Protein N-Terminomics Reveals Major Proteases in Regulating Beige Adipocyte Differentiation

Chih-Hsiang Chang[1], Hsin-Yi Chang[1, 2,*] and Yasushi Ishihama[1, 3, 4,*]

[1]Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan.
[2]Graduate Institute of Medical Sciences, National Defense Medical Center, Taipei 11490, Taiwan.
[3]Laboratory of Clinical and Analytical Chemistry, National Institute of Biomedical Innovation, Health and Nutrition, Ibaraki, Osaka, 567-0085, Japan
[4]Lead contact

Running title: Beige adipocyte differentiation by N-terminomics

*Co-corresponding authors

Dr. Hsin-Yi Chang
Graduate Institute of Medical Sciences, National Defense Medical Center
161, Sec. 6, Minquan E. Rd., Neihu Dist., Taipei, 11490, Taiwan (R.O.C.)
Tel: +886-2-8792-3100;18525, E-mail: hsinyi.chang@mail.ndmctsgh.edu.tw

Dr. Yasushi Ishihama (Lead)
Graduate School of Pharmaceutical Sciences, Kyoto University
46-29 Yoshidashimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan
Tel: +81-75-753-4555, Fax: +81-75-753-4601, E-mail: yishiham@pharm.kyoto-u.ac.jp

Summary

We have developed an analytical platform to measure the temporal profiles of the proteome and the protein N-terminome during beige adipocyte differentiation. Our recently developed N-terminomics approach was applied to isolate protein N-terminal peptides with up to 97% purity. In total, we identified 7,171 unique N-terminal peptides, of which 3,043 were from canonical proteins (CanNt peptides) and 4,129 were from proteins with neo-N-termini (NeoNt peptides). We found that 44% of proteins showed different temporal profiles between the global proteome and protein N-terminome. Our results indicate that proteolysis plays a major role in regulating beige adipocyte differentiation. Knockdown of Pmpbc, Plg or Cstd in preadipocytes impeded thermogenesis by decreasing Cidea, Pgc1a and Ucp1 expression and elevating adipogenic markers such as Fasn and Fabp4, thus promoting the beige-to-white adipocyte transition. Our approach reveals that proteases and the generated proteolysis products play major roles in beige adipocyte differentiation.

Keywords: N-terminomics, adipocyte differentiation, proteomics, temporal profiling, proteolysis, protease

Introduction

It is well established that adipose tissue is primarily responsible for energy regulation in the mammalian body, and recent studies have revealed that it also plays an important role in the regulation of the endocrine system. Indeed, metabolic dysregulation, which leads to obesity, insulin resistance, type 2 diabetes, dyslipidemia, and cancer cachexia, is invariably accompanied by adipocyte dysfunction (Kershaw and Flier, 2004). Mammals have two types of adipocytes, white adipocytes and brown adipocytes, which are responsible for regulating metabolic homeostasis throughout the body through adipogenesis and adaptive thermogenesis, respectively. White adipocytes store and release energy in the form of fatty acids in response to systemic demand, while brown adipocytes generate heat by oxidizing substrates such as fatty acids and glucose. Some white adipocytes have the ability to generate heat, and these cells, named beige, have a high mitochondrial mass and express thermogenesis-specific genes such as Ucp1, Cidea, and Pgc1a (Giralt and Villarroya, 2013; Shapira and Seale, 2019; Wang and Seale, 2016). Beige adipocytes were found to have a similar function to brown adipocytes, but express characteristic markers that distinguish them from brown adipocytes (Harms and Seale, 2013; Ikeda et al., 2018).

Adipocytes are derived from mesenchymal stem cells, which differentiate into brown and white adipocytes early in their development. Brown adipocyte progenitors express the muscle marker myogenic factor 5 (MYF5), whereas white adipocytes do not. Brown adipocytes have multivesicular fat droplets and numerous mitochondria filling the cytoplasm, whereas white adipocytes have large univesicular fat droplets and a limited number of mitochondria. Beige adipocytes can arise by preadipocyte commitment or dedifferentiation from white adipocytes in response to certain environmental conditions or external cues, including chronic cold acclimation, exercise, or long-term treatment with peroxisome proliferator-activated receptor gamma (PPARg) agonists or beta-3-adrenergic receptor (AR) agonists (Giralt and Villarroya, 2013). Beige adipocytes express UCP1 and other genes related to mitochondrial biogenesis, and exhibit multivesicular lipid droplets and a large number of mitochondria in the cytoplasm. Stimulation of beige adipocyte differentiation improves glucose and insulin homeostasis and has anti-obesity and antidiabetic effects in rodent models (Liu et al., 2013; Rodriguez-López et al., 2021; Whitehead et al., 2021). Recently, several studies have identified transcriptional regulators involved in beige adipocyte development (Guo et al., 2015; Marcon et al., 2017). Genomic and transcriptomic analysis of single adipocytes established extensive cellular heterogeneity of adipocytes and identified a variety of cell types from obese patients (Deutsch et al., 2020; Henriques et al., 2020; Vijay et al., 2020). Several groups have also profiled adipocyte proteomes by mass spectrometry and reported functional annotations (Adachi et al., 2007; Choi et al., 2020; Du et al., 2021; Forner et al., 2009; Giansanti et al., 2022). Shinoda et al. identified CK2 as a negative regulator of heat production in beige adipocytes by shotgun

phosphoproteomics (Shinoda et al., 2015). Furthermore, cathepsins, which belong to the cysteine protease family, play a major role in the regulation of adipocyte functions (Han et al., 2009; Taleb et al., 2006; Yang et al., 2007). However, the translational and post-translational regulation of differentiation into beige adipocytes is still poorly understood at the proteome level.

Proteomic techniques based on modern mass spectrometry continue to evolve, making it possible to quantitatively measure not only translation products, but also post-translational modification products and, more recently, their spatial and temporal distribution in unprecedented depth (Bantscheff et al., 2012; Bekker-Jensen et al., 2017; Gao and Yates, 2019). In particular, N-terminomics, which focuses on the N-terminus of proteins, can identify canonical and noncanonical N-termini of proteins and provide information on alternative splicing, alternative translation initiation, cotranslation, and post-translational modifications (Bludau et al., 2021; Wang et al., 2021). So far, a number of techniques have been developed to enrich protein N-terminal peptides for shotgun proteomics by means of positive or negative selection strategies. Both strategies require chemical or enzymatic modification of N-terminal peptides. Positive selection methods capture N-terminal peptides by affinity purification (Chen et al., 2016; Mahrus et al., 2008; Vaca Jacome et al., 2015; Weeks et al., 2021), whereas negative selection methods such as Terminal Amine Isotopic Labeling of Substrates (TAILS) (Kleifeld et al., 2010), Combined/Charge-based FRActional DIagonal Chromatography (CO-/Cha-FRADIC) (Gevaert et al., 2003; Venne et al., 2015) and hydrophobic tagging-assisted N-termini enrichment (HYTANE) (Chen et al., 2016) deplete non-N-terminal peptides. In any case, tedious steps and relatively large amounts of samples are required. Recently, we developed a new method, which we called CHAMP (CHop and throw to AMPlify the terminal peptides), to isolate protein N-terminal peptides in only two steps, that is, TrypN or LysargiNase digestion and low-pH SCX chromatography. This afforded protein N-terminal peptides in up to 97% purity, and we identified over 1,700 protein N-termini from 20 µg of peptides (Chang et al., 2021). This CHAMP approach was applied to profile ectodomain shedding from ten human cell lines, and we identified 6,181 N-termini from proteolytic products as well as canonical proteins (Tsumagari et al., 2021a, 2021b).

In this study, we measured the temporal profiles of the global proteome and N-terminome during differentiation into beige adipocytes. The N-terminome data enabled us to analyze the expression changes of each proteoform during the maturation process of beige adipocytes. This is the first N-terminome study that covers in detail regulation at the transcriptional, translational, and post-translational levels during cell differentiation.

Results

N-Terminal peptide enrichment and LC/MS/MS analysis

To investigate changes in the global and N-terminal proteomes during beige adipocyte maturation, immortalized preadipocytes were induced to differentiate and were harvested every 2 days during differentiation. For label-free quantitative analysis of the global proteome, extracted proteins from five time points were digested with TrypN and the peptides obtained were analyzed by nanoLC/MS/MS. For N-terminome analysis, the protein N-terminal peptides were isolated by SCX chromatography from TrypN digests, using the CHAMP protocol to enrich canonical protein N-terminal peptides as well as non-canonical N-terminal peptides without chemical derivatization (Chang et al., 2021). For in-parallel multiple sample preparation, we further adapted the method to incorporate StageTip-based SCX chromatography (Figure 1A). Applying match-between-runs identification and label-free quantitation, we found that most proteins were regulated in a time-dependent fashion with a larger proportion decreasing during cell differentiation (Figure 1B). In total, we identified 3,081 proteins and 7,171 unique N-terminal peptides from 3,053 proteins, with 1,757 proteins commonly identified between the two datasets (Figure 1C).

Annotation of protein N-terminal peptides

We categorized the N-termini according to their features (Document S1, Figure S1 and Figure 2A). Using these definitions, we annotated 3,043 unique CanNt peptides, 4,129 NeoNt peptides and 378 internal peptides from 7,550 unique peptides (Figure 2B, Table S1A). By applying the CHAMP protocol with StageTip-based SCX chromatography, we achieved more than 95% purity for both CanNt and NeoNt peptides. We compared these N-terminal peptides with the identified proteins from global proteomics, and found that only 423 proteins were commonly identified, showing the importance of N-terminomics to profile the protein N-termini (Figure 1C).

Co-translational modifications and protein N-terminal processing events are important determinants of the N-terminome. Excision of initial methionine and acetylation of the N-terminal α-amino group are the two most common modifications coupled in tandem to protein translation, and are reported to have specific sequence preferences (Bonissone et al., 2013) (Helbig et al., 2010) (Figure 2C, D). Not only the CanNt peptides, but also a part of the NeoNt peptides have N-terminal acetylation, suggesting the presence of a significant number of unreported translation initiation sites. Several peptide processing events were observed with N-terminally unmodified NeoNt peptides, including the removal of 11 propeptides, 77 transit peptides, and 49 signal peptides. We considered that the remaining 3,692 N-terminally unmodified NeoNt peptides with undefined peptide processing sites are proteolytic products generated by various endogenous proteases. To simplify the temporal analyses of these proteolytic peptides, we removed the sequence redundancy to obtain 1,313 unique N-termini for the proteolysis analysis (Figure S1A, D). In most cases, the number of

N-termini matched the number of proteins, but in the "other" category of proteolytic peptides, there was a discrepancy between the numbers of N-termini and proteins (Figure 2B). This was due to the fact that multiple cleavages can occur in one protein.

Temporal profiles of N-terminome and proteome

We applied MaxLFQ to quantify 3,032 proteins from the global proteome (Table S1B). In the N-terminome, 2,124 CanNt and 2,231 NeoNt peptides were quantified (Table S1C). We evaluated the temporal dynamics of the global proteome, CanNt peptides and NeoNt peptides by calculating PCC between the expression level and the differentiation time points (Figure 1B). The global proteome exhibits a symmetric distribution of increased (PCC > 0.4, red) and decreased (PCC < -0.4, blue) time-courses of protein expression, whereas CanNt peptides showed a greater proportion in the decreasing fraction. This might be due to the increase of proteins during differentiation accompanied by translational or post-translational modifications of protein N-termini, resulting in diverse protein N-termini and protein NeoNt arising from alternative splicing, proteolysis or other unknown mechanisms. Regarding NeoNt peptides, it is difficult to assess the trend in Figure 1B because of multiple cleavages of a single protein.

In the global proteome, we performed a ChIP-Atlas transcription factor enrichment analysis (Oki et al., 2018) and confirmed that the temporal proteomic data in this study recaptured several critical regulations during adipocyte differentiation (Document S1, Table S1D). Moreover, we found down-regulation of mRNA processing/splicing, and ribonucleoproteins during differentiation (Figure S2A). However, cytosolic ribosomal proteins as well as translation initiation factors increased until day 6 and then suddenly decreased by day 8, whereas mitochondrial ribosomal proteins continued to increase throughout the differentiation process. These results indicate that the low rate of protein synthesis in non-differentiated cells maintains their overall homeostasis, whereas the stimulus to differentiation activates protein synthesis. The activation ends upon completion of differentiation, with the exception of mitochondrial proteins.

We further investigated the functions of the proteins commonly identified in the global proteome and CanNt peptides based on their temporal profiles. A total of 797 commonly quantified proteins in the global proteome and CanNt peptides were divided into seven groups and functionally annotated (Figure 3A and S2B). About 56% of the proteins showed the same profiles (groups 5-7), approximately 14% of the proteins exhibited negative regulation (groups 1-3), and 30% of them had irregular patterns (group 4) during cell differentiation (Figure S2B). For each group, we applied functional enrichment analysis using UniProt Keywords. In group 1, ubiquitin and proteasome were increased in the global proteome but decreased in CanNt peptides; group 2 exhibiting the reverse pattern to group 1 was enriched in the annotations of spliceosome, nucleus, and methylation; group 6 where the protein expression and CanNt peptides commonly decreased was enriched in shared functions of group 2 and to a greater extent in cytoplasm and

mRNA splicing; group 5 with an increasing pattern in both protein and CanNt peptides was enriched in mitochondrial proteins.

Since groups 1 and 2 exhibited reverse patterns between CanNt and the global proteome, for which the major regulations were proteolysis and alternative splicing, respectively, we next investigated proteome homeostasis during adipocyte differentiation. To this end, we utilized the NeoNt data to verify the alternative splicing and neo-translational initiation evidence at the transcriptional and translational levels, respectively, as described below.

Dynamic profiling of splicing variants

By combining the canonical and isoform sequences from Swiss-Prot and TrEMBL for peptide identification, we identified 1,697, 68 and 149 protein N-termini from Swiss-Prot canonical proteins, Swiss-Prot protein isoforms and TrEMBL protein N-termini, respectively. Among them, 41 proteoform pairs were identified and 23 pairs were quantifiable. In most cases, similar temporal profiles were observed during differentiation, except for only two pairs, Pym1 and Tnfaip8. By qPCR analysis, we confirmed that the protein variant expression is regulated at the transcriptional level (Figure 4). However, the contribution of splicing regulation with N-terminal change would not be a major contributor to beige adipocyte differentiation, as only 2 of the 41 cases were involved in alternative splicing, although previous studies focusing on genome-wide splicing events found a significant impact on the differentiation (Fiszbein and Kornblihtt, 2017; Salomonis et al., 2010).

Translational regulation

The protein N-terminus is considered to be the starting point of protein translation when it is acetylated. Therefore, the new TISs can be precisely identified as NeoNt peptides from our N-terminomics results. Among the 300 acetylated NeoNt peptides, preservation of the first methionine was found in 100 NeoNt peptides and excision of the first methionine was observed in 75 NeoNt peptides. The acetylated NeoNt peptides with N-terminal methionine had a strong preference for Glu, Asp and Asn at the P2 position, and the acetylated NeoNt peptides with methionine excision showed a strong preference for Ala, Ser and Thr at the P2 position (Figure 2D). The frequency logos of N-terminal amino acids of these acetylated NeoNt peptides were similar to those of the acetylated CanNt peptides (Figure 2C). Thus, they appear to be newly discovered protein N-termini, which have not yet been annotated in Uniprot. For acetylated NeoNt peptides without N-terminal preceding methionine, we extracted the nucleotide sequences from the upstream 180 bp of the putative TISs corresponding to the identified NeoNt peptides and employed a TIS prediction tool PreTIS to evaluate the confidence of alternative non-AUG TISs (Reuter et al., 2016). For the nucleotide sequences of the acetylated NeoNt peptides without Met at the −1 or 1 amino acid positions, PreTIS successfully predicted that the 28 acetylated NeoNts would be neoTISs, indicating that our N-terminal CHAMP method is a useful tool

to discover authentic TISs (Table S1E).

Among 300 newly identified acetylated NeoNt peptides, more than 45% of them were generated from the first 49 residues of the canonical protein N-termini (Figure S3A), and 34 of them were quantified together with the corresponding CanNt peptides. For all of the pairs of quantified acetylated NeoNt and CanNt peptides, similar temporal profiles were observed, suggesting that the ambiguous translational initiations reflected the canonical translation activity and were not regulated during beige maturation (Table S1F).

For the acetylated NeoNt peptides starting more than 50 residues from the canonical N-termini, only 5 NeoNt peptides were quantified together with the corresponding CanNt peptides and 4 of these pairs showed similar temporal profiles. Interestingly, an acetylated NeoNt peptide derived from the histone protein family showed a reversed temporal profile compared with 30 CanNt peptides (Figure S3B). The NeoNt peptide started at Ala68 and was reported to be cleaved by cathepsin E (Impens et al., 2010). However, we identified this NeoNt peptide as having α-amine acetylation, suggesting that it is likely to be a co-modification during protein translation. In view of the Kozak sequence motif near Ala68, we considered this NeoNt peptide to be not a proteolytic product, but a translational product. However, the reason why the temporal profile was reversed compared to those of 30 CanNt peptides remains unclear.

In addition, there were 27 acetylated NeoNt peptides starting from the third amino acid of the canonical protein sequences (Figure S3C, top), of which 13 lacking the first two [Met-Ala] residues were likely formed by removal of Ala concomitantly with first Met excision (Figure S3C, bottom). Thus, N-terminal acetyltransferases may acetylate not only the N-termini of the canonical sequences, but also the neoNts after the removal of the first two [Met-Ala] residues. Since we observed [Met-Ala] removal in many cases, this phenomenon should be a common event in protein translation processes, although the regulatory mechanisms involved, including the responsible aminopeptidase, are unknown.

Temporal profile correlation between NeoNt peptides and proteases

Since the unmodified NeoNt peptides were considered to be products of proteolysis, we characterized the unmodified NeoNt sequences with UniProt annotation and obtained 11, 77 and 49 cleavage sites for propeptides, signal peptides and transit peptides, respectively. Nevertheless, there were 3692 unmodified NeoNt peptides not characterized in UniProt, presumably in part because products generated by proteases, including exoproteases, with unrestricted cleavage preference are not registered in UniProt (Figure 2B). To evaluate the regulation of protease activity during beige adipocyte maturation, we assumed that the levels of proteases and the corresponding substrates were correlated. We extracted the expression values of annotated proteases from global proteome data and candidate NeoNt sequences from the N-terminome data, and correlated the protease-substrate information by considering the

cleavage preference, temporal profile, and subcellular localization.

By aligning the NeoNt peptides derived from the same protein identifier, we found a series of sequentially truncated N-termini, which might have resulted from proteolysis that occurred after endoprotease-driven cleavages. For example, the NeoNt peptides from protein Q8BHD7 have a series of different N-termini starting from the 72nd to 75th residues of the parent protein shown in Figure S1D. We next checked whether the subsequent proteolysis was regulated by a specific endopeptidase or represented spontaneous exopeptidase cleavage. We assumed that the regulated proteolysis should be abundance-independent and exhibit sequence preference. The proteins with sequential truncation that were also identified in the global proteome were selected to unravel the relationship between sequential truncation and protein abundance. As shown in Figure S4A, sequential truncation tended to occur in the higher abundance proteins, especially in the top 100 list, with up to 90% of the proteins having one or more amino acids sequentially truncated. We considered that the subsequent N-terminal truncations corresponded to exopeptidase-derived products formed in a protein abundance-dependent manner, which is also sensitive to peptide enrichment.

