

High throughput ‘omics’ approaches to assess the effects of phytochemicals in human health studies

Jaroslava Ovesná^{1*}, Ondřej Slabý², Olivier Toussaint³, Milan Kodíček⁴, Petr Maršík⁵, Vladimíra Pouchová¹ and Tomáš Vaněk⁵

¹*Crop Research Institute, Drnovská 507, 161 06 Prague 6, Ruzyně, Czech Republic*

²*Masaryk Memorial Cancer Institute, Zlutý kopec 7, 656 53 Brno, Czech Republic*

³*Department of Biology, Unit of Cellular Biochemistry and Biology, University of Namur (FUNDP), 5000 Namur, Belgium*

⁴*Institute of Chemical Technology, Technická 3, Praha 6, 160 00 Prague 6, Czech Republic*

⁵*Institute of Experimental Botany, Suchbátka 1, 161 06 Prague 6, Czech Republic*

Human health is affected by many factors. Diet and inherited genes play an important role. Food constituents, including secondary metabolites of fruits and vegetables, may interact directly with DNA via methylation and changes in expression profiles (mRNA, proteins) which results in metabolite content changes. Many studies have shown that food constituents may affect human health and the exact knowledge of genotypes and food constituent interactions with both genes and proteins may delay or prevent the onset of diseases. Many high throughput methods have been employed to get some insight into the whole process and several examples of successful research, namely in the field of genomics and transcriptomics, exist. Studies on epigenetics and RNome significance have been launched. Proteomics and metabolomics need to encompass large numbers of experiments and linked data. Due to the nature of the proteins, as well as due to the properties of various metabolites, experimental approaches require the use of comprehensive high throughput methods and a sufficiency of analysed tissue or body fluids. In this contribution, we describe the basic tools currently used in nutrigenomics studies and indicate the general requirements for future technology methodological routings.

Nutrigenomics: Phytochemicals: High throughput platforms: -Omics: Human health

Nutrigenomics/nutrigenetics is defined as the application of the genomics, transcriptomics, proteomics and metabolomics studies to human responses to nutrition, especially the relationship between nutrition and health. Nutrition and health research is focused on the prevention of disease by optimizing and maintaining cellular, tissue, organ and whole-body homeostasis.

Results of nutrigenetics/nutrigenomics studies could help people to modify their diets to avoid or include certain nutrients and food combinations in a more personalized way than the common state of knowledge of the ordinary consumer currently allows. The genetic basis of each individual differs. Genes and their variants (alleles) can put people at higher or lower risk of developing chronic disease, e.g. cardiovascular diseases, obesity, or cancer. Inherited genes, if known, may also help to predict the risks and benefits of certain foods. An example is the effect of ingesting junk foods, which trigger genes to send out messages that slow the body's metabolism⁽¹⁾, or the effect of diet insulin from cow milk. Consumption of insulin might cause type 1 diabetes, underpinned by genetic factors⁽²⁾. Dietary chemicals can cause damage to DNA through mutations, DNA rearrangements, epigenetic changes, or gene expression; some others can protect genome functions^(3,4).

Those cuisines based on large amounts of plants (vegetables, grains, fruits, nuts, beans and legumes), such as

found the Mediterranean diet, are believed to be health promoting. However, it was not until 1933 that a direct relationship between the consumption of fruit and vegetables and diseases, such as cancer, was demonstrated⁽⁵⁾. It is well known that plants produce chemicals to protect themselves, but recent research has demonstrated that they can also protect humans against diseases. A few of the well-known phytochemicals are lycopene in tomatoes, isoflavones in soy, and flavonoids in fruits⁽⁶⁾. The thousands of different types of phytochemicals can be categorized into various families, according to their protective functions⁽⁷⁾. However, medical research has indicated that there is no single magical phytochemical, but there are synergistic or ‘team effort’ to their medicinal values⁽⁸⁾. Also, there is increasing evidence that consumption of whole foods is better than isolated food components, such as dietary supplements and nutraceuticals. For example, increased consumption of carotenoid-rich fruits and vegetables was more effective than carotenoid dietary supplements⁽⁹⁾.

Nutrigenomics is expected to increase our fundamental knowledge of the interactions between life processes and our diet. Impacts of the diet upon gene stability and gene expression is basically studied using ‘-omics’ approaches. Measuring nutritionally responsive genome activity involves (1) model object(s), (2) type of nutrition or diet, (3)

* Corresponding author: Jaroslava Ovesná, fax +42 02 33 02 22 86, email ovesna@vurv.cz

methodological platforms, and (4) data evaluation/compilation, involving the analysis of hundreds of samples. This constitutes the advantages of exploring high throughput technologies⁽¹⁰⁾. This review article attempts to summarise the available high throughput technological platforms used in nutrigenomics studies.

