A new method for sensitive and accurate proteome quantification in neat plasma

Chuanping Zhao[1,†], Huiting Hu[3,†], Xiaoxi Dong[2], Xiaohong Qian[1], Wanjun Zhang[1]* and Weijie Qin[1,3]*

[1] State Key Laboratory of Medical Proteomics, National Center for Protein Sciences (Beijing), Beijing 102206, China.

[2] Beijing Proteome Research Center, Beijing 102206, China.

[3] School of Basic Medical Science, Anhui Medical University, Hefei 230032, China

† Chuanping Zhao and Huiting Hu contributed equally to this work.

* Corresponding author email: zwj2004zwj@126.com for Zhang W. J. and aunp_dna@126.com for Qin W. J.

Abstract

Blood is the primary source of protein biomarkers because it is readily accessible and reflects both physiological and pathological changes. However, ion suppression by high-abundance plasma proteins has long limited the coverage of mass spectrometry (MS)-based proteomics. Recent advances in nanomaterial-based enrichment have expanded blood proteome coverage to more than 6000 proteins. Nevertheless, whether the quantities of the enriched proteins proportionally reflect their true levels in neat blood – and thereby faithfully represent disease-associated variation – remains debated. We therefore developed NCTPro (Nanomaterial-enhanced Cross-Channel Transfer strategy for neat blood Proteomics), which enables in-depth plasma proteome quantification without enrichment-induced perturbation of endogenous protein concentrations. Using non-isobaric mass tags, NCTPro combines a reference channel

of enriched low-abundance plasma proteins with a neat plasma channel in a single data-independent acquisition mass spectrometry (DIA-MS) run. Deep proteome coverage in the reference channel is transferred to the neat plasma channel through MS1 feature matching and MS2-assisted elucidation. In this way, NCTPro identifies more than 3000 proteins in neat plasma using a 23-min elution gradient – at least a three-fold increase in coverage over recent reports at similar throughput. Spike-in experiments further demonstrate that NCTPro provides accurate fold-change quantification, high precision, and high linearity, making it well suited to a wide range of clinical plasma proteomics studies.

Keywords

Proteomics, Neat plasma, Mass-spectrometry, Nanomaterials, Enrichment

1. Introduction

Blood is a critical source for screening disease biomarkers and investigating long-term effects of environmental conditions and aging on the human body. It is valued primarily because it is readily accessible and uniquely reflects physiological changes[1]. Because proteins are central to most biological processes, blood proteomics has become a focus of medical diagnostics and research[2-5]. Recent advances in multi-platform technologies, including affinity- and MS-based methods have significantly improved the proteome coverage of circulating proteins[6-8]. While affinity binder technologies like SomaScan and Olink provide superior coverage and throughput relative to MS-based proteomics, they are less suited to discovery-driven studies and cannot detect protein isoforms or post-translational modifications[9]. Moreover, their high cost limits widespread use in routine clinical research.

In contrast, MS-based methods profile the proteome in an untargeted manner, enabling unbiased discovery of proteins whose abundances change with physiological or pathological states, at relatively low cost[9]. While human blood theoretically contains more than 10000 distinct proteins, 22 high-abundance proteins account for nearly 99% of the total protein mass in plasma/serum[10-12]. Because these proteins dominate scan resources and suppress the ionization of low-abundance species during liquid chromatography–mass spectrometry (LC-MS) analysis[7], they severely limit the detection of low-abundance proteins. To address this, several strategies have been developed–including chromatographic fractionation, depletion of high-abundance proteins, and protein corona-based enrichment of medium- and low-abundance proteins [6,13-16]. However, these approaches introduce new complications. Specifically, depletion strategies using either antibodies affinity or perchloric acid (PerCA) precipitation remove only a subset of the high-abundance proteins, yielding a modest gain in protein identifications relative to untreated samples[15,17,18]. More importantly, loss of low-abundance proteins during depletion is inevitable, owing to the limited selectivity of these strategies and to interactions between high- and low-abundance proteins[15,19,20]. Recent advances in protein corona-based enrichment have markedly improved identification of medium- and low-abundance proteins[6]. However, protein corona formation on nanoparticle surfaces is influenced by multiple plasma factors, which may be highly complex in large clinical cohorts[21,22]. Moreover, protein quantification relies on measuring their counterparts in the protein coronas, assuming

nanoparticle-protein adsorption coefficients remain constant across concentrations and are unaffected by variations of other biomolecules in the blood. However, validity of this assumption remains unverified and deviation from theoretical abundance have been observed in recent literature -where the magnitude of deviation increased with greater fold-changes[8,23,24].

Recent advances have introduced multiplexed data-independent acquisition (DIA) technologies that use non-isobaric isotopic labels to enable single-cell proteomics by transferring identifications from bulk to single-cell channels in the same LC-MS run[25,26]. Inspired by this approach, we developed NCTPro (Nanomaterial-enhanced Cross-Channel Transfer strategy for neat blood Proteomics), a workflow that combines low-abundance protein enrichment with neat plasma analysis via multiplexed DIA to achieve ultra-deep proteome coverage and non-disturbed quantification simultaneously (Figure 1). In this workflow, one enrichment channel and two neat plasma channels, each labeled with a different dimethyl reagent, are analyzed in the same DIA-MS run. Deep proteome coverage is obtained in the enrichment channel through spectral library-based database searching, and peptide identifications are then transferred to each neat plasma channel by MS1 feature matching and MS2-assisted elucidation. In this way, both sensitive protein identification and accurate quantification can be achieved in the neat plasma channels without enrichment or depletion. Compared with conventional label-free DIA (LF-DIA) strategies, NCTPro achieved an approximately six-fold increase in protein coverage from neat plasma using an Orbitrap Exploris 480 mass spectrometer. Spike-in experiments confirmed accurate fold-change quantification under stringent false discovery rate (FDR) control in all channels. On the Orbitrap Astral platform, NCTPro enables a mapping depth of more than 3000 proteins at a throughput of 60 samples per day (SPD) from neat plasma, representing at least a three-fold enhancement in proteome coverage at comparable throughput relative to recent reports[27]. Additionally, NCTPro achieves an average KEGG pathway coverage of more than 30%, comparable to that of the Olink Explore HT platform, demonstrating its potential for disease biomarker discovery and mechanistic studies.

