



HAL
open science

i2MassChroQ: full native timsTOF PASEF-enabled quantitative proteomics

Olivier Langella, Thierry Balliau, Marlène Davanture, Mélisande Blein-Nicolas, Filippo Rusconi

► **To cite this version:**

Olivier Langella, Thierry Balliau, Marlène Davanture, Mélisande Blein-Nicolas, Filippo Rusconi. i2MassChroQ: full native timsTOF PASEF-enabled quantitative proteomics. EuBIC-MS Winter School 2024, Jan 2024, Winterberg, Germany. hal-04321638

HAL Id: hal-04321638

<https://hal.science/hal-04321638v1>

Submitted on 10 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

i2MassChroQ : full native timsTOF PASEF-enabled quantitative proteomics

Olivier Langella*, Thierry Balliau, Marlène Davanture, Mélisande Blein-Nicolas, Filippo Rusconi

olivier.langella@universite-paris-saclay.fr

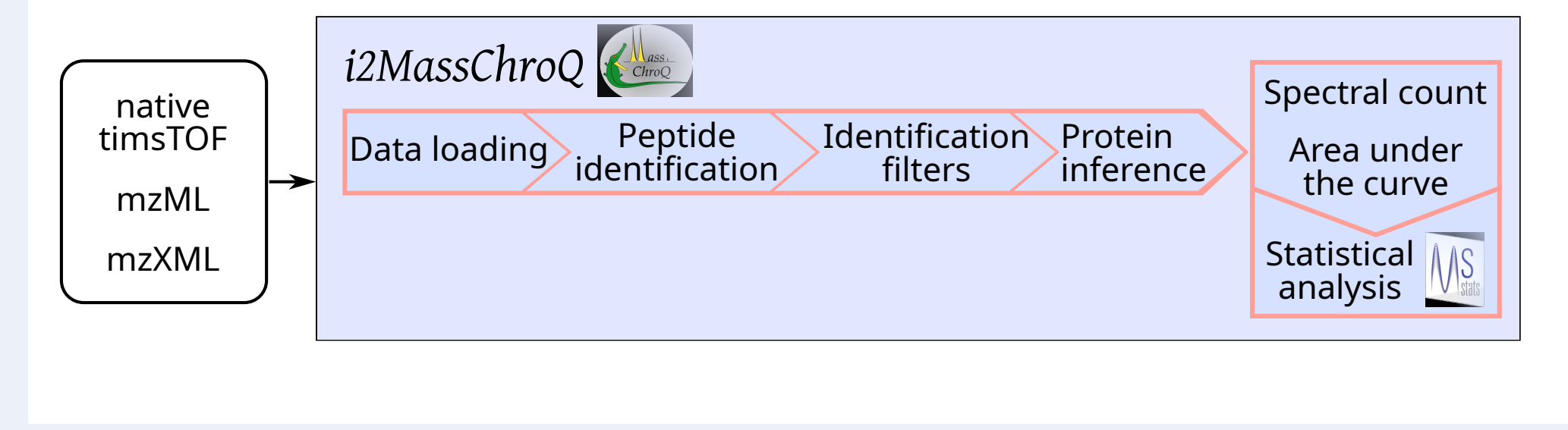


IDEEV - Institut diversité, écologie et évolution du vivant
GQE-Le Moulon, PAPPSO
Université Paris-Saclay, INRAE, CNRS, AgroParisTech,
F-91190 Gif-sur-Yvette, France

