Cell Systems, Volume *2*

Supplemental Information

Plasma Proteome Profiling

to Assess Human Health and Disease

Philipp E. Geyer, Nils A. Kulak, Garwin Pichler, Lesca M. Holdt, Daniel Teupser, and Matthias Mann

Supplemental Figures

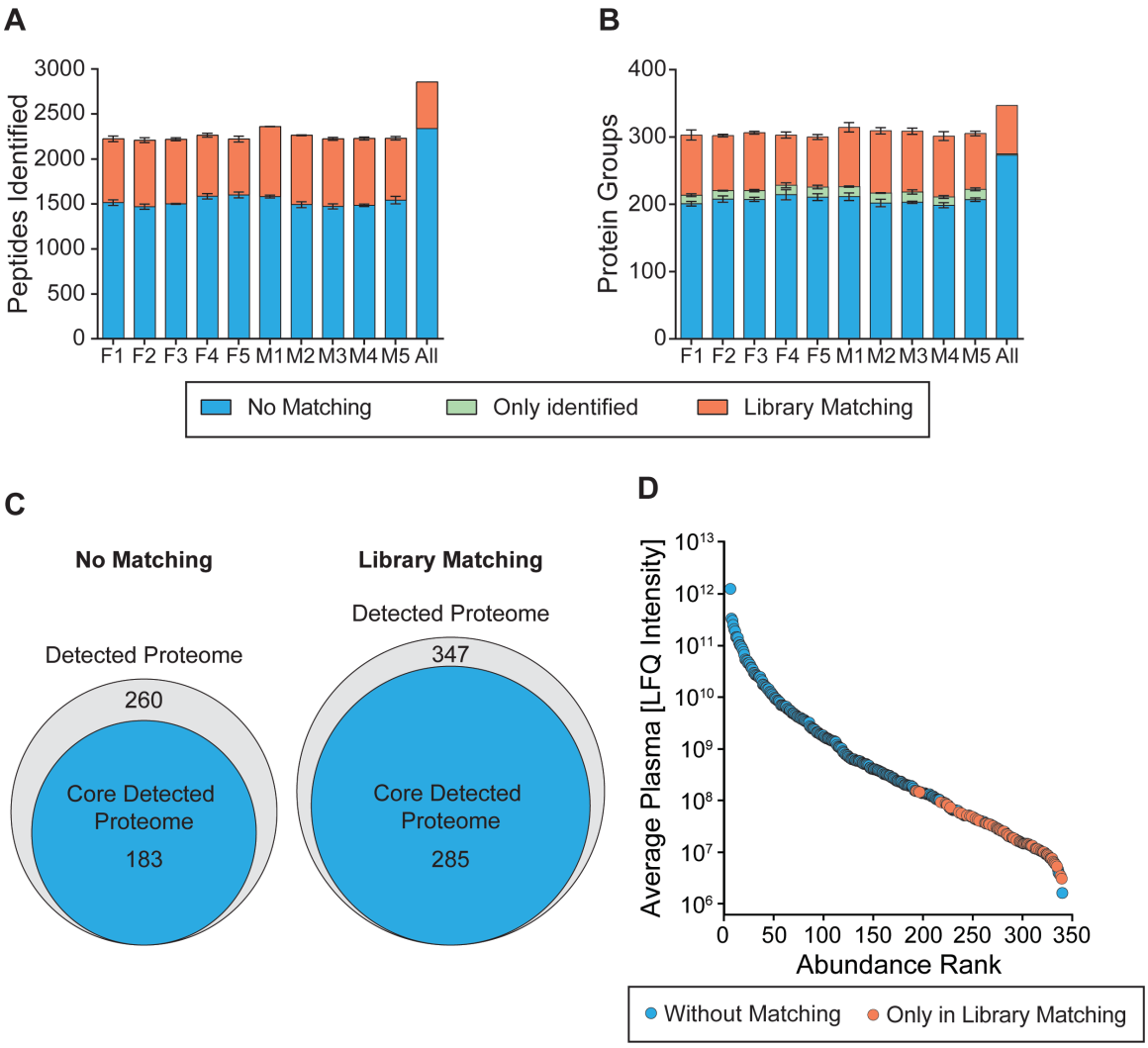


Figure S1. Related to Figure 1. Gain of using a library for match between runs.

A Number of peptides identified in ten different individuals with and without employing match between runs.

B Number of proteins that were identified and quantified with and without the gain of a matching library. Proteins that were only identified, but not quantified are indicated in green.

C 285 out of 347 proteins were identified in all ten individuals, reflecting the core detected proteome in this dataset.

D The additional proteins after library matching are all present in the lower concentration rage and are shown in orange. Blue dots represent proteins that were also present in the analysis without matching.