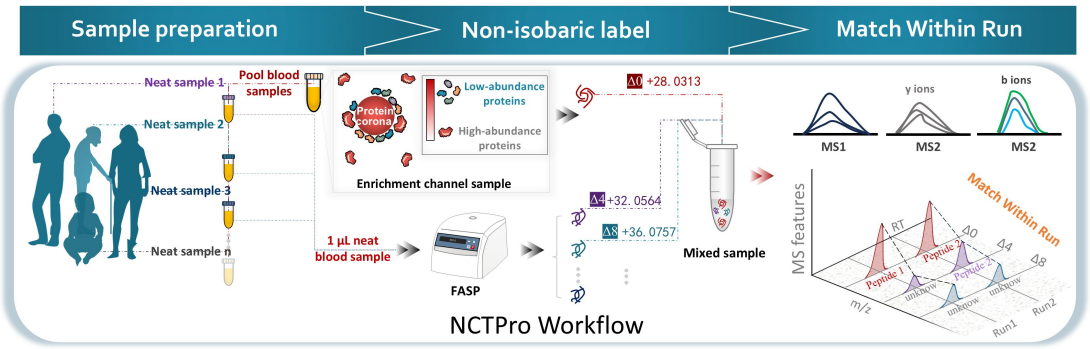

Figure 1. Workflow of NCTPro. Peptides from the enriched sample (Δ0, enrichment channel) and neat plasma (Δ4/Δ8, neat plasma channels) are labeled using a 3-plex dimethyl labeling strategy. In the enrichment channel, alleviation of signal suppression from high-abundance proteins and enhanced signal intensity of low-abundance peptides facilitates robust peptide identification. This provides crucial information–retention time (RT), precursor ion features, and fragment ion patterns – for cross-channel matching with the neat plasma data, enabling peptide sequence propagation. By employing this "match-within-run" approach–analogous to match-between-runs (MBR) but performs within the same run – peptide identification is transferred between the enrichment and neat plasma channels, whereas quantification across channels remains independent.

2. Methods

2.1 Reference plasma sample collection and ethics statement

Blood samples were collected from ten healthy volunteers. The Institutional Ethics Committee of the Beijing Proteome Research Center approved the study protocol. All procedures involving human participants complied with the Declaration of Helsinki, applicable Chinese laws and regulations, and institutional guidelines. All participants provided written informed consent before blood collection. Plasma was prepared by collecting blood into EDTA tubes, followed by centrifugation at 3000 g for 15 minutes, subsequently pooled to generate a reference plasma sample. To minimize freeze-thaw cycles, the plasma was aliquoted and stored at -80 °C until further use.

2.2 Sample preparation

2.2.1 Low-abundance protein enrichment and digestion

Reference plasma was treated with the surface functionalized magnetic MoS_2 nanomaterials, as previously described[28], with modifications. Initially, 100 μL of Reference plasma was mixed with MoS_2 resuspension at a concentration of 1 μg/μL

(prepared in wash buffer: 10 mM Tris-HCl, pH 7.4, 150mM KCl, 0.05% CHAPS). The mixture was incubated at 37 °C on a shaker (1500 rpm) for 1 hour. After incubation, the MoS_2 material was magnetically separated, followed by washing three times with 400 μL wash buffer to enrich low-abundance plasma proteins. Next, 20 μL of lysis buffer (50mM TEAB, 0.5% SDC, 10mM TCEP, 40mM CAA) was added to elute the low-abundance proteins from the magnetic MoS_2 material. The collected low-abundance protein supernatant was heated at 90 °C for 10 minutes, cooled to room temperature, then 10 μL of 50mM TEAB and 0.5 μg of trypsin were added for digestion at 37 °C for 4 hours. An additional 0.5 μg of trypsin was added, and digestion was continued for 8 hours. The digestion was quenched by adding Trifluoroacetic acid (TFA) to a final concentration of 1%, followed by centrifugation at 16000 g for 10 minutes to retain the supernatant. The obtained supernatant was desalted using a C18 SPE cartridge, and the low-abundance peptides were dried at 45 °C, then stored at -80 °C until further use.

2.2.2 Sample preparation of neat plasma and depleted plasma

Neat plasma was digested using Filter-Aided Sample Preparation (FASP) method[29]. One microliter of neat plasma was mixed with 200 μL of 8M urea (UA) and transferred to a 30 kDa ultrafiltration tube. The mixture was centrifuged at 12000g for 10 minutes to remove UA. Next, 8M UA containing 10mM TCEP and 40mM CAA was added, mixed thoroughly, and left at room temperature for 1 hour. The solution was centrifuged at 12000g, and the filter was washed three times with 200 μL of 8M UA, followed by three washes with 200 μL of 50mM TEAB. Finally, 200 μL of 100mM TEAB containing 1 μg of trypsin was added for digestion at 37°C for 4 hours. An additional 1 μg of trypsin was then added, and the digestion was continued for another 8 hours. The peptides were collected by centrifugation at 12000g for 10 minutes, desalted, and stored at -80°C.

The Top14 high-abundant protein depletion method follows the High-Select Top14 Abundant Protein Depletion column protocol (Thermo Fisher Scientific, https://assets.thermofisher.com/TFS-Assets/LSG/manuals). To begin, 10 μL of reference plasma was added to the depletion column and incubated at room temperature for 10 minutes with gentle inversion. Next, the column was centrifuged at 12000g for 10 minutes to collect depleted proteome in the flow-through for enzymatic digestion using the FASP method. The resulting peptides were desalted, dried, and stored at -80°C until further use.

2.2.3 Sample preparation for NCTPro

Dimethyl Labeling: Low-abundance peptides enriched using the magnetic MoS2 nanomaterials, along with digested peptides from undisturbed neat plasma, were labeled using dimethyl labeling[30]. The enriched low-abundance peptides were dissolved in 100 mM TEAB, and 20 μg of these peptides were labeled with 4 μL of formaldehyde (4% CH_2O v/v) and 4 μL of sodium cyanoborohydride (0.6M, $NaBH_3CN$), designated as Δ0 channel. Two portions (20 μg each) of peptides from the neat plasma, also dissolved in 100mM TEAB, were labeled separately. One portion was labeled with 4 μL of 4% formaldehyde-d2 solution (CD_2O) and 4 μL of sodium cyanoborohydride (0.6M), designated as Δ4 channel. The other portion was labeled with 4 μL of 4% formaldehyde-13C-d2 solution ($^{13}CD_2O$) and 4 μL of 0.6M sodium cyanoborodeuteride ($NaBD_3CN$), designated as Δ8 channel. The three channels were shaken at 20°C at 1000 rpm for 1 hour, followed by the addition of 2 μL of 10% $NH_3 \cdot H_2O$, vortexed, and left to stand for 2 minutes. The reaction was stopped by adding 10 μL of 20% TFA. The peptides were desalted, dried, and stored at -80°C until further use.

2.2.4 Spike-in experiment

Count-based FDR estimation of NCTPro: The low-abundance peptides were labeled as Δ0, Δ4, and Δ8 following the dimethyl labeling method. These labeled peptides were then mixed in different combinations to create various sample mixtures: Δ0:Δ4:Δ8 = 2:0:0, Δ0:Δ4:Δ8 = 2:1:0, Δ0:Δ4:Δ8 = 2:0:1, and Δ0:Δ4:Δ8 = 2:1:1. The resulting mixtures were stored at -80°C until needed.

Quantitative accuracy benchmarks of NCTPro: Fetal bovine serum (FBS) was digested using the FASP method and were labeled as Δ4 and Δ8 channels. The Δ4-labeled human neat plasma peptides were mixed with the Δ4-labeled FBS peptides in ratios of 1:0, 1:1, 2:3, 1:2, 1:9, 1:99, and 0:1. Similarly, the Δ8-labeled human neat plasma peptides were mixed with the Δ8-labeled FBS peptides in the same ratios. These multi-species dimethyl-labeled peptides were then combined with Δ0-labeled low-abundant human plasma peptides (enriched using magnetic MoS_2 nanomaterials) in ratios of Δ0:Δ4:Δ8 = 2:1:1, Δ0:Δ4:Δ8 = 2:1:0, and Δ0:Δ4:Δ8 = 2:0:1. The final mixtures were stored at -80°C.

