

[1] *Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA*
[2] *Rowland Institute at Harvard, Cambridge, MA, USA*
[3] *Division of Invertebrate Zoology, American Museum of Natural History, New York, NY, USA*
[4] *Department of Mathematics and Computer Science, Philipps University, Marburg, Germany*
[5] *Department of Botany, Connecticut College, New London, CT, USA*

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
Serge Delrot, University of Bordeaux, France
Ranjan Swarup, University of Nottingham, UK

****Correspondence:***
Rachel Spicer, Department of Botany, Connecticut College, 270 Mohegan Avenue, New London, CT 06320, USA.
e-mail: rspicer@conncoll.edu

Intercellular transport of the plant hormone auxin is mediated by three families of membrane-bound protein carriers, with the *PIN* and *ABCB* families coding primarily for efflux proteins and the *AUX/LAX* family coding for influx proteins. In the last decade our understanding of gene and protein function for these transporters in *Arabidopsis* has expanded rapidly but very little is known about their role in woody plant development. Here we present a comprehensive account of all three families in the model woody species *Populus*, including chromosome distribution, protein structure, quantitative gene expression, and evolutionary relationships. The *PIN* and *AUX/LAX* gene families in *Populus* comprise 16 and 8 members respectively and show evidence for the retention of paralogs following a relatively recent whole genome duplication. There is also differential expression across tissues within many gene pairs. The *ABCB* family is previously undescribed in *Populus* and includes 20 members, showing a much deeper evolutionary history, including both tandem and whole genome duplication as well as probable gene loss. A striking number of these transporters are expressed in developing *Populus* stems and we suggest that evolutionary and structural relationships with known auxin transporters in *Arabidopsis* can point toward candidate genes for further study in *Populus*. This is especially important for the ABCBs, which is a large family and includes members in *Arabidopsis* that are able to transport other substrates in addition to auxin. Protein modeling, sequence alignment and expression data all point to ABCB1.1 as a likely auxin transport protein in *Populus*. Given that basipetal auxin flow through the cambial zone shapes the development of woody stems, it is important that we identify the full complement of genes involved in this process. This work should lay the foundation for studies targeting specific proteins for functional characterization and *in situ* localization.

Keywords: auxin, PIN, AUX/LAX, ABCB, *Populus*

INTRODUCTION

Plant development is highly plastic owing to growth via meristems, and this plasticity is fundamental to the ability of plants, as sessile organisms, to adapt to changing environments. Developmental flexibility is particularly important for trees, which can live for thousands of years in the same place, growing massive bodies that must face a multitude of environmental challenges. The plant hormone auxin is well established as a key regulator of plant morphogenesis and in recent years the molecular mechanisms of transport and action have been elucidated. With the publication of the *Populus trichocarpa* genome (Tuskan et al., 2006), new tools to improve our understanding of secondary growth – the type of vascular growth that defines woody plants – became available. *Populus* is not only the dominant model species for woody plant growth, but also a valuable crop for pulp, bioenergy production, and carbon sequestration. Thus, understanding the mechanisms that underlie auxin transport in *Populus* is of interest both in the context of the evolution of plant development and as a means to manipulate plant architecture, biomass production, and fiber quality.

The auxins as a group include several molecules, with the most abundant natural form in plants being indole-3-acetic acid (IAA). Auxin synthesis occurs in young, actively growing tissues including shoot tips, young leaves, and germinating seeds (Ljung et al., 2001a,b), and increasing evidence suggests that synthesis takes place in the roots as well (Ljung et al., 2005). Auxin moves from the sites of production throughout the plant via two routes: long distance transport of conjugated forms in the phloem and short distance transport of "free" (non-conjugated) auxin via polar auxin transport (PAT). By far the better studied route, PAT is a form of active intercellular transport mediated by proteins inserted in the plasma membrane that belong to three distinct families. The *PIN* and *ABCB* families encode efflux proteins (i.e., proteins that facilitate movement out of cells), whereas members of the

AUX/LAX family facilitate auxin entry into cells, along with passive diffusion. PAT is relatively slow (5–20 mm/h; Lomax et al., 1995), saturable and can be impaired by the application of both competitive inhibitors and inhibitors of protein synthesis (Katekar and Geissler, 1980; Sussman and Goldsmith, 1981). This form of transport is considered polar because the protein carriers are often asymmetrically positioned in the plasma membrane such that transport is directional. Transport directionality can then be altered on relatively short timescales in response to repositioning of the protein carriers. Feedback mechanisms also exist such that PAT is often self-reinforcing, with multiple transport proteins themselves being upregulated by auxin (Sauer et al., 2006; Titapiwatanakun and Murphy, 2009).

The PIN proteins have been studied extensively in *Arabidopsis thaliana* (Chen et al., 1998; Luschnig et al., 1998; Müller et al., 1998; Utsuno et al., 1998; Friml et al., 2002a,b, 2003) and show dynamic polar localization at the plasma membrane (PIN1, PIN2, PIN3, PIN7) or in the endoplasmic reticulum (ER) (PIN5, PIN6, PIN8; Mravec et al., 2009; Friml and Jones, 2010). PIN1 was first described as mediating PAT and determining organ outgrowth at the inflorescence (Okada et al., 1991; Gälweiler et al., 1998; Vernoux et al., 2011). Subsequently its role in embryogenesis, vein patterning, vascular development, and root development were established (Friml et al., 2003; Vieten et al., 2005; Scarpella et al., 2006; Petrásek and Friml, 2009). The characterization of PIN genes has been expanded to include the monocotyledons *Zea mays* and *Oryza sativa*, both of which express several PINs thought to be specific to the monocots. In maize, ZmPIN1a, b, and c are responsible for directing auxin transport in the male and female inflorescences and in the floret meristems (Carraro et al., 2006; Wu and McSteen, 2007). They are also involved in endosperm and embryonic development (Forestan et al., 2010) and in the maintenance of phyllotaxy (Lee et al., 2009). The monocot-specific PINs from rice (OsPIN9, OsPIN10a, and OsPIN10b) are highly expressed in adventitious root primordia and pericycle cells at the stem-base, suggesting that they may have evolved to promote adventitious root development (Wang et al., 2009).

Members of the *AUXIN/LIKE AUXIN* (*AUX/LAX*) family in *Arabidopsis* (Bennett et al., 1996; Yemm et al., 2004) are largely responsible for auxin influx, although the protonated form of auxin (IAAH) is able to passively diffuse into cells. The founder member *AUX1* encodes a plasma membrane protein that belongs to the amino acid permease family of proton-driven transporters and functions as an anionic symporter (Swarup et al., 2005; Yang et al., 2006). AUX1-mediated IAA uptake is implicated in gravitropic response, as the agravitropic phenotype of the *aux1* mutant can be phenocopied in wild-type seedlings by applying the auxin influx carrier inhibitor 1-naphthoxyaceticacids (1-NOA) and rescued using the membrane-permeable auxin 1-naphthaleneacetic acid (NAA; Swarup et al., 2001; Yemm et al., 2004). The paralogs of *AUX1*, *LAX1*, *LAX2*, and *LAX3* encode proteins that maintain a correct phyllotactic pattern at the shoot apical meristem (SAM), as they act together with PIN1-mediated auxin efflux (Bainbridge et al., 2008). *LAX3* is also involved in the development of lateral root primordia (Swarup et al., 2008).

The involvement of ABCB [ATP-binding cassette (ABC) transporters of the B class, previously known as multidrug resistance (MDR)/Phosphoglycoprotein (PGP)] proteins in auxin transport was first hypothesized when expression of ABCB1/PGP1 in *Arabidopsis* was found to regulate hypocotyl elongation in a light-dependent fashion (Sidler et al., 1998). Subsequently, ABCB1 was shown to function with ABCB19/PGP19/MDR1 in mediating PAT (Noh et al., 2001). ABCB1 and ABCB19 are the closest *Arabidopsis* orthologs of mammalian ABCB1-type MDR transporters and although specificity for auxin is not assured (Lee et al., 2008), some appear to transport auxin with relatively high substrate specificity (Titapiwatanakun and Murphy, 2009; Yang and Murphy, 2009). ABCB14 and ABCB15 promote auxin transport along the inflorescence of *Arabidopsis*, where they are expressed in vascular tissue and interfascicular fibers. Inflorescence stems in both knockout mutants show a reduction in PAT (Kaneda et al., 2011). ABCB4 from *Arabidopsis* is involved in basipetal PAT in the root (Terasaka et al., 2005; Wu et al., 2007; Kubeš et al., 2011) and, although most ABCBs studied to date function as efflux carriers, heterologous expression of ABCB4 suggests that it functions as an auxin influx carrier under low concentrations of IAA and reverses to efflux when IAA concentrations increase (Yang and Murphy, 2009). The *ABCB1/PGP1* ortholog has been cloned in maize (*Brachytic2/ZmPGP1*) and in *Sorghum bicolor* (*dwarf3/SbPGP1*) and shown to be responsible for IAA transport along the stem (Multani et al., 2003; Knöller et al., 2010).

Our understanding of PAT and its role in development has advanced considerably in *Arabidopsis* and to a lesser extent in monocots, but the functional significance of these transport proteins — particularly the ABCBs — remain largely unknown in woody plants. Woody plants are defined by the production of secondary vascular tissue, specifically secondary xylem and phloem. These vascular tissues are derived from a lateral meristem called the vascular cambium that encircles the stem, adding new cells that will ultimately differentiate into xylem toward the inside of the stem and phloem toward the outside. Given the demonstrated role of PAT in vascular development in herbaceous plants it seems logical to expect a role in secondary growth. Indeed, the vascular cambium contains high levels of IAA in both *Pinus* and *Populus*, with a peak concentration occurring either in the cambial initials themselves, or perhaps more likely, in the earliest differentiating xylem elements (Uggla et al., 1996, 1998; Tuominen et al., 1997; Hellgren et al., 2004). Concentrations rapidly decline through the regions of cell differentiation to near zero in mature secondary xylem and phloem. Auxin transport in the cambium is basipetal (Lachaud and Bonnemain, 1984; Uggla et al., 1998; Kramer et al., 2008) and several members of the *PIN* and *AUX/LAX* gene families are expressed in developing *Populus* stems (Schrader et al., 2003, 2004; Nilsson et al., 2008). Furthermore, expression of one or more *PIN* and *AUX/LAX* genes is downregulated with the onset of dormancy (Schrader et al., 2003, 2004) and upregulated following exogenous application of IAA and/or gibberellins (Schrader et al., 2003; Björklund et al., 2007). Despite several excellent studies in *Populus*, our knowledge of the molecular mechanisms that regulate PAT in woody plants is essentially restricted to the expression patterns of just three *PIN* and *AUX/LAX* genes. A more comprehensive understanding of PAT gene and protein function in *Populus* will help to clarify the molecular mechanisms controlling vascular pattering in woody plants and explain the link(s) between

short and long distance auxin transport in species with extensive stem development.

Here we present the first comprehensive account of the *PIN*, *AUX/LAX*, and *ABCB* gene families in *Populus*, which contain 16, 8, and 20 members respectively. We investigate the history of gene family members relative to each other within *Populus* and relative to proposed orthologs in *Arabidopsis*. Through phylogenetic analysis we describe the timing of the diversification of the *PIN*, *AUX/LAX*, and *ABCB* gene families relative to when plants colonized land. Because the transport function of the ABCB proteins is less understood and their specificity for auxin has not been completely elucidated, we model the protein structures for *Populus* ABCBs and compare these to known *Arabidopsis* ABCB transporters. We then provide expression data for all putative auxin transporters in *Populus*, including presence or absence data for each gene in the cortex, phloem, cambial zone, and xylem of mature stems. We present quantitative RT-PCR expression levels for whole plantlets, internodes just beginning to form secondary vascular tissue, roots and developing xylem from mature stems. Lastly, in order to determine the most likely contributors to the positive feedback mechanism driving "canalization" of auxin flow during vascular development, we test the response of *PIN*, *ABCB*, and *AUX/LAX* genes to exogenous IAA application. These findings should lay the foundation for the functional characterization of members of each family and suggest which proteins are likely to be important regulators of secondary growth.

MATERIALS AND METHODS

PLANT MATERIAL

Populus tremula × *alba* hybrid clone INRA 717-1B4 was chosen for all experimental procedures. *In vitro* plants were grown on half-strength Murashige and Skoog (MS) supplemented with 2% sucrose, 0.25 mg ml^{-1}MES, 0.04 mg ml^{-1} glycine, and 0.2 mg ml^{-1} myo-inositol at 25 ± 2°C under 16 h day length conditions using GE 20W F20T12 growth lamps. Greenhouse plants were grown in 2:1:1 promix HP: perlite:vermiculite supplemented with 19–6–12 N–P–K slow release fertilizer. Greenhouse temperatures were maintained around 22 ± 5°C and day light supplemented to achieve a 16 h day length using metal halide lamps.

IDENTIFICATION OF *PIN*, *AUX/LAX*, AND *ABCB* GENE AND PROTEIN FAMILIES

Populus trichocarpa gene and protein sequences were retrieved from the Joint Genome Institute's (JGI) *P. trichocarpa* v.1.1 database[1]. Henceforth we refer to these genes and gene families as *PtrPIN*, *PtrAUX*, and *PtrABCB*. When reporting expression data, we will refer to the same genes from *P. tremula* × *alba* (abbreviated as *Pta*, i.e., *PtaPIN1*). The *PIN* and *AUX/LAX* sequences had been previously annotated and we maintained the original nomenclature including the *AUX* and *LAX* names for every member of the *AUX/LAX* family from *P. trichocarpa* (i.e., *PtrAUX1–LAX5*). Every sequence was used as query with the BLASTn algorithm to search the National Centre for Biotechnology Information (NCBI) nucleotide collection database to confirm sequence identity. Putative *ABCB* genes in the *P. trichocarpa* genome were identified in the same database using 22 *Arabidopsis ABCB* gene sequences retrieved from the *Arabidopsis* Genome Initiative Research database (TAIR)[2]. The JGI *P. trichocarpa* v.1.1 database was also searched using the terms "MDR" and "ATP" as queries. A third search was conducted using the retrieved sequences to interrogate the *Populus* DataBase (PopulusDB)[3]. Finally all retrieved sequences were confirmed as encoding putative auxin transporters by searching the phytozome v.7.0 database[4]. All the remaining *PIN*, *AUX/LAX* and *ABCB* sequences from other species were retrieved from phytozome v.7.0, TAIR10, The Rice Genome Annotation Project[5], and MaizeGDB[6]. The complete list of retrieved genes is provided in **Table A4** in Appendix. All sequences were inspected for redundancy and presence of pseudogenes and invalid gene models were discarded. ABCB protein sequences were used as queries to search the PROSITE database[7] to confirm the presence of the TMD–NBD–TMD–NBD (transmembrane domain, nucleotide-binding domain) structure and the ABC C-motif. This allowed to rule out the presence of ABC half transporters and other ABC proteins not belonging to class B (Sanchez-Fernandez et al., 2001) and to classify the genes according to their full length structure, conserved motifs, sequence similarity, and EST support. Intron–exon structures of *P. trichocarpa PIN*, *AUX/LAX*, and *ABCB* genes were produced using the online tool GSDS, Gene Structure Display Server (Guo et al., 2007)[8]. The genome representation for *Populus* was created using the online tool SyMAP v.3.5[9]

PtrABCB, PIN, AND AUX/LAX STRUCTURE ANALYSIS AND PtrABCB MODELING

Transmembrane domains were predicted using the online tools TMHMM Server v.2.0[10] and Aramemnon[11]. The protein structure of Sav1866 and MDR1 were obtained from the PDB (Protein Data Bank) database[12]. The predicted protein structures of AtABCB1 and 4 have been previously generated by Yang and Murphy (2009). *Arabidopsis* templates (ABCB1 or 4) were chosen based on closest sequence identity. To generate the alignment files of *Populus* ABCB protein sequences and *Arabidopsis* ABCB sequences, Multialin[13] was used with default settings. The output file was manually edited to meet Modeller 9v5 requirements[14]. The predicted 3D protein structure was generated using the python script Modeller 9v5. Three structures were generated and the quality was determined according to the manual (Wiederstein and Sippl, 2007). The best model was used for substrate docking. Furthermore, the

[1] http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html

[2] www.arabidopsis.org
[3] http://www.populus.db.umu.se
[4] http://www.phytozome.org
[5] http://rice.plantbiology.msu.edu/
[6] http://www.maizegdb.org/
[7] http://ca.expasy.org/prosite/
[8] http://gsds.cbi.pku.edu.cn/index.php
[9] http://www.symapdb.org/.
[10] http://www.cbs.dtu.dk/services/TMHMM/
[11] http://aramemnon.uni-koeln.de/
[12] http://www.rcsb.org/pdb/home/home.do
[13] http://bioinfo.genotoul.fr/multalin/multalin.html
[14] http://salilab.org/modeller/release.html

quality of the protein model was tested using the program ProSA[15]. Substrate docking was performed using MEDOCK[16]. PDB files of all proteins were translated into pdbq files using the PDB2PQR server[17]. For substrate docking prediction, the nucleotide-binding folds (NBFs) were removed. All loops connecting the TMDs were removed to reduce the size of the file. Finally, the pdbq file of IAA was produced with the Dundee PRODRG2 Server (Dolinsky et al., 2004, 2007)[18]. Each run had a docking repeat of five times and four runs were performed, resulting in a total of 20 molecules docked to the protein structure. Protein models were displayed using PyMol[19].

PHYLOGENETIC ANALYSIS

Phylogenic reconstruction was conducted using the coding sequences of 18 species, including 3 monocotyledonous and 10 dicotyledonous plants. Sequences from the green algae *Chlamydomonas reinhardtii* (Merchant et al., 2007) and *Volvox carteri* (Prochnik et al., 2010), the moss *Physcomitrella patens* (Rensing et al., 2008) and the lycopod *Selaginella moellendorffii* (Banks et al., 2011) were also included. For each coding sequence, three types of trees were retrieved from two different alignments. The first alignment was generated in concert with the tree search, a method called "dynamic homology" (Wheeler, 1996). 149, 68, and 245 unaligned coding sequences from the *PIN*, *AUX/LAX*, and *ABCB* families (**Table A4** in Appendix) were read into the phylogenetic program POY v.4.1.2 (Varón et al., 2009) and trees and alignments were searched simultaneously for the least costly sequence alignment and tree topology combination under the parsimony criterion. A second alignment was generated in the program MAFFT (Katoh et al., 2009), where the same sequences were aligned under a gap opening cost of 4 and a gap extension cost of 0.05. This alignment was then input to the program Gblocks v.0.91b (Castresana, 2000; Talavera and Castresana, 2007), which removes regions with multiple gaps and of dubious homology. Gblocks was run with default settings, except that gaps were allowed in all parts of the resulting alignment (such as in cases where one or a few sequences have a clear insertion or deletion). The alignment output by Gblocks was then used for tree searching in POY, where it was read as pre-aligned. Both unaligned and aligned POY tree searches were immediately followed by bootstrap searches, where 100 pseudoreplicates were searched starting with one Wagner tree each. Tree searches were conducted on a parallel computing cluster, using 24 processors searching for a maximum of 6 h of automated searching (in which POY decides on the best combination of builds, swapping, ratchet, and fusing) with dynamic homology and 16 processors for the pre-aligned data. For dynamic homology, in both the tree searches and the bootstrap calculations, the data were divided by the program into seemingly homologous blocks before searching using the command "auto_sequence_partition," which greatly increases search speed. For all POY searches, the costs of transitions, transversions, and insertion/deletion events were the same.

The alignment from Gblocks was also used for a maximum likelihood search in RaxML (Stamatakis et al., 2008) on the CIPRES Science Gateway (Miller et al., 2010)[20]. The alignment was first uploaded and converted to relaxed Phylip format and then tree searches were performed with likelihood bootstrap in which the best tree is reported along with the results of a 100-pseudoreplicate bootstrap calculation. The program was allowed to determine the best model (the GAMMA Model was chosen) and other parameters automatically before tree searching. All trees were visualized and edited using FigTree v.1.3.1[21]

DNA AND RNA ISOLATION AND cDNA SYNTHESIS

Total RNA from whole *in vitro*-grown plantlets, internodes, roots, and developing xylem was extracted using the Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) according to manufacturer's instructions. Aliquots of approximately 100 mg developing xylem tissue were homogenized with a Mini Bead Beater (BioSpec Products Inc., Bartlesville, OK, USA) and stainless steel beads. mRNA from 20 μm-thick frozen sections from the cortex, secondary phloem, cambium, and secondary xylem was extracted using the DynaBeads mRNA Direct Kit (Invitrogen, Carlsbad, CA, USA) according to manufacturer's instructions. DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions using approximately 100 mg fresh leaf tissue. DNA and RNA concentrations were measured with a NanoDrop 2000™(Thermo Scientific, Waltham, MA, USA). Total RNA was treated with TURBO DNA-*free*™(Ambion, Austin, TX, USA) according to manufacturer's instructions. cDNA was synthesized from 1.5 μg of total RNA using SuperscriptII reverse transcriptase (Invitrogen, Carlsbad, CA, USA) with the oligodt$_{20}$ primer. RT-PCR reaction cycles were carried out according to manufacturer's instructions including a final 20 min incubation step with RNAseH (Invitrogen, Carlsbad, CA, USA). cDNA concentration was measured with a Nanodrop 2000™and the cDNA was diluted to 170 ng μl^{-1}.

AMPLIFICATION, CLONING AND SEQUENCING OF 3′ END PCR PRODUCTS

In order to amplify the 3′ end untranslated region (UTR) of transcripts that could not be detected in quantitative real time PCR (qRT-PCR) reactions with at least three different primer pairs, reverse transcription reactions were carried out using the Adp1-dt17 primer (Kramer et al., 1998) and SuperscriptII reverse transcriptase according to manufacturer's instructions. cDNA was amplified using the Adp1 primer coupled to the corresponding forward primer specifically designed to amplify the 3′ end of the transcript (the complete list of primers is provided in **Table A5** in Appendix). The PCR amplifications were carried out with Taq DNA polymerase (SIGMA, St. Louis, MO, USA) or Amplitaq® Gold DNA polymerase (Applied Biosystems™, Foster City, CA, USA) according to manufacturer's instructions. PCR

[15] http://www.came.sbg.ac.at/typo3/index.php?id = prosa
[16] http://medock.csbb.ntu.edu.tw
[17] http://pdb2pqr.sourceforge.net
[18] http://davapc1.bioch.dundee.ac.uk/prodrg,
[19] http://pymol.sourceforge.net

[20] http://www.phylo.org/news/raxml.php
[21] http://tree.bio.ed.ac.uk/software/figtree/.

products were run on 1% agarose gels, gel purified using the Zymoclean™Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA) and cloned into the *pGEM*®-T Easy Vector Systems (Promega, Madison, WI, USA). Colonies were grown on LB plates containing 100 mg/ml ampicillin. Following PCR amplification, positive colonies were grown in 4 ml of LB medium containing 100 mg/ml ampicillin, at 37˚C, over night. Plasmid DNA was extracted using the Qiagen Plasmid Mini Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions. Plasmids were sequenced by Eurofins MWG Operon (Huntsville, AL, USA). Sequences were aligned using the Vector NTI Advance™10.3.0 AlignX module (Invitrogen, Carlsbad, CA, USA).

QUANTITATIVE RT-PCR

Quantitative real time PCR was carried out on the MX3000P and MX3005P systems (Stratagene, La Jolla, CA, USA) using Brilliant™SYBR® Green QPCR Master Mix (Stratagene, La Jolla, CA, USA) according to manufacturer's instructions. The SYBR® Green (with dissociation curve) experimental setup was used. Plates were manually loaded and reactions were carried out in a total volume of 20 µl, using 75 ng of cDNA per reaction. Reactions were run in triplicate. Primer pairs were designed using Primer3 software[22], analyzed with OligoAnalyzer 3.1 software[23] for melting temperature, oligo-, hetero-dimer, and hairpin structure formation, synthesized by Integrated DNA Technologies (IDT, IA) and tested with conventional PCR to verify amplification of a single product. Following primer titration, a final concentration of 250 nM for each primer was chosen. In qRT-PCR experiments the following thermal cycling conditions were used: activation step of 10 min at 95˚C; 40 cycles of 30 s at 95˚C, 25 s at 57˚C, 25 s at 72˚C; fluorescence was collected at the end of each extension step. A melting curve analysis was performed.

Efficiency-corrected expression values were calculated based on standard curves for all genes (Livak and Schmittgen, 2001; Pfaffl, 2001). Standard curves were run in triplicate for every gene in every cDNA batch and amplification efficiencies were calculated from the standard curve slopes. Baseline-subtracted and ROX-normalized fluorescence readings were collected with the MX3005P software v.4.01. Expression values were normalized to the geometric mean of four housekeeping genes (*PtaPD-E1*, *PtaUBQ1*, *PtaTUA2*, *PtaACT2*) that were found, in our hands, to have the highest amplification efficiency and most stable expression across different tissues (Vandesompele et al., 2002; Brunner et al., 2004; Gutierrez et al., 2008). For expression following exogenous IAA application, the same set of normalizers was used in a comparative quantitation experiment comparing treated and untreated control tissues.

IAA TREATMENTS

Two-month-old *P. tremula* × *alba* was grown in the greenhouse. Approximately 1-cm-long segments of internodes between four and eight nodes beneath the shoot apex and actively growing root tips were collected and incubated at room temperature in 30 µM IAA in liquid growth media (half-strength MS salts, 2% sucrose, 0.25 mg ml^{-1} MES, 0.04 mg/ml glycine, and 0.2 mg ml^{-1} myo-inositol) for 6 h in the dark following a 15 min vacuum infiltration. The same conditions were used for negative controls (no IAA). Tissues were frozen in liquid N_2 and ground for RNA extraction.

RESULTS

CHROMOSOMAL DISTRIBUTION AND GENE DUPLICATION IN THE *PIN*, *AUX/LAX*, AND *ABCB* FAMILIES OF *POPULUS*

Nearly every locus coding for a PIN, AUX/LAX, or ABCB protein has a corresponding paralogous locus in another chromosomal block (**Figure 1**). *Populus* has exactly twice the number of *PIN* (16) and *AUX/LAX* (8) genes as *Arabidopsis* (eight and four, respectively) and these genes form pairs with highly similar coding sequences, which may be the consequence of the relatively recent genome duplication (**Figures 1**, **2**, and **3**). Neither the *PIN* loci nor the *AUX/LAX* loci appear to be derived from tandem duplications. In contrast, three tandem duplicated *ABCB* loci pairs (*PtrABCB2–PtrABCB8*, *PtrABCB10–PtrABCB11*, and *PtrABCB13–PtrABCB14*) are present in the *Populus* genome. Unlike the *PIN* and *AUX/LAX* families, the *ABCB* genes are more randomly distributed between corresponding and non-corresponding duplicated regions, with nine members that do not present any paired gene on another chromosome (**Figure 1**).

GENE AND PROTEIN STRUCTURE OF THE *PIN*, *AUX/LAX*, AND *ABCB* FAMILIES OF *POPULUS*

We identified a total of 44 *Populus* genes encoding putative auxin transport proteins, including 16 *PIN*, 8 *AUX/LAX*, and 20 *PtrABCB* loci. The complete list of *P. trichocarpa PIN*, *AUX/LAX*, and *ABCB* gene names, gene models, and loci can be found in **Table A2** in Appendix. The PIN genes of *Populus* present a conserved intron–exon organization which is illustrated in **Figure A1** in Appendix. The same structural characteristics are present across PINs from different plant species including *Arabidopsis* (Mravec et al., 2009; Wang et al., 2009; Shen et al., 2010). The proteins belonging to the PtrPIN family range from 347 to 650 amino acids in length. In *Populus*, seven, three, and six PIN proteins present long, reduced and short central hydrophilic domains respectively. In general, there is no strict correlation between the length of the genomic sequence of loci coding for auxin transporters and their protein product length (**Figure A1** and **Table A3** in Appendix). One locus (*PtrPIN14*) is classified as encoding a pseudogene. The proteins for the PtrAUX/LAX family range from 465 to 492 amino acids and present the most conserved sequence among the three families of putative auxin transporters. Their primary sequence is generally conserved across the plant kingdom and *Populus* has twice the number of *AUX/LAX* coding loci compared to *Arabidopsis*. All of the PtrAUX/LAX proteins have 11 predicted transmembrane domains. All the *ABCB* loci from *P. trichocarpa* encode proteins with a repeated TMD–NBD structure and carry a predicted nucleotide-binding domain signature ([AG]- × (4)-G-K-[ST]; Rea, 2007; Verrier et al., 2008). Their length varies between 1141 and 1578 amino acids and the two regions integral to the plasma membrane are highly hydrophobic and comprise 7–12 transmembrane helices. In addition to these two conserved modules, a more variable and less hydrophobic linker region connects the first NBD to the second TMD in all PtrABCB proteins.

[22] http://frodo.wi.mit.edu/primer3
[23] http://www.idtdna.com/analyzer/Applications/OligoAnalyzer

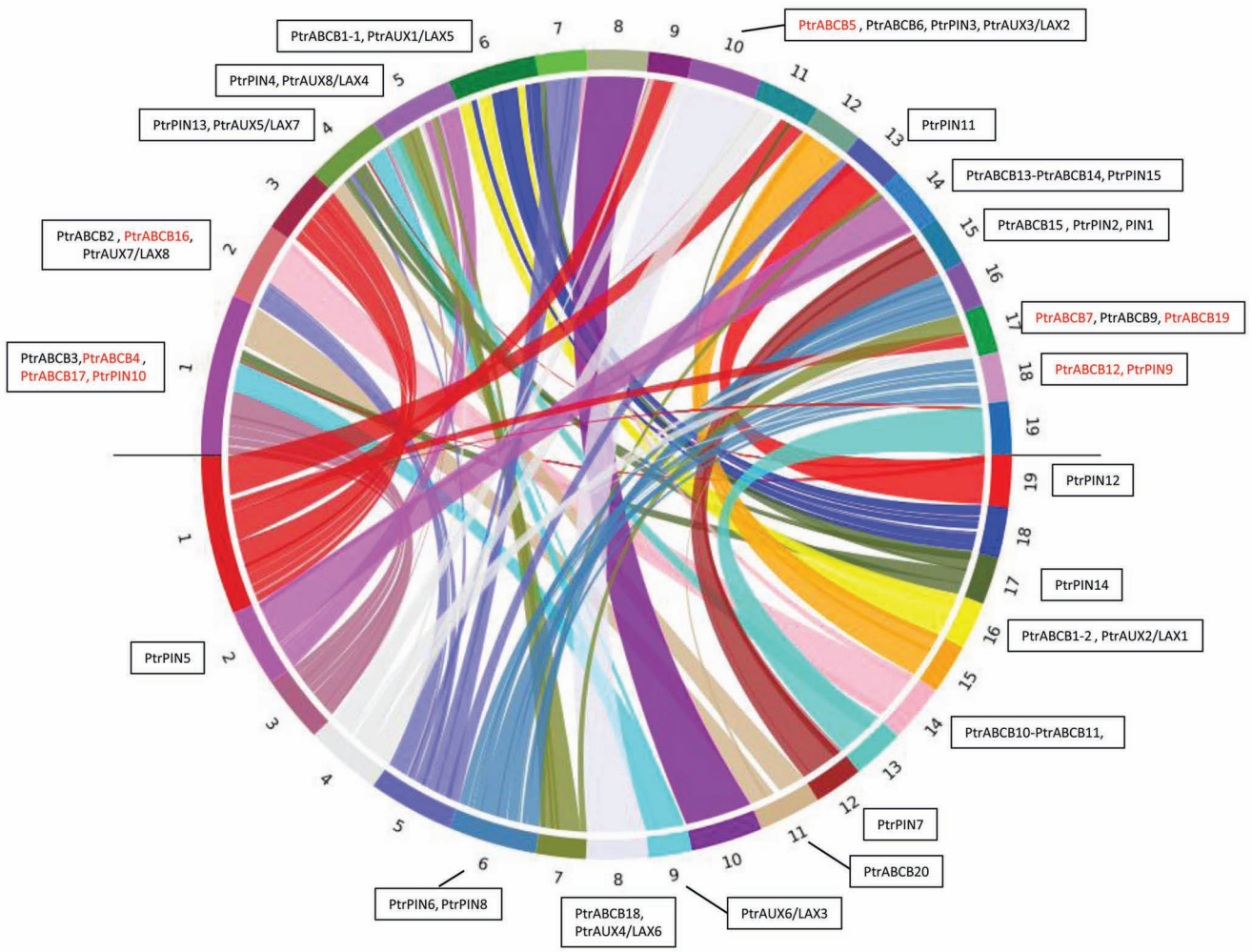


FIGURE 1 | Chromosome distribution of *PtrPIN, PtrAUX/LAX,* and *PtrABCB* genes. The online tool symap v.3.5 was used to blast the *Populus trichocarpa* genome against itself and find duplicated regions. *Populus* has 19 chromosomes in the haploid state, shown here mapped onto a circle with homologous pairs along the upper and lower semi-circumferences. The color coded ribbons link one region with the correspondent homologous chromosomal segments. All *PIN, AUX/LAX,* and *ABCB* genes are assigned to a chromosome based on their map position. Red coded genes do not have any unique match on another locus in the genome. For a detailed list of these genes, see **Table A2** in Appendix.

IDENTIFICATION OF PREDICTED IAA MEMBRANE TRANSPORTERS FROM THE ABCB FAMILY OF *POPULUS*

After analysis of the primary structure of the PtrABCB proteins, models of tertiary structure were produced using all 20 ABCB amino acid sequences. Structural models were displayed using PyMol (**Figure A2** in Appendix) in order to determine which PtrABCBs are the most likely candidates for IAA transport. Although pairwise comparison of amino acid sequences can provide a first estimate of which proteins are the true orthologs of confirmed *Arabidopsis* auxin transporters (AtABCB1, AtABCB19, and AtABCB4), this information should be supported with the identification of IAA docking sites and transmembrane barrel structure predictions (Yang and Murphy, 2009). Among all PtrABCBs, 10 are predicted to have one or more IAA binding sites (**Figure A2** in Appendix). In *Arabidopsis*, IAA is primarily docked at two binding sites in the TMDs of ABCB19 while ABCB4 has a unique additional binding site (Yang and Murphy, 2009). In *Populus*, ABCB1.1/ABCB1.2 and ABCB19 have the most similar sequence to AtABCB1 and AtABCB19 and have two, five, and three predicted binding pockets respectively.

RECONSTRUCTION OF THE PHYLOGENETIC RELATIONSHIPS IN THE *PIN, AUX/LAX,* AND *ABCB* GENE FAMILIES OF *POPULUS*

All three phylogenetic analyses (parsimony using unaligned and aligned sequences and maximum likelihood with aligned sequences) generally resulted in well resolved, reasonable, highly supported trees, indicating considerable phylogenetic signal in the sequence data, which was robust to different methods of analysis. Here we show the trees for all three gene families found under maximum likelihood and the tree found under dynamic homology

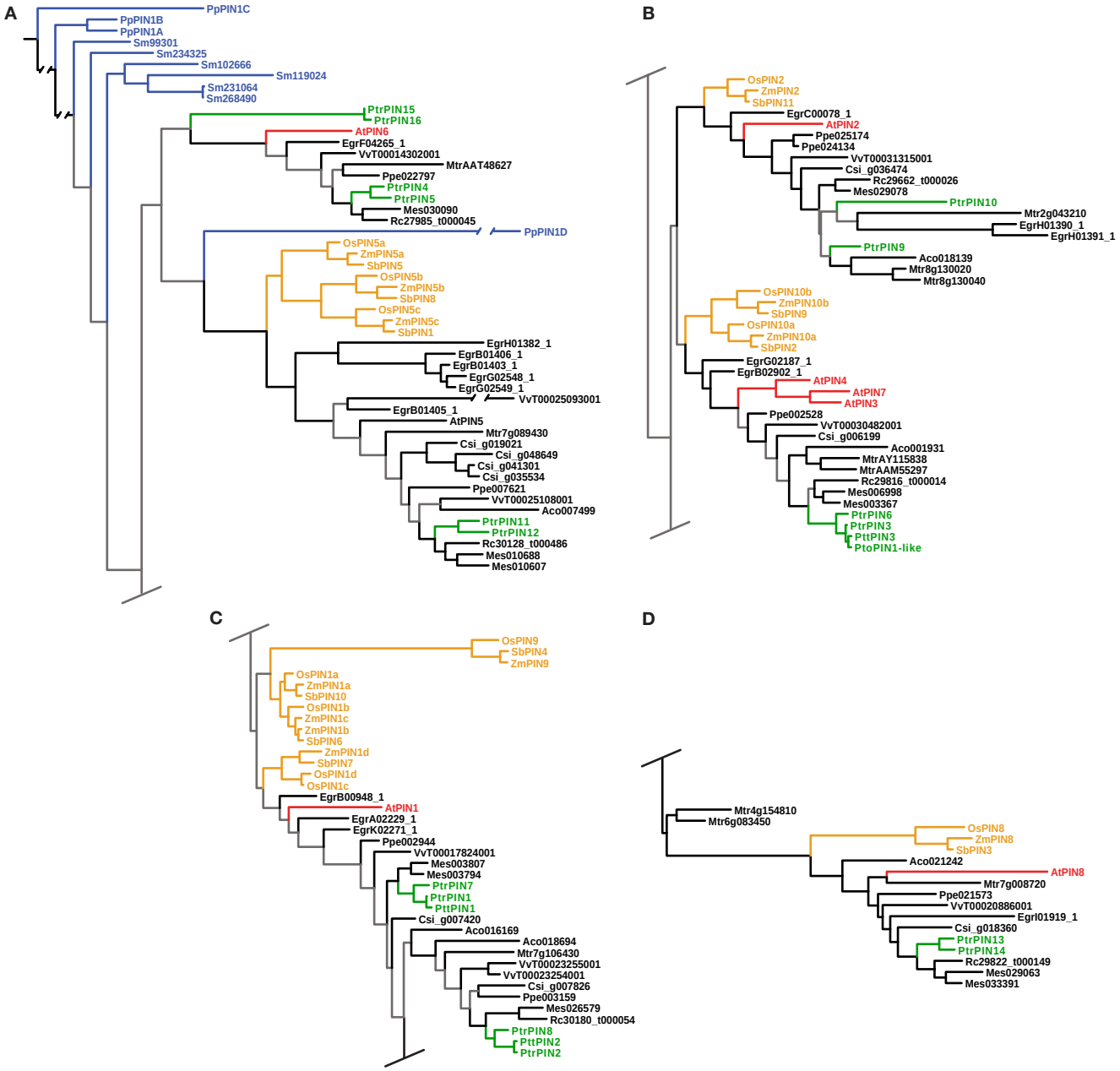


FIGURE 2 | Phylogeny of the *PIN* genes. Maximum likelihood phylogeny of *PIN* genes from land plants, based upon coding sequences from the loci listed in **Table A4** in Appendix. Gray branches indicate nodes with bootstrap support lower than 50%. Basal land plant *PINs* are colored blue, *Populus* green, *Arabidopsis* red, and monocots yellow. Abbreviated names of each species are listed in **Table A1** in Appendix.

and parsimony for the ABCB family (**Figures 2**, **3**, and **4**; **Figure A3** in Appendix). The three different analyses showed the same general patterns in each gene family, although the *PIN* analysis was more sensitive to the difference between likelihood and parsimony, the latter producing long, pectinate clades containing a mixture of taxonomic groups.

The *PIN* genes of basal land plants (*Physcomitrella* and *Selaginella* in our analysis) cluster at the base of the tree, with the exception of *PpPIN1D* (**Figure 2A**). The placement of *PpPIN1D* may indicate an erroneous or highly derived sequence, as its placement was unstable and with low bootstrap support and it was recovered in the likelihood tree on an extremely long branch. The angiosperm *PINs* initially split into two large clades, with subsequent splits that show the monocot/dicot divergence four or five times, although support for several of these nodes is weak (**Figure 2**). There is also the frequent occurrence of clear sister pairs of *PINs* in *Populus*.

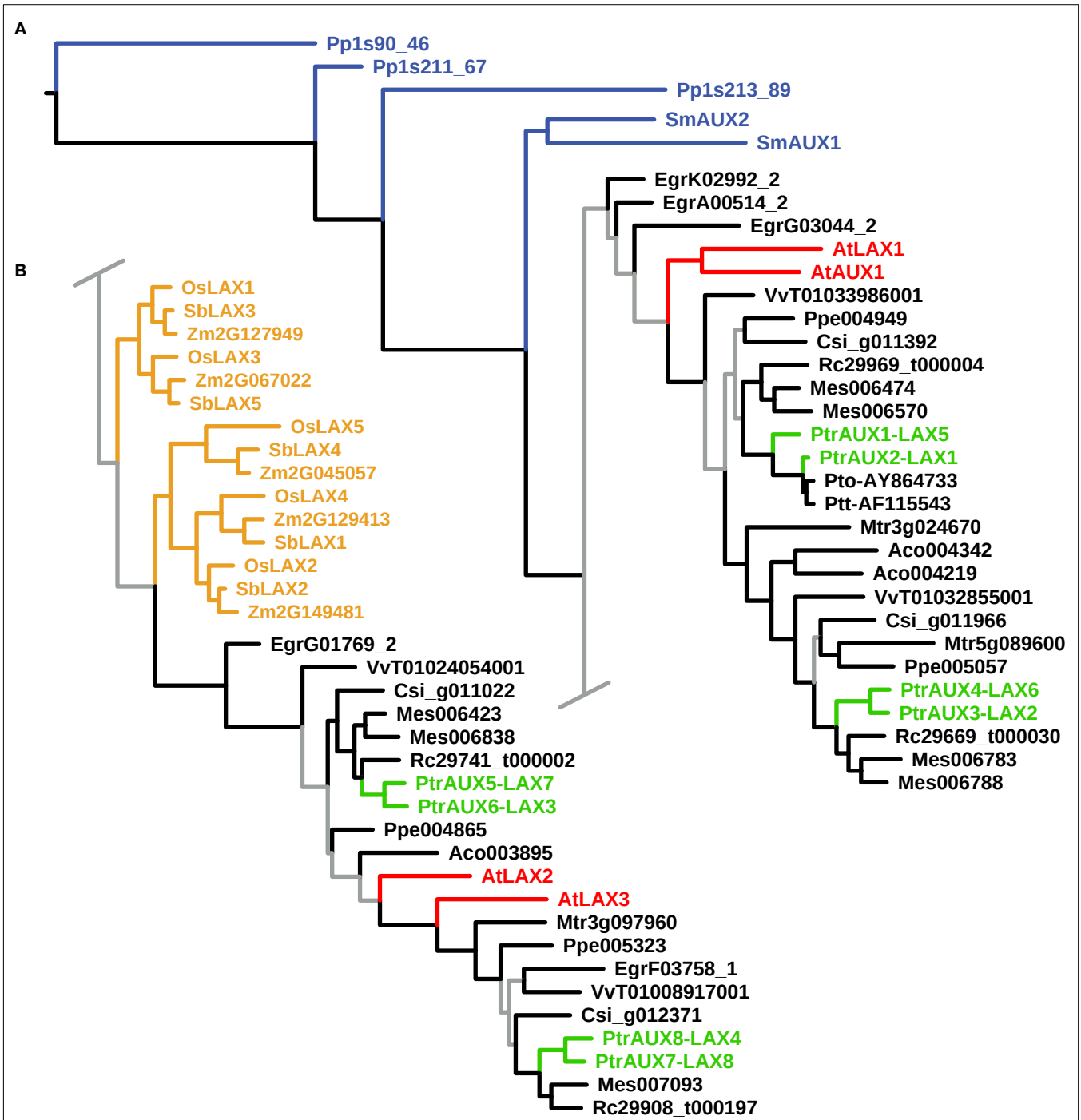


FIGURE 3 | Phylogeny of the *AUX/LAX* genes. Maximum likelihood phylogeny of *AUX/LAX* genes from land plants, based upon coding sequences from the loci listed in **Table A4** in Appendix. Gray branches indicate nodes with bootstrap support lower than 50%. Basal land plant *AUX/LAX* genes are colored blue, *Populus* green, *Arabidopsis* red, and monocots yellow. Abbreviated names of each species are listed in **Table A1** in Appendix.

The *AUX/LAX* analysis similarly places the basal land plant *AUX/LAX* genes in a grade at the base of the tree followed by two large clades of angiosperms (albeit with weak support; **Figure 3**). The monocot *AUX/LAX* genes were recovered as two closely related clades under maximum likelihood (**Figure 3B**) but were recovered as a single clade when the aligned data were analyzed under parsimony (trees not shown). All *Populus AUX/LAX* genes were recovered as sister pairs or, in the case of *PtrAUX1–LAX5* and *PtrAUX2–LAX1*, as closely related in a clade with the *P. tomentosa* and *P. tremula* × *tremuloides AUX/LAXs*.

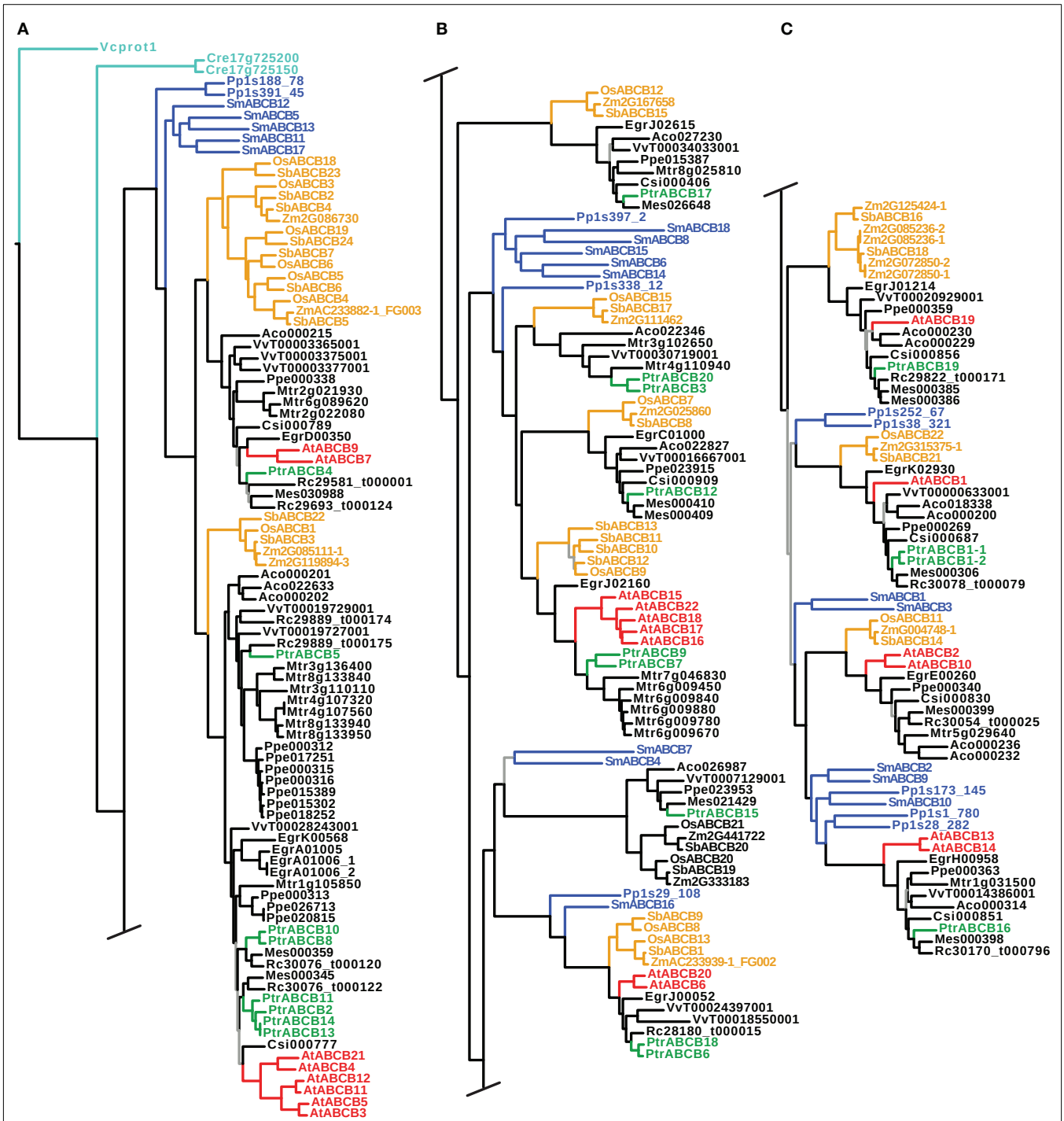


FIGURE 4 | Phylogeny of the *ABCB* genes. Maximum likelihood phylogeny of *ABCB* genes from land plants, based upon coding sequences from the loci listed in **Table A4** in Appendix. Gray branches indicate nodes with bootstrap support lower than 50%. Algal *ABCBs* are colored light blue–green, basal land plants blue, *Populus* green, *Arabidopsis* red, and monocots yellow. Abbreviated names of each species are listed in **Table A1** in Appendix. An alternative phylogeny for the *ABCBs* based on dynamic homology and parsimony, generated with the program POY v.4.1.2, is shown in **Figure A3** in Appendix.

In contrast to the *PIN* and *AUX/LAX* trees, clades, or paraphyletic grades of basal land plant *ABCBs* were recovered in several different locations throughout each tree, often as sister to angiosperm clades that subsequently showed the monocot/dicot split (**Figure 4**). We included coding sequences from the green algae in our *ABCB* analysis: two putative *ABCB* transporters

from *C. reinhardtii* (Cre17_g725200 and Cre17_g725150) and one *ABCB*-like sequence from *V. carteri* (*Vcprot1*), the latter used to root each *ABCB* tree. The inclusion of the algal sequences and the use of *Volvox* as a root appear valid, as they are not recovered on especially long branches, and *Physcomitrella* and *Selaginella* are appropriately placed on the first branches of each tree. In the maximum likelihood tree, we recovered 10 separate clades of monocot *ABCB*s, as well as an apparent expansion of the *ABCB*s in several angiosperm species, including *Medicago truncatula* and *Prunus persica* (**Figures 4A,B**). Among the *Populus ABCBs*, only few were recovered in clear sister pairs. The tree found under dynamic homology for the *ABCBs* recovered almost identical groupings of basal land plant, monocot, and dicot *ABCBs* as those trees found using aligned sequences, but the relationships among these clades or groups differed. For example, a clade containing *OsABCB12* and *Mes026648* (top of **Figure 4B**) was recovered as a paraphyletic grade immediately after the algal sequences in the dynamic homology tree (**Figure A3A** in Appendix).

TISSUE-SPECIFIC AND IAA-INDUCED EXPRESSION OF *PtaPINs*, *PtaAUX/LAXs*, AND *PtaABCBs*

Expression of all *PIN*, *AUX/LAX*, and *ABCB* gene family members in *P. tremula* × *alba* was characterized for whole plantlets, roots, and stem tissues from several developmental stages through qRT-PCR (**Figures 6–8**). Whole *in vitro*-grown plantlets that were old enough to have initiated secondary growth were used as an initial screen and showed that over half of the *PtaPINs* and *PtaAUX/LAX* genes were expressed at above-trace levels, while only four or five *PtaABCBs* showed above-trace expression. Internodes that spanned the region of secondary growth initiation in greenhouse-grown plants should reflect combined expression in several distinct tissues, including cortex, vascular cambium, developing secondary vasculature, and primary xylem parenchyma. Here *PtaPIN1*, *6*, and *PtaABCB1.1* show high expression levels, with lower levels of *PtaPIN7*, *11*, *15*, *16*, and *PtaABCB7* (**Figures 6** and **8**). Developing secondary xylem removed from beneath the bark in 6-month-old greenhouse-grown trees showed high expression of *PtaPIN1* and *PtaABCB1.1*, with lower levels of *PtaABCB7*. Roots showed low expression levels of most genes, which may simply reflect the fact that the roots collected were relatively mature and composed largely of parenchyma, rather than a concentration of actively growing root tips. *PtaAUX/LAX* genes were expressed at relatively uniform levels across all tissues and developmental stages (**Figure 7**), although expression levels were highest for developing xylem, where very high levels of *PtaAUX2* were detected.

In order to perform an expression screen (RT-PCR) with higher spatial resolution in developing woody stems, basal internodes approximately 100 nodes and 2.5 m down from the stem apex of 6-month-old *Populus* were freeze-sectioned and tissue collected from the cortex, secondary phloem, cambial zone (restricted to cambial initials and mother/daughter cells), and secondary xylem. Developing secondary xylem and phloem were discarded in order to obtain the most pure collections of tissues possible. Given that, the number of members of all families that are expressed in each tissue is striking (**Figures 5–8**). Only *PtaPIN9*, *10*, and *12* and *PtaABCB5* and *10* were not expressed in any tissue (**Figures 6** and **8**), and although some of the transcripts detected through RT-PCR are likely expressed at very low levels, it is clear that expression of many previously undescribed members (e.g., *PtaPIN6*, *7*, *15*, and *16* and *PtaABCB1.1* and *7*) is widespread in *Populus* stems. Also striking is the fact that several members of all three transport families are expressed in mature secondary xylem, from which all mRNA is derived from living ray parenchyma cells.

Because a positive feedback mechanism is fundamental to the canalization of auxin flow during vascular development, we also tested the auxin response of members of the *PtaPIN*, *PtaAUX/LAX*, and *PtaABCB* gene families in roots and internodes from 2-month-old plants, following exogenous IAA application, via qRT–PCR. *PtaPIN1*, *2*, and *7* and *PtaAUX5* and *6* were strongly upregulated in developing internodes, with *PtaPIN15* and *16* showing a more moderate increase (**Figure 9**). In contrast, *PtaPIN3* and *8* were strongly upregulated in roots, with *PtaAUX6* and *PtaABCB7* showing a lower expression level.

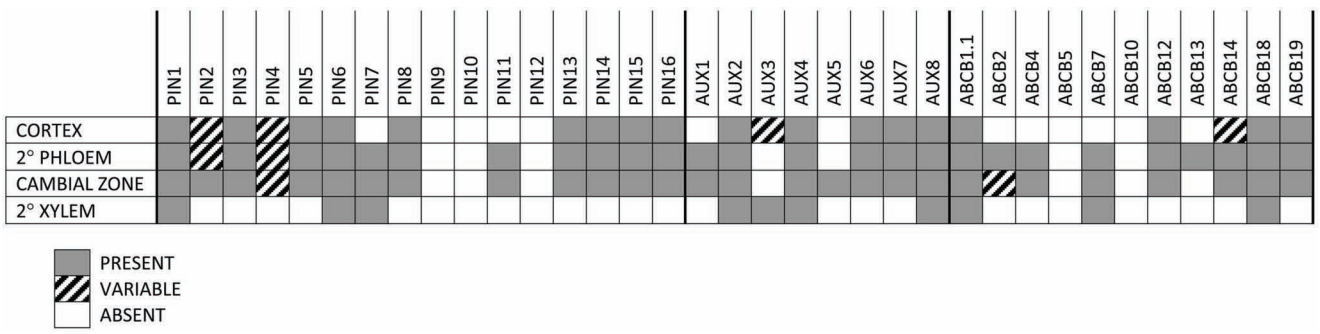


FIGURE 5 | Analysis of tissue-specific expression of *PIN, AUX/LAX,* and *ABCB* transcripts. Presence or absence of transcripts of genes coding for putative auxin transport proteins in the cortex, secondary phloem, cambial zone (i.e., initials and mother/daughter cells), and mature secondary xylem of *Populus tremula* × *alba* as determined by RT-PCR. Consensus of four biological replicates is shown, where GRAY = PRESENT, WHITE = ABSENT, and CROSS-HATCHED = VARIABLE among biological replicates. Samples were taken from the base of 6-month-old trees during active growth, approximately 100 internodes down from the top of the tree at a diameter of about 2 cm.

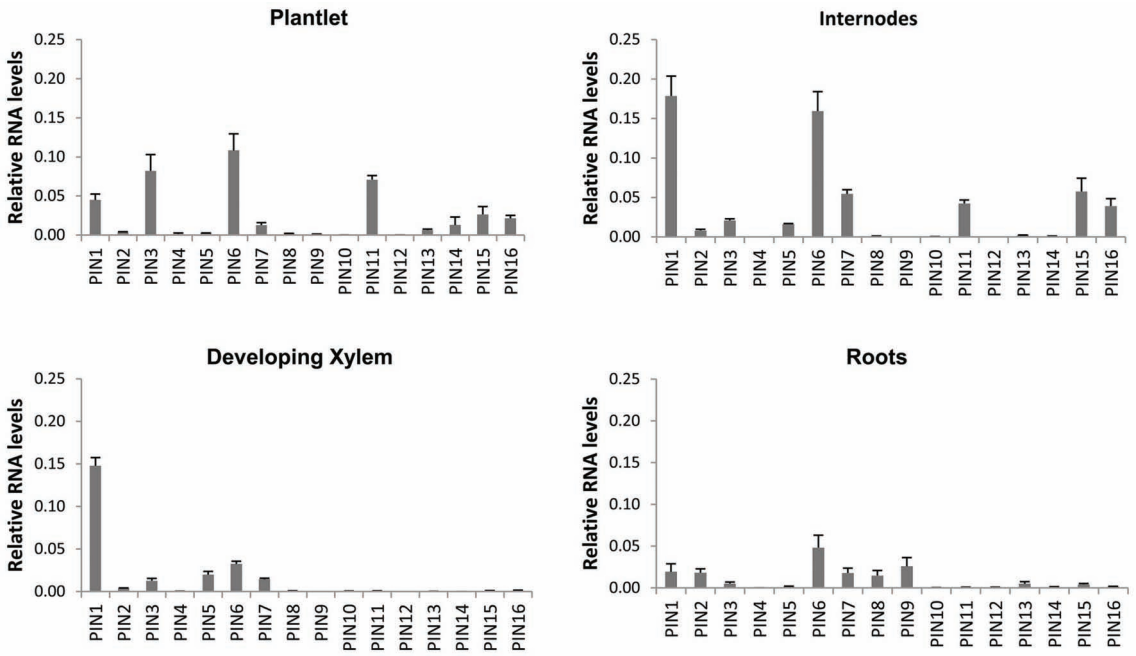


FIGURE 6 | Quantification of *PIN* transcripts expression by qRT-PCR. *PIN* genes show tissue-specific expression profiles that may reflect a role in directional auxin transport in developing vasculature, with *PtaPIN1* highly expressed across all tissues. *PtaPIN6, 7, 15,* and *16* were expressed in internodes and have not been described before. Total RNA was extracted from four biological replicates and qRT-PCR standard curves and assays were run in triplicate. Expression values were calculated via the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001; Pfaffl, 2001) and baseline-corrected fluorescence values were normalized against the geometric mean of *PtaPD-E1, PtaTUA2, PtaUBQ, PtaACT2*. These reference genes were stably expressed across all tissues with the exception of developing xylem; this means that it is permissible to compare expression levels within any single tissue as well as across whole plantlets, internodes, and roots. Error bars represent the SEM.

DISCUSSION

THE ARRAY OF PUTATIVE AUXIN TRANSPORTERS IN *POPULUS* REFLECTS BOTH PRE-EXISTING DIVERSITY AND EXPANSION DUE TO GENOMIC AND SEGMENTAL DUPLICATIONS

There are twice as many members of the *PIN* and *AUX/LAX* gene families in *Populus* as there are in *Arabidopsis* and both families show a number of clear pairs based on coding sequence (e.g., *PtrPIN4/5*, *PtrAUX3/4*; **Figures 2** and **3**). With no clear evidence for any tandem duplication in the *PIN* and *AUX/LAX* gene families, it is possible that all gene copies were retained following the "salicoid" genome duplication (Tuskan et al., 2006). Although the functional role of these proteins has not been demonstrated in *Populus*, given the conserved protein structure and known specificity for IAA for most PINs in *Arabidopsis* (and to a lesser extent, AUX/LAX proteins), it seems likely that they have retained a function in auxin transport. To what extent new PINs have developed specialized roles in PAT in *Populus* is not known and the added redundancy for such an important developmental mechanism may be beneficial enough to warrant retention. Indeed, redundancy in *Arabidopsis* allows single PIN mutants to complete embryogenesis, whereas quadruple mutants are required before severe defects are observed (Benková et al., 2003; Friml et al., 2003). At the same time it is interesting to note that there are clear differences in expression among presumed paralogs. For instance, *PtaPIN1* is expressed at much higher levels than *PtaPIN7* in internodes and developing xylem. Predictions about PIN function in *Populus* may also be informed by structural comparisons with *Arabidopsis*. The "long" PINs in *Arabidopsis* are localized to the plasma membrane and function in PAT, whereas those with shorter structure are found in the ER (Mravec et al., 2009; Friml and Jones, 2010). PtrPIN1–3 and PtrPIN6–9 are all classified as "long" PINs (**Table A3** in Appendix), but it is not known whether similar localization patterns exist in *Populus*.

In contrast to the *PIN* and *AUX/LAX* gene families, the number of *ABCBs* in *Populus* is not expanded relative to *Arabidopsis* (both species include about 20 members; **Table A2** in Appendix) and only a few appear as closely related gene pairs. This is perhaps not surprising given that this gene family has a much deeper history and that ABCB proteins transport a number of substrates in addition to IAA. There also appears to be expansion in a number of angiosperms included in our phylogeny, such as *Z. mays*, *M. truncatula*, *P. persica*, and *Arabidopsis* (**Figure 4**). Although there has been retention of *ABCB* copies from both tandem duplication and whole genome duplication events in *Populus*, there also appears to have been loss. Much functional work is needed on *Populus* ABCB genes and proteins before any role in PAT can be ascribed.

CANDIDATE ABCBs FOR IAA TRANSPORT FUNCTION IN *POPULUS* ARE SUGGESTED BY PHYLOGENETIC PLACEMENT AND PROTEIN STRUCTURE PREDICTION

ATP-binding cassette proteins constitute a very large superfamily that has representatives across the bacteria, plant, and animal

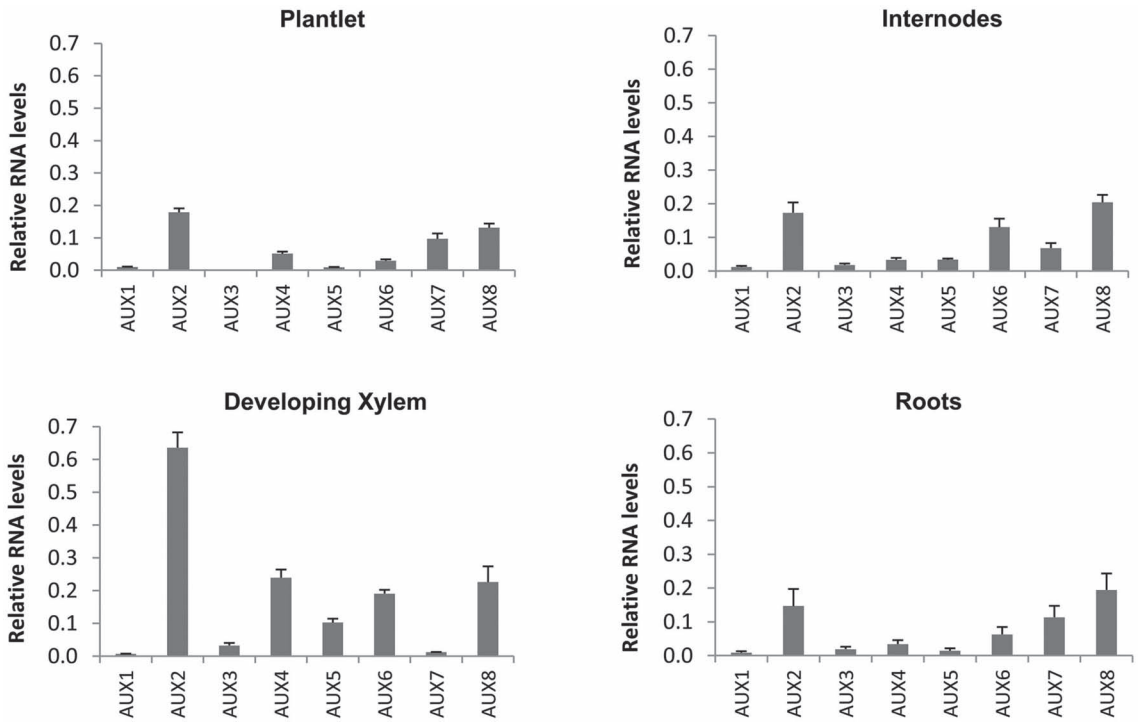


FIGURE 7 | Quantification of *AUX/LAX* transcripts expression by qRT-PCR. Most *AUX/LAX* transcripts showed broad expression across plant tissues, including the previously undescribed *PtaAUX4–8*. *PtaAUX2* and *PtaAUX8* were highly expressed in internodes and developing xylem. Error bars represent the SEM.

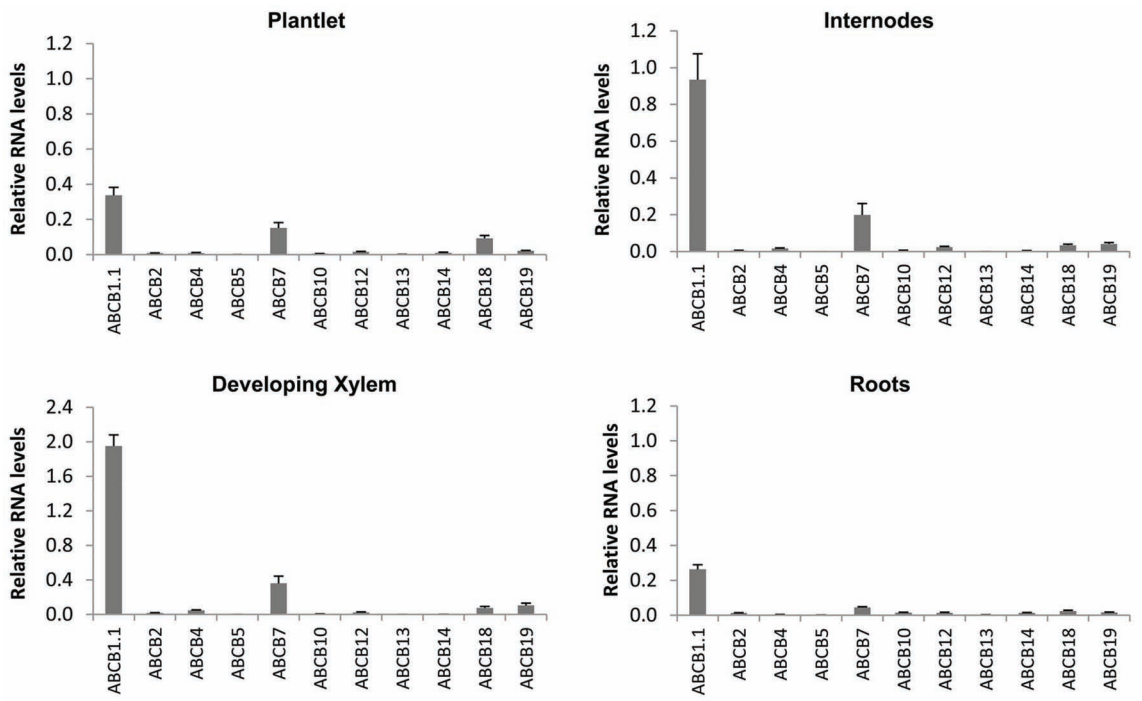


FIGURE 8 | Quantification of *ABCB* transcripts expression by qRT-PCR. Most notable among the *ABCB* family is *PtaABCB1.1*, which was highly expressed in internodes and developing xylem and whose ortholog in *Arabidopsis* (*AtABCB1*) has been demonstrated to transport auxin. Expression patterns of all PtaABCB genes are previously undescribed. Error bars represent the SEM.

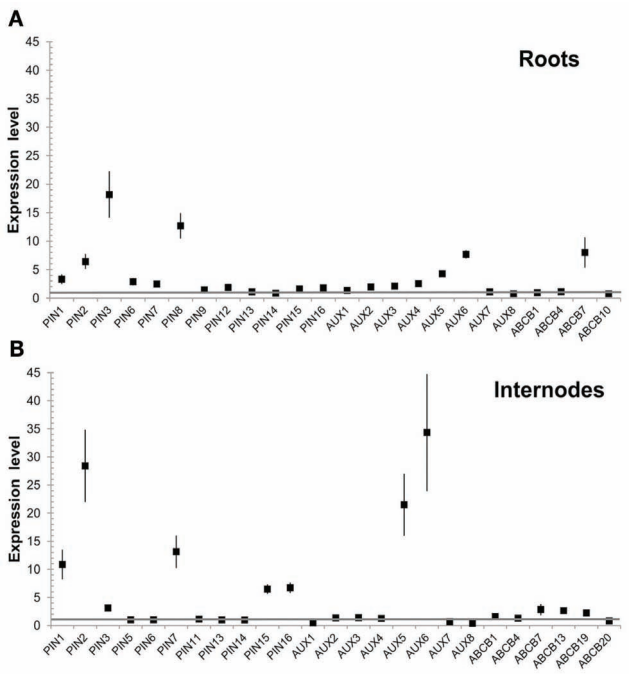


FIGURE 9 | Upregulation of putative auxin transporters expression following IAA treatment. Several *PtaPIN, PtaAUX/LAX,* and *PtaABCB* genes showed increased transcript levels in response to exogenous IAA in both roots and internodes. Two-month-old *Populus tremula* × *alba* were grown in the greenhouse and root tips and internodes were collected and incubated at room temperature in liquid growth media with or without 30 μM IAA for 6 h in the dark. Assays were run in triplicates. Bars represent SEM. Gene expression in the IAA treated tissue is reported relative to the untreated tissue according to the comparative quantitation methodology.

kingdoms (Jasinski et al., 2003; Verrier et al., 2008) and, as a group, are able to transport a wide array of different molecules (Geisler et al., 2005; Bandyopadhyay et al., 2007). Among the ABCs, the subclass B includes proteins that are able to bind and transport auxin across the plasma membrane in *Arabidopsis*, whereas other members transport other substrates in addition to IAA (e.g., AtABCB14 functions primarily as a malate transporter (Lee et al., 2008)). There has been no functional characterization of the ABCBs in *Populus* to date and given the large size of the family and the likely role of one or more members in IAA transport, we sought to identify candidate PtrABCBs with this function. Our phylogenetic analysis shows that the coding sequences of *PtrABCB1.1*, *PtrABCB1.2*, and *PtrABCB19* cluster together with *AtABCB1* and *AtABCB19* respectively, both of which are known IAA transporters with high specificity for IAA (Zazímalová et al., 2010). Interestingly, although 10 of the 20 PtaABCBs are predicted to have one or more IAA binding sites based on tertiary structure, both PtrABCB1 and PtrABCB19 have only one clearly defined binding pocket for IAA. All but one of the remaining ABCBs with putative IAA binding sites (PtrABCB2, PtrABCB5, PtrABCB6, PtrABCB8, PtrABCB11, PtrABCB14) cluster together in the same clade, which includes *AtABCB4*, a gene coding codes for another membrane protein capable of IAA transport (Terasaka et al., 2005; Kubeš et al., 2011). Similarly, *PtrABCB16* occurs in the same clade as *AtABCB13* and *AtABCB14*, where AtABCB14 has been recently determined as responsible for auxin transport in the inflorescence stem of *Arabidopsis* (Kaneda et al., 2011).

We found *PtrABCB1.1* to be highly expressed in most *Populus* tissues, particularly in internodes and developing xylem. *PtrABCB7* was also expressed in these same tissues and was strongly upregulated in response to IAA, although most notably in roots. However, although coding sequence similarity places *PtrABCB7* as a close relative of a presumed IAA transporter in *Arabidopsis* (AtABCB15; Kaneda et al., 2011), the protein was not predicted to contain an IAA binding site. We suggest therefore that PtrABCB1.1 and its nearly identical paralog PtrABCB1.2 are the most logical candidates for initial functional characterization, both in heterologous expression systems (e.g., *Schizosaccharomyces pombe*) and in planta, given their phylogenetic placement relative to *AtABCB1* and predicted IAA binding sites. It is interesting to note that in contrast to *AtABCB1* (Geisler et al., 2005), we did not find *PtaABCB1.1* to be upregulated by exogenous IAA treatment. Lastly, we did not observe strong expression of *PtaABCB19* in any *Populus* tissues nor was it upregulated by IAA. The expression of its presumed ortholog in *Arabidopsis*, *AtABCB19*, is induced by IAA treatments (Noh et al., 2001) and the protein often co-localizes with AtPIN1 (Bandyopadhyay et al., 2007), suggesting that the relationship of these two proteins may have changed. Clearly there is much to be learned about the role of these ABCBs in IAA transport in *Populus*.

AUXIN TRANSPORTERS IN *POPULUS* STEM DEVELOPMENT

That auxin regulates vascular development in woody plants is clear, but our understanding of the genetic mechanisms and the role of specific proteins in basipetal transport is limited. The expression of *PttPIN1–3* and *PttLAX1–3* has already been characterized in detail across the developing stem tissues of *P. tremula* × *tremuloides* (Schrader et al., 2003), but our results suggest that a far greater number of putative transporters are expressed in young internodes where cambial growth is being initiated. In particular, *PtaPIN1*, *PtaPIN6*, and *PtaABCB1.1* are highly expressed in internodes, a complex tissue that includes primary xylem parenchyma, primary phloem, cortex, and a nascent vascular cambium. In developing xylem, *PtaPIN1*, *PtaAUX2*, and *PtaABCB1.1* are highly expressed, with the latter likely to function in auxin transport given its protein sequence similarity to *AtABCB1*. Similarly, several previously uncharacterized transporters are strongly upregulated by auxin, including *PtaPIN8*, *PtaAUX6*, and *PtaABCB7* in roots and *PtaPIN7*, *PtaPIN15*, *PtaPIN16*, *PtaAUX5*, and *PtaAUX6* in internodes. Given the retention of copies of auxin transporters following duplication events, there is likely to be both redundancy and neo-functionalization for PAT proteins in *Populus*.

The vascular cambium and the secondary xylem and phloem that it produces are often viewed as distinct from primary growth, but it is important to remember that vascular development forms a continuum between stem and leaf (Spicer and Groover, 2010). We know a great deal about the role of PAT in venation patterning in leaves of *Arabidopsis* (Scarpella et al., 2006). Here, AtPIN1 directs auxin flow up through the epidermis toward a convergence point, from where it is channeled down through the center of a developing leaf primordium, establishing the location of the

first central vascular bundle. This vascular bundle differentiates from a strand of procambium that is continuous with the vascular cambium below, such that the basipetal transport of auxin out of developing primordia is likely continuous with the basipetal stream moving down through the cambium (Lachaud and Bonnemain, 1984; Uggla et al., 1998; Kramer et al., 2008). Based on a combination of our results and published work in both *Arabidopsis* and *Populus*, we suggest that PtaPIN1, PtaAUX2, and PtaABCB1.1 are the best initial candidates for the maintenance of PAT in the cambial zone, although additional transporters are very likely involved. Given the slow time course and laborious nature of transformation in woody plants, our hope is that this work will provide a starting point for work in planta by identifying candidate IAA transporters involved in woody stem development. Functional studies, transport assays and protein localization are all needed to resolve the action of specific transporters in shaping the distribution of auxin across the cambial zone.

Finally, it is interesting to note that several members of the *PIN*, *AUX/LAX*, and *ABCB* gene families are expressed in the mature xylem. Although the bulk of this tissue is dead (e.g., vessels and fibers), ray parenchyma cells remain alive for many years (Spicer and Holbrook, 2007) and serve as a route of transport between xylem and phloem (Van Bel, 1990). In particular, *PtaPIN1*, *PtaAUX2*, *PtaAUX3*, *PtaAUX4* and *PtaABCB1*, *PtaABCB7*, *PtaABCB20* were found to be expressed in these cells. In addition to their role in carbohydrate transport and storage, xylem parenchyma cells are able to exchange solutes with the transpiration stream and function in wound response. What is puzzling however is that these cells are symplasmically connected, at least in the radial direction, whereas PAT requires transport across a membrane. Furthermore, there is no evidence for free IAA in mature xylem (Uggla et al., 1996; Tuominen et al., 1997). Although conjugated forms of IAA are transported in the phloem (Baker, 2000) no studies to date have looked for conjugated IAA in ray or axial parenchyma in secondary xylem. Given their role in wound response, some capacity for IAA transport (or even IAA synthesis) would not be surprising, but transport assays and protein localization are needed to clarify any potential role these cells might play in IAA transport.

THE *ABCB* GENE FAMILY DIVERSIFIED PRIOR TO THE *PIN* AND *AUX/LAX* FAMILIES AND PRIOR TO THE DIVERSIFICATION OF LAND PLANTS

It is clear from our phylogenetic analysis that the *ABCB* gene family existed before the diversification of land plants, whereas the *PIN* and *AUX/LAX* families arose within the land plant clade. This is supported by the fact that *ABCB* genes from a moss (*P. patens*) and a lycopod (*S. moellendorffii*) consistently occur nested within multiple, well-supported clades that also include higher plants (**Figure 4**; **Figure A3** in Appendix). It also confirms previous work reconstructing the evolutionary history of this family (Bandyopadhyay et al., 2007; Krecek et al., 2009). In contrast, diversification of the *PIN* and *AUX/LAX* gene families occurred after the origin of land plants, as suggested by the well-supported and exclusively basal position of both *Physcomitrella* and *Selaginella PIN* and *AUX/LAX* genes (**Figures 2** and **3**). There was already considerable diversity in the *ABCB* gene family at the time of the monocot/dicot divergence, dated at approximately 130–150 Myr ago (Wolfe et al., 1989; Chaw et al., 2004; Bell et al., 2010), as we recovered as many as 10 distinct *ABCB* gene clades that contain a clear monocot/dicot split with strong support. The picture is not as clear for the *PIN* and *AUX/LAX* genes due to weak support at some nodes, but there may have been five copies of the *PIN* and likely just two copies of the *AUX/LAX* genes at the time of the monocot/dicot divergence. It is not clear at this time whether all *AUX/LAX* genes in monocots descended from a single original copy, as suggested by the tree found using aligned sequences under parsimony, since monocot *AUX/LAX* genes were not recovered in a single clade in other trees (**Figure 3**).

In conclusion, we show that the deep history of the ABCB family of transporters coupled with the expansion of the PIN and AUX/LAX families following a genome duplication has led to a diverse array of over 40 putative auxin transport proteins in *Populus*. Given this large number and the inherent difficulties in working with a woody plant (e.g., long generation times, slow transformation process, difficult nucleic acid extraction), it is important to establish a comprehensive picture of gene expression profiles and predict their protein structures. By considering both evolutionary relationships and structural similarities to known auxin transporters, we can choose the most appropriate candidates for future study. One of the main goals in the short term should be to develop a set of tools for protein localization, including antibodies and protein fusions for stable plant transformation. Although technically difficult for trees, these findings should be coupled with functional studies with knockout mutants. Lastly, it will be important to determine the transport capacity and substrate specificity of target proteins of *Populus* by expressing them in heterologous systems such as *S. pombe*. We hope that this work provides a foundation on which to build an improved understanding of auxin transport in *Populus*, as knowing the role of specific transport proteins in secondary vascular development is likely key to enhanced utilization of woody plants.

ACKNOWLEDGMENTS

The authors would like to thank the laboratories of Noel M. Holbrook and Elena M. Kramer (Harvard University, OEB) for providing space and access to equipment, technical support, and for helpful discussion. The authors are also grateful to Angus S. Murphy and Wendy A. Peer (Purdue University) for helpful discussion of the manuscript; Serena Varotto and Cristian Forestan for sharing sequences and for helpful discussion. This work was supported by a Rowland Junior Fellowship awarded to Rachel Spicer from 2007 to 2010.

REFERENCES

Bainbridge, K., Guyomarc'h, S., Bayer, E., Swarup, R., Bennett, M., Mandel, T., and Kuhlemeier, C. (2008). Auxin influx carriers stabilize phyllotactic patterning. *Genes Dev.* 22, 810–823.

Baker, D. A. (2000). Long-distance vascular transport of endogenous hormones in plants and their role in source: sink regulation. *Isr. J. Plant Sci.* 48, 199–203.

Bandyopadhyay, A., Blakeslee, J. J., Lee, O. R., Mravec, J., Sauer, M., Titapiwatanakun, B., Makam, S. N., Bouchard, R., Geisler, M., Martinoia, E., Friml, J., Peer, W. A., and Murphy, A. S. (2007). Interactions of PIN and PGP auxin transport mechanisms. *Biochem. Soc. Trans.* 35, 137–141.

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., DePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., DeBarry, J., Dreyer, I., Elias, M., Engstrom, E. M., Estelle, M., Feng, L., Finet, C., Floyd, S. K., Frommer, W. B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K. G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T. P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T., Mueller-Roeber, B., Nelson, D. R., Obara, M., Oguri, Y., Olmstead, R. G., Onodera, N., Petersen, B. L., Pils, B., Prigge, M., Rensing, S. A., Riaño-Pachón, D. M., Roberts, A. W., Sato, Y., Scheller, H. V., Schulz, B., Schulz, C., Shakirov, E. V., Shibagaki, N., Shinohara, N., Shippen, D. E., Sørensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theissen, G., Ulvskov, P., Wakazuki, S., Weng, J. K., Willats, W. W., Wipf, D., Wolf, P. G., Yang, L., Zimmer, A. D., Zhu, Q., Mitros, T., Hellsten, U., Loqué, D., Otillar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D., and Grigoriev, I. V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963.

Bell, C. D., Soltis, D. E., and Soltis, P. S. (2010). The age and diversification of the angiosperms re-revisited. *Am. J. Bot.* 97, 1296–1303.

Benková, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertová, D., Jürgens, G., and Friml, J. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* 115, 591–602.

Bennett, M. J., Marchant, A., Green, H. G., May, S. T., Ward, S. P., Millner, P. A., Walker, A. R., Schulz, B., and Feldmann, K. A. (1996). *Arabidopsis* AUX1 gene: a permease-like regulator of root gravitropism. *Science* 273, 948–950.

Björklund, S., Antti, H., Uddestrand, I., Moritz, T., and Sundberg, B. (2007). Cross-talk between gibberellin and auxin in development of Populus wood: gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. *Plant J.* 52, 499–511.

Brunner, A. M., Yakovlev, I. A., and Strauss, S. H. (2004). Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol.* 4, 14. doi:10.1186/1471-2229-4-14

Carraro, N., Forestan, C., Canova, S., Traas, J., and Varotto, S. (2006). ZmPIN1a and ZmPIN1b encode two novel putative candidates for polar auxin transport and plant architecture determination of maize. *Plant Physiol.* 142, 254–264.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.

Chaw, S.-M., Chang, C.-C., Chen, H.-L., and Li, W.-H. (2004). Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* 58, 424–441.

Chen, R., Hilson, P., Sedbrook, J., Rosen, E., Caspar, T., and Masson, P. H. (1998). The *Arabidopsis thaliana* AGRAVITROPIC 1 gene encodes a component of the polar-auxin-transport efflux carrier. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15112–15117.

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 35, W522–W525.

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32, W665–W667.

Forestan, C., Meda, S., and Varotto, S. (2010). ZmPIN1-mediated auxin transport is related to cellular differentiation during maize embryogenesis and endosperm development. *Plant Physiol.* 152, 1373–1390.

Friml, J., Benková, E., Blilou, I., Wisniewska, J., Hamann, T., Ljung, K., Woody, S., Sandberg, G., Scheres, B., Jürgens, G., and Palme, K. (2002a). AtPIN4 mediates sink-driven auxin gradients and root patterning in *Arabidopsis*. *Cell* 108, 661–673.

Friml, J., Wisniewska, J., Benková, E., Mendgen, K., and Palme, K. (2002b). Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. *Nature* 415, 806–809.

Friml, J., and Jones, A. R. (2010). Endoplasmic reticulum: the rising compartment in auxin biology. *Plant Physiol.* 154, 458–462.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., and Jürgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* 426, 147–153.

Gälweiler, L., Guan, C., Müller, A., Wisman, E., Mendgen, K., Yephremov, A., and Palme, K. (1998). Regulation of polar auxin transport by AtPIN1 in *Arabidopsis* vascular tissue. *Science* 282, 2226–2230.

Geisler, M., Blakeslee, J. J., Bouchard, R., Lee, O. R., Vincenzetti, V., Bandyopadhyay, A., Titapiwatanakun, B., Peer, W. A., Bailly, A., Richards, E. L., Ejendal, K. F., Smith, A. P., Baroux, C., Grossniklaus, U., Müller, A., Hrycyna, C. A., Dudler, R., Murphy, A. S., and Martinoia, E.(2005). Cellular efflux of auxin catalyzed by the *Arabidopsis* MDR/PGP transporter AtPGP1. *Plant J.* 44, 179–194.

Guo, A.-Y., Zhu, Q.-H., Chen, X., and Luo, J.-C. (2007). GSDS: a gene structure display server. *Yi Chuan* 29, 1023–1026.

Gutierrez, L., Mauriat, M., Guénin, S., Pelloux, J., Lefebvre, J.-F., Louvet, R., Rusterucci, C., Moritz, T., Guerineau, F., Bellini, C., and Van Wuytswinkel, O. (2008). The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnol. J.* 6, 609–618.

Hellgren, J. M., Olofsson, K., Plant, U., Centre, S., and Sciences, A. (2004). Patterns of auxin distribution during gravitational induction of reaction wood in poplar and pine. *Plant Physiol.* 135, 212–220.

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

Kaneda, M., Schuetz, M., Lin, B. S. P., Chanis, C., Hamberger, B., Western, T. L., Ehlting, J., and Samuels, A, L. (2011). ABC transporters coordinately expressed during lignification of *Arabidopsis* stems include a set of ABCBs associated with auxin transport. *J. Exp. Bot.* 62, 2063–2077.

Katekar, G. F., and Geissler, A. E. (1980). Auxin transport inhibitors. *Plant Physiol.* 66, 1190–1195.

Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537, 39–64.

Knöller, A. S., Blakeslee, J. J., Richards, E. L., Peer, W. A., and Murphy, A. S. (2010). Brachytic2/ZmABCB1 functions in IAA export from intercalary meristems. *J. Exp. Bot.* 61, 3689–3696.

Kramer, E. M., Dorit, R. L., and Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages. *Genetics* 149, 765–783.

Kramer, E. M., Lewandowski, M., Beri, S., Bernard, J., Borkowski, M., Borkowski, M. H., Burchfield, L. A., Mathisen, B., and Normanly, J. (2008). Auxin gradients are associated with polarity changes in trees. *Science* 320, 1610.

Kubeš, M., Yang, H., Richter, G. L., Cheng, Y., Mlodzinska, E., Wang, X., Blakeslee, J. J., Carraro, N., Petrášek, J., Zažímalová, E., Hoyerová, K., Peer, W. A., and Murphy, A. S. (2011). The *Arabidopsis* concentration-dependent influx/efflux transporter ABCB4 regulates cellular auxin levels in the root epidermis. *Plant J.* [Epub ahead of print].

Krecek, P., Skupa, P., Libus, J., Naramoto, S., Tejos, R., Friml, J., and Zažímalová, E. (2009). Protein family review The PIN-FORMED (PIN) protein family of auxin transporters. *Genome Biol.* 10, 1–11.

Lachaud, S., and Bonnemain, J. L. (1984). Seasonal variations in the polar transport pathways and retention sites of [^{3}H]indole-3-acetic acid in young branches of *Fagus sylvatica* L. *Planta* 161, 207–215.

Lee, B. H. A., Johnston, R., Yang, Y., Gallavotti, A., Kojima, M., Travençolo, B. A. N., Costa, L. D. F., Sakakibara, H., and Jackson, D. (2009). Studies of aberrant phyllotaxy1 mutants of maize indicate complex interactions between auxin and cytokinin signaling in the shoot apical meristem. *Plant Physiol.* 150, 205–216.

Lee, M., Choi, Y., Burla, B., Kim, Y.-Y., Jeon, B., Maeshima, M., Yoo, J.-Y., Martinoia, E., and Lee, Y. (2008). The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO2. *Nat. Cell Biol.* 10, 1217–1223.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25, 402–408.

Ljung, K., Bhalerao, R. P., and Sandberg, G. (2001a). Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth. *Plant J.* 28, 465–474.

Ljung, K., Ostin, A., Lioussanne, L., and Sandberg, G. (2001b). Developmental regulation of indole-3-acetic acid turnover in Scots pine seedlings. *Plant Physiol.* 125, 464–475.

Ljung, K., Hull, A. K., Celenza, J., Yamada, M., Estelle, M., and Normanly, J. (2005). Sites and regulation of auxin biosynthesis in *Arabidopsis* roots. *Plant Cell* 17, 1090–1104.

Lomax, T., Muday, G. K., and Rubery, P. H. (1995). *Plant Hormones: Physiology, Biochemistry, and Molecular Biology.* Dordrecht: K. A. Publishers.

Luschnig, C., Gaxiola, R. A., Grisafi, P., and Fink, G. R. (1998). EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in *Arabidopsis thaliana. Genes Dev.* 12, 2175–2187.

Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Maréchal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riaño-Pachón, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245–250.

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). "Creating the CIPRES science gateway for inference of large phylogenetic trees," in *Gateway Computing Environments Workshop (GCE)*, New Orleans.

Mravec, J., Skupa, P., Bailly, A., Hoyerová, K., Krecek, P., Bielach, A., Petrásek, J., Zhang, J., Gaykova, V., Stierhof, Y.-D., Dobrev, P. I., Schwarzerová, K., Rolcík, J., Seifertová, D., Luschnig, C., Benková, E., Zazímalová, E., Geisler, M., and Friml, J. (2009). Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. *Nature* 459, 1136–1140.

Müller, A., Guan, C., Gälweiler, L., Tänzler, P., Huijser, P., Marchant, A., Parry, G., Bennett, M., Wisman, E., and Palme, K. (1998). AtPIN2 defines a locus of *Arabidopsis* for root gravitropism control. *EMBO J.* 17, 6903–6911.

Multani, D. S., Briggs, S. P., Chamberlin, M. A, Blakeslee, J. J., Murphy, A. S., and Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science* 302, 81–84.

Nilsson, J., Karlberg, A., Antti, H., Lopez-Vernaza, M., Mellerowicz, E., Perrot-Rechenmann, C., Sandberg, G., and Bhalerao, R. P. (2008). Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell* 20, 843–855.

Noh, B., Murphy, A. S., and Spalding, E. P. (2001). Multidrug resistance-like genes of *Arabidopsis* required for auxin transport and auxin-mediated development. *Plant Cell* 13, 2441–2454.

Okada, K., Ueda, J., Komaki, M. K., Bell, C. J., and Shimura, Y. (1991). Requirement of the auxin polar transport system in early stages of *Arabidopsis* floral bud formation. *Plant Cell* 3, 677.

Petrásek, J., and Friml, J. (2009). Auxin transport routes in plant development. *Development* 136, 2675–2688.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, e45.

Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., Ferris, P., Kuo, A., Mitros, T., Fritz-Laylin, L. K., Hellsten, U., Chapman, J., Simakov, O., Rensing, S. A., Terry, A., Pangilinan, J., Kapitonov, V., Jurka, J., Salamov, A., Shapiro, H., Schmutz, J., Grimwood, J., Lindquist, E., Lucas, S., Grigoriev, I. V., Schmitt, R., Kirk, D., and Rokhsar, D. S. (2010). Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri. Science* 329, 223–226.

Rea, P. A. (2007). Plant ATP-binding cassette transporters. *Annu. Rev. Plant Biol.* 58, 347–375.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.-F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin-I, T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., Yamaguchi, K., Sugano, S., Kohara, Y., Fujiyama, A., Anterola, A., Aoki, S., Ashton, N., Barbazuk, W. B., Barker, E., Bennetzen, J. L., Blankenship, R., Cho, S. H., Dutcher, S. K., Estelle, M., Fawcett, J. A., Gundlach, H., Hanada, K., Heyl, A., Hicks, K. A., Hughes, J., Lohr, M., Mayer, K., Melkozernov, A., Murata, T., Nelson, D. R., Pils, B., Prigge, M., Reiss, B., Renner, T., Rombauts, S., Rushton, P. J., Sanderfoot, A., Schween, G., Shiu, S. H., Stueber, K., Theodoulou, F. L., Tu, H., Van de Peer, Y., Verrier, P. J., Waters, E., Wood, A., Yang, L., Cove, D., Cuming, A. C., Hasebe, M., Lucas, S., Mishler, B. D., Reski, R., Grigoriev, I. V., Quatrano, R. S., and Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69.

Sanchez-Fernandez, R., Davies, T. G., Coleman, J. O., and Rea, P. A. (2001). The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. *J. Biol. Chem.* 276, 30231–30244.

Sauer, M., Balla, J., Luschnig, C., Wisniewska, J., Reinöhl, V., Friml, J., and Benková, E. (2006). Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Genes Dev.* 20, 2902–2911.

Scarpella, E., Marcos, D., Friml, J., and Berleth, T. (2006). Control of leaf vascular patterning by polar auxin transport. *Genes Dev.* 20, 1015–1027.

Schrader, J., Baba, K., May, S. T., Palme, K., Bennett, M., Bhalerao, R. P., and Sandberg, G. (2003). Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10096–10101.

Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P., and Bhalerao, R. P. (2004). Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. *Plant J.* 40, 173–187.

Secchi, F., MacIver, B., Zeidel, M. L., and Zwieniecki, M. A. (2009). Functional analysis of putative genes encoding the PIP2 water channel subfamily in *Populus trichocarpa. Tree Physiol.* 29, 1467–1477.

Shen, C., Bai, Y., Wang, S., Zhang, S., Wu, Y., Chen, M., Jiang, D., and Qi, Y. (2010). Expression profile of PIN, AUX/LAX and PGP auxin transporter gene families in *Sorghum bicolor* under phytohormone and abiotic stress. *FEBS J.* 277, 2954–2969.

Sidler, M., Hassa, P., Hasan, S., Ringli, C., and Dudler, R. (1998). Involvement of an ABC transporter in a developmental pathway regulating hypocotyl cell elongation in the light. *Plant Cell* 10, 1623–1636.

Spicer, R., and Groover, A. (2010). Evolution of development of vascular cambia and secondary growth. *New Phytol.* 186, 577–592.

Spicer, R., and Holbrook, N. M. (2007). Parenchyma cell respiration and survival in secondary xylem: does metabolic activity decline with cell age? *Plant Cell Environ.* 30, 934–943.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* 57, 758–771.

Sussman, M. R., and Goldsmith, M. H. M. (1981). The action of specific inhibitors of auxin transport on uptake of auxin and binding of N-1-naphthylphthalamic acid to a membrane site in maize coleoptiles. *Planta* 13–18.

Swarup, K., Benková, E., Swarup, R., Casimiro, I., Péret, B., Yang, Y., Parry, G., Nielsen, E., De Smet, I., Vanneste, S., Levesque, M. P., Carrier, D., James, N., Calvo, V., Ljung, K., Kramer, E., Roberts, R., Graham, N., Marillonnet, S., Patel, K., Jones, J. D., Taylor, C. G., Schachtman, D. P., May, S., Sandberg, G., Benfey, P., Friml, J., Kerr, I., Beeckman, T., Laplaze, L., and Bennett, M. J. (2008). The auxin influx carrier LAX3 promotes lateral root emergence. *Nat. Cell Biol.* 10, 946–954.

Swarup, R., Friml, J., Marchant, A., Ljung, K., Sandberg, G., Palme, K., and Bennett, M. (2001). Localization of the auxin permease AUX1 suggests two functionally distinct hormone transport pathways operate in the *Arabidopsis* root apex. *Genes Dev.* 15, 2648–2653.

Swarup, R., Kramer, E. M., Perry, P., Knox, K., Leyser, H. M. O., Haseloff, J., Beemster, G. T. S., Bhalerao, R., and Bennett, M. J. (2005). Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. *Nat. Cell Biol.* 7, 1057–1065.

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577.

Terasaka, K., Blakeslee, J. J., Titapiwatanakun, B., Peer, W. A., Bandyopadhyay, A., Makam, S. N., Lee, R., Richards, E. L., Murphy, A. S., Sato, F., and Yazaki, K. (2005). PGP4, an ATP binding cassette P-glycoprotein, catalyzes auxin transport in *Arabidopsis thaliana* roots. *Plant Cell* 17, 2922–2939.

Titapiwatanakun, B., and Murphy, A. S. (2009). Post-transcriptional regulation of auxin transport proteins: cellular trafficking, protein phosphorylation, protein maturation, ubiquitination, and membrane composition. *J. Exp. Bot.* 60, 1093–1107.

Tuominen, H., Puech, L., Fink, S., and Sundberg, B. (1997). A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in hybrid aspen. *Plant Physiol.* 115, 577–585.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G. L., Cooper, D., Coutinho, P. M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Déjardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjärvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J. C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D. R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouzé, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C. J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y., and Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313, 1596–1604.

Uggla, C., Mellerowicz, E., and Sundberg, B. (1998). Indole-3-acetic acid controls cambial growth in scots pine by positional signaling. *Plant Physiol.* 117, 113–121.

Uggla, C., Moritz, T., Sandberg, G., and Sundberg, B. (1996). Auxin as a positional signal in pattern formation in plants. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9282–9286.

Utsuno, K., Shikanai, T., Yamada, Y., and Hashimoto, T. (1998). Agr, an Agravitropic locus of *Arabidopsis thaliana*, encodes a novel membrane-protein family member. *Plant Cell Physiol.* 39, 1111–1118.

Van Bel, A. J. E. (1990). Xylem-phloem exchange via the rays: the undervalued route of transport. *J. Exp. Bot.* 41, 631–644.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, RESEARCH0034.

Varón, A., Vinh, L. S., and Wheeler, W. C. (2009). POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics* 26, 72–85.

Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J., Oliva, M., Das, P., Larrieu, A., Wells, D., Guédon, Y., Armitage, L., Picard, F., Guyomarc'h, S., Cellier, C., Parry, G., Koumproglou, R., Doonan, J. H., Estelle, M., Godin, C., Kepinski, S., Bennett, M., De Veylder, L., and Traas, J. (2011). The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* 7, 508.

Verrier, P. J., Bird, D., Burla, B., Dassa, E., Forestier, C., Geisler, M., Klein, M., Kolukisaoglu, U., Lee, Y., Martinoia, E., Murphy, A., Rea, P. A., Samuels, L., Schulz, B., Spalding, E. J., Yazaki, K., and Theodoulou, F. L. (2008). Plant ABC proteins – a unified nomenclature and updated inventory. *Trends Plant Sci.* 13, 151–159.

Vieten, A., Vanneste, S., Wisniewska, J., Benková, E., Benjamins, R., Beeckman, T., Luschnig, C., and Friml, J. (2005). Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. *Development* 132, 4521–4531.

Wang, J.-R., Hu, H., Wang, G.-H., Li, J., Chen, J.-Y., and Wu, P. (2009). Expression of PIN genes in rice (*Oryza sativa* L.): tissue specificity and regulation by hormones. *Mol. Plant* 2, 823–831.

Wheeler, W. (1996). Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* 12, 1–9.

Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, W407–W410.

Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M., and Li, W. H. (1989). Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6201–6205.

Wu, G., Lewis, D. R., and Spalding, E. P. (2007). Mutations in *Arabidopsis* multidrug resistance-like ABC transporters separate the roles of acropetal and basipetal auxin transport in lateral root development. *Plant Cell* 19, 1826–1837.

Wu, X., and McSteen, P. (2007). The role of auxin transport during inflorescence development in maize (*Zea mays*, Poaceae). *Am. J. Bot.* 94, 1745–1755.

Yang, H., and Murphy, A. S. (2009). Functional expression and characterization of *Arabidopsis* ABCB, AUX 1 and PIN auxin transporters in *Schizosaccharomyces pombe*. *Plant J.* 59, 179–191.

Yang, Y., Hammes, U. Z., Taylor, C. G., Schachtman, D. P., and Nielsen, E. (2006). High-affinity auxin transport by the AUX1 influx carrier protein. *Curr. Biol.* 16, 1123–1127.

Yemm, A., May, S., Williams, L., Millner, P., Tsurumi, S., Moore, I., Napier, R., Kerr, I. D., and Bennett, M. J. (2004). Structure-function analysis of the presumptive *Arabidopsis* auxin permease AUX1. *Plant Cell* 16, 3069–3083.

Zazímalová, E., Murphy, A. S., Yang, H., Hoyerová, K., and Hosek, P. (2010). Auxin transporters – why so many? *Cold Spring Harb. Perspect. Biol.* 2, 1–14.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 October 2011; accepted: 17 January 2012; published online: 07 February 2012.

Citation: Carraro N, Tisdale-Orr TE, Clouse RM, Knöller AS and Spicer R (2012) Diversification and expression of the PIN, AUX/LAX, and ABCB families of putative auxin transporters in Populus. Front. Plant Sci. ***3****:17. doi: 10.3389/fpls.2012.00017*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

APPENDIX

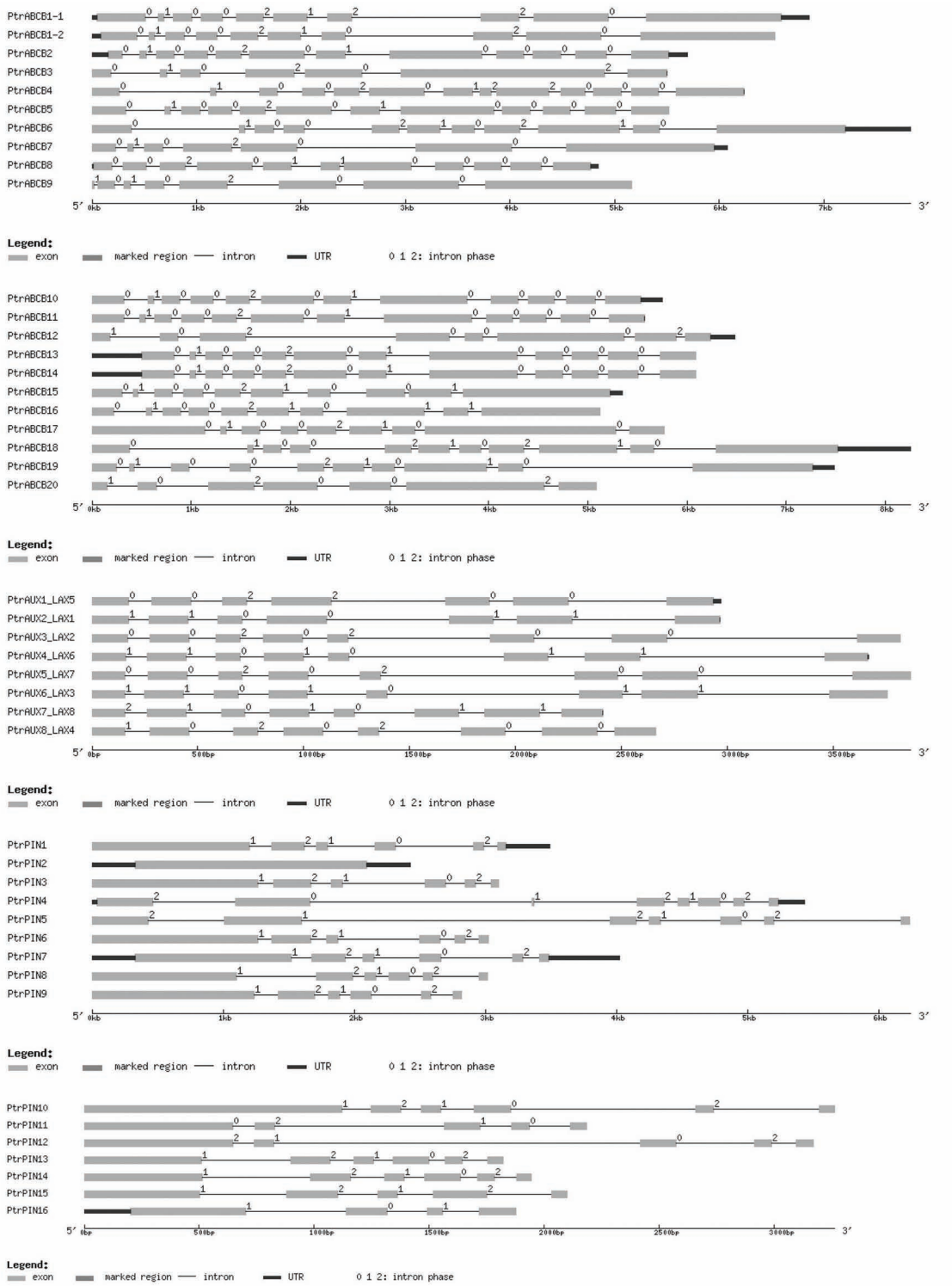


FIGURE A1 | Intron–exon structure of *PIN, AUX/LAX,* and *ABCB* genes from *Populus trichocarpa*.

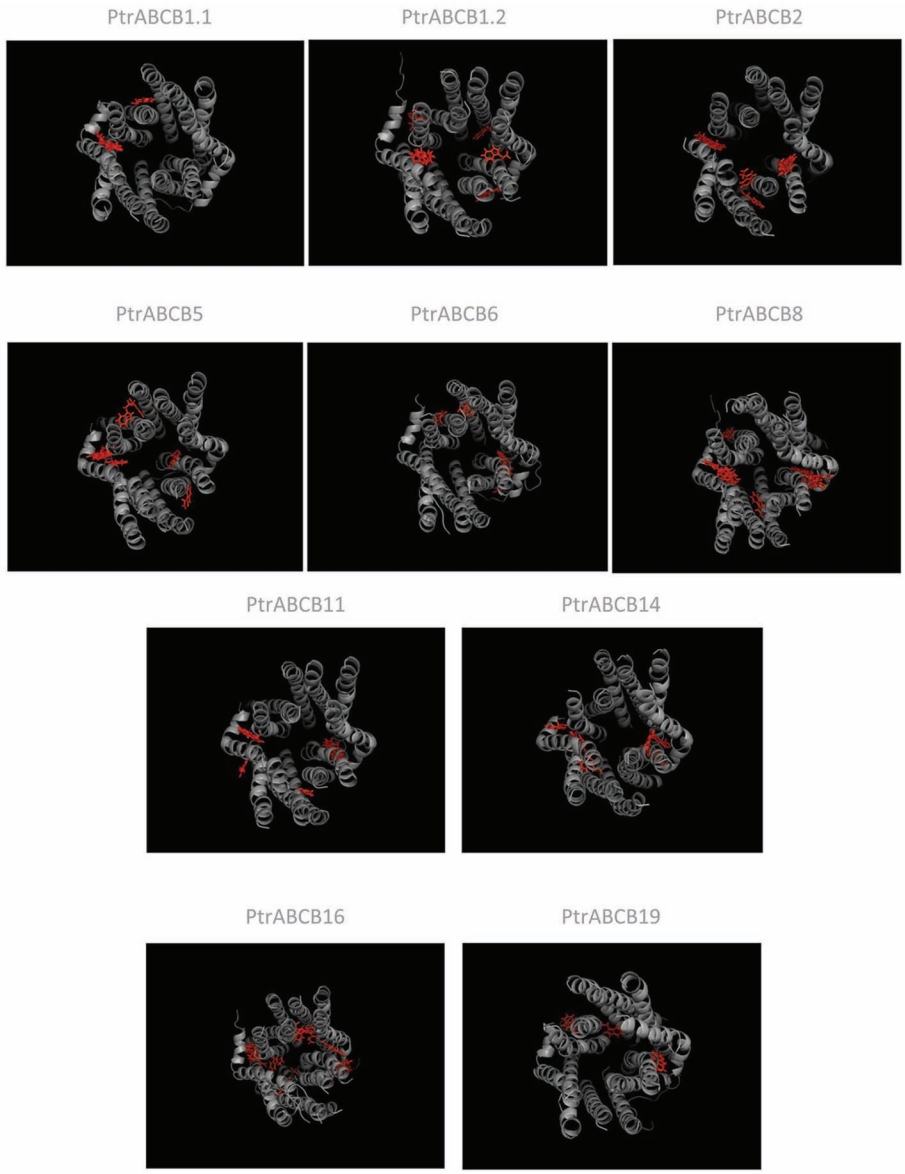


FIGURE A2 | Predicted model structures of putative auxin transport ABCBs from *Populus trichocarpa*. Tertiary protein structures have been generated using the python script Modeller 9v5. Predicted IAA docking sites are depicted in red.

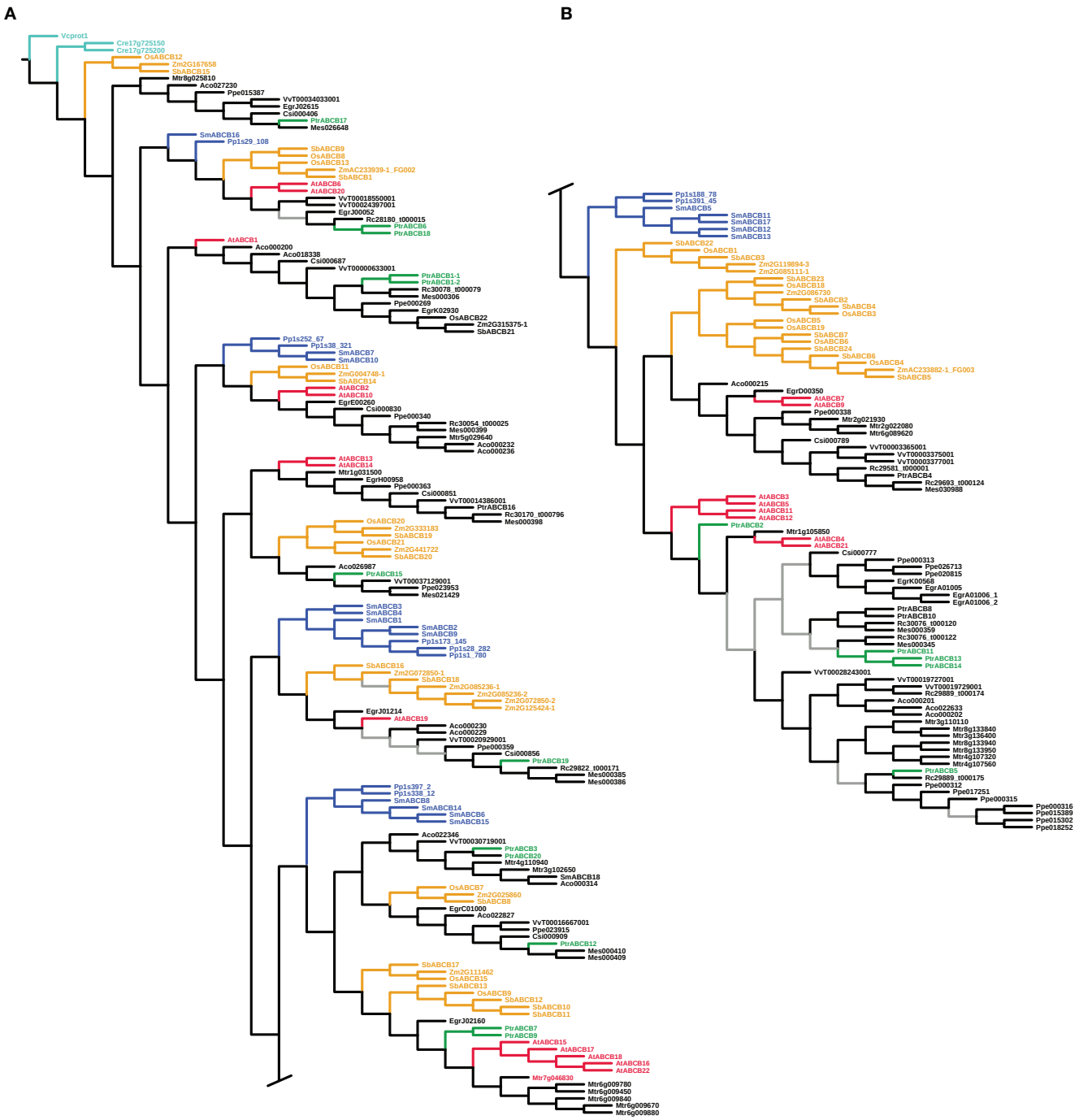


FIGURE A3 | Phylogeny of *ABCB* genes from land plants, based upon coding sequences from the loci listed in Table A4, analyzed using dynamic homology under the parsimony criterion. Gray branches indicate nodes with bootstrap support lower than 50%. Algal *ABCBs* are colored light blue–green, basal land plants blue, *Populus* green, *Arabidopsis* red, and monocots yellow. Abbreviated names of each species are listed in Table A1.

