Maui and Maltcms - A Graphical User Interface and Application Framework for High-Throughput Metabolomics

Nils Hoffmann, Mathias Wilhelm, Kai Stadermann and Jens Stoye
Genome Informatics Group, Faculty of Technology, Bielefeld University, Germany
Nils.Hoffmann@CeBiTec.Uni-Bielefeld.DE

August 15th, 2011

1 Introduction

Recent advances in analytical technology used for gas chromatography-mass spectrometry based metabolomics produce larger and larger amounts of data in need of processing. Furthermore, metabolomics is being used in clinical diagnostics and studies, generating multiple measurements for hundreds or even thousands of individuals over a period of years, which require subsequent comparison and analysis. File sizes are approaching and exceeding gigabytes per measurement and thus require efficient algorithms and data storage techniques for raw data preprocessing and analysis to be of practical use.

2 Methods

We present Maui, the maltcms user interface, a rich client application for interactive visualization and analysis of large datasets from metabolomics experiments. Maui is based on Maltcms, our modular application toolkit for chromatography-mass spectrometry that allows to define and execute typical data preprocessing and analysis workflows in metabolomics. It is focused on processing of data from one-dimensional gas chromatography-mass spectrometry (GC-MS) and comprehensive two-dimensional GCxGC-MS. It provides methods for smoothing, baseline removal, peak detection, peak integration, and peak alignment [1]. Visualizations of aligned and unaligned data, as well as statistical evaluation of peak areas with different normalization schemes between multiple conditions and sample groups are also covered by the framework. Maui supports the generation and execution of user-defined workflows in an intuitive way and provides tools in order to create user databases of reference compounds for targeted or comparative analysis.

Maltcms includes libraries for local and remote parallelization of tasks on a local network of connected computers or within a grid or cloud infrastructure, allowing researchers to scale up their analysis with increasing computing capacities.

It supports the major formats for chromatography mass-spectrometry data as input, like netCDF, mzXML, mzData and mzML. Databases can be imported from msp (AMDIS) compatible text data. The export of user databases to msp is also possible.

3 Results

Maui and Maltcms are implemented using the JAVA programming language allowing them to be extended and adapted easily to custom requirements. Extensions can be developed either in JAVA or in scripting languages like Groovy.

Maltcms is freely available at http://maltcms.sourceforge.net under the L-GPL v3 license. Maui will be officially released early next year and will be available from the same location.

References

[1] Nils Hoffmann and Jens Stoye. ChromA: signal-based retention time alignment for chromatography-mass spectrometry data. *Bioinformatics*, 25(16):2080–2081, August 2009.