Journal of proteome research

pubs.acs.org/jpr

Article

Quantification of 782 Plasma Peptides by Multiplexed Targeted Proteomics

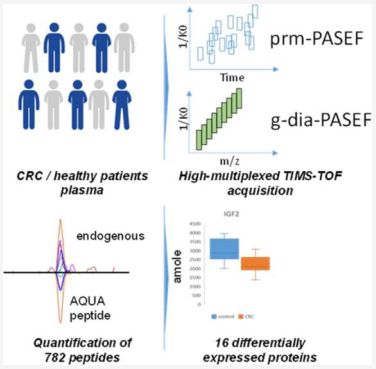
Antoine Lesur, François Bernardin, Eric Koncina, Elisabeth Letellier, Gary Kruppa, Pierre-Olivier Schmit, and Gunnar Dittmar*

Cite This: *J. Proteome Res.* 2023, 22, 1630−1638

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Blood analysis is one of the foundations of clinical diagnostics. In recent years, the analysis of proteins in blood samples by mass spectrometry has taken a jump forward in terms of sensitivity and the number of identified proteins. The recent development of parallel reaction monitoring with parallel accumulation and serial fragmentation (prm-PASEF) combines ion mobility as an additional separation dimension. This increases the proteome coverage while allowing the use of shorter chromatographic gradients. To demonstrate the method's full potential, we used an isotope-labeled synthetic peptide mix of 782 peptides, derived from 579 plasma proteins, spiked into blood plasma samples with a prm-PASEF measurement allowing the quantification of 565 plasma proteins by targeted proteomics. As a less time-consuming alternative to the prm-PASEF method, we describe guided data independent acquisition (dia)-PASEF (g-dia-PASEF) and compare its application to prm-PASEF for measuring blood plasma. To demonstrate both methods' performance in clinical samples, 20 patient plasma samples from a colorectal cancer (CRC) cohort were analyzed. The analysis identified 14 differentially regulated proteins between the CRC patient and control individual plasma samples. This shows the technique's potential for the rapid and unbiased screening of blood proteins, abolishing the need for the preselection of potential biomarker proteins.



KEYWORDS: *targeted proteomics, PASEF, PRM, DIA, SIL peptides, colorectal cancer, CRC, plasma, biomarker*

INTRODUCTION

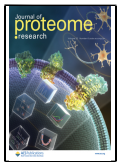
Blood-based diagnostics is one of the pillars of modern medicine. It allows the diagnosis of diseases at an early or asymptomatic stage, predicting a treatment outcome or supporting decisions for the most appropriate disease treatment. Blood components, including the serum and the plasma, can be collected with minimal patient discomfort and are readily available for prospective and retrospective studies. Although blood-based assays have been used for a long time, many diagnostics still rely on measuring a single protein (e.g., immunoassay) or the distribution of cells contained within the blood sample (e.g., cytometry).

Recently, several studies showed that protein panels provide a better diagnostic or predictive power than single protein markers.[1−5] As many described biomarkers are derived from the direct analysis of the affected tissue or tumor material, they usually do not directly translate into valid biomarkers in blood samples, making further studies and method development necessary.[1,6,7]

As a multiplexed technique, mass spectrometry plays an integrated role in discovering and verifying biomarker candidates and is an alternative to antibody-based analysis techniques.[8−10] While shotgun proteomics measurements are excellent tools for the initial discovery of regulated proteins, the inherent risk of not detecting proteins in samples of a cohort due to the random selection of ions for fragmentation is mitigated by switching to targeted acquisition. In targeted proteomics, the monitored set of peptides is set and is measured in every sample of a cohort. Focusing on a selected set of peptides minimizes the risk of missing acquisitions and increases the sensitivity of the measurement.[11−13]

The targeted mass spectrometric method spectrum includes selected/multiple reaction monitoring (RM) using a triple quadrupole or a quadrupole-ion trap and, more recently, PRM (parallel RM), which takes advantage of the high resolution and mass accuracy of an orbitrap or time-of-flight (TOF) mass analyzer. PRM increases the selectivity in complex matrices like plasma,[7] and it has been successfully used to confirm proteins as biomarkers.[14−16] In addition to its advantages, the maximum number of peptides that can be measured in a single acquisition by PRM is restricted by the number of data points defining the chromatography peaks. If the cycle time for the selected peptides is too long, the shape of the chromatographic

Received: September 23, 2022
Published: April 3, 2023



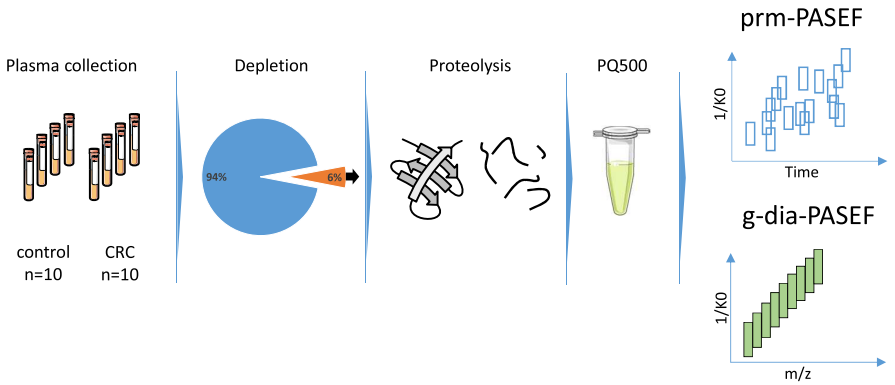


Figure 1. Plasma samples were depleted, digested by trypsin, and spiked with PQ500 isotope-labeled synthetic peptides. Samples were analyzed using prm-PASEF and g-dia-PASEF, and data were processed with Skyline.

peak is not described sufficiently. Targeted proteomics also requires a significant time investment to select and validate proteotypic peptides, and the cost of synthetic peptide standards can limit large-scale screening projects.

To overcome the time investment limitation, we combined our recently developed targeted acquisition method (PRM with parallel accumulation and serial fragmentation (prm-PASEF)) with a commercial set of 782 peptide-derived plasma proteomics studies (PQ500, Biognosys). The PQ500 set contains 804 isotope-labeled peptides in known amounts representing 582 plasma proteins. Targeting a given group of peptides reduces the development time as the methodology needs to be optimized only once. prm-PASEF uses trapped ion mobility (TIMS) in combination with a quadrupole TOF mass spectrometer to increase the multiplexing capability without losing the chromatographic peak definition and sensitivity.[11] Using the defined concentrations of the labeled peptide from the PQ500 set, the mass spectrometry (MS) signal of the endogenous peptides can be converted into a concentration, thus allowing the comparison of the quantification results across different mass spectrometry systems or acquisition methods.