Fold-change accuracy of the nanomaterials-enrichment method and LF-DIA of neat plasma: label-free FBS peptides were mixed with the label-free neat plasma peptides

in ratios of 1:0, 1:1, 2:3, 1:2, 1:9, 1:99, and 0:1. The final mixtures were stored at -80°C until further use.

2.3 Data acquisition

2.3.1 Orbitrap Exploris 480

LC conditions: The experiment was performed on an Easy nLC-1200 system. Peptides were resuspended in buffer A (0.1% FA in LC-MS-grade water) and injected onto a C18 reversed-phase analytical column (75 μm × 20 cm, 1.9 μm, 120 Å) at 60 °C. Two LC gradients were applied on Orbitrap Exploris 480. 1) 75-minuates gradient: 5-12% buffer B (80% ACN and 0.1% FA mixed with LC-MS-grade water) from 0 to 4 minutes, 12-30% buffer B from 4 to 49 minutes, 30-45% buffer B from 49 to 65 minutes, and 45-95% buffer B from 65 to 75 minutes. 2) 120-minuates gradient: 7-12% buffer B from 0 to 6 minutes, 12-30% buffer B from 6 to 102 minutes, 30-45% buffer B from 102 to 114 minutes, and 45-95% buffer B from 114 to 120 minutes. The flow rate was maintained at 300 nL/min throughout the run.

MS/MS conditions: 1) Data-dependent acquisition (DDA) was performed to assess dimethyl labeling efficiency, with MS1 full scans from 300 to 1500 m/z at 60000 resolution and HCD MS2 fragmentation at 27% normalized collision energy. 2) DIA was used to analyze the NCTPro, neat plasma, and depleted plasma proteome obtained by the Top14 high-abundant proteins depletion column. For MS1 analysis, the full scan range was set to 300-1650 m/z with a resolution of 120000 in the Orbitrap. The AGC target was set to 300%. For MS2 analysis, fragment-ion spectra were acquired in the Orbitrap at a resolution of 30000 using variable-width isolation windows. A total of 44 and 64 windows were used for the 75- and 120-min gradients, respectively, as described in previous studies[31]

2.3.2 Orbitrap Astral

LC conditions: Approximately 200ng peptide was loaded onto an analytical column (Aurora Ultimate 25 cm × 75 μm ID, 1.7 μm, C18) using 23-min gradient. The gradient began at 4% buffer B, increasing to 12% over 2 min, then gradually rising to 25% over the next 7 min. Buffer B continued to increase to 40% over the subsequent 9.5 min before quickly ramping up to 99% within 0.4 min. Finally, the column was held at 99% buffer B for an additional 4 min.

MS/MS conditions: DIA was performed on an Orbitrap Astral mass spectrometer (Thermo Fisher Scientific, Germany) operated without FAIMS. Full MS1 scans were

acquired in the Orbitrap at a resolution of 240000 (FWHM), spanning an m/z range of 380–980. The RF lens was set to 40%, and the AGC target was normalized to 500%, with a custom-defined maximum injection time. DIA MS2 scans were acquired across a precursor m/z range of 398–980 using different isolation windows of 2, 4, and 8 Th. Fragmentation was achieved using HCD with a normalized collision energy of 25%. MS2 spectra were acquired with a normalized AGC target of 800%, also employing a custom injection time. Fragment ions were detected over an m/z range of 150–2000. All MS1 data were acquired in profile mode, while MS2 data were recorded in centroid mode.

2.4 DIA analysis

All raw files were converted to mzML format using MSConvert (ProteoWizard release: 3.0.23163 [09bc765]).

Neat plasma and Top14-depleted plasma: The in silico predicted spectral library for LF-DIA analysis (neat plasma and Top14-depleted plasma) was generated by DIA-NN (version 1.8.1) using deep learning to predict spectra, retention times, and ion mobility, based on the same FASTA file mentioned earlier. For DIA-NN searches, “C carbamidomethylation” was set as a fixed modification, while “Ox (M)” and “Ac (N-term)” were set as variable modifications. “Trypsin/P” was used as the protease, allowing up to 1 missed cleavage. The precursor FDR threshold was set at 1%, and “MBR”was enabled. All other settings were kept as DIA-NN defaults.

NCTPro: NCTPro raw data were analyzed using a two-step workflow. An *in silico* spectral library was generated with AlphaPeptDeep, using the predefined dimethyl labeling models from MannLabs[25,32]. The raw data were then searched against this library with DIA-NN (v1.8.1) in plexDIA mode.

Data processing and quality filtering: Data analysis was conducted using R (version 4.3.0) and Python (version 3.9.18). For neat plasma and Top14-depleted plasma samples, search results by DIA-NN were filtered with thresholds of Q.Value < 0.01, Lib.Q.Value < 0.01, Lib.PG.Q.Value < 0.01, and PG.Q.Value < 0.01. For NCTPro, the filtering criteria included Lib.PG.Q.Value < 0.01, Q.Value < 0.01, and Channel.Q.Value < 0.15 to ensure a protein-level FDR below 1% for reliable protein identification.

3. Results

3.1 The principle of NCTPro

In proteomics analysis, low-abundance proteins/peptides often produce low-quality MS2 spectra that cannot be annotated, owing to the limited number of fragment ions and their poor intensity. To promote spectral elucidation, MS1 feature matching and MS2 identification transfer across LC-MS runs – collectively known as "match between runs" (MBR)–have been widely adopted[33,34]. This approach relies heavily on accurate alignment of RT between runs. However, variability in chromatographic performance across runs complicates precise peptide matching by RT and increases the risk of false-positive identifications[35-37]. To address these limitations, within-run identification transfer strategies have recently been demonstrated in single-cell proteomics using non-isobaric multiplexing techniques such as multiplex-DIA and plex-DIA[25,26]. Because labeled channels are compared within a shared chromatographic context, this design eliminates the need for cross-run RT alignment and markedly improves the accuracy and sensitivity of identification transfer, particularly for low-abundance proteins.

We therefore propose the NCTPro strategy, which leverages nanomaterials-enriched plasma and its in-depth proteome coverage to facilitate low-abundance protein identification in neat plasma through multiplexed DIA technologies with MS1 and MS2 feature matching. NCTPro employs dimethyl labeling to create three channels with a 4 Da mass shift between each. A reference channel containing peptides from the nanomaterials-enriched plasma is mixed with two target channels containing neat plasma peptides, allowing the three channels to be analyzed jointly. The enriched plasma is prepared using our previously developed surface-functionalized magnetic MoS_2-based corona enrichment method, which detects more than 3000 plasma proteins–an approximately four- to five-fold enhancement over direct analysis of neat plasma[28].

In NCTPro, the reference channel (e.g., Δ0), which provides rich high-quality MS2-level identifications of low-abundance proteins, serves as an effective reference database for the target channels. Leveraging the high RT consistency of co-eluting, co-fragmented isotopic-labeled peptides within the same LC-MS run, annotation of low-quality MS2 spectra in the target channels benefits significantly from the reference channel via MS1 feature alignment, MS2 fragment matching within the

same spectrum, and stringent FDR filtering (Figure 1). In this way, accurate proteome quantification with deep coverage can be achieved in neat plasma without enrichment-induced perturbation of original protein concentrations.

