

## How to Avoid Pre-Analytical Drivers of Bias in Nanoparticle-Enriched Plasma Proteomics Studies?

Nanomics Biotechnology us@nanomics.bio

### Keywords:

Proteomics Studies, Proteonano™ Ultrplex Proteomics Platform, Sample-Related Biases

### Objective:

Our aim is to prevent inter-batch variability in the Proteonano™ ultrplex proteomics platform that may arise from sample-related factors.

### Sample-related factors

Proteomics is a leading field for research and clinical translation in the post-genomic era. Early efforts in the Human Proteome Project show that proteomics excels in applications such as discovering new biomarkers, identifying novel drug targets, early cancer screening, companion diagnostics, patient stratification, and selecting candidates for clinical trials.

Plasma proteomics is one of the most challenging tasks in clinical proteomics due to the inherent complexity of the samples. Factors such as extreme heterogeneity of proteins at highly dynamic concentrations and improper collection and storage conditions of plasma/serum samples can adversely affect the results and lead to erroneous conclusions. Therefore, it is crucial to evaluate sample quality before initiating a project.

Potential sample related factors include:

- ① Hemolysis
- ② Sample contamination
- ③ Duration of sample storage
- ④ Pre-treatment methods for plasma samples
- ⑤ Centrifugation

#### 1. Hemolysis

We collected 10 plasma samples with varying levels of

hemolysis, categorized as follows: 0 (no hemolysis), 1 (mild hemolysis), and 2 (severe hemolysis). After processing these samples using the Proteonano™ platform—through steps of enrichment, reduction, alkylation, enzymatic digestion, desalting, and lyophilization—we analyzed them with the timsTOF Pro 2. The results, shown in Figure 1, indicate that the number of identified proteins increased with the severity of hemolysis, with the severe hemolysis group having almost twice as many identified proteins as the no hemolysis group.

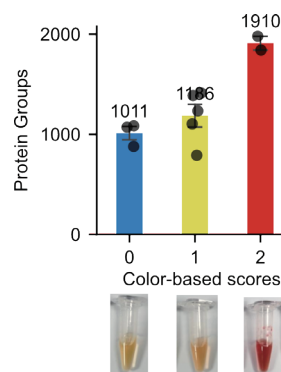


Figure 1: Impact of sample hemolysis on protein identification

#### 2. Potential sample contamination

Professor Matthias Mann from the Max Planck Institute has highlighted that platelets, red blood cells, and clotting factors are key contributors to plasma quality issues[1]. Plasma

tors are key contributors to plasma quality issues[1]. Plasma samples, which are typically collected during routine clinical procedures, can vary in handling, potentially impacting research outcomes. In case-control studies, any inconsistencies in sample collection and processing can introduce systematic biases, complicating the identification of true biomarkers and possibly leading to erroneous conclusions.

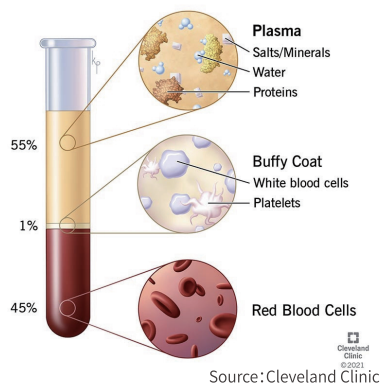


Figure 2: Blood plasma components

Protein	Source of contamination
MYH9	
ACTB	Platelet
TLN1	
HBA1	
HBB	Red Blood Cells
HBD	
FGA	
FGB	Coagulation
FGG	

Table 1: Three quality control panels containing key indicative marker proteins of contamination from platelets, erythrocytes, and coagulation.

A Reference Proteomes

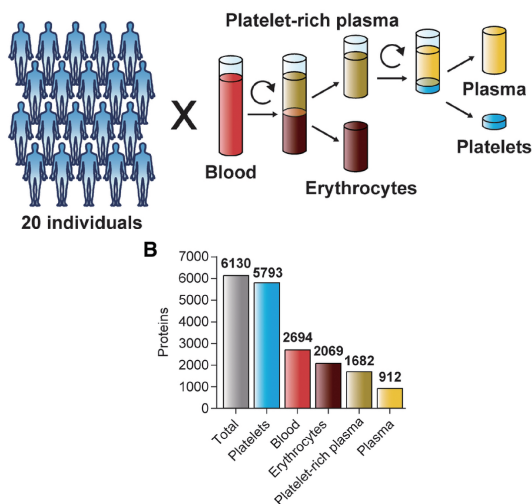


Figure 3: PGs identified in different sample types

A previous study developed a biomarker panel to systematically evaluate hemolysis, platelet contamination, and coagulation in plasma samples. This panel is useful for assessing the quality of individual samples as well as entire clinical studies [1]. If the intensity of any protein exceeds the average intensity across all samples by more than two standard deviations, it is considered to fail the quality control standards, as illustrated in Figure 4.

