nature portfolio

https://doi.org/10.1038/s41596-022-00780-w

Supplementary information

High-quality and robust protein quantification in large clinical/ pharmaceutical cohorts with IonStar proteomics investigation

In the format provided by the authors and unedited

Supplementary Information

Supplementary Methods. Reagents, equipment, and procedures for sample procurement and pre-processing.

Supplementary Figure 1. Example of column pressure profiles for trapping and separation columns during a 2.5-hour gradient.

Supplementary Figure 2. Base-peak ion chromatograms (BICs) from 3 benchmark sample (BS) runs (at the beginning, in the middle and at the end of LC-MS analysis, over a 10-day period).

Supplementary Figure 3. Alignment of chromatographic profiles from 5 quality control sample (QS) runs over an 8-day period.

Supplementary Figure 4. Extracted bovine serum albumin (BSA) ion chromatogram for performance evaluation of a new separation column.

Supplementary Table 1. The summary of the basic principles, advantages, and limitations of different quantification techniques.

Supplementary Table 2. Summary of the most popular and/or state-of-the-art software tools for label-free quantitative proteomics.

Supplementary Table 3. An example of LC gradient for off-line immunoaffinity chromatography for depletion of high-abundance proteins from human serum using a MARS14 column.

Supplementary Methods

Reagents, equipment, and procedures for sample procurement and pre-processing, including:

A. Cryo-pulverization of tissue samples; B. Cell harvesting from Petri dishes; C. Depletion of high-abundance proteins from plasma/serum samples; D. Protein bicinchoninic acid assay (BCA).

REAGENTS

- Buffer A, Equilibrate/Load/Wash, for MARS14 (Agilent, part no.: 5185-5987)
- Buffer B, Elution, for MARS14 (Agilent, part no.: 5185-5987)
- 10X phosphate-buffered saline (PBS, Corning®, VWR, cat. no. 45001-130)
- HyClone™ Fetal Bovine Serum (FBS, Cytiva, formerly GE healthcare, cat. no. SH3007002)
- TrypLE™ Express Enzyme (1×), phenol red (Gibco®, Thermo Fisher Scientific, cat. no. 12605010)
- Bovine serum albumin (BSA, Thermo Fisher Scientific, cat. no. NC9771311)

EQUIPMENT

Sample processing

- Disposable Pellet Mixers (VWR, cat. no. 47747-358)
- Cordless Pestle Motor (VWR, cat. no. 47747-370)
- 96 well clear flat bottom plate (Falcon, product no. 353072)
- Mortar 70 × 90 mm (VWR, cat. no. 89038-148)
- Pestle 114 mm (VWR, cat. no. 89038-160)
- Incubator (VWR, cat. no. 9120979)
- Vortex Mixer (VWR, cat. no. 97043-562 or similar, speed range 300-3200 RPM)

Off-line immunoaffinity chromatography for serum depletion

- LC system for immunoaffinity chromatography (Waters 2796 Bioseparations Module, Waters, model no. BA1)
- PEEK tubing (125 µm I.D., IDEX Health & Science, cat. no. 1535L)
- Autosampler (2500 µL syringe, Waters, cat. no. WAR077342; 100 µL sample loop, Waters, cat. no. 430000783; 3000 µL holding loop, Waters, cat. no. 700000557)
- Waters 2996 Photodiode Array Detector (Waters, model no. 2996)
- Multiple Affinity Removal Column, Human 14 (MARS14, Agilent, cat. no. 5188-6557)
- 5K MWCO spin concentrator (Agilent, cat. no. 5185-5991)
- 0.22 µm spin filter (Agilent, cat. no. 5185-5990)

Software

- Waters Empower 2 (Version 6.10.00.00, Waters)

REAGENT SETUP

- 1× PBS buffer: To prepare 1 L, dilute 100 mL 10× PBS by adding 900 mL water.

