The Nanoparticle-Based Plasma Proteomics Workflow enables the Investigation of Glycoproteome

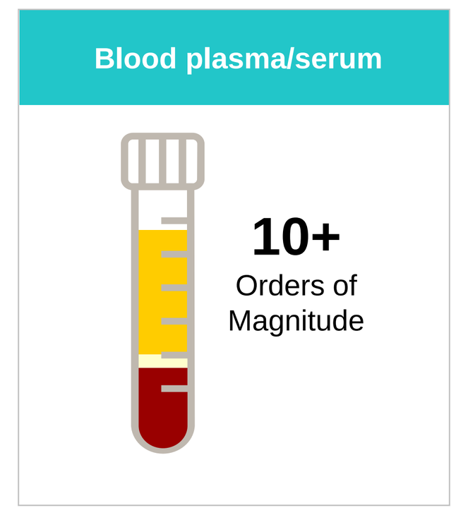
Shadi Ferdosi*, Moaraj Hasan, Eric Chen, Alexey Stukalov, Iman Mohtashemi, Jian Wang, Khatereh Motamedchaboki, and Daniel Hornburg

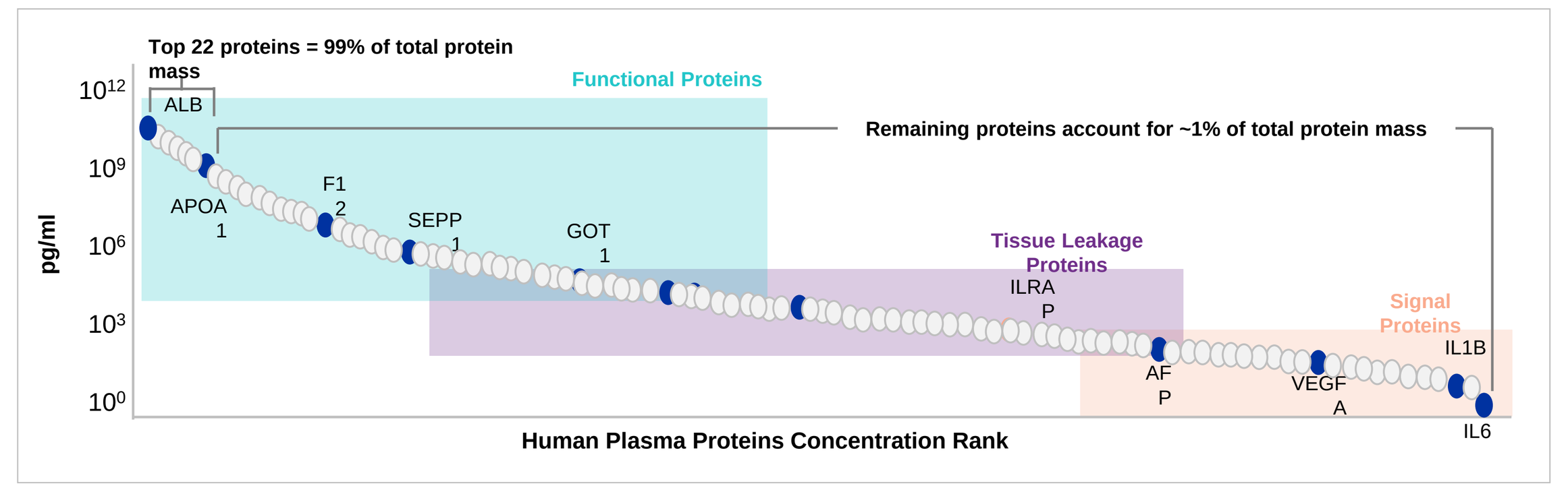
Investigating Plasma Glycoproteome with Proteograph™ Product Suite

