UC San Diego
MOORES CANCER CENTER
An NCI-designated Comprehensive Cancer Center

American Cancer Society – Institutional Research Grants 2013

"Translational bioinformatics of the NCI 60 proteome"

Nuno Bandeira, PhD

Tandem mass spectrometry is the technology of choice in high throughput proteomics enabling the daily identification and quantification of tens of thousands of peptides and proteins per experiment from hundreds of millions of spectra generated worldwide on a daily basis. But despite significant achievements, the dominant computational paradigm for automated peptide identification still 1) ignores the prior knowledge of many millions of spectra in the public domain and 2) processes every new spectrum in isolation as if it is the first and only spectrum ever acquired by 3) matching against exponentially large search spaces of all possible variants of post-translationally modified peptide sequences. Such intrinsic fundamental limitations result in a significant majority of the spectra (sometimes over 90%) being discarded as unidentified and dramatically restrict biomedical research in high throughput analysis of post-translational modifications and of more complex cancer samples which need to be searched against aberrant genomes/transcriptomes.

In contrast, the proposed comparative mass spectrometry approach builds on the many millions of publicly available NCI60 cancer proteomics spectra to overcome the fundamental limitations of the dominant database search approach. This changes the focus towards a spectrum/library-centric framework where algorithms and statistical models focus primarily on spectrum matching operations and build on these to substantially improve spectrum/sequence matching against aberrant sequence databases.