

February 21, 2013 | By Doug Ramsey

UC San Diego Computer Scientist Among Young Faculty Recipients of Sloan Research Fellowships

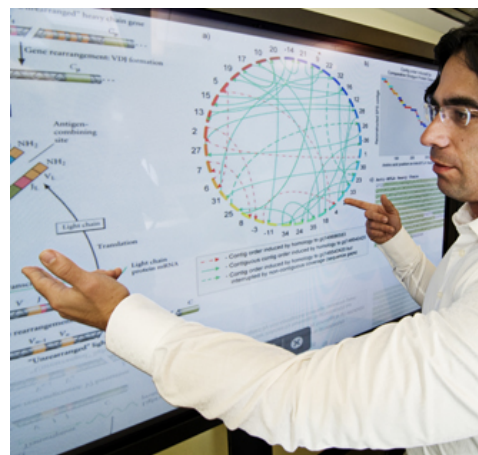
An expert in bioinformatics and computational mass spectrometry at the University of California, San Diego is among the 2013 crop of young faculty members identified by the Alfred P. Sloan Foundation as “rising stars, the next generation of scientific leaders.”

Nuno Bandeira – who leads a repository with the ambitious goal of collecting and organizing all mass spectrometry data in the world – has dual faculty appointments in the Computer Science and Engineering (CSE) department, and the Skaggs School of Pharmacy and Pharmaceutical Sciences.

Bandeira is one of four UC San Diego assistant professors selected to receive Sloan Research Fellowships in 2013 (from among 126 recipients across the U.S. and Canada). The other UC San Diego awardees are Adrian Ioana (Mathematics), Eva-Maria Schötz Collins (Physics and Biology), and Scripps Institution of Oceanography assistant professor Martín Tresguerres.

“Nuno Bandeira joins a distinguished set of faculty in the CSE department who have received this honor, including professors Russell Impagliazzo, Daniele Micciancio, Henrik Jensen, Stefan Savage, Serge Belongie, Alin Deutsch and Alex Snoeren,” said CSE chair Rajesh Gupta in UC San Diego’s Jacobs School of Engineering. “We are proud of Nuno’s achievements as a researcher pushing the frontiers of pharmaceutical sciences through his tremendous knowledge and work at the intersection of biology and computer sciences.”

Bandeira did his undergraduate and Master’s degrees in Computer Science and Applied Artificial Intelligence, respectively, at the New University of Lisbon, in Portugal. He was accepted into CSE’s doctoral program in bioinformatics, where he earned his Ph.D. under advisor Pavel Pevzner. Later, as a postdoctoral researcher in CSE, Bandeira was appointed executive director of a new Center for Computational Mass Spectrometry (CCMS), what is believed to be the only National Institutes of Health-sponsored center in a computer science department in the nation.



Skaggs School and CSE professor Nuno Bandeira has received Sloan Research Fellowship for 2013-2015.

“The university has allowed me to push forward and help develop a novel paradigm that is changing the way that researchers interpret and learn from mass spectrometry,” said Bandeira. “My hope is that the Sloan Research Fellowship will make more scientists aware of the work we are doing, and how researchers around the world can take advantage of the new algorithms, tools and the repository we are building.”



far left) with fellow members of the Center for Computational Mass Spectrometry: CSE professors Pavel Pevzner, [unclear], and Vineet Bafna.

Bandeira has been at the epicenter of a seismic paradigm shift in computational mass spectrometry. Instead of interpreting each spectrum in isolation, he began developing algorithms for so-called ‘spectral networks.’ “We are demonstrating that interpreting *multiple* spectra from related peptides as a consensus is much more powerful than interpreting each spectrum in isolation,” he said. “Together these fundamental changes will probably redefine the range of computational tools and software to interpret mass-spec data over the next three to five years.”

“In the same way that genomics was profoundly transformed by the enabling power of Genbank, BLAST and advanced alignment algorithms,” Bandeira added, “we expect that the spectral networks paradigm could have a similar impact on our ability to understand proteomics biology via high-throughput mass spectrometry.”

The CCMS is developing a repository for all mass-spectrometry data, effectively replacing the Tranche repository at the University of Michigan, which no longer accepts new data uploads or user requests due to an end of major funding. However, rather than simply storing or redistributing the data as the Tranche system did, CCMS’s Mass Spectrometry Interactive Virtual Environment (MassIVE) repository cross-connects all the data.

