

<https://doi.org/10.1038/s42004-025-01665-1>

Current landscape of plasma proteomics from technical innovations to biological insights and biomarker discovery



Douglas Y. Kirsher , Shreya Chand, Aron Phong, Bich Nguyen, Balazs G. Szoke & Sara Ahadi

Plasma is a rich source of biomolecules, including proteins, that reflect both health and disease. Due to their key roles in biological processes, proteins hold significant potential as biomarkers, fueling the rise of plasma proteome profiling in recent years. However, comprehensive comparisons of the performance of different plasma proteomics platforms are limited. Our study addresses this gap by conducting a direct comparison of eight platforms, representing both affinity-based and diverse mass spectrometry approaches, and covering over 13,000 proteins. By applying these platforms to the same cohort, we systematically assess their performance, identifying key differences and complementary strengths. Our findings offer valuable insights for researchers, highlighting trade-offs in coverage and their implications for biomarker discovery and clinical applications. This study serves as an essential resource, offering both technical evaluation and biological insights to support the development of novel diagnostics and therapeutics through plasma proteomics.

Plasma is a key component of blood, containing a diverse array of molecules, including proteins, nucleic acids (DNA and RNA), lipids, small molecule metabolites, and electrolytes. Many of these molecules are essential for maintaining various physiological processes within the body¹. Changes in their concentrations can signal or drive underlying health issues, making plasma a valuable resource for discovering biomarkers or therapeutic targets for diseases. One of the main benefits of using plasma for this purpose is the relative ease of collection. Plasma samples are inexpensive to collect and minimally invasive to obtain, and the procedure is well-tolerated by patients, making it ideal for routine use in clinical settings. This simplicity enables longitudinal collection, which provides richer data for biomarker discovery, target identification, and insights into system biology and disease progression.

While plasma is a promising source for clinical diagnostics and biomarker discovery, plasma proteins are particularly significant due to their essential roles in biological processes and disease mechanisms^{2,3}. Given their direct involvement in cellular functions and potential for therapeutic targeting, proteins often emerge as key indicators of health and disease states⁴. Moreover, analyzing plasma protein abundance alongside genetic and phenotypic information can create a comprehensive picture that enhances our understanding of health status and improves disease management.

However, utilizing the plasma proteome as a source of biomarkers presents several challenges^{3,5–7}. The plasma proteome spans a wide dynamic range, with some biomarkers present at very low concentrations, complicating their detection and precise quantification. Additionally, data

interpretation requires careful consideration of confounding factors such as age^{8,9}, sex⁹, BMI^{9,10}, and fasting status, in addition to technical aspects^{6,11} such as storage duration^{12–14}, temperature, and blood collection factors such as time of the day, the use of different anticoagulants¹⁵, and blood processing protocols^{13,14}. Studies have also indicated that while individual proteome profiles tend to remain stable over time, there is significant variability between individuals⁷. These challenges emphasize the need for robust and complementary methodologies, along with a thorough understanding of the data, to effectively leverage the plasma proteome for biomarker and therapeutics development.

Currently, the two primary approaches used to analyze the plasma proteome are affinity-based techniques and mass spectrometry (MS)-based methods. Affinity-based platforms, such as SomaScan, Olink Explore and NULISA use binding probes—aptamers or antibodies—to detect proteins. In contrast, MS methods typically derive protein-level information by measuring proteolytic peptides of proteins through a bottom-up approach. To address the dynamic range challenge, MS workflows may include high-abundance protein depletion¹⁶, peptide pre-fractionation¹⁷, protein precipitation^{16,18}, ion mobility-based separation¹⁶, or protein enrichment via beads¹⁹ or nanoparticles²⁰.

Each method measures distinct characteristics of plasma proteins and has its own unique advantages and disadvantages. SomaLogic and Olink utilize large panels of targeted assays developed for pre-selected proteins. These techniques allow for high-throughput measurements and multiplexing of analytes, enabling the analysis of thousands of proteins from small

sample volumes. However, the specificity of affinity-based assays is variable as it depends on many factors, such as the uniqueness of the binding epitope, the affinity of the probe, and the composition of the sample (i.e., the matrix), among others. SomaLogic assays rely on the unique engagement of a single high-affinity aptamer binder for each target protein. Although aspects of the interaction, such as affinity and low dissociation rate, are exploited for specificity, this can introduce bias based on the matrix in which the protein is measured. Publicly available information on the binding reagents used in SomaScan assays (SOMAmers) and their availability for pull-down assays can help improve our understanding of the specificity of each aptamer²¹. Olink's proximity extension assays mitigate the specificity issue by requiring two different antibodies to bind to the target protein in close proximity²². The more recently introduced NULISA technology, while not in the same league in terms of proteome coverage as SomaScan or Olink assays, has higher sensitivity and lower limit of detection²³. In contrast, MS-based proteomics is less matrix-sensitive and can be performed via both untargeted or targeted approaches. MS-based assays typically measure multiple peptides of one protein and are capable of identifying post-translational modifications (PTMs) and isoforms of proteins, offering unique specificity in protein identification²⁴. However, its limited depth makes it challenging to use for plasma, which has a wide protein concentration range spanning 10 orders of magnitude. Despite this limitation, advancements in mass spectrometry have improved proteome coverage and throughput, allowing for the identification and quantification of a broader range of proteins, including clinically relevant ones present in lower abundance²⁵. Nanoparticle-based enrichment approaches, such as the Seer Proteograph™ XT introduced recently, use surface-modified magnetic nanoparticles to enrich proteins based on physicochemical properties, significantly increasing soluble proteome coverage^{20,26}. Other effective enrichment or depletion strategies include perCA depletion¹⁸, ENRICH/ENRICHplus kits

from PreOmics, and Biognosys's P2 Plasma Enrichment method. These innovations have collectively enabled the detection of previously elusive low-abundance plasma proteins.

This study offers a comprehensive comparison of state-of-the-art plasma proteome analysis techniques by applying each technology to the same cohort of plasma samples. This allowed for a detailed comparison of the platforms, contrasting their advantages and disadvantages. We examined multiple versions of affinity-based platforms (SomaScan 11K, SomaScan 7K, Olink Explore 3072, Olink Explore HT, and NULISA) alongside discovery MS workflows, including the nanoparticle-based Seer Proteograph™ platform and the high-abundance protein depletion-based Biognosys TrueDiscovery™ platform, both utilizing Data-Independent Acquisition (DIA) to generate deep, unbiased MS data. Additionally, we incorporated a targeted MS workflow, SureQuant™ Internal Standard Triggered Parallel Reaction Monitoring, as a “gold standard” MS platform of high reliability and absolute quantification due to the use of internal standards and optimized detection. To the best of our knowledge, this represents the largest number of plasma proteomic technologies compared using a single cohort. By juxtaposing these methods, we aim to provide a detailed technical evaluation and biological insight that will serve as a valuable resource for researchers in plasma proteomics. Understanding the unique characteristics of each platform will assist investigators in selecting the most appropriate method for their research objectives and guide future biomarker discovery efforts.

Results and discussion

Cohort and platforms

In our analysis, we obtained plasma protein profiles from a cohort of 78 individuals, maintaining an equal sex ratio of 1:1 (male to female), consisting of 40 aged (55–65 years old) and 38 young (18–22 years old) individuals. A detailed summary of cohort demographics is listed in Table 1. Plasma samples were collected via plasmapheresis and analyzed using eight proteomic platforms, as shown in Fig. 1. For clarity, we refer to Olink Explore 3072 and Olink Explore HT as Olink 3K and Olink 5K, respectively, for the remainder of this discussion. Similarly, MS-based workflows are designated as MS-Nanoparticle (Seer Proteograph™ XT), MS-HAP Depletion (Biognosys TrueDiscovery™ platform using high-abundance protein depletion), and MS-IS Targeted (Thermo Scientific SureQuant™ Internal Standard Triggered - Parallel Reaction Monitoring using Biognosys PQ500™ Reference Peptides). NULISA refers to the combination of the platform's inflammation and CNS panels (Alamar Bio's NULISAseq™ Inflammation Panel 250 and NULISAseq™ CNS Disease Panel 120). SomaScan 11K includes 10,776 human protein assays targeting 9852 unique proteins corresponding to 9645 distinct UniProt IDs. SomaScan 7K includes 7288 human protein assays targeting 6467 unique proteins, corresponding to 6401 distinct UniProt IDs. Olink 5K and 3K assays target 5416 and 2925 unique human proteins, respectively. NULISA's combined inflammation and CNS panels include 377 assays in total, targeting 325 unique proteins represented by 319 unique UniProt IDs. In our study, the number of proteins identified by MS-based platforms were 5943 in MS-Nanoparticle, 3575 in MS-HAP Depletion and 551 in MS-IS Targeted. Each quantified 68,527, 42,581, and 766 peptides, respectively. A complete list of the proteins identified in this study, including UniProt IDs, can be found in Supplementary Data 1.

In the rest of our paper, we refer to SomaScan and Olink, along with MS-Nanoparticle and MS-HAP Depletion, as “discovery” platforms since they cover large numbers of proteins despite SomaScan and Olink being targeted approaches. In addition, for simplicity in our protein coverage comparisons, we used UniProt IDs instead of protein names to represent unique proteins.

Technical assessments

Across all eight platforms, we identified a total of 13,011 unique plasma proteins, as represented by unique UniProt IDs, in our healthy plasma samples. As illustrated in Fig. 2a, the SomaScan 11K and SomaScan 7K

Table 1 | Summary statistics of study cohort

	Aged	Young	P-val
N	40	38	
Age	58.85 ± 2.9	20.34 ± 1.38	2.43E-14***
Sex			1
Male	20	19	
Female	20	19	
Race			1
African American	10	9	
Caucasian	30	29	
Smoking Status			7.20E-02
Smoker	13	21	
Non Smoker	27	17	
Physiological			
Hematocrit	0.42 ± 0.03	0.43 ± 0.04	6.77E-01
TotalProtein	6.73 ± 0.58	6.70 ± 0.51	9.40E-01
Height (cm)	170.31 ± 10.78	171.78 ± 11.23	5.98E-01
Weight (kg)	87.73 ± 22.62	81.26 ± 28.7	7.19E-02
BMI	30.36 ± 7.8	27.46 ± 8.4	6.96E-02
Systolic BP	131.6 ± 18.35	124.95 ± 17.3	1.08E-01
Diastolic BP	78.38 ± 8.55	72.32 ± 8.42	2.86E-03***
BPM	78.25 ± 9.99	79.89 ± 13.19	3.24E-01

Numbers indicate Mean ± Standard deviation for continuous variables or N for categorical variables. The Wilcoxon Rank Sum test was performed for categorical variables. Chi-Squared test was performed for continuous variables. P-value significance is indicated at *P < 0.05, **P < 0.01, and ***P < 0.005.

