PNAS
www.pnas.org

Supplementary Information for

Sample multiplexing for targeted pathway proteomics: application to aging mice

Qing Yu, Haopeng Xiao, Mark P. Jedrychowski, Devin K. Schweppe, Jose Navarrete-Perea, Jeffrey Knott, John Rogers, Edward T. Chouchani, Steven P. Gygi

Correspondence to Steven P. Gygi

Email: Steven_Gygi@hms.harvard.edu

This PDF file includes:

Supplementary text

Figs. S1 to S8

Other supplementary materials for this manuscript include the following:

pdf file - Quick guide to Tomahto

Dataset S1 - Target peptide information for 77 kinases

Dataset S2 - Target peptide information for proteins in metabolism- and inflammation-related processes.

Dataset S3 - Tomahto and full proteome analysis of 77 kinases in three cell lines

Dataset S4 - Tomahto analysis of 260 proteins in 90 mouse tissues

Dataset S5 - Proteome analysis of 90 mouse tissues

SI Experimental Procedures

Experimental Model and Subject Details

Human colorectal carcinoma (HCT116), human embryonic kidney 293 (HEK293) and Michigan cancer foundation-7 (MCF7) cells were used for method development and benchmarking.

Ten young and ten old male C57BL/6J mice were purchased from the Jackson Laboratories and housed in a temperature-controlled (20–22 °C) room on a 6am to 6pm light/dark cycle for a month until age of 16 and 80 weeks. The age groups were littermate matched. All animal-related experiments were approved by Institutional Animal Care and Use Committee of the Beth Israel Deaconess Medical Center.

Method Details

Standard Peptide Preparation

Peptides (Dataset S2) were synthesized by Cell Signaling Technologies. Peptides were purified using a SepPak 96-well plate. 1 nmol of each peptide was combined into a mixture of 520 peptides and the pH was adjusted to 8.5 in 200 mM EPPS. HPLC grade acetonitrile was added to a final concentration of 10% (v/v). TMTsh reagents (Thermo Fisher Scientific; Catalog #A43073) were then added at a ratio of 2:1 (TMT:Peptide) by mass. The reaction proceeded at room temperature for 1 hr before quenching with a final volume of 0.5% hydroxylamine (Sigma). The peptides were then vacuum-dried and purified using a 50 mg SepPak. The peptides were then reconstituted in 1% formic acid in water and ready for MS analysis.

Mouse Tissue Sample Preparation

Mice were euthanized by cervical dislocation, and all tissues were rapidly extracted within thirty seconds after euthanasia. Wollenberger tongs were prechilled in liquid nitrogen and were used to freeze-clamp each tissue immediately after extraction (1). Clamped tissues were then placed in a liquid nitrogen dewar. To ensure the extraction speed, in each age group, brown fat, kidney, skeletal muscle, and white fat (epididymal fat) were extracted from a set of five mice, and brain, liver, heart, spleen, and lung were from the other set of five mice.

Tissues were homogenized by bead beating in 8 M urea buffer [8 M urea, 2% SDS w/v, 200 mM EPPS (pH 8.5), 1 x protease inhibitor tablet (Thermo Fisher Scientific) and 1 x phosphatase inhibitor tablet (Thermo Fisher Scientific)]. Suspensions were centrifuged at maximum speed for 15 min at 4°C and lysates were transferred to clean Eppendorf tubes. Protein concentrations were measured by BCA assay (Thermo Fisher Scientific). Samples were reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP) and alkylated with 10 mM iodoacetamide that was quenched with 10 mM DTT. A total of 100 µg of protein was chloroform−methanol precipitated. Protein was reconstituted in 200 mM EPPS at pH 8.5 and digested by Lys-C overnight and trypsin for 6 h, both at a 1:100 protease-to-peptide ratio. Resulting peptides were labeled with TMT10 at a ratio of 2:1 (TMT:Peptide) by mass. The reaction proceeded at room temperature for 1 hr before quenching with a final volume of 0.5% hydroxylamine (Sigma). Peptides were then combined at a 1:1 across all channels and desalted using a 100 mg SepPak cartridge. For TOMAHAQ analysis, 100 fmol TMTsh-labeled trigger peptides and 2 µg TMT10-labeled endogenous were loaded on column. For shotgun proteome analysis, 600 µg of labeled peptides were loaded and fractionated with basic-pH reverse-phase (BPRP) high-performance liquid chromatography (HPLC) , collected in a 96-well plate and consolidated to a final of 24 fractions, out of which only alternating fractions (a total of 12) were analyzed (2).

Liquid Chromatography-Mass Spectrometry Analysis

All data were collected on an Orbitrap Fusion Lumos mass spectrometer coupled with a Proxeon NanoLC1200 UHPLC. The 100 µm capillary column was packed with 35 cm of Accucore 150 resin (2.6 µm, 150 Å; ThermoFisher Scientific). Mobile phase were 5% acetonitrile, 0.125% formic acid (A) and 95% acetonitrile, 0.125% formic acid (B).

For the standard DDA-SPS-MS3 method with online real-time database search (RTS), each fraction was eluted using a 120-min method over a gradient from 4% to 30% B. Peptides were ionized with a spray voltage of 2,600 kV. The instrument method included Orbitrap MS1 scans (resolution of 120,000; mass range 400−1400 m/z; automatic gain control (AGC) target $2e^5$, max injection time of 50 ms) and ion trap MS2 scans (CID collision energy of 35%; AGC target $1e^4$; rapid scan mode; max injection time of 120 ms). RTS was enabled and quantitative SPS-MS3 scans (resolution of 50,000; AGC target $2.5e^5$; max injection time of 250 ms) were inserted

through Orbiter with a real-time false discovery rate (rtFDR) filter implementing a modified linear discriminant analysis (3).

For Tomahto, each sample was eluted using a 180-min method over a gradient from 5% to 35% B. The instrument method only included Orbitrap MS1 scans (resolution of 120,000; mass range 400−1400 m/z; automatic gain control (AGC) target $2e^5$, max injection time of 50 ms). Peptide targets were imported into Tomahto, and all possible fragment ions were auto-populated, obviating the priming run required by the original version of TOMAHAQ.

Tomahto listens to each collected MS1 scan. When a precursor ion matched a potential trigger peptide (±10 ppm mass accuracy; matched charge state; minimal intensity of $5e^4$), the software will insert up to four scan to be collected by the Orbitrap.

1) Tomahto prompts insertion of an Orbitrap MS2 scan with the trigger peptide's precursor mass (0.5 m/z isolation window; resolution of 15,000; AGC target $1e^4$; max injection time of 120 ms; CID collision energy 35%). This scan has priority after insertion and is almost immediately collected. Once collected, a real-time peak matching strategy (RTPM) is used to confirm the identity of the trigger peptide (must match >6 fragment peaks within ±10 ppm).

