SMART: Statistical Metabolomics Analysis – an R Tool

Hsin-Chou Yang
Institute of Statistical Science, Academia Sinica
hsinchou@stat.sinica.edu.tw

Abstract

Metabolomics data provide unprecedented opportunities to decipher metabolic mechanisms by analyzing hundreds to thousands of metabolites. Data quality concerns and complex batch effects in metabolomics must be appropriately addressed through statistical analysis. This study developed an integrated analysis tool for metabolomics studies to streamline the complete analysis flow from initial data preprocessing to downstream association analysis. We developed Statistical Metabolomics Analysis – An R Tool (SMART), which can analyze input files with different formats, visually represent various types of data features, implement peak alignment and annotation, conduct quality control for samples and peaks, explore batch effects, and perform association analysis. A pharmacometabolomics study of antihypertensive medication was conducted and data were analyzed using SMART. Neuromedin N was identified as a metabolite significantly associated with angiotensin-converting-enzyme inhibitors in our metabolome-wide association analysis ($p = 1.56 \times 10^{-4}$ in an analysis of covariance (ANCOVA) with an adjustment for unknown latent groups and $p = 1.02 \times 10^{-4}$ in an ANCOVA with an adjustment for hidden substructures). This endogenous neuropeptide is highly related to neurotensin and neuromedin U, which are involved in blood pressure regulation and smooth muscle contraction. The SMART software, a user guide, and example data can be downloaded from
http://www.stat.sinica.edu.tw/hsinchou/metabolomics/SMART.htm.

Joint work with Yu-Jen Liang, Yu-Ting Lin, Chia-Wei Chen, Chien-Wei Lin, Kun-Mao Chao and Wen-Harn Pan.