

The emergence of metabolomics as a key discipline in the drug discovery process

Marianne Fillet¹ and Michel Frédérich^{2,*}

¹Laboratory for the Analysis of Medicines, Center for Interdisciplinary Research on Medicines (CIRM), University of Liège, Liège, Belgium.

²Laboratory of Pharmacognosy, Center for Interdisciplinary Research on Medicines (CIRM), University of Liège, Liège, Belgium.

* Corresponding author

Abstract

Metabolomics is a recent science that could be defined as the comprehensive qualitative and quantitative analysis of all small molecular weight compounds present in a cell, organ (including biofluids) or organism at a specific time point. More and more applications have been found these last years to metabolomics in the pharmaceutical field. Specifically in the drug discovery process, metabolomics open new perspectives, in new targets identification, in toxicological studies and in bioactive natural products discovery. The challenge in metabolomics is to find a technological approach allowing the reproducible identification and quantitation of as much metabolites as possible. In this context, mass spectrometry and NMR are emerging as key and complementary technologies.

Introduction

Metabolomics is the newest of the “omics” sciences, part of “functional genomics” with genomics, transcriptomics and proteomics (Figure 1). As defined by Fiehn in 2001, metabolomics could be considered as the “comprehensive qualitative and quantitative analysis of all metabolites (= small molecular weight compounds = metabolome) present in a cell, an organ or an organism at a specific time point” [1]. The compounds generally considered as metabolites are molecules with a molecular weight inferior to 2000 Da. It could be primary metabolites such as sugars, amino acids, organic acids, fatty acids, energetic metabolites, nucleotides, urea metabolites,... In plants or microorganisms, secondary metabolites presenting a very huge structural diversity could also be found. A metabolome, and more particularly the human metabolome,

consists in a mixt of both endogenous and exogenous compounds [2]. Actually, several thousands of human metabolites have been described and could be accessed freely in the “Human Metabolome Database” which offers detailed informations about around 40,000 endogenous and exogenous metabolites [2]. Up to 200 000 metabolites are estimated to occur in the plant kingdom [1].

The study of metabolites is in reality very old, and saccharose, isolated for the first time, from beets, in 1747 by Marggraf (but its structure only determined at the end of the XIXth Century) could be considered as one of the first metabolite isolated [3]. Even if metabolites are studied in various organisms since numerous years, ‘metabolomics’ in its actual acceptation is very recent. The first paper that could be considered as a real “metabolomics” analysis is the paper from Arthur B. Robinson and Linus Pauling in 1971, dedicated to the analysis of urine vapour and breath metabolites using gas chromatography. In this paper, 250 metabolites were analysed and identified in breath, and 280 in urine vapour [4]. Nevertheless, the first mention of the term “metabolome” in the literature was only in 1998 (as sorted by bibliographic engines) in a paper of S. Oliver about the functional analysis of the yeast genome [5]. The word “metabolomics” was then only used in about 5 scientific publications in 2001, but already in 2004 was created the “metabolomics society”¹. Since then, there was an exponential increase in the number of publications per year about metabolome, metabolomic or metabolomics, reaching a number of around 2168 publications in 2013 (Figure 2).

In 1999, the term “metabonomics” was also introduced by Nicholson and colleagues in a paper where it was defined as “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to physiopathological stimuli or genetic modifications”[6]. Nowadays, the term metabolomic(s) is the most currently used, as it could clearly be seen in Figure 2. The concept metabonomics is nevertheless still used, particularly for the study of dynamic changes in the metabolome, mainly in the toxicological field.

