Chapter 4

GC-MS-Based Metabolic Phenotyping

Olga Deda*, Helen Gika†, Nikolaos Raikos† and Georgios Theodoridis*
*Laboratory of Analytical Chemistry, School of Chemistry, Aristotle University of Thessaloniki, Thessaloniki, Greece, †Laboratory of Forensic Medicine and Toxicology, Medical School, Aristotle University of Thessaloniki, Thessaloniki, Greece

Chapter Outline

1 INTRODUCTION

Metabolic phenotyping focuses on the study of low-molecular-weight primary metabolites such as amino acids, organic acids, alcohols, carbohydrates, vitamins, nucleotides, as well as secondary metabolites such as phenolic compounds, alkaloids antioxidants, and others [1], usually in complex matrices. At its core, it aims at the reliable detection and quantification of small molecules derived from biochemical reactions in cells [2]. The approach has already been proven as a powerful tool in the service of biochemistry. Metabolic phenotyping as "Biochemistry's new look" can shed light on the small molecule metabolites that flood the human body, ranging (in numbers) from 3000 to 100,000 different molecules [3]. Metabolic phenotyping is now considered to be one of the three main omics subfields of systems biology [4].

By developing and applying robust, sensitive, modern analytical methods, metabolic phenotyping is capable of providing a reliable "snapshot" of the concentrations of a great number of metabolites in biological samples. Simultaneous identification of a variety of metabolites can answer the question of

The Handbook of Metabolic Phenotyping. https://doi.org/10.1016/B978-0-12-812293-8.00004-9

what the metabolic response of the studied organisms is to any stimulus (stress, diet, physical exercise, disease, medication, or other).

Three approaches are usually encountered in metabolism-based studies. Metabolic phenotyping refers to both absolute and relative quantitative and qualitative analysis of as many metabolites as possible, which belong in predefined common metabolic or structural groups (carbohydrates, amino acids, organic acids, and several other groups) [5].

A more accurate quantification is accomplished with targeted methods. Strictly defined metabolites involved in the biochemical pathway(s) of interest are extracted with selected sample preparation protocols and are analyzed using optimal parameters in order to obtain reliable quantification.

The last and more general approach includes the version of metabolic phenotyping that does not target specific metabolites but provides a general map of the samples' content, to quickly assess whether or not any biological differentiation of the samples is present [6]. This is a key point, which makes metabolic phenotyping stand out over the other "omics" technologies and to be considered as a core technique for the diagnosis of diseases [7].

Although human metabolites are estimated to be present in lower numbers than proteins, mRNAs, and genes [7], these are still too many to be covered by one analytical technique. In addition, both the primary and the secondary metabolites extracted from the various analyzed biological samples are characterized by a great molecular diversity and heterogeneity, while they are present in a great range of concentration levels.

Although efforts are made to detect, quantify, annotate/identify more metabolites with as few analytical platforms as possible [2], in fact, this is very challenging and to this day rather unachievable.

It is often imperative to utilize more than one analytical technique, in order to fulfill the analysis requirements of a large scale of various structural and chemical diversities [1]. When looking for the effect of an intervention on a patients' group, changes in the metabolites triggered by the intervention may not be directly revealed by the applied technique, thus leading the researchers to unfitting and wrong results. If the same samples are analyzed using an additional technique, a clear and unambiguous change may be revealed, due to different metabolome coverage provided by the complementary technique.

A good compromise between compound specificity and the number of metabolites detected, is provided by mass spectrometry (MS), usually coupled to gas (GC) or liquid (LC) chromatography [2].

2 GC-MS AS AN ANALYTICAL TECHNIQUE

The definition of MS (mass spectrometry), provided by the International Union of Pure and Applied Chemistry (IUPAC), is "the study of systems by a process of forming gaseous ions, with or without fragmentation, which are then characterized by their mass/charge ratios (m/z) and their relative abundances" [8].

Typically in coupling of gas chromatography to mass spectrometry (GC-MS), the process in the mass spectrometer begins with the formation of compounds' gaseous ions via electron ionization. Chemical reactions in the gas phase determine the production of ionic and neutral species. The separation of the multitude of ions, generated by each sample, is based on their specific mass-to-charge ratio (m/z). When tandem MS (MS/MS) is applied, ions are submitted to fragmentation, and the resulting product ions are separated according to their mass-to-charge ratio [9]. Ions move under the influence of electric and/or magnetic fields, in high vacuum conditions, in order to complete their journey through a free pathway without collisions. Apart from the ion separation achieved under the influence of electro/magnetic fields, separation can also be achieved in the absence of field, as in a time-of-flight (TOF) analyzer [10]. The analytical process ends with the detection of ions derived from the analyzer and the measurement of their relative abundance.

MS analysis concludes to a plot, called a mass spectrum, that illustrates the intensity versus mass-to-charge ratio, representing the distribution of ions in the analyzed sample [11], providing information about the nature and the structure of their precursors.

A typical configuration of a mass spectrometer is a sample introduction system (injector), an ion source providing gaseous ions, the mass analyzer and the detector, which can provide quantitative results of specific ions in a default mass window or time period. The required software allows the analyst to acquire, collect, and handle the obtained data. This system is loaded with spectral libraries to help identify the analyzed molecules [12].

The coupling of the MS with the GC (gas chromatography) is one of the oldest and most widespread, tested, and successful techniques used in metabolite analysis. The strong advantages of capillary GC, i.e., the high separation resolution and the relatively few problems of linkage to the MS interface, have greatly enhanced its use in metabolic phenotyping [13]. A schematic diagram of a typical GC-MS system is illustrated in Fig. 1A.

A sample subjected to GC-MS analysis becomes gaseous in the heated injector of the GC and its compounds are separated as they migrate in the heated column. The different volatility and polarity of the analyzed compounds are the key parameters that drive analyte separation; due to their diverse properties, analytes exhibit dissimilar interactions with the stationary phase. Those that show increased interaction migrate slowly and are eluted with long retention times. Analytes will small interactions migrate faster and elute at early retention times. The analyzed compounds are distributed between the mobile phase (carrier gas, an inert gas, in most cases helium) and the stationary phase (column) [8].

Biological specimens, which are considered complex matrices, are usually separated by applying a programmed temperature gradient and the more apolar, volatile compounds are eluted first. The last part of the GC column links through a heated transfer line to the entrance of the MS ion source. Of the

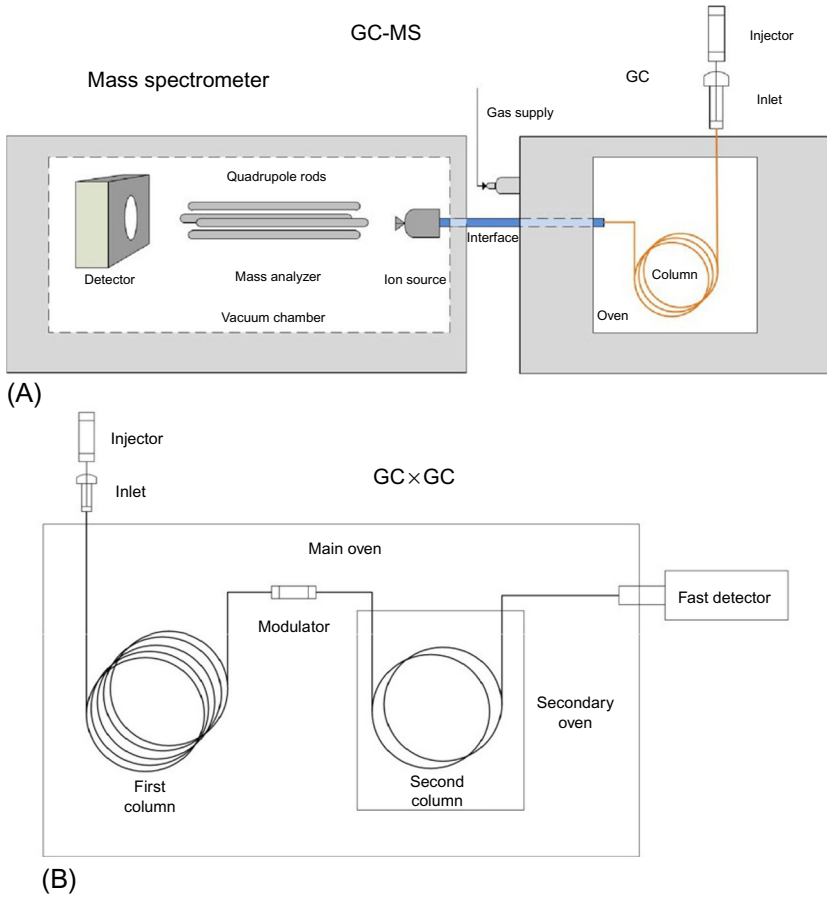

FIG. 1 (A) A schematic diagram of a typical GC-MS system. (B) A schematic diagram of a typical GC × CC system.

two ionization methods usually applied in GC-MS, namely, electron impact ionization (EI) and chemical ionization (CI) [13], electron ionization by far dominates the field.

GC-MS is considered to be the "gold standard" in metabolic phenotyping analysis [1, 14] and is still extensively used in studies due to its undisputed advantages, which include high chromatographic separation, reliable and efficient analyte identification, and precise quantification of metabolites [15]. Moreover, compared to other analytical techniques, GC-MS requires only a relatively low cost for instrument acquisition, analysis, and system maintenance [16].

Another reason that bolsters the use of this technique is the existence of validated methods and protocols for the analysis of a plethora of specimens

[17, 18]. Detailed instructions for instrument maintenance, as discussed in Section 8, data processing (Section 5) as well as the interpretation of the findings (Section 7) are provided through the literature, greatly assisting the analyst [19]. Additionally, both public and commercially available spectral libraries provided for the identification of metabolites, facilitate data handling are probably a key reason why GC-MS remains an indispensable tool for small molecule analysis and thus metabolic phenotyping applications [2, 20, 21].

The most important aspect of the specific technique is that the fragmentation patterns of the analyzed metabolites are highly reproducible and independent of the instrument used. In a more conceptual meaning, this attribute is rather similar to NMR spectroscopy, in contrast to liquid chromatography-mass spectrometry (LC-MS), where the analysis is very dependent on analytical parameters such as the column, the mobile phases, the gradient eluent system, etc. As a result, the robustness of GC-MS made the existence of spectral libraries a reality [22]. Chapters 2 and 3 of this volume cover the principles, theory, and practice of NMR spectroscopy and LC-MS, respectively.

Over the last decade, developments in comprehensive two-dimensional chromatography represent a central element of analytical separations development. In this aspect, two-dimensional GC methods (GC × GC) have been developed, often combined with TOF-MS, thus offering a very powerful analytical tool for very complex mixtures, with unprecedented capabilities and perspectives [4, 23]. A typical configuration of GC × GC system is presenting in Fig. 1B.

In GC × GC analysis, compounds that elute from the first column are stationed for a very short time in an appropriate modulator. These stationed compounds are subsequently introduced to a second column for additional separation. The first column is usually nonselective, whereas the second column, which is typically much shorter, allows for compound separation via an orthogonal mechanism (e.g., if the first separation is based on the boiling point the second column allows separation due to compound polarity) [4]. The first column is usually of a typical GC column length (e.g., 30 m), but the second column is much shorter to allow rapid elution of the stationed analytes before the elution of a subsequent peak in the first column (1st dimension of separation). As a result in GC × GC experiment, the total analysis time may be shorter due to the increased chromatographic resolution. Peak capacities are increased allowing larger numbers of compounds to be detected in a single injection. Unfortunately, the instrument cost is higher and the robustness of this approach still lags behind one-dimensional applications. Furthermore data treatment/mining is more demanding compared to GC-MS, as the read out of such systems is not a two-dimensional plot but more complicated three-dimensional or four-dimensional matrices. A comparative evaluation of software for the deconvolution of metabolic phenotyping GC-TOF-MS data can be found in [24].

