Urine proteomics analysis of patients with neuronal ceroid lipofuscinoses

**HIGHLIGHTS**

- The urine proteome is altered in humans and animals with NCL.
- Hexosaminidase A and LAMP1 are increased in patients with NCL.
- Betaine-homocysteine S-methyltransferase 1 is elevated in CLN2 patients.
- Proteins altered in CLN5 and CLN6 sheep models are not affected in humans.
Urine proteomics analysis of patients with neuronal ceroid lipofuscinoses

Katharina Iwan, 1 Robert Clayton, 1 Philippa Mills, 1, 2 Barbara Csányi, 3 Paul Gissen, 1, 2, 3 Sara E. Mole, 1, 4 David N. Palmer, 5 Kevin Mills, 1, 2 and Wendy E. Heywood 1, 2, 6, *

SUMMARY
The neuronal ceroid lipofuscinoses (NCL) are a group of 13 rare neurodegenerative disorders characterized by accumulation of cellular storage bodies. There are few therapeutic options, and existing tests do not monitor disease progression and treatment response. However, urine biomarkers could address this need. Proteomic analysis of CLN2 patient urine revealed activation of immune response pathways and pathways associated with the unfolded protein response. Analysis of CLN5 and CLN6 sheep model urine showed subtle changes. To confirm and investigate the relevance of candidate biomarkers a targeted LC-MS/MS proteomic assay was created. We applied this assay to additional CLN2 samples as well as other patients with NCL (CLN1, CLN3, CLN5, CLN6, and CLN7) and demonstrated that hexosaminidase-A, aspartate aminotransferase-1, and LAMP1 are increased in NCL samples and betaine-homocysteine S-methyltransferase-1 was specifically increased in patients with CLN2. These proteins could be used to monitor the effectiveness of future therapies aimed at treating systemic NCL disease.

INTRODUCTION
The neuronal ceroid lipofuscinoses (NCLs, Batten disease) are a heterogeneous group of rare neurodegenerative diseases that present mainly in childhood and as a combined group represent the most common cause of childhood neurodegeneration and dementia. The group share phenotypic characteristics including seizures, visual loss, loss of motor skills, and cognitive function resulting in premature death, but each subtype is a distinct genetic entity with unique pathophysiological characteristics (Kohlschütter et al., 2019). Brain pathology reveals a massive loss of neurons, and most cells throughout the body contain lysosome-derived fluorescent storage bodies, the main component of which is the c subunit of mitochondrial ATP synthase or Saposins A and D (Palmer, 2015). So far, at least 13 genes have been identified that lead to NCL (CLN1–8, 10–14) (Butz et al., 2020; Williams and Mole, 2012). Of all the NCLs, CLN2 disease is one of the most commonly reported with an incidence in Europe varying between 0.15 and 0.78 per 100,000 live births. Autosomal recessive mutations in the TPP1 gene lead to loss or deficiency of the lysosomal enzyme tripeptidyl peptidase 1 (TPP1). TPP1 is a lysosomal serine protease that cleaves tripeptides from the N termini of small polypeptides during protein degradation inside the lysosomes (Golabek et al., 2003; Wlodawer et al., 2003). Five other NCLs occur due to a lack or defect of lysosomal-associated proteins, namely, palmitoyl protein thioesterase 1 (PPT1) in CLN1, CLN5 in CLN5, cathepsin D (CTSD) in CLN10, progranulin in CLN11, and cathepsin F (CTSF) in CLN13 diseases. Other NCL forms resulting from missing or defective lysosomal membrane proteins are CLN3, CLN7/MFSD8, and CLN12/ATP13A2 diseases. CLN6 and CLN8 diseases result from deficiencies of endoplasmic reticulum (ER) membrane proteins CLN6 and CLN8 (Di Ronza et al., 2018; Mole et al., 2004). The very rare CLN4 disease results from defects in a chaperone protein DNAJC5 (Henderson et al., 2016), and CLN14 disease results from defects in the plasma membrane potassium channel KCTD7 gene (Staropoli et al., 2012).

To date, the only clinically approved treatment that has slowed disease progression is Brineura (cerliponase alfa) for CLN2 disease. This TPP1 enzyme replacement therapy (ERT) is delivered intra-cerebroventriculally into the brain ventricles of patients every fortnight (Schulz et al., 2018). Other treatment approaches in development for NCL include gene therapy, various small molecules, and the administration of immunosuppressive drugs to prevent neuroinflammation (Kohlschütter et al., 2019; Mole et al., 2019). Initial animal studies...
have shown promising results for an ERT with the recombinant proenzyme PPT1 for patients with CLN1 disease (Lu et al., 2015; Hu et al., 2012). ERT administered directly into the brain still cannot prevent loss of vision, hence additional intraocular/intravitreal treatment is necessary (Katz et al., 2014; Tracy et al., 2016; Whiting et al., 2014; Griffey et al., 2005; Mole et al., 2019). Clinical trials for gene therapy for patients with CLN2, CLN3, and CLN6 disease have begun (Liu et al., 2020; Kohlschütter et al., 2019), and promising data for an early clinical trial for CLN2 disease suggest an attenuation of neurological impairment. So far, preclinical and clinical studies indicate that a single treatment strategy will not suffice, as therapeutic strategies have to target the brain, the retina, and peripheral organs (Kohlschütter et al., 2019; Mole et al., 2019).

Clinical diagnoses of specific NCL forms are still challenging with investigations mainly only being conducted after NCL disease is suspected after the first symptoms are detected. Early or new-born diagnosis is crucial if a small therapeutic window exists where successful interventions can more effectively prevent or slow the neurological disease progression. Diagnosis of NCL used to require microscopic analysis of storage material using a skin biopsy. Nowadays, genetic testing and enzyme activity assays are the methods of choice. Although genetic tests can be used to diagnose the NCL form, they cannot monitor disease progression or response to treatments. Typically disease progression is determined using baseline clinical scores like the Unified Batten Disease Rating Scale (UBDRS) (de Blieck et al., 2013; Adams and Mink, 2013), the Weill Cornell CNS scale, and the Hamburg CLN2 scale to quantify seizures, loss of language, motor skills, and visual function. To accompany the clinical presentation, assessing what is happening at the molecular level will be crucial in monitoring early stages of disease and for preclinical studies or ongoing clinical trials. Robust molecular biomarkers are therefore needed to monitor even subtle changes in physiological or pathological disease progression.

Various biomarker studies in patients with different NCLs have previously been conducted to identify possible biomarkers of disease progression. Using proteomics, Sleat et al. (2017) investigated cerebrospinal fluid (CSF) and postmortem brain samples of patients with CLN1, CLN2, and CLN3 disease. Initially they investigated lysosomal proteins, i.e., TPP1, beta-hexosaminidase subunit alpha (HEXA), and cathepsins, as they are known to be altered in many lysosomal storage disorders (Sleat et al., 2017). Other studies analyzed CSF from patients with CLN2 disease and identified metabolites that correlated to disease severity scores (Sindelar et al., 2018). Urine is a highly desirable body fluid for biomarker analysis because collection is non-invasive, feasible in relatively large volumes, and at frequent intervals; because it is inexpensive to preserve; and, due to its high urea content, urine proteins are considered more stable than serum or plasma proteins. The urinary proteins present derive from glomerular filtrations of plasma, renal tubule excretion, and urogenital tract secretion, with fluctuations representing normal physiological changes as well as pathological changes due to renal or systemic disease (Harpole et al., 2016; Kalantari et al., 2015).

One of the greatest challenges in biomarker research in rare diseases is accessing enough samples for a viable study. Often researchers turn to animal or cell models to address this. One such model is the slime mold Dictyostelium, which has revealed how secretion is impaired in NCL (Huber et al., 2014; Huber, 2020). Proteins observed to be altered in Dictyostelium were also observed to be altered in human patient tissues (Sleat et al., 2017).

**Experimental design**

We have used both human and animal models in a two-phase proteomic strategy to identify proteins altered in patients with NCL.

Phase 1 the “discovery” step is an unbiased eyes-open approach that uses label-free proteomic profiling to study urine from patients with CLN2 disease and CLN5 and CLN6 disease sheep models. The technology is semi-quantitative and lengthy, therefore small sample cohorts are analyzed using this approach.

Phase 2 involves a targeted analysis using multiple reaction monitoring mass spectrometry coupled with liquid chromatography (LC-MS). A candidate biomarker peptide(s) is accurately quantitated using this technique, and multiple peptides can be measured in one analysis, therefore creating a multiplex approach. Proteins of interest identified from the phase 1 profiling analysis are considered as potential biomarker candidates for corresponding CLN2, CLN5, and CLN6 diseases in human patients. As additional biomarkers can be included we also selected peptides from other biomarker candidates described in previous NCL omics studies (Sleat et al., 2017, 2019) to confirm these in our analyses. Using the targeted
multiplex assay we applied this to a larger separate cohort of CLN2 disease patient samples and patient samples with CLN5 and CLN6 disease to confirm phase 1 findings. To see if the biomarkers may be specific for the NCL type we also included analyses of the NCL diseases CLN1, CLN3, and CLN7.

RESULTS

Phase 1: Proteomic screening of urine from human and animal models of NCL disease identify candidate biomarkers and affected pathways

Untargeted proteomic analyses of two patients with CLN2 disease and comparison with four age-matched healthy control samples demonstrated an increase in 61 proteins (p value < 0.05 with a normalized fold change >2) and 11 proteins with reduced levels (p value < 0.05 and normalized fold change < -2, Table S1). Additionally, proteomic analyses of urine from 3 CLN5-affected sheep and 13 CLN6-affected sheep were compared with seven normal control sheep (Tables S2 and S3). The volcano plot in Figure 1A summarizes the alterations found in the urinary proteome of patients with CLN2 disease and the CLN5 and CLN6 sheep models. The CLN5 profile shows the least number of significant proteins altered among the three NCL urine profiles.

Figure 1. Label-free proteomic analysis of urine from CLN2 patients and CLN5 and CLN6 sheep models

(A) Volcano plots summarizing the proteomic data. Proteins highlighted in red are significantly upregulated, proteins highlighted in blue are downregulated, whereas the proteins in gray are not significantly changed (p (ANOVA) < 0.05). (B) Summary of the affected pathways in CLN2 patient urine as determined by Ingenuity Pathway Analysis (Qiagen). This shows which cellular pathways are altered significantly in urine of patients with CLN2 disease compared with healthy controls. Only proteins that were significantly (p (ANOVA) < 0.05) changed with a fold change >2 or < -2 were used for this analysis. The y axis depicts how significantly changed a pathway is in a negative logarithmic scale of the p values. The cutoff is 1.3 as this is the equivalent to a p value of 0.05. The activity pattern is color coded: bars in gray indicate it is unknown if the pathway is activated or inactivated, bars in blue label negative z values and therefore inhibited pathways, whereas positive z values and therefore activated pathways are highlighted in orange.
To gain some insight into the effects of CLN2 disease on biological pathways an Ingenuity Pathway Analysis (IPA, QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) was conducted on the differentially expressed proteins (Figure 1B). Pathways showing no activity pattern (z = 0) were omitted. The most significantly activated pathway is the BAG2 signaling pathway, with altered levels of proteins Annexin A2, Heat shock 70 kDa protein 1A/B, and Heat shock 70 kDa proteins 2 and 8. Interestingly, the neuroinflammation signaling pathway and the synaptogenesis signaling pathway are both significantly activated, whereas the natural killer cell signaling pathway is inhibited. A functional enrichment gene ontology analysis for cellular components (CC) performed using the DAVID bioinformatics website (Huang et al., 2008, 2009) confirmed that the differentially expressed proteins detected in the urine proteome were derived from the lysosomal lumen and the lysosome (Table S4). The proteins affected primarily in patients with CLN2 disease were GOT1 aspartate aminotransferase and FMO3 dimethylaniline monooxygenase, which were absent in controls. IPA analysis was also performed on the CLN5 and CLN6 disease sheep urine proteomic data (Figure S1).

