Genes controlling the development and function of plant vacuoles

Li, Y.

Publication date
2017

Document Version
Other version

License
Other

Citation for published version (APA):
Li, Y. (2017). Genes controlling the development and function of plant vacuoles. [Thesis, fully internal, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3

Evolution of tonoplast P-ATPase transporters involved in vacuolar acidification*

Yanbang Li1,2,#, Sofia Provenzano1,3,#, Mattijs Blick1,2,#, Cornelis Spelt1,2, Ingo Appelhagen4, Laura Machado de Faria2, Walter Verweij2, Andrea Schubert3, Martin Sagasser4, Thorsten Seidel5, Bernd Weisshaar4, Ronald Koes1,2,* and Francesca Quattrocchio1,2,*.

1Swammerdam Institute for Life Sciences, Dept. of Plant Development and (Epi)Genetics, University of Amsterdam, 1098XH Amsterdam, The Netherlands.
2Department of Molecular and Cell Biology, VU-University, De Boelelaan 1081, 1071HK Amsterdam, The Netherlands.
3Department of Agricultural, Forestry and Food Sciences, University of Turin, 10095 Grugliasco, Italy.
4Genome Research, Faculty of Biology, Bielefeld University, 33615 Bielefeld, Germany.
5Dynamic Cell Imaging, Faculty of Biology, Bielefeld University, 33501 Bielefeld, Germany.
#Y.L., S.P., and M.B. equally contributed to this work.

*Published in New Phytologist 211, 1092–1107 (2016).
Summary

- Petunia mutants with blue flowers defined a novel vacuolar proton pump consisting of two interacting P-ATPases, PH1 and PH5, that hyper-acidify vacuoles of petal cells. PH5 is similar to plasma membrane H^+ P_{3A}-ATPase, whereas PH1 is the only known eukaryotic P_{3B}-ATPase. Since there were no indications that this tonoplast pump is widespread in plants, we investigated the distribution and evolution of PH1 and PH5.
- We combined database mining, phylogenetic and synteny analyses of PH1- and PH5-like proteins from all kingdoms, with functional analyses (mutant complementation, and intracellular localization) of homologs from diverse Angiosperms.
- We identified functional PH1 and PH5 homologs in divergent Angiosperms. PH5 homologs evolved from plasma membrane P_{3A}-ATPases acquiring a N-terminal tonoplast-sorting sequence and new cellular function before angiosperms appeared. PH1 is widespread among seed plants and related proteins are found in some groups of bacteria and fungi and in one moss, but is absent in most algae suggesting that its evolution involved several cases of gene loss and possibly horizontal transfer events.
- PH1 and PH5 distribution in the plant kingdom suggests that vacuolar acidification by P-ATPases appeared in Gymnosperms before flowers. This implies that, next to flower color determination, vacuolar hyper-acidification is required for yet unknown processes.
Chapter 3

Introduction

The vacuole is used for storage of proteins, sugars, ions, and a variety of secondary metabolites, including anthocyanin pigments (Eisenach et al., 2015). The transport of metabolites across the vacuolar membrane (tonoplast) and other endomembranes is in many cases energized by an electrochemical H\(^+\)-gradient that is generated by proton pumps. In most cells the V-ATPase, together with pyrophosphatase (PPase) proton pumps (Drozdowicz & Rea, 2001), mildly acidifies the vacuole and other cellular compartments (Gaxiola et al., 2007; Schumacher & Krebs, 2010; Shen et al., 2013). However, in certain cell-types, like lemon fruit juice cells, the vacuole can be very acidic (pH 2) (Muller & Taiz, 2002).

In petunia, vacuoles of petal epidermal cells are also hyper-acidified as compared to leaf cells, which confers a red-violet color to the anthocyanin pigments in the vacuole (Faraco et al., 2014). Mutations in seven distinct loci, named PH1 to PH7, reduce vacuolar acidification in these cells, resulting in a bluish flower color and reduced acidity of petal homogenates (de Vlaming et al., 1983; Verweij et al., 2008; Faraco et al., 2014). PH3 and PH4 encode MYB and WRKY type proteins, which interact with the bHLH protein ANTHOCYANIN1 (AN1) and the WD40 protein AN11 to activate the transcription of downstream genes including PH1 and PH5 (Spelt et al., 2002; Quattrocchio et al., 2006). PH1 and PH5 encode P-ATPase ion pumps that are essential and sufficient for vacuolar hyper-acidification (Verweij et al., 2008; Faraco et al., 2014).

PH5 belongs to the P\(_{3A}\)-ATPase family of proton pumps found in fungi and plants (Verweij et al., 2008). Yeast and primitive plants contain two or three P\(_{3A}\)-ATPases, while in higher plants 10-15 proteins have been identified, divided in subgroups I to V (Arango et al., 2003; Baxter et al., 2003; Pedersen et al., 2012). P\(_{3A}\)-ATPases of groups I, II and IV reside in the plasma membrane (DeWitt & Sussman, 1995; DeWitt et al., 1996; Kim et al., 2001; Lefebvre et al., 2004; Lefebvre et al., 2005) to energize various transporters and to regulate cytoplasmic pH (Palmgren, 2001). However, PH5 of petunia, which is similar to group III proteins, resides in the tonoplast to acidify the vacuolar lumen (Verweij et al., 2008; Faraco et al., 2014).

PH1 belongs to the distinct subfamily of P\(_{3B}\)-ATPases (Faraco et al., 2014), which were identified as facilitators of Mg\(^{2+}\) import in bacteria (Smith & Maguire, 1998), and long thought to be absent from plants, fungi and animals (Kuhlbrandt, 2004; Thever & Saier, 2009; Pedersen et al., 2012). When expressed in leaf cells PH1 has no electrogenic activity on its own, but can bind to PH5 and boost its H\(^+\)-pumping activity (Faraco et al., 2014). Ectopic co-expression of PH1 and PH5 lowers the pH in leaves and, interestingly, suppresses the activity of the V-ATPase providing direct evidence that the complex of these two P-ATPases alone can acidify vacuoles to a greater extend than V-ATPases.

The identification of the PH1/PH5 tonoplast pump broke with the long-held view that P\(_{3A}\)-ATPases all reside in the plasma membrane, and that P\(_{3B}\)-ATPases are lacking in plants (Thever & Saier, 2009; Pedersen et al., 2012; Schumacher, 2014). The universality of these findings remained, however,
Evolution of tonoplast P-ATPase transporters

ambiguous. Even though breeders and geneticists collected over centuries in numerous (ornamental) species, flower color mutants with defects in (conserved) anthocyanin genes, \( ph \) mutants were only described in petunia, and decades of (electro)physiological studies did not uncover a vacuolar P-ATPase pump in other species either. This raises the question whether the PH1/PH5 vacuolar hyper-acidification pathway is an eccentric feature that recently evolved only in a few species.

In this study we investigated the distribution of PH1 and PH5 homologs in all kingdoms of life. We show that homologs of PH1 and PH5 are widespread among Angiosperms and appeared already in Gymnosperms. Phylogenetic and functional analyses indicate that PH5 originated by gene duplication from plasma membrane \( P_{3\alpha} \)-ATPases about 300MYA and that PH1 evolved from bacterial \( Mg^{2+} \) transporters, apparently losing its capability to bind ions about at the same time point. This indicates that these two pumps, and the process of vacuolar hyper-acidification that they control, play a role in mechanisms unrelated to color display.

Results

PH5 homologs from different species

We identified proteins from rose (\( R. \) hybrida) and carnation (\( D. \) caryophyllus) with high similarity to PH5 by cDNA library screening and PCR amplification respectively and retrieved similar sequences of other species from various (genome) databases. Phylogenetic analysis showed that these proteins, including those from the gymnosperms \( Picea \) and \( Pinus \), constitute a monophyletic clade distinct from the type I and II plasma membrane \( P_{3\alpha} \)-ATPases (Fig. 1a). However, in a few angiosperms, including the nightshades tomato, potato (\( S. \) lycopersicum and \( S. \) tuberosum) and pepper (\( C. \) annuum), several grasses (\( S. \) bicolor, \( Z. \) mays and \( P. \) virgatum) and cucumber (\( C. \) sativus), the protein with the highest similarity to PH5 grouped with type I or type II \( P_{3\alpha} \)-ATPases, indicating that a true PH5 ortholog is lacking in these species. More primitive plants, such as the lycophyte \( S. \) moellendorfii and the bryophytes \( P. \) patens and \( S. \) fallax also lack a clear PH5 homolog, suggesting that the gene duplication giving rise type III \( P_{3\alpha} \)-ATPases took place in early seed plants. The type III clade of \( P_{3\alpha} \)-ATPases contains proteins all sharing high homology with the petunia PH5 (and the Arabidopsis AHA10).

\( P_{3\alpha} \)-ATPases display high similarity throughout the protein sequence (Fig. S1a), except for the N- and C- terminal regions. These are conserved among PH5-like proteins, but diverged from the type I and type II proteins (Fig. 2a-b). Members of the PH5 clade share similar intron/exon architecture. Eudicot \( PH5 \) homologs contain 20 introns at identical positions (Fig. S1b), while the monocot genes (e.g. the rice \( O.sAHA9 \) and \( S.titalica \) \( S.t034109m.g \)) lack intron 13 (Fig. S1b), whereas \( P_{3\alpha} \)-ATPase genes from different subclades contain distinct subsets of the 20 introns found in eudicot \( PH5 \) homologs (Arango et al., 2003; Baxter et al., 2003), consistent with a monophyletic origin of angiosperm \( PH5 \) homologs.
Fig. 1 PH5 homologs in various species. (a) Phylogenetic analysis of selected P3A-ATPase proteins. For each species the protein(s) with most similarity to AtAHA2 (type II) and PhPH5 (type III) are shown. In addition we added some type I proteins (like AtAHA4). Note that for some species the protein with most similarity to PH5 belongs to a different P3A-ATPase clade (marked by red dots), indicating that a true PH5 ortholog is missing in that species. Proteins used for localization studies are indicated in green and those that complemented the ph5 mutation in petunia with asterisks. Branch support is calculated on the basis of 500 bootstraps. The protein alignment from which this tree was generated is reported in Note S1. (b) Complementation of the petunia ph5 mutant with PH5 homologs from different species. The bars beside each flower denote the average pH of the crude petal extract (n ≥ 5, ±SD). (c) Confocal images showing the localization of GFP-tagged AHA2 and PH5 homologs (green signal) in petunia petal epidermis protoplasts. RFP-AtSYP122 marks the plasma membrane (red signal). The insets (white boxes) show a detail at higher magnification. AtAHA2-GFP and PhAHA2-GFP colocalize with RFP-AtSYP122, while the PhPH5 homologs result in green fluorescence on the tonoplast (as better visible in the inserts). Bar, 10 µm.
To assess if and when type III P$_{3A}$-ATPase became functionally different from the type I and II plasma membrane proteins, we examined PH5 homologs from divergent angiosperms. We have previously shown that the expression of the petunia PH5 protein, when expressed from the 35S promoter, is able to fully complement the ph5 (Verweij et al., 2008). We therefore constitutively expressed carnation DcPH5, AHA10, grape VvPH5 and the GFP fusion RhPH5-GFP in a petunia ph5 mutant. We produced 20-25 transgenic plants for each construct and in those with high transgene expression (Fig. S1c) the flower color and pH of the crude petal extracts were restored (Fig. 1b). However, ph5 transformants that expressed AtAHA2 at a similar level (Fig. S1c) all retained a mutant (ph5) phenotype (Fig. 1b).

To examine whether the (in)capability of these P$_{3A}$-ATPases to replace PH5 related to differences in their subcellular localization we transiently expressed GFP-fusions in petunia petal protoplasts. Since RhPH5-GFP corrects the ph5 phenotype (Fig. 1b) without being cleaved (Fig. S1d), similar to PH5-GFP (Verweij et al., 2008), these fusions faithfully reflect the sorting of the native proteins. PH5-GFP was visible on the tonoplast 48 hours after transformation (Fig. 1c), consistent with previous results (Verweij et al., 2008). AtAHA10-GFP, VvPH5-GFP, DcPH5-GFP and RhPH5-GFP showed a similar localization, whereas AtAHA2-GFP and the petunia homolog PhAHA2-GFP both localized in the plasma membrane together with the plasma membrane marker RFP-AtSYP122 (Fig. 1c).

Together these results show that PH5 homologs evolved from plasma membrane P$_{3A}$-ATPases prior to the appearance of seed plants, estimated at some 300 million years ago (MYA) (Palmer et al., 2004), and acquired a new cellular localization and function in vacuolar acidification at least 180 MYA, prior to the appearance of angiosperms, and presumably even earlier, before gymnosperms (see Discussion).

**Functional divergence of tonoplast and plasma membrane P$_{3A}$-ATPases**

To identify the molecular changes by which PH5 homologs diverged from plasma membrane P$_{3A}$-ATPases and evolved into tonoplast pumps, we examined the localization of GFP-tagged chimeras of PH5 and PhAHA2. We focused on the N- and C-terminal cytoplasmic domains (Arango et al., 2003; Pedersen et al., 2007), because these differ the most between PH5 homologs and the plasma membrane P$_{3A}$-ATPases of subfamilies I and II (Fig. 2a, b). PhAHA2-GFP marked the circumference of transiently transformed epidermal petal protoplasts and co-localized with the plasma membrane marker RFP-AtSYP122, indicating that PhAHA2 resides in the plasma membrane, like AtAHA2 (Fig. 1c, 3a). Chimeras in which either the N- and C-terminal domains of PhAHA2 (PhAHA2$^{N=PH5; C=PH5}$), or the 25-amino acid N-terminal domain alone (PhAHA2$^{N=PH5}$), were exchanged with the corresponding PH5 domain(s), marked the circumference of the central vacuole (Fig. 3b, c). Exchange of the C-terminus only (PhAHA2$^{C=PH5}$) by contrast, left the co-localization with RFP-SYP122 in the plasma membrane unaffected (Fig. 3d). Replacement of both the N- and C-terminus of PH5 (PH5$^{N=PhAHA2; C=PhAHA2}$) or the N-terminal domain alone (PH5$^{N=PhAHA2}$) with the corresponding domains of PhAHA2, (re-)directed PH5 to the plasma membrane (Fig. 3f, g) instead of the vacuolar membrane (Fig. 3e), whereas replacement
of the C-terminus alone did not change the tonoplast localization of PH5 (Fig. 3b). These data indicate that changes in the sequence of the N-terminal domain of a plasma membrane P3A-ATPase during evolution were sufficient to acquire tonoplast localization.

The N-terminal domain of AHA10 shows an insertion of eight amino acids missing in AtAHA2 (Fig. 2a and 4a). This insertion (EDLDKPLL in AHA10, and EDLERPLL in PH5, Fig. 2a) resembles a dileucin motif identified in several vacuolar Arabidopsis proteins and lysosomal proteins from animals (Pedrazzini et al., 2013). The presence of this motif is, however, not always accompanied by vacuolar localization. We fused the ten amino-terminal amino acids from AHA10 to the plasma membrane protein AtAHA2 and found that the AHA101-10-AtAHA2 fusion was in tobacco BY protoplasts sorted to the tonoplast (Fig. 4b). Next, we generated AHA10 mutants in which this 8 amino acid-motif was deleted (AHA10 D1-10-GFP) or replaced in part (AHA108-10>Ala-GFP) or completely (AHA103-10>Ala-GFP) with alanine residues (Fig. 4a). In petunia petal protoplasts all three AHA10-GFP mutants no longer sorted to the tonoplast and instead accumulated mostly in the plasma membrane (Fig. 4c), while some fluorescence was seen in punctae localized between both membranes (Fig. 4d, e). The latter could be due to slow sorting of the mutant AHA10-GFP protein or to miss-folding of some of it and subsequent aggregation or retention in the ER or vesicles.

Fig. 2 Comparison of N- and C-termini of P3A-ATPases. (a) Alignment of the N-terminal domain of several P3A-ATPases. The red asterisks indicate the amino acids that have been mutated to produce the AHA10 constructs in Fig.4. In red are the proteins used in this study for subcellular localization analysis. (b) Alignment of the C-terminal portion of several plant P3A-ATPases. The red asterisks mark the conserved threonine residue (T) that is phosphorylated in AtAHA2 homologs to enable interaction with 14-3-3 proteins and the tyrosine (Y) that is missing in the proteins of the PH5-clade. In red are the proteins used for the Y2H analysis of the interaction with 14-3-3 proteins. A list of the species is available in the Tables S1 and S2.
Evolution of tonoplast P-ATPase transporters

Fig. 3 Functional divergence of N- and C-terminal regions between PhAHA2 (Petunia hybrida H’ ATPase 2) and PH5. (a–h) Confocal images of petunia petal protoplasts transiently expressing the plasma membrane marker Red Fluorescent Protein (RFP)-AtSYP122 (SYP: Syntaxin of Plants 122) (left panels) and green fluorescent protein (GFP) fusions (middle panels) of (a) PhAHA2, (e) PH5 and chimeras in which the N- and/or C-terminal cytoplasmic domains are exchanged, as indicated in the diagrams (b–d, f–h). Bars, 10 µm. Numbers on the protein diagrams denote the amino acid residues at the points of fusion.