PROCEDURES

A. Cryo-pulverization of tissue samples ●TIMING ~2 h (per 50 samples)

!CAUTION It is recommended to perform Step i) – v) in a hood to avoid exposure to spilled tissue particulates. Use appropriate protection measures and avoid direct contact with liquid nitrogen.

i) Cover the work area with an underpad and place all equipment on top. Put all sample tubes into an insulated container filled with liquid nitrogen.
ii) Weigh each frozen tissue sample. Carefully examine the appearance, color, shape, and size of each tissue and record any abnormality; consider excluding samples with excessive blood contamination, abnormal appearance, color, or shape.
iii) Pre-chill a clean set of mortar and pestle with liquid nitrogen for 1 min.
iv) Transfer the tissue to a pre-chilled mortar (containing enough liquid nitrogen to immerse the tissue) using tweezers. Pulverize the tissue into a fine powder using the pestle. Pour more liquid nitrogen into the mortar during the procedure if necessary.

▲**CRITICAL** Make sure that there is always liquid nitrogen in the mortar during the pulverization procedure, which allows facile tissue pulverization with minimal protein degradation.

v) Pre-chill a spatula and an Eppendorf tube in liquid nitrogen for about 40 s. Collect the pulverized tissue powder into the tube.

▲**CRITICAL** Transfer tissue powder into the tube as fast as possible to avoid moisture condensation. Leave the Eppendorf tube open for ~3 s after collection to ensure that all liquid nitrogen evaporates.

■**PAUSE POINT** The pulverized tissue powder can be stored under -80 °C until further processing.

▲**CRITICAL** Cryo-pulverization may cause substantial sample loss for small amounts of soft tissues (*e.g.,* <5 mg of brain or liver tissues). Instead, the user could choose to homogenize the tissue chunks for 30 s using a cordless pestle motor and disposable pellet mixers.

B. Cell harvesting from Petri dishes ●**TIMING ~5 h (per 50 samples)**

i) Collect cell medium containing floating cells into a 15 mL centrifuge tube (if floating cells are needed in the project; otherwise discard cell medium).
ii) Gently add 2 mL 1× PBS to wash the dish and remove any residual medium (which will lower trypsin activity). Aspirate and discard PBS.
iii) Add 1 mL of pre-warmed TrypLE™ to the dish and incubate at 37 °C for 3 min (time depends on the adherence properties of different cell lines).
iv) Add 4 mL of cell culture medium with FBS to terminate trypsinization and collect the detached cells into the same centrifuge tube with floating cells (collected in step i).
v) Centrifuge at 400 g for 5 min at 4°C to pellet the cells and remove the supernatant carefully. (Optional) Before centrifugation, aliquot 200 µL for cell counting.
vi) Add 5 mL of cold 1× PBS into the tube and gently pipet to wash the pellet. Centrifuge and discard the supernatant. Repeat this step two additional times and remove the supernatant as much as possible.

▲**CRITICAL** Minimize cell rupture during cell harvesting, *e.g.,* by avoiding higher centrifugation speed and vigorous pipetting.

■**PAUSE POINT** Cell pellets can be stored under -80 °C until further processing.

C. Depletion of high-abundance proteins from plasma/serum samples ●**TIMING ~24 h (per 50 samples)**

The depletion of high-abundance proteins in human plasma/serum is performed using Multi Affinity Removal Column, Human 14 (MARS14), which showed exceptional reproducibility in our benchmark studies. Procedures in this protocol are adapted from the manufacturer's manual (https://www.agilent.com/cs/library/usermanuals/public/5990-7226.pdf) with adjustments to achieve optimal performance.

i) Thaw plasma/serum samples at room temperature. Take 20 µL (for 4.6×50 mm MARS14 column) and dilute 4-fold with Buffer A with protease inhibitor tablets (1 tablet in 10 mL Buffer A).
ii) Under 4 °C, centrifuge the diluted plasma/serum samples on 0.22 µm-spin filters at 16,000 g for 1.5 min to remove particulates.
iii) Inject 80 µL diluted plasma/serum sample and run the LC gradient (an example is in **Supplementary Table 3**).
iv) Acquire chromatogram at 280-nm absorbance. Collect the flow-through fraction in a ±2 min window around the apex of the peak of depleted proteins (may vary depending on the LC system).
v) Centrifuge the collected fraction using a 5K MWCO spin concentrator at 2,000 - 7,500 g for 20 - 30 min at 4 °C to concentrate.

