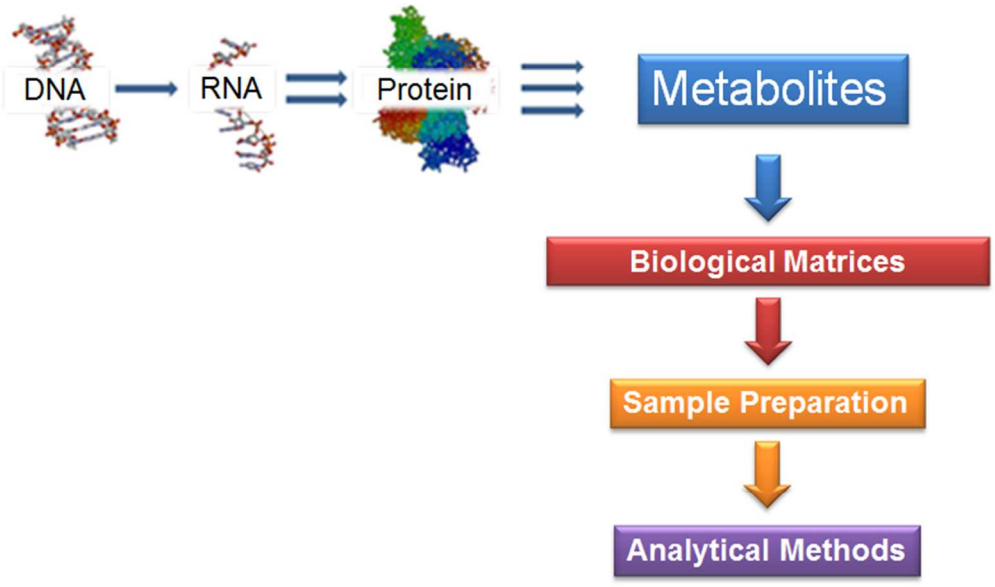
Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst



278x177mm (72 x 72 DPI)

Sampling and analysis of metabolome in biological fluids

Maria José Nunes Paiva.[a,b], Helvécio Costa Menezes.[a], Zenilda de Lourdes Cardeal.[a1]

[a] Departamento de Química, ICEx, Universidade Federal de Minas Gerais, Avenida Antônio Carlos, 6627 – 31270901 Belo Horizonte/MG Brazil.

[b] Universidade Federal de São João Del Rei, Avenida Sebastião Gonçalves

Abstract

Metabolome analysis involves the study of small molecules that are involved the metabolic responses that occur through patho-physiological changes caused by genetic stimuli or chemical agents. Qualitative and quantitative metabolome analyses are used for the diagnosis of various diseases or chemical exposure. This article presents an overview of the different analytical methods available for performing the determination of the metabolome, including sampling, sample preparation and processing and interpretation of data. Critical comments are aimed at emphasizing the extraction methods as well as the biological samples used for metabolome analysis and data processing.

KEY WORDS: Metabolome, environmental exposure, methods.

[1] Corresponding Author: Zenilda@ufmg.br

1 Introduction

By analogy with other ‘‘omes’’, the metabolome is defined as the quantitative complement of low-molecular-weight metabolites present in a biological fluid, cell or organism under a given set of physiological conditions. The concept of the metabolome as a global view of metabolites produced by an organism was introduced in the 70s. Thus, the existing methods would be insufficient to analyze the profile of the human metabolome. Even with the advancement of analytical techniques, we still face the problem of low sensitivity analyses [1]. Metabolomics involves the study of small molecules from cells, tissues, organisms or other biological fluids. These small molecules involve primary and intermediary metabolites as well as exogenous compounds, such as drugs and other chemical compounds [2, 3].

In establishing the metabolic profile of these small molecules with recognition of transcription and expression/protein activity, the different metabolites found can be studied to explain possible causes of diseases or even to identify substances that are beneficial to the body. An example is the discovery of alkaptonuria as a result of an inborn error in metabolism, which accumulates homogentisic acid, causing browning of the urine [4]. The fingerprint of a specific disease can be defined by the interaction of the methods that are used for high-definition genome, proteome and metabolome studies [5]. The use of post-genomic technologies in large-scale molecular epidemiology is important for identifying associations between molecular markers (genes, proteins, metabolites, etc.) and disease [6]. It is increasingly clear that our ability to modify the metabolic profile will help to discover many previously inaccessible areas of biology [7].

The analysis of small molecules is a problem because the samples are extremely complex and their chemical identification can be difficult [2]. However, the identification and quantification of specific metabolites may indicate normal growth or the presence of cellular changes caused by pollutants present in the workplace, drug use or drug abuse, and exposure to environmental pollutants. Initial metabolomic signatures have already been reported for several disease states, including schizophrenia [8], neuropsychiatric diseases [5], Alzheimer's disease [9, 10], cardiovascular and coronary artery disease [11], hypertension [12], type 2 diabetes [13, 14], liver cancer [15], ovarian cancer [16], breast cancer [17], and Huntington's disease [18]. Metabolomics technology has been used as a tool for the elucidation of the mechanism of various cardiovascular and neurological diseases [5, 11, 19], the toxic effects of drugs, nutrition and cancer [20-23] and provides additional information beyond the genomic and proteomic evidence [24]. This profile is composed of dozens of metabolites that are altered qualitatively and quantitatively in pathological conditions or in various situations of exposure to chemicals [3]. The existing strategies for metabolomic analysis employ metabolomics, metabolite profiling, metabolite fingerprinting, and metabolite target analysis. Table 1 shows the analytical differences in these strategies based on the particular situation being investigated. The endogenous metabolites include sugars, organic acids, amino acids, steroids and many other intermediary metabolites, which present wide variation in polarity, molecular weight and volatility. Analysis of all metabolites is difficult because of their diversity and their concentrations, which range between pmol to mmol in various biological fluids. The diversity of physicochemical properties shows the complexity (involution) in the choice of analytical method. Thus, in the

metabolomics experiment, analytical methods and instrument validation is essential to ensure the applicability, accuracy and precision of the methods. Analytical methods in metabolomics include sample preparation, instrumental analysis, data processing and interpretation of data, all of which are performed in different ways depending on factors such as the nature of the sample and the analytical technique available. One concern in metabolome analysis of microorganisms is the separation of intra- and extracellular metabolites. Extraction methods are used for the extraction of intracellular metabolites and plasma; however, it is important to interrupt metabolism accomplished rapidly with temperature change or pH change in the determination of metabolites from microorganisms or cell cultures. The selection of metabolites to be identified and quantified is performed according to the metabolic pathway to be studied [25], which is one of the important points for an accurate interpretation of the data obtained [26]. Another relevant issue is the use of animals for the evaluation of metabolome. The majority of studies that are performed to evaluate the effects of chemicals in humans have been conducted in animals, a practice that has been greatly reduced by substituting cultured animal cells [24]. The advantages of using cell cultures include the elimination of the use of animals, a significant decrease in cost and a faster response. Identification and quantification of metabolites is difficult because of their kinetic behavior, which hampers collection protocols, and the various classes of chemical compounds, such as peptides, carbohydrates, lipids, nucleosides and products of the catabolism of exogenous compounds [27-29]. Multiple changes in the structure of the metabolites may occur within hours to fractions of seconds [28]. This variability of metabolites poses challenges for the analytical procedures and confidence in

the quality of the data generated, showing that the metabolites can be recovered and quantified. Thus, considering the variety of possible metabolites, the first requirement for the metabolome analysis is that the available techniques should be comprehensive [7]. This variability of metabolites poses challenges for the analytical procedures and confidence in the quality of the data generated, showing that the metabolites can be recovered and quantified. Analytical technologies based on liquid chromatography coupled to mass spectrometry (LC/MS) [30], gas chromatography coupled to mass spectrometry (GC/MS) [1, 31], capillary electrophoresis coupled to mass spectrometry (EC/MS) [32, 33] and nuclear magnetic resonance have been applied in the analysis of the metabolic profiles of organisms and human cells. This review describes the current methods for sampling and analysis of the metabolome in biological fluids.

