Advances in Nutrigenomics research: Novel and future analytical approaches to investigate the biological activity of natural compounds and food functions

V. García-Cañas, C. Simó, C. León, A. Cifuentes*

Institute of Industrial Fermentations (CSIC), Juan de la Cierva 3, Madrid, Spain

*Corresponding author: acifuentes@ifi.csic.es

1

ABSTRACT

In recent years, nutrition research has moved from classical epidemiology and physiology to molecular biology and genetics. Following this trend, Nutrigenomics has emerged as a novel and multidisciplinary research field in nutritional science that aims to elucidate how diet can influence human health. It is already well-known that bioactive food compounds can interact with genes affecting transcription factors, protein expression and metabolite production. The study of these complex interactions requires the development of advanced analytical approaches combined with bioinformatics. Thus, to carry out these studies Transcriptomics, Proteomics and Metabolomics approaches are employed together with an adequate integration of the information that they provide. In this article, an overview of the current methodologies and a thorough revision of the advances in analytical technologies and their possibilities for future developments and applications in the field of Nutrigenomics is provided.

1. Introduction to Nutrigenomics.

2. Advances in Transcriptomics.

       2.1 Gene expression microarray technology.

       2.2. Sequencing-based technologies.

       2.3. Bioinformatics and Gene Ontology database.

3. Advances in Proteomics.

       3.1 Bottom-up approach,

       3.2 Top-down approach.

       3.3 Other proteomic approaches

4. Advances in Metabolomics.

       4.1. Analytical techniques in Metabolomics.

       4.2. Data analysis in Metabolomics.

5. Systems Biology.

6. Future needs and developments in Nutrigenomics.

1. Introduction to Nutrigenomics.

Diet is a key environmental factor affecting health and the incidence of many chronic diseases [1]. Nutrition research has traditionally explored the importance of this relation by classical approaches based on human intervention studies and the use of biomarkers. The effects of nutrient deficiencies, imbalance of macronutrients or toxic concentrations of certain food compounds on health have been main topics in Nutrition research [2, 3]. On the other hand, other bioactive food constituents such as certain polyphenols, vitamins, carotenoids and terpenoids have significant beneficial effects for health promotion and disease prevention by reducing the process of sustained inflammation that accompanies chronic disease [4, 5]. Most foods are composed of diverse constituents, many of which have specific biological activity. Owing to the lack of effective analytical approaches that take into account multiple aspects of the biological effects of food compounds, little is known about their molecular functions and the biological processes involved. In addition to this, bioactive food compounds also interact with each other, making even more difficult any effort to identify their biological activity [6].

Nowadays, it is recognized that understanding the effect of diet on health requires the study of the mechanisms of nutrients and other bioactive food constituents at the molecular level. This is supported by the increasingly growing number of studies in humans, animals and cell cultures demonstrating that nutrients and other bioactive compounds in food can regulate gene expression in diverse ways [7]. More specifically, nutrients and other bioactive food constituents have been referred to as signals that are detected by cellular sensor systems and affect the expression of the genome at several levels (mRNA and proteins) and subsequently, the production of metabolites [8, 9]. These aspects have motivated current trends in Nutrition research to study how diet affects the balance between health and disease by altering the expression of an individual's genetic makeup (Nutrigenomics) or how the genetic variability among individuals can influence their predisposition to suffer specific illness related to diet as e.g., obesity (Nutrigenetics). As a consequence, Nutritional Genomics (Nutrigenomics and Nutrigenetics) has emerged as a new field that focuses on the study of the interaction between nutrition and human genome [10, 11]. The topic of this review, Nutrigenomics,

4

focuses on the study of the impact of specific nutrients and diets on health through the expression of genetic information by the integration of "omics" technologies such as Transcriptomics, Proteomics and Metabolomics [12, 13].

One of the main interests in Nutrigenomics research relates to health and prevention of chronic diseases (such as e.g., cardiovascular diseases, metabolic syndromes, cancer, etc) through diet. These disorders are complex and multifactorial in their origin, involving genetic factors but also a number of behavioural and environmental factors such as exposure to certain food components [14]. Also, it has been underlined the importance of the biological homeostasis maintenance for disease prevention. The loss of homeostasis and altered biochemical composition of cells and tissues can be primary cause in disease [15]. In line with this, there is a necessity for understanding the molecular mechanisms that describe homeostasis at biochemical, cellular and organ levels, associated with a healthy status and how diet affects this homeostatic control [16].

As a result, there is a need for molecular biomarkers that allow early detection of the onset of disease or, ideally, the pre-disease state [3]. These early effect biomarkers should accurately reflect subtle changes in homeostasis and the efforts of the body to maintain it [16]. However, the discovery of such biomarkers is not easy since diet-gene interactions are complex. Unlike the comparative simplicity of the single-gene disorders, chronic diseases are likely the result of multiple genes and multiple variants of each gene interacting with multiple environmental factors, each combination making a relatively small contribution to overall homeostasis, function and health. Thus, to determine health status and reflecting the functional response to a bioactive food component [8], complete biomarker profiles of gene expression, protein expression and metabolite production will be more useful than single markers. To do this, the availability of advanced analytical techniques will be essential for the investigation of complete biomarker profiles of gene expression, protein expression and metabolite production.

The mentioned profiles can be considered as 'dietary signatures' that Nutrigenomics will use to understand the cellular functions of nutrients and other active food components and how they affect homeostasis in specific tissues within the whole organism. To attain this, Nutrigenomics makes use of an integrated analytical approach including the latest developments in high-throughput –omics techniques for the comprehensive study of different aspects of this biological complexity. This integrated analytical approach puts together Transcriptomics, Proteomics and Metabolomics, and might also be ideal for elucidating the effects of novel functional foods and nutraceuticals on global expression of genetic information and cell function without making assumptions about what to look for in terms of risk, and providing new means for discovering biomarkers for efficacy testing of bioactive functional food ingredients [1].

The development of Genomics, Transcriptomics, Proteomics and Metabolomics [17] has created extraordinary opportunities for increasing our understanding about (i) the biochemical, molecular and cellular mechanisms that underlies the beneficial or adverse effects of certain bioactive food components; (ii) the identity of genes that are involved in the previous stage to the onset of the disease, and therefore, possible molecular biomarkers; and (iii); the effect of bioactive food constituents on crucial molecular pathways [8, 18]. In the following sections, the main characteristics of these –omics techniques and their use in Nutrigenomics are reviewed.

2. Advances in Transcriptomics.

In the expression process of genomic information, several steps may be regulated by nutrients and other bioactive compounds in food. Consequently, the analysis of changes in mRNA expression by nutrients and bioactive food constituents is often the first step to study the flow of molecular information from the genome to the proteome and metabolome and one of the main goals in Nutrigenomics research [8]. For years, the expression of individual genes has been determined by quantification of mRNA with Northern blotting. This classical technique has gradually been replaced by more sensitive techniques such as real-time PCR. Both techniques, however, can only analyse gene

expression for a limited number of candidate genes at a time. This is an important limitation for their application in Nutrigenomics research since the analysis of a reduced number of genes may not provide insights about the causative relationship between the bioactive food constituent and its biological effect [19].

On the opposite, the analysis of global gene expression may offer better opportunities to identify the effect of bioactive food constituents on metabolic pathways and homeostatic control and how this regulation is potentially altered in the development of certain chronic diseases [13]. In the past decade, two conceptually different analytical approaches have emerged to allow quantitative and comprehensive analysis of changes in mRNA expression levels of hundreds or thousands of genes. One approach is based on microarray technology, and the other group of techniques is based on DNA sequencing [20].

2.1 Gene expression microarray technology.

During the last years, owing to the extensive optimization and standardization, gene expression microarray has become a leading analytical technology in Nutrigenomics research for the investigation of the interactions between nutrients and other bioactive food compounds and genes [15, 19, 21].

Typically, DNA microarrays are collections of oligonucleotides or probes, representing thousands of genes, attached to a substrate, usually a glass slide, at predefined locations within a grid pattern. This technique is based on specific nucleic acids hybridization and it can be used to measure the relative quantities of specific mRNAs in two or more samples for thousands of genes simultaneously.

Regardless the platform used for the analysis, the typical experimental procedure is based on the same analytical steps: RNA is extracted from a source of interest (tissue, cells, or other materials), labelled with a detectable marker (typically, fluorescent dye) and allowed to hybridize to the microarrays with individual DNA sequences hybridizing to their complementary gene-specific probes on the

microarray. Once hybridization is complete, samples are washed and imaged using a confocal laser scanner (**Figure 1**). Theoretically, the fluorescent signal of derivatized-nucleic acids bound to any probe is a function of their concentration. The relative fluorescence intensity for each gene is extracted and transformed to a numeric value [22].

Gene expression microarrays are powerful, but variability arising throughout the measurement process can obscure the biological signals of interest. In order to fully exploit the possibilities of DNA microarrays, a careful study of the experimental design is required [23]. Data acquisition and pre-procesing are important post-technical steps in microarray experiments. The latter involves data normalization, which consists of adjusting the individual hybridization intensities in order to remove variation derived from unequal quantities of starting RNA, differences in labelling or detection efficiencies between the fluorescent dyes used, and systematic biases in the measured expression levels. The most common approaches adopted for normalization are total intensity normalization, Locally Weighted Scatterplot Smoothing (LOWESS) and the use of housekeeping genes [24, 25].

Differences in precision, sensitivity, or specificity between different microarray platforms, makes comparisons between experimental platforms unreliable. Moreover, the annotations across different platforms are not represented by exactly the same gene sequence regions. In order to alleviate in part cross-comparison problems among distinct microarray platforms, the Microarray Gene Expression Data (MGED) organization [26] has established Minimum Information About a Microarray Experiment (MIAME) guidelines for microarray data annotation and reporting. These guidelines, aimed to improve the data sharing between platforms, have been adopted by a number of scientific journals. Furthermore, in an effort to make MIAME-compliant data publicly available, the ArrayExpress microarray database [27] has been created as a repository for Transcriptomics experimental data. In addition, this database accepts processed data files generated with the very recent RNA-Seq technologies that will be discussed later. Also, ArrayExpress Atlas [28] is a recent tool developed to allow the user to query for conditions-specific gene expression across multiple data sets within the increasingly growing ArrayExpress database [29].

The analysis of the vast amount of microarray data for extracting biologically meaningful information is perhaps the most challenging and daunting tasks [30]. The fundamental goal of microarray expression profiling is to identify genes that are differentially expressed in the condition of interest [31]. Data analysis process can be divided into three main parts: Identification of significantly regulated genes, identification of global patterns of gene expression, and determination of the biological meaning of both individual genes and group genes. First, filtering criteria, including fold change and statistical significance determined by comparison statistics is necessary in order to identify candidate genes that are differentially expressed. After filtering data, the differentially expressed genes are then classified into discrete groups or clusters, based on expression pattern. This step allows the identification of groups of co-regulated genes. To achieve this, sophisticated bioinformatics tools such as unsupervised clustering, principal component analysis and self-organizing maps are the most widely used [32, 33].

Most of microarray studies published to date corroborate the gene expression data by an alternative sensitive technique. In general, there is a good correlation between microarray results and other more traditional methods, indicating the ability of the technology to produce reliable results [33]. The validation is generally performed by analysing a selection of differentially expressed genes of interest in the samples using real-time, quantitative reverse transcriptase polymerase chain reaction (RT-PCR) [35-38], although other methods such as Northern blotting [39], quantitative radioactive *in situ* hybridization and verification at the level of the protein, through routine western blotting [40], immunohistochemical or immunocytochemical assays has been also used [31, 41]. In RT-PCR analysis, mRNA is converted into cDNA and subsequently amplified in a PCR by specific primers in the presence of a fluorescent dye or internal probes. The expression level of the target gene is computed relative to the expression level of one or more reference genes, often housekeeping genes. Selecting proper housekeeping genes is one of the most critical aspects of the analysis, since they need to be constantly expressed between all samples and conditions in the experiment. In a recent study, the application of a 96-well plate PCR-array analysis for validation of microarray data suggested that most

differences between the results obtained by two different technologies were attributed to saturation problems in the microarray, strict quality criteria for array-PCR data analysis, and possible differences in isoforms detected by the two technologies [21].

Early applications of microarray to Nutrigenomics were related to the effects of caloric restriction on aging [42-45]. Soon, the technology was extended to study other interesting aspects in Nutrigenomics including the effects of dietary protein in the gene expression of cells [46, 47], the mechanisms dietary long chain polyunsaturated fatty acids in molecular function cancer and normal cells [48, 49] and the effects on transcriptome of a high- or low-carbohydrate intake [50, 51]. The molecular mechanisms of certain bioactive food constituents have also been investigated by microarray technology. A summary of some representative applications of microarray in Nutrigenomics field is given in **Table 1**.

