Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

After the feature presentation: technologies bridging untargeted metabolomics and biology

Kevin Cho[1,2,3], Nathaniel G Mahieu[1,2,3], Stephen L Johnson[1] and Gary J Patti[1,2,3]

Liquid chromatography/mass spectrometry-based untargeted metabolomics is now an established experimental approach that is being broadly applied by many laboratories worldwide. Interpreting untargeted metabolomic data, however, remains a challenge and limits the translation of results into biologically relevant conclusions. Here we review emerging technologies that can be applied after untargeted profiling to extend biological interpretation of metabolomic data. These technologies include advances in bioinformatic software that enable identification of isotopes and adducts, comprehensive pathway mapping, deconvolution of MS^2 data, and tracking of isotopically labeled compounds. There are also opportunities to gain additional biological insight by complementing the metabolomic analysis of homogenized samples with recently developed technologies for metabolite imaging of intact tissues. To maximize the value of these emerging technologies, a unified workflow is discussed that builds on the traditional untargeted metabolomic pipeline. Particularly when integrated together, the combination of the advances highlighted in this review helps transform lists of masses and fold changes characteristic of untargeted profiling results into structures, absolute concentrations, pathway fluxes, and localization patterns that are typically needed to understand biology.

Addresses
[1] Department of Genetics, Washington University School of Medicine, St. Louis, MO, United States
[2] Department of Chemistry, Washington University, St. Louis, MO, United States
[3] Department of Medicine, Washington University School of Medicine, St. Louis, MO, United States

Corresponding author: Patti, Gary J (gjpattij@wustl.edu)

Current Opinion in Biotechnology 2014, **28**:143–148

This review comes from a themed issue on **Systems biology**

Edited by **Christian M Metallo** and **Victor Sourjik**

For a complete overview see the Issue and the Editorial

Available online 6th May 2014

0958-1669/$ – see front matter, © 2014 Elsevier Ltd. All rights reserved.

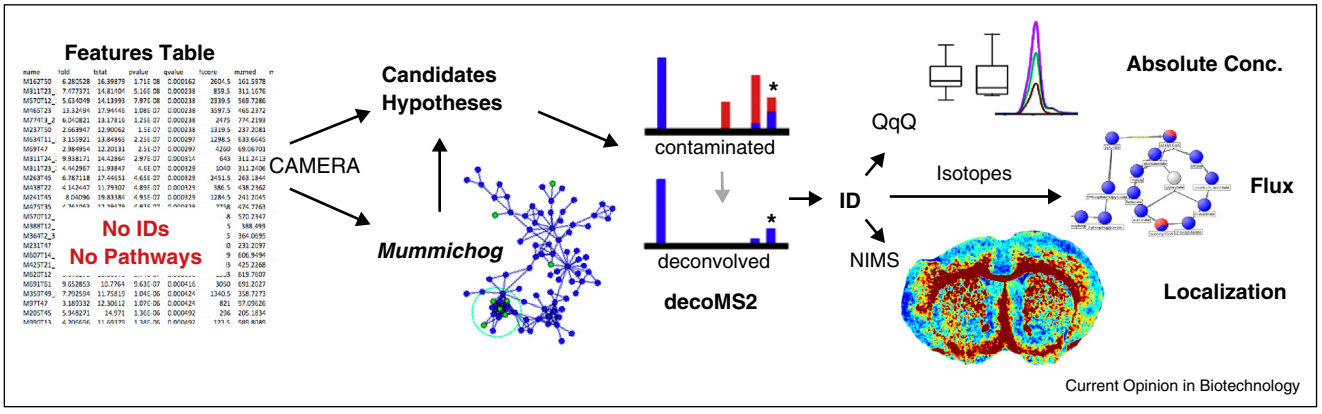
http://dx.doi.org/10.1016/j.copbio.2014.04.006

Introduction

Liquid chromatography/mass spectrometry (LC/MS) provides a robust analytical platform to assay a physiochemically diverse group of small molecules and is therefore widely used to study the metabolome [1]. By using reversed-phase and hydrophilic interaction liquid chromatography together with quadropole time-of-flight or Orbitrap mass spectrometers, thousands of peaks are detected in the metabolic extract of biological samples [2]. Each of these peaks, often referred to as a 'feature', has a unique pair of retention-time and mass-to-charge ratios. Although experimental strategies to optimize metabolome coverage are still being developed, the process of measuring thousands of metabolite features in a biological specimen is now routine and has been discussed in detail [3]. In contrast, the interpretation of these large untargeted metabolomic datasets remains a challenge for many laboratories. This review focuses on emerging technologies that can be applied downstream of untargeted metabolite profiling to drive biological discovery.

