Plasma Proteomics Identifies Insulin Resistance Signatures Predictive of Cardiometabolic Disease Risk

Jeppe Kjærgaard[1,10*], Nigel Kurgan[1*], Justus F. Gräf[1], Mafalda Custódio[1], Farina Schlabs[1], Paola Pisano[7], Lisa Ottander[1], Maria Houborg Petersen[5], Martin Eisemann de Almeida[5,11], Jacob Stidsen[5,6], Niels Ørtenblad[11], Mikael Rydén[4,8], Simon Rasmussen[1], Anna Krook[2], Ruth J.F. Loos[1,9], Roelof A.J. Smit[1,9], Kurt Højlund[5,6], Juleen R. Zierath[1,2,3], Michael Wierer[7], Atul S Deshmukh[1,12,#]

Supplementary Note:

Technical comparison of plasma proteomics workflows

Across the MS-based workflows, median protein identification was highest for Mag-Net (2,725), followed by Depleted (2,209), PCA (1,326), and Neat (1,185) (Fig. 1a). Completeness curves (10–100%) illustrate differences in quantitative coverage across workflows; where at the 100% completeness threshold, 1,150, 1,227, 669, and 517 proteins were quantified across all 150 samples using the MagNet, Depleted, PCA, and Neat workflows, respectively (Fig. 1b). Each proteomic workflow offered distinct advantages for capturing proteins of metabolic relevance (Fig. 1c). For example, PCA selectively captured hormones such as insulin, LEAP2, and ghrelin, while Mag-Net enriched for extracellular vesicle (EV) marker proteins such as CD47 and TSG101. In contrast, the depleted plasma workflow effectively reduced the dominance of highly abundant plasma proteins, enabling the detection of lower-abundance proteins like RET and Leptin. Albumin protein intensity was effectively reduced from a median of 14.9% in the Neat workflow to 3.6%, 1.8%, and 0.007% in the MagNet, PCA, and Depleted workflows, respectively (Fig. 1d). Although the PCA and MagNet workflows may enrich for proteins with lower molecular weight (MW), no difference was observed in the median MW across proteins identified in each workflow, with a median of approximately 50 kDa (Fig. 1e). Enrichment analysis of proteins uniquely identified by Olink, MagNet, and Neat revealed distinct biological processes, with Olink primarily capturing protein–DNA complex-associated proteins, MagNet enriching for mitochondrial proteins, and Neat identifying a higher proportion of ribosomal proteins (Fig. 1f). Proteins uniquely detected by MagNet were enriched in neurological and genetic disorders, supporting its potential for studies on Alzheimer's disease, dementia, and age-related conditions (Fig. 1g).

Reproducibility of plasma protein measurements across proteomics workflows

Discordance in quantitative protein measurements between proteomics workflows has recently been reported. We identified 348 proteins overlapping across all workflows, with 155

of these consistently quantified in all 150 samples (Fig. 2a). Spearman's correlation analysis between workflows showed that the Neat–Depleted pair exhibited the highest correlation (Rho = 0.61), followed by Neat–MagNet (Rho = 0.46) (Fig. 2a-b). The Neat and Depleted workflows also showed robust correlations with proteins quantified by Olink. In contrast, correlations with Olink, PCA, and MagNet revealed subsets of anti-correlated proteins ranging from 3.2% to 9.7% of the total. For example, lipoprotein(a) (LPA) displayed a strong positive correlation between Neat and Olink (Rho = 0.95), whereas vitamin D–binding protein (GC) showed no correlation between these methods (Fig. 2c). Notably, GC was strongly correlated between Neat and Depleted workflows, underscoring the value of applying complementary proteomics strategies for validation (Fig. 2d).

We compared the coefficients of variation (CV) across the 5 workflows. The Depleted and Neat workflows exhibited the lowest (CV) across biological samples (Fig. 2e). Notably, we identified clusters of proteins where PCA, MagNet, and Olink showed high CVs, while Depleted and Neat remained low, indicating workflow-specific analytical variability and potential challenges in quantification (Fig. 2f). In conclusion, these findings demonstrate that technical variability in one workflow can limit accurate quantification for specific protein subsets, which may be resolved by a complementary workflow. This multi-workflow approach therefore maximizes biological yield, a particular advantage when studying precious or limited clinical samples.