2.2. Sequencing-based technologies.

In contrast to microarray technology, sequencing-based techniques consist of counting tags of DNA fragments to provide "digital" representation of gene expression levels using sequencing. These techniques include Serial Analysis of Gene Expression (SAGE) [57] and some of its variants, including LongSAGE [58], Cap Analysis of Gene Expression (CAGE) [59], Gene Identification Signature (GIS) [60]. In SAGE-based methods, the abundance of a particular mRNA species is estimated from the count of tags derived from one of its ends. First, restriction enzymes are used to obtain short tags of 14-21 bp, usually derived from one end of an mRNA, which are then concatenated, cloned and sequenced to determine the expression profiles of their corresponding mRNAs. Despite the more statistical robustness and less stringent standardization and replication requirements than those used for microarrays, the size of the sample tags should be increased in order to improve the precision and accuracy for detecting rare mRNAs [61]. Unfortunately, most of these methods are based on expensive Sanger sequencing method and, therefore, sequenced SAGE library rarely exhibits saturating tag counts that would indicate complete representation of the cellular transcriptome [62]. An additional limitation of these techniques is that owing to short length, many

SAGE tags are shared by the transcripts from different genes, which can complicate gene identification [63]. Moreover, the many PCR amplifications, cloning and bacterial cell propagations that these techniques involve may result in a quantitative bias for different tags [64]. Massively Parallel Signature Sequencing (MPSS) is a more sophisticated "clone-and-count" technique that also generates small tags of each mRNA species; however, it uses a different strategy that does not involve propagation in bacteria neither Sanger sequencing [62]. Nevertheless, the MPSS technology has been restricted to only a few specialized laboratories [64].

In the past few years, the development of "next-generation" sequencing methods, also referred to as either "massive parallel" or "ultra-deep" sequencing, is changing the way in which gene expression is studied [20, 65]. These novel technologies have had an enormous impact on research in a short time period, and it is likely to increase further in the future. Although they have not been applied in Nutrigenomics yet, their potential is underlined by their adoption in studies of transcriptomes during physiological changes, and for the comparative analysis between different disease states or conditions [66].

Next-generation sequencing technologies apply distinct concepts and procedures aimed to increase sequencing throughput in a cost-effective and rapid manner. Last advances in emulsion PCR and non-cloning based methods for DNA amplification, together with improved imaging instruments have facilitated the development of sequencers capable of reading up to tens of million of bases per run in a massively parallel fashion [67-69]. The read lengths typically achievable by these technology range from 30 to 300 bp, depending on the sequencing platform used. These short reads will represent a challenge for either mapping them to reference genome or *de novo* assembling without genomic reference [69].

The new features of next-generation sequencers have stimulated the development of new techniques that have expanded their applications, for example, to comprehensively map and quantify transcriptomes, which with Sanger sequencing would not have been economically or logistically

practical before [70, 71]. Similarly to the concept adopted in SAGE and MPSS, the principle behind these novel techniques for Transcriptomics, which have been termed RNA-Seq methods, is based on a global sequence census approach. Generally, a complex RNA sample is converted to a library of cDNA with adaptors attached to one or both ends. cDNA molecules, with or without amplification (depending on the sequencing platform), are subjected to massively parallel sequencing in short reads without bacterial cloning as a pre-requisite. Following sequencing, powerful bioinformatics tools are needed in order to align the resulting short reads (30-300 bp), depending on the sequencing platform and method used, to a reference genome or reference transcripts. That is, with reference genomes available, short reads are sufficient to map their locations rather than sequencing the entire large genome. Once mapped, the sequence hits are counted to determine their density and distribution and then, quantify the gene expression [72]. RNA-Seq is still a technology under active development that is being evaluated in multiple laboratories for RNA profiling [73-77] and other applications such as discovery of small RNAs, mapping RNA splice isoforms and discovery of small RNAs [78-80].

2.3. Bioinformatics and Gene Ontology database.

As mentioned, high-throughput technologies in transcriptomics usually generate large lists of differentially expressed genes as final output. However, the biological interpretation of such results is very challenging. Over the last years, the use of biological knowledge accumulated in public databases by means of bioinformatics, allows to systematically analyse large gene lists in an attempt to assemble a summary of the most enriched and significant biological aspects [81]. The principle behind enrichment analysis is that if a certain biological process is occurring in a given study, the co-functioning genes involved should have a higher (enriched) potential to be selected as a relevant group by high-throughput screening technologies. This approach increases the probability for researchers to identify the correct biological processes most pertinent to the biological mechanism under study [82]. Thus, a variety of high-throughput enrichment tools (e.g. DAVID, Onto-Express, FatiGO, GOminer, EASE, ProfCom, etc.) have been developed since 2002 in order to assist microarray end user to understand the biological mechanisms behind the large set of regulated genes. These bioinformatics

resources systematically map the list of interesting (differentially expressed) genes to the associated biological annotation terms and then statistically examine the enrichment of gene members for each of the terms by comparing them to a control (or reference). A recent work by Leong *et al.* shows the potential of exploiting bioinformatics tools for extracting valuable biological information from microarray experiments. In their work, the use of Affymetrix microarray platform and stringent data analysis provided interesting insights into a homeostatic mechanism, based on arginine-sensitive regulation that coordinates aspects related to nutrient availability in hepatic cells [38].

Enrichment analysis would not be possible without appropriately structured databases such as Gene Ontology [83]. More specifically, Gene Ontology provide a systematic and controlled language, or ontology, for the consistent description of attributes of genes and gene products, in three key biological domains that are shared by all organisms: molecular function, biological process and cellular component. Thus, standard biological phrases, referred to as terms, which are applied to genes and proteins are then, linked or associated with other Gene Ontology terms by trained curators at genome databases [84]. A recent study of the effects of soy isoflavones, based on gene expression microarray data, has revealed potential mechanisms of action for genistein by the combined use of GoMiner as the enrichment tool, Gene Ontology database and information from Dragon Estrogen Responsive Genes Database [24]. In addition to the aforementioned Gene Ontology database, recent enrichment tools also integrate information extracted from other databases as for instance, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (metabolites) in order to improve the comprehensiveness of this type of studies [82].

3. Advances in Proteomics.

The proteome is the set of expressed proteins at a given time under defined conditions, it is dynamic and varies according to the cell type and functional state. In the case of a Nutrigenomic study, the proteome provides a "picture" of the impact of specific bioactive nutrients and diets in a certain organism, tissue or cell in a particular moment. Thus, the importance of Proteomics as a tool to

understand the effect of the diet on health has already been recognised by several authors [18, 85-90], however, up to now the application of Proteomics in Nutrigenomics is still rather limited, being the number of review papers higher than research papers [91].

One of the main differences when working with proteins is that there is not an amplification methodology for proteins comparable to PCR. Physical and chemical diversity of proteins are also higher than nucleic acids. They differ among individuals, cell types, and within the same cell depending on cell activity and state. In addition, there are hundreds of different types of post-translational modifications (PTMs), which evidently will influence chemical properties and functions of proteins. PTMs are key to the control and modulation of many processes inside the cell. The selected analytical strategy for the detection of PTMs will depend on the type of modification: acetylation, phosphorylation, ubiquitination, sumoylation, glycosylation, etc. It is already known that dietary components can also modify the translation of RNA to proteins and the post-translational procedures [92]. Thus, it has been observed that dietary components such as diallyl disulfide, acompound found in processed garlic, has been shown to post-translationally modify proteins [92]. In other cases post-translational regulation of proteins by dietary components can involve the modification of their thiol groups [93].

Another important drawback in the proteome study is the huge dynamic concentration range of proteins in biological fluids or tissues. Namely, a 10^{10} dynamic range has been estimated in serum for protein concentration [94]. This situation causes many detection difficulties due to a large number of proteins are below the level of sensitivity of the most advanced instruments. For this reason, in Proteomics fractionation and subsequent concentration of the proteome is often needed [95, 96]. In a first instance, in order to reduce the complexity of the sample, fractionation of the proteome can be done by differential centrifugation into different populations of organelles [97] and also attending solubility properties into different cellular compartments. The chemical tagging strategies, involving the modification of functional groups of amino acid residues, including PTM in proteins (and peptides) is the methodology of choice when a specific population of the proteome is under study [98].

Another strategy for proteome fractionation that is gaining more and more significance is the sequential or simultaneous immuno-affinity depletion of the most-abundant proteins [99]. It has become the alternative in certain application, however, the main drawback regarding this approach is that co-depletion of a certain fraction of the proteome with the abundant proteins can occur. Chromatographic [100] and electrophoretic [101] strategies have also been applied to the pre-fractionation of the proteome. In general, the main problem of the mentioned strategies is that after fractionation or depletion, the remaining proteins (in case of depletion) or the proteins in the fractions of interest remain dilute, being still in not enough concentration for subsequent analysis. Most recently, it has been described the use of a library of combinatorial ligands, acting by reducing the signal of high-abundance proteins while increasing the level of the low-abundance ones to bring their signal within the detection limit of the present-day analytical instruments [102].

Because of the complexity of proteome the use and development of high-resolving separation techniques as well as highly accurate mass spectrometers is nowadays critical in Proteomics [103, 104]. Currently, more than a single electrophoretic or chromatographic step is used to separate the thousands of proteins found in a biological sample. This separation step is followed by analysis of the isolated proteins (or peptides) by mass spectrometry (MS) via the so-called "soft ionization" techniques, such as electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), combined with the everyday more powerful, easy to use and affordable mass spectrometers.

Although Proteomics has been scarcely applied to study the effect of nutrients on health, two fundamental analytical strategies can be employed: the *bottom-up* and the *top-down* approach. Both methodologies differ on the separation requirements and the type of MS instrumentation.

3.1 Bottom-up approach.

The *bottom-up* approach is the most widely used in Proteomics since it can apply conventional or modern methodologies. In a conventional approach, large-scale analyses of proteomes is accomplished

by the combination of two-dimensional gel electrophoresis (2DE) followed by MS analysis [105]. 2DE is the methodology that currently provides the highest protein species resolution capacity with low-instrumentation cost. However, 2DE is laborious, time-consuming and presents low sensitivity, depending strongly on staining and visualization techniques. 2DE has also some limitations to separate highly hydrophobic biomolecules or proteins with extreme isoelectric point or molecular weight values. Moreover, one of the major sources of error in 2DE is gel-to-gel variation. In this sense, the introduction of difference gel electrophoresis (DIGE), in which gel variation is eliminated by loading different samples in the same gel [106], has brought about an important improvement. DIGE methodology is based on the use of novel ultra high sensitive fluorescent dyes (typically, Cy3, Cy5 and Cy2) to label up to three different protein samples that will be separated in the same 2DE run. After image analysis of the 2DE gels, the protein spots of interest are subsequently submitted to an in-gel digestion step with a protease enzyme. MS and databases of protein sequences are then use in different ways for protein identification, following: (i) the peptide mass fingerprint (PMF) approach using MS data, in which the molecular masses of the peptides from the protein digest are compared with those simulated of already sequenced proteins, or (ii) by tandem MS in a sequence tag search in which a few peptidic sequences obtained from MS/MS analysis (using collisional activation or some other energy deposition process) can be used to search in protein databases. The main limitation of the *bottom-up* approach is that information obtained is related to a fraction of the protein, loosing information about PTM.

There are limited studies on the effect of specific natural compounds, nutrients or diets on the proteome, being most of them based in the *bottom-up* approach, more precisely in classical 2DE-MS, although also LC-MS/MS has been applied with this purpose. These studies are summarized in **Table 2**.

Shotgun Proteomics is an advanced *bottom-up* approach in which a complex protein mixture is digested with endoproteinases of known specifity [119], providing a comprehensive, rapid and automatic identification of complex protein mixtures. In spite of the higher sample complexity due to

the huge number of peptidic species obtained from the protein mixture hydrolysate, peptides are in general better separated and analyzed by MS than proteins.

An essential part of shotgun Proteomics is the use of powerful and highly resolving separation methods prior to MS analysis, allowing the study of complicated biological samples while avoiding suppression ionization problems. Liquid chromatography (LC) [120] and capillary electrophoresis (CE) [121, 122] in their different modes are the main methodologies applicable to the separation of complex peptide mixtures due to their high resolving power and their potential for full automation and high sampling rates. Although physico-chemical diversity of peptides makes them well suited to be separated by different liquid-based separation modes, reversed phase (RP) and aqueous mobile phases are normally used for the chromatography of peptides. Special attention has been put in miniaturization of columns through the reduction of column diameters and sorbent particles, as in the case of ultra performance LC systems (an LC system using sub-2 μm packing columns combined with high operating pressures), as well as the use of instruments capable of delivering nanoliter/min flow rates such us nano-scale liquid chromatography (nLC). On the other hand, CE combined with MS has also come out as an effective strategy for proteomic studies, since CE is a robust system that provides fast and high resolution separations employing inexpensive capillaries [123]. Special attention is currently being paid to the use of CE in microchips since it provides much shorter separation times [124], lower sample and reagent consumption than nLC and CE. Due to the analytical characteristics of these micro-devices further high-throughput in shotgun Proteomics is expected in the near future.

To carry out shotgun Proteomics most efforts have been focused on the development of on-line combination of various chromatography and/or electrokinetic separation methods, as well as in multidimensional protein identification technology (MudPIT), coupled with MS or tandem MS [125, 126] (**Figure 2**). The vast majority is based on the combination of ion exchange chromatography and reversed-phase (RP) chromatography. It has also been described the great potential of monolithic compared to particle-based columns due to monolithic can be used together with higher flows and gradients [127]. This is of special interest in a multidimensional chromatography setup where the

separation and identification of a vast number of analytes in a short time is mandatory mainly in the second dimension. Other novel couplings have been described, as for instance those based on hydrophilic interaction liquid chromatography (HILIC) [128]. The capabilities of multidimensional systems involving on-line combination of capillary isoelectric focusing (CIEF) with nano-RPLC have also been demonstrated to be effective in the analysis of proteins and peptides [129] creating new and encouraging perspectives in the discovery of biomarkers in Nutrigenomics studies.