Among 2,093 quantified unmodified NeoNt peptides, 780 were putative exoprotease substrates with sequentially truncated peptides and 1,313 were putative endoprotease substrates (Figure S1A, Table S1G). To reveal the proteolytic events participating in adipocyte differentiation, we hierarchically clustered the quantitative results of protease substrates by temporal trends and extracted the sequence motifs as the cleavage preference (Figure 5A and S7B). Five main clusters were obtained (Figure S4B). Clusters 1 and 2 were increased, clusters 3 and 5 were decreased, and cluster 4 showed an unregulated profile. Sequences in cluster 1, which markedly increased on day 8, exhibited a cleavage preference of Lys/Arg at the P1’ position, whereas sequences in cluster 2, which increased from the early stage of beige adipocyte differentiation, showed a cleavage preference of Arg at the P2’ and P3’ positions (Figure 5A).

To identify the proteases responsible for generating the peptides in clusters 1 and 2, we extracted protease expression levels from the global proteome dataset based on the annotations registered in the peptidase database MEROPS (Rawlings et al., 2018) and matched their cleavage preferences to those of clusters 1 and 2. A total of 87 quantified proteases were obtained, of which 15 proteases have more than 100 reported substrates in MEROPS and were suitable for constructing models to predict proteases for given substrates (Figure 5B). A position weight matrix (PWM) per protease was constructed from known substrates in the MEROPS database, and the PWMs were applied to calculate the specificity of the substrates (Imamura et al., 2017; Tsumagari et al., 2021a).

Assuming that proteases and their substrates should be co-expressed, we focused on mitochondrial processing peptidase subunit beta (Pmpcb), plasminogen (Plg) and cathepsin D (Ctsd), which showed increasing trends during adipocyte differentiation. The PWM matching results of Pmpcb to

substrates of cluster 2 and Plg to substrates of cluster 1 are shown in Figure 5C. Pmpcb, a mitochondrial processing protease, is activated after cleavage by PINK1 to remove the transit peptide from the precursor protein newly imported into the mitochondria (Greene et al., 2012). The activation involves N-terminal cleavage between the 43rd-44th residues, and indeed, the N-terminal peptide starting at the 44th residue was also identified in cluster 2 with an increasing profile (Figure 5A). Furthermore, Pmpcb had a strong cleavage sequence preference for Arg at the P2' and P3' positions, which are appropriate features to construct the PWM. From the PWM matching result for Pmpcb, 15 proteolytic peptides with PWM scores larger than 4.0 were obtained (Figure 5C). To assess the correlation of these peptides and Pmpcb, we further considered their subcellular locations. All of the 15 proteolytic peptides in cluster 2 were located in mitochondria, and therefore are potential substrates of Pmpcb. We next compared the experimentally observed cleavage sites with the putative mitochondrial presequences registered in Uniprot or predicted by the bioinformatic tools TargetP-2.0 and MitoFates (Almagro Armenteros et al., 2019; Fukasawa et al., 2015). Figure 5D shows the results for 15 Pmpcb candidate substrates. All of the experimentally observed cleaved sites were matched to the computationally predicted sites, indicating that our approach provided accurate data on mitochondria presequence cleavage positions. This is the first example of experimental identification and quantification of the cleavage of mitochondrial transit peptides during mouse beige adipocyte differentiation via N-terminome profiling.

We also applied this approach to Plg and Ctsd to find their potential substrates (Figure 5B). Cstd, a lysosomal protease, is a regulator of UCP1 turnover and mitochondrial translation (Moazed and Desautels, 2002). Plg, located on the cell surface, is an activator of extracellular chemokines by proteolysis of the extracellular matrix. The downstream cascade of Plg proteolysis was proposed to regulate Wnt signaling, a switcher transcription factor for adipogenesis (Lilla et al., 2002). However, based on the cleavage specificity of the protease, it was difficult to construct an effective PWM for predicting Cstd substrates. On the other hand, Plg had a strong cleavage sequence preference for Arg and Lys at the P1' position and was used to construct the PWM. NeoNt peptides in Cluster 1 were evaluated by Plg PWM and 21 proteolytic peptides with high PWM scores were identified in the tubulin family and other proteins distributed throughout the cytoplasm, membrane, nucleus, and extracellular space (Figure 5C). However, taking the subcellular localization into account, these 21 proteolytic peptides were considered to be cleaved by other intracellular trypsin-like proteases, rather than Plg. In our current experimental design, the proteolytic products of extracellular proteins were difficult to collect, and another procedure would be required to isolate Plg substrates from the extracellular fractions (Tsumagari et al., 2021a). In brief, we successfully identified a number of endoprotease substrates with high confidence without manipulation of the proteases.

siRNA knockdown of Cstd, Plg or Pmpcb potentiates beige-to-white differentiation

To inspect the function of selected proteases in adipocyte differentiation, we performed siRNA knockdown of Cstd, Plg, or Pmpbc in preadipocytes prior to differentiation induction and examined the relative mRNA expression of various markers (Figure 5E). With the knockdown of Cstd, Plg, or Pmpcb, the adipocyte differentiation markers Pparg2 and Adipoq were unchanged or marginally increased, while the adipogenesis markers Fasn and Fabp4 were significantly enhanced, and the thermogenesis markers Cidea, Pgc1a and Ucp1 were markedly suppressed. The knockdown of each protease did not dampen cell differentiation, but directed differentiation in favor of white adipocytes rather than beige adipocytes. Thus, these three proteases, Cstd, Plg and Pmpbc, cleaved proteins in the lysosome, extracellular matrix and mitochondria, and appear to be key regulators of cell fate determination in beige adipocyte maturation.

Discussion

Adipose tissue dysfunction plays a significant role in obesity and related metabolic diseases. Recently, energy-dissipating beige adipocytes have emerged as a promising therapeutic target for treating metabolic diseases (Lizcano, 2019; Tsagkaraki et al., 2021). Studies employing MS-based proteomics to investigate fundamental regulatory mechanisms in adipogenesis and adipose biology have implicated alternative splicing (Chao et al., 2021; Park et al., 2020), PTMs such as phosphorylation (Ahmad et al., 2020; Shinoda et al., 2015), and proteolysis (Chung et al., 2010; Han et al., 2009; Lilla et al., 2002). However, little is known about the temporal profiles of these factors at the proteome scale in beige adipocyte differentiation. Here, we applied global proteome and N-terminome profiling at five time points to identify transcriptional, translational and post-translational modifications associated with beige adipocyte differentiation.

In this study, we employed our previously established N-terminal peptide enrichment method (Tsumagari et al., 2021b) to investigate N-terminome alterations during beige adipocyte differentiation. We introduced SCX-based enrichment to isolate protein N-terminal peptides with high specificity, including both acetylated and non-acetylated peptides, and confirmed that this method enables the isolation of enzymatically cleaved peptides. By considering semi-specific cleavage for identification, more than 4,200 Neo-N-terminal peptides were identified independently of N-terminal acetylation. If we replaced SCX-HPLC with SCX-StageTip, we could identify only 378 TrypN-digested peptides, which were regarded as contaminants, so that the method achieved an enrichment specificity of up to 95%. In contrast to N-TAILS, a method for purifying N-terminal peptides with a specificity of about 90-40% (Pablos et al., 2021; Prudova et al., 2016; Yeom et al., 2017), our method provides a simple and rapid approach without compromising the specificity of N-terminal enrichment.

The results of N-terminome data acquisition depend on the methods used to prepare protein sequence databases, such as minimizing the size of the database (Willems et al., 2017) and customizing the database for translational initiation (Nakahigashi et al., 2016). Here, protein N-terminal peptides were identified against the full-length protein sequence from Uniprot, covering protein isoforms and computer-annotated proteoforms. The database could be extended using ribo-seq datasets or customized databases to identify protein N-termini from upstream open reading frames or ncRNAs in the future.

Protein N-terminal methionine excision and acetylation are ribosome-associated co-translational modifications mediated by methionine aminopeptidases and acetyltransferases. We identified 591 and 1,220 acetylated protein N-termini at the 1st methionine and 2nd amino acid, respectively. Interestingly, among 27 peptides with acetylation at the 3rd amino acid, alanine was the 2nd amino acid in 13 peptides, suggesting that additional aminopeptidases may remove the second amino acid after first Met excision prior to acetylation.

All unmodified NeoNt peptides were considered to be proteolytic peptides generated by exo- or endopeptidase, but without control experiments it was difficult to distinguish whether the unmodified NeoNt peptides were cleaved by exopeptidases or endopeptidases. A significant feature of this group was sequentially cleaved peptides. It is noteworthy that most sequentially cleaved peptides show high abundance, suggesting that we still need to improve our N-terminal peptide enrichment method to uncover low abundant proteolytic peptides during protein degradation. Prudova et al. reported the degradation of cathepsins under *in vivo* conditions, clarifying the processing of specific substrates (Prudova et al., 2016). For peptides lacking sequential cleavage and for the most extended sequence among sequentially cleaved peptides, we annotated them as the primary proteolytic products and subjected them to substrate-protease analysis.

The temporal profiles of proteases and primary proteolytic products were retrieved from the quantitative global proteome and N-terminome, respectively. The extracellularly localized products were removed based on our experimental design before inferring the protease-substrate relationship. We evaluated the sequence specificity at cleavage sites of selected substrates by protease-based PWM scoring, which we previously employed to evaluate protein kinase-substrate phosphorylation (Imamura et al., 2017) and metalloproteinase-substrate cleavage (Tsumagari et al., 2021a). The PWM scoring successfully identified 14 cleavage sites in mitochondrial proteins recognized by protease Pmpcb and provided authentic site information (Figure 5C and 5D). However, the current protease database MERPOS is limited in terms of the number of substrates and the specificity of proteolysis. Once a comprehensive protease database has been constructed, it should be possible to annotate the responsible protease for given substrates.

The N-terminome analysis revealed key roles of highly active proteases in

beige adipocyte differentiation. The activation of lysosomal protease (Cstd), extracellular protease (Plg), and mitochondrial processing protease (Pmpbc) was necessary for thermogenesis. Cstd was markedly increased by phosphatidylinositol 3-kinase (PI3K) activation in brown adipocytes during differentiation, and Cstd proteolysis controls lysosome-dependent UCP1 turnover and mitochondrial translation (Moazed and Desautels, 2002). Plg has a role in regulating the expression of PPARγ and other adipogenic molecules in the adipogenic program (Samad et al., 2022). Proteolysis of Pmpbc was essential for controlling protein stability and transport into the mitochondria and is a key step in mitochondrial biogenesis (Anand et al., 2013). Although the relationships between proteases and their endogenous substrates are not yet completely understood, they could be resolved in the future by protease targeting analysis.

In conclusion, the CHAMP method identifies indispensible proteases promoting beige adipocyte differentiation. Our multiomics approach enables systematic investigation of the endogenous relationship between proteases and their substrates, and should be applicable to other models and also to proteogenomics.

Limitations of the study

In this study, we analyzed the N-terminus of the protein, focusing on changes in splicing, translation initiation, and post-translational proteolysis. Therefore, we focused only on splicing products that alter the protein N-termini and missed those that alter the internal domains or C-termini of the proteins. For TIS, we also identified and searched only the database consisting of canonical protein sequences, and missed the neoNt peptides generated from the untranslated regions of the canonical sequences. In addition, post-translational proteolysis was also missed for changes at the C-terminus. The samples used were cell pellets, and the supernatant fraction was not analyzed. If the supernatant contained cleaved fragments from post-translational proteolysis, they could have been omitted from the present results. The effects at these splicing, translation, and proteolysis steps on adipocyte differentiation are the subject of future studies.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.xxxx/.

Acknowledgments

We would like to thank all lab members for fruitful discussions. The immortalized preadipocytes derived from inguinal white adipose tissue were kindly provided by Prof. Sakai (The University of Tokyo/Tohoku University) according to the protocol from Prof. Kajimura (Harvard Medical School).

This work was supported by the JST Strategic Basic Research Program, CREST (grant No. 18070870) to Y.I., by a JSPS Grants-in-Aid for Scientific Research (No. 21H02459) to Y.I., and the Ministry of Science of Technology

grants (Taiwan, MOST 109-2320-B-038-048; MOST 111-3111-Y-016-004) to H-Y.C.

Author contribution

Conceptualization, Y.I.; Methodology, C-H.C., H-Y.C., and Y.I.; Investigation, C-H.C. and H-Y.C.; Writing – Original Draft, C-H.C.; Writing –Review & Editing, C-H.C., H-Y.C., and Y.I.; Funding Acquisition, H-Y.C. and Y.I.; Supervision, Y.I.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Main figure legends

Figure 1. Global proteome and N-terminome profiling during adipocyte differentiation. (A) Experimental workflow for the global proteome and N-terminome. Temporal expression profiles of proteins and peptides during adipocyte differentiation were measured by nanoLC/MS/MS. The canonical N-terminal (CanNt) and neo-N-terminal (NeoNt) peptides were enriched by CHAMP prior to nanoLC/MS/MS. (B) Cells were collected every two days during beige adipocyte differentiation and analyzed by global proteomics and N-terminomics. The protein expression data obtained by global proteomics and the peptide expression data obtained by N-terminomics are shown in color from green to red according to the normalized abundance from low to high across the five time points. The N-terminome was divided into CanNt and NeoNt peptides. The relationship between expression level and differentiation time points was evaluated by calculating Pearson's correlation coefficient (PCC) and is shown in blue to red corresponding to PCC from −1 to 1. (C) Comparison of identified proteins in the global proteome and N-terminome datasets.

Figure 2. Characterization of N-terminomics datasets. (A) Decision tree of peptide annotation for N-terminome datasets. (B) Annotation summary of protein N-terminal peptides identified by N-terminomics. Identification numbers of unique peptides, N-termini and proteins are listed. (C) Sequence logo showing the amino acid frequencies at each position of the CanNt peptides. (D) Sequence logo showing the amino acid frequencies at each position of the acetylated NeoNt peptides.

Figure 3. Temporal profiles and functionally enriched terms of global proteins and CanNt peptides. (A) PCC between temporal profiles of protein expression and CanNt. We found a total of 797 commonly identified and also quantified proteins in the global proteome and CanNt peptides. PCCs larger than 0.5, less than −0.5, and between −0.5 and 0.5 were defined as positively coexpressed, negatively coexpressed and independent pairs. Expression levels

with increasing or decreasing patterns during differentiation are annotated in red or green, respectively. (B) Temporal profiles of proteins and CanNt peptides in each group. (C) Functional enrichment analysis based on UniProt keywords in the global proteome and CanNt peptides. The terms acetylation and phosphorylation, which were commonly enriched in all groups, have been removed.

Figure 4. Temporal profiles of alternatively spliced variants during adipocyte differentiation. (A) Temporal profiling of protein N-terminal peptides expression, according to canonical and isoform (Pym1 and Tnfaip8). (B) Expression levels of selected mRNA variants (orange) of Pym1 and Tnfaip8 were detected by qPCR. Primers (blue blocks) were located upstream and downstream of the TIS regions of the coding sequences for the canonical protein and isoforms (green). (C) Relative expression of Pym1 variants and Tnfaip8 variants normalized to the level of 36b2.

Figure 5. Proteases and their substrates during beige adipocyte differentiation. (A) Differentiation-dependent clusters of representative substrates. Two clusters with an increasing tendency showed a distinct cleavage preference. (B) Expression patterns of protease from global proteome. The heatmap from green to red corresponds to the normalized abundance from low to high. (C) Distribution of PWM scores for proteolytic peptides from group 1 substrates with protease Pmpcb (top) and group 2 substrates with protease Plg (bottom). (D) Potential substrates of Pmpcb. Cleaved site, the position of cleavage; N/A, without cleavage; *experimental result from related homologs; $ cleavage predicted from Uniprot. (E) The mRNA expression of selected adipocyte differentiation markers with specific siRNA knockdown (Day 8). *P<0.5; **P<0.1; ***P<0.05. Adipoq, adiponectin; Fasn, fatty acid synthase; Fabp4, fatty acid binding protein 4; Pparg2, peroxisome proliferator-activated receptor gamma; Cidea, cell death activator CIDE-A; Pgc1a, peroxisome proliferator-activated receptor gamma coactivator 1-alpha; Ucp1, mitochondrial brown fat uncoupling protein 1.

STAR Methods text

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-Pparγ	Cell Signaling Technology	Cat#: 2435S
Rabbit monoclonal anti-C/ebpα	Cell Signaling Technology	Cat#: 8178S
Rabbit monoclonal anti-Acetyl-CoA Carboxylase	Cell Signaling Technology	Cat#: 3676S
Rabbit monoclonal anti-Fatty Acid Synthase	Cell Signaling Technology	Cat#: 3180S
Biological samples		
Mouse preadipocytes	(Aune et al., 2013)	N/A
Chemicals, peptides, and recombinant proteins		
Ammonium bicarbonate	FUJIFILM Wako	Cat#: 012-21745
2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride	FUJIFILM Wako	Cat#: 015-20995
Sodium deoxycholate (SDC)	FUJIFILM Wako	Cat#: 194-08311
Sodium lauroyl sarcosinate (SLS)	FUJIFILM Wako	Cat#: 194-10381
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	FUJIFILM Wako	Cat#: 040587
2-chloroacetamide (CAA)	FUJIFILM Wako	Cat#: 036-09765
Calcium chloride	FUJIFILM Wako	Cat#: 035-24995
Ethyl acetate	FUJIFILM Wako	Cat#: 055-03115

Acetonitrile	FUJIFILM Wako	Cat#: 018-19853
Acetic acid	FUJIFILM Wako	Cat#: 018-20061
Trifluoroacetic acid	FUJIFILM Wako	Cat#: 205-14101
Insulin	FUJIFILM Wako	Cat#: 093-06476
TrypN	PROTIFI	Cat#: K01-20
Dulbecco's Modified Eagle Medium (DMEM)	Thermo Fisher Scientific	Cat#: 11965092
Rosiglitazone	Sigma-aldrich	Cat#: PHR2932
Indomethacin	Sigma-aldrich	Cat#: PHR1247
Dexamethasone	Sigma-aldrich	Cat#: PHR1526
3-Isobutyl-1-methylxanthine	Sigma-aldrich	Cat#: I7018
3,3',5-triiodo-L-thyronine (T3)	Sigma-aldrich	Cat#: T-075
Empore SDB – XC membrane disks	CDS	Cat#: 98-0604-0224-9
Empore cation exchange – SR membrane disks	CDS	Cat#: 98-0604-0232-2
Critical commercial assays		
Bicinchoninic acid (BCA) assay kit	Thermo Fisher Scientific	Cat#: 23227
ZR Urine RNA Isolation Kit	Zymo Research	Cat#: R1038
HiScript II Q RT Supermix	Vazyme Biotech	Cat#: R222
Fast SYBR™ Green Master Mix	Thermo Fisher Scientific	Cat#: 4385612
Deposited data		

jPOST Repository	This paper	JPST001099 (PXD024334)
Oligonucleotides		
Primer: Pgc1a Forward: TGAGGACCGCTAGCAAGTTT	IDT Corporation	N/A
Primer: Pgc1a Reverse: TGTAGCGACCAATCGGAAAT	IDT Corporation	N/A
Primer: Ucp1 Forward: CTTTGCCTCACTCAGGATTGG	IDT Corporation	N/A
Primer: Ucp1 Reverse: ACTGCCACACCTCCAGTCATT	IDT Corporation	N/A
Primer: Cidea Forward: CATGATCTTGGAAAAGGGACAG	IDT Corporation	N/A
Primer: Cidea Reverse: ATCGTGGCTTTGACATTGAGAC	IDT Corporation	N/A
Primer: Fabp4 Forward: TGATGCCTTTGTGGGAACCT	IDT Corporation	N/A
Primer: Fabp4 Reverse: ATGCCTGCCACTTTCCTTGT	IDT Corporation	N/A
Primer: Adipoq Forward: GCCGCTTATGTGTATCGCTC	IDT Corporation	N/A
Primer: Adipoq Reverse: AGAGGCTCACCTTCACATCTT	IDT Corporation	N/A
Primer: Fasn Forward: CGTGCTCATCCACTCAGGTT	IDT Corporation	N/A
Primer: Fasn Reverse: CCTGGAGGTATGCTCGCTTC	IDT Corporation	N/A
Primer: 36b4 Forward: GGGCATCACCACGAAAATCTC	IDT Corporation	N/A
Primer: 36b4 Reverse: CTGCCGTTGTCAAACACCT	IDT Corporation	N/A
Primer: Pym_v1 Forward: TCAGTCATGGAAACCGCCAG	IDT Corporation	N/A
Primer: Pym_v1 Reverse: GGGTCTCTGTGTTGAGGCAA	IDT Corporation	N/A

