Florian Breitwieser

Research Group: **Jacques Colinge group**

Department: **CeMM / Bioinformatics Dept.**

Current academic degree: Dipl.-Ing. (FH)

Previous University: Upper Austrian University of Applied Sciences

PhD Thesis: Computational Approaches for Quantifying Proteins and Posttranslational Modifications from Labeled Mass Spectrometry Data

Project title:

Project description:

With modern mass spectrometry equipment, researchers routinely identify thousands of proteins in biological samples. To compare relative changes across samples, stable-isotope coded labels are typically employed. The isobaric labels iTRAQ and TMT are a popular choice to compare up to 10 samples. Differentially labeled samples are pooled and co-analyzed in the mass spectrometer (MS). The MS readout, then, has reporter ions in the peptide fragment spectra, whose intensities reflect the relative abundances of the peptides across the samples.

The analysis of isobarically tagged data is complicated by missing data, variable preciscion of the ratios, and unspecific peptides. My project thus concerned the implementation of a bioinformatical tool and statistical models for better quantification of this type of data. I developed the *isobar* R package, which is now part of Bioconductor, with fundamental classes for proteomics in S4 and protein identification import from Mascot, Phenyx and MzIdentML formats. We developed statistical models to capture the technical variability of protein ratios on the spectra and sample level. As the computations are performed in the R environment, this facilitates data exploration and visualization. User-oriented LATEX and Excel quality-control and analysis reports can easily be generated via scripts.

Furthermore, I implemented extensions to the software for the analysis of changes in post-translational modifications (PTM), specifically: interface to external software for site localization, integration of complementary proteome experiments to derive motification-state changes, and harvesting public PTM databases.