

November 09, 2015 | By Ioana Patringenaru

## **This New Method Identifies Up to Twice as Many Proteins and Peptides in Mass Spectrometry Data**

An international team of researchers developed a method that identifies up to twice as many proteins and peptides in mass spectrometry data than conventional approaches. The method can be applied to a range of fields, including clinical settings and fundamental biology research for cancer and other diseases. The key to the new method's improved performance is its ability to compare data to so-called spectral libraries—essentially a pattern-matching exercise—rather than individual spectra or a database of sequences.

The team describes their results in the Nov. 9 issue of *Nature Methods*. “You can integrate our method with existing pipelines to increase performance by up to three- to four-fold,” said Nuno Bandeira, a professor at the Jacobs School of Engineering and the Skaggs School of Pharmacy at UC San Diego, the study's senior author.

The advance is particularly important because many research teams are now switching to an approach known as data-independent acquisition, which captures a wealth of raw data instead of running data analysis on a few elements at random, (selected from a distribution based on peptide intensities). The amount of data collected has created a bottleneck for existing computational tools. “We needed a new data analysis method,” said Bandeira.

Bandeira's research group focuses on building spectral network algorithms, which analyze pairs of overlapping peptides spectra produced during mass spectrometry experiments. The spectra are produced when an enzyme digests proteins, breaking them down into subcomponents, including peptides. The algorithms detect patterns that the pairs have in common and then look for these patterns in other spectra. This considerably speeds up the process of identifying peptides and by extension, proteins.

Traditional methods compare spectra either against databases or against spectra that have already been identified.

In the *Nature Methods* study, researchers showed that they were able to identify twice as many peptides in human samples when compared to traditional methods. They also observed 40 percent more protein-on-protein interactions. “The results are more stable and easier to reproduce” when compared to traditional methods, Bandeira said.

Next steps include speeding up the process, which takes roughly twice as long as traditional methods. The method also will have to be fine-tuned for the next generation of mass spectrometers.

“We also want to start seeing how we can capitalize on this to analyze a full cohort of data,” Bandeira said.

In addition to Bandeira, authors are Jian Wang of the Center for Computational Mass Spectrometry and the Department of Computer Science and Engineering at UC San Diego; Monika Tucholska, James D.R. Knight, Jean-Phillipe Lambert and Brett Larsen of the Lunefeld-Tanenbaum Research Institute, Sinai Health Systems, in Toronto, Canada; Stephen Tate of SCIEX, Canada, a company that delivers cutting-edge computational solutions to the mass spectrometry community; and Anne-Claude Gingras, of the Lunefeld-Tanenbaum Research Institute and the Department of Molecular Genetics at the University of Toronto.

## **MSPLIT-DIA: sensitive peptide identification for data-independent acquisition**

Jian Wang, Monika Tucholska, James D R Knight, Jean-Philippe Lambert, Stephen Tate, Brett Larsen, Anne-Claude Gingras and Nuno Bandeira

---

### **MEDIA CONTACT**

**Ioana Patringenaru**, 858-822-0899, [ipatrin@ucsd.edu](mailto:ipatrin@ucsd.edu)

UC San Diego’s [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via [studio@ucsd.edu](mailto:studio@ucsd.edu). To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.