2) If successful, Tomahto prompts the insertion of an Orbitrap MS2 scan using the target peptide as the precursor m/z value (0.5 m/z isolation window; resolution of 15,000; AGC target $1e^5$; max injection time of 900 ms; CID collision energy 35.1). The target peak m/z is a mixture of multiplexed endogenous peptides. At the same time, the MS2 fragment ions and their intensities for the trigger peptide are stored in memory as a template library spectrum. After collection, the target MS2 scan is used to confirm that the target peptide is present at levels sufficient for detection. This is accomplished via RTPM where fragment ions must be present in the spectrum (±10 ppm) and rank ordered by intensity from the trigger MS2. SPS fragment ions are now selected from this scan. Only *b*- and *y-type* ions are considered for selection provided they have a TMT modification. SPS candidates are required to match the fragmentation pattern of the stored library spectrum, meaning fragment ratios relative to the highest fragment were within ±50% of that in the stored spectrum. In addition, each SPS candidate undergoes a purity filter of 0.5 (at least 50% of the signal attributed to the fragment ion within a 3 m/z window) to be included in the final list.

3) Upon confirmation of target peptide presence and successful selection of SPS ions, Tomahto next triggers an ion trap SPS-MS3 prescans (normal scan mode; AGC target of $1e^6$; max injection time of 10 ms). This is used to quickly estimate the signal strength for the TMT reporter ions. This estimate is used to set the lengthy injection times need for the SPS-MS3 scan detected in the Orbitrap.

4) Following the prescan, Tomahto prompts the insertion of the SPS-MS3 quantification scan (resolution of 50,000; SPS ions from part 2, 0.5 m/z window, max injection time of 5,000 ms).

The results of each scan are plotted in real-time through the data visualization module in Tomahto. A close-out on a trigger peptide is initiated when 3 MS3 scans are collected on any given trigger peptide with a minimum of 1,000 summed signal-to-noise value. Two additional technical injection replicates were performed for each tissue with close-out carried through to test if Tomahto was able to target >500 peptides (>1000 precursors) in single shots **(Fig. S3E)**. See Supplemental Information **(Fig. S1 and S2)** for detailed workflow.

Data Processing

For shotgun DDA-SPS-MS3 analyses, raw data were converted to mzXML format using a modified version of RawFileReader and searched against a mouse target-decoy protein database (Uniprot, December 21, 2018) using the SEQUEST algorithm (4). Searches were performed with a 50 ppm precursor mass tolerance, 0.9 Da fragment ion mass tolerance, trypsin digest with up to 2 missed cleavages. Allowed modifications included cysteine carboxyamidomethylation (+57.02146), static TMT on lysine and peptide N temini (+229.16293) and up to 3 variable methionine oxidation (+15.99491). Peptide spectral matches were filtered with a linear discriminant analysis (LDA) method to a 1% FDR (5) and protein-level FDR of 1% was enforced at the entire dataset level of 108 runs (9 tissues of 12 fractions each)(6).

For Tomahto, raw data were analyzed by the data analysis module of Tomahto. RawFileReader was used to read files and spectra were matched to synthetic trigger peptide or endogenous target peptides respectively.

Quantification and Statistical Analysis

For quantification, TMT reporter ion signal-to-noise (SN) values were extracted from MS3 scans and those with a summed SN<100 across ten channels were removed from final dataset. Column normalization was performed to correct for different protein loading in each channel. Lastly, for each protein signal to noise measurements of the peptides were summed and then normalized to 100 across the 10 samples yielding a “relative abundance” measurement. Further quantitative analysis was completed in R (http://www.R-project.org). T tests were performed on each protein. Resulting p values were corrected for multiple hypothesis testing to generate q values (7). GSEA analysis was performed using clusterProfiler (8). Protein functional networks were generated and displayed using Cytoscape (9).

Data and Software Availability

The software and user instructions will be freely available at https://gygi.med.harvard.edu/software (also smarttmt.org). All MS raw files will be available through the PRIDE archive. The use of Tomahto requires an API license (https://github.com/thermofisherlsms/iapi) from Thermo Fisher Scientific.

