Computational Proteomics and Applied Bioinformatics

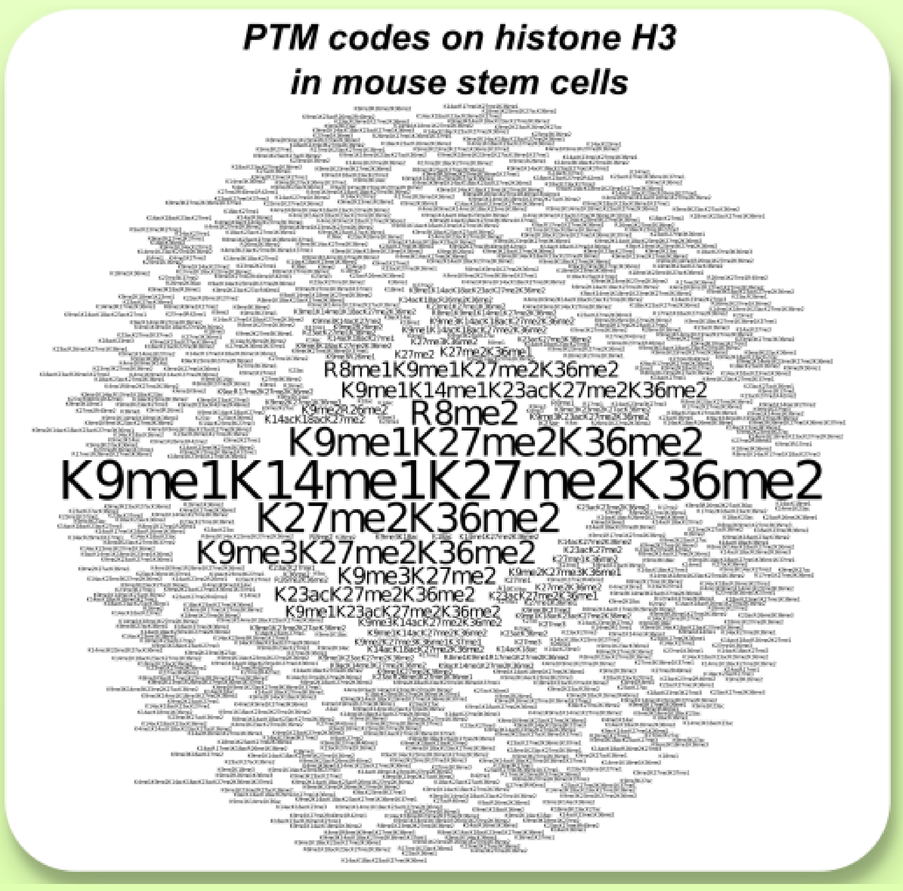
Forskningsleder Veit Schwämmle

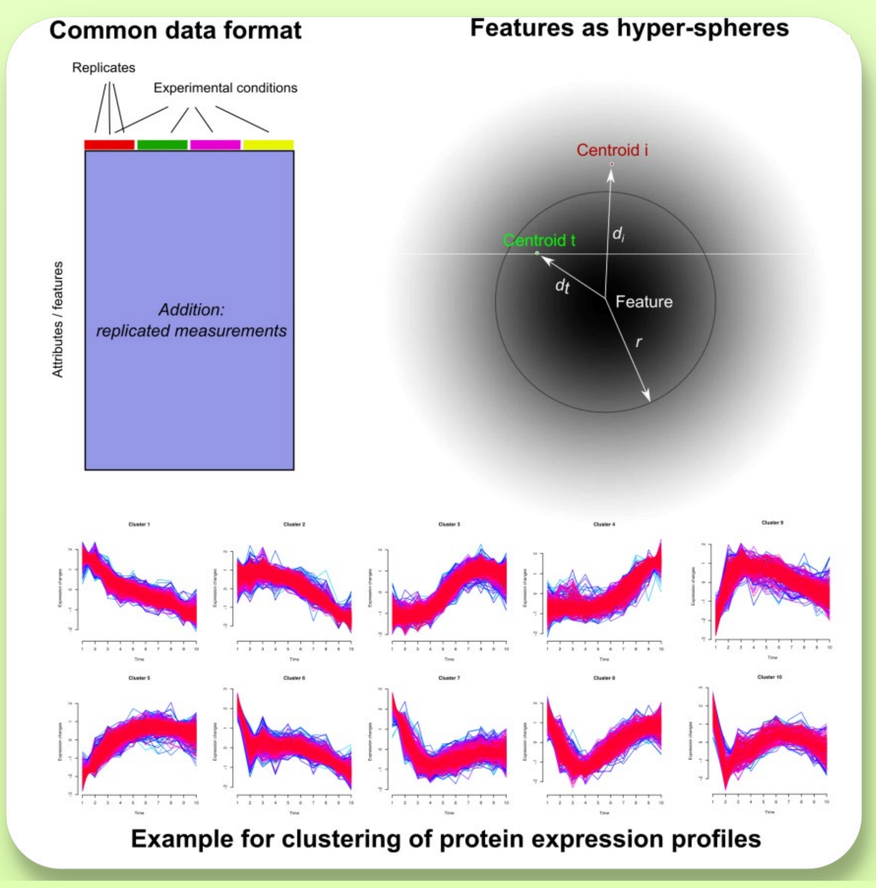
Gruppens kerneforskningsområder

The Computational Proteomics Group develops and applies computational solutions for improved data analysis in large-scale omics experiments with focus on proteins and their post-translational modifications (PTMs).
The aim is to better understand the functional protein states in order to determine, confirm and predict their contribution to cell behavior and disease. This becomes particularly challenging (and interesting) when multiple PTMs on the same protein communicate with each other (so-called PTM crosstalk).

Main research lines:
- *Analysis of data from protein mass spectrometry experiments*: development of software tools, implementation and benchmarking of workflows, PTM quantification
- *Tools for quantification and interpretation of omics data*: data clustering, statistical testing, smart visualization of complex data
- *Re-processing and integration of public data sets*: quantification of protein complexes, phosphorylations, behavior across all human cells
- *Simulation of molecular pathways*: computer models for PTM crosstalk, moderated protein function and biological pathways.

(Experience in programming, for example in R, is advised but not necessary)





Er du interesseret i at skrive projekt i gruppen, så kontakt : veits@bmb.sdu.dk

Projekter	Beskrivelse
Animated visualization of PTMomics data	Histone proteins can be simultaneously decorated by many PTMs. Changes of the PTM landscape in chromatin define different cell types, can induce cancer or lead to dysregulation of gene groups. Our experimental and bioinformatics pipeline allows measuring and quantifying