2 Sampling

The sampling step is a continuous challenge in metabolomics studies that requires previous knowledge of the biological system and involves the extraction of diverse biological sample types. In human metabolomics studies, any biological fluid or tissue can be used for the qualitative and quantitative measurement of endogenous metabolites. The most frequently used samples for exploring the modification of the metabolome are urine and blood. However, other matrices, such as cerebrospinal fluid [34], saliva [35] and erythrocytes [36] are also used to analyze the metabolites. The choice of sample type and method of sample preparation are critical aspects of metabolomics studies. These aspects interfere directly in the data quality and, therefore, the accuracy of the biological interpretation [37] . Other factors that should be considered in sampling design

include age, gender, diurnal variation, and diet of subjects. The selection of an appropriate control group must be performed carefully, especially in studies with patients suffering from disease and receiving medical treatment [38]. The vials used to collect biological samples should be chosen and cleaned carefully to minimize the introduction of contaminants, such as surfactants, which can cause interference in the subsequent steps of analysis. Non-specific binding or container surface adsorption of metabolites should be assessed in metabolomic analysis. In the Whole blood, plasma or serum are present at approximately 8% of protein and lipids though urine does not normally contain these compounds. This absence can be associated with the non-specific binding or container surface adsorption of these molecules that can be evaluated often by the low extraction recovery of the analytes of interest. To prevent the loss of analyte due to non-specific binding or container surface adsorption it can be recommended a post-addition of reagents such as bovine serum albumin (BSA) and Tween-80. This is a simple procedure in the trouble shooting non-specific binding or container surface adsorption of quantitative analysis of biological samples [39]. Ideally, sample collection should strive to be non-invasive and representative. Urine, serum and plasma are the bio fluids that are usually employed by incorporating the functions and phenotypes of many different parts of the body. Moreover, the choice of appropriate biological sample is of great significance in the development of protocols suitable for the preparation and storage of samples. The storage step should be included when the immediate analysis of the sample is not feasible; this procedure is commonly performed with liquid nitrogen (-80 ^{o}C). To prevent changes in the composition of the

samples during storage, the number of cycles of freezing and thawing should be minimized [40].

2.1 Plasma

Blood is composed of two parts: a cellular component (erythrocytes, leukocytes and platelets) and plasma (liquid in which blood cells are suspended), which accounts for approximately 50–55% of the blood volume [41]. Plasma and serum (with and without the addition of anti-coagulants, respectively) are obtained from blood samples by centrifugation, where the supernatant portion is used for the analysis. Blood serum is a primary carrier of small molecules in the body and plays a critical role in transporting nutrients, hormones and metabolites. Interferents present in the whole blood sample can also be found in the membrane and nucleotides, which can interfere with the analysis. Some studies show that the treatment of samples using organic solvent extraction, solid phase extraction and solid phase micro-extraction can decrease or exclude this interference [42-44]. The blood irrigates every tissue and organ in the body, providing a pathway for all the molecules being secreted and excreted by different tissues. Thus, tissue dysfunctions and pathological states can both alter the chemical composition of blood. Plasma and serum contain a range of metabolites, which can be used in the diagnosis of physiological alterations or pathological states [45]. The choice of sample (plasma or serum) is related to the clotting process, which stimulates blood fatty acid biosynthesis where the serum levels of these compounds do not represent physiological concentrations [46]. A variety of substances is present in plasma, including proteins, peptides and electrolytes, and may interfere with the analysis. To reduce the presence of plasma proteins the precipitation with trichloroacetic acid and cold methanol is

the most common practice and reduces the interference in instrumental analysis [47]. In order to minimize this interference in the analysis of the blood and plasma the addition of anti-coagulant and a preservative is required. The use of heparin in samples of blood show an absence of interference in polysaccharides analyses [48]. In metabolome analysis of dried blood spots EDTA is also used to maintain the stability of analytes [49]. In the comparison of 4 different blood collection tubes, EDTA and Na^+-fluoride show stability of analytes in relation to heparin and serum blood that demonstrate clear patterns of noise signals. Thus, in the metabolome studies test the suitability of the blood collection tubes is necessary. However, the anti-coagulant K^+-EDTA plasma is mostly used [50]. The plasma composition directly reflects the catabolic and anabolic processes occurring in the whole organism. However, small metabolic changes from a specific part of the body can dilute the original profiling [51].

2.2 Urine

Urine provides a pattern of polar metabolites discarded from the body as a result of catabolic processes. Urine samples that are collected in large quantities by noninvasive sampling, have lower protein content and lower sample complexity with fewer intermolecular interactions [52]. Urinary metabolomic approaches are likely to be used to screen for potentially earlier diagnostic and prognostic biomarkers of disease. Metabolites present in blood are filtered at the glomerulus and may remain in the tubular lumen and be excreted through the urine. These metabolites can also be excreted from plasma into urine by passive diffusion through the tubule. Another possibility is excretion into urine by active secretion, which involves the uptake of toxicants from the blood into the cells of the renal proximal tubule. There are numerous

transporters, such as specific glucose transporters or nucleotide transporters that play a predominant role in the flux of endogenous substances. The advantages of human urine analysis, compared to blood or cerebrospinal fluid analysis, include noninvasive sample collection and the ease of sampling repetition. Another advantage of urine analysis is its less complex sample preparation because urine contains lower protein and peptide level. In contrast, the number of conjugated compounds can be higher in urine than in serum [53]. Urine samples are generally incubated with urease and glucuronidase because urinary metabolites can be conjugated [31]. The time of urine sample collection appears to play a significant role in clinical studies because dietary, lifestyle and diurnal variation can influence metabolite concentrations. Urine samples are typically collected as random samples, timed samples or 24 h samples. Urine analysis usually presents significant inter-individual variability (representing normal genetic variation) but less pronounced intra-individual variability. The first void urines present more variability and may reflect differences in the lifestyles and diets of the subjects. However, this variability is acceptable and may highlight biomarkers of disease or toxicity [54, 55]. Sampling urine day and night shows that the metabolic profile may differ by the presence of given biomarkers [53]. The storage condition is very important in metabolomic studies because it reflects the stability of the urinary metabolites. Studies of stability become necessary when the data of a few analytes are interpreted as the degradation of all the metabolites, which may be greater than a biological response. The integrity of the sample is ensured and the original state of the biological system is maintained when storage conditions are suitable and can be reflected accurately in the metabolomic study. In studies using urine for

metabolome analyses, the storage conditions before sample preparation are room temperature, refrigerator (4 ^{o}C) or ultra-freezer (-80 ^{o}C) [55-58]. Targeted analysis requires a procedure for sample collection and a container of polypropylene that do not degrade the compounds of interest and are not endowed with special characteristics or additives. When the physicochemical properties of the analytes are known, maintaining them under optimal conditions is simple. However, it is more complicated to prevent deterioration in metabolomics fingerprinting because many compounds are analyzed and, to prevent the deterioration of compounds, the samples are stored at a cold temperature (4 °C) and immediately sent to the laboratory for analysis or for proper storage prior to analysis [59].

2.3 Cerebrospinal Fluid

Cerebrospinal fluid is the secretion product of the central nervous system that fills the ventricles and the subarachnoid space of the brain and spinal column. The composition of Cerebrospinal Fluid (CSF) is directly dependent upon metabolite production rates in the brain. Analysis of the CSF metabolome can offer biochemical insights into central nervous system disorders, such as brain injury, Parkinson's disease and multiple sclerosis [34]. However, given the invasiveness of sample collection, the use of CFS in routine diagnostics may be limited. Early diagnostic biomarkers may be useful for neurodegenerative diseases, in which damage to the central nervous system (CNS) may occur long before symptoms develop. Although blood is the most commonly used bio fluid for clinical chemistry diagnoses, cerebrospinal fluid (CSF) may represent a better sample for diseases of the CNS and the neurochemical state. The analysis of cerebrospinal fluid is used in biomarker discovery studies for various

neurodegenerative central nervous system (CNS) disorders [60, 61]. Cerebrospinal fluid is the secretion product of the central nervous system, where it serves several functions including protection, circulation of nutrients, and transport and excretion of biological substances, which contain lower concentrations of proteins and metabolites than blood [40]. The low protein content of the CSF limits the movement of water-insoluble compounds by cellular transport, which is possible only when such compounds are bound to proteins. Moreover, in most cases, toxicants achieve concentrations in the CSF that are no higher than the concentration of the unbound compounds in the plasma [62]. In a comparative study of bio fluids with similar characteristics, human CSF presented a coefficient of variation (CV) 11% lower than that found in the analysis of plasma [63]. Processing and storage of CSF can affect the results of the amino acid analysis. In untreated CSF samples, compounds were significantly increased after 2 days at temperatures of -20 ^{o}C and -80 ^{o}C but decreased progressively after the third day of storage. CSF is easily contaminated by blood during the sampling process and is one of the most important pre-analytical factors in CSF analysis. Because of the high protein concentration in blood, contamination with a small amount of a blood will affect the protein concentration in CSF. In order to decrease the effect of contamination of CSF with protein of blood, centrifugation of the fluid is suggested before freezing in order to remove the highest amount of these proteins [64]. In a sample of CSF that was treated with trichloroacetic acid (TCA), the metabolites presented better results (both qualitative and quantitative) compared to the analysis of untreated aliquots. CSF samples treated with TCA to remove proteins, where the supernatant pH was adjusted to 7.3 and stored at

