Nanoparticle-based method identifies over 2200 proteins in a cardiovascular disease study covering known biomarkers among other differentially expressed proteins

Michael Burgess[1], Hasmik Keshishian[1], Michelle Dubuke[2], Juan Cruz Cuevas[2], Laurie Farrell[3], Debby Ngo[3], Karsten Krug[1], DR Mani[1], Robert Gerszten[3], Steven A Carr[1]

[1] Broad Institute of MIT and Harvard, Cambridge, MA; [2] Seer Inc., Redwood City, CA; [3] Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

Overview

- Plasma is an ideal sample for clinical proteomics but analysis is confounded by the intrinsic high dynamic range of plasma proteins making confident and reproducible measurement of low abundance proteins difficult.
- In a previous cardiovascular disease study, our proteomics workflow involving abundant protein depletion, iTRAQ-labeling, off-line basic reversed-phase fractionation and analysis of 30 fractions by LC-MS/MS quantified over 3500 proteins including the measurement of clinical markers such as troponins I and T, but at the cost of extended labor, cost and low throughput[1,2]. Specifically, to process 16 patient samples required 7-10 days of lab work followed by 120 LC-MS/MS runs.
- Here we compare this strategy with the newly developed Proteograph™ Product Suite (Seer) which simplifies sample processing prior to MS analysis by using nanoparticles in a robotic platform with much higher throughput.

Methods: Seer Proteograph Workflow

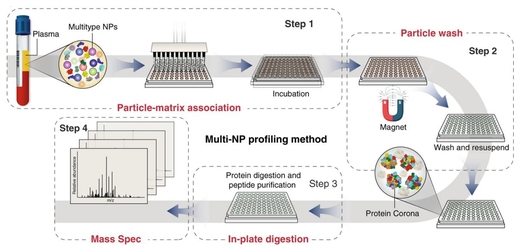


Figure 1: Seer Proteograph Plasma Sample Processing Scheme[3]. This plate-based workflow includes the processing of 16 total samples across 5 NPs along with Plasma Controls for digestion, NP-binding and SPE. Samples and reagents are loaded in 30min followed by the 6.5 hour process. Each patient analyses is composed of 5 LC-MS/MS runs or 80 total for all 16 samples.

- Initial evaluation of the Proteograph plasma processing system was done using a plasma control sample provided by Seer. Method reproducibility was evaluated across 15 samples processed for 3 consecutive days.
- Patient plasma samples included in this study were comprised of peripheral blood draws at 5 time points (baseline-pre-heparin, BL-post heparin, 10min, 1hr and 4hr after alcohol ablation) across 3 patients undergoing a therapeutic septal alcohol ablation heart surgical procedure[2]. These planned myocardial infarctions (PMI) act as a human model for longitudinal myocardial injury.
- Following the standard protocol, 200uLs of each sample is arrayed across 5 wells prior to incubation with Seer's panel of 5 nanoparticles (NP) (Step 1). After extensive washing to remove nonspecific binders (Step 2), samples undergo an on-bead, 3 hour Trypsin/Lys-C digestion (Step 3) using an adapted EasyPep 96 MS Sample Prep Kit (ThermoFisher).
- Each sample was then desalted using an on-board positive pressure manifold for sample cleanup.
- Total protein recovered post process ranged from few hundred nanograms to 5 microgram per NP.
- Data were acquired on a Thermo Exploris LC-MS/MS system using a 1 hour gradient. Each sample analysis is comprised of 5 LC-MS/MS runs, one for each NP subpopulation. Data were searched with Spectrum Mill software using the label free quantification (LFQ) mode for differential analysis of proteins identified with 2 or more peptides.

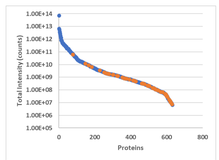
Initial Evaluation

Enriching plasma with the NP panel increases depth of protein coverage by 5X relative to direct analysis; proteins span 8 orders of magnitude

Table 1: Proteins identified using each of the 5 individual nanoparticles versus direct analysis of unprocessed Plasma Control sample. Each individual NP showed a 2 to 3-fold increase in the number of proteins ID'ed with a combined 5-fold increase in distinct proteins (2 or more peptides/protein) using them as a panel (ca. 1400 proteins vs. 300 proteins). Each individual NP captured unique proteins with most showing some overlap with at least another member of the panel.

Nanoparticle	Protein ID's	Fold Increase	Unique ID's
No NP	296	-	
NP1	549	1.9	31
NP2	625	2.1	51
NP3	797	2.7	137
NP4	608	2.1	90
NP5	930	3.1	254
Combined	1376	4.6	

Figure 2: Dynamic range of total intensities for an example NP particle. Total ion current intensities for all proteins observed (2 or more peptides) using nanoparticle 2 (NP2) for Plasma Control showing a span of 8 orders of magnitude. Those proteins observed to be unique to NP2 only are highlighted in orange. There appears to be no protein abundance bias for any of the NP enrichments.



