Metabolomics in Cardiovascular Disease: Towards Clinical Application

Ignasi Barba and David Garcia-Dorado

Institut de Recerca, Àrea del Cor, Hospital Universitari Vall d’Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain

1. Introduction

Metabolomics were initially defined as the global analysis of all metabolites present in a sample (metabolomics) and the analysis of metabolic responses to drugs and diseases (metabonomics) [Nicholson et al., 1999, Fiehn 2002, Weckwerth and Morgenthal 2005]. These two definitions have an historical origin, while metabolomics has its foundations in microbial and plant sciences metabonomics originated in toxicology studies of mammalian systems. Today the two terms are becoming synonyms [Nicholson and Lindon 2008]. Other useful definitions have appeared in the literature [Ellis et al., 2007, Dunn et al., 2011] Table 1. In the particular setting of disease diagnosis it is relevant to talk about metabolic fingerprinting and metabonomics, both strategies aimed mainly at differentiating between two or more sets of data (i.e., disease/healthy) in a quantitative (metabonomics) or qualitative approach (fingerprinting).

As with other ‘omics’ technologies metabolomics aims to measure metabolite dynamics in an unbiased manner; the advent of analytical and statistical methods has made possible to apply metabolic analysis to a wide range of applications. ‘Omic’ technologies allows the simultaneous measure of thousands of parameters from a single sample; together with new statistical methods able to tackle very large databases and systems or network approaches we should be able to extract the knowledge in order to better phenotype disease and help in the diagnostic process. Being able to apply all this newly found knowledge to the clinic is the next big challenge at the beginning of the 21st century.

Genomics, proteomics and metabolomics are closely related approaches as they are all aimed at the quantitative, non-biased study of biological systems although not at the same level. Genomics was the first to appear, the relative chemical simplicity of DNA allowed the implementation of high throughput approaches first; it was followed by proteomics and finally metabolomics. Although there is a continuum between genomics and metabolomics with multiple interactions between genes, proteins and metabolites (Figure 1) the study of the metabolome is relevant because the it is downstream of the genome it is thereby amplified both in theory [Mendes et al., 1996] and in practice [Raamsdonk et al., 2001]. It has also been shown that small changes in the concentration of enzymes have only small effects on the fluxes through metabolic pathways but the changes in metabolite intermediates may be substantial [Cascante et al., 2002].
Term | Definition
---|---
Metabolomics | The nonbiased identification and quantification of all the metabolites present in a biological system
Metabolome | The complete set of low-molecular-weight metabolites present in a biological sample (i.e., biofluid, organism, bacterial community)
Metabolic profiling | Identification and quantification of a selective number of metabolites, usually related to a specific metabolic pathway.
Metabolic fingerprint | Global, high-throughput analysis aimed at sample classification. Also used as a screening tool.
Metabolic footprint | Analysis of the metabolites secreted/excreted by an organism; it may include environmental and growth substances.
Metabolite target analysis | Qualitative and quantitative analysis of one, or several, metabolites related to a specific metabolic reaction.
Metabonomics | Quantitative analysis of metabolites in response of biological perturbations or genetic modification.
Lipidomics | Analysis of all lipids, and the molecules with they interact, and their function within a biological system.
Metabolite flux analysis | Also known as fluxomics. Labeled metabolites are fed into a biosystem and the destination of the label is assessed, usually in a time-dependent manner.
Metabotype | The metabolic phenotype.

Table 1. Definitions of terms used in metabolomics. Modified from [Ellis et al 2007, Goodacre 2007].

Metabolism is downstream of the genome, transcriptome and proteome leading to an amplification of the signal. Only 2766 metabolites are estimated to be derived from man [Duarte et al., 2007], compared with 31896 genes [http://eugenes.org] and $10^6$ different proteins estimated from gene expression, alternative splicing and post-translational modifications [Oh 2004, Goodacre 2007] makes the metabolome relatively simple.

Also, another characteristic of the metabolome is that it is affected by disease and external perturbations, such as age [Fredman et al., 2004, Athertorn et al., 2009], nutrition, disease, gut microflora, toxins and drugs [Bollard et al., 2005]. This favours the study of the metabolome as a relevant representation of the phenotype [Ellis et al., 2007]. While the genome does provide information about the future (for example, the probability of suffering a disease) the metabolome reflects what is actually happening to the system (i.e. is it diseased? How is responding to treatment?).