▲**CRITICAL** It is necessary to achieve 5× concentration. Increase centrifugation speed or prolong the centrifugation time if needed.

vi) Add 500 µL 50 mM Tris-FA (pH 8.5) to the concentrator and repeat step v). Repeat this step at least once to efficiently reduce nonvolatile salts in buffer A.
vii) Recover the concentrated sample in Tris-FA to a final volume of 100 µL.
▲**CRITICAL** For a brand-new column or a used column stored for over 2 weeks, depletion efficiency should be evaluated before application on real samples by consecutively running a pooled sample for 3 times and calculating depletion efficiency based on BCA results before and after depletion using the following formula: (protein amount after depletion/ protein amount in non-depleted serum/plasma before injection) × 100%. Mean depletion efficiency should be 95±2% to ensure efficiency and reproducible depletion.

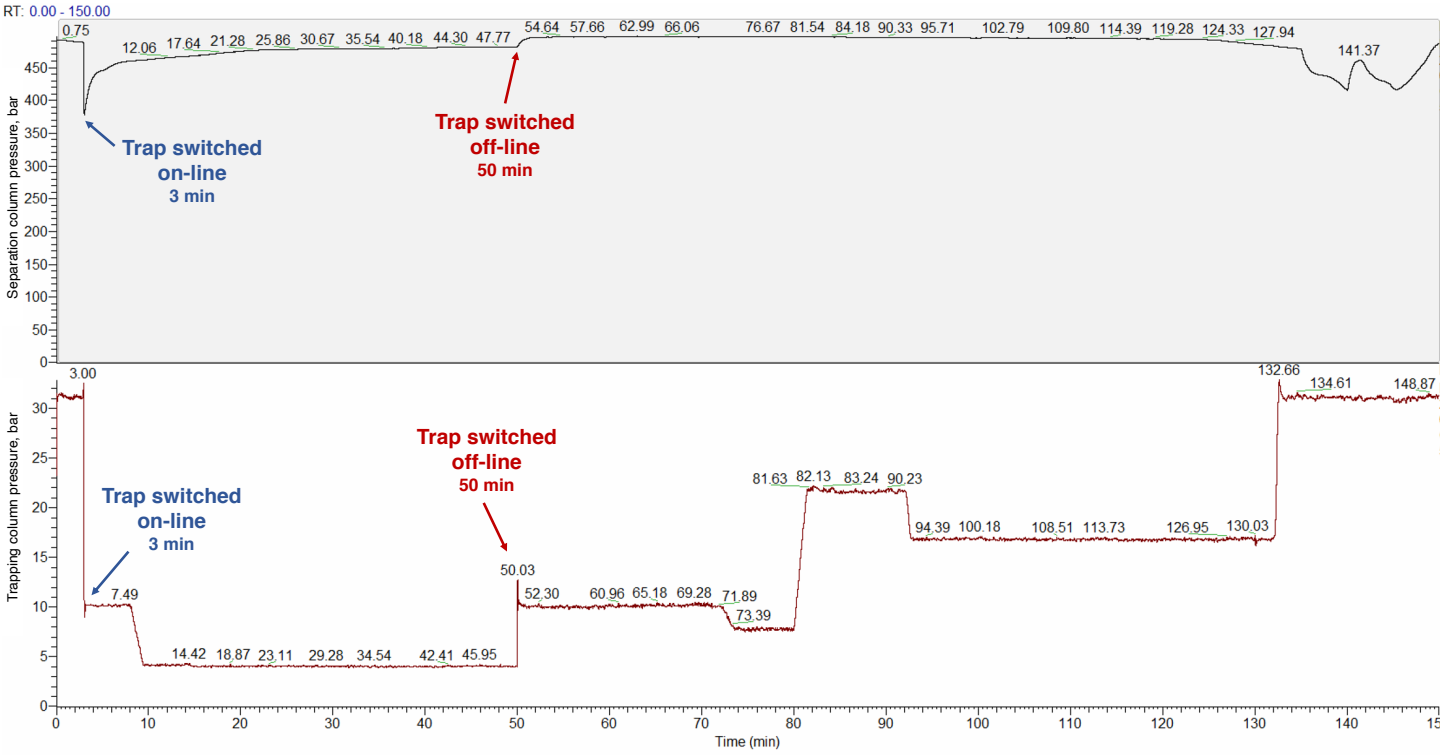
■**PAUSE POINT** Depleted samples can be stored under -80 °C until further processing.

