

# CLN3 is required for the clearance of glycerophosphodiester from lysosomes

<https://doi.org/10.1038/s41586-022-05221-y>

Received: 25 April 2020

Accepted: 10 August 2022

Published online: 21 September 2022

 Check for updates

Noof N. Laqtom<sup>1,2,3</sup>, Wentao Dong<sup>1,2,3</sup>, Uche N. Medoh<sup>1,2,3,4</sup>, Andrew L. Cangelosi<sup>5,6,7</sup>, Vimisha Dharamdasani<sup>7</sup>, Sze Ham Chan<sup>7,11</sup>, Tenzin Kunchok<sup>7</sup>, Caroline A. Lewis<sup>7</sup>, Ivonne Heinze<sup>8</sup>, Rachel Tang<sup>9</sup>, Christian Grimm<sup>9</sup>, An N. Dang Do<sup>10</sup>, Forbes D. Porter<sup>10</sup>, Alessandro Ori<sup>9</sup>, David M. Sabatini<sup>12,13</sup> & Monther Abu-Remaileh<sup>1,2,3,12</sup>✉

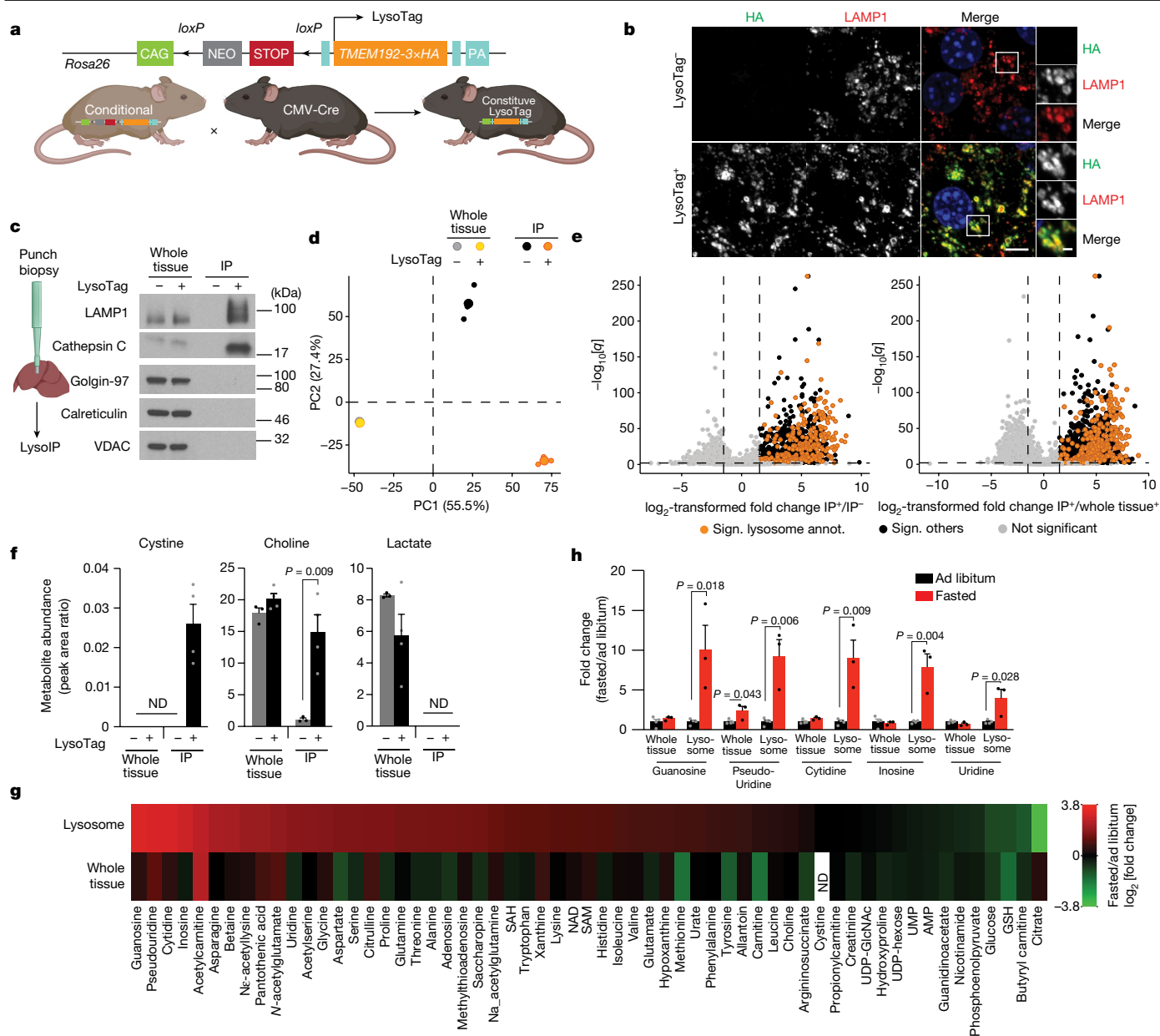
Lysosomes have many roles, including degrading macromolecules and signalling to the nucleus<sup>1</sup>. Lysosomal dysfunction occurs in various human conditions, such as common neurodegenerative diseases and monogenic lysosomal storage disorders (LSDs)<sup>2–4</sup>. For most LSDs, the causal genes have been identified but, in some, the function of the implicated gene is unknown, in part because lysosomes occupy a small fraction of the cellular volume so that changes in lysosomal contents are difficult to detect. Here we develop the LysoTag mouse for the tissue-specific isolation of intact lysosomes that are compatible with the multimodal profiling of their contents. We used the LysoTag mouse to study CLN3, a lysosomal transmembrane protein with an unknown function. In children, the loss of *CLN3* causes juvenile neuronal ceroid lipofuscinosis (Batten disease), a lethal neurodegenerative LSD. Untargeted metabolite profiling of lysosomes from the brains of mice lacking CLN3 revealed a massive accumulation of glycerophosphodiester (GPDs)—the end products of glycerophospholipid catabolism. GPDs also accumulate in the lysosomes of CLN3-deficient cultured cells and we show that CLN3 is required for their lysosomal egress. Loss of CLN3 also disrupts glycerophospholipid catabolism in the lysosome. Finally, we found elevated levels of glycerophosphoinositol in the cerebrospinal fluid of patients with Batten disease, suggesting the potential use of glycerophosphoinositol as a disease biomarker. Our results show that CLN3 is required for the lysosomal clearance of GPDs and reveal Batten disease as a neurodegenerative LSD with a defect in glycerophospholipid metabolism.

Lysosomes are membrane-bound organelles that degrade macromolecules and clear damaged organelles as well as regulate cellular signalling, such as the activation of the mTORC1 pathway by nutrients<sup>1,5,6</sup>. Mutations in genes encoding lysosomal proteins cause severe disorders that are collectively known as LSDs<sup>2,3,7–9</sup>. Many LSDs are associated with neurodegeneration, indicating a critical role for the lysosome in neuronal homeostasis<sup>3,4</sup>. Consistent with this, human genetic studies have implicated impaired lysosomal function in age-associated neurodegenerative diseases, including Alzheimer disease, Parkinson's disease and frontotemporal dementia<sup>2,10,11</sup>. The function of several lysosomal proteins implicated in disease is incompletely understood, in part because lysosomes usually occupy only 1–3% of the cellular volume and changes in lysosomal contents can therefore be difficult to ascertain in whole cells or tissues<sup>12</sup>.

## The development of the LysoTag mouse

Given this and the increasing interest in the roles of lysosomes in physiology and disease, we adapted our LysoIP method for the rapid purification of lysosomes from cultured cells<sup>12</sup> for use in mouse tissues. We generated a transgenic mouse line by integrating the cDNA encoding the lysosomally localized TMEM192–3×HA fusion protein into the *Rosa26* locus downstream of a lox-stop-lox (LSL) cassette<sup>13</sup> (Fig. 1a). By crossing these mice with those expressing a Cre recombinase driven by the CMV promoter, we generated LysoTag mice that constitutively express TMEM192–3×HA across tissues (Fig. 1a and Extended Data Fig. 1a). TMEM192–3×HA marks lysosomes with a triple HA epitope, enabling their immunoprecipitation with anti-HA antibodies. Immunofluorescence-based analyses of liver and brain sections, two

<sup>1</sup>Department of Chemical Engineering, Stanford University, Stanford, CA, USA. <sup>2</sup>Department of Genetics, Stanford University, Stanford, CA, USA. <sup>3</sup>The Institute for Chemistry, Engineering & Medicine for Human Health (ChEM-H), Stanford University, Stanford, CA, USA. <sup>4</sup>Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA. <sup>5</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. <sup>6</sup>Koch Institute for Integrative Cancer Research, Cambridge, MA, USA. <sup>7</sup>Whitehead Institute for Biomedical Research, Cambridge, MA, USA. <sup>8</sup>Leibniz Institute on Aging—Fritz Lipmann Institute, Jena, Germany. <sup>9</sup>Walther Straub Institute of Pharmacology and Toxicology, Faculty of Medicine, Ludwig-Maximilians-Universität, Munich, Germany. <sup>10</sup>Division of Translational Medicine, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, Bethesda, MD, USA. <sup>11</sup>Present address: Department of Pharmacology, University of Virginia School of Medicine, Charlottesville, VA, USA. <sup>12</sup>These authors contributed equally: David M. Sabatini, Monther Abu-Remaileh. <sup>13</sup>Unaffiliated: David M. Sabatini. ✉e-mail: monther@stanford.edu



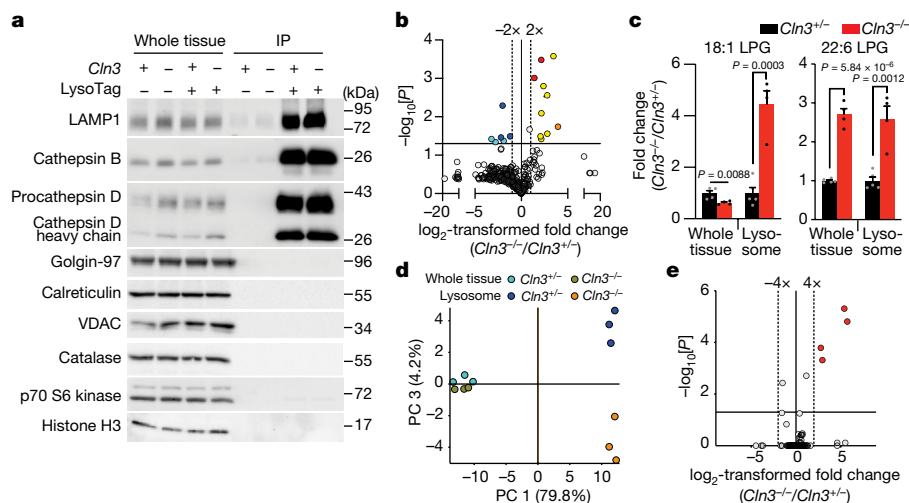
**Fig. 1 | LysoTag mouse for proteomic and metabolite profiling of tissue lysosomes.** **a**, Schematic of the *TMEM192-3×HA* fusion gene (LysoTag) in the *Rosa26* locus and the generation of the constitutive LysoTag line.

**b**, Immunofluorescence analyses of *TMEM192-3×HA* (HA, green) and lysosomes (LAMP1, red) in the liver. Nuclei are shown in blue. Scale bars, 5  $\mu$ m (main images) and 1  $\mu$ m (insets). Micrographs are representative of three independent experiments. **c**, Immunoblot analyses of protein markers of subcellular compartments in whole tissues, purified lysosomes (immunoprecipitate: IP LysoTag<sup>+</sup>), and control immunoprecipitates (IP LysoTag<sup>-</sup>). Golgin-97, VDAC and calreticulin were used as markers for the Golgi, mitochondria and endoplasmic reticulum, respectively, and LAMP1 and cathepsin C for lysosomes. Immunoblots are representative of three independent experiments. **d, e**, Proteomic analyses of liver LysoIPs. **d**, PCA comparing the levels of the detected proteins. **e**, Volcano plot showing that previously annotated lysosomal proteins (orange dots) are enriched in lysosomes. Enriched unannotated proteins are shown in black. The vertical dashed lines indicate a  $\log_2$ -transformed fold change of >1.5 in LysoTag<sup>+</sup>

immunoprecipitates either over control immunoprecipitates (left) or whole tissue (right). The horizontal dashed lines indicate  $q < 0.01$ .  $n = 3$  male mice with an average age of 6.5 weeks. Statistical analysis was performed using paired *t*-tests, performed at the peptide precursor level; *P* values were corrected for multiple testing according to ref. <sup>42</sup>. Annot., annotation; sign., significant. **f**, The relative abundance of the indicated metabolites in whole tissue and immunoprecipitates. ND, not detected. Data are mean  $\pm$  s.e.m.  $n = 4$  (LysoTag<sup>+</sup>) and  $n = 3$  (LysoTag<sup>-</sup>). **g**, Heat map of the mean of  $\log_2$ -transformed fold changes in metabolite levels after a 16 h fast (Supplementary Table 2). **h**, Fold changes in the levels of nucleosides after fasting. For **g** and **h**, data are mean and for **h**  $\pm$  s.e.m. of  $n = 4$  (ad libitum) or  $n = 3$  (fasted). For **f** and **h**, statistical analysis was performed using two-tailed unpaired *t*-tests. For **f-h**, male mice with an average age of 5.5 weeks were used. Gel source data are provided in Supplementary Fig. 1. The drawings in the schematics in **a** and **c** were created using Biorender. GlcNAc, N-acetylglucosamine; GSH, Glutathione; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine.

tissues examined in this study, confirmed that *TMEM192-3×HA* colocalizes with the lysosomal marker LAMP1 (Fig. 1b and Extended Data Fig. 2b). Other than a 10–15% decrease in body weight, the constitutive

LysoTag mice appear indistinguishable from their wild-type (WT) counterparts (Extended Data Fig. 1b). Importantly, comprehensive characterizations confirm that LysoTag expression has no effect on



**Fig. 2 | Loss of CLN3 leads to significant alterations in metabolite levels in brain lysosomes.** **a**, Immunoblot analyses of organellar protein markers in whole-tissue fractions, purified lysosomes and control immunoprecipitates. LAMP1 and cathepsin B and D were used as markers for lysosomes. Golgin-97, calreticulin, VDAC, catalase, S6 kinase and histone H3 were used as markers for the Golgi, endoplasmic reticulum, mitochondria, peroxisomes, the cytosol and the nucleus, respectively. Immunoblots are representative of three independent experiments. **b**, Volcano plot comparing untargeted LysoIP lipidomic data from lysosomes derived from brains of *Cln3<sup>+/+</sup>* and *Cln3<sup>-/-</sup>* mice. Significantly changing lipids belong to phospholipids (yellow), LPG (red), lysophosphatidylcholine (orange) and BMP (blue). BMP/PG annotation was used when MS/MS fragmentation was not acquired (cyan). Data are provided in Supplementary Table 3. The horizontal line indicates  $P = 0.05$ , and the vertical dashed lines indicate a fold change of 2.  $P$  values were calculated using

two-tailed unpaired  $t$ -tests. **c**, Targeted analyses of LPGs. Fold changes in LPGs in lysosomes from *Cln3<sup>+/+</sup>* and *Cln3<sup>-/-</sup>* mice were calculated after subtracting the background from control immunoprecipitates. Data are mean  $\pm$  s.e.m. Statistical analysis was performed using two-tailed unpaired  $t$ -tests. For **b** and **c**,  $n = 5$  (*Cln3<sup>+/+</sup>*) and  $n = 4$  (*Cln3<sup>-/-</sup>*) female mice (aged 7 months). **d**, e, Untargeted polar metabolite analyses of brain lysosomes and tissues after CLN3 loss (Methods and Supplementary Table 4). **d**, PCA comparing the levels of the 189 unique compounds (see the main text) between whole-tissue or lysosome samples from the brains of *Cln3<sup>+/+</sup>* and *Cln3<sup>-/-</sup>* mice. **e**, The lysosomal changes in these compounds. The horizontal line indicates  $P = 0.05$ , and the vertical dashed lines represent fold changes of 4.  $n = 3$  male mice (aged 7 months) per genotype.  $P$  values were calculated using analysis of variance (ANOVA) with Tukey HSD test;  $P$  values were corrected using the Benjamini–Hochberg method; FDR = 5%. Gel source data are provided in Supplementary Fig. 1.

lysosomal proteins, enzymatic activity or lysosomal ultrastructure (Extended Data Fig. 1a,c–f).

We modified our protocol<sup>12</sup> for preparing anti-HA immunoprecipitates from cultured cells for use with tissues and applied it to livers from control or LysoTag mice. Immunoprecipitates from LysoTag mice (hereafter, LysoIPs), but not from control mice, were substantially enriched for the lysosomal membrane and luminal proteins LAMP1 and cathepsin C, respectively, but free of markers of other organelles, including the Golgi, endoplasmic reticulum and mitochondria (Fig. 1c). Furthermore, data-independent acquisition-based proteomics confirmed that the LysoIPs (Fig. 1d,e and Supplementary Table 1) were highly enriched for proteins that were previously annotated as lysosomal regardless of whether we determined lysosomal enrichment relative to the abundance of these proteins in control immunoprecipitates or whole-liver lysates (Fig. 1e). The LysoIPs were mostly depleted for proteins from other compartments, except for a select set of proteins that are mostly associated with the Golgi, suggesting a possible interaction between the organelles (Supplementary Table 1). Lysosomal membrane and luminal proteins were enriched to similar degrees, indicating that lysosomes purified by the LysoIP protocol are intact (Supplementary Table 1). Finally, we observed enrichment of some endosomal markers, although to a lower extent than those for lysosomes (Supplementary Table 1 and Extended Data Fig. 1g).