Taken together, these observations suggest that alterations in the N-terminal sequence of a plasma membrane P₃A-ATPase changed the PH5 ancestor from plasma membrane into a tonoplast protein. Because PH5 homologs from species other than petunia and Arabidopsis have similar cytoplasmic N-terminal domains (Fig. 2a), it is likely that this characteristic was acquired early in the diversification of the PH5 ancestor from the other members of the P₃A-ATPase family (Fig. 2a).

Plasma membrane-based P₃A-ATPases contain an (auto)inhibitory C-terminal domain, which is regulated by phosphorylation and interaction with 14-3-3 proteins. The C-terminus of type I and II proteins consists of a conserved YTV motif (Fig. 2b). Phosphorylation of the penultimate threonine (T) enables binding to 14-3-3 proteins and the formation of an active hexameric P-ATPase complex (Jahn et al., 1997; Fuglsang et al., 1999; Svennelid et al., 1999; Maudoux et al., 2000; Kanczewska et al., 2005; Ottmann et al., 2007). In PH5 homologs this motif is replaced by a sequence, - HTV in eudicots or QTV in the monocots *Oryza sativa* (rice) and *Setaria italic*, that lacks the antepenultimate tyrosine residue. Interestingly, the PH5 homologs of the gymnosperms *Picea abies* and *Pinus taeda* contain the
antepenultimate tyrosine (AYTV, Fig. 2b), supporting the origin of PH5 from other P₃A-ATPases and indicating that the C-termini of PH5-like proteins differentiated after the separation of gymnosperms and angiosperms.

**Fig. 4** Identification of vacuolar sorting signal in AtAHA10. (a) Diagram of different *AtAHA10* mutants in which amino acid residues in the N-terminus were deleted or replaced by alanine residues, as indicated in red. (b) Localization in tobacco BY2 protoplasts of the chimera consisting of the first ten amino acids of AHA10 fused to AHA2. RFP-AtSYLP122 marks the plasma membrane. (c) Confocal micrographs of the localization of the Arabidopsis AHA2 PM proton pump. AHA10 and the AHA10 mutants in petunia petal protoplasts. RFP-AtSYLP122 marks the plasma membrane. (b) Localization in tobacco BY2 protoplasts of the chimera consisting of the first ten amino acids of AHA10 fused to AHA2. RFP-AtSYLP122 marks the plasma membrane. (d) Petunia petal protoplast expressing the AHA10₈₉₋₁₀>A₁₀ mutant. (e) Scan of the fluorescence localization in the cell in (d). The green fluorescent signal from AHA10₈₉₋₁₀>A₁₀-GFP is present on the plasma membrane (together with the red signal of RFP-AtSYP122) and in punctae in the cytoplasm. The autofluorescence of the anthocyanins (shown here in blue) marks the lumen of the vacuole. (f) Yeast two-hybrid analysis of the interactions of P₃A-ATPases with 14-3-3 proteins. Interactions are seen as growth on plates lacking histidine (LTH) and lacZ reporter activity (blue stain). The petunia AAA-ATPase protein was used as a positive control for the interaction with the C-terminal region of PH5. PhBEH2 (BES1HOMOLOG2), a member of the BES1/BZR1 (Bri1-EMS Suppressor /BRASSI-NAZOLE-RESISTANT 1) family of transcription factors, was shown previously to interact with all the 14-3-3 proteins used in this assay (Verhoeve *et al.*, 2013) and served here as a positive control.
Evolution of tonoplast P-ATPase transporters

Deletion of the YTV motif from AtAHA2 abolishes 14-3-3 binding and replacement of any of the three ultimate amino acids (Y, T or V) with alanine severely reduces activity (Fuglsang et al., 1999). In PH5, however, deletion of this motif does not affect activity (Verweij et al., 2008). We used a yeast two-hybrid (Y2H) assay, similar to those used to study the interaction of 14-3-3 proteins with type I and II P₃A-ATPases (Jahn et al., 1997; Fuglsang et al., 1999), and found that the C-terminal domain of AtAHA2 and PhAHA2 can interact with 14-3-3 proteins, in particular isoforms chi and psi, whereas the same domain of PhPH5 cannot (Fig. 4f). By contrast, the PH5 C-terminal domain can interact, albeit weakly, with a putative AAA-ATPase (ATPases Associated with diverse cellular Activities) (Iyer et al., 2004) that we identified in a Y2H cDNA library screen, whereas AtAHA2 and PhAH2 cannot (Fig. 4f). These results suggest that the divergence of the C-terminal motif altered the interactions with other proteins involved in post-translational regulation.

Distribution of PH1 homologs

P₃B-ATPases, such as MgtA from E. coli and Salmonella, were discovered for their involvement in magnesium uptake and were long thought to be specific for prokaryotes (Smith & Maguire, 1998; Kuhlbrandt, 2004). In early database searches we uncovered a putative grape homolog (VvPH1) (Faraco et al., 2014), and by PCR amplification a similar gene in rose (RhPH1). The VvPH1 and RhPH1 proteins display high (70%) amino acid similarity with PhPH1, and the genes display conservation of the intron/exon structure (Fig. S2a). Although P₃B-ATPases are thought to be most related to P₃A-ATPases (Axelsen & Palmgren, 1998), the intron/exon organization of the P₃B and P₃A-ATPase genes is entirely different.

We transformed p35S:PhPH1, p35S:VvPH1 and p35S:RhPH1 in ph1 petunia plants (Faraco et al., 2014) and obtained about 20 transgenics for each construct. Those expressing the transgene (Fig. S1c) showed a fully complemented petal phenotype (red-violet color and a low pH of the crude extract) while the untransformed controls have blue-violet petals and higher pH values (Faraco et al., 2014) (Fig. 5a). These results indicate that PH1 is functionally conserved in asterids (petunia) and distantly related rosids (grape and rose).

To assess the distribution of P₃B-ATPases in eukaryotes and retrace their evolutionary origin, we analyzed both predicted protein sequences and translated genome sequences. We found P₃B-ATPases, putative PH1 homologs, in a broad range of angiosperms, including numerous eudicot species and three monocots (date palm Phoenix dactylifera, duckweed Spirodela polyrhiza and banana Musa acuminata), a bryophyte (peatmoss, Sphagnum fallax), an algae (Chlorella variabilis) (Fig. 5b and S3). Outside the plant kingdom, we found PH1/MgtA-like proteins in some groups of bacteria (Firmicutes and Proteobacteria) and archaea (in the phyla Euryarchaeota and Crenarchaeota), as expected, but also in several slime molds (Dictyostelium species and Polysphondylium pallidum) and in a few classes of fungi (Fig. 5b and S3). In all these species the PH1 homolog appears to be a single gene. However, the genomes of all metazoans and most of the prokaryotes and fungi, including all true yeasts, lack PH1 homologs.
Fig. 5 PH1 homologs in different species. (a) Phenotype of ph1 mutant petunia flowers complemented with transgenes expressing PH1 homologs from different species. The bar next to the flowers represents the average pH of the crude petal extract (N≥5, ±SD). (b) Phylogenetic tree of proteins from distinct species with highest similarity to PH1. The proteins functionally tested by complementation of the ph1 petunia mutant (a) are marked by an asterisk. The proteins in the strongly supported clade on the top are the proteins with most similarity to PH1 from species (apparently) lacking a true PH1 homolog and belong to different P-ATPase families of P-ATPases. The branch support was calculated on the basis of 500 bootstraps. The protein alignment from which this tree was generated is reported in Note S2. A larger version of this tree encompassing more species is shown in Fig. S3. (c) Synteny analysis of the genomic regions containing PH1 homologs. The arrows denote genes (and their orientation) flanking PH1 homologs in different species. Identical number and color indicate similar genes. Syntenic regions from Arabidopsis thaliana, Capsella rubella (Cr), Solanum lycopersicum (Sl) and Solanum tuberosum (St), which lack a PH1 homologous genes, are also depicted. The maps are not drawn to scale. In Arabidopsis a genomic fragment containing some genes belonging to this region is found elsewhere in the genome. In potato we could also identify such...
genomic fragment, but here only a portion is left of the gene immediately downstream of PH1 (7= Glucan 1,3-beta-glucosidase), indicating that the breakpoint, which lead to the loss of PH1, is located within the gene sequence. (d) Comparison of the putative transmembrane domains 5 and 6 of PH1 homologs from different plant species. The asterisk marks the aspartate (D, in one-letter code) residue within the cation-binding site, which is conserved among different ion transporters but is instead mutated in the petunia PH1 protein. All PH1 homologs from plants, with the only exception of the protein from the algae Chlorella variabilis, bear a substitution at this position.

The identified P$_{3\beta}$-ATPases make up a phylogenetic clade that is clearly separated from other P-ATPases, and consists of several sister groups roughly comprising the plant, fungal, slime mold and prokaryotic proteins respectively (Fig. S3). Moreover, we observed clear similarity (micro synteny) among the genes flanking the angiosperm P$_{3\beta}$-ATPase genes, strengthening the view that these genes and the encoded proteins are orthologous (Fig. 5b and S4).

Interestingly, several early sequenced angiosperm genomes lack PH1-related sequences, which contributed to the idea that these proteins were absent from plants. All sequenced Brassicaceae lack P$_{3\beta}$-ATPase sequences, as the encoded proteins with most similarity to PH1 belong to distinct P-ATPase families (Fig. 5b and 6). The same holds true for tomato (S. lycopersicum), potato (S. tuberosum) and pepper (C. annuum) (Figs. 5b and 6). The absence of PH1 homologs in these two families seems due to two independent losses, as related species do contain a PH1 homolog (Fig. 6). In S. lycopersicum the loss of PH1 is associated with the deletion of part of a neighboring gene and in S. tuberosum with the complete loss of that gene (gene #7 in Fig. 5c and S4). Also in Arabidopsis thaliana and Capsella rubella (both Brassicaceae) additional genes appear to be missing around the position where PH1 is located in related species. As the “gap” is identical in Capsella and Arabidopsis, the loss of PH1 was most likely a single event in a common ancestor. Independent losses of PH1 have occurred within the Caryophyllales, because no homolog is found in the Beta vulgaris genome sequence (Dohm et al., 2014) and PCR reactions with primers that amplified PH1 homologs from various species did not yield any product from carnation (Fig. 6). Although PH1 is present in the monocots Phoenix dactylifera, Spirodela polyrrhiza, and Musa acuminata, it is absent from all the grasses (Poaceae) sequenced so far (Sorghum, Panicum, rice, maize, Brachypodium), due to yet another deletion event early in the grass lineage (Fig. 6).

To study the evolutionary origin of the monocot and eudicot PH1 homologs, we examined more distantly related species (Fig 5b and 6). We found a partial EST sequence for the early angiosperm (Magnoliid) Persea americana (avocado), and a complete gene sequence in Amborella trichopoda, two gymnosperms, Picea abies and Pinus taeda, and in Sphagnum fallax (peat moss), a bryophyte. Phylogenetic analyses indicated that the Persea, Amborella, Picea and Sphagnum proteins are more similar, and apparently ancestral to the angiosperm PH1 homologs than to the proteins from fungi, slime molds or prokaryotes (Fig. 5b, S2b and S3). However, the lycophyte Selaginella moellendorfii, the bryophyte Physcomitrella patens, and four algae (Chlamydomonas reinhardtii, Volvox carteri, Coccomyxa subellipsoidea, and Micromonas pusilla), apparently lack PH1 homologs (Fig. 6). The green algae Chlorella variabilis is an exception as its genome encodes a P$_{3\beta}$-ATPase (Fig. 6). However,
the evolutionary relationship of this *Chlorella* protein with spermatophyte PH1 homologs is ambiguous. In most phylogenetic trees the *Chlorella* protein clusters with very weak support with the angiosperm PH1 homologs. Small changes in the selection of proteins used (which affects the curation of the sequence alignment), decrease the bootstrap support or completely disrupt clustering with the other plant proteins (Fig. 5b, S2b and S3), and in some cases even results in weakly supported clustering of the *Chlorella* protein with fungal P$_{3B}$-ATPases. Furthermore, the intron-exon architecture of the *Chlorella* gene (21 exons) is entirely different from that of the analyzed angiosperms (eight exons in various asterids, rosids, *Amborella* and in the monocot *Phoenix dactylifera*), fungal (e.g. species with eight or thirteen exons) or slime mold species (two or three exons) (Fig. S5). Hence, it appears unlikely that this *Chlorella* protein is an early ancestor (ortholog) from which the Angiosperm proteins derived by vertical evolution.

**Evolution of cation-binding site in P$_{3B}$-ATPases**

P$_{3B}$-ATPases of most subfamilies, including proton-pumping P$_{3A}$-ATPases such as AHA2 and PH5, and the bacterial P$_{3B}$-ATPase Mg$^{2+}$ transporters, contain a conserved aspartate (D) in their cation binding site (Pedersen *et al.*, 2007; Eisenach *et al.*, 2014). Replacement of this aspartate in AHA2 with asparagine (N) blocks the transition from the E(1) to E(2)P conformation and completely abolishes ion translocation activity (Pedersen *et al.*, 2007). Interestingly, the conserved aspartate is in PH1 replaced by asparagine (N), which suggests that PH1 is not able to transport any ion at all (Eisenach *et al.*, 2014) and may explain why PH1 has no obvious electrogenic activity on its own (Faraco *et al.*, 2014).

Analysis of the different protein sequences showed that prokaryotic, fungal and slime mold P$_{3B}$-ATPases, as well as the proteins from *Chorella* and *Sphagnum*, all contain this conserved aspartate, whereas it is replaced in all land plants by either an asparagine (N, like in petunia) or a serine (S) (Fig. 5d and S6). A D to N replacement requires a G to A transition of the first base in the encoding triplet (GAU or GAT), whereas a D to S replacement requires transition of both the first (G to A) and second base (A to G). Together with the phylogenetic distribution of the amino acid substitutions, this suggests that the S substitution resulted from a second mutation in an ancestral protein already containing the D to N replacement.

These findings suggest that PH1 homologs have lost their ion transport capacity in early seed plants, coinciding with the appearance of PH5 homologs, which are their binding partners.
Evolution of tonoplast P-ATPase transporters

**Fig. 6** Distribution of *PH1* and *PH5* homologs among plant species. The presence or absence of *PH1* and *PH5* homologs is indicated by green or red ovals respectively. Question marks (?) indicate uncertainty because of incomplete or poorly annotated genome sequence information. The green open circle for *C. variabilis* indicates the presence of a P$_{3}$P$_{4}$-ATPase which is unlikely to be an homolog of PH1.
Expression of PH1 and PH5 homologs in grape and rose

In petunia the combination of PH1 and PH5 is required for the reddish color of the petals necessary for the attraction of pollinators. PH5, but not PH1, is required for the accumulation of tannins in the seeds in both petunia (Verweij et al., 2008; Faraco et al., 2014) and Arabidopsis (Baxter et al., 2005; Appelhagen et al., 2015). To assess the function of PH5 and PH1 homologs in species with different pollination syndromes, we studied their expression in rose, which has brightly colored flowers, and grapevine, which has self-pollinating uncolored flowers, but produces colored berries. In Barbera and Pinot noir, the accumulation of anthocyanins (mainly malvidin 3-O-glycosides) in berries skin marks the onset of ripening (véraison) (Fig. S7), while the berries of Chardonnay stay uncolored due to a mutation in VvMYBa (Kobayashi et al., 2004), a homolog of petunia AN2, a regulator of anthocyanin accumulation in petals (Quattrocchio et al., 1999). RT-PCR analysis showed that VvPH5 and VvPH1 are expressed in the berry skin of Barbera and Pinot, peaking around véraison, (Fig. 7 a,b), whereas their expression remains low and shows no obvious peak at véraison in the unpigmented Chardonnay berries.