Supplemental Figures

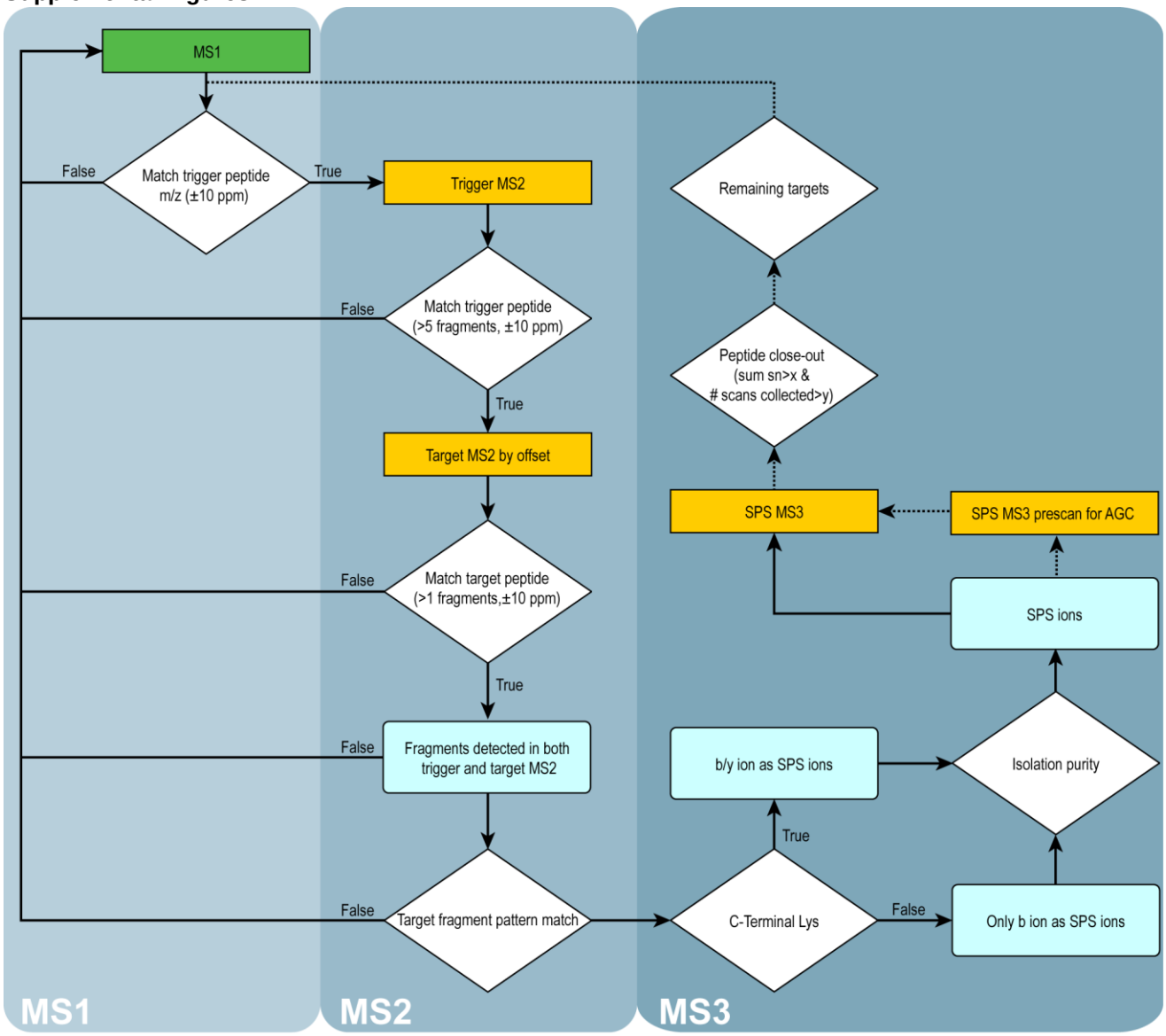
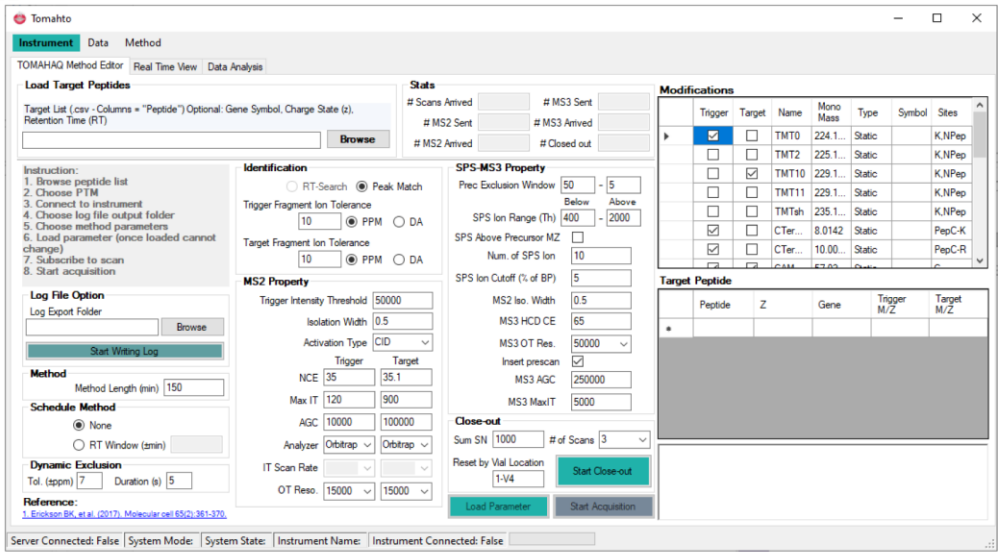


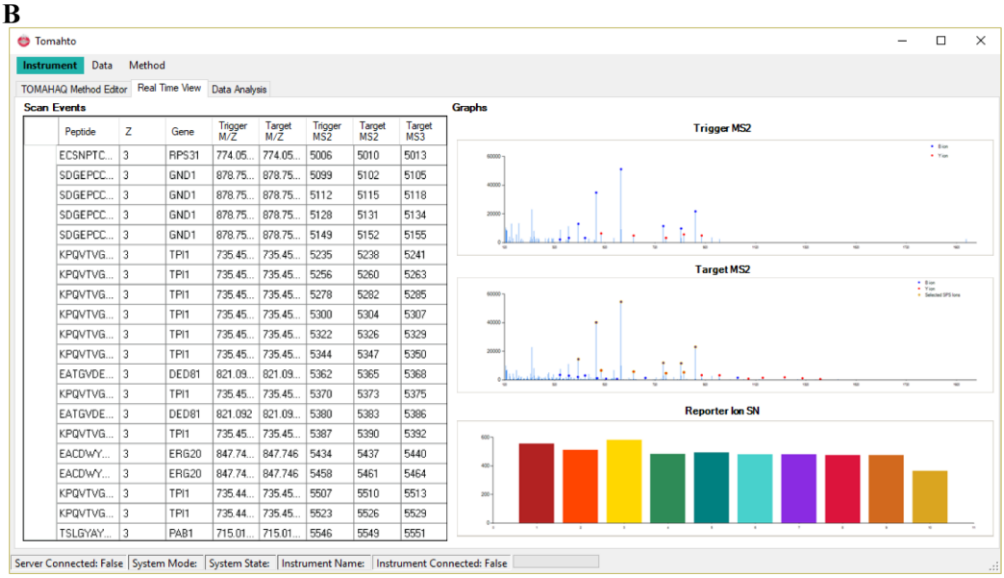
Figure S1. A flowchart of Tomahto scan sequence and filters. Scans scheduled by instrument method are indicated by green squares whereas scans inserted by Tomahto are in orange squares. Filters are presented as diamonds and lists of fragment ions in light blue squares. Optional scans and functionalities are connected by dotted lines. 1) An MS1 scan is acquired through the instrument method to determine the presence of the trigger peptide labeled with TMTsh (precursor mass error ±10 ppm). 2) Upon detection of the peak of a potential trigger peptide, Tomahto inserts a trigger MS2 scan (Orbitrap MS2) by isolating and fragmenting the precursor. 3) Tomahto performs real-time peak matching (RTPM) on the trigger MS2 scan to confirm the sequence (>5 matched peaks, mass error ±10 ppm). 4) If trigger peptide confirmed, Tomahto prompts the insertion of a target MS2 (Orbitrap MS2) using the theoretical m/z of the target peptide. 5) Tomahto sequences the target MS2 by RTPM. 6) Only fragment ions common to trigger and target MS2 scans are included for further assessment. 7) Fragment ion peak abundances relative to the highest fragment ion peak are assessed. Only those ions with <50% relative abundance change compared to the preceding trigger MS2 are included for next step. 8) b- and y-ions are kept for consideration as SPS ions if lysine is at the C terminus, otherwise

only b-ions. 9) The isolation purity is calculated for each remaining SPS ion candidate and only those having a purity value greater than 0.5 (at least 50% of the signal attributed to the fragment ion within a 3 m/z window) are included as final SPS ions. 10) If MS3 prescan is enabled, Tomahto inserts an ion trap SPS-MS3 prescan (normal scan mode; AGC target of $1e^6$; max injection time of 10 ms; scan range 125-132). This is used to estimate the injection time needed to accumulate sufficient TMT reporter ions. 11) The estimated injection time is used to set the lengthy injection time for the quantitative SPS-MS3 scan acquired in the Orbitrap (resolution of 50,000; SPS ions from part 2, 0.5 m/z window, max injection time of 5,000 ms). 11) A close-out on a trigger peptide is initiated when 3 MS3 scans are collected on any given target peptide with a minimum of 1,000 summed signal-to-noise value. This excludes the peptide from triggering any further scan events.

