A Proteomics Dissection of Arabidopsis thaliana Vacuoles Isolated from Cell Culture.

Michel Jaquinod, Florent Villiers, Sylvie Kieffer-Jaquinod, Véronique Hugouvieux, Christophe Bruley, Jérôme Garin, Jacques Bourguignon

This manuscript has been published online in Molecular and Cellular Proteomics (Dec 2006) - doi:10.1074/mcp.M600250-MCP200

A proteomic dissection of *Arabidopsis thaliana* vacuoles isolated from cell culture

Running title: A proteomic dissection of *Arabidopsis thaliana* vacuoles

Michel Jaquinod‡§¶, Florent Villiers‖§, Sylvie Kieffer-Jaquinod‡, Véronique Hugouvieux‖, Christophe Bruley‡, Jérôme Garin‡ and Jacques Bourguignon‖¶

‡CEA, DSV, DRDC, Laboratoire de Chimie des Protéines, Grenoble, F-38054, France, INSERM, ERM 0201, Grenoble, F-38054, France, Université Joseph Fourier, ERM 0201, Grenoble, F-38054, France.

‖CEA, DSV, DRDC, Laboratoire de Physiologie Cellulaire Végétale, UMR 5168, CEA/CNRS/Université Joseph Fourier/INRA, CEA-Grenoble, 17 avenue des martyrs, 38054 Grenoble cedex 9, France.

This work was funded by the CEA, CNRS, INRA and INSERM scientific program: « Toxicologie Nucléaire Environnementale »

§ These authors contributed equally to this work.

¶ Authors to whom correspondence should be addressed

- Michel Jaquinod, Laboratoire de Chimie des Protéines, DRDC, CEA-Grenoble, 17 avenue des martyrs, 38054 Grenoble cedex 9, France. (e mail : mjaquinod@cea.fr)
- Jacques Bourguignon, Laboratoire de Physiologie Cellulaire Végétale, UMR 5168, CEA/CNRS/Université Joseph Fourier/INRA, DRDC, CEA-Grenoble (e mail : jacques.bourguignon@cea.fr)

ABSTRACT

To better understand the mechanisms governing cellular traffic, storage of various metabolites and their ultimate degradation, *Arabidopsis thaliana* vacuoles proteomes were established. To this aim, a procedure was developed to prepare highly purified vacuoles from protoplasts isolated from *Arabidopsis* cell cultures using Ficoll density gradients. Based on the specific activity of the vacuolar marker α-mannosidase, the enrichment factor of the vacuoles was estimated at approximately 42 fold with an average yield of 2.1%. Absence of significant contamination by other cellular compartments was validated by western blot using antibodies raised against specific markers of chloroplasts, mitochondria, plasma membrane and endoplasmic reticulum. Based on these results, vacuole preparations showed the necessary degree of purity for proteomic study. Therefore, a proteomic approach was developed in order to identify the protein components present in both the membrane and soluble fractions of the *Arabidopsis* cell vacuoles. This approach includes: (i) a mild oxidation step leading to the transformation of cysteine residues into cysteic acid and methionine to methionine sulfoxide, (ii) an in-solution proteolytic digestion of very hydrophobic proteins, (iii) a pre-fractionation of proteins by short migration on SDS-PAGE followed by analysis by liquid chromatography coupled to tandem mass spectrometry. This procedure allowed the identification of more than 650 proteins, 2/3 of which copurify with the membrane hydrophobic fraction and 1/3 with the soluble fraction. Among the 416 proteins identified from the membrane fraction, 195 were considered integral membrane proteins based on the presence of one or more predicted transmembrane domains, and 110 transporters and related proteins were identified (91 putative transporters and 19 proteins related to the V-ATPase pump). With regard to function, about 20% of the proteins identified were previously known to be associated with vacuolar activities. The proteins identified are involved in: ion and metabolite transport (26%), stress response (9%), signal transduction (7%), metabolism (6%) or have been described to be involved in typical vacuolar activities, such as protein- and sugar-hydrolysis. The sub-cellular localization of several putative vacuolar proteins was confirmed by transient expression of GFP-fusion constructs.

Abbreviation

AAAP	Amino Acid / Auxin Permease
ABA	Abscissic Acid
ABC	ATP binding cassette
ABRC	Arabidopsis Biological Resource Center
ACAc	calcium transporting ATPase
AVP	H^+-pumping pyrophosphatase
CaMV	Cauliflower Mosaïc Virus
CAT	Cationic Amino acid Transporter
CAX	Ca^{2+}/H^+ antiporters
CCD1	Carotenoid Cleavage enzyme D1
COPT	Copper Transporter
DIM/DWF	Diminuto / Dwarf
DMT	Drug / Metabolite Transporter
DRM	Detergent Resistant Microdomains
ERp57	protein disulfide isomerase of the Endoplasmic Reticulum
GSTF	Glutathione S-transferase, type F
H^+-ATPase	vacuolar-type H^+-pumping ATP hydrolase
H^+-PPase	H^+-pumping pyrophosphatase
id	inner diameter
IRT	Iron Transporter
KUP	K^+ Uptake Permease
LAT	L-type Amino acid Transporter
LHC	Light Harvesting Complex
MATE	Multidrug and Toxin Extrusion
MDR	Multidrug Resistance Protein
MEC	Multifunctional Extracellular Protein
MFS	Major Facilitator Superfamily protein
MHX	Mg^{2+}/H^+ antiporters
MOP	Multidrug / Oligosaccharidyl-lipid / Polysaccharide exporter

MRP	Multidrug resistance Associated Protein
MTP	Microsomal Triglyceride transfer Protein
MudPiT	Multidimensional Protein Identification Technology
NAP	Non-intrinsic ABC Protein
NCED1	Neoxanthin Cleavage Enzyme D1 (=9-cis-epoxycarotenoid dioxygenase)
NHX	Na^+/H^+ antiporter
NPC1	Niemann-Pick C1 protein
NRAMP	Natural Resistance-Associated Macrophage Protein
od	outer diameter
OEP	Outer Envelop Protein
OPT	Oligopeptide Transporter
p22HBP	Heme Binding Protein
PDR	Pleiotropic Drug Resistance
PMA	Plasma Membrane H^+-ATPase
POT	Proton dependant Oligopeptide Transporter
PSV	Protein Storage vacuole
PTR2B	Peptide Transporter 2B
RND	Resistance / Nodulation / Division
RT	Room Temperature
SPFH	Stomatin Prohibitin Flotilin Hbc
TAP	Transporter associated with Antigen Processing
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TIP	Tonoplast Intrinsic Protein
TOM	Translocase of the mitochondrial Outer Membrane
TRH	Tiny Root Hair protein (AtTRH1 = KUP4)
TWD	Twisted Dwarf protein
VHA	vacuolar-type H^+-pumping ATP hydrolase
YS	Yellow Stripe
YSL	Yellow Stripe-Like
ZAT	Zinc Transporter of Arabidopsis
ZIP	Zinc / Iron Permease

INTRODUCTION

Plant cell vacuoles are multifunctional organelles that play a key role in plant physiology. Vacuoles are considered as the main storage site in plant cells and can occupy up to 90% of the cellular volume in mature cells. Vacuoles are involved in the storage of a plethora of metabolites essential for plant function, including water, inorganic anions and cations, organic and amino acids, sugars, proteins and a diverse group of soluble and insoluble compounds including anthocyanin and anthoxanthin pigments (1). Vacuoles are also involved in the sequestration of toxic molecules including metal ions, drugs, xenobiotic molecules. They are important for maintenance of turgor pressure, digestion of cytoplasmic constituents, pH regulation and ion homeostasis. The vacuole dynamically changes its function and shape according to developmental and physiological conditions (2). In addition to the large central vacuole present in mature vegetative cells that are considered as lytic vacuoles, plant cells also contain Protein Storage Vacuoles (PSV) (3). Lytic vacuoles are analogues of the yeast vacuole or animal lysosome. PSV are particularly prominent in developing seeds (4). PSV contain vacuolar storage proteins to be used for anabolism during seedling growth. The function of the different vacuoles seems to be correlated with the presence of specific tonoplast intrinsic protein (TIP) isoforms (5,6). Current knowledge of the cellular traffic of higher eukaryotes indicates that vacuole biogenesis is closely related to the traffic of proteins resident in these compartments (7-10). Resident vacuolar proteins, as well as proteins intented for degradation, are delivered to the vacuole via the secretory pathway which includes the biosynthetic, autophagic and endocytotic transport routes (2,10-14). Although some aspects have been studied in detail (13,15), many questions remain unanswered concerning autophagy, transport and fusion with small vacuoles. The analogy with the processes set up by yeast or higher eukaryotes seems to indicate that these mechanisms may be common to all kingdoms (16).

Most of the compounds present in vacuoles have to be transported in a passive or active manner across the tonoplast (the vacuolar membrane) for storage or degradation, but they also need to be exported in response to plant cell demands. Surprisingly, the number of transporters that have been identified on the tonoplast is quite low (17). Two proton pumps known as primary active H^+ transport systems are present in this membrane: the vacuolar-type H^+-pumping ATP hydrolase (H^+-ATPase, VHA) (18,19) and the H^+-pumping pyrophosphatase (pyrophosphate-

energised H^+-PPase, AVP1) (17,20). They are responsible for the acidification of the vacuolar lumen, thus creating proton concentration and electrical gradients across the tonoplast. It was recently shown that H^+-PPase also controls auxin transport and consequently auxin-dependent development (21). The tonoplast also contains secondary active transporters energised by the proton motive force and several other pumps (1,17). A Na^+/H^+ antiporter (AtNHX1) is present in the tonoplast and mediates Na^+ sequestration in the vacuole. This transporter contributes to the plant salt tolerance of transgenic *Arabidopsis* overexpressing AtNHX1 (22,23) and was recently shown to be regulated by calmodulin (24). The free cytosolic Ca^{2+} concentration must also be strictly regulated as it controls many essential cellular responses (25). The tonoplast contains Ca^{2+}/H^+ antiporters (CAX1 and CAX2) (26-28), that are responsible, in conjunction with a Ca^{2+} pump (P_{2B}-type ATPase, ACA4) (29), for the sequestration of Ca^{2+} in the vacuolar sap (30). It was recently proposed that CAX1 regulates several plant processes including ion homeostasis, development and hormonal responses (28). Other metal transporters have also been identified in the tonoplast. These include: an Mg^{2+}/H^+ exchanger (AtMHX); a cation diffusion facilitator family member MTP1 (ZAT) and the AtNRAMP3 and AtNRAMP4 transporters. AtMHX functions as an electrogenic exchanger of protons with Mg^{2+} and Zn^{2+} ions (31). By sequestering excess cellular Zn in the *Arabidopsis thaliana* vacuole, MTP1 is involved in Zn homeostasis and detoxification (32-34). This transporter is probably involved in Zn tolerance in the Zn hyperaccumulator *Arabidopsis halleri* (35). AtNRAMP3 and AtNRAMP4 have recently been shown to be present in the tonoplast and to participate specifically in Fe mobilization from vacuolar metal stores during seed germination (36,37). Some ATP binding cassette (ABC) transporters are also present in the tonoplast, such as MRP2 that has been shown to be not only competent in the transport of glutathione conjugates but also glucuronate conjugates following its heterologous expression in yeast (38). AtMRP1 is also localized to the vacuolar membrane of *Arabidopsis* and interacts with an immunophilin-like protein (TWD1) through a calmodulin-binding domain present in the C-terminus of AtMRP1 (39).

Key steps in understanding the transport process of substrates to the vacuole and their storage depends on the identification of additional membrane proteins. Recently, proteomic analyses of the tonoplast have been published (40-42). Shimaoka *et al*. (40) identified a large number of mostly soluble proteins within their vacuolar fractions. Forty two of the 163 proteins were annotated with one or more transmembrane domains and 39 proteins were predicted to have

more than two transmembrane domains, 17 of which were putative transporters. Szponarski *et al.*'s procedure (41) allowed characterization of 70 proteins from an *Arabidopsis* tonoplast-enriched fraction, including only a small number of transporters. The most complete study published so far identified 402 proteins (42). However, almost half of the proteins listed were identified by a single peptide hit which is often insufficient for certain identification. From these proteins, 29 were putative or known transporters and 17 were related to the H^+-ATPase complex. Taken together, all these previously published results indicated the need to extend the knowledge of the vacuolar proteome of *Arabidopsis*.

In the present study, intact vacuoles were isolated from *Arabidopsis* suspension cells. Potential cross contaminations were examined by western blot analyses and the quality of the vacuole preparations led us to a proteomic investigation. A proteomic approach was developed in order to identify the protein components present in the membrane and the soluble fractions of the *Arabidopsis* cell vacuoles. This approach includes: a mild oxidation step leading to the transformation of the cysteinyl residues into cysteic acid and methionine into methionine sulfoxide, which facilitates peptide assignment; an in-solution proteolytic digestion of membrane proteins; and/or a pre-fractionation of proteins by SDS-PAGE. Peptides were identified using liquid chromatography coupled to tandem mass spectrometry. The combined results of these approaches spectrometry allowed the identification of over 650 proteins. Among these, 195 were considered as integral membrane proteins based on the presence of one or more predicted transmembrane domains and 91 transporters were identified. The sub-cellular localization of several putative vacuolar proteins was confirmed by transient expression in *Arabidopsis* protoplasts overexpressing GFP-fusion proteins.

EXPERIMENTAL PROCEDURES

Cell material and growth conditions- Arabidopsis thaliana cells (var. Columbia) were cultivated at 22°C, under permanent light (80 µmol photons.m^{-2}.s^{-1}) and shaking (135 rpm) in a Murashige & Skoog medium (MS basal medium, Sigma, ref. M5519, pH 5.5) containing 88 mM sucrose and 2,4-dichlorophenoxyacetic acid (0.02 g.L-1). Every seven days, 100 mL of fresh medium were inoculated with an aliquot of the culture of 7-day-old cells (5 ml of packed cells after centrifugation at 110g for 5 min).

