PLANT AND FOOD METABOLOMICS IN THE POST-GENOMIC ERA: CONCEPTS, METHODOLOGIES AND APPLICATIONS

(review)

Carmen Socaciu

Department of Chemistry and Biochemistry, University of Agricultural Sciences and Veterinary Medicine, 3-5 Manastur street, 3400 Cluj-Napoca, Romania, E-mail: csocaciu@usamvcluj.ro

Key words: phytochemistry, plant metabolomics, fingerprint, metabolic profile, chromatography, spectroscopy, Infrared, Raman, Mass spectrometry, plant extracts

Abstract: The metabolome (by analogy to genome, transcriptome, proteome) represents generally the total metabolite pool of a living organism, the entire complement of all the low molecular weight metabolites (small organic molecules, i.e. sugars, amino acids, flavours, acids, pigments, hormones) in biological samples such as a leaf, fruit, food, blood, etc. Metabolomics refers either to plants, microorganisms, food or animal and human organisms. Different specific aspects of fingerprinting, metabolic profile and metabolite target analysis are presented. A critical discussion of methodologies used in metabolomics is presented. Finally, there are mentioned the advantages offered by metabolomics versus genomics and their applications, i.e. a tremendous number of measurements to be done in short time and with high resolution and sensitivity, to realize “maps” of plants and their derived products.

Abbreviations: HPTLC – High Performance Thin Layer Chromatography; HPLC-High Performance Liquid Chromatography; GC/MS – Gas Chromatography coupled with Mass Spectrometry (MS); Nuclear Magnetic Resonance (NMR); IR-Infrared; NIR-Near Infrared; CE-Capillary Electrophoresis.

INTRODUCTION

In the post-genomic era, as addition to the family of ‘omics’ is metabolomics as an integrative representation of the interactions of the genome, transcriptome and proteome with the environment (Fiehn, 2002; Fridman, 2005). The combination of the four „omics” provides a platform to investigate the regulation of the cellular environment from the specific codes stored in the DNA, through its expression by proteins (e.g. enzymes, structural proteins) to the chemical conversion of small molecules within biochemical pathways.

The development of these ‘omics’ is depended on analytical advances in spectroscopy, chromatography and electrophoresis, sensitive and specific analytical techniques to permit the handling of large numbers of samples. While micro-chip arrays permit testing of RNA and
DNA fragments against libraries of several hundred standards at the same time, the multiparallel chromatographic, IR/Raman and mass spectrometry as well nuclear magnetic resonance methods perform the separation, identification and quantification of small molecules and fragments of proteins (Brown, 2002; Sumner, 2003).

Due to the high quantity of informations cumulated by measurements, mathematicians have joined the chemists, biochemists and molecular biologists to develop algorithms that allow data analysis representing the results in a comprehensible manner, via bioinformatics (Taylor, 2002).

This area represents one of the main scientific world priorities and is a component of the FP7 priorities program on European Agriculture-Food-Health fields.

The metabolome (by analogy to genome, transcriptome, proteome) represents a pool of total metabolites of a living organism, the entire complement of all the low molecular weight metabolites (small organic molecules, i.e. sugars, amino acids, organic acids, pigments, hormones) in a biological sample such as a leaf, fruit, food, blood, etc. The evaluation of genes perturbations by transcriptome, proteome, and metabolome analysis, represents a valuable tool to study their functions (Sumner, 2003).

Metabolomics is also called metabolic profiling, biochemical profiling, metabolome analysis, metabolic phenotyping and realizing a high-throughput analysis of metabolites released by biochemical processes in living organisms. This field of research has evolved from the increasing amount of genetic (genomics) and protein (proteomics) data, generated from recent advances in molecular biology techniques (Bender, 2005). Metabolomics can be considered a holistic description of the metabolic complement of a defined genotype, and refers also to a holistic analytical approach of metabolism. Metabolomics sets out to determine how the metabolite levels respond to genetic or environmental changes or may be a reflection of molecular changes induced during growth, post-harvest and processing.

Metabolomics refers either to plants, microorganisms, food or animal and human organisms.

Plant metabolomics (PM) is the large-scale phytochemical fingerprint of plant metabolome (around 200000 metabolites found in plants), in relation to functional genomics, proteomics and systems biology (Trethewey, 2001; Sumner, 2003; Ott, 2003; Oksman-Caldentey, 2005). Plant metabolic profile can be a key-way to identify and preserve different plant or seed species, from different areas, detect genetically-modified plants, detect plant products and their safety, degradation and traceability (Noteborn, 2000; Abel, 2002; Hall, 2002).

Food metabolomics (FM) aims to encourage application of plant metabolomics-inspired approaches to foods, to improve understanding of the molecular basis for food quality, prediction of structure/function relationships needed for important quality attributes such as flavour, colour and added-value nutrition (by functional food components with health protecting ability). Metabolomics and nutrigenomics are future, prioritary scientific areas to allow the characterization of plant-based food's chemical profile in relation to our individual metabolism. For example, some phytochemicals in food promote genetic stability in less stable genomes, or modify proteins/enzymes, involved in tumor incidence or suppression.