Associations between plasma proteins and clinical features

To explore whether these proteomics datasets capture relevant clinical phenotypes, we conducted a Spearman correlation analysis (adjusted for sex) between protein abundance and multiple clinical traits in the fasted state. Across all workflows, we identified 3,050 unique protein–phenotype associations (FDR < 0.05) (Fig. 3a), including traits linked to adiposity, insulin sensitivity, and glycemic control. The majority of associations were observed in the Olink workflow (1,351), followed by the Depleted workflow (820), illustrating that individual workflow captures distinct and complementary biological signals shaped by their unique enrichment strategies and proteome coverage.

In summary, MWP substantially expands plasma proteome coverage and uncovers clinically meaningful, individual-specific biological signatures. Each workflow enriches distinct protein subsets, including hormone regulators, extracellular vesicle components, and low-abundance signaling proteins. Together, MWP establishes a robust foundation for personalized metabolic profiling and biomarker discovery.

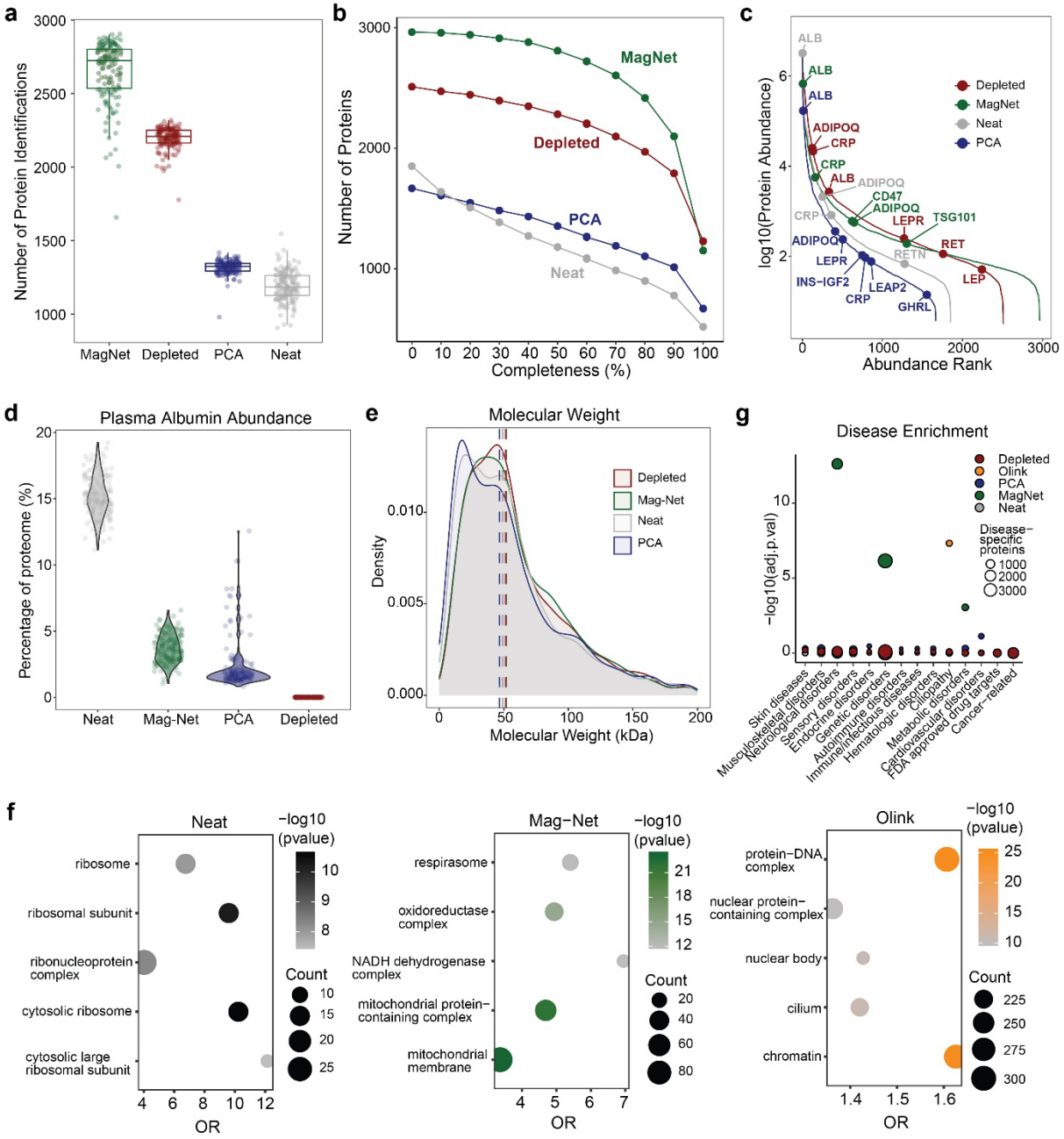

Figure 1 - Technical comparison of plasma proteomics workflows: a) Number of proteins identified per sample by mass spectrometry-based plasma proteome profiling. b) Data completeness across the 150 samples and four mass spectrometry-based workflows.c) Ranked plasma protein abundance across four MS-based workflows. d) Plasma albumin abundance (%) across the four mass spectrometry-based workflows. e) Density plot of plasma protein molecular weight across the four mass spectrometry-based workflows. f) Gene Ontology Cellular Components overrepresentation analysis of the proteins uniquely identified by each workflow. All the proteins identified in the plasma proteome were used as background. g) Disease overrepresentation analysis of the proteins uniquely identified by each workflow.

