Matrix-Matched Calibration Curves Provide Verification of Quantitative Data-Independent Acquisition Techniques for Deep Plasma Proteomics

Seth Just*, *Drew Nichols, Iman Mohtashemi, Ting Huang, Jian Wang, Margaret Donovan, Shadi Ferdosi, and Daniel Hornburg*

Calibration Curves at Proteome Scale

- **Deep and unbiased proteomics** of biofluids by Data Independent Acquisition Mass Spectrometry (DIA-MS) and the Proteograph™ Product Suite[1] offers significant **opportunities for protein biomarker discovery at scale.**
- Proteomics analysis techniques **vary** in their consistency, **accuracy**, and precision. **Quantitative verification** is **essential** to ensure reliable results.
- **High complexity** and extreme **dynamic range** across proteins in biofluids can cause notable **matrix effects**, where a single analyte's measurement is affected by the presence of other analytes.
- We show **the Proteograph™ workflow gives well-calibrated quantitative results** and characterize matrix effects in neat plasma and Proteograph samples.

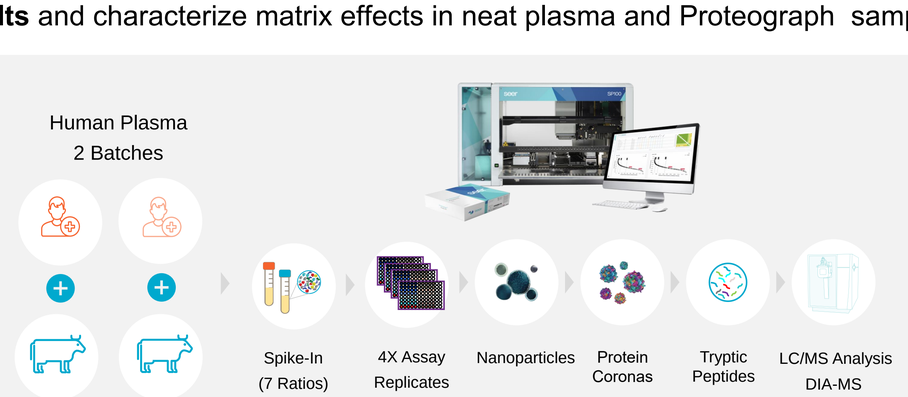

Figure 1. Experimental design for calibration curve experiment. Using a matched matrix ensures similar sample complexity throughout the calibration range, allowing accurate estimation of the noise and linearity of the system's quantitative response to protein concentration.

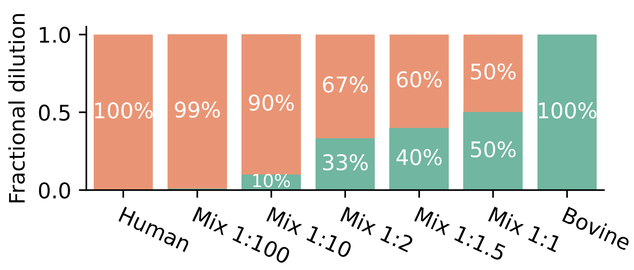

Figure 2. Dilution series of 7 ratios. Asymmetric dilution allows analyzing Bovine-specific proteins to determine the quantitative response of low-concentration proteins in a complex matrix like human plasma to evaluate quantitative performance of our analysis pipeline. This is of particular interest for biomarker discovery because most biomarkers are expected to have low concentration.