ESI has been the interface of choice to ionize peptides eluted from LC or CE columns and to analyze them by MS: from the lower resolving power ion trap instruments able to generate MS^n spectra, to the more advanced hybrid instrumentation such us quadrupole-time of flight (Q-TOF), TOF-TOF and the recently introduced hybrid linear ion trap (LTQ)-Orbitrap (Orbitrap^{TM}) [130]. Some proteomic studies have also focused on the use of MALDI interface for chromatographic separations and MS. The technical problems generated by the non-atmospheric working conditions required by MALDI interface could be compensated for a lower interference with sample matrix and chromatographic eluents and an improvement on proteome coverage [131].

3.2 Top-down approach.

A growing number of researchers are focusing on the use of *top-down* Proteomics, a relatively new approach compared to *bottom-up*, in which structure of proteins is studied through measurement of their intact mass followed by direct ion dissociation in the gas phase [132]. The main advantages over the *bottom-up* approach are that higher sequence coverage is obtained, it permits the study of PTM and it makes possible to discern between biomolecules with a high degree of sequence identity.

Today, Fourier transform ion cyclotron resonance (FTICR) MS offers the highest mass resolution, resolving power, accuracy and sensitivity among present MS technologies [133]. Nevertheless, its high purchase, and maintenance costs, mostly derived from the expensive superconducting magnet and liquid helium supply required, precludes its general use. Thus, a variety of instruments have been

used for *top-down* Proteomics, such us MALDI-TOF/TOF, ESI-Q/TOF, ESI-IT and the novel Orbitrap^{TM}, suitable for routine *top-down* Proteomics as recently reported [134]. In those cases in which the molecular mass of proteins exceeds the analytical power of the mass spectrometer, they can be subjected to a limited proteolysis to produce large polypeptides that are then analyzed employing *top-down* Proteomic approaches. This has been called the *middle-down* approach, maintaining the advantages of high sequence coverage and PTM information.

A key to the *top-down* approach is the capacity to fragment intact proteins inside the mass spectrometer. Usually, low-energy multiple collision-induced dissociation (CID) is used, although ultraviolet photodissociation (UVPD) [135], infrared multiphoton dissociation (IRMPD) [136], and blackbody infrared radiative dissociation (BIRD) [137] can be found in literature. Alternative dissociation techniques such us electron capture dissociation (ECDβ) represents one of the most recent and significant advances in tandem MS [138]. This alternative technique is based on the fragmentation of multiple charged protein cations due to interaction with low-energy electrons. A newer dissociation method similar to ECD is electron transfer dissociation (ETD), which also uses charge reduction by electron transfer [139]. Both charge reduction dissociation techniques, ECD and ETD initially used in FTICR-MS and now extended to (LTQ)-Orbitrap instruments provide more extensive sequence fragments over the entire protein backbone preserving the PTM after cleavage, resulting in easier identification of the modification sites.

Top-down Proteomics approaches are usually limited to simple protein mixtures since very complex spectra are usually generated by multiple charged proteins. In general, a protein isolate from a previous fractionation or purification step is directly infused to the high-resolution MS. In this regard, due to the capabilities of CE for complex protein separation, CE coupling with high-resolution MS is one of the most promising methodology in *top-down* studies [140]. However, improved ECD and ETC speed will be needed for future high-throughput CE-MS (and also LC-MS) applications in *top-down* Proteomics.

3.3. Other proteomic approaches

New Proteomic approaches are under development, with a growing increase of works focused on the detection of protein biomarkers. One of the approaches that is gaining popularity is based on the array technology. Protein microarrays can be composed by recombinant protein molecules or antibodies immobilized in a high-density format on the surface of a substrate material. There are two major classes of protein micro- (or nano-) arrays: analytical and functional protein microarrays, being the antibody-based microarray the most common platform in proteomic studies [141]. This miniaturized arrays can be fabricated with an almost infinite number of antibodies (in the particular case of antibody microarrays) carrying the desired specificities and being capable of simultaneously profiling numerous low-abundant protein analytes in complex proteomes, while consuming only a few microliters of sample. The microarray patterns generated can then be transformed into proteomic maps, or detailed molecular fingerprints, revealing the composition of the proteome. After import data files from the microarray scanning sources and normalization and quantification of protein microarrays, a statistical analysis has to be carried out.

Surface enhanced laser desorption ionization time-of-flight mass spectrometry (SELDI-TOF MS), an emerging tool for protein profiling and biomarker discovery applicable to minute amounts of starting material [142] can also be used as a complementary, rapid and high-throughput proteomic approach to identify differentially expressed peptides and proteins and, therefore, can help to save time and research effort in studies about the effect of bioactive food components on the proteome. This technology enables to speed up the process significantly with the capacity to run and analyze several hundreds samples in parallel per day. Different SELDI chip types have different surfaces, depending on the type of proteins to be analyzed, range from chromatographic chemistries, to surfaces with a specific biomolecular affinity (e.g. antibodies, receptors, enzymes and ligands) that bind one specific protein or group of proteins. The efficacy of the SELDI-TOF technology has been proved in the analysis of plasma samples (previously fractionated by anion-exchange chromatography) by quickly

identifying the proteins whose expression changed in response to vitamin A status from retinol-sufficient and retinol-deficient rats [143] (**Figure 3**).

4. Advances in Metabolomics.

As defined by Trujillo *et al* [144], the metabolome can be described as the full set of endogenous or exogenous low molecular weight metabolic entities of approximately < 1000 Da (metabolites), and the small pathway motifs that are present in a biological system (cell, tissue, organ, organism or species). The most common metabolites are amino acids, lipids, vitamins, small peptides or carbohydrates. Metabolites are the real endpoints of gene expression and of any physiological regulatory processes. Therefore, according to some authors [12], any change in metabolite concentration may describe better the biochemical state of a biological system than proteomic or transcriptomic variations.

The objective of Metabolomics within the frame of Nutrigenomics is to investigate the metabolic alterations produced by the effect of nutrients or bioactive food constituents in the different metabolic pathways. Its importance not only lays on the information obtained about the molecular events involved in nutrition and how the body adapts through metabolic pathways to different nutrient fluxes, but also on the identification of certain metabolites such as cholesterol or glucose as biomarkers for health or disease status [145].

There are three basic approaches used in Metabolomic research; target analysis, metabolic profiling and metabolic fingerprinting. Target analysis aims the quantitative measurement of selected analytes, such as a specific biomarker or reaction product. Metabolic profiling is a non targeted strategy that focuses on the study of a group of related metabolites or a specific metabolic pathway. It is one of the basic approaches to phenotyping as the study of metabolic profiles of a cell gives a more accurate description of a phenotype [146]. Meanwhile, metabolic fingerprinting does not aim to identify all metabolites, but to compare patterns of metabolites that change in response to the cellular environment [147].

Metabolomics has diverse applications such as biomarker discovery, determination of the metabolic effects of environmental changes in the body or early disease detection [148]. The most relevant application of Metabolomics in Nutrigenomics is the possible health benefits provided by the ingest of functional compounds. In this regard, the effects of phytochemicals in human health are the most studied. Some examples are the potential benefits of flavones in heart diseases, of stannols in cholesterol metabolism, and soy-based estrogen analogues in cancer [149]. Because of the extensive consumption of polyphenols in the diet, and its association with health benefits, the biological activity of these compounds is an important topic of investigation in Nutrigenomics [150].

Application of Metabolomics in Nutrigenomics (or Nutritional Metabolomics) is, however, even more complex than in other areas of research. Humans feed on other organisms, each with its own metabolome, so the number of different metabolomes that conform our diet is significant enough to make Nutrigenomics a very complex discipline [151]. Therefore, there are several challenges that Metabolomics has to face in human nutrition. In general, metabolism is dynamic and depends on the different cellular environments and physiological situation, so it is hard to understand the effects separately and to link directly metabolites to genes and proteins [152]. In Nutritional Metabolomics the complex interactions between gut microflora and host metabolism as well as other extrinsic factors such as food habits, diet and other lifestyle parameters give rise to high metabolic variability [153-156].

One of the advantages of Nutritional Metabolomics is that there are several accessible body fluids that contain possible biomarkers in human body. To date, most Metabolomics studies in mammals have used urine as the vehicle for investigation and, in the field of Nutrigenomics blood and urine are the most likely candidates for sample choice [157].

4.1. Analytical techniques in Metabolomics.

Unlike Transcriptomics or Proteomics, which intend to determine a single chemical class of compounds (mRNA or proteins), Metabolomics has to deal with very different compounds of very diverse chemical and physical properties. Moreover, the relative concentration of metabolites in the biofluids vary from millimolar level (or higher) to picomolar, making it easy to exceed the linear range of the analytical technique employed. As no single technique can be expected to meet all these requirements, many Metabolomics approaches can employ several analytical techniques [12, 151].

Metabolomics is an emerging technology, so new analytical techniques and methods are continuously being developed and will continue in the near future in order to achieve its goals. So far, Metabolomics is proving to be very useful for the analysis of metabolic patterns and changes in the metabolism derived from different situations in the cellular environment. This is certainly interesting in the nutrition field as can determine variations in different metabolic pathways due to the consumption of different compounds in the diet.

In order to gain a detailed knowledge on human metabolism by means of Metabolomic tools, the ideal steps should be first to elucidate the nature and concentration of the searched metabolites and second to share the information in accessible databases [88] (**Figure 4**). The two most common analytical techniques used so far in Metabolomics are nuclear magnetic resonance (NMR) and MS. Thus, as a general result it can be stated that the influence of diet in metabolism has been clearly established using these techniques. For example, the effect of diet on urea cycle and purine metabolic pathways was demonstrated using both ^{1}H-NMR and ion trap MS [158]. Another study involving the use of both ^{1}H-NMR spectroscopy and high performance liquid chromatography (HPLC) coupled to a Q-TOF MS reported that acute changes in human urinary metabolomic profiles occur after the consumption of dietary phytochemicals [159].

4.1.1 Mass spectrometry (MS) approaches.

MS can be used as a stand alone technique or most commonly combined with a preliminary chromatographic separation technique, either gas chromatography (GC), high performance chromatography (HPLC) with MS (HPLC-MS) or CE-MS [151]. Direct injection is performed for Metabolomics with high or ultra high resolution mass analyzers as TOF-MS (mass accuracy < 10 ppm) or FTICR-MS which provides a mass accuracy < 1 ppm and detection limits lower than attomole or femtomole levels. These characteristics make them ideal tools for metabolomic studies as was shown by Rosello-Mora *et al* [160] which stated that different strains of *Salinibacter ruber* could be differentiated by some characteristic metabolites. Hybrid analyzers, as for example Q-TOF, have the advantages of mass accuracy given by a TOF analyzer combined with the possibility to fragment the ions and thus, provide information about the structure of the detected metabolites [161]. These kind of analyzers should lead the metabolomic studies in the near future. As for Proteomics, the most usual ionization techniques are the ones that use atmospheric pressure ionization, especially ESI interfaces, as it can work with a wide range of polarities [162].

Mass accuracy and resolution provided by MS instruments are usually not enough to undoubtedly separate and identify all metabolites in a sample by direct infusion. A complete methodology for metabolite elucidation is provided by Chen *et al.* [163] in the metabolomic identification of biomarkers of individuals at risk of diabetes. To overcome these limitations MS analyzers are usually coupled to separation techniques for metabolomic studies. Thus, GC-MS are so far the most common couplings for metabolomic studies. A GC-MS method for the analysis of flavonoids, a sub class of polyphenols, in urine and plasma has been developed [164]. In this method, liquid-liquid extraction, derivatization and analysis by GC-TOF-MS have been optimized, providing limits of detection below 0.1 μg/ml for most phenolic acids. This method identified 11 phenolic acids in urine samples whose concentration increased after tea consumption.

Apart of GC, other separation techniques have also been successfully coupled to MS instruments to carry out Nutritional Metabolomics. Thus, Yin *et al*. [165] have recently developed a reliable metabolic profiling method for human serum analysis using a C-18 column coupled to a Q-TOF-MS

analyzer. The aim of this LC-Q-TOF-MS strategy was to identify potential biomarkers and the metabolic pathways altered by an herbal preparation used to treat some diseases such as cardiovascular problems or cancer.

The most recent introduction into the field of separation techniques with enormous potential in Metabolomics research is the ultra performance liquid chromatography coupled to MS (UPLC[TM]-MS) technology. UPLC[TM]-MS has been used in the determination of metabolic profiles in human urine [166]. The small particle size and the high pressure allow to reduce analysis time approximately ten times respect to conventional LC while increasing efficiency and maintaining resolution.