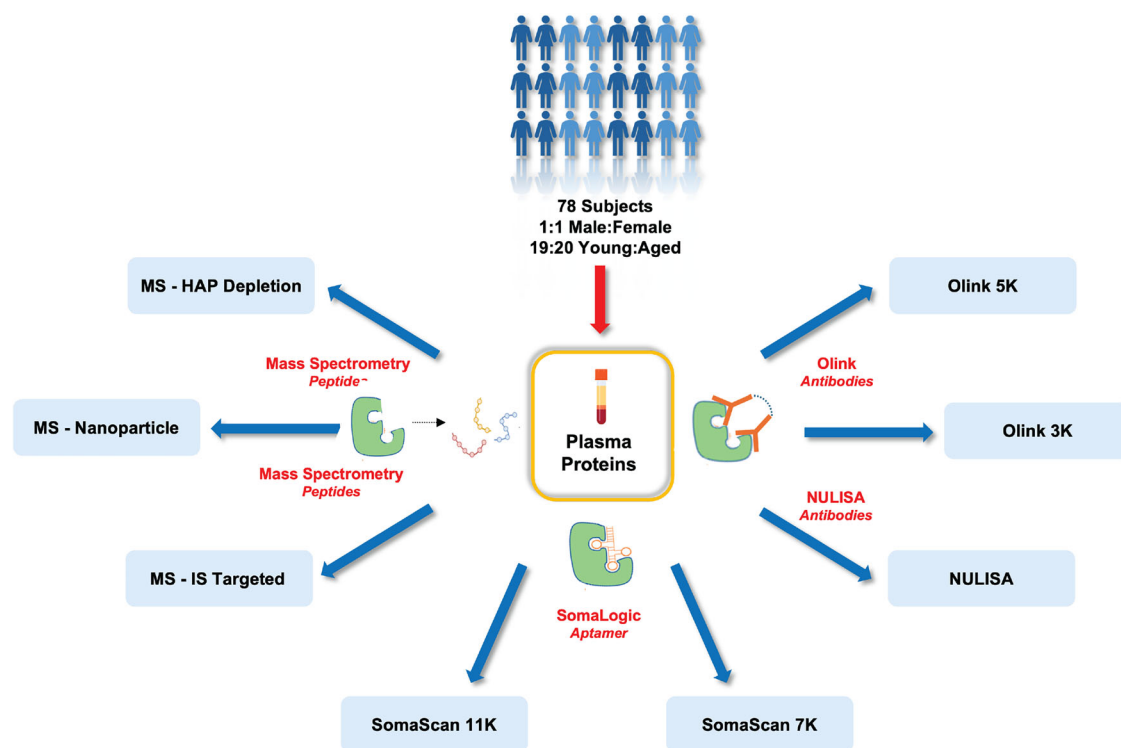


Fig. 1 | Overview of study cohort and proteomic platforms. Plasma samples collected via plasmapheresis from aged ($n = 40$) and young ($n = 38$) subjects were analyzed using eight proteomic platforms: SomaScan 11K, SomaScan 7K, Olink 5K (Olink Explore HT), Olink 3K (Olink Explore 3072), MS-HAP Depletion (high-

abundance protein depletion), MS-Nanoparticle (Seer Proteograph™ XT), MS-IS Targeted (SureQuant™ Internal Standard Triggered - Parallel Reaction Monitoring) and NULISA (NULISAseq™ Inflammation Panel 250 and NULISAseq™ CNS Disease Panel 120).

platforms provided the most comprehensive proteomic coverage, detecting 9645 and 6401 proteins, respectively. MS-Nanoparticle followed by 5943 unique proteins. Each platform contributed a set of exclusive proteins that were not identified by the others. The two SomaLogic platforms contributed the largest number of exclusive proteins, 3600 proteins, followed by the two Olink panels combined with 1227 exclusive protein targets. Across all eight platforms, there was an overlap of only 36 proteins. Excluding NULISA, which has so far focused on inflammation and CNS proteins, the seven remaining platforms with broader protein lists share 259 proteins for which we have absolute quantitation values. This overlap includes MS-IS Targeted data, which was used as a reference to evaluate the other platforms. Excluding MS-IS Targeted and NULISA, the remaining six discovery platforms detected 961 shared proteins. To assess the variability in quantification across platforms, technical replicates were utilized. As shown in Fig. 2b, SomaScan exhibited the highest precision among all platforms, with the lowest technical coefficients of variation (CV). The median CVs for SomaScan 11K and 7K were 5.3% and 5.8%, respectively, indicating that the addition of over 3000 assays to the latest version of SomaScan has not compromised the platform's excellent precision. In contrast, the median CV for Olink 5K was more than twice that of Olink 3K, at 26.8% and 11.4%, respectively. There have been many technical assessments of SomaScan and Olink assays in literature in the past few years^{25,27–30}. Our CV results are consistent with these reports and observations. Median CV of NULISA platform is 6.6% for its combination of inflammation and CNS markers. For the MS-based platforms, the technical CVs of discovery-based approaches were higher than those for SomaScan 11K, SomaScan 7K and Olink 3K, with median CVs of 26.4% for MS-Nanoparticle and 29.8% for MS-HAP Depletion. In contrast to discovery MS platforms, MS-IS Targeted had a median CV of 8.3%, due to the optimized targeted analysis of the proteins and the use of internal controls for each analyte. This same trend was also observed when we restricted the results to analytes that were observed across all platforms (Supplementary Fig. 1a). The median CVs of Olink 5K and 3K were again higher than those of SomaScan 11K and 7K. The MS-based

platforms continued to show higher CVs than affinity platforms, except for MS-IS Targeted that had comparable precision to the affinity platforms (Supplementary Data 1). Filtering to analytes for which at least two-thirds of measurements were above the platform-specific estimated Limit of Detection (eLOD) for affinity-based assays, or to those detected in at least two-thirds of the samples for MS-based platforms, resulted in subtle CV changes for some platforms but more pronounced changes for others (Supplementary Fig. 1b). The most notable change occurred for Olink 5 K, where limiting data to measurements that were above eLOD for healthy plasma improved the CV from 26.8% to 12.4%. It is important to note that this was also accompanied by a 40% reduction in the number of analytes used to calculate the CVs after filtering (Supplementary Data 2).

As expected, the median total CV values across all samples, incorporating both biological and technical variability, were higher than the corresponding technical CVs for all platforms. The gap between technical and biological CV varied widely among the different proteomic approaches. For the two SomaScan platforms, Olink 3K, NULISA and MS-IS Targeted, technical variability was small compared to combined technical plus biological variability, while for the two MS discovery approaches and Olink 5K, the technical variability accounted for a much larger portion of the total observed variability.

For the affinity probe-based platforms, SomaScan and Olink, we carried out a simple linearity assessment by diluting pooled plasma samples 3 and 9 times and checking for the linearity of the measured protein signals. Pearson correlation coefficient (r) was calculated for the dilution data and used to characterize the linearity of each protein assay (Supplementary Fig. 2). Our results showed that 97% of all SomaScan assays (both 7K and 11K) detected normal plasma protein levels in their linear range ($r > 0.9$), while the same measure was found to be 42% for the assays of the Olink 5 K platform. Limiting the analysis to assays in which all 3 dilutions yielded values above eLOD, we found that high proportion of the Olink assays also showed linear behavior (85% of assays with $r > 0.9$).

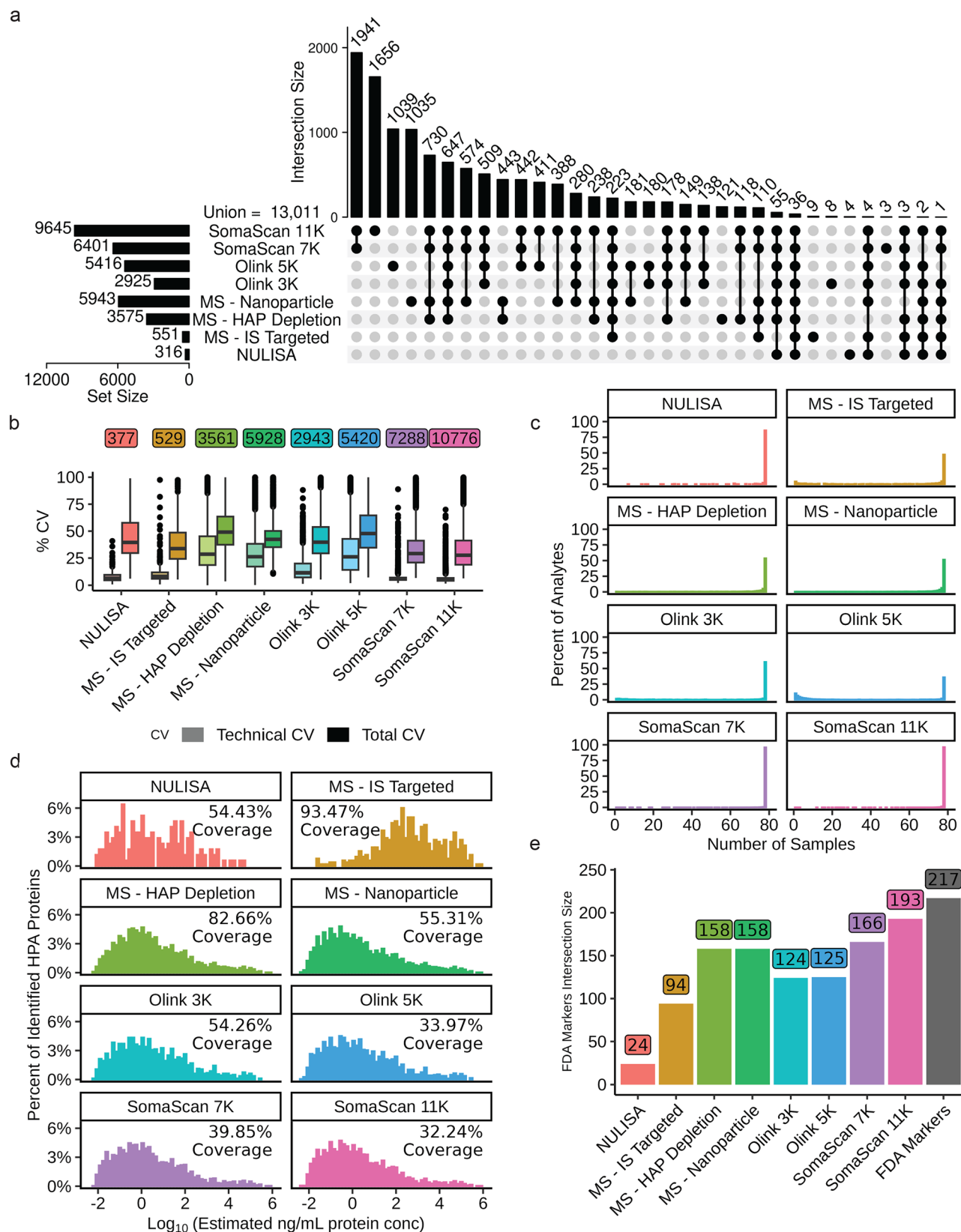


Fig. 2 | Cross-platform comparison of protein quantification and detection metrics. a UpSet plot showing set size and intersection of identified proteins by each platform. 259 shared proteins are quantifiable across all platforms, excluding NULISA. Inclusion of NULISA reduces the number of quantifiable proteins to 36. **b** Technical and total CVs shown for each analyte on each platform. Technical CVs were calculated using set of technical replicates. Total CVs were calculated using all

78 subjects. **c** Percentage of analytes that were detected in each number of samples. Each dataset was filtered based on 2/3rds data completeness from each age group. **d** The detected proteins present in the HPA dataset were binned according to the estimated concentrations from the HPA dataset. The percentage make-up of proteins belonging to these bins are plotted. **e** The number of FDA approved protein biomarkers identified by each platform.

To assess data completeness across different platforms, we examined the number of missing values for each and plotted the distribution of data completeness in Fig. 2c. For MS data, missing values were defined as either a recorded value of 0 or failure to be detected, and for affinity platforms, missing values were those falling below the platform-specific estimated Limit of Detection (eLOD). Our analysis revealed that SomaLogic had the highest data completeness, with SomaScan 11K and 7K showing 96.2% and 95.8% completeness, respectively, across all 78 samples. The Olink 3K platform followed with 60.3% completeness, and MS-HAP Depletion had 53.6% completeness. Notably, the latest version of Olink platform, 5K, had significantly lower data completeness at 35.9%, compared to the previous 3K version at 60.3%. This finding is consistent with the recent literature on comparison between affinity platforms³⁰.

We visualized the abundance of proteins identified in each platform by plotting their distribution based on estimated concentrations of 4,392 plasma proteins from the Human Proteome Atlas (HPA, <https://v21.proteinatlas.org/humanproteome/blood+protein>). It is important to note that the HPA dataset is limited in the number of proteins, and this comparison does not consider a large number of proteins detected by only affinity-based platforms for which there is no absolute concentration estimate in the HPA dataset. As shown in Fig. 2d, all platforms, with the exception of MS-IS Targeted, identified analytes across a wide concentration range (10^5 – 10^{-2} ng/mL) with a strikingly similar concentration distribution for the proteins overlapping between each platform and the HPA. A breakdown of secretome-specific proteins based on the HPA dataset is presented in Supplementary Fig. 3. Deeper proteome coverage together with the increased number of proteins detected by recent MS methodologies, suggest that technological advancements in mass spectrometry, such as depletion or enrichment methods, have significantly expanded the MS coverage of the plasma proteome, especially in the low abundance range. In this sense, we found the MS-Nanoparticle platform identified more proteins than MS-HAP Depletion as it detected almost twice as many proteins with very similar concentration distribution. However, nanoparticle enrichment resulted in a significantly higher platelet contamination than using neat plasma (same plasma samples as enriched plasma, prepared side by side for comparison). For the same plasma samples, the platelet contamination index, defined as the sum of intensities of platelet signature proteins divided by the sum of intensities of non-signature proteins³¹ averaged at 0.0059 (with a standard deviation of 0.0016) for the neat plasma, significantly (p -val < 2.2×10^{-16}) higher at average 0.0240 (with a standard deviation of 0.0011) for nanoparticle-enriched plasma.

