We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,400
Open access books available

117,000
International authors and editors

130M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter 4

Gas Chromatography in Metabolomics Study

Yunping Qiu and Deborah Reed

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/57397

1. Introduction

The metabolome of a cell, tissue, organ, or organism is represented by its small molecular metabolites (molecular weight less than 1000 Da), which are the end products of cellular processes. Variation of the metabolome reflects the interaction of changes in upstream molecules, such as genes and proteins with environmental factors. Metabolic profiling, which is the high-throughput characterization of the metabolome, can be used to assess health status and is a potential diagnostic tool for human diseases. Historically, the sweetness of urine was used to diagnose diabetes. In late 1940s and early 1950s, Roger Williams and colleagues introduced the concept of a "metabolic pattern" to explain the results from their paper chromatography studies that distinguished saliva and urine components among individuals [1]. However, this concept was not widely accepted until the 1960s, when gas chromatography (GC) was successfully used for profiling a complex biological matrix. In the 1960s and 1970s, GC and gas chromatography-mass spectrometry (GC-MS) were successfully used to diagnose metabolic disorders, including maple syrup urine disease [2] and phenylketonuria [3]. In the early 1970s, the term "metabolic profiles" was conceived by Hornings to describe the chromatographic pattern associated with bio-fluid analysis [4].

With the improvement of the sensitivity for the instrument and statistical analysis tools, the field of metabolomics (metabonomics) consequently developed at the end of last century and has been widely used in the past decade. In addition to GC-MS, nuclear magnetic resonance (NMR) and other chromatography coupled mass spectrometry, such as liquid chromatography-mass spectrometry (LC-MS) and capillary electrophoresis–mass spectrometry (CE-MS), are widely used as analytical tools for metabolomics. Compared with other analytical tools, GC-MS is one of the most efficient, sensitive, and reliable tools for metabolomics studies. GC-MS produces reproducible molecular fragmentation patterns making it an integral tool for metabolite identification. Although the application of GC is limited to volatile compounds (before or after derivatization), a large portion of small molecular metabolites are within the

INTECH
open science | open minds

© 2014 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

range of GC separation. For their different metabolic windows among NMR, LC-MS, CE-MS, and GC-MS, the combination of two or more analytical platforms has been used in many metabolomics studies.

In this chapter we introduce GC-MS-based metabolomics. We will begin with the introduction of GC-MS-based metabolomics procedures, and then describe each section of the procedures in more detail. GC-MS-based metabolomics have been utilized in various areas (such as biomedical research, plant research, and microbial research) as well as for different types of samples. We will also provide some examples of using GC-MS-based metabolomics studies in these described fields.

2. Overview of GC-MS-based metabolomics

Prior to GC-MS analysis, effective sample collection and processing methods are necessary to optimize analyte yield and instrument performance. The collected samples can be liquid, solid, or gas, and different sample pretreatment procedures can be used on different sample types. For complex matrices and solid samples such as blood, tissue, and cell pellets, metabolite extraction is necessary to diminish the matrix effect arising from the complex biological materials. Unlike LC-MS and NMR-based metabolomics, compound derivatization is necessary to increase the volatility of metabolites containing polar functional groups, such as carboxylic and amino groups. The volatile compounds (pre- or post-derivatization) are subjected to GC-MS analysis and data collection. To understand the complex mass signals, a data processing procedure should be performed to identify the true signals, assign the signals into different compounds, and align these compounds from different samples. The aligned data can then reveal the differential variables with different sample groups based on statistical analysis. After peak annotation, the metabolic pathway associated with certain physiological or pathological variations can be obtained. Therefore, the entire process is comprised of sample collection, metabolites extraction, compound derivatization, instrument analysis, data analysis, and metabolite annotation and pathway analysis (Figure 1).

3. Sample collection

Typical samples for GC-MS-based metabolomics can be bio-fluids, tissue, or cell samples. For biomedical or clinical use, bio-fluids are widely used due to the non-invasive collection procedures. Bio-fluids have also been considered to be good sources of diagnostic biomarkers. The most widely used bio-fluids are serum/ plasma and urine. Saliva, tears, and sweat are also used in metabolomics studies. Tissue samples can be obtained from animals, human beings, or plants; and cell samples can be from mammals, plants, or microorganisms. Specific sample collection procedures may impact reproducibility, sensitivity and yield of identified metabolites. Here, we will use blood sample collection and treatment as an example to show how the procedure can impact the metabolomics results.

