Deep and Unbiased Plasma Protein Profiling of Alzheimer's and Mild Cognitive Impairment Subjects with a Novel Multi-nanoparticle Approach

Siddiqui A*, *Benz RB, Goldberg M, Cuevas JC, Manning W, Zhao X, Platt TL, Ko, M, Ha HJ, Liou H, Elgierari EM, Figa M, Xia H, and Farokhzad OC*

Deep and unbiased plasma proteomics for disease cohort studies at scale

Introduction

Blood plasma is a rich source of protein biomarkers for early detection of diseases, but its large dynamic range of protein concentrations necessitates complex workflows and trade-offs between throughput, scalability, coverage, and precision. Here we use a deep and quantitative proteome profiling platform, Proteograph™ Product Suite, which leverages multiple nanoparticles, engineered with distinct physicochemical properties to provide broad coverage of the plasma proteome at scale. In this study, we aim to identify protein biomarkers for Alzheimer's disease (AD) from blood with this untargeted plasma protein profiling approach, for AD and Mild Cognitive Impairment (MCI) condition.

These analyses identified novel a combinations of known and potential new candidate plasma protein markers, confirming the Proteograph platform's ability to generate profiling data in a deep, broad, and rapid fashion, enabling large-scale studies to detect novel insights with clinically relevant potential.

Proteograph solution

Proteograph Product Suite provides untargeted, deep, and rapid proteomics at scale

From sample to peptides, ready for analysis on most LC-MS instruments with a variety of proteomics methods

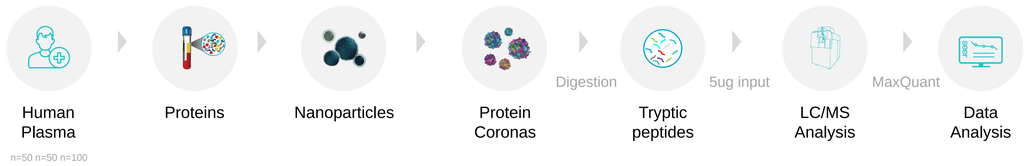
Methods

Plasma samples from 200 subjects comprising 50 AD, 50 MCI, and 100 healthy controls were analyzed using Proteograph plasma protein profiling platform[1].

Using 5 injections per sample, proteins were quantified using two liquid-chromatography mass-spectrometry (LC-MS) methods:

1. SWATH DIA workflow (50 AD, 50 MCI, 100 healthy samples); and
2. ddaPASEF workflow (100 AD, 100 healthy samples). Normalized peptide intensities were used to develop models for class discrimination.

A SWATH DIA Workflow



B ddaPASEF Workflow

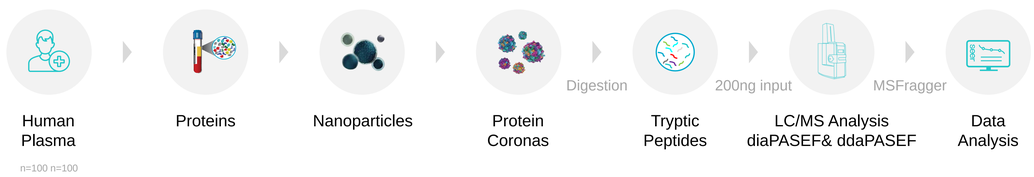


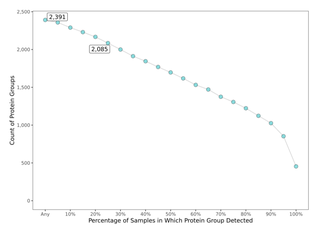
Figure 1. Variable Window (vw) SWATH DIA and ddaPASEF Workflows.

A) in the label-free DIA workflow, n=100 healthy, n=50 AD, and n=50 MCI plasma samples were processed using the Proteograph workflow. Following processing, the 96-well plate with digested peptides was analyzed with a microflow vw SWATH on a TripleTOF 6600 (SCIEX). **B)** In the ddaPASEF workflow, n=100 healthy human plasma and n= 100 AD plasma samples were processed using the Proteograph workflow. Following processing, LCMS analysis of digested peptides were performed with a microflow ddaPASEF on a Bruker's timsTOF pro. DDA data were analyzed with MSFragger, applying 1% FDR cutoff at the protein and peptide levels.

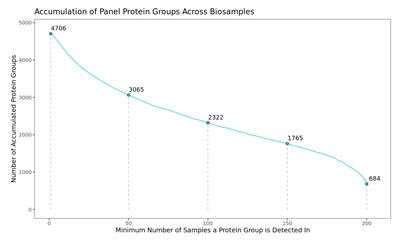
Deep proteomics enhances disease classification and biomarker characterization of AD cohort

Data and classifier summaries

A **Original study (DIA):**
2,391 protein groups are found across all 200 subjects and **2,085** in **25%** of all the subjects



B **New study (DDA):**
4,706 protein groups are found across all 200 subjects and **3,065** in **25%** of all the subjects.