Vacuole isolation- Protoplasts were obtained after digestion of 5-day-old cells in 0.6 M mannitol, 2% [w/v] cellulase, 0.5% [w/v] pectolyase, 25 mM MES, pH 5.5. After 2 hours, protoplasts were filtered through a 50-µm nylon net and the suspension was centrifuged for 1 min at 1270 g. The protoplast pellet was washed with a rinsing medium (0.7 M mannitol, 10 mM Tris, 15 mM MES, pH 7.0) and adjusted to a final concentration of 52 (± 4)x10^6 protoplasts.mL^{-1}. Vacuoles were then purified following a protocol adapted from Frangne *et al*. (43). The protoplast suspension was diluted 4-fold in a lysis medium pre-warmed to 42°C (medium A: 0.2 M mannitol, 10% [w/v] Ficoll 400, 20 mM EDTA, 5 mM HEPES-KOH, pH 7.5). After 15 min incubation, vacuoles were isolated using a three-step gradient. The suspension of protoplasts was loaded at the bottom of a centrifuge tube and covered with two volumes of a mix (1: 2; medium A: medium B) (medium B: 0.4 M betaine, 30 mM KCl, 20 mM HEPES-KOH, pH 7.5) and one volume of medium B. The gradient was then centrifuged at 1500 *g* for 20 min and vacuoles were collected at the interface of the first and second layers corresponding to 0% and 3.3% Ficoll. The vacuoles were then concentrated by centrifugation (1800 *g*, 20 min).

Enzyme activities and western blot analyses- Purification yield of vacuole preparations was followed by enzymatic assay of the vacuolar specific marker, α-mannosidase, according to Boller & Kende (44). Protein (10-100µg) was added to a medium containing: 50 mM citric acid – NaOH (pH 4.5) and 1 mM p-nitrophenyl-α-D-mannopyrannoside. After 20, 40 and 60 min of incubation at 37°C, the reaction was stopped by adding 0.8 mL of 1 M Na_2CO_3 per 0.5 mL

assay medium. After centrifugation (16,000 *g*, 15 min), enzyme activity was avaluated by detecting the product, p-nitrophenol, at 405nm ($\varepsilon_{pNP(405nm)}$ = 18.5 $mM^{-1}.cm^{-1}$).

Purity of the vacuole was controlled by western blot analyses. Fifteen micrograms of proteins from protoplast or vacuole extracts were loaded on a 12% SDS-PAGE and, after migration at 200 V, transferred to a nitrocellulose membrane. The presence of different characteristic proteins was assessed using western blot analyses and primary polyclonal antibodies, raised against: the outer envelop protein 21 (OEP21, (45)) and the light harvesting complex b (LHC; O. Vallon, IBPC-Paris, France) (plastid markers); the preprotein translocase of the mitochondrial outer membrane (TOM40, channel forming subunit; (46)); the plasma membrane P-type H^+-ATPase (PMA2; (47)) and the HDEL domain of the endoplasmic reticulum (ER) proteins (48,49). Antibodies raised against tonoplastic markers such as the tobacco alpha tonoplast intrinsic protein (α-TIP; (50)) and the cauliflower gamma tonoplast intrinsic protein (γ-TIP; (51)) were used to control for vacuole-enrichment.

Vacuolar protein preparation and trypsin digestion for mass spectrometry analyses- Vacuole were disrupted by a freeze / thaw cycle in nitrogen. The suspension was then centrifuged for 60 min at 100,000 *g.* The pellet containing the membranes was resuspended in 100 µL of 100 mM ammonium bicarbonate (pH 8.2). The mix was then adjusted to 1500 µL with 0.55 M NaCl and incubated at 4°C for 60 min under permanent shaking.

For in-gel digestion, 20µg of proteins from either the soluble or the membrane fractions were separated by a short electrophoresis (2.5 cm) by 12 % SDS-PAGE. After coloration with Coomassie Blue R 250, the gel was cut into 1.5 mm slices. Each band was further cut and washed twice in 100µL destaining solution (50 mM NH_4HCO_3/ CH_3CN) 50/50 (v/v) at room temperature for 30 min, before dehydratation with 100 µL of pure ACN. The solution was then removed, the gel pieces dried in a speed vacuum and rehydrated in 100 µL 7% H_2O_2, at room temperature for 15 min, in the dark. The oxidizing solution was removed and the gel slices were rinsed in water as previously described above. After complete drying, the bands were rehydrated in 20 µL digestion buffer (150 mM Tris-HCl, 10 mM $CaCl_2$, 100 mM urea pH 8.1/CH_3CN 95/5 (v/v), containing 150 ng sequencing grade modified trypsin (Promega, Madison, WI, USA). After 15 min of incubation at 4°C, 30 µL of digestion buffer was added and the digestion reaction carried out at 37°C for 5 h, under permanent shaking. The digestion

solution was then collected and peptides were extracted from the gel by diffusion in 50 µL of 0.3M urea, 90 % (v/v) CH_3CN for 30 min with sonication. Digestion and extraction solutions were pooled and dried under speed vacuum. Peptide mixtures were redissolved in 25 µl of water/CH_3CN 95/5 (v/v) containing 0.2 % Formic acid (FA) prior to LC-MS/MS analysis.
For in-solution digestion, the vacuolar membrane fraction was re-suspended in 80 µL of 25 mM NH_4HCO_3, pH 8.1, and heated at 90°C for 10 min. Denaturation was stopped by adding 120 µL of cold (-20°C) methanol, thus avoiding protein refolding which can occur during slow cooling. Digestion was carried out overnight at 35°C, with an enzyme/substrate ratio of 1:100 (w/w). Finally, TCEP was added, to a final concentration of 10 mM, and reduction was carried out at 35°C for 30 min. The digestion solution was then dried under speed vacuum and peptide mixtures redissolved in 25 µl of Water/CH_3CN 95/5 (v/v) containing 0.2 % Formic acid (FA) prior to LC-MS/MS analysis.

NanoLC-ESI-MS/MS- Injected samples (6 µl), from in gel digestion, were first trapped and desalted isocratically on a PepMap µC18 precolumn cartridge 65 mm (300 µm id, 5 µm and 100 Å, Dionex, Sunnyvale, CA, USA). Chromatographic separation was accomplished by loading peptide samples onto a 15 cm fused silica C18 column (75 µm id, 3 µm, 100 Å and 360 µm od; Dionex) using an autosampler. Sequential peptide elution was achieved using the following linear gradient: (i) from 10% to 40% solvent B [CH_3CN/Water 90/10 (v/v) containing 0.1% FA] for 40 min, (ii) from 40% to 90% solvent B for 5 min, (iii). The remaining percentage of the elution solvent is made of solvent A [Water/CH_3CN 95/5 (v/v) containing 0.1% FA]. Flow rate through the nanoLC column is set to 200-300 nL/min. The mass spectrometer was calibrated using the product ions generated from fragmentation of the doubly-charged molecular ion of Glu-fibrinopeptide B. Raw data were processed using PeptideAuto (ProteinLynx, MassLynx 4.0) using smooth 3/2, Savitzky Golay). The mass spectrometer was operated in the positive ion electrospray ionization mode with a resolution of 9,000–11,000 full-width half-maximum. For automatic LC-MS/MS analysis the QTOF Ultima instrument was run in data dependent mode (DDA) with the following parameters: 1 s scan time and 0.1 s interscan delay for MS survey scans; 400-1400 and 50-2000 m/z mass ranges for the survey and the MS/MS scans respectively; 5 components; MS/MS to MS switch after 5 s; switchback threshold:30 counts/s; include charge states 2, 3 and 4 with the corresponding optimized

collision energy profiles. A list of the m/z corresponding to the most intense peptides of trypsin was set as an exclude list. Peptide identification from the resulting MS/MS dataset was achieved using an in-house MASCOT server (version 2.0) (Matrix Sciences, London, UK). Chromatographic separation of in-solution digested proteins was accomplished by loading 0.15 μg peptide mixture on the column. Sequential elution of peptides was performed using the following linear gradient: (i) from 15% to 60% solvent B for 90 min, (ii) from 60% to 90% solvent B for 5 min, (iii). Mass spectrometer was set as described above.

The acquired data were post-processed to generate peak lists (.pkl) using PeptideAuto which is a part of proteinLynx from Masslynx 4.0. The following parameters were used: QA threshold: 10; Smooth window: 3 (2 times in Savitzky Golay mode) and Centroid at min peak width at half height: 4; centroid top, 80 %. The peak lists are appended as a single file.

Each sample was submitted to consecutive searches against the Swiss-Prot Trembl database and the specific *A. thaliana* database using MASCOT 2.0. MASCOT search parameters used with MS/MS data were: database = *A. thaliana* (nuclear, mitochondrial, and chloroplastic genome), enzyme = trypsin/P, one missed cleavage allowed, a peptide tolerance = 0.25 Da, MS/MS tolerance = 0.25 Da, peptide charge = 2+/3+ and variable modifications. For the in-gel digestion procedure, these were acetyl-Nter /oxidized methionine under sulfone and sulfoxide form/ FMA+1 / FMA-1 / Cysteic acid. For the in-solution digestion procedure, variable modifications were acetyl-Nter /oxidized methionine under sulfone form / FMA+1 / FMA-1. Proteins identified by at least two peptides with a MASCOT MOWSE score higher than 50 were automatically validated. When this criterion was not respected, the fragmentation spectrum from each peptide was manually interpreted using the conventional fragmentation rules. In particular, we looked for a succession of at least five *y*- and/or *b*-ions, specific immonium ions, specific fragment ions (proline, glycine), and signatures of any modifications carried by the peptides. In cases where the protein was mainly identified by a single peptide match, the MS/MS ions were manually examined for notable sequence tag and independently verified using the PEAKS studio program (http://www.bioinformaticssolutions.com/). The algorithm can efficiently choose the best amino acid sequence, from all possible amino acid combinations, to interpret the MS/MS spectrum according to the same chemical modifications defined in Mascot. An additional search was carried out for the identification of the possible N-

terminal peptide of proteins using semi-trypsin and N-terminal acetylation as Mascot parameters (see supplemental Table III).

Specialized databases for Arabidopsis membrane proteins (Aramemnon) (http://crombec.botanik.uni-koeln.de/) (52) and Tair database (http://www.arabidopsis.org/) (53)(http://www.membranetransport.org/) were used to facilitate interpretation of protein sequence identified by the proteomic approach. Complementary information was obtained using Psort II prediction program (http://psort.hgc.jp) (54).

Expression of protein fusion in plants and protoplasts- GFP::Protein fusions were generated using Gateway technology (Invitrogen). cDNA encoding the protein to be tested for sub-cellular localization (At3g19820, At1g19450, At3g63520, At5g58070, At3g16240, At1g69840) were provided in entry clones (pENTR/SD/D-TOPO respectively U15125, U16253, U16861, U17005, U17252 and U25581), by the ABRC stock center. LR reactions were performed following manufacturer's instructions, using the destination vector pK7WGF2 (55), kindly provided by the Flanders Interuniversity Institute for Biotechnology. The resulting expression vectors were introduced in *Agrobacterium tumefaciens* C58 strain. For transient expression in tobacco leaves, 24 h-old culture of *A. tumefaciens* C58 were diluted five-times in an induction medium as described in (56). After 6 hours, the bacterial suspension was adjust to an OD_{600} of 0.5 and leached into leaves of 9 week-old tobacco (*Nicotiana bentamiana*). In parallel and as a positive control for tonoplastic localization, tobacco plants were also transformed with the vector pNB96 containing the Nramp3::GFP construction (kindly provided by S. Thomine). For transient protoplast transformation, the SmaI-DraI cassettes from expression vectors, containing the CaMV 35S, the GFP::protein fusion and the NOS terminator were sub-cloned into the pCR-BLUNT II-TOPO (Invitrogen) for more efficent transformation. Protoplasts were prepared as described for vacuole isolation, except that the rinsing medium was replaced by W5 medium (154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM glucose, 1.5 mM MES-KOH, pH 5.6). In a microcentrifuge tube, 100 µL of protoplasts (corresponding to 10^5 protoplasts) were added to 50 µg of plasmid DNA in 10 µL of water. PEG solution (110 µL) [0.4 M mannitol, 100 mM $CaCl_2$, 40% (w/v) PEG 4000] was added to the mix. After 25 min of incubation at room temperature, protoplasts were diluted in 440 µL W5 medium, centrifuged for 1 min at 100 *g*, resuspended in 1 mL of W5 buffer and incubated at 22°C for 4 days, before observation. Sub-

cellular localization was assessed by scanning confocal microscopy (TCS SP2, Leica, Germany), GFP was excited at a wave-length of 488 nm and fluorescence collected between 500 nm and 550 nm.

RESULTS AND DISCUSSION

Isolation and evaluation of purity of the vacuole preparations isolated from cultured Arabidopsis cells

Vacuoles from *Arabidopsis thaliana* suspension cultures grown in the light were purified after protoplast preparation, as described (43) with some modifications. An average of 120 ± 35 µg of vacuole proteins could be obtained from 18 g of 5-day-old *Arabidopsis* cells. The specific activity of the vacuolar marker α-mannosidase was 4.6 ± 0.5 µmol $h^{-1}mg^{-1}$ proteins in purified vacuoles. This represented an enrichment factor of 42-fold when compared to the specific activity of the same marker in crude protoplast extract (0.11 ± 0.02 µmol $h^{-1}mg^{-1}$ proteins) and an average yield of 2.1 % based on this activity. To estimate cross contamination, vacuole preparations were visually checked using epifluorescence microscopy (set for DAPI coloration). In these conditions very few red-fluorescing chloroplasts or protoplasts were seen among the blue (due to the presence of phenolic compounds in the lumen) vacuoles. This low contamination by chloroplasts and protoplasts was confirmed by the absence of detectable chlorophyll when analyzed after acetone extraction (not shown).