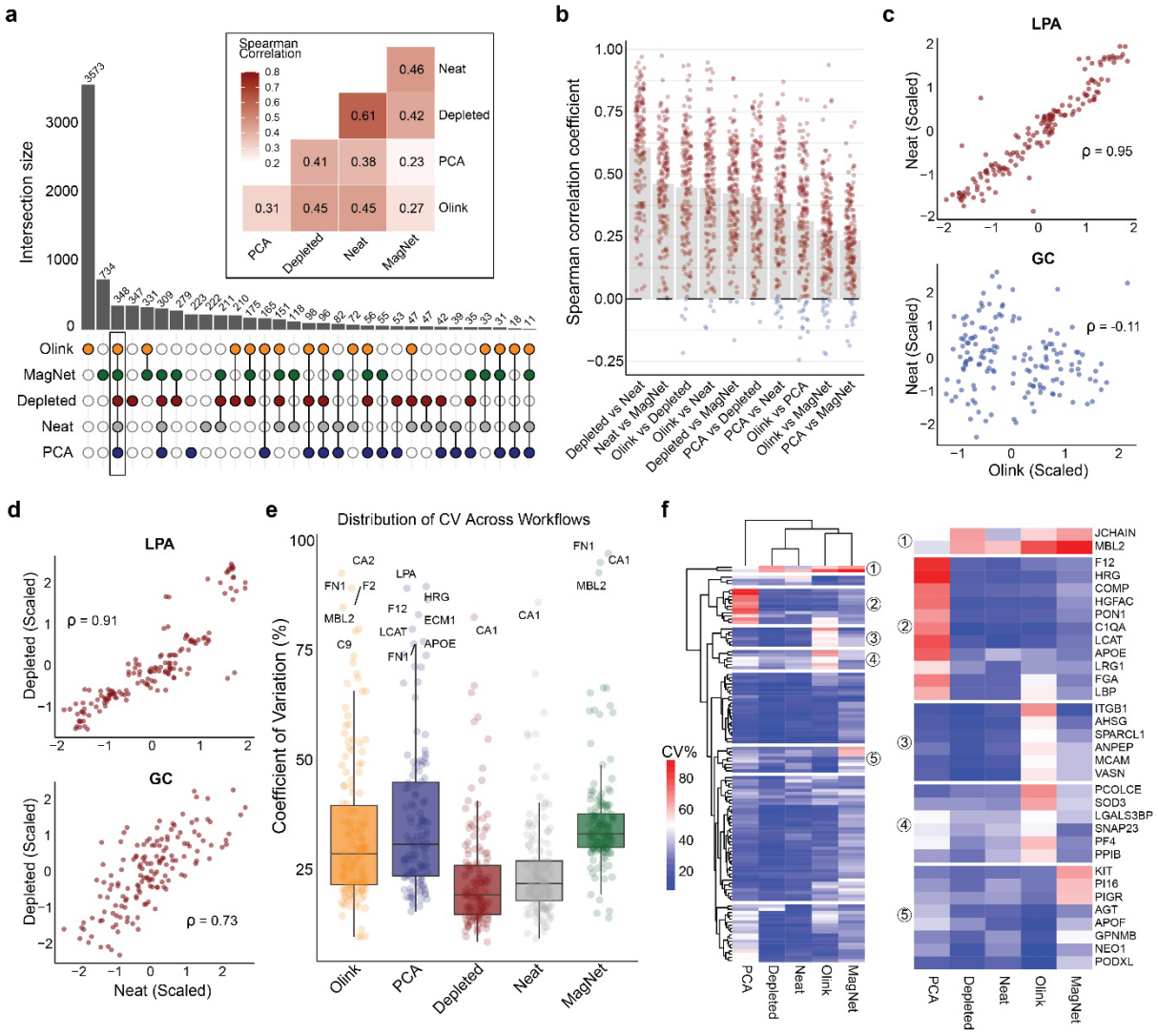

Figure 2 - Reproducibility of plasma protein measurements across proteomics workflows: a) Upset plot of unique and overlapping proteins identified across workflows, including Spearman correlation between shared 155 proteins quantified across all 150 samples. b) Individual protein Spearman correlation of the shared 155 proteins across workflows. c) Example of proteins with high and low correlation between Neat and Olink-based proteomics. d) Example of proteins between Neat and Depleted-based proteomics. e) Coefficient of variation of the 155 shared fully