Proteonano™: A Robust Platform for Large Cohort Deep Plasma Proteomic Studies

Hao Wu*[,1], Yonghao Zhang[1], Yi Wang[1], Xiehua Ouyang[1]

[1]Nanomics Biotechnology Co., Ltd.

*Corresponding author: haowu@nanomics.bio

Keywords: Proteonano Ultraplex Proteomics Platform, Deep Plasma Proteomics, Robust and Fast MS-based Proteomics Workflow

Abstract

Complete profiling of human plasma proteome is an immerse source for disease biomarker discovery. Cutting-edge mass spectrometers, like ThermoFisher's Orbitrap Astral, have promised unprecedented insights into the exploration of multiple protein biomarkers from human plasma samples. However, large-scale, deep profiling of the human plasma proteome, especially low-abundant proteins (LAPs, <10 $ng.mL^{-1}$), in a robust and fast way remains challenging. This is largely due to the lack of standardized and automated workflows including LAPs enrichment, reduction, and enzymatic digestion procedures. Until now, these complex procedures have not been incorporated into a streamlined workflow to achieve reproducibility, high-throughput, and deep proteome coverage.

Here we report the Proteonano Ultraplex Proteomics (PUP) Platform for large cohort plasma proteomics studies with robustness and fast throughput by standardizing workflow by incorporating the Proteonano platform and high-resolution mass spectrometers, including Orbitrap Exploris™ 480, Orbitrap Astral™, and timsTOF Pro. This pipeline demonstrates excellent stability and reproducibility, with tunable balance between proteome coverage and throughput. We further demonstrate the utility of this platform for potential biomarker discovery in a neurodegenerative disease cohort. This harmonized method enables robust, fast and large-cohort plasma proteomics studies to meet the need to discover new biomarkers.

Introduction

The development of untargeted bottom-up mass spectrometry-based proteomics has evolved over the past two decades. Proteomic studies have played a pivotal role in identifying disease biomarkers, which are now integral to precision medicine, contributing to disease detection, patient stratification, therapeutic treatment monitoring, and predictive analysis[1–9]. Blood plasma, a rich repository of novel protein biomarkers, presents unique challenges due to the wide dynamic range (>10 orders of magnitude) in protein concentrations. High-abundance plasma proteins often overshadow the detection of medium- and low-abundance proteins during mass spectrometry-based proteomic analyses[10–12].

Therefore, prior to mass spectrometry, enrichment and fractionation of low-abundance proteins from plasma samples are commonly employed to mitigate the interference caused by high-abundance proteins[10,11,13]. One widely used approach involves the immunodepletion of the most abundant plasma proteins using immunoassay columns, such as the High-Select Top 14 mini columns. Immunodepletion techniques typically enhance protein detection depth, achieving a median 4-fold increase in identified protein groups. However, studies have shown that depleted plasma primarily contains medium- to high-abundance proteins, with only 5–6% identified as low-abundance[12]. Additionally, immunodepletion poses challenges such as non-specific binding leading to unintended loss of low-abundance proteins, complex sample handling, high costs, and batch effects, limiting its scalability for large-cohort plasma proteomics studies[12].

Another substantial barrier to the widespread adoption of mass spectrometry (MS)-based proteomics in large cohort studies is the reproducibility of distinct mass spectrometry configurations across laboratories. Ruedi et al.[14] assessed the quantitative reproducibility, dynamic range, and sensitivity of 11 mass

58 spectrometers globally. Their study quantified over 4,000 proteins from HEK293 cells, achieving a median
59 coefficient of variation (CV) of 57.6% before normalization and 22.0% after normalization. For plasma
60 proteomics, sample-related biases also greatly impact data reproducibility and subsequent biomarker
61 discovery. To address this issue, Geyer et al.[15] developed three quality marker panels to evaluate whether
62 plasma samples were contaminated by coagulation, platelets, or erythrocytes.

63

64 The aforementioned challenges have resulted in most MS-based proteomic studies being small and
65 specialized. In contrast, affinity binder technologies, despite their higher cost, have been widely adopted
66 for larger proteomics cohorts[16]. Consequently, there is a pressing need for robust, efficient, and high-
67 throughput MS-based proteomic workflows that are cost-effective and suitable for large-cohort studies.

68

69 Recently, the advancement of several cutting-edge technologies has driven substantial progress in MS-
70 based plasma proteomics. These technologies include nanoparticle enrichment, automated sample
71 preparation, and high-throughput mass spectrometry. Firstly, surface-engineered superparamagnetic
72 nanoparticles have emerged as a promising technology for enriching low-abundance plasma
73 proteins[10,11,17–19]. These nanoparticles interact with proteins in biological fluids to form a thin layer known
74 as the protein corona, and extensive studies have shown the critical role of protein corona in nanoparticle
75 pathophysiology[20–27]. Secondly, unlike traditional manual methods such as immunodepletion or
76 chromatographic fractionation, magnetic nanoparticle reagent-based protein capture, combined with
77 peptide identification via mass spectrometry, is well-suited for automation using common liquid handling
78 workstations equipped with magnet modules and additional components such as heater-shaker units.
79 Notably, affinity binder-based protein capture integrated with signal readout through next-generation
80 sequencing (NGS)-based proteomics technologies like Olink PEA has already been widely employed in

large cohort plasma proteomic studies, offering highly automated workflows. Lastly, newly commercialized high-resolution mass spectrometers, including ThermoFisher's Orbitrap Astral, have significantly enhanced detection depth and proteome coverage, achieving a maximum throughput of 180 samples per day per instrument[28].

Thanks to these advancements, mass spectrometry-based nanoparticle-enriched plasma proteomics has seen remarkable advancements in recent years. Blume et al.[10] introduced a 5-nanoparticle panel, functionalized with chemical ligands, that successfully identified over 2,000 proteins from 141 plasma samples using an automated sample preparation workflow. However, the reliance on multiple nanoparticles, sample fractionation, and high starting plasma volume imposes limitations on throughput and substantially increases overall costs. To address such challenges, MacCoss and colleagues[11] developed the Mag-Net enrichment strategy, which leverages strong-anion exchange magnetic microparticles to enrich membrane-bound particles, enabling the profiling of over 4,000 plasma proteins. Nevertheless, the heterogeneity of extracellular vesicles across plasma samples and impact of centrifugation conditions may pose challenges for achieving reproducibility between batches, a topic that receives minimal attention in the manuscript. For instance, Beimers et al.[12] compared six plasma proteomics methods and reported that the Mag-Net strategy achieved a detected depth of 1,200 protein groups.

While significant progress has been achieved, notable challenges continue to persist. A recent meta-analysis of 2,134 published manuscripts in the proteomics analyses of nanoparticle corona revealed substantial variability in proteome coverage[18]. Gharibi et al.[17] highlighted this issue by distributing three technical replicates of a protein corona sample to 17 mass spectrometry core facilities across the United States, uncovering notable discrepancies in plasma proteome coverage. They further demonstrated that

104 employing a standardized database search could effectively minimize these discrepancies across distinct
105 sites[19]. Therefore, there is a critical need to establish standardized nanoparticle protein corona workflows
106 encompassing nanoparticle characterization, sample preparation, and automation, along with mass
107 spectrometry instrumentation, data analysis pipelines, and achieving an optimal balance between
108 proteome coverage and throughput.

109

110 In this study, we introduced the Proteonano Ultraplex Proteomics (PUP) Platform, a comprehensive system
111 integrating proprietary monodispersed magnetic nanoparticle-based reagents for low-abundance protein
112 enrichment, an automated sample preparation workstation for parallel sample preparation (Nanomation
113 G1), and high-resolution mass spectrometry instruments for measurement, as well as built-in quality
114 control procedures to ensure consistent and reliable data monitoring throughout the workflow. The PUP
115 platform was rigorously evaluated for key features, including proteome coverage, the balance between
116 protein depth and throughput, compatibility with various mass spectrometry configurations, and batch-
117 to-batch quantitative reproducibility. Notably, it demonstrated the capacity to profile up to 6,000 proteins
118 from pooled human plasma samples, with an intra-batch quantitative median CV of 15%. Its versatility was
119 further tested across ten LC-MS/MS systems, showcasing its ability to maintain a balance between
120 throughput and proteome coverage. Finally, the platform's feasibility was validated in a cohort of
121 participants with cognitive impairment, demonstrating its potential for novel biomarker discovery in
122 plasma proteomics studies. To our knowledge, this study presents a fully integrated, robust, and
123 standardized mass spectrometry-based deep plasma proteomics platform. We believe this innovation
124 holds great promise for fostering the widespread adoption of nanoparticle-enriched MS-based platforms,
125 thereby accelerating advancements in plasma biomarker discovery on a broader scale.

126

Results

Overview of the Proteonano workflow

The Proteonano platform is intended to be a standardized, high-throughput solution for large-cohort plasma studies, combining robustness with efficiency. It integrates proprietary plasma proteome enrichment reagents, the Nanomation G1 automated workstation, and high-resolution LC-MS/MS systems (Fig. 1). The reagents feature hierarchically structured peptide-conjugated nanoparticles (PCNPs) to selectively enrich low-abundance proteins while depleting medium- and high-abundance proteins during sample preparation. Details of peptide properties are included in the Supplementary Table 1. The Nanomation G1 workstation includes a heater-shaker module for nanoparticle-plasma incubation, a magnetic module for protein corona separation, and an 8-channel liquid handler for parallel pipetting. The top view of the workstation layout can be seen in the Supplementary Information Fig. S1). For peptide identification, the platform incorporates three high-resolution mass spectrometers—Orbitrap 480, Astral instruments (ThermoFisher Scientific), and timsTOF Pro 2 (Bruker Corporation)—ensuring compatibility with widely established proteomics infrastructure (Fig. 1)

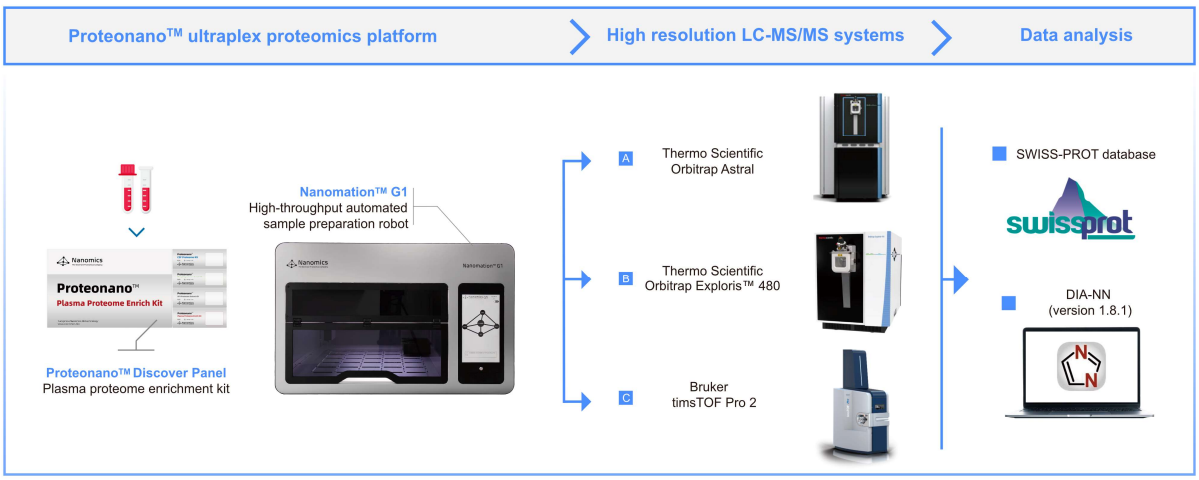


Figure 1. An overall view of the Proteonano Ultraplex Proteomics platform.