Fields of application

¹ <http://metabolomicssociety.org/>

In all cases, a metabolomic analysis consists in the identification of modifications in a metabolome, modifications that could be related to internal or external causes. Internal causes are generally correlated with genetic variations, i.e. the study of two genetically close organisms (taxonomic studies) or study of the influence of a gene modification on the metabolome. External causes of modification for the metabolome could be in a cell or an organism the administration of a drug or a toxic, or pathology. An external cause could also be different environmental conditions (including lifestyle and alimentation, or culture conditions for plants), different conservation conditions or different conditions of extraction. Ideally, a complete metabolomics study should identify the metabolites that are modified and this could lead to the comprehension of a metabolic pathway, to the identification of bioactive compounds or to the determination of the mode of action of a drug. The fields of application of metabolomics are then huge, ranging from food analysis and nutrition [7,8] (where the techniques are also routinely used by the food industry) to functional genomics applications. More particularly in the pharmaceutical field, metabolomic techniques and applications find scope from drug discovery to clinical development of new medicines. This includes target identification and validation, elucidation of the mechanisms of action of drugs, lead identification [9,10], identification of drug efficacy or pathology biomarkers [11–13], toxicological analysis by identification of safety biomarkers [14–16] but also quality control, particularly of medicinal plants [17,18] and bioactive drug discovery from natural sources [19]. Recently, several reviews have highlighted the potential of metabolomics in the drug discovery process [20–22]. The challenge in metabolomics is then to find an analytical technique or a technological approach, combined with adequate statistical analysis, that allows the detection and the identification of the highest possible number of metabolites.

Technological approaches

As described above, one of the objectives of metabolomics in the pharmaceutical field is to identify and quantify specific metabolic markers susceptible to improve quality control, improve early diagnosis, survey therapeutic outcomes and facilitate the development of novel drugs candidates [23–25].

The methodology relies on differential metabolic expression profiling. The fundamental approach is based on the assumption that the pathology of concern or the therapeutic treatment (i.e. drug or drug candidate) will affect some physiological processes causing changes in the metabolic expression levels. Metabolic generating similar signals in both sample groups are ignored while significantly up- and down regulated metabolites become potential biomarkers. Differential expression profiling requires both a sensitive technology to discern any tiny differences and a high throughput system in order to process large series of samples required to reach statistical significance.

Besides “metabolic fingerprinting” that consists in the comparison of metabolic fingerprints, without identification of particular metabolites, there are two general approaches, target specific and global or non-directed, for metabolomic biomarker discovery. Target-specific, or biology driven, approach frequently uses mass spectrometry and is useful for validation and routine clinical analysis. In that case, tested metabolites are typically selected on the basis of current knowledge of disease mechanism and pathophysiology. Global/non-directed approaches may have more potential for biomarker discovery because they are unbiased. They allow the simultaneous evaluation of hundreds to thousands metabolites in a blind manner without *a priori* increasing the potential number of candidate biomarkers.

One- or two dimensional liquid or gaz chromatography (LC or GC) and capillary electrophoresis (CE) coupled to mass spectrometry (MS), as well as nuclear magnetic resonance (NMR), are regarded as the most powerful tools for establishing fingerprint profiles. Many reports regarding the use of these technologies in the context of metabolomic studies have already been published [26-28]. The separation techniques (i.e. LC, GC and CE) are very versatile approaches due to the multiple choices of stationary and mobile phases, as well as background electrolytes that are available. However that makes them consequently rather complex, time consuming and requires experienced staff.

The earliest researches on metabolite profiling were described by GC-MS. GC-MS is highly selective, relatively inexpensive and provides reproducible analysis with an added advantage of comprehensive structured databases for GC-MS metabolite identification. However, it is not a universal approach since it is limited to volatile

compounds or appropriately derivatized metabolites. The use two-dimensional chromatography further enhances metabolite quantification and gives metabolomics platforms even greater analytic capability [29].

Liquid chromatography-mass spectrometry (LC-MS) is the most widely used technology in metabolomic studies, due to its ability to separate and detect a diverse range of molecules with high sensitivity. Reverse-phase (RP)-LC is generally used for the separation of metabolites of medium to low polarity (i.e. fatty acids, flavonoids, triglycerides...); although normal-phase (NP)-LC and hydrophilic interaction liquid chromatographic (HILIC) modes are better suited for the analysis of highly polar metabolites (i.e. glycols, sugars, amino-acids,...). The combination of HILIC and RPLC coupled with an MS detector expands the number of detected analytes and provides more comprehensive metabolite coverage than use of only RP chromatography. Multidimensional LC is also frequently used; it combines generally two orthogonal LC modes providing a better selectivity. More recently ultra-high pressure (UHPLC) and superficially porous particles substantially increased the peak capacity of liquid chromatography, potentially reducing coelution and enabling equivalent separation in shorter analysis times [30-31].

