Pushing DIA proteomics analyses of neat plasma to 1000 protein groups ID/h

Ines Metatla[1], Kevin Roger[1], Cerina Chhuon[1], Maria de Moraes[2], Anne Jamet[2], Pierre-Olivier Schmit[3], Manuel Chapelle[3], MoonJu Cho[4] and Chiara Guerrera[1].

[1]Necker Proteomics Platform, Faculty of Medecine, University Paris Cité, Paris, France, [2]Institut Necker Enfants Malades, Paris [3]Bruker France SAS, Wissembourg, France, [4]Bruker Korea, Coventry, UK

Introduction

Direct proteomics analysis of neat plasma remains a challenge because of the huge dynamic range of the plasma proteins, but is also very appealing because it requires very small volumes and is unexpensive compared to new depletions technology. We focused on workflow optimization of neat plasma analysis using dia-PASEF® approach to maximize the number of proteins groups (PG) identified and quantified, while minimizing both gradient time and missing value levels. Here we describe the comparisons of different LC settings, and the optimization of data analysis using DIANN.

We report a reference “deep workflow” to quantify up to 811 proteins using a 40 min gradient and a “fast workflow” leading to 490 proteins using a 21 min gradient.

Methods

The plasma sample cohort consisted in 15 patients affected by a rare genetic disease and 18 age-matched controls. Samples were digested with trypsin using STRAP columns, separated either by nano-HPLC (nanoElute, Bruker Daltonics) using a 25cm column on a 40min gradient (IonOpticks, Australia), or with an EvosepOne ran with the 60 Samples Per Day (60SPD) method (Endurance column, 8cm). Both LC’s were connected to a timsTOF Pro™ instrument (Bruker Daltonics) via its Captive Spray source. The timsTOF Pro was in dia-PASEF acquisition mode: a 19-windows method using 33da windows covering a 0.65-1.3 V/cm2 mobility ranges over a 400-1050m/z range. DIA-NN1.8.1 (Demichev group), or Spectronaut®17 (Biognosys).

Results

Minimizing technical variability and obtaining reproducible intensity information is a must to enable label-free plasma-based proteomics research, in order to avoid adding technical variability to the already high biological (individual to individual) variability. We have compared the technical and biological reproducibilities for Parallel Accumulation Serial Fragmentation (PASEF) and dia-PASEF appraoches (Fig1). The higher number of ID’s comes along with a better technical reproducibility. The higher sensitivity of the approach does however reveal more of the biological variability. We also went through an optimization of the dia-PASEF acquisition method. Unlike what was observed for cell culture analysis, the use of a broader ion mobility range for acquisition allowed to increase the number of protein group ID’s. **(Fig2)**. In order to find a sweet spot in terme of number of ID’s / time we have evaluated different gradient length on different chromatographic setups, namely a nanoElute UHPLC system, designed to offer a maximized analytical flexibility, and an Evosep One system, designed for best robustness and optimized injection to injection headtimes. The gradient length ended having more influence than the column size **(Fig3)**. Using an optimized spectral library (built from 100 neat plasma and 10 prefractionnated plasma) we could get up to 854 protein group ID’s from the same experiment (data not shown).