Protoplast and vacuolar proteins were separated by SDS-PAGE (Fig. 1A) and western blots were performed in order to analyze cross contaminations by other subcellular compartments including plastids but also mitochondria, plasmalemma and endoplasmic reticulum (Fig. 1B). Western blotting experiments using antibodies raised against the tobacco alpha tonoplast intrinsic protein (α-TIP)(50) and the cauliflower gamma tonoplast intrinsic protein (γ-TIP)(51) were also performed and showed that these TIP proteins were both highly enriched in purified vacuoles compared to the protoplast protein extract in which they were hardly visible under our experimental conditions (Fig. 1B). In contrast, the plasmalemma, the plastid and the mitochondrial protein markers were not detectable in the vacuolar protein extract (Fig. 1B). In the case of the ER marker, a thin band of 70 kDa (corresponding to the most abundant HDEL protein) (49) was detectable in the vacuole fraction, reflecting minor ER contaminations. Taken together these results showed that our vacuole preparations were of high quality and that contaminations by other organelle membranes were very low, allowing further proteomic investigation.

Identification of the major vacuolar proteins isolated from Arabidopsis suspension cells

Use of an in-organic-aqueous digestion method to identify membrane proteins

We first undertook the identification of the most abundant vacuolar proteins. The most intensely stained bands of the total vacuolar proteins separated by SDS-PAGE presented in Figure 1A were trypsin digested and peptide mixtures were submitted to nanoLC-ESI-MS/MS analysis (see also Fig. 2A). This proteomic analysis led essentially to the identification of soluble proteins (annotations Fig 1A). Among them, we identified: the tripeptidyl peptidase II (At4g20850, a subtilase family protein); the A and B subunits of the vacuolar type H^+-ATPase (VHA-A, At1g78900; VHA-B1, At1g76030), a protein similar to the bacterial TolB protein, member of the Tol system (At1g21680)(57); a putative leucine aminopeptidase (At2g24200); a putative pectin methylesterase (At1g11580); the glycosyl hydrolase family 17 (At4g16260), similar to the glucan endo-1,3-beta-glucosidase vacuolar isoform from *Hevea brasiliensis*; 3 band 7 family proteins (At5g62740, At5g51570, At1g69840), bearing strong similarities to hypersensitive-induced response protein from Zea mays; a basic endochitinase (At3g12500) involved in the ethylene/jasmonic acid mediated signaling pathway during systemic acquired resistance; a glutathione S-transferase (At2g30870, *At*GSTF10) induced by dehydration (58) and a SOUL heme-binding family protein (At1g17100). Interestingly, the latter protein, which has never been described in plants, is similar to the mammalian and chicken members of the SOUL/p22HBP family. P22HBP, a cytosolic protein, has been suggested to be involved in heme utilization for hemoprotein synthesis (59).

In order to identify the major membrane proteins present in the tonoplast, vacuoles were first frozen / thawed as described in "experimental procedures". The membrane fraction was then separated from the soluble material by ultracentrifugation and salt-washed in order to eliminate membrane-associated soluble proteins. At this step, to prevent the use of detergent to solubilize hydrophobic proteins, we used an alternative method which consists of an in-solution trypsin digestion in the presence of methanol. Membrane proteins were first denatured at a high temperature (90°C) followed by rapid cooling by the addition of cold (–20°C) MeOH to 60% (v/v) final concentration and tryptic digestion was then carried out in this solution. The resulting peptide mixture was analyzed by nanoLC-ESI-MS/MS (Fig. 2B) and 122 proteins were identified from a very low amount of proteins (2 x 0.15 µg). Peptides identified with a Mascot score

sufficient for protein identification are listed in supplemental Table I. This procedure allowed the identification of the most hydrophobic tonoplast proteins. Thirty of the most abundant membrane proteins identified in the tonoplast are presented in Table I (the others are presented in supplemental Table I). This indicative classification is based on a simplified exponentially modified protein abundance index (emPAI) proposed by Ishihama *et al.* (60). Relative protein abundance was estimated by normalising the number of peptides per protein by the theoretical number of peptides per protein. A caveat applies to this estimation, as extremely abundant proteins may affect the efficiency of protein identification because of ionization suppression and detector saturation. Highly hydrophobic proteins may also be under-represented because they generate a limited number of peptides due to their sequence characteristics. Despite these limitations, approximate comparison of the relative abundance of tonoplastic membrane proteins is possible because of the common characteristics they share. It is noteworthy that a total of 70% of the proteins were identified by two peptide hits or more. The majority of the proteins identified are transporters and the most representative proteins, as expected, are Tonoplast Intrinsic Protein (δ-TIP), subunits of the vacuolar H^+-pumping ATPase and the H^+-pumping pyrophosphatase (H^+-PPase AtVP1). Among the 30 most abundant proteins, there are also 3 ABC transporters (MRP 1, 4 and 10), 5 MATE efflux family proteins, a sodium/calcium exchanger family protein and a calcium-transporting ATPase (ACAc), a peptide transporter (PTR2-B), a copper transporter family protein (COPT5) and a cationic amino acid transporter (CAT2) (Table I).

This preliminary proteomic work not only confirmed the quality of our vacuole preparations but also revealed the identification of several proteins that had never been identified before by classical approaches (40-42). On the other hand, it also showed the lack of proteins, such as the thioglucosidases (At1g54010, At1g54000), that are known to be the main components of other vacuolar systems (42). Taken together, these results indicated the need to extend the knowledge of the vacuolar proteome of cultured *Arabidopsis* cells.

Toward an extended vacuolar proteome study

In order to increase the number of proteins identified in the tonoplast, pre-fractionation of the membrane and soluble proteins was performed by SDS-PAGE (Fig. 2C & D). A short migration (2.5 cm) was carried out to obtain efficient separation and to avoid diffusion effect

across the gel and keep the total number of gel bands to be analyzed reasonable. The gel was cut into 15 bands. A peroxide treatment was performed and proteins were then in-gel digested with trypsin and identified by nanoLC-ESI-MS/MS analysis. Extended coverage of the protein sequences was obtained using both: i) dynamic exclusion during the MS/MS process; and ii) a second nanoLC-ESI-MS/MS analysis using an exclusion list to limit re-fragmentation of peptides fragmented during the first run. The peroxide treatment applied before trypsin digestion is a mild oxidation procedure that transforms the cysteinyl residues into cysteic acids (+48 Da). The advantage of this method lies not only in the increased peptide coverage (not shown), but also in the speed of sample preparation compared with classical reduction/alkylation approaches. This treatment also converts all the methionines to the maximal oxidation state (sulfone, +32 Da), providing better quality MS/MS spectra and prevents the apparition of the intermediate fragmentation pattern of - 64 Da that complicates the mass attributions obtained with the Mascot software.

Using this procedure, 387 proteins were identified from the membrane fraction (supplemental Table II). Using both the in-gel and in-solution digestion procedures, a total of 416 non-redundant proteins were identified from the tonoplast. Among them, 50 had been previously demonstrated to be localized to the vacuole; 195 were integral membrane proteins, based on the presence of one or more predicted transmembrane domains (TMHMM2.0) (supplemental Table I and II); 5 had transmembrane beta barrel structures such as porins; 29 did not have any transmembrane domain but were known to be part of membrane complexes such as H^+-ATPase; 31 were predicted or known to have covalent lipid modifications leading to their insertion into the membrane (myristoylation and prenylation sites were predicted using http://plantsp.sdsc.edu/myrist.html and Psort II) (54). Among the latter, we found several band 7 family proteins, a putative glycosylphosphatidylinositol (GPI)-anchored protein, a multi-copper oxidase and Ras-related GTP-binding proteins. The 351 soluble proteins identified from the vacuolar sap will not be discussed here. The whole proteomic work presented here represents the identification of more than 650 non-redundant proteins which is the most complete study done to-date. The proteins found both in our study and in previously published ones (40-42) are given in Table I, II and supplemental Tables I and II.

The most complete study previously published identified 402 proteins (42). In general, our respective fractions are quite different and the overlap rate of the two studies is around 26%.

This weak overlap, in terms of identification is almost certainly related to differences in starting material. Carter *et al*. (42) used *Arabidopsis* leaves while we used *Arabidopsis* suspension cultures. It is noteworthy that in our study around 70% of the 650 proteins were identified by two or more peptide hits. Contrary to Carter *et al.*'s analysis of the total vacuolar fraction, our identification did not reveal the presence of a 42 kDa protein representing the bulk of vacuolar protein content (identified as a myrosinase-associated protein, At3g14210 and the corresponding myrosinase gene products, At5g25980 and At5g26000). This difference is probably due to the presence of myrosin cells in the *Arabidopsis* leaves which accumulate the thioglucoside glucohydrolase (61). Our membrane protein data was much more complete because the transporters identified in the previous study were well overlapped and our analysis identified a larger number of transporters with better coverage rates (and associated Mascot scores). To be complete, our membrane protein tonoplastic set must be compared to Shimaoka *et al*.'s results (40) where 163 proteins including well-characterized tonoplast proteins such as V-type H^+-ATPases and V-type H^+-PPases were identified. Both identifications were obtained from suspension cultured *Arabidopsis* cells. However, Shimaoka *et al.*'s data led to the identification of 163 proteins, of which 42 were membrane proteins (annotated with one or more transmembrane domains). Of these 42, 39 were predicted to have more than two transmembrane domains and 17 are possible transporters. Shimaoka *et al.* (40) and Szponarski *et al.* (41) identified a large number of mostly soluble proteins within their vacuolar fractions. This may be because of a lack of thorough washing of the vacuolar membranes or due to different LC or MS analysis protocols. Figure 3 shows the cross-correlation of the different proteome analyses of the vacuolar membrane system.

Proteins were categorized into thirteen major groups: transporters, stress response, signal transduction, metabolism, cellular transport, protein synthesis and degradation, cytoskeleton, glycosyl hydrolase, RNA degradation, unclassified and contaminants. It is of note that the main contamination of our preparations seems to be with cytosolic proteins, in particular with ribosomal proteins. Few possible contaminating mitochondrial or chloroplast proteins were detected. Among the classes of proteins identified, transporters are of particular interest since few vacuolar transporters have been identified and fully characterized so far.

Pumps and Transporters

Import and export of metabolites and ions require their corresponding transport systems across the tonoplast and many transporters and channels in this membrane remain to be identified. In the present work, we identified 91 proteins having demonstrated or predicted transporter activities. In addition, the two major components of the tonoplast, the vacuolar H^+-ATPase and the H^+-PPase were identified. The vacuolar H^+-ATPase is a large oligomeric protein complex (>700 kDa) composed of 12 subunits that form a hydrophilic V_1-subdomain (VHA-A to VHA-H) localized in the cytoplasm and an integral-membrane V_0-subdomain (VHA-a, VHA-c, VHA-d and VHA-e) (19,62). In *Arabidopsis*, 5 subunits are encoded by alternative splicing of a single gene (VHA-A, VHA-C, VHA-D, VHA-F and VHA-H), the others being encoded by multiple genes. Among the 28 possible subunits, 19 proteins were clearly identified. Table II summarizes the H^+-ATPase subunits identified in the present work alongside those identified in other published works (40,42). The high quality of the vacuole preparation and the analytical methods we used have allowed coverage of a high percentage of the protein sequences (12% to 75%). Corroborating this, very high Mascot score values, reaching for example 2038 in the case of the VHA-A subunit identification (vs. 222 in Carter *et al.*'s published data), were obtained. The identification of the other vacuolar pump, the H^+-PPase (AVP3) (63), was also achieved with a high Mascot score value (1949 and 1229 obtained from in-solution and in-gel digestion, respectively vs. 88 in previously published data).

The main representative superfamily of transporters that we identified was the ABC transporter family. Substrates assigned to members of this large family of transporters include compounds as diverse as peptides, sugars, lipids, heavy metal chelates, polysaccharides, alkaloids, steroids, inorganic acids and glutathione-conjugated compounds (64). The 129 ABC proteins encoded by the *Arabidopsis* genome are classified into 12 subfamilies based on their size, orientation, domain organization, and resemblance to ABC proteins from other organisms (65,66). In the present proteomic work, 14 different members belonging to 5 different subfamilies: multidrug resistance protein (MDR), multidrug resistance-associated proteins (MRP), pleiotropic drug resistance (PDR), non-intrinsic ABC protein (NAP) and transporter associated with antigen processing (TAP) were identified (Table III). Ten were clearly identified as MRP: MRP1-8, MRP10 and MRP14. The degree of identities shared between these MRP subfamily members ranges from 33% to 87% (67) making their unambiguous identification

difficult unless a high protein coverage is obtained (41,42). Only three of them: AtMRP1, AtMRP2 and AtMRP4 have been identified and localized to the tonoplast so far by classical methods (38,39). Our data confirmed their presence in the tonoplast membrane and 6 others were identified. Table III shows once again the high Mascot score and coverage values obtained in our analyses even for these difficult-to-identify proteins. As an example, MRP10 (At3g62700) was identified with a Mascot score of 2613 and 32% sequence coverage. Based on the protein coverage found in the in-solution digestion protocol, MRP10 seems to be the most abundant ABC subclass transporter in cultured *Arabidopsis* cells. Analysis of AtMRP10 (At3g62700) using the ARAMEMNON database (52) predicts the presence of a strong chloroplast-targeting signal and a weaker mitochondrion signal peptide. These predictions show that the *in silico* subcellular localization prediction algorithms are probably not robust enough. Surprisingly, we also find MRP4 in the tonoplast whereas Klein *et al.* (68) using confocal microscopy analysis of onion epidermal cells transiently expressing AtMRP4–EGFP showed fluorescence at the periphery of cells, suggesting AtMRP4 protein to be located at the plasma membrane. In support of our data, previous proteomic studies confirm the presence of MRP4 in the purified vacuolar fractions (40,42) but both localizations could be possible. Moreover, in a recent study, Dunkey *et al*. have unambiguously assigned 18 transporters to specific organelles by LOPIT technology (69). The authors have shown that the vacuolar membrane class is dominated by proteins involved in membrane transport, including eight ABC transporters (MRP1-6, MRP10, and TAP2).