To assess the clinical utility of each platform, we examined their coverage of known protein biomarkers in human plasma approved by the U.S. Food and Drug Administration (FDA)³². This is illustrated in Fig. 2e, which shows the distribution of circulating biomarkers across platforms using a previously published list³³. SomaScan platforms demonstrated the highest coverage of these biomarkers. Specifically, 11K and 7K covered 88% and 76% of FDA-approved biomarkers, respectively. This was followed by the discovery MS methods (73%) and Olink (57%). Although MS-IS Targeted had the lowest coverage at 43%, this still represented substantial detection of the 217 FDA-approved markers, considering the relatively small number of total proteins identified using this platform.

Correlations between shared proteins among platforms

With the increase in the depth of proteomics platforms, the overlap of proteins among all increase, and correlation between overlapping proteins can be indicative of the platform's similarities and contrasts in the identification and quantitation of proteins. Figure 3a illustrates the number of overlapping proteins between different platforms. The largest overlap between inter-proteomic technologies is 3978 proteins between SomaScan 11K and MS-Nanoparticle, followed by 3720 proteins between SomaScan 11K and Olink 5K.

To further investigate these intersections, we analyzed the correlation of protein intensities across the various platforms. Spearman correlation values for unique protein assays shared between pairs of platforms are

plotted in Fig. 3b and available in Supplementary Data 3. As expected, two versions of each affinity platform correlated highly with each other, SomaScan 11K and 7K (0.79), followed by Olink 5K and 3K (0.74). When comparing Olink 5K and 3K with other platforms, 5K showed poorer correlation than 3K, while SomaScan 11K and 7K had roughly the same median correlation with other platforms. A similarly high correlation was found between shared proteins of NULISA and the Olink platforms (0.75 for Olink 5K and 0.79 for Olink 3K), which can partly be due to the technical similarities of the two methodologies but also provides mutual specificity validation of the shared assays, considering that they likely use different antibody pairs. Among the MS technologies, MS-IS Targeted demonstrated the strongest correlations with other platforms due to the high specificity of targeted MS with the internal standard approach (Spearman correlation of 0.35, 0.46, 0.46, 0.49, 0.50, 0.56 and 0.62 with MS-Nanoparticle, MS-HAP Depletion, NULISA, SomaScan 7K, SomaScan 11K, Olink 5K and Olink 3K, respectively). Notably, MS-IS Targeted showed a correlation of 0.62 with Olink 3K, which was the highest correlation between any of the affinity-based and MS-based technologies compared in this study. High correlation with MS-IS Targeted, as a reference for assessing other platforms, indicates higher specificity of the measurements in Olink 3K, which has also been reported in literature via *cis*-pQTL validations of affinity assays³⁴. Additionally, all correlations exhibited a bimodal pattern, which was more pronounced in some overlaps than in others. This suggests the presence of two groups of proteins with distinct distributions of high and low correlations. Pairwise correlations between each two platforms are presented in Supplementary Figs. 4 and 5.

To better understand the reasons behind the bimodal distribution of platforms' pairwise correlations, we focused on the 259 proteins shared by seven platforms. We excluded NULISA from this analysis to have a substantial number of overlapped proteins for robust analysis. The lack of correlation between platforms can be related to several factors, most likely differences in identification specificity and quantification precision. We examined the relationship between the technical CVs of these proteins within each platform and their correlation with all other platforms. Proteins were categorized into two groups based on their CV values: less than 20%, and higher than 20%. Figure 3c displays the distribution of Spearman correlation values for each CV category. This plot revealed that proteins with CVs less than 20% tend to have higher correlations with other platforms. Conversely, proteins with CVs higher than 20% showed a shift towards lower correlations with other platforms. This trend is true for all platforms, except for MS-Nanoparticle, where there is no clear distinction in correlation values with technical CVs of proteins. This exception may be due to the unique characteristics of the nanoparticle-based protein enrichment strategy. Factors such as nanoparticle surface chemistry, selective binding affinities, and the specific mass spectrometry workflow can introduce additional variability. This variability can affect protein quantification, and because even the most precise assays show limited correlation, differences between high and low CV groups can be difficult to detect.

We also examined the connection between the technical CV and data completeness for each platform. Figure 3d is a scatterplot showing the percentage of data completeness versus CV, illustrating a clear negative correlation between the two. This negative correlation was more pronounced in MS-HAP Depletion, MS-Nanoparticle and Olink platforms, which also had the highest number of missing values. A similar correlation between CV and data completeness has been reported for Olink 5K, where CV was strongly inversely correlated ($r = -0.77$) with protein detectability, consistent with our findings³⁰.

Differences in proteoform selectivity are also likely to play a significant role in the low correlation observed between platforms. In Fig. 4a, we show correlation data for ApoE and its biologically significant proteoforms, ApoE2, ApoE3, and ApoE4, one of the very few targets for which some proteoform-specific data is available across the platforms. Most platforms measure only total ApoE (non-proteoform specific), but SomaScan 7 K and 11 K also include assays for all three isoforms, and NULISA also measures ApoE4 specifically. Total ApoE correlates strongly across the platforms

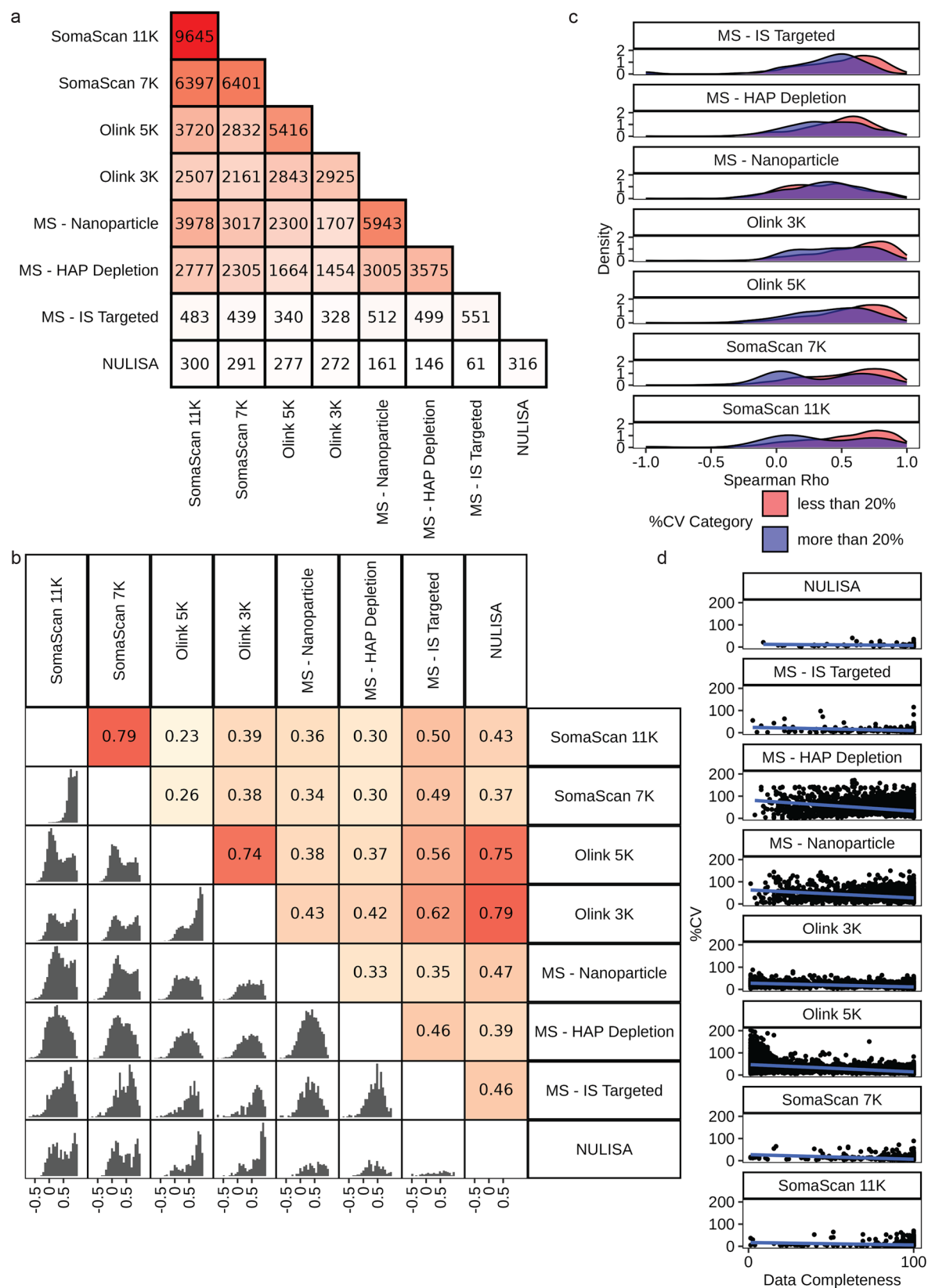


Fig. 3 | Comparative analysis of shared proteins across platforms. a Intersection sizes of detected proteins between each pair of platforms. **b** Histograms showing the distribution of Spearman Rho correlation coefficient calculated on a per-protein basis. The median Spearman Rho correlation coefficient is shown in the upper triangle. **c** Density plot of the Spearman Rho correlation coefficient between a single

platform and all other platforms, for the shared 259 proteins across all platforms, excluding NULISA. The density plot is colored by technical CV of the baseline platform. **d** Technical CV versus Data Completeness for each analyte across each of the platforms.

