# An Extended Proteome Map of the Lysosomal Membrane Reveals Novel Potential Transporters*[S]

Agnès Chapel‡§¶, Sylvie Kieffer-Jaquinod‡§¶, Corinne Sagné‖, Quentin Verdon‖§§, Corinne Ivaldi‡§¶, Mourad Mellal‡§¶, Jaqueline Thirion**, Michel Jadot**, Christophe Bruley‡§¶, Jérôme Garin‡§¶, Bruno Gasnier‖, and Agnès Journet‡§¶‡‡

**Lysosomes are membrane-bound endocytic organelles that play a major role in degrading cell macromolecules and recycling their building blocks. A comprehensive knowledge of the lysosome function requires an extensive description of its content, an issue partially addressed by previous proteomic analyses. However, the proteins underlying many lysosomal membrane functions, including numerous membrane transporters, remain unidentified. We performed a comparative, semi-quantitative proteomic analysis of rat liver lysosome-enriched and lysosome-nonenriched membranes and used spectral counts to evaluate the relative abundance of proteins. Among a total of 2,385 identified proteins, 734 proteins were significantly enriched in the lysosomal fraction, including 207 proteins already known or predicted as endo-lysosomal and 94 proteins without any known or predicted subcellular localization. The remaining 433 proteins had been previously assigned to other subcellular compartments but may in fact reside on lysosomes either predominantly or as a secondary location. Many membrane-associated complexes implicated in diverse processes such as degradation, membrane trafficking, lysosome biogenesis, lysosome acidification, signaling, and nutrient sensing were enriched in the lysosomal fraction. They were identified to an unprecedented extent as most, if not all, of their subunits were found and retained by our screen. Numerous transporters were also identified, including 46 novel potentially lysosomal proteins. We expressed 12 candidates in HeLa cells and observed that most of them colocalized with the lysosomal marker LAMP1, thus confirming their lysosomal residency. This list of candidate lysosomal proteins substantially increases our knowledge of the lysosomal membrane and provides a basis for further characterization of lysosomal functions.** *Molecular & Cellular Proteomics 12: 10.1074/mcp.M112.021980, 1572–1588, 2013.*

Lysosomes are membrane-bound intracellular organelles that are key players in the degradation and recycling of biological material. Their crucial role in cell physiology is underlined by the existence of ~50 lysosomal storage diseases caused by genetic defects in lysosomal proteins or proteins involved in lysosome biogenesis (1). The degradative function is carried out in the lysosomal lumen by the concerted action of over 60 hydrolases and accessory proteins (2). Although these soluble lysosomal proteins have been extensively studied, knowledge about membrane proteins remains rather limited, despite the multiple and crucial functions fulfilled by the membrane. It is indeed responsible for establishing and maintaining pH and ionic gradients, transporting degradation substrates and products from/into the cytosol, and maintaining lysosome integrity. Additionally, the lysosomal membrane is subjected to multiple fusion and fission events with other endocytic or biosynthetic compartments. Substrates for degradation are conveyed to lysosomes from the extracellular milieu, the plasma membrane, or the cytoplasm through the endocytic, phagocytic, and autophagic routes. Delivery of newly synthesized material to lysosomes requires exchanges between endocytic or biosynthetic organelles on the one hand and lysosomes on the other hand. These numerous trafficking events are supported by molecular machineries that associate with the lysosomal membrane (3).

In the last decade, large scale mass spectrometry-based approaches have been exploited to study the lysosome protein composition. The soluble content has first been analyzed by the use of an affinity purification protocol based on the mannose 6-phosphate modification (4–11) that is characteristic of soluble lysosomal proteins (12). This has resulted in the identification of about 60 known luminal lysosomal proteins, as well as of many mannose 6-phosphate-containing proteins that were not previously thought to carry out a lysosomal

From the ‡Commissariat à l'Energie Atomique, Institut de Recherches en Technologies et Sciences du Vivant, Laboratoire Biologie à Grande Echelle, F-38054 Grenoble, France, §INSERM, U1038, F-38054 Grenoble, France, the ¶Université Joseph Fourier, Grenoble 1, F-38000, France, the ‖Université Paris Descartes, Sorbonne Paris Cité, CNRS, UMR 8192, Centre Universitaire des Saints-Pères, 45 Rue des Saints-Pères, F-75006 Paris, France, §§Graduate School ED 419, Université Paris-Sud 11, Hôpital Bicêtre, F-94276 Le Kremlin Bicêtre France, and the **Unité de Recherche en Physiologie Moléculaire, Namur Research Institute for Life Sciences, University of Namur (FUNDP), 61, Rue de Bruxelles B,-5000, Namur, Belgium

Received July 14, 2012, and in revised form, February 1, 2013
Published, MCP Papers in Press, February 24, 2013, DOI 10.1074/mcp.M112.021980

function (13). To gain insight into the membrane composition, several groups have used preparative subcellular fractionation to recover samples enriched in lysosomes (14–18). Despite the experimental limitations of the latter methods that are unable to completely separate organelles, the use of comparative strategies and statistical tools (14, 16, 17) allowed the identification of novel putative resident lysosomal membrane proteins, including a few potential transporters, such as SLC12A4, SLC44A2, C19ORF28 (MFSD12), SIDT2, and MFSD1 (14, 16). More recently, the coupling of selective lysosome density shift and MS quantification was shown to allow simultaneous identification and validation of lysosomal candidates (19). The efficiency of these various approaches in identifying candidates was highlighted by the demonstration of the effective lysosomal residency of several selected proteins (16, 20–29). Concerning membrane proteins, these studies have led to a list of about 45 integral membrane lysosomal proteins for which evidence of the lysosomal localization has been obtained by at least overexpression of epitope-tagged fusion proteins (30).

However, despite the expanded knowledge provided by these recent studies, many lysosomal actors are still missing. For instance, although more than 20 lysosomal transport activities have been biochemically described (31, 32), many of these transport functions remain orphans because the underlying proteins have not been identified yet (33). The aim of the present proteomic study was to gain deeper insight into the characterization of the lysosomal membrane and its associated proteins, with a particular interest in novel potential lysosomal transporters, given their major role in lysosomal physiology. Transporters are integral membrane proteins (IMPs)[1] displaying multiple transmembrane domains, and such IMPs are usually difficult to identify by mass spectrometry because of their high hydrophobicity and low abundance (34, 35). Therefore, to extend our protein identification capacities, we used a combination of subcellular and biochemical fractionations prior to MS analysis. We first established an overall list of 2,385 gene products from lysosome-enriched and lysosome-nonenriched fractions. Then, a comparative proteomics analysis based on spectral counts led to the selection of 734 candidate proteins. They included on the one hand 94 novel potentially lysosomal proteins and, on the other hand, 46 established or putative transporters for which lysosomal residency is suggested by this study. The lysosomal localization has been validated for nine candidates, including five transporters. Moreover, we recently showed elsewhere that another candidate identified during this proteomic study, PQLC2, is a novel lysosomal amino acid transporter (36).

[1] The abbreviations used are: IMP, integral membrane protein; CM, chloroform/methanol; Nsc, normalized spectral count; SpI, spectral index; PM, plasma membrane; EL, endo-lysosome; MFS, major facilitating superfamily; SLC, solute carrier; mTOR, mammalian target of rapamycin; CMI, CM-insoluble; CMS, CM-soluble; ER, endoplasmic reticulum; TGN, *trans*-Golgi network; L, light mitochondrial.

## EXPERIMENTAL PROCEDURES

*Subcellular Fractionation*—All experiments involving rats were conducted in compliance with approved Institutional Animal Use Committee protocols. Livers were obtained from male Wistar rats. Each preparation was performed on four rat livers essentially as described previously (37). Briefly, fractionation of subcellular organelles by differential centrifugation produced nucleus and heavy mitochondrial (NM), light mitochondrial (L), and microsomal and soluble (PS) fractions. The L fraction was subjected to isopycnic centrifugation on a discontinuous Nycodenz density gradient. Conditions of the gradient were essentially the same as described in the original publication (37), except that Nycodenz® was used instead of metrizamide. The following density layers were successively loaded on top of the L fraction: 1.16 (7 ml), 1.145 (6 ml), 1.135 (7 ml), and 1.10 (7 ml). Centrifugation was performed at 83,000 × *g* for 2 h 30 min in an SW28 Beckman rotor. Fraction 2 (the interface between the layers of respective densities, 1.10 and 1.135 g/ml) was recovered as the L+ ("lysosome-enriched") fraction. Fractions 1, 3, and 4 (upper and lower fractions) were pooled as the L− ("lysosome-nonenriched") fraction. Organelles from both L+ and L− fractions were separately diluted in 0.25 M sucrose, pelleted by ultracentrifugation (100,000 × *g*, 4°C, 1 h), and subjected to a hypoosmotic shock in buffer A (10 mM Hepes, pH 7.8, supplemented with protease inhibitors (Complete, Roche Applied Science)). Membranes (MbL+ and MbL−) were recovered by ultracentrifugation (100,000 × *g*, 4°C, 1 h), extensively washed in buffer A, and resuspended in 200 μl of buffer A before storage at −80°C.

Recovery of lysosomes in the fractions resulting from differential centrifugation and from the Nycodenz gradient was followed by β-galactosidase activity measurement (38). These data along with the protein amounts recovered in each fraction allowed calculation of purification factors as compared with the initial homogenate. Protein concentration was evaluated using a Micro BCA™ protein assay kit (Thermo Scientific).

*Chloroform/Methanol Extraction*—Chloroform/methanol (CM) fractionation of proteins was performed according to Salvi *et al.* (39). Briefly, 250 μg of organelle membranes (1–10 mg/ml) in buffer A were sonicated, left for 15 min on ice, and ultracentrifuged for 40 min at 100,000 × *g* and at 4°C. Membranes pellets were then gently resuspended in 100 μl of buffer A and slowly diluted in 900 μl of cold CM (5:4, v/v) on ice. The mixture was left 15 min on ice, with periodic agitation, and then centrifuged (15 min, 15,000 × *g*, 4°C) to produce a pellet (the CM-insoluble fraction, CMI) and a supernatant (the CM-soluble fraction, CMS, containing the most hydrophobic proteins). Solvent from the CMS fraction was evaporated under nitrogen, down to 100 μl, and proteins were acetone-precipitated. Proteins from the CMI and CMS pellets were dissolved in Laemmli buffer for SDS-PAGE separation.

*Triton X-114 Phase Separation*—Triton X-114 phase separation was performed according to Donoghue *et al.* (40). Briefly, 250 μg of organelle membranes (1–10 mg/ml) in buffer A were sonicated, left for 15 min on ice, and ultracentrifuged for 40 min, at 100,000 × *g* and at 4°C. Membranes pellets were then gently resuspended in 800 μl of cold PBS, and 200 μl of 10% Triton X-114 (Sigma-Aldrich) was added. The mixture was gently agitated on a rotating wheel overnight at 4°C and cleared by centrifugation (30 min, 20,000 × *g*, 4°C). The supernatant was warmed at 37°C for 30 min and centrifuged (30 min, 5,000 × *g*, 25°C) for phase separation. The upper aqueous phase (AQ) and lower detergent phase (DT) were mixed, respectively, with 200 μl of 10% Triton X-114 and 800 μl of cold PBS, incubated for 15 min at 37°C, and centrifuged as described previously. This step was repeated three times, before recovering the final AQ and DT phases as well as a pellet present in the DT fraction (DTP). AQ and DT proteins were acetone-precipitated, and all samples were finally dissolved in Laemmli buffer for SDS-PAGE separation. AQ samples from the first replicate were not kept.

*SDS-PAGE Protein Separation and Western Blots*—For Western blots, proteins were separated by SDS-PAGE as described previously (41), transferred to PVDF membranes (Immobilon P, Millipore), and immunodetected with the following antibodies directed against subcellular organelles markers: monoclonal mouse anti-rat LAMP2, 1:5 (10D10; homemade); rabbit polyclonal anti-OSCP, 1:50,000 (kind gift from G. Brandolin), rabbit polyclonal anti-GRP78, 1:250 (BiP; Abcam, ab2902); mouse monoclonal anti-TGN38, 1:250 (Transduction Laboratories, T69020); rabbit polyclonal anti-58K protein, 1:500 (FTCD; Abcam, ab5820); rabbit polyclonal anti-Rab5, 1:2,000 (StressGen, KAP-GP006); and mouse monoclonal anti-$\alpha$1 sodium potassium ATPase, 1:5,000 (Abcam, ab7671). Proteins were revealed with the Western Lightning Plus-ECL reagent (PerkinElmer Life Sciences) and visualized on autoradiography films (Kodak Biomax XAR).

For MS analysis, SDS-PAGE separation of the reduced proteins was performed on 4–12% gradient acrylamide gels (NuPAGE, Invitrogen). Proteins were stained by Bio-Safe Coomassie stain or Coomassie Brilliant Blue R-250 (Bio-Rad). The amount of loaded proteins and the migration length were adapted to the protein sample complexity.

*MS Sample Preparation*—For protein digestion each SDS-polyacrylamide gel lane was systematically cut into 1-mm bands that were washed several times by successive incubations in 25 mM $NH_4HCO_3$ for 15 min and in 50% (v/v) acetonitrile, 25 mM $NH_4HCO_3$ for 15 min. Gel pieces were dehydrated by 100% acetonitrile and then incubated with 7% $H_2O_2$ for 15 min before being washed again with the destaining solutions described above. 0.15 $\mu$g of modified trypsin (Promega, sequencing grade) in 25 mM $NH_4HCO_3$ was added to the dehydrated gel pieces for an overnight incubation at 37°C. Peptides were extracted from gel pieces in three 15-min sequential extraction steps in 30 $\mu$l of 50% acetonitrile, 30 $\mu$l of 5% formic acid, and 30 $\mu$l of 100% acetonitrile. The pooled supernatants were finally dried under vacuum.

*NanoLC-MS/MS Analysis*—The dried extracted peptides were resuspended in 30 $\mu$l in 4% acetonitrile, 0.5% trifluoroacetic acid and analyzed by on-line nanoLC-MS/MS (Ultimate 3000 and LTQ-Orbitrap, Thermo Fisher Scientific). The nanoLC method consisted of a 40-min gradient ranging from 5 to 55% acetonitrile in 0.1% formic acid at a flow rate of 300 nl/min. Peptides were sampled on a 300-$\mu$m $\times$ 5-mm PepMap C18 precolumn and separated on a 75-$\mu$m $\times$ 150-mm C18 column (Gemini C18, Phenomenex). MS and MS/MS data were acquired using Xcalibur (Thermo Fischer Scientific) and processed automatically using Mascot Daemon software (version 2.1, Matrix Science).