Four different Tonoplast Intrinsic Proteins, also known as aquaporins, were identified on the tonoplast membrane (Table IV) (AtTIP1.1 and AtTIP1.2, corresponding to γ-TIP, and AtTIP2.1 and AtTIP2.2, corresponding to δ-TIP). We thus demonstrated that aquaporins are well expressed on *Arabidopsis* cultured cells and showed that one of the most abundant proteins present on the tonoplast is the AtTIP2.1 (Table I). Previous studies argued that plant cells have the ability to generate and maintain separate vacuole organelles, with each being marked by a different TIP subtype (5,6). For example, the fully differentiated cell types PSVs are marked by α-TIP plus δ-TIP, and lytic vacuoles are marked by γ-TIP. The presence of both markers, in our proteomic work, could be explained by the presence of different kinds of vacuoles in our samples, or the presence of vacuoles that come from fusion of storage and lytic vacuoles combining properties of both vacuoles as mentioned in (4).

Other transporters identified in the tonoplast proteome are given in Table IV. Among them, transporters that mediate the efflux of a broad range of compounds have been identified. For example, 17 members of detoxifying efflux transporters which include 8 major facilitator superfamily (MFS) members, 6 multidrug and toxin compound extrusion (MATE) family members, 3 examples of drug/metabolite transporter (DMT) superfamily proteins and 1 member of the resistance/nodulation/division (RND) superfamily. Members of the MFS family are secondary transporters that represent the largest group of ion-coupled transporters involved in the symport, antiport, or uniport of various substrates such as sugars, Krebs cycle intermediates, phosphate esters, oligosaccharides, and antibiotics (70). Members of the Multidrug / Oligosaccharidyl-lipid / Polysaccharide (MOP) family were also identified here. These are interesting because *Arabidopsis* expresses many members of this family not found in the animal kingdom. The MATE family, which is a MOP subtype, shows homologies with bacterial efflux transporters. The *Arabidopsis* genome codes for at least 54 members of this family and only a few members of the MATE family are characterized functionally. Their contribution to drug resistance has been shown only for a few isolated cases (71). Hydropathy analysis suggests that proteins of the MATE family have a common topology consisting of 12 transmembrane domains. Six different MATE efflux transporters were characterized in the tonoplast fraction suggesting the potential contribution of the vacuole system to drug resistance (Table IV).

Five peptide transporters were identified, 3 members of the Proton-dependent oligopeptide Transporter (POT) family and 2 Oligopeptide Transporters (OPT). Interestingly, one of the best characterized plant OPT members is the maize 'Yellow Stripe1' (YS1) transporter, a Fe^{3+}-phytosiderophore:H^+ symporter involved in Fe^{3+} uptake (72) but also in the uptake of various other metal cations complexed with either phytosiderophores or nicotianamine (73). The 2 OPT family members that we have identified (AtYSL4 and AtYSL6) in the tonoplast are classed as Yellow Stripe-Like (YSL) transporters (72) and present 69% and 70% similarity with YS1. We also confirmed the presence of other metal transporters such as NRAMP3 (36) and AtMTP1 (AtZAT1)(32,33). We did not identify the Na^+/H^+ antiporter (AtNHX1)(74) but found an isoform, AtNHX4, which is also a member of the Monovalent Cation: Proton Antiporter-1 (CPA1) Family. A putative iron (Fe^{2+}) transporter (AtIRT3), a member of the Zinc-Iron Permease (ZIP) family, was also identified (75). The vacuolar Calmodulin-regulated Ca^{2+}-ATPase 4 (ACA4) (30) was also found in the tonoplast. Three K^+ Uptake Permeases, members of one of the

five potassium transporter families were identified in the present work (AtKUP7, AtKUP5 and AtKUP4) (75). The latter, AtKUP4 (or AtTRH1), has recently been shown to be required for auxin transport in *Arabidopsis* roots (76).

The tonoplast contains at least 7 amino acid transporters from the 50 distinct amino acid transporter genes encoded by the *Arabidopsis* genome. We identified 4 of the 9 Cationic Amino acid Transporters (AtCAT2, AtCAT4, AtCAT8 and AtCAT9) present in *Arabidopsis*. Interestingly, a recent molecular and functional characterization of this family shows that AtCAT2 is localized to the tonoplast whereas AtCAT5 (not found in this proteomic work) is present in the plasma membrane (77). AtCAT2 and AtCAT4 were identified in the tonoplast in of Carter *et al's* study. (42). We also found 5 Amino Acid/Auxin Permease (AAAP) Family members in the tonoplast (Table IV).

Moreover, a few other unexpected proteins such as "expressed protein similar to TolB protein precursor" (At1g21680) or "Niemann-Pick C1 similar protein" (At4g38350) were found in the tonoplast membrane. TolB, a periplasmic protein found in most Gram-negative proteomes, is one of the Tol proteins of *Escherichia coli* and is involved in the translocation of group A colicins. TolB also forms a complex with Pal, an outer membrane peptidoglycan-associated lipoprotein anchored to the outer membrane by its N-terminal lipid moiety (57). The exact role of the Tol system remains to be determined. Its presence in the tonoplast is probably linked to membrane biogenesis as supposed for the Gram-negative organisms. Niemann-Pick C1 protein is a large multitransmembrane glycoprotein that was shown to reside primarily in mamalian late endosomes. Its cytoplasmic tail contains a dileucine endosome-targeting motif and it transiently associates with lysosomes and the *trans*-Golgi network. The function of the NPC1 protein is unclear. However, a number of observations suggest that NPC1 may be related to a family of prokaryotic efflux pumps and thus it may also act as a molecular pump (78), a cholesterol transporter (79) or play a role in docking/fusion events (80).

Other proteins

From our data, we can pinpoint another important uncharacterized class of proteins: the band 7 protein family. Twenty members of this family (also known as Stomatin Prohibitin Flotilin Hbc, SPFH domain proteins) are predicted from the *Arabidopsis* genome, 10 of which

were identified in the vacuole fraction (Table V). Co-fractionation with the tonoplastic membrane in the presence of sodium carbonate at pH 11.5 followed by salt wash strongly suggests that the band 7 proteins are true integral membrane proteins. Interestingly, this observation is in agreement with the fact that the band 7 proteins are putatively myristoylated or have one TMD. In most cases, this modification is essential for protein function to mediate membrane association or protein protein interaction. The SPFH protein-domain is characteristic of the prohibitins and the stomatins which are putatively involved in cell cycle and ion channel control. Multiple stomatin orthologues from bacteria, plants and animals have been identified. In *Caenorhabditis elegans*, MEC-2, a stomatin-like protein is involved in mechanosensation (81). Another example is the protein UNC-1, a close *Caenorhabditis elegans* homologue of the mammalian protein stomatin, which may be involved in anesthetic sensitivity and could represent a molecular target for volatile anesthetics (82). Cumulative evidence suggests that band 7 stomatins may modulate membrane functions especially those of Detergent Resistant Microdomains (DRM) or lipid rafts (reviewed in (83)). The presence of this family in the vacuolar membrane may play a crucial role in the regulation of either the biogenesis of the tonoplast membrane or the regulation of transporters and metabolite flux across the tonoplast. Consequently, we confirmed the sub-cellular localization of one member of this family using a GFP fusion protein approach (see below).

In eukaryotes, peptides derived from proteasomal degradation of intracellular proteins have been shown to translocate from the cytosol into the ER via the transporter associated with antigen processing (TAP) (84). TAP is a macromolecular peptide-loading machinery composed of TAP1, TAP2, tapasin and several auxiliary factors (e.g. calreticulin and ERp57). According to the proteins characterized in our tonoplastic fraction, similar machinery for protein degradation and peptide import into the vacuole may be present in *Arabidopsis*. Indeed, we have identified a TAP2 homologue, a calreticulin, ERp57 and other members of the TAP machinery and this could explain the number of proteasome subunits observed in our sample (supplemental Table II).

Confirmation of the vacuolar localization of selected proteins

The quality of the purified vacuolar fraction, as assessed by enzymatic and immunological assays, was the first criterion for establishing the likelihood of vacuolar localization of the

proteins identified in this study. Indeed, mass spectrometry analyses did not reveal the presence of major proteins from other organelles. Moreover, several proteins identified (V^+-ATPase, tonoplast intrinsic proteins (TIPs), ABC transporter MRP1 and MRP2, etc.) were already known as vacuolar proteins. To further characterize the vacuolar system studied here, the sub-cellular localization of several proteins was investigated by transient expression in transfected tobacco plants and in *Arabidopsis* protoplasts of GFP-fusion proteins. We chose the following proteins from Tables I, V and supplemental Table II: the cell elongation protein DWARF1/DIMINUTO (DWARF1, At3g19820), a band 7 family protein (At1g69840), an integral membrane protein (At1g19450), a putative lipocalin (At5g58070) and the 9-cis-epoxycarotenoid dioxygenase (CCD1, At3g63520). Confocal microscopy imaging was carried out on transfected tobacco plants or on protoplasts isolated from *Arabidopsis* cell suspensions (Fig. 4). As a positive control in accordance with previously published data (36), we expressed GFP-NRAMP3 (At2g23150) and confirmed its vacuolar cellular location in tobacco cell leaves (Fig. 4A). The GFP-TIP2.1 (δ-TIP, At3g16240) construct was also used as vacuolar protein control and is presented in Figure 4B. Inspection of the GFP-DWARF1 transformed plant using fluorescence microscopy confirms its presence in the tonoplast compartment (Fig. 4D). Analysis of the At3g19820 DWARF1 amino acid sequence predicts a transmembrane domain at the N terminus of the protein. It has been proposed that DWARF1 is a peripheral or an integral membrane protein (85). The *Arabidopsis DWARF1* gene encodes a protein involved in steroid synthesis and a vacuolar sub-cellular localization is in good agreement with the proposed role for the DWARF1 protein as a biosynthetic enzyme, because sterols as well as steroid hormones are relatively hydrophobic moieties, one would expect that synthesis may occur in a membrane environment. The At1g69840 gene product is a band 7 family protein presenting similarity to hypersensitive-induced response protein with relatively high homology to regions of stomatin and prohibitin (see above). Its presence in the vacuolar membrane was also confirmed (Fig. 4E).

According to Aramemnon database prediction, the At1g19450 gene product is an integral membrane protein with 12 transmembrane domains. Among the biological processes to which this protein may contribute are: phosphate transport, oligopeptide transport, cell adhesion, carbohydrate transport, phospholipid biosynthesis and tissue regeneration. The At1g19450 gene product shows 71% similarity to a putative sugar transporter from sugar beet which was cloned and localized to the vacuole in transgenic yeast and tobacco (86). This putative transporter is a

member of a subgroup of a large gene family, currently termed the MFS (70). Figure 4F confirmed the vacuolar localization of this transporter in tobacco cells.

The At5g58070 gene product is a putative lipocalin family member or outer membrane lipoprotein-like. Lipocalins are widely distributed among vertebrates and a few have been isolated from invertebrates and plants. Predominantly, lipocalins are small secreted proteins. Some members of the family exhibit high affinity and selectivity for hydrophobic molecules such as cholesterol, pheromones or fatty acids. Others have been shown to bind to specific cell-surface receptors and to form macromolecular complexes. The lipocalins have been classified as transport proteins. The non-solubilization of the protein during our membrane washing procedure indicated that this protein could be an anchor-linked protein, and the fluorescence microscopic analysis presented in Figure 4G demonstrated the presence of lipocalin protein on the tonoplast membrane. As the localisation of these latter two proteins (i.e. lipocalin and the sugar transporter) presented some heterogeneity in the leaf model, we performed a transient expression analysis in *Arabidopsis* protoplasts (Fig 4I-K). It is interesting to note that lipocalin seems not to be exclusively localised on the tonoplast (Fig. 4I), while its presence on the vacuolar membrane is confirmed by observing a fluorescent vacuole being released from a cell (Fig. 4J). Inversely, the putative sugar transporter is clearly exclusively targeted to the tonoplast (Fig. 4K).

The last protein that we chose to study as a GFP-fusion construct was the carotenoid cleavage dioxygenase (CCD1, At3g63520). Apocarotenoids, derived from the oxidative cleavage of carotenoids, are involved in important metabolic and hormonal functions in diverse organisms (87,88). Since the discovery of VP14 in maize (87), the "prototypic nine-*cis*-epoxy carotenoid dioxygenase" (NCED), several carotenoid cleavage enzymes have been characterized and grouped in different classes according to substrate specificities (88). Class 1 is represented by the NCEDs which are involved in abscisic acid (ABA) biosynthesis (89). CCD1 is a member of the class 2 family, catalysing a dioxygenase reaction leading to the synthesis of β-ionone and C_{14} dialdehyde (90). So far, nine carotenoid cleavage dioxygenase genes have been identified in the complete *Arabidopsis* genome and among them, only five NCEDs (2, 3, 5, 6, and 9) have been shown to be targeted to plastids (91). However, no information was available concerning CCD1's localization; as shown in Figure 4H, our analyses clearly identified CCD1 in the vacuole membrane. As expected, and taking into account the overall results presented previously, the vacuolar localization of all the GFP tagged proteins was confirmed.

CONCLUSION

The analysis of a proteome at the level of subcellular structures represents an analytical strategy which combines traditional biochemical methods of fractionation and tools for protein identification. We have shown the strategy developed in this work to be very efficient to get a broader view of the vacuole membrane proteome. We identified several new membrane proteins such as channels and transporters that mediate the translocation of molecules and ions across tonoplast membranes. We consider that most of the proteins identified in this study are genuine constituents of the tonoplast. This is based on the identification of key tonoplastic components, the verification of vacuolar localization for selected proteins using GFP-fusion proteins and the low contamination of our preparations by other organelles. Through our vacuolar proteomic strategy, it seems clear that the tonoplast membrane is a complex structure that receives membrane and protein contributions from a variety of subcellular sources and pathways. The high degree of protein diversity of the tonoplast membrane is indicative of a highly complex organelle. Further validation of the results presented here by relevant functional studies should provide a better explanation for the biogenesis and maintenance of these unique organelles.