CE is nowadays a powerful alternative to chromatographic techniques such as GC or HPLC to carry out metabolomic studies. Today CE is increasingly used for several reasons: its high efficacy, versatility and feasibility of incorporating a large number of selectors in the background electrolyte that greatly facilitates method development, rapidity, cheapness (a low consumption of reagents), and minimum environmental impact with respect to HPLC or GC techniques. The high peak efficiency that can be obtained in CE (between 100,000-200,000 plates in a typical capillary) is one of the major advantages [32].

When a high selectivity and sensitivity is required, ultraviolet detection (UV) is not the preferred detection system. Mass spectrometry (MS) detection offers not only an improvement in the sensitivity but also offers the possibility to unambiguously identify an analyte by its mass to charge ratio (m/z). In addition, when MS/MS experiments are performed, structural information can also be obtained from the fragmentation mass spectrum which offers an excellent selectivity.

Traditional MS/MS analysis consists of a first acquisition of a MS survey scan followed by selection of precursor ions by the first analyser for MS/MS fragmentation and analysis by the second. More recently, targeted quantitative MS has been considered for validation step with the emergence of single- and multiple- reaction monitoring (SRM and MRM). The analytical system most often are constituted of a liquid chromatography coupled to an ESI-triple quadrupole mass spectrometer which allows the absolute quantification of one or several metabolites in hundred complex samples. In a single reaction monitoring (SRM), the first and the third analysers act as filters to specifically select predefined m/z values corresponding to the ion and a specific fragment ion of this molecule respectively, whereas the second quadrupole serves as a collision cell. This approach has the advantage to offer the possibility of multiplexing without affecting the sensitivity and selectivity. For metabolic profiling, Q-TOF mass analysers are most commonly used, due to their ability to provide the accurate high mass resolution of MS/MS analysis [33].

Given the abundant and ubiquitous presence of ^1H in most metabolites, $1\text{D } ^1\text{H-NMR}$ spectroscopy is the primary source of NMR-based metabolomics data, used almost exclusively to generate metabolomics data until the recent years. In spite of numerous advantages, which are presented in Table 1 in comparison with MS, the main drawback of metabolite analysis by NMR platforms is the sensitivity that does not match the one of MS based platforms. Nevertheless, the highly reproducible, quantitative results of NMR, the easy sample preparation step required for the method, ensure it will remain extremely valuable in metabolomic studies, especially when combined with MS based techniques [34-36]. NMR is particularly valuable in the natural products area where the method is highly helpful to identify unknown secondary metabolites [37-38].

Vibrational spectroscopy and more particularly Fourier transform infrared spectroscopy (FT-IR) were found particularly interesting for metabolomic studies due to their rapid, high-throughput and non-destructive analysis of a wide range of sample types. The resultant infrared absorbance spectrum is characteristic of the functional groups within the sample and thus constitutes an unique fingerprint [39]. This is a promising analytical tool for metabolomic studies however, it is important to note that the presence of water in a sample may influence the bands at certain specific wave number and that

environmental conditions around the FT-IR instrument can cause variations in the spectra. Moreover a strong expertise in the chemometric analysis of spectra is required. It is worth noting that an intrinsic disadvantage of *in vitro* and *ex vivo* metabolomic experiments is the loss of information regarding the subcellular and cellular localization of the metabolites. Recently, imaging mass spectrometry (IMS) and magnetic resonance imaging (MRI) have been introduced providing information about spatial distribution. This technique allow spatial mapping of lipids, peptides or drugs in fixed samples, without extraction and loss of cellular architecture. This approach is very promising improving our understanding of cellular physiology and enzyme function [40].

All those analytical tools provide complementary metabolic information and it is thus recommended to use them in combination.

Other challenges

Recent interest in the field has yielded a large number of candidate biomarkers in various diseases. However, the small size and poor design of some studies drove validation of these biomarkers quite challenging [41]. Critical steps that should be undertaken to avoid any bias, to maximize reproducibility and detection sensitivity, with the final aim to find relevant, specific and robust biomarkers should be addressed.

