X!TandemPipeline++: Software for Ion Mobility-Enabled Quantitative Proteomics in timsTOF Data Format

Olivier Langella*, Thomas Renne, Thierry Balliau, Marlène Davanture, Sven Brehmer, Michel Zivy, Mélisande Blein-Nicolas and Filippo Rusconi

olivier.langella@universite-paris-saclay.fr - filippo.rusconi@universite-paris-saclay.fr

IDEEV - Institut diversité, écologie et évolution du vivant
GQE-Le Moulon, PAPPSO
Université Paris-Saclay, INRAE, CNRS, AgroParisTech,
F-91190 Gif-sur-Yvette, France

ABSTRACT

X!TandemPipeline (Langella et al. 2017) is a free and open source Java software program for proteomics initially designed to filter and group peptide/protein identifications from MS/MS mass spectra. After a complete rewrite in C++17, the new *X!TandemPipeline*++ program features peptide/protein quantification, in particular with support for native timsTOF raw data.

X!TandemPipeline++ performs peptide identifications and area-under-the-curve XIC-based quantifications using the Bruker's native timsTOF raw data format. Using a common HeLa data set, published by Meier *et al.* (2018, PXD010012), we demonstrate that *X!TandemPipeline*++ identifies and quantifies significantly more proteins than the *MaxQuant* and *MSFragger* competitors, while being also significantly faster.

Please, note that *X!TandemPipeline*++ will soon be renamed to ***i2MassChroQ*** to reflect the identification & inference — mass chromatogram-based quantification capabilities.

RESULTS

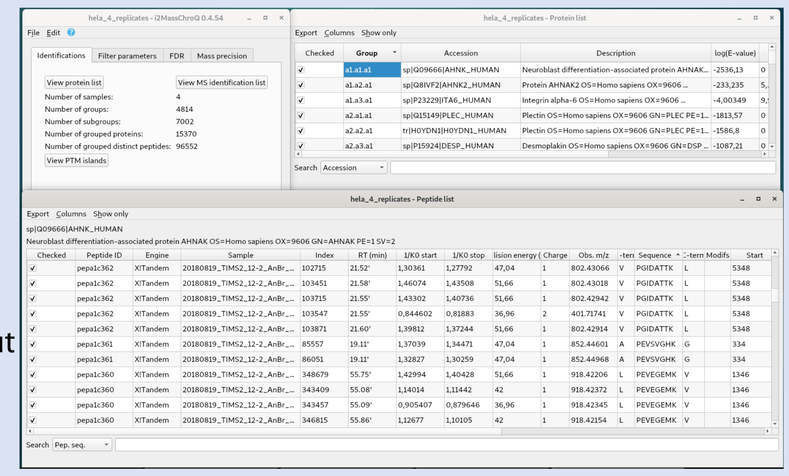
Identification and protein inference

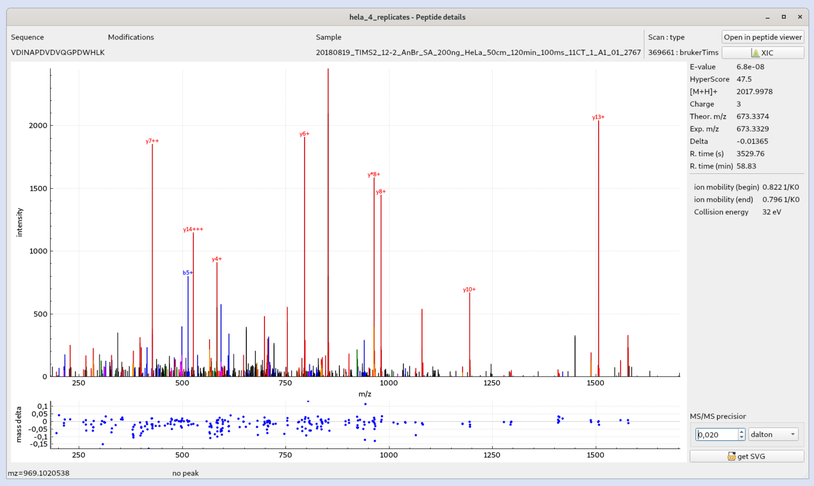
X!TandemPipeline++ features high throughput and high resolution proteomics and metaproteomics features.