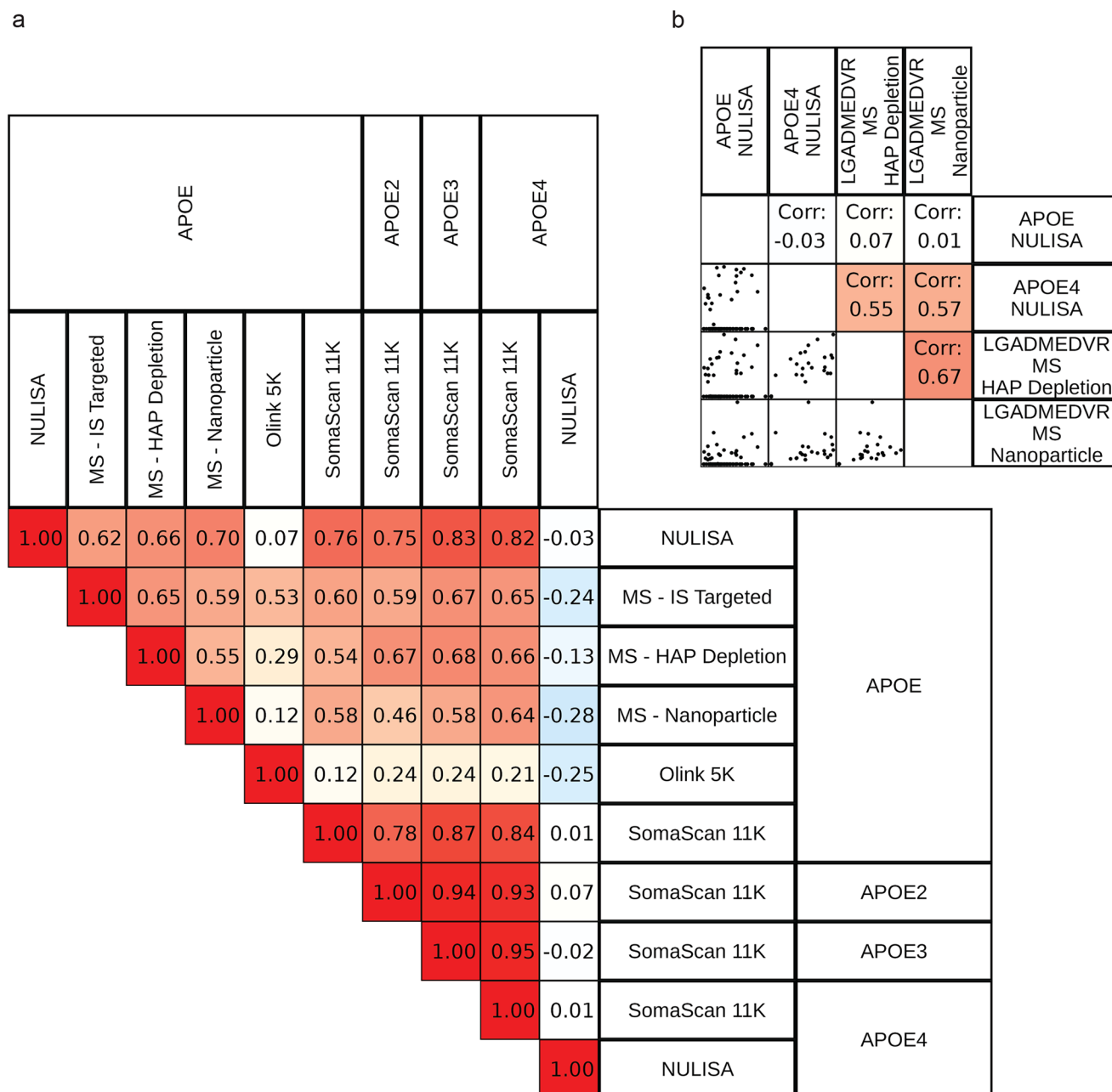


Fig. 4 | Comparison of APOE isoform measurements across diverse proteomic platforms. a Spearman correlation of APOE proteoform measurements across platforms. **b** Scatter plot and Spearman correlation of APOE proteoform measurements between APOE4 specific peptides by MS and total APOE and APOE4 by NULISA.

except for Olink 3 K and 5 K. There is low or no correlation between NULISA ApoE4 and ApoE assays, or with most other platform's total ApoE assays, which is fully consistent with the expectations based on the low prevalence of the ApoE epsilon4 allele (~24%) in the general population³⁵. However, there is also no correlation between NULISA ApoE4 and SomaScan ApoE4 assays, which is not expected. In addition, there is a strong correlation between all four SomaScan ApoE assays, whether supposedly isoform selective or not, suggesting poor isoform selectivity. This is confirmed by information provided on the SomaLogic website (menu.soma-logic.com) for the ApoE4 assay stating that "Binding was observed with Apolipoprotein E, Apolipoprotein E (isoform E2) and Apolipoprotein E (isoform E3) with similar affinities." This example illustrates how differences in proteoform selectivity can affect the correlation between multiple assays targeting the same protein.

To further validate our findings, we searched for and extracted quantification of an ApoE4-specific peptide from our two discovery MS datasets -

MS-HAP Depletion and MS-Nanoparticle. We assessed the correlation of this ApoE4-specific peptide with both the NULISA ApoE4 and total ApoE assays. ApoE4 peptide data of both MS approaches showed convincingly strong correlation with the NULISA ApoE4 assay but no correlation at all with the isoform-agnostic total ApoE assay (Fig. 4b). These results strongly support the isoform specificity of the NULISA ApoE4 assay and highlight the power of the unbiased MS proteomics approaches for discovery at proteoform-level resolution.

We then specifically examined the correlation between the "redundant" SomaScan 11 K assays, i.e. assays utilizing different SOMAmers for measuring the same protein for 834 distinct protein targets. While the median of the Spearman correlation coefficients for the redundant assays is 0.54, the distribution is bimodal with one group of assays showing excellent correlation ($\rho \sim 0.9$), and a second group showing weak correlation ($\rho \sim 0.4$) (Supplementary Fig. 6). While the result discrepancies are most likely related to differences in the selectivity profiles of the redundant SOMAmers,

interpretation of the results is difficult without knowing those selectivity profiles. Identifying the exact target selectivity of each aptamer requires experimental approaches such as pull-down assays and protein structure determination, ultimately at the proteoform level.

Supplementary Fig. 7 presents three example proteins with different correlation patterns across six discovery and the MS-IS Targeted platforms. P11226 (MBL2), for example, is one of the proteins that is highly correlated among all platforms, with Spearman correlation values ranging from 0.86–0.99. For some other proteins, correlations are high between affinity platforms, but not with the discovery MS platforms. P07359 (GP1BA) illustrates this trend, showing high correlation values (0.79–0.93) among affinity platforms and low correlation values between either of the MS discovery platforms and any of the other platforms (0.10–0.24). In this case, MS-IS Targeted is well correlated with affinity platforms, indicating the accuracy of their measurements (0.76–0.81). P01008 (SERPINC1), on the other hand, is an example of a protein whose measurements do not correlate between any of the platforms, even showing poor correlation between two different versions of SomaScan and Olink platforms (0.17–0.50).

Biological relevance and variance analysis of identified proteins

To examine the biology covered by these proteomics platforms, proteins measured by each platform were categorized into protein classes defined by the PANTHER Classification System. Figure 5a shows a heatmap representing the distribution of PANTHER protein classes across different platforms. There is a vast difference in protein class coverage of these platforms, primarily due to the number of covered proteins or targeted/untargeted nature of each platform (Supplementary Data 4). As expected, SomaScan 11 K covered the highest number of protein classes compared to the other platforms. There are 48 protein classes common across all platforms, with at least one protein representation in the class. Globin, cytokine, and complement component are the three highest represented protein classes across all platforms. The Runt transcription factor class was uniquely characterized by SomaScan platforms, whereas the class of mitochondrial carrier proteins was uniquely characterized by SomaScan 11 K and MS-Nanoparticle platforms. While DNA methyltransferase was uniquely characterized by SomaScan and MS-Nanoparticle platforms, DNA photolyase was uniquely characterized by SomaScan 11 K with 100% representation (class size of 2 proteins). Further, SomaScan 11 K was able to characterize seven additional protein classes (DNA ligase, DNA photolyase, adenylate cyclase, centromere DNA-binding protein, deacetylase, mitochondrial carrier protein, tubulin), compared to SomaScan 7K. Olink 5K characterized 13 additional protein classes (DNA ligase, MADS box transcription factor, RNA methyltransferase, adenylate cyclase, amino acid transporter, centromere DNA-binding protein, gene-specific transcriptional regulator, glucosidase, mRNA capping factor, mRNA polyadenylation factor, primase, replication origin binding protein, storage protein) compared to Olink 3K. The transketolase class was identified in all SomaScan and MS-based platforms but not antibody-based platforms, whereas the phosphatase activator class was identified in SomaScan, MS-HAP Depletion and MS-Nanoparticle platforms.

To assess the biological relevance of these proteins, we employed a linear model that incorporated available metadata (age, sex, race, hemato-crit, total protein, BMI, and smoking status) to identify significant markers associated with each factor across the various platforms. Figure 5b illustrates the number of significant markers identified by each platform, based on either p -values (p -val < 0.05) or Benjamini–Hochberg adjusted p -values (p -adj < 0.05). SomaScan 11K, which covered the highest number of proteins and protein classes, identified the greatest number of biologically relevant markers compared to other platforms for age (2170 p -val and 685 p -adj), BMI (5239 p -val and 4040 p -adj), and sex (2726 p -val and 1074 p -adj). MS-Nanoparticle also identified large number of sex related markers and majority of these markers were significant after the p -value adjustment (2427 p -val and 1873 p -adj). Supplementary Fig. 8 highlights the distribution of linear model coefficients categorized by their significance level. After SomaScan platforms, Olink 3K and 5K emerged as the next most

comprehensive assays for identifying markers related to age, sex, and BMI. Supplementary Data 5 and Supplementary Fig. 9 summarize number of shared and unique biological markers for each platform. For all platforms, only a few proteins are shared as age, sex and BMI markers and mostly are unique to each of the biological factors.

Our results indicate that the number of significant markers (p -adj < 0.05) varied across platforms, independent of their overall protein coverage. The distribution of p -values strongly influenced the extent of reduction in significant proteins on multiple comparison adjustment, with some platforms showing disproportionate reductions compared to others.

A variance decomposition analysis (Supplementary Fig. 10a) revealed that each platform captured unique biological factors but to varying degrees. By accounting for known factors of age, sex, race, hematocrit levels, total protein BMI and smoking status in our analysis, we evaluated the percentage of explained variance specific to each platform, as shown in Fig. 5c. A substantial portion of the variance remained unexplained, suggesting the influence of additional biological factors beyond those included in our model (Supplementary Data 6). Two main factors, disease status and genetics, have been reported⁹ to be highly important in explaining the variance of plasma proteins. However, these are absent from our dataset as our cohort consists of healthy individuals, and due to the small number of subjects, a genomics association study was not considered for this cohort. Figure 5d lists the top 20 protein classes that are representative of proteins with more than 50% unexplained variance shared among all discovery platforms. In our cohort with information on biological factors available, NULISA, MS-IS Targeted, Olink 3K, MS-Nanoparticle, SomaScan 7K and 11K explained about the same amount of variance (19.9–22.9%) which is higher than Olink 5K, and MS-HAP depletion platforms (13.8–14.5%). Notably, despite covering fewer proteins, NULISA and MS-IS Targeted explained a significant amount of variance.

Supplementary Fig. 10b highlights examples of proteins, all exhibiting more than 40% explained variance contribution in all discovery platforms. These examples demonstrate similar variance contribution by the different platforms for candidate biomarkers of sex, age and BMI. For instance, Leptin (P41159), a circulating adipokine involved in regulating appetite, food intake, and fat distribution³⁶, is known to be influenced by sex hormones. Sex-related differences in leptin levels are well-documented, with women generally having higher concentrations than men³⁷. In obese individuals, leptin levels are elevated and correlate with BMI and body fat percentage³⁸. This aligns with our data, where biological factors such as sex and BMI explained 38.4% and 40.5% (on average among all platforms) of the variance associated with this protein across all platforms. Another example of sex biomarker is Pregnancy Zone Protein (P20742), where the sex factor alone explained up to 46.3% of the variance in PZP levels. This is consistent with strong sex-related differences in PZP plasma levels, with females having significantly higher levels than males¹¹. Chromogranin-A (P10645) and Insulin-like growth factor-binding protein 2 (P18065), two known markers of age³⁹, explain age variable with 22.9% and 13.3%, respectively, or higher variance in all platforms. As expected, IGFBP2, as an insulin regulatory protein, explains BMI the most⁴⁰, and again this observation is consistent between all platforms. Similar patterns of variance decomposition were also observed across platforms for Neurocan core protein (O14594), a predictive marker of age⁴¹, and L-xylulose reductase (Q7Z4W1) which participates in glucose metabolism.