Primer: Pym_v2 Forward: CGGGTTCTCAAGATGGCGA	IDT Corporation	N/A
Primer: Pym_v2 Reverse: AATGTACTTGCCGCCGGTC	IDT Corporation	N/A
Primer: Tnfaip8_v1 Forward: GCGCACATCTCACTCGTCTA	IDT Corporation	N/A
Primer: Tnfaip8_v1 Reverse: CTGAACGGCCAGGTTTTTGG	IDT Corporation	N/A
Primer: Tnfaip8_v2 Forward: CAGTCCAACTCCGGGGAAC	IDT Corporation	N/A
Primer: Tnfaip8_v2 Reverse: GACATCTGTAGCCATCGGAGG	IDT Corporation	N/A
ON-TARGETplus Non-targeting Control Pool	Dharmacon	Cat#: D-001810-10
ON-TARGETplus-smartpool mouse Ctsd siRNA	Dharmacon	Cat#: L-051673-01
ON-TARGETplus-smartpool mouse Plg siRNA	Dharmacon	Cat#: L-062291-01
ON-TARGETplus-smartpool mouse Pmpcb siRNA	Dharmacon	Cat#: L-051430-01
Software and algorithms		
Mascot 2.6.1	PMID: 10612281	https://www.matrixscience.com/
Maxquant 1.6.2.10	PMID: 27809316	https://www.maxquant.org/
Perseus 1.6.14.0	PMID: 27348712	https://maxquant.net/perseus/
WebLogo 2.8.2	PMID: 15173120	https://weblogo.berkeley.edu/
PreTIS	PMID: 27768687	https://service.bioinformatik.uni-saarland.de/pretis/
iceLogo	PMID: 19876014	https://iomics.ugent.be/icelogoserver/
Other		

Protein knowledgebase	UniProtKB	https://www.uniprot.org/ (version 2019_2, Mus musculus, 93,387 sequences)
Resource for peptidases	MEROPS PMID: 29145643	https://www.ebi.ac.uk/merops/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Yasushi Ishihama (yishiham@pharm.kyoto-u.ac.jp).

Materials availability

This study did not generate new unique reagents or materials. The TrypN can be replaced with LysargiNase (Merck Millipore).

Data and code availability

- All LC/MS/MS data that support the findings of this study have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the jPOST partner repository (https://jpostdb.org) (Okuda et al., 2017). The accession number is listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Beige adipocyte differentiation

To differentiate beige adipocytes, the preadipocytes from the inguinal white adipose tissue were cultured in DMEM supplemented with 10% Fetal Bovine Serum (FBS) to a confluency of 80% to 90% and incubated in the induction medium consisting of 0.5 μM rosiglitazone, 125 μM indomethacin, 2 μg/mL dexamethasone, 0.5 mM 3-Isobutyl-1-methylxanthine, 5 μg/mL insulin, and 1 nM T3 in DMEM supplemented with 10% FBS. After 2 days of induction, medium was renewed with the maintenance medium (DMEM containing 10% FBS, 0.5 μM rosiglitazone, 5 μg/mL insulin, and 1 nM T3) every 2 days. Preadipocytes were fully differentiated into beige cells after 8 days after induction.

METHOD DETAILS

Cell culture and protein extraction

To capture the alterations during the beige adipocyte maturation, we harvested the cells at day 0, day 2, day 4, day 6, and day 8 after inducing cell differentiation. We collected the cells from three biological replicates in the lysis buffer containing 12 mM SDC, 12 mM SLS, 10 mM TCEP, 40 mM CAA, and protease inhibitors in 100 mM Tris buffer (pH 8.5). The lysates were heated at 95 °C for 5 minutes and sonicated in iced water for 20 minutes. The protein concentration was determined using the BCA protein assay.

Protein digestion

The proteins were digested by a previously described method with some modifications (Chang et al., 2021). Equal amounts of protein were subsequently 10-fold diluted with 10 mM $CaCl_2$ and digested by TrypN (1: 50 w/w) overnight at 37 °C. After enzymatic digestion, an equal volume of ethyl acetate was added to the protein digests, and the mixture was acidified by adding trifluoroacetic acid to a final concentration of 0.5%. The mixture was vortexed for 1 min and centrifuged at 15,700 x g for 2 minutes to separate ethyl acetate from the mixture. The aqueous part was collected and desalted by using StageTips filled with poly(styrene-divinylbenzene) SDB-XC disk membrane (Rappsilber et al., 2007). The proteolytic peptides were quantified by LC-UV at 214 nm with standard BSA peptides and stored in 80% ACN and 0.5% TFA at −20 °C until nanoLC/MS/MS analysis or N-terminal peptides enrichment.

N-terminal peptides enrichment by SCX StageTip separation

N-terminal peptides were enriched as described previously (Chang et al., 2021). The strong cation exchange (SCX) StageTip was prepared by packing a layer of SCX disk membrane into 200-µL tips. The activation buffer consisted of 500 mM KCl, 30% ACN, and 7.5 mM phosphate buffer, pH 2.2, and the elution buffer consisted of 12.5 mM KCl, 30% ACN, and 7.5 mM phosphate buffer, pH 2.2. Conditioning and equilibration were done by sequentially passing 100 µL buffers and centrifuging at 1000 × g for 1 min of the following reagents: methanol, activation buffer and elution buffer. Twenty µg of digested peptides were loaded into equilibrated SCX StageTips by centrifugation. We collected the flowthrough and sequentially eluted peptides with 100 µL of elution buffer into the same tube as enriched N-terminal peptides. The samples were evaporated by SpeedVac SPD121P (Thermo Scientific), and resuspended in 50 uL of 0.1% TFA and desalted by using StageTips with SDB-XC disk membrane. The desalted samples were dried in a vacuum evaporator and dissolved in 0.1% TFA and 4% ACN, and subjected to nanoLC/MS/MS analysis.

NanoLC/MS/MS analysis

NanoLC/MS/MS analyses were performed on an Orbitrap Fusion LUMOS (Thermo Fisher Scientific, Waltham, MA) mass spectrometer, which was connected to a Thermo Scientific Ultimate 3000 system (Germering, Germany)

and a HTC-PAL autosampler (CTC Analytics, Zwingen, Switzerland). Peptides were separated by self-pulled needle columns (150 mm length × 100 μm ID, 6 μm opening) packed with Reprosil-C18 3 μm reserved phase material (Dr. Maisch, Ammerbuch, Germany). The injection volume was 5 μL and the flow rate was 500 nL/min. The mobile phases consisted of (A) 0.5% acetic acid and (B) 0.5% acetic acid and 80% ACN. A three-step linear gradient of 10-40% B in 100 min, 40-50% B in 10 min, 50-99% B in 5 min, and 99% B for 5 min was employed for N-terminal peptides and 8-35% B in 100 min, 35-50% B in 10 min, 50-99% B in 5 min, and 99% B for 5 min was employed for global proteome samples. A 2400 V was applied to spray voltage, the mass scan range was *m/z* 300-1500, with an automatic gain control value of 1.00e + 06, a max injection time of 50 ms and detected at a mass resolution of 60,000 at *m/z* 200 in orbitrap analyzer. The top ten precursor ions were selected in each MS scan for subsequent MS/MS scans with an automatic gain control value of 5.00e + 04 and a max injection time of 300 ms. Dynamic exclusion was set for 20 s. The normalized HCD was set to 30 and detected at a mass resolution of 15,000 at *m/z* 200 in Orbitrap analyzer. A lock mass of *m/z* 445.12 was set to obtain constant mass accuracy during the analyses.

Proteomics Data Processing

For N-terminome samples, the peak list mgf file of each run was generated from the MS/MS spectra by MaxQuant. The peptides and proteins were identified by Mascot (Matrix Science, London, U.K.) and MaxQuant against the Universal Protein Resource Knowledgebase (UniProtKB, version 2019_2, 93,387 sequences; 25,228 entries form Swiss-Prot including isoforms and 68,159 entries from TrEMBL (Translated EMBL Nucleotide Sequence Data Library)). In MaxQuant searches, 20 ppm and 4.5 ppm of mass tolerance was set for first and main search in both precursors and fragments, respectively. N-terminally semi-TrypN specificity was allowed for up to 2 missed cleavages for each peptide with a minimum length of 7 amino acids for specific cleavage, 8 amino acids for semi-specific cleavage. The maximum peptide mass was limited to 4,600 Da and peptides were identified through the match-between-runs function. In Mascot searching, a precursor mass tolerance of 10 ppm, a fragment ion mass tolerance of 20 ppm, C- and N-terminally semi-TrypN specificity allowing up to 2 missed cleavages per peptide with a minimum length of 7 amino acids for specific cleavage, 8 amino acids for semi-specific cleavage, and a removal of C-terminally semi-specific peptides were set. Both search engines applied the same parameters: fixed carbamidomethylation of cysteine, and variable methionine oxidation and N-terminal acetylation. A reversed sequence library was employed to control the false discovery rate (FDR) less than 1% in sequence level for both MaxQuant and Mascot.

For the global proteome dataset, the peptides and proteins were identified by MaxQuant against UniProtKB (version 2019_2, 93,387 sequences). A 20 ppm mass tolerance for precursor and fragment, strict TrypN specificity allowing for

up to 2 missed cleavages in mono proteolytic peptides, 7 amino acids for the required minimum peptide sequence length, 4600 Da for the maximum peptide mass, less than 1% FDR for the peptide spectrum matches and protein group identifications were set. Carbamidomethylation of cysteine was set as a fixed modification, methionine oxidation and protein N-terminal acetylation were allowed as variable modifications. Match-between-runs and label-free protein quantification were performed via MaxLFQ algorithm with at least one identified peptide.

Peptides characterization and protein identifier assignment for the N-terminomic dataset

To improve the identified sequence coverage, we utilized Mascot and MaxQuant for N-terminome annotation. The identified peptides were mainly classified into three categories: the canonical N-termini (CanNt), the neo-N-termini (NeoNt), and the internal peptides. Canonical N-terminus is the protein N-terminal peptide matching to the first or second amino acid of the known translational initiation sites (TISs) recorded in the sequence database. In the case of sequences matching multiple proteins, we defined the master protein by the following rule: the sequence matching the protein from the first one or the first two amino acids was defined as CanNt; otherwise, the sequence was annotated as NeoNt, or the internal peptides. If the sequences were semi-cleaved at N-terminal, we defined them as NeoNt. Those sequences might be the products derived from the endopeptidase or exopeptidase cleavage and also possible to be the protein N-termini of unreported RNA isoforms or unreported TISs. For peptides containing identical C-terminal sequence but sequentially cleaved from the N-termini, we also assigned them as NeoNt and proposed that these peptides were initially cleaved by endopeptidase and processed by exopeptidase in advance. Lastly, we defined the group of internal peptides from the peptides with a first N-terminal residue of lysine or arginine, resulting from TrypN digestion at N-termini. The number of internal peptides represented as the impurity of the N-terminal peptide enrichment. We assigned the protein identifier from the longest protein isoform in the UniProt database if any ambiguous accession numbers existed (more detailed descriptions are in Figure S1D).

Quantitative analysis of proteome and N-terminome data

Data processing and statistical analysis were performed using Perseus. To quantify the N-terminome, we analyzed the sequences commonly identified in Mascot and MaxQuant searches. We used the peptide intensities and protein MaxLFQ values to quantify the N-terminome and global proteome, respectively. The values were logarithm-transformed, the proteins or peptides with at least two quantifiable values among three technical replicates at one time point out of 5 points were included, and the missing values were replaced by the random values with a normal distribution (width = 0.3, shift = 1.8), assuming these values were close to the minimum of detection.

Position weight matrix (PWM) for protease-substrate prediction

Flanking sequences surrounding substrate cleavage sites for protease binding were obtained from the MEROPS database. Sequences including non-natural amino acids were excluded. The constructed protease-dependent PWM was applied to PWM score computation for NeoNt peptides as described previously (Imamura et al., 2017; Tsumagari et al., 2021a). Briefly, the PWM score was computed by summing log2 weights for the flanking ±4 residues of each cleavage site.

siRNA transfections

Preadipocytes were seeded in the 12-well plates and induced to differentiate by incubating the cells in the beige adipocyte differentiation medium as described above for 2 days. 25 pmol siRNA targeting Ctsd, Plg, Pmpcb, and the non-targeting control was transfected into cells by Lipofectamine RNAiMAX for 2 days according to the manual. The medium was changed every 2 days after transfection with the maintained medium. On day 8 after differentiation induction, cells were harvested in TRI reagent for RNA purification.

RNA isolation and real-time PCR (qPCR)

The total RNA was purified with the RNA Isolation Kit. One μg RNA was reverse transcribed into cDNA using HiScript II Q RT Supermix. The expression of mRNA in differentiated beige adipocytes was determined using the Fast SYBR™ Green Master Mix on a StepOnePlus™ System (Thermo Fisher Scientific). The primers were listed in the key resources table. The fold change of mRNAs relative to 36b4 and siControl was calculated using the 2−ΔΔCt method. Differential expression was determined by Student's t-test. $^{*}P < 0.05$; $^{***}P < 0.001$.

Supplemental table (excel) title and figure legends

Table S1. Quantified peptide list of CanNt and NeoNt peptides.xlsx

(A) List of identified unique peptides from N-terminome, (B) Quantified protein list from global proteome, (C) ChIP-Atlas transcription factor enrichment analysis, (D) Quantified peptide list of CanNt and NeoNt peptides, (E) List of acetylated NeoNt peptides without N-terminal preceding Met, (F) Quantified pairs of acetylated NeoNt and CanNt peptides, (G) Quantified peptide list of putative endogenous protease substrates.

Figure S1. Data processing steps for the N-terminome and global proteome datasets, related to Figure 2. (A) Data processing steps for the N-terminome datasets. MS raw files were searched with N-terminal semi-specific TrypN cleavage by Mascot and Maxquant. The peptide and protein annotations were described in Fig S1D. In terms of peptide quantification, only high-confident peptides (identified by both Mascot and MaxQuant) were included, and peptides

with at least two quantitative values in triplicates at one time point were employed for temporal analyses even if values at other time points were missing. All N-terminal peptides were divided into CanNt and NeoNt. (B) Data processing steps for the global proteome datasets. MS raw files of the global proteome datasets were searched with TrypN-specific cleavage by MaxQuant. The MaxLFQ was used to quantify proteins with at least one quantified peptide per protein. Proteins with at least two quantitative values in triplicates at one time point were employed for temporal analyses even if values at other time points were missing. (C) Identification of unique peptides from the N-terminome datasets via Mascot and MaxQuant. (D) Annotation of N-terminal peptides and their master proteins. The CanNt peptide was determined by matching its N-terminal residue to the canonical N-terminal residue of the protein sequence. Semi-specific N-termini peptides are annotated as NeoNt peptides. Sequences with N-terminal K/R are annotated as internal peptides. For proteins with multiple protein entries, the protein of CanNt peptide is defined as the master protein in preference to the Swiss-Prot canonical protein, Swiss-Prot protein isoform, and TrEMBL protein. If more than one accession numbers were listed in the same category, the alphabetical order was used. For NeoNt peptides and proteins by global proteomics, we gave priority to proteins identified in common with CanNt, otherwise applied the same rules as for CanNt peptides.

Figure S2. Temporal profiles of partial N-terminome and global proteome dataset, related to Figure 3. (A) Temporal profiles of mRNA processing/splicing, ribonucleoproteins, translation initiation factors, ribosomal proteins and mitochondrial ribosomal proteins identified in the global proteome. Each category was defined according to the UniProt keyword and the protein quantification was performed by MaxLFQ and Z-score normalized across time points. (B) Temporal profiles of proteins and CanNt peptides in each group. (C) The UniProt keywords functional enrichment analysis of the global proteome and CanNt peptides. Note that the terms acetylation and phosphorylation, which were commonly enriched in all groups, were removed.

Figure S3. Transcriptional or translational modification of protein N-termini in adipocyte differentiation. (A) The starting site of the acetylated NeoNt-peptide is within the first 49 amino acid residues. (B) Temporal profiles of 30 CanNt-peptides (black box) and an acetylated putative TIS peptide (red) during adipocyte differentiation. (C) Sequence logo of the amino acid frequency for acetylated NeoNt peptides starting with the third residue of the canonical protein sequence (top). Of these 27 acetylated NeoNt peptides, there were 13 peptides with the removal of the first two [Met-Ala] residues (bottom).

Figure S4. Temporal profiling of substrates during beige adipocyte differentiation, related to Figure 4. (A) Frequency of the number of unmodified protein Neo-N-termini generated by sequential cleavage. Data was binned by the

rank of protein expression abundance. Proteins without, with one, or with more than one N-terminal sequential cleavage was indicated by 0, 1, and > 1. (B) Expression patterns of potential substrates from NeoNt peptides. Five clusters were determined by hierarchical clustering. The heatmap from green to red corresponds to the normalized abundance from low to high. (C) Dynamic clustering of representative substrates and the amino acid frequency of cleaved sites.

REFERENCES

Adachi, J., Kumar, C., Zhang, Y., and Mann, M. (2007). In-depth analysis of the adipocyte proteome by mass spectrometry and bioinformatics. Mol. Cell. Proteomics *6*, 1257–1273.

Ahmad, B., Serpell, C.J., Fong, I.L., and Wong, E.H. (2020). Molecular Mechanisms of Adipogenesis: The Anti-adipogenic Role of AMP-Activated Protein Kinase. Front Mol Biosci *7*, 76.

Almagro Armenteros, J.J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance *2*. https://doi.org/10.26508/lsa.201900429.

Anand, R., Langer, T., and Baker, M.J. (2013). Proteolytic control of mitochondrial function and morphogenesis. Biochim. Biophys. Acta *1833*, 195–204.

Aune, U.L., Ruiz, L., and Kajimura, S. (2013). Isolation and differentiation of stromal vascular cells to beige/brite cells. J. Vis. Exp. *73,* 50191.

Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Analytical and Bioanalytical Chemistry *404*, 939–965. https://doi.org/10.1007/s00216-012-6203-4.

Bekker-Jensen, D.B., Kelstrup, C.D., Batth, T.S., Larsen, S.C., Haldrup, C., Bramsen, J.B., Sørensen, K.D., Høyer, S., Ørntoft, T.F., Andersen, C.L., et al. (2017). An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. Cell Syst *4*, 587–599.e4.

