Unbiased and deep proteomic analysis of secretome samples using the Proteograph™ XT workflow

Shao-Yung (Eric) Chen, Gabriel Castro, Evangelina Bahu, and Aaron S. Gajadhar

Introduction

Proteomic analysis of biospecimens, encompassing the thorough characterization of proteins within non-blood biofluids like cell line conditioned media ("CM"), holds the potential to provide novel insights into the intricacies of human health and disease. Despite recent advancements in sample collection techniques and mass spectrometry technology, which have broadened our knowledge of CM secretome proteomes, the field grapples with the absence of standardized sample preparation protocols and the intricacies of analytical workflows necessary for achieving both comprehensive depth and scalability in proteomic characterizations.

CM secretomes represent an immensely valuable class of samples for investigating cell-cell signaling, communication, the development of innovative biotherapeutics, elucidating disease mechanisms, and the discovery of potential biomarkers. However, their analysis is beset by several inherent challenges. CM secretomes present a vast spectrum of protein concentrations, spanning orders of magnitude, rendering the detection of lower-abundant secreted targets a complex endeavor. Moreover, CM samples are often dominated by a limited number of highly abundant proteins, thereby hindering the identification of less abundant species.

In CM studies, researchers must face the challenge of dealing with heightened concentrations of growth supplements, notably fetal bovine serum ("FBS"), which significantly obstruct the assessment of proteins across the dynamic concentration spectrum. Consequently, the proteomics field has devised intricate preprocessing and sample manipulation strategies to surmount this obstacle, encompassing techniques such as ultracentrifugation, concentration, ultrafiltration, buffer exchange, fractionation, and protein depletion. In other cases, the process of serum starvation is employed to maximize proteome depth of coverage. While effective in mitigating the influence of dominant, highly abundant proteins, these approaches are marked by their time- and cost-intensive nature, propensity to introduce variability, susceptibility to experimental artifacts from cell stress and manipulation, and demand for a substantial degree of technical expertise.

In this study, we have assessed the efficacy of the Proteograph XT Product Suite in conjunction with conventional nano-flow liquid chromatography coupled with state-of-the-art mass spectrometry (nanoLC-MS/MS). Our investigation encompassed the analysis of diverse CM secretomes utilizing the Proteograph XT workflow. This included the examination of five distinct cell lines both with and without prior fetal bovine serum (FBS) starvation. The LC-MS/MS data were analyzed with the Proteograph™ Analysis Suite ("PAS"). Our evaluation assessed multiple parameters, including peptide yield, performance relative to conventional direct digestion methods, coverage in the identification of protein groups, and quantitative reproducibility of the Proteograph XT workflow.

The results obtained from our study highlight the significant performance enhancements and convenience benefits afforded by the Proteograph XT workflow when conducting secretome analysis. Importantly, these advantages are realized without the necessity for traditional preprocessing and sample manipulation protocols, which are often imperative for investigating the wide-ranging dynamic concentration gradients inherent to CM proteomes. Importantly, the results suggest the Proteograph XT workflow can extend detection of biologically meaningful classes of proteins with disease-association, including growth factors and cytokines.

About this application note

Easily accessible biofluids like plasma, serum, and conditioned media (CM) serve as important sample types for proteomic analysis in biomarker discovery, biotherapeutic properties and understanding human health and disease states. We evaluated the application of the Proteograph™ XT workflow, a standardized, automated, nanoparticle-based deep proteomics workflow optimized for plasma samples, to interrogate proteomic performance in a variety of secretome sample types under various cell line conditions, with no additional modifications to the standard protocol. We demonstrate the feasibility of using the standard Proteograph XT workflow to support comprehensive proteomic characterization of conditioned media secretomes.

Study design

Sample information

CM samples were procured from a third-party supplier (HD Biosciences, San Diego). CM from different cancer cell lines (LNCap, BT549, Hela, and DLD-1) were acquired and obtained under two distinct growth conditions: samples cultured with fetal bovine serum (FBS) and those subjected to FBS deprivation for 48 hours prior to collection.