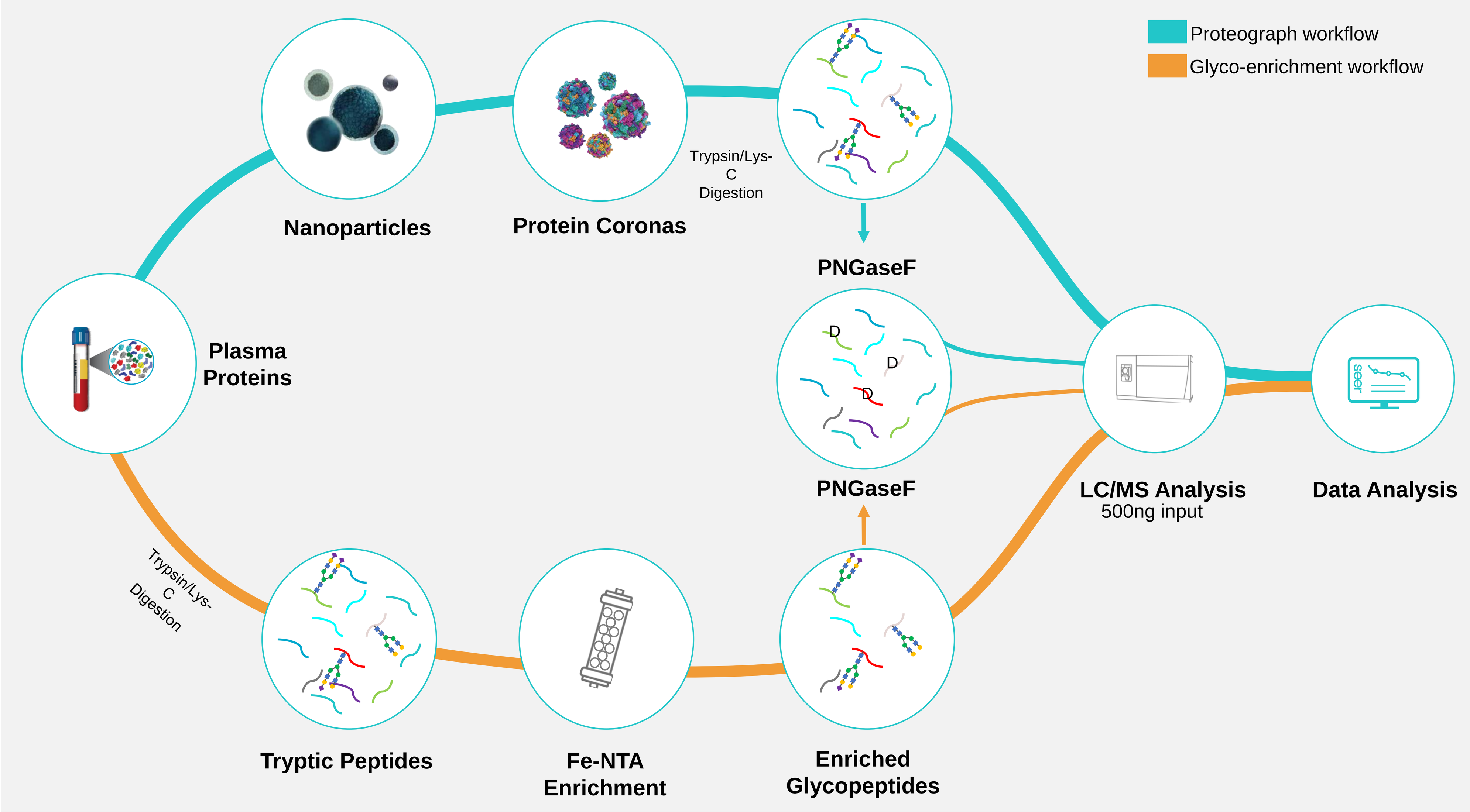
To overcome limitation of deep plasma proteomics in large cohorts, we have developed a fast and scalable technology that employs intricate protein-nano interactions[1-3]. Introducing a nanoparticle (NP) into a biofluid such as blood plasma leads to the formation of a selective, specific, and reproducible protein corona at the nano-bio interface driven by the relationship between protein-NP affinity, protein abundance, and protein-protein interactions. Since NP-coronas differentially interrogate complex samples at the proteoform level and plasma proteins are often glycosylated, here we investigate whether an automated NP-corona based proteomics workflow can interrogate the plasma glycoproteome without subsequent enrichment of glycosylated peptides. Protein glycosylation states can provide diagnostic evidence where total protein abundance is uninformative. Improved methods for profiling the plasma glycoproteome, specially at the lower abundance level thus can have a major impact in biomarker discovery.