Table A1 | List of all species with their abbreviated names used in the present work.

Species	Abbreviation
Aquilegia caerulea	Aco
Arabidopsis thaliana	At
Chlamydomonas reinhardtii	Cre
Eucalyptus grandis	Egr
Manihot esculenta	Mes
Medicago truncatula	Mtr
Oryza sativa	Os
Physcomitrella patens	Pp
Populus tomentosa	Pto
Populus tremula × *tremuloides*	Ptt
Populus trichocarpa	Ptr
Prunus persica	Ppe
Ricinus communis	Rc
Selaginella moellendorffii	Sm
Sorghum bicolor	Sb
Vitis vinifera	Vv
Volvox carteri	Vc
Zea mays	Zm

Table A2 | List of putative auxin transport genes identified in the *Populus trichocarpa* genome.

Genes	JGI v1.1 gene model	JGI v1.1 locus
PtrPIN1	estExt_fgenesh4_pg.C_LG_XV0366	LG_XV:3955456–3958939
PtrPIN2	estExt_Genewise1_v1.C_LG_XVI1213	LG_XVI:2023747–2028247
PtrPIN3	gw1.X.6584.1	LG_X:11493441–11496545
PtrPIN4	estExt_fgenesh4_pm.C_LG_V0399	LG_V:12604974–12610191
PtrPIN5	fgenesh4_pm.C_LG_II000334	LG_II:4970467–4976705
PtrPIN6	fgenesh4_pm.C_LG_VIII000556	LG_VIII:8394273–8397294
PtrPIN7	estExt_Genewise1_v1.C_LG_XII1068	LG_XII:3820572–3824595
PtrPIN8	eugene3.00060333	LG_VI:2296469–2299715
PtrPIN9	fgenesh4_pm.C_LG_XVIII000434	LG_XVIII:12913539–12916356
PtrPIN10	fgenesh4_pm.C_LG_I000524	LG_I:12290101–12293363
PtrPIN11	estExt_fgenesh4_pg.C_870067	scaffold_87:1004073–1006598
PtrPIN12	fgenesh4_pg.C_LG_XIX000547	LG_XIX:6900262–6903432
PtrPIN13	fgenesh4_pg.C_LG_IV001142	LG_IV:12489496–12491318
PtrPIN14	gw1.XVII.929.1	LG_XVII:3836316–3838259
PtrPIN15	fgenesh4_pg.C_LG_XIV000875	LG_XIV:7307054–7309154
PtrPIN16	gw1.5147.2.1	scaffold_5147:1–1679
PtrAUX1/LAX5	grail3.0023028402	LG_VI:6769035–6772003
PtrAUX2/LAX1	eugene3.00161081	LG_XVI:10707443–10710997
PtrAUX3/LAX2	estExt_fgenesh4_pg.C_LG_X1704	LG_X:17003105–17007090
PtrAUX4/LAX6	estExt_Genewise1_v1.C_LG_VIII1679	LG_VIII:3795803–3800287
PtrAUX5/LAX7	estExt_fgenesh4_pg.C_LG_IV1437	LG_IV:15662320–15666183
PtrAUX6/LAX3	grail3.0001031001	LG_IX:2231536–2235747
PtrAUX7/LAX8	estExt_fgenesh4_pg.C_LG_V0933	LG_V:11098424–11101148
PtrAUX8/LAX4	grail3.0003074001	LG_II:6104679–6107343
PtrABCB1.1	gw1.28.733.1	scaffold_28:2297969–2304256
PtrABCB1.2	fgenesh4_pg.C_LG_XVI000833	LG_XVI:7805788–7812322
PtrABCB2	estExt_Genewise1_v1.C_LG_II3719	LG_II:16940658–16946357
PtrABCB3	eugene3.00130846	scaffold_1: 44776038–44781535
PtrABCB4	fgenesh4_pg.C_scaffold_204000026	scaffold_204:388201–394437
PtrABCB5	gw1.X.3657.1	LG_X:276730–282241
PtrABCB6	estExt_fgenesh4_pm.C_LG_X0835	LG_X:18271669–18278875
PtrABCB7	gw1.XVII.765.1	LG_XVII:3190614–3196509
PtrABCB8	estExt_fgenesh4_pm.C_LG_II0929	LG_II:16965413–16970969
PtrABCB9	fgenesh4_pg.C_LG_XVII000406	LG_XVII:4919010–4924173
PtrABCB10	eugene3.00140575	LG_XIV:4755266–4761017
PtrABCB11	eugene3.00140576	LG_XIV:4765985–4771483
PtrABCB12	gw1.XVIII.2596.1	LG_XVIII:8860516–8866795
PtrABCB13	eugene3.00140578	LG_XIV:4778008–4781195
PtrABCB14	estExt_fgenesh4_pm.C_LG_XIV0249	LG_XIV:4781910–4787506
PtrABCB15	fgenesh4_pm.C_LG_XV000001	LG_XV:12903–18128
PtrABCB16	fgenesh4_pm.C_LG_II000094	LG_II:1130589–1135712
PtrABCB17	eugene3.01580034	scaffold_158:318976–324742
PtrABCB18	fgenesh4_pg.C_LG_VIII000415	LG_VIII:2748354–2755879
PtrABCB19	estExt_fgenesh4_pg.C_LG_XVII0355	LG_XVII:4160851–4168120
PtrABCB20	fgenesh4_pm.C_LG_XI000351	scaffold_11:16,395,988.0.16,402,087

(Continued)

Table A2 | Continued

Genes	Phytozome v.7.0 locus	GenBank accesion number	Chrom.	Closest similar sequence
PtrPIN1	POPTR_0015s04570	XM_002322068	chr.15	PtrPIN7
PtrPIN2	POPTR_0016s03450	XM_002322578	chr.16	PtrPIN8
PtrPIN3	POPTR_0010s12320	XM_002314774	chr.10	PtrPIN6
PtrPIN4	POPTR_0005s20990	XM_002306642	chr.5	PtrPIN5
PtrPIN5	POPTR_0002s07310	XM_002302160	chr.2	PtrPIN4
PtrPIN6	POPTR_0008s12830	XM_002312400	chr.8	PtrPIN3
PtrPIN7	POPTR_0012s04470	XM_002317838	chr.12	PtrPIN1
PtrPIN8	POPTR_0006s03540	XM_002307930	chr.6	PtrPIN2
PtrPIN9	POPTR_0018s13610	XM_002324641	chr.18	No clear match
PtrPIN10	POPTR_0001s21230	XM_002298168	chr.1	No clear match
PtrPIN11	POPTR_0013s08510	XM_002328968	chr.13	PtrPIN12
PtrPIN12	POPTR_0019s07990	XM_002325430	chr.19	PtrPIN11
PtrPIN13	POPTR_0004s12310	XM_002305335	chr.4	PtrPIN14
PtrPIN14	POPTR_0017s11440	NC_008483	chr.17	PtrPIN13
PtrPIN15	POPTR_0014s14390[a]	XM_002320399	chr.14	No clear match
PtrPIN16	POPTR_0014s14390[a]	XM_002336619	chr.2	No clear match
PtrAUX1/LAX5	POPTR_0006s09940	XM_002309092	chr.6	PtrAUX2/LAX1
PtrAUX2/LAX1	POPTR_0016s12100	XM_002322933	chr.16	PtrAUX1/LAX5
PtrAUX3/LAX2	POPTR_0010s19840	XM_002316190	chr.10	PtrAUX4/LAX6
PtrAUX4/LAX6	POPTR_0008s06630	XM_002311172	chr.8	PtrAUX3/LAX2
PtrAUX5/LAX7	POPTR_0004s17860	XM_002306139	chr.4	PtrAUX6/LAX3
PtrAUX6/LAX3	POPTR_0009s13470	XM_002312937	chr.9	PtrAUX5/LAX7
PtrAUX7/LAX8	POPTR_0005s16020	XM_002306579	chr.5	PtrAUX8/LAX4
PtrAUX8/LAX4	POPTR_0002s08750	XM_002302217	chr.2	PtrAUX7/LAX8
PtrABCB1.1	POPTR_0006s12590	XM_002323449	chr.6	PtrABCB1.2
PtrABCB1.2	POPTR_0016s09680	XM_002519442	chr.16	PtrABCB1.1
PtrABCB2	POPTR_0002s18860	XM_002301511	chr.2	PtrABCB10
				PtrABCB11
				PtrABCB13
				PtrABCB14
PtrABCB3	POPTR_0001s44320	XM_002319243	chr.1	PtrABCB20
PtrABCB4	POPTR_0001s34280	XM_002331841	chr.1	No clear match
PtrABCB5	POPTR_0010s00540	XM_002314297	chr.10	No clear match
PtrABCB6	POPTR_0010s21720	XM_002316273	chr.10	PtrABCB18
PtrABCB7	POPTR_0017s11030	XM_002323983	chr.17	No clear match
PtrABCB8	POPTR_0002s18850	XM_002301514	chr.2	PtrABCB10
				PtrABCB11
PtrABCB9	POPTR_0017s12120	XM_002323830	chr.17	POPTR_0004s12180
PtrABCB10	POPTR_0014s10860	XM_002320902	chr.14	PtrABCB2, PtrABCB8
PtrABCB11	POPTR_0014s10870	XM_002320903	chr.14	PtrABCB2, PtrABCB8
PtrABCB12	POPTR_0018s09420	XM_002324987	chr.18	No clear match
PtrABCB13	POPTR_0014s10880.1	XM_002320905	chr.14	PtrABCB2, PtrABCB8
PtrABCB14	POPTR_0014s10880.2	XM_002320906	chr.14	PtrABCB2, PtrABCB8
PtrABCB15	POPTR_0015s00250	XM_002321303	chr.15	POPTR_0012s00290[c]
				POPTR_0012s00360[b]
				POPTR_0012s00370[c]

(Continued)

Table A2 | Continued

Genes	Phytozome v.7.0 locus	GenBank accesion number	Chrom.	Closest similar sequence
PtrABCB16	POPTR_0002s02110	XM_002301925	chr.2	No clear match
PtrABCB17	POPTR_0001s16560	XM_002331169	chr.1	No clear match
PtrABCB18	POPTR_0008s05020	XM_002311108	chr.8	PtrABCB6
PtrABCB19	POPTR_0017s11750	XM_002323811	chr.17	No clear match
PtrABCB20	POPTR_0011s13720	XM_002316941	chr.11	PtrABCB3

Gene models, accession numbers, chromosome position, and the closest most similar match for each gene are reported.

[a]*These genes are distinct in GenBank but they retrieve the same entry in the phytozome database (www.phytozome.org).*

[b]*Very short protein classified as ATP-binding transporter.*

[c]*Uncharacterized conserved protein.*

Table A3 | Summary of the protein characteristics of the PIN, AUX/LAX, and ABCB families of *Populus trichocarpa, Populus tomentosa, Populus tremula* × *tremuloides,* and *Arabidopsis*.

Gene	length cds (bp)	Length Protein (aa)	*n* TMHs	Type
AtPIN1	1869	622	11	Long
AtPIN2	1944	647	10	Long
AtPIN3	1923	640	10	Long
AtPIN4	1851	616	10	Long
AtPIN5	1056	351	10	Short
AtPIN6	1713	570	10	Reduced
AtPIN7	1860	619	10	Long
AtPIN8	1104	367	10	Short
PtrPIN1	1845	614	10	Long
PtrPIN2	1767	588	11	Long
PtrPIN3	1905	634	10	Long
PtrPIN4	1338	446	9	Reduced
PtrPIN5	1110	369	8	Reduced
PtrPIN6	1950	650	10	Long
PtrPIN7	1830	610	10	Long
PtrPIN8	1764	588	10	Long
PtrPIN9	1902	634	10	Long
PtrPIN10	1644	548	10	Reduced
PtrPIN11	1041	347	9	Short
PtrPIN12	1041	347	10	Short
PtrPIN13	1068	356	8	Short
PtrPIN14	1071	357	8	Short
PtrPIN15	1113	371	8	Short
PtrPIN16	912	304	6	Short
PttPIN1	1845	614	10	Long
PttPIN2	1767	588	10	Long
PttPIN3	1923	640	10	Long
PtoPIN1	1860	619	9	Long
AtAUX1	1458	485	11	
AtLAX1	1467	489	11	
AtLAX2	1452	484	11	
AtLAX3	1413	471	11	
PtrAUX1/LAX5	1443	481	11	
PtrAUX2/LAX1	1434	478	11	
PtrAUX3/LAX2	1422	474	11	
PtrAUX4/LAX6	1416	472	11	
PtrAUX5/LAX7	1476	492	11	
PtrAUX6/LAX3	1476	492	11	
PtrAUX7/LAX8	1395	465	11	
PtrAUX8/LAX4	1398	466	11	
PttLAX1	1434	477	10	
PttLAX2	1422	473	11	
PttLAX3	1476	491	11	
PtoAUX1	1434	477	10	
AtABCB1	3861	1286	12	
AtABCB2	3822	1273	12	
AtABCB3	3690	1229	11	
AtABCB4	3861	1286	9	
AtABCB5	3693	1230	9	
AtABCB6	4224	1407	13	

(Continued)

Table A3 | Continued

Gene	Length cds (bp)	Length Protein (aa)	*n* TMHs
AtABCB7	3747	1248	11
AtABCB8	3723	1241	12
AtABCB9	3711	1236	9
AtABCB10	3684	1227	10
AtABCB11	3837	1278	9
AtABCB12	3822	1273	9
AtABCB13	3738	1245	11
AtABCB14	3744	1247	11
AtABCB15	3723	1240	11
AtABCB16	3687	1228	7
AtABCB17	3723	1240	9
AtABCB18	3678	1225	9
AtABCB19	3759	1252	10
AtABCB20	4227	1408	13
AtABCB21	3891	1296	9
AtABCB22	3666	1221	7
PtrABCB1.1	4074	1357	12
PtrABCB1.2	3975	1324	12
PtrABCB2	3687	1228	10
PtrABCB3	3756	1251	9
PtrABCB4	3768	1255	10
PtrABCB5	3882	1294	9
PtrABCB6	4194	1398	12
PtrABCB7	3780	1260	11
PtrABCB8	3828	1276	11
PtrABCB9	3717	1239	9
PtrABCB10	3864	1287	9
PtrABCB11	3882	1294	9
PtrABCB12	3693	1230	8
PtrABCB13	3597	1199	7
PtrABCB14	3885	1294	9
PtrABCB15	3828	1276	10
PtrABCB16	3660	1220	11
PtrABCB17	4644	1548	12
PtrABCB18	4197	1399	12
PtrABCB19	3756	1252	10
PtrABCB20	3516	1171	10

All proteins are classified according to their sequence length, number of predicted transmembrane helices, and length of the central hydrophilic loop (short, reduced, long).

Table A4 | List of all the sequences used in the reconstruction of *PIN, AUX/LAX,* and *ABCB* families phylogenies.

Phytozome database locus or GenBank accession number	Assigned name
ABCBs	
ppa000359m.g	Ppe000359
ppa000340m.g	Ppe000340
ppa000269m.g	Ppe000269
ppa000313m.g	Ppe000313
ppa000316m.g	Ppe000316
ppa023953m.g	Ppe023953
ppa000315m.g	Ppe000315
ppa015302m.g	Ppe015302
ppa000363m.g	Ppe000363
ppa015387m.g	Ppe015387
ppa015389m.g	Ppe015389
ppa017251m.g	Ppe017251
ppa023915m.g	Ppe023915
ppa018252m.g	Ppe018252
ppa000312m.g	Ppe000312
ppa026713m.g	Ppe026713
ppa000338m.g	Ppe000338
ppa0208157m.g	Ppe020815
POPTR_0006s12590	PtrABCB11
POPTR_0016s09680	PtrABCB12
POPTR_0002s18860	PtrABCB2
POPTR_0001s44320	PtrABCB3
POPTR_0001s34280	PtrABCB4
POPTR_0010s00540	PtrABCB5
POPTR_0010s21720	PtrABCB6
POPTR_0017s11030	PtrABCB7
POPTR_0002s18850	PtrABCB8
POPTR_0017s12120	PtrABCB9
POPTR_0014s10860	PtrABCB10
POPTR_0014s10870	PtrABCB11
POPTR_0018s09420	PtrABCB12
POPTR_0014s10880.1	PtrABCB13
POPTR_0014s10880.2	PtrABCB14
POPTR_0015s00250	PtrABCB15
POPTR_0002s02110	PtrABCB16
POPTR_0001s16560	PtrABCB17
POPTR_0008s05020	PtrABCB18
POPTR_0017s11750	PtrABCB19
POPTR_0011s13720	PtrABCB20
GRMZM2G315375_T01	Zm2G315375-1
GRMZM2G085236_T01	Zm2G085236-1
GRMZM2G085236_T02	ZmG085236-2
GRMZM2G004748_T01	ZmG004748-1
GRMZM2G119894_T01	Zm2G119894-1
GRMZM2G119894_T03	Zm2G119894-3
GRMZM2G086730_T01	Zm2G086730
AC233882.1_FGT003	ZmAC233882-1_FG003
GRMZM2G025860_T01	Zm2G025860
GRMZM2G167658_T01	Zm2G167658
GRMZM2G111462_T01	Zm2G111462
GRMZM2G085111_T02	Zm2G085111-1

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
GRMZM2G333183_T01	Zm2G333183
AC233939.1_FGT002	ZmAC233939-1_FG002
GRMZM2G441722_T01	Zm2G441722
Eucrg.J2160.1	EgrJ02160
Eucgr.D00350.1	EgrD00350
Eucgr.K00568.1	EgrK00568-1
Eucgr.K02930.1	EgrK02930
Eucgr.E00260.1	EgrE00260
Eucgr.C01000.1	EgrC01000
Eucgr.A01005.1	EgrA01005
Eucgr.A01006.1	EgrA01006-1
Eucgr.A01006.2	EgrA01006-2
Eucgr.J01214.1	EgrJ01214
Eucgr.J02615.1	EgrJ02615
Eucgr.H00958.1	EgrH00958
Eucgr.J00052.1	EgrJ00052
cassava4.1_000398m.g	Mes000398
cassava4.1_000345m.g	Mes000345
cassava4.1_000359m.g	Mes000359
cassava4.1_030988m.g	Mes030988
cassava4.1_000410m.g	Mes000410
cassava4.1_000306m.g	Mes000306
cassava4.1_000385m.g	Mes000385
cassava4.1_000386m.g	Mes000386
cassava4.1_000399m.g	Mes000399
cassava4.1_000409m.g	Mes000409
cassava4.1_026648m.g	Mes026648
cassava4.1_021429m.g	Mes021429
Medtr5g029640.1	Mtr5g029640
Medtr1g031500.1	Mtr1g031500
Medtr2g022080.1	Mtr2g022080
Medtr6g089620.1	Mtr6g089620
Medtr2g021930.1	Mtr2g021930
Medtr1g105850.1	Mtr1g105850
Medtr8g078020.1	Mtr8g078020
Medtr6g009670.1	Mtr6g009670
Medtr8g133940.1	Mtr8g133940
Medtr3g110110.1	Mtr3g110110
Medtr8g133950.1	Mtr8g133950
Medtr8g133840.1	Mtr8g133840
Medtr4g107320.1	Mtr4g107320
Medtr4g107560.1	Mtr4g107560
Medtr6g009780.1	Mtr6g009780
Medtr6g009880.1	Mtr6g009880
Medtr6g009840.1	Mtr6g009840
Medtr3g136400.1	Mtr3g136400
Medtr7g046830.1	Mtr7g046830
Medtr6g009450.1	Mtr6g009450
Medtr3g102650.1	Mtr3g102650
Medtr8g025810.1	Mtr8g025810
Medtr4g110940.1	Mtr4g110940
GSVIVT00000633001	VvT00000633001

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
GSVIVT00003365001	VvT00003365001
GSVIVT00003375001	VvT00003375001
GSVIVT00003377001	VvT00003377001
GSVIVT00014386001	VvT00014386001
GSVIVT00016667001	VvT00016667001
GSVIVT00018550001	VvT00018550001
GSVIVT00019727001	VvT00019727001
GSVIVT00019729001	VvT00019729001
GSVIVT00020929001	VvT00020929001
GSVIVT00024397001	VvT00024397001
GSVIVT00028243001	VvT00028243001
GSVIVT00030719001	VvT00030719001
GSVIVT00034033001	VvT00034033001
GSVIVT00037129001	VvT00037129001
Sb01g039110.1	SbABCB1
Sb02g019540.1	SbABCB2
Sb03g011860.1	SbABCB3
Sb03g023740.1	SbABCB4
Sb03g031990.1	SbABCB5
Sb03g032000.1	SbABCB6
Sb03g032030.1	SbABCB7
Sb03g033290.1	SbABCB8
Sb03g047490.1	SbABCB9
Sb04g006087.1	SbABCB10
Sb04g006090.1	SbABCB11
Sb04g006100.1	SbABCB12
Sb04g022480.1	SbABCB13
Sb04g031170.1	SbABCB14
Sb06g001440.1	SbABCB15
Sb06g018860.1	SbABCB16
Sb06g020350.1	SbABCB17
Sb06g030350.1	SbABCB18
Sb07g003510.1	SbABCB19
Sb07g003520.1	SbABCB20
Sb07g023730.1	SbABCB21
Sb09g002940.1	SbABCB22
Sb09g027320.1	SbABCB23
Sb09g027330.1	SbABCB24
e_gw1.13.597.1	SmABCB1
fgenesh1_pm.C_scaffold_6000062	SmABCB2
fgenesh2_pg.C_scaffold_13000013	SmABCB3
e_gw1.6.146.1	SmABCB4
estExt_Genewise1Plus.C_350372	SmABCB5
fgenesh1_pm.C_scaffold_42000045	SmABCB6
e_gw1.0.369.1	SmABCB7
fgenesh2_pg.C_scaffold_9000128	SmABCB8
estExt_Genewise1.C_210058	SmABCB9
fgenesh1_pm.C_scaffold_2000054	SmABCB10
e_gw1.73.37.1	SmABCB11
estExt_Genewise1Plus.C_90010	SmABCB12
e_gw1.0.1863.1	SmABCB13
e_gw1.22.307.1	SmABCB14

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
fgenesh1_pm.C_scaffold_0000169	SmABCB15
estExt_Genewise1.C_00569	SmABCB16
e_gw1.73.196.1	SmABCB17
fgenesh1_pm.C_scaffold_15000068	SmABCB18
LOC_Os01g18670.1	OsABCB1
LOC_Os01g35030.1	OsABCB3
LOC_Os01g50080.1	OsABCB4
LOC_Os01g50100.1	OsABCB5
LOC_Os01g50160.1	OsABCB6
LOC_Os01g52550.1	OsABCB7
LOC_Os01g74470.1	OsABCB8
LOC_Os02g09720.1	OsABCB9
LOC_Os02g46680.1	OsABCB11
LOC_Os03g08380.1	OsABCB12
LOC_Os03g17180.1	OsABCB13
LOC_Os04g40570.1	OsABCB15
LOC_Os05g47490.1	OsABCB18
LOC_Os05g47500.1	OsABCB19
LOC_Os08g05690.1	OsABCB20
LOC_Os08g05710.1	OsABCB21
LOC_Os08g45030.1	OsABCB22
Rco30078.t000079	Rc30078_t000079
Rco30054.t000025	Rc30054_t000025
Rco30076.t000120	Rc30076_t000120
Rco30076.t000122	Rc30076_t000122
Rco28180.t000015	Rc28180_t000015
Rco30170.t000796	Rc30170_t000796
Rco29581.t000001	Rc29581_t000001
Rco29693.t000124	Rc29693_t000124
Rco29822.t000171	Rc29822_t000171
Rco29889.t000174	Rc29889_t000174
Rco29889.t000175	Rc29889_t000175
Pp1s252_67V6.1	Pp1s252_67
Pp1s38_321V6.1	Pp1s38_321
Pp1s28_282V6.1	Pp1s28_282
Pp1s173_145V6.1	Pp1s173_145
Pp1s1_780V2.1	Pp1s1_780
Pp1s397_2V6.1	Pp1s397_2
Pp1s188_78V6.1	Pp1s188_78
Pp1s391_45V6.1	Pp1s391_45
Pp1s338_12V6.1	Pp1s338_12
Pp1s29_108V2.1	Pp1s29_108
Vc_estExt_fgenesh4_pg.C_30286	VcProt1
Cre17.g725200	Cre17_g725200
Cre17.g725150	Cre17_g725150
AT2G36910	AtABCB1
AT4G25960	AtABCB2
AT4G01820	AtABCB3
AT2G47000	AtABCB4
AT4G01830	AtABCB5
AT2G39480	AtABCB6
AT5G46540	AtABCB7

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
AT3G30875	AtABCB8
AT4G18050	AtABCB9
AT1G10680	AtABCB10
At1g02520	AtABCB11
AT1G02530	AtABCB12
AT1G27940	AtABCB13
AT1G28010	AtABCB14
AT3G28345	AtABCB15
AT3G28360	AtABCB16
AT3G28380	AtABCB17
AT3G28390	AtABCB18
AT3G28860	AtABCB19
AT3G55320	AtABCB20
AT3G62150	AtABCB21
AT3G28415	AtABCB22
orange1.1g000851m.g	Csi_g000851
orange1.1g000777m.g	Csi_g000777
orange1.1g000789m.g	Csi_g000789
orange1.1g000909m.g	Csi_g000909
orange1.1g000830m.g	Csi_g000830
orange1.1g000406m.g	Csi_g000406
orange1.1g000687m.g	Csi_g000687
orange1.1g000856m.g	Csi_g000856
AcoGoldSmith_v1.000232m.g	Aco000232
AcoGoldSmith_v1.022827m.g	Aco022827
AcoGoldSmith_v1.027230m.g	Aco027230
AcoGoldSmith_v1.000200m.g	Aco000200
AcoGoldSmith_v1.018338m.g	Aco018338
AcoGoldSmith_v1.000314m.g	Aco000314
AcoGoldSmith_v1.022346m.g	Aco022346
AcoGoldSmith_v1.026987m.g	Aco026987
AcoGoldSmith_v1.022633m.g	Aco022633
AcoGoldSmith_v1.000202m.g	Aco000202
AcoGoldSmith_v1.000201m.g	Aco000201
AcoGoldSmith_v1.000230m.g	Aco000230
AcoGoldSmith_v1.000215m.g	Aco000215
AcoGoldSmith_v1.000236m.g	Aco000236
AcoGoldSmith_v1.000229m.g	Aco000229
AUX/LAXs	
ppa005323m.g	Ppe005323
ppa005057m.g	Ppe005057
ppa004949m.g	Ppe004949
ppa004865m.g	Ppe004865
POPTR_0006s09940	PtrAUX1/LAX5
POPTR_0016s12100	PtrAUX2/LAX1
POPTR_0010s19840	PtrAUX3/LAX2
POPTR_0008s06630	PtrAUX4/LAX6
POPTR_0004s17860	PtrAUX5/LAX7
POPTR_0009s13470	PtrAUX6/LAX3
POPTR_0005s16020	PtrAUX7/LAX8
POPTR_0002s08750	PtrAUX8/LAX4
GRMZM2G067022_T01	Zm2G067022

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
GRMZM2G127949_T01	Zm2G127949
GRMZM2G045057_T01	Zm2G045057
GRMZM2G149481_T01	Zm2G149481
GRMZM2G129413_T01	Zm2G129413
Eucgr.F03758.1	EgrF03758_1
Eucgr.K02992.2	EgrK02992_2
Eucgr.G03044.2	EgrG03044_2
Eucgr.G01769.2	EgrG01769_2
Eucgr.A00514.2	EgrA00514_2
cassava4.1_006838m.g	Mes006838
cassava4.1_006423m.g	Mes006423
cassava4.1_006788m.g	Mes006788
cassava4.1_006570m.g	Mes006570
cassava4.1_006783m.g	Mes006783
cassava4.1_006474m.g	Mes006474
cassava4.1_007093m.g	Mes007093
Medtr3g024670.1	Mtr3g024670
Medtr3g097960.1	Mtr3g097960
Medtr5g089600.1	Mtr5g089600
GSVIVT01008917001	VvT01008917001
GSVIVT01024054001	VvT01024054001
GSVIVT01032855001	VvT01032855001
GSVIVT01033986001	VvT01033986001
Sb01g026240.1	SbLAX1
Sb01g041270.1	SbLAX2
Sb03g040320.1	SbLAX3
Sb05g004250.1	SbLAX4
Sb09g021990.1	SbLAX5
estExt_Genewise1Plus.C_20968	SmAUX1
estExt_fgenesh2_pg.C_50586	SmAUX2
LOC_Os01g63770.1	OsLAX1
LOC_Os03g14080.1	OsLAX2
LOC_Os05g37470.1	OsLAX3
LOC_Os10g05690.1	OsLAX4
LOC_Os11g06820.1	OsLAX5
Rco29669.t000030	Rc29669_t000030
Rco29741.t000002	Rc29741_t000002
Rco29908.t000197	Rc29908_t000197
Rco29969.t000004	Rc29969_t000004
Pp1s90_46V6.1	Pp1s90_46
Pp1s213_89V6.1	Pp1s213_89
Pp1s211_67V6.1	Pp1s211_67
AT2G38120.1	AtAUX1
AT5G01240.1	AtLAX1
AT2G21050.1	AtLAX2
AT1G77690.1	AtLAX3
orange1.1g011392m.g	Csi_g011392
orange1.1g011022m.g	Csi_g011022
orange1.1g012371m.g	Csi_g012371
orange1.1g011966m.g	Csi_g011966
AcoGoldSmith_v1.004219m.g	Aco004219
AcoGoldSmith_v1.004342m.g	Aco004342

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
AcoGoldSmith_v1.003895m.g	Aco003895
AY864733	Pto-AY864733
AF115543	Ptt-AF115543
PINs	
ppa022797m.g	Ppe022797
ppa003159m.g	Ppe003159
ppa024134m.g	Ppe024134
ppa002528m.g	Ppe002528
ppa025174m.g	Ppe025174
ppa002944m.g	Ppe002944
ppa021573m.g	Ppe021573
ppa007621m.g	Ppe007621
POPTR_0015s04570	PtrPIN1
POPTR_0016s03450	PtrPIN2
POPTR_0010s12320	PtrPIN3
POPTR_0005s20990	PtrPIN4
POPTR_0002s07310	PtrPIN5
POPTR_0008s12830	PtrPIN6
POPTR_0012s04470	PtrPIN7
POPTR_0006s03540	PtrPIN8
POPTR_0018s13610	PtrPIN9
POPTR_0001s21230	PtrPIN10
POPTR_0013s08510	PtrPIN11
POPTR_0019s07990	PtrPIN12
POPTR_0004s12310	PtrPIN13
POPTR_0017s11440	PtrPIN14
POPTR_0014s14390	PtrPIN15
XM_002336619.1	PtrPIN16
ZmPIN1a_GRMZM2G098643	ZmPIN1a
ZmPIN1b_GRMZM2G074267	ZmPIN1b
ZmPIN1c_GRMZM2G149184	ZmPIN1c
ZmPIN1d_GRMZM2G171702_T01	ZmPIN1d
ZmPIN2	ZmPIN2
ZmPIN5a-GRMZM2G025742	ZmPIN5a
ZmPIN5b-GRMZM2G148648	ZmPIN5b
ZmPIN5c-GRMZM2G040911	ZmPIN5c
ZmPIN8_GRMZM5G839411	ZmPIN8
ZmPIN9_GRMZM5G859099	ZmPIN9
ZmPIN10a-GRMZM2G126260	ZmPIN10a
ZmPIN10b-GRMZM2G160496	ZmPIN10b
Eucgr.F04265.1	EgrF04265_1
Eucgr.K02271.1	EgrK02271_1
Eucgr.G02187.1	EgrG02187_1
Eucgr.G02549.1	EgrG02549_1
Eucgr.B01406.1	EgrB01406_1
Eucgr.B02902.1	EgrB02902_1
Eucgr.B00948.1	EgrB00948_1
Eucgr.C00078.1	EgrC00078_1
Eucgr.A02229.1	EgrA02229_1
Eucgr.H01390.1	EgrH01390_1
Eucgr.H01391.1	EgrH01391_1
Eucgr.I01919.1	EgrI01919_1

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
Eucgr.G02548.1	EgrG02548_1
Eucgr.B01405.1	EgrB01405_1
Eucgr.B01403.1	EgrB01403_1
Eucgr.H01382.1	EgrH01382_1
cassava4.1_003807m.g	Mes003807
cassava4.1_030090m.g	Mes030090
cassava4.1_029078m.g	Mes029078
cassava4.1_003367m.g	Mes003367
cassava4.1_006998m.g	Mes006998
cassava4.1_026579m.g	Mes026579
cassava4.1_003794m.g	Mes003794
cassava4.1_029063m.g	Mes029063
cassava4.1_033391m.g	Mes033391
cassava4.1_010688m.g	Mes010688
cassava4.1_010607m.g	Mes010607
Medtr2g043210	Mtr2g043210
Medtr4g154810	Mtr4g154810
Medtr6g083450	Mtr6g083450
Medtr7g008720	Mtr7g008720
Medtr7g089430	Mtr7g089430
Medtr7g106430	Mtr7g106430
Medtr8g130020	Mtr8g130020
Medtr8g130040	Mtr8g130040
MtrAAM55297	MtrAAM55297
MtrAY115838	MtrAY115838
MtrAAT48627	MtrAAT48627
GSVIVT00014302001	VvT00014302001
GSVIVT00017824001	VvT00017824001
GSVIVT00020886001	VvT00020886001
GSVIVT00023254001	VvT00023254001
GSVIVT00023255001	VvT00023255001
GSVIVT00025093001	VvT00025093001
GSVIVT00025108001	VvT00025108001
GSVIVT00030482001	VvT00030482001
GSVIVT00031315001	VvT00031315001
Sb02g029210.1	SbPIN1
Sb03g029320.1	SbPIN2
Sb03g032850.1	SbPIN3
Sb03g037350.1	SbPIN4
Sb03g043960.1	SbPIN5
Sb04g028170.1	SbPIN6
Sb05g002150.1	SbPIN7
Sb07g026370.1	SbPIN8
Sb10g004430.1	SbPIN9
Sb10g008290.1	SbPIN10
Sb10g026300.1	SbPIN11
e_gw1.26.13.1	Sm102666
e_gw1.59.169.1	Sm119024
fgenesh1_pm.C_scaffold_9000007	Sm231064
fgenesh1_pm.C_scaffold_59000022	Sm234325
estExt_fgenesh1_pm.C_500006	Sm268490

(Continued)

Table A4 | Continued

Phytozome database locus or GenBank accession number	Assigned name
e_gw1.21.81.1	Sm99301
Os01g45550.1	OsPIN10a
Os01g51780	OsPIN8
Os01g58860	OsPIN9
Os01g69070	OsPIN5a
Os02g50960.1	OsPIN1b
Os05g50140	OsPIN10b
Os06g12610	OsPIN1a
Os06g44970	OsPIN2
Os08g41720	OsPIN5b
Os09g32770	OsPIN5c
Os11g04190	OsPIN1c
Os12g04000	OsPIN1d
Rco27985.t000045	Rc27985_t000045
Rco29662.t000026	Rc29662_t000026
Rco29816.t000014	Rc29816_t000014
Rco30180.t000054	Rc30180_t000054
Rco29822.t000149	Rc29822_t000149
Rco30128.t000486	Rc30128_t000486
Pp1s10_17V6.1	PpPIN1A
Pp1s18_186V6.1	PpPIN1B
Pp1s32_43V6.1	PpPIN1C
Pp1s79_126V6	PpPIN1D
AT1G73590	AtPIN1
AT5G57090	AtPIN2
AT1G70940	AtPIN3
AT2G01420	AtPIN4
AT5G16530	AtPIN5
AT1G77110	AtPIN6
AT1G23080	AtPIN7
AT5G15100	AtPIN8
orange1.1g006199m.g	Csi_g006199
orange1.1g007826m.g	Csi_g007826
orange1.1g036474m.g	Csi_g036474
orange1.1g041301m.g	Csi_g041301
orange1.1g048649m.g	Csi_g048649
orange1.1g035534m.g	Csi_g035534
orange1.1g007420m.g	Csi_g007420
orange1.1g018360m.g	Csi_g018360
orange1.1g019021m.g	Csi_g019021
AcoGoldSmith_v1.001931m.g	Aco001931
AcoGoldSmith_v1.018694m.g	Aco018694
AcoGoldSmith_v1.018139m.g	Aco018139
AcoGoldSmith_v1.016169m.g	Aco016169
AcoGoldSmith_v1.007499m.g	Aco007499
AcoGoldSmith_v1.021242m.g	Aco021242
AY302060	PtoPIN1-like
AF190881	PttPIN1
AF515435	PttPIN2
AF515434	PttPIN3

Table A5 | List of all primers used in the present work.

Name	Direction	Sequence (5′–3′)	Tm (°C)[a]	Amplicon (bp)
PIN1 RT-F3	Forward	AAGCTGAAGATGGTAGGGACCTT	58	94
PIN1 RT-R3	Reverse	TGGGCGCCATAATCATGAC	59	
PIN2 RT-F4	Forward	GATCAATGTTCAGGGATCAACAGA	59	81
PIN2 RT-R4	Reverse	GTTGTTGGTGGAAATGAAGTGAAA	59	
PIN3 RT-F3	Forward	CTTCACGTTGCTATTGTTCAGG	54.1	238
PIN3 RT-R3	Reverse	TGACACACGACCAGCAAGTAA	56.5	
PIN4 RT-F4	Forward	CGTTGGAATGAGAGGAGTGC	55	204
PIN4 RT-R4	Reverse	AATCTAAATTCCCCCTCTAATTCATGG	54.8	
PIN5 RT-F2	Forward	GACTAATGCAACCAACACACCTTT	58	67
PIN5 RT-R2	Reverse	TGGATGCCGGGATATTTTACC	59	
PIN6 RT-F2	Forward	CCATTCCACAAGCTGGAAATT	53.7	166
PIN6 RT-R2	Reverse	CCGGAATCTGGAGCGCCGA	62.6	
PIN7 RT-F4	Forward	TCAGTGCTCGGGCATCAA	58	81
PIN7 RT-R4	Reverse	GGATCATTAGTAGATATGAAGTGGAAAGAG	58	
PIN8 RT-F2	Forward	CTTCATTTGCTGTTGGACTACG	54.1	192
PIN8 RT-R2	Reverse	GTCCAAGCAAAATATAGTAAACCAGTGT	55.6	
PIN9 RT-F2	Forward	GCTGCTTTTCAACCTGAATCCG	57	173
PIN9 RT-R2	Reverse	TCTGCTGCCATATCCATCTTCTTTTG	57.3	
PIN10 RT-F4	Forward	GGCAGACACACCTACCCTGATC	59.4	100
PIN10 RT-R4	Reverse	CCGGAGGCATCTGTTGTTTC	56.3	
PIN11 RT-F3	Forward	CAGCATTGCCACAGTCAATTACATC	56.8	196
PIN11 RT-R3	Reverse	GCCGAGCTATATTCCTCCTTCAAG	57	
PIN12 RT-F6	Forward	GCTACGGCTGGTCCATTACC	58	100
PIN12 RT-R6	Reverse	ACTGCCGTCGGCCCATA	59.6	
PIN13 RT-F2	Forward	GGATACATTGAGCACAGGGGTAA	56.6	199
PIN13 RT-R2	Reverse	TGGACGGGACAGACTTCTATGATTC	57.9	
PIN14 RT-F3	Forward	ATAGTGATATTGTCAACAGGAGGG	54.1	175
PIN14 RT-R3	Reverse	CCAGTCTAACGGCGAAGGAAG	57.6	
PIN15 RT-F2	Forward	TTTGCTGGGCTAATTTCTCAAGA	55.5	188
PIN15 RT-R1	Reverse	AGTGGGATCCCCATCACAAG	54.9	
PIN16 RT-F4	Forward	GGTAACAATCTTGTCAAAGGCAGGT	57.3	199
PIN16 RT-R4	Reverse	GGATAGTTTCAACATGGTCCCTCTCA	58.2	
AUX1 RT-F1	Forward	TCCCTTTATGCCAAGCTGGA	56.5	217
AUX1 RT-R1	Reverse	ATGTAGTCAGCTCACTCAGCG	56.6	
AUX2 RT-F3	Forward	CGTTCGGACTCTTCGCAAAG	56.3	100
AUX2 RT-R3	Reverse	TCTTGGGACTGATTTGCTTCAG	55.1	
AUX3 RT-F2	Forward	GTTCACGGCCAGGTTGATG	56.6	100
AUX3 RT-R2	Reverse	CATGCCCACCAAAAGTGTAGAG	56.1	
AUX4 RT-F4	Forward	AGGGTGGGCTAGTATGTCCAA	57.7	191
AUX4 RT-R4	Reverse	AAACACAATGCAGAGGAGATGC	55.9	
AUX5 RT-F1	Forward	AGCCATCAAAGTACACGGGA	56.3	174
AUX5 RT-R1	Reverse	TCTGAGGTGGGCATTGGTAA	56.1	
AUX6 RT-F4	Forward	CCTGTGGTTATTCCCATTTGGTT	55.6	180
AUX6 RT-R4	Reverse	GTACTTTGGTGGTTGCTCCA	55.2	
AUX7 RT-F2	Forward	CGTCAGATTGATTCATTTGGTCTATTC	54.2	213
AUX7 RT-R2	Reverse	ATCACACCTTTTCAAGAACCAACA	55.2	
AUX8 RT-F1	Forward	GAGAGAATGCTGTGGAGAGAC	54.8	182
AUX8 RT-R1	Reverse	ACACTGGTAGCACTTGGTGA	56.2	
ABCB1 RT-F4	Forward	GATGGTAAAGTAGCAGAGCAAGGAT	56.7	212
ABCB1 RT-R4	Reverse	ATGGGATATACTCCTCTTACTGGTGT	56.5	
ABCB2 RT-F3	Forward	CAAGCATGAGACTCTGATTCATATCA	54.7	100

(Continued)

Table A5 | Continued

Name	Direction	Sequence (5′–3′)	Tm (°C)[a]	Amplicon (bp)
ABCB2 RT-R3	Reverse	AATATTGCAGGTGGTGACTCAAGA	56.4	
ABCB4 RT-F2	Forward	GGGCAATCCTAAAGAATCCGAAAAT	55.7	264
ABCB4 RT-R4	Reverse	TATGAAGGGCGACCAAGGATG	56.9	
ABCB5 RT-F3	Forward	TCGCAATACCTCCCGGTACA	58.1	100
ABCB5 RT-R3	Reverse	GCGTGCGGGTCGTAAAAC	57.3	
ABCB7 RT-F2	Forward	GTGGTTTTGCTGTTAGATGAGGC	56.5	269
ABCB7 RT-R2	Reverse	ACTGTTTTGTGTTGTCCTCTGG	55.4	
ABCB10 RT-F4	Forward	CAG AAG CAA AGG GTA GCC ATT	55.4	211
ABCB10 RT-R4	Reverse	CTCCATTTTTAACCACTGCGATTAGA	56.4	
ABCB13 RT-F3	Forward	CAAGAGCAATTCTGAAAGATCCACG	56.3	206
ABCB13 RT-R3	Reverse	ACCTTTTTCCACTATCTTGCCATG	55.6	
ABCB14 RT-F1	Forward	GACAGTCAAGTCAAAGAATCTCATTG	54.2	221
ABCB14 RT-R1	Reverse	TGGAACCTCTGGCTTGTTAAGA	56	
ABCB13 RT-F2	Forward	CAAGAAGCACTGGACCGAATCAT	57.4	229
ABCB13 RT-R2	Reverse	TAAACACACGGAGGTGCTACAAT	56.4	
ABCB18 RT-F3	Forward	AGCTCATCCATCGAATCTGAATCAA	56.3	211
ABCB18 RT-R3	Reverse	GCATCAGACGGACATACAAACCAT	57.4	
ABCB19 RT-F3	Forward	TCTTAAGGACCCAGCAATCCTACT	57.3	100
ABCB19 RT-R3	Reverse	CCTCATTAGCCTCTCGAGTGCTT	58.5	
ACT2 RT-F1[b]	Forward	GCAACTGGGATGATATGGAGA	54.3	213
ACT2 RT-R1	Reverse	TACGACCACTGGCATACAGG	56.5	
UBQ RT-F1[b]	Forward	CAGCTTGAAGATGGGAGGAC	55.4	154
UBQ RT-R1	Reverse	CAATGGTGTCTGAGCTCTCG	55.5	
TUA2 RT-F1	Forward	CCTACTGTAGTACCTGGGGGTG	58.2	230
TUA2 RT-R1	Reverse	CCAACTTCCTCGTAATCCTTCTCA	56.2	
PD-E1 RT-F1	Forward	ATGAGAACTGGTGGTATTGGTGC	57.3	164
PD-E1 RT-R1	Reverse	GTCACAATCTGGGCAGGTTGAAC	58.5	
CLONING AND SEQUENCING				
M13F	Forward	TTGTAAAACGACGGCCAGT	54.7	
M13R	Reverse	CAGGAAACAGCTATGACC	50.1	
adp1-dT17[c]		CCGGATCCTCTAGAGCGGCCGC$(T)_{17}$	64.6	
adp1		CCGGATCCTCTAGAGCGGCC	61.9	
PIN3 RT-F3	Forward	CTTCACGTTGCTATTGTTCAGG	54.1	
PIN4 RT-F3	Forward	CTTCAGCCTCGGATAATTGTATGC	55.1	
PIN11A RT-F3	Forward	GCGATGTCTTACGTGTTGCTA	55.1	
PIN13 RT-F2	Forward	GGATACATTGAGCACAGGGGTAA	56.6	
AUX4 RT-F3	Forward	CCGACTCCTGCAAAACATCATTA	55.4	
ABCB1 RT-F3	forward	CGCATGATACAGTTACAAAGGTTCA	55.5	

[a] *Melting temperatures were calculated with the online tool OlygoAnalyzer v.3.1 from Integrated DNA Technologies.*

[b] *These primer pairs have been first published in Secchi et al. (2009).*

[c] *This primer sequence has been first published in Kramer et al. (1998).*

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 20 February 2012
doi: 10.3389/fpls.2012.00033

Annotation of *Selaginella moellendorffii* major intrinsic proteins and the evolution of the protein family in terrestrial plants

Hanna I. Anderberg, Per Kjellbom and Urban Johanson*

Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Center for Chemistry and Chemical Engineering, Lund University, Lund, Sweden

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Christophe Maurel, National Institute for Agricultural Research, France
Ian Wallace, The University of California Berkeley, USA

****Correspondence:***
Urban Johanson, Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Center for Chemistry and Chemical Engineering, Lund University, PO Box 124, S-221 00 Lund, Sweden.
e-mail: urban.johanson@biochemistry.lu.se

Major intrinsic proteins (MIPs) also called aquaporins form pores in membranes to facilitate the permeation of water and certain small polar solutes across membranes. MIPs are present in virtually every organism but are uniquely abundant in land plants. To elucidate the evolution and function of MIPs in terrestrial plants, the MIPs encoded in the genome of the spikemoss *Selaginella moellendorffii* were identified and analyzed. In total 19 MIPs were found in *S. moellendorffii* belonging to 6 of the 7 MIP subfamilies previously identified in the moss *Physcomitrella patens*. Only three of the MIPs were classified as members of the conserved water specific plasma membrane intrinsic protein (PIP) subfamily whereas almost half were found to belong to the diverse NOD26-like intrinsic protein (NIP) subfamily permeating various solutes. The small number of PIPs in *S. moellendorffii* is striking compared to all other land plants and no other species has more NIPs than PIPs. Similar to moss, *S. moellendorffii* only has one type of tonoplast intrinsic protein (TIP). Based on ESTs from non-angiosperms we conclude that the specialized groups of TIPs present in higher plants are not found in primitive vascular plants but evolved later in a common ancestor of seed plants. We also note that the silicic acid permeable NIP2 group that has been reported from angiosperms appears at the same time. We suggest that the expansion of the number MIP isoforms in higher plants is primarily associated with an increase in the different types of specialized tissues rather than the emergence of vascular tissue *per se* and that the loss of subfamilies has been possible due to a functional overlap between some subfamilies.

Keywords: water channels, AQP, SIP, XIP, HIP, GIP, phylogeny

INTRODUCTION

The major intrinsic proteins (MIPs) constitute an ancient protein family and representatives can be found in virtually all types of organisms (Preston et al., 1992; Heymann and Engel, 1999). These proteins form pores in biological membranes and facilitate the passive transport of a range of small polar molecules, most notably water and glycerol. Several structural features are conserved throughout the protein family such as the six transmembrane helices (H1–H6) connected by five loops (LA–LE) as well as two conserved NPA motifs (Asn-Pro-Ala) at the N-terminal end of two short helices in LB and LE (Fu et al., 2000; Sui et al., 2001). These oppositely oriented short helices connected by the two NPA-motifs facing each other in the middle of the membrane form a seventh helical transmembrane structure. The dipole moment of each of the half transmembrane helices creates partial positive charges centered at the NPA motifs and these charges prevent the passage of protons through the pore (de Groot et al., 2003). Four amino acid residues, known as the aromatic/arginine-filter (ar/R filter), located toward the extracellular entrance form the narrowest part of the pore and are thought to determine the substrate selectivity (Fu et al., 2000; Beitz et al., 2006).

The MIPs are especially abundant and multifaceted in terrestrial plants suggesting that the expansion of this protein family has contributed to the adaptation of green plants to life on land (Danielson and Johanson, 2010; Anderberg et al., 2011). The green plants (viridiplantae) consist of two major clades, the chlorophytes, including the majority of all green algae, and the streptophytes (Leliaert et al., 2011). The streptophytes comprise a number of classes of freshwater algae, collectively known as the charophytes, and the land plants (embryophyta, **Figure A1** in Appendix). The earliest diverging groups of embryophytes, the liverworts, mosses, and hornworts are all primitive in the sense that they are non-vascular plants (Shaw et al., 2011). Following the evolution of vascular tissue in plants the lycophytes emerged and came to dominate the flora during the Carboniferous period with some species growing to heights of 30 m (Banks, 2009). Even though all members of this group have leaf-like structures, these structures are not considered to be true leaves and only have a central vein (Banks, 2009). True leaves evolved in the lineage leading to ferns and spermatophytes and by the time angiosperms split from gymnosperms the ability to produce seeds had emerged in spermatophytes.

The MIPs of terrestrial plants can be divided into seven subfamilies, the plasma membrane intrinsic proteins (PIPs), the tonoplast intrinsic proteins (TIPs), the nodulin 26-like intrinsic proteins (NIPs), the small basic intrinsic proteins (SIPs), the X intrinsic proteins (XIPs), the hybrid intrinsic proteins (HIPs), and the GlpF-like intrinsic proteins (GIPs). While the moss *Physcomitrella patens* contains representatives from all seven groups the dicots at most have five of these subfamilies, the PIPs, TIPs, NIPs, SIPs, and XIPs and monocots only four, having lost also the XIPs (Danielson and Johanson, 2008). Whereas the number of subfamilies has decreased during the evolution of land plants the number of MIP isoforms present in each species seems to have increased. In *P. patens* only 23 isoforms are present while in higher angiosperms like *A. thaliana* and *O. sativa* 35 and 31 isoforms have been identified respectively (Johanson et al., 2001; Sakurai et al., 2005). This difference is mainly due to an expansion of the PIP, TIP, and NIP subfamilies which is also manifested by the appearance of new subgroups within the two latter subfamilies. The current understanding of the evolution of the MIP family in terrestrial plants is based on analyses of MIPs in a moss and in higher angiosperms, representing lineages which split more than 440 million years ago. In order to get a more comprehensive picture of the evolution of the plant MIP family it is necessary to investigate MIPs encoded in genomes of species belonging to lineages diverging later than mosses but before the emergence of higher angiosperms.

In the present study we have classified and annotated the MIPs encoded in the recently sequenced genome of the lycophyte *Selaginella moellendorffii* and complemented the analysis by MIPs encoded in ESTs of gymnosperms and basal angiosperms to provide a more detailed picture of the evolution of MIPs in terrestrial plants.

RESULTS

IDENTIFICATION AND CLASSIFICATION OF *S. MOELLENDORFFII* MIPs

The diploid genome of *S. moellendorffii* available at Joint Genome Institute was searched using the complete set of *Physcomitrella patens* MIPs as queries. Out of 20 unique hits one was deemed to be a pseudo NIP gene after manual inspection, containing just one NPA-box and only parts of helix 1–4. No MIP encoding sequence was found at this locus on the homologous chromosome. Of the remaining 19 sequences approximately half already had a satisfactory gene model, defining the coding sequence and exon/intron boarders, whereas new models encoding more typical MIP features were created for the rest (**Table 1**; **Table A1** in Appendix).

The identified *S. moellendorffii* MIPs (*Sm*MIPs) were then subjected to phylogenetic analyses (**Figure 1**). Of the 19 *Sm*MIPs, one clearly grouped with the SIP1 isoforms (SIP1s), three with the XIPs, and two with the HIPs. Hence, no further analysis was required for the classification of these MIPs and they were named *Sm*SIP1;1, *Sm*XIP1;1–*Sm*XIP1;3, *Sm*HIP1;1, and *Sm*HIP1;2. Three of the *Sm*MIPs clustered with the PIPs, whereof one was firmly nested within the PIP1s whereas the remaining two together with the PIP2s and *Pp*PIP3;1 ended up basal to the PIP1s with an unresolved internal relationship. Two sequences clustered with the TIPs and eight with the NIPs. Although the classification into the seven different subfamilies was straightforward, the phylogenies within the PIPs, TIPs, and NIPs were largely unresolved. In an attempt to achieve a more precise classification of these MIPs, individual phylogenetic analyses based on separate alignments, were performed. These analyses also included the available relevant MIPs from plants lacking sequenced genomes but since these are likely to represent only a fraction of the MIPs encoded in these species, no names were suggested to leave the classification of these MIPs open for future more comprehensive analyses.

PHYLOGENY OF PIPs

All three *S. moellendorffii* PIPs have ar/R filters identical or very similar to those in other PIPs and are likely to be specific for water (**Table 1**; **Figure A2** in Appendix). The PIP consensus tree (**Figure 2**) shows a clearly defined PIP1 clade with a bootstrap support of 99% and include representatives from the gymnosperms *P. glauca*, *P. taeda* and one sequence from *S. moellendorffii*. This is consistent with the initial analysis (**Figure 1**), alas the phylogeny of the other PIPs in part remained unresolved. On the node basal to the PIP1s, the PIP2s of *A. thaliana*, *O. sativa*, and *P. patens*, as well as 2 sequences from *S. moellendorffii* and 11 from *P. glauca* clustered into 7 separate clades. Based on the phylogenetic analysis and a manual comparison to the reference PIPs, these *S. moellendorffii* PIPs were clearly not PIP1s and in contrast to the *Pp*PIP3;1 they had the N- and C-terminal lengths characteristic of PIP2s, containing the among PIP2s conserved C-terminal phosphorylation site SFRS (Johansson et al., 2000). They were therefore classified as PIP2s and hence the *S. moellendorffii* PIPs were named *Sm*PIP1;1, *Sm*PIP2;1, and *Sm*PIP2;2. In accordance with earlier analyses the two PIPs from the algae *Coccomyxa* C-169 were basal to all other PIPs (Anderberg et al., 2011).

PHYLOGENY OF TIPs

The TIP6s of *P. patens* together with two sequences from the liverwort *Marchantia polymorpha* and one from the quillwort *Isoetes lacustris* formed three separate groups on the most basal node of the TIP consensus tree (**Figure 3**). On the next node two sequences from *S. moellendorffii* formed a sister clade to the TIPs of seed plants and to indicate an orthologous relationship with the TIP6s in *P. patens*, the *Sm*TIPs were named TIP6;1 and TIP6;2. The TIPs of seed plants formed five stable clades, TIP1 to TIP5. The clades of TIP1s and TIP2s included sequences from gymnosperms, basal angiosperms, monocots, and dicots, and the TIP4 clade sequences from gymnosperms, monocots, and dicots. In contrast, the TIP3s and TIP5s contained monocot and dicot sequences only. As shown by the bootstrap support the TIP3s and the TIP5s are clearly associated with the TIP1s and the TIP2s, respectively. To facilitate a comparison between phylogenetic relationship and possible substrate specificity the ar/R filters of the different subgroups are shown in **Figure 3**.

PHYLOGENY OF NIPs

Previous studies of NIPs in higher plants have established at least three major groups, NIP1, NIP2, and NIP3 (Sakurai et al., 2005; Zardoya, 2005; Gupta and Sankararamakrishnan, 2009; Danielson and Johanson, 2010). In our analysis the NIP2s and NIP3s were well supported but to resolve the NIP1 group in the consensus tree it was necessary to lower the cut off value for collapsing nodes slightly from the standard 50% to 49% (**Figure 4**). Of these three

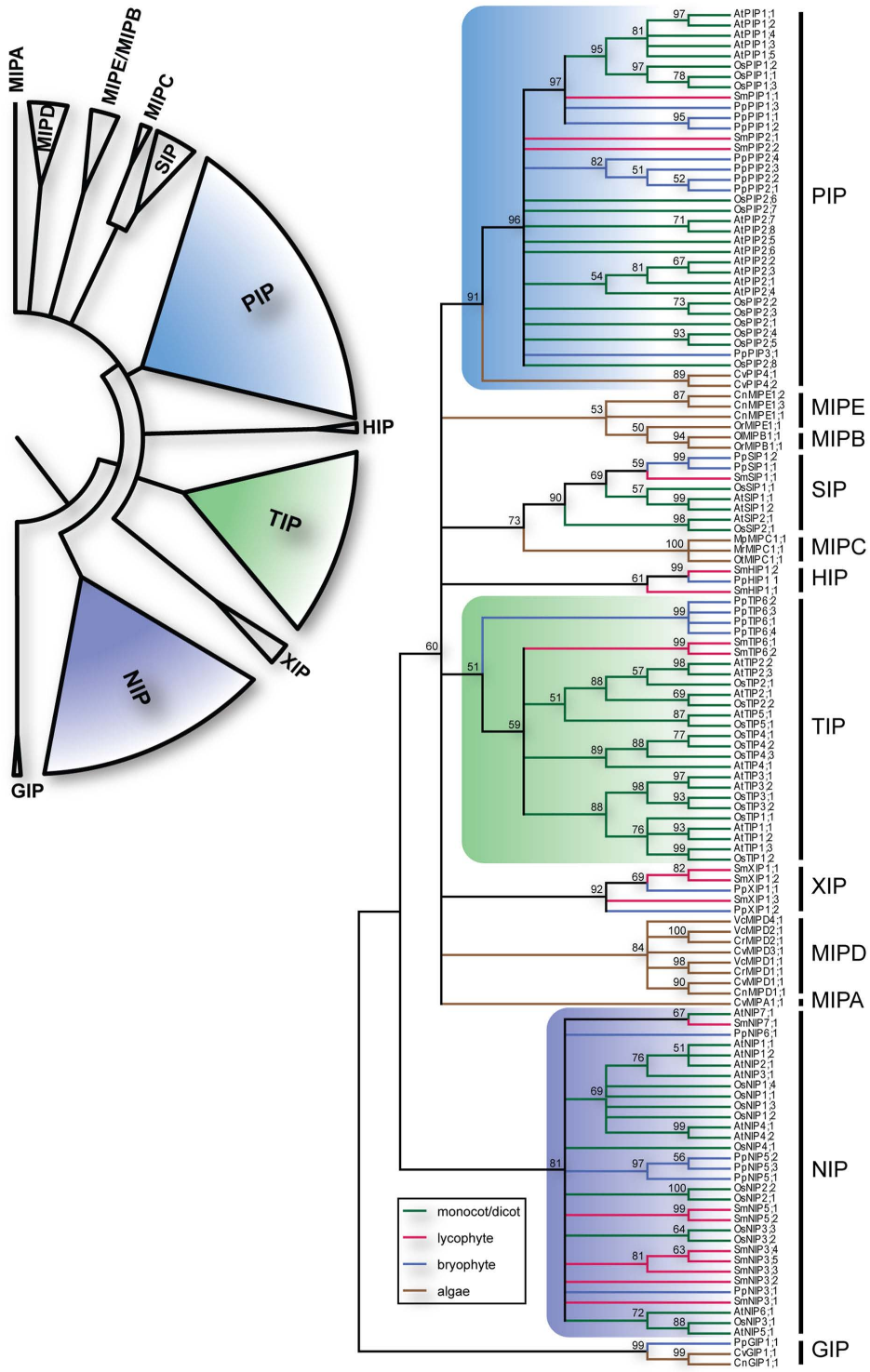


FIGURE 1 | Phylogeny of MIPs. The upper left panel summarize the different subfamilies of plant MIPs. The right panel depicts the bootstrap consensus tree of the complete set of MIPs from *O. sativa* (Os), *A. thaliana* (At), *S. moellendorffii* (Sm), *P. patens* (Pp) and the chlorophyte algae *Chlamydomonas reinhardtii* (Cr), *Volvox carteri* (Vc), *Coccomyxa* sp. C-169 (Cc), *Chlorella* sp. NC64A (Cn), *Micromonas pusilla* CCMP1545 (Mp), *Micromonas* sp. RCC299 (Mr), *Ostreococcus lucimarinus* (Ol), *Ostreococcus* sp. RCC809 (Or), and *Ostreococcus tauri* (Ot) using the Maximum Likelihood method. The branches are colored according to from what phyla the sequences are derived. The numbers by the nodes are bootstrap support in percentage and nodes with less the 50% support are collapsed. The vertical lines to the right delimit the different subfamilies.

Table 1 | Selectivity filters and NPA boxes of *S. moellendorffii* MIPs.

Subfamily	Protein[a]	ar/R selectivity filter[b]				NPA motifs[c]	
		H2	H5	LE_1	LE_2	Loop B	Loop E
PIP	*Sm*PIP1;1	F	H	T	R	NPA	NPA
	*Sm*PIP2;1	F	H	S	R	NPA	NPA
	*Sm*PIP2;2	F	H	T	R	NPA	NPA
	***Pp*PIP1;1**	**F**	**H**	**T**	**R**	**NPA**	**NPA**
TIP	*Sm*TIP6;1	H	I	G	R	NPA	NPA
	*Sm*TIP6;2	H	I	G	R	NPA	NPA
	***Pp*TIP6;1**	**H**	**I**	**A**	**R**	**NPA**	**NPG**
NIP	*Sm*NIP3;1	S	I	A	R	NPS	NPV
	*Sm*NIP3;2	P	N	A	R	NPS	NPA
	*Sm*NIP3;3	A	N	A	R	NPA	NPC
	*Sm*NIP3;4	A	N	A	R	NPA	NPI
	*Sm*NIP3;5	S	I	A	R	NPA	NPV
	***Pp*NIP3;1**	**A**	**I**	**A**	**R**	**NPA**	**NPV**
	*Sm*NIP5;1	F	A	A	R	NPA	NPA
	*Sm*NIP5;2	N	N	A	R	NPA	NPA
	***Pp*NIP5;1**	**F**	**A**	**A**	**R**	**NPA**	**NPA**
	*Sm*NIP7;1	A	V	G	R	NPS	NPA
	***Pp*NIP6;1**	**G**	**V**	**A**	**R**	**NPA**	**NPM**
SIP	*Sm*SIP1;1	I	I	P	N	NPT	NPA
	***Pp*SIP1;1**	**V**	**V**	**P**	**N**	**NPT**	**NPA**
XIP	*Sm*XIP1;1	H	V	A	R	NPC	NPA
	*Sm*XIP1;2	H	I	A	R	NPC	NPA
	*Sm*XIP1;3	L	T	A	R	NPI	NPA
	***Pp*XIP1;1**	**Q**	**I**	**A**	**R**	**NPC**	**NPG**
HIP	*Sm*HIP1;1	H	H	A	R	NPA	NPA
	*Sm*HIP1;2	H	H	A	R	NPA	NPA
	***Pp*HIP1;1**	**H**	**H**	**A**	**R**	**NPA**	**NPA**

[a] *MIPs from P. patens are included as a reference and are written in bold.*

[b] *The ar/R filter is defined by one amino acid residue in helix 2, one in helix 5, and two in loop E. See* ***Figure A2*** *in Appendix.*

[c] *The sequences corresponding to the NPA motifs in loops B and E.*

groups the NIP3s included sequences from *P. patens*, *S. moellendorffii*, gymnosperms and angiosperms, the NIP2s sequences from gymnosperms and angiosperms, and the NIP1s sequences from angiosperms only. The *S. moellendorffii* sequences in the NIP3 group were named *Sm*NIP3;1 to *Sm*NIP3;5. In addition to the three major clades, two sequences from *S. moellendorffii* and one from the fern *Adiantum capillus-veneris* formed a separate clade as did the *P. patens* NIP5s. The ar/R filters suggest that at least one of these *Sm*NIPs and the fern sequence *Ac*MIP1 are functionally equivalent to the *Pp*NIP5s and in accordance with the phylogenetic analysis, the *Sm*NIPs of this group were both classified as NIP5s and were hence named *Sm*NIP5;1 and *Sm*NIP5;2. However it should be noted that these two sequences are not very similar and the constriction region of the latter deviates from other NIP5s and instead show some resemblance to *Sm*NIP3;2–3;4. Finally, a single sequence from *S. moellendorffii* was named *Sm*NIP7;1 since it clustered with NIP7;1 from *A. thaliana* and also share the same ar/R filter. In contrast to the *Sm*NIPs, *Pp*NIP6;1, *Os*NIP4;1 and two gymnosperm sequences showed no apparent association with any other sequence.

DISCUSSION

Selaginella moellendorffii constitute the smallest genome of a sequenced land plant but nevertheless this spikemoss overall encode more protein families than the about twofold larger genome of the moss *P. patens* (Banks et al., 2011), consistent with having a more intricate physiology. Despite this general increase in the number of protein families the number of MIPs in *P. patens* and in *S. moellendorffii* are very similar, 23 and 19 respectively. Thus there is no evidence of an overall expansion of the MIP family coinciding with the evolution of primitive vascular plants like the lycophytes. This suggests that the number of encoded MIPs instead increased at a later stage in the evolution of terrestrial plants with the emergence of spermatophytes (**Table 2**). In our analysis six of the seven MIP subfamilies previously identified in the moss *P. patens* are confirmed in *S. moellendorffii*. This indicates that the seventh subfamily, the GIPs, was lost early on in the lineage leading to vascular plants, already before the appearance of the lycophytes. Interestingly, also the number of members within some of the subfamilies differs substantially between *P. patens* and *S. moellendorffii*. Each of the

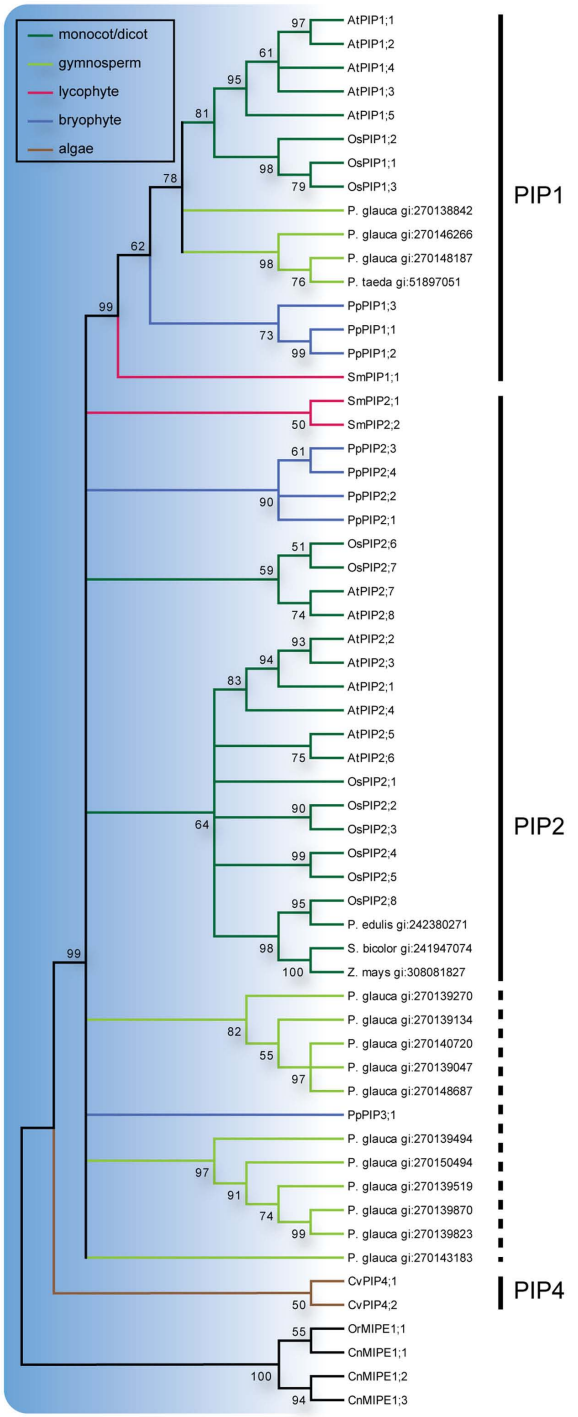


FIGURE 2 | Phylogeny of PIPs. Maximum Likelihood bootstrap consensus tree of PIPs from plant species with sequenced genomes together with PIP-like sequences retrieved from the NCBI EST database. EST sequences are named with species name followed by their GenBank gi numbers. The PIP branches are color coded according to from what phyla the sequences are derived. Subgroups are delimited by the vertical lines to the left where the dashed line indicates sequences for which classification is uncertain. The robustness of nodes is denoted with bootstrap support in percentage and nodes with less than 50% support are collapsed.

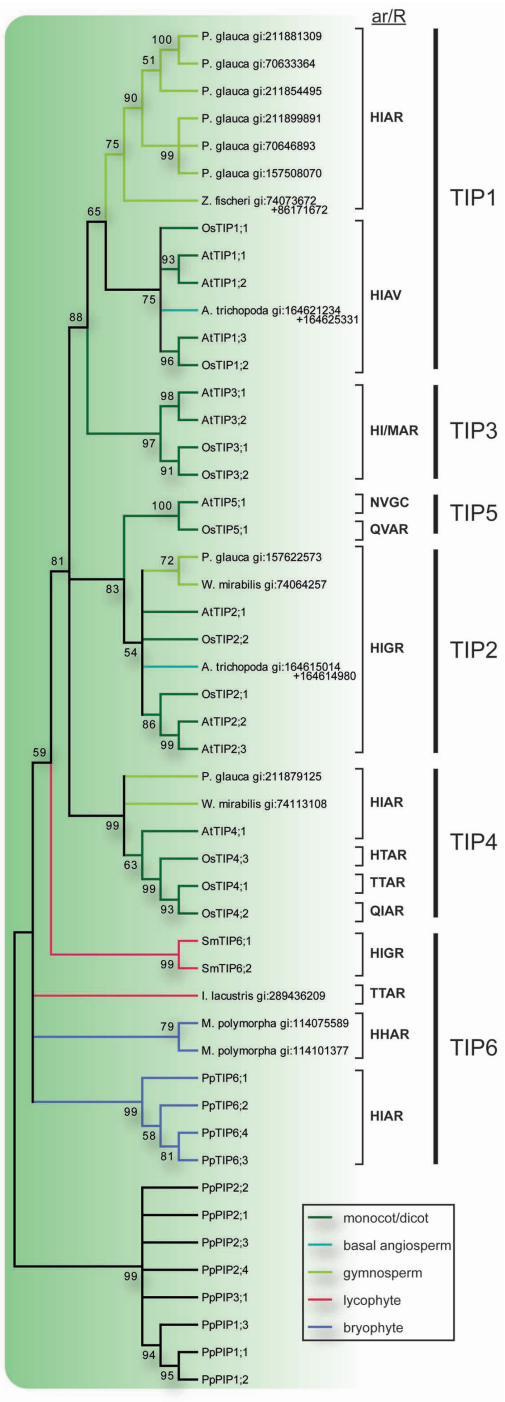


FIGURE 3 | Phylogeny of TIPs. The bootstrap consensus tree resulting from a Maximum Likelihood analysis of the TIPs of species with sequenced genome along with full length TIP-like EST sequences retrieved from the NCBI EST database. EST sequences are named with species name and GenBank gi number. In the case where two ESTs were used to compile a full length MIP sequence both GenBank gi numbers are provided. Selectivity filters (ar/R) are displayed next to the sequence names and the vertical lines to the right indicate the subgroup to which they belong. The TIP branches are color coded according to from what phyla the sequences are derived. Nodes with less than 50% bootstrap support are collapsed.

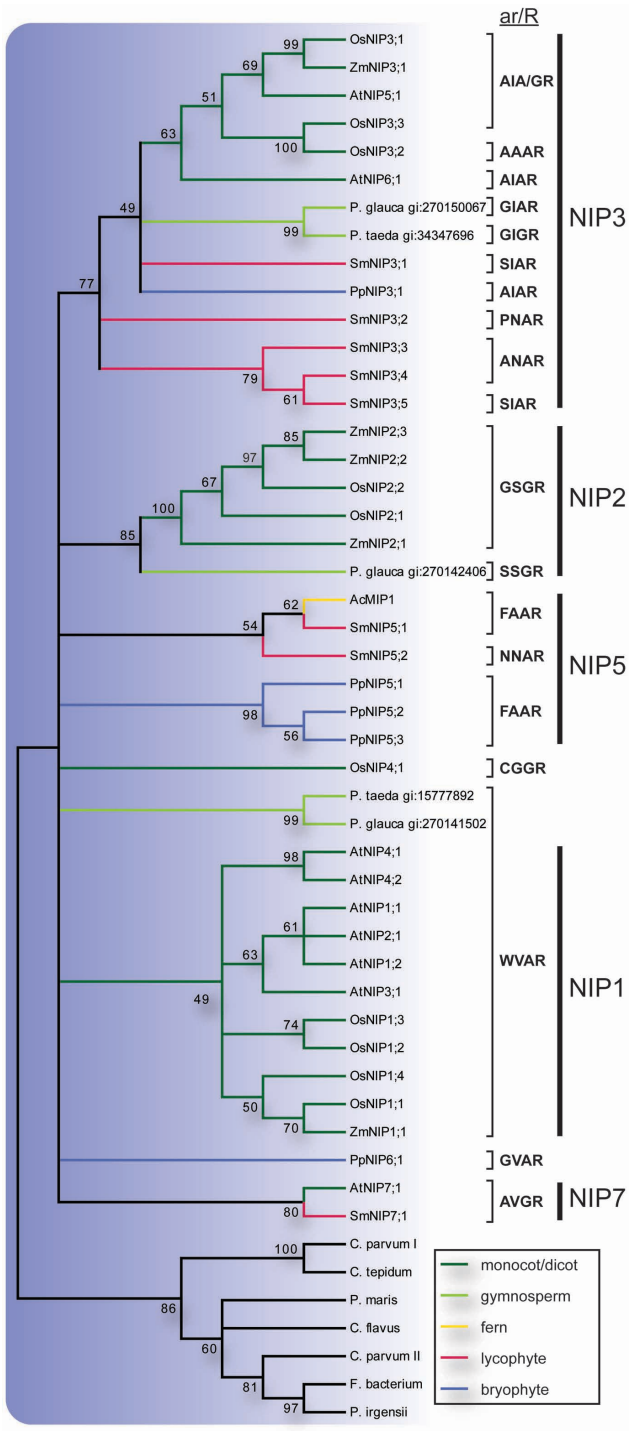


FIGURE 4 | Phylogeny of NIPs. The bootstrap consensus tree resulting from a Maximum Likelihood analysis of the NIPs of species with sequenced genome along with full length NIP-like EST sequences retrieved from the NCBI EST database. EST sequences are named with species name and GenBank gi number. Selectivity filters (ar/R) are displayed next to the sequence names and the vertical lines to the right indicate the subgroup to which they belong. The NIP branches are color coded according to from what phyla the sequences are derived. Nodes with less than 49% bootstrap support are collapsed.

six subfamilies found in *S. moellendorffii* is discussed in detail below.

SURPRISINGLY FEW PIPs IN *S. MOELLENDORFFII* AND UNRESOLVED PIP2 GROUPS IN SEED PLANTS

The PIPs are the most conserved plant MIPs and form a monophyletic clade including the basal algal PIP4s. Within this clade the PIP1 group is well defined and has a congruent topology with the exception of *Sm*PIP1;1 which appear more basal than the moss PIP1s. Despite this small inconsistency with the phylogeny of the species, the present result suggests that all terrestrial plants have PIP1 orthologs, implying a fundamental function in land plants. Unexpectedly, compared to the three *PIP1* genes found in *P. patens*, the *S. moellendorffii* genome only encode one and may therefore provide a useful naturally non-redundant system for addressing the physiological function of PIP1s. In addition to PIP1s it is clear that all of the terrestrial plants have at least one other type of PIPs, even though the phylogeny of these PIPs is not fully resolved. In *P. patens* PIP2s and a PIP3 have earlier been defined as separate groups although the functional integrity of the latter has been questioned (Danielson and Johanson, 2008) and therefore the remaining two *Selaginella* sequences were annotated as PIP2s. Interestingly, both the angiosperm and gymnosperm non-PIP1s form two distinct clades with unresolved relationship and it is conceivable that these clades correspond to two distinct isoforms of PIP2s with specialized functions which evolved already in a common ancestor to gymnosperms and angiosperms. All the PIPs appear to be classical water channels when the ar/R filter is considered; the algal PIP4s have the residues FHCR at the filter, the *Pp*PIPs and PIPs in higher plants FHTR, and the *Sm*PIPs FHTR or FHSR. Thus, any functional difference between different clades of PIPs is more likely to relate to regulatory features rather than substrate specificity. However, it should be noted that all PIPs in terrestrial plants are likely to share some regulatory features as key amino acid residues of the proposed gating mechanism are conserved (D28, E31, S115, H193 in SoPIP2;1; Tornroth-Horsefield et al., 2006; Anderberg et al., 2011).

TIPs DIVERSIFIED AFTER THE EMERGENCE OF PRIMITIVE VASCULAR PLANTS

In higher plants like *Arabidopsis* and rice five distinct subgroups of TIPs, TIP1 to TIP5, are discerned, whereas in more primitive plants like mosses and spikemosses only one type of TIP, TIP6, is found (Danielson and Johanson, 2008). Although there are several isoforms of TIP6s in the latter species, they are all very similar within each species, indicative of a specific distinct physiological role of the TIP6s at least within the species. The *SmTIP6* genes share the same gene structure with two introns which is also conserved in *TIPs* of higher plants (Johanson et al., 2001), whereas only one of these introns is present in *PpTIP6s*, suggesting that the conserved *TIP* gene structure evolved in early vascular plants. Interestingly, ESTs corresponding to *TIP6s* in the drought resistant spikemoss *S. lepidophylla* were the seventh most abundant transcript, estimated to make up 0.5% of the total ESTs after 2.5 h of dehydration (Iturriaga et al., 2006). This implies that the TIP6s might have a direct role in water relations.

Table 2 | Overview of MIP isoforms in terrestrial plants.

	Number of isoforms in subfamilies							Total isoforms	Total subfamilies	Reference
	GIP	HIP	XIP	SIP	PIP	TIP	NIP			
Physcomitrella patens	**1** (4.3)[b]	**1** (4.3)	**2** (8.7)	**2** (8.7)	**8** (34.8)	**4** (17.4)	**5** (21.7)	**23**	**7**	Danielson and Johanson (2008)
Selaginella moellendorffii	–	**2** (10.5)	**3** (15.8)	**1** (5.3)	**3** (15.8)	**2** (10.5)	**8** (42.1)	**19**	**6**	This work
Solanum lycopersicum[a]			**1** (2.7)	**3** (8.1)	**18** (48.6)	**9** (24.3)	**6** (16.2)	**37**	**5**	Sade et al. (2009)
Populus trichocarpa	–	–	**6** (10.9)	**6** (10.9)	**15** (27.3)	**17** (30.9)	**11** (20.0)	**55**	**5**	Gupta and Sankararamakrishnan (2009)
Gossypium hirsutum	–	–	**1** (1.4)	**7** (9.9)	**28** (39.4)	**23** (32.4)	**12** (16.9)	**71**	**5**	Park et al. (2010)
Arabidopsis thaliana	–	–	–	**3** (8.6)	**13** (37.1)	**10** (28.6)	**9** (25.7)	**35**	**4**	Johanson et al. (2001)
Oryza sativa	–	–	–	**2** (6.1)	**11** (33.3)	**10** (30.3)	**10** (30.3)	**33**	**4**	Sakurai et al. (2005)

[a]*The occurrence of MIPs in S. lycopersicum is based on an extensive analysis of ESTs.*

[b]*Relative size of subfamily in the species (%).*

The presented phylogenetic analysis suggests that the first subgroups of TIPs appeared after the lycophytes had diverged from other vascular plants, when the TIP1s, TIP2s, and TIP4s evolved in the lineage leading to higher plants. This diversification happened before the angiosperm/gymnosperm split and these three groups are thus expected to be present in all seed plants. Furthermore, the well supported clustering of TIP1s with TIP3s and TIP2s with TIP5s suggest that each pair evolved from a common ancestor. Since TIP3s and TIP5s are basal to TIP1s and TIP2s, respectively, this implies that TIP3s and TIP5s were also present in the lineage leading to the gymnosperms but later lost. However, it should be noted that the apparent absence of TIP3s and TIP5s in gymnosperms is based on searches in EST libraries and the picture might change as genomic sequences become available for analyses.

Typically the ar/R filters of TIPs have a characteristic histidine at the first position followed by a large aliphatic residue (I/M/V) at the second, a small amino acid residue (A/G) at the third and the among MIPs conserved large charged arginine at the fourth position. Based on the frequencies of occurrence, the original filter in the last common ancestor of all TIPs was suggested to be HIAR (Danielson and Johanson, 2008). There are however some deviations from this consensus, e.g., the TIP6s from *M. polymorpha* similar to the HIPs have a histidine also at the second position of the filter, TIP5s and *Os*TIP4;2 have an asparagine or a glutamine at the first position, and one of the TIP4s from rice together with a TIP from quillwort have a threonine at the first and second positions. Most notably the TIP1s in angiosperms lack the arginine at the fourth position and instead have a valine (HIAV). Functional studies suggest that TIPs in general are permeable for both water and ammonia, whereas the TIPs missing the arginine have a wider range of substrates including hydrogen peroxide and urea (Soto et al., 2008; Azad et al., 2011). Based on the ar/R filter we speculate that the gymnosperm TIP1s might functionally correspond to the TIP3s but only partly to the TIP1s in angiosperms since they are not expected to permeate either hydrogen peroxide or urea. A similar physiological function is supported by the fact that the expression of the TIP3s have been reported to be seed specific (Maurel et al., 1997; Gattolin et al., 2011) and at least one gymnosperm TIP1-like MIP (MIPFG) is also expressed in the protein storage vacuoles of seeds (Hakman and Oliviusson, 2002).

NIPs CONSTITUTE THE MOST NUMEROUS SUBFAMILY IN *S. MOELLENDORFFII*

Similar to the TIPs, the NIPs form a highly divergent subfamily with large variation in ar/R filters (**Figure 4**). Based on the ar/R residues at least three functional groups have been recognized (Wallace and Roberts, 2005; Mitani et al., 2008; Rouge and Barre, 2008; Ali et al., 2009) and substrate specificities have been reviewed extensively (Ma et al., 2008; Ali et al., 2009; Bienert and Jahn, 2010). It has been noted that the NIPs are surprisingly numerous in primitive plants and at least one of the phylogenetical subgroups (NIP3) had evolved already in a common ancestor of mosses and higher plants (Danielson and Johanson, 2008). The analyzed *Sm*MIPs support this finding and almost half (8) of the 19 *Sm*MIPs are NIPs, where of 5 are classified as NIP3s. Two of the *Sm*NIPs are proposed to be orthologous to *Pp*NIP5s based on the common ar/R-filter FAAR. This filter is also present in bacterial NIPs supporting that this represents the ancestral specificity of the plant NIPs (Danielson and Johanson, 2010). Interestingly, no NIP5 has yet been identified in higher plants. The somewhat similar ar/R filter of the NIP1s (WVAR) suggests that they have a similar substrate specificity and could therefore have replaced the NIP5s in higher plants. A single *Sm*NIP grouped with *At*NIP7;1 and both are likely to correspond to *Pp*NIP6;1 since the ar/R filters are related (AVGR and GVAR). Thus similar to NIP3s, the NIP6/7 function might have evolved already in a common ancestor of mosses and higher plants or at the latest in a common ancestor of vascular plants.

The low expression of NIPs in *Arabidopsis* suggests that transcripts encoding NIPs may be rare also among ESTs from other species (Alexandersson et al., 2005). This indicates that many NIPs are yet to be identified in species for which a complete genomic sequence is not available. Nevertheless, we do find gymnosperm ESTs encoding NIPs that firmly cluster with NIP3s and NIP2s. Although the NIP3 group is well supported there is a large variation in the ar/R filters. The *Sm*NIP3s can be divided into three types based on the ar/R filters (SIAR, PNAR, ANAR). Whether

these also translate into differences in substrate specificity and if they differ in specificity from the other NIP3s remain to be seen. The NIP2 group contains the best characterized NIPs regarding their physiological function and have been shown to have a role in the uptake and distribution of silicic acid within the plant (Ma et al., 2006). NIP2 homologs have previously only been reported from monocots and dicots (Ma et al., 2008) but the, in this study identified, gymnosperm NIP2 suggest that this physiological function was present in a common ancestor of gymnosperms and angiosperms. Two other gymnosperm NIPs appear to correspond to the NIP1s judging from a strictly conserved ar/R filter (WVAR), although the bootstrap support for such an association is weak. Hence, both NIP1s and NIP2s may be present in many seed plants although most dicots seem to have lost the NIP2s.

SIP2s EVOLVED LATER

The SIPs cluster together with the algal MIPC with high bootstrap support which may indicate an orthologous relationship. Intriguingly, previous analyses have indicated that MIPCs are actually more related to AQP11/12 in mammals than SIPs and this have hampered the classification of the MIPCs as SIPs (Anderberg et al., 2011). The two SIPs of *P. patens* and the single SIP of *S. moellendorffii* all clearly belong to the SIP1 group whereas the angiosperms also have a SIP2 group. The hydrophobic ar/R filter of the SIPs from moss and spikemoss also supports the SIP1 classification. Our result is consistent with previous analyses of ESTs (Johanson and Gustavsson, 2002) suggesting that SIP2s have only evolved in the angiosperms whereas SIP1s go deeper and are present in all terrestrial plants sharing a common ancestor with the mosses.

XIPs – ALTERNATIVE ALIGNMENTS AND FUNCTIONAL REDUNDANCY

There are three XIPs encoded in the genome of *S. moellendorffii* compared to two in *P. patens*. Previous analyses have shown that XIPs from moss and spikemoss are basal to all XIPs from higher plants (Gupta and Sankararamakrishnan, 2009). So far XIPs from higher plants have only been found in dicot plants and these cluster into XIP1s and XIP2s, although the latter have only been identified in two species (*Ricinus communis* and *Populus trichocarpa*). There have been several alternative suggestions for how H5 should be aligned to reference sequences and this has resulted in alternative ar/R filters, deviating at the second position. Here we suggest yet another alignment of H5, based on hydrophobicity plots (data not shown), where the predicted transmembrane region is shifted four residues toward the N-terminus of XIPs. Like earlier alignments this would preserve a glycine or an alanine at the position corresponding to G203 in SoPIP2;1, which is important for the close packing of H5 and H2 (Bansal and Sankararamakrishnan, 2007). More importantly, all XIPs will now have a conserved glycine also at the position corresponding to F207 in SoPIP2;1, which will release the structural constraint in H2 at the position corresponding to G82 in SoPIP2;1. This is consistent with the large size variation of amino acids (G/A/S/V/F) found at this position in the XIPs. This new alignment results in a slightly more hydrophobic filter with valine at the second position in the dicot XIPs. At this position *Sm*XIP1;3 has a threonine whereas the *Pp*XIP1;1 and *Sm*XIP1;1–1;2 have isoleucine or valine similar to dicot XIPs. Nevertheless, the *Pp*XIP1;1 and *Sm*XIP1;1–1;2 deviate from other XIPs in having a TIP-like ar/R filter with histidine in the first position. Interestingly, the XIPs have recently been lost in both monocots and in *Arabidopsis*. These evolutionary events as well as the change of ar/R filters in higher plants may have been possible due to a functional redundancy with TIPs. However, with the present understanding of MIP subcellular localization this scenario is unlikely since the XIPs of higher plants have been shown to be targeted to the plasma membrane (Bienert et al., 2011) whereas TIPs reside in the tonoplast (Jauh et al., 1999).

HIPs AND TIPs – SHARED ANCESTRY AND OVERLAPPING FUNCTION?

So far HIPs have only been identified in moss and spikemoss. The ar/R filter in all three HIPs is conserved (HHAR) and appears to be a hybrid of the PIP (FHTR) and TIP (HIAR) filters, having histidine at both the first and second position (Danielson and Johanson, 2008). To the best of our knowledge the substrate specificity of such a filter has not been tested. However, we note that it is identical to the ar/R filter in a TIP6 from *Marchantia polymorpha*, leading us to speculate that it is functionally equivalent to the filter in other basal TIPs (HIA/GR). Although the subcellular localization of HIPs is not known, a redundant function with the TIPs would explain why the HIPs were lost in higher plants and would fit in time with the expansion of the TIP subfamily. We also note that the lengths of the two first exons of *SmHIP1;1* are identical to *SmTIP6s*, and if an alternative gene model is used for this position in *SmHIP1;2* and *PpHIP1;1* (data not shown) the first intron position of *SmTIP6s* and *TIPs* of higher plants is conserved in both the *SmHIPs* and *PpHIP1;1*. This is a strong indication of a shared ancestry of this region in *HIPs* and *TIPs* where the first position of the ar/R filter is encoded. The precise details of this common evolutionary history are difficult to reconstruct from our current data.

GENERAL REMARKS AND FUTURE PERSPECTIVES

Land plants encode more MIP isoforms than any other type of organism. It has been suggested that the major expansion of MIPs occurred as an adaption to life on land, enabling the plant to exploit concentration gradients as well as to tolerate drought and hypo-osmotic stress (Anderberg et al., 2011). Here we have identified and analyzed MIPs encoded in the genome of the primitive vascular plant *S. moellendorffii* to investigate how the MIP family has evolved in terrestrial plants. The total number of *Sm*MIPs is similar to that of moss and six of the seven MIP subfamilies in moss are also found in *S. moellendorffii*. Of the six subfamilies, the NIP subfamily is dominating in *S. moellendorffii* whereas both the PIPs and TIPs are uniquely few compared to all other land plants. Based on EST data a second expansion of the MIP family appears to coincide with the evolution of spermatophytes. It seems likely that this later expansion was a consequence of the development of more specialized types of tissue such as those present in the seed and the flower organs, requiring a more complex regulation of water and solute transport. Future analyses of sequenced genomes will hopefully corroborate the more precise timing of this event and provide more information on the transition from fresh water to life on land.

MATERIALS AND METHODS

IDENTIFICATION AND ANNOTATION OF *S. MOELLENDORFFII* MIPs

The diploid *S. moellendorffii* genome, available at Joint Genome Institute[1], was searched for MIPs using tblastn with *P. patens* MIPs as queries. In subsequent rounds of searches identified *S. moellendorffii* MIPs were included until no more MIPs could be found. The allele to be included in subsequent analyses was chosen randomly.

The genomic sequence around the hits was then checked for existing gene models. The models were evaluated and kept if they were found to accurately represent a MIP. New gene models were created if they were found to correspond poorly to the sequences of known MIPs with respect to conserved residues and lengths of predicted loops and transmembrane helices as well as conserved intron positions (**Table 2**). If no satisfactory gene model existed or could be created for a hit and no allelic variant could be found it was deemed to be a pseudo-gene and was excluded from further analyses.

SEQUENCE ALIGNMENTS

The alignments used for the phylogenetic analyses was created in the program MEGA5 (Tamura et al., 2011) and was based on a structural alignment of *Bt*AQP1 (1J4N, Sui et al., 2001), *Ec*GLPF (1LDA, Fu et al., 2000), *Ec*AQPZ (1RC2, Savage et al., 2003), *So*PIP2;1 (1Z98, Tornroth-Horsefield et al., 2006), *Rn*AQP4 (2D57, Hiroaki et al., 2006), *Mm*AQPM (2F2B, Lee et al., 2005), *Pf*AQP (3C02, Newby et al., 2008), *Hs*AQP5 (3D9S, Horsefield et al., 2008), and *Bt*AQP0 (2B6P, Gonen et al., 2005) made in DeepView/Swiss-PdbViewer v4.0.1. All MIP sequences from *Arabidopsis thaliana*, *Physcomitrella patens* and from the chlorophyte algae species *Chlamydomonas reinhardtii*, *Volvox carteri*, *Coccomyxa* sp. C-169, *Chlorella* NC64A, *Micromonas pusilla* CCMP1545, *Micromonas* RCC299, *Ostreococcus lucimarinus*, *Ostreococcus* RCC809, and *Ostreococcus tauri* (Anderberg et al., 2011) as well as those of *Oryza sativa* were then added and manually aligned to the initial structural alignment. Accession numbers are given in **Table A2** in Appendix. Finally the MIP sequences identified in *S. moellendorffii* were added and aligned. Since the N- and C-terminal regions were not included in the subsequent phylogenetic analyses no effort was put into aligning these.

SUBSET ALIGNMENTS

Based on the resulting alignment three subsets including only PIPs, TIPs, or NIPs were created. The PIP alignment included all PIP sequences from the original alignment with the addition of PIPs encoded by full length ESTs from the gymnosperms *Picea glauca* and *Pinus taeda*. In preliminary analyses *Os*PIP2;8 associated with *Pp*PIP3;1 however this was not observed in the final analysis when PIP2;8-like sequences from the monocots *Phyllostachys edulis*, *Sorghum bicolor*, and *Zea mays* were included. The algal MIPEs were included as an out-group. The NIP alignment, including all the NIP sequences from the original alignment, was supplemented with NIPs encoded by full length EST sequences identified in blast searches with selected NIPs against the NCBI database[2]. The searches were restricted to species belonging to viridiplantae, excluding monocots and dicots. To this alignment bacterial NIPs (bNIPs) which are the closest homologs to plant NIPs (Danielson and Johanson, 2010) were added as out-group. The TIP alignment was created in the same way but here the *P. patens* PIPs were included as out-group. If two ESTs from the same organism were found to overlap in such a way that they covered a whole coding sequence, the encoded MIP was also included.

PHYLOGENETIC ANALYSES

The phylogenetic analyses were performed using the Maximum Likelihood algorithm in MEGA5. The best substitution model for each alignment was determined within the MEGA 5 program and were rtREV + G + I + F for the alignment including all subfamilies, JTT + G for the PIP alignment and WAG + G + I + F for both the TIP and NIP alignments. For all alignments the number of discrete gamma categories was set to 5, all sites were used, the heuristic method was set to Nearest-Neighbor-Interchange and the initial tree was made automatically. For all analyses the robustness of the resulting best trees were assessed by 1000 bootstrap replications.

ACKNOWLEDGMENTS

We are grateful to the U.S. Department of Energy Joint Genome Institute for sequencing the genome of *S. moellendorffii* and making the sequences available to the public. We would also like to thank Jonas Danielson for sharing valuable insights on the evaluation of gene models. This work was supported by funding from the Swedish Research Council.

[1]http://www.jgi.doe.gov/

[2]http://www.ncbi.nlm.nih.gov/

REFERENCES

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

Ali, W., Isayenkov, S. V., Zhao, F. J., and Maathuis, F. J. (2009). Arsenite transport in plants. *Cell. Mol. Life Sci.* 66, 2329–2339.

Anderberg, H. I., Danielson, J. A., and Johanson, U. (2011). Algal MIPs, high diversity and conserved motifs. *BMC Evol. Biol.* 11, 110.

Azad, A. K., Yoshikawa, N., Ishikawa, T., Sawa, Y., and Shibata, H. (2011). Substitution of a single amino acid residue in the aromatic/arginine selectivity filter alters the transport profiles of tonoplast aquaporin homologs. *Biochim. Biophys. Acta* 1818, 1–11.

Banks, J. A. (2009). *Selaginella* and 400 million years of separation. *Annu. Rev. Plant Biol.* 60, 223–238.

Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., Depamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G., Dacre, M., Debarry, J., Dreyer, I., Elias, M., Engstrom, E. M., Estelle, M., Feng, L., Finet, C., Floyd, S. K., Frommer, W. B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K. G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T. P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T., Mueller-Roeber, B., Nelson, D. R., Obara, M., Oguri, Y., Olmstead, R. G., Onodera, N., Petersen, B. L., Pils, B., Prigge, M., Rensing, S. A., Riano-Pachon, D. M., Roberts, A. W., Sato, Y., Scheller, H. V., Schulz, B., Schulz, C., Shakirov, E. V., Shibagaki, N., Shinohara, N., Shippen, D. E., Sorensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theissen, G., Ulvskov, P., Wakazuki, S., Weng, J. K., Willats, W. W., Wipf, D., Wolf, P. G., Yang, L., Zimmer, A. D., Zhu, Q., Mitros, T., Hellsten, U., Loque, D., Otillar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D.,

and Grigoriev, I. V. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963.

Bansal, A., and Sankararamakrishnan, R. (2007). Homology modeling of major intrinsic proteins in rice, maize and *Arabidopsis*: comparative analysis of transmembrane helix association and aromatic/arginine selectivity filters. *BMC Struct. Biol.* 7, 27. doi:10.1186/1472-6807-7-27

Beitz, E., Wu, B., Holm, L. M., Schultz, J. E., and Zeuthen, T. (2006). Point mutations in the aromatic/arginine region in aquaporin 1 allow passage of urea, glycerol, ammonia, and protons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 269–274.

Bienert, G. P., Bienert, M. D., Jahn, T. P., Boutry, M., and Chaumont, F. (2011). Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates. *Plant J.* 66, 306–317.

Bienert, G. P., and Jahn, T. P. (2010). Major intrinsic proteins and arsenic transport in plants: new players and their potential role. *Adv. Exp. Med. Biol.* 679, 111–125.

Danielson, J. A., and Johanson, U. (2008). Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biol.* 8, 45. doi:10.1186/1471-2229-8-45

Danielson, J. A., and Johanson, U. (2010). Phylogeny of major intrinsic proteins. *Adv. Exp. Med. Biol.* 679, 19–31.

de Groot, B. L., Frigato, T., Helms, V., and Grubmuller, H. (2003). The mechanism of proton exclusion in the aquaporin-1 water channel. *J. Mol. Biol.* 333, 279–293.

Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J., and Stroud, R. M. (2000). Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* 290, 481–486.

Gattolin, S., Sorieul, M., and Frigerio, L. (2011). Mapping of tonoplast intrinsic proteins in maturing and germinating *Arabidopsis* seeds reveals dual localization of embryonic TIPs to the tonoplast and plasma membrane. *Mol. Plant* 4, 180–189.

Gonen, T., Cheng, Y., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S. C., and Walz, T. (2005). Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. *Nature* 438, 633–638.

Gupta, A. B., and Sankararamakrishnan, R. (2009). Genome-wide analysis of major intrinsic proteins in the tree plant *Populus trichocarpa*: characterization of XIP subfamily of aquaporins from evolutionary perspective. *BMC Plant Biol.* 9, 134. doi:10.1186/1471-2229-9-134

Hakman, I., and Oliviusson, P. (2002). High expression of putative aquaporin genes in cells with transporting and nutritive functions during seed development in Norway spruce (*Picea abies*). *J. Exp. Bot.* 53, 639–649.

Heymann, J. B., and Engel, A. (1999). Aquaporins: phylogeny, structure, and physiology of water channels. *News Physiol. Sci.* 14, 187–193.

Hiroaki, Y., Tani, K., Kamegawa, A., Gyobu, N., Nishikawa, K., Suzuki, H., Walz, T., Sasaki, S., Mitsuoka, K., Kimura, K., Mizoguchi, A., and Fujiyoshi, Y. (2006). Implications of the aquaporin-4 structure on array formation and cell adhesion. *J. Mol. Biol.* 355, 628–639.

Horsefield, R., Norden, K., Fellert, M., Backmark, A., Tornroth-Horsefield, S., Terwisscha Van Scheltinga, A. C., Kvassman, J., Kjellbom, P., Johanson, U., and Neutze, R. (2008). High-resolution x-ray structure of human aquaporin 5. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13327–13332.

Iturriaga, G., Cushman, M. A. F., and Cushman, J. C. (2006). An EST catalogue from the resurrection plant *Selaginella lepidophylla* reveals abiotic stress-adaptive genes. *Plant Sci.* 170, 1173–1184.

Jauh, G. Y., Phillips, T. E., and Rogers, J. C. (1999). Tonoplast intrinsic protein isoforms as markers for vacuolar functions. *Plant Cell* 11, 1867–1882.

Johanson, U., and Gustavsson, S. (2002). A new subfamily of major intrinsic proteins in plants. *Mol. Biol. Evol.* 19, 456–461.

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

Johansson, I., Karlsson, M., Johanson, U., Larsson, C., and Kjellbom, P. (2000). The role of aquaporins in cellular and whole plant water balance. *Biochim. Biophys. Acta* 1465, 324–342.

Lee, J. K., Kozono, D., Remis, J., Kitagawa, Y., Agre, P., and Stroud, R. M. (2005). Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 Å. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18932–18937.

Leliaert, F., Verbruggen, H., and Zechman, F. W. (2011). Into the deep: new discoveries at the base of the green plant phylogeny. *Bioessays* 33, 683–692.

Ma, J. F., Tamai, K., Yamaji, N., Mitani, N., Konishi, S., Katsuhara, M., Ishiguro, M., Murata, Y., and Yano, M. (2006). A silicon transporter in rice. *Nature* 440, 688–691.

Ma, J. F., Yamaji, N., Mitani, N., Xu, X. Y., Su, Y. H., Mcgrath, S. P., and Zhao, F. J. (2008). Transporters of arsenite in rice and their role in arsenic accumulation in rice grain. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9931–9935.

Maurel, C., Chrispeels, M., Lurin, C., Tacnet, F., Geelen, D., Ripoche, P., and Guern, J. (1997). Function and regulation of seed aquaporins. *J. Exp. Bot.* 48, 421–430.

Mitani, N., Yamaji, N., and Ma, J. F. (2008). Characterization of substrate specificity of a rice silicon transporter, Lsi1. *Pflugers Arch.* 456, 679–686.

Newby, Z. E., O'Connell, J. III, Robles-Colmenares, Y., Khademi, S., Miercke, L. J., and Stroud, R. M. (2008). Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite *Plasmodium falciparum*. *Nat. Struct. Mol. Biol.* 15, 619–625.

Park, W., Scheffler, B. E., Bauer, P. J., and Campbell, B. T. (2010). Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). *BMC Plant Biol.* 10, 142. doi:10.1186/1471-2229-10-142

Preston, G. M., Carroll, T. P., Guggino, W. B., and Agre, P. (1992). Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein. *Science* 256, 385–387.

Rouge, P., and Barre, A. (2008). A molecular modeling approach defines a new group of Nodulin 26-like aquaporins in plants. *Biochem. Biophys. Res. Commun.* 367, 60–66.

Sade, N., Vinocur, B. J., Diber, A., Shatil, A., Ronen, G., Nissan, H., Wallach, R., Karchi, H., and Moshelion, M. (2009). Improving plant stress tolerance and yield production: is the tonoplast aquaporin SlTIP2;2 a key to isohydric to anisohydric conversion? *New Phytol.* 181, 651–661.

Sakurai, J., Ishikawa, F., Yamaguchi, T., Uemura, M., and Maeshima, M. (2005). Identification of 33 rice aquaporin genes and analysis of their expression and function. *Plant Cell Physiol.* 46, 1568–1577.

Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O'Connell, J. D. III, and Stroud, R. M. (2003). Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z. *PLoS Biol.* 1, e72. doi:10.1371/journal.pbio.0000072

Shaw, A. J., Szovenyi, P., and Shaw, B. (2011). Bryophyte diversity and evolution: windows into the early evolution of land plants. *Am. J. Bot.* 98, 352–369.

Soto, G., Alleva, K., Mazzella, M. A., Amodeo, G., and Muschietti, J. P. (2008). AtTIP1;3 and AtTIP5;1, the only highly expressed *Arabidopsis* pollen-specific aquaporins, transport water and urea. *FEBS Lett.* 582, 4077–4082.

Sui, H., Han, B. G., Lee, J. K., Walian, P., and Jap, B. K. (2001). Structural basis of water-specific transport through the AQP1 water channel. *Nature* 414, 872–878.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

Tornroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., Neutze, R., and Kjellbom, P. (2006). Structural mechanism of plant aquaporin gating. *Nature* 439, 688–694.

Wallace, I. S., and Roberts, D. M. (2005). Distinct transport selectivity of two structural subclasses of the nodulin-like intrinsic protein family of plant aquaglyceroporin channels. *Biochemistry* 44, 16826–16834.

Zardoya, R. (2005). Phylogeny and evolution of the major intrinsic protein family. *Biol. Cell* 97, 397–414.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 October 2011; accepted: 01 February 2012; published online: 20 February 2012.

*Citation: Anderberg HI, Kjellbom P and Johanson U (2012) Annotation of Selaginella moellendorffii major intrinsic proteins and the evolution of the protein family in terrestrial plants. Front. Plant Sci. **3**:33. doi: 10.3389/fpls.2012.00033*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

APPENDIX

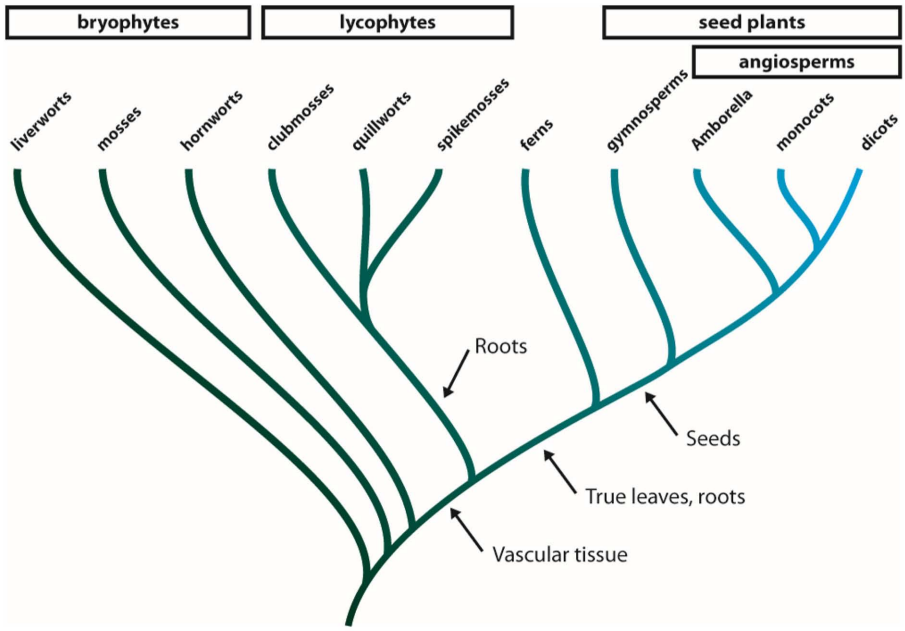


FIGURE A1 | Phylogeny of land plants. The land plants (embryophytes) can be divided into four groups. The bryophytes are a paraphyletic group comprised of the three earliest diverging lineages of land plants, i.e., the liverworts, mosses, and hornworts. The next group to diverge is the monophyletic lycophytes that consists of clubmosses, quillworts, and spikemosses and further up the tree the ferns represent a lineage diverging just before the emergence of seed plants. The seed plants are divided into two groups, the gymnosperms and the angiosperms (flowering plants). Within the angiosperms *Amborella* is thought to be the earliest diverging genus. Arrows indicate important morphological changes during the evolution of land plants.

Table A1 | Gene structure of *S. moellendorffii* MIPs.