4.1.2 Nuclear magnetic resonance (NMR) approaches.

NMR is a high reproducibility technique that has demonstrated its great potential in Metabolomics studies [167]. Using this analytical technique, Wang *et al.* have studied the relationship between the consumption of chamomile tea and some human biological responses. Statistical differences in three different excreted metabolites between high-resolution ^{1}H-NMR analyses of urine samples taken before and after chamomile tea consumption were found [168]. In another study, the metabolic profiling approach using high resolution ^{1}H-NMR spectroscopy has been applied to study metabolite changes in human feces by the intake of grape juice and ethanol-free wine extracts [169] (see **Figure 5**). This study showed changes in the levels of isobutyrate when the mixture of juice and wine was taken but could not find any difference in the metabolic profile due to the intake of juice. This could be explained through the modulation of microbial gut metabolism produced by the polyphenols present in wine. Metabolic profiling after dietary intervention with soy isoflavones has also been determined by ^{1}H-NMR differences in lipoprotein, amino acids and carbohydrate levels in plasma from five healthy pre-menopauseal women, thus suggesting a metabolic alteration due to soy consumption. [170]. In another ^{1}H-NMR study on the effect of soy isoflavones, an improvement in renal function and an increase in the urinary excretion of the osmolyte trimethylamine-N oxide has

been proven [171]. A summary of main Metabolomic applications in Nutrigenomics field are listed in **Table 3**.

Coupling of NMR with separation techniques is less extended than in MS, but can also give interesting results in Metabolomics. Offline coupling with solid-phase extraction (SPE) or on-line coupling with HPLC are the most extended in metabolomic research. For example HPLC-NMR has been used for the identification of hippuric acid, 3-(3-hydroxyphenyl) propionic acid, and 3-hydroxycinnamic acid in rat urine using reverse phase gradient HPLC coupled with ^{1}H-NMR [172].

The lack of sensitivity of NMR, in the range of μg, has been one of the main problems of this analytical technique. However, the recent introduction of new instrumentation, like [1]H-detection triple resonance cold probes, or different approaches such as dynamic nuclear polarization, have provided an improvement in detection sensitivity that will increase even further the number of applications of NMR in metabolomic studies [173].

4.2. Data analysis in Metabolomics.

Due to the huge amount of data obtained in metabolomic studies, it has been necessary to develop strategies to convert the complex raw data obtained into useful information. Metabolomic data reduction is normally accomplished using principal component analysis (PCA) or other related techniques. PCA is a tool for exploratory data analysis that determines correlation differences among sample sets, which can be caused by either a biological difference or a methodological bias. It is usually used as a first step to have information about the quality of the data. After data reduction a multivariate analysis is usually performed. The most common is partial least square discriminant analysis (PLS-DA). The aim of PLS-DA is to discriminate the complete peak list and reduce it with the most relevant ones [151].

Once an unknown compound is determined, the following challenge is to identify it and determine its biological significance [148]. For this matter, there are different large-scale databases available on the web. One of the most accessed databases available online is KEGG [174]. In particular for Nutritional Metabolomics the most interesting biochemical databases are the ones operated by the European Nutrigenomics Organization [175] that allows to upload and edit scientific information, and the Human Metabolome Database [176, 177], which provides information on more then 2180 metabolites from human. For all these reasons, Metabolomics has been considered in the NIH roadmap as an interesting tool for the overall initiative to carry out possible solutions to human metabolic diseases [88].

5. Systems Biology.

Nutrigenomics has to face important difficulties derived, among others, from food complexity, the large number of different nutrients and bioactive food compounds, their very different concentrations and the numerous targets with different affinities and specificities that they may have. As described above, Transcriptomics, Proteomics and Metabolomics represents powerful analytical platforms developed for the analysis of genes, proteins and metabolites. However, 'omics' platforms need to be integrated in order to obtain optimal means to understand the influences of bioactive food components on the investigated system (e.g., cell, tissue, organ) giving rise to the growing of a new area of biology called Systems Biology. Modern Systems Biology is the analysis of the relationships among the elements in a system in response to genetic or environmental perturbations, with the goal of understanding the system or the emergent properties on the system [178-180]. Thus, Systems Biology approaches may encompass molecules, cells, organs, individuals, or even ecosystems and it is regarded as an integrative approach of all information at the different levels of genomic expression (mRNA, protein, metabolite). However, in Nutrigenomic studies biologic responses to a bioactive food component may be subtle and, therefore, careful attention will need to be given to the methodologies used to identify these responses. Unlike any reductionist approach that would take these techniques individually, Systems Biology exploits global data sets to derive useful information

[181]. Each large data set contains sufficient noise to preclude the identification of multiple minor but relevant changes that could be unnoticed without adequate statistical tools since the researcher is focused on the changes that are really significant within the whole data set. Systems Biology, however, by confining the information can provide a filter for "distracting" noise generated in each individual platform and minimize the data to be interpreted by focusing on only those endpoints common between the various experimental platforms [178, 182]. To achieve this, appropriate statistical models have to be used in order to filter through the large data sets and highlight only those important changes. Although Systems Biology has been scarcely applied in Nutrigenomic studies, their potential is underlined by their adoption by other disciplines. For instance, a Systems Biology approach has been applied to improve our knowledge about carbohydrate metabolism in yeast [178]. In a recent work, Kohanski *et al.* used the context likelihood of relatedness (CLR) algorithm (gene network analysis) in combination with gene expression microarrays and Gene Ontology-based enrichment analysis to construct and filter gene connectivity maps of bacteria under antibiotic treatment [183]. The gene networks were further enriched with data derived from antibiotic growth high-throughput screening to provide insight into the pathway whereby the antibiotic under study triggers its bactericide action.

An adequate Systems Biology approach in Nutrigenomics should provide a holistic view of the molecular mechanisms underlying the beneficial or adverse effects of certain bioactive food components. Also, it should help in the discovery of key genes and proteins that function to regulate metabolic pathways and whose expression is affected by specific bioactive food compounds. This will aid in rapidly identifying new biomarkers for nutritional status and disease progression, and designing a novel concept for dietary prevention and intervention of disease [16].

6. Future needs and developments in Nutrigenomics.

Despite the use of DNA microarrays is a powerful analytical approach, there are technical limitations that have to be addressed for optimal implementation. Some of these problems are associated with the

high background noise that specially hinders the detection of low signals (i.e., low signal-to-noise ratios). Novel approaches focused on the use of electrochemical transducers in combination with either enzymatic, redox-active indicators, or nanoparticle labels, as well as with label-free hybridization strategies are being investigated as cheaper and sensitive alternatives to current optical detection systems [184, 185]. A second problem is related to the efficiency and specificity of the hybridization reaction that is subject to variability and cross-hybridization and thus, it affects the ability of the technique to accurately detect differences in gene expression levels [41]. In this regard, some interesting alternatives to typical linear probes have been proposed such as molecular beacon probes and peptide nucleic acids [185]. These probes offer high specificity and appear as good candidates for mismatch discrimination. Still, problems derived from element-to-element differences within a microarray and microarray-to-microarray differences need also to be solved [23, 186]. For that, the MicroArray Quality Control (MAQC) Consortium establishes quality control criteria to ensure data quality, to identify critical factors affecting data quality, and to optimize and standardize microarray procedures [187].

Although microarrays is currently the technique of choice for profiling RNA populations under different conditions, some advantages of RNA-Seq over existing technologies make this technique valuable for comprehensive transcriptome studies. First, unlike hybridization-based approaches, which rely on continuous signals, RNA-Seq consists of absolute numbers of reads that provide accurate estimates of the relative abundance of given transcripts and it is exempt from cross-hybridization problems [72, 188]. In addition, dynamic range of expression level in RNA-Seq depends of the sequencing depth total number of all the sequences reads or base pairs represented in the experiment) and it has been reported to be larger (up to five orders of magnitude) than the one provided by microarray scanners (few-hundredfold) [189]. Contrary to microarray, RNA-Seq has very low background signal because DNA sequences can be unambiguously mapped to unique regions of the genome, allowing detection of very low expressed mRNA, provided that sequencing depth is sufficient [190]. Another important advantage of RNA-Seq relative to gene expression microarray is their ability to identify, without prior knowledge, spliced transcript isoforms [191] and sequence

variations such as single nucleotide polymorphisms (SNPs) in the transcribed regions [66]. Despite of these advantages, RNA-Seq currently faces several challenges including those concerning library construction issues, bioinformatics and complete coverage at reasonable costs. RNA-Seq does not require cloning, however, the library construction involves some manipulation stages and procedures that could originate some bias [66]. Also, one of the main limitations of RNA-Seq stems from the short length of the sequence reads provided. For large and complex transcriptomes, mapping the short reads to the reference transcriptome can be complicated, especially if an important fraction of sequence reads match multiple locations in the genome. In such case, longer reads and paired-end read sequencing, based on sequencing both ends of each interrogated DNA fragment, have been proposed to help to alleviate the problem [72]. Sequencing depth is another important aspect that affects sequence coverage (percentage of transcripts surveyed), determines the number of expressed genes and rare spliced isoforms detected and is directly proportional to the sequencing costs [192].

Next-generation methods in Transcriptomics will undoubtedly continue to technically improve in several ways within the next years. New improvements will probably include the establishment of routine data analysis methods and increases in the numbers and lengths of sequence reads as well [193, 194]. Fortunately, the cost is likely to keep falling, allowing next-generation sequencing methods demonstrate full potential for the study of transcriptomes and provide new applications and extensive use of these technologies in Nutrigenomics research.

In Proteomics, 2DE, LC, CE and MS have become the most used methodologies. There is an evident need of development of improved or alternative technologies to become into a reality the routine analysis for proteome research. In this sense it has been proved the robustness of MudPIT for the analysis of complex mixtures of peptides; however improvements in the resolution of peptides are desirable to provide increased protein coverage. On the other hand, although much less employed than the methodologies used today in Proteomics, it is expected that relatively new protein microarray technology will play an important role in Proteomics in the near future due to the possibility to simultaneously analyze a large number of different proteins, providing unique opportunities to

establish interrelationship between dietary components and the development and/or progression of diseases. Apart of the everyday more sophisticated sample treatments and separation techniques, MS is essential for the systematic investigation in Proteomics. Rapid advances in MS instrumentation, both in terms of hardware and software, are improving current technologies and have catalyzed the development of new proteomic approaches. In this sense, conventional mass spectrometers are giving way to the more sophisticated and compact mass spectrometers, most of them hybrid instruments in a combination of two or more analyzers. As can be seen in the low number of proteomic applications in Nutrigenomic studies, it is expected that new innovations in proteomic technology will help proteomic profiling to become standard practice in the Nutrigenomic field.

In addition, a great advance in Metabolomics is expected related to the Nutrigenomics field with the incorporation of new interfaces such as desorption electrospray ionization (DESI) [195], SELDI [196], nanostructure-initiator mass spectrometry (NIMS) or with system miniaturization (nanoESI) [197]. These interfaces have the advantage that they need nearly no sample preparation. On the other hand, capillary electrokinetic techniques and their coupling to mass spectrometry (CE and CE-MS) are ideal tools for Metabolomics, due to the wide variety of applications, great efficiency and resolution, and low sample consumption. Although CE and CE-MS have not been widely used, they have already been identified as a very promising tool for metabolomic studies [161, 198, 199]. Thus, CE has been recently used to evaluate the response to vitamin E and D of diabetic rats metabolite urine profile with two different CE approaches (one based in micellar electrokinetic capillary chromatography (MEKC) and another using capillary zone electrophoresis (CZE)). With these new analytical methods it could be proved that the metabolic profile of diabetic rats fed with a mixture of these antioxidants trend to be similar to that from non diabetic control rats [200].

Comprehensive multidimensional techniques, such as GCxGC or LCxLC, are also a revolutionary improvement in separation techniques that will be implemented in Nutritional Metabolomics studies in the near future. They not only provide enhanced resolution and a huge increase in the peak number but also an increase in selectivity and sensitivity in comparison with conventional separation techniques.

As an example, comprehensive GCxGC coupled to TOF-MS has demonstrated to be a promising tool for metabolic profiling [201], and it has been successfully used for the detection of biomarkers in obese mouse tissue [202].

The challenge in Systems Biology approach is not on the technological level, as great improvements, as discussed here, are being made in the 'omics' technologies. Instead, it will be the bioinformatics side (data processing, clustering, dynamics, integration of the various 'omics' levels, etc.) that will have to progress for systems biology to mature and expand into Nutrigenomics.

REFERENCES

[1] R. Elliot, T.J. Ong, Br. J. Med. 324 (2002) 1438-1442.

[2] B. van Ommen, Nutrition 20 (2004) 4-8.

[3] L. Afman, M. Müller, J. Am. Diet. Assoc. 106 (2006) 569- 576.

[4] D. Evans, J.B. Hirsch, S. Dushenkov, J. Sci. Food Agric. 86 (2006) 2503-2509.

[5] N. Chalabi, D.J. Bernard-Gallon, M.P. Vasson, Y.J Bignon, Personalized Med. 5 (2008) 25-36

[6] T. Mizoguchi, I. Takehara, T. Masuzawa, T. Saito, Y. Naoki, J. Med. Food 11 (2008) 395-404.

[7] M.N. Mead, Environm. Health Persp. 115 (2007) 584-589.

[8] M. Müller, S. Kersten, Nat. Rev. Gen. 4 (2003) 315-322.

[9] Y. Liu, S.L. Sturley, Nutrition 20 (2004) 166-172.