Acknowledgements

This work was funded by the « Programme inter-organismes CEA CNRS INRA INSERM de Toxicologie Nucléaire Environnementale ». We thank Jean-Jacques Leguay, Marie-Thérèse Ménager and Eric Ansoborlo for their support and Francis Marty, Christophe Maurel, Marc Boutry, Maryse Block and Béatrice Satiat-Jeunemaitre for providing antibodies. We also thank Sébastien Thomine for providing the plasmid harbouring the GFP-Nramp3 fusion protein. We are also grateful to Maighread Gallagher-Gambarelli for critical reading of the manuscript.

REFERENCES

1. De, D. N. (2000), pp 79-114
2. Marty, F. (1999) *Plant Cell* **11**(4), 587-600
3. Paris, N., Stanley, C. M., Jones, R. L., and Rogers, J. C. (1996) *Cell* **85**(4), 563-572
4. Vitale, A., and Raikhel, N. V. (1999) *Trends in plant science* **4 (4)**, 149-155
5. Jauh, G. Y., Fischer, A. M., Grimes, H. D., Ryan, C. A., Jr., and Rogers, J. C. (1998) *Proc Natl Acad Sci* **95**(22), 12995-12999
6. Jauh, G. Y., Phillips, T. E., and Rogers, J. C. (1999) *Plant Cell* **11**, 1867-1882
7. Phillipson, B. A., Pimpl, P., daSilva, L. L. P., Crofts, A. J., Taylor, J. P., Movafeghi, A., Robinson, D. G., and Denecke, J. (2001) *Plant Cell* **13**(9), 2005-2020
8. Rojo, E., Zouhar, J., Carter, C., Kovaleva, V., and Raikhel, N. V. (2003) *Proc Natl Acad Sci U S A* **100**(12), 7389-7394
9. Neumann, U., Brandizzi, F., and Hawes, C. (2003) *Annals Of Botany* **92**(2), 167-180
10. Vitale, A., and Hinz, G. (2005) *Trends Plant Sci* **10**(7), 316-323
11. Pimpl, P., Movafeghi, A., Coughlan, S., Denecke, J., Hillmer, S., and Robinson, D. G. (2000) *Plant Cell* **12**(11), 2219-2235
12. DaSilva, L. L. P., Snapp, E. L., Denecke, J., Lippincott Schwartz, J., Hawes, C., and Brandizzi, F. (2004) *Plant Cell* **16**(7), 1753-1771
13. Park, M., Kim, S. J., Vitale, A., and Hwang, I. (2004) *Plant Physiology* **134**(2), 625-639
14. Pimpl, P., Hanton, S. L., Taylor, J. P., Pinto daSilva, L. L., and Denecke, J. (2003) *Plant Cell* **15**(5), 1242-1256
15. Rojo, E., Zouhar, J., Kovaleva, V., Hong, S., and Raikhel, N. V. (2003) *Mol Biol Cell* **14**(2), 361-369
16. Huang, W. P., and Klionsky, D. J. (2002) *Cell Struct Funct* **27**(6), 409-420
17. Maeshima, M. (2001) *Annual Review Of Plant Physiology And Plant Molecular Biology* **52**, 469-NIL_433,473-497
18. Sze, H., Li, X., and Palmgren, M. G. (1999) *Plant Cell* **11**(4), 677-690
19. Kluge, C., Lahr, J., Hanitzsch, M., Bolte, S., Golldack, D., and Dietz, K. J. (2003) *J Bioenerg Biomembr* **35**(4), 377-388

20. Rea, P. A., and Poole, R. J. (1993) *Annual Review Of Plant Physiology And Plant Molecular Biology* **44**, 157-180
21. Li, J., Yang, H., Peer, W. A., Richter, G., Blakeslee, J., Bandyopadhyay, A., Titapiwantakun, B., Undurraga, S., Khodakovskaya, M., Richards, E. L., Krizek, B., Murphy, A. S., Gilroy, S., and Gaxiola, R. (2005) *Science* **310**(5745), 121-125
22. Blumwald, E. (2000) *Curr Opin Cell Biol* **12**(4), 431-434
23. Apse, M. P., Aharon, G. S., Snedden, W. A., and Blumwald, E. (1999) *Science* **285**(5431), 1256-1258
24. Yamaguchi, T., Aharon, G. S., Sottosanto, J. B., and Blumwald, E. (2005) *Proc Natl Acad Sci U S A* **102**(44), 16107-16112
25. Hirschi, K. D. (2004) *Plant Physiol* **136**(1), 2438-2442
26. Hirschi, K. D., Zhen, R. G., Cunningham, K. W., Rea, P. A., and Fink, G. R. (1996) *Proc Natl Acad Sci U S A* **93**(16), 8782-8786
27. Ueoka-Nakanishi, H., Nakanishi, Y., Tanaka, Y., and Maeshima, M. (1999) *Eur J Biochem.* **262**(2), 417-425
28. Cheng, N. H., Pittman, J. K., Barkla, B. J., Shigaki, T., and Hirschi, K. D. (2003) *Plant Cell* **15**(2), 347-364
29. Geisler, M., Frangne, N., Gomes, E., Martinoia, E., and Palmgren, M. G. (2000) *Plant Physiology* **124**(4), 1814-1827
30. Baxter, I., Tchieu, J., Sussman, M. R., Boutry, M., Palmgren, M. G., Gribskov, M., Harper, J. F., and Axelsen, K. B. (2003) *Plant Physiol* **132**(2), 618-628
31. Shaul, O., Hilgemann, D. W., de-Almeida-Engler, J., Van Montagu, M., Inze, D., and Galili, G. (1999) *Embo J* **18**(14), 3973-3980
32. Kobae, Y., Uemura, T., Sato, M. H., Ohnishi, M., Mimura, T., Nakagawa, T., and Maeshima, M. (2004) *Plant Cell Physiol* **45**(12), 1749-1758
33. Desbrosses-Fonrouge, A. G., Voigt, K., Schroder, A., Arrivault, S., Thomine, S., and Kramer, U. (2005) *FEBS Lett* **579**(19), 4165-4174
34. Kramer, U. (2005) *Trends Plant Sci* **10**(7), 313-315
35. Drager, D. B., Desbrosses-Fonrouge, A. G., Krach, C., Chardonnens, A. N., Meyer, R. C., Saumitou-Laprade, P., and Kramer, U. (2004) *Plant J* **39**(3), 425-439

36. Thomine, S., Lelievre, F., Debarbieux, E., Schroeder, J. I., and Barbier Brygoo, H. (2003) *Plant Journal* **34**(5), 685-695
37. Lanquar, V., Lelievre, F., Bolte, S., Hames, C., Alcon, C., Neumann, D., Vansuyt, G., Curie, C., Schroder, A., Kramer, U., Barbier-Brygoo, H., and Thomine, S. (2005) *Embo J* **24**(23), 4041-4051
38. Liu, G., Sanchez-Fernandez, R., Li, Z. S., and Rea, P. A. (2001) *J Biol Chem.* **276**(12), 8648-8656
39. Geisler, M., Girin, M., Brandt, S., Vincenzetti, V., Plaza, S., Paris, N., Kobae, Y., Maeshima, M., Billion, K., Kolukisaoglu, U. H., Schulz, B., and Martinoia, E. (2004) *Mol Biol Cell* **15**(7), 3393-3405
40. Shimaoka, T., Ohnishi, M., Sazuka, T., Mitsuhashi, N., Hara Nishimura, I., Shimazaki, K. I., Maeshima, M., Yokota, A., Tomizawa, K. I., and Mimura, T. (2004) *Plant And Cell Physiology* **45**(6), 672-683
41. Szponarski, W., Sommerer, N., Boyer, J. C., Rossignol, M., and Gibrat, R. (2004) *Proteomics* **4**(2), 397-406
42. Carter, C., Pan, S., Zouhar, J., Avila, E. L., Girke, T., and Raikhel, N. V. (2004) *Plant Cell.* **16**(12), 3285-3303
43. Frangne, N., Eggmann, T., Koblischke, C., Weissenbock, G., Martinoia, E., and Klein, M. (2002) *Plant Physiology* **128**(2), 726-733
44. Boller, T., and Kende, H. (1979) *Plant Physiol* **63**(6), 1123-1132
45. Bolter, B., Soll, J., Hill, K., Hemmler, R., and Wagner, R. (1999) *EMBO J.* **18**(20), 5505-5516
46. Jansch, L., Kruft, V., Schmitz, U. K., and Braun, H. P. (1998) *J Biol Chem.* **273**(27), 17251-17257
47. Morsomme, P., Dambly, S., Maudoux, O., and Boutry, M. (1998) *J Biol Chem* **273**(52), 34837-34842
48. Napier, R. M., Fowke, L. C., Hawes, C., Lewis, M., and Pelham, H. R. (1992) *J Cell Sci* **102 (Pt 2)**, 261-271
49. Baluska, F., Samaj, J., Napier, R., and Volkmann, D. (1999) *Plant journal* **19 (4)**, 481-488
50. Gerbeau, P., Guclu, J., Ripoche, P., and Maurel, C. (1999) *Plant journal* **18 (6)**, 577-587

51. da Silva Conceicao, A., Marty-Mazars, D., Bassham, D. C., Sanderfoot, A. A., Marty, F., and Raikhel, N. V. (1997) *Plant Cell* **9**(4), 571-582
52. Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W. B., Flugge, U. I., and Kunze, R. (2003) *Plant Physiology* **131**(1), 16-26
53. Huala, E., Dickerman, A. W., Garcia Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M. Z., Huang, W., Mueller, L. A., Bhattacharyya, D., Bhaya, D., Sobral, B. W., Beavis, W., Meinke, D. W., Town, C. D., Somerville, C., and Rhee, S. Y. (2001) *Nucleic Acids Research* **29**(1), 102-105
54. Nakai, K., and Horton, P. (1999) *Trends Biochem Sci.* **24**(1), 34-36
55. Karimi, M., De Meyer, B., and Hilson, P. (2005) *Trends Plant Sci* **10**(3), 103-105
56. Lavy, M., Bracha-Drori, K., Sternberg, H., and Yalovsky, S. (2002) *Plant Cell* **14**(10), 2431-2450
57. Lazzaroni, J. C., Dubuisson, J. F., and Vianney, A. (2002) *Biochimie* **84**(5-6), 391-397
58. Kiyosue, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1993) *FEBS Lett* **335**(2), 189-192
59. Taketani, S., Adachi, Y., Kohno, H., Ikehara, S., Tokunaga, R., and Ishii, T. (1998) *J Biol Chem* **273**(47), 31388-31394
60. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005) *Mol Cell Proteomics* **4**(9), 1265-1272
61. Ueda, H., Nishiyama, C., Shimada, T., Koumoto, Y., Hayashi, Y., Kondo, M., Takahashi, T., Ohtomo, I., Nishimura, M., and Hara-Nishimura, I. (2006) *Plant Cell Physiol* **47**(1), 164-175
62. Sze, H., Schumacher, K., Muller, M. L., Padmanaban, S., and Taiz, L. (2002) *Trends Plant Sci* **7**(4), 157-161
63. Sarafian, V., Kim, Y., Poole, R. J., and Rea, P. A. (1992) *Proc Natl Acad Sci U S A* **89**(5), 1775-1779
64. Theodoulou, F. L. (2000) *Biochimica Et Biophysica Acta Biomembranes* **1465**(1-2), 79-103
65. Sanchez Fernandez, R., Davies, T. G. E., Coleman, J. O. D., and Rea, P. A. (2001) *Journal Of Biological Chemistry* **276**(32), 30231-30244
66. Garcia, O., Bouige, P., Forestier, C., and Dassa, E. (2004) *J Mol Biol* **343**(1), 249-265

67. Kolukisaoglu, H. U., Bovet, L., Klein, M., Eggmann, T., Geisler, M., Wanke, D., Martinoia, E., and Schulz, B. (2002) *Planta* **216**(1), 107-119
68. Klein, M., Geisler, M., Suh, S. J., Kolukisaoglu, H. U., Azevedo, L., Plaza, S., Curtis, M. D., Richter, A., Weder, B., Schulz, B., and Martinoia, E. (2004) *Plant Journal* **39**(2), 219-236
69. Dunkley, T. P., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., Griffin, J. L., Bessant, C., Brandizzi, F., Hawes, C., Watson, R. B., Dupree, P., and Lilley, K. S. (2006) *Proc Natl Acad Sci U S A* **103**(17), 6518-6523
70. Saier, M. H., and Paulsen, I. T. (2001) *Seminars In Cell And Developmental Biology* **12**(3), 205-213
71. Diener, A. C., Gaxiola, R. A., and Fink, G. R. (2001) *Plant Cell* **13**(7), 1625-1637
72. Curie, C., Panaviene, Z., Loulergue, C., Dellaporta, S. L., Briat, J. F., and Walker, E. L. (2001) *Nature* **409**(6818), 346-349
73. Schaaf, G., Ludewig, U., Erenoglu, B. E., Mori, S., Kitahara, T., and von Wiren, N. (2004) *J Biol Chem* **279**(10), 9091-9096
74. Gaxiola, R. A., Rao, R., Sherman, A., Grisafi, P., Alper, S. L., and Fink, G. R. (1999) *Proc Natl Acad Sci U S A.* **96**(4), 1480-1485
75. Maser, P., Thomine, S., Schroeder, J. I., Ward, J. M., Hirschi, K., Sze, H., Talke, I. N., Amtmann, A., Maathuis, F. J., Sanders, D., Harper, J. F., Tchieu, J., Gribskov, M., Persans, M. W., Salt, D. E., Kim, S. A., and Guerinot, M. L. (2001) *Plant Physiol* **126**(4), 1646-1667
76. Vicente-Agullo, F., Rigas, S., Desbrosses, G., Dolan, L., Hatzopoulos, P., and Grabov, A. (2004) *Plant J* **40**(4), 523-535
77. Su, Y. H., Frommer, W. B., and Ludewig, U. (2004) *Plant Physiol* **136**(2), 3104-3113
78. Davies, J. P., Chen, F. W., and Ioannou, Y. A. (2000) *Science* **290**(5500), 2295-2298
79. Amigo, L., Mendoza, H., Castro, J., Quinones, V., Miquel, J. F., and Zanlungo, S. (2002) *Hepatology* **36**(4 Pt 1), 819-828
80. Ioannou, Y. A. (2005) *Trends Biochem Sci* **30**(9), 498-505
81. Huang, M., Gu, G., Ferguson, E. L., and Chalfie, M. (1995) *Nature London* **378 (6554)**, 292-295