Name	Sc[a]	Str.[b]	Start pos[c]	E1[d]	I1[d]	E2	I2	E3	I3	E4	I4	E5
*Sm*PIP1;1	sc 43	–	543709	322	206	296	195	141	79	87		
*Sm*PIP2;1	sc 266	+	14192	292	244	296	255	141	69	126		
*Sm*PIP2;2	sc 17	–	1553601	289	71	302	107	141	76	141		
*Sm*TIP6;1	sc 1	+	4467056	127	83	248	61	360				
*Sm*TIP6;2	sc 1	–	4470607	127	58	248	104	360				
*Sm*NIP3;1	sc 23	+	391409	156	145	225	61	195	79	62	68	193
*Sm*NIP3;2	sc 21	–	95220	96	64	222	63	195	56	62	50	196
*Sm*NIP3;3	sc 30	+	1231132	162	49	225	2408	195	59	62	64	181
*Sm*NIP3;4	sc 3	+	1364278	120	78	225	68	263	129	151		
*Sm*NIP3;5	sc 7	–	3041227	108	70	225	59	198	54	62	56	187
*Sm*NIP5;1	sc 2	+	2404270	180	70	225	159	195	59	62	88	193
*Sm*NIP5;2	sc 75	–	422849	162	60	225	95	198	50	62	196	226
*Sm*NIP7;1	sc 13	–	1722259	147	62	228	56	198	59	62	52	175
*Sm*SIP1;1	sc 51	+	562196	288	127	273	61	150				
*Sm*XIP1;1	sc 1	–	2139783	123	58	828						
*Sm*XIP1;2	sc 1	–	2093139	102	74	792						
*Sm*XIP1;3	sc 4	–	630059	111	56	936						
*Sm*HIP1;1	sc 3	–	3858367	127	57	248	56	417				
*Sm*HIP1;2	sc 9	+	2872669	48	61	118	113	191	55	187	58	197

[a] *Scaffold on which the gene is located.*

[b] *Encoding strand (+ or –).*

[c] *Position of first nucleotide on scaffold.*

[d] *Length of exons (E) and introns (I).*

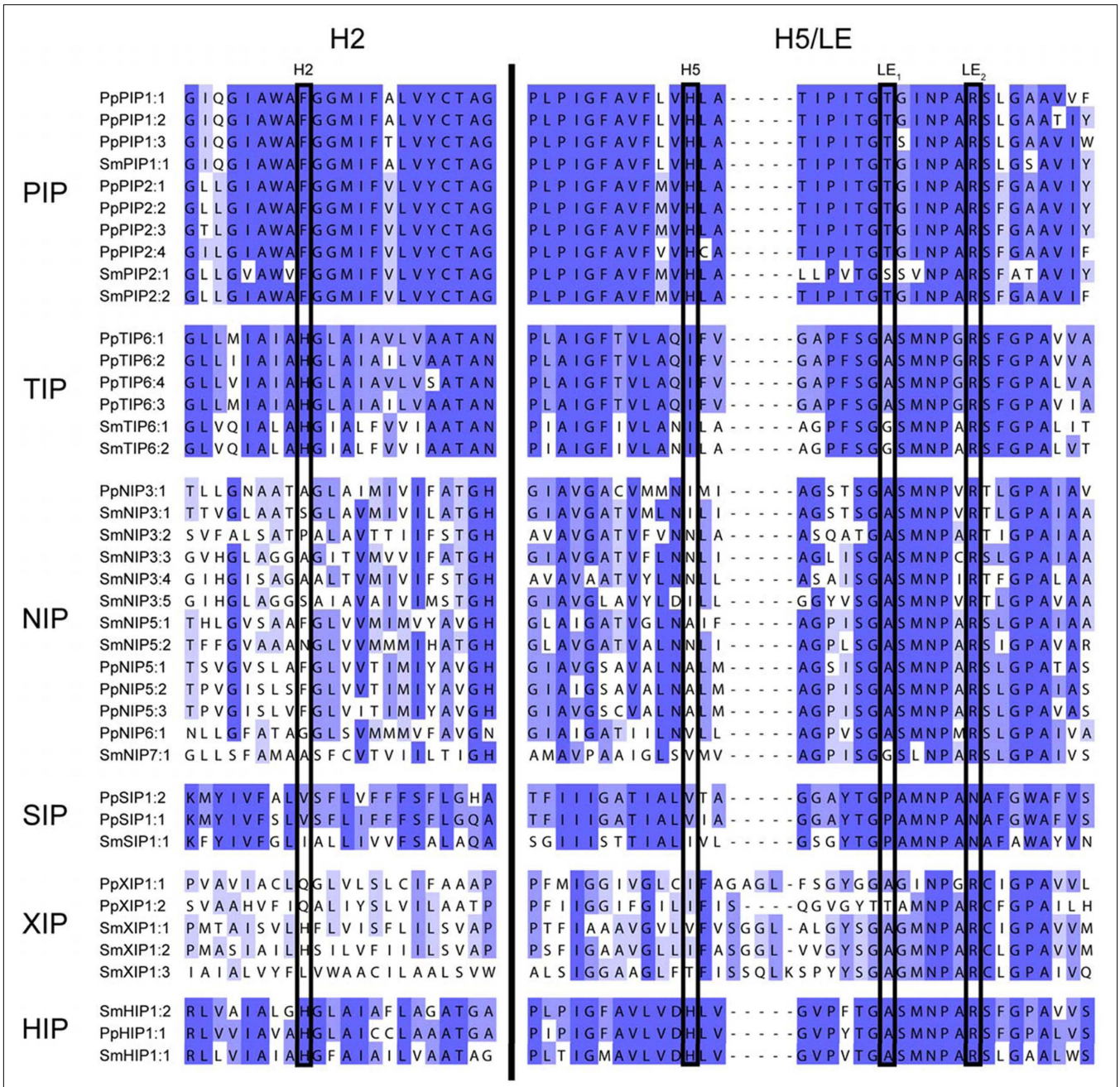


FIGURE A2 | Alignment of regions determining the ar/R filter. Two regions of the alignment, helix 2 (H2) and helix 5/loop E (H5/LE) separated by a black bar, containing the four residues of the ar/R selectivity filter (boxed). All MIPs of *S. moellendorffii* and *P. patens* are included in the alignment. The blue shading reflects the degree of conservation within each subfamily.

Table A2 | Major intrinsic protein accession numbers.

Taxa name	Species	Accession No.		Taxa name	Species	Accession No.	
		GenBank[a]	Phytozome[b]			GenBank[a]	Phytozome[b]
*Pp*PIP1;1	*Physcomitrella patens*		Pp1s1_535V6.4	*At*NIP3;1		AAG50717	
				*At*NIP4;1		BAB10360	
*Pp*PIP1;2			Pp1s102_107V6.1	*At*NIP4;2		BAB10361	
*Pp*PIP1;3			Pp1s305_12V6.1	*At*NIP5;1		CAB39791	
*Pp*PIP2;1			Pp1s8_151V6.1	*At*NIP6;1		AAF14664	
*Pp*PIP2;2			Pp1s55_301V6.2	*At*NIP7;1		AAF30303	
*Pp*PIP2;3			Pp1s267_61V6.1	*At*SIP1;1		AAF26804	
*Pp*PIP2;4			Pp1s118_199V6.1	*At*SIP1;2		BAB09487	
*Pp*PIP3;1			Pp1s17_281V6.1[c]	*At*SIP2;1		CAB72165	
*Pp*TIP6;1			Pp1s44_31V6.1	*Os*PIP1;1	*Oryza sativa*	BAD28398	
*Pp*TIP6;2			Pp1s156_153V6.1	*Os*PIP1;2		EAZ31617	
*Pp*TIP6;3			Pp1s101_226V6.1	*Os*PIP1;3		BAD22920	
*Pp*TIP6;4			Pp1s184_96V6.1	*Os*PIP2;1		BAC15868	
*Pp*NIP3;1			Pp1s258_69V6.1	*Os*PIP2;2		BAD23735	
*Pp*NIP5;1			Pp1s13_445V6.1	*Os*PIP2;3		CAD41442	
*Pp*NIP5;2			Pp1s91_35V6.1	*Os*PIP2;4		BAC16113	
*Pp*NIP5;3			Pp1s37_249V6.1	*Os*PIP2;5		BAC16116	
*Pp*NIP6;1			[d]	*Os*PIP2;6		CAE05002	
*Pp*SIP1;1			Pp1s3_429V6.1	*Os*PIP2;7		BAD46581	
*Pp*SIP1;2			Pp1s475_9V6.3	*Os*PIP2;8		AAP44741	
*Pp*XIP1;1			Pp1s31_73V6.1	*Os*TIP1;1		AAK98737	
*Pp*XIP1;2			Pp1s32_353V6.1	*Os*TIP1;2		BAB63833	
*Pp*HIP1;1			[d]	*Os*TIP2;1		BAD25765	
*Pp*GIP1;1			Pp1s283_16V6.1	*Os*TIP2;2		BAD61899	
*At*PIP1;1	*Arabidopsis thaliana*	CAB71073		*Os*TIP3;1		AAG13544	
				*Os*TIP3;2		CAE05657	
*At*PIP1;2		AAC28529		*Os*TIP4;1		AAS98488	
*At*PIP1;3		AAF81320		*Os*TIP4;2		BAA92993	
*At*PIP1;4		AAF02782		*Os*TIP4;3		BAA92991	
*At*PIP1;5		CAA20461		*Os*TIP5;1		BAF15407	
*At*PIP2;1		CAB67649		*Os*NIP1;1		BAD27715	
*At*PIP2;2		AAD18142		*Os*NIP1;2		BAD73177	
*At*PIP2;3		AAD18141		*Os*NIP1;3		AAV44140	
*At*PIP2;4		BAB09839		*Os*NIP1;4		BAD53665	
*At*PIP2;5		CAB41102		*Os*NIP2;1		BAD16128	
*At*PIP2;6		AAC79629		*Os*NIP2;2		BAD37471	
*At*PIP2;7		CAA17774		*Os*NIP3;1		AAG13499	
*At*PIP2;8		AAC64216		*Os*NIP3;2		BAC99758	
*At*TIP1;1		AAD31569		*Os*NIP3;3		BAC65382	
*At*TIP1;2		BAB01832		*Os*NIP4;1		BAB61180	
*At*TIP1;3		AAC62778		*Os*SIP1;1		BAB32914	
*At*TIP2;1		BAB01264		*Os*SIP2;1		ABF95655	
*At*TIP2;2		CAB10515		*Cc*MIPA1;1	*Coccomyxa C-169*	[e]	
*At*TIP2;3		BAB09071					
*At*TIP3;1		AAG52132		*Cc*MIPD1;1			
*At*TIP3;2		AAF97261		*Cc*MIPD3;1			
*At*TIP4;1		AAC42249		*Cc*PIP4;1			
*At*TIP5;1		CAB51216		*Cc*PIP4;2			
*At*NIP1;1		CAA16760		*Cc*GIP1;1			
*At*NIP1;2		CAA16748		*Cn*MIPD1;1	*Chlorella NC64A*	[e]	
*At*NIP2;1		AAC26712					

(Continued)

Table A2 | Continued

Taxa name	Species	Accession No.	
		GenBank[a]	Phytozome[b]
*Cn*MIPE1;1			
*Cn*MIPE1;2			
*Cn*MIPE1;3			
*Cn*GIP1;1			
*Cr*MIPD1;1	*Chlamydomonas reinhardtii*	[e]	
*Cr*MIPD2;1			
*Mp*MIPC1;1	*Micromonas pusilla CCMP1545*	[e]	
*Mr*MIPC1;1	*Micromonas RCC299*	[e]	
*Ol*MIPB1;1	*Ostreococcus lucimarinus*	[e]	
*Or*MIPB1;1	*Ostreococcus RCC809*	[e]	
*Or*MIPE1;1			
*Ot*MIPC1;1	*Ostreococcus tauri*	[e]	
C. parvum I	*Chlorobaculum parvum*	ACF10864	
F. bacterium	*Flavobacteriales bacterium*	EDP70495	
P. irgensii	*Polaribacter irgensii*	EAR12182	
C. flavus	*Chthoniobacter flavus*	EDY20715	
P. maris	*Planctomyces maris*	EDL59061	
C. parvum II	*Chlorobaculum parvum*	ACF11962	
C. tepidum	*Chlorobium tepidum*	AAM72699	
*Zm*NIP1;1	*Zea mays*	AAK26750	
*Zm*NIP2;1		AAK26751	
*Zm*NIP2;2		AAK26752	
*Zm*NIP2;3		AAK26849	
*Zm*NIP3;1		AAK26753	
*Ac*MIP1	*Adiantum capillus-veneris*	BAB12437	

[a] *GenBank accession number.*

[b] *Protein identifier Phytozome.*

[c] *The sequence at Phytozome is three nucleotides longer.*

[d] *Annotated in Danielson and Johanson (2008).*

[e] *Annotated in Anderberg et al. (2011).*

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 12 September 2011
doi: 10.3389/fpls.2011.00050

The metabolite transporters of the plastid envelope: an update

Fabio Facchinelli and Andreas P. M. Weber*

Institut für Biochemie der Pflanzen, Heinrich-Heine Universität Düsseldorf, Düsseldorf, Germany

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
Guillaume Pilot, Virginia Tech, USA
Mary E. Rumpho, University of Maine, USA

****Correspondence:***
Andreas P. M. Weber, Institut für Biochemie der Pflanzen, Heinrich-Heine Universität Düsseldorf, Universitätstrasse 1, D-40225 Düsseldorf, Germany.
e-mail: andreas.weber@uni-duesseldorf.de

The engulfment of a photoautotrophic cyanobacterium by a primitive mitochondria-bearing eukaryote traces back to more than 1.2 billion years ago. This single endosymbiotic event not only provided the early petroalgae with the metabolic capacity to perform oxygenic photosynthesis, but also introduced a plethora of other metabolic routes ranging from fatty acids and amino acids biosynthesis, nitrogen and sulfur assimilation to secondary compounds synthesis. This implicated the integration and coordination of the newly acquired metabolic entity with the host metabolism. The interface between the host cytosol and the plastidic stroma became of crucial importance in sorting precursors and products between the plastid and other cellular compartments. The plastid envelope membranes fulfill different tasks: they perform important metabolic functions, as they are involved in the synthesis of carotenoids, chlorophylls, and galactolipids. In addition, since most genes of cyanobacterial origin have been transferred to the nucleus, plastidial proteins encoded by nuclear genes are post-translationally transported across the envelopes through the TIC–TOC import machinery. Most importantly, chloroplasts supply the photoautotrophic cell with photosynthates in form of reduced carbon. The innermost bilayer of the plastidic envelope represents the permeability barrier for the metabolites involved in the carbon cycle and is literally stuffed with transporter proteins facilitating their transfer. The intracellular metabolite transporters consist of polytopic proteins containing membrane spans usually in the number of four or more α-helices. Phylogenetic analyses revealed that connecting the plastid with the host metabolism was mainly a process driven by the host cell. In *Arabidopsis*, 58% of the metabolite transporters are of host origin, whereas only 12% are attributable to the cyanobacterial endosymbiont. This review focuses on the metabolite transporters of the inner envelope membrane of plastids, in particular the electrochemical potential-driven class of transporters. Recent advances in elucidating the plastidial complement of metabolite transporters are provided, with an update on phylogenetic relationship of selected proteins.

Keywords: endosymbiosis, envelope membrane, translocator

INTRODUCTION

Oxygenic photosynthesis is the process by which plants are able to convert the solar energy into stable chemical bonds. By trapping the ephemeral energy contained in the photon, photosynthetic organisms are able to form and stabilize chemical bonds.

In the oxygenic photosynthesis reduced carbon is formed, using electrons extracted from water. The energy stored in the C–H bonds can be released during respiration, where the electrons are transferred back to oxygen, leading to the formation of water. In this way photosynthesis and respiration define a water–oxygen cycle, thereby re-distributing the solar energy to the whole biosphere (Hohmann-Marriott and Blankenship, 2011).

Long before oxygenic photosynthesis arose, the solar energy was trapped by non-oxygenic organisms, using hydrogen, ferrous iron, and hydrogen sulfide as the main electron source. Some of today's bacteria still use this source of electrons instead of water. These are the purple non-sulfur bacteria (e.g., *Rhodopseudomonas*), which perform cyclic electron flow around a PSII-like protein complex, and the green sulfur bacteria (e.g., *Chlorobium*), which possess a PSI-like reaction center able to extract electrons from hydrogen sulfide and transferring them to NAD^+ via linear electron transport (Allen and Martin, 2007; Hohmann-Marriott and Blankenship, 2011). Different hypotheses exist for explaining the events that gave rise to the oxygen-evolving protocyanobacterium, possessing both types of reaction centers and using water as electron donor. One states that lateral gene transfer occurred between a purple non-sulfur bacterium with a quinone-based (Type II) reaction center and a green sulfur bacterium with a FeS-based (Type I) reaction center, while in an alternate hypothesis the structural similarity of the two photosystems indicates that they might have evolved from a single ancestor via gene duplication and divergence (Xiong et al., 2000; Mulkidjanian et al., 2006; Allen and Martin, 2007).

Whatever model best explains the integration of the two photosystems, the key in evolving oxygenic photosynthesis is the development of the water-splitting oxygen-evolving complex (OEC),

containing four Mn atoms and a Ca atom capable of sequentially removing four electrons from two water molecules. This process ultimately releases molecular oxygen as by-product. The accumulation of O_2 converted the atmosphere into an oxidizing one, permitting the development of more complex life forms that utilize oxygen as electron acceptor during the aerobic respiration.

This all posed the bases for the next level of complexity: endosymbiotic associations. Eukaryotic origin is defined as being concomitant with the acquisition of mitochondria, likely an α-proteobacterium, either by a nucleus-bearing amitochondriate cell, or by a prokaryotic host cell, later evolving eukaryote-specific features (Embley and Martin, 2006).

Later on the engulfment of a photoautotrophic cyanobacterium by a primitive mitochondriate eukaryote established the lineage of the Archaeplastida, also known as the Plantae (Gould et al., 2008; Keeling, 2010). All plastids are believed to have originated after a single endosymbiotic event, more than 1.2 billion years ago (Björn, 2009). One exception is provided by *Paulinella chromatophora*, which represents a case of ongoing primary endosymbiosis between an ameba and a cyanobacterium-like endosymbiont (**Figure 1**; Kies, 1974). Soon after the primary endosymbiosis, three major autotrophic lineages arose: the glaucophytes, the green algae, and the red algae (**Figure 1**; Adl et al., 2005).

Glaucophytes are a small group of unicellular, chlorophyll *a* containing freshwater algae consisting of four genera with at least seven species. They were the first lineage branching off and contain a plastid (muroplast) with a thin peptidoglycan wall between the two envelope membranes (Martin and Herrmann, 1998; Steiner and Löffelhardt, 2002; Sato et al., 2009). The peptidoglycan layer is believed to be a relic of the cyanobacterial cell wall, thus supporting the endosymbiotic theory of plastid origin. Other similarities to the cyanobacterial progenitors are the maintenance of phycobilisomes as light-harvesting antenna and a carboxysomal-like carbon-concentrating mechanism (Burey et al., 2007). Red algae (rhodophytes) also possess plastids with phycobilisomes but

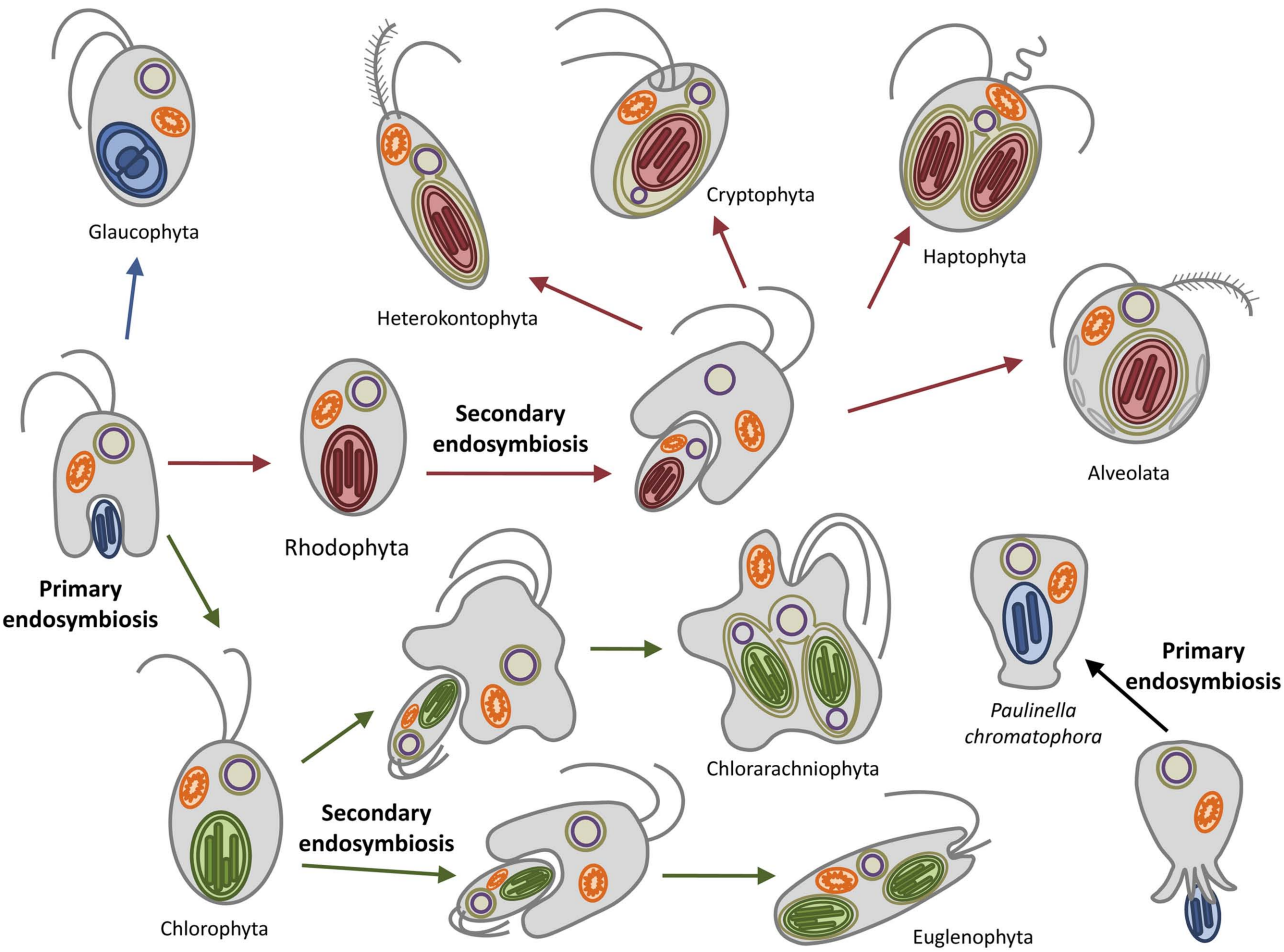


FIGURE 1 | Plastid evolution in photosynthetic eukaryotes. The uptake of a cyanobacterium resulted in a photosynthetic plantae ancestor which subsequently diverged in the three lineages containing primary plastids: the chlorophytes (including the land plants), the rhodophytes, and the glaucophytes. The subsequent secondary endosymbioses of green and red algae engulfed by different hosts resulted in the euglenophyta and chlorarachniophyta (greens) and in the possibly monophyletic chromalveolates (reds). *Paulinella chromatophora* represents an independent primary endosymbiosis in progress.

lack any remnant of the peptidoglycan wall and carboxysomes. In addition, they possess membrane-intrinsic light-harvesting complexes (LHCs) similar to those of the chloroplasts of the Viridiplantae associated with the PSI, thus representing a transitional state between cyanobacteria and chloroplasts (Wolfe et al., 1994). Rhodophyta have furthermore contributed to at least six other algal lineages by secondary endosymbiosis, as indicated by the presence of up to four membranes surrounding their plastids. These include members from the chromalveolates supergroup: the cryptophytes (still retaining a remnant nucleus, the nucleomorph), the haptophytes, the plastid-containing heterokontophyta (e.g., diatoms), and the alveolates. (**Figure 1**; Reyes-Prieto et al., 2007; Archibald, 2009).

The green lineage (chlorophyta) is represented by the green algae and the land plants. This is the most derived lineage. In contrast to glaucophytes and red algae (collectively termed bilihytes) their photosynthetic plastids lack the phycobilisomes; they possess instead chlorophyll *b*-containing LHCs associated with both photosystems as well as other accessory pigments (Keeling, 2010). A secondary endosymbiotic event involving one member of the green algae has been proposed to have occurred once, leading to the Cabozoa which include the euglenophyta and the chlorarachniophyta (**Figure 1**).

Even if the presence of the plastid is often associated with the capability to perform photosynthesis, a diverse number of lineages reduced the plastids metabolic capability to the point where photosynthesis was lost. These include for example the red algal-derived apicoplast of the malaria parasite *Plasmodium* (which still retains important anabolic pathways providing the host cell with fatty acids, isoprene units, and haem; Ralph et al., 2004) and secondary endosymbionts containing plastids of green origin occurring in the excavata, such as the colorless euglenozoa.

Primary plastids are delimited by two membranes, whose composition reflects their bacterial origin: besides the already mentioned peptidoglycan layer, the outer envelope membrane contains galactolipids and β-barrel proteins (Jarvis et al., 2000; Schleiff et al., 2003). However, envelope membranes were subjected to extensive modifications: whereas the inner envelope derives from the cyanobacterial plasma membrane, the outer envelope is of chimeric origin, with the outer leaflet containing lipids derived from the host endomembrane system (Cavalier-Smith, 2000).

The events occurring at the interface between the endosymbiont and the host were of crucial importance for the establishment of the plastid. Insertion of host protein translocators into the envelope was necessary to take full advantage of the newly acquired metabolic entity. Besides the photosynthetic reactions that lead to carbon fixation, endosymbiosis introduced into the petroalgae a set of metabolic pathways, including nitrogen and sulfur assimilation, the biosynthesis of amino acids, fatty acids, vitamins, hormones, and a plethora of secondary compounds (Weber and Flügge, 2002; Weber et al., 2006; Weber and Fischer, 2007). Metabolite translocators were therefore of utmost importance for integrating and coordinating the flux of precursors and products between the plastid and other cellular compartments. The metabolite traffic between plastids and other cellular compartments is predominantly facilitated by antiporters embedded in the inner membrane. The choice of antiporter systems ensures that the exchange of a substrate on one side of the membrane is always accompanied by the presence of another counter-exchange substrate on the *trans*-side (Flügge, 1995). This antiport transport mechanism is common to other subcellular compartments such as mitochondria, peroxisomes, and the endomembrane system, and is important to link different metabolite pools subjecting them to a strict control. In the context of endosymbiosis, the regulated exchange of metabolites would have been of crucial importance in maintaining the metabolic homeostasis between the cyanobacterium and the host (Weber and Linka, 2011).

The outer envelope contains porin-like channels and has long been considered as not acting as a permeability barrier, proposing that solute movement across the outer envelope membrane is limited solely by size (Flügge, 2000). The discovery of low-affinity, high-specific channels in the outer envelope points to a more active role of the outer envelope, even if their regulatory role *in vivo* is still a matter of debate (Soll et al., 2000; Bölter and Soll, 2001; Duy et al., 2007). The reader is referred to recent reviews describing the outer envelope proteins (e.g., Duy et al., 2007; Inoue, 2011).

This review focuses on the metabolite transporters of the inner envelope membrane of plastids, in particular the electrochemical potential-driven class of transporters. Recent advances in elucidating the plastidial complement of metabolite transporters are provided, with an update on phylogenetic relationship of selected proteins.

PLASTIDIC PHOSPHATE TRANSLOCATORS

The plastidic phosphate translocator (pPT) family comprises four members of the drug/metabolite transporters (DMT) superfamily (TC 2.A.7; Saier et al., 2006, 2009). These antiporters catalyze the exchange of phosphorylated C3-, C5-, and C6-compounds for inorganic phosphate (P_i). The homo-exchange guarantees a balance in the phosphate content of the stroma and the cytosol and ensures a constant provision of phosphate to sustain ATP synthesis (Weber et al., 2005).

Phylogenetic analyses revealed that pPTs are monophyletic and originated from an existing endomembrane translocator (Weber et al., 2006). The major plant plastid translocators co-localize in branches containing homologs from red and green algae, supporting their monophyly and an origin of the translocators in their common ancestor (Weber et al., 2006). pPTs are more closely related to the nucleotide sugar transporters (NSTs) than to other families of the DMT superfamily (Knappe et al., 2003a).

This is in agreement with the need of the petroalgae to export the adenosine-diphosphoglucose (ADP-Glc) produced by the cyanobiont into the host cytosol. The first metabolic connection between the two compartments would thus have been through an ADP-Glc translocator. Indeed, it was recently shown that members of the NSTs possess an innate ability to transport ADP-Glc, filling the link between the carbon metabolism of endosymbiont and host cell (Colleoni et al., 2010). The establishment of a carbon flux was thus a host-driven process, by recruiting a translocator of its endomembrane system and redirecting it to the endosymbiont, later on evolving into the pPT family of transporters (Weber et al., 2006; Tyra et al., 2007; Colleoni et al., 2010). This hypothesis, however, awaits support from analyzing the plastid permeome of a member of the third lineage of the Archaeplastida, the

glaucophyta. The host contribution to the establishment of the endosymbiosis was further supported by a work in which transporter proteins broadly distributed in the plantae kingdom were subjected to phylogenetic analysis: out of 83 proteins scrutinized, 58% were shown to be of host origin (Tyra et al., 2007).

THE TRIOSE PHOSPHATE/PHOSPHATE TRANSLOCATOR

The plastid provides the cell with reduced carbon compounds by the assimilation of carbon dioxide during photosynthesis. In plants the photoassimilates are either exported in their phosphorylated form by members of the pPT family to fuel sucrose biosynthesis in the cytosol, or they are retained in the plastid to drive starch synthesis (**Figure 2**). The daily path for carbon export occurs in form of triose phosphates (TP) via the triose phosphate/phosphate translocator (TPT). TPT accepts TPs (dihydroxyacetone phosphate, DHAP, and glyceraldehyde 3-phosphate) as well as 3-phosphoglycerate (3-PGA) as substrates (**Table 1**; Fliege et al., 1978; Flügge and Heldt, 1984). In addition to its role in allocating the carbon between the stroma and the cytosol, the TPT has the ability to export reducing equivalents to the cytosol. The TP/3-PGA shuttle exports TPs, which are then converted by the glyceraldehyde phosphate dehydrogenase to 3-PGA, reducing one molecule of $NADP^+$ to NADPH, required for biosynthetic processes. The 3-PGA is then transported back to the chloroplast stroma by the TPT and can re-enter the Calvin–Benson cycle (**Figure 2**; Flügge and Heldt, 1984).

In the red alga *Galdieria sulphuraria* the TPT shows narrower substrate specificity, being able to transport TPs but not 3-PGA (**Figure 3B**). Moreover, the GsTPT displays a three-fold higher affinity to P_i and the substrate DHAP has a two-fold higher affinity to the P_i binding site than its counterpart from the chlorophytes (**Table 1**; Linka et al., 2008a). While in organisms descending from the green lineage the products of the carbon fixation are stored in insoluble starch granules inside the chloroplasts, rhodophytes accumulate a starch-like polymer in the cytosol through the use of UDP-glucose, called floridean starch, resembling amylopectin in its structure (Ball and Morell, 2003; Patron and Keeling, 2005). The heteroside floridoside is synthesized from UDP-Gal and glyceraldehyde 3-phosphate (Gly-3-P) in the cytosol as well and represents the main soluble pool of fixed carbon in analogy with sucrose in chlorophytes (**Figure 3B**; Viola et al., 2001). This means that the partitioning of the photoassimilates in red algae occurs exclusively in the cytosol, with GsTPT acting only as TP-exporter. This also explains the higher affinity of GsTPT for its substrates, enabling the export of TPs even at low concentrations (Linka et al., 2008a).

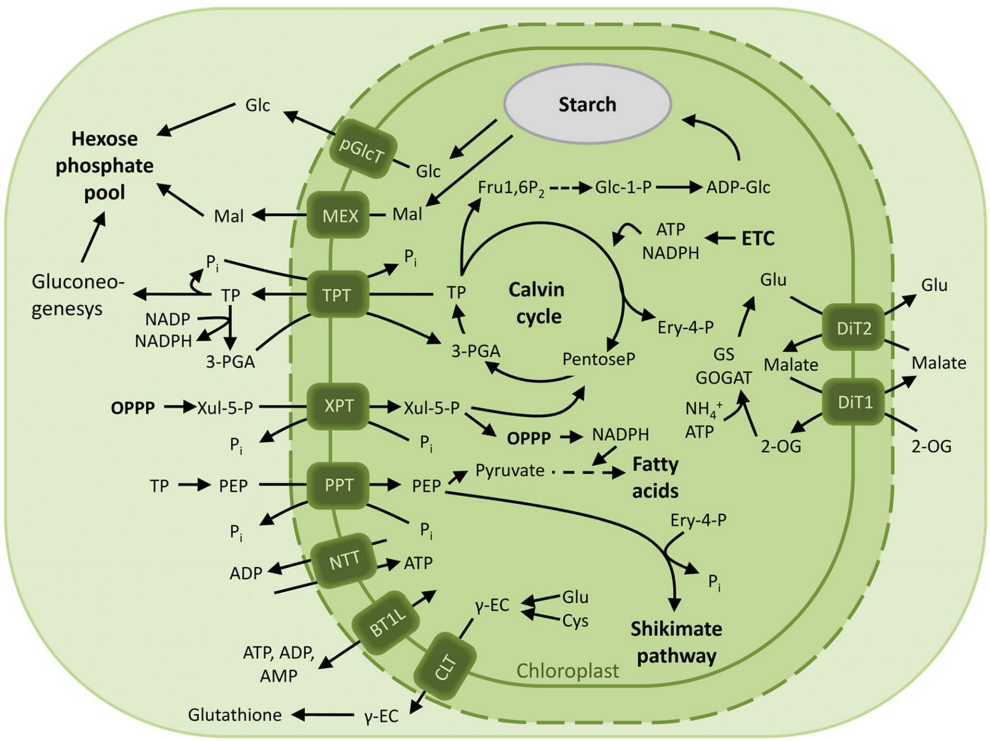


FIGURE 2 | Overview of the characterized chloroplast envelope metabolite transporters. Transport processes of carbon, nitrogen compounds and energy across the envelope membrane of photosynthetic plastids in green plants are depicted. Abbreviations: 2-OG, 2-oxoglutarate; ADP-Glc, ADP-glucose; BT1L, BT1-like transporter; CLT, CRT-like transporter; Cys, cysteine; Ery 4-P, Erythrose 4-phosphate; ETC, electron transport chain; $Fru1,6P_2$, fructose-1,6-bisphosphate; γ-EC, γ-glutamylcysteine; Glc, glucose; Glc 1-P, glucose 1-phosphate; Glc 6-P, glucose 6-phosphate; Glu, glutamine; GPT, glucose 6-phosphate/phosphate translocator; Mal, maltose; MEX, maltose exporter; NTT, nucleoside triphosphate transporter; OPPP, oxidative pentose phosphate pathway; 3-PGA, 3-phosphoglyceric acid; PEP, phospho*eno*l/pyruvate; pGlcT, plastidic glucose transporter; Pi, inorganic phosphate; PPT, PEP/phosphate translocator; TP, triose phosphate; TPT, triose phosphate/phosphate translocator; XPT, xylulose 5-phosphate/phosphate translocator; Xul 5-P, xylulose 5-phosphate.

Table 1 | Phosphate translocators of the plastidic inner envelope membrane.

		TPT (triose phosphate/phosphate translocator)	PPT (phospho-*enol*-pyruvate/phosphate translocator)	GPT (glucose-6-phosphate/phosphate translocator)	XPT (xylulose-5-phosphate/phosphate translocator
AGI number		At5g17630	At5g33320, At3g01550	At5g54800, At1g61800	At5g17630
TC number			TC 2.A.7.9		
Substrates	Green plants	TP, 3-PGA	PEP	Glc-6-P, TP, 3-PGA	TP, Xul-5-P, Ru-5-P, Ery-4-P
	Red algae	TP	PEP	n.d.	–
	Glaucophytes	–	–	–	–
	Apicomplexas cryptomonads	TP, PEP	–	–	–
Kinetic constant	Substrate				
K_m (app)	Phosphate	1.0	0.8	1.1	1.0
K_i (app)	TrioseP	1.0	8.0	0.6	0.4
K_i (app)	3-PGA	1.0	4.6	1.8	10.6
K_i (app)	PEP	3.3	0.3	2.9	2.1
K_i (app)	Glc-6-P	>50	>50	1.1	6.6
K_i (app)	Ery-4-P				3.3
K_i (app)	Xul-5-P				0.8
K_i (app)	Ru-5-P				3.5

Apparent K_m (phosphate) and K_i values expressed in millimolar for various phosphorylated metabolites. Transport activities of the recombinant translocators from green plants were measured in a reconstituted system using proteoliposomes pre-loaded with inorganic phosphate (Fischer et al., 1997; Kammerer et al., 1998; Eicks et al., 2002).

In addition to its ability to grow heterotrophically, *G. sulphuraria* can grow on more than 50 different carbon sources (Gross and Schnarrenberger, 1995; Barbier et al., 2005). Unlike land plants, where TPT is expressed in photosynthetic active tissues, GsTPT is also expressed in heterotrophically grown cells (Schulz et al., 1993; Linka et al., 2008a). Whereas plants provide heterotrophic plastids with glucose 6-phosphate (Glc6P) through the Glucose 6-phosphate/phosphate translocator (GPT), red algae do not appear to have pentose phosphate or hexose phosphate translocators (Oesterhelt et al., 2007).

Since the GsTPT is unable to transport 3-PGA, its role in the abovementioned NADPH shuttle can be excluded. While in plants the enzyme fructose-1,6-bisphosphatase is redox regulated and inactivated by oxidation, the *G. sulphuraria* homolog lacks redox modulation (Reichert et al., 2003). In this way the TPs can be converted into hexose phosphates in heterotrophic growth conditions and in the dark by the plastidic FBPase, ensuring the production of reduction equivalents (Reichert et al., 2003; Oesterhelt et al., 2007).

The expression pattern of GsTPT thus reflects the route for carbon import into heterotrophically grown cells with GsTPT as unique mediator of the carbon flux between plastid and cytosol (Linka et al., 2008a).

The lineage of the glaucophytes is best studied in the freshwater alga *Cyanophora paradoxa* (Löffelhardt et al., 1997). As for the red algae, glaucophytes synthesize and store the starch in the cytosol through an UDP-glucose pathway (Plancke et al., 2008). Transport assays with isolated cyanelles demonstrated the activity of a translocator transferring phosphate, DHAP, and 3-PGA, in analogy with chloroplasts of higher plants (**Figure 3A**; Schlichting and Bothe, 1993). It is expected that the ongoing *C. paradoxa* genome sequencing project will soon provide clues as to whether a TPT ortholog also exists in glaucophytes.

The plastids of organisms derived by secondary endosymbiosis can have up to four envelope membranes, like the apicoplast of the malaria parasite *Plasmodium falciparum*. The apicoplast is no longer able to perform photosynthetic carbon fixation, but retained the capacity to synthesize fatty acids, heme, iron sulfur clusters, and isoprene subunits (Ralph et al., 2004). The compartmentalization of these biosynthetic pathways poses as prerequisite the net import of carbon skeletons, ATP, and reducing equivalents. Indeed, *P. falciparum* has two TPTs, PfoTPT, and PfiTPT, previously shown to reside in the outermost and innermost membranes of the apicoplast (**Figure 4A**; Mullin et al., 2006).They both have a substrate preference for DHAP, 3-PGA, and phospho*enol*pyruvate (PEP), suggesting that they operate in tandem (**Table 1**; Lim et al., 2010). The discovery of PEP as substrate of a member of the TPTs is unprecedented and provides a plausible strategy to solve the needs of the apicoplast anabolism. PEP can be converted to pyruvate yielding one molecule of ATP by substrate-level phosphorylation. Pyruvate can then be channeled to fatty acid biosynthesis further producing NADH (Lim et al., 2010). It still remains unclear how the C3 compounds cross the two internal membranes. In another apicomplexan, *Toxoplasma gondii*, a pfoTPT homolog (TgAPT1) has been proposed to reside in more than one membrane, providing a solution to the issue of transport (Karnataki et al., 2007).

The cryptomonad *Guillardia theta* harbors a complex plastid which retained a vestigial nucleus (nucleomorph) from the

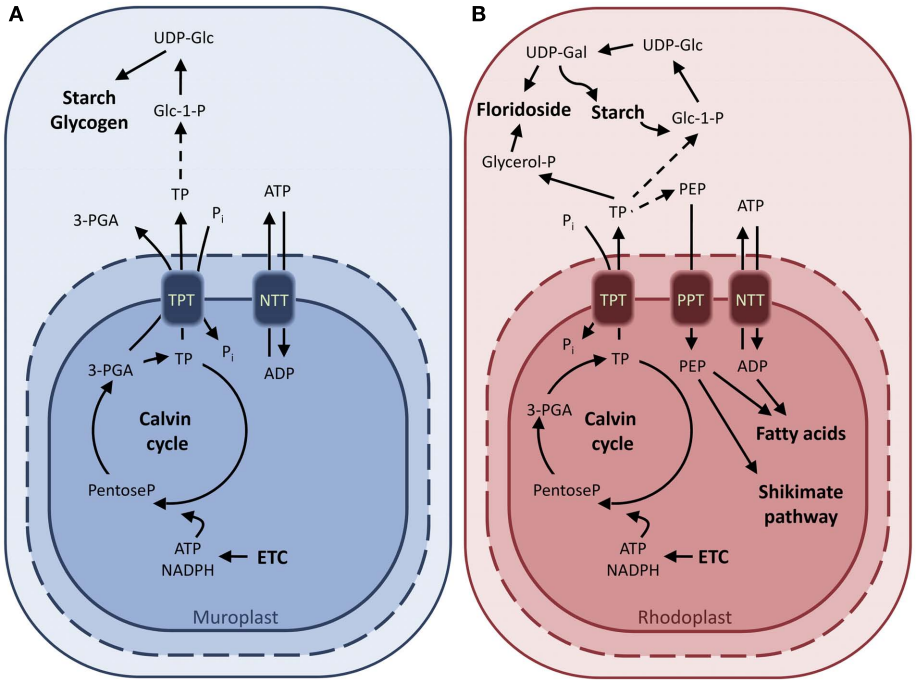


FIGURE 3 | Characterized and putative metabolite transporters of the glaucophaye *Cyanophora paradoxa* (A) and the rhodophyte Galdieria sulphuraria (B). Proposed model of carbon export from *C. paradoxa* based on uptake experiments with isolated cyanelles. The identity of the transporter responsible for the TPT activity is not elucidated yet **(A)**. Characterized metabolite transporters of *G. sulphuraria* and their substrate specificities. The GsTPT, unlike for the green plant chloroplasts, is unable to transport 3-PGA pointing to the presence of an alternate reduction shuttle in the red algae **(B)**.

Abbreviations: ADP-Glc, ADP-glucose; ETC, electron transport chain; Glc 1-P, glucose 1-phosphate; Glc 6-P, glucose 6-phosphate; GPT, glucose 6-phosphate/phosphate translocator; NTT, nucleoside triphosphate transporter; OPPP, oxidative pentose phosphate pathway; 3-PGA, 3-phosphoglyceric acid; PEP, phospho*enol*pyruvate; Pi, inorganic phosphate; PPT, PEP/phosphate translocator; TP, triose phosphate; TPT, triose phosphate/phosphate translocator; UDP-Glc, UDP-glucose; UDP-Gal, UDP-galactose.

engulfed red alga. The nucleomorph resides in the periplastid compartment, the former algal cytoplasm. This intermembrane space is also the site of starch accumulation (Haferkamp et al., 2006). The genome of *G. theta* encodes for two putative TPTs, able to import P_i in counter-exchange to DHAP but not 3-PGA (**Table 1**). Interestingly, the genes encoding these phosphate translocators (TPT1 and TPT2) are regulated in opposite directions. *TPT1* transcripts accumulate during the night phase, whereas *TPT2* accumulates during the light phase (Haferkamp et al., 2006). It has been proposed that that this expression pattern reflects the metabolic involvement of the corresponding carrier proteins, if TPT2 resides in the innermost membrane and favors the export of TPs in the intermembrane space during the day. On the other hand, TPT1 could reside in the first or second outermost membrane and function as exporter of the starch degradation products during the night (**Figure 4C**; Haferkamp et al., 2006).

THE GLUCOSE 6-PHOSPHATE/PHOSPHATE TRANSLOCATOR

In contrast to photosynthetic plastids which are able to synthesize carbon skeletons and reducing equivalents, heterotrophic plastids of sink tissues rely on the import of photosynthates. The phosphorylated C3 sugar phosphates synthesized in the stroma of photosynthetically active chloroplasts are exported to the cytosol through the TPT and converted into sucrose, which is in turn loaded into the phloem and redirected to sink tissues such as tubers, fruits, and seeds. Here sucrose is cleaved by the action of invertase or sucrose synthase and converted into hexose phosphates. The primary hexose phosphate, glucose 6-phosphate, is ultimately imported into the plastid through the (GPT; **Figure 5**).

Glucose 6-phosphate/phosphate translocator is preferentially expressed in non-green tissues and mediates the import of glucose 6-P (Glc6P) as well as TPs and 3-PGA (**Table 1**; Kammerer et al., 1998). Glc6P is further channeled into the starch biosynthetic pathway or into the oxidative pentose phosphate pathway (OPPP), producing reduction equivalents required for diverse anabolic pathways (**Figure 5**; Flügge et al., 2011). The importance of GPT in plant metabolism is demonstrated by the lethal phenotype resulting after disruption of the AtGPT1 homolog of *Arabidopsis thaliana* (Niewiadomski et al., 2005). Mutant plants are severely impaired in the gametogenesis, due to the lack of NADPH produced by the OPPP to sustain fatty acid biosynthesis. Shortage of Glc6P as substrate for the OPPP results in reduced formation of lipid bodies and non-physiological cell death (Niewiadomski et al., 2005). In a recent report the GPT2 ortholog of oil palm (*Elaeis guineensis* Jacq) could be associated with a function in fatty acid biosynthesis and oil accumulation in the mesocarp (Bourgis et al., 2011). Oil palm can accumulate up to 90% oil in its mesocarp. This is achieved through a high flux of glycolytic intermediates

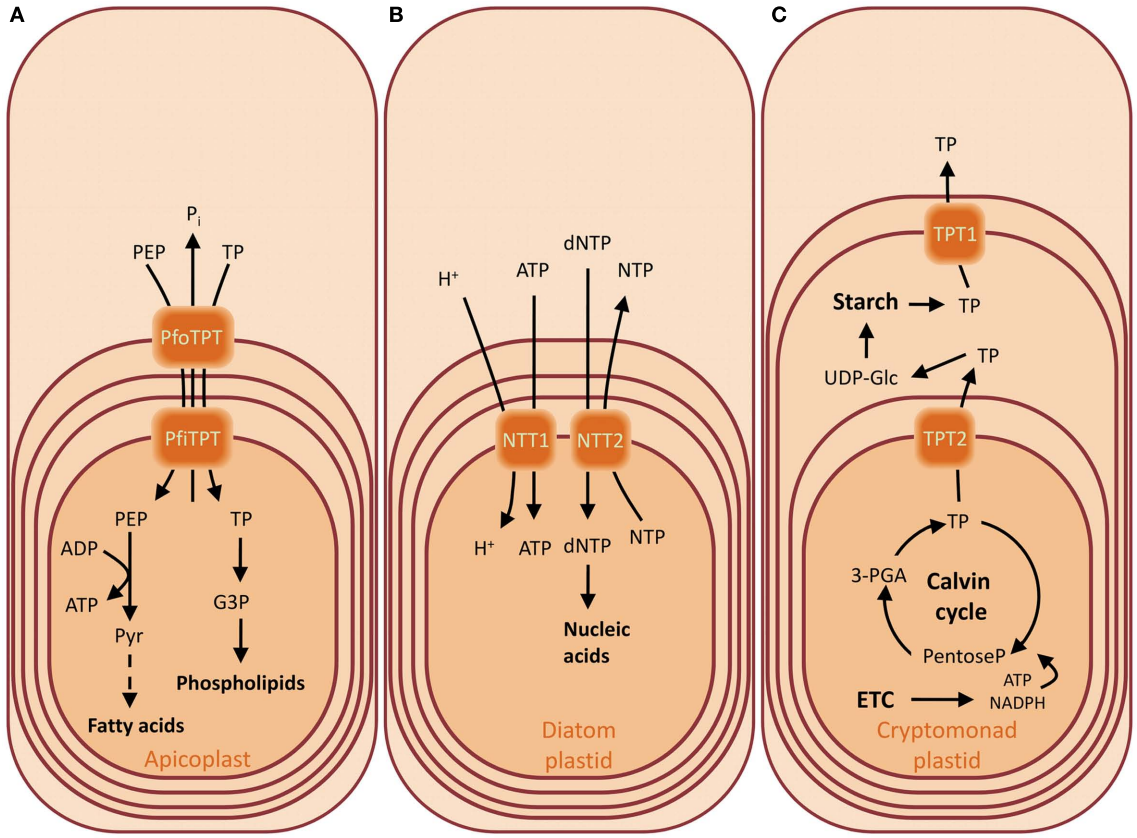


FIGURE 4 | Substrates of the triose phosphate/phosphate translocators (TPTs) and nucleoside triphosphate transporters (NTTs) of the apicoplast of *Plasmodium falciparum* (A), of the diatom plastids from *Thalassiosira pseudonana* and *Phaedactylum tricornutum* (B), and of the crytomonad plastid from *Guillardia theta* (C). Overview of the metabolite transport processes for the characterized transporters of the chromalveolates. The presence of four envelope membranes reflects their secondary endosymbiotic origin. The TPT of the apicoplast of *P. falciparum* has a broader substrate specificity compared to the green counterpart, accepting also PEP **(A)**. Diatom plastids are not able to synthesize nucleotides and therefore they depend on the import of nucleotides from the cytosol. They possess NTTs which catalyze a uniport mode of transport as in the chlamydial intracellular parasites from where these transporters were acquired **(B)**. The cryptomonads contain a less reduced secondary endosymbiont, still harboring a vestigial nucleus in the periplasmic compartment, the former algal cytoplasm, where the starch is synthesized. The day and night path of carbon metabolism are regulated by differential expression of TPT genes whose products localize in different envelopes **(C)**. Abbreviations: (d)NTP, (deoxy)nucleotide triphosphate; G3P, glycerol 3-phosphate; PEP, phospho*enol*pyruvate; PfiTPT, *P. falciparum* innermost envelope TPT; PfoTPT, *P. falciparum* outermost envelope TPT; 3-PGA, 3-phosphoglycerate; Pi, inorganic phosphate; TP, triose phosphate; UDP-Glc, UDP-glucose.

providing pyruvate for oil synthesis. Up-regulation of GPT2 and PPT supplies glycolytic substrates and intermediates to the plastid by transport from the cytosol, funneling carbon compounds toward pyruvate to sustain the high demand of substrates for fatty acid biosynthesis during mesocarp ripening (Bourgis et al., 2011).

Reconstituted *G. sulphuraria* membranes showed high activity of phosphate/phosphate and phosphate/DHAP counter-exchange, whereas 3-PGA, 2-PGA, PEP, and Glc 6-P transport rates did not vary from that of the negative controls (Weber et al., 2004). Indeed, uptake experiments with reconstituted liposomes containing a recombinant protein with significant homology to higher plant GPT demonstrated that Glc6P is not a relevant substrate and suggests that *G. sulphuraria* and likely red algae in general do not possess a plastidic hexose-P importer (Linka et al., 2008a). As mentioned above, GPT supplies heterotrophic plastids of green plants due to the absence of FBPase activity. Since in *G. sulphuraria* the FBPase is not redox regulated, import of TPs mediated by GsTPT could ensure the supply of hexose-P for NADPH production during the night or under heterotrophic growth conditions.

THE XYLULOSE 5-PHOSPHATE/PHOSPHATE TRANSLOCATOR

The XPT of *Arabidopsis* was the last member of plant pPT to be identified and is able to transport inorganic phosphate, TPs, 3-P-glycerate, xylulose 5-P (Xul 5-P), and, to a lesser extent, erythrose 4-P (Ery 4-P) and ribulose 5-P (Ru 5-P; **Table 1**; Eicks et al., 2002). The XPT protein displays high homology to the GPT, and most likely derived from the latter by retrotranscription and genome insertion, as suggested by the lack of introns in the XPT gene (Knappe et al., 2003a). Phylogenetic analysis of the phosphate transporters of the red algae *G. sulphuraria* and *Cyanidioschyzon merolae* showed that members of this family are present which localize within the branch of the GPT/XPT group of transporters (Weber et al., 2006).

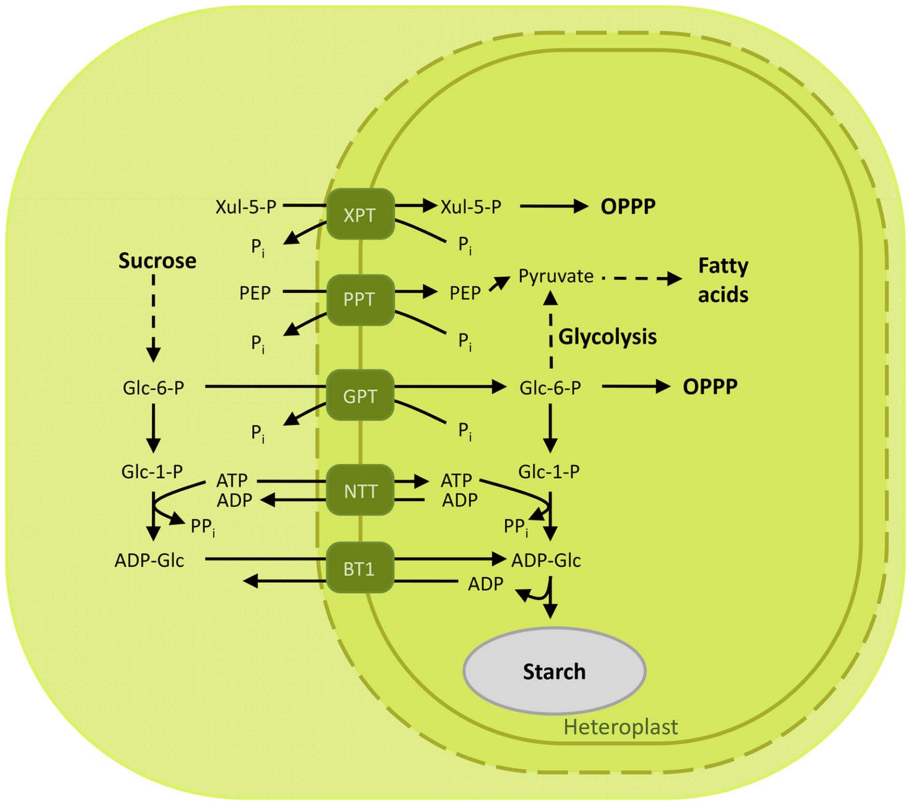


FIGURE 5 | Overview of the characterized metabolite transporters of the heterotrophic plastid. Non-photosynthetic plastids rely on the import of photosynthates to drive the metabolic reactions in the sink tissues. Glucose 6-phosphate is transported by GPT to provide heterotrophic plastids with substrates further used to produce reducing equivalents and storage compounds. In maize, the substrate for starch synthesis, ADP-glucose, is produced in the cytoplasm and shuttled to the seeds by the brittle-1 transporter. Abbreviations: ADP-Glc, ADP-glucose; BT1, Brittle-1 ADP-glucose carrier; Glc 1-P, glucose 1-phosphate; Glc 6-P, glucose 6-phosphate; GPT, glucose 6-phosphate/phosphate translocator; Mal, maltose; NTT, nucleoside triphosphate transporter; OPPP, oxidative pentose phosphate pathway; 3-PGA, 3-phosphoglyceric acid; PEP, phospho*enol*pyruvate; Pi, inorganic phosphate; PPi, pyrophosphate; PPT, PEP/phosphate translocator; XPT, xylulose 5-phosphate/phosphate translocator; Xul 5-P, xylulose 5-phosphate.

The proposed function of XPT is to import pentose phosphates into the stroma in the form of Xyl 5-P to be further metabolized in the Calvin cycle and the OPPP (**Figures 2** and **5**). A further function of the XPT would be the supply of carbon skeletons intermediates withdrawn from the Calvin cycle and the OPPP and channeled into other biosynthetic pathways such as the synthesis of nucleotides and the shikimate pathway requiring Rib 5-P and Ery 4-P, respectively (Eicks et al., 2002).

THE PHOSPHO*ENOL*PYRUVATE/PHOSPHATE TRANSLOCATOR

The phospho*enol*pyruvate/phosphate translocator (PPT) mediates the import of cytosolic PEP into the plastid stroma in exchange with P_i (**Table 1**; Fischer et al., 1997).

Since the chloroplasts and most plastids of non-green tissues lack the enzymes required to form PEP from TPs *via* the glycolytic pathway, PEP has to be imported from the cytosol (Stitt and Aprees, 1979; Borchert et al., 1993). The energy-rich compound PEP serves as substrate for the shikimate pathway, whose starting point is the condensation of one molecule of PEP and one of Ery 4-P (**Figure 2**). The shikimate pathway is exclusively localized within the plastid and is involved in the biosynthesis of aromatic amino acids and of a variety of secondary compounds such as lignin, alkaloids, and flavonoids (Herrmann, 1995; Herrmann and Weaver, 1999). PEP can also be metabolized to pyruvate and acetyl-CoA which serves as precursor for the biosynthesis of fatty acids, as already mentioned for the GPT (**Figures 2** and **4**; Rajasekharan and Nachiappan, 2010; Bourgis et al., 2011).

Arabidopsis contains two *PPT* genes (*PPT1* and *PPT2*) which differ in their expression pattern (Knappe et al., 2003b). *PPT1* is expressed in the leaf vasculature and in some root tissues. The abolishment of its activity in the *cue1* mutant causes a reticulate leaf phenotype with abnormally developed mesophyll cells (MC) (Streatfield et al., 1999). The expression pattern of *PPT1* and the observation that the biosynthesis of aromatic amino acids is not severely affected in the *cue1* mutant led to the suggestion that it might be involved in the development of the MCs trough a not yet identified signaling molecule likely derived from the shikimate pathway. AtPPT1 is presumably involved in the synthesis of secondary metabolites that regulate leaf development, whereas AtPPT2 which is ubiquitously expressed in leaves could have a more housekeeping function in supplying the plastidic shikimate pathway with PEP (Knappe et al., 2003b; Voll et al., 2003).

Another way to provide plastids with PEP is from plastidic pyruvate by pyruvate, orthophosphate dikinase (PPDK). This reaction is an essential step in C_4- and CAM-plants where PEP is exported to the cytosol of MCs and serves as substrate for the PEP-carboxylase during CO_2 fixation (Bräutigam et al., 2008). In C_4 and CAM-plants PPT is therefore an exporter of PEP (Weber and Von Caemmerer, 2010).

In the red alga *G. sulphuraria*, preliminary uptake experiments with isolated total membranes did not show pronounced PPT activity (Weber et al., 2004). However, phylogenetic analysis revealed the presence of a candidate ortholog in the red algal genome, suggesting the presence of a potential coding capability for PPT (Weber et al., 2006). Indeed, the recombinant protein was functionally characterized in reconstituted liposome membranes and displayed the signature of PPT activity (Linka et al., 2008a). As its green counterpart, the *G. sulphuraria* PPT is highly specific for P_i and PEP (**Table 1**).

PEP synthesis inside the rhodoplast is assumed to be unlikely since the isozymes of both routes for synthesizing PEP in the plastid (i.e., the glycolytic pathway through phosphoglyceromutase and enolase, and the conversion of pyruvate to PEP by the PPDK) do not possess obvious subcellular targeting sequences and are predicted to be cytosolic (Linka et al., 2008a). Since fatty acid biosynthesis and the shikimate pathway in red algae are predicted to be similar to those of the green lineage of Archaeplastida, it is assumed that conversion of TP to PEP is not possible in *G. sulphuraria*, thus depending on PEP import from the cytosol to drive PEP-depended reactions in the stroma (**Figure 3B**; Weber et al., 2004; Richards et al., 2006; Linka et al., 2008a).

THE GLUTATHIONE TRANSPORTER

Recently, a new transporter member of the DMT superfamily has been characterized as glutathione transporter in *Arabidopsis*. In plants, glutathione is in the center of an antioxidant network that buffers the reactive oxygen species (ROS) originating as by-product of the aerobic energy metabolism. Glutathione is a tripeptide enzymatically synthesized by the glutamate–cysteine ligase (GSH1) and the glutathione synthase (GSH2). In *Arabidopsis*, GSH1 is plastid localized, while GSH2 is targeted to both plastid and cytosol (Meister, 1995; Wachter et al., 2005). Consequently, the pathway intermediate, γ-glutamylcysteine (γ-EC), must be exported from the plastid to allow for cytosolic GSH biosynthesis (**Figure 2**). In a mutant screening designed to select mutants resistant to L-buthionine-SR-sulfoximine (BSO), an inhibitor of GSH1, the gene At5g19380, encoding a predicted membrane transport protein, was identified as responsible for the revertant phenotype (Maughan et al., 2010). The gene was found to be related to the chloroquine resistance transporter (PfCRT) of the malaria parasite *P. falciparum* and was consequently designated CRT-like transporter1 (CLT1). *Arabidopsis* mutants lacking CLT1 are heavy metal-sensitive, GSH-deficient, and hypersensitive to *Phytophthora* infection. CLT1 is a member of a small gene family comprising three members in *Arabidopsis* and all localize to the plastid envelope. Expression of the transporter in *Xenopus* oocytes confirmed that it can mediate GSH uptake (Maughan et al., 2010). Homology search revealed related genes in higher plants, in the moss *Physcomitrella patens*, in green algae and in parasitic protozoans, suggesting an ancient ancestry for this transporter.

THE ATP:ADP ANTIPORTER FAMILY

Members of the ATP:ADP Antiporter (AAA) Family (TC 2.A.12), also called nucleoside triphosphate transporters (NTT) catalyze a counter-exchange mode of transport and may be distantly related to members of the major facilitator superfamily (MFS; TC 2.A.1). They are not related to the mitochondrial ATP/ADP exchangers of the mitochondrial carrier family (MCF; TC 2.A.5), which pump ATP out of mitochondria in accordance with the polarity of the mitochondrial membrane potential. In addition, the plant mitochondrial adenylate translocator is highly sensitive to specific inhibitors, whereas the plastidic ATP/ADP translocator is only slightly sensitive to them (Vignais et al., 1976; Ardila et al., 1993; Schunemann et al., 1993; Winkler and Neuhaus, 1999).

In prokaryotes these transporters enable "energy parasitism," as in *Rickettsia* and *Chlamydia* (Neuhaus et al., 1997). In addition, putative nucleotide transport proteins with weak sequence similarities to the NTT family of transporters have been identified in the nuclear genome of the obligate intracellular parasite *Encephalitozoon cuniculi* (Katinka et al., 2001). The bacterial nucleotide carriers have conspicuous sequence similarity to the NTT proteins from chloroplasts of higher plants (Neuhaus et al., 1997).

In plants, the main function proposed for the NTT family of transporters is to supply ATP-dependent reactions in non-photosynthetic plastids. Amyloplasts are heterotrophic plastids of sink tissues and contain various anabolic pathways, like the biosyntheses of starch and fatty acids or nitrogen assimilation. During starch biosynthesis in potato tubers, glucose 1-phosphate (Glc1P) is converted to ADP-Glc by the AGPase, a reaction which is ATP-dependent and relies on the supply of cytosolic ATP by NTT (**Figure 5**). This is corroborated by transgenic plants where repression of the potato NTT leads to decreased tuber yield and starch contents (Tjaden et al., 1998). In contrast to heterotrophic plastids, the role of NTTs in chloroplasts is not obvious. Rather than exporting ATP out of the photosynthetically active plastid during the day, NTT import ATP during the night or when photosynthetic ATP biosynthesis is not sufficient to fuel anabolic reactions requiring ATP (Reinhold et al., 2007).

Arabidopsis knock-out mutants demonstrated the involvement of NTT and ATP import in the assembly of magnesium chelatase, an enzyme involved in chlorophyll biosynthesis (Kobayashi et al., 2008). Plants repressing NTTs display necrotic lesions caused by photo-oxidation due to an accumulation of high amounts of the phototoxic chlorophyll biosynthesis intermediate protoporphyrin IX, which induces ROS production and photooxidative damage. Furthermore, the severity of the stress depended on the size of the plastidic starch pool, with plants grown in short day conditions being more affected (Reinhold et al., 2007). It has been proposed that the sensitivity to photooxidative stress in plants with limited plastidic ATP supply may link the appearance of the light harvesting complexes in the green lineage with the return of the storage polysaccharide to the plastid (Deschamps et al., 2008b). In plants, the main pathway of starch breakdown does not allow for the direct generation of

ATP within the plastid. However, hexose phosphates can be generated in the stroma by plastidial starch phosphorylase and further metabolized within the plastid to generate the required ATP pools, thereby circumventing protoporphyrin IX-induced oxidative stress (Deschamps et al., 2008b; Rathore et al., 2009). The authors propose that a sufficient supply of ATP was a potentially important protective innovation necessary for "safe" proliferation of LHCs.

The glaucophyte *C. paradoxa* also possesses a NTT, although the activity for ATP/ADP translocation could not be demonstrated in isolated muroplasts (Schlichting et al., 1990; Linka et al., 2003). In *C. paradoxa*, both fatty acids and starch biosynthesis are cytosol-localized, indicating that ATP import into the muroplast is not required for these reactions (**Figure 3A**; Ma et al., 2001; Plancke et al., 2008). Functional expression and reconstitution into membranes of the glaucophyte NTT ortholog will be required to elucidate its physiological role and demonstrate whether the protein is able to mediate ATP/ADP exchange.

The red algae *C. merolae* and *G. sulphuraria* contain each one gene coding for NTT (Tyra et al., 2007). In *Galdieria*, similarly to the plant NTT, the ortholog transporter is able to exchange ATP for ADP (Linka et al., 2003). In red algae the synthesis of storage carbon is confined in the cytoplasmic compartment, however, other anabolic reactions such as fatty acid biosynthesis require energy supply (Viola et al., 2001). The synthesis of fatty acids in *G. sulphuraria* has been postulated to occur in the plastid, as indicated by the presence of the subunits of the plastidic acetyl-coA carboxylase, an ATP-dependent enzyme (Weber et al., 2004). As mentioned before, the occurrence of a glycolytic pathway in the plastids of red algae is unlikely and ATP import would be required to drive fatty acid biosynthesis (**Figure 3B**; Linka et al., 2008a).

In contrast to higher plants, the secondary derived plastids of the diatoms are not able to synthesize nucleotides. Plants export newly synthesized nucleotides through a member of the MCF, an adenine nucleotide uniporter (Leroch et al., 2005). In diatoms, the synthesis of nucleotides is confined to the cytosol, implying the existence of an import mechanism. *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* possess diverse members of the NTTs mediating the import of purine and pyrimidine nucleotides (Ast et al., 2009). In these organisms the NTTs are therefore involved in the nucleotide supply to maintain organellar metabolism (such as nucleic acids synthesis) rather than energy provision. Diatoms possess therefore plastidial NTT proteins with biochemical properties similar to those of some bacterial intracellular parasites, which catalyze nucleotide transport in a unidirectional mode, a feature not observed for any eukaryotic NTT (**Figure 4B**; Tjaden et al., 1999). This role in providing the plastid with externally synthesized nucleotides is similar to that of metabolically impaired intracellular bacteria such as *Chlamydia* which depend on host-derived metabolites. The functional relatedness of *Chlamydia* and plant NTTs reflects their common origin. The prokaryote transporters with the highest sequence similarity to plant NTT proteins occur in *Chlamydia* relatives and phylogenetic analyses of plant, *Chlamydiae* and *Rickettsiae* NTTs suggest that the translocator was transferred from *Chlamydia* to plants (Schmitz-Esser et al., 2004). The *Arabidopsis* AtNTT1 and AtNTT2 ATP/ADP translocases were shown to be monophyletic with the chlamydial translocators (Tyra et al., 2007). Moreover, free-living cyanobacteria are devoid of genes encoding for NTTs. Although *Chlamydiae* are not found in plants, an unexpected number of chlamydial genes display a high degree of similarity to their plant homologs. An analysis of bacteria–eukaryote protein similarities revealed that, beyond the well explainable abundance in cyanobacterial and α-proteobacterial sequences, a notable number of plant-like genes were found in the Chlamydiaceae genomes and, most intriguingly, the vast majority of these genes contain a plastid-targeting signal (Brinkman et al., 2002; Horn et al., 2004). A phylogenetic analysis of the *C. merolae* genome aimed at searching for genes that are evolutionary related to *Chlamydia* retrieved 21 genes which are broadly present in primary photosynthetic eukaryotes. The direction of transfer is proposed to be from the bacteria to the plants, since the genes are predominantly distributed in bacteria and the cyanobacterial homologs form a well supported group distinct from the chlamydial homologs (Huang and Gogarten, 2007). Genes of chlamydial origin are mainly restricted to photosynthetic eukaryotes and non-photosynthetic plastid-bearing lineages, an observation that supports an association between *Chlamydiae* and the plant ancestor. A deeper phylogenomic analysis of 17 plant genomes extended the list to 55 proteins of chlamydial origin, two-thirds of which are of plastid function (Moustafa et al., 2008). The scenario evolving from these comparative genomics surveys suggests a three-way partnership between the host, the cyanobacterial endosymbiont and an environmental *Chlamydia*. The chlamydial contribution may have consisted in facilitating the establishment of the cyanobacterial endosymbiont by providing the necessary transport systems required for the equilibration of their metabolism, like the ADP/ATP translocator needed for the energy flux into the organelle. Later on, when the organelle was fully established in the host cell, the bacterial parasite became dispensable and was eventually lost.

Chlamydia could have thus given a contribution in the early steps of endosymbiosis by providing the petroalgae with the genetic tools necessary to establish a favorable interaction.

MITOCHONDRIAL CARRIER FAMILY

Members of the MC family (TC 2.A.29) catalyze the passage of hydrophobic compounds across the inner mitochondrial membrane and preferentially catalyze the exchange of one solute acting as antiporters (Haferkamp, 2007; Palmieri et al., 2011). Members of the MCF are involved in transporting keto acids, amino acids, nucleotides, inorganic ions, and co-factors across the mitochondrial inner membrane. Although most are found in mitochondria, recently members of the MCF were found in peroxisomes, in hydrogenosomes of anaerobic fungi, and chloroplasts (Picault et al., 2004; Haferkamp, 2007; Linka et al., 2008b). In *Arabidopsis*, 58 genes encode for MCF proteins (Picault et al., 2004; Palmieri et al., 2011). The number in other sequenced plant genomes varies from 37 to 125, thus being larger than that of *Saccharomyces cerevisiae*, which encodes for 35 MCs and comparable with that of humans which contain 53 members (Palmieri et al., 1996; Palmieri, 2004).

Members of the MCF located in the plastid membrane represent transporters which are of host origin and were likely relocated to the envelope to substitute the pre-existing cyanobacterial transporters with the endomembrane-derived host transporters (Tyra et al., 2007). No homologs of mitochondrial carriers have ever been found in prokaryotes or archaea, meaning that MCFs are a later addition during the development of the eukaryotic cell (Kunji, 2004). In plants plastids, they mediate the transport of different substrates like folates, *S*-adenosylmethionine, NAD, ADP-glucose or ATP, ADP, and AMP (Haferkamp, 2007).

THE ADP-GLUCOSE TRANSPORTER BRITTLE-1

Besides the abovementioned NTT transporter, plastids contain another type of nucleotide transporter, the ADP-glucose transporter Brittle-1 (BT1), first identified in maize endosperm (Sullivan et al., 1991). The *brittle1* mutant was named after the reduced starch content in the endosperm, resulting in collapsed, brittle kernels (Mangelsdorf and Jones, 1926). In the cereals endosperm, ADP-glucose pyrophosphorylase is localized in the cytosol (Denyer et al., 1996). The enzyme generates ADP-glucose, which must be imported into the plastid to drive starch biosynthesis. BT1 is localized to the inner envelope membrane of the amyloplast and transports ADP-glucose in antiport with the ADP released during starch biosynthesis, ensuring the balance of adenylates between stroma and cytoplasm (**Figure 5**; Sullivan and Kaneko, 1995; Kirchberger et al., 2007).

Another maize Brittle transporter protein, Brittle-1-2, and homologs of these transporters are found in other plants (Kirchberger et al., 2007). Interestingly, two types of Brittle-1 transporters can be identified: the BT1 from maize endosperm is present only in monocotyledonous plants and catalyzes ADP-glucose-ADP antiport, whereas the BT1-2 isoform has homologs also in dicotyledonous plants and transports nucleotides but not ADP-glucose (Leroch et al., 2005; Kirchberger et al., 2008). The StBT1 from potato and the AtBT1 from *Arabidopsis* locate to the inner plastid envelope and transport AMP, ADP, and ATP. Moreover, they show a unidirectional mode of transport rather than antiport (**Figure 2**; Leroch et al., 2005; Kirchberger et al., 2008).

TRANSPORTERS INVOLVED IN THE ONE-CARBON METABOLISM

Several cellular processes such as the synthesis of amino acids, purines, secondary metabolites, and the photorespiratory pathway depend upon the supply or removal of one-carbon units by tetrahydrofolate (THF) or *S*-adenosylmethionine (SAM).

Tetrahydrofolate and its derivatives are a family of co-factors that are essential for all cellular one-carbon transfer reactions: they are required for the synthesis of purines, thymidylate, methionine, pantothenate, and the interconversion of serine and glycine (Hanson and Roje, 2001). Folates are tripartite molecules composed of a pterin, a *p*-aminobenzoic acid (pABA), and a glutamate chain with a variable number of glutamate moieties.

In plants, folate synthesis presents a complex spatial organization, involving three subcellular compartments: the cytosol, the plastid, and the mitochondria.

In *Arabidopsis*, one member of the MCF was identified as folate carrier of the plastid envelope (Bedhomme et al., 2005). The AtFOLT1 protein is the closest homolog of the human mitochondrial folate transporter. Functional complementation of a glycine auxotroph Chinese hamster ovary (CHO) cell line deficient for the transport of folates into mitochondria could identify the folate carriers both in human and *Arabidopsis* (Titus and Moran, 2000; Bedhomme et al., 2005). The expression of AtFOLT1 in *E. coli* confers to the bacteria the ability to uptake exogenous folate, whereas the disruption of the gene did not led to a deficient folate supply to the plastid, since wild-type levels of folates were detected in the chloroplasts and the enzymatic capacity to catalyze folate-dependent reactions was preserved (Bedhomme et al., 2005). This suggests that alternative routes for folate uptake into plastids are present.

SAM serves as methyl-group donor and it is exclusively synthesized in the plant cytosol. Consequently, it has to be imported into plastids and mitochondria where it acts as substrate for the methyltransferases. The product resulting from the transmethylation reaction, *S*-adenosylhomocysteine (SAHC) has to be in turn exported to undergo re-methylation in the cytosol. The carrier that catalyzes the counter exchange of SAM for SAHC has been recently identified in *Arabidopsis* (Bouvier et al., 2006; Palmieri et al., 2006). The SAM transporter (SAMT) exhibits a broad expression in various plant tissues, is localized to the plastid envelope membrane and, interestingly, the complete sequence contains additional information for mitochondrial targeting (Palmieri et al., 2006).

OTHER TRANSPORTERS

THE DICARBOXYLATE TRANSLOCATORS

For plants, nitrogen is available in the form of nitrate, which is sequentially reduced to nitrite in the cytosol and finally to ammonia in the chloroplast stroma. Ammonia is further assimilated into nitrogen compounds via the plastidic glutamine synthetase/glutamate synthetase (GS/GOGAT) cycle. Glutamate serves as a universal organic nitrogen donor for the biosynthesis of a vast variety nitrogen-containing compounds. Assimilation of ammonia in the plastids involves two distinct dicarboxylate translocators (DiTs) with partially overlapping substrate specificities, the 2-oxoglutarate/malate-translocator (DiT1), and the glutamate/malate-translocator (DiT2). DiT1 imports the precursor of ammonia assimilation, 2-oxoglutarate, in exchange with malate; DiT2 exports the end product of ammonia assimilation, glutamate, to the cytosol, importing back the malate (**Figure 2**; Weber et al., 1995; Weber and Flügge, 2002; Renne et al., 2003). These transporters play also an important role in recycling the ammonia lost during photorespiration by the mitochondrial glycine decarboxylase, as indicated by the photorespiratory phenotype displayed by *Arabidopsis* plants lacking DiT2 and tobacco DiT1 antisense plants (Renne et al., 2003; Linka and Weber, 2005; Schneidereit et al., 2006). DiT1 also displays a high affinity for oxaloacetate (OAA) and has been proposed as the long-sought OAA/malate valve (Taniguchi et al., 2002). The malate valve functions as redox sink to dissipate excess of electrons and protect the photosynthetic apparatus from photoinhibition by exporting reducing equivalents thus balancing the stromal ATP/NADPH ratio (Heineke et al., 1991; Scheibe, 2004; Scheibe et al., 2005). The role as malate valve for DiT1 was recently proven by mutant

analysis: CO_2-dependent O_2 evolution assays showed that cytosolic oxaloacetate is efficiently transported into chloroplasts mainly by DiT1 and O_2 evolution is not enhanced by addition of OAA in the mutant, supporting the absence of additional oxaloacetate transporters (Kinoshita et al., 2011).

Dicarboxylate translocators belong to the Divalent Anion:Na^+ Symporter (DASS) Family (TC 2.A.47) which transport organic di- and tri-carboxylates of the Krebs Cycle as well as dicarboxylate amino acid, inorganic sulfate, and phosphate. These proteins are found in Gram-negative bacteria, cyanobacteria, archaea, plant chloroplasts, yeast, and animals. DiTs are present only in the green lineage of plants and in bacteria, but not in red algae or glaucophytes. Phylogenetic surveys indicate that the dicarboxylate transporters are of chlamydial origin, reinforcing the hypothesis on the importance these intracellular parasites had in the establishment of endosymbiosis (Tyra et al., 2007). In isolated *C. paradoxa* muroplasts, transport activities for glutamine and 2-oxoglutarate could be measured (Kloos et al., 1993). In uptake experiments with isolated cyanelles by silicone oil filtering it could be shown that glutamate poorly penetrated into cyanelles, whereas glutamine was enriched in the plastid fraction by 1.7 fold within 10 min. Glutamine uptake proceeded in two phases and was stimulated by 2-oxoglutarate (Kloos et al., 1993). Ammonia formed by nitrite reduction inside the cyanelles is incorporated by GS into glutamine and then exported jointly with oxoglutarate (Kloos et al., 1993). The export of glutamine appears to be the route for providing the host cell with fixed nitrogen, a situation that encounters a parallel in the nitrogen-fixing filamentous cyanobacteria. In cyanobacteria however, the transporter seems not to discriminate between glutamine and glutamate, indicating that the glutamine export route is unique to *C. paradoxa* (Chapman and Meeks, 1983; Flores and Muropastor, 1988; Kloos et al., 1993).

THE MALTOSE TRANSPORTER

In the dark, the starch which has accumulated during the day is degraded to supply sucrose biosynthesis. Maltose and glucose are the main products of starch degradation, with maltose being the most prominent sugar exported to the cytosol at night or during increased sink demand (**Figure 2**; Weise et al., 2004). The maltose transporter was identified by a screening of *Arabidopsis* mutants displaying a maltose excess (MEX) phenotype (Niittyla et al., 2004). MEX is unrelated to any other membrane transporter and belongs to a family represented solely by the *A. thaliana* member (TC 2.A.84). The maltose exporter is found only in the genomes of chloroplastida and is absent in red algae, glaucophytes, and photosynthetic organisms derived by secondary endosymbiosis. The presence of MEX orthologs in the dinoflagellate *Karlodinium micrum* and in *Synechococcus* likely is the result of independent horizontal gene transfers (**Figure 6**; Tyra et al., 2007). The pathway of starch synthesis and degradation is presumed to have been cytosolic in the common ancestor of the three Archaeplastida lineages (Deschamps et al., 2008a). MEX appears therefore to be a plant-specific transporter, evolved as a consequence of the relocation of the starch metabolism to the stroma of the green algal ancestor (Tyra et al., 2007; Deschamps et al., 2008a).

THE PLASTIDIC GLUCOSE TRANSPORTER

Starch breakdown initiates by the action of the phosphorylating enzymes glucan, water dikinase (GWD) and phosphoglucan, water dikinase (PDW) and, after involvement of a debranching enzyme, results in the release linear glucans (Smith et al., 2005). β-amylases catalyze the production of maltose from linear glucans, and, to a lesser extent, maltotriose which is further metabolized to glucose by the disproportionating enzyme DPE1 (Critchley et al., 2001; Smith et al., 2005). A transporter involved in carbon mobilization mediating glucose export has been initially characterized and two decades later cloned and identified (Schafer et al., 1977; Weber et al., 2000). The glucose transporter (pGlcT) mediates the export of glucose released during the onset of starch degradation and has additional and still unknown roles suggested by its expression in some heterotrophic tissues that do not contain starch (Butowt et al., 2003). pGlcT contributes to the export of photoassimilates from chloroplasts, even though it is not the major route for the export of starch degradation products (**Figure 2**). The *pglct-1/mex1* double knock-out *Arabidopsis* line displays a severe growth retardation phenotype suggesting that the effects of pGlcT loss on photoassimilate partitioning and plant growth are observable when the major route for exporting starch degradation products is blocked (Hahn et al., 2011). pGlcT belongs to the (MFS, TC 2.A.1) a very old, large and diverse superfamily whose members catalyze uniport, solute:cation symport and/or solute:H^+ or solute:solute antiport (Law et al., 2008). This superfamily is ubiquitous in all kingdoms of life, however, close relatives of pGlcT are found only within the Chloroplastida and, as for the maltose exporter, it reflects an acquisition by the green lineage imposed by the compartmentalization of the starch metabolism within the chloroplast (Tyra et al., 2007).

THE PLASTIDIC BILE ACID:Na SYMPORTER

Glucosinolates (GSLs) are sulfur- and nitrogen-containing plant secondary metabolites derived from amino acids and sugar. They are mainly present in the Brassicaceae family, including *Arabidopsis*, and have repellent activity against herbivore insects and pathogens (Fahey et al., 2001).

The biosynthesis of methionine-derived glucosinolates involves both the chloroplast and the cytosolic compartments and starts in the cytosol where Met is transaminated by the aminotransferase BCAT4. The resulting 2-keto acid, 4-methylthio-2-oxobutanoate (MTOB) is imported into the chloroplast where it undergoes repeated condensations with acetyl-CoA and subsequent decarboxylations to the corresponding α-keto acid. The chain-elongated keto acid is aminated to the amino acid and transported back to the cytosol for the following core structure synthesis (Fahey et al., 2001; Knill et al., 2008; Nour-Eldin and Halkier, 2009). Thus, the methionine derivatives must be transported at least twice across the chloroplast membranes. The first transporter involved in the biosynthesis of Met-derived glucosinolates was recently identified as member of the Bile Acid:Na Symporter (BASS) Family (TC 2.A.28) due to significant sequence similarities to mammalian sodium-coupled bile acid transporters (Gigolashvili et al., 2009; Sawada et al., 2009). The BAT5 protein localizes to the plastid and loss-of-function mutants defective in BAT5 function contained strongly reduced levels of aliphatic glucosinolates. Feeding

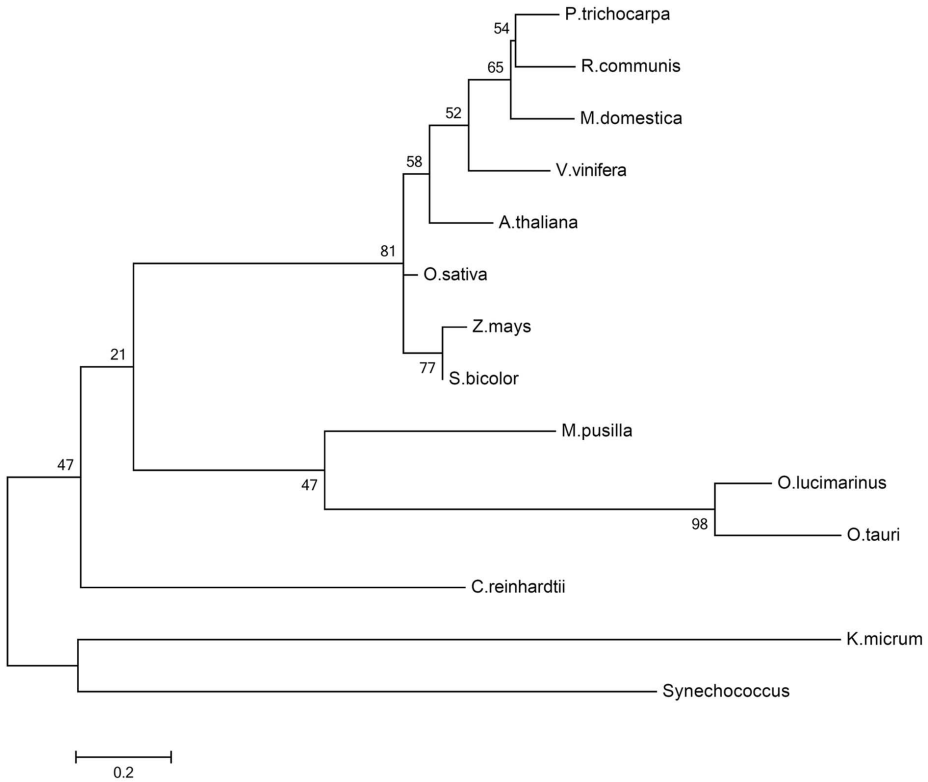


FIGURE 6 | Molecular phylogenetic analysis of the MEX1 proteins. The evolutionary history was inferred by using the maximum likelihood method based on the WAG model (Whelan and Goldman, 2001). Sequence alignment was performed with MUSCLE included in MEGA5 (Tamura et al., 2011). Alignment quality was assessed by GUIDANCE giving an alignment score of 0.783. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. A discrete gamma distribution was used to model evolutionary rate differences among sites (five categories, $+G$, parameter = 3.3094). The rate variation model allowed for some sites to be evolutionarily invariable ([$+I$], 5.4532% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 14 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 173 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).

experiments suggest that BAT5 transports MTOB and chain-elongated 2-keto acids across the chloroplast envelope membrane, indicating a role for BAT5 in the transport of GS intermediates (Gigolashvili et al., 2009; Sawada et al., 2009).

Homologs of the sodium-coupled bile acid transporters are also found in the algal genomes, however no data about their localization or function are to date available.

THE PLASTIDIC FOLATE CARRIER

In addition to the plastidic folate carrier belonging to the MCF, a second folate carrier was identified which belongs to the folate-biopterin transporter (FBT) Family (TC 2.A.71). The FBT family includes functionally characterized members from protozoa and homologs are found in cyanobacteria (*Synechocystis*) and plants (*A. thaliana*). The *Synechocystis* and *Arabidopsis* FBTs were identified based on their homology to the folate transporter of the trypanosomatid protozoa *Leishmania* (Klaus et al., 2005). These transporters are able to complement *E. coli* mutants impaired in folate uptake or production; in particular they are capable of transporting the monoglutamyl form of folate (Klaus et al., 2005). The plant FBT likely originated by endosymbiotic gene transfer (EGT) from the captured cyanobacterium and contributes to a redundant folate uptake system involving also the abovementioned host-derived FOLT1 transporter (Tyra et al., 2007).

THE PYRUVATE CARRIER

Plants that perform C_4 photosynthesis require a massive flux of metabolites across the membranes of their MC and bundle sheat cells (BSC). In particular, pyruvate is formed in the chloroplasts of the BSC by decarboxylation of the C_4 acid malate and has to be transported back to the MC chloroplasts to regenerate the CO_2 acceptor PEP (Bräutigam and Weber, 2011). Pyruvate transport in C_4 plants has been shown to occur either in a proton- or sodium-dependent manner (Ohnishi and Kanai, 1987, 1990). Accordingly, a putative plastidic proton:sodium symporter (NHD1) is 16-fold up-regulated in the C_4 species *Cleome gynandra* with respect to the C_3 relative *Cleome spinosa* (Bräutigam et al., 2011). It is proposed that NDH1 functions in exporting sodium out of the chloroplast to sustain a gradient necessary for pyruvate import. In addition, a strong up-regulation of a bile acid:sodium cotransporter was observed in *C. gyandra* (Bräutigam et al., 2011) and in C_4 Flaveria species (Gowik et al., 2011). This cotransporter is a member

of the BASS Family (TC 2.A.28) and is not found in maize, a species which displays a proton-dependent transport of pyruvate (Bräutigam et al., 2008). The biochemical function NDH1 and BASS2 was recently demonstrated by expression in an heterologous system of *BASS2* alone and in presence of *NDH1* (Furumoto et al., 2011). Indeed, these proteins represent a two-translocator system, consisting of the sodium:pyruvate cotransporter BASS2 and the sodium:proton antiporter NHD1 (**Figure 7**; Furumoto et al., 2011).

CONCLUSION

The transport of organic compounds across the cyanobacterial membrane was of crucial importance for establishing endosymbiosis and insertion of protein translocators into the envelope was necessary to take full advantage of the newly acquired metabolic entity. The pPTs were most likely pioneers in settling a connection between the host and the cyanobiont in order to tap the photosynthates, as suggested by their ubiquitous distribution among the plastid-bearing organisms. The majority of the plastid translocators, including the PTs, are of host origin, indicating that establishment of the plastid was a host-driven process. However, evidence is accumulating which implies a small but substantial contribution from the intracellular parasite *Chlamydia*. The requirement of three organisms in the early steps of endosymbiosis could explain why this event occurred only once during evolution. Understanding the metabolic connection between plastid and cytosol is also crucial for future attempts to generate synthetic organelles, such as synthetic plastids that harbor tailor-made metabolic capacities and that are introduced into plastid-free recipient cells (Weber and Osteryoung, 2010).

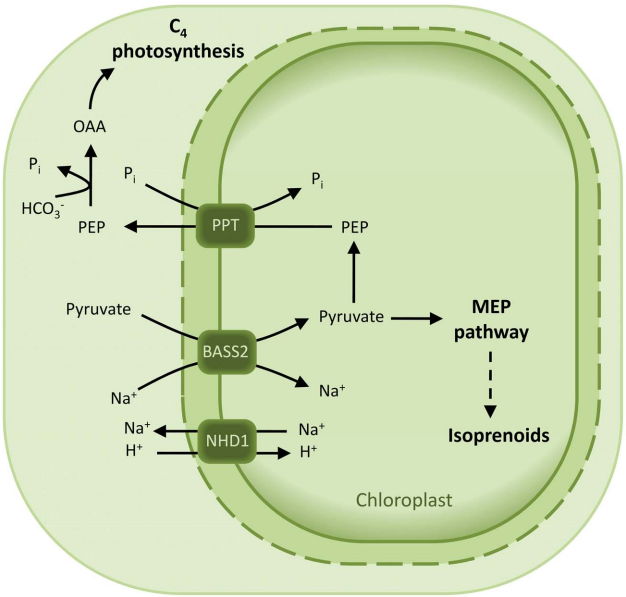


FIGURE 7 | Model for the pyruvate uptake into the plastid. The two-translocator system responsible for pyruvate uptake into the plastid of certain C_4 species is based on a sodium:pyruvate cotransporter (BASS2) and a sodium:proton antiporter (NHD1). NHD1 is able to establish a sodium gradient across the envelope membrane, which in turn drives pyruvate import by the cotransporter BASS2. Abbreviations: BASS2, BILE ACID:SODIUM SYMPORTER FAMILY PROTEIN 2; NHD1, sodium:proton antiporter; OAA, oxaloacetate; PEP, phospho*enol*pyruvate; PPT, PEP/phosphate translocator.

REFERENCES

Adl, S. M., Simpson, A. G., Farmer, M. A., Andersen, R. A., Anderson, O. R., Barta, J. R., Bowser, S. S., Brugerolle, G., Fensome, R. A., Fredericq, S., James, T. Y., Karpov, S., Kugrens, P., Krug, J., Lane, C. E., Lewis, L. A., Lodge, J., Lynn, D. H., Mann, D. G., Mccourt, R. M., Mendoza, L., Moestrup, O., Mozley-Standridge, S. E., Nerad, T. A., Shearer, C. A., Smirnov, A. V., Spiegel, F. W., and Taylor, M. F. (2005). The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* 52, 399–451.

Allen, J. F., and Martin, W. (2007). Evolutionary biology: out of thin air. *Nature* 445, 610–612.

Archibald, J. M. (2009). The puzzle of plastid evolution. *Curr. Biol.* 19, R81–R88.

Ardila, F., Pozuetaromero, J., and Akazawa, T. (1993). Adenylate uptake by proplastids from cultured-cells of tobacco (nicotiana-tabacum-L Cv-By2) indicates that an adenylate translocator is present in all types of plastid. *Plant Cell Physiol.* 34, 237–242.

Ast, M., Gruber, A., Schmitz-Esser, S., Neuhaus, H. E., Kroth, P. G., Horn, M., and Haferkamp, I. (2009). Diatom plastids depend on nucleotide import from the cytosol. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3621–3626.

Ball, S. G., and Morell, M. K. (2003). From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. *Annu. Rev. Plant Biol.* 54, 207–233.

Barbier, G., Oesterhelt, C., Larson, M. D., Halgren, R. G., Wilkerson, C., Garavito, R. M., Benning, C., and Weber, A. P. (2005). Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol.* 137, 460–474.

Bedhomme, M., Hoffmann, M., Mccarthy, E. A., Gambonnet, B., Moran, R. G., Rebeille, F., and Ravanel, S. (2005). Folate metabolism in plants – an *Arabidopsis* homolog of the mammalian mitochondrial folate transporter mediates folate import into chloroplasts. *J. Biol. Chem.* 280, 34823–34831.

Björn, L. O. (2009). The evolution of photosynthesis and chloroplasts. *Curr. Sci.* 96, 1466–1474.

Bölter, B., and Soll, J. (2001). Ion channels in the outer membranes of chloroplasts and mitochondria: open doors or regulated gates? *EMBO J.* 20, 935–940.

Borchert, S., Harborth, J., Schunemann, D., Hoferichter, P., and Heldt, H. W. (1993). Studies of the enzymic capacities and transport properties of Pea Root Plastids. *Plant Physiol.* 101, 303–312.

Bourgis, F., Kilaru, A., Cao, X., Ngando-Ebongue, G. F., Drira, N., Ohlrogge, J. B., and Arondel, V. (2011). Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12527–12532.

Bouvier, F., Linka, N., Isner, J. C., Mutterer, J., Weber, A. P. M., and Camara, B. (2006). *Arabidopsis* SAMT1 defines a plastid transporter regulating plastid biogenesis and plant development. *Plant Cell* 18, 3088–3105.

Bräutigam, A., Hoffmann-Benning, S., and Weber, A. P. (2008). Comparative proteomics of chloroplast envelopes from C_3 and C_4 plants reveals specific adaptations of the plastid envelope to C_4 photosynthesis and candidate proteins required for maintaining C_4 metabolite fluxes. *Plant Physiol.* 148, 568–579.

Bräutigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K. L., Carr, K. M., Gowik, U., Mass, J., Lercher, M. J., Westhoff, P., Hibberd, J. M., and Weber, A. P. (2011). An mRNA blueprint for C_4 photosynthesis derived from comparative transcriptomics of closely related C_3 and C_4 species. *Plant Physiol.* 155, 142–156.

Bräutigam, A., and Weber, A. P. M. (2011). "Transport processes: connecting the reactions of C_4 photosynthesis," in *C_4 Photosynthesis and Related CO_2 Concentrating Mechanisms*, eds A. S. Raghavendra and R. F. Sage (Berlin: Springer), 199–219.

Brinkman, F. S., Blanchard, J. L., Cherkasov, A., Av-Gay, Y., Brunham, R. C., Fernandez, R. C., Finlay, B. B., Otto, S. P., Ouellette, B. F., Keeling, P. J., Rose, A. M., Hancock, R. E., Jones, S. J., and Greberg, H. (2002). Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. *Genome Res.* 12, 1159–1167.

Burey, S. C., Poroyko, V., Ergen, Z. N., Fathi-Nejad, S., Schüller, C., Ohnishi, N., Fukuzawa, H., Bohnert, H. J., and Löffelhardt, W. (2007). Acclimation to low [CO_2] by an inorganic carbon-concentrating mechanism in *Cyanophora paradoxa*. *Plant Cell Environ.* 30, 1422–1435.

Butowt, R., Granot, D., and Rodriguez-Garcia, M. I. (2003). A putative plastidic glucose translocator is expressed in heterotrophic tissues that do not contain starch, during olive (Olea europea L.) fruit ripening. *Plant Cell Physiol.* 44, 1152–1161.

Cavalier-Smith, T. (2000). Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* 5, 174–182.

Chapman, J. S., and Meeks, J. C. (1983). Glutamine and glutamate transport by Anabaena-variabilis. *J. Bacteriol.* 156, 122–129.

Colleoni, C., Linka, M., Deschamps, P., Handford, M. G., Dupree, P., Weber, A. P., and Ball, S. G. (2010). Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. *Mol. Biol. Evol.* 27, 2691–2701.

Critchley, J. H., Zeeman, S. C., Takaha, T., Smith, A. M., and Smith, S. M. (2001). A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in *Arabidopsis*. *Plant J.* 26, 89–100.

Denyer, K., Dunlap, F., Thorbjornsen, T., Keeling, P., and Smith, A. M. (1996). The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial. *Plant Physiol.* 112, 779–785.

Deschamps, P., Haferkamp, I., D'Hulst, C., Neuhaus, H. E., and Ball, S. G. (2008a). The relocation of starch metabolism to chloroplasts: when, why and how. *Trends Plant Sci.* 13, 574–582.

Deschamps, P., Moreau, H., Worden, A. Z., Dauvillee, D., and Ball, S. G. (2008b). Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* 178, 2373–2387.

Duy, D., Soll, J., and Philippar, K. (2007). Solute channels of the outer membrane: from bacteria to chloroplasts. *Biol. Chem.* 388,879–889.

Eicks, M., Maurino, V., Knappe, S., Flügge, U.-I., and Fischer, K. (2002). The plastidic pentose phosphate translocator represents a link between the cytosolic and the plastidic pentose phosphate pathways in plants. *Plant Physiol.* 128, 512–522.

Embley, T. M., and Martin, W. (2006). Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630.

Fahey, J. W., Zalcmann, A. T., and Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51.

Fischer, K., Kammerer, B., Gutensohn, M., Arbinger, B., Weber, A., Häusler, R. E., and Flügge, U. I. (1997). A new class of plastidic phosphate translocators: a putative link between primary and secondary metabolism by the phosphoenolpyruvate/phosphate antiporter. *Plant Cell* 9, 453–462.

Fliege, R., Flügge, U. I., Werdan, K., and Heldt, H. W. (1978). Specific transport of inorganic phosphate, 3-phosphoglycerate and triosephosphates across the inner membrane of the envelope in spinach chloroplasts. *Biochim. Biophys. Acta* 502, 232–247.

Flores, E., and Muropastor, M. I. (1988). Uptake of glutamine and glutamate by the dinitrogen-fixing cyanobacterium anabaena Sp Pcc7120. *FEMS Microbiol. Lett.* 56, 127–130.

Flügge, U. I. (1995). Phosphate translocation in the regulation of photosynthesis. *J. Exp. Bot.* 46, 1317–1323.

Flügge, U. I. (2000). Transport in and out of plastids: does the outer envelope membrane control the flow? *Trends Plant Sci.* 5, 135–137.

Flügge, U. I., and Heldt, H. W. (1984). The phosphate-triose phosphate-phosphoglycerate translocator of the chloroplast. *Trends Biochem. Sci.* 9, 530–533.

Flügge, U.-I., Häusler, R. E., Ludewig, F., and Gierth, M. (2011). The role of transporters in supplying energy to plant plastids. *J. Exp. Bot.* 62, 2381–2392.

Furumoto, T., Yamaguchi, T., Ohshima-Ichie, Y., Nakamura, M., Tsuchida-Iwata, Y., Shimamura, M., Ohnishi, J., Hata, S., Gowik, U., Westhoff, P., Bräutigam, A., Weber, A. P. M, and Izui, K. (2011). A plastidial sodium-dependent pyruvate transporter. *Nature* 476, 472–475.

Gigolashvili, T., Yatusevich, R., Rollwitz, I., Humphry, M., Gershenzon, J., and Flügge, U. I. (2009). The plastidic bile acid transporter 5 is required for the biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*. *Plant Cell* 21, 1813–1829.

Gould, S. B., Waller, R. F., and Mcfadden, G. I. (2008). Plastid evolution. *Annu. Rev. Plant Biol.* 59, 491–517.

Gowik, U., Bräutigam, A., Weber, K. L., Weber, A. P. M., and Westhoff, P. (2011). Evolution of C4 photosynthesis in the genus Flaveria - how many and which genes does it take to make C4? *Plant Cell* 23, 2087–2105.

Gross, W., and Schnarrenberger, C. (1995). Heterotrophic growth of two strains of the acido-thermophilic red alga *Galdieria sulphuraria*. *Plant Cell Physiol.* 36, 633.

Haferkamp, I. (2007). The diverse members of the mitochondrial carrier family in plants. *FEBS Lett.* 581, 2375–2379.

Haferkamp, I., Deschamps, P., Ast, M., Jeblick, W., Maier, U., Ball, S., and Neuhaus, H. E. (2006). Molecular and biochemical analysis of periplastidial starch metabolism in the cryptophyte *Guillardia theta*. *Eukaryot. Cell* 5, 964–971.

Hahn, T. R., Cho, M. H., Lim, H., Shin, D. H., Jeon, J. S., Bhoo, S. H., and Park, Y. I. (2011). Role of the plastidic glucose translocator in the export of starch degradation products from the chloroplasts in *Arabidopsis thaliana*. *New Phytol.* 190, 101–112.

Hanson, A. D., and Roje, S. (2001). One-carbon metabolism in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 119–137.

Heineke, D., Riens, B., Grosse, H., Hoferichter, P., Peter, U., Flügge, U. I., and Heldt, H. W. (1991). Redox transfer across the inner chloroplast envelope membrane. *Plant Physiol.* 95, 1131–1137.

Herrmann, K. M. (1995). The Shikimate pathway: early steps in the biosynthesis of aromatic compounds. *Plant Cell* 7, 907–919.

Herrmann, K. M., and Weaver, L. M. (1999). The Shikimate pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 473–503.

Hohmann-Marriott, M. F., and Blankenship, R. E. (2011). Evolution of photosynthesis. *Annu. Rev. Plant Biol.* 62, 515–548.

Horn, M., Collingro, A., Schmitz-Esser, S., Beier, C. L., Purkhold, U., Fartmann, B., Brandt, P., Nyakatura, G. J., Droege, M., Frishman, D., Rattei, T., Mewes, H. W., and Wagner, M. (2004). Illuminating the evolutionary history of *Chlamydiae*. *Science* 304, 728–730.

Huang, J. L., and Gogarten, J. P. (2007). Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol.* 8, R99.

Inoue, K. (2011). Emerging roles of the chloroplast outer envelope membrane. *Trends Plant Sci.* doi: 10.1016/j.tplants.2011.06.005

Jarvis, P., Dörmann, P., Peto, C. A., Lutes, J., Benning, C., and Chory, J. (2000). Galactolipid deficiency and abnormal chloroplast development in the *Arabidopsis* MGD synthase 1 mutant. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8175–8179.

Kammerer, B., Fischer, K., Hilpert, B., Schubert, S., Gutensohn, M., Weber, A., and Flügge, U. I. (1998). Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate/phosphate antiporter. *Plant Cell* 10, 105–117.

Karnataki, A., Derocher, A., Coppens, I., Nash, C., Feagin, J. E., and Parsons, M. (2007). Cell cycle-regulated vesicular trafficking of Toxoplasma APT1, a protein localized to multiple apicoplast membranes. *Mol. Microbiol.* 63, 1653–1668.

Katinka, M. D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., and Vivares, C. P. (2001). Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414, 450–453.

Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 729–748.

Kies, L. (1974). Electron microscopical investigations on *Paulinella chromatophora* Lauterborn, a thecamoeba containing blue-green endosymbiionts (cyanelles). *Protoplasma* 80, 69–89.

Kinoshita, H., Nagasaki, J., Yoshikawa, N., Yamamoto, A., Takito, S., Kawasaki, M., Sugiyama, T., Miyake, H., Weber, A. P. M., and Taniguchi, M. (2011). The chloroplastic 2-oxoglutarate/malate transporter has dual function as the malate valve and in carbon/nitrogen metabolism. *Plant J.* 65, 15–26.

Kirchberger, S., Leroch, M., Huynen, M. A., Wahl, M., Neuhaus, H. E., and Tjaden, J. (2007). Molecular and biochemical analysis of the plastidic ADP-glucose transporter (ZmBT1) from Zea mays. *J. Biol. Chem.* 282, 22481–22491

Kirchberger, S., Tjaden, J., and Neuhaus, H. E. (2008). Characterization of the *Arabidopsis* Brittle1 transport protein and impact of reduced activity on plant metabolism. *Plant J.* 56, 51–63.

Klaus, S. M., Kunji, E. R., Bozzo, G. G., Noiriel, A., De La Garza, R. D., Basset, G. J., Ravanel, S., Rebeille, F., Gregory, J. F., and Hanson, A. D. (2005). Higher plant plastids and cyanobacteria have folate carriers related to those of trypanosomatids. *J. Biol. Chem.* 280, 38457–38463.

Kloos, K., Schlichting, R., Zimmer, W., and Bothe, H. (1993). Glutamine and glutamate transport in cyanophora-paradoxa. *Bot. Acta* 106, 435–440.

Knappe, S., Flügge, U. I., and Fischer, K. (2003a). Analysis of the plastidic phosphate translocator gene family in *Arabidopsis* and identification of new phosphate translocator-homologous transporters, classified by their putative substrate-binding site. *Plant Physiol.* 131, 1178–1190.

Knappe, S., Lottgert, T., Schneider, A., Voll, L., Flügge, U. I., and Fischer, K. (2003b). Characterization of two functional phosphoenolpyruvate/phosphate translocator (PPT) genes in *Arabidopsis*-AtPPT1 may be involved in the provision of signals for correct mesophyll development. *Plant J.* 36, 411–420.

Knill, T., Schuster, J., Reichelt, M., Gershenzon, J., and Binder, S. (2008). *Arabidopsis* branched-chain aminotransferase 3 functions in both amino acid and glucosinolate biosynthesis. *Plant Physiol.* 146, 1028–1039.

Kobayashi, K., Mochizuki, N., Yoshimura, N., Motohashi, K., Hisabori, T., and Masuda, T. (2008). Functional analysis of *Arabidopsis thaliana* isoforms of the Mg-chelatase CHLI subunit. *Photochem. Photobiol. Sci.* 7, 1188–1195.

Kunji, E. R. (2004). The role and structure of mitochondrial carriers. *FEBS Lett.* 564, 239–244.

Law, C. J., Maloney, P. C., and Wang, D. N. (2008). Ins and outs of major facilitator superfamily, antiporters. *Annu. Rev. Microbiol.* 62, 289–305.

Leroch, M., Kirchberger, S., Haferkamp, I., Wahl, M., Neuhaus, H. E., and Tjaden, J. (2005). Identification and characterization of a novel plastidic adenine nucleotide uniporter from Solanum tuberosum. *J. Biol. Chem.* 280, 17992–18000.

Lim, L., Linka, M., Mullin, K. A., Weber, A. P., and Mcfadden, G. I. (2010). The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite. *FEBS Lett.* 584, 549–554.

Linka, M., Jamai, A., and Weber, A. P. M. (2008a). Functional characterization of the plastidic phosphate translocator gene family from the thermo-acidophilic red alga *Galdieria sulphuraria* reveals specific adaptations of primary carbon partitioning in green plants and red algae. *Plant Physiol.* 148, 1487–1496.

Linka, N., Theodoulou, F. L., Haslam, R. P., Linka, M., Napier, J. A., Neuhaus, H. E., and Weber, A. P. M. (2008b). Peroxisomal ATP import is essential for seedling development in *Arabidopsis thaliana*. *Plant Cell* 20, 3241–3257.

Linka, M., and Weber, A. P. (2005). Shuffling ammonia between mitochondria and plastids during photorespiration. *Trends Plant Sci.* 10, 461–465.

Linka, N., Hurka, H., Lang, B. F., Burger, G., Winkler, H. H., Stamme, C., Urbany, C., Seil, I., Kusch, J., and Neuhaus, H. E. (2003). Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eukaryotes. *Gene* 306, 27–35.

Löffelhardt, W., Bohnert, H. J., and Bryant, D. A. (1997). The cyanelles of *Cyanophora paradoxa*. *CRC Crit. Rev. Plant Sci.* 16, 393–413.

Ma, Y., Jakowitsch, J., Maier, T. L., Bayer, M. G., Muller, N. E., Schenk, H. E., and Löffelhardt, W. (2001). ATP citrate lyase in the glaucocystophyte alga *Cyanophora paradoxa* is a cytosolic enzyme: characterisation of the gene for the large subunit at the cDNA and genomic levels. *Mol. Genet. Genomics* 266, 231–238.

Mangelsdorf, P. C., and Jones, D. F. (1926). The expression of mendelian factors in the gametophyte of maize. *Genetics* 11, 423–455.

Martin, W., and Herrmann, R. G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* 118, 9–17.

Maughan, S. C., Pasternak, M., Cairns, N., Kiddle, G., Brach, T., Jarvis, R., Haas, F., Nieuwland, J., Lim, B., Muller, C., Salcedo-Sora, E., Kruse, C., Orsel, M., Hell, R., Miller, A. J., Bray, P., Foyer, C. H., Murray, J. A., Meyer, A. J., and Cobbett, C. S. (2010). Plant homologs of the *Plasmodium falciparum* chloroquine-resistance transporter, PfCRT, are required for glutathione homeostasis and stress responses. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2331–2336.

Meister, A. (1995). Glutathione biosynthesis and its inhibition. *Biothiols Pt B* 252, 26–30.

Moustafa, A., Reyes-Prieto, A., and Bhattacharya, D. (2008). *Chlamydiae* has contributed at least 55 genes to plantae with predominantly plastid functions. *PLoS ONE* 3, e2205. doi: 10.1371/journal.pone.0002205

Mulkidjanian, A. Y., Koonin, E. V., Makarova, K. S., Mekhedov, S. L., Sorokin, A., Wolf, Y. I., Dufresne, A., Partensky, F., Burd, H., Kaznadzey, D., Haselkorn, R., and Galperin, M. Y. (2006). The cyanobacterial genome core and the origin of photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13126–13131.

Mullin, K. A., Lim, L., Ralph, S. A., Spurck, T. P., Handman, E., and Mcfadden, G. I. (2006). Membrane transporters in the relict plastid of malaria parasites. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9572–9577.

Neuhaus, H. E., Thom, E., Mohlmann, T., Steup, M., and Kampfenkel, K. (1997). Characterization of a novel eukaryotic ATP/ADP translocator located in the plastid envelope of *Arabidopsis thaliana* L. *Plant J.* 11, 73–82.

Niewiadomski, P., Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U. S., Rosso, M. G., Ache, P., Flügge, U. I., and Schneider, A. (2005). The *Arabidopsis* plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. *Plant Cell* 17, 760–775.

Niittyla, T., Messerli, G., Trevisan, M., Chen, J., Smith, A. M., and Zeeman, S. C. (2004). A previously unknown maltose transporter essential for starch degradation in leaves. *Science* 303, 87–89.

Nour-Eldin, H. H., and Halkier, B. A. (2009). Piecing together the transport pathway of aliphatic glucosinolates. *Phytochem. Rev.* 8, 53–67.

Oesterhelt, C., Klocke, S., Holtgrefe, S., Linke, V., Weber, A. P. M., and Scheibe, R. (2007). Redox regulation of chloroplast enzymes in *Galdieria sulphuraria* in view of eukaryotic evolution. *Plant Cell Physiol.* 48, 1359–1373.

Ohnishi, J., and Kanai, R. (1990). Pyruvate uptake induced by a Ph jump in mesophyll chloroplasts of maize and sorghum, Nadp-malic enzyme type-C_4 species. *FEBS Lett.* 269, 122–124.

Ohnishi, J. I., and Kanai, R. (1987). Na+-induced uptake of pyruvate into mesophyll chloroplasts of a C_4 plant, *Panicum miliaceum*. *FEBS Lett.* 219, 347–350.

Palmieri, F. (2004). The mitochondrial transporter family (SLC25): physiological and pathological implications. *Pflugers Arch.* 447, 689–709.

Palmieri, F., Palmieri, L., Arrigoni, R., Blanco, E., Carrari, F., Zanor, M. I., Studart-Guimaraes, C., and Fernie, A. R. (2006). Molecular identification of an *Arabidopsis* S-adenosylmethionine transporter. Analysis of organ distribution, bacterial expression, reconstitution into liposomes, and functional characterization. *Plant Physiol.* 142, 855–865.

Palmieri, F., Pierri, C. L., De Grassi, A., Nunes-Nesi, A., and Fernie, A. R. (2011). Evolution, structure and function of mitochondrial carriers: a review with new insights. *Plant J.* 66, 161–181.

Palmieri, L., Palmieri, F., Runswick, M. J., and Walker, J. E. (1996). Identification by bacterial expression and functional reconstitution of the yeast genomic sequence encoding the mitochondrial dicarboxylate carrier protein. *FEBS Lett.* 399, 299–302.

Patron, N. J., and Keeling, P. J. (2005). Common evolutionary origin of starch biosynthetic enzymes in green and red algae. *J. Phycol.* 41, 1131–1141.

Picault, N., Hodges, M., Paimieri, L., and Palmieri, F. (2004). The growing family of mitochondrial carriers in *Arabidopsis*. *Trends Plant Sci.* 9, 138–146.

Plancke, C., Colleoni, C., Deschamps, P., Dauvillee, D., Nakamura, Y., Haebel, S., Ritte, G., Steup, M., Buleon, A., Putaux, J. L., Dupeyre, D., D'Hulst, C., Ral, J. P., Löffelhardt, W., and Ball, S. G. (2008). Pathway of cytosolic starch synthesis in the model glaucophyte *Cyanophora paradoxa*. *Eukaryot. Cell* 7, 247–257.