[10] J. Kaput, J.M. Ordovas, L. Ferguson, B. van Ommen, R.L. Rodríguez, L. Allen, B.N. Ames, K. Dawson, B. German, R. Krauss, W. Malyj, M.C. Archer, S. Barnes, A. Bartholome, R. Bir, P. van Bladeren, K.J. Bradford, K.H. Brown, R. Caetano, D. Castle, R. Chadwick, S. Clarke, K. Clément, C.A. Cooney, D. Corella, I.B. Manica da Cruz, H. Daniel, T. Duster, S.O. Ebbesson, R. Elliott, S. Fairweather-Tait, J. Felton, M. Fenech, J.W. Finley, N. Fogg-Johnson, R. Gill-Garrison, M.J. Gibney, P.J. Gillies, J.A. Gustafsson, J.L. Hartman Iv, L. He, J.K. Hwang, J.P. Jais, Y. Jang, H. Joost, C. Junien, M. Kanter, W.A. Kibbe, B. Koletzko, B.R. Korf, K. Kornman, D.W. Krempin, D. Langin, D.R. Lauren, J. Ho Lee, G..A Leveille, S.J. Lin, J. Mathers, M. Mayne, W. McNabb, J.A. Milner, P. Morgan, M. Muller, Y. Nikolsky, F. van der Ouderaa, T. Park, N. Pensel, F. Perez-Jimenez, K. Poutanen, M. Roberts, W.H. Saris, G. Schuster, A.N. Shelling, A.P. Simopoulos, S. Southon, E.S. Tai, B. Towne, P. Trayhurn, R. Uauy, W.J. Visek, C. Warden, R. Weiss, J. Wiencke, J. Winkler, G.L. Wolff, X. Zhao-Wilson, J.D. Zucker, Br. J. Nutrition 94 (2005) 623-632.

[11] H.M. Roche, J. Sci. Food Agr. 86 (2006) 1156-1163.

[12] G. Rimbach, C. Boesch-Saadatmandi, J. Frank, D. Fuchs, U. Enzel, H. Daniel, W.L. Hall, P.D. Weinberg, Food Chem. Toxicol. 46 (2008) 1308-1319.

[13] R. Hu, A.T. Kong, Nutrition 20 (2004) 23-88.

[14] J.M. Ordovás, J. Kaput, D. Corella, Mol. Nut. Food Res. 51 (2007) 1293-1299.

[15] P. Anderle, P. Farmer, A. Berger, M.A. Roberts, Nutrition 20 (2004) 103-108.

[16] B. van Ommen, R. Stierum, Food Biotechnol. 13 (2002) 517-521.

[17] L.R. Ferguson, M. Philpott, P. Dryland, Cell. Mol. Life Sci. 63 (2007) 3105.

[18] C.D. Davis, J. Milner, Mut. Res. 554 (2004) 51-64.

[19] K. Gohil, A.A. Chakraborty, Nutrition 20 (2004) 50-55.

[20] O. Morozova, M.A. Marra, Genomics 92 (2008) 255-264.

[21] S. Gaj, L. Eijssen, R.P. Mensink, C.T. Evelo, Genes Nutr. 3 (2008) 153-157.

[22] J.J. Storhoff, S,S. Marla, V.G., C.A. Mirkin in: U.R. Müller, D.V. Nicolau (Eds.), Microarray Tecnology and its applications. Springer-Verlag, Berlin, Heidelgerg, 2005, 147-180.

[23] B. Spielbauer, F. Stahl, Mol. Nut. Food Res. 49 (2005) 908-917.

[24] M.D. Niculescu, E.A. Pop, L.M. Fischer and S.H. Zeisel, J. Nutr. Biochem. 18 (2007) 380–390.

[25] P. Stafford, Y. Tak in: P. Safford (Ed.), Methods in microarray normalization, CRC Press, Boca Raton, FL, 2008, pp. 151-203.

[26] http://www.mged.org/

[27] http://www.ebi.ac.uk/arraexpress/

[28] http://www.ebi.ac.uk/microarray-as/atlas/

[29] H. Parkinson, M. Kapushesky, N. Kolesnikov, G. Rustici, M. Shojatalab, N. Abeygunawardena, H. Berube, M. Dylag, I. Emam, A. Farne, E. Holloway, M. Lukk, J. Malone, R. Mani, E. Pilicheva, T.F. Rayner, f. Rezwan, A. Sharma, E. Williams, X.Z. Bradley, T. Adamusiak, M. Brandizi, T. Burdett, R. Coulson, M. Krestyaninova, P. Kurnosov, E. Maguire, S.G. Neogi, P. Rocca-Serra, S.A. Sansone, N. Sklyar, m. Zhao, U. Sarkans, A. Brazma, Nucleic Acids Res. 37 (2009) D868-D872.

[30] G. Burton, R.E. McGehee, Nutrition 20 (2004) 109-114.

[31] Dos Santos in: S.Q. Ye (Ed.), Bioinformatics. A practical approach, Chapman & Hall/CRC, Boca Raton, FL, 2008, pp. 131-187.

[32] S. Klaus, J. Keijer, Nutrition 20 (2004) 115-120.

[33] F.A. Middleton, E.J. Ramos, Y.Xu. H. Diab, X. Zhao, U.N. Das, M. Menguid, Nutrition 20 (2004) 14-25.

[34] R.D. Canales, Y. Luo, J.C. Willey, B. Austermiller, C.C. Barbacioru, C. Boysen, K. Hunkapiller,

R.V. Jensen, C.R. Knight, K.Y. Lee, Y. Ma, B. Maqsodi, A. Papallo, E.H. Peters, K. Poulter, P.L. Ruppel, R.R. Samaha, L. Shi, W. Yang, L. Zhang, F.M. Goodsaid, Nat. Biotechnol. 24 (2006) 1115-1122.

[35] R. Vittal, Z.E. Selvanayagam, Y. Sun, J. Hong, F. Liu, K.V. Chin, C.S. Yang, Mol. Cancer. Ther. 3 (2004) 1091–1099.

[36] V. Noé, S. Peñuelas, R. M. Lamuela-Raventós, J. Permanyer, C.J. Ciudad, M. Izquierdo-Pulido, J. Nutrition 134 (2004) 2509-2516.

[37] T. Tsuda, Y. Ueno, T. Yoshikawa, H. Kojo, T. Osawa, Biochem. Pharmacol. 71 (2006) 1184-1197.

[38] H.X. Leong, C. Simkevich, A. Lesieur-Brooks, B.W. Lau, C. Fugere, E. Sabo Nutr. Metab. 3 (2006) 37.

[39] R.K. Thimmulappa, K.H. Mai, S. Srisuma, T.W. Kensler, M. Yamamoto, S. Biswal, Cancer Res. 62 (2002) 5196–5203.

[40] P. McLoughlin, M. Roengvoraphoj, C. Gissel, J. Hescheler, U. Certa, A. Sachinidis, Genes Cells 9 (2004) 661–669.

[41] K. Mills, in: G. Rimbach, J. Fuchs, L. Packer (Eds.), Nutrigenomics, CRC Press, Boca Raton, FL, 2005, pp. 43-66.

[42] C.K. Lee, R.G. Klopp, R. Weindruch, R.A. Prolla Science 285 (1999) 1390-1393.

[43] T. Kayo, D.B. Allison, R. Weindruch, T.A. Prolla, Proc. Natl. Acad. Sci. USA (2001) 5093-5098.

[44] R. Weindruch, T. Kayto, C. Lee, T.A. Prolla, J. Nutr. 131 (2001) 918S-923S.

[45] S.X. Cao, J.M. Dhabbi, P.L. Spindler. Proc. Natl. Acad. Sci. USA 98 (2001) 10630-10635.

[46] H. Kato, T. Kimura, J. Nutrition 133 (2003) 2073S-2077S.

[47] Y. Endo, Z.W. Fu, K. Abe, S. Arai, H. Kato J. Nutr. 132 (2002) 3632-3637.

[48] A. Berger, D.M. Mutch, J.B. German, M.A. Roberts, Lipids in Health and Disease 1 (2002) 2.

[49] B.A. Narayanan, N.K. Narayanan, B. Simi, and B.S. Reddy, Cancer Res. 62 (2003) 972-979.

[50] M.J. van Erk, W.A. Blom, V. Van Ommen, H. F. Hendriks, Am. J. Clin. Nutr. 84 (2006) 1233-1241.

[51] L.M. Sparks, H. Xie, R.A. Koza, R. Mynatt, G.A. Bray and S.R. Smith, Metabolism 55 (2006) 1457–1463.

[52] Y. Naito, K. Uchiyama, K. Mizushima, M. Kuroda, S. Akagiri, T. Takagi, O. Handa, S. Kokura, N. Yoshida, H. Ichikawa, J. Takahashi, T. Yoshikawa, Int. J. Mol. Med. 18 (2006) 685-695.

[53] S.I. Wang, H. Mukhar, Cancer Lett. 182 (2002) 43–51.

[54] L.G. Puskás, Z.B. Nagya, Z. Giricz, A. Onody, C. Csonkab, K. Ktajkac, L. Hackler, A. Zvara, P. Ferdinandy, FEBS Lett. 562 (2004) 99-104.

[55] D. Parra, A.B. Crujeiras, F.I. Milagro, E. Goyenechea, E. Larrarte, J. Margareto, J.A. Martínez, OMICS 12 (2008) 1-11.

[56] L. Murtaza, G. Marra, R. Schlapbach, A. Patrignani, M. Kunzli, U. Wagner, Biotechnol. Appl. Biochem. 45 (2006) 29–36.

[57] Velculescu, Zhang, Bofelstein, Science 270 (1995) 484-487.

[58] S. Saha, A. B. Sparks, C. Rago, V. Akmaev, C. J. Wang, B. Vogelstein, D.W. Kinzler, V. E. Velculescu, Nat. Biotechnol. 19 (2002) 508-512.

[59] T. Shiraki, S. Katayama, D. Waki, T. Kasukawa, H. Kawaji, R. Kodzius, A. Watahiki, M. Nakamura, T. Arakawa, S. Fukuda, D. Sasaki, Proc. Natl. Acad. Sci U.S.A. 100 (2003) 15776-15781.

[60] P. Ng, C.L. Wei, W.K. Sung, K.P. Chiu, L. Lipovich, C.C. Ang, S. Gupta, A. Shahab, A. Ridwan, C. Hong, Nat. Methods 2 (2005) 105-111.

[61] S.M. Wang, Trends Genet. 23 (2006) 42-50.

[62] S. Brenner, M. Johnson, J. Bridgham, G. Golda, D. H Lloyd, D. Johnson, S. Luo, S. McCurdy, M. Foy, M. Ewan, R. Roth, D. George, S. Eletr, G. Albrecht, E. Vermaas, S.R. Williams, K. Moon, T. Burcham, M. Pallas, R.B. DuBridge, J. Kirchner, K. Fearon, J. Mao, K. Corcoran, Nat. Biotechnol. 18.

[63] R. Tuteja in: S.Q. Ye (Ed.), Bioinformatics. A practical approach, Chapman & Hall/CRC, Boca Raton, FL, 2008, pp. 189-218.

[64] K. L. Nielsen, A. L Hogh, J. Emmergen, Nucl. Acids Res. 34 (2006) e133.

[65] E. R. Mardis, Annu. Rev. Genomics Hum. Genet 9 (2008) 387-402.

[66] Z. Wang, M. Gerstein, M. Snyder, Nat. Rev. 10 (2009) 57-63.

[67] M. Margulies, M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemden, J. Berka, M.S. Braverman, Y.J. Chen, Z. Chen, S.B. Dewell, L. Du, J.M. Fierro, X.V. Gomes, B.C. Godwin, W. He, S. Helgesen, C.H. Ho, G.P. Irzyk, S.C. Jando, M.L. Alenquer, T.P. Jarvie, K.B. Jirage, J.B. Kim, J.R. Knight, J.R. Lanza, J.H. Leamon, S.M. Lefkowitz, M. Lei, J. Li , K.L. Lohman, H. Lu, V.B. Makhijani, K.E. McDade, M.P. McKenna, E.W. Myers, E. Nickerson, J.R. Nobile, R. Plant, B.P. Puc, M.T. Ronan, G.T. Roth, G.J. Sarkis, J.F. Simons, J.W. Simpson, M. Srinivasan, K.R. Tartaro, A. Tomasz, K.A. Vogt, G.A. Volkmer, S.H. Wang, Y. Wang, M.P. Weiner, P. Yu, R.F. Begley, J.M. Rothberg, Nature 437 (2005) 376-380.

[68] S.T. Bennett, C. Barnes, A. Cox, L. Davies, C. Brown, Pharmacogenomics 6 (2005) 373-382.

[69] J. Schendure, G.J. Porreca, N.B. Reppas, X. Lin, J.P. McCutcheon, A.M. Rosenbaum, M.D. Wang, K. Zhang, R.D. Mitra, G.M. Church, Science 309 (2005) 1728-1732.

[70] S. Marguerat, B. T. Wilhelm, J. Bahler, Biochem. Soc. Trans. 36 (2008) 1091-1096.

[71] C.A. Hutchison, Nucl. Acids Res. 35 (2007) 6227-6237.

[72] B. Wold, R.M. Myers, Nat. Methods 5 (2008) 19-21.

[73] J.B. Kim, G.J. Porreca, L. Song, S.C. Greenway, J.M. Gorham, G.M. Church, C.E. Seidman, J.G. Seidman, Science 316 (2007) 1481-1484.

[74] V.E. Velculescu, K.W. Kinzler, Nat. Biotechnol. 25(2007) 878-880.

[75] T.T. Torres, M. Metta, B. Ottenwalder, C. Schlotterer, Genome Res. 18 (2008) 172-177.