82. Rajaram, S., Sedensky, M. M., and Morgan, P. G. (1998) *Proc Natl Acad Sci U S A.* **95**(15), 8761-8766
83. Kurzchalia, T. V., and Parton, R. G. (1999) *Curr Opin Cell Biol* **11**(4), 424-431
84. Abele, R., and Tampe, R. (2006) *FEBS Lett* **580**(4), 1156-1163
85. Klahre, U., Noguchi, T., Fujioka, S., Takatsuto, S., Yokota, T., Nomura, T., Yoshida, S., and Chua, N. H. (1998) *Plant Cell.* **10**(10), 1677-1690
86. Chiou, T. J., and Bush, D. R. (1996) *Plant Physiology* **110**(2), 511-520
87. Schwartz, S. H., Bao Cai, T. A. N., Gage, D. A., Zeevaart, J. A. D., and McCarty, D. R. (1997) *Science Washington DC* **276 (5320)**, 1872-1874
88. Bouvier, F., Isner, J. C., Dogbo, O., and Camara, B. (2005) *Trends Plant Sci* **10**(4), 187-194
89. Nambara, E., and Marion-Poll, A. (2005) *Annu Rev Plant Biol* **56**, 165-185
90. Simkin, A. J., Underwood, B. A., Auldridge, M., Loucas, H. M., Shibuya, K., Schmelz, E., Clark, D. G., and Klee, H. J. (2004) *Plant Physiol* **136**(3), 3504-3514
91. Tan, B. C., Joseph, L. M., Deng, W. T., Liu, L. J., Li, Q. B., Cline, K., and McCarty, D. R. (2003) *Plant Journal* **35**(1), 44-56
92. Ren, Q. H., Kang, K. H., and Paulsen, I. T. (2004) *Nucleic Acids Research* **32 Special Iss. SI**, D284-D288

Table I Compilation of the proteins identified in the vacuole tonoplast proteome using an in-solution digestion procedure developed for hydrophobic protein identification.

Name	Gene	pI	Mr (Da)	Score	Cov. (%)	Nb peptide	N_{obsbl}	Rel ab	TMD	Function
delta Tonoplast Intrinsec Protein (δ-TIP)	At3g16240	5.3	25011	311	21.2	2	2	1	6	Transporter
Vacuolar H^+-pumping ATPase (VHA-c2)	At1g19910	8.62	16632	96	10.9	1	1	1	3	Transporter
Expressed protein	At2g34585	4	8565	97	50.6	2	2	1	1	Unclassified
Prenylated rab acceptor (PRA1) family protein	At1g55190	7	21058	117	10.6	2	2	1	4	Cellular trans
H^+-translocating (pyrophosphate-energized) inorganic pyrophosphatase (H^+-PPase AtVP3)	At1g15690	5.1	80768	1948	32.5	18	20	0.90	15	Transporter
MATE efflux family protein	At3g21690	5.4	54916	173	4.3	4	5	0.80	12	Transporter
Cell elongation protein DWARF1/DIMINUTO	At3g19820	8	65352	1119	43.3	19	24	0.79	1	Micellaneous
Vacuolar H^+-pumping ATPase (VHA-a3)	At4g39080	5.7	92774	2027	39.8	22	29	0.75	6	Transporter
Expressed protein	At4g20150	8.1	9202	193	28.4	2	3	0.66	0	Unclassified
Vacuolar H^+-pumping ATPase (VHA-a2)	At2g21410	5.4	93045	1048	25.8	17	27	0.63	7	Transporter
Acid phosphatase class B family protein (glycol/protein?)	At1g04040	9.2	31076	625	47.2	9	15	0.60	1	Signal Trans.
Sodium/calcium exchanger family protein	At1g53210	5.2	63376	574	21.4	9	15	0.60	10	Transporter
Transporter-related (hexose transporter?)	At4g35300	5.4	79674	354	12.7	6	10	0.60	10	Transporter
Glutathione-conjugate transporter (MRP10) similar to AtMRP4	At3g62700	8.8	172027	2080	29.2	37	63	0.58	15	Transporter
Peptide transporter (PTR2-B)	At2g02040	5.5	64380	567	18.6	7	13	0.54	10	Transporter
Glutathione-conjugate transporter (MRP4)	At2g47800	7.6	168972	1738	21.3	32	59	0.54	16	Transporter
MATE efflux family protein	At4g25640	6.1	53190	210	8.6	4	8	0.50	12	Transporter
Polyubiquitin (UBQ9)	At5g37640	5.8	36234	110	15.2	2	4	0.50	0	Stress resp
Copper transporter family protein (COPT5)	At5g20650	9.1	15774	75	13.0	2	4	0.50	2	Transporter
Glutathione S-conjugate transporter (MRP1)	At1g30400	5.9	181811	1549	22.9	27	61	0.44	14	Transporter
Band 7 family protein (strong similarity to hypersensitive-induced response protein (Zea mays) GI:7716470)	At5g62740	5.3	31411	312	30.4	7	16	0.44	Myr	Stress resp

Cationic amino acid Transporter 2 (CAT2)	At1g58030	5.6	67072	240	11.5	4	9	0.44	14	Transporter
Vacuolar H^+-pumping ATPase (VHA-G1)	At3g01390	5.8	12389	163	30.9	2	5	0.40	MB	Transporter
Nodulin MtN21 family protein	At1g75500	9.2	42544	204	7.5	3	8	0.38	10	Transporter
Integral membrane protein, putative	At1g75220	8.7	52867	206	8.2	3	9	0.33	12	Transporter
Integral membrane protein, putative (sugar transporter ?)	At1g19450	8.7	52753	122	7.0	3	9	0.33	12	Transporter
MATE efflux family protein	At3g26590	5.7	54286	119	8.6	3	9	0.33	12	Transporter
Expressed protein	At1g73650	9.1	32939	177	11.4	3	9	0.33	6	Transporter
MATE efflux protein-related	At5g52450	8.4	52439	166	7.2	3	10	0.30	12	Transporter
Calcium-transporting ATPase	At3g57330	6	111874	593	11.2	11	37	0.30	11	Transporter
Ripening-responsive protein	At1g47530	6.8	52389	173	8.5	3	10	0.30	12	Transporter

pI and Mr represent the theoretical isoelectric point and molecular mass, respectively. Score, the Mascot score of in-solution digestion protocol. Cov., the protein sequence coverage in %. Nb peptide represents the number of peptides assigned; N_{obsbl} is the number of observable peptides per protein in our experimental conditions. Rel ab is an estimation of relative protein abundance obtained using the number of peptides detected per protein (Nb peptide) normalized by the theoretical number of detectable peptides (N_{obsbl}). This classification is based on a simplified Exponentially Modified Protein Abundance Index (emPAI) proposed by Ishihama Y *et al*. (60). TMD, the number of transmembrane domains predicted by ARAMEMNON (http://crombec.botanik.uni-koeln.de/) (52). Function, the predicted protein function, as assigned by the MIPS Functional Catalogue (FunCat) (http://mips.gsf.de/proj/funcatDB/search_main_frame.html).

Table II: Compilation of the H^+-ATPase subunits identified in the vacuole proteome.

Subunit	Gene	AA / Mr (kDa)	pI	TMD	Score gel & cov. (%)	Nb gel peptides (cys)	Score solution	Score Carter *et al.* (42)	Cov(%) Shimaoka *et al* (40)	PTM
VHA-A	At1g78900	623 / 68.8	5.11	0	2038 (66)	34 (2)	157	222	65	
VHA-B1	At1g76030	483 / 54.1	4.98	0	1455 (68)	22 (2)	Not found	422	50	
VHA-B2	At4g38510	487 / 54.3	5.03	0	1364 (60)	20 (2)	140	152	44	
VHA-B3	At1g20260	468 / 52.1	4.84	0	1402 (65)	21 (3)	Not found	Not found	19	
VHA-C	At1g12840	370 / 42.0	5.4	0	998 (46)	17 (1)	139	117	44	AcNt
VHA-D	At3g58730	261 / 29.0	10.2	0	471 (33)	10	92	89	29	
VHA-E1	At4g11150	230 / 26.1	6.04	0	888 (75)	18 (5)	51	100	58	
VHA-E2	At3g08560	235 / 26.8	9.2	0	144 (12)	3	Not found	Not found	Not found	
VHA-E3	At1g64200	237 / 27.1	5.82	0	422 (32)	8 (2)	Not found	174	35	
VHA-F	At4g02620	128 / 14.3	6.08	0	250(52)	5	Not found	318	54	
VHA-G1	At3g01390	110 / 12.4	5.77	0	389 (66)	6	163	164	32	
VHA-G2	At4g23710	106 / 11.8	5.5	0	145 (23)	2	Not found	107	Not found	AcNt
VHA-H	At3g42050	441 / 50.3	6.58	0	857 (50)	15 (3)	81	98	19	AcNt
VHA-a2	At2g21410	821 / 93.1	5.39	7	1270 (32)	22 (1)	1048	90	17	
VHA-a3	At4g39080	843 / 95.2	5.66	6	2368 (47)	35 (1)	1920	185	21	
VHA-c1	At4g34720	165 / 16.6	8.67	3	Not found	Not found	Not found	Not found	10	
VHA-c2	At1g19910	164 / 16.6	8.67		Not found	Not found	97	Not found	Not found	
VHA-c3	At4g38920	165 / 16.6	8.67	3	Not found	Not found	Not found	91	Not found	
VHA-c5	At2g16510	164 / 16.6	8.61	3	172 (32)	2	Not found	Not found	Not found	
VHA-d1	At3g28710	351 / 40.8	5.04	0	579 (35)	11	Not found	78	46	
VHA-d2	At3g28715	343 / 39.8	4.99	0	539 (33)	10	Not found	148	47	

The table gives the protein acronym (subunit), gene designation (Gene), number of amino acid (AA) residues in the deduced translation product and the molecular mass (Mr), the predicted isoelectric point (pI), the number of transmembrane domains (TMD)

predicted by ARAMEMNON (52). Score gel & cov. (%) represents the Mascot score and protein sequence coverage using the in-gel digestion protocol. Nb gel peptide is the number of peptides assigned using the in-gel digestion protocol and (cys) represents the number of peptides found with cysteic acid. Score solution is the Mascot Score using the in-solution digestion protocol; Score Carter *et al*., the Mascot score found by Carter *et al.* (42); Cov (%) *Shimaoka et al*, sequence coverage found by *Shimaoka et al*. (40). PTM represents post translational modifications found using independent Mascot searches with semi-trypsin and N-terminal acetylation (AcNt) as additional parameters.

Table III: List of the ABC transporters identified on the tonoplast of vacuoles isolated from *Arabidopsis* cell cultures.

Name	Protein Id	Gene (TAIR)	AA	TMD	Topology	Score gel & cov. (%)	Nb gel peptides	Score sol & cov. (%)	Nb sol peptides	Score Carter *et al.* (42)	Cov (%) Shimaoka *et al* (40)
MDR SUBFAMILY (22 MEMBERS)											
AtMDR17	CAB71875	At3g62150	1292	11	(TMD-NBF)2	53 (2)	2	62	2	Not found	Not found
MRP SUBFAMILY (15 MEMBERS)											
AtMRP1	AAG51096	At1g30400	1622	16	(TMD-NBF)2	1980(26)	33	1549 (23)	27	80 (Not found)	5%
AtMRP2	AAC16268	At2g34660	1623	15	(TMD-NBF)2	1337 (23)	25	887 (14)	15	Not found	4%
AtMRP3	BAB01399	At3g13080	1515	15	(TMD-NBF)2	474 (8)	8	312 (6)	6	Not found	Not found
AtMRP4	AAC63634	At2g47800	1516	15	(TMD-NBF)2	2155 (27)	37	1738 (21)	30	55 (42)	5%
AtMRP5	AAC16754	At1g04120	1514	12	(TMD-NBF)2	657 (11)	11	474 (9)	9	Not found	Not found
AtMRP6	BAB01717	At3g21250	1294	10	(TMD-NBF)2	357 (7)	7	273 (7)	6	Not found	Not found
AtMRP7	CAD45086	At3g13100	1493	12	(TMD-NBF)2	Not found	Not found	86	1	Not found	Not found
AtMRP8	AAL14776	At3g13090	1466	13	(TMD-NBF)2	100(2)	2	Not found	0	Not found	Not found
AtMRP10	CAB83120	At3g62700	1539	15	(TMD-NBF)2	2613 (32)	46	2080 (29)	41	55 (171)	6%
AtMRP14	CAB86942	At3g59140	1453	14	(TMD-NBF)2		2	39 (2)	1	Not found	Not found
PDR SUBFAMILY (13 MEMBERS)											
AtPDR8	AAD39329	At1g59870	1469	13	(NBF-TMD)2	230 (5)	8	Not found		Not found	Not found
NAP SUBFAMILY (15 MEMBERS)											
AtNAP3	AAG52004	At1g67940	263	0	NBF	109 (9)	3	Not found		Not found	Not found
TAP SUBFAMILY (3 MEMBERS)											
AtTAP2	BAB10828	At5g39040	639	5	TMD-NBF	80 (5)	2	87 (3)	2	(38)	9%

The table gives the protein acronym (Name); protein identification number (Protein Id); gene designation in TAIR (Gene); number of amino acid residues in the deduced translation product (AA); TMD, putative transmembrane domains predicted by ARAMEMNON; Topology, number and orientation of transmembrane domain-nucleotide binding fold (TMD and NBF); Score Gel & cov, Mascot Score and protein sequence coverage (in %) obtained using the in-gel digestion protocol; Nb Gel peptides, the number of peptides assigned using the in-gel digestion protocol; Score Sol and cov., Mascot Score and protein sequence coverage (in %) obtained with the

in-solution digestion protocol; Nb Sol Peptides, number of peptides assigned with the in-solution digestion protocol; Score Carter *et al*., Mascot score found by Carter *et al*. (42); Cov (%) Shimaoka *et al*., sequence coverage found by Shimaoka *et al*. (40).