As an alternative to the prm-PASEF method, which requires less setup time before the measurement, we developed guided data independent acquisition (dia)-PASEF (g-dia-PASEF). Here, we combined the PQ500 peptide set with dia-PASEF.[17] This allows the rapid measurement of peptide samples while using the internal standard to confirm the identification, normalize the signal, and quantify the peptides (Figure 1). Both methods allowed measuring and quantifying a significant fraction of the blood proteome in a single assay. Using the plasma from 20 patients of a colon carcinoma cohort, we showed the potential for the rapid identification of protein panels as potential disease markers. We demonstrated that combining ion mobility with high-resolution mass spectrometry paves the way to highly multiplexed quantitative assays in clinical samples.

MATERIAL AND METHODS

Patient Samples

Patient samples were donated willingly under informed consent and were handled in accordance with institutional guidelines. Ethical approval was obtained from the Comité National d'Ethique de Recherche, Luxembourg (reference 201009/09) followed by institutional approval by the Ethics Review Panel of the University of Luxembourg (ERP-16-032).

Plasma Depletion and Processing

Aliquots of 20 μL of human plasma were depleted on a 1260 Infinity Bio-inert liquid chromatography (LC) system (Agilent) coupled to a depletion column (human 14 multiple affinity removal column; 4.6 × 50 mm; Agilent) according to the manufacturer's procedure. After depletion, the buffer was exchanged for 100 mM NH_4HCO_3, and the volume was concentrated to 100 μL using a spin concentrator with a 5 kDa cutoff (Pall). Proteins were denatured with 1% sodium deoxycholate (SDC), reduced with 10 mM dithiothreitol for 30 min at 37 °C, and alkylated with 25 mM iodoacetamide for 30 min at room temperature. All reagents were prepared in a freshly made 50 mM ammonium bicarbonate buffer pH 8.0. The denatured samples were diluted to 0.5% SDC and proteolysis was performed by adding 6.5 μg of sequencing grade trypsin (Promega) for 16 h at 37 °C with a ratio protease protein of approximately 1/20 (w/w). Potential N-glycan chains were trimmed by adding 5 U of PNGase F for 1 h at 37 °C followed by an additional step of trypsin digestion with 1 μg of trypsin for 3 h at 37 °C. The SDC was removed by precipitation with 1% formic acid and centrifugation. Digested samples were cleaned on Sep-Pak C18 cartridges (Waters) and dried in a vacuum centrifuge. Samples were reconstituted with 200 μL of 0.1% formic acid/4% acetonitrile. Sample concentrations were normalized based on the absorbance at 205 nm measured with a nanodrop (Thermo Scientific)[2,18] (concentrations are available in Supplemental Table 10). Samples were spiked with the PQ500 kit (Biognosys), consisting of a mixture of 804 stable isotopes labeled (SIL) peptides of known concentration determined by amino acid analysis. SIL peptides were isotope labeled with arginine, $^{13}C_6$, $^{15}N_4$, Δm = 10 u or lysine $^{13}C_6$, $^{15}N_2$, Δm = 8 u, with an isotope purity of >99.5% (according to the manufacturer). The SIL peptide mixture was provided in lyophilized form, and samples were exclusively spiked with a new SIL peptide solution suspended according to the manufacturer's protocol. For reproducibility experiments, each sample was re-injected in triplicate.

LC–MS Data Acquisition

The samples were analyzed on a nano-UHPLC (nanoElute, Bruker Daltonics) coupled to a tims-TOF Pro mass spectrometer (Bruker Daltonics). Samples (4 μL, 0.042 μg/μL) were directly injected onto a pulled emitter column (250 mm × 75 μm, 1.6 μm, C18; IonOptiks) which was heated to 50 °C in a column oven. Mobile phases consisted of 0.1% (v/v) formic acid in water (phase A) and acetonitrile (phase B).

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

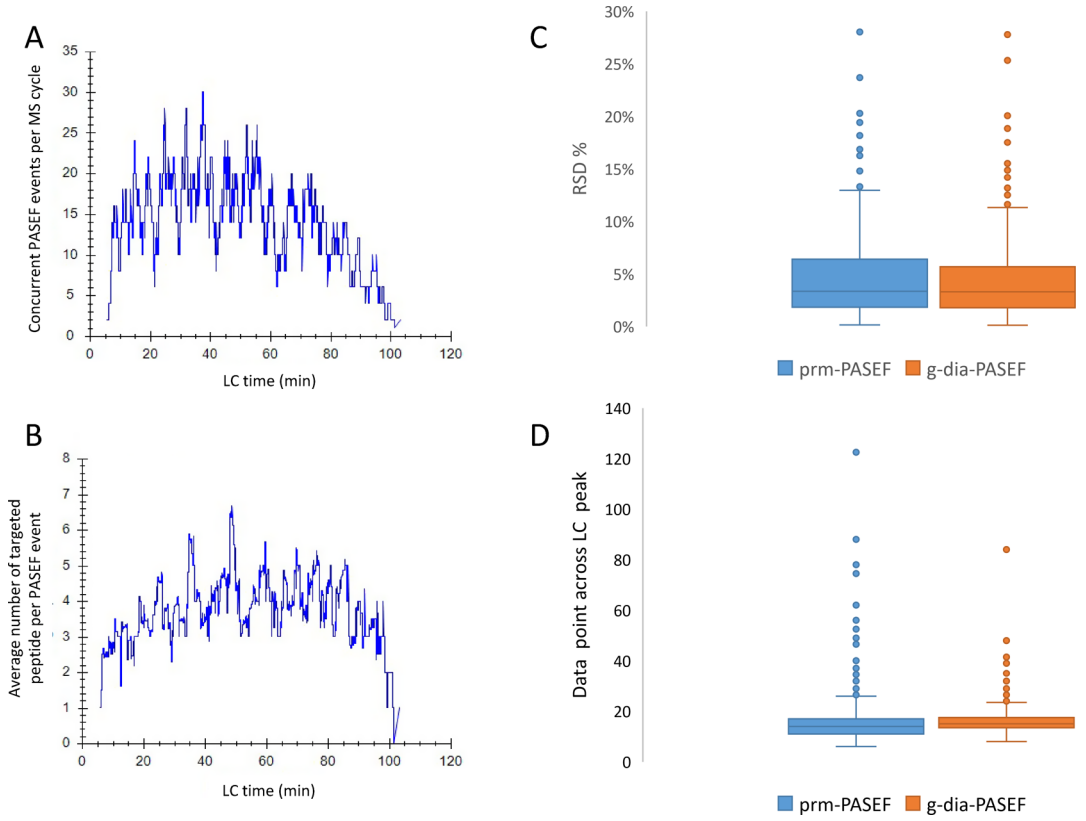


Figure 2. (A) Number of PASEF events (100 ms) per MS cycle across the chromatography separation, with the prm-PASEF method. (B) Averaged number of PASEF events per MS cycle during the prm-PASEF acquisition. (C) Number of data points per LC peak profile. (D) Relative standard deviation (RSD) of the endogenous to heavy ratios of the prm-PASEF and g-dia-PASEF analysis of a pooled plasma sample (n = 3).