The CLN5 and CLN6 sheep datasets had one significant pathway in common, the LRX/RXR activation pathway. Interestingly, the ER stress response pathway was affected in the CLN6 sheep, as was an unfolded protein response pathway and pathways associated with the immune response. In CLN5-affected sheep, however, there were greater effects on pathways associated with carbohydrate metabolism. Significant proteins affected in CLN5 sheep urine showed small fold changes. Only one protein, beta enolase, a muscle-specific form of enolase, was increased more than 2-fold. The number of significantly affected proteins in CLN6 sheep urine was greater than that seen in CLN5 urine, whereas the fold changes were similarly small. The most affected protein, ERICH6B, glutamate-rich protein 6B of unknown function, showed a >2-fold reduction in amount.

Phase 2: confirmation of human and animal model discovery targets in an extended NCL patient cohort

A multiplexed, targeted proteomic assay was developed to confirm if the changes in pathways and protein levels observed in the proteomic profiling analyses could be detected in a separate and larger number of 9 CLN2 patient samples and 13 controls. A selection of proteins from the sheep proteomic analyses was also included to see if they were relevant in human CLN5 (n = 2) and CLN6 (n = 2) urine. Candidate selection criteria are given in methods, and the proteins included in the multiplex assay are shown in Table 1. Urine from patients with other NCL disorders, CLN1 (n = 2), CLN3 (n = 1), and CLN7 (n = 1), was also analyzed to determine whether our initial discovery findings were specific to CLN2, CLN5, and CLN6 or if similar findings could be attributed to all NCL disorders. Further sample details are given in Table S7 in methods.

Multivariate-principal component analysis (PCA) was performed to determine if the proteins in the biomarker panel are able to collectively distinguish between the control and patient samples. Figure 2 shows a PCA plot of all samples apart from one omitted outlier (patient CLN2 (4)-post). The first two components (t [1] and t[2]) explain 27% and 19.2% of the variance in urine. The variation within the dataset roughly separates the patient samples from the control samples. Some patient samples were present within the cluster of control samples (particularly CLN6 and CLN7) indicating they have a normal urine profile. The other NCL urines show a dispersed profile indicating a large degree of variation, which is away from the cluster of control samples. This indicates that these NCL urine profiles are abnormal compared to controls.

The loading plot of the multivariate analysis indicated that lysosome-associated membrane glycoprotein 1 (LAMP1), peptidyl-glycine alpha-amidating monoxygenase (PAM), and aspartate aminotransferase, cytoplasmic (GOT1) are the biggest drivers of the difference observed in the two clusters, i.e., controls and patients.

Proteins affected in CLN2 disease: Figure 3A shows univariate analysis of proteins in CLN2 patient urine compared with age- and sex-matched controls. Peptide ratios representative of betaine-homocysteine S-methyltransferase 1 (BHMT1), LAMP1, and TPP1 were significantly altered in CLN2 disease samples. As patients with CLN2 disease have mutations in the TPP1 gene this result was expected, with the level of TPP1 approaching zero. Conversely, levels of urinary BHMT1 and LAMP1 were increased, on average 17.8- and 3.4-fold, respectively, compared with the control samples. Two of these three proteins, TPP1 and BHMT1, are specific for CLN2 (Figure 3B): BHMT1 is increased 16-fold and TPP1 is decreased up to
9-fold relative to patients with other NCL types (CLN 1, CLN3, CLN5, CLN6, and CLN7). Comparison of the levels of BHMT1, LAMP1, and TPP1 in pre-treatment and ongoing treatment (post) samples did not show any visible trend (only two paired samples available so statistical testing is not possible), although for one patient, TPP1 did increase to a measurable amount (Figure 3 D), suggesting that the ERT delivered into the brain ventricles may not stay confined to the CNS. However, the other proteins in this patient did not decrease suggesting that the levels of TPP1 coming from the CNS are not at a level sufficient to correct the rest of the diseased organ systems.

ERT administered into the brain might not alter the proteins that come from the periphery, i.e., kidney lysosomes. Levels of BHMT1 have increased in Patient ID-1 with ongoing treatment, but levels of LAMP1 and TPP1 have remained constant, whereas patient ID-3 reveals a constant BHMT1 level but increasing levels for LAMP1 and TPP1 with ongoing treatment.

Proteins assessed in CLN5 and CLN6 sheep models: liver-enriched antimicrobial peptide-2 (LEAP2), desmocollin 2 (DSC2), cathepsin B (CTSB), secretogogin (SCGN), and lysosomal-associated membrane protein 2 (LAMP2) were identified as significantly altered in the sheep urine proteomic analysis and included in the targeted LC-MS/MS assay (Tables S2 and S3). CTSB and SCGN were not detectable in human urine. Figure 4 A shows label-free proteomics data for increased levels of DSC2 in CLN5 and CLN6 (p < 0.017), increased LAMP2 in CLN5 (p < 0.028), and reduced LEAP2 in the CLN5 (p < 0.01) and CLN6 (p < 0.003) sheep urine. Targeted analysis (Figure 4 B) of LAMP2, LEAP2, and DSC2 for CLN5 and CLN6 patients did not show a difference relative to controls.

| Gene          | CLN2 patients (label-free proteomics)       | Location                  |
|---------------|--------------------------------------------|---------------------------|
| GOT1          | Aspartate aminotransferase, cytoplasmic     | Cytoplasm                 |
| CTSH          | Pro-cathepsin H                            | Lysosome                  |
| PEBP1         | Phosphatidylethanolamine-binding protein 1  | Cytoplasm                 |
| CTZ2          | Cathepsin Z                                | Lysosome                  |
| AKT1          | RAC-alpha serine/threonine-protein kinase   | Nucleosome/cytoplasm       |
| GALNS         | N-acetylglactosamine-6-sulfatase            | Lysosome                  |
| BHMT1         | Betaine-homocysteine S-methyltransferase 1 | Cytoplasm                 |
| LAMP1         | Lysosome-associated membrane glycoprotein 1| Lysosome/endosome          |
| Hexas        | Beta-hexosaminidase subunit alpha           | Lysosome                  |

Table 1. Multiplex panel of proteins listed with their corresponding gene names and cellular location

| Gene   | CLN5 and CLN6 sheep (label free proteomics)                        |
|--------|---------------------------------------------------------------------|
| CTS1   | Cathepsin L1*                                                       | Lysosome                  |
| TPP1   | Tripeptidyl-peptidase 1*                                            | Lysosome                  |
| PAM    | Peptidyl-glycine alpha-amidating monoxygenase*                      | Secreted                  |
| GM2A   | Ganglioside GM2 activator*                                          | Secreted                  |
| NAGA   | Alpha-N-acetylgalactosaminidase*                                    | Lysosome                  |

| Gene   | CLN6 sheep (label-free proteomics)                                |
|--------|-------------------------------------------------------------------|
| CTSB   | Cathepsin B                                                       | Lysosome                  |

The list is subdivided by the origin of the proteins, i.e., if they were selected from the label-free proteomics study, found in literature (* Sleat et al., 2017; #Doll et al., 2017; /C14 Sjödin et al., 2019) or from the label-free proteomics study of the urine of CLN5 and CLN6 sheep. The gene names labeled in bold were reliably detected in human urine.
Proteins affected in NCL: Figure 5A shows the comparison of all NCL patient samples as a disease group with controls. HEXA, LAMP1, TPP1, and GOT1 are significantly altered (p < 0.05). LAMP1 and HEXA are increased 3-fold on average, and GOT1 is elevated 6-fold. TPP1, although the affected protein in CLN2 disease, is also reduced in all other NCL forms relative to controls. The CLN7 patient sample, as indicated in the multivariate analysis, did not show many protein changes apart from reduced TPP1.

DISCUSSION

Using an animal model and human patient samples, we have employed a two-phase proteomic discovery and validation strategy to identify and confirm urinary proteins altered in patients with NCL. Urine is an

Figure 2. Multivariate analysis of targeted proteomics data of NCL patient urine

Principle-component analysis of all patient (except from one outlier, namely, CLN2(4)-post) and control samples using SIMCA presents two almost separate clusters to distinguish between NCL and control samples.

(A and B) (A) Score plot of the samples with the control samples cluster highlighted. Samples outside the control cluster are also highlighted. (B) Loading plot of the measured proteins indicating which proteins are driving the variation/change in the score plot.
easily accessible tissue for monitoring possible biomarkers for disease progression, and treatment and is a valuable tissue resource as the kidneys present a high proportion of mitochondria and lysosomes, both known to be affected in NCL disease.

The proteomic profiling analyses were performed on samples from two patients with CLN2 disease. Proteins observed to be altered included many proteins associated with the lysosome and is consistent with CLN2 being caused by a deficiency in a lysosomal enzyme involved in protein degradation. Other pathways also potentially affected in CLN2 urine include signaling pathways, e.g., those involved in inactivation of BAG2 signaling and inflammatory response-associated pathways. The BAG2 signaling pathway, which
includes heat shock proteins involved in the control of folding or degradation of misfolded proteins, are typically either increased or decreased in neurodegenerative disorders and implicated in the pathogenesis of various neurodegenerative diseases (Soto and Pritzcow, 2018). Many neurodegenerative diseases are
Figure 5. Proteins affected across multiple NCL forms

For the box and whisker plot: error bars are the 95% confidence interval, the bottom and top of the box are the 25th and 75th percentiles, the line inside the box is the 50th percentile (median), and any outliers are shown as open circles.
primarily disorders of protein misfolding and accumulation of storage material. Protein misfolding would be expected to be affected in CLN2 disease due to the lysosomal accumulation of mitochondrial ATP-synthase subunit-c, which has been shown to be stored in lysosome-derived organelles in many but not all NCL, including CLN2, 5, 6 diseases (Palmer, 2015; Chen et al., 2004; Palmer et al., 1989, 1992). Protein misfolding is clearly a pathological pathway in the brain and other tissues that we can observe in patient urine, and is consistent with the systemic pathology of CLN2 disease that affects nearly all cell types in the body. Other potentially affected pathways also include integrin-like kinase (ILK) signaling, which promotes cell growth. Altered cell growth has previously been described in NCL disease models (Guo et al., 1999; Huber et al., 2014; Mahmood et al., 2013; Savchenko et al., 2017; Sleat et al., 2017). Other pathways that may be affected involve inflammation reflective of neuroinflammation, which is common to the ovine diseases and several other NCLs (Palmer et al., 2013; Kay et al., 2006; Oswald et al., 2005; Nelvagal et al., 2020), that is known to occur due to stress of the ER from the unfolded protein response (Marotta et al., 2017). Proteins affected in patients with CLN2 disease included FMO3, which is likely to be present due to its role in metabolism of administered drugs. Phosphatidylethanolamine-binding protein 1 (PEBP1) is a protease inhibitor with multiple functions; why it is increased in CLN2 is unknown. GALNS N-acetylgalactosamine-6-sulfatase and beta-hexosaminidase alpha are part of the glycosaminoglycan degradation pathway, which has been previously implicated in NCL disease (Sleat et al., 2019).