The generation of mass spectrometry-based proteomic data is often influenced by various technical and nontechnical factors. To address these challenges, we introduce the Quality Control System (QCS) for the Proteonano workflow—a framework designed to evaluate and ensure data quality and reproducibility. As illustrated in Fig. 2, the QCS encompasses three key control steps: the **Incubation Control** characterizes the proteome enrichment process via nanoparticle-plasma incubation; the **Detection Control** assesses and monitors systematic variations from LC-MS/MS instruments using premade peptide samples; and the **Data Control** includes algorithms for missing data imputation and normalization are employed to detect and correct batch-to-batch variations effectively.

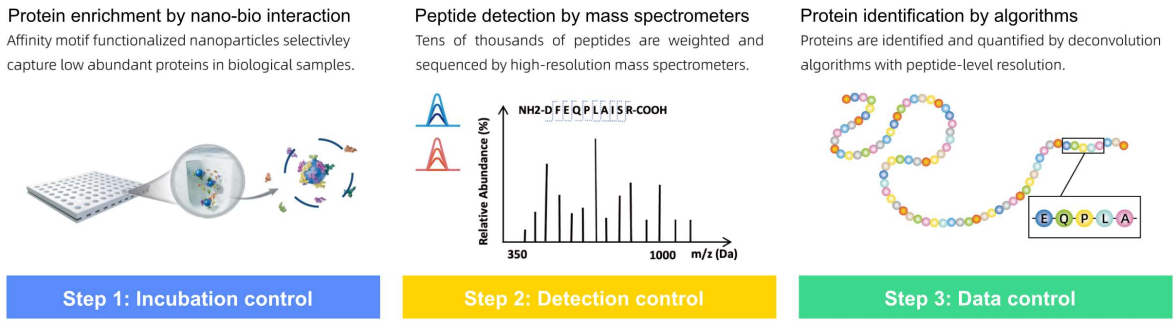


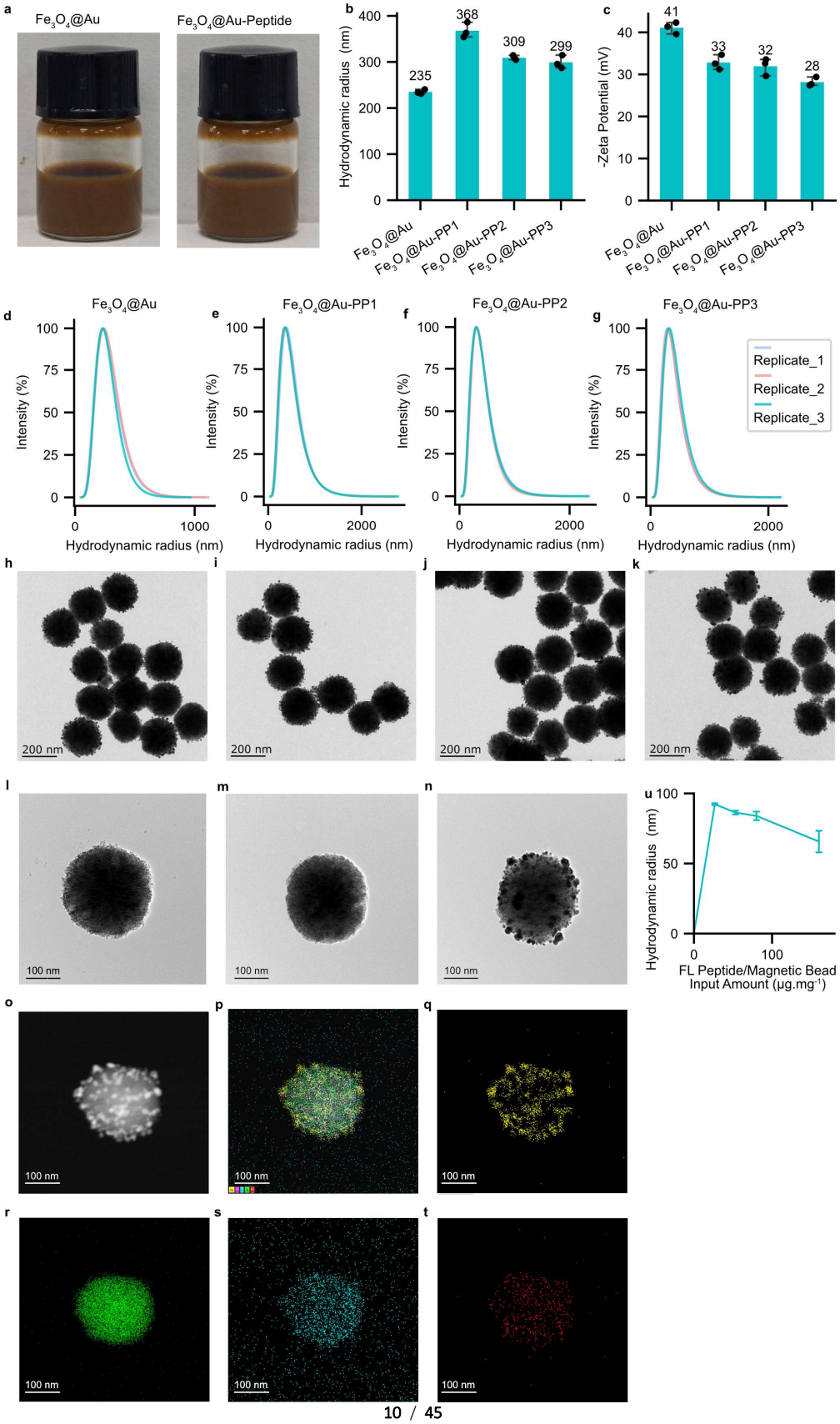
Fig. 2: The Quality Control System of the Proteonano Platform for batch-to-batch data monitoring.

In a 96-well plate experiment, six QC samples are used to monitor workflow performance. QC 1 consists of a lyophilized peptide mix derived from pooled plasma, processed with nanoparticle enrichment followed by conventional proteomic preparation steps. It is included in triplicate per plate to assess LC-MS/MS reliability. QC 2 uses neat plasma, untreated by nanoparticles, that undergoes standard sample preparation to evaluate the quality of conventional procedures. QC 3 employs pooled plasma processed

nanoparticle enrichment and standard preparation steps to monitor the performance of the complete pipeline. By comparing results from replicates of QC1, QC2 and QC3, the protein enrichment process's performance can be evaluated. The recommended positions of the six QC samples on the 96-well plate can be seen in the Supplementary Fig. S2.

Preparation and Characterization of Nanoparticles

The Proteonano nanoparticle is designed to enrich low-abundance proteins from biological samples through nano-bio interactions. As illustrated in Fig. 3a, photographs depict the dispersion of peptide-conjugated nanoparticles in water both before and after peptide conjugation. To fully characterize the multi-layered structures of the PCNPs, we conducted comprehensive measurements on key features, including size, size distribution, surface charge, elemental distribution, and surface ligand density. These analyses were achieved using a combination of advanced techniques including dynamic light scattering (DLS, Fig. 3b, d-g), transmission electron microscopy (TEM, Fig. 3h-n), , zeta potential measurements (Fig. 3c), and high-angle annular dark-field imaging (HAADF-STEM, Fig. 3o-t), and fluorescein isothiocyanate analysis (FTIC, Fig. 3u). The experimental details can be found in the Methods section.


a
Fe3O4@Au
Fe3O4@Au-Peptide
b
Hydrodynamic radius (nm)
235
368
309
299
Fe3O4@Au
Fe3O4@Au-PP1
Fe3O4@Au-PP2
Fe3O4@Au-PP3
c
-Zeta Potential (mV)
41
33
32
28
d
e
f
g
Intensity (%)
Hydrodynamic radius (nm)
Replicate_1
Replicate_2
Replicate_3
h
i
j
k
200 nm
l
m
n
100 nm
u
Hydrodynamic radius (nm)
FL Peptide/Magnetic Bead Input Amount (µg.mg-1)
o
p
q
r
s
t

178 **Figure 3. Characterization of peptide-conjugated nanoparticles (PCNPs). a** Photos of the PCNPs reagent (right) and
179 its precursor (left). **b-k** The hydrodynamic radius and zeta potential (**b**), size distribution (**d-g**), and TEM images (**h-k**)
180 of the Fe3O4@PDA@Au, Fe_3O_4@PDA@Au-PP1, Fe_3O_4@PDA@Au-PP2, and Fe_3O_4@PDA@Au-PP3. **i-n** Representative
181 TEM images of nanoparticles Fe_3O_4, Fe_3O_4@PDA, and Fe_3O_4@PDA@Au. **o Representative** HAADF-STEM images
182 Fe_3O_4@PDA@Au. **p-t** EDS elemental mapping of the dispersion of all elements combined (**p**) and individual element
183 Au (**q**), Fe (**r**), C (**s**) and N (**t**). **u** FITC characterization showing the peptide conjugation density on the surface of
184 nanoparticles.

185

186 First, TEM, DLS and zeta potential measurements were utilized to characterize three peptide-conjugated
187 nanoparticles, namely Fe_3O_4@PDA@Au-PP1, Fe_3O_4@PDA@Au-PP2, and Fe_3O_4@PDA@Au-PP3, as well as
188 their precursor (Fe_3O_4@PDA@Au). The Fe_3O_4 nanoparticles exhibit a relatively coarse surface and a regular
189 spherical morphology with an average size of 200 nm (Fig. 3b), which is in line with the hydrodynamic
190 radius and size distribution shown in Fig. 3b, d. A thin layer of polydopamine (PDA) was coated on the
191 surface of Fe_3O_4 nanoparticles to stabilize it. Upon peptide modification, the particle size increased to
192 approximately 300 nm (Fig. 3b, e-g, i-k). The surface properties of the PCNPs were further assessed using
193 zeta-potential measurement. Initially, the nanoparticles displayed a negative zeta-potential of -41.09 ±
194 1.38 mV, which shifted to around -30 mV following peptide modification, indicating changes in surface
195 charge due to peptide conjugation (Fig. 3c).

196

197 Second, HAADF-STEM imaging (Fig. 3l–o) and EDS elemental mapping were performed to investigate the
198 nanoscale structures and elemental distributions of the PCNPs. HAADF-STEM images (Fig. 3l, m, n) show
199 the progression from Fe_3O_4 nanoparticles (l) to Fe_3O_4@PDA nanoparticles (m) to Fe_3O_4@PDA@Au
200 nanoparticles (n). A noticeable surface smoothing was observed after coating Fe_3O_4 with a thin PDA layer.
201 The subsequent attachment of Au onto Fe_3O_4@PDA was evident as discrete nanospheres, rather than a

continuous Au layer. EDS elemental maps (Fig. 3q–t) reveal that Au (Fig. 3q) and N (Fig. 3t) are predominantly located on the nanoparticle's outer surface, while Fe (Fig. 3r) and C (Fig. 3s) are distributed throughout. The combined elemental map (Fig. 3p) and the analysis effectively validate the structure and composition of the synthesized nanoparticles.

Third, the density of peptide ligands on the outer layer was quantified using FITC. As shown in Fig. 2u, at an initial FITC-polypeptide concentration of 24.70 $\mu g.mg^{-1}$, the input level of fluorescent polypeptide was relatively low, yet the binding efficiency to magnetic beads was high, achieving a conjugation rate of 93%. As the FITC-polypeptide concentration increased, the effective binding efficiency gradually declined. At 53 $\mu g.mg^{-1}$—consistent with the ligand density on the NP surface (PP1, PP2, PP3)—the conjugation rate decreased to 86%, yielding an estimated ligand density of approximately 46 $\mu g.mg^{-1}$.

