Comparison of different processing pipelines for dia-PASEF® data

MP 248

Manuel Chapelle[1] Celine Henry[2], Aaron Millan-Oropeza[2], Pierre-Olivier Schmit[1] and Stephanie Kaspar[3]

[1] Bruker France S.A., Wissembourg, France

[2] Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, PAPPSO, Jouy-en-Josas, France.

[3]Bruker Daltonik GmbH, Bremen, Germany

Introduction

dia-PASEF promises better ion utilization as compared to traditional DIA approaches, as well as the PASEF –specific increase of specificity and sensitivity. This allows for optimum data completeness in label-free quantification approaches. Nevertheless the use of the peptide's collisional cross section as a characteristic value requires to modify both the library generation and comparison processes. The importance of the processing pipeline to the final analytical outcome is high. We have used different processing pipelines to process datasets generated with the dia-PASEF approach, comparing the computational results.

Methods

Tryptic digests from *Arabidopsis Thaliana (ARATH)*, *Yarrowia Lipolytica (YWL29)* and *Streptomyces Coelicolor (SCoeli)* sources were pooled in two distinct proportions (0.33/1.5/1 for plant, yeast and bacteria, respectively) to create two distinct and complex samples.

Eight fractions were generated from those samples (using a Pierce kit) to generate the libraries. All samples were separated by nano-HPLC (nanoElute, Bruker Daltonics) on 250 mm pulled emitter column (IonOpticks, Australia) with a 90 min gradient and analyzed on a timsTOF Pro instrument (Bruker Daltonics).

The timsTOF Pro was operated in data dependent PASEF mode as well as in dia-PASEF(Fig.1). Collisional Cross Section (CCS) aware library building and data processing have been performed by using either combination of MaxQuant (2019) and Mobi-DYK (Toronto Univ) with OpenSwath (OpenMS) or the Spectronaut 14 software suite (Biognosys) or the Peaks StudioX+ software suite (BSI).

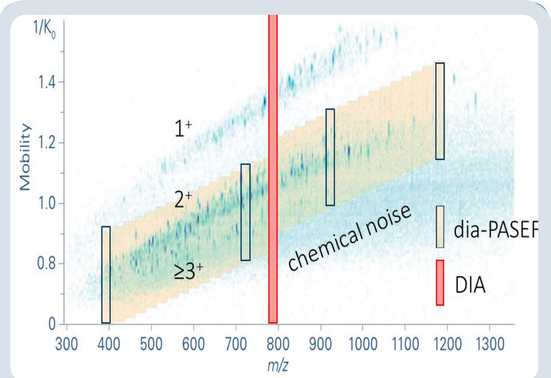


Fig. 1: dia-PASEF method overview :
Each dia-PASEF step uses four 25 Da * 0.4 1/k0 windows per PASEF event. 16 Steps (for a total of 64 windows) are used in a cycle to cover the m/z and 1/k0 range that contains all 2+ and3+ precursors. The resulting cycle time is 1.7sec.
In comparison to standard DIA (in red), dia-PASEF allows a focus on the ion of interest (here 2+ and 3+) and an increased ion usage.