D. Protein bicinchoninic acid assay (BCA) ●**TIMING ~1 h (per 50 samples)**

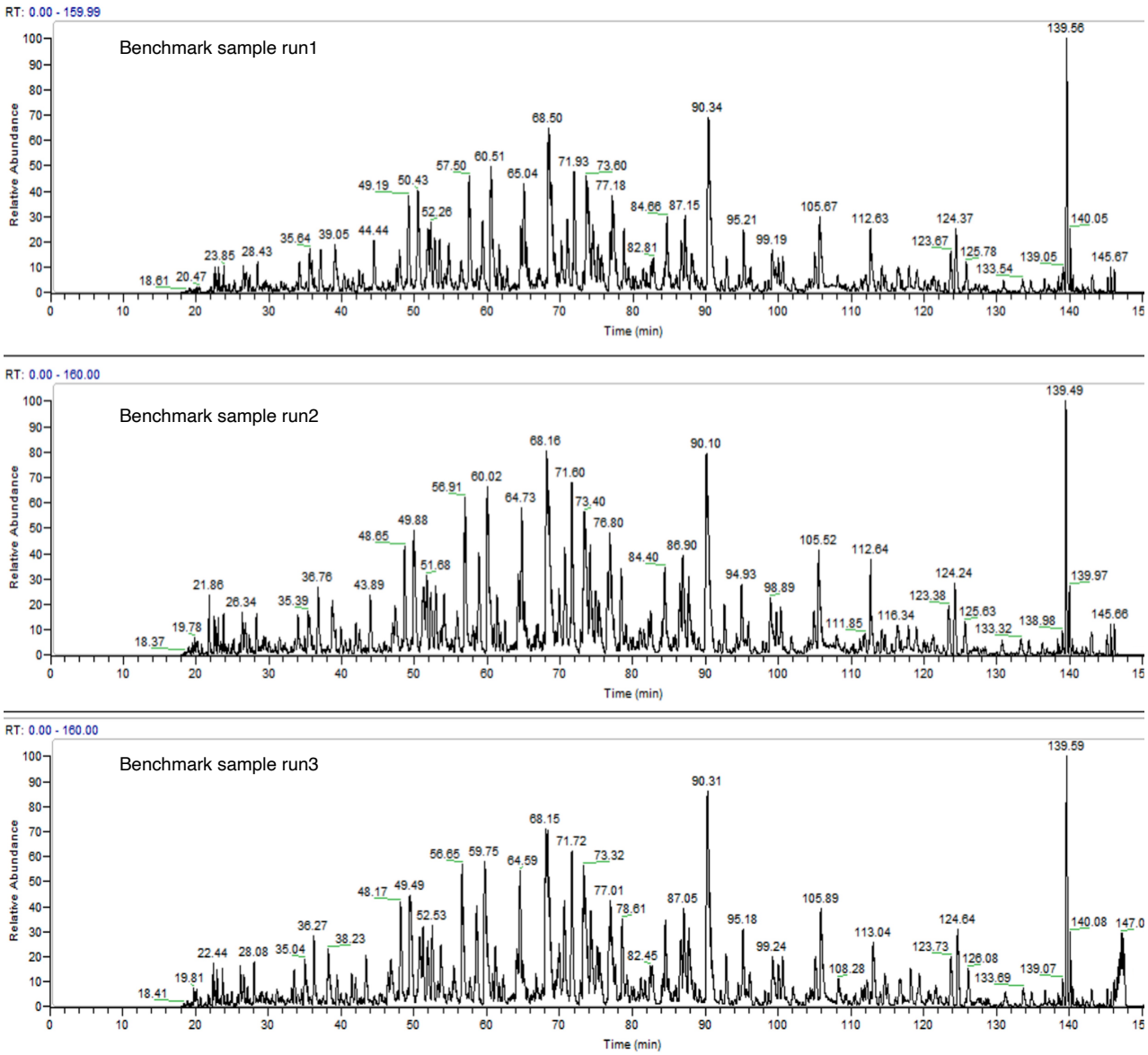
Protein assay is necessary to normalize the amount of total protein prior to proteolytic digestion. In this protocol, we use bicinchoninic acid assay (BCA), but any protein assay compatible with surfactants and the sample matrix can be used. Procedures in this protocol are adapted from the manufacturer's manual. (https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011430_Pierce_BCA_Protein_Asy_UG.pdf)