http://dx.doi.org/10.5772/57397

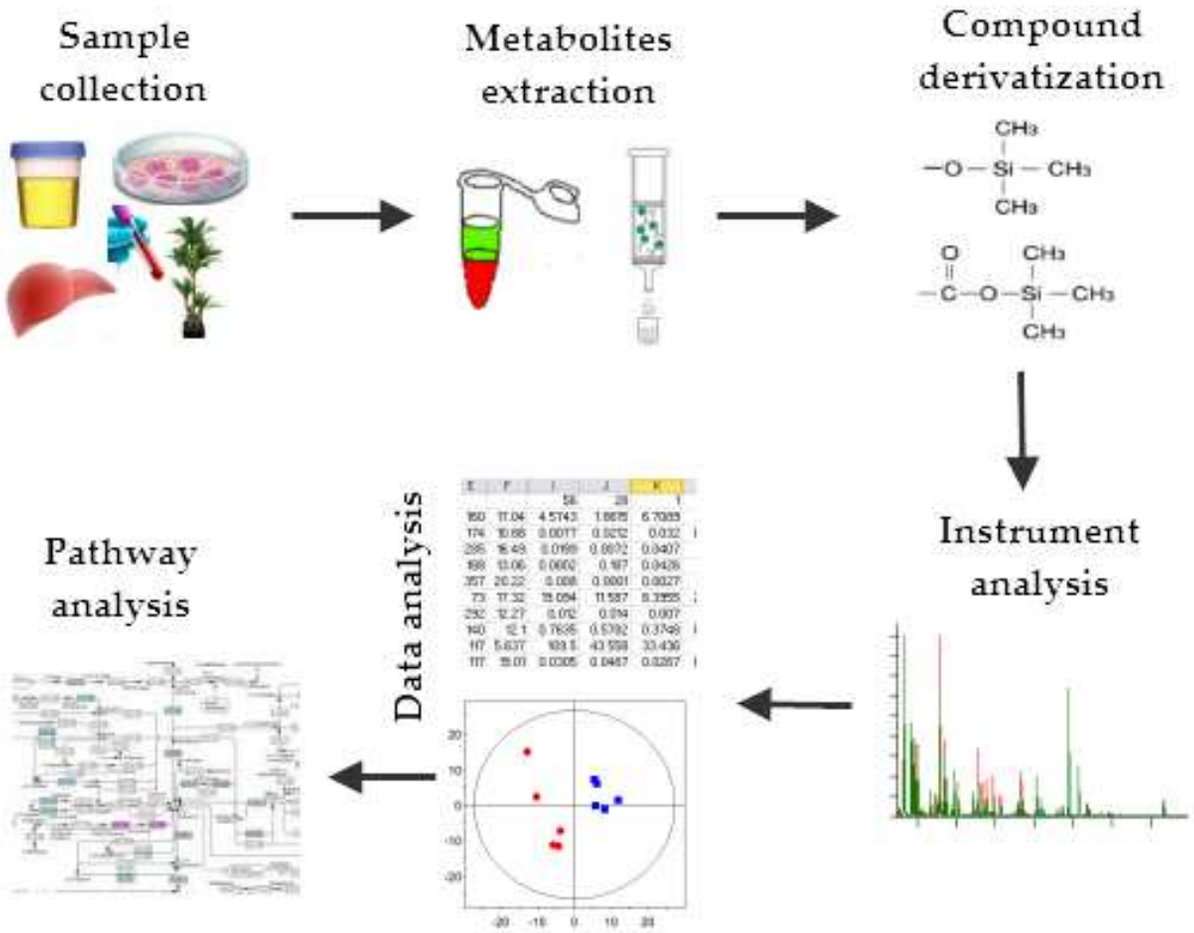

Figure 1. The procedure for GC-MS-based metabolomics.

3.1. Serum or plasma

For blood samples, both serum and plasma can be obtained using different biochemical processes after collection. Serum samples are subjected to a clotting process, which mainly removes fibrinogens, while plasma samples need an anticoagulant during the separation from blood cells. Do these different processes cause variation in metabolites? Is serum or plasma better for metabolomic studies? To answer these questions, previous studies were performed using different analytical platforms. The results showed that there were only slight differences between plasma and serum samples. Using NMR, Teahan et al. showed minor peak shifts between serum and plasma (heparin used as the anticoagulant) samples, such as intensity differences in the resonance dominated by lipid/ triglyceride peaks [5]. In an LC-MS-based untargeted metabolomics study, differences between the serum and plasma (heparin used as the anticoagulant) samples were found to be peptide/protein fragments and lysophophatidylinositol. A lower concentration of lysophophatidylinositol in serum samples was due to degradation in the blood clotting cascade [6]. In an LC-MS/MS-based targeted metabolomics study [7], the authors compared 122 metabolites (focus on lipids and amino acids) between serum and plasma (EDTA used as the anticoagulant) samples. That study found a high correlation between serum and plasma concentrations in all of the targeted metabolites and a

slightly higher concentration of metabolites in serum samples. There was also better reproducibility with the plasma samples, but higher sensitivity with the serum samples.

In addition to NMR and LC-MS, GC-MS-based metabolomics has also been used to evaluate the differences between serum and plasma samples. Dettmer et al. showed slightly higher concentrations in most of the 26 quantified metabolites (19 amino acids, 6 organic acids, and glucose), with the exception of pyruvate, malate, citrate, and glucose in serum samples compared to plasma samples (EDTA used as the anticoagulant) [8]. Only lactate and citrate had variation greater than 15%, which could be attributed to serum and plasma sample preparation. The clotting time for serum preparation may result in a higher level of lactate in the serum due to glycolysis occurring in the blood cells. The conjugation of cations (such as Ca^{2+} and Mg^{2+}) with EDTA (the anticoagulant) lowers the concentration of cations in the plasma, which conjugate with citrate during the metabolite extraction process [8]. Slightly higher metabolite concentrations in the serum samples were observed using different analytical platforms, which may be due to lower protein levels in the serum samples as a result of fibrinogen removal during the clotting procedure. Taken together, data from multiple analytical platforms indicate that serum and plasma samples generate similar metabolomics results.

3.2. Pre-analytical variations in serum and plasma samples

In both serum and plasma samples, metabolite variations can be derived during the sample collection and preparation procedures. Therefore in order to acquire high quality samples for metabolomics studies, a suitable protocol for sample collection is very important. In a NMR-based metabolomics study, Teahan et al. analyzed metabolite variations under different processing conditions, and found slight differences in serum samples processed with different clotting times (up to 180 min) regardless of the temperature used (ice or room temperature) [5]. NMR spectrometry has also detected variations between samples caused by freeze-thaw cycles, and therefore it is important to minimize the number of freeze-thaw cycles during sample handling. Serum samples were found to be stable for 24 hours at 4^0C and 4 hours at room temperature [9].

4. Metabolite extraction

Due to the complexity of metabolites with respect to their chemical structures and properties, molecular weight, and concentrations, it is impossible to extract all metabolites using a single extraction procedure. Different extraction procedures target different metabolites. Although many extraction procedures are currently used, such as liquid-liquid extraction, solid-phase extraction, and supercritical fluid extraction, liquid-liquid extraction is used more universally for metabolomic studies. For polar metabolites, commonly used extraction solvents include water, isopropanol, methanol, and acetonitrile, while for lipophilic metabolites, chloroform and ethyl acetate are commonly used. To increase the extracted metabolites range, combina-

tions of extraction solvents, such as methanol/ chloroform or methanol/ chloroform/ water, are used in many metabolite extraction procedures.

Methanol or methanol/chloroform combination is widely used for GC-MS-based metabolomics studies of blood samples (serum/ plasma). Using the combination of a D-optimal experimental design and partial least squares projections to latent structures analysis, Jiye and colleagues developed a methanol/water-based metabolites extraction procedure from plasma [10]. In this study, the authors compared different solvents and their combinations including water, methanol, ethanol, acetonitrile, acetone, and chloroform. Using methoxyamine and N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) derivatization, the authors resolved 501 peaks in a gas chromatography-Time-of-Flight mass spectrometry (GC-TOF-MS) analysis and identified 80 metabolites from these peaks. Their study showed 100 μL of plasma mixed with 900 μL of methanol and water (8:1, v/v) provided optimal metabolites extraction results. Our laboratory has found that the combination of methanol and chloroform (3:1, v/v) is better for serum metabolite extraction than methanol alone [11]. Alternative ratios of methanol and chloroform have also been used in other metabolomics studies. For example, Nishiumi and colleagues used 250 μl of a solvent mixture (MeOH: H_2O: $CHCl_3$ = 2.5:1:1) for extracting 50 μl of serum [12].

Liquid extraction is necessary for metabolomics studies with solid samples, such as tissue and cell pellets. Tissue samples (from mammalian or plant) stored in liquid nitrogen need to be ground to a fine powder prior to solvent extraction or homogenized in a solvent with blenders. A combination of chloroform/methanol/water has been used extensively in metabolite extraction from tissue. Denkert et al. used a combination of chloroform/methanol/water (2:5:2, v/v/v) for colon tissue extraction [13]. Pan et al. developed a two-step metabolite extraction procedure with a mixed solvent of chloroform/methanol/water (1:2:1, v/v/v) in the first step and methanol in the second step for GC-TOFMS-based metabonomics study of liver tissue [14]. In addition, Gullberg et al. optimized an extraction protocol that yielded 66 metabolites from *Arabidopsis thaliana* samples. The authors found that the optimal extraction mixture for plant tissue was 200 μL of chloroform in a vibration mill followed by 800 μl of a methanol/water mixture (75:25, v/v) to form a monophase [15].