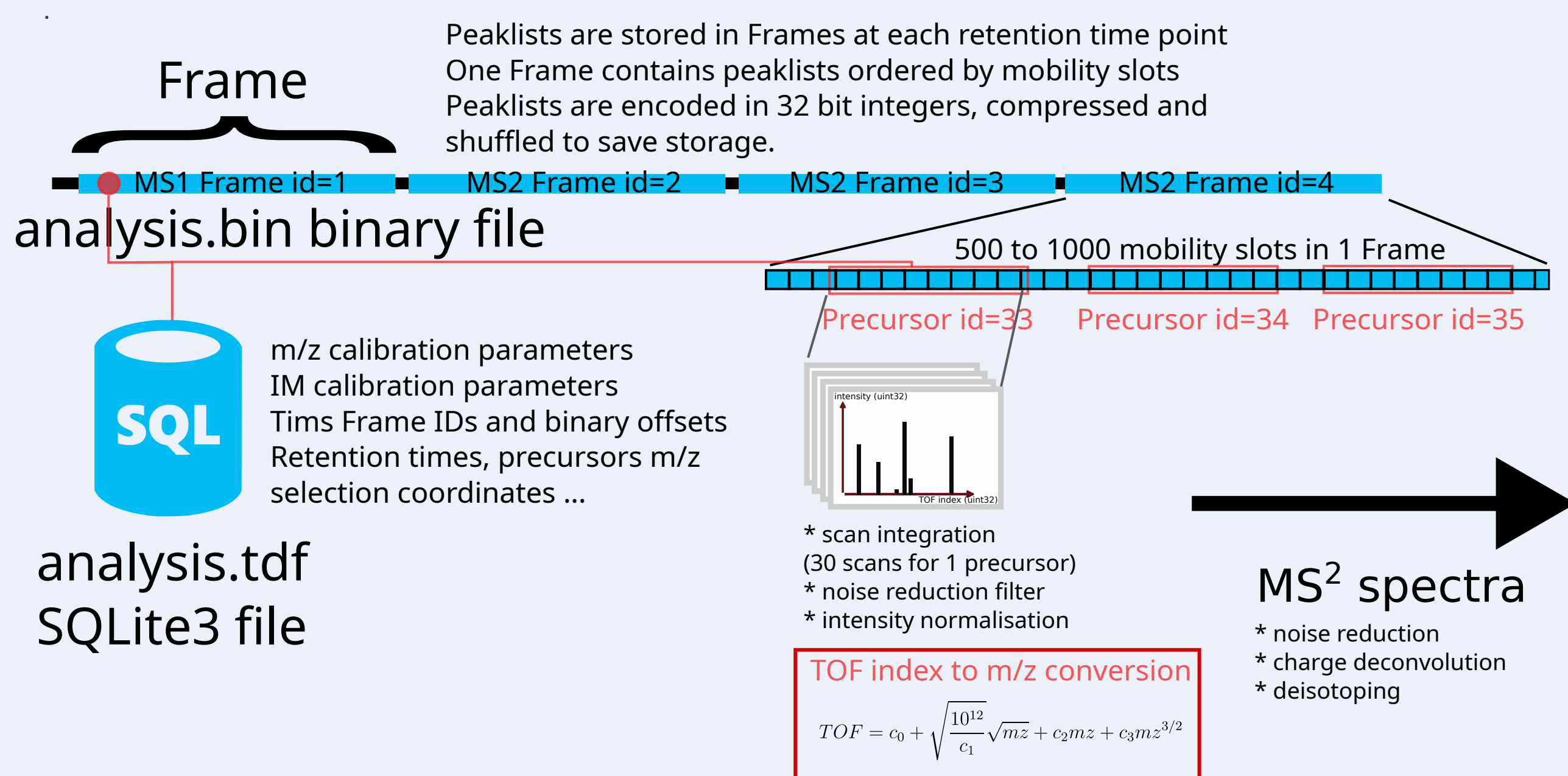


ABSTRACT

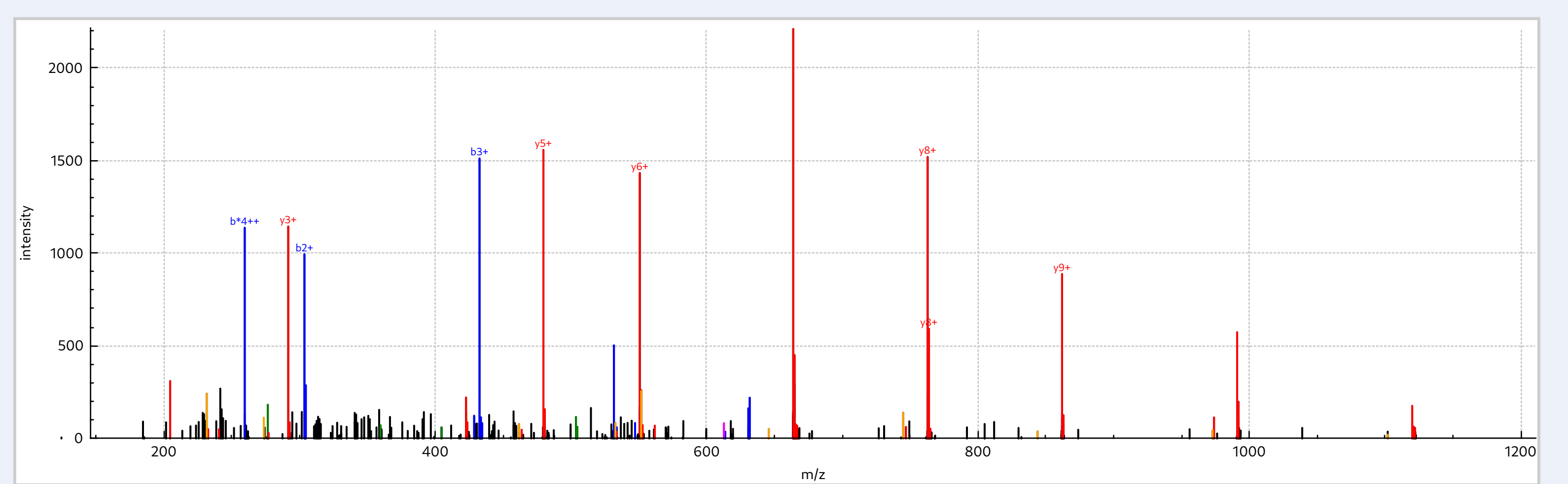
X!TandemPipeline (Langella *et al.* 2017) is a proteomics free and open source Java software program designed to filter and group peptide/protein identifications from MS/MS data that is used with the *MassChroQ* quantitative proteomics software (Valot *et al.* 2011). After a complete rewrite in C++17, the new software, named *i2MassChroQ*, now features native support for the timsTOF raw data format, peptide/protein quantification by merging the features initially in *X!TandemPipeline* and *MassChroQ*, and ultimately statistical analysis using either the *MSstats* or the *MCQR* GNU R package.