3.2 The feasibility of NCTPro

As a first step toward the NCTPro workflow, we assessed the dimethyl labeling efficiency of plasma peptides and its impact on plasma protein analysis. Figure S1A (Supporting Information) shows that dimethyl labeling efficiency exceeded 99% for both MoS_2-enriched low-abundance plasma peptides (labeled with Δ0) and neat plasma peptides (labeled with Δ4 and Δ8), indicating near-complete labeling. Compared with 1-plex and 3-plex DIA, proteome coverage of the enriched plasma decreases by 8.6% and 11.4%, respectively (Figure S1B, Supporting Information). Despite this reduction, more than 3000 proteins are still identified in the enriched plasma—demonstrating its capability to serve as a reference database—whereas direct analysis of neat plasma yields only 459 proteins, owing to signal suppression by high-abundance proteins. As shown in Figure S1C (Supporting Information), although increasing the number of dimethyl labels leads to a slight decline in the reproducibility of protein identification between channels, the consistency remains high. This modest reduction may be attributed to the substantially higher MS/MS spectral complexity of multiplex DIA compared with label-free DIA. Moreover, a coefficient of variation (CV) below 5% is achieved for the dimethyl-labeled samples using either 1-plex or 3-plex tags, comparable to LF-DIA (Figure S1D, Supporting Information), indicating that quantification precision is not impaired by dimethyl labeling. This stability of MoS_2 enrichment ensures a reliable reference database for the NCTPro workflow.

3.3 NCTPro exhibits ultra-deep protein coverage in neat plasma

We compared plasma protein identification by NCTPro on an Orbitrap Exploris 480 under 45-, 75-, and 120-min gradients. The three gradients yield approximately 1300, 1800, and 2100 proteins, respectively, at a protein- and precursor-level FDR of 1% in neat plasma (Figure 2A). In contrast, LF-DIA of the same neat plasma samples

identified only 459 proteins on average under the 75-min gradient (Figure 2A). In terms of reproducibility, NCTPro shows strong consistency across replicates, with over 85% of proteins identified in all three runs for each elution gradient. This level of overlap is comparable to the results obtained using neat or enriched plasma samples (Figure 2B). The 3-channel design (Δ0, Δ4, Δ8) offers higher throughput but produces more complex MS/MS spectra, complicating spectral deconvolution. To address this, we adopted a simplified 2-plex setup. As shown in Figure 2C, the MS/MS spectral identification rate–defined as the number of identified MS/MS spectra divided by the total number of MS/MS spectra, at a precursor-level FDR of 1%–increased from 35.2%–39.0% in the 3-plex format to 57.7%–66.8% in the 2-plex format. This improvement translates into higher proteome coverage: under a 200-ng loading and a 75-min gradient, the 2-plex design identified 2469 proteins (Δ0/Δ4) and 2279 proteins (Δ0/Δ8), corresponding to 36.5% and 27.2% more proteins than the 3-plex format (1809 and 1791 proteins), respectively (Figure 2D). Proteome coverage is further improved with the increased sample loading: 600 ng of peptides yielded the highest number of protein identifications, and increasing the loading to 800 ng provided no further benefit (Figure 2E). Extending the elution gradient to 120 min further increases protein identification for the 2-plex design. Under a 600-ng loading and a 120-min gradient, NCTPro (Δ0/Δ4] and NCTPro (Δ0/Δ8] identified 3207 and 2809 proteins in neat plasma–48.6% and 41.2% more than the 2158 and 1990 proteins identified using the 3-plex format (Figure 2F). After matching the identified proteins to the absolute protein concentrations in the Human Protein Atlas (HPA) database[38,39], NCTPro identified 910–1146 low-abundance proteins in the 1 ng/L–1 μg/L low concentration range (Figure 2G). In contrast, LF-DIA of the same neat plasma samples identified 459 proteins overall but only five proteins in this range, highlighting NCTPro's sensitivity for low-abundance proteins without altering their endogenous concentrations.

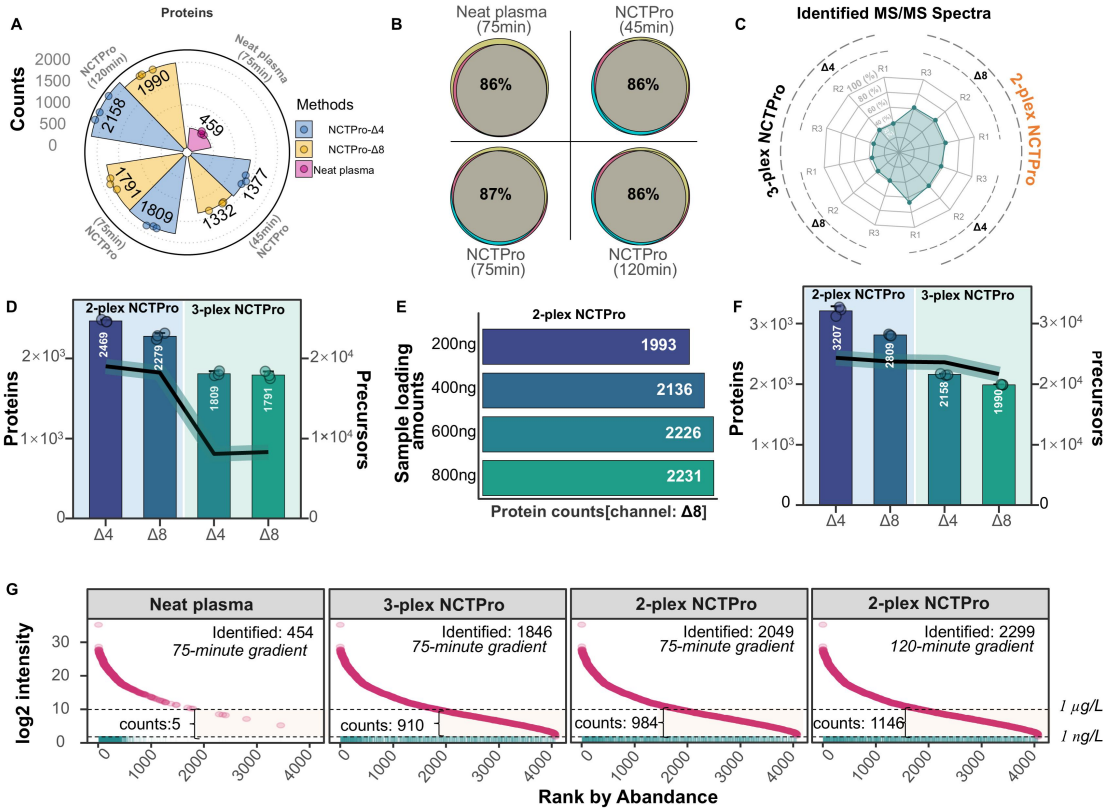

Figure 2. NCTPro expands proteomic coverage in neat plasma. (A) Number of proteins identified by NCTPro and LF-DIA. (B) Venn diagrams showing the reproducibility of protein identification by NCTPro and LF-DIA. (C) MS/MS spectral identification rates for 2-plex and 3-plex DIA formats. (D) Numbers of proteins and precursors identified by NCTPro under a 75-min gradient in 2-plex and 3-plex formats. (E) Number of proteins identified with increasing sample loading. (F) Numbers of proteins and precursors identified by NCTPro under a 120-min gradient in 2-plex and 3-plex formats (600 ng peptide loading). (G) Absolute concentrations of proteins identified from neat plasma by LF-DIA and NCTPro. All identifications were filtered at a protein- and precursor-level FDR of 1%.