The extraction procedures for collected cell pellets have been widely tested with different solvents and various types of cells. Yeast is one of the most investigated species used for targeted metabolites analysis. One commonly used intracellular procedure for metabolite extraction is based on boiling ethanol, which was firstly described by Entian et al. in 1977 [16]. However, boiling ethanol may be harmful for some metabolites. Villas-Boas et al. tested 6 extraction conditions for yeast, including chloroform/methanol/buffer, boiling ethanol, perchloric acid, potassium hydroxide, methanol/water, and pure methanol. Their results indicated that pure methanol is the most suitable solvent and perchloric acid is the least suitable for metabolite extraction [17]. After testing five extraction protocols, including methanol, methanol/chloroform, perchloric acid, boiling ethanol and potassium hydroxide, Winder et al. also found that cold methanol with repeated freeze/thaw cycles was most effective for extracting intracellular metabolites from *E. coli*. [18]. For mammalian cells, methanol, methanol/water, and methanol/chloroform have been identified as suitable agents for metabolite

extraction. Methanol/chloroform and boiling methanol with tricine buffer have been shown to sufficiently extract metabolites involved in cellular energy metabolism from canine kidney cells [19]. To increase the range of extracted metabolites, Sellick et al. reported a procedure using two extractions in Chinese hamster ovary cells: the first with pure ethanol and the second with water [20]. In addition, Dettmer et al. found that a methanol/water extraction yielded the best results compared to pure methanol, pure acetone, acetone/water, methanol/chloroform/water, methanol/isopropanol/water, or acid–base methanol in a colorectal cancer cell line (SW480). They also found that direct scraping with methanol/water extraction markedly improved the concentration of metabolites compared to trypsin/EDTA detachment [21].

5. Derivatization

GC-MS is designed to analyze volatile compounds that can pass through a gas chromatography column heated to approximately 300°C. Most metabolites have polar functional groups, such as hydroxyl, amino, or carboxyl groups, and therefore are not volatile at the highest temperature allowed by the GC system. Although some of the metabolites may pass through the GC system, their peak shapes and/or recoveries are often unacceptable due to column absorption. Therefore, protection of those polar functional groups from metabolites with chemical derivatization is necessary.

To analyze as many metabolites as possible, wide range silylation is still the most popular derivatization procedure prior to GC-MS metabolomics analysis. Silylation agents can react with nearly all polar functional groups, including –COOH, –OH, –NH, and –SH. In general, the silylation reaction and stability of the derivatized metabolites decrease in the following order: –hydroxy > hydroxyl (phenol) > carboxylic acid > amine > amide; and primary > secondary > tertiary for alcohols and amines. After silylation, metabolites are more volatile and stable to high temperature. The most commonly used silylation agents for metabolomics include N,O-bis-(trimethylsilyl)-trifluoroacetamide (BSTFA), MSTFA, and N-methyl-N-tert-butyldimethylsilyltrifluoroacetamide (MTBSTFA) (Fig. 2). Typically, 1% trimethylchlorosilane (TMCS) is added as a catalyst for the reaction. Fiehn et al. recommended MSTFA instead of BSTFA for metabolomics studies to avoid high volatility and volatile by-products, including trifluoroacetamide, which causes some interference with early eluting peaks [22]. For keto (-oxo) groups, direct derivatization with silylation will result in multiple products via enolization reactions, which can result in complicated chromatograms with quantification and identification problems. Methoximation is always used to protect keto groups prior to silylation in metabolomic studies, although it will generate two peaks with *syn* and *anti* isoforms for the same keto-containing metabolite. For better separation between the *syn* and *anti* isoforms, Fiehn et al. recommended alkoxyamines instead of hydroxylamines [22].

Since silylation agents are sensitive to moisture, it is important to dry the samples thoroughly prior to the derivatization and prevent moisture exposure during the derivatization procedure. The methoximation reaction can be completed in 1-2 hours with heat or overnight at room temperature [10, 15]. Silylation reactions are usually finished in an hour at room temperature or with heat at 60-70 ^{0}C.

An alternative to TMS silylation is chloroformate derivatization, which has also been used in metabolomic studies. Compare to silylation, the reaction with chloroformates can be conducted in aqueous media, which is very convenient for bio-fluids such urine, serum, and saliva. Reaction in the aqueous phase allows derivatization and analysis of certain volatile polar metabolites, such as short chain fatty acids [23], which would evaporate from the sample during the drying procedure prior to TMS derivatization. In addition, the reaction with chloroformates is very quick; under ultra-sonication, the reaction can be finished in seconds [24]. Chloroformate derivatization is widely used in quantitative measurements of amino acids, organic acids, and amines prior to GC-MS analysis [24]. Qiu et al. successfully adapted an ethyl chloroformate derivatization procedure for urine metabolomics analysis, which was validated with 25 standards (amino acids, amines, and organic acids) and urine samples from a precancerous rat model [25]. The authors used a two-step derivatization to increase the metabolite range and intensity, and found the method to be efficient and robust. Tao et al. subsequently optimized an ethyl chloroformate derivatization procedure for a serum metabolomics study and successfully discriminated between uremic patients and normal subjects [26]. However, only 50 metabolites were identified in that analysis, and thus this method has not been widely used in serum metabolomic studies.

Chloroformate derivatization-based metabolomic analysis has also been used for microbial cells. As reported by Smart et al., a methyl chloroformate (MCF) derivatization-based GC-MS metabolomics analysis of microbial cells was able to profile the end products and/or intermediates from specific metabolic pathways, including glycolysis, the Kreb cycle, amino acid metabolism, and fatty acid metabolism [27]. The authors described the advantages of MCF derivatization as being an economical reagent with a fast derivatization procedure (1 min) and less damage of the GC capillary column by the extraction solvent (chloroform).

Silylation and chloroformate derivatization are the two predominantly used procedures for GC-MS-based metabolomic studies. Figure 2 illustrates examples of some of the derivatization agents react with glycine.

6. Data analysis in GC-MS-based metabolomics

Data analysis for metabolomics study includes two key steps. The first step is to acquire a readable dataset from mass spectrometry signals. The other step is to perform statistical analysis, which allows comparisons of the same data generated from different analytical platforms. However, different types of instruments generate different signal patterns. Therefore, in order to obtain better interpretation of the signals, specific computational programs should be designed.