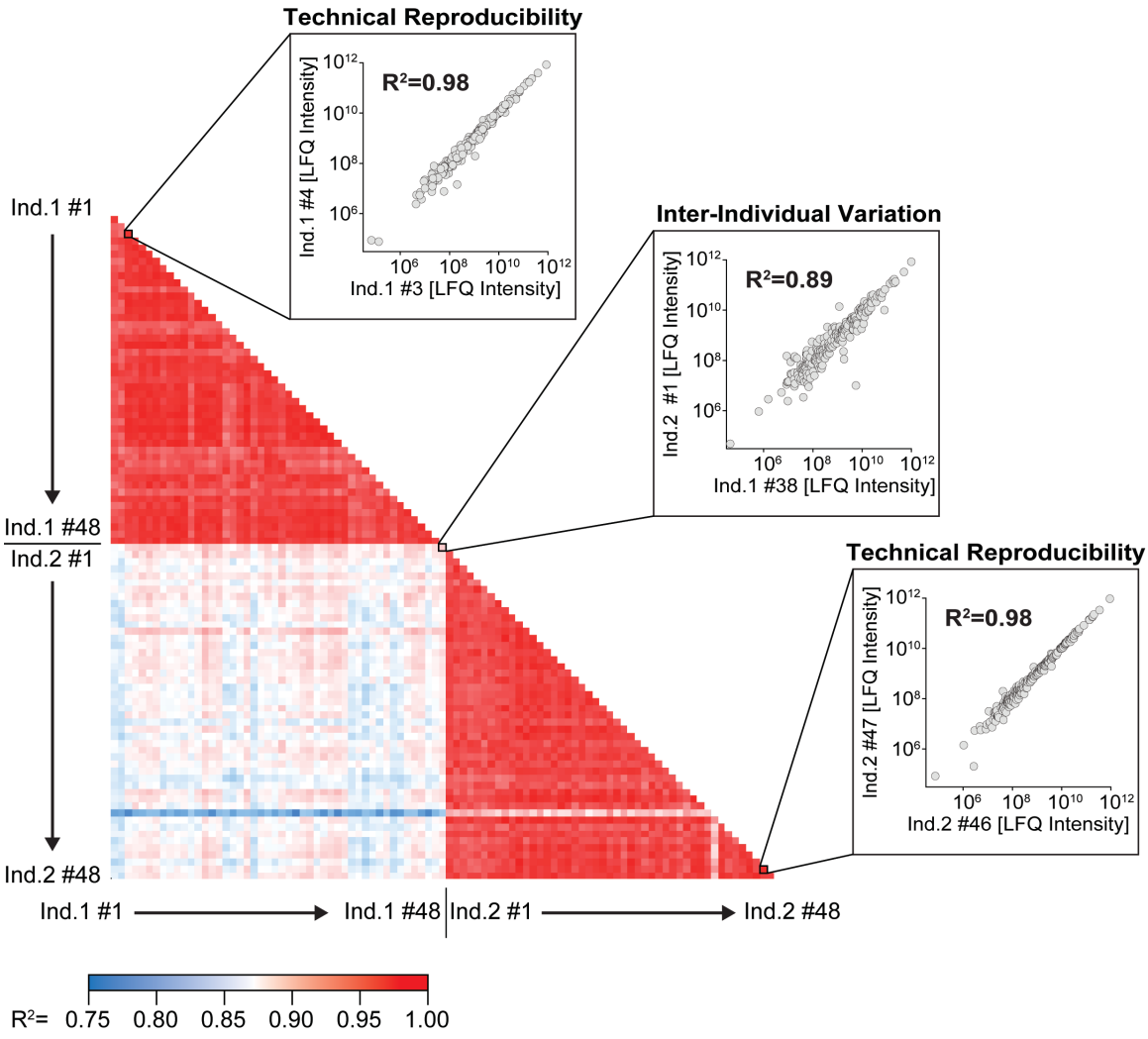


Figure S2. Related to Figure 1. Reproducibility of an automated preparation of 96 plasma samples.

Plasma of two individuals was distributed on a 96 well plate (48 plasma samples for each individual) and samples were prepared on a liquid handling platform. The figure shows 4,560 binary comparisons between the samples with color-coded R^2 values. For illustration, two correlations of technical replicates and one of the two different individuals are zoomed and displayed in the insets.