Metabolism is organized as a scale-free network [Jeong et al., 2000]. These networks are composed of highly interconnected nodes (metabolites) that link together the various parts
of the network making them are very robust against random errors [Albert et al., 2000]. Also, any perturbation will be rapidly and widely spread throughout the network, thus, the detection of a few highly interconnected nodes allows to detect differences originating in distant parts of the metabolic network [Brindle KM 2003] (Figure 2). The way metabolic networks spread perturbations allow for an easy detection but it is difficult to retrace the steps back to the origin unless there is some prior knowledge as in the case of the FANCY approach [Raamsdonk et al., 2001].

The properties of a system are different from those of its individual components; systems biology aims to study all the components of a systems and their interaction; according to Sauer and cols. [2007], complex biological networks are “better addressed by observing, through quantitative measures, multiple components simultaneously, and by rigorous data integration with mathematical models” [Sauer et al., 2007]. Accordingly, the integration of all the information obtained from the ‘omics’ approaches will be essential when trying to understand how biological networks change during disease; a new field of network analysis is emerging in order to tackle the integration of various ‘omics technologies and in silico models [Loscalzo et al., 2007].

Fig. 1. Schematic representation of the interaction between different ‘omes’. The metabolome or metabolic phenotype is both downstream of the genome, proteome and transcriptome and also affected by the environment including disease and drug treatments; it is thus well suited as a diagnostic and treatment follow-up tool.
Fig. 2. Organization of scale free networks, perturbations (marked here as star) in the network are transferred to the central hubs of metabolism (Gray circles) allowing the detection of many different situations even though the metabolites directly responsible for them are beyond the limits of technical detection. Once a change in one or more hubs is detected, knowing the metabolic reactions ultimately responsible for them may not be straightforward unless there is some prior knowledge as in the FANCY approach [Raamsdonk et al., 2001].

1.1 Methodologies

The chemical diversity of metabolites has promoted the use of various analytical methods (Table 2), the most common ones being NMR spectroscopy and Mass spectrometry. There are other analytical methods including infrared, visible, and Raman spectroscopy reviewed elsewhere [Ellis et al., 2007].

Mass spectrometry operates by forming ions (positively or negatively charged) and separating them according to their mass to charge ratio (m/z). It is a high sensitivity method that allows high mass accuracy and resolution thus allowing chemical identification of metabolites; it has
been proposed as the real metabolomic approach [Kell 2004]. Mass spectrometer detectors are usually placed after a chromatography (gas or liquid) step; this introduces a bias as there will be some metabolites that will not be able to reach the detector. If there is prior knowledge of which group of metabolites may be affected the bias may turn to be an advantage and mass spectra the ideal tool to use as it is possible to focus interest in a group of metabolites (Figure 3A).

NMR spectroscopy looks into magnetic properties of atomic nuclei when placed in a magnetic field; the molecular environment (electronic cloud) is able modify the magnetic field thus each nuclei within a molecule will sense a slightly different magnetic field that in turn allow to differentiate between them in a spectra (Figure 3B). A more detailed description of the applications of biological NMR spectroscopy can be found elsewhere [Bothwell and Griffin 2011]

When sample classification, as in the case of disease diagnosis, is the primary objective, metabonomics or metabolic fingerprinting is the desired approach. In this case, NMR spectroscopy is the preferred analytical technique because it is fast, reproducible and cheap on a per-sample basis. NMR spectroscopy is only able to detect a limited number of metabolites but it is able to pick metabolic differences arising from very different parts of the metabolic network; the reason behind this lies on the fact that most of the central hubs of metabolic networks are 'NMR visible', by monitoring these hubs it is possible to indirectly detect a huge number of metabolic alterations [Griffin 2003].

Mass and NMR spectroscopy can be performed on easy to obtain biofluids (serum, urine, saliva, etc) that integrate metabolites from various processes and organs thus providing an integrative metabolic snapshot.

**1.2 Data analysis**

Metabolomic, and other 'omic' technical advances has been mirrored by advances in data processing, statistical methods able to work with large data sets comprising many more variables than subjects under study. The objective of the various pattern recognition techniques is two-fold; on the one hand they aim at classification between two or more groups of samples but also to extract information about which variables (metabolites) are relevant for each classification.