3 GC-MS METABOLIC PHENOTYPING WORKFLOW

In metabolic phenotyping studies, experimental design represents an important and essential matter. Providing detailed descriptions and rules of experimental design goes beyond the scope of this chapter, but here we will only highlight the need for careful experimental planning in relation to the kinetics of derivatization and the stability of the formed derivatives (see Section 4). Such issues may prove very important for the validity of the findings and should be carefully considered at the beginning of the study, before planning for the length and the number of analytical batches needed to conclude the analysis of real test samples. Pilot studies should be performed first to check the capabilities of the analytical system (chromatographic resolution, "cleanliness" of the chromatograms) and personnel, the stability of the system and the total analytical turnover.

The next step is the collection of the samples and their storage. Usually raw samples are immediately placed at −80°C, after collection. Special treatments depending on the kind of sample are required. For example, different procedures are followed for whole blood [25] and serum or plasma samples [17]. Sodium azide could be added as preservative [26], in the case of fecal and urine sample collection. Snap freezing using dry ice or liquid nitrogen is preferable for tissues in order to directly inhibit ongoing biochemical reactions. Good practices even at the early stage of sample collection and storage could result in obtaining a reliable and accurate snapshot of the metabolism, by avoiding sample contamination and deterioration [27].

The next crucial step, which may determine the quality of the obtained results, is sample preparation. Sample preparation is considered the part of the analytical process that introduces the biggest number of errors in the workflow. Although in metabolic phenotyping the best sample preparation is the minimum sample preparation, in GC-MS metabolic profiling, this cannot be applied. Only apolar, volatile/semivolatile, thermally stable analytes can be directly analyzed using GC-MS [4]. Thus, GC-MS is more suitable for low-molecular-weight metabolites, such as organic acids, fatty acids, alcohols, sterols, and other types of molecules [28] that become volatile in the conditions of the experiment. For all other nonvolatile molecules, derivatization is required, which results in thermally stable, volatile derivatives. Derivatization also reduces the polarity of the polar hydroxyl-, amine-, carboxyl-, and thiol groups [1] present in the analyzed molecules, decreasing supramolecular interactions among them and between them and the stationary phase, thus improving chromatographic separation. The derivatization process is considered to be the hallmark of GC-MS metabolic profiling and may determine to a greater extent the quality of the results. Both sample preparation and derivatization processes will be discussed more extensively in the next section (Section 4). As a result, GC-MS is not practical for global profiles but is limited to the analysis of apolar, volatile to semivolatile molecules (or those that can be made volatile), as mentioned. In addition, the molecular weights of the analyzed should be below 600 amu.

Analysis should accommodate the elution of all compounds of interest and their detection with the highest possible sensitivity and efficiency. The majority of the published works and protocols involve GC in "universal" columns, with a thermal gradient covering the elution temperature of the majority of the sample constituents. Detection is typically performed in single-quadrupole MS analyzer following electron ionization. Full scan mode is employed from 20 to 600 amu. This instrumental combination, which is by far dominating the GC-MS arsenal, offers high level of performance, high stability in addition to exhaustive documentation that includes thousands of publications, technical reports, and wide instrument availability. TOF-MS analyzers are not as common, and so have found less use in such metabolic studies.

Injection occurs in split or splitless mode. The former is better for instrument protection, while the latter is expected to provide higher detection sensitivity for metabolites occurring in trace levels.

At the time that samples are injected on the instrument, many issues should be taken into account. The whole analytical batch should be considered in detail. The production of derivatives through the derivatization process is time dependent. The duration of the applied analytical method and the number of samples subjected to analysis, should be predesigned carefully, in order to avoid variations between the derivatized samples [29]. Smaller batches are preferable, but introduce batch effects, a fact that needs demanding data handling to be overcome [30]. Smaller batches also allow maintaining good control of the analytical system and more easily correcting problematic issues (troubleshooting) with regard to instrumentation.

The stability and the good functional condition of the analytical system, and the repeatability (aiming at repeatable retention times and compound signals) can easily be monitored by using quality control (QC) samples [31, 32]. QC samples are usually pooled aliquots of tested samples, prepared (derivatized) in exactly the same way with the real samples. A good practice is to analyze QC samples in the beginning ("conditioning"), in the middle and in the end of every batch [33]. QC samples are also a very useful tool in data processing [34, 35]. They could be used to detect and correct time trends, retention time drifts and to adjust for batch-to-batch differences [29].

GC-MS data quality should be in accordance to certain proposed criteria, which are, for example, the number of zero values (<40%), and CV% of the peak areas (<30% for QC samples). Such measures secure the quality and the validity of the obtained results [29, 31]. Comparison of the chromatographic peak areas of reference standards injected within the batch could also be used complementarily in the evaluation of analytical figures of merit.

Finally, injector-related parameters (mode and temperature) and the oven temperature program should be optimized in each case. A post run backflush program or application of high temperatures (bake-out) acts as a protection for the column and is always necessary in order to avoid introducing impurities to the MS analytical system [29].

Data processing or data handling is the next step in a GC-MS metabolic profiling workflow. It mainly aims to filter acquired data, remove or correct errors, and generate peak tables for further statistical analysis. Chromatographic peaks should be effectively identified and their abundance could be quantified. This highly important step is discussed in detail (Section 5).

Statistical evaluation (multivariate and univariate) of processed data and biochemical interpretation of the results complete the puzzle. A clear and reliable picture of the analyzed sample differentiations has been obtained. In Fig. 2, a typical workflow followed in GC-MS-based metabolic phenotyping is illustrated.

4 SAMPLE PREPARATION FOR GC-MS-BASED METABOLIC PHENOTYPING

Sample preparation mainly aims to improve the chromatographic behavior and the detectability of the compounds as well as to isolate compounds from the matrix and remove them from the analytical system [36]. Every type of sample should be switched to a form compatible with GC-MS analysis. Sample preparation for GC-MS analysis evolves on two axes. The first one is the handling needed to prepare the samples/extracts, and the second is the production of derivatives of the metabolites contained in the samples.

Usual practices for sample preparation include liquid-liquid extraction (LLE), solid-phase extraction (SPE), protein precipitation, homogenization, centrifugation, sonication, and others. The nature (type) of sample determines the process to be applied. Liquid extraction or SPE is often inappropriate for volatile samples because incomplete extraction and losses may occur if the extracts are subsequently concentrated prior to analysis [37]. Hence, headspace-solid phase microextraction (HS-SPME) is preferable for volatile and semivolatile compounds [37, 38]. SPE is extensively used to process biological samples, which typically are complex matrices where the substrates may interfere in the analysis and thus lead to low sensitivity and selectivity. LLE is commonly used in many protocols, assisted by vortexing, homogenization, and other methods. Polar metabolites are preferably extracted with isopropanol, methanol, acetonitrile, water, or mixtures of these solvents. More lipophilic metabolites are better extracted with chloroform, ethyl acetate, or other apolar solvents. The adjustment of pH value may have dramatic effect in the distribution of certain ionizable molecules in one or other of the extracting phase. Adjustment, control, or at least monitoring of pH value of samples is highly recommended if such extraction steps are applied.

Many metabolic phenotyping GC-MS-based studies use blood plasma or serum. Addition of organic solvent (usually acetonitrile or methanol) is necessary for protein precipitation from blood-derived samples [39]. In urine, sample proteins do not generally become an issue as they are practically absent in the specimen (unless the patient is proteinuric as can occur as a consequence of

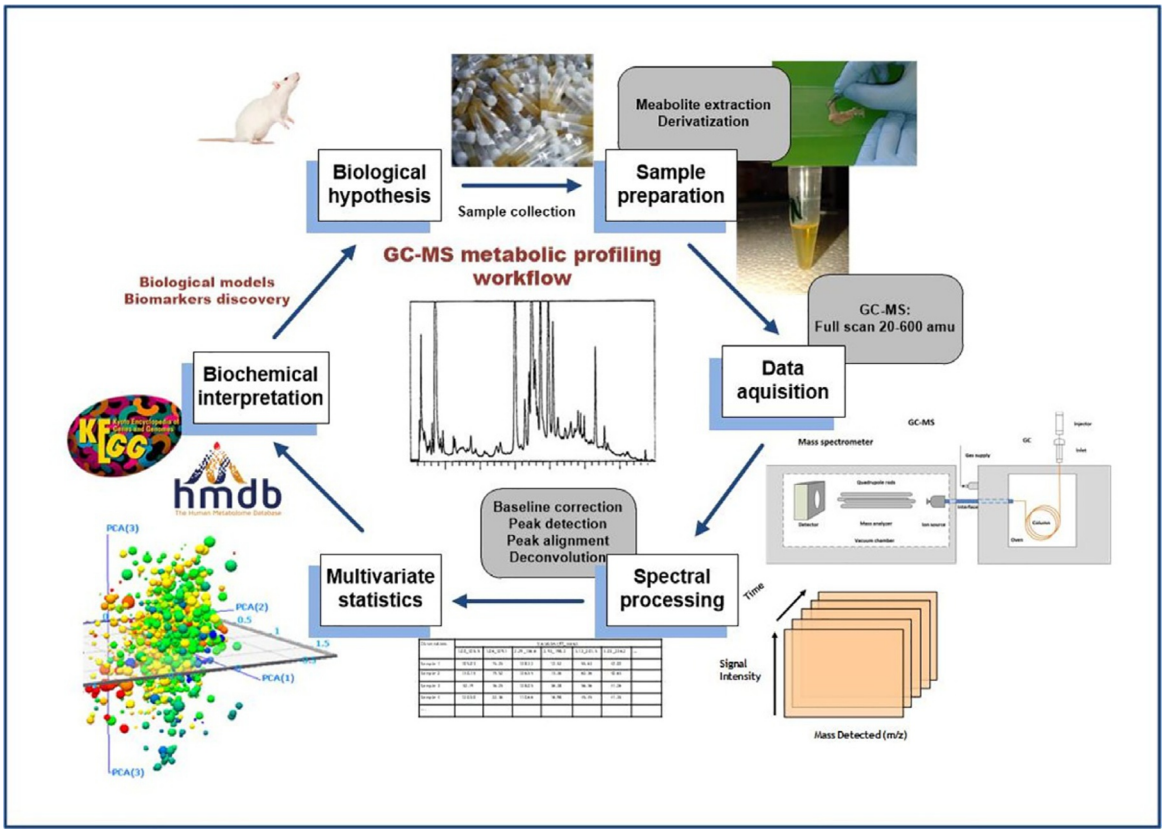

FIG. 2 A typical GC-MS metabolic profiling workflow is presented.

some disease). However, high variability of the water content could prove a serious factor of variance that could introduce analytical bias. Thus, normalization techniques are often required in order to overcome this issue [40]. Fecal samples need a more demanding process in order to extract the plurality of contained metabolites. The existing protocols of fecal sample preparation for metabolic profiling purposes have been recently reviewed [26]. Our group also worked on the optimization of a fecal sample preparation protocol for GC-MS analysis [41]. Although lyophilization reduces the contained moisture of fecal samples, it is not recommended prior to GC-MS analysis as it may result in the loss of volatile compounds such as short-chain fatty acids [42].

In work from our group [41], a pooled fecal sample was mechanically homogenized and treated under three different ratios of sample weight to buffer volume ratios (W_f/V_s) and pH values. Both aqueous and organic extracts were evaluated using untargeted metabolic profiling. We therefore suggested that a ratio of 0.5 W_f/V_s in neutral pH is the best compromise for the treatment of fecal sample prior to untargeted metabolic profiling using both NMR and GC-MS techniques.

Derivatization follows metabolite extraction in the GC-MS sample preparation process. In most protocols, derivatization should take place in the complete absence of moisture. Thus, the clear supernatant obtained after the addition of organic solvents to the samples (as described earlier) should be evaporated to dryness, typically under a gentle nitrogen stream. In the majority of reported GC-MS metabolic phenotyping studies, a two-step derivatization is conducted on biological samples. This includes methoxyamination and sylilation reactions.

A well-established derivatization protocol comprises the addition of methoxylamine hydrochloride (40 mg/mL in pyridine) at 28–37°C, up to 120 min and then the addition of *N*-methyl-*N*-(trimethylsilyl)-trifluoracetamide (MSTFA) at 37°C for 30 min to 1 h [18, 39].

