Optimizing a dia-PASEF acquisition and data analysis for non-depleted plasma for rare disease research

Ines Metatla[1], Cerina Chhuon[1], Kévin Roger[1], Antonietta Tarallo[2], Giancarlo Parenti[2], Pierre-Olivier Schmit[3], Manuel Chapelle[3], Ida Chiara Guerrera[1]

[1]Faculty of Medecine, University of Paris, Paris, France; [2]Telethon Institute of Genetics and Medicine Department of Translational Medical Science Federico II University of Naples, Pozzuoli, Italy; [3]Bruker Daltonique S.A., Wissembourg, France

Introduction

The accessibility of non-depleted plasma makes it an ideal fluid for clinical proteomics studies, despite its huge dynamic range and the low abundance of organ-specific secretions. We worked with **non-depleted plasma** samples to minimize the pre-analytical variability, reduce the costs and therefore increase the number of patients analyzed.

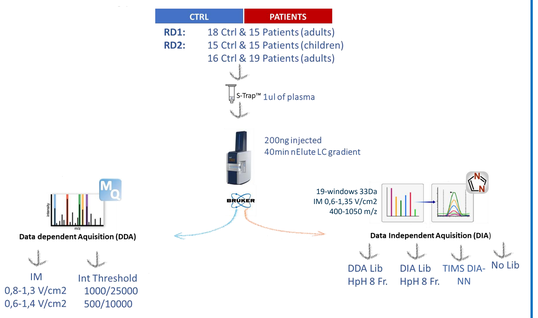
AIMS

- To compare dda-PASEF® and **dia-PASEF®** for non-depleted plasma samples
- To compare different **DIA-NN based pipelines** dia-PASEF® data analysis
- To test the optimized pipelines in two case studies of **plasma proteomics in clinical research in rare diseases**

Methods

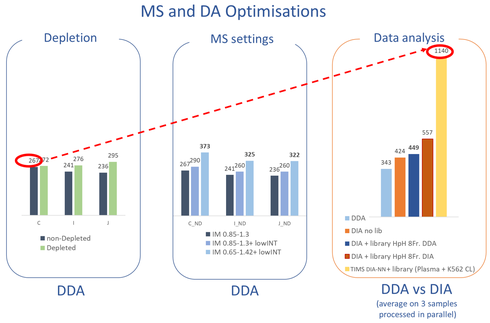
Plasma samples from 16 healthy adult donors were used for optimization. Two plasma sample cohorts from 2 separate clinical project on rare disease were used: Rare disease 1 and 2 (RD1and RD2). Samples (1ul of plasma) were digested with trypsin using STRAP columns, separated by nano-HPLC (nanoElute, Bruker Daltonics) on a pulled emitter column (IonOpticks, Australia) using a 40 min gradient. Peptides were analyzed on a timsTOF Pro™ instrument (Bruker Daltonics) operated both in PASEF and dia-PASEF modes.

Experimental strategy



Results

MS and DA Optimisations



Plasma dda-PASEF:

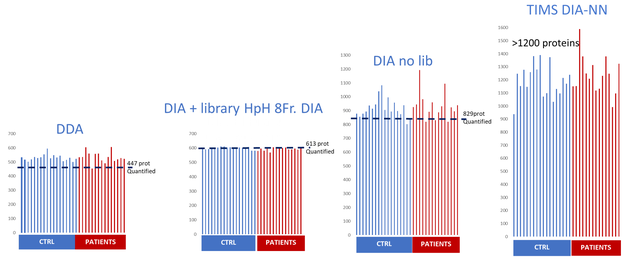
- ✓ Depletion improves IDs by only 15%
- ✓ Lowering the MS/MS triggering threshold from 1000 to 500 increases IDs by 8%
- ✓ Widening the mobility range from a 0,8-1,3 V/cm2 to 0,6-1, 4 V/cm2 allowed to increase the number of protein group ID by another 20%.

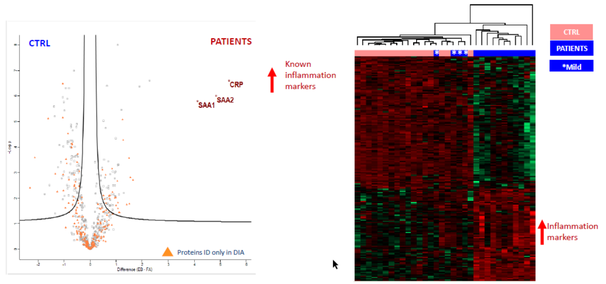
Plasma dia-PASEF

- ✓ DIA of HpH 8 Fr. allowed the ID of **2438 prot**.
- ✓ DDA of HpH 8 Fr. allowed the ID of **1277 prot**.
- ✓ DIANN library-free vs DDA increased IDs by 23%
- ✓ DIA-NN search using libraries from HpH 8 Fr. (DIA or DDA) did not improve the results dramatically
- ✓ TIMS DIA-NN (Bruker) lead to >1200 proteins IDs using plasma + K562 cell lysate library

Case study RD 1

Dermatological rare genetic disease





Technology key points:

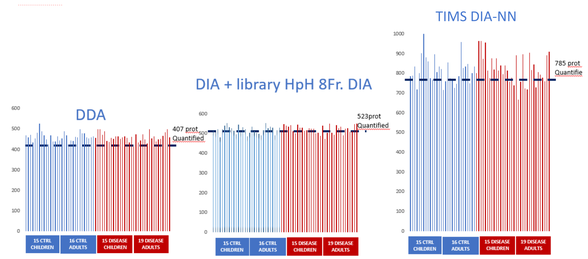
- ✓ 2x more proteins IDs in DIA than DDA
- ✓ >850 proteins per sample with DIA-NN Library-free pipeline
- ✓ >1200 proteins per sample with TIMS DIA-NN pipeline
- ✓ Library free runs works best on large series of samples (MBR)

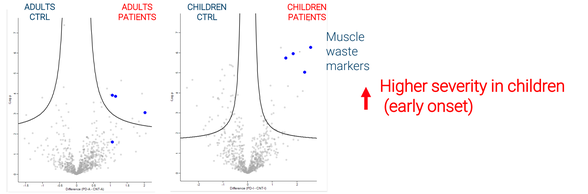
Clinical project key points:

- ✓ Detection of markers from derma
- ✓ Upregulation of known inflammation proteins
- ✓ Stratification of patients according to severity

Case study RD 2

Neuromuscular rare genetic disease





Clinical key points:

- ✓ Upregulation of expected proteins implicated in the metabolism defect
- ✓ Higher signature in patient with early onset of the disease (children) compared to late onset (adults)

Conclusion

High-throughput analysis of non-depleted, non fractionated, non pre-treated, plasma leads to over 1200 proteins ID. The gradients lengths used here (40min) can still be reduced, but is sufficiently short to handle cohorts of patients affected by rare genetic disease (usually <100 samples).

dia-PASEF® and **PASER data analysis** (TIMS DIA-NN) maximize the number of proteins groups identified and quantified, allowing the identification of proteins from tissue leakage pertinent to the disease.