In conclusion, PCNPs can be visualized as a composition of three functional layers. The core layer consists of superparamagnetic Fe_3O_4 nanoparticles with a uniform diameter of 200 nm and remarkable homogeneity. The middle layer features a thin coating of Au particles, deposited as discrete nanospheres. The outer layer comprises chemically modified peptide ligands with a density of approximately 46.0 $\mu g.mg^{-1}$. These features collectively endow each nanoparticle with excellent size uniformity and a high density of binding sites, enabling the selective enrichment of low-abundance proteins within an extremely small volume. Simultaneously, high-abundance proteins are depleted during sample preparation, facilitating the detection of peptides from rare proteins in subsequent MS measurements.

Assessment of the plasma proteome coverage

To assess the proteomic detection depth of plasma samples processed with the Proteonano platform, pooled healthy donor plasma samples were analyzed using both the conventional neat plasma pipeline and the Proteonano workflow, followed by LC-MS/MS on the Orbitrap Astral instrument. As shown in Fig. 4a, the Proteonano workflow significantly improved plasma proteome coverage, identifying nearly 4000 protein groups compared to 900 with the conventional pipeline—a ~330% increase. Among these, ~2800 protein groups overlapped with proteins cataloged in the Human Plasma Proteome Project (HPPP), whereas only ~800 overlapped with the neat plasma pipeline results (Fig. 4b). Additionally, the Proteonano workflow detected 123 FDA-cleared circulating plasma biomarkers spanning eight orders of magnitude from a single pooled plasma sample, markedly outperforming the neat plasma pipeline (Fig. 4c).

Given the plasma proteome's composition of secreted proteins from various tissues—both healthy and diseased—it serves as an ideal matrix for disease biomarker discovery and pathogenesis research[29]. We further investigated the secreted proteins linked to the hallmarks of cancer, ten defining traits of cancerous cells: evading growth suppression, immune destruction, genome instability, replicative immortality, reprogrammed energy metabolism, resisting cell death, sustained angiogenesis, proliferative signaling, tissue invasion and metastasis, and tumor-promoting inflammation[30,31]. As shown in Fig. 4d, for each hallmark, the Proteonano workflow identified 3–5 times more secreted proteins compared to the unenriched process. The broader coverage of the plasma secretome associated with cancer progression may position the Proteonano platform as a promising proteomic strategy for early diagnosis, subtyping, and prognosis of various cancers.

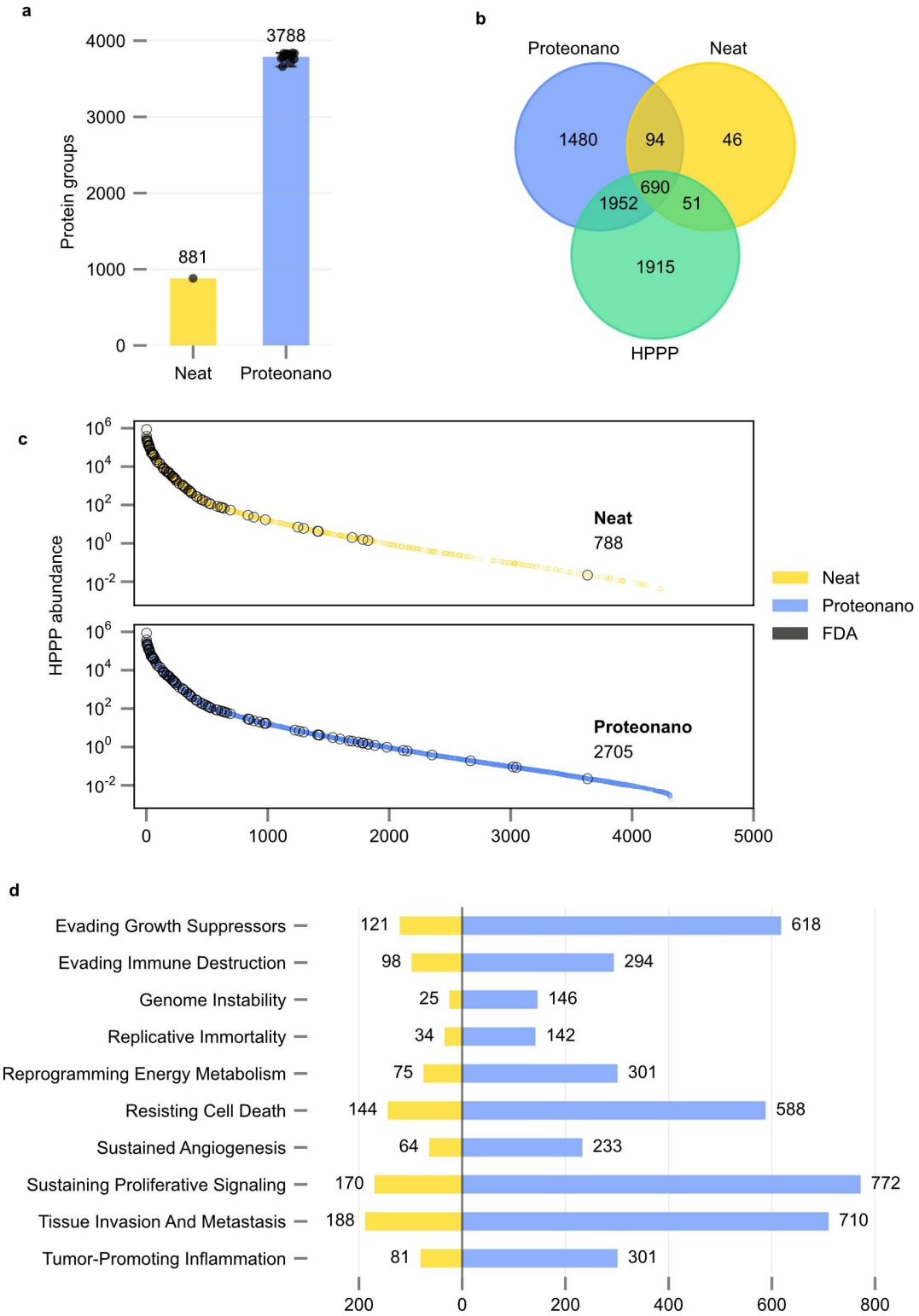

a
Protein groups
4000
3000
2000
1000
0
3788
881
Neat
Proteonano
b
Proteonano
Neat
HPPP
1480
94
46
690
1952
51
1915
c
HPPP abundance
Neat
788
Proteonano
2705
Neat
Proteonano
FDA
0
1000
2000
3000
4000
5000
d
Evading Growth Suppressors
Evading Immune Destruction
Genome Instability
Replicative Immortality
Reprogramming Energy Metabolism
Resisting Cell Death
Sustained Angiogenesis
Sustaining Proliferative Signaling
Tissue Invasion And Metastasis
Tumor-Promoting Inflammation
121
618
98
294
25
146
34
142
75
301
144
588
64
233
170
772
188
710
81
301
200
0
200
400
600
800

Figure 4. Proteomic coverage of the Proteonano workflow in comparison to neat digestion workflow. a Protein groups identified from a pooled healthy donor plasma sample processed using the Proteonano workflow and the neat plasma pipeline. DIA LC-MS/MS measurement was conducted on an Orbitrap Astral instrument (1% protein and peptide FDR; details in the Methods section). **b** Venn diagram comparing plasma proteomic coverage, highlighting overlaps between protein groups identified by the two workflows and those cataloged in the HPPP. **c** Intensity distribution of proteins groups detected matching the HPPP database. Intensities for proteins from neat plasma are shown in the upper panel. Intensities for the Proteonano are shown in the bottom panel. Black circles indicated FDA-approved protein biomarkers. **d** A comparison of the Proteonano workflow and the neat workflow regarding the coverage of secreted proteins associated with 10 hallmarks of cancer progression.

Balance between proteome coverage and throughput on ten LC-MS/MS systems.

Besides proteome coverage, throughput (typically defined as samples per day, SPD) is a crucial consideration when adopting proteomics platforms, as it directly affects the speed and scalability of data generation. High-throughput platforms are essential for large-scale or time-sensitive studies, such as biomarker discovery or population analyses. However, there is often a trade-off between throughput and proteome coverage; some platforms prioritize speed, while others focus on achieving comprehensive proteomic analysis at a slower pace. Selecting the right LC-MS/MS system depends on the specific goals, budget, and requirements of the study.

To demonstrate the versatility of the Proteonano workflow, we assessed its coverage-throughput balance across 10 distinct LC-MS/MS systems. These systems featured various combinations of liquid chromatography columns and high-resolution mass spectrometry instruments. Detailed specifications are provided in Supplementary Information Table S2.

271 We first evaluated the plasma proteome coverage of the Proteonano workflow using the Orbitrap Astral
272 instrument paired with three widely used LC columns (ES906, ES75550, µPac 110 cm), varying throughput
273 from 180SPD to 7SPD. As illustrated in Fig. 5a, even at the shortest gradient (180SPD), the workflow
274 identified 3168±9 protein IDs (AVG±SE, n=3) from a single pooled healthy plasma sample. Gradual
275 increases in gradient time enhanced proteome coverage, yielding 4945±5 protein IDs (AVG±SE, n=3) at
276 24SPD and 6410±4 protein IDs (AVG±SE, n=3) at 7SPD. Furthermore, with MS/MS parameters kept
277 constant, the proteomic detection depth was evaluated using the ES75500 and µPac 110 cm long LC
278 columns. At comparable throughputs (14SPD), both columns exhibited similar numbers of detected
279 protein groups, indicating the stability of the Proteonano workflow.