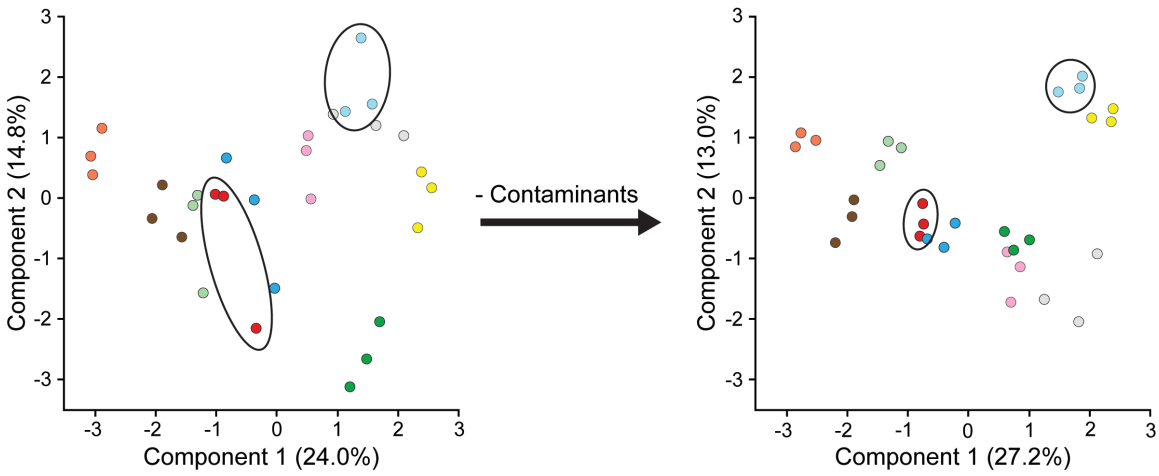
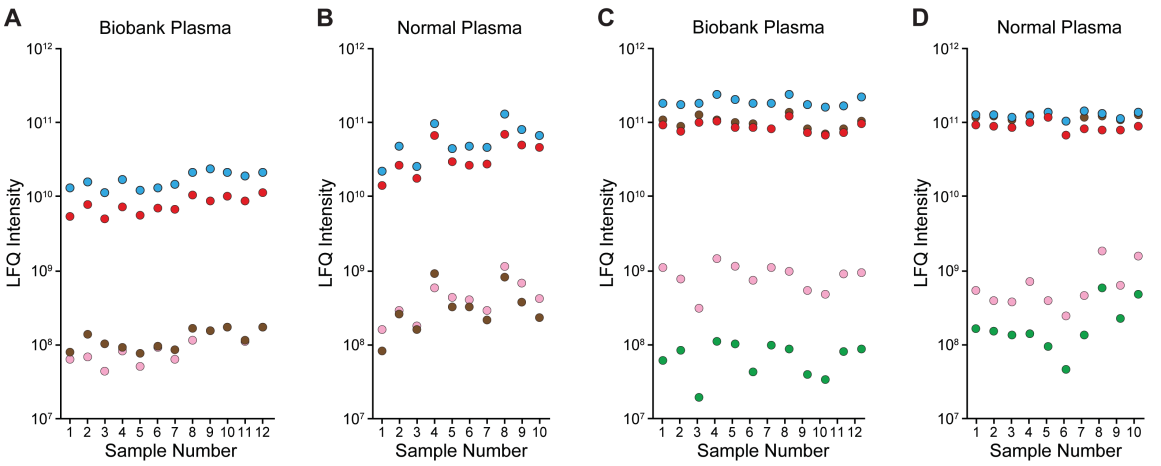


Figure S3. Related to Figure 2. Effect of removal of specific contaminants.

Removal of typical contaminants like keratins and high abundant erythrocyte specific proteins from the analysis results in a stronger clustering of workflow triplicates of ten individuals in a two-dimensional PCA. The grey circles exemplify the stronger clustering for two individuals measured in triplicates.



	HBB	HBA1	HBD	CA1
Biobank CV [%]	25	26	31	47
Normal CV [%]	55	53	75	63

	FGA	FGB	FGG	PPBP	PF4V1
Biobank CV [%]	14	17	20	37	42
Normal CV [%]	10	16	6	76	80

Figure S4. Related to Figure 3. Sample quality marker of a plasma reference panel.

A Variations of markers of erythrocyte lysis within twelve high quality reference plasma samples from a blood biobanks compared to randomly chosen normal plasma samples from ten individuals, in which variation is much higher. HBA, HBB, HBD: Hemoglobin subunit alpha, beta, delta; CA1: Carbonic anhydrase 1.

B Protein marker for coagulations in plasma samples from a biobank compared to normal plasma. FGA, FGB, FGG: Fibrinogen alpha, beta, gamma chain; PPBP: Platelet basic protein; PF4V1: Platelet factor 4 variant.