Thousands of DDA identification results can be loaded at once as part of the same project.

Our software supports *Mascot*-, *X!Tandem*-generated identification data or any data set in mzIdentML and pepXML formats (from *OpenMS*, *Comet*, *PeptideProphet*). Its Occam's razor-based grouping algorithm can infer protein groups very quickly, in particular by filtering-out non informative redundancies (Van Den Bossche et al., nature communications, 2021).

Grouping might be performed by processing each sample separately or by creating a consensus protein list starting from multiple combined samples.

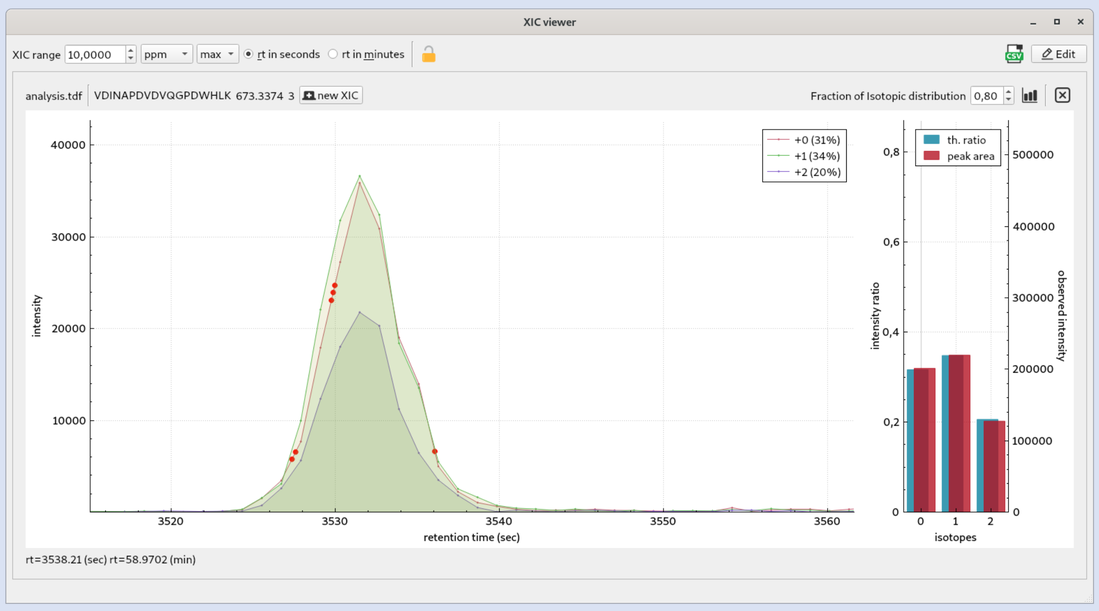




Quantitative analysis

Peptide quantification is handled by calculating extracted ion current chromatograms from precursor ion MS data (MassChroQ module).

Protein quantification is performed by the MCQR module that computes protein intensities by filtering and aggregating peptide intensities and carrying over cluster analysis, ANOVA, and comparisons between treatments.