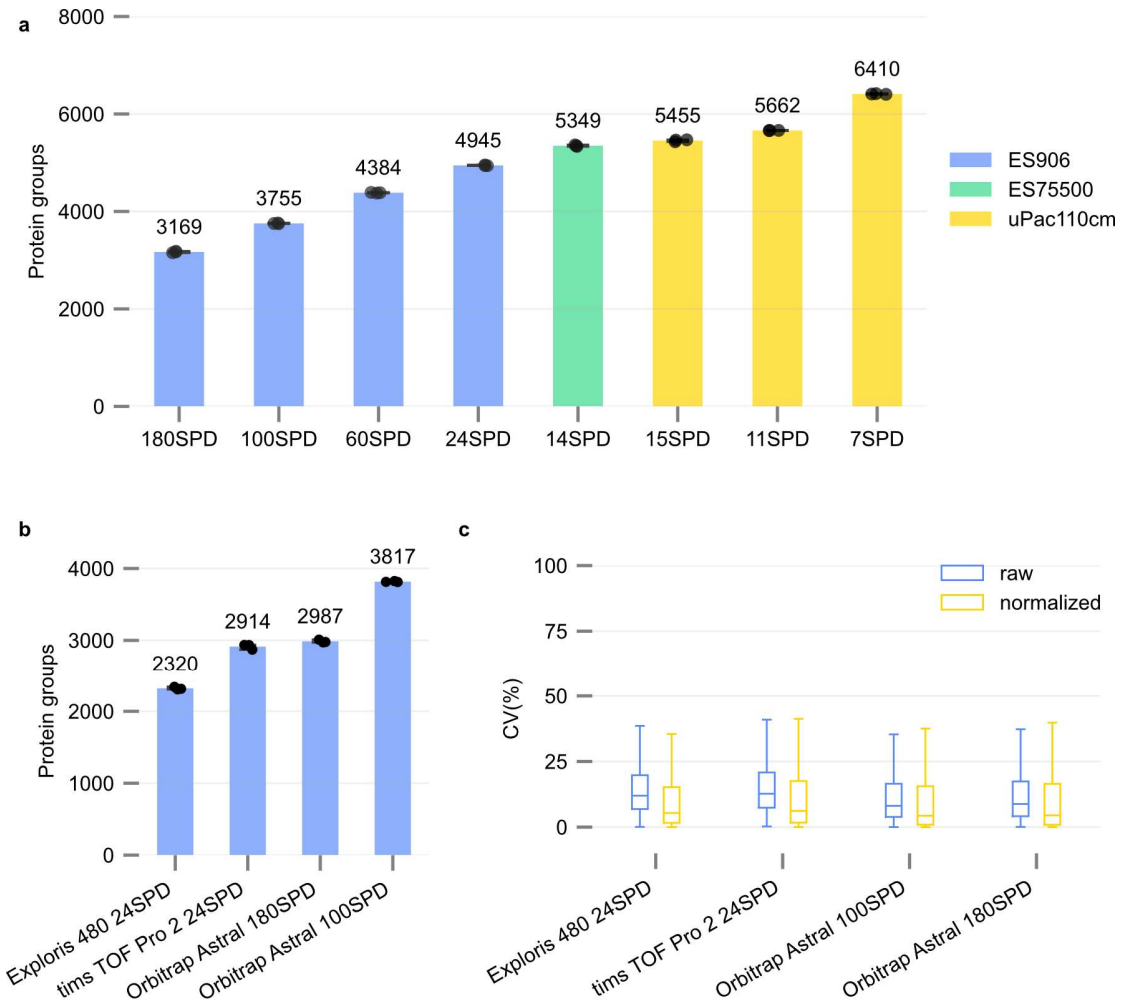
280

281 We further evaluated the impact of the Proteonano workflow on proteomic detection depth and
282 throughput using three mass spectrometers—Orbitrap Exploris 480, Orbitrap Astral, and timsTOF Pro 2—
283 all coupled with ES906 liquid chromatography. Technical assessment was conducted using three replicates
284 of a pooled healthy plasma sample. As shown in Fig. 5b, at 24SPD, Orbitrap Exploris 480 identified 2,320
285 ± 11 protein IDs (AVG ± SE, n=3), while timsTOF Pro 2 detected 2,914 ± 22 protein IDs (AVG ± SE, n=3).
286 Orbitrap Astral exhibited the best performance, identifying 2,987 ± 12 protein groups (AVG ± SE, n=3) at
287 180SPD and 3,817 ± 4 protein groups (AVG ± SE, n=3) at 100SPD. All mass spectrometers showed low and
288 similar inter-sample quantitative CVs, with Orbitrap Astral achieving slightly lower CVs (8.10% at 100SPD
289 and 8.80% at 180SPD) compared to timsTOF Pro 2 (12.71% at 24SPD) and Orbitrap Exploris 480 (11.97%
290 at 24SPD) (See Fig. 5c).

291

292 These results demonstrate the excellent stability of the Proteonano workflow across various LC-MS/MS
293 systems. Additionally, its tunable and broad depth-throughput bandwidth, ranging from over 3000 protein

294 IDs at 180SPD and up to 6400 at 7SPD, makes it a desirable MS-based platform for deep plasma profiling
295 across different instrument configurations and study requirements.



296

297 **Figure 5. Tunability of proteome coverage and throughput across 10 distinct LC-MS/MS systems. a** Number of
298 proteins groups identified using different LC columns coupled with the Orbitrap Astral Instrument (DIA mode, 1%
299 protein and peptide FDR). Bar height indicates the mean protein groups detected of three replicates of a pooled
300 plasma sample. Individual dots represent protein groups detected in each replicate. **b** Venn diagram of protein groups
301 of identified in A. **b** Proteome coverage of different mass spectrometry coupled with the same ES906 liquid
302 chromatography column. **c** Distribution of CV values of protein intensity across different LC-MS/MS systems
303 described in **b**.

304

305 **Technical Reproducibility of Proteonano Workflow**

306 Affinity-based, non-MS multiplexed proteomics platforms have established comprehensive quality control
307 measures as a fundamental approach to ensure data reproducibility. For instance, Olink implements a
308 rigorous quality control system built on its Proximity Extension Assay (PEA) technology, incorporating
309 validated biomarker panels and automated processes to reduce manual handling while maintaining data
310 integrity. These measures include protein expression normalization and stringent panel performance
311 validation[32]. Similarly, SomaLogic adopts a Total Quality Management approach focused on safety,
312 efficacy, and reproducibility. Their quality control framework integrates normalization techniques,
313 calibration, and adaptive signal adjustments to ensure consistent and accurate proteomic data[33].

314

315 We first evaluated whether variations between PCNP batches influenced proteomic analysis, as
316 reproducibility of the enrichment reagent is vital for quality control. Supplementary Fig. 4 shows high
317 pairwise Pearson correlation coefficients between two batches, confirming the strong reproducibility and
318 consistency of PCNPs in enrichment-based plasma proteomics workflows.

319

320 We next assessed the baseline reliability of the LC-MS/MS system (Vanquish Neo liquid chromatography
321 loaded with ES906 column, coupled with Orbitrap Astral) using 12 equal aliquots of the QC1 plasma sample
322 from six batches, which were processed in parallel with the Proteonano workflow. The resulting freeze-
323 dried peptides were dissolved, pooled, and consecutively injected for LC-MS/MS measurements (detailed
324 in the Methods section). As shown in Fig. 6a and 6b, 3641±15.9 protein groups (AVG±SE, n=12) were
325 identified, with an inter-plate coefficient of variation (CV) of 7.69% before and 3.38% after quantile
326 normalization, based on the number of identified protein groups. We further compared the technical CVs
327 of the low abundance proteins versus the medium- to high- abundance proteins. As shown in the boxplot

328 in Fig. 6c, low-abundance proteins displayed higher median CVs (6.65%, <10 $ng.mL^{-1}$) compared to
329 medium- (4.32%, 10–100 $ng.mL^{-1}$) and high-abundance proteins (3.86%, ⩾100 $ng.mL^{-1}$). It is anticipated
330 that the higher abundance proteins would be more quantitatively reproducible, as there is a lower noise-
331 to-signal ratio for more reliable quantification.
332

333 Second, we evaluated the inter-batch reproducibility of the complete Proteonano workflow across six
334 independent batches with each batch including three replicates of QC3 sample. As the results shown in
335 Fib. 6d and 6e, the average number of protein groups identified in each batch kept relatively consistent,
336 ranging from the lowest of 4858 to the highest of 5170. We next compared the median CV values of the 6
337 batches. The median CV of the raw unadjusted intensities were in the range of 18.2% to 26.9%. After
338 quantile normalization, the median CVs dropped to 7.4% to 10.2%. Compared to the CVs in Fig. 6c, we
339 noticed a clear increase, which is probably due to the additional variations introduced by the sample
340 preparation steps prior to mass spectrometry measurements. It is reasonable to infer that the coefficient
341 of variation for the complete Proteonano workflow (Fig. 6e) represents the combined variability stemming
342 from the sample preparation process and the LC-MS/MS measurement (Fig. 6c). Based on this rationale,
343 the average CV attributed to the sample preparation procedure was determined to be 16.5%.
344

345 Third, the reproducibility of differentially expressed proteins (DEPs) across two independent projects was
346 assessed. Project 1 comprised six batches, each containing three replicates of the QC3 sample (P1QC3),
347 while Project 2 included three batches, each with three replicates of a different QC3 sample (P2QC3). Nine
348 P1QC3 samples were randomly divided into three groups, with the same grouping applied to the nine
349 P2QC3 samples. Pairwise comparisons were conducted between the three groups from Project 1 and the
350 three groups from Project 2. DEPs were identified based on $|\log_2(\text{fold change})| > 1.2$, and categorized as

Log2FC_1, Log2FC_2, and Log2FC_3, as illustrated in Fig. 6f. The analysis showed that over 90% of DEPs were reproducibly measured across both projects. Furthermore, the Jaccard Similarity Index (JSI), a measure of quantitative similarity between datasets, was calculated for the three Log2FC groups, resulting in a value of 0.718. This high JSI score reflects excellent quantitative reproducibility of DEPs between the independent projects.

In addition, all the replicates of QC samples analyzed here were examined for potential contamination risks from platelets, erythrocytes, or coagulation, as highlighted by Geyer et al[15]. Quality marker panels across all replicates of the two projects were systematically checked, and none exhibited signs of obvious contamination (Supplementary Fig. S5, S6). Moreover, we compared the quantitative accuracy and fold-change accuracy between the reference neat digestion workflow and the Proteonano workflow, the results are Detailed results are presented in Supplementary Information, Fig. S7 and S8.

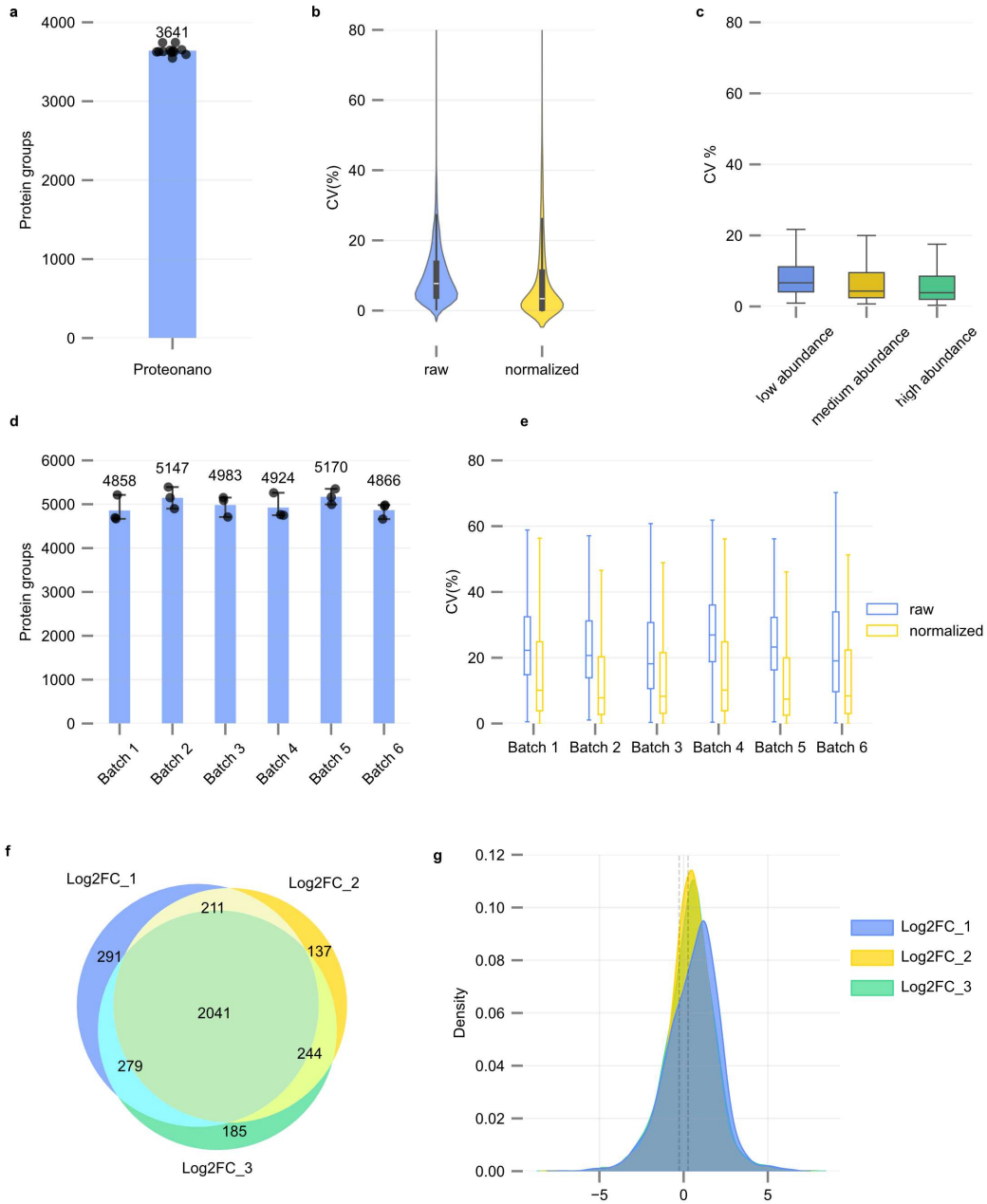


Figure 6. Investigation on the batch-to-batch reproducibility of the Proteonano workflow. a Baseline LC-MS/MS system reliability—number of identified protein groups (AVG±SE, n=12). **b** Inter-plate CVs before and after quantile normalization. **c** Technical variability comparison—boxplot illustrating median CVs of low-, medium-, and high-abundance proteins. **d** Inter-batch protein identification consistency—average number of protein groups across six independent batches. **e** Median CVs across batches and assessment of sample preparation variability. **g** Differentially

369 expressed proteins (DEPs)-pairwise comparisons of Log2FC groups across two independent projects. **f** Evaluation of
370 reproducibility using Jaccard Similarity Index.