3.4 Potential clinical implications of the NCTPro identified plasma proteome

We next assessed the clinical relevance of the NCTPro plasma proteome by evaluating its coverage of FDA-approved protein biomarkers. NCTPro identified 123 FDA-approved protein biomarkers[40] from neat plasma (Figure 3A), exceeding the 79 detected by direct LF-DIA of neat plasma. For comparison, we examined two Olink Explore HT protein sets: (i) the complete assay target panel, representing its maximum nominal coverage, and (ii) the subset of panel targets detected in plasma in an independent study[41]. The complete panel, designated “Olink Theoretical” in Figure

3A, comprises 5416 protein targets, including 125 FDA-approved biomarkers. Beimers *et al.*[41] detected 2631 of these targets in plasma; this plasma-detected subset, designated "Olink Measured," included 105 FDA-approved biomarkers. We next examined the overlap among the biomarker sets identified by the four approaches (Figure S2A). The 123 FDA-approved biomarkers identified by NCTPro from neat plasma completely encompass the 79 biomarkers detected by direct LF-DIA of the same samples. In contrast, NCTPro and the Olink platform share only 65 of the 105 biomarkers measured by Olink, with 58 and 40 biomarkers unique to NCTPro and Olink, respectively. At the proteome-wide level, NCTPro shared only 1224 of its 3232 proteins with the 5416 proteins theoretically covered by the Olink platform – approximately 16% of the combined set (Figure S2B); 2008 and 4192 proteins were unique to NCTPro and Olink, respectively. These results highlight the marked complementarity between the LC-MS-based NCTPro workflow and the antibody-based Olink platform. To assess NCTPro's ability to capture tissue-associated biological signals, we annotated the identified proteins using the HPA tissue-enhanced classification. Tissue-enhanced proteins are defined as proteins encoded by genes showing at least fourfold higher expression in one or more tissues than the average expression across all other tissues[42]. After merging and deduplicating the male and female HPA annotations, NCTPro identifies 1152 unique tissue-enhanced proteins, compared with 996 among the proteins measured by the Olink Explore HT platform[41] (Figure 3B). In addition, NCTPro identified 824 additional tissue-enhanced proteins that were absent from the published Olink plasma-detected set (Olink Measured; Figure S3A, Supporting Information)[41], including 706 proteins not represented in the complete Olink Explore HT target panel (Olink Theoretical; Figure S3B, Supporting Information). The top 20 tissues represented by NCTPro spanned diverse physiological systems, including the immune and hematopoietic, nervous, digestive, musculoskeletal, cardiovascular, endocrine, urinary, and reproductive systems (Figure 3B). Seventeen of the top 20 tissues were shared between NCTPro and measured Olink, including highly represented tissues such as the lymph node, brain, liver, bone marrow, and skeletal muscle. This result demonstrates that NCTPro provides broad tissue representation comparable to that of the measured Olink data[41], while detecting 15.7% more tissue-enhanced proteins overall (1152 versus 996). To assess NCTPro's pathway mapping capability, we evaluated KEGG pathway coverage using the Rich

Factor (RF) metric. As shown in Figure 3C, the median RF for KEGG pathways covered by proteins identified using NCTPro was 0.333–significantly higher than the 0.047 median RF achieved by direct LF-DIA of neat plasma. Notably, NCTPro exhibits analytical performance comparable to that of the Olink Explore HT platform under its theoretical configuration. We calculated an RF of 0.334 for the 5416 proteins nominally covered by the Olink panel, whereas the median RF for NCTPro was 0.333. However, this value was derived from idealized coverage assumptions; when restricted to the 2631 proteins quantified[41], Olink's median RF decreased to 0.193. These findings highlight NCTPro's ability to provide rich biological information.

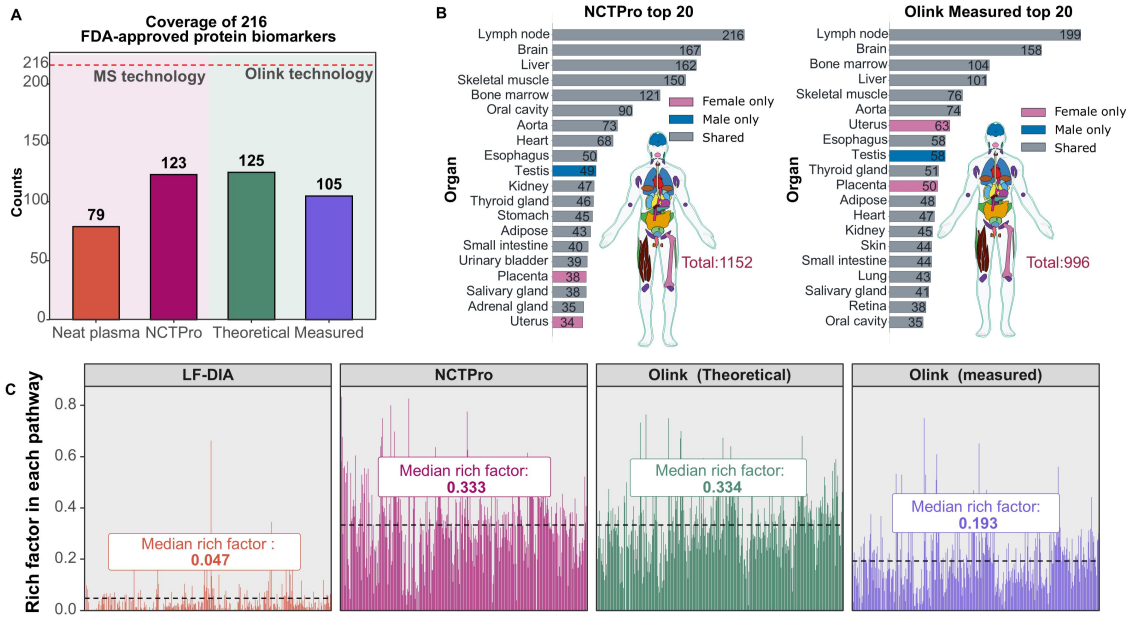

Figure 3. NCTPro captures biological information enriched plasma proteome. (A) Numbers of the 216 FDA-approved protein biomarkers covered by neat-plasma LF-DIA, NCTPro, the full Olink Explore HT target panel, and an experimentally measured Olink plasma dataset. “Olink (Theoretical)” denotes the complete set of 5416 protein targets included in the Olink Explore HT assay, whereas “Olink (measured)” denotes the 2631 proteins detected in plasma using Olink Explore HT in the study by Beimers *et al.*[41] (B) Top 20 organs ranked by the number of HPA tissue-enhanced proteins detected by NCTPro and Olink (measured). Male and female HPA organ tables were merged and deduplicated. Colors indicate organs exclusive to the female table, exclusive to the male table, or shared between both tables. Numbers on the bars indicate protein counts for individual organs, whereas the values beside the anatomograms indicate the total numbers of unique tissue-enhanced proteins detected by each method. (C) KEGG pathway coverage of different methods.