Genetic make-up and nutrition

Several chronic diseases and cancers result from the connections between genetic susceptibility and both environmental and lifestyle factors, including diet. Genetics thus plays an important role in the health status of man. The human genome was successfully sequenced, and more than a thousand genes and their forms – alleles for rare, highly heritable ‘Mendelian’ disorders have been identified; variations in which are both necessary and sufficient to cause disease and/or susceptibility to a disease^(11–13). Single nucleotide polymorphisms (SNPs) in gene sequences are often underlying such disorders. The knowledge on the resulting amino acid substitution(s) is very valuable for the investigating of the genetics behind common diseases, and is collected in the International HapMap Project⁽¹⁴⁾. Over 7 million SNPs are expected to be in the human genome. The SNP data are available through the Human Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>). Up to now, over 150 000 validated, genome-wide SNPs have been genotyped in the same set of individuals. Such information constitutes goal-directed studies explaining interactions of nutrition, and especially phytochemicals with individual human haplotypes⁽¹⁵⁾. One of the best-known examples of the applicability of nutrigenetics are the 2 SNPs, in the methylene tetrahydrofolate reductase (MTHFR) gene, causing serious health problems. Treatment with folic acid supplementation helps to overcome the negative health effect of these SNPs⁽¹⁶⁾. Other epidemiological studies indicate that human exposure to isothiocyanates and indoles through cruciferous vegetable consumption may decrease cancer risk, but the protective effects are influenced by individual genetic make-up underlying metabolism and elimination⁽¹⁷⁾. Currently, hundreds of genes have been described that have been associated or linked with diseases; and possible interventions by pharmaceuticals and nutrients including botanicals have been under investigation⁽¹⁸⁾.

The genetic make-ups of individuals differ, and there is a need to discern these differences. New technological platforms are required to speed-up the investigation of existing haplotypes. Genomics use either classical DNA-sequencer technology or other advanced technologies, such as DNA arrays⁽¹⁹⁾. Recently, new approaches based on new scientific findings have been launched to promote high throughput sequencing, combining, e.g. elements of emulsion and oligo-ligation strategy^(20,21). These next-generation sequencing technologies are generating great excitement in the life-science community with the offer of high throughputs at a lower cost. Once haplotypes are described, methods are in place for the efficient haplotyping, including high throughput ones, such as DNA arrays^(22,23), pyrosequencing⁽²⁴⁾, or other PCR based approaches. Human epidemiology and intervention studies may thus effectively employ the information

on haplotypes, and drive an evolution toward personalised nutrition.

Gene expression and its regulation

Early investigations of gene regulation revealed that nutrients could modulate gene expression; flavonoids are a well-known example⁽²⁵⁾. Transcription of genes involved in chronic diseases or involved in inflammation may be affected by enriching our diet with specific foods, inherently rich in phytochemicals⁽²⁶⁾. Initiation of gene transcription itself requires the presence of transcription factors, enhancers/repressors, accessible promoters, and RNA polymerase.

Transcription is driven by, among others, the methylation status of the DNA string, specifically the promoter regions and histone decoration. Methylation patterns are affected by several factors – tissue specificity, developmental stage, as well by diet. Considerable epidemiologic evidence supports the hypothesis that susceptibility to chronic diet-related diseases is influenced by nutrition *in utero*, resulting in different methylation patterns. In fact, an entire research field of epigenomics has grown-up around the idea that epigenetic drugs and foods might be able to promote healthier methylation patterns in the otherwise intact genes of cancer patients, whose methylation switches may simply need tweaking, or resetting. The Human Epigenome Project (HEP) aims to identify and interpret the DNA methylation patterns of all human genes, in all major tissues. Epigenetic changes have been implicated in many diseases, particularly autoimmune diseases such as diabetes, rheumatoid arthritis, or cancer. The results available suggest that prevention and therapy of diseases by individualised tailoring of optimal epigenetic diets is conceivable⁽²⁷⁾.

The fundamental concept of epigenetic research is to combine an initial, sequence-based mapping of CpG and a prediction of CpG island strength⁽²⁸⁾. Epigenetics researchers using DNA methylation analysis are facing methodological problems, because of high CG content in the sequences. Companies have tried to offer solutions to overcome these problems by developing special chemistry and approaches. Pyrosequencing is an accurate and fast analysis method for quantifying CpG methylation in epigenetic studies. Bioinformatics analysis and direct mRNA profiling constitute yet another option.

Many techniques exist to study gene expression at the mRNA level. These techniques can be sorted into two classes. Either they are aimed at finding new genes that are differentially transcribed (high-density array, differential display RT-PCR, etc.), or the goal is to know whether a given interesting set of genes are differentially transcribed (low-density arrays, etc.).

High throughput techniques, producing a snapshot image of genes that are translated at any given moment, are currently employed to get an insight into transcriptome changes, such as reactions upon nutrition changes. SAGE (serial analysis of gene expression) or whole genome DNA arrays are currently used. SAGE has been employed to show the effects of phenolic compounds upon gene expression. Over 30 000 transcripts may be easily detected in reactions. The requirements for large amounts of mRNA inputs, limits SAGE utilization, irregardless of further improvements in the technique (e.g. SAGE-lite or PCR-SAGE or super SAGE)^(29,30).

Various DNA arrays/chips are available, such as those produced by photolithography, spotted oligonucleotides, or cDNA arrays. Commercial chips exhibit some advantages over in house prepared arrays; the number of probes, accuracy, and reproducibility constitute their major advantages. High-density oligonucleotide microarrays have been used for pioneer studies on the multiple gene expression effects exhibited (e.g. by Ginkgo biloba leaf extract), changing even traditional pharmacology and medicine concepts⁽³¹⁾. Several other examples of DNA arrays applicability in nutrigenomics studies have been documented by studies of flavonoids or garlic constituents. DNA array technology has proven that the high consumption of soy isoflavones in Asian diets has been correlated to a lower incidence of clinically important cases of prostate cancer⁽³²⁾. Although thousands of data points have been acquired using DNA arrays, scientists still call attention to the fact that the mode of action of several phytochemicals, that prevent the growth of cancer cells or poses anti-inflammatory activities, at the gene transcription level, is not completely understood. The complexity of the human genome, the difficulties in setting-up cohort study experiments, plus the price of the chips represent the main obstacles.