All these difficulties can be addressed only by substantial reduction in sample complexity and the application of a rigorous standardization program of the entire analytical process. This involves a fully operational and calibrated instrument and the use of suitable QC samples, similar in nature and complexity to the studied samples. Indeed, the use of relevant quality controls (QCs) is highly recommended and even mandatory in such applications. One of the most challenging aspects in studying body fluids metabolic profiles is the sample preparation which remains a laborious and time consuming task as it usually needs enrichment and purification before sequencing by MS-MS. Ideally, sample preparation methods should be high-throughput, reproducible and preserve quantitative information. In addition, they should avoid loss and dilution of samples as well as to be compatible with downstream analytic techniques. In order to enlarge dynamic range and thus metabolite coverage, several complementary

techniques of preparation can be combined. The use of orthogonal techniques is also mandatory to provide valuable information for the identification of unknown metabolites.

Another important methodological source of artefacts is the data analysis of metabolite profiles. The data pre-processing (calibration, baseline correction, normalization, peak detection and peak alignment) represents a key step for metabolomic analysis [41].

Besides the instrumentation and the methodologies related to chromatography-mass spectrometry analysis, the nature, quality and number of clinical samples to process and statistical data treatment are also key elements to be considered for any metabolomic approaches.

Conclusions

Taking into account herein and previously described recommendations, metabolomic approaches offer very exciting opportunities to discover not only diagnostic, but prognostic and also mechanistic markers for a number of major diseases. This could lead to the identification of new targets and then new drugs with original modes of action. It is also expected that the ability to identify markers of drug toxicity/efficacy will significantly accelerate drug discovery and help define the appropriate clinical plan.

However, some challenges still remain, as for all other “omic” approaches, due in part to the complexity and the wide dynamic range of the samples. Sample fractionation and/or enrichment procedure will certainly be the solution to visualise the deep metabolome.

Conflicts of interests

The authors declare that they have no conflict of interest

References

- 1 Fiehn, O. (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp. Funct. Genomics* 2, 155–68
- 2 Wishart, D.S. *et al.* (2013) HMDB 3.0--The Human Metabolome Database in 2013. *Nucleic Acids Res.* 41, D801–7
- 3 Marggraf, A. (1749) “Experiences chimiques faites dans le dessein de tirer un veritable sucre de diverses plantes, qui croissent dans nos contrées.” In *Histoire de*

l'Académie royale des sciences et des belles lettres de Berlin, Volume 3 (Google eBook) pp. 79–90, Deutsche Akademie der Wissenschaften

- 4 Pauling, L. *et al.* (1971) Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proc. Natl. Acad. Sci.* 68, 2374–2376
- 5 Oliver, S. (1998) Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 16, 373–378
- 6 Nicholson, J.K. *et al.* (1999) “Metabonomics”: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 29, 1181–1189
- 7 Astarita, G. and Langridge, J. (2013) An emerging role for metabolomics in nutrition science. *J. Nutrigenet. Nutrigenomics* 6, 181–200
- 8 Brennan, L. (2014) NMR-based metabolomics: From sample preparation to applications in nutrition research. *Prog. Nucl. Magn. Reson. Spectrosc.* 83, 42–49
- 9 Creek, D.J. and Barrett, M.P. (2014) Determination of antiprotozoal drug mechanisms by metabolomics approaches. *Parasitology* 141, 83–92
- 10 Wu, M. *et al.* (2012) Identification of drug targets by chemogenomic and metabolomic profiling in yeast. *Pharmacogenet. Genomics* 22, 877–86
- 11 Kuc, S. *et al.* (2014) Metabolomics profiling for identification of novel potential markers in early prediction of preeclampsia. *PLoS One* 9, e98540
- 12 Nkuipou-Kenfack, E. *et al.* (2014) Assessment of metabolomic and proteomic biomarkers in detection and prognosis of progression of renal function in chronic kidney disease. *PLoS One* 9, e96955
- 13 Mathé, E.A. *et al.* (2014) Noninvasive urinary metabolomic profiling identifies diagnostic and prognostic markers in lung cancer. *Cancer Res.* 74, 3259–70
- 14 Lin, Z. *et al.* (2014) Short-term oral atrazine exposure alters the plasma metabolome of male C57BL/6 mice and disrupts α -linolenate, tryptophan, tyrosine and other major metabolic pathways. *Toxicology* 326C, 130–141
- 15 Dudka, I. *et al.* (2014) Metabonomic analysis of serum of workers occupationally exposed to arsenic, cadmium and lead for biomarker research: a preliminary study. *Environ. Int.* 68, 71–81
- 16 Liang, Y.-J. *et al.* (2013) A metabonomic investigation of the effects of 60 days exposure of rats to two types of pyrethroid insecticides. *Chem. Biol. Interact.* 206, 302–8

