Final report

Bio&Data was the first workshop of the newly established MOVISS – "Mountain Village Science Series". It took place in Vorau (Austria) from September 20-23, 2017. It was a small, problem-driven meeting, full of discussions and questions about how to deal with metabolomics data reasonably. A lot of questions from the field of metabolomics and statistics were discussed. The whole meeting was divided into four sessions, every focused on one specific part of the metabolomic experiments. The first part called "From samples to peaks" was targeted on pitfalls connected with the design of experiment. In the second session called "From number to peaks" was discussed peak detection and the other parts of the initial metabolomics analysis including preprocessing and data normalization. The third part "From numbers to pictures" was focused on statistical analysis including univariate and multivariate methods. The topic of Log-Ratio methodology of compositional data was discussed. The last session "From pictures to understanding" was directed to interpretation of metabolomics experiments with a demonstration of real data sets. A summary of all discussions is planned to be produced as a paper for publication to share within the wider metabolomics community in the next year.

Against expectations, the workshop had truly international character since the invited speakers came from EU countries. The program of contributions of invited speakers was as follows:

I. From samples to peaks
   - I1: Oscar Yanes (Universitat Rovira i Virgili & CIBERDEM, Reus, Spain), From spectrometric data to metabolic networks: An integrated quantitative view of cell metabolism
   - I2: Karl Burgess (University of Glasgow, Glasgow, United Kingdom), A communications problem: Context in metabolomics

II. From peaks to numbers
   - I3: Steffen Neuman (IPB Halle, Halle, Germany), From numbers to patterns and back to structures
   - I4: Beata Walczak (University of Silesia, Katowice, Poland), Pre-processing of metabolomic data as a challenging enterprise

III. From numbers to pictures
   - I5: Age Smilde (University of Amsterdam, Amsterdam, The Netherlands), Numerical Representations of Metabolic Systems

IV. From pictures to understanding
   - I6: Nicola Zamboni (ETH Zurich, Zurich, Switzerland), The journey from data to testable hypotheses

All presentations and posters are fully available for participants of the project on google disc: https://drive.google.com/drive/folders/0B31kxt5Nv55jTU1kS0pqS2JYYmM

The aim of this meeting was to bring together undergraduate students, Ph.D. students and early career postdocs from the field of chemistry, biochemistry and statistics with prominent experts in the respective fields and to teach the attendees all necessary skills for metabolomic

data analysis. This purpose has been fulfilled. Students and postdocs were acquainted with the whole process of the metabolomic experiment from its planning, through processing to evaluation. They realized the complexity and the importance of planning all the processes. They also learned new approaches that they can now apply to their own projects. They improved their statistical knowledge which can help in their future scientific career, because these skills are not so common in biochemical sciences, but they become very important. All of the participants engaged extensively in the discussions, so they also practiced communicative skills. In order to strengthen the statistical part of the discussions four PhD students from Technical University Wien: Šárka Brodinová, Irene Hoffmann, Fabian Schroeder and Jan Walach were included in the project.

The meeting was very successful, discussions were very fruitful and all participants are interested in its continuation in the following year. The cooperation of both parts was very rewarding, because both groups are focused on the specific issues connected with chemometrics and metabolomics that can be involved in future cooperation projects. Both Czech and Austrian groups shared the experiences with metabolomic data processing, evaluation and data interpretation. Statistical methods based on univariate and multivariate approaches were also shown with the application on data from clinical metabolomics.

The meeting was concluded as very successful and we plan to repeat this meeting next year in the same place.

List of participants supported by project

Research participants from Czech Republic and their contributions:

- **prof. RNDr. Tomáš Adam, Ph.D.** - He moderated the discussion in the first session connected with the analytical part of the metabolomics project. He also talked about biological interpretation of metabolomic experiments.
- **RNDr. David Friedecký, Ph.D.** - LC/MS metabolomics in diagnosing and treatment of rare diseases (oral presentation).
- **doc. RNDr. Karel Hron, Ph.D.** - He moderated the discussion in the second sessions connected with the data preprocessing and normalization techniques used in metabolomics.
- **Mgr. Alžběta Gardlo, Ph.D.** – Weighted pivot coordinates and their use in metabolomics (poster), Compositional data analysis used in metabolomics (oral presentation).
- All researchers also participated on contributions of students from Czech Republic as supervisors.

Student participants from Czech Republic and their contributions:

- **Mgr. Lucie Mádrová** - Impact of high resolution of mass spectrometry in metabolomics (oral presentation).

- **Ing. Jan Václavík** - LC-HRMS untargeted metabolomics in drug metabolism studies (poster).
- **Mgr. Radana Karlíková** - - LC/MS targeted metabolomics based on HILIC aminopropyl stationary phase (oral presentation).
- **Mgr. Julie Rendlová** - Bayesian counterpart to t-tests in compositional analysis of metabolomic data (poster).
- **Mgr. Nikola Štefelová** - Regression analysis with compositional covariates in the presence of cellwise outliers (poster).
- **Mgr. Štěpán Kouřil** – He joined the meeting as novice with no contribution, because he is in the first year of Ph.D. study. He learned the basic approaches of biostatistics and the whole process of planning the metabolomic experiment.

Research participant from Austria and his contribution:

- **Univ. Prof., Dr. techn., Dipl.-Ing. Peter Filzmoser** - He moderated the discussion in the third sessions connected with the statistical analysis of the metabolomics project. He also participated on contributions of students from Austria as supervisor.

Student participants from Austria and their contributions:

- **Mgr. Jan Walach** - Cell-wise outlier diagnostics and its use for biomarker identification (poster).
- **Mgr. Šárka Brodinová** - Finding groups in large and high-dimensional data using k-means-based algorithm (poster).
- **Dipl.-Ing. Irene Hoffmann** - Robust and sparse estimation methods for high dimensional linear and logistic regression (oral presentation).
- **Dipl.-Ing. Fabian Schroeder** - On an exact nonparametric test for class separability for the purpose of filter-type model selection (oral presentation).

Univ. Prof., Dr. techn., Dipl.-Ing. Peter Filzmoser
Institute of Statistics & Mathematical Methods in Economics, Faculty of Mathematics and Geoinformation, Vienna University of Technology
Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

prof. RNDr. Tomáš Adam, Ph.D.
Pediatric Clinic, Faculty of Medicine
Palacky University Olomouc
I. P. Pavlova 6, 779 00 Olomouc, Czech Republic