-20 °C or -80 °C, presented no significant difference in their metabolite concentrations during a 30 day period [65]. Another pre-analysis influence is the addition of stabilizing agents (protease inhibitors, deproteinization). However, the addition of these agents alters the samples by introducing additional compounds, which lead to new peaks in the chromatographic analysis [66].

2.4 Saliva

Human saliva is a clear, slightly acidic, oral fluid that is secreted mainly from the salivary glands, including the parotid, submandibular, sublingual and other minor glands beneath the oral mucosa and is composed of various secretory products such as proteins and metabolites [67]. This complex matrix plays a role in several physiological processes, such as oral digestion, food swallowing and tasting, tissue lubrication, maintenance of tooth integrity, and antibacterial and antiviral protection [68]. Furthermore, saliva is an ideal medium to be explored for health and disease surveillance. Similar to blood, saliva is a complex fluid containing a variety of enzymes, hormones, antibodies, antimicrobial constituents and cytokines that can be influenced by the physiological state of an individual. Therefore, most compounds found in the blood are also present in saliva and, thus, are used for disease diagnosis and prognosis [69]. Many biomarkers of cancer, cardiovascular, and other diseases can potentially be detected in saliva [70]. Compared to other bio fluids, saliva collection is easy, noninvasive, safe and cost-effective. Additionally, collection of saliva samples can be performed privately, in remote sites, or in clinically challenging situations where blood sampling is not possible [67]. Saliva is an easily accessible bio fluid that can be sampled noninvasively without stress or pain. Metabolomic

approaches using saliva are emerging as a promising clinical strategy for diagnostics [70, 71].

2.5 Exhaled breath, Bronchoalveolar lavage fluid (BALF) and epithelial lining fluid (ELF)

The respiratory tract offers a matrix, exhaled breath, which presents high potential for metabolomic analysis. Exhaled breath contains different chemical species, such as small inorganic molecules (nitric oxide, carbon monoxide, carbon dioxide) and volatile organic compounds (VOCs). Exhaled breath can be collected noninvasively, in both the liquid and gaseous phases, and requires minimal sample preparation [72]. Exhaled breath metabolomics is applied to a living matrix in the absence of externally induced perturbations and has been used to differentiate metabolic profiles of patients with asthma, chronic obstructive pulmonary disease, or cystic fibrosis [73]. The advance in the discovery of new lung disease biomarkers by the omics science and metabolomic techniques has allowed further understanding of the pathological processes. Various biological matrices have been used for the analysis and for the biomarker lung disease discovery [74]. The respiratory epithelium is found in the lining of the respiratory tract and has the function moistening it and it also so protects the airways preventing infection and tissue injury functioning as a barrier to pathogens and foreign particles. Bronchoalveolar lavage fluid (BALF) and epithelial lining fluid (ELF) are biological matrices used for the study of lung physiology or pathogenesis [75] such as allergic asthma [76], acute respiratory distress syndrome (ARDS) [77], chronic obstructive pulmonary disease (COPD) [78] and pneumonia [79]. Bronchoalveolar lavage (BAL) is a medical procedure in which a bronchoscope is inserted through the mouth or nose into the lungs and

fluid is squirted into a small part of the lung and then collected for examination. The ELF is obtained indirectly by sampling BAL fluid (BALF) or using a bronchoscopic microsampling probe, obtaining the direct recovery of the matrix [80, 81]. BALF can be obtained also by non-bronchoscopic methods [82]. The BALF has been the most used technique for clinical and experimental investigation of lung disease because the composition of BALF is considered to represent the ELF [80]. In metabolomics studies the determination of these molecules in BALF may provide access to novel disease biomarkers [78]. Chromium (Cr) and arsenic (As) can potentially disrupt the redox balance and cause respiratory diseases and cancer in humans. Analysis in BALF of samples of animals exposed to these metals showed the combined toxic effects of these metals in the lungs evidenced by an increase in the production of reactive oxygen species and decrease of glutathione content [83]. In the study performed by Evans et al (2014) the 37 different metabolites in the BALF associated with amino acid metabolism, glycolysis and gluconeogenesis, fatty acid biosynthesis, phospholipids, and purine metabolism show that these molecules can be used for prognostic biomarkers in the acute respiratory distress syndrome (ARDS) [77]

2.6 Other samples

Biles are of major importance for the maintenance of cholesterol homeostasis. In addition to their functions in lipid absorption, bile acids/bile alcohols are regulatory molecules for a great number of metabolic processes. Their effects are structure-dependent, and numerous metabolic conversions result in a complex mixture of biologically active and inactive forms. Characterization of bile acid metabolome should also increase the understanding of how different bile acid structures might participate in regulatory processes. Advanced

methods are required to characterize and quantify individual bile acids in these mixtures [84]. Determination of metabolites in tissues provides novel aspects of pathological conditions that cannot be obtained from target-specific fluid measurements. The measurement of metabolites in tissues is of great importance in metabolomic research. For solid tissues, blood should be removed through rapid rinsing before storage; otherwise, the obtained profile may reflect that of both the blood and the solid tissue of interest. Among the most commonly used tissues is the liver, which is involved in the major biochemical reactions of the organism with a number of important functions, such as glycogen storage, glycogenesis, and decomposition of erythrocytes [85-88]. Metabolites, such as amino acids, fatty acids, organic acids, and carbohydrates, reflect changes in liver metabolism [89]. Other tissues can be used in metabolomics studies, such as kidney homogenates containing products of glycolysis and amino acids, as well as organic osmolytes. Many metabolites revealed changes in their levels, including decreased levels of organic osmolytes and amino acids in the inner medulla [90]. The study of cell cultures is one of the most extensive approaches in metabolomics. The level of metabolites in a cell represents integrative information, which is an important advantage when biological functions are to be assessed or phenotypes are to be defined in response to genetic or environmental changes [91]. Metabolome analysis of cells can usually be performed with extracellular or intracellular liquid, representing the endometabolome and the exometabolome, respectively. Metabolic profiling of feces can also be developed, mainly as a noninvasive method, for the diagnosis of gastrointestinal diseases. For metabolome analysis

of feces, the extraction method should be optimized such that a large number of metabolites may be extracted [23].