To test the feasibility of our method for studying lysosomal metabolites, we measured cystine, which is known to be enriched in lysosomes<sup>12,14</sup>. Indeed, the liver LysoIPs had much higher cystine levels compared with the control immunoprecipitates or whole-liver samples, in which cystine levels were below the limit of detection (Fig. 1f). The LysoIPs also contained metabolites that we previously detected in lysosomes from cultured cells<sup>12</sup>, including choline (Fig. 1f and Supplementary Table 2). Importantly, abundant cytosolic metabolites, such as lactate, were either not detected in the LysoIPs or equally abundant in

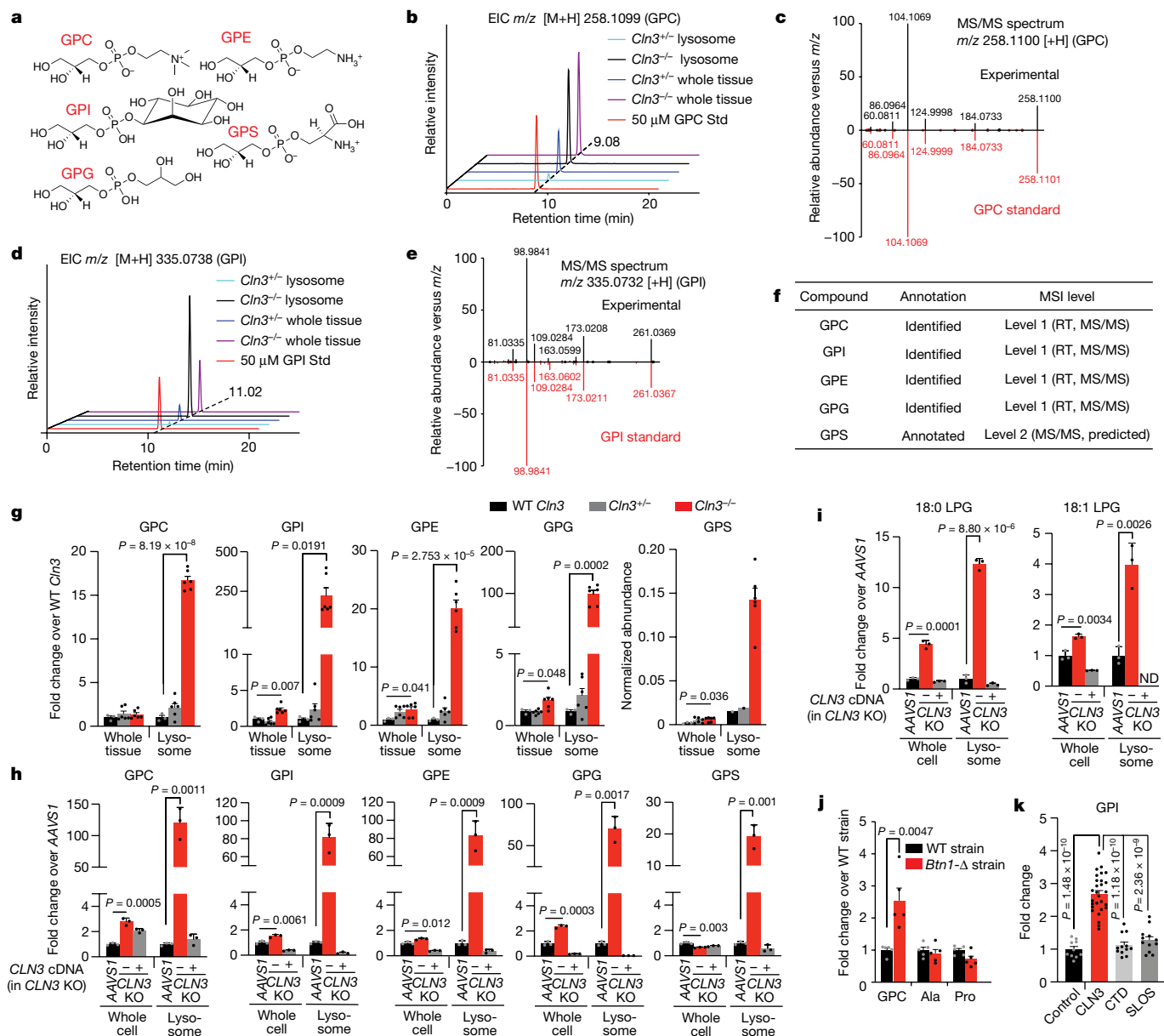
the control immunoprecipitates that were used to reveal metabolites that bind non-specifically to the magnetic beads (Fig. 1f and Supplementary Table 2).

To test how lysosomal contents change under different physiological states, we generated targeted metabolite profiles of lysosomes from the livers of mice fed ad libitum or fasted. Whereas some of the fasting-induced changes in lysosomal metabolites were also detected at the whole-tissue level (such as acetyl-carnitine), others were observed only in lysosomes (Fig. 1g and Supplementary Table 2). For example, fasting increased the levels in lysosomes of all nucleosides, without affecting those in whole-liver samples, except for pseudouridine (Fig. 1g,h and Supplementary Table 2). Collectively, these data show that the LysoIP can be used to rapidly isolate intact and highly pure lysosomes from mouse organs and to study metabolite changes that are not detectable using traditional tissue-based metabolite profiling.

### CLN3 loss alters lysosomal metabolome

To determine the ability of the LysoIP approach for deorphaning gene function, we turned to *CLN3*, the gene that is mutated in the lethal neurodegenerative disorder Batten disease<sup>15,16</sup>. The function of *CLN3* remains unclear<sup>17,18</sup>. The *CLN3* gene product is a multi-pass transmembrane protein that localizes to the lysosome<sup>19–23</sup> and its loss is associated with defects in multiple cellular functions, including autophagy, membrane fusion, vesicular trafficking and arginine transport<sup>17,18</sup>. Although its localization might suggest that it transports molecules across the lysosomal membrane, only recently has CLN3 been suggested to be an atypical solute carrier based on structural predictions<sup>24</sup>.

To identify changes in the lysosomal metabolome caused by CLN3 loss, we crossed the LysoTag mice with *Cln3<sup>-/-</sup>* mice<sup>25,26</sup> (Extended Data Fig. 2a). As in the liver, TMEM192–3×HA colocalized with LAMP1-positive lysosomes in the brains of all genotypes (Extended Data Fig. 2b). Consistent



**Fig. 3 | GPDs accumulate in lysosomes after CLN3 loss.** **a**, The chemical structure of GPDs in this study. **b**, Extracted ion chromatogram (EIC) analysis of GPC across a range of samples alongside an authentic standard (Std) showing a matching retention time. **c**, Mirror plot for an MS/MS spectrum of GPC in a representative sample (lysosomes from *Cln3*<sup>-/-</sup> mouse brain) and an authentic standard. **d**, Extracted ion chromatogram as described in **b** but for GPI. **e**, Mirror plot for an MS/MS spectrum as described in **c** but for GPI. **f**, The level of confidence in the identification of GPDs in this study according to MSI guidelines. **g**, Targeted analyses of GPDs in brain lysosomes. Fold changes between lysosomes were calculated after subtracting the background signal in control immunoprecipitates and normalizing to methionine. Glycerophosphoserine (GPS) was detected in only one WT and *Cln3*<sup>-/-</sup> lysosomal sample; thus, its abundance was reported instead. Data are mean  $\pm$  s.e.m.  $n = 3$  (WT *Cln3*),  $n = 6$  (*Cln3*<sup>-/-</sup>) and  $n = 6$  (*Cln3*<sup>-/-</sup>). *P* values presented are between

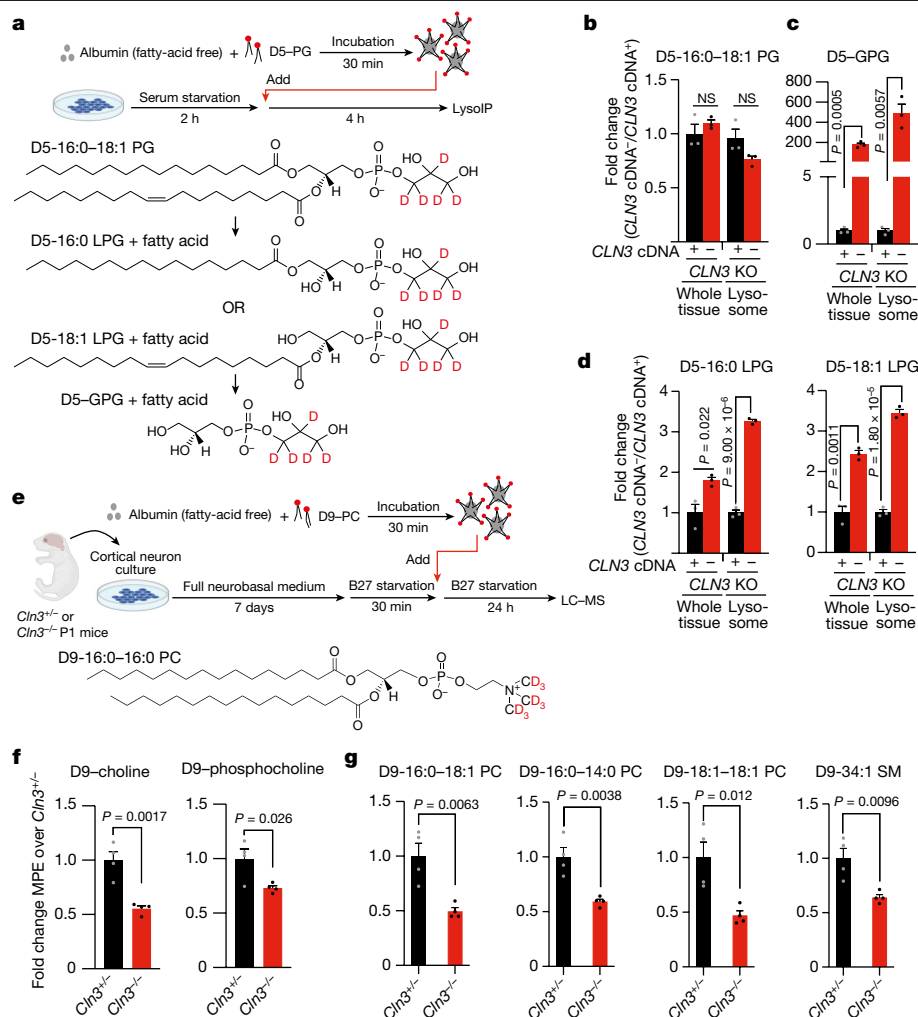
WT *Cln3* and *Cln3*<sup>-/-</sup>. Male mice from two independent experiments, including that in Fig. 2e. **h**, Fold change in the levels of GPDs between HEK293T cells and their lysosomes after *CLN3* knockout and rescue. sgRNA targeting the *AAVS1* locus was used as a control. **i**, Targeted analyses of LPGs from cells as described in **h**. For **h** and **i**, data are mean  $\pm$  s.d. of  $n = 3$  biologically independent samples. **j**, The fold change in the levels of GPC, alanine (Ala) and proline (Pro) between WT and *Btn1* $\Delta$  yeast strains. Data are mean  $\pm$  s.e.m.  $n = 5$  biologically independent samples. **k**, The fold change in GPI levels in the cerebrospinal fluid from patients with *CLN3* Batten disease (*CLN3*,  $n = 28$ ), creatine transporter deficiency (CTD,  $n = 12$ ) and Smith–Lemli–Opitz Syndrome (SLOS,  $n = 12$ ) compared with anonymized paediatric control individuals (control,  $n = 10$ ). Data are mean  $\pm$  s.e.m. Characteristics of the study participants are provided in Supplementary Table 5. For **g**–**k**, statistical analysis was performed using two-tailed unpaired *t*-tests.

with previous reports, LAMP1-positive compartments in neurons lacking CLN3 were noticeably more abundant than in control cells<sup>27</sup> (Extended Data Fig. 2b). Moreover, our protocol successfully yielded pure and intact lysosomes from the brain, regardless of CLN3 status (Fig. 2a).

As alterations in lysosome-mediated lipid catabolism have been associated with multiple forms of LSDs<sup>28,29</sup>, we used untargeted lipidomics (Methods) to analyse brain lysosomes from the mice lacking CLN3.

Consistent with previous reports<sup>30</sup>, we detected a significant decrease in tissue and lysosomal levels of bis(monoacylglycero)phosphates (BMPs), a class of lysosomal lipids (Fig. 2b, Extended Data Fig. 2c and Supplementary Table 3). Interestingly, we also found significant increases in various glycerophospholipids (hereafter phospholipids) (Supplementary Table 3), including the degradation intermediate lysophosphatidylglycerol (LPG) (Fig. 2b,c and Extended Data Fig. 2c). Although these changes





**Fig. 4 | CLN3 is required for the efflux of GPDs out of lysosomes.**

**a**, An approach to deliver exogenously added deuterated phospholipids to lysosomes. The cell culture drawing in the schematic was created using BioRender. **b**, CLN3 loss does not affect the cellular uptake of exogenous D5-PG (deuterated phosphatidylglycerol D5-16:0-18:1 PG). Fold changes in the whole-cell and lysosomal abundances of D5-PG in CLN3-KO cells relative to those in the CLN3-KO cells rescued with the CLN3 cDNA. **c**, Loss of CLN3 blocks the efflux of GPG from lysosomes. Data are presented as described in **b** but for D5-GPG. **d**, LPGs generated in lysosomes accumulate upon CLN3 loss. Data are presented as in **b** but for the two possible D5-LPGs that are generated from the degradation of the exogenously added D5-PG. For **b-d**, data are mean  $\pm$  s.d. of  $n = 3$ . **e-g**, GPC-derived choline contributes to lipid biosynthesis in a CLN3-dependent manner in primary cortical neurons. **e**, The approach to trace the metabolic fate of the choline moiety of GPC, which is generated from

phosphatidylcholine (PC) degradation in the lysosome. PC with nine deuterium atoms in its choline headgroup (D9-16:0-16:0 PC) was delivered to lysosomes in primary cultured neurons isolated from the cerebral cortex of *Cln3*<sup>+/+</sup> (control) or *Cln3*<sup>-/-</sup> mice. The mouse and cell culture drawings are from BioRender. **f**, CLN3 loss significantly reduces the cellular levels of deuterated phosphocholine and its precursor choline. **g**, CLN3 loss leads to a reduced contribution of choline derived from lysosomal GPC to the biosynthesis of PC and sphingomyelin (SM) in the cell. Data are presented as fold changes in the whole-cell molar percent enrichment (MPE) of D9-choline-containing metabolites in cortical neuron cultures prepared from *Cln3*<sup>-/-</sup> mice relative to those from *Cln3*<sup>+/+</sup> animals. For **f** and **g**, data are mean  $\pm$  s.e.m. of  $n = 4$ . For **b-d**, **f** and **g**, statistical analysis was performed using two-tailed unpaired *t*-tests; and  $n$  values represents biologically independent samples. NS, not significant.

could represent a link between CLN3 function and phospholipid catabolism in the lysosome, which was previously noted in yeast models<sup>31</sup>, it is not clear how they are directly linked to CLN3 loss.

Thus, we turned to untargeted polar metabolomics to search for additional lysosomal metabolites that could be affected by CLN3 loss (Methods and Extended Data Fig. 3a). We detected 1,073 putative compounds in lysosomes after background removal and signal thresholding (Supplementary Table 4). After further filtering for compounds in known compound databases and with reliable fragmentation data, we arrived at 189 unique compounds for which we had various degrees of confidence in their identities based on the Metabolomics Standards Initiative (MSI) guidelines (Extended Data Fig. 3a and Supplementary Table 4). Principal component analyses (PCA) of the 189 compounds revealed that, although the metabolic landscapes of the *Cln3*<sup>-/-</sup> and

*Cln3*<sup>+/+</sup> whole-brain samples were similar, those in their respective lysosomes were distinct (Fig. 2d); four compounds at high levels in lysosomes lacking CLN3 were the most differentially abundant (Fig. 2e). Critically, these same four compounds were the most significantly affected by CLN3 loss independently of how we filtered the datasets (Supplementary Table 4), indicating that CLN3 function affects a specific set of metabolites in lysosomes rather than the whole lysosomal metabolome.