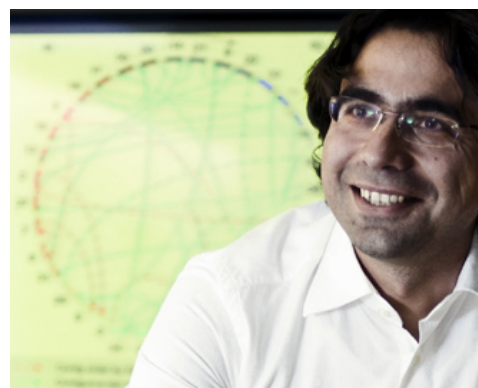
“Every time new mass spectrometry data is generated, people still go back and reinterpret every spectrum as if it’s the first time it has been observed,” explained Bandeira. “This can take months, and as much as two-thirds or three-quarters of the data goes uninterpreted. With the new platform we are developing, if a spectrum is confidently identified once by one lab, then all other researchers can make use of that information.”

Bandeira says that the \$50,000, two-year Sloan Research Fellowship will help him transition proof-of-concept algorithms to “production-grade algorithms to deal with billions of spectra generated worldwide.”

“As a community, we are also building a social network on top of this resource, where we can converge on annotating all of the data,” said Bandeira, who is also affiliated with the California Institute for Telecommunications and Information Technology (Calit2). “Each group focuses on their own area of expertise, but at the time of publication, submission to the repository will allow them to seamlessly share that knowledge with everyone else. By connecting researchers through matching spectra in their datasets, we facilitate a conversation between parties who may have differing interpretations of the same data.”

When identified spectra are committed, the new identifications are automatically propagated to earlier data sets, so the percentage of outstanding data that has not been identified can be reduced over time.

The pioneering efforts of Tranche and now CCMS got a boost from rules at NIH and the leading role of the journal *Molecular and Cellular Proteomics*. They require all authors to deposit their raw data in an independently-managed public access repository at the time of publication. The MassIVE repository allows labs to meet that requirement. “The data that they deposit immediately adds to a growing body of knowledge and any remaining unidentified spectra may continuously be annotated by other people’s efforts,” adds Bandeira. “So six months later, a researcher may receive an email that says 10 percent more data has been annotated. You could go back and use the data for possible new projects.”



Bandeira is one of four UC San Diego assistant professors with Sloan Research Fellowships in 2013.

“Our vision is that every spectrum should only be identified once,” said Bandeira, “and then that knowledge should be seamlessly re-usable by everyone else.”

The MassIVE repository is just that: massive. It is already provisioned to store roughly ten times as much as the 25 Terabytes located in the outgoing Tranche repository. But Bandeira says the computer cluster located in the CSE building at UC San Diego will eventually have the capacity to store 700 TB of data.

CCMS is coming up on its fifth birthday in June 2013. “This first cycle allowed us to develop a proteomics platform for scalable, accessible and flexible – ProteoSAFe – computational mass spectrometry,” said Bandeira. “It allows us to run mass-spec algorithms on thousands of cores at the click of a button, and to allow access via a simple Web interface that is open to anyone. It already exists, and almost 3,000 users have already searched over a billion spectra on this platform.”

Indeed, the CCMS cluster runs practically non-stop. Noted Bandeira: “We have 1,150 cores, and we have clear spikes when new algorithms are released.”

One of the most exciting applications of mass spectrometry, argues Bandeira, is in the discovery of drugs and therapeutic compounds – a primary reason for his faculty appointment in the Skaggs School. “Together with Pieter Dorrestein at UCSD, we published two papers last year in the Proceedings of the National Academy of Sciences, where we showed that spectral networks can be used for antibiotic discovery for natural products and, for the first time ever, that they enable mass-spectrometry analysis of living organisms in microbial colonies – in real time.”

The technology effectively allows a researcher to put two microbial species next to each other and then to “listen in” on the “conversation” as they fight each other with molecules. Changing secretions occur if the microbe is threatened by an invading species. “So we learn what ‘weapons’ they use for fighting each other,” concluded Bandeira. “Those could be interesting precursors to possible antibiotics.”

The Alfred P. Sloan Research Fellowships seek to stimulate fundamental research by early-career scientists and scholars of outstanding promise. The two-year fellowships are awarded yearly to researchers in recognition of distinguished performance and a unique potential to make substantial contributions to their respective fields.

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. 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>.