3 Sample preparation

The sample preparation method should be simple and fast to prevent metabolite loss during the preparation procedure and to enable high-throughput processing [92]. Furthermore, the preparation method should be reproducible and include a metabolism-quenching step to represent the true metabolome composition at the time of sampling. Significant variations in the concentration and the chemical and physical properties of metabolites must be prevented by evaluating the residual enzyme activity. In metabolomic analysis, the choice of the sample-preparation method is crucial to the success of a given analysis. Focused identification of metabolites improves the efficiency of metabolomics research. Sample preparation is one of the stages of analytical chemistry that requires a significant amount of time and is underestimated in metabolome analysis. In a biological system, the number and variety of chemical classes of compounds can exceed one million substances [93, 94]. However, the number of metabolites in simple microorganisms can vary from 240 in simple bacteria, such as *Mycoplasma pneumonia,* to 800 as in *Escherichia coli* [95]. Because there are still many undiscovered genes in the genomes of many organisms with broad substrate specificity, this number may even be two or three times this amount. In plant species, the number of metabolites is higher because of secondary metabolites, which are presented in greater numbers than the primary metabolites. Even using methods to identify a simple race from 100 to 500 compounds does not guarantee that the complete metabolome will be realized. Many methods are used to extract the metabolites so that all the

metabolites are recovered with an adequate sensitivity to be detected. Another concern is the extraction of metabolites from various matrices, including cells, tissues, organisms and biological fluids because the extraction procedures are similar but each has its own peculiarities. The concentration of metabolites is prone to rapid and significant changes with the continued growth of the cells. Because of this feature of biological systems, techniques are needed to promote the mitigation or even termination of metabolic activity. The first step in sample preparation of microorganisms and cell cultures is to freeze the metabolism because the pattern of the metabolites may change prior to analysis, resulting in a different profile in vivo. One concern in metabolome analysis of microorganisms is the separation of intra- and extracellular metabolites. It is possible to obtain a rapid analysis of the metabolic pathways of intracellular metabolites [25]. The treatment of samples with methanol cooling has often been used as a simple and rapid method to achieve termination of metabolism and rupturing of cells, promoting leakage of intracellular metabolites. This approach should be avoided during the washing and isolation of cells because the cells are usually separated from the medium prior to extraction of the metabolites [25, 94, 96]. For cells in the quenching suspension, this treatment is performed using 0.85% methanol/bicarbonate (6:4 v/v) and 0.9% NaCl (0.5 °C). This method prevents leakage of the metabolite, which is not the case when using pure methanol or 60% methanol (-40 °C) and 60% methanol buffer (-40 °C) [97, 98]. For adherent cells, this step is performed before quenching [94, 97]. The metabolites are extracted after quenching; however, the extraction step is very important for evaluating the amount of metabolites and depends on the selectivity of the method. The extraction method should be as similar as

possible for all metabolites that have the same physico-chemical properties, especially in the case of metabolic profiling and targeted metabolomics. Following the quenching step, it is necessary to separate the cells from the medium containing the extracellular metabolites. Alternatively, intracellular metabolites can be extracted by permeation through the cell membrane. When the goal is metabolite targeting or profile analysis of solid samples, the addition of an extraction step for the metabolites of interest is required [99]. For solid biological samples, such as feces and tissues, more elaborate protocols are necessary for their preparation. The most common steps include cooling, homogenization and extraction with the addition of solvents of different polarities. As the polarity of the solvent increases, the range of metabolites extracted also increases. Moreover, extractions should be performed under suitable conditions of the pH, which leads to a high recovery of the metabolites and minimal extraction of interfering materials [100]. The effect of pH on the extraction is important, particularly for the ionizable metabolites, as pH does not interfere with neutral molecules, which therefore need to be co-extracted with compounds of acidic or basic character [101]. Solid phase extraction (SPE) is also employed in studies of the metabolome [57, 102, 103]. This extraction procedure presents many advantages, such as an improved signal response of most metabolites in blood fluids and a quick and simple extraction method, allowing for multi-sample, simultaneous preparation. The SPE pre-concentration technique presents very good recovery and allows for the extraction of low and high levels of amino acids [57]. In SPE, free, conjugated and neutral metabolites of steroids are extracted using different eluents, such as ethyl acetate-methanol, methanol and ammonium hydroxide in methanol, sequentially [104].

Sampling of exhaled breath and exhaled breath condensate (EBC) is creating new opportunities for investigating unknown metabolites and novel biomarkers, which may provide insight into respiratory disease mechanisms [105, 106]. In this type of sampling, the proper breath is crucial to obtaining repeatable and reliable results. The extraction methods must be highly efficient, non-selective, reproducible and cause no degradation of the analytes [107]. However, for the analysis of target metabolites or metabolic profiles, a selective method of sample preparation is necessary to decrease other compounds that may interfere with the analysis. A promising alternative for the study of the global metabolome is Solid Phase Microextraction (SPME) [108]. This technique involves solvent-free, sampling in which pre-concentration and conjugation occurs in one step. Furthermore, this technique can be used in headspace (HS) mode or direct immersion (DI). In HS mode, it is possible to assess the metabolomic composition of volatile compounds from solid and liquid samples. SPME direct immersion in biological fluids and direct sampling *in vivo* also offer the possibility of capturing unstable metabolites and the detection of biotransformation intermediates [109]. Recently, in vivo SPME was applied in brain studies to measure neurotransmitters in freely moving rats [110], in patients undergoing cardiac surgery with cardiopulmonary by-pass [111], and short-lived metabolites included in energetic pathways [112]. The common approach is to collect exhaled air in Tedlar bags or gas-tight syringes, followed by subsequent extraction with an SPME fiber [113]. After sampling by SPME, metabolites can be thermally desorbed, if the analysis technique is Gas Chromatography (GC), or placed in an appropriate solvent or solvent mixture for Liquid Chromatography analysis [114]. There is a variety of commercially available fiber coatings, including

polyacrylate (PA), polydimethylsiloxane (PDMS), carbowax, polyethylene glycol (PEG), and mixed phases of carboxen (CAR)-PDMS [115]. However, in metabolomic analyses, diverse types of analytes are generally determined, thus a single SPME fiber coating may be insufficient for a comprehensive analysis of the complex mixtures present in biological samples. Additionally, there is the problem of recurrent obstruction of the polymeric coating of the fiber by the biological matrix. New studies for evaluating SPME coatings for untargeted metabolomics profiling were tested for compounds from different classes within a broad polarity range [116]. Mixed-mode coatings, such as [(C18 + benzenesulfonic acid), polar-enhanced polystyrene-divinylbenzene (HRP), and phenylboronic acid (PBA)], were found to be the best sorbents, allowing for simultaneous extraction of hydrophobic and hydrophilic metabolites. Ultrafiltration is a simple procedure of sample preparation and has been indicated mainly for polar metabolites. Small molecules can physically separate macromolecules using a special filter that allows the passage of molecules with a given molecular weight. Utrafiltration has been recommended for use prior to LC and Nuclear Magnetic Resonance (NMR) [117, 118 119]. Ultrafiltration is more efficient for protein precipitation than the use of acids or solvents. In addition to ultrafiltration may improve the stability of the post-collection metabolites [120]. However the efficiency of extraction of hydrophobic metabolites can be significantly reduced in the ultrafiltration. To increase the recovery of non-polar metabolites a step of washing the membrane with suitable solvents has been proposed. Recently they have been used to analyze dried blood spots (DBS) for extracting target metabolites and global metabolomics. This extraction technique consists in placing a small drop of blood on a strip of filter paper and

waiting for it to dry [121]. After this step the metabolites are extracted with small amounts of a suitable solvent. Chromatography with off-line or online Tandem Mass Spectrometry [122] and Direct Mass Spectrometry [123] are the main analytical tools that can be used in DBS sampling. Among the main advantages of DBS it is noteworthy to cite the potential for automation, the use of small volumes of sample and solvents, and simplicity of collection and storage [49]. However interference with the role and difficulties in extracting strongly adsorbed metabolites, and high uncertainties in sample volumes are some of the main disadvantages [124]. High uncertainties in the volume of samples originate from the non-uniform distribution of blood spot and changes in the hematocrit levels. To minimize this problem one can use a performed DBS (PDBS) [125]. In PDBS a predetermined (5 to 10 μL) microsampling volume is achieved by spotting the blood with a micropipet on the precut 0.6 cm filter paper disk. The PDBS is inserted into the well for online extraction, concentration, and subsequent analysis [126]. According to the strategy for the analysis of metabolome, a selected number of pre-defined metabolites, generally related to a specific metabolic pathway(s) in metabolite profiling or analysis of one or a few metabolites related to a specific metabolic reaction in metabolite targeted analysis, are used for analytical separation. Physicochemical differences in metabolites present in organisms, such as molecular weight, polarity and charge, contribute to the chemical heterogeneity of these compounds causing many metabolites to remain undetected. Another factor leading to the failure to recognize these compounds is a very large dynamic of concentration ranges [98]. Therefore, both the strategies for metabolome analysis and the physicochemical properties indicate the appropriate sampling method.