371

372 **Use of Proteonano workflow for neurodegenerative biomarker discovery**

373 Emerging immunoassay-based and immuno-MS-based technologies have demonstrated excellent
374 accuracy in detecting well-established blood-based biomarkers linked to cognitive decline, such as plasma
375 Aβ, p-Tau, GFAP, NfL, and TDP-43[34,35]. However, these approaches are suitable for detecting a few validated
376 biomarkers for the purpose of clinical diagnosis for a specific disease. Typically, they fall short in providing
377 an unbiased discovery of novel plasma biomarkers, which could enable more nuanced patient stratification
378 and facilitate early screening—not solely for Alzheimer's pathology, but for neurodegenerative diseases in
379 general.

380

381 Affinity-based multiplexed proteomics platforms like Olink and SomaLogic have been employed in the
382 discovery of blood-based plasma biomarkers associated with cognitive decline[36]. However, these two
383 approaches generally rely on panels of pre-selected proteins and are unable to differentiate post-
384 translational modifications (PTMs), such as phosphorylation. In contrast, mass spectrometry offers an
385 unbiased approach to profiling the plasma proteome and PTMs and has the potential to discover novel
386 biomarkers. Nevertheless, it often faces challenges in detecting low-abundance blood-based biomarkers.
387 Therefore, MS is traditionally applied to brain tissues and cerebrospinal fluids rather than plasma for
388 neurological biomarker detection[37]. Thus, there remains a critical need for discovery tools that provide
389 deep proteomic coverage, robust performance, and cost-effectiveness, addressing current gaps in
390 biomarker identification.

391

We conducted a proof-of-concept study to demonstrate the feasibility of utilizing the Proteonano workflow for the discovery of potential plasma biomarker candidates. 183 plasma samples were collected from the Hubei Memory and Aging cohort Study (HAMCS) (ChiCTR; Registration number: ChiCTR1800019164), which included two tests for global cognitive function (the Mini-Mental State Examination [MMSE] and the basic Montreal Cognitive Assessment). The details of the population characteristics can be seen in the previous publications[38,39] and Supplementary Table 3. In this experiment, a total of 195 plasma samples (Control: n = 125; Cognitive decline: n = 58; QC: n=12) were processed in six batches using the Proteonano workflow. The complete workflow was visualized in Fig. 7a.

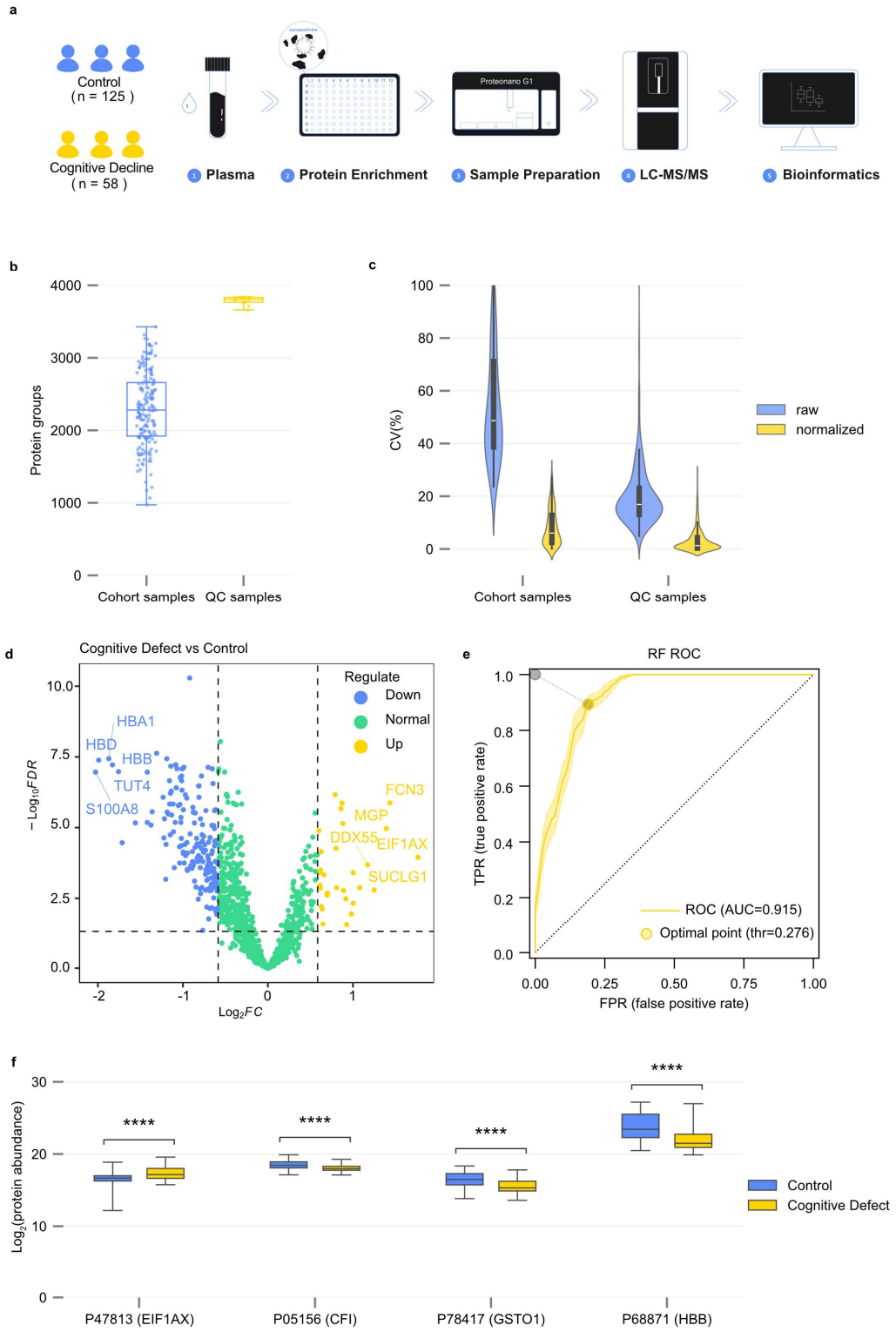
We first assessed the quality of data obtained from 12 QC samples. As shown in Fig. 7b, an average of 3788 ± 16 protein groups (AVG±SE, n=12) were identified, with a median CV of 16.9%, indicating high quantitative consistency. For the cohort samples, an average of 2298 ± 38 protein groups (AVG±SE, n=183) were identified. However, the median CV for protein raw intensities prior to quantile normalization was **48.7%** (Fig. 7c), reflecting significant biological variability compared to technical variability. In their study on biomarker discovery for Parkinson's disease via cerebrospinal fluid proteome profiling, Mann and colleagues[40] highlighted that achieving technical variability lower than biological variability is crucial for identifying disease-specific proteomic signatures. This observation emphasizes the importance of producing high-quality and consistent data in biomarker research.

In addition, through differential protein expression analysis (Fig. 7d), we identified 8 upregulated proteins groups and 49 downregulated proteins groups (FDR corrected p value < 0.05 and |log2FC| > 1) between samples with and without cognitive function decline. Additionally, we matched these 57 DEPs with the abundance data from the HPPP database and revealed that 13 of these proteins fall within the low

415 abundance range, defined as less than 10 ng mL^{-1}. The DEPs are listed in Supplementary Table 4. For
416 comparison, only 5 of the 13 low-abundance DEPs were detected in neat plasma without the use of
417 Proteonano workflow. Further, we conducted a cross-examination to determine whether these 13 low
418 abundance proteins have been detected by other plasma proteome profiling platforms. Notably, Soni et
419 al.[41] reported that 12 of the 13 DEPs were detectable using Seer's Proteograph XT platform.

420

421


a
Control
(n = 125)
Cognitive Decline
(n = 58)
Proteonano G1
1 Plasma
2 Protein Enrichment
3 Sample Preparation
4 LC-MS/MS
5 Bioinformatics
b
Protein groups
Cohort samples
QC samples
c
CV(%)
raw
normalized
Cohort samples
QC samples
d
Cognitive Defect vs Control
Regulate
Down
Normal
Up
HBA1
HBD
HBB
TUT4
S100A8
FCN3
MGP
DDX55
EIF1AX
SUCLG1
$-Log_{10}FDR$
Log_2FC
e
RF ROC
TPR (true positive rate)
FPR (false positive rate)
ROC (AUC=0.915)
Optimal point (thr=0.276)
f
Log_2(protein abundance)
****
Control
Cognitive Defect
P47813 (EIF1AX)
P05156 (CFI)
P78417 (GSTO1)
P68871 (HBB)

422

423 **Figure 7. The feasibility of using the Proteonano workflow for the discovery of potential biomarker candidates. a**
424 Study design. **b** Number of proteins identified in the Cognitive Decline sample group and Control group. **c**
425 Quantification precision assessed by calculating the CVs of the proteins in samples of the cognitive decline and
426 control group. **d** Volcano plot of differentially expressed proteins in the control group and the cognitive decline group.
427 **e** ROC curves of best multivariate model based on features selected by random forest method. **f** Illustration of
428 selected proteins that are either increased (EIF1AX) or decreased (CFI, GSTO1, and HBB) in plasma with cognitive
429 decline when compared with the ones in the Control Group.

430

431 Furthermore, we found that several proteins among the DEPs have been previously linked to cognitive
432 decline in earlier research, indicating their potential significance in neurodegenerative processes. We
433 highlight seven proteins—GPX1, RAC1, RAC2, AZU1, S100A4, SUCLG1, and CDC42, all of which have been
434 implicated in various neurodegenerative processes. GPX1, a selenium-dependent antioxidant enzyme, has
435 shown protective effects against oxidative stress in neurodegenerative disorders[42]. RAC2, known for its
436 role in cytoskeletal dynamics, regulates neurite outgrowth, and its function is impaired when γ-adducin
437 cleavage leads to its downregulation[43]. CAP37 (AZU1), a protein released by neutrophils, promotes
438 microglial activation and neuroinflammation in Alzheimer's disease, with elevated levels in AD brains
439 correlating with disease progression[44]. Similarly, S100A4 is associated with mitochondrial dysfunction and
440 the accumulation of reactive oxygen species (ROS) in astrocytes affected by Alzheimer's disease,
441 contributing to neurodegenerative pathology[45]. SUCLG1, an essential enzyme in the TCA cycle, shows
442 decreased modifications in Alzheimer's disease; however, treatment with β-hydroxybutyrate has been
443 shown to enhance its activity and ATP production, suggesting potential therapeutic avenues[46]. Lastly,
444 CDC42, a key regulator of cell signaling, is involved in Alzheimer's disease-related cognitive deficits and
445 dendritic spine loss, with its decline directly correlating with disease progression and cognitive
446 impairment[47,48]. These findings emphasize the significant roles of these proteins in neurodegenerative

processes, highlighting them as potential targets and biomarkers for further research and therapeutic development.