i) Prepare the standard curve as described in the user manual and dilute the samples to reach the linear concentration range of the kit (*e.g.,* 0.2~2 µg/µL). A volume of 40 µL is recommended for both calibration standards and samples.

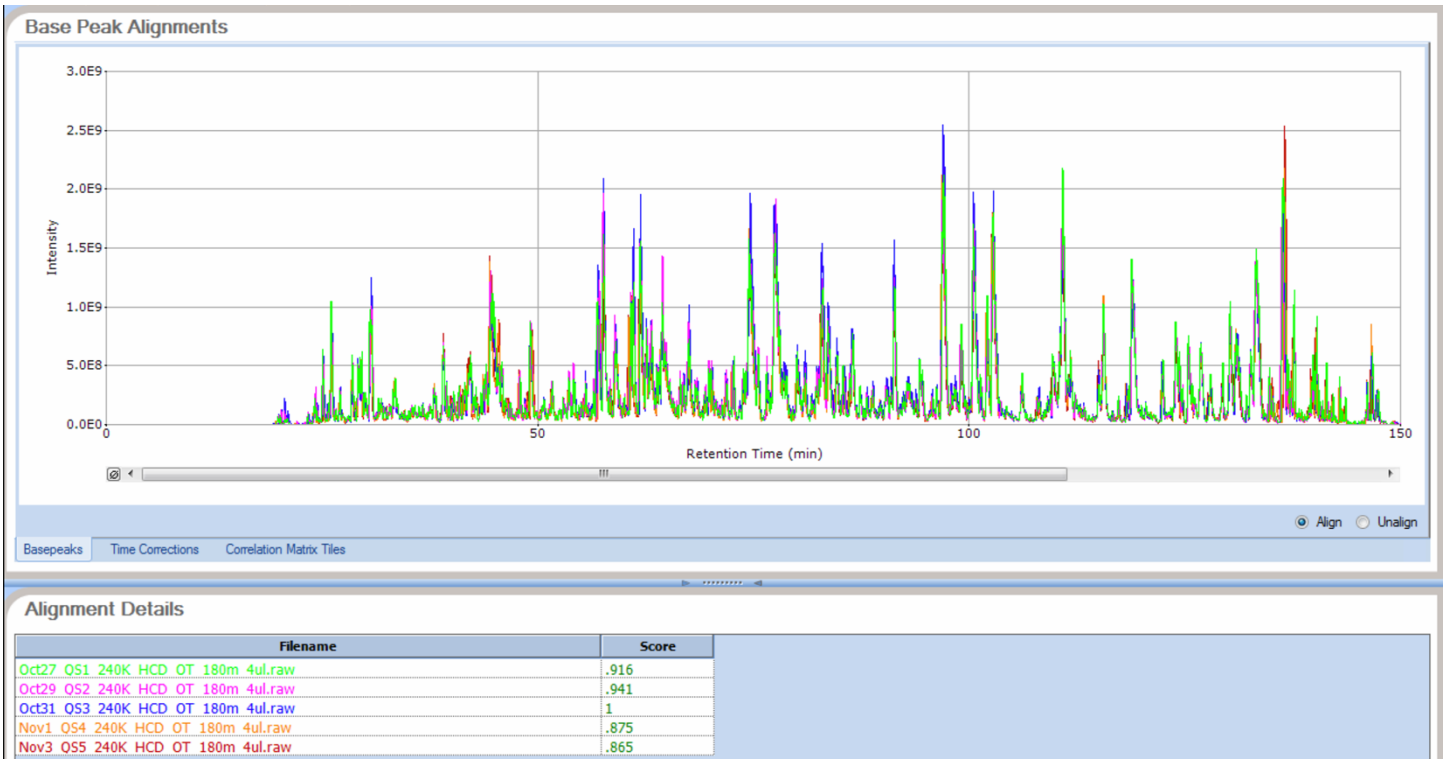
▲**CRITICAL** Prior to dilution, protein concentrations can be estimated by this rule of thumb: 10 -15% wet weight of cell pellets or tissues are proteins extractable by the SC buffer. As to plasma, usually, a 50-time dilution is a reasonable start.
ii) Add 10 µL of each calibration standard and sample in triplicate to a 96-well plate.
iii) Mix BCA Reagent A and B (50:1, vol/vol). Add 190 µL to each well.
iv) Incubate at 37 ºC for 30 min, and then read the plate with a microplate reader at 562 nm absorbance. Establish the calibration curve and calculate the protein concentration of each sample.