Fig. 7 Expression of rose and grape homologs of PH1 and PH5. (a, b) Relative expression of VvPH1 and VvPH5 in developing Chardonnay, Pinot and Barbera berries. The arrow indicates véraison. The accumulation of anthocyanin at the different stages of berry development is reported in Fig. S7. Expression in berries of Chardonnay 20 days after anthesis has been chosen as reference. (c) Expression of RhPH1 and RhPH5 in the last stages of development of the flower buds (stage 4 and stage 6) and in the leaves of rose. The error bars indicate ±SE.
In developing rose petals *RhPH5* and *RhPH1* mRNAs are expressed, simultaneously with the accumulation of pigments, and their expression decreases later in development (Fig. 7c). *RhPH1* and *RhPH5* mRNAs are also detectable in leaves, which in most varieties, including the ones used here, do not accumulate anthocyanin (Fig. 7c).

Although limited to two species, this analysis indicates that PH1 and PH5 appear to have functions outside petals in both pigmented (e.g. berry) and unpigmented tissues (e.g. rose leaves).

**Discussion**

The capacity to differentially regulate the pH in cellular compartments and to generate electrochemical H\(^+\) gradients across membranes is essential for a range of cellular functions. The V-ATPases in various endomembranes and the P\(_{3A}\)-ATPases in the plasma membrane are very ancient proton pumps that appeared already before the separation of green plants from fungi and animals (Kuhlbrandt, 2004; Schumacher & Krebs, 2010), about 1500 MYA (Hedges *et al.*, 2004). Here we present evidence that the heteromeric P\(_{3A/3B}\)-ATPase proton pump encoded by *PH1* and *PH5* came up in early seed plants, prior to the appearance of angiosperms, some 180-200 MYA (Palmer *et al.*, 2004), or even of gymnosperms (300 MYA) and is nowadays wide spread among higher plants.

During evolution the P\(_{3A}\)-ATPase family expanded from 2-3 genes in algae to five distinct classes (I-V) in angiosperms that together comprise some 10-15 genes (Arango *et al.*, 2003; Baxter *et al.*, 2003; Pedersen *et al.*, 2012). Phylogenetic evidence showed that genes belonging to subfamily III, such as PH5 and *AHA10*, originate from a gene duplication well before the appearance of angiosperms and colored flowers, and diverged (neo-functionalized) from the ancestral genes by changes in expression pattern, post-translational regulation and intracellular localization of the encoded protein.

The finding that AHA2 cannot replace the function of PH5, provides direct evidence that the acquisition of a tonoplast localization was a key-step in the neo-functionalization of type III proteins. The sequences that determine sorting of type I and II P\(_{3A}\)-ATPases via the endoplasmatic reticulum (ER) and Golgi to the plasma-membrane lie in the middle part of the protein, in particular in the large cytoplasmic loop, whereas the cytoplasmic N- and C-terminal domains are devoid of such sequences (Lefebvre *et al.*, 2004). Our data indicate that Type III proteins retained these sorting sequences and became tonoplast proteins through the acquisition of a di-leucine motif in the N-terminal cytoplasmic domain that is necessary and, when added to a type II protein, sufficient to send the protein to the tonoplast instead of the plasma membrane. A similar scenario appears to have caused the divergence of plasma membrane and tonoplast-based inositol transporters (Wolfenstetter *et al.*, 2012). It is likely, though not demonstrated, that PH5 homologs move, like type I and II proteins, via the ER and Golgi, thanks to signals in the (highly conserved) middle part of the protein and that the sorting of type I/II and III proteins diverges thereafter, into a (default) pathway towards the plasma membrane and a distinct pathway to the tonoplast, taken if a vacuolar sorting sequence is present.
We infer that the tonoplast localization was acquired early during the evolution of type III proteins, because PH5 homologs from distantly related eudicots, belonging to asterids (petunia), rosids (Arabidopsis, rose, grape), and Caryophyllales (carnation), are all targeted to the tonoplast and can functionally substitute PH5. Most likely this is true also for the PH5 homologs from monocots and gymnosperms, because the sequence motif in the N-terminal domain of PH5 and AHA10 that is required for tonoplast localization is at least partially conserved in all these proteins.

Several observations indicate that type II and III P3α-ATPases diverged also with regard to their post-translational regulation. First, the C-terminal YTV motif of plasma membrane P3α-ATPases, which is important for phosphorylation, subsequent binding of 14-3-3 proteins, multimerization and pump activity (Jahn et al., 1997; Fuglsang et al., 1999; Maudoux et al., 2000; Kanczewska et al., 2005; Ottmann et al., 2007), is in angiosperm PH5 homologs mutated into HTV or QTV and the C-terminal domain of PH5 fails to bind to 14-3-3 proteins in yeast. It is unlikely that the latter is a false negative result, because this PH5 domain can interact in yeast with an AAA-ATPase. In addition, the full PH5 protein is biologically active when expressed in yeast and can form homodimers and bind to PH1, as it does in petunia cells (Verweij et al., 2008; Faraco et al., 2014). Interestingly, the Picea and Pinus PH5 homologs still retained the penultimate tyrosine, while its N-terminal domain contains the sequence for vacuolar localization (Fig 2), suggesting that this gymnosperm protein is a transition stage in the divergence of PH5 homologs from other P3α-ATPases. Second, PH5 is up-regulated by binding to PH1 (Faraco et al., 2014). Although it is unknown whether the structure of type I and II plasma membrane P3α-ATPases would allow binding to PH1, this will not happen in vivo, because PH1 is absent from the plasma membrane (Faraco et al., 2014).

The finding that PH1 encodes a P3β-ATPase discredited the long-held view that P3α-ATPases are unique for prokaryotes (Thever & Saier, 2009; Pedersen et al., 2012). Our data show that P3β-ATPases are, in fact, more omnipresent in seed plants, where they are found in most angiosperms and the two sequenced gymnosperms, than in prokaryotes, fungi, mosses and protists, where they are found in a few groups of species only. Remarkably, P3β-ATPases remained in all analyzed species a single gene, in contrast to the other P-ATPases, which expanded into gene families (Pedersen et al., 2012).

The phylogenetic and syntenic relationships indicate that the seed plant P3β-ATPases have a monophyletic origin and evolved by vertical transmission from a common ancestor that already existed in the earliest seed plants. The P3β-ATPase of Sphagnum fallax groups invariably with the seed plant proteins and the encoding genes share a highly similar intron-exon structure, suggesting that the Sphagnum protein is directly related (ancestral) to the seed plant proteins. If so, the absence of P3β-ATPases in Physcomitrella a bryophyte and Selaginella (a lycophyte) is probably due to a gene loss, similar to those in several groups of angiosperms.

It is hard to reconstruct the deeper evolutionary origin of the P3β-ATPases in detail, especially because several findings are difficult to fit with a simple vertical evolution model. First, only a few groups of prokaryote, fungal and protist species possess a P3β-ATPase. Second, only one of the five
analyzed algae contains a P\textsubscript{3B}-ATPase (*Chlorella*). It is unlikely that the *Chlorella* proteins is orthologous (ancestral) to PH1, because it does not properly cluster with the land plant proteins and the encoding genes have entirely different intro-exon structures. If P\textsubscript{3B}-ATPases evolved by vertical transmittance from a common ancestor, one would have expected to find them in more than a few sporadic fungi and primitive plants, unless they were lost in the large majority of species. Third, the P\textsubscript{3B}-ATPase from the Glomeromycete *Rhizofagus irregularis* does not cluster with P\textsubscript{3B}-ATPase from other fungal phyla (Supplemental Fig. S2 and S3). Taken together, this indicates that the deep evolution of P\textsubscript{3B}-ATPases probably involved one or more horizontal gene transfer events, followed more recently by independent losses in several angiosperm families.

PH1 has, other than boosting the activity of PH5, no known activity on its own and might even not transport any ion as inferred by the substitution of a highly conserved residue in the cation-binding site of the protein (Eisenach *et al.*, 2014; Faraco *et al.*, 2014). Interestingly, this key aspartate residue in the cation-binding site of P-ATPases, which is essential to translocate ions (Buch-Pedersen *et al.*, 2000; Buch-Pedersen & Palmgren, 2003), is conserved in the proteins from prokaryotes, fungi, *Chlorella* and *Sphagnum* (bryophyte) but mutated in all spermatophytic PH1 homologs. This mutation, and the presumed loss of ion translocation capacity, happened early in the evolution of spermatophyte PH1 homologs, coinciding with the appearance of type III P\textsubscript{3A}-ATPase binding partners.

The reason that seed plants co-opted, besides the more ancient V-ATPase, a P-ATPase pump for the acidification of endomembrane compartments may be that P-ATPases can generate steeper trans-membrane proton gradients than V-ATPase, because they hydrolyze more ATP per translocated proton (Rea & Sanders, 1987; Eisenach *et al.*, 2014). In petunia and Arabidopsis PH5/AHA10 is required for the accumulation of tannins in seeds. Since, condensed tannins evolved before flowers and anthocyanins, this function of PH5 is older than that in flower pigmentation. Although PH1 is co-expressed with PH5 in petunia seeds, it is not essential for tannin accumulation, as tannin synthesis is not disrupted in petunia ph1 seeds, or in Arabidopsis, which lacks PH1 While PH1 has probably no activity by its own, PH5 alone is sufficient for tannin accumulation in *Petunia ph1* mutants and Arabidopsis (Appelhagen *et al.*, 2015), which lacks PH1 homologs. This may explain why several angiosperms lost *PH1* but maintained *PH5*, whereas loss of PH5 alone was not observed in plants (Fig. 6b) and suggests that the oldest role of PH5 was the facilitation of vacuolar tannin accumulation. Interestingly, a few species, including for example tomato and potato, lost both *PH1* and *PH5*. That does not necessarily imply that these species lack tonoplast P-ATPase proton pumps altogether, as only for a subset of the 10-20 plant P\textsubscript{3A}-ATPases the intracellular localization is known.

The old age and widespread distribution of PH1 and PH5 homologs indicates that the hyperacidification of vacuoles is a common and ancient phenomenon in seed plants that was only recently, in (some) angiosperms co-opted for the coloration of flowers and attraction of pollinators. Therefore, we can anticipate that additional biological functions for this vacuolar hyper-acidification pathway wait to be uncovered.
Materials and methods

Identification of P-ATPase genes

Sequences of PH1 and PH5 homologs from different species were obtained by the screening of dedicated databases reported in Table S1. We used the full PH1 and PH5 protein to query predicted protein sequences. Because we noticed that PH1 homologs were often overlooked in (early) genome annotations, we also searched with the PH1 protein sequence entire (translated) genome sequences using the tblastn search function. For each species we selected the one or two proteins with the highest similarity to the query (E values equal to zero for angiosperms and higher for less related species) for further phylogenetic analyses. In Table S2 are reported the sequences which annotation we improved as compared to the original data contained in the databases.

A complete list of the species appearing in this analysis is given in Tables S3 and S4. The PH1 and PH5 homologs from rose and carnation we cloned and sequenced with primers reported in table S5.

Phylogeny analysis

Sequence alignments were generated with Multiple Sequence Comparison by Log- Expectation (MUSCLE) (Edgar, 2004) and phylogenetic trees were constructed with Maximum likelihood (PhyML) (Guindon et al., 2010), after curation of the alignments with the G-blocks tool (Castresana, 2000), and then rendered by TreeDyn (Chevenet et al., 2006) using the online tools at www.Phylogeny.fr and http://phylogeny.lirmm.fr (Dereeper et al., 2008). For G-blocks we used the rather stringent default settings of the online tool (minimum number of sequences for a conserved position: 50% + 1, minimum number of sequences for a flank position: 85%, maximum number of contiguous non conserved positions: 8, minimum length of a block: 10, allowed gap positions: none). For PhyML we used the LG substitution model and applied 300 or 500 bootstraps replicates (depending on the number of sequences) to calculate the branch support. Synteny was studied using the Phytozome tool (http://phytozome.jgi.doe.gov/pz/portal.html) and analysis of a 200-kb scaffold from the Petunia genome containing the PhPH1 gene in the middle (Petunia Platform, accession to the petunia genome in preparation).

The alignments relative to the construction of the phylogenetic trees in Fig. 1, 5, S2 and S3 are reported in Note S1 to S4.

Genetic stocks

The ph5 P. hybrida host used for transformation was an F1 hybrid of the ph5 lines V69 and R159. As a transformable ph1 host we used F2 progeny bearing blue-violet (ph1) flowers from the cross V23 (ph1) X V30 (PH1). Carnation (cv. ‘Moonshadow’) was provided by Florigene/Suntory (Osaka, Japan). Grape (Vitis vinifera L.) cv. Chardonnay, Barbera, and Pinot plants from a vineyard at Agliano (Piedmont, Italy) were used for gene expression analysis. RNA was extracted as previously reported (Carra et al., 2007). Roses were purchased at the local market. Primers for the cloning of the PH1 and PH5 homologs from different species are given in Table S5.
Expression analysis
Quantitative RT-PCR was done as described (Faraco et al., 2014), using primers shown in Table S6.

pH measurement of petunia corolla homogenates
Measurements of crude petal extract pH were performed as previously described (Quattrocchio et al., 2006) and were repeated for at least five flowers per plant.

Constructs for expression of PH5 and PH1 homologs
The coding sequence from ATG to STOP (including all introns) of PH1 and PH5 homologs from Arabidopsis, carnation and rose were amplified from genomic DNA by PCR and cloned between the p35S promoter and t35S terminator of Cauliflower Mosaic Virus as described previously for the p35S:PH5 and p35S:PH1 constructs (Verweij et al., 2008; Faraco et al., 2014). Because the genomic fragment containing the VvPH5 coding sequence (15 Kb) was too long, we fused (spliced) exons 1 to 9 from the cDNA to a genomic fragment containing exons 10, 11 and 12, and the four contiguous introns by gene SOEing (Horton et al., 1990), using primers shown in Table S5. For the expression of VvPH1 we used a chimeric fragment consisting of the spliced exons 1 to 4 from cDNA, the central part (exons 5, 6 and 7, and the intervening introns) from genomic DNA and exon 8 from cDNA using primers shown in Table S5. As full length AtAHA2 and PhAHA2 cDNA were not toxic for E. coli, we could use the contiguous (cDNA) coding sequence to generate p35S:AtAHA2 and p35S:PhAHA2. All constructs were prepared in Gateway vectors (destination vector pK2GW7.0) (Karimi et al., 2002). We prepared similar constructs for the expression of the PH5 homologs as GFP fusions. Sequences of primers used to make gene constructs are given in7. Immunoblot analysis of total petal protein was performed as described (Verweij et al., 2008).

The PhAHA2 coding sequence was amplified from petunia R27 petal cDNA with primers 6341(+attB1) and 6342(+attB2) and used to create a Gateway Entry clone by BP reaction with pDONR P1-P2 (Gateway system, LifeTechnologies) and then recombined into pK7FWG2.0 (Karimi et al., 2002) to obtain p35S:PhAHA2-GFP. The analysis of the transgene expression level in the transgenic lines was done by RT-PCR using the primers in Table S8.

For the NPH5-PhAHA2-CPH5 fusion we amplified by PCR three fragments encoding: (i) NPH5 (amino acids 1-29) with primers 4699 (+attB1) and 6348, (ii) PhAHA2 (amino acids 25-880) with primers 6347 and 6350 and (iii) CPH5 (amino acids 885-950) with primers 6349 and 5220(+attB2). The PCR products were purified from gel, mixed in equal molar ratio and used as template for a final PCR amplification with primers 4699(+attB1) and 5220(+attB2). The full-size fragment was inserted in pDONR P1-P2 and subsequently moved to pK7FWG2.0 by Gateway recombination (Life Technologies). All primer sequences are reported in Table S9.

Constructs for the expression of mutated p35S:AtAHA10-GFP fusions were generated starting from the p35S:AtAHA10-GFP clone by conventional restriction/ligation cloning. Fragments containing the respective mutations were amplified in a series of overlap extension PCRs, using primers shown in
Table S9, and inserted into SpeI and PmlII sites of the AtAHA10 sequence in the p35S:AtAHA10-GFP plasmid.

The construct coding for AHA10

AHA2-GFP was generated from AHA2-GFP in pK2GW7.0 by conventional restriction/ligation cloning. The fragment containing the putative AHA10-TSS was amplified in a series of overlap extension PCRs and inserted into the SpeI and BglII sites of AHA2-GFP-pK2GW7. MJ294 and I408 were used as outer primers, I406 and I407 as inner primers for the addition of the MAEDLDKPLL sequence to the N-terminus of AHA2 (for primer sequences see Table S9).