Samples were separated on a 105 min stepped gradient ranging from 2−30% B at a flow rate of 400 nL/min. The gradient is built of sequential steps ranging from 2 to 15% B in 60 min, 15 to 22% B in 20 min, 22 to 26% in 10 min, 26 to 33% in 10 min, and finally, 33 to 45% in 5 min. The column was cleaned with 10 min step at 95% and equilibrated at initial conditions with seven column volumes. The mass spectrometer was fully calibrated, and the ion mobility was automatically calibrated between each run. The performance of the mass spectrometer was monitored by standardized measurements of a HeLa tryptic digest (Thermo), where the number of peptides and the total ion chromatogram (TIC) shape were considered.

The nano-UHPLC was coupled to a tims-TOF Pro instrument (Bruker Daltonics) operated in prm-PASEF mode or dia-PASEF mode. The prm-PASEF method was defined with a range of mobility values of 0.6−1.6 $1/K_0$, a TIMS accumulation time fixed at 50 ms, while the ion mobility separation was fixed to 100 ms. The time and mobility scheduled acquisition boxes were set with 2 min of tolerance on retention time and 0.05 $1/K_0$ on the ion mobility.

The g-dia-PASEF method is based on dia-PASEF with 32 isolation windows of 26 m/z width, including a margin of 0.5 m/z. Isolation windows were associated with ion mobility windows of 0.3 $1/K_0$ to cover the peptide-ions distribution on both m/z and mobility dimensions as closely as possible. The TIMS accumulation and separation were both set at 100 ms.

Data Processing

All MS data were processed with Skyline daily,[19] and extracted fragment ion chromatograms (XICs) were extracted with a TOF resolution tolerance set to 60,000. For the g-dia-PASEF method, ion mobility data were filtered with tolerance windows of 0.05 $1/K_0$ centered on the experimental mobility values. Only fragment ions of the y-series were allowed, avoiding b-series crosstalk between the endogenous and the isotope-labeled standard peptides. For g-dia-PASEF data, retention time prediction was performed using iRT peptides and Biognosys spectral library values. The predicted retention time windows were set to 20 min. Both prm-PASEF and g-dia-PASEF transitions were manually curated to remove interfering transitions and correct wrong peak picking or peak integration boundaries. Light to heavy peptide area ratios were exported from Skyline and filtered with a similarity score (normalized ratio dot product) higher or equal to 0.98. The dot product was calculated with Skyline v21.1. We manually inspected the chromatograms of the 10% lowest abundant peptides and eliminated peptides detected with poor signal quality. Examples of selected peptide TICs are shown in Figure 4. The manually curated data were finally processed using R (R-project.org). Peptides not found in all samples were removed, and the significantly regulated peptides were identified using a paired t-test corrected for multiple testing.[20] Peptides that showed an adjusted p-value <0.05 were considered significantly regulated. The significantly regulated peptides were mapped back to the related protein. The gene ontology (GO) term annotations were created using the g:profiler software library for R.[21]

prm-PASEF and g-dia-PASEF data are accessible in the public Proteome Exchange database (https://www.ebi.ac.uk/pride/) with the identifier: PXD036594.

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

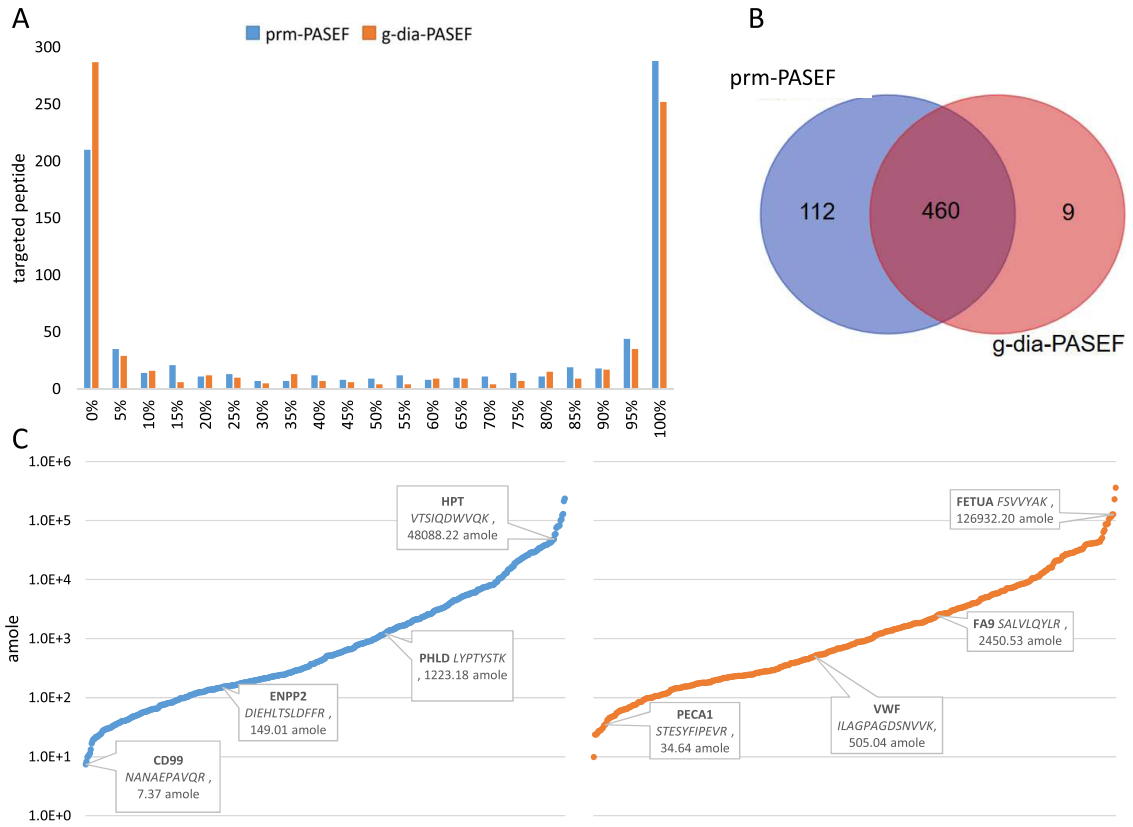


Figure 3. (A) Histogram of the completeness for the prm-PASEF experiment. Bins represent the percentage of successful detection across the 20 samples. (B) Peptide detection overlap between both methods. (C) Dynamic range of the peptide detection. Peptides with different abundances are shown with the detected amounts.