[76] M.N. Bainbridge, R.L. Warren, M. Hirst, T. Romanuik, T. Zeng, A. Go, A. Delaney, M. Griffith, M. Hickenbotham, V. Magrini, BMC Genomics 7 (2006) 246.

[77] A.L. Toth, K. Varala, T.C. Newman, F.E. Miguez, S.K. Hutchison, D.A. Willoughby, J.F. Simons, M. Egholm, J.H. Hunt, M.E. Hudson, G.E. Robinson, Science 318 (2007) 441-444.

[78] C. Lu, K. Kulkarni, F.F. Souret, R. MuthuValliappan, S.S. Tej, R.S. Poethig, I.R. Henderson, S.E. Jacobsen, W. Wang, P.J. Green, B.C. Meyers, Genome Res. 16 (2006) 1276-1288.

[79] J.G. Ruby, C. Jan, C. Player, M.J. Axtell, W. Lee, C. Nusbaum, H. Ge, D.P. Bartel, Cell 127 (2006) 1193-1207.

[80] A. Stark, N. Bushati, C.H. Jan, P. Kheradpour, E. Hodges, J. Brennecke, D.P. Bartel, S.M. Cohen, M. Kellis, Genes Dev. 22 (2008) 8-13.

[81] A.S. Waagmeester, T. Kelder, C.T.A. Evelo, Genes Nutr. 3 (2008) 139-142.

[82] D.W. Huang, B.T. Sherman, R.A. Lempicki, Nucleic Acids Res. 37 (2009) 1-13.

[83] The Gene Ontology Consortium, Nucleic Acids Res. 36 (2008) D440-D444.

[84] The Gene Ontology Consortium, Nat. Gen. 25 (2000) 25-29.

[85] J.J. Moresco, M-Q. Dong, J.R. Yates III, Am. J. Clin. Nutr. 88 (2008) 597-604.

[86] J. Ovesná, O. Slabý, O. Toussaint, M. Kodicek, P. Marsík, V. Pouchová, T. Vanek, Brit. J. Nutr. 99 (2008) ES127-ES134.

[87] F.J. Schweigert, Ann. Nutr. Metab. 51 (2007) 99-107.

[88] J. Astle, J.T. Ferguson, J.B. German, G.G. Harrigan, N.L. Kelleher, T. Kodadek, B.A. Parks, M.J. Roth, K.W. Singletary, C.D. Wenger, G.B. Mahady, J. Nutr. 137 (2007) 2787-2793.

[89] D. Fuchs, I. Winkelmann, I.T. Johnson, E. Mariman, U. Wenzel, H. Daniel. Brit. J. Nutr. 94 (2005) 302-314."

[90] H. Kim, G.P. Page, S. Barnes, Nutrition 20 (2004) 155-165.

[91] B. de Roos, H.J. McArdle, Brit. J. Nutr. 99 (2008) S66-S71.

[92] L.M. Knowles, J.A. Milner, Carcinogenesis 21 (2000) 1129-1134.

[93] A.T. Dinkova-Kostova, Mini Rev. Med. Chem. 2 (2002) 595-610).

[94] N.L. Anderson, N.G. Anderson, Mol. Cell. Proteomics 1 (2002) 845-867.

[95] X. Fang, W-W. Zhang, J. Proteomics 71 (2008) 284-303.

[96] M. Pernemalm, L.M. Orre, J. Lengqvist, P. Wikström, R. Lewensohn, J. Lehtiö. J. Proteome Res. 7 (2008) 2712-2722.

[97] L.A. Huber, K. Pfaller, I. Vietor, Circ. Res. 92 (2003) 962-968.

[98] A. Leitner, W. Lindner, Proteomics 6 (2006) 5418-5434.

[99] B. de Roos, S.J. Duthie, A.C.J. Polley, F. Mulholland, F.G. Bouwman, C. Heim, G.J. Rucklidge, I.T. Johnson, E.C. Mariman, H. Daniel, R.M. Elliott. J. Proteome Res. 7 (2008) 2280-2290.

[100] P. Lescuyer, D.F. Hochstrasser, J-C. Sanchez, Electrophoresis 25 (2004) 1125-1135.

[101] M. Fountoulakis, P. Dimitraki in: A. Posch (Ed.), 2D PAGE: Sample preparation and fractionation, Humana Press, Totowa, NJ, 2008 pp 301-313.

[102] E. Boschetti, P.G. Righetti, J. Proteomics 71 (2008) 255-264.

[103] C.H. Chen, Anal. Chim. Acta. 624 (2008) 16-36.

[104] X. Han, A. Aslanian, J.R Yates III, Curr. Opin. Chem. Biol. 12 (2008) 483-490.

[105] B. Wittmann-Liebold, H-R. Graack, T. Pohl. Proteomics 6 (2006) 4688-4703.

[106]R. Marouga, S. David, E. Hawkins, Anal. Bioanal. Chem. 382 (2005) 669-678.

[107] M.F. Mouat, K. Kolli, R. Orlando, J.L. Hargrove, A. Grider, Nutr. J. 4 (2005), 1-7.

[108] B. de Roos, I. Duivenvoorden, G. Rucklidge, M. Reid, K. Ross, R-J.A.N. Lamers, P.J. Voshol,

L.M. Havekes, B. Teusink, FASEB J. 19 (2005) 813-815

[109] D. Fuchs, B. Dirscher, J.H. Schroot, H. Daniel, U. Wenzel, Mol. Nutr. Food. Res. 50 (2006), 58-

69.

[110] D. Fuchs, S. de Pascual-Teresa, G. Rimbach, F. Virgili, R. Ambra, R. Turner, H. Daniel, U.

Wenzel, Eur. Nutr. 44 (2005) 95-104.

[111] D. Fuchs, P. Erhard, R. Turner, G. Rimbach, H. Daniel, U. Wenzel, J. Proteome Res. 4 (2005)

369-376.

[112] D. Fuchs, P. Erhard, G. Rimbach, H. Daniel, U. Wenzel, Proteomics 5 (2005) 2808-2818.

[113] D. Fuchs, K. Vafeiadou, W.L. Hall, H. Daniel, C.M. Williams, J.H, Schroot, U. Wenzel, Am. J.

Clin. Nutr. 86 (2007) 1369-1375.

[114] W.Y. Lin, C-Y. Song, T-M. Pan, J. Agric. Food Chem. 54 (2006) 6192-6200.

[115] H-S. Seo, M. Hirano, J. Shibato, R. Rawkal, I.K. Hwang, Y. Masuo, J. Agric. Food. Chem., 56

(2008) 4665-4673.

[116] Q-Y. Lu, Y.S. Jin, Z-F. Zhang, A.D. Le, D. Heber, F.P. Li, S.M. Dubinett, J.Y. Rao, Lab. Invest.

87 (2007) 456-465.

[117] G. Breikers, S.G.J. van Breda, F.G. Bouwman, M.H.M. van Herwijnen, J. Renes, E.C.M.

Mariman, J.C.S. Kleinjans, J.H.M. van Delft, Proteomics 6 (2006) 2844-2852.

[118103] M.C.Y. Wong, P.W. Emery, V.R. Preedy, H. Wiseman, Inflammopharmacology 16 (2008)

235-239.

[119] A.I. Nesvizhskii, R. Aebersold, Mol. Cell. Prot. 4 (2005) 1419-1440.

[120] L. Hu, M. Ye, X. Jiang, S. Fng, H. Zou, Anal. Chim. Acta 598 (2007) 193-204.

[121] M. Herrero, E. Ibañez, A. Cifuentes, Electrophoresis 29 (2008) 2148-2160.

[122] V. Kasicka, Electrophoresis 29 (2008) 179-206.

[123] E. Schiffer, H. Mischak, J. Novak, Proteomics 20 (2006) 5615-5627.

[124] V. Dolnik, S. Liu, J. Sep. Sci. 28 (2005) 1994-2009.

[125] M.L. Fournier, J.M. Gilmore, S.A. Martin-Brown, M.P. Washburn, Chem. Rev. 107 (2007) 3654-3686.

[126] Motoyama, J.R. Yates III, Anal. Chem. 80 (2008) 7187-7193.

[127] R. Wu, L. Hu, F. Wang, M. Ye, H. Zou, J. Chromatogr. A 1184 (2008) 369-392.

[128] P.J. Boersema, S. Mohammed, A.J.R. Heck, Anal. Bioanal. Chem. 391 (2008) 151-159.

[129] T. Guo, C.S. Lee, W. Wang, D.L. de Voe, B.M. Balgley, Electrophoresis 27 (2006) 3523-3532.

[130] Q. Hu, R.J. Noll, H. Li, A. Makarov, M. Hardmanc, R.G. Cooks, J. Mass Spectrom. 40 (2005) 430-443.

[131] Y. Yang, S. Zhang, K. Howe, D.B. Wilson, F. Moser, D. Irwin, T.W. Thannhauser, J. Biomol. Tech. 18 (2007) 226-237.

[132] K. Breuker, M. Jin, X. Han, H. Jiang, F.W. McLafferty, J. Am. Soc. Mass Spectrom. 19 (2008) 1045-1053.

[133] B. Bogdanov, R.D. Smith, Mass. Spec. Rev. 24 (2005) 168-200.

[134] B. Macek, L.F. Waanders, J.V. Olsen, M. Mann, Mol. Cell. Prot. 5 (2006) 949-958.

[135] J.W. Morgan, J.M. Hettick, D.H. Russell, 402 (2005) 186-209.

[136] S.A. Raspopov, A. El-Faramawy, B.A. Thomson, K.W.M. Siu, Anal. Chem. 78 (2006), 4572-4577.

[137] Y. Gea, D.M. Horna, F.W. McLafferty, Int. J. Mass Spectrom. 210-211 (2001) 203-214.

[138] S.M. Sweet, H.J. Cooper, Expert Rev. Proteomics 4 (2007) 149-59.

[139] J. Wiesner, T. Premsler, A. Sickmann, Proteomics 8 (2008) 4466-4483.

[140] D.C. Simpson, S. Ahn, L. Pasa-Tolic, B. Bogdanov, H.M. Mottaz, A.N. Vilkov, G.A. Anderson, M.S. Lipton, R.D. Smith, Electrophoresis 27 (2006) 2722-2733.

[141] C.A.K. Borrebaeck, C. Wingren, Expert Rev. Mol. Diagn. 7 (2007) 673-686.

[142] T.C.W. Poon, Expert Rev. Proteomics 4 (2007) 51-65.

[143] T. Linke, A.C. Ross, E.H. Harrison, J. Chromatogr. A 1043 (2004) 65-71.

[144] E. Trujillo, C. Davis, J. Millner, J. Am. Diet. Assoc. 106, 3 (2006) 403-413.

[145] M. Kussman, S. Rezzi, H. Danelore, Curr. Opin. Biotech. 19 (2008) 83-99.

[146] Wai-Nang, P. Lee, V. L. Go, J. Nutr. (2005) 3027S-3032S.

[147] O. Fiehn, Plant Mol. Biol. 48 (2002) 155-171.

[148] M.C. Walsh, A. Nugent, L. Brennan, M.J. Gibney, Nutr. Bull. 33 (2008) 316-323.

[149] A. Cassidy, F.S. Dalais, in: M.J. Gibney, I. Macdonald, H.M. Roche (Eds.), Nutrition & Metabolism, Oxford, United Kingdom, 2003, pp307-317.

[150] J.D. Lambert, J. Hong, G. Yang, J. Liao, C. Yang., Am. J. Clin. Nutr. 81 (2005) 284S-291S.

[151] A. Koulman, D.A. Volmer, Nutr. Bull. 33 (2008) 324-330.

[152] S. Rezzi, Z. Ramadan, L.B. Fay, S. Kochhar, J. Prot. Res. 6 (2007) 513-525.

[153] J.K. Nicholson, E. Holmes, J.C. Lindon, I.D. Wilson, Nat. Biotechnol, 22 (2004) 1268-1274.

[154] R. Goodacre, J. Nutr. 137 (2007) 259S-266S.

[155] M.C. Walsh, L. Brennan, J. Malthouse, H.M. Roche, M.J. Gibney, Am. J. Clin. Nutr. 84 (2006) 531-539.

[156] X. Zhang, Y. Yap, D. Wei, G. Chen, F. Chen, Biotechnol. Adv. 26 (2008) 169-176.

[157] M.J. Gibney, M. Walsh, L. Brennan, H.M. Roche, B. German, B. Van Ommen, Am. J. Clin. Nutr. 82 (2005) 497-503.

[158] H. Gu, H. Chen, Z. Pan, A. U. Jackson, N. Talaty, B. Xi, C. Kissinger, C. Duda, D. Mann, D. Raftery, R. G. Cooks, Anal. Chem. 79 (2007) 89-97.

[159] M.C Walsh, L. Brennan, E. Pujos-Guillot, J-L. Sébédio, A. Scalbert, A. Fagan, D.G Higgins, M.J. Gibney, Am. J. Clin. Nutr. 86 (2007) 1687-1693.

[160] R. Roselló-Mora, M. Lucio, A. Pea, J. Brito-Echeverría, A. López-López, M. Valens-Vadell, M. Frommberger, J. Antón, P. Schmitt-Kopplin, ISME J. 2 (2008) 242-253.