To investigate the innate variation in a dataset, unsupervised techniques such as principal component analysis (PCA) [Hotellin 1933] could be used. In this approach, the sources of variation within the dataset are organized as orthogonal, thus independent, vectors in a descending manner so the first principal component (PC) is the one that explains more variation, followed by the second, third, etc. It is possible that the condition under study, for example a disease, is one of the main sources of metabolic variation; then one of the first PCs will be related to the effect of disease to the metabolic profile and allow for a discrimination between healthy and diseased samples. The great advantage of unsupervised techniques is that they don't require prior knowledge of the condition under study. Also, unsupervised approaches are useful in removing out-layers from the dataset.

However, when unsupervised approaches are not sufficient, supervised techniques may be better suited to answer specific questions. These approaches require the input of the group of each sample in order to create a statistical model that maximizes the differences between groups and minimizes inter-group variation. In the case of supervised models, it is necessary
to test the robustness and predictability of the models produced in order to avoid creating models that will only fit the data used to create them. The ideal situation is when two different datasets are used, one to create the model and an independent one to test it; when the number of samples is limited, internal validation protocols like 'leave-one-out' may be used.

Fig. 3. Examples of UPLC MS-MS spectra (A) from rat urine treated with phenilacelilglutamine where each peak corresponds to an ion and an $^1\text{H}$$^1\text{H}$ COSY NMR spectra of human serum (B) showing the correlation between close (three bond) protons in a molecule.
1.3 Metabolic profiling

Sample classification is one of the main targets of any 'omics' technology, when the groups of samples include a healthy (control) and at least one diseased group (case) the method can be used directly as a diagnostic tool.

Metabolite levels, at least some of them, change between healthy and diseased states and may respond to therapeutic intervention (Figure 4). Perturbations in the metabolic pathways may occur prior to clinical symptoms, those would be the early prognostic markers. It is likely that at the early stage of disease a few metabolites may be slightly altered but not far away from normal values; in this case fingerprinting (Figure 5A), measuring a variety of metabolites, may be the tool to use in order to detect early markers of diagnosis. On the other hand, as a disease progresses it is more likely that one (or more) metabolites alter their concentration significantly, consequently looking for individual metabolites associated to a disease may be the better choice.

![Fig. 4. Metabolite levels during disease evolution. In the pre-clinical stage some metabolites may be altered but likely within the normal deviation of the population, in this setting metabolic profiling of various metabolites may be a good approach towards finding discriminant models; on the other hand in the clinical stages it is possible that one or a few metabolites are highly modified from their normal values in this case looking for biomarkers would be relevant. Some metabolites may change after treatment allowing for unbiased follow-up and pre-treatment metabolomic pattern may also be useful to predict treatment response.](www.intechopen.com)
1.4 Biomarker discovery

On top of sample classification, metabolomic analysis is able to provide information regarding which metabolites are more relevant for a classification. It is thus possible to identify individual metabolites that could be used as biomarkers and use metabolomic analysis can be used as a tool for biomarker discovery. Once a metabolite is identified (Figure 5A) as a biomarker using high throughput metabolomic methodologies it may be detected by classical analytical methods for clinical practice. For example, Sreekumar and cols. [2009] were able to identify sarcosine in urine as potentially important metabolic intermediary of cancer cell invasion and aggressiveness after profiling 1126 metabolites from 262 clinical samples.

Fig. 5. Metabolomics in clinical applications. Panel A shows the result of a discriminant analysis of NMR spectra from serum obtained from patients in the coronary unit and healthy controls. Circles correspond to 95% confidence intervals. Note that dispersion in samples from patients is higher that that of controls. Pattern recognition analysis allows to identify the most relevant variables in an analysis and it is possible to identify the metabolite responsible. Panel B shows a $^1$H-$^{13}$C HSQC NMR spectra; a useful approach to identify metabolites. Processing shown in panel A was obtained using Metaboanalyst software [Xia et al., 2009].