### CLN3-deficient lysosomes accumulate GPDs

To determine the chemical identity of the accumulating metabolites, we used MSI as well as tandem mass spectrometry (MS/MS) fragmentation data (Extended Data Fig. 3a). This effort indicated that the four

metabolites were probably all GPDs, which are formed by the deacylation of phospholipids (Fig. 3a and Supplementary Table 4). We tentatively annotated the compounds as glycerophosphoglycerol (GPG), glycerophosphocholine (GPC), glycerophosphoinositol (GPI) and glycerophosphoethanolamine (GPE), which represent four of the five major GPDs (Fig. 3a). To validate the compounds as GPDs, we showed that commercially available standards for GPC (Fig. 3b,c and Extended Data Fig. 3b) and GPI (Fig. 3d,e and Extended Data Fig. 3c) had retention times and MS/MS spectra matching those in biological samples. For GPE and GPG, owing to the lack of commercial standards, we generated these metabolites in house, and validated that the retention times, *m/z* values and MS/MS spectrum from the biological samples matched those of the standards (Extended Data Fig. 3d–f) and, when available, previously experimentally reported fragments<sup>32</sup>. Although the fifth major GPD, glycerophosphoserine (GPS), was not detected in the original untargeted metabolite profiles, follow-up work showed that its lysosomal levels are also increased after CLN3 loss (Fig. 3g). Its identity was validated by comparing *m/z* values and MS/MS spectrum in our samples with those of *in silico* fragmentation patterns generated using competitive fragmentation modelling-ID (CFM-ID)<sup>33</sup> (Extended Data Fig. 3g). Together, our approach rigorously identified and annotated the metabolites that accumulate in the lysosome after CLN3 loss according to MSI guidelines<sup>34</sup> (Fig. 3f). We also manually extracted ion chromatograms for GPDs, which confirmed large and significant accumulations of the GPDs in brain lysosomes from the Batten disease model, but also showed almost no or only very modest increases in their corresponding levels in the whole-brain samples (Fig. 3g). These data are consistent with lysosomes occupying a small fraction of the total cellular volume and underscore the value of directly measuring metabolites in lysosomes.

To rule out that the increase in GPDs in the brain lysosomes of the Batten disease mice is a consequence of neurodevelopment defects, we generated and analysed HEK293T cells (Extended Data Fig. 3h) that are null for *CLN3* and also express TMEM192–3×HA. In lysosomes from these cells, all five GPDs also accumulated to very high levels and reverted to control amounts or below after expression of the *CLN3* cDNA (Fig. 3h and Extended Data Fig. 3i). Importantly, we found that, regardless of the tagging system used to purify lysosomes, similar levels of GPD accumulation were observed after CLN3 loss (Extended Data Fig. 3j). As in the brain, the HEK293T lysosomes lacking CLN3 also had increases in LPG levels (Fig. 3i).

Human CLN3 can complement the functions of its yeast homologue Btn1p in multiple lysosomal assays<sup>17</sup>, indicating that they share the same function. Indeed, yeast lacking Btn1p significantly accumulated GPC, the only detectable GPD in yeast, suggesting that the function of CLN3 in GPD metabolism is conserved (Fig. 3j). Importantly, we found elevated levels of GPI in the cerebrospinal fluid (CSF) of individuals with CLN3 Batten disease compared to that from individuals without CLN3 disease<sup>35</sup> (Fig. 3k), suggesting its potential use as a clinical biomarker.

Thus, the use of untargeted metabolomics in combination with the LysoIP approach reveals that, in the mouse brain and in human cells in culture, the loss of CLN3 causes a large lysosomal accumulation of GPDs along with a more modest increase in lysophospholipids, which are the metabolites directly upstream of GPDs in the pathway for phospholipid catabolism. Although the specific lysosomal phospholipases that generate GPDs from phospholipids have not been defined, we note that lysosomes do contain the phospholipase A/B activities required for their release from phospholipids<sup>36</sup>.

### CLN3 is required for the efflux of GPDs

The acidic lysosomal lumen establishes a proton gradient with the cytosol that drives the efflux of several metabolites, including GPDs<sup>12</sup> (Extended Data Fig. 4c,d). A plausible explanation for GPD accumulation is that CLN3 loss increases the lysosomal pH, which attenuates

GPD egress from the lysosome. However, consistent with previous reports<sup>37</sup>, CLN3 loss in HEK293T cells does not increase lysosomal pH and, if anything, leads to a small decrease (Extended Data Fig. 4a,b). Consistent with this, after CLN3 loss, we observed no lysosomal accumulation of amino acids (proline, alanine and glutamate) of which the efflux from the lysosome is dependent on pH—functionally confirming our pH measurements (Extended Data Fig. 4d). Thus, the lysosomal accumulation of GPDs in CLN3-deficient lysosomes is independent of perturbations in lysosomal pH.

Given these data and that CLN3 is a multi-pass transmembrane protein, we hypothesized that CLN3 is either a lysosomal glycerophosphodiesterase (GDE) required for hydrolysing GPDs into glycerol 3-phosphate and their corresponding headgroups—an activity that is mostly associated with membrane proteins<sup>38</sup>—or, alternatively, required for the efflux from lysosomes of GPDs produced by the lysosomal catabolism of phospholipids. To test the former, we purified CLN3 and, as a positive control, the established glycerophosphodiesterase GDE1. Although GDE1 readily hydrolysed deuterated GPC *in vitro*, we found no such activity for CLN3 (Extended Data Fig. 4e). These results are consistent with previous reports indicating that GPDs are the final products of phospholipid degradation in mammalian lysosomes<sup>36</sup>.

To test the possibility that CLN3 is required for the lysosomal efflux of GPDs, we used an assay in HEK293T cells in which, through endocytosis, we delivered to lysosomes a phospholipid with a labelled head group that, after lysosomal catabolism, produces a labelled GPD. We monitored the fate of the labelled GPD using mass spectrometry (Fig. 4a). Specifically, we conjugated phosphatidylglycerol with a 5×deuterium-labelled head group (D5–PG) to bovine serum albumin (BSA) and added it to the medium of serum-starved HEK293T cells expressing TMEM192–3×HA (Fig. 4a). The relative amounts of the deuterated phosphatidylglycerol were the same in the whole-cell and lysosomal fractions of control and *CLN3*-null cells, indicating that loss of CLN3 does not affect the uptake of the phospholipid into lysosomes (Fig. 4b). Notably, we could detect only a small signal for deuterated GPG—the GPD released from the labelled phosphatidylglycerol—in the whole-cell and lysosomal samples from the control cells, consistent with GPG being quickly effluxed from normal lysosomes and further catabolized in the cytosol. By contrast, in the *CLN3*-null cells, deuterated GPG readily accumulated in lysosomes and was also easily detected in whole-cell samples (Fig. 4c). These results indicate that, in cells lacking CLN3, the labelled GPG is trapped in lysosomes and is therefore not catabolized in the cytosol (Fig. 4c). Consistent with our previous metabolite profiles (Fig. 3i), in the *CLN3*-null cells, we also observed a significant accumulation of labelled LPGs (Fig. 4d)—the metabolites directly upstream of GPD. Thus, we conclude that CLN3 is required for transporting the GPDs released from phospholipids out of the lysosome.

To further validate this model and to understand the effect of impaired GPD recycling in neurons, we used phosphatidylcholine (PC) with a deuterium-labelled head group (D9–PC) to trace the metabolic fate of its choline moiety in the cell, which is generated by the catabolism of GPC by GDE activity in the cytoplasm (Fig. 4e). CLN3 loss significantly reduced the pools of labelled choline and phosphocholine in primary neuronal cultures (Fig. 4f). As a consequence, we observed a significant reduction in the deuterated isotopomers of multiple choline-containing lipid species, including those of PC and sphingomyelin (Fig. 4g). Importantly, loss of CLN3 did not decrease the uptake of the tracer into neuronal cells (Extended Data Fig. 4f), and did not affect lipid biosynthesis or turnover as indicated by the similar labelling patterns, regardless of CLN3 status, when free D9–choline was used as a tracer (Extended Data Fig. 4g). Together, these data further support a critical role for CLN3 in the egress of GPDs from the lysosome for their catabolism and use by the rest of the cell.

## Discussion

Interestingly, over 60 years ago Christian deDuve discovered that GPDs are terminal products of phospholipid degradation in the lysosome<sup>36</sup>. This finding implies that a mechanism must exist to efflux GPDs from the lysosome so that they can be degraded by GDEs in the cytosol<sup>36,38</sup>. However, no transporter for GPDs has been identified in vertebrates, although a plasma membrane GPD transporter has been reported in yeast<sup>39</sup>. Here we show that CLN3, the loss of which causes a severe neurodegenerative disease in children, is required for the efflux of GPDs from the lysosome. The membrane localization and predicted atypical solute carrier structure<sup>24</sup> of CLN3 suggests that it is a lysosomal GPD exporter or a key component of one, although further research using purified CLN3 will be needed to test this possibility. Moreover, the substantial accumulation of GPDs caused by CLN3 loss might inhibit as yet unidentified enzymes upstream of GPD production to cause the increase that we observed in lysophospholipids, which are known to be neurotoxic<sup>40</sup>. Indeed, GPDs have previously been reported to inhibit lysophospholipase activities *in vitro*<sup>41</sup>. Furthermore, our tracing experiments suggest that CLN3 loss might contribute to neurodegeneration by causing the extralysosomal deprivation of nutrients that are essential for brain cell function and membrane integrity such as choline in addition to the potentially toxic effects of storage material in the lysosome<sup>2</sup>.

Our work provides a framework for future studies into how the loss of CLN3 might affect cellular homeostasis and, more generally, for using the *in vivo* LysoIP method in combination with untargeted metabolite profiling for understanding the function of disease-associated lysosomal proteins.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-05221-y>.

1. Ballabio, A. & Bonifacino, J. S. Lysosomes as dynamic regulators of cell and organismal homeostasis. *Nat. Rev. Mol. Cell Biol.* **21**, 101–118 (2020).
2. Medoh, U. N., Chen, J. Y. & Abu-Remaileh, M. Lessons from metabolic perturbations in lysosomal storage disorders for neurodegeneration. *Curr. Opin. Syst. Biol.* **29**, 100408 (2022).
3. Platt, F. M., d'Azzo, A., Davidson, B. L., Neufeld, E. F. & Tiffet, C. J. Lysosomal storage diseases. *Nat. Rev. Dis. Primers* **4**, 27 (2018).
4. Ferguson, S. M. Neuronal lysosomes. *Neurosci. Lett.* **697**, 1–9 (2019).
5. Perera, R. M. & Zoncu, R. The lysosome as a regulatory hub. *Annu. Rev. Cell Dev. Biol.* **32**, 223–253 (2016).
6. Savini, M., Zhao, Q. & Wang, M. C. Lysosomes: signaling hubs for metabolic sensing and longevity. *Trends Cell Biol.* **29**, 876–887 (2019).
7. Ballabio, A. & Gieselmann, V. Lysosomal disorders: from storage to cellular damage. *Biochim. Biophys. Acta* **1793**, 684–696 (2009).
8. Boustany, R. M. Lysosomal storage diseases—the horizon expands. *Nat. Rev. Neurol.* **9**, 583–598 (2013).
9. Marques, A. R. A. & Saftig, P. Lysosomal storage disorders—challenges, concepts and avenues for therapy: beyond rare diseases. *J. Cell Sci.* **132**, jcs221739 (2019).
10. Wallings, R. L., Humble, S. W., Ward, M. E. & Wade-Martins, R. Lysosomal dysfunction at the centre of Parkinson's disease and frontotemporal dementia/amyotrophic lateral sclerosis. *Trends Neurosci.* **42**, 899–912 (2019).
11. Wang, C., Telpoukhovskaia, M. A., Bahr, B. A., Chen, X. & Gan, L. Endo-lysosomal dysfunction: a converging mechanism in neurodegenerative diseases. *Curr. Opin. Neurobiol.* **48**, 52–58 (2018).
12. Abu-Remaileh, M. et al. Lysosomal metabolomics reveals V-ATPase- and mTOR-dependent regulation of amino acid efflux from lysosomes. *Science* **358**, 807–813 (2017).
13. Thai, T. H. et al. Regulation of the germinal center response by microRNA-155. *Science* **316**, 604–608 (2007).
14. Pisoni, R. L., Acker, T. L., Lisowski, K. M., Lemons, R. M. & Thoene, J. G. A cysteine-specific lysosomal transport system provides a major route for the delivery of thiol to human

fibroblast lysosomes: possible role in supporting lysosomal proteolysis. *J. Cell Biol.* **110**, 327–335 (1990).

15. Eiberg, H., Gardiner, R. M. & Mohr, J. Batten disease (Spielmeyer-Sjogren disease) and haptoglobins (HP): indication of linkage and assignment to chr. 16. *Clin. Genet.* **36**, 217–218 (1989).
16. Lerner, T. J. et al. Isolation of a novel gene underlying Batten disease, *CLN3*. *Cell* **82**, 949–957 (1995).
17. Mirza, M. et al. The *CLN3* gene and protein: what we know. *Mol. Genet. Genomic Med.* **7**, e859 (2019).
18. Butz, E. S., Chandrachud, U., Mole, S. E. & Cotman, S. L. Moving towards a new era of genomics in the neuronal ceroid lipofuscinoses. *Biochim. Biophys. Acta Mol. Basis Dis.* **1866**, 165571 (2019).
19. Jarvela, I. et al. Biosynthesis and intracellular targeting of the *CLN3* protein defective in Batten disease. *Hum. Mol. Genet.* **7**, 85–90 (1998).
20. Storch, S., Pohl, S. & Braulke, T. A dileucine motif and a cluster of acidic amino acids in the second cytoplasmic domain of the batten disease-related *CLN3* protein are required for efficient lysosomal targeting. *J. Biol. Chem.* **279**, 53625–53634 (2004).
21. Mao, Q., Foster, B. J., Xia, H. & Davidson, B. L. Membrane topology of *CLN3*, the protein underlying Batten disease. *FEBS Lett.* **541**, 40–46 (2003).
22. Ezaki, J. et al. Characterization of *Cln3p*, the gene product responsible for juvenile neuronal ceroid lipofuscinosis, as a lysosomal integral membrane glycoprotein. *J. Neurochem.* **87**, 1296–1308 (2003).
23. Oetjen, S., Kuhl, D. & Hermey, G. Revisiting the neuronal localization and trafficking of *CLN3* in juvenile neuronal ceroid lipofuscinosis. *J. Neurochem.* **139**, 456–470 (2016).
24. Perland, E., Bagchi, S., Klaesson, A. & Fredriksson, R. Characteristics of 29 novel atypical solute carriers of major facilitator superfamily type: evolutionary conservation, predicted structure and neuronal co-expression. *Open Biol.* **7**, 170142 (2017).
25. Mitchison, H. M. et al. Targeted disruption of the *Cln3* gene provides a mouse model for Batten disease. The Batten Mouse Model Consortium [corrected]. *Neurobiol. Dis.* **6**, 321–334 (1999).
26. Kovacs, A. D. & Pearce, D. A. Finding the most appropriate mouse model of juvenile *CLN3* (Batten) disease for therapeutic studies: the importance of genetic background and gender. *Dis. Model. Mech.* **8**, 351–361 (2015).
27. Lojewski, X. et al. Human iPSC models of neuronal ceroid lipofuscinosis capture distinct effects of TPPI and *CLN3* mutations on the endocytic pathway. *Hum. Mol. Genet.* **23**, 2005–2022 (2014).
28. Platt, F. M. Sphingolipid lysosomal storage disorders. *Nature* **510**, 68–75 (2014).
29. Fuller, M. & Futerman, A. H. The brain lipidome in neurodegenerative lysosomal storage disorders. *Biochem. Biophys. Res. Commun.* **504**, 623–628 (2018).
30. Hobert, J. A. & Dawson, G. A novel role of the Batten disease gene *CLN3*: association with BMP synthesis. *Biochem. Biophys. Res. Commun.* **358**, 111–116 (2007).
31. Padilla-Lopez, S., Langager, D., Chan, C. H. & Pearce, D. A. *BTNI*, the *Saccharomyces cerevisiae* homolog to the human Batten disease gene, is involved in phospholipid distribution. *Dis. Model. Mech.* **5**, 191–199 (2012).
32. Kopp, F. et al. The glycerophospho metabolome and its influence on amino acid homeostasis revealed by brain metabolomics of *GDE1(-/-)* mice. *Chem. Biol.* **17**, 831–840 (2010).
33. Allen, F., Pon, A., Wilson, M., Greiner, R. & Wishart, D. CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res.* **42**, W94–W99 (2014).
34. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis. Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211–221 (2007).
35. Dang Do, A. N. et al. Neurofilament light chain levels correlate with clinical measures in *CLN3* disease. *Genet. Med.* **23**, 751–757 (2021).
36. Fowler, S. & De Duve, C. Digestive activity of lysosomes. 3. The digestion of lipids by extracts of rat liver lysosomes. *J. Biol. Chem.* **244**, 471–481 (1969).
37. Schmidtko, C. et al. Lysosomal proteome analysis reveals that *CLN3*-defective cells have multiple enzyme deficiencies associated with changes in intracellular trafficking. *J. Biol. Chem.* **294**, 9592–9604 (2019).
38. Corda, D. et al. The emerging physiological roles of the glycerophosphodiesterase family. *FEBS J.* **281**, 998–1016 (2014).
39. Patton-Vogt, J. Transport and metabolism of glycerophosphodiesterases produced through phospholipid deacylation. *Biochim. Biophys. Acta* **1771**, 337–342 (2007).
40. Rigoni, M. et al. Equivalent effects of snake PLA2 neurotoxins and lysophospholipid-fatty acid mixtures. *Science* **310**, 1678–1680 (2005).
41. Fallbrook, A., Turenne, S. D., Mamas, N., Kish, S. J. & Ross, B. M. Phosphatidylcholine and phosphatidylethanolamine metabolites may regulate brain phospholipid catabolism via inhibition of lysophospholipase activity. *Brain Res.* **834**, 207–210 (1999).
42. Storey, J. D. A direct approach to false discovery rates. *J. R. Stat. Soc. B* **64**, 479–498 (2002).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

## Methods

### Mouse studies

Mice were maintained on a standard light–dark cycle with access to food and water ad libitum in a room with a controlled temperature (22 °C) and humidity (around 50%). Cages were cleaned every 4–5 days and supplies of water and food were checked daily. For the generation of *Rosa26;lox-stop-lox-TMEM192-3×HA*, the *TMEM192* coding sequence containing a C-terminal 3×HA epitope tag was inserted into the CTV vector (Addgene plasmid, 15912)<sup>13</sup>, linearized with AsiSI and electroporated into embryonic stem cells of the 129 background. Embryonic stem cell colonies were picked, and successfully targeted clones were identified by long-range PCR amplifying the regions spanning each homology arm. Positive embryonic stem cell clones with proper integration into the *Rosa26* locus were then injected into blastocysts and transferred into pseudopregnant female mice to obtain chimeric mice. Chimeras were bred to C57BL/6J mice for at least three generations before maintaining the line in a homozygous form. To generate a constitutive LysoTag expressing line, *Rosa26;lox-stop-lox-TMEM192-3×HA* mice were crossed with CMV-Cre transgenic mice of the C57BL/6J background, excision of *lox-stop-lox* was validated using our genotyping PCR method (details are provided below), and positive progeny were bred to C57BL/6J mice for at least three generations. The conditional LysoTag mouse was deposited at the Jackson Laboratory (strain, 035401) with the common name: conditional LysoTag (*Rosa26;lox-stop-lox-TMEM192-3×HA*).

