Unbiased human biofluids analysis using a scalable, deep, automated, multi-nanoparticle-based proteomics workflow

Aaron S. Gajadhar*, Shao-Yung (Eric) Chen, Lucy Williamson, and Taher Elgierari

Seer, Inc., Redwood City, CA 94065, USA

The Proteograph™ Product Suite enables rapid sample processing for reproducible, deep biofluid proteomic analysis

Biospecimen proteomics including characterization of non-blood biofluids like urine are non-invasive and longitudinal sampling source, and other biofluids like cerebral spinal fluid (CSF), and cell line conditioned media (CM) have the potential to reveal new insights to human health and disease. While recent advances in sample collection and mass spectrometry have deepened our understanding of the proteomes for these samples, the field is plagued by non-standardized sample preparation and analytically complex workflows to characterize proteomes from these biofluids at acceptable depth and throughput, required for large cohort studies. In this work, we evaluated the utility of the Proteograph™ workflow, a standardized, automated multi-nanoparticle-based deep plasma proteomics approach, to interrogate a variety of conditions and sample types spanning urine, CSF, and CM.

Proteograph Product Suite

Seer core technology provides unbiased, deep, and rapid proteomics at scale

From biospecimen sample to peptides, ready for analysis on most LC-MS instruments with a variety of proteomics methods including; label-free and TMT workflows

Methods

Sample Preparation and DIA LC-MS Data Acquisition

To evaluate the performance of the Proteograph workflow, biospecimens were sourced from commercial biobanks to create representative panels of healthy and diseased samples. For human CSF, we profiled samples from donors with Parkinson's disease (PD), Amyotropic Lateral Sclerosis (ALS), and healthy individuals. Human urine samples were profiled from normal, prostate cancer, and bladder cancer donors. To assess *in vitro* secretome models, CM samples from iPSC, breast, colon, cervical, and prostate cancer cell lines were analyzed. Samples were processed directly using the SP100 Automation Instrument with Proteograph™ workflow (Seer Inc.), and in parallel with conventional neat sample digestion workflow as a control. Tryptic peptides were analyzed by DIA LC-MS analysis using an Orbitrap Exploris 480 mass spectrometer, and data processing was performed using Proteograph™ Analysis Suite (PAS).

Data Acquisition / Analysis

For Data-Independent Acquisition (DIA), 200 – 400 ng of dried, desalted peptides were reconstituted in in 4 µL of a solution of 0.1% formic acid (FA) and 3% acetonitrile (ACN) spiked with 5 fmol/µL PepCalMix from SCIEX for constant mass MS injection between samples regardless of starting volume. DIA LC-MS data was processed using Proteograph Analysis Suite 2.1. Raw LC-MS data was processed using the DIA-NN search engine (version 1.8.1) in library-free mode searching MS/MS spectra against an *in silico* generated spectral library of human protein entries (UP000005640_9606) and or human + bovine (UP000009136_9913) *in silico* generated spectral library (conditioned media).