A


Tomahto
Instrument
Data
Method
TOMAHAQ Method Editor
Real Time View
Data Analysis
Load Target Peptides
Target List (.csv - Columns = "Peptide") Optional: Gene Symbol, Charge State (z), Retention Time (RT)
Browse
Stats
Scans Arrived
MS2 Sent
MS2 Arrived
MS3 Sent
MS3 Arrived
Closed out
Modifications
Trigger
Target
Name
Mono Mass
Type
Symbol
Sites
TMT0
TMT2
TMT10
TMT11
TMTsh
CTer...
PepC-K
PepC-R
Instruction:
1. Browse peptide list
2. Choose PTM
3. Connect to instrument
4. Choose log file output folder
5. Choose method parameters
6. Load parameter (once loaded cannot change)
7. Subscribe to scan
8. Start acquisition
Identification
RT-Search
Peak Match
Trigger Fragment Ion Tolerance
PPM
DA
Target Fragment Ion Tolerance
SPS-MS3 Property
Prec Exclusion Window
SPS Ion Range (Th)
Below
Above
SPS Above Precursor MZ
Num. of SPS Ion
SPS Ion Cutoff (% of BP)
MS2 Iso. Width
MS3 HCD CE
MS3 OT Res.
Insert prescan
MS3 AGC
MS3 MaxIT
Target Peptide
Peptide
Z
Gene
Trigger M/Z
Target M/Z
Log File Option
Log Export Folder
Start Writing Log
Method
Method Length (min)
Schedule Method
None
RT Window (±min)
Dynamic Exclusion
Tol. (±ppm)
Duration (s)
MS2 Property
Trigger Intensity Threshold
Isolation Width
Activation Type
NCE
Max IT
AGC
Analyzer
IT Scan Rate
OT Reso.
Close-out
Sum SN
# of Scans
Reset by Vial Location
Start Close-out
Load Parameter
Start Acquisition
Reference:
1. Erickson BK, et al. (2017). Molecular cell 65(2):361-370.
Server Connected: False
System Mode:
System State:
Instrument Name:
Instrument Connected: False

B


Tomahto
Instrument
Data
Method
TOMAHAQ Method Editor
Real Time View
Data Analysis
Scan Events
Peptide
Z
Gene
Trigger M/Z
Target M/Z
Trigger MS2
Target MS2
Target MS3
Graphs
Trigger MS2
Target MS2
Reporter Ion SN
Server Connected: False
System Mode:
System State:
Instrument Name:
Instrument Connected: False

C

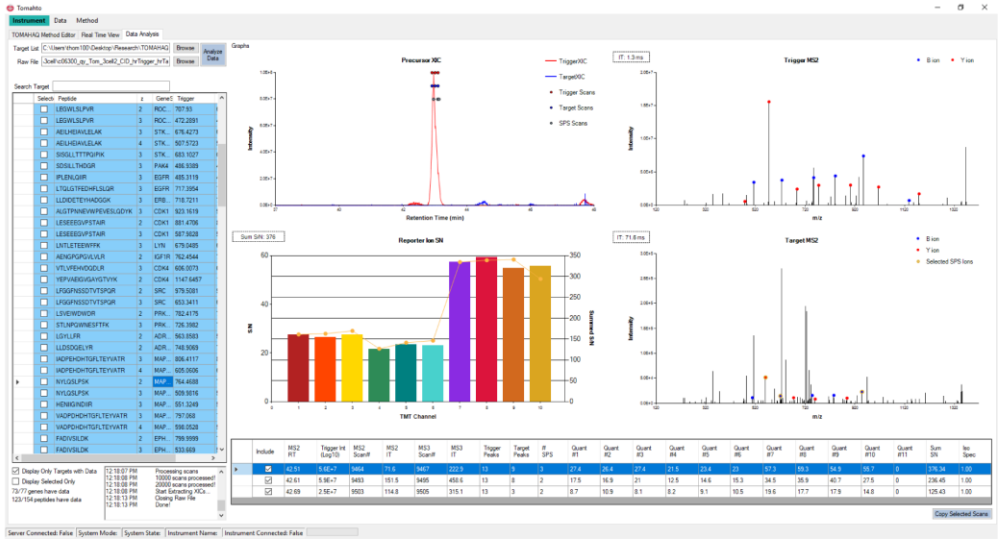

Tomahto
Instrument
Data
Method
Precursor XIC
Trigger MS2
Reporter Ion SN
Target MS2

Figure S2. Graphic user interface of Tomahto. (A) Method setup page allows direct control of the instrument and greatly facilitates experiment setup. Steps to set up a Tomahto experiment include: 1). Browse and select peptide list; 2). Select peptide modifications; 3) Choose MS parameters; 4). Load method and start acquisition. **(B)** Real-time data visulazation page. When Tomahto acquires a quantification event, consisting of a trigger MS2, a target MS2 and an SPS-MS3 scan, it plots the MS2 scans with matched peaks labeled, and TMT reporter signal-to-noise values to enable data visualization in real-time. **(C)** Post-acquisition data analysis. Tomahto loads raw data file and lists all peptide targets. Selecting one peptide will show all quantification events for it and selecting one scan event will plot four figures, including MS1 extracted chromatogram, trigger MS2, target MS and TMT quantification values. Data can be exported to .csv or .txt files for further analysis.