*Database Searching and Criteria for Protein Identification*—Consecutive searches against the IPI_rat_decoy_database (based on the IPI-Rat version 3.48 database; 80,082 entries including the reverse ones) were performed for each sample using Mascot 2.1 (Matrix Science, London, UK). ESI-TRAP was chosen as the instrument and trypsin as the enzyme, and two missed cleavages were allowed. Precursor and fragment mass error tolerance were set respectively at 15 ppm and 1 Da. Peptide variable modifications allowed during the search were: acetyl (N-terminal), oxidation (M), dioxidation (M), and trioxidation (C). Proteins identified with a minimum of one unique peptide and with a score higher than the query threshold (for a $p$ value of peptide $<0.01$) were automatically validated using IRMa (42). The filtered results were downloaded into an MS identification database, in which the peptide false discovery rate (FDR) was of 2.38%. (FDR = 2 $\times$ reverse/(reverse + forward)). A homemade tool[2] was used for the compilation, grouping of proteins identified by a same set or subset of peptides (according to the principle of parsimony) and final comparison. Peptides shorter than hexamers were rejected at the grouping step. A last filtering step retained only protein groups identified by at least two unique peptides. All keratin isoforms and trypsin were deleted from the results. All MS data are available on the Pride database site (43) as Pride project 22847.

*Protein Annotation*—Gene names were retrieved from the IPI-Rat or Uniprot databases. Uncharacterized proteins (IPI sequence set) underwent a Blastp process against the mammalian Uniprot database section (released February, 2011). Top protein hits with at least 10e-05 e-value and a query coverage greater than 50% were kept and manually checked for relevance. The query coverage represents the percent of the query length that is included in the aligned segments and is calculated over all segments. When several protein groups corresponded to the same gene name, they were all kept.

The TMHMM version 2.0 server (Center for Biological sequence analysis, Lyngby, DK) was used for predictions of membrane-spanning regions (*i.e.* transmembrane domains). Protein functional annotation and subcellular localization information, either experimental or predicted, were collected from the IPI, Uniprot, or QuickGO sites and from the bibliography.

*Spectral Counting and Semi-quantitative Analysis*—For each identified protein $p$, the spectral count values ($sc_{p,s}$ = number of spectra assigned per protein in a given sample $s$) were determined with the homemade hEIDI software (supplemental Tables S1 and S2). All spectra pointing to a given protein after the filtering steps were considered. Spectra matching the protein isoforms were counted once for each protein group containing one or several of the isoforms. These spectral count values were then normalized to the equivalent amount (in micrograms) of total membrane protein prior to CM or Triton X-114 extraction (Fig. 1), which had been injected in the spectrometer. The normalized spectral count (Nsc) thus corresponds to a number of spectra per microgram of total MbL+ or MbL− proteins. For each identified protein $p$, the Nsc value was first calculated for each sample $s$ ($Nsc_{p,s}$), then for each fraction $f$ (MbL+ or MbL−; $Nsc_{p,f}$) in each replicate, as the sum of the $Nsc_{p,s}$ in the CMS, CMI, AQ, DTP, and DT samples and at last for each of the MbL+ or MbL− fractions by summing the $Nsc_{p,f}$ obtained for the three replicates. Evaluation of the relative abundance of a protein in a given sample or fraction was based on the label-free spectral counting method (44), and performed by dividing $Nsc_{p,s}$ or $Nsc_{p,f}$ by the total Nsc of the considered sample or fraction.

For each protein, a spectral index (SpI) comprising both relative protein abundance and number of samples containing this protein was then calculated as indicated in Fu *et al.* (45) to allow comparison between MbL− and MbL+ samples. Confidence intervals were established through permutation analysis (45) and used for determination of proteins significantly enriched in MbL+ (lysosomal protein candidates).

*Molecular Cloning*—IMAGE or ORFEOME clones coding for the following proteins were obtained from Source Bioscience: LOH12CR1, MFSD1, PTTG1IP, SLC37A2, SLC38A7, SLC46A3, SLCO2B1, STARD10, TMEM104, TMEM175, TTYH2, and TTYH3. Inserts were amplified by PCR using the Phusion polymerase (New England Biolabs) and the commercial plasmids as template and cloned for heterologous expression of GFP or YFP fusion proteins. Original plasmids, DNA accession numbers, primers, and expression vectors are given in supplemental Table S3.

*Cell Culture and Fluorescence Studies*—HeLa cells were from the American Type Culture Collection (ATCC) and were grown in DMEM/GlutaMAXI supplemented with 10% FBS. Media and serum were from PAA Laboratories and Invitrogen, respectively. Cells were transiently transfected using electroporation or lipofection with Lipofectamine 2000 and processed for epifluorescence 2 days after transfection. Cells were fixed at room temperature in 4% paraformaldehyde. Antibodies were used at the following dilutions: mouse monoclonal anti-human LAMP1, 1:2,000 (H4A3, Developmental Studies Hy-

[2] hEIDI: Hesse, A.-H., Adam, A., Dupierris, V., Court, M., Barthe, D., Emadali, A., Masselon, C., Ferro, M., and Bruley C., manuscript in preparation.

bridoma Bank); Cy3-conjugated donkey anti-mouse, 1:1,000 (The Jackson Laboratory). Fluorescence was examined using a Nikon TE2000 epifluorescence microscope. Images were deconvoluted after acquisition with the PSF-based Iterative 3D Deconvolution module of Metamorph software (Universal Imaging Corp.).

## RESULTS

To extend our knowledge of the lysosomal protein content, with a particular focus on membrane proteins and especially transporters, we performed a semi-quantitative and comparative proteomics analysis of membranes from rat liver fractions enriched and nonenriched in lysosomes. As novel proteins remaining to be discovered have a low abundance, we maximized protein resolution and coverage by analyzing several biological replicates and by reducing sample complexity using membrane protein subfractionation and SDS-PAGE. The label-free spectral counting method, based on the number of redundant peptides that identify a protein (44, 46), was used to evaluate the relative abundance of each protein. Further selection of lysosomal protein candidates resulted from a statistical comparison between lysosome-enriched and -nonenriched fractions (45). Finally, novel candidate lysosomal transporters were identified among the multipass transmembrane proteins.

*Preparation of Samples from Lysosome-enriched and -non-enriched Fractions*—We essentially followed the well established protocol of Wattiaux *et al.* (37) for preparation of lysosomal fractions (Fig. 1). Rat liver homogenates were first fractionated by differential centrifugation, and the primary lysosome-enriched fraction (fraction L) was further separated on a discontinuous Nycodenz density gradient (47), resulting in secondary L+ and L− fractions. Nycodenz is a gradient medium that displays very similar banding density for various organelles as the originally described metrizamide medium (48). Three independent preparations (biological replicates) were made. Our protocol aimed at improving the identification of IMPs, because their hydrophobicity and usually low abundance hinder their MS detection and identification in highly complex protein samples (34, 35). However, we also attempted to retain peripheral membrane associated proteins. These membrane associated proteins indeed include various trafficking machineries and cytoskeleton-associated proteins, which are crucial for the biogenesis and function of endo-lysosomes. We thus avoided harsh treatments that would have removed membrane-associated proteins, such as alkaline washes, to prepare L+ and L− membranes (respectively MbL+ and MbL− fractions). We then subfractionated both fractions according to protein hydrophobicity by two independent treatments, chloroform/methanol extraction (39) and Triton X-114 phase separation (Fig. 1) (40, 49). All resulting samples were finally separated by 1D electrophoresis, before being processed for MS analysis.

The $\beta$-galactosidase activity was measured to follow the recovery of lysosomes along the fractionation process, for the three replicates. These measurements indicated that the L, L+, and L− fractions were enriched 9–13-, 65–75-, and 7–9.5-fold, respectively, in lysosomes relative to the initial liver homogenate. These values, consistent with published data (37, 47), demonstrated the enrichment and nonenrichment of L+ and L−, respectively, as compared with L, and the much higher concentration (~7–9-fold) of lysosomes in L+ as compared with L−. These fractions are thus described as lysosome-enriched and lysosome-nonenriched, respectively, and evaluation of "enrichment" will hereafter always be based on the comparison between L+ and L− fractions.

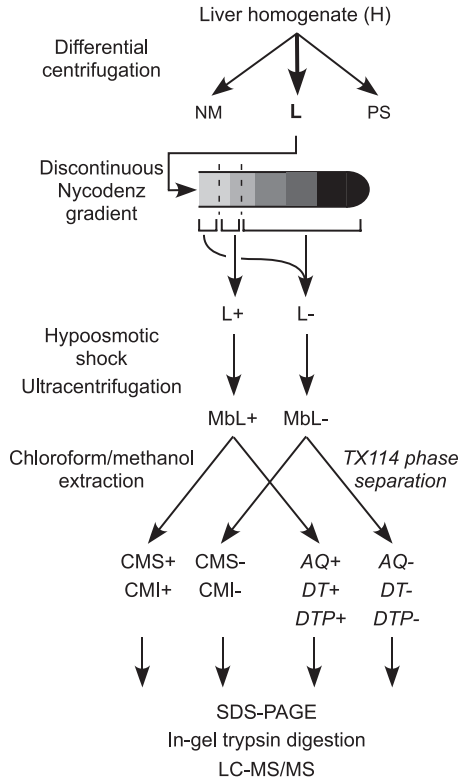


FIG. 1. **Workflow of the sample preparation for MS analysis.** Differential centrifugation of rat liver homogenates (*H*) produced a light mitochondrial fraction L, which was submitted to Nycodenz gradient centrifugation. This step allowed separation of a lysosome-enriched fraction (*L+*) from the rest of the gradient (*L−*). Organelles from L+ and L− were broken by hypoosmotic shock, and membranes were recovered by ultracentrifugation. Membrane pellets (*MbL+* and *MbL−*) were split in two equal parts that were separately fractionated by independent methods based on protein hydrophobicity (chloroform/methanol extraction or Triton X-114 (*TX114*) phase separation). All resulting samples were subsequently separated by SDS-PAGE prior to LC-MS/MS analysis of in-gel digested samples.

We then analyzed the enrichment of several subcellular compartments by immunodetection of organelle markers in the NM, L, and PS fractions resulting from differential centrifugation, as well as in the membranes of the L, L+, and L− fractions (Fig. 2). The lysosomal protein LAMP2 was the only protein that was strongly enriched in both L and MbL+ fractions. Rab5 (early endosome), the $\alpha$1 subunit of the sodium potassium ATPase (plasma membrane), TGN38 (TGN), and

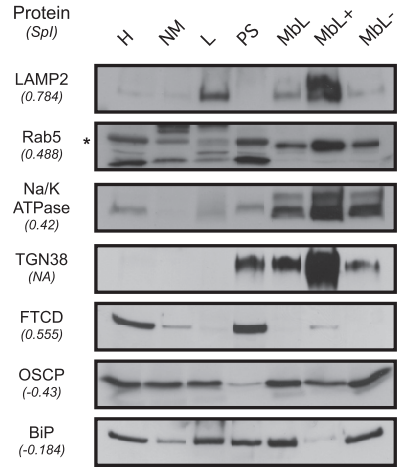


FIG. 2. **Western blot analysis of subcellular fractions.** The relative abundance of organelle protein markers was examined by Western blot analysis, in the differential centrifugation fractions (*H, NM, L,* and *PS*) and in the membranes recovered from the L fraction and the L+ and L− samples (respectively *MbL, MbL+,* and *MbL−*). Identical protein amounts (25 $\mu$g) have been loaded for each sample. Subcellular markers are as follows: *LAMP2*, lysosomes; *Rab5*, early endosomes; $\alpha$1 subunit of the sodium potassium ATPase (*Na/K-ATPase*), plasma membrane; *TGN38*, *trans*-Golgi network; *FTCD*, Golgi; *OSCP* subunit of the ATP synthase, mitochondria; and *BiP*, endoplasmic reticulum. For each protein tested, the SpI (see "Results" and Fig. 4) is indicated in *italics*.

FTCD (Golgi) were depleted from L and enriched in PS to different extents. Although they all display some enrichment in MbL+, TGN38 is the only one for which this enrichment is comparable with that of LAMP2. Both the mitochondrial ATP synthase OSCP subunit and the endoplasmic reticulum BiP were depleted in MbL+.

*Protein Identification*—From the three biological replicates, we generated 959 MS analyses. After first pass filters, this resulted in 1,398,920 spectra, 368,147 of which could be assigned to 4,097 nonredundant rat gene products from the IPI-Rat database. According to the principle of parsimony, protein isoforms that could not be segregated by the identified peptides were counted as one unique gene product. The 4,097 gene products corresponded to 24,316 nonredundant peptide sequences. All corresponding protein and peptide information is available in the Pride database under project number 22847 and in supplemental Table S2. Further filtering excluding trypsin and keratins as contaminants and retaining proteins identified by at least two unique peptides led to a list of 2,385 nonredundant gene products, hereafter named the MbL2385 list. In this list, 528 proteins were present in the MbL+ fraction only, 157 in the MbL− fraction only, and 1,700 in both fractions (supplemental Tables S4a and S5a). Thus, most of the proteins were common to both MbL+ and MbL− fractions, in agreement with the limited resolution power of subcellular fractionation and the high sensitivity of mass spectrometers.

To evaluate the content in IMPs identified in our samples, transmembrane domains were predicted by use of the TMHMM 2.0 server (supplemental Table S4a). This led to the identification of 762 IMPs (32%), including 361 polytopic proteins (proteins with at least two transmembrane domains, 15.1%).

*Extraction of Semi-quantitative Data*—Despite the high enrichment factor obtained by the well established Nycodenz gradient method used in this study, cofractionation of other organelles, such as mitochondria, challenges the identification of true lysosomal residents, including proteins with dual or multiple localization. We thus compared the protein sets from lysosome-enriched and -nonenriched fractions to identify the subset associated with lysosomes. Because most proteins were common to both MbL+ and MbL− fractions, comparing their number was less informative and relevant than comparing their abundance (compare Fig. 3 with supplemental Fig. S1). Abundance information was extracted from spectral count data (supplemental Table S1), according to the spectral counting semi-quantitative approach (44, 46). The relative abundance for any given protein was derived from the Nsc calculated as indicated under "Experimental Procedures" using merged data issued from all MbL+ or MbL− samples (supplemental Table S1).