Rajasekharan, R., and Nachiappan, V. (2010). Fatty acid biosynthesis and regulation in plants. *Plant Dev. Biol. Biotechnol. Perspect.* 105–115.

Ralph, S. A., Van Dooren, G. G., Waller, R. F., Crawford, M. J., Fraunholz, M. J., Foth, B. J., Tonkin, C. J., Roos, D. S., and Mcfadden, G. I. (2004). Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat. Rev. Microbiol.* 2, 203–216.

Rathore, R. S., Garg, N., Garg, S., and Kumar, A. (2009). Starch phosphorylase: role in starch metabolism and biotechnological applications. *Crit. Rev. Biotechnol.* 29, 214–224.

Reichert, A., Dennes, A., Vetter, S., and Scheibe, R. (2003). Chloroplast fructose 1,6-bisphosphatase with changed redox modulation: comparison of the *Galdieria* enzyme with cysteine mutants from spinach. *Biochim. Biophys. Acta* 1645, 212–217.

Reinhold, T., Alawady, A., Grimm, B., Beran, K. C., Jahns, P., Conrath, U., Bauer, J., Reiser, J., Melzer, M., Jeblick, W., and Neuhaus, H. E. (2007). Limitation of nocturnal import of ATP into *Arabidopsis* chloroplasts leads to photooxidative damage. *Plant J.* 50, 293–304.

Renne, P., Dressen, U., Hebbeker, U., Hille, D., Flügge, U. I., Westhoff, P., and Weber, A. P. M. (2003). The *Arabidopsis* mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. *Plant J.* 35, 316–331.

Reyes-Prieto, A., Weber, A. P. M., and Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. *Annu. Rev. Genet.* 41, 147–168.

Richards, T. A., Dacks, J. B., Campbell, S. A., Blanchard, J. L., Foster, P. G., Mcleod, R., and Roberts, C. W. (2006). Evolutionary origins of the eukaryotic shikimate pathway: gene fusions, horizontal gene transfer, and endosymbiotic replacements. *Eukaryot. Cell* 5, 1517–1531.

Saier, M. H., Tran, C. V., and Barabote, R. D. (2006). TCDB: the transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Res.* 34, D181–D186.

Saier, M. H., Yen, M. R., Noto, K., Tamang, D. G., and Elkan, C. (2009). The transporter classification database: recent advances. *Nucleic Acids Res.* 37, D274–D278.

Sato, M., Mogi, Y., Nishikawa, T., Miyamura, S., Nagumo, T., and Kawano, S. (2009). The dynamic surface of dividing cyanelles and ultrastructure of the region directly below the surface in *Cyanophora paradoxa*. *Planta* 229, 781–791.

Sawada, Y., Toyooka, K., Kuwahara, A., Sakata, A., Nagano, M., Saito, K., and Hirai, M. Y. (2009). *Arabidopsis* bile acid:sodium symporter family protein 5 is involved in methionine-derived glucosinolate biosynthesis. *Plant Cell Physiol.* 50, 1579–1586.

Schafer, G., Heber, U., and Heldt, H. W. (1977). Glucose transport into spinach chloroplasts. *Plant Physiol.* 60, 286.

Scheibe, R. (2004). Malate valves to balance cellular energy supply. *Physiol. Plant* 120, 21–26.

Scheibe, R., Backhausen, J. E., Emmerlich, V., and Holtgrefe, S. (2005). Strategies to maintain redox homeostasis during photosynthesis under changing conditions. *J. Exp. Bot.* 56, 1481–1489.

Schleiff, E., Eichacker, L. A., Eckart, K., Becker, T., Mirus, O., Stahl, T., and Soll, J. (2003). Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. *Protein Sci.* 12, 748–759.

Schlichting, R., and Bothe, H. (1993). The cyanelles (organelles of a low evolutionary scale) possess a phosphate-translocator and a glucose-carrier in cyanophora-paradoxa. *Bot. Acta* 106, 428–434.

Schlichting, R., Zimmer, W., and Bothe, H. (1990). Exchange of metabolites in cyanophora-paradoxa and its cyanelles. *Bot. Acta* 103, 392–398.

Schmitz-Esser, S., Linka, N., Collingro, A., Beier, C. L., Neuhaus, H. E., Wagner, M., and Horn, M. (2004). ATP/ADP translocases: a common feature of obligate intracellular amoebal symbionts related to *Chlamydiae* and *Rickettsiae*. *J. Bacteriol.* 186, 683–691.

Schneidereit, J., Häusler, R. E., Fiene, G., Kaiser, W. M., and Weber, A. P. M. (2006). Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. *Plant J.* 45, 206–224.

Schulz, B., Frommer, W. B., Flügge, U. I., Hummel, S., Fischer, K., and Willmitzer, L. (1993). Expression of the triose phosphate translocator gene from potato is light-dependent and restricted to green tissues. *Mol. Gen. Genet.* 238, 357–361.

Schunemann, D., Borchert, S., Flügge, U. I., and Heldt, H. W. (1993). Adp/Atp translocator from pea root plastids – comparison with translocators from spinach-chloroplasts and pea leaf mitochondria. *Plant Physiol.* 103, 131–137.

Smith, A. M., Zeeman, S. C., and Smith, S. M. (2005). Starch degradation. *Annu. Rev. Plant Biol.* 56, 73–98.

Soll, J., Bölter, B., Wagner, R., and Hinnah, S. C. (2000). . . .response: the chloroplast outer envelope: a molecular sieve? *Trends Plant Sci.* 5, 137–138.

Steiner, J. M., and Löffelhardt, W. (2002). Protein import into cyanelles. *Trends Plant Sci.* 7, 72–77.

Stitt, M., and Aprees, T. (1979). Capacities of pea-chloroplasts to catalyze the oxidative pentose-phosphate pathway and glycolysis. *Phytochemistry* 18, 1905–1911.

Streatfield, S. J., Weber, A., Kinsman, E. A., Häusler, R. E., Li, J. M., Post-Beittenmiller, D., Kaiser, W. M., Pyke, K. A., Flügge, U. I., and Chory, J. (1999). The phosphoenolpyruvate/phosphate translocator is required for phenolic metabolism, palisade cell development, and plastid-dependent nuclear gene expression. *Plant Cell* 11, 1609–1621.

Sullivan, T. D., and Kaneko, Y. (1995). The maize brittle1 gene encodes amyloplast membrane polypeptides. *Planta* 196, 477–484.

Sullivan, T. D., Strelow, L. I., Illingworth, C. A., Phillips, R. L., and Nelson, O. E. (1991). Analysis of maize brittle-1 alleles and a defective suppressor-mutator-induced mutable allele. *Plant Cell* 3, 1337–1348.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* doi: 10.1093/molbev/msr121

Taniguchi, M., Taniguchi, Y., Kawasaki, M., Takeda, S., Kato, T., Sato, S., Tahata, S., Miyake, H., and Sugiyama, T. (2002). Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana*. *Plant Cell Physiol.* 43, 706–717.

Titus, S. A., and Moran, R. G. (2000). Retrovirally mediated complementation of the glyB phenotype – cloning of a human gene encoding the carrier for entry of folates into mitochondria. *J. Biol. Chem.* 275, 36811–36817.

Tjaden, J., Mohlmann, T., Kampfenkel, K., Henrichs, G., and Neuhaus, H. E. (1998). Altered plastidic ATP/ADP-transporter activity influences potato (*Solanum tuberosum* L.) tuber morphology, yield and composition of tuber starch. *Plant J.* 16, 531–540.

Tjaden, J., Winkler, H. H., Schwoppe, C., Van Der Laan, M., Mohlmann, T., and Neuhaus, H. E. (1999). Two nucleotide transport proteins in *Chlamydia trachomatis*, one for net nucleoside triphosphate uptake and the other for transport of energy. *J. Bacteriol.* 181, 1196–1202.

Tyra, H. M., Linka, M., Weber, A. P. M., and Bhattacharya, D. (2007). Host origin of plastid solute transporters in the first photosynthetic eukaryotes. *Genome Biol.* 8, R212.

Vignais, P. V., Douce, R., Lauquin, G. J., and Vignais, P. M. (1976). Binding of radioactively labeled carboxyatractyloside, atractyloside and bongkrekic acid to Adp translocator of potato mitochondria. *Biochim. Biophys. Acta* 440, 688–696.

Viola, R., Nyvall, P., and Pedersen, M. (2001). The unique features of starch metabolism in red algae. *Proc. Biol. Sci.* 268, 1417–1422.

Voll, L., Häusler, R. E., Hecker, R., Weber, A., Weissenbock, G., Fiene, G., Waffenschmidt, S., and Flügge, U. I. (2003). The phenotype of the *Arabidopsis* cue1 mutant is not simply caused by a general restriction of the shikimate pathway. *Plant J.* 36, 301–317.

Wachter, A., Wolf, S., Steininger, H., Bogs, J., and Rausch, T. (2005). Differential targeting of GSH1 and GSH2 is achieved by multiple transcription initiation: implications for the compartmentation of glutathione biosynthesis in the Brassicaceae. *Plant J.* 41, 15–30.

Weber, A., and Flügge, U.-I. (2002). Interaction of cytosolic and plastidic nitrogen metabolism in plants. *J. Exp. Bot.* 53, 865–874.

Weber, A., Menzlaff, E., Arbinger, B., Gutensohn, M., Eckerskorn, C., and Flügge, U. I. (1995). The 2-oxoglutarate malate translocator of chloroplast envelope membranes – molecular-cloning of a transporter containing a 12-helix motif and expression of the functional protein in yeast-cells. *Biochemistry* 34, 2621–2627.

Weber, A., Servaites, J. C., Geiger, D. R., Kofler, H., Hille, D., Groner, F., Hebbeker, U., and Flügge, U. I. (2000). Identification, purification, and molecular cloning of a putative plastidic glucose translocator. *Plant Cell* 12, 787–801.

Weber, A. P. M., and Fischer, K. (2007). Making the connections-the crucial role of metabolite transporters at the interface between chloroplast and cytosol. *FEBS Lett.* 581, 2215–2222.

Weber, A. P. M., Linka, M., and Bhattacharya, D. (2006). Single, ancient origin of a plastid metabolite translocator family in plantae from an endomembrane-derived ancestor. *Eukaryot. Cell* 5, 609–612.

Weber, A. P. M., and Linka, N. (2011). Connecting the plastid: transporters of the plastid envelope and their role in linking plastidial with cytosolic metabolism. *Annu. Rev. Plant Biol.* 62, 53–77.

Weber, A. P. M., Oesterhelt, C., Gross, W., Bräutigam, A., Imboden, L. A., Krassovskaya, I., Linka, N., Truchina, J., Schneidereit, J., Voll, H., Voll, L. M., Zimmermann, M., Jamai, A., Riekhof, W. R., Yu, B., Garavito, R. M., and Benning, C. (2004). EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. *Plant Mol. Biol.* 55, 17–32.

Weber, A. P. M., and Osteryoung, K. W. (2010). From endosymbiosis to synthetic photosynthetic life. *Plant Physiol.* 154, 593–597.

Weber, A. P. M., Schwacke, R., and Flügge, U.-I. (2005). Solute transporters of the plastid envelope membrane. *Annu. Rev. Plant Biol.* 56, 133–164.

Weber, A. P. M., and Von Caemmerer, S. (2010). Plastid transport and metabolism of C_3 and C_4 plants – comparative analysis and possible biotechnological exploitation. *Curr. Opin. Plant Biol.* 13, 257–265.

Weise, S. E., Weber, A. P. M., and Sharkey, T. D. (2004). Maltose is the major form of carbon exported from the chloroplast at night. *Planta* 218, 474–482.

Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18, 691–699.

Winkler, H. H., and Neuhaus, H. E. (1999). Non-mitochondrial ATP transport. *Trends Biochem. Sci.* 24, 64–68.

Wolfe, G. R., Cunningham, F. X., Durnford, D., Green, B. R., and Gantt, E. (1994). Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. *Nature* 367, 566–568.

Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M., and Bauer, C. E. (2000). Molecular evidence for the early evolution of photosynthesis. *Science* 289, 1724–1730.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 July 2011; accepted: 23 August 2011; published online: 12 September 2011.

*Citation: Facchinelli F and Weber APM (2011) The metabolite transporters of the plastid envelope: an update. Front. Plant Sci. **2**:50. doi: 10.3389/fpls.2011.00050*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 09 January 2012
doi: 10.3389/fpls.2011.00110

Phylogenetic analysis of the thylakoid ATP/ADP carrier reveals new insights into its function restricted to green plants

Cornelia Spetea[1]*, Bernard E. Pfeil[1] and Benoît Schoefs[2]

[1] *Department of Plant and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden*
[2] *Mer, Molécules, Santé, Faculté des Sciences et Techniques, Université du Maine à Le Mans, Le Mans, France*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Daniel Hofius, University of Copenhagen, Denmark
Uener Kolukisaoglu, University of Tuebingen, Germany

***Correspondence:**
Cornelia Spetea, Department of Plant and Environmental Sciences, University of Gothenburg, 40530 Gothenburg, Sweden.
e-mail: cornelia.spetea.wiklund@dpes.gu.se

ATP is the common energy currency of cellular metabolism in all living organisms. Most of them synthesize ATP in the cytosol or on the mitochondrial inner membrane, whereas land plants, algae, and cyanobacteria also produce it on the thylakoid membrane during the light-dependent reactions of photosynthesis. From the site of synthesis, ATP is transported to the site of utilization via intracellular membrane transporters. One major type of ATP transporters is represented by the mitochondrial ADP/ATP carrier family. Here we review a recently characterized member, namely the thylakoid ATP/ADP carrier from *Arabidopsis thaliana* (AtTAAC). Thus far, no orthologs of this carrier have been characterized in other organisms, although similar sequences can be recognized in many sequenced genomes. Protein Sequence database searches and phylogenetic analyses indicate the absence of TAAC in cyanobacteria and its appearance early in the evolution of photosynthetic eukaryotes. The TAAC clade is composed of carriers found in land plants and some green algae, but no proteins from other photosynthetic taxa, such as red algae, brown algae, and diatoms. This implies that TAAC-like sequences arose only once before the divergence of green algae and land plants. Based on these findings, it is proposed that TAAC may have evolved in response to the need of a new activity in higher photosynthetic eukaryotes. This activity may provide the energy to drive reactions during biogenesis and turnover of photosynthetic complexes, which are heterogeneously distributed in a thylakoid membrane system composed of appressed and non-appressed regions.

Keywords: green alga, chloroplast, plant, photosynthesis, ADP/ATP carrier, thylakoid, TAAC phylogeny

INTRODUCTION

Oxygenic photosynthesis is a biophysicochemical process that converts carbon dioxide into organic compounds using sunlight as a source of energy. It occurs in the chloroplasts of land plants and algae, and also in the cytoplasm of cyanobacteria, and uses water as a source of electrons, releasing oxygen as a waste product (for a recent review, see Hohmann-Marriott and Blankenship, 2011). The chloroplast in algae and plants has evolved from a cyanobacterial ancestor via endosymbiosis with a primitive eukaryotic host. It is a highly compartmentalized organelle, with three membrane systems (outer envelope, inner envelope, and thylakoid) and three soluble spaces (intermembrane space, stroma, and thylakoid lumen). A wide variety of solute and metabolite transporters reside within the different types of chloroplast membranes and mediate communication between the cytosol, stroma, and lumen. Several excellent reviews on the identification and functional characterization of these transporters have recently become available (Spetea and Schoefs, 2010; Breuers et al., 2011; Flügge et al., 2011; Weber and Linka, 2011).

Regarding their evolution, the majority of the inner envelope metabolite transporters have been shown to have a host origin, driven by the requirement to establish communication between the host cytosol and the cyanobiont (Facchinelli and Weber, 2011). There are so far no evolutionary studies dedicated to solute transporters from the outer envelope or thylakoid membranes, but it is believed that they originate from proteins in the ancestral cyanobacterial outer and thylakoid membranes, respectively.

Mitochondria and chloroplasts are the two organelles able to synthesize ATP, which is the universal energy currency of cellular metabolism in all living organisms. The difference between these organelles in this respect is that in all eukaryotes mitochondria produce ATP via oxidative phosphorylation on the inner membrane, to be used during cell metabolism. Chloroplasts and also cyanobacteria use sunlight as a source of energy to produce ATP (photophosphorylation) on the thylakoid membrane, which is consumed during CO_2 fixation in the stroma. In addition, ATP is used for energy-dependent reactions on the envelope and thylakoid membrane or inside the thylakoid lumen (Spetea and Thuswaldner, 2008).

ATP is the largest and most highly charged solute transported across organellar membranes. Two structurally and phylogenetically different types of ATP transporters are represented in chloroplasts, namely ATP/ADP antiporters (AAA, TC #2.1.12, according to Saier et al., 2009) and mitochondrial ADP/ATP carriers (AAC,

TC #2.A.29.1.1). There are two ATP/ADP antiporters (AATPs) in the inner envelope membrane of photosynthetic and heterotrophic plastids, supplying cytosolic ATP to the stroma (for recent reviews, see Haferkamp et al., 2011; Traba et al., 2011). The transport is electroneutral since the counter-ions for ATP^{4-} are ADP^{3-} together with $H_2PO_4^-$ (Trentmann et al., 2008). AATPs possess 12 putative transmembrane helices, and share a common origin with the ATP/ADP antiporters found in the parasite bacteria *Rickettsia prowazekii* and *Chlamydia psittaci* (Haferkamp et al., 2011).

The mitochondrial ADP/ATP carriers are the first and most studied members of the mitochondrial carrier (MC) family and are present only in eukaryotic cells (Haferkamp et al., 2011; Traba et al., 2011). Like the members of the AAA family, AACs are exchangers, but the transport is electrogenic (ATP^{4-}/ADP^{3-}) and proceeds in the opposite direction, since they transport matrix ATP through the intermembrane space out into the cytosol. Yet another functional difference from AAA is that AACs are sensitive to specific inhibitors, such as bongkrekic acid and carboxyatractyloside (Klingenberg, 2008). The 3D structure of the bovine AAC has been resolved at 2.2 Å resolution in a conformation stabilized with carboxyatractyloside (Pebay-Peyroula et al., 2003). The structure revealed six transmembrane helices and a selectivity filter for adenine nucleotides, whose sequence could be used to predict other AACs and even other MCs (Nury et al., 2010). From a total of 58 MC in *Arabidopsis thaliana*, three classical mitochondrial AACs have been characterized, with at least four more paralogous sequences awaiting validation (Palmieri et al., 2011). In addition to the mitochondrion, AAC members have also been found in peroxisomes, endoplasmic reticulum, amyloplasts, and chloroplasts (Haferkamp et al., 2011; Traba et al., 2011).

Using western blotting and activity inhibition with an antibody against the bovine AAC, the activity of an AAC was reported in the spinach thylakoid membrane (Spetea et al., 2004). BLAST searches with the bovine AAC against the *Arabidopsis* protein database combined with prediction of chloroplast transit peptides revealed one putative chloroplast AAC, encoded by the *At5g1500* gene. The corresponding protein was localized to the thylakoid membrane of *Arabidopsis* (Thuswaldner et al., 2007; Zybailov et al., 2008), and was annotated as the thylakoid ATP/ADP carrier (AtTAAC). Subfractionation as well as immunocytochemical experiments indicated the non-appressed regions as the precise location of TAAC within the thylakoids (Thuswaldner et al., 2007). Another chloroplast AAC is encoded by the *At3g51870* gene and was initially localized to the envelope using western blotting (Spetea, C., unpublished data) and mass-spectrometry based proteomics (Ferro et al., 2010).

The AtTAAC sequence contains 415 amino acids that include a predicted transit peptide of 60 amino acids. The sequence of the processed form is 80 residues longer than that of bovine AAC, explaining the observed difference between the two proteins in the reported size in SDS gels (Spetea et al., 2004). The extra 80 residues are distributed as 50 in the N-terminus and 30 in the C-terminus – regions containing many charged residues and a five-glycine repeat that could play a role in the regulation of TAAC activity. TAAC shares about 30% identity with bovine AAC, which is concentrated in the six putative transmembrane helices and to a lesser degree in the connecting loops (Thuswaldner et al., 2007). The selectivity filter for adenine nucleotides, represented by residues K-130, R-186, Y-282, and K-369, is fully conserved, indicating adenine nucleotides as the most likely substrates for the transport activity of TAAC.

Arabidopsis TAAC was characterized in *E. coli* as an ATP importer in exchange for cytosolic ADP (Thuswaldner et al., 2007), and its activity was found sensitive to bongkrekic acid (Thuswaldner, S., and Spetea, C., unpublished data). Pi is not a substrate for transport by TAAC (Thuswaldner et al., 2007), implying the requirement for a separate thylakoid Pi transporter. Indeed, such a protein has been identified in *Arabidopsis*, and functionally characterized in yeast and *E. coli* (Guo et al., 2008; Pavón et al., 2008). When assessed in thylakoid membranes, TAAC transports stromal ATP into the thylakoid lumen in exchange for ADP. The direction of TAAC-mediated transport determined in both *E. coli* and thylakoids is opposite to the direction of transport by mitochondrial AACs (Thuswaldner et al., 2007). Therefore, to distinguish it from the mitochondrial ADP/ATP carrier, the thylakoid protein has been named ATP/ADP carrier. Through adenine nucleotide exchange, TAAC was proposed to supply ATP for nucleotide-dependent reactions in the thylakoid lumen (Spetea et al., 2004). An extensive review on the structure, function, and evolution of the MC family has become recently available and provides insights into their roles in plants (Palmieri et al., 2011). This review focuses on the evolutionary origin of the TAAC subfamily of the mitochondrial AACs, which aids in elucidating its function in the thylakoid membrane.

WHEN AND WHERE IN THE TREE OF LIFE DID TAAC ORIGINATE?

We assembled a set of protein sequences with which to place TAAC in an evolutionary context. We extracted protein sequences from the curated gene families at ARAMEMNON (Schwacke et al., 2003)[1] and from search results to specific clades and genomes at NCBI[2] and PHYTOZOME (Goodstein et al., 2011)[3], respectively. We added the best BlastP matches to AtTAAC from each major eukaryotic division of life with an *E*-value $<10^{-55}$. We also added all 17 members of the *Arabidopsis* MC family (including AtTAAC) that were listed as protein sequences related to TAAC at ARAMEMNON. We also added the best BlastP hits to proteins from several green plant clades, including Bryophyta, Chlorophyta, Lycopodiophyta, and Pinophyta. An alignment of these sequences was made using MUSCLE (Edgar, 2004, implemented at http://www.ebi.ac.uk). From the alignment (not shown) we could see that only two copies in *Arabidopsis* (encoded by the *At5g01500* and *At3g51870* genes) and all other non-*Arabidopsis* land plant proteins in this sample possessed a partly conserved N-terminal motif, consisting of 19 amino acid residues directly upstream from the first transmembrane helix delimited by Thuswaldner et al. (2007).

For the first phylogenetic analysis we included amino acid positions 121–369 in TAAC (spanning the six transmembrane

[1] http://aramemnon.uni-koeln.de/
[2] http://blast.ncbi.nlm.nih.gov/
[3] http://www.phytozome.net/

domains), which resulted in 340 aligned positions. The analysis was carried out using a Bayesian approach. We used four parallel chains of a reverse model jump protein Bayesian analysis in MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001). Each chain was run for two million generations and sampled every 1,000 generations. The tree in **Figure 1** shows a well-supported [posterior probability (PP) = 1.0] N-terminal containing clade (the "TAAC clade") that includes land plants and some copies from Chlorophyta, but no proteins from other taxa. Sister to this clade with strong support (PP = 1.0) is a clade (also PP = 1.0) containing proteins from Chlorophyta that lack the N-terminal motif, including proteins from *Chlamydomonas*, *Chlorella*, and *Volvox*.

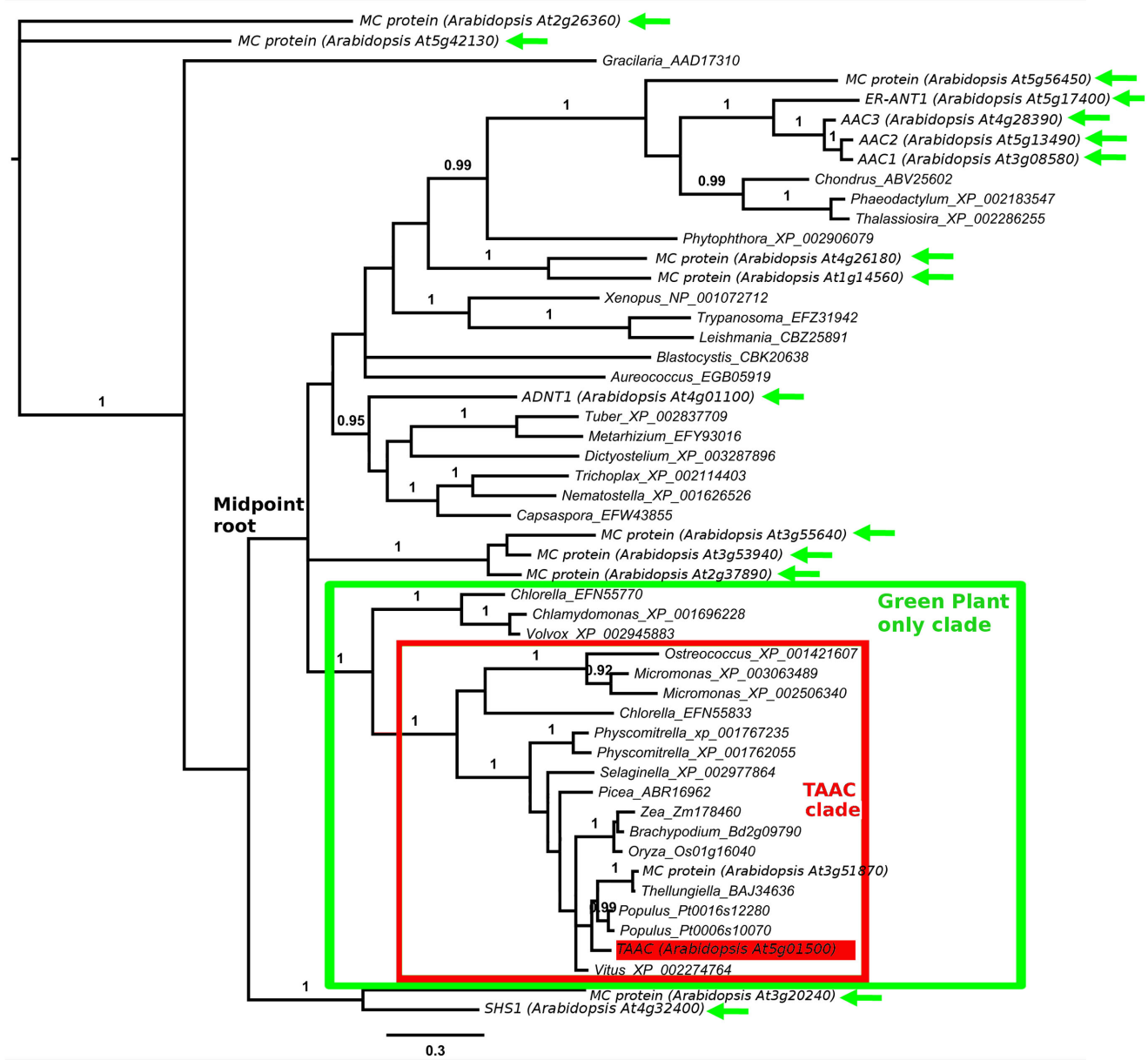


FIGURE 1 | Protein phylogeny of 340 Muscle aligned positions of selected members of the MC family. The analysis involved 51 protein sequences, including 17 from Arabidopsis. The root position shown is arbitrary. The mid-point root position is marked, as is the clade containing all TAAC-like sequences (boxed in red). The clade of Green Plants is marked in green. All plant sequences outside this clade are indicated by green arrows. Accession numbers follow the gene annotation (where available) and names of genera within which the proteins are found. The position of *Arabidopsis* TAAC is highlighted in red. The scale bar indicates substitutions per site. Posterior probabilities ≥0.90 of clades summarizing two million Markov chain Monte Carlo generations (of a reversible model jump protein Bayesian analysis in MrBayes 3.1.2) are shown above branches. The results discussed are robust to an alternative alignment (MAFFT, not shown).

The best BlastP matches to proteins from the following taxa did not contain the N-terminal motif, although they each possessed the conserved six transmembrane domain structure: Amoebozoa, Animalia, Euglenozoa, Fungi, Opisthokonta, Rhodophyta, and Stramenopiles – the latter includes brown algae and diatoms (nomenclature based on NCBI's taxonomy). All other eukaryote taxa at one hierarchical level below Eukaryota in NCBI's taxonomy were searched, but did not contain BlastP matches to AtTAAC with *E*-values $<10^{-55}$. Apart from TAAC, only one other protein from *Arabidopsis* is also a member of the TAAC clade and contains this motif, namely the one encoded by the *At3g51870* gene. The amino acid sequences of these two proteins are 67% identical.

The TAAC clade contains relatively short branches in relation to the remainder of the phylogeny, making it unlikely that the root resides within the clade. This is also consistent with mid-point rooting (marked in **Figure 1**) as well as the phylogenetic pattern expected among the plants that carry these proteins. Together, this indicates that the TAAC clade contains several proteins orthologous to AtTAAC. This clade also contains duplicated copies (paralogs) from several plant taxa (more are seen in our expanded sample in **Figure 2**, details below), but these are inferred to have arisen only after green plants diverged from other eukaryotes and are expected due to the action of recent gene and/or genome duplication (e.g., in *Arabidopsis*, Bowers et al., 2003; legumes, Pfeil et al., 2005; *Physcomitrella*, Rensing et al., 2007).

The position of the remaining 15 *Arabidopsis* MC proteins (**Figure 1**) indicates that the MC protein family was present very early in eukaryote evolution. Scattered across the tree are, for example, homologs from Metazoa (*Nematostella*), Fungi (*Tuber*), Amoebozoa (*Dictyostelium*), Stramenopiles (*Aureococcus*), Euglenozoa (*Leishmania*), and Rhodophyta (*Chondrus*). However, none of these homologs are part of the TAAC clade, which contains only proteins from Chlorophyta and Streptophyta. No matches within the BlastP cut-off used were found to other photosynthetic organisms, such as Bacillariophyta (diatoms), Cyanophyta (cyanobacteria), Glaucocystophyceae, Phaeophyceae (brown algae), or Rhodophyta (red algae).

Therefore, we conclude that TAAC-like sequences arose only a single time, after the clade including both Chlorophyta and Streptophyta (the latter including land plants) diverged from the other taxa in our sample, but earlier than the divergence of Chlorophyta and Streptophyta. The phylogenetic results are robust to an alternative alignment based on MAFFT (Katoh et al., 2002), although this is not shown here.

HOW DID THE TAAC COPIES IN *ARABIDOPSIS* AND OTHER EUDICOTS EVOLVE?

We increased the sample of TAAC-like proteins in a more focused phylogenetic inference in order to place the two *Arabidopsis* members of the TAAC subfamily into context. For this purpose, we sampled all copies from the species listed in **Table 1** that contain the N- and C-terminal motifs similar to TAAC and that were also part of the TAAC clade. We focused on eudicots, because preliminary analyses indicated that the two *Arabidopsis* TAAC subfamily members were placed in a clade containing proteins from eudicot species only.

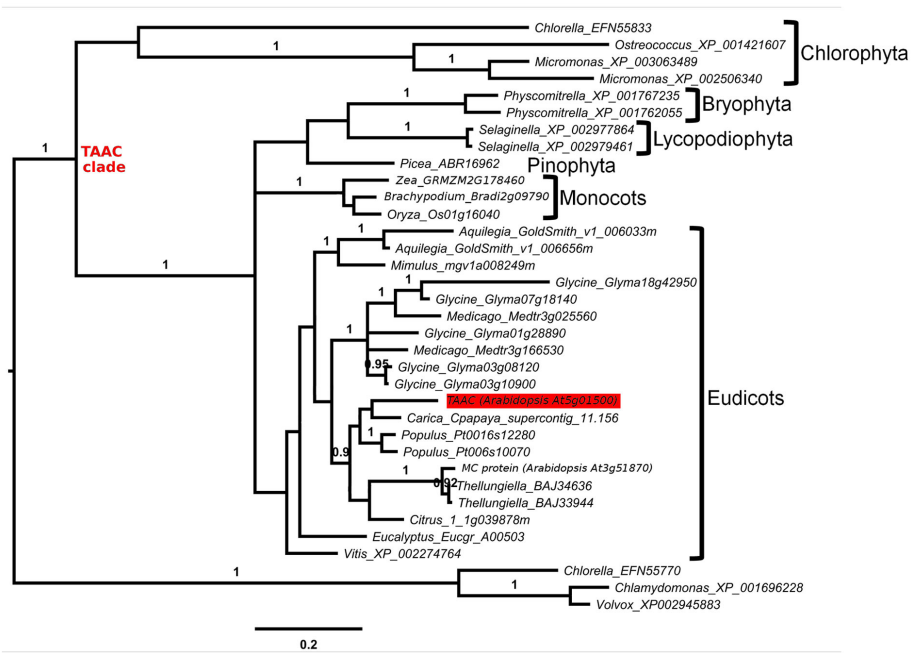


FIGURE 2 | Protein phylogeny of 379 Muscle aligned positions of selected members of the MC family, focusing on the TAAC clade (marked). The analysis involved 35 protein sequences, including two from *Arabidopsis*. The root position is based on **Figure 1**. Accession numbers follow the names of genera within which the proteins are found. The taxon to which each subclade belongs is indicated. The position of *Arabidopsis* TAAC is highlighted in red. The scale bar indicates substitutions per site. Posterior probabilities ≥0.90 of clades, analyzed as in **Figure 1**, are shown above branches. The results discussed are robust to an alternative alignment (MAFFT, not shown).

Table 1 | Copy number of TAAC-like proteins found in various eudicot plants.

Taxon	Group (NCBI taxonomy)	TAAC-like copies
Arabidopsis thaliana (L.) Heynh	Malvids	2
Thellungiella halophila O. E. Schulz	Malvids	2
Eucalyptus grandis W. Hill	Malvids	1
Carica papaya L.	Malvids	1
Citrus sinensis Osbeck	Malvids	1
Populus trichocarpa Torr and A. Gray	Fabids	2
Medicago truncatula Gaertn	Fabids	2
Glycine max (L.) Merr	Fabids	5
Vitis vinifera L.	Rosids incertae sedis	1
Mimulus guttatus DC	Asterids	1
Aquilegia caerulea E. James	Stem eudicots	2

We expanded our search for land plant sequences containing the N-terminal motif by using BlastP with AtTAAC as the query to specific genomes in Phytozome, as well as to *Thellungiella* (Brassicaceae), because in earlier BlastP searches a *Thellungiella* sequence closely related to TAAC was detected. We excluded possible allelic variants that differed by only a single amino acid from our enumeration. The alignment of this set of sequences was performed as before. From this alignment (Figure S1 in Supplementary Material), we used amino acids 107–482 in TAAC (corresponding to 375 aligned positions and including the N-terminal and C-terminal motifs) in our second phylogenetic analysis and also added the sister group of the TAAC clade (from **Figure 1**) as an outgroup. The phylogenetic analysis was performed as described for the data shown in **Figure 1**, with sample information summarized in **Table 2** and the results presented in **Figure 2**.

The copy number of TAAC-like proteins in the eudicots examined here appears to be rather conservative, ranging from one to five copies (**Table 1**). *Arabidopsis*, *Medicago*, and *Populus* each have only two copies, but these are perhaps the best-known eudicot genomes. Therefore, the copy number for the other species may be underestimated. The lack of support for some key nodes in the phylogeny (**Figure 2**) makes it difficult to pinpoint the timing of divergence of the two copies found in *Arabidopsis*. They almost certainly diverged sometime before this genus diverged from *Thellungiella* (perhaps 40–50 million years ago: Amtmann, 2009; Beilstein et al., 2010), but it is difficult to say how much earlier this divergence might have been.

Unlike the *Arabidopsis* copies, one of the copies found in *Aquilegia* differs from its closest relative by a long branch (**Figure 2**). In contrast, the rather recently diverged (in comparison to *Arabidopsis*) multiple copies found in *Selaginella* and *Physcomitrella* do not show this pattern. The long branch observed in *Aquilegia* may indicate functional shifts and positive selection may be involved, as has been shown in, e.g., monkey pancreatic ribonucleases (Zhang et al., 2002). These possibilities could be examined further and may lead to novel research questions. However, in the case of one *Glycine* sequence, poor assembly is a more likely explanation for the long branch observed in that taxon (NB: alignment positions 308–389 in Figure S1 in Supplementary Material), so caution needs to be taken in specific cases.

HOW DOES THE N-TERMINAL MOTIF VARY WITHIN THE TAAC SUBFAMILY?

The N-terminal motif may be helpful to identify potential members of the TAAC subfamily. Furthermore, this clade shows complete conservation of the ring of four positively charged residues (Figure S1 in Supplementary Material) proposed to act as a selectivity filter for adenine nucleotides (Thuswaldner et al., 2007), confirming that these aligned proteins are all AACs.

Potentially important is that significant variation in the N-terminal sequences (beyond just the 19 residue motif described earlier) occurs among the proteins sampled here. For example, several proteins contain many charged residues, including paired charged ones, and a five-glycine repeat (Figure S1 in Supplementary Material). These may be implicated in ligand (e.g., calcium) binding and large conformational changes that could regulate their activity, as in the case of Ca^{2+}-dependent members of the MC family (Weber et al., 1997). However, some of these proteins do not contain either of these features. For example, the *Selaginella* sequences have only 13 residues upstream of the more conserved N-terminal motif, few charged residues, and lack the five-glycine repeat. One *Micromonas* protein appears to lack the N-terminal motif altogether (Figure S1 in Supplementary Material). Variation in the N-terminal motif among the land plant TAAC-like proteins is smaller (**Table 2**), but so is the phylogenetic distance. It is tempting to speculate as to what the precise function of the N-terminal residues might be, moreover, interesting research questions can be formulated around these sequences. However, the sequences (especially those apparently lacking the N-terminal motif) should of course be verified before pursuing these lines of investigation, as technical problems (e.g., with the assembly), deposition of incomplete sequences, etc., need to be ruled out first.

WHAT ARE THE POSSIBLE FUNCTIONS OF TAAC IN LAND PLANTS AND GREEN ALGAE?

The phylogenetic analyses performed in this work indicate that TAAC is absent in cyanobacteria and that the most recent common ancestor of green plants (Streptophyta + Chlorophyta) was the earliest photosynthetic organism that we can identify to have carried a TAAC-like protein. Although TAAC is a thylakoid protein and the thylakoid membrane originated with the thylakoid membrane of an ancestral cyanobacterium, TAAC appears to be eukaryote-specific, as are all mitochondrial AACs (Palmieri et al., 2011). Nevertheless, in contrast to mitochondrial AACs, which are found in both photosynthetic and heterotrophic eukaryotes (Palmieri et al., 2011), only some photosynthetic eukaryotes, more specifically, Chlorophyta and Streptophyta, carry TAAC-like copies. Other photosynthetic eukaryotes (red algae, brown algae, and diatoms) and, interestingly, even some Chlorophyta (*Chlamydomonas* and *Volvox*) do not appear to have such a protein, i.e., not one with a similar N-terminus nor placed in the TAAC clade, whereas other Chlorophyta do (e.g., *Chlorella*; **Table 2**).

A potential answer to the question "why do not all photosynthetic organisms carry TAAC-like proteins?" could be that TAAC may fulfill specialized functions in the thylakoid membrane of

Table 2 | The presence (green) or absence (red) of TAAC-like proteins within the surveyed genomes is shown.

Clade	Species	Accessions used (Figure 2 only)	Annotation[1]	N-terminal[2] % AA identity to TAAC	Accession source
Eudicots	*Aquilegia caerulea*	AcoGoldSmith_v1.006656m	MC[3] protein	68	Phytozome
		AcoGoldSmith_v1.006033m	MC protein	58	Phytozome
	Arabidopsis thaliana	**At5g01500**	**Thylakoid ATP/ADP carrier (TAAC)**	**100**	Phytozome
		At3g51870	MC protein	68	Phytozome
	Carica papaya	evm.TU.supercontig_11.156	MC protein	58	Phytozome
	Citrus sinensis	orange1.1g039878m	MC protein	Not available	Phytozome
	Eucalyptus grandis	Eucgr.A00503	MC protein	53	Phytozome
	Glycine max	Glyma03g10900	MC protein	Not available	Phytozome
		Glyma03g08120	MC protein	58	Phytozome
		Glyma01g28890	MC protein	Not available	Phytozome
		Glyma07g18140	MC protein	53	Phytozome
		Glyma18g42950	MC protein	58	Phytozome
	Medicago truncatula	Medtr3g166530	MC protein	Not available	Phytozome
		Medtr3g025560	MC protein	58	Phytozome
	Mimulus guttatus	mgv1a008249m	MC protein	47	Phytozome
	Populus trichocarpa	Pt0006s10070	MC protein	68	Aramemnon
		Pt0016s12280	MC protein	63	Aramemnon
	Thellungiella halophila	BAJ33944	MC protein	68	NCBI
		BAJ34636	MC protein	68	NCBI
	Vitis vinifera	XP_002274764	MC protein	47	NCBI
Monocots	*Oryza sativa*	LOC_Os01g16040	MC protein	26	Phytozome
	Brachypodium distachyon	Bradi2g09790	MC protein	37	Phytozome
	Zea mays	GRMZM2G178460	MC protein	53	Phytozome
Pinophyta	*Picea sitchensis*	ABR16962	MC protein (NCBI)	42	NCBI
Lycopodiophyta	*Selaginella moellendorffii*	XP_002977864	Putative MC protein (NCBI)	42	NCBI
		XP_002979461	Putative MC protein (NCBI)	42	NCBI
Bryophyta	*Physcomitrella patens*	XP_001762055	Putative MC protein (NCBI)	37	NCBI
		XP_001767235	Putative MC protein (NCBI)	42	NCBI
Chlorophyta	*Micromonas* sp. *RCC299*	XP_002506340	Putative MC protein (NCBI)	15	NCBI
	Micromonas pusilla	XP_003063489	Amyloplast brittle-1 (BT1) protein homolog (NCBI)	Not available	NCBI
	Ostreococcus lucimarinus	XP_001421607	ADP/ATP transporter on adenylate translocase; provisional (NCBI)	26	NCBI
	Chlorella variabilis	EFN55833	As above	26	NCBI
		EFN55770	ADP/ATP transporter on adenylate translocase; provisional (NCBI)	Not available	NCBI
	Volvox carteri	XP_002945883	ADP/ATP transporter on adenylate translocase; provisional (NCBI)	Not present	NCBI
	Chlamydomonas reinhardtii	XP_001696228	ADP/ATP transporter on adenylate translocase; provisional (NCBI)	Not present	NCBI
Alveolata			No hits $<10^{-55}$		
Amoebozoa	*Dictyostelium purpureum*	XP_003287896	Putative MC protein (NCBI)	Not present	NCBI
Apusozoa			No hits $<10^{-55}$		
Centroheliozoa			No hits $<10^{-55}$		
Cryptophyta			No hits $<10^{-55}$		

(Continued)

Table 2 | Continued

Clade	Species	Accessions used (Figure 2 only)	Annotation[1]	N-terminal[2] % AA identity to TAAC	Accession source
Euglenozoa	*Leishmania mexicana*	CBZ25891	Putative MC protein (NCBI)	Not present	NCBI
Fornicata			No hits $<10^{-55}$		
Glaucocystophyceae			No hits $<10^{-55}$		
Haptophyceae			No hits $<10^{-55}$		
Jakobida			No hits $<10^{-55}$		
Katablepharidophyta			No hits $<10^{-55}$		
Malawimonadidae			No hits $<10^{-55}$		
Opisthokonta	*Capsaspora owczarzaki*	EFW43855	MC protein (NCBI)	Not present	NCBI
	Nematostella vectensis	XP_001626526	Putative MC protein (NCBI)	Not present	NCBI
	Trichoplax adhaerens	XP_002114403	Putative MC protein (NCBI)	Not present	NCBI
	Tuber melanosporum	XP_002837709	Putative MC protein (NCBI)	Not present	NCBI
	Xenopus tropicalis	NP_001072712	MC protein (NCBI)	Not present	NCBI
	Metarhizium acridum	EFY93016	Putative MC protein (NCBI)	Not present	NCBI
Oxymonadida			No hits $<10^{-55}$		
Parabasalia			No hits $<10^{-55}$		
Rhizaria			No hits $<10^{-55}$		
Rhodophyta	*Chondrus crispus*	ABV25602	Putative MC protein (NCBI)	Not present	NCBI
	Gracilaria gracilis	AAD17310	Putative MC protein (NCBI)	Not available	NCBI
Stramenopiles	*Phytophthora infestans*	XP_002906079	Putative MC protein (NCBI)	Not present	NCBI
	Bacillariophyta (diatoms)		No hits $<10^{-55}$		
	Phaeophyceae (brown algae)		No hits $<10^{-55}$		
Bacteria (including cyanobacteria)			No hits $<10^{-55}$		

The taxon, species of origin, accession numbers, annotation, and sequence identity of the N-terminus relative to the Arabidopsis TAAC are indicated. The information about Arabidopsis TAAC is highlighted in bold. The source of these data is given in the last column: Phytozome, ARAMEMNON, or NCBI. Not available, incomplete gene model (no start codon); not present, the sequence appears to be complete, but lacks anything resembling the 19 amino acid N-terminal motif common to most sequences in the TAAC clade. Green = taxa with genes in the TAAC clade. Red = taxa without genes in the TAAC clade. E-value for BlastP (http://blast.ncbi.nlm.nih.gov) against Arabidopsis thaliana TAAC. For clades where TAAC type sequences are not found, the best hit (September 2011) is recorded instead, down to a threshold of 10^{-55}. We assume that matches with E-value $<10^{-55}$ will not be closely related to the TAAC clade, given that many matches better than this are also not part of this clade. "Putative" annotations include "hypothetical" annotations and annotations based on conceptual translations. [1]Annotations come from Phytozome unless stated otherwise in the table. [2]N-terminal refers to the 19 residue N-terminal motif discussed in the text. [3]MC, mitochondrial carrier.

land plants and most green algae. These membranes are organized in highly stacked (appressed) regions interconnected by stroma-exposed (non-appressed) regions (**Figure 3**). Other photosynthetic organisms (cyanobacteria, red algae, brown algae, and diatoms) display unstacked or weakly stacked thylakoids (for details and references, see the legend to **Figure 3**). To explain the apparent difference within green algae, it is relevant to consider that *Chlamydomonas* has appressed thylakoids, but they do not form regular grana stacks (de Vitry and Vallon, 1999). To our knowledge, the thylakoid structure has not, thus far, been studied in green algae other than *Chlamydomonas*, but they are expected to have a similar organization to the land plant thylakoids. The thylakoid organization depends on the type and arrangement of light-harvesting antennae and on the distribution of photosynthetic complexes (**Figure 3**). It was held for a long time that the different macrocomplexes comprising the photosynthetic apparatus were organized linearly along thylakoid membranes. This view is no longer valid since it has been established that these complexes are laterally distributed, i.e., localized exclusively in the appressed membranes (photosystem II), exclusively in the stroma-exposed thylakoids (photosystem I, ATP synthase) or in both types of membranes (cytochrome b_6/f complex; Anderson, 2002). This heterogenous composition of thylakoids is restricted to the green algae and land plants (**Figure 3**).

The differences in thylakoid membrane organization among cyanobacteria, green algae, and land plants may have implications for biogenesis and turnover of photosynthetic complexes, such as photosystem II (PSII). The biosynthesis of complexes in cyanobacteria occurs at contact sites with the plasma membrane, whereas PSII repair takes place within thylakoid membranes (Zak et al., 2001). In green algae, the site of PSII assembly during *de novo* D1 synthesis is around a so-called pyrenoid, a specialized sub-compartment within the chloroplast for CO_2 fixation that is different from the site of PSII repair – the stroma-regions of the thylakoid membrane (Uniacke and Zerges, 2007). In land plants, the location of thylakoid membrane complex biosynthesis is uncertain. A widely accepted model is that PSII subunits assemble in the non-appressed region of the thylakoid membrane (Baena-González and Aro, 2002). The mechanism of PSII repair in plants has been studied in detail and shown to have important

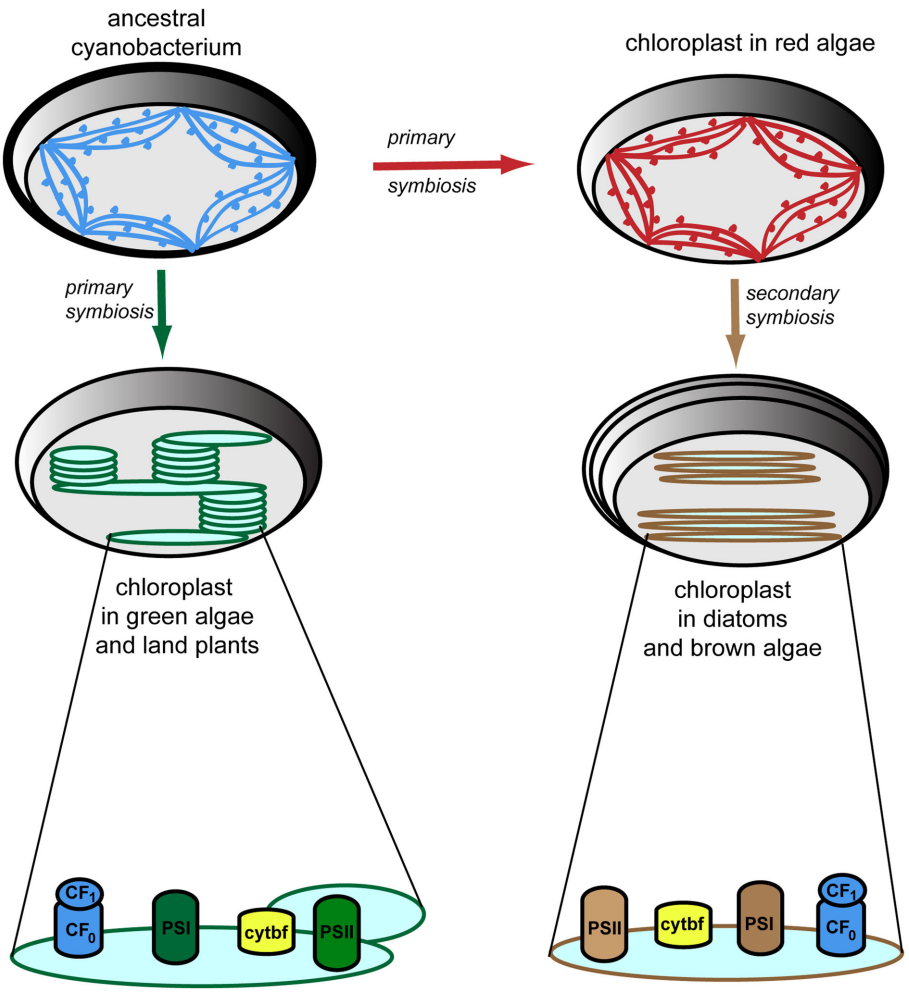


FIGURE 3 | Evolution of thylakoid membrane organization and composition. Cyanobacterial thylakoid membranes (shown in blue) are located directly in the cytoplasm, are arranged in multiple (three to eight) layers, and make contacts with the plasma membrane. They have attached phycobilisomes for harvesting light, which protrude toward the cytoplasm, thus preventing membrane stacking (Liberton and Pakrasi, 2008). Chloroplasts have evolved via primary symbiosis from a cyanobacterial relative. Thus, similar arrangement of the thylakoid membranes is observed in red algae (shown in red), which also have attached phycobilisomes (Tsekos et al., 1996). Thylakoid membranes in diatoms and brown algae (shown in brown) are arranged in groups of three loosely appresed vesicles and contain fucoxanthin–chlorophyll-protein complexes for harvesting light (Bertrand, 2010). Note the four-layer envelopes surrounding the chloroplast as compared to the types found in other algae and plants, as a result of secondary symbiosis. Thylakoid membranes in green algae and higher plants (shown in green) are located inside the chloroplast, are organized in grana stacks (5–20 vesicles) interconnected by stroma-exposed lamellae, and contain chlorophyll–protein complexes for harvesting light (Mustardy and Garab, 2003). Various 3D models of the organization of these thylakoid membranes have been proposed and debated (Daum and Kühlbrandt, 2011). According to the most accepted model, based on electron tomography (Daum et al., 2010; Austin and Staehelin, 2011), grana should be seen as stacks of straight, exactly parallel pairs of thylakoid membranes, each pairs of grana disks merging with stroma lamellae by staggered membranes protrusions. Importantly, but not shown in this figure, the plane of the stromal thylakoids is tilted by 10–25° relatively to the plane of the grana stack so that the stromal thylakoids display a helix around the granum (Mustardy et al., 2008). The photosynthetic apparatus is composed of four multisubunit complexes, namely the water-oxidizing photosystem II (PSII), cytochrome b_6/f (cytbf), photosystem I (PSI), and the H^+-translocating ATP synthase (CF_0F_1; Nelson and Ben-Shem, 2004). These complexes are laterally distributed in green algae and plants, whereas in other algae and diatoms, they display a more uniform distribution.

differences from that in cyanobacteria, such as the shuttling of PSII subcomplexes between the grana and stroma membranes, which is accompanied by changes in the oligomeric structure of the complex (Mulo et al., 2011).

The question arising is whether TAAC plays a role in the biogenesis and turnover of plant photosynthetic complexes. Its tissue expression pattern has been studied in detail, and supports this possibility, as described below. AtTAAC was found highly expressed in young photosynthetic organs, such as developing leaves, flower buds, and green siliques. Furthermore, it was found expressed in etiolated seedlings, similar to the thylakoid HCF136 protein, which was proven to be required for biogenesis of thylakoids and PSII assembly (Meurer et al., 1998; Plücken et al., 2002). TAAC expression is strongly upregulated in leaves

undergoing senescence or exposed to wounding, light stress, oxidative stress, salt stress, and desiccation, pointing to an additional role in supplying ATP for energy-dependent processes (e.g., proteolysis, folding) during turnover of photosynthetic complexes.

Thus far, no dedicated studies have been undertaken on the role of TAAC activity during thylakoid biogenesis. Nevertheless, its role during PSII repair has been studied in more detail by taking advantage of *taac* mutants. Remarkably, the mutants grew slower and were more sensitive to high light stress due to an inability to degrade the reaction center D1 protein (Yin et al., 2010). A model for its function during PSII repair cycle in plants under high light stress is presented in **Figure 4**. Briefly, ATP translocated by TAAC into the lumen is converted to GTP, which is then bound to the PsbO lumenal extrinsic subunit of the PSII dimeric complex (Spetea et al., 2004; Lundin et al., 2007a). The GTPase activity of this protein regulates the monomerization and partial disassembly of PSII, pre-requisite steps for the proteolytic degradation of the D1 protein (Lundin et al., 2007b). Replacement of the damaged D1 with a new copy requires coordination between its degradation, synthesis, insertion, and assembly of the new copy in the PSII complex. We propose that the novel mechanism of ATP transport and GTP signaling across thylakoid membranes, discussed above, may be a plant-specific feature for the following reasons: the D1 degradation is GTP dependent in plants but not in cyanobacterial thylakoids (Spetea et al., 1999); GTP binding to PsbO may be a plant-specific feature, since putative GTP binding sequence domains are not found conserved in cyanobacteria or the green alga *Chlamydomonas* PsbOs (Lundin et al., 2007a); TAAC belongs to the MC family, which is eukaryotic specific (Palmieri et al., 2011). Thus, we speculate that TAAC is associated with the need for highly controlled regulation of PSII repair in a highly stacked thylakoid membrane system requiring shuttling of complexes between the appressed and non-appressed regions.

CONCLUSION

Study of the adenine nucleotide carrier TAAC in chloroplast thylakoids is interesting because it has revealed similarities but also differences from classical mitochondrial AACs. Most MC (including AACs) are present in both phototrophic and heterotrophic eukaryotes, suggesting functions required for basic eukaryotic processes that take place inside the mitochondria and other organelles. On the other hand, TAAC-like proteins have been found only in green plants, indicating at the earliest an appearance after the endosymbiosis of a phototrophic prokaryote with the most recent common ancestor of land plants and green algae. Essentially, TAAC is an AAC with additional N (and C) motifs, apart from the chloroplast targeting sequence. The pathway by which TAAC arrived in the chloroplast and the precise function of the N-motif in TAAC activity remain to be investigated. Nevertheless, based on the phylogenetic study presented here, we propose that TAAC may fulfill functions that may be required in the context of the more complex organization of thylakoid membranes in green algae and land plants, such as biogenesis and repair of photosynthetic complexes. Some of the lumenal enzymes participating in the biogenesis/turnover processes may require ATP or other nucleotides resulting from inter-conversion, and therefore ATP must be translocated. This would not be unexpected, taking into consideration the increasing evidence for a complex role of the thylakoid lumen in photosynthetic regulation and plant cell signaling, expanding its function beyond an energetic perspective

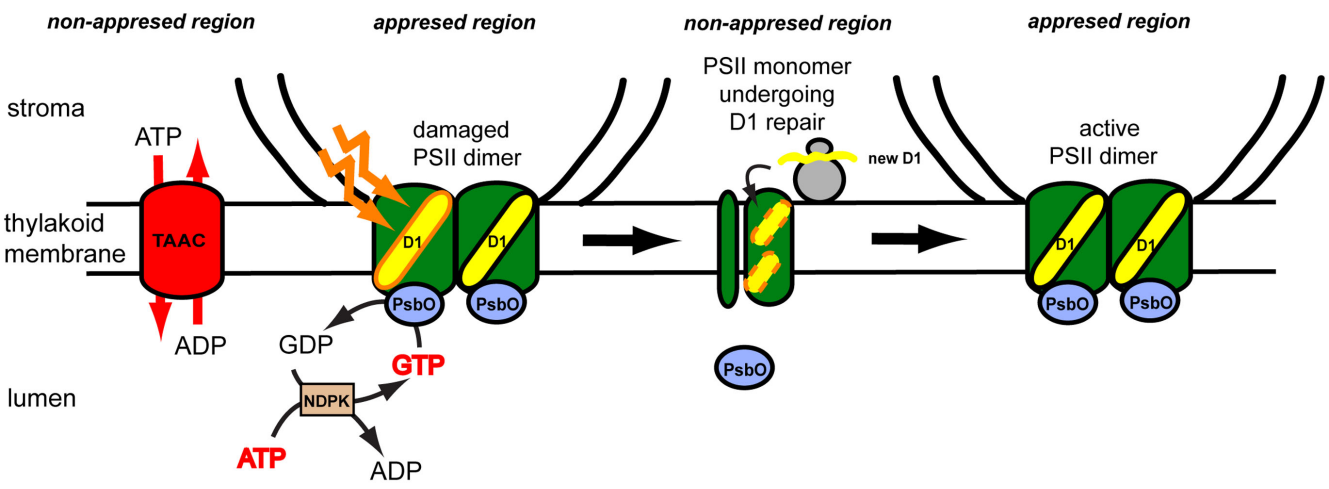


FIGURE 4 | Proposed model for role of the thylakoid ATP/ADP carrier (TAAC) during photosystem II (PSII) repair cycle in plants. The plant thylakoid membrane is organized in grana stacks (appressed region) and stroma-exposed (non-appressed region) lamellae. The composition of the thylakoid membrane is heterogenous, although its lumenal space is continuous. TAAC is located in the stroma lamellae and exchanges stromal ATP for lumenal ADP. ATP is inter-converted to GTP by the lumenal nucleoside diphosphate kinase (NDPK). The active PSII dimer is located in the grana regions and contains lumenal extrinsic PsbO proteins. During illumination with excess light, the reaction center D1 protein may be oxidatively damaged and needs to be replaced. The PsbO subunit of the damaged monomer binds and hydrolyzes GTP, leading to its dissociation and partial disassembly of the monomeric complex on the way to the stroma-exposed regions. Here the D1 protein is degraded and replaced with a new copy synthesized by the chloroplast ribosomes and co-translationally inserted in the membrane. Monomers assemble into the dimers and migrate to the grana regions.

(for a recent review, see Spetea, 2011). Thus, learning about TAAC evolution can be very important to understand its functional importance in green plants.

ACKNOWLEDGMENTS

Research in the C.S. laboratory has been supported by grants from the Swedish Research Council and the Swedish Research Council for Environment, Agriculture and Space Planning (Formas). Bernard E. Pfeil acknowledge funding from: Swedish Research Council, Royal Swedish Academy of Sciences, Royal Physiographic Society in Lund, Lars Hiertas Minne fund, Helge Ax:son Johnsons Foundation, and Lundgrenska fund. Research in the Benoît Schoefs laboratory has been supported by grants from Université du Maine à Le Mans.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Physiology/10.3389/fpls.2011.00110/abstract

Figure S1 | MUSCLE alignment of TAAC subfamily of full-length amino acid sequences.

REFERENCES

Amtmann, A. (2009). Learning from evolution: *Thellungiella* generates new knowledge on essential and critical components of abiotic stress tolerance in plants. *Mol. Plant* 2, 3–12.

Anderson, J. M. (2002). Changing concepts about the distribution of photosystems I and II between grana-appressed and stroma-exposed thylakoid membranes. *Photosynth. Res.* 73, 157–164.

Austin, J. R., and Staehelin, L. A. (2011). Three-dimensional architecture of grana and stroma thylakoids of higher plants as determined by electron tomography. *Plant Physiol.* 155, 1601–1611.

Baena-González, E., and Aro, E. M. (2002). Biogenesis, assembly and turnover of photosystem II units. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1451–1459.

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R., and Mathews, S. (2010). Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18724–18728.

Bertrand, M. (2010). Carotenoid biosynthesis in diatoms. *Photosynth. Res.* 106, 89–102.

Bowers, J. E., Chapman, B. A., Rong, J., and Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438.

Breuers, F., Braeutigam, A., and Weber, A. P. M. (2011). The plastid outer envelope – a highly dynamic interface between plastid and cytoplasm. *Front. Plant Sci.* 2:97. doi:10.3389/fpls.2011.00097

Daum, B., and Kühlbrandt, W. (2011). Electron tomography of plant thylakoid membranes. *J. Exp. Bot.* 62, 2393–2402.

Daum, B., Nicastro, D., Austin, J. R., McIntosh, J. R., and Kühlbrandt, W. (2010). Arrangement of photosystem II and ATP synthase in chloroplast membranes of spinach and pea. *Plant Cell* 22, 1299–1312.

de Vitry, C., and Vallon, O. (1999). Mutants of *Chlamydomonas*: tools to study thylakoid membrane structure, function and biogenesis. *Biochimie* 81, 631–643.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Facchinelli, F., and Weber, A. P. M. (2011). The metabolite transporters of the plastid envelope. An update. *Front. Plant Sci.* 2:50. doi:10.3389/fpls.2011.00050

Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S., Kieffer-Jaquinod, S., Bruley, C., Garin, J., Joyard, J., Masselon, C., and Rolland, N. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell Proteomics* 9, 1063–1084.

Flügge, U. I., Häusler, R. E., Ludewig, F., and Gierth, M. (2011). The role of transporters in supplying energy to plant plastids. *J. Exp. Bot.* 62, 2381–2392.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D. S. (2011). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186.

Guo, B., Jin, Y., Wussler, C., Blancaflor, E. B., Motes, C. M., and Versaw, W. K. (2008). Functional analysis of the *Arabidopsis* PHT4 family of intracellular phosphate transporters. *New Phytol.* 177, 889–898.

Haferkamp, I., Fernie, A. R., and Neuhaus, H. E. (2011). Adenine nucleotide transport in plants: much more than a mitochondrial issue. *Trends Plant Sci.* 16, 507–515.

Hohmann-Marriott, M. F., and Blankenship, R. E. (2011). Evolution of photosynthesis. *Annu. Rev. Plant Biol.* 62, 515–548.

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755.

Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.

Klingenberg, M. (2008). The ADP and ATP transport in mitochondria and its carrier. *Biochim. Biophys. Acta* 1778, 1978–2021.

Liberton, M., and Pakrasi, H. (2008). "Membrane systems in cyanobacteria," in *The Cyanobacteria: Molecular Biology, Genomics and Evolution*, eds A. Herrero, and E. Flores (Norwich: Caister Academic Press), 271–289.

Lundin, B., Thuswaldner, S., Shutova, T., Eshaghi, S., Samuelsson, G., Barber, J., Andersson, B., and Spetea C. (2007a). Subsequent events to GTP binding by the plant PsbO protein: structural changes, GTP hydrolysis and dissociation from the photosystem II complex. *Biochim. Biophys. Acta* 1767, 500–508.

Lundin, B., Hansson, M., Schoefs, B., Vener, A. V., and Spetea, C. (2007b). The *Arabidopsis* PsbO2 protein regulates dephosphorylation and turnover of the photosystem II reaction centre D1 protein. *Plant J.* 49, 528–539.

Meurer, J., Plücken, H., Kowallik, K. V., and Westhoff, P. (1998). A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in *Arabidopsis thaliana*. *EMBO J.* 7, 5286–5297.

Mulo, P., Sakurai, I., and Aro, E. M. (2011). Strategies for psbA gene expression in cyanobacteria, green algae and higher plants: from transcription to PSII repair. *Biochim. Biophys. Acta* 1817, 247–257.

Mustardy, L., Buttle, K., Steinbach, G., and Garab, G. (2008). The three dimensional network of the thylakoid membranes in plants: quasihelical model of the granum–stroma assembly. *Plant Cell* 20, 2552–2557.

Mustardy, L., and Garab, G. (2003). Granum revisited. A three dimensional model–where things fall into place. *Trends Plant Sci.* 8, 117–122.

Nelson, N., and Ben-Shem, A. (2004). The complex architecture of oxygenic photosynthesis. *Nat. Rev. Mol. Cell Biol.* 5, 971–982.

Nury, H., Blesneac, I., Ravaud, S., and Pebay-Peyroula, E. (2010). Structural approaches of the mitochondrial carrier family. *Methods Mol. Biol.* 654, 105–117.

Palmieri, F., Pierri, C. L., De Grassi, A., Nunes-Nesi, A., and Fernie, A. R. (2011). Evolution, structure and function of mitochondrial carriers: a review with new insights. *Plant J.* 66, 161–181.

Pavón, R. L., Lundh, F., Lundin, B., Mishra, A., Persson, B. L., and Spetea, C. (2008). *Arabidopsis* ANTR1 is a thylakoid Na^+-dependent phosphate transporter: functional characterization in *Escherichia coli*. *J. Biol. Chem.* 283, 13520–13527.

Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trézéguet, V., Lauquin, G. J., and Brandolin, G. (2003). Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* 426, 39–44.

Pfeil, B. E., Schlueter, J. A., Shoemaker, R. C., and Doyle, J. J. (2005). Placing paleopolyploidy in relation to taxon divergence: a phylogenetic analysis in legumes using 39 gene families. *Syst. Biol.* 54, 441–454.

Plücken, H., Müller, B., Grohmann, D., Westhoff, P., and Eichacker, L. A. (2002). The HCF136 protein is essential for assembly of the photosystem II reaction center in *Arabidopsis thaliana*. *FEBS Lett.* 532, 85–90.

Rensing, S. A., Ick, J., Fawcett, J. A., Lang, D., Zimmer, A., Van de Peer, Y., and Reski, R. (2007). An ancient genome duplication contributed to the abundance of metabolic genes in the moss *Physcomitrella* patens. *BMC Evol. Biol.* 7, 130. doi:10.1186/1471-2148-7-130

Saier, M. H. Jr, Yen, M. R., Noto, K., Tamang, D. G., and Elkan, C. (2009). The transporter classification database: recent advances. *Nucleic Acids Res.* 37, D274–D278.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W. B., Flügge, U. I., and Kunze, R. (2003). Aramemnon, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol.* 131, 16–26.

Spetea, C. (2011). "Role of chloroplast thylakoid lumen in photosynthetic regulation and plant cell signaling," in *Progress in Botany*, Vol. 73, eds U. Lüttge, W. Beyschlag, B. Büdel, D. Francis (Heidelberg: Springer-Verlag), 207–230.

Spetea, C., Hundal, T., Lohmann, F., and Andersson, B. (1999). GTP bound to chloroplast thylakoid membranes is required for light-induced, multienzyme degradation of the photosystem II D1 protein. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6547–6552.

Spetea, C., Hundal, T., Lundin, B., Heddad, M., Adamska, I., and Andersson, B. (2004). Multiple evidence for nucleotide metabolism in the chloroplast thylakoid lumen. *Proc. Natl. Acad. Sci. U.S.A.* 101, 1409–1414.

Spetea, C., and Schoefs, B. (2010). Solute transporters in plant thylakoid membranes: key players during photosynthesis and light stress. *Commun. Integr. Biol.* 3, 122–129.

Spetea, C., and Thuswaldner, T. (2008). "Update in nucleotide-dependent processes in plant chloroplasts," in *Plant Cell Compartments – Selected Topics*, ed. B. Schoefs (Trivandrum: Research Signpost), 105–149.

Thuswaldner, S., Lagerstedt, J. O., Rojas-Stütz, M., Bouhidel, K., Der, C., Leborgne-Castel, N., Mishra, A., Marty, F., Schoefs, B., Adamska, I., Persson, B. L., and Spetea, C. (2007). Identification, expression, and functional analyses of a thylakoid ATP/ADP carrier from *Arabidopsis*. *J. Biol. Chem.* 282, 8848–8859.

Traba, J., Satrústegui, J., and del Arco, A. (2011). Adenine nucleotide transporters in organelles: novel genes and functions. *Cell. Mol. Life Sci.* 68, 1183–1206.

Trentmann, O., Jung, B., Neuhaus, H. E., and Haferkamp, I. (2008). Nonmitochondrial ATP/ADP transporters accept phosphate as third substrate. *J. Biol. Chem.* 283, 36486–36493.

Tsekos, I., Reiss, H. D., Orfanidisi, S., and Orologas, N. (1996). Ultrastructure and supramolecular organization of photosynthetic membranes of some marine red algae. *New Phytol.* 133, 543–551.

Uniacke, J., and Zerges, W. (2007). Photosystem II assembly and repair are differentially localized in *Chlamydomonas*. *Plant Cell* 19, 3640–3654.

Weber, A. P., and Linka, N. (2011). Connecting the plastid: transporters of the plastid envelope and their role in linking plastidial with cytosolic metabolism. *Annu. Rev. Plant Biol.* 62, 53–77.

Weber, F. E., Minestrini, G., Dyer, J. H., Werder, M., Boffelli, D., Compassi, S., Wehrli, E., Thomas, R. M., Schulthess, G., and Hauser, H. (1997). Molecular cloning of a peroxisomal $Ca2^{+}$-dependent member of the mitochondrial carrier superfamily. *Proc. Natl. Acad. Sci. U.S.A.* 94, 8509–8514.

Yin, L., Lundin, B., Bertrand, M., Nurmi, M., Solymosi, K., Kangasjärvi, S., Aro, E. M., Schoefs, B., and Spetea, C. (2010). Role of thylakoid ATP/ADP carrier in photoinhibition and photoprotection of photosystem II in *Arabidopsis*. *Plant Physiol.* 153, 666–677.

Zak, E., Norling, B., Maitra, R., Huang, F., Andersson, B., and Pakrasi, H. B. (2001). The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13443–13448.

Zhang, J., Zhang, Y., and Rosenberg, H. F. (2002). Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat. Genet.* 30, 411–415.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and van Wijk, K. J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3, e1994. doi:10.1371/journal.pone.0001994

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 September 2011; accepted: 17 December 2011; published online: 09 January 2012.

Citation: Spetea C, Pfeil BE and Schoefs B (2012) Phylogenetic analysis of the thylakoid ATP/ADP carrier reveals new insights into its function restricted to green plants. Front. Plant Sci. ***2****:110. doi: 10.3389/fpls.2011.00110*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 18 January 2012
doi: 10.3389/fpls.2012.00002

The plant mitochondrial carrier family: functional and evolutionary aspects

Ilka Haferkamp[1]* and Stephan Schmitz-Esser[2]

[1] *Zelluläre Physiologie/Membrantransport, Technische Universität Kaiserslautern, Kaiserslautern, Germany*
[2] *Institut für Milchhygiene, Veterinärmedizinische Universität Wien, Wien, Austria*

Edited by:
Markus Geisler, University of Fribourg, Switzerland

Reviewed by:
Nicole Linka, Heinrich-Heine Universität Düsseldorf, Germany
Karsten Fischer, University of Tromsø, Norway

****Correspondence:***
Ilka Haferkamp, Biologie, Zelluläre Physiologie/Membrantransport, Technische Universität Kaiserslautern, Erwin-Schrödinger-Str. 22, 67653 Kaiserslautern, Germany.
e-mail: haferk@rhrk.uni-kl.de

Mitochondria play a key role in respiration and energy production and are involved in multiple eukaryotic but also in several plant specific metabolic pathways. Solute carriers in the inner mitochondrial membrane connect the internal metabolism with that of the surrounding cell. Because of their common basic structure, these transport proteins affiliate to the mitochondrial carrier family (MCF). Generally, MCF proteins consist of six membrane spanning helices, exhibit typical conserved domains and appear as homodimers in the native membrane. Although structurally related, MCF proteins catalyze the specific transport of various substrates, such as nucleotides, amino acids, dicarboxylates, cofactors, phosphate or H^+. Recent investigations identified MCF proteins also in several other cellular compartments and therefore their localization and physiological function is not only restricted to mitochondria. MCF proteins are a characteristic feature of eukaryotes and bacterial genomes lack corresponding sequences. Therefore, the evolutionary origin of MCF proteins is most likely associated with the establishment of mitochondria. It is not clear whether the host cell, the symbiont, or the chimerical organism invented the ancient MCF sequence. Here, we try to explain the establishment of different MCF proteins and focus on the characteristics of members from plants, in particular from *Arabidopsis thaliana*.

Keywords: mitochondrial carrier family, MCF, solute carriers, evolution, function, mitochondria, plant, *Arabidopsis thaliana*

ENDOSYMBIOSIS AND THE ESTABLISHMENT OF PLASTIDS AND MITOCHONDRIA

Already in 1883, the botanist A. Schimper observed that the division process of chloroplasts resembles that of bacteria and he stated briefly that plants represent a consortium of two different organisms (Schimper, 1883; Ward, 1883). Mereschkowsky (1905) focused on plastid evolution in his publication "On the nature and origin of the chromatophores in the plant kingdom" and again proposed that chloroplasts are "semi-autonomous," endosymbiotic entities of plant cells. About 20 years later Wallin (1923, 1927) postulated that also mitochondria represent a kind of foreign/bacterial cell that thrives in the eukaryotic host. Many striking observations supported a relatedness of bacteria and mitochondria, and in particular of cyanobacteria and chloroplasts however, the pioneering idea of an endosymbiotic origin of these organelles was disregarded for a long time. The endosymbiotic theory was reinvented 1967 by Lynn Margulis (Sagan, 1967) and gained broader scientific attention in the 1970s and 1980s. Detailed electron microscopical data, microbiological investigations, and most importantly the identification of the mitochondrial and the plastidial genome unequivocally clarified that these organelles originated from formerly free-living bacteria that were taken up by another cell (Gray, 1992; Delwiche, 1999). Deciphering organellar and nuclear genomes demonstrated that mitochondria and plastids each evolved by a single endosymbiotic event. Mitochondria were derived from an alpha-proteobacterium whereas chloroplasts originated from a photosynthetic cyanobacterium (Lang et al., 1997; Delwiche, 1999; Martin et al., 2002; Raven and Allen, 2003; Gray et al., 2004; Rodriguez-Ezpeleta et al., 2005; de Duve, 2007). In contrast to the origin of mitochondria, plastid evolution is rather complicated due to additional endosymbioses (secondary, tertiary endosymbioses), due to secondary loss of the plastid, or due to replacement of existing plastids by new ones (Delwiche, 1999; Keeling, 2010).

Although mitochondria and chloroplasts arose by independent evolutionary processes, both share important similarities concerning reduction, modification, and integration of the symbiont into the context of the host cell and have relict bacterial features in common. Establishment of these organelles was accompanied by massive gene transfer from the symbiont to the host nucleus (Martin and Herrmann, 1998; Kurland and Andersson, 2000). Accordingly, the symbiont became metabolically highly impaired due to the loss of genes (and their corresponding functions) and had to be genetically and physiologically integrated within the host cell. Moreover, plastids and mitochondria are surrounded by two membranes. This feature is traced back to their endosymbiotic ancestry from Gram-negative bacteria. The inner organelle membrane represents the former bacterial plasma membrane and the outer organelle membrane descended from the outer bacterial membrane (Cavalier-Smith, 2000; Gross and Bhattacharya, 2009a). Similar to the situation in Gram-negative bacteria, the outer membrane of mitochondria and plastids harbors mainly β-barrel proteins and allows a rather unselective

passage of several, small molecules. The inner organellar or bacterial membrane constitutes the permeability barrier and harbors carriers mediating a highly specific and selective solute transfer (Cavalier-Smith, 2000; Gross and Bhattacharya, 2009a). Accordingly, bacterial transport systems were already present in the ancient organelle. Because of the genetic reduction the symbiont lost its independence and relied on additional protein and metabolite provision from the host (Martin and Herrmann, 1998; Kurland and Andersson, 2000). Specific transport systems for solute and protein import into the symbiont were required to compensate for its genetic deficits and to functionally embed the organelle into the host. Pre-existing transport systems in the two surrounding membranes had to be adapted or new systems had to be established prior to or at least synchronously with the transfer of symbiotic genes to the nucleus (Gross and Bhattacharya, 2009a,b; Alcock et al., 2010; Bohnsack and Schleiff, 2010). Moreover, targeting of nuclear encoded proteins to the mitochondrion or plastid and their passage through the two membranes had to be established, guaranteed, and controlled. Protein import systems of mitochondria and plastids consist of specialized translocase complexes in the inner and in the outer membrane (Gross and Bhattacharya, 2009a,b; Alcock et al., 2010; Bohnsack and Schleiff, 2010; Lithgow and Schneider, 2010). An N-terminal extension generally characterizes nuclear encoded proteins with plastidial destination and also several mitochondrial proteins are initially synthesized with mitochondrial presequences that serve as matrix-targeting signals (Lithgow, 2000; Bruce, 2001; Patron and Waller, 2007).

MITOCHONDRIA AND MITOCHONDRIA-RELATED ORGANELLES

Mitochondria are the main sites of cellular respiration and ATP supply and thus represent an important and characteristic component of eukaryotes. To date no lineage has been shown to have lost or replaced this organelle and also additional endosymbioses apparently did not occur during mitochondrial evolution. However, mitochondria are not homogenous at all; they exhibit substantial genetic, functional, and morphological differences. Interestingly, certain anaerobic or microaerophilic protists possess "anaerobic mitochondria" or mitochondria-related organelles, called hydrogenosomes or mitosomes (Embley and Martin, 2006; Hackstein et al., 2006). Similar to the aerobic "powerhouses," also "anaerobic" mitochondria and hydrogenosomes are involved in energy production. In the respiratory chain of mitochondria oxygen acts as electron acceptor and water is produced, whereas in hydrogenosomes electrons are transferred to protons resulting in hydrogen generation (Embley and Martin, 2006; Hackstein et al., 2006). "Anaerobic" mitochondria use other electron acceptors than oxygen or protons, like nitrate or fumarate (Embley and Martin, 2006; Hackstein et al., 2006). Mitosomes represent extremely reduced mitochondrial relatives not capable for ATP generation. Accordingly, these mitochondrial remnants rely on ATP import to fuel interior processes (at least iron–sulfur cluster biosynthesis) with energy (Chan et al., 2005; Tsaousis et al., 2008; Williams et al., 2008).

There is ample evidence that mitochondria, hydrogenosomes, and mitosomes evolved from one common facultative anaerobic ancestor (van der Giezen et al., 2002; Voncken et al., 2002; Embley et al., 2003; Embley and Martin, 2006; Hackstein et al., 2006; Williams et al., 2008). However, the nature of the host is uncertain and still debated (de Duve, 2007; Gross and Bhattacharya, 2009a). The hydrogen hypothesis describes metabolic interaction and tight association as a possible reason for the fusion of a facultative aerobic alpha-proteobacterium with an anaerobic methane-producing Archaean (Martin and Müller, 1998). In presence of O_2 alpha-proteobacteria respire organic compounds and produce carbon dioxide and water, whereas under anaerobic conditions they perform fermentation and hydrogen instead of water is delivered. Archaeal bacteria can use the H_2 and CO_2 as sole energy and carbon sources. It is imaginable that ancient methanogenic archaea – due to tight interaction with alpha-proteobacteria – became independent on abiotic H_2 and thus were able to colonize new niches. Gain of H_2 and CO_2 was probably enhanced by enlargement of the contact area between the two species and led to complete surrounding and engulfment of the proteobacterium (Martin and Müller, 1998). During evolution the archaebacterium converted from autotrophy (using H_2 and CO_2) to heterotrophy (using organic molecules from the environment), gained typical eukaryotic features and the alpha-proteobacterium was reduced to an organelle. Absence of O_2 supported the establishment of "early" hydrogenosomes and its presence caused the generation of mitochondria. The conversion of typical aerobic mitochondria into anaerobic organelles and massive reduction to mitochondrial remnants apparently happened several times and might explain why "anaerobic mitochondria," hydrogenosomes, and mitosomes occur in phylogenetically diverse lineages (Tjaden et al., 2004; Embley and Martin, 2006; Hackstein et al., 2006).

The hydrogen hypothesis tightly connects the establishment of the first eukaryote with the establishment of the first mitochondrion. A second hypothesis bases on a more separated, successive evolution of eukaryotes and mitochondria (Cavalier-Smith, 1983). This hypothesis suggests that primitive early branching eukaryotes ("Archaezoa") evolved in a pre-mitochondrial era. These "Archaezoa" harbored an internal membrane system and were able to perform phagocytosis and thus possessed important prerequisites for the capture of the symbiont.

STRUCTURAL FEATURES OF MCF CARRIERS

The mitochondrial carrier family (MCF) is a large family of proteins with about 30 members in yeast and more than 50 in humans and plants (Palmieri et al., 2000b; Picault et al., 2004; Wohlrab, 2006). MCF proteins are highly heterogenous in terms of substrate specificity and transport mode but all possess a molecular mass of about 30–35 kDa, and exhibit an identical basic structure with six transmembrane domains. MCF proteins are composed of three repetitive modules that take a tilted position in the membrane (Kuan and Saier, 1993; Pebay-Peyroula et al., 2003; Nury et al., 2006). The three repeated homologous regions are of about 100 amino acids in length and each repeat comprises two membrane spanning domains (Saraste and Walker, 1982). Furthermore, the odd-numbered transmembrane domains are kinked and characterized by a conserved MCF motif at the kink (Jezek and Jezek, 2003; Nury et al., 2006). Three loops connect the six

transmembrane spanning regions at the matrix side and each loop contains a short amphipathic helical domain with a second conserved MCF motif (**Figure 1**; Nury et al., 2006). The N- and the C-terminal parts as well as the hydrophilic loops between transmembrane helix 2 and 3 and helix 4 and 5 are exposed to the intermembrane space.

Although MCF proteins are described to exhibit a dimeric structure in their native membrane, more recent analyses suggest that the monomer might be one functional entity (Kunji and Crichton, 2010). The crystal structure of the monomeric ADP/ATP carrier (AAC) in complex with the inhibitor carboxyatractyloside represents the overexpanded form of the cytosolic-state (c-state; Pebay-Peyroula et al., 2003; Klingenberg, 2008). It resembles a cup or a basket with the closed bottom directed to the matrix and a wide opening facing the intermembrane space and possesses a channel-like interior structure. Two salt-bridge networks are suggested to act as gates of the carrier (Nury et al., 2006). In the c-state, one network closes the carrier at the matrix side and the open entry faces the cytosol whereas in the matrix-state (m-state), the other network closes the cytosolic part of the carrier and the internal cavity becomes accessible from the matrix. The mitochondria-specific anionic phospholipid cardiolipin resides between two monomers and was shown to play an import role in the biogenesis and stability of MCF proteins (Jiang et al., 2000; Klingenberg, 2009). Furthermore, transport studies demonstrated that cardiolipin addition enhances the activity of the recombinant mitochondrial AAC and also of other MCF proteins (Jiang et al., 2000; Heimpel et al., 2001). It can be assumed that dimer formation occurred early in MCF evolution because different present-day MCF proteins appear as dimers in their native membranes.

EVOLUTION OF MCF PROTEINS

The common basic structure of all present-day MCF proteins suggests that all carriers arose from one single ancestral sequence that was duplicated and differentiated in several rounds. So far, the evolutionary origin of the first MCF protein is not clarified. The threefold repetitive structure of MCF carriers (Kuan and Saier, 1993) most likely reflects also the structure of the first MCF protein. During its evolution, pre-existing gene fragments were probably fused to a basic module encoding a protein with two transmembrane helices (**Figure 1**). Two subsequent gene duplications of this basic module and fusion of the three identical copies probably gave rise to the ancestor of all MCF proteins (pre-MCF protein, **Figure 1**). It is imaginable that the postulated single basic module itself also arose from gene duplication of a fragment encoding a single transmembrane domain (Palmieri et al., 2011). The ancestral MCF protein might have originated prior to, during or after establishment of the mitochondrial organelle. It might have been introduced by the proteobacterial symbiont or by the host (archaebacteria or "Archaezoa") or was newly invented in the consortium. The fact that MCF proteins or sequences that substantially resemble the postulated basic module have never been identified in bacterial species argues against a bacterial origin. Whereas, the presence of MCF proteins in almost all eukaryotes (very few exceptions exist) tightly connects the MCF ancestry with the eukaryotic cell.

Taken the "Archaezoa" hypothesis as a possible evolutionary basis, the first MCF sequence might have been established in the early amitochondriate eukaryote to fulfill a function in a different organelle. In fact, MCF proteins are not only restricted to mitochondria but also present in many other cellular compartments (Bedhomme et al., 2005; Bouvier et al., 2006; Thuswaldner

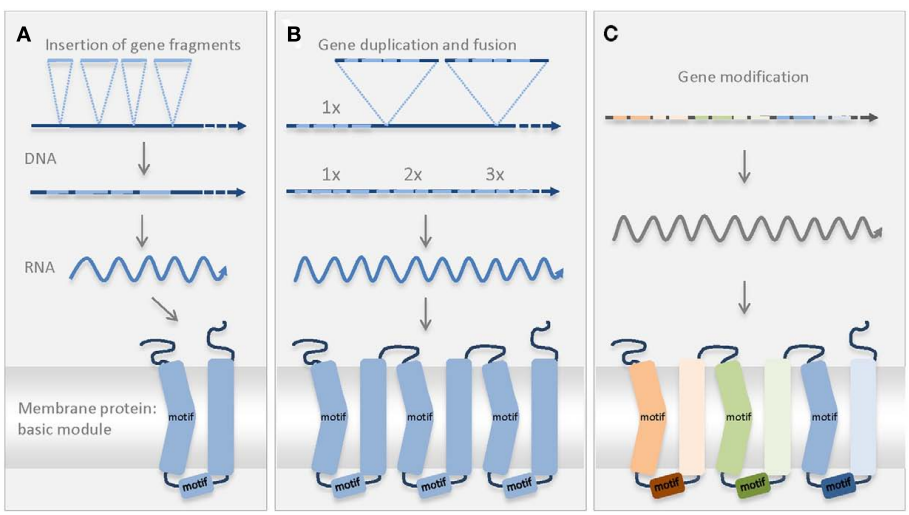


FIGURE 1 | Schematic diagram of the possible scenario explaining the establishment of the first MCF protein. Conserved MCF motives in the secondary structure of the protein are indicated (motif). The odd-numbered transmembrane regions contain one conserved MCF motif (PxD/ExxK/RxK/R) and the proline residue induces the kinked form (Nury et al., 2006). A second MCF motif (D/EGxxxxaK/RG) is located in the helical region of the connecting loop (Nury et al., 2006). **(A)** In the genome of the host or engulfed bacterium a basic MCF sequence originated by fusion of different smaller gene fragments. The new sequence encodes the postulated basic module of the MCF consisting of two transmembrane helices and of a connecting loop with a helical region. **(B)** Twofold duplication of the basic sequence and insertion in the genome resulted in the coding sequence of the pre-MCF protein. **(C)** Gene modification resulted in the establishment of further/different MCF carriers that still exhibit the conserved structure and the MCF motives.

et al., 2007; Kirchberger et al., 2008; Leroch et al., 2008; Linka et al., 2008; Palmieri et al., 2009; Bernhardt et al., 2011; Rieder and Neuhaus, 2011). However, the majority of MCF proteins reside in the inner mitochondrial membrane. Accordingly, their main metabolic role and thus probably also their origin is rather connected with mitochondria. Moreover, organisms that harbor relicts of mitochondria contain only few or no MCF proteins (Chan et al., 2005; Tsaousis et al., 2008; Williams et al., 2008). The important relevance of MCF proteins with respect to mitochondria is further supported by the fact that these proteins represent the only solute carriers in the inner mitochondrial membrane whereas membranes from other organelles harbor a mixture of different transport proteins which belong to different families and apparently originated due to different events. For example, carriers of the inner envelope of the chloroplast were derived from proteins of other organelles, some have clearly cyanobacterial ancestors (endosymbiotic gene transfer), others were acquired by horizontal gene transfer (from other bacteria, such as chlamydial species; Tyra et al., 2007).

It is tempting to speculate that the basic module of MCF proteins was newly created from the scratch by fusion of smaller DNA segments (**Figure 1**) shortly before or during endosymbiosis and that establishment of the first MCF protein drove or at least supported mitochondrial genesis. Gene duplications and modifications were apparently required to obtain diverse MCF proteins with different characteristics. Carriers of the MCF mediating the exchange of adenine nucleotides are a characteristic feature of all present-day mitochondria, of mitochondria-derived hydrogenosomes and were also detected in mitosomes (Chan et al., 2005; Tsaousis et al., 2008; Williams et al., 2008; Klingenberg, 2009). Because of this wide distribution and because of the important role of mitochondria in energy production it can be hypothesized that the first MCF protein was also involved in energy passage (ADP/ATP exchange). However, due to the fact that present-day MCF proteins exhibit a broad range of different substrate spectra (**Figure 2**) one might also assume that the first MCF protein was not a highly specific carrier but rather an "all-rounder" mediating the translocation of various solutes. A supposed relatively broad substrate spectrum of the first MCF protein might have provided an advantageous basis for fast establishment of a subset of different, more specific carriers. Some present-day MCF carriers exhibit relatively broad substrate specificities; although certain substrates are often preferred and others are transported with lower affinity or velocity. This suggests that the carriers binding pocket is not always highly selective and restrictive and that several substrates loosely fit in (Nury et al., 2006; Klingenberg, 2008). Interestingly, recent studies of metazoan MCF carriers showed that the ratio of hydrophobic amino acids at specific helix–helix interfaces increased during their evolution which indicates a trend toward a higher selectivity and to smaller substrates (Gong et al., 2010).

Most eukaryotic organisms possess a large set of MCF proteins and phylogenetic analyses of the corresponding amino acid sequences demonstrate that the various MCF carriers from plants (*Arabidopsis thaliana*), animals (*Homo sapiens*), and fungi (*Saccharomyces cerevisiae*) build independent clades (**Figure 3**; **Figures A1–A5** in Appendix; Palmieri et al., 2011). The branching pattern of these main clades generally does not reflect the affiliation to the different organismic groups but is rather caused by differences in functional properties (Palmieri et al., 2011). For example the cluster containing basic amino acid and carnitine/acylcarnitine carriers is clearly separated from the clusters comprising phosphate carriers (PiC), iron transporters, adenine nucleotide carriers, etc. (**Figure 3**). Almost every functional cluster comprises carriers from animals, yeast, and plants. This fact suggests that functionally different carriers probably evolved by multiple events of gene duplication, sequence alterations, and protein specification before the separation of the three eukaryotic kingdoms. A basic set of MCF carriers with different transport properties, substrate specificities, and localization might have already existed in ancient protists and these carriers were the evolutionary basis for the functional groups that are still required and retained in present-day eukaryotes (**Figure 3**). Multi-gene families are often arranged in physical clusters in the respective genomes. A collective grouping of the respective members is indicative for their establishment by recent gene duplications (Sappl et al., 2004). However, sequences encoding different eukaryotic MCF carriers are generally widely distributed across the different chromosomes and most are separated by a high number of other genes. This pattern also argues against recent duplication events as basis of many MCF proteins.

Some functional groups are not present in all eukaryotic lineages (Palmieri et al., 2011). Succinate/fumarate carriers are apparently absent in animals or yeasts lack uncoupling proteins (UCPs; **Figure A1** in Appendix). The presence of a functional group in two of the three lineages suggests that the remaining lineage lost the respective carrier function because it was compensated by other transporters and/or no more required. The presence of a carrier function in solely one lineage however, might be indicative for its new invention after the separation of the eukaryotes. A mitochondrial GTP/GDP carrier exists only in yeasts and the subgroup of plastidial adenine nucleotide carriers and Brittle1 proteins is restricted to plants (**Figure 3**). In the functional clades, paralogs/isoforms of one species often form a highly related subgroup (**Figures A1–A5** in Appendix). The relation in between these paralogs is more pronounced than that to the corresponding orthologs from other species. The three AACs from *Arabidopsis* for example, form a small sub-cluster distinct from that of the yeast isoforms and from the human orthologs (**Figure A2** in Appendix). This phylogenetic positioning suggests that the additional isoforms evolved by gene duplications independently in the different eukaryotic lineages (Palmieri et al., 2011).

In certain functional clades, plants exhibit a higher number of proven or predicted MCF paralogs. This might indicate that gene duplication and/or gene retention occurred more often in plants. Immotile plants might exhibit additional *mcf* genes that are differentially regulated (expressed) to react more flexible to changing environmental conditions, biotic or abiotic stresses. However, also mobile algae possess a higher number of MCF paralogs than human or yeast, at least in some functional clades (Palmieri et al., 2011). It is tempting to assume that plants require additional MCF proteins because they possess an additional organelle, the plastid. MCF proteins that are directly or indirectly associated with plastid function might have originated by gene duplication of the mitochondrial pendant and rerouting to the plastid, like the

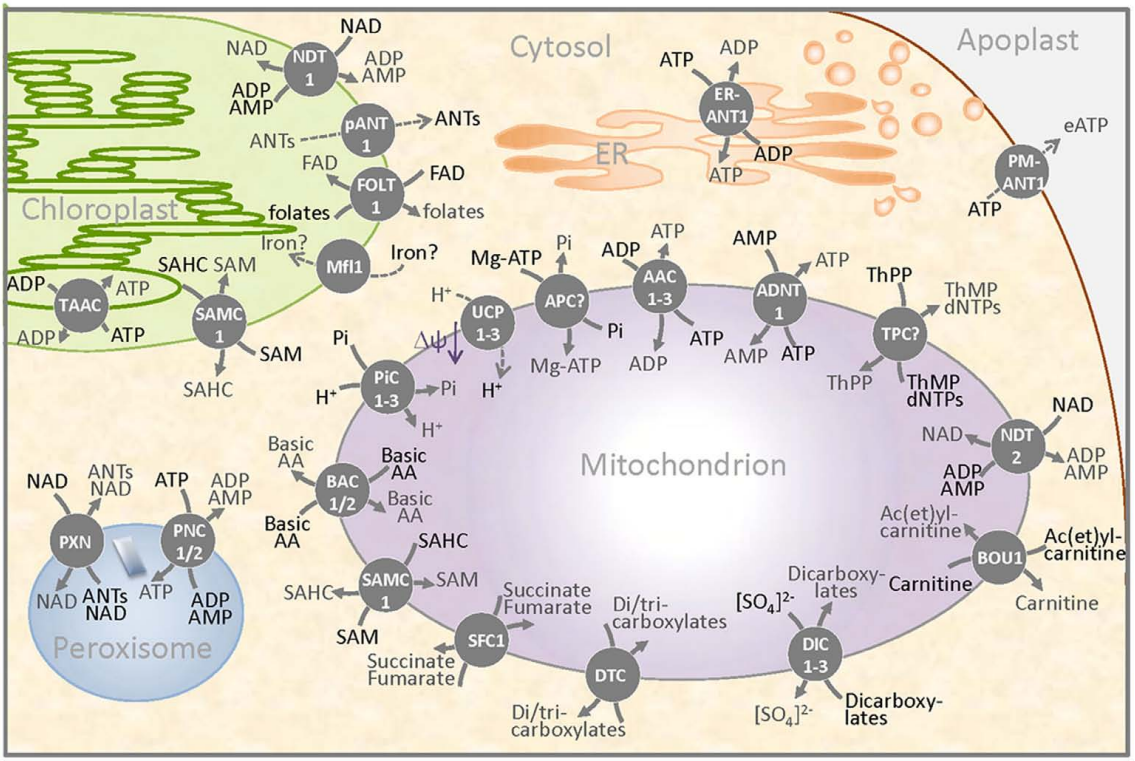


FIGURE 2 | Schematic overview of a plant cell detail presenting the occurrence of MCF carriers in membranes of different compartments. MCF proteins were identified in the inner membrane of mitochondria and plastids, in the thylakoids, in the endoplasmic reticulum (ER), in peroxisomes, and in the plasma membrane (separating the cytosol and the apoplast). Three ADP/ATP carriers (AAC1–3), the ATP/AMP carrier (ADNT1), putative ATP–Mg/Pi carriers (APC?), and the NAD/adenine nucleotide exchanger mediate the passage of ATP, ADP, AMP, or of the adenylated cofactor NAD across the inner mitochondrial membrane. The mitochondrial transport of carboxylates is mediated by three dicarboxylate carriers (DIC1–3), by a dicarboxylate/tricarboxylate carrier (DTC), and by a carrier probably specific for succinate/fumarate exchange. The controlled flux of protons reduces the mitochondrial membrane potential ($\Delta\Psi$) and is catalyzed by two proven and one predicted uncoupling protein (UCP1–3). The transport of the cofactor thiamine pyrophosphate in exchange with thiamine monophosphate or (d)NTPs is probably catalyzed by a postulated mitochondrial thiamine pyrophosphate carrier (TPC?). Basic amino acids leave and enter the mitochondrion via the carriers BAC1 and BAC2 (BAC1–2) and the carnitine/ac(et)ylcarnitine carrier and the *S*-adenosylmethionine carrier (SAMC1) provide carbon moieties to the organelle. The three carriers PiC1–3 allow the net import of phosphate that can act as substrate of ATP synthesis and as counter exchange substrate of other MCF proteins. Plastids also harbor MCF proteins. An adenine nucleotide exporter (pANT1), a NAD/adenine nucleotide exchanger (NDT1), a possible folate carrier (FOLT1), and a Mitoferrin-like protein reside in the inner envelope membrane and provide important substrates and cofactors to the plastid. In addition to its localization in mitochondria, the carrier SAMC1 was also identified in the plastid. ATP provision to the thylakoid and to the ER is probably facilitated by the respective ATP/ADP exchangers – TAAC and ER-ANT1. Peroxisomes contain specific adenine nucleotide carriers (PNC1/2) as well as the NAD transporter PXN. In the plasma membrane, the protein PM-ANT1 might be involved in the provision of extracellular ATP.

plastidial NAD transporter (NDT1) which is a close homolog to the mitochondrial NAD transporter (NDT2; **Figure A3** in Appendix; Palmieri et al., 2009). Furthermore, plastids might have gained MCF proteins also due to dual targeting of previously mitochondrial localized carriers. Recent studies revealed that several MCF proteins exhibit dual targeting and thus might fulfill functions in different organelles, like mitochondria as well as plastids (Agrimi et al., 2004; Bouvier et al., 2006; Palmieri et al., 2006a; Bahaji et al., 2011a,b). The existence of certain plastid-specific MCF proteins suggests that the respective gene duplications and modifications occurred after separation of the plant lineage. It is important to mention that plant specific MCF proteins (e.g., the adenine nucleotide carriers ER-ANT1 and PM-ANT1) do not only comprise plastidial carriers (Leroch et al., 2008; Rieder and Neuhaus, 2011).

THE FUNCTIONAL SUBGROUPS OF THE MCF FROM *ARABIDOPSIS*

Plant and algal genomes contain a high number of MCF coding sequences. In the past decades, the function of more and more carriers was identified; however, the physiological role of the majority of plant MCF proteins is still unclear. Not even one-half of the 58 MCF proteins of the frequently analyzed model plant *Arabidopsis thaliana* are biochemically characterized in detail. In rare cases, the physiological impact of carriers was deduced from metabolic defects of corresponding loss-of-function mutants. Import studies performed with intact yeast or *E. coli* cells expressing the respective carriers, or with the purified, recombinant proteins reconstituted into lipid vesicles, allowed the identification of specific inhibitors and the determination of detailed biochemical properties, like substrate specificities, transport kinetics, modes, and

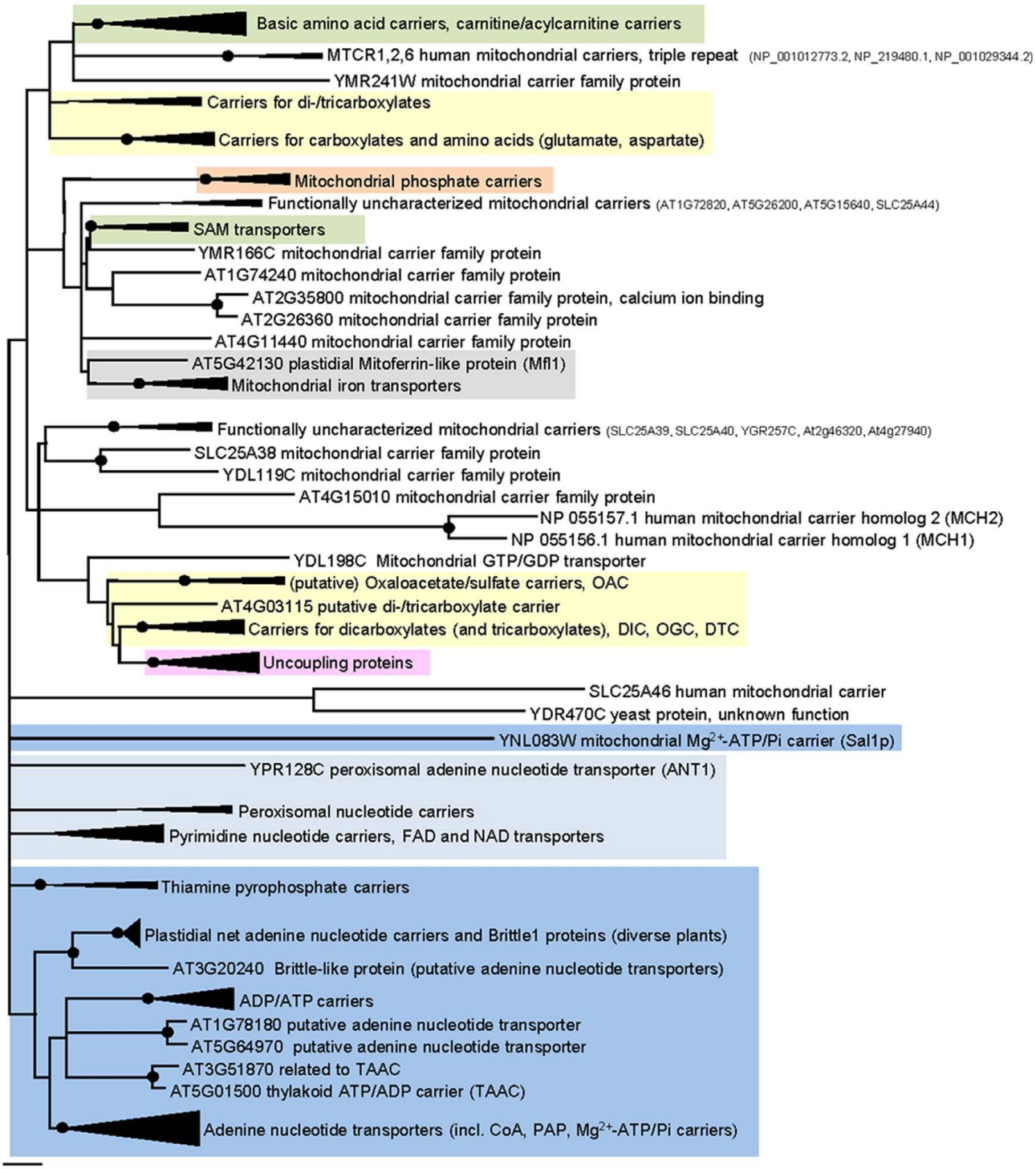


FIGURE 3 | Phylogenetic relationships of MCF proteins. A phylogenetic tree based on amino acid sequences of the MCF proteins from *Homo sapiens*, *Saccharomyces cerevisiae*, and *Arabidopsis thaliana* is shown. The tree was calculated with MEGA5 using the maximum likelihood treeing method and the JTT model of amino acid substitution. The analysis involved 164 amino acid sequences (Table S1 in Supplementary Material). All positions containing gaps and missing data were eliminated. There were a total of 147 positions in the final dataset. Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots. Different functional groups (clusters or sub-clusters) are shaded: basic amino acid carriers, carnitine/acylcarnitine carriers, and SAM transporters in green, carriers for carboxylates, dicarboxylates, tricarboxylates, glutamate, and aspartate in yellow, mitochondrial phosphate carriers in orange, iron transporters in gray, uncoupling proteins in purple, peroxisomal ADP/ATP carriers, pyrimidine nucleotide, FAD, and NAD carriers in very pale blue, adenine nucleotide and thiamine pyrophosphate transporters transporters in blue). The shaded subsets are displayed in detail in **Figures A1–A5** in Appendix.

electrogenicity (Palmieri et al., 2000a; Picault et al., 2004). According to the transport characteristics, four functional subfamilies of the MCF from human, yeast, and *Arabidopsis* were recently proposed (Palmieri et al., 2011). The first subfamily comprises nucleotide and nucleotide derivate transporters. MCF proteins mediating the passage of di- and tricarboxylates or keto acids constitute the second subfamily. Amino acid carriers and carnitine/acylcarnitine carriers are affiliated to a third functional group and transporters with other substrates like UCPs (H^+ passage) or PiCs represent the fourth subdivision. The MCF subfamilies can be further sub-divided into functionally related groups (for example AACs, ATP–Mg/Pi transporters, and NAD transporters, etc., constitute the subfamily of nucleotide and dinucleotide carriers; Palmieri et al., 2011). Each of these MCF groups is characterized by typical triplets in conserved important domains (such as regions extending into the translocation cavity). Subgroups that share some substrates often also share some of these characteristic triplets (Palmieri et al., 2011). The affiliation of a biochemically yet uncharacterized MCF protein to one functional group seems to be a proper basis allowing the prediction of its putative transport substrates and thus of a possible physiological function. In the following, we focus on functional groups and subgroups of our phylogenetic analysis (**Figure 3**; Table S1 in Supplementary Material) and describe the properties of biochemically characterized members from *Arabidopsis*.

MITOCHONDRIAL ADP/ATP TRANSPORTERS AND HIGHLY RELATED PROTEINS

Arabidopsis possesses three typical AACs (AAC1: At3g08580, AAC2: At5g13490, and AAC3: At4g28390; **Figure 2**) involved in mitochondrial energy export (Haferkamp et al., 2002). These carriers are structurally and functionally highly related to their orthologs from animals or yeast, they exhibit identical substrate specificities and inhibitor sensitivities and their transport characteristics are similarly influenced by the membrane potential (Haferkamp et al., 2002). However, in contrast to the mitochondrial AACs from other organisms, plant AACs are characterized by an N-terminal extension (approximately 60 amino acid residues in length) that is not essential for but might support correct targeting to the mitochondrion. In phylogenetic analyzes of MCF carriers AACs cluster together (Palmieri et al., 2011). However, the plant paralogs are separated from fungal and animal AACs (**Figure A2** in Appendix). Interestingly, the *Arabidopsis* genome encodes two additional AAC-related proteins (At5g17400, At5g56450) that share significant homology to the AACs (**Figure A2** in Appendix). Recently, the physiological function of these carriers was clarified.

The first AAC-related carrier (At5g17400) lacks the mitochondrial targeting sequence of plant AACs. This carrier was shown to reside in the plant ER (Leroch et al., 2008) and therefore was called ER-located adenine nucleotide transporter 1 (ER-ANT1; **Figure 2**). When heterologously expressed in *E. coli*, the recombinant protein was able to mediate ATP/ADP exchange and thus ER-ANT1 might fuel ATP dependent processes in the ER lumen, like protein translocation, folding, or maturation. Interestingly, yeast and animals do not possess a definite ER-ANT1 ortholog or AAC-related proteins (with comparably high sequence similarities to mitochondrial AACs) and apparently use another MCF member or a completely different carrier-type for energy provision to the ER. The existence of other/additional ER-located ATP importers is also supported by the fact that ER-ANT1 *Arabidopsis* knock out mutants (although impaired in growth) are still able to survive and to produce fertile seeds (Leroch et al., 2008).

The second functionally investigated AAC-related carrier (PM-ANT1: At5g56450) is N-terminally slightly longer (plus 20 amino acids) than ER-ANT1 but a sequence extension comparable to that of the plant mitochondrial AACs is missing (Rieder and Neuhaus, 2011). The GFP-fusion-protein was targeted to the plant plasma membrane, the recombinant carrier (PM-ANT1) showed ATP and ADP uptake into *E. coli* and probably also accepts a broad range of (desoxy) nucleotides as substrates. Until now the transport mode of PM-ANT1 is not clarified, however, its relatedness to AAC proteins might suggest a function in nucleotide exchange. In animal systems but also in plant cells extracellular ATP (eATP) represents an important signal molecule and the localization of PM-ANT1 in the plasma membrane indicates a possible role in eATP supply and metabolism (**Figure 2**; Rieder and Neuhaus, 2011). PM-ANT1 from *Arabidopsis* is highly expressed in developing pollen and mutant plants reduced in PM-ANT1 activity are impaired in flower development, in particular in anther dehiscence. Therefore, PM-ANT1 was suggested to mediate ATP export specifically from pollen cells and that enhanced eATP acts as a signal received by the anther stomium cells. In addition to the PM-ANT1 also other proteins or mechanisms have to be involved in eATP release, because solely the flowers of the mutant plants show developmental or metabolic defects whereas other plant organs seem to be unaffected (Rieder and Neuhaus, 2011). In fact, several other transport processes were identified to be involved in ATP release from animal but also plant cells. The lack of clear orthologs of ER-ANT1 and PM-ANT1 in yeast or animals (**Figure A2** in Appendix) suggests that these proteins were newly invented or retained to fulfill specific functions in plants.

CARRIERS (POSSIBLY) INVOLVED IN ADENINE NUCLEOTIDE TRANSPORT

Several MCF carriers from *Arabidopsis* (At1g78180, At5g64970, At3g51870 At5g01500, At1g14560, At4g26180, At4g01100, At3g53940, At3g55640, At2g37890, At5g61810, At5g07320, and At5g51050) show lower but still important amino acid similarities to the AACs and the AAC-related proteins and thus can be supposed to mediate adenine nucleotide transport (**Figure 3**; **Figure A2** in Appendix). Two of these carriers were characterized and proven to catalyze the exchange of adenine nucleotides or related compounds.