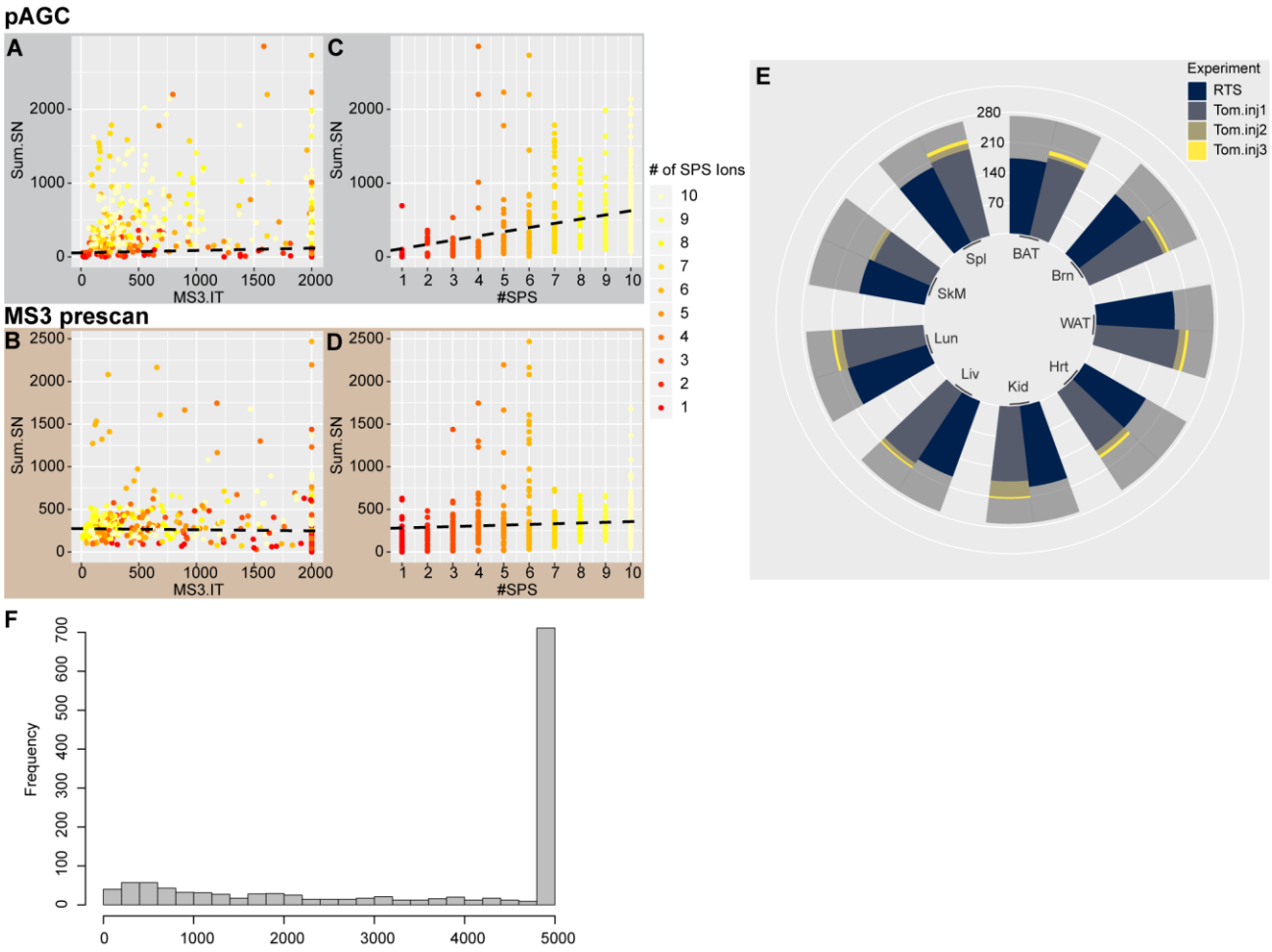


Figure S3. Evaluation of the Tomahto method, related to Figure 2 and 3. (A, B) Summed signal-to-noise (SN) versus injection time. Results suggested an increasing trend of summed SN with increasing MS3 injection time using instrument method-based pAGC whereas summed SN was relatively constant when MS3 prescans were used, suggesting better injection time estimate with Tomahto-inserted prescans. **(C, D)** Summed signal-to-noise (SN) versus number of SPS ions. Results suggested heavy dependence of summed SN on number of SPS ions when using instrument method-based pAGC whereas it was more consistent when Tomahto-inserted MS3 prescans were used. **(E)** Stress test of Tomahto. Number of protein targets quantified in DDA-RTS and three technical replicates of Tomahto, with close-out enabled. Quantifiable targets were defined as protein targets quantified in any of the 3 Tomahto technical replicates. Tomahto was able to cover 90% of all quantifiable targets with single shots and 97% of all quantifiable targets with two injections. Tomahto was able to target >500 peptides (>1000 precursors) in single shots. **(F)** Frequency distribution of SPS-MS3 scan injection times with prescans. A max injection time of 5,000 ms was allowed and ~54% of all MS3 scans met the maximum, suggesting targeted peptides were present at low levels.