The samples were cultivated in distinct initial media compositions, comprising (1) Roswell Park Memorial Institute 1640 Medium ("RPMI 1640") + 10 % FBS + 1% Penicillin-Streptomycin ("1% pen/strep") for BT549, DLD-1, and LNCap, Eagle's Minimum Essential Medium ("EMEM") + 10% FBS + 1% pen/strep for Hela, and mTeSR™1 + 10% FBS for induced pluripotent stem cells (iPSCs). Additionally, iPSCs were acquired exclusively from media containing FBS. A minimum of 275 µL was collected for all the samples to ensure compatibility with the standard Proteograph XT workflow with starting volume (240 µL). All samples were received and stored at –80°C before sample processing.

Sample preparation

Sample preparation using the Proteograph XT is shown in **Figure 1** with the following steps. (1) Upon addition of biologic samples to Seer's NP suspensions, a stable and reproducible protein corona is formed based on the particle physicochemical properties. Protein corona-containing NPs are pulled down and washed, taking advantage of the paramagnetic core. (2) Proteins are then denatured, reduced, alkylated, and digested directly on the particles using an automated one-pot sample preparation workflow, resulting in tryptic peptides released into the supernatant. The resulting peptide mixture is then desalted using solid phase extraction on the SP100 Automation Instrument. Peptides are then quantified using a fluorescence spectrometer, dried, and resuspended on the SP100 Automation Instrument before injection onto a (3) LC-MS system. (4) LC-MS raw data were transferred directly to the Proteograph PAS for peptide and protein identification, quantification, and other biological insights. For direct digestion, 20 µL of samples are added to the SP100 Automation Instrument without nanoparticles and processed in parallel with the standard Proteograph XT assay sample processing beginning at the protein denaturation step prior to enzymatic digestion **(step 3 in Figure 1)**.

LC-MS data acquisition

For Data-Independent Acquisition ("DIA"), 240 ng of peptides in 4 µL were reconstituted in a solution of 0.1% formic acid and 3% acetonitrile spiked with 5 fmol/µL PepCalMix from SCIEX (Framingham, MA) for constant mass injection for all samples.

Each sample was analyzed with a Themo Fisher Scientific™ Ultimate™ 3000 nanoLC system coupled with the Exploris™ 480 Mass Spectrometer. First, the peptides were loaded on an Acclaim™ PepMap™ 100 C18 (0.3 mm ID × 5 mm) trap column and then separated on a 50 cm µPAC analytical column (Thermo Fisher Scientific) at a flow rate of 1 µL/min using a gradient of 5–25% solvent B (99.9% ACN + 0/1 % formic acid) mixed into solvent A (99.9% water + 0.1 % formic acid) over 26 mins, and a 33 min sample-to-sample run time.

The mass spectrometer was operated in DIA mode using 10 m/z isolation windows from 380-1000 m/z and 3s cycle time. MS1 scans were acquired at 60k resolution and MS2 at 30k resolution.

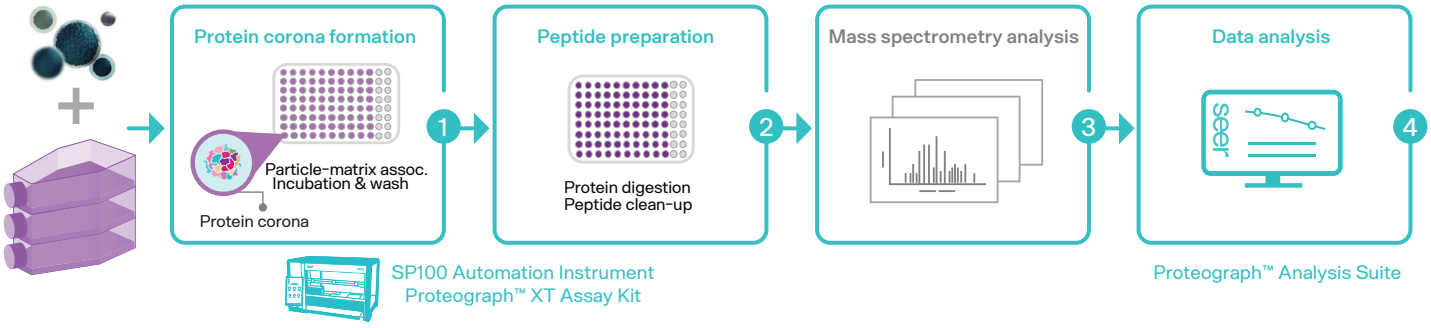


Figure 1. The Proteograph XT workflow. **(1)** Samples are added to Seer's NP suspensions and a protein corona is formed. **(2)** Proteins are processed then the resulting peptide mixture is desalted and quantified. **(3)** Peptides are injected onto an LC-MS system. **(4)** LC-MS raw data are transferred to the PAS for peptide and protein identification.

