



HAL
open science

Open Modification Search using SpecOMS and i2MassChroQ

Olivier Langella, Luciana de Oliveira, Michel Zivy, Melisande Blein-Nicolas,
Dominique Tessier

► **To cite this version:**

Olivier Langella, Luciana de Oliveira, Michel Zivy, Melisande Blein-Nicolas, Dominique Tessier. Open Modification Search using SpecOMS and i2MassChroQ. analytics 2022, Sep 2022, Nantes, France. hal-03735820

HAL Id: hal-03735820

<https://hal.science/hal-03735820v1>

Submitted on 2 Sep 2022

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 4.0 International License

Open Modification Search using SpecOMS and i2MassChroQ

Olivier Langella*¹, Luciana De Oliveira¹, Michel Zivy¹, Mélisande Blein-Nicolas¹, Dominique Tessier²

olivier.langella@universite-paris-saclay.fr



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



2 - INRAE, UR1268 BIA, F-44316 Nantes, France
INRAE, PROBE Research Infrastructure, BIBS Facility,
F-44316 Nantes, France

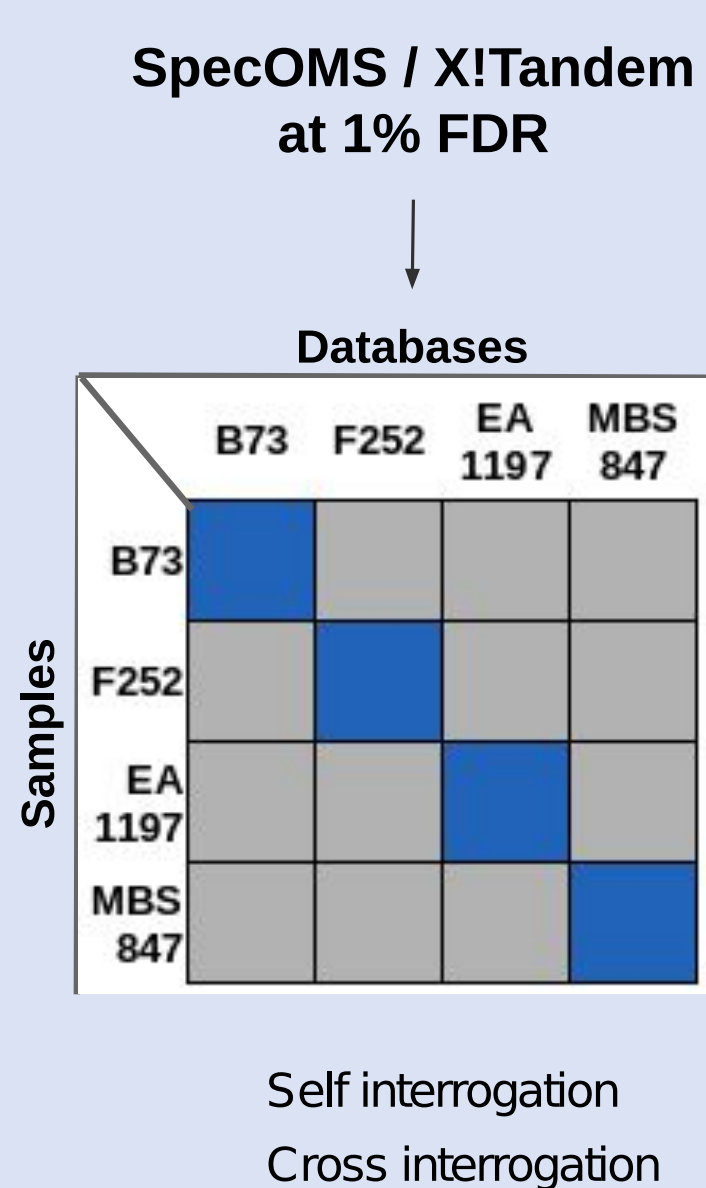


ABSTRACT

SpecOMS is an Open Modification Search method published originally in David et al. 2017. SpecOMS compares experimental spectra generated by a discovery proteomics experiment to a whole set of theoretical spectra deduced from a protein database in a few minutes on a standard workstation. This allows the identification of unknown modifications on peptides and reduces the proportion of undecided peptide spectrum matches compared to classical identification engines. We present here an evaluation of the SpecOMS accuracy to identify modified peptides and a dedicated graphical user interface that helps to browse and visualize peptides identified with this method.