Other derivatization methods include the use of alternative reaction agents such as ethyl chloroformate (ECF) [43], methylchloroformate (MCF) [44], and heptafluorobutyl chloroformate (HFBCF) derivatization reagents [45, 46] for organic acids, amines, and amino acids [47]. In contrast to trimethylsilyl (TMS) derivatizing reagents, chloroformate reagents react also in aqueous medium [39], thus providing a promising alternative for straightforward application in biological or food samples. It should be pointed that, to best of our knowledge the existing commercial, validated data/spectral libraries include data for nonderivatized molecules and/or TMS-derivatized molecules.

Derivatization could be considered as the bottleneck of GC-MS, making this technique seemingly unsuited to metabolic phenotyping, if the process is not fully controllable. If not properly studied and planned, the derivatization process can be long and capable of introducing errors in the analysis. The sources of biases derived from this process and the impact on the obtained results have been studied in detail [16] and possible corrections have been proposed [16, 48]. In Fig. 3, the derivatization process as well as the production of the

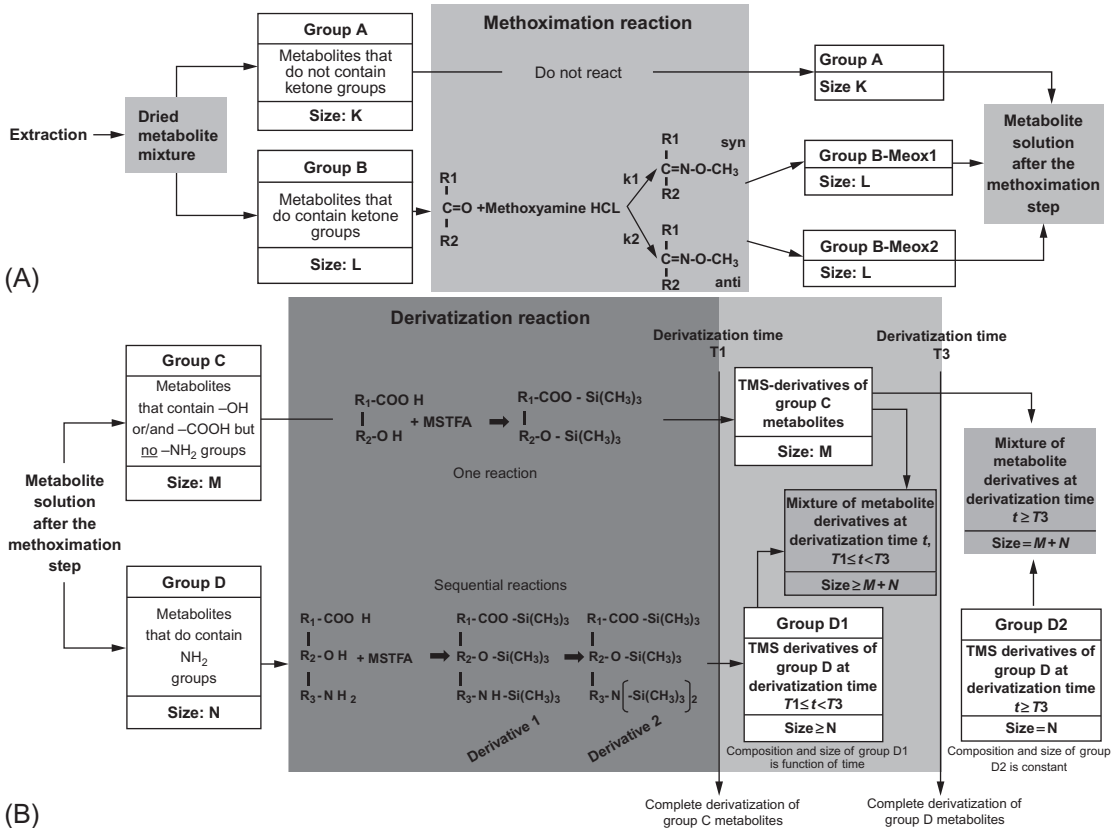

FIG. 3 Schematic diagram of the derivatization process for the production of the (methoxime)-trimethylsilyl derivatives of the free metabolites of a biological sample mixture; see [16]. *(From H. Kanani, P.K. Chrysanthopoulos, M.I. Klapa, Standardizing GC-MS metabolomics, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 871 (2008) 191–201. doi:10.1016/j.jchromb.2008.04.049, with permission from Elsevier.)*

(methoxime)-trimethylsilyl derivatives from a biological mixture are illustrated [16]. According to the presence of different chemical moieties, four metabolite groups were identified by the authors. These groups exhibit different kinetics in the formation of TMS derivatives, thus reaching equilibrium (derivative steady stage) in different time scales.

Such problems may hinder the straightforward unattended analysis of large number of samples: samples cannot be derivatized together and then be left in the autosampler for analysis, if the reaction is still ongoing, as different proportions of TMS derivatives are still forming. The biases derived from the derivatization process affect each metabolite to a different extent, thus increasing the complexity of the data treatment process. For example, it should be stressed that one molecule that contains three active hydrogen moieties (OH, NH_2, COOH species) may form one, two, or three derivatives. For example, asparagine (Asn) will form Asn-TMS, Asn bis-TMS, and Asn tris-TMS derivatives. Production of all three derivatives occurs in parallel but with different kinetics, thus resulting in different product ratios (Asn-TMS: Asn bis-TMS: Asn tris-TMS) depending on the time the reaction is left to proceed.

Another issue is the stability of the formed derivatives. This factor along with the selection of incubation time and ideal temperature and other related parameters has been investigated by different researchers [16, 49–51]. The rule found was that derivatization reagents should be in excess in order to allow for all the metabolites to be derivatized. Critical factors that determine this process are the reaction time and temperature. The end point of derivatization (zero point) should be the time that either no more derivatives are formed or derivative degradation does not yet happen. In this time, samples should be subjected to GC-MS analysis.

Recently, Moros et al. [52] investigated methoxyamination and silylation conditions, as well as the stability of derivatized metabolites at low temperature. Higher numbers of identified metabolites and peak areas were obtained using MSTFA with the catalyst 1% TMCS. Moros et al. also suggested 24-h methoxyamination at room temperature followed by 2-h silylation, as the optimal conditions. The behavior of derivatives of different chemical classes compounds according to postreaction time is illustrated in Fig. 4A (a–d) [52]. Based on the ideal hypothesis that derivatives should remain stable through the time, the authors determined and used a deviation of 20% as a threshold of stability. The authors' general consideration was that amino acids should be analyzed within 18 h from derivatization, while carbohydrates and organic acids should be analyzed within 24 h from derivatization. In addition, in Fig. 4B [52], the stability of 33 metabolite derivatives as a function of storage time is demonstrated. Similar findings to the previously mentioned were extracted, as most amino acids derivatives were stable for 18 h while most carbohydrates derivatives remain stable for 24 h.

Despite these issues, GC-MS remains the platform of choice for volatile and semivolatile compounds. The major advantages offered by this technique are

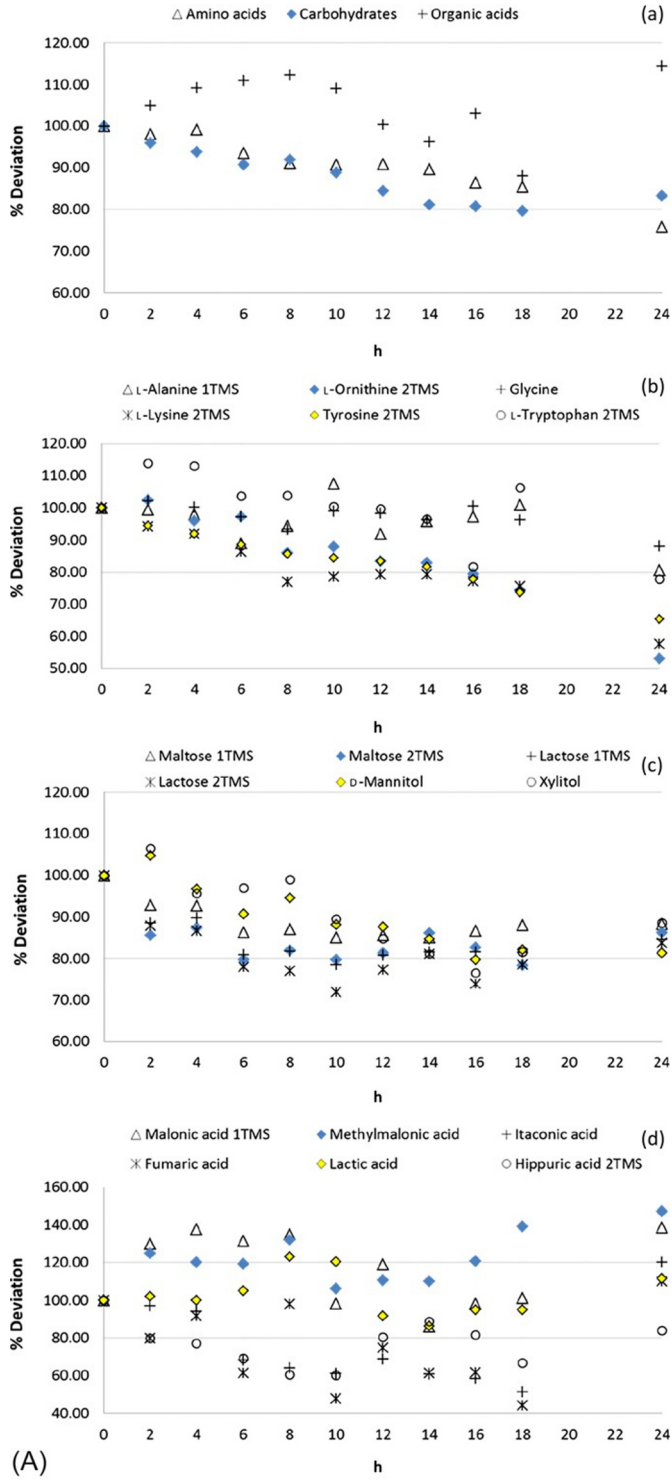

FIG. 4 (A) Stability of TMS derivatives (a) deviation % of sum of means ($n=2$) of normalized peak areas of 22 amino, 15 carbohydrates, and 11 organic acids derivatives through time at 4°C. (b) Deviation % of six representative amino acids derivatives, (c) deviation % of six representative carbohydrates derivatives, (d) deviation % of six representative organic acids derivatives; see [52].

(Continued)

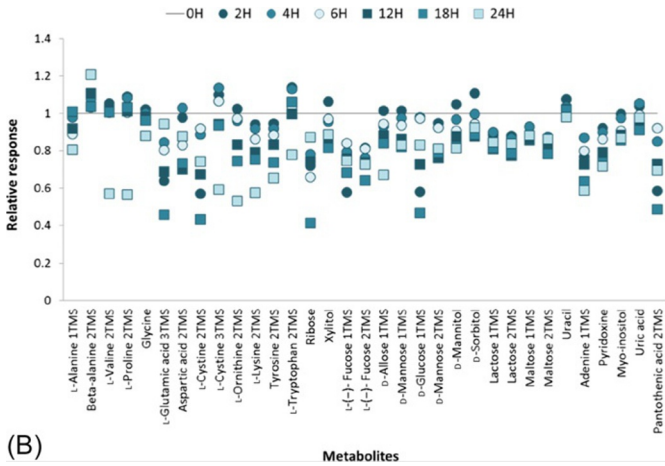

FIG. 4, CONT'D (B) Stability of 33 derivatives over time at 4°C before analysis after normalization with the mean of their corresponding normalized peak area at 0 h immediately after derivatization; see [52]. *(From G. Moros, A.C. Chatziioannou, H.G. Gika, N. Raikos, G. Theodoridis, Investigation of the derivatization conditions for GC-MS metabolomics of biological samples, Bioanalysis 9 (2017) 53–65. doi:10.4155/bio-2016-0224. © Future Science.)*

considered to be the widespread availability of GC-MS instrumentation and the large numbers of trained users, the robustness of the method, and the availability of commercial electron impact (EI) spectral libraries [4].