Table IV: Other transporters present on the tonoplast of *Arabidopsis* cell vacuoles

PROTEIN	Gene (Tair)	Mr (Da)	pI	TMD	Score sol & cov (%)	Score Gel	Family	type	Score Carter *et al.* (42)	Cov % Shimaoka *et al* (40)
Integral membrane protein	At1g75220	52867	8.7	12	206 (8)	239	MFS	2d Transp.	N F	N F
Integral membrane protein	At1g19450	52753	8.7	12	122 (7)	163	MFS	2d Transp.	91	N F
Sugar transporter	At2g48020	49667	9	12		34	MFS	2d Transp.	97	N F
Transporter-related similarity to D-xylose proton-symporter	At1g20840	79433	5.1	11	86 (4)	78	MFS	2d Transp.	91	N F
Sugar transporter family protein	At5g17010	53504	6.4	12	32 (2)		MFS	2d Transp.	N F	N F
Sugar transporter family protein	At3g03090					31	MFS	2d Transp.	N F	N F
Transporter-related, low similarity to hexose transporter	At4g35300	79674	5.4	10	354 (13)	354	MFS	2d Transp.	50	N F
Transporter-related, similar to cAMP inducible 2 protein and glycerol 3-phosphate permease	At4g17550	59037	7.1	13	82 (2)	124	MFS	2d Transp.	N F	N F
MATE efflux family protein	At3g21690	54916	4.2	12	127(4.5)	131	MOP	2d Transp.	N F	6
MATE efflux family protein	At3g26590	54286	5.7	12	119 (9)	103	MOP	2d Transp.	33	N F
MATE efflux family protein	At4g25640	53190	6.5	12	128 (9)	112	MOP	2d Transp.	N F	N F
MATE efflux protein	At5g52450	52439	8.1	12	166 (7)	168	MOP	2d Transp.	N F	N F
Ripening-responsive protein	At1g47530	52389	7.2	12	173 (9)	209	MOP	2d Transp.	N F	N F
Ripening-responsive protein	At1g75530	52426	6.4	11		65	MOP	2d Transp.	N F	N F
Nodulin MtN21 family protein	At1g75500	42544	9.2	10	204(7)	291	DMT	2d Transp.	N F	N F
Glucose-6-phosphate/phosphate translocator 2 (GTP2)	At1g61800	42726	9.6	8		52	DMT	2d Transp.	N F	N F
Mechanosensitive ion channel domain-containing protein	At1g78610	96543	8.7	6	68 (2)	102	DMT	2d Transp.	N F	N F
Patched family protein similar to Niemann-Pick C1 protein	At4g38350	116900	5.3	13	283 (6)	438	RND	2d Transp.	66	5
Proton-dependent oligopeptide transport family protein	At1g22570	62897	9.1	12		44	POT	2d Transp.	N F	N F
Proton-dependent oligopeptide transport family protein	At1g72140	61347	8.2	12		34	POT	2d Transp.	N F	N F
Peptide transporter (His transporter)(PTR2-B)	At2g02040	64380	5.5	10	567 (19)	807	POT	2d Transp.	103	7
Oligopeptide transporter OPT family protein, putative Fe(III)-phytosiderophore uptake mediator (YSL6)	At3g27020	73525	5.9	15	184 (9)	236	OPT	2d Transp.	N F	N F
Oligopeptide transporter OPT family protein, putative Fe(III)-phytosiderophore uptake mediator (YSL4)	At5g41000	73526	8.3	14	65 (3)		OPT	2d Transp.	N F	N F

NRAMP metal ion transporter 3 (NRAMP3)	At2g23150	56102	4.8	11	84 (11)	91	Nramp	2d Transp.	N F	4
Zinc transporter (ZAT1, MTP1)	At2g46800	43810	6.6	7	47 (2)	48	CDF	2d Transp.	N F	3
Sodium proton exchanger (NHX4, putative)	At3g06370	55570	7.4	14	93 (7)	92	CPA	2d Transp.	N F	N F
Metal transporter, putative (IRT3)	At1g60960	45060	6.2	7	47 (3)	47	ZIP	2d Transp.	N F	N F
Potassium transporter family protein (KUP7, HAK7, KT7, POT7)	At5g09400	95292	5.1	13	99 (4)	76	KuP	2d Transp.	N F	N F
Potassium transporter family protein (KUP5/KT5)	At4g33530	94677	5.2	13	40 (2)	44	KuP	2d Transp.	N F	N F
Potassium transporter / tiny root hair 1 protein/K carrier required for auxin transport (TRH1/KT3/KUP4)	At4g23640	86786	9.6	13	48 (3)	46	KuP	2d Transp.	N F	N F
Lysosomal Cystine Transporter	At5g40670	31007	7.4	6		37	LCT	2d Transp.	N F	N F
Cationic amino acid transporter (CAT9)	At1g05940	60138	7.7	15	51 (2)	52	APC	2d Transp.	N F	N F
Cationic amino acid transporter (CAT2)	At1g58030	67072	5.8	14	240 (12)	287	APC	2d Transp.	58	N F
Cationic amino acid transporter (CAT4)	At3g03720	63597	5.8	14	145 (6)	180	APC	2d Transp.	40	N F
Cationic amino acid transporter (CAT8)	At1g17120	64863	8.26	12		32	APC	2d Transp.	N F	N F
Amino acid transporter family protein	At3g30390	49485	6.1	11	50 (3)	50	AAAP	2d Transp.	30	N F
Amino acid transporter family protein	At2g41190	58825	4.8	10	179 (6)	185	AAAP	2d Transp.	N F	N F
Amino acid transporter family protein	At2g39130	60070	5.6	10		62	AAAP	2d Transp.	N F	N F
Amino acid transporter family protein	At2g40420	47639	6.3	10	62 (5)		AAAP	2d Transp.	N F	N F
Amino acid transporter family protein	At3g28960	44645	8.9	10		32	AAAP	2d Transp.	N F	N F
Proton-dependent concentrative adenosine transporter (ENT1)	At1g70330	49312	6.8	11	74 (4)	75	ENT	2d Transp.	N F	N F
ADP, ATP carrier protein 2	At5g13490	41720	9.8	3		198	MC	2d Transp.	N F	N F
Putative ADP, ATP carrier protein 1	At3g08580	41449	9.8	3		424	MC	2d Transp.	N F	N F
Mitochondrial phosphate transporter (PHT3-1)	At5g14040	40064	9.3	3		83	MC	2d Transp.	N F	N F
Peroxisomal membrane protein (PMP36)	At2g39970	36190	9.9	3		43	MC	2d Transp.	N F	N F
Dicarboxylate/tricarboxylate carrier (DTC)	At5g19760	31891	9.4	2		324	MC	2d Transp.	N F	N F
Putative mitochondrial substrate carrier family protein	At4g01100	38301	9.6	1		155	MC	2d Transp.	N F	N F
Putative plant uncoupling mitochondrial protein	At3g54110	32641	9.6	1		227	MC	2d Transp.	N F	N F
Two-pore calcium channel (TPC1)	At4g03560	84819	4.9	10	268 (6)	413	VIC	Ion Channels	167	N F
Zinc finger (DHHC type) family protein	At3g51390	39202	8.4	4	45	45	VIC	Ion Channels	N F	N F
Major intrinsic family protein gamma (TIP1.1)	At2g36830	25604	6	7		88	MIP	Ion Channels	N F	N F
Tonoplast intrinsic protein gamma (TIP1.2)	At3g26520	25832	4.7	7		40	MIP	Ion Channels	108	N F

Tonoplast integral protein delta (TIP2.1)	At3g16240	25011	5.3	7	311 (21)	447	MIP	Ion Channels	234	N F
Major intrinsic family protein delta (TIP2.2)	At4g17340	25064	4.5	6		45	MIP	Ion Channels	108	N F
Chloride channel protein (CLC-c)	At5g49890	85177	8.7	10		79	CLC	Ion Channels	N F	N F
Putative plasma membrane P3A-type H+-ATPase (AHA2)	At4g30190	104335	6.5	10	117 (2)	159	P-ATPase	ATP-Dep	N F	N F
Putative Ca2+-transporting P2B-type ATPase (ACA11)	At3g57330	111874	6	11	593 (11)	731	P-ATPase	ATP-Dep	91	2
Putative plasma membrane P3A-type H+-ATPase (AHA9)	At1g80660	105142	6	10	51 (2)	84	P-ATPase	ATP-Dep	N F	N F
Calmodulin-regulated Ca2+-ATPase 4 (ACA4)	At2g41560	112678	5.8	9	200 (4)	196	P-ATPase	ATP-Dep	105	N F
Ca2+/Mn2+-transporting P2A-type ATPase (ECA1/ACA3)	At1g07810	116291	5	8		38	P-ATPase	ATP-Dep	N F	N F
PoriN Family protein	At3g20000	34228	6.3	b		177	MPT	ATP-Dep	N F	N F
Vacuolar proton-translocating pyrophosphatase (AVP-3)	At1g15690	80768	5.1	15	1948 (32)	2193	H-PPase	ATP-Dep	88	17
Copper transporter family	At5g20650	15774	9.2	2	75 (13)	86	CTR2	UNK	58	N F
Expressed protein	At1g73650	32939	9.1	6	177 (11)	194	UNC	UNK	68	N F
NoduliN Family protein	At5g14120	63108	7.2	12		108	UNC	UNK	68	N F
Epressed protein	At1g64650	50625	6.2	11		31	UNC	UNK	N F	N F
Expressed protein	At2g20230	29702	5.1	4	48 (5)				N F	N F
Endomembrane protein 70, putative	At5g10840	74418	8.2	10		103	UNC	UNK	N F	N F
Endomembrane protein 70, putative	At4g12650	59670	5.9	10		39	UNC	UNK	N F	N F
Endomembrane protein 70, putative	At3g13772	74185	8.7	9		59	UNC	UNK	N F	N F
Endomembrane protein 70, putative	At2g01970	68005	6.7	9		71	UNC	UNK	N F	N F
Transmembrane protein-related (TOM1)	At4g21790	32988	8.7	7	73(1)	72	UNC	UNK	N F	N F
Putative drought-induced protein	At1g53210	63376	5.2	10	574 (21)	703	UNC	UNK	88	N F
Early-responsive to dehydration stress protein (ERD4)	At1g30360	81883	9.3	10	121 (5)	225	UNC	UNK	64	N F
Putative outer envelope protein	At3g46740	89133	8.9	β		151	Omp IP	UNK	N F	N F
Porin, putative	At5g67500	29577	8.9	β		155	MPP	UNK	N F	N F
Porin, putative	At3g01280	29407	8.8	β		356	MPP	UNK	103	29
Porin, putative, Voltage-dependent anion-selective channel protein	At5g15090	29193	7.9	β		464	MPP	UNK	99	35
Porin	At3g20000	34250	6.8	β		177	MPP	UNK	N F	N F

The table gives the protein acronym (protein), gene designation (Gene), deduced molecular mass Mr (Da); theoretical isoelectric point (pI); the number of transmembrane domain given by ARAMEMNON (52) (TMD); Mascot score and protein sequence coverage

obtained using the in-solution digestion protocol (Score sol & cov (%)); Mascot Score obtained using the in-gel digestion protocol (Score gel); Transporter families and types were classified according to Transport DB: a relational database of cellular membrane transport systems (http://www.membranetransport.org) proposed by Ren *et al.* (92). MFS, Major Facilitator ; MOP, Multidrug/Oligosaccharidyl-lipid/Polysaccharide; DMT, Drug/Metabolite Transporter; RND, Resistance-Nodulation-Cell Division; POT, Proton-dependent Oligopeptide Transporter; OPT, Oligopeptide Transporter; Nramp, Metal Ion (Mn^{2+}-iron) Transporter; CDF, Cation DiffusioN Facilitator; CPA, Monovalent Cation:Proton Antiporter; ZIP, Zinc (Zn^{2+})-Iron (Fe^{2+}) Permease; KuP, K^+ Uptake Permease; BASS, Bile Acid: Na^+ Symporter; APC, Amino Acid-Polyamine-Organocation; AAAP, Amino Acid/Auxin Permease; ENT, Equilibrative Nucleoside Transporter; MC, Mitochondrial Carrier; VIC, Voltage-gated Ion Channel; MIP, Major Intrinsic Protein; CLC, Chloride Channel; P-ATPase, P-type ATPase; MPT, Mitochondrial Protein Translocase; H-PPase, H^+-translocating Pyrophosphatase; CTR2, Copper Transporter; UNK, unknown; Omp IP, Outer Membrane Protein Insertion Porin; MPP, Mitochondrial and Plastid Porin. Score Carter *et al*. represents Mascot score found by Carter *et al*. (42); Shimoaka *et al*. Cov %, sequence coverage found by Shimaoka *et al*. (40). N F., Not Found.