PTM crosstalk to reveal the rules that control writing and deleting these combinatorial marks. The many data dimensions (age, tissue, histone variant, PTM, crosstalk) requires smart visualization techniques to reveal and understand the underlying mechanisms and patterns. Animations provide an elegant way to find co-regulatory patterns and changes that for example propagate through a biological pathway.
Extraction of relevant information from large data collections	Proteomics studies focus on solving particular biological questions by adequate experimental designs. The results are mostly based on a small subgroup of the measured proteins leaving a large number of valuable, yet unharvested data. We aim to extract knowledge by integrating large numbers of proteomics data sets, therefore yielding sufficient statistical power to achieve confident data interpretations. Currently, our main interests are: - High-quality estimation of PTM stoichiometry and crosstalk in the proteome - Assess protein complex composition and regulation by changes of subunits and PTMs
Cluster analysis with missing values	Omics data sets usually consist of thousands of measurements, containing quantitative values for features such as proteins or transcripts. Consideration of multiple cell states or multiple time points allows distinguishing different patterns based on expression profiles. A common problem in the analysis are mising values which can be the result of missed measurements or molecule abundance below the detection limit. These unknowns so far either lead to the removal of relevant data or the imputation of most likely incorrect values. An algorithm for better inclusion of missing values is developed and tested on multiple artificial and real data set from different experimental studies.