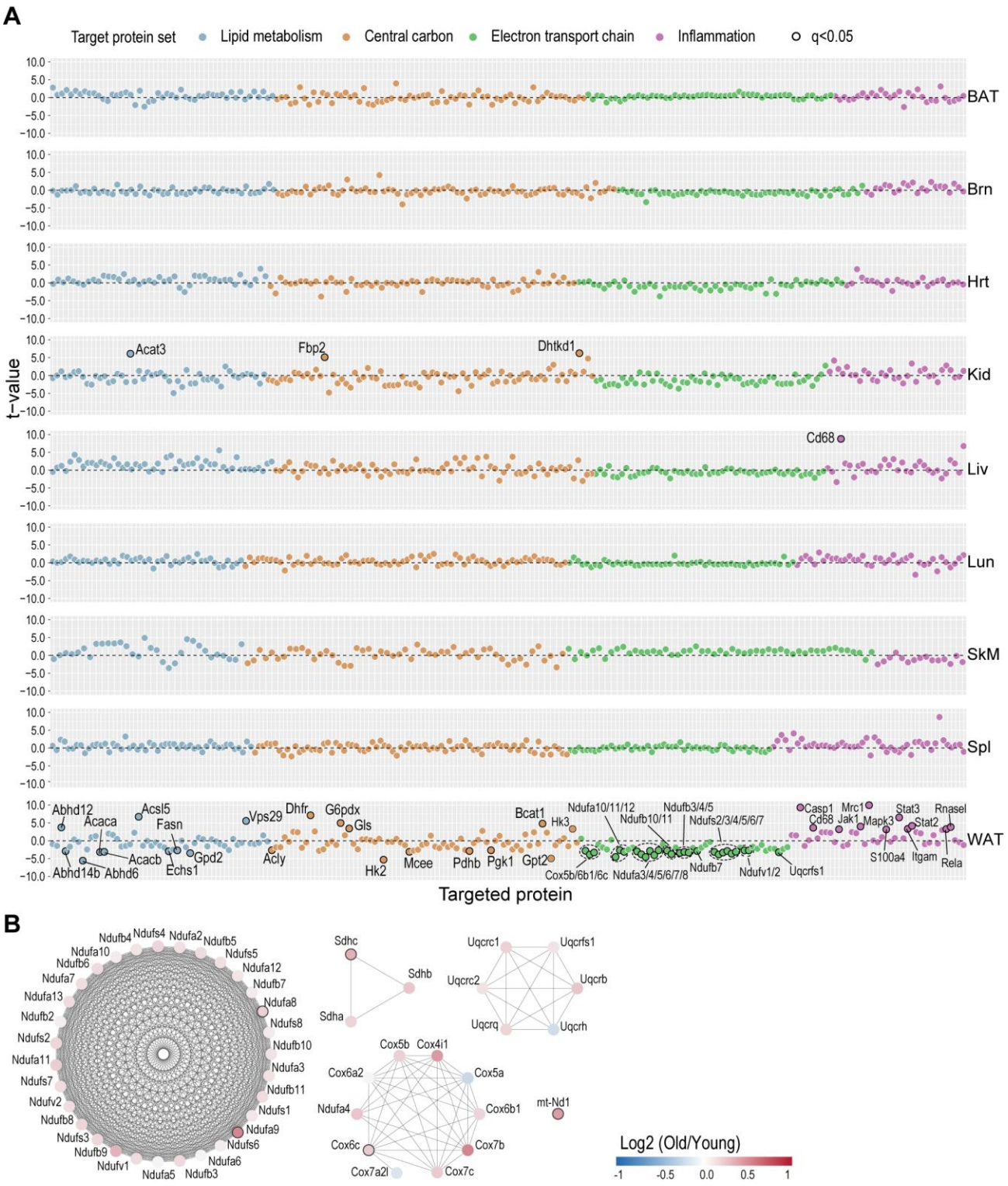


Figure S4. Targeted quantification with Tomahto, related to Figure 3. (A) T statistics of each target in 9 tissues. Protein quantifications were evaluated using t-tests and resulting p values were corrected for multiple hypothesis testing. Significantly regulated proteins (q<0.05) were circled and labeled. Most tissues presented minimal disturbance whereas WAT underwent significant changes in multiple targeted pathways. **(B)** Log_2 ratios of ETC complex subunits in skeletal muscle. Though 7 out of 9 investigated

tissues presented down-regulation of ETC complexes, SkM and BAT showed the opposite trend. Quantification of ETC complex subunits in SkM by Tomahto suggested up-regulation.

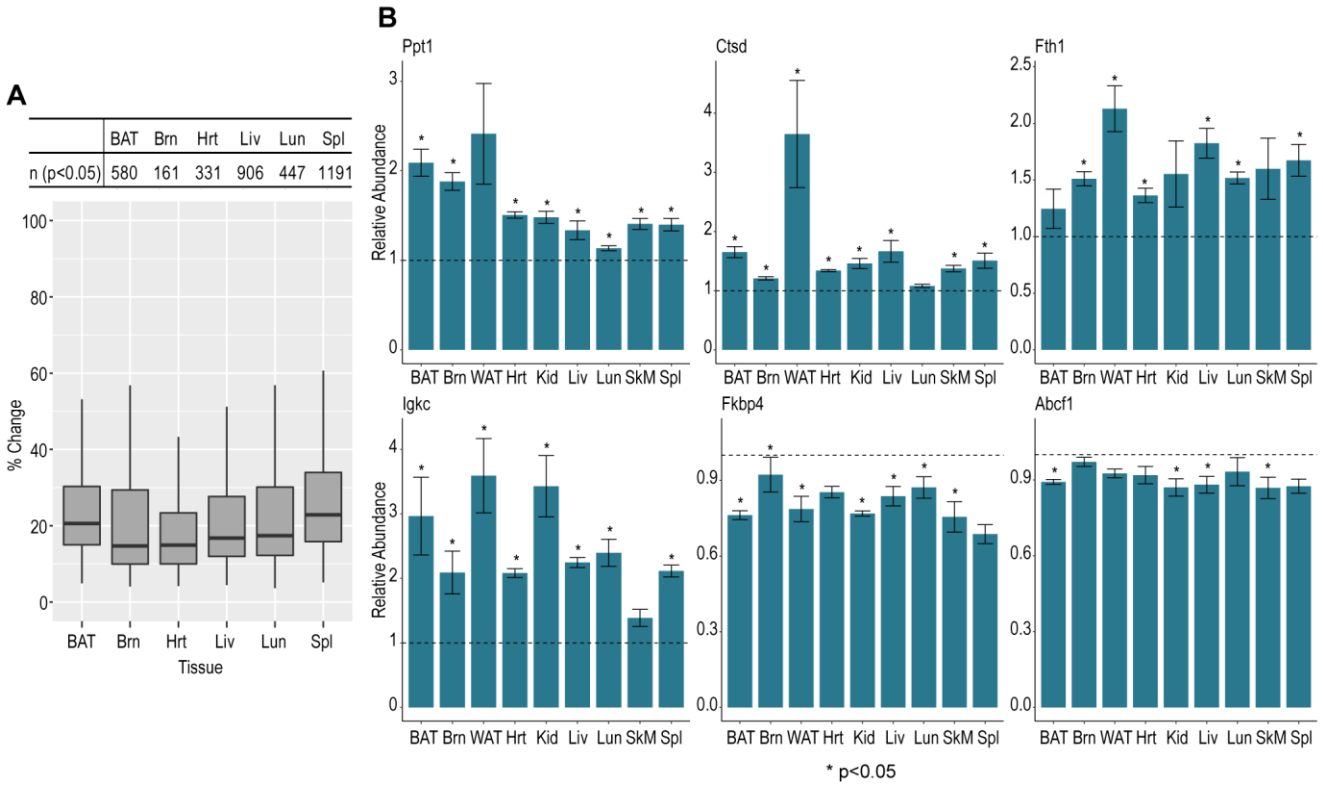


Figure S5. Example protein changes, related to Figure 5. (A) Protein (p<0.05) fold change distributions in BAT, Brn, Hrt, Liv, Lun and Spl. The majority of the included proteins only had ~20% change in old mice. **(B)** Example proteins showing consistent changes in all tissues Ppt1, Ctsd and Fth1 are lysosome-residing proteins involved in pathways such as lipid metabolism and autophagy and they showed up-regulation in 9 tissues. Igkc is a secreted protein involved in immune response. Other examples include Fkbp4, an immunophilin protein with prolyl isomerase and co-chaperone activities, and Abcf1, mediator of mRNA translation initiation.