6.1. Signal interpretation for GC-MS-based metabolomics

The field of metabolomics studies the entire metabolome in a given biological sample. Compared to conventional quantitative analytical studies, which target a limited number of compounds, interpretation of mass signals from metabolomics is substantially more complex

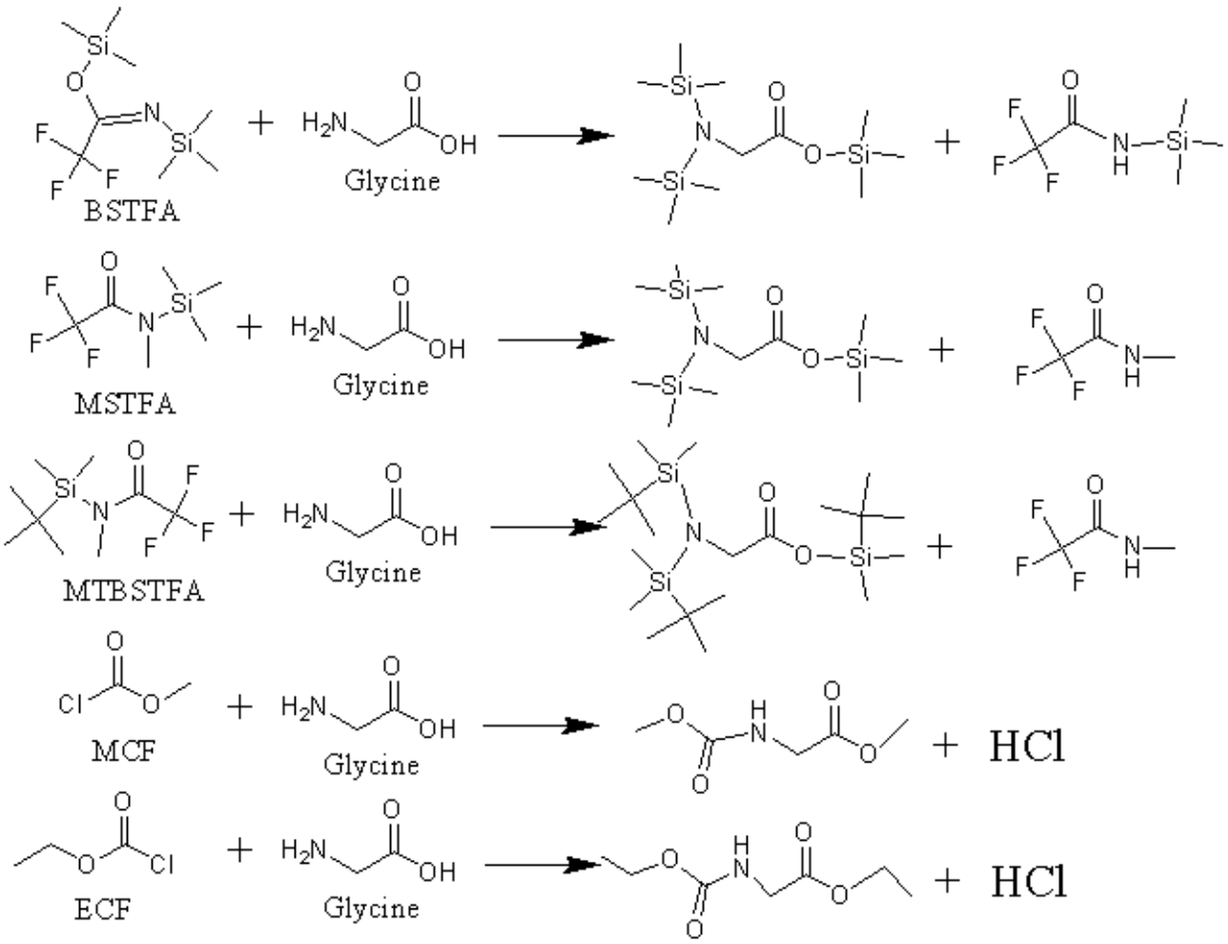

Figure 2. Silylation and chloroformates derivatization with glycine.

and difficult. This complexity lies in two main aspects: deconvolution of coeluting metabolites and alignment of metabolites across multiple samples. An automated mass spectral deconvolution and identification system (AMDIS) developed by Stephen E. Stein is regarded as one of the most powerful software programs for peak deconvolution [28]. This software is designed to deconvolute coeluting compounds from a single sample and does not provide alignment across samples. More recently, several software programs have been designed for GC-MS-based metabolomic data analysis.

Hierarchical Multivariate Curve Resolution (H-MCR) is one of the pioneering software programs that focuses on GC-MS-based metabolomics data [29]. Unlike conventional procedures, this software used a multiprocessing approach to process all interested samples simultaneously. The generated results include the aligned peak information for all samples. However, it will slow down the speed or even exceed computer's capacity when simultaneously processing a large size of samples. An improved procedure was subsequently developed that allows the software to process a subset of representative training samples and generate H-MCR parameters. Using the same parameters, the remaining samples can then be processed very quickly [30].

In 2008, Kopka and colleagues from the Max Planck Institute published a data processing procedure for GC-MS-based metabolomics called TagFinder [31], which is based on the Java programming language. In their process, the raw chromatography files with interchanged format (such as the NetCDF format) or pre-processed peak lists are acceptable. Based on the retention index or retention time, the fragment ions from different chromatograms are binned to "mass tags". After testing of the time groups, the selective mass tags can then be extracted. This process supports both non-targeted fingerprinting and targeting profiling, and the software is freely available for academic use from http://www-en.mpimp-golm.mpg.de/03-research/researchGroups/01-dept1/Root_Metabolism/smp/TagFinder/index.html.

The Automated Data Analysis Pipeline (ADAP; available at http://www.du-lab.org) was developed by Wenxin et al. in 2010. ADAP uses data generated by GC-TOFMS and contains features such as peak detection, deconvolution, peak alignment, and library search [32]. This software is written in standard C++ and R language. To increase the speed of the peak detection and deconvolution steps, parallel computing via Message Passing Interface (MPI) was used. The software was subsequently modified with improved algorithms mainly for peak deconvolution of coeluting metabolites, and substantial improvements in ADAP-GC 2.0 compared to ADAP 1.0 have been reported [33].

In addition to software designed for GC-MS-based metabolomic studies, some software originally developed for LC-MS data has also been used for GC-MS data analysis. XCMS is one of the most popular software programs for MS-based metabolomic studies, and it was originally developed based on LC-MS data [34]. This software was built using an R environment. After data input, a filter and peak identification are performed. The identified peaks are subjected to peak matching across samples. In this step, a group of metabolites are used as standards for a nonlinear retention time correction. Missing values during the peak matching step are filled in during the following step. Statistical analysis and peak visualization are the last two steps. XCMS is freely available under an open-source license at http://metlin.scripps.edu/download/. An online version includes all the same features and is available at https://xcmsonline.scripps.edu [35]. Based on the results from XCMS, a meta-comparison among different groups or complex subgroups can be subsequently performed with metaXCMS (http://metlin.scripps.edu/metaxcms/) [36].