Methods

We have investigated the utility of NPs in capturing glycoproteins in comparison to neat plasma and a conventional glyco-enrichment workflow using magnetic Fe-NTA beads. Samples were analyzed with Orbitrap Lumos mass spectrometer and UltiMate3000 Dionex LC system using 60 min DDA sample-to-sample runs. We evaluated depth, dynamic range, coverage, and precision of quantification at a wide range of concentrations for each NP.



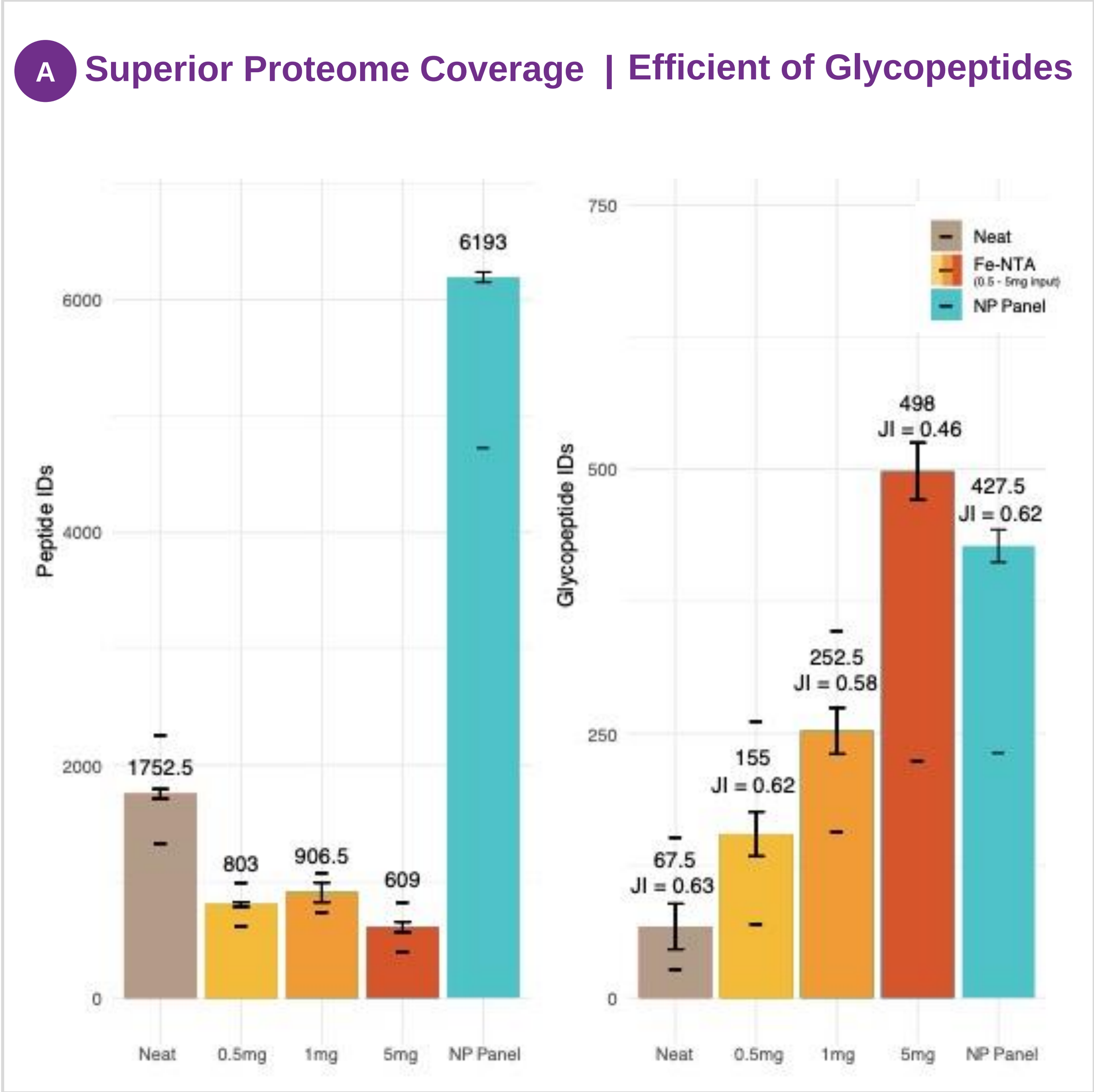




Results

By compressing the dynamic range and making the low abundance proteins and corresponding peptides more visible to the downstream LC-MS, NPs facilitate the detection of peptides with increased sensitivity, efficiently, and robustly even without subsequent enrichment. Based on the physicochemical properties of NPs, they capture distinct sets of glycoproteins at the lower abundance range, significantly enhancing the coverage of these proteins in blood plasma.