Although many proteins require further investigation, we explored whether a small panel of proteins could reliably distinguish individuals with cognitive decline from those without. Differential protein expression data were analyzed using a random forest approach, which identified key features separating the two cohort groups. Based on these features, optimal multivariate models were constructed using the Akaike Information Criterion. The best model incorporated four proteins—EIF1AX, CFI, GSTO1, and HBB—and achieved an excellent AUC value of 0.92. Fig. 7e presents a receiver operating characteristic (ROC) curve, demonstrating the model's ability to distinguish participants in the Cognitive Decline Group from the Control. Notably, GSTO1's role in modulating oxidative stress stands out as a significant factor that may contribute to the underlying pathophysiological mechanisms of neurodegeneration[49]. This suggests its potential as a proteomic signature for patient stratification. However, transition from biomarker discovery to validation is a lengthy, time-consuming, and oftentimes expensive process. It involves critical steps of study design, including cohort selection, evaluation of statistical power, sample blinding, randomization, and rigorous data quality control[50–53]. While a detailed discussion of these best practices exceeds the scope of this study. Further efforts on the validation of their candidacy as potential biomarkers are necessary and will be the focus of our future work.

For cross-referencing, another nanoparticle-based MS proteomics platform, known as Proteograph XT, was utilized by Lucar et al.[54] to analyze plasma samples from 379 AD patients and 240 healthy controls using a linear mixed-effects model to identify AD-associated proteins. Their findings revealed that proteins upregulated in AD included DKK2, MAPK3, HLA-DRA, LPL, and ACSL4, among others, while downregulated

470 proteins included ETFA, TFRC, ADH1B, MGST3, and TNFRSF6B. Conversely, the Mag-Net enrichment
471 strategy examined plasma samples from 10 AD patients and 10 healthy controls, applying a two-sided
472 Mann-Whitney U test to identify AD-associated proteins[11]. This approach identified upregulated proteins
473 in AD such as INPPL1, VPS13A, DNAJC10, S100B, and HTT, while downregulated proteins included VCP,
474 MAP2K2, USP15, JMJD6, and CD36. Notably, our study observed a consistent trend in the differential
475 expression of VCP with results reported by the Mag-Net strategy. However, none of the well-established
476 blood biomarkers—such as Aβ, p-Tau, GFD, and TDP-42—were identified by any of those three
477 nanoparticle-enriched mass spectrometry workflows. We hypothesize that factors like cohort
478 characteristics, sample size, population diversity, and nanoparticle enrichment methodology may have
479 contributed to this outcome. Further investigations with refined experimental designs are warranted to
480 explore these observations more comprehensively.

481 **Discussion**

482 This study developed and systematically evaluated the Proteonano Ultraplex Platform for robust, deep
483 plasma proteomic analysis. When paired with advanced mass spectrometers, the platform demonstrated
484 exceptional capability for deep proteomic analysis of low-input plasma samples, achieving extensive
485 proteome coverage, high throughput, a dynamic range spanning nine orders of magnitude, and minimized
486 batch-to-batch variability. Furthermore, we demonstrated that the Proteonano workflow exhibited
487 remarkable stability and reproducibility across ten distinct and widely adopted DIA LC-MS/MS systems,
488 showcasing flexibility in balancing protein depth and sample processing speed. Finally, the platform's
489 performance was validated using plasma samples from an elderly community cohort, including individuals
490 with neurodegenerative diseases.

491

492 The PUP Platform described here offers several advantages for MS-based plasma proteomics studies,
493 especially in the discovery phase of biomarker screening. Unlike traditional immunodepletion methods,
494 our platform is highly standardized, robust, high-throughput, user-friendly, cost-effective, and compatible
495 with various mass spectrometry configurations across different laboratories. Notably, the workflow
496 requires only 40 µL of plasma per sample to achieve optimal performance, whereas previously reported
497 technologies typically require at least 100 µL[12]. This critical feature, which has not been emphasized
498 enough, plays a key role in the practical adoption of new technologies, especially given the common
499 limitations on sample volumes. In comparison, Olink PEA technology, widely adopted by dozens of
500 biopharmaceutical companies and applied to millions of samples globally, requires only 10 µL of plasma.

501

502 Secondly, the PUP Platform achieved cutting-edge deep plasma proteome coverage and data quality by
503 effectively combining multiple nanoparticles into a single, uniform reagent, a strategy akin to those utilized
504 in multiplexed immunoassays but without the challenges of antibody cross-reactions. This innovative
505 approach removes the need for sample fractionation and multiple mass spectrometry runs for individual
506 samples. As a result, it significantly minimizes time, cost, and batch effect in plasma proteomics studies,
507 making it particularly advantageous for large cohorts involving hundreds or even thousands of samples. In
508 the future, we plan to enhance nanoparticle technology by integrating a wider range of compositions into
509 a single reagent to enrich a broader range of low abundance plasma proteins. This innovation aims to
510 achieve deeper plasma proteome coverage through MS-based proteomics, bringing its capabilities closer
511 to those of non-MS multiplex technologies regarding proteome coverage, such as SomaScan 7K and 11K[55].

512

513 Thirdly, the PUP Platform has implemented a fully standardized workflow, starting with nanoparticle-
514 plasma incubation at the beginning of the experiment and extending to data analysis across various

515 batches. This carefully structured system incorporates a robust quality control framework, guarantees
516 compatibility with multiple mainstream mass spectrometers, supports deep plasma coverage, and
517 minimizes batch-to-batch effect over independent studies. This level of standardization significantly
518 improves reliability and efficiency, solidifying its applicability to large-scale plasma proteomics studies.

519

520 That said, we have not included discussions of other considerations that may be associated with the
521 proteome coverage and data reproducibility of nanoparticle-enriched plasma proteomics. For instance,
522 the effect of nanoparticle size and size distribution on the compositions of protein corona is less studied.
523 Notably, Tenzer et al.[23] demonstrated the significance of nanoparticle size in shaping corona layers by using
524 monodispersed silica nanoparticles of varying sizes (20, 30, and 100 nm). Their findings revealed that
525 nanoparticle size critically influenced the composition of the corona, which accounted for 37% of the
526 identified proteins in their study[23]. Yet, the effect of size uniformity on the coverage and the reproducibility
527 of proteomic analysis remains unexamined. In this context, the surface-modified nanoparticles presented
528 here offer an average diameter of 300 nm and exhibit exceptional size uniformity (PDI < 0.05, Fig. 3b). In
529 comparison, particles documented by Blume[10] and Wu[11] are either characterized by relatively high
530 polydispersity or, in the case of microparticles, lack detailed data on size distribution. Alternative
531 approaches, such as graphene oxide-modified materials, have also been explored for proteome
532 enrichment[56]. However, the highly heterogeneous nature of these flake-like nanomaterials may pose
533 substantial challenges on quality control, particularly when employed in large cohort plasma studies.
534 Although the precise mechanisms remain unclear, we hypothesize that employing nanoparticles with
535 defined size and low polydispersity is crucial for ensuring quality control and enhancing reproducibility in
536 plasma proteomics applications[57].

537

538 In addition, we recognize that the proof-of-concept study using the PUP Platform for biomarker discovery
539 is subject to several limitations. First, mass spectrometers inherently face significant challenges in
540 detecting low-abundance proteins in plasma. This limitation is particularly evident for biomarkers of
541 neuropathology, which are present at very low concentrations in plasma (typically a few $pg.mL^{-1}$). That
542 explains partly why none of the three reported nanoparticle-enriched and MS-based workflows detected
543 low abundant biomarkers like p-Tau181. We cannot rule out the possibility that these AD-related
544 biomarkers were also present in this cohort but at concentrations below the detection sensitivity of mass
545 spectrometry, even after enrichment. While the PrecivityAD2 assay also relies on mass spectrometry, it
546 incorporates the use of antibody-conjugated microbeads to enrich specific AD biomarkers from plasma
547 prior to LC-MS/MS measurements. Second, it is evident that the cohort is relatively small and lacks
548 diversity. Mann and colleagues proposed a "rectangular" plasma proteome profiling strategy that
549 suggested [29] typically 100 to 1000 samples should be used in the phase of plasma biomarker discovery.

550

551 Therefore, we anticipate that PUP Platform could serve as a valuable discovery-phase tool, enabling the
552 profiling of thousands of plasma proteins to identify potential biomarkers. These identified biomarkers can
553 then undergo further validation in larger cohorts using targeted proteomics techniques such as MS-based
554 Parallel Reaction Monitoring (PRM) and single-molecule immunoassays like Simoa, or even low-cost ELISA.

555

556 In summary, our study highlights the Proteonano Ultraplex Proteomics Platform as an innovative solution
557 for MS-based deep plasma proteomics. This platform offers extensive proteomic coverage, high-
558 throughput capabilities, cost efficiency, and a standardized data production workflow. It supports minimal
559 plasma input (40 µL), demonstrates compatibility with various LC-MS/MS systems and throughput levels,
560 and ensures batch-to-batch consistency, reproducibility, and precision in protein identification and analysis

in large-cohort studies. These advancements offer exceptional potential to push the frontiers of plasma proteomics and pave the way for the discovery of novel protein biomarkers and therapeutic targets across a range of diseases. In future endeavors, we aim to expand our research by evaluating the performance of the PUP platform across multiple laboratories and conducting larger, multi-centered plasma proteomics studies to further validate its robustness and scalability.

Methods

Blood Samples

For most of the experiments, pooled healthy donor blood samples were obtained from Oricells (Ori Biotech, Shanghai, China), Whole blood was collected in K2-EDTA containing tubes. Samples were centrifuged at 1500 *g* for 15 min at 4 °C. Supernatant was transferred from centrifuge tubes. Plasma

580 samples from multiple donors were pooled, aliquoted into microfuge tubes, and stored at −80 °C until
581 analysis.

582

583 Plasma samples collected from a community cohort aimed at understanding cognitive deficiencies in
584 elderly patients in Hubei Province (Hubei Memory and Aging Cohort Study, HMACS) were utilized. Ethical
585 approval for blood sample collection was received at respective institutions for the study.

586 **Preparation and Characterization of Nanoparticles**

587 The magnetic Fe_3O_4 particles were prepared according to a solvothermal approach. Typically, $FeCl_3$·6H2O
588 (3.91 g) and trisodium citrate (0.46 g) were first dissolved in ethylene glycol (70 mL). Then, NaAc (9.42 g)
589 was added, and the mixture was vigorously stirred to form a transparent solution. Afterward, the solution
590 was transferred to a 150 mL Teflon-lined stainless-steel autoclave. The autoclave was sealed and heated
591 at 200 °C and maintained for 12 h, and then it was allowed to cool to room temperature. The products
592 were washed with ethanol and deionized water several times and dried at 60 °C for 12 h.

593

594 In order to prepare monodispersed Fe_3O_4@PDA (polydopamine) core-shell nanospheres, 180 mg of the
595 hydrochloride dopamine was dispersed in 90 ml Boric acid-borax buffer (0.5 mol.L^{-1}, pH=8.0). Then, 90
596 mg of the as-prepared Fe_3O_4 was added, stirring overnight at room temperature. The resultant product
597 was separated and collected with a magnet and then washed with ultrapure water three times.

598

599 In a typical preparation of Fe_3O_4@PDA@Au, firstly, 20 mg of the Fe_3O_4@PDA was added into 12.5 mL
600 deionized water, sonicate for 10 min. Then, 0.5 mL dopamine solution (0.76 mg of the hydrochloride

601 dopamine was dispersed in 1 mL deionized water) and 83.5 µL sodium hydroxide standard solution (0.6M)
602 were added in sequence. Typically, 0.8 mL 1 wt% aqueous solution of $HAuCl_4$ (1 g $HAuCl_4$ was dispersed in
603 100 mL deionized water) was quickly added and reaction was allowed for overnight at room temperature
604 under shaking at 1000 rpm. Finally, the product was separated by a magnet and washed with water.

