

High-throughput metabolomics - challenges for bioinformatics

R. Wehrens⁽¹⁾

⁽¹⁾Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, TN, Italy

For fruits and vegetables, the metabolome can be seen as the relevant phenotype, directly related to nutritional value and health effects. Worldwide, considerable efforts are being put in characterising the metabolomes of many different plant species. Even though the experimental work is labour-intensive and requires highly specialized people and material, the bottleneck very often lies in the elaboration of the data. The usual methods relying on hyphenated techniques lead to complex data, that even in small experiments need appreciable time for full analysis. Larger experiments, containing hundreds of samples, can no longer be analysed manually and automatic pipelines need to be developed. Several examples of such pipelines will be presented, focusing on LC-MS, GC-MS and HPLC-DAD data. These pipelines basically provide more abstract data summaries that are amenable for interpretation at a biological level. Fundamental differences with the classical manual analysis will be discussed. One should realize that the large amount of data obtained from typical untargeted experiments also has consequences for subsequent statistical analyses; also this aspect will be touched upon.