5 DATA PROCESSING

GC-MS data analysis in metabolic phenotyping studies can be treated in two modes: either by first identifying the metabolites by analyzing all peaks observed in the chromatogram or as holistic analysis, where data processing is performed with special algorithms developed by the vendor or in the context of open source code/programing.

Taking the first mode, peak annotation and metabolite identification is not trivial and remains a challenging step. Over the last decade, this issue has been the focus of much research and development. In fact, of the three main analytical platforms applied in metabolic phenotyping, GC-MS is the most advanced and effective for metabolite identification. This is because years of experimentation in electron ionization (EI) GC-MS analysis for thousands of analytes have accumulated impressive knowledge about ion fragmentation and have populated spectral databases and libraries. As already mentioned, the comparative advantage of GC-MS analysis is the very high reproducibility of the EI process, which generates the same fragments of a given analyte. This high reproducibility (yet unmet by other techniques) enables the creation of EI spectral libraries

and databases, which are either local or web-based. General-purpose spectral databases (e.g., National Institute of Standards and Technology (NIST), Wiley) include thousands of chemical compounds covering various categories. In addition, more specialized libraries include the Agilent-Fiehn Metabolomics Database or the Pfleger-Maurer-Weber-Drugs-and-Pesticides-Library for toxicology [29]. The former is focused on endogenous metabolites and provides data for TMS derivatives. The correct use of such libraries can facilitate the identification of a wide range of metabolites. Several papers have described this part of the analysis, shedding light on the steps of this complex process [53–57].

Another reason that makes peak annotation in GC-MS more manageable than LC-MS is the use of the retention index of a metabolite [17, 29]. Retention indices are calculated by linking metabolite retention time with the retention times of standard *n*-alkanes or other standards (e.g., fatty acid methyl esters (FAMES)), analyzed under the same conditions [4]. In all modes of separation science (liquid or gas chromatography as well as electro-driven capillary separation modes such as capillary electrophoresis), retention (or migration) times are strongly dependent on column and instrument conditions. Hence, reporting retention times shows little promise for enhancing metabolite identification. Small changes in the mobile phase or the aging of the column will result in retention time changes. Retention indices on the other hand are not instrument specific and can be used to compare data across systems or laboratories. The FiehnLib GC/MS library utilizes fatty-acid methyl esters (FAMEs) retention index system from GC-TOF and GC-Quad for over 1000 small metabolites, including lipids, amino acids, fatty acids, amines, alcohols, sugars, amino sugars, sugar alcohols, sugar acids, organic phosphates, hydroxyl acids, aromatics, purines, and sterols, as methoximated and trimethylsilylated mass spectra under electron impact ionization [2]. In practice, the use of RI necessitates a separate injection of the RI standard, which provides very little additional effort; however, using RI provides the means for crossinstrument and crosslaboratory comparisons, while in addition it promotes data reporting.

The reader should be aware that data processing in GC-MS metabolic phenotyping becomes complex as metabolites often exist in more than one form. As mentioned earlier, an amino acid may be derivatized in one or more functional groups, and these forms will elute in different peaks at different retention times. In MS analysis, each eluting molecule (peak) produces several fragments [4, 15]. Libraries often include data for both free (nonderivatized) compounds and their derivatives and capture single or double TMS derivatives [29].

In most cases, the way to quantify metabolites is to use only one (quantifier ion) of the numerous generated signals and manually identify them based on the appropriate library (for example, using the Fiehn Library for endogenous compounds). This is a time-consuming procedure when it comes to detecting many metabolites in many samples. The other ions are used to corroborate metabolite identification and reporting (qualifier ions).

The alternative choice is the use of special platform for spectral deconvolution and library matching using special software such as Automated Mass Spectral Deconvolution and Identification System (AMDIS) linked to the NIST database. The setting parameters for deconvolution using AMDIS software should be suitably selected to allow only the well-formed peaks to be detected and to exclude the background noise peaks. The processing steps of AMDIS software are initially the noise analysis, then the "capturing" of the compounds and peaks deconvolution, and finally the compounds identification using an appropriate spectral library [58].

Lately, researchers have shown an increased use of holistic data analysis tools that perform peak picking and alignment and additional data processing and mining steps. Such software includes "MetaboAnalyst 3.0" [59], which processes MS data on a handy web-based platform, XCMS, which uses R language [60], MetSign based on MatLab [61], MetDAT [62], Metabolome express [63], and Metabolite detector [64] for GC-MS [53]. XCMS is a well-established platform for both LC-MS and GC-MS data. Filtering and peak identification followed by identification across samples, nonlinear retention time correction and filling of missing values are performed. Finally, statistical analysis and peak visualization complete the processes on the platform [65]. Metacomparisons of complex data are performed using metaXCMS [66]. Vendor-specific tools have been developed and depending on the instrument also enable data processing to be performed. Examples of vendor software include MSD ChemStation and a range of different add-ons (Agilent), Tracefinder, and Excalibur software (Thermofisher).

These tools, either freely or commercially available, are used for peak detection and integration, alignment, warping of the *m*/*z*-retention time plane, and subtraction of the background noise, deisotoping of data, and finally extraction of metabolite peaks by deconvolution, and the use of spectroscopic libraries, etc. [67]. Despite their effectiveness, the platforms' typical problems include lost peaks, inconsistency with databases, noisy signals counted as real peaks, bad alignment, errors in the integration, etc. These errors are intensified in areas with a small signal-to-noise ratio, where the metabolites of interest are present in low concentrations.

Peak detection and deconvolution, for qualitative and quantitative purposes, are of great importance for further statistical analysis, in order to reach the final goal of the discovery of biomarkers [68]. Background noise constitutes a major issue and overcoming this issue is critical for correct detection and identification of the metabolites of interest. The appropriate method should be able to detect all metabolites without counting the background noise. Increased number of peaks or features could often only mean false parameters in peak picking and false final results. Traditionally, detection algorithms follow two strategies, "derivative techniques" or "matched filter response" [57].

Normalization is an appropriate step to reduce the variance attributed to experimental sources, but it should be applied with caution as it is also possible to reduce the variance due to the biological response [69].

In a recent study, our group used both AMDIS and XCMS in order to handle and compare the metabolic fingerprinting of fecal samples prepared with different conditions [41]. The number of detected features and the sum of the intensities of all features were compared for both aqueous and organic fecal extracts. Based on both approaches, more concentrated organic fecal extracts in neutral pH and aqueous ones in basic pH gave higher scores.

In the same line of study, cecal samples from exercised and sedentary rats were analyzed by GC-MS. Data were processed with both instrument software and XCMS [70]. After the statistical analysis, features found to be altered in cecal samples from the exercise group of rats were annotated using the Fiehn library.

The typical problem of multiple signals per metabolite and the abundance of artifacts originating from the derivatizing agents [4] could be bypassed by annotating only the compounds that contribute to differentiation of the samples. Such a process may promote handling of complex data matrices.

Misinterpretation of data could derive from even small changes in RT and to overcome this correct alignment of spectra is required. In XCMS, MetAlign and other software [39] peak alignment, baseline correction, and noise subtraction are performed. Although AMDIS is considered to be one of the most powerful software packages for GC-MS data processing, it is not capable of performing alignment across samples [65]. The output obtained is most often in the open source MS exchange format netCDF [39]. Peak picking, alignment, and deconvolution-related parameters should be selected carefully in order to avoid false-positive and -negative results [39]. Another procedure for GC-MS-based metabolic phenotyping data analysis, called TagFinder, was presented in 2008 [71]. It is based on Java and performs the connection of fragment ions from different chromatograms to “mass tags” [65].

6 STATISTICAL EVALUATION

Univariate and multivariate statistical analysis follows the data processing. In univariate statistical analysis (*t*-test, ANOVA, etc.), each variable is examined individually, providing the advantage of easy handling and interpretation of the outcomes. The disadvantage is that multiple interactions between various metabolites are not estimated in this way [53].

On the contrary, multivariate statistical analysis deals with the evaluation and interpretation of complex data matrices, taking into account simultaneously all contributing factors and therefore can recognize the complex interactions between the groups [53]. Multivariate statistical analysis is mainly used in

metabolic phenotyping due to its ability to deal with highly collinear and noisy data [72] and to create models that are able to predict individual samples.

In unsupervised methods (PCA), the aim is to find trends in a data matrix (usually denoted as *X*), while in supervised methods (PLS, OPLS) the patterns found in the *X* matrix are correlated with additional data (e.g., clinical data) contained in a matrix *Y* [73]. In supervised models, information about the desired distinction of groups of samples is already given [74, 75].

One of the most valuable tools of multivariate statistical analysis is PCA (principal component analysis), in which data from the complex data matrix are compressed, in order for new latent variables to be produced (principal components), the number of which is less, or at least equal to the initial number of variables [73].

The most widely used supervised/discriminant approach is based on partial least squares regression (PLS), which when the *Y* matrix is class information can be partial least squares-discriminant analysis (PLS-DA). Based on multiple linear regression models, this method is used to describe the linear relationship of the outcome variables *Y* to the *X* data variables. Discriminant analysis is mostly applied when the aim is the best possible differentiation of two or more groups of samples. An expansion of the PLS methods is orthogonal partial least squares (OPLS), in which the elucidated covariance between *X* and *Y* variables is maximized by orthogonal correction [76, 77]. Chapters 8 and 9 of this volume provide the full theory and background to univariate and multivariate methods, respectively.

Both PCA and PLS methods are integral tools of metabolic phenotyping used to clarify the separations between the various experimental group samples as analyzed either by NMR spectroscopy or MS [53, 57]. However, without the appropriate validation, such methods could lead to unreliable and even erroneous results.

In fact, this issue is in highly important when PLS models are created to explain the separation of samples. PLS models are often created when PCA fails to expose group separation [78]; in addition PLS is used to better identify the features (metabolites) with the highest contribution to the group separation. However, caution is needed as group separation is predefined in biased models and without rigorous validation, the results could be groundless.

Permutation tests are usually used to validate a PLS model. Data from the *Y* data matrix are repeatedly and randomly permuted, and a new classification is calculated based on new data [79].

Commonly using software for multivariate statistical analysis is SIMCA (Umetrics). In addition online platforms such as Metaboanalyst [59] can also perform multivariate statists.

In Fig. 5, PCA (A & C) and OPLS-DA (B & D) scores plots from multivariate statistical analysis of the time-dependent metabolic-based changes in zebrafish embryogenesis [80] are presented. OPLS-DA scores plots of both GC-MS and LC-MS analysis data separated the observations into five clusters instead of

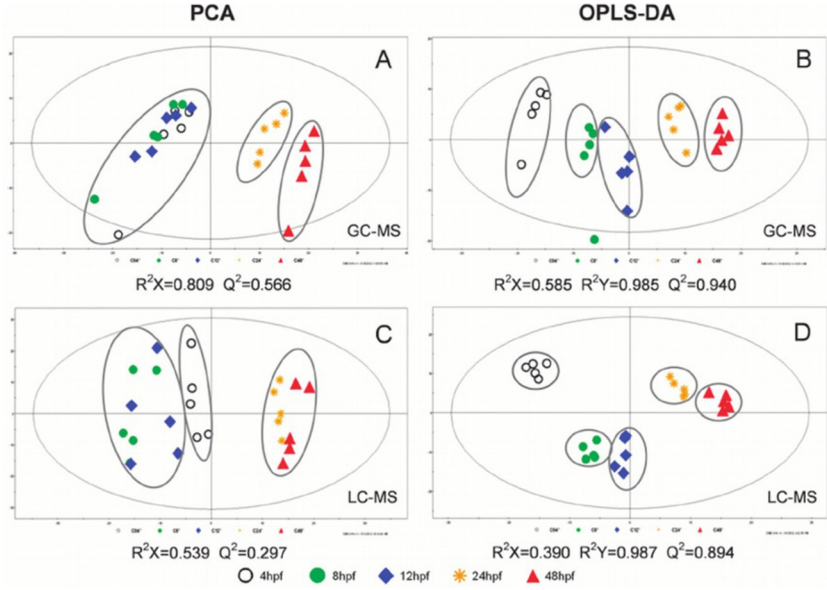

FIG. 5 PCA (A and C) and OPLS-DA (B and D) scores plots of GC-MS and LC-MS data, of the time-dependent metabolic-based changes in zebrafish embryogenesis are presented; see [80]. *(From S.-M. Huang, F. Xu, S.H. Lam, Z. Gong, C.N. Ong, Metabolomics of developing zebrafish embryos using gas chromatography- and liquid chromatography-mass spectrometry, Mol. Biosyst. 9 (2013) 1372–1380. doi:10.1039/c3mb25450j, with permission from Royal Society of Chemistry.)*

three clusters in PCA. Zebrafish embryogenesis has a fast progression and metabolic changes occur in the different stages (4, 8, 12, 24, 48 h postfertilization). Mean centering and unit variance scaling was performed on the obtained data. A sevenfold crossvalidation was applied to both PCA & OPLS-DA as well as a further validation by permutation tests, in SIMCA-P 11.0 (Umetrics).