Data analysis

DIA data were processed using PAS. For CM with or without FBS starvation, raw MS data were processed with *in silico* generated spectral library from combined entries from bovine protein entries (UP000009136_9913) and human proteins (UP000005640_9606).

Library-free search parameters include trypsin protease, one missed cleavage, N-terminal Met excision, fixed modification of Cys carbamidomethylation, no Met oxidation, peptide length of 7-30 amino acids, precursor range of 300-1800 m/z, and fragment ion range of 200-1800 m/z.

Heuristic protein inference was enabled, MS1 and MS2 mass accuracy was set to 10 ppm. Precursor FDR was set to 0.01, and PG q-value was set to 0.01. Quantification was performed on summed abundances of all unique peptides considering only precursors passing the q-value cutoff. PAS summarizes all nanoparticle values for a single protein into a single quantitative value. Specifically, a single protein may have been measured both times, once for each nanoparticle mixture.

To derive the single measurement value, PAS uses a maximum representation approach, whereby the single quantification value for a particular peptide or protein group represents the quantitation value of the NP most frequently measured across all samples.

Results

Proteograph XT workflow exceeds proteome depth of coverage

To assess the efficacy of the Proteograph XT workflow relative to direct digestion, we conducted a comparative analysis of proteome identifications. The evaluation encompassed the performance of both the Proteograph XT workflow and the direct digestion method. Notably, the results unveiled a profound enhancement, approximately 3-fold across the panel of cell lines, in protein identification performance when utilizing the Proteograph XT assay in comparison to the conventional direct digestion approach for CM samples containing FBS. Similarly, CM samples subjected to a 48-hour FBS starvation regimen exhibited a discernible improvement, approximately 1.3-fold across the panel, with the Proteograph XT workflow as opposed to direct digestion **(Figure 2A)**.