The Batten disease mouse model (*Cln3<sup>-/-</sup>*) was purchased from the Jackson Laboratory (029471, B6.129S6-Cln3<sup>tm1Nbn/J</sup>) and maintained in the homozygous state. For LysoIP experiments, *Cln3<sup>-/-</sup>* mice were bred with LysoTag mice to generate *LysoTag;Cln3<sup>+/-</sup>* (used as controls) and the *LysoTag;Cln3<sup>-/-</sup>* experimental group. Statistical methods were not used to predetermine the sample size. Sample blinding was not performed for most of the experiments as they did not involve qualitative scoring (see the Reporting summary for details).

All of the procedures involving mice were carried out in accordance with the approved guidelines by the Institutional Animal Care and Use Committee at MIT, the Whitehead Institute for Biomedical Research and Stanford University.

### Mouse genotyping

Genomic DNA was isolated from ear biopsies using QuickExtract DNA Extraction Solution (Lucigen) and PCR was performed using GoTaq DNA polymerase (Promega). To check for the presence of the *lox-stop-lox-TMEM192-3×HA* cassette in the *Rosa26* locus, two pairs of primers were designed. The non-targeted allele is amplified by primer pair P1, which produces a PCR product of 588 bp; by contrast, if the *lox-stop-lox-TMEM192-3×HA* cassette is present, it is amplified by primer pair P2, which gives a 215 bp product. The primer sequences are as follows: P1F, 5'-CAGTAAGGGAGCTGCAGTGG-3'; P1R, 5'-GCAGAAGGAGCGGGAGAAAT-3'; P2F, 5'-TACCTGAAGCGACACAATGC-3' (within *TMEM192* sequence); P2R, 5'-CCGCTCCCTCCAGCATAATC-3' (within HA sequence). The PCR reaction was performed with four primers in multiplex using the following conditions: 95 °C for 3 min followed by 35 cycles of 95 °C for 30 s, 56 °C for 30 s and 72 °C for 2 min, and a final extension step was added for 72 °C for 5 min. To test for the excision of the *lox-stop-lox* fragment by the Cre recombinase, two forward primers were designed; one within the *loxp* stop cassette and another upstream the first *loxp* site, both having a common reverse primer within the *TMEM192* sequence. The non-excised allele generates a 331 bp product, whereas successful excision of the *loxp-stop-loxp* site gives a 193 bp product. The primer sequences were as follows: P3F1, 5'-GGGCAACGTGCTGGTTATTG-3'; P3F2, 5'-CCTCCCCCTGAACCTGAAAC-3'; P3R, 5'-GAAGCTGGGCATC CAGAAGT-3'. The PCR reaction was performed with three primers using the following conditions: 95 °C for 3 min followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1:20 min, and a final extension step was added for 72 °C for 5 min.

### Antibodies, chemicals and plasmids

Detailed information on the antibodies, chemicals and plasmids used in this study is provided in Supplementary Table 6. Additional reagents included the following secondary antibodies for immunofluorescence (used at 1:1,000): goat anti-rat Alexa Fluor 488 (Thermo Fisher Scientific, A-11006), goat anti-rabbit Alexa Fluor 488 (Thermo Fisher Scientific, A-21206), goat anti-mouse Alexa Fluor 647 (Thermo Fisher Scientific, A-21240), goat anti-mouse Alexa Fluor 488 (Thermo Fisher Scientific, A-11001), and goat anti-rabbit Alexa Fluor 594 (Thermo Fisher Scientific, A32740); the following secondary antibodies for immunoblotting (used at 1:3,000): anti-rabbit HRP (Cell Signaling Technology (CST), 7074), anti-mouse HRP (CST, 7076), anti-rat HRP (CST, 7077), and normal donkey serum, which were purchased from CST; Inactivated fetal calf serum (IFS) from Invitrogen; DMEM high glucose, anti-HA magnetic beads and ECL2 western blotting substrate from Thermo Fisher Scientific; XtremeGene9 DNA Transfection Reagent, Complete Protease Inhibitor Cocktail and PhosSTOP phosphatase inhibitor from Roche; and QuickExtract DNA Extraction Solution from Lucigen.

### Rapid isolation of lysosomes from brain and liver tissues (LysoIP)

Immediately after euthanasia, mouse brains were collected and cerebral hemispheres dissected on an ice-cold plastic dish. For liver samples, we obtained a small round piece using a biopsy punch (4 mm diameter), which helped to reduce variability between samples and minimize processing time prior to the LysoIP. The liver tissues and cerebral hemispheres were used immediately without freezing following similar steps as described in the original LysoIP protocol<sup>12</sup>.

**Protein analyses.** Liver biopsy or half cerebral hemisphere samples (4 mm) were gently homogenized in 1 ml ice-cold PBS supplemented with protease and phosphatase inhibitor cocktails (Roche) using a 2 ml Dounce homogenizer (Douncer). Both liver and brain samples required 25 strokes for optimal homogenization. Homogenates were collected and 25 µl was saved on ice as the input fraction, while the remaining was centrifuged at 1,000g for 2 min at 4 °C. The supernatant was then incubated with anti-HA beads that were prewashed three times with the same buffer, incubated for 20 min on a gentle rotator shaker at 4 °C and then washed three times in the same cold buffer. Washes were performed by pipetting the beads three times in each wash, then discarding the supernatant by placing the tube on a magnet to trap the beads. To increase the purity of the lysosomal preparation, we changed the tube during the last wash and always used protein low binding tubes. The input and lysosomal fractions were resuspended in ice-cold Triton lysis buffer (50 mM HEPES pH 7.4, 40 mM NaCl, 1% Triton X-100, 10 mM β-glycerol phosphate, 10 mM pyrophosphate, 30 mM sodium fluoride, 1.5 mM Sodium orthovanadate, 2 mM EDTA and protease and phosphatase inhibitor cocktails). The samples were incubated for 10 min at 4 °C and centrifuged at maximum speed in cold. The supernatant was collected and saved for analysis. The lysosomal fraction was not centrifuged but rather incubated on the magnet and transferred to new tubes twice. For immunoblots, an equivalent of 0.08% and 15% of the whole tissue input and captured lysosomes were run, respectively.

**Polar metabolite and lipid analyses.** The procedure was similar to the steps above, but processing was performed in ice-cold KPBS (136 mM KCl, 10 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.25 in Optima LC–MS water) with no protease or phosphatase inhibitors. The homogenates were incubated with prewashed anti-HA beads for 5 min for polar metabolites and 20 min for lipid experiments. We noticed that incubation for 3–15 min still returns similar results for polar metabolites, but incubation for 5 min was used for convenience. For polar metabolites, the input fractions were resuspended in 225 µl, while the lysosomal fractions were resuspended in 50 µl of ice-cold 80% methanol in LC–MS water containing



500 nM isotope-labelled amino acids used as internal standards (Cambridge Isotope Laboratories). The samples were stored at  $-80^{\circ}\text{C}$  until analysed. On the day of analysis, the samples were vortexed for 10 min at  $4^{\circ}\text{C}$  and centrifuged at  $17,000g$ . For lipid extraction,  $800\ \mu\text{l}$  of chloroform:methanol at ratio of 2:1 (v/v) containing  $750\ \text{ng}\ \text{ml}^{-1}$  of SPLASH LipidoMIX internal standard mix (Avanti) was added to the sample, followed by the addition of  $160\ \mu\text{l}$  of 0.9% (w/v) NaCl and vortexing for 10 min at  $4^{\circ}\text{C}$ . The mixture was centrifuged at  $3,000g$  for 15 min at  $4^{\circ}\text{C}$ . The lower phase containing lipids was collected and dried using a SpeedVac. At this stage, the samples were stored at  $-80^{\circ}\text{C}$ . On the day of analysis, dried lipid extracts were reconstituted in  $50\ \mu\text{l}$  of acetonitrile (ACN):isopropyl alcohol: $\text{H}_2\text{O}$  at a ratio of 13:6:1 (v/v/v) and vortexed for 10 min at  $4^{\circ}\text{C}$ . The samples were centrifuged for 10 min at  $4^{\circ}\text{C}$  and transferred to glass insert vials for LC-MS.

### Proteomics

Eluates from LysoIPs and aliquots of the matched starting lysates (corresponding to approximately  $50\ \mu\text{g}$  of protein extract) were processed as described previously<sup>43</sup>. In brief, proteins were solubilized by addition of SDS to a final concentration of 2% (w/v), followed by sonication using the Bioruptor Plus (Diagenode) and heating for 10 min at  $95^{\circ}\text{C}$ . After reduction and alkylation, proteins were precipitated by cold-acetone precipitation. The resulting pellets were resuspended in digestion buffer (3 M urea in 100 mM HEPES pH 8.0) and digested by consecutive addition of LysC (Wako, 3 h at  $37^{\circ}\text{C}$ ) and trypsin (Promega, 16 h at  $37^{\circ}\text{C}$ ). The obtained digested peptides were acidified and desalted using a Waters Oasis HLB  $\mu\text{Elution}$  Plate  $30\ \mu\text{m}$  (Waters) according to the manufacturer's instructions. The desalted peptides were dissolved in 5% (v/v) ACN, 0.1% (v/v) formic acid to a peptide concentration of approximately  $1\ \mu\text{g}\ \mu\text{l}^{-1}$  and spiked with iRT peptides (Biognosys) before analysis using LC-MS/MS. Approximately  $1\ \mu\text{g}$  of reconstituted peptides were analysed by data-independent acquisition (DIA) using the UltiMate 3000 UPLC system (Thermo Fisher Scientific) fitted with a trapping (Waters nanoEase M/Z Symmetry C18,  $5\ \mu\text{m}$ ,  $180\ \mu\text{m} \times 20\ \text{mm}$ ) and an analytical column (Waters nanoEase M/Z Peptide C18,  $1.7\ \mu\text{m}$ ,  $75\ \mu\text{m} \times 250\ \text{mm}$ ). The outlet of the analytical column was coupled directly to a QExactive HF (Thermo Fisher Scientific) system using the Proxeon nanospray source. Solvent A was water, 0.1% (v/v) formic acid and solvent B was 80% (v/v) ACN, 0.08% (v/v) formic acid. Peptides were eluted using a non-linear gradient from 1% to 62.5% B in 131 min. The total runtime was 150 min, including clean-up and column re-equilibration. The S-lens RF value was set to 60. MS acquisition parameters were set as follows: full scan MS spectra with a mass range of  $350\text{--}1,650\ m/z$  were acquired in profile mode in the Orbitrap with resolution of 120,000 FWHM. The filling time was set at a maximum of 60 ms with an AGC target of  $3 \times 10^6$  ions. DIA scans were acquired with 40 mass window segments of differing widths across the MS1 mass range. The default charge state was set to  $3^+$ . Higher energy collisional dissociation (HCD) fragmentation (stepped normalized collision energy; 25.5, 27, 30%) was applied and MS/MS spectra were acquired with a resolution of 30,000 FWHM with a fixed first mass of  $200\ m/z$  after accumulation of  $3 \times 10^6$  ions or after a filling time of 35 ms (whichever occurred first). Data were acquired in profile mode. For data acquisition and processing, we used Tune v.2.9 and XCalibur v.4.1.

Acquired data were processed using Spectronaut Professional v.13.10 (Biognosys). Raw files were searched by directDIA search with Pulsar (Biognosys) against the mouse UniProt database (*Mus musculus*, entry only, release 2016\_01) with a list of common contaminants appended, using the default settings. For quantification, the default BGS factory settings were used, except for the following settings: proteotypicity filter = only protein group specific; major group quantity = median peptide quantity; major group top N = OFF; minor group quantity = median precursor quantity; minor group top N = OFF; data filtering = Q value percentile with fraction = 0.2 and imputing strategy = global imputing; normalization strategy = local normalization; row selection = automatic.

The candidates and protein report tables were exported from Spectronaut and were used for volcano plot generation and PCA, respectively, using R v.3.4.1 and RStudio server v.1.1.463. Protein groups were considered to be significantly enriched in LysoIP samples if they displayed a Q value of  $<0.01$  and an average  $\log_2$  ratio of  $>1.5$ . Known lysosomal proteins were defined based on Gene Ontology cellular component or UniProt subcellular localization annotation. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>44</sup> partner repository with the dataset identifier PXD018624.

### Polar metabolomics

The initial metabolite profiling was conducted on the QExactive benchtop orbitrap mass spectrometer equipped with an Ion Max source and a heated electrospray ionization (HESI) II probe, which was coupled to a Dionex UltiMate 3000 HPLC system (Thermo Fisher Scientific). To prevent any systematic bias from possible column carryover, sample order was randomized in LC-MS runs. External mass calibration was performed using the standard calibration mixture every 7 days. Typically,  $5\ \mu\text{l}$  was injected onto a SeQuant ZIC-pHILIC  $150\ \text{mm} \times 2.1\ \text{mm}$  analytical column equipped with a  $2.1\ \text{mm} \times 20\ \text{mm}$  guard column (both  $5\ \text{mm}$  particle size; EMD Millipore). Buffer A was 20 mM ammonium carbonate, 0.1% ammonium hydroxide; buffer B was ACN. The column oven and autosampler tray were held at  $25^{\circ}\text{C}$  and  $4^{\circ}\text{C}$ , respectively. The chromatographic gradient was run at a flow rate of  $0.150\ \text{ml}\ \text{min}^{-1}$  as follows: 0–20 min, linear gradient from 80–20% B; 20–20.5 min, linear gradient from 20–80% B; 20.5–28 min, hold at 80% B. The mass spectrometer was operated in full-scan, polarity-switching mode, with the spray voltage set to 3.0 kV, the heated capillary was held at  $275^{\circ}\text{C}$ , and the HESI probe held at  $350^{\circ}\text{C}$ . The sheath gas flow was set to 40 units, the auxiliary gas flow was set to 15 units, and the sweep gas flow was set to 1 unit. MS data acquisition was performed in a range of  $m/z = 70\text{--}1,000$ , with the resolution set at 70,000, the AGC target at  $1 \times 10^6$  and the maximum injection time at 20 ms. To improve sensitivity of detection of LysoTracker DND-99 and deuterated GPG (D5-GPG) in lipid-uptake experiments, additional targeted selected ion monitoring scans were included in positive mode, with  $m/z$  centred on 400.21150 and 252.08920, respectively. For these scans, the resolution was set to 70,000, the AGC targeted was  $1 \times 10^5$ , the max IT was 250 ms and the isolation window around each target mass was set to  $2.0\ m/z$ .

Relative quantification of polar metabolites was performed by extracting ion chromatograms with XCalibur QuanBrowser v.4.1 (Thermo Fisher Scientific) using a 5 ppm mass tolerance and referencing an in-house library of chemical standards. Metabolites and runs were subject to predefined quality-control parameters: CV (s.d./mean peak area across multiple injections of a representative (pooled) biological sample) below 0.25; and  $R^2$  (linear correlation across a three-point dilution series of the representative (pooled) biological sample) greater than 0.90.

For a metabolite to be lysosomal in the targeted analysis performed in Fig. 1, its abundance in the immunoprecipitates from LysoTag<sup>+</sup> mice had to be twice higher than that in a control immunoprecipitate in at least 3 out of the 4 samples. Furthermore, all of the fold change calculations in the abundance of lysosomal metabolites, where targeted analyses are presented, were determined after subtracting the abundance of the metabolites in the control immunoprecipitates. For polar targeted metabolites from cultured cells, we also used LysoTracker detected with LC-MS to account for any gross changes in lysosomal mass or capturing efficiency by normalizing metabolite intensities to the total captured LysoTracker divided by the total cellular LysoTracker signal (capturing efficiency). Whole-cell abundances were normalized to total cell volume.