Other LC-MS-based software programs originally designed for LC-MS metabolomics data can also be used for GC-MS-based metabolomic data analysis, including xMSanalyzer [37], MZmine [38], and metAlign [39]. Vendors have also improved their software by adding metabolomic tool kits. For example, ChromaTOF (LECO, St. Joseph, MI) now provides an alignment function with its "Statistical Compare" section.

6.2. Statistical analysis

Metabolomic studies typically contain different groups with multiple samples in each group. The aim of metabolomics is to identify metabolites (variables) that are highly correlated with different conditions, such as gender, stimuli, and pathological status by making comparisons among groups. Due to the complex dataset generated from the chromatography processing, multivariate statistical analysis is typically used in addition to traditional univariate statistics.

Multivariate analyses based on projection methods are popular in metabolomics studies [40, 41]. An unsupervised Principal Component Analysis (PCA) is usually applied to provide a general overview of the acquired dataset and check for outliers. For more specific correlations between the observations and variables, supervised pattern recognitions may be used, such as Partial least-squares (PLS), Orthogonal-PLS (OPLS), PLS-discriminant analysis (PLS-DA), and OPLS-DA. By setting a specific Y matrix, such as gender and pathological status, the variables with high impact on the selected observations can be identified. Those metabolites with high variable importance in the projection (VIP) value can then be selected as differentiating variables. In addition to projection-based methods, other pattern recognition technologies, such as support vector machines (SVM) [42], hierarchical cluster analysis (HCA) [43], and random forest (RF) [44] have also been used in metabolomic data processing.

To further characterize differentiating metabolites, univariate statistics, such as a Student's *t*-test, analysis of variance (ANOVA) test, and non-parametric Mann-Whitney V-tests, are used to compare the selected differential variables from the multivariate statistics [45, 46]. For a large number of distinct variables in the metabolomics data, the false discovery rate (FDR) is used to correct p values from the univariate statistics [47].

7. Metabolite identification and interpretation

Accurate and reproducible compound identification is one of the most important reasons for the popularity of GC-MS in metabolomics studies. However, because of the diversity in chemical structures and concentrations, identification of metabolites remains a challenge. A comparison with commercial libraries, such as the NIST and Wiley libraries, can tentatively identify some of the metabolites. As discussed above, coeluting peaks are a major problem for metabolite identification, and improvements in the separation conditions and advanced deconvolution software (such as ADMIS and ChromaTOF) aid in the accuracy of compound identification. In addition, due to the limited number of metabolite entries in current commercial libraries, even well separated metabolites cannot be identified. Many metabolomic laboratories have developed their own metabolite libraries, which are also used to confirm the identifications based on the retention time (or retention index). FiehnLib is one of the major metabolomic databases and uses a fatty acid methyl ester retention index system. Using GC-MS and GC-TOFMS (LECO, St. Joseph, MI), more than 1,000 primary metabolites, including lipids, amino acids, fatty acids, amines, alcohols, sugars, amino-sugars, sugar alcohols, sugar acids, organic phosphates, hydroxyl acids, aromatics, purines, and sterols have been entered into this library [48].

In addition to biomarker discovery, metabolomics also aims to understand the metabolic variations behind physiological or pathological changes. To obtain the entire picture of metabolic changes, the detected metabolic variations are connected in the metabolic pathways. In a pathway analysis, differential metabolites can be linked with each other as well as to the corresponding genes and enzymes. One of the widely used integrated maps for metabolic pathways, called the 22nd (2003) edition of the IUBMB-Sigma-Nicholson Metabolic Pathways

Chart, is available online (http://www.sigmaaldrich.com/img/assets/4202/MetabolicPathways_6_17_04_.pdf). Other online sources for the metabolic pathway analysis include Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg), BioCyc (http://www.biocyc.org/), Pathway Commons (PC, http://www.pathwaycommons.org/pc/), Small Molecule Pathway Database (SMPDB, http://www.smpdb.ca/), Reactome (http://www.reactome.org/), The Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org), Expert Protein Analysis System (ExPASy, http://web.expasy.org/pathways/), the University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD, http://umbbd.msi.umn.edu/), and others.

8. Applications of GC-MS-based metabolic profiling

Advanced GC-MS-based metabolic profiling techniques, which evolved with the development of GC-MS instrumentation, were originally applied in biomedical research. However, in the past several decades metabolic profiling has been widely used in plant, microorganism, and environment research as well. In this section, we will introduce applications of GC-MS-based metabolomics in these fields.

8.1. GC-MS-based metabolomics used in biomedical research

Variations in the metabolome are associated with both genetic alterations and environmental interactions. Therefore, the field of metabolomics is an effective tool for biomedical research. GC-MS-based metabolomics approaches are extensively used in biomarker discoveries and biological understandings for human diseases, such as cancers, psychosis, and cardiovascular diseases. Samples can be obtained from both patients and animal models.

Colorectal cancer (CRC) is among the most common types of cancer worldwide. GC-MS-based metabolomics approaches have been used with various types of CRC samples, including urine, serum, tissue, and stool. Qiu et al. used GC-MS analysis of urine from CRC patients and a 1,2-dimethylhydrazine (DMH)-induced precancerous colon lesion rat model and their healthy control counterparts [46]. Good separation was observed between CRC patients and healthy controls, pre- and post-operative patients, and between precancerous colon lesion rats and control rats. The authors observed significantly increased tryptophan metabolism as well as disturbed tricarboxylic acid (TCA) cycle and gut microflora metabolism in both the CRC patients and the rat model. In particular, recovery of tryptophan metabolism toward a healthy state was observed in post-surgical samples. A similar study assessed sera metabolomics from CRC patients and found perturbations in glycolysis, arginine and proline metabolism, fatty acid metabolism, and oleamide metabolism [45]. Ong et al. detected enhancement of the purine salvage pathway in CRC tumor tissues compared to adjacent normal epithelia [49]. Higher concentrations of purines and pyrimidines in CRC tissues were also observed by Denkert et al. based on GC-MS analysis [13]. They also observed higher levels of amino acids and lower levels of intermediates of the TCA cycle and lipids in CRC tissues compared to adjacent normal colon mucosa. In another study, GC-MS analysis was used for metabolic profiling of stool

samples from 11 CRC patients and 10 healthy controls, and, higher concentrations of amino acids and lower concentrations of fatty acids and ursodeoxycholic acid were observed in stool samples from CRC patients compared to those from healthy adults[50].