According to the authors, early development stages were distinguished from the posterior ones. The evaluation of OPLS-DA and loading plots and the analysis of VIP (variable importance in the projection) values (>1) demonstrated several differentiated metabolites between the stages [80].

7 BIOCHEMICAL INTERPRETATION

The ultimate goal of a metabolic phenotyping-based study is often achieved by correlating the discovered biomarkers to the related biochemical pathways. Along with the identification of metabolites, biochemical interpretation of the results is often the most challenging step.

Important questions include the following: (1) how are the biochemical pathways changed in response to examined stimulus or treatment, (2) in which biochemical pathway(s) are the differentiated metabolites involved, (3) have

the related biochemical pathway(s) changed according to the studied effect, (4) are there crosstalk effects that alter the related biochemical pathways? All these questions are difficult to answer and represent a constant "headache" for the researchers. To be deciphered, they require in-depth knowledge of the biochemistry of the studied system. It is indeed difficult to decide the effect that a stimulus has on the whole metabolism, from studying only a snapshot of the metabolome. This snapshot is imprinted on the analyzed sample, when this is captured at the moment of sampling. As the metabolome is affected by various (confounding) factors, this analysis may lead to poor interpretation.

This great need is assisted by online databases that have been developed and used to facilitate biochemical interpretation. A review of available databases is given in Chapter 11 of this volume. Some of the most widely used databases include KEGG (kyoto encyclopedia of genes and genomes), [81] which is a reference database for pathway mapping consisting of 18,111 metabolites and small molecules, MetaCyc [82, 83] consisting of 2526 pathways from 2844 different organisms, the SMPDB "small molecule pathway database" [84, 85] consisting of more than 700 small molecule pathways found in humans (around 400 are not found in any other pathway database), WikiPathways [86], which contains over 1600 pathways and supports 21 species. Valuable information is also extracted from HMDB "The Human Metabolome Database" [87, 88], available online, which contains detailed information about small molecule metabolites found in the human body (114,100 metabolite entries).

In addition, either commercial pathway analysis software tools, such as Pathway Studio, MetaCore etc., or online free available platforms such as MetaboAnalyst [59] facilitate the biochemical interpretation. For example, MetaboAnalyst is bridged with KEGG metabolic pathways and the contribution of the involved pathways can be visualized in the platform.

The importance and the complexity of that metabolic phenotyping technology is dealing with is imprinted on the integrated map of metabolic pathways, called the 22nd (2003) edition of the IUBMB-Sigma-Nicholson Metabolic Pathways Chart, which is also available online.

8 GC-MS TROUBLESHOOTING

An issue that the analyst usually has to deal with is the handling of problems during the GC-MS analysis. Problems are usually related either to the baseline, the quality of chromatographic peaks and results, or to the performance of the instrumentation.

Common baseline problems include drifting, falling, rising, and often irregular shapes. High standing current and high-frequency noise can be also observed. Very common peak-related problems are broadening or double peaks. Other problems are associated with peaks fronting, tailing and clipping, or having irregular shape. Solvent-related issues could also appear, for example no

peaks after a solvent peak or solvent peak tailing. Ghost peaks, the absence of peaks, or high-intensity contaminant peaks are some of the unpleasant situations that the analyst could meet [89, 90].

Such problems can be observed and identified more easily in targeted analysis, where the analyst is looking for a certain number of peaks. Peak failures can be efficiently identified either in total ion current or (even better) in extracted ion mode plots. In untargeted analysis, however, identification of instrument failure becomes less straightforward. The analyst does not know the content of the samples analyzed, so deformed peak shapes cannot be directly attributed to column contamination or to wrong settings in the injector. It is highly advised that researchers add injection standards to the real test samples, spiking with octafluoronaphthalene (OFN) or other such molecule not expected to be present in the real samples. Spiking with an isotope-labeled compound such as a deuterated standard provides additional levels of control of the analysis. In addition, the periodic analysis of a well-known sample, either a QC sample or a sample not related to the test samples, can also provide an efficient way to monitor system performance. Keeping track record of these samples run over several weeks or months allows for fast control of system suitability by comparison of retention times, peak areas, and peak shapes.

If examination of the obtained data shows low reproducibility, poor sensitivity, retention time fluctuations, and inconsistency, these are all markers that something went wrong during data acquisition.

Periodic maintenance by the analyst and authorized technical personnel is essential in order to keep the analytical system in optimal operation function.

Injectors either split/split-less or PTV (programmed temperature vaporizing) injectors represent an important part of the GC-MS system, the role of which is often underestimated. Keeping the injector in good operational order requires the replacement of the septum and the liner regularly by the analyst. The correct setting of temperature of injection, especially in PTV injectors, may provide a solution to problems related to solvent/reagent peaks, reduced peak capacity, and column overloading.

A useful practice when planning instrumental configurations is the incorporation of backflush. Metabolic phenotyping GC-MS analysis uses an excess of derivatization reagents, while injections may be performed in splitless mode to enhance the detection of trace analytes. Application of a backflush period of ca. 10 min after each run cleans the whole chromatographic path that the analytes follow, eliminating contaminants, thus improving stability and reproducibility. Carryover, ghost peaks, or detection drifts (in retention time and MS signal, e.g., slope in baseline) are also overcome. In contrast without backflush, highly retained components are not removed from the system and a long bake-out period is needed. Even if this is applied after each analysis and for at least the same duration of the analysis (e.g., ca. 60 min), matrix peaks can appear sporadically in various analyses. Over several batches, such components will still remain in the system and may cause column contamination, detection problems,

and erroneous quantification of trace analytes. These factors also cause reduced MS performance and increase the need for ion source cleaning and instrument maintenance.

Leaks could be another worrisome issue. Depending on the instrument configuration, automatic system or manual controls are used to check the condition of the analytical system. The corrective actions that can be followed in order to fix possible leaks are also dependent on the instrument.

Finally, MS rather than chromatography-related problems could be low sensitivity, high baseline, peak tailing, and low reproducibility of the obtained data. Some of the errors can be solved by an experienced analyst without resorting directly to expert technicians.

For example, baseline drifting could be eliminated by removing the last part of the column. Stabilization of the column possibly corrects baseline falling. Conditioning of the column and cleaning of the detector are usually the "remedies" of baseline rising. The broadening of the peaks is maybe due to low flow of the column or dirty injector. Peak tailing can be reduced by increasing the column or oven temperature. In all cases, care should be taken not to reach or overcome the maximum temperature recommended for the stationary phase.

More complex problems such as low reproducibility need careful examination of the various steps of the analytical system and the analytical process. Unsuitable injection technique or parameters and nonoptimal sample preparation process may be a source of error. Again, it is good practice in this case to compare the results with a series of standard injections.

9 CLINICAL APPLICATIONS

The ideal analytical method should provide high sensitivity, the ability to cover a wide range of concentrations (which may actually span up to eight orders of magnitude) for a broad spectrum of metabolites, and provide both quantitative and qualitative analysis. The latter should extend from peak annotation to obtaining structural information of unknown analytes that would help in metabolite identification. Such results are of utmost importance in order to promote biomarker discovery, which is the overall end point in metabolic phenotyping studies [13]. Chapter 16 of this volume covers the applications in clinical diagnosis and prognosis using all analytical techniques. Here, we focus on providing an overview of the use of GC-MS.

As already mentioned, there is no technique that can fully cover the whole range of metabolites in a given sample. Each technique performs better for certain aspects such as the analysis of a type of sample and different metabolites. However, each technique has its own limiting factors and only the combination of complementary techniques can achieve the "complete" analysis of a sample.

For example in NMR spectroscopy, signals from highly concentrated analytes could dominate spectra and interfere with signals from lower

concentrations analytes. LC-MS is a very sensitive and versatile analytical system. Many parameters should be carefully considered to ensure reproducible and reliable results, but unfortunately such systems can suffer from poor cross system reproducibility. This does not allow for efficient combination of data and the generation of commercial curated spectral libraries. The final outcome is the very slow, painstaking identification of unknown peaks (metabolites).

Due to all these factors, GC-MS has been widely used in the last decades in contemporary bioanalysis and in metabolic phenotyping in particular. Compared to LC-MS/MS, GC-MS lacks sensitivity in most cases. However, the recent rise in the use of GC-MS/MS brings the technologies shoulder to shoulder with regard to detection sensitivity.

Kanani et al. describe (Table 1, in their paper, 2008) the applied fields for GC-MS-based metabolic phenotyping. These range from basic research (e.g., discovery of new biochemical pathways), to applied research, such as study of disease biomarkers, toxicology, agricultural and nutrition applications, etc. [16].

GC-MS metabolic phenotyping constitutes a multipurpose tool, which responds ideally to demands of biomedical science. Evaluation of pathological phenotypes for both prognostic and diagnostic purposes of several diseases has been achieved, using GC-MS metabolic analysis. A plethora of studies document the investigation of multiple kinds of biological specimens. Metabolic profiling for several diseases, such as cancer and cardiovascular diseases, is in the first line of GC-MS-based metabolic phenotyping [91–93].

For targeted metabolite-based analyses GC is usually coupled to single quadrupole MS detectors, focusing on the detection and quantification of several endogenous compounds (amino acids, organic acids, and sugars) in urine, serum, plasma, feces, tissues, yeast cells, and plants. Begou et al. have recently reviewed the field of targeted metabolic phenotyping [47]. The instrumental methods used in targeted studies and their application are very similar to other multianalyte methods such as those used in antidoping control, drug of abuse toxicology analysis, or pesticide analysis in environmental and food samples. A key difference is that in targeted metabolic phenotyping the analytical tools are used for the identification or confirmation of potential biomarkers and indeed detect and (semi)quantitate metabolites over several orders of magnitude in concentration in all samples. In comparison, in the aforementioned multianalyte applications, prohibited substances under investigation are not detected in the majority of the analyzed samples, and in the positive samples detection or quantitation of one or a handful (in rare cases) of analytes occurs. Overall targeted metabolic phenotyping represents a demanding but upcoming research line with increasing application in various fields.

In our group, the cost-effective and reliable GC-MS technique has been applied in several studies in the analysis of complex matrices and has proven to be a powerful tool. Samples collected via celiocentesis, during the first trimester of gestation, were analyzed using GC-MS among other platforms.

In total, 52 and 42 metabolites were identified in celomic fluid and cells, respectively, revealing their potential use in mammalian system biology studies [94].

GC-MS metabolic profiling has been also used to decipher the biochemistry of exercise [70, 95]. In rat plasma samples from four groups, including an exercise group and a group treated with allopurinol and controls, 86 metabolites were annotated. Several compounds were altered, affected either by exercise or the administration of allopurinol. Metabolites found to be responsible for the differentiation of the exercise group were lactic acid, pyruvic acid, 2-hydroxybutyric acid, uracil, oxalic acid, pyroglutamic acid, and stearic acid. The allopurinol administration effect was reflected more in purine catabolic pathway compounds and amino acid levels [95].