Note that some metabolite profiling data were acquired on an ID-X tribrid mass spectrometer (Thermo Fisher Scientific) at Stanford University. The same HILIC method and column parts were used for a Vanquish Horizon UPLC system. The different mass spectrometer parameters were as follows: Orbitrap resolution, 120,000; maximum

injection time, 80 ms. Note that, due to the change of instrumentation, there is an approximately 1 min retention time shift for those polar metabolite data collected at Stanford University, hence the difference in the retention time of GPE and GPG standards from those reported in the experimental samples in Supplementary Table 4. Experimental samples were re-run and metabolite identities were validated using the new instrumentation alongside our standards.

### Untargeted polar metabolomics

Data were acquired as described above, with additional data-dependent MS/MS (ddMS/MS) collected on pooled samples to help with unknown metabolite identification. For ddMS/MS, the top 10 ions in each full scan were isolated with a 1.0 Da window, fragmented with a step-wise collision energy of 15, 30 and 45 units, and analysed at a resolution of 17,500 with an AGC target of  $2 \times 10^5$  and a maximum injection time of 100 ms. The underfill ratio was set to 0. The selection of the top 10 ions was set to isotopic exclusion, a dynamic exclusion window of 5.0 s and an exclusion list of background ions based on a solvent blank. Data were analysed using Compound Discoverer v.3.1 (Thermo Fisher Scientific) and by including a custom mass list. Normalization was performed using the default total ion count (TIC) parameter. Identification and annotation of metabolite species was confirmed by matching retention times and MS/MS data to authentic standards. In cases in which there was no authentic standard available, MS/MS data were matched first to published experimental MS/MS data (that is, through the Thermo Fisher Scientific mzCloud spectral library), or to predicted MS/MS data produced using CFM-ID<sup>33</sup>. Details of each metabolite annotated/identified in this study are provided in Fig. 3 and Supplementary Table 4.

**Compound discoverer workflow.** Features were extracted in the  $m/z = 70$ –1,000 mass range, from 0.1–16 min of the PHILIC chromatography runs. The signal to noise threshold for feature detection was set to 1.5. Compounds were detected using the following parameter settings: mass tolerance was set to 10 ppm, retention time tolerance was set to 0.4 min, minimum peak intensity was set to 100,000 AU and the signal to noise threshold for compound detection was set to 3. Isotopes and adducts were grouped using the default ions and base ion lists. Missing peaks were gap-filled and mock extraction blank samples were used to determine the background compounds for subsequent subtraction. Areas were normalized to constant median excluding the blank samples.

To predict elemental compositions of the compounds, the relative intensity tolerance was set to 30% for isotope pattern matching. For the mzCloud database search, both the precursor and fragment mass tolerance were set to 10 ppm. The databases used for matching compounds in ChemSpider for structural search were BioCyc, the Human Metabolome Database and KEGG, and the mass tolerance in ChemSpider Search was set to 5 ppm. The mass tolerance for matching compounds in Metabolika pathways was set to 5 ppm. Compounds were also matched to our in-house mass list, with mass tolerance set to 5 ppm and retention time tolerance set to 1 min. mzLogic was applied to score structure candidates based on MS/MS, with fragment mass tolerance set to 10 ppm and match factor threshold set to 30. Compounds were assigned by comparing annotations using the following nodes in order of priority: (1) mass list search; (2) mzCloud search; (3) predict compositions; (4) ChemSpider search; (5) Metabolika search.

The differential analysis was performed on  $\log_{10}$ -transformed values using defined sample ratios (*Cln3*<sup>-/-</sup> versus *Cln3*<sup>+/-</sup>, both in LysoIP fraction and whole cell fraction), and *P* values were calculated. A results table of 1,073 detected compounds was generated from this workflow (Supplementary Table 4) after (1) background compound subtraction and (2) applying a filter of maximum area intensity greater than 500,000 AU to filter out low signal peaks.

The 1,073 detected compounds include possible artifacts as well as redundant or mis-annotated compounds. To focus on unique compounds with good peak quality and sufficient information

for validation studies, we performed additional filtering, which included the following: (1) at least 1 compound match in the ChemSpider search; (2) at least 1 compound match in the mzCloud search. This resulted in 237 compounds. We then manually removed unreliable compounds: those with poorly aligned chromatograms and/or those that were obviously redundant. The final filtered compound list contained 189 compounds with assigned chemical formula and fragmentation results to help with compound identification (Supplementary Table 4). Among the 189 compounds after filtering, 3.2% had at least 1  $\log_2$  fold difference between the *Cln3*<sup>+/-</sup> versus *Cln3*<sup>-/-</sup> LysoIP samples, with adjusted *P* values less than 0.05. The top compounds are structurally related and can be grouped as a class of compounds termed GPDs. PCA in Fig. 2d was performed using PC1 and PC3, because PC3 includes the metabolites that separate the lysosomes of different genotypes. Whole tissues were not separated regardless of the PC used. *P* values were calculated by running the Tukey HSD test (post hoc) after ANOVA (for two group comparison, the ANOVA was simplified to unpaired *t*-tests). The *P* values were corrected using the Benjamini–Hochberg method for the false-discovery rate. These analyses were performed using the differential analysis node in compound discoverer v.3.1.

For targeted analyses, Thermo Fisher Scientific TraceFinder software v.5.0 was used to help in processing targeted metabolites screening and obtaining quantitative data from LC–MS.

In the final result table with 189 detected compounds, compound identification was performed based on accurate mass, retention time, and MS/MS spectra comparisons. We assigned to each compound one of four different levels of identification/annotation according to guidelines issued by the MSI<sup>45</sup>. Level 1 defines compounds identified with both experimental retention time and MS/MS fragmentation patterned matched to authentic chemical standards that were run under the same chromatographic conditions. Level 2 defined putatively annotated compounds, with only one of the orthogonal properties (such as retention time) matched with an authentic chemical standard, or an MS2 match to only the online in silico spectra library or the Thermo Fisher Scientific mzCloud database. Level 3 defined putatively characterized compound classes. Finally, level 4 defined unknown compounds.

### Lipidomics

Lipids were separated on an Ascentis Express C18 2.1 mm  $\times$  150 mm 2.7  $\mu$ m column (Sigma-Aldrich) connected to a Vanquish Horizon UPLC system and an ID-X tribrid mass spectrometer (Thermo Fisher Scientific) equipped with a HESI probe. External mass calibration was performed using the standard calibration mixture every 7 days. Between 2  $\mu$ l and 5  $\mu$ l of sample were injected onto the column, with separate injections for positive and negative ionization modes. Mobile phase A in the chromatographic method consisted of 60:40 water:ACN with 10 mM ammonium formate and 0.1% formic acid, and mobile phase B consisted of 90:10 isopropanol:ACN, with 10 mM ammonium formate and 0.1% formic acid. The chromatographic gradient was adapted from ref. <sup>46</sup>. In brief, the elution was performed with a gradient of 40 min; during 0–1.5 min isocratic elution with 32% B; from 1.5 to 4 min, increase to 45% B; from 4 to 5 min, increase to 52% B; from 5 to 8 min, increase to 58% B; from 8 to 11 min, increase to 66% B; from 11 to 14 min, increase to 70% B; from 14 to 18 min, increase to 75% B; from 18 to 21 min, increase to 97% B; during 21 to 35 min, 97% B is maintained; from 35–35.1 min, solvent B was decreased to 32% and then maintained for another 4.9 min for column re-equilibration. The flow rate was set to 0.260 ml min<sup>-1</sup>. The column oven and autosampler were held at 55 °C and 15 °C, respectively. The mass spectrometer parameters were as follows: the spray voltage was set to 3.25 kV in positive mode and 3.0 kV in negative mode, and the heated capillary and the HESI were held at 300 °C and 375 °C, respectively. The S-lens RF level was set to 45, and the sheath and auxiliary gas were set to 40 and 10 units, respectively. These conditions were held constant for both positive- and negative-ionization mode acquisitions.

The mass spectrometer was operated in full-scan ddMS/MS mode with an orbitrap resolution of 120,000 (MS1) and 30,000 (MS/MS). Internal calibration using Easy IC was enabled. Quadrupole isolation was enabled, the AGC target was  $1 \times 10^5$ , the maximum injection time was 50 ms and the scan range was  $m/z = 200-2,000$ . For ddMS/MS, the cycle time was 1.5 s, the isolation window was 1 and an intensity threshold of  $1 \times 10^3$  was used. HCD fragmentation was achieved using a step-wise collision energy of 15, 25 and 35 units, and detected in the orbitrap with an AGC target of  $5 \times 10^4$  and a maximum injection time of 54 ms. Isotopic exclusion was on, a dynamic exclusion window of 2.5 s was used and a top 100 exclusion list was generated using a solvent bank.

High-throughput annotation and relative quantification of lipids was performed using LipidSearch v.4.2.21 (Thermo Fisher Scientific/Mitsui Knowledge Industries)<sup>47,48</sup> using the HCD and 'Labelled' databases (the latter was searched to detect the Avanti Splash LipidoMix that was used as an internal standard). Precursor ion tolerance was set to 5 ppm and product ion tolerance was set to 8 ppm. LipidSearch nomenclature uses underscores to separate the fatty acyl chains to indicate the lack of *sn* positional information (for example, PC(16:0\_18:1)) and not (PC(16:0/18:1)). Raw peak areas for all annotated lipids were exported to Microsoft Excel and filtered according to the following predetermined quality control criteria: Rej (reject parameter calculated by LipidSearch) equal to 0; PQ (peak quality parameter calculated by LipidSearch software) greater than 0.75; CV (s.d./mean peak area across triplicate injections of a represented (pooled) biological sample) below 0.4. Raw peak areas of the filtered lipids were normalized to well-detected deuterated standards found in the Avanti Splash LipidoMix as a control for lipid extraction and efficiency, as well as sample loading. Individual lipid species shown in figures were validated by manually checking the peak alignment and matching the MS/MS spectra to the characteristic fragmentation patterns found in the LIPID-MAPS database ([www.lipidmaps.org](http://www.lipidmaps.org)). Notably, BMP lipid class is not included in the Lipidsearch target database—the PG lipid species from Lipidsearch analysis might be misannotated because BMP and PG are structural isomers. To distinguish between PG and BMP lipid species, annotation for those two lipid classes was performed manually based on the retention time and positive mode MS2 fragmentation. The lipid standards (18:1 ( $\Delta^9$ -Cis) PG and 18:1 BMP (S,R)) were purchased from Avanti Polar Lipids and found to be separable under our lipid chromatographic conditions. The BMP standard elutes before its PG isomer, either when the pure standards were prepared in isopropanol:methanol (1:1) solvent or when spiked into a bovine liver lipid extract. Although BMP and PG have better sensitivity in negative-ionization mode, their positive ammonium adducts were useful precursors for generating distinct neutral loss fragments critical for annotation. Consistent with a previous report<sup>49</sup>, the PG standard generated loss of the polar head group leading to a diglyceride-like fragment. By contrast, the BMP standard generated loss of the monoacylglycerolphosphate moiety leading to monoglyceride-like fragments. When MS2 fragmentation was not acquired, the species was annotated as BMP/PG.

### Protein extractions from mouse tissues

Immediately after euthanasia, tissues were collected and snap-frozen in liquid nitrogen. Before protein extraction, tissues were powdered in liquid nitrogen using a mortar and a pestle. RIPA lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease and phosphatase inhibitors was added to 25 mg of tissue powder and the samples were incubated on shaker for 1 h at 4 °C. Lysates were centrifuged for 30 min at maximum speed and the protein concentrations were determined using the Pierce BCA Protein Assay kit (Thermo Fisher Scientific).

### Immunoblotting

Lysates were resolved by SDS-PAGE (Thermo Fisher Scientific) at 120 V. Resolved proteins were transferred for 2 h at 40 V to PVDF

membranes. Membranes were blocked with 5% non-fat dry milk in TBST (Tris-buffered saline with Tween-20) for 1 h, then incubated overnight with primary antibodies in 5% BSA in TBST at 4 °C. A list of all primary antibody dilutions is provided in Supplementary Table 6. After incubation, membranes were washed three times with TBST for 5 min per wash and then incubated with the appropriate secondary antibodies diluted 1:3,000 in 5% BSA for 1 h at room temperature. Membranes were then washed three times with TBST and visualized using ECL2 western blotting substrate (Thermo Fisher Scientific).

### Electron microscopy

For electron microscopy studies, mice were perfused with 2% paraformaldehyde and 2.5% glutaraldehyde. Brains and livers were then isolated and postfixed in the same fixing solution. Coronal 200  $\mu$ m sections were cut using the Leica VT-1000 (Leica) vibratome. Sections were post-fixed with 2% osmium, rinsed, dehydrated and embedded in Durcupan resin (Sigma-Aldrich). Ultra-thin sections (0.07–0.09  $\mu$ m) were cut with a diamond knife with an ultramicrotome (Leica EM UC6, Leica), stained with lead citrate (Reynolds solution) and examined with a transmission electron microscope (FEI Tecnai G2 Spirit BioTwin, Thermo Fisher Scientific) using a digital camera (Morada, Olympus Soft Image Solutions).

### Cathepsin B activity assay

Cathepsin B activity was assayed fluorometrically using 5  $\mu$ M of the fluorogenic CTSB substrate, Z-RR-AMC (Sigma) in 100  $\mu$ l of the cathepsin B activation buffer (KH<sub>2</sub>PO<sub>4</sub> (87.7 mM), Na<sub>2</sub>HPO<sub>4</sub> (12.3 mM), EDTA (4 mM), DTT (2.6 mM) and cysteine (8 mM) in pH 6.0). The plate was incubated in the Tecan Spark multimode microplate reader (Tecan Trading) at 37 °C and read at an excitation of 380 nm and emission of 460 nm every 5 min. The relative fluorescence unit (RFU) values of all of the samples were corrected by subtracting the value of a blank sample.

### $\beta$ -Hexosaminidase activity assay

$\beta$ -Hexosaminidase activity was assayed using a *N*-acetyl- $\beta$ -D-glucosaminide (sugar-substrate) conjugated to the fluorophore 4-methylumbelliferone (4-MU). 4-MU-*N*-acetyl- $\beta$ -D-glucosaminide (0.6 mM) in 50 mM sodium acetate buffer, pH 5.4, was used with 5  $\mu$ g of brain or liver homogenates and incubated at 37 °C for several time points. The reaction was stopped by adding 0.1 M glycine buffer, pH 10. The fluorescence of released 4-MU was measured in a Tecan Spark multimode microplate reader (Tecan Trading) at an emission wavelength of 448 nm and excitation at 362 nm. The RFU values of all of the samples were corrected by subtracting the value of a blank sample.

### Immunofluorescence assays in cells and tissues

**Cells.** A total of 50,000 cells were plated on fibronectin-coated glass cover slips, fixed with 4% paraformaldehyde and blocked with 5% donkey serum in PBS for 1 h at room temperature. This was followed by overnight incubation at 4 °C with primary antibodies. The primary antibodies used in this study were mouse anti-LAMP2 (1:500) and rabbit anti-Flag (1:500). Secondary antibodies (Alexa-Fluor) were then applied in 1% BSA, 0.3% Triton X-100 for 2 h. Images were acquired in z-series on a spinning-disk confocal system and maximum-intensity projections were processed using ImageJ v.1.52.

**Tissues.** Immediately after euthanasia, brains and livers were removed and fixed by immersion in 4% paraformaldehyde overnight, and embedded in paraffin. For experiments including *Cln3*-knockout mice, after fixation, tissues were transferred to 30% sucrose overnight, embedded in OCT (TissueTek) and frozen. For immunofluorescence staining of paraffin-embedded tissues, 5  $\mu$ m sections were deparaffinized. Heat-induced antigen retrieval was performed in Borg Decloaker (Biorcare Medical) buffer using a pressure cooker (Dako). For cryo-sections, antigen retrieval was omitted. Sections were blocked in 5% donkey

# Article

serum, 0.3% Triton X-100 for 2 h. Primary antibodies (rat anti-LAMP1, mouse anti-HA and rabbit anti-NeuN) were applied in 1% BSA, 0.3% Triton X-100 for 48 h. Secondary antibodies were then applied in 1% BSA, 0.3% Triton X-100 for 2 h. The sections were covered by cover slips with Vectashield (Vector Laboratories). Images were acquired in z-series on a spinning-disk confocal system.

Images were acquired using the Airyscan2 LSM980 microscope or Zeiss AxioVert200M microscope using ZEN v.3.2 imaging software or MetaMorph 7, respectively, and maximum-intensity projections were processed using ImageJ v.1.52.

## Cell lines and plasmids

The pMXs-IRES-Bsd vector was obtained from Cell Biolabs. HEK293T cells were purchased from ATCC and authenticated by STR profiling. The sequence of the human *CLN3* cDNA was mutated to remove the proto-spacer adjacent motif, as well as within the seed region of the sgRNA used in this study. Other mutations were also introduced to the sequence targeting another sgRNA site that was not included in the current study.

## Cell culture

HEK293T cells and their derivatives were cultured in DMEM base medium with 10% inactivated fetal calf serum (Thermo Fisher Scientific) supplemented with 2 mM glutamine, penicillin and streptomycin (Thermo Fisher Scientific). All of the cell lines were maintained at 37 °C and 5% CO<sub>2</sub>. For all LysoIP experiments using HEK293T cell lines, cells were cultured in RPMI base medium supplemented with 4 nM of LysoTracker Red DND-99 for 1 h before processing as described in ref.<sup>12</sup>.