605

606 The nanomaterial sample was ultrasonically dispersed in ethanol, and a drop of the solution was then
607 placed onto an ultra-thin carbon film. After drying, the sample was analyzed using a transmission electron
608 microscope (TEM). The instrument used was an FEI Tecnai G2 F20 with a maximum acceleration voltage
609 of 200 kV. It has a point resolution of 0.24 nm, a line resolution of 0.102 nm, and an information resolution
610 of ≤0.14 nm. Energy dispersive spectroscopy (EDS) was performed using an Oxford XPLORE system.

611

612 Well dispersed Fe_3O_4 microspheres were prepared by a solvothermal method[58]. Fe_3O_4 microspheres were
613 dispersed in chloroauric acid containing sodium citrate aqueous solution and stirred at 90 °C for 0.5 h.
614 Resulting Au NPs-deposited Fe_3O_4 (Fe_3O_4@PDA@Au) were magnetically separated from the suspension
615 and subsequently washed dried under vacuum.

616

617 For peptide conjugation, peptides were dissolved in water. Dissolved peptide was incubated with Tris(2-
618 carboxyethyl) phosphine hydrochloride (TCEP) solution, followed by adding of Fe_3O_4@PDA@Au
619 synthesized above. The reaction mixture was incubated at 25°C overnight. Reaction product was washed
620 by ethanol then deionized water and stored at refrigerator at 4°C. Three peptides, PP1
621 (HKAATKIQASFRGHITRKKLC), PP2 (DIEEVEVRSKYFKKNERTVEC), and PP3 (DIEEVEVRSKYFKKNERTVEC) were

custom synthesized by GenScript (Suzhou, China). The details can be seen in the Supporting Information S1 section.

Following peptide conjugated Fe_3O_4@PDA@Au synthesis, transmission electron microscopy (TEM) was performed by using a Tecnai 12 electron microscope (ThermoFisher Scientific) at an accelerating voltage of 200 kV. DLS and Zeta potential of the particles were examined by using a NanoBrook 90Plus PALS instrument (Brookhaven Instruments, Holtsville, NY).

Peptide Conjugation Density

First, 125 µL of TCEP solution (200 mM) and 5 mL of FITC-peptide solution (0.2 $mg.mL^{-1}$) were added to a centrifuge tube, mixed thoroughly, and incubated in the dark for 2 hours. Then, varying volumes (0, 200, 400, 600, and 1200 µL) of the TCEP-treated FITC-peptide solution (fluorescence intensity denoted as I_o) were transferred into separate 5 mL centrifuge tubes. Each tube was supplemented with 1000 µL of Fe_3O_4@PDA@Au NPs (1.5 $mg.mL^{-1}$). Ultra-pure water was then added to bring the final volume to 2300 µL, followed by ultrasonic mixing for 5 minutes. The reaction mixtures were incubated at 25°C under shaking conditions for 17 hours. After the reaction, the samples were washed three times with ultra-pure water, and the supernatants from the three washes were collected (fluorescence intensity denoted as I_1). All fluorescence intensity values were measured in triplicate using a Tecan SPARK multifunctional microplate reader.

642 **Sample Preparation by the Proteonano Workflow**

643 Automatic sample preparation was performed in batches utilizing Nanomation G1 workstation, equipped
644 with both a magnetic module and a heater shaker module. Plasma samples were dispensed on 96 well flat
645 bottom assay plates for subsequent processing. For most experiments, 40 µL of human plasma was diluted
646 to a final volume of 100 µL by using 1 x PBS at pH7.4 and was subsequently combined with PCNPs (as
647 synthesized above) from the Proteonano reagent in a 1:1 volumetric ratio. The mixture was incubated at
648 25 °C and agitated for 60 min. Following magnetic immobilization, PCNPs were washed thrice with 1 x PBS.
649 Proteins captured on the PCNPs were reduced by 20 mM DTT at 37 °C for 60 min. Alkylation was performed
650 using 50 mM iodoacetamide (IAA) at room temperature in darkness for 30 min. Trypsin (Promega
651 Corporation, Madison, WI, USA) digestion was carried out at 37°C for a duration of 16 hours with shaking
652 at 1000 rpm. Post-digestion, peptides were purified using desalting C18 columns in micropipette tip format
653 (ThermoFisher Scientific) and lyophilized with a LyoQuest freeze dryer (LyoQuest, Telstar, Terrassa, Spain).
654 Lyophilized peptides were then reconstituted in 0.1 % formic acid prior to mass spectrometry. Peptide
655 concentrations were measured with a Nano300 microvolume spectrophotometer (Allsheng Instruments,
656 China).

657

658 **LC-MS/MS Experiments**

659 Multiple LC-MS/MS instruments were used in the study. For most studies, Orbitrap Astral (ThermoFisher
660 Scientific) mass spectrometer was used. In some cases, Orbitrap Exploris 480 (ThermoFisher Scientific)
661 mass spectrometer equipped with FAIMS, or timsTOF Pro 2 mass spectrometer (Bruker Instruments) were
662 used.

663

For studies using Orbitrap Astral instrument, 300 ng or 500 ng peptides dissolved in 0.1% formic acid were separated by a Vanquish™ Neo UHPLC system (ThermoFisher Scientific) followed by mass spectrometry. Data were acquired in data independent acquisition mode (Detailed experimental setup are listed in Supplementary Table S5-14).

For studies using Orbitrap Exploris 480, 500 ng of peptides were separated by an Easy-nLC1200 reverse-phase HPLC system (ThermoFisher Scientific) using a precolumn (homemade, 0.075 mm × 2 cm, 1.9 µm, C18) and a self-packed analytical column (0.075 mm × 20 cm, 1.9 µm, C18) over a 48 min gradient before nano-electrospray on Orbitrap Exploris 480 mass spectrometer equipped with FAIMS (ThermoFisher Scientific). Solvent A was 0.1 % formic acid and solvent B was 80 % acetonitrile (ACN)/0.1 % formic acid. Gradient conditions were 3–7 % B (0–1 min), 7–30 % B (1–36 min), 30–95 % B (36–38 min), and 95 % B (38–48 min). Mass spectrometer was operated in DIA mode. Spray voltage was set to 2.4 kV, RF Lens level at 40 %, and heated capillary temperature at 320 °C. Full MS resolutions were set to 60,000 at m/z 200 and full MS automatic gain control (AGC) target was 100 % with an injection time (IT) of 50 ms. Mass range was set to m/z 350–1200. The AGC target value for fragment spectra was set at 1,000 %. Resolution was set to 30,000 and IT to 54 ms. Normalized collision energy was set at 30 %. Default settings were used for FAIMS with voltages applied as -45 V, except gas flow, which was applied with 3.5 L/min.

For studies using timsTOF Pro 2 instrument, 200 ng of peptides was subjected to a nanoElute® 2 nanoLC system coupled with a Bruker timsTOF Pro 2 mass spectrometer using a trap-and-elute configuration. First, the peptides were loaded on an Acclaim PepMap 100 C18 trap column (0.3 mm ID × 5 mm), then separated on a self-packed analytical column (0.075 mm × 25 cm, 1.8 µm, C18) at a flow rate of 300 nL min^{-1}. Solvent A was 0.1 % formic acid and solvent B was 80 % ACN/0.1 % formic acid. Gradient conditions were 2–22 %

687 B (0–45 min), 22–35 % B (45–50 min), 35–80 % B (50–55 min), 80 % B (55–60 min). The spray voltage was
688 set to 1.5 kV. The mass spectrometer was operated in diaPASEF mode using ion mobility range of 0.76–
689 1.29 Vs/cm^2 with 100 ms accumulation time. MS1 mass spectrometry scans 452-1152 m/z with peak
690 height above 2,500 before being detected. The 452-1152 m/z range was divided into four steps. Each step
691 was divided into seven windows, the Number of Mobility Windows was set to 2, and a total of 56 windows
692 for continuous window shattering and information gathering. The splitting mode was CID and the splitting
693 energy was set 20-59 eV. The mass width of each window was 25 Da and the cycle time for a DIA scan was
694 1.59 s.

695 **Proteomic Data Analysis**

696 The *.raw* or *.d* files were obtained directly from mass spectrometers, or *mzML* files converted from .raw
697 files by msConvert[59] (Version 3.0) software were searched using DIA-NN (Version 1.8.1) in library free mode.
698 For each sample, spectra were searched against a UniProt *Homo sapiens* reviewed proteome dataset, or
699 other datasets. DIA-NN search parameters were: 10 ppm mass tolerance for mass accuracy, one missed
700 cleavages of trypsin, carbamidomethylation of cysteine as fixed modification, and methionine oxidation as
701 the only variable modification. The rest of the parameters were set to default. The FDR cutoffs at both
702 precursor and protein level were set to 0.01.

703

704 The .raw or .d files generated directly from mass spectrometers, or mzML files converted from .raw files
705 using msConvert (Version 3.0) software, were analyzed with DIA-NN (Version 1.8.1) in library-free mode.
706 For each sample, spectra were matched against a UniProt Homo sapiens reviewed proteome dataset or
707 other relevant datasets. The DIA-NN search parameters included a 10 ppm mass tolerance for mass
708 accuracy, one missed cleavage for trypsin, carbamidomethylation of cysteine as a fixed modification, and

709 methionine oxidation as the sole variable modification. All other parameters were set to default, with false
710 discovery rate (FDR) cutoffs of 0.01 applied at both the precursor and protein levels.

711

712 **Statistical Analysis and Data Visualization**

713 Coefficient of variation (CV) for each protein was determined by dividing its empirical standard deviation
714 by its empirical mean, and CV analyses were performed on raw intensities and quantile-normalized
715 intensities, respectively. Median values were reported as overall coefficient of variation. Pearson
716 correlation analyses and linear regression were conducted using Pingouin[59]. No missing value imputation
717 was performed during the above analysis, and calculations were performed only for proteins that were
718 identified in all samples. For replicate experiments, protein identifications are expressed as AVG±SE.
719 Seaborn was used to generate bar charts and violin charts[60]. The matplotlib-venn[61] package
720 (https://pypi.org/project/matplotlib-venn/) and the Venn[62] (https://pypi.org/project/venn/) package
721 were used to plot VENN[39] plots. MS-DAP was used to screen for differentially expressed proteins, and
722 ggVolcano[63] (https://github.com/BioSenior/ggVolcano) was used to generate volcano plot. KEGG pathway
723 enrichment analysis was performed using the GSEApy[64] software package in Python. Feature selection was
724 performed using MetaboAnalystR[65]. Multivariate analysis was performed using MASS[66] in R. Variables
725 were first processed by glm (MASS package in R), followed by Akaike information criterion (AIC)
726 determination using stepAIC (MASS package in R) with forward selection and Wald test. Multivariate
727 Receiver operating characteristic(ROC) analysis was subsequently performed using roc (glmtoolbox[67]
728 package in R) to assess the predictive performance of the logistic regression model selected by step AIC.
729 ROC curves were graphed using ROC-utils[68].

Data Availability

All raw and processed proteomics data generated in this study have been deposited to the iProX repository and are available for review. The data can be accessed at https://www.iprox.cn/page/PSV023.html?url=1744530654097Tmnc using the password dO25. The data are currently under controlled access for peer review and will become publicly available after the embargo period. Alternatively, the data can be made available upon reasonable request by contacting the corresponding author.

References

1. Zhong, W. *et al.* Next generation plasma proteome profiling to monitor health and disease. *Nat. Commun.* **12**, 2493 (2021).

2. Deng, Y.-T. *et al.* Atlas of the plasma proteome in health and disease in 53,026 adults. *Cell* **188**, 253-271.e7 (2025).

3. He, B., Huang, Z., Huang, C. & Nice, E. C. Clinical applications of plasma proteomics and peptidomics: Towards precision medicine. *Proteomics Clin. Appl.* **16**, e2100097 (2022).