Bludau, I., Frank, M., Dörig, C., Cai, Y., Heusel, M., Rosenberger, G., Picotti, P., Collins, B.C., Röst, H., and Aebersold, R. (2021). Systematic detection of functional proteoform groups from bottom-up proteomic datasets. Nat. Commun. *12*, 3810.

Bonissone, S., Gupta, N., Romine, M., Bradshaw, R.A., and Pevzner, P.A. (2013). N-terminal protein processing: a comparative proteogenomic analysis. Mol. Cell. Proteomics *12*, 14–28.

Chang, C.-H., Chang, H.-Y., Rappsilber, J., and Ishihama, Y. (2021). Isolation of Acetylated and Unmodified Protein N-Terminal Peptides by Strong Cation Exchange Chromatographic Separation of TrypN-Digested Peptides. Mol. Cell. Proteomics *20*, 100003.

Chao, Y., Jiang, Y., Zhong, M., Wei, K., Hu, C., Qin, Y., Zuo, Y., Yang, L., Shen, Z., and Zou, C. (2021). Regulatory roles and mechanisms of alternative RNA splicing in adipogenesis and human metabolic health. Cell & Bioscience *11*, 66. https://doi.org/10.1186/s13578-021-00581-w.

Chen, L., Shan, Y., Weng, Y., Sui, Z., Zhang, X., Liang, Z., Zhang, L., and Zhang, Y. (2016). Hydrophobic Tagging-Assisted N-Termini Enrichment for In-Depth N-Terminome Analysis. Anal. Chem. *88*, 8390–8395. .

Choi, S., Goswami, N., and Schmidt, F. (2020). Comparative Proteomic Profiling of 3T3-L1 Adipocyte Differentiation Using SILAC Quantification. J. Proteome Res. *19*, 4884–4900. .

Chung, S.S., Ahn, B.Y., Kim, M., Choi, H.H., Park, H.S., Kang, S., Park, S.G., Kim, Y.-B., Cho, Y.M., Lee, H.K., et al. (2010). Control of adipogenesis by the SUMO-specific protease SENP2. Mol. Cell. Biol. *30*, 2135–2146. .

Deutsch, A., Feng, D., Pessin, J.E., and Shinoda, K. (2020). The Impact of Single-Cell Genomics on Adipose Tissue Research. Int. J. Mol. Sci. *21*. https://doi.org/10.3390/ijms21134773.

Du, Y., Wang, Y., Xu, Q., Zhu, J., and Lin, Y. (2021). TMT-based quantitative proteomics analysis reveals the key proteins related with the differentiation process of goat intramuscular adipocytes. BMC Genomics *22*, 417. .

Fiszbein, A., and Kornblihtt, A.R. (2017). Alternative splicing switches: Important players in cell differentiation. Bioessays *39*. https://doi.org/10.1002/bies.201600157.

Forner, F., Kumar, C., Luber, C.A., Fromme, T., Klingenspor, M., and Mann, M. (2009). Proteome differences between brown and white fat mitochondria reveal specialized metabolic functions. Cell Metab. *10*, 324–335. .

Fukasawa, Y., Tsuji, J., Fu, S.-C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol. Cell. Proteomics *14*, 1113–1126. .

Gao, Y., and Yates, J.R. (2019). Protein Analysis by Shotgun Proteomics. In Mass Spectrometry‐Based Chemical Proteomics,Wiley, Hoboken, NJ , pp. 1–38.

Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G.R., and Vandekerckhove, J. (2003). Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. Nat. Biotechnol. *21*, 566–569.

Giansanti, P., Samaras, P., Bian, Y., Meng, C., Coluccio, A., Frejno, M., Jakubowsky, H., Dobiasch, S., Hazarika, R.R., Rechenberger, J., et al. (2022). Mass spectrometry-based draft of the mouse proteome. Nat. Methods *19*, 803–811.

Giralt, M., and Villarroya, F. (2013). White, Brown, Beige/Brite: Different Adipose Cells for Different Functions? Endocrinology *154*, 2992–3000. https://doi.org/10.1210/en.2013-1403.

Greene, A.W., Grenier, K., Aguileta, M.A., Muise, S., Farazifard, R., Haque, M.E., McBride, H.M., Park, D.S., and Fon, E.A. (2012). Mitochondrial processing peptidase regulates PINK1 processing, import and Parkin recruitment. EMBO Rep. *13*, 378–385.

Guo, L., Li, X., and Tang, Q.-Q. (2015). Transcriptional regulation of adipocyte differentiation: a central role for CCAAT/enhancer-binding protein (C/EBP) β. J. Biol. Chem. *290*, 755–761.

Han, J., Luo, T., Gu, Y., Li, G., Jia, W., and Luo, M. (2009). Cathepsin K regulates adipocyte differentiation: possible involvement of type I collagen degradation. Endocr. J. *56*, 55–63.

Harms, M., and Seale, P. (2013). Brown and beige fat: development, function and therapeutic potential. Nat. Med. *19*, 1252–1263.

Helbig, A.O., Gauci, S., Raijmakers, R., van Breukelen, B., Slijper, M., Mohammed, S., and Heck, A.J.R. (2010). Profiling of N-acetylated protein termini provides in-depth insights into the N-terminal nature of the proteome. Mol. Cell. Proteomics *9*, 928–939.

Henriques, F., Bedard, A.H., Guilherme, A., Kelly, M., Chi, J., Zhang, P., Lifshitz, L.M., Bellvé, K., Rowland, L.A., Yenilmez, B., et al. (2020). Single-Cell RNA Profiling Reveals Adipocyte to Macrophage Signaling Sufficient to Enhance Thermogenesis. Cell Rep. *32*, 107998.

Ikeda, K., Maretich, P., and Kajimura, S. (2018). The Common and Distinct Features of Brown and Beige Adipocytes. Trends in Endocrinology & Metabolism *29*, 191–200.

Imamura, H., Wagih, O., Niinae, T., Sugiyama, N., Beltrao, P., and Ishihama, Y. (2017). Identifications of Putative PKA Substrates with Quantitative Phosphoproteomics and Primary-Sequence-Based Scoring. J. Proteome Res. *16*, 1825–1830.

Impens, F., Colaert, N., Helsens, K., Ghesquière, B., Timmerman, E., De Bock, P.-J., Chain, B.M., Vandekerckhove, J., and Gevaert, K. (2010). A quantitative proteomics design for systematic identification of protease cleavage events. Mol. Cell. Proteomics *9*, 2327–2333.

Kershaw, E.E., and Flier, J.S. (2004). Adipose tissue as an endocrine organ. J. Clin. Endocrinol. Metab. *89*, 2548–2556.

Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat. Biotechnol. *28*, 281–288.

Lilla, J., Stickens, D., and Werb, Z. (2002). Metalloproteases and adipogenesis: a weighty subject. Am. J. Pathol. *160*, 1551–1554.

Liu, X., Zheng, Z., Zhu, X., Meng, M., Li, L., Shen, Y., Chi, Q., Wang, D., Zhang, Z., Li, C., et al. (2013). Brown adipose tissue transplantation improves whole-body energy metabolism. Cell Res. *23*, 851–854.

Lizcano, F. (2019). The Beige Adipocyte as a Therapy for Metabolic Diseases. International Journal of Molecular Sciences *20*, 5058.

Mahrus, S., Trinidad, J.C., Barkan, D.T., Sali, A., Burlingame, A.L., and Wells, J.A. (2008). Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell *134*, 866–876.

Marcon, B.H., Holetz, F.B., Eastman, G., Origa-Alves, A.C., Amorós, M.A., de Aguiar, A.M., Rebelatto, C.K., Brofman, P.R.S., Sotelo-Silveira, J., and Dallagiovanna, B. (2017). Downregulation of the protein synthesis machinery is a major regulatory event during early adipogenic differentiation of human adipose-derived stromal cells. Stem Cell Res. *25*, 191–201.

Moazed, B., and Desautels, M. (2002). Differentiation-dependent expression of cathepsin D and importance of lysosomal proteolysis in the degradation of UCP1 in brown adipocytes. Can. J. Physiol. Pharmacol. *80*, 515–525.

Nakahigashi, K., Takai, Y., Kimura, M., Abe, N., Nakayashiki, T., Shiwa, Y., Yoshikawa, H., Wanner, B.L., Ishihama, Y., and Mori, H. (2016). Comprehensive identification of translation start sites by tetracycline-inhibited ribosome profiling. DNA Res. *23*, 193–201.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19,* e46255.

Okuda, S., Watanabe, Y., Moriya, Y., Kawano, S., Yamamoto, T., Matsumoto, M., Takami, T., Kobayashi, D., Araki, N., Yoshizawa, A.C., et al. (2017). jPOSTrepo: an international standard data repository for proteomes. Nucleic Acids Res. *45*, D1107–D1111.

Pablos, I., Machado, Y., de Jesus, H.C.R., Mohamud, Y., Kappelhoff, R., Lindskog, C., Vlok, M., Bell, P.A., Butler, G.S., Grin, P.M., et al. (2021). Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CL substrate degradome. Cell Rep. *37*, 109892.

Park, J.W., Fu, S., Huang, B., and Xu, R.-H. (2020). Alternative splicing in mesenchymal stem cell differentiation. Stem Cells *38*, 1229–1240.

Prudova, A., Gocheva, V., Auf dem Keller, U., Eckhard, U., Olson, O.C., Akkari,

L., Butler, G.S., Fortelny, N., Lange, P.F., Mark, J.C., et al. (2016). TAILS N-Terminomics and Proteomics Show Protein Degradation Dominates over Proteolytic Processing by Cathepsins in Pancreatic Tumors. Cell Rep. *16*, 1762–1773.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research *46*, D624–D632.

Reuter, K., Biehl, A., Koch, L., and Helms, V. (2016). PreTIS: A Tool to Predict Non-canonical 5' UTR Translational Initiation Sites in Human and Mouse. PLOS Computational Biology *12*, e1005170.

Rodriguez-López, L.A., Torre-Villalvazo, I., Aleman-Escondrillas, G., Flores-López, A., Guevara-Cruz, M., Sánchez-Tapia, M., Torre-Anaya, E.A., Martínez-López, V., Vasquez-Reyes, S., Torres-Villalobos, G.M., et al. (2021). The capacity of differentiation of stromal vascular fraction cells into beige adipocytes is markedly reduced in subjects with overweight/obesity and insulin resistance: effect of genistein. Int. J. Obes. *45*, 2471–2481.

Salomonis, N., Schlieve, C.R., Pereira, L., Wahlquist, C., Colas, A., Zambon, A.C., Vranizan, K., Spindler, M.J., Pico, A.R., Cline, M.S., et al. (2010). Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. Proc. Natl. Acad. Sci. U. S. A. *107*, 10514–10519.

Samad, F., Bai, H., Baik, N., Haider, P., Zhang, Y., Rega-Kaun, G., Kaun, C., Prager, M., Wojta, J., Bui, Q., et al. (2022). The plasminogen receptor Plg-R regulates adipose function and metabolic homeostasis. J. Thromb. Haemost. *20*, 742–754.

Shapira, S.N., and Seale, P. (2019). Transcriptional Control of Brown and Beige Fat Development and Function. Obesity *27*, 13–21. https://doi.org/10.1002/oby.22334.

Shinoda, K., Ohyama, K., Hasegawa, Y., Chang, H.-Y., Ogura, M., Sato, A., Hong, H., Hosono, T., Sharp, L.Z., Scheel, D.W., et al. (2015). Phosphoproteomics Identifies CK2 as a Negative Regulator of Beige Adipocyte Thermogenesis and Energy Expenditure. Cell Metab. *22*, 997–1008. .

Taleb, S., Cancello, R., Clément, K., and Lacasa, D. (2006). Cathepsin s promotes human preadipocyte differentiation: possible involvement of fibronectin degradation. Endocrinology *147*, 4950–4959. .

Tsagkaraki, E., Nicoloro, S.M., DeSouza, T., Solivan-Rivera, J., Desai, A., Lifshitz, L.M., Shen, Y., Kelly, M., Guilherme, A., Henriques, F., et al. (2021). CRISPR-enhanced human adipocyte browning as cell therapy for metabolic

disease. Nat. Commun. *12*, 6931.

Tsumagari, K., Chang, C.-H., and Ishihama, Y. (2021a). Exploring the landscape of ectodomain shedding by quantitative protein terminomics. iScience *24*, 102259.

Tsumagari, K., Chang, C.-H., and Ishihama, Y. (2021b). A protocol for analyzing the protein terminome of human cancer cell line culture supernatants. STAR Protoc *2*, 100682.

Vaca Jacome, A.S., Rabilloud, T., Schaeffer-Reiss, C., Rompais, M., Ayoub, D., Lane, L., Bairoch, A., Van Dorsselaer, A., and Carapito, C. (2015). N-terminome analysis of the human mitochondrial proteome. Proteomics *15*, 2519–2524.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics *15*, 2458–2469.

Vijay, J., Gauthier, M.-F., Biswell, R.L., Louiselle, D.A., Johnston, J.J., Cheung, W.A., Belden, B., Pramatarova, A., Biertho, L., Gibson, M., et al. (2020). Single-cell analysis of human adipose tissue identifies depot and disease specific cell types. Nat Metab *2*, 97–109.

Wang, W., and Seale, P. (2016). Control of brown and beige fat development. Nature Reviews Molecular Cell Biology *17*, 691–702. https://doi.org/10.1038/nrm.2016.96.

Wang, L.-B., Karpova, A., Gritsenko, M.A., Kyle, J.E., Cao, S., Li, Y., Rykunov, D., Colaprico, A., Rothstein, J.H., Hong, R., et al. (2021). Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell *39*, 509–528.e20.

Weeks, A.M., Byrnes, J.R., Lui, I., and Wells, J.A. (2021). Mapping proteolytic neo-N termini at the surface of living cells. Proc. Natl. Acad. Sci. U. S. A. *118*, e2018809118.

Whitehead, A., Krause, F.N., Moran, A., MacCannell, A.D.V., Scragg, J.L., McNally, B.D., Boateng, E., Murfitt, S.A., Virtue, S., Wright, J., et al. (2021). Brown and beige adipose tissue regulate systemic metabolism through a metabolite interorgan signaling axis. Nat. Commun. *12*, 1905.

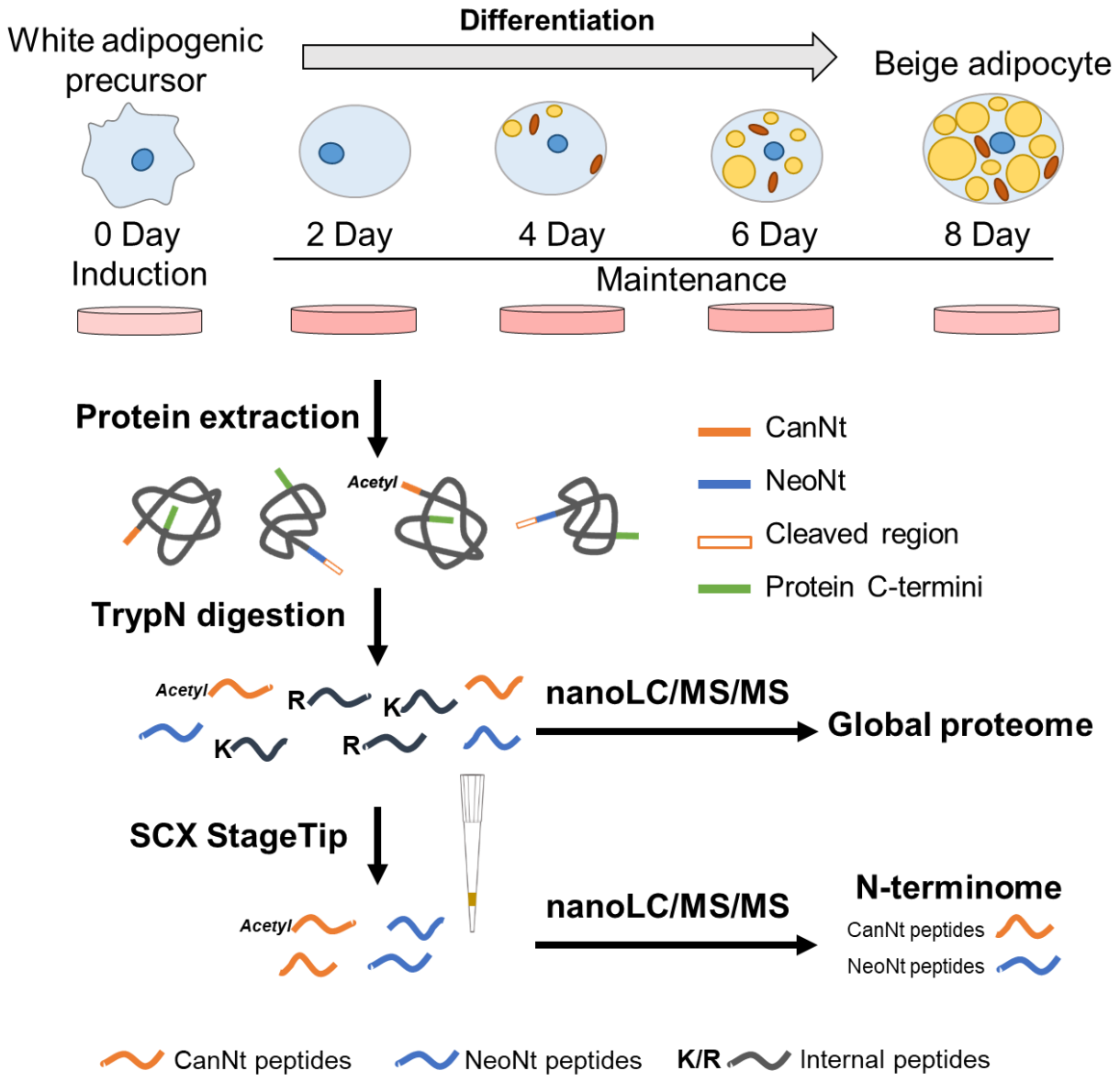
Willems, P., Ndah, E., Jonckheere, V., Stael, S., Sticker, A., Martens, L., Van Breusegem, F., Gevaert, K., and Van Damme, P. (2017). N-terminal Proteomics Assisted Profiling of the Unexplored Translation Initiation Landscape in Arabidopsis thaliana. Mol. Cell. Proteomics *16*, 1064–1080.

Yang, M., Zhang, Y., Pan, J., Sun, J., Liu, J., Libby, P., Sukhova, G.K., Doria, A., Katunuma, N., Peroni, O.D., et al. (2007). Cathepsin L activity controls adipogenesis and glucose tolerance. Nat. Cell Biol. *9*, 970–977.