*MbL+ and MbL− Fractions Display Different Organellar Profiles*—We then analyzed the known or predicted subcellular localization of proteins from the MbL2385 list by manually collecting this information in protein databases (IPI, UniprotKB, and QuickGO) and bibliography. For IPI entries without any attributed gene name, homologs were previously searched in a mammalian subset of the Uniprot database using Blastp. This allowed comparison of protein abundances in MbL+ and MbL− according to the following subcellular categories: endo-lysosomes (EL), plasma membrane (PM), mitochondria, peroxisomes and nucleus (MPN), endoplasmic reticulum (ER), Golgi (G), cytoplasm (C), cytoskeleton (CS), secreted (S), miscellaneous (Misc; vesicles, granules, and multiple localizations) and Unknown. The comparison of protein abundances in MbL+ and MbL− according to the subcellular distribution showed striking differences (Fig. 3, *left panel*; supplemental Table S5c); EL and PM proteins were clearly enriched in MbL+, as they altogether accounted for 34% of the material, as compared with 4.7% only in MbL−. By contrast, proteins from the ER and MPN compartments were depleted from MbL+ relative to MbL− (34.2 and 73.7%, respectively). The similar behaviors of EL and PM proteins on the one hand and ER and MPN proteins on the other hand were systematically observed in subsequent analyses. Despite its slight enrichment in the MbL+ fraction (0.59% of the abundance in MbL+ *versus* 0.15% in MbL−), the small set of Golgi proteins ($n$ = 30) has been ranked as "Contaminants," along with ER and MPN proteins, in subsequent quantitative analyses (see below). Except for the cytoplasm, the other subcellular constituents (CS, S, and Misc) were slightly enriched in the MbL+ fraction (15.6 *versus* 7.1%). Proteins of unknown localization represented 11.2 and 10.0% in number but 6.3 and 3.8% in abundance in MbL+ and MbL−, respec-

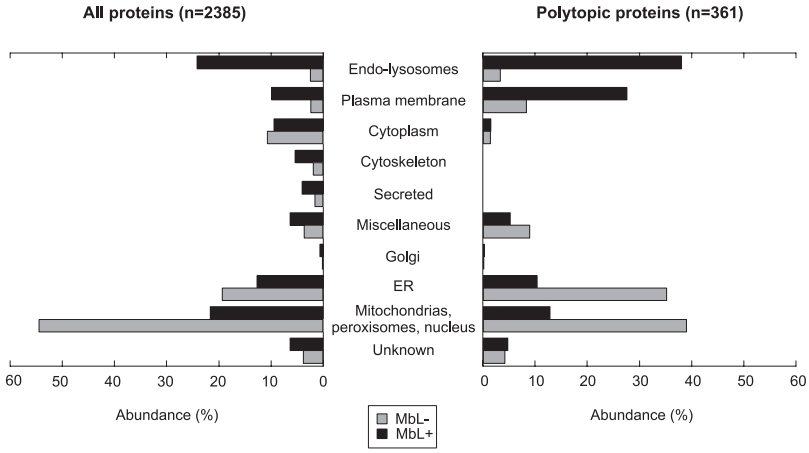


FIG. 3. **Subcellular distribution of the proteins in MbL+ and MbL−.** The relative abundance of proteins in MbL+ (*black bars*) or in MbL− (*gray bars*) is compared, according to their subcellular classification. *Left panel,* distribution of all proteins. *Right panel,* distribution of polytopic proteins (*i.e.* harboring at least two transmembrane domains).

tively, indicating that the average relative abundance of such proteins is low (Fig. 3, *left panel*, and supplemental Fig. S1 and supplemental Table S5c).

Thus, the subcellular distribution features that stem from our spectral count data were consistent with qualitative expectations based on a restricted set of organelle markers (Fig. 2) (37). The substantial presence of contaminant organelles in MbL+ was expected as a known characteristic of subcellular fractions. Our results therefore validate the use of the spectral count-based semi-quantitative method to describe and analyze these fractions.

*Assignment of Proteins to Lysosomes*—The next step in our study was to identify which proteins identified in MbL+ were indeed novel potential lysosomal proteins. We thus aimed at identifying those significantly enriched in MbL+ relative to MbL−, similarly, for instance, to the observed enrichment of the typical lysosomal marker $\beta$-galactosidase in L+ relative to L−. Proteins from the MbL+ fraction were either exclusively detected in MbL+ or common to both fractions (supplemental Table S4a). Among the 528 proteins exclusively present in MbL+, we chose to consider as potentially lysosomal only those present in at least two out of the three biological replicates and identified by at least five spectra (356 proteins; supplemental Table S4b). Among the proteins common to MbL+ and MbL−, lysosomal candidates were selected according to their SpI (45), a parameter that takes into account both the relative protein abundance (estimated by normalized spectral counts) and the number of replicates in which the protein has been found (supplemental Tables S1 and S4a). SpI values range from −1 to +1, the lower and upper extreme values corresponding to proteins almost exclusively detected in MbL− and MbL+ fractions, respectively. These values displayed a roughly bimodal distribution in the MbL2385 list, with a massive peak covering negative values and a second subset rising toward an SpI of +1 (Fig. 4*A*). The SpI analysis highlighted the different distributions between MbL+ and MbL− of proteins from various annotated subcellular categories (Fig. 4*B*). Indeed, proteins from contaminants (essentially mitochondrial, ER and peroxisomal proteins) were the main contributors to the massive peak of negative SpI, whereas EL and PM proteins demonstrated a strong tendency to score high SpI values, with respective median values of 0.77 and 0.60. Confidence intervals were established through permutation analysis (45). Proteins were considered as significantly enriched in MbL+ when their SpI was higher than the 95th percentile cutoff value (SpI ≥0.594), a level reached by 378 proteins out of 1,700 (supplemental Table S4b).

Altogether, our selection criteria for significant enrichment in MbL+ led us to sort 734 proteins (Lys-734 list; supplemental Table S4b) out of 2,385. This selection included 79.3% ($n$ = 207) of the EL-annotated proteins, 56% ($n$ = 132) of the PM-annotated ones, and only 3.2% ($n$ = 28) of the contaminant proteins (Fig. 5*A* and supplemental Table S5d). The C, CS, S, and Misc categories were represented by a total of 273 proteins. To our knowledge, 38 of the 94 proteins without any subcellular localization annotation (Table I) were completely novel lysosomal candidates, because they have not been identified in previous proteomic studies of lysosomes (14–17, 19).

As it was recently shown that most known lysosomal genes exhibit a coordinated transcriptional behavior regulated by the transcription factor TFEB, we compared our Lys-734 list to the list of 291 genes up-regulated following TFEB overexpression in HeLa cells (50). This comparison pointed to 38 common proteins, among which 30 were EL-annotated proteins and one, the product of the *Wdr81* gene, was of unknown annotated localization.

Extraction from protein databases or bibliography, and analysis of known or predicted functional annotation showed that all defined functional processes were represented in the Lys-734 list, although only two "polypeptide transport" annotated proteins remained (Fig. 5*B*). Transporters, channels, and pumps of ions and small molecules represented the most

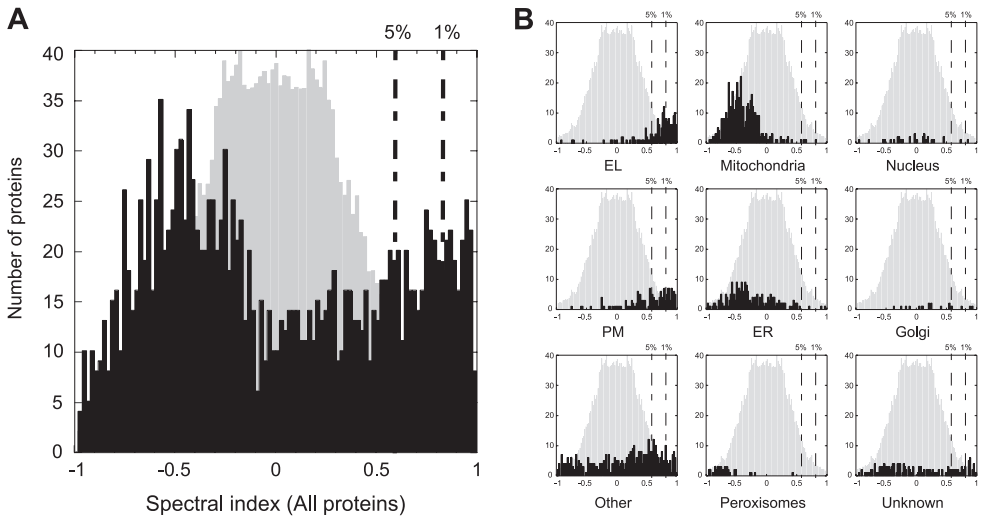


FIG. 4. **Distribution of spectral indexes.** The distribution of spectral indexes of identified proteins is represented (in *black*) along with a random distribution (*light gray*) derived from permutation analysis. Thresholds for the 5 and 1% confidence intervals (indicative of a significant enrichment in MbL+) are indicated by *dotted lines*. *A,* SpI distribution for all identified proteins common to MbL+ and MbL−. *B,* SpI distribution for proteins of specific subcellular annotation. The number of proteins in each category is as follows: *All*, $n = 1{,}700$; *EL*, $n = 160$; *ER*, $n = 228$; *Golgi*, $n = 19$; *Mitochondria*, $n = 419$; *Nucleus*, $n = 32$; other, $n = 467$; *Peroxisomes,* $n = 44$; PM, $n = 163$; *Unknown*, $n = 168$; *x* axis, SpI (range from −1 to +1); *y* axis, number of proteins.

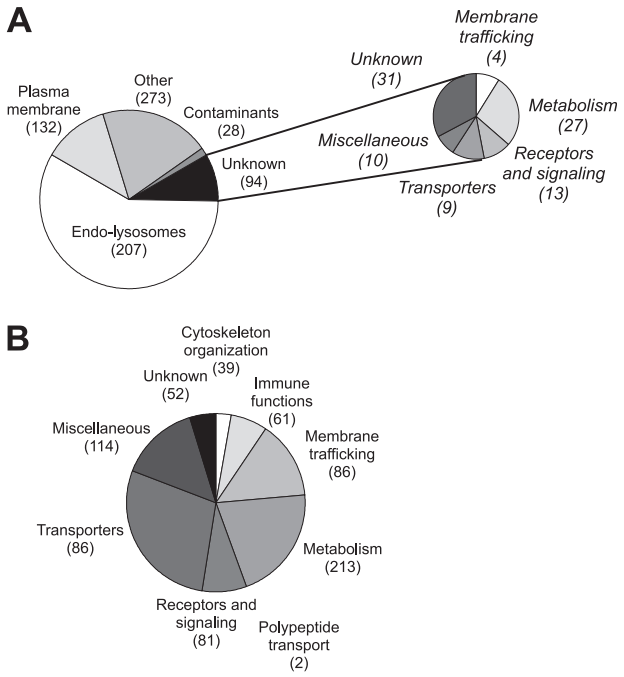


FIG. 5. **Subcellular and functional distribution of the Lys-734 proteins.** *A,* distribution of the proteins according to their subcellular annotation. The functional annotation of proteins of unknown localization is shown. "Cytoplasm," "cytoskeleton," "secreted," and "miscellaneous" are merged in the *Other* category. *B,* distribution of the proteins according to their functional annotation. Each *pie section* represents the relative abundance of proteins from the corresponding category. Numbers of proteins are indicated.

abundant functional class, despite its third position by protein number. Metabolism-associated proteins were the most numerous but ranked second in abundance (supplemental Table S5e). As for the 94 proteins without subcellular annotation, one-third had no functional annotation either; more than one-quarter were various metabolic enzymes; and the remaining were distributed between the "miscellaneous," "transporters, channels, and ion pumps," and "receptors and signaling" classes with rather similar abundances (Fig.5*A*).

*Identification of Novel Putative Lysosomal Transporters*—In addition to extending the current list of known lysosomal proteins, our interest was focused on the discovery of potential novel lysosomal transporters. As transporters display multiple membrane spanning domains (35), we filtered the Lys-734 list for polytopic proteins. Among the 136 MbL+-enriched polytopic proteins, 10% ($n = 11$) had no attributed function and more than half ($n = 72$; 67.5% of the Lys-734 IMPs abundance) belonged to the transporters, channels, and ion pumps class. This protein set contains numerous subunits of ATPases (v-ATPase ($n = 6$); P-ATPases ($n = 5$)), ATP-binding cassette (ABC) transporters ($n = 9$), channels ($n = 10$), and secondary active transporters ($n = 42$). The latter include the recently discovered potential or effective lysosomal transporters C2ORF18 (21, 51), DIRC2 (27), LMBD1 (52), and MFSD8 (53) (Fig. 6). During the revision of our manuscript, the lysosomal localization of the ABC transporter ABCD4 was established (29), and we showed in a separate study the lysosomal localization and transport function of the PQLC2 protein (36).