Subcellular localization
Transformation of petunia petal protoplasts and visualization of GFP-fusion proteins and the co-expressed plasma membrane marker RFP-AtSYP122 by confocal laser scanning microscopy were done as reported previously (Faraco et al., 2011). Tobacco BY2 protoplasts were isolated, transformed and analyzed as described before (Appelhagen et al., 2015). Fluorescence was analyzed after 24h incubation in the dark. Epifluorescence images were obtained with a Leica DM5500 microscope equipped with automated DIC condenser and a fluorescein filter set (Leica filter cube I3). Pictures were taken with Diskus digital imaging software (Technisches Buero Hilgers, Koenigswinter, Germany).

Yeast two-hybrid assay
The C-terminal regions of PhPH5, AtAHA2 and PhAHA2 were amplified by PCR using primers shown in Table S8 and cloned into pENTR/D and pDONR207 by a BP reaction (Life Technologies). Entry clones were used to recombine the fragments into pAD-GAL4 and pBD-GAL4 vectors (Stratagene) modified with an attR1-attR2 cassette to make them Gateway compatible. The PhPH5 C-terminal domain chosen consists of the last 100 amino acids of the protein (amino acid 838 to end), that of PhAHA2 of 106 amino acids (amino acid 847 to end) and that of AtAHA2 of 105 amino acids of the PhPH5 protein (amino acid 844 to end). The Ph14-3-3 chi, kappa, omega, omicron and psi coding sequences in the pAD-GAL4 vector were described before. (Verhoef et al., 2013).

PhAAA-ATPase coding sequence in the pAD-GAL4 vector was obtained by screening a cDNA library of petunia line R27 petals with the PhPH5 C-terminal region in pBD-GAL4.

Accession numbers
See Table S1.
Acknowledgments

We thank Drs. Yoshikazu Tanaka and Masako Fukuchi-Mizutani (Suntory Global Innovation Center, Osaka Japan) for providing the sequence of the rose PH5 homolog, the Petunia community for access to the (unpublished) petunia genome sequence, and Erik Manders and Ronald Breedijk (Center of Advanced Microscopy, University of Amsterdam) for help with confocal microscopy.

Author contributions

YL, SP and MB performed experiments, the data mining and participated to write the manuscript. CS, AI, LMdF, WV and TS performed part of the experiments, BW, MS and AS contributed to write the manuscript, RK and FQ conceived the project, performed the phylogenetic analysis and wrote the paper.
References


**Supporting Information**

![Image of phylogenetic tree and alignment](image_url)

**Fig. S1** Comparison of PH5 homologs and other P$_{\text{ATP}}$-ATPases from different plant species. (a) Alignment of the protein sequences of PH5 from petunia (PhPH5) and homologs from Arabidopsis (AtAHA10), rose (RhPH5), grape (VvPH5) with type I and type II Arabidopsis plasma membrane protein AtAHA4 and AtAHA2, plus homologs from Petunia inflata.
(PiAHA2) and grape (VvAHA12). (b) Diagram depicting the gene architecture of several members of the P1A family. Arrows indicate exons. (c) Expression of transgenes in ph5 and ph1 mutants transformed respectively with constructs for the expression of PH5 homologs, PH1 homologues and AtAHA2 in petaloid transgenic plants. RT-PCR was carried out using a limited number of cycles and amplification products were detected by gel-blot hybridization. Although the AtAHA2 transcript is expressed at levels comparable to those of the other transgenes, it is not able to rescue the ph5 phenotype (the color of the flowers are given by the dots above the gel lanes). The expression of the endogenous PH1 or PH5 is undetectable as the primers used for the amplification (see Appendix Table 3) are specific for the transgene. Actin expression is used as reference. (d) Immunoblot of proteins from plants expressing PhPH5-GFP or RhPH5-GFP probed with anti-GFP. Red asterisks indicate the bands of expected size.

**Fig. S2** Comparison of PH1 homologs from different species. (a) Alignment of PH1 homologs from petunia (Ph), grape (Vv) and rose (Rh). In addition to the high conservation of the protein sequence also the position of the introns is conserved. (b) Phylogeny of PH1 homologs from different species. This phylogenetic tree shows clustering of partial *Picea abies* and *Persea americana* P1B-ATPase sequences with Angiosperm homologs. Because the sequences of *Picea abies* and *Persea americana* proteins are incomplete this tree is based on an alignment of partial sequences. The red numbers indicate bootstrap support (% of 300 replicates). The protein alignment from which this tree was generated is reported in Note S3.
Evolution of tonoplast P-ATPase transporters

Fig. S3 Phylogenetic tree of PH1 homologs from a large selection of plants, protozoa, algae, bacteria and archaea. The red numbers indicate bootstrap support (% of 300 replicates). The protein alignment from which this tree was generated is reported in Note S4.
**Fig. S4** Genes appearing in the synteny analysis in **Fig. 5.** Homologous genes in different species have the same number and same color.
**Evolution of tonoplast P-ATPase transporters**

**Fig. S5** Exon structure of *PHI* homologs from different species. Regions encoded by distinct exons are indicated by alternating shades.
Conservation of the cation-binding site in PH1 homologs from fungi and slime molds. Alignment of PH1 sequences (similar to the alignment in Fig. 5d) from different species of fungi and slime molds. The asterisk indicates the aspartate (D) residue that is conserved in different types of ion transporters, but mutated in an asparagine (N) residue in the petunia PH1 protein (marked in red).

![Fig. S6](image)

Accumulation of anthocyanin pigments during development of grape berries of the varieties Barbera and Pinot. In Barbera véraison (arrow) occurs some two weeks later than in Pinot. The error bar indicates ± SD.

![Fig. S7](image)
Table S1: Proteins used in the construction of the phylogenetic trees in Figure 1, 5 and S3.

<table>
<thead>
<tr>
<th>species</th>
<th>database</th>
<th>URL</th>
<th>ID</th>
<th>modified</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PH1 homologs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amborella trichopoda</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>AmTr_v1.0_scaffold_00008.112</td>
<td>yes</td>
</tr>
<tr>
<td>Aquilegia coerulea</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Aqco00080152</td>
<td>yes</td>
</tr>
<tr>
<td>Carica papaya</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>supercontig_55.62</td>
<td>no</td>
</tr>
<tr>
<td>Chlorella variabilis</td>
<td>KEGG</td>
<td><a href="http://www.genome.jp/kegg-bin/show_organism">http://www.genome.jp/kegg-bin/show_organism</a></td>
<td>CHLNCDRAFT_22_876</td>
<td>no</td>
</tr>
<tr>
<td>Coffea canephora</td>
<td>Coffee Genome Hub</td>
<td><a href="http://coffee-genome.org/">http://coffee-genome.org/</a></td>
<td>Ch7_11123000 -&gt; 11135999</td>
<td>manually annotated (partial)</td>
</tr>
<tr>
<td>Eucalyptus grandis</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Eucgr.H04736</td>
<td>no</td>
</tr>
<tr>
<td>Glycine maximum</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Glyma05g30905</td>
<td>no</td>
</tr>
<tr>
<td>Gossypium raimondii</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Gorai.007G235300</td>
<td>no</td>
</tr>
<tr>
<td>Malus domestica</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>MDP0000319016</td>
<td>no</td>
</tr>
<tr>
<td>Malus domestica</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>MDP0000254558</td>
<td>yes</td>
</tr>
<tr>
<td>Manihot esculenta</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Manes.12G129300</td>
<td>no</td>
</tr>
<tr>
<td>Musa acuminate</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>P009399682</td>
<td>no</td>
</tr>
<tr>
<td>Phaseolus vulgaris</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Phvl.002G260400</td>
<td>no</td>
</tr>
<tr>
<td>Phoenix dactylifera</td>
<td>NCBI and KEGG</td>
<td><a href="http://www.genome.jp/kegg-bin/show_organism">http://www.genome.jp/kegg-bin/show_organism</a></td>
<td>XP008775198.1 and 103695607</td>
<td>manually annotated</td>
</tr>
<tr>
<td>Picea abies</td>
<td>CONGENIE</td>
<td><a href="http://congenie.org/">http://congenie.org/</a></td>
<td>lcl:MA_102379 and lcl:MA_1031743</td>
<td>manually annotated from two scaffolds</td>
</tr>
<tr>
<td>Pinus taeda</td>
<td>CONGENIE</td>
<td><a href="http://congenie.org/">http://congenie.org/</a></td>
<td>lcttscaffold8448</td>
<td>no</td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Poiri.005G174790</td>
<td>yes</td>
</tr>
<tr>
<td>Prunus persica</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Prupc.6G194600</td>
<td>no</td>
</tr>
<tr>
<td>Ricinus communis</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>29810.000008</td>
<td>yes</td>
</tr>
<tr>
<td>Sphagnum fallax</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Splfalx0010s0101</td>
<td>no</td>
</tr>
<tr>
<td>Spirodela polyrhiza</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Sspopol15G0017800</td>
<td>no</td>
</tr>
<tr>
<td>Theobroma cacao</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Thecc1EG001029</td>
<td>yes</td>
</tr>
<tr>
<td><strong>PH5 homologs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aquilegia coerulea</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Aqua02000609</td>
<td>no</td>
</tr>
<tr>
<td>Brachypodium distachyon</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Bradi1g72417</td>
<td>no</td>
</tr>
<tr>
<td>Capsella grandiflora</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Cagra.08240094.1</td>
<td>no</td>
</tr>
<tr>
<td>Capsella rubella</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Carubv10008225m</td>
<td>no</td>
</tr>
<tr>
<td>Carica papaya</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>supercontig_1377.1</td>
<td>yes</td>
</tr>
<tr>
<td>Coffee canephora</td>
<td>Coffee Genome Hub</td>
<td><a href="http://coffee-genome.org/">http://coffee-genome.org/</a></td>
<td>Cc04g003560</td>
<td>no</td>
</tr>
<tr>
<td>Dacca carota</td>
<td>CarrotDB</td>
<td><a href="http://apiaceae.njau.edu.cn/carrotdb">http://apiaceae.njau.edu.cn/carrotdb</a></td>
<td>Dck10800</td>
<td>no</td>
</tr>
<tr>
<td>Eucalyptus grandis</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Eucgr.102171</td>
<td>no</td>
</tr>
<tr>
<td>Eutrema</td>
<td>Phytozone</td>
<td><a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a></td>
<td>Thhalv10006710m</td>
<td>no</td>
</tr>
<tr>
<td>Plant Name</td>
<td>Original Source</td>
<td>URL</td>
<td>ID</td>
<td>Note</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>-----</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td><em>Linum usitatissimum</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Lus10037466</td>
<td>No</td>
</tr>
<tr>
<td><em>Gossypium raimondii</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Gorai.011G053900</td>
<td>No</td>
</tr>
<tr>
<td><em>Malus domestica</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>mdp0000303799</td>
<td>No</td>
</tr>
<tr>
<td><em>Manihot esculenta</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>scaffold05875:13106-80</td>
<td>No (two genes on same scaffold)</td>
</tr>
<tr>
<td><em>Salix purpurea</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>SapurV1A.0273s011-0.1</td>
<td>No</td>
</tr>
<tr>
<td><em>Setaria italica</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>S034109</td>
<td>No</td>
</tr>
<tr>
<td><em>Spirodela polyrhiza</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Spipo2G0033700 and Spipo2G0033800</td>
<td>Yes, from two gene fragments</td>
</tr>
<tr>
<td><em>Sphagnum fallax</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Sphfalx0286s0011.1</td>
<td>No</td>
</tr>
</tbody>
</table>

**Other P34-ATPases**

<table>
<thead>
<tr>
<th>Plant Name</th>
<th>Original Source</th>
<th>URL</th>
<th>ID</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Brachypodium distachyon</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Bradi5g24690</td>
<td></td>
</tr>
<tr>
<td><em>Capsella rubella</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Carubv10025825m</td>
<td>No</td>
</tr>
<tr>
<td><em>Carica papaya</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Supercontig1377:2570</td>
<td>No</td>
</tr>
<tr>
<td><em>Eucalyptus grandis</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Eucgr.C02386</td>
<td>No</td>
</tr>
<tr>
<td><em>Eutrema salsugineum</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Thhalv10003603m</td>
<td>No</td>
</tr>
<tr>
<td><em>Gossypium raimondii</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Gorai.013G212200</td>
<td>No</td>
</tr>
<tr>
<td><em>Linum usitatissimum</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Lus10001631</td>
<td>No</td>
</tr>
<tr>
<td><em>Mimulus guttatus</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Mgvl1a000867m</td>
<td>No</td>
</tr>
<tr>
<td><em>Petunia inflata</em></td>
<td>Genome</td>
<td><a href="http://petunia.sgn.cornell.edu/blast/blast.html">http://petunia.sgn.cornell.edu/blast/blast.html</a></td>
<td>Pinf1012scf03264</td>
<td>No</td>
</tr>
<tr>
<td><em>Panicum virgatum</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Pairv00044612</td>
<td>No</td>
</tr>
<tr>
<td><em>Setaria italica</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Si06g031240</td>
<td>No</td>
</tr>
<tr>
<td><em>Sorghum bicolor</em></td>
<td>Phytozone</td>
<td><a href="https://phytozone.jgi.doe.gov/pz/portal.html">https://phytozone.jgi.doe.gov/pz/portal.html</a></td>
<td>Sbh08g023070</td>
<td>No</td>
</tr>
</tbody>
</table>

The table reports the database containing the original sequence, its URL, and the ID of the protein. For proteins that have been reannotated manually to remove mistakes, the corrected sequence is given in Table S2.
Table S2  Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Organism</th>
<th>Accession</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGSIGT...</td>
<td>Malus domestica</td>
<td>XP_004504112</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MGPWKT...</td>
<td>Coffea canephora</td>
<td>XP005848152</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MGNING...</td>
<td>Chlorella variabilis</td>
<td>XP_001884015</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MGD...</td>
<td>Aquilegia coerulea</td>
<td>XP_000802168</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MG...</td>
<td>Cicer arietinum</td>
<td>XP_000802168</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MGA...</td>
<td>Chlorella variabilis</td>
<td>XP001884015</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
<tr>
<td>MGF...</td>
<td>Cicer arietinum</td>
<td>XP_000802168</td>
<td>ATPase transporters</td>
<td>Table S2 Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.</td>
</tr>
</tbody>
</table>

**Table S2** Sequences of proteins which sequence was (partially) reannotated on basis of the gene structure in other species.
Table S3 List of plant species appearing in the paper.