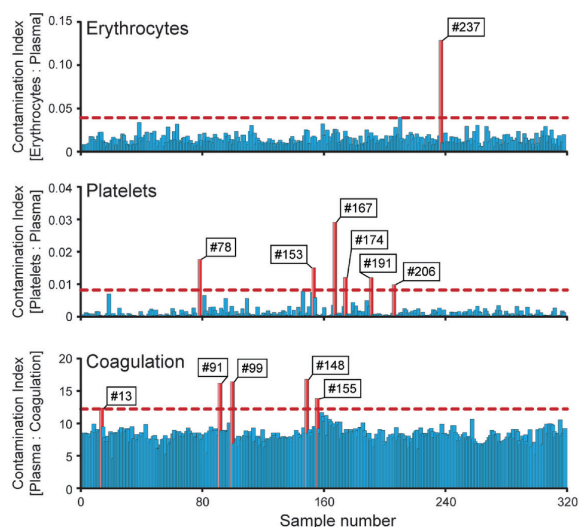


Figure 4: Assessment of potential contaminants in samples [1]. The horizontal red line marks the threshold where the contamination index exceeds two standard deviations above the mean for all samples.

3. Impact of sample storage duration

Sample storage duration is a crucial factor to consider. Enroth S. and colleagues studied how protein levels change with individual size, using storage time as an independent variable. We tested Proteonano™ plasma enrich kit on samples stored at -80°C for 0.5, 4, 7, and 12 years. Our findings showed that extended storage significantly reduced the

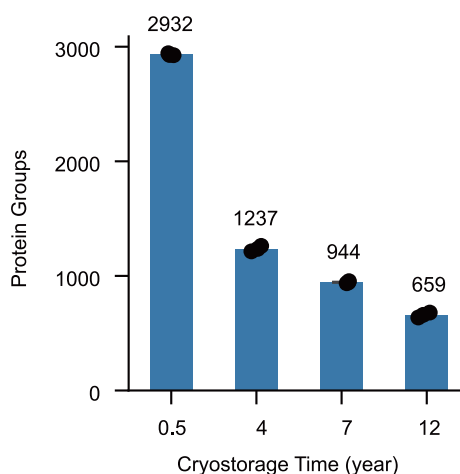
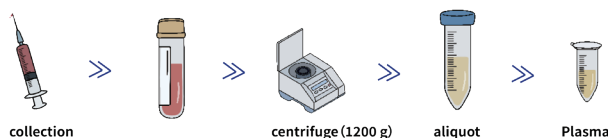


Figure 5: Effect of sample storage duration on PGs identified

number of detectable proteins. On average, 2,932 PGs were identified in samples stored for 0.5 years (see Figure 5). This number decreased to 1,237 PGs after 4 years and further dropped to 659 PGs after 12 years. This decline is likely due to protein structural changes and degradation over long storage periods.

#### 4. Plasma handling

Plasma is typically separated from blood using centrifugation (Figure 6). Research by Shen et al. [2] has shown that the delay before centrifugation can significantly affect plasma proteomics. They found that longer delays before centrifugation lead to higher levels of certain plasma proteins (see Figure 7). This increase is due to proteins leaking from lysed blood cells. When samples are processed at 4°C and centrifuged within 8 hours, this effect is minimal. However, if samples are handled at room temperature (22°C), protein levels can rise noticeably within just one hour. This suggests that protein degradation is not the only concern with temperature changes. The impact on protein levels becomes more significant if the delay exceeds this timeframe. Therefore, it is essential to keep the delay time before centrifugation as short and consistent as possible to avoid affecting the experimental data.



Tips: Centrifuge newly collected samples within 1 hour at room temperature or within 8 hours if kept at 4°C.

Figure 6: Sample preprocessing workflow

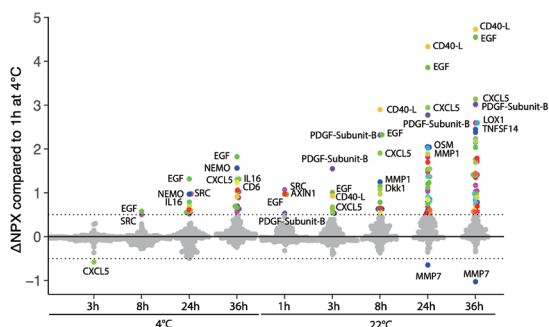


Figure 7: Protein level changes [2]

#### Recommended Plasma Sample Processing Procedure

Plasma: Centrifuge newly collected samples within 1 hour at room temperature or within 8 hours if kept at 4°C.

Serum: Allow newly collected samples to clot at room temperature for 30-60 minutes before centrifugation.