Removal of the transporters already annotated as endo-lysosomal led to a set of 46 novel potentially lysosomal transporters that notably included 27 plasma membrane proteins

TABLE I
*List of potential novel lysosomal proteins*

Proteins from the Lys-734 list without any subcellular localization annotation are given. The functions are as follows: I, immune function; M, metabolism; Misc, miscellaneous; MT, membrane trafficking; R/S, receptors and signaling; T, transporters, channels, and ion pumps; U, unknown; TM, number of transmembrane domains; SpI, spectral index; NA, not applicable (protein identified in MbL+ exclusively). The lysosomal localization of MFSD1 has been shown during the course of this work (66).

| Gene name | Description | Accession no. | TM | Function | SpI | No. of peptides | Coverage |
|---|---|---|---|---|---|---|---|
| *Abca6* | Abca6, similar to ATP-binding cassette, subfamily A (ABC1), member 6 | IPI00762951 | 13 | T | 0.76 | 6 | 4.61 |
| *Acp1* | Acp1, isoform 1 of low molecular weight phosphotyrosine protein phosphatase | IPI00206664 | 0 | R/S | NA | 2 | 12.66 |
| *Afmid* | LOC688283, similar to kynurenine formamidase | IPI00882532 | 0 | M | NA | 7 | 27.45 |
| *Ahcy* | Ahcy, adenosylhomocysteinase | IPI00476295 | 0 | M | 0.60 | 16 | 38.66 |
| *Akr1c12l1* | RGD1559604, similar to protein RAKd | IPI00557070 | 0 | M | NA | 4 | 13.31 |
| *Akr1c13* | LOC364773, 17$\beta$-hydroxysteroid dehydrogenase | IPI00387641 | 0 | M | NA | 5 | 19,38 |
| *Aox3* | Aox3, aldehyde oxidase 1 | IPI00205560 | 0 | M | 0.88 | 26 | 25.11 |
| *Ap5z1* | Kiaa0415, hypothetical protein LOC641386 | IPI00363750 | 0 | MT | NA | 8 | 11.03 |
| *Atp11c-ps1* | 128-kDa protein | IPI00370178 | 8 | T | 0.63 | 25 | 27.32 |
| *C10orf32* | RGD1311783, hypothetical protein | IPI00371685 | 0 | U | NA | 10 | 59.06 |
| *C16orf62* | LOC361635, UPF0505 protein C16orf62 homolog | IPI00569226 | 0 | U | 0.80 | 5 | 5.61 |
| *C17orf59* | LOC497934, uncharacterized protein C17orf59 homolog | IPI00188598 | 0 | U | 0.87 | 11 | 22.16 |
| *C1galt1* | C1galt1, glycoprotein-*N*-acetylgalactosamine 3$\beta$-galactosyltransferase 1 | IPI00200858 | 1 | M | NA | 7 | 19.83 |
| *C1galt1c1* | C1galt1c1, C1GALT1-specific chaperone 1 | IPI00197034 | 1 | Misc | NA | 3 | 11.08 |
| *C2orf72* | 18-kDa protein | IPI00188569 | 0 | U | NA | 4 | 36.9 |
| *C6orf58* | RGD1311933, hypothetical protein | IPI00358842 | 0 | U | 0.95 | 9 | 29.97 |
| *C9orf91* | RGD1304595, hypothetical protein LOC298104 | IPI00357901 | 2 | U | NA | 7 | 22.51 |
| *Ca2* | Car2, carbonic anhydrase 2 | IPI00230787 | 0 | M | NA | 2 | 8.85 |
| *Cacfd1* | RGD1311501, hypothetical protein LOC296599 | IPI00363948 | 3 | T | NA | 2 | 14.62 |
| *Ccdc22* | Ccdc22, similar to coiled-coil domain containing 22 | IPI00362580 | 0 | U | NA | 11 | 23,6 |
| *Ccdc93* | Ccdc93, coiled-coil domain-containing protein 93 | IPI00371846 | 0 | U | NA | 10 | 18.28 |
| *Cd302* | Cd302, CD302 antigen | IPI00372762 | 1 | R/S | NA | 4 | 15.79 |
| *Clec4f* | Clec4f, C-type lectin domain family 4 member F | IPI00193212 | 1 | R/S | 0.98 | 20 | 33.64 |
| *Clec4g* | Clec4g, similar to C-type lectin domain family 4, member g | IPI00764324 | 1 | U | NA | 4 | 14.65 |
| *Cnp* | Cnp, 2′,3′-cyclic-nucleotide 3′-phosphodiesterase | IPI00199394 | 0 | M | 0.78 | 5 | 13.81 |
| *Commd10* | Commd10, COMM domain containing 10 | IPI00365123 | 0 | U | NA | 8 | 36.63 |
| *Commd2* | Commd2, COMM domain containing 2 | IPI00372217 | 0 | U | NA | 2 | 8.54 |
| *Commd3* | Commd3, COMM domain-containing protein 3 | IPI00400613 | 0 | U | NA | 6 | 36.92 |
| *Commd7* | Commd7, COMM domain containing 7 | IPI00373166 | 0 | Misc | NA | 5 | 31 |
| *Commd9* | Commd9, COMM domain containing 9 | IPI00210812 | 0 | U | NA | 8 | 58.59 |
| *Cpne5* | Cpne5, copine V | IPI00360489 | 0 | MT | 0.60 | 2 | 3.04 |
| *Crip2* | Crip2, cysteine-rich protein 2 | IPI00200352 | 0 | Misc | NA | 3 | 33.65 |
| *Csad* | Csad, cysteine sulfinic acid decarboxylase | IPI00214394 | 0 | M | 0.74 | 12 | 32.86 |
| *Csnk2a1* | Csnk2a1, casein kinase II subunit $\alpha$ | IPI00192586 | 0 | R/S | NA | 4 | 10.23 |
| *Dak* | Dak, dihydroxyacetone kinase | IPI00372498 | 0 | M | 0.81 | 22 | 46.71 |
| *Dnajc13* | Dnajc13, 108-kDa protein | IPI00366703 | 0 | U | NA | 3 | 5.59 |
| *Dnajc13* | Dnajc13, DnaJ (Hsp40) homolog, subfamily C, member 13 | IPI00870706 | 0 | U | NA | 6 | 4.14 |
| *Dnajc5* | Dnajc5, DnaJ homolog subfamily C member 5 | IPI00210881 | 0 | Misc | 0.63 | 6 | 31.31 |
| *Enpp4* | Enpp4, ectonucleotide pyrophosphatase/phosphodiesterase 4 | IPI00371761 | 1 | M | 0.89 | 5 | 11.07 |
| *Eprs* | Eprs, LRRGT00050 | IPI00476855 | 0 | M | NA | 4 | 2.76 |
| *Fth1* | Fth1, ferritin heavy chain | IPI00777061 | 0 | Misc | 0.70 | 14 | 59.34 |
| *Gna11* | Gna11, guanine nucleotide-binding protein $\alpha$-11 subunit | IPI00200437 | 0 | R/S | 0.86 | 13 | 36.21 |
| *Gna13* | Gna13, G$\alpha$13 | IPI00422053 | 0 | R/S | 0.79 | 13 | 34.22 |
| *Gna14* | Gna14, guanine nucleotide-binding protein, $\alpha$14 | IPI00360645 | 0 | R/S | 0.86 | 3 | 12.68 |

Table I—*continued*

| Gene name | Description | Accession no. | TM | Function | SpI | No. of peptides | Coverage |
|---|---|---|---|---|---|---|---|
| *Gnai3* | Gnai3, guanine nucleotide-binding protein G(k) subunit $\alpha$ | IPI00231726 | 0 | R/S | 0.88 | 12 | 43.22 |
| *Gpr155* | Gpr155 G protein-coupled receptor 155 | IPI00365274 | 17 | R/S | 0.89 | 9 | 14.29 |
| *Grhpr* | Grhpr, Grhpr protein | IPI00767591 | 0 | M | 0.87 | 7 | 24.18 |
| *Haao* | Haao 3-hydroxyanthranilate 3,4-dioxygenase | IPI00339188 | 0 | M | NA | 9 | 22.03 |
| *Iah1* | Iah1, isoamyl acetate-hydrolyzing esterase 1 homolog | IPI00421610 | 0 | M | NA | 5 | 16.47 |
| *Igtp* | Igtp Ac2-233 | IPI00369234 | 0 | M | NA | 5 | 5.73 |
| *Itfg3* | Itfg3 protein ITFG3 | IPI00372350 | 1 | U | NA | 5 | 11.78 |
| *Jak1* | Jak1, similar to tyrosine-protein kinase JAK1 | IPI00212981 | 0 | R/S | 0.75 | 3 | 2.84 |
| *Kctd12* | Kctd12, similar to potassium channel tetramerization domain-containing protein 12 | IPI00767085 | 0 | Misc | NA | 3 | 9.77 |
| *Kxd1* | LOC498606, UPF0459 protein C19orf50 homolog | IPI00197953 | 0 | U | NA | 7 | 28.25 |
| *Lgals5* | Lgals5, galectin-5 | IPI00231663 | 0 | Misc | 0.94 | 8 | 55.17 |
| *Loh12cr1* | Loh12cr1, loss of heterozygosity, 12, chromosomal region 1 homolog | IPI00189639 | 0 | U | 0.94 | 22 | 80.51 |
| *Mef2bnb* | LOC684626, similar to K11B4.2 | IPI00364627 | 0 | U | 0.97 | 6 | 38.66 |
| *Mfsd1* | Mfsd1, similar to major facilitator superfamily domain-containing 1 | IPI00373212 | 11 | T | NA | 6 | 18.75 |
| *Mic1* | RGD1311805, similar to RIKEN cDNA 2400010D15 | IPI00363472 | 0 | U | NA | 25 | 41.55 |
| *Mios* | RGD1308432, similar to missing oocyte CG7074-PA | IPI00199953 | 0 | U | NA | 7 | 8.11 |
| *Mon1b* | Mon1b, MON1 homolog b | IPI00359673 | 0 | MT | 0.93 | 23 | 53.86 |
| *Mpeg1* | Mpeg1, macrophage-expressed gene 1 protein | IPI00204417 | 1 | U | 0.80 | 18 | 39.92 |
| *Myadm* | Myadm, myeloid-associated differentiation marker | IPI00339007 | 8 | U | 0.76 | 2 | 8.49 |
| *Napg* | LOC682827, 35-kDa protein | IPI00367524 | 0 | MT | 0.92 | 28 | 52.24 |
| *Oplah* | Oplah, 5-oxoprolinase | IPI00326436 | 0 | M | NA | 5 | 5.05 |
| *Pah* | Pah, phenylalanine-4-hydroxylase | IPI00193258 | 0 | M | 0.66 | 11 | 36.64 |
| *Pbld* | Pbld, phenazine biosynthesis-like domain-containing protein | IPI00200041 | 0 | M | NA | 6 | 32.64 |
| *Pebp1* | Pebp1, phosphatidylethanolamine-binding protein 1 | IPI00230937 | 0 | Misc | 0.68 | 5 | 43.85 |
| *Pik3ca* | Pik3ca, similar to phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit $\alpha$ isoform | IPI00955176 | 0 | R/S | 0.64 | 2 | 2.06 |
| *Pk* | Similar to pyruvate kinase 3 | IPI00561880 | 0 | M | NA | 18 | 35.78 |
| *Pklr* | Pklr, isoform R-type of pyruvate kinase isozymes R/L | IPI00202549 | 0 | M | 0.86 | 18 | 43.55 |
| *Pla2g2a* | Pla2g2a, phospholipase A2, membrane-associated | IPI00205248 | 1 | M | NA | 3 | 28.08 |
| *Ppa1* | Ppa1, similar to pyrophosphatase | IPI00371957 | 0 | M | NA | 3 | 11.63 |
| *Prkag1* | Prkag1, 5′-AMP-activated protein kinase subunit $\gamma$-1 | IPI00196645 | 0 | R/S | 0.72 | 5 | 20 |
| *Rbp1* | Rbp1, retinol-binding protein 1 | IPI00231825 | 0 | T | NA | 5 | 29.63 |
| *RGD1308461* | RGD1308461, similar to CG5149-PA | IPI00359821 | 0 | U | NA | 9 | 24.35 |
| *Sec14l4* | Sec14l4, SEC14-like 4 | IPI00204634 | 0 | Misc | 0.66 | 11 | 31.07 |
| *Slc38a7* | Slc38a7, putative sodium-coupled neutral amino acid transporter 7 | IPI00421684 | 11 | T | NA | 7 | 10.8 |
| *Slc46a3* | Slc46a3, solute carrier family 46 member 3 | IPI00364398 | 11 | T | NA | 2 | 5.86 |
| *Slco2b1* | Slco2b1, solute carrier organic anion transporter family member 2B1 | IPI00230858 | 12 | T | 0.75 | 3 | 5.56 |
| *Sord* | Sord, sorbitol dehydrogenase | IPI00760137 | 0 | M | 0.68 | 16 | 45.1 |
| *Stard10* | Stard10, START domain containing 10 | IPI00555188 | 0 | U | NA | 5 | 16.21 |
| *Tagln3* | Tagln3, transgelin-3 | IPI00210532 | 0 | U | NA | 5 | 11.47 |
| *Tgm2* | Tgm2, transglutaminase 2, C polypeptide | IPI00205135 | 0 | M | NA | 6 | 11.81 |
| *Tm9sf4* | Tm9sf4, transmembrane 9 superfamily member 4 | IPI00373155 | 9 | Misc | NA | 9 | 13.84 |
| *Tmem104* | Tmem104, similar to CG5262-PA | IPI00778760 | 11 | T | NA | 6 | 11.29 |
| *Tmem144* | Tmem144, transmembrane protein 144 | IPI00373219 | 10 | U | 0.95 | 5 | 23.85 |
| *Tmem175* | Tmem175, transmembrane protein 175 | IPI00211068 | 9 | U | NA | 5 | 18.44 |

Table I—*continued*

| Gene name | Description | Accession no. | TM | Function | SpI | No. of peptides | Coverage |
|---|---|---|---|---|---|---|---|
| *UBB* | LOC679594;LOC682397 similar to polyubiquitin | IPI00763565 | 0 | M | 0.88 | 3 | 38.96 |
| *Ubl3* | Ubl3, ubiquitin-like protein 3 | IPI00358637 | 0 | U | NA | 6 | 47.01 |
| *Uroc1* | Uroc1, similar to urocanase domain containing 1 | IPI00388707 | 0 | M | NA | 8 | 14.35 |
| *UST4r* | UST4r, putative integral membrane transport protein | IPI00202688 | 7 | T | 0.89 | 6 | 10.33 |
| *Vsig4* | Vsig4, V-set and immunoglobulin domain-containing 4 | IPI00372986 | 1 | R/S | NA | 5 | 17.19 |
| *Wdr81* | Wdr81, similar to $\alpha$2-plasmin inhibitor | IPI00370309 | 0 | U | 0.65 | 6 | 3.58 |
| *Wdr91* | Wdr91, Wdr91 protein | IPI00373314 | 0 | U | NA | 4 | 7.62 |