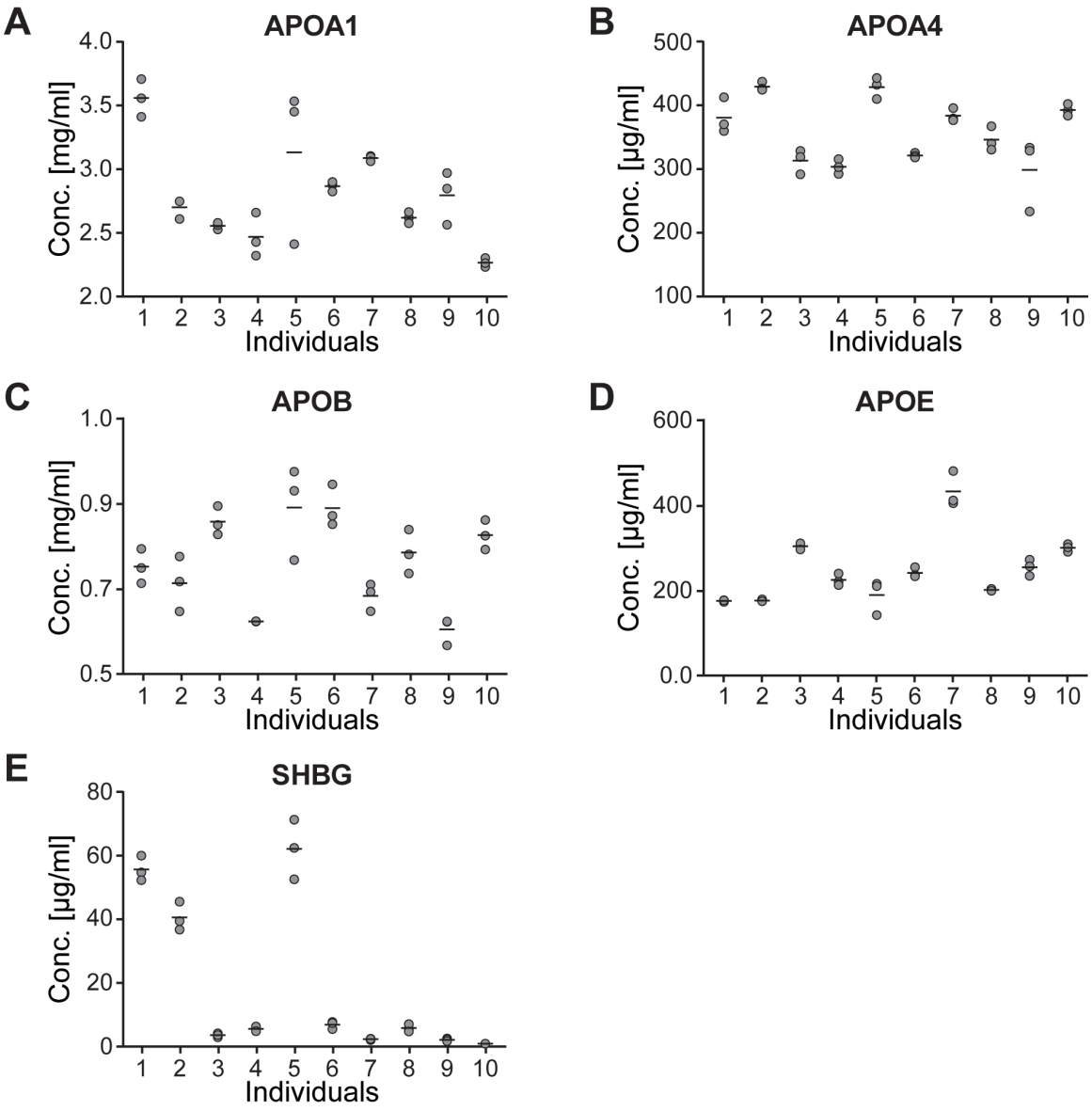


Figure S5. Related to Table S4. Plasma PrESTs as internal standards for protein quantification

A Ratios of heavy labeled PrESTs to light endogenous apolipoprotein A1 (APOA1) for ten individuals in triplicates.

B Ratios of heavy labeled PrESTs to light endogenous apolipoprotein A4 (APOA4) for ten individuals in triplicates.

C Ratios of heavy labeled PrESTs to light endogenous apolipoprotein B (APOB) for ten individuals in triplicates.

D Ratios of heavy labeled PrESTs to light endogenous apolipoprotein E (APOE) for ten individuals in triplicates.

E Ratios of heavy labeled PrESTs to light endogenous sex hormone-binding globulin (SHBG) for ten individuals in triplicates. Individuals 1-5 are women and 6-10 are male.

Supplemental Table Legends

Table S1. Related to Figure 1. Comparison of 1h versus overnight digestions and 20 min versus 100 min gradients.

Table S2. Related to Figure 1. Reproducibility of protein quantification.

Table S3. Related to Figure 2. CVs of the technical replicates of all individuals.

Table S4. Related to Figure S5. CVs and median concentrations of plasma PrESTs.

Table S5. Related to Figure 4. Deep proteome data.

Table S6. Related to Figure 4. Statistically significant features identified by '1D annotation' of plasma proteome.