Bioinformatic Meta-Analysis of Transcriptome Data

We have set up a meta-analysis and used it as previously described.[22] Briefly, we integrated all individual CEL files from selected data sets profiled on HG-U133 plus 2.0 (Affymetrix, Santa Clara, CA, USA), retrieved from GEO (GSE14333, GSE17538, GSE21510, GSE8671, GSE9254, GSE20916, GSE10714, GSE15960, GSE4183, and GSE10961) and corresponding to different studies into one single global analysis covering expression data on 829 patients. The suitability of potential markers to discriminate between colorectal cancer (CRC) and normal colorectal samples was assessed by receiver operating characteristic curves as previously described.[22]

RESULTS AND DISCUSSION

prm-PASEF and dia-PASEF Acquisition Methods: Design and Performance

The PQ500 isotope-labeled standard peptide mix comprised 804 isotope-labeled peptides mixed at known concentrations and covered 578 blood proteins. Some strongly hydrophilic peptides did not elute reproducibly and thus were removed from the prm-PASEF method. The prm-PASEF method allowed the detection of 782 synthetic peptides spiked into the plasma samples, covering a total of 565 proteins with a 100 min chromatography gradient (Supplemental Table S1). Including the corresponding unlabeled peptides from the endogenous proteins, this sums up to 1564 precursor ions analyzed with a targeted acquisition window of 2 min and ion mobility window of 0.05 $1/K_0$.

For the g-dia-PASEF method, variations in the retention times were not a problem as the method did not require scheduled acquisition. However, few peptides were not detected because they were outside of the m/z or ion mobility scanning range of the method. We manually re-analyzed all missing peptides without retention time filtering to ensure that no peptide was missing due to inaccurate retention time prediction. We finally detected 756 internal standard peptides covering 549 proteins with the g-dia-PASEF method (Supplemental Table S1).

A PASEF event covers the accumulation of the incoming ions from the source into the first TIMS cell, the ion mobility separation of the peptide ions in the second TIMS cell followed by the MS/MS analysis in the Q-TOF section of the instrument. The prm-PASEF MS cycle consisted of a variable number of PASEF events designed to acquire all the targeted peptides across the chromatographic separation.[11] A prm-PASEF method is scheduled in a way that target peptide-ions with a similar retention time, but nonoverlapping ion mobility values can be processed from the same ion mobility separation. If the ion mobility of a peptide overlaps with another peptide, an extra PASEF event is added to the cycle time. We recorded a maximum of 30 prm-PASEF events per MS cycle (Figure 2A), which correspond to an MS cycle of approximately 3 s. A maximum average of 6.8 precursor-ions targeted per prm-PASEF event in an MS cycle was measured (Figure 2B). Despite the high density of targeted precursor-ions, it was possible to maintain a controlled cycle time with the prm-PASEF method and to acquire the peptides with a median number of 14 data points per peak and a minimum of 6 points

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

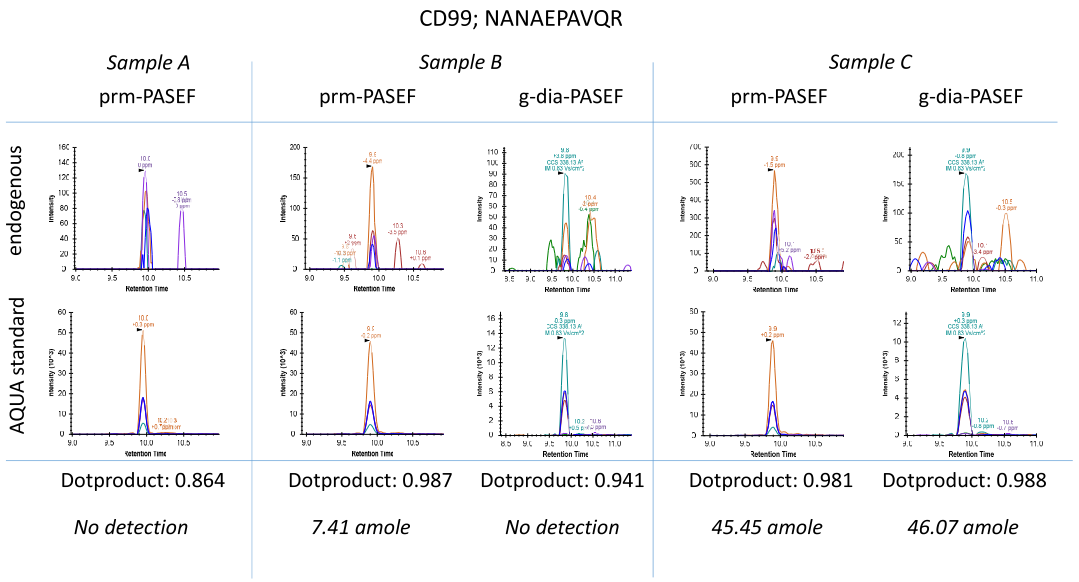


Figure 4. XICs of the peptide NANAEPAVQR (CD99) in prm- and g-dia-PASEF mode. Peptide identification was confirmed by similarity scoring (dot products) between the fragmentation pattern of the endogenous and the internal standard. Quantitation results are expressed in amol injected onto the column.

(Figure 2C), ensuring a proper quantification performance as illustrated by a median coefficient of variation of 3% (Figure 2D and Supplemental Tables S5, S6).

We observed that both prm-PASEF and g-dia-PASEF methods were performed with similar reproducibility and a median RSD of 3%. This value was calculated for each peak area of unlabeled/labeled product ion pair detected in the three replicates (Supplemental Table S7). Previously we showed that prm-PASEF measurements are linear over 3.5 orders of magnitude and have an average precision of 3% when combined with an internal standard.[9] Because the MS cycle time was constant with g-dia-PASEF (1.8 s), a sharper dispersion of the number of data points per peak was expected, and the actual variation reflected the actual chromatographic peak width of the different peptides.

prm-PASEF data were processed with the Skyline software, and area ratios of the fragment-ion chromatograms of the endogenous and heavy-isotope-labeled standard, as well as their respective dot-product similarity scores, were exported. Data points associated with a dot product score below 0.98 were filtered out, and we manually inspected the chromatograms of the 10% lowest intense peptides and eliminated peptides detected with poor signal quality. Figure 3A shows the data completeness distribution across the entire dataset showing a distribution where peptides are primarily detected in all or no samples for both methods. This distribution can be expected for a targeted acquisition method (i.e., PRM, prm-PASEF) or a targeted data processing (i.e., DIA, dia-PASEF) which are driven mainly by limits of detection. The risk of missing a peptide during acquisition is not as prevalent as with data-dependent acquisition (DDA). We estimated the actual concentration of the 572 endogenous peptides covering 378 proteins with the prm-PASEF method. The comparison with the g-dia-PASEF method (469 peptides quantified covering 308 proteins) revealed that 112 peptides were only detected in prm-PASEF. This higher performance can be explained by the fact that prm-PASEF relies on narrow quadrupole isolation windows that improve the signal-to-noise ratio. The g-dia-PASEF method detected nine unique peptides that were typically peptides with variable retention times (Figure 3B) and thus were not detected in prm-PASEF (Supplemental Tables S8 and S9).