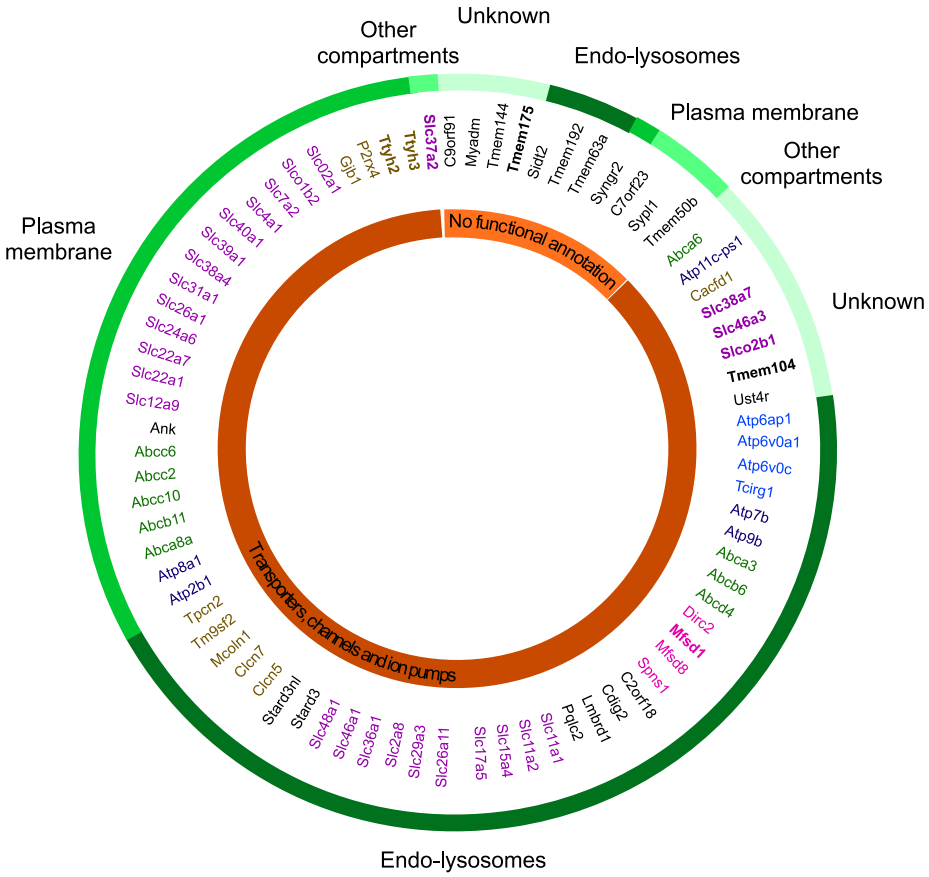


FIG. 6. **Lysosomal transportome.** All known and potential transporters or channels retained as selectively enriched in MbL+ are represented. These proteins display at least two transmembrane domains, they either belong to the functional class "transporters, channels, and ion pumps" or have no functional annotation. They are classified according to their functional and subcellular annotations. Different categories of transporters are depicted by different colors (ABC transporters, *green*; MFS transporters, *pink*; SLC family members, *purple*; ATPases, *deep blue*; V-ATPase subunits, *light blue*; channels, *brown*; miscellaneous, *black*). Validated candidates are in *boldface*.

and 12 proteins of unknown localization (Table II). To our knowledge, 9 out of these 46 proteins (ABCA6, C7ORF23, C9ORF91, CACFD1, SLC26A1, SLC38A7, SLC40A1, SLC46A3, and TMEM50b) have not been identified in previous proteomic analyses of mammalian lysosomes, phagosomes, or lysosome-related organelles (14–17, 19, 54–61).

*Validation of Selected Candidates*—Twelve candidates, LOH12CR1, STARD10, PTTG1IP, MFSD1, SLC37A2, SLC38A7, SLC46A3, SLCO2B1, TMEM104, TMEM175, TTYH2 and TTYH3, were chosen to validate independently the proteomic data. Peptides allowing their identification are given in supplemental Table S6. LOH12CR1 and STARD10 are putative cytosolic proteins. PTTG1IP is predicted to be an integral membrane protein with a role in cellular trafficking. Its subcellular localization is unclear as it has been observed in cytosol and nucleus by some authors (62) and in late endosomes by others (63). All other candidates are multispanning transmembrane proteins. TMEM175 has no homology with

TABLE II

*List of potential novel lysosomal transporters*

Candidates with two TM or more, of unknown function or with an attributed transport function, are shown. The localizations are as follows: ER, endoplasmic reticulum, Golgi; EL, endo-lysosomes; Misc, miscellaneous; PM, plasma membrane; U, unknown. The functions are as follows: T, transporters, channels, and ion pumps; U, unknown; TM, number of transmembrane domains; SpI, spectral index; NA, not applicable (proteins identified in MbL+ exclusively). The lysosomal localization of MFSD1 has been shown during the course of this (66).

| Gene name | Description | Accession no. | TM | Localization | Function | SpI | No. of peptides | Coverage |
|---|---|---|---|---|---|---|---|---|
| *Abca6* | Abca6, similar to ATP-binding cassette, subfamily A (ABC1), member 6 | IPI00762951 | 13 | U | T | 0.76 | 6 | 4.61 |
| *Abca8a* | Abca8a, similar to ATP-binding cassette, subfamily A (ABC1), member 8a | IPI00763783 | 11 | PM | T | NA | 4 | 1.18 |
| *Abcb11* | Abcb11, bile salt export pump | IPI00195615 | 9 | PM | T | 0.90 | 15 | 14.76 |
| *Abcc10* | Abcc10, ATP-binding cassette, subfamily C (CFTR/MRP), member 10 | IPI00371742 | 14 | PM | T | 0.84 | 15 | 10.85 |
| *Abcc2* | Abcc2, canalicular multispecific organic anion transporter 1 | IPI00205806 | 14 | PM | T | 0.94 | 28 | 24.08 |
| *Abcc6* | Abcc6, multidrug resistance-associated protein 6 | IPI00207513 | 12 | PM | T | 0.97 | 17 | 17.58 |
| *Ank* | Ank, progressive ankylosis protein homolog | IPI00765376 | 8 | PM | T | NA | 2 | 6.63 |
| *Atp11c-ps1* | 128-kDa protein | IPI00370178 | 8 | U | T | 0.63 | 25 | 27.32 |
| *Atp2b1* | Atp2b1, isoform D of plasma membrane calcium-transporting ATPase 1 | IPI00231268 | 7 | PM | T | NA | 5 | 4.26 |
| *Atp8a1* | Atp8a1, similar to ATPase, aminophospholipid transporter (APLT), class I, type 8A, member 1 | IPI00952342 | 8 | PM | T | NA | 10 | 9.08 |
| *C7orf23* | RGD1562351, hypothetical protein LOC499990 | IPI00565669 | 3 | MPN | U | 0.60 | 4 | 23.73 |
| *C9orf91* | RGD1304595, hypothetical protein LOC298104 | IPI00357901 | 2 | U | U | NA | 7 | 22.51 |
| *Cacfd1* | RGD1311501, hypothetical protein LOC296599 | IPI00363948 | 3 | U | T | NA | 2 | 14.62 |
| *Gjb1* | Gjb1, Gap junction $\beta$-1 protein | IPI00207191 | 4 | PM | T | NA | 4 | 15.19 |
| *Mfsd1* | Mfsd1, similar to major facilitator superfamily domain containing 1 | IPI00373212 | 11 | U | T | NA | 6 | 18.75 |
| *Myadm* | Myadm, myeloid-associated differentiation marker | IPI00339007 | 8 | U | U | 0.76 | 2 | 8.49 |
| *P2rx4* | P2rx4, P2X purinoceptor 4 | IPI00324987 | 2 | PM | T | 0.87 | 13 | 38.4 |
| *Sidt2* | Sidt2, SID1 transmembrane family, member 2 | IPI00369576 | 9 | EL | U | 0.82 | 4 | 6.23 |
| *Slc12a9* | Slc12a9, solute carrier family 12 member 9 | IPI00198772 | 11 | PM | T | NA | 6 | 9.85 |
| *Slc22a1* | Slc22a1, isoform 1 of solute carrier family 22 member 1 | IPI00213324 | 10 | PM | T | 0.77 | 6 | 15.81 |
| *Slc22a7* | Slc22a7, solute carrier family 22 member 7 | IPI00203971 | 11 | PM | T | 0.83 | 12 | 25.23 |
| *Slc24a6* | Slc24a6, sodium/potassium/calcium exchanger 6 | IPI00464527 | 12 | PM | T | 0.92 | 5 | 12.27 |
| *Slc26a1* | Slc26a1, sulfate anion transporter 1 | IPI00207298 | 9 | PM | T | 0.74 | 10 | 25.6 |
| *Slc31a1* | Slc31a1, high affinity copper uptake protein 1 | IPI00231403 | 3 | PM | T | 0.79 | 5 | 16.58 |
| *Slc37a2* | Slc37a2, similar to solute carrier family 37 (glycerol 3-phosphate transporter), member 2 | IPI00569704 | 12 | ER | T | 0.98 | 5 | 16.4 |
| *Slc38a4* | Slc38a4, sodium-coupled neutral amino acid transporter 4 | IPI00189469 | 11 | PM | T | 0.95 | 7 | 14.63 |
| *Slc38a7* | Slc38a7, putative sodium-coupled neutral amino acid transporter 7 | IPI00421684 | 11 | U | T | NA | 7 | 10.8 |
| *Slc39a1* | Slc39a1, similar to zinc transporter ZIP1 | IPI00373235 | 6 | PM | T | 0.96 | 6 | 19.75 |
| *Slc40a1* | Slc40a1, solute carrier family 40 member 1 | IPI00326002 | 10 | PM | T | NA | 2 | 5.61 |
| *Slc46a3* | Slc46a3, solute carrier family 46 member 3 | IPI00364398 | 11 | U | T | NA | 2 | 5.86 |
| *Slc4a1* | Slc4a1, solute carrier family 4, member 1 | IPI00231379 | 10 | PM | T | NA | 6 | 8.41 |
| *Slc7a2* | Slc7a2, cationic amino acid transporter-2 | IPI00608159 | 16 | PM | T | 0.92 | 9 | 14.63 |
| *Slc7a2* | Slc7a2, cationic amino acid transporter-2A | IPI00206144 | 14 | PM | T | 0.95 | 6 | 8.98 |
| *Slco1b2* | Slco1b2, isoform 1 of solute carrier organic anion transporter family member 1B2 | IPI00215390 | 11 | PM | T | 0.73 | 14 | 27.91 |
| *Slco2a1* | Slco2a1, solute carrier organic anion transporter family member 2A1 | IPI00231272 | 11 | PM | T | 0.76 | 3 | 5.13 |
| *Slco2b1* | Slco2b1, solute carrier organic anion transporter family member 2B1 | IPI00230858 | 12 | U | T | 0.75 | 3 | 5.56 |
| *Syngr2* | Syngr2, synaptogyrin 2 | IPI00200093 | 4 | PM | U | 0.72 | 2 | 4.7 |
| *Sypl1* | Sypl, synaptophysin-like protein | IPI00471762 | 3 | Misc | U | 0.95 | 4 | 20.16 |
| *Tmem104* | Tmem104 similar to CG5262-PA | IPI00778760 | 11 | U | T | NA | 6 | 11.29 |
| *Tmem144* | Tmem144, transmembrane protein 144 | IPI00373219 | 10 | U | U | 0.95 | 5 | 23.85 |

Table II—*continued*

| Gene name | Description | Accession no. | TM | Localization | Function | SpI | No. of peptides | Coverage |
|---|---|---|---|---|---|---|---|---|
| *Tmem175* | Tmem175, transmembrane protein 175 | IPI00211068 | 9 | U | U | NA | 5 | 18.44 |
| *Tmem192* | Tmem192, transmembrane protein 192 | IPI00364640 | 4 | EL | U | NA | 4 | 12.78 |
| *Tmem50b* | Tmem50b, transmembrane protein 50B | IPI00373040 | 4 | ER | U | NA | 2 | 12.03 |
| *Tmem63a* | Tmem63a, similar to transmembrane protein 63a | IPI00363369 | 11 | EL | U | NA | 4 | 5.1 |
| *Ttyh2* | Ttyh2, similar to tweety 2 | IPI00763162 | 6 | PM | T | NA | 5 | 4.81 |
| *Ttyh3* | Ttyh3, tweety homolog 3 | IPI00363776 | 5 | PM | T | NA | 4 | 7.44 |
| *UST4r* | UST4r, putative integral membrane transport protein | IPI00202688 | 7 | U | T | 0.89 | 6 | 10.33 |

functionally known proteins, and there is no other clue about its function. TTYH2 and TTYH3 may represent large conductance anion channels (64). The remaining candidates correspond to orphan members from distinct transporter families.

SLC46A3 (SoLute Carrier family 46 member 3), SLC37A2 (SoLute Carrier 37 family member 2), and MFSD1 (Major Facilitator Superfamily Domain-containing protein 1) belong to distant families within the Major Facilitator Superfamily of secondary transporters (65). MFSD1, which has previously been identified in proteomics analyses of lysosomes and phagosomes (14, 16, 19, 54, 59, 60), is responsive to the transcription factor TFEB, and it was considered as a promising lysosomal protein candidate (16, 50). Its lysosomal localization has been confirmed independently during the course of our study (66). SLC37A2 mediates sugar-phosphate/phosphate and phosphate/phosphate exchange in proteoliposomes (67).

SLC38A7 (SoLute Carrier family 38 member 7) belongs to the amino acid/polyamine/organocation superfamily (68). It has been reported to transport neutral and cationic amino acids at the plasma membrane during the course of this study (69), but signal-to-noise ratios were intriguingly low, suggesting that the actual role of SLC38A7 deserves further investigation. SLCO2B1/SLC21A9/OATP2B1 is an organic anion transporter that is stimulated at acidic pH (70). TMEM104 is an orphan member of the amino acid and auxin permease transporter family.

These candidates were transiently expressed as GFP or YFP fusion proteins in HeLa cells, and their intracellular distribution was compared with the lysosomal/late endosomal marker LAMP1. Interestingly, only three candidates did not overlap with LAMP1 but localized instead at the plasma membrane (STARD10 and SLCO2B1) or in LAMP1-negative puncta (LOH12CR1; data not shown). By contrast, the distribution of the nine other candidates extensively overlapped with LAMP1 (Fig. 7), thus confirming that they reside in lysosomes and validating the predictive value of the Lys-734 list.