Age-related markers

Among all biological factors analyzed in the study, we closely examined age-related markers. Figure 6a shows the intersection of significant (p -adj < 0.05) age markers identified across all platforms. Olink 3K identified the highest number of age markers (669), followed by SomaScan 11K (628). Among the MS-based platforms, MS-Nanoparticle had the most age-related markers. In terms of the exclusive age markers identified by a platform, SomaScan 11K identified the greatest number of 282 markers that were not found by any other platform, followed by Olink 3K and Olink 5K, which identified 176 and 99 exclusive markers, respectively. Nine proteins (P07998, P10645,

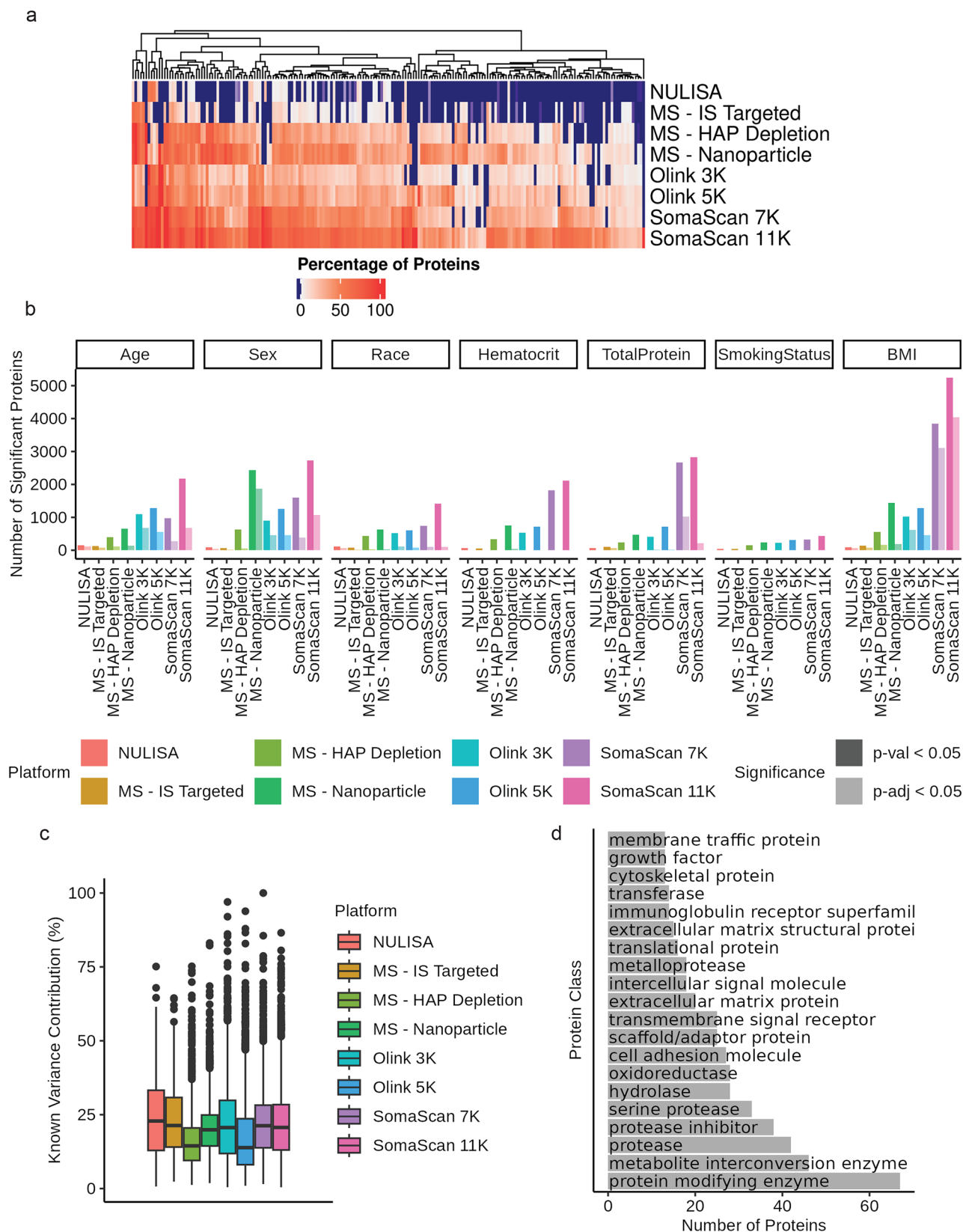


Fig. 5 | Platform-specific protein classes, predictive markers, and variance analysis. a Percent class representation of PANTHER protein classes across the platforms. **b** Number of protein markers identified for each of the seven biological predictors by the different platforms at p -val < 0.05 and p -adj < 0.05 significance.

c Percent variance explained by the biological factors included in the model across the eight platforms. **d** PANTHER protein classes for proteins with >50% unexplained variance across six platforms (MS-IS Targeted and NULISA excluded).

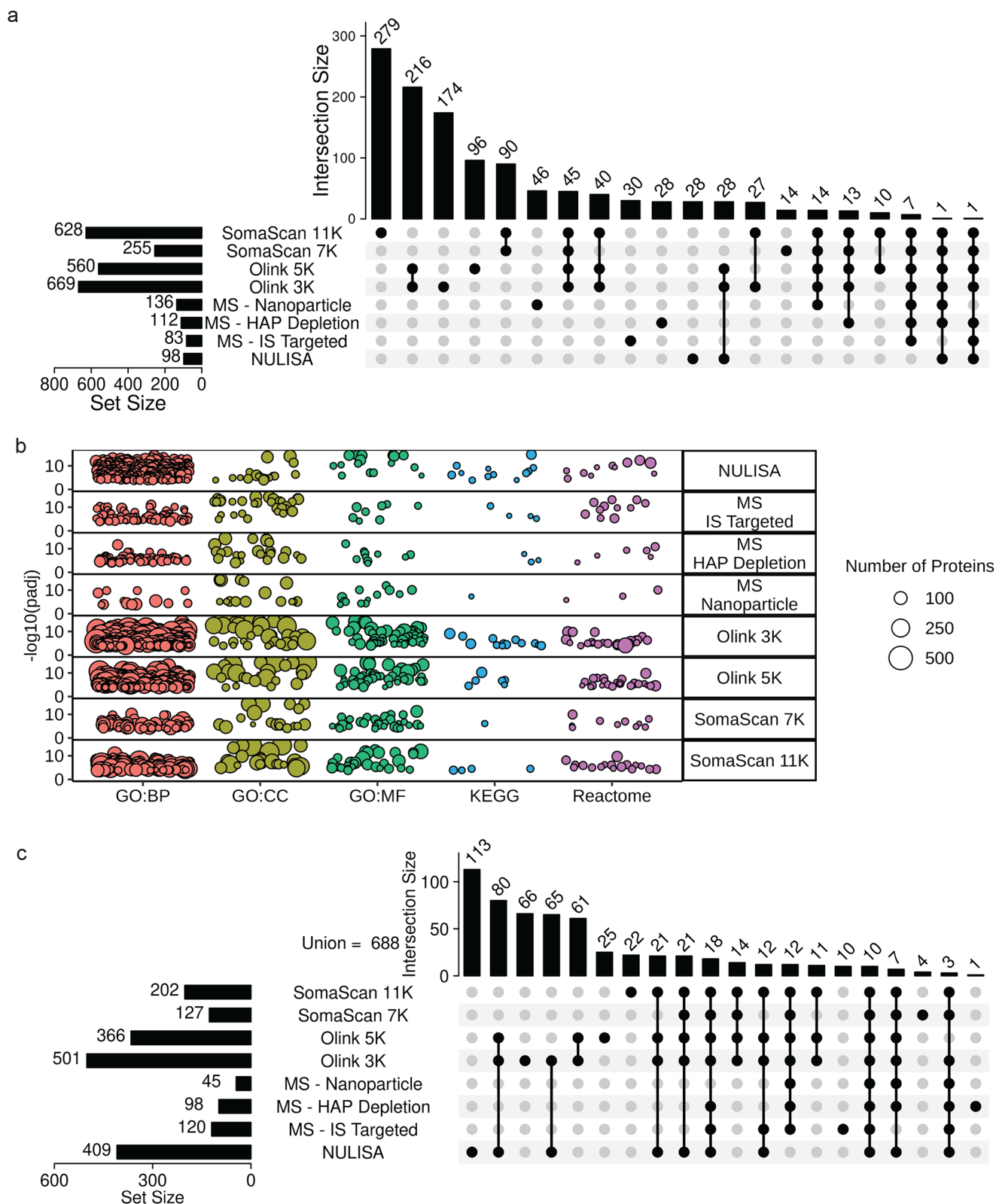


Fig. 6 | Cross-platform analysis of age-associated proteins and pathways. a UpSet plot showing the intersection of significant ($p\text{-adj} < 0.05$) age markers identified by each platform. **b** Results of pathway enrichment analysis using significant age markers from each platform and the union of proteins across all platforms as a background. GO, KEGG and Reactome pathway analysis of significant age markers

from each of the platforms. Each bubble represents a term, and the size of the bubble indicates the number of markers involved. Only terms with significant ($p\text{-adj} < 0.01$) and at least 10 markers are shown. **c** UpSet plot showing the intersection of significant age-related pathways identified across platforms. Only terms with significant ($p\text{-adj} < 0.01$) and at least 10 markers are shown.

P17936, P18065, P49747, Q15113, Q9NQ79, P08254, and P36222) appeared in at least seven of the eight tested platforms. P18065 (IGFBP2, Insulin-like growth factor-binding protein 2) and P17936 (IGFBP3, Insulin-like growth factor-binding protein 3) are well-known markers of aging and have been identified in plasma proteomics studies as associated with aging³⁹. Both proteins are linked to cellular senescence and are likely involved in age-related physiological and pathological processes. P10645 (CHGA, Chromogranin-A), P49747 (COMP, Cartilage oligomeric matrix protein) and Q9NQ79 (CKTAC1, Cartilage acidic protein 1), markers of extreme agers, have been recently reported as age-associated proteins³⁹. Q15113 (PCOLCE, Procollagen C-endopeptidase enhancer 1) has also been reported as an aging protein discovered via quantitative mass spectrometry⁴². P07998 (RNASE1, Ribonuclease pancreatic) is among the top 20 most significant SOMAmer targets associated with chronological age⁴³. P08254 (MMP3, Stromelysin-1) is a secreted protein associated with skin elasticity and aging⁴⁴ and P36222 (CHI3L1, Chitinase-3-like protein 1) has been identified as a protein associated with age waves of 10-year windows of 80 and 90 years old⁴⁵.

Pathway analysis of significant age markers revealed distinct patterns across the various platforms with substantial differences in the number of pathways enriched in GO terms of molecular function, cellular component and biological process. We used the union of detected proteins across all platforms as a background (Fig. 6b). Using this method, we identified 409, 120, 98, 45, 501, 366, 127, and 202 enriched terms in NULISA, MS-IS Targeted, MS-HAP Depletion, MS-Nanoparticle, Olink 3K, Olink 5K, SomaScan 7K and SomaScan 11K respectively at the $p\text{-adj} < 0.01$ level with at least 10 markers present in the pathway, of which 113, 10, 1, none, 66, 25, 4, and 22 were unique to each platform (Fig. 6c). Ten of these terms were commonly enriched across all platforms, mainly involved in cell signaling and communication, cellular structure and tissue and organ function (Supplementary Data 7).

In addition, we compared the aging markers identified by SomaScan 11K and Olink 5K to those detected by the earlier versions of the platforms. Based on our findings, SomaScan 11K and Olink 5K contribute a broader array of age-related markers compared to their predecessors. Specifically, SomaScan 11K identifies 118 aging markers that were not present in SomaScan 7K, while Olink 5K detects 68 unique aging proteins not found in the previous version. Overall, affinity-based platforms contributed the largest number of both unique and shared significant aging markers. In contrast, discovery and targeted MS platforms identified fewer significant markers, but each added valuable proteins that were not covered by the affinity proteomics platforms (Supplementary Fig. 11). Additionally, while we identified markers present in both old and new SomaScan or Olink versions, their statistical significance was sometimes observed in only one version. This discrepancy is likely attributable to differences in protein numbers impacting multiple hypothesis testing corrections and/or variations in the technical characteristics between platform versions, as well as to assay-to-assay variation.

We calculated linear model coefficients of age for all proteins in each platform and examined their correlation between any two platforms. Although correlation values are low (between 0.34 to 0.75) for all shared proteins, correlation between coefficients of shared significant age markers is high (0.66 to 0.95), indicating that when a marker is biologically significant, it is likely to exhibit a strong correlation across multiple platforms (Fig. 7).

To assess the relevance of the age markers identified in this study, we compared the significant markers across all platforms with those reported in the UK Biobank Pharma Proteomics Project (UKB-PPP). The UKB-PPP dataset contains plasma proteomic profiles of 54,219 participants using the Olink Explore 3072 platform⁴⁶. This dataset is widely used in plasma proteomics research. Supplementary Fig. 12 displays the number of markers identified by each platform that were also present in the UKB dataset. Despite the limited size of the cohort included in our analysis, we observed overlap between the markers identified by each platform and those in the

UKB dataset, reinforcing the biological relevance and reliability of the age-related markers identified across the different platforms.

Conclusions

As plasma proteomics gains importance for biomarker discovery, both affinity-based and mass spectrometry (MS) platforms have become indispensable tools. We benchmarked eight technologies in a healthy cohort, spanning over 13,000 proteins, and found that affinity assays detect low-abundance proteins effectively while targeted MS, despite lower coverage, uniquely provides absolute quantification.