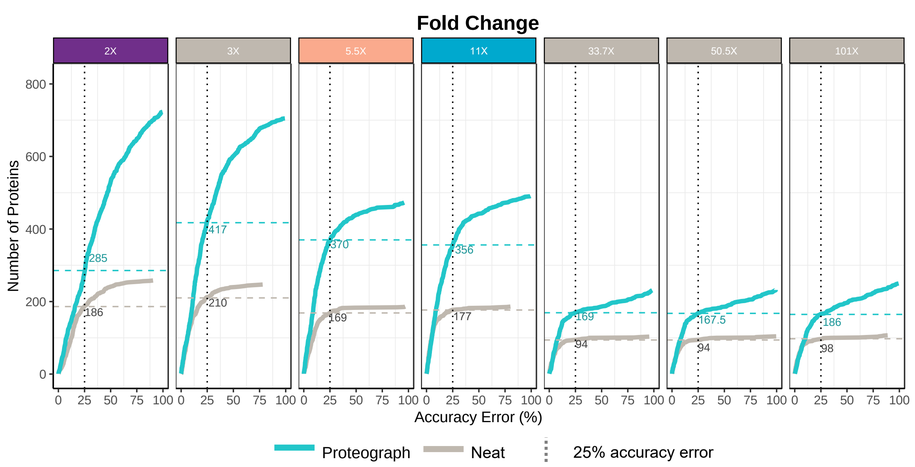

Figure 3. Proteograph workflow outperforms neat plasma identification performance at any level of accuracy and across all ratios. Cumulative protein counts across accuracy cutoffs (MAPE) for dilution pairs in Proteograph workflow and neat plasma.

Quantitative Plasma Proteomics with the Proteograph Workflow

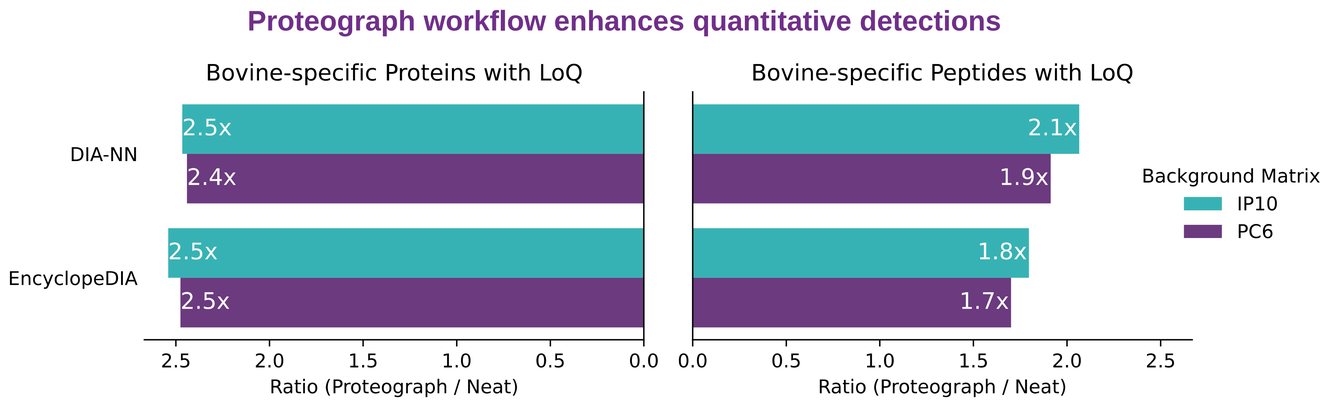

Figure 4. Proteograph workflow increases the number of quantitative signals over neat plasma. Calibration curve results determine lower limits of quantification (LoQ) for proteins with a range of linear intensity response exceeding noise by two standard deviations and with CV < 20%. **Significantly more proteins (2.5x) have reliable quantitative responses** with the Proteograph workflow than in neat plasma.