Budczies et al. reported a GC-TOFMS-based metabolic profiling approach for comparing breast cancer tissues to normal ones [51]. That study detected changes in several metabolic pathways, including glycolysis, TCA cycle, nucleotide metabolism, and catabolism of amino acids. They also found that the ratio of cytidine-5-monophosphate to pentadecanoic acid was able to discriminate breast cancer tissues from controls with a sensitivity of 94.8% and a specificity of 93.9%. Using GC-MS analysis, Nam et al. detected four metabolic biomarkers (homovanillate, 4-hydroxyphenylacetate, 5-hydroxyindoleacetate, and urea) in the urine of breast cancer patients, which were validated with gene expression data from the NCBI GEO database [52]. GC-MS-based metabolomics approaches have also been used in several other cancer types, including ovarian carcinoma [53], hepatocellular cancer [54], and osteosarcoma [55].

Several GC-MS-based metabolomics studies on diabetes have also been reported in recent years. To identify metabolic signatures for a relatively new type of diabetes, fulminant type 1 diabetes (FT1DM), Lu et al. profiled the serum metabolome in FT1DM patients together with healthy control subjects, type 2 diabetes, classic type 1 diabetes, and diabetic ketoacidosis patients. Their study found that three metabolite markers (homocysteine, 5-oxoproline, and glutamate) had diagnostic potential for diagnosing FT1DM [56]. GC-MS was also used to evaluate metabolic variations of three drugs, including rosiglitazone, metformin, and repaglinide, in type 2 diabetic patients using multivariate statistical analysis [57]. In another study, two-dimensional GC-TOFMS-based metabolomics coupled with data from LC-MS analysis implicated involvement of the gut microbiota in the development of type 1 diabetes [58].

In addition to cancer and diabetes, GC-MS-based metabolomics approaches have been used in other diseases such as cardiovascular disease and psychosis. To identify metabolite biomarkers for heart failure, Dunn et al. used a GC-MS technique to profile serum metabolites from 52 patients with systolic heart failure and 57 controls [59]. Their observations revealed two promising biomarkers for the failing heart: pseudouridine and 2-oxoglutarate. Yang et al. profiled the serum metabolome in schizophrenia patients and healthy controls using GC-TOFMS and found that fatty acid catabolism was up-regulated with elevated fatty acids and beta-hydroxybutyrate levels in schizophrenia patients [60].

8.2. GC-MS-based metabolomics used in plant research

The plant cell metabolome has not been completely described due to the complexity of primary and secondary metabolites in plants. New metabolites are continuously reported with the development of advanced technologies. GC-MS is widely used in plant metabolomics studies with its powerful separation and compound identification abilities. For example, Fiehn et al. identified 15 relatively new metabolites in *Arabidopsis* extracts by calculating the elemental compositions, and some were found to be completely novel in plant metabolism [22].

One focus in plant metabolomics is the elucidation of the correlation between genotypes and metabolomic phenotypes. The effect of gene modification, identification of key genes and

enzymes involved in existing metabolic pathways, or discovery of a new metabolic pathway are widely investigated with metabolomics technology. One example of metabolomics in genetically-modified plants was published in 2000 [61]. Fiehn et al. analyzed 326 distinct compounds in *Arabidopsis thaliana* leaf extracts with GC-MS technology and successfully distinguished four *Arabidopsis* genotypes (two homozygous ecotypes and a mutant of each ecotype) from each other. In another study, Weckwerth et al. applied GC-TOFMS-based metabolomics to a potato plant line with suppressed sucrose synthase isoform II expression levels and found metabolic variations in carbohydrate and amino acid metabolism compared to its parental background line [62].

A more specific link between metabolic regulation and gene or protein expression levels was achieved with the combination of metabolomics and transcriptomic analysis. Using a GC–TOFMS-based metabolomics technique, Albinsky et al. screened the metabolome of *Arabidopsis* lines that overexpressed rice full-length (FL) cDNAs (rice FOX Arabidopsis lines) [63]. Combined with the transcriptome data, they found that Os-LBD37/ASL39 is a new transcriptional regulator for nitrogen metabolism in rice. Kusano et al. also revealed the importance of cytosolic glutamine synthetase 1;1 (*OsGS1*) in the metabolic balance of rice plant growth [64]. In addition, Sulpice et al. attempted to link enzyme activities and metabolite levels to biomass in *Arabidopsis* accessions [65]. Although the association was weak, it is a good example of a metabolomics study using enzyme analysis.

Plant metabolism is largely influenced by genes and genetic modifications; however, the environment can also impact changes in metabolism. Metabolomics is an effective tool for investigating plant adaptive responses against environmental changes. One good example of a link between environment changes and genetic modification was shown in a study using metabolomics technology in *Arabidopsis*. Urano et al. analyzed metabolic variations of *Arabidopsis* with a knock-out mutant of *NCED3* gene wide type compared to the wide type ones under dehydration stress using the combination of GC-MS and CE-MS [66]. Their results showed the dependence of the accumulation of amino acids to abscisic acid production, and the independence of the level of raffinose to abscisic acid production under dehydration stress. Metabolomics technology has also been used to reveal how plants adapt to defend against insects. Using GC-MS analysis, Kant et al. detected a significant increase in the emission of volatile terpenoids in spider mite-infected tomato plants [67].

8.3. GC-MS-based metabolomics used in microbial research

Microorganisms are an essential component of both environmental and human intestinal ecosystems, and they exhibit marked similarities with human beings in terms of metabolism. Microbial metabolomic approaches are important tools for understanding the interactions between human and environmental changes as well as cellular functions.

Escherichia coli represent a commonly found microorganism in the human gut and thus have received extensive attention in previous metabolomics studies. In one time course response study, *E. coli* were exposed to five different conditions (cold, heat, oxidative stress, lactose diauxie, and stationary phase) [68]. The GC-MS-based metabolomic results showed similar global impact on cell metabolism under different stresses, and metabolic signatures revealed

a consistent decrease in the levels of metabolites related to glycolysis, the pentose phosphate pathway, and the TCA cycle, which were in agreement with the transcriptomic data. In another example, *E. coli* were used to investigate metabolic responses to different biofuel products, including ethanol, butanol, and isobutanol [69]. The authors observed variations in amino acids and osmoprotectants, including isoleucine, valine, glycine, glutamate, and trehalose in *E. coli* cells exposed to biofuel stress. Stress-derived metabolic variations were also investigated with other microorganisms including yeast. Using GC-MS analysis, Aggio et al. investigated the metabolic changes in yeast cells exposed to different frequencies of sonic vibration and silence and found that metabolic perturbations occurred at different sound frequencies [70].

GC-MS-based metabolomics technology can also be used to rapidly detect microorganisms present in the environment or food. For example, Cevallos-Cevallos et al. developed a rapid GC-MS-based procedure that could simultaneously detect 4 microorganisms in ground beef or chicken: *Escherichia coli O157:H7, Salmonella Typhimurium, Salmonella Muenchen,* and *Salmonella Hartford* [71]. In addition, a study by Sue et al. rapidly detected microbial contamination in fermentation processes using a metabolic footprint analysis approach [72].