## DISCUSSION

The main concern in lysosome-oriented proteomic studies based on subcellular fractionation is the identification of true lysosomal residents, because of cofractionation of other organelles (37, 71). Thus, identification of lysosomal candidates requires comparison of lysosome-enriched and -nonenriched fractions. A pioneer study performed by Callaghan and coworkers (15) aimed at identifying lysosomal membrane proteins from Triton WR1339 density-shifted lysosomes, also referred to as tritosomes. However, the actual lysosomal residency of several proteins identified in this study could not be established, because of the lack of comparative approach. Later on, a study of placental lysosomal proteins took advantage of the comparison between successive steps of the preparation and used a semi-quantitative label-free spectral counting method to select 86 lysosomal candidates (16). More recently, Lobel and coworkers (19) demonstrated the potential of coupling the selective lysosome density shift induced by Triton WR-1339 injection in rats with MS quantification by isobaric peptide labeling, for simultaneous identification and validation of lysosomal candidates. In this work, we compared lysosome-enriched and -nonenriched fractions obtained from rat liver by differential centrifugation and isopycnic density gradient centrifugation, followed by detergent or organic solvent extraction steps to reduce sample complexity prior to MS analysis. Our spectral count-based analysis provided us with an extensive list of 2,385 proteins (MbL2385 list), including 32% of IMPs. Among these proteins, 734 were selected as significantly enriched in the lysosomal fraction (Lys-734 list).

To our knowledge, the MbL2385 list is the most extensive published to date for lysosomes (15–17, 19, 72), phagosomes (54–56, 59, 60, 73, 74), or lysosome-related organelles (57, 58, 75–78). Its IMP content (32%) is much higher than that commonly obtained if no specific subfractionation treatment is performed (5–15% IMPs (34)), but it is very similar to that obtained in a study of placental lysosomal membranes that also used an organic solvent treatment (16). Because of our preparation protocol, we identified altogether IMPs and membrane-associated proteins, but soluble proteins as well, such as luminal lysosomal hydrolases. Indeed, centrifugation of the lysosomal membranes leads to sedimentation of aggregated inclusions from the lysosomal matrix and thus induces the presence of soluble lysosomal enzymes and of proteins being degraded (30). Moreover, soluble proteins might also be retained as entrapped in membrane fragments generated upon hypotonic lysis and subsequent resealing of the organelles. The Lys-734 list is also longer than those pre-

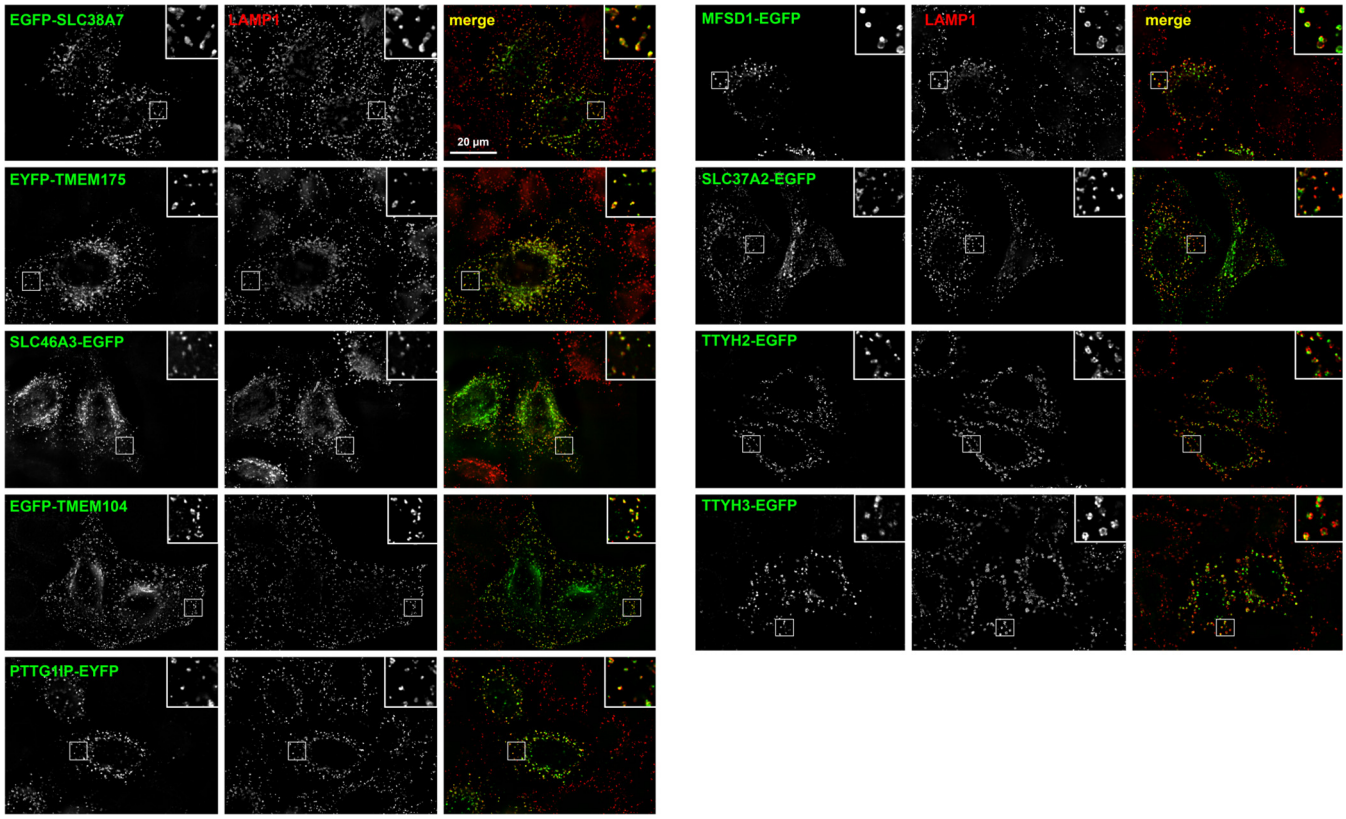


FIG. 7. **Subcellular localization of candidates in HeLa cells.** HeLa cells transiently expressing GFP- or YFP-tagged candidates were fixed, immunostained with an antibody directed against LAMP1, a late endosome and lysosome marker, and imaged by epifluorescence followed by deconvolution. Fluorescent protein-associated fluorescence, LAMP1 immunostaining, and merged images are shown from *left to right*. *Insets* are ×3 magnification of the squared area. *Scale bar,* 20 µm.

sented in other comparative proteomic studies (124 proteins in Ref. 16 and 76 in Ref. 19). Comparison of these datasets indicates nevertheless rather important overlaps of 69 and 51 proteins, respectively.

As ~80% of the EL-annotated proteins but only 3.2% of contaminant proteins were recovered in the Lys-734 list, our semi-quantitative approach was able to strongly discriminate endo-lysosomal proteins from those of recognized contaminating organelles, such as mitochondria or endoplasmic reticulum. Nevertheless, the presence of proteins annotated as non-endo-lysosomal questions the significance of their selection, beside the possibility of false-positive retention. Additionally, among the EL proteins themselves, lysosomal proteins are not distinguished from proteins from other endocytic compartments (early or late endosomes).

The presence of proteins annotated to other compartments than lysosomes may represent true lysosomal residents with multiple subcellular locations, the lysosomal residency being either predominant or secondary. Indeed, as our data were restricted to fractions issued from the isopycnic density gradient, we do not know if the "lysosome-like" behavior observed for a given protein is representative of the whole cellular pool of protein or restricted to a small, specific subset. For instance, the TGN marker TGN38 is depleted from the L fraction and mainly recovered in the PS fraction after differential centrifugation (Fig. 2). However, the minority of TGN38 proteins that cosegregated with lysosomes during differential centrifugation was concentrated in the MbL+ fraction after the subsequent centrifugation on a Nycodenz gradient (Fig. 2). A surprisingly high number of PM proteins (56%) was retained in the Lys-734 list. The presence of PM in the lysosome-enriched fraction has been discussed previously (37); it was shown that the small amount of PM proteins recovered in the L fraction (~5%) behaves like lysosomes on a metrizamide gradient, either as true PM residents or as lysosomal proteins. Migration of PM proteins between PM and lysosomes is conceivable. Indeed, the endocytic pathway constitutes a link between these two compartments, as numerous fusion/fission events occur between the various entities of the pathway (PM, endocytic vesicles, early and late endosomes, and lysosomes). Moreover, lysosomes are known to directly fuse with the PM in given circumstances (3). Such a dual localization has been suggested by observations of 5′-nucleotidase reactivity on the cytoplasmic face of lysosomes (37). This protein, which is considered as a PM marker, is notably present in the Lys-734 list.

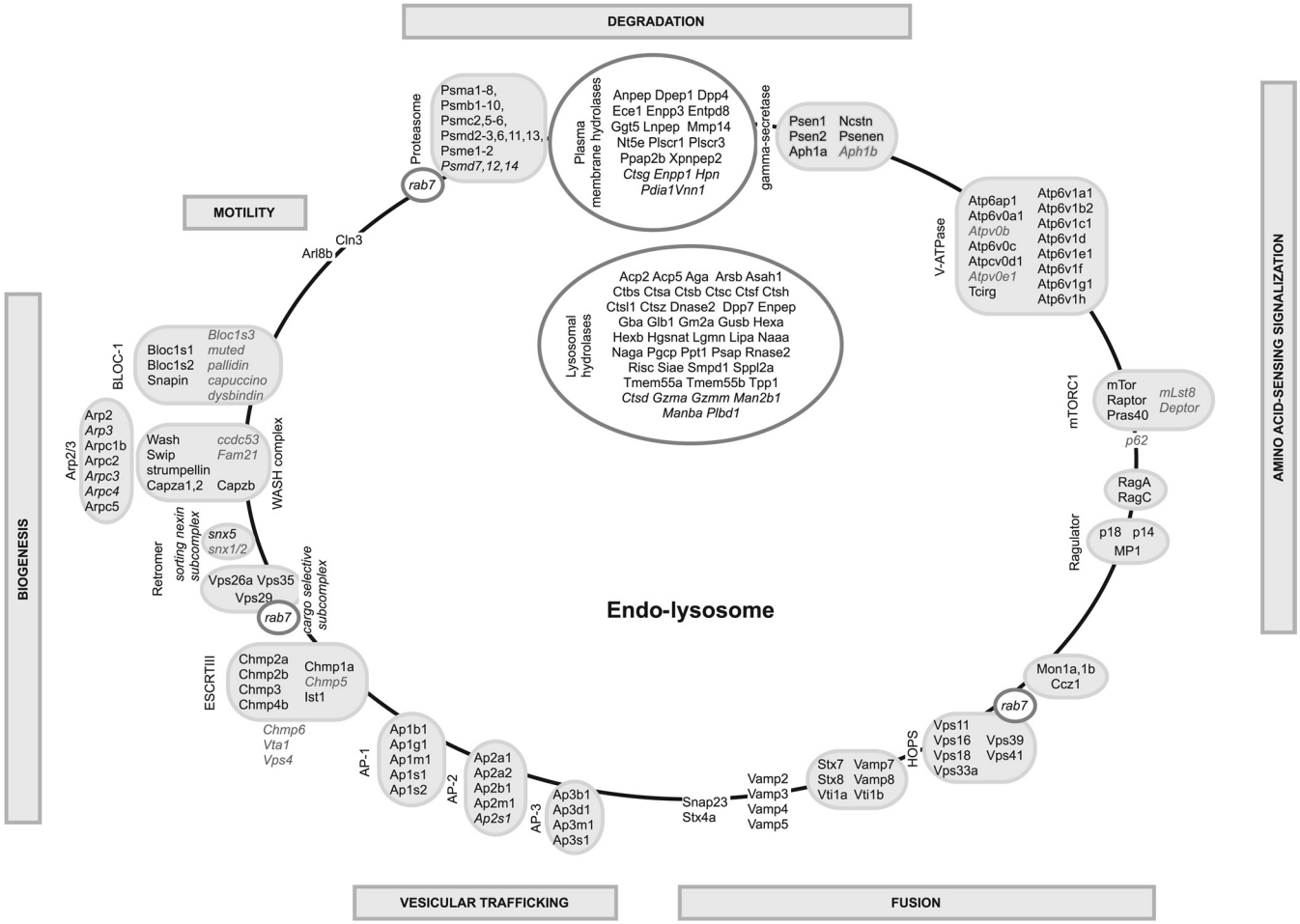


FIG. 8. **Schematic representation of the identified actors of chosen lysosomally associated processes.** A schematic endo-lysosome is drawn with the names of identified proteins implicated in chosen lysosome-associated processes. Transmembrane transport is not considered here. Well established complexes are represented on a *gray* background. All their described components are indicated, whether identified in this study or not. The Rab7 protein is indicated near the diverse complexes requiring Rab7 interaction for their endo-lysosome membrane association. *Black,* selected proteins; *black italics,* proteins identified with at least five spectra but not selected; *gray italics,* proteins neither identified nor selected.

The non-EL annotation of a candidate may also be too restrictive. For example, numerous proteins annotated as cytoplasmic or belonging to the cytoskeleton might in fact be associated with endo-lysosomes as constituents of membrane trafficking machineries that allow membrane exchange between lysosomes and other organelles or as belonging to the microtubules along which endo-lysosomes move inside the cell (79, 80). Finally, if not true lysosomal residents, the candidate proteins may be targeted to lysosomes for degradation through endocytosis or autophagy. For instance, many PM tyrosine kinase receptors, such as the EGF receptor, are down-regulated by this process (81). Only a few receptors of this type, including the EGF receptor, were, however, identified in our work.

The robustness and sensitivity of our approach were underscored by the close-to-completion identification in the Lys-734 list of numerous structural (*i.e.* vacuolar ATPase) or functional (*i.e.* $\gamma$-secretase, trafficking and nutrient-sensing machineries) complexes, known to act at the late endosomal and lysosomal membranes (detailed in Fig. 8). Many subunits of these complexes have been identified in previous proteomic studies of lysosomes (15–17, 19, 72), phagosomes (54–56, 59, 60, 73, 74), or lysosome-related organelles (57, 58, 75–78). However, these complexes were most often not as extensively documented as in this work. A few specific points are discussed below.