C **False positive fraction**

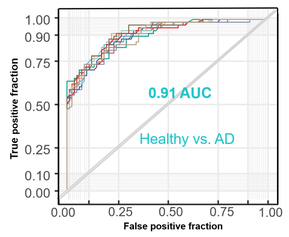
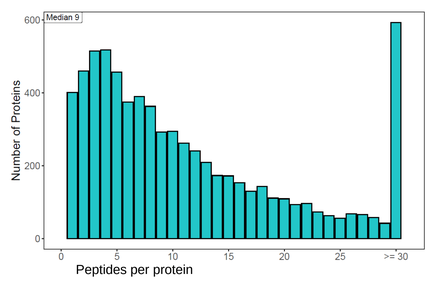


Figure 2. Protein groups identification across initial and new study with the resulting classifier.

A) Number of protein groups detected across various fractions of the 200 initial subjects (100 healthy, 50 MCI, and 50 AD). **B)** Number of protein groups detected across various fractions of the 200 initial subjects (100 healthy and 100 AD). **C)** ROC plot of the 10x10 cross validation of the results in the initial study. Further verification studies are required to validate the model.

Multiple peptides per protein detected for most proteins

A



B

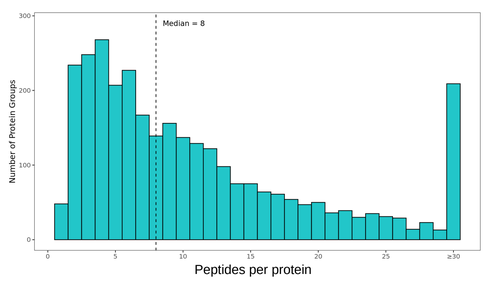
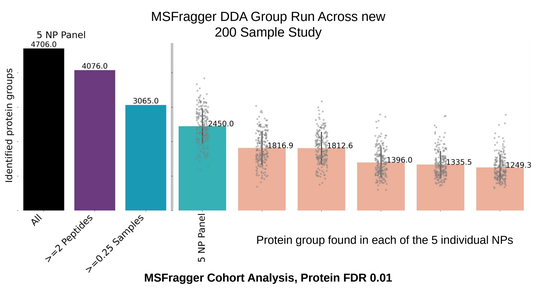


Figure 3. Number of peptides per protein detected for each study.

A) The original study found a median number of 9 peptides per protein. **B)** The new study found a median number of 8 peptides per protein, computed from proteins present in ≥25% of bio-samples..

1,359 unique variant peptides detected across the 200 samples; scaled analysis in ~ 24 hours

A



B

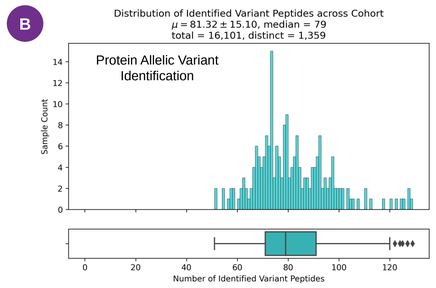


Figure 4. In addition to 4706 reference protein groups, an additional 1,359 unique peptide variants were detected yielding > 6,000 protein groups.

A) at a 1% FDR we find 4,706 protein groups across all samples, 4076 proteins at least with 2 unique peptides and 2450 per sample.
B) a histogram of the number of peptide variants found in each sample.

Conclusion

200 samples
1000 injections
4700 protein groups
24 hours

- The 200-sample DIA study was processed in approx. 6 weeks, generating unbiased proteomics data including 2,085 proteins and 15,661 peptides quantified in at least 25% of the samples.
- The 200 sample DDA study found > 6000 reference and variant protein groups at a 1% FDR. Scaled cloud processing allowed this analysis to complete in 24 hours.
- Proteograph is enabling deep, broad, and rapid processing of samples enabling larger and more powerful studies per unit time and resource

References
[1] Blume et al. *Nat. Comm.* (2020)

Acknowledgements
Becky Carlyle[1],Sudeshna Das[1], Matthijs B. De Geus[1], Bradley T. Hyman[1], Clemens R. Scherzer[2], Pia Kivisäkk[1], Steven E. Arnold[1]

[1] Massachusetts General Hospital
[2] Brigham and Women's Hospital

Publications

SBIR · STTR America's Seed Fund™ POWERED BY SBA
Funded by SBIR grant

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