Once a variable in the analysis is found to be relevant in a classification model it can be associated to a metabolite because mass to charge ratio and chemical shift are specific for each metabolite. There are public databases that help metabolite identification http://www.hmdb.ca/. In order to obtain an absolute identification, though, it is important to perform co-resonance or co-elution experiments.
Figure 6 shows a typical metabolomic experimental design aimed at disease diagnosis. Metabolomic profiling and biomarker discovery are not excluding strategies; moreover there is a synergy between them and also with the use of metabolic and pathway information that are not directly associated to diagnostic applications.

Fig. 6. Typical work-flow in a metabolomic study aimed at disease diagnosis. Once the database is obtained, supervised and unsupervised statistical methods are applied to obtain discriminant models that could be applied directly in the clinic or to find biomarkers. The future of metabolomics will be the implementation of information obtained at different stages of the analysis, particularly relating to metabolites and metabolic pathways into systems biology approaches.
2. Metabolomic applications in cardiovascular diseases

2.1 Basic research applications/metabolic phenotyping

Metabolic analysis has been widely used in laboratory research applications. One of the main uses in this field was the metabolic phenotyping of mouse models of cardiovascular diseases; this approach was pioneered by the group of Julian Griffin mainly using models of Duchenne muscular distrophy where they were able to show different metabolic profiles associated with the expression of dystrophin and utrophin in heart muscle [Griffin et al., 2002, Griffin and Des Rosiers 2009]. In a later work the same group applied the FANCY approach (Functional Analysis by Co-responses in Yeast) to mouse models of cardiac diseases and showed that although the background strain of mice was an important source of metabolic variation, multivariate statistics were able to separate each disease model from the control strain [Jones et al., 2005].

Another mouse model well characterized form the metabolic point of view is the PPAR-α null mouse where authors demonstrate that age exacerbates the changes associated with genotype in this mouse model lacking the peroxisome proliferator activated receptor-α [Athertorn et al., 2009]. Recent reports describing the phenotype of genetically modified mouse models tend to include data regarding the metabolism [Inserte et al., 2009, Rodriguez-Sinovas et al., 2010, Ashrafian et al., 2010]. This is a trend that will become increasingly popular as it becomes clear that the the metabotype is a relevant part of the phenotype.

In 2003 a method to hyperpolarize $^{13}$C and increase the signal to noise ratio 10000 times was described [Ardenkjaer-Larsen et al., 2003], this opened the possibility of following metabolism in vivo through $^{13}$C NMR spectroscopy. It has been possible to measure energy production, particularly the Krebs cycle turnover, by following the label in 1-$^{13}$C-pyruvate [Golman et al., 2008, Schroeder et al., 2009, 2010] in animal models. Also, pH can be monitored in vivo by measuring hyperpolarised bicarbonate signals [Gallagher et al., 2008, Schroeder et al., 2010]. This extremely sensitive method may play a great role in clinical practice, specially in heart failure where metabolism and its modification through medical treatment is relevant [Lee et al., 2005, Abozguia et al., 2006]

Because NMR is non-invasive, there is the possibility of obtaining localized spectra. While this is technically feasible in brain [Sundgren et al., 2009] it is more complex in the case of heart due mainly to cardio-respiratory motion. However, we were able to differentiate between necrotic, at risk and control areas in an animal model using high resolution magic angle spinning NMR (HR-MAS) [Barba et al., 2007], in consequence, if chemical shift imaging becomes feasible it would be possible to necrosis and at risk areas after myocardial infarction.

2.2 Treatment monitoring and follow up

One of the key areas of metabolomics application is treatment monitoring and follow-up; actually the term metabonomics was introduced by Nicholson and cols [Nicholson et al., 1999] as a toxicology tool to evaluate the effect of drugs. A decade later the same authors have introduced the term 'pharmacometabonomics' in drug treatment defined as an approach “which uses metabolite profiling and chemometrics to model and predict responses of individual subjects” [Clayton et al., 2006, Nicholson et al., 2011]
A great variety of studies have been published in the literature, mostly in the field of cancer [Spratlin et al., 2009]. But in the area of cardiovascular diseases the amount of reports has been very limited, possibly associated to the few new drug treatments appearing for cardiovascular diseases as compared to, for example, cancer.

So far, reports regarding heart disease and drug treatments in the literature are focused on animal models either studying drug toxicity [Perrine et al., 2009, Andreadou et al., 2009] or the effect of antioxidant intervention [Constantinou et al., 2009]. This relative lack of reports may be due to the fact that the number of new drugs under development in the cardiovascular area is not as high as in other fields such as oncology.