## Virus production and transduction

Lentiviruses were produced by transfecting HEK293T cells with lentiviral plasmids of the lysosomal tags in combination with the packaging plasmids VSV-G envelope and ΔVPR. For retrovirus production, pMXs Flag-CLN3 plasmid and retroviral packaging plasmids Gag-Pol and VSV-G were co-transfected using XtremeGene9 transfection reagent. The culture medium was changed to DMEM supplemented with 30% inactivated fetal calf serum 16 h after transfection. The virus-containing supernatant was collected 48 h after transfection and centrifuged for 5 min at 400g to remove cells and then frozen at -80 °C. To establish cell lines with stable expression, 500,000 WT or *CLN3*-knock-out HEK293T cells were plated in six-well plates in DMEM with 10% inactivated fetal calf serum and 8 μg ml<sup>-1</sup> polybrene and 100–250 μl of virus-containing medium. Spin-infection was performed at 2,200 rpm for 45 min at 37 °C. Cells with the virus were then incubated for 16 h before adding fresh culture medium containing blasticidin (InvivoGen) for selection for at least 72 h. Stable cell lines were tested for proper localization of the lysosomal tag or Flag-CLN3 using immunofluorescence.

## Generation of knock-out cell lines using CRISPR-Cas9 technology

Human *CLN3* was depleted using the pLentiCRISPRv1 system. The following sense (S) and antisense (AS) oligo-nucleotides were cloned into pLentiCRISPRv1: (Upper case bases are for target gene and lower case bases are for cloning) sgCLN3 (S): caccgGTACAGCTCAGCAATACCGC; sgCLN3 (AS): aaacGCGGTATTGCTGAGCGTGACc; sgAAVS (S): caccgTCCCCTCCACCCACAGTG; and sgAAVS (AS): aaacACTGTGGGGTGGAGGGGAc.

To generate knock-out cells, HEK293T cells were transduced as described above and after puromycin selection, cells were single-cell sorted into 96-well plates containing 200 μl of DMEM supplemented with 30% inactivated fetal calf serum. Cell clones with the desired knock-outs were identified by Illumina amplicon deep sequencing at Massachusetts General Hospital. Control cells were generated by targeting the non-coding *AAVS1* locus as described previously<sup>50</sup>.

## Generation and purification of GPG and GPE

To obtain GPG and GPE, phosphatidylglycerol (16:0 PG) and phosphatidylethanolamine (16:0 PE) (Avanti) were saponified as follows: 25 mg of each was dissolved in a 3 ml 2:1 chloroform:methanol solution in a clean 20 ml scintillation vial provided with a magnetic stirring bar. NaOH (1 ml of 2 M) was added to the mixture and the reaction was stirred at room temperature for 2 h. HCl (1 ml of 2 M) was used to neutralize the reaction and GPDs were extracted with 2 ml water twice into separate vessels. Aqueous phases were combined, washed with 2 ml chloroform, frozen in liquid nitrogen and lyophilized. To separate the crude product from inorganic salts, the lyophilates were dissolved in a minimum amount of methanol, and the desired products were separated from the precipitants by paper filtration.

## Metabolite extraction from yeast

*Saccharomyces cerevisiae* (WT strain BY6741 and *Btn1Δ* strain P35H10) were obtained by the Jarosz laboratory at Stanford University. Yeast strains were grown in yeast nitrogen base (with amino acids and 2% glucose) and collected during the early growth phase (OD<sub>600</sub> of 0.5–1.0). Cells were centrifuged for 3 min at 2,700g at the lowest temperature and the cell pellets were washed twice with cold PBS. Cells were then resuspended in 1 ml of prechilled quenching solution (2:2:1 mixture of ACN:MeOH:H<sub>2</sub>O) and transferred to 2 ml vials half filled with 0.5 mm glass beads. The cells were homogenized by 10 pulses of 10–20 s each at 3,800–5,000 rpm with 2 min rest on ice between each pulse. The homogenates were removed and centrifuged at 2,700g in the cold for 2 min. The supernatants were removed to prechilled tubes and centrifuged for 30 min at 17,000g in the cold. The supernatants were transferred to new tubes and stored at -80 °C for subsequent analysis using LC-MS. This protocol was adapted from ref.<sup>51</sup>

## Glycerophosphodiester phosphodiesterase activity assay

HEK293T cells stably expressing 3×Flag-GDE or 3×Flag-CLN3 were collected and homogenized on ice. The resulting supernatant was immunoprecipitated with anti-Flag-conjugated agarose beads for 4 h at 4 °C. Tagged proteins were competitively eluted using Flag peptide. The activity assay was carried out in 50 μl buffer containing 100 mM Tris (pH 7.5), 10 mM MgCl<sub>2</sub>, 2 μg fatty-acid free BSA, 50 μM labelled D9-GPC, and 2 μg of 3×Flag-GDE or 3×Flag-CLN3 protein<sup>52</sup>. The incubation was performed at 37 °C and terminated by the addition of 80% cold methanol. GDE activity was assessed by the change in the level of D9-GPC and the generation of D9-choline using MS analysis.

## Measurement of lysosomal pH

Lysosomal pH was measured using the LysoSensor Yellow/Blue DND-160 indicator dye. In brief, the cells were incubated with 5 μM LysoSensor Yellow/Blue DND-160 in complete cell growth medium for 10 min at 37 °C. The cells were then washed three times with PBS. The cells were subsequently analysed using microplate reader and fluorescence emissions were collected at 440 nm and 540 nm for excitations at 340 nm and 380 nm, respectively. LysoSensor Yellow/Blue is ratiometric, so the ratio of light excited at 340 nm to 380 nm is proportional with lysosomal pH that is measured using a calibration curve. To create a standard curve, the cells with LysoSensor were equalized by incubating them in a series of MES calibration buffer (20 mM MES (2-(*N*-morpholino) ethanesulfonic acid), 110 mM KCl, 20 mM NaCl, 10 μM monesin, 20 μM nigericin) at pH values ranging from 3.5 to 7.5 for 10 min at 37 °C.

## In vitro lipid-uptake assay

Fatty-acid-free BSA (1.25%; FAF-BSA) was dissolved in DMEM medium at 37 °C and subsequently filtered. Then, 0.1 mM of 16:0–18:1 D5-PG lipid (Avanti) was incubated with FAF-BSA at room temperature for 30 min on a rotating shaker. The molar ratio of D5-PG to FAF-BSA was 1:1. This mixture of D5-PG conjugated to FAF-BSA was considered as 10× stock

and experiments were performed at a 1× working concentration. Before supplementation with D5-PG:FAF-BSA, cells were serum-starved for 2 h to enhance endocytosis. Cells were then incubated with the mixture for 4 h. Treated and untreated cells were subjected to LysoIP followed by LC-MS analysis as described above.

### D9-dipalmitoylphosphatidylcholine (16:0-16:0) and D9-choline labelling in primary neurons

*Cln3<sup>-/-</sup>* and *Cln3<sup>-/-</sup>* postnatal (day 1) mouse pups were euthanized by decapitation, and the cortices were collected by microdissection. Fresh cortices were then washed by dissection medium (Thermo Fisher Scientific, 14170161) before tissue dissociation using scissors, 0.25% trypsin and pipette trituration. Cell strainers (70 µm) were used to remove tissue chunks from digested cortices. Dissociated neurons were plated with minimum essential medium (Thermo Fisher Scientific, 21010046) containing 10% inactivated fetal calf serum (Thermo Fisher Scientific, 10438026) supplemented with 2 mM glutamine, penicillin and streptomycin (Thermo Fisher Scientific, SV30010) in 24-well plates with coverslips coated with poly-L-lysine (Newcomer Supply, 1339A). A complete medium change to neurobasal medium (Thermo Fisher Scientific, 21103049) supplemented with B-27 (Invitrogen, 17504044) and 2 mM GlutaMax (Thermo Fisher Scientific, 35050061) was performed 24 h after initial seeding. Primary neuronal cells were maintained at 37 °C and 5% CO<sub>2</sub>. Neurons (day 7) were first starved in neurobasal medium without B-27 for 30 min, and were then labelled with 37.6 µM d9-dipalmitoylphosphatidylcholine (16:0-16:0) conjugated with 0.25% fatty acid-free bovine serum albumin (BSA) or 37.6 µM d9-choline chloride in neurobasal medium and B-27 for 24 h. Molar percent enrichment for trimethyl-d9 labelling was calculated based on the following equation:

$$\text{MPE} = \frac{\sum_{i=0}^n i \times m_i}{n \times \sum_{i=0}^n m_i} \times 100\%$$

where  $n$  is the number of hydrogen atoms,  $m_i$  the abundance of a mass isotopomer and  $i$  the labelling state (M+i). As a choline-containing metabolite can only be present as unlabelled (M+0) or D9-labelled (M+9), the equation can be simplified as follows:

$$\text{MPE} = \frac{\text{D9-labelled species}}{\text{D9-labelled species} + \text{unlabelled species}} \times 100\%$$

### Human studies

**Ethics statement.** The Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health IRB approved the natural history studies. Individuals of any age, sex and race who have at least one variant in *CLN3* and clinical symptoms suggestive of CLN3 disease were enrolled following appropriate provision of consent and assent. For children, parents or legally authorized representatives provided permission and consent.

**CSF sample collection.** Collection of CSF samples used in this study was as previously described<sup>35</sup>. We performed lumbar puncture under sedation, in participants unable to cooperate with the procedure when awake, to collect CSF from individuals with CLN3 (Investigations of Juvenile Neuronal Ceroid Lipofuscinosis (CLN3); NCT033007304), creatine transporter deficiency (NCT02931682) or Smith-Lemli-Opitz syndrome (NCT00001721). Anonymized residual CSF samples from outside clinical laboratories were also collected (NCT00344331). Participants in the CLN3 natural history study have a consistent identifier (SP\_...) across publications to facilitate data comparisons. These are included in Supplementary Table 5.

### Data preparation and statistics

Displays of quantitative data were prepared in Microsoft Excel v.16.61 and 16.62 and GraphPad Prism v.9.0. Statistical comparisons were performed using two-tailed unpaired *t*-tests in Prism or Excel, unless stated otherwise in the figure legends. All of the displayed measurements represent samples generated independently or biological replicates. Immunoblot and immunofluorescence data are representative experiments and the number of independent experiments performed is indicated in the legend. Drawings used in creating the illustrations in Fig. 1a,c and Extended Data Figs. 2a and 4a,e were created using BioRender and were used with permission.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

The MS proteomics data were deposited to the ProteomeXchange Consortium through the PRIDE partner repository with the dataset identifier PXD018624. Known lysosomal proteins in our proteomics analysis were defined based on Gene Ontology Cellular Component or UniProt subcellular localization annotation (<https://www.uniprot.org/>). The conditional LysoTag mouse was deposited at the Jackson Laboratory (strain O35401). Other unique biological materials in the form of plasmids or cell lines are available from the corresponding author on request. Detailed lipidomics and metabolomics data are provided in Supplementary Tables 3 and 4, respectively. Other data generated are available from the corresponding author on request. Source data are provided with this paper.

- Wyant, G. A. et al. NUFIP1 is a ribosome receptor for starvation-induced ribophagy. *Science* **360**, 751–758 (2018).
- Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
- Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R. & Dunn, W. B. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Gigascience* **2**, 13 (2013).
- Bird, S. S., Marur, V. R., Sniatynski, M. J., Greenberg, H. K. & Kristal, B. S. Serum lipidomics profiling using LC-MS and high-energy collisional dissociation fragmentation: focus on triglyceride detection and characterization. *Anal. Chem.* **83**, 6648–6657 (2011).
- Taguchi, R. & Ishikawa, M. Precise and global identification of phospholipid molecular species by an Orbitrap mass spectrometer and automated search engine Lipid Search. *J. Chromatogr. A* **1217**, 4229–4239 (2010).
- Yamada, T. et al. Development of a lipid profiling system using reverse-phase liquid chromatography coupled to high-resolution mass spectrometry with rapid polarity switching and an automated lipid identification software. *J. Chromatogr. A* **1292**, 211–218 (2013).
- Hankin, J. A., Murphy, R. C., Barkley, R. M. & Gijon, M. A. Ion mobility and tandem mass spectrometry of phosphatidylglycerol and bis(monoacylglycerol)phosphate (BMP). *Int. J. Mass Spectrom.* **378**, 255–263 (2015).
- Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
- Krink-Koutsoubelis, N. et al. Engineered production of short-chain acyl-coenzyme A esters in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **7**, 1105–1115 (2018).
- Zheng, B., Berrie, C. P., Corda, D. & Farquhar, M. G. GDE1/MIR16 is a glycerophosphoinositol phosphodiesterase regulated by stimulation of G protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **100**, 1745–1750 (2003).

**Acknowledgements** We thank the members of the Abu-Remaileh and Sabatini laboratories for insights; J. M. Kirkpatrick and the FLI Core Proteomics; and the participants and families in this study for their invaluable contributions. This work was supported by grants from the NIH (DP2-CA271386), Stanford Alzheimer's Disease Research Center (ADRC), Charles King fellowship, BeatBatten (Netherlands) and NCL-Stiftung (Germany) foundations to M.A.-R. and from the NIH (R01 CA103866 and R01 CA129105) to D.M.S. N.N.L. and W.D. were also supported by the BeatBatten (Netherlands) and NCL-Stiftung (Germany) foundations; U.N.M. by the Stanford ChEM-H Chemistry/Biology Interface Program, O'Leary-Thiry Graduate Fellowship and NIH T32 training grant (T32GM120007); A.L.C. by the F31 predoctoral NRSA (5F31DK113665); and C.G. and R.T. were partially supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Projektnummer 23928380—TRR 152 and the NCL-Stiftung (Germany); A.O. by DFG through the Research Training Group ProMoAge (GRK 2155),



# Article

the Else Kröner Fresenius Stiftung (no. 2019\_A79), the Fritz-Thyssen foundation (no. 10.20.1.022MN) and the Chan Zuckerberg Initiative Neurodegeneration Challenge Network (nos 2020-221617 and 2021-230967). The FLI is a member of the Leibniz Association and is financially supported by the Federal Government of Germany and the State of Thuringia. The CLN3 natural history study is funded by an NIH Clinical Center Bench to Bedside award and the NICHD intramural research programme (ZIA HD008989). M.A.-R. is a Terman Faculty Fellow at Stanford University.

**Author contributions** M.A.-R. and D.M.S. initiated the project and, with help from N.N.L., designed the research plan. N.N.L. and W.D. performed most of the experiments and analysed the data with help from A.L.C., V.D. and S.H.C.; S.H.C., T.K. and C.A.L. had a critical role in LC-MS analysis. W.D. also helped with analysing LC-MS data and U.N.M. generated GPD chemical standards. I.H. and A.O. designed the proteomic experiments and analysed the lysosomal proteomic data. R.T. and C.G. were consulted on CLN3 function and edited the manuscript.

A.N.D.D. and F.D.P. supervised the CLN3 natural history study and provided the CSF samples. M.A.-R. and N.N.L. wrote the manuscript and D.M.S. edited it.

**Competing interests** M.A.-R. is a scientific advisory board member of Lycia Therapeutics. A.N.D.D. and F.D.P. have a collaborative research agreement with Amicus Therapeutics. The other authors declare no competing interests.

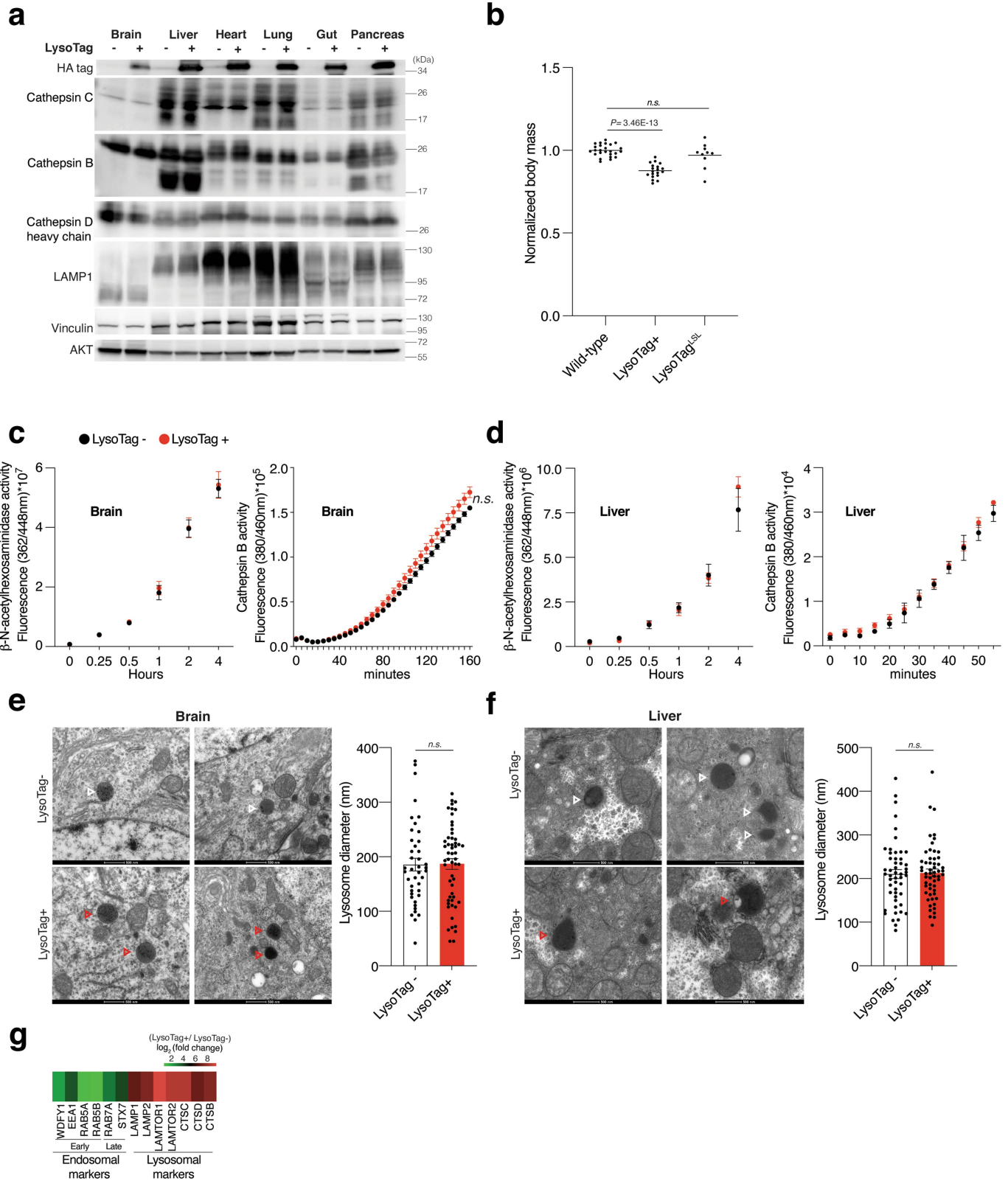
#### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-05221-y>.