quantified proteins across the 150 samples for the four mass spectrometry-based workflows. Only the coefficient of variation < 100% is shown in the plot. f) Heatmap with hierarchical clustering of the 155 shared fully quantified proteins. 5 highly variable clusters across workflows are shown.

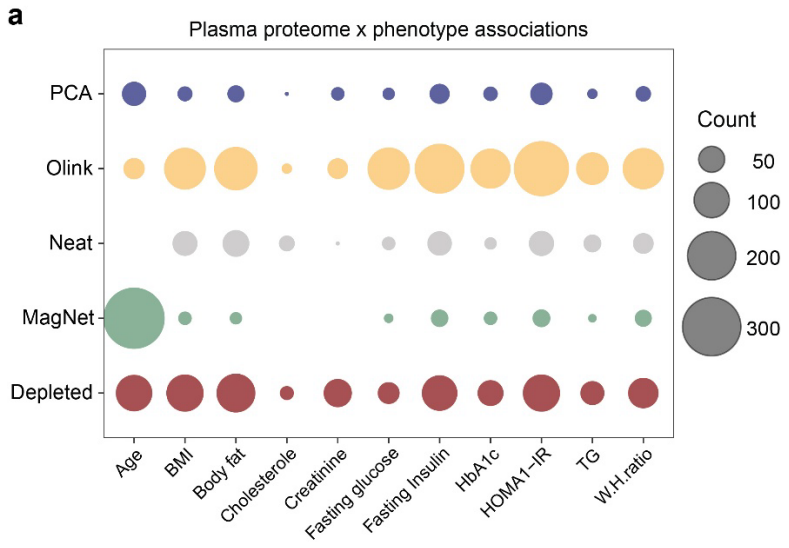

Figure 3 - Associations between plasma proteins and clinical features: a) Global plasma proteome associations with clinical features in the fasted state. Spearman correlation adjusted for sex was used. Proteins with a Benjamini-Hochberg corrected $P < 0.05$ are displayed.