<table>
<thead>
<tr>
<th>Scientific name</th>
<th>Common name</th>
<th>Group</th>
<th>Class</th>
<th>Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquilegia coerulea</td>
<td>Indian snakeroot</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Ranunculaceae</td>
</tr>
<tr>
<td>Amborella trichopoda</td>
<td>-</td>
<td>Angiosperm</td>
<td>Amborellales</td>
<td>Amborellaceae</td>
</tr>
<tr>
<td>Arabidopsis lyrata</td>
<td>-</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>thale cress</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Beta vulgaris</td>
<td>sugar beet</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Amaranthaceae</td>
</tr>
<tr>
<td>Brachypodium distachyon</td>
<td>purple false brome</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
<tr>
<td>Brassica rapa</td>
<td>field mustard</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Camelia sinensis</td>
<td>tea plant</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Theaceae</td>
</tr>
<tr>
<td>Camptotheca acuminata</td>
<td>happy tree</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Cornaceae</td>
</tr>
<tr>
<td>Capsella rubella</td>
<td>pink shepherd's purse</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Capsella grandiflora</td>
<td>-</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Carica papaya</td>
<td>papaya</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Caricaceae</td>
</tr>
<tr>
<td>Catharanthus roseus</td>
<td>Madagascan periwinkle</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Apocynaceae</td>
</tr>
<tr>
<td>Chlamydomonas reinhardtii</td>
<td>-</td>
<td>Chlorophyte</td>
<td>Chlorophyceae</td>
<td>Chlamydomonadaceae</td>
</tr>
<tr>
<td>Chlorella variabilis</td>
<td>-</td>
<td>Chlorophyte</td>
<td>Chlorophyceae</td>
<td>Chlorophyceae</td>
</tr>
<tr>
<td>Coccomyza subellipsoides</td>
<td>-</td>
<td>Chlorophyte</td>
<td>Chlorophyceae</td>
<td>Coccomyceae</td>
</tr>
<tr>
<td>Coffea canephora</td>
<td>coffee robusta</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Rubiaceae</td>
</tr>
<tr>
<td>Cucumis sativus</td>
<td>cucumber</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Cucurbitaceae</td>
</tr>
<tr>
<td>Dianthus caryophyllus</td>
<td>carnation</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Caryophylaceae</td>
</tr>
<tr>
<td>Eucalpytus grandis</td>
<td>eucalyptus</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Myrtaceae</td>
</tr>
<tr>
<td>Eutrema salsugineum</td>
<td>salt cress</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Brassicaceae</td>
</tr>
<tr>
<td>Fragaria vesca</td>
<td>strawberry</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Rosaceae</td>
</tr>
<tr>
<td>Glycine max</td>
<td>soybean</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Gossypium raimondii</td>
<td>diploid cotton</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Malvaceae</td>
</tr>
<tr>
<td>Lotus japonicus</td>
<td>Lotus, wild legume</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Linum usitatissimum</td>
<td>flax</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Linaceae</td>
</tr>
<tr>
<td>Malus domestica</td>
<td>apple</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Rosaceae</td>
</tr>
<tr>
<td>Manihot esculenta</td>
<td>cassava</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Euphorbiaceae</td>
</tr>
<tr>
<td>Medicago truncatula</td>
<td>clover</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Micromonas pusilla</td>
<td>-</td>
<td>Chlorophyte</td>
<td>Prasinophyceae</td>
<td>Mamiellaceae</td>
</tr>
<tr>
<td>Mimulus guttatus</td>
<td>monkey flower</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Phrymaceae</td>
</tr>
<tr>
<td>Musa acuminata</td>
<td>wild banana</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Musaceae</td>
</tr>
<tr>
<td>Nicotiana plumbaginifolia</td>
<td>tex-mex tobacco</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Solanaceae</td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>rice</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
<tr>
<td>Nelumbo nucifera</td>
<td>Sacred lotus</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Nelumboaceae</td>
</tr>
<tr>
<td>Panicum virgatum</td>
<td>switchgrass</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
<tr>
<td>Persea americana</td>
<td>avocado</td>
<td>Angiosperm</td>
<td>Magnolids</td>
<td>Lauraceae</td>
</tr>
<tr>
<td>Phaseolus vulgaris</td>
<td>kidney bean</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Fabaceae</td>
</tr>
<tr>
<td>Physochlaina patens</td>
<td>moss</td>
<td>Bryophyte</td>
<td>Bryopsida</td>
<td>Funariaceae</td>
</tr>
<tr>
<td>Phoenix dactylifera</td>
<td>date palm</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Araceae</td>
</tr>
<tr>
<td>Picea abies</td>
<td>Norway spruce</td>
<td>Gymnosperm</td>
<td>Pinopsida</td>
<td>Pinaceae</td>
</tr>
<tr>
<td>Pimus taeda</td>
<td>loblolly pine</td>
<td>Gymnosperm</td>
<td>Pinopsida</td>
<td>Pinaceae</td>
</tr>
<tr>
<td>Populus trichocarpa</td>
<td>poplar</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Salicaceae</td>
</tr>
<tr>
<td>Prunus persica</td>
<td>peach</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Rosaceae</td>
</tr>
<tr>
<td>Rauvolfia serpentina</td>
<td>Indian snakeroot</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Apocynaceae</td>
</tr>
<tr>
<td>Rosa hybrida</td>
<td>rose</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Rosaceae</td>
</tr>
<tr>
<td>Salix purpurea</td>
<td>purple willow</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Salicaceae</td>
</tr>
<tr>
<td>Setaria italica</td>
<td>foxtail millet</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
<tr>
<td>Selaginella moellendorfii</td>
<td>spike moss</td>
<td>Lycophyte</td>
<td>Isoetopsida</td>
<td>Selaginellaceae</td>
</tr>
<tr>
<td>Solanum lycopersicum</td>
<td>tomato</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Solanaceae</td>
</tr>
<tr>
<td>Solanum tuberosum</td>
<td>potato</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Solanaceae</td>
</tr>
<tr>
<td>Sorghum bicolor</td>
<td>sorghum</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
<tr>
<td>Sphagnum fallax</td>
<td>peat moss</td>
<td>Bryophyte</td>
<td>Sphagnopsida</td>
<td>Sphagnaceae</td>
</tr>
<tr>
<td>Spriodela polyrhiza</td>
<td>duckweed</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Araceae</td>
</tr>
<tr>
<td>Theobroma cacao</td>
<td>cacao plant</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Malvaceae</td>
</tr>
<tr>
<td>Volvax carteri</td>
<td>-</td>
<td>Chlorophyte</td>
<td>Chlorophyceae</td>
<td>Volvaceae</td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>grape</td>
<td>Angiosperm</td>
<td>Eudicots</td>
<td>Vitaceae</td>
</tr>
<tr>
<td>Zea mays</td>
<td>corn</td>
<td>Angiosperm</td>
<td>Monocots</td>
<td>Poaceae</td>
</tr>
</tbody>
</table>
Table S4 List of non-plant species appearing in the paper.

<table>
<thead>
<tr>
<th>Scientific name</th>
<th>Kingdom</th>
<th>Phylum/class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danio rerio (zebrafish)</td>
<td>Metazoa</td>
<td>Chordata</td>
</tr>
<tr>
<td>Dictyostelium discoideum</td>
<td>Amoebozoa</td>
<td>Mycetozoa/Dictyostelia</td>
</tr>
<tr>
<td>Polysphondylum pullidum</td>
<td>Amoebozoa</td>
<td>Mycetozoa/Dictyostelia</td>
</tr>
<tr>
<td>Claviceps purpurea</td>
<td>Fungi</td>
<td>Ascomycota</td>
</tr>
<tr>
<td>Coccidioides immitis</td>
<td>Fungi</td>
<td>Ascomycota</td>
</tr>
<tr>
<td>Cordyceps militaris</td>
<td>Fungi</td>
<td>Ascomycota</td>
</tr>
<tr>
<td>Fusarium fujikuroi</td>
<td>Fungi</td>
<td>Ascomycota</td>
</tr>
<tr>
<td>Laccaria bicolor</td>
<td>Fungi</td>
<td>Agaricomycetes</td>
</tr>
<tr>
<td>Metarhizium acridum</td>
<td>Fungi</td>
<td>Ascomycota</td>
</tr>
<tr>
<td>Rhizopagus irregularis</td>
<td>Fungi</td>
<td>Glomeromycota</td>
</tr>
<tr>
<td>Methanogenium boonei</td>
<td>Archaea</td>
<td>Euryarchaeota</td>
</tr>
<tr>
<td>Bacillus cereus</td>
<td>Bacteria</td>
<td>Firmicutes</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>Bacteria</td>
<td>Proteobacteria</td>
</tr>
<tr>
<td>Enterococcus gallinarum</td>
<td>Bacteria</td>
<td>Firmicutes</td>
</tr>
<tr>
<td>Shigella flexneri</td>
<td>Bacteria</td>
<td>Proteobacteria</td>
</tr>
<tr>
<td>Salmonella enterica</td>
<td>Bacteria</td>
<td>Proteobacteria</td>
</tr>
</tbody>
</table>

Table S5 Primers for the cloning of PH1 and PH5 homologues

<table>
<thead>
<tr>
<th>Primers</th>
<th>Gene</th>
<th>Sequence</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4563</td>
<td>VvPH5</td>
<td>ATGGCGGAGGATCTGGGAC</td>
<td>F</td>
</tr>
<tr>
<td>4803</td>
<td>VvPH5</td>
<td>TGCTGATCCAAAGGAGGTAGG</td>
<td>R</td>
</tr>
<tr>
<td>4790</td>
<td>VvPH5</td>
<td>GCCAAAGATATGGACAAAGACACA</td>
<td>F</td>
</tr>
<tr>
<td>4807</td>
<td>VvPH5</td>
<td>TGATGGGTTCGCTGGTGTATT</td>
<td>R</td>
</tr>
<tr>
<td>4806</td>
<td>VvPH5</td>
<td>AGTTGGCAATCGAAAGGAG</td>
<td>F</td>
</tr>
<tr>
<td>4654</td>
<td>VvPH5</td>
<td>ATCCAGGCAGCTCATACGGTATGA</td>
<td>R</td>
</tr>
<tr>
<td>117</td>
<td>VvPH5</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGCAAC3</td>
<td>F</td>
</tr>
<tr>
<td>118</td>
<td>VvPH5</td>
<td>GGGGACCAGTTTGTACAGAAAAGCTGGGTCATTACGTGCGTCG</td>
<td>R</td>
</tr>
<tr>
<td>4836</td>
<td>VvPH1</td>
<td>ATGGCAACTCCAGATTTCCTCA</td>
<td>F</td>
</tr>
<tr>
<td>4934</td>
<td>VvPH1</td>
<td>AGATCAGAGCAGCTCGTCTTAGA</td>
<td>R</td>
</tr>
<tr>
<td>4933</td>
<td>VvPH1</td>
<td>CACTAACAGGGAGGTCTGGAGT</td>
<td>F</td>
</tr>
<tr>
<td>4938</td>
<td>VvPH1</td>
<td>CACTTACACCTCAGATGGTCTCCAC</td>
<td>R</td>
</tr>
<tr>
<td>4935</td>
<td>VvPH1</td>
<td>TCACCTCGAGGGTTGTGGTAA</td>
<td>F</td>
</tr>
<tr>
<td>4837</td>
<td>VvPH1</td>
<td>TTACAGCCATTTGTGTA</td>
<td>R</td>
</tr>
<tr>
<td>119</td>
<td>VvPH1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTTATTAGCCAGGTCA</td>
<td>F</td>
</tr>
<tr>
<td>120</td>
<td>VvPH1</td>
<td>GGGGACCAGTTTGTACAGAAAAGCTGGGTA</td>
<td>R</td>
</tr>
</tbody>
</table>

1The orientation of the primer compared to the gene is indicated as Forward (F) or Reverse (R).
**Table S6** Primer for Real Time PCR analysis of the expression of different genes.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Gene</th>
<th>Sequence</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>95</td>
<td>RhPH5</td>
<td>AATGGCGGGAATGGATGTC</td>
<td>F</td>
</tr>
<tr>
<td>96</td>
<td>RhPH5</td>
<td>AGTGAGGGCGGTTCCAGGTAAG</td>
<td>R</td>
</tr>
<tr>
<td>101</td>
<td>DcPh5</td>
<td>GCCTCTACTAAGTCCCGAGGAACT</td>
<td>F</td>
</tr>
<tr>
<td>102</td>
<td>DcPh5</td>
<td>CTCCAAAGGCAAACGTCCA</td>
<td>R</td>
</tr>
<tr>
<td>65</td>
<td>VsPh5</td>
<td>GGATCTGGACAAAAACATTGGCTT</td>
<td>F</td>
</tr>
<tr>
<td>66</td>
<td>VsPh5</td>
<td>GATATCGTCCAGATCAATGC</td>
<td>R</td>
</tr>
<tr>
<td>103</td>
<td>PhActin</td>
<td>TCCCGTGTCCAGAAGTACTGTT</td>
<td>F</td>
</tr>
<tr>
<td>104</td>
<td>PhActin</td>
<td>TCGTGAATGCCACGACCTT</td>
<td>R</td>
</tr>
<tr>
<td>63</td>
<td>VsPH1</td>
<td>CGCAATTCGTTTCAGTTTT</td>
<td>F</td>
</tr>
<tr>
<td>64</td>
<td>VsPH1</td>
<td>CGGTGACCTCCATCTATTTCCTT</td>
<td>R</td>
</tr>
<tr>
<td>105</td>
<td>RhPH1</td>
<td>TCCAACAGGGTGTTTATGCTGGTG</td>
<td>F</td>
</tr>
<tr>
<td>106</td>
<td>RhPH1</td>
<td>TCGACCCCCATCAATTTTT</td>
<td>R</td>
</tr>
<tr>
<td>107</td>
<td>PhPH1</td>
<td>AGTTTTTGCCCCGACTACCTACT</td>
<td>F</td>
</tr>
<tr>
<td>108</td>
<td>PhPH1</td>
<td>CATGATCACCCTTTGTGTGCA</td>
<td>R</td>
</tr>
<tr>
<td>75</td>
<td>VvUbi</td>
<td>TCTGAGGCTGTTGCTGTT</td>
<td>F</td>
</tr>
<tr>
<td>76</td>
<td>VvUbi</td>
<td>AGGCGTGCAATGATGCTTGGCGT</td>
<td>R</td>
</tr>
<tr>
<td>91</td>
<td>VvActin</td>
<td>GCCCGCTGCTCAGTTGACATG</td>
<td>F</td>
</tr>
<tr>
<td>92</td>
<td>VvActin</td>
<td>CCTGGCCGACCCCAATA</td>
<td>R</td>
</tr>
<tr>
<td>113</td>
<td>RhActin</td>
<td>CTCTTAACCCAAAGGCCAACA</td>
<td>F</td>
</tr>
<tr>
<td>114</td>
<td>RhActin</td>
<td>GGGCACTTGAGGTCGTCTAA</td>
<td>R</td>
</tr>
<tr>
<td>115</td>
<td>DcActin</td>
<td>GAGAAGCTGCGTCATGCTGCT</td>
<td>F</td>
</tr>
<tr>
<td>116</td>
<td>DcActin</td>
<td>TGAGCTGTTATGGCCGTTTT</td>
<td>R</td>
</tr>
</tbody>
</table>
Table S7 Primers used to prepare constructs for ectopic expression, expression of GFP fusion proteins and Y2H assay.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Gene/motif</th>
<th>Sequence 5’-3’</th>
<th>sites</th>
<th>Cloned in</th>
</tr>
</thead>
<tbody>
<tr>
<td>2039</td>
<td>PhPH5</td>
<td>TTAGCCATGGCCAGATCTGGAGAGACC</td>
<td>NcoI</td>
<td>pEntrGFP</td>
</tr>
<tr>
<td>2040</td>
<td>PhPH5</td>
<td>CATGACCTATGGACAAACTGTAAGCCTGTGGTT</td>
<td>NcoI</td>
<td>pEntrGFP</td>
</tr>
<tr>
<td>2078</td>
<td>PhPH5</td>
<td>CGGGATCTGGAGACCAAGATTTTATAAGGAGG</td>
<td>BamHI</td>
<td>pEntrGm</td>
</tr>
<tr>
<td>4484</td>
<td>RhPH5</td>
<td>GGGGACAAGGTTGTCACAAAAGCAAGGCAAGATCTGGAGAGACC</td>
<td>attB1</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4485</td>
<td>RhPH5</td>
<td>GGGGACCACTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4486</td>
<td>RhPH5</td>
<td>GGGGACCAAGTTTGTACAAAAGCAAGGCAAGATCTGGAGAGACC</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4787</td>
<td>DcPH5</td>
<td>GGGGACAAGTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB1</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4788</td>
<td>DcPH5</td>
<td>GGGGACCACTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4789</td>
<td>VvPH5</td>
<td>GGGGACAAGTTTGTACAAAAGCAAGGCAAGATCTGGAGAGACC</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4790</td>
<td>VvPH5</td>
<td>GGGGACCACTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4762</td>
<td>AtAHA10</td>
<td>GGGGACAAGTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4763</td>
<td>AtAHA10</td>
<td>GGGGACCACTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4764</td>
<td>AtAHA10</td>
<td>GGGGACACCTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>3013</td>
<td>AtAHA2</td>
<td>GGGCCATGGCGACTGTTGCGAGATCTGGAGAGACC</td>
<td>BamHI</td>
<td>pEntrGFP</td>
</tr>
<tr>
<td>3014</td>
<td>AtAHA2</td>
<td>CGGGACTCCTGCTGACTGTTGCGAGATCTGGAGAGACC</td>
<td>EcoRI</td>
<td>pEntrGFP</td>
</tr>
<tr>
<td>3742</td>
<td>AtAHA2</td>
<td>CGGGACTCCTGCTGACTGTTGCGAGATCTGGAGAGACC</td>
<td>EcoRI</td>
<td>pEntrGm</td>
</tr>
<tr>
<td>6341</td>
<td>AtAHA2</td>
<td>GGGGACAAGTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB1</td>
<td>pDONR207</td>
</tr>
<tr>
<td>6342</td>
<td>AtAHA2</td>
<td>GGGGACCACTTTTGTACAAAGAAGCTGGGTACTGTTGCGGGAAGAAATAA</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>3700</td>
<td>PhPh5 Cterm</td>
<td>cacCCCTGCTGAGAAGCGTGGTG</td>
<td>Topo</td>
<td>pENTR/D</td>
</tr>
<tr>
<td>3701</td>
<td>PhPH5 Cterm</td>
<td>TCAAACTGTAGCTGTTGGAATTTTGT</td>
<td>Topo</td>
<td>pENTR/D</td>
</tr>
<tr>
<td>6943</td>
<td>PhAHA2 Cterm</td>
<td>ggggacaaggtgtaaaaaagctggcctGGGAAAGGCTTT</td>
<td>attB1</td>
<td>pDONR207</td>
</tr>
<tr>
<td>6944</td>
<td>PhAHA2 Cterm</td>
<td>ggggaccacctgtagcaaaaaagctggcctGGGAAGGCTTT</td>
<td>attB2</td>
<td>pDONR207</td>
</tr>
<tr>
<td>4464</td>
<td>AtAHA2 Cterm</td>
<td>cacACTTCTGCGAAAGGCGTGAGC</td>
<td>Topo</td>
<td>pENTR/D</td>
</tr>
<tr>
<td>4465</td>
<td>AtAHA2 Cterm</td>
<td>CTACACACTGTAGCTGAGGAG</td>
<td>Topo</td>
<td>pENTR/D</td>
</tr>
</tbody>
</table>
## Table S8 Primers for RT-PCR analysis of transgene expression in transgenic plants

<table>
<thead>
<tr>
<th>Primers</th>
<th>Gene</th>
<th>Sequence 5'-3'</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4775</td>
<td>PH5 deg</td>
<td>ATAGCAGTGGCAGTGNCACAGATGCTNACAGATGCT</td>
<td>F</td>
</tr>
<tr>
<td>53a</td>
<td>attR2</td>
<td>ACCACTTTGTACAAAGAAAGCTG</td>
<td>R</td>
</tr>
<tr>
<td>1676</td>
<td>PhPH5</td>
<td>TCATGGCAGATCATAGGCGTCTCAG</td>
<td>R</td>
</tr>
<tr>
<td>4329</td>
<td>P1 deg1</td>
<td>GCN GAA GGA GTA AAA GCN AAA GT</td>
<td>F</td>
</tr>
<tr>
<td>629</td>
<td>35S term</td>
<td>TGCTCAACACATGAGCGAAAC</td>
<td>R</td>
</tr>
</tbody>
</table>

*1The orientation of the primer compared to the gene is indicated as Forward (F) or Reverse (R).