Low-density DNA arrays have become available as a tool offering several specific advantages. They allow for an optimized design of each capture probe, which increases greatly the reproducibility of DNA arrays. The selection of the genes is based upon a scrupulous analysis of the literature and in-house investigations, specifically by proteomics, high-density arrays and RT-PCR differential display. As low-density cDNA arrays are cheaper than high-density arrays, low-density DNA arrays allow kinetic and dose/response studies. Examples of applications are low-density cDNA arrays that allow one to monitor the expression of *MAGE-A* genes, rat gene expression profiling of drug metabolism and toxicology markers⁽³³⁾, human inflammation, human apoptosis, etc.

As it is recommended to validate microarray data by real-time PCR, a combination of customer-designed oligo-arrays in a 384 well plate (measured in a real-time PCR device) may be an effective step to the next generation of high throughput platforms for precise gene expression, together with low-density arrays.

RNome

RNome represents a new scientific word, the RNA equivalent of the proteome or genome, with the subject referred to as RNomics. RNomics is a newly emerging discipline that studies the structure, function and processes of non-coding RNAs and the mechanism of RNA interference (RNAi) in a cell. RNA interference is ancient and an evolutionarily conserved gene regulation mechanism, triggered by the presence of double-stranded RNA (dsRNA). Key players in this mechanism are microRNAs (miRNAs) and short interfering RNAs (siRNAs), which in association with Argonaute proteins (AGOs) form multimeric RNA-induced silencing complexes (RISC) that elicit mRNA decay, translational repression, or transcriptional silencing, respectively, of the mRNA targets selected, even by the complementarity with siRNAs or miRNAs⁽³⁴⁾.

Short interfering RNAs are mainly derived from exogenous dsRNA and they are supposed to have evolved in plants as a defence mechanism against viruses. Since the 'coincidental' discovery of RNAi during attempts to alter the pigmentation of petunia flowers, intensive studies have focused on developing RNAi technologies, because of their possible usage for treating human diseases. As a consequence, RNAi technologies are common molecular biological techniques today, widely used in biomedical research focused on virology, oncology and inflammatory diseases.

MicroRNAs (miRNAs) are an abundant class of non-protein-coding small endogenous RNAs that are present in diverse eukaryotes. Bioinformatics and cloning studies have estimated that miRNAs may regulate up to 30% of all human genes, and each miRNA can control hundreds of gene targets⁽³⁵⁾. This led to enthusiastic interest in profiling miRNA expression as well as developments in Northern blot cloning PCR (bead-based, SAGE-based and microarray-based methods). The major part of miRNAs studies were done in the oncology research area^(36–38).

Most recent reports have shown that distinct miRNA expression patterns are associated with various solid tumour types, lung cancer prognosis, and chronic lymphocytic leukaemia progression. Particular miRNAs were studied in association with metabolic diseases such as type 2 diabetes^(39,40). It is evident that the pathogenesis of these multi-factorial and complex diseases are at least partially linked to dietary patterns. Studies of specific RNomes, associated with bioactive food components, and together with human diseases genomics and RNomics, will provide efficient forces leading to remarkable future improvements of disease prevention.

Changes of protein profiles and their investigations

Although microarray studies can reveal the relative amounts of different mRNAs in the cell, levels of mRNA are not directly proportional to the expression level of the proteins that they code for. The human genome consists of 24 000–30 000 genes, which may generate ~100 000 proteins – due to mRNA splice variants, protein processing, and post-translational modifications⁽⁴¹⁾. Protein profiles reflect the physiological status of a person, including health or disease processes. Protein profiles and their characteristics in physiological fluids are excellent indicators of nutritional status, as well. For example, plasma sub-proteome released by organs and cell populations might be useful for diagnostics⁽⁴²⁾ and perhaps the investigation of the effects of botanicals upon human health.

Proteomics are often considered the next step in the study of biological systems, after genomics. However, protein analysis is usually very complicated, requiring a higher amount of biological material, compared with RNA profiling, for which *in vitro* amplification steps are possible. Unlike protein analysis, a single cell may be enough for some RNA profiling experiments. Proteomics is also much more complicated than genomics. This is primarily because while an organism's genome is rather constant, a proteome differs from cell to cell and constantly changes through its biochemical interactions with the genome and environment, as well. Moreover, food constituents not only alter gene expression, including translation, but also affect interactions among proteins in

the cells. Thus, nutritional proteomics or nutriproteomics is not only the application of proteomics methodology to nutrition-related research, but it also represents the interaction of bioactive food ingredients with proteins; whereby the interactions with proteins occurs in specific ways⁽⁴³⁾. Non-covalent and covalent modifications have been observed for the interaction of selected secondary plant products (e.g. glucosinolate breakdown products, phenolic compounds) with a series of proteins, addressing both food technological and the physiological aspects of human nutrition.