Beside its major role as a cytosolic proteolytic machine, the proteasome is also required for endocytic transport and sorting of receptors toward inner membranes of the multivesicular bodies (82, 83), through a specific interaction between Rab7 and the proteasome $\alpha$-subunit PMSA7 (84). Accordingly, numerous subunits ($n = 28$) of the proteasome were present in the Lys-734 list. In a previous study, 24 proteasome subunits had been found in placental lysosome membranes, although not considered as lysosomal candidates (16). Proteasome subunits had also been identified, although to a lesser extent,

in proteomic studies of phagosomes (54), lysosome-related organelles (57), or *Arabidopsis thaliana* vacuoles.[3]

Biogenesis of the lysosomes and delivery of endocytic cargo to these organelles involve numerous and highly dynamic membrane fusion and fission events between compartments of the endocytic pathway and with the secretory pathway, thus allowing protein import to, or retrieval from, lysosomes (3, 85). All complexes involved in these processes were present in the Lys-734 list (Fig. 8). Whereas numerous components of the ESCRT-III (Endosomal Sorting Complex Required for Transport-III) complex, which mediates the abscission of the newly forming intraluminal vesicles (86), were selected, none of the components of the ESCRT-0, -I, or -II complexes was identified. This was already the case in our recent proteomic analysis of the endocytic pathway of *Dictyostelium discoideum* (87) or in studies performed on the vacuolar membrane of *Arabidopsis thaliana*.[3] The origin of this apparently "tighter" association of ESCRT-III with endo-lysosomal membranes (in contrast, ESCRT-0 and -I were detected in phagosomes (54)) deserves further investigation. As for the process of homo- or heterotypic fusion between late endosomes and lysosomes, it implies an initial tethering step mediated by the HOPS (homotypic fusion and vacuole protein sorting) complex (88). Very similarly, tethering in early endosomes homotypic fusion is performed by the CORVET (class C core vacuole-endosome transport) complex, which shares four subunits with HOPS (89). Interestingly, all HOPS components were present in the Lys-734 list, although none of the specific CORVET subunits could be detected. This is similar to what was observed in a proteomic study of the yeast vacuolar membrane (90).

Similarly to endosomes (91), lysosomes are now emerging as signaling platforms with the capability to detect modifications of the cell environment, such as energy, growth factors, and nutrient levels (92). Accordingly, many actors of signaling processes were enriched in the MbL+ fraction, such as receptors, $\alpha$ subunits of the heterotrimeric G proteins, protein kinases and a few Ras-related proteins. As half of these signaling proteins were PM-annotated, their additional endocytic localization might have been ignored until now. A key signaling pathway in nutrient sensing involves the master cell growth regulator mTOR that controls autophagy in response to a wide range of signals, including amino acid availability (93). Recent studies showed that the lysosome acts as an assembly site for a sensing device, the "nutrisome," which is composed of the RagA/B-RagC/D heterodimer, the Ragulator and mTORC1 complexes, the Rheb GTPase, and the V-ATPase (92, 94). Most proteins from this pathway were present in the Lys-734 list (Fig. 8).

*Conclusions and Perspectives*—Almost a hundred proteins, in which subcellular localization had never been described nor predicted, were sorted out as novel putative lysosomal proteins in this study. Concerning molecular transporters, 46 candidates were selected, most of which were either devoid of subcellular annotation or annotated as plasma membrane proteins, suggesting a dual localization for the latter. The lysosomal subcellular localization was validated for nine candidates, including five secondary transporters, further supporting the relevance of our list of candidate lysosomal proteins. The numerous novel candidates revealed by this work should promote new research and help with understanding the cell biology, physiology, and pathophysiology of this important organelle.

*Acknowledgments*—*We thank L. Aubry and G. Klein for their stimulating support and L. Aubry for critical reading of the manuscript. We thank A. Kraut for technical help; C. Adam, D. Barthe, and V. Dupierris for the homemade software support, and Y. Vandenbrouck for Blast searches. We also thank M. Court and C. Masselon for writing the SpI analysis script. Author contributions: J.G., B.G., M.J., and A.J. conceived the project. J.T. and A.C. prepared samples for proteomics analysis. S.K.J. performed the mass spectrometry analysis. A.J. analyzed proteomic data. M.M. and A.J. performed statistical analysis. A.J., C.S., C.I., and Q.V. performed molecular biology experiments and subcellular localization analyses. C.B. supervised software developments for protein identification and parsimony analysis. A.J., B.G., M.J., and J.G. wrote the manuscript.*

* This work was supported by Fonds de la Recherche Scientifique Grant 2.4543.08 and a grant from the Cystinosis Research Foundation (to B.G.).

S This article contains supplemental material.

‡‡ To whom correspondence should be addressed: iRTSV/BGE/ODyCell, CEA-Grenoble, 17, Rue des Martyrs, 38054 Grenoble Cedex 9, France. Tel.: 33-438-785-419; Fax: 33-438-783-065; E-mail: ajournet@cea.fr.

[3] N. Jarno and M. Jaquinod, personal communication.

REFERENCES

1. Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., and Childs, B. (eds) (2001) *The Metabolic and Molecular Basis of Inherited Disease*, 8th Ed., McGraw-Hill Inc., New York
2. Lübke, T., Lobel, P., and Sleat, D. (2009) Proteomics of the lysosome. *Biochim. Biophys. Acta* **1793,** 625–635
3. Luzio, J. P., Pryor, P. R., and Bright, N. A. (2007) Lysosomes: fusion and function. *Nat. Rev. Mol. Cell Biol.* **8,** 622–632
4. Jaquinod, S. K., Chapel, A., Garin, J., and Journet, A. (2008) Affinity purification of soluble lysosomal proteins for mass spectrometric identification. *Methods Mol. Biol.* **432,** 243–258
5. Sleat, D. E., Della Valle, M. C., Zheng, H., Moore, D. F., and Lobel, P. (2008) The Mannose 6-phosphate glycoprotein proteome. *J. Proteome Res.* **7,** 3010–3021
6. Sleat, D. E., Zheng, H., Qian, M., and Lobel, P. (2006) Identification of sites of mannose 6-phosphorylation on lysosomal proteins. *Mol. Cell. Proteomics* **5,** 686–701
7. Czupalla, C., Mansukoski, H., Riedl, T., Thiel, D., Krause, E., and Hoflack, B. (2006) Proteomic analysis of lysosomal acid hydrolases secreted by osteoclasts: implications for lytic enzyme transport and bone metabolism. *Mol. Cell. Proteomics* **5,** 134–143
8. Sleat, D. E., Lackland, H., Wang, Y., Sohar, I., Xiao, G., Li, H., and Lobel, P. (2005) The human brain mannose 6-phosphate glycoproteome: A complex mixture composed of multiple isoforms of many soluble lysosomal proteins. *Proteomics* **5,** 1520–1532
9. Kollmann, K., Mutenda, K. E., Balleininger, M., Eckermann, E., von Figura, K., Schmidt, B., and Lübke, T. (2005) Identification of novel lysosomal matrix proteins by proteome analysis. *Proteomics* **5,** 3966–3978
10. Journet, A., Chapel, A., Kieffer, S., Roux, F., and Garin, J. (2002) Proteomic analysis of human lysosomes: Application to monocytic and breast cancer cells. *Proteomics* **2,** 1026–1040
11. Journet, A., Chapel, A., Kieffer, S., Louwagie, M., Luche, S., and Garin, J.

(2000) Towarda human repertoire of monocytic lysosomal proteins. *Electrophoresis* **21,** 3411–3419
12. Kornfeld, S. (1986) Trafficking of lysosomal enzymes in normal and disease states. *J. Clin. Invest.* **77,** 1–6
13. Sleat, D. E., Jadot, M., and Lobel, P. (2007) Lysosomal proteomics and disease. *Proteomics Clin. Appl.* **1,** 1134–1146
14. Zhang, H., Fan, X., Bagshaw, R. D., Zhang, L., Mahuran, D. J., and Callahan, J. W. (2007) Lysosomal membranes from beige mice contain higher than normal levels of endoplasmic reticulum proteins. *J. Proteome Res.* **6,** 240–249
15. Bagshaw, R. D., Mahuran, D. J., and Callahan, J. W. (2005) A proteomic analysis of lysosomal integral membrane proteins reveals the diverse composition of the organelle. *Mol. Cell. Proteomics* **4,** 133–143
16. Schröder, B., Wrocklage, C., Pan, C., Jäger, R., Kösters, B., Schäfer, H., Elsässer, H. P., Mann, M., and Hasilik, A. (2007) Integral and associated lysosomal membrane proteins. *Traffic* **8,** 1676–1686
17. Duclos, S., Clavarino, G., Rousserie, G., Goyette, G., Boulais, J., Camossetto, V., Gatti, E., LaBoissière, S., Pierre, P., and Desjardins, M. (2011) The endosomal proteome of macrophage and dendritic cells. *Proteomics* **11,** 854–864
18. Stahl, S., Reinders, Y., Asan, E., Mothes, W., Conzelmann, E., Sickmann, A., and Felbor, U. (2007) Proteomic analysis of cathepsin B- and L-deficient mouse brain lysosomes. *Biochim. Biophys. Acta* **1774,** 1237–1246
19. Della Valle, M. C., Sleat, D. E., Zheng, H., Moore, D. F., Jadot, M., and Lobel, P. (2011) Classification of subcellular location by comparative proteomic analysis of native and density-shifted lysosomes. *Mol. Cell. Proteomics* **10,** M110.006403
20. Jensen, A. G., Chemali, M., Chapel, A., Kieffer-Jaquinod, S., Jadot, M., Garin, J., and Journet, A. (2007) Biochemical characterization and lysosomal localization of the mannose 6-phosphate protein p76 (hypothetical protein LOC196463). *Biochem. J.* **402,** 449–458
21. Boonen, M., Hamer, I., Boussac, M., Delsaute, A. F., Flamion, B., Garin, J., and Jadot, M. (2006) Intracellular localization of p40, a protein identified in a preparation of lysosomal membranes. *Biochem. J.* **395,** 39–47
22. Deuschl, F., Kollmann, K., von Figura, K., and Lübke, T. (2006) Molecular characterization of the hypothetical 66.3-kDa protein in mouse: Lysosomal targeting, glycosylation, processing and tissue distribution. *FEBS Lett.* **580,** 5747–5752
23. Della Valle, M. C., Sleat, D. E., Sohar, I., Wen, T., Pintar, J. E., Jadot, M., and Lobel, P. (2006) Demonstration of lysosomal localization for the mammalian ependymin-related protein using classical approaches combined with a novel density-shift method. *J. Biol. Chem.* **281,** 35436–35445
24. Campomenosi, P., Salis, S., Lindqvist, C., Mariani, D., Nordström, T., Acquati, F., and Taramelli, R. (2006) Characterization of RNASET2, the first human member of the Rh/T2/S family of glycoproteins. *Arch. Biochem. Biophys.* **449,** 17–26
25. Jialin, G., Xuefan, G., and Huiwen, Z. (2010) SID1 transmembrane family, member 2 (Sidt2): a novel lysosomal membrane protein. *Biochem. Biophys. Res. Commun.* **402,** 588–594
26. Schieweck, O., Damme, M., Schröder, B., Hasilik, A., Schmidt, B., and Lübke, T. (2009) NCU-G1 is a highly glycosylated integral membrane protein of the lysosome. *Biochem. J.* **422,** 83–90
27. Savalas, L. R., Gasnier, B., Damme, M., Lübke, T., Wrocklage, C., Debacker, C., Jézégou, A., Reinheckel, T., Hasilik, A., Saftig, P., and Schröder, B. (2011) Disrupted in renal carcinoma 2 (DIRC2), a novel transporter of the lysosomal membrane, is proteolytically processed by cathepsin L. *Biochem. J.* **439,** 113–128
28. Lang, C. M., Fellerer, K., Schwenk, B. M., Kuhn, P. H., Kremmer, E., Edbauer, D., Capell, A., and Haass, C. (2012) Membrane orientation and subcellular localization of transmembrane protein 106B (TMEM106B), a major risk factor for frontotemporal lobar degeneration. *J. Biol. Chem.* **287,** 19355–19365
29. Coelho, D., Kim, J. C., Miousse, I. R., Fung, S., du Moulin, M., Buers, I., Suormala, T., Burda, P., Frapolli, M., Stucki, M., Nürnberg, P., Thiele, H., Robenek, H., Höhne, W., Longo, N., Pasquali, M., Mengel, E., Watkins, D., Shoubridge, E. A., Majewski, J., Rosenblatt, D. S., Fowler, B., Rutsch, F., and Baumgartner, M. R. (2012) Mutations in ABCD4 cause a new inborn error of vitamin B12 metabolism. *Nat. Genet.* **44,** 1152–1155
30. Schröder, B. A., Wrocklage, C., Hasilik, A., and Saftig, P. (2010) The proteome of lysosomes. *Proteomics* **10,** 4053–4076
31. Pisoni, R. L., and Thoene, J. G. (1991) The transport systems of mammalian lysosomes. *Biochim. Biophys. Acta* **1071,** 351–373
32. Mancini, G. M., Havelaar, A. C., and Verheijen, F. W. (2000) Lysosomal transport disorders. *J. Inherit. Metab. Dis.* **23,** 278–292
33. Sagné, C., and Gasnier, B. (2008) Molecular physiology and pathophysiology of lysosomal membrane transporters. *J. Inherit. Metab. Dis.* **31,** 258–266
34. Speers, A. E., and Wu, C. C. (2007) Proteomics of integral membrane proteins–theory and application. *Chem. Rev.* **107,** 3687–3714
35. Tan, S., Tan, H. T., and Chung, M. C. (2008) Membrane proteins and membrane proteomics. *Proteomics* **8,** 3924–3932
36. Jézégou, A., Llinares, E., Anne, C., Kieffer-Jaquinod, S., O'Regan, S., Aupetit, J., Chabli, A., Sagné, C., Debacker, C., Chadefaux-Vekemans, B., Journet, A., André, B., and Gasnier, B. (2012) Heptahelical protein PQLC2 is a lysosomal cationic amino acid exporter underlying the action of cysteamine in cystinosis therapy. *Proc. Natl. Acad. Sci. U.S.A.* **109,** E3434–E3443
37. Wattiaux, R., Wattiaux-De Coninck, S., Ronveaux-dupal, M. F., and Dubois, F. (1978) Isolation of rat liver lysosomes by isopycnic centrifugation in a metrizamide gradient. *J. Cell Biol.* **78,** 349–368
38. Peters, T. J., Müller, M., and De Duve, C. (1972) Lysosomes of the arterial wall. I. Isolation and subcellular fractionation of cells from normal rabbit aorta. *J. Exp. Med.* **136,** 1117–1139
39. Salvi, D., Rolland, N., Joyard, J., and Ferro, M. (2008) Purification and proteomic analysis of chloroplasts and their sub-organellar compartments. *Methods Mol. Biol.* **432,** 19–36
40. Donoghue, P. M., Hughes, C., Vissers, J. P., Langridge, J. I., and Dunn, M. J. (2008) Nonionic detergent phase extraction for the proteomic analysis of heart membrane proteins using label-free LC-MS. *Proteomics* **8,** 3895–3905
41. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685
42. Dupierris, V., Masselon, C., Court, M., Kieffer-Jaquinod, S., and Bruley, C. (2009) A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa. *Bioinformatics* **25,** 1980–1981
43. Vizcaíno, J. A., Côté, R., Reisinger, F., Foster, J. M., Mueller, M., Rameseder, J., Hermjakob, H., and Martens, L. (2009) A guide to the Proteomics Identifications Database proteomics data repository. *Proteomics* **9,** 4276–4283
44. Liu, H., Sadygov, R. G., and Yates, J. R., 3rd (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **76,** 4193–4201
45. Fu, X., Gharib, S. A., Green, P. S., Aitken, M. L., Frazer, D. A., Park, D. R., Vaisar, T., and Heinecke, J. W. (2008) Spectral index for assessment of differential protein expression in shotgun proteomics. *J. Proteome Res.* **7,** 845–854
46. Gilchrist, A., Au, C. E., Hiding, J., Bell, A. W., Fernandez-Rodriguez, J., Lesimple, S., Nagaya, H., Roy, L., Gosline, S. J., Hallett, M., Paiement, J., Kearney, R. E., Nilsson, T., and Bergeron, J. J. (2006) Quantitative proteomics analysis of the secretory pathway. *Cell* **127,** 1265–1281
47. Olsson, G. M., Svensson, I., Zdolsek, J. M., and Brunk, U. T. (1989) Lysosomal enzyme leakage during the hypoxanthine/xanthine oxidase reaction. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **56,** 385–391
48. Ford, T., Rickwood, D., and Graham, J. (1983) Buoyant densities of macromolecules, macromolecular complexes, and cell organelles in Nycodenz gradients. *Anal. Biochem.* **128,** 232–239
49. Bordier, C. (1981) Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* **256,** 1604–1607
50. Sardiello, M., Palmieri, M., di Ronza, A., Medina, D. L., Valenza, M., Gennarino, V. A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R. S., Banfi, S., Parenti, G., Cattaneo, E., and Ballabio, A. (2009) A gene network regulating lysosomal biogenesis and function. *Science* **325,** 473–477
51. Boonen, M., Rezende de Castro, R., Cuvelier, G., Hamer, I., and Jadot, M. (2008) A dileucine signal situated in the C-terminal tail of the lysosomal membrane protein p40 is responsible for its targeting to lysosomes. *Biochem. J.* **414,** 431–440
52. Rutsch, F., Gailus, S., Miousse, I. R., Suormala, T., Sagné, C., Toliat, M. R., Nürnberg, G., Wittkampf, T., Buers, I., Sharifi, A., Stucki, M., Becker, C., Baumgartner, M., Robenek, H., Marquardt, T., Höhne, W., Gasnier, B., Rosenblatt, D. S., Fowler, B., and Nürnberg, P. (2009) Identification of a putative lysosomal cobalamin exporter altered in the cblF defect of