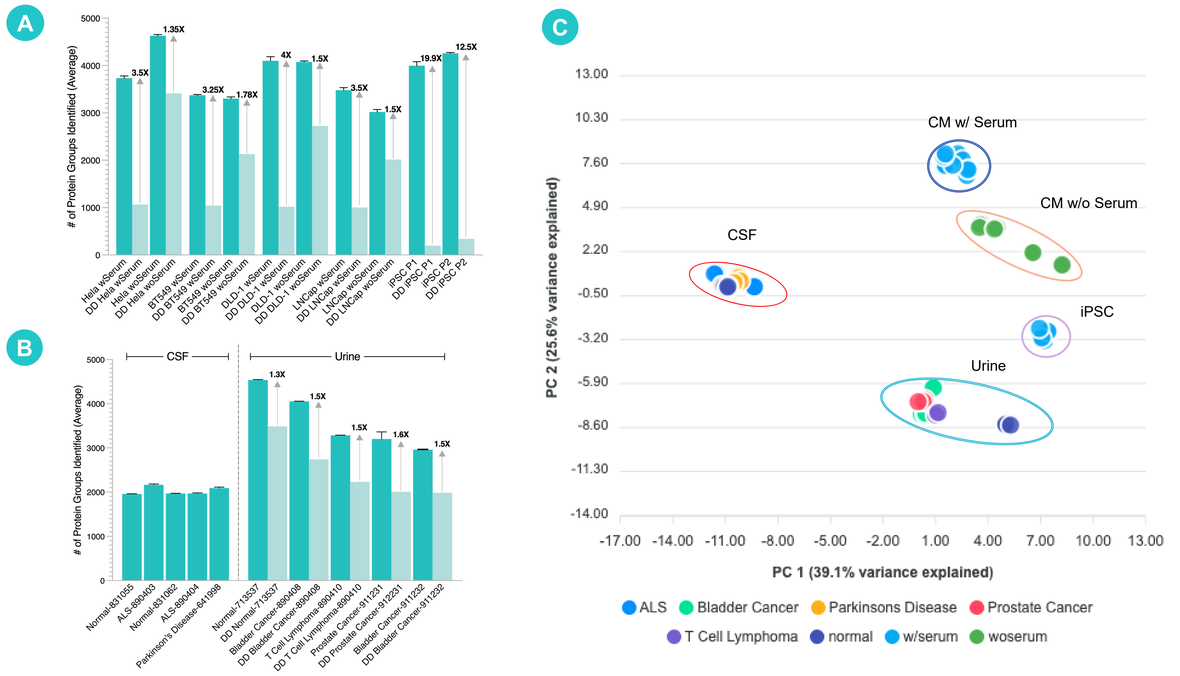

Figure 1. Proteograph Biofluid Performance Comparison and Quantitative Expression Profiles.

Protein identifications for CM samples (A), CSF and urine samples (B) using a standard Proteograph Assay protocol compared to baseline direct digest (DD), with error bars denoting standard deviation (n= 3). (C) Principal component analysis (PCA) of complete panel of biofluids indicating high reproducibility and distinct profiles between biofluid types.

Proteograph analysis of biofluids enables detection of quantitative proteome expression differences and biofluid specific functional pathway enrichments

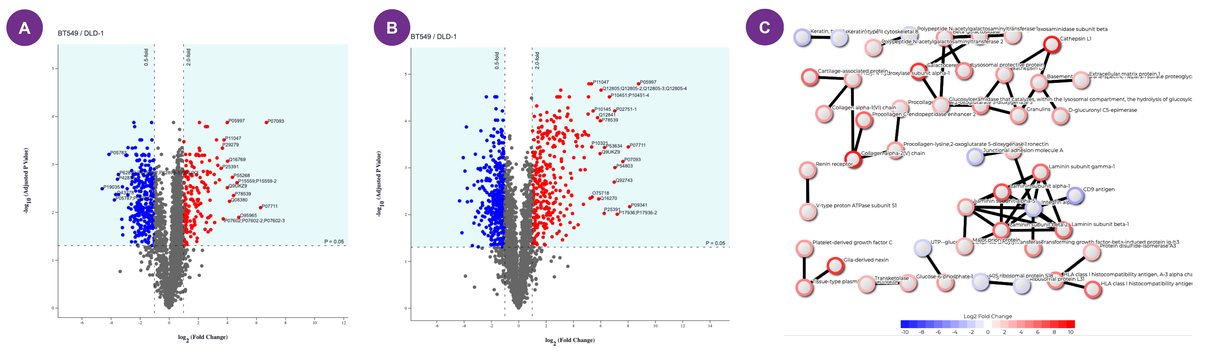

Figure 2. Differential Expression Analysis of Conditioned Media Proteomes.

(A) Example of significantly regulated protein groups detected in serum containing condition media, BT549 and DLD-1 cancer cell lines shown (>2.0 fold and adjusted p value <= 0.05) (B) Example of significantly regulated protein groups detected in serum-free condition media, BT549 and DLD-1 cancer cell lines shown. (C) STRING protein-interactome plot of significantly regulated proteins from serum-free condition media of BT549 and DLD-1 cancer cell lines.