4. Data interpretation

Metabolomics studies produce large amounts of data, which require high capacity for analysis and interpretation. Often it is necessary to distinguish the metabolic variability inherent in the biological system of interest, genetic and environmental variability. The handling of data in metabolomics takes place in two stages, processing and analysis. In the processing step the raw data are processed to simplify subsequent analysis. In the analysis stage the processed data are subject to interpretation by the use of multivariate statistical tools [127]. Usually the data processing includes filtering, feature detection, alignment and normalization. The filtering aims to remove the noise. In chromatography data noise random source is attributed mainly to the detection system, while the noise of chemical origin can be related to the mobile and stationary phases [128]. Usually the data processing includes filtering, feature detection, alignment and normalization. The filtering aims to remove the noise. In chromatography data noise random source is attributed mainly to the detection system, while the noise of chemical origin can be related to the mobile and stationary phases. Several algorithms are proposed with the aim of removing noise, among the main ones; filtering with window moving average, median filter in chromatographic direction, Savitzky-Golay type of local polynomial fitting and wavelet transformation[129]. The purpose of the feature detection step of data processing is to identify all signals caused by true ions and avoid detection of false positives. A simple approach is to compare data points in raw data directly. Direct comparison of the raw data is a simple approach, for this purpose of statistical tests that can be used to compare differences between two or multiple datasets to a two-dimensional plot [130]. Deconvolution algorithms

are usually used in MS detectors, because different fragments of the same molecule have the same retention time as well as assumption that their profiles in multiple samples are highly correlated as they are subject to the same biological variation and systematic error. In addition several metabolites may be subject to the same mechanisms in the biological system, therefore their levels are highly correlated [131]. Alignment is required for correcting the retention time differences between runs and combining data from different samples. Correlation optimized warping (COW) and Fast Fourier Transform have been applied to the alignment of chromatographic data [132, 133]. Normalization aims to remove the unwanted systematic bias in ion intensities between measurements, while retaining the interesting biological variation. Statistical models can be used to derive optimal scaling factors for each sample based on a complete dataset, such as normalization by unit norm, median of intensities, or the maximum likelihood method [134, 135]. After processing the data should be analyzed to highlight relevant biological information. However, due to the high complexity and large volume of data in metabolomics, data-dimensionality reduction it is necessary to use multivariate statistical analysis [136]. The algorithms used for this purpose can be grouped into unsupervised clustering algorithms, and supervised classifiers [137]. In supervised methods there must be some initial information about the identity of the samples for training classes and the aim is to develop a model based on information contained in the samples. On the other hand, the unsupervised methods, the separation of classes occurs without the need for initial information on the nature of the samples and the goal is to identify natural groupings among the samples [138]. The unsupervised methods commonly used in metabolomics are Hierarchical

cluster analysis (HCA) [139] and Principal components analysis (PCA) [140]. These methods are simple and provide the reduction of the dimensionality of data NMR-based or MS-based studies, revealing the inherent data structure. In HCA dataset is clustered into different groups creating representative dendograms and hierarchical relationships of data points with respect to the predefined distance measure. PCA is an unsupervised method that can project the metabolomic dataset into lower dimensional space or principal components (PC) revealing inherent data structure, and providing a simple visualization of dataset. However, PCA has several shortcomings. Most significantly, PCA does not have an associated probabilistic model, which makes assessing the fit of PCA to the data difficult and limits the potential to extend the scope of application of PCA. Furthermore, PCA can fail to reveal underlying groups of subjects in the data, therefore providing a spurious view of the underlying data structure [141]. To get around the problem of probabilistic model for PCA a probabilistic PCA (PPCA) was proposed [142]. Among the methods supervised discriminant analysis (DA) is the most widely used in metabolomics. In this case, particular care must be taken to ensure an appropriately large number of observations in order to reduce the possibility of generating false positives [143].

5 Conclusion

This review has presented biological samples, and the sampling and analytical methods that are used in different strategies of metabolomics. The objectives of this study were to highlight the several factors that are considered important for achieving analysis of the metabolome. The choice of sample depends on the analysis goal; for instance, cerebrospinal fluid is used in the diagnosis of CNS pathologies, and the relevance of this phase is fundamental to the quality of the

final results. In the general preparation of biological samples of metabolic fingerprinting studies, the solvent dilution of the sample is enough for analysis. Alternatively, when targeted metabolic or metabolic profiling is performed, solid phase extraction or solid phase microextraction is very useful. However, analytical methods, such as GC/MS, LC/MS or GCxGC/MS, that are used in metabolome studies exhibit the sensitivity and precision required for the identification of new biomarkers.

References

1. E. N. Tiziana Pacchiarotta, Alegria Carrasco-Pancorbo, Andre´ M. Deelder, Oleg A. Mayboroda, *J. Biomol. Tech*, 2010, **21**, 205 - 213.
2. T. R. Sana, K. Waddell and S. M. Fischer, *J. Chromatogr. B*, 2008, **871**, 314-321.
3. R. Kaddurah-Daouk, B. S. Kristal and R. M. Weinshilboum, *Annu. Rev. Pharmacol*, 2008, **48**, 653-683.
4. A. E. Garrod, *Inborn errors of metabolism*, Oxford University Press, London, 1923.
5. M. P. Quinones and R. Kaddurah-Daouk, *Neurobiol. Dis.*, 2009, **35**, 165-176.
6. P. Goldsmith, H. Fenton, G. Morris-Stiff, N. Ahmad, J. Fisher and K. R. Prasad, *J. Surg. Res.*, 2010, **160**, 122-132.
7. D. B. Kell, *Curr. Opin. Microbiol.*, 2004, **7**, 296-307.
8. M. Bicikova, M. Hill, D. Ripova, P. Mohr and R. Hampl, *J. Steroid. Biochem.*, 2013, **133**, 77-83.
9. S. F. Graham, O. P. Chevallier, D. Roberts, C. Hoelscher, C. T. Elliott and B. D. Green, *Anal. Chem.*, 2013, **85**, 1803-1811.

10. A. N. Fonteh, R. J. Harrington, A. Tsai, P. Liao and M. G. Harrington, *Amino Acids*, 2007, **32**, 213-224.

11. G. D. Lewis, A. Asnani and R. E. Gerszten, *J Am. coll. Cardiol.*, 2008, **52**, 117-123.

12. Y. Zheng, B. Yu, D. Alexander, T. H. Mosley, G. Heiss, J. A. Nettleton and E. Boerwinkle, *Hypertension*, 2013, **62**, 398-403.

13. M. M. Koek, F. M. van der Kloet, R. Kleemann, T. Kooistra, E. R. Verheij and T. Hankemeier, *Metabolomics*, 2011, **7**, 1-14.

14. H. Tsutsui, T. Maeda, J. Z. Min, S. Inagaki, T. Higashi, Y. Kagawa and T. Toyo'oka, *Clin. Chim. Acta*, 2011, **412**, 861-872.

15. O. Abu Aboud and R. H. Weiss, *Clin. Chem.*, 2013, **59**, 138-146.

16. M. Y. Fong, J. McDunn and S. S. Kakar, *Plos One*, 2011, **6**, e19963.

17. C. Yang, A. D. Richardson, A. Osterman and J. W. Smith, *Metabolomics*, 2008, **4**, 13-29.

18. P. M. Joyner, R. M. Matheke, L. M. Smith and R. H. Cichewicz, *J. Proteome Res.*, 2010, **9**, 404-412.

19. D. C. Wang, C. H. Sun, L. Y. Liu, X. H. Sun, X. W. Jin, W. L. Song, X. Q. Liu and X. L. Wan, *Neurobiol. Aging*, 2012, **33**, 1057-1066.

20. C. D. Davis and J. Milner, *Mutat. Res. Fund. Mol. M.*, 2004, **551**, 51-64.

21. M. Oresic, *Nutr. Metab. Cardiovas.*, 2009, **19**, 816-824.

22. M. Kato, Y. Onda, M. Sekimoto, M. Degawa and T. Toyo'oka, *J. Chromatogr. A*, 2009, **1216**, 8277-8282.

23. D. M. Jacobs, N. Deltimple, E. van Velzen, F. A. van Dorsten, M. Bingham, E. E. Vaughan and J. van Duynhoven, *Nmr Biomed.*, 2008, **21**, 615-626.