In a life-long exercise metabolic phenotyping study, gut tissue samples (cecal samples) from rats swimming for 21 months and sedentary ones were analyzed using GC-MS. The results showed that the samples were clearly different metabolically, and the impact of exercise on cecal tissues, probably derived from gut microbe exercise alterations, was demonstrated [70]. In Fig. 6, GC-MS chromatograms of cecal extracts of lifelong exercise (black) and sedentary (blue) rats are illustrated [70]. The annotated metabolites consisted of fatty acids, amino acids, carbohydrates, purines derivatives, and other compounds such as uracil, creatinine, inositol, glycerol, lactic, dehydroascorbic, phosphoric, and 2-ketoisocaproic acids. Based on two data-processing methods (peak areas and XCMS features) arachidonic, palmitic, stearic and linoleic acids, urea, inositol, phosphoric and lactic acids were changed under the impact of exercise [70].

Plasma samples from mice treated for 12 or 15 months with low (10 kcal%) or high (60 kcal%) fat diets were also analyzed by GC-MS. Upon data processing, from the 77 detected metabolites, 48 were further used in the interpretation. A reduction in energy catabolism along with aging was observed by the authors [96].

The ever-increasing demands for collection of reliable quantitative metabolic profiles of clinically valuable samples led to the idea of Human Metabolome Project, in order for metabolite references to be created, based on the available analytical platforms. As part of this effort, Psychogios et al. were focused on the human serum metabolome, using five metabolic profiling methods [97]. In total, 4229 confirmed metabolites that are possible to be found in human serum and their relative information are available. In Fig. 7, a typical high-resolution GC-MS total ion chromatogram of the polar extracts from a pooled serum sample of healthy adult subjects is presented [97]. The samples were extracted separately to obtain polar and lipophilic metabolites under different protocols. By both untargeted and targeted GC-MS analysis, 90 compounds were identified. Furthermore, 45 compounds were detected only by GC-MS and not by NMR spectroscopy, including three very high abundance compounds (phosphoric acid/phosphate, uric acid, and *N*-acetyl-glycine). The authors provide the possible reasons that these compounds were not

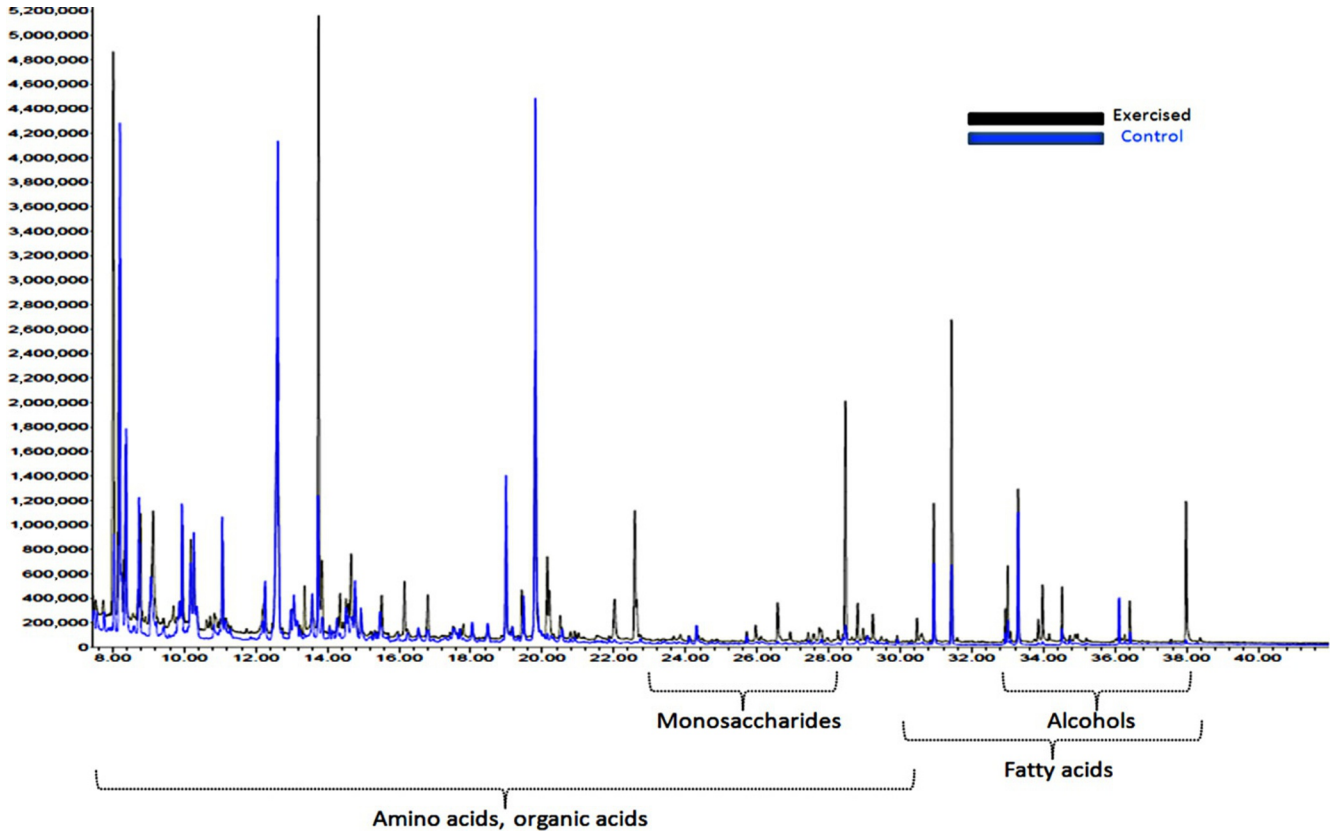

FIG. 6 Overlay of representative GC-MS cecal extract chromatograms of exercise for 21 months *(black)* and sedentary *(blue)* rats; see [70]. *(From O. Deda, H. Gika, T. Panagoulis, I. Taitzoglou, N. Raikos, G. Theodoridis, Impact of exercise on fecal and cecal metabolome over aging: a longitudinal study in rats, Bioanalysis 9 (2017) 21–36. doi:10.4155/bio-2016-0222. © Future Science.)*

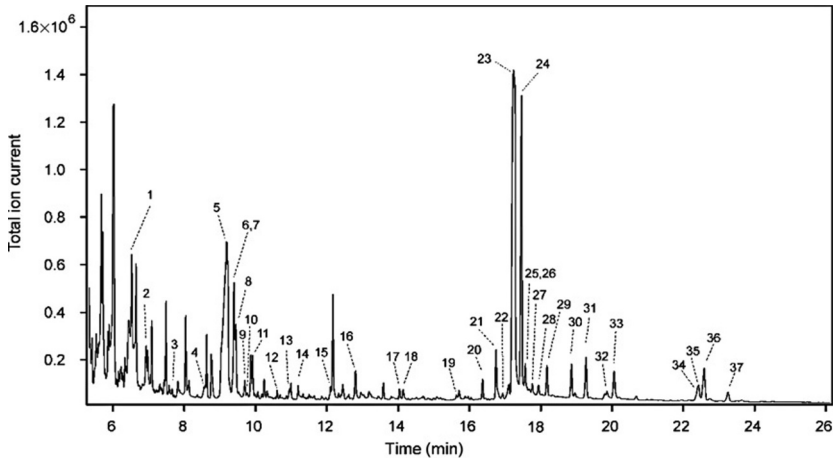

FIG. 7 Typical total ion chromatogram of serum from a healthy subject. Numbers indicate the following metabolites: 1, L-lactic acid; 2, L-alanine; 3, oxalic acid; 4, L-valine; 5, urea; 6, L-L-L-leucine; 7, glycerol; 8, phosphoric acid; 9, L-isoleucine; 10, L-proline; 11, glycine; 12, L-L-L-serine; 13, L-threonine; 14, L-methionine/L-aspartic acid; 15, aminomalonic acid; 16, pyroglutamic acid/L-glutamine; 17, L-glutamic acid; 18, L-phenylalanine; 19, L-ornithine; 20, citric acid; 21, D-erythrofuranose; 22, D-fructose; 23, D-glucose; 24, D-galactose; 25, L-histidine; 26, L-lysine; 27, L-tyrosine; 28, gulonic acid/mannonic acid; 29, D-glucopyranose; 30, 6-deoxy mannose; 31, palmitelaidic acid; 32, palmitic acid; 33, myo inositol; 34, uric acid; 35, L-tryptophan; 36, linoleic acid; 37, oleic acid; 38, stearic acid; see [97].

detectable by NMR spectroscopy and conclude that GC-MS and NMR spectroscopy are complementary techniques in human serum metabolic phenotyping [97].

Another important application of the GC-MS technique is the quantitative analysis of complex mixtures of fatty acids in biological samples [98], which constitutes an individual -omics technology, called lipidomics, the details of which go beyond the interests of this chapter.

10 CONCLUSION

GC-MS still dominates the analysis of volatile and semivolatile molecules and remains the platform of choice in many metabolic phenotyping studies. Applying meticulous and validated protocols can lead to reliable and accurate results. Untargeted investigations in practice mean the metabolic profiling of several biochemical pathways in one analysis. In this aspect, GC-MS technology is still very useful to the researcher. The wide range of GC-MS application results from the large number of GC-MS instruments installed and the large number of trained users in laboratories around the globe. Decades after it was firstly performed, the GC-MS technique provides unsurpassed advantages, such as the robustness and the availability of EI spectral libraries. Since quantification and identification of the whole metabolome, using one analytical platform is

still utopian. GC-MS remains highly useful side by side with other analytical platforms for the characterization of biosample metabolomes rendering it an indispensable tool for the analysis of small molecules.

REFERENCES

[1] C.B. Hill, U. Roessner, Metabolic Profiling of Plants by GC–MS, in: W. Weckwerth, G. Kahl (Eds.), The Handbook of Plant Metabolomics, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2013, pp. 1–23, https://doi.org/10.1002/9783527669882.ch1.

[2] T. Kind, G. Wohlgemuth, D.Y. Lee, Y. Lu, M. Palazoglu, S. Shahbaz, O. Fiehn, FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, Anal. Chem. 81 (2009) 10038–10048, https://doi.org/10.1021/ac9019522.

[3] N. Blow, Metabolomics: biochemistry's new look, Nature 455 (2008) 697–700, https://doi.org/10.1038/455697a.

[4] G. Theodoridis, H.G. Gika, I.D. Wilson, Mass spectrometry-based holistic analytical approaches for metabolite profiling in systems biology studies, Mass Spectrom. Rev. 30 (2011) 884–906, https://doi.org/10.1002/mas.20306.

[5] M.M. Koek, R.H. Jellema, J. van der Greef, A.C. Tas, T. Hankemeier, Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives, Metabolomics Off. J. Metabolomic Soc. 7 (2011) 307–328, https://doi.org/10.1007/s11306-010-0254-3.

[6] W.B. Dunn, D.I. Ellis, Metabolomics: current analytical platforms and methodologies, TrAC Trends Anal. Chem. 24 (2005) 285–294, https://doi.org/10.1016/j.trac.2004.11.021.

[7] D.I. Ellis, W.B. Dunn, J.L. Griffin, J.W. Allwood, R. Goodacre, Metabolic fingerprinting as a diagnostic tool, Pharmacogenomics 8 (2007) 1243–1266, https://doi.org/10.2217/14622416.8.9.1243.

[8] A. Agin, D. Heintz, E. Ruhland, J.M. Chao de la Barca, J. Zumsteg, V. Moal, A.S. Gauchez, I.J. Namer, Metabolomics—an overview. From basic principles to potential biomarkers (part 1), Médecine Nucl. 40 (2016) 4–10, https://doi.org/10.1016/j.mednuc.2015.12.006.

[9] P. Agarwal, A. Goyal, A review on analyzers for mass spectrometry, Int. J. Pharma Bio Sci. 8 (4) (2017). http://ijpbs.net/abstract.php?article=6228. Accessed 21 October 2017.

[10] M.A. Baldwin, Mass spectrometers for the analysis of biomolecules. in: Methods in Enzymology, Academic Press-Elsevier, Cambridge, MA, 2005, pp. 3–48, https://doi.org/10.1016/S0076-6879(05)02001-X.