The first one (At4g01100) is affiliated to a phylogenetic subgroup containing mitochondrial CoA and phospho-adenosyl-phosphate (PAP) transporters from human (including the Graves' disease carrier) and yeast (Fiermonte et al., 2009; Palmieri et al., 2011; **Figure A2** in Appendix). This carrier was named adenine nucleotide transporter (ADNT1: At4g01100) because the reconstituted recombinant protein transported several adenylated (desoxy) nucleotides and adenosine 5-sulfophosphate (APS) in a counter exchange mode, whereas CoA transport was not measurable (Palmieri et al., 2008b). Whether also PAP represents a further substrate of ADNT1 was not investigated. Localization

studies revealed that the ADNT1–GFP-fusion-protein is targeted to mitochondria. Several facts supported the idea that ADNT1 might stimulate mitochondrial energy provision of heterotrophic tissues by facilitating the import of cytosolic AMP in exchange with mitochondrial ATP (**Figure 2**). First, the carrier favors AMP instead of ADP as counter exchange substrate of ATP. Second, the *adnt1* gene is highly expressed in non-photosynthetic as well as in fast growing tissues. Third, roots of ADNT1 loss-of-function mutants are shorter in size than that of wild type plants and are impaired in mitochondrial respiration (Palmieri et al., 2008b).

The second characterized protein with moderate similarities to the AAC proteins (At5g01500) does not clearly affiliate to the AACs or to the CoA and PAP carrier subgroup (**Figure 3**). This carrier was shown to mediate ATP/ADP exchange across the bacterial plasma membrane when heterologously expressed in *E. coli* and localization studies revealed a presence in the thylakoid membrane (**Figure 2**; Thuswaldner et al., 2007). Investigations of TAAC loss-of-function mutants led to the assumption that TAAC mediated ATP provision to the thylakoid lumen plays a role in photoinhibition and photoprotection of photosystem II and in the regulation of the electrochemical H^+ gradient across the thylakoid membrane (Yin et al., 2010). However, it is important to mention that TAAC is not only located in thylakoids but also in the plastid envelope and thus a dual targeting of the carrier is likely (Thuswaldner et al., 2007). Furthermore, increased transcript and protein abundance in growing and heterotrophic tissues points to an alternative/additional function of TAAC in developing and "thylakoid free" plastids. Due to certain important structural similarities to CoA/PAP transporters and ADNT1 it was recently suggested that TAAC accepts further substrates than solely ATP and ADP (Palmieri et al., 2011). The high sequence similarity of one functionally uncharacterized carrier (At3g51870) and TAAC indicates a similar biochemical function (**Figure 3**).

In human and yeast, Mg^{2+}–ATP/Pi carriers (APC, Sal1p) are involved in the replenishment of the mitochondrial purine nucleotide pool in accordance to organellar demands (Fiermonte et al., 2004; Traba et al., 2008, 2009). These carriers catalyze an electroneutral Mg^{2+}–ATP/phosphate exchange and hence, allow net provision of ATP. Three of the remaining putative adenine nucleotide carriers (At5g07320, At5g51050, At5g61810) form a well supported sub-cluster with the human APCs (**Figure A2** in Appendix). Accordingly, these proteins are supposed to catalyze net nucleotide provision however, their biochemical and physiological characteristics have not been analyzed yet.

PLASTIDIAL NET ADENINE NUCLEOTIDE TRANSPORTERS AND BRITTLE PROTEINS

Two further proteins from *Arabidopsis* (At4g32400, At3g20240) exhibit at least moderate similarities to the different groups of adenine nucleotide carriers. These carriers are absent in yeast and human and thus probably originated after the establishment of the plant kingdom (Leroch et al., 2005; Comparot-Moss and Denyer, 2009). One of these carriers (pANT1: At4g32400) was shown to reside in plastids and to mediate a net transport of adenylated nucleotides (Kirchberger et al., 2008). In yeast and human the *de novo* synthesis of nucleotides takes place in the cytosol whereas in plants these pathways are located in the plastid (Zrenner et al., 2006). Accordingly, pANT1 might possibly supply newly synthesized adenine nucleotides to the cytosol (**Figure 2**). The potential *Arabidopsis* APC proteins are supposed to mediate net uptake of these molecules into mitochondria. Recent investigations demonstrated that pANT1 exhibits also a mitochondrial localization (Bahaji et al., 2011b). The N-terminal extension alone possesses the capacity to guide GFP to the plastid whereas the N-terminally truncated (mature carrier) is targeted to mitochondria. Until now, the physiological role of such a net adenine nucleotide transporter in mitochondria is unclear (Bahaji et al., 2011a). In the context of a possible evolution of this protein by gene duplication of a mitochondrial adenine nucleotide carrier sequence and by addition of a plastidial targeting sequence one might speculate that the interior targeting information represents a relict of an ancient mitochondrial localization.

Both carriers (the plastidial adenine nucleotide transporter pANT1 and At3g20240) form a well supported subgroup with functionally related carriers from other plants including, ADP glucose carriers (Brittle1, BT1) from cereals (**Figure 3**; Kirchberger et al., 2007; Comparot-Moss and Denyer, 2009). In most plants, ADP glucose, the carbohydrate precursor for starch production, is synthesized in the plastid. However, culture forms of cereals with high starch levels possess a cytosolic variant of the ADP glucose pyrophosphorylase (AGPase; Denyer et al., 1996; Beckles et al., 2001). Accordingly, in these plants ADP glucose has to be transported from the cytosol into the plastids. The protein BT1 is highly abundant in plastids of maize endosperm (Cao et al., 1995). Furthermore, a maize mutant with a defect in the brittle1 locus is impaired in starch synthesis and possesses enhanced cytosolic ADP glucose levels in immature kernels (Shannon et al., 1998). BT1 shows important similarities to pANT1. Actually, they exhibit 79% identity and 89% similarity in the transport domain and identical substrate contact points (Traba et al., 2011). However, BT1 facilitates the exchange of ADP glucose with ADP or AMP (Kirchberger et al., 2007) whereas pANT1 catalyzes a unidirectional transport of AMP, ADP, and ATP but not of ADP glucose. Accordingly, modification of the substrate spectrum and the change from uniporter to antiporter apparently requires only minor modifications. In fact, the transport mode of some MCF carriers can be changed solely by treatment with reducing agents and hence modifications of thiol groups are sufficient to convert the transport properties (Dierks et al., 1990).

ADP glucose transporters are restricted to monocotyledonous cereals with starch-rich tissues (Comparot-Moss and Denyer, 2009) and form a phylogenetic clade closely related to two clades containing proven or suggested plastidial adenine nucleotide transporters (clade pANT1 and clade pANT2; **Figure A2** in Appendix). Clade one consists of pANT1 proteins from diverse plant lineages whereas the pANT2s clade comprises solely members from cereals and grasses. The monocot-specific genes of pANT2 and the ADP glucose carriers are suggested to have originated by segmental or whole-genome duplication (Comparot-Moss and Denyer, 2009). The first pANT was probably established before the separation of monocotyledonous and dicotyledonous plants and *pant* gene duplication in the monocotyledonous line probably gave rise to the homologs pANT1 and pANT2. In ancestral grasses an additional gene duplication of a pANT gene (maybe pANT2)

occurred and provided the basis for the establishment of the ADP glucose transporter (Comparot-Moss and Denyer, 2009).

POTENTIAL THIAMINE PYROPHOSPHATE CARRIERS IN *ARABIDOPSIS*

Thiamine pyrophosphate is a coenzyme involved in diverse metabolic pathways. In insects, human, and yeast MCF proteins specific for this cofactor (thiamine pyrophosphate carriers, TPC) have been identified (Marobbio et al., 2002; Lindhurst et al., 2006; Iacopetta et al., 2010). Generally, nucleotides and deoxynucleotides represent additional substrates of the TPCs and in fact, the human TPC was initially described to act as a deoxynucleotide carrier (Iacobazzi et al., 2001). TPCs from the different eukaryotic lineages exhibit overlapping substrate spectra, however, they also differ in certain aspects (like in the preference of different nucleotides). The proteins from human and *Drosophila melanogaster* act in a strict antiport mode (Iacopetta et al., 2010) whereas the yeast pendant is also capable for substrate uniport (Marobbio et al., 2002). In contrast to the other characterized TPC the insect homolog does not accept thiamine monophosphate (ThMP) as substrate (Marobbio et al., 2002). The *Arabidopsis* genome encodes two homologous sequences (At3g21390 and At5g48970) similar and related to the TPCs and thus, one or both of the corresponding proteins might act as a transporter of this cofactor in plant mitochondria (**Figure 2**; Palmieri et al., 2011). The TPCs clearly constitute a phylogenetic group distantly related to the groups of adenine nucleotide carriers (**Figure 3**).

PEROXISOMAL MCF PROTEINS, FAD, AND NAD CARRIERS IN *A. THALIANA*

Further phylogenetic groups of MCF carriers also rather distantly related to the mitochondrial adenine nucleotide carriers comprise nucleotide transporters from peroxisomes, mitochondrial pyrimidine nucleotide carriers (PyrNC), NAD/ADP exchangers, and FAD transporters (Palmieri et al., 2011; **Figure 3**; **Figure A3** in Appendix). Until now, in *Arabidopsis*, three MCF proteins with peroxisomal localization (At3g05290, At5g27520, and At2g39970) have been identified. The peroxisomal nucleotide carriers PNC1 (At3g05290) and PNC2 (At5g27520) are functionally similar to their peroxisomal counterpart from yeast (ANT1; Palmieri et al., 2002; Linka et al., 2008). They catalyze the specific exchange of ATP with ADP or AMP (**Figure 2**) and are able to complement a yeast mutant (Δ*ant*) deficient in peroxisomal ATP uptake (Linka et al., 2008). Investigations of transgenic *Arabidopsis* lines demonstrated that PNC1 and PNC2 play an essential role in energy provision to plant peroxisomes and that lack of these proteins significantly impairs fatty acid breakdown and other peroxisomal reactions (Linka et al., 2008).

Nicotinamide adenine dinucleotide (NAD) and its phosphorylated derivative NADP are characteristic factors of redox homeostasis. In addition to their basic role in cellular metabolism these pyridine nucleotides were shown to control intracellular processes, like Ca^{2+} signaling or transcriptional regulation. NAD is synthesized in the cytosol but is required in several compartments and thus has to be delivered to the respective organelles. Recently, two plant homologs (NDT1: At2g47490 and NDT2: At1g25380) of the mitochondrial NAD transporters from yeast (NDT1: YIL006W and NDT2: YEL006W) were identified (Todisco et al., 2006; Palmieri et al., 2009). Complementation of yeast double mutants lacking its mitochondrial NDTs suggested that also the plant MCF homologs mediate NAD import into mitochondria. Import studies with reconstituted plant NDTs demonstrated that both proteins perform a fast NAD counter exchange with various nucleotides – preferentially with ADP or AMP as well as a slow unidirectional NAD transport. However, GFP-based targeting analyzes revealed a mitochondrial localization only for plant NDT2, whereas plant NDT1 apparently resides in plastids (**Figure 2**; Palmieri et al., 2009). Plastids and mitochondria of fast growing tissues require high amounts of NAD. Significant net NAD uptake via the plant NDTs is probably facilitated by interaction with a net adenine nucleotide transporter that provides the counter exchange adenylates to the respective organelle.

Just recently the function of the third peroxisomal MCF protein (At2g399970) was identified. Because of certain similarities and to the peroxisomal adenine nucleotide carriers (PNCs from *Arabidopsis* and ANT1 from yeast) this protein initially was assumed to catalyze ATP transport. However, it was not able to restore ATP import into peroxisomes of the yeast Δ*ant* mutant (Linka et al., 2008; Bernhardt et al., 2011). Biochemical studies demonstrated that this peroxisomal protein (called PXN) catalyzes NAD uptake in exchange with mainly AMP but also with NADH, ADP, and with nicotinate adenine dinucleotide, an intermediate of NAD synthesis (**Figure 2**). The absence of a functional peroxisomal NAD carrier in *Arabidopsis* mutants led to an impaired NAD-dependent β-oxidation and to the accumulation of oil bodies in the seedlings (Bernhardt et al., 2011). Apart from its possible role as a peroxisomal redox shuttle (NAD/NADH exchange) this carrier might also be (directly or indirectly) involved in net uptake of NAD and in the establishment of the peroxisomal NAD pool.

Similar to NAD also FAD operates as an important cofactor in multiple cellular pathways and in different organelles. Moreover, FAD synthesis involves cytosolic reactions, mitochondrial, and plastidial enzymes and thus precursors and intermediates must enter and leave the respective compartments. Complementation studies with a yeast mutant deficient for its mitochondrial folic acid carrier suggested that the *Arabidopsis* MCF protein FOLT1 (At5g66380) acts as a transporter for folic acid and folate derivates (Tzagoloff et al., 1996; Bedhomme et al., 2005). FOLT1 was targeted to plastids when fused to GFP (**Figure 2**). At first glance *Arabidopsis* does not encode a clearly identifiable second FOLT isoform. The absence of a visible phenotype in *Arabidopsis* mutant plants lacking functional FOLT1 however is indicative for the existence of an alternative plastidial folate transport system (Bedhomme et al., 2005). Interestingly, a transporter of the major facilitator superfamily was shown to mediate FAD transport into plastids (Klaus et al., 2005).

Generally, the plastidial NAD and FAD carriers form a phylogenetic sub-cluster together with their functional mitochondrial counterparts as well as with mitochondrial PyrNC from yeast and human (**Figure 3**; **Figure A3** in Appendix). The peroxisomal NAD and adenine nucleotide transporters from different species exhibits low (plant and human) or rather insignificant support (yeast; **Figure 3**; **Figure A3** in Appendix). Nevertheless, the phylogenetic relation of functional homologs from one species is

well supported and argues for their establishment by recent gene duplications.

TRANSPORTERS FOR DI- AND TRICARBONIC ACIDS

Di- and tricarboxylates are important substrates and intermediates of the mitochondrial Krebs cycle and they are required in several other metabolic processes, for example in the *de novo* synthesis of amino acids and nucleotides, in gluconeogenesis or in the glyoxylate cycle.

In *Arabidopsis*, three mitochondrial dicarbonic acid carriers (DIC1: At2g22500, DIC2: At4g24570, and DIC3: At5g09470; **Figure 2**) were recently identified (Palmieri et al., 2008a). DIC1 and DIC2 are highly expressed in different tissues, whereas *dic3*-RNA is hardly detectable (restricted to flower buds and siliques). Similar to their counterparts from yeast or animals (Palmieri et al., 1996; Fiermonte et al., 1998), DIC1–3 are able to transport malate, oxaloacetate, succinate, maleate, malonate, phosphate, sulfate, and thiosulfate; however they exhibit slightly different substrate preferences than their non-plant orthologs (Palmieri et al., 2008a). DIC mediated transport of dicarboxylates in exchange with phosphate or sulfate allows the provision of substrates for mitochondrial reactions/respiration. Whereas malate/oxaloacetate exchange by the DIC proteins might be an important factor of the mitochondrial redox shuttle (metabolic interaction with cytosolic and mitochondrial NAD-dependent malate dehydrogenases).

The dicarboxylate/tricarboxylate carrier from *Arabidopsis* (DTC: At5g19760; **Figure 2**) exhibits high similarities to the DICs and was shown to mediate an electroneutral transport of different single protonated tricarboxylates (citrate, isocitrate, and aconitate) in exchange with a broad variety of unprotonated dicarboxylates (2-oxoglutarate, oxaloacetate, malate, maleate, succinate, and malonate; Picault et al., 2002). Due to these exchange capacities DTC is able to adapt and control the amount of Krebs cycle intermediates according to the cellular metabolism. In fact, in every plant tissue *dtc* transcripts were detectable, however at different levels (Picault et al., 2002). Apart from a proposed basic metabolic role in the Krebs cycle DTC was additionally supposed to be involved in the mitochondrial provision of organic acids for ammonium assimilation (in the plastids) because the *dtc* gene expression was significantly increased by nitrate application (after a nitrate starvation phase; Picault et al., 2002). Interestingly, plants seem to contain "hybrid" carriers when compared to animals and yeast. The DTC from *Arabidopsis* possesses characteristics overlapping with that of the mammalian oxoglutarate and citrate carriers and the DIC proteins from *Arabidopsis* exhibit combined properties of the DICs and oxaloacetate/sulfate carrier (OAC) from yeast (Catoni et al., 2003a,b).

One MCF sequence (At5g01340: SFC1) from *Arabidopsis* forms a subgroup with the succinate/fumarate carrier from yeast (ACR1) but also exhibits moderate similarities to the citrate and oxodicarboxylate carriers from yeast and human (Fiermonte et al., 2001; Palmieri et al., 2001a, 2011; Catoni et al., 2003b). The yeast carrier ACR1 accepts a broad spectrum of di- and tricarboxylates with preference of succinate and fumarate (Palmieri et al., 1997). The plant carrier SFC1 was suggested to act in a similar way (**Figure 2**) and to exhibit overlapping biochemical properties as ACR1 because it complemented the growth phenotype of a yeast mutant lacking this mitochondrial protein (Catoni et al., 2003b). The expression pattern of the *sfc1* gene indicates a possible involvement of the carrier in gluconeogenesis, in particular of young (germinating) seedlings and pollen, as well as in ethanolic fermentation (Catoni et al., 2003b).

Phylogenetic analyses of members of the MCF showed that carriers transporting di- and tricarboxylates form two distinct clusters (**Figure 3**; **Figure A1** in Appendix; Catoni et al., 2003b). One cluster comprises the diverse DIC proteins, the DTC from plant and the human oxoglutarate carrier (OGC). Moreover, plant DIC isoforms are highly homologous and clearly separated from the DICs of yeast and animals and therefore, most likely arose by gene duplication of one common ancestral plant *dic* sequence (**Figure A1** in Appendix; Catoni et al., 2003b). UCPs from different organisms as well as a group consisting of an OAC from yeast, two putative homologs from human and a so far biochemically uncharacterized transporter from *Arabidopsis* (At4g03115) are more distantly related to the DIC proteins but also affiliated to the first cluster (Catoni et al., 2003b; Palmieri et al., 2011). The succinate/fumarate carriers from plant and yeast (SFC1, ACR1) are clearly separated from the cluster containing the DIC proteins and are rather related to citrate carriers and different carbonic acid and amino acid transporters from yeast and human (citrate, oxodicarboxylate oxoglutarate, glutamate, aspartate carriers; Catoni et al., 2003b; Palmieri et al., 2011).

UNCOUPLING PROTEINS IN *ARABIDOPSIS*

Uncoupling proteins are generally known to mediate the passage of protons across the inner mitochondrial membrane (**Figure 2**) in a nucleotide-sensitive and fatty-acid-dependent manner (Klingenberg, 1990; Garlid et al., 1996). UCP activity reduces the proton gradient and hence ATP production, it stimulates respiratory (enzyme) activity and finally leads to heat generation. Accordingly, UCPs were thought to play an exclusive role in thermogenesis of newborn, cold-acclimated, and hibernating animals and to be absent in yeast or plants. However, UCP proteins were identified in diverse plant lineages and UCP1 from *Arabidopsis* (also called PUMP1: At3g54110) and related proteins from other plants were demonstrated to play a role in proton transport at low temperatures or in thermogenesis (Vercesi et al., 1995; Borecky et al., 2001; Ito et al., 2003, 2006). The expression of the remaining two *Arabidopsis* UCPs (UCP2: At5g58970 and UCP3: At1g14140) indicates a different physiological function (Watanabe et al., 1999; Borecky et al., 2006). The activation of plant UCPs by reactive oxygen species and their expression pattern in non-thermogenic plants suggests that the primary role of most plant UCPs is not the organ specific production of heat but rather the regulation of mitochondrial energy synthesis in response to stress factors (Laloi, 1999; Ito et al., 2003, 2006; Borecky et al., 2006; Vercesi et al., 2006). Plant UCP1 and UCP2 form a subgroup related to human UCP1–3 whereas the third UCP from *Arabidopsis* clusters with human UCP4 and UCP4-related proteins and thus was supposed to represent a rather ancient UCP form (**Figure A1** in Appendix; Hanak and Jezek, 2001). As mentioned in the previous chapter UCPs and DICs exhibit important amino acid sequence similarities and phylogenetic relation (**Figure 3**; Borecky et al., 2006). And in fact, prior to their biochemical investigation it was controversially discussed

whether the DIC proteins from plants represent dicarbonic acids carriers or UCPs (Millar and Heazlewood, 2003; Picault et al., 2004; Borecky et al., 2006).

BASIC AMINO ACID AND CARNITINE CARRIERS

In plants, basic amino acids serve as an important nitrogen source and are mobilized from storage proteins during seed germination. Following protein degradation the basic amino acid arginine enters plant mitochondria where it is converted by arginase to urea and ornithine (Goldraij and Polacco, 2000). Ornithine acts as precursor for the synthesis of glutamine, proline, polyamines and alkaloids (and also arginine) and accordingly has to leave the mitochondrion.

In *Arabidopsis* two carriers for basic amino acids (BAC1: At2g33820 and BAC2: At1g79900; **Figure 2**) were identified by complementation of the yeast mutant *arg11* defective in mitochondrial ornithine/arginine transport (carrier: ORT1; Catoni et al., 2003a; Hoyos et al., 2003; Palmieri et al., 2006b). Transport studies with reconstituted recombinant proteins revealed that BAC1 and BAC2 mediate the exchange of the basic amino acids arginine, lysine, ornithine, and histidine (Hoyos et al., 2003; Palmieri et al., 2006b). BAC2 has a broader substrate spectrum than BAC1; it additionally transports citrulline and is less stereospecific because it also allows the passage of D-amino acids at high rates. BAC1 is most likely preferentially involved in the mobilization of storage proteins, whereas BAC2 was shown to contribute to proline accumulation in response to hyperosmotic stress (Hoyos et al., 2003; Toka et al., 2010).

In phylogenetic analyzes, the *Arabidopsis* carriers BAC1, BAC2, and BOU1 ("A BOUT DE SOUFFLE"; At5g46800) are positioned in an MCF subgroup consisting of basic amino acid carriers and carnitine carriers as well as of so far biochemically non-characterized MCF proteins from animals or yeast (**Figure 3**; **Figure A4** in Appendix). The BOU1 protein exhibits important structural similarities to BAC1 and to the ornithine carriers from yeast and human and was suggested to fulfill a role in the degradation of storage lipids during seedling germination. In peroxisomes fatty acids are mobilized and oxidized and subsequently delivered to mitochondria to undergo respiration. BOU1 was hypothesized to catalyze the exchange of ac(et)ylcarnitine and carnitine at the mitochondrial membrane and by this to supply carbons to the mitochondrion (**Figure 2**; Lawand et al., 2002). Because plant peroxisomes export the majority of carbons in form of citrate and not in form of ac(et)ylcarnitine and because an acyl-CoA carnitine acetyltransferase has not yet been identified in plants the primarily supposed function is questionable (Pracharoenwattana et al., 2005). Accordingly, a detailed investigation of the functional characteristics of BOU1 is required to further clarify its physiological role. Just recently, the mammalian carnitine–acylcarnitine carrier-like protein (SLC25A29) with palmitoylcarnitine transporting activity (Sekoguchi et al., 2003) was shown to additionally transport basic amino acids, because it was able to rescue impaired ornithine transport when overexpressed in fibroblasts lacking the ornithine carrier (ORT1; hyperornithinemia–hyperammonemia–homocitrullinuria syndrome; Camacho and Rioseco-Camacho, 2009).

It is important to mention that MCF carriers for the transport of non-basic amino acids are known in animals and fungi but not in plants. Carriers specific for glutamate were identified in human (Fiermonte et al., 2002) and for glutamate and aspartate in human and yeast (Palmieri et al., 2001b; Cavero et al., 2003). These carriers are phylogenetically related to carboxylate carriers from different organisms (**Figure A1** in Appendix).

S-ADENOSYLMETHIONINE CARRIERS

In nearly all cellular methylation-processes the amino acid derivative *S*-adenosylmethionine (SAM) acts as a C1-group donor, however, in specific reactions it also operates as a provider of amino, ribosyl, or aminopropyl groups. SAM is exclusively synthesized in the plant cytosol and has to be imported into mitochondria as well as into plastids where it is required as substrate for, e.g., methylation of organellar DNA, RNA, and proteins (Hanson and Roje, 2001). For re-methylation the resulting product *S*-adenosylhomocysteine (SAHC) has to be exported from the respective organelle into the cytosol. *Arabidopsis* possesses two homologs (SAMC1: At4g39460 and SAMC2: At1g34065) to the SAM transporters from yeast and mammalia (Marobbio et al., 2003; Agrimi et al., 2004). The SAM transporters form a (sub)-cluster clearly separated from other amino acid carriers (**Figure 3**).

The amino acid sequence of both plant proteins possesses an N-terminal extension. The N-terminal sequence of SAMC1 targets GFP to the plastid, whereas the full-length SAMC1 resides in both, chloroplasts and mitochondria (**Figure 2**; Bouvier et al., 2006; Palmieri et al., 2006a). SAMC1 was shown to mediate the counter exchange of SAM and SAHC, to be of higher abundance than SAMC2 and to be expressed in various plant tissues (Bouvier et al., 2006; Palmieri et al., 2006a). Furthermore, *Arabidopsis* plants lacking functional SAMC1 exhibit a dwarf phenotype and are impaired prenyl-lipid metabolism (Bouvier et al., 2006). In comparison to SAMC1 the paralog SAMC2 is rather poorly investigated; its biochemical properties, subcellular localization, and physiological function are still unclear. However, due to high sequence similarities to SAMC1 and phylogenetic relation to SAM carriers from other organisms, a function of SAMC2 in SAM transport seems very likely (**Figure A4** in Appendix).

PHOSPHATE TRANSPORTERS AND MITOFERRIN-LIKE PROTEINS

Apart from ADP also phosphate has to be imported to fuel energy regeneration in mitochondria with substrates. Accordingly, mitochondrial ATP/ADP exchange requires a concerted phosphate transport. PiCs in the inner mitochondrial membrane physiologically interact with AAC proteins, they catalyze a Pi/H^+ symport (or Pi/OH^- antiport) and by this supply phosphate for ATP synthesis (Pratt et al., 1991; Stappen and Krämer, 1994). *Arabidopsis* encodes three proteins (At5g14040: PiC1, At3g48850: PiC2, and At2g17270: PiC3; **Figure 2**) related to mitochondrial PiC from human and yeast. In phylogenetic analyses all PiCs proteins form a (sub)-cluster separated from functionally different MCF proteins (**Figure 3**; Palmieri et al., 2011).

Complementation studies with a yeast mutant deficient in mitochondrial phosphate import confirmed that PiC1 and PiC2 from *Arabidopsis* act as PiCs (Hamel et al., 2004). The function

of the third putative *Arabidopsis* PiC (more distantly related to the PiC1 and 2 plant isoforms, **Figure A5** in Appendix) remains to be experimentally proven. AAC1 and PiC1 are supposed to be the predominant carriers providing the substrates for ATP synthesis because they are expressed in several tissues and show highest abundance in the mitochondrial membrane (Millar and Heazlewood, 2003).

Two so far uncharacterized carriers from *Arabidopsis* (At1g07030 and At2g30160) exhibit high similarities and well supported phylogenetic relation to mitochondrial iron transporters from animals (Mitoferrin 1 and 2) and yeast (MRS3 and 4; Mühlenhoff et al., 2003; Shaw et al., 2006; Froschauer et al., 2009; Paradkar et al., 2009). Therefore, these carriers might represent the plant homologs (**Figure A5** in Appendix). The functional cluster of iron transporters is separated from functionally different MCF carriers but contains an additional member that exhibits lower similarities and a more distant relationship (with low bootstrap support) to the Mitoferrins (**Figure 3**). The corresponding protein is called Mitoferrin-like transporter (AtMfl1: At5g42130), is a component of the inner plastid envelope and investigations of mutant plants suggested its involvement in iron metabolism (Tarantino et al., 2011). Accordingly, AtMfl1 was supposed to mediate iron uptake into the chloroplast (**Figure 2**).

Interestingly, just recently the PyrNC from yeast (Marobbio et al., 2006) that clusters with the NAD and FAD carriers (**Figure 3**; **Figure A3** in Appendix) was also suggested to be involved in mitochondrial iron transport (Yoon et al., 2011).

CONCLUSION

Until now the biochemical properties of many MCF members from plants are not or only poorly investigated. The vast majority of biochemically characterized plant MCF proteins comprises carriers from *Arabidopsis*, whereas carriers from other plant species or algae are clearly underrepresented. However, the investigation of MCF proteins that are restricted to plants and algae allows elucidating plant specific functions and properties. Phylogenetic analyses demonstrate that the various MCF carriers from *Arabidopsis*, human, and yeast build independent clades. Generally, the branching pattern of these main clades reflects differences in their biochemical and physiological properties and MCF proteins with similar functions often form distinct clusters. Several functional MCF groups can be deduced from phylogenetic analysis (Table S1 in Supplementary Material). These functional groups are generally in line with the groups recently proposed by Palmieri et al. (2011). The affiliation to a phylogenetic cluster or sub-cluster (containing functionally characterized carriers from other organisms) might help to get an idea about putative transport substrates of functionally uncharacterized pendants. Proteins of unknown function that possess functionally characterized paralogs or orthologs apparently are the most promising candidates for fast elucidation of their biochemical properties. However, proteins are not always clearly affiliated to a functional MCF cluster or form a cluster with solely biochemically uncharacterized MCF proteins. The identification of the functional properties of these proteins (of no distinct cluster, or of an uncharacterized cluster) is a particular challenge and of significant importance. It will extend our knowledge about the capacities of MCF proteins; it will provide new and probably unexpected insights into the general capacities of MCF proteins and represents an important basis for the characterization of related carriers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Traffic_and_Transport/10.3389/fpls.2012.00002/abstract

REFERENCES

Agrimi, G., Di Noia, M. A., Marobbio, C. M., Fiermonte, G., Lasorsa, F. M., and Palmieri, F. (2004). Identification of the human mitochondrial S-adenosylmethionine transporter: bacterial expression, reconstitution, functional characterization and tissue distribution. *Biochem. J.* 379, 183–190.

Alcock, F., Clements, A., Webb, C., and Lithgow, T. (2010). Evolution. Tinkering inside the organelle. *Science* 327, 649–650.

Bahaji, A., Munoz, F. J., Ovecka, M., Baroja-Fernandez, E., Montero, M., Li, J., Hidalgo, M., Almagro, G., Sesma, M. T., Ezquer, I., and Pozueta-Romero, J. (2011a). Specific delivery to mitochondria of AtBT1 complements the aberrant growth and sterility phenotype of homozygous Atbt1 *Arabidopsis* mutants. *Plant J.* 68, 1115–1121.

Bahaji, A., Ovecka, M., Barany, I., Risueno, M. C., Munoz, F. J., Baroja-Fernandez, E., Montero, M., Li, J., Hidalgo, M., Sesma, M. T., Ezquer, I., Testillano, P. S., and Pozueta-Romero, J. (2011b). Dual targeting to mitochondria and plastids of AtBT1 and ZmBT1, two members of the mitochondrial carrier family. *Plant Cell Physiol.* 52, 597–609.

Beckles, D. M., Smith, A. M., and ap Rees, T. (2001). A cytosolic ADP-glucose pyrophosphorylase is a feature of graminaceous endosperms, but not of other starch-storing organs. *Plant Physiol.* 125, 818–827.

Bedhomme, M., Hoffmann, M., McCarthy, E. A., Gambonnet, B., Moran, R. G., Rebeille, F., and Ravanel, S. (2005). Folate metabolism in plants: an *Arabidopsis* homolog of the mammalian mitochondrial folate transporter mediates folate import into chloroplasts. *J. Biol. Chem.* 280, 34823–34831.

Bernhardt, K., Wilkinson, S., Weber, A. P., and Linka, N. (2011). A peroxisomal carrier delivers NAD^+ and contributes to optimal fatty acid degradation during storage oil mobilisation. *Plant J.* 69, 1–13.

Bohnsack, M. T., and Schleiff, E. (2010). The evolution of protein targeting and translocation systems. *Biochim. Biophys. Acta* 1803, 1115–1130.

Borecky, J., Maia, I. G., Costa, A. D., Jezek, P., Chaimovich, H., de Andrade, P. B., Vercesi, A. E., and Arruda, P. (2001). Functional reconstitution of *Arabidopsis thaliana* plant uncoupling mitochondrial protein (AtPUMP1) expressed in *Escherichia coli*. *FEBS Lett.* 505, 240–244.

Borecky, J., Nogueira, F. T., de Oliveira, K. A., Maia, I. G., Vercesi, A. E., and Arruda, P. (2006). The plant energy-dissipating mitochondrial systems: depicting the genomic structure and the expression profiles of the gene families of uncoupling protein and alternative oxidase in monocots and dicots. *J. Exp. Bot.* 57, 849–864.

Bouvier, F., Linka, N., Isner, J. C., Mutterer, J., Weber, A. P., and Camara, B. (2006). *Arabidopsis* SAMT1 defines a plastid transporter regulating plastid biogenesis and plant development. *Plant Cell* 18, 3088–3105.

Bruce, B. D. (2001). The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochim. Biophys. Acta* 1541, 2–21.

Camacho, J. A., and Rioseco-Camacho, N. (2009). The human and mouse SLC25A29 mitochondrial transporters rescue the deficient ornithine metabolism in fibroblasts of patients with the hyperornithinemia-hyperammonemia-homocitrullinuria (HHH) syndrome. *Pediatr. Res.* 66, 35–41.

Cao, H., Sullivan, T. D., Boyer, C. D., and Shannon, J. C. (1995). Bt1, a structural gene for the major 39-44 kDa amyloplast membrane polypeptides. *Physiol. Plant* 95, 176–186.

Catoni, E., Desimone, M., Hilpert, M., Wipf, D., Kunze, R., Schneider, A., Flügge, U. I., Schumacher, K., and Frommer, W. B. (2003a). Expression pattern of a nuclear encoded mitochondrial arginine-ornithine translocator gene from *Arabidopsis*. *BMC Plant Biol.* 3, 1. doi:10.1186/1471-2229-3-1

Catoni, E., Schwab, R., Hilpert, M., Desimone, M., Schwacke, R., Flügge, U. I., Schumacher, K., and Frommer, W. B. (2003b). Identification of an *Arabidopsis* mitochondrial succinate-fumarate translocator. *FEBS Lett.* 534, 87–92.

Cavalier-Smith, T. (1983). "A 6 kingdom classification and a unified phylogeny," in *Endocytobiology II*, eds W. Schwemmler and H. E. A. Schenk (Berlin: De Gruyter), 1027–1034.

Cavalier-Smith, T. (2000). Membrane heredity and early chloroplast evolution. *Trends Plant Sci.* 5, 174–182.

Cavero, S., Vozza, A., del Arco, A., Palmieri, L., Villa, A., Blanco, E., Runswick, M. J., Walker, J. E., Cerdan, S., Palmieri, F., and Satrustegui, J. (2003). Identification and metabolic role of the mitochondrial aspartate-glutamate transporter in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 50, 1257–1269.

Chan, K. W., Slotboom, D. J., Cox, S., Embley, T. M., Fabre, O., van der Giezen, M., Harding, M., Horner, D. S., Kunji, E. R., Leon-Avila, G., and Tovar, J. (2005). A novel ADP/ATP transporter in the mitosome of the microaerophilic human parasite *Entamoeba histolytica*. *Curr. Biol.* 15, 737–742.

Comparot-Moss, S., and Denyer, K. (2009). The evolution of the starch biosynthetic pathway in cereals and other grasses. *J. Exp. Bot.* 60, 2481–2492.

de Duve, C. (2007). The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet.* 8, 395–403.

Delwiche, C. F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* 154, 164–177.

Denyer, K., Dunlap, F., Thorbjørnsen, T., Keeling, P., and Smith, A. E. (1996). The major form of ADP-glucose pyrophosphorylase in maize endosperm is extra-plastidial. *Plant Physiol.* 112, 779–785.

Dierks, T., Salentin, A., Heberger, C., and Krämer, R. (1990). The mitochondrial asparate/glutamate and ADT/ATP carrier switch from obligate counterexchange to unidirectional transport after modification by SH-reagents. *Biochim. Biophys. Acta* 1028, 268–280.

Embley, T. M., and Martin, W. (2006). Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630.

Embley, T. M., van der Giezen, M., Horner, D. S., Dyal, P. L., Bell, S., and Foster, P. G. (2003). Hydrogenosomes, mitochondria and early eukaryotic evolution. *IUBMB Life* 55, 387–395.

Fiermonte, G., De Leonardis, F., Todisco, S., Palmieri, L., Lasorsa, F. M., and Palmieri, F. (2004). Identification of the mitochondrial ATP-Mg/Pi transporter. Bacterial expression, reconstitution, functional characterization, and tissue distribution. *J. Biol. Chem.* 279, 30722–30730.

Fiermonte, G., Dolce, V., Palmieri, L., Ventura, M., Runswick, M. J., Palmieri, F., and Walker, J. E. (2001). Identification of the human mitochondrial oxodicarboxylate carrier. Bacterial expression, reconstitution, functional characterization, tissue distribution, and chromosomal location. *J. Biol. Chem.* 276, 8225–8230.

Fiermonte, G., Palmieri, L., Dolce, V., Lasorsa, F. M., Palmieri, F., Runswick, M. J., and Walker, J. E. (1998). The sequence, bacterial expression, and functional reconstitution of the rat mitochondrial dicarboxylate transporter cloned via distant homologs in yeast and *Caenorhabditis elegans*. *J. Biol. Chem.* 273, 24754–24759.

Fiermonte, G., Palmieri, L., Todisco, S., Agrimi, G., Palmieri, F., and Walker, J. E. (2002). Identification of the mitochondrial glutamate transporter. Bacterial expression, reconstitution, functional characterization, and tissue distribution of two human isoforms. *J. Biol. Chem.* 277, 19289–19294.

Fiermonte, G., Paradies, E., Todisco, S., Marobbio, C. M., and Palmieri, F. (2009). A novel member of solute carrier family 25 (SLC25A42) is a transporter of coenzyme A and adenosine 3′,5′-diphosphate in human mitochondria. *J. Biol. Chem.* 284, 18152–18159.

Froschauer, E. M., Schweyen, R. J., and Wiesenberger, G. (2009). The yeast mitochondrial carrier proteins Mrs3p/Mrs4p mediate iron transport across the inner mitochondrial membrane. *Biochim. Biophys. Acta* 1788, 1044–1050.

Garlid, K. D., Orosz, D. E., Modriansky, M., Vassanelli, S., and Jezek, P. (1996). On the mechanism of fatty acid-induced proton transport by mitochondrial uncoupling protein. *J. Biol. Chem.* 271, 2615–2620.

Goldraij, A., and Polacco, J. C. (2000). Arginine degradation by arginase in mitochondria of soybean seedling cotyledons. *Planta* 210, 652–658.

Gong, M., Li, J., Wang, M., Wang, J., Zen, K., and Zhang, C. Y. (2010). The evolutionary trajectory of mitochondrial carrier family during metazoan evolution. *BMC Evol. Biol.* 10, 282. doi:10.1186/1471-2148-10-282

Gray, M. W. (1992). The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* 141, 233–357.

Gray, M. W., Lang, B. F., and Burger, G. (2004). Mitochondria of protists. *Annu. Rev. Genet.* 38, 477–524.

Gross, J., and Bhattacharya, D. (2009a). Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. *Nat. Rev. Genet.* 10, 495–505.

Gross, J., and Bhattacharya, D. (2009b). Revaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci.* 14, 13–20.

Hackstein, J. H., Tjaden, J., and Huynen, M. (2006). Mitochondria, hydrogenosomes and mitosomes: products of evolutionary tinkering! *Curr. Genet.* 50, 225–245.

Haferkamp, I., Hackstein, J. H., Voncken, F. G., Schmit, G., and Tjaden, J. (2002). Functional integration of mitochondrial and hydrogenosomal ADP/ATP carriers in the *Escherichia coli* membrane reveals different biochemical characteristics for plants, mammals and anaerobic chytrids. *Eur. J. Biochem.* 269, 3172–3181.

Hamel, P., Saint-Georges, Y., de Pinto, B., Lachacinski, N., Altamura, N., and Dujardin, G. (2004). Redundancy in the function of mitochondrial phosphate transport in *Saccharomyces cerevisiae* and *Arabidopsis thaliana*. *Mol. Microbiol.* 51, 307–317.

Hanak, P., and Jezek, P. (2001). Mitochondrial uncoupling proteins and phylogenesis – UCP4 as the ancestral uncoupling protein. *FEBS Lett.* 495, 137–141.

Hanson, A. D., and Roje, S. (2001). One-carbon metabolism in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 119–137.

Heimpel, S., Basset, G., Odoy, S., and Klingenberg, M. (2001). Expression of the mitochondrial ADP/ATP carrier in *Escherichia coli*. Renaturation, reconstitution, and the effect of mutations on 10 positive residues. *J. Biol. Chem.* 276, 11499–11506.

Hoyos, M. E., Palmieri, L., Wertin, T., Arrigoni, R., Polacco, J. C., and Palmieri, F. (2003). Identification of a mitochondrial transporter for basic amino acids in *Arabidopsis thaliana* by functional reconstitution into liposomes and complementation in yeast. *Plant J.* 33, 1027–1035.

Iacobazzi, V., Ventura, M., Fiermonte, G., Prezioso, G., Rocchi, M., and Palmieri, F. (2001). Genomic organization and mapping of the gene (SLC25A19) encoding the human mitochondrial deoxynucleotide carrier (DNC). *Cytogenet. Cell Genet.* 93, 40–42.

Iacopetta, D., Carrisi, C., De Filippis, G., Calcagnile, V. M., Cappello, A. R., Chimento, A., Curcio, R., Santoro, A., Vozza, A., Dolce, V., Palmieri, F., and Capobianco, L. (2010). The biochemical properties of the mitochondrial thiamine pyrophosphate carrier from *Drosophila melanogaster*. *FEBS J.* 277, 1172–1181.

Ito, K., Abe, Y., Johnston, S. D., and Seymour, R. S. (2003). Ubiquitous expression of a gene encoding for uncoupling protein isolated from the thermogenic inflorescence of the dead horse arum *Helicodiceros muscivorus*. *J. Exp. Bot.* 54, 1113–1114.

Ito, K., Matsukawa, K., and Kato, Y. (2006). Functional analysis of skunk cabbage SfUCPB, a unique uncoupling protein lacking the fifth transmembrane domain, in yeast cells. *Biochem. Biophys. Res. Commun.* 349, 383–390.

Jezek, P., and Jezek, J. (2003). Sequence anatomy of mitochondrial anion carriers. *FEBS Lett.* 534, 15–25.

Jiang, F., Ryan, M. T., Schlame, M., Zhao, M., Gu, Z., Klingenberg, M., Pfanner, N., and Greenberg, M. L. (2000). Absence of cardiolipin in the crd1 null mutant results in decreased mitochondrial membrane potential and reduced mitochondrial function. *J. Biol. Chem.* 275, 22387–22394.

Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 729–748.

Kirchberger, S., Leroch, M., Huynen, M. A., Wahl, M., Neuhaus, H. E., and Tjaden, J. (2007). Molecular and biochemical analysis of the plastidic ADP-glucose transporter (ZmBT1) from *Zea mays*. *J. Biol. Chem.* 282, 22481–22491.

Kirchberger, S., Tjaden, J., and Neuhaus, H. E. (2008). Characterization of the *Arabidopsis* Brittle1 transport protein and impact of reduced activity on plant metabolism. *Plant J.* 56, 51–63.

Klaus, S. M., Kunji, E. R., Bozzo, G. G., Noiriel, A., de la Garza, R. D., Basset, G. J., Ravanel, S., Rébeillé, F., Gregory, J. F. III, and Hanson, A. D. (2005). Higher plant plastids and cyanobacteria have folate carriers related to those of trypanosomatids. *J. Biol. Chem.* 280, 38457–38463.

Klingenberg, M. (1990). Mechanism and evolution of the uncoupling protein of brown adipose tissue. *Trends Biochem. Sci.* 15, 108–112.

Klingenberg, M. (2008). The ADP and ATP transport in mitochondria and

its carrier. *Biochim. Biophys. Acta.* 1778, 1978–2021.

Klingenberg, M. (2009). Cardiolipin and mitochondrial carriers. *Biochim. Biophys. Acta* 1788, 2048–2058.

Kuan, J., and Saier, M. H. Jr. (1993). The mitochondrial carrier family of transport proteins: structural, functional, and evolutionary relationships. *Crit. Rev. Biochem. Mol. Biol.* 28, 209–233.

Kunji, E. R., and Crichton, P. G. (2010). Mitochondrial carriers function as monomers. *Biochim. Biophys. Acta* 1797, 817–831.

Kurland, C. G., and Andersson, S. G. (2000). Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* 64, 786–820.

Laloi, M. (1999). Plant mitochondrial carriers: an overview. *Cell. Mol. Life Sci.* 56, 918–944.

Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M. W. (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387, 493–497.

Lawand, S., Dorne, A. J., Long, D., Coupland, G., Mache, R., and Carol, P. (2002). *Arabidopsis* A BOUT DE SOUFFLE, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. *Plant Cell* 14, 2161–2173.

Leroch, M., Kirchberger, S., Haferkamp, I., Wahl, M., Neuhaus, H. E., and Tjaden, J. (2005). Identification and characterization of a novel plastidic adenine nucleotide uniporter from *Solanum tuberosum*. *J. Biol. Chem.* 280, 17992–18000.

Leroch, M., Neuhaus, H. E., Kirchberger, S., Zimmermann, S., Melzer, M., Gerhold, J., and Tjaden, J. (2008). Identification of a novel adenine nucleotide transporter in the endoplasmic reticulum of *Arabidopsis*. *Plant Cell* 20, 438–451.

Lindhurst, M. J., Fiermonte, G., Song, S., Struys, E., De Leonardis, F., Schwartzberg, P. L., Chen, A., Castegna, A., Verhoeven, N., Mathews, C. K., Palmieri, F., and Biesecker, L. G. (2006). Knockout of Slc25a19 causes mitochondrial thiamine pyrophosphate depletion, embryonic lethality, CNS malformations, and anemia. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15927–15932.

Linka, N., Theodoulou, F. L., Haslam, R. P., Linka, M., Napier, J. A., Neuhaus, H. E., and Weber, A. P. (2008). Peroxisomal ATP import is essential for seedling development in *Arabidopsis thaliana*. *Plant Cell* 20, 3241–3257.

Lithgow, T. (2000). Targeting of proteins to mitochondria. *FEBS Lett.* 476, 22–26.

Lithgow, T., and Schneider, A. (2010). Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 799–817.

Marobbio, C. M., Agrimi, G., Lasorsa, F. M., and Palmieri, F. (2003). Identification and functional reconstitution of yeast mitochondrial carrier for S-adenosylmethionine. *EMBO J.* 22, 5975–5982.

Marobbio, C. M., Di Noia, M. A., and Palmieri, F. (2006). Identification of a mitochondrial transporter for pyrimidine nucleotides in *Saccharomyces cerevisiae*: bacterial expression, reconstitution and functional characterization. *Biochem. J.* 393, 441–446.

Marobbio, C. M., Vozza, A., Harding, M., Bisaccia, F., Palmieri, F., and Walker, J. E. (2002). Identification and reconstitution of the yeast mitochondrial transporter for thiamine pyrophosphate. *EMBO J.* 21, 5653–5661.

Martin, W., and Herrmann, R. G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* 118, 9–17.

Martin, W., and Müller, M. (1998). The hydrogen hypothesis for the first eucaryote. *Nature* 392, 37–41.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12246–12251.

Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralbl.* 25, 593–596. [English translation in Martin, W., Kowallik, K. V. (1999). Annotated English translation of Mereschkowsky's 1905 paper Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Eur. J. Phycol.* 34, 287–295].

Millar, A. H., and Heazlewood, J. L. (2003). Genomic and proteomic analysis of mitochondrial carrier proteins in *Arabidopsis*. *Plant Physiol.* 131, 443–453.

Mühlenhoff, U., Stadler, J. A., Richhardt, N., Seubert, A., Eickhorst, T., Schweyen, R. J., Lill, R., and Wiesenberger, G. (2003). A specific role of the yeast mitochondrial carriers MRS3/4p in mitochondrial iron acquisition under iron-limiting conditions. *J. Biol. Chem.* 278, 40612–40620.

Nury, H., Dahout-Gonzalez, C., Trezeguet, V., Lauquin, G. J., Brandolin, G., and Pebay-Peyroula, E. (2006). Relations between structure and function of the mitochondrial ADP/ATP carrier. *Annu. Rev. Biochem.* 75, 713–741.

Palmieri, F., Pierri, C. L., De Grassi, A., Nunes-Nesi, A., and Fernie, A. R. (2011). Evolution, structure and function of mitochondrial carriers: a review with new insights. *Plant J.* 66, 161–181.

Palmieri, F., Rieder, B., Ventrella, A., Blanco, E., Do, P. T., Nunes-Nesi, A., Trauth, A. U., Fiermonte, G., Tjaden, J., Agrimi, G., Kirchberger, S., Paradies, E., Fernie, A. R., and Neuhaus, H. E. (2009). Molecular identification and functional characterization of *Arabidopsis thaliana* mitochondrial and chloroplastic NAD^+ carrier proteins. *J. Biol. Chem.* 284, 31249–31259.

Palmieri, L., Agrimi, G., Runswick, M. J., Fearnley, I. M., Palmieri, F., and Walker, J. E. (2001a). Identification in *Saccharomyces cerevisiae* of two isoforms of a novel mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate. *J. Biol. Chem.* 276, 1916–1922.

Palmieri, L., Pardo, B., Lasorsa, F. M., del Arco, A., Kobayashi, K., Iijima, M., Runswick, M. J., Walker, J. E., Saheki, T., Satrustegui, J., and Palmieri, F. (2001b). Citrin and aralar1 are Ca^{2+}-stimulated aspartate/glutamate transporters in mitochondria. *EMBO J.* 20, 5060–5069.

Palmieri, L., Arrigoni, R., Blanco, E., Carrari, F., Zanor, M. I., Studart-Guimaraes, C., Fernie, A. R., and Palmieri, F. (2006a). Molecular identification of an *Arabidopsis* S-adenosylmethionine transporter. Analysis of organ distribution, bacterial expression, reconstitution into liposomes, and functional characterization. *Plant Physiol.* 142, 855–865.

Palmieri, L., Todd, C. D., Arrigoni, R., Hoyos, M. E., Santoro, A., Polacco, J. C., and Palmieri, F. (2006b). *Arabidopsis* mitochondria have two basic amino acid transporters with partially overlapping specificities and differential expression in seedling development. *Biochim. Biophys. Acta* 1757, 1277–1283.

Palmieri, L., Lasorsa, F. M., de Palma, A., Palmieri, F., Runswick, M. J., and Walker, J. E. (1997). Identification of the yeast ACR1 gene product as a succinate-fumarate transporter essential for growth on ethanol or acetate. *FEBS Lett.* 417, 114–118.

Palmieri, L., Lasorsa, F. M., Vozza, A., Agrimi, G., Fiermonte, G., Runswick, M. J., Walker, J. E., and Palmieri, F. (2000a). Identification and functions of new transporters in yeast mitochondria. *Biochim. Biophys. Acta* 1459, 363–369.

Palmieri, L., Runswick, M. J., Fiermonte, G., Walker, J. E., and Palmieri, F. (2000b). Yeast mitochondrial carriers: bacterial expression, biochemical identification and metabolic significance. *J. Bioenerg. Biomembr.* 32, 67–77.

Palmieri, L., Palmieri, F., Runswick, M. J., and Walker, J. E. (1996). Identification by bacterial expression and functional reconstitution of the yeast genomic sequence encoding the mitochondrial dicarboxylate carrier protein. *FEBS Lett.* 399, 299–302.

Palmieri, L., Picault, N., Arrigoni, R., Besin, E., Palmieri, F., and Hodges, M. (2008a). Molecular identification of three *Arabidopsis thaliana* mitochondrial dicarboxylate carrier isoforms: organ distribution, bacterial expression, reconstitution into liposomes and functional characterization. *Biochem. J.* 410, 621–629.

Palmieri, L., Santoro, A., Carrari, F., Blanco, E., Nunes-Nesi, A., Arrigoni, R., Genchi, F., Fernie, A. R., and Palmieri, F. (2008b). Identification and characterization of ADNT1, a novel mitochondrial adenine nucleotide transporter from *Arabidopsis*. *Plant Physiol.* 148, 1797–1808.

Palmieri, L., Rottensteiner, H., Girzalsky, W., Scarcia, P., Palmieri, F., and Erdmann, R. (2002). Identification and functional reconstitution of the yeast peroxisomal adenine nucleotide transporter. *EMBO J.* 20, 5049–5059.

Paradkar, P. N., Zumbrennen, K. B., Paw, B. H., Ward, D. M., and Kaplan, J. (2009). Regulation of mitochondrial iron import through differential turnover of mitoferrin 1 and mitoferrin 2. *Mol. Cell. Biol.* 29, 1007–1016.

Patron, N. J., and Waller, R. F. (2007). Transit peptide diversity and divergence: a global analysis of plastid targeting signals. *Bioessays* 29, 1048–1058.

Pebay-Peyroula, E., hout-Gonzalez, C., Kahn, R., Trezeguet, V., Lauquin, G. J., and Brandolin, G. (2003). Structure of mitochondrial ADP/ATP

carrier in complex with carboxyatractyloside. *Nature* 426, 39–44.
Picault, N., Hodges, M., Palmieri, L., and Palmieri, F. (2004). The growing family of mitochondrial carriers in *Arabidopsis*. *Trends Plant Sci.* 9, 138–146.
Picault, N., Palmieri, L., Pisano, I., Hodges, M., and Palmieri, F. (2002). Identification of a novel transporter for dicarboxylates and tricarboxylates in plant mitochondria. Bacterial expression, reconstitution, functional characterization, and tissue distribution. *J. Biol. Chem.* 277, 24204–24211.
Pracharoenwattana, I., Cornah, J. E., and Smith, S. M. (2005). *Arabidopsis* peroxisomal citrate synthase is required for fatty acid respiration and seed germination. *Plant Cell* 17, 2037–2048.
Pratt, R. D., Ferreira, G. C., and Pedersen, P. L. (1991). Mitochondrial phosphate transport. Import of the H^+/Pi symporter and role of the presequence. *J. Biol. Chem.* 266, 1276–1280.
Raven, J. A., and Allen, J. F. (2003). Genomics and chloroplast evolution: what did cyanobacteria do for plants? *Genome Biol.* 4, 209.
Rieder, B., and Neuhaus, H. E. (2011). Identification of an *Arabidopsis* plasma membrane-located ATP transporter important for anther development. *Plant Cell* 23, 1932–1944.
Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S. C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H. J., Philippe, H., and Lang, B. F. (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325–1330.
Sagan, L. (1967). On the origin of mitosing cells. *J. NIH Res.* 5, 65–72.
Sappl, P. G., Heazlewood, J. L., and Millar, A. H. (2004). Untangling multi-gene families in plants by integrating proteomics into functional genomics. *Phytochemistry* 65, 1517–1530.
Saraste, M., and Walker, J. E. (1982). Internal sequence repeats and the path of polypeptide in mitochondrial ADP/ATP translocase. *FEBS Lett.* 144, 250–254.
Schimper, A. F. W. (1883). Über die Entwicklung der Chlorophyllkörner und Farbkörper. *Bot. Zeitung* 41, 105–162. [For an English publication concerning this topic see Ward, H. M. (1883)].
Sekoguchi, E., Sato, N., Yasui, A., Fukada, S., Nimura, Y., Aburatani, H., Ikeda, K., and Matsuura, A. (2003). A novel mitochondrial carnitine-acylcarnitine translocase induced by partial hepatectomy and fasting. *J. Biol. Chem.* 278, 38796–38802.
Shannon, J. C., Pien, F. M., Cao, H., and Liu, K. C. (1998). Brittle-1, an adenylate translocator, facilitates transfer of extraplastidial synthesized ADP – glucose into amyloplasts of maize endosperms. *Plant Physiol.* 117, 1235–1252.
Shaw, G. C., Cope, J. J., Li, L., Corson, K., Hersey, C., Ackermann, G. E., Gwynn, B., Lambert, A. J., Wingert, R. A., Traver, D., Trede, N. S., Barut, B. A., Zhou, Y., Minet, E., Donovan, A., Brownlie, A., Balzan, R., Weiss, M. J., Peters, L. L., Kaplan, J., Zon, L. I., and Paw, B. H. (2006). Mitoferrin is essential for erythroid iron assimilation. *Nature* 440, 96–100.
Stappen, R., and Krämer, R. (1994). Kinetic mechanism of phosphate/phosphate and phosphate/OH^- antiports catalyzed by reconstituted phosphate carrier from beef heart mitochondria. *J. Biol. Chem.* 269, 11240–11246.
Tarantino, D., Morandini, P., Ramirez, L., Soave, C., and Murgia, I. (2011). Identification of an *Arabidopsis* mitoferrinlike carrier protein involved in Fe metabolism. *Plant Physiol. Biochem.* 49, 520–529.
Thuswaldner, S., Lagerstedt, J. O., Rojas-Stutz, M., Bouhidel, K., Der, C., Leborgne-Castel, N., Mishra, A., Marty, F., Schoefs, B., Adamska, I., Persson, B. L., and Spetea, C. (2007). Identification, expression, and functional analyses of a thylakoid ATP/ADP carrier from *Arabidopsis*. *J. Biol. Chem.* 282, 8848–8859.
Tjaden, J., Haferkamp, I., Boxma, B., Tielens, A. G., Huynen, M., and Hackstein, J. H. (2004). A divergent ADP/ATP carrier in the hydrogenosomes of *Trichomonas gallinae* argues for an independent origin of these organelles. *Mol. Microbiol.* 51, 1439–1446.
Todisco, S., Agrimi, G., Castegna, A., and Palmieri, F. (2006). Identification of the mitochondrial NAD^+ transporter in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 281, 1524–1531.
Toka, I., Planchais, S., Cabassa, C., Justin, A. M., De Vos, D., Richard, L., Savoure, A., and Carol, P. (2010). Mutations in the hyperosmotic stress-responsive mitochondrial BASIC AMINO ACID CARRIER2 enhance proline accumulation in *Arabidopsis*. *Plant Physiol.* 152, 1851–1862.
Traba, J., Froschauer, E. M., Wiesenberger, G., Satrustegui, J., and del Arco, A. (2008). Yeast mitochondria import ATP through the calcium-dependent ATP-Mg/Pi carrier Sal1p, and are ATP consumers during aerobic growth in glucose. *Mol. Microbiol.* 69, 570–585.
Traba, J., Satrustegui, J., and del Arco, A. (2009). Characterization of SCaMC-3-like/slc25a41, a novel calcium-independent mitochondrial ATP-Mg/Pi carrier. *Biochem. J.* 418, 125–133.
Traba, J., Satrustegui, J., and del Arco, A. (2011). Adenine nucleotide transporters in organelles: novel genes and functions. *Cell. Mol. Life Sci.* 68, 1183–1206.
Tsaousis, A. D., Kunji, E. R., Goldberg, A. V., Lucocq, J. M., Hirt, R. P., and Embley, T. M. (2008). A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. *Nature* 453, 553–556.
Tyra, H. M., Linka, M., Weber, A. P. M., and Bhattacharya, D. (2007). Host origin of plastid solute transporters in the first photosynthetic eukaryotes. *Genome Biol.* 8, R212.
Tzagoloff, A., Jang, J., Glerum, D. M., and Wu, M. (1996). FLX1 codes for a carrier protein involved in maintaining a proper balance of flavin nucleotides in yeast mitochondria. *J. Biol. Chem.* 271, 7392–7397.
van der Giezen, M., Slotboom, D. J., Horner, D. S., Dyal, P. L., Harding, M., Xue, G. P., Embley, T. M., and Kunji, E. R. (2002). Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. *EMBO J.* 21, 572–579.
Vercesi, A. E., Borecky, J., Maia, I. G., Arruda, P., Cuccovia, I. M., and Chaimovich, H. (2006). Plant uncoupling mitochondrial proteins. *Annu. Rev. Plant Biol.* 57, 383–404.
Vercesi, A. E., Martins, I. S., Silva, M. A. P., Leite, H. M. F., Cuccovia, I. M., and Chaimovich, H. (1995). Pumping plants. *Nature* 375, 24.
Voncken, F., Boxma, B., Tjaden, J., Akhmanova, A., Huynen, M., Verbek, F., Tielens, A. G. M., Haferkamp, I., Neuhaus, H. E., Vogels, G., Veenhuis, M., and Hackstein, J. H. P. (2002). Multiple origins of hydrogenosomes: functional and phylogenetic evidence from the ADP/ATP carrier of the anaerobic chytrid *Neocallimastix* sp. *Mol. Microbiol.* 44, 1441–1454.
Wallin, I. E. (1923). The mitochondria problem. *Am. Nat.* 57, 255–261.
Wallin, I. E. (1927). *Symbionticism and the Origin of Species*. Baltimore: Williams & Wilkins.
Ward, H. M. (1883). Chlorophyll corpuscles and pigment bodies in plants. *Nature* 28, 267–268.
Watanabe, A., Nakazono, M., Tsutsumi, N., and Hirai, A. (1999). AtUCP2: a novel isoform of the mitochondrial uncoupling protein of *Arabidopsis thaliana*. *Plant Cell Physiol.* 40, 1160–1166.
Williams, B. A., Haferkamp, I., and Keeling, P. J. (2008). An ADP/ATP-specific mitochondrial carrier protein in the microsporidian *Antonospora locustae*. *J. Mol. Biol.* 375, 1249–1257.
Wohlrab, H. (2006). The human mitochondrial transport/carrier protein family. Nonsynonymous single nucleotide polymorphisms (nsSNPs) and mutations that lead to human diseases. *Biochim. Biophys. Acta* 1757, 1263–1270.
Yin, L., Lundin, B., Bertrand, M., Nurmi, M., Solymosi, K., Kangasjarvi, S., Aro, E. M., Schoefs, B., and Spetea, C. (2010). Role of thylakoid ATP/ADP carrier in photoinhibition and photoprotection of photosystem II in *Arabidopsis*. *Plant Physiol.* 153, 666–677.
Yoon, H., Zhang, Y., Pain, J., Lyver, E. R., Lesuisse, E., Pain, D., and Dancis, A. (2011). Rim2, pyrimidine nucleotide exchanger, is needed for iron utilization in mitochondria. *Biochem. J.* 440, 137–146.
Zrenner, R., Stitt, M., Sonnewald, U., and Boldt, R. (2006). Pyrimidine and purine biosynthesis and degradation in plants. *Annu. Rev. Plant Biol.* 57, 805–836.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 October 2011; accepted: 03 January 2012; published online: 18 January 2012.
*Citation: Haferkamp I and Schmitz-Esser S (2012) The plant mitochondrial carrier family: functional and evolutionary aspects. Front. Plant Sci. **3**:2. doi: 10.3389/fpls.2012.00002*
This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

APPENDIX

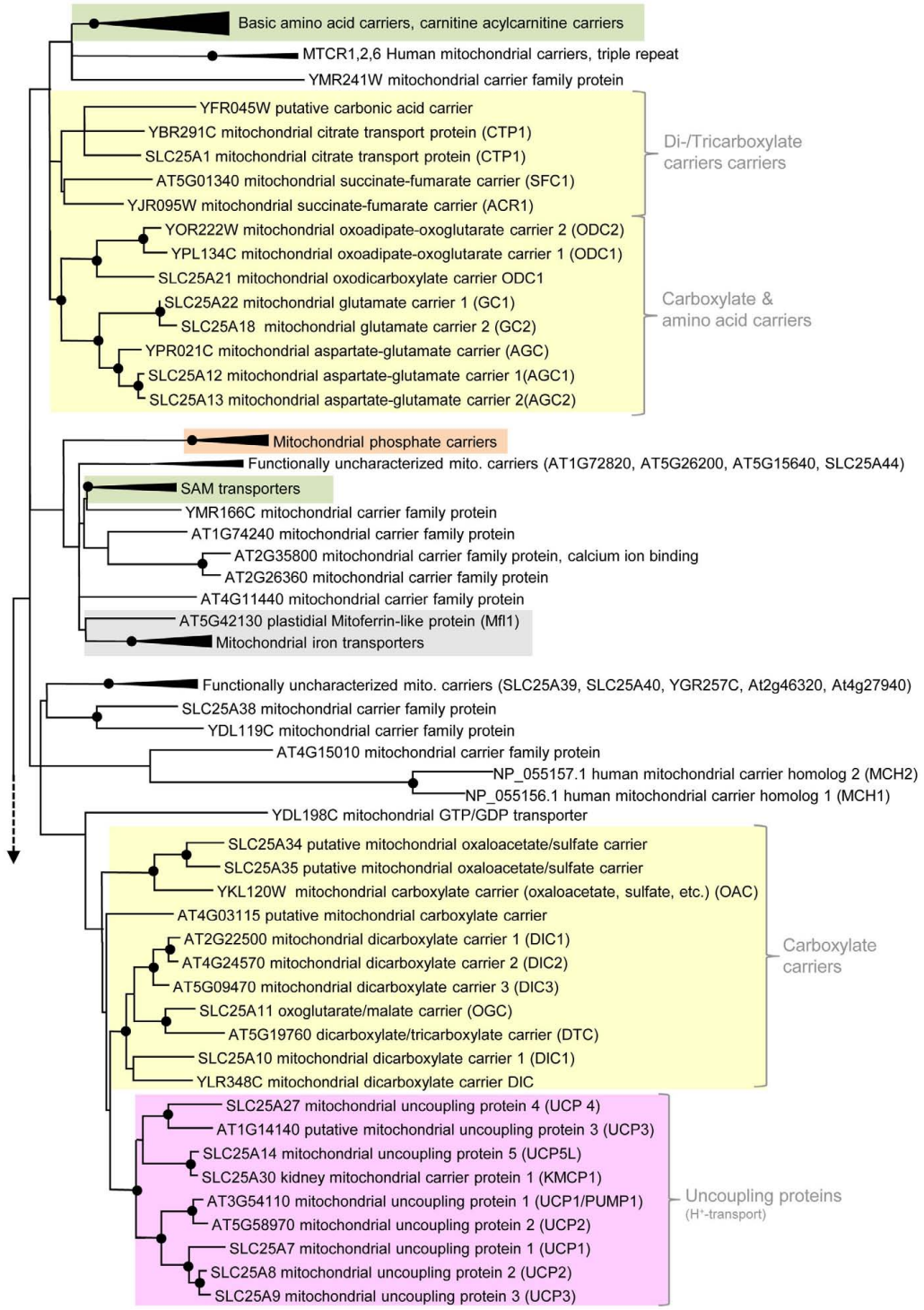


FIGURE A1 | Phylogenetic relationships of putative and proven carbonic acid carriers and uncoupling proteins. Detailed subsets of the phylogenetic tree shown in **Figure 3** are displayed (yellow and purple shaded subsets). Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots.

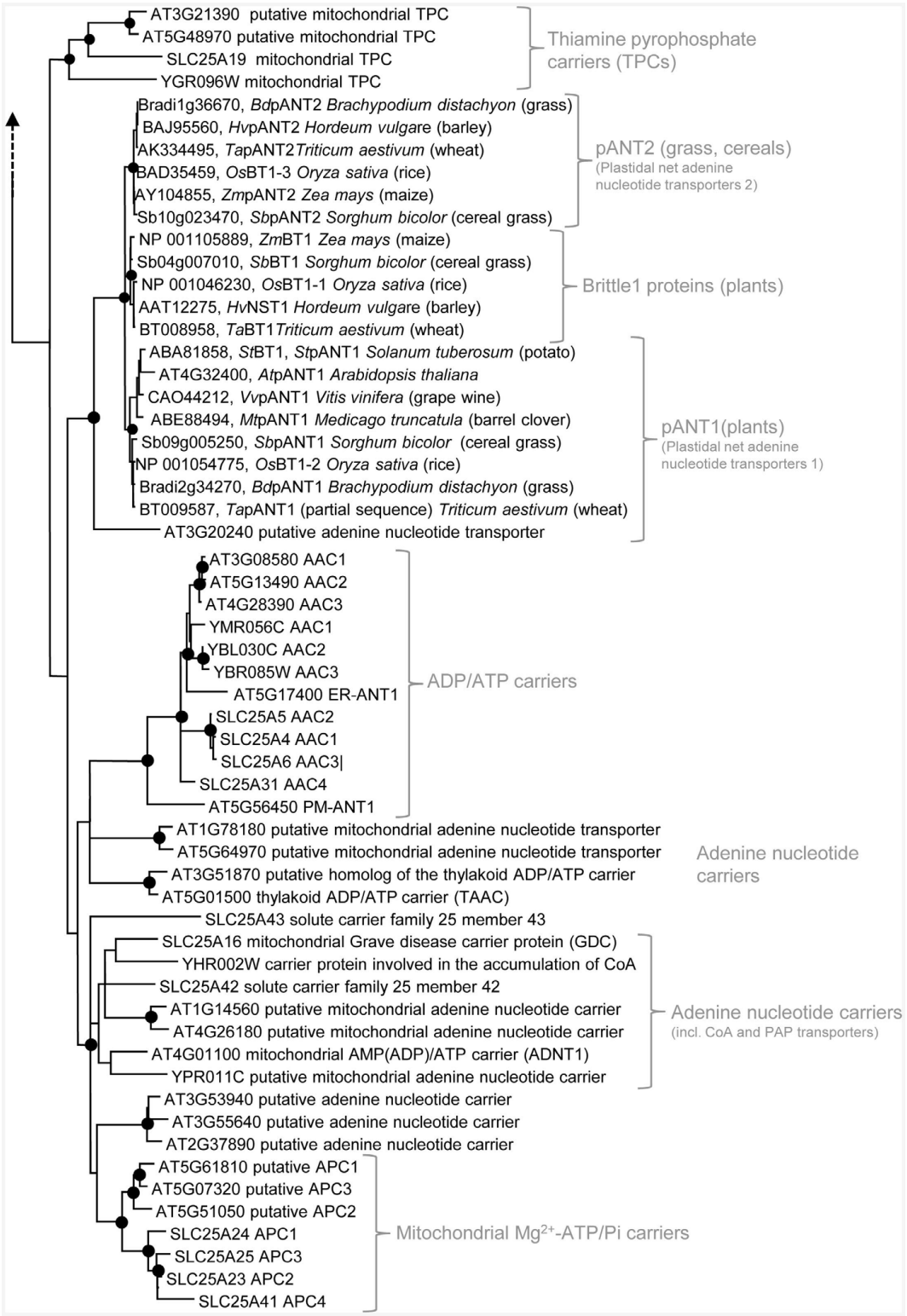


FIGURE A2 | Phylogenetic relationships of putative and proven adenine nucleotide transporters. A subset of the phylogenetic tree shown in **Figure 3** is displayed [majority of the blue shaded subset, the mitochondrial Mg^{2+}–ATP/Pi carrier (Sallp) is included in the subset of **Figure A3**]. Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots.

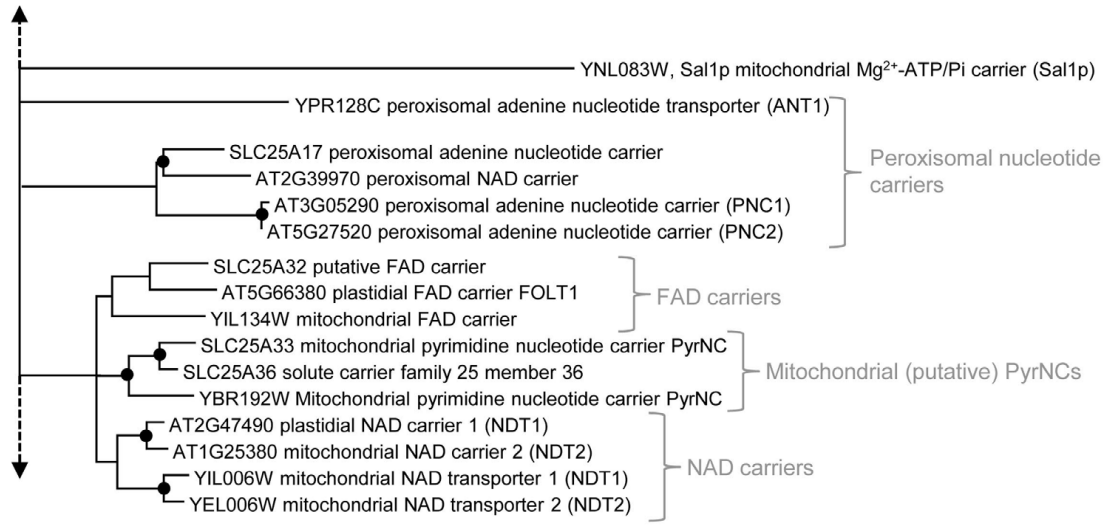


FIGURE A3 | Phylogenetic relationships of peroxisomal adenine nucleotide carriers, pyrimidine nucleotide, FAD, and NAD carriers. A subset of the phylogenetic tree shown in **Figure 3** is displayed [pale blue shaded subset and the mitochondrial Mg^{2+}–ATP/Pi carrier (Sallp)]. Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots.

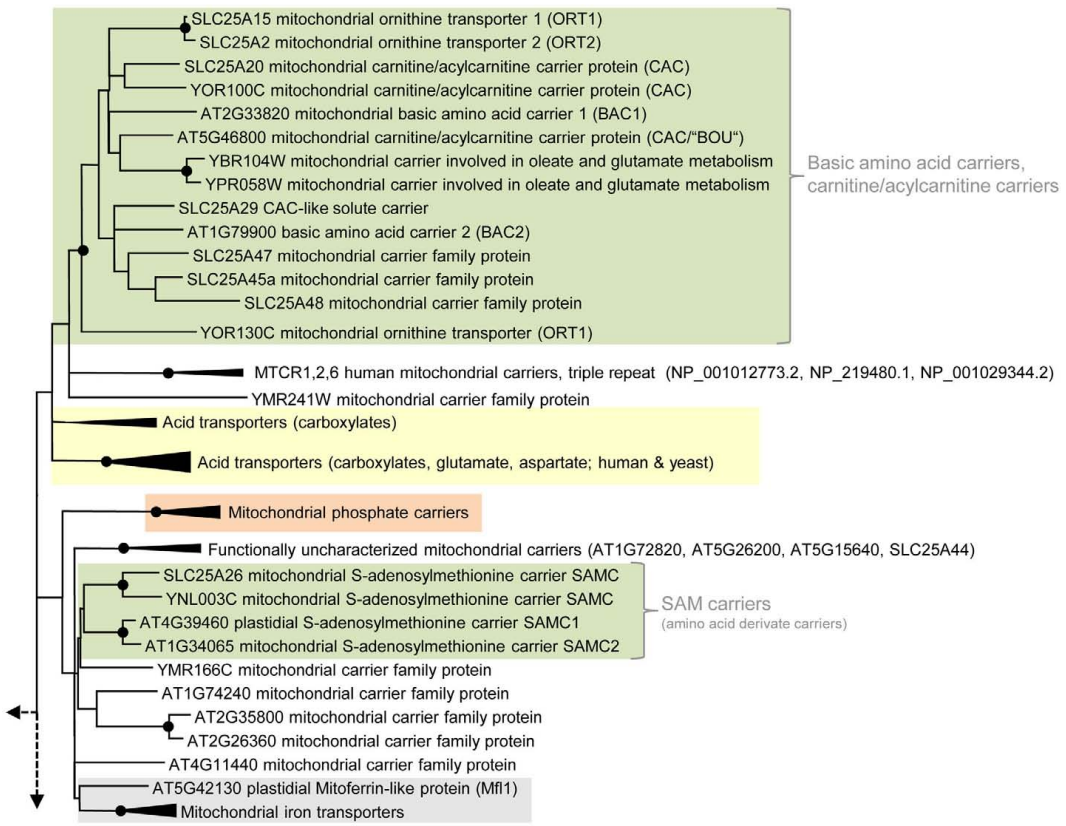


FIGURE A4 | Phylogenetic relationships of putative and proven basic amino acid carriers, carnitine/acylcarnitine carriers, and *S*-adenosylmethionine carriers. Detailed subsets of the phylogenetic tree shown in **Figure 3** are displayed (green shaded subsets). Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots.

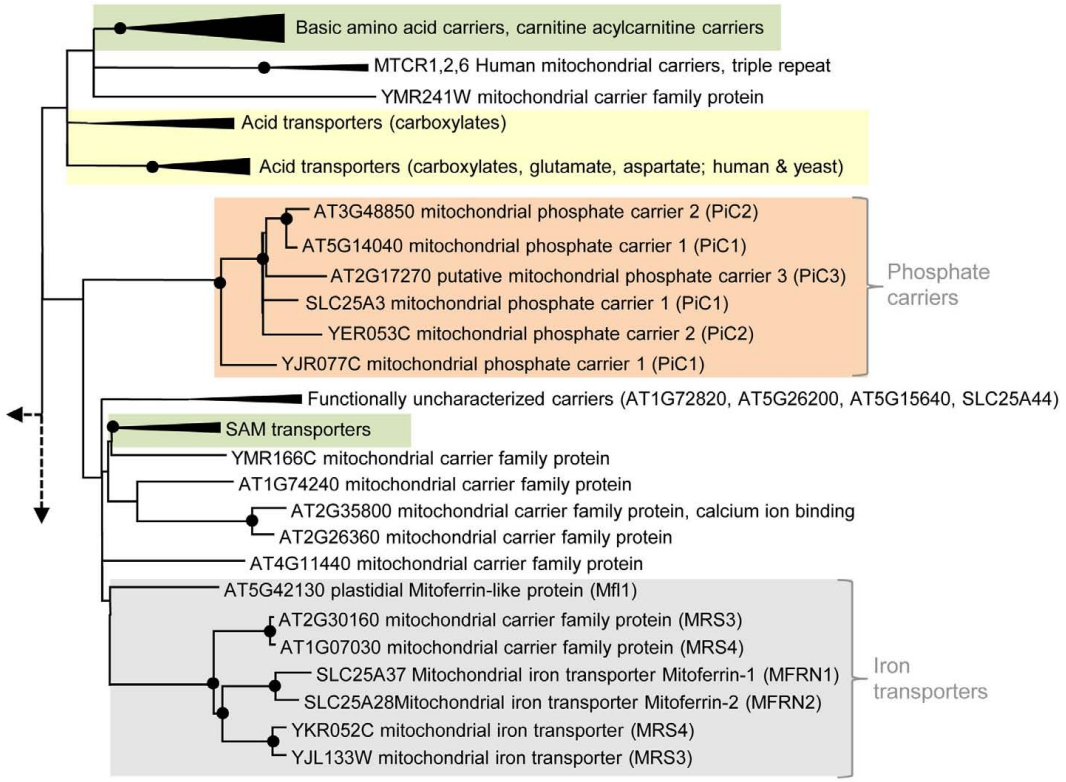


FIGURE A5 | Phylogenetic relationships of putative and proven phosphate and iron carriers. Detailed subsets of the phylogenetic tree shown in **Figure 3** are displayed (orange and gray shaded subsets). Nodes that are supported by bootstrap values (maximum likelihood and maximum parsimony, 1000 replications) higher than 70 are indicated by black dots.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 16 January 2012
doi: 10.3389/fpls.2012.00003

Transport proteins regulate the flux of metabolites and cofactors across the membrane of plant peroxisomes

***Nicole Linka*[1]* and *Christian Esser*[2]**

[1] *Department of Plant Biochemistry, Heinrich Heine University, Düsseldorf, Germany*
[2] *Department of Bioinformatics, Heinrich Heine University, Düsseldorf, Germany*

Edited by:
Markus Geisler, University of Fribourg, Switzerland

Reviewed by:
Rosario Vera-Estrella, Universidad Nacional Autonoma de Mexico, Mexico
Frederica Louise Theodoulou, Rothamsted Research, UK

***Correspondence:**
Nicole Linka, Department of Plant Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstrasse 1, Building 26.03.01, 40225 Düsseldorf, Germany.
e-mail: nicole.linka@uni-duesseldorf.de

In land plants, peroxisomes play key roles in various metabolic pathways, including the most prominent examples, that is lipid mobilization and photorespiration. Given the large number of substrates that are exchanged across the peroxisomal membrane, a wide spectrum of metabolite and cofactor transporters is required and needs to be efficiently coordinated. These peroxisomal transport proteins are a prerequisite for metabolic reactions inside plant peroxisomes. The entire peroxisomal "permeome" is closely linked to the adaption of photosynthetic organisms during land plant evolution to fulfill and optimize their new metabolic demands in cells, tissues, and organs. This review assesses for the first time the distribution of these peroxisomal transporters within the algal and plant species underlining their evolutionary relevance. Despite the importance of peroxisomal transporters, the majority of these proteins, however, are still unknown at the molecular level in plants as well as in other eukaryotic organisms. Four transport proteins have been recently identified and functionally characterized in *Arabidopsis* so far: one transporter for the import of fatty acids and three carrier proteins for the uptake of the cofactors ATP and NAD into plant peroxisomes. The transport of the three substrates across the peroxisomal membrane is essential for the degradation of fatty acids and fatty acids-related compounds via β-oxidation. This metabolic pathway plays multiple functions for growth and development in plants that have been crucial in land plant evolution. In this review, we describe the current state of their physiological roles in *Arabidopsis* and discuss novel features in their putative transport mechanisms.

Keywords: plant, peroxisomes, transport proteins, metabolites

INTRODUCTION

Peroxisomes are eukaryotic organelles, which have been identified in various photosynthetic organisms, from single algal cells to land plants (Igamberdiev and Lea, 2002). Overall our knowledge of how peroxisomes are distributed across the plant kingdom and their versatile roles is limited (Igamberdiev and Lea, 2002; Gabaldon, 2010). Peroxisomes adapted coincidently as land plants derived from the freshwater green algal ancestor that is closely related to the modern Charophytes. Going from water to terrestrial habits led to the development of a complex plant body composed of specialized organs, such as leaves, roots, flowers, and seeds (Langdale, 2008). Plant peroxisomes underwent specialization during land plant evolution and thus contain highly variable and dynamic enzymatic content depending on the specific organ function (Igamberdiev and Lea, 2002).

In vascular plants, peroxisomes are assigned to many processes, including storage oil mobilization via β-oxidation of fatty acids coupled with glyoxylate cycle, photorespiration, membrane lipids turnover, and branched amino acid breakdown during leaf senescence, purine catabolism for nitrogen remobilization, biosynthesis of plant hormones, photomorphogenesis, and pathogen defense (Kaur et al., 2009; Reumann, 2011). Most of these pathways are shared between peroxisomes and other organelles. Peroxisomes are required to scavenge oxidative reactions catalyzed by flavin-containing oxidoreductases (oxidases) that produce highly toxic hydrogen peroxide (del Rio et al., 2006; Nyathi and Baker, 2006). Peroxisomes compartmentalize these lethal steps of metabolism, because they contain efficient ROS-scavenging systems, like catalases, to prevent poisoning of the cell (del Rio et al., 2006; Nyathi and Baker, 2006). Thus, peroxisomes have been described as organelles at the crossroad (Erdmann et al., 1997). Most of our current knowledge is based on detailed studies of fatty acid oxidation and photorespiration, which are both present in algae as well in plants.

During land plant evolution a major innovation was the development of seeds allowing spermatophytes to proliferate and spread to drier areas (Linkies et al., 2010). In seed oil storing dicotyledonous plants (dicots), peroxisomes are involved in storage reserve mobilization to support seedling growth and development until the seedlings become photoautotrophic (Graham, 2008). Upon germination, the fatty acids are released from the seed oil triacylglycerols (TAGs) and degraded via β-oxidation to acetyl-CoA, which is subsequently condensated via glyoxylate cycle to 4-carbon compounds (**Figure 1**). The resulting dicarbonic acids are further converted in mitochondria to malate, which is then either exported for sucrose synthesis or used as a substrate for respiration (Graham, 2008).

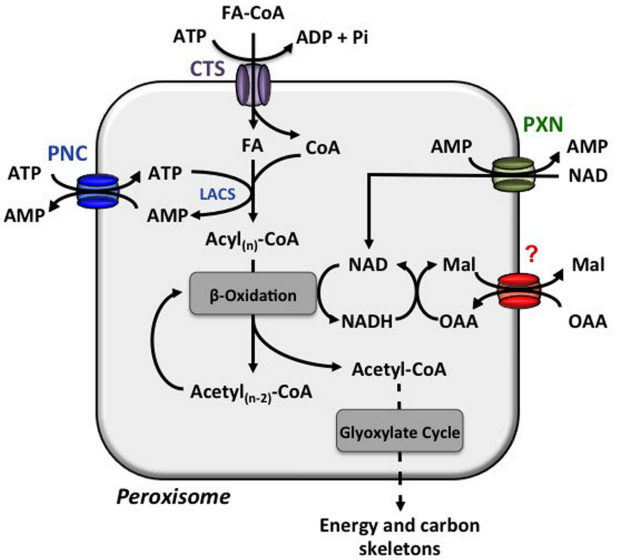


FIGURE 1 | Transport processes across the peroxisome membrane involved in β-oxidation. The COMATOSE (CTS) transporter facilitates the import of fatty acids from the cytosol to the peroxisomal matrix. The peroxisomal long-chain acyl-coenzyme A synthetase (LACS) activates the free fatty acid (FA) to enter β-oxidation for a stepwise release of acetyl-CoA. For the energy consuming esterification of FAs with coenzyme A (CoA), ATP is required and provided by the import via the PCN proteins. The released AMP is the counter-substrate for the next round of ATP import. The reduction equivalent NAD is reduced during β-oxidation. The PXN transport most likely imports NAD to set-up a sufficient NAD pool in the peroxisomes and the main route for NADH regeneration is enabled via the not yet molecularly identified oxaloacetate (OAA)–malate (Mal) shuttle system.

Many monocotyledonous plants (monocots), such as cereal grasses, store starch in the endosperm of their seeds to fuel seedling establishment. In this case, β-oxidation coupled with the glyoxylate cycle plays another critical role (Igamberdiev and Lea, 2002). Both pathways together metabolize TAGs that are present in considerable amounts in the aleurone layer and in the scutellum, not only to provide energy or carbon skeletons for the synthesis of starch-hydrolyzing enzymes secreted into the endosperm, but also to allow the export of dicarbonic acids, which causes an acidification of the endosperm required for α-amylase activities (Drozdowicz and Jones, 1995).

Beside the breakdown of fatty acids during seedling development, β-oxidation enzymes are also present in other plant organs, such as seeds, leaves, roots, and flowers (Pracharoenwattana and Smith, 2008). Fatty acid oxidation and the glyoxylate cycle contribute to multiple processes, including seed development, seed dormancy, flower fertility, turnover of membrane lipids, and branched amino acids during senescence, the generation of jasmonic acid in response to wounding, and auxin biosynthesis to drive root hair and cotyledon cell expansion during seedling development (Baker et al., 2006; Poirier et al., 2006; Strader et al., 2010).

Many algae that mainly store starch as their storage reserve are capable of accumulating TAGs, particularly under abiotic stress conditions, such as nutrient deprivation or high-light exposure (Thompson, 1996; Hu et al., 2008). The enzymes involved in the TAG degradation have not been intensively studied at the biochemical and molecular levels in algae (Thompson, 1996). It was demonstrated that the charophyte *Mougeotia* metabolizes fatty acids via the β-oxidation pathway located in the peroxisomes as it is in embryophytes (Stabenau et al., 1984). It is proposed that storage oil turnover in algae contributes primarily to the assembly of membrane lipids to drive rapid cell division after the cessation of nutrient limitations (Thompson, 1996).

Leaves, which are the major photosynthetic parts of the plant, developed after the conquest of land by leafless plants (Dolan, 2009). When land plants appeared, atmospheric oxygen had risen and carbon dioxide (CO_2) fallen substantially. In such a high-oxygen containing environment the CO_2-fixing enzyme Ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase (Rubisco) favors the oxygenation of RuBP, producing the toxic compound phosphoglycolate (2PG; Bauwe et al., 2010). Plants scavenge 2PG to the Calvin cycle intermediate 3-phoshoglycerate (3PGA) by a sequence of reactions called the photorespiratory C_2 cycle. Thus a high photorespiration rate is a necessary precondition for plants to invade terrestrial habitats (Bauwe et al., 2010).

The metabolic steps are distributed between chloroplasts, peroxisomes, and mitochondria (Bauwe et al., 2010). Leaf-type peroxisomes contain many key enzymes of the photorespiratory pathway, most notably glycolate oxidase and hydroxypyruvate reductase (Reumann and Weber, 2006). A defect in one of these peroxisomal enzymes causes lethality under low CO_2 conditions (Reumann and Weber, 2006). A complete C_2 photorespiratory cycle is also found in C4 species, although they recently innovated a sophisticated CO_2 concentration mechanism (CCM; Bauwe et al., 2010). It seems that despite the existence of an active CCM, photorespiration is still essential for C4 metabolism.

Since the photorespiratory pathway was transferred from cyanobacteria via endosymbiosis, a complete C_2 cycle has been demonstrated in certain green algae (Bauwe et al., 2010). In algae, photorespiration proceeds with lower rates. Due to their ability to concentrate CO_2 around Rubisco via carbonic anhydrase, the glycolate production is suppressed (Stabenau and Winkler, 2005). Few algae metabolize glycolate via the mitochondrial glycolate (D-lactate) dehydrogenase. This reaction is coupled with electron transport in mitochondria (Paul and Volcani, 1976). In case of high photorespiration rates, however, most algae excrete glycolate into the surrounding medium (Stabenau and Winkler, 2005).

Both pathways, the breakdown of fatty acids and the photorespiratory C_2 cycle, necessitate an efficient flux of metabolites and cofactors between peroxisomes and other cellular compartments. These transport processes across the single bilayer are important, because they connect and control the peroxisomal metabolism with that of the other organelles (Theodoulou et al., 2011). A "two-channel" concept describes the permeability of the peroxisome membrane that consists of channel-forming proteins or "porins" along with specific transport proteins (Antonenkov and Hiltunen, 2006; Visser et al., 2007).

Peroxisomal channels allow free diffusion of small compounds with molecular masses less than 300 kDa, as it known for the outer membrane from mitochondria and plastids (Antonenkov and

Hiltunen, 2006; Visser et al., 2007). In addition, selective transport proteins are responsible for the exchange of "bulky" substrates, such as cofactors (ATP, NAD) or fatty acids and their derivatives (Antonenkov and Hiltunen, 2006; Visser et al., 2007). Despite the importance of these porins and transporters, our knowledge is limited in plants, as well as in other eukaryotic organisms. The molecular identity of the channel-forming pore and most of the transporters has not been discovered so far (Theodoulou et al., 2011).

The present review describes the current state of knowledge of the peroxisomal transport processes in plants. Considerable progress has recently been made in the identification of peroxisomal transporters in *Arabidopsis* (Theodoulou et al., 2011). To date, the proteins of two transporter families are known to reside in the peroxisomal membrane: The peroxisomal ATP-binding cassette (ABC) transporter involved in the import of substrates for β-oxidation and three members of the mitochondrial carrier (MC) family required for the influx of the cofactors ATP or NAD, respectively. This review summarizes their biochemical transport properties and their metabolic role for peroxisomes. In the context of land plant evolution, we analyzed the presence of these peroxisomal transporters in algae and other plant species.

THE PEROXISOMAL ABC TRANSPORTER

The peroxisomal ABC transporter imports the substrates for β-oxidation, such as fatty acids or related derivatives, into plant peroxisomes (Theodoulou et al., 2006, 2011). Several independent forward genetic screens identified the peroxisomal ABC transporter in *Arabidopsis*, here referred to as COMATOSE (CTS, At4g39850; Theodoulou et al., 2006, 2011). Pro- and eukaryotic ABC transporters (TC 3.A.1), in general, are composed of two transmembrane domains and two ABC. The energy from the ATP hydrolysis drives the transport of various molecules across the membrane, often against concentration gradients (Higgins, 1992).

The *Arabidopsis* CTS encoded by a single gene is functional as a full-length ABC protein with two dissimilar halves. Interestingly, the non-plant peroxisomal ABC transporter from human and yeast are half-transporters (Wanders et al., 2007). Among the *Arabidopsis* ABC superfamily, that consists of more than 100 members, CTS belongs to the subgroup D (Verrier et al., 2008). In contrast, the second protein of this subgroup (At1g54350) represents a half-size transporter, but its function and putative plastidic localization remain to be analyzed (Verrier et al., 2008).

FUNCTION OF THE PEROXISOMAL ABC TRANSPORTER IN PLANTS

Analysis of *Arabidopsis cts* null mutants demonstrated that CTS is a transporter with relatively broad spectrum of substrates for β-oxidation. It mediates the uptake of several biologically important molecules into peroxisomes, including fatty acids or fatty acid-derived signaling molecules, such as precursors of auxin and jasmonic acid (Baker et al., 2006; Theodoulou et al., 2006). The loss-of-function led to variable phenotypes, reflecting different roles of β-oxidation in a number of developmental processes in *Arabidopsis* (Baker et al., 2006; Theodoulou et al., 2006).

The *cts* mutant alleles exhibit a classical sucrose-dependent phenotype, as mutants defective in β-oxidation. The mutant seedlings are arrested in growth and development due to a block in storage oil mobilization (Zolman et al., 2001; Footitt et al., 2002; Hayashi et al., 2002). High levels of fatty acids in these *cts* seedlings suggest that CTS transports TAG-derived fatty acids for peroxisomal β-oxidation to fuel seedling establishment. This seedling growth phenotype can be rescued in the presence of an exogenous carbon source, such as sucrose (Zolman et al., 2001; Footitt et al., 2002; Hayashi et al., 2002).

Besides mobilization of storage oil during early seedling growth, CTS provides peroxisomal β-oxidation with fatty acids that are hydrolyzed from membrane lipids during lipid turnover (Baker et al., 2006). This process has not only house-keeping function; it plays a crucial role under conditions, when carbon and energy status are low (Kunz et al., 2009; Slocombe et al., 2009). For instance, under extended darkness conditions, β-oxidation respires fatty acids as an energy source. In case of the *cts* mutant, free fatty acids, most likely derived from plastidial lipids, dramatically accumulate, which causes a rapid lethal phenotype compared to wildtype (Kunz et al., 2009; Slocombe et al., 2009).

The root growth of the *cts* mutants is resistant against the protoauxin indole-3-butyric acid (IBA) or the proherbicide 2,4-dichlorophenoxybutyric acid (2,4-DB; Zolman et al., 2001; Hayashi et al., 2002). Due to the loss of CTS both compounds cannot be taken up into peroxisomes, where they are converted into the active auxin indole-3-acetic acid (IAA) and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D), respectively, which both severely inhibit primary root elongation and cotyledon cell expansion (Zolman et al., 2001; Hayashi et al., 2002; Strader et al., 2010). Since several other pathways contribute to auxin biosynthesis, β-oxidation-dependent conversion of IBA to IAA is not essential for other auxin-dependent cellular responses. Notably, in the *cts* mutant the elongation of stamen filaments is inhibited, which could be restored by auxin application. This observation indicates an involvement of peroxisomal β-oxidation to supply auxin (Footitt et al., 2007).

COMATOSE also plays a role in jasmonic acid (JA) biosynthesis, demonstrated by the fact that the levels of both basal and wound-inducible JA are reduced in the *cts* mutants (Theodoulou et al., 2005). It is assumed that CTS imports 12-oxo-phytodienoic acid (OPDA), an intermediate of the JA biosynthesis, into the peroxisomes, where it is further converted by three rounds of β-oxidation to JA (Theodoulou et al., 2005). Because the *cts* mutant still contains residual JA levels, it suggests an alternative route for the peroxisomal OPDA uptake (Theodoulou et al., 2005). Possible transport mechanisms could be either the existence of an unidentified transporter or a passive transport by anion trapping (Theodoulou et al., 2005).

Unlike other mutants involved in JA biosynthesis, *cts* plants are not male-sterile, implying that they have sufficient residual JA to produce fertile pollen (Theodoulou et al., 2005). However, the reduced fertility, observed for *cts* mutants, is caused by an inability to mobilize reserve lipids in both pollen and female gametophytic tissues (Footitt et al., 2007). Therefore, during fertilization, CTS is required to import fatty acids released from stored oil into peroxisomes, where they are broken down via β-oxidation to support energy for pollen germination and tube growth (Footitt et al., 2007).

Mutations in the CTS locus results in seeds that fail to germinate, even in the presence of sucrose, which led to the name COMATOSE (Russell et al., 2000; Footitt et al., 2002; Pinfield-Wells et al., 2005). Only when the seed coat was nicked could *cts* seeds germinate. *cts* embryos have the potential to germinate, but they are unable to rupture the seed coat (Kanai et al., 2010). Transcriptome analysis of the *cts* seeds revealed the molecular mechanism for this phenomenon. The transcription factor ABSCISIC ACID-INSENSITIVE 5 (ABI5) is up-regulated in the mutant and as a consequence, *cts* seeds contain high transcript levels of polygalacturonase inhibiting proteins (Kanai et al., 2010). These proteins inhibit the degradation of pectin in the seed coat, and thus prevent seed coat rupture during germination (Kanai et al., 2010).

Since other β-oxidation mutants also showed a seed dormant phenotype, CTS might import a molecule into the peroxisomes, which is processed via β-oxidation to generate an as-yet unidentified signal molecule for ABI5 gene repression (Kanai et al., 2010). Recently, it has been reported that the *cts* seeds contain elevated levels of 12-oxo-phytodienoic acid (OPDA; Dave et al., 2011). Given the fact that CTS imports OPDA into peroxisomes (Theodoulou et al., 2005), OPDA might be the regulator that triggers, along with ABA, ABI5 protein abundance in *cts* seeds (Dave et al., 2011).

TRANSPORT FUNCTION OF THE PEROXISOMAL ABC TRANSPORTER

Based on the phenotypic mutant analysis, CTS imports various β-oxidation substrates into peroxisomes; but it is still a matter for debate whether the accepted substrates are free fatty acids or activated acyl-CoA esters (Theodoulou et al., 2006, 2011). Supportive evidence for the uptake of free acids via CTS is the presence of several acyl-activating enzymes (AAEs) in peroxisomes from *Arabidopsis* (Fulda et al., 2004; Koo et al., 2006; Wiszniewski et al., 2009). These proteins differ in their substrate specificity and catalyze the esterification of fatty acids or other related molecules, such as the JA precursor or IBA, with coenzyme A (CoA). Such an ATP-dependent activation is essential prior to entering β-oxidation (Fulda et al., 2004; Koo et al., 2006; Wiszniewski et al., 2009). Accordingly, a mutation in these activating enzymes inhibits peroxisomal β-oxidation, resulting in arrested seedling growth, reduced jasmonic acids levels, or IBA resistance (Fulda et al., 2004; Koo et al., 2006; Wiszniewski et al., 2009). Since the *cts* mutant displays the same defects, this implies that CTS and activation reaction operate in the same – rather than parallel – pathways. Thus, CTS delivers the unesterified fatty acids to the peroxisomal matrix, where they are subsequently activated.

Alternatively, it cannot be excluded that CTS imports the activated CoA esters, as it is known for the yeast homolog. In yeast, the peroxisomal ABC transporter, which consists of two heterodimers named PXA1 and PXA2, is involved in the transport of long-chain acyl-CoA esters (Hettema et al., 1996). Expression of CTS in the *pxa1*Δ *pxa2*Δ double mutant rescued the growth phenotype on oleic acid (Nyathi et al., 2010). This complementation assay indicates that CTS is able to transport acyl-CoA esters across the peroxisomal membrane, which is required to metabolize oleic acids as the sole carbon source via β-oxidation. Another implication for activated fatty acids as potential substrates is that acyl-CoA esters, but not fatty acids, stimulate the basal ATPase activity of CTS (Nyathi et al., 2010). In case CTS transports acyl-CoA esters, the β-oxidation substrates have to be activated in the cytosol. However, the corresponding cytosolic enzymes have not been identified so far.

A third model combines the opposite assumptions: CTS imports CoA esters into peroxisomes, the CoA moiety is cleaved off either by CTS or by peroxisomal thioesterases inside the peroxisome to drive the import, and the peroxisomal acyl-CoA activating enzymes re-esterify the β-oxidation substrates (Fulda et al., 2004). At this point the key experiment will be the *in vitro* transport studies using reconstituted CTS protein to determine the biochemical identity of its substrates. In summary, CTS is the entry point for the substrates of the peroxisomal β-oxidation, which is a crucial pathway for plant growth and development.

THE PEROXISOMAL MC-TYPE CARRIER

Transport proteins that belong to the large MC family (TC 2.A.29) reside in the peroxisomal membrane (Picault et al., 2004; Haferkamp, 2007; Palmieri et al., 2011). This family is named after its prominent member the mitochondrial ATP/ADP carrier (AAC), which is highly abundant in the inner mitochondrial membrane. AAC was one of the first transporter identified at molecular levels and up to now the best biochemically characterized transporter. Members of this family are present in all eukaryotes. The MC members have common structural features. They arose by tandem intragenic triplication. Each replicate contains two a-helical TMDs and a characteristic signature motif (PFAM PF00153; Picault et al., 2004; Haferkamp, 2007; Palmieri et al., 2011). Recently, the three dimensional structure of AAC has been solved, suggesting that MCs are functional as a monomer (Kunji and Crichton, 2010). Despite their related structure, MCF members cover a wide range of transported substrates, exhibit different transport mechanisms, and are distributed to various subcellular membranes.

The *Arabidopsis* genome encodes for 58 MCs, while three members are localized into peroxisomes (Picault et al., 2004; Haferkamp, 2007; Palmieri et al., 2011). These peroxisomal MC-type transporters are involved in the import of cofactors, such as ATP and NAD (Arai et al., 2008; Linka et al., 2008; Bernhardt et al., 2012). These striking findings for specific ATP and NAD uptake systems refute early biochemical data that the peroxisomal membrane is impermeable for these cofactors (Antonenkov et al., 2004). These transport proteins allow plant peroxisomes to share a common pool of cofactors with the cytosol.

THE PEROXISOMAL ATP TRANSPORTERS

The peroxisomal adenine nucleotide carrier (PNC) transports ATP into the matrix of plant peroxisomes to fuel ATP-dependent reactions. *Arabidopsis* peroxisomes contain two PNC proteins, PNC1 (At3g05290) and PNC2 (At5g27520; Arai et al., 2008; Linka et al., 2008). A candidate gene approach identified the two related *Arabidopsis* PNC proteins based on sequence similarity to the soybean and yeast homolog, which both are prominent membrane proteins in peroxisomes from *Glycine max* and *Saccharomyces cerevisiae* (Arai et al., 2008; Linka et al., 2008).

Functional complementation studies verified the ATP transport function of the PNC proteins using a yeast mutant lacking the endogenous peroxisomal ATP carrier Ant1p (Palmieri et al., 2001).

This yeast mutant is unable to use medium-chain fatty acids as sole carbon source. Yeast metabolizes medium-chain fatty acids exclusively via peroxisomal β-oxidation. A prerequisite is the import of medium-chain fatty acids and their activation in the peroxisomal matrix. Thus, a lack of the Ant1p inhibits fatty acid breakdown (Palmieri et al., 2001) and expression of PNC1 or PNC2 rescued the growth phenotype of the yeast mutant, demonstrating that both are able to supply yeast peroxisomes with ATP (Linka et al., 2008).

Further, *in vitro* ATP uptake experiments using recombinant proteins provide a direct proof that PNCs indeed facilitate ATP import (Arai et al., 2008; Linka et al., 2008). Expression and functional integration of PNC2 into the *Escherichia coli* membrane revealed ATP and ADP uptake activities into intact bacterial cells, while PNC1 could not be expressed (Arai et al., 2008). Interestingly, the presence of AMP did not inhibit the uptake of external ATP (Arai et al., 2008). On the other hand, reconstitution of the two *Arabidopsis* carriers expressed in yeast led to high ATP uptake activities into liposomes, but only when lipid vesicles were preloaded with another adenine nucleotide (Linka et al., 2008). This clearly shows that PNC1 and PNC2 function as an antiporter that catalyzes the strict counter-exchange of adenine nucleotides (Linka et al., 2008). In contrast to the *E. coli* uptake studies, the reconstitution systems strongly suggest that PNC proteins accept all adenine nucleotides, including AMP (Linka et al., 2008).

FUNCTION OF THE PEROXISOMAL ATP TRANSPORTER IN PLANTS

Plant peroxisomes need ATP to drive their energy consuming reactions. For instance, the activation of β-oxidation substrates depends on peroxisomal ATP, before they are introduced into this cycle and in *Arabidopsis*, PNC1 and PNC2 supply peroxisomes with cytosolic ATP. The action of β-oxidation is severely impaired when both *PNC* genes are repressed via RNA interference (RNAi; Arai et al., 2008; Linka et al., 2008).

The RNAi lines were compromised in seedling establishment in the absence of sucrose (Arai et al., 2008; Linka et al., 2008). The seed oil-derived fatty acids cannot enter the β-oxidation pathway and dramatically accumulate in the *iPNC1/2* lines (Arai et al., 2008; Linka et al., 2008). The apparent seedling phenotype is consistent with that of the double knockout mutant for the two acyl-activating enzymes LACS6 and LACS7 in *Arabidopsis* (Fulda et al., 2004).

The PNC-mediated ATP import is also crucial for peroxisomal enzymes that activate IBA and 2,4-DB (Arai et al., 2008; Linka et al., 2008). An impaired ATP uptake into peroxisomes results in partial resistance against these compounds, which inhibit root growth when converted by a full functional β-oxidation (Arai et al., 2008; Linka et al., 2008). Whether additional ATP-dependent activation steps for β-oxidation are affected in the *PNC1/2i* plants, such as JA precursor, is currently under investigation (*Linka*, personal communication).

In sum, the phenotypic analysis of the *iPNC1/2* plants shows that (i) there are no other ATP-generating systems in plant peroxisomes, such as substrate-level phosphorylation, and that (ii) the PNC-mediated transport pathway is the only source for intraperoxisomal ATP. Interestingly, the sucrose-dependent phenotype is only partially rescued by exogenous sucrose, indicating a further function of the peroxisomal ATP pool beyond β-oxidation. Latest peroxisomal proteomic approaches detected kinases and other ATP-consuming enzymes, e.g., chaperones, in *Arabidopsis*, but their functions remains to be determined (Reumann et al., 2007, 2009). A future aim will be to elucidate whether PNC proteins are required to supply these yet-to-be defined steps with external ATP.

TRANSPORT FUNCTION OF THE PEROXISOMAL ATP TRANSPORTER

Consistent with the substrate specificities obtained by *in vitro* uptake studies and their *in planta* role for β-oxidation, PNC proteins most likely facilitate exchange of cytosolic ATP for peroxisomal AMP under physiological conditions: PNCs import ATP into peroxisomes, which is then hydrolyzed to AMP by the acyl-CoA synthetases. In turn, PNCs export the resulting AMP back to the cytosol, where it is recycled to ATP (Linka et al., 2008). Although PNC catalyze the ATP/AMP exchange *in vivo*, the following important physiological questions regarding the transport mechanism are not yet clarified.

(i) How is the nucleotide pool in plant peroxisomes loaded in the first place? In yeast, the peroxisomal ATP transporter Ant1p catalyzes *in vitro* a unidirectional net import of ATP, in addition to the antiport. Such a uniport mechanism allows a net influx of ATP into yeast peroxisomes early in their genesis (Lasorsa et al., 2004). Both *Arabidopsis* ATP transporters, however, facilitate under the experimental conditions only the exchange of adenine nucleotides (Linka et al., 2008). So it is tempting to speculate that an unidentified uniporter for adenine nucleotides resides in the membrane of plant peroxisomes. Alternatively, PNCs are able *in vivo* to mediate unidirectional transport since they rescued the loss of the yeast carrier without any obvious remaining phenotype.

(ii) Is the transport of ATP and AMP mediated by PNCs electrogenic or electroneutral? An exchange of ATP^{4-} versus AMP^{2-} would result in a net transfer of two negative charges. In case of the AAC, the electrogenic nature of ADP/ATP transport is crucial for its physiological role. The membrane potential of the inner mitochondrial membrane generated by the respiration chain favors the export of ATP synthesized in the mitochondrial matrix and the entry of cytosolic ADP (Traba et al., 2010). For peroxisomes, there are no reports describing that an ATPase or electron transport chain energized the membranes. The uptake of ATP into peroxisomes might be driven by a concentration gradient. Especially during postgerminative growth, the LACS proteins directly consume ATP and produce AMP in the peroxisomal matrix and favor AMP export to the cytosol (Stitt et al., 1982). Under these conditions, the ATP import catalyzed by PNCs is favored. This is also supported by the kinetic properties of the PNC proteins. The apparent inhibition constant (K_i) of AMP is three times higher than for ATP, indicating that AMP inhibits ATP import only at high concentrations of AMP in the cytosol (Linka et al., 2008).

If the PNCs catalyze an electroneutral transport, counter-ions are required to compensate the charge difference of the ATP/AMP exchange (**Figure 2**). PNCs might perform a proton-coupled transport, like it was demonstrated for the yeast carrier ANT1p (Lasorsa et al., 2004). Consequently, PNCs transfer protons into peroxisomes and as a result generate a pH gradient across the peroxisomal membrane. To validate this scenario, the pH of plant peroxisomes has to be determined. For human fibroblasts, Chinese hamster ovary cells, and yeast reports about the peroxisomal pH are controversial describing a basic, acidic or neutral lumen for peroxisomes (Dansen et al., 2000; Jankowski et al., 2001; Lasorsa et al., 2004; van Roermund et al., 2004). Notably, calculated isoelectric points of the peroxisomal proteins found by proteomic approaches suggest that the peroxisomal lumen is basic (Reumann et al., 2004). To avoid acidification due to the proton-compensated ATP/AMP exchange, maintenance of the peroxisomal pH is mandatory to ensure a constant internal environment for optimal enzyme activities. Despite the extensive studies of the peroxisomal proteome, there is no evidence for a proton-pumping ATPase at the peroxisomal membrane in any eukaryotic organisms. Alternatively, the "Donnan equilibrium" might maintain the alkaline pH in peroxisomes. The positively net charge of proteins inside peroxisomes drives the free penetration of hydroxyl ions across the peroxisomal membrane. As a result the pH inside peroxisomes is kept more basic than outside (Antonenkov and Hiltunen, 2006; Rokka et al., 2009).

Conversely, an export of pyrophosphate (PP_i) from the peroxisomal matrix could balance the electrogenic ATP/AMP exchange. Beside the electroneutrality, this co-transport also prevents the accumulation of PP_i inside peroxisomes, which is produced during the intraperoxisomal ATP-dependent activation of β-oxidation substrates (Fulda et al., 2004). The plastidial ATP transporter, for example, mediates such a phosphate-coupled transport. It imports ATP from the cytosol into plastids in exchange with stromal ADP and inorganic phosphate (P_i) and hence maintains the cellular phosphate homeostasis (Trentmann et al., 2008). For the PNCs, P_i as a third efflux substrate can be excluded, because none of the *Arabidopsis* encoded pyrophosphatases are targeted to peroxisomes, which could decompose PP_i into two P_i molecules. Thus the export of PP_i from the peroxisomal matrix is essential. The reconstitution of peroxisomal membranes from mammals revealed phosphate/pyrophosphate transport activities (Visser et al., 2005). The existence of a phosphate carrier would reject the hypothesis that PNCs catalyze an ATP/AMP + PP_i antiport. However, the corresponding gene for this postulated transport process has not yet been assigned (Visser et al., 2005).