24. Q. Teng, W. L. Huang, T. W. Collette, D. R. Ekman and C. Tan, *Metabolomics*, 2009, **5**, 199-208.

25. S. G. Villas-Boas, J. Hojer-Pedersen, M. Akesson, J. Smedsgaard and J. Nielsen, *Yeast*, 2005, **22**, 1155-1169.

26. J.-F. Xia, Q.-L. Liang, P. Hu, Y.-M. Wang and G.-A. Luo, *Chinese J. Anal. Chem.*, 2009, **37**, 136-143.

27. A. Saghatelian and B. F. Cravatt, *Curr. Opin. Chem. Biol.*, 2005, **9**, 62-68.

28. M. Stitt and A. R. Fernie, *Curr. Opin. Biotech.*, 2003, **14**, 136-144.

29. D. S. Wishart, M. J. Lewis, J. A. Morrissey, M. D. Flegel, K. Jeroncic, Y. P. Xiong, D. Cheng, R. Eisner, B. Gautam, D. Tzur, S. Sawhney, F. Bamforth, R. Greiner and L. Li, *J. Chromatogr. B*, 2008, **871**, 164-173.

30. P. Kiefer, J.-C. Portais and J. A. Vorholt, *Anal. Biochem.*, 2008, **382**, 94-100.

31. K. K. Pasikanti, P. C. Ho and E. C. Y. Chan, *J. Chromatogr. B*, 2008, **871**, 202-211.

32. I. Garcia-Perez, M. Vallejo, A. Garcia, C. Legido-Quigley and C. Barbas, *J. Chromatogr. A*, 2008, **1204**, 130-139.

33. M. R. N. Monton and T. Soga, *J. Chromatogr. A*, 2007, **1168**, 237-246.

34. D. S. Wishart, M. J. Lewis, J. A. Morrissey, M. D. Flegel, K. Jeroncic, Y. Xiong, D. Cheng, R. Eisner, B. Gautam, D. Tzur, S. Sawhney, F. Bamforth, R. Greiner and L. Li, *J. Chromatogr. B*, 2008, **871**, 164-173.

35. E. Neyraud, M. Tremblay-Franco, S. Gregoire, O. Berdeaux and C. Canlet, *Metabolomics*, 2013, **9**, 213-222.

36. T. Nishino, A. Yachie-Kinoshita, A. Hirayama, T. Soga, M. Suematsu and M. Tomita, *J. Biotechnol.*, 2009, **144**, 212-223.

37. X. Duportet, R. B. M. Aggio, S. Carneiro and S. G. Villas-Boas, *Metabolomics*, 2012, **8**, 410-421.

38. G. A. Theodoridis, H. G. Gika, E. J. Want and I. D. Wilson, *Anal. Chim. Acta*, 2012, **711**, 7-16.

39. W. Li, S. Luo, H. T. Smith and F. L. S. Tse, *J. Chromatogr. B*, 2010, **878**, 583-589.

40. T. Rosenling, M. P. Stoop, A. Smolinska, B. Muilwijk, L. Coulier, S. Shi, A. Dane, C. Christin, F. Suits, P. L. Horvatovich, S. S. Wijmenga, L. M. C. Buydens, R. Vreeken, T. Hankemeier, A. J. van Gool, T. M. Luider and R. Bischoff, *Clin. Chem.*, 2011, **57**, 1703-1711.

41. S. Fox, in *Human physiology*. McGraw-Hill, Boston 1999, pp 364–367.

42. Z. Ying, A. Jiye, G. Wang, H. Qing, Y. Bei, Z. Weibin, G. Shenghua, L. Linsheng, R. Hongcan, R. Meiting and S. Longsheng, *J. Chromatogr. B*, 2009, **877**, 1751-1757.

43. J. A, J. Trygg, J. Gullberg, A. I. Johansson, P. Jonsson, H. Antti, S. L. Marklund and T. Moritz, *Anal. Chem.*, 2005, **77**, 8086-8094.

44. D. Vuckovic, X. Zhang, E. Cudjoe and J. Pawliszyn, *J. Chromatogr. A*, 2010, **1217**, 4041-4060.

45. T. Soga, M. Sugimoto, M. Honma, M. Mori, K. Igarashi, K. Kashikura, S. Ikeda, A. Hirayama, T. Yamamoto, H. Yoshida, M. Otsuka, S. Tsuji, Y. Yatomi, T. Sakuragawa, H. Watanabe, K. Nihei, T. Saito, S. Kawata, H. Suzuki, M. Tomita and M. Suematsu, *J. Hepatol.*, 2011, **55**, 896-905.

46. I. Beheshti, L. M. Wessels and J. H. Eckfeldt, *Clin. Chem.*, 1994, **40**, 2088-2092.

47. M. J. N. Paiva, H. C. Menezes, P. P. Christo, R. R. Resende and Z. L. Cardeal, *J. Chromatogr. B*, 2013, **931**, 97-102.

48. J. Pinto, M. R. M. Domingues, E. Galhano, C. Pita, M. D. Almeida, I. M. Carreira and A. M. Gil, *Analyst*, 2014, **139**, 1168-1177.

49. P. A. Demirev, *Anal. Chem.*, 2013, **85**, 779-789.

50. P. Y. Yin, A. Peter, H. Franken, X. J. Zhao, S. S. Neukamm, L. Rosenbaum, M. Lucio, A. Zell, H. U. Haring, G. W. Xu and R. Lehmann, *Clin. Chem.*, 2013, **59**, 833-845.

51. D. B. Kell, M. Brown, H. M. Davey, W. B. Dunn, I. Spasic and S. G. Oliver, *Nat. Rev. Microbiol.*, 2005, **3**, 557-565.

52. D. Ryan, K. Robards, P. D. Prenzler and M. Kendall, *Anal. Chim. Acta*, 2011, **684**, 17-29.

53. M. A. Fernández-Peralbo and M. D. Luque de Castro, *TrAC Trends in Anal. Chem.*, 2012, **41**, 75-85.

54. E. M. Lenz, J. Bright, I. D. Wilson, S. R. Morgan and A. F. P. Nash, *J. Pharmaceut. Biomed.*, 2003, **33**, 1103-1115.

55. L. G. Rasmussen, H. Winning, F. Savorani, H. Toft, T. M. Larsen, L. O. Dragsted, A. Astrup and S. B. Engelsen, *Nutrients*, 2012, **4**, 112-131.

56. C. Napoli, N. Sperandio, R. T. Lawlor, A. Scarpa, H. Molinari and M. Assfalg, *J. Proteome Res.*, 2012, **11**, 1274-1283.

57. L. C. Soliman, Y. Hui, A. K. Hewavitharana and D. D. Y. Chen, *J. Chromatogr. A*, 2012, **1267**, 162-169.

58. Y.-Z. Zhou, X.-Y. Zheng, X.-J. Liu, Z.-Y. Li, X.-X. Gao, H.-F. Sun, L.-Z. Zhang, X.-Q. Guo, G.-H. Du and X.-M. Qin, *Chromatographia*, 2012, **75**, 157-164.

59. M. C. Walsh, L. Brennan, E. Pujos-Guillot, J.-L. Sebedio, A. Scalbert, A. Fagan, D. G. Higgins and M. J. Gibney, *Am. J. Clin. Nutr.*, 2007, **86**, 1687-1693.

60. M. P. Stoop, L. Coulier, T. Rosenling, S. Shi, A. M. Smolinska, L. Buydens, K. Ampt, C. Stingl, A. Dane, B. Muilwijk, R. L. Luitwieler, P. Smitt, R. Q. Hintzen, R. Bischoff, S. S. Wijmenga, T. Hankemeier, A. J. van Gool and T. M. Luider, *Mol. Cell. Prot.*, 2010, **9**, 2063-2075.

61. A. Wuolikainen, P. M. Andersen, T. Moritz, S. L. Marklund and H. Antti, *Mol. Genet. Metab.*, 2012, **105**, 472-478.

62. B. Toczylowska, M. Chalimoniuk, M. Wodowska and E. Mayzner-Zawadzka, *Brain Res.*, 2006, **1104**, 183-189.

63. B. Crews, W. R. Wikoff, G. J. Patti, H.-K. Woo, E. Kalisiak, J. Heideker and G. Siuzdak, *Anal. Chem.*, 2009, **81**, 8538-8544.

64. E. Aasebo, J. A. Opsahl, Y. Bjorlykke, K. M. Myhr, A. C. Kroksveen and F. S. Berven, *Plos One*, 2014, **9**, e90429.