vitamin $B_{12}$ metabolism. *Nat. Genet.* **41,** 234–239
53. Siintola, E., Topcu, M., Aula, N., Lohi, H., Minassian, B. A., Paterson, A. D., Liu, X. Q., Wilson, C., Lahtinen, U., Anttonen, A. K., and Lehesjoki, A. E. (2007) The novel neuronal ceroid lipofuscinosis gene *MFSD8* encodes a putative lysosomal transporter. *Am. J. Hum. Genet.* **81,** 136–146
54. Boulais, J., Trost, M., Landry, C. R., Dieckmann, R., Levy, E. D., Soldati, T., Michnick, S. W., Thibault, P., and Desjardins, M. (2010) Molecular characterization of the evolution of phagosomes. *Mol. Syst. Biol.* **6,** 423
55. Buschow, S. I., Lasonder, E., Szklarczyk, R., Oud, M. M., de Vries, I. J., and Figdor, C. G. (2012) Unraveling the human dendritic cell phagosome proteome by organellar enrichment ranking. *J. Proteomics* **75,** 1547–1562
56. Li, Q., Singh, C. R., Ma, S., Price, N. D., and Jagannath, C. (2011) Label-free proteomics and systems biology analysis of mycobacterial phagosomes in dendritic cells and macrophages. *J. Proteome Res.* **10,** 2425–2439
57. Raymond, A. A., Gonzalez de Peredo, A., Stella, A., Ishida-Yamamoto, A., Bouyssie, D., Serre, G., Monsarrat, B., and Simon, M. (2008) Lamellar bodies of human epidermis: proteomics characterization by high throughput mass spectrometry and possible involvement of CLIP-170 in their trafficking/secretion. *Mol. Cell. Proteomics* **7,** 2151–2175
58. Schmidt, H., Gelhaus, C., Nebendahl, M., Lettau, M., Lucius, R., Leippe, M., Kabelitz, D., and Janssen, O. (2011) Effector granules in human T lymphocytes: proteomic evidence for two distinct species of cytotoxic effector vesicles. *J. Proteome Res.* **10,** 1603–1620
59. Shui, W., Sheu, L., Liu, J., Smart, B., Petzold, C. J., Hsieh, T. Y., Pitcher, A., Keasling, J. D., and Bertozzi, C. R. (2008) Membrane proteomics of phagosomes suggests a connection to autophagy. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 16952–16957
60. Shui, W., Petzold, C. J., Redding, A., Liu, J., Pitcher, A., Sheu, L., Hsieh, T. Y., Keasling, J. D., and Bertozzi, C. R. (2011) Organelle membrane proteomics reveals differential influence of mycobacterial lipoglycans on macrophage phagosome maturation and autophagosome accumulation. *J. Proteome Res.* **10,** 339–348
61. Trost, M., English, L., Lemieux, S., Courcelles, M., Desjardins, M., and Thibault, P. (2009) The phagosomal proteome in interferon-gamma-activated macrophages. *Immunity* **30,** 143–154
62. Chien, W., and Pei, L. (2000) A novel binding factor facilitates nuclear translocation and transcriptional activation function of the pituitary tumor-transforming gene product. *J. Biol. Chem.* **275,** 19422–19427
63. Smith, V. E., Read, M. L., Turnell, A. S., Sharma, N., Lewy, G. D., Fong, J. C., Seed, R. I., Kwan, P., Ryan, G., Mehanna, H., Chan, S. Y., Darras, V. M., Boelaert, K., Franklyn, J. A., and McCabe, C. J. (2012) PTTG-binding factor (PBF) is a novel regulator of the thyroid hormone transporter MCT8. *Endocrinology* **153,** 3526–3536
64. Suzuki, M., and Mizuno, A. (2004) A novel human $Cl^-$ channel family related to *Drosophila* flightless locus. *J. Biol. Chem.* **279,** 22461–22468
65. Pao, S. S., Paulsen, I. T., and Saier, M. H., Jr. (1998) Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62,** 1–34
66. Palmieri, M., Impey, S., Kang, H., di Ronza, A., Pelz, C., Sardiello, M., and Ballabio, A. (2011) Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum. Mol. Genet.* **20,** 3852–3866
67. Pan, C. J., Chen, S. Y., Jun, H. S., Lin, S. R., Mansfield, B. C., and Chou, J. Y. (2011) SLC37A1 and SLC37A2 are phosphate-linked, glucose 6-phosphate antiporters. *PLoS ONE* **6,** e23157
68. Jack, D. L., Paulsen, I. T., and Saier, M. H. (2000) The amino acid/polyamine/organocation (APC) superfamily of transporters specific for amino acids, polyamines, and organocations. *Microbiology* **146,** 1797–1814
69. Hägglund, M. G., Sreedharan, S., Nilsson, V. C., Shaik, J. H., Almkvist, I. M., Bäcklin, S., Wrange, O., and Fredriksson, R. (2011) Identification of SLC38A7 (SNAT7) protein as a glutamine transporter expressed in neurons. *J. Biol. Chem.* **286,** 20500–20511
70. Kobayashi, D., Nozawa, T., Imai, K., Nezu, J., Tsuji, A., and Tamai, I. (2003) Involvement of human organic anion transporting polypeptide OATP-B (SLC21A9) in pH-dependent transport across intestinal apical membrane. *J. Pharmacol. Exp. Ther.* **306,** 703–708
71. Bergeron, J. J., Au, C. E., Desjardins, M., McPherson, P. S., and Nilsson, T. (2010) Cell biology through proteomics–ad astra per alia porci. *Trends Cell Biol.* **20,** 337–345
72. Nylandsted, J., Becker, A. C., Bunkenborg, J., Andersen, J. S., Dengjel, J., and Jäättelä, M. (2011) ErbB2-associated changes in the lysosomal proteome. *Proteomics* **11,** 2830–2838
73. Garin, J., Diez, R., Kieffer, S., Dermine, J. F., Duclos, S., Gagnon, E., Sadoul, R., Rondeau, C., and Desjardins, M. (2001) The phagosome proteome: insight into phagosome functions. *J. Cell Biol.* **152,** 165–180
74. Rogers, L. D., and Foster, L. J. (2007) The dynamic phagosomal proteome and the contribution of the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 18520–18525
75. Schmidt, H., Gelhaus, C., Nebendahl, M., Lettau, M., Lucius, R., Leippe, M., Kabelitz, D., and Janssen, O. (2011) Effector granules in human T lymphocytes: the luminal proteome of secretory lysosomes from human T cells. *Cell Commun. Signal.* **9,** 4
76. Ridsdale, R., Na, C. L., Xu, Y., Greis, K. D., and Weaver, T. (2011) Comparative proteomic analysis of lung lamellar bodies and lysosome-related organelles. *PLoS ONE* **6,** e16482
77. Lominadze, G., Powell, D. W., Luerman, G. C., Link, A. J., Ward, R. A., and McLeish, K. R. (2005) Proteomic analysis of human neutrophil granules. *Mol. Cell. Proteomics* **4,** 1503–1521
78. Basrur, V., Yang, F., Kushimoto, T., Higashimoto, Y., Yasumoto, K., Valencia, J., Muller, J., Vieira, W. D., Watabe, H., Shabanowitz, J., Hearing, V. J., Hunt, D. F., and Appella, E. (2003) Proteomic analysis of early melanosomes: identification of novel melanosomal proteins. *J. Proteome Res.* **2,** 69–79
79. Caviston, J. P., and Holzbaur, E. L. (2006) Microtubule motors at the intersection of trafficking and transport. *Trends Cell Biol.* **16,** 530–537
80. Krendel, M., and Mooseker, M. S. (2005) Myosins: tails (and heads) of functional diversity. *Physiology* **20,** 239–251
81. Sorkin, A., and von Zastrow, M. (2009) Endocytosis and signalling: intertwining molecular networks. *Nat. Rev. Mol. Cell Biol.* **10,** 609–622
82. Melman, L., Geuze, H. J., Li, Y., McCormick, L. M., Van Kerkhof, P., Strous, G. J., Schwartz, A. L., and Bu, G. (2002) Proteasome regulates the delivery of LDL receptor-related protein into the degradation pathway. *Mol. Biol. Cell* **13,** 3325–3335
83. Longva, K. E., Blystad, F. D., Stang, E., Larsen, A. M., Johannessen, L. E., and Madshus, I. H. (2002) Ubiquitination and proteasomal activity is required for transport of the EGF receptor to inner membranes of multivesicular bodies. *J. Cell Biol.* **156,** 843–854
84. Dong, J., Chen, W., Welford, A., and Wandinger-Ness, A. (2004) The proteasome $\alpha$-subunit XAPC7 interacts specifically with Rab7 and late endosomes. *J. Biol. Chem.* **279,** 21334–21342
85. McGough, I. J., and Cullen, P. J. (2011) Recent advances in retromer biology. *Traffic* **12,** 963–971
86. Adell, M. A., and Teis, D. (2011) Assembly and disassembly of the ESCRT-III membrane scission complex. *FEBS Lett.* **585,** 3191–3196
87. Journet, A., Klein, G., Brugière, S., Vandenbrouck, Y., Chapel, A., Kieffer, S., Bruley, C., Masselon, C., and Aubry, L. (2012) Investigating the macropinocytic proteome of *Dictyostelium amoebae* by high-resolution mass spectrometry. *Proteomics* **12,** 241–245
88. Pryor, P. R., and Luzio, J. P. (2009) Delivery of endocytosed membrane proteins to the lysosome. *Biochim. Biophys. Acta* **1793,** 615–624
89. Epp, N., Rethmeier, R., Krämer, L., and Ungermann, C. (2011) Membrane dynamics and fusion at late endosomes and vacuoles–Rab regulation, multisubunit tethering complexes and SNAREs. *Eur. J. Cell Biol.* **90,** 779–785
90. Wiederhold, E., Gandhi, T., Permentier, H. P., Breitling, R., Poolman, B., and Slotboom, D. J. (2009) The yeast vacuolar membrane proteome. *Mol. Cell. Proteomics* **8,** 380–392
91. Scita, G., and Di Fiore, P. P. (2010) The endocytic matrix. *Nature* **463,** 464–473
92. Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A. L., Nada, S., and Sabatini, D. M. (2010) Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* **141,** 290–303
93. Sengupta, S., Peterson, T. R., and Sabatini, D. M. (2010) Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Mol. Cell* **40,** 310–322
94. Zoncu, R., Bar-Peled, L., Efeyan, A., Wang, S., Sancak, Y., and Sabatini, D. M. (2011) mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar $H^+$-ATPase. *Science* **334,** 678–683