- 17 Frederich, M. *et al.* (2011) Quality assessment of *Polygonum cuspidatum* and *Polygonum multiflorum* by ¹H NMR metabolite fingerprinting and profiling analysis. *Planta Med.* 77, 81–86
- 18 Van der Kooy, F. *et al.* (2009) Quality Control of Herbal Material and Phytopharmaceuticals with MS and NMR Based Metabolic Fingerprinting. *Planta Med.* 75, 763–775
- 19 Yuliana, N.D. *et al.* (2011) Metabolomics for bioactivity assessment of natural products. *Phyther. Res.* 25, 157–169
- 20 Powers, R. (2009) NMR metabolomics and drug discovery. *Magn. Reson. Chem.* 47 Suppl 1, S2–11
- 21 Ross, B. *et al.* (2011) Application of NMR Spectroscopy in Medicinal Chemistry and Drug Discovery. *Curr. Top. Med. Chem.* 11, 93–114
- 22 Powers, R. (2014) The current state of drug discovery and a potential role for NMR metabolomics. *J. Med. Chem.* 57, 5860–70
- 23 Kumar, B. *et al.* (2014) Prakash A, Ruhela RK, Medhi B. Potential of metabolomics in preclinical and clinical drug development. *Pharmacol Rep.* 66,956-963.
- 24 Ahmed, K.A. *et al.* (2014) Applying metabolomics to understand the aggressive phenotype and identify novel therapeutic targets in glioblastoma. *Metabolites.* 27, 740-750.
- 25 Cox, D.G. *et al.* (2014) The utility of metabolomics in natural product and biomarker characterization. *Biochim Biophys Acta.* 1840, 3460-3474.
- 26 Ramautar, R. *et al.* (2014) Recent developments in liquid-phase separation techniques for metabolomics. *Bioanalysis.* 6, 1011-1026.
- 27 Wolfender, J.L. *et al.* (2014) Current approaches and challenges for the metabolite profiling of complex natural extracts. *J Chromatogr A.* S0021-9673(14)01703-8.
- 28 Yin, P. *et al.* (2014) Current state-of-the-art of nontargeted metabolomics based on liquid chromatography-mass spectrometry with special emphasis in clinical applications. *J. Chromatogr. A.* S0021-9673(14)01808-1.
- 29 Han, J. *et al.* (2009) Mass spectrometry-based technologies for high-throughput metabolomics. *Bioanalysis.* 1, 1665-1684.

30 Gika, H.G. *et al* (2014) LC-MS-based holistic metabolic profiling. Problems, limitations, advantages, and future perspectives. *J Chromatogr B Analyt Technol Biomed Life Sci.* 966, 1-6.

31 Tang, D.Q. *et al.* (2014) HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. *Mass Spectrom Rev.* Oct 3. doi: 10.1002 / mas.21445.

32 Ramautar, R. *et al.* (2014) CE-MS for metabolomics: Developments and applications in the period 2012-2014. *Electrophoresis.* 2014 Oct 7. doi: 10.1002/elps.201400388.

33 Buzatto, A.Z. *et al.* (2014) Metabolomic investigation of human diseases biomarkers by CE and LC coupled to MS. *Electrophoresis.* 35, 1285-1307.