65. A. Anesi, M. Rondanelli and G. M. d'Eril, *Clin. Chem.*, 1998, **44**, 2359-2360.

66. T. Rosenling, C. L. Slim, C. Christin, L. Coulier, S. Shi, M. P. Stoop, J. Bosman, F. Suits, P. L. Horvatovich, N. Stockhofe-Zurwieden, R. Vreeken, T. Hankemeier, A. J. van Gool, T. M. Luider and R. Bischoff, *J. Proteome Res.*, 2009, **8**, 5511-5522.

67. J. Zheng, R. A. Dixon and L. Li, *Anal. Chem.*, 2012, **84**, 10802-10811.

68. A. Zhang, H. Sun, P. Wang, Y. Han and X. Wang, *J. Proteomics*, 2012, **75**, 1079-1088.

69. N. Spielmann and D. T. Wong, *Oral Diseases*, 2011, **17**, 345-354.

70. H. C. Bertram, N. Eggers and N. Eller, *Anal. Chem.*, 2009, **81**, 9188-9193.

71. M. Aimetti, S. Cacciatore, A. Graziano and L. Tenori, *Metabolomics*, 2012, **8**, 465-474.

72. M. Sofia, M. Maniscalco, G. de Laurentiis, D. Paris, D. Melck and A. Motta, *J. Biomed. Biotechnol.*, 2011, 1-7.

73. P. Montuschi, D. Paris, D. Melck, V. Lucidi, G. Ciabattoni, V. Raia, C. Calabrese, A. Bush, P. J. Barnes and A. Motta, *Thorax*, 2012, **67**, 222-228.

74. U. P. Kodavanti, in *Biomarkers in Toxicology*, ed. R. C. Gupta, Academic Press, Boston, 2014, pp. 217-239.

75. M. Drent and J. A. Jacobs, in *Encyclopedia of Respiratory Medicine*, eds. G. J. Laurent and S. D. Shapiro, Academic Press, Oxford, 2006, pp. 275-284.

76. W. E. Ho, Y. J. Xu, F. Xu, C. Cheng, H. Y. Peh, S. R. Tannenbaum, W. S. Wong and C. N. Ong, *Am. J. Resp. Cell. Mol.*, 2013, **48**, 204-211.

77. C. R. Evans, A. Karnovsky, M. A. Kovach, T. J. Standiford, C. F. Burant and K. A. Stringer, *J. Proteome Res.*, 2014, **13**, 640-649.

78. B. Barroso and R. Bischoff, *J. Chromatogr. B*, 2005, **814**, 21-28.

79. L.-C. Fan, H.-W. Lu, K.-B. Cheng, H.-P. Li and J.-F. Xu, *Plos One*, 2013, **8,** e73099.

80. E. Kipnis, K. Hansen, T. Sawa, K. Moriyama, A. Zurawel, A. Ishizaka and J. Wiener-Kronish, *CHEST Journal*, 2008, **134**, 338-345.

81. A. J. Henderson, *Arch. Dis. Child.*, 1994, **70**, 167-169.

82. R. K. Rai, A. Azim, N. Sinha, J. N. Sahoo, C. Singh, A. Ahmed, S. Saigal, A. K. Baronia, D. Gupta, M. Gurjar, B. Poddar and R. K. Singh, *Metabolomics*, 2013, **9**, 667-676.

83. H. Tajima, T. Yoshida, A. Ohnuma, T. Fukuyama, K. Hayashi, S. Yamaguchi, R. Ohtsuka, J. Sasaki, M. Tomita, S. Kojima, N. Takahashi, Y. Kashimoto, M. Kuwahara, M. Takeda, T. Kosaka, N. Nakashima and T. Harada, *Toxicology*, 2010, **267**, 118-124.

84. W. J. Griffiths and J. Sjovall, *Biochem. Bioph. Res. Co.*, 2010, **396**, 80-84.

85. S. R. Hao, J. J. Xin, J. S. Lian, Q. Xie, D. Y. Chen, Y. Z. Guo, Y. F. Lu, G. P. Sheng, W. Xu, J. R. Huang and L. J. Li, *Metabolomics*, 2011, **7**, 400-412.

86. C. C. Huang, W. T. Lin, F. L. Hsu, P. W. Tsai and C. C. Hou, *Eur. J. Appl. Physiol.*, 2010, **108**, 557-566.

87. S. Hao, J. Xin, J. Lian, Q. Xie, D. Chen, Y. Guo, Y. Lu, G. Sheng, W. Xu, J. Huang and L. Li, *Metabolomics*, 2011, **7**, 400-412.

88. C.-C. Huang, W.-T. Lin, F.-L. Hsu, P.-W. Tsai and C.-C. Hou, *Eur. J. Appl. Physiol.*, 2010, **108**, 557-566.

89. S. Gu, A. Jiye, G. Wang, W. Zha, B. Yan, Y. Zhang, H. Ren, B. Cao and L. Liu, *Biomed. Chromatogr.*, 2010, **24**, 245-252.

90. K. Suhre, S.-Y. Shin, A.-K. Petersen, R. P. Mohney, D. Meredith, B. Waegele, E. Altmaier, P. Deloukas, J. Erdmann, E. Grundberg, C. J.

Hammond, M. Hrabe de Angelis, G. Kastenmueller, A. Koettgen, F. Kronenberg, M. Mangino, C. Meisinger, T. Meitinger, H.-W. Mewes, M. V. Milburn, C. Prehn, J. Raffler, J. S. Ried, W. Roemisch-Margl, N. J. Samani, K. S. Small, H. E. Wichmann, G. Zhai, T. Illig, T. D. Spector, J. Adamski, N. Soranzo, C. Gieger and CardioGram, *Nature*, 2011, **477**, 54-U60.

91. V. Mapelli, L. Olsson and J. Nielsen, *Trends Biotechnol.*, 2008, **26**, 490-497.

92. G. Theodoridis, H. G. Gika and I. D. Wilson, *Trac-Trends Anal. Chem.*, 2008, **27**, 251-260.

93. S. Moco, R. J. Bino, R. C. H. De Vos and J. Vervoort, *Trac-Trends Anal. Chem.*, 2007, **26**, 855-866.

94. A. P. H. Danielsson, T. Moritz, H. Mulder and P. Spegel, *Anal. Biochem.*, 2010, **404**, 30-39.

95. M. J. van der Werf, K. M. Overkamp, B. Muilwijk, L. Coulier and T. Hankemeier, *Anal. Biochem.*, 2007, **370**, 17-25.

96. S. N. Martin Volmer, Sebastian Scholz, Tobias Thute, Heino Buntemeyer, Thomas Noll, *Biotechnol. Lett*, 2010, **33**, 495-502.

97. S. Dietmair, N. E. Timmins, P. P. Gray, L. K. Nielsen and J. O. Kromer, *Anal. Biochem.*, 2010, **404**, 155-164.

98. C. A. Sellick, R. Hansen, A. R. Maqsood, W. B. Dunn, G. M. Stephens, R. Goodacre and A. J. Dickson, *Anal. Chem.*, 2009, **81**, 174-183.

99. M. M. Koek, R. H. Jellema, J. van der Greef, A. C. Tas and T. Hankemeier, *Metabolomics*, 2011, **7**, 307-328.

100. N. Ileri and P. Calik, *Biotechnol. Progr.*, 2006, **22**, 411-419.

101. K. Sulek, H. L. Frandsen, J. Smedsgaard, T. H. Skov, A. Wilcks and T. R. Licht, *Metabolomics*, 2012, **8**, 244-252.

102. S. Tulipani, R. Llorach, M. Urpi-Sarda and C. Andres-Lacueva, *Anal. Chem.*, 2013, **85**, 341-348.

103. A. E. Lewis-Stanislaus and L. Li, *J Am. Soc. Mass. Spectr.*, 2010, **21**, 2105-2116.

104. A. M. Flores-Valverde and E. M. Hill, *Anal. Chem.*, 2008, **80**, 8771-8779.

105. S. Carraro, S. Rezzi, F. Reniero, K. Heberger, G. Giordano, S. Zanconato, C. Guillou and E. Baraldi, *Am. J. Resp. Crit. Care*, 2007, **175**, 986-990.