**Correspondence and requests for materials** should be addressed to Monther Abu-Remaileh.

**Peer review information** *Nature* thanks David Pearce, Peter Vangheluwe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

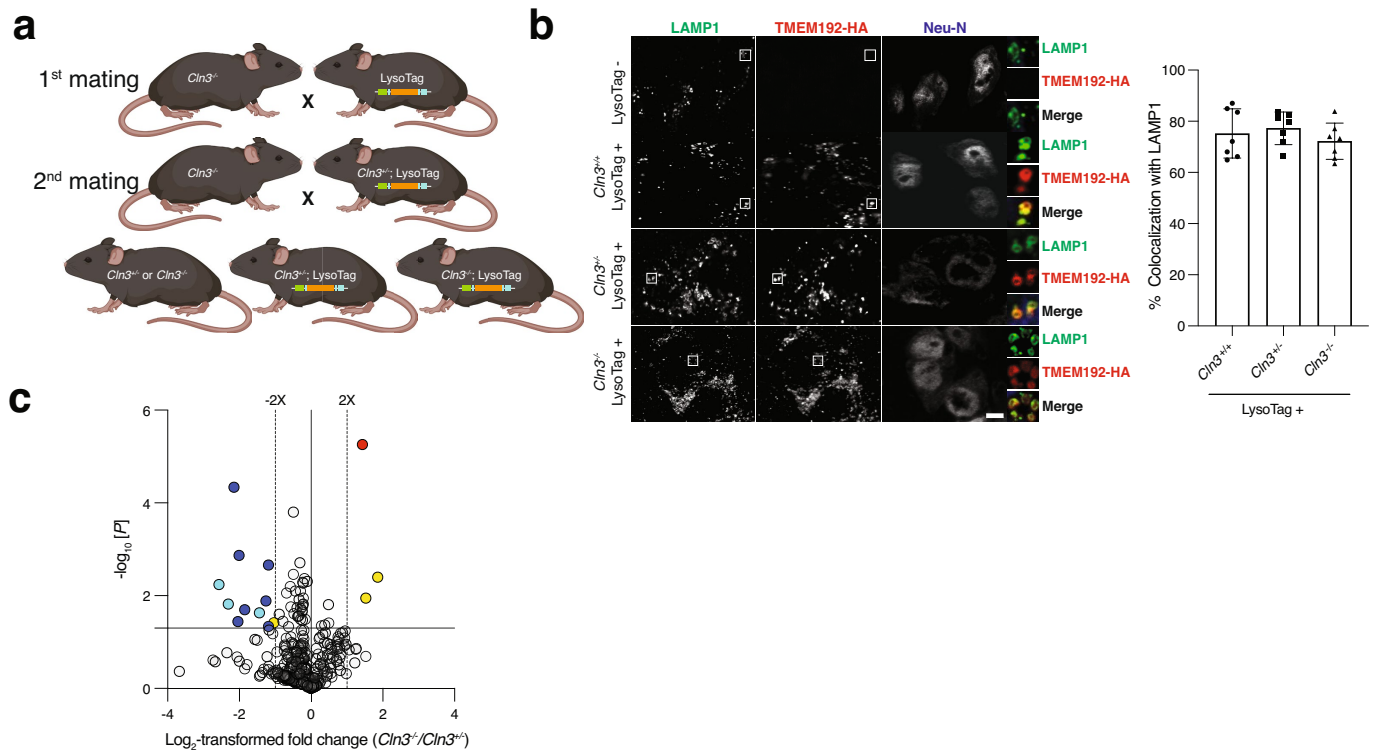


Extended Data Fig. 1 | See next page for caption.

# Article

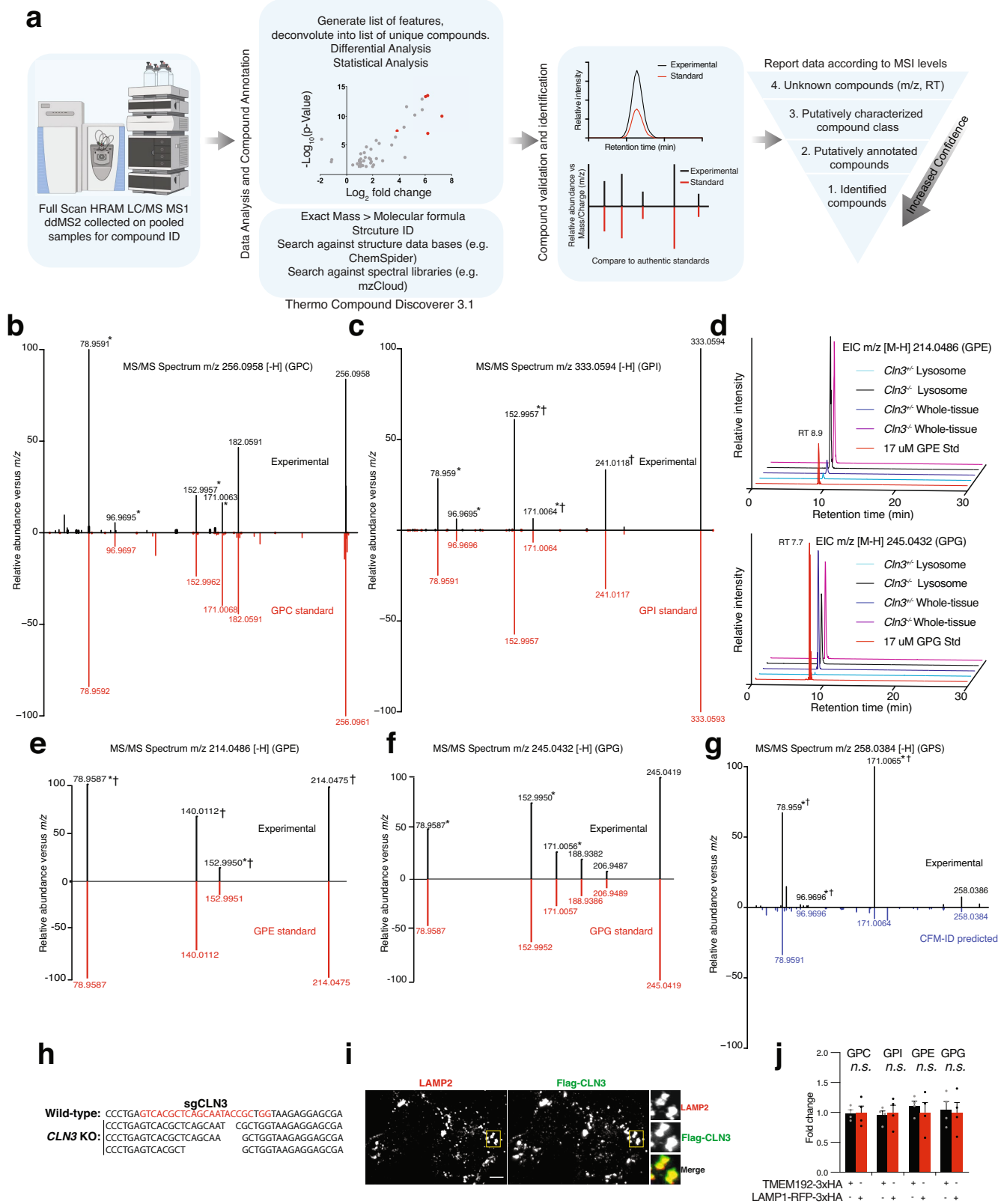
**Extended Data Fig. 1 | The effects of LysoTag expression on lysosomes in vivo.** **a**, In the constitutive LysoTag mice, TMEM192-3×HA is expressed across all tissues examined as determined by immunoblotting using antibodies to the HA epitope. No difference in the expression of lysosomal markers LAMP1, Cathepsin B, C, and D is observed upon the expression of TMEM192-3×HA. Vinculin and AKT were used as loading controls. Numbers indicate molecular weights in kDa according to protein standards run on the same gel. Immunoblot is representative of two independent experiments. **b**, Normalized body weights of mice with the indicated genotypes at 5-9 weeks of age. Mice were weighed individually. The body mass of each mouse was normalized to the mean mass of sex-matched wild-type littermates before pooling by genotype for statistical analyses using a two-tailed unpaired *t*-test. ( $n = 25, 19,$  and  $10$  for Wild type, LysoTag<sup>+</sup>, and LysoTag<sup>LSL</sup> mice, respectively.). **c,d**, Measurement of lysosomal hydrolase activity in brain and liver tissues of the constitutive LysoTag mice shows no effect of TMEM192-3×HA expression on lysosomal function. For measuring Cathepsin B activity, its fluorogenic substrate was incubated with homogenates of the corresponding tissues (see methods). For determining  $\beta$ -hexosaminidase activity, fluorogenic 4-methylumbelliferyl (MUF)-N-acetyl- $\beta$ -Glucosaminide was used as a substrate (see methods).

Fluorescence was measured at 37 °C and data were shown following the subtraction of background fluorescence observed in the absence of homogenate (mean  $\pm$  s.e.m.;  $n=4$ ). **e,f**, Representative transmission electron micrographs showing lysosomes from brain and liver tissues of LysoTag<sup>-</sup> (control) and LysoTag<sup>+</sup> mice. Lysosomes were identified by the presence of a single membrane and granular, electron-dense appearance. White and Red arrowheads indicate lysosomes in cells from LysoTag<sup>-</sup> and LysoTag<sup>+</sup> mice, respectively.  $n = 2$  mice per group and scale bar, 500 nm. The graph showing the mean diameter of lysosomes in multiple quantified cells. The diameter of lysosome in each group was generated by measuring lysosomes using ImageJ v1.52. (For brain,  $n = 42$  and  $n = 55$  measured lysosomes from LysoTag<sup>-</sup> and LysoTag<sup>+</sup> mice, respectively. For liver,  $n = 51$  and  $n = 56$  measured lysosomes LysoTag<sup>-</sup> and LysoTag<sup>+</sup> mice, respectively. Data presented as mean  $\pm$  s.e.m. n.s: non-significant; Two-tailed unpaired *t*-test (**e,f**). **g**, Heatmap presentation of the enrichment of early- and late-endosomal markers as well as those for lysosome in the LysoIP performed from mouse liver tissue. Enrichment values are derived from Supplementary Table 1. For gel source data, see Supplementary Fig. 1.



**Extended Data Fig. 2 | Generation and validation of LysoTag mouse model for Batten disease studies.** **a**, Schematic depicting the breeding strategy to generate LysoTag models for Batten disease studies. *Cln3*<sup>-/-</sup> mice were crossed with heterozygous constitutive LysoTag mice and their progeny were then crossed back to *Cln3*<sup>-/-</sup> animals to generate the indicated experimental mouse genotypes. Mouse drawing was created with BioRender. **b**, Expression and localization of the TMEM192-3×HA fusion protein to lysosomes in brain neurons of *Cln3*<sup>+/+</sup>, *Cln3*<sup>+/-</sup> and *Cln3*<sup>-/-</sup> LysoTag mice. TMEM192-3×HA, lysosomes and neurons were detected in an immunofluorescence assay using antibodies to the HA epitope the lysosomal marker, LAMP1 and the neuronal marker Neu-N, respectively. Scale bar = 5 μm, and magnified insets are labeled with white boxes in the main image. Percentage of colocalization between TMEM192-3×HA and LAMP1 shown in the right panel was measured in *n* = 7 cells per genotype. Data are shown as mean ± s.e.m. Micrographs are representative

of at least three independent experiments. **c**, Volcano plot comparing the lipidomic data from whole brain homogenates of *Cln3*<sup>+/+</sup> and *Cln3*<sup>-/-</sup> mice shows that only few lipids change significantly and they belong to phospholipids (yellow) and lysophosphatidylglycerol (LPG) in red. Loss of CLN3 also led to a decrease in the brain of bis(monoacylglycerol)phosphate (BMP), a class of lysosomal lipids (blue). BMP/PG (cyan) annotation was used when MS/MS fragmentation was not acquired (see Supplementary Table 3 and methods for details). Horizontal line indicates a p-value of 0.05, and vertical dotted lines a fold change of 2. Each dot represents a lipid species. Data were acquired in negative ion mode and normalized to internal lipid standards for the best-matched lipid class (*n* = 5 and 4 for *Cln3*<sup>+/+</sup> and *Cln3*<sup>-/-</sup> mice, respectively). Female mice with an average age of 7 months were used. Data are in Supplementary Table 3. Two-tailed unpaired *t*-test was used.

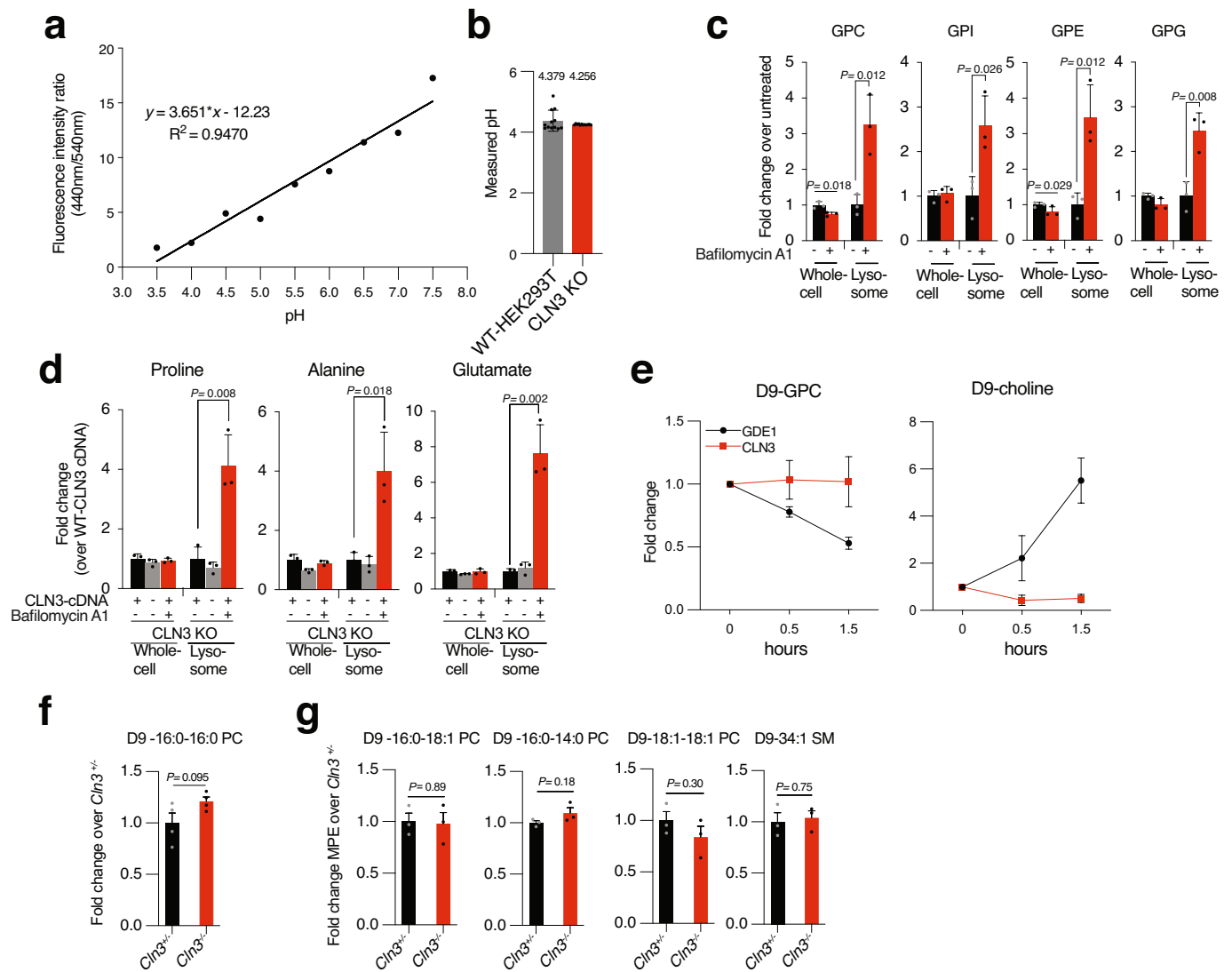


Extended Data Fig. 3 | See next page for caption.