Reproducibility

Consistent protein ID's, tight clustering of each NP and median CV of 25%

Inter-day analysis over 3 consecutive days (5 replicates/day)

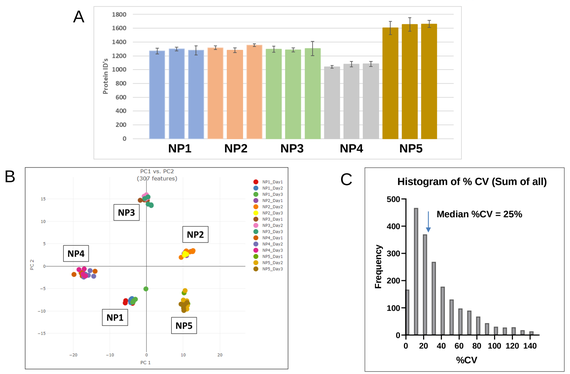
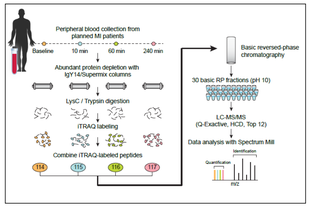


Figure 3: (A) Proteins identified in all 5 individual NPs across 3 consecutive days for Plasma Control with 5 replicates/day. NP4 showed consistently the lowest number of identifications while NP5 always resulted in the most. **(B)** PCA analysis of same data showing tight clustering of individual surface chemistries and separation of them on PC1 and PC2. **(C)** Histogram of %CV for all NPs combined across all three plates. CVs are calculated for the protein intensities across all 5 nanoparticles per replicate. Median CV was 25% while over 50% of the data has a CV of less than 30%.

Proteograph Applied in Planned Myocardial Infarction

Figure 4: PMI deep profile plasma processing workflow. The original iTRAQ study comprised of peripheral blood draws for 4 patients taken at 4 time points: baseline taken pre-administration of heparin block (BLpre), and then 10min, 1hr and 4hr post alcohol ablation heart surgical procedure (figure 4)[2].

In the present study employing the Proteograph, we analyzed the same 4 plasma timepoint samples from 3 of the 4 original patients, and we also analyzed a 5th time point sample taken immediately post heparin application but pre-alcohol ablation from these same patients (BLpost).



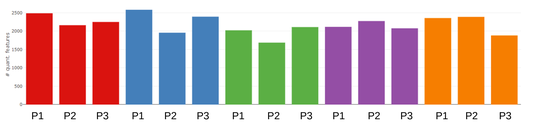


Figure 5: Bar graph protein ID's across 3 PMI patients. On average, 2200 proteins were detected in each individual patient sample with a total of over 4000 observed across all 15 PMI samples processed using the Proteograph.

Correlation of patient samples for every time point is high, PC1 separates patients, and PC2 time points within each patient

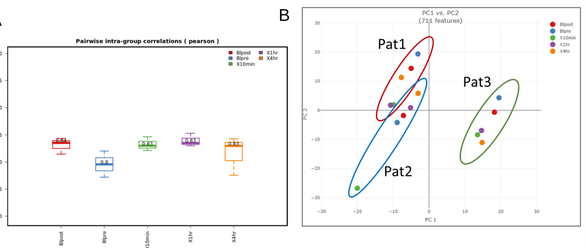


Figure 6: (A) Box-Whisker plot of correlation coefficient distribution grouped by timepoints. Each group shows a similar correlation of ca. 0.83 and tight distributions. **(B)** PCA analysis of the samples shows separation of the patients on PC1 and timepoints on PC2.

Cardiac Troponins robustly detected in all 3 PMI patient samples as early as 10min post-injury – much earlier than in our original deepscale iTRAQ experiments

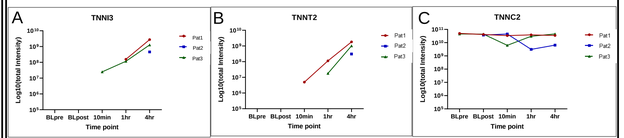


Figure 7: Longitudinal intensity changes for 3 troponin markers, TNNI3 (A), TNNT2 (B), TNNC2 (C) Cardiac Troponin I and T were detected and observed to rise early while skeletal Troponin C2 remained flat as previously observed. Moreover, cardiac Troponin I and T were detected robustly across the patient samples as early as 10 minutes after the procedure in one of the PMI patients and at 1 hour after the injury in two out of three patient samples.

Correlation of Seer Proteograph results to iTRAQ discovery

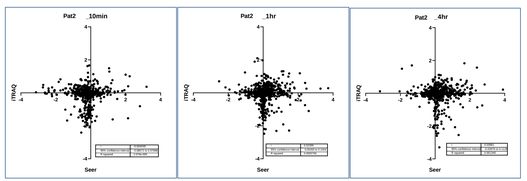


Figure 8: Scatter plots of ~ 450 proteins detected in Proteograph and iTRAQ discovery for one of the PMI patients at 10min A), 1hr B) and 4hr C) relative to baseline. More than 150 of the previously observed ca. 300 regulated proteins were detected in the current analysis and showed similar trends.

Conclusions

- An average of 2200 proteins were detected in each patient plasma sample using the Proteograph system; a total of ca. 4000 proteins across the 15 PMI samples.
- Cardiac Troponins, known markers of myocardial infarction, were robustly detected as early as 10 min post ablation using the Proteograph system and increased with time as expected.
- Troponins were detected at earlier times post ablation than in the original iTRAQ study and with many more peptides/protein.
- More than 150 of the previously observed ca. 300 regulated proteins were detected in the current study and showed similar trends.
- Comparison of pre- and post- heparin baseline samples revealed a large number of heparin binding proteins in the Proteograph data, and AEBP1 and FHL1 proteins were detected and showed trends across the time course similar to the iTRAQ discovery study (data not shown).
- New potential markers of PMI were detected using the Proteograph that need to be validated.

References

[1] Keshishian et al., *Nature Protocols* Vol.12 No.8, 2017.

[2] Keshishian et al. *Molecular and Cellular Proteomics* 2015

[3] Blume et al. *Nature Communications* **11,** 3662 2020

Disclosure: S.A.C. is a scientific advisor to Seer