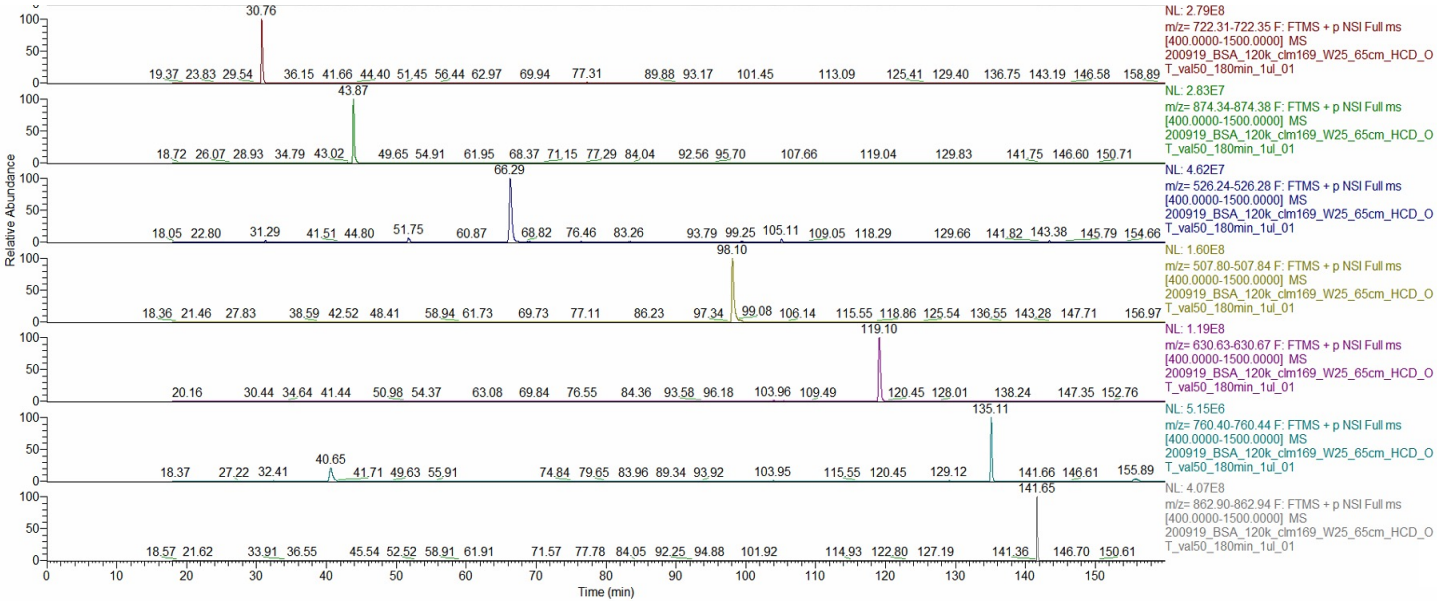
Supplementary Figure 1. Example of column pressure profiles for trapping and separation columns during a 2.5-hour gradient.