106. S. Snowden, S. E. Dahlen and C. E. Wheelock, *Bioanalysis*, 2012, **4**, 2265-2290.

107. D. Vuckovic, *Analytical and BioAnal. Chem.*, 2012, **403**, 1523-1548.

108. D. Vuckovic, *Bioanalysis*, 2011, **3**, 1305-1308.

109. D. Vuckovic, S. Risticevic and J. Pawliszyn, *Angew. Chem. Int. Edit.*, 2011, **50**, 5618-5628.

110. C. Hamani, M. Diwan, C. E. Macedo, M. L. Brandao, J. Shumake, F. Gonzalez-Lima, R. Raymond, A. M. Lozano, P. J. Fletcher and J. N. Nobrega, *Biol. Psychiat.*, 2010, **67**, 117-124.

111. V. Sharma, J. Fan, A. Jerath, K. S. Pang, B. Bojko, J. Pawliszyn, J. M. Karski, T. Yau, S. McCluskey and M. Wasowicz, *Anaesthesia*, 2012, **67**, 1242-1250.

112. D. Vuckovic, I. de Lannoy, B. Gien, R. E. Shirey, L. M. Sidisky, S. Dutta and J. Pawliszyn, *Angew. Chem. Int. Edit.*, 2011, **50**, 5344-5348.

113. B. Bojko, E. Cudjoe, M. Wasowicz and J. Pawliszyn, *Trac-Trends Anal. Chem.*, 2011, **30**, 1505-1512.

114. 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 and R. Goodacre, *Nat. Protocols*, 2011, **6**, 1060-1083.

115. E. Dixon, C. Clubb, S. Pittman, L. Ammann, Z. Rasheed, N. Kazmi, A. Keshavarzian, P. Gillevet, H. Rangwala and R. D. Couch, *Plos One*, 2011, **6,** e18471.

116. D. Vuckovic and J. Pawliszyn, *Anal. Chem.*, 2011, **83**, 1944-1954.

117. S. Tiziani, A. H. Einwas, A. Lodi, C. Ludwig, C. M. Bunce, M. R. Viant and U. L. Gunther, *Anal. Biochem.*, 2008, **377**, 16-23.

118. M. Jupin, P. J. Michiels, F. C. Girard, M. Spraul and S. S. Wijmenga, *J. Magn. Reson.*, 2013, **228**, 81-94.

119. S. Stolzenburg, M. B. Lauridsen, H. Toft, P. A. Zalloua and D. Baunsgaard, *Metabolomics*, 2011, **7**, 270-277.

120. R. Lee and P. Britz-McKibbin, *Anal. Chem.*, 2009, **81**, 7047-7056.

121. R. Guthrie and A. Susi, *Pediatrics*, 1963, **32**, 338-343.

122. B. G. Keevil, *Clin. Biochem.*, 2011, **44**, 110-118.

123. J. Deglon, A. Thomas, P. Mangin and C. Staub, *Analytical and BioAnal. Chem.*, 2012, **402**, 2485-2498.

124. P. Denniff and N. Spooner, *Bioanalysis*, 2010, **2**, 1817-1822.

125. F. M. Li, J. Zulkoski, D. Fast and S. Michael, *Bioanalysis*, 2011, **3**, 2321-2333.

126. Y. Y. Li, J. Henion, R. Abbott and P. Wang, *Bioanalysis*, 2012, **4**, 1445-1456.

127. R. Steuer, *Briefings in Bioinformatics*, 2006, **7**, 151-158.

128. M. Hilario, A. Kalousis, C. Pellegrini and M. Muller, *Mass. Spectrom. Rev.*, 2006, **25**, 409-449.

129. D. Radulovic, S. Jelveh, S. Ryu, T. G. Hamilton, E. Foss, Y. Y. Mao and A. Emili, *Mol. Cell. Prot.*, 2004, **3**, 984-997.

130. R. Baran, H. Kochi, N. Saito, M. Suematsu, T. Soga, T. Nishioka, M. Robert and M. Tomita, *Bmc Bioinformatics*, 2006, **7,** 530-539.

131. Y. Ni, Y. P. Qiu, W. X. Jiang, K. Suttlemyre, M. M. Su, W. C. Zhang, W. Jia and X. X. Du, *Anal. Chem.*, 2012, **84**, 6619-6629.

132. A. M. van Nederkassel, M. Daszykowski, P. H. C. Eilers and Y. V. Heyden, *J. Chromatogr. A*, 2006, **1118**, 199-210.

133. Y. B. Zheng, Z. M. Zhang, Y. Z. Liang, D. J. Zhan, J. H. Huang, Y. H. Yun and H. L. Xie, *J. Chromatogr. A*, 2013, **1286**, 175-182.

134. M. Katajamaa and M. Oresic, *J. Chromatogr. A*, 2007, **1158**, 318-328.

135. S. M. Kohl, M. S. Klein, J. Hochrein, P. J. Oefner, R. Spang and W. Gronwald, *Metabolomics*, 2012, **8**, S146-S160.

136. C. D. Wijetunge, Z. P. Li, I. Saeed, J. Bowne, A. L. Hsu, U. Roessner, A. Bacic and S. K. Halgamuge, *Metabolomics*, 2013, **9**, 1311-1320.

137. K. Dettmer, P. A. Aronov and B. D. Hammock, *Mass Spectrom. Rev.*, 2007, **26**, 51-78.

138. A. Krastanov, *Biotechnol. Biotec. Eq.*, 2010, **24**, 1537-1543.

139. Y. Sawada, K. Akiyama, A. Sakata, A. Kuwahara, H. Otsuki, T. Sakurai, K. Saito and M. Y. Hirai, *Plant Cell Physiol.*, 2009, **50**, 37-47.

140. W. C. Chang, *J. Roy. Stat. Soc. C-App.*, 1983, **32**, 267-275.

141. G. Nyamundanda, L. Brennan and I. C. Gormley, *Bmc Bioinformatics*, 2010, **11**, 571-581.

142. M. E. Tipping and C. M. Bishop, *Neural Comput.*, 1999, **11**, 443-482.

143. D. Nagrath, C. Caneba, T. Karedath and N. Bellance, *BBA-Bioenergetics*, 2011, **1807**, 650-663.

Table 1. Analytical strategies for metabolic research

Criteria	Global metabolomics	Metabolic profiling	Fingerprinting	Target Metabolic
Purpose	Analysis of broad range of metabolites	describe metabolic patterns for a group of metabolites	screening of samples to provide classification	analysis of one or few target metabolites
Selectivity of metabolites	No	Yes	No	Yes
Specific metabolic pathway	No	Yes	No	Yes
Biologycal samples	cells, body fluids, tissues	cells, body fluids, tissues	cells, body fluids, tissues	cells, body fluids, tissues
Sample preparation	generic	selective	generic	selective
Analysis	Qualitative and quantitative	Qualitative and quantitative	Qualitative	Qualitative and quantitative
analytical methods	GC/MS, LC/MS, CE/MS, NMR	GC/MS, LC/MS, CE/MS, NMR	GC/MS, LC/MS, CE/MS, NMR	GC/MS, LC/MS, CE/MS, NMR
Relative time required	long	long	short	long

Table 2. Methods used to metabolome analysis

Reference	Compounds	Matrix	Sample preparation	Method
Flores-Valverde, 2008	steroids, glucocortoids, and progestagens	Ovaries, testes and liver	SPE	UPLC-TOF MS
WANG, D. C. et al. 2012	Serum fatty	Serum	LLE	GC-MS
Paiva et al 2013	Amino acids	CSF	Metanol (−20 °C)	GC-MS
Wu et al 2009	Carbohydrates, amino acids, Fatty acid	urine	Metanol	GC-MS
Zheng et al 2012	Amino acids	salive	acetone (−20 °C)	LC−FTICR-MS.
Uhl et al 2011	glycerophospholipid	Plasma	methanol	LC-MS/MS
Snyder et al 2010	L-β-Methylamino-alanine	human tissue	LLE	GC × GC-ToFMS
Rocha et al 2011	hydrocarbons, amines, amides, esters, ketones, aldehydes, alcohols, carboxylic acids, ethers, nitriles, thiols, terpenoids, and heterocyclic compounds	human urine	HS-SPME	GC × GC-ToFMS
Caldeira et al. 2012	Cycloalkanes, Aromatic aldehyde, Ketones	Exhaled breath	HS-SPME	GC × GC-ToFMS