Native support for timsTOF raw data format

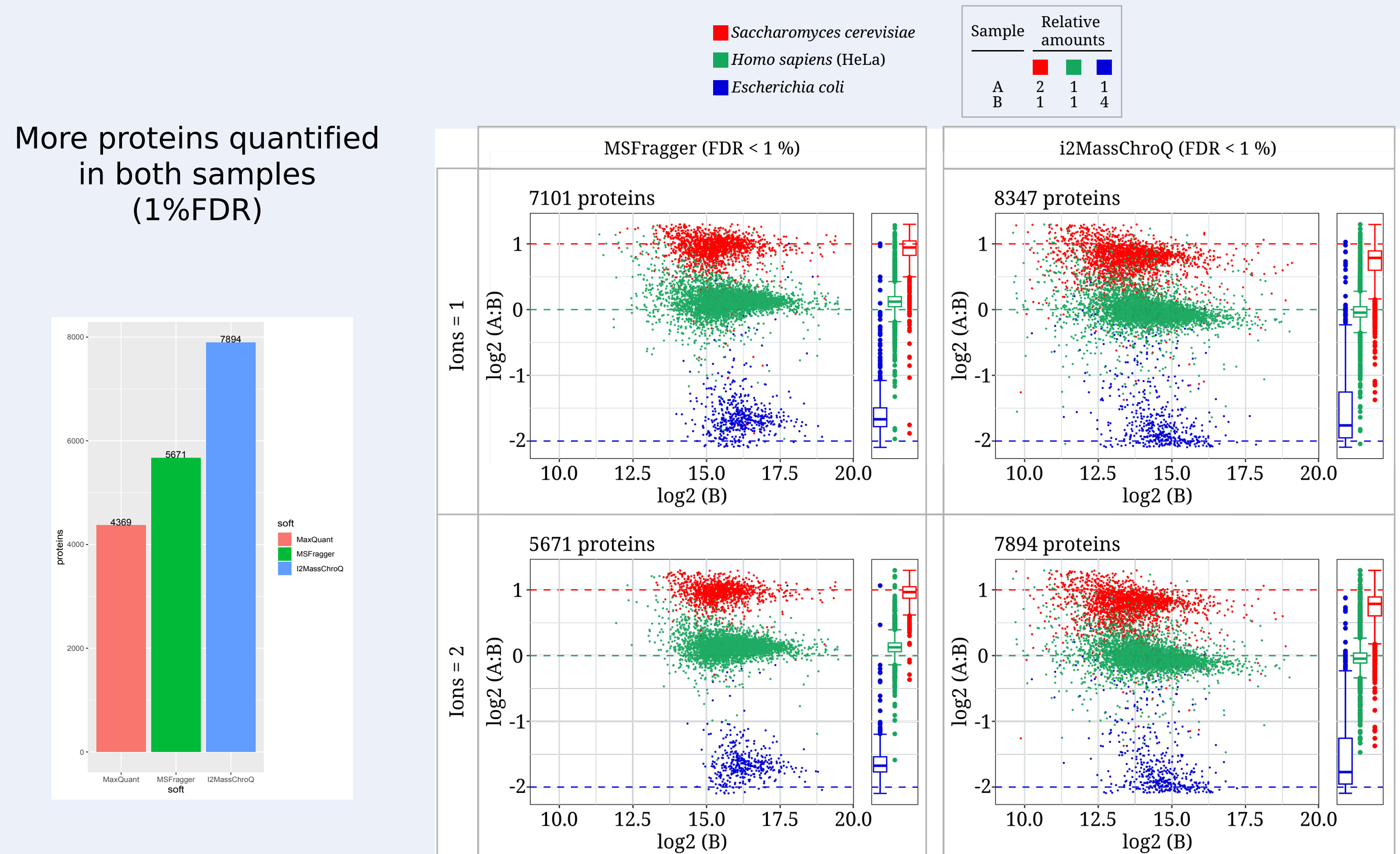


The complete control over raw data with our parser allows for optimizations at each stage, e.g. from input/output cache (very efficient over the network), polynomial solving, multi-tasking, MS² data-specific filters depending on the identification engine.



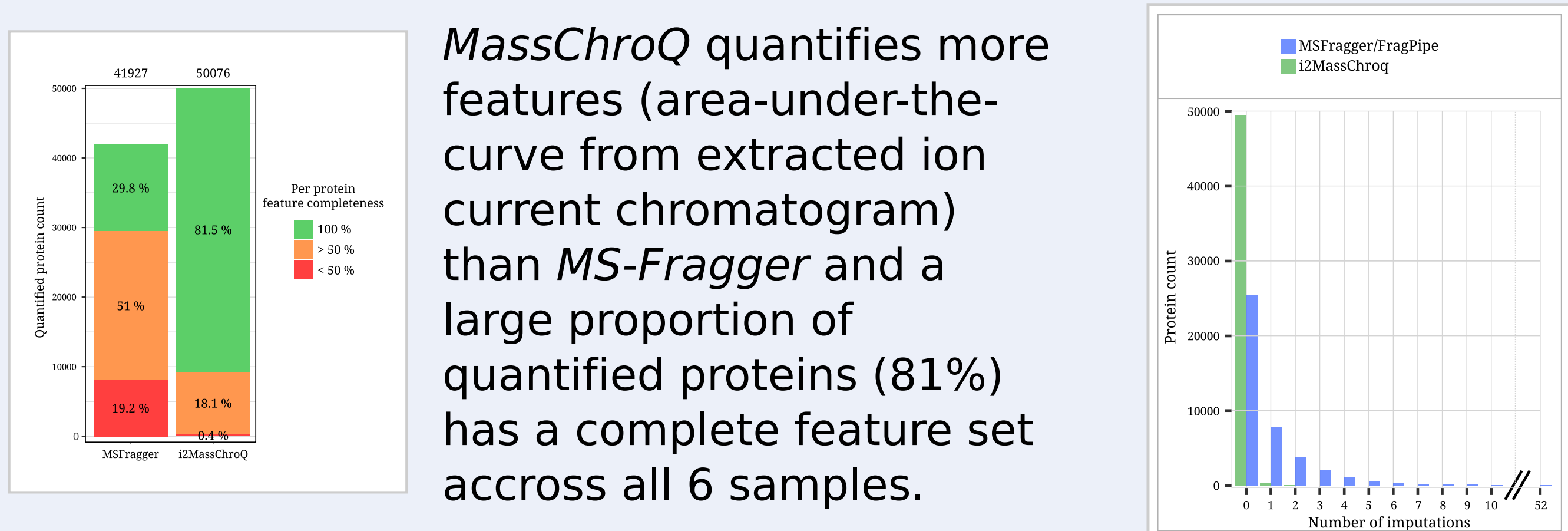
Benchmarking software with timsTOF dataset PXD014777