Reproducible Identification and Quantification of Plasma Glycopeptides



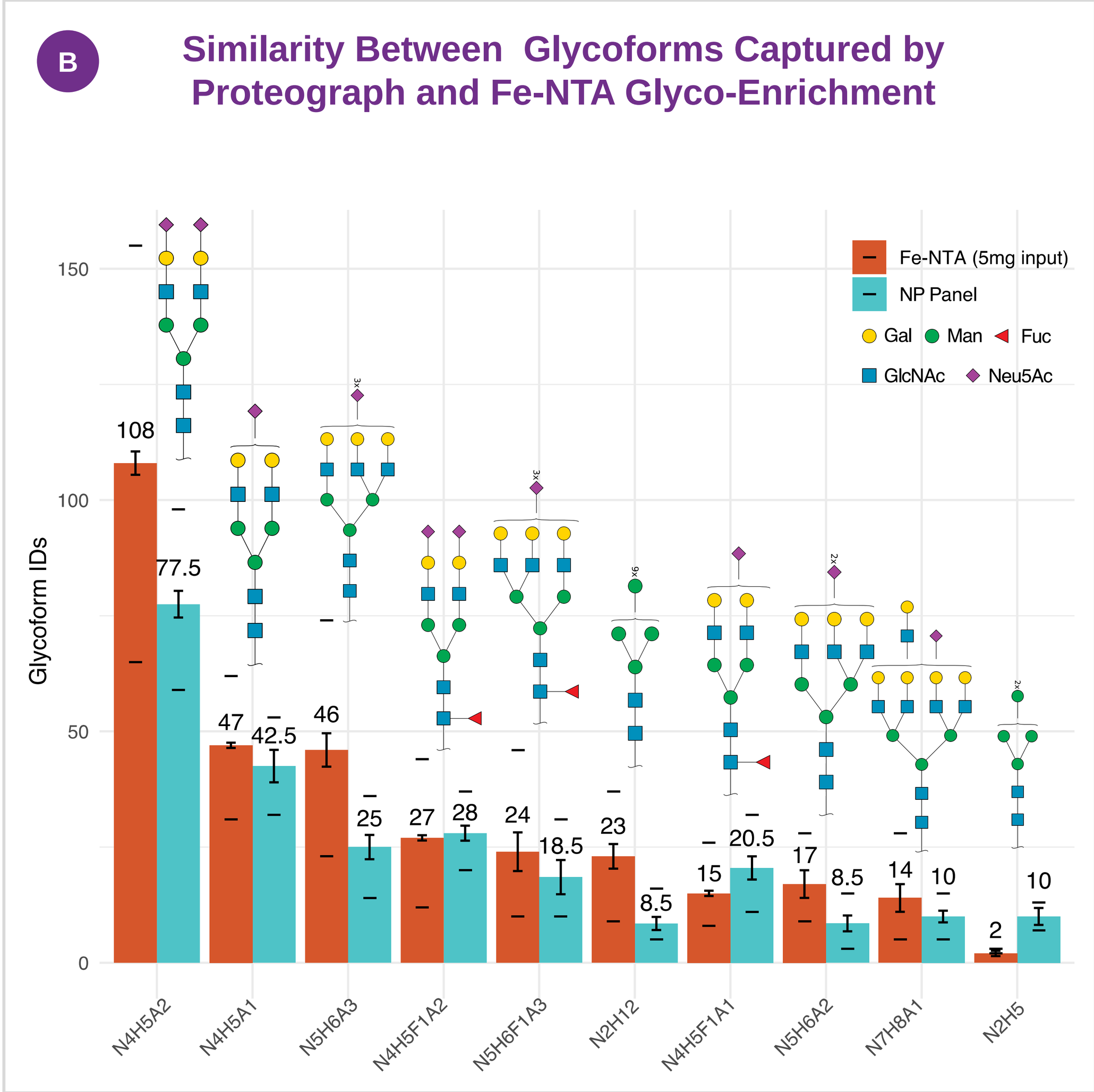