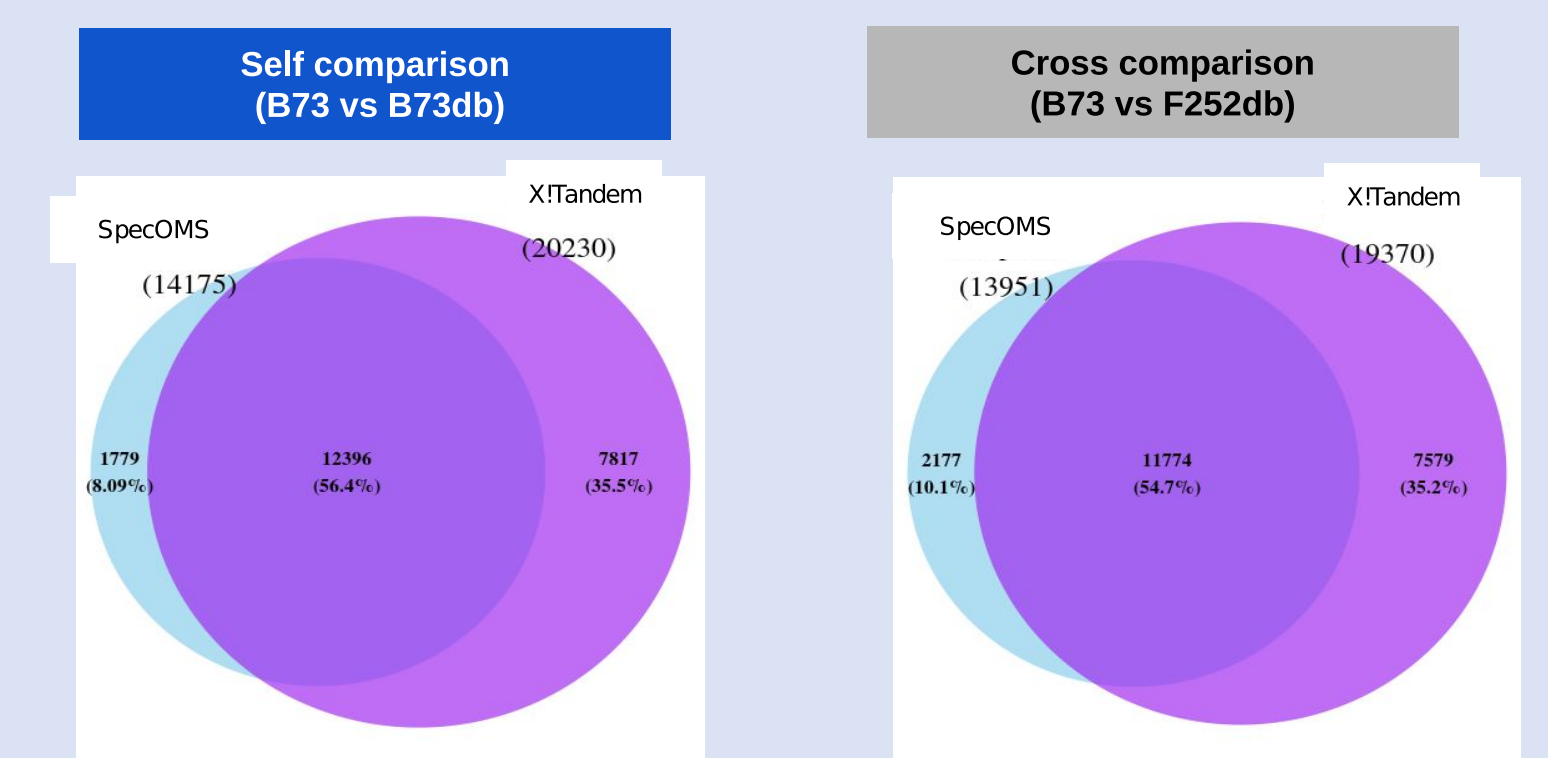
RESULTS

Identification of Single Amino acid Variants (SAVs)