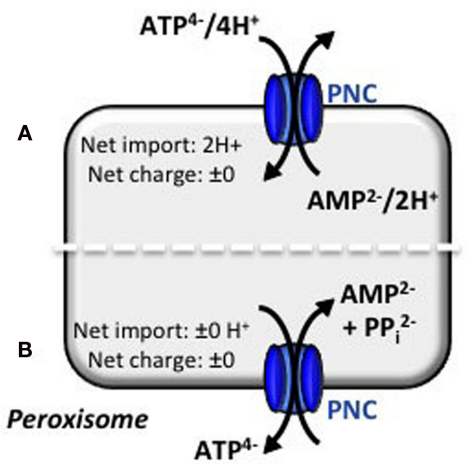


FIGURE 2 | Electroneutral transport mechanisms for the peroxisomal ATP carrier. Based on the literature concerning similar transport mechanisms, two modes of transport actions for ATP can be hypothesized. **(A)** Electroneutral ATP/AMP counter-exchange is accomplished by co-transport of protons. The import of $2H^+$ must be compensated by the Donnan equilibrium. **(B)** Electroneutral ATP/AMP counter-exchange is maintained by a three substrate-model. AMP and PP_i are exchanged for ATP and no proton net import is needed. Both models are under investigation (N. Linka, personal communication).

Concisely stated, additional transport studies are required to understand in more detail how the PNC proteins transport ATP across the peroxisomal membrane, e.g., influence of protonophores on ATP/AMP exchange or stimulation of internal PP_i on ATP uptake against AMP.

THE PEROXISOMAL NAD CARRIER IN *ARABIDOPSIS*

The transport of NAD into plant peroxisomes is mediated via the peroxisomal NAD carrier PXN (At2g39970; Bernhardt et al., 2012). The existence of a transport protein that imports cytosolic NAD into peroxisomes has been controversially discussed for decades (Rottensteiner and Theodoulou, 2006). The biosynthesis of NAD in the cytosol necessitates the import of NAD into peroxisomes for numerous reduction/oxidation (redox) reactions (Noctor et al., 2006; Hashida et al., 2009).

Proteomic studies detected an abundant membrane protein with an apparent mass of 38 kDa in the purified peroxisomal membrane fraction of etiolated pumpkin cotyledons by Edman degradation (Fukao et al., 2001). The peptide sequence obtained showed a high sequence similarity to the PXN. Since its discovery 10 years ago, it has been assumed that PXN represents the peroxisomal ATP transport due to its close relationship to the mitochondrial AAC (Fukao et al., 2001). This hypothesis was recently disproved, considering that PXN was unable to complement the growth phenotype of the *ant1*Δ yeast mutant that is impaired in the peroxisomal ATP uptake (Linka et al., 2008). Further phylogenetic analysis exhibited that At2g39970 clusters with recently identified NAD carrier from plastids and mitochondria (Palmieri et al., 2009), proposing a NAD transport function for PXN (Bernhardt et al., 2012).

The *in vitro* synthesized PXN protein demonstrated the uptake of NAD into liposomes, which were preloaded with NAD (Bernhardt et al., 2012). No significant NAD uptake activities were detectable in the absence of a counter-exchange substrate. This compelling result points out that PXN transports NAD in an antiport mechanism, like the plastidial and mitochondrial NAD transporters (Bernhardt et al., 2012). Further analysis identified NADH, AMP, and ADP as suitable efflux substrates for the NAD import (Bernhardt et al., 2012). The highest NAD uptake rates against AMP indicated that PXN preferentially mediates the NAD/AMP antiport, like the other previously characterized NAD

carriers from *Arabidopsis* and yeast (Todisco et al., 2006; Palmieri et al., 2009). However, PXN is unique in terms of its capability to transport *in vitro* NAD in its reduced form (Bernhardt et al., 2012).

Based on the *in vitro* transport data, PXN imports NAD against AMP. But how does PXN achieve a net influx of NAD to fuel NAD-dependent reactions? To balance the loss of peroxisomal AMP, two scenarios can be proposed for a net NAD influx (Bernhardt et al., 2012). An unknown peroxisomal carrier might be involved in the unidirectional re-import of cytosolic AMP into peroxisomes. The presence of an adenylate uniporter could refill the peroxisomal adenine nucleotide pool (**Figure 3A**). In *Arabidopsis* a peroxisomal Nudix hydrolase with a typical peroxisomal targeting signal might provide the efflux substrates for PXN (Reumann et al., 2004). This enzyme favors *in vitro* the hydrolysis of NADH to AMP and NMNH (Ge et al., 2007; Ge and Xia, 2008). The net import of one NAD molecule occurs, when PXN imports two NAD molecules against one AMP molecule and one NMNH molecule. Both efflux substrates derived form the hydrolysis of one NADH molecule (**Figure 3B**). In the cytosol they can be fed in the salvage pathway to regenerate NAD (Noctor et al., 2006; Hashida et al., 2009).

FUNCTION OF THE PEROXISOMAL NAD CARRIER IN PLANTS

Redox reactions play fundamental roles in plant peroxisomes (Cousins et al., 2008; Pracharoenwattana and Smith, 2008). For instances, the multifunctional protein strictly requires NAD as a cofactor for the oxidative degradation of fatty acids (Graham, 2008; Quettier and Eastmond, 2009). The peroxisomal NAD carrier might be involved in the uptake of NAD to fuel β-oxidation.

A link between the function of PXN and fatty acid breakdown is indicated by the publicly available microarray database *Arabidopsis* eFP Browser (Winter et al., 2007). *PXN* is ubiquitously expressed in all tissues, but dry mature seeds store high amounts of *PXN* mRNA, indicating a possibly involvement in early developmental stages upon germination (Nakabayashi et al., 2005). The steady state transcript level of *PXN* is also up-regulated in senescent leaves and thus might be required for the process of senescence (van der Graaff et al., 2006).

Consistent with the gene expression on the transcript level are the results of the immunoblot analysis using the related pumpkin protein of PXN (PMP38). Like the enzymes of the β-oxidation, PMP38 is highly abundant in young etiolated pumpkin seedlings (Fukao et al., 2001). Once the transition to photoautotrophic metabolism occurred, the protein amount was markedly decreased. In addition, an accumulation of PMP38 protein was observed during *in vivo* senescence when green pumpkin cotyledons were kept in extended darkness (Fukao et al., 2001). For both processes, NAD-dependent β-oxidation plays a pivotal role in mobilization of storage or membrane lipids (Baker et al., 2006; Graham, 2008).

Surprisingly, *Arabidopsis* loss-of-function mutants did not display an apparent sucrose-dependent phenotype, indicating that the storage oil mobilization was not completely blocked during seedling establishment (Bernhardt et al., 2012). A detailed analysis, revealed that the fatty acid breakdown was delayed, because fatty acids accumulated and oil bodies were still present in the *pxn* seedlings (Bernhardt et al., 2012). Nevertheless, the *pxn* mutants were able to metabolize a sufficient portion of the seed storage oil that enabled normal seedling growth. Taken together, PXN delivers the peroxisomal β-oxidation with NAD to ensure its optimal operation during storage oil mobilization (Bernhardt et al., 2012).

The intermediate seedling phenotype suggests, for instances, the existence of a second NAD import system in the peroxisomal membrane that builds up the peroxisomal NAD pool together with PXN. Alternatively, plant peroxisomes already contain a "catalytic" amount of NAD for their metabolism, when they derive from the ER (Hoepfner et al., 2005). Since peroxisomes are highly dynamic organelles and they multiply via fission according to their metabolic demand (Kaur and Hu, 2009), each division would dilute the intraperoxisomal NAD content. Hence, PXN might replenish a certain NAD level in peroxisomes instead of setting it up. Thirdly, NAD is taken up together with its NAD-dependent enzymes from the cytosol via the peroxisomal protein importomer. This large pore allows the passage of fully folded and assembled proteins across the membrane (Meinecke et al., 2010). In yeast, it has been shown that the FAD-dependent acyl-CoA oxidase binds its cofactor in the cytosol, where it is

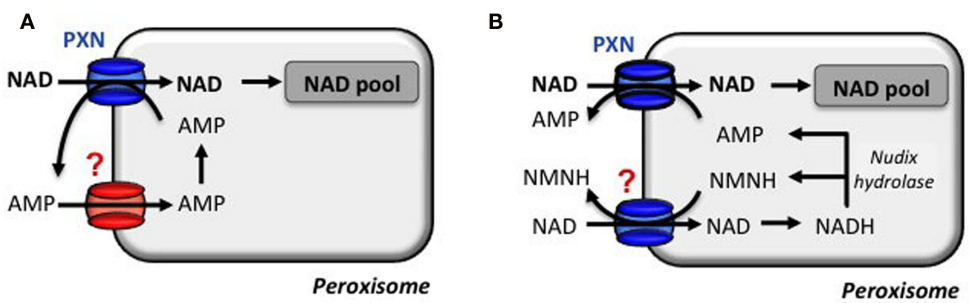


FIGURE 3 | Transport modes for a net NAD influx via the peroxisomal NAD carrier. The proposed net import of NAD by the transport protein PXN is strictly accompanied by an export of AMP. **(A)** NAD import is hypothesized to be supported by an AMP uniporter to provide a peroxisomal AMP counter-substrate pool in the first place. **(B)** Alternatively, an *in silico* predicted peroxisomal Nudix hydrolase is postulated to degrade NADH to AMP and reduced nicotinamide mononucleotide (NMNH). Both function as a counter-substrate and deliver a net import of NAD. A NMNH/NAD carrier has not been described and NMNH has not been tested for PXN (N. Linka, personal communication).

synthesized, and is then targeted to the peroxisomes (Titorenko et al., 2002).

In contrast to its proposed function as a net NAD importer, PXN might catalyze *in vivo* the import of NAD in exchange with NADH. This unique transport feature of PXN would lead *in vivo* to the transfer reducing equivalents across the peroxisomal membrane (Bernhardt et al., 2012). It is widely accepted that the peroxisomal malate/oxaloacetate shuttle is the exclusive route for the exchange of the oxidized and reduced forms of NAD with the cytosol (Mettler and Beevers, 1980). In the course of β-oxidation NAD is reduced to NADH, which is re-oxidized by the peroxisomal malate dehydrogenase (PMDH) via the reversible reduction of oxaloacetate to malate (Pracharoenwattana et al., 2007, 2010). This enzyme is part of this redox shuttle together with the so far unidentified transporters exchanging these two dicarbonic acids. Since PXN and the peroxisomal redox shuttle are redundant in their function, it would explain the partial inhibition of β-oxidation in the *pxn* mutants as well as in plants deficient in PMDH activities (Pracharoenwattana et al., 2007, 2010; Bernhardt et al., 2012).

Future analyses are required to elucidate the *in vivo* role of the peroxisomal NAD carrier in plants. For example, further investigations will address whether other peroxisomal NAD-dependent processes, such as photorespiration, depend on PXN. The *pxn* plants that were grown under photorespiratory conditions did not exhibit an obvious phenotype (Linka, personal communication); however, a detailed metabolic analysis remains to be done.

DISTRIBUTION OF THE PEROXISOMAL TRANSPORTER PROTEINS WITHIN THE PLANT KINGDOM

To date, members of the ABC transporter and MC families represent carrier proteins of the peroxisome membrane in plants. Their function and peroxisomal localization have so far been investigated only in *Arabidopsis*. We conducted a comprehensive BLASTP analysis (Altschul et al., 1997) using protein

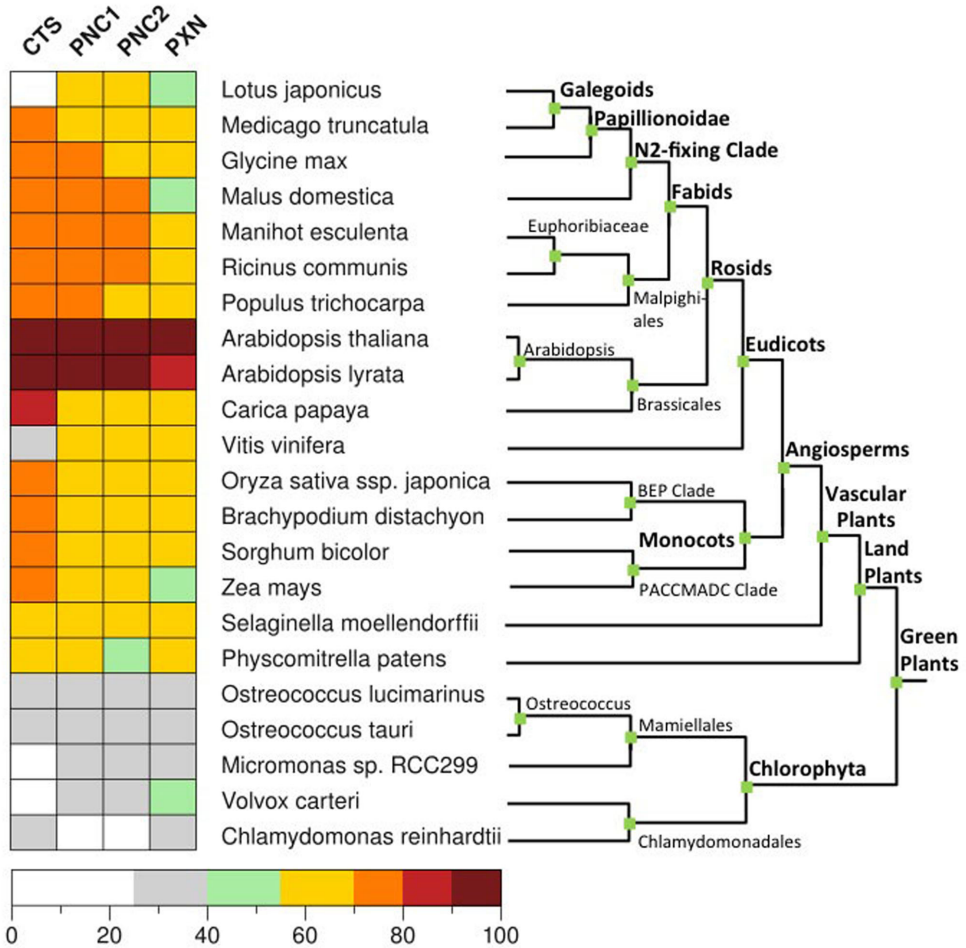


FIGURE 4 | Distribution of the peroxisomal transporter involved in the uptake of fatty acids, ATP, and NAD in green plants. Homologs genes for CTS (At4g39850), PNC1 (At3g05290), PNC2 (At5g27520), and PXN (At2g39970) were identified in 22 fully sequenced plant genomes from PLAZA 2.0 (http://bioinformatics.psb.ugent.be/plaza) using BLASTP (Altschul et al., 1997). For rice only *Oryza sativa* ssp. *japonica* was taken into account. The reciprocal hit was chosen if it was in both directions among the three best hits in the other genome. As a cut-off value we used a normalized sequence identity greater than 25%.

database of 22 fully sequenced plant genomes from PLAZA 2.0 (http://bioinformatics.psb.ugent.be/plaza), including 5 algae, 1 moss, and 17 land plants. To identify homologs for CTS, PNCs, and PXN, the reciprocal best hit was chosen if it was in both directions among the three best hits in the other genome (**Figure 4**). This finding indicates that the evolution of these peroxisomal transporters was an early event before plants evolved. The four plant-specific peroxisomal transporters were already present in the eukaryotic host cells before the cyanobacterial endosymbiont was engulfed. This is consistent with the fact that peroxisomes are eukaryotic organelles and present in animals, plants, and fungi. The peroxisomal transport proteins from plants most likely derived from already existing host membrane proteins and are conserved in green plants.

It is not surprisingly that members of the ABC transporter and MC families reside in the peroxisomal membrane. Both families have extensively expanded during evolution. The members are highly abundant in various eukaryotes, display a broad substrate spectrum, and are widely distributed in all types of cellular membranes, such as mitochondria, plastids, peroxisomes. The diversity within a species is driven by duplication of genes that allow for the acquisition of new functions. This contributes to plant complexity and diversification, allowing the plant peroxisomes to adapt during land plant evolution.

Among most of the analyzed land plants, the putative CTS genes encode for a full size ABC protein. This gene structure appears to be conserved in land plants with a few exceptions. In algae the candidates for the peroxisomal ABC transporter represent a half-size transporter, like the human and yeast homologs (**Table 1**). The presence of CTS-related proteins in algae is contradictory to earlier reports, stating that no peroxisomal ABC transporter genes were found in the algal genomes (Schulze and Kolukisaoglu, 2006).

Regarding the peroxisomal MC-type proteins, at least two different candidates have been identified within the green plants that might be putative ATP or NAD transporters (**Table 2**). One exception is the single-cell chlorophyte *Chlamydomonas*, the genome of which encodes only for one peroxisomal MCF-type gene that is related to PXN. In contrast, two paralogs for the peroxisomal ATP carrier are found in *Arabidopsis*, the apple, and the moss *Physcomitrella*. In case of *Arabidopsis*, PNC1 and PNC2 are derived from the recent chromosome re-arrangement in the genome (Palmieri et al., 2011).

The conducted candidate gene approach was based on protein sequence similarity. None of the predicted candidates have been experimentally investigated up to now. Future efforts are required to confirm the predicted function of the MC- or ABC proteins from algae, moss, and land plants, including measurements

Table 1 | Plant homologs of the peroxisomal ABC transporter from *Arabidopsis*.

Organism	CTS (PLAZA gene ID)	Identity (%)	Length (AA)	Molecular weight (kDa)	Sub-group
L. japonicus	No hit	–	–	–	–
M. trunculata	MT3G36770	77	1335	150.15	FS
G. max	GM11G38160	77	1287	144.79	FS
M. domestica	MD07G005530	76	1952	220.18	FS
M. esculenta	ME05923G00060	76	796	89.49	HS
R. communis	RC30186G00190	75	1339	151.05	FS
P. trichocarpa	PT05G07590	78	987	110.4	HS
A. thaliana	**AT4G39850**	**100**	**1337**	**149.6**	**FS**
A. lyrata	AL7G02070	98	1337	149.55	FS
C. papaya	CP00213G00160	81	465	51.76	HS
V. vinifera	VV19G09290	26	697	78.83	HS
O. sativa	OS01G73530	72	1326	147.95	FS
B. distachyon	BD2G61780	74	1251	140.38	FS
S. bicolor	SB03G047010	72	1324	148.03	FS
Z. mays	ZM08G17060	72	1326	148.26	FS
S. moellendorffii	SM00029G00480	62	1306	145.81	FS
P. patens	PP00150G00750	65	1271	141.49	FS
O. lucimarinus	OL02G00720	32	583	64.69	HS
O. tauri	OT02G06480	26	722	79.43	HS
Micromonas. sp. *RCC299*	No hit	–	–	–	–
V. carteri	No hit	–	–	–	–
C. reinhardtii	CR15G00650	26	753	81.34	HS

*The table summarizes the information of the genes used for the distribution analysis of the CTS gene (**Figure 4**). The following information is listed: PLAZA gene identity (ID) number of the reciprocal hit in the individual genome, relative identity of the genes to the Arabidopsis homolog, molecular weight, amino acid (AA) length, and molecular structure half or full size (HS or FS). The pairwise normalized identities were calculated by dividing the number of identical positions by the length of the shorter protein. As a cut-off value we used a normalized sequence identity greater than 25%. In case for A. thaliana the best match was the query sequence itself and has been highlighted in bold.*

Table 2 | Plant homologs of the peroxisomal MC-type transporter from *Arabidopsis*.

Organism	PNC1 (PLAZA gene ID)	Identity (%)	PNC2 (PLAZA gene ID)	Identity (%)	PXN (PLAZA gene ID)	Identity (%)
L. japonicus	LJ5G027490	63	LJ5G027490	63	LJ0G084980	55
M. trunculata	MT5G33690	70	MT5G33690	64	MT5G64320	66
G. max	GM02G11800	72	GM02G11800	67	GM18G03400	66
M. domestica	MD00G459280	71	MD00G150490	72	MD01G023270	51
M. esculenta	ME03905G00190	72	ME03905G00190	71	ME04268G00320	70
R. communis	RC30128G01600	76	RC30128G01600	71	RC29669G00050	69
P. trichocarpa	PT13G02220	76	PT13G02220	70	PT10G18860	69
A. thaliana	**AT3G05290**	**100**	**AT5G27520**	**100**	**AT2G39970**	**100**
A. lyrata	AL3G05130	98	AL6G27740	99	AL4G26590	90
C. papaya	CP00994G00020	69	CP00994G00020	68	CP00091G00750	60
V. vinifera	VV05G04420	68	VV05G04420	64	VV13G03870	69
O. sativa	OS05G32630	69	OS05G32630	67	OS03G15860	65
B. distachyon	BD2G26480	68	BD2G26480	65	BD1G67180	63
S. bicolor	SB09G019430	68	SB09G019430	68	SB01G040130	63
Z. mays	ZM08G13190	67	ZM08G13190	66	ZM08G15920	47
S. moellendorffii	SM00005G05600	57	SM00005G05600	59	SM00006G03610	58
P. patens	PP00086G00350	58	PP00022G01420	55	PP00090G00790	57
O. lucimarinus	OL01G02450	39	OL01G02450	37	OL20G02530	39
O. tauri	OT01G02440	35	OT01G02440	33	OT14G02730	37
Micromonas. sp. RCC299	MRCC299_16G02440	27	MRCC299_16G02440	26	MRCC299_13G00830	38
V. carteri	VC00102G00130	27	VC00102G00130	26	VC00025G00570	42
C. reinhardtii	No hit	–	No hit	0	CR07G08520	30

*The table summarizes the information of the genes used for the distribution analysis of the MC-type transporters PNC1, PNC2, and PXN (**Figure 4**). Following information is listed: PLAZA gene identity (ID) number of the reciprocal best hit in the individual genome, relative identity of the genes to the Arabidopsis homolog. The pairwise normalized identities were calculated by dividing the number of identical positions by the length of the shorter protein. As a cut-off value we used a normalized sequence identity greater than 25%. In case for A. thaliana the best matches were the query sequences itself and have been highlighted in bold.*

of substrate specificity using the recombinant proteins and verification of the peroxisomal localization by GFP fusion.

CONCLUSION

Future efforts will discover additional peroxisomal transport proteins. Recent proteome studies using purified peroxisomes revealed only the identification of highly abundant membrane proteins that have been already known, such as CTS, PXN, PNC, and a number of biogenesis proteins. This indicates that the other peroxisomal transporters are underrepresented in the peroxisome membrane. A challenge will be to enrich sufficient amounts of peroxisomal membrane proteins for upcoming proteomic approaches. Complementary strategies bear good prospects of elucidating the peroxisomal permeome, including co-expression analyses, mutant studies in combination with metabolic profiling, and genome-wide localization studies of putative transport proteins tagged with GFP.

For decades, several researchers have searched for the postulated transporter that transfers the intermediates of the glyoxylate cycle and photorespiration in plants. This missing link is important to understanding how the flux through the photorespiration is mediated. Recently, it has been hypothesized that the peroxisomal membrane protein with a molecular mass of 22 kDa (PMP22) might be a candidate for peroxisomal porin in plants (Reumann, 2011). To address this, the channel activities have to be determined using electrophysiological techniques in the near future.

Transporters connect the highly compartmentalized metabolic network in plant cells, tissues, and organs and unite the separate parts. A deeper knowledge of the organism-wide permeome will have a major impact on our understanding and designing of metabolic fluxes. This is also true for the subset of transporters in the peroxisomal membrane. Important pathways such as β-oxidation, glyoxylate cycle, photorespiration, branched amino acid breakdown, purine catabolism, biosynthesis of plant hormones, and pathogen defense resides completely or in part in peroxisomes. Fine-tuning the flux via transporters is a promising approach for applied sciences. Peroxisomal research should be intensified in order to show the biotechnological potential and also to close the knowledge gap compared to other organelles. There is still a lot to learn about peroxisomal transporters and a comprehensive understanding of the transport mechanism is required in model organisms prior the transfer into crop plants.

ACKNOWLEDGMENTS

This work was supported by the DFG-grant 1781/1-1 and 1781/1-2 (to Nicole Linka) and DFG-iGRAD plant GRK1525 (to Christian Esser). We thank Andreas P. M. Weber for helpful discussions. We are also thankful that Kristin Bernhardt, Martin Schroers, Sarah K. Vigelius, Jan Wiese, and Thomas J. Wrobel committed to elucidating of the peroxisomal permeome in plants.

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.

Antonenkov, V. D., and Hiltunen, J. K. (2006). Peroxisomal membrane permeability and solute transfer. *Biochim. Biophys. Acta* 1763, 1697–1706.

Antonenkov, V. D., Sormunen, R. T., and Hiltunen, J. K. (2004). The rat liver peroxisomal membrane forms a permeability barrier for cofactors but not for small metabolites in vitro. *J. Cell Sci.* 117, 5633–5642.

Arai, Y., Hayashi, M., and Nishimura, M. (2008). Proteomic identification and characterization of a novel peroxisomal adenine nucleotide transporter supplying ATP for fatty acid β-oxidation in soybean and *Arabidopsis. Plant Cell* 20, 3227–3240.

Baker, A., Graham, I. A., Holdsworth, M., Smith, S. M., and Theodoulou, F. L. (2006). Chewing the fat: β-oxidation in signalling and development. *Trends Plant Sci.* 11, 124–132.

Bauwe, H., Hagemann, M., and Fernie, A. R. (2010). Photorespiration: players, partners and origin. *Trends Plant Sci.* 15, 330–336.

Bernhardt, K., Wilkinson, S., Weber, A. P. M., and Linka, N. (2012). A peroxisomal carrier delivers NAD^+ and contributes to optimal fatty acid degradation during storage oil mobilisation. *Plant J.* 69, 1–13.

Cousins, A. B., Pracharoenwattana, I., Zhou, W., Smith, S. M., and Badger, M. R. (2008). Peroxisomal malate dehydrogenase is not essential for photorespiration in *Arabidopsis* but its absence causes an increase in the stoichiometry of photorespiratory CO2 release. *Plant Physiol.* 148, 786–795.

Dansen, T. B., Wirtz, K. W., Wanders, R. J., and Pap, E. H. (2000). Peroxisomes in human fibroblasts have a basic pH. *Nat. Cell Biol.* 2, 51–53.

Dave, A., Hernandez, M. L., He, Z., Andriotis, V. M., Vaistij, F. E., Larson, T. R., and Graham, I. A. (2011). 12-Oxo-phytodienoic acid accumulation during seed development represses seed germination in *Arabidopsis. Plant Cell* 23, 583–599.

del Rio, L. A., Sandalio, L. M., Corpas, F. J., Palma, J. M., and Barroso, J. B. (2006). Reactive oxygen species and reactive nitrogen species in peroxisomes. Production, scavenging, and role in cell signaling. *Plant Physiol.* 141, 330–335.

Dolan, L. (2009). Body building on land: morphological evolution of land plants. *Curr. Opin. Plant Biol.* 12, 4–8.

Drozdowicz, Y. M., and Jones, R. L. (1995). Hormonal regulation of organic and phosphoric acid release by barley aleurone layers and scutella. *Plant Physiol.* 108, 769–776.

Erdmann, R., Veenhuis, M., and Kunau, W. H. (1997). Peroxisomes: organelles at the crossroads. *Trends Cell Biol.* 7, 400–407.

Footitt, S., Dietrich, D., Fait, A., Fernie, A. R., Holdsworth, M. J., Baker, A., and Theodoulou, F. L. (2007). The COMATOSE ATP-binding cassette transporter is required for full fertility in *Arabidopsis. Plant Physiol.* 144, 1467–1480.

Footitt, S., Slocombe, S. P., Larner, V., Kurup, S., Wu, Y., Larson, T., Graham, I., Baker, A., and Holdsworth, M. (2002). Control of germination and lipid mobilization by COMATOSE, the *Arabidopsis* homologue of human ALDP. *EMBO J.* 21, 2912–2922.

Fukao, Y., Hayashi, Y., Mano, S., Hayashi, M., and Nishimura, M. (2001). Developmental analysis of a putative ATP/ADP carrier protein localized on glyoxysomal membranes during the peroxisome transition in pumpkin cotyledons. *Plant Cell Physiol.* 42, 835–841.

Fulda, M., Schnurr, J., Abbadi, A., Heinz, E., and Browse, J. (2004). Peroxisomal acyl-CoA synthetase activity is essential for seedling development in *Arabidopsis thaliana. Plant Cell* 16, 394–405.

Gabaldon, T. (2010). Peroxisome diversity and evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 765–773.

Ge, X., Li, G. J., Wang, S. B., Zhu, H., Zhu, T., Wang, X., and Xia, Y. (2007). AtNUDT7, a negative regulator of basal immunity in *Arabidopsis*, modulates two distinct defense response pathways and is involved in maintaining redox homeostasis. *Plant Physiol.* 145, 204–215.

Ge, X., and Xia, Y. (2008). The role of AtNUDT7, a Nudix hydrolase, in the plant defense response. *Plant Signal. Behav.* 3, 119–120.

Graham, I. A. (2008). Seed storage oil mobilization. *Annu. Rev. Plant Biol.* 59, 115–142.

Haferkamp, I. (2007). The diverse members of the mitochondrial carrier family in plants. *FEBS Lett.* 581, 2375–2379.

Hashida, S. N., Takahashi, H., and Uchimiya, H. (2009). The role of NAD biosynthesis in plant development and stress responses. *Ann. Bot.* 103, 819–824.

Hayashi, M., Nito, K., Takei-Hoshi, R., Yagi, M., Kondo, M., Suenaga, A., Yamaya, T., and Nishimura, M. (2002). Ped3p is a peroxisomal ATP-binding cassette transporter that might supply substrates for fatty acid β-oxidation. *Plant Cell Physiol.* 43, 1–11.

Hettema, E. H., van Roermund, C. W., Distel, B., van den Berg, M., Vilela, C., Rodrigues-Pousada, C., Wanders, R. J., and Tabak, H. F. (1996). The ABC transporter proteins Pat1 and Pat2 are required for import of long-chain fatty acids into peroxisomes of *Saccharomyces cerevisiae. EMBO J.* 15, 3813–3822.

Higgins, C. F. (1992). ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* 8, 67–113.

Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P., and Tabak, H. F. (2005). Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* 122, 85–95.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., and Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J.* 54, 621–639.

Igamberdiev, A. U., and Lea, P. J. (2002). The role of peroxisomes in the integration of metabolism and evolutionary diversity of photosynthetic organisms. *Phytochemistry* 60, 651–674.

Jankowski, A., Kim, J. H., Collins, R. F., Daneman, R., Walton, P., and Grinstein, S. (2001). In situ measurements of the pH of mammalian peroxisomes using the fluorescent protein pHluorin. *J. Biol. Chem.* 276, 48748–48763.

Kanai, M., Nishimura, M., and Hayashi, M. (2010). A peroxisomal ABC transporter promotes seed germination by inducing pectin degradation under the control of ABI5. *Plant J.* 62, 936–947.

Kaur, N., and Hu, J. (2009). Dynamics of peroxisome abundance: a tale of division and proliferation. *Curr. Opin. Plant Biol.* 12, 781–788.

Kaur, N., Reumann, S., and Hu, J. (2009). "Peroxisome biogenesis and function," in *The Arabidopsis Book*, eds C. R. Somerville and E. M. Meyerowitz (Rockville: American Society of Plant Biologists). doi: 10.1199/tab.0123

Koo, A. J., Chung, H. S., Kobayashi, Y., and Howe, G. A. (2006). Identification of a peroxisomal acyl-activating enzyme involved in the biosynthesis of jasmonic acid in *Arabidopsis. J. Biol. Chem.* 281, 33511–33520.

Kunji, E. R. S., and Crichton, P. G. (2010). Mitochondrial carriers function as monomers. *Biochim. Biophys. Acta* 1797, 817–831.

Kunz, H. H., Scharnewski, M., Feussner, K., Feussner, I., Flügge, U. I., Fulda, M., and Gierth, M. (2009). The ABC transporter PXA1 and peroxisomal β-oxidation are vital for metabolism in mature leaves of *Arabidopsis* during extended darkness. *Plant Cell* 21, 2733–2749.

Langdale, J. A. (2008). Evolution of developmental mechanisms in plants. *Curr. Opin. Genet. Dev.* 18, 368–373.

Lasorsa, F. M., Scarcia, P., Erdmann, R., Palmieri, F., Rottensteiner, H., and Palmieri, L. (2004). The yeast peroxisomal adenine nucleotide transporter: characterization of two transport modes and involvement in DeltapH formation across peroxisomal membranes. *Biochem. J.* 381, 581–585.

Linka, N., Theodoulou, F. L., Haslam, R. P., Linka, M., Napier, J. A., Neuhaus, H. E., and Weber, A. P. M. (2008). Peroxisomal ATP import is essential for seedling development in *Arabidopsis thaliana. Plant Cell* 20, 3241–3257.

Linkies, A., Graeber, K., Knight, C., and Leubner-Metzger, G. (2010). The evolution of seeds. *New Phytol.* 186, 817–831.

Meinecke, M., Cizmowski, C., Schliebs, W., Kruger, V., Beck, S., Wagner, R., and Erdmann, R. (2010). The peroxisomal importomer constitutes a large and highly dynamic pore. *Nat. Cell Biol.* 12, 273–277.

Mettler, I. J., and Beevers, H. (1980). Oxidation of NADH in glyoxysomes by a malate-aspartate shuttle. *Plant Physiol.* 66, 555–560.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., and Nambara, E. (2005). Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *Plant J.* 41, 697–709.

Noctor, G., Queval, G., and Gakiere, B. (2006). NAD(P) synthesis and pyridine nucleotide cycling in plants and their potential importance in stress conditions. *J. Exp. Bot.* 57, 1603–1620.

Nyathi, Y., and Baker, A. (2006). Plant peroxisomes as a source of signalling molecules. *Biochim. Biophys. Acta* 1763, 1478–1495.

Nyathi, Y., De Marcos Lousa, C., van Roermund, C. W., Wanders, R. J., Johnson, B., Baldwin, S. A.,

Theodoulou, F. L., and Baker, A. (2010). The *Arabidopsis* peroxisomal ABC transporter, comatose, complements the *Saccharomyces cerevisiae* pxa1 pxa2 delta mutant for metabolism of long chain fatty acids and exhibits fatty acyl-CoA stimulated ATPase activity. *J. Biol. Chem.* 285, 29892–29902.

Palmieri, F., Pierri, C. L., De Grassi, A., Nunes-Nesi, A., and Fernie, A. R. (2011). Evolution, structure and function of mitochondrial carriers: a review with new insights. *Plant J.* 66, 161–181.

Palmieri, F., Rieder, B., Ventrella, A., Blanco, E., Do, P. T., Nunes-Nesi, A., Trauth, A. U., Fiermonte, G., Tjaden, J., Agrimi, G., Kirchberger, S., Paradies, E., Fernie, A. R., and Neuhaus, H. E. (2009). Molecular identification and functional characterization of *Arabidopsis thaliana* mitochondrial and chloroplastic NAD carrier proteins. *J. Biol. Chem.* 284, 31249–31259.

Palmieri, L., Rottensteiner, H., Girzalsky, W., Scarcia, P., Palmieri, F., and Erdmann, R. (2001). Identification and functional reconstitution of the yeast peroxisomal adenine nucleotide transporter. *EMBO J.* 20, 5049–5059.

Paul, J. S., and Volcani, B. E. (1976). A mitochondrial glycolate: cytochrome C reductase in *Chlamydomonas reinhardtii*. *Planta* 129, 59–61.

Picault, N., Hodges, M., Palmieri, L., and Palmieri, F. (2004). The growing family of mitochondrial carriers in *Arabidopsis*. *Trends Plant Sci.* 9, 138–146.

Pinfield-Wells, H., Rylott, E. L., Gilday, A. D., Graham, S., Job, K., Larson, T. R., and Graham, I. A. (2005). Sucrose rescues seedling establishment but not germination of *Arabidopsis* mutants disrupted in peroxisomal fatty acid catabolism. *Plant J.* 43, 861–872.

Poirier, Y., Antonenkov, V. D., Glumoff, T., and Hiltunen, J. K. (2006). Peroxisomal β-oxidation – a metabolic pathway with multiple functions. *Biochim. Biophys. Acta* 1763, 1413–1426.

Pracharoenwattana, I., Cornah, J. E., and Smith, S. M. (2007). *Arabidopsis* peroxisomal malate dehydrogenase functions in β-oxidation but not in the glyoxylate cycle. *Plant J.* 50, 381–390.

Pracharoenwattana, I., and Smith, S. M. (2008). When is a peroxisome not a peroxisome? *Trends Plant Sci.* 13, 522–525.

Pracharoenwattana, I., Zhou, W., and Smith, S. M. (2010). Fatty acid β-oxidation in germinating *Arabidopsis* seeds is supported by peroxisomal hydroxypyruvate reductase when malate dehydrogenase is absent. *Plant Mol. Biol.* 72, 101–109.

Quettier, A. L., and Eastmond, P. J. (2009). Storage oil hydrolysis during early seedling growth. *Plant Physiol. Biochem.* 47, 485–490.

Reumann, S. (2011). Toward a definition of the complete proteome of plant peroxisomes: where experimental proteomics must be complemented by bioinformatics. *Proteomics* 11, 1764–1779.

Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G. E., Rasche, N., Luder, F., Weckwerth, W., and Jahn, O. (2007). Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19, 3170–3193.

Reumann, S., Ma, C., Lemke, S., and Babujee, L. (2004). AraPerox. A database of putative *Arabidopsis* proteins from plant peroxisomes. *Plant Physiol.* 136, 2587–2608.

Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., Linka, N., Switzenberg, R., Wilkerson, C. G., Weber, A. P. M., Olsen, L. J., and Hu, J. (2009). In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.* 150, 125–143.

Reumann, S., and Weber, A. P. (2006). Plant peroxisomes respire in the light: some gaps of the photorespiratory C2 cycle have become filled – others remain. *Biochim. Biophys. Acta* 1763, 1496–1510.

Rokka, A., Antonenkov, V. D., Soininen, R., Immonen, H. L., Pirilä, P. L., Bergmann, U., Sormunen, R. T., Weckström, M., Benz, R., and Hiltunen, J. K. (2009). Pxmp2 is a channel-forming protein in mammalian peroxisomal membrane. *PLoS ONE* 4, e5090. doi:10.1371/journal.pone.0005090

Rottensteiner, H., and Theodoulou, F. L. (2006). The ins and outs of peroxisomes: co-ordination of membrane transport and peroxisomal metabolism. *Biochim. Biophys. Acta* 1763, 1527–1540.

Russell, L., Larner, V., Kurup, S., Bougourd, S., and Holdsworth, M. (2000). The *Arabidopsis* COMATOSE locus regulates germination potential. *Development* 127, 3759–3767.

Schulz, B., and Kolukisaoglu, H. U. (2006). Genomics of plant ABC transporters: the alphabet of photosynthetic life forms or just holes in membranes? *FEBS Lett.* 580, 1010–1016.

Slocombe, S. P., Cornah, J., Pinfield-Wells, H., Soady, K., Zhang, Q., Gilday, A., Dyer, J. M., and Graham, I. A. (2009). Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways. *Plant Biotechnol. J.* 7, 694–703.

Stabenau, H., and Winkler, U. (2005). Glycolate metabolism in green algae. *Physiol. Plant.* 123, 235–245.

Stabenau, H., Winkler, U., and Säftel, W. (1984). Enzymes of β-oxidation in different types of algal microbodies. *Plant Physiol.* 75, 531–533.

Stitt, M., Lilley, R. M., and Heldt, H. W. (1982). Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protoplasts. *Plant Physiol.* 70, 971–977.

Strader, L. C., Culler, A. H., Cohen, J. D., and Bartel, B. (2010). Conversion of endogenous indole-3-butyric acid to indole-3-acetic acid drives cell expansion in *Arabidopsis* seedlings. *Plant Cell* 153, 1577–1586.

Theodoulou, F. L., Holdsworth, M., and Baker, A. (2006). Peroxisomal ABC transporters. *FEBS Lett.* 580, 1139–1155.

Theodoulou, F. L., Job, K., Slocombe, S. P., Footitt, S., Holdsworth, M., Baker, A., Larson, T. R., and Graham, I. A. (2005). Jasmonic acid levels are reduced in COMATOSE ATP-binding cassette transporter mutants. Implications for transport of jasmonate precursors into peroxisomes. *Plant Physiol.* 137, 835–840.

Theodoulou, F. L., Zhang, X., De Marcos Lousa, C., Nyathi, Y., and Baker, A. (2011). "Peroxisomal transport systems: roles in signaling and metabolism," in *Transporters and Pumps in Plant Signaling (Signaling and Communications in Plants)*, eds M. Geisler and K. Venema (Berlin: Springer), 327–351.

Thompson, G. A. (1996). Lipids and membrane function in green algae. *Biochim. Biophys. Acta* 1302, 17–45.

Titorenko, V. I., Nicaud, J. M., Wang, H., Chan, H., and Rachubinski, R. A. (2002). Acyl-CoA oxidase is imported as a heteropentameric, cofactor-containing complex into peroxisomes of *Yarrowia lipolytica*. *J. Cell Biol.* 156, 481–494.

Todisco, S., Agrimi, G., Castegna, A., and Palmieri, F. (2006). Identification of the mitochondrial NAD transporter in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 281, 1524–1531.

Traba, J., Satrustegui, J., and del Arco, A. (2010). Adenine nucleotide transporters in organelles: novel genes and functions. *Cell. Mol. Life Sci.* 68, 1183–1206.

Trentmann, O., Jung, B., Neuhaus, H. E., and Haferkamp, I. (2008). Non-mitochondrial ATP/ADP transporters accept phosphate as third substrate. *J. Biol. Chem.* 283, 36486–36493.

van der Graaff, E., Schwacke, R., Schneider, A., Desimone, M., Flügge, U. I., and Kunze, R. (2006). Transcription analysis of *Arabidopsis* membrane transporters and hormone pathways during developmental and induced leaf senescence. *Plant Physiol.* 141, 776–792.

van Roermund, C. W., De Jong, M., Ijlst, L., van Marie, J., Dansen, T. B., Wanders, R. J., and Waterham, H. R. (2004). The peroxisomal lumen in *Saccharomyces cerevisiae* is alkaline. *J. Cell Sci.* 117, 4231–4237.

Verrier, P. J., Bird, D., Burla, B., Dassa, E., Forestier, C., Geisler, M., Klein, M., Kolukisaoglu, U., Lee, Y., Martinoia, E., Murphy, A., Rea, P. A., Samuels, L., Schulz, B., Spalding, E. P., Yazaki, K., and Theodoulou, F. L. (2008). Plant ABC proteins – a unified nomenclature and updated inventory. *Trends Plant Sci.* 13, 151–159.

Visser, W. F., van Roermund, C. W., Ijlst, L., Hellingwerf, K. J., Wanders, R. J., and Waterham, H. R. (2005). Demonstration and characterization of phosphate transport in mammalian peroxisomes. *Biochem. J.* 389, 717–722.

Visser, W. F., van Roermund, C. W., Ijlst, L., Waterham, H. R., and Wanders, R. J. (2007). Metabolite transport across the peroxisomal membrane. *Biochem. J.* 401, 365–375.

Wanders, R. J., Visser, W. F., van Roermund, C. W., Kemp, S., and Waterham, H. R. (2007). The peroxisomal ABC transporter family. *Pflugers Arch.* 453, 719–734.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2, e718. doi:10.1371/journal.pone.0000718

Wiszniewski, A. A. G., Zhou, W. X., Smith, S. M., and Bussell, J. D. (2009). Identification of two *Arabidopsis* genes encoding a peroxisomal oxidoreductase-like protein and an acyl-CoA synthetase-like protein that are required for responses to pro-auxins. *Plant Mol. Biol.* 69, 503–515.

Zolman, B. K., Silva, I. D., and Bartel, B. (2001). The *Arabidopsis* pxa1 mutant is defective in an ATP-binding cassette transporter-like protein required for peroxisomal fatty acid β-oxidation. *Plant Physiol.* 127, 1266–1278.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 October 2011; accepted: 03 January 2012; published online: 16 January 2012.

*Citation: Linka N and Esser C (2012) Transport proteins regulate the flux of metabolites and cofactors across the membrane of plant peroxisomes. Front. Plant Sci. **3**:3. doi: 10.3389/fpls.2012.00003*

This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 30 November 2011
doi: 10.3389/fpls.2011.00083

Phosphate import in plants: focus on the PHT1 transporters

Laurent Nussaume[1]*, Satomi Kanno[2], Hélène Javot[1], Elena Marin[1], Nathalie Pochon[1], Amal Ayadi[1], Tomoko M. Nakanishi[2] and Marie-Christine Thibaud[1]

[1] IBEB-SBVME Laboratoire de Biologie du Développement des Plantes, UMR6191 CNRS-Commissariat à l'Energie Atomique et aux Energies Alternatives Cadarache, Université Aix-Marseille, F-13108 Saint-Paul-lez-Durance, France

[2] Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo, Japan 113-8657

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
Bronwyn Jane Barkla, Universidad Nacional Autónoma de México, Mexico
Joshua Blakeslee, The Ohio State University, USA

***Correspondence:**
Laurent Nussaume, IBEB-SBVME Laboratoire de Biologie du Développement des Plantes, UMR 6191 CNRS-Commissariat à l'Energie Atomique et aux Energies Alternatives Cadarache, Université d'Aix-Marseille, F-13108 Saint-Paul-lez-Durance, Cedex, France.
e-mail: lnussaume@cea.fr

The main source of phosphorus for plants is inorganic phosphate (Pi), which is characterized by its poor availability and low mobility. Uptake of this element from the soil relies heavily upon the PHT1 transporters, a specific family of plant plasma membrane proteins that were identified by homology with the yeast PHO84 Pi transporter. Since the discovery of PHT1 transporters in 1996, various studies have revealed that their function is controlled by a highly complex network of regulation. This review will summarize the current state of research on plant PHT1 multigenic families, including physiological, biochemical, molecular, cellular, and genetics studies.

Keywords: phosphate, phosphorus, plant, transporter, uptake, PHT1, transcriptional and post-translational regulation

INTRODUCTION

Phosphorus is an essential macro-element for life. It plays a key role in many crucial processes such as heredity (DNA, RNA), cellular compartmentalization (membrane lipids), energy metabolism (ATP), and phosphorylation-based signaling mechanisms (Poirier and Bucher, 2002; Vance et al., 2003; Misson et al., 2005; Jouhet et al., 2007). Despite its abundance in the environment (ranked as the 11th most abundant element), phosphorus is neither easily accessible nor evenly distributed in most soils. For example, 20–80% of phosphorus in soil can be present as organic pools (mostly composed of plant and microorganism residues), which cannot be absorbed directly by the plants (Richardson, 1994). In plants, phosphorus represents 0.1–0.5% of the dry weight, and this element is acquired in the form of phosphate [inorganic phosphate (Pi)], an inorganic form of P. Depending on the pH, several anions can exist (PO_4^{3-}, HPO_4^{2-}, $H_2PO_4^-$), however between pH 5–6 the predominant form is dihydrogen phosphate ion ($H_2PO_4^-$). Analysis of Pi uptake indicates that this latter form is favored for transport into the plants, since maximum Pi uptake is typically observed within this pH range (Ullrich-Eberius et al., 1984; Furihata et al., 1992; Schachtman et al., 1998). Two factors restrict the availability of this anion in soil: (i) its assimilation by microbes and (ii) its capacity to strongly interact with most of the cations. In acidic soils (30% of soils worldwide), aluminum and iron are abundant and often combined with Pi. This interaction can strongly limit their bioavailability for biological processes (Von Vexhull and Mutert, 1998). Conversely, *in vitro* experiments have demonstrated that a high Pi level can reduce the pool of some metals in *Arabidopsis*, by a factor of 4–5 (Hirsch et al., 2006). The importance of Pi chelation by cations or organic compounds is apparent, when juxtaposed against the poor efficiency of phosphate fertilizers, where up to 80% of provided Pi cannot be acquired by the crops. Furthermore, Pi is poorly mobile in the soil, as its diffusion coefficient is very low (from 10^{-12} to 10^{-15} m^2 s^{-1}). As a consequence, Pi uptake by the plant creates a depleted area around the roots (Ullrich-Eberius et al., 1984; Furihata et al., 1992; Schachtman et al., 1998). Therefore, Pi can easily be considered as one of the least available plant macronutrients (Raghothama, 1999).

Before the modern intensification of agriculture, farmers relied mostly on phosphorus naturally present in soils or from manure. However, as a growing world population has put high demands on food production (with adverse effects on soil degradation), agriculture has become dependent on external sources of phosphate fertilizers (such as guano, ground bones, and phosphate rock) to increase crop yields (Cordell et al., 2009). Phosphate rock is an abundant and cheap source of phosphorus, and has quickly become the main source of Pi for fertilizer production. Despite this abundance, phosphate rock is still a non-renewable resource, which required 10–15 million years for its production (Cordell et al., 2009). As of 2009, the total mineable phosphate rock reserves were estimated at 15–16 billion tons (containing approximately 30% Pi), according to the US Geological Survey (an organization that collects data worldwide to estimate levels of our natural resources). The gravity of the situation is apparent when it is taken into consideration that 170 million tons of phosphate rock were extracted that same year for fertilizer production. The demand for phosphate rock will only steadily increase, with the mounting concerns to feed a growing global population, and

the development of plant culture dedicated to biofuel production. These facts clearly strengthen the estimate of remaining Pi rock reserves. According to these data, our Pi rock reserves should be exhausted in only 50–100 years (Gilbert, 2009; Gross, 2010). A 700% increase in the price of Pi was observed in 2008, foreshadowing the impact that Pi scarcity can have on the economy (Cordell et al., 2009). This underscores an urgent need for an alternative to the use of phosphate rock reserves, which are largely devoted (82%) to fertilizer production. Moreover, this production process demands the exploitation of high quality Pi reserves in order to avoid contamination by toxic heavy metals such as cadmium or isotopes of the natural decay series of uranium or thorium (Othman and Al-Masri, 2007; Casacuberta et al., 2011; Da Conceicao et al., 2011). Improving our knowledge of Pi acquisition and use by plants will undoubtedly have a positive effect on reducing the dependency on fertilizer supply for crop production.

Inorganic phosphate concentration in plant tissues has been measured at 5–20 mM (Raghothama, 1999), whereas the level available in soils is typically less than 10 μM (Bieleski, 1973). This sharp concentration gradient between the plant and the soil illustrates the crucial role of Pi transporters. The study of the PHT1 family of phosphate transporters is particularly appropriate, due to their presence in the plant plasma membrane (at the interface between the cell and the external medium). Several reviews have focused on this family of transporters (Raghothama, 1999; Smith et al., 2000; Rausch and Bucher, 2002; Smith and Barker, 2002; Bucher, 2007), since the identification of phosphate transporters in *Arabidopsis thaliana* in 1996 (Muchhal et al., 1996). More recent advances have revealed novel processes that control this family of transporters. The present article will synthesize this literature with special attention to PHT1 regulation and the plant model *A. thaliana*.

IDENTIFICATION OF PHT1 PHOSPHATE TRANSPORTERS

Beginning in the early 1990s, several high affinity phosphate transporters were cloned in various fungi. The first one, PHO84, was identified in the yeast *Saccharomyces cerevisiae* (Bun-Ya et al., 1991). Several years later, transporter homologs were found in *Neurospora crassa* (Versaw, 1995; Versaw and Metzenberg, 1995) and in the mycorrhizal fungus *Glomus versiforme* (Harrison and Van Buuren, 1995). Their discovery enabled the identification of phosphate transporters in plants. Two separate approaches were used for their identification in *Arabidopsis*, either by the heterologous complementation of the yeast *pho84* mutant by plant cDNA (Muchhal et al., 1996) or by the identification of *Arabidopsis* ESTs closely related to these proteins (Mitsukawa et al., 1997; Smith et al., 1997). Later, their function as phosphate transporters was confirmed *in planta* by the analysis of *Arabidopsis* mutants such as *pht1;1* and *pht1;4* and the corresponding double mutant (Misson et al., 2004; Shin et al., 2004). The nearly simultaneous identification of these genes by several teams introduced some initial confusion to the literature, as different nomenclatures were proposed (PT or PHT) by these authors. Adding to this confusion, as the genes turned out to belong to a wide multigenic family, several authors published descriptions of different homologs under the same name (Muchhal et al., 1996; Smith et al., 1997; Okumura et al., 1998). Not until the complete sequencing of the *A. thaliana* genome in 2000 was it revealed that the *PHT1* family is comprised of nine members (PHT1;1 through PHT1;9; **Figure 1**). Their names have now been unified according to the rule of the Commission on Plant Gene Nomenclature (Rausch and Bucher, 2002).

Homologs of PHT1 transporters have been characterized in a wide range of species, since their initial identification in *Arabidopsis*. This non-exclusive list includes: potato (Leggewie et al., 1997), *Lupinus albus* (Liu et al., 2001), tomato (Daram et al., 1998; Liu et al., 1998a), *Catharanthus roseus* (Kai et al., 1997), *Medicago truncatula* (Liu et al., 1998b; Xiao et al., 2006), barley (Smith et al., 1999), tobacco (Baek et al., 2001), lotus (Nakamori et al., 2002), rice (Paszkowski et al., 2002; Ming et al., 2005), maize (Nagy et al., 2006), wheat (Tittarelli et al., 2007), *Populus trichocarpa* (Loth-Pereda et al., 2011) as well as more distant organisms such as *Chlamydomonas* (Chang et al., 2005).

PHT1 STRUCTURE

Sequence analysis of different PHT1 transporters highlights an amino acid sequence conserved from fungi to plants. Hydrophobicity analysis predicts that these transporters share a common topology with 12 membrane-spanning domains, which are separated into two groups of six domains by a charged hydrophilic loop (Raghothama, 1999). Both C- and N-termini are expected to be oriented inside the cell, with the protein inserted in the plasma membrane (Raghothama, 1999). The *PHT1* family members encode closely related proteins, as indicated by a greater than 76% identity in their protein sequences between various plant species (*A. thaliana*, tomato, *M. truncatula*, potato, and *C. roseus*; Raghothama, 1999; cf. **Figure 1** for examples). Strong homologies for this protein family are also observed between plants and yeast: *Arabidopsis* PHT1 and yeast PHO84 proteins share 34% identity and around 50% similarity.

In *Arabidopsis*, these transporters range in size from 520 to 550 amino acids, with an estimated molecular weight around 58 kDa. As previously reported in Medicago (Chiou et al., 2001) or in *Arabidopsis* (**Figure 2**) this weight is lower than expected. Moreover, these western blots exhibit also higher weight molecular bands that are compatible with the presence of multimeric proteins (**Figure 2**; Chiou et al., 2001). This hypothesis will require further experiments for validation, such as those performed for NRT2.1 (Wirth et al., 2007). In these experiments multimers were also identified on SDS PAGE gel before validation by crosslinking. Nevertheless, this view is also supported by the semi-dominant character of the *pht1;1–3* mutation in *Arabidopsis* (Catarecha et al., 2007), a trait generally associated with multimeric transporters (Ludewig et al., 2003; Loque et al., 2007) in which incorporating the inactive subunits into the multimeric pore affects its global structure.

PHT1 ARE Pi:H^+ CO-TRANSPORTERS

The study of Pi uptake kinetics *in planta* using radioactive Pi isotopes (P^{33} or P^{32}) identified two distinct phases, which were interpreted as the co-existence of a high and a low affinity system (Cogliati and Clarkson, 1983; Drew and Saker, 1984; Drew et al., 1984). Pi uptake kinetics have been measured for several species (*Lemna gibba*, *Brassica nigra*, *C. roseus*, tobacco, *Arabidopsis*), with

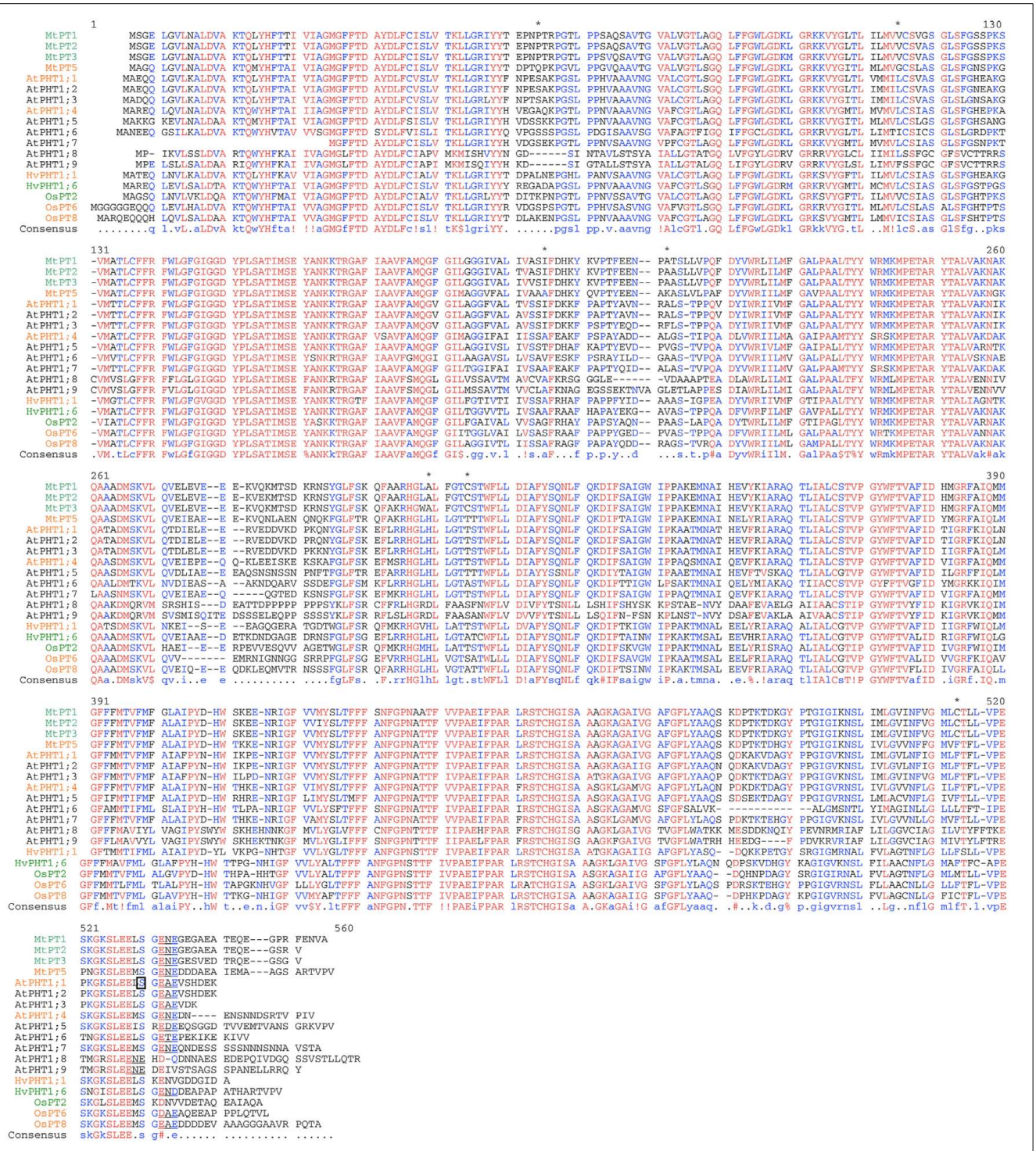


FIGURE 1 | Sequence alignment of several PHT1 transporters from *Medicago truncatula* (Mt), *Arabidopsis thaliana* (At), *Hordeum vulgare* (Hv), and *Oryza sativa* (Os). High or low affinity characterized PHT1 are indicated by orange or green colors respectively. Stars indicate the 7 amino acids with non-conservative changes between the three low affinity (MtPT1, Mt PT2, MtPT3) and the high affinity transporters (MtPT5) identified by Liu et al. (2008). The serine identified in AtPHT1.1 (Bayle et al., 2011) that regulates exit from the ER is boxed and the putative ER exit site is underlined.

K_m values for the high affinity system in the range of 2.5–12.3 μM, whereas K_m values for the low affinity system were usually observed between 50 and 100 μM (Ullrich-Eberius et al., 1984; Lefebvre et al., 1990; Furihata et al., 1992; Shimogawara and Usuda, 1995; Dunlop et al., 1997). In rare cases, only high affinity phosphate transporters could be detected (Drew and Saker,

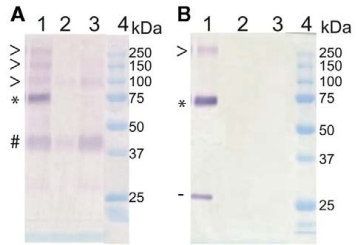


FIGURE 2 | Western blot analysis of PHT1;1, PHT1;2, and PHT1;3 on root membrane protein extracts reveal the presence of potential multimers. Plants were grown *in vitro* on medium supplemented with high (500 μM) or low (10 μM) Pi content (+P or −P respectively). Five micrograms of membrane protein were loaded per lane, on 10% acrylamide SDS PAGE gels before protein transfer onto nitrocellulose membranes. **(A)** The primary antibody was designed against two peptides common to PHT1;1, PHT1;2, and PHT1;3. In parallel, an anti GFP antibody (Roche) was used for control western blots at the same dilution **(B)**. Lanes were loaded as follows: (1) Transgenic line expressing *35S:PHT1;1:GFP* construct; and wild type control on +P (2) or −P (3) and Bio-Rad Precision Plus Protein Standards (4). GFP tagged (*) and untagged (#) PHT1 monomeric form, putative multimers (>) and cleaved GFP (–).

1984), and it remains to be determined if this is linked to the plant samples or the experimental design.

Inorganic phosphate uptake is an energy-mediated process that promotes the alkalinization of the medium, which points toward a Pi/H^+ co-transport (Ullrich and Novacky, 1990). The measurements of phosphate absorption and proton flux revealed a stoichiometry of two to four protons for each phosphate ion transported across the plasma membrane (Ullrich-Eberius et al., 1981, 1984; Sakano, 1990).

It should be noted that phosphate transporters (in particular PHT1 proteins) can also transport Pi analogs such as phosphite (Ticconi et al., 2001; Varadarajan et al., 2002) and arsenate (Lee et al., 2003; Catarecha et al., 2007; Wu et al., 2011). This last compound is highly toxic and can be found in the soil or water of certain areas. Interestingly, both arsenate and phosphite, like Pi, down-regulate PHT1 transcription (Ticconi et al., 2001; Varadarajan et al., 2002; Catarecha et al., 2007).

A RANGE OF AFFINITIES

Several approaches have been used to investigate the Pi transport properties of the various PHT1 family members. One approach relies on the complementation by plant phosphate transporters of the yeast *pho84* mutant or *pho84/pho89* double mutant (PHO84 and PHO89 are the two major high affinity Pi transporters present in yeast). The K_m values measured for several PHT1 family members were: 280 and 130 μM, for two of the potato PHT1 transporters (Muchhal et al., 1996; Leggewie et al., 1997; Liu et al., 1998b); 97 μM for the rice OsPT6 (Ai et al., 2008) or 23 μM for OsPT8 (Jia et al., 2011); and 192 μM for *Medicago* MtPT1 (Liu et al., 1998b).

Although a range of affinities can be determined from yeast complementation assays, in most cases these measurements greatly exceed the expected K_m calculated from *in planta* radioactive Pi uptake experiments (from 2.5 to 100 μM as mentioned in a previous section). This suggests that yeast complementation can only partially reflect the behavior of PHT1 proteins in plants, and it is important to keep this information in mind when predicting the affinities of PHT1 proteins in the plant.

Because of the discrepancies between the *in planta* affinities and the values obtained in yeast, alternative methods may be needed for precise Pi uptake measurements. One such technique is the heterologous transformation of *Xenopus* oocytes with PHT1 transporters, although measured affinities based on this method can also largely differ from the expected values *in planta*. For instance, the rice Pi transporter OsPT2 exhibits low affinity characteristics in the mM scale, when expressed in oocytes (Ai et al., 2008). By contrast, using the same method reveals a K_m value of 19 μM for the barley high affinity Pi transporter HvPHT1 (Preuss et al., 2011). One revealing potential study could be to compare values obtained for both systems (yeast vs. oocytes) for several transporters.

To avoid using a heterologous system, the kinetic parameters can be investigated by gene over-expression in plant cell culture. This technique was used to estimate K_m values for AtPHT1;1 (3.1 μM, Mitsukawa et al., 1997), and for barley HvPHT1;1 (9.06 μM) and HvPHT1;6 (385 μM, Rae et al., 2003). These data are much closer to the affinities expected for Pi uptake measurements *in planta*, although the high Km determined for HvPHT1;6 reinforces the idea that PHT1 members could encode both high and low affinity transporters.

One study has attempted to identify which amino acid alterations are responsible for the variation observed in transporter affinity, in *M. truncatula*. The results indicate that slight variations in the protein sequence of the PHT1 proteins could influence their affinity for Pi transport (Liu et al., 2008): when several *M. truncatula* PHT1 transporters with an 84% amino acid identity were expressed in yeast, their K_m ranged broadly from 13 to 858 μM. Analysis of the transporter sequences revealed that only seven amino acids showed non-conservative changes between the three low affinity (MtPT1, Mt PT2, MtPT3) and the high affinity transporters (MtPT5). These amino acid positions are located all along the protein sequence (**Figure 1**). The predicted structural model hypothesizes that they are clustered in two regions of the protein: one on the extracellular surface, and one within the membrane (Liu et al., 2008). The authors posit that these particular extracellular membrane areas could be directly responsible in controlling the PHT1 affinities. Nevertheless, sequence comparisons with other PHT1 transporters that exhibit distinct affinities for Pi could not reveal a correlation with these results (**Figure 1**). This suggests either independent evolution in separate species, or that the amino acids identified from the *M. truncatula* data cannot account for affinity changes in other species.

SPATIAL DISTRIBUTION

Most members of the PHT1 family exhibit strong expression in roots, and this property is shared between monocotyledonous and dicotyledonous species (Muchhal et al., 1996; Leggewie et al., 1997; Smith et al., 1997; Liu et al., 1998a,b; Schunmann et al., 2004a; Koyama et al., 2005; Nagy et al., 2006; Xiao et al., 2006; Tittarelli et al., 2007; Jia et al., 2011; Loth-Pereda et al., 2011). However, PHT1 proteins are also detected in aerial organs such as leaves or flowers (**Figure 3**).

More precisely, *in situ* hybridization performed in tomato revealed the presence of LePT1 transcript in the root cap and in the external layers of the root (Daram et al., 1998; Liu et al., 1998a). Similar experiments in *M. truncatula* and *A. thaliana* localized MtPT1 and AtPHT1;4 in the root epidermis and root hairs (Chiou et al., 2001; Misson et al., 2004). Immunoblot analyses of tomato LePT1 (Muchhal and Raghothama, 1999) and potato StPT2 (Gordon-Weeks et al., 2003), or use of translational GFP fusions with either AtPHT1;1, AtPHT1;2, or AtPHT1;4 (Gonzalez et al., 2005; Bayle et al., 2011) have all provided similar results, localizing these transporters mainly in epidermis and root hairs. Additionally, translational fusion experiments between PHT1 and

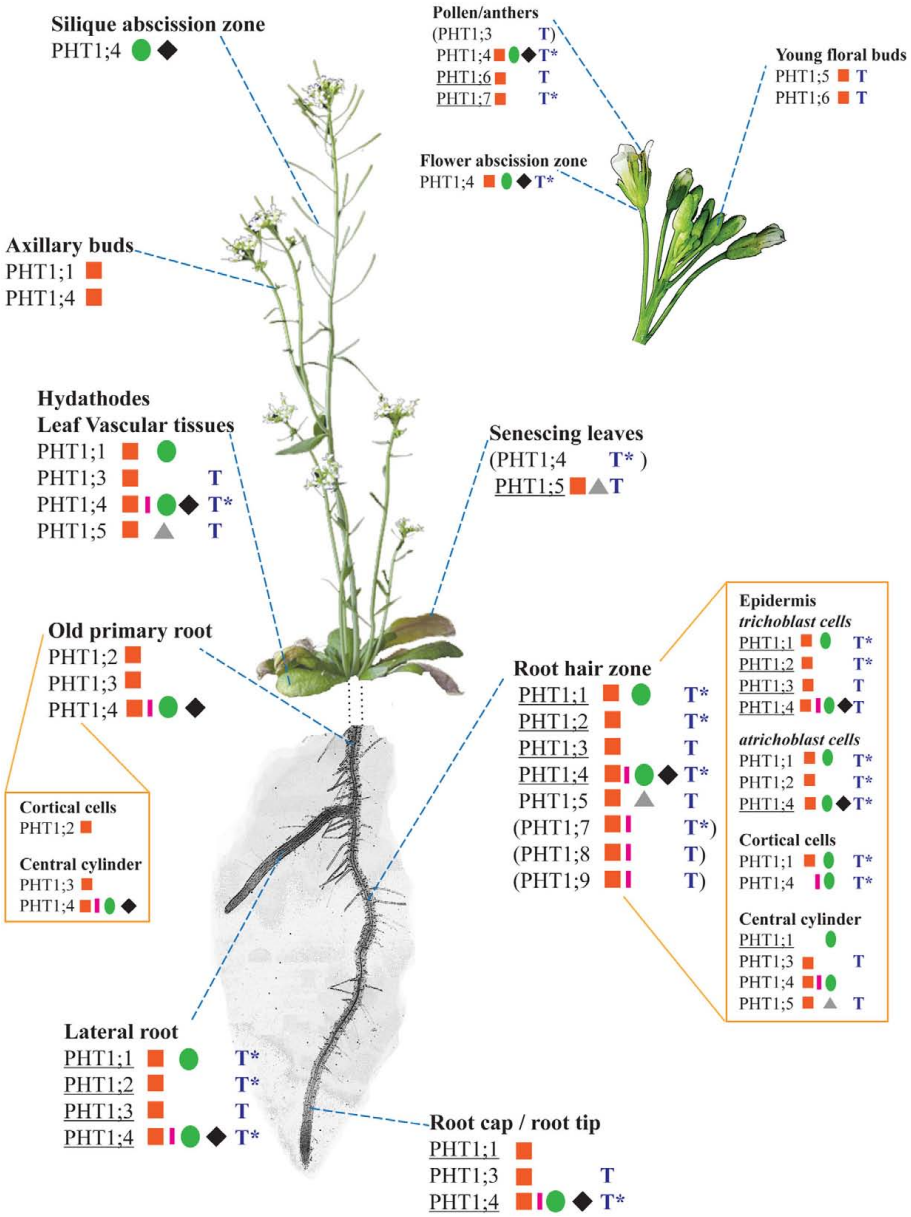


FIGURE 3 | Expression patterns of the nine PHT1 transporters from *Arabidopsis thaliana*, based on a compilation of histological and transcriptomics data (combining information for both +P and −P conditions). The different plant parts are not presented to scale; in particular, the root system has been simplified so that only a single primary root and lateral root are represented. References corresponding to histological studies (promoter::GFP/GUS fusions or *in situ* hybridization) are identified on the figure by the following symbols: red square (Mudge et al., 2002), green circle (Karthikeyan et al., 2002), pink rectangle (Misson et al., 2004), black rhombus (Karthikeyan et al., 2009), and gray triangle (Nagarajan et al., 2011). Results based on transcriptomics studies are indicated by a blue T, and they correspond to a summary of several data sets presented on the eFP server (Winter et al., 2007). *Due to the design of most transcriptomic chips, it is not possible to distinguish the expression patterns of PHT1;1/PHT1;2 as well as PHT1;4/PHT1;7. When a PHT1 transporter is predominantly (but not exclusively) expressed in one or several tissues, its name is underlined. The subset of expression data indicated between brackets () are based on RT-PCR and/or transcriptomics only (i.e., not validated by promoter::reporter studies).

GFP reporter genes confirmed that these phosphate transporters are targeted to the plasma membrane in various species (Gonzalez et al., 2005; Bayle et al., 2011; Jia et al., 2011). Transcriptional fusions between PHT1 promoters and the GUS or GFP reporter genes have also facilitated the study of the expression patterns of PHT1 members from different species, along with the identification of promoter elements involved in the control of these expression patterns (Karthikeyan et al., 2002; Mudge et al., 2002; Schunmann et al., 2004a,b; Misson et al., 2005; Jia et al., 2011). In the specific case of the plant model *A. thaliana*, the full genome sequence permitted extensive analyses of all *Arabidopsis* PHT1 family members, using either RT-PCR (Mudge et al., 2002; Misson et al., 2005) or promoter fusions with GUS or GFP reporter genes (Karthikeyan et al., 2002; Mudge et al., 2002). A landmark contribution to this field was provided by Mudge et al. (2002), who performed a systematic study of the nine PHT1 transporters. Subsequently, the expression patterns were confirmed or analyzed more precisely by several other laboratories (Karthikeyan et al., 2002, 2009; Misson et al., 2004; Nagarajan et al., 2011). In addition, the publication of multiple transcriptomics studies allows us to take into account data sets that were not initially developed for the study of phosphate transport. Powerful tools such as the eFP browser provide an efficient way of analyzing this vast amount of data (Winter et al., 2007). **Figure 3** summarizes the expression patterns for the nine PHT1 members of *A. thaliana*, and compares their published expression patterns to values contained within transcriptomics data sets.

It appears that the vast majority of the transporters (eight of the nine *Arabidopsis* PHT1 members) are expressed at least in roots, in line with their major role in Pi uptake from the soil. Within the roots, the expression of phosphate transporters is mostly concentrated in root epidermis and central cylinder, particularly in the root hair zone (Karthikeyan et al., 2002; Mudge et al., 2002; Misson et al., 2004; Koyama et al., 2005; Xiao et al., 2006; Hirsch et al., 2011).

Most PHT1 transporters are not present in a single cell type, but are found in diverse tissues from various organs, often in overlapping patterns with other PHT1 members, suggesting a greater complexity in their roles (Karthikeyan et al., 2002; Mudge et al., 2002; **Figure 3**). However, each PHT1 member taken individually is defined by a preferential expression in a given tissue, which may indicate the location of its predominant role. It should be noted that this location is often strongly influenced by environmental or developmental factors. For instance, strong AtPHT1;6 expression in pollen could reveal a role for PHT1 transporters during flower development (Karthikeyan et al., 2002; Mudge et al., 2002). Functional confirmation of these expression patterns can be performed by reverse genetic analyses: the role of PHT1 protein in soil Pi uptake was demonstrated with *Arabidopsis pht1;4* and *pht1;1* single and double mutants, which are dramatically attenuated in their phosphate uptake capacities (Misson et al., 2004; Shin et al., 2004). A similar strategy was recently applied to understand the importance of the strong expression of PHT1;5 in senescing leaves (Nagarajan et al., 2011; **Figure 2**): analysis of the *pht1;5* mutant confirmed a role in senescing leaves, and revealed its role in source-to-sink Pi mobilization. Altogether, this literature suggests a major role for PHT1 transporters in Pi uptake from soil, and additional roles in Pi re-mobilization from other organs.

One category of PHT1 transporters which cannot be studied in *Arabidopsis* are the PHT1 transporters that are induced by mycorrhizal symbiosis (endo and ectomycorhizal associations). *Arabidopsis* is among the minority of plants which is incapable of establishing any symbiosis with mycorrhizal fungi. This symbiosis is however widespread among the plant kingdom (80% of vascular plants), and its impact on the capacity of plants to acquire Pi from the soil has been well established (Smith and Read, 1997). This adaptation of plants has probably evolved to benefit from the ability of the fungus to explore a greater part of the soil (exploring up to 100-fold more soil volume) and to have access to mineralized organic phosphorus (Bucher, 2007).

The induction of specific plant PHT1 transporters in response to mycorrhization is a widely reported feature of these symbioses (Javot et al., 2007b; Hata et al., 2010; Loth-Pereda et al., 2011). In the case of the endomycorrhizal symbiosis, the fungus penetrates into the root system and develops hyphal structures (referred to as arbuscules) in the cortex. These arbuscules are composed of a trunk and arbuscular branches, which are enveloped by the plasma membrane of the cortical cell. The expression of a subset of PHT1 members is induced in cells containing arbuscules, which are believed to be the site of nutrient exchange between the plant and the fungus (Smith and Read, 1997; Javot et al., 2007a; Loth-Pereda et al., 2011). The symbiosis-specific PHT1 phosphate transporter MtPT4 from the plant *M. truncatula* is only detected in cells containing arbuscules, and it has exclusively been located in the plasma membrane that surrounds arbuscular branches (Pumplin and Harrison, 2009). The fine regulation of arbuscule development and associated stimulation of PHT1 expression was recently illustrated by the study of the dynamics of the rice transporter OsPT11 fused to GFP (Kobae and Hata, 2010).

Promising techniques using uptake of radioactive Pi isotopes *in planta* can be applied to study specific PHT1 transporters. It is now possible to visualize and analyze the precise dynamics of Pi distribution *in planta*. This technique, using X-ray films or imaging plates, has revealed that the main location for phosphate uptake in *Arabidopsis* is found close to the root tips, and that PHT1 proteins are involved in this process (Misson et al., 2004). A recent improvement in visualizing Pi flux dynamics is the use of a CsI scintillator to convert β-rays into visible light (Kanno et al., 2007). As illustrated in **Figure 4**, it is now possible to compare precisely and in a non-destructive way the entry of Pi into various genotypes of plants at the same time (in particular for mutations affecting PHT1 transporters).

TRANSCRIPTIONAL REGULATION OF PHT1 TRANSPORTERS

The tissue specificity of PHT1 described above appears to be mostly controlled at the transcriptional level. It is likely that the spatial distribution of PHT1 proteins requires regulatory elements located in the promoters, which remain poorly characterized (Schunmann et al., 2004a,b; Karthikeyan et al., 2009). Nevertheless, they are probably well conserved between monocots and dicots, since fusion of the AtPHT1;1 promoter with the GUS reporter gene has given fairly similar expression patterns

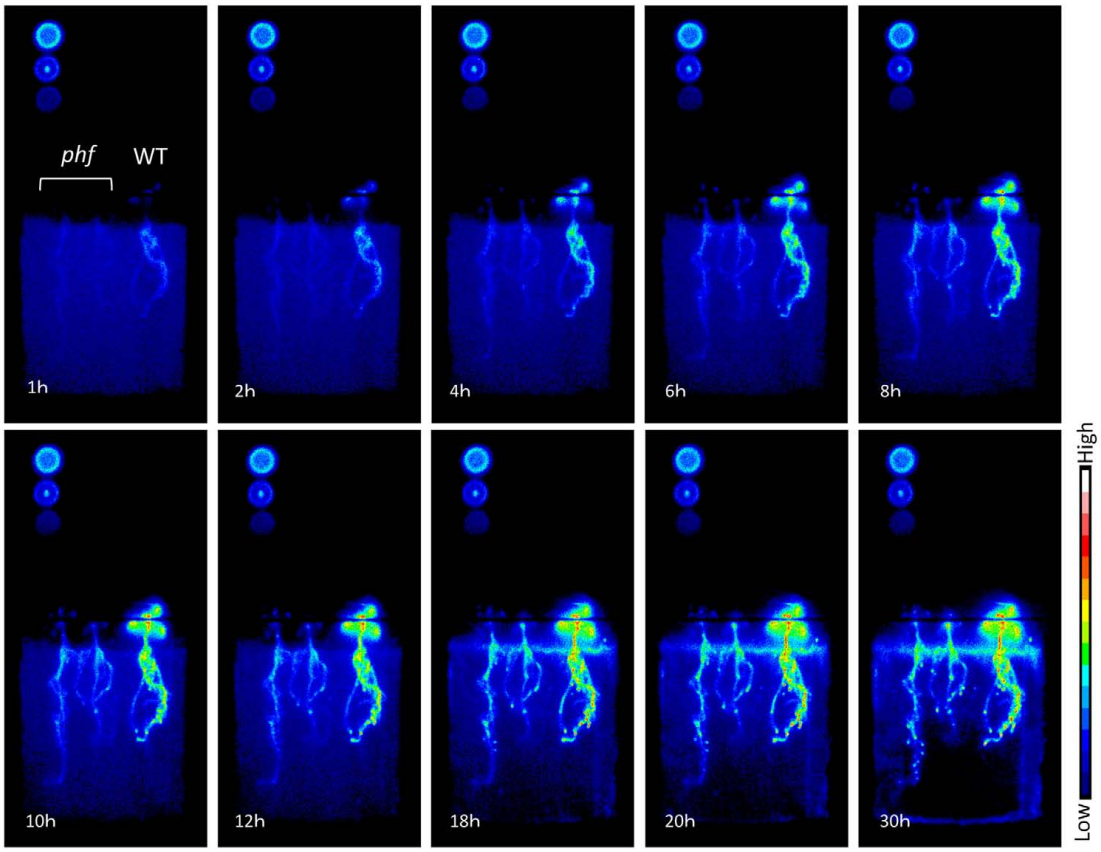


FIGURE 4 | Time course for phosphate absorption based on ^{32}P detection in *Arabidopsis*. Plants were supplied with 10 μM Pi with ^{32}P (H_3PO_4, 12.5 kBq/ml). Images of radioactivity were obtained every 3 min; 10 images are presented here. For each panel, two *phf* mutants (left) and one wild type (Col-0; right) plantlet are shown. Circles on the left are ^{32}P standards:descending from the upper left corner, they represent 2.5, 1, and 0.5k Bq/μl. The plants shown are 11 days old, grown *in vitro* using 1/10 MS medium containing 500 μM Pi and subsequently transferred to 10 μM Pi for 3 days before analysis.

in both rice and *Arabidopsis* transformants (Koyama et al., 2005).

The expression of specific plant PHT1 transporters that are induced in roots in response to mycorrhizal symbiosis is also under transcriptional control in species as diverse as: rice (Paszkowski et al., 2002), *Medicago* (Harrison et al., 2002; Javot et al., 2007a), potato (Nagy et al., 2005), maize (Nagy et al., 2006), various cereals (Glassop et al., 2005), tomato (Nagy et al., 2005; Balestrini et al., 2007; Xu et al., 2007), and trees such as poplar (Loth-Pereda et al., 2011).

The internal Pi concentration of cells is a major factor in the control of *PHT1* expression. This role has been clearly illustrated by various transcriptomics experiments (Hammond et al., 2003; Wu et al., 2003; Misson et al., 2005; Morcuende et al., 2007; Winter et al., 2007; Calderon-Vazquez et al., 2008; Bustos et al., 2010; Thibaud et al., 2010), or with the assistance of reporter genes (Karthikeyan et al., 2002; Mudge et al., 2002; Misson et al., 2004.). Interestingly, plants can adjust *PHT1* transcription when challenged with a broad spectrum of Pi concentrations (from 0 to 1250 μM) that are outside of the expected range in soils, typically observed below 10 μM (Muchhal and Raghothama, 1999; Karthikeyan et al., 2002; Misson et al., 2004). Low basal levels of gene expression are usually observed when the Pi concentration of the medium exceeds 250 μM (Muchhal and Raghothama, 1999; Karthikeyan et al., 2002; Misson et al., 2004).

Variation of Pi concentration in the medium promotes a rapid modulation of *PHT1* transcripts. In *Arabidopsis*, the detection of Pi deficiency can occur within 12–24 h after removal of Pi from the medium (Muchhal and Raghothama, 1999; Misson et al., 2004, 2005). Not surprisingly, re-supply of Pi triggered a faster response (Karthikeyan et al., 2002), presumably since Pi accumulation occurs more rapidly than its depletion from the vacuole (Stefanovic et al., 2011). This phenomenon can be greatly affected by the protocol for cultivating plants (hydroponics *vs.* solid medium), as gelling agents are often contaminated with traces of Pi, which can impact the responses of the plant (Jain et al., 2009). The analysis of promoter sequences from barley, wheat, and *Arabidopsis PHT1* transporters prompted the identification of regulatory boxes involved in the response to Pi starvation (Schunmann et al., 2004b; Tittarelli et al., 2007; Karthikeyan et al., 2009; Thibaud et al., 2010). These studies revealed the presence of the P1BS motif, which is the binding site for the R2R3 Myb transcription

factor PHR1 (PHOSPHATE STARVATION RESPONSE 1, Rubio et al., 2001; Nilsson et al., 2007) in *Arabidopsis*. This protein is a homolog of *Chlamydomonas reinhardtii* protein PSR1 (PHOSPHORUS STARVATION RESPONSE 1, Wykoff et al., 1999). PHR1 belongs to a wide multigenic family that includes PHL1 (Bustos et al., 2010). Both PHR1 and PHL1 bind to the P1BS DNA motif (GNATATNC; Rubio et al., 2001), and play a crucial role in the Pi responsiveness of PHT1 transporters as demonstrated with both *phr1* and *phr1/phl1* mutants (Rubio et al., 2001; Bustos et al., 2010). This DNA binding motif appears to be conserved between monocots and dicots (Zhou et al., 2008). In addition, the P1BS motif participates in the systemically regulated *PHT1* response to internal Pi concentration, as revealed by the full genome analysis in split root experiments (Thibaud et al., 2010).

Additional regulatory sequences have been proposed from *in silico* analyses (including HLH, and NIT2 motifs highly present in the promoters of the *PHT1* genes), although no mechanism has been characterized yet (Mukatira et al., 2001). The genes *PHO1* and *PHO2* have also been identified as critical to the control of Pi homeostasis (Hamburger et al., 2002; Aung et al., 2006; Bari et al., 2006). The corresponding *pho1* and *pho2* mutants accumulate Pi in either roots or leaves, respectively (Poirier et al., 1991; Delhaize and Randall, 1995), and affect the transcription of PHT1;8 and PHT1;9 (Bari et al., 2006; Rouached et al., 2011). It is likely that PHO1 (membrane protein of unknown function) and PHO2 (E2 conjugase protein) act along with additional elements, since neither of these two proteins are transcription factors.

Several transcription factors have been found in addition to PHR1 that regulate PHT1 expression. For example, genetic dissection of the AtPHT1;4 promoter identified putative binding sequences for AtMyb2 and AtMyc2 (two positive regulators responding to water deprivation and abscisic acid stimuli) and also WRKY proteins (Karthikeyan et al., 2009), which are involved in biotic and abiotic stress responses. Reducing WRKY75 expression through RNAi silencing (Devaiah et al., 2007a) or the over-expression of MYB62 (Devaiah et al., 2009) produces plants that exhibit a decreased level of several PHT1 transporters, although there is no evidence for a direct interaction between these transcription factors and PHT1 promoters.

Further regulatory components were identified based on genetic evidence and transcriptomic analysis of Pi starvation. These include the zinc finger transcription factor ZAT6 (Devaiah et al., 2007b), the regulatory protein SPX3 (Duan et al., 2008), the nuclear actin-related protein ARP6, and the histone H2A.Z (Smith et al., 2010). These various proteins are reported to modulate *PHT1* genes (see Chiou and Lin, 2011 for a recent review of the Pi regulatory pathway), although a mechanism of action remains to be proposed.

Factors other than Pi concentration also influence PHT1 expression, including active photosynthesis or sugars (Jain et al., 2007; Karthikeyan et al., 2007; Hammond and White, 2008, 2011; Lei et al., 2011). Cytokinins have also been observed to prevent PHT1 induction during Pi deficiency; this action relies on the cytokinin receptors CRE1/AHK3 and AHK4 (Franco-Zorrilla et al., 2002, 2005). Many genes responding to Pi deficiency are stimulated by sugars or are enhanced in the *cre1/ahk4* mutant (regardless of cytokinin applications), suggesting cross-talk between Pi deficiency, sugars, and cytokinin (Franco-Zorrilla et al., 2005).

POST-TRANSCRIPTIONAL REGULATION OF PHT1

Beyond transcriptional control, PHT1 targeting and accumulation in the plasma membrane are also modulated by multiple steps of post-translational regulations. A mutant allele of the *PHF1* locus (PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR1) has been identified, which displays a severe decrease in Pi influx (Gonzalez et al., 2005). In accordance with its status as an endoplasmic reticulum (ER) resident protein, the loss of PHF1 results in the abnormal accumulation of PHT1;1, PHT1;2, and PHT1;4 in the ER, suggesting that PHF1 could be involved in intracellular trafficking of several PHT1 family proteins (Gonzalez et al., 2005; Bayle et al., 2011). These observations can explain how the *phf1-1* mutant shows a strong reduction (80%) of Pi uptake capacity compared to WT when grown in low Pi (Gonzalez et al., 2005; **Figure 4**). Recently, similar results were found in rice, where a mutant of OsPHF1 was shown to affect several members of PHT1 family, including the low affinity Pi transporter OsPT2 and the high affinity Pi transporter OsPT8 (Chen et al., 2011). A similar targeting defect was described for *axr4*, an auxin-resistant mutant that is affected in the expression of an ER accessory protein; this defect promotes the retention of the auxin influx facilitator AUX1 in the ER (Dharmasiri et al., 2006). These results suggest that several membrane proteins may require assistance in transiting through the ER compartment. PHF1 exhibits structural homologies with yeast SEC12 protein, an ER resident component involved in the formation of COPII vesicles (Gonzalez et al., 2005). These vesicles are involved in the export of newly synthesized cargo proteins, including PHT1, from the ER to the Golgi apparatus in eukaryotic cells. Although PHF1 shares structural homologies with SEC12, multiple amino acids that have been conserved between all SEC12 proteins are lost in the plant PHF1 homologs. This suggests a functional specialization that has diverged from that of SEC12 (Gonzalez et al., 2005). Indeed, cell biology experiments in *Arabidopsis* revealed the absence of co-localization between COPII markers and PHF1, demonstrating the independence of PHF1's role from COPII formation (Bayle et al., 2011).

Phospho-proteomic studies have determined that phosphorylation events occur at the C-end of PHT1 proteins when Pi is abundant in the environment (Nuhse et al., 2004; Hem et al., 2007). A phosphorylation-mimicking mutagenesis of PHT1;1 (at Ser 514, **Figure 1**) resulted in its accumulation in the ER (Bayle et al., 2011), implicating phosphorylation as an additional way for plant cells to modulate PHT1 exit from the ER. Phosphorylation events on serine residues located at the C-end of the PHT1;1 protein could provide additional negative charges in the vicinity of the putative ER export site (D/E-X-D/E), thereby altering the recognition of this motif and resulting in the accumulation of PHT1;1 in the ER (**Figure 1**). This mechanism is likely conserved between the different PHT1 transporters, since phosphorylation of other *Arabidopsis* members of this family (PHT1;4, PHT1;5; PHT1;7, PHT1;9) at the C-end of the protein has been reported (Bayle et al., 2011). This regulation recalls the previously identified phosphorylation

of the C-terminal serine residue during subcellular trafficking of PIP2;1 toward the plasma membrane (Prak et al., 2008). However, phosphorylation has distinct effects on these proteins. Whereas phosphorylation of PIP2;1 is required for its proper localization, in the case of PHT1;1, phosphorylation was found to prevent the protein from reaching its final destination.

Finally, a new type of post-transcriptional regulation has recently been identified, which specifically degrades plasma membrane PHT1 in the presence of phosphate (Bayle et al., 2011). This mechanism is dependent upon endocytosis and subsequent degradation of the protein in the vacuole. Similar regulatory mechanisms have been identified for many types of plasma membrane proteins in yeast, plants, and animals (Lagerstedt et al., 2002; Persson et al., 2003; Takano et al., 2005). As for the broader family of Pi transporters, this regulation has already been reported for the yeast PHO84 transporter (Lagerstedt et al., 2002; Persson et al., 2003). An important distinction here is that the regulatory mechanism identified in yeast requires the phosphorylation of an amino acid to observe this degradation phenomenon, whereas this particular amino acid is not conserved in plant PHT1 (Lundh et al., 2009). Thus, despite the high homologies between PHO84 and PHT1 proteins, similar regulatory mechanisms could rely on distinct types of machineries. This probably reflects adaptive steps used by plants to cope with a complex environment and a multicellular organization.

CONCLUSION

Great strides have been made in the characterization of the PHT1 family of Pi transporters since their identification in 1996. The fact that these transporters belong to broad multigenic family often exhibiting overlapping expression patterns limits the analysis of their role *in planta*. Another obstacle to the characterization of PHT1 proteins is their hydrophobic nature, which has impeded the study of their biochemistry and structural biology. This issue will eventually be surmounted as improved techniques in genetics, cell and molecular biology become more readily available. In view of the ever-expanding array of tools available to researchers, rapid advances are anticipated in our understanding of the finely regulated, sophisticated mechanisms of the PHT1 family.

ACKNOWLEDGMENTS

We thank Dr. Brandon Loveall (from the "Improvence" editing company) for English proof reading of the manuscript.

REFERENCES

Ai, P., Sun, S., Zhao, J., Fan, X., Xin, W., Guo, Q., Yu, L., Shen, Q., Wu, P., Miller, A. J., and Xu, G. (2008). Two rice phosphate transporters, OsPht1;2 and OsPht1;6, have different functions and kinetic properties in uptake and translocation. *Plant J.* 57, 798–809.

Aung, K., Lin, S. I., Wu, C. C., Huang, Y. T., Su, C. L., and Chiou, T. J. (2006). pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol.* 141, 1000–1011.

Baek, S. H., Chung, I. M., and Yun, S. J. (2001). Molecular cloning and characterization of a tobacco leaf cDNA encoding a phosphate transporter. *Mol. Cells* 11, 1–6.

Balestrini, R., Gomez-Ariza, J., Lanfranco, L., and Bonfante, P. (2007). Laser microdissection reveals that transcripts for five plant and one fungal phosphate transporter genes are contemporaneously present in arbusculated cells. *Mol. Plant Microbe Interact.* 20, 1055–1062.

Bari, R., Datt Pant, B., Stitt, M., and Scheible, W. R. (2006). PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol.* 141, 988–999.

Bayle, V., Arrighi, J. F., Creff, A., Nespoulous, C., Vialaret, J., Rossignol, M., Gonzalez, E., Paz-Ares, J., and Nussaume, L. (2011). *Arabidopsis thaliana* high-affinity phosphate transporters exhibit multiple levels of posttranslational regulation. *Plant Cell* 23, 1523–1535.

Bieleski, R. L. (1973). Phosphate pools, phosphate transport, and phosphate availability. *Annu. Rev. Plant Physiol.* 24, 225–252.

Bucher, M. (2007). Functional biology of plant phosphate uptake at root and mycorrhiza interfaces. *New Phytol.* 173, 11–26.

Bun-Ya, M., Nishimura, M., Harashima, S., and Oshima, Y. (1991). The PHO84 gene of *Saccharomyces cerevisiae* encodes an inorganic phosphate transporter. *Mol. Cell. Biol.* 11, 3229–3238.

Bustos, R., Castrillo, G., Linhares, F., Puga, M. I., Rubio, V., Perez-Perez, J., Solano, R., Leyva, A., and Paz-Ares, J. (2010). A central regulatory system largely controls transcriptional activation and repression responses to phosphate starvation in *Arabidopsis*. *PLoS Genet.* 6, e1001102. doi:10.1371/journal.pgen.1001102

Calderon-Vazquez, C., Ibarra-Laclette, E., Caballero-Perez, J., and Herrera-Estrella, L. (2008). Transcript profiling of *Zea mays* roots reveals gene responses to phosphate deficiency at the plant- and species-specific levels. *J. Exp. Bot.* 59, 2479–2497.

Casacuberta, N., Masque, P., and Garcia-Orellana, J. (2011). Fluxes of (238)U decay series radionuclides in a dicalcium phosphate industrial plant. *J. Hazard. Mater.* 190, 245–252.

Catarecha, P., Segura, M. D., Franco-Zorrilla, J. M., Garcia-Ponce, B., Lanza, M., Solano, R., Paz-Ares, J., and Leyva, A. (2007). A mutant of the *Arabidopsis* phosphate transporter PHT1;1 displays enhanced arsenic accumulation. *Plant Cell* 19, 1123–1133.

Chang, C. W., Moseley, J. L., Wykoff, D., and Grossman, A. R. (2005). The LPB1 gene is important for acclimation of *Chlamydomonas reinhardtii* to phosphorus and sulfur deprivation. *Plant Physiol.* 138, 319–329.

Chen, J. Y., Liu, Y., Ni, J., Wang, Y. F., Bai, Y. H., Shi, J., Gan, J., Wu, Z. C., and Wu, P. (2011). OsPHF1 regulates the plasma membrane localization of low- and high-affinity inorganic phosphate transporters and determines inorganic phosphate uptake and translocation in rice. *Plant Physiol.* 157, 269–278.

Chiou, T. J., and Lin, S. I. (2011). Signaling network in sensing phosphate availability in plants. *Annu. Rev. Plant Biol.* 62, 185–206.

Chiou, T. J., Liu, H., and Harrison, M. J. (2001). The spatial expression patterns of a phosphate transporter (MtPT1) from *Medicago truncatula* indicate a role in phosphate transport at the root/soil interface. *Plant J.* 25, 281–293.

Cogliati, D. H., and Clarkson, D. T. (1983). Physiological changes in, and phosphate uptake by potato plants during development of and recovery from phosphate deficiency. *Physiol. Plant* 58, 287–294.

Cordell, D., Drangert, J.-O., and White, S. (2009). The story of phosphorus: global food security and food for thought. *Glob. Environ. Change* 19, 292–305.

Da Conceicao, F. T., Antunes, M. L., and Durrant, S. F. (2011). Radionuclide concentrations in raw and purified phosphoric acids from Brazil and their processing wastes: implications for radiation exposures. *Environ. Geochem. Health.* doi: 10.1007/s10653-011-9394-2

Daram, P., Brunner, S., Persson, B. L., Amrhein, N., and Bucher, M. (1998). Functional analysis and cell-specific expression of a phosphate transporter from tomato. *Planta* 206, 225–233.

Delhaize, E., and Randall, P. J. (1995). Characterization of a phosphate-accumulator mutant of *Arabidopsis thaliana*. *Plant Physiol.* 107, 207–213.

Devaiah, B. N., Karthikeyan, A. S., and Raghothama, K. G. (2007a). WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *Arabidopsis*. *Plant Physiol.* 143, 1789–1801.

Devaiah, B. N., Nagarajan, V. K., and Raghothama, K. G. (2007b). Phosphate homeostasis and root development in *Arabidopsis* are synchronized by the zinc finger transcription factor ZAT6. *Plant Physiol.* 145, 147–159.

Devaiah, B. N., Madhuvanthi, R., Karthikeyan, A. S., and Raghothama, K. G. (2009). Phosphate starvation responses and gibberellic acid biosynthesis are regulated by the MYB62 transcription factor in *Arabidopsis*. *Mol. Plant* 2, 43–58.

Dharmasiri, S., Swarup, R., Mockaitis, K., Dharmasiri, N., Singh, S. K., Kowalchyk, M., Marchant, A., Mills, S., Sandberg, G., Bennett, M. J.,

and Estelle, M. (2006). AXR4 is required for localization of the auxin influx facilitator AUX1. *Science* 312, 1218–1220.

Drew, M. C., and Saker, L. R. (1984). Uptake and long-distance transport of phosphate, potassium and chloride in relation to internal ion concentrations in barley – evidence of non-allosteric regulation. *Planta* 160, 500–507.

Drew, M. C., Saker, L. R., Barber, S. A., and Jenkins, W. (1984). Changes in the kinetics of phosphate and potassium absorption in nutrient-deficient barley roots measured by a solution-depletion technique. *Planta* 160, 490–499.

Duan, K., Yi, K., Dang, L., Huang, H., Wu, W., and Wu, P. (2008). Characterization of a sub-family of *Arabidopsis* genes with the SPX domain reveals their diverse functions in plant tolerance to phosphorus starvation. *Plant J.* 54 965–975.

Dunlop, J., Phung, H. T., Meeking, R., and White, D. W. R. (1997). The kinetics associated with phosphate absorption by *Arabidopsis* and its regulation by phosphorus status. *Aust. J. Plant Physiol.* 24, 623–629.

Franco-Zorrilla, J. M., Martin, A. C., Leyva, A., and Paz-Ares, J. (2005). Interaction between phosphate-starvation, sugar, and cytokinin signaling in *Arabidopsis* and the roles of cytokinin receptors CRE1/AHK4 and AHK3. *Plant Physiol.* 138, 847–857.

Franco-Zorrilla, J. M., Martin, A. C., Solano, R., Rubio, V., Leyva, A., and Paz-Ares, J. (2002). Mutations at CRE1 impair cytokinin-induced repression of phosphate starvation responses in *Arabidopsis*. *Plant J.* 32, 353–360.

Furihata, T., Suzuki, M., and Sakurai, H. (1992). Kinetic characterization of two phosphate uptake systems with different affinities in suspension-cultured *Catharanthus roseus* protoplants. *Plant Cell Physiol.* 33, 1151–1157.

Gilbert, N. (2009). Environment: the disappearing nutrient. *Nature* 461, 716–718.

Glassop, D., Smith, S. E., and Smith, F. W. (2005). Cereal phosphate transporters associated with the mycorrhizal pathway of phosphate uptake into roots. *Planta* 222, 688–698.

Gonzalez, E., Solano, R., Rubio, V., Leyva, A., and Paz-Ares, J. (2005). Phosphate transporter traffic facilitator1 is a plant-specific SEC12-related protein that enables the endoplasmic reticulum exit of a high-affinity phosphate transporter in *Arabidopsis*. *Plant Cell* 17, 3500–3512.

Gordon-Weeks, R., Tong, Y., Davies, T. G., and Leggewie, G. (2003). Restricted spatial expression of a high-affinity phosphate transporter in potato roots. *J. Cell. Sci.* 116, 3135–3144.

Gross, M. (2010). Fears over phosphorus supplies. *Curr. Biol.* 20, R386–R387.

Hamburger, D., Rezzonico, E., Macdonald-Comber Petetot, J., Somerville, C., and Poirier, Y. (2002). Identification and characterization of the *Arabidopsis* PHO1 gene involved in phosphate loading to the xylem. *Plant Cell* 14, 889–902.

Hammond, J. P., Bennett, M. J., Bowen, H. C., Broadley, M. R., Eastwood, D. C., May, S. T., Rahn, C., Swarup, R., Woolaway, K. E., and White, P. J. (2003). Changes in gene expression in *Arabidopsis* shoots during phosphate starvation and the potential for developing smart plants. *Plant Physiol.* 132, 578–596.

Hammond, J. P., and White, P. J. (2008). Sucrose transport in the phloem: integrating root responses to phosphorus starvation. *J. Exp. Bot.* 59, 93–109.

Hammond, J. P., and White, P. J. (2011). Sugar signalling in root responses to low P availability. *Plant Physiol.* 156, 1033–1040.

Harrison, M. J., Dewbre, G. R., and Liu, J. (2002). A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* 14, 2413–2429.

Harrison, M. J., and Van Buuren, M. L. (1995). A phosphate transporter from the mycorrhizal fungus *Glomus versiforme*. *Nature* 378, 626–629.

Hata, S., Kobae, Y., and Banba, M. (2010). Interactions between plants and arbuscular mycorrhizal fungi. *Int. Rev. Cell Mol. Biol.* 281, 1–48.

Hem, S., Rofidal, V., Sommerer, N., and Rossignol, M. (2007). Novel subsets of the *Arabidopsis* plasmalemma phosphoproteome identify phosphorylation sites in secondary active transporters. *Biochem. Biophys. Res. Commun.* 363, 375–380.

Hirsch, J., Marin, E., Floriani, M., Chiarenza, S., Richaud, P., Nussaume, L., and Thibaud, M. C. (2006). Phosphate deficiency promotes modification of iron distribution in *Arabidopsis* plants. *Biochimie* 88, 1767–1771.

Hirsch, J., Misson, J., Crisp, P. A., David, P., Bayle, V., Estavillo, G. M., Javot, H., Chiarenza, S., Mallory, A. C., Maizel, A., Declerck, M., Pogson, B. J., Vaucheret, H., Crespi, M., Desnos, T., Thibaud, M. C., Nussaume, L., and Marin, E. (2011). A novel fry1 allele reveals the existence of a mutant phenotype unrelated to 5'→3' exoribonuclease (XRN) activities in *Arabidopsis thaliana* roots. *PLoS ONE* 6, e16724. doi:10.1371/journal.pone.0016724

Jain, A., Poling, M. D., Karthikeyan, A. S., Blakeslee, J. J., Peer, W. A., Titapiwatanakun, B., Murphy, A. S., and Raghothama, K. G. (2007). Differential effects of sucrose and auxin on localized phosphate deficiency-induced modulation of different traits of root system architecture in *Arabidopsis*. *Plant Physiol.* 144, 232–247.

Jain, A., Poling, M. D., Smith, A. P., Nagarajan, V. K., Lahner, B., Meagher, R. B., and Raghothama, K. G. (2009). Variations in the composition of gelling agents affect morphophysiological and molecular responses to deficiencies of phosphate and other nutrients. *Plant Physiol.* 150, 1033–1049.

Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., and Harrison, M. J. (2007a). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1720–1725.

Javot, H., Pumplin, N., and Harrison, M. J. (2007b). Phosphate in the arbuscular mycorrhizal symbiosis: transport properties and regulatory roles. *Plant Cell Environ.* 30, 310–322.

Jia, H., Ren, H., Gu, M., Zhao, J., Sun, S., Zhang, X., Chen, J., Wu, P., and Xu, G. (2011). The phosphate transporter gene OsPht1;8 is involved in phosphate homeostasis in rice. *Plant Physiol.* 156, 1164–1175.

Jouhet, J., Marechal, E., and Block, M. A. (2007). Glycerolipid transfer for the building of membranes in plant cells. *Prog. Lipid Res.* 46, 37–55.

Kai, M., Masuda, Y., Kikuchi, Y., Osaki, M., and Tadano, T. (1997). Isolation and characterization of a cDNA from *Catharanthus roseus* which is highly homologous with phosphate transporter. *Soil Sci. Plant Nutr.* 43, 227–235.

Kanno, S., Rai, H., Ohya, T., Hayashi, Y., Tanoi, K., and Nakanishi, T. M. (2007). Real-time imaging of radioisotope labeled compounds in a living plant. *J. Radioanal. Nucl. Chem.* 272, 565–570.

Karthikeyan, A. S., Ballachanda, D. N., and Raghothama, K. G. (2009). Promoter deletion analysis elucidates the role of cis elements and 5'UTR intron in spatiotemporal regulation of AtPht1;4 expression in *Arabidopsis*. *Physiol. Plant* 136, 10–18.

Karthikeyan, A. S., Varadarajan, D. K., Jain, A., Held, M. A., Carpita, N. C., and Raghothama, K. G. (2007). Phosphate starvation responses are mediated by sugar signaling in *Arabidopsis*. *Planta* 225, 907–918.

Karthikeyan, A. S., Varadarajan, D. K., Mukatira, U. T., D'Urzo, M. P., Damsz, B., and Raghothama, K. G. (2002). Regulated expression of *Arabidopsis* phosphate transporters. *Plant Physiol.* 130, 221–233.

Kobae, Y., and Hata, S. (2010). Dynamics of periarbuscular membranes visualized with a fluorescent phosphate transporter in arbuscular mycorrhizal roots of rice. *Plant Cell Physiol.* 51, 341–353.

Koyama, T., Ono, T., Shimizu, M., Jinbo, T., Mizuno, R., Tomita, K., Mitsukawa, N., Kawazu, T., Kimura, T., Ohmiya, K., and Sakka, K. (2005). Promoter of *Arabidopsis thaliana* phosphate transporter gene drives root-specific expression of transgene in rice. *J. Biosci. Bioeng.* 99, 38–42.

Lagerstedt, J. O., Zvyagilskaya, R., Pratt, J. R., Pattison-Granberg, J., Kruckeberg, A. L., Berden, J. A., and Persson, B. L. (2002). Mutagenic and functional analysis of the C-terminus of *Saccharomyces cerevisiae* Pho84 phosphate transporter. *FEBS Lett.* 526, 31–37.

Lee, D. A., Chen, A., and Schroeder, J. I. (2003). Ars1, an *Arabidopsis* mutant exhibiting increased tolerance to arsenate and increased phosphate uptake. *Plant J.* 35, 637–646.

Lefebvre, D. D., Duff, S. M., Fife, C. A., Julien-Inalsingh, C., and Plaxton, W. C. (1990). Response to phosphate deprivation in *Brassica nigra* suspension cells: enhancement of intracellular, cell surface, and secreted phosphatase activities compared to increases in pi-absorption rate. *Plant Physiol.* 93, 504–511.

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

Lei, M., Liu, Y., Zhang, B., Zhao, Y., Wang, X., Zhou, Y., Raghothama, K. G., and Liu, D. (2011). Genetic and genomic evidence that sucrose is a global regulator of plant response to phosphate starvation in *Arabidopsis*. *Plant Physiol.* 156, 1116–1130.

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

Liu, H., Trieu, A. T., Blaylock, L. A., and Harrison, M. J. (1998b). Cloning and characterization of two phosphate transporters from *Medicago truncatula* roots: regulation in response to phosphate and to colonization by arbuscular mycorrhizal (AM) fungi. *Mol. Plant Microbe Interact.* 11, 14–22.

Liu, J., Uhde-Stone, C., Li, A., Vance, C., and Allan, D. (2001). A phosphate transporter with enhanced expression on proteoid roots of white lupin (*Lupinus albus* L.). *Plant Soil* 237, 257–266.

Liu, J., Versaw, W. K., Pumplin, N., Gomez, S. K., Blaylock, L. A., and Harrison, M. J. (2008). Closely related members of the *Medicago truncatula* PHT1 phosphate transporter gene family encode phosphate transporters with distinct biochemical activities. *J. Biol. Chem.* 283, 24673–24681.

Loque, D., Lalonde, S., Looger, L. L., von Wiren, N., and Frommer, W. B. (2007). A cytosolic trans-activation domain essential for ammonium uptake. *Nature* 446, 195–198.

Loth-Pereda, V., Orsini, E., Courty, P. E., Lota, F., Kohler, A., Diss, L., Blaudez, D., Chalot, M., Nehls, U., Bucher, M., and Martin, F. (2011). Structure and expression profile of the phosphate Pht1 transporter gene family in mycorrhizal *Populus trichocarpa*. *Plant Physiol.* 156, 2141–2154.

Ludewig, U., Wilken, S., Wu, B., Jost, W., Obrdlik, P., El Bakkoury, M., Marini, A. M., Andre, B., Hamacher, T., Boles, E., von Wiren, N., and Frommer, W. B. (2003). Homo- and hetero-oligomerization of ammonium transporter-1 NH4 uniporters. *J. Biol. Chem.* 278, 45603–45610.

Lundh, F., Mouillon, J. M., Samyn, D., Stadler, K., Popova, Y., Lagerstedt, J. O., Thevelein, J. M., and Persson, B. L. (2009). Molecular mechanisms controlling phosphate-induced downregulation of the yeast Pho84 phosphate transporter. *Biochemistry* 48, 4497–4505.

Ming, F., Mi, G. H., Lu, Q., Yin, S., Zhang, S. S., Guo, B., and Shen, D. L. (2005). Cloning and characterization of cDNA for the *Oryza sativa* phosphate transporter. *Cell. Mol. Biol. Lett.* 10, 401–411.

Misson, J., Raghothama, K. G., Jain, A., Jouhet, J., Block, M. A., Bligny, R., Ortet, P., Creff, A., Somerville, S., Rolland, N., Doumas, P., Nacry, P., Herrerra-Estrella, L., Nussaume, L., and Thibaud, M. C. (2005). A genome-wide transcriptional analysis using *Arabidopsis thaliana* affymetrix gene chips determined plant responses to phosphate deprivation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11934–11939.

Misson, J., Thibaud, M. C., Bechtold, N., Raghothama, K., and Nussaume, L. (2004). Transcriptional regulation and functional properties of *Arabidopsis* Pht1;4, a high affinity transporter contributing greatly to phosphate uptake in phosphate deprived plants. *Plant Mol. Biol.* 55, 727–741.

Mitsukawa, N., Okumura, S., Shirano, Y., Sato, S., Kato, T., Harashima, S., and Shibata, D. (1997). Overexpression of an *Arabidopsis thaliana* high-affinity phosphate transporter gene in tobacco cultured cells enhances cell growth under phosphate-limited conditions. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7098–7102.