On the whole, studies currently addressing proteomic aspects, in relation to nutritional research, use both classical methods and modern approaches (Fig. 1) to look at the changes in peptide and protein patterns (such as protein mapping or profiling). There are currently four major proteomics technologies that involve either specific digestion of proteins [2D-PAGE, MS and multidimensional protein identification technology (MudPIT)] or the direct analysis of proteins after their chromatographic separation [the top-down approach, and surface-enhanced laser desorption ionization (SELDI)]. The choice of technology depends on the goal of the study and facilities availability^(44,45) (Fig. 2). Since all of these techniques are very different in terms of their sensitivity, the usefulness for quantitative determinations varies significantly⁽⁴⁶⁾.

Current research in proteomics requires first that proteins be resolved, sometimes on a massive scale. Protein separation can be performed using two-dimensional gel electrophoresis. Once proteins have been separated and quantified, they are identified by mass spectrometry, specifically, matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) mass spectrometry. Mass spectrometry provides key tools for the analysis of proteins. The limiting step for this technology in human diagnostics lies mainly with the data analysis, i.e. computing power and the lack of adequate software tools (for more detail, see proteomics tools <http://www.expsy.ch/tools/>).

Three proteins were identified with the use of 2D electrophoresis and MS, as potential molecular targets for the proposed action against colon cancer of quercetin, a plant

flavonoid. The application of proteomic techniques demonstrated the response of these proteins to quercetin treatment of colon cancer cells⁽⁴⁷⁾.

Protein microarrays offer another tool that has been developed. Instead of many copies of a particular segment of a gene on a spot, many copies of a protein antibody occupy a spot. The protein is then exposed to the protein microarray. If a particular protein is present, it will adhere to the antibody and can be visualized by the addition of non-specific fluorescently-labelled antibodies to the proteins that adhere to the microarray. Such protein microarrays allow scientists to screen for protein–protein interactions or identify protein targets of small molecules. While investigators are still fine-tuning protein microarray production, there is enormous promise in these protein microarrays as future diagnostic tools⁽⁴⁸⁾.

When applying proteomics to nutrition-related research, one has to keep in mind that apart from the analytical complexity the limitations of these methods still particularly lie in the cost, resolution, reproducibility, and throughput. Integrated with other advanced technologies and systems biology, proteomics could facilitate the discovery of key proteins that function to regulate metabolic pathways; and whose synthesis, degradation, and modifications are affected by specific nutrients (or other dietary factors) including phytochemicals. This will aid in rapidly enhancing our knowledge of the complex mechanisms responsible for nutrient utilization, identifying new biomarkers for nutritional status and disease progression, and in designing a contemporary paradigm for dietary prevention and intervention of disease. Thus, the proteome analysis holds great promise in improving human health⁽⁴⁴⁾.

Metabolomics

Metabolomics originated from metabolic profiling, which appeared at the beginning of the 1970s. Although metabolic profiling is more easily done with bacteria, yeast or plants because of the accessibility of the materials, the topic has also concerned qualitative and quantitative analyses of complex mixtures of human origins, for medicinal purposes.

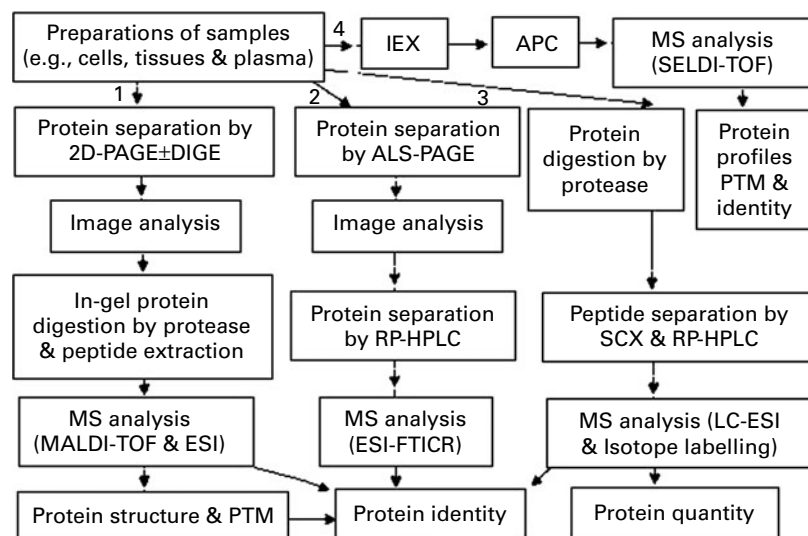


Fig. 1. Workflow and methods used in nutriproteomics studies as described in Wang *et al.*⁽⁴⁴⁾.