## Table S9 Primers used to prepare constructs for expression of GFP fusion to mutated versions of PH5 and AHA10.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene or mutation</th>
<th>Sequence</th>
<th>Cloned in</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1294</td>
<td>AHA10 external-FW</td>
<td>CGCAAGACCTCTCTCTATATAAGGAAGTTC</td>
<td>35S:AHA10-Gl</td>
</tr>
<tr>
<td>1938</td>
<td>AHA10 external rev</td>
<td>TGAATGTTAAGAATGAAGTTC</td>
<td>35S:AHA10-Gl</td>
</tr>
<tr>
<td>1400</td>
<td>PLL-&gt;AAA</td>
<td>TTTGGAAGCCAGACCTGCTGATCTGATATTTTCAAC</td>
<td></td>
</tr>
<tr>
<td>1401</td>
<td>PLL-&gt;AAA</td>
<td>AGCAGCTGCTTGTCCAAAATCTCGGACCAGT</td>
<td></td>
</tr>
<tr>
<td>1402</td>
<td>ELDKPLL-&gt;AAAAAA AA</td>
<td>GCTGCTGCTGCTGCTGAGCTGCTGATCTGATATTTTCAAC</td>
<td></td>
</tr>
<tr>
<td>1403</td>
<td>ELDKPLL-&gt;AAAAAA AA</td>
<td>CTCGACACAGACAGACGGCCATGAGCCCTGCTTTC</td>
<td></td>
</tr>
<tr>
<td>1404</td>
<td>Deletion of ELDKPLL</td>
<td>CTCCATGGCAGATCTGATATTTTCAACAG</td>
<td></td>
</tr>
<tr>
<td>1405</td>
<td>Deletion of ELDKPLL</td>
<td>ATCGAGACCTGCCGAGCTGCTTTC</td>
<td></td>
</tr>
<tr>
<td>1406</td>
<td>AHA1010-AH2A2</td>
<td>CTTGTTTGAGTGAACCAAG</td>
<td></td>
</tr>
<tr>
<td>1407</td>
<td>AHA1010-AH2A2</td>
<td>ATGGCTGCTCAGATCTGAGAG</td>
<td></td>
</tr>
<tr>
<td>6341 (+attB1)</td>
<td>PhAHA2 (aa.1) Fw</td>
<td>ggggacagttgtacaanaaacagggagttgattataaggaagctcagcctgyagaga</td>
<td>pDONRP1-P2</td>
</tr>
<tr>
<td>6342 (+attB2)</td>
<td>PhAHA2 (aa.952) Rv</td>
<td>ggggacactctcagacaagacagtcaactgtaataagatctgctgagtctgtgccagagatctgagag</td>
<td>pDONRP1-P2</td>
</tr>
<tr>
<td>4699 (+attB1)</td>
<td>Ph5(aa.1) Fw</td>
<td>ggggacagttgtacaanaaacagggagttgattataaggaagctcagcctgyagaga</td>
<td>pDONRP1-P2</td>
</tr>
<tr>
<td>5220 (+attB2)</td>
<td>Ph5(aa.950) Rv</td>
<td>ggggacactctcagacaagacagtcaactgtaataagatctgctgagtctgtgccagagatctgagag</td>
<td></td>
</tr>
<tr>
<td>6348 (+attB1)</td>
<td>Ph5(aa.29) Rv</td>
<td>GTACATTTCAGCTGCTCAGATCTGAGAGTCGTCC</td>
<td></td>
</tr>
<tr>
<td>6347</td>
<td>PhAHA2(aa.25) Fw</td>
<td>GAAAAGTGTTGCGGACTGATAGTTCAGATCGTAC</td>
<td></td>
</tr>
<tr>
<td>6350</td>
<td>PhAHA2(aa.880) Rv</td>
<td>CGCTGTAAACTTCTCTGAGAAATGGCCCATGTGAGCTT</td>
<td></td>
</tr>
<tr>
<td>6349 (+attB1)</td>
<td>Ph5(aa.885) Fw</td>
<td>GGAAGCTCAATGGGCACTTTCTCAGAGAGATCTGACG</td>
<td></td>
</tr>
<tr>
<td>6343</td>
<td>Ph5(aa.30) Fw</td>
<td>TGGAGAAAATTCCTCAGATGTTGAGACAGCTGATAAGATCCAAG</td>
<td></td>
</tr>
<tr>
<td>6344 (+attB1)</td>
<td>PhAHA2(aa.24) Rv</td>
<td>CGAGAGGCGAGTCGGCGGCTGTGAGAGTCGTCC</td>
<td></td>
</tr>
<tr>
<td>6345 (+attB1)</td>
<td>PhAHA2(aa.881) Fw</td>
<td>ATGGGAGAGGCGAGTCGGCGGCTGTGAGAGTCGTCC</td>
<td></td>
</tr>
<tr>
<td>6346</td>
<td>Ph5(aa.884) Rv</td>
<td>AGTCCCATGTTAAGTTCTGAGCAACCCACTGGCCTTCCT</td>
<td></td>
</tr>
</tbody>
</table>

92
Note S1. Protein alignment (in Fasta format) from which the tree in Figure 1 was generated. The alignment was cured by G-boxes.

Carnation_ph5
IVLGFMILAIWWEYDFPPFMVIIILAILNDGTIMTISKDRVKSPSPSDWKLQIEIFATGVVGITYLALVTNLVFWVANGTD
FKEKHFSVSLSAIYLQVISSIQALIFVTRTSQWSFLERPGILLSAFVAVQQALATLIAYEAINLQSVGSDSLQLVLL
YSIV YFPPLD

Spaghnum_fallax_Sphfalx0163s0015.1
IVLGFMILAIWKFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGJIVLGYALAMMTIVFVWAIKTD
FFSSHIFHVPISAIYLVPSLQVSAFLVFTRTSRSWSFVERPGLLMLMPSAFVIAQLATflaterVANWFGVQIGGWGAVWLV
YSIVYIPPLD

AtAHAA2(NP_001190870)
IVLGFMILAIWCEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPTPDWKLKEIFATGVVGGYQAIMTVFVWAIKTD
FFSDTFGVRSIAVYLQVISSIQALIFVTSRSRSWSVERPALLMIAFLIAQLIAVAINWSFIRGIGGWGAVWVL
YSIVYFPLD

Brachypodium_distachyon (Bradi5g24690.1)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLNEIFATGIVVLGGTYLALMTVWWFVWAIKTD
FNTNKFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLTTAFLAQLLATLAVAINWFEKIGIGGWGAVWVL
YSIVYFPPLD

Oryza_sativa (NP001054118)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIKIGGWGAVWVL
YSIVYFPLD

Zea_mays (GRMZM2G019404)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIGIGGWGAVWVL
YSIVYFPLD

Zea_mays(GRAMZM2G06894)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIKIGGWGAVWVL
YSIVYFPLD

Sorghum_bicolor (Sb06g031240.1)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIKIGGWGAVWVL
YSIVYFPLD

Setaria_italica (Si021113m.g)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FPEKEFGKPSSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIGIGGWGAVWVL
YSIVYFPLD

Panicum_virgatum(Pavirv00044632m)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEFIRIGGWGAVWVL
YSIVYFPLD

Gossypium_raimondii(Gorai.013G212200.1)
IVLGFMILAIWKFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FPEKEFGKPSSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEKIGIGGWGAVWVL
YSIVYFPLD

Cucumis_sativus_ph5_like2(XP004146327)
IVLGFMILAIWKEFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD
FFSDTFGRVSIALYLVPSLQVSAFLVFTRTSRSWSFVERPGLLLLLTAFMAILAQVATLAVAINWFEFIRIGGWGAVWVL
YSIVYFPLD

Linum_usitatissimum(Lus10001631)
IVLGFMILAIWKFDFSPFMVIIILAILNDGTIMTISKDRVKPSPLPDWKLKEIFATGIVLGSYALMTVFWAIKTD

93
Evolution of tonoplast P-ATPase transporters

\[
\text{FPDKFGVRPSAALYQVSISQLIFVTRSRSSFSVERPGLLLVMVAFAIQLATIAVANWGFIKGIGGWAGVWVL}
\]

\[
\text{YSIVFYFPDL}
\]

- \text{Medicago truncatula (XP003610080) VIVFGFMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

- \text{Paeonia lactiflora (XP004242908) IVFGLGMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

- \text{Nicotiana plumbaginifolia (XP004154026) IVFGLGMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

- \text{Prunus persica (EMJ26692) IVFGLGMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

- \text{Nicotiana plumbaginifolia (XP004154026) IVFGLGMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

- \text{Carica papaya (supercontig 1377:2570) IVFGLGMIALIKWKFIPSPFMVILIALNDGTMITISKDRKPSPMDPSWKLKEIFATGIVLGYGGYLMALMTVFWMHDTS FFTEKFGVSIALYQVSISQLIFVTRSRSSFSVERPGLLLVIAFIFIQLIAVANWGFIKGIGGWAGVWVL YSIVFYFPDL}

94
Chapter 3

>Medicago truncatula(XP003594954)
IVFGMFIALIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGVMLGGYLAMMTTVIFFWGQVIDK
FPFERFGVRHIAALYLQVSIQLALIFVRSSWSYVERPGMLLGMQAVIAQLIALIAVANWGFIQGIGWGAGSI
YSIFYPLD

>Panicum vulgaiis(ESW10800)
IVFGMFIALIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLKEIFATGIVLGLYLLMLMTTVIFFWA
FPDKFGRVRHIAALYLQVSIQLALIFVRSSWSFIERPGMLLGLQAVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Glycine max(XP006586984)
IVFGFIALIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGVGVLGGYLAMMTTVIFFWA
FPFRFVGKSLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Mimulus guttatus(mgv1a000867m)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPKIFHVSLASYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Sorgum bicolor(Sb08g023070.1)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Panicum virgatum(Pavirv00044632m.g)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Solub XP_006367061.1
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Fragaria vesca(XP004299601)
IVMGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>AtAH4(NP190378.2)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Capsella rubella(Caruvb10025825m)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>Brassica rapa(A02:26321550)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

>E_salsugineum(Thhalv10003603m)
IVLGFLAILIWKDFSPFMVLLAILNDGTMTISKDVRKPSPLPDWSKLEIFATGIVLGLYLLMLMTTVIFFW
FPFRFVGVRSTLSAIYLQVSIQLALIFVRSSWSFIERPGMLLGLQVIALIALIYANWGFISIQGIGWGAGSI
YSIFYPLD

95
Evolution of tonoplast P-ATPase transporters

YNIVFYIPLD

>Setaria_italica_ph5(Si034109m)
IVLGFLVLSIWEYDFPPFMVLHAILNDGTIMTISKDRVKPSPRPDRWKLNEIFATGVVMGTYLALVTLYFWAYVRTTT175
FEHSHFKVRSLAMYLVQISIALIFVTRSHGLSLFRPGALICAFVVAQLVATLTVYATIGFISAIGWRAGWIWL
YSLFYVFYPLD

>Brachypodium(Bradl1g72417.1)
IVLGFLVLSIWEYDFPPFMVLHAILNDGTIMTISKDRVKPSPPDSWKLNEIAAGGVGIVGYTLALVTLYFWAYVTTT
FEHSHFVNRSLAVYLQVI3ALIFVTRSRGISFLDPRGALLLCDFVAAVQLVATLVAYATVAFAIAAVGWRAGVWWL
YSLVSYAPLD

>Oryza_sativa(Oud_09682)
IVGFVLVLSIWEYDFPPFMVLHAILNDGTIMTISKDRVKPSPPDSWKLNEIAAGGVGIVGYTLALVTLYFWAYVTTT
FEHSHFKVRSLAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFILAVLTriAATVAYATISFIAGWGAGVWIL
YSLFYVFYPLD

>AtAHAl0(NP173169)
IVLGFLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVSTT
FEKFEHVKSISAAMYLVQISIALIFVTRSRGWSSFERPTGTLFFAFLAQALATTIAYANISFINTGWRAGWVIWL
YSLFYVFYPLD

>Capsella_rubella(Carubv10008225m)
IVLGFLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFILAQLVATLIAVYANISFINTGWRAGWVIWL
YSLFYVFYPLD

>Capsella_grandiflora_Cagra0824s0094
IVLGFLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFILAQLVATLIAVYANISFINTGWRAGWVIWL
YSLFYVFYPLD

>E_salsugineum(Thhalv10006710m)
IVLGFLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFILAQLVATLIAVYANISFINTGWRAGWVIWL
YSLFYVFYPLD

>Linum_usitatissimum(Lus10037466)
IVVAFSILALIWKYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
IFETCFQVSVLSAQLVQISIALIFVTRSRGSQTERPGTLLICAFVMAQLVATLIAVYHLSFISGIGWAGWVWVL
YCLFYVFYPLD

>Platycerium_biohyza_Spolz2G0033700
IVLGFLVLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFVMAQLVATLIAAHITFIIHIGWAGWVIWL
YSLFYVFYPLD

>Picea_abies_PAB00017861
IVLGFLVLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFVMAQLVATLIAAHITFIIHIGWAGWVIWL
YSLFYVFYPLD

>PhPH5(ABI98399)
IVLGFLVLLALIWEYDFPPFMVLHAILNDGTIMTISKDRVRPSPTPSWKLNQIFATGIVGYTLALVTLYFWIIVFTT
FEKFEHVKSISAAMYLVQISIALIFVTRSGQGLSFLERPGALLICAFVMAQLVATLIAAHITFIIHIGWAGWVIWL
YSLFYVFYPLD

96
Chapter 3

FFEHFHKLCSAVLVLSAVLQLSFALFVTRSQWFSFEMALLMFAFVAAQLVATLIAVAYHSFVRGIGWGWAGVIVL
YSLIFYPLD

>Nicotiana_plumbaginifolia(AAD46188)
JVLGFLMLAILIWEYDPFPPNMVLIALNLDGNTIMTSDKRVDKPSGPDPDWSWLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Caneaphora(Ce04_g00560)
JVLGFLMLAILIWKYDFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Phaseolus_vulgaris(ESW11044)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Glycine_max_(XP_003521833)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Musa_acuminata(XP_009386770.1)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Carica_papaya_ph5(supercontig_229:199945)22intrans?
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Phoenix_dactylifera_(XP008783452)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Persea_americana(FD503901)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Fragaria_vesca(XP004303775)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Rosa_hybrida(KU896052)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

>Malus_domestica(MDP0000303799)
JVLGFLMLAILIWEYDPFPPPMVLALNLDGNTIMTSDKRVDKPSPTPDWLSKLNEIFATGIVIGTYLALVSVLFWLADSTL
FFETTHFHVLSAIIYLQVSASQALIFVTRSQWFSGQLLMLMAFVAAQLVATLIAVAYAHIEFSGIGWGWAGVIVL
YSLIFYPLD

97
Evolution of tonoplast P-ATPase transporters

> *Malus domestica* (MDP0000290422)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVDTD
FFERTFNVRDLASAYLVQISIQLAVFTRSQWSFLERPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPDL

> *Populus trichocarpa* (XP006369068)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFETHFHVRSISAVYLQVISIQLAVFTRSQWSFLERPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPLD

> *Salix purpurea* (SapurV1A.0273s0110)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFETHFHVRSISAVYLQVISIQLAVFTRSQWSFLERPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPLD

> *Manihot esculenta* (cassava) scaffold05875:1310680
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDSTT
FFQTHFHVSTLISAYLVQISIQLAVFTRSQWSFMRPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIW
YSLVFYVPDL

> *Theobroma cacao* (EOY27421)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFETHFHVRSISAVYLQVISIQLAVFTRSQWSFMRPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPLD

**Note S2.** Protein alignment (in Fasta formate) from which the tree in Figure 5 was generated. The alignment was cured by G-boxes.