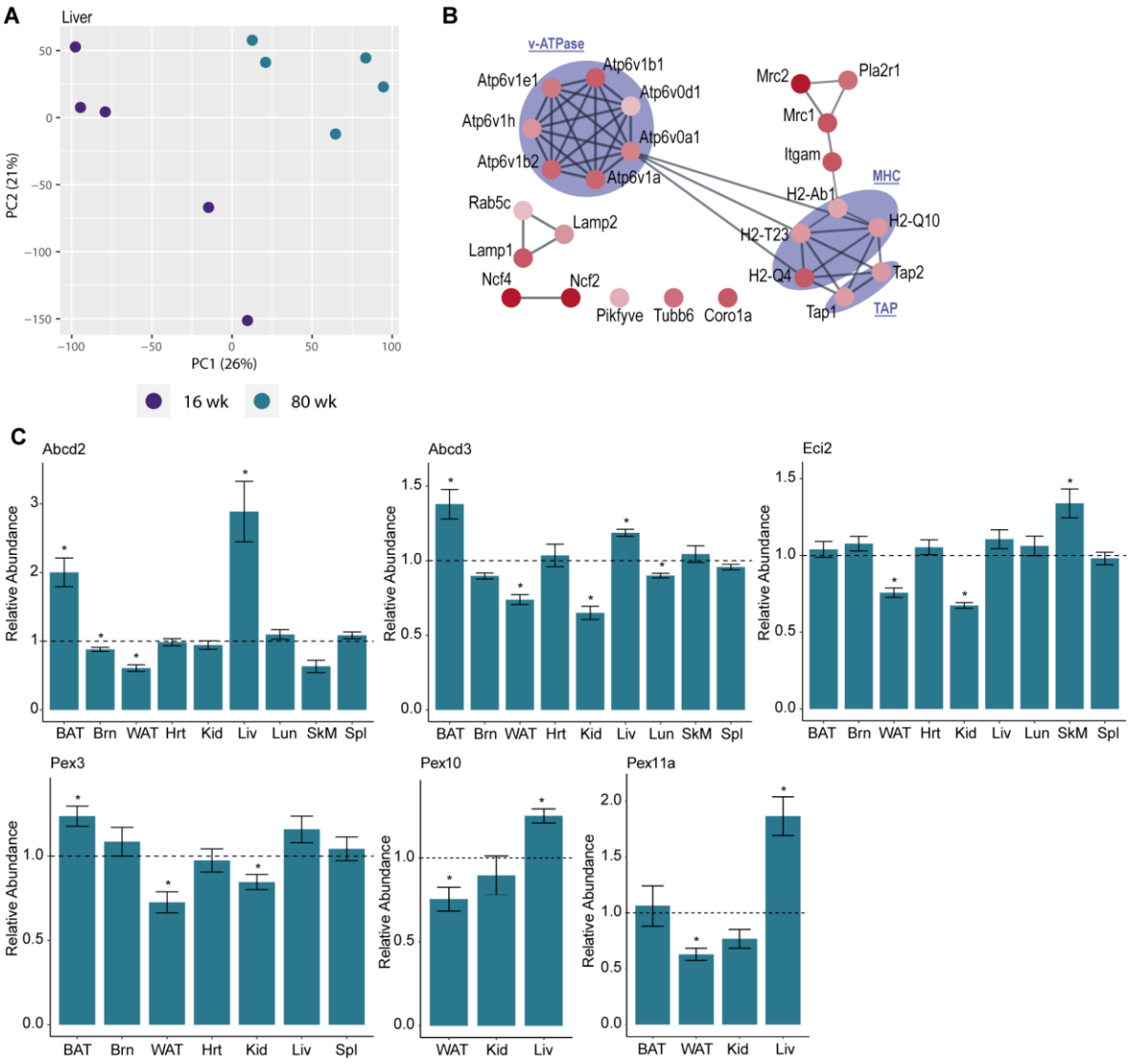


Figure S6. Full proteome characterization of aging mouse tissues, related to Figure 6. (A) Major components separating old from young in liver**.** (**B)** Significantly regulated (q<0.05) phagosomal proteins in WAT. Subunits of vacuolar ATPase, MHC and TAP complexes are highlighted in purple. (**C)** Bar plots for representative peroxisomal proteins. * indicates p<0.05. Error bars represent means ± SEM (n=5).