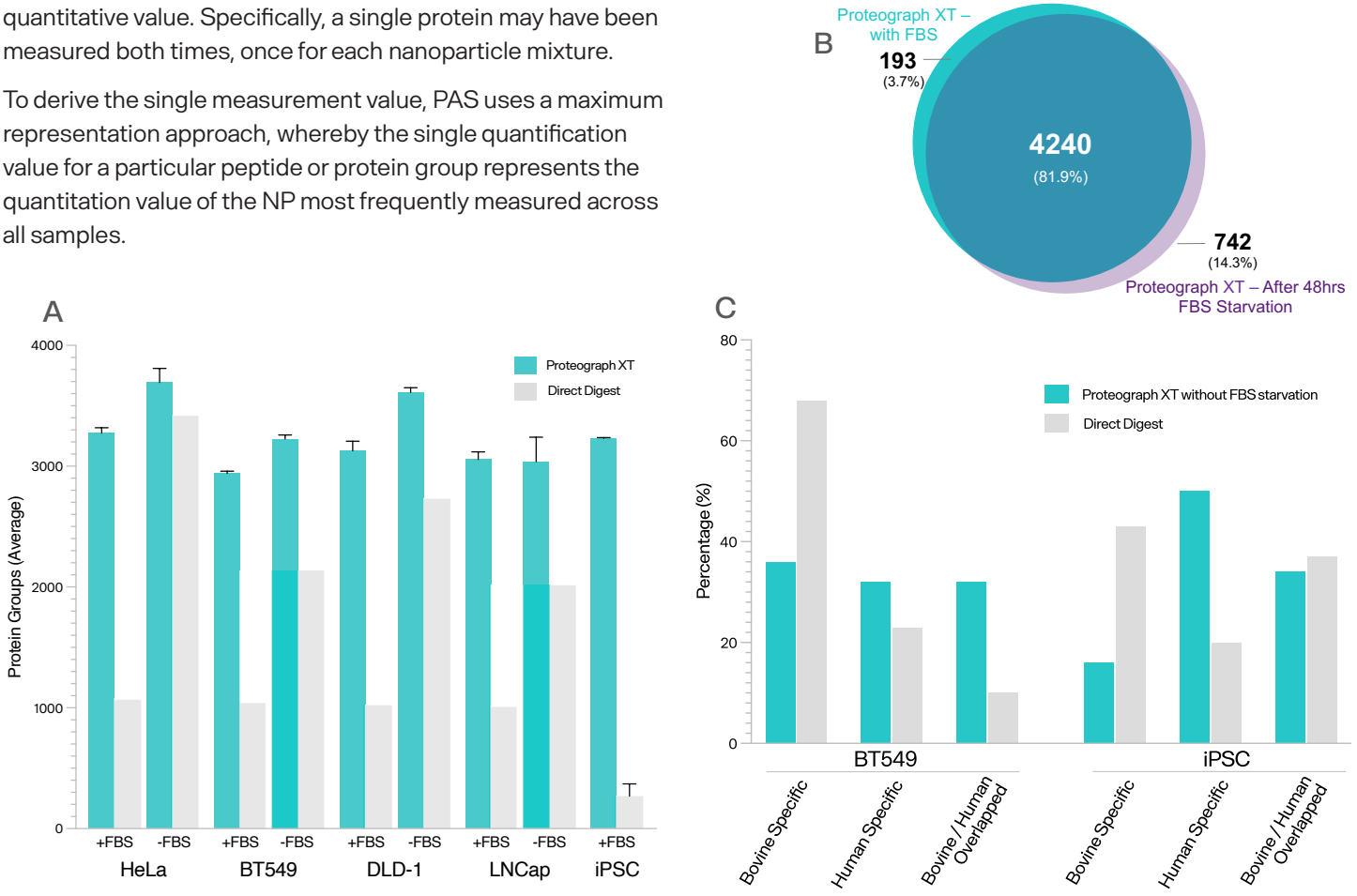


Figure 2. (A) Protein identifications for conditioned media with and without FBS using the standard Proteograph XT workflow and standard assay input volumes and comparison to direct digestion workflow. Mean protein identification numbers shown with error bars denoting standard deviation (n= 3). **(B)** Protein identification overlap for Proteograph XT workflows with and without 48 hr FBS starvation (results from four cancer cell lines). **(C)** Proportion of human and bovine derived proteins for non-depleted secretome of one representative cancer cell line (BT549) and iPSC comparing direct digestion and Proteograph XT approaches.

Of particular interest, when analyzing iPSCs, a substantial advancement, approximately 12.5-fold, was observed in favor of the Proteograph XT workflow over the direct digestion workflow. This observation suggests a selective advantage inherent to this specific secretome, which may potentially be linked to the unique properties of the media employed and its impact on protein corona formation.

Proteograph XT delivers robust analysis of non-manipulated CM secretomes

In the context of CM, we observed Proteograph XT delivers comparable identifications of protein groups across the same cell lines, regardless of whether cell lines underwent FBS starvation or not (Figure 2A). This observation holds paramount importance, as it implies that Seer's Proteograph XT workflow can facilitate in-depth proteome analysis of the secretome without necessitating FBS starvation, thereby averting potential loss of secreted factors during the experimental process (Figure 2B).

As expected, further examination of the protein identifications between bovine and human protein databases revealed that the detected direct digest proteome was composed primarily of bovine derived proteins (from FBS growth supplement) rather than the desired human proteome (Figure 2C).

Importantly, the Proteograph XT workflow significantly improved the proportion of detected human proteins while reducing the contribution from unwanted bovine proteins. Notably, these CM samples were cultivated in distinct initial media compositions ("Study design – Sample information" section). The collective evidence suggests that the Proteograph XT workflow can be used with a diverse array of media formulations and their associated characteristics.