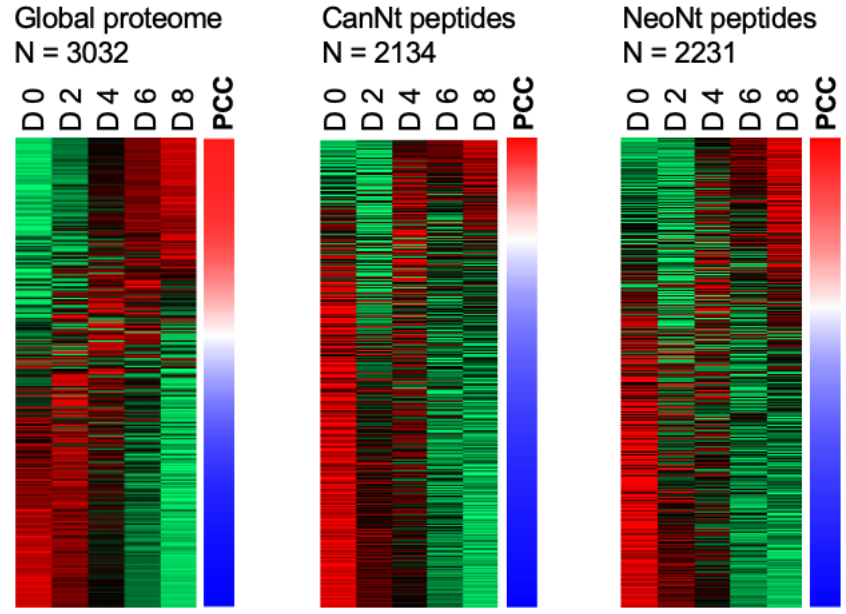
Yeom, J., Ju, S., Choi, Y., Paek, E., and Lee, C. (2017). Comprehensive

analysis of human protein N-termini enables assessment of various protein forms. Sci. Rep. *7*, 6599.

Figure 1 (A)


Differentiation
White adipogenic precursor
Beige adipocyte
0 Day
Induction
2 Day
4 Day
6 Day
8 Day
Maintenance
Protein extraction
Acetyl
CanNt
NeoNt
Cleaved region
Protein C-termini
TrypN digestion
Acetyl
R
K
K
R
nanoLC/MS/MS
Global proteome
SCX StageTip
Acetyl
nanoLC/MS/MS
N-terminome
CanNt peptides
NeoNt peptides
CanNt peptides
NeoNt peptides
K/R
Internal peptides

(B)


Global proteome
N = 3032
CanNt peptides
N = 2134
NeoNt peptides
N = 2231
D0
D2
D4
D6
D8
PCC

(C)

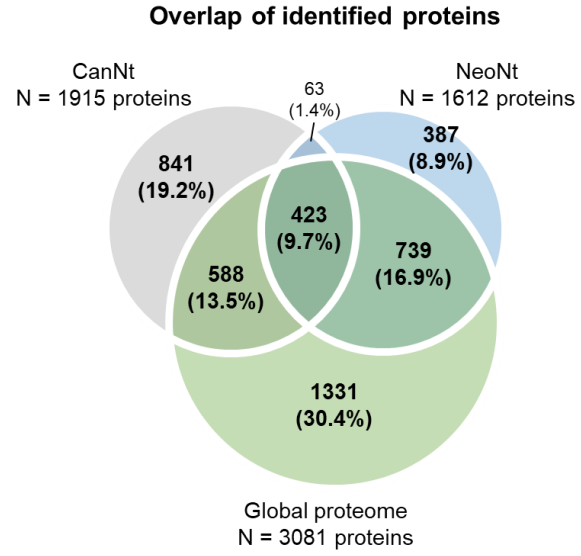
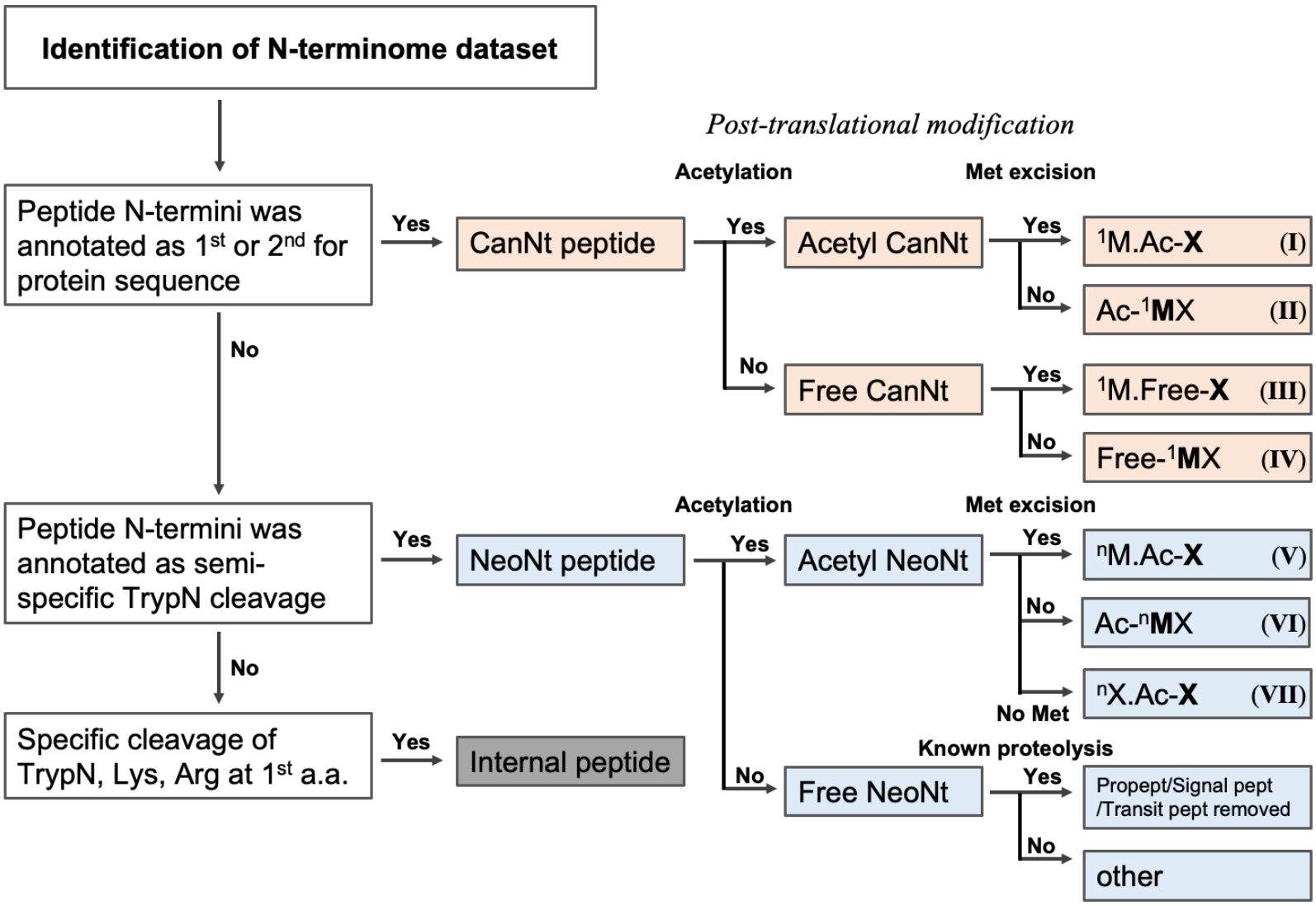

Overlap of identified proteins
CanNt
N = 1915 proteins
NeoNt
N = 1612 proteins
63
(1.4%)
841
(19.2%)
387
(8.9%)
423
(9.7%)
588
(13.5%)
739
(16.9%)
1331
(30.4%)
Global proteome
N = 3081 proteins

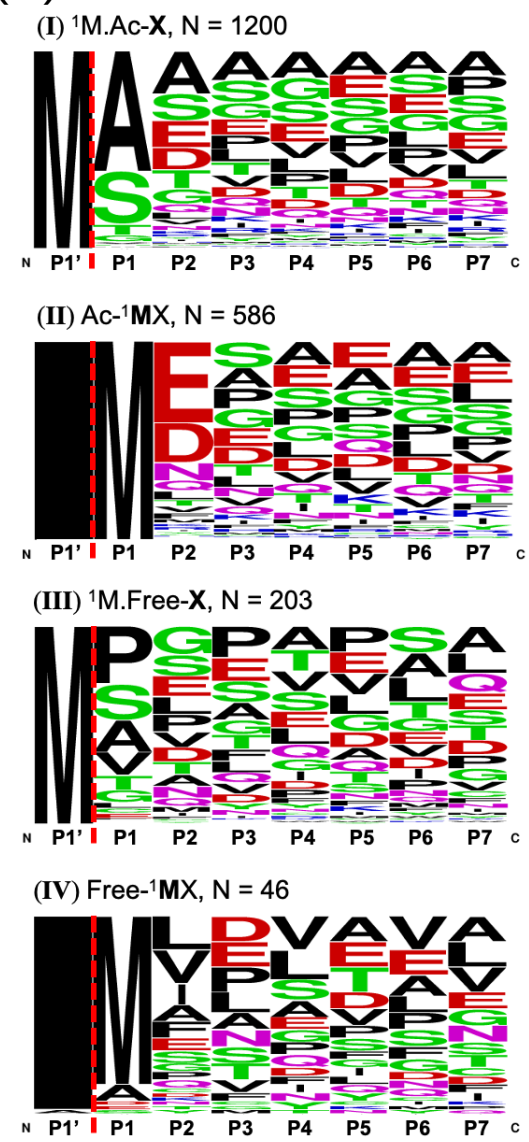
Figure 2 (A)



(B)

Type of peptides	Unique peptides	N-terminal modification	Unique peptides			Unique peptides	N-termini	Proteins
Canonical protein N-terminal peptides (CanNt peptides)	3043	Acetylated	2728		(I) Met removed (1M.Ac-X)	1628	1220	1220
					(II) Met intact (Ac-1MX)	1100	592	592
		Unmodified	315		(III) Met removed (1M.Free-X)	242	203	203
					(IV) Met intact (Free-1MX)	73	47	47
Neo-N-terminal peptides (NeoNt peptides)	4129	Acetylated	300		(V) Met removed (nM.Ac-X)	75	67	67
					(VI) Met intact (Ac-nMX)	100	62	62
					(VII) No preceding Met (nX.Ac-X)	125	118	117
		Unmodified	3829	Proteolytic peptides	Propeptide removed	11	9	9
					Signal peptide removed	77	65	65
					Transit peptide removed	49	37	37
					Other	3692	3306	1377
Internal peptide	378	Unmodified	378	Internal peptide		378		250
					Total	7550		

(C)



(D)

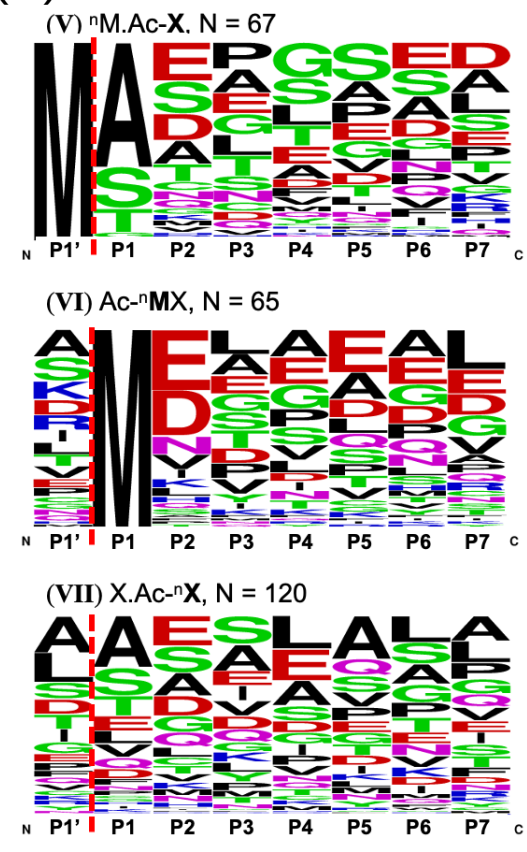
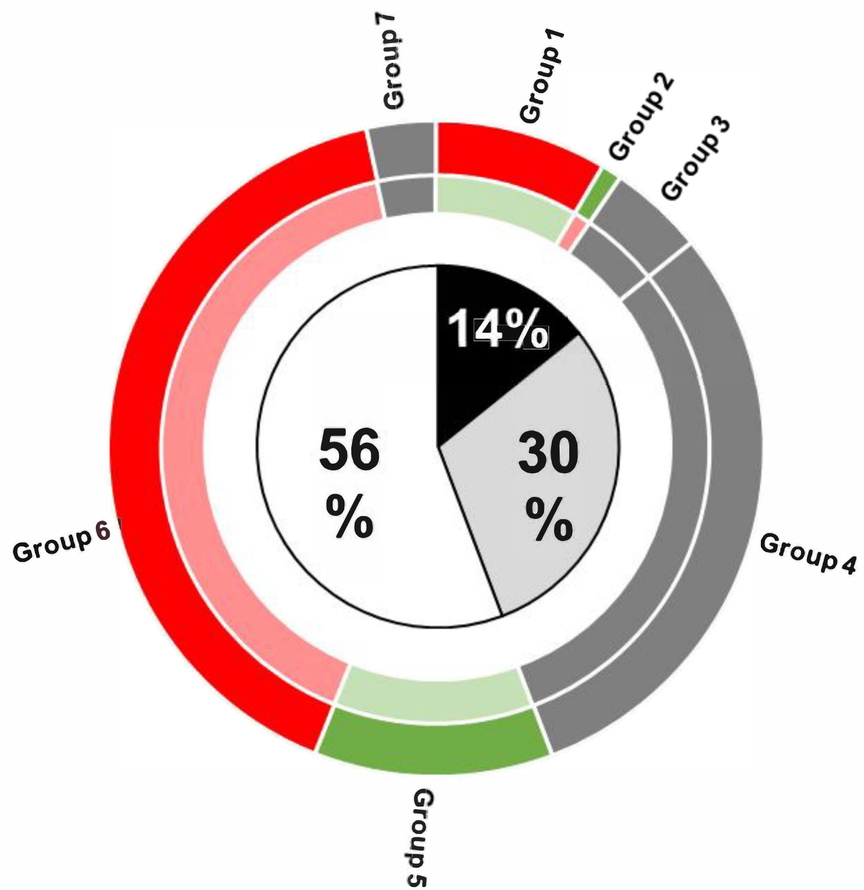