**Extended Data Fig. 3 | Validation of glycerophosphodiester (GPDs) as the metabolites that accumulate in lysosomes upon CLN3 loss.** **a**, Identification and annotation of metabolites using untargeted metabolomics analyses. Overview of our untargeted metabolomics experimental workflow. MS1 data were collected for every sample using full scan mode on a high-resolution accurate-mass instrument. MS/MS data were collected from pooled samples only to aid with compound identification. Data were then analysed using Thermo Compound Discoverer v3.1 to generate a list of unique features which were deconvoluted into a list of unique compounds. Differential and statistical analyses were performed. Compounds which were statistically significant in the differential analyses were validated and identified using authentic standards, when available. If no authentic standard was available, compounds were annotated by comparing MS/MS data to a database, or *in silico* fragmentation prediction (competitive fragmentation modelling-ID, CFM-ID). Data were then reported according to the Metabolomics Standards Initiative (MSI) guidelines. **b,c**, Mirror plots for GPC and GPI in negative ion mode, respectively. Fragments common to the glycerol-phosphate group are indicated with an asterisk\*. Fragments common to those found in the MS/MS spectra reported by Kopp et al.<sup>32</sup> are indicated with †. **d**, EICs for GPE and GPG across a range of samples were analysed alongside in-house generated standards (Std) and showed a matching retention time (RT). **e,f**, Mirror plots for

GPE and GPG, respectively. Fragments common to the glycerol-phosphate group are indicated with an asterisk\*. Fragments common to those found in the MS/MS spectra reported by Kopp et al.<sup>32</sup> are indicated with †. **g**, Mirror plots showing MS/MS spectra for glycerophosphoserine from experimental samples compared to MS/MS spectra generated *in silico* using CFM-ID. Fragments common to the glycerol-phosphate group are indicated with an asterisk\*. Fragments common to those found in the MS/MS spectra reported in Kopp et al.<sup>32</sup> are indicated with †. **h**, Amplicon sequencing was used to validate *CLN3* knock out cells generated using CRISPR-Cas9. Three mutant alleles were identified and all were found to generate frameshift deletions in the *CLN3* coding sequence as compared to the wild-type sequence. **i**, Localization of Flag-*CLN3* protein to lysosomes. Flag-*CLN3* and lysosomes were detected in an immunofluorescence assay using antibodies to the Flag epitope and the lysosomal marker LAMP2, respectively. Scale bar = 5  $\mu$ m. Micrographs are representative images of three experiments. **j**, Changes in GPD levels in *CLN3*-deficient cells expressing TMEM192-3 $\times$ HA tag or LAMP1-3 $\times$ HA tag. Data presented as a comparison between the increase in the lysosomal abundance of GPDs upon *CLN3* loss in LAMP1-3 $\times$ HA tagged lysosomes relative to that in TMEM192-3 $\times$ HA tagged lysosomes (mean  $\pm$  s.e.m.;  $n = 4$  biologically independent samples). n.s: non-significant; Two-tailed unpaired t-test.



**Extended Data Fig. 4 | Testing the effects of CLN3 loss on lysosomal pH and cellular lipid metabolism.** **a** to **d**, CLN3 loss does not increase lysosomal pH. **a**, A standard calibration curve of the ratiometric dye LysoSensor Yellow/Blue DND-160. See methods for experimental details. **b**, Lysosomal pH in wild-type and CLN3 KO HEK293T cells as calculated using the standard curve in **a**. Data are presented as mean  $\pm$  s.d.,  $n = 12$  biologically independent samples. **c**, Targeted analyses of the fold changes in whole-cell and lysosomal levels of GPDs upon treatment with 500 nM Bafilomycin A1 for 6 h in CLN3 expressing HEK293T cells. Data are presented as mean  $\pm$  s.d.,  $n = 3$  biologically independent samples, (Two-tailed unpaired  $t$ -test). **d**, The levels of several amino acids whose egress from lysosomes is sensitive to the proton gradient across the lysosomal membrane are not affected by CLN3 loss. The levels of proline, alanine, and glutamate in whole cells and lysosomes were compared between CLN3 KO cells with and without CLN3 cDNA addback and upon treatment with 500 nM Bafilomycin A1 (BafA1) for 6 h. Data are presented as mean  $\pm$  s.d.,  $n = 3$  biologically independent samples, (Two-tailed unpaired  $t$ -test). **e**, The recombinant CLN3

protein does not have glycerophosphodiesterase activity. Deuterated GPC (D9-GPC) was incubated with recombinant 3 $\times$ Flag-CLN3 or the positive control glycerophosphodiesterase1 (3 $\times$ Flag-GDE-1) for the indicated time points. The extent of D9-GPC hydrolysis was determined by measuring the decline in its level and the increase in the levels of the product D9-choline. Data presented as mean  $\pm$  s.e.m. of  $n = 3$  biologically independent samples. **f**, CLN3 loss does not affect tracer uptake. Fold change in tracer levels (D9-16:0-16:0 PC) normalized to total protein for samples measured in Fig. 4f and g. Data presented as mean  $\pm$  s.e.m. of  $n = 4$  biologically independent samples. **g**, CLN3 loss does not affect the biosynthesis or turnover of PC and sphingomyelin (SM) in cells. Free D9-choline was used as a tracer. Data are presented as fold changes in the whole-cell molar percent enrichment (MPE) of D9-choline-containing lipids in cortical neuron cultures prepared from *Cln3*<sup>-/-</sup> mice relative to those from *Cln3*<sup>+/+</sup> animals (mean  $\pm$  s.e.m.,  $n = 3$  biologically independent samples, (Two-tailed unpaired  $t$ -test)).

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- |                                     |                                     |                                                                                                                                                                                                                                                            |
|-------------------------------------|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of all covariates tested                                                                                                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/>            | Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated                                                                                                                                                         |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection All softwares for data collection are publicly or commercially available: Tune version 2.9, Xcalibur 4.1, MetaMorph 7, ZEN 3.2

Data analysis All softwares for data analysis are publicly or commercially available: Microsoft Excel v16.61 and 16.62, GraphPad Prism v9.0, XCalibur v4.1, ImageJ v1.52, CompoundDiscoverer v3.1, Spectronaut Professional v13.10, LipidSearch v4.2.21, TraceFinder software v5.0, R version 3.4.1 and RStudio server version 1.1.463,

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018624. Known lysosomal proteins in our proteomics analysis were defined based on Gene Ontology Cellular Component or UniProt sub-cellular localization annotation (<https://www.uniprot.org/>). The conditional LysoTag mouse was deposited at the Jackson Laboratory (Strain #035401). Other unique biological materials in the form of plasmids or cell lines are available from corresponding author upon request. Detailed lipidomics and metabolomics data are provided in

Supplementary Tables 3 and 4, respectively. Other data generated are available from the corresponding author upon request. Source data are provided with this paper.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No formal sample-size calculations were performed. However, based on previously published work and literature in the field that have conducted similar experiments as well as the number of replicates sufficient for statistical analyses, we used a minimal sample size of $n = 3$ for quantitative experiments. For some experiments, where we expected higher variance, generally due to lower amounts of starting material (e.g. LysolP), we included more samples per condition. Taking this into account for in vivo experiments, the final exact numbers were determined based on the availability of each genotype, experimental design and animal housing conditions. We observed that sample sizes of 3 to 6 animals per group was adequate sample size and sufficient to obtain statistical power. For cell based experiments, each independent culture of cells was defined as biological replicate.
Data exclusions	In rare cases where we had very poor detection across the majority of metabolites the whole sample was excluded.
Replication	For most of our figures, at least three independent experiments were performed and have not encountered discrepancy between results and conclusions. Extended data figure 1A was reproduced only two times. Extended data figures. 1E and 1F, the transmission electron microscopy (TEM) provide the same observations from two animals per group.
Randomization	To ensure rapid processing, LysolP protocols were performed one sample at a time and usually by two people and samples were randomized across conditions. To prevent any systematic bias from possible column carryover, sample order was randomized in LC/MS runs. For pairwise comparisons, samples were processed in internally controlled pairs (i.e. Wild-type vs. KO), while matching sex and age. For other experiments where samples were not allocated randomly, group allocation and randomization were unnecessary because all samples were measured independently in the same way in an internally controlled manner.
Blinding	Sample blinding was not performed for most experiments as they did not involve formal qualitative scoring. Furthermore, all measurement modalities were quantitative and all samples and experiments were performed in the same way regardless of group or treatment. For in vivo metabolomic and proteomic experiments as well as for analysis of patient CSF samples, the investigators were blinded during data acquisition since mouse or patient number was used as identifier, respectively.  Immunofluorescence (IF) assessments were qualitative and based on identifiable features. In at least one replicate for each IF experiment, samples were independently evaluated by at least one other investigator whose interpretation supported the conclusions in the manuscript.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Included in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

### Antibodies used

Antibody Target:Company Catalogue number/clone number Dilution used in this study  
 LAMP2: Santa Cruz Biotechnology: sc-18822/ H4B4 1:1000  
 LAMP1: Developmental Studies Hybridoma Bank: 1D4B 1:1000  
 VDAC1: Cell Signaling Technology: CST-4661/ D73D12 1:2000

GOLGIN-97: Cell Signaling Technology: CST-13192/ D8P2K 1:500  
 CALR: Cell Signaling Technology: CST-12238/ D3E6 1:2000  
 CTSB: Cell Signaling Technology: CST-31718/ D1C7Y 1:1000  
 CTSC: Santa Cruz Biotechnology: sc-74590/ D-6 1:1000  
 HA epitope: Cell Signaling Technology: CST-2367/ 6E2 1:2000  
 FLAG epitope: Cell Signaling Technology: CST-14793/ D6W5B 1:2000  
 Catalase: Cell Signaling Technology: CST-14097/ D5N7V 1:1000  
 CTSD: R&D Systems: AF1029/ polyclonal 1:1000  
 NeuN: Cell Signaling Technology: CST-24307/ D4G4O 1:1000  
 AKT: Cell Signaling Technology: CST-4685/ 11E7 1:3000  
 Vinculin: Santa Cruz Biotechnology: sc-73614/ 7F9 1:3000  
 p70 S6 Kinase: Cell Signaling Technology: CST-9202/ polyclonal 1:1000  
 Histone H3: Cell Signaling Technology: CST-4499/ D1H2 1:500

Secondary antibodies for IF: Goat anti-Rat Alexa Fluor 488 (Thermo A-11006), Goat anti-rabbit Alexa Fluor 488 (Thermo A-21206), Goat anti-Mouse Alexa Fluor 647 (ThermoA-21240), Goat anti-Mouse Alexa Fluor 488 (Thermo, A-11001), and Goat anti-rabbit Alexa Fluor 594 (Thermo, A32740).

Secondary antibodies for immunoblotting: Anti-rabbit HRP (CST, 7074), anti-mouse HRP (CST, 7076), anti-rat HRP (CST, 7077)

## Validation

All antibodies used in this study are acquired from commercial sources and were validated for specificity and species reactivity by the manufacturer. The information are readily available on the website of the manufacturer listed here:

Antibody Target Direct link to manufacturer's website

LAMP2 <https://www.scbt.com/p/lamp-2-antibody-h4b4>

LAMP1 <https://dshb.biology.uiowa.edu/1D4B>

VDAC1 <https://www.cellsignal.com/products/primary-antibodies/vdac-d73d12-rabbit-mab/4661>

GOLGIN-97 [https://www.cellsignal.com/products/primary-antibodies/golgin-97-d8p2k-rabbit-mab/13192?site-search-type=Products&N=4294956287&Ntt=golgin-97&fromPage=plp&\\_requestid=789344](https://www.cellsignal.com/products/primary-antibodies/golgin-97-d8p2k-rabbit-mab/13192?site-search-type=Products&N=4294956287&Ntt=golgin-97&fromPage=plp&_requestid=789344)

CALR <https://www.cellsignal.com/products/primary-antibodies/calreticulin-d3e6-xp-rabbit-mab/12238>

CTSB <https://www.cellsignal.com/products/primary-antibodies/cathepsin-b-d1c7y-xp-rabbit-mab/31718>

CTSC <https://www.scbt.com/p/cathepsin-c-antibody-d-6>

HA epitope [https://www.cellsignal.com/products/primary-antibodies/ha-tag-6e2-mouse-mab/2367?site-search-type=Products&N=4294956287&Ntt=ha+epitope+2367&fromPage=plp&\\_requestid=789757](https://www.cellsignal.com/products/primary-antibodies/ha-tag-6e2-mouse-mab/2367?site-search-type=Products&N=4294956287&Ntt=ha+epitope+2367&fromPage=plp&_requestid=789757)

FLAG epitope [https://www.cellsignal.com/products/primary-antibodies/dykdddk-tag-d6w5b-rabbit-mab-binds-to-same-epitope-as-sigma-s-anti-flag-m2-antibody/14793?site-search-type=Products&N=4294956287&Ntt=flag+epitope+14793&fromPage=plp&\\_requestid=789909](https://www.cellsignal.com/products/primary-antibodies/dykdddk-tag-d6w5b-rabbit-mab-binds-to-same-epitope-as-sigma-s-anti-flag-m2-antibody/14793?site-search-type=Products&N=4294956287&Ntt=flag+epitope+14793&fromPage=plp&_requestid=789909)

Catalase [https://www.cellsignal.com/products/primary-antibodies/catalase-d5n7v-rabbit-mab/14097?site-search-type=Products&N=4294956287&Ntt=catalase+14097&fromPage=plp&\\_requestid=790017](https://www.cellsignal.com/products/primary-antibodies/catalase-d5n7v-rabbit-mab/14097?site-search-type=Products&N=4294956287&Ntt=catalase+14097&fromPage=plp&_requestid=790017)

CTSD [https://www.rndsystems.com/products/mouse-cathepsin-d-antibody\\_af1029](https://www.rndsystems.com/products/mouse-cathepsin-d-antibody_af1029)

NeuN [https://www.cellsignal.com/products/primary-antibodies/neun-d4g4o-xp-rabbit-mab/24307?site-search-type=Products&N=4294956287&Ntt=neun+24307&fromPage=plp&\\_requestid=790222](https://www.cellsignal.com/products/primary-antibodies/neun-d4g4o-xp-rabbit-mab/24307?site-search-type=Products&N=4294956287&Ntt=neun+24307&fromPage=plp&_requestid=790222)

AKT [https://www.cellsignal.com/products/primary-antibodies/akt-pan-11e7-rabbit-mab/4685?site-search-type=Products&N=4294956287&Ntt=akt+4685%5C&fromPage=plp&\\_requestid=790281](https://www.cellsignal.com/products/primary-antibodies/akt-pan-11e7-rabbit-mab/4685?site-search-type=Products&N=4294956287&Ntt=akt+4685%5C&fromPage=plp&_requestid=790281)

Vinculin <https://www.scbt.com/p/vinculin-antibody-7f9>

p70 S6 Kinase [https://www.cellsignal.com/products/primary-antibodies/p70-s6-kinase-antibody/9202?site-search-type=Products&N=4294956287&Ntt=p70+s6+kinase+9202&fromPage=plp&\\_requestid=790500](https://www.cellsignal.com/products/primary-antibodies/p70-s6-kinase-antibody/9202?site-search-type=Products&N=4294956287&Ntt=p70+s6+kinase+9202&fromPage=plp&_requestid=790500)

Histone H3 [https://www.cellsignal.com/products/primary-antibodies/histone-h3-d1h2-xp-rabbit-mab/4499?site-search-type=Products&N=4294956287&Ntt=histone+h3+4499&fromPage=plp&\\_requestid=790630](https://www.cellsignal.com/products/primary-antibodies/histone-h3-d1h2-xp-rabbit-mab/4499?site-search-type=Products&N=4294956287&Ntt=histone+h3+4499&fromPage=plp&_requestid=790630)

Lamp1 1D4B was confirmed via IF to overlap with over-expressed TMEM192, a marker of lysosomes.

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

HEK-293T cells were from ATCC

Authentication

STR profiling



Mycoplasma contamination

Commonly misidentified lines (See [ICLAC](#) register)

## Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals   
 Age and sex are all indicated in the figure legend for each experiment. Male and female mice were used with an average age of 5.5 weeks to 7 months and age- and sex-matched in each experiment.  
 All mice were maintained on a standard light-dark cycle with access to food and water ad libitum. Animals were fasted for the fasting experiment in Fig. 1. The animal room had a controlled temperature of 22°C, and humidity (around 50%). Cages were cleaned every 4-5 days and supplies of water and food were checked daily.

Wild animals

Field-collected samples

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Recruitment

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Study protocol

Data collection

Outcomes