PROTEIN IDENTIFICATION VIA SPECTRAL NETWORKS ANALYSIS

PAVEL PEVZNER

Department of Computer Science and Engineering
University of California, San Diego

Advances in tandem mass-spectrometry (MS/MS) steadily increase the rate of generation of MS/MS spectra. As a result, the existing approaches that compare spectra against databases are already facing a bottleneck, particularly when interpreting spectra of modified peptides. We introduce a new idea that allows one to perform MS/MS database search without ever comparing a spectrum against a database. We propose to slightly change the experimental protocol to intentionally generate spectral pairs - pairs of spectra obtained from overlapping (often non-tryptic) peptides or from unmodified and modified versions of the same peptide. While seemingly redundant, spectral pairs open up computational avenues that were never explored before. Having a spectrum of a modified peptide paired with a spectrum of an unmodified peptide, allows one to separate the prefix and suffix ladders, to greatly reduce the number of noise peaks, and to generate a small number of peptide reconstructions that are likely to contain the correct one. The MS/MS database search is thus reduced to extremely fast pattern matching (rather than time-consuming matching of spectra against databases). In addition to speed, our approach provides a new paradigm for identifying post-translational modifications and shotgun protein sequencing via spectral networks analysis. This is a joint work with Nuno Bandeira, Karl Clauser, Ari Frank and Dekel Tsur.