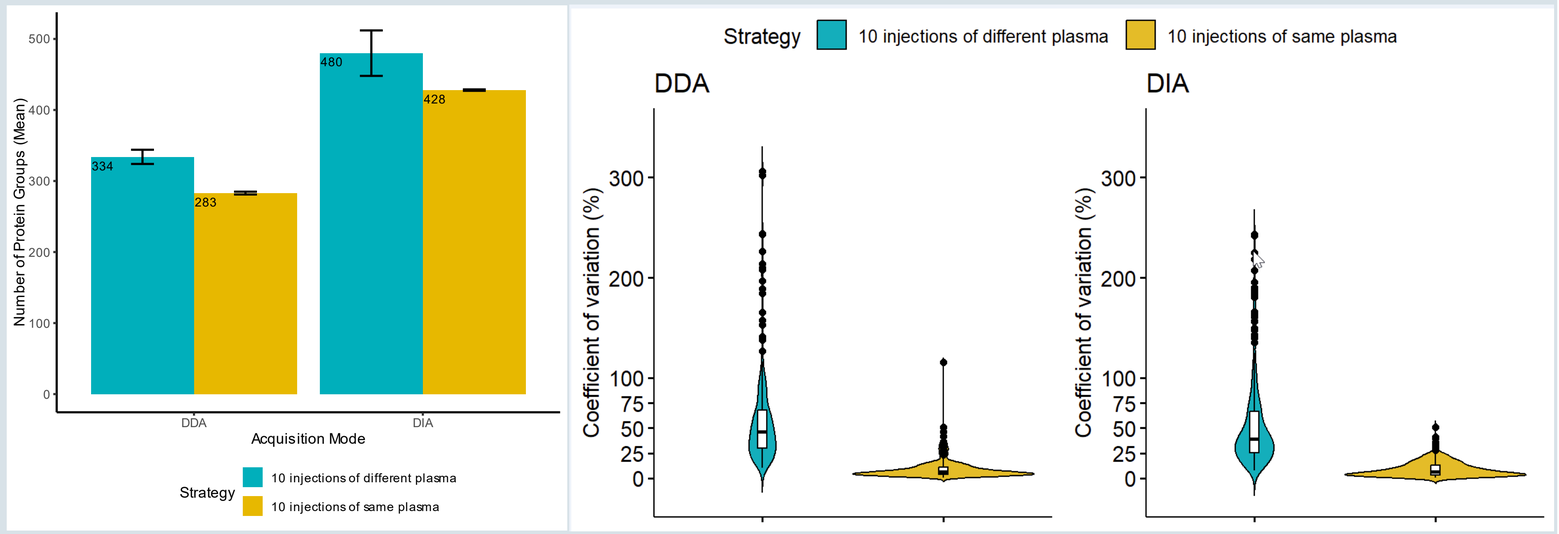

Fig. 1: Technical vs biological reproducibility.
Results from 10X injection of 1 neat plasma sample (technical variability) or 10 injections of different neat plasma sample; Separation with a 60 SPD method on an EvosepOne. Acquisition with optimized PASEF and dia-PASEF acquisition methods.

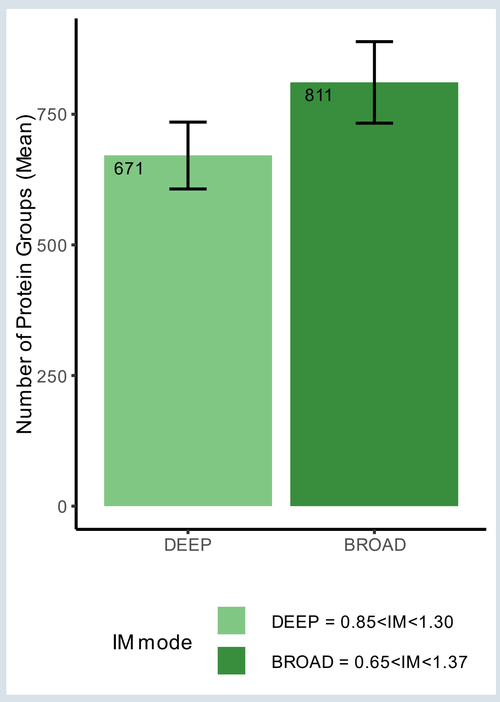

Fig. 2: dia-PASEF method optimization.
Outcome of a direct DIA search for 33 neat plasma samples acquired with dia-PASEF methods covering the same mass range but different ion mobility ranges

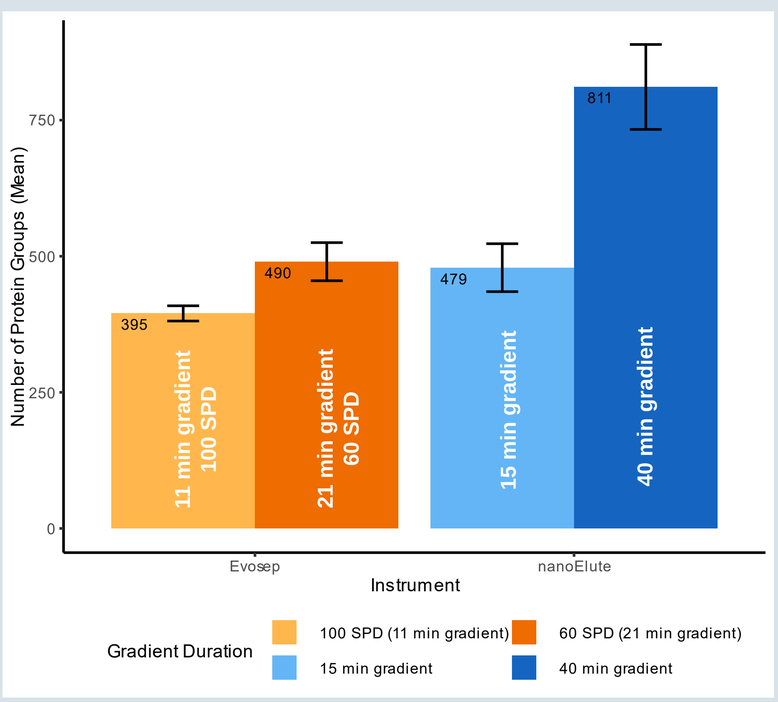

Fig. 3: gradient and column length influence
Outcome of a direct DIA search for 33 neat plasma sample. Evosep was used with an 8 cm Performance column while the nanoElute was equipped with the 25cm IonOptiks Aurora column.

Conclusions

- Both PASEF and dia-PASEF allow to obtain excellent technical reproducibility from undepleted plasma samples.
- The CV’s obtained from different biological replicates allows to get useful label free quantitation information for a majority of the detected proteins
- We have optimized two workflows which are allowing to work on more than 400 protein groups from a high throughput analysis (60 SPD) or more than 800 protein groups from a lower throughput analysis (12SPD).

timsTOF PRO