Subsequent targeted analysis was performed on a larger cohort of samples from patients with CLN2 disease undergoing ERT treatment (n = 7, for two patients pre- and post-treatment paired samples were available) and a small number of other patients with NCL including two patients with CLN1, one with CLN3, two with CLN5, two with CLN6, and one with CLN7 disease. We compared data of the NCL disease group, involving all samples, and that of the CLN2 disease group with controls. Proteins that were found specific for CLN2 disease included BHMT1 as shown in Figures 3A and 5B. BHMT1 is a zinc-dependent cytosolic enzyme that catalyzes the transfer of a methyl group from betaine to homocysteine. This reaction results in increased dimethylglycine and methionine (Teng et al., 2011). High concentrations of BHMT may indicate higher production of the substrate homocysteine. High concentrations of homocysteine in plasma are known to occur due to stress of the ER from the unfolded protein response (Marotta et al., 2017). Proteins affected in patients with CLN2 disease included FMO3, which is likely to be present due to its role in metabolism of administered drugs. Phosphatidylethanolamine-binding protein 1 (PEBP1) is a protease inhibitor with multiple functions; why it is increased in CLN2 is unknown. GALNS N-acetylgalactosamine-6-sulfatase and beta-hexosaminidase alpha are part of the glycosaminoglycan degradation pathway, which has been previously implicated in NCL disease (Sleat et al., 2019).

The lack of TPP1 in the brain is known to cause neurological disease, but the deficiency of TPP1 outside the CNS could also indicate a wider functional impairment. This may be associated with the progression of cardiac pathology suggested to occur in some NCL forms, including CLN2 disease (Fukumura et al., 2012; Katz et al., 2017; Østergaard et al., 2011). It is assumed that upon administration of ERT directly into the CSF the neurological burden is relieved, and a longer lifespan becomes possible, which may result in pathology developing in the untreated organs becoming apparent (Katz et al., 2017). Treatments that target not just the CNS will likely also be needed, and the ability to monitor urinary TPP1 could be of benefit in this scenario. Another protein significantly increased in urine of patients with CLN2 disease is LAMP1, but this is not specific to CLN2. This lysosomal-associated membrane protein is situated in a dynamic equilibrium between lysosomes, endosomes, and the plasma membrane and is partially responsible for maintaining lysosomal integrity, pH, and catabolism (Xicoy et al., 2019; Eskelinen, 2006). LAMP1 is a widely used biomarker for lysosomes and lysosomal storage disorders but is usually measured in plasma or cells (Meikle et al., 1997). The elevated levels of LAMP1 in urine of patients with NCL may indicate the higher burden on the lysosome. Other urinary proteins increased in NCL disorders include GOT1 and HEXA. HEXA is increased in most patients with NCL; however, it is not increased sufficiently in patients with CLN2 disease to be significant when compared with the corresponding controls. HEXA is part of the glycosphingolipid degradation pathway (Breiden and Sandhoff, 2020; Hepbildikler et al., 2002). Specific deficiency of HEXA results in the lysosomal storage disorder called Tay Sachs disease (Okada and O’Brien, 1969). This observation of HEXA in urine here confirms a previous study that showed that the enzymatic activity of HEXA in CSF from patients with CLN2 and CLN3 disease is significantly increased (Sleat et al., 2017).
HEXA was also found to be altered in the Dictyostelium NCL model of CLN3 disease, indicating this enzyme is affected across different species model systems and may be a suitable biomarker candidate. Like HEXA, CLN5 has been described to also be a glycolytic hydrolase (Huber and Mathavarajah, 2018), indicating this particular molecular function is relevant in NCL disease pathology.

The fact that HEXA is increased but most other lysosomal enzymes are not indicates that it may not derive from lysosomes of disrupted cells. The HEXA could come from mast cells as it is a known marker of mast cell degranulation (Moon et al., 2014). Mast cell activity has not been specifically assessed in NCL, but mast cells have been observed to surround the optic nerve in a mouse model for CLN3 disease indicating compromise of the blood-brain barrier (Sappington et al., 2003). Proteomic analysis of CLN1 and CLN2 mouse model brain tissue has shown increased levels of CD63 (Sleat et al., 2019), which is thought to play a role in activation of mast cells and basophils (Kabashima et al., 2018). Therefore it is possible that the presence of HEXA is indicative of a mast cell-associated immune response in NCL.

Glutamic-oxaloacetic transaminase 1 is a cytosolic enzyme that plays a role in amino acid metabolism and the urea and tricarboxylic acid cycles. In principle GOT1 is a bidirectional enzyme, which turns it into an important regulator of glutamate concentration, the major excitatory neurotransmitter of the CNS, and also of aspartic acid levels. The form seen in urine is likely derived from kidney and plasma, which itself contains GOT1 from liver and muscle. Clinically GOT1 in plasma is routinely measured as aspartate aminotransferase 1 (ALT) and is a known marker of tissue damage. Increased plasma ALT has been observed in a canine model of CLN2 disease undergoing intracerebroventricular gene therapy (Katz et al., 2017), which also noted that treated dogs went on to develop cardiac disease, which is the likely source of the raised ALT. The elevated GOT1/ALT seen in NCL is probably due to tissue damage. It is plausible that this could be from the heart as progressive cardiac disease has been reported in children and adults with NCL (Ostergaard et al., 2011; Katz et al., 2017; Fukumura et al., 2012).

We considered the specificity of these proteins as biomarkers for NCL or Lysosomal Storage Disorders (LSDs). LAMPI is already well documented as a general marker of lysosome dysfunction and lysosomal disease (Meikle et al., 1997). We have profiled urine from patients with mucopolysaccharidosis and Niemann Pick C disease (unpublished data, Table S9). GOT1 was not affected, but interestingly, levels of BHMT are significantly reduced in NPC patient urine, the opposite of that observed in CLN2 disease. Why this occurs is unknown, but it does indicate that methionine metabolism is perturbed in LSDs.

ALT (GOT1) and HEXA are routinely measured in the clinical setting to monitor tissue damage and mast cell degranulation, so it would be possible to include these proteins when monitoring the effect of future novel systemic treatments. In summary, we have identified BHMT1 as a new biomarker in CLN2 disease and GOT1 (ALT) for systemic NCL disease as well as the known marker LAMPI. Furthermore, although we were unable to demonstrate that biomarkers in the sheep CLN5 and CLN6 models can translate to humans, we have confirmed the involvement of HEXA as observed in previous studies on the eukaryotic model system Dictyostelium (Huber et al., 2014).

Limitations of the study
The scarcity of available patient samples due to these disorders being rare means it will always be difficult to ascertain statistical significance in such studies. The targeted analysis unfortunately did not confirm the proteins detected in the sheep as altered in human CLN5 and CLN6 disease samples. CTSB, alpha-N-acetylgalactosaminidase (NAGA), and RAC-alpha serine/threonine-protein kinase (AKT1) could not be detected in the human urine. Animal models can be very useful for understanding disease pathology and designing treatments, but their application to identify biomarkers in humans as in this scenario is limited so far.

Resource availability
Lead contact
Wendy E Heywood wendy.heywood@ucl.ac.uk.

Materials availability
This study did not generate new unique reagents.
Data and code availability
Original data have been deposited to PeptideAtlas at http://www.peptideatlas.org/PASS/PASS01636.

METHODS
All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.isci.2020.102020.

ACKNOWLEDGMENTS
The authors would like to thank the Batten Disease Support and Research Association (BDSRA), USA, Noah’s Hope/Hope for Bridget and Drew’s Hope for funding this project. Medical Research Council (Grant No MR/R025134/1) supported research (S.E.M.). All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Center. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. We would also like to thank Martin Wellby and Dr. Katharina Russell, Lincoln University, for collecting the ovine samples.

AUTHOR CONTRIBUTIONS
K.I. created the targeted assay, performed the targeted analysis, and wrote the manuscript. R.C. performed the sheep urine analysis. P.M. enabled provision of patient samples, study design, and concept. B.C. provided patient samples. P.G. enabled provision of patient samples, study design, and concept. S.E.M., study design and concept and expertise. D.N.P. provided sheep model samples. K.M., study design and concept and resources. W.E. H., funding, concept, study design, and wrote manuscript. All authors contributed to manuscript writing and revision.

DECLARATION OF INTERESTS
P.G. is an investigator on BioMarin-sponsored studies and has received research grants and speaker honoraria. S.E.M. receives financial support from BioMarin Pharmaceutical Inc. to maintain the NCL Mutation Database. W.E. H. has received a travel award from BioMarin Pharmaceutical Inc. The remaining authors declare no competing interests.

REFERENCES
Adams, H.R., and Mink, J.W. (2013). Neurobehavioral features and natural history of juvenile neuronal ceroid lipofuscinosis (batten disease). J. Child Neurol. 28, 1128–1136.

Breiden, B., and Sandhoff, K. (2020). Mechanism of secondary ganglioside and lipid accumulation in lysosomal disease. Int. J. Mol. Sci. 21, 2566.

Butz, E.S., Chandrachud, U., Mole, S.E., and Cotman, S.L. (2020). Moving towards a new era of genomics in the neuronal ceroid lipofuscinoses. Biochim. Biophys. Acta Mol. Basis Dis. 1866, 165571.

Chen, R., Fearnley, I.M., Palmer, D.N., and Walker, J.E. (2004). Lysine 43 is trimethylated in subunit C from bovine mitochondrial ATP synthase and in storage bodies associated with batten disease. J. Biol. Chem. 279, 21883–21887.

de Blieck, E.A., Augustine, E.F., Marshall, F.J., Adams, H., Cialone, J., Dure, L., Kwon, J.M., Newhouse, N., Rose, K., Rothberg, P.G., et al. (2013). Methodology of clinical research in rare diseases: development of a research program in juvenile neuronal ceroid lipofuscinosis (JNCL) via creation of a patient registry and collaboration with patient advocates. Contemp. Clin. Trials 35, 48–54.

di Ronza, A., Bajaj, L., Sharma, J., Sanagasetti, D., Lotfi, P., Adamski, C.J., Collette, J., Palmieri, M., Amawi, A., Popp, L., et al. (2018). CLN8 is an endoplasmic reticulum cargo receptor that regulates lysosome biogenesis. Nat. Cell Biol. 20, 1370–1377.

Doll, S., Dreßlen, M., Geyer, P.E., Itzhak, D.N., Braun, C., Doppler, S.A., Meier, F., Deutsch, M.-A., Lahm, H., Lange, R., et al. (2017). Region and cell-type resolved quantitative proteomic map of the human heart. Nat. Commun. 8, 1469.

Eskelinen, E.-L. (2006). Roles of LAMP-1 and LAMP-2 in lysosome biogenesis and autophagy. Mol. Aspects Med. 27, 495-502.

Fukumura, S., Saito, Y., Saito, T., Komaki, H., Nakagawa, E., Sugai, K., Sasaki, M., Oka, A., and Takamisawa, I. (2012). Progressive conduction defects and cardiac death in late infantile neuronal ceroid lipofuscinosis. Dev. Med. Child Neurol. 54, 663-666.

Golabek, A.A., Kida, E., Walus, M., Wu, P., Mehria, P., and Wasiowski, K.E. (2003). Biosynthesis, glycosylation, and enzymatic processing in vivo of human tripeptidyl-peptidase I. J. Biol. Chem. 278, 7135–7145.

Griffey, M., Macauley, S.L., Ogilvie, J.M., and Sands, M.S. (2005). AAV2-mediated ocular gene therapy for infantile neuronal ceroid lipofuscinosis. Mol. Ther. 12, 413–421.
Mammalian models and the social amoeba

neuronal ceroid lipofuscinoses: insights from

Huber, R.J. (2020). Molecular networking in the

toward the comprehensive functional analysis of

(2008). Bioinformatics enrichment tools: paths

Loss of Cln3 function in the social amoeba

Huber, R.J., Myre, M.A., and Cotman, S.L. (2014).