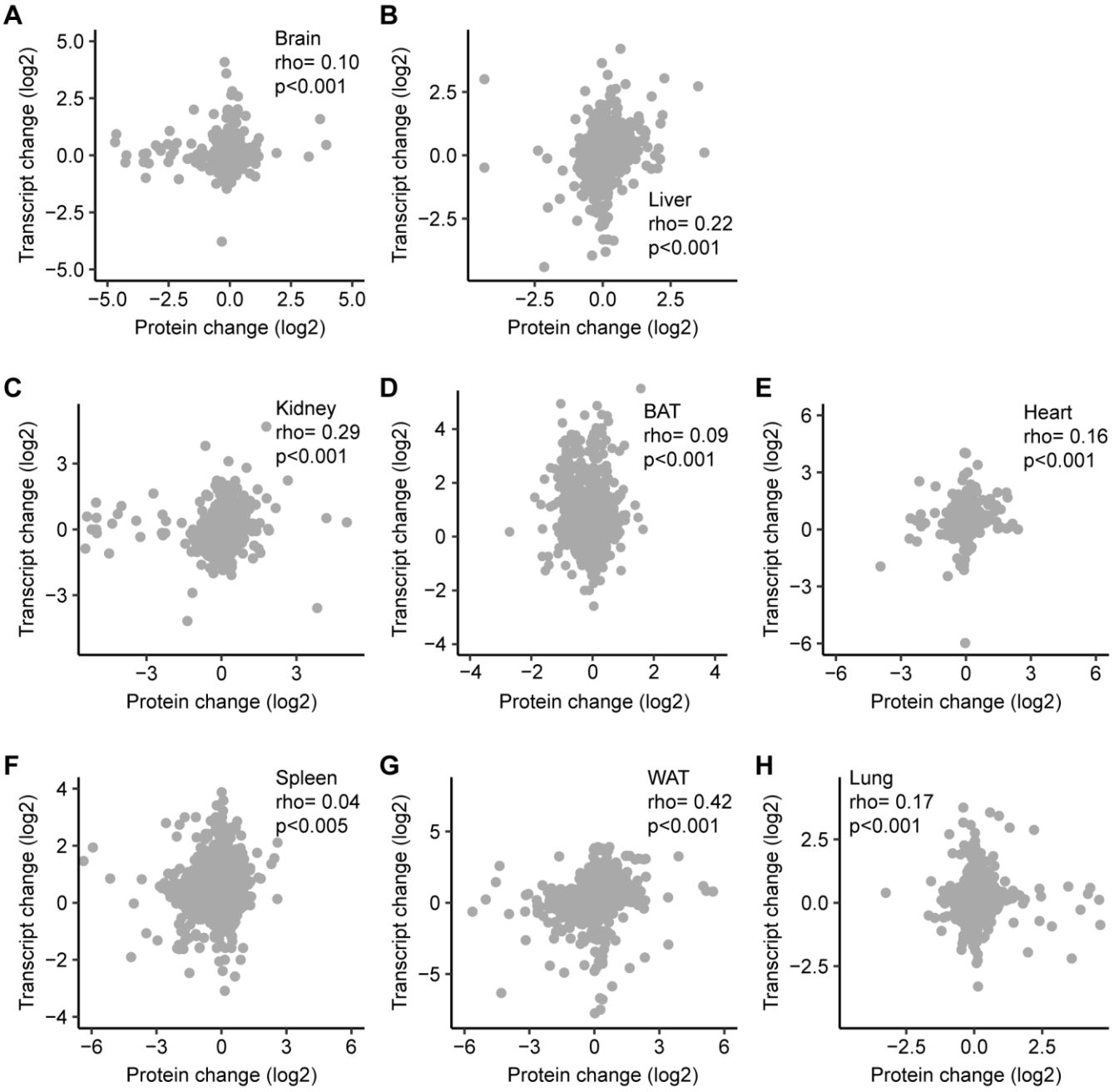


Figure S7. Correlations between changes in transcript and protein abundance. Protein ratio changes were compared to transcript ratio changes obtained from the published dataset (10). Due to the absence of identical age groups, transcript data from 3 month- and 21 month-old male mice were used in the comparison. Brain **(A)** and liver **(B)** showed comparable proteome and transcriptome correlation to the previous study in rat (rho =0.10 vs 0.13 in brain; rho =0.22 vs 0.25 in liver), though different organisms and age groups were studied (11). WAT **(H)** had the best correlation among all tissues investigated (rho= 0.42).

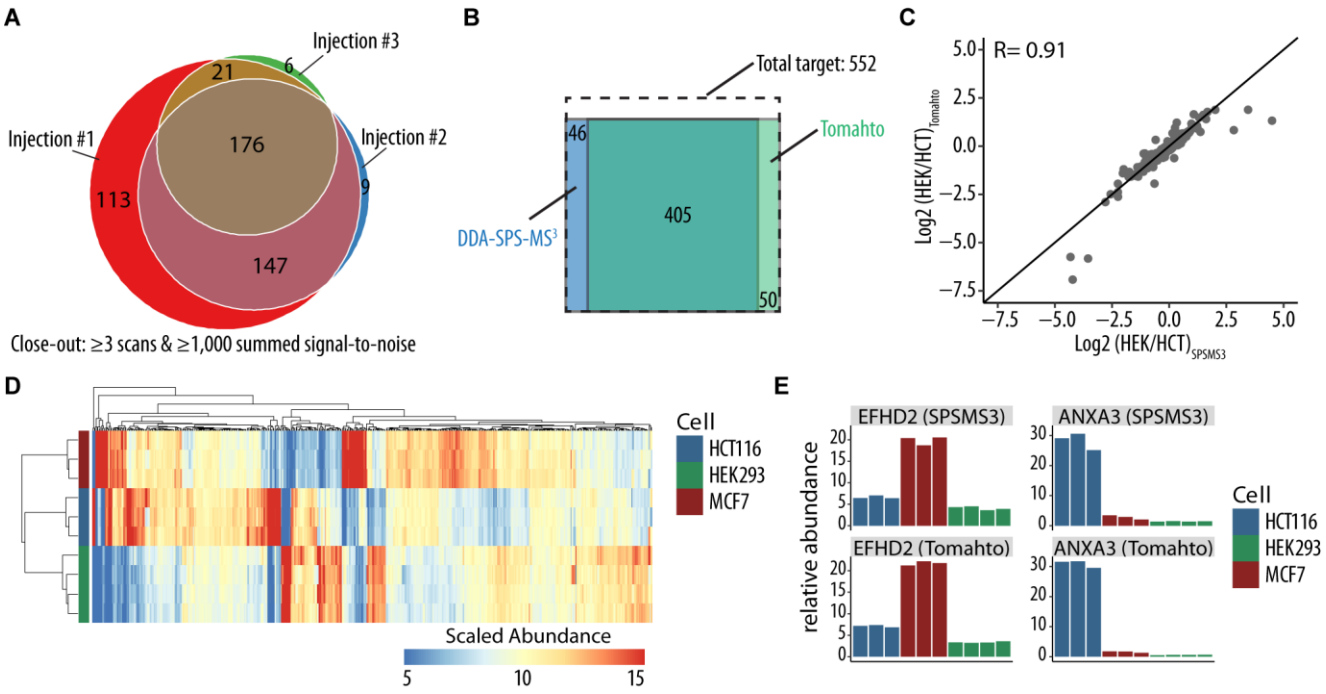


Figure S8. Tomahto analysis of 1,000 cross-species proteotypic peptides. 1000 proteotypic peptides, corresponding to 552 proteins, were obtained from JPT Peptide Technologies (Catalog #: SPT-ABRF-POOL-L). The peptides were labeled with TMTsh and spiked into the three cell line proteome sample and analyzed using Tomahto. Three injection replicates were performed with the close-out option (⩾3 scans & ⩾1,000 summed signal-to-noise). Most proteins were quantified by the first replicate **(A)** and the subsequent two replicates added marginal number of proteins. **(B)** The coverage were comparable to the regular DDA-SPS-MS^3 experiment acquired with 36-hour instrument time. **(C)** Quantitative correlation between Tomahto and SPSMS^3. Log2 ratios of HEK293T to HCT116 calculated. Correlation coefficient between the two methods was 0.91. **(D)** Heatmap showed clear clustering of three cell lines based on the quantified target proteins. **(E)** Example proteins (EFHD2 and ANXA3) quantified by SPSMS^3 and Tomahto.

Reference

1. Wollenberger A, Ristau O, & Schoffa G (1960) [A simple technic for extremely rapid freezing of large pieces of tissue]. *Pflugers Arch. Gesamte Physiol. Menschen Tiere* 270:399-412.
2. Navarrete-Perea J, Yu Q, Gygi SP, & Paulo JA (2018) Streamlined Tandem Mass Tag (SL-TMT) Protocol: An Efficient Strategy for Quantitative (Phospho)proteome Profiling Using Tandem Mass Tag-Synchronous Precursor Selection-MS3. *J. Proteome Res.* 17(6):2226-2236.
3. Schweppe DK*, et al.* (2019) Full-featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics. *bioRxiv*:668533.
4. Eng JK, McCormack AL, & Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass. Spectrom.* 5(11):976-989.
5. Huttlin EL*, et al.* (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 143(7):1174-1189.
6. Savitski MM, Wilhelm M, Hahne H, Kuster B, & Bantscheff M (2015) A Scalable Approach for Protein False Discovery Rate Estimation in Large Proteomic Data Sets. *Mol. Cell. Proteomics* 14(9):2394-2404.
7. Strimmer K (2008) fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* 24(12):1461-1462.

8. Yu G, Wang LG, Han Y, & He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5):284-287.
9. Shannon P*, et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13(11):2498-2504.
10. Schaum N*, et al.* (2019) The murine transcriptome reveals global aging nodes with organ-specific phase and amplitude. *bioRxiv*:662254.
11. Ori A*, et al.* (2015) Integrated Transcriptome and Proteome Analyses Reveal Organ-Specific Proteome Deterioration in Old Rats. *Cell Syst.* 1(3):224-237.