Two samples, A and B, containing a mix of protein extracts from three different organisms. 3 replicates each.

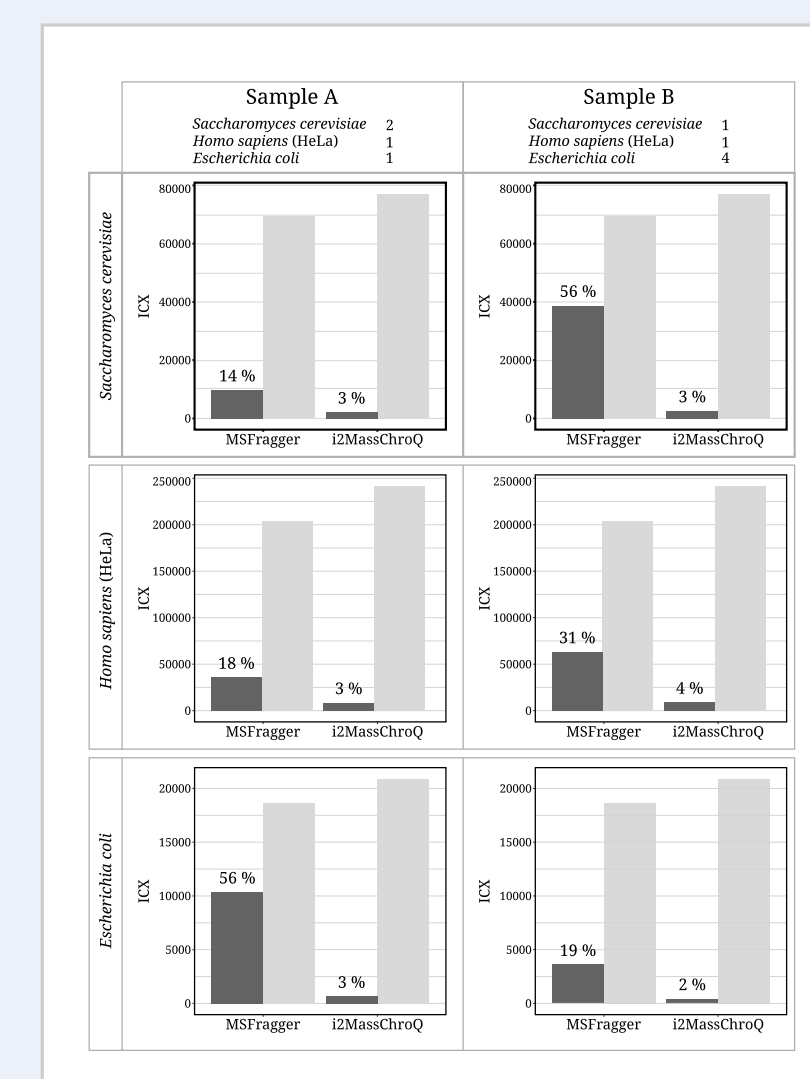


i2MassChroQ quantifies more proteins and on a wider dynamic range.

Data completeness : MSFragger 72%, i2MassChroQ 96%



MassChroQ quantifies more features (area-under-the-curve from extracted ion current chromatogram) than *MS-Fragger* and a large proportion of quantified proteins (81%) has a complete feature set across all 6 samples.



