Mass spectrometry

SpecOMS: software for exploring the protein universe

The protein universe is still largely unchartered territory: 99.8% of the proteins referenced in the Uniprot database have been predicted *in silico*—and so not strictly identified—based on the genomic sequence data available. Tandem MS/MS-mode spectrometry is the go-to technique most widely used for protein characterization, but the quality of the results depends on the sample preparation protocol, on the analytical mass spectrometry expertise mobilized, and on the ability of the software to handle and interpret the tens of thousands of spectra generated. This interpretation of the spectra is a time-consuming step, and only around 25% of tandem spectra get successfully interpreted. This low rate is widely explained by the presence of modifications carried by proteins and that are a *priori* unknown. These modifications can reflect post-translational modifications that are essential for protein activity, or they can result from variants that could explain certain phenotypes or certain disease processes. Deciphering and unravelling these modifications, which are thought to be displayed on practically all proteins, is a major scientific challenge for progress in healthcare and biology.

► RESULTS

The SpecOMS software demonstrates that the fragment ion accuracy reachable with the latest generations of mass spectrometers paves the way to a new generation of spectra interpretation algorithms. By exploiting a reorganization of the spectral data at all-sample level in an appropriately-geared data structure and with efficient data access modes, SpecOMS can handle pairwise comparisons of the tens of thousands of experimental spectra from tandem mass analysis against hundreds of thousands of spectra, such as the spectra corresponding to the human proteome. SpecOMS is currently the world's fastest mass spectral analysis software (taking just minutes on a standard desktop PC) and the least-memory intensive, which makes it easy to use on any and all mass spectrometry platforms. The method provides a profile of the modifications brought by the proteins in a sample, and can even reveal modifications that are simply impossible to get with the other approaches available.

► FUTURE OUTLOOK

SpecOMS meets a whole set of the science community's needs, as its rapid adoption in an array of laboratories goes to show. A handful of issues still need to be resolved to ensure the software can exhaustively meet increasingly complex proteomics needs (metaproteomics, foodomics).

Participants

Laboratory of Digital Sciences—Nantes (LS2N) under the Griote programme, financed by the Pays de la Loire region.

Thesis by Matthieu David, under thesis supervisor Guillaume Fertin

Read more

SpecOMS: A Full Open Modification Search Method Performing All-to-All Spectra Comparisons within Minutes

(2017) Journal of Proteome Research

David M, Fertin G, Rogniaux H, Tessier D

SpecTrees: an efficient without a priori data structure for MS/MS spectra identification

(2016) 16th Workshop on Algorithms in Bioinformatics

David M, Fertin G, Tessier D

CONTACTS

Dominique Tessier
dominique.tessier@inra.fr
Hélène Rogniaux
helene.rogniaux@inra.fr

Biopolymers, Interactions, Assemblies (BIA)