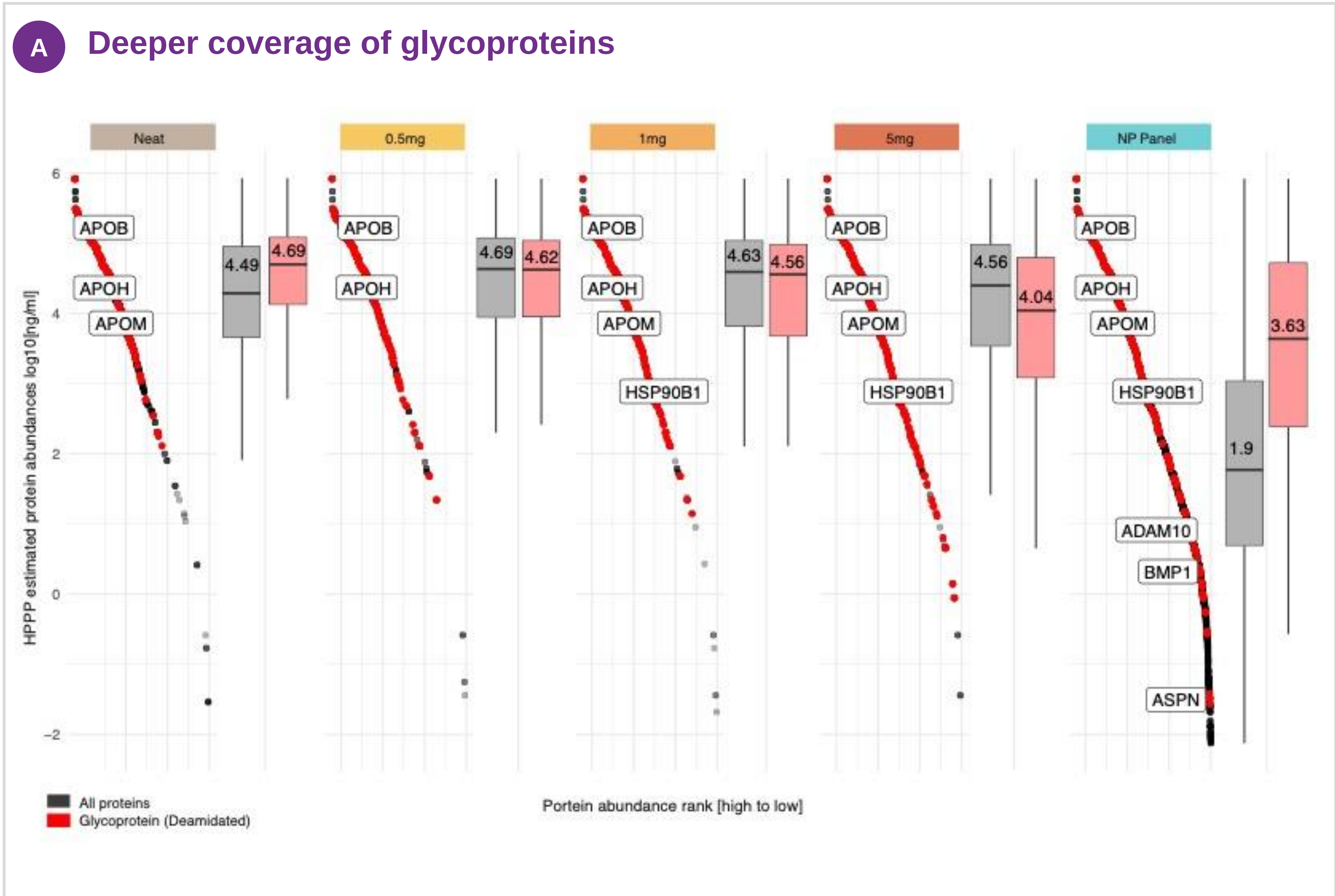
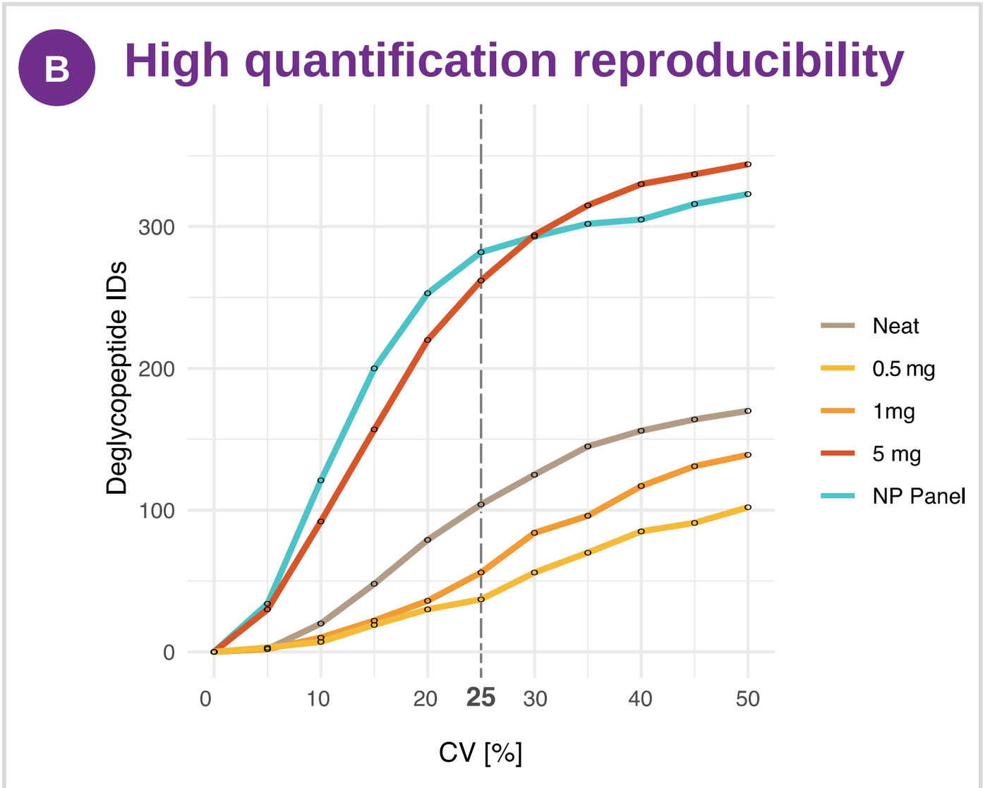