3.5 Quantification accuracy of NCTPro in neat plasma

To assess the quantification accuracy of NCTPro, we conducted a spike-in experiment using mixtures of bovine and human plasma at varying ratios to simulate inter-individual protein concentration difference in the target channel, with nanomaterial-enriched human plasma serving as the reference channel (Figure 4A). Figure 4B showed NCTPro maintains high quantification precision, with CV consistently below 15%. Additionally, NCTPro exhibited high quantification consistency, achieving correlation coefficients greater than 0.95 in three replicates (Figure S4, Supporting Information). More importantly, NCTPro demonstrates quantification accuracy remarkably similar to that of the perturbation-free LF-DIA, with both closely approaching theoretical fold change (FC) values (Figure 4C). This holds true across the entire tested ratio range, though both showed moderate deviation at extremely high ratios (e.g., 100:1). In contrast to the accuracy achieved by NCTPro, the nanomaterials-enrichment method exhibits consistent deviations from theoretical expectations. Protein-level quantitative accuracy is more clearly displayed in Figure S5 (Supporting Information) by comparing the observed and expected log_2FC for proteins quantified at five or more dilution points. Among the evaluable proteins, 72.17% quantified by NCTPro meet the predefined fold-change accuracy criterion, approaching the proportion achieved by LF-DIA 81.07% and exceeding the 39.30% obtained by nanomaterials-enrichment method . Although the accurately qualified proportion of NCTPro is slightly lower than that of LF-DIA, NCTPro yields 2.56 times more proteins in absolute terms within the expected FC. These results indicate that NCTPro maintains quantitative accuracy close to that of LF-DIA while providing protein coverage approaching that of the enrichment-based method. Collectively, these results demonstrate the advantage of using non-disturbed NCTPro to achieve both deep proteome coverage and high quantification accuracy from neat plasma.

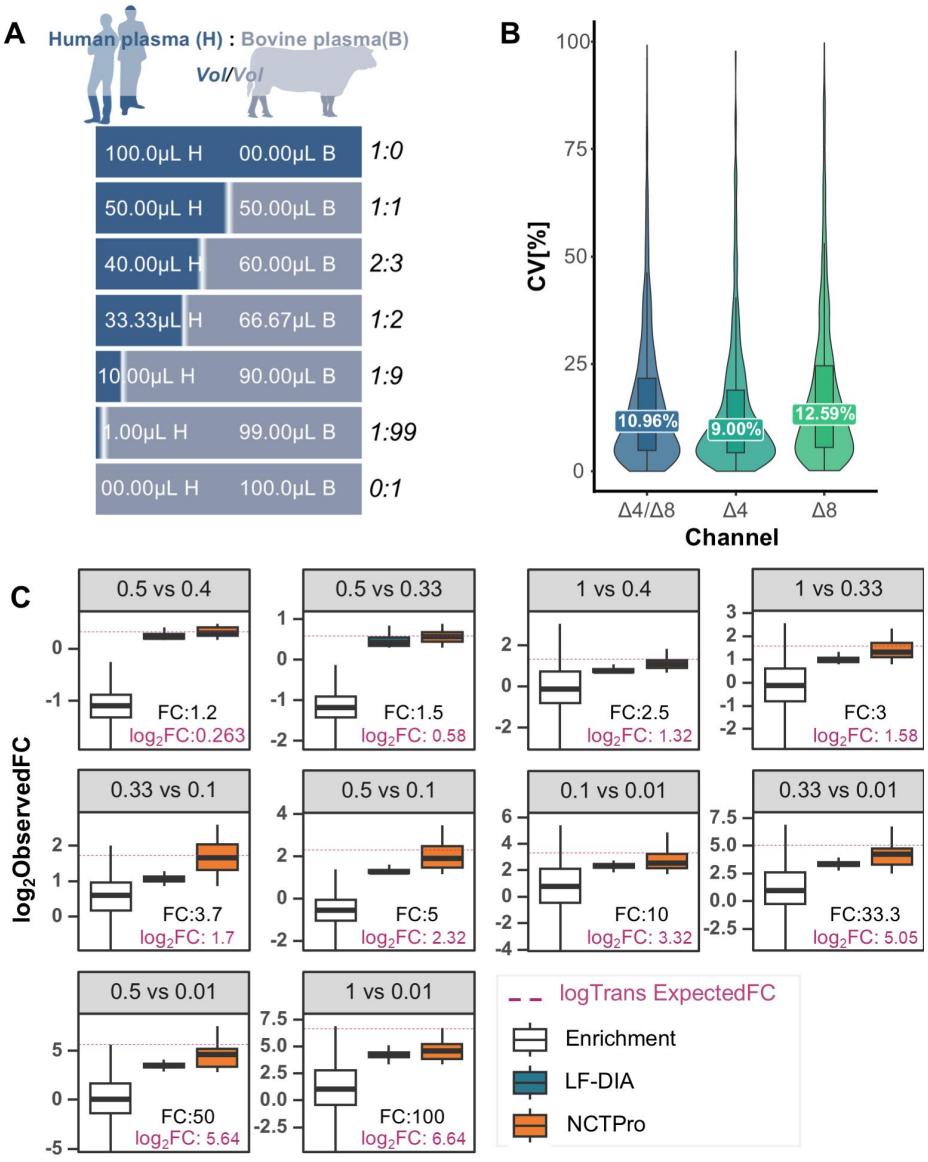

Figure 4. Evaluation of quantification accuracy in spike-in experiments using NCTPro. (A) Experimental design in which human plasma proteins are spiked into a bovine plasma background at defined ratios (human (vol): bovine (vol). (B) CV of protein quantification by NCTPro from technical replicates. The x-axis labeled with Δ4/Δ8 indicates CV calculated across all channels from technical replicates, Δ4 or Δ8 indicates CV calculated between the Δ4 or Δ8 channels in different runs, respectively. (C) Experimental FC distributions versus theoretical expectation. Dashed lines denote the expected log_2FC. The corresponding expected FC and log_2FC are shown in each panel.