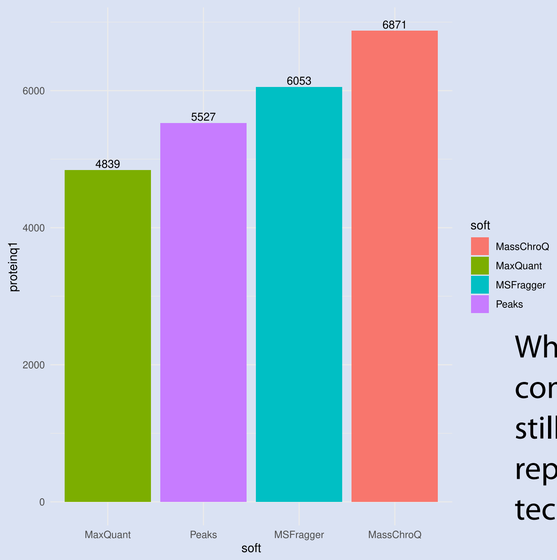
timsTOF data native support

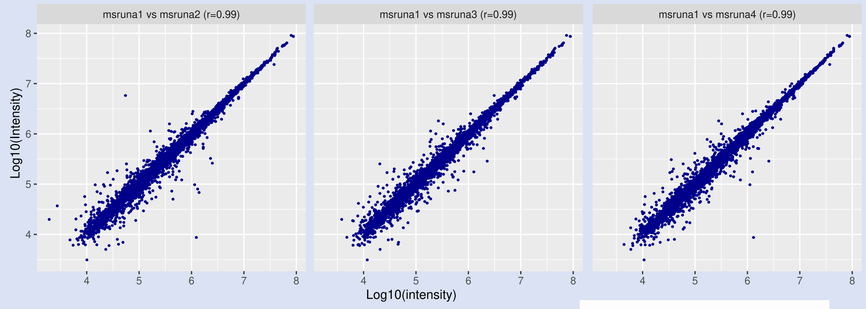
Our native C++-based timsTOF raw data reader was developed from scratch thanks to the technical specifications provided by Bruker. Tight optimizations were needed to ensure super-fast access to the Bruker-encoded binary data.

The current version provides real time MS/MS peptide annotation and performs extremely fast ion current extraction and XIC chromatogram visualization.

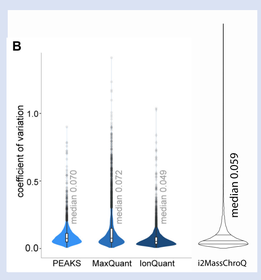
Full processing of a 2 hour PASEF run can be performed in less than 10 s for 1 peptide, less than 30 m for 100 000 peptides+isotopes XICs, peak detection and area under the curve computation.

TimsTOF PASEF quantification reproducibility

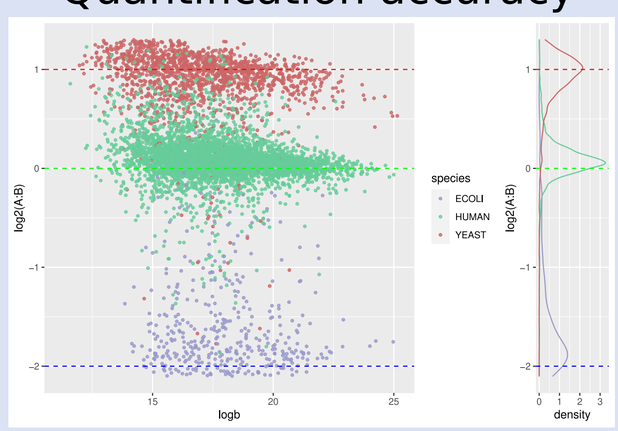




While quantifying more proteins than competitors (1 ion or more), X!TandemPipeline still shows excellent correlations between replicates and very low median CV (4 Hela technical replicates, PXD10012)



Quantification accuracy



Plot of the 6428 proteins quantified (2 ions or more) and compared to ground truth ratio (2:1 for Yeast, 1:1 for Human, 1:4 for Ecoli, PXD01477)

AVAILABILITY

i2MassChroQ and its user manual are available as Free Software, licensed under the GNU General Public License (v3 +).
The user manual is available in PDF or HTML format. *i2MassChroQ* is developed on Debian GNU/Linux.
Binary packages are built for GNU/Linux and MS Windows.
GIT repository: https://forgemia.inra.fr/pappso/i2masschroq.

Main website

User manuals HTML PDF

* presenting and corresponding author