Table V Stomatin ProhibitiN Flotilin Hbc (SPFH) family proteins (SPFH domain proteins)

PROTEIN	Gene (Tair)	Accession	Mr (Da)	pI	TMD or PTM	Score gel	Cov (%)	Nb peptides	Score Carter *et al*. (42)	Cov (%) Shimaoka *et al* (40)	Function
Band 7 family protein	At2g03510	NP_027545	40571	5.65	1	328	18	6	Not found	Not found	Unclassified
Band 7 family protein	At3g01290	NP_566135	31301	5.67	N-myristoylation	269	23	6	Not found	12	Unclassified
Band 7 family protein	At5g51570	NP_199970	32358	5.28	N-myristoylation	495	39	10	298	39	Stress response
Band 7 family protein	At5g62740	NP_201080	31411	5.29	N-myristoylation	641	55	13	237	29	Stress response
Band 7 family protein	At1g69840	NP_974117	31386	5.3	N-myristoylation	214	18	4	98	14	Unclassified
Expressed protein	At5g25250	NP_197907	52292	5.83	N-myristoylation	183	7	3	Not found	Not found	Unclassified
Prohibitin	At5g40770	NP_198893	30381	6.99	N-myristoylation	473	38	7	109	24	Stress response
Prohibitin, putative	At2g20530	NP_179643	31617	9.65	1	218	12	2	Not found	Not found	Miscellaneous
Prohibitin, putative	At4g28510	NP_194580	31687	9.26	*	373	30	6	Not found	7	Miscellaneous
Prohibitin, putative	At1g03860	NP_171882	24881	9.09	N-myristoylation	305	22	4	Not found	8	Stress response
Prohibitin, putative	At3g27280	NP_189364	30619	6.93	*	457	38	7	Not found	10	Stress response

The table gives the protein acronym (Protein); gene designation (Gene); protein identification number (Accession), deduced molecular mass (Mr (kDa)); Theoretical isoelectric point (pI); the number of transmembrane domains (TMD) or post translational modification (PTM). Myristoylation sites were predicted using http://plantsp.sdsc.edu/myrist.html and Psort II (54), Score gel represents Mascot score of in-gel digestion protocol; Cov (%) protein sequence coverage; Nb peptides, number of peptides assigned using the in-gel digestion protocol; Score Carter *et al*., Mascot score found by Carter *et al*.(42); Cov %, Shimaoka *et al*. sequence coverage found by Shimaoka *et al*. (40); Function, possible function assigned by The MIPS Functional Catalogue (FunCat) (http://mips.gsf.de/proj/funcatDB/search_main_frame.html).

FIGURE LEGENDS

Figure 1 Major protein constituents and evaluation of purity of vacuoles isolated from *Arabidopsis* cell culture. (A) SDS-PAGE analysis of proteins (15 μg) isolated from protoplasts (P) and purified vacuoles (V) separated on 12% acrylamide gel stained with Coomassie blue (R250). The identification of the proteins present in the most intensely stained bands was determined by LC-MS/MS analysis. Numbers, on the left, represent the size of the molecular weight markers (MW) in kDa. (B) Enrichment and purity of vacuole samples were estimated by western blots. The vacuolar markers, Tonoplast Intrinsic Protein (TIP) (α and γ isoforms), were revealed using specific antibodies; cross-contaminations were evaluated using antibodies raised against: the outer envelop protein 21 (OEP 21) and the light harvesting complex b (LHC) of the chloroplast; the preprotein translocase of the mitochondrial outer membrane (TOM 40); the HDEL domain of the endoplasmic reticulum proteins and the plasma membrane P-type H^+-ATPase.

Figure 2 Strategy used for the identification of vacuolar proteins purified from *Arabidopsis* cell culture. Vacuoles were purified from *Arabidopsis thaliana* suspension cultures. After vacuole purification, proteins were either separated by 12% SDS-PAGE and stained with Coomassie blue (A) or subjected to centrifugation to obtain 2 different enriched fractions: membranes (B, C) and vacuolar sap (D). Major bands from the whole vacuole SDS-PAGE migration were cut out and subjected to in-gel trypsin digestion (A, Fig. 1A). Proteins from the membrane fraction were either subjected to in-solution trypsin digestion (B) (see text and experimental procedures) or separated by a short SDS-PAGE migration and digested in-gel (C). Proteins from the soluble fraction were digested in-gel following a short SDS-PAGE migration (D). Peptides from (A) to (D) were separated by liquid chromatography prior to MS/MS analysis.

Figure 3 Cross-correlation of the different proteome analyses of the vacuolar membrane system. (A) This Venn diagram presents the proteins identified in our study with those presented in the different published studies. The combined data from Shimaoka *et al*., Szponarski *et al*. and Carter *et al*. and our dataset identified 815 non redundant proteins. Overlap of the different protein sets is shown. Numbers in parentheses indicate the total number of proteins found by a

particular study. **(B)** This Venn diagram presents the overlap of transporters and H^+-ATPase subunits identified in our study with those presented in the different published studies. The combined data from Shimaoka *et al*., Szponarski *et al*. and Carter *et al.* and our dataset identified 123 non redundant transporters and related proteins.

Figure 4 Sub-cellular localisation of selected proteins by transient expression of GFP-fusion proteins. GFP-fusion proteins were expressed either in tobacco leaf cells (A to H) or in *Arabidopsis* protoplasts (I to K) (1 represents the GFP fluorescence alone and 2 the overlay of transmission and GFP fluorescence). Nramp3-GFP (A1 & A2) and TIP2.1-GFP constructions (B1) were used as vacuolar protein controls and GFP-GUS protein (C1) as a cytosolic protein control. The tested fusion proteins were: (D1) Dwarf1, At1g19820; (E1) a band 7 family protein, At1g69840; (F1) putative sugar transporter, At1g19450; (G1) lipocalin, At5g58070 and (H1) CCD1, At3g63520. I-K: transient expression in *Arabidopsis* protoplasts, lipocalin (I & J), putative sugar transporter (K). The bar corresponds to 15 µm.

Figure 1

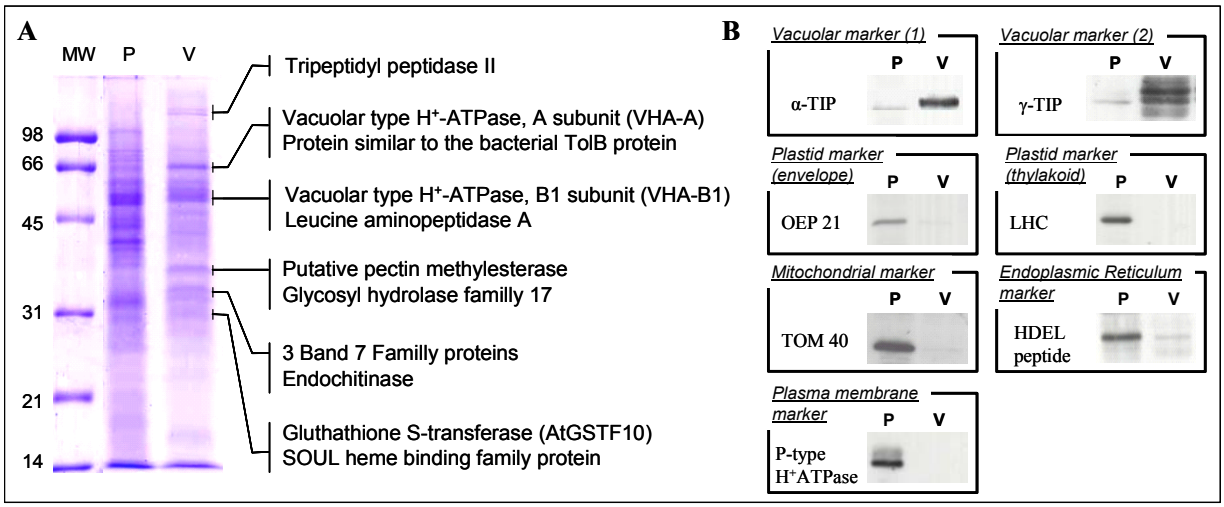
A
MW
P
V
98
66
45
31
21
14
Tripeptidyl peptidase II
Vacuolar type H+-ATPase, A subunit (VHA-A)
Protein similar to the bacterial TolB protein
Vacuolar type H+-ATPase, B1 subunit (VHA-B1)
Leucine aminopeptidase A
Putative pectin methylesterase
Glycosyl hydrolase familly 17
3 Band 7 Familly proteins
Endochitinase
Gluthathione S-transferase (AtGSTF10)
SOUL heme binding family protein
B
Vacuolar marker (1)
α-TIP
Vacuolar marker (2)
γ-TIP
Plastid marker (envelope)
OEP 21
Plastid marker (thylakoid)
LHC
Mitochondrial marker
TOM 40
Endoplasmic Reticulum marker
HDEL peptide
Plasma membrane marker
P-type H+ATPase
P
V

Figure 2

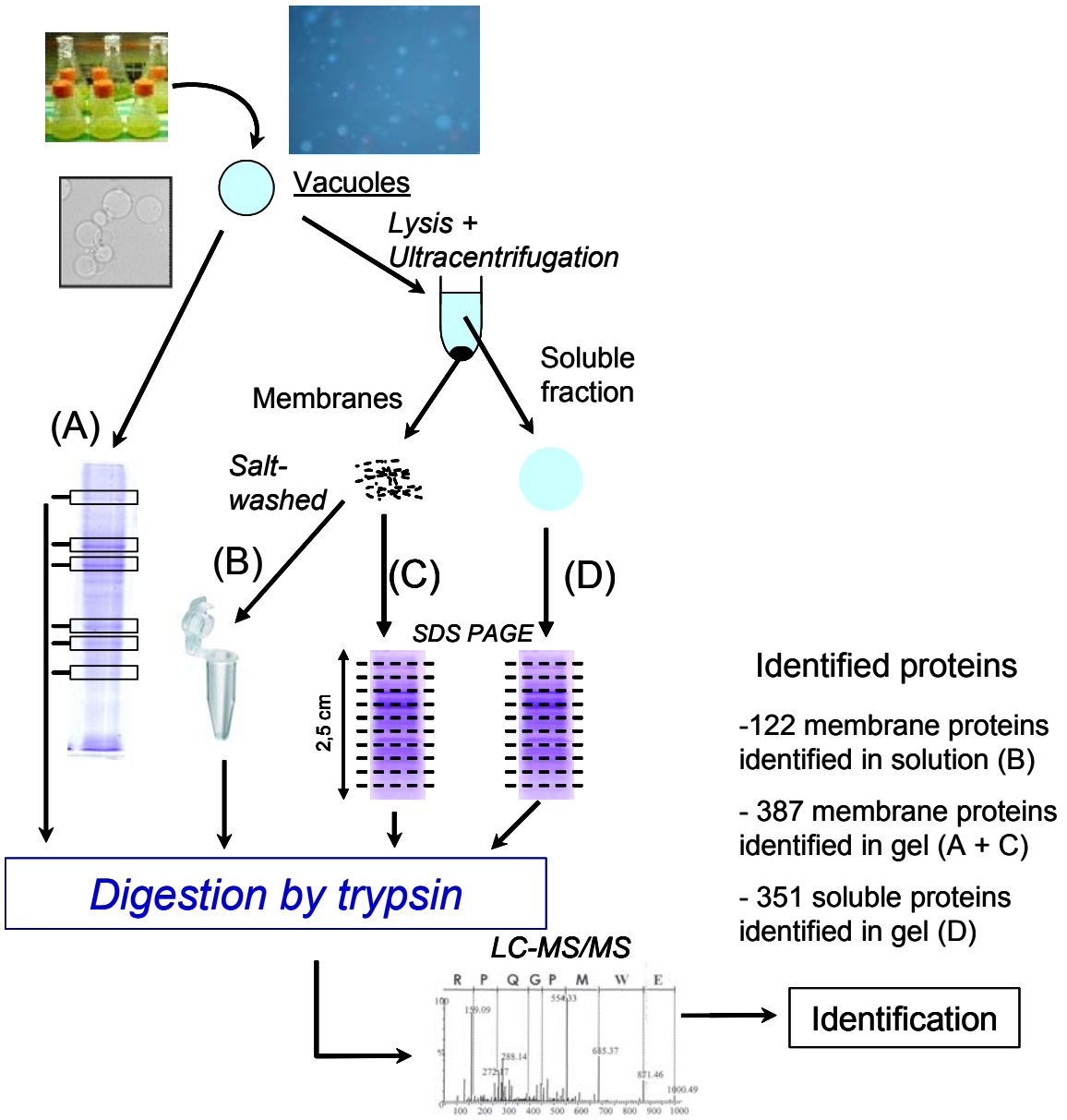
Vacuoles
Lysis +
Ultracentrifugation
Soluble
fraction
Membranes
(A)
Salt-
washed
(B)
(C)
(D)
SDS PAGE
2,5 cm
Digestion by trypsin
LC-MS/MS
Identification
Identified proteins
-122 membrane proteins identified in solution (B)
- 387 membrane proteins identified in gel (A + C)
- 351 soluble proteins identified in gel (D)

Figure 3

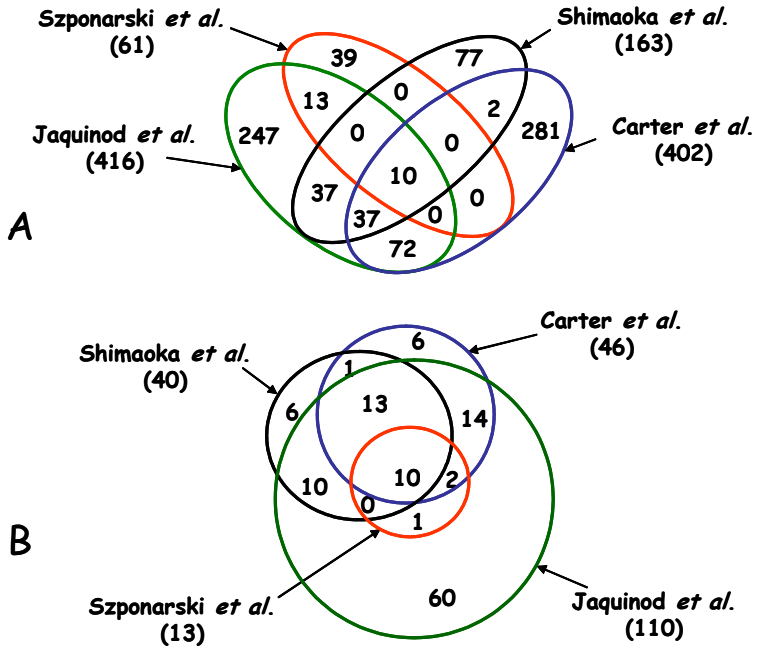
Szponarski et al.
(61)
Shimaoka et al.
(163)
Jaquinod et al.
(416)
Carter et al.
(402)
39
77
13
0
2
247
0
0
281
10
37
0
37
0
72
A
Carter et al.
(46)
Shimaoka et al.
(40)
6
1
13
6
14
10
10
2
0
1
60
Szponarski et al.
(13)
Jaquinod et al.
(110)
B

Figure 4