Figure 3

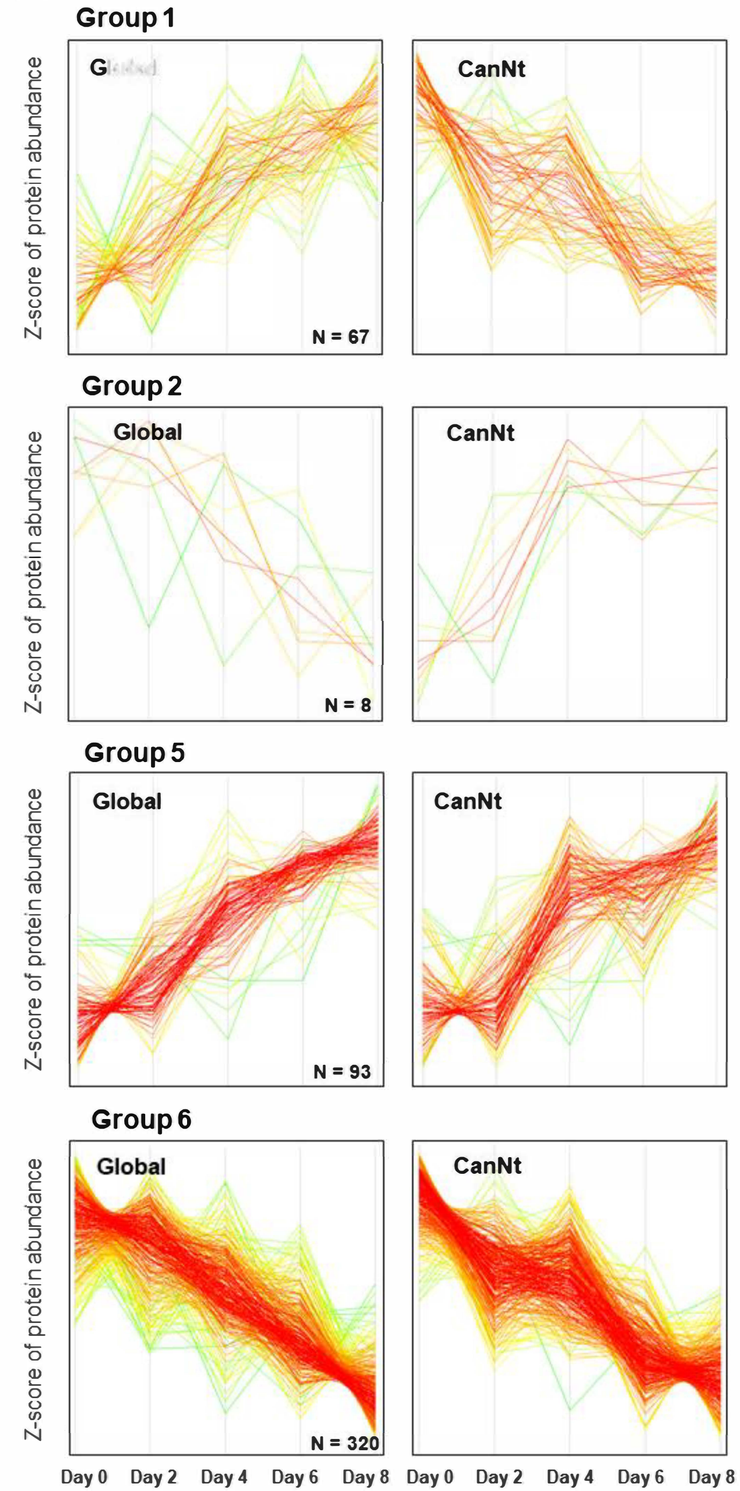


Pearson correlation of temporal profiles

Negative Independent Positive

Temporal profiles

Global	CanNt
Decreasing	Decreasing
Increasing	Increasing
None-linear	None-linear



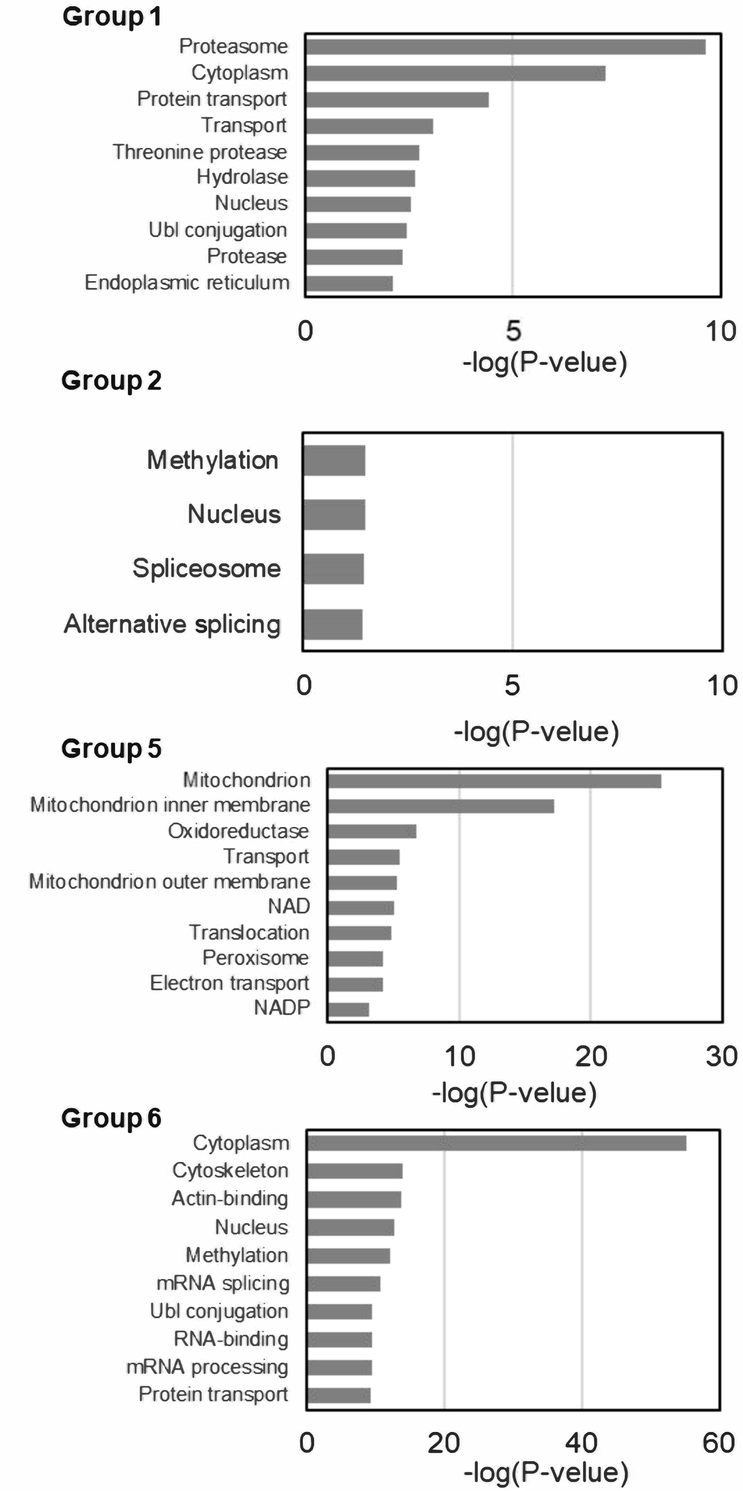
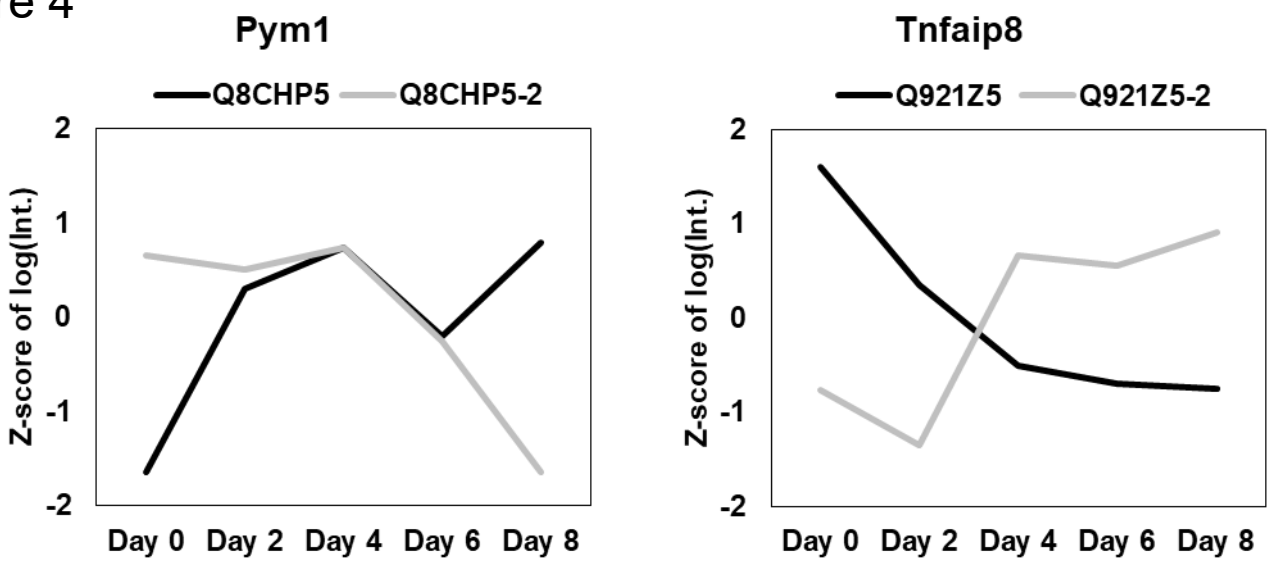
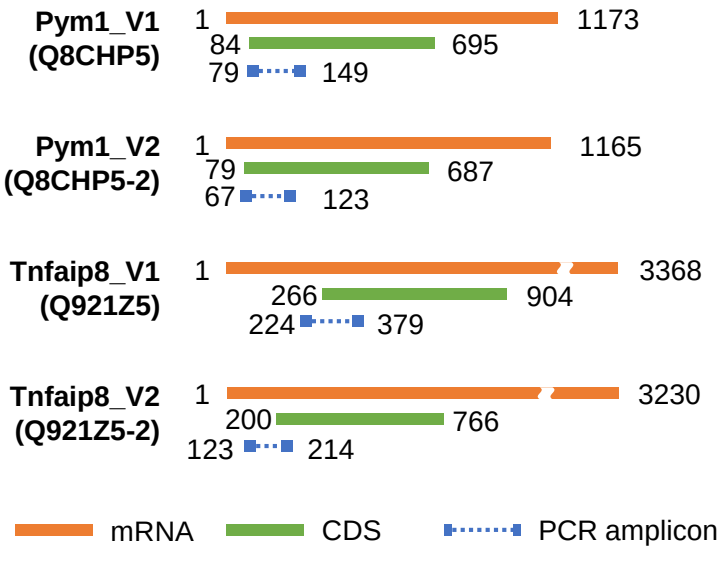


Figure 4

(A)


Pym1
Q8CHP5
Q8CHP5-2
Z-score of log(Int.)
2
1
0
-1
-2
Day 0
Day 2
Day 4
Day 6
Day 8
Tnfaip8
Q921Z5
Q921Z5-2
Z-score of log(Int.)
2
1
0
-1
-2
Day 0
Day 2
Day 4
Day 6
Day 8

(B)


Pym1_V1
(Q8CHP5)
1
1173
84
695
79
149
Pym1_V2
(Q8CHP5-2)
1
1165
79
687
67
123
Tnfaip8_V1
(Q921Z5)
1
3368
266
904
224
379
Tnfaip8_V2
(Q921Z5-2)
1
3230
200
766
123
214
mRNA
CDS
PCR amplicon

(C)

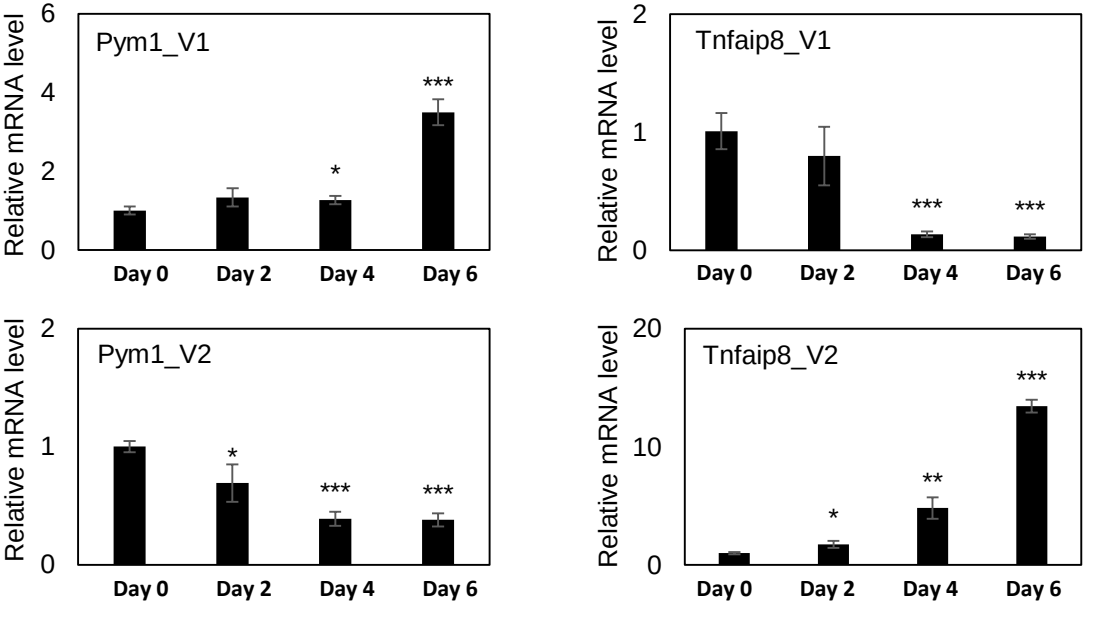
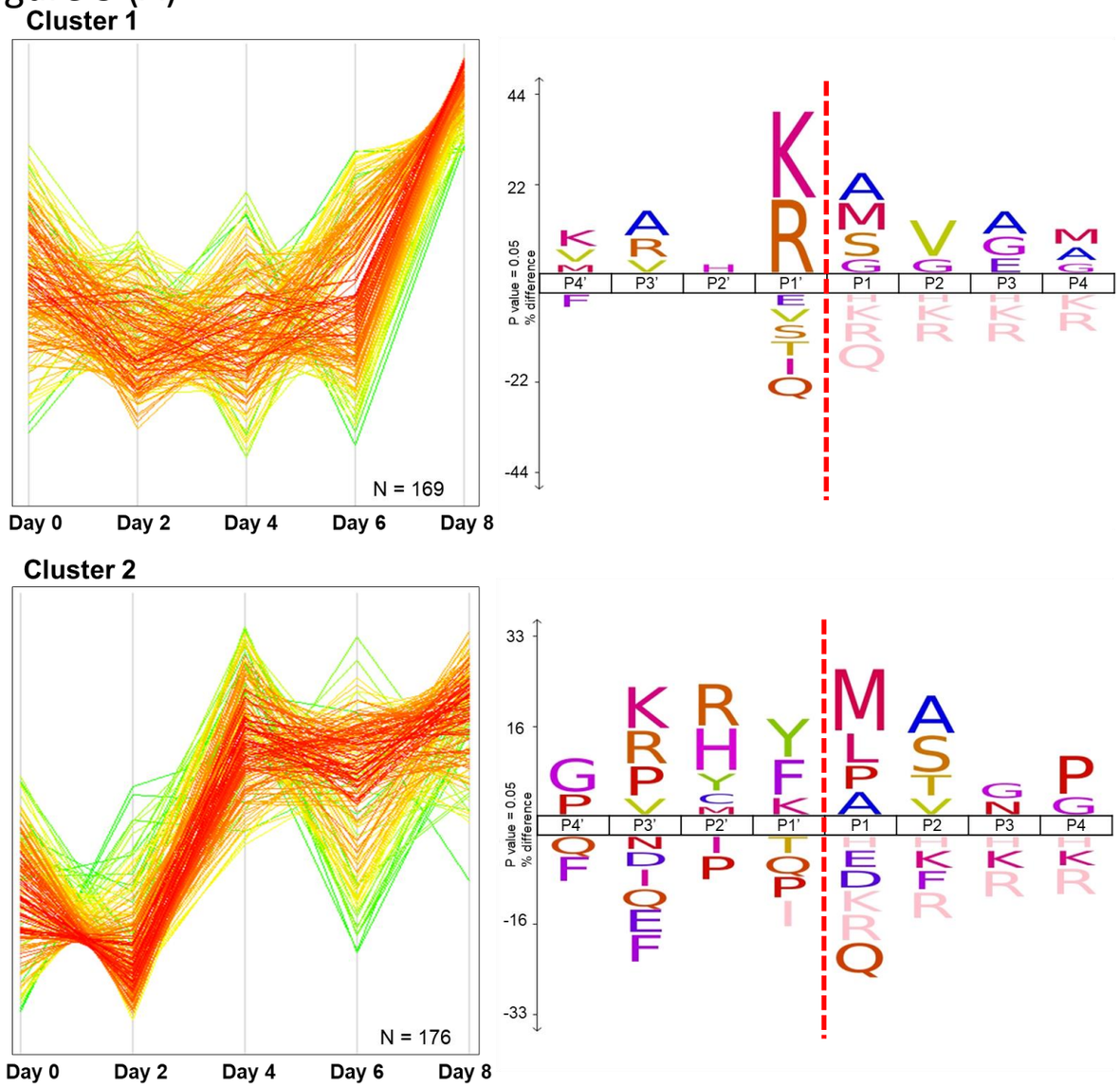
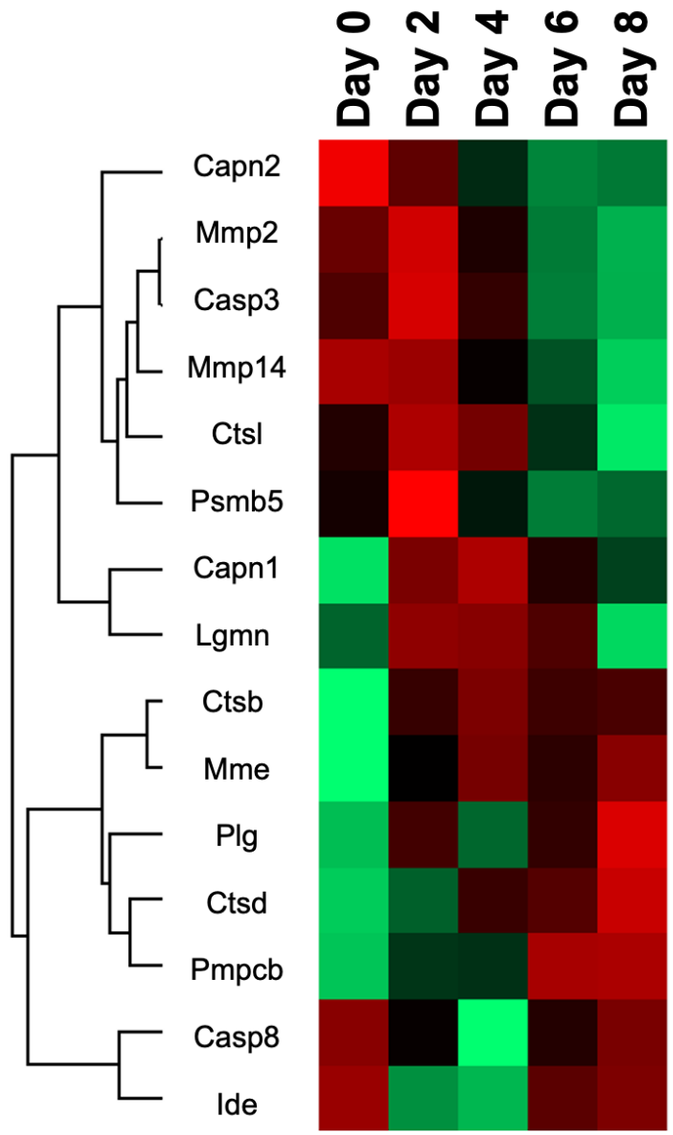

Pym1_V1
Relative mRNA level
6
4
2
0
*
***
Day 0
Day 2
Day 4
Day 6
Tnfaip8_V1
Relative mRNA level
2
1
0
***
***
Day 0
Day 2
Day 4
Day 6
Pym1_V2
Relative mRNA level
2
1
0
*
***
***
Day 0
Day 2
Day 4
Day 6
Tnfaip8_V2
Relative mRNA level
20
10
0
*
**
***
Day 0
Day 2
Day 4
Day 6

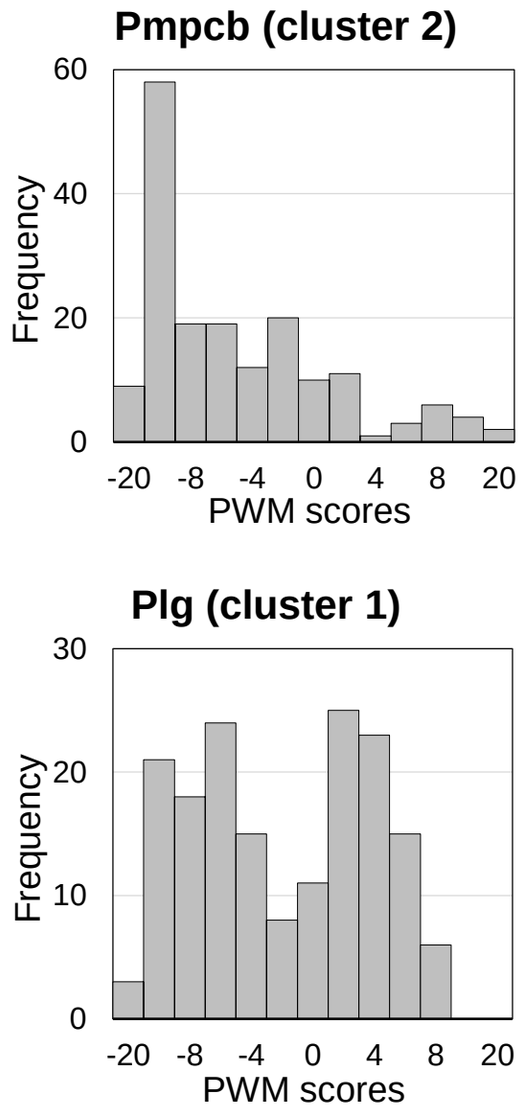
Figure 5 (A)



(B)



(C)



(D)

GENE	Protein accession	Cleaved Site (Identified)	Cleaved Site (Identified)	Cleaved Site (Uniport)	Cleaved Site (TargetP 2.0)	Cleaved Site (MitoFates)
Atpaf1	Q811I0	PVRR–[51]GSGL	50–51	54–55$	50–51	24–25
Bola1	Q9D8S9	VSRG–[23]SAGS	22–23	N/A	22–23	22–23
Echs1	Q8BH95	LRRF–[28]ASGA	27–28	27–28*	27–28	27–28
Fam210a	Q8BGY7	SKRV–[96]LSSS	95–96	N/A	-	95–96
Fxn	O35943	HLRN–[78]LGTL	77–78	77–78*	77–78	40–41
Gfm2	Q8R2Q4	TRNY–[46]SSAP	45–46	N/A	45–46	44–45
Hmgcl	P38060	SLRA–[21]VSTS	20–21	27–28*	20–21	20–21
Idh3a	Q9D6R2	VTRG–[27]FAGG	26–27	27–28*	26–27	26–27
Idh3g	P70404	PRRS–[39]ISSQ	38–39	39–40*	38–39	38–39
Mars2	Q499X9	CRRY–[23]GSCS	22–23	46–47$	22–23	21–22
Mrpl49	Q9CQ40	GRRK–[26]LSQT	25–26	N/A	26–27	25–26
Mrps28	Q9CY16	PLRN–[29]ASTE	28–29	70–71*	28–29	28–29
Pc	Q05920	GVRR–[19]SSSA	18–19	20–21$	18–19	18–19
Pdha1	P35486	ASRN–[30]FAND	29–30	29–30*	30–31	29–30

(E)