Traditionally, untargeted metabolomics is performed by analyzing metabolic extracts derived from two or more sample groups in MS^1 mode. These raw data are then processed with bioinformatic software and a 'features table' containing all detected compounds is produced. The most popular software for processing untargeted metabolomic data is the freely accessible and platform-independent XCMS, but other programs are also available [4–6]. A features table includes mass-to-charge ratios, retention time, statistical comparisons, and relative peak intensities [7••]. Current software, however, does not provide metabolite identifications. Therefore, while the features table can be used to identify potential biomarkers or to broadly compare the similarity of samples, the value of the features table is relatively limited [8,9]. The question that inevitably arises after this initial processing of untargeted metabolomic data is what are the next steps. Most investigators perform targeted MS^2 analysis on peaks of interest with the objective of making structural identifications [10]. Given the time required for metabolite characterization and quantitation by LC/MS, generally only a small number of features are pursued. When comparing samples in which there are many metabolic differences, choosing the most relevant peaks to target for identification is a challenge. Moreover, even once structures are determined, biological interpretation is complicated because untargeted metabolomics only provides a relative comparison of metabolite levels. Additionally, untargeted metabolomics does not provide insight into pathway dynamics or spatial information with respect to tissue, cell type, or organelles. Here we review

Figure 1



Schematic showing the possible integration of emerging, mass spectrometry-based technologies into the untargeted metabolomic pipeline. The workflow starts with a features table that is output from LC/MS-based untargeted metabolomics. Features likely representing naturally occurring isotopes, adducts, and fragments of a single compound can be grouped by using the CAMERA software. After data reduction, mass-to-charge ratios can be searched in databases and interesting candidates chosen as targets for structural identification. Alternatively, the data can be analyzed by *mummichog* to find candidates that map onto related pathways. After acquiring MS^2 data for features of interest, the spectra can be deconvolved by decoMS2 and matched against the MS^2 of authentic standards. These pathway enrichments and identifications can be used to guide further experimentation that involve construction of standard curves to determine absolute concentrations, stable-isotope analysis to calculate metabolic flux, and mass spectrometry-based imaging to localize metabolites within biological tissues.

technologies that can be readily integrated with the untargeted metabolomic workflow to address these issues and facilitate data interpretation (Figure 1).

Post-processing of untargeted metabolomic data

A common strategy applied when prioritizing features to target for structural characterization is data filtering. Often, for example, features that are not changing within defined statistical thresholds are discarded. Additionally, features that have too weak of a signal intensity to obtain high-quality MS^2 data or features that are not of biological origin can be removed [11•,12•]. In experiments where the sample groups being compared differ greatly, there may still be thousands of features that meet the specified criteria. Here, recently introduced software such as CAMERA and *mummichog* may facilitate further feature refinement [7••,13••]. CAMERA is a Bioconductor package that is designed to accept untargeted profiling data initially processed by XCMS. From the list of features detected by XCMS and the raw LC/MS data, CAMERA identifies features that likely correspond to the same metabolite. Given the tendency of metabolites to be detected as multiple features due to naturally occurring isotopes, in-source fragmentation, and adduct formation, CAMERA enables significant data reduction. In some cases, CAMERA reduces the number of features by ~50% [14]. It is important to emphasize that removal of these features from the data improves the likelihood of targeting a molecular ion for structural characterization with a fragmentation pattern that is in a metabolomic MS^2 database. Although here we have described the filtering of a features table by manual inspection of statistical values and signal intensity followed by CAMERA processing, a freely available software package called MetShot was introduced last year to automate these steps [15•].

Even after this filtering, the remaining list of features is often too large for targeted MS^2 analysis. The next post-processing step is to search each feature's mass-to-charge ratio in metabolite databases. Databases with the largest number of mass-to-charge ratios for metabolites include METLIN, the Human Metabolome Database, LIPID MAPS, and the Madison Metabolomics Consortium Database [16–20]. These repositories can be manually searched on an individual basis or searched simultaneously by using a recently developed resource called MetaboSearch [21•]. Database hits provide only putative feature assignments that must be validated by subsequent MS^2 analysis, but these candidate matches can be assessed based on user interest and biochemical relevance. One strategy to assess features based on biochemical relevance is to prioritize features that have candidates belonging to the same metabolic pathway, a process that can be performed computationally by a program called *mummichog*. In brief, the input of *mummichog* is the mass-to-charge ratios of features determined to be unaltered between sample groups as well as those found to be statistically different. The *mummichog* software then determines possible candidate matches, maps all candidate matches onto a metabolic network (derived from KEGG, Recon1, and Edinburgh human metabolic network), and searches unique pathways for enrichment. Features within enriched pathways of interest may then be selectively targeted for MS^2 analysis.

Download English Version:

https://daneshyari.com/en/article/15635

Download Persian Version:

https://daneshyari.com/article/15635

Daneshyari.com