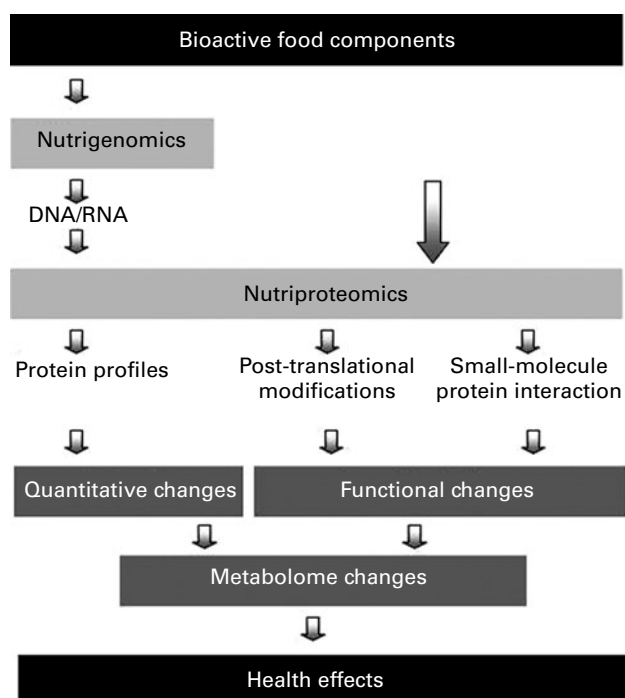


Fig. 2. Position of nutriproteomics in food–cell interactions. Adopted from Schweigert⁽⁴³⁾.

Metabolites (the end-products of the activity of metabolic enzyme cascades) come in all shapes and sizes, and include everything from small basic precursors (such as diacyl glycerol) to larger molecules (such as hormones). Global metabolic profiling (metabonomics/metabolomics) has already shown particular promise in the area of toxicology and drug development⁽⁴⁹⁾. The principal aim of this method consists in the elucidation of physiological processes in organisms through changes in the occurrence and concentrations of these metabolic products. It can contribute, together with other ‘-omics’ data, to complex biological processes⁽⁵⁰⁾.

Some authors have stressed⁽⁵¹⁾ that human metabolic profiling can be done unlike genomics and transcriptomics on human fluids, such as urine, serum or plasma. However, metabolic studies on human tracing effects of diets or specific nutrients are sparse, and are based mostly on the ¹H-NMR method. The metabolomic approach was applied to evaluate all of the biochemical changes occurring after dietary intervention by soy consumption. Despite the presence of substantial inter-subject variability, the metabolomic analysis enabled the identification of plasma components related to the dietary intervention, suggesting soy-induced alterations in fat, protein, and carbohydrate metabolism⁽⁵²⁾.

Metabolomic approaches, in fact, utilize various analytical technologies for the detection and quantification of as many compounds as possible. The most popular and basic analytical method used in metabolic profiling is mass spectrometry (MS), preceded by separation by gas (GC) or liquid chromatography (LC). It is very effective tool with good sensitivity and relative sufficient selectivity. LC/MS is a suitable highly sensitive method for complex biological mixtures, which can be combined with other detection methods, such as UV or NMR⁽⁵³⁾. GC/MS is very useful technique but is limited by

volatility of the analyzed compounds. Poor volatility of some constituents can be partially compensated for by chemical derivation (silylation, methylation), but at the cost of processing time and variance⁽⁵⁴⁾. A subset of the metabolome is the lipidome, which is composed of the lipids found in a given biological fluid, such as plasma. Metabolomics and lipidomics, basically, use two core methods: mass spectrometry (MS) – which measures all ionisable, low-abundance metabolites in a sample; and nuclear magnetic resonance spectroscopy (NMR) – which can quantify all ionisable, high-abundance metabolites. The impact of botanicals upon lipid profiles may lead to identification of fruits and vegetables lowering the risk of, e.g., arteriosclerosis in certain genotypes.

Nuclear magnetic resonance (NMR) is a very specific method for the analysis of most organic compounds. However, most NMR-based approaches are focused on the analysis of biomarker substances whose occurrence and concentration signalled the changes in the organism, and are not useful for the analysis of chemically complicated biological extracts. It is also used for the metabolic fingerprinting of extracts⁽⁵⁵⁾. Nevertheless, the combination of HPLC and detection NMR is a very promising approach for the metabolomic analysis, although the expensiveness of this method (due to the consumption of deuterated mobile phases) might be its most important disadvantage⁽⁵³⁾.

There are a number of other analytical techniques used for metabolite profiling and metabolomics, such as optical spectroscopic methods, such as ultraviolet (UV) and infrared (IR) spectroscopy, evaporative light scattering detection (ELSD), or laser-induced fluorescence (LIF); usually coupled with various separation methods, such as HPLC, TLC, or capillary electrophoresis (CE). They are very different in respect to their sensitivity. For example, ELSD is a non-selective method with a relatively low sensitivity, but suitable for all non-volatile compounds, regardless of chromophores. IR and UV provide sufficient sensitivity together with a certain selectivity, and LIF has a very high sensitivity and lower selectivity. With the use of all of these techniques, it is necessary for the additional identification of specific metabolites by other more selective methods. Each of these methods is suited for the analysis of different molecules, so only the integration of various metabolic profiling methods could pretend to characterise the whole metabolome⁽⁵⁶⁾.