We estimated the concentration of the endogenous peptides by single-point calibration with the spiked isotope-labeled standard (PQ500, Biognosys) (Figure 3C). We calculated a peptide detection ranging from 7.4 (CD99) to 234,232 (HEMO) amol injected onto the column for prm-PASEF. The limit of detection and dynamic range are in line with our previous technical evaluation of the prm-PASEF method on plasma and cell extract samples.[11,23] Figure 4 shows prm-PASEF and g-DIA-PASEF elution profiles for the lowest abundant protein detected in the set (CD99). prm-PASEF detected 27 low abundant peptides below 30 amol injected, whereas the dia-PASEF detected eight peptides below that threshold. The correlation between the two methods for the overlapped quantification results was high (Supplemental Figure S1).

We finally compared the peptides and associated proteins detectable with label-free DIA data processing. It is possible to process the DIA data without being restricted by the peptides included in the PQ500 kit, and this approach detected 5773 peptides associated with 774 protein groups. Interestingly, we found that if the number of detected peptides is about ten times higher than in the targeted approaches, the number of detected protein groups stayed in the same range. In total, 453 unique proteins, primarily associated with extracellular vesicles, the complement system, and immunoglobins, were detected (Supplemental Figure S2 and Supplemental Table S2). Even if the label-free approach does not provide the same degree of identification and quantification confidence as SIL peptide-based DIA, it still allows estimating the profiles of proteins not covered by the PQ500 set.

Correlation of the prm-PASEF and g-diaPASEF Data

The correlation between the two measurement methods is high, with an R^2 ranging between 0.92 and 0.97 for all samples (Supplemental Figure S1). The main differences are the higher

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

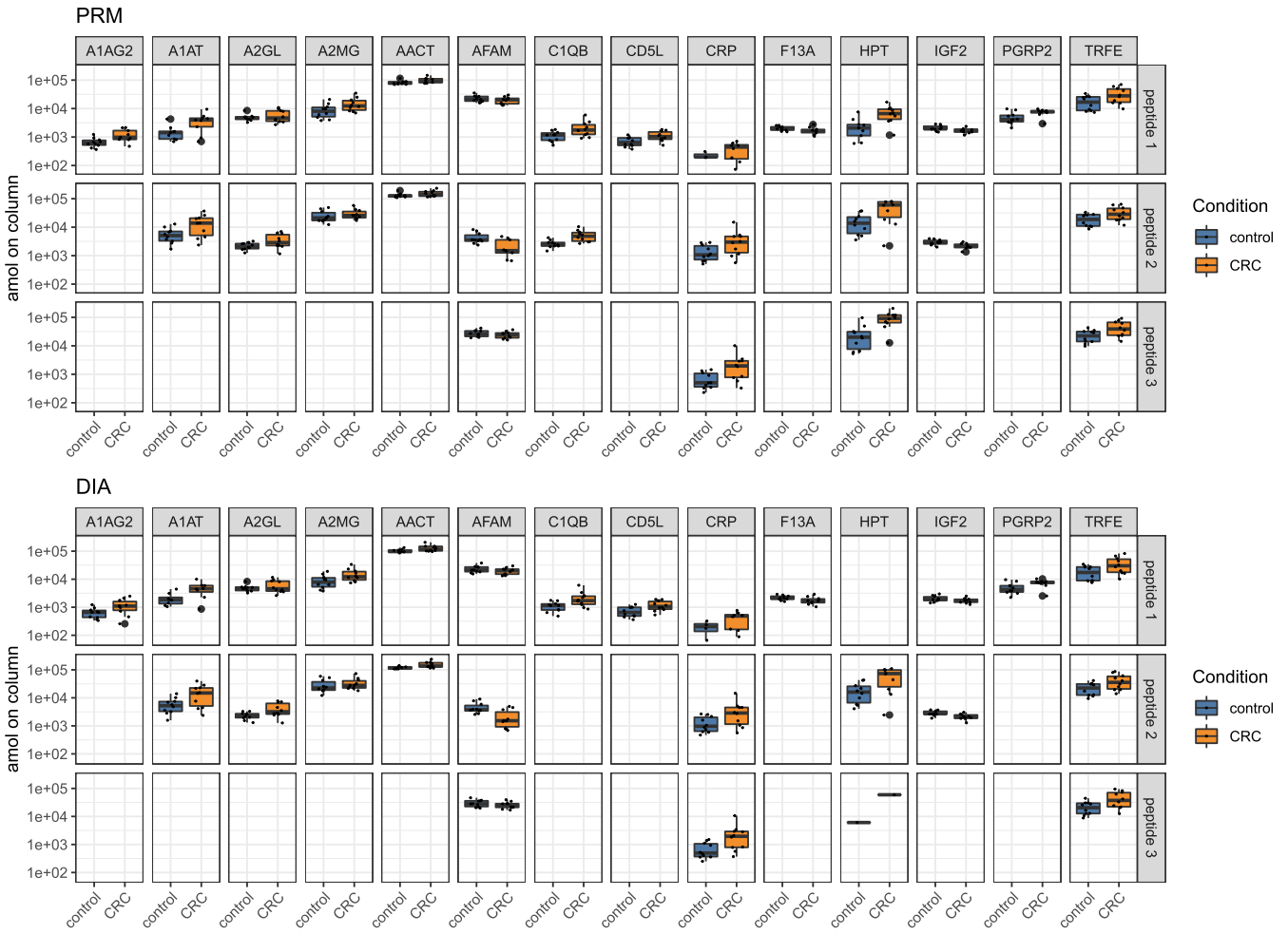


Figure 5. Boxplot for the proteins significantly regulated between the CRC and control patients shown for prm-PASEF (panel A) and g-dia-PASEF (panel B). Measured data are overlayed on the boxplot. In the PQ500 standard, the number of peptides per protein varies. For the regulated proteins, the number of peptides per protein varies between one and three peptides. Each boxplot shows the distribution of the quantification separated by each peptide. CRC patients are shown in orange, and controls in dark blue. Proteins were considered significantly regulated with a Benjamini−Hochberg-corrected *p*-value of 0.05.

sensitivity of the prm-PASEF measurement, while the g-dia-PASEF method has a higher tolerance to retention time shifts.