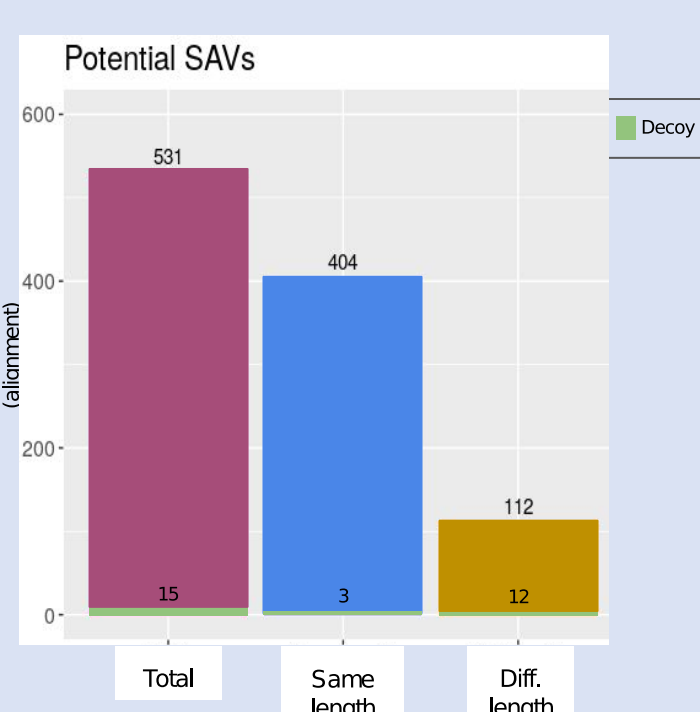
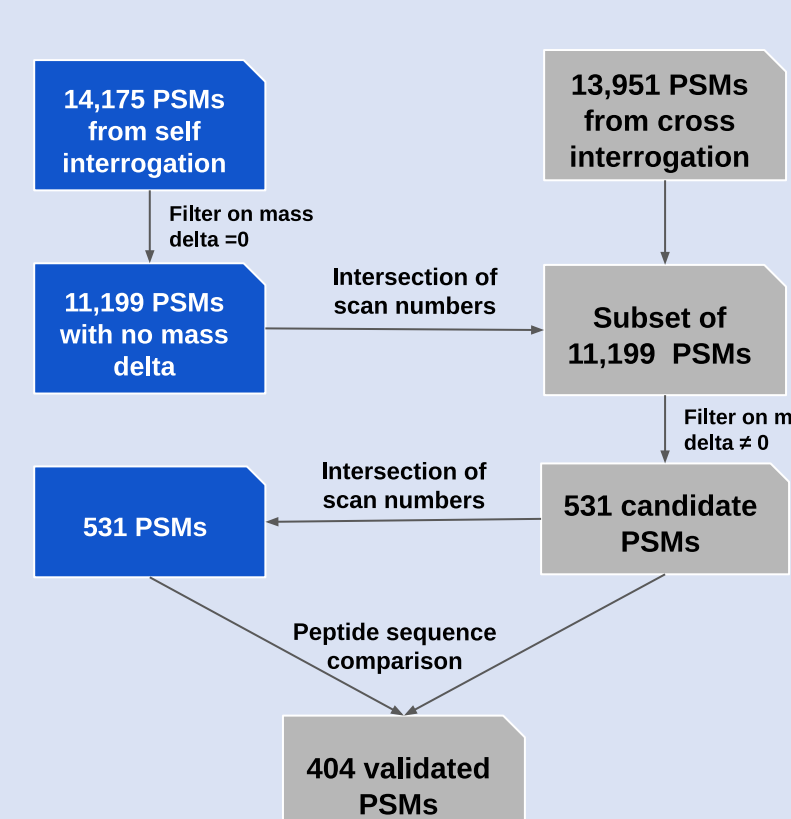
The ability of SpecOMS to identify slightly modified peptide sequences compared to the search sequence database has been tested using 4 well known maize genotypes. In each genotypes, same protein with different sequences wearing Single Amino acid Variants (SAV) are present.

Identifications on a biological sample containing B73 material are made using the B73 search database or a search database from another genotype (F252, EA1197 or MBS847). In the first case, this is a "self interrogation" and the other a "cross interrogation". Identification results from a classical identification engine (X!Tandem) and from SpecOMS are compared (FDR < 1%).