> *Methanoregula borkumensis* (YP001404446)
NAVDIAKESADIILLKNDLRILHDGVIEGRKTVGNTMKYILMNTSSNFVMFSVAMLPIQILLNNLLYSQVAIPTDNVA
DYPPKWDIGFKDFTILF

> *Laccaria bicolor* (XP_001873732)
SGTEIAKEAADVILLEKSLDVIAHGVLQGRQTTNTIKYIKMATSSNFVMFSVAMLPIQILLNNLLYSQVAIPFDNVDD

> *Eucalyptus grandis* (Eucgr.I02171.1)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFEDHFGVSSLSAYLVQISIQLAVFTRSQWSFMRPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPDL

> *Theobroma cacao* (EOY27421)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFETHFHVRSISAVYLQVISIQLAVFTRSQWSFMRPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPLD

> *Gossypium raimondii* (Gorni.011G053900.2)
IVLGFVLLLALIWEYDFPPFMLLIALILDGTVMLKKEIFATGIVGTLYLAVTLVLFYWVVIDTD
FFEDHFGVSSLSAYLVQISIQLAVFTRSQWSFMRPGLLKCMAFVAQVLAVTIAVYAHISFIRGIGWAGVWIWL
YSLVFYVPLD

> *Enterococcus gallinarum* (ERE43931)
GAVDIAREAADIILLEKSLMVLEQGVRQITQTVKKYIKMTASSNFVMFSVAMLPIQLLQLLQILYSAIPFDNVDD
C

chapter 3

eqpqrwpnlgrfvmvff

> salmonella enterica(wp_023203322)
gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

gavdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> e.coli(yp_672334)
gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> shigella flexneri(np_839642)gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> fusarium fujikuroi(cct70598)gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> claviceps purpurea(cece32637)gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> metarhizium acridum(efy89523)gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> polysphondylium pallidum(efz88373)gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> sphaignum falax_sphfalx0010s0101gevdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> picea abies

eqeqeqeqeqeqeqeqe

gavdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> pinus taeda_ph1_letiscaffold8448

eqeqeqeqeqeqeqeqe

gavdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> aquilegia coerulea(aquca00800152)

> amborella trichopoda(erm95363)

> petunia hybrida(gba34483)

> phoenix dactylifera(p008775198)

> sausaphaknlaadiiddlekdnlvlgvhrgrthngnsiyikyikmaassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe
gavdiareaddileekslmvlveegygrfrtsnmlkyikmtassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe

> amborella trichopoda(erm95363)

> petunia hybrida(gba34483)

> cyphrellasclamftswn

> persea americana(fd08035)

> sausaphaknlaadiiddlekdnlvlgvhrgrthngnsiyikyikmaassnfgnvsfvlmlplhliliqnnlllysqvaipfdnvde
eqeqeqeqeqeqeqeqe
Evolution of tonoplast P-ATPase transporters

>Carica_Papaya(CP00055G00600)
SgyvskdFadIjelleKldNvLVLVAYVQG VLTFGN TIKYIKAVSAVQG VLSISSLTLTVPQPLTQNFLSGQIAIPW DKMEE
DypqvrsKglpmFilwn

>Cicer_arietinum(XP_004504112)
SgvsvvAkDlAdIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IVIPWDKID
EypkhsresarGlsMfilwn

>Glycine_maximunm(XP_004504112)
SgyvskdFadIjelleKldNvLVLVAVHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Phaseolus_vulgaris_PH1(ESW31700)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Populus_trichocarpa(XP_002306511)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Manihot_esculenta(ME08265G01360)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Prunus_persica_PH1(EMJ09296)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Malus_domesticus(MD12G002970)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Eucalyptus_grandis(Eucgr.H04736)
SgyvskdFadIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Theobroma_cacao_PH1(EOX91997)_MGSLKFFRSFFRFRTQFDGLRYSNPVQLNSSNHHDQDFSNS
SasaSvIndAdIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Gossypium_raimondii(Gorai.007G235300)
SasaSvIndAdIjelleKldNvLVLVAGVEHG RLTFGNTMKKYKMSVIANLGVSISLLTSRQLTQNFIYQ8IAPWDKDE
EypkhsresarGlsMfilwn

>Medicago_trucatula(XP_00594954.1)
DatdaargasadivltepGlsViasvltSraIFQRMKNTTYYTAYSITIRVFMSGFSPFMVLiaINLMITISKDRVKP
Chapter 3

Note S3. Protein alignment (in Fasta format) from which the tree in Figure S2 was generated. The alignment was cured by G-boxes.

> Methanoregula_boone(YP001404446)
KESARESVELLSRIKILKGDNLVTRKTCSELGILSRLVNVDITITFARYTPVQKNRNMLKHHVVGGMDGGINDAPSIREADVGISV
ANAVDIASKADEILIKNLRLHIHDGVIEGRKTVSMTKYLMTLCTSNFFMNFSVAIFLPMLPQILLLNLNSQVAINPTDVAD
DYTKPKCWIDGIFKDTLTILGFLSISRTYFTQSWFTVESICSTQTLFIVIFVRPPSYRPSKLLVIGACIGAVGILFTAGGFGFV
QPPLFYAAALILIVGGYILLVELAKRW

> Laccaria_bicolor_Phi1(XP_001873732)_fungus
KSDASEIDRLAKLGQVVRILTDPAVAAKVRDLGFDAALERCIIFAKVSPHQKLQVVEGLRVVAFLDGVNDAALAIADVG
ISVSDGTEIAKEDVILEKSLDVIAHGVQGRQFTINTYIKIRMATSNFGNDFSLAWLPTQPLQQLFLBNFTSDQATIPW
VDPEYLPATTWNARSIRFMIQFLGPTSSVIDCTFAQTNWEEGATLQFIHHTLGTRKIPFlFQRSASVSVSVALTLSGVIFAIYPISA
LGMTRPKPEFYGFLVMVTGYAILVHVIVVIY

> Chlorocapsa_pulchella_Phi1(CCT70598)_fungus
KSDAOAQLKDVGQVQVQLGLDGVNDAALAIADVG
ISVSDGTEIAKEDVILEKSLDVIAHGVQGRQFTINTYIKIRMATSNFGNDFSLAWLPTQPLQQLFLBNFTSDQATIPW
VDPEYLPATTWNARSIRFMIQFLGPTSSVIDCTFAQTNWEEGATLQFIHHTLGTRKIPFlFQRSASVSVSVALTLSGVIFAIYPISA
LGMTRPKPEFYGFLVMVTGYAILVHVIVVIY

> Fusarium_fujikuroi_Phi1(CCT70598)_fungus
KSDAOAQLKDVGQVQVQLGLDGVNDAALAIADVG
ISVSDGTEIAKEDVILEKSLDVIAHGVQGRQFTINTYIKIRMATSNFGNDFSLAWLPTQPLQQLFLBNFTSDQATIPW
VDPEYLPATTWNARSIRFMIQFLGPTSSVIDCTFAQTNWEEGATLQFIHHTLGTRKIPFlFQRSASVSVSVALTLSGVIFAIYPISA
LGMTRPKPEFYGFLVMVTGYAILVHVIVVIY

> Chlorella_variabilis_XP_005848152
KETARQAVQQLQDKAVQLKVLTGDSLAVACSVCDSDFGCVCAVRASQMGKLTQPSMRKARVAALKHTVFLGDGVNDAALRA
DVGVTGDGSAADKNECAVKEKLISLVRVSRGRHTHNTIKYKMVASSNFNFNSILALWLPFPTMSQIALQNLMDISQAIWPDR
EEYLTQPRWRNMARDKLFJIVLGPQTSVATMTFSQATHWLQGQLTTQTMIVHLLRTRAKIPFISFQSRATVPLAALSJMIVMIGFVIT
WIPQRALNFQPSPTFVGLFAELLLLCEVQVQVMMY

> Dictyostelium_discoideum_Phi1(XP_646983)_slimemold
KSDAOAQLKDVGQVQVQLGLDGVNDAALAIADVG
ISVSDGTEIAKEDVILEKSLDVIAHGVQGRQFTINTYIKIRMATSNFGNDFSLAWLPTQPLQQLFLBNFTSDQATIPW
VDPEYLPATTWNARSIRFMIQFLGPTSSVIDCTFAQTNWEEGATLQFIHHTLGTRKIPFlFQRSASVSVSVALTLSGVIFAIYPISA
LGMTRPKPEFYGFLVMVTGYAILVHVIVVIY
Evolution of tonoplast P-ATPase transporters

DEEYLKIPHFWSVRLKFVMFLGPISSIFDFIVGQTFQGTWYVEGLIQTFQVIFHMRVTKVPIFQQRWSAQVTNVTIYAIAYTPSPYLLMEKLPAMYPGLASTVVFYFLTQTVKX

>Polypholium_pallidum_Ph1(EFA84373_slimemold)
KEFVKEALWRLAEKGVKVLGTVLPAVIKCEKVEGFHVEQVTKTVIATLQPKLQVQVLKHSKVHLGFLGIDSLALKAGADV

>Picea_abies
KESVKEALWRLAEKGVKVLGTVLPAVIKCEKVEGFHVEQVTKTVIATLQPKLQVQVLKHSKVHLGFLGIDSLALKAGADV

>Amborella_trichopoda_(ERM95363)

>petunia_pH1_gb(AHH24342)

>persea_americana_ph1(FD508035)

>Rhizophagus_irregularis_Ph1(ESA12740_fungus)

>Salmonella_enterica_ph1(WP_023203322)

>EcoliMgt(YP_672334)

>Rhizophasus_irregularis_Phi1(2003532222)_(fungus)

KPSKTPAQLQFYNKVEQKVLTVGDTSPACRCKVEEEINELIAESQFTAKLTPKANIKVLKHKHVFLGIDGIDAPAIRESDCGISVD

EGTDIAKESADIILLEKSMVLGVEHRGFTNMYKLMKSVANIGSVSITLLCYPLETPMQLUTQNYLNGQIAIPDKMEDCYYKVPQKRSLLKGLAMMTFSGNSALCSADIATLRSVAGFMELQMTLIIHRTEKIIPIFQEVASAWPVPCATILLISSIGIVIPPT

GKLMFGTALSYFGVLFLFGYFSGQIIKXKGY

>persea_americana_Phi1(FD508035)

KSHASKAQARLLWRLAEKGVKVLGTVLPAVIKCEKVEGFHVEQVTKTVIATLQPKLQVQVLKHSKVHLGFLGIDSLALKAGADV

GIVSDASGASKDFIANILLKDLVLTVLIAVEQGRTNMYKLMKSVANIGSISITLLCYPLETPMQLUTQNYLNGQIAIPDKMEDCYYKVPQKRSLLKGLAMMTFSGNSALCSADIATLRSVAGFMELQMTLIIHRTEKIIPIFQEVASAWPVPCATILLISSIGIVIPPT

TPGMGLNLPSYFGVLFVLGYYFSGQIIKXKGY

>Vitis_vinifera_Phi1(CBI41039)

KDSAKQALWRLAEKGVKVLGTVLPAVIKCEKVEGFHVEQVTKTVIATLQPKLQVQVLKHSKVHLGFLGIDSLALKAGADV

GIVSDASGASKDFIANILLKDLVLTVLIAVEQGRTNMYKLMKSVANIGSISITLLCYPLETPMQLUTQNYLNGQIAIPDKMEDCYYKVPQKRSLLKGLAMMTFSGNSALCSADIATLRSVAGFMELQMTLIIHRTEKIIPIFQEVASAWPVPCATILLISSIGIVIPPT

PKGMGLNLPSYFGVLFVLGYYFSGQIIKXKGY
Note S4. Protein alignment (in Fasta formate) from which the tree in Figure S3 was generated. The alignment was cured by G-boxes.
Evolution of tonoplast P-ATPase transporters

DSWLKNEIFATG
>Solanum tuberosum(XP006343689)
ENSAGNAAASALNPKTKVLPVDDLISIKLGDIVPADARLLDVDLPKFQGTVQGEICVFQVKVLTAIGFCICSLILLGIPIAMPTVLSMTAIGHSLAQGGAICTRMAITEAMAGMDVLCSDKTGLTLEVHLPFNPVPEKRTIAKAGEPIEICDP
PRHDASETRKALGAVKMITDQGLAGKETARLLGDVELEDIKAQDFGPVEHPKFYEEVKKLDQRKCIHMGTDVDGDNADALKK
ADIGIAVDAATDARSADIVLTPEGLSIVSAVLSRTAIFRQMRKNTYIAYVISITIRVMGMFSFPMVLIAILNDGTMITSKDRVKPSPD
DSWKLNEIFATG
>Fusarium fujikuroi(CCT70598)
EFQSGVAFIRLOSLQPVFRRPVGDIVLPVGAIPVPADCLILSLIRLSQSAWTGGESPEPGVKLMGTNIVSNQHVGVLVFKGQILNYTVMLIGFMVAVGLPVEMPAIVANLANLRAAAHQLSQKAIQVRDLSVQNLGMTVLCSDKTGLTTLTVAAIPNFEERRSACKGAAEVELRLCDPPDDAAQSIADKELGVIQLKTGDALPVPAVNQCRLEFDVVTRCTVFALKTPQNLAVGSLRKSCHVCGMLGDKGINDGALR
DADVIGSVDGAVKACDLITKEGLISIVRSVILGRTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEEPRRWRNARDILKFI
>Claviceps purpurea(CCE32637)
EYRSSVAFVKLQASVSTNLERVVRPGDVVLSPSVMAPDACLLFLRLISQTWTGTEDTPVFKMGTSVISGNGVALVFQGIRHVSYMLIGFMVAVGLPVEMPAIVANLANLARAYVLSMKAIKLDSVQNLGMTVLCSDKTGLTTVAAIPNFEERRSACKGAAEVELLRVCDPPDDAAEIAQLKSLGIEVKLGTDLTLPVNQCRLEFDVVTRCTVFALKTPQNLAVGSLRKSCHVCGMLGDKGINDGALR
CIARADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>Metsarhizium acridum(EFY95253)
EYRSSVAFVKLQASVSCNLDVVRRPGDVVLSPSVMAPDACLLFLRLISQTWTGTEDTPVFKMGTSVISGNGVALVFQGIRHVSYMLIGFMVAVGLPVEMPAIVANLANLARAYVLSMKAIKLDSVQNLGMTVLCSDKTGLTTVAAIPNFEERRSACKGAAEVELLRVCDPPDDAAEIAQLKSLGIEVKLGTDLTLPVNQCRLEFDVVTRCTVFALKTPQNLAVGSLRKSCHVCGMLGDKGINDGALR
CIARADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>Laccaria bicolor(CX01873732)
ELKSVQSLAQLNLVVLRPVGDLVVLSGVDVFPGDCVFLGTVAQASTLGELMPVEKLASTSVAGTSGRAMVMQVGVRRVSYLLLFAMVAVGTPPEMLPMMVAANLALSARVKKVIVKRDFIAIQSLGAVNILCSDKTGTLTVAAIFPNNFERRSACKGAAEVELLRVCDPPDDAAEIAQLKSLGIEVKLGTDLTLPVNQCRLEFDVVTRCTVFALKTPQNLAVGSLRKSCHVCGMLGDKGINDGALR
CIALRAADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>Amphibius tricophodes(JE959269)
EHSSKAKAMKSLNSPAPSIVQVPDGIDILFFSSGDLFGPDVRLLDDLSVLVSQSSLTGESMTKFMGTSGTVIASSGTGFLEFKVRRVSYALICMVAMFLGTVPLPPLMLNLKALKELAMLADKCRIVSKOSVAIQNMAGMDIICLDKTLTLTNIEPSFARRMSSGVSLOGEELVIRCPDCPPDVKGLQALR{}{}{}{}
EADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>KDMPHPQWSPKGIIMF
>Spirolella polymorpha
EYKSSRAAMKSLNSPAPSIVQVPDGIDILFFSSGDLFGPDVRLLDDLSVLVSQSSLTGESMTKFMGTSGTVIASSGTGFLEFKVRRVSYALICMVAMFLGTVPLPPLMLNLKALKELAMLADKCRIVSKOSVAIQNMAGMDIICLDKTLTLTNIEPSFARRMSSGVSLOGEELVIRCPDCPPDVKGLQALR{}{}{}{}
EADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>Medusophycus calcaratus(P009399682)
EVSSKAAAMVSIPFIVQVVPDIILFSSGDLFGPDVRLLDDLSVLVSQSSLTGESMTKFMGTSGTVIASSGTGFLEFKVRRVSYALICMVAMFLGTVPLPPLMLNLKALKELAMLADKCRIVSKOSVAIQNMAGMDIICLDKTLTLTNIEPSFARRMSSGVSLOGEELVIRCPDCPPDVKGLQALR{}{}{}{}
EADVIGSVDGASVAKCDADLVTEKLGLIHIMVASVTTTRGTHGNTIKYIKMVASSNFNGVFSILMTSQAQLNLYIDISQAIHPDWRVDEPWSKWKTWDLLRFV
>Acquilegia coerulea(Accuca_008_00152)
>Campanula coerulea
Evolution of tonoplast P-ATPase transporters