The lower limit of quantification of the prm-PASEF method is based on the quadrupole peptide-ion isolation window, which is at unit resolution. The gain in sensitivity comes at the cost of a more demanding method development. The prm-PASEF method requires additional measurements for the determination of retention times and ion mobility values to set up the acquisition method. The acquisition method must balance the number of peptides, the type of gradient, and MS measurement requirements to keep the number of data points per chromatographic peak acceptable for quantification.

For this analysis, we used a 25 cm column packed with 1.6 μm particles to increase the peak capacity while maintaining sharp elution profiles (i.e., 7.4 s full-width half-maximum on average) on a 100 min gradient. The reproducibility of the retention was extremely important and required constant monitoring during the sample acquisition. Ion mobility values are susceptible to change with time and affect the synchronization of the TIMS with the quadrupole; we employed an automated recalibration of the ion mobility trap between each sample injection to alleviate this problem. The recently developed on-the-fly retention time correction might improve the implementation of prm-PASEF, allowing for narrower acquisition windows and thus shorter LC gradients.[24]

Conversely, the g-dia-PASEF method is more tolerant to retention time changes and does not require as much care and monitoring as prm-PASEF during the acquisition process. Since g-dia-PASEF is based on dia-PASEF, it allows the in silico identification of peptides that are not included in the isotope-labeled standard peptide mix. Peptides without a SIL can still be quantified between samples using label-free quantification methods. The GO-analysis of the additionally identified proteins revealed that these mainly belong to the immunoglobulins and complement proteins (Supplemental Table S2).

One of the most time-consuming steps is the data processing and curation for prm-PASEF and g-dia-PASEF. The signal integration quality control of 1564 precursor-ions is the main time-consuming step for both approaches. In our hands, Skyline was a practical software for the visualization and correction of peak integration boundaries.[19] Skyline draws customizable and interactive graphics and tables that visualize critical metrics such as retention times, XIC areas, mass error, and the dot-product similarity score. Graphics and tables are clickable and linked to the peptides XICs, allowing for a swift

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

verification and correction of the outliers. We used the dot-product to filter true peptide signals from background noise.

Additionally, each peak was manually reviewed, particularly for low-abundance peptides. For these peptides, background noise is particularly prone to false-positive dot product scores and is more likely to happen when one transition strongly dominates the others. We did not use Spectronaut (Biognosys) for the g-diaPASEF result processing because it tended to overestimate the peptides' detection with this experimental design.

Detection of Differentially Expressed Proteins between CRC and Control Patients

To demonstrate the power of the prm-PASEF and g-dia-PASEF techniques, we applied the measurement to the plasma of 20 patients of the CRC cohort.[25] The patients of the two groups were age-matched (Supplemental Figure S3A), and the 10 CRC patients were equally distributed across the different disease stages (Supplemental Figure S3B).

For the analysis of the prm-PASEF and the g-DIA-PASEF data, only peptides, which were detected in all PRM or in all DIA samples, were included. Of the 288 proteins detected in all samples, 14 were significantly regulated (Benjamini–Hochberg corrected p-value <0.05) between the CRC patients and the control group.[20] The proteins were further filtered for consistent regulation if several peptides per protein were measured (examples are shown in Figure 5), and 11 proteins were upregulated in the cancer groups, while three proteins were downregulated (Figure 5 and Supplemental Table S3). We used GO term analysis to see if these proteins belong to a similar pathway. Most of the proteins are associated with the activation of the immune system (Supplemental Table S4), especially the activation of the antibody-mediated immune response. This is expected as many CRC patients have predisease inflammations of the intenstine.[26] Of particular interest is the enrichment of the CORUM complex term (IGF2-IGFBP2 complex), which is a known CRC-associated regulatory complex.[27] We used data obtained from different public tumor databases containing transcriptomics data from tumor material[25] to analyze if the genes were found to be associated with CRC in other studies (Supplemental Figure S4). The analysis showed A2M (A2MG), C1QB, CD5L, CRP, IGF2, and Serpin A1(A1AT) mRNAs are upregulated in comparison to tissue samples of healthy control patients and show the same regulation as the detected proteins. F13A1-(F13A) and IGF2 proteins and mRNA were downregulated. Interestingly, the PGLYRP2 (PGRP2) showed no regulation of its mRNA while the protein was upregulated and for serotransferin the mRNA was downregulated, while the protein was upregulated in CRC patients (Supplemental Figure S4). An analysis of the current literature revealed that A1AT,[28] IGF2,[29] and F13A[30] were previously associated with colon cancer. Even with the low number of patients, we can see a trend for the stage-specific regulation of the selected proteins (Supplemental Figure S5).

CONCLUSIONS

Here, we describe the use of a highly complex isotope-labeled peptide standard in combination with prm-PASEF and dia-PASEF. It allowed us to develop the new g-dia-PASEF technique, which is less complex to set up than the prm-PASEF method while still allowing the relative quantification of the plasma proteins between samples or as absolute quantification with the known restrictions.

The prm-PASEF analysis allowed us to measure and estimate the concentration of 782 peptides in plasma samples with a high dynamic range of four orders of magnitude (Supplemental Table S5). The comparison of the prm-PASEF with the g-dia-PASEF method shows that the highest number of peptides in a plasma sample is identified with prm-PASEF, but it requires a significant workload to set up the acquisition method. The g-dia-PASEF method combines the ease of setting up a dia-PASEF method with the possibility of quantifying peptides with an isotope-labeled standard peptide mix. Using g-dia-PASEF offers the important advantage of cross-experiment comparison over label-free dia-PASEF, which is essential for using biomarker measurements.

The careful analysis of each peptide's XIC revealed that the automated software algorithms for peak detection, integration, and identity validation are still prone to false positives. A more stringent reevaluation of the results is necessary for the targeted measurement using prm-PASEF and g-dia-PASEF. This is particularly important for potential clinical applications.

Using the prm-PASEF and g-dia-PASEF methods for analyzing the CRC-cohort samples, we showed that despite the relatively low number of samples, we could distinguish CRC patients from the control group with an unbiased measurement of the 579 selected plasma proteins, which represent a significant portion of the plasma proteome. The significantly regulated proteins contain proteins associated with colon carcinoma or immune responses, which are typical of CRC patients. For validation of these proteins as a biomarker, more measurements in a larger cohort will be necessary, but the 14 proteins identified as CRC-associate already underline the potential of the measurement as a rapid preselection tool.

Based on this study, it should be possible to further utilize this technique as a general measurement standard. This means that a patient sample can be measured using the PQ500 or a similar standard, and the results can be directly compared to other patient samples measured using the same approach. The diagnostic potential would be extended by measuring 500 proteins and will allow the definition of protein panels as biomarkers of different diseases. A single plasma proteome measurement using the PQ500 standard with either prm-PASEF or g-diaPASEF could be the unified diagnostic tool for many different diseases.

In conclusion, our study monitored the highest number of peptides by PRM in plasma samples so far. The results show that PRM-based techniques are now capable of competing with discovery methods without the drawbacks of DDA-based measurements in plasma.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.2c00575.