Figure 2. Identification of Glycopeptides Using Proteograph Compared to Traditional Glyco-enrichment Workflows. **(A)** The number of peptides and glycopeptides identified in a single pooled plasma. NP panel identifies more than 3.5x and 6x number of peptides compared to neat plasma and Fe-NTA enrichment, respectively. In addition, we identified comparable number of glycopeptides in the NP panel and Fe-NTA enrichment with 5mg plasma digest input. The glycopeptide identifications in the NP panel are more reproducible with JI (Jaccard Index) of 0.62 compared to 0.46 for the best Fe-NTA enrichment. The bar plots show the median number of IDs for three replicates. The top dash shows the number of features identified in any of the replicates and the bottom dash line shows the number of features commonly identified in all three replicates.

(B) The top 10 glycoforms identified in NP panel (142 overall) in comparison to the Fe-NTA enrichment with 5mg input. Data show similar distribution of glycopeptides between the two methods. In addition, for each glycoform, the NP panel shows more reproducible identification (bottom dash) compared to Fe-NTA. Since in the NP panel we are not specifically enriching for glycopeptides, it enables glycopeptide identification with minimum bias toward any glycan structure.





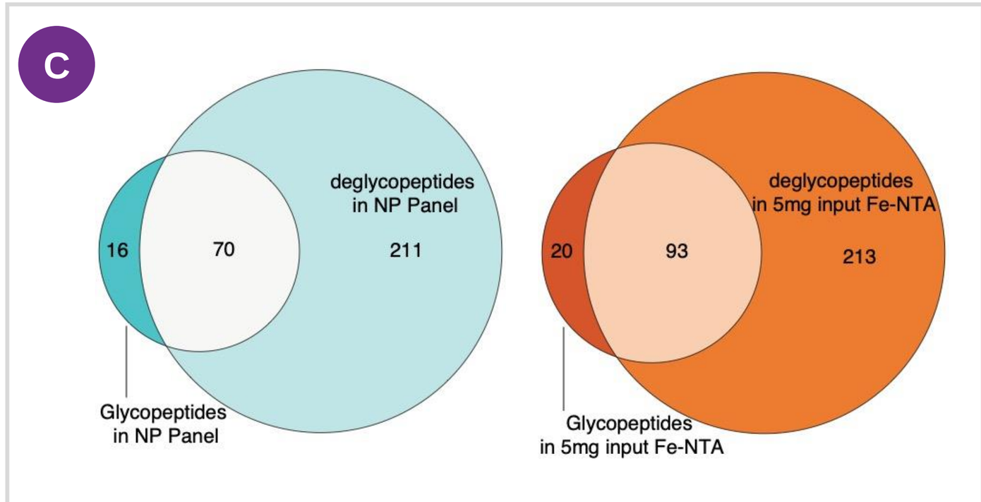


Figure 3. Coverage of Proteome and Glycoproteome with Different Workflows. **(A)** Depth of coverage in each workflow is shown based on the estimated protein abundance in the HPPP reference library[4]. With Proteograph, the median abundance of the identified peptides is about 2 orders of magnitude lower than the other workflows. Consequently, using Proteograph, we identify proteins with deglycopeptide evidence at the lower abundance range compared to neat plasma and the glyco-enriched workflows. In each workflow a few proteins that have at least one deglycopeptide evidence are highlighted. These glycoproteins are known to have roles in lipid metabolism, neurodegeneration, heatshock and growth.

(B) Using the NP Panel, we identify more deglycopeptides compared to the glyco-enriched workflow with high quantification reproducibility and it is only at higher CVs that the number of deglycopeptide identifications in glyco-enriched workflow surpasses the NP panel.

(C) The overlap between the glycopeptide and deglycopeptide sequences in both the NP panel and the glyco-enrichment workflow confirms that most of glycopeptide identifications are reliable, but many glycopeptides are not identified prior to PNGaseF treatment.

Conclusion

- NP-protein coronas capture subpopulations of the glycoproteome efficiently without subsequent glycopeptide enrichment, minimizing the bias toward specific glycan structures.
- Distinctly functionalized NPs offer complementary views of the plasma glycoproteome due to their specificities for different subsets of the proteome, and likely, different glycosylated proteoforms.
- NPs can identify similar number of glycopeptides to a glyco-enrichment workflow, but simultaneously capture 6x more peptides.

References

1. Blume et al. (2020). Nature Communications.
2. Ferdosi et al. (2022). PNAS
3. Ferdosi et al. (2022) Advanced Materials.
4. Deep plasma reference proteomes: Schwenk, et al. (2017). Journal of Proteome Research

Publications