42

in Dictyostelium. Cell Signal.

Huber, R.J., and Mathavarajah, S. (2018). Cln5 is

attenuates disease progression in a canine model

of late-infantile neuronal ceroid lipofuscinosis (CLN2 disease). J. Neurosci. Res. 92, 1591–1598.

Katz, M.L., Johnson, G.C., Leach, S.B., Williamson, B.G., Coates, J.R., Whiting, R.E.H., Vaništekníkství, D.P., and Whitney, M.S. (2017). Extraneuronal pathology in a canine model of CLN2 neuronal ceroid lipofuscinosis after intracerebroventricular gene therapy that delays neurological disease progression. Gene Ther. 24, 215–223.

Kay, G.W., Palmer, D.N., Rezaie, P., and Cooper, J.D. (2006). Activation of non-neuronal cells within the prenatally developing brain of sheep with neuronal ceroid lipofuscinosis. Brain Pathol. 16, 110–116.

Kohlschutter, A., Schulz, A., Bartsch, U., and Storch, S. (2019). Current and emerging treatment strategies for neuronal ceroid lipofuscinoses. CNS Drugs 33, 315–325.

Liu, W., Kleine-Holthaus, S.M., Herranz-Martín, S., Arstotzka, M., Mole, S.E., Smith, A.J., Al, R.R., and Rahim, A.A. (2020). Experimental gene therapies for the NCLs. Biochim. Biophys. Acta. Mol. Basis Dis. 1866, 165772.

Lu, J.-Y., Nelvalag, H.R., Wang, L., Binbaum, S.G., Cooper, J.D., and Hofmann, S.L. (2015). Intrahepatic enzyme replacement therapy improves motor function and survival in a preclinical mouse model of infantile neuronal ceroid lipofuscinosis. Mol. Genet. Metab. 116, 98–105.

Mahmood, F., Fu, S., Cooke, J., Wilson, S.W., Cooper, J.D., and Russell, C. (2013). A zebrafish model of CLN2 disease is deficient in tripeptidyl peptidase 1 and displays progressive neurodegeneration accompanied by a reduction in proliferation. Brain 138, 1488–1507.

Marotta, D., Tinelli, E., and Mole, S.E. (2017). NCLs and ER: a stressful relationship. Biochim. Biophys. Acta. Mol. Basis Dis. 1863, 1273–1281.

Meike, P.J., Brooks, D.A., Ravencroft, E.M., Yan, M., Williams, R.E., Jaunzems, A.E., Chataway, T.K., Karageorgos, L.E., Davel, R.C., Boulter, C.D., et al. (1997). Diagnosis of lysosomal storage disorders: evaluation of lysosome-associated membrane protein LAMP-1 as a diagnostic marker. Clin. Chem. 43, 1235–1335.

Mole, S.E., Anderson, G., Band, H.A., Berkovic, S.F., Cooper, J.D., Kleine Holthaus, S.M., McKay, T.R., Medina, D.L., Rahim, A.A., Schulz, A., and Smith, A.J. (2019). Clinical challenges and future therapeutic approaches for neuronal ceroid lipofuscinosis. Lancet Neurol. 18, 107–116.

Mole, S.E., Michaux, G., Codlin, S., Wheeler, R.B., Sharp, J.D., and Cutler, D.F. (2004). CLN6, which is associated with a lysosomal storage disease, is involved in the secretory pathways involved. Front. Immunol. 5, 82.

Okada, S., and O’Brien, J.S. (1969). Tay-Sachs disease: generalized absence of a beta-D-N-acetyhexosaminidase component. Science 165, 698–700.

Ostergaard, J.R., Rasmussen, T.B., and Molgaard, H. (2011). Cardiac involvement in juvenile neuronal ceroid lipofuscinosis (Batten disease). Neurology 76, 1245–1251.

Oswald, M.J., Palmer, D.N., Kay, G.W., Shemilt, S.J., Rezaie, P., and Cooper, J.D. (2005). Glial activation spreads from specific cerebral foci and precedes neurodegeneration in presymptomatic ovine neuronal ceroid lipofuscinosis (CLN6). Neurobiol. Dis. 20, 49–63.

Palmer, D.N. (2015). The relevance of the storage of subunit c of ATP synthase in different forms and models of Batten disease (NCLs). Biochim. Biophys. Acta. 1852, 2287–2291.

Palmer, D.N., Barry, L.A., Tynjela, J., and Cooper, J.D. (2013). NCL disease mechanisms. Biochim. Biophys. Acta. 1832, 1882–1893.

Palmer, D.N., Feenley, I.M., Meded, S.M., Walker, J.E., Martinus, R.D., Bayliss, S.L., Hall, N.A., Lake, B.D., Wolfe, L.S., and Jolly, R.D. (1989). Lysosomal storage of the DCDD reactive protopсид subunit of mitochondrial ATP synthase in human and ovine ceroid lipofuscinoses. Adv. Exp. Med. Biol. 266, 211–222, discussion 223.

Palmer, D.N., Feenley, I.M., Walker, J.E., Hall, N.A., Lake, B.D., Wolfe, L.S., Haltia, M., Martinus, R.D., and Jolly, R.D. (1992). Mitochondrial ATP synthase subunit c storage in the cerebral lipofuscinoses (Batten disease). Am. J. Med. Genet. 42, 561–567.

Sappington, R.M., Pearce, D.A., and Calkins, D.J. (2003). Optic nerve degeneration in a murine model of juvenile ceroid lipofuscinosis. Invest. OphthalmoL Vis. Sci. 44, 3725–3731.

Savchenko, E., Singh, Y., Kontinen, H., Lejava, K., Medivilla Santos, L., Grubman, A., Karkkainen, V., Keksa-Goldsteine, V., Naumenko, N., Tavi, P., et al. (2017). Loss of CIN causes altered neurogenesis in a mouse model of a childhood neurodegenerative disorder. Dis. Model. Mech. 10, 1089–1100.

Schulz, A., Ajayi, T., Specchio, N., de Los Reyes, E., Gissen, P., Ballon, D., Dyke, J.P., Cahan, H., Slator, P., Jacoby, D., et al. (2018). Study of intraventricular cerilipidina for CLN2 disease. N. Engl. J. Med. 379, 1896–1907.

Sindelar, M., Dyke, J.P., Deeb, R.S., Sondhi, D., Kaminsky, S.M., Kosofsky, E.B., Ballon, D.J., Crystal, R.G., and Gross, S.S. (2018). Untargeted metabolite profiling of cerebrospinal fluid uncovers biomarkers for severity of late infantile neuronal ceroid lipofuscinosis (CLN2, batten disease). Sci. Rep. 8, 15229.

Sjödin, S., Brinkmalm, G., Ohrfelt, A., Parnetti, L., Paciotti, S., Hansson, O., Hardy, J., Blennow, K., Zetterberg, H., and Brinkmalm, A. (2019). Endolysosomal proteins and ubiquitin CSF concentrations in Alzheimer’s and Parkinson’s disease. Alzheimers Res. Ther. 11, 82.

Sleaf, D.E., Tannous, A., Sohar, I., Wiseman, J.A., Zheng, H., Gian, M., Zhao, C., Xin, W., Barone, R., Sims, K.B., et al. (2017). Proteomic analysis of brain and cerebrospinal fluid from the three...
major forms of neuronal ceroid lipofuscinosis reveals potential biomarkers. J. Proteome Res. 16, 3787–3804.

Slet, D.E., Wiseman, J.A., El-Banna, M., Zheng, H., Zhao, C., Soherwardy, A., Moore, D.F., and Lobel, P. (2019). Analysis of brain and cerebrospinal fluid from mouse models of the three major forms of neuronal ceroid lipofuscinosis reveals changes in the lysosomal proteome. Mol. Cell Proteomics 18, 2244–2261.

Soto, C., and Pritzkow, S. (2018). Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. Nat. Neurosci. 21, 1332–1340.

Staropoli, J.F., Karaa, A., Lim, E.T., Kirby, A., Elbalalesy, N., Romansky, S.G., Leydiker, K.B., Coppel, S.H., Barone, R., Xin, W., et al. (2012). A homozygous mutation in KCTD7 links neuronal ceroid lipofuscinosis to the ubiquitin-proteasome system. Am. J. Hum. Genet. 91, 202–208.

Teng, Y.-W., Mehedint, M.G., Garrow, T.A., and Zeisel, S.H. (2011). Deletion of betaine-homocysteine S-methyltransferase in mice perturbs choline and 1-carbon metabolism, resulting in fatty liver and hepatocellular carcinomas. J. Biol. Chem. 286, 36258–36267.

Tracy, C.J., Whiting, R.E.H., Pearce, J.W., Williamson, B.G., Vansteenkiste, D.P., Gillespie, L.E., Castaner, L.J., Bryan, J.N., Coates, J.R., Jensen, C.A., and Katz, M.L. (2016). Intravitreal implantation of TPP1-transduced stem cells delays retinal degeneration in canine CLN2 neuronal ceroid lipofuscinosis. Exp. Eye Res. 152, 77–87.

Whiting, R.E.H., Nafirstrom, K., Yao, G., Pearce, J.W., Coates, J.R., Castaner, L.J., Jensen, C.A., Dougherty, B.N., Vuillemenot, B.R., Kennedy, D., et al. (2014). Enzyme replacement therapy delays pupillary light reflex deficits in a canine model of late infantile neuronal ceroid lipofuscinosis. Exp. Eye Res. 125, 164–172.

Williams, R.E., and Mole, S.E. (2012). New nomenclature and classification scheme for the neuronal ceroid lipofuscinoses. Neurology 79, 183–191.

Wlodawer, A., Durell, S.R., Li, M., Oyama, H., Oda, K., and Dunn, B.M. (2003). A model of tripeptidyl-peptidase I (CLN2), a ubiquitous and highly conserved member of the sedolisin family of serine-carboxyl peptidases. BMC Struct. Biol. 3, 8.