SomaScan 11 K achieved the broadest proteome coverage and near-complete detectability of healthy plasma proteins, making it well suited for discovery studies. However, affinity assays target predefined proteins and are not fully unbiased, while MS coverage—though improving—still struggles with robust quantification of low-abundance proteins. Untargeted MS excels at identifying biomarkers beyond known panels, but preparation methods like nanoparticle enrichment may introduce contaminants such as platelets^{47,48}. Pre-analytical factors, including plasma processing, also influence results, underscoring the need for harmonization^{49,50}.

Technical precision and data completeness were key determinants of performance. Platforms such as SomaScan 7K/11K, Olink 3K, NULISA, and MS-IS Targeted had low technical CVs and favorable CV gaps, indicating strong biological signal detection. Olink 5K showed higher CVs comparable to discovery MS, though filtering by detection limits improved precision at the cost of coverage. Across all platforms, assays with more missing data—often low-abundance proteins—had higher variability, complicating interpretation. Interestingly, different SomaScan assays for the same protein sometimes preferentially detected distinct proteoforms, causing sample-dependent differences that likely reflect underlying biological heterogeneity rather than platform limitations.

Although cross-platform correlations were generally modest, consistent with prior reports^{27,29,34,51,52} and reflective of fundamental methodological differences, concordance was higher for biologically significant proteins. Pairwise comparisons of shared proteins showed strong agreement in model coefficients for significant markers ($p < 0.05$), such as age-associated proteins. Cross-referencing with UK Biobank data confirmed overlaps, with Olink 3K and 5K showing the highest match to their age-related markers. Our analysis also captured 82% of UK Biobank aging clock proteins, >90% of SomaScan markers from large cohorts, and 79% of MS-derived centenarian markers^{39,53,54}. This underscores how multiple platforms, each with distinct strengths, collectively provide a more comprehensive view of biomarker landscapes and reveal complementary aspects of marker biology.

In summary, each proteomics platform offers unique and complementary strengths. Selection should be guided by study goals, balancing coverage, precision, and discovery potential. Continued cross-platform evaluations will be essential as technologies advance, enabling improved biomarker discovery and deeper insight into the complexity of the plasma proteome.

Methods

Ethics

Informed consent was obtained from all subjects at Access Biologicals.

Study enrollment

Subjects were selected from among regular plasma donors of Access Biologicals. 78 subjects (39 female and 39 male) were selected in two groups of young (18–22 years old) and aged (55–65 years old). Subjects were considered healthy since they were from qualified plasma donor participants. Information on their medications, medical condition, fasting time, smoking status as well as biometric data (BMI), blood pressure, temperature, hematocrit level, total protein, age, gender and race were collected.

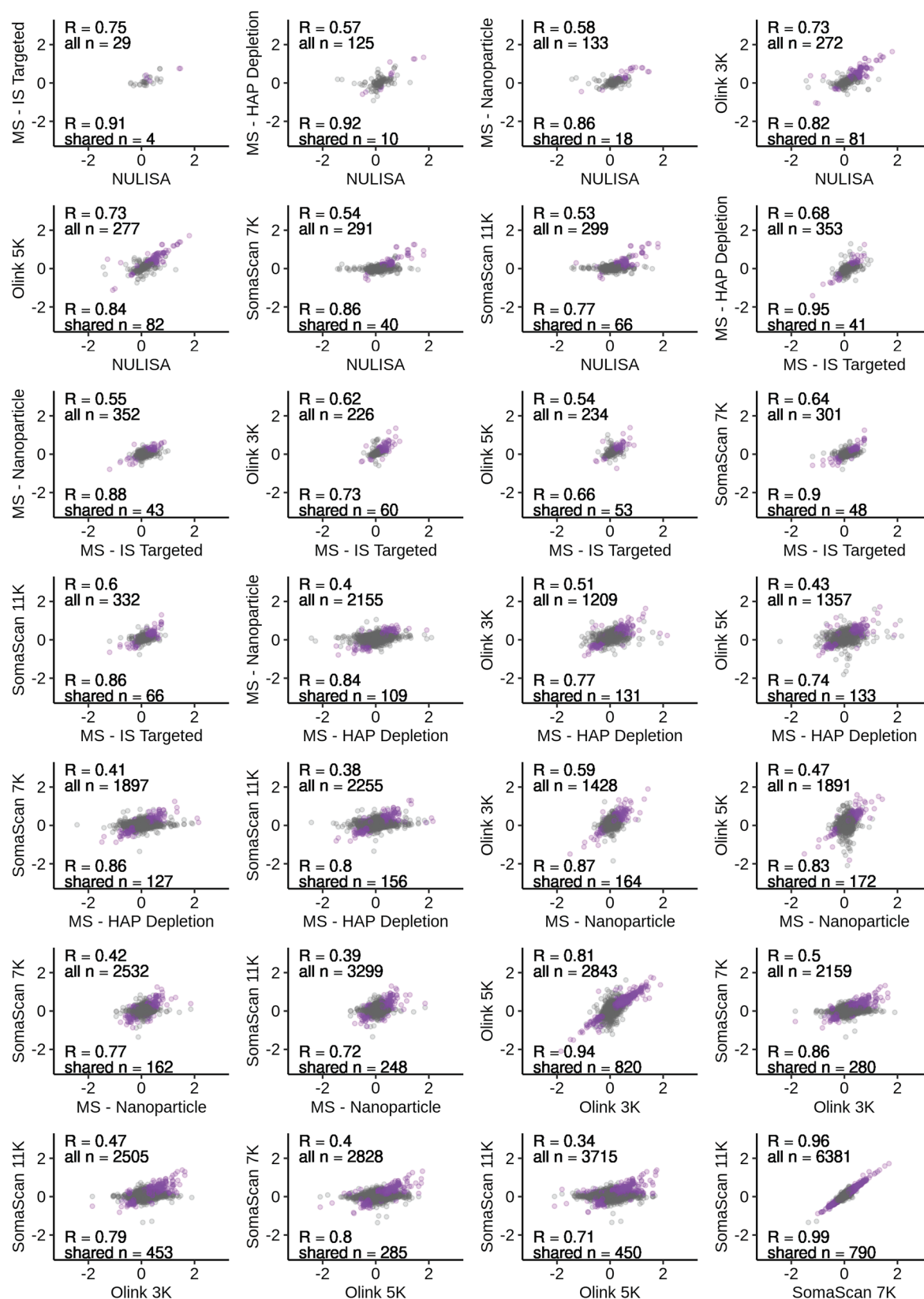


Fig. 7 | Platform agreement in linear model coefficients for shared aging markers. Purple symbols represent shared significant age markers, dark gray symbols are used for all other shared proteins.

Plasma collection

Plasma was collected through plasmapheresis with sodium citrate as anticoagulant from each participant.

Plasma profiling—SomaLogic

Plasma samples were analyzed in two versions of the SomaScan proteomics platform, SomaScan v4.1 (7K) and v5.0 (11K), at Somalogic (Boulder, CO,

USA). The two platform versions share the same assay technology but differ in the number of assays included (7288 vs 10,776 human protein assays, respectively).

The technology is based on Somalogic's proprietary Slow Off-rate Modified Aptamers (SOMAmers) that bind to structural epitopes on proteins. SOMAmers include a fluorophore and a photocleavable biotin moiety in addition to the binding probe single-stranded DNA sequence that contains modified bases to promote protein binding. During the assay, the SOMAmer reagents are pre-immobilized onto streptavidin beads and used to capture target proteins from biological samples. Unbound proteins are washed away, and captured proteins are biotinylated using NHS-biotin. UV light is used to cleave the photosensitive linker, releasing complexes back into solution in the presence of a high concentration of universal polyanionic competitor. Complexes and some free proteins that dissociated are captured onto new streptavidin beads. After washing, SOMAmer reagents are eluted from the beads by denaturing the proteins. The eluate is placed onto a custom Agilent microarray with probes complementary to each SOMAmer reagent for overnight hybridization. Slides are washed and read in an Agilent microarray scanner. The resulting RFU values reflect the amount of target epitope in the initial samples.

SomaScan Assay data used in our studies were first normalized using hybridization controls to mitigate variation that originates from the readout steps, followed by median signal normalization across calibrator replicates within the run. The plate scale factors and calibration scale factors based on the calibrator replicates and external reference values were used to adjust for overall signal intensity differences between runs and SOMAmer reagent-specific assay differences, respectively. Finally, median signal normalization was performed using Adaptive Normalization by Maximum Likelihood (ANML) to scale data to an external reference exclusively based on QC control samples but not based on test samples.

Protein profiling—Olink

Plasma samples were analyzed in two versions of the Olink proteomics platform, Olink® Explore 3072 and Explore HT (Olink Proteomics AB, Uppsala, Sweden) at Olink Analysis Services in Boston (MA, USA). The two platform versions share the same assay technology but differ in the number of assays (3072 vs 5416, respectively) as well as sample requirement and throughput.

The shared underlying technology is based on Proximity Extension Assay (PEA)⁵⁵ coupled with next-generation sequencing (NGS) as readout. In brief, pairs of oligonucleotide-labeled antibody probes against the same protein bind to their target, bringing the complementary oligonucleotides in close proximity and allowing for their hybridization. The addition of a DNA polymerase leads to the extension of the hybridized oligonucleotides, generating a unique protein identification "barcode". Next, library preparation adds sample identification indexes and the required nucleotides for Illumina sequencing. Prior to sequencing using the Illumina® NovaSeq™ 6000/NextSeq™ 550/NextSeq™ 2000, libraries go through a bead-based purification step and the quality is assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The raw output data is quality controlled, normalized and converted into Normalized Protein expression (NPX) values, Olink's proprietary unit of relative abundance. Three internal controls are spiked into every sample and are used to monitor the performance of the three main steps in the protocol: an incubation control, an extension control and an amplification control. In parallel with the samples, the protocol is performed on a set of external controls: Two sample controls, three negative controls and three plate controls (PCs).

Quality control is performed for each sample plate on both the samples (using the spiked internal controls) and the external controls. All assay validation data (detection limits, intra- and inter-assay precision data, predefined values, etc.) are available on the manufacturer's website (www.olink.com).

Olink data used in our studies were normalized by the standard Olink process. This included normalization based on the Extension controls to correct for intra-plate technical variation, followed by normalization based

on the PCs to adjust for batch variability (plate-to-plate variation). As in the case of SomaScan data, no test sample-based normalization was used.

Protein profiling—Biognosys TrueDiscovery™ Pipeline

Plasma samples were shipped frozen by Alkagest and subsequently processed and analyzed at Biognosys. Samples were depleted using a Multiple Affinity Removal Column Human 14 (Agilent) column. Samples were prepared for LC-MS/MS according to Biognosys' SOP which includes reduction, alkylation and digestion to peptides using trypsin (Promega, 1:50 protease to total protein ratio) per sample overnight at 37 °C. Peptides were desalted using a C18 HLB µElution plate (Waters) according to the manufacturer's instructions and dried down using a SpeedVac system. Peptides were resuspended in 1% acetonitrile and 0.1% formic acid and spiked with Biognosys' iRTkit calibration peptides. Peptide concentrations were determined using a UV/VIS Spectrometer at 280 nm (SPECTROstarNano, BMG Labtech).

For DIA LC-MS/MS measurements, 3.5 µg of peptides per sample were injected into an in-house packed reversed phase column on a ThermoScientific™ EASY-nLC™ 1200 nano-liquid chromatography system connected to a ThermoScientific™ Orbitrap™ Exploris480™ mass spectrometer equipped with a NanosprayFlex™ ion source and a FAIMS Pro™ ion mobility device (ThermoScientific™). LC solvents were A: water with 0.1% FA; B: 80% acetonitrile, 0.1% FA in water. The nonlinear LC gradient was 1–50% solvent B in 210 min followed by a column washing step in 90% B for 10 minutes, and a final equilibration step of 1% B for 8 minutes at 60 °C with a flow rate set to 250 nL/min. The FAIMS-DIA method consisted per applied compensation voltage of one full range MS1 scan and 34 DIA segments as adopted from Tognetti et al.¹⁶.