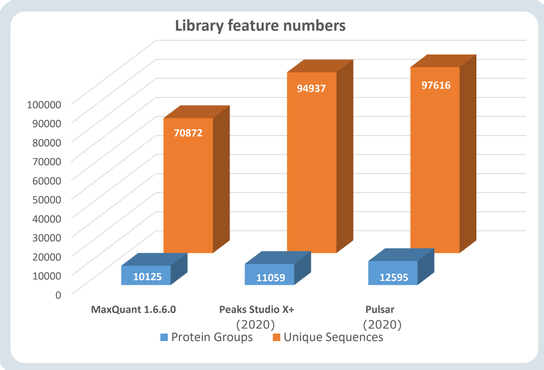
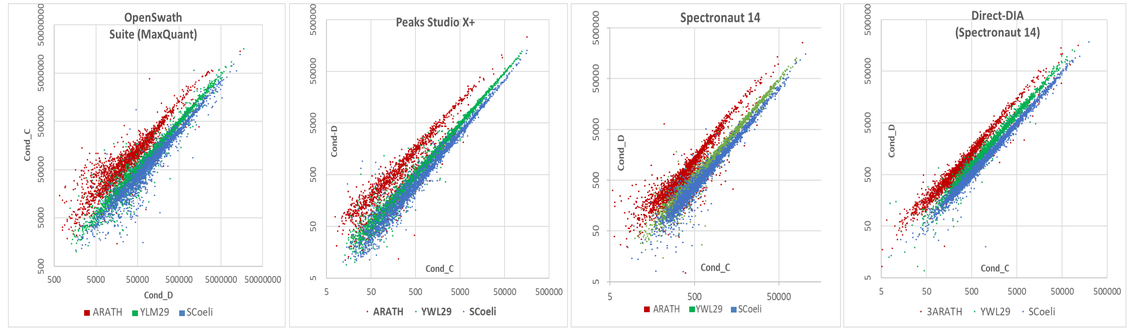


Fig. 2: Library generation :
MaxQant 1.6.6.0 was used for the Mobi-DYK/OpenSwath processing pipeline. Peak's and Spectronaut's (Pulsar) own search engines were used to generate the libraries in the PeaksX+ and Spectronaut 14 software suites, respectively.



Ratio D/C	YWL29	SCoeli	ARATH
Expected value.	1.07	1.61	0.32
Mobi-DYK OpenSwath	1.08	1.66	0.44
PEAKS Studio X+	1.07	1.57	0.61
Spectronaut 14	1.03	1.44	0.55
Spectronaut 14 Direct-DIA	1.03	1.48	0.42

Fig. 4: Quantification results
Quantification settings were kept similar across pipelines. Plotting has been performed using Excel

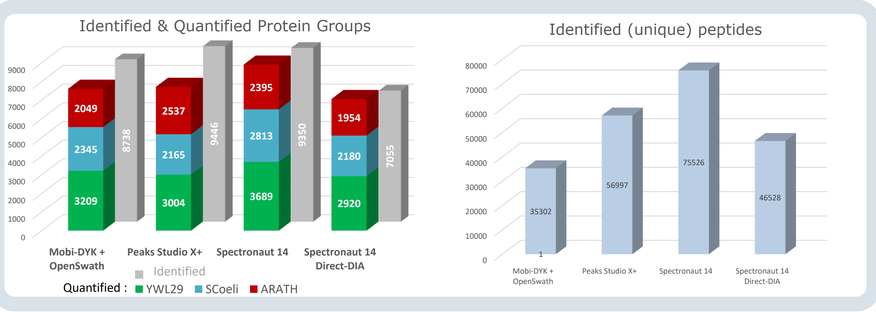


Fig. 3: Overall Identification results
All identification settings were kept similar across pipelines, and a 1% FDR filtering was applied at peptide level to all results.

Results

All search engines allowed to build CCS-aware libraries containing an excess of 10 000 protein groups and 70 000 unique peptide sequence (>90 000 for Peaks and Spectronaut) @ 1% FDR (Fig.2)

The percentage of library recovery for a 90min gradient is exceeding 70% for all library-based approaches and remains superior to 55% for the Direct-DIA approach (Fig.3). All workflows yielded correct ratio determination (Fig4). It must be underlined that these results have evolved significantly as new versions of those software were released. The Mobi-DYK/OpenSwath results might benefit from the use of a more recent version of MaxQuant.

Conclusions

- The ion mobility dimension is fully supported in these three processing pipelines. Consequently, the use of the 4th dimension is beneficial both at the acquisition and the processing stages, especially ensuring a much improved selectivity and result certainty.
- All available processing pipelines allow to generate Collisional Cross Section (CCS) aware libraries and process dia-PASEF data successfully.
- The dia-PASEF approach is now supported by a variety of freely available and commercial softwares for data processing and valuation.

timsTOF PRO