2.3 Clinical application of metabolomics

Since the beginning of the 21st century the term 'personalized medicine' has continuously gained popularity and is now considered an essential trait of present and future medicine. For personalized medicine to be successful, it is necessary to properly identify subjects at increased risk of developing a disease, which patients will respond to a given therapy or how a disease will evolve in each case. In other words it is important to genotype and phenotype the individual patient so that its individual response to disease and treatment can be predicted.

Metabolomics is a promising technology for personalized medicine, it is fast, reproducible, easy to automate and not very expensive on a per sample basis. Most important, unlike the genome, the metabolome is modified by external factors such as disease. Today the only field where personalized medicine is flourishing is cancer; by definition cancer does imply modifications of the genome and genomics is technically more advanced than metabolomics. For diseases that does not imply the modification of the genome, such as cardiovascular pathology, metabolomics may be the tool of choice and extensive research has been conducted in recent years towards achieving this goal.

In 2002, Brindle and cols. published a paper in which they describe "a techniques capable of providing an accurate, noninvasive and rapid diagnosis of coronary heart disease" [Brindle et al., 2002]. 1H NMR spectra from serum samples from patients diagnosed with no coronary artery disease or with one, two or three vessel disease was used to create discriminant models able to differentiate the severity of coronary artery disease. Unfortunately, the authors did not take into account the effects of medical treatments (particularly statins) and patient gender in their analysis and was later demonstrated that those two factors have a strong influence in the NMR pattern. As a result, the specificity and sensitivity first predicted of above 90% was reduced, once corrected by gender and statin use to approximately 60% [Kirschenlohr et al., 2006]. Also, age has also been described as a differential factor regarding lipoprotein subclasses measured by NMR spectroscopy [Freedman 2004]. Nevertheless, the paper by Brindle and cols [Brindle et al., 2002] has been extremely important in the field because it quick started pre-clinical metabolomic studies.

Sabatine and cols [Sabatine et al., 2005] showed that it was possible to apply metabolomic analysis in a carefully characterized cohort of patients undergoing exercise stress testing and to differentiate between patients that developed inducible ischemia from the ones that did not [Sabatine et al., 2005]. This work was done by analyzing serum samples
obtained before, during and after stress testing by high performance liquid chromatography coupled to mass spectroscopy; ischaemic patients had higher circulating levels of metabolites belonging to the citric acid pathway. Later, we were able to show that it was possible to predict which patients would suffer from a positive stress test by analyzing serum samples obtained before the stress test using $^1$H NMR spectroscopy [Barba et al., 2008]. Ischemic patients had relatively higher lactate levels than non ischemic suggesting an underlying ischemic process although it could not be directly related to myocardial ischemia. Differences between the two groups were seen in samples obtained prior to the stress test thus, at least in theory, it would be possible to use metabolic profiling instead of stress testing but, unfortunately, the accuracy of the predictions was around 85%, clearly not high enough as to be used in the clinical setting. Searching for early markers of ischemia, Lewis and cols [Lewis et al., 2008] studied patients undergoing planed alcohol septal ablation and were able to detect several metabolites that changed early after myocardial injury and report a ROC curve for a composite metabolite score with an area under the curve of 0.88 ± 0.07.

It has been known for a time that patients with heart failure (HF) have an altered heart metabolism [Clark et al., 1996] and that metabolic modulation (shifting the main substrate from free fatty acids to glucose) improved VO2max, left ventricular ejection fraction, symptoms, resting and peak stress myocardial function, and skeletal muscle energetics [Lee et al., 2005]. Metabolic modulation as a tool to treat patients with heart failure has attracted interest [Abozguia et al., 2006] but the metabolomic analysis has not followed suit until recently when Kang and cols showed by profiling urine by NMR spectroscopy that it was possible to detect changes between HF patients and controls [Kang et al., 2011]. It could be interesting to evaluate possible changes in the urine metabolic profile of patients treated with drugs targeting heart metabolism for example, perhexiline or trimetazidine.