9. Conclusions

With its powerful separation and identification ability, GC-MS is widely used in metabolomic studies today. The technology is capable of analyzing non-polar metabolites directly and polar metabolites through derivatization. Many GC-MS-based metabolomics protocols have been developed and validated with various applications, and these applications have resulted in new knowledge regarding metabolic networks as well as enhanced systemic understanding of diseases, metabolic regulations, metabolic response to stresses, and cellular functions. Promising diagnostic biomarkers have also been identified in biomedical research studies, and novel microorganism detection procedures have been developed based on microbial metabolomics. Although great achievements have been made with this technology, challenges still remain. For instance, the processing of huge datasets generated from high-throughput metabolomic studies and identification of those unknown peaks are two of the top challenges for GC-MS-based metabolomics. With continuous development and application, GC-MS-based metabolomics will be a powerful tool that benefits to our daily life.

Author details

Yunping Qiu* and Deborah Reed

*Address all correspondence to: qyp29@163.com

Center for Translational Biomedical Research at University of North Carolina at Greensboro, Kannapolis, NC, USA

References

[1] Gates, S.C. and C.C. Sweeley, *Quantitative metabolic profiling based on gas chromatography.* Clin Chem, 1978. 24(10): p. 1663-73.

[2] Melvin Greera, C.M.W., *Diagnosis of branched-chain ketonuria (maple syrup urine disease) by gas chromatography.* Biochemical Medicine, 1967. 1(1): p. 87–91.

[3] Blau, K., H.H. Cameron, and G.K. Summer, *Diagnosis of phenylketonuria by gas chromatography.* Methods Med Res, 1970. 12: p. 100-5.

[4] E. C. Horning, M.G.H., *Human Metabolic Profiles Obtained by GC and GC/MS.* Journal of Chromatographic Science, 1971. 9(3): p. 129-140.

[5] Teahan, O., et al., *Impact of analytical bias in metabonomic studies of human blood serum and plasma.* Anal Chem, 2006. 78(13): p. 4307-18.

[6] Denery, J.R., A.A. Nunes, and T.J. Dickerson, *Characterization of differences between blood sample matrices in untargeted metabolomics.* Anal Chem, 2011. 83(3): p. 1040-7.

[7] Yu, Z., et al., *Differences between human plasma and serum metabolite profiles.* PLoS One, 2011. 6(7): p. e21230.

[8] Dettmer, K., et al., *Comparison of serum versus plasma collection in gas chromatography--mass spectrometry-based metabolomics.* Electrophoresis, 2010. 31(14): p. 2365-73.

[9] Fliniaux, O., et al., *Influence of common preanalytical variations on the metabolic profile of serum samples in biobanks.* J Biomol NMR, 2011. 51(4): p. 457-65.

[10] A, J., et al., *Extraction and GC/MS analysis of the human blood plasma metabolome.* Analytical chemistry, 2005. 77(24): p. 8086-94.

[11] Qiu, Y., *Metabonomics Study on Colorectal Cancer Using Combined Chromatography-mass Spectrometry Strategy* 2008, Shanghai Jiao Tong University: Shanghai. p. 92-95.

[12] Nishiumi, S., et al., *A novel serum metabolomics-based diagnostic approach for colorectal cancer.* PLoS One, 2012. 7(7): p. e40459.

[13] Denkert, C., et al., *Metabolite profiling of human colon carcinoma--deregulation of TCA cycle and amino acid turnover.* Mol Cancer, 2008. 7: p. 72.

[14] Pan, L., et al., *An optimized procedure for metabonomic analysis of rat liver tissue using gas chromatography/time-of-flight mass spectrometry.* J Pharm Biomed Anal, 2010. 52(4): p. 589-96.

[15] Gullberg, J., et al., *Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry.* Anal Biochem, 2004. 331(2): p. 283-95.

[16] Entian, K.D., F.K. Zimmermann, and I. Scheel, *A partial defect in carbon catabolite repression in mutants of Saccharomyces cerevisiae with reduced hexose phosphyorylation.* Mol Gen Genet, 1977. 156(1): p. 99-105.

[17] Villas-Boas, S.G., et al., *Global metabolite analysis of yeast: evaluation of sample preparation methods.* Yeast, 2005. 22(14): p. 1155-69.

[18] Winder, C.L., et al., *Global metabolic profiling of Escherichia coli cultures: an evaluation of methods for quenching and extraction of intracellular metabolites.* Anal Chem, 2008. 80(8): p. 2939-48.

[19] Ritter, J.B., Y. Genzel, and U. Reichl, *Simultaneous extraction of several metabolites of energy metabolism and related substances in mammalian cells: optimization using experimental design.* Anal Biochem, 2008. 373(2): p. 349-69.

[20] Christopher A. Sellick, D.K., Alexandra S. Croxford, Arfa R. Maqsood, Gill M. Stephens, Royston Goodacre, Alan J. Dickson, *Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets.* Metabolomics, 2010. 6: p. 427–438.

[21] Dettmer, K., et al., *Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols.* Anal Bioanal Chem, 2011. 399(3): p. 1127-39.

[22] Fiehn, O., et al., *Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry.* Anal Chem, 2000. 72(15): p. 3573-80.

[23] Zheng, X., et al., *A targeted metabolomic protocol for short-chain fatty acids and branched-chain amino acids.* Metabolomics, 2013. 9(4): p. 818-827.

[24] Husek, P., *Chloroformates in gas chromatography as general purpose derivatizing agents.* J Chromatogr B Biomed Sci Appl, 1998. 717(1-2): p. 57-91.

[25] Qiu, Y., et al., *Application of ethyl chloroformate derivatization for gas chromatography-mass spectrometry based metabonomic profiling.* Anal Chim Acta, 2007. 583(2): p. 277-83.

[26] Tao, X., et al., *GC-MS with ethyl chloroformate derivatization for comprehensive analysis of metabolites in serum and its application to human uremia.* Anal Bioanal Chem, 2008. 391(8): p. 2881-9.

[27] Smart, K.F., et al., *Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography-mass spectrometry.* Nature protocols, 2010. 5(10): p. 1709-29.

[28] Stein, S., *An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data* J. Am. Soc. Mass Spectrom, 1999. 10(8): p. 770-781.

[29] Jonsson, P., et al., *High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses.* Anal Chem, 2005. 77(17): p. 5635-42.

[30] Jonsson, P., et al., *Predictive metabolite profiling applying hierarchical multivariate curve resolution to GC-MS data--a potential tool for multi-parametric diagnosis.* J Proteome Res, 2006. 5(6): p. 1407-14.

[31] Luedemann, A., et al., *TagFinder for the quantitative analysis of gas chromatography--mass spectrometry (GC-MS)-based metabolite profiling experiments.* Bioinformatics, 2008. 24(5): p. 732-7.

[32] Jiang, W., et al., *An automated data analysis pipeline for GC-TOF-MS metabonomics studies.* Journal of proteome research, 2010. 9(11): p. 5974-81.

