ONE-DAY COURSE, Sunday-only

Computational Approaches and Practical Data Analysis for Untargeted Metabolomics

Instructors

Mingxun Wang
University of California, Riverside

Daniel Petras
University of California, Riverside

Prerequisites

- All participants will need to bring a laptop to interactively perform exercises during the course
- Basic knowledge of mass spectrometry and data analysis is recommended
- No extensive experience with untargeted metabolomics computational tools is required

Untargeted metabolomics has grown to be a key data acquisition approach in the annotation of known and discovery of new small molecules from varied fields including clinical metabolomics, natural products discovery, and chemical ecology. However, the richness and complexity of liquid chromatography (LC) mass spectrometry (MS) and tandem mass spectrometry (MS/MS) datasets are a key hurdle in their analysis. This course will introduce the state-of-the-art computational approaches for the analysis of untargeted LC-MS/MS data. Our goal is to provide a basic scaffold of what kinds of biological and/or chemical questions we can ask of the data, and how these computational metabolomics approaches can help answer these questions. The course will have a large practical hands-on component to realize the analysis on real data. This will help researchers have the practical skills to analyze untargeted metabolomics data, understand where the approaches may be limiting, and broadly understand what is possible in the computational analysis of untargeted LC-MS/MS data to help guide future experimental design and data acquisition.

This short course will cover the fundamentals of untargeted LC-MS/MS based metabolomics, theoretical computational concepts for analyzing metabolomics data, and hands-on exercises that bring these theoretical approaches into practical application. *The specific topics covered are as follows:*

- Introduction
  - Untargeted vs Targeted Metabolomics
  - Data Acquisition Strategies in Untargeted Metabolomics
- Data preprocessing and Relative Quantification and Statistical Analysis
  - Feature Finding - MZmine and GNPS2 Dashboard
  - Statistical Analysis (experimental design, multivariate and univariate analysis, multiple hypothesis testing) - Hitchhiker's Guide to Statistical analysis in metabolomics

- Compound Annotation
    - MS/MS Library Matching (Cosine, Spectral Entropy, False Discovery Rate Estimation) - GNPS2 MS/MS Spectral Library Search and ModiFinder
    - In Silico Approaches
        - Database Searching - Sirius, CSI:FingerID
        - MS/MS Spectrum Prediction - CFM-ID, ICEBERG
    - Molecular Networking Approaches
        - MS/MS Alignment - Modified Cosine/Entropy
        - Network Topology - Transitive Alignments
        - Annotation Propagation - ModiFinder
        -

After completing the course, participants will:

- Will understand the limitations of untargeted metabolomics approaches for quantitative and annotation analysis
- Understand the broad computational classes of tools for the analysis of untargeted metabolomics
- Gain familiarity with state of the art tools to develop practical analysis workflow in their own labs
- Gain confidence to provide constructive feedback on the use of computational untargeted metabolomics tools in manuscripts they review