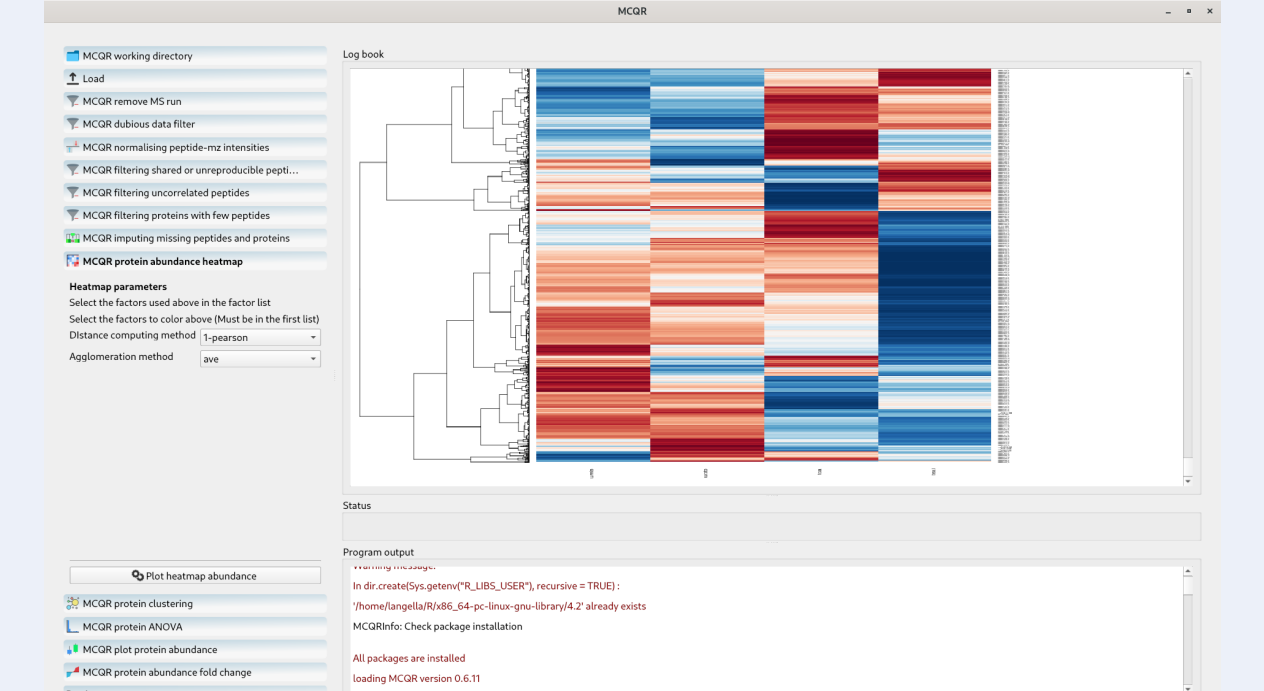
The *MassChroQ* match-between-run algorithm, specifically designed for timsTOF data, lowers dramatically the proportion of missing values, especially in unfavourable contrasted conditions.

MSstats imputation:

For each protein quantification in all 6 samples, the feature set may be incomplete and *MSstats* need to recourse to imputations. This is the case for more than 50% of the protein quantification values in the case of *MS-Fragger*, compared to less than 1% for *MassChroQ*

User interface

i2MassChroQ has a complete graphical user interface to handle your proteomics projects, from raw MS data to statistical analysis (*MSstats* and *MCQR*).



CLI tools

mzxmlconverter

Fast conversion with customisable settings from timsTOF raw data directory to mzXML standard file to use with any identification engine.

tandemwrapper

Transparent use of the X!Tandem identification engine on timsTOF data repository or using non supported mzXML encoded files. It first transcodes timsTOF data to mzXML, runs X!Tandem and rewrites results as if X!Tandem had been used in the first place.

xicextractor

Single xic extraction tool using customisable settings, working on any MS data file, including timsTOF directory (ion mobility enabled).

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



* presenting and corresponding author