Correlation quantification prm-PASEF and g-dia-PASEF, overlap of peptide detection dia-PASEF, g-dia-PASEF, and prm-PASEF, characteristics of the CRC cohort, mRNA expression of the corresponding genes, and regulated proteins by CRC stage (PDF)

List of detectable SIL peptides with prm- and g-dia-PASEF, GO term analysis of the proteins identified in the label-free dia-PASEF analysis, summary of the

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638

detected amount (amol onto column), concentration (fmole/μg of total protein), p-value, fold change, RSD in a technical triplicate for the 14 significantly differentially expressed proteins, GO term analysis of the significantly regulated proteins from the prm-PASEF analysis, quantification values in amol onto the column for all detected peptides by prm-PASEF, quantification values in amol onto the column for all detected peptides by g-dia-PASEF, RSD values for the technical repetitions, coefficient of variation for the biological samples measured by prm-PASEF, coefficient of variation for the biological samples measured by g-dia-PASEF, and concentrations of the depleted patient plasma samples (XLSX)

AUTHOR INFORMATION

Corresponding Author

Gunnar Dittmar − *Luxembourg Institute of Health, Strassen L-1445, Luxembourg; Department of Life Sciences and Medicine, University of Luxembourg, Belvaux L-4367, Luxembourg;* orcid.org/0000-0003-3647-8623; Email: gunnar.dittmar@lih.lu

Authors

Antoine Lesur − *Luxembourg Institute of Health, Strassen L-1445, Luxembourg*

François Bernardin − *Luxembourg Institute of Health, Strassen L-1445, Luxembourg*

Eric Koncina − *Department of Life Sciences and Medicine, University of Luxembourg, Belvaux L-4367, Luxembourg*

Elisabeth Letellier − *Department of Life Sciences and Medicine, University of Luxembourg, Belvaux L-4367, Luxembourg*

Gary Kruppa − *Bruker Daltonics, Billerica, Massachusetts 01821, United States*

Pierre-Olivier Schmit − *Bruker Daltonique S.A., Wissembourg F-67160, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.2c00575

Author Contributions

A.L. developed the method implementation. A.L. and F.B. performed the sample preparation and measurement. E.K. performed the analysis of the tumor expression data. E.L. provided the CRC samples. A.L., G.K., P.O.S., and G.D. developed the methodology. A.L. wrote the first draft. A.L. and G.D. finalized the manuscript.

Notes

The authors declare the following competing financial interest(s): Two of the authors Gary Kruppa and Pierre-Olivier Schmit are employees of Bruker. The other authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank the Foundation Cancer for the support with grant FC/2022/01 SOCS to E.L. The authors would like to thank Dr. Guy Berchem and Dr. Keipes for their support and the IBBL (Integrated Biobank of Luxembourg) for the overall setup of the sample collection, processing, and provision. We thank Komal Baig for the setup of the meta-analysis. We thank Tony Kaoma and Dr. Petr Nazarov from the LIH bioinformatics platform and Dr. Daniel Perez-Hernandez for the fruitful discussions on the statistical analysis.

REFERENCES

(1) Coll-de la Rubia, E.; Martinez-Garcia, E.; Dittmar, G.; Nazarov, P. V.; Bebia, V.; Cabrera, S.; Gil-Moreno, A.; Colás, E. In Silico Approach for Validating and Unveiling New Applications for Prognostic Biomarkers of Endometrial Cancer. *Cancers* **2021**, *13*, 5052.

(2) Martinez-Garcia, E.; Lesur, A.; Devis, L.; Cabrera, S.; Matias-Guiu, X.; Hirschfeld, M.; Asberger, J.; van Oostrum, J.; Casares de Cal, M.; de los, Á.; Gómez-Tato, A.; Reventos, J.; Domon, B.; Colas, E.; Gil-Moreno, A. Targeted Proteomics Identifies Proteomic Signatures in Liquid Biopsies of the Endometrium to Diagnose Endometrial Cancer and Assist in the Prediction of the Optimal Surgical Treatment. *Clin. Cancer Res.* **2017**, *23*, 6458−6467.

(3) Coll-de la Rubia, E.; Martinez-Garcia, E.; Dittmar, G.; Gil-Moreno, A.; Cabrera, S.; Colas, E. Prognostic Biomarkers in Endometrial Cancer: A Systematic Review and Meta-Analysis. *J. Clin. Med.* **2020**, *9*, 1900.

(4) Griffith, O. L.; Chiu, C. G.; Gown, A. M.; Jones, S. J.; Wiseman, S. M. Biomarker Panel Diagnosis of Thyroid Cancer: A Critical Review. *Expert Rev. Anticancer Ther.* **2008**, *8*, 1399−1413.

(5) Mun, S.; Lee, J.; Park, M.; Shin, J.; Lim, M.-K.; Kang, H.-G. Serum Biomarker Panel for the Diagnosis of Rheumatoid Arthritis. *Arthritis Res. Ther.* **2021**, *23*, 31.

(6) Bhardwaj, M.; Terzer, T.; Schrotz-King, P.; Brenner, H. Comparison of Proteomic Technologies for Blood-Based Detection of Colorectal Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 1189.

(7) Surinova, S.; Schiess, R.; Hüttenhain, R.; Cerciello, F.; Wollscheid, B.; Aebersold, R. On the Development of Plasma Protein Biomarkers. *J. Proteome Res.* **2011**, *10*, 5−16.

(8) Nakayasu, E. S.; Ansong, C.; Brown, J. N.; Yang, F.; Lopez-Ferrer, D.; Qian, W.-J.; Smith, R. D.; Adkins, J. N. Evaluation of Selected Binding Domains for the Analysis of Ubiquitinated Proteomes. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1214−1223.

(9) Liu, Y.; Hüttenhain, R.; Collins, B.; Aebersold, R. Mass Spectrometric Protein Maps for Biomarker Discovery and Clinical Research. *Expert Rev. Mol. Diagn.* **2013**, *13*, 811−825.

(10) Hüttenhain, R.; Malmström, J.; Picotti, P.; Aebersold, R. Perspectives of Targeted Mass Spectrometry for Protein Biomarker Verification. *Curr. Opin. Chem. Biol.* **2009**, *13*, 518−525.

(11) Lesur, A.; Schmit, P.-O.; Bernardin, F.; Letellier, E.; Brehmer, S.; Decker, J.; Dittmar, G. Highly Multiplexed Targeted Proteomics Acquisition on a TIMS-QTOF. *Anal. Chem.* **2021**, *93*, 1383−1392.

(12) Picotti, P.; Rinner, O.; Stallmach, R.; Dautel, F.; Farrah, T.; Domon, B.; Wenschuh, H.; Aebersold, R. High-Throughput Generation of Selected Reaction-Monitoring Assays for Proteins and Proteomes. *Nat. Methods* **2010**, *7*, 43−46.