Very promising application areas of FM are the identification of effects of whole food (or extracts) on cellular level and nutritional benefits (1), the characterization of the molecular basis for sensory (e.g., taste, aroma) perception from complex foods (2), tracking the complex
post-harvest and food processing molecular changes, (3), consideration of genetic and/or environmental factors affecting metabolomes (4), important for final quality.

The metabolomics is a fundamental scientific field which offers informations, complement ary to transcriptomics and proteomics for characterizing the metabolic functions. The transcriptome and proteome are NOT adequate monitors of cell function, due to complex relationships between mRNA, protein levels and metabolic function. The metabolite levels change according to developmental, physiological, and pathological states, so it is a complex and difficult area to investigate and make accurate interpretations (Weckwerth, 2003).

Metabolic fingerprinting is a high-throughput, rapid, global analysis of samples to provide a specific recognition-shape of a sample, without proper quantification and metabolite. It is also a screening tool to discriminate between samples of different biological status (a mirror of genomic characteristics) or origin (species/varieties, geographic origin). Sample preparation is simple and analysis is rapid (1 min or less)

The metabolic profile is determined accurately by simultaneous, parallel identification and quantification of a selected number of pre-defined metabolites level, generally related to a specific metabolic pathway(s). The sample analysis is difficult and need advanced instrumentation, different procedures to isolate subclasses of metabolites prior to detection, using chromatographic separation prior detection with spectroscopic methods like UV-Vis, NMR or MS. Metabolic profiling is now rapidly expanding to catch up with other multiparallel analytical techniques (mRNA profiling, proteome analysis), particularly in plants. It is far less developed than mRNA profiling or proteomics. Metabolic profiling is now growing fast in both public and private sectors, and plants are major targets (Fiehn, 2000, 2001, 2002).

The metabolite target analysis is a qualitative and quantitative evaluation of one or a few metabolites related to a specific metabolic reaction. Extensive sample preparation and separation from other metabolites is required, especially when low limits of detection are required. Generally, chromatographic separation is used followed by sensitive MS or UV detection.

METHODOLOGIES

During the last 10 years, a new and advanced analytical methodologies and technologies were introduced related to the analysis of plant metabolomes (Dunn, 2005). Sample preparation depends on the sample and metabolomics strategy employed. In many applications, no further isolation of metabolites from the sample matrix is performed, and samples are diluted and analysed directly or analysed after chemical derivatisation. Many of the advanced technologies are used in metabolomics, including high-performance chromatography or electrophoresis, optical spectroscopy, nuclear magnetic resonance, and mass spectrometry. Bioinformatics and various methods of data visualization plays a critical role.

Extraction procedures. No single extraction procedure works for all metabolites, due to their different solubility, stability, interferences in extraction solvents (i.e. pigments, versus alkaloids, phytohormones or polyphenols) because conditions that stabilize one type of compound will destroy other types or interfere with their analysis (Dunn, 2005). Therefore the extraction protocol has to be tailored to the metabolites to be profiled. Further fractionations by chromatography may split metabolites into classes prior to analysis (Fiehn et al, 2000a).
Metabolic profiling has up to now largely been confined to fairly stable compounds that can be extracted together, such as major primary metabolites (sugars, amino acids, etc) but few secondary metabolites (e.g., phenylpropanoids, alkaloids) which are regulatory, minor components with high impact on health as antioxidants, anti-cholesterol, natimutagenic or anticarcinogenic compounds. A huge efforts is done now in advanced laboratories all over the world to realize the metabolic profile and fingerprint for polyphenols, antioxidant carotenoids, anthocyanins, unsaturated fatty acids, which give the functionality of our food for beneficial actions on the health. A comprehensive profiling would covers all these classes of compounds (300-500 compounds), in fresh plants up to their profile in food ready to eat, including bioavailability to humans, according to their genomic profile. A few hundred such compounds most but not all of which are identified.

The “minimum” instrumentation useful for such studies include either direct spectroscopy or chromatography (fingerprinting) or high-performance chromatography (GC- or HPLC-) combined with spectroscopy (IR, NIR, MS) for single or multiple metabolite profiling (Malmquist, 1994).

Gas Chromatography/Mass-Spectrometry (GC/MS) is used for metabolites with high thermal stability and volatility (Fiehn et al., 2000b; Roessner et al., 2001) which are separated after a previous derivatization and the eluting compounds are detected and identified by electron-impact mass spectrometers. In metabolomics, GC-MS has been described as the gold standard (Brown, 2005; Dunn, 2005)

Direct-injection MS analysis may also be applied for phenotyping plants, i.e. FT-MS, provides ultimate limit of detection and precision of mass measurement to enable metabolomic analyses (Aharoni et al., 2002).