Morcuende, R., Bari, R., Gibon, Y., Zheng, W., Pant, B. D., Blasing, O., Usadel, B., Czechowski, T., Udvardi, M. K., Stitt, M., and Scheible, W. R. (2007). Genome-wide reprogramming of metabolism and regulatory networks of *Arabidopsis* in response to phosphorus. *Plant Cell Environ.* 30, 85–112.

Muchhal, U. S., Pardo, J. M., and Raghothama, K. G. (1996). Phosphate transporters from the higher plant *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 93, 10519–10523.

Muchhal, U. S., and Raghothama, K. G. (1999). Transcriptional regulation of plant phosphate transporters. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5868–5872.

Mudge, S. R., Rae, A. L., Diatloff, E., and Smith, F. W. (2002). Expression analysis suggests novel roles for members of the Pht1 family of phosphate transporters in *Arabidopsis*. *Plant J.* 31, 341–353.

Mukatira, U. T., Liu, C., Varadarajan, D. K., and Raghothama, K. G. (2001). Negative regulation of phosphate starvation-induced genes. *Plant Physiol.* 127, 1854–1862.

Nagarajan, V. K., Jain, A., Poling, M. D., Lewis, A. J., Raghothama, K. G., and Smith, A. P. (2011). *Arabidopsis* Pht1;5 mobilizes phosphate between source and sink organs, and influences the interaction between phosphate homeostasis and ethylene signaling. *Plant Physiol.* 156, 1149–1163.

Nagy, R., Karandashov, V., Chague, V., Kalinkevich, K., Tamasloukht, M., Xu, G., Jakobsen, I., Levy, A. A., Amrhein, N., and Bucher, M. (2005). The characterization of novel mycorrhiza-specific phosphate transporters from *Lycopersicon esculentum* and *Solanum tuberosum* uncovers functional redundancy in symbiotic phosphate transport in solanaceous species. *Plant J.* 42, 236–250.

Nagy, R., Vasconcelos, M. J., Zhao, S., Mcelver, J., Bruce, W., Amrhein, N., Raghothama, K. G., and Bucher, M. (2006). Differential regulation of five Pht1 phosphate transporters from maize (*Zea mays* L.). *Plant Biol. (Stuttg.)* 8, 186–197.

Nakamori, K., Takabatake, R., Umehara, Y., Kouchi, H., Izui, K., and Hata, S. (2002). Cloning, functional expression, and mutational analysis of a cDNA for *Lotus japonicus* mitochondrial phosphate transporter. *Plant Cell Physiol.* 43, 1250–1253.

Nilsson, L., Muller, R., and Nielsen, T. H. (2007). Increased expression of the MYB-related transcription factor, PHR1, leads to enhanced phosphate uptake in *Arabidopsis thaliana*. *Plant Cell Environ.* 30, 1499–1512.

Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2004). Phosphoproteomics of the *Arabidopsis* plasma membrane and a new phosphorylation site database. *Plant Cell* 16, 2394–2405.

Okumura, S., Mitsukawa, N., Shirano, Y., and Shibata, D. (1998). Phosphate transporter gene family of *Arabidopsis thaliana*. *DNA Res.* 5, 261–269.

Othman, I., and Al-Masri, M. S. (2007). Impact of phosphate industry on the environment: a case study. *Appl. Radiat. Isot.* 65, 131–141.

Paszkowski, U., Kroken, S., Roux, C., and Briggs, S. P. (2002). Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13324–13329.

Persson, B. L., Lagerstedt, J. O., Pratt, J. R., Pattison-Granberg, J., Lundh, K., Shokrollahzadeh, S., and Lundh, F. (2003). Regulation of phosphate acquisition in *Saccharomyces cerevisiae*. *Curr. Genet.* 43, 225–244.

Poirier, Y., and Bucher, M. (2002). "Phosphate transport and homeostasis in *Arabidopsis*," in *The Arabidopsis Book*, eds C. R. Somerville and E. M. Meyerowitz (Rockville, MD: American Society of Plant Biologists), 1–35.

Poirier, Y., Thoma, S., Somerville, C., and Schiefelbein, J. (1991). Mutant of *Arabidopsis* deficient in xylem loading of phosphate. *Plant Physiol.* 97, 1087–1093.

Prak, S., Hem, S., Boudet, J., Viennois, G., Sommerer, N., Rossignol, M., Maurel, C., and Santoni, V. (2008). Multiple phosphorylations in the C-terminal tail of plant plasma membrane aquaporins: role in subcellular trafficking of AtPIP2;1 in response to salt stress. *Mol. Cell. Proteomics* 7, 1019–1030.

Preuss, C. P., Huang, C. Y., and Tyerman, S. D. (2011). Proton-coupled high-affinity phosphate transport revealed from heterologous characterization in *Xenopus* of barley-root plasma membrane transporter, HvPHT1;1. *Plant Cell Environ.* 34, 681–689.

Pumplin, N., and Harrison, M. J. (2009). Live-cell imaging reveals periarbuscular membrane domains and organelle location in *Medicago truncatula* roots during arbuscular mycorrhizal symbiosis. *Plant Physiol.* 151, 809–819.

Rae, A. L., Cybinski, D. H., Jarmey, J. M., and Smith, F. W. (2003). Characterization of two phosphate transporters from barley; evidence for diverse function and kinetic properties among members of the Pht1 family. *Plant Mol. Biol.* 53, 27–36.

Raghothama, K. G. (1999). Phosphate acquisition. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 665–693.

Rausch, C., and Bucher, M. (2002). Molecular mechanisms of phosphate transport in plants. *Planta* 216, 23–37.

Richardson, A. E. (1994). "Soil microorganisms and phosphorus availability," in *Soil Biota: Management in Sustainable Farming Systems*, eds C. E. Pankhurst, B. M. Doube, V. V. S. R. Gupta, and P. R. Grace, 50–62.

Rouached, H., Stefanovic, A., Secco, D., BulakArpat, A., Gout, E., Bligny, R., and Poirier, Y. (2011). Uncoupling phosphate deficiency from its major effects on growth and transcriptome via PHO1 expression in *Arabidopsis*. *Plant J.* 65, 557–570.

Rubio, V., Linhares, F., Solano, R., Martin, A. C., Iglesias, J., Leyva, A., and Paz-Ares, J. (2001). A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev.* 15, 2122–2133.

Sakano, K. (1990). Proton/phosphate stoichiometry in uptake of inorganic phosphate by cultured cells of *Catharanthus roseus* (L.) G. Don. *Plant Physiol.* 93, 479–483.

Schachtman, D. P., Reid, R. J., and Ayling, S. M. (1998). Phosphorus uptake by plants: from soil to cell. *Plant Physiol.* 116, 447–453.

Schunmann, P. H., Richardson, A. E., Smith, F. W., and Delhaize, E. (2004a). Characterization of promoter expression patterns derived from the Pht1 phosphate transporter genes of barley (*Hordeum vulgare* L.). *J. Exp. Bot.* 55, 855–865.

Schunmann, P. H., Richardson, A. E., Vickers, C. E., and Delhaize, E. (2004b). Promoter analysis of the barley Pht1;1 phosphate transporter gene identifies regions controlling root expression and responsiveness to phosphate deprivation. *Plant Physiol.* 136, 4205–4214.

Shimogawara, K., and Usuda, H. (1995). Uptake of inorganic phosphate by suspension-cultured tobacco cells: kinetics and regulation by Pi starvation. *Plant Cell Physiol.* 36, 341–351.

Shin, H., Shin, H. S., Dewbre, G. R., and Harrison, M. J. (2004). Phosphate transport in *Arabidopsis*: Pht1;1 and Pht1;4 play a major role in phosphate acquisition from both low- and high-phosphate environments. *Plant J.* 39, 629–642.

Smith, A. P., Jain, A., Deal, R. B., Nagarajan, V. K., Poling, M. D., Raghothama, K. G., and Meagher, R. B. (2010). Histone H2A.Z regulates the expression of several classes of phosphate starvation response genes but not as a transcriptional activator. *Plant Physiol.* 152, 217–225.

Smith, F. W., Cybinski, D. H., and Rae, A. L. (eds). (1999). *Regulations of Expression of Genes Encoding Phosphate Transporters in Barley Roots.* Dordrecht: Kluwer Academic Publishers.

Smith, F. W., Ealing, P. M., Dong, B., and Delhaize, E. (1997). The cloning of two *Arabidopsis* genes belonging to a phosphate transporter family. *Plant J.* 11, 83–92.

Smith, F. W., Rae, A. L., and Hawkesford, M. J. (2000). Molecular mechanisms of phosphate and sulphate transport in plants. *Biochim. Biophys. Acta* 1465, 236–245.

Smith, S. E., and Barker, S. J. (2002). Plant phosphate transporter genes help harness the nutritional benefits of arbuscular mycorrhizal symbiosis. *Trends Plant Sci.* 7, 189–190.

Smith, S. E., and Read, D. J. (1997). *Mycorrhizal Symbiosis*, 2nd Edn. San Diego, CA: Academic Press.

Stefanovic, A., Arpat, A. B., Bligny, R., Gout, E., Vidoudez, C., Bensimon, M., and Poirier, Y. (2011). Over-expression of PHO1 in *Arabidopsis* leaves reveals its role in mediating phosphate efflux. *Plant J.* 66, 689–699.

Takano, J., Miwa, K., Yuan, L. X., Von Wiren, N., and Fujiwara, T. (2005). Endocytosis and degradation of BOR1, a boron transporter of *Arabidopsis thaliana*, regulated by boron availability. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12276–12281.

Thibaud, M. C., Arrighi, J. F., Bayle, V., Chiarenza, S., Creff, A., Bustos, R., Paz-Ares, J., Poirier, Y., and Nussaume, L. (2010). Dissection of local and systemic transcriptional responses to phosphate starvation in *Arabidopsis*. *Plant J.* 64, 775–789.

Ticconi, C. A., Delatorre, C. A., and Abel, S. (2001). Attenuation of phosphate starvation responses by phosphite in *Arabidopsis*. *Plant Physiol.* 127, 963–972.

Tittarelli, A., Milla, L., Vargas, F., Morales, A., Neupert, C., Meisel, L. A., Salvo, G. H., Penaloza, E., Munoz, G., Corcuera, L. J., and Silva, H. (2007). Isolation and comparative analysis of the wheat TaPT2 promoter: identification in silico of new putative regulatory motifs conserved between monocots and dicots. *J. Exp. Bot.* 58, 2573–2582.

Ullrich, C. I., and Novacky, A. J. (1990). Extra- and intracellular pH and membrane potential changes induced by K, Cl, H(2)PO(4), and NO(3) uptake and fusicoccin in root hairs of *Limnobium stoloniferum*. *Plant Physiol.* 94, 1561–1567.

Ullrich-Eberius, C. I., Novacky, A., Fischer, E., and Luttge, U. (1981). Relationship between energy-dependent phosphate uptake and the electrical membrane potential in *Lemna gibba* G1. *Plant Physiol.* 67, 797–801.

Ullrich-Eberius, C. I., Novacky, A., and Van Bel, A. J. E. (1984). Phosphate uptake in *Lemna gibba* G1: energetics and kinetics. *Planta* 161, 46–52.

Vance, C. P., Uhde-Stone, C., and Allan, D. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New Phytol.* 157, 423–447.

Varadarajan, D. K., Karthikeyan, A. S., Matilda, P. D., and Raghothama, K. G. (2002). Phosphite, an analog of phosphate, suppresses the coordinated expression of genes under phosphate starvation. *Plant Physiol.* 129, 1232–1240.

Versaw, W. K. (1995). A phosphate-repressible, high-affinity phosphate permease is encoded by the pho-5+ gene of *Neurospora crassa*. *Gene* 153, 135–139.

Versaw, W. K., and Metzenberg, R. L. (1995). Repressible cation-phosphate symporters in *Neurospora crassa*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3884–3887.

Von Vexhull, H. R., and Mutert, E. (1998). "Global extent, development and economic impact of acid soils," in *Plant-Soil Interactions at Low pH: Principles and Management*, eds R. A. Date, N. J. Grundon, G. E. Rayment, and M. E. Probert (Dordrecht: Kluwer), 5–19.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2, e718. doi:10.1371/journal.pone.0000718

Wirth, J., Chopin, F., Santoni, V., Viennois, G., Tillard, P., Krapp, A., Lejay, L., Daniel-Vedele, F., and Gojon, A. (2007). Regulation of root nitrate uptake at the NRT2.1 protein level in *Arabidopsis thaliana*. *J. Biol. Chem.* 32, 23541–23552.

Wu, P., Ma, L., Hou, X., Wang, M., Wu, Y., Liu, F., and Deng, X. W. (2003). Phosphate starvation triggers distinct alterations of genome expression in *Arabidopsis* roots and leaves. *Plant Physiol.* 132, 1260–1271.

Wu, Z., Ren, H., McGrath, S. P., Wu, P., and Zhao, F.-J. (2011). Investigating the contribution of the phosphate transport pathway to arsenic accumulation in rice. *Plant Physiol.* 157, 498–508.

Wykoff, D. D., Grossman, A. R., Weeks, D. P., Usuda, H., and Shimogawara, K. (1999). Psr1, a nuclear localized protein that regulates phosphorus metabolism in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 15336–15341.

Xiao, K., Liu, J., Dewbre, G., Harrison, M., and Wang, Z. Y. (2006). Isolation and characterization of root-specific phosphate transporter promoters from *Medicago truncatula*. *Plant Biol. (Stuttg.)* 8, 439–449.

Xu, G. H., Chague, V., Melamed-Bessudo, C., Kapulnik, Y., Jain, A., Raghothama, K. G., Levy, A. A., and Silber, A. (2007). Functional characterization of LePT4: a phosphate transporter in tomato with mycorrhiza-enhanced expression. *J. Exp. Bot.* 58, 2491–2501.

Zhou, J., Jiao, F., Wu, Z., Li, Y., Wang, X., He, X., Zhong, W., and Wu, P. (2008). OsPHR2 is involved in phosphate-starvation signaling and excessive phosphate accumulation in shoots of plants. *Plant Physiol.* 146, 1673–1686.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 August 2011; accepted: 03 November 2011; published online: 30 November 2011.

Citation: Nussaume L, Kanno S, Javot H, Marin E, Pochon N, Ayadi A, Nakanishi TM and Thibaud M-C (2011) Phosphate import in plants: focus on the PHT1 transporters. Front. Plant Sci. ***2****:83. doi: 10.3389/fpls.2011.00083*

This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 05 January 2012
doi: 10.3389/fpls.2011.00108

Plant lessons: exploring ABCB functionality through structural modeling

***Aurélien Bailly*[1,2], *Haibing Yang*[3], *Enrico Martinoia*[2], *Markus Geisler*[1,2]* *and Angus S. Murphy*[3]**

[1] *Plant Biology, Department of Biology, University of Fribourg, Fribourg, Switzerland*
[2] *Institute of Plant Biology, Zurich–Basel Plant Science Center, University of Zurich, Zurich, Switzerland*
[3] *Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA*

Edited by:
Heven Sze, University of Maryland, USA

Reviewed by:
Fatima Cvrckova, Charles University in Prague, Czech Republic
Frantisek Baluska, University of Bonn, Germany

***Correspondence:**
Markus Geisler, Plant Biology, Department of Biology, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland.
e-mail: markus.geisler@unifr.ch

In contrast to mammalian ABCB1 proteins, narrow substrate specificity has been extensively documented for plant orthologs shown to catalyze the transport of the plant hormone, auxin. Using the crystal structures of the multidrug exporters Sav1866 and MmABCB1 as templates, we have developed structural models of plant ABCB proteins with a common architecture. Comparisons of these structures identified kingdom-specific candidate substrate-binding regions within the translocation chamber formed by the transmembrane domains of ABCBs from the model plant *Arabidopsis*. These results suggest an early evolutionary divergence of plant and mammalian ABCBs. Validation of these models becomes a priority for efforts to elucidate ABCB function and manipulate this class of transporters to enhance plant productivity and quality.

Keywords: ABCB exporter, ABCB importer, structural modeling, auxin, substrate docking

INTRODUCTION

Members of the ATP-binding cassette (ABC) transphyletic protein superfamily are predominantly transporters that function in movement of a wide variety of substrates across cellular membranes. The B subclass of this superfamily comprises the eukaryotic P-glycoproteins (PGPs; Hrycyna and Gottesman, 1998; Higgins, 2001). The most notorious ABCB transporter is human P-glycoprotein/multiple drug resistance1 (P-GP, MDR1, HsABCB1) due to its contribution to cellular resistance toward multiple cytotoxic chemotherapeutic agents when overexpressed in tumor cells (O'Connor et al., 2007). Elucidation of the mechanistic basis of ABCB multi-substrate specificity has been a longstanding biomedical research priority (Borowski et al., 2005) and has been accelerated by the publication of a high-resolution crystal structure of murine ABCB1 complexed with a cyclic tetrapeptide inhibitor (3G5U, 3G60, 3G61) in a ligand-binding conformation (Aller et al., 2009). Publication of detailed structures of the bacterial ABC transporters Sav1866, MsbA, and MalGFK (Dawson and Locher, 2006, 2007; Dawson et al., 2007; Oldham et al., 2007; Ward et al., 2007; Aller et al., 2009; Oldham and Chen, 2011a) have provided important insights into the conserved mechanisms of ATP hydrolysis, domain organization, and membrane interactions in ABC transporters. The crystal structures of MmABCB1 and Sav1866 represent the putative ligand-binding and ligand-releasing states of the transport cycle, thus allowing modeling of both states in the context of the commonly accepted ABCB export mechanism (Li et al., 2010). Increasingly, these crystal structures are utilized by researchers to inform experimental analyses of specific details of hydrophobic substrate translocation in mammalian ABCBs.

Plant genomes contain an expanded family of ABC-transporter genes: over 120, compared to 50–60 in other organisms of equivalent genome size (Theodoulou, 2000; Sanchez-Fernandez et al., 2001a,b; Martinoia et al., 2002; Jasinski et al., 2003). Despite their high numerous representation, only a few ABCB orthologs have been extensively characterized in plants and shown to catalyze the transport of structurally diverse substrates, such as phytohormones, secondary metabolites, and xenobiotics (Geisler et al., 2005; Sugiyama et al., 2006; Knoller et al., 2010). However, the most extensive analysis of ABCB function has taken place in the model plant *Arabidopsis* and has focused on AtABCB1 (PGP1) and AtABCB19 (PGP19/MDR1; Geisler et al., 2005; Geisler and Murphy, 2006). Lesions in the genes encoding these proteins result in reductions in long distance transport of the phytohormone auxin and consequent dwarfism in mutant plants (**Figure A1** in Appendix). Most notably, the agriculturally important *brachytic2* and *dwarf3* mutants in maize and sorghum were shown to result from loss-of-function mutations in ABCB1 genes (**Figures A1A,B** in Appendix; Multani et al., 2003). AtABCB1 and AtABCB19 do not transport standard HsABCB1 substrates, even when overexpressed in human HeLa cells, but directly transport the natural auxin indole-3-acetic acid (IAA) and the artificial auxin 1-NAA (AtABCB1 transports some additional artificial auxins, and, to a lesser extent, auxin-like compounds; Noh et al., 2001; Geisler et al., 2003, 2005; Bouchard et al., 2006; Bailly et al., 2008). A third ABCB transporter, AtABCB4, is a conditional auxin im/exporter with specificity for auxins similar to ABCB1 (Santelia et al., 2005; Terasaka et al., 2005; Kubes et al., 2011). AtABCB14 that is included in this study as non-auxin transporting ABCB is a malate/citrate transporter functioning in *Arabidopsis* guard cells (Lee et al., 2008).

Abbreviations: ABCB, ATP-binding cassette protein subfamily B; ADP, adenosine-5′-diphosphate; ATP, adenosine-5′-triphosphate; IAA, indole-3-acetic acid; ICL, intracellular loop; MDR, multidrug resistance; NMD, nucleotide-binding domain; PGP, P-glycoprotein; TM, transmembrane; TMD, transmembrane domain.

The relative substrate specificity of these plant transporters in comparison to mammalian orthologs suggests that plant transporters of the ABCB1/19 subgroup contain either very different amino acid compositions in substrate-binding sites identified in HsABCB1 or additional sites that dictate narrower substrate specificity (Shapiro and Ling, 1997). However, amino acid sequences variation associated with these differential specificities are not readily deduced from phylogenetic sequence comparisons alone (Knoller et al., 2010). Robust comparative homology modeling and sequence analysis of a variety of plant and mammalian ABCB1 proteins with dissimilar substrate specificities is therefore a strategy of choice to predict kingdom-specific ligand-recognition patterns within the protein subfamily. The results presented here support evolutionary divergence within the recently proposed substrate-binding domains and indicate that electrostatic changes in surface residues within the translocation chamber dictate substrate specificity.

MATERIALS AND METHODS

HOMOLOGY MODELING OF ABCB STRUCTURES

The high-resolution P-glycoprotein *Mus musculus* ABCB1 (MmABCB1) structure (3G5U), representing the ligand-binding competent conformation, was utilized as a homology modeling template for *Arabidopsis* ABCB1, 19, 4, and 14 transporters (GenBank accession numbers: ABCB1:NP_181228, ABCB4: NP_182223, ABCB14: NP_174122, ABCB19: NP_189528). The sequences of AtABCBs and MmABCB1 were used to generate a multiple alignment with MultAlin (Corpet, 1988). The alignments used to build the models are shown in Table S2 in Supplementary Material. The Modeller9v7 was used to generate the AtABCB models based on the alignments and MmABCB1 crystal structure (Sali and Blundell, 1993; Eswar et al., 2006). The N-terminal and linker regions of AtABCBs connecting nucleotide-binding domain (NBD) 1 with transmembrane domain (TMD) 2 were removed since these regions in MmABCB1 are missing in the crystal structure. A second set of homology models was generated using the nucleotide-bound state Sav1866 crystal structure as template (2HYD). AtABCB4, 14, and 19 models were described in Yang and Murphy (2009). Each half of ABCBs were aligned with Sav1866 using BLAST2[1] and MultAlin (Corpet, 1988). The N-terminal and linker regions were used to blast protein sequence by selecting Protein Data Bank proteins as database. The most similar structures were chosen for templates for N-terminal and linker regions. As an example, the complete template alignments for HsABCB1 are shown in Table S3 in Supplementary Material in two runs of modeling. For each ABCB model, five models were generated and the ones with best molecular probability density function (molpdf) and discrete optimized protein energy (DOPE) scores were selected. The evaluation of the selected models was carried out by analyses of ERRAT (Colovos and Yeates, 1993), Qmean (Benkert et al., 2008), PROCHECK (Laskowski et al., 1993), and WHAT_CHECK (Hooft et al., 1996). The TMDs and NBDs regions in both Sav1866- and MmABCB1-based models showed good scores in assessment with these analyses while the N-terminal, linker, and C-terminal regions in Sav1866-based models showed less satisfied scores in these analyses. These regions were not modeled in MmABCB1-based models due to difficulties to model these regions without disturbing the flexible NBDs. All ABCB models from this study are available as PDB files as listed in the Supplementary Material information.

ALIGNMENTS AND STRUCTURE DISPLAY

Multiple sequence alignments of the primary polypeptide sequences of identified HsABCB1 and AtABCB1 orthologs (listed in Table S1 in Supplementary Material) and the generation of bootstrapped N-J trees were performed using ClustalX v2.1 (Larkin et al., 2007) in its default settings. In order to map the relative degree of residue conservation onto protein model surfaces, the ConSurf server tool[2] (Ashkenazy et al., 2010) has been next used with the relevant PDB templates or generated models and guiding trees, employing the JTT evolutionary substitution model and the Bayesian calculation method. Output figures have been generated using PyMOL v1.3[3].

SUBSTRATE DOCKING TO ABCB TMDs

For the docking process 35 different auxins and auxin-related compounds have been selected according to their different structure–effect relationships in hormonal activity (Ferro et al., 2006, 2010) and were docked into both homology models. Thousand poses for IAA and 100 poses for each of the other compounds were generated using the PyMOL embedded AutoDock Vina toolset (Seeliger and de Groot, 2010). In order to avoid any bias, the binding site was defined either as the whole molecule or the complete transmembrane region for each model conformer. This region in AtABCB1 shows patterns of polar and apolar residues suitable for IAA binding (**Figure 2**; Tan et al., 2007; Ferro et al., 2010) and, since there is indication that the protein's translocation chamber is water filled during the catalytic cycle (Gutmann et al., 2010), the ligands were docked in their ionized state.

ELECTROSTATICS

Electrostatic surfaces were generated for both inward- and outward-facing conformations of the structural models using the APBS Tools2 plugin (Baker et al., 2001) implemented within PyMOL. The solving of the non-linear Poisson–Boltzmann equation had been performed with the package default settings and molecular surfaces were colored by the potential on solvent accessible surface using a temperature spectrum ranging from −10 to +10 kT.

RESULTS

PREDICTIONS FROM STRUCTURAL MODELING

ABCB proteins share a common architecture

ATP-binding cassette transporters share a common architecture consisting of two TMDs and two cytosolic NBDs (constituted by a dimer in the "half transporter" subgroup). Models derived from high-resolution crystal structures of MmABCB1, Sav1866, and $MalGFK_2$ are shown in **Figures 1A–C** (Dawson and Locher, 2006, 2007; Dawson et al., 2007; Oldham et al., 2007; Aller et al., 2009;

[1] http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi

[2] http://consurftest.tau.ac.il

[3] http://www.pymol.org

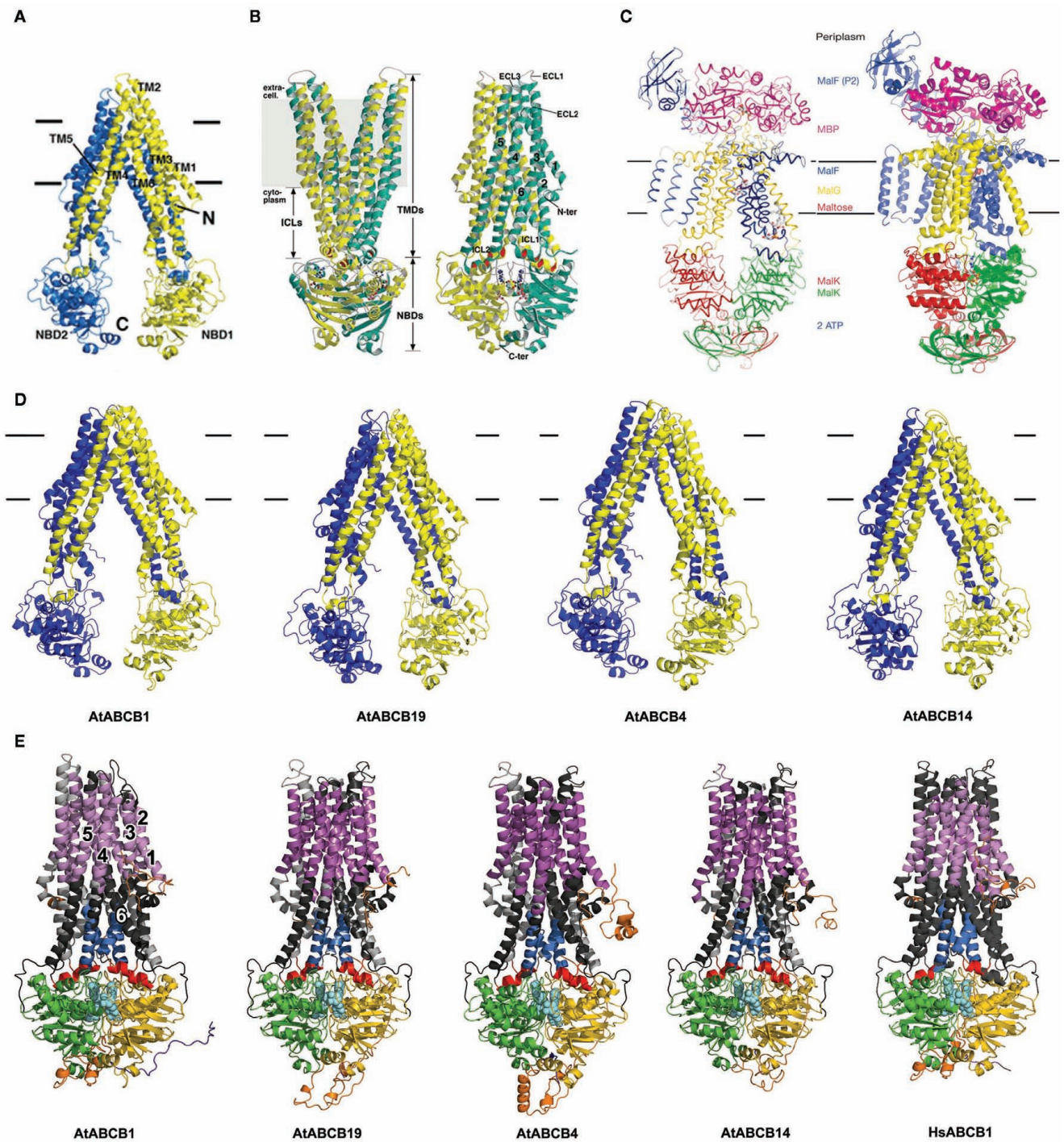


FIGURE 1 | ABC structures and *Arabidopsis* ABCB models. **(A)** MmABCB1 (Aller et al., 2009). **(B)** Sav1866 (Dawson and Locher, 2006). **(C)** Maltose transporter: left, pretranslocation intermediate state (Oldham and Chen, 2011a), right, translocation intermediate state (Oldham et al., 2007). **(D)** *Arabidopsis* ABCB structural models based on MmABCB1. **(E)** ABCB models based on Sav1866 including AtABCB19 and AtABCB4 models (Yang and Murphy, 2009).

Oldham and Chen, 2011a,b). Prokaryotic ABC importers contain a fifth component, a periplasmic or cell-surface-associated binding protein that binds specific solutes with high affinity as shown in maltose transporter in both pretranslocation intermediate state (**Figure 1C** left) and translocation intermediate state (**Figure 1C** right; Oldham et al., 2007; Oldham and Chen, 2011a,b).

Arabidopsis ABCB transporters, ABCB1, ABCB4, and ABCB19 that were independently and unambiguously demonstrated to function as auxin transporters (Geisler et al., 2005; Santelia et al., 2005; Terasaka et al., 2005; Bouchard et al., 2006; Cho et al., 2007) were chosen for our analyses. Moreover, AtABCB14 as non-auxin transporter (Lee et al., 2008) was included in this study (see Introduction for details). Models of the *Arabidopsis* unidirectional exporters ABCB1 and ABCB19 and conditional exporters ABCB4 and 14 were threaded onto theMmABCB1 and Sav1866 crystal structures (**Figures 1D,E**). AtABCB4/14 aligned better with exporters MmABCB1 and Sav1866 than ABC importers with reported crystal structures, suggesting that all *Arabidopsis* full-length ABCB transporters share ABC exporter-like architectures and are dissimilar to prokaryotic importers. Instead, import activity observed in some plant ABCB transporters appears to be determined by N-terminal linker regions and additional substrate binding or regulatory sites (Yang and Murphy, 2009; Kubes et al., 2011). The linker domains of ABCBs were constructed in outward-facing models based on structures searched from PDB database (described in Materials and Methods; **Figure 1E**), and these domains may function to stabilize the NBDs as in the maltose transporter $MalGFK_2$ (Oldham et al., 2007).

The canonical structure of ABCB-type proteins can be extended to all generated models and adopts a pseudo-symmetric arrangement of the two TMD–NBD modules (**Figures 1D,E**). This forms a large internal cavity exposed to the cytosol with separated NBDs in the nucleotide-free conformation, and an asymmetric distribution in the nucleotide-bound state that reveals the conduction chamber to the outer medium (Rosenberg et al., 2005; Dawson and Locher, 2007). In the commonly accepted ABC transport mechanism, helices from each TMD participate to the binding site(s), with a predominant role for TMH6 and TMH12 (Martin et al., 2001; Pleban et al., 2005) and form the translocation pathway for the substrate, while both NBDs transmit the necessary ATP-dependent "power stroke" to perform a complete transport cycle (Hopfner et al., 2000; Hollenstein et al., 2007; Linton and Higgins, 2007). The overall Sav1866 and MmABCB1 architectures and the subsequent structure-based allosteric mechanisms proposed in Hollenstein et al. (2007) are therefore relevant for all close ABCB-type exporters (Dawson and Locher, 2006; Hollenstein et al., 2007). As such, these computational predictions suggest that substrate specificity in plant ABCBs is determined primarily by characteristics of the TMDs. Moreover, biochemical and structural evidences designate the top of the inward-facing translocation chamber as the paradigm for the competent ligand-binding region that could provide access to either lipophilic or cytosolic substrates with minimum energies (Gutmann et al., 2010). For simplicity, further analyses concentrated on comparison of ABCB1 with other characterized plant ABCB transporters and MmABCB1 and Sav1866.

PLANT AND MAMMALIAN ABCB1 PROTEINS DISPLAY KINGDOM-SPECIFIC AND SUBSTRATE-SPECIFIC TRANSLOCATION CHAMBERS

Most of the divergences observed in ABCB substrate recognition can be attributed to residues facing the translocation chamber space. However the differences between poly- and mono-specific transporters can only be identified by analyzing the 3-D feature of the translocation chamber space. The combination of homology modeling and surface mapping of conservation scores employed in this study stands as an initial effort to pinpoint substrate specificity in large transmembrane proteins to domains composed by the dynamic arrangement of multiple transmembrane helices. Alignment and projection of plant ABCB transporter sequences onto the protein model surfaces of MmABCB1 and Sav1866 allowed us to assess the conservation of amino acid residues facing the chamber cavity. Not surprisingly, plant and animal proteins share a high degree of conservation between functional and structural domains such as nucleotide-binding folds and coupling helices, but low degree of conservation in the residues exposed to the transport cavity (**Figure 2**). Although precursors of plant and animal *ABCB* genes appear to have diverged during early plant evolution (Rea, 2007; Knoller et al., 2010; De Smet et al., 2011), comparisons of plant and animal sequences indicated considerable

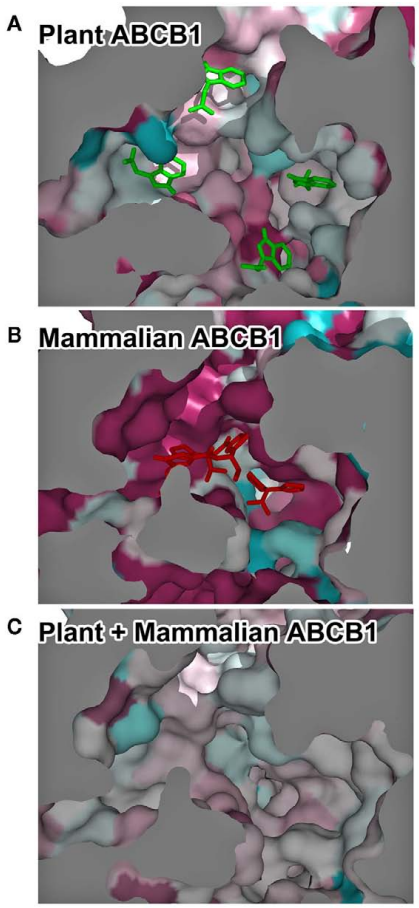


FIGURE 2 | Kingdom-specific conservation of the translocation chamber suggests early specialization of the transporters. **(A–C)** Temperature-coded residue conservation between plant ABCB1 orthologs **(A)**, mammalian ABCB1 orthologs **(B)**, and all ABCB1 sequences **(C)**. Dark-blue, lowest conservation score; dark red, highest conservation score. Bottom view of the inward-facing translocation chamber. For illustration purposes, representative IAA docking poses are displayed as green sticks and QZ59 crystallographic coordinates are displayed in red sticks.

conservation within three regions within the translocation chamber and a fourth region exhibiting conservation to a lesser extent (**Figure 2**). Comparison of mammalian ABCB1 sequences showed a strong and large spatial conservation of the chamber residues that correlates with the poly-specific drug-binding model suggested for HsABCB1 based on published drug-binding data (Aller et al., 2009; Klepsch et al., 2011). This conservation reflects the relatively recent evolution of multi-substrate specificity in mammalian ABCBs and the earlier evolution of plant ABCB auxin transporters (Knoller et al., 2010). Although only a small numbers of ABCB transporters have been characterized *in planta* (reviewed in Knoller et al., 2010; Zazimalova et al., 2010), the respective auxin transport functions of AtABCB1 and 19 have been conserved in monocots and dicots (Knoller et al., 2010), suggesting that the selective specificity of ABCB auxin transporters appeared early in plant evolution.

DOCKING OF INDOLE-3-ACETIC ACID TO AtABCB1 STRUCTURAL MODELS

The generation of 1000 IAA docking runs into the entire inward- and outward-facing structural models of AtABCB1 and AtABCB19 resulted in homologous clusters of docking poses concentrated within the translocation chambers of the transporters (**Figures 3A,B**). The poses generated for the outward-facing conformer (**Figure 3B**) targeted the intracellular loops (ICLs) region and the inner leaflet regions of the transporter, which are similar to the sites predicted in AtABCB19 (Yang and Murphy, 2009). Four binding regions were predicted in inward-facing AtABCB1 model with two regions (1 and 2) at the position of inner leaflet membrane, region 3 at the outer leaflet of membrane and region 4 out of the cell is not an entry binding sites (**Figure 3A**). The large majority of the resulting poses in outward-facing conformer was distributed at the region 3 (**Figure 3A**), which overlaps with the binding region of the cyclic peptide QZ59 in the co-crystal structures of MmABCB1 published by Aller et al. (2009). Analysis of the surface residues surrounding the region 3 showed that residues primarily located on TM helices 5, 6, 7, 11, and 12 were involved in binding (**Table 1**). The importance of a large majority of these residues was formerly assessed in different studies (Loo and Clarke, 2000, 2001; Loo et al., 2006a,b; Pajeva et al., 2009; Bessadok et al., 2011). Numerous residues involved have been previously shown to influence HsABCB1 transport activity (see **Table 1**). Interestingly, our analysis indicates that most of the identified residues listed in **Table 1** are indeed evolutionary conserved among ABCB1 orthologs taken from Table S1 in Supplementary Material (results not shown). This holds also true for Asp974, shown to be highly critical for transport activity of AtABCB1 functionally expressed in yeast (see below).

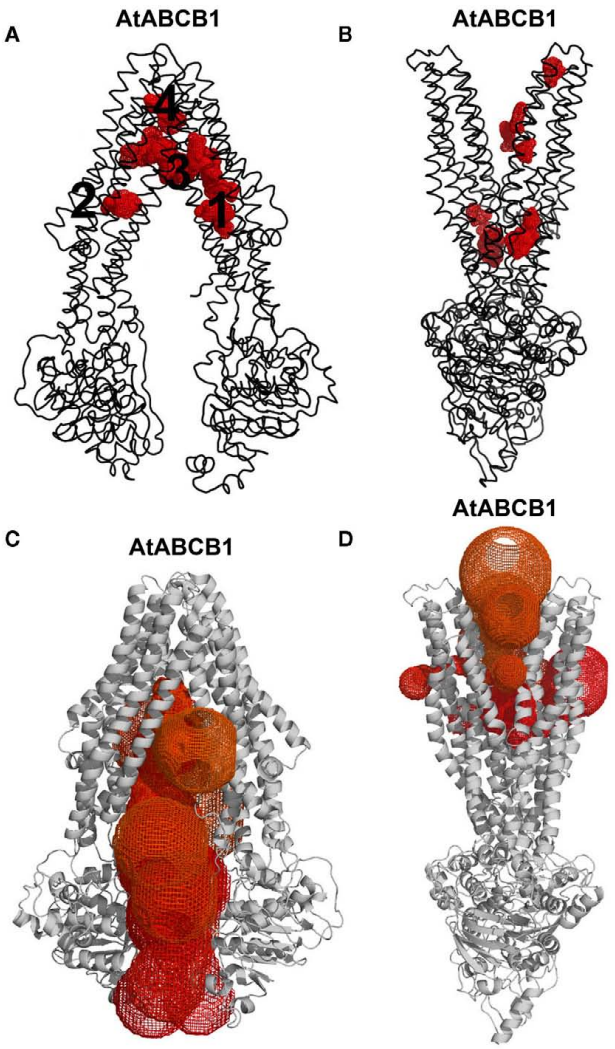


FIGURE 3 | Exhausting *in silico* docking revealed auxin-binding hotspots in AtABCB1. (A,B) Thousand IAA docking poses (red meshes) segregated in discrete regions of the transporter (black ribbon) mainly located at the bottom of the funnel-shaped translocation chambers in the inward- **(A)** and outward-facing **(B)** conformations. **(C,D)** MOLE-computed access routes (colored meshes) to the bottom of the translocation chambers in the inward- **(C)** and outward-facing **(D)** conformations of the transporter (gray cartoon).

PUTATIVE AtABCB1 AUXIN-BINDING SITES ARE CONSISTENT WITH THE EFFLUX PROCESS

The transport of the major natural auxin, IAA, across biological membranes has been extensively studied (reviewed in Petrasek and Friml, 2009). Due to the acidic pH in the apoplastic space (around pH 5.5), a part of IAA can as protonated IAAH readily diffuse over the plasma membrane. When IAA reaches the neutral cytosolic compartment, the IAA^- anion then prevails consequently trapping the hormone inside the cell (Rubery and Sheldrake, 1974; Raven, 1975). Dedicated plasma membrane transporters, such as AtABCB1/19 or members of the PIN-FORMED family (Petrasek and Friml, 2009) are then needed to control auxin efflux from the cell. This process infers that AtABCB1 recognition sites for the polar IAA^- are required to face the cytosolic space. The AtABCB1 inward-facing structure model therefore represents the auxin-binding competent state of the exporter. The two predicted binding sites associated with the inner leaflet are likely to bind IAA^- from the cytosolic side of the membrane and are predicted by docking analyses to be more specific. The third predicted region corresponding to the position of QZ59 binding sites of MmABCB1 may be the site for exclusion of IAAH in the lipid bilayer, which is

Table 1 | ABCB1–ligand interaction.

TMD	Site	Predicted residues AtABCB1	Predicted residues AtABCB19	Referenced residues HsABCB1	HsABCB1/MmABCB1 reference
1	4	L58	M55	L65	Loo and Clarke (2001), Aller et al. (2009)
	4	F65	F62	F72	Loo and Clarke (2001), Loo et al. (2003a,b), Pajeva et al. (2009)
4	3	V198	I195	S222	Loo and Clarke (2001)
5	3	Y279	Y276	F303	Klepsch et al. (2011)
	1	F280	G277	L304	Aller et al. (2009), Pajeva et al. (2009), Klepsch et al. (2011)
	3	V282	A279	I306	Loo and Clarke (2001), Pajeva et al. (2009)
	3	F283	C280	Y307	Loo and Clarke (2001), Loo et al. (2003a), Pajeva et al. (2009), Klepsch et al. (2011)
	3	Y286	W283	Y310	Pajeva et al. (2009), Klepsch et al. (2011)
	3	A287	A284	A311	n.d.
6	4	G305	G302	G329	n.d.
	3	M311	I308	F335	Loo and Clarke (1995), Aller et al. (2009)
	4	F312	F309	F336	Klepsch et al. (2011)
	3	M315	I312	L339	Loo and Clarke (2001), Aller et al. (2009)
	4	I316	V313	I340	Loo and Clarke (2001), Loo et al. (2003a), Aller et al. (2009), Klepsch et al. (2011)
	3	G318	G315	A342	Loo and Clarke (2001)
	3	L319	M316	F343	Aller et al. (2009), Klepsch et al. (2011)
	1	Q323	Q320	Q347	n.d.
7	1	C710	S697	N721	n.d.
	1	G711	G698	G722	Loo et al. (2004)
	1	S712	F699	G723	Loo et al. (2006b), Klepsch et al. (2011)
	1	L713	I700	L724	Loo et al. (2006b), Klepsch et al. (2011)
	1	S714	G701	Q725	Aller et al. (2009)
	1	F716	T703	A727	Klepsch et al. (2011)
	1	F717	F704	F728	Loo et al. (2006b), Aller et al. (2009)
	1	V720	V707	I731	Klepsch et al. (2011)
8	1	S752	L739	I764	n.d.
	1	A753	Y740	I765	Klepsch et al. (2011)
	1	V756	G743	I768	n.d.
9	1	A829	T816	A841	Loo et al. (2004)
10	2	V852	T839	I864	Loo and Clarke (2001), Aller et al. (2009)
	2	F853	F840	V865	Gruol et al. (2001)
	2	V856	L843	I868	Loo and Clarke (2001)
	2	T860	N847	G872	Loo and Clarke (2001)
11	4	Y930	F917	F942	Loo and Clarke (2001)
	2	A933	S920	T945	Loo and Clarke (2001)
	4	Q934	Q921	Q946	n.d.
	2–4	L937	L924	M949	n.d.
	4	Y938	Y925	Y950	Loo et al. (2005)
	2	S940	S927	S952	Klepsch et al. (2011)
	2–4	Y941	E928	Y953	Loo et al. (2005), Aller et al. (2009)
	4	L945	L932	F957	Loo et al. (2005)
12	1	I963	I950	L975	Loo and Clarke (2001)
	2	F966	F953	F978	Aller et al. (2009)
	2	L969	L956	V981	Klepsch et al. (2011)
	1	M970	V957	V982	Loo and Clarke (2001), Aller et al. (2009)
	1	V971	I958	F983	Dey et al. (1999)
	2	S972	T959	G984	Klepsch et al. (2011)
	2	A973	A960	A985	Loo and Clarke (2001)
	1	N974	N961	M986	Sakurai et al. (2007)

(Continued)

Table 1 | Continued

TMD	Site	Predicted residues AtABCB1	Predicted residues AtABCB19	Referenced residues HsABCB1	HsABCB1/MmABCB1 reference
	2	A976	V963	V988	n.d.
	1	A977	A964	G989	n.d.
	2	L980	V967	S992	n.d.

Amino acids involved in the binding cavity of the inward-facing conformation of AtABCB1 and AtABCB19 after 1000 IAA docking poses and their corresponding residues in HsABCB1. Site refers to the binding hotspots described in the test.

less specific as proposed in Aller et al. (2009): substrates wrapped by lipids enter the central binding sites (site 3) and are flipped out to the outer leaflet of the membrane during the change to outward-facing conformation. Another possibility is that the third region functions as an intermediate site for IAA translocation rather than as an entry site. Apolar localization of AtABCB1 and AtABCB19 suggests that, unlike PINs, AtABCB1, and AtABCB19 (Blakeslee et al., 2007; Wu et al., 2007; Mravec et al., 2008) function primarily in exclusion of IAAH from cellular membranes via site 3. However, the higher substrate specificity observed in the plant ABCB1/19 transporters compared to MmABCB1 favors interactions at the inner leaflet sites in region 1 and 2 (**Figure 3A**).

Unfortunately, due to the limitations of the docking algorithm it was technically not possible to run the simulations with IAA^- as AutoDock Vina ignores the user-supplied partial charges. On the other side, we would like to stress that at the generally assumed apoplastic pH of around 5.5, the majority of IAA ($pK_a = 4.75$) would be in its deprotonated form (roughly 83%, Zazimalova et al., 2010).

In AtABCB1, binding region 3 contains four apparent hotspots. Interestingly, top-scoring IAA docking poses in region 3 cluster in a homologous, if not identical, area in AtABCB19. Mutations in HsABCB1 residues equivalent to AtABCB1 M970 and N974 (V982 and M986, respectively) have biochemically been associated with substrate binding in the human transporter (Loo and Clarke, 2000, 2001; Loo et al., 2003a,b, 2006a,b). An Asp974 to Ala mutation in AtABCB1 functionally expressed in *Saccharomyces cerevisiae* resulted in a substantial decrease in the auxin efflux activity (Bailly and Geisler, unpublished data). The functional conservation of IAA-binding sites tends to argue for a paradigm of auxin-binding modus preserved in plants under the evolution pressure. Subsequent structure analysis using MOLE (Petrek et al., 2007) revealed that the center cavity of AtABCB1 in inward-facing conformation is only accessible from the cytosol and the inner leaflet of the lipid bilayer, thus allowing the separation between the intra- and extracellular compartments (**Figure 3C**). In the outward-facing AtABCB1 conformer, upon the large rearrangement of the transmembrane domain and especially demonstrated TMH6 movements (Loo and Clarke, 1997; Rothnie et al., 2004; Oldham et al., 2007; Storm et al., 2008; Becker et al., 2010), important residues for ligand binding to the inward-facing conformation move away from the binding cavity and face the apoplastic space, get buried in the TMD or face the lipid bilayer (**Figure 3D**). Therefore all putative binding sites are destroyed and favorable protein–ligand interactions are likely to be disrupted, thus representing the step for IAA release outside of the cell (Oldham et al., 2007; Gutmann et al., 2010). Indeed, substrate paths generated by MOLE in the outward-facing conformers exit the hotspots binding region to the extracellular space and outer plasma membrane leaflet (**Figure 3D**). Furthermore, in our simulations, IAA hardly docked in the upper parts of the transmembrane region when challenged against the outward-facing conformer (**Figure 3B**). Taken together, these results infer a conserved translocation mechanism in which IAA binds to the high-affinity internal transporter chamber from the cytosol in the closed conformation and is subsequently excluded from the low-affinity, opened binding sites to the outer space in the course of TMD reorganization (Becker et al., 2010).

AtABCB1 CAN DISTINGUISH DISSIMILAR AUXINS

The *Arabidopsis* ABCB1 and ABCB19 proteins have been recently extensively characterized and a large body of physiological, cellular, and biochemical data confirmed their function as auxin-transport components specialized in the export of the hormone across the plasma membrane (Geisler et al., 2005; Wang and Lin, 2005). However, investigators have focused their efforts on few well-studied auxins, thus leaving old questions about other auxin-related compounds unanswered. Therefore we here took the opportunity to challenge our structural models with 34 auxin-related compounds distributed within three classes categorized as Class 1 (active auxins), Class 2 (weak auxins and competitive inhibitors), and Class 3 (inactive, but structurally related compounds; **Table 2**; Ferro et al., 2006). Best docking results showed IAA poses mainly distributed in hotspots 1, 2, and 3, with a close-fitting clustering in hotspot 1 (**Figure 4A**). Remarkably, Class 1 molecules followed the same docking pattern as IAA and most of the docking poses were in the same magnitude of computed docking energy as IAA (**Table 2**) and preferentially targeted hotspots 2, 3, and 1, in this order respectively (**Figure 4B**). More interestingly, docking results from Class 2 and Class 3 showed a significantly different patterning: although these inactive or transport-inhibiting compounds share close structures with the active auxins, they exhibited binding to hotspots 2 and 3 sites, but almost excluded hotspot 1 (**Figure 4C**). This suggests a dominant role for hotspot 1 in auxin recognition and this motif may hold the ability to specifically select active molecules while the other binding regions may exhibit less selectivity. Moreover, Class 2 molecules binding energies scored generally lower and Class 3 higher than IAA, thus suggesting that the behavior of those chemicals is linked to the affinity to their binding site(s) (**Table 2**). Remarkable examples are illustrated by 1-naphtalene acetic acid (1-NAA), 2-napthalene acetic

Table 2 | Docking scores of auxin-related compounds used in this study.

Compound name	Class no	Top pose score	Mean of 5*
1-Naphthalene acetic acid	Class 1	-6.9	-6.72
2-Me, 4-Cl-phenoxy acetic acid	Class 1	-5.9	-5.62
2,5-Cl_2-phenoxy acetic acid	Class 1	-5.9	-5.60
2,4-Cl_2-phenoxy acetic acid	Class 1	-5.8	-5.58
Indole-3-acetic acid	Class 1	-6.4	-6.14
4-Cl-indole-3-acetic acid	Class 1	-6.5	-6.44
5-Cl-indole-3-acetic acid	Class 1	-6.6	-6.40
6-Cl-indole-3-acetic acid	Class 1	-6.8	-6.44
5,6-Cl_2-indole-3-acetic acid	Class 1	-6.9	-6.58
6-F-indole-3-acetic acid	Class 1	-6.8	-6.46
4-F-indole-3-acetic acid	Class 1	-6.7	-6.42
4-Me-indole-3-acetic acid	Class 1	-6.6	-6.46
2,3,4-Cl_3-phenoxy acetic acid	Class 1	-6	-5.80
5-F-indole-3-acetic acid	Class 1	-6.8	-6.42
4-Et-indole-3-acetic acid	Class 1	-6.8	-6.54
5-Me-indole-3-acetic acid	Class 1	-6.8	-6.42
7-F-indole-3-acetic acid	Class 1	-6.7	-6.44
4,6-Cl_2-indole-3-acetic acid	Class 1	-6.3	-6.26
2,4,5-Cl_3-phenoxy acetic acid	Class 1	-5.9	-5.68
6,7-Cl_2-indole-3-acetic acid	Class 1	-7.1	-6.38
4-Indole-3-butyric acid	Class 1	-6.6	-6.38
R-2,5-Me2-phenoxy-2-propionic acid	Class 1	-6.4	-6.22
7-Cl-indole-3-acetic acid	Class 2	-6.7	-6.42
4,7-Cl_2-indole-3-acetic acid	Class 2	-7.1	-6.56
3-Indole-3-propionic acid	Class 2	-6.4	-6.34
5,7-Cl_2-indole-3-acetic acid	Class 2	-6.8	-6.40
2-Naphthalene acetic acid	Class 2	-7.2	-6.86
2-Cl-Benzoic acid	Class 3	-5.3	-5.24
2-F-Benzoic acid	Class 3	-5.6	-5.54
2-I-Benzoic acid	Class 3	-5.3	-5.16
2,6-Cl_2-benzoic acid	Class 3	-5.5	-5.48
Benzoic acid	Class 3	-5.5	-5.38
4-Cl-benzoic acid	Class 3	-5.9	-5.60
Beta-naphthoic acid	Class 3	-6.8	-6.70
2-Me-benzoic acid	Class 3	-6.1	-5.68
2,3,5-I_3-benzoic acid	Class 3	-5	-4.28
1-*N*-naphthylphthalamic acid	–	-8.5	-7.90

*Color scaled binding affinities obtained in AutoDock Vina simulations expressed in kcal/mol. *Mean of the five best docking clusters.*

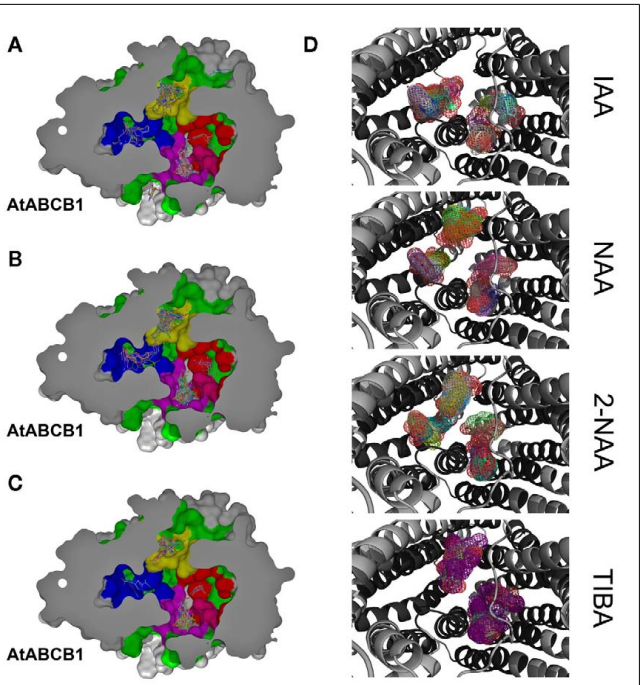


FIGURE 4 | Docking of auxinic compounds to AtABCB1 suggests functional substrate selectivity. (A–C) Docking poses of IAA **(A)**, active **(B)**, and inactive **(C)** auxinic compounds in the inward-facing cavity. IAA-binding hotspots are colored on the protein surface as 1, blue; 2, magenta; 3, yellow; 4, red. **(D)** Close-up views of docking poses (shown as colored meshes) of commonly used auxin-related compounds. AtABCB1 is shown as a gray cartoon and predicted transmembrane residues are colored in black.

acid (2-NAA), and 2,3,5-triiodobenzoic acid (TIBA; **Figure 4D**; **Figure A2** in Appendix).

Synthetic 1-NAA acts as an active lipophilic auxin and mimics IAA docking pattern by occupying the four hotspots (**Figure 4D**) with lower binding energies than IAA, probably due to the stronger hydrophobicity generated by the naphthalene ring compared to the indole ring in absence of an obvious hydrogen-binding partner for the indole amine (**Table 2**). Surprisingly, its close enantiomer, the anti-auxin, 2-NAA, an inactive auxin analog, docks to the same regions in AtABCB1 and AtABCB19 but with higher affinities (**Table 2**). This is further supported by analogous behaviors from auxinic competitors in early radiolabeled auxin-binding protein studies (Lobler and Klambt, 1985). Expression of AtABCB1 in yeast and human heterologous systems lead to enhanced export capacities of IAA and NAA but not 2-NAA (Geisler et al., 2005), inferring that the inefficient transport might lie in its high affinity to the binding pocket. Class 2 compounds displayed the highest binding scoring consistent with the role of transport inhibitors by competing with active auxins for substrate-binding sites (**Table 2**).

Despite its clearly different ring structure, the weak auxin, TIBA, is thought to act as auxin-transport inhibitor by mimicry of auxin *in planta* (Keitt and Baker, 1966). However, in computed dockings against AtABCB1 and AtABCB19, TIBA docks to hotspots 2, 3, and 4 with very low scores in the same range as benzoic acid and its derivatives (**Figure 4D**; **Table 2**). Moreover, TIBA is excluded from the first IAA-binding hotspot, suggesting that the transporters probably do not recognize the molecule as a potent substrate (**Figure 4D**). This suggests that the inhibitory effect of TIBA on AtABCB1 activity might be conferred not by direct interference with the substrate-binding site of ABCB1 as found for allosteric inhibitors or by indirect events.

Following the idea that ABCB-mediated polar auxin-transport inhibitors may directly interact with the auxin-binding sites, we included NPA (1-*N*-naphthylphthalamic acid) in our analysis. Interestingly, in addition to the docking poses in the TMD–TMD interface previously reported (Kim et al., 2010), NPA docked into

hotspot 1 and 3 with the highest binding score observed in this study (**Table 2**), which is in line with its reported striking inhibitory effect on ABCB1-mediated auxin transport (Wang and Lin, 2005; Bailly et al., 2008; Nagashima et al., 2008) and NPA binding studies (Sussman and Gardner, 1980; Muday et al., 1993). These data however also support the concept that ABCBs are primary NPA targets (Rojas-Pierce et al., 2007; Kim et al., 2010).

SURFACE ELECTROSTATICS OF TRANSLOCATION CHAMBERS REFLECT SUBSTRATE SPECIFICITY

In order to understand the functional influence of evolution on the translocation pathways of ABCB proteins and its consequence on substrate specificity, electrostatic potentials were computed for a choice of related plant and mammalian sequences from different species and were mapped to each corresponding chamber surface. Further, the overall distribution of apolar residues and polar residues on both the *Arabidopsis* and murine ABCB1 binding chambers (**Figure 5**) is not sufficient to clarify the observed discrepancies in substrate specificity. Elements of answer may come from the analysis of the electrostatic potentials computed for both translocation chambers.

In its inward-facing modeling, AtABCB1 displayed predominantly lateral negative potentials near the entrance of the cavity that weakens to a more neutral environment in the region where the binding hotspots were described (**Figure 4A**). In a greater magnitude, AtABCB19 retained the same electrostatic distribution with a solvent accessible surface that remained principally negative, with values $>10\ kT$ near the putative binding pocket (**Figure 6A**). This conserved pattern seems to be functionally important as other models of plant ABCB1 orthologs, despite their divergence in residues aligned in the translocation chamber, kept the same negative to neutral distribution of the surface electrostatic potential (data not shown). Strikingly, AtABCB4 and AtABCB14, described to function as auxin and malate importers (Terasaka et al., 2005; Oldham and Chen, 2011b), respectively, displayed a common neutral to positive electrostatic surface throughout the whole chamber (**Figure 6A**), arguing for a functional evolution of the plant ABCB subfamily toward a specialization of their transport activities.

Human and murine B1 showed a less homogenous electrostatic potential ranging from strong negative and positive stretches toward the substrate entrance of the translocation pathway to a more neutral electrostatic character with small patches of strong charges closer to the substrate-binding site (**Figure 6A**). Remarkably, the electrostatic surface for the mammalian ABCB1 seems to mirror the various physico-chemical properties of the lipophilic substrates described for this transporter (**Figure A2** in Appendix, see Sharom, 2008) in comparison to the hydrophilic IAA^-.

It is worth mentioning that the transmembrane region of the TMDs concentrates most of the differences observed between the different categories of transporters included in this study. Furthermore, the observed neutral to charged environments described for the ligand-accepting conformers are tremendously altered with the TMDs conformational rearrangements associated with ATP binding and the exposition of the cavity to the extracellular space. Indeed, all molecules tested here switched to fully charged surfaces in their outward-facing conformers (**Figure 6A**). These changes of electrostatic properties may represent an important component of the ligand release step by presenting charged and polar residues to the cavity that were previously buried in the protein and thus creating an incompatible drug-binding situation.

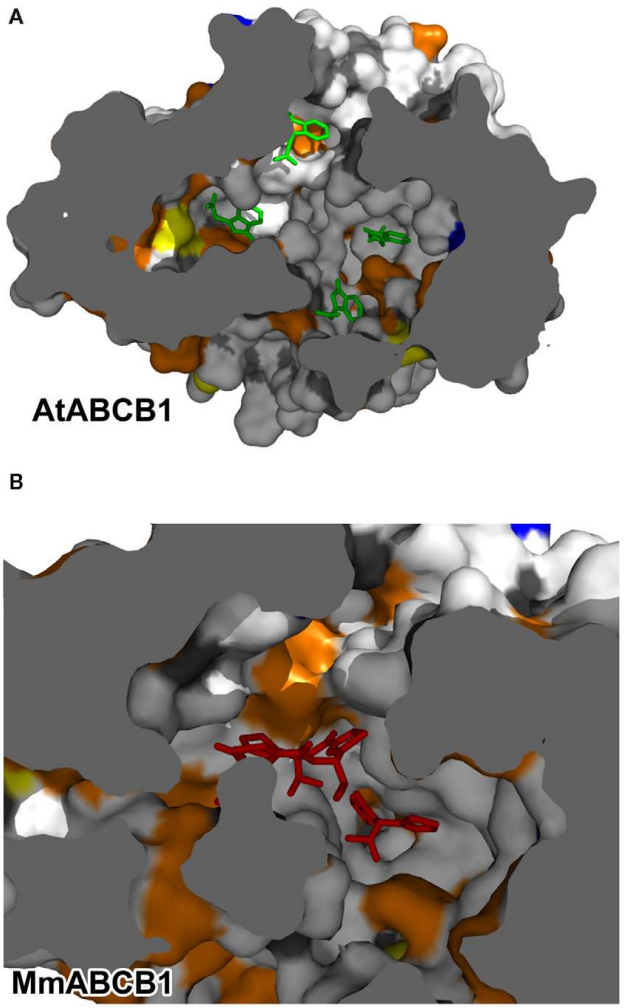


FIGURE 5 | Surface distribution of residue properties in the ABCB translocation chamber. (A) AtABCB1; representative IAA docking poses are depicted as green sticks. **(B)** MmABCB1; Co-crystallized compounds coordinates are shown as red sticks. Residue property color code is as follow: white, apolar; orange, polar; red, acidic; blue, alkaline; yellow, cysteine.

To summarize, computation of surface electrostatics revealed a great level of divergence in electrostatic properties that correlated with the substrate specificity of the transporters. While the mammalian B1 substrate promiscuity is reflected by a broad range distribution of charges, narrow-specific plant transporters show a high level of functional conservation in surface electrostatic potentials.

The network of electrostatic interactions between the surface-facing residues is expected to provide a basis for the distinction between discrete sites or common functional regions. Plant ABCB models clearly show a neutral setting with weak patches of charges in the putative auxin-binding region (**Figure 6B**). The weak static

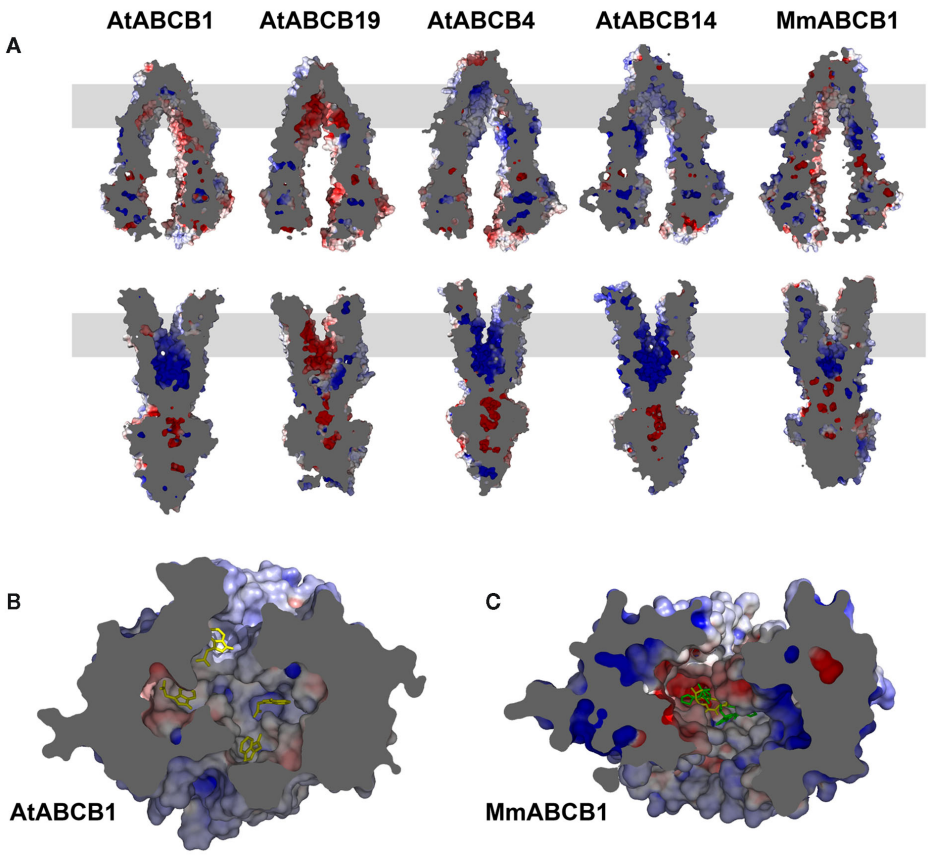


FIGURE 6 | Surface electrostatic potential of the ABCB translocation chamber reflects its substrate specificity. (A) Cut views of *Arabidopsis* (At) and mouse (Mm) ABCB proteins' surface electrostatics in both inward- (up) and outward-facing conformers (down). The gray box represents the estimated position of the lipid bilayer. **(B)** Bottom view of the AtABCB1 inward-facing translocation chamber. Representative IAA docking poses are shown as yellow sticks. **(C)** Bottom view of the MmABCB1 inward-facing translocation chamber. Co-crystallized compound coordinates are shown as yellow and green sticks. Scale is red −10 *kT*, blue +10 *kT*.

electric fields thus represent the ideal environment for the transient interaction of such small molecules. Surface electrostatic potentials in the crystal structure of MmABCB1 correlate with the binding of the cyclic peptides QZ59-RRR and QZ59-SSS where patches of negative, neutral, and positive fields provide accommodation to these complex molecules that build hydrophobic and aromatic interaction within the chamber (**Figure 6C**; Aller et al., 2009). It is worth mentioning that the increased negative potential in AtABCB19 is correlating with increased IAA transport rates reflected finally by the more severe null mutant phenotype (Geisler et al., 2005; Wang and Lin, 2005; Geisler and Murphy, 2006). On the contrary, the strong negative potential concentrated at the top of the MmABCB1 translocation chamber is surrounded by a more neutral environment that may accommodate its large and lipophilic substrates. Taken with the dominantly apolar residues lying in its cavity (**Figure 4**), the electrostatic properties of the mammalian transporter may account for the diversity in ligands reported so far. The size of the binding region is large enough to receive bulky or long-chained compounds and the flexibility of the amino acid side chains facing the cavity is confidently sufficient to build the binding plasticity or adaptability that forms the canonical multisite previously hypothesized (Corpet, 1988; Garrigos et al., 1997; Knoller et al., 2010). Though Bachas et al. (2011) have proposed a model for multidrug recognition based on a single rigid binding pocket in the *Bacillus subtilis* BmrR, such a scenario seems unlikely to be true in mammalian ABCB molecules where the different structural and biochemical evidences point to a larger platform for recognition of dissimilar molecules (Lee et al., 2008; Newberry et al., 2008).

Although it represents a loose cavity when compared to enzymatic or receptor's binding pockets, from all 4 putative binding hotspots proposed in this work, only hotspot 1 forms a close-fitting site for IAA recognition (**Figure 4A**). It is therefore difficult to sustain that information obtained *in silico* may stay valid when challenged by *in vivo* experimentation. However, our first attempts to use the knowledge gathered via this analysis has been the successful site-directed mutagenesis of site 1 by exchanging I963 and N974 to alanines that led to drastic decreases in AtABCB1-mediated IAA efflux from *S. cerevisiae* cells (not shown). Those regions may therefore form *bona fide* binding sites congruent with the transport cycle (Oldham et al., 2007) as each pocket described in the inward-facing conformer vanishes in the outward-facing

conformer along with the large transmembrane helices movements. However, the molecular dynamics of the ABC-type cavity and its remodeling during the catalytic cycle remain difficult to predict and, although simulations have been recently performed (Oldham et al., 2007; Becker et al., 2010).

Obviously, the long evolutionary period separating plant and mammalian ABCB1 transporters has resulted in different electrostatic properties, which are likely to reflect differences in the recognition and transport of their respective substrate(s).

DISCUSSION

Despite diverse substrate specialization, bacterial, mammalian, and plant ABCB transporters share a close common architecture (Locher, 2004; Hollenstein et al., 2007), underlining the essential function of pumping specific drugs in or out cellular compartments. Our analysis identified candidate substrate regions in the transporter's transmembrane domains that appear to describe a paradigm for substrate recognition and subsequent translocation. Comparisons of plant and animal ABCB central binding sites suggest that the multi-substrate specificity associated with this region is not due to a randomized distribution of amino acid side chains and is functionally conserved. However, the putative IAA-binding sites and translocation surfaces of the plant ABCBs examined herein are relatively conserved, although they lack a high degree of sequence identity (**Figure 4**). This suggests that evolutionary pressure may have favored a plastic, multi-site substrate-binding environment over a rigid substrate-binding pocket as found for the auxin receptor (Tan et al., 2007).

The kingdom-specific differences presented here raise the question of how the large ABCB binding pocket can selectively recognize and transport molecules as large as hydrophobic HsABCB1 substrates and as small as auxins? Indeed, IAA binding may not generate sufficient energy to cause the expected "induced-fit" sterical shift by itself. Kimura et al. (2007) proposed a model for human ABC transporters in which sterols or phospholipids would be needed to fill the remaining space and play a role in small substrate recognition. This "cholesterol fill-in" concept is an attractive alternative to a water-filled cavity and is not incompatible with auxin transport but lacks biochemical evidence for plant ABCBs. In support of this concept, activity and membrane stability of *Arabidopsis* ABCB19 have been shown to be positively modulated by structural sterols (Rojas-Pierce et al., 2007; Titapiwatanakun and Murphy, 2009).

Despite the large space of the translocation chamber, most of the auxinic compounds tested in this analysis share the same binding sites (**Figure 4**) with different binding energies (**Table 2**). Most of these compounds display a common structural basis, comprising the ring system and the carboxylic function, but differ in their substituents, inferring that the protein would recognize rather a family of ligands than a definite molecule. The docking scores obtained here are weak and within a close range, probably because of the small size and chemical similarity of auxins, but all three compound classes showed distinct behaviors: Class 1 active molecules displayed midrange scores, Class 2 anti-auxins seemed to have the highest binding affinities and Class 3 inactive auxins had the lowest scores (**Table 2**). Those properties could partially explain the physiological effects of auxin-like molecules since active compounds bind preferentially to site 1 while inactive and anti-auxins do not (**Figure 4**). One might speculate that hotspot 1 may represent an active and selective site for active auxins where inhibitors with high affinity could block the translocation mechanism. Interestingly the efficient polar auxin-transport inhibitor, NPA, targeted hotspot 1 and 3 in our docking simulations, with the highest scores obtained in this study (**Table 2**). The modeling presented here suggests that the ABCB1 chamber could accommodate multiple auxin molecules at the same time. As functional cooperation between multiple sites with different affinities has been observed in HsABCB1 (Shapiro and Ling, 1997), this prediction is a priority target for experimental testing.

CONCLUSION AND FURTHER DIRECTION

The solution of the Sav1886, MsbA, and MmABCB1 crystal structures and an avalanche of sequence information derived from genome sequencing projects have provided the basis for sophisticated structural models of plant ABCB transporters that can be rigorously tested. As plant genomes contain an expanded family of ABC-transporter genes (Theodoulou, 2000; Sanchez-Fernandez et al., 2001a,b; Martinoia et al., 2002; Jasinski et al., 2003), multiple narrow-specificity transporters in these sessile organisms reflects an apparent need (Geisler and Murphy, 2006). Structurally and functionally, ABCB substrate specialization seems to be recent when compared to the ancient origin of ABC proteins. Some genetic studies from water- to land-plants indicate that the complexity of the ABC-transporter family arose with the emergence of auxin-controlled development (Kelch et al., 2004), although this expansion requires further validation as more plant genomes are sequenced. The task at hand is the experimental testing of the models shown here and the identification of the cofactors required for efficient plant ABCB transport.

ACKNOWLEDGMENTS

We would like thank Jue Chen and her lab members for technical advices, fruitful discussion, and comments on the manuscript. This work was supported by grants from the Division of Energy Biosciences, US Department of Energy (DE-FG02-06ER15804 to Angus S. Murphy and Haibing Yang), from the Forschungskredit of the University of Zurich (to Aurélien Bailly), and from the Swiss National Funds (to Markus Geisler).

SUPPLEMENTARY MATERIAL

The Supplementary Material (Tables and zipped folder of homology models in PDB format) for this article can be found online at http://www.frontiersin.org/plant_traffic_and_transport/10.3389/fpls.2011.00108/abstract

Table S1 | ABCB1/PGP1 orthologs used in this study. All protein sequences used in the present work have been obtained from BLAST analyzes (blastp, psi-blast, phi-blast, and tblastn) using HsABCB1 and AtABCB1 translated cDNAs and a scoring threshold of 1000 points, 90% coverage, and *e*-value <0.01.

Table S2 | Alignments of AtABCBs and MmABCB1 sequences that were used to build homology models for this study. For details, see Section "Material and Methods."

Table S3 | Complete template alignments for HsABCB1 are shown in two runs of HsABCB1 modeling based on Sav1866. For details, see Section "Material and Methods."

REFERENCES

Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P. M., Trinh, Y. T., Zhang, Q., Urbatsch, I. L., and Chang, G. (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323, 1718–1722.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* 38, W529–W533.

Bachas, S., Eginton, C., Gunio, D., and Wade, H. (2011). Structural contributions to multidrug recognition in the multidrug resistance (MDR) gene regulator, BmrR. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11046–11051.

Bailly, A., Sovero, V., Vincenzetti, V., Santelia, D., Bartnik, D., Koenig, B. W., Mancuso, S., Martinoia, E., and Geisler, M. (2008). Modulation of P-glycoproteins by auxin transport inhibitors is mediated by interaction with immunophilins. *J. Biol. Chem.* 283, 21817–21826.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and Mccammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10037–10041.

Becker, J. P., Van Bambeke, F., Tulkens, P. M., and Prevost, M. (2010). Dynamics and structural changes induced by ATP binding in SAV1866, a bacterial ABC exporter. *J. Phys. Chem. B* 114, 15948–15957.

Benkert, P., Tosatto, S. C., and Schomburg, D. (2008). QMEAN: a comprehensive scoring function for model quality assessment. *Proteins* 71, 261–277.

Bessadok, A., Garcia, E., Jacquet, H., Martin, S., Garrigues, A., Loiseau, N., Andre, F., Orlowski, S., and Vivaudou, M. (2011). Recognition of sulfonylurea receptor (ABCC8/9) ligands by the multidrug resistance transporter P-glycoprotein (ABCB1): functional similarities based on common structural features between two multispecific ABC proteins. *J. Biol. Chem.* 286, 3552–3569.

Blakeslee, J. J., Bandyopadhyay, A., Lee, O. R., Mravec, J., Titapiwatanakun, B., Sauer, M., Makam, S. N., Cheng, Y., Bouchard, R., Adamec, J., Geisler, M., Nagashima, A., Sakai, T., Martinoia, E., Friml, J., Peer, W. A., and Murphy, A. S. (2007). Interactions among PIN-FORMED and P-glycoprotein auxin transporters in *Arabidopsis*. *Plant Cell* 19, 131–147.

Borowski, E., Bontemps-Gracz, M. M., and Piwkowska, A. (2005). Strategies for overcoming ABC-transporters-mediated multidrug resistance (MDR) of tumor cells. *Acta Biochim. Pol.* 52, 609–627.

Bouchard, R., Bailly, A., Blakeslee, J. J., Oehring, S. C., Vincenzetti, V., Lee, O. R., Paponov, I., Palme, K., Mancuso, S., Murphy, A. S., Schulz, B., and Geisler, M. (2006). Immunophilin-like TWISTED DWARF1 modulates auxin efflux activities of *Arabidopsis* P-glycoproteins. *J. Biol. Chem.* 281, 30603–30612.

Cho, M., Lee, S. H., and Cho, H. T. (2007). P-glycoprotein4 displays auxin efflux transporter-like action in *Arabidopsis* root hair cells and tobacco cells. *Plant Cell* 19, 3930–3943.

Colovos, C., and Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2, 1511–1519.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890.

Dawson, R. J., Hollenstein, K., and Locher, K. P. (2007). Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Mol. Microbiol.* 65, 250–257.

Dawson, R. J., and Locher, K. P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185.

Dawson, R. J., and Locher, K. P. (2007). Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett.* 581, 935–938.

De Smet, I., Voss, U., Lau, S., Wilson, M., Shao, N., Timme, R. E., Swarup, R., Kerr, I., Hodgman, C., Bock, R., Bennett, M., Jurgens, G., and Beeckman, T. (2011). Unraveling the evolution of auxin signaling. *Plant Physiol.* 155, 209–221.

Dey, S., Hafkemeyer, P., Pastan, I., and Gottesman, M. M. (1999). A single amino acid residue contributes to distinct mechanisms of inhibition of the human multidrug transporter by stereoisomers of the dopamine receptor antagonist flupentixol. *Biochemistry* 38, 6630–6639.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* Chap. 5, Unit 5.6.1–5.6.30.

Ferro, N., Bredow, T., Jacobsen, H. J., and Reinard, T. (2010). Route to novel auxin: auxin chemical space toward biological correlation carriers. *Chem. Rev.* 110, 4690–4708.

Ferro, N., Gallegos, A., Bultinck, P., Jacobsen, H. J., Carbo-Dorca, R., and Reinard, T. (2006). Coulomb and overlap self-similarities: a comparative selectivity analysis of structure-function relationships for auxin-like molecules. *J. Chem. Inf. Model.* 46, 1751–1762.

Garrigos, M., Mir, L. M., and Orlowski, S. (1997). Competitive and non-competitive inhibition of the multidrug-resistance-associated P-glycoprotein ATPase – further experimental evidence for a multisite model. *Eur. J. Biochem.* 244, 664–673.

Geisler, M., Blakeslee, J. J., Bouchard, R., Lee, O. R., Vincenzetti, V., Bandyopadhyay, A., Titapiwatanakun, B., Peer, W. A., Bailly, A., Richards, E. L., Ejendal, K. F., Smith, A. P., Baroux, C., Grossniklaus, U., Muller, A., Hrycyna, C. A., Dudler, R., Murphy, A. S., and Martinoia, E. (2005). Cellular efflux of auxin catalyzed by the *Arabidopsis* MDR/PGP transporter AtPGP1. *Plant J.* 44, 179–194.

Geisler, M., Kolukisaoglu, H. U., Bouchard, R., Billion, K., Berger, J., Saal, B., Frangne, N., Koncz-Kalman, Z., Koncz, C., Dudler, R., Blakeslee, J. J., Murphy, A. S., Martinoia, E., and Schulz, B. (2003). TWISTED DWARF1, a unique plasma membrane-anchored immunophilin-like protein, interacts with *Arabidopsis* multidrug resistance-like transporters AtPGP1 and AtPGP19. *Mol. Biol. Cell* 14, 4238–4249.

Geisler, M., and Murphy, A. S. (2006). The ABC of auxin transport: the role of p-glycoproteins in plant development. *FEBS Lett.* 580, 1094–1102.

Gruol, D. J., Bernd, J., Phippard, A. E., Ojima, I., and Bernacki, R. J. (2001). The use of a novel taxane-based P-glycoprotein inhibitor to identify mutations that alter the interaction of the protein with paclitaxel. *Mol. Pharmacol.* 60, 104–113.

Gutmann, D. A., Ward, A., Urbatsch, I. L., Chang, G., and Van Veen, H. W. (2010). Understanding polyspecificity of multidrug ABC transporters: closing in on the gaps in ABCB1. *Trends Biochem. Sci.* 35, 36–42.

Higgins, C. F. (2001). ABC transporters: physiology, structure and mechanism – an overview. *Res. Microbiol.* 152, 205–210.

Hollenstein, K., Dawson, R. J., and Locher, K. P. (2007). Structure and mechanism of ABC transporter proteins. *Curr. Opin. Struct. Biol.* 17, 412–418.

Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996). Errors in protein structures. *Nature* 381, 272.

Hopfner, K. P., Karcher, A., Shin, D. S., Craig, L., Arthur, L. M., Carney, J. P., and Tainer, J. A. (2000). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 101, 789–800.

Hrycyna, C. A., and Gottesman, M. M. (1998). Multidrug ABC transporters from bacteria to man: an emerging hypothesis for the universality of molecular mechanism and function. *Drug Resist. Updat.* 1, 81–83.

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

Keitt, G. W., and Baker, R. A. (1966). Auxin activity of substituted benzoic acids and their effect on polar auxin transport. *Plant Physiol.* 41, 1561.

Kelch, D. G., Driskell, A., and Mishler, B. (2004). "Inferring phylogeny using genomic characters: a case study using land plant plastomes," *Molecular Systematics of Bryophytes*, eds V. H. B. Goffinet and R. Magill (St. Louis, MO: Missouri Botanical Garden Press), 3–12.

Kim, J. Y., Henrichs, S., Bailly, A., Vincenzetti, V., Sovero, V., Mancuso, S., Pollmann, S., Kim, D., Geisler, M., and Nam, H. G. (2010). Identification of an ABCB/P-glycoprotein-specific inhibitor of auxin transport by chemical genomics. *J. Biol. Chem.* 285, 23309–23317.

Kimura, Y., Kodan, A., Matsuo, M., and Ueda, K. (2007). Cholesterol fill-in model: mechanism for substrate recognition by ABC proteins. *J. Bioenerg. Biomembr.* 39, 447–452.

Klepsch, F., Chiba, P., and Ecker, G. F. (2011). Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Comput. Biol.* 7, e1002036. doi:10.1371/journal.pcbi.1002036

Knoller, A. S., Blakeslee, J. J., Richards, E. L., Peer, W. A., and Murphy, A. S. (2010). Brachytic2/ZmABCB1 functions in IAA export from intercalary meristems. *J. Exp. Bot.* 61, 3689–3696.

Kubes, M., Yang, H., Richter, G. L., Cheng, Y., Mlodzinska, E., Wang, X.,

Blakeslee, J., Carraro, N., Petrasek, J., Zazimalova, E., Hoyerova, K., Peer, W. A., and Murphy, A. S. (2011). The *Arabidopsis* concentration-dependent influx/efflux transporter ABCB4 regulates cellular auxin levels in the root epidermis. *Plant J.* doi: 10.1111/j.1365-313X.2011.04818.x

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., Mcwilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.

Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993). Procheck – a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.

Lee, M., Choi, Y., Burla, B., Kim, Y. Y., Jeon, B., Maeshima, M., Yoo, J. Y., Martinoia, E., and Lee, Y. (2008). The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to CO2. *Nat. Cell Biol.* 10, 1217–1223.

Li, Y., Yuan, H., Yang, K., Xu, W., Tang, W., and Li, X. (2010). The structure and functions of P-glycoprotein. *Curr. Med. Chem.* 17, 786–800.

Linton, K. J., and Higgins, C. F. (2007). Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch.* 453, 555–567.

Lobler, M., and Klambt, D. (1985). Auxin-binding protein from coleoptile membranes of corn (*Zea mays* L).10.1. Purification by immunological methods and characterization. *J. Biol. Chem.* 260, 9848–9853.

Locher, K. P. (2004). Structure and mechanism of ABC transporters. *Curr. Opin. Struct. Biol.* 14, 426–431.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2003a). Drug binding in human P-glycoprotein causes conformational changes in both nucleotide-binding domains. *J. Biol. Chem.* 278, 1575–1578.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2003b). Substrate-induced conformational changes in the transmembrane segments of human P-glycoprotein. Direct evidence for the substrate-induced fit mechanism for drug binding. *J. Biol. Chem.* 278, 13603–13606.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2004). Processing mutations located throughout the human multidrug resistance P-glycoprotein disrupt interactions between the nucleotide binding domains. *J. Biol. Chem.* 279, 38395–38401.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2005). ATP hydrolysis promotes interactions between the extracellular ends of transmembrane segments 1 and 11 of human multidrug resistance P-glycoprotein. *Biochemistry* 44, 10250–10258.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2006a). Transmembrane segment 1 of human P-glycoprotein contributes to the drug-binding pocket. *Biochem. J.* 396, 537–545.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. (2006b). Transmembrane segment 7 of human P-glycoprotein forms part of the drug-binding pocket. *Biochem. J.* 399, 351–359.

Loo, T. W., and Clarke, D. M. (1995). Rapid purification of human P-glycoprotein mutants expressed transiently in HEK 293 cells by nickel-chelate chromatography and characterization of their drug-stimulated ATPase activities. *J. Biol. Chem.* 270, 21449–21452.

Loo, T. W., and Clarke, D. M. (1997). Drug-stimulated ATPase activity of human P-glycoprotein requires movement between transmembrane segments 6 and 12. *J. Biol. Chem.* 272, 20986–20989.

Loo, T. W., and Clarke, D. M. (2000). Identification of residues within the drug-binding domain of the human multidrug resistance P-glycoprotein by cysteine-scanning mutagenesis and reaction with dibromobimane. *J. Biol. Chem.* 275, 39272–39278.

Loo, T. W., and Clarke, D. M. (2001). Defining the drug-binding site in the human multidrug resistance P-glycoprotein using a methanethiosulfonate analog of verapamil, MTS-verapamil. *J. Biol. Chem.* 276, 14972–14979.

Martin, C., Higgins, C. F., and Callaghan, R. (2001). The vinblastine binding site adopts high- and low-affinity conformations during a transport cycle of P-glycoprotein. *Biochemistry* 40, 15733–15742.

Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kolukisaoglu, U., Muller-Rober, B., and Schulz, B. (2002). Multifunctionality of plant ABC transporters – more than just detoxifiers. *Planta* 214, 345–355.

Mravec, J., Kubes, M., Bielach, A., Gaykova, V., Petrasek, J., Skupa, P., Chand, S., Benkova, E., Zazimalova, E., and Friml, J. (2008). Interaction of PIN and PGP transport mechanisms in auxin distribution-dependent development. *Development* 135, 3345–3354.

Muday, G. K., Brunn, S. A., Haworth, P., and Subramanian, M. (1993). Evidence for a single naphthylphthalamic acid binding site on the zucchini plasma membrane. *Plant Physiol.* 103, 449–456.

Multani, D. S., Briggs, S. P., Chamberlin, M. A., Blakeslee, J. J., Murphy, A. S., and Johal, G. S. (2003). Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science* 302, 81–84.

Nagashima, A., Uehara, Y., and Sakai, T. (2008). The ABC subfamily B auxin transporter AtABCB19 is involved in the inhibitory effects of N-1-naphthyphthalamic acid on the phototropic and gravitropic responses of *Arabidopsis* hypocotyls. *Plant Cell Physiol.* 49, 1250–1255.

Newberry, K. J., Huffman, J. L., Miller, M. C., Vazquez-Laslop, N., Neyfakh, A. A., and Brennan, R. G. (2008). Structures of BmrR-drug complexes reveal a rigid multidrug binding pocket and transcription activation through tyrosine expulsion. *J. Biol. Chem.* 283, 26795–26804.

Noh, B., Murphy, A. S., and Spalding, E. P. (2001). Multidrug resistance-like genes of *Arabidopsis* required for auxin transport and auxin-mediated development. *Plant Cell* 13, 2441–2454.

O'Connor, R., Clynes, M., Dowling, P., O'donovan, N., and O'driscoll, L. (2007). Drug resistance in cancer – searching for mechanisms, markers and therapeutic agents. *Expert Opin. Drug Metab. Toxicol.* 3, 805–817.

Oldham, M. L., and Chen, J. (2011a). Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* 332, 1202–1205.

Oldham, M. L., and Chen, J. (2011b). Snapshots of the maltose transporter during ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15152–15156.

Oldham, M. L., Khare, D., Quiocho, F. A., Davidson, A. L., and Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450, 515–521.

Pajeva, I. K., Globisch, C., and Wiese, M. (2009). Comparison of the inward- and outward-open homology models and ligand binding of human P-glycoprotein. *FEBS J.* 276, 7016–7026.

Petrasek, J., and Friml, J. (2009). Auxin transport routes in plant development. *Development* 136, 2675–2688.

Petrek, M., Kosinova, P., Koca, J., and Otyepka, M. (2007). MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure* 15, 1357–1363.

Pleban, K., Kopp, S., Csaszar, E., Peer, M., Hrebicek, T., Rizzi, A., Ecker, G. F., and Chiba, P. (2005). P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. *Mol. Pharmacol.* 67, 365–374.

Raven, J. A. (1975). Transport of indoleacetic-acid in plant-cells in relation to Ph and electrical potential gradients, and its significance for polar IAA transport. *New Phytol.* 74, 163–172.

Rea, P. A. (2007). Plant ATP-binding cassette transporters. *Annu. Rev. Plant Biol.* 58, 347–375.

Rojas-Pierce, M., Titapiwatanakun, B., Sohn, E. J., Fang, F., Larive, C. K., Blakeslee, J., Cheng, Y., Cutler, S. R., Peer, W. A., Murphy, A. S., and Raikhel, N. V. (2007). *Arabidopsis* P-glycoprotein19 participates in the inhibition of gravitropism by gravacin. *Chem. Biol.* 14, 1366–1376.

Rosenberg, M. F., Callaghan, R., Modok, S., Higgins, C. F., and Ford, R. C. (2005). Three-dimensional structure of P-glycoprotein: the transmembrane regions adopt an asymmetric configuration in the nucleotide-bound state. *J. Biol. Chem.* 280, 2857–2862.

Rothnie, A., Storm, J., Campbell, J., Linton, K. J., Kerr, I. D., and Callaghan, R. (2004). The topography of transmembrane segment six is altered during the catalytic cycle of P-glycoprotein. *J. Biol. Chem.* 279, 34913–34921.

Rubery, P. H., and Sheldrake, A. R. (1974). Carrier-mediated auxin transport. *Planta* 118, 101–121.

Sakurai, A., Onishi, Y., Hirano, H., Seigneuret, M., Obanayama, K., Kim, G., Liew, E. L., Sakaeda, T., Yoshiura, K., Niikawa, N., Sakurai, M., and Ishikawa, T. (2007). Quantitative structure – activity relationship analysis and molecular dynamics simulation to functionally validate nonsynonymous polymorphisms of human ABC transporter ABCB1 (P-glycoprotein/MDR1). *Biochemistry* 46, 7678–7693.

Sali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.

Sanchez-Fernandez, R., Davies, T. G., Coleman, J. O., and Rea, P. A. (2001a). The *Arabidopsis* thaliana ABC protein superfamily, a complete inventory. *J. Biol. Chem.* 276, 30231–30244.

Sanchez-Fernandez, R., Rea, P. A., Davies, T. G., and Coleman, J. O. (2001b). Do plants have more genes than humans? Yes, when it comes to ABC proteins. *Trends Plant Sci.* 6, 347–348.

Santelia, D., Vincenzetti, V., Azzarello, E., Bovet, L., Fukao, Y., Duchtig, P., Mancuso, S., Martinoia, E., and Geisler, M. (2005). MDR-like ABC transporter AtPGP4 is involved in auxin-mediated lateral root and root hair development. *FEBS Lett.* 579, 5399–5406.

Seeliger, D., and de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.* 24, 417–422.

Shapiro, A. B., and Ling, V. (1997). Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur. J. Biochem.* 250, 130–137.

Sharom, F. J. (2008). ABC multidrug transporters: structure, function and role in chemoresistance. *Pharmacogenomics* 9, 105–127.

Storm, J., Modok, S., O'mara, M. L., Tieleman, D. P., Kerr, I. D., and Callaghan, R. (2008). Cytosolic region of TM6 in P-glycoprotein: topographical analysis and functional perturbation by site directed labeling. *Biochemistry* 47, 3615–3624.

Sugiyama, A., Shitan, N., Sato, S., Nakamura, Y., Tabata, S., and Yazaki, K. (2006). Genome-wide analysis of ATP-binding cassette (ABC) proteins in a model legume plant, *Lotus japonicus*: comparison with *Arabidopsis* ABC protein family. *DNA Res.* 13, 205–228.

Sussman, M. R., and Gardner, G. (1980). Solubilization of the receptor for N-1-naphthylphthalamic acid. *Plant Physiol.* 66, 1074–1078.

Tan, X., Calderon-Villalobos, L. I., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M., and Zheng, N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 446, 640–645.

Terasaka, K., Blakeslee, J. J., Titapiwatanakun, B., Peer, W. A., Bandyopadhyay, A., Makam, S. N., Lee, O. R., Richards, E. L., Murphy, A. S., Sato, F., and Yazaki, K. (2005). PGP4, an ATP binding cassette P-glycoprotein, catalyzes auxin transport in *Arabidopsis thaliana* roots. *Plant Cell* 17, 2922–2939.

Theodoulou, F. L. (2000). Plant ABC transporters. *Biochim. Biophys. Acta* 1465, 79–103.

Titapiwatanakun, B., and Murphy, A. S. (2009). Post-transcriptional regulation of auxin transport proteins: cellular trafficking, protein phosphorylation, protein maturation, ubiquitination, and membrane composition. *J. Exp. Bot.* 60, 1093–1107.

Wang, H. Y., and Lin, R. C. (2005). Two homologous ATP-binding cassette transporter proteins, AtMDR1 and AtPGP1, regulate *Arabidopsis* photomorphogenesis and root development by mediating polar auxin transport. *Plant Physiol.* 138, 949–964.

Ward, A., Reyes, C. L., Yu, J., Roth, C. B., and Chang, G. (2007). Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19005–19010.

Wu, G., Lewis, D. R., and Spalding, E. P. (2007). Mutations in *Arabidopsis* multidrug resistance-like ABC transporters separate the roles of acropetal and basipetal auxin transport in lateral root development. *Plant Cell* 19, 1826–1837.

Yang, H., and Murphy, A. S. (2009). Functional expression and characterization of *Arabidopsis* ABCB, AUX 1 and PIN auxin transporters in *Schizosaccharomyces pombe*. *Plant J.* 59, 179–191.

Zazimalova, E., Murphy, A. S., Yang, H., Hoyerova, K., and Hosek, P. (2010). Auxin transporters – why so many? *Cold Spring Harb. Perspect. Biol.* 2, a001552.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 October 2011; accepted: 17 December 2011; published online: 05 January 2012.

*Citation: Bailly A, Yang H, Martinoia E, Geisler M and Murphy AS (2012) Plant lessons: exploring ABCB functionality through structural modeling. Front. Plant Sci. **2**:108. doi: 10.3389/fpls.2011.00108*

This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

APPENDIX

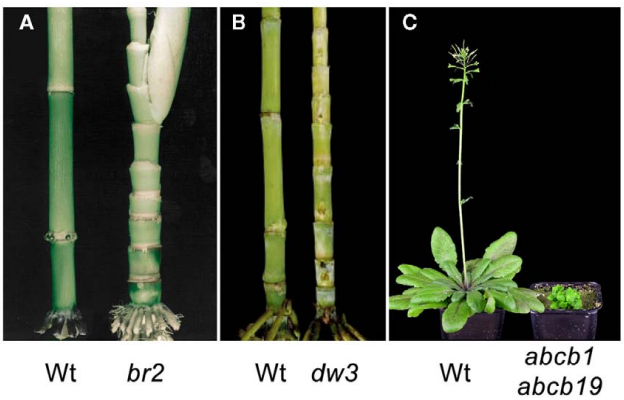


FIGURE A1 | Mutations in ABCB1 orthologs result in reduced basipetal auxin transport and dwarfism. (A,B) Compressed stalk internodes in maize *br2* (*brachytic2/zmpgp1*) and sorghum *dw3* (*dwarf3/sbpgp1*) mutants (Multani et al., 2003). **(C)** Drastic dwarfing phenotype in *Arabidopsisabcb1 abcb19* double mutant (Geisler et al., 2005). Wt, wild type.

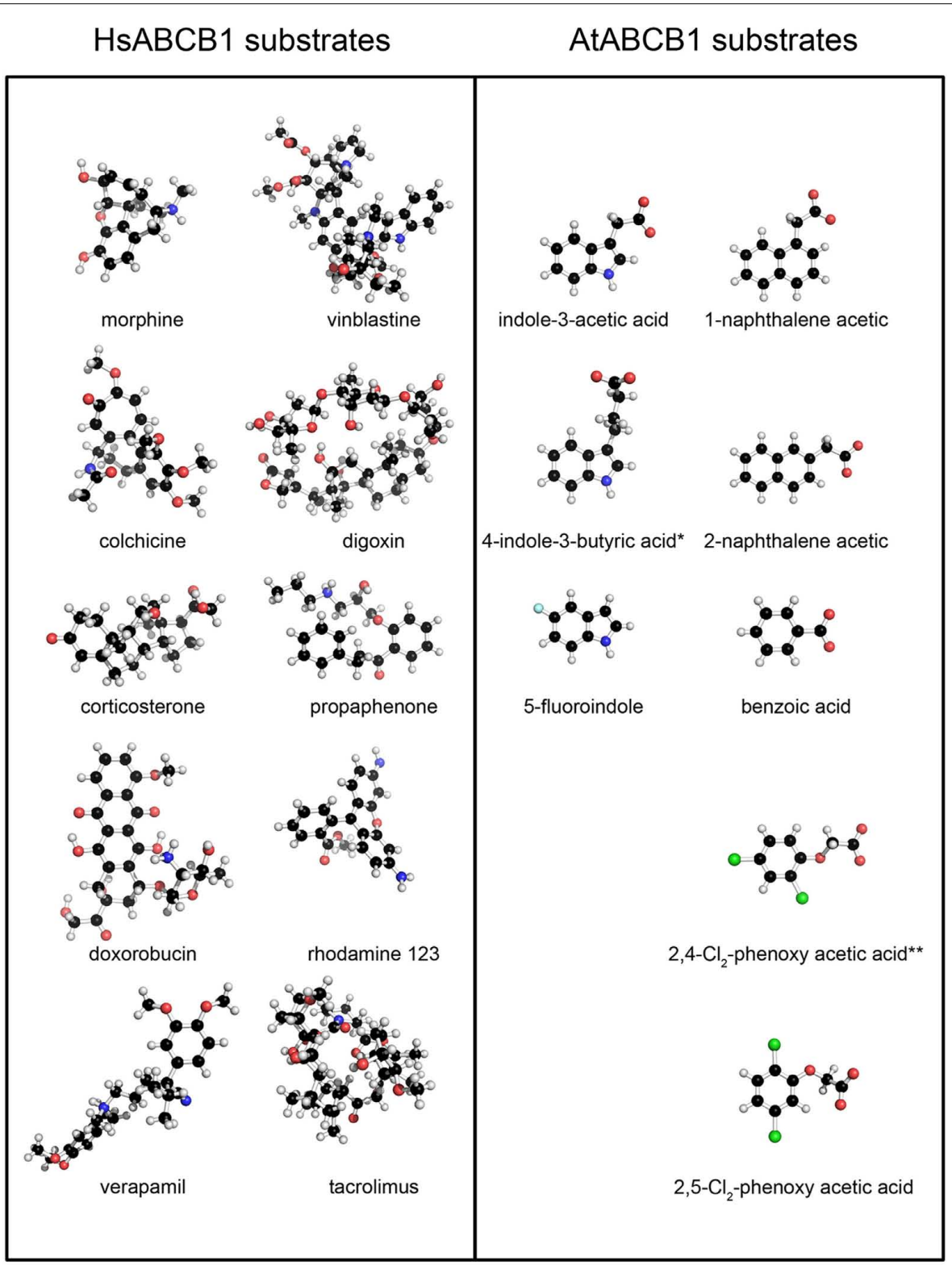


FIGURE A2 | Mammalian and plant ABCB1 proteins show divergent substrate selectivities. Structural representations of discrete HsABCB1 and AtABCB1 substrates evidence the extreme variance in the reported mammalian substrates when compared to the similarity of AtABCB1 substrates. *No experimental evidence for those compounds has been reported to date, although their biological activity in polar auxin transport has been extensively documented.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 28 March 2012
doi: 10.3389/fpls.2012.00034

High affinity ammonium transporters: molecular mechanism of action

Omar Pantoja*

Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, México

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
Yi-Fang Tsay, Academia Sinica, Taiwan
Li-Qing Chen, Carnegie Institution for Science, USA

****Correspondence:***
Omar Pantoja, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, Cuernavaca, Morelos 62250, México.
e-mail: omar@ibt.unam.mx

The importance of the family of high affinity ammonium transporters is demonstrated by the presence of these proteins in all domains of life, including bacteria, archaea, fungi, plants, and humans. The majority of the proteins that have been studied from this family show high affinity and selectivity for ammonium, are impermeable to alkaline cations, saturate rapidly at low millimolar concentrations and most of them, are also permeable to methylammonium. Crystallization of homologue proteins from bacteria and archaea has demonstrated that the functional entity corresponds to a trimer, with each monomer maintaining a conductive pore. Through molecular modeling, it has been demonstrated that even though the identity of the proteins between bacteria/archaea with those from plants is below 25%, the latter seem to maintain similar tertiary and quaternary structures, an observation that has helped to address the functionality of conserved residues by means of mutational analysis. Results have shown that changes in the extracellular binding site of some plant transporters may result in their inhibition or reduction in transport activity, while in *Escherichia coli*, dissimilar replacements like Phe/Ala or Ser/Leu that eliminate possible π-interactions or H-bonds with ammonium, respectively, lead to more active transporters. Active mutants with changes in the pair of conserved His in the center of the transporter suggest these residues are dispensable. Additional mutations have identified other important amino acids, both in the entrance of the pore and in cytoplasmic loops. Regulation of this family of transporters can be achieved by interactions of the C-terminal with cytoplasmic loops within the same monomer, or with a neighbor in the trimer. Depending on the interacting residues, these contacts may lead to the activation or inhibition of the protein. The aim of this review is to critically evaluate the newest findings on the role of the proposed amino acids that structure the ammonium pathway, as well as highlight the importance of additional residues that have been identified through mutational analyses.

Keywords: ammonium, transport, AMT, high affinity

INTRODUCTION

Ammonium is the preferred nitrogen source for many plants, and its transport across cellular membranes is carried out by a family of integral membrane proteins that is conserved throughout all domains of life, the AMT/MEP/Rh family (Ammonium Transporter/Methylammonium Permease/mammalian Rhesus proteins; Wirén and Merrick, 2004; Winkler, 2006). Members of this family of transporters share several properties, among them, high affinity (K_m in the micromolar range) and high selectivity for ammonium [transporting exclusively ammonium and its methylated analog, methylammonium (MeA), but no other monovalent cation] and saturation at ammonium concentrations ≤1 mM (Gazzarrini et al., 1999; Shelden et al., 2001; Ludewig et al., 2002; D'Apuzzo et al., 2004; Wood et al., 2006; Couturier et al., 2007; Boeckstaens et al., 2008; Ortiz-Ramirez et al., 2011). In spite of these common characteristics, it is not clear what is the chemical species that is transported, ammonium or ammonia, or the actual transport mechanism represented by these proteins, either a channel or a transporter (Khademi et al., 2004; Conroy et al., 2007). However, important progress in the understanding of ammonium transporters has been derived from the crystal structure of AmtB from *Escherichia coli* (Khademi et al., 2004; Zheng et al., 2004; Conroy et al., 2007) and Amt1 from *Archaeoglobus fulgidus* (Andrade et al., 2005). These studies have demonstrated that the ammonium transporters are conformed as homotrimers, with each monomer containing a central hydrophobic channel, proposed to be the pathway through which the uncharged ammonia is transported (Khademi et al., 2004; Zheng et al., 2004; Andrade et al., 2005). Taking as a reference the structure for the *E. coli* ammonium transporter (*Ec*AmtB), it is proposed that the putative selective ammonium pathway is structured by a binding/recruitment site facing the periplasmic side; a so-called phenylalanine gate composed by two conserved Phe residues; a central section characterized by the presence of two highly conserved His and a cytoplasmic vestibule (Khademi et al., 2004; Zheng et al., 2004). It is suggested that in the recruitment site, S219 and W148, and well conserved in the AMT/MEP/Rh family, can bind ammonium through a hydrogen bonding by the former and through the π system of the latter (Khademi et al., 2004), although in the crystal structure of the Amt1 homologue from *A. fulgidus* (*Af*Amt1;1) this site is occupied

by water (Andrade et al., 2005). The Phe gate is structured by the F107 and F215 pair that seems to block the entrance to the putative channel. Deeper into the channel of *Ec*AmtB, the central section is partially structured by His168 and His318 which are proposed to deprotonate ammonium and to bind the substrate ammonia, facilitating thus the transport of the latter. Finally, in the cytoplasmic vestibule, the side chain of V314 occupies the pore exit, and depending on the crystal structure, it may exist in a closed (P63 structure) or open (R3 structure) state, which may suggest V314 plays a role in controlling ammonium transport, that in the case of *Ec*AmtB, may induce the binding of the regulatory GlnK protein inhibiting ammonium transport (Zheng et al., 2004).

SEQUENCE ANALYSIS OF AMT/MEP/RH TRANSPORTERS

The genome of higher plants contain two families of ammonium transporters, AMT1 and AMT2 (Sohlenkamp et al., 2000; Simon-Rosin et al., 2003; Couturier et al., 2007). Both family members show a high affinity for ammonium (in the micromolar range), but those from the AMT2 family are unable to transport the ammonium analog, methylammonium; possess a smaller transport capacity, and ammonium transport by *At*AMT2 has been shown to be electroneutral, suggesting that the transported species corresponds to the uncharged NH_3 and not NH_4^+ (Sohlenkamp et al., 2000, 2002; Simon-Rosin et al., 2003; Neuhäuser et al., 2009). From a pair wise analysis (**Figure 1**) of the sequences shown in **Figure 2**, it can be observed that among the plant AMT1 proteins there exists a high degree of identity between them, varying between 65 and 91%, with the exception of *Lycopersicon esculentum Le*AMT1;3 which on average, shows a slightly lower identity with the other members, ≤60%. In comparison, identity between the AMT1 and AMT2 families is lower and accounting for only 20–25%, as has previously been reported (Sohlenkamp et al., 2000). Identity of the plant AMT1 proteins with their homologues from fungi and bacteria is also low with values close to 20–25% (**Figure 1**). Identity between AMT2 homologues from plants and those from fungi and bacteria falls within similar low values (≤20%). These observations are better appreciated in the phylogenetic tree in **Figure 3**, which clearly shows that the AMT1 family members group in a closely related cluster, while those from the AMT2, fungi, archaea and bacteria branch off and form a separate cluster. In the latter, the low identity levels among these members (**Figure 3**), as represented by the length of the branches, clearly indicates that plant AMT2, archaea, bacteria, and fungi proteins form distantly related groups. In spite of these differences, and as mentioned before, the majority of these proteins show high affinity and high selectivity for ammonium, which indicates that among the 20–25% conserved residues must reside those directly involved in the mechanism of transport that may correspond to those proposed to structure the ammonium pathway (Khademi et al., 2004; Zheng et al., 2004; Andrade et al., 2005).