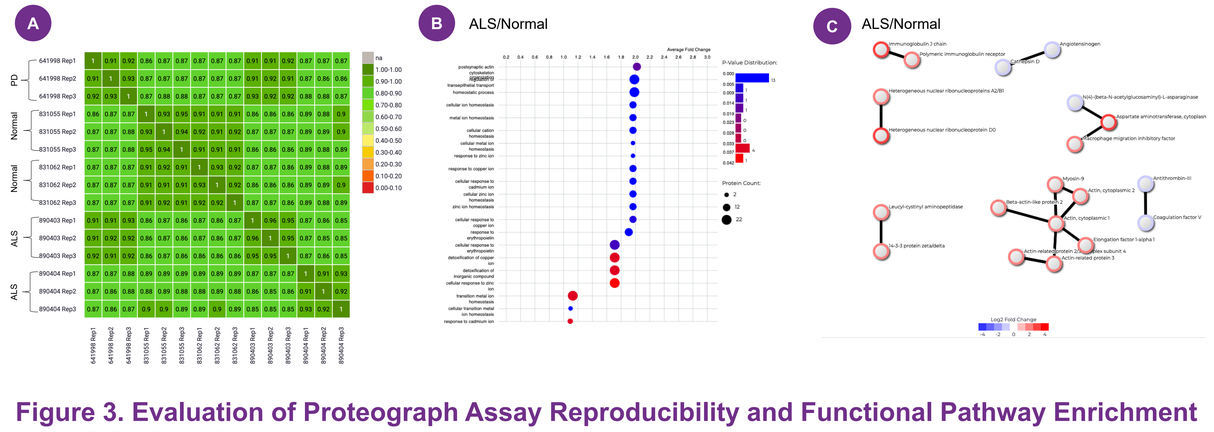

Figure 3. Evaluation of Proteograph Assay Reproducibility and Functional Pathway Enrichment in CSF Biofluid Samples.

(A) Jaccard Index analysis of triplicate samples from CSF biospecimens representing ALS, PD and normal donors (similarity coefficient shown). (B) Gene Ontology enrichment analysis of significantly regulated proteins in ALS compared to normal CSF samples. (C) STRING protein-interactome plot of significantly regulated proteins in ALS compared to normal CSF samples.

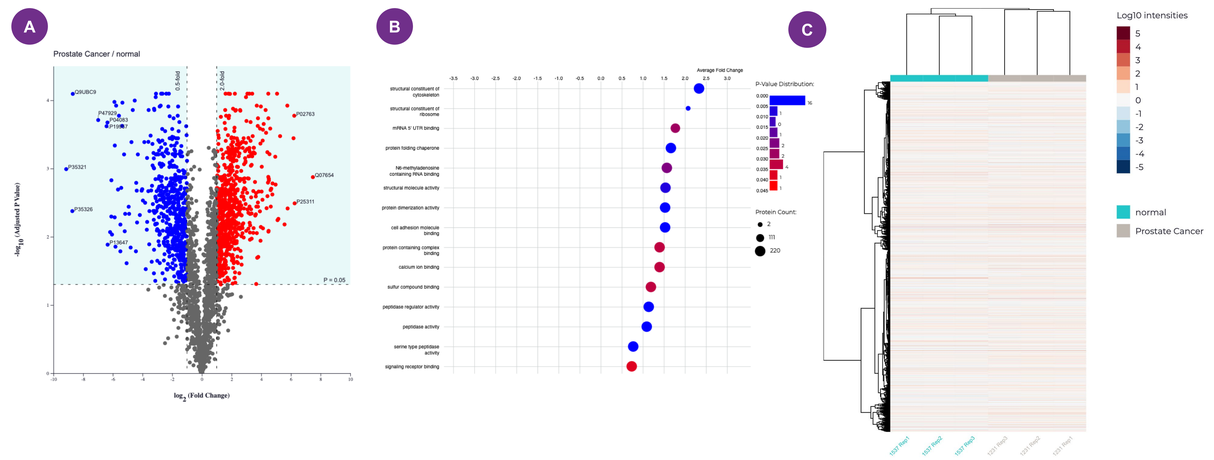

Figure 4. Differential Expression Analysis of Urine Proteomes from Prostate Cancer, Bladder Cancer and T-Cell Lymphoma Specimens.

(A) Example of significantly regulated protein groups detected between urine biospecimens representing prostate cancer compared to normal donors (>2.0 fold and adjusted p value <= 0.05). (B) Gene Ontology enrichment analysis of significantly regulated proteins in urine samples from prostate cancer patients compared to normal donors (C) Agglomerative hierarchical clustering of triplicates of normal urine samples and prostate cancer urine samples.

Conclusions

- The standard Proteograph Assay for plasma enables biofluid sample processing, including serum, CM, urine, and CSF with no additional pre-processing steps.
- Proteograph workflow improves proteome coverage for CM, Urine, and CSF biofluids compared to conventional neat digest workflows (1.5x - 20x improvement).
- Proteograph conditioned media processing workflow is compatible with serum containing conditions, obviating the need for FBS starvation and additional sample manipulations for deep proteomic analysis.
- Deep proteomic coverage, high assay reproducibility, and proteome pathway insights are possible with the standard Proteograph Assay for CSF, Urine, and condition media samples.

Publications

Seer, Inc., Redwood City, CA 94065, USA | *agajadhar@seer.bio