[33] Ni, Y., et al., *ADAP-GC 2.0: deconvolution of coeluting metabolites from GC/TOF-MS data for metabolomics studies.* Anal Chem, 2012. 84(15): p. 6619-29.

[34] Smith, C.A., et al., *XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification.* Anal Chem, 2006. 78(3): p. 779-87.

[35] Tautenhahn, R., et al., *XCMS Online: a web-based platform to process untargeted metabolomic data.* Anal Chem, 2012. 84(11): p. 5035-9.

[36] Tautenhahn, R., et al., *metaXCMS: second-order analysis of untargeted metabolomics data.* Anal Chem, 2011. 83(3): p. 696-700.

[37] Uppal, K., et al., *xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data.* BMC bioinformatics, 2013. 14: p. 15.

[38] Katajamaa, M., J. Miettinen, and M. Oresic, *MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data.* Bioinformatics, 2006. 22(5): p. 634-6.

[39] Lommen, A., *MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing.* Anal Chem, 2009. 81(8): p. 3079-86.

[40] Holmes, E. and H. Antti, *Chemometric contributions to the evolution of metabonomics: mathematical solutions to characterising and interpreting complex biological NMR spectra.* Analyst, 2002. 127(12): p. 1549-57.

[41] Trygg, J., E. Holmes, and T. Lundstedt, *Chemometrics in metabonomics.* J Proteome Res, 2007. 6(2): p. 469-79.

[42] Mahadevan, S., et al., *Analysis of metabolomic data using support vector machines.* Anal Chem, 2008. 80(19): p. 7562-70.

[43] Tikunov, Y., et al., *A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles.* Plant Physiol, 2005. 139(3): p. 1125-37.

[44] Chen, T., et al., *Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection.* Evid Based Complement Alternat Med, 2013. 2013: p. 298183.

[45] Qiu, Y., et al., *Serum metabolite profiling of human colorectal cancer using GC-TOFMS and UPLC-QTOFMS.* J Proteome Res, 2009. 8(10): p. 4844-50.

[46] Qiu, Y., et al., *Urinary metabonomic study on colorectal cancer.* J Proteome Res, 2010. 9(3): p. 1627-34.

[47] Boudonck, K.J., et al., *Discovery of metabolomics biomarkers for early detection of nephrotoxicity.* Toxicol Pathol, 2009. 37(3): p. 280-92.

[48] Kind, T., et al., *FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry.* Anal Chem, 2009. 81(24): p. 10038-48.

[49] Ong, E.S., et al., *Metabolic profiling in colorectal cancer reveals signature metabolic shifts during tumorigenesis.* Mol Cell Proteomics, 2010.

[50] Weir, T.L., et al., *Stool Microbiome and Metabolome Differences between Colorectal Cancer Patients and Healthy Adults.* PLoS One, 2013. 8(8): p. e70803.

[51] Budczies, J., et al., *Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue - a GC-TOFMS based metabolomics study.* BMC Genomics, 2012. 13: p. 334.

[52] Nam, H., et al., *Combining tissue transcriptomics and urine metabolomics for breast cancer biomarker identification.* Bioinformatics, 2009. 25(23): p. 3151-7.

[53] Denkert, C., et al., *Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors.* Cancer Res, 2006. 66(22): p. 10795-804.

[54] Chen, T., et al., *Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma.* Mol Cell Proteomics, 2011. 10(7): p. M110 004945.

[55] Zhang, Z., et al., *Serum and urinary metabonomic study of human osteosarcoma.* J Proteome Res, 2010. 9(9): p. 4861-8.

[56] Lu, J., et al., *Serum metabolic signatures of fulminant type 1 diabetes.* J Proteome Res, 2012. 11(9): p. 4705-11.

[57] Bao, Y., et al., *Metabonomic variations in the drug-treated type 2 diabetes mellitus patients and healthy volunteers.* J Proteome Res, 2009. 8(4): p. 1623-30.

[58] Oresic, M., et al., *Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes.* J Exp Med, 2008. 205(13): p. 2975-84.

[59] Warwick B. Dunn, D.I.B., Sasalu M. Deepak, Mamta H. Buch, Garry McDowell, Irena Spasic, David I. Ellis, Nicholas Brooks, Douglas B. Kell, Ludwig Neyses, *Serum metab-*

olomics reveals many novel metabolic markers of heart failure, including pseudouridine and 2-oxoglutarate. Metabolomics, 2007. 3(4): p. 413–426.

[60] 60.Yang, J., et al., *Potential metabolite markers of schizophrenia.* Mol Psychiatry, 2013. 18(1): p. 67-78.

[61] Fiehn, O., et al., *Metabolite profiling for plant functional genomics.* Nat Biotechnol, 2000. 18(11): p. 1157-61.

[62] Weckwerth, W., et al., *Differential metabolic networks unravel the effects of silent plant phenotypes.* Proc Natl Acad Sci U S A, 2004. 101(20): p. 7809-14.

[63] Albinsky, D., et al., *Metabolomic screening applied to rice FOX Arabidopsis lines leads to the identification of a gene-changing nitrogen metabolism.* Mol Plant, 2010. 3(1): p. 125-42.

[64] Kusano, M., et al., *Metabolomics data reveal a crucial role of cytosolic glutamine synthetase 1;1 in coordinating metabolic balance in rice.* Plant J, 2011. 66(3): p. 456-66.

[65] Sulpice, R., et al., *Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions.* Plant Cell, 2010. 22(8): p. 2872-93.

[66] Urano, K., et al., *Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics.* Plant J, 2009. 57(6): p. 1065-78.

[67] Kant, M.R., et al., *Differential timing of spider mite-induced direct and indirect defenses in tomato plants.* Plant Physiol, 2004. 135(1): p. 483-95.

[68] Jozefczuk, S., et al., *Metabolomic and transcriptomic stress response of Escherichia coli.* Mol Syst Biol, 2010. 6: p. 364.

[69] Wang, J., et al., *Global Metabolomic and Network analysis of Escherichia coli Responses to Exogenous Biofuels.* J Proteome Res, 2013.

[70] Raphael Bastos Mereschi Aggio, V.O., Silas Granato Villas-Boas, *Sonic vibration affects the metabolism of yeast cells growing in liquid culture: a metabolomic study.* Metabolomics, 2012. 8(4): p. 670-678.

[71] Cevallos-Cevallos, J.M., M.D. Danyluk, and J.I. Reyes-De-Corcuera, *GC-MS based metabolomics for rapid simultaneous detection of Escherichia coli O157:H7, Salmonella Typhimurium, Salmonella Muenchen, and Salmonella Hartford in ground beef and chicken.* J Food Sci, 2011. 76(4): p. M238-46.

[72] Sue, T., et al., *An exometabolomics approach to monitoring microbial contamination in microalgal fermentation processes by using metabolic footprint analysis.* Appl Environ Microbiol, 2011. 77(21): p. 7605-10.

IntechOpen

IntechOpen