A directDIA™ spectral library was generated by searching the HRM mass spectrometric data using Spectronaut (Biognosys, version 16.2), the false discovery rate on peptide and protein level was set to 1%. A human UniProt. Fasta database (Homo sapiens, 2022-07-01) was used for the search engine, allowing for 2 missed cleavages and variable modifications (N-term acetylation, methionine oxidation, deamidation (NQ) and ammonia-loss). The results were combined with a proprietary deep spectral library for human plasma using Spectronaut.

Raw mass spectrometric data were first converted using the HTRMS Converter (version 15.6, Biognosys) and then analyzed using Spectronaut (Biognosys, version 16.2) with the default settings, but Qvalue filtering with background signal as imputation method was enabled and the hybrid spectral library generated in this study was used. Default settings included peptide and protein level false discovery rate control at 1% and cross-run normalization using global normalization on the median.

For testing of differential protein abundance, protein intensities for each protein were analyzed using a two-sample Student's *t* test. *P*-values were corrected for overall FDR using the *q*-value approach. The following thresholds were applied for candidate identification: *q*-value < 0.05; absolute average log2 ratio > 0.58 (fold-change > 1.5). Distance in heat maps was calculated using the "manhattan" method, and clustering was performed using "ward.D" for both axes. Principal component analysis was conducted in R using pcomp and a modified ggbiplot function for plotting, and partial least squares discriminant analysis was performed using mixOMICS package. Functional analysis was performed using String-db (string-db.org, version 11.5). Topology of candidate proteins was visualized using Protter. General plotting was done in R using ggplot2 package.

Protein profiling—Seer Proteograph

Plasma samples were processed with the Proteograph XT Assay at Seer Inc in Redwood City (CA, USA)^{20,26}. In brief, 240 µL from each sample was transferred to Seer Sample Tubes for processing with the Proteograph XT Assay kit. Plasma proteins were quantitatively captured in nanoparticle (NP) associated protein coronas. Proteins were subsequently denatured, reduced, alkylated and subjected to proteolytic digestion (trypsin and LysC). Peptides were purified and yields were determined using Pierce™

Quantitative Peptide Digest Fluorescent Standards (Thermo Fisher Scientific catalog #23290). Peptides were dried down overnight with a vacuum concentrator and reconstituted with a reconstitution buffer to a concentration of 50 ng/ μ L.

For Data-Independent Acquisition (DIA), 8 μ L of reconstituted peptide mixture from each NP preparation was analyzed, resulting in a constant 400 ng mass MS injection between NP A and NP B samples. Each sample was analyzed with a VanquishTM Neo UHPLC system coupled with a OrbitrapTM AstralTM (Thermo Fisher, Germany) mass spectrometer using a trap-and-elute configuration. First, the peptides were loaded onto an AcclaimTM PepMapTM 100 C18 (0.3 mm ID x 5 mm) trap column and then separated on a 50 cm μ PACTM analytical column (PharmaFluidics, Belgium) at a flow rate of 1 μ L/min using a gradient of 5 – 25% solvent B (0.1% FA, 100% ACN) mixed into solvent A (0.1% FA, 100% water) over 22 min, resulting in a 33 min total run time. The mass spectrometer was operated in DIA mode with MS1 scanning and MS2 precursor isolation windows between 380–980 *m/z*. MS1 scans were performed in the Orbitrap detector at 240,000 R every 0.6 seconds with a 5 ms ion injection time or 500% AGC (500,000 ion) target. Two-hundred fixed window MS2 DIA scans were collected at the Astral detector per cycle with 3 Th precursor isolation windows, 25% normalized collision energy, and 5 ms ion injection times with a 500% (50,000 ion) active gain control maximum. MS2 scans were collected from 150–2000 *m/z*.

Raw mass spectral files were processed using DIA-NN search engine, v1.8.1 using a Homo Sapiens FASTA file containing canonical reviewed and unreviewed proteins. Library-free search was performed in silico based on the input UniProt reference database listed above with Match Between Runs (MBR) enabled. A 1% FDR filtering for identification on peptide/protein group level and a quantification strategy of Robust LC (high precision) was used. DIA-NN search parameters included N-term M excision fixed modification, C carbamidomethylation fixed modification, minimum Peptide Length 7, maximum Peptide Length 30, minimum Precursor Charge 1, maximum Precursor Charge 4, minimum Precursor *m/z* 300, maximum Precursor *m/z* 1800, minimum Fragmentation Ion *m/z* 200, and maximum Fragment Ion *m/z* 1800. Panel Rollup was performed using MaxLFQ treating each nanoparticle as independent species per precursor via R package iq⁵⁶. Precursor-level data were normalized using median intensities from a set of calibration peptides (PepCal) spiked into each MS run to correct for MS drift. The PepCal normalized precursor intensities were rolled up to protein group level using the MaxLFQ algorithm implemented by Precursors observed on multiple nanoparticle wells were kept as separate feature inputs to MaxLFQ. Post-MaxLFQ was applied using the removeBatchEffect function from the R package Limma (<https://rdrr.io/bioc/limma/man/removeBatchEffect.html>)⁵⁷ with plate_id as the batch term to correct for plate-level batch effects for each protein group.

Protein profiling—NULISA

Plasma samples were analyzed using two NULISAseq panels from Alamar Biosciences: the CNS Disease Panel 120 (targeting neurodegenerative disease-related targets as well as inflammatory and immune response cytokines/chemokines) and the Inflammation Panel 250 (targeting primarily inflammatory and immune response cytokines/chemokines). Both panels utilize the same assay technology but differ in the number and nature of targets. NULISAseq assays were performed at Alamar Biosciences as previously described²³. Briefly, -80°C stored samples were thawed on ice and centrifuged at 2,200 g for 10 minutes. 25 μ L of supernatant from each sample was plated in 96-well plates and analyzed. The ARGOTM HT platform was used for the NULISAseq workflow, which included: immunocomplex formation with DNA-barcoded capture and detection antibodies; capture and washing of immunocomplexes on paramagnetic oligo-dT beads; release of immunocomplexes into a low-salt buffer followed by capture and washing on streptavidin beads; and ligation of the proximal ends of DNA strands to generate DNA reporter molecules containing target- and sample-

specific barcodes. These reporter molecules were then pooled, amplified by PCR, purified, and sequenced on an Illumina NextSeq 2000.

Sequencing data were processed using the NULISAseq algorithm (Alamar Biosciences). Sample- (SMI) and target-specific (TMI) barcodes were quantified, allowing for up to two base mismatches or one indel and one mismatch. Intraplate normalization was performed by dividing target counts by the well's internal control counts. Interplate normalization used interplate control (IPC) wells, dividing counts by target-specific medians of the three IPC wells per plate. Finally, data were rescaled, a value of 1 was added, and the data were log2 transformed to generate NULISA Protein Quantification (NPQ) units for downstream statistical analysis.

Note: Most platform-specific analyses, including data processing, were performed by the technology providers using their standard pipelines. As such, comparisons reflect the full analytical workflow, including both experimental procedures and data processing steps.

Protein profiling—SureQuant

SureQuant Analysis was performed on site at Alkahest Inc. Undepleted human plasma was processed using the Thermo ScientificTM EasyPepTM MS Sample Prep Kit. A set of 804 SIL peptides from PQ500 (Biognosys PN# Ki-3019-96) was spiked at approximately 80 fmol (median value) into 1 μ g of plasma tryptic digest. The volume corresponding to 1 μ g of the digest was injected for LC-MS/MS analyses on an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific) coupled with a Vanquish Neo UHPLC system (Thermo Fisher Scientific).

Chromatographic separations were performed using a 0.5 cm C18 PepMapTM Neo Trap Cartridge column (5 μ m, 100 Å, 300 μ m inner diameter; Thermo Fisher Scientific Cat# 174500) and a 15 cm C18 EASY-SprayTM HPLC column (2 μ m, 100 Å, 150 μ m inner diameter; Thermo Fisher Scientific Cat# ES906). Peptides were separated over a gradient from 2% to 31.5% acetonitrile (80%) with 0.1% formic acid over 60 minutes.

To implement this method, the custom SureQuant acquisition template available in the Thermo Orbitrap Exploris was utilized. In the 'watch' mode of the SureQuant method, the MS1 resolution was set to 120k to monitor the predefined optimal precursor ions of the internal standard (IS), which were included in the targeted mass filter. This was followed by the recognition of heavy peptide precursors and fragments from the list at a low resolution of 7.5k, with HCD collision energy set to 27% and a maximum injection time of 10 ms. The detection of the IS triggered the 'quant' mode, which required at least five product ions to initiate an offset scan at a resolution of 60k, with HCD collision energy at 27%, a normalized AGC target of 1000%, and a maximum injection time of 116 ms in profile mode. Data analysis from SureQuant acquisitions was performed using SpectroDiveTM (Biognosys).

Dilution linearity

Two pooled plasma samples, one from young (18–22 year-old, pool of 20) and one from aged (55–65-year old, pool of 20) healthy males, were diluted 3x and 9x with PBS. At least three technical replicates of each of the undiluted samples, their two dilutions, and the buffer were tested by the proteomic platforms. Pearson correlation coefficient (*r*) of the resulting protein level versus sample relative concentration data was calculated for each measured protein and used to characterize the linearity of the platforms in the range of healthy plasma protein levels. Estimated Limit of Detection (eLOD) was calculated as mean + 3*SD of buffer controls.

Data analysis

Multi-UniProt Accession ID analytes were addressed by sorting the UniProt Accession IDs of these analytes in order to maintain consistency between platforms.

For the affinity-based platforms, eLOD values for each protein were calculated depending on the platform. For SomaScan 7K and 11 K platforms, eLOD values were calculated from buffer controls using mean

+ 3.3*SD. For Olink 3K, eLOD values were provided with the data. For Olink 5K, eLOD values were calculated using the `olink_lod` function from the OlinkAnalyze 3.8.2 R⁵⁸ package using the fixed LOD method.

Data completeness for each analyte was calculated across all platforms, defined by the percentage of analytes that were detected in each sample. For MS-based platforms, raw measurements were used. For affinity-based platforms, pre and post eLOD filtered values were used.

Technical and total CVs were calculated for each analyte across all platforms. For the SomaScan 7K assay dataset, the technical CV was calculated using duplicates of eight samples. For the SomaScan 11K assay dataset, the technical CV was calculated using four samples with four technical replicates each. This included one aged female bridging sample, one young female bridging sample, a bridging sample pool of 20 aged male subjects, and a bridging sample pool of 20 young male subjects. For Olink 3K, depleted Mass Spectrometry data, and SureQuant data, a technical CV was calculated using three samples tested in triplicate. For Olink 5K and Seer, a technical CV was calculated using two pools of samples, aged male and young male, each tested in six replicates. Total CV percentages were calculated for each analyte across all platforms, using 78 shared samples. Unique proteins were determined using unique UniProt Accession IDs. For platforms in which a single UniProt Accession ID was represented by multiple analytes, all analytes were kept separate in the analysis until UniProt Accession IDs were necessary to work with. To compare platforms to one another, we calculated the correlation for analytes between platforms using Spearman's rank correlation. For instances in which there was a one-to-many or many-to-many relationship of analytes representing the same UniProt Accession IDs, correlations were calculated using all combinations of analytes. For SomaScan 7K and 11K, due to the analytes of 11K being a superset of the 7K analytes, we performed a one-to-one comparison of the same analytes.

To assess platelet contamination for the MS—Nanoparticle technology, a platelet index was computed for paired neat and nanoparticle-enriched plasma samples. The platelet contamination index was computed as the sum of intensities of previously identified platelet contaminant proteins, divided by the sum of intensities of the non-platelet contaminant proteins³¹.

$$PI = \frac{\sum \text{Platelet protein intensity}}{\sum \text{Non-Platelet protein intensity}}$$

Multivariate linear modeling was performed in R v4.3.2 with `lm` according to the formula $\text{Log}_2(\text{Protein Measurement}) \sim \text{Age} + \text{Gender} + \text{Race} + \text{Hematocrit} + \text{TotalProtein} + \text{SmokingStatus} + \text{BMI}$. The protein measurement used was platform dependent. The mass spectrometry datasets were filtered for at least 2/3rds data completeness within each age group. The Olink datasets were left unfiltered. The SomaScan datasets were filtered based on eLOD values. Pathway enrichment was performed using `topGO` v2.54.0, `ReactomePA` v1.46.0 and `clusterProfiler` v4.10.1 using proteins identified with a *p*-adj less than 0.05. The total set of detected proteins for each platform was used as a background. Pathways with at least 10 markers present and *p*-adj less than 0.01 were highlighted. Plots were generated using the R package `ggplot2` v3.5.1, Heatmaps and UpSet plots were generated using the R package `ComplexHeatmap` v2.18.0.