3.6 Extending NCTPro to the Orbitrap Astral platform

The emergence of new MS instrumentation has expanded the depth of proteome coverage achievable from plasma, prompting us to evaluate NCTPro on an Orbitrap Astral mass spectrometer. Using a 2-plex NCTPro configuration with narrow-window

data-independent acquisition (nDIA), an 8 Da isolation window, and a 23-min elution gradient results in 1600-1800 plasma proteins (Figure 6A). Narrowing the isolation window to 4 Da or 2 Da markedly increased proteome coverage, yielding 2300-2900 proteins at a throughput of 62 SPD. In comparison, direct LF-DIA of neat plasma with nDIA acquisition identified only 570 proteins under otherwise identical instrument conditions (Figure 6B). Consistent with this coverage advantage, NCTPro identified approximately five times more proteins than LF-DIA under equivalent conditions (Figure 6C). Moreover, shortening the LC gradient time to 14 min still enabled detection of approximately 2000 proteins per run, increasing throughput to 102 SPD (Figure 6C). We next assessed NCTPro's quantitative performance on the Orbitrap Astral platform. The median CV for protein quantification remained below 10% at both 62 and 102 SPD, demonstrating that the high precision of NCTPro is not impaired by the shorter LC gradient or fast MS acquisition (Figure 6D). Pearson correlation analyses further confirmed high reproducibility across replicates under both gradient conditions, with correlation coefficients exceeding 0.89 (Figure 6E). Together, these results establish NCTPro as a robust, high-throughput plasma proteomics solution compatible with the Orbitrap Astral platform, offering both depth and precision without compromising efficiency.

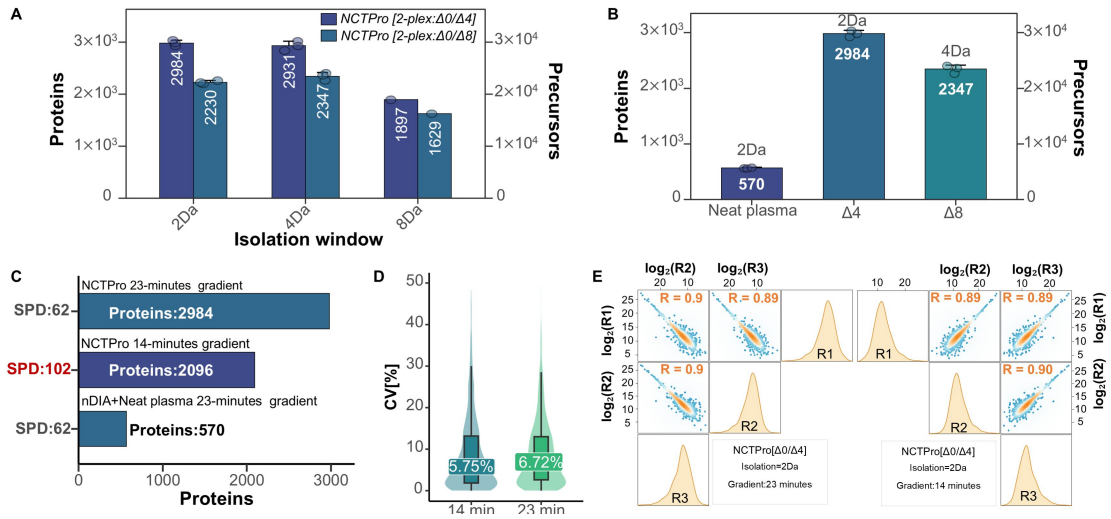

Figure 6. Performance evaluation of NCTPro on the Orbitrap Astral. (A) Number of identified proteins and precursors using different isolation window for NCTPro. (B) Number of identified proteins and precursors using optimized isolation window setting for different methods. (C) SPD and identified proteins achieved by NCTPro and LF-DIA using different LC-gradients. Quantification CV (D) and correlation (E) of NCTPro obtained using different LC-gradient.

4. Discussion

In this study, we introduced NCTPro to increase proteome coverage and analytical throughput in neat plasma analysis. NCTPro uses nanomaterial-based enrichment of low-abundance plasma proteins to generate a high-intensity reference channel, while neat plasma samples are analyzed in parallel target channels. Dimethyl isotopic labeling combined with multiplexed DIA generates isotopically distinct forms of the same peptides with predictable mass offsets and closely aligned elution profiles. These properties enable high-confidence, reference-assisted transfer of peptide identifications from the enriched reference channel to the neat plasma channels within the same DIA run, supported by corresponding MS1 precursor and MS2 fragment-ion evidence. Under otherwise identical instrument conditions, NCTPro identified 5 times as many plasma proteins as LF-DIA. Because quantification is based on channel-specific ion intensities extracted directly from the neat plasma channels, the reference channel enhances peptide detection without substituting for target-channel quantitative signals. This combination of depth, throughput, and direct neat plasma proteome quantification is particularly advantageous for clinical biomarker studies.

Compared with the Olink Explore HT target panel, NCTPro provides substantial complementary proteome coverage. Of the 3232 proteins identified by NCTPro, 2008 are not represented among the 5416 protein targets in the Olink Explore HT panel, whereas 1224 are shared between the two protein sets. This complementarity extends to tissue-enhanced proteins: 706 of the 1152 HPA tissue-enhanced proteins identified by NCTPro are absent from the Olink Explore HT target panel, whereas 446 are shared. These substantial differences highlight the necessity of employing plasma proteome methods with distinct detection mechanisms to achieve comprehensive coverage. Notably, proteins enhanced in the Lymph node, brain and liver comprise the largest proportions within NCTPro's unique identifications, suggesting its particular utility for non-invasive mirroring of organ-level proteome changes in these critical systems.

5. Conclusion

NCTPro holds significant potential for advancing clinical and research applications in neat plasma proteomics. Its in-depth proteome coverage, accurate quantification and

cost-effectiveness–provide a practical alternative for large-scale biomarker studies. Compatibility with the edge-cutting Orbitrap Astral platform largely improves its throughput, making it a powerful tool for biomarker discovery and validation.

Reference:

1 Tebani, A. *et al.* Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nature Communications* **11** (2020). https://doi.org/10.1038/s41467-020-18148-7

2 Dayon, L., Cominetti, O. & Affolter, M. Proteomics of human biological fluids for biomarker discoveries: technical advances and recent applications. *Expert Review of Proteomics* **19**, 131–151 (2022). https://doi.org/10.1080/14789450.2022.2070477

3 Melani, R. D. *et al.* The Blood Proteoform Atlas: A reference map of proteoforms in human hematopoietic cells. *Science* **375**, 411–418 (2022). https://doi.org/doi:10.1126/science.aaz5284

4 Eldjarn, G. H. *et al.* Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature* **622**, 348–358 (2023). https://doi.org/10.1038/s41586-023-06563-x

5 Hampel, H. *et al.* Blood-based biomarkers for Alzheimer's disease: Current state and future use in a transformed global healthcare landscape. *Neuron* **111**, 2781–2799 (2023). https://doi.org/10.1016/j.neuron.2023.05.017

6 Blume, J. E. *et al.* Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nature Communications* **11** (2020). https://doi.org/10.1038/s41467-020-17033-7

7 Ferdosi, S. *et al.* Engineered nanoparticles enable deep proteomics studies at scale by leveraging tunable nano–bio interactions. *Proceedings of the National Academy of Sciences* **119**, e2106053119 (2022). https://doi.org/doi:10.1073/pnas.2106053119

8 Huang, T. *et al.* Protein Coronas on Functionalized Nanoparticles Enable Quantitative and Precise Large-Scale Deep Plasma Proteomics. *bioRxiv* (2023). https://doi.org/10.1101/2023.08.28.555225

