Statistical Workflows in PhenoMeNaI

Etienne Thévenot [1]

- Chemical biology
- Beginner
- 0.5 hour

Statistical analysis offers powerful approaches to mine complex and high-dimension datasets, find significant features, and build prediction models with high-prediction performance. Due to the high number of methods available, their complex mathematical background, and the potential pitfalls due to biases and overfitting, a good understanding of each step as well as an environment allowing to efficiently manage the whole workflow, are of major importance.

In this webinar, we will see how to build a statistical workflow within the user-friendly Galaxy environment, including: normalization (signal drift and batch effect), quality control, univariate hypothesis testing, multivariate modelling with (Orthogonal) Partial Least Squares, and feature selection (with Partial Least Squares, Random Forest and Support Vector Machines).

Our example dataset (MTBLS404 [2]) can be downloaded from the MetaboLights [2] repository, the Sacurine study, aims at discovering physiological variations of the human urine metabolome with age, body mass index, and gender.

We will be using the PhenoMeNaI platform [3]. We will also see how data from the MetaboLights repository can be uploaded directly into Galaxy workflows. Additional statistical modules and public analyses are available on the Workflow4Metabolomics platform [4].

This webinar was recorded on 18th April 2018 and was presented by Etienne Thévenot. It is best viewed in full screen mode using Google Chrome. The slides from this webinar can be downloaded below.

See the EMBL-EBI training pages for a list of upcoming webinars [5].

This webinar is aimed at metabolomics researchers and bioinformaticians.

Learning objectives:

- List statistical steps in analysis of metabolomics data in PhenoMeNaI
- Identify statistical parameters and options for analysis of metabolomics data in PhenoMeNaI

Contributors
Etienne Thévenot [1]

CEA, LIST, Laboratory for data analysis and systems; intelligence, MetaboHUB
Computational metabolomics and high-throughput phenotyping

After an education in mathematics and physics at the MINES [6] ParisTech graduate school, Etienne Thévenot [7] obtained his PhD in molecular and cellular neurobiology (CNRS [8]). He has been working at CEA [9] for more than 12 years as a biostatistician and bioinformatician. His research in computational metabolomics [10] focuses on signal processing, analysis, modeling and integration of metabolomics data for biomarker [11] discovery. He is coordinator of the “bioinformatics” work package from the French infrastructure for metabolomics and fluxomics (MetaboHUB [12]).

Your feedback

Please tell us what you thought about this webinar. Your feedback is invaluable and helps us to improve our courses and thus enhance your learning experience.

Source URL: https://www.ebi.ac.uk/training/online/course/statistical-workflows-phenomenal

Links
[1] https://www.ebi.ac.uk/training/online/trainers/Etienne.THEVENOT_17130
[2] https://www.ebi.ac.uk/metabolights/MTBLS404
[10] https://www.ebi.ac.uk/training/online/glossary/metabolomics
[11] https://www.ebi.ac.uk/training/online/glossary/biomarker