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eucalyptus grandis (Eucgr.H04736)</td>
<td>EHSSSKAKMSEFVKCPVKVRPVGIDIEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Populus trichocarpa (XP002306511)</td>
<td>EYSSSKAKMSEFVRPCPKVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Manihot esculenta (ME08265G01360)</td>
<td>ENSSSKAKMSEFVKCPVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Ricinus communis (XP_002533565)</td>
<td>ENSSSKAKMSEFVKCPVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Theobroma cacao (EOX91997)</td>
<td>EYSSSKAKMSEFVRPCPKVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Gossypium raimondii (Gorai.007G235300)</td>
<td>EYSSSKAKMSEFVKCPVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Vitis vinifera (XP004288155)</td>
<td>EYSSSKAKMSEFVRPCPKVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
<tr>
<td>Fragaria vesca (XP004288155)</td>
<td>EYSSSKAKMSEFVRPCPKVKVRPVRPGDIFEPFGDLFDVRLKLHKLVSQSSLTGESWTEKFMTGNVSVSGLTGFEENGIRQISYVG</td>
<td>DAANIGISVDGVSVAKDFADIILLEKDLNLVLVGVEQGRLTFGNTIKYIKASVIANVGSILLLTPKQLLTQNFLSSIGQIAIPWDKMEEDPQWIKPLMFN</td>
</tr>
</tbody>
</table>

106
Chapter 3

> Rosa_hybrida_PH1

EYGGSKAAAMELSFVCPRVCQVPGDIIFPDLGDPGVRLKLHVLVSQASLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Malus_domesticum(MD15G024990)

EYGGSKAAAMELSFVCPRVCQVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Prunus_persica(EMJ09296)

EYGGSKAAAMELSFVCPRVCQVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Methanoregula_boone(YP001404446)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Chlorella.variabilis(XP005848152)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Dictyostelium_discoideum(XP646983)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Polysphondylium_pallidum(EFA84373)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Rhizophagus_irregularis(ESA12740)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Enterococcus_dispar(WP016171755)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Polyphondium_discoides(IPM646807)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Rhizophagus_irregularis(ESA12740)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF

> Enterococcus_dispar(WP016171755)

EYKSAFQAAELLRKKIITNASVLRVPGDIVIFEPDLGDPGVRLKLHVLVSQSSLTGESWTTEKGMGNTVNSGSGSLVFEDGVRRIYSYVPVAMVSAALTQPMVLPVNTSLAKGALAMARDCICKLSAIRNMSMIDLAMCIDLKTGTLKLPDIFEDRRVSIGKALEEMVKCDPPKDSQAKLWRLAEKVGKAVLVTGDSLSSLIRVCVEFGHETVKTVLALTQPLKLVRQSLQTIGHIVGFLGVDGNVSLALDAHVGGSDAGVAKDFADILLEKLVDNLIAGVEHGRLTEFGNTMKYIKMVSIANLGSVSLILLTARQLLTQNFSLYSVQQAIPWDKMDDEPDQPKWKSLPKLPIF
Evolution of tonoplast P-ATPase transporters

CDPAKESAIKSLHEHGTVKVLTDNAIVAKVCDVDGLMGQVDETNLFAKLNPMQKSRLIEVIQKKGHTVGFMGDGINDAPA
LRKADVGSVDADITKADSAIIKLSLELNVLGTIEGVRERVSMMKVKITFSSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDNV
DEEPYRVEWIDFLKFT

>Carnobacterium_maltaromaticum(YP006992121)
EYRSQASLALKLETTLETACTVRPQDGVLSVTGDMPADARLIKDLFVNPQSSLGSEMPVEFKMGTVDVLSQGRVVFDRGVTNVS
L1RIFMVAGLTPEMPMLTSAIKLAIAMSKVKEILNQAIQLGNDMLICTDKTGTKTIDIFIPFDSRRLTVKGAEEVMKICD
PAKKSITAIKLHEHGTVKLTDNIEVSKKCDVGLTLASTRNLFAKLAINTQKARJIALLQQAKHTVGFMGDGINDAPALRKA
DVGISVDAADITKEDSAIIKLSLELNVLGEDIIGRVSFGNFMMKVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDREE
PAKWDTSNLKMT

>Enterococcus_malodoratus(WP010739787)
SVRSKNAEKLSLKVLKTVATTAAVRVCGDLKLSADGMDAPAIRLKDILFSQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPVTTLNKLQATGMNAGINTNQLNMLCNDTFKLTGTLTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Enterococcus_raffinosus(WP010746437)
SVRSKNAEKLSLKVLKTVATVATVVRVCGDLKLSADGMDAPAIRLKDILFSQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPVTTLNKLQATGMNAGINTNQLNMLCNDTFKLTGTLTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Lactococcus_lactis(YP006999610)
ETRSDNAENLKLKMTTITTVHRLVGDIELLSADGMDAPAIRLKDILFSQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SWVLIRFMVAGLTPEMPMLPMTSNLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SWVLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Eubacterium_limosum(YP003960913)
EQKSDRAAERLGEMVETTAAVDRVVGDIIYLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SWVLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Rahnella_aquatilis(YP005198626)
EARSNKAADALKAMVSNTATVIRVPGDIIKLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Serratia_proteamaculans(YP001476771)
EARSNKAADALKAMVSNTATVIRVPGDIIKLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>Serratia_liquefaciens(YP008228385)
EARSNRAADALKAMVSNTATVIRVPGDIIKLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>DNPVDDQPQRWNNSGIDFRGFM

>Serratia_proteamaculans(YP001476771)
EARSNKAADALKAMVSNTATVIRVPGDIIKLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

>DNPVDDQPQRWNNSGIDFRGFM

>Serratia_liquefaciens(YP008228385)
EARSNKAADALKAMVSNTATVIRVPGDIIKLKLAAGDMIPADVLRILKDLFISQALTGESEPVEFKMGNSIIGAYIVGFVEKGVNVS
SVWLIRFMVAGLTPEMPMLPMIASKAKAVLSKVGEILNQAGMDVLICNTDKTGTKTKEIPDFDQFRRMSVKGAAEMLIDCD
PKETKTKQALKEHGTVKLTDNLVSKQVQVGELAIENVFKLNPQOKKSKFSLRLANTHGVGFMGDGINDAPAMAKA
ADVIGISVDAVDAKEDSAVIIKLSLELNVLGDIEGRSFVSNMGMVKMTISSNFGNFVSILMLISQILVQNLIYMDAQLTIPWDMDK
KKWTASSIGKFM

108
Chapter 3

Yersinia enterocolitica (CCQ38899)

```
Yersinia enterocolitica (CCQ38899)
```

Yersinia ruckeri (WP004719319)

```
Yersinia ruckeri (WP004719319)
```

Dickeya zeae (WP02363915)

```
Dickeya zeae (WP02363915)
```

Dickeya dianthicola (WP024104954)

```
Dickeya dianthicola (WP024104954)
```

Cedecea davisae (WP016538339)

```
Cedecea davisae (WP016538339)
```

Kosakonia radicincitans (WP007372828)

```
Kosakonia radicincitans (WP007372828)
```

Escherichia albertii (WP_000471895)

```
Escherichia albertii (WP_000471895)
```

Klebsiella pneumoniae (CDK77607)

```
Klebsiella pneumoniae (CDK77607)
```

Shigella dysenteriae (WP_000471869)

```
Shigella dysenteriae (WP_000471869)
```
Evolution of tonoplast P-ATPase transporters

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Shigella_sonnei(YP313138)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Escherichia_coli_MgtA(YP672334)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Shigella_flexneri(NP_839642)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Citrobacter_rodentium(YP003366794)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Citrobacter_freundii(WP003839602)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Salmonella_enterica(WP023203322)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Enterobacter_cloacae(WP023305328)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM

Enterobacter_hormaechei(WP017694637)

VCDPPKETTAPALKASGITVKILTGDSELVAAKCHEVGLANLAQRTTLFARLTPMHKERIVTLKREGHVVGFMGDGINDAPA
LRAADIGISVDADIAEAADIIELKSLMVEEGVIEGRRTFANMLKYIKMTASSNFNVFSVLSMLPHLIIQNLLYDVSQVAIPFDN
VDDEPQRWNPDAILGRFM
Chapter 3

Yokenella_regensburgei (WP003839602)

EARSTKAADALKAMVSNTATVLRVPGDIVKLSAGDMIPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFD
NVDEPQRWNPAPDLGRFM

Raoultella_ornithinolytica (YP007875494)

EARSTKAADALKAMVSNTATVLRVPGDIVKLSAGDMIPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFD
NVDEPQRWNPAPDLGRFM

Klebsiella_variicola (WP022065403)

EARSTKAADALKAMVSNTATVLRVPGDIVKLSAGDMIPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFD
NVDEPQRWNPAPDLGRFM

Enterococcus_gallinarum (ERE43931)

EARSTKAADALKAMVSNTATVLRVPGDIVKLSAGDMIPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFD
NVDEPQRWNPAPDLGRFM

Bacillus_cereus (WP_000933490)

EFRLKAFQAKADKLKAMVRTTASVFRVPGDIISLSAGDIVPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFDN
VDEPQRWNPAPDLGRFM

Bacillus_thuringiensis (YP006830256)

EFRLKAFQAKADKLKAMVRTTASVFRVPGDIISLSAGDIVPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFDN
VDEPQRWNPAPDLGRFM

Paenibacillus_dendritiformis (WP006677380)

EFRLKAFQAKADKLKAMVRTTASVFRVPGDIISLSAGDIVPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFDN
VDEPQRWNPAPDLGRFM

Paenibacillus_alvei (WP005543552)

EFRLKAFQAKADKLKAMVRTTASVFRVPGDIISLSAGDIVPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFDN
VDEPQRWNPAPDLGRFM

Brevibacillus_laterosporus (WP003336336)

EFRLKAFQAKADKLKAMVRTTASVFRVPGDIISLSAGDIVPADLRILFRDLFVAQASLTSGESLPVEKFMTGVNSVSAQAMVFQKGISRVS
SMIIIRFVMVGLPLMVPMTSTLARAVGKLQKIVVHKHDAIQNFAGMDLCTDKTLTVKIDEIPFDFERRMSVKGALQEIL
NVSDPPKETTAPALAKAGSTIKVTLGDSEVaAAVKCNEVDVGLALLARITLLFRALTPMKIERIVVTKRKEHGFVGMGIDGINDAP
ALRAADIGISVDAVIAREAADILIEKLMLVEEGVIEGRRTFANMLKYIKMTASSSNFGNVFSVLMLPHLLIQNLLYDSVQAIPFDN
VDEPQRWNPAPDLGRFM
Evolution of tonoplast P-ATPase transporters

CDPPKESASAMQKHLAEGVNVKVTGDNAVVRKCTDLGLADLAEEKTTYFAKLNLPLQKARIVRVLQSGKHTGVFGMVDGINA
ASLREADVSVDAVIAKESADIIKLESLVLQEDGILEGRRTFIGNIKYIKMTASSNFVGVFSLMPLHIIQNLYFDISQISIPWDNM
DKPARKWDAKSVGRF

> Aneurinibacillus_aneurinilyticus (WP021620844)
EFSRMKTAEKLKAMVKTITT0TVIRPGDIIHLASAGDLVAPAVRLLKDLFVSEALTGEALPKVEKYMNTIVSGSATAIVDFKGNVNSVT
VLIRFMVAGLTPLEMVVTANLAKGAVVMARNKVMIKRLHAIQNGFAMDILDTCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
CDPPKDSAAAIKALRENGVNIKLTGDNETVRKVCVQDVGADLAERTTVFAKLNLPLQKARIVRVLQSKGHTVGFMVDGINA
ASLREADVSVDAVIAKESADIIKLESLVLQEDGQIEGRRTFIGNIKYIKMTASSNFVGVFSLMPLHIIQNLYFDISQISIPWDNM
DKPARKWDKADSRIFM

> Paenibacillus_elgii (WP010496117)
EYRSITAEKLKAMVKTTTATTVRSVGDIHIHASAGDLVAPAVRLLKDLFVSEALTGEALPKVEKYMNTISGTATATIVDFKGNVHNTVF
VLIRFIVAGLTPLEMVVTANLAKGAVVMARNKVMIKRLHAAIQNGFAMNCLTDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
CDPPKETAAAIKALQEGVVEKVTGDNAGAVAKVCQDVGADLAERTTVFAKLNLPLQKARIVRVLQSKGHTVGFMVDGINA
ASLREADVSVDAVIAKESADIIKLESLVLQEDGQIEGRRTFIGNIKYIKMTASSNFVGVFSLMPLHIIQNLYFDISQISIPWDNM
DKPARKWDKADSRIFM

>Caldanaerobacter_subterraneus (WP022588790)
EYRSNIEAEKLKAMVHTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIGVVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Caldicellulosiruptor_kronotskyensis (YP004023281)
EYRSNIEAEKLKAMVHTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIGVVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Thermoanaerobacter_pseudethanolicus (YP001664222)
EYRSNVEAEKLKALVYTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIAYVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Thermoanaerobacter_italicus (YP003477957)
EYRSNVEAEKLKALVYTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIAYVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Clostridium_botulinum (YP001788195)
ELKSNKAAEKLKQLVRTTAAAVYRPDIHLAAGMDPAVLIRYKLINQATLTGESEPKVEKFMGTNSGSAIAFVEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Clostridium_acidurici (YP006789559)
EYRSNVEAEKLKAMVHTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIAYVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Clostridium_butyricum (YP001788189)
EYRSNVEAEKLKAMVHTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIAYVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

> Clostridium_acyetobutylicum (YP001788189)
EYRSNVEAEKLKAMVHTTAAAVKRPDIHLAAGMDPAVLIRVKDLYNQATLTGESEPKVEKFMGTVSVGSAIAYVFEGIDNVD
SLIKFMVAGLTPLEMVVTANLAKGAVMKHKTVIKRLDAIQNFAMLDCTKTLTGVDEIPFDFRSMVSGKAGEEILSV
ICDPPKESAPYAIKALKHEGVKVTGDNIFTKVCVEGGLAIAEERTIATFLPMQAIKIKRLTGGVFGFLGIDNAPAM
READVSVDAVIAKESADIIKLESLVLEEGVVEGEKRGFIGNIMKYYAITASSNFVGVFSLMPLAQLQLNLTYDLSMAPWDRM
DREPRKWDANAIGNFM

112