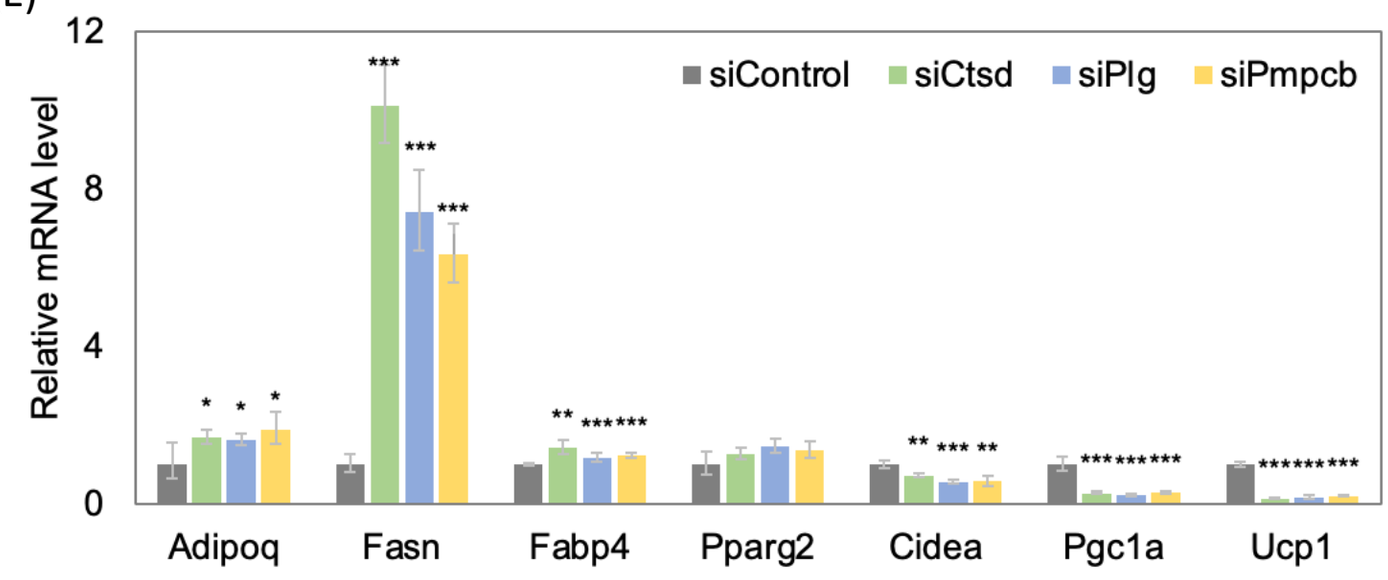
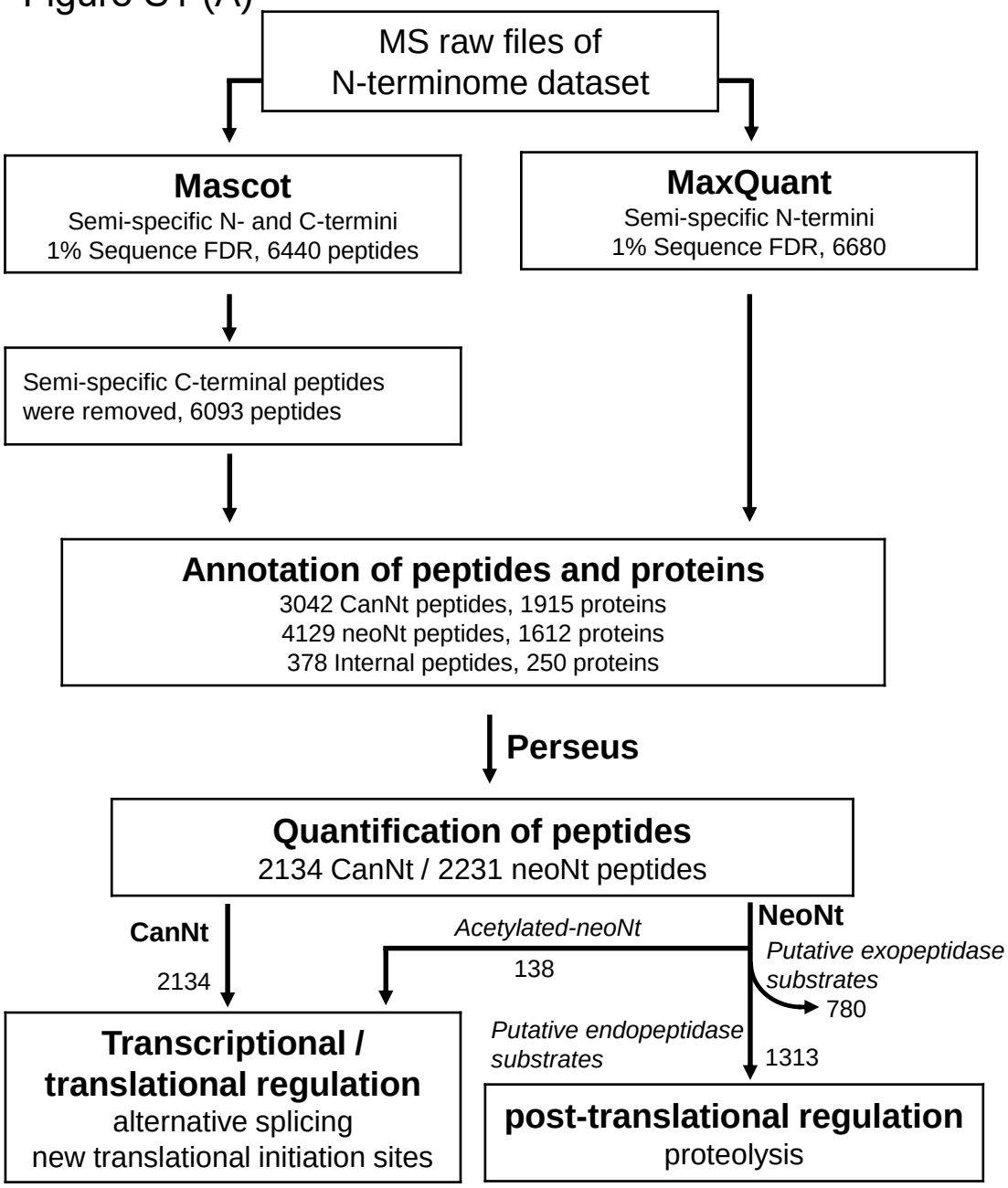
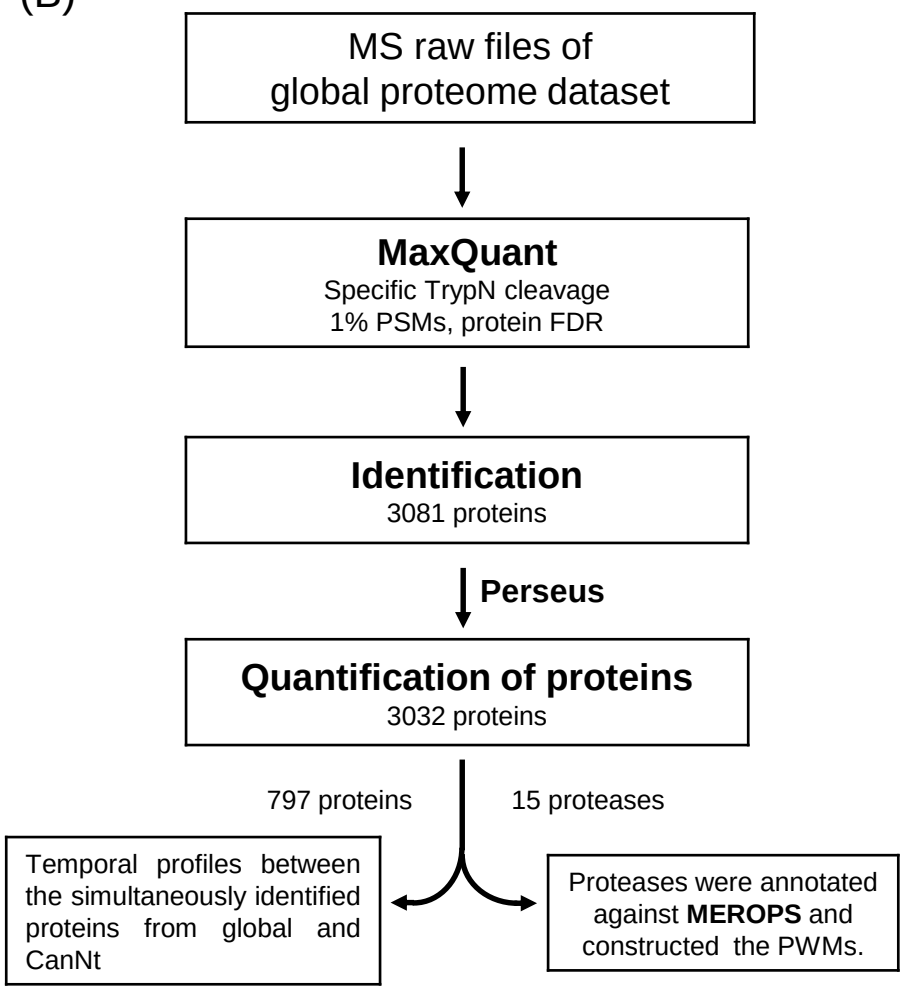


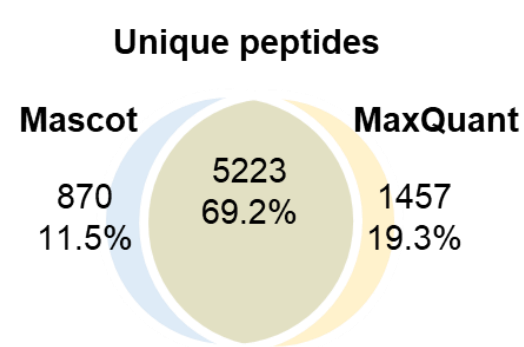
Figure S1 (A)



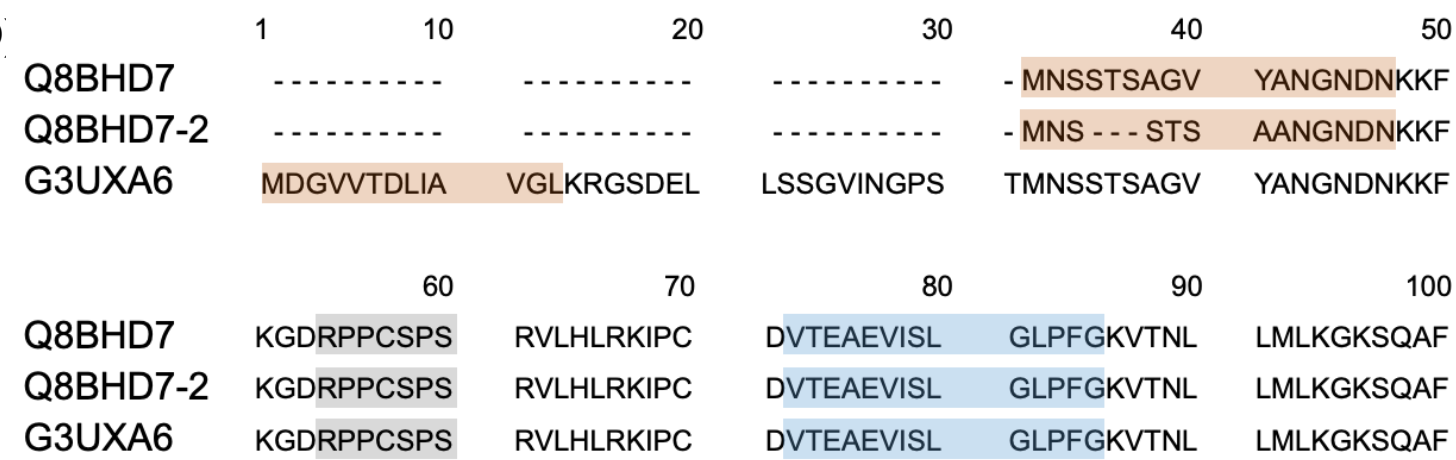
(B)



(C)

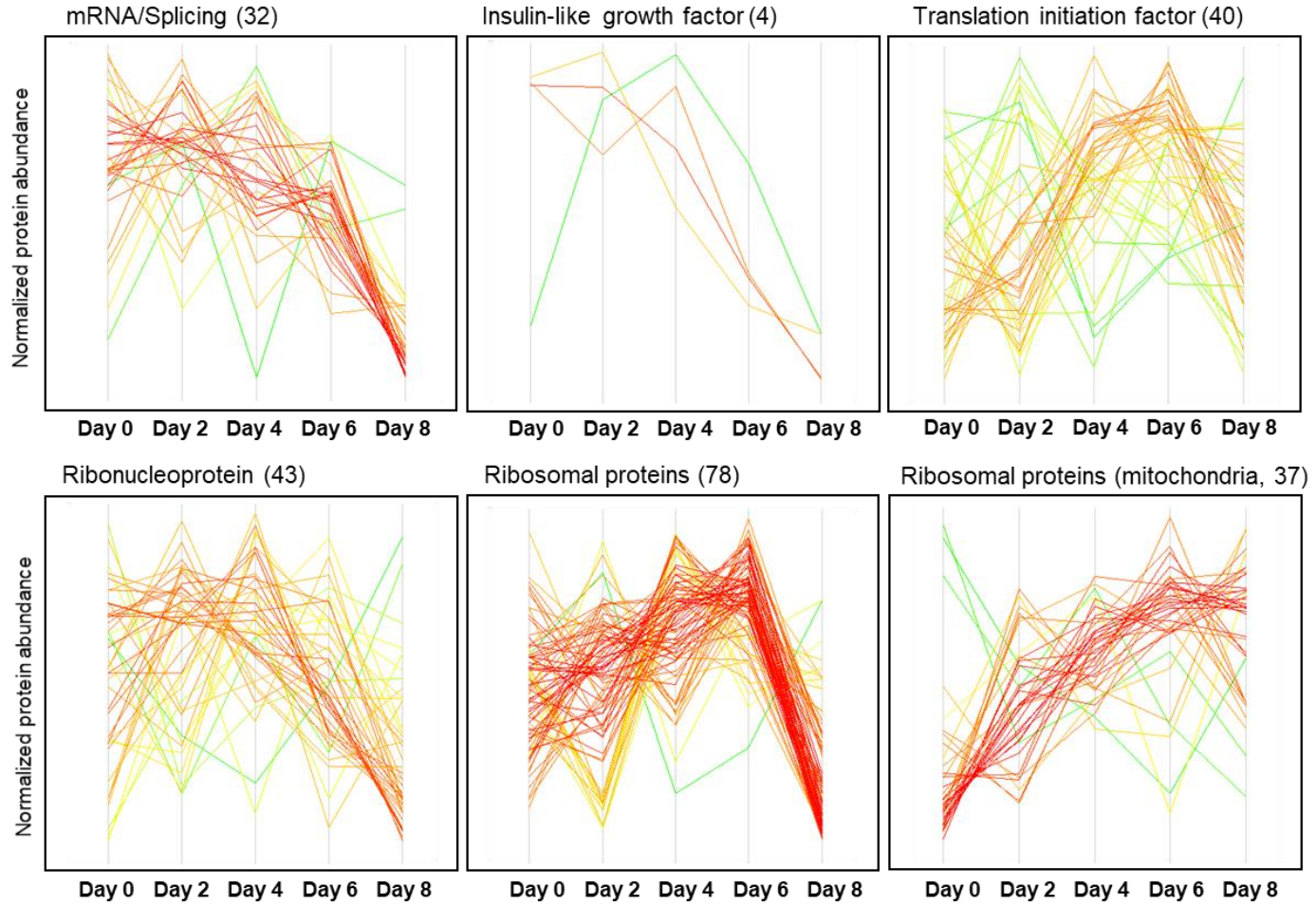


(D)

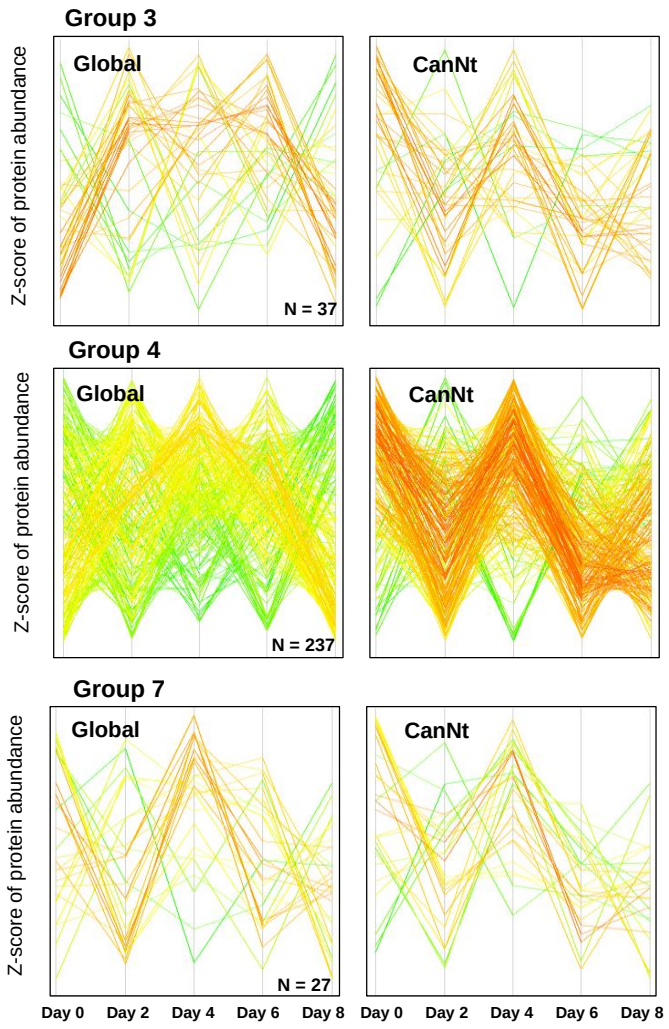


Type of pept	Start site	Sequence	Master protein	Protein Groups
CanNt-pept	1	MNSSTSAGVYANGNDN	Q8BHD7	Q8BHD7
CanNt-pept	2	NSSTSAGVYANGNDN	Q8BHD7	Q8BHD7
NeoNt-pept	3	SSTSAGVYANGNDN	Q8BHD7	Q8BHD7
CanNt-pept	1	MNSSTSAANGNDN	Q8BHD7-2	Q8BHD7-2
CanNt-pept	1	MDGVVTDLIAVGL	G3UXA6	G3UXA6
Internal-pept	54	RPPCSPS	Q8BHD7	Q8BHD7, Q8BHD7-2, G3UXA6
NeoNt-pept	72	VTEAEVISLGLPFG	Q8BHD7	Q8BHD7, Q8BHD7-2, G3UXA6
NeoNt-pept	73	TEAEVISLGLPFG	Q8BHD7	Q8BHD7, Q8BHD7-2, G3UXA6
NeoNt-pept	74	EAEVISLGLPFG	Q8BHD7	Q8BHD7, Q8BHD7-2, G3UXA6
NeoNt-pept	75	EVISLGLPFG	Q8BHD7	Q8BHD7, Q8BHD7-2, G3UXA6

Figure S2 (A)



(B)



(C)