Proteograph XT workflow distinguishes quantitative expression patterns

In our study, we systematically assessed quantitative protein expression profiles among various cell lines and growth conditions, comparing CM subjected to FBS starvation with those harvested in the presence of FBS. Principal component

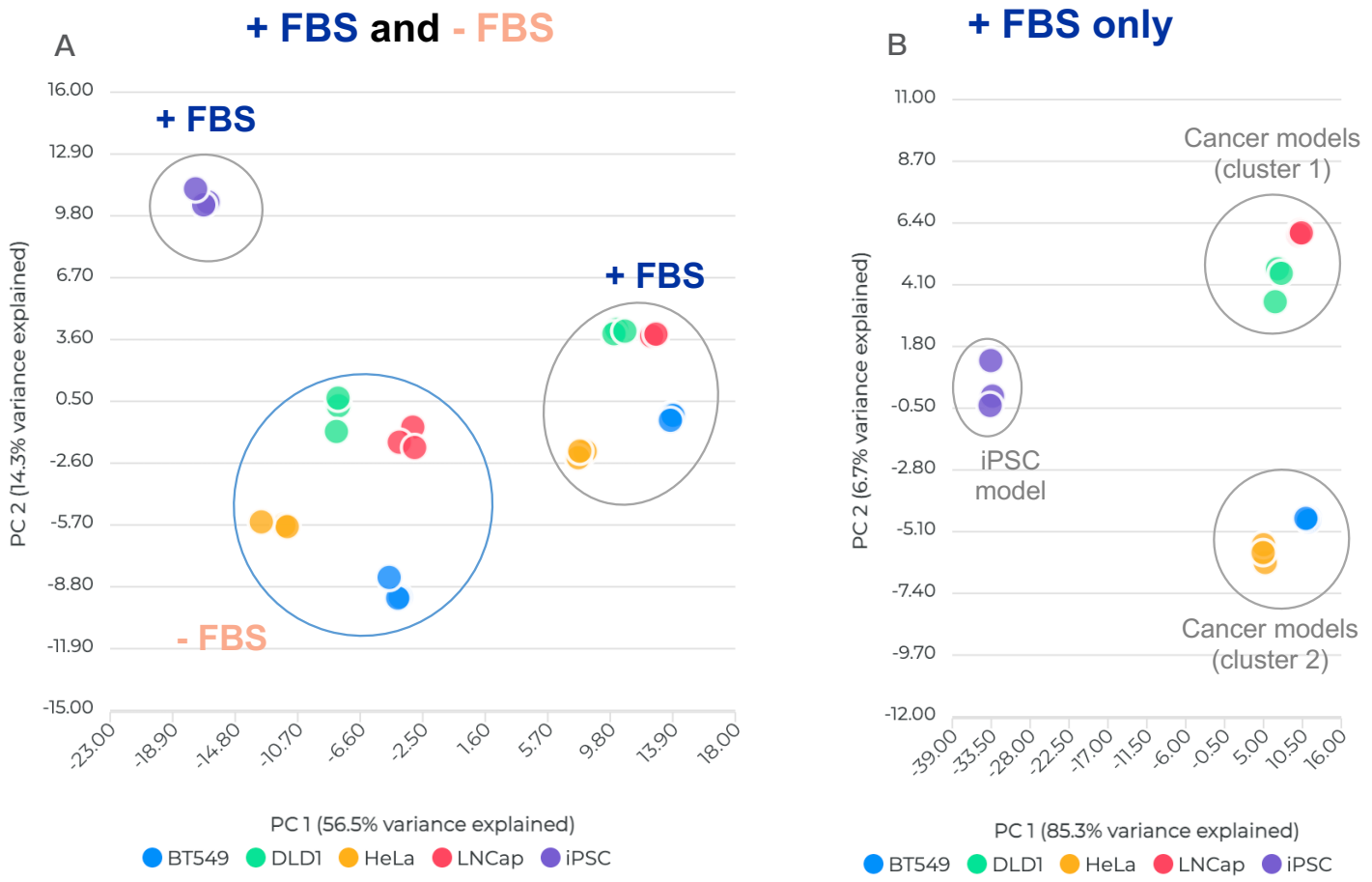


Figure 3. PCA plot for the samples processed with Proteograph XT **(A)** with and without FBS, and **(B)** with FBS only.

analysis ("PCA") was instrumental in unveiling notable distinctions in expression patterns, prominently segregating cell lines based on their response to FBS starvation **(Figure 3A)**. Moreover, Proteograph-based quantitation provided further resolution, successfully distinguishing unique profiles corresponding to each cell line, effectively classifying cancer cell line sub-groups and an additional grouping representing the iPSC profile **(Figure 3B)**.