4. Giudice, G. & Petsalaki, E. Proteomics and phosphoproteomics in precision medicine: applications and challenges. *Brief. Bioinform.* **20**, 767–777 (2017).

5. Yu, X., Schwenk, J. M., Xu, P. & LaBaer, J. Advances in plasma proteomics: Moving from technology to precision medicine. *PROTEOMICS – Clin. Appl.* **16**, 2200083 (2022).

6. Soni, R. K. Frontiers in plasma proteome profiling platforms: innovations and applications. *Clin. Proteomics* **21**, 43 (2024).

7. Schmidt, I. M. *et al.* Plasma proteomics of acute tubular injury. *Nat. Commun.* **15**, 7368 (2024).

8. Liu, W.-S. *et al.* Plasma proteomics identify biomarkers and undulating changes of brain aging. *Nat. Aging* **5**, 99–112 (2025).

9. Guo, T., Steen, J. A. & Mann, M. Mass-spectrometry-based proteomics: from single cells to clinical applications. *Nature* **638**, 901–911 (2025).

10. Blume, J. E. *et al.* Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat. Commun.* **11**, 3662 (2020).

11. Wu, C. C. *et al.* Mag-Net: Rapid enrichment of membrane-bound particles enables high coverage quantitative analysis of the plasma proteome. *BioRxiv Prepr. Serv. Biol.* 2023.06.10.544439 (2024) doi:10.1101/2023.06.10.544439.

12. Beimers, W. F. *et al.* A Technical Evaluation of Plasma Proteomics Technologies. *BioRxiv Prepr. Serv. Biol.* 2025.01.08.632035 (2025) doi:10.1101/2025.01.08.632035.

13. Liu, X. *et al.* Plasma-to-tumour tissue integrated proteomics using nano-omics for biomarker discovery in glioblastoma. *Nat. Commun.* **16**, 3412 (2025).

14. Collins, B. C. *et al.* Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat. Commun.* **8**, 291 (2017).

15. Geyer, P. E. *et al.* Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol. Med.* **11**, e10427 (2019).

16. Gan, Y.-H. *et al.* Large-scale proteomic analyses of incident Parkinson's disease reveal new pathophysiological insights and potential biomarkers. *Nat. Aging* (2025) doi:10.1038/s43587-025-00818-0.

17. Ashkarran, A. A. *et al.* Measurements of heterogeneity in proteomics analysis of the nanoparticle protein corona across core facilities. *Nat. Commun.* **13**, 6610 (2022).

18. Hajipour, M. J. *et al.* An Overview of Nanoparticle Protein Corona Literature. *Small Weinh. Bergstr. Ger.* **19**, e2301838 (2023).

19. Gharibi, H. *et al.* A uniform data processing pipeline enables harmonized nanoparticle protein corona analysis across proteomics core facilities. *Nat. Commun.* **15**, 342 (2024).

20. Walkey, C. D. & Chan, W. C. W. Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. *Chem. Soc. Rev.* **41**, 2780–2799 (2012).

21. Walkey, C. D. *et al.* Protein Corona Fingerprinting Predicts the Cellular Interaction of Gold and Silver Nanoparticles. *ACS Nano* **8**, 2439–2455 (2014).

22. Tenzer, S. *et al.* Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nat. Nanotechnol.* **8**, 772–781 (2013).

23. Tenzer, S. *et al.* Nanoparticle size is a critical physicochemical determinant of the human blood plasma corona: a comprehensive quantitative proteomic analysis. *ACS Nano* **5**, 7155–7167 (2011).

24. Ritz, S. *et al.* Protein Corona of Nanoparticles: Distinct Proteins Regulate the Cellular Uptake. *Biomacromolecules* **16**, 1311–1321 (2015).

25. Hofmann, D. *et al.* Mass Spectrometry and Imaging Analysis of Nanoparticle-Containing Vesicles Provide a Mechanistic Insight into Cellular Trafficking. *ACS Nano* **8**, 10077–10088 (2014).

26. Landgraf, L. *et al.* A plasma protein corona enhances the biocompatibility of Au@Fe3O4 Janus particles. *Biomaterials* **68**, 77–88 (2015).

27. Chithrani, B. D. & Chan, W. C. W. Elucidating the Mechanism of Cellular Uptake and Removal of Protein-Coated Gold Nanoparticles of Different Sizes and Shapes. *Nano Lett.* **7**, 1542–1550 (2007).

28. Heil, L. R. *et al.* Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. *J. Proteome Res.* **22**, 3290–3300 (2023).

29. Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Mol. Syst. Biol.* **13**, 942 (2017).

30. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).

31. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).

32. Wik, L. *et al.* Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. *Mol. Cell. Proteomics MCP* **20**, 100168 (2021).

33. Candia, J. *et al.* Assessment of Variability in the SOMAscan Assay. *Sci. Rep.* **7**, 14248 (2017).

34. Chatterjee, P. *et al.* Plasma Aβ42/40 ratio, p-tau181, GFAP, and NfL across the Alzheimer's disease continuum: A cross-sectional and longitudinal study in the AIBL cohort. *Alzheimers Dement. J. Alzheimers Assoc.* **19**, 1117–1134 (2023).

35. Schindler, S. E. *et al.* Head-to-head comparison of leading blood tests for Alzheimer's disease pathology. *Alzheimers Dement.* **20**, 8074–8096 (2024).

36. Teunissen, C. E. *et al.* Methods to Discover and Validate Biofluid-Based Biomarkers in Neurodegenerative Dementias. *Mol. Cell. Proteomics MCP* **22**, 100629 (2023).

37. Johnson, E. C. B. *et al.* Large-scale proteomic analysis of Alzheimer's disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).

38. Li, L. *et al.* The Hubei Memory and Aging Cohort Study: Study Design, Baseline Characteristics, and Prevalence of Cognitive Impairments. *J. Alzheimers Dis. JAD* **85**, 561–571 (2022).

39. Cheng, G.-R. *et al.* Prevalence and risk factors for subjective cognitive decline and the correlation with objective cognition among community-dwelling older adults in China: Results from the Hubei memory and aging cohort study. *Alzheimers Dement. J. Alzheimers Assoc.* **19**, 5074–5085 (2023).

40. Karayel, O. *et al.* Proteome profiling of cerebrospinal fluid reveals biomarker candidates for Parkinson's disease. *Cell Rep. Med.* **3**, 100661 (2022).

41. Soni, R. K. Frontiers in plasma proteome profiling platforms: innovations and applications. *Clin. Proteomics* **21**, 43 (2024).

42. Sharma, G. *et al.* Glutathione peroxidase-1 and neuromodulation: Novel potentials of an old enzyme. *Food Chem. Toxicol.* **148**, 111945 (2021).

43. Xiong, M. *et al.* A γ-adducin cleavage fragment induces neurite deficits and synaptic dysfunction in Alzheimer's disease. *Prog. Neurobiol.* **203**, 102074 (2021).

827 44. Stock, A. J., Kasus-Jacobi, A. & Pereira, H. A. The role of neutrophil granule proteins in
828 neuroinflammation and Alzheimer's disease. *J. Neuroinflammation* **15**, 240 (2018).

829 45. Lampinen, R. *et al.* Neuron-astrocyte transmitophagy is altered in Alzheimer's disease.
830 *Neurobiol. Dis.* **170**, 105753 (2022).

831 46. Han, W. *et al.* Ketogenic β-hydroxybutyrate regulates β-hydroxybutyrylation of TCA cycle-
832 associated enzymes and attenuates disease-associated pathologies in Alzheimer's mice. *Aging Cell* **24**,
833 e14368 (2025).

834 47. Zhu, M. *et al.* Cdc42GAP deficiency contributes to the Alzheimer's disease phenotype. *Brain* **146**,
835 4350–4365 (2023).

836 48. Zhang, Y. & Niu, C. Relation of CDC42, Th1, Th2, and Th17 cells with cognitive function decline in
837 Alzheimer's disease. *Ann. Clin. Transl. Neurol.* **9**, 1428–1436 (2022).

838 49. Allen, M. *et al.* Glutathione S-transferase omega genes in Alzheimer and Parkinson disease risk,
839 age-at-diagnosis and brain gene expression: an association study with mechanistic implications. *Mol.*
840 *Neurodegener.* **7**, 13 (2012).

841 50. Nakayasu, E. S. *et al.* Tutorial: best practices and considerations for mass-spectrometry-based
842 protein biomarker discovery and validation. *Nat. Protoc.* **16**, 3737–3760 (2021).

843 51. Frantzi, M., Bhat, A. & Latosinska, A. Clinical proteomic biomarkers: relevant issues on study
844 design & technical considerations in biomarker development. *Clin. Transl. Med.* **3**, 7 (2014).

845 52. Mischak, H. *et al.* Recommendations for Biomarker Identification and Qualification in Clinical
846 Proteomics. *Sci. Transl. Med.* **2**, 46ps42-46ps42 (2010).

847 53. Ignjatovic, V. *et al.* Mass Spectrometry-Based Plasma Proteomics: Considerations from Sample
848 Collection to Achieving Translational Data. *J. Proteome Res.* **18**, 4085–4097 (2019).

849 54. Lacar, B. *et al.* Identification of Novel Biomarkers for Alzheimer's Disease and Related Dementias
850 Using Unbiased Plasma Proteomics. 2024.01.05.574446 Preprint at
851 https://doi.org/10.1101/2024.01.05.574446 (2024).

852 55. Kirsher, D. Y. *et al.* The Current Landscape of Plasma Proteomics: Technical Advances, Biological
853 Insights, and Biomarker Discovery. 2025.02.14.638375 Preprint at
854 https://doi.org/10.1101/2025.02.14.638375 (2025).

855 56. Zhang, B. *et al.* Nitrogen-modified reduced graphene oxide for serum enrichment of N-glycans
856 and MALDI-TOF MS-based identification of HCC biomarkers. *The Analyst* **150**, 650–660 (2025).

857 57. Mahmoudi, M. The need for improved methodology in protein corona analysis. *Nat. Commun.*
858 **13**, 1–4 (2022).

859 58. Deng, H. *et al.* Monodisperse Magnetic Single-Crystal Ferrite Microspheres. *Angew. Chem. Int.*
860 *Ed.* **44**, 2782–2785 (2005).

861 59. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).

60. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).

61. Tretyakov, K. konstantint/matplotlib-venn. (2024).

62. Grigorev, K. LankyCyril/pyvenn. (2024).

63. BioSenior/ggVolcano: The goal of `ggVolcano` is to help users make a beautiful volcano map more easily, including general volcano plot(`ggvolcano`), gradient color volcano plot(`gradual_volcano`) and GO term volcano plot(`term_volcano`). https://github.com/BioSenior/ggVolcano.

64. Fang, Z., Liu, X. & Peltz, G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinforma. Oxf. Engl.* **39**, btac757 (2023).

65. Pang, Z., Chong, J., Li, S. & Xia, J. MetaboAnalystR 3.0: Toward an Optimized Workflow for Global Metabolomics. *Metabolites* **10**, 186 (2020).

66. Ripley, B. *et al.* MASS: Support Functions and Datasets for Venables and Ripley's MASS. (2024).

67. Vanegas, L. H., Rondón, L. M. & Paula, G. A. glmtoolbox: Set of Tools to Data Analysis using Generalized Linear Models. (2024).

68. hirsch-lab/roc-utils. Digital Health Group (2024).

Acknowledgements

This work was financially supported by Nanomics Biotechnology., Ltd.

Author contributions

H.W. designed the study and drafted the manuscript. Y.W. led product development. Y.H.Z. performed the proteomics data analysis. X.H.O. conducted the sample preparation and mass spectrometry experiments.

Competing interests statement

Y.W., Y.H.Z., X.H.O. and H.W. are employees of Nanomics Biotechnology.

886

887

888

889

890