For data analysis, in the metabolomic as well as for other ‘-omics’ approaches, a wide range of statistical and deconvolution methods are utilized. Deconvolution algorithms (AMDIS, Mzmine) are applied to the identification and processing of outputs from chromatographic data of MS analysis⁽⁵⁷⁾. These tools allow improved deconvolution and the identification of overlapping chromatographic peaks. For statistical analysis, similar statistical methods as in the other ‘-omics’ approaches are used. They involve a number of multivariate statistical methods, such as principal component analysis (PCA), factor analysis (FA), various cluster analyses, or algorithms based on neural networks as self-organizing maps (SOMs). One of most commonly used is principal component analysis (PCA)⁽⁵⁸⁾, which consists of description of variance in a set of multivariate data by means of set of orthogonal variables (principal components). The original variables (i.e. concentrations) are described as a linear combination of particular principal components. Cluster analysis is a

Table 1. Possible analyses used in fruit and vegetable constituent's interaction with human cells (adopted according to Young)⁽⁶⁵⁾

Level	Definition	Example of analysis
1. Genome	Genomic fingerprinting	Nucleotide sequencing, DNA arrays
2. Methylome	DNA methylation status	DNA arrays, pyrosequencing
3. RNome	Set of RNAs in a cell	DNA arrays
4. Transcriptome	mRNA profiling	DNA chips, RTqPCR
5. Proteome	Set(s) of cellular proteins	2D gel, MS
6. Metabolome	Metabolites in cells/organs	μ TAS: IR, NMR
7. Populome	Complete information on group of people	Epidemiological studies

2D gel, two-dimensional electrophoresis; RTqPCR, quantitative PCR coupled with reverse transcription; MS, mass spectrometry; μ TAS, micro-total analytic systems; IR, infrared; NMR, nuclear magnetic resonance.

method of grouping data by their similarity. There are a number of clustering methods, in which hierarchical clustering analysis is most frequently used (HCA). SOMs are algorithms, which are able to be calibrated ('trained') with known data patterns, and then compared with newly added data sets⁽⁵⁹⁾. Aside from the statistical analysis tools, there are also various programs used which allow for the visualization of metabolomic data and their connections with proteomic or genomic datasets^(60,61).

Recent efforts combining metabolomics data sets with RNA microarray or shotgun proteomics data offer new opportunities to strive toward a comprehensive understanding of biological processes. However, currently, several methodological limitations hamper any swift advancement toward this important goal, and have to be addressed in the near future. To strengthen the confidence in any metabolic profiling techniques, standard validation protocols will have to be applied, in the future. Analytical reliability will have to be proven; especially both intra- and inter-laboratory reproducibility will have to be demonstrated if the broad applicability of these assays is desired⁽⁶²⁾.

A full description of human metabolism is far from being complete, reflecting the different: tissues, age, genotype, gender, diet, stress, health; and other internal and external factors. Researchers at the University of Alberta, Canada, announced at the beginning of 2007 the completion of the first draft of the human metabolome, the chemical equivalent of the human genome. The scientists have catalogued and characterized 2500 metabolites, 1200 drugs, and 3500 food components that can be found in the human body⁽⁶³⁾; many others have yet to be identified.

The application of metabolomics to nutritional research will meet unique challenges. Little is known of the extent to which changes in the nutrient content of the human diet elicit changes in metabolic profiles. Notwithstanding possible confounding effects, there exists every reason to be optimistic about the potential of metabolomics for the assessment of various biofluids in nutrition research. Metabolomics have the potential to distinguish between different dietary treatments, which would not have been identified with conventional techniques⁽⁶⁴⁾.

Conclusions

It is generally accepted that food consumption has impacts on dynamic changes in the human genome, gene expression, metabolome, and finally in human health, generally. It has

been indicated that a diet rich in phytochemicals, linked with the daily intake of fruits and vegetables, may have a protective effect upon the human genome. However, scientists need to gather more comprehensive knowledge on the variability of the human genome, with respect to allelomorphism and possibly epigenetic changes, protein, and metabolite levels (as a consequence of fruit and vegetable consumption); and thus the chance to prevent diseases and prolong human life by personalised nutrition. To realize this aim, high throughput methods have to be employed to fully understand the mechanism of food's constituents' interactions with individual genotypes under different life situations (Table 1). While DNA and RNA analyses are quite advanced, especially in human substances, protein and metabolite analyses still cope with a range of analysis, speed, and cost effectiveness. Therefore, the studies of human proteome and metabolome responses treated with fruit and vegetable constituents, employing high throughput techniques, are not numerous.

However, because of rapid technological developments of platforms used for disease and pharmaceutical investigations (as well as subsequent nutrionomics, as the emerging face of nutrition) will be able to provide the necessary stepping-stones to achieve this ambitious goal of optimizing an individual's health via nutritional intervention.

Acknowledgements

The publication of this paper was made possible by the financial support of the European Co-operation in the field of Scientific and Technical (COST) Research Action 926 'Impact of new technologies on the health benefits and safety of bioactive plant compounds' (2004–2008), linked to project 1P05OC054 of the Ministry of Education, Youth and Sport of the Czech Republic, and project 002700602 of the Ministry of Agriculture of the Czech Republic. The authors had no conflicts of interest to disclose.

References

1. Kaput J & Rodriguez RL (2004) Nutritional genomics: the next frontier for post genomic era. *Physiol Genomics* **16**, 166–177.
2. Vaarala O (2006) Is it dietary insulin? *Ann N Y Acad Sci* **1079**, 350–359.
3. Kaput J & Rodriguez RL (2004) Nutritional genomics: the next frontier for post genomic era. *Physiol Genomics* **16**, 166–177.
4. Ames BN (2005) Increasing longevity by tuning-up metabolism. *EMBO Rep* **6**, S1, S20–S24.