[161] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass Spectrom. Rev. 26 (2007) 51-78.

[162] E. Werner, J-F. Heilier, C. Ducruix, E. Ezan, C. Junot, J-C. Tabet, J. Chromatogr. B 871 (2008) 143-163.

[163] J. Chen, X. Zhao, J. Fritsche, P. Yin, P. Schmitt-Kopplin, W. Wang, X. Lu, H.U. Haring, E.D. Schleicher, R. Lehmann, G. Xu, Anal. Chem. 80 (2008) 1280-1289.

[164] C.H. Grün, F.A. van Dorsten, D.M. Jacobs, M. Le Belleguic, E.J. van Velzen, M.O. Bingham, H-G. Janssen, J.P. van Duynhoven, J. Chromatogr. B 871 (2008) 212-219.

[165] P. Yin, P. Mohemaiti, J. Chen, X. Zhao, X. Lu, A. Yimiti, H. Upur, G. Xu J. Chromatogr. B 871 (2008) 322-327.

[166] P.A. Guy, I. Tavazzi, S.J. Bruce, Z. Ramadan, S. Kochhar, J. Chromatogr. B. 871 (2008) 253-260.

[167] M.E. Dumas, E.C. Maibaum, C. Teague, H. Ueshima, B. Zhou, J.C. Lindon, J.K. Nicholson, J. Stamler, P. Elliot, Q. Chan, E. Holmes, Anal Chem. 78 (2006) 2199-2208.

[168] Y. Wang, H. Tang, J.K. Nicholson, P.J. Hylands, J. Sampson, E. Holmes., J. Agric. Food Chem. 53 (2005) 191-196.

[169] D.M. Jacobs, N. Deltimple, E. van Velzen, F.A. van Dorsten, M. Bingham, E.E. Vaughan, J. van Duynhoven, NMR Biomed. 21 (2008) 615-626.

[170] K.S. Solanky, N.J. Bailey, B.M. Beckwith-Hall, A. Davis, S. Bingham, E.Holmes, J.K. Nicholson, A. Cassidy, Anal. Biochem. 323 (2003) 197-204.

[171] K.S. Solanky, N.J. Bailey, B.M. Beckwith-Hall, S. Bingham, A. Davis, E. Holmes, J.K. Nicholson, A. Cassidy, J. Nutr. Biochem. 16 (2005) 236-244.

[172] C.L Gavaghan, J.K. Nicholson, S.C. Connor, I.D. Wilson, B. Wright, E. Holmes, Anal. Biochem. 291 (2001) 2245-2252.

[173] T.W-M. Fan, A.N. Lane, Prog. Nucl. Mag. Res. Sp. 52 (2008) 69-117.

[174] M. Kanehisa, S. Goto, Nucleic Acids Res. 28 (2000) 27-30.

[175] http://www.nugowiki.org

[176] http://www.hmdb.ca

[177] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. MacInnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, Nucleic Acids Res. 35 (2007) D521-D526.

[178] A.D. Weston, L. Hood, J. Proteome Res. 3 (2004) 179-196.

[179] L. Hood, R.M. Perlmutter, Nat. Biotechnol. 22 (2004) 1215-1216.

[180] I. Rigoutsos in: I. Rigoutsos (Ed.), Systems Biology, Oxford University Press, Cary, NC, USA, 2006, pp xiii-xxi.

[181] X. Feng, X. Liu, Q. Luo, B.F. Liu, Mass Spectrom. Rev. 27 (2008) 635-660.

[182] D.M. Mutch, W. Wahli, G. Williamson, FASEB J. 12 (2005) 1602-1616.

[183] M.A. Kohanski, D.J. Dwyer, J. Wierzbowski, G. Cottarel, J.J. Collins, Cell 135 (2008) 679-690.

[184] B.J. Privett, J.H. Shin, M.H. Schoenfisch, Anal. Chem. 80 (2008) 4499-4517.

[185] A. Sassolas, B.D. Leca-Bouvier, L.J. Blum, Chem. Rev. 108 (2008) 109-139.

[186] J.A. Stamatoyannopoulos, Genomics 84 (2004) 449-457.

[187] D.A. Casciano, J. Woodcock, Nat. Biotechnol. 24 (2006) 113.

[188] P.A. Hoen, Y. Ariyurek, H.H. Thygesen, E. Vreugdenhil, R.A. Vossen, R.X. de Menezes, J.M. Boer, G.B van Ommen, J.T. den Dunnen, Nucleic Acids Res. 36 (2008) e141.

[189] A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, B. Wold, Nat. Methods 5 (2008) 621-628.

[190] B.T. Wilhelm, S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C.J. Penkett, J.Rogers, J. Bahler, Nature 453 (2008) 1239-1243.

[191] M. Sultan, M.H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, D. Schmidt, Science 321 (2008) 956-960.

[192] C.P. Fredlake, D.G. Hert, E.R. Mardis, A.E. Barron, Electrophoresis 27 (2006) 3689-3702.

[193] M.A. Quail, I. Kozarewa, F. Smith, A. Scally, P.J. Stephens, R. Durbin, H. Swerdlow, D.J. Turner, Nat. Methods 5 (2008) 1005-1100.

[194] T. Maricic, S. Paabo, BioTechniques 46 (2009) 51-57.

[195] H. Chen, Z. Pan, N. Talaty, D. Raftery, R.G. Cooks, Rapid Commun. Mass Spectrom. 20 (2006) 1577-1584.

[1

-                                    tter, E. Holmes, S. Bahn, Plos ONE 2 (2007) e756.

[197] X. Feng, X. Liu, Q. Luo, B-F. Liu, Mass Spectrom. Rev. 27 (2008) 635-660.

-                         -   -

-        , A. Cifuentes, J. Chromatogr. A 1195 (2008) 164-173.

-        , A. Cifuentes, Anal. Chem. 80 (2008) 6329-6335.

[200] C. Barbas, M. Vallejo, A. García, D. Barlow, M. Hanna-Brown., J. Pharmac. Biomed. Anal. 47 (2008) 388-398.

[201] K.K. Pasikanti, P.C. Ho, E.C. Chan, J. Chromatogr. B. 871 (2008) 202-211.

[202] R.A. Shellie, W. Welthagen, J. Zrostliková, J. Spranger, M. Ristow, O. Fiehn, R. Zimmermann. J. Chromatogr. A 1086 (2005) 83-90.

FIGURE LEGENDS

Figure 1. Scanned image of a microarray composed of a total of 250 kinases and phosphatases showing the expression profile obtained from human prostate carcinoma LNCaP cells treated with 12 mM epigallocatechin gallate (EGCG) or water-only for 12 h. Redrawn from [53].

Figure 2. Workflow of multidimensional separation-MS/MS-based shotgun proteomics strategy. N is the sample fraction. Redrawn from [126].

Figure 3. SELDI–TOF mass spectra representing protein expression differences in plasma of vitamin A-deficient (n=3) and vitamin A-sufficient (n=3) rats at three different m/z: 10693, 15203 and 18720. Bold line, vitamin A-sufficient Mr profile; dotted line, vitamin A-deficient Mr profile. Redrawn from [143].

Figure 4. Metabolomics workflow in Nutrigenomics research.