Xicoy, H., Peñuelas, N., Vila, M., and Laguna, A. (2019). Autophagic- and lysosomal-related biomarkers for Parkinson’s disease: lights and shadows. Cells 8, 1317.
Supplemental Information

Urine proteomics analysis of patients with neuronal ceroid lipofuscinoses

Katharina Iwan, Robert Clayton, Philippa Mills, Barbara Csanyi, Paul Gissen, Sara E. Mole, David N. Palmer, Kevin Mills, and Wendy E. Heywood
Supplemental Figure S1: Significant pathways affected in urine CLN5 and CN6 affected sheep. Top panel IPA of CLN5. Bottom panel IPA of CLN6. Related to figure 4A.
Supplemental Figure 2: All calibration curves for the peptides of interest that were valid for quantification are shown. Related to figures 2-5.
| Gene Name | Peptide count | Unique peptides | Confidence score | Anova (p) | fold change |
|-----------|--------------|----------------|-----------------|-----------|-------------|
| GOT1      | 2            | 2              | 9.6             | 1.97E-05  | 5129.65     |
| FMOD      | 2            | 2              | 7.3             | 6.95E-03  | 2060.44     |
| PEBP1     | 3            | 2              | 5.28            | 3.37E-03  | 69.47       |
| GAK       | 2            | 2              | 10.3            | 7.76E-04  | 62.99       |
| NBPF8     | 2            | 2              | 7.8             | 1.19E-03  | 43.89       |
| WDR19     | 2            | 2              | 10.2            | 1.30E-02  | 27.89       |
| BMK6      | 7            | 6              | 58.1            | 1.33E-02  | 27.68       |
| IGAV3-49  | 2            | 2              | 13.4            | 1.21E-02  | 25.61       |
| HEXA      | 7            | 6              | 32.7            | 1.01E-03  | 22.32       |
| CTHH      | 4            | 4              | 16.3            | 4.64E-04  | 20.00       |
| HSPA1A    | 3            | 3              | 16.5            | 9.39E-03  | 18.74       |
| CTSS      | 2            | 2              | 8.2             | 4.73E-03  | 18.17       |
| TP1       | 2            | 2              | 38.6            | 2.04E-02  | 15.26       |
| IG1       | 3            | 3              | 23.4            | 2.13E-02  | 13.13       |
| PPIA      | 4            | 4              | 26.2            | 1.59E-02  | 12.57       |
| ALDOB     | 7            | 7              | 47.6            | 4.33E-03  | 11.45       |
| COL6A1    | 17           | 16             | 133.8           | 2.40E-03  | 11.30       |
| HSPA8     | 6            | 6              | 37.8            | 1.20E-02  | 10.31       |
| KRTFL7    | 11           | 3              | 71.4            | 4.47E-03  | 9.04        |
| SALL3     | 3            | 3              | 10.9            | 1.88E-02  | 9.12        |
| NSFL1C    | 3            | 2              | 18.2            | 2.47E-02  | 9.12        |
| GDF5      | 2            | 2              | 15.7            | 2.69E-02  | 8.94        |
| ANK2A     | 3            | 2              | 5.9             | 1.80E-02  | 7.37        |
| PRCP      | 5            | 5              | 32.4            | 9.95E-03  | 7.35        |
| SERPINA6  | 4            | 3              | 30.2            | 4.73E-02  | 7.35        |
| PGM1      | 5            | 4              | 30.1            | 4.73E-02  | 7.35        |
| LAMPI     | 5            | 4              | 35.2            | 7.33E-03  | 6.21        |
| NAGLU     | 4            | 4              | 18.3            | 7.22E-02  | 6.04        |
| SAE       | 5            | 4              | 29.6            | 2.83E-02  | 5.81        |
| RECO5L    | 7            | 7              | 23.7            | 4.62E-02  | 5.69        |
| ICAM1     | 3            | 3              | 19.2            | 4.87E-03  | 5.63        |
| DDB1      | 4            | 4              | 20.6            | 4.20E-02  | 5.56        |
| GUSB      | 7            | 7              | 49.1            | 3.49E-02  | 5.33        |
| ACP2      | 8            | 7              | 60.0            | 1.82E-02  | 5.32        |
| IGA2      | 5            | 1              | 48.9            | 6.69E-03  | 5.30        |
| PEPIB     | 3            | 3              | 14.6            | 2.11E-03  | 5.20        |
| DAG1      | 3            | 3              | 31.1            | 1.66E-03  | 5.13        |
| C3SP64    | 4            | 2              | 20.4            | 3.95E-02  | 5.10        |
| G3PTP1    | 4            | 4              | 26.1            | 3.23E-04  | 4.53        |
| SCDCB     | 2            | 2              | 36.5            | 1.89E-02  | 3.96        |
| SOS3      | 4            | 4              | 29.7            | 4.08E-02  | 3.95        |
| GP6       | 4            | 4              | 25.8            | 4.42E-02  | 3.72        |
| PRO2      | 9            | 8              | 76.2            | 1.85E-02  | 3.69        |
| LCN2      | 5            | 5              | 34.3            | 4.12E-02  | 3.55        |
| GLB1      | 7            | 7              | 45.6            | 9.02E-02  | 3.31        |
| RHNR1L2   | 3            | 3              | 28.4            | 2.65E-03  | 3.18        |
| CLEC4C    | 5            | 4              | 25.0            | 3.35E-02  | 3.16        |
| HSPA2A    | 4            | 4              | 27.6            | 1.14E-02  | 3.13        |
| ACTB      | 8            | 2              | 59.8            | 2.57E-03  | 2.78        |
| CEACAM8   | 2            | 2              | 21.8            | 3.12E-02  | 2.67        |
| TIGAM2    | 2            | 2              | 8.3             | 1.82E-02  | 2.66        |
| ANK45     | 2            | 2              | 10.9            | 3.20E-02  | 2.55        |
| CREB1L3   | 3            | 3              | 17.8            | 4.87E-02  | 2.55        |
| DLK2      | 3            | 3              | 15.8            | 4.99E-02  | 2.54        |
| ACTC1     | 6            | 6              | 41.0            | 4.12E-02  | 2.47        |
| EPHB2     | 8            | 7              | 44.6            | 4.75E-02  | 2.42        |
| HAVCR2    | 8            | 8              | 78.6            | 8.50E-03  | 2.39        |
| MMRN2     | 8            | 7              | 50.7            | 3.12E-02  | 2.33        |

Supplemental Table S1: Untargeted proteomics analysis of 2 CLN2 vs. 4 control samples Related to figure 1.
**Supplemental Table S2**: Significant proteins from untargeted proteomics analysis of \( n=3 \) CLN5 sheep vs. 8 control sheep urine samples (\( p \)-value (Welch test after transformation) < 0.05 and normalised fold change > 1). Related to figure 1.

| Gene       | Name                                           | Species       | Ranking | Unique peptide | Confidence score | \( p \) value | Fold change |
|------------|------------------------------------------------|---------------|---------|----------------|------------------|-------------|-------------|
| ENO3       | Beta-enolase                                   | Homo sapiens  | A       | 3              | 63.35            | 0.0178      | 2.3578      |
| FANCA      | Fanconi anemia group A protein                 | Homo sapiens  | B       | 2              | 8.00             | 0.0155      | 1.7416      |
| uncharacterised | Uncharacterized protein                       | Ovis aries   | B       | 1              | 51.31            | 0.0077      | 1.4584      |
| B3SP       | Integrin binding sialoprotein                 | Ovis aries   | A       | 5              | 56.62            | 0.0361      | 1.2218      |
| RB8P8      | DNA endonuclease RB8P8                        | Homo sapiens  | A       | 7              | 42.81            | 0.0358      | 1.1808      |
| ILDH1A1    | Retinal dehydrogenase 1                       | Homo sapiens  | A       | 2              | 19.19            | 0.0456      | 1.1206      |
| LAMP2      | Lysosomal associated membrane protein 2       | Ovis aries   | A       | 9              | 48.28            | 0.0279      | 1.0842      |
| DS1C2      | Desmocollin 2                                 | Ovis aries   | A       | 7              | 74.71            | 0.0172      | 1.0737      |
| CTNNA1     | Catenin alpha-1                               | Homo sapiens  | A       | 4              | 30.50            | 0.0478      | 1.0644      |
| COX41      | Cytochrome c oxidase subunit 4I1              | Ovis aries   | B       | 2              | 11.44            | 0.0186      | 1.0620      |
| VWA3B      | von Willebrand factor A domain-containing protein 3B | Homo sapiens  | B       | 2              | 9.10             | 0.0010      | 1.0494      |
| SERPINF1   | Serpin family F member 1                      | Ovis aries   | A       | 4              | 43.93            | 0.0156      | 1.0422      |
| uncharacterised | Uncharacterized protein                       | Ovis aries   | A       | 19             | 188.99           | 0.0114      | -1.0398     |
| CD93       | CD93 molecule                                  | Ovis aries   | A       | 22             | 159.51           | 0.0415      | -1.0627     |
| ORM1       | Alpha-1-acid glycoprotein                     | Ovis aries   | A       | 22             | 143.97           | 0.0190      | -1.0718     |
| SCGN       | Secretagogin, EF-hand calcium binding protein | Ovis aries   | A       | 2              | 27.50            | 0.0451      | -1.0799     |
| NECTIN1    | Nectin cell adhesion molecule 1               | Ovis aries   | B       | 2              | 10.80            | 0.0448      | -1.1338     |
| GS5S       | Immunoglobulin superfamily member 5          | Ovis aries   | A       | 14             | 99.15            | 0.0169      | -1.2003     |
| LEAP2      | Liver enriched antimicrobial peptide 2        | Ovis aries   | A       | 4              | 19.97            | 0.0102      | -1.2378     |
| NEFL       | Neurofilament light polypeptide               | Homo sapiens  | A       | 6              | 62.09            | 0.0302      | -1.3340     |
### Supplemental Table S3: Significant proteins from untargeted proteomics analysis of n=13 CLN6 sheep vs. n=8 control sheep samples (p-value (Welch test after transformation) < 0.05 and normalised fold change > 1). Related to figure 1.

| Gene Name | Species   | Ranking | Unique peptides | Confidence score | p value | fold change |
|-----------|-----------|---------|----------------|------------------|---------|-------------|
| I4FR1C   | Ovis aries| A       | 2              | 18.14            | 0.0488  | 1.438       |
| CTSB     | Homo sapiens| A       | 2              | 36.15            | 0.0294  | 1.342       |
| GPR1A    | Ovis aries| A       | 2              | 23.78            | 0.0110  | 1.230       |
| C21orf69 | Ovis aries| A       | 2              | 11.20            | 0.0133  | 1.203       |
| CYP19A1  | Homo sapiens| A       | 4              | 23.80            | 0.0187  | 1.130       |
| TPRP1    | Homo sapiens| B       | 1              | 16.98            | 0.0093  | 1.101       |
| CASC5    | Homo sapiens| B       | 2              | 11.77            | 0.0097  | 1.094       |
| ATP5A1   | Ovis aries| B       | 1              | 31.92            | 0.0362  | 1.064       |
| MRPS51   | Ovis aries| B       | 2              | 10.02            | 0.0429  | 1.063       |
| C4BPA    | Ovis aries| A       | 3              | 22.65            | 0.0039  | 1.057       |
| ENA11    | Homo sapiens| A       | 2              | 15.86            | 0.0333  | 1.051       |
| ZFC2     | Ovis aries| A       | 7              | 14.71            | 0.0164  | 1.049       |
| LC4C     | Ovis aries| A       | 6              | 35.15            | 0.0198  | 1.045       |
| WASSA    | Ovis aries| B       | 2              | 9.10             | 0.0062  | 1.043       |
| APG2     | Ovis aries| A       | 3              | 17.96            | 0.0452  | 1.038       |
| PLA2G7   | Ovis aries| A       | 8              | 69.74            | 0.0437  | 1.033       |
| PI16     | Ovis aries| A       | 7              | 54.38            | 0.0367  | 1.038       |
| CATHL3   | Ovis aries| A       | 6              | 99.46            | 0.0185  | 1.038       |
| XPC     | Ovis aries| A       | 13             | 174.09           | 0.0168  | 1.039       |
| JNCL     | Ovis aries| A       | 19             | 188.99           | 0.0429  | 1.041       |
| CT6      | Ovis aries| B       | 46             | 235.83           | 0.0488  | 1.043       |
| NUCB1    | Ovis aries| A       | 22             | 196.46           | 0.0287  | 1.059       |
| IL1B     | Ovis aries| A       | 3              | 25.15            | 0.0538  | 1.073       |
| FABP4    | Ovis aries| A       | 16             | 164.25           | 0.0173  | 1.077       |
| LOC544331| Homo sapiens| A       | 9              | 82.03            | 0.0311  | 1.077       |
| LOC101112304| Homo sapiens| A       | 5              | 70.60            | 0.0299  | 1.079       |
| SFAP     | Ovis aries| A       | 5              | 70.60            | 0.0299  | 1.079       |
| SLC14A1  | Homo sapiens| A       | 12             | 104.23           | 0.0471  | 1.125       |
| AKAP13   | Ovis aries| A       | 10             | 139.58           | 0.0468  | 1.139       |
| NOTCH1   | Ovis aries| B       | 1              | 40.51            | 0.0459  | 1.152       |
| RAB6A-GEF| Homo sapiens| B       | 1              | 29.92            | 0.0227  | 1.161       |
| OXACTAN  | Ovis aries| A       | 13             | 116.55           | 0.0373  | 1.170       |
| CHD1     | Homo sapiens| A       | 2              | 25.33            | 0.0445  | 1.186       |
| ABHD11B  | Ovis aries| A       | 2              | 23.87            | 0.0141  | 1.188       |
| NOTCH1   | Ovis aries| B       | 15             | 163.84           | 0.0288  | 1.193       |
| TBC1D14  | Homo sapiens| B       | 2              | 10.45            | 0.0227  | 1.210       |
| TIGIT    | Homo sapiens| A       | 4              | 35.73            | 0.0042  | 1.210       |
| ZSP5    | Homo sapiens| B       | 2              | 10.80            | 0.0884  | 1.215       |
| LEAP2    | Ovis aries| A       | 4              | 19.97            | 0.0302  | 1.246       |
| BLG      | Ovis aries| A       | 4              | 48.55            | 0.0287  | 1.282       |
| NOTCH1   | Ovis aries| A       | 6              | 62.69            | 0.0292  | 1.356       |
| DNMAB    | Homo sapiens| B       | 2              | 10.39            | 0.0215  | 1.417       |
| CUB1     | Homo sapiens| A       | 3              | 21.81            | 0.0471  | 1.417       |
| NCF50    | Homo sapiens| B       | 3              | 15.71            | 0.0365  | 1.460       |
| NOTCH1   | Ovis aries| B       | 1              | 25.39            | 0.0005  | 1.537       |
| RIM5     | Ovis aries| A       | 2              | 16.56            | 0.0371  | 1.547       |
| DDDK     | Ovis aries| B       | 1              | 24.78            | 0.0067  | 1.591       |
| TIDEIP   | Ovis aries| B       | 2              | 16.02            | 0.0097  | 1.597       |
| NOTCH1   | Ovis aries| B       | 2              | 10.00            | 0.0086  | 1.674       |
| DDDK     | Ovis aries| B       | 1              | 66.78            | 0.0236  | 1.705       |
| ERICH6B  | Ovis aries| B       | 1              | 33.31            | 0.0004  | 3.645       |
**Supplemental Table S4: DAVID analysis of the target list, i.e. the significant proteins from the label-free proteomics analysis, and the background list, i.e. the non-significant proteins from the label-free proteomics analysis. Only intracellular components are listed that show a Benjamini value of < 0.05. The intracellular components highlighted in green were of interest and proteins located in that intracellular location were chosen for further targeted proteomics analysis. Related to figure 1**