Other methods as Fourier transform ion cyclotron resonance (FTICR)-MS, provides high resolving power. All possible metabolite peaks can be resolved and their molecular formula calculated very accurately. Generally, quadrupole or TOF instruments are employed, although FTICR instruments have also been used. Another area of growing interest is the application of matrix-assisted laser desorption ionisation (MALDI), laser desorption ionisation (LDI) or direct ionisation on silicon (DIOS) to provide ionisation of metabolite solutions spotted directly on a target plate, allowing minimal sample preparation and high-throughput analysis.

Liquid Chromatography/Mass-Spectrometry LC/MS is more suitable than GC/MS for labile compounds, for those that are hard to derivatize, or hard to render volatile. LC/MS is less developed as a profiling tool than GC/MS. A closely related method is capillary electrophoresis (CE)/MS (Malmquist, 1994; Dunn, 2005).

Capillary electrophoresis coupled with MS is also a new type of analysis with high resolution and reliability (Terabe, 2001; Soga, 2003)

FT-IR, NIR and Raman, and more recently magnetic resonance (NMR) spectroscopy are constantly developing techniques, rapid, nondestructive, reagentless and high-throughput for a diverse range of sample types. Very recently, FT-IR has also been introduced as a metabolic fingerprinting technique within the plant sciences (Ellis, 2003; Brown, 2005, Baranska, 2004, 2006). Raman spectroscopy coupled with microscopy is recently used with great success, making possible the identification and quantification of phytochemicals distribution directly from the plant tissues (Roessner, 2001, Quilitzsch, 2005; Schultz, 2006).

NMR techniques require pure samples and a minimum quantity which depass the requirements of other routine methods. It becomes difficult for minor derivatives in complex mixtures.
The performances of these various metabolomic techniques are different (Fig.2): NMR has rapid analysis times but lower sensitivity. GC/MS and HPLC/MS provide good selectivity and sensitivity, CE/LIF (laser induced fluorescence) provides high sensitivity, lower selectivity (Sumner, et al., 2003).

To give an example of the complex interpretation given by Raman spectroscopy and microscopy (Fig.3.): the microscopic image of chamomile inflorescence (A), is correlated with the FT-Raman spectrum around red cross region, (B) and Raman mapping shows the distribution of polyacetylenes (C) and carotenoids (D)(Sumner, 2003).

Data Analysis. An avalanche of profiling data for metabolites are yet accumulated, but difficult to analyze and compare. The problems in moving from the descriptive level to large data sets are similar for all forms of profiling, the objective is to recognize non-random patterns and to identify or quantify specific markers. Various statistical methods (‘data mining tools’) are used in order to reduce data complexity by focusing on the information content of a given data set. Data mining tools include cluster analysis (CA), principal components analysis (PCA), and discriminant function analysis (DFA). Cluster analysis (CA) is a collection of statistical methods that group (‘cluster’) similar data together (Fiehn et al., 2000; Roessner et al., 2001) and have certain properties in common. Data are presented in dendrograms that emphasize naturally related groups.

Principal Component Analysis (PCA) uses the metabolite data from a sample to compute an individual metabolic profile that is afterwards compared to the other profiles, resulting a cloud of data points, represented as two- or three-dimensional plots (Fiehn et al., 2000a,b;). Discriminant Function Analysis (DFA) is used to determine which variables discriminate between two or more groups, similarly to variance analysis (Roessner et al., 2001).

CONCLUSION

In conclusion, to mention the most important advantages to study metabolomics (Brown, 2005):

1. The metabolome is the downstream product of gene expression so reflects the functional level of the cell more appropriately and changes in the metabolome are expected to be amplified relative to proteome or transcriptome.
2. The number of metabolites is lower than the number of genes and proteins in a cell, so becomes more informative and easier the use of metabolic profile than the DNA profile.
3. Considering that metabolic fluxes are regulated not only by gene expression but also by environmental conditions, the measurement of the metabolites is more appropriate to characterize a plant metabolome.
4. The advanced methodologies available now make possible a tremendous number of measurements to be done in short time and with high resolution and sensitivity to realize “maps” of plants and their derived products.
5. Metabolomic experiments cost 2-3 times less than proteomic and transcriptomic experiments.

So, fundamental studies to elucidate plant and food metabolome and fingerprint can develop new applications and technological approaches.
Fig.1. Evaluation of genes perturbations by transcriptome, proteome, and metabolome analysis, a valuable tool to study their functions (Sumner, 2003).

Fig.2. Performances of various metabolomic techniques: NMR has rapid analysis times but lower sensitivity. GC/MS and HPLC/MS provide good selectivity and sensitivity, CE/LIF (laser induced fluorescence) provides high sensitivity, lower selectivity (Sumner, et al., 2003).

Fig.3. Microscopic image of chamomile inflorescence ((A), the FT-Raman spectrum at the red cross region, (B) and Raman mapping showing the distribution of polyacetylenes (C) and carotenoids (D). (Sumner, 2003)
REFERENCES