THE AMMONIUM BINDING/RECRUITING SITE

As postulated by several groups, among the conserved amino acids are W148 and S219, in *Ec*AmtB, proposed to structure the ammonium binding site (Khademi et al., 2004; Zheng et al., 2004; Andrade et al., 2005). According to the alignment shown in **Figure 2**, most of these residues are well conserved, with the exception of S219 which is replaced by Ala or Asp in plant AMT2 proteins. The more detailed analyses on the putative pore have been made on the *E. coli* homologue *Ec*AmtB. Most results on the

Number of Identical Residues

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
LjAMT1;1	1		415	407	395	386	377	386	406	369	365	381	371	299	163	117	126	116	120	117	116	112
PvAMT1;1-2	2	82.67		393	370	366	372	381	396	351	358	375	365	300	166	114	122	117	122	114	113	113
LeAMT1;1	3	80.91	78.44		381	382	363	364	394	357	356	365	361	307	160	111	120	111	124	116	113	113
LeAMT1;2	4	76.55	71.71	73.84		400	362	361	372	373	369	366	369	299	168	115	117	117	125	119	122	114
LjAMT1;2	5	74.09	70.25	73.60	76.78		366	367	375	365	370	371	370	308	172	111	119	115	125	116	120	112
AtAMT1;5	6	74.07	73.08	71.32	70.29	70.25		454	410	351	362	356	356	288	170	123	124	121	130	122	119	114
AtAMT1;3	7	75.69	74.71	71.37	70.10	70.44	91.16		419	349	363	357	353	291	168	117	124	116	127	118	115	114
AtAMT1;1	8	79.45	77.65	77.71	72.23	72.12	81.03	82.81		352	367	370	367	305	169	118	123	118	131	116	113	114
LjAMT1;3	9	72.35	68.82	70.00	71.87	70.06	68.42	67.90	68.22		386	359	366	296	168	115	116	114	131	111	105	109
AtAMT1;4	10	71.43	70.06	69.67	70.96	70.88	70.57	70.62	70.99	75.83		358	359	290	164	113	114	112	132	113	114	113
TeAMT1;2	11	72.99	71.98	70.19	69.32	70.27	68.20	68.39	71.02	68.64	68.45		433	302	161	114	119	120	121	122	116	117
TeAMT1;1	12	72.04	70.87	70.51	70.83	71.02	69.13	68.54	71.26	70.93	69.57	85.91		298	161	114	117	120	121	115	116	113
LeAMT1;3	13	58.17	58.48	60.43	57.50	58.67	56.25	56.84	59.22	57.70	56.20	58.75	58.78		158	115	115	115	125	108	113	108
AfAMT2	14	31.29	31.86	30.95	31.94	32.70	33.01	32.56	32.50	31.94	31.48	30.67	31.14	30.80		105	103	117	128	102	96	95
Yeast MEP3	15	20.97	20.43	19.89	20.35	19.47	21.96	20.86	21.03	20.35	19.96	20.54	20.54	21.42	19.23		386	197	152	130	118	124
Yeast MEP1	16	22.50	21.79	21.43	20.74	20.99	22.14	22.10	21.96	20.53	20.14	21.17	21.08	21.26	18.97	77.35		201	146	127	117	126
Yeast MEP2	17	21.05	21.23	20.15	20.82	20.28	21.84	20.90	21.18	20.58	20.18	21.51	21.78	21.58	21.47	37.74	38.65		149	127	121	129
EcAmtB	18	23.21	23.64	24.36	23.58	23.50	25.29	24.61	25.05	25.19	25.58	23.05	23.45	25.25	26.12	29.69	28.52	29.62		137	133	141
AtAMT2	19	21.12	20.61	21.21	21.06	20.35	22.02	21.26	20.79	19.96	20.29	22.10	20.91	20.57	19.03	24.39	23.65	23.92	27.18		364	289
LjAMT2;1	20	20.60	20.11	20.25	21.25	20.69	21.14	20.39	19.86	18.58	20.14	20.53	20.68	21.12	17.61	21.69	21.43	22.41	25.93	74.59		307
LjAMT2;2	21	20.36	20.58	20.73	20.32	19.72	20.58	20.54	20.50	19.71	20.32	21.16	20.58	20.57	17.69	23.26	23.51	24.57	27.81	58.38	60.79	

% Identity

FIGURE 1 | Pairwise analysis of AMT/MEP transporters. Employing the sequence from *Pv*AMT1;1 and the CLC Main Workbench software (CLC bio, Katrinebjerg, Aarhus N, Denmark) 106 most conserved sequences related to AMTs were obtained from NCBI using BLAST. Twenty-one sequences that are mentioned in the text were aligned, and compared (Neighbor Joining; 100 bootstraps), to obtain the number of conserved residues, *upper half* and percent identity between the proteins, *bottom half*. The shades of red and blue color represent, respectively, the maximal and minimal number of identical residues, *upper half*; or per cent of identity, *lower half*, between the AMT/MEP transporters.

role of the residues localized in the ammonium recruiting/binding site have demonstrated that substitution of F107, W148, or S219 with either Ala or Leu, do not modify the expression level of *Ec*AmtB, although some changes were observed in the transport properties (Fong et al., 2007; Javelle et al., 2008; Hall and Kustu, 2011). Mutations in F107A, W148A-L, or S219A increased transport activity between 2- and 10-fold while the rest of the mutants showed similar levels to those of the wild type (*wt*; Javelle et al., 2008; Hall and Kustu, 2011). Inhibition of MeA transport by thallium in all these mutants was similar to the *wt*, indicating that the microenvironment is not severely affected for MeA binding to occur. Inhibition by thallium however, was reduced (15–20%) in the triple mutant F107A/W148A/S219A. Crystal analyses of some of these variants showed that the mutations did not modify the general structure of the transporter, with the changes being limited to the mutated residue or with minor changes on neighboring amino acids (Javelle et al., 2008). These results indicate that the proposed role for the aromatic amino acids in NH_4^+/MeA binding through cation-π interactions or to S219 by H-bonding (Khademi et al., 2004) is dispensable, and that these substrates can apparently, move freely through the periplasmic/extracellular region. An additional residue, D160 has been shown to play an

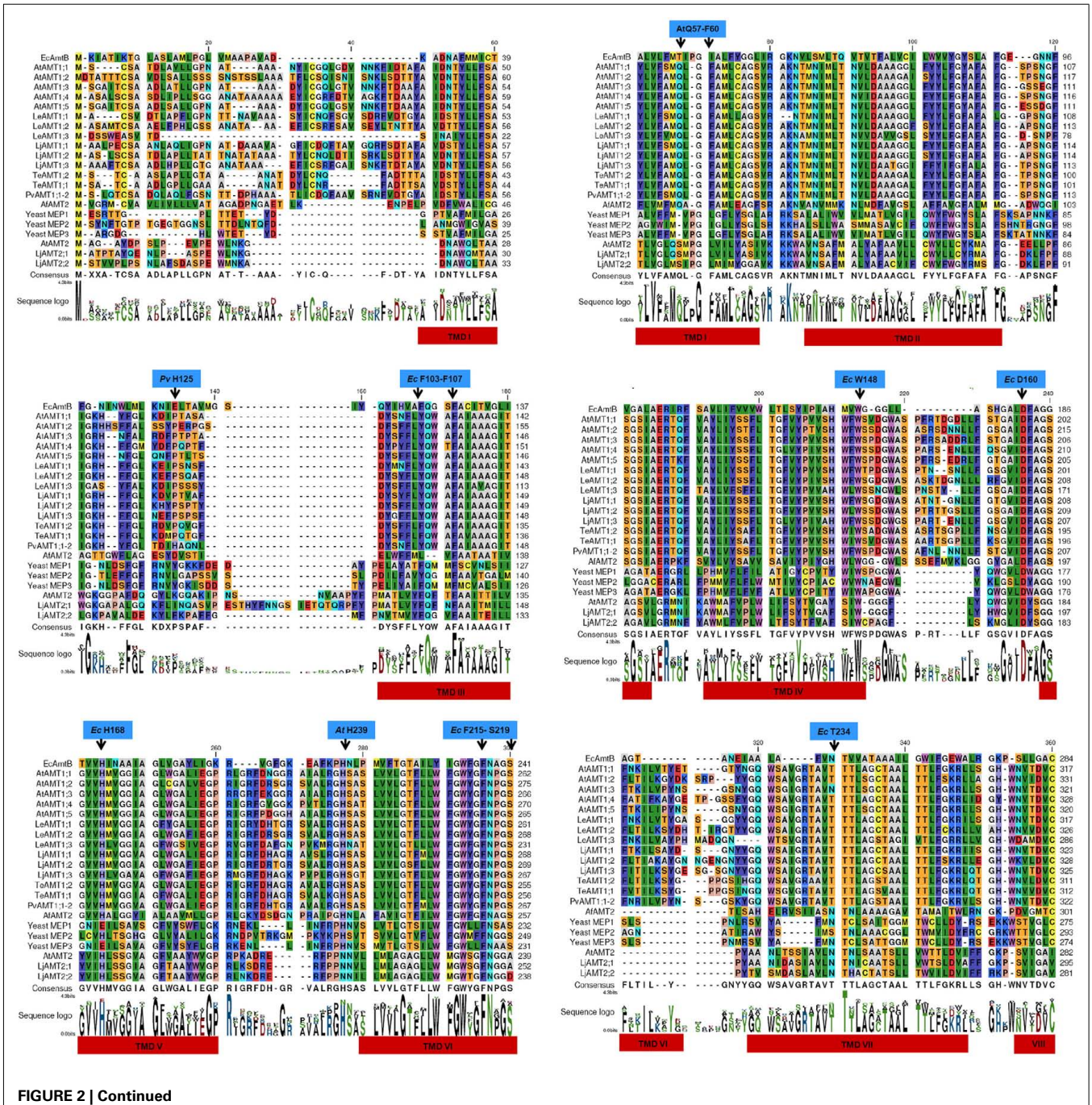


FIGURE 2 | Continued

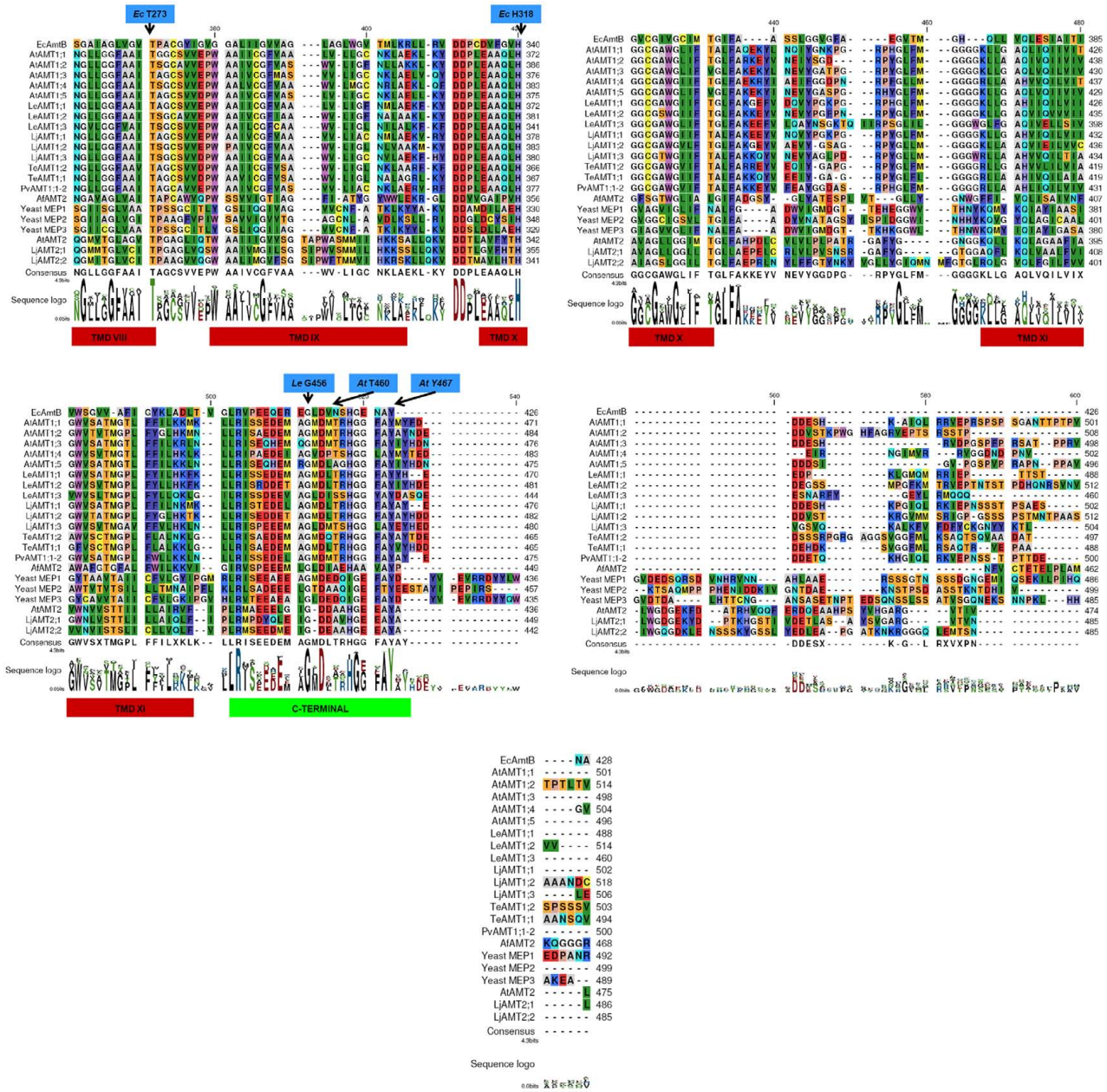


FIGURE 2 | Alignment of 21 AMT/MEP transporters. Alignment of the 21 proteins reviewed in the text were aligned with the CLC Main Workbench software (CLC bio, Katrinebjerg, Aarhus N, Denmark). Amino acids identified as constitutive of the pore in *E. coli* AmtB are marked at the top of the sequences in blue squares; transmembrane domains are shown in red rectangle at the bottom of the sequences; the C-terminal is highlighted by a green rectangle.

important role by mutational analysis, where substitution by Ala renders the protein inactive, and thus, proposed to be the ammonium binding site (Javelle et al., 2004). However, in the crystal structure of *Ec*AmtB (Khademi et al., 2004), a density peak associated with the presence of high ammonium or MeA concentrations was not observed close to D160, that besides being inaccessible to the extracellular solution, suggest that the role of this residue may be structural (Khademi et al., 2004). Mutations in the equivalent residue in yeast (*Saccharomyces cerevisiae*) MEP transporters (*Sc*MEP2 D186N; Marini et al., 2006) and *Arabidopsis thaliana* (*At*AMT1;1 D198N; Loqué et al., 2007) also inhibited ammonium transport, thus confirming the importance of this Asp located in the periplasmic/extracellular side of the proteins.

In plants, attempts have been made to ascertain the importance of W178 and Y133 in *Le*AMT1;1 from tomato (W148 and F103 in *Ec*AmtB). The mutation W178L transformed the transporter into an inactive mechanism, unable to generate currents with ammonium or to transport the analog MeA (Mayer et al.,

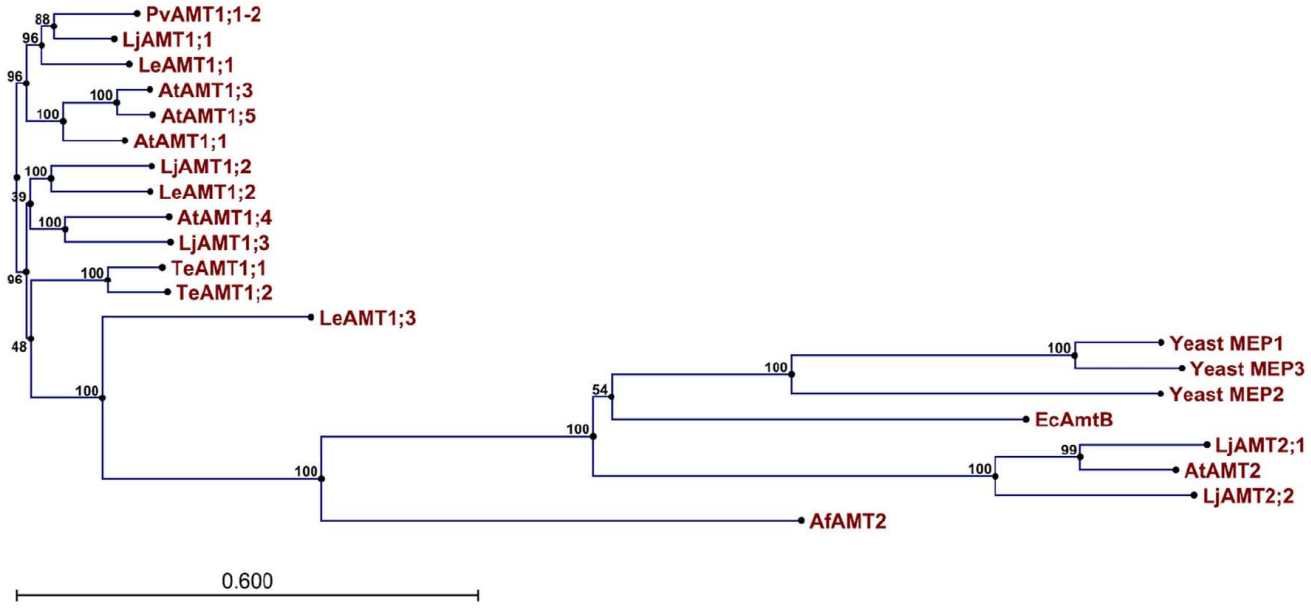


FIGURE 3 | Phylogenetic tree of AMT/MEP transporters. The phylogenetic tree was calculated with the CLC Main Workbench software (CLC bio, Katrinebjerg, Aarhus N, Denmark) employing the 21 sequences described in the text.

2006). On the contrary, the change Y133I induced an increase in ammonium affinity (10-fold), together with a significant decrease in transport capacity (10-fold smaller currents; Mayer et al., 2006). The mutations W178L and Y133I were made to mimic the binding site found in the human Rh homologue glycoproteins, transporters that carry out low affinity and electroneutral ammonium transport (Mayer et al., 2006). The properties shown by the W178L and Y133I mutants indicate that substitutions of these two amino acids does not transform the tomato transporter into an Rh-like protein and that the transport properties of the latter must depend on residues located elsewhere.

Changes in the putative ammonium binding site occur naturally in the members of the AMT2 plant family, where the residue equivalent to S219 in *Ec*AmtB and proposed to form an H-bond with ammonium, is substituted either, by Ala in *At*AMT2;1 (A239) and (*Lotus japonicus*) *Lj*AMT2;1 (A252) or Asp in *Lj*AMT2;2 (D238; see **Figure 2**). *Lj*AMT2;1 and *At*AMT2;1 seem to transport uncharged ammonia, with a low transport capacity (low V_{max}) and are impermeable to MeA (Simon-Rosin et al., 2003; Neuhäuser et al., 2009). Interestingly, *Lj*AMT2;2 a mycorrhizal specific ammonium transporter is also electroneutral, but it does transport MeA with a relatively high affinity ($K_m = 0.83$ mM). The natural substitution of S219 by Ala in *At*AMT2;1 and *Lj*AMT2;1 together with the results from the *Ec*AmtBS219A mutant challenge the proposed role of S219 in ammonium binding through an H-bond and suggests that other amino acids must be involved in ammonium binding.

THE PHENYLALANINE GATE

The Phe gate, structured in *Ec*AmtB by the F107 and F215 pair, it seems to block the entrance to the putative channel (Khademi et al., 2004; Andrade et al., 2005). Both residues are well conserved in all the analyzed proteins (**Figure 2**). Interestingly, while the single mutant F215A, was well expressed in the membrane, did not carry out MeA uptake (Javelle et al., 2008) nor grew on low ammonium concentrations (Hall and Kustu, 2011), the F107A mutant behaved very much like the *wt* transporter (Javelle et al., 2008; Hall and Kustu, 2011). Further mutations in F215 (H, Q, S, or W) did not transport MeA (Javelle et al., 2008), just like the mutant F215L (Javelle et al., 2008; Hall and Kustu, 2011), which however, grew slightly slower than the *wt* in micromolar ammonium (Hall and Kustu, 2011). It is important to point out the contrasting results with the F215L mutant which led to the conclusion by one group that F215 plays a crucial role in the functioning of the transporter based on its inability to transport MeA (Javelle et al., 2008), while another group proposed that F215 and the associated π-interactions that they may establish with ammonium are not required for the functioning of *Ec*AmtB, as indicated by the similar growth rates shown by the F215L mutant and the *wt* (Hall and Kustu, 2011). Interestingly, the double (W148L/F215L) and triple (F107A/W148L/F215L) mutants did not differ much from the wt when grown in low ammonium, although both did not achieve MeA transport (Hall and Kustu, 2011).

Intuitively, mutation on the Phe gate to F107A/F215A has potential structural effects, generating an open pore, which would suggest that ammonium and other ions with similar radius could freely permeate through the mutant transporter. However, neither MeA, K^+, H^+ nor water were transported by the F107A/F215A double mutant or by *Ec*AmtB. It remains to be demonstrated if this mutant is capable to grow at low ammonium concentrations before reaching a definitive conclusion on the effects of this double substitution (Javelle et al., 2008).

THE ROLE OF His168 AND His318 IN AMMONIUM DEPROTONATION AND pH-DEPENDENCE

The pair of His postulated to deprotonate NH_4^+ and facilitate the movement of uncharged NH_3 (H168 and H318) is also well conserved, excluding the yeast *Sc*MEP1 and *Sc*MEP3 proteins where the residue equivalent to His168 corresponds to E181 and E180, respectively (**Figure 3**). Analysis of 14 mutants in *Ec*AmtB where the His residues were replaced by aliphatic, aromatic, or polar residues, demonstrated that most of them were expressed in the membrane and maintained their trimeric structure. Six of the polar mutants were also expressed in the membrane but the apparent formation of monomers and dimers indicated that they were rather unstable (Javelle et al., 2006). From those mutants that were crystallized, it was observed that the structural changes were restricted to the mutated His, with small effects on adjacent residues (Javelle et al., 2006). From all the analyzed mutants, only H168E and H168D showed reduced MeA uptake that accounted for 15–25% of the *wt* activity (Javelle et al., 2006; Hall and Kustu, 2011), although growth in low ammonium was normal for the H168E mutant (Hall and Kustu, 2011), Functionality of the H168E mutant should not be surprising since this substitution is naturally present in the *Sc*MEP1 and *Sc*MEP3 homologues from yeast (**Figure 2**), both being able to transport ammonium and MeA (Marini et al., 2000). Recently, it was reported that the same substitution (H211E) in *Pv*AMT1;1 from bean (*Phaseolus vulgaris*), caused an increase in ammonium current and a large decrease in MeA transport (Ortiz-Ramirez et al., 2011). *Pv*AMT1;1, like most of the AMT1 family members shows high affinity ($K_m = 15\,\mu M$ and $V_{max} = 140$ nA), and high selectivity for ammonium, and saturates at concentrations below 1 mM; however, it is pH-dependent, being more active in acidic extracellular environments. Associated to the higher activity at acidic external pH, ammonium transport (currents) in *Xenopus* oocytes was accompanied by cytoplasm acidification. Changes in both, extracellular ammonium and proton concentration induced changes in the reversal potential of the ammonium currents with slopes of 33 and 26 mV/decade, respectively, values that were used to calculate an stoichiometry of $1\,NH_4^+/1\,H^+$, indicating that *Pv*AMT1;1 operates as an NH_4^+/H^+ symporter (Ortiz-Ramirez et al., 2011). The mutation H211E transformed the protein into a pH-independent mechanism, with a decreased affinity and increased transport capacity for ammonium ($K_m = 1.3$ mM and $V_{max} = 300$ nA), accounting for 40- and 2-fold increases in these two parameters, respectively. Additionally, the reversal potential of the ammonium currents changed according to the Nernst equation, indicating that ammonium was the only ion transporter by the mutant *Pv*AMT1;1H211E, and confirmed by the absence of cytoplasm acidification (Ortiz-Ramirez et al., 2011). These results indicate that H211 positioned in the middle of the channel can be substituted by Glu, as it is naturally for other transporters. The *Sc*MEP1 and *Sc*MEP3 homologues from yeast present a similar arrangement to the H211E mutant, in which the first His is substituted by Glu, and show similar transport properties. *Sc*Mep1 and *Sc*Mep3 are pH-independent, have a low MeA affinity and apparently, larger transport capacity as evidenced by the toxic effects of MeA (Boeckstaens et al., 2008). In contrast, *Sc*Mep2 maintains the first conserved His, like the wt bean transporter, and also shares the high affinity for ammonium and pH-dependence (Boeckstaens et al., 2008). These properties become similar to those of its paralogs or the H211E bean mutant if the first His is mutated to Glu (Boeckstaens et al., 2008). Analogous manipulation in the *Ec*AmtB protein lead to similar results to those observed for *Sc*Mep2 (Boeckstaens et al., 2008). All these observations confirm that the first conserved histidine can be replaced by Glu without inhibiting ammonium transport, and suggest that this residue in association with as yet, unidentified amino acids, are responsible for the pH sensitivity of some members of this family.

Analysis of the four double mutants (H168E/H318E, H168D/H318D, H168D/H318E, and H168E/H318D) demonstrated that all were severely affected, unable to grow in low ammonium or to transport MeA, as a result of low levels of expression. Employing these mutants to isolate suppressor mutants led to the identification of I110N as an important modification (Hall and Kustu, 2011). The mutants H168D/H318D, H168D/H318E were both rescued by I110N, showing a slightly longer duplication rates in low ammonium medium than the *wt*, but significantly better that the parental lines. It was proposed that Asn, being present in close proximity to the two acidic amino acids, at the other side of the pore, may help to decrease the strong electrostatic effect of Asp and Glu by means of H-bond interactions (Hall and Kustu, 2011). It is important to point out that neither of the rescued mutants was able to transport MeA, demonstrating, that the two assays for the evaluation of AMT activity, growth in low ammonium and MeA transport, do not give similar information. The studies on the pair of His show that when one of them is mutated to an acidic amino acid, the transporter is functional, but not when both are, implying that presence of two carboxy groups in the center, destabilize the protein, as suggested by the low levels of expression of the H168D/H318D, H168D/H318E mutants. The fact that ammonium and MeA can be transported by the suppressor I110N in the presence of hydrophilic residues in the pore suggests that presence of Asn may help to the stabilization of the protein and that charged NH_4^+ seems to be the transported species, implying that an hydrophobic pore is not a requisite for ammonium transport to occur.

THE IMPORTANCE OF THE C-TERMINAL IN THE REGULATION OF AMMONIUM TRANSPORTERS

Although the similarity of plant ammonium transporters with those from *E. coli* or *A. fulgidus* is somewhat low (20–25%), it is remarkable that molecular modeling has shown that plant proteins seem to maintain similar tertiary and quaternary structures, particularly the trans membrane domains (TMD), the cytoplasmic loops and the carboxy terminal (C-terminal), with the extracellular loops showing some divergence, mainly as a result of some of them being longer (Mayer et al., 2006; Loqué et al., 2007; Neuhäuser et al., 2007, 2009; Guether et al., 2009; Ortiz-Ramirez et al., 2011).

Confirmation that plant ammonium homologues can be modeled very closely to the crystal structure of bacteria and archaea, has open up the possibility to perform directed mutational analysis for the identification of additional amino acids that may be involved in the functioning of these transporters, as well as to

verify the importance of the conserved residues present in the putative ammonium pathway.

One of the first studies where the knowledge derived from the quaternary structure obtained for the ammonium transporter from *A. fulgidus* was applied, is that reported by Loqué et al. (2007) where modeling of *A. thaliana At*AMT1;1 demonstrated that the C-terminal established interactions with cytoplasmic loops of the same monomer, as well as with loops from neighbor monomers in the homotrimer. Based on the observed interactions of the C-terminal with cytoplasmic loops between TMDI–II, TMDIII–IV, and TMDV–VI of its own monomer, as well as with cytoplasmic loops connecting TMDI–II, TMDV–VI, and TMDVII–VIII of a neighboring monomer reported for *A. fulgidus* (Andrade et al., 2005), Loqué et al. (2007) analyzed the predicted interaction between Y467 in *At*AMT1;1 (Y448 in *Af*AMT1) located at the C-terminal and H239 (H234 in *Af*AMT1) positioned in the loop between TMDV–VI by mutating both residues. These authors observed that both mutants, Y467F and H239C, independently, were inactive and trans-inhibitory. More importantly, through a targeted suppressor screen, the interaction between these two residues was confirmed, since Y467F was only rescued by H239C, and vice versa. Thus, demonstrating that the interaction between the two regions of the protein (C-terminal and the cytoplasmic loop linking TMDV–VI), either through H or hydrophobic bonding is important in the activation of the transporter. Conservation of these two amino acids in most of the AMT proteins from plants (**Figure 2**), indicates that this regulatory mechanism is well conserved, and presence of the same pair of residues in *E. coli* and *A. fulgidus* suggests that it may be more widely common. Similar analyses on Thr460 from *At*AMT1;1, that seems to establish H-bonding within the C-terminal, showed that the partially inactive mutant T460A was rescued by mutations in residues located at cytoplasmic loops between TMDIII–IV (I146M and Q151E) of the same monomer and TMDV–VI (S242) and TMDVII–VIII (V313L) from a neighbor monomer, confirming the interactions observed in the crystal structure of *A. fulgidus*, and reinforcing the role of these associations in the regulation of the transporter (Loqué et al., 2007).

The importance of T460 in *At*AMT1;1 has been shown to be a result of its post-translational modification by phosphorylation (Lanquar et al., 2009). Initial studies had demonstrated that proteins carrying the mutation T460D were inactive, that when co-expressed with *wt* functioning transporters *At*AMT1;1 (Loqué et al., 2007) or *At*AMT1;2 (Neuhäuser et al., 2007), caused the trans-inhibition of the latter. Phosphorylation of T460 has been demonstrated to be dependent on ammonium concentration and time of exposure, as a mechanism proposed to prevent the continuous accumulation of this cation, that may become toxic if accumulated in the cytoplasm (Lanquar et al., 2009). Importantly, phosphorylation of T460 was caused rapidly and at low ammonium concentration (50 μM), and only induced by extracellular ammonium but not by any other cation, suggesting the existence of an ammonium receptor that directly or indirectly modifies the activity of the transporter. Alternatively, the same transporter could be functioning as a transceptor that relies the signal to an associated membrane kinase (Lanquar et al., 2009). Conservation of this regulatory mechanism may be common among the plant AMT1 homologues, where T460 is well conserved, with the exception of *At*AMT1;5 (A463). An Ala is also present in the equivalent position of plant AMT2 family members which indicates that in these homologues the regulatory role associated with T460 is not present. It remains to be shown if other mechanisms or residues in the C-terminal of the AMT2 transporters or *At*AMT1;5 participate in a similar regulatory mechanism.

An additional residue located at the C-terminal of plant transporters that plays an important role in maintaining the functionality of these proteins is G456 in *Le*AMT1;1 (Ludewig et al., 2003), first identified in the yeast *Sc*MEP1 homologue (G413; Marini et al., 2000). Mutation of G456 to Asp transforms *Le*AMT1;1 into an inactive transporter that also trans-inactivates its own *wt* or its paralog *Le*AMT1;2, evidence that first suggested that the AMT transporters could oligomerize. In contrast to T460, G456 is highly conserved among all the members of the MEP/AMT family (**Figure 2**), highlighting the importance of this residue in establishing trans-interactions between monomers that results in the activation of the transporters.

In view of all these results, it is important to highlight the role of the C-terminal in regulating the activity of plant AMT1 ammonium transporters, that as has been described above, it seems to be through direct interactions within the same or an adjacent monomer (Loqué et al., 2007; Neuhäuser et al., 2007). In bacteria, in contrast, the role of the C-terminal is dependent on its association with the non-uridylylated form of the cytoplasmic protein GlnK, interaction that results in the inhibition of *Ec*AmtB and ammonium transport, a mechanism that serves to regulate the uptake of this potentially toxic nutrient (Coutts et al., 2002). Interestingly, deletion of the C-terminal reduces the activity of *Ec*AmtB while completely inhibiting the functioning of *At*AMT1;1 (Marini et al., 2000; Loqué et al., 2007).

IDENTIFICATION OF ADDITIONAL AMINO ACIDS IMPORTANT IN THE FUNCTIONING OF PLANT AMT's

Among the plant AMT1 family members, an additional common property is the pH-independence of their activity (Ludewig et al., 2002, 2003; Wood et al., 2006), with the exception, so far, of *Pv*AMT1;1 and *Triticum aestivum Ta*AMT1;1 (Søgaard et al., 2009; Ortiz-Ramirez et al., 2011). In the wheat homologue, transport was stimulated by acidic external pH, with a half-maximal response (EC_{50}) of 50 μM ammonium that was constant at all levels of pH. Moreover, *Te*AMT1;1 does not seem to transport H^+, as indicated by the unchanged reversal potential in response to external pH changes (Søgaard et al., 2009). A mutation in the cytoplasmic side of the pore (L56F), caused the transporter to become pH-independent, to lose affinity for ammonium, and also associated with an increase in transport activity (Søgaard et al., 2009). From these results, Søgaard et al. (2009) postulated that the increased ammonium transport was a result of a larger diameter of the pore in the cytoplasmic side together with a facilitated H^+ transport. However, if this were the case, it would be expected that changes in both ammonium and H^+ concentrations would cause changes in the reversal potential of the currents, as has been reported for *Pv*AMT1;1 (Ortiz-Ramirez et al., 2011). Unfortunately, this information was not presented in the report by Søgaard et al. (2009).

What is intriguing is that a point mutation on the cytoplasmic side of the protein senses pH changes on the extracellular side, when it is expected that those residues directly involved in binding and/or sensing of regulatory factors, should be located near to the exposure side.

Studies on the role of pore structuring residues is still limited in plant ammonium transporters, however, participation of other amino acids in the mechanism of action of these proteins has been obtained by isolating suppressors of non-functional mutants. Recovering of the mutant T460A and the truncated *At*AMT1;1 (G456stop) protein was achieved individually by the mutations F60S, I136F, A137D, postulated to transform the transporter into a deregulated constantly active mechanism (Loqué et al., 2007). Two further suppressor mutants, Q57H and A138P, recovered the non-functional T460A and T460D proteins, showing a twofold larger ammonium uptake, and as a consequence, sensitivity of the yeast transformants to relative low concentrations of ammonium (Loqué et al., 2007, 2009). Interestingly, the Q57H and A138P mutants also conferred resistance to the toxic effects of MeA. Confirmation of these properties was obtained for *At*AMT1;1Q57H in *Xenopus* oocytes, were large and non-saturable ammonium currents were recorded, while recording an important decrease in the magnitude of the currents (10-fold lower) generated with MeA (Loqué et al., 2009). The kinetic properties of Q57H were substantially different to those of *At*AMT1;1, showing a 65-fold larger K_m for ammonium and a 10-fold larger V_{max} (Loqué et al., 2009). All these modified residues are located in the middle of TMDI (Q57 and F60) or TMDIII (I136, A137, and A138), forming part of the putative pore, an observation that was used to suggest that modifications to the structure of the pore not only release the regulation of the proteins by the C-terminal, but also change the mechanism of action, transforming the transporter into an ion channel (Loqué et al., 2007). A proposal that was partially confirmed by the electrical properties of the Q57H mutant, although failure in recording single channel activity from this protein (Loqué et al., 2009), leaves still open the confirmation of such proposal.

Identification of a unique His in *Pv*AMT1;1, which is replaced by Pro in all other plant homologues (**Figure 2**), prompted the authors to ascertain if this residue was responsible for the pH-dependence showed by the bean transporter. The mutation H125A transformed the transporter into a less active mechanism, but still pH-dependent, with higher activity under an acidic environment. In contrast, the mutant H125R, generated a more active protein with a V_{max} twice as large as that observed for the *wt*, but maintaining its stimulation by low pH (Ortiz-Ramirez et al., 2011). Interestingly, H125 is located in the extracellular loop joining TMDII and TMDIII indicating that substantial changes in the properties of the transporter can be caused by modification of residues not directly lining or structuring the putative conductive pore.

Continuing with the analysis of other conserved amino acids, it is important to recall an observation made by Zheng et al. (2004) who have pointed out that T234 and T273 are also well conserved as shown in **Figure 2**, and suggested that these residues may participate in structuring the side chains of the conserved residues surrounding the periplasmic side of the pore (ammonium binding site). Studying the crystal structure 1U7C and 1XQE of *Ec*AmtB (Khademi et al., 2004; Zheng et al., 2004), it was observed that

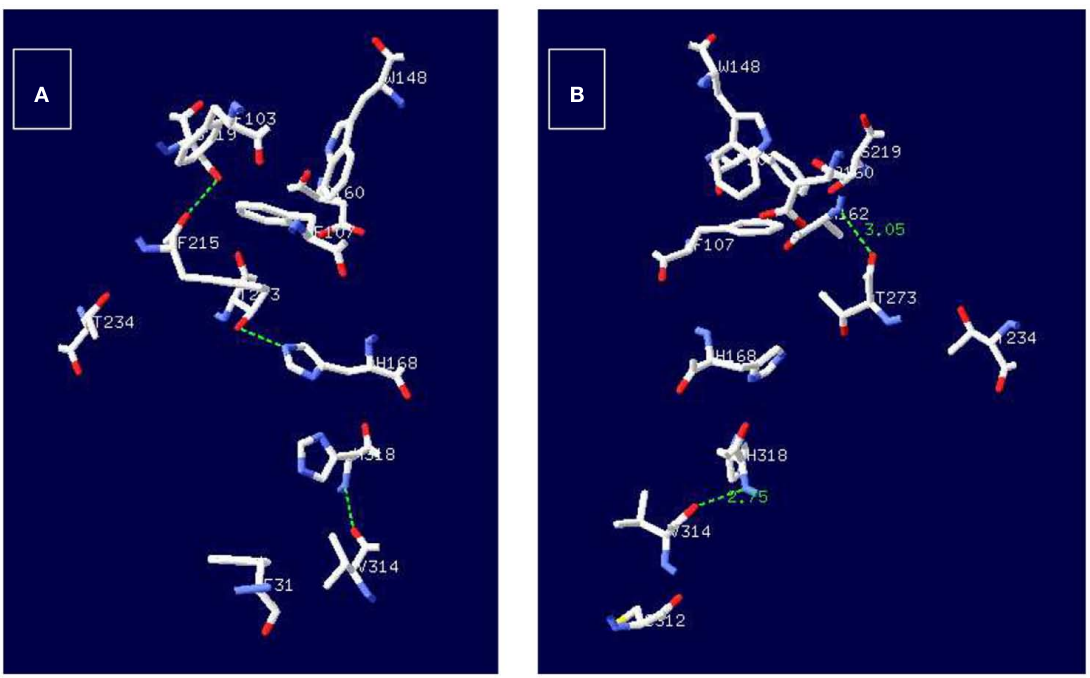


FIGURE 4 | Potential H-bond interactions of T273 with amino acids structuring the conductive ammonium pore in *Ec*AmtB. Employing the structures 1XQE **(A)** and 1U7C **(B)** of *Ec*AmtB and the program DeepView/Swiss-PdbViewer (http://www.expasy.org.spdbv) H-bonds were calculated and are shown as broken green lines between T273 and H128 **(A)** or T273 and A162.

T273 may form an H-bond with A162 located in the periplasmic side as well as with H168 (**Figure 4**), which indicates that T273 may have an important role in the functioning of the transporter. Unfortunately, so far, there is no evidence on the participation of T273 or A162 in the functioning of any of the AMT/MEP/Rh proteins which leaves open the analysis of the importance of these two residues. According to the pseudo twofold symmetry of the ammonium transporters (Khademi et al., 2004; Zheng et al., 2004; Andrade et al., 2005), the amino acid equivalent to D160, D310, is also very well conserved (**Figure 2**), and preceded by another well conserved D309, both located in the cytoplasmic side of the protein where it has been postulated to participate in structuring the backbone of trans membrane domain X (TMDX) through H-bonding (Zheng et al., 2004). So far, evidence on the importance of D310 (or D309) in *Ec*AmtB or its equivalent in any of the homologues is still lacking.

OBSERVATIONS ON THE USE OF MeA AS AN ANALOG OF AMMONIUM

Based on the results that have recently been reported on the modified properties of mutant transporters, where differences have been observed in their capacity to grow on low ammonium but unable to take up MeA, it is important to draw attention to the latter response. Radio-labeled MeA has been widely used for the evaluation of ammonium transport activity in a variety of organisms due to the availability of ^{14}C-labeled MeA and the short life of radioactive ^{13}N. The use on this ammonium analog is also attractive because it allows the parallel measurements of ^{14}C-MeA uptake and the currents associated to this transport, when electrogenic, allowing the estimation of the stoichiometry of the transporter, as it has been used by Ludewig et al. to demonstrated that NH_4^+ is the transported substrate by plant AMT1's (Mayer et al., 2006). However, the results reported for mutants like *Ec*AmtBH138E or H138D (Hall and Kustu, 2011), *At*AMT1;1Q57H (Loqué et al., 2009) and *Pv*AMT1;1H211E (Ortiz-Ramirez et al., 2011) that show an increased selectivity toward ammonium in comparison to MeA, underscore potential misinterpretations if the results are obtained exclusively from MeA uptake experiments. Further work on the analysis of AMT/MEP transporters or mutants should include evidence obtained from different methodologies to clearly demonstrate that the proteins do or do not carry out ammonium transport.

CONCLUSION

Our knowledge on the functioning of AMT/MEP/Rh proteins has advanced considerably in the last few years, and resolution of the crystal structure from the bacteria and archaea homologues, has also been of great help in guiding the research focused on identifying crucial amino acids in the functioning of these transporters. It seems that the residues initially postulated as directly involved in the mechanism of ammonium transport may not play a central role, as it has been reviewed here. Modifications to the putative binding site by residues with properties substantially different to the conserved aromatic amino acids that structure this extracellular domain, result in changes that do not completely abolished the activity of the transporter or that generated a more active mechanism (Loqué et al., 2009; Hall and Kustu, 2011). Likewise, it now appears that the deprotonation role assigned to the twin His may not be necessary, and that a combination of more hydrophilic residues in the pore is also functional, as demonstrated by the triple mutants H168D/H318D/I110N and H168D/H318E/I110N (Hall and Kustu, 2011) which also suggest that the charged NH_4^+ is the transported molecule. However, it is also possible that the properties of an individual ammonium transporter are not due to the presence or absence of a particular amino acid, but most likely, are a result of a concerted group of residues that establish singular conditions proper to each member of this family.

ACKNOWLEDGMENTS

The support from CONACyT-México (grants 42664 and 79191) and Dirección General de Asuntos del Personal Académico, Universidad Nacional Autónoma de México (Grant IN218308) is acknowledged.

REFERENCES

Andrade, S. L. A., Dickmanns, A., Ficner, R., and Einsle, O. (2005). Crystal structure of the archaeal ammonium transporter Amt-1 from *Archaeoglobus fulgidus*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14994–14999.

Boeckstaens, M., Bruno, A., and Marini, A. M. (2008). Distinct transport mechanisms in yeast ammonium transport/sensor proteins of the Mep/Amt/Rh family and impact on filamentation. *J. Biol. Chem.* 283, 21362–21370.

Conroy, M. J., Durand, A., Lupo, D., Li, X.-D., Bullough, P. A., Winkler, F. K., and Merrick, M. (2007). The crystal structure of the *Escherichia coli* AmtB–GlnK complex reveals how GlnK regulates the ammonia channel. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1213.

Coutts, G., Thomas, G., Blakey, D., and Merrick, M. (2002). Membrane sequestration of the signal transduction protein GlnK by the ammonium transporter AmtB. *EMBO J.* 21, 536–545.

Couturier, J., Montanini, B., Martin, F., Brun, A., Blaudez, D., and Chalot, M. (2007). The expanded family of ammonium transporters in the perennial poplar plant. *New Phytol.* 174, 137–150.

D'Apuzzo, E., Rogato, A., Simon-Rosin, U., El Alaoui, H., Barbulova, A., Betti, M., Dimou, M., Katinakis, P., Marquez, A., Marini, A. M., Udvardi, M. K., and Chiurazzi, M. (2004). Characterization of three functional high-affinity ammonium transporters in *Lotus japonicus* with differential transcriptional regulation and spatial expression. *Plant Physiol.* 134, 1763–1774.

Fong, R. N., Kim, K.-S., Yoshihara, C., Inwood, W. B., and Kustu, S. (2007). The W148L substitution in the *Escherichia coli* ammonium channel AmtB increases flux and indicates that the substrate is an ion. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18706–18711.

Gazzarrini, S., Lejay, L., Gojon, A., Ninnemann, O., Frommer, W. B., and von Wirén, N. (1999). Three functional transporters for constitutive, diurnally regulated, and starvation-induced uptake of ammonium into *Arabidopsis* roots. *Plant Cell* 11, 937–948.

Guether, M., Neuhäuser, B., Balestrini, R., Dynowski, M., Ludewig, U., and Bonfante, P. (2009). A mycorrhizal-specific ammonium transporter from *Lotus japonicus* acquires nitrogen released by arbuscular mycorrhizal fungi. *Plant Physiol.* 150, 73–83.

Hall, J. A., and Kustu, S. (2011). The pivotal twin histidines and aromatic triad of the *Escherichia coli* ammonium channel AmtB can be replaced. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13270–13274.

Javelle, A., Lupo, D., Ripoche, P., Fulford, T., Merrick, M., and Winkler, F. K. (2008). Substrate binding, deprotonation, and selectivity at the periplasmic entrance of the *Escherichia coli* ammonia channel AmtB. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5040–5045.

Javelle, A., Severi, E., Thornton, J., and Merrick, M. (2004). Ammonium sensing in *Escherichia coli*. Role of the ammonium transporter AmtB

and AmtB-GlnK complex formation. *J. Biol. Chem.* 279, 8530–8538.

Javelle, A., Lupo, D., Zheng, L., Li, X. D., Winkler, F. K., and Merrick, M. (2006). An unusual twin-his arrangement in the pore of ammonia channels is essential for substrate conductance. *J. Biol. Chem.* 281, 39492–39498.

Khademi, S., O'Connell, J. III, Remis, J., Robles-Colmenares, Y., Miercke, L. J. W., and Stroud, R. M. (2004). Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science* 305, 1587–1594.

Lanquar, V., Loqué, D., Hörmann, F., Yuan, L., Bohner, A., Engelsberger, W. R., Lalonde, S., Schulze, W. X., von Wirén, N., and Frommer, W. B. (2009). Feedback inhibition of ammonium uptake by a phospho-dependent allosteric mechanism in *Arabidopsis*. *Plant Cell* 21, 3610–3622.

Loqué, D., Lalonde, S., Looger, L. L., von Wirén, N., and Frommer, W. B. (2007). A cytosolic trans-activation domain essential for ammonium uptake. *Nature* 446, 195–198.

Loqué, D., Mora, S. I., Andrade, S. L. A., Pantoja, O., and Frommer, W. B. (2009). Pore mutations in ammonium transporter AMT1 with increased electrogenic ammonium transport activity. *J. Biol. Chem.* 284, 24988–24995.

Ludewig, U., von Wirén, N., and Frommer, W. B. (2002). Uniport of NH_4^+ by the root hair plasma membrane ammonium transporter *Le*AMT1; 1. *J. Biol. Chem.* 277, 13548–13555.

Ludewig, U., Wilken, S., Wu, B., Jost, W., Obrdlik, P., El Bakkoury, M., Marini, A. M., André, B., Hamacher, T., Boles, E., Von Wirén, N., and Frommer, W. B. (2003). Homo-and hetero-oligomerization of ammonium transporter-1 uniporters. *J. Biol. Chem.* 278, 45603–45610.

Marini, A. M., Boeckstaens, M., Benjelloun, F., Chérif-Zahar, B., and André, B. (2006). Structural involvement in substrate recognition of an essential aspartate residue conserved in Mep/Amt and Rh-type ammonium transporters. *Curr. Gen.* 49, 364–374.

Marini, A. M., Springael, J. Y., Frommer, W. B., and André, B. (2000). Cross-talk between ammonium transporters in yeast and interference by the soybean SAT1 protein. *Mol. Microbiol.* 35, 378–385.

Mayer, M., Dynowski, M., and Ludewig, U. (2006). Ammonium ion transport by the AMT/Rh homologue *Le*AMT1; 1. *Biochemical J.* 396, 431–437.

Neuhäuser, B., Dynowski, M., and Ludewig, U. (2009). Channel-like NH_3 flux by ammonium transporter *At*AMT2. *FEBS Lett.* 583, 2833–2838.

Neuhäuser, B., Dynowski, M., Mayer, M., and Ludewig, U. (2007). Regulation of NH_4^+ transport by essential cross talk between AMT monomers through the carboxyl tails. *Plant Physiol.* 143, 1651–1659.

Ortiz-Ramirez, C., Mora, S. I., Trejo, J., and Pantoja, O. (2011). *Pv*AMT1;1, a highly selective ammonium transporter that functions as H^+/NH_4^+ symporter. *J. Biol. Chem.* 286, 31113–31122.

Shelden, M. C., Dong, B., de Bruxelles, G. L., Trevaskis, B., Whelan, J., Ryan, P. R., Howitt, S. M., and Udvardi, M. K. (2001). *Arabidopsis* ammonium transporters, *At*AMT1; 1 and *At*AMT1; 2, have different biochemical properties and functional roles. *Plant Soil* 231, 151–160.

Simon-Rosin, U., Wood, C., and Udvardi, M. K. (2003). Molecular and cellular characterisation of *Lj*AMT2;1, an ammonium transporter from the model legume *Lotus japonicus*. *Plant Mol. Biol.* 51, 99–108.

Søgaard, R., Alsterfjord, M., Macaulay, N., and Zeuthen, T. (2009). Ammonium ion transport by the AMT/Rh homolog TaAMT1;1 is stimulated by acidic pH. *Pflügers Arch. Eur. J. Physiol.*458, 733–743.

Sohlenkamp, C., Shelden, M., Howitt, S., and Udvardi, M. (2000). Characterization of *Arabidopsis At*AMT2, a novel ammonium transporter in plants. *FEBS Lett.* 467, 273–278.

Sohlenkamp, C., Wood, C. C., Roeb, G. W., and Udvardi, M. K. (2002). Characterization of *Arabidopsis At*AMT2, a high-affinity ammonium transporter of the plasma membrane. *Plant Physiol.* 130, 1788–1796.

Winkler, F. K. (2006). Amt/MEP/Rh proteins conduct ammonia. *Pflügers Arch. Eur. J. Physiol.* 451, 701–707.

Wirén, N. V., and Merrick, M. (2004). Molecular mechanisms controlling transmembrane transport. *Topics in Current Genetics* 1–26.

Wood, C. C., Porée, F., Dreyer, I., Koehler, G. J., and Udvardi, M. K. (2006). Mechanisms of ammonium transport, accumulation, and retention in ooyctes and yeast cells expressing *Arabidopsis At*AMT1; 1. *FEBS Lett.* 580, 3931–3936.

Zheng, L., Kostrewa, D., Bernèche, S., Winkler, F. K., and Li, X.-D. (2004). The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 17090–17095.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 October 2011; accepted: 06 February 2012; published online: 28 March 2012.

*Citation: Pantoja O (2012) High affinity ammonium transporters: molecular mechanism of action. Front. Plant Sci. **3**:34. doi: 10.3389/fpls.2012.00034*

This article was submitted to Frontiers in Plant Physiology, a specialty of Frontiers in Plant Science.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 29 May 2012
doi: 10.3389/fpls.2012.00095

Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels

Alice K. Zelman[1], Adam Dawe[2], Christoph Gehring[3] and Gerald A. Berkowitz[1]*

[1] *Agricultural Biotechnology Laboratory, Department of Plant Science, University of Connecticut, Storrs, CT, USA*
[2] *Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*
[3] *Division of Chemistry, Life Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

Edited by:
Angus S. Murphy, Purdue University, USA

Reviewed by:
Gerald Schoenknecht, Oklahoma State University, USA
Joshua Blakeslee, The Ohio State University, USA

****Correspondence:***
Gerald A. Berkowitz, Agricultural Biotechnology Laboratory, Department of Plant Science, University of Connecticut, 1390 Storrs Road, Storrs, CT 06269-4163, USA.
e-mail: gerald.berkowitz@uconn.edu

Ligand-gated cation channels are a frequent component of signaling cascades in eukaryotes. Eukaryotes contain numerous diverse gene families encoding ion channels, some of which are shared and some of which are unique to particular kingdoms. Among the many different types are cyclic nucleotide-gated channels (CNGCs). CNGCs are cation channels with varying degrees of ion conduction selectivity. They are implicated in numerous signaling pathways and permit diffusion of divalent and monovalent cations, including Ca^{2+} and K^+. CNGCs are present in both plant and animal cells, typically in the plasma membrane; recent studies have also documented their presence in prokaryotes. All eukaryote CNGC polypeptides have a cyclic nucleotide-binding domain and a calmodulin binding domain as well as a six transmembrane/one pore tertiary structure. This review summarizes existing knowledge about the functional domains present in these cation-conducting channels, and considers the evidence indicating that plant and animal CNGCs evolved separately. Additionally, an amino acid motif that is only found in the phosphate binding cassette and hinge regions of plant CNGCs, and is present in all experimentally confirmed CNGCs but no other channels was identified. This CNGC-specific amino acid motif provides an additional diagnostic tool to identify plant CNGCs, and can increase confidence in the annotation of open reading frames in newly sequenced genomes as putative CNGCs. Conversely, the absence of the motif in some plant sequences currently identified as probable CNGCs may suggest that they are misannotated or protein fragments.

Keywords: amino acid motif, CNGC, cyclic nucleotide-binding domain, calmodulin binding domain, Ca^{2+} signaling, ion channel, cyclic nucleotide, channel evolution

Plant cyclic nucleotide-gated cation channels (CNGCs) comprise a large family of non-selective cation-conducting channels in plants. Of the 56 coding sequences identified at present as cation-conducting channels in the *Arabidopsis thaliana* genome, 20 are members of the CNGC family (Ward et al., 2009). CNGCs have been functionally characterized by expression in heterologous systems, or by analysis of cation-related phenotypes of mutant plants (typically *A. thaliana*) that have specific CNGC genes silenced (Talke et al., 2003; Ma et al., 2010). All relevant experimental evidence indicates that plant CNGCs are specifically localized to the plasma membrane, although CNGC20 may be targeted to the chloroplast envelope according to a putative annotation in UniProt[1] (accessed August 28, 2011). Functional analyses of members of this channel family have associated many of them with inward K^+ and Ca^{2+}currents, and at least in several cases, Na^+ conductance (Ma et al., 2010). It may be relevant to their function in plants that in one case, K^+ conductance by a CNGC was restricted in the presence of millimolar concentrations of external Ca^{2+} (Leng et al., 2002). Ca^{2+} block of monovalent cation conductance is central to the function of animal CNGCs (Alam et al., 2007). CNGC function in plant biology may be more related to their ability to conduct Ca^{2+} rather than monovalent cations into plant cells. Abdel-Hamid et al. (2011) provide additional evidence consistent with this conjecture.

Plant CNGCs were first identified as such by the presence of a cyclic nucleotide-binding domain (CNBD), and by their overall structural similarity to animal CNGCs (Talke et al., 2003). CNGCs have also recently been identified in prokaryotes (Nimigean et al., 2004; Kuo et al., 2007), although the prokaryote channel primary sequences show some notable differences in their functional domains (Cukkemane et al., 2011). The most conserved region of the CNGC CNBD is a phosphate binding cassette (PBC) which binds the sugar and phosphate moieties of the cyclic nucleotide (cNMP) ligand (Cukkemane et al., 2011). A "hinge" region adjacent to the PBC is also conserved and is thought to contribute to ligand binding efficacy and selectivity (Young and Krougliak, 2004). The hinge affects cAMP and cGMP activation by modifying the energy requirement of ligand binding and does so independently from other regions (i.e., structural motifs) of the animal CNBD (Young and Krougliak, 2004). Young and Krougliak (2004) modified residues in the hinge, affecting cNMP dependent gating of the channel without causing structural changes in other regions of the CNBD. Additional features of the CNBD of

[1]http://www.uniprot.org/uniprot/Q9LD37

plant CNGCs will be discussed below. Like animal CNGCs, plant CNGCs are members of the "P-loop" superfamily of cation channels present in all prokaryotic and eukaryotic cells (Ward et al., 2009). P-loop channels arose early in evolution and their structure provides a basic form that has evolved to give rise to channels capable of conducting various cations, thereby fulfilling a broad range of functions in cells (Zhorov and Tikhonov, 2004). The core primary structure of a P-loop channel polypeptide (**Figure 1A**) can be represented by two α helices forming membrane-spanning domains (M1 and M2 in **Figure 1A**) surrounding a membrane re-entering pore loop (P-loop; "P" in **Figure 1A**). Alternatively, this core structure can also be formed by a polypeptide with six transmembrane (TM) regions (S1–S6 in **Figure 1B**); in this case the P-loop is present between the fifth (S5) and sixth (S6) transmembrane regions. The amino (N)- and carboxyl (C)-termini are both on the same side of the membrane, and typically cytosolic. The highly conserved P-loop contains a short α helix, a turn, and a random coil (Zhorov and Tikhonov, 2004). It dips into and then out of the membrane from the exterior side of the membrane, forming the ion conducting pathway across the membrane, and includes the amino acids that form an ion selective filter (Ward et al., 2009). The quaternary structure of a P-loop cation conduction pathway of a P-loop channel is formed by four P-loops assembling into an "inverted teepee" structure across the membrane (**Figure 2**). The M1–P-loop–M2 structure of the P-loop polypeptide is present in bacterial cation channels, while the six TM polypeptide is characteristic of many cation channels in eukaryotes (Hua et al., 2003b).

The super families of voltage-gated and ligand-gated channels share this six TM P-loop structure (Mäser et al., 2001). The fact that many channel families share this primary polypeptide structure with common folds is consistent with the concept that the number of ion channel proteins with specific functional properties is substantially larger than the possible folding patterns that could accommodate the generation of an ion-conducting pathway across membranes (Zhorov and Tikhonov, 2004). In animals, voltage-gated Na^+, K^+, and Ca^{2+}-selective channels evolved from

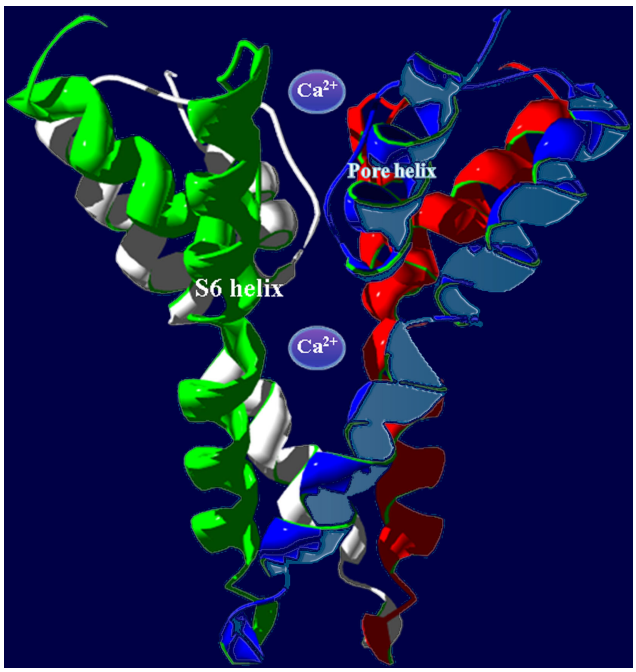


FIGURE 2 | In animals, P-loop cation conduction pathways are known to be formed from four P-loops in an "inverted teepee" structure stretching across the membrane. Shown is an example of a pore structure generated from the corresponding regions of four *A. thaliana* CNGC2 polypeptides illustrating the "inverted teepee" conformation formed by four P-loops arranged symmetrically. Ca^{2+} ions are shown moving through the conduction pathway of the channel.

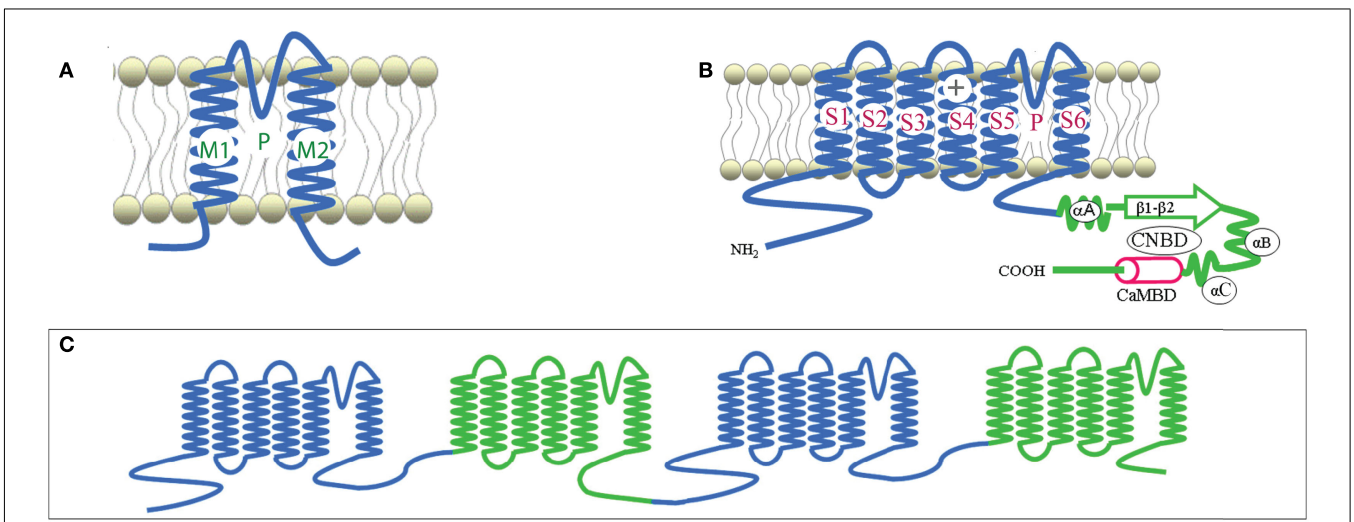


FIGURE 1 | Representative structures of P-loop cation channels. Adapted from Hua et al., 2003a. **(A)** Representation of the core primary loop of a two transmembrane/one pore P-loop channel. **(B)** A six transmembrane/one pore P-loop plant CNGC channel polypeptide. P designates the P-loop of the channel. **(C)** A representation of the secondary structure of an animal Ca^{2+} channel, containing four repeats of the six transmembrane/one pore structure in one large polypeptide.

the rudimentary progenitor P-loop channels. In higher animals, voltage-gated Na^+- and Ca^{2+}-selective channels are formed by one large polypeptide that has four repeat sections; each section contains the six TM P-loop structure (Zhorov and Tikhonov, 2004). The secondary structure of a polypeptide that forms animal Ca^{2+}-selective ion channels is shown in **Figure 1C**. Genes encoding canonical "four repeat P-loop" Ca^{2+}-selective channels present in animal genomes are absent in plants.

Potassium channels are formed by four separate polypeptides. In each case, four P-loops are aligned to form the ion-conducting pathway (Zheng and Zagotta, 2004). **Figure 2** portrays the ion-conducting pathway formed by four P-loops of four different polypeptides (in different colors); also shown is the S6 TM domain of each of the four polypeptides. The configuration of voltage-gated channels is affected by the membrane potential (E_m) such that the pore is physically occluded and therefore closed at some membrane potentials (blocking conductance), and open at other potentials. Evenly spaced, positively charged amino acids in the S4 domain of P-loop channels (represented as a "+" symbol on S4 in **Figure 1B**) act as a voltage sensor and provide a mechanism for the channel to sense and then respond to the E_m by altering the gating of the pore (Hua et al., 2003b). Ligand-gated ion channels, including CNGCs, are also represented in the six TM P-loop superfamily.

Conductance through the pore of channels can also be gated by the binding of ligands to the N- and C-terminal extensions into the cytosol (Biel, 2009). Animal CNGCs are gated allosterically by calmodulin (CaM) and cyclic nucleotides (cAMP and/or cGMP). CaM (in the presence of cytosolic Ca^{2+}) binds to a CaM binding domain (CaMBD) of animal CNGCs at the N-terminus; CaM binding closes the CNGC channel thus preventing ion conductance. Cyclic nucleotides can bind at the C-terminus of both plant and animal CNGCs. CNMP binding to the CNBD of CNGC polypeptides activates the channel, opening it and thereby enabling cation conductance. The aforementioned S4 loop voltage sensor of voltage-gated channels is also present in plant CNGCs.

Plant CNGCs have been demonstrated to function as channels with hyperpolarization-activated (i.e., inwardly rectified), voltage-dependent conductance (Hua et al., 2003b). There are two classes of animal channels whose conductance is regulated by cyclic nucleotides (Biel, 2009). Animal CNGCs are directly activated upon binding of cyclic nucleotides and their open probability does not change at varying E_m. Hyperpolarization-activated cyclic nucleotide-gated channels (HCNs) conduct ions (i.e., are open) at hyperpolarizing E_m, and binding of cyclic nucleotides increases the percentage of open channels at a given hyperpolarizing E_m. Thus, conductance of these channels changes at varying E_m as well as in response to changes in the level of cyclic nucleotides in the cytosol. In plants CNGCs may be incorrectly classified as solely "ligand-gated" channels without response to voltage when in fact they may function like animal HCNs.

As indicated above, the quaternary structure of a member of the P-loop channel family is dependent on the assembly of four gene products in some cases (e.g., in animal and plant K^+-selective voltage-gated channels) where only one P-loop "cassette" is encoded by the gene (i.e. S1–S6 with the P-loop between S5 and S6). It is therefore conceivable that plant CNGC channel complexes are formed by the assembly of four polypeptides, with each polypeptide corresponding to a structure similar to that shown in **Figure 1B**. Threading regions (S6 and pore loop, an example is shown in **Figure 3** below) of plant CNGC coding sequences through the crystal structure of a quaternary P-loop channel indicates that the plant CNGC polypeptide may be capable of forming the tetrameric structure common to P-loop channels (Hua et al., 2003b). However, this model has not been verified experimentally. It should be noted that functional animal CNGC channel proteins are in all cases generated from such a tetrameric assembly of P-loop cassette polypeptides (Zheng and Zagotta, 2004) and

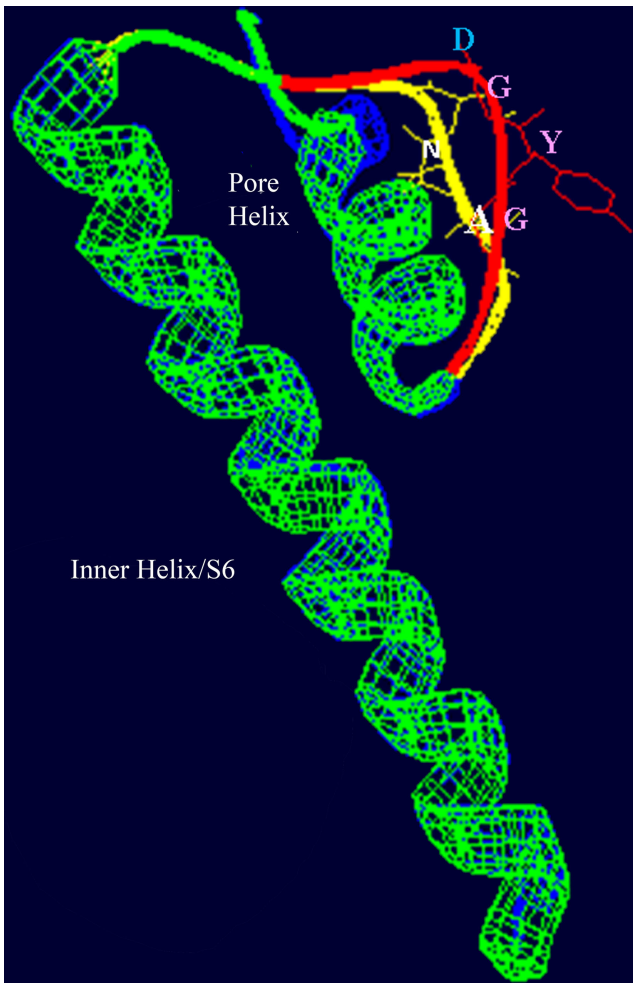


FIGURE 3 | Superimposed images (redrawn from Leng et al., 2002) of the S6 helix and pore domains of the plant (*A. thaliana*) CNGC2 sequence and the bacterial (*Streptomyces lividans*) K^+-selective channel KcsA (Doyle et al., 1998; PDB record 1BL8A). Threading of the *A. thaliana* CNGC2 S6 transmembrane helix and pore region through the KcsA structure identifies the residues "AND" as in the ion conducting pathway where the residues "GYG" (shown in pink) that form the ion selectivity filter of a K^+-selective channel are positioned. This suggests that plant CNGCs may have a mechanism for cation conductance that is conserved in P-loop channels. In plants, four separate CNGC peptides are presumed to form a tetrameric structure that shares the "inverted teepee" quaternary structure of animal P-loop channels.

FIGURE 4 | Predicted three-dimensional structures of the plant CNGC CNBD. (A) A proposed three-dimensional structure of the *A. thaliana* CNGC2 CNBD generated by threading the sequence through the structure of the cAMP-binding domain of PDB record 1RGS, bovine cAMP-dependent protein kinase A (RIα; Su et al., 1995). **(B)** Three-dimensional structure of the CNBD of RIα that was used to create the model of the CNGC2 CNBD. The CNBD structures of CNGC2 and RIα share an N-terminal α-helix (αA) and a β-barrel (depicted as ribbons with arrows denoting antiparallel strands) that together bind the cNMP ligand. Republished from Hua et al. (2003a). **(C)** The tertiary structure of the *A. thaliana* CNGC6 CNBD (Chikayama et al., 2004; PDB record 1WGP). Amino acid residues comprising the hinge domain are highlighted in yellow; α-helices are depicted as purple cylinders. The β-barrel is shown as straight antiparallel purple ribbons. Two cGMP molecules are shown with the structure, one in the pocket formed between an αA helix and the β-barrel, and a second cGMP molecule (not bound) is depicted at the α-C helix adjacent to the hinge domain.

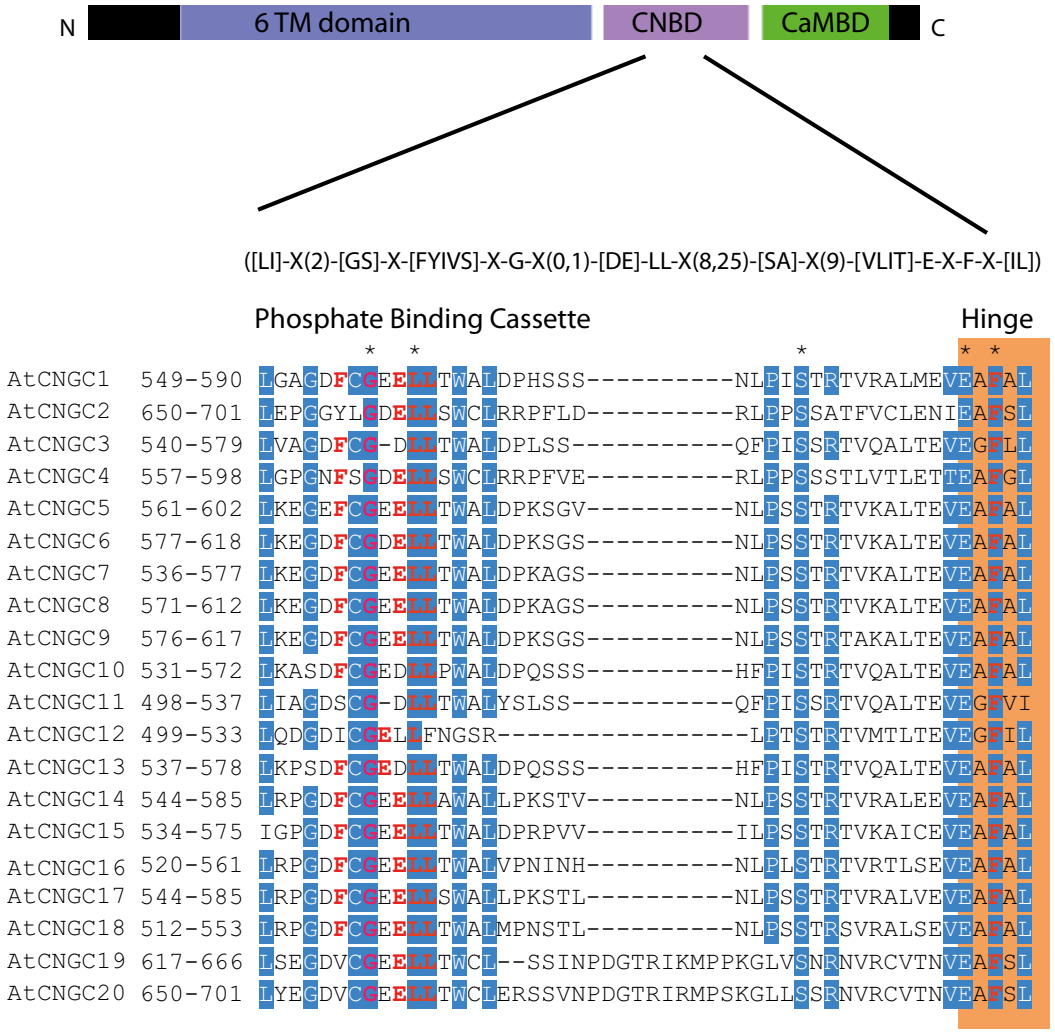


FIGURE 5 | The *A. thaliana* CNGC-specific motif spans the putative PBC and the hinge within the CNBD of the 20 CNGCs. The diagram at the top portrays three regions of plant CNGCs: the six transmembrane domains (TM), a CNBD containing a PBC and the hinge which presumably make direct contact with the cyclic nucleotide, and a calmodulin binding domain (CaMBD) toward the C-terminus. The CNGC-specific amino acid motif is shown below the cartoon. In the square brackets "[]" are the amino acids allowed in this position of the motif, "X" represents any amino acid and the round brackets "()" indicate the number of amino acids. The red amino acids in the PBC and the hinge are conserved in both animal and plant CNBDs based on comparison with Jackson et al. (2007). Below the CNGC-specific motif is an alignment of the motif regions of the 20 *A. thaliana* CNGCs. For each of the *A. thaliana* CNGCs, the residues at the N- and C-termini are indicated to the left of the motif. The green shaded box demarcates the putative PBC region of the CNBDs (corresponding to the animal CNBD PBC). The orange shaded box marks the presumed hinge region. Residues in white highlighted in blue indicate >90% identity among the 20 *A. thaliana* CNGCs (in this case "N" is counted as a dissimilarity) indicating a high level of conservation. A "∗" above the alignment marks a position with 100% conservation between the *A. thaliana* sequences shown. Residues in red denote conservation with the animal CNBD alignment generated by Jackson et al. (2007). Note, there is experimental evidence identifying the residues which interact with cyclic nucleotides in animal CNBDs according to Jackson et al., but there is no empirical confirmation of which residues in plant CNBDs bind cyclic nucleotides. The alignment of the *A. thaliana* CNGC-specific motif was generated by the MEGA5 program and ClustalW (Saitou and Nei, 1987).

that plant and animal CNGC polypeptides share a similar general S1–S6 topography. Thus, the aforementioned conjecture about the tetrameric structure of plant CNGCs is entirely consistent with the general concept of ion channel assembly and function.

The amino acids that line the pore selectivity filter of P-loop channels determine the relative conductance for Na^+, K^+, and Ca^{2+}. This region of a plant CNGC (*A. thaliana*, CNGC2, isoform 2) is shown superimposed over the corresponding region of a K^+-selective channel in **Figure 3**. The triplet "GYG" forms the selectivity filter of the K^+ channel. The corresponding residues in CNGC2 are "AND," with the negatively charged aspartic acid (D) residue positioned at the outer mouth of the pore. No other plant CNGC has a selectivity filter with an acidic D or glutamic acid (E) at this position in the pore. Animal voltage-gated Ca^{2+}-selective channels have an "EEEE" ring at the extracellular end of the ion conducting pathway formed by the corresponding residues in each of the four P-loop repeats that form the functional channel. This E ring of Ca^{2+} channels has been experimentally verified

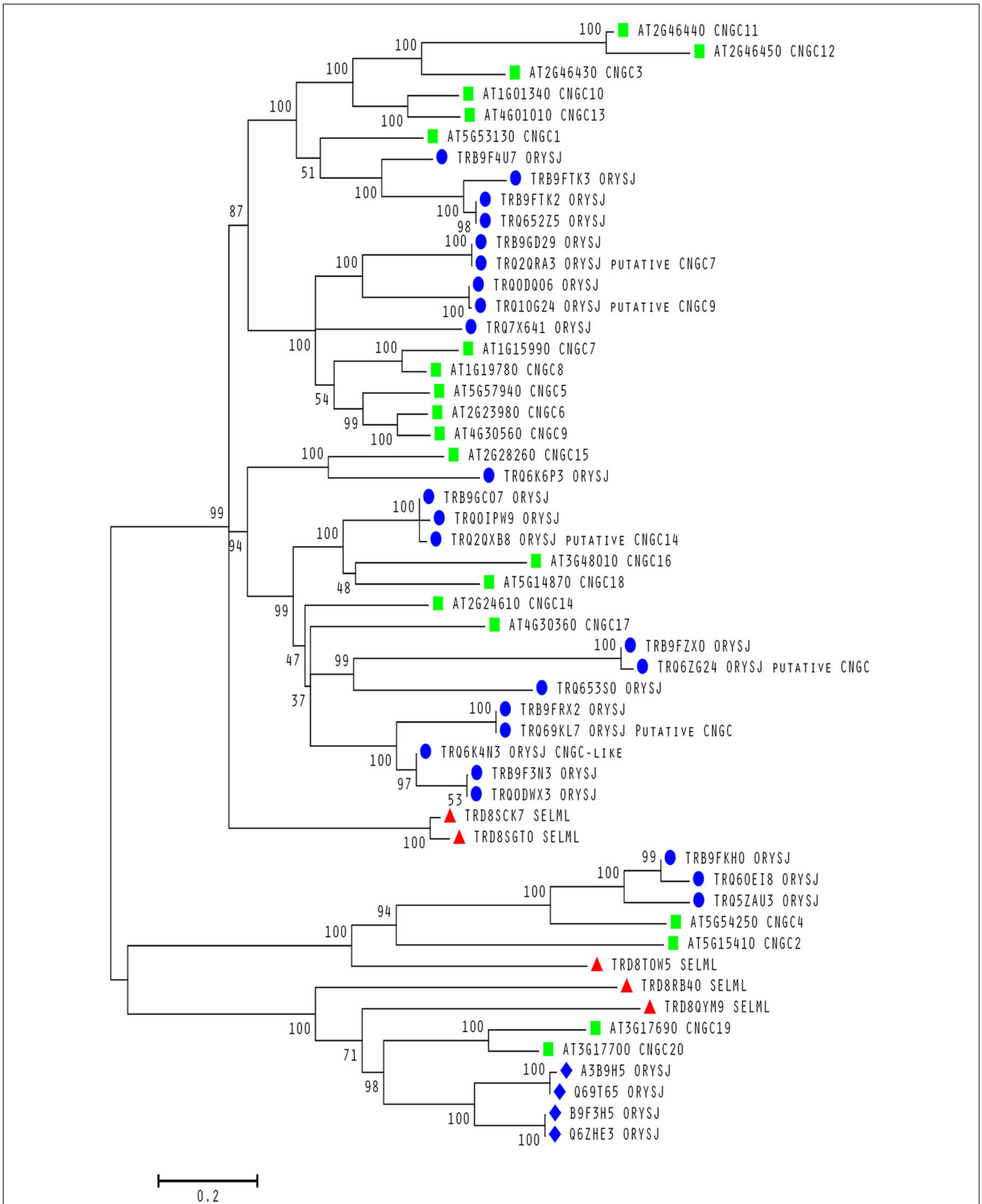


FIGURE 6 | *(Continued)*

FIGURE 6 | Continued
Evolutionary relationships within a conserved motif in plant CNGCs. Molecular phylogenetic analysis using the Maximum Likelihood method for 53 cyclic nucleotide-gated channel (CNGC) amino acid sequences from *A. thaliana* (green squares), *S. moellendorffii* (red triangles) and *O. sativa* var. *japonica* (blue circles and diamonds). All sequences contain a conserved S between the two motif subsequences except for the rice sequences labeled with a diamond. All *A. thaliana* sequences are from confirmed CNGCs while the remaining sequences from the other two species are from uncharacterized proteins unless stated otherwise as putative CNGCs. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The sequences used to create the tree were obtained using BLASTP (Camacho et al., 2009) with an *E*-value <0.01 against the UniProt Knowledgebase (The UniProt Consortium, 2012) for Viridiplantae from the European Bioinformatics Institute website (http://www.ebi.ac.uk/Tools/sss/ncbiblast/) for all 20 annotated CNGCs from *A. thaliana*. All sequences from *S. moellendorffii*, *O. sativa* var. *japonica*, and *A. thaliana* were then extracted from the hits with exact duplicates (same identifiers, same sequence and different identifiers, same sequence) as well as annotated fragments being removed. The remaining sequences containing the motif ([LI]-X(2)-[GS]-X-[FYIVS]-X-G-X(0,1)-[DE] -LL-X(8,25)-[SA]-X(9)-[VLIT]-E-X-F-X-[IL]) were then aligned using MUSCLE (Edgar, 2004) and inspected for isoforms (near 100% identity) and unannotated fragments. The bootstrap consensus tree was generated using the JTT matrix-based model with discrete gamma distribution (five categories (+ G, parameter = 1.0524); Jones et al., 1992) by MEGA5 (Tamura et al., 2011) from 1000 bootstraps (Felsenstein, 1985). It is drawn to scale with branch lengths measured in the number of substitutions per site and all ambiguous positions removed for each sequence pair. It is based on 1082 positions in the final data set. Branches corresponding to less than 50% bootstrap replicates are collapsed and the percentage of replicate trees in which the associated taxa clustered together are shown next to the branches. The BioNJ method with the MCL distance matrix was used to generate the initial trees except where the common sites were less than 100 or a quarter of the total sites, in which case, maximum parsimony was used.

to coordinate divalent Ca^{2+} ions as they enter the pore; and cause Ca^{2+} selectivity in animal channels (Tikhonov and Zhorov, 2011). The selectivity filter of plant CNGCs does not appear to be conserved among isoforms within this protein family, and is different from the conserved pore of animal CNGCs. However, some plant CNGCs have an acidic E two residues away from the D of CNGC2 (the D residue can be seen close to the outer mouth of the channel pore in **Figure 3**). The D in CNGC2, or alternatively the two E residues close to the outer mouth of the pore may conceivably provide a basis for Ca^{2+} conductance by some of the plant CNGCs. In *A. thaliana*, CNGCs 2, 4, 11, 12, 5, 6, and 9 have the E at the outer mouth of the pore and CNGCs 1, 2, 10, 11, 12, and 18 have been experimentally demonstrated (either directly or indirectly) to be involved in Ca^{2+} conductance (Leng et al., 1999; Ali et al., 2006; Frietsch et al., 2007; Urquhart et al., 2007; Guo et al., 2010). The speculative functional assignment of this acidic residue of the plant CNGC selectivity filter awaits further experimental confirmation.

A defining structural aspect of plant CNGCs is their CNBD, which is unique for a number of reasons. The presence of CNGC coding sequences with CNBDs in animal and plant is intriguing from an evolutionary perspective. Unicellular fungi lack channels with such CNBDs (Talke et al., 2003). Considering that plants branched from animal progenitors on the evolutionary tree before the animal–fungus division, CNGCs might have evolved independently in these two lineages. Alternatively, fungi might have lost CNGC channel genes. Some aspects of the structural features of the plant CNBD correspond to that of their bacterial and animal cyclic-nucleotide-binding analogs. The CNBD of plant CNGCs overlaps with a CaMBD near the C-terminus, and CaM (in the presence of Ca^{2+}) prevents cNMP activation of plant CNGCs (Hua et al., 2003a). However there is no corresponding CaMBD in bacterial CNGC sequences (Cukkemane et al., 2011). In animals the functional CaMBD is located distal of the CNBD, near the N-terminus (Ungerer et al., 2011). As mentioned above, CaM binds to and regulates the conductance of animal CNGCs and this is also the case in plant CNGCs. However, in plant CNGCs, the amino acids that form the CaM binding domain overlap with the region of the polypeptide that forms the CNBD (**Figure 1B**). The predicted three-dimensional structure of a plant CNGC CNBD (**Figure 4A**) has been generated by threading this portion of the polypeptide through the crystal structure of a bovine CNBD shown in **Figure 4B** (Hua et al., 2003a). Similar structures can be found in the literature (Chikayama et al., 2004; Bridges et al., 2005; Kaplan et al., 2007). The Protein Data Bank contains a solved structure for the calmodulin binding domain (CaMBD) of a small conductance Ca^{2+}-activated potassium channel (SK2) in *Rattus norvegicus* (Wissmann et al., 2002; PDB record 1KKD). Threading the plant CNGC CaMBD sequences through the rat CaMBD structure may shed light on the possible convergent evolution of animal and plant CaMBDs present in ion channels.

It should be noted that plants contain two known domains, GAF and CNBD, which bind cyclic nucleotides. The GAF domain evolved billions of years ago and was first identified as a structural component of some cGMP phosphodiesterases and some light-sensing proteins in bacteria, yeast, humans, *A. thaliana*, and several species of cyanobacteria (Aravind and Ponting, 1997). Besides cAMP and cGMP, GAF domains can bind a number of other small molecules and are not specific to cNMPs (Bridges et al., 2005). The CNBD found in all plant CNGCs is also present in members of three plant K^+-selective channel families; the "Shaker-like" (so named for the shaking phenotype of some animals with null mutations) KAT and AKT channel families, and the outwardly conducting channels SKOR and GORK. However, these K^+-selective channels are affected by cyclic nucleotides in a very different manner than CNGCs. The voltage threshold for activation is shifted to more negative potentials in the presence of cGMP (experimentally verified for isoforms of the KAT and AKT families (Hoshi, 1995; Gaymard et al., 1996); the outward rectifiers have not been studied). This effectively means that cNMP elevation in the cytosol would reduce conductance through plant K^+ channels that have CNBDs. Electrophysiological analysis of plant CNGCs has shown that the same elevation of cNMP activates the channel thereby increasing conductance (Leng et al., 1999, 2002; Lemtiri-Chlieh and Berkowitz, 2004; Ali et al., 2007).

To date, no CNGC-specific motifs have been reported, however, Jackson et al. (2007) aligned different animal CNGCs, and in particular their PBC and hinge regions from which the following consensus motif can be derived: FGE-[IT]-[CIA]-LL-X(3,4)-[RK]-R-X-A-SV-X(11)-[SH]-[VRA]-[FY]-[HNQ]-X-[LV]-[LA] (the animal CNGC hinge sequence spans from the conserved serine (S) to

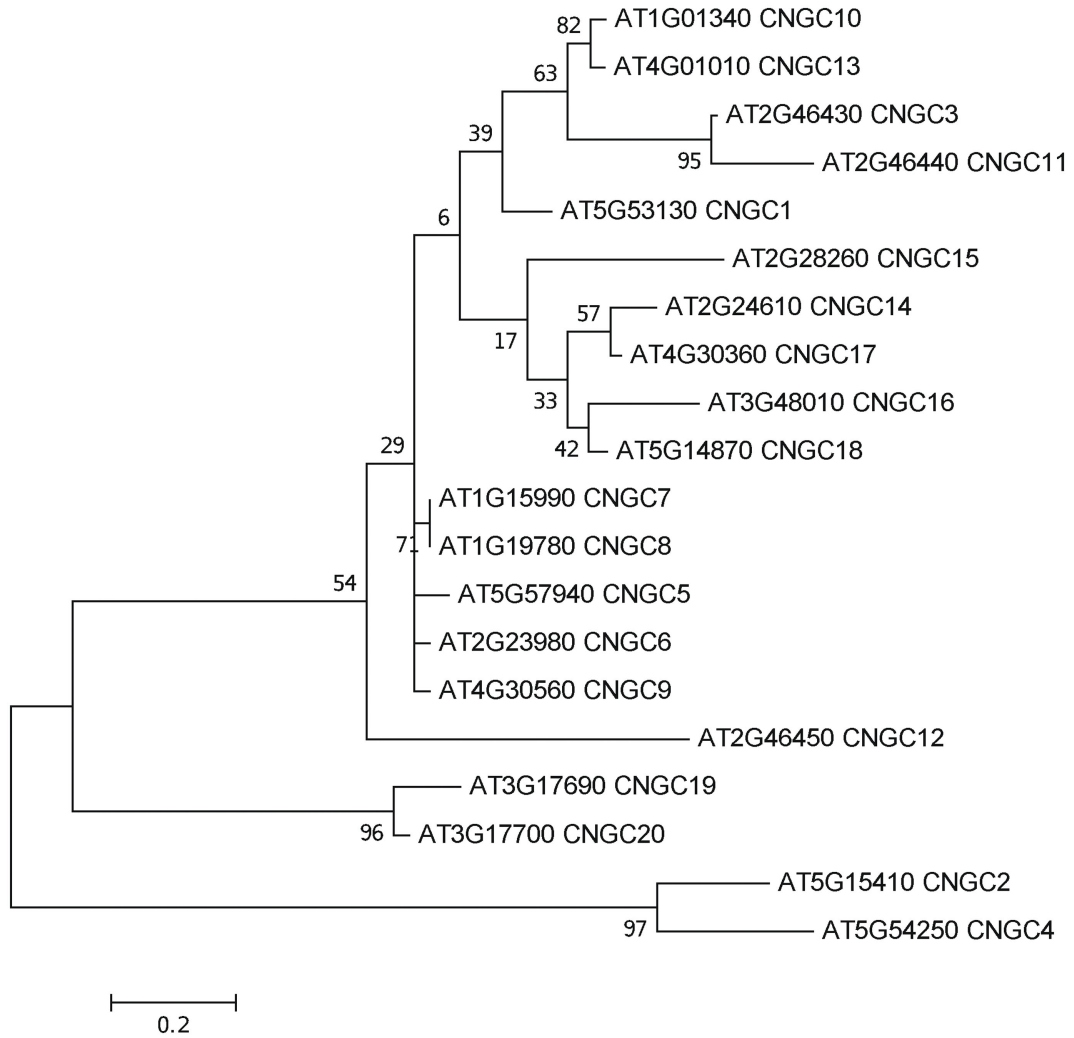


FIGURE 7 | Molecular phylogenetic history of the CNGC-specific conserved motif ([LI]-X(2)-[GS]-X-[FYIVS]-X-G-X(0,1)-[DE]-LL-X(8,25)-[SA]-X(9)-[VLIT]-E-X-F-X-[IL]) in *A. thaliana*. The analysis used the Maximum Likelihood method based on the JTT matrix-based model with a discrete gamma distribution [five categories (+G, parameter = 1.0852); Jones et al., 1992] in MEGA5 (Tamura et al., 2011) with 1000 bootstraps (Felsenstein, 1985). The initial heuristic trees were obtained automatically using the BioNJ method with MCL distance, unless the number of common sites was less than 100 or one quarter of the total number of sites, in which case the maximum parsimony method was used. The tree with the highest log likelihood (−771.1381) is shown, drawn to scale with branch lengths measured in the number of substitutions per site. All ambiguous positions were removed for each sequence pair with a total of 52 positions in the final dataset of 20 sequences.

the C-terminus of the motif). We noted that such a motif does not occur in plants but hypothesized that some of the functionally critical residues might be conserved. We therefore aligned *A. thaliana* CNGCs and identified putative PBCs and hinges. Within the putative PBCs we identified a conserved phenylalanine (F), a stabilizing glycine (G) and an acidic residue (either D or E) followed by two aliphatic leucines (L). We also observed that the putative hinge also contains an E, F, and one aliphatic residue much like in the animal CNGCs. The plant CNGC hinge occurs in between the CNBD and CaMBD regions (**Figure 4C**). We subsequently built a stringent motif ([LI]-X(2)-[GS]-X-[FYIVS]-X-G-X(0,1)-[DE]-LL-X(8,25)-[SA]-X(9)-[VLIT]-E-X-F-X-[IL]) that recognizes 20 *A. thaliana* CNGC proteins and no other sequences in *A. thaliana*. This subsequence includes the hinge domain and PBC (**Figure 5**). This conserved sequence differs from the animal CNBD; it occurs in between the CNBD and CaMBD while in animals the hinge occurs within the CNBD itself. Additionally it lacks, for example, the conserved proline (P) that was shown to affect gating in animal CNGCs (Jackson et al., 2007). Possible functional similarities may derive from the correspondence between the F in the center of the hinge in *A. thaliana* CNGCs and the F/Y in a bovine and a catfish CNGC (Young and Krougliak, 2004). The animal PBC region differs from the *A. thaliana* PBC conserved residues as well (Young and Krougliak, 2004). A scan of UniProt (Swiss-Prot/TrEMBL, The UniProt Consortium, 2012) including splice variants and excluding fragments using the ScanProSite tool[2] (Sigrist et al., 2010) further revealed that the motif is restricted

[2]http://prosite.expasy.org/scanprosite/

to land plants. Notably, the three predicted algal CNGC sequences from *Chlamydomonas reinhardtii* (Verret et al., 2010) lacked the motif. CNGC-related sequences matched against the 20 *A. thaliana* CNGCs using BLASTP (*E*-value <0.01) in UniProt for rice (*Oryza sativa* var. *japonica*) and the pin-cushion spikemoss (*Selaginella moellendorffii)* were then extracted, along with all hits for *A. thaliana*. Duplicates were removed and the remaining 53 sequences checked for the motif, before being used to construct a cladogram (**Figure 6**). The genes included in this cladogram are listed in **Table A1** in Appendix. In this cladogram it can be seen that the majority of the rice and *A. thaliana* sequences tend to group together intraspecifically, suggesting numerous gene duplication events. The sequences across all three species partition into two clusters: a larger, more diverse cluster containing the bulk of the CNGC gene family, and a smaller cluster containing CNGC2, 4, 19, and 20. Of particular interest is that the only sequences to contain a conserved alanine (A) in place of the more frequent S in the motif are all from rice (denoted by blue diamonds) and all group together closely with CNGC19 and 20. The cladogram shown in **Figure 7** is based upon only the conserved CNGC-specific motif. This phylogeny has striking congruence to previously published cladograms generated from aligning the full-length *A. thaliana* CNGC gene sequences (Mäser et al., 2001; Talke et al., 2003; Ward et al., 2009), as well as alignments of the pore domain amino acid sequences and the CNDB domain amino acid sequences (Kaplan et al., 2007). This supports the idea that the motif is informative for identifying and comparing putative CNGCs in other plants. **Figure 7** also suggests that mutations in the motif have occurred concurrently with the branching of the CNGC gene family.

Interestingly, the motif recognizes all currently annotated CNGCs in higher plants only and identifies new unannotated candidate CNGCs. Two CNGC isoforms in *A. thaliana* [Swiss-Prot: Q94AS9-2 (an isoform of CNGC4) and Q8GWD2-2 (an isoform of CNG12)] lack the motif. However, these sequences are not experimentally confirmed and are truncated. The few annotated CNGCs in other plants (e.g., barley and poplar) that lack the motif – even if we relax it by allowing one or several mismatches – may suggest that they are either misannotated, or fragments that show insufficient coverage and similarity to be confidently classified as CNGCs. Furthermore, the motif recognizes neither any non-plant sequences nor any algal sequences. We therefore contend that the motif may serve as a supporting method for identifying CNGCs in other land plant species.

Scanning the rice and pin-cushion spikemoss genomes with the ScanProsite tool yields putative CNGC homologs and uncharacterized proteins containing the motif (**Figure 6**). Since CNBDs are found in other proteins besides the CNGC ion channels in the *A. thaliana* proteome (as noted above, i.e., the K^+-selective channels), it may be useful to develop motifs that can distinguish these sequences. Several recent publications predict the number of CNGCs present in the genomes of organisms ranging from brown and red algae to green algae and multicellular plants. For example, Ward et al. (2009) predict that no CNGCs exist in the green alga *C. reinhardtii*. In contrast, Verret et al. (2010) predict that *C. reinhardtii* has three CNGCs. These three sequences do not contain the land plant-specific CNGC motif. The fact that these authors disagree on the number of CNGCs attests to the need for alternative ways to assess the likelihood that a protein for which the primary sequence is known is likely to function as a CNGC. The motif we present here does not recognize any non-plant sequences or any algal sequences, and thus lends weight to the view that the *C. reinhardtii* proteins identified as putative CNGCs by Verret et al. (2010) may contain algal-specific CNGC motifs and we argue that higher plants may have evolved a specific cNMP-binding domain.

A functional motif-based means of recognizing CNGC proteins will also contribute to our understanding the evolutionary history of this gene family including the possible role of genome and gene duplication events (Seoighe and Gehring, 2004). Analyzing orthologous CNGCs in the new Whole Genome Shotgun sequence of *Arabidopsis lyrata*, a cress whose genome is approximately 50% larger than that of *A. thaliana,* is the next obvious step in learning more about the evolutionary effects of ion channel gene duplication and subsequent sub-functionalization and changing expression patterns. These phenomena have already been partly addressed (Mäser et al., 2001), and it was remarked that CNGC12 has a CaMBD that was the least conserved compared to the other CNGCs this group (Group I). CNGC12 was the last gene in a sequence of three syntenic CNGCs: CNGC3 and CNG11 precede it. CNGC12 is likely the least homologous indicating that gene duplication allowed the original function of the Group I gene CNGC3 to be maintained, easing selection pressure on the duplicate copies and perhaps allowing for sub-functionalization.

CONCLUDING REMARKS

The pore and CNBD sequences of the plant cyclic nucleotide-gated channels differ from the pore and CNBDs in other plant ion channel families. They are also different from the pore and CNBD regions of animal CNGCs. Phylogenies based on alignments of the pore, CNBD, or full-length CNGC sequences are similar, showing that the evolution of these functional domains preceded the expansion of plant CNGC genes that occurred after the split between green algae and higher plants, and that the split between dicots and monocots occurred after the advent of the higher plant CNGCs. A protein motif specific to plant CNGCs is located within the CNBD. An alignment of just this motif region from the CNGCs produces a cladogram that is almost identical to published phylogenies of the full-length sequences. This motif may therefore be used to refine tentative annotations and to find novel CNGCs in newly sequenced plant genomes. Whether this motif represents differences in the function or regulation of plant CNGCs has yet to be evaluated.

ACKNOWLEDGMENTS

The authors thank Nadia Burger and Nathan Wojtyna for assistance in preparing images. This work is funded by a grant from the National Science Foundation (Award 1146827).

REFERENCES

Abdel-Hamid, H., Chin, K., Moeder, W., and Yoshioka, K. (2011). High throughput chemical screening supports the involvement of Ca^{2+} in cyclic nucleotide-gated ion channel-mediated programmed cell death in *Arabidopsis*. *Plant Signal. Behav.* 11, 1817–1819.

Alam, A., Shi, N., and Jiang, Y. (2007). Structural insight into Ca^{2+} specificity in tetrameric cation channels. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15334–15339.

Ali, R., Ma, W., Lemtiri-Chlieh, F., Tsaltas, D., Leng, Q., von Bodman, S., and Berkowitz, G. A. (2007). Death don't have no mercy and neither does calcium: *Arabidopsis* cyclic nucleotide gated channel 2 and innate immunity. *Plant Cell* 19, 1081–1095.

Ali, R., Zielinski, R., and Berkowitz, G. A. (2006). Expression of plant cyclic nucleotide-gated cation channels in yeast. *J. Exp. Bot.* 57, 125–138.

Aravind, L., and Ponting, C. P. (1997). The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem. Sci.* 22, 458–459.

Biel, M. (2009). Cyclic nucleotide-regulated cation channels. *J. Biol. Chem.* 284, 9017–9021.

Bridges, D., Fraser, M. E., and Moorhead, G. B. G. (2005). Cyclic nucleotide binding proteins in the *Arabidopsis thaliana* and *Oryza sativa* genomes. *BMC Bioinformatics* 6, 6–18. doi:10.1186/1471-2105-6-6

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421–430. doi:10.1186/1471-2105-10-421

Chikayama, E., Nameki, N., Kigawa, T., Koshiba, S., Inoue, M., Tomizawa, T., Kobayashi, N., and Yokoyama, S. (2004). Solution structure of the cNMP-binding domain from *Arabidopsis thaliana* cyclic nucleotide-regulated ion channel. RIKEN Structural Genomics/Proteomics Initiative. *PDB ID:1WGP.* doi:10.2210/pdb1wgp/pdb

Cukkemane, A., Seifert, R., and Kaupp, U. B. (2011). Cooperative and uncooperative cyclic-nucleotide-gated ion channels. *Trends Biochem. Sci.* 36, 55–64.

Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* 280, 69–77.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113–132. doi:10.1186/1471-2105-5-113

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.

Frietsch, S., Wang, Y. F., Sladek, C., Poulsen, L. R., Romanowsky, S. M., Schroeder, J. I., and Harper, J. F. (2007). A cyclic nucleotide-gated channel is essential for polarized tip growth of pollen. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14531–14536.

Gaymard, F., Cerutti, M., Horeau, C., Lemaillet, G., Urbach, S., Ravallec, M., Devauchelle, G., Sentenac, H., and Thibaud, J. B. (1996). The baculovirus/insect cell system as an alternative to *Xenopus oocytes*. First characterization of the AKT1 K+ channel from *Arabidopsis thaliana*. *J. Biol. Chem.* 271, 22863–22870.

Guo, K. M., Babourina, O., Christopher, D. A., Borsic, T., and Rengel, Z. (2010). The cyclic nucleotide-gated channel AtCNGC10 transports Ca^{2+} and Mg^{2+} in *Arabidopsis*. *Physiol. Plant* 139, 303–312.

Hoshi, T. (1995). Regulation of voltage dependence of the KAT1 channel by intracellular factors. *J. Gen. Physiol.* 105, 309–328.

Hua, B.-G., Mercier, R. W., Zielinski, R. E., and Berkowitz, G. A. (2003a). Functional interaction of calmodulin with a plant cyclic nucleotide gated cation channel. *Plant Physiol. Biochem.* 41, 945–954.

Hua, B.-G., Mercier, R. W., Leng, Q., and Berkowitz, G. A. (2003b). Plants do it differently: a new basis for potassium/sodium selectivity in the pore of an ion channel. *Plant Physiol.* 132, 1353–1361.

Jackson, H. A., Marshall, C. R., and Accili, E. A. (2007). Evolution and structural diversification of hyperpolarization-activated cyclic nucleotide-gated channel genes. *Physiol. Genomics* 29, 231–245.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282.

Kaplan, B., Sherman, T., and Fromm, H. (2007). Cyclic nucleotide-gated channels in plants. *FEBS Lett.* 581, 2237–2246.

Kuo, M. M., Saimi, Y., Kung, C., and Choe, S. (2007). Patch clamp and phenotypic analyses of a prokaryotic cyclic nucleotide-gated K^+ channel using *Escherichia coli* as a host. *J. Biol. Chem.* 282, 24294–24301.

Lemtiri-Chlieh, F., and Berkowitz, G. A. (2004). Cyclic adenosine monophosphate regulates calcium channels in the plasma membrane of *Arabidopsis* leaf guard and mesophyll cells. *J. Biol. Chem.* 279, 35306–35312.

Leng, Q., Mercier, R. W., Hua, B.-G., Fromm, H., and Berkowitz, G. A. (2002). Electrophysiological analysis of cloned cyclic nucleotide gated ion channels. *Plant Physiol.* 128, 400–410.

Leng, Q., Mercier, R. W., Yao, W. Z., and Berkowitz, G. A. (1999). Cloning and first functional characterization of a plant cyclic nucleotide-gated cation channel. *Plant Physiol.* 121, 753–761.

Ma, W., Yoshioka, K., Gehring, C. A., and Berkowitz, G. A. (2010). "The function of cyclic nucleotide gated channels in biotic stress," in *Ion Channels and Plant Stress Responses*, eds V. Demidchik and F. J. M. Maathuis (Berlin: Springer-Verlag), 159–174.

Mäser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J., Sanders, D., Harper, J. F., Tchieu, J., Gribskov, M., Persans, M. W., Salt, D. E., Kim, S. A., and Guerinot, M. L. (2001). Phylogenetic relationships within cation transporter families of *Arabidopsis*. *Plant Physiol.* 126, 1646–1667.

Nimigean, C. M., Shane, T., and Miller, C. (2004). A cyclic nucleotide modulated prokaryotic k^+ channel. *J. Gen. Physiol.* 124, 203–210.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Seoighe, C., and Gehring, C. (2004). Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends Genet.* 20, 461–464.

Sigrist, C. J. A., Cerutti, L., de Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., and Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res.* 38, 161–166.

Su, Y., Dostmann, W. R., Herberg, F. W., Durick, K., Xuong, N. H., Ten Eyck, L., Taylor, S. S., and Varughese, K. I. (1995). Regulatory subunit of protein kinase a: structure of deletion mutant with cAMP binding domains. *Science* 269, 807–813.

Talke, I. N., Blaudez, D., Maathuis, F. J. M., and Sanders, D. (2003). CNGCs: prime targets of plant cyclic nucleotide signalling? *Trends Plant Sci.* 8, 286–293.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.

The UniProt Consortium. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 40, D71–D75.

Tikhonov, D. B., and Zhorov, B. S. (2011). Possible roles of exceptionally conserved residues around the selectivity filters of sodium and calcium channels. *J. Biol. Chem.* 286, 2998–3006.

Ungerer, N., Mücke, N., Broecker, J., Keller, S., Frings, S., and Möhrlen, F. (2011). Distinct binding properties distinguish LQ-type calmodulin-binding domains in cyclic nucleotide-gated channels. *Biochem.* 50, 3221–3228.

Urquhart, W., Gunawardena, A. H., Moeder, W., Ali, R., Berkowitz, G. A., and Yoshioka, K. (2007). The chimeric cyclic nucleotide-gated ion channel ATCNGC11/12 constitutively induces programmed cell death in a Ca^{2+} dependent manner. *Plant Mol. Biol.* 65, 747–761.

Verret, F., Wheeler, G., Taylor, A. R., Farnham, G., and Brownlee, C. (2010). Calcium channels in photosynthetic eukaryotes: implications for evolution of calcium-based signalling. *New Phytol.* 187, 23–43.

Ward, J. M., Mäser, P., and Schroeder, J. I. (2009). Plant ion channels: gene families, physiology, and functional genomics analyses. *Annu. Rev. Physiol.* 71, 59–82.

Wissmann, R., Bildl, W., Neumann, H., Rivard, A. F., Klöcker, N., Weitz, D., Schulte, U., Adelman, J. P., Bentrop, D., and Fakler, B. (2002). A helical region in the C terminus of small-conductance Ca^{2+}-activated K^+ channels controls

assembly with apo-calmodulin. *J. Biol. Chem.* 277, 4558–4564.

Young, E. C., and Krougliak, N. (2004). Distinct structural determinants of efficacy and sensitivity in the ligand-binding domain of cyclic nucleotide-gated channels. *J. Biol. Chem.* 279, 3553–3562.

Zheng, J., and Zagotta, W. N. (2004). Stoichiometry and assembly of olfactory cyclic nucleotide-gated channels. *Neuron* 42, 411–421.

Zhorov, B. S., and Tikhonov, D. B. (2004). Potassium, sodium, calcium and glutamate-gated channels: pore architecture and ligand action. *J. Neurochem.* 88, 782–799.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 September 2011; accepted: 24 April 2012; published online: 29 May 2012.

Citation: Zelman AK, Dawe A, Gehring C and Berkowitz GA (2012) Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels. Front. Plant Sci. ***3****:95. doi: 10.3389/fpls.2012.00095*

This article was submitted to Frontiers in Plant Traffic and Transport, a specialty of Frontiers in Plant Science.

APPENDIX

Table A1 | Accessions for sequences used to construct Figures 6 and 7.

Key	Database	Species	Accession	Description/Gene Name
AT1G01340 CNGC10	TAIR	*A. thaliana*	AT1G01340	CNGC10
AT1G15990 CNGC7	TAIR	*A. thaliana*	AT1G15990	CNGC7
AT1G19780 CNGC8	TAIR	*A. thaliana*	AT1G19780	CNGC8
AT2G23980 CNGC6	TAIR	*A. thaliana*	AT2G23980	CNGC6
AT2G24610 CNGC14	TAIR	*A. thaliana*	AT2G24610	CNGC14
AT2G28260 CNGC15	TAIR	*A. thaliana*	AT2G28260	CNGC15
AT2G46430 CNGC3	TAIR	*A. thaliana*	AT2G46430	CNGC3
AT2G46440 CNGC11	TAIR	*A. thaliana*	AT2G46440	CNGC11
AT2G46450 CNGC12	TAIR	*A. thaliana*	AT2G46450	CNGC12
AT3G17690 CNGC19	TAIR	*A. thaliana*	AT3G17690	CNGC19
AT3G17700 CNGC20	TAIR	*A. thaliana*	AT3G17700	CNGC20
AT3G48010 CNGC16	TAIR	*A. thaliana*	AT3G48010	CNGC16
AT4G01010 CNGC13	TAIR	*A. thaliana*	AT4G01010	CNGC13
AT4G30360 CNGC17	TAIR	*A. thaliana*	AT4G30360	CNGC17
AT4G30560 CNGC9	TAIR	*A. thaliana*	AT4G30560	CNGC9
AT5G14870 CNGC18	TAIR	*A. thaliana*	AT5G14870	CNGC18
AT5G15410 CNGC2	TAIR	*A. thaliana*	AT5G15410	CNGC2
AT5G53130 CNGC1	TAIR	*A. thaliana*	AT5G53130	CNGC1
AT5G54250 CNGC4	TAIR	*A. thaliana*	AT5G54250	CNGC4
AT5G57940 CNGC5	TAIR	*A. thaliana*	AT5G57940	CNGC5
TRB9F3N3 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9F3N3_ORYSJ	Putative uncharacterized protein
TRB9F4U7 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9F4U7_ORYSJ	Putative uncharacterized protein
TRB9FKH0 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9FKH0_ORYSJ	Putative uncharacterized protein
TRB9FRX2 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9FRX2_ORYSJ	Putative uncharacterized protein
TRB9FTK2 ORYSJ	UniProt	*O. sativa* var. *japonica*	TRB9FTK2_ORYSJ	Putative uncharacterized protein
TRB9FTK3 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9FTK3_ORYSJ	Putative uncharacterized protein
TRB9FZX0 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9FZX0_ORYSJ	Putative uncharacterized protein
TRB9GC07 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9GC07_ORYSJ	Putative uncharacterized protein
TRB9GD29 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9GD29_ORYSJ	Putative uncharacterized protein
TRQ0DQ06 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q0DQ06_ORYSJ	Os03g0646300 protein (CNGC7-like)
TRQ0DWX3 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q0DWX3_ORYSJ	Os02g0789100 protein (CNGC17-like)
TRQ0IPW9 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q0IPW9_ORYSJ	Os12g0163000 protein (CNGC18-like)
TRQ10G24 ORYSJ putative CNGC9	UniProt	*O. sativa* var. *japonica*	Q10G24_ORYSJ	Putative CNGC9
TRQ2QRA3 ORYSJ putative CNGC7	UniProt	*O. sativa* var. *japonica*	Q2QRA3_ORYSJ	Putative CNGC7
TRQ2QXB8 ORYSJ putative CNGC14	UniProt	*O. sativa* var. *japonica*	Q2QXB8_ORYSJ	Putative CNGC14
TRQ5ZAU3 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q5ZAU3_ORYSJ	Os01g0782700 protein (CNGC4-like)
TRQ60EI8 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q60EI8_ORYSJ	Os05g0502000 protein (CNGC4-like)
TRQ652Z5 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q652Z5_ORYSJ	Os06g0527100 protein (CNGC1-like)
TRQ653S0 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q653S0_ORYSJ	Os09g0558300 protein (ATCNGC14-like)
TRQ69KL7 ORYSJ Putative CNGC	UniProt	*O. sativa* var. *japonica*	Q69KL7_ORYSJ	Putative CNGC
TRQ6K4N3 ORYSJ CNGC-like	UniProt	*O. sativa* var. *japonica*	Q6K4N3_ORYSJ	CNGC-like
TRQ6K6P3 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q6K6P3_ORYSJ	Os02g0627700 protein (CNGC15-like)
TRQ6ZG24 ORYSJ putative CNGC	UniProt	*O. sativa* var. *japonica*	Q6ZG24_ORYSJ	putative CNGC
TRQ7X641 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q7X641_ORYSJ	OSJNBa0033G05.7 protein (CNGC-like)
TRA3B9H5 ORYSJ	UniProt	*O. sativa* var. *japonica*	A3B9H5_ORYSJ	Putative uncharacterized protein
TRQ69T65 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q69T65_ORYSJ	Putative cyclic nucleotide-regulated ion channel
TRB9F3H5 ORYSJ	UniProt	*O. sativa* var. *japonica*	B9F3H5_ORYSJ	Putative uncharacterized protein
TRQ6ZHE3 ORYSJ	UniProt	*O. sativa* var. *japonica*	Q6ZHE3_ORYSJ	Putative cyclic nucleotide-binding transporter 1
TRD8QYM9 SELML	UniProt	*S. moellendorffii*	D8QYM9_SELML	Putative uncharacterized protein

(Continued)

Table A1 | Continued

Key	Database	Species	Accession	Description/Gene Name
TRD8RB40 SELML	UniProt	*S. moellendorffii*	D8RB40_SELML	Putative uncharacterized protein
TRD8SCK7 SELML	UniProt	*S. moellendorffii*	D8SCK7_SELML	Putative uncharacterized protein
TRD8SGT0 SELML	UniProt	*S. moellendorffii*	D8SGT0_SELML	Putative uncharacterized protein
TRD8T0W5 SELML	UniProt	*S. moellendorffii*	D8T0W5_SELML	Putative uncharacterized protein

Database accessions and gene names were derived from UniProt (S. moellendorffii and O. sativa var. japonica; The UniProt Consortium, 2012) and The Arabidopsis Information Resource (TAIR; A. thaliana; Swarbreck et al., 2008).

REFERENCE

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., Li, D., Meyer, T., Muller, R., Ploetz, L., Radenbaugh, A., Singh, S., Swing, V., Tissier, C., Zhang, P., and Huala, E. (2008). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res.* 36, D1009–D1014.