[11] M. Nič, J. Jirát, B. Košata, A. Jenkins, A. McNaught (Eds.), IUPAC Compendium of Chemical Terminology: Gold Book, 2.1.0, IUPAC, Research Triagle Park, NC, 2009, https://doi.org/10.1351/goldbook.

[12] A.L. Burlingame, Biological Mass Spectrometry, Academic Press-Elsevier, Cambridge, MA, 2005.

[13] I. Wilson, Chapter 5—Chromatographic and electrophoretic separations combined with mass spectrometry for metabonomics. in: The Handbook of Metabonomics and Metabolomics, Elsevier Science B.V., Amsterdam, 2007, pp. 149–169, https://doi.org/10.1016/B978-044452841-4/50006-0

[14] G.G. Harrigan, R. Goodacre, Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis, Springer Science & Business Media, Berlin/Heidelberg, Germany, 2012.

[15] G. Theodoridis, H.G. Gika, I.D. Wilson, Analytical methodology for clinical metabolic profiling: sample preparation, sample analysis and statistical evaluation. in: Global Metabolic Profiling: Clinical Application, Future Science Ltd, London, 2014, pp. 20–34, https://doi.org/10.4155/ebo.13.379.

[16] H. Kanani, P.K. Chrysanthopoulos, M.I. Klapa, Standardizing GC-MS metabolomics, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 871 (2008) 191–201, https://doi.org/10.1016/j.jchromb.2008.04.049.

[17] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, T.H.S.M. (HUSERMET) Consortium, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, Nat. Protoc. 6 (2011) 1060–1083, https://doi.org/10.1038/nprot.2011.335.

[18] J. Lisec, N. Schauer, J. Kopka, L. Willmitzer, A.R. Fernie, Gas chromatography mass spectrometry-based metabolite profiling in plants, Nat. Protoc. 1 (2006) 387–396, https://doi.org/10.1038/nprot.2006.59.

[19] S. Osorio, P.T. Do, A.R. Fernie, Profiling primary metabolites of tomato fruit with gas chromatography/mass spectrometry. in: Plant Metabolomics, Humana Press, New York City, NY, 2011, pp. 101–109, https://doi.org/10.1007/978-1-61779-594-7_7.

[20] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K.P. Patel, P.D. Fraser, P.M. Bramley, Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS, J. Exp. Bot. 56 (2005) 219–243, https://doi.org/10.1093/jxb/eri069.

[21] N. Schauer, D. Steinhauser, S. Strelkov, D. Schomburg, G. Allison, T. Moritz, K. Lundgren, U. Roessner-Tunali, M.G. Forbes, L. Willmitzer, A.R. Fernie, J. Kopka, GC–MS libraries for the rapid identification of metabolites in complex biological samples, FEBS Lett. 579 (2005) 1332–1337, https://doi.org/10.1016/j.febslet.2005.01.029.

[22] A. Garcia, C. Barbas, Gas chromatography-mass spectrometry (GC-MS)-based metabolomics, Methods Mol. Biol. (Clifton NJ) 708 (2011) 191–204, https://doi.org/10.1007/978-1-61737-985-7_11.

[23] M.F. Almstetter, I.J. Appel, M.A. Gruber, C. Lottaz, B. Timischl, R. Spang, K. Dettmer, P.J. Oefner, Integrative normalization and comparative analysis for metabolic fingerprinting by comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry, Anal. Chem. 81 (2009) 5731–5739, https://doi.org/10.1021/ac900528b.

[24] H. Lu, Y. Liang, W.B. Dunn, H. Shen, D.B. Kell, Comparative evaluation of software for deconvolution of metabolomics data based on GC-TOF-MS, TrAC Trends Anal. Chem. 27 (2008) 215–227, https://doi.org/10.1016/j.trac.2007.11.004.

[25] O. Deda, H.G. Gika, I. Taitzoglou, N. Raikos, G. Theodoridis, Impact of exercise and aging on rat urine and blood metabolome. An LC-MS based metabolomics longitudinal study, Metabolites 7 (2017), https://doi.org/10.3390/metabo7010010.

[26] O. Deda, H.G. Gika, I.D. Wilson, G.A. Theodoridis, An overview of fecal sample preparation for global metabolic profiling, J. Pharm. Biomed. Anal. 113 (2015) 137–150, https://doi.org/10.1016/j.jpba.2015.02.006.

[27] A. Mastrangelo, A. Ferrarini, F. Rey-Stolle, A. García, C. Barbas, From sample treatment to biomarker discovery: a tutorial for untargeted metabolomics based on GC-(EI)-Q-MS, Anal. Chim. Acta 900 (2015) 21–35, https://doi.org/10.1016/j.aca.2015.10.001.

[28] O. Fiehn, Metabolomics by gas chromatography-mass spectrometry: the combination of targeted and untargeted profiling, Curr. Protoc. Mol. Biol. 114 (2016) 30.4.1–30.4.32. F.M. Ausubel (ed.), https://doi.org/10.1002/0471142727.mb3004s114.

[29] H.G. Gika, I.D. Wilson, G.A. Theodoridis, Chapter 9—The role of mass spectrometry in non-targeted metabolomics. in: C. Simó, A. Cifuentes, V. García-Cañas (Eds.), Comprehensive Analytical Chemistry, Elsevier, Amsterdam, 2014, pp. 213–233, https://doi.org/10.1016/B978-0-444-62651-6.00009-X.

[30] M. Chen, R.S.P. Rao, Y. Zhang, C.X. Zhong, J.J. Thelen, A modified data normalization method for GC-MS-based metabolomics to minimize batch variation, SpringerPlus 3 (2014), https://doi.org/10.1186/2193-1801-3-439.

[31] H.G. Gika, G.A. Theodoridis, M. Earll, I.D. Wilson, A QC approach to the determination of day-to-day reproducibility and robustness of LC–MS methods for global metabolite profiling in metabonomics/metabolomics, Bioanalysis 4 (2012) 2239–2247, https://doi.org/10.4155/bio.12.212.

[32] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, Within-day reproducibility of an HPLC − MS-based method for metabonomic analysis: application to human urine, J. Proteome Res. 6 (2007) 3291–3303, https://doi.org/10.1021/pr070183p.

[33] L. Lai, F. Michopoulos, H. Gika, G. Theodoridis, R.W. Wilkinson, R. Odedra, J. Wingate, R. Bonner, S. Tate, I.D. Wilson, Methodological considerations in the development of HPLC-MS methods for the analysis of rodent plasma for metabonomic studies, Mol. Biosyst. 6 (2009) 108–120, https://doi.org/10.1039/B910482H.

[34] M.A. Kamleh, T.M.D. Ebbels, K. Spagou, P. Masson, E.J. Want, Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies, Anal. Chem. 84 (2012) 2670–2677, https://doi.org/10.1021/ac202733q.

[35] R. Wehrens, J.A. Hageman, F. van Eeuwijk, R. Kooke, P.J. Flood, E. Wijnker, J.J.B. Keurentjes, A. Lommen, H.D.L.M. van Eekelen, R.D. Hall, R. Mumm, R.C.H. de Vos, Improved batch correction in untargeted MS-based metabolomics, Metabolomics 12 (2016), https://doi.org/10.1007/s11306-016-1015-8.

[36] S. de Koning, H.-G. Janssen, U.A.T. Brinkman, Modern methods of sample preparation for GC analysis, Chromatographia 69 (2009) 33, https://doi.org/10.1365/s10337-008-0937-3.

[37] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass Spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51–78, https://doi.org/10.1002/mas.20108.

[38] T. Živković Semren, I. Brčić Karačonji, T. Safner, N. Brajenović, B. Tariba Lovaković, A. Pizent, Gas chromatographic-mass spectrometric analysis of urinary volatile organic metabolites: optimization of the HS-SPME procedure and sample storage conditions, Talanta 176 (2018) 537–543, https://doi.org/10.1016/j.talanta.2017.08.064.

[39] K.K. Pasikanti, P.C. Ho, E.C.Y. Chan, Gas chromatography/mass spectrometry in metabolic profiling of biological fluids, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 871 (2008) 202–211, https://doi.org/10.1016/j.jchromb.2008.04.033.

[40] L. Le Moyec, L. Mille-Hamard, M.N. Triba, C. Breuneval, H. Petot, V.L. Billat, NMR metabolomics for assessment of exercise effects with mouse biofluids, Anal. Bioanal. Chem. 404 (2012) 593–602, https://doi.org/10.1007/s00216-012-6165-6.

[41] O. Deda, A.C. Chatziioannou, S. Fasoula, D. Palachanis, N. Raikos, G.A. Theodoridis, H.G. Gika, Sample preparation optimization in fecal metabolic profiling, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 1047 (2017) 115–123, https://doi.org/10.1016/j.jchromb.2016.06.047.

[42] T. Bezabeh, R.L. Somorjai, I.C.P. Smith, MR metabolomics of fecal extracts: applications in the study of bowel diseases, Magn. Reson. Chem. 47 (2009) S54–S61, https://doi.org/10.1002/mrc.2530.

[43] X. Gao, E. Pujos-Guillot, J.-F. Martin, P. Galan, C. Juste, W. Jia, J.-L. Sebedio, Metabolite analysis of human fecal water by gas chromatography/mass spectrometry with ethyl chloroformate derivatization, Anal. Biochem. 393 (2009) 163–175, https://doi.org/10.1016/j.ab.2009.06.036.

[44] B. Vancompernolle, K. Croes, G. Angenon, Optimization of a gas chromatography-mass spectrometry method with methyl chloroformate derivatization for quantification of amino acids in

plant tissue, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 1017–1018 (2016) 241–249, https://doi.org/10.1016/j.jchromb.2016.02.020.

[45] P. Hušek, Z. Švagera, D. Hanzlíková, L. Řimnáčová, H. Zahradníčková, I. Opekarová, P. Šimek, Profiling of urinary amino-carboxylic metabolites by in-situ heptafluorobutyl chloroformate mediated sample preparation and gas chromatography-mass spectrometry, J. Chromatogr. A 1443 (2016) 211–232, https://doi.org/10.1016/j.chroma.2016.03.019.

[46] P. Simek, P. Hušek, H. Zahradníčková, Heptafluorobutyl chloroformate-based sample preparation protocol for chiral and nonchiral amino acid analysis by gas chromatography, Methods Mol. Biol. (Clifton NJ) 828 (2012) 137–152, https://doi.org/10.1007/978-1-61779-445-2_13.

[47] O. Begou, H.G. Gika, I.D. Wilson, G. Theodoridis, Hyphenated MS-based targeted approaches in metabolomics, Analyst 142 (2017) 3079–3100, https://doi.org/10.1039/C7AN00812K.

[48] H.H. Kanani, M.I. Klapa, Data correction strategy for metabolomics analysis using gas chromatography-mass spectrometry, Metab. Eng. 9 (2007) 39–51, https://doi.org/10.1016/j.ymben.2006.08.001.

[49] J. A, J. Trygg, J. Gullberg, A.I. Johansson, P. Jonsson, H. Antti, S.L. Marklund, T. Moritz, Extraction and GC/MS analysis of the human blood plasma metabolome, Anal. Chem. 77 (2005) 8086–8094, https://doi.org/10.1021/ac051211v.

[50] E.A. Bekele, C.E.P. Annaratone, M.L.A.T.M. Hertog, B.M. Nicolai, A.H. Geeraerd, Multi-response optimization of the extraction and derivatization protocol of selected polar metabolites from apple fruit tissue for GC-MS analysis, Anal. Chim. Acta 824 (2014) 42–56, https://doi.org/10.1016/j.aca.2014.03.030.

[51] J. Gullberg, P. Jonsson, A. Nordström, M. Sjöström, T. Moritz, 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. 331 (2004) 283–295, https://doi.org/10.1016/j.ab.2004.04.037.

[52] G. Moros, A.C. Chatziioannou, H.G. Gika, N. Raikos, G. Theodoridis, Investigation of the derivatization conditions for GC-MS metabolomics of biological samples, Bioanalysis 9 (2017) 53–65, https://doi.org/10.4155/bio-2016-0224.

[53] A. Alonso, S. Marsal, A. Julià, Analytical methods in untargeted metabolomics: state of the art in 2015, Front. Bioeng. Biotechnol. 3 (2015), https://doi.org/10.3389/fbioe.2015.00023.