34 Powers, R. (2014) The current state of drug discovery and a potential role for NMR metabolomics. *J. Med. Chem.* 57, 5860-5870.

35 Rankin, N.J. *et al.* (2014) The emergence of proton nuclear magnetic resonance metabolomics in the cardiovascular arena as viewed from a clinical perspective. *Atherosclerosis.* 237, 287-300.

36 Bingol, K. *et al.* (2014) Multidimensional approaches to NMR-based metabolomics. *Anal Chem.* 86, 47-57.

Anal Chem. 83, 555-562.

37 Wolfender, J.-L. *et al.* (2014) Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* DOI: 10.1016/j.chroma.2014.10.091

38 Wolfender, J.-L. *et al.* (2013) Plant metabolomics: from holistic data to relevant biomarkers. *Curr. Med. Chem.* 20, 1056-90

39 Schattka, B. *et al.* (2011) Metabolic fingerprinting of biofluids by infrared spectroscopy: modeling and optimization of flow rates for laminar fluid diffusion interface sample preconditioning.

40 Huang, H.M. *et al.* (2014) Pushing CT and MR imaging to the molecular level for studying the "omics": current challenges and advancements. *Biomed Res Int.*, 2014:365812.

41 Naz, S. *et al.* (2014) Method validation strategies involved in non-targeted metabolomics. *J. Chromatogr. A.* 1353, 99-105.

Table 1. Comparison of main advantages and disadvantages of NMR and MS techniques in metabolomics.

| Key points | NMR techniques | MS techniques |
|--------------------------------|---|---|
| Detected metabolites | Universal detector | Universal detector for ionisable molecules |
| Reproducibility | High reproducibility | Lower reproducibility than NMR (especially when combined with LC) |
| Sensitivity | Low sensitivity (μg) | High sensitivity (ng) |
| Resolution | Low resolution but could be improved by 2D-methods | High resolution |
| Quantitativity | Quantitation possible (no standard required) | Quantitation possible (standard required) |
| Sample preparation | Minimal | Sample preparation required |
| Structural informations | Gives structural informations to identify metabolites | Gives structural informations to identify metabolites |
| Other | Non destructive method | Destructive method |

Figure 1:

Place of metabolomics amongst other “omics” sciences.

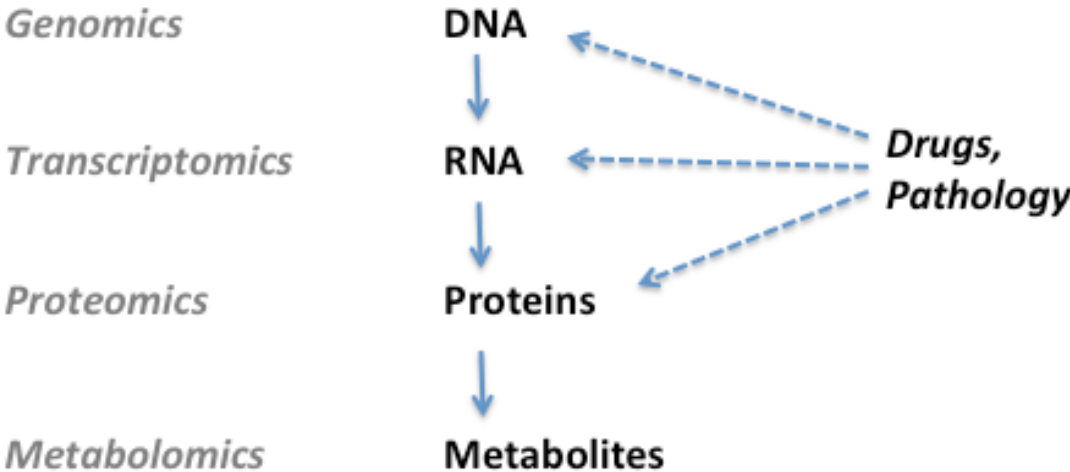


Figure 2 :

Evolution of the number of publications per year using the terms “metabolomics”, or “metabolomics” and “metabonomics” or “metabonomic”. The data were obtained using the search engine Pubmed.