9 Suhre, K., McCarthy, M. I. & Schwenk, J. M. Genetics meets proteomics: perspectives for large population-based studies. *Nature Reviews Genetics* **22**, 19–37 (2020). https://doi.org/10.1038/s41576-020-0268-2

10 Hortin, G. L. & Sviridov, D. The dynamic range problem in the analysis of the plasma proteome. *Journal of Proteomics* **73**, 629–636 (2010). https://doi.org/10.1016/j.jprot.2009.07.001

11 Anderson, N. L. & Anderson, N. G. The Human Plasma Proteome. *Molecular & Cellular Proteomics* **1**, 845–867 (2002). https://doi.org/10.1074/mcp.R200007-MCP200

12 Tirumalai, R. S. *et al.* Characterization of the Low Molecular Weight Human Serum Proteome. *Molecular & Cellular Proteomics* **2**, 1096–1103 (2003). https://doi.org/10.1074/mcp.M300031-MCP200

13 Bi, X. *et al.* Proteomic and metabolomic profiling of urine uncovers immune responses in patients with COVID-19. *Cell Reports* **38** (2022). https://doi.org/10.1016/j.celrep.2021.110271

14 Lee, P. Y., Osman, J., Low, T. Y. & Jamal, R. Plasma/serum proteomics: depletion strategies for reducing high-abundance proteins for biomarker discovery. *Bioanalysis* **11**, 1799–1812 (2019). https://doi.org/10.4155/bio-2019-0145

15 Viode, A. *et al.* A simple, time- and cost-effective, high-throughput depletion strategy for deep plasma proteomics. *Science Advances* **9**, eadf9717 (2023). https://doi.org/doi:10.1126/sciadv.adf9717

16 Qian, L. *et al.* Proteomic landscape of epithelial ovarian cancer. *Nature Communications* **15**, 6462 (2024). https://doi.org/10.1038/s41467-024-50786-z

17 Bellei, E. *et al.* High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* **40**, 145–156 (2010). https://doi.org/10.1007/s00726-010-0628-x

18 Tu, C. *et al.* Depletion of Abundant Plasma Proteins and Limitations of Plasma Proteomics. *Journal of Proteome Research* **9**, 4982–4991 (2010). https://doi.org/10.1021/pr100646w

19 Bader, J. M., Albrecht, V. & Mann, M. MS-Based Proteomics of Body Fluids: The End of the Beginning. *Molecular & Cellular Proteomics* **22** (2023). https://doi.org/10.1016/j.mcpro.2023.100577

20 Ashkarran, A. A. *et al.* Measurements of heterogeneity in proteomics analysis of the nanoparticle protein corona across core facilities. *Nature Communications* **13** (2022). https://doi.org/10.1038/s41467-022-34438-8

21 Wang, X. & Zhang, W. The Janus of Protein Corona on nanoparticles for tumor targeting, immunotherapy and diagnosis. *Journal of Controlled Release* **345**, 832–850 (2022). https://doi.org/10.1016/j.jconrel.2022.03.056

22 Monopoli, M. P., Åberg, C., Salvati, A. & Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nature Nanotechnology* **7**, 779–786 (2012). https://doi.org/10.1038/nnano.2012.207

23 Korff, K. *et al.* Pre-Analytical Drivers of Bias in Bead-Enriched Plasma Proteomics. *bioRxiv*, 2025.2005.2007.652495 (2025). https://doi.org/10.1101/2025.05.07.652495

24 Gao, H. *et al.* Systematic evaluation of blood contamination in nanoparticle-based plasma proteomics. *EMBO Mol Med* **18**, 275–296 (2026). https://doi.org/10.1038/s44321-025-00346-9

25 Thielert, M. *et al.* Robust dimethyl-based multiplex-DIA doubles single-cell proteome depth via a reference channel. *Molecular Systems Biology* **19** (2023). https://doi.org/10.15252/msb.202211503

26 Derks, J. *et al.* Increasing the throughput of sensitive proteomics by plexDIA. *Nature Biotechnology* **41**, 50–59 (2022). https://doi.org/10.1038/s41587-022-01389-w

27 Fu, Q. *et al.* A Proteomics Pipeline for Generating Clinical Grade Biomarker Candidates from Data-Independent Acquisition Mass Spectrometry (DIA-MS)

Discovery. *Angew Chem Int Ed Engl* **63**, e202409446 (2024). https://doi.org/10.1002/anie.202409446

28 Liu, Y. *et al.* Synthesis of Surface-Functionalized Molybdenum Disulfide Nanomaterials for Efficient Adsorption and Deep Profiling of the Human Plasma Proteome by Data-Independent Acquisition. *Analytical Chemistry* **94**, 14956–14964 (2022). https://doi.org/10.1021/acs.analchem.2c02736

29 Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359–362 (2009). https://doi.org/10.1038/nmeth.1322

30 Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* **4**, 484–494 (2009). https://doi.org/10.1038/nprot.2009.21

31 Steger, M. *et al.* Time-resolved in vivo ubiquitinome profiling by DIA-MS reveals USP7 targets on a proteome-wide scale. *Nat Commun* **12**, 5399 (2021). https://doi.org/10.1038/s41467-021-25454-1

32 Zeng, W.-F. *et al.* AlphaPeptDeep: a modular deep learning framework to predict peptide properties for proteomics. *Nature Communications* **13** (2022). https://doi.org/10.1038/s41467-022-34904-3

33 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513–2526 (2014). https://doi.org/10.1074/mcp.M113.031591

34 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301–2319 (2016). https://doi.org/10.1038/nprot.2016.136

35 Liu, Y. *et al.* DeepRTAlign: toward accurate retention time alignment for large cohort mass spectrometry data analysis. *Nat Commun* **14**, 8188 (2023). https://doi.org/10.1038/s41467-023-43909-5

36 Yu, F., Haynes, S. E. & Nesvizhskii, A. I. IonQuant Enables Accurate and Sensitive Label-Free Quantification With FDR-Controlled Match-Between-Runs. *Mol Cell Proteomics* **20**, 100077 (2021). https://doi.org/10.1016/j.mcpro.2021.100077

37 Kalxdorf, M., Muller, T., Stegle, O. & Krijgsveld, J. IceR improves proteome coverage and data completeness in global and single-cell proteomics. *Nat Commun* **12**, 4787 (2021). https://doi.org/10.1038/s41467-021-25077-6

38 Uhlen, M. *et al.* A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science* **366**, eaax9198 (2019). https://doi.org/doi:10.1126/science.aax9198

39 Atlas, H. P. *Protein detected in human plasma by mass spectrometry*, <https://www.proteinatlas.org/humanproteome/blood/proteins+detected+in+ms> (2025).

40 Bhowmick, P., Roome, S., Borchers, C. H., Goodlett, D. R. & Mohammed, Y. An Update on MRMAssayDB: A Comprehensive Resource for Targeted Proteomics Assays in the Community. *Journal of Proteome Research* **20**, 2105–2115 (2021). https://doi.org/10.1021/acs.jproteome.0c00961

41 Beimers, W. F. *et al.* Technical Evaluation of Plasma Proteomics Technologies. *J Proteome Res* **24**, 3074–3087 (2025). https://doi.org/10.1021/acs.jproteome.5c00221

42 Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015). https://doi.org/doi:10.1126/science.1260419