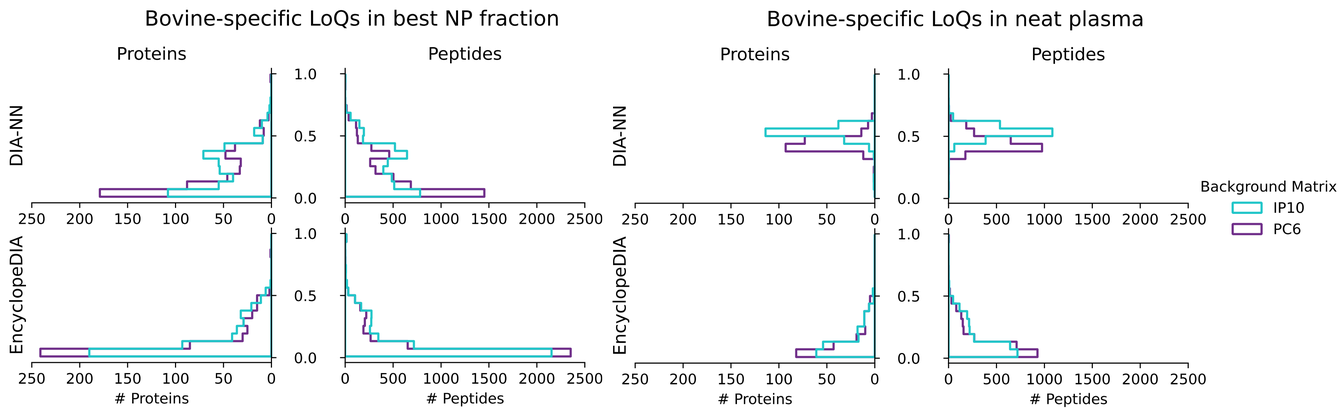

Figure 5. Lower limits of quantification (LoQ) improve with Proteograph workflow (left) vs. neat plasma (right). LoQ is the lowest level of dilution with a reliable quantitative response (lower is better). Analytes are harder to quantify in a higher-complexity background matrix (IP10), but this effect is reduced by the Proteograph. EncyclopeDIA gives lower LoQs and less difference between backgrounds for neat plasma, suggesting it's better able to reject interfering signals.

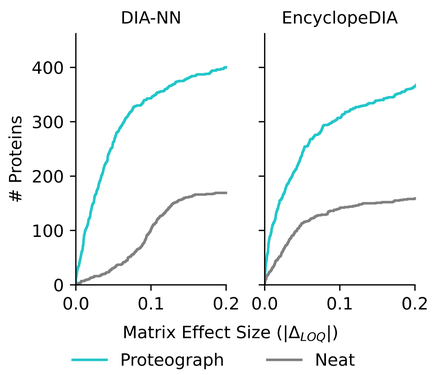

Figure 6. Proteograph reduces quantitative matrix effects, suggesting that **nanoparticles can remove signals that confound quantification in neat plasma**. Each protein's quantitative response to concentration is affected by the background matrix due to effects such as ionization suppression, fragment ion interference, or (in the Proteograph workflowl) competition during nanoparticle corona formation. Changes in LoQ between background matrices represent the total effect of a different matrix on each protein's range of reliable quantitative response. **The number of proteins with small changes between backgrounds is significantly increased by the Proteograph workflow.** EncyclopeDIA shows smaller matrix effects in neat plasma but gives similar results to DIA-NN for the Proteograph data.

Methods

Two batches of pooled control human plasma ("IP10" and "PC6") were mixed with Bovine plasma at seven ratios and processed in quadruplicate by the Proteograph Product Suite, using five physiochemically distinct nanoparticles (NPs). Each biosample/NP/ratio combination was analyzed by DIA-MS on a Orbitrap Exploris 480 MS with an UltiMate 3000 HPLC (Thermo Fisher Scientific) employing a 30-minute LC gradient at 1µL/min flow rate. We additionally analyzed samples prepared without nanoparticle enrichment ("neat plasma") using the same digestion and desalting technique.

All 373 MS acquisitions were searched by DIA-NN and EncyclopeDIA using a sample-specific Data Dependent Acquisition (DDA) spectral library and employing a 1% FDR threshold at the level of PSMs (DIA-NN only), peptides, and proteins. Proteograph workflow results are from a search of all acquisitions, with re-run, control, and neat samples excluded from consideration after FDR control. Neat plasma results are from a separate search. Figures of merit for each biosample/peptide/NP combination were computed by the matrix-matched calibration curve approach[2] which requires excluding peptides shared between proteomes from consideration (after FDR control).

- **The Proteograph workflow enables quantitative plasma proteomics with a combination of depth and scale**
- **Seer multi-nanoparticle workflow uniquely enables deep, unbiased biomarker discovery in biofluids**

References

1. *Blume et al. Nature Communications (2020)*
2. *Pino et al. J. Proteome Res. (2020)*

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