[54] T. Kind, O. Fiehn, Advances in structure elucidation of small molecules using mass spectrometry, Bioanal. Rev. 2 (2010) 23–60, https://doi.org/10.1007/s12566-010-0015-9.

[55] M.P. Styczynski, J.F. Moxley, L.V. Tong, J.L. Walther, K.L. Jensen, G.N. Stephanopoulos, Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery, Anal. Chem. 79 (2007) 966–973, https://doi.org/10.1021/ac0614846.

[56] D.S. Wishart, Computational strategies for metabolite identification in metabolomics, Bioanalysis 1 (2009) 1579–1596, https://doi.org/10.4155/bio.09.138.

[57] L. Yi, N. Dong, Y. Yun, B. Deng, D. Ren, S. Liu, Y. Liang, Chemometric methods in data processing of mass spectrometry-based metabolomics: a review, Anal. Chim. Acta 914 (2016) 17–34, https://doi.org/10.1016/j.aca.2016.02.001.

[58] AMDIS (NIST), (n.d.). http://chemdata.nist.gov/mass-spc/amdis/explanation.html (Accessed 22 October 2017).

[59] J. Xia, I.V. Sinelnikov, B. Han, D.S. Wishart, MetaboAnalyst 3.0—making metabolomics more meaningful, Nucleic Acids Res. 43 (2015) W251–W257, https://doi.org/10.1093/nar/gkv380.

[60] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, Anal. Chem. 78 (2006) 779–787, https://doi.org/10.1021/ac051437y.

[61] A. Lommen, H.J. Kools, MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware, Metabolomics 8 (2012) 719–726, https://doi.org/10.1007/s11306-011-0369-1.

[62] A. Biswas, K.C. Mynampati, S. Umashankar, S. Reuben, G. Parab, R. Rao, V.S. Kannan, S. Swarup, MetDAT: a modular and workflow-based free online pipeline for mass spectrometry data processing, analysis and interpretation, Bioinformatics (Oxf., Engl.) 26 (2010) 2639–2640, https://doi.org/10.1093/bioinformatics/btq436.

[63] A.J. Carroll, M.R. Badger, A. Harvey Millar, The MetabolomeExpress project: enabling web-based processing, analysis and transparent dissemination of GC/MS metabolomics datasets, BMC Bioinformatics 11 (2010) 376, https://doi.org/10.1186/1471-2105-11-376.

[64] K. Hiller, J. Hangebrauk, C. Jäger, J. Spura, K. Schreiber, D. Schomburg, MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis, Anal. Chem. 81 (2009) 3429–3439, https://doi.org/10.1021/ac802689c.

[65] Y. Qiu, D. Reed, Gas chromatography in metabolomics study, Adv. Gas Chromatogr.Intech, Rijeka, 2014, https://doi.org/10.5772/57397.

[66] R. Tautenhahn, G.J. Patti, E. Kalisiak, T. Miyamoto, M. Schmidt, F.Y. Lo, J. McBee, N.S. Baliga, G. Siuzdak, metaXCMS: second-order analysis of untargeted metabolomics data, Anal. Chem. 83 (2011) 696–700, https://doi.org/10.1021/ac102980g.

[67] A. Scalbert, L. Brennan, O. Fiehn, T. Hankemeier, B.S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart, S. Wopereis, Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research, Metabolomics 5 (2009) 435–458, https://doi.org/10.1007/s11306-009-0168-0.

[68] G.A. Nagana Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah, D. Raftery, Metabolomics-based methods for early disease diagnostics: a review, Expert Rev. Mol. Diagn. 8 (2008) 617–633, https://doi.org/10.1586/14737159.8.5.617.

[69] Y. Wu, L. Li, Sample normalization methods in quantitative metabolomics, J. Chromatogr. A 1430 (2016) 80–95, https://doi.org/10.1016/j.chroma.2015.12.007.

[70] O. Deda, H. Gika, T. Panagoulis, I. Taitzoglou, N. Raikos, G. Theodoridis, Impact of exercise on fecal and cecal metabolome over aging: a longitudinal study in rats, Bioanalysis 9 (2017) 21–36, https://doi.org/10.4155/bio-2016-0222.

[71] A. Luedemann, K. Strassburg, A. Erban, J. Kopka, TagFinder for the quantitative analysis of gas chromatography—mass spectrometry (GC-MS)-based metabolite profiling experiments, Bioinformatics (Oxf., Engl.) 24 (2008) 732–737, https://doi.org/10.1093/bioinformatics/btn023.

[72] P.S. Gromski, H. Muhamadali, D.I. Ellis, Y. Xu, E. Correa, M.L. Turner, R. Goodacre, A tutorial review: metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding, Anal. Chim. Acta 879 (2015) 10–23, https://doi.org/10.1016/j.aca.2015.02.012.

[73] B. Worley, R. Powers, Multivariate analysis in metabolomics, Curr. Metabolomics 1 (2013) 92–107, https://doi.org/10.2174/2213235X11301010092.

[74] H.J. Issaq, T.D. Veenstra, Proteomic and Metabolomic Approaches to Biomarker Discovery, Academic Press-Elsevier, Cambridge, MA, 2013.

[75] P. Jonsson, Multivariate processing and modelling of hyphenated metabolite data, (2005). http://lnu.diva-portal.org/smash/record.jsf?pid=diva2:144165 (Accessed 8 January 2017).

[76] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), J. Chemom. 16 (2002) 119–128, https://doi.org/10.1002/cem.695.

[77] T. Verron, R. Sabatier, R. Joffre, Some theoretical properties of the O-PLS method, J. Chemom. 18 (2004) 62–68, https://doi.org/10.1002/cem.847.

[78] B. Worley, R. Powers, PCA as a practical indicator of OPLS-DA model reliability, Curr. Metabolomics 4 (2016) 97–103, https://doi.org/10.2174/2213235X04666160613122429.
[79] J.A. Westerhuis, H.C.J. Hoefsloot, S. Smit, D.J. Vis, A.K. Smilde, E.J.J. van Velzen, J.P.M. van Duijnhoven, F.A. van Dorsten, Assessment of PLSDA cross validation, Metabolomics 4 (2008) 81–89, https://doi.org/10.1007/s11306-007-0099-6.
[80] S.-M. Huang, F. Xu, S.H. Lam, Z. Gong, C.N. Ong, Metabolomics of developing zebrafish embryos using gas chromatography- and liquid chromatography-mass spectrometry, Mol. Biosyst. 9 (2013) 1372–1380, https://doi.org/10.1039/c3mb25450j.
[81] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30.
[82] R. Caspi, H. Foerster, C.A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S.Y. Rhee, A.G. Shearer, C. Tissier, T.C. Walk, P. Zhang, P.D. Karp, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases, Nucleic Acids Res. 36 (2008) D623–D631, https://doi.org/10.1093/nar/gkm900.
[83] A. Karnovsky, T. Weymouth, T. Hull, V.G. Tarcea, G. Scardoni, C. Laudanna, M.A. Sartor, K.A. Stringer, H.V. Jagadish, C. Burant, B. Athey, G.S. Omenn, Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data, Bioinformatics (Oxf., Engl.) 28 (2012) 373–380, https://doi.org/10.1093/bioinformatics/btr661.
[84] A. Frolkis, C. Knox, E. Lim, T. Jewison, V. Law, D.D. Hau, P. Liu, B. Gautam, S. Ly, A.C. Guo, J. Xia, Y. Liang, S. Shrivastava, D.S. Wishart, SMPDB: the small molecule pathway database, Nucleic Acids Res. 38 (2010) D480–D487, https://doi.org/10.1093/nar/gkp1002.
[85] T. Jewison, Y. Su, F.M. Disfany, Y. Liang, C. Knox, A. Maciejewski, J. Poelzer, J. Huynh, Y. Zhou, D. Arndt, Y. Djoumbou, Y. Liu, L. Deng, A.C. Guo, B. Han, A. Pon, M. Wilson, S. Rafatnia, P. Liu, D.S. Wishart, SMPDB 2.0: big improvements to the small molecule pathway database, Nucleic Acids Res. 42 (2014) D478–D484, https://doi.org/10.1093/nar/gkt1067.
[86] T. Kelder, M.P. van Iersel, K. Hanspers, M. Kutmon, B.R. Conklin, C.T. Evelo, A.R. Pico, WikiPathways: building research communities on biological pathways, Nucleic Acids Res. 40 (2012) D1301–D1307, https://doi.org/10.1093/nar/gkr1074.
[87] D.S. Wishart, T. Jewison, A.C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorndahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner, A. Scalbert, HMDB 3.0—the human metabolome database in 2013, Nucleic Acids Res. 41 (2013) D801–D807, https://doi.org/10.1093/nar/gks1065.
[88] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.-A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. Macinnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, HMDB: the human metabolome database, Nucleic Acids Res. 35 (2007) D521–D526, https://doi.org/10.1093/nar/gkl923.
[89] GC-thermo-troubleshooting, (n.d.). http://www.chromacademy.com/troubleshooter-gc/resources/gc-thermo-troubleshooting-10.pdf (Accessed 22 October 2017).
[90] CFT Backflush Agilent Technologies, (n.d.). https://www.agilent.com/cs/library/brochures/5989-9804EN.pdf (Accessed 22 October 2017).
[91] X. Chen, L. Liu, G. Palacios, J. Gao, N. Zhang, G. Li, J. Lu, T. Song, Y. Zhang, H. Lv, Plasma metabolomics reveals biomarkers of the atherosclerosis, J. Sep. Sci. 33 (2010) 2776–2783, https://doi.org/10.1002/jssc.201000395.

[92] T.L. Weir, D.K. Manter, A.M. Sheflin, B.A. Barnett, A.L. Heuberger, E.P. Ryan, Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults, PLoS One 8 (2013), https://doi.org/10.1371/journal.pone.0070803.

[93] H. Wu, R. Xue, L. Dong, T. Liu, C. Deng, H. Zeng, X. Shen, Metabolomic profiling of human urine in hepatocellular carcinoma patients using gas chromatography/mass spectrometry, Anal. Chim. Acta 648 (2009) 98–104, https://doi.org/10.1016/j.aca.2009.06.033.

[94] C. Virgiliou, I. Sampsonidis, H.G. Gika, N. Raikos, G.A. Theodoridis, Development and validation of a HILIC-MS/MS multitargeted method for metabolomics applications, Electrophoresis 36 (2015) 2215–2225, https://doi.org/10.1002/elps.201500208.

[95] A. Pechlivanis, A.C. Chatziioannou, A.S. Veskoukis, D. Kouretas, V. Mougios, G.A. Theodoridis, GC-MS analysis of blood for the metabonomic investigation of the effects of physical exercise and allopurinol administration on rats, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 966 (2014) 127–131, https://doi.org/10.1016/j.jchromb.2014.02.005.

[96] K. Spagou, G. Theodoridis, I. Wilson, N. Raikos, P. Greaves, R. Edwards, B. Nolan, M.I. Klapa, A GC-MS metabolic profiling study of plasma samples from mice on low- and high-fat diets, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 879 (2011) 1467–1475, https://doi.org/10.1016/j.jchromb.2011.01.028.

[97] N. Psychogios, D.D. Hau, J. Peng, A.C. Guo, R. Mandal, S. Bouatra, I. Sinelnikov, R. Krishnamurthy, R. Eisner, B. Gautam, N. Young, J. Xia, C. Knox, E. Dong, P. Huang, Z. Hollander, T.L. Pedersen, S.R. Smith, F. Bamforth, R. Greiner, B. McManus, J.W. Newman, T. Goodfriend, D.S. Wishart, The human serum metabolome, PLoS One 6 (2011), https://doi.org/10.1371/journal.pone.0016957.

[98] O. Quehenberger, A.M. Armando, E.A. Dennis, High sensitivity quantitative lipidomics analysis of fatty acids in biological samples by gas chromatography-mass spectrometry, Biochim. Biophys. Acta 1811 (2011) 648–656, https://doi.org/10.1016/j.bbalip.2011.07.006.