5. Premier R (2002) Phytochemical composition: paradigm shift for food-health considerations. *Asia Pac J Clin Nutr* **11**, S197–S201.
6. Rotelli AE, Guardia T, Juarez AO, de la Rocha NE & Pelzer LE (2003) Comparative study of flavonoids in experimental models of inflammation. *Pharmacol Res* **48**, 601–606.
7. Williamson G & Manach C (2005) Bioavailability and bioefficacy studies of polyphenols in humans. II. Review of 93 intervention studies. *Am J Clin Nutr* **81**, 243S–255S.
8. Liu RH (2003) Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *Am J Clin Nutr* **78**, Suppl., S17S–S20S.
9. Heber D (2004) Vegetables, fruits and phytoestrogens in the prevention of diseases. *J Postgrad Med* **50**, 145–149.
10. Galitski T (2004) Molecular Networks In Model systems. *Annu Rev Genomics Hum Genet* **5**, 177–187.
11. Stein LD (2004) Human genome: end of the beginning. *Nature* **431**, 915–916.
12. Kaput J (2006) Diet–disease interactions at the molecular level. An experimental paradigm. In *Phytochemicals: Nutrient–Gene Interactions*, pp. 23–39 [M Meskin, WR Bidlack and RK Randolph, editors]. Boca Raton: CRC Press.
13. Serre D & Hudson TJ (2006) Resources for genetic variation studies. *Annu Rev Genomics Hum Genet* **7**, 443–457.
14. The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* **437**, 1299–1320.
15. Hesketh J, Wybranska I, Dommels Y, King M, Elliott R, Pico C & Keijer J (2006) Nutrient–gene interactions in benefit–risk analysis. *Br J Nutr* **95**, 1232–1236.
16. Molloy AM & Scott JM (2001) Folates and prevention of disease. *Public Health Nutr* **4**, 601–609.
17. Higdon JV, Delage B, Williams ED & Dashwood RH (2007) Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. *Pharmacol Res* **55**, 224–236.
18. Meskins MS & Bidlack WR (2006) Nutrient–gene interactions. Lavoisier 2000–2007 <http://www.lavoisier.fr/notice/gb282010.html>
19. Daniel H (2002) Genomics and proteomics: importance for the future of nutrition research. *Br J Nutr* **87**, S305–S311.
20. Hardenbol P, Baner J, Jain M, *et al.* (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol* **21**, 673–678.
21. Bonetta L (2006) Genome sequencing in the fast lane. *Nat Methods* **3**, 141–147.
22. Dove A (2005) The SNPs are down: genotyping for the rest of us. *Nat Methods* **2**, 989–994.
23. Herr A, Grutzmann R, Matthaei A, Artelt J, Schrock E, Rump A & Pilarsky C (2005) High-resolution analysis of chromosomal imbalances using the Affymetrix 10K SNP genotyping chip. *Genomics* **85**, 392–400.
24. Elahi E & Ronaghi M (2004) Pyrosequencing: a tool for DNA sequencing analysis. *Methods Mol Biol* **255**, 211–220.
25. Frigo DE, Bich ND, Melnik LI, Schief LS, Collins-Burrow BM, Pace DK, McLachlan JA & Burow ME (2002) Flavonoid phytochemicals regulate activator protein-1 signal transduction pathways in endometrial and kidney stable cell lines1. *J Nutr* **132**, 1848–1853.
26. Gross-Steinmeyer K, Stapleton PL, Liu F, *et al.* (2004) Phytochemical-induced changes in gene expression of carcinogen-metabolizing enzymes in cultured human primary hepatocyt. *Xenobiotica* **34**, 7, 619–632.
27. Gallou-Kabani C, Vigé A, Gross MS & Junien C (2007) Nutri-epigenomics: lifelong remodelling of our epigenomes by nutritional and metabolic factors and betone. *Clin Chem Lab Med* **3**, 321–327.
28. Laird PW (2005) Cancer epigenetics. *Hum Mol Genet* **14**, R65–R76.
29. Velculescu VE, Zhang L, Vogelstein B & Kinzler KW (1995) Serial analysis of gene expression. *Science* **270**, 484–487.
30. Matsumura H, Nasir KHB, Yoshida K, Ito A, Kahl G, Krüger DH & Terauchi R (2006) SuperSAGE array: the direct use of 26-base-pair transcript tags in oligonucleotide arrays. *Nat Methods* **3**, 421.
31. Christen Y, Olano-Martin E & Packer L (2002) Egb 761 in the postgenomic era: new tools from molecular biology for the study of complex products such as Ginkgo biloba extract. *Cell Mol Biol* **48**, 539.
32. Rice L, Handayani R, Cui Y, Medrano T, Von Samedi, Baker H, Szabo NJ, Rozset CHJ, Goodison S & Shiverick KT (2007) Soy isoflavones exert differential effects on androgen responsive genes in LNCaP human prostate cancer cells. *J Nutr* **137**, 964–972.
33. De Longueville F, Atienzar FA, Marcq L, *et al.* (2003) Use of a low-density microarray for studying gene expression patterns induced by hepatotoxicants on primary cultures of rat hepatocytes. *Toxicol Sci* **75**, 378–392.
34. Montgomery MK (2004) RNA interference: historical overview and significance. *Methods Mol Biol* **265**, 3–21.
35. Grunweller A & Hartmann RK (2005) RNA interference as a gene-specific approach for molecular medicine. *Curr Med Chem* **12**, 3143–3161.
36. Lim LP, Lau NC, Garrett-Engel P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS & Johnson JM (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773.
37. Barad O, Meiri E, Avniel A, *et al.* (2004) MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res* **14**, 2486–2494.
38. Lu J, Getz G, Miska EA, *et al.* (2005) MicroRNA expression profiles classify human cancers. *Nature* **435**, 834–838.
39. Cummins JM, He Y, Leary RJ, *et al.* (2006) The colorectal microRNAome. *Proc Natl Acad Sci USA* **103**, 3687–3692.
40. Poy MN, Eliasson L, Krutzfeldt J, *et al.* (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432**, 226–230.
41. Kim H, Page GP & Barnes S (2004) Proteomics and mass spectrometry in nutrition research. *Nutrition* **20**, 155–165.
42. Ping P, Vondriska TM, Creighton CJ, *et al.* (2005) A functional annotation of subproteomes in human plasma. *Proteomics* **5**, 3506–3519.
43. Schweigert FJ (2007) Nutritional proteomics methods and concepts for research in nutritional science. *Ann Nutr Metab* **51**, 99–107.
44. Wang J, Li D, Dangott LJ & Wuy G (2006) Proteomics and its role in nutrition research. *J Nutr* **136**, 1759–1762.
45. Fuchs D, Winkelmann I, Johnson IT, Mariman E, Wenzel U & Daniel H (2005) Proteomics in nutrition research: principles, technologies and applications. *Br J Nutr* **94**, 302–314.
46. Smith L (2006) Proteomics: challenges and emerging technologies. *Expert Rev Proteomics* **3**, 6, 573–577.
47. Mouat MF, Kollu K, Orlando R, Hargrove JL & Grider A (2005) The effects of quercetin on SW480 human colon carcinoma cells: a proteomic study. *Nutr J* **4**, 11.
48. Brooksbarak C (2000) This way up and handle with care. *Nat Rev Mol Cell Biol* **1**, October 2000.
49. Keun HC (2006) Metabonomic modeling of drug toxicity. *Pharmacol Ther* **109**, 92–106.
50. Quackenbush J (2006) From 'omes' to biology. *Anim Genet* **37**, Suppl. 1, 48–56.
51. Rist MJ, Wenzel U & Daniel H (2006) Nutrition and food science go genomic. *Trends Biotechnol* **24**, 4, 172–178.