Of particular significance, the observed clustering exhibited a remarkable level of reproducibility. Triplicates of each cell line consistently formed coherent clusters, effectively demarcating them from other cell lines even within the same subgroup. This technology consistently delivers a reproducible proteomic expression profile for samples subjected to the same growth condition, providing compelling evidence underscoring the robust reproducibility achieved with Proteograph XT.

Proteograph XT demonstrates reliable inter-replicate quantitative performance

In order to gain deeper insights into Proteograph XT's distinctive quantitative proteomic profiles characterizing the proteins within the CM secretome, we conducted a comprehensive examination of the overarching protein expression patterns across diverse starting conditions, encompassing both the presence and absence of FBS starvation.

Our findings underscored a remarkable degree of reproducibility in quantification, both evident in the context of conditions involving the presence of FBS, and after 48 hrs FBS starvation. **(Figure 4A, 4C)** This outcome is indicative of the robustness and precision of Proteograph XT quantification, as reflected by the high inter-replicate correlation coefficients observed within each individual cell line, consistently ranging from 0.97 to 0.99 across the entirety of the study. **(Figure 4B, 4D)**.

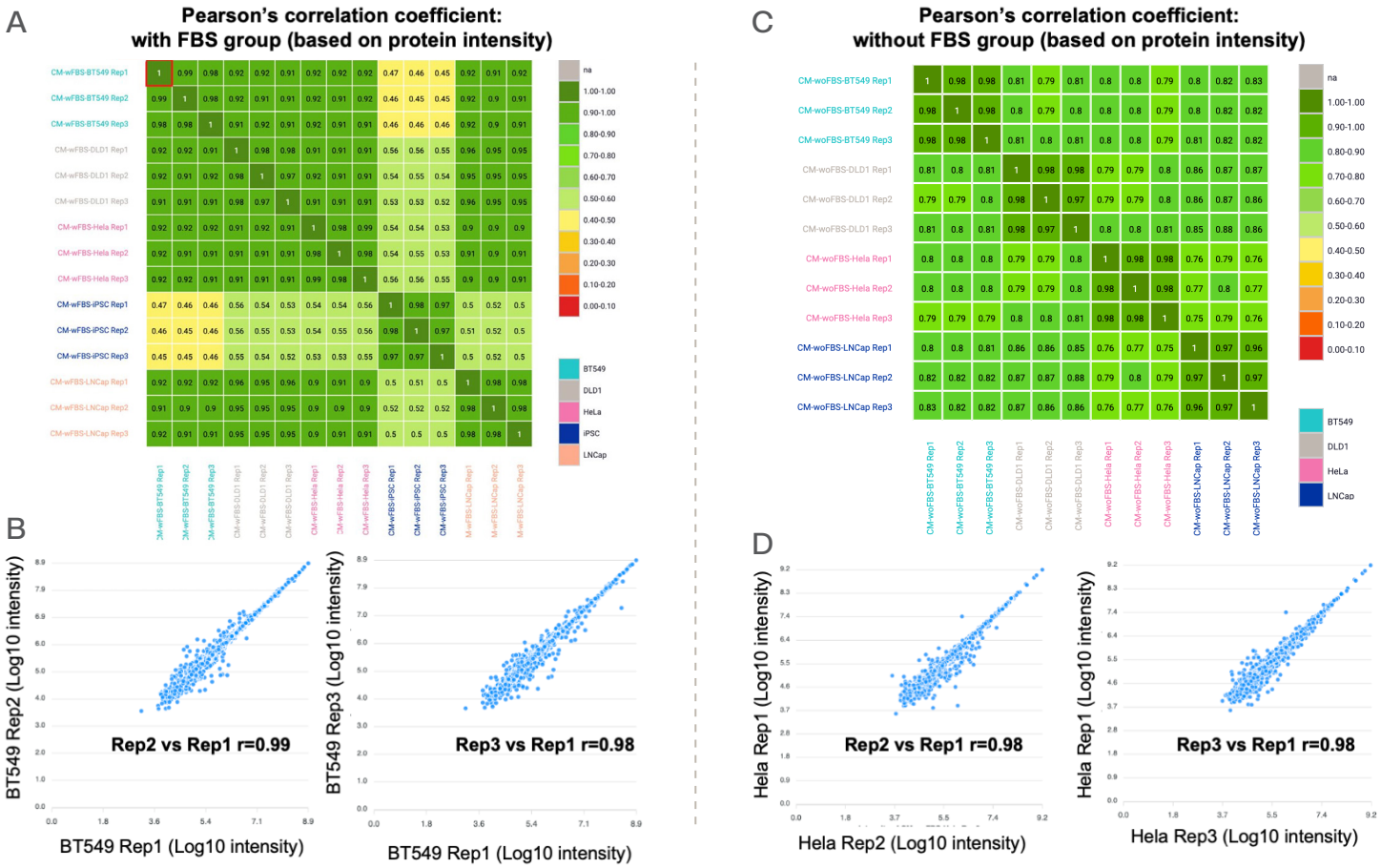


Figure 4. Pearson Correlation coefficient to assess quantitative performance of **(A)** CM with FBS groups and **(B)** CM with 48-hour FBS starvation between different cell line conditions. Log 10 intensity correlation analysis between inter-replicates for **(C)** BT549 cells with FBS, and **(D)** Hela cells with 48-hour FBS starvation.