| Intracellular Component | Target list | Count | %   | PValue   | List Total | Fold Enrichment | Benjamini FDR |
|-------------------------|-------------|-------|-----|----------|------------|----------------|---------------|
| lysosomal lumen         |             | 8     | 0.076694 | 0.0000 | 70 | 24.50 | 1.82E-06 | 3.90E-05 |
| sarcolemma              |             | 5     | 0.047934 | 0.0003 | 70 | 15.31 | 0.00736 | 0.368754 |
| blood microparticle     |             | 8     | 0.076694 | 0.0000 | 70 | 13.70 | 7.31E-05 | 0.002089 |
| melanosome              |             | 5     | 0.047934 | 0.0000 | 70 | 12.89 | 0.01102 | 0.709453 |
| lysosome                |             | 9     | 0.086281 | 0.0000 | 70 | 10.37 | 6.15E-05 | 0.002633 |
| membrane raft           |             | 6     | 0.057521 | 0.0011 | 70 | 7.58 | 0.01863 | 1.337883 |
| external side of plasma membrane | | 6 | 0.057521 | 0.0013 | 70 | 7.33 | 0.01963 | 1.545576 |
| focal adhesion          |             | 11    | 0.105455 | 0.0000 | 70 | 7.32 | 3.94E-05 | 0.002478 |
| extracellular space     |             | 29    | 0.278017 | 0.0000 | 70 | 5.61 | 1.20E-12 | 1.71E-11 |
| extracellular exosome   |             | 54    | 0.517688 | 0.0000 | 70 | 5.00 | 1.46E-27 | 1.04E-26 |

| Intracellular Component | Background list | Count | %   | PValue   | List Total | Fold Enrichment | Benjamini FDR |
|------------------------|-----------------|-------|-----|----------|------------|----------------|---------------|
| platelet dense granule lumen |             | 5     | 1.404494 | 1.16E-04 | 352 | 18.49 | 0.001579 | 0.158153 |
| platelet alpha granule lumen |             | 19    | 5.337079 | 6.26E-18 | 352 | 17.89 | 3.55E-16 | 8.53E-15 |
| blood microparticle     |             | 45    | 12.64045 | 1.25E-07 | 352 | 15.33 | 1.06E-37 | 1.70E-36 |
| anchored component of external side of plasma membrane | | 6 | 1.685393 | 5.18E-05 | 352 | 14.12 | 8.00E-04 | 0.070507 |
| lysosomal lumen         |             | 17    | 4.775281 | 6.02E-12 | 352 | 10.35 | 2.56E-10 | 1.70E-09 |
| extracellular vesicle   |             | 10    | 2.808989 | 4.11E-07 | 352 | 10.33 | 9.32E-06 | 5.60E-04 |
| extracellular matrix    |             | 41    | 11.51685 | 1.63E-22 | 352 | 10.33 | 9.32E-06 | 5.60E-04 |
| basement membrane       |             | 10    | 2.808989 | 2.07E-05 | 352 | 6.55 | 3.52E-04 | 0.02817 |
| melanosome              |             | 12    | 3.370787 | 3.92E-06 | 352 | 6.15 | 7.40E-05 | 0.005333 |
| anchored component of membrane |             | 13    | 3.651685 | 1.87E-06 | 352 | 5.96 | 3.73E-05 | 0.002541 |
| extracellular space     |             | 150   | 42.13483 | 3.24E-76 | 352 | 7.17 | 3.47E-06 | 5.26E-05 |
| vesicle                 |             | 14    | 3.892854 | 1.16E-06 | 352 | 5.56 | 2.47E-05 | 0.00158 |
| collagen trimer         |             | 10    | 2.808989 | 7.01E-05 | 352 | 5.63 | 0.001036 | 0.09548 |
| external side of plasma membrane | | 23    | 6.460674 | 1.63E-10 | 352 | 5.59 | 6.17E-09 | 2.23E-07 |
| extracellular exosome   |             | 280   | 78.65169 | 3.53E-16 | 352 | 5.16 | 1.20E-15 | 4.81E-15 |
| lysosome                |             | 22    | 6.179775 | 2.97E-09 | 352 | 5.04 | 1.01E-07 | 4.05E-06 |
| endoplasmic reticulum lumen |             | 16    | 4.494382 | 4.92E-06 | 352 | 4.31 | 8.80E-05 | 0.006694 |
| extracellular region    |             | 130   | 36.51685 | 3.73E-48 | 352 | 4.18 | 4.22E-46 | 5.08E-45 |
| proteinaceous extracellular matrix | | 21    | 5.898876 | 2.73E-07 | 352 | 4.06 | 6.63E-06 | 3.72E-04 |
| cell surface            |             | 41    | 11.51685 | 3.17E-13 | 352 | 3.92 | 1.54E-11 | 4.32E-10 |
| apical plasma membrane  |             | 22    | 6.179775 | 2.40E-07 | 352 | 3.91 | 6.28E-06 | 3.27E-04 |
| membrane raft           |             | 15    | 4.213483 | 4.82E-05 | 352 | 3.77 | 7.81E-04 | 0.065679 |
| focal adhesion          |             | 26    | 7.303171 | 1.74E-07 | 352 | 3.44 | 4.94E-06 | 2.37E-04 |
| integral component of plasma membrane | | 49    | 13.76404 | 8.76E-05 | 352 | 1.79 | 0.001241 | 0.119259 |
| plasma membrane         |             | 126   | 35.39326 | 1.63E-08 | 352 | 1.58 | 5.05E-07 | 2.22E-05 |
### Supplemental Table S5: Evaluation of the quality controls (QCs) made from pooled urine digests. Related to figures 3-5.

|       | GOT1  | TPP1   | LAMP2  | PAM    | CTSH   | LEAP   | GM2A   | CTSZ   |
|-------|-------|--------|--------|--------|--------|--------|--------|--------|
| QC1   | 194.068 | 1.015804 | 0.311667 | 0.0409633 | 0.491935 | 5.678913 | 1.844095 | 0.611566 |
| QC2   | 123.839 | 0.960723 | 0.29239 | 0.0353521 | 0.44485 | 5.506657 | 1.747883 | 0.558724 |
| QC3   | 215.653 | 0.996755 | 0.288965 | 0.0511751 | 0.435797 | 5.283002 | 1.830296 | 0.588697 |
| mean  | 177.8533 | 0.991094 | 0.297674 | 0.0424968 | 0.457528 | 5.489524 | 1.807425 | 0.586329 |
| SD    | 39.19727 | 0.02284 | 0.009993 | 0.0065501 | 0.024609 | 0.162083 | 0.042478 | 0.021637 |
| CV    | 22.0391 | 2.304543 | 3.357011 | 15.41308 | 5.378711 | 2.952591 | 2.350178 | 3.690281 |

|       | LAMP1 | HEXA   | BHMT1  | KRT1   | GALNS  | PEBP1  | DSC2   | CTSL1  |
|-------|-------|--------|--------|--------|--------|--------|--------|--------|
| QC1   | 0.567994 | 0.081765 | 1.599123 | 1.8241034 | 0.049222 | 0.678218 | 0.189882 | 0.085533 |
| QC2   | 0.548064 | 0.077686 | 1.598073 | 1.9268526 | 0.050182 | 0.700525 | 0.182958 | 0.109676 |
| QC3   | 0.5466 | 0.084786 | 1.801216 | 1.9256295 | 0.05663 | 0.652403 | 0.172949 | 0.126183 |
| mean  | 0.554219 | 0.081412 | 1.666137 | 1.8921951 | 0.052011 | 0.677049 | 0.18193 | 0.10713 |
| SD    | 0.009759 | 0.002909 | 0.095516 | 0.0481507 | 0.00329 | 0.019663 | 0.006951 | 0.016693 |
| CV    | 1.760781 | 3.573634 | 5.732779 | 2.5447016 | 6.324699 | 2.90422 | 3.820796 | 15.5816 |
Supplemental Table S6: Unpublished data of untargeted proteomics analyses of urine samples from ten controls and four NPC patients. Depicted are the two possible biomarkers for LSD (LAMP1) and CLN2 (BHMT1). Related to figure 3-5.

|     | BHMT1 | LAMP1 |
|-----|-------|-------|
| ctrl| 11.12 | 18.33 |
| ctrl| 9.41  | 17.47 |
| ctrl| 9.50  | 17.83 |
| ctrl| 11.93 | 22.06 |
| ctrl| 13.91 | 24.53 |
| ctrl| 12.67 | 23.17 |
| ctrl| 11.19 | 15.71 |
| ctrl| 13.05 | 21.27 |
| ctrl| 13.06 | 17.56 |
| ctrl| 13.39 | 12.07 |
| NPC | 10.94 | 18.46 |
| NPC | 9.34  | 12.75 |
| NPC | 6.62  | 21.03 |
| NPC | 6.76  | 13.62 |
| mean (ctrl) | 11.92 | 19.00 |
| mean (NPC) | 8.42  | 16.47 |
| t-test | 0.0049 | 0.2830 |
**Transparent Methods**

**Ethics Approval**
The collection of samples for this study has ethical approval (13/LO/0168; IRAS ID 95005; London-Bloomsbury Research Ethics Committee) and Health Research Authority (HRA) approval.