52. Solanky KS, Bailey NJ, Beckwith-Hall BM, Davis A, Bingham S, Holmes E, Nicholson JK & Cassidy A (2003) Application of biofluid ¹H nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. *Anal Biochem* **323**, 197–204.
53. Bailey NJ, Stanley PD, Hadfield ST, Lindon JC & Nicholson JK (2000) Mass spectrometrically detected directly coupled high performance liquid chromatography/nuclear magnetic resonance spectroscopy/mass spectrometry for the identification of xenobiotic metabolites in maize plants. *Rapid Commun Mass Spectrom* **14**, 679–684.
54. Halket JM, Przyborowska A, Stein SE, Mallard WG, Down S & Chalmers RA (1999) Deconvolution gas chromatography/mass spectrometry of urinary organic acids—potential for pattern recognition and automated identification of metabolic disorders. *Rapid Commun Mass Spectrom* **13**, 279–284.
55. Choi YH, Tapias EC, Kim HK, Lefeber AWM, Erkelens C, Verhoeven JTJ, Brzin J, Zel J & Verpoorte R (2004) Metabolic discrimination of *Catharanthus roseus* leaves infected by *Phytoplasma* using ¹H-NMR spectroscopy and multivariate data analysis. *Plant Physiol* **135**, 2398–2410.
56. Dunn WB & Ellis DI (2005) Metabolomics: current analytical platforms and methodologies. *Trends Analyt Chem* **24**, 4, 285–294.
57. Katajamaa M & Orešič M (2005) Processing methods for differential analysis of LC/MS profile data <http://www.biomedcentral.com/content/pdf/1471-2105-6-179.pdf>
58. Jansen JJ, Hoefsloot H CJ, Boelens HFM, Van der Greef J & Smilde AK (2004) Analysis of longitudinal metabolomics data. *Bioinformatics* **20**, 15, 2438–2446.
59. Kohonen T (1982) Self-organized formation of topologically correct feature maps. *Biol Cybern* **43**, 59–69.
60. Wolf D, Gray CP & de Saizieu A (2000) Visualising gene expression in its metabolic context. *Brief Bioinform* **1**, 297–304.
61. Mendes P (2002) Emerging bioinformatics for the metabolome. *Brief Bioinform* **3**, 134–145.
62. Seger CH & Sturm S (2007) Analytical aspects of plant metabolite profiling platforms: current standings and future aims. *J Proteome Res* **6**, 480–497.
63. Wishart DS, Tzur D, Knox C, *et al.* (2007) HMDB: the Human Metabolome Database. *Nucleic Acids Res* **35**, Database issue, D521–D526.
64. Gibney MJ, Walsh M, Brennan L, Roche HM, German B & Van Ommen B (2005) Metabolomics in human nutrition: opportunities and challenges. *Am J Clin Nutr* **82**, 497–503.
65. Young VR (2002) 2001 W.O. Atwater Memorial Lecture and the 2001 ASNS President's Lecture: Human Nutrient Requirements: The challenge of the Post Genome Era. *J Nutr* **132**, 621–629.