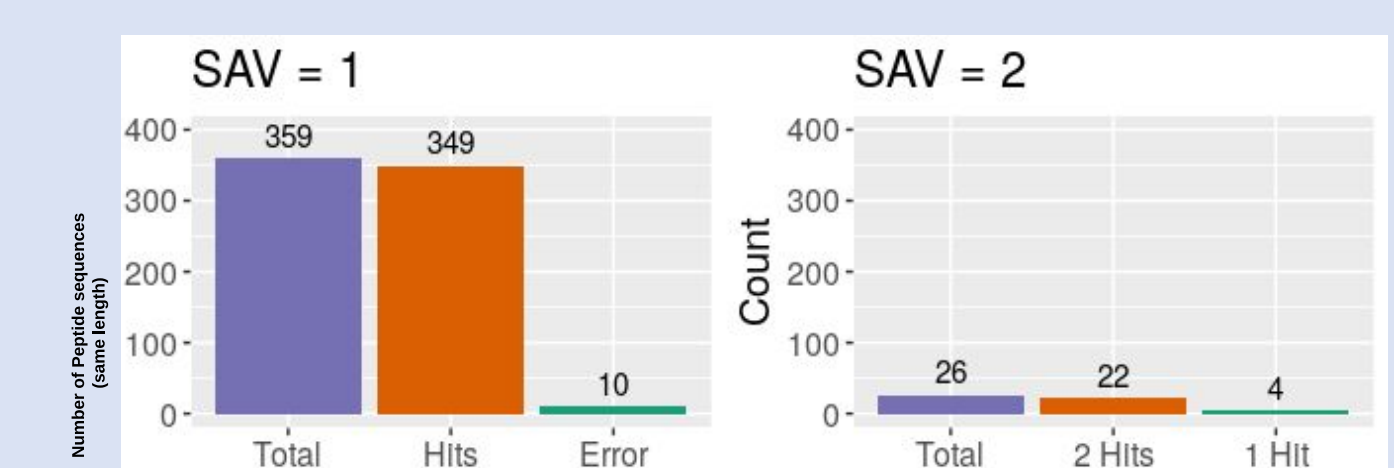
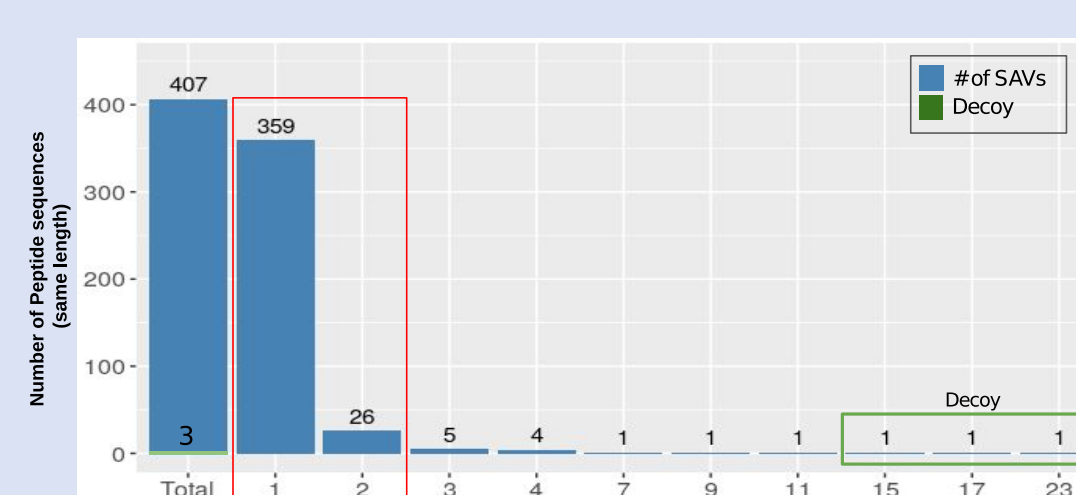


Peptide Spectrum Match (PSMs) comparison between X!Tandem and SpecOMS in self or cross interrogation condition : X!Tandem identifies more PSMs than SpecOMS, in self or cross interrogations, but X!Tandem loses more PSMs in cross interrogations than SpecOMS. This means that SpecOMS can still identifies peptides that are not exactly represented in the search database.

SAV analysis and position prediction



A strong majority has the same same length and has been found on self interrogation : These are potential SAVs.