Metabolomics have been used to study diseases related to the cardiovascular system like diabetes [Wang et al., 2011], and obesity [van de Woestinje et al., 2011] reviewed in [Griffin and Nichols 2006, Muller 2010] but this is beyond the scope of the present work.

3. Future directions

3.1 Next hurdles before clinical application

There has been interest to apply metabolomic analysis to the clinical setting and it has been shown in the literature that various cardiovascular-related pathologies could be differentiated using metabolic analysis [Lee et al., 2005, Lewis et al., 2008, Barba et al., 2008, Kang et al., 2011] but there is still some work to be done for metabolomics in order becoming a clinical reality.

There are two main problems to solve before metabolomics becomes a widespread clinical reality; the first is to make metabolomics technically easy from both data acquisition and processing point of view as to be used in daily routine by non-specialists. The second problem will be more difficult to solve and it relates to population variation; so far studies have focused in small, controlled populations; studies with large cohorts of patients are not yet present in the literature.
Technical problems will be solved soon, there are new advances in automation procedures, mainly in the NMR field, that will allow to use metabolomics routinely for screening purposes in the near future. Moreover, once metabolic biomarkers are found using any of the hypothesis free metabolomics approaches their detection may be done using standard analytical methodologies already implemented in the clinical laboratory; in this case the application of metabolomics derived results would be straight forward.

Solving problems related to population variability is much more challenging; first of all 'omics' data processing is focused on finding sources of variation that can be later correlated with disease, treatment response, etc. Adding sources of variation to the dataset makes it difficult to identify specific patterns. Studies at population level should provide evidence for the background variation in metabolic profile and, thus tell how well metabolomics can perform in the clinical setting.

One of the areas where the application of metabolomics may become a reality soon is the evaluation of response to pharmacologic treatment. Although this is one of the first areas were metabolomics were applied, in the case of cardiovascular diseases so far it has only been reported in animal studies (see above). Still, it has been show that drugs do influence the metabolic pattern in human studies of plasma [Kirschenlohr et al., 2006] and that altering metabolism could be a therapeutic target [Lee et al., 2005] thus pharmacogenomics application may not be far away in the area of cardiovascular disease.

Work has started in the integration of various 'omics' technologies in global systems biology approaches [Mayr et al., 2007, 2008, 2009a, 2009b. Wheelock et al., 2009, De Souza et al., 2010]. Systems biology should provide with a complete and definitive view of a biomedical problem [Barabasi et al., 2011] but there are some problems associated that need to be solved. For example, just to integrate all the data it has to be considered that response time after a stimulus is not the same in transcripts, proteins or metabolites. Also, fold change in metabolite levels does not have to mimic changes in mRNA or protein levels.

4. Conclusion

In conclusion, the future of metabolomics is now. It is clear that metabolomics can be applied to various cardiovascular related diseases, although its clinical value in different settings remains to be determined; this is the next big challenge in the field.

Metabolomics, being a high throughput methodology is very well suited for biomarker discovery. The detection of biomolecules is an established methodology in the everyday clinical practice thus once a new biomarker has been found its application to the clinic should be straight forward. In our opinion, this approach will be successful in the sort term if it is able to address specific and relevant clinical situations.

Also, efforts should focus on the integration of metabolomics with systems biology analysis as in the long term, for the success of personalized medicine, the treatment of diseases will give way to the treatment of patients each with individual co-morbidities and environment. How well the integration of systems biology is archived will define the future not only of metabolomics but of all the 'omics' approaches.
5. Acknowledgment

This work has been founded by the Spanish ministry of Science SAF-2008-03067 and Instituto de Salud Carlos III, RECAVA DR06/0014/0025. I Barba is recipient of a Ramón y Cajal Fellowship.

6. References


Hotelling, H. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, Vol 24(6), 1933, 417-441


Coronary Artery disease is one of the leading causes of death in industrialized countries and is responsible for one out of every six deaths in the United States. Remarkably, coronary artery disease is also largely preventable. The biggest challenge in the next years is to reduce the incidence of coronary artery disease worldwide. A complete knowledge of the mechanisms responsible for the development of ischaemic heart disease is an essential prerequisite to a better management of this pathology improving prevention and therapy. This book has been written with the intention of providing new concepts about coronary artery disease pathogenesis that may link various aspects of the disease, going beyond the traditional risk factors.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