**Urine Samples**
Table 2 lists all the urine samples used in this study. Patient samples were collected from NCL patients after obtaining informed consent and were anonymised and stored at -80 °C within 24 hours of collection. Two CLN2 patients were on cerliponase alpha treatment and pre- and post-treatment samples were obtained. The urine of 3 (3 x male) CLN5 sheep (Frugier et al., 2008) and 13 (7 female, 6 male) CLN6 sheep (Tammen et al., 2006) as well urine from 7 unaffected control sheep (2 female, 5 male) were collected and provided by Prof David Palmer Lincoln University, New Zealand from flocks maintained at Lincoln University under protocols approved by the Lincoln University Animal Ethics Committee and in accordance with the New Zealand Animal Welfare Act, 1999 and NIH guidelines. Samples were frozen at -80 °C within 24 hours of collection.
Table S7: Human cohort sample information for CLN type, age, sex and the corresponding control samples. All samples were analysed and the NCL type was determined by the Great Ormond Street Diagnostic Enzyme Laboratory. Two of the CLN2 patients are subdivided into pre- and post- which indicates whether the sample was taken pre-treatment or with ongoing ERT treatment (post). Separate patient and control samples were used in the label free proteomics discovery analysis.

| patient samples | age [years] | sex |
|-----------------|------------|-----|
| **Label Free Proteomics** | | |
| CLN2a           | 5          | M   |
| CLN2b           | 7          | M   |
| 4 x Control     | 3-10       | M   |
| **Targeted Proteomics** | | |
| CLN1(1)         | 7.5        | M   |
| CLN1(2)         | 7.3        | F   |
| CLN2(1)-pre     | 4.3        | M   |
| CLN2(1)-post    | 9          | M   |
| CLN2(2)         | 5.7        | M   |
| CLN2(3)-pre     | 13.8       | F   |
| CLN2(3)-post    | 17.8       | F   |
| CLN2(4)         | 7.9        | F   |
| CLN2(5)         | 3          | M   |
| CLN2(6)         | 6.9        | F   |
| CLN2(7)         | 7          | M   |
| CLN3(1)         | 28.5       | F   |
| CLN5(1)         | 10.2       | M   |
| CLN5(2)         | 10.2       | M   |
| CLN6(1)         | 5.1        | M   |
| CLN6(2)         | 8.5        | M   |
| CLN7(1)         | 6.6        | M   |
| **control samples** | | |
| ctrl(1)         | 2.4        | M   |
| ctrl(2)         | 5.8        | F   |
| ctrl(3)         | 6.2        | M   |
| ctrl(4) | 6.2 | M |
|---------|-----|---|
| ctrl(5) | 7.2 | M |
| ctrl(6) | 9.2 | F |
| ctrl(7) | 8.5 | M |
| ctrl(8) | 12.4 | F |
| ctrl(9) | 15.4 | F |
| ctrl(10) | 16 | F |
| ctrl(11) | 31.5 | F |
| ctrl(12) | 7.5 | M |
| ctrl(13) | 11.2 | M |

**Label Free Proteomics**

Human and sheep urine was prepared as previously described. (Heywood et al., 2015) Urine samples were thawed, 2.5 mL of each sample were aliquoted into 5 mL centrifugation tubes and centrifuged for 30 min at 4600 rpm and room temperature to separate out from particulates and cell debris. 2 mL of the supernatant were filtered using an Amicon Ultra-15 filters (3 kDa). Filters were centrifuged for 60 min at 4600 rpm to desalt and concentrate the urine samples to a residual volume of approx. 200 µL transferred to a new plastic tube and combined with the wash solution (100 µL of 50 mM ammonium bicarbonate) of the filter. Protein was precipitated with cold acetone (800 µL) over-night at -20 °C. Protein in acetone was centrifuged at 16900 x g for 10 min at 4 °C. The supernatant was discarded, 50 µL of ddH$_2$O added to each sample, which was vortexed thoroughly and freeze-dried overnight. To each dried protein sample 20 µL digest buffer (6M urea, 2M thioruea, 2% ASB14, 100mM Tris-HCL pH 7.8) was added and left to shake for 1 h. Next, 1.5 µL of DTE (30 mg/mL) was added to break the disulphide bonds then samples were shaken for 1 h then 3.0 µL of IAA added followed by shaking for another 45 min covered from light to carboximidomethylate all cysteine residues. To this solution 165 µL of ddH$_2$O and 10 µL of trypsin gold (0.1 µg/µL) were added followed by incubation overnight (16 h) at 37 °C to digest the proteins.

Peptide samples underwent an SPE clean-up before LC-MS/MS analysis. Samples were diluted with 200 µL 0.2% TFA to adjust to a final concentration of 0.1% TFA. SPE cartridges (ISOLUTE®C18, 1 mL) were equilibrated with 2 x 1 mL 70% ACN, 0.1%TFA and then 2 x 1 mL H$_2$O, 0.1% TFA. The samples were loaded and allow to drip by gravity, cartridges were washed with 2 x 1 mL H$_2$O, 0.1% TFA before the peptides were eluted in 2 x 250 µL 70% ACN into a new plastic tubes. The solvents were evaporated, and dried peptides were stored at -20 °C until further usage. Concentrated urine was acetone precipitated and the protein pellet subjected to tryptic digestion followed by desalting using
C18 columns (Agilent, UK). Peptides were analysed over a 60-minute gradient of increasing acetonitrile using a SYNAPT G2-Si (Waters, Manchester) mass spectrometer coupled to a nanoACQUITY UPLC (Waters, Manchester) as previously described. (Bliss et al., 2016) Raw data was processed using Progenesis Qi (Nonlinear dynamics, UK) software.

Proteins were identified at first pass against a downloaded UniProt sheep reference proteome to which the sequence of P00761 porcine trypsin were added manually. Due to lack of annotation in the sheep database a second pass identification was performed using the human reference proteome database. Data was searched with fixed modifications of carboxamidomethylation of cysteines, dynamic modifications of deamidation of asparagine/glutamine and oxidation of methionine, one missed cleavage sites, and a false discovery rate set at 1%. Only protein identifications with >95% confidence and more than one peptide were used to determine significance in protein expression between groups.

**Statistical and bioinformatics analysis**

Multivariate analysis of the targeted proteomic data was performed using SIMCA v 15 (Umetrics, Sweden). Analysis of the affected pathways was conducted with Ingenuity Pathway Analysis (IPA, Qiagen). Cellular compartment locations of proteins were analysed with DAVID Bioinformatics Resources. (Huang et al., 2008, Huang et al., 2009) Exported data were analysed using Excel and R studio. The results were depicted using R studio.

**Targeted MRM LC-MS/MS Assay**

**Candidate selection:** Proteins were selected from the label free analysis by ranking proteins by their number of detected unique peptides and their confidence score resulted in A ranked proteins with > 2 unique peptides and a confidence score > 15, B ranked proteins with either > 2 unique peptides or a confidence score > 15 and the residual proteins that were labelled as C. Only A and B ranked proteins were selected based on significance (ANOVA p < 0.05) and normalised fold change (> 2, < -2). Known lysosomal proteins observed altered in the profiling analysis were selected. Significantly decreased protein AKT1 was chosen for the targeted MRM panel since it was previously mentioned in neurodegenerative diseases, where a disruption of Akt signalling was found to contribute to the pathogenesis. (Palmieri et al., 2017). Proteins from the sheep analysis were selected based on shared significance between CLN5 and 6 which included 3 proteins (LEAP2, secretagogin and desmocollin 2). Cathepsin B and LAMP2 had a fold change greater than 2 in the analysis. Additional proteins we thought could possibly detected in urine were selected from previous published studies. Those are
proteins directly found in NCL patients (Sleat et al., 2017, Sleat et al., 2019) like Cathepsin L1 (CTSL1), TPP1 and alpha-N-acetylgalactosaminidase (NAGA) or added out of interest, i.e. PAM (Doll et al., 2017) because it was detected in a heart proteome study. Since NCLs are thought to affect other organs like the heart, this protein was added as a potential marker for heart impairment. The ganglioside GM2 activator (GM2A) was included to gain further knowledge in the ganglioside pathway.

**MRM-LC-MS/MS assay development:** Two representative quantotypic peptides for the human version of all proteins were identified using original label free proteomics data and the MRM database (www.mrm/thegpm.org). The peptides of choice were custom synthesized by Genscript, USA to identify the correct m/z in urine, optimise the detection and determine the retention times. Based on the levels of detection in urine, i.e. noninterfering peaks, and reflection of quantitative values the optimal peptides per protein with two transitions each were selected. Details of those peptides are given in Supplemental Table S7 as well as calibration curves in Supplemental Figure 3.

**Sample preparation:** For targeted proteomic analysis urine was prepared as described above after the addition of 7.5 µL of whole yeast enolase protein (30 µg/µL) (Sigma UK) the internal standard during sample preparation and for quantitation.

**LC-MS/MS analysis:** All samples were injected onto a Waters CORTECS UPLC C18+ column (90 Å, 1.6 µm, 3 mm x 100 mm) attached to a C18+ VanGuard precolumn (Supplemental Table S5). Multiple reaction monitoring was performed over a 12 min gradient (Supplemental Table S6) on a Waters Xevo TQ-S mass spectrometer as previously described and quantitative data acquired using the transitions given in Supplemental Table S7.

**Quality control and standardisation:** Calibration curves of all peptides were prepared using pooled urine and ddH₂O with the same amount of yeast enolase as in the pooled urine samples. Calibration curves were measured together with a urine pool as quality controls (QCs) (Supplemental Table S8) at the beginning, in the middle and at the end of the patient and control urine samples. Urine creatinine was measured by LC-MS for every sample to be used as normalisation factor for a better comparison of the individual urine samples. (Mills et al., 2005a)

**Data processing:** The areas under the curve for each peptide of interest were analysed using Waters TargetLynx software. Those areas were normalised to the areas of the yeast enolase peptide internal standard. Those ratios were translated into absolute values with the linear equation of the corresponding standard curves. The absolute values of each peptide of interest in a sample were normalised to the creatinine amounts in this sample. (Mills et al., 2005b)
### Supplemental Table S8: Table of mass spectral parameters used for analysis in this study. Related to transparent methods

| Parameter                        | Value                                      |
|----------------------------------|--------------------------------------------|
| column                           | Cortecs C18 (90 Å) Waters Corp.            |
| length                           | 50 mm                                      |
| internal diameter                | 2.1 mm                                     |
| particle diameter                | 1.6 μm                                     |
| column temperature               | 40 °C                                      |
| guard column                     | Cortecs C18 VanGuard Pre-column (90 Å, 1.6 μm, 2.1 mm x 5 mm) |
| weak wash solvent                | H₂O, 0.1% TFA                              |
| strong wash solvent              | H₂O:ACN:MeOH:IPA 1:1:1:1                   |
| mobile phase A                   | H₂O, 0.1% FA                               |
| mobile phase B                   | CAN, 0.1% FA                               |
| gradient                         | yes                                        |
| flow rate                        | 0.5 mL/min                                 |
| injection volume                 | 3 μL                                       |
| injection mode                   | partial loop                               |
| autosampler temperature          | 10 °C                                      |
| capillary voltage                | 2.8 kV                                     |
| nebulizer gas flow               | 8.0 bar                                    |
| source temperature               | 150 °C                                     |
| desolvation temperature          | 600 °C                                     |