Supplementary Figure 2. Base-peak ion chromatograms (BICs) from 3 benchmark sample (BS) runs (at the beginning, in the middle and at the end of LC-MS analysis, over a 10-day period).



Supplementary Figure 3. Alignment of chromatographic profiles from 5 quality control sample (QS) runs over an 8-day period.



Supplementary Figure 4. Extracted bovine serum albumin (BSA) ion chromatogram for performance evaluation of a new separation column. The panels from top to bottom are the extracted ion chromatograms of representative peptides at *m/z* = 722.33, 874.36, 526.26, 507.82, 630.65, 760.42, and 862.92 within an *m/z* window of ± 0.02 Th.

Supplementary Table 1. The summary of the basic principles, advantages, and limitations of different quantification techniques.

Labeling condition	Representative methods	Quantification method	Identification method	Advantages	Limitations
Labeling methods	Isobaric labeling: the most popular (*e.g.,* TMT/iTRAQ)	MS2 and or MS3 AUC or intensity	MS2, Database-searching	- Excellent depth of analysis - High intra-batch reproducibility	- High cost - Limited number of samples that can be quantitatively compared in one labeling batch - Multiple batches are needed for large cohorts, which cause elevated missing data
Non-labeling method	Spectral counting (*e.g.,* SpC, NSAF)	MS2 frequency	MS2, Database-searching	- Straightforward data processing	- Under-sampling due to MS2 stochasticity - Low quantitative quality for low abundance proteins and high levels of missing data
	MS1-based	MS1 AUC or intensity	MS2, Database-searching	- High sensitivity of MS1 - higher quality than spectral counting - Streamlined data processing	- Under-sampling due to MS2 stochasticity - Suboptimal reproducibility for low abundance proteins and high levels of missing data for large cohort
	IonStar	UHR-MS1 AUC	MS2, Database-searching	- High sensitivity and selectivity by UHR-MS1, enabling accurate and reproducible quantification for low abundance proteins in large cohort samples	- More stringent requirement on the MS resolution and LC reproducibility
	MS2-DIA (*i.e.,* SWATH)	MS2 AUC or intensity	MS2, Spectral library-based	- Reproducible protein measurement and reduced missing value in analysis of large cohort samples	- Difficulties in interpreting MS2 spectra from multiple co-fragmented precursors - Elevated false-positives - Suboptimal quantitative accuracy for low abundance proteins

Supplementary Table 2. Summary of the most popular and/or state-of-the-art software tools for label-free quantitative proteomics.

Method	Quantification method	Availability	DOI
OpenSWATH	MS2-DIA-based	Open source	10.1038/nbt.2841
Spectronaut	MS2-DIA-based	Commercial	10.1074/mcp.M114.044305 10.1007/s11192-009-0145-4
DIA-Umpire	MS2-DIA-based	Open source	10.1038/nmeth.3255
Skyline	MS2-DIA-based; Targeted quantification with PRM/SRM; MS1-based	Open source	10.1038/nprot.2015.055
Scaffold	MS2-DDA-based	Commercial	10.1002/pmic.200900437
Progenesis QI	MS1-based	Commercial	10.1007/s00216-016-9482-3
PEAKS	MS1-based; MS2-DIA-based	Commercial	10.1074/mcp.M111.010587
MaxQuant	MS1-based	Open source	10.1038/nbt.1511
Proteome Discoverer	MS1-based	Commercial	N/A
OpenMS	MS1-based	Open source	10.1093/bioinformatics/btl299 10.1186/1471-2105-9-163 10.1038/nmeth.3959
IonStar/UHR-IonStar	MS1-based	Open source	10.1073/pnas.1800541115

Supplementary Table 3. An example of LC gradient for off-line immunoaffinity chromatography for depletion of high-abundance proteins from human serum using a MARS14 column.

Time (min)	Flow rate (mL/min)	Gradient (% Buffer B)
0.0	0.16	0
9.50	0.16	0
9.51	1	0
11.50	1	0
11.51	1	100
18.00	1	100
18.01	1	0
27.00	1	0