Additionally, the exceptionally low levels of analytical variance demonstrated in our analyses significantly amplify our ability to detect nuanced differences in protein quantification arising from genuine biological effects. This heightened sensitivity is exemplified by the differentiation observed between iPSC conditions and those representatives of cancer cell lines, shedding light on the capacity of our analytical approach to discern subtle proteomic variations of biological significance.

These outcomes collectively affirm the robustness and reliability of Proteograph XT's ability to conduct CM quantitative proteomic analysis, offering a foundation for the elucidation of meaningful distinctions in the conditioned media secretome across varying experimental conditions.

Proteograph XT unlocks deeper secretome access than direct digestion

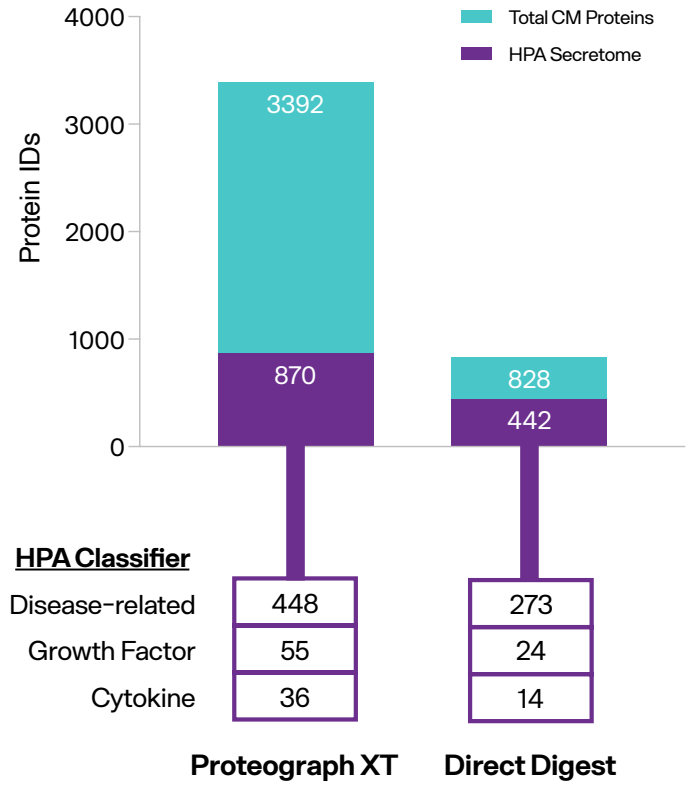
Subsequently, we subjected the protein group identifications derived from the Proteograph XT and direct digestion methodologies to annotation with the publicly accessible Human Protein Atlas (HPA) secretome database, thereby facilitating an in-depth examination of secretome proteins.

Our investigation revealed that the implementation of the Proteograph XT platform has yielded substantial enhancements, thereby allowing us an unparalleled level of access to the deeply hidden layers of the secretome, an achievement unattainable by conventional direct digestion methods.