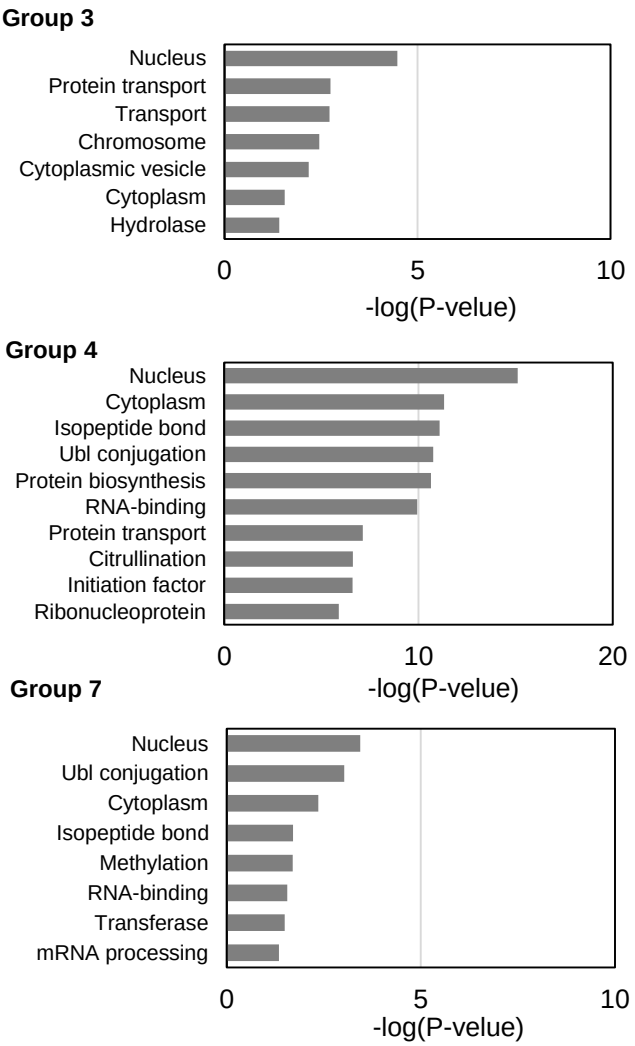
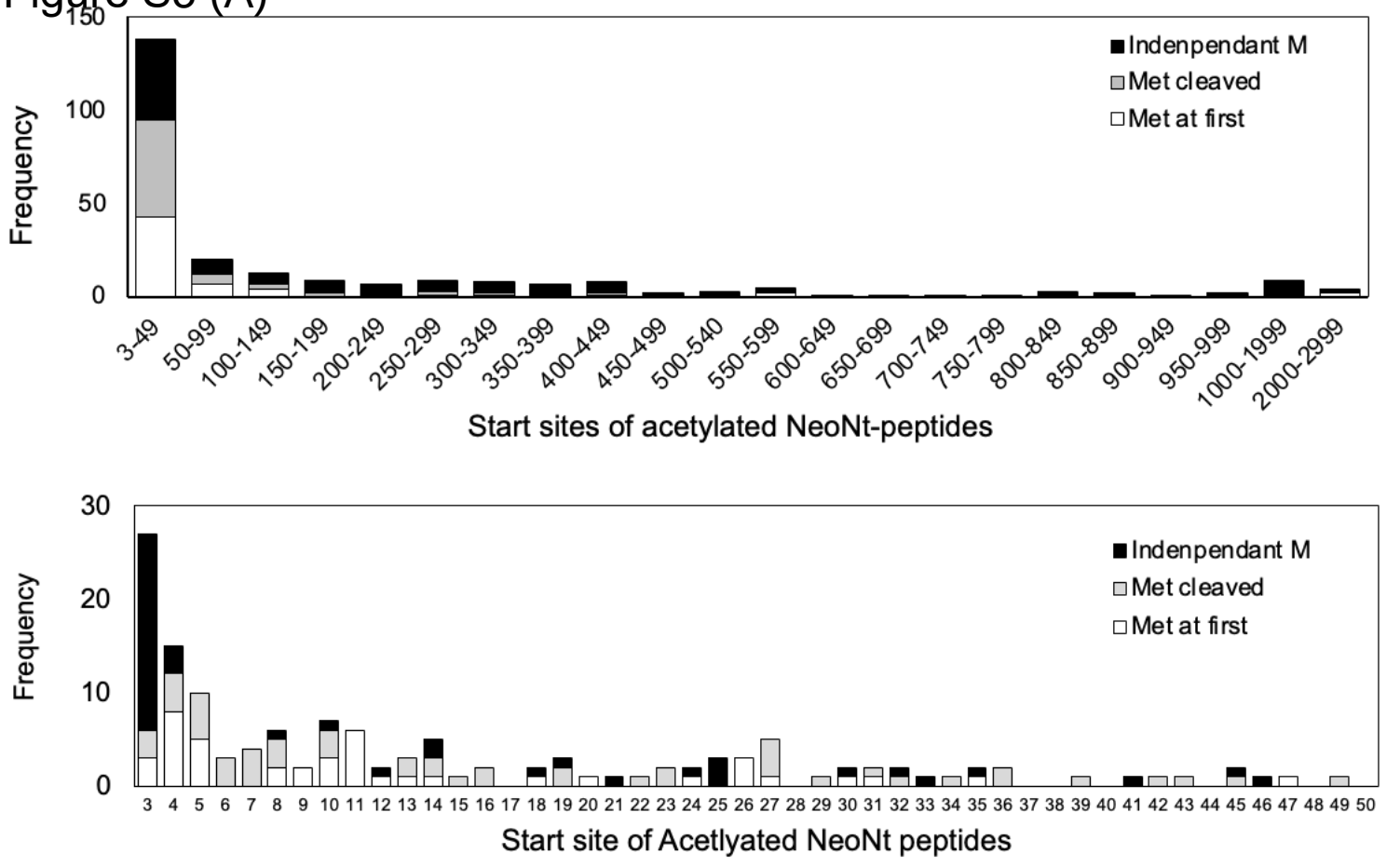
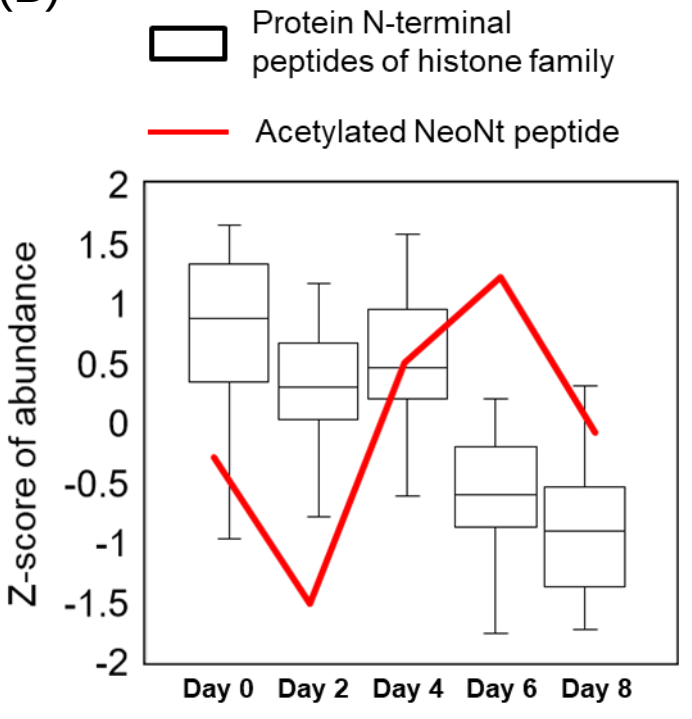


Figure S3 (A)



(B)



(C)

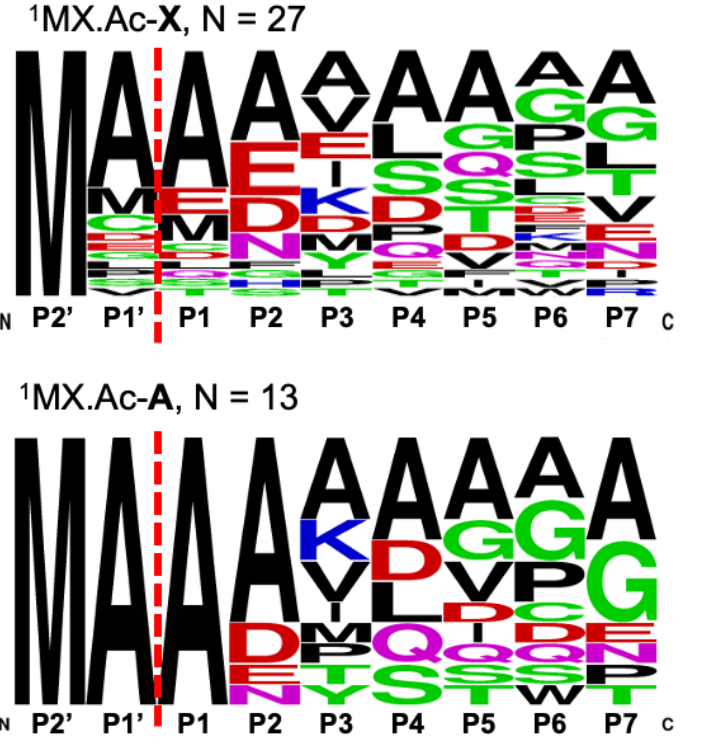
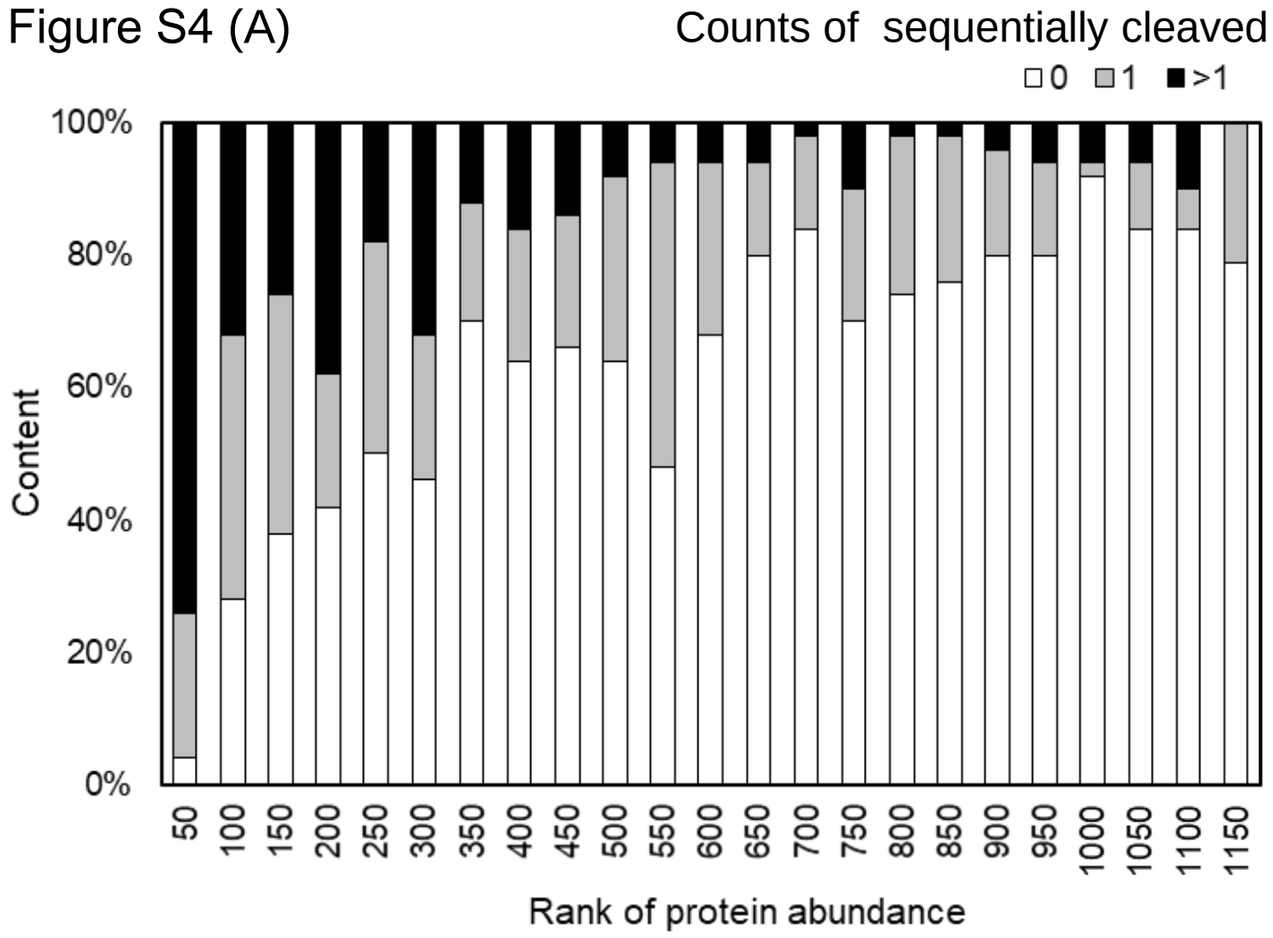
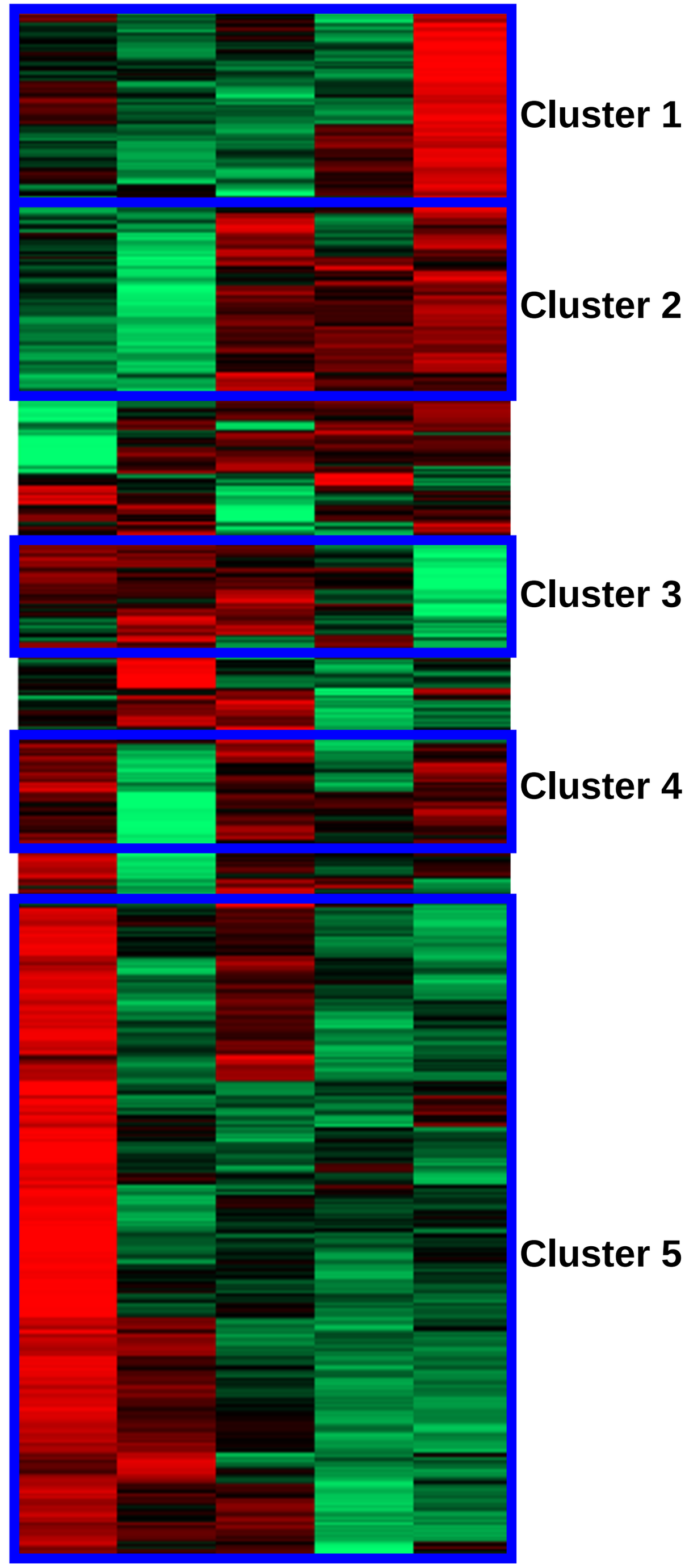
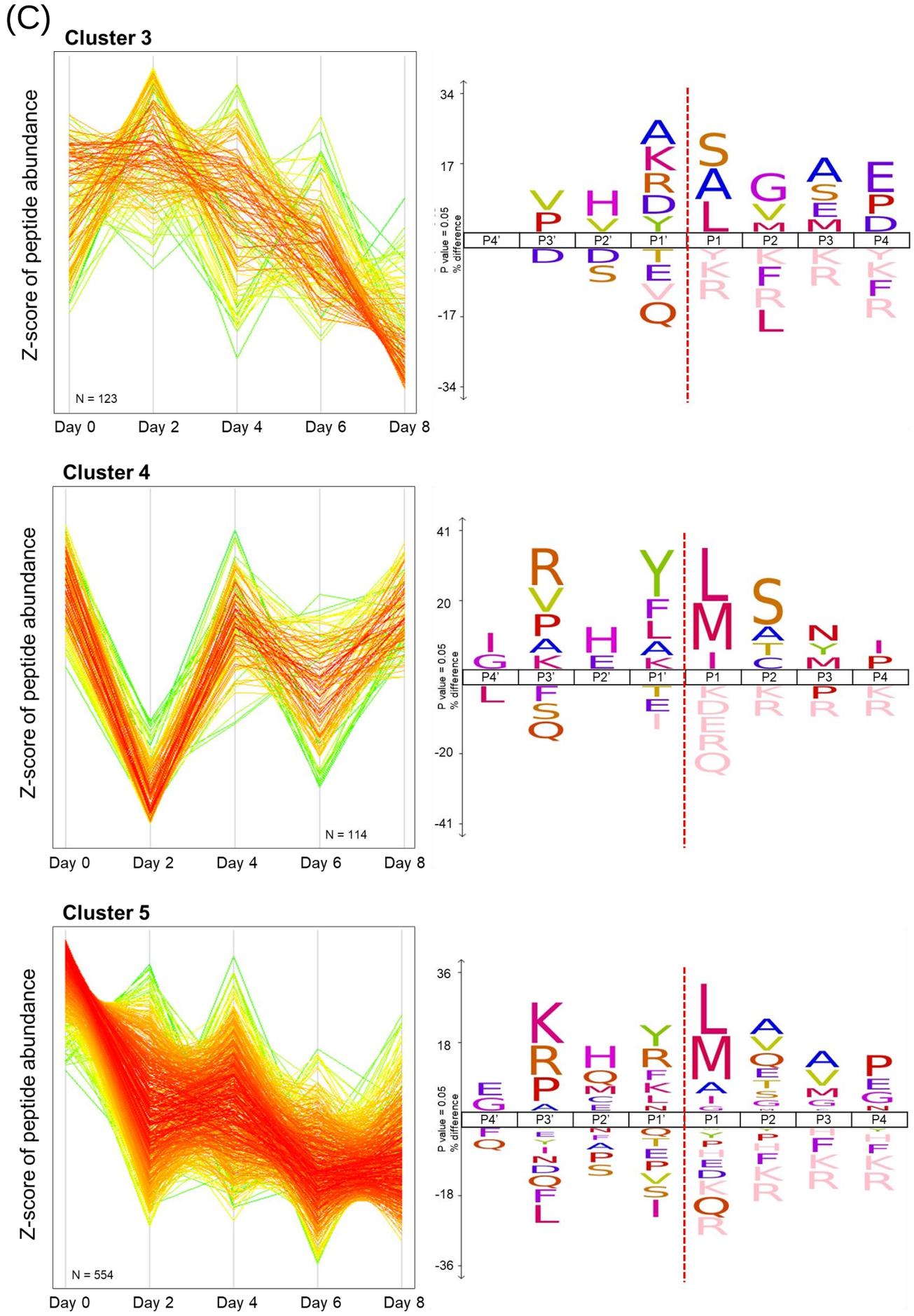


Figure S4 (A)



(B)



(C)