### Supplemental Table S9: Liquid chromatography gradient parameters used in this study. Related to transparent methods

| Time   | Flow [mL/min] | %A  | %B  | Curve |
|--------|---------------|-----|-----|-------|
| 0.0    | 0.5           | 97  | 3   | Initial |
| 0.9    | 0.5           | 97  | 3   | 6     |
| 7.8    | 0.5           | 70  | 30  | 6     |
| 8.5    | 0.5           | 30  | 70  | 6     |
| 9.5    | 0.5           | 30  | 70  | 6     |
| 10.01  | 0.5           | 97  | 3   | 1     |
| 12     | 0.5           | 97  | 3   | 1     |
Supplemental Table S10: Mass transition and conditions for all peptides of interest used in this study. Related to transparent methods

| Gene | peptide sequence | precursor ion [m/z] | product ion [m/z] | cone energy | collision energy | fragmentation |
|------|------------------|---------------------|-------------------|-------------|------------------|--------------|
| AKT1 | EAPLNNFSVAQCQLMK | 611.2992 2+ | 621.2991* | 35 | 21 | b6 |
| AKT1 | EAPLNNFSVAQCQLMK | 611.2992 2+ | 679.3266* | 35 | 21 | y5 |
| BHMT1 | ISGQEVNEAACDIAR | 816.8859 2+ | 705.3348* | 35 | 29 | y6 |
| BHMT1 | ISGQEVNEAACDIAR | 816.8859 2+ | 905.4145* | 35 | 29 | y8 |
| CTSB | VMFTEDLK | 491.7493 2+ | 752.3825* | 35 | 17 | y6 |
| CTSB | VMFTEDLK | 491.7493 2+ | 883.423* | 35 | 17 | y7 |
| CTSB | QGDHCIESEVVAGIPR | 608.6283 2+ | 513.3144* | 35 | 21 | y5 |
| CTSB | QGDHCIESEVVAGIPR | 608.6283 2+ | 711.4512* | 35 | 21 | y7 |
| CTSH | TGIYSSTSCHK | 414.1924 2+ | 485.2084 2+ | 35 | 14 | y8 |
| CTSH | TGIYSSTSCHK | 414.1924 2+ | 570.2611 2+ | 35 | 14 | y10 |
| CTSH | LQTFASNWDR | 561.7882 2+ | 780.3787* | 35 | 20 | y6 |
| CTSH | LQTFASNWDR | 561.7882 2+ | 881.4264* | 35 | 20 | y7 |
| CTSL1 | LYGMNEEGWR | 627.7822 2+ | 921.3883* | 35 | 22 | y7 |
| CTSL1 | LYGMNEEGWR | 627.7822 2+ | 978.4098* | 35 | 22 | y8 |
| CTSZ | VGDYGSLGSR | 505.7487 2+ | 576.31* | 35 | 18 | y6 |
| CTSZ | VGDYGSLGSR | 505.7487 2+ | 854.4003* | 35 | 18 | y8 |
| DSC2 | ENAEVGTTSNGYK | 685.3151 2+ | 827.3894* | 35 | 24 | y8 |
| DSC2 | ENAEVGTTSNGYK | 685.3151 2+ | 926.4578* | 35 | 24 | y9 |
| GALNS | YYEEFPINLK | 658.3321 2+ | 860.4876* | 35 | 23 | y7 |
| GALNS | YYEEFPINLK | 658.3321 2+ | 989.5302* | 35 | 23 | y8 |
| GM2A | EGTYSLPK | 447.7324 2+ | 444.2817* | 35 | 16 | y4 |
| GM2A | EGTYSLPK | 447.7324 2+ | 607.745* | 35 | 16 | y5 |
| GOT1 | HIYLLPSGR | 528.3035 2+ | 805.4567* | 35 | 18 | y7 |
| GOT1 | HIYLLPSGR | 528.3035 2+ | 918.5407* | 35 | 18 | y8 |
| GOT1 | EPESILQVLSQMEK | 815.9215 2+ | 962.4975* | 35 | 29 | y8 |
| GOT1 | EPESILQVLSQMEK | 815.9215 2+ | 1075.5816* | 35 | 29 | y9 |
| HEXA | ISYGPDK | 483.24 2+ | 545.2718* | 35 | 17 | y4 |
| HEXA | ISYGPDK | 483.24 2+ | 765.3566* | 35 | 17 | y6 |
| Gene  | Peptide            | Msec1 | Msec2 | Score1 | Score2 | y   |
|-------|--------------------|-------|-------|--------|--------|-----|
| KRT1  | TNAENEFVTIK        | 633.3223 | 979.5095 | 35     | 22     | y7  |
| KRT1  | TNAENEFVTIK        | 633.3223 | 850.4669 | 35     | 22     | y8  |
| LAMP1 | ALQATVGNSYK        | 576.3064 | 768.3887 | 35     | 20     | y7  |
| LAMP1 | ALQATVGNSYK        | 576.3064 | 839.4258 | 35     | 20     | y8  |
| LAMP2 | IPLNDLFR           | 494.2847 | 664.3413 | 35     | 17     | y5  |
| LAMP2 | IPLNDLFR           | 494.2847 | 874.4781 | 35     | 17     | y7  |
| LAMP2 | GILTVDPELLAIR      | 656.8954 | 829.4778 | 35     | 23     | y7  |
| LAMP2 | GILTVDPELLAIR      | 656.8954 | 928.5462 | 35     | 23     | y8  |
| LEAP2 | DDSECTIR           | 498.2086 | 276.1666 | 35     | 17     | y2  |
| LEAP2 | DDSECTIR           | 498.2086 | 549.2813 | 35     | 17     | y4  |
| NAGA  | NCISEQLFMEMADR     | 581.9215 | 621.2661 | 35     | 20     | y5  |
| NAGA  | NCISEQLFMEMADR     | 581.9215 | 752.3066 | 35     | 20     | y6  |
| NAGA  | TISAQNMQLNPLMIK    | 644.0094 | 715.3807 | 35     | 22     | y6  |
| NAGA  | TISAQNMQLNPLMIK    | 644.0094 | 843.4757 | 35     | 22     | y7  |
| PAM   | IVQFSPSGK          | 481.7689 | 622.3195 | 35     | 17     | y6  |
| PAM   | IVQFSPSGK          | 481.7689 | 750.3781 | 35     | 17     | y7  |
| PAM   | NNLVIFHR           | 506.7882 | 572.3303 | 35     | 18     | y4  |
| PAM   | NNLVIFHR           | 506.7882 | 671.3988 | 35     | 18     | y5  |
| PEBP1 | LYTLVTDPDAPSR      | 780.9172 | 858.3952 | 35     | 28     | y8  |
| PEBP1 | LYTLVTDPDAPSR      | 780.9172 | 971.4792 | 35     | 28     | y9  |
| SCGN  | IFAYYDVK           | 553.2819 | 448.2402 | 35     | 19     | y4  |
| SCGN  | IFAYYDVK           | 553.2819 | 845.404 | 35     | 19     | y7  |
| TPP1  | LYQQHAGLFDVTR      | 535.6108 | 637.3304 | 35     | 18     | y13 |
| TPP1  | LYQQHAGLFDVTR      | 535.6108 | 746.371 | 35     | 18     | y5  |
| TPP1  | LFGGNFAHQASVAR     | 492.2565 | 607.8052 | 35     | 16     | y12 |
| TPP1  | LFGGNFAHQASVAR     | 492.2565 | 768.4111 | 35     | 16     | y7  |
| YE    | GNPTVEELTTEK       | 709.062 | 623.492 | 35     | 18     | y11 |
| YE    | GNPTVEELTTEK       | 709.062 | 948.68 | 35     | 20     | y8  |
References

Bliss, E., Heywood, W. E., Benatti, M., Sebire, N. J., Mills, K. (2016). An optimised method for the proteomic profiling of full thickness human skin. Biological procedures online, 18, 15-15.

Doll, S., Dreßen, M., Geyer, P. E., Itzhak, D. N., Braun, C., Doppler, S. A., Meier, F., Deutsch, M.-A., Lahm, H., Lange, R., Krane, M., Mann, M. (2017). Region and cell-type resolved quantitative proteomic map of the human heart. Nature Communications, 8, 1469.

Frugier, T., Mitchell, N. L., Tammen, I., Houweling, P. J., Arthur, D. G., Kay, G. W., van Diggelen, O. P., Jolly, R. D., Palmer, D. N. (2008). A new large animal model of CLN5 neuronal ceroid lipofuscinosis in Borderdale sheep is caused by a nucleotide substitution at a consensus splice site (c.571+1G>>>A) leading to excision of exon 3. Neurobiology of Disease, 29, 306-315.

Heywood, W. E., Camuzeaux, S., Doykov, I., Patel, N., Preece, R.-L., Footitt, E., Cleary, M., Clayton, P., Grunewald, S., Abulhoul, L., Chakrapani, A., Sebire, N. J., Hindmarsh, P., de Koning, T. J., Heales, S., Burke, D., Gissen, P., Mills, K. (2015). Proteomic Discovery and Development of a Multiplexed Targeted MRM-LC-MS/MS Assay for Urine Biomarkers of Extracellular Matrix Disruption in Mucopolysaccharidoses I, II, and VI. Analytical Chemistry, 87, 12238-12244.

Huang, D. W., Sherman, B. T., Lempicki, R. A. (2008). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research, 37, 1-13.

Huang, D. W., Sherman, B. T., Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4, 44-57.

Mills, K., Morris, P., Lee, P., Vellodi, A., Waldek, S., Young, E., Winchester, B. (2005b). Measurement of urinary CDH and CTH by tandem mass spectrometry in patients hemizygous and heterozygous for Fabry disease. Journal of Inherited Metabolic Disease, 28, 35-48.

Palmieri, M., Pal, R., Nelvagal, H. R., Lotfi, P., Stinnett, G. R., Seymour, M. L., Chaudhury, A., Bajaj, L., Bondar, V. V., Bremner, L., Saleem, U., Tse, D. Y., Sanagasetti, D., Wu, S. M., Neilson, J. R., Pereira, F. A., Pautler, R. G., Rodney, G. G., Cooper, J. D., Sardiello, M. (2017). mTORC1-independent TFEB activation via Akt inhibition promotes cellular clearance in neurodegenerative storage diseases. Nature Communications, 8, 14338.

Sleat, D. E., Tannous, A., Sohar, I., Wiseman, J. A., Zheng, H., Qian, M., Zhao, C., Xin, W., Barone, R., Sims, K. B., Moore, D. F., Lobel, P. (2017). Proteomic Analysis of Brain and Cerebrospinal Fluid from the Three Major Forms of Neuronal Ceroid Lipofuscinosis Reveals Potential Biomarkers. Journal of Proteome Research, 16, 3787-3804.

Sleat, D. E., Wiseman, J. A., El-Banna, M., Zheng, H., Zhao, C., Soherwardy, A., Moore, D. F., Lobel, P. (2019). Analysis of Brain and Cerebrospinal Fluid from Mouse Models of the Three Major Forms of Neuronal Ceroid Lipofuscinosis Reveals Changes in the Lysosomal Proteome. Molecular & Cellular Proteomics, 18, 2244-2261.

Tammen, I., Houweling, P. J., Frugier, T., Mitchell, N. L., Kay, G. W., Cavanagh, J. A. L., Cook, R. W., Raadsma, H. W., Palmer, D. N. (2006). A missense mutation (c.184C>T) in ovine CLN6 causes neuronal ceroid lipofuscinosis in Merino sheep whereas affected South Hampshire sheep have reduced levels of CLN6 mRNA. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 1762, 898-905.