Within the purview of this CM study, we accomplished the detection of approximately 30% of the secretome proteins annotated in the HPA secretome database. This achievement translates to the identification of 870 out of the 2793 annotated secretome identifiers (IDs) in HPA, marking an approximate 2-fold increase in the number of identified HPA secretome proteins when compared against the conventional direct digestion approach (Figure 5A).

Moreover, it is noteworthy that beyond Proteograph XT's capacity to provide a deeper dive in the identification of secreted proteins, the Proteograph XT exhibited greater coverage of disease-implicated, low abundant proteins including those categorically classified as growth factors or cytokines (Figure 5B). These results collectively underscore the profound value intrinsic to the Proteograph XT workflow.

A



B **Representative cytokines and growth factors detected in serum-containing CM**

AIMP1	AMH	BMP1	BMP4	BMP6
CCN3	CLEC11A	CSF1	CXCL1	CXCL10
EFEMP1	FGF2	FGFBP1	FGFBP3	GDF11
HTRA1	IGF1	IGF2	IGFBP1	IGFBP2
IL36A	IL6	INHBB	INHBE	LEFTY1
MDK	MIA	MIF	MSTN	NAMPT
PTN	SPP1	TAFA5	TDGF1	TGFB2
C1QTNF4	CCL14	CCL2	CCL3	CCL5
CXCL12	CXCL2	CXCL3	CXCL5	CXCL6
GDF15	GDF2	GPI	GREM1	GRN
IGFBP3	IGFBP4	IGFBP5	IGFBP6	IGFBP7
LEFTY2	LTBP1	LTBP2	LTBP3	LTBP4
OGN	PDGFA	PDGFC	PDGFD	PF4
THBS4	TIMP1	VEGFA	VEGFC	VGF
CCN1	CXCL8	HDGF	IL17C	MANF
PPBP				

Unique to Proteograph XT, not in neat digest results

Figure 5. **(A)** HPA Classifier analysis of disease related, growth factors, and cytokines by Proteograph XT and direct digestion, with Protein ID from CM with FBS conditions pulled across all cell lines, and referenced from the HPA database (HPA secretome database). **(B)** Representative cytokines and growth factors detected, with teal denoting unique identification in Proteograph XT.

extending the frontiers of our capacity to interrogate the secretome at greater depths. This capacity amplifies our ability to potentially elucidating the pivotal roles of secretome proteins in disease mechanisms and intercellular communication.

Summary

In this study, we systematically assessed the efficacy of processing secretome data derived from a variety of cancer cell lines employing the standard Proteograph XT protocol and traditional direct digestion. Our findings revealed a substantial enhancement, approximately a 3-fold increase, in the number of identified protein groups compared to the conventional direct digestion workflow for CM samples without prior 48 hrs FBS starvation. Notably, this heightened performance in protein group identification remained consistent between conditioned media samples subjected to FBS starvation and those without, suggesting the capacity of the Proteograph XT's potential to obviate the necessity of conventional FBS starvation procedures. Minimizing conditioned media handling and manipulations mitigates the potential risk of losing valuable secreted factors, which might otherwise constitute an accurate representation of the secretome proteome across distinct cell lines. It is important to emphasize that the proteome coverage achievable with the Proteograph XT workflow may vary contingent upon the specific characteristics of the conditioned media employed.

The Proteograph XT platform emerges as a robust solution for achieving deep, unbiased, rapid, and scalable proteomic analyses, all while circumventing the demands of laborious, time-consuming workflows and sample pre-processing procedures. Significantly, by simplifying the processing pipeline for CM samples, the streamlined Proteograph XT workflow preserves the high-fidelity capture of intrinsic biological signals, that might otherwise be compromised by the inclusion of additional processing and analytical stages like FBS starvation, buffer exchange, depletion, and fractionation.

In summary, our exploration of various biofluids utilizing the Proteograph XT workflow has unveiled unparalleled advancements in depth and proteome coverage compared to traditional direct digestion methods. This underscores the transformative potential of the Proteograph XT platform in the realm of secretome analysis.

Learn more at
seer.bio/products/proteograph-product-suite