Storage and Transport: Store samples at -80°C and use dry ice for shipping.

#### Additional Guidelines:

Record the procedures for sample collection and processing.

Ensure uniform processing methods for all samples within a study. For example, keep centrifugation temperatures and speeds consistent, and use the

#### 5. Centrifugation

The necessity of pre-enrichment centrifugation for plasma proteomics remains inconclusive, with varying practices across laboratories. To systematically evaluate its impact, plasma samples were allocated into three groups and subjected to differential centrifugation at 0 g (no centrifugation), 2000 g, and 12000 g, respectively. All samples were subsequently processed using the standard Proteonano™ enrichment protocol. Quantitative proteomic analysis revealed a stepwise reduction in the number of identified proteins with increasing centrifugal force (Figure 7). These data demonstrate that high-speed centrifugation prior to enrichment markedly decreases protein identification yield in plasma samples

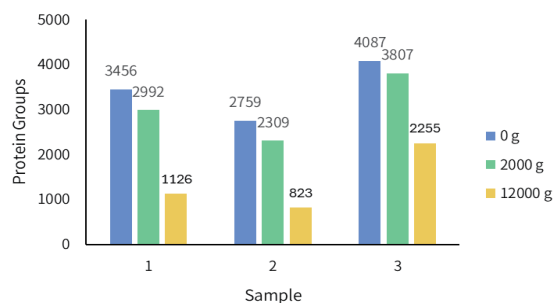
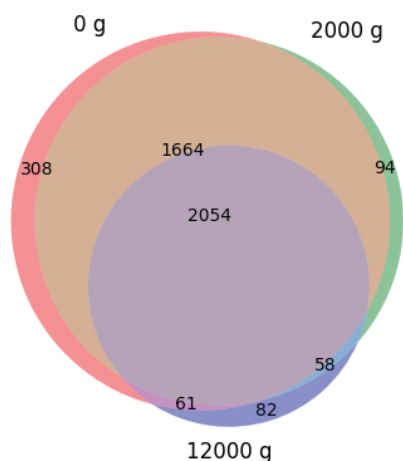


Figure 8: Protein identifications under different centrifugation conditions

To further investigate the effect of centrifugation on protein recovery, proteomic profiles of enriched serum and plasma samples under different centrifugation conditions were

compared. In plasma samples, where high overlap was noted between the non-centrifuged and 2000 g conditions, whereas significant differences emerged when compared to the 12000 g group (Figure 9). These results collectively indicate that pre-enrichment centrifugation is not necessary and may in fact reduce proteome coverage in both serum and plasma.



**Figure 9:** Venn diagram of protein overlap across centrifugation conditions

Maintaining consistency in sample collection, processing, storage, and transport throughout the experiment is essential to avoid data disturbances or misinterpretations. Use high-quality laboratory techniques and reagents to minimize the risk of contamination from external proteins, which could interfere with the genuine biological signals you are studying. Additionally, ensure your sample size is large enough to produce statistically significant results, as small samples can lead to unreliable findings or fail to reveal true biological differences. For studies involving clinical samples or samples in specific conditions, meticulously document and control these variables to exclude other potential confounding factors.

In summary, approach sample handling with rigor, including thorough experimental design, proper quality control, and robust statistical analysis, to ensure reliable and meaningful results from proteomics research.

## Conclusion

In plasma proteomics experiments, improper sample handling is a major factor contributing to variability between samples. To ensure the highest quality of proteomics data, especially in large-cohort studies, we recommend the following:

Avoid contamination from blood cells, platelets, and coagulation factors during sample collection.

Prolonged sample storage can lead to nontrivial protein degradation. We recommend store samples for no more than 3 years, and avoid repeated freeze-thaw cycles.

We recommend use the same protocol for plasma sample collection at different sites and complete plasma collection within 8 hours at 4°C.

Modest sample (blood) centrifugation speed is 1200 g.

## Reference

- [1] Geyer PE, Voytik E, Treit PV, Doll S, Kleinhempel A, Niu L, Müller JB, Buchholtz ML, Bader JM, Teupser D, Holdt LM, Mann M. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med*. 2019 Nov 7;11(11):e10427. doi: 10.15252/emmm.201910427. Epub 2019 Sep 30. PMID: 31566909; PMCID: PMC6835559.
- [2] Shen Q, Björkstén J, Galli J, Ekman D, Broberg J, Nordberg N, Tillander A, Kamali-Moghaddam M, Tybring G, Landegren U. Strong impact on plasma protein profiles by precentrifugation delay but not by repeated freeze-thaw cycles, as analyzed using multiplex proximity extension assays. *Clin Chem Lab Med*. 2018 Mar 28;56(4):582-594. doi: 10.1515/cclm-2017-0648. PMID: 29040064.