Isoform analysis of Apolipoprotein E

Apolipoprotein E (ApoE) isoforms were analyzed in the respective MS-HAP Depletion (Biognosys) and MS-Nanoparticle (Seer) datasets. Raw MS-Nanoparticle mass spectral files were processed using DIA-NN search engine, using a Homo Sapiens FASTA file containing canonical reviewed proteins as previously described and including ApoE isoform sequences. Raw MS-HAP Depletion mass spectral files were processed using Spectronaut as previously described with Homo Sapiens FASTA file containing canonical reviewed proteins, with the addition of ApoE isoform sequences. Spearman correlation coefficient of the resulting MS data versus

NULISA proteomic data was calculated for each measured ApoE isoform-specific peptide sequence and used to characterize the relationship of MS quantitation versus NULISA quantitation of total ApoE and the ApoE4 isoform.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw and search results of datasets of MS-Nanoparticle, MS-HAP Depletion, MS-IS Targeted have been deposited to the PRIDE repository (<https://www.ebi.ac.uk/pride/archive/>) with the data set identifier PXD067119, PXD067064, PXD067061, respectively. A map of the raw data and search files is included in each PRIDE submission. Affinity proteomics data is available upon request from corresponding author. Data on Subjects is available in Supplemental Data 10, limited to age and sex due to consent restrictions.

Received: 2 April 2025; Accepted: 18 August 2025;

Published online: 25 September 2025

References

- Yurkovich, J. T. & Hood, L. Blood is a window into health and disease. *Clin. Chem.* **65**, 1204–1206 (2019).
- Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**, 845–867 (2002).
- Geyer, P. E., Holdt, L. M., Teupser, D. & Mann, M. Revisiting biomarker discovery by plasma proteomics. *Mol. Syst. Biol.* **13**, 942 (2017).
- Birhanu, A. G. Mass spectrometry-based proteomics as an emerging tool in clinical laboratories. *Clin. Proteomics* **20**, <https://doi.org/10.1186/s12014-023-09424-x> (2023).
- Deutsch, E. W. et al. Advances and utility of the human plasma proteome. *J. Proteome Res.* **20**, 5241–5263 (2021).
- Ignjatovic, V. et al. Mass spectrometry-based plasma proteomics: considerations from sample collection to achieving translational data. *J. Proteome Res.* **18**, 4085–4097 (2019).
- Geyer, P. E. et al. The Circulating Proteome—technological developments, current challenges, and future trends. *J. Proteome Res.* <https://doi.org/10.1021/acs.jproteome.4c00586> (2024).
- Lehallier, B. et al. Undulating changes in human plasma proteome profiles across the lifespan. *Nat. Med.* **25**, 1843–1850 (2019).
- Carrasco-Zanini, J. et al. Mapping biological influences on the human plasma proteome beyond the genome. *Nat. Metab.* <https://doi.org/10.1038/s42255-024-01133-5> (2024).
- Zaghlool, S. B. et al. Revealing the role of the human blood plasma proteome in obesity using genetic drivers. *Nat. Commun.* **12**, 1279 (2021).
- Geyer, P. E. et al. Plasma proteome profiling to assess human health and disease. *Cell Syst.* **2**, 185–195 (2016).
- Zander, J. et al. Effect of biobanking conditions on short-term stability of biomarkers in human serum and plasma. *Clin. Chem. Lab Med.* **52**, 629–639 (2014).
- Daniels, J. R. et al. Stability of the human plasma proteome to pre-analytical variability as assessed by an aptamer-based approach. *J. Proteome Res.* **18**, 3661–3670 (2019).
- Cao, Z. et al. An integrated analysis of metabolites, peptides, and inflammation biomarkers for assessment of preanalytical variability of human plasma. *J. Proteome Res.* **18**, 2411–2421 (2019).
- Lan, J. et al. Systematic evaluation of the use of human plasma and serum for mass-spectrometry-based shotgun proteomics. *J. Proteome Res.* **17**, 1426–1435 (2018).
- Tognetti, M. et al. Biomarker candidates for tumors identified from deep-profiled plasma stem predominantly from the low abundant area. *J. Proteome Res.* **21**, 1718–1735 (2022).

17. Keshishian, H. et al. Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry. *Nat. Protoc.* **12**, 1683–1701 (2017).
18. Viode, A. et al. A Simple, Time-and Cost-Effective, High-Throughput Depletion Strategy for Deep Plasma Proteomics. *Sci. Adv.* **9**, <https://doi.org/10.1126/sciadv.adf9717> (2023).
19. Soni, R. K. Frontiers in plasma proteome profiling platforms: innovations and applications. *Clin. Proteomics* **21**, 43 (2024).
20. Blume, J. E. et al. Rapid, deep and precise profiling of the plasma proteome with multi-nanoparticle protein corona. *Nat. Commun.* **11**, 3662 (2020).
21. Sun, B. B. et al. Genomic Atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
22. Lundberg, M. et al. Multiplexed homogeneous proximity ligation assays for high-throughput protein biomarker research in serological material. *Mol. Cell Proteomics* **10**, <https://doi.org/10.1074/mcp.M110.004978> (2011).
23. Feng, W. et al. NULISA: a proteomic liquid biopsy platform with attomolar sensitivity and high multiplexing. *Nat. Commun.* **14**, 7238 (2023).
24. Van Eyk, J. E. & Sobhani, K. Precision medicine: establishing proteomic assessment criteria from discovery to clinical diagnostics. *Circulation* **138**, 2172–2174 (2018).
25. Suhre, K. et al. Nanoparticle enrichment mass-spectrometry proteomics identifies protein-altering variants for precise pQTL mapping. *Nat. Commun.* **15**, 989 (2024).
26. Ferdosi, S. et al. Engineered nanoparticles enable deep proteomics studies at scale by leveraging tunable nano-bio interactions. <https://doi.org/10.1073/pnas> (2022).
27. Katz, D. H. et al. Proteomic profiling platforms head to head: leveraging genetics and clinical traits to compare aptamer-and antibody-based methods. *Sci. Adv.* **8**, eabm5164 (2022).
28. Raffield, L. M. et al. Comparison of proteomic assessment methods in multiple cohort studies. *Proteomics* **20**, e1900278 (2020).
29. Candia, J. et al. Assessment of variability in the SOMAscan Assay. *Sci. Rep.* **7**, 14248 (2017).
30. Rooney, M. R. et al. Plasma proteomic comparisons change as coverage expands for SomaLogic and Olink. <https://doi.org/10.1101/2024.07.11.24310161> (2024).
31. Geyer, P. E. et al. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol. Med.* **11**, e10427 (2019).
32. Hendricks, N. G. et al. An inflection point in high-throughput proteomics with Orbitrap Astral: analysis of biofluids, cells, and tissues. <https://doi.org/10.1101/2024.04.26.591396> (2024).
33. 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. *J. Proteome Res.* **20**, 2105–2115 (2021).
34. Eldjarn, G. H. et al. Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature* **622**, 348–358 (2023).
35. Wang, Y. Y. et al. The proportion of APOE4 carriers among non-demented individuals: a pooled analysis of 389,000 community-dwellers. *J. Alzheimer's Dis.* **81**, 1331–1339 (2021).
36. Clemente-Suárez, V. J. et al. The role of adipokines in health and disease. *Biomedicines* **11**, <https://doi.org/10.3390/biomedicines11051290> (2023).
37. Hellström, L., Wahrenberg, H., Hruska, K., Reynisdottir, S. & Arner, P. Mechanisms behind gender differences in circulating leptin levels. *J. Intern. Med.* **247**, 457–462 (2000).
38. Shah, N. R. & Braverman, E. R. Measuring adiposity in patients: the utility of body mass index (BMI), percent body fat, and leptin. *PLoS ONE* **7**, e33308 (2012).
39. Reed, E. R. et al. Cross-platform proteomics signatures of extreme old age. <https://doi.org/10.1101/2024.04.10.588876> (2024).
40. Rauzier, C., Lamarche, B., Tremblay, A. J., Couture, P. & Picard, F. Associations between insulin-like growth factor binding protein-2 and lipoprotein kinetics in men. *J. Lipid Res.* **63**, 100269 (2022).
41. Llobet, M. O., Johansson, Å., Gyllenstein, U., Allen, M. & Enroth, S. Forensic prediction of sex, age, height, body mass index, hip-to-waist ratio, smoking status and lipid lowering drugs using epigenetic markers and plasma proteins. *Forensic Sci. Int. Genet.* **65**, 102871 (2023).
42. Wang, H. et al. Quantitative iTRAQ-based proteomic analysis of differentially expressed proteins in aging in human and monkey. *BMC Genomics* **20**, 725 (2019).
43. Sathyan, S. et al. Plasma proteomic profile of age, health span, and all-cause mortality in older adults. *Aging Cell* **19**, e13250 (2020).
44. Naval, J., Alonso, V. & Herranz, M. A. Genetic polymorphisms and skin aging: the identification of population genotypic groups holds potential for personalized treatments. *Clin. Cosmet. Investig. Dermatol.* **7**, 207–214 (2014).
45. Siino, V. et al. Plasma proteome profiling of healthy individuals across the life span in a Sicilian cohort with long-lived individuals. *Aging Cell* **21**, e13684 (2022).
46. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
47. Gao, H. et al. Systematic evaluation of blood contamination in nanoparticle-based plasma proteomics. <https://doi.org/10.1101/2025.04.26.650757> (2025).
48. Korff, K. et al. Pre-analytical drivers of bias in bead-enriched plasma proteomics. <https://doi.org/10.1101/2025.05.07.652495> (2025).
49. Beimers, W. F. et al. Technical evaluation of plasma proteomics technologies. *J. Proteome Res.* **24**, 3074–3087 (2025).
50. Sun, Z. et al. Deep blood proteomics identifies over 12,000 proteins providing valuable information about the state of the human body. <https://doi.org/10.1101/2025.05.15.654311> (2025).
51. Rooney, M. R. et al. Comparison of proteomic measurements across platforms in the Atherosclerosis Risk in Communities (ARIC) study. *Clin. Chem.* **69**, 68–79 (2023).
52. Candia, J. et al. Variability of 7K and 11K SomaScan plasma proteomics assays. *J. Proteome Res.* **23**, 5531–5539 (2024).
53. Argentieri, M. A. et al. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat. Med.* <https://doi.org/10.1038/s41591-024-03164-7> (2024).
54. Coenen, L., Lehallier, B., de Vries, H. E. & Middeldorp, J. Markers of aging: unsupervised integrated analyses of the human plasma proteome. *Front. Aging* **4**, 1112109 (2023).
55. Assarsson, E. et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE* **9**, e95192 (2014).
56. Pham, T. V., Henneman, A. A. & Jimenez, C. R. Iq: An R package to estimate relative protein abundances from ion quantification in DIA-MS-based proteomics. *Bioinformatics* **36**, 2611–2613 (2020).
57. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
58. Nevola K SMGJFSCPPZBSMDKKACLOCK. OlinkAnalyze: Facilitate analysis of proteomic data from Olink. <https://github.com/Olink-Proteomics/OlinkRPackage>.

Acknowledgements

The authors are grateful to Don Macagba for his help in organizing files and facilitating data submission. This study was funded by Alkahest Inc., a Grifols company.

Author contributions

D.K., A.P., B.S., S.A. analysis of data, S.C., B.S., S.A. writing the paper, B.N. data acquisition. All authors critically revised, edited, and approved the final manuscript.

Competing interests

Authors of this manuscript are full-time employees of Alkahest Inc.

Additional information

Supplementary information The online version contains supplementary material available at

<https://doi.org/10.1038/s42004-025-01665-1>.

Correspondence and requests for materials should be addressed to Sara Ahadi.

Peer review information *Communications Chemistry* thanks Rajesh Soni, Rui Vitorino, and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025