

An intelligent Hybrid-DIA data acquisition strategy for cracking the clinical sample complexity challenge in translational proteotyping

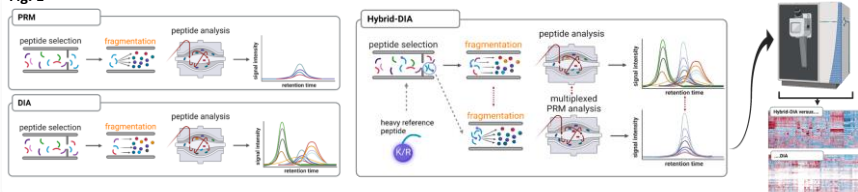
Sandra Goetze^{1,2}, Kyle Fort³, Jesse D. Canterbury⁴, Audrey van Drogen^{1,2}, Alexander Makarov³, Bernd Wollscheid^{1,2}, Yue Xuan³

¹ Swiss Federal Institute of Technology Zurich (ETHZ), Switzerland; ² Swiss Institute of Bioinformatics (SIB), Switzerland; ³ Thermo Fisher Scientific Bremen, Germany; ⁴ Thermo Fisher Scientific, San Jose, USA

Introduction

MS-based proteotyping has been widely employed for biomarker discovery over the past decades, yet the clinical/translational proteotyping community requires strategies that not only enable the discovery of novel biomarker candidates but can also boost the probability of establishing protein-based biomarker assays, enhance analytical and clinical validation speed, and resolve the issue of data missingness in a quantitative protein matrix. Here we present an intelligent data acquisition Hybrid-DIA strategy enabling the comprehensive digitization of a clinical specimen on the proteotype level while at the same time enhancing measurement sensitivity for a specific set of markers of clinical interest (Fig. 1).

Fig. 1



Results

We tested Hybrid-DIA on a pool of 252 representative proteotypic peptides for tumor associated antigens (TAA) derived from 61 annotated human proteins. We generated mixes containing both the heavy reference peptide as well as its synthetic light isotope. Whereas the heavy reference peptide was kept constant in all samples at approximately 100 femtomole (injected onto column), its light counterpart was measured in a dilution series ranging from 100 femtomole to 100 attomole. We monitored 185 TAA peptides at a time in a scheduled fashion using high-resolution msxPRM while at the same time recording the DIA traces in Hybrid-DIA. Data analysis was performed using Qual Browser, Spectronaut and Skyline. Preliminary data show that for some of the peptides we monitored we observed a lower LOD/LOQ for msxPRM than for DIA, as well as a lower CVs at lower peptide concentrations (Fig. 3A,B).

Fig. 3A

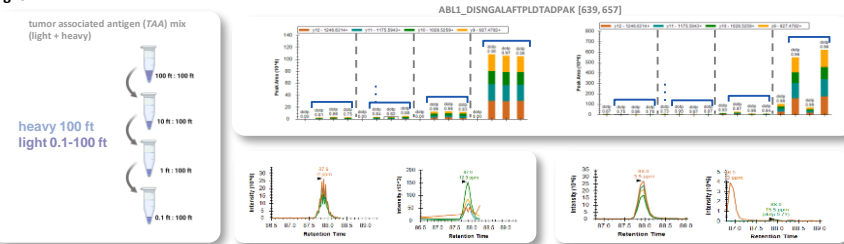


Fig. 3B

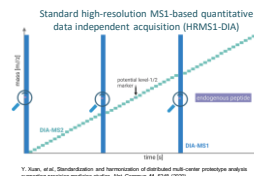


Method

Hybrid-DIA - A strategy to reproducibly quantify clinical marker proteins while at the same time allowing comprehensive sample proteotyping (Fig. 2A, B)

1. The Hybrid-DIA strategy consists of a standard DIA scan cycle, where MS1 scans are followed by several DIA MS/MS scans. Fast (multiplexed) PRM MS/MS scans are triggered based on the detection of isotope-labeled reference peptides and serve as a second layer of confirmation (Fig. 2B).
2. Successful isotope-labeled peptide detection (3x in a DIA-MS1 scan) triggers the high-quality measurement of the corresponding endogenous counter-peptide, multiplexed (msx) with the isotope-labeled peptide through msxPRM MS/MS scans acquired with narrower isolation window width and maximizing ion injection time for each species.
3. The Hybrid-DIA method is programmed in C#, utilizing the Exploris API on an Orbitrap Exploris 480 MS.

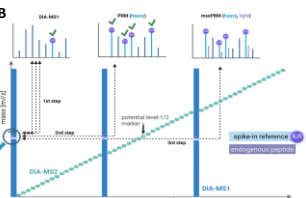
Fig. 2A



2A. Standard data independent acquisition (DIA): all ions are fragmented and detected in DIA MS/MS scans with a wide isolation window width and ion injection time is calculated for all the ions in the DIA window simultaneously.

2B. Hybrid-DIA intelligently acquires targeted peptide species using isotope-labeled references (IS) to on-the-fly trigger msxPRM scans with a narrow isolation window and maximizing the ion injection time of the endogenous peptide. Simultaneously, the entire sample is proteotyped using DIA.

Fig. 2B



Conclusion

We could show that Hybrid-DIA has the potential to monitor clinical marker peptides at a better sensitivity and specificity than DIA alone.

Using Hybrid-DIA we expect to reliably provide clinically actionable information and at the same time new research insights into disease development and treatment trajectories as the basis for discoveries.



- Level 1 / Level 2 diagnostic markers
 - Level-1 clinical data (e.g. drug targets)
 - Level-2 data (markers/pathways of potential clinical relevance)
 - Reporting in molecular summary report / tumor board
- Exploratory Research Loop
 - Level-3 data (proteins/markers of currently unknown clinical values)
 - molecular data (level-3)