(13) Gallien, S.; Duriez, E.; Crone, C.; Kellmann, M.; Moehring, T.; Domon, B. Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Mol. Cell. Proteomics* **2012**, *11*, 1709−1723.

(14) Kim, Y. J.; Gallien, S.; El-Khoury, V.; Goswami, P.; Sertamo, K.; Schlesser, M.; Berchem, G.; Domon, B. Quantification of SAA1 and SAA2 in Lung Cancer Plasma Using the Isotype-Specific PRM Assays. *Proteomics* **2015**, *15*, 3116−3125.

(15) Henderson, C. M.; Bollinger, J. G.; Becker, J. O.; Wallace, J. M.; Laha, T. J.; MacCoss, M. J.; Hoofnagle, A. N. Quantification by Nano Liquid Chromatography Parallel Reaction Monitoring Mass Spectrometry of Human Apolipoprotein A-I, Apolipoprotein B, and Hemoglobin A1c in Dried Blood Spots. *Proteomics: Clin. Appl.* **2017**, *11*, No. 1600103.

(16) Sinha, A.; Hussain, A.; Ignatchenko, V.; Ignatchenko, A.; Tang, K. H.; Ho, V. W. H.; Neel, B. G.; Clarke, B.; Bernardini, M. Q.; Ailles, L.; Kislinger, T. N-Glycoproteomics of Patient-Derived Xenografts: A Strategy to Discover Tumor-Associated Proteins in High-Grade Serous Ovarian Cancer. *Cell Syst.* **2019**, *8*, 345−351.

(17) Meier, F.; Brunner, A.-D.; Frank, M.; Ha, A.; Bludau, I.; Voytik, E.; Kaspar-Schoenefeld, S.; Lubeck, M.; Raether, O.; Bache, N.; Aebersold, R.; Collins, B. C.; Röst, H. L.; Mann, M. DiaPASEF: Parallel Accumulation−Serial Fragmentation Combined with Data-Independent Acquisition. *Nat. Methods* **2020**, *17*, 1229−1236.
(18) El-Khoury, V.; Schritz, A.; Kim, S.-Y.; Lesur, A.; Sertamo, K.; Bernardin, F.; Petritis, K.; Pirrotte, P.; Selinsky, C.; Whiteaker, J. R.; Zhang, H.; Kennedy, J. J.; Lin, C.; Lee, L. W.; Yan, P.; Tran, N. L.; Inge, L. J.; Chalabi, K.; Decker, G.; Bjerkvig, R.; Paulovich, A. G.; Berchem, G.; Kim, Y. J. Identification of a Blood-Based Protein Biomarker Panel for Lung Cancer Detection. *Cancers* **2020**, *12*, 1629.
(19) Pino, L. K.; Searle, B. C.; Bollinger, J. G.; Nunn, B.; MacLean, B.; MacCoss, M. J. The Skyline Ecosystem: Informatics for Quantitative Mass Spectrometry Proteomics. *Mass Spectrom. Rev.* **2020**, *39*, 229−244.
(20) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B: Methodol.* **1995**, *57*, 289−300.
(21) Reimand, J.; Arak, T.; Adler, P.; Kolberg, L.; Reisberg, S.; Peterson, H.; Vilo, J. G:Profiler—a Web Server for Functional Interpretation of Gene Lists (2016 Update). *Nucleic Acids Res.* **2016**, *44*, W83−W89.
(22) Letellier, E.; Schmitz, M.; Baig, K.; Beaume, N.; Schwartz, C.; Frasquilho, S.; Antunes, L.; Marcon, N.; Nazarov, P. V.; Vallar, L.; Even, J.; Haan, S. Identification of SOCS2 and SOCS6 as Biomarkers in Human Colorectal Cancer. *Br. J. Cancer* **2014**, *111*, 726−735.
(23) Lesur, A.; Dittmar, G. The Clinical Potential of Prm-PASEF Mass Spectrometry. *Expert Rev. Proteomics* **2021**, *18*, 75−82.
(24) Brzhozovskiy, A.; Kononikhin, A.; Bugrova, A. E.; Kovalev, G. I.; Schmit, P.-O.; Kruppa, G.; Nikolaev, E. N.; Borchers, C. H. The Parallel Reaction Monitoring-Parallel Accumulation-Serial Fragmentation (Prm-PASEF) Approach for Multiplexed Absolute Quantitation of Proteins in Human Plasma. *Anal. Chem.* **2022**, *94*, 2016−2022.
(25) Letellier, E.; Schmitz, M.; Ginolhac, A.; Rodriguez, F.; Ullmann, P.; Qureshi-Baig, K.; Frasquilho, S.; Antunes, L.; Haan, S. Loss of Myosin Vb in Colorectal Cancer Is a Strong Prognostic Factor for Disease Recurrence. *Br. J. Cancer* **2017**, *117*, 1689−1701.
(26) Galon, J.; Bruni, D. Approaches to Treat Immune Hot, Altered and Cold Tumours with Combination Immunotherapies. *Nat. Rev. Drug Discovery* **2019**, *18*, 197−218.
(27) Giurgiu, M.; Reinhard, J.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Ruepp, A. CORUM: The Comprehensive Resource of Mammalian Protein Complexes—2019. *Nucleic Acids Res.* **2019**, *47*, D559−D563.
(28) Kwon, C. H.; Park, H. J.; Choi, J. H.; Lee, J. R.; Kim, H. K.; Jo, H.; Kim, H. S.; Oh, N.; Song, G. A.; Park, D. Y. Snail and SerpinA1 Promote Tumor Progression and Predict Prognosis in Colorectal Cancer. *Oncotarget* **2015**, *6*, 20312−20326.
(29) Unger, C.; Kramer, N.; Unterleuthner, D.; Scherzer, M.; Burian, A.; Rudisch, A.; Stadler, M.; Schlederer, M.; Lenhardt, D.; Riedl, A.; Walter, S.; Wernitznig, A.; Kenner, L.; Hengstschläger, M.; Schüler, J.; Sommergruber, W.; Dolznig, H. Stromal-Derived IGF2 Promotes Colon Cancer Progression via Paracrine and Autocrine Mechanisms. *Oncogene* **2017**, *36*, 5341−5355.
(30) Picault, F.-X.; Chaves-Almagro, C.; Projetti, F.; Prats, H.; Masri, B.; Audigier, Y. Tumour Co-Expression of Apelin and Its Receptor Is the Basis of an Autocrine Loop Involved in the Growth of Colon Adenocarcinomas. *Eur. J. Cancer* **2014**, *50*, 663−674.

https://doi.org/10.1021/acs.jproteome.2c00575
J. Proteome Res. 2023, 22, 1630−1638