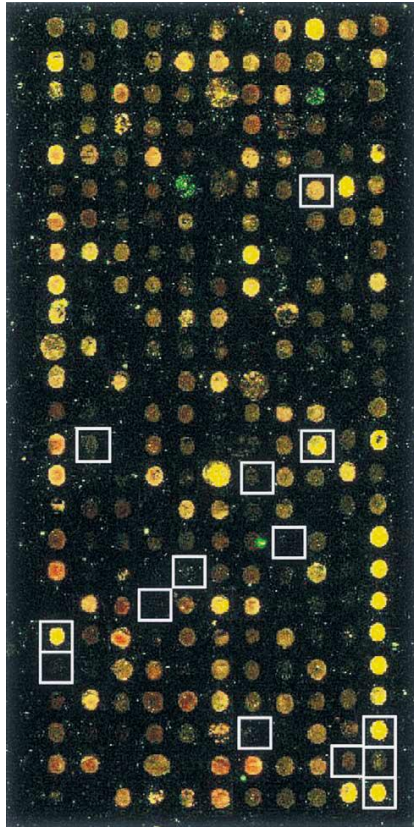**Figure 5.** NMR profiles of faeces from different subjects using two different extraction methods (water and methanol). Redrawn from [169].
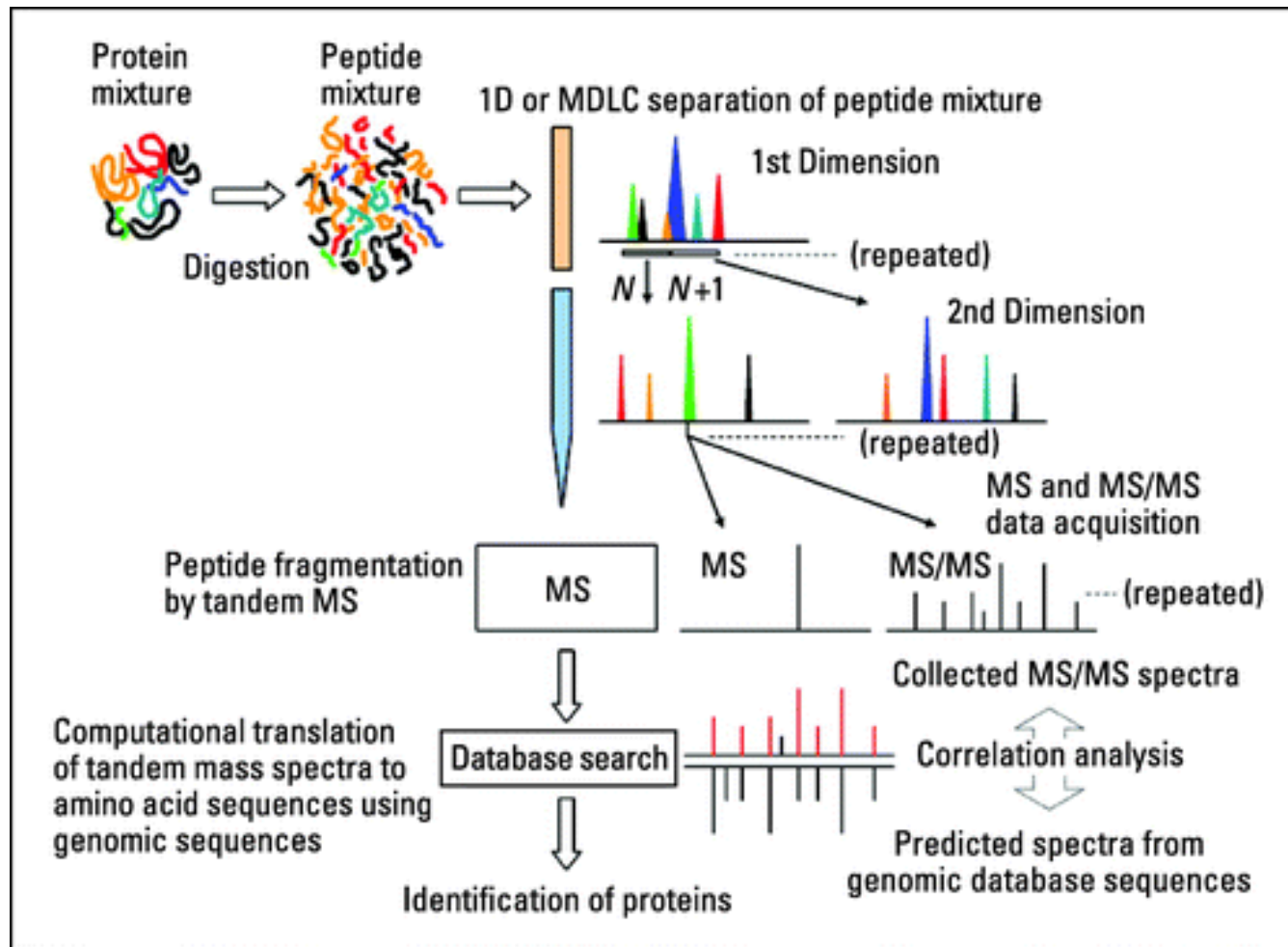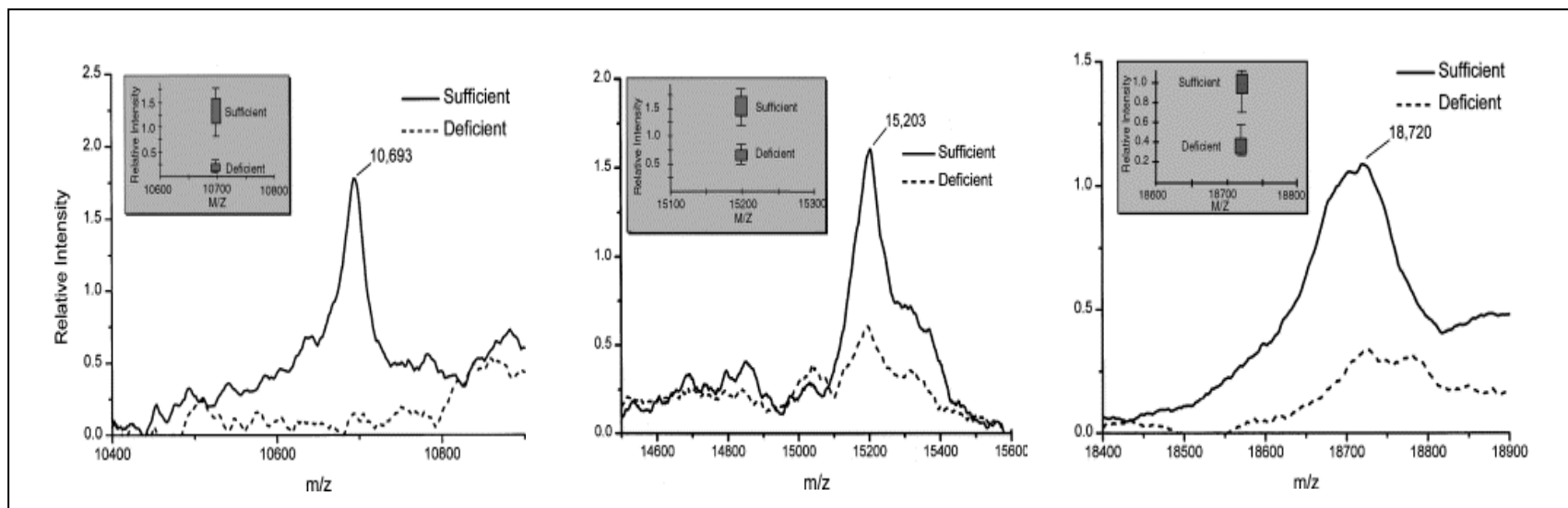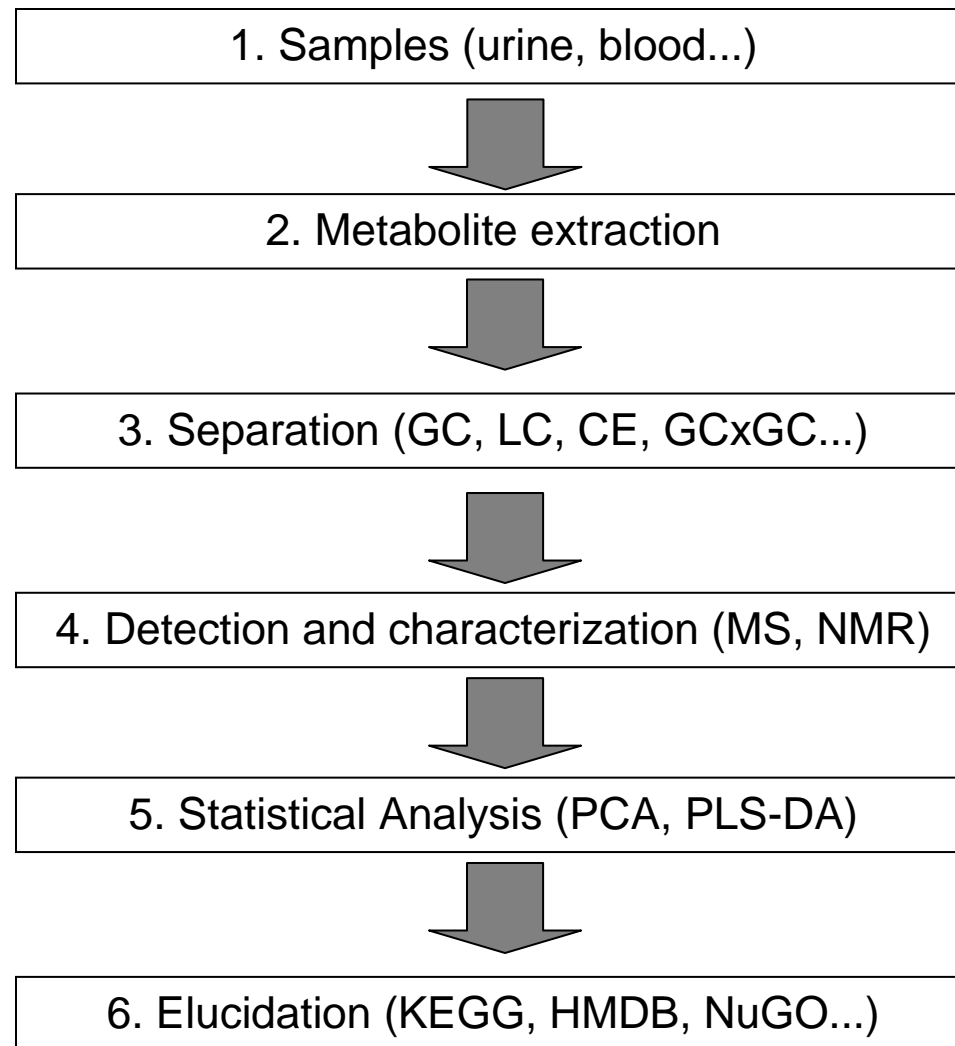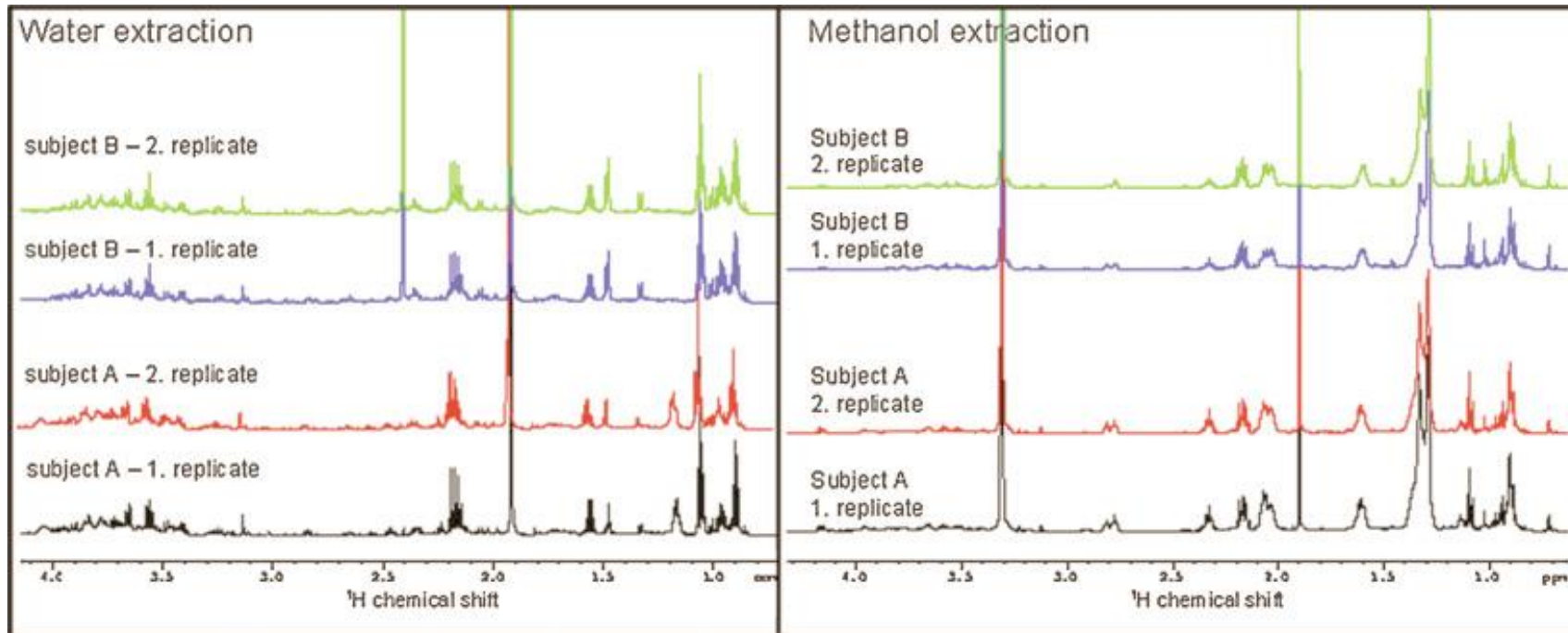
Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Table 1. Applications of expression DNA microarrays in Nutrigenomics research.

Bioactive compound/ Food ingredient	Food/ beverage	Studied model	Expected effect/target illness	Analytical methodology	Ref.
Anthocyanins C3G and cyanidin Cy	Fruits, vegetables, red wine	Human adipocyte cells	Regulation of adipocyte function	Affymetrix Human Genome microarray	[37]
Astaxanthin	Fish, algae	Mice	Regulation of oxidative phosphorylation and oxidative stress	Affymetrix Mouse Expression microarray	[52]
Chlorella algae intake	Diet	Healthy men (blood cells)	Regulation of fat and glucose metabolism. Modulation of glucose sensitivity in humans	Custom diabetes-related microarray (Hitachi)	[6]
Epicatechin	Cocoa	Human colon adenocarcinoma Caco-2 cells	Prevention of the oxidative DNA damage, reduction of inflammatory response	Clontech Human Haematology microarray	[36]
Epigallocatechin-3 gallate	Green tea	Human bronchial epithelial 21BES cells	Chemopreventive agent in cancer	Custom printed Human microarray	[35]
		Human prostate carcinoma LNCaP cells	Anti-proliferative action	cDNA microarray	[53]
		Human HT 29 colon carcinoma cells	Anti-proliferative action	Affymetrix Human Genome microarray	[40]
Genistein	Soybean	Postmenopausal women (peripheral lymphocytes)	Regulation cAMP signalling and cell differentiation	Human oligo microarrays	[24]
High-Cholesterol intake	Diet	Rats	Effects in cardiovascular disease	Custom printed cDNA microarray	[54]
High-protein / high-carbohydrate intakes	Dairy-based breakfasts	Healthy men (blood cells)	Regulation of glycogen metabolism and protein biosynthesis	Affymetrix Human Genome microarray	[50]
Long chain polyunsaturated fatty acids	Fish oil	Mice	Regulation of hepatic beta-oxidation and gluconeogenesis	Affymetrix Murine Genome microarray	[48]
Low-calorie intake	Diet	Obese men (blood cells)	Regulation of oxidative stress and inflammation	Agilent Human Oligo microarray	[55]
Omega-3 fatty acid	Fish oil	Human colon adenocarcinoma Caco-2 cells	Chemopreventive agent in cancer	Clontech Human Atlas Glass Arrays	[49]
Quercetin	Fruits and vegetables	CO115 colon-adenocarcinoma cells	Chemopreventive agent in cancer	Affymetrix Human Genome microarray	[56]
Sulforaphane	Cruciferous vegetables	Mice	Chemopreventive agent in cancer	Affymetrix Murine Genome microarray	[39]

Table 2. Applications of Proteomics in Nutrigenomics research.

Bioactive compound/ Food ingredient	Food/ beverage	Studied model	Expected effect/ target illness	Analytical methodology	Ref.
Quercetin	-	Human SW480 colon carcinoma cells	Colorectal cancer prevention	2DE, MALDI-TOF MS	107
Fatty acids	Dietary fish oil	Leiden transgenic mice	Lipid and glucose metabolism study	2DE, MALDI-TOF MS	108
Ginestein and daidzein isoflavones	Soy extract	Human endothelial cells	Atherosclerosis-preventive activities	2DE, MALDI-TOF MS	[109-113]
Isoflavones	Cereal bars	Human peripheral blood mononuclear cells	Atherosclerosis-preventive activities	2DE, MALDI-TOF MS	[113]
Monacolin K	Fungus	Caco-2 cells	Chemopreventive agent in cancer	2DE, MALDI-TOF/TOF MS	[114]
Volatile compounds	Coffe bean	Rat brain	Antioxidart and stress relaxation	2DE, MALDI-TOF MS, MALDI-Q-TOF MS/MS	[115]
Polyphenols	Green tea	Human lung adenocarcinoma A549 cells	Anticancer activity	2DE, nLC-ESI-Q-TOF MS/MS	[116]
No specific bioactive compound	Diets with different % of vegetable	Mouse colon mucosal cells	Colorectal cancer prevention	2DE, MALDI-TOF MS	[117]
Isoflavones	Soya foods	Human serum	Vascular protection	DIGE, LC-MS/MS	[118]
Vitamin A	-	Rat plasma	Changes in nutritional status	SELDI-TOF	[143]

Table 3. Metabolomics applications in Nutrigenomics research.

Bioactive compound/ Food ingredient	Food/ beverage	Studied model	Expected effect/target illness	Analytical methodology	Ref.
Phytochemicals	Low and standard phytochemical diet	Human urine samples	Acute effects in urinary metabolism	^{1}H-NMR and HPLC-ESI-Q-TOF	[159]
Flavonoids/Phenolic compounds	Red wine/Red grape juice/ Tea	Human urine, plasma and fecal samples..	Reduce risk of cardiovascular disease via study of the gut microbial impact in polyphenols	GC-TOF-MS	[164]
Polyphenols	Chamomile tea	Human urine samples	Increased excretion of hippurate and glycine and depleted of creatinine	^{1}H-NMR	[168]
Polyphenols	Wine/Grape Juice	Human fecal samples	Modulation of the gut microflora to prevent inflammatory bowel	^{1}H-NMR	[169]
Isoflavones	Soy enriched diets	Human plasma samples	Changes in carbohydrate and energy metabolism.	^{1}H-NMR	[170]
Isoflavones	Soy enriched diets	Human urine samples	Improved glomerular function or general kidney function	^{1}H-NMR	[171]