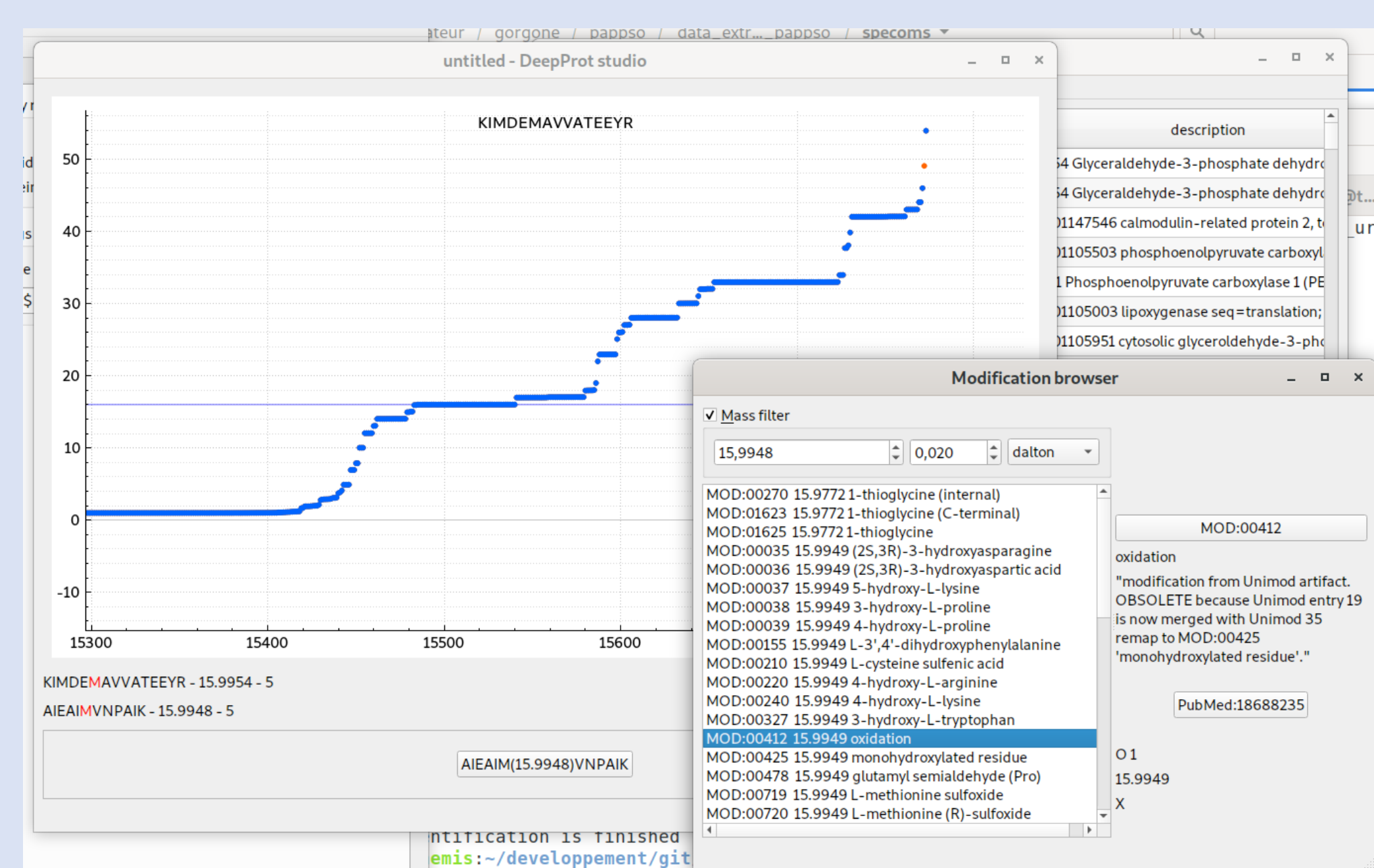


To assess that SpecOMS specific PSMs obtained by cross interrogation are relevant, predicted peptide sequences have been blasted on the original genotypes.

Among PSM with potential SAVs, very few were found on the decoy database.

SpecOMS is able to predict the SAV position on the peptide sequence. The original algorithm called SpecFit is being replaced by a more powerful one using dynamic programming : SpecGlob. This new algorithm will be more reliable especially to infer peptide sequence with multiple mass delta (2 SAVs or more).

