Mobilis in MoBiLe: Students in a Dynamic Research Field

Oleg A. Mayboroda, Aswin Verhoeven and Magnus Palmblad
Center for Proteomics and Metabolomics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

One of the most quoted definitions of bioinformatics describes it as a discipline which "encompassesalmostallcomputerapplicationsin biological sciences" (Attwood, 1999), but at the same time points out that the term "was originally coined in the mid-1980s for the analysis of the biological sequence data". Indeed, the analysis of the biological sequence data had dominated the field until analytical instrumentation such as mass spectrometry and NMR has become an everyday reality of a biological/medical laboratory. Today, translation of the instrumental raw data into a "computationally efficient" format, matching tandem mass spectra to peptide sequences (derived from DNA sequences) or 1D and 2D NMR data to metabolic libraries is probably most dynamic part of the field. Here we present the data processing solutions for proteomics and metabolomics implemented into routine analysis over the last years. An essential element of the selected workflows is a contribution of the MoBiLe program students to their development. Presented with small but intellectually challenging tasks, the students have often had to better define the problem as a first step towards solving it. Examples include a format converter for a novel algorithm matching tandem mass spectra to peptides that had to infer enzyme specificity, and a model for species identification that ended up simulating and comparing bottom-up proteomics experiments. This year, a pair of students (or "Dimitries") will combine anatomical and stage ontologies, controlled vocabularies with mass spectrometry, RNA-Seq or NMR data to produce visual, molecular, maps projecting the NMR, RNA-Seq or MS data onto representative anatomical drawings of our model systems. Finally, using an ongoing collaboration with Department of Nephrology as an example, we show how the tools developed by our students can lead to clinically meaningful results.