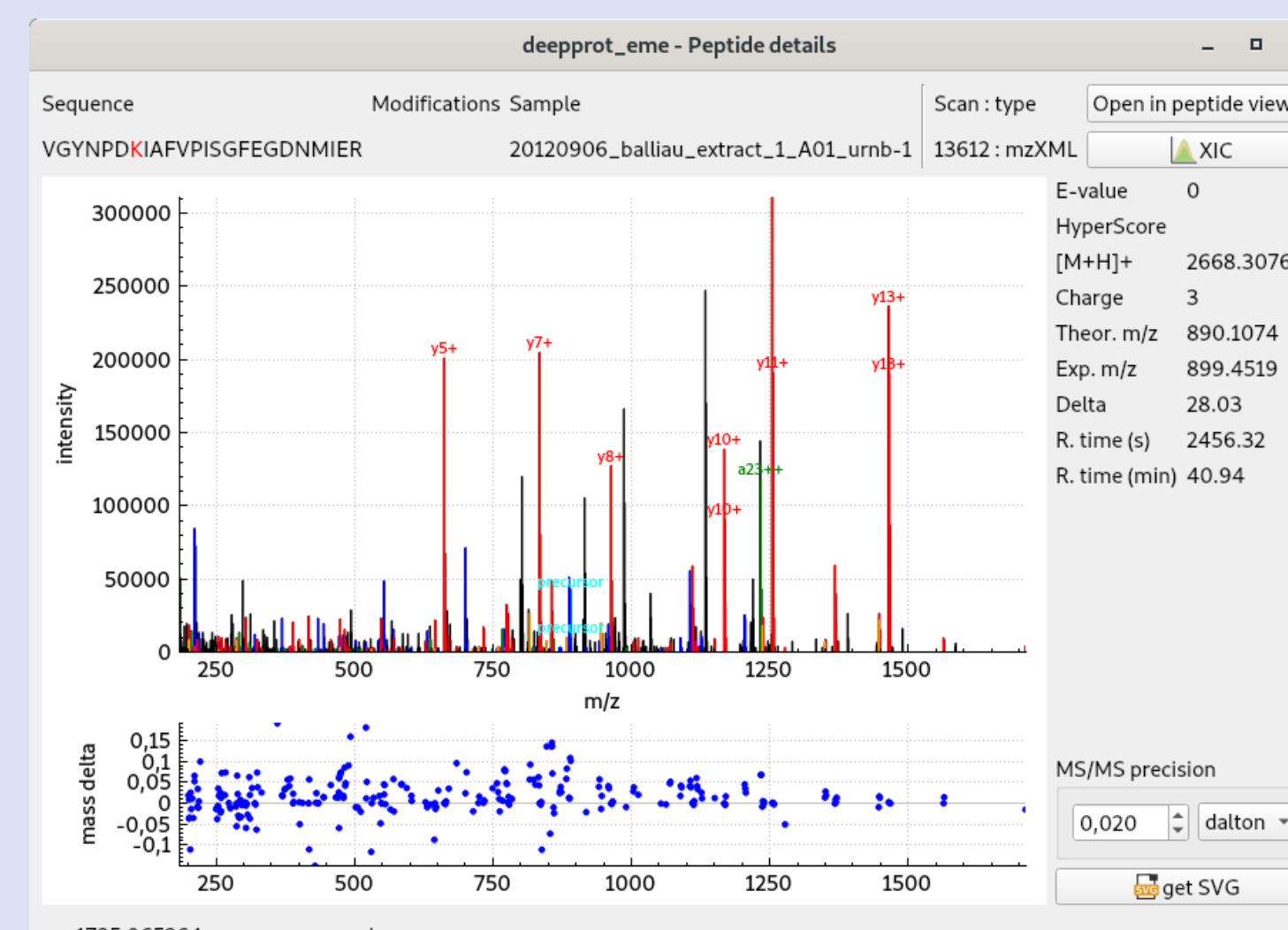
i2MassChroQ integration



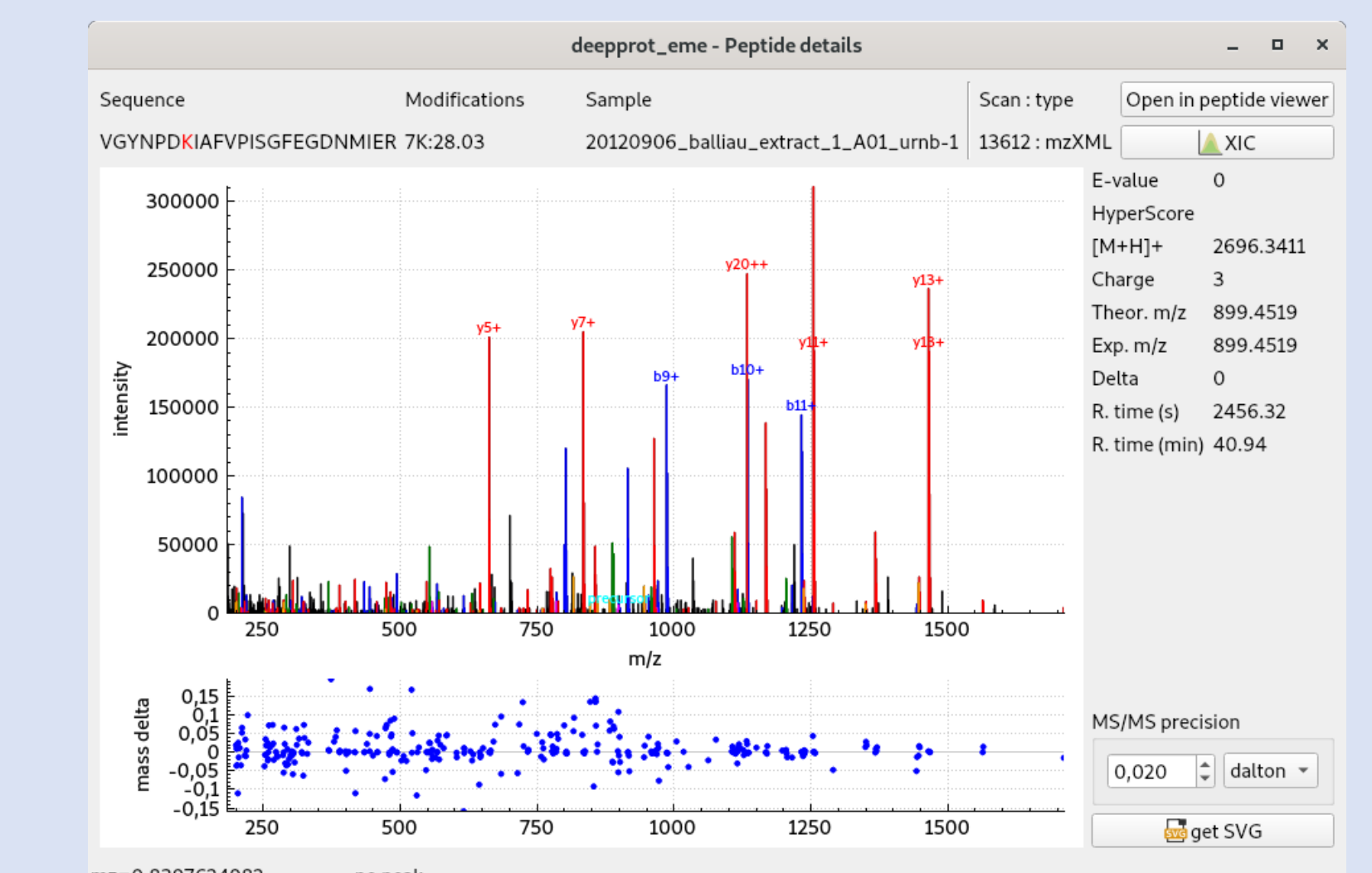
SpecOMS identification engine exports results in the standard mzIdentML file format. The graphical interface of i2MassChroQ reads the SpecOMS file and provide an interactive mass delta browser dedicated to Open Modification Search engines.

Each blue dot represents a PSM, sorted by his mass delta (mass difference between the database hit and the precursor mass).

Spectrum annotation before SpecFit



Spectrum annotation after SpecFit



i2MassChroQ provides a list of the amino acid modifications matching PSM mass delta. Each modification can be placed on the PSM and the user can visualize the fit on the annotated mass spectrum.

SpecOMS can use directly mzML, mzXML, MGF or TimsTOF raw data files.

AVAILABILITY

SpecOMS C++ identification engine source code or binary packages for Windows and Linux are available on demand to the PAPPSO facility or on the ForgeMIA project (<https://forgemia.inra.fr/pappso/specoms>).

The source code is licensed under the GNU General Public Licence v3+. Binary packages are built for GNU/Linux and MS Windows.

* presenting and corresponding author

this work is funded by ANR project ANR-18-CE45-0004