Part I
Introduction – Definitions

1
Nutritional Genomics: Concepts, Tools and Expectations

Hannelore Daniel and Uwe Wenzel

1.1
Nutrigenomics: Just Another "omic"?

The age of nutrigenomics is already upon us. Various new programs in molecular nutrition research have been launched in Europe, Asia, and the US under the heading of nutrigenomics. We may for this review consider nutrigenomics as the science that seeks to provide a molecular understanding for how diets and common dietary constituents affect mammalian metabolism and health by altering gene/protein expression on basis of an individual's genetic makeup.

Although nutrigenomics represents in the first place just another "omic," it clearly induces a conceptual shift in nutritional sciences by moving the genome into the center of all the processes that essentially determine mammalian metabolism in health and disease. Moreover, for the first time, nutritional science speaks the same language and uses the same tools as the other biomedical sciences and this is going to change the face of nutrition research. Nutritional sciences is functional genomics "par excellence" and will thereby move the discipline into the heart of biological sciences. Unlike other environmental factors, nutrients, non-nutrient components of foods, and xenobiotics in foods have huge variability in dose and time and hit a rather static genome, affecting the function of a large number of proteins encoded by the respective mRNA molecules that are expressed in a certain cell, organ or organism. Alterations of mRNA levels and in turn of the corresponding protein levels are critical parameters in controlling the flux of a nutrient or metabolite through a biochemical pathway. Nutrients and non-nutrient components of foods, diets, and lifestyle can affect essentially every step in the flow of genetic information from gene expression control to protein synthesis, protein degradation, and allosteric control and consequently alter metabolic functions in the most complex ways (Fig. 1.1).

The advent of high-throughput technologies has led to the rapid accumulation of biological data, ranging from complete genomic sequences, transcripts, proteome and metabolome profiles as well as the first protein–protein interaction maps. Referred to as "omics", these parallel approaches are usually classified by the measured target molecules. Transcriptomics determines the transcript levels or pat-

Nutritional Genomics. Edited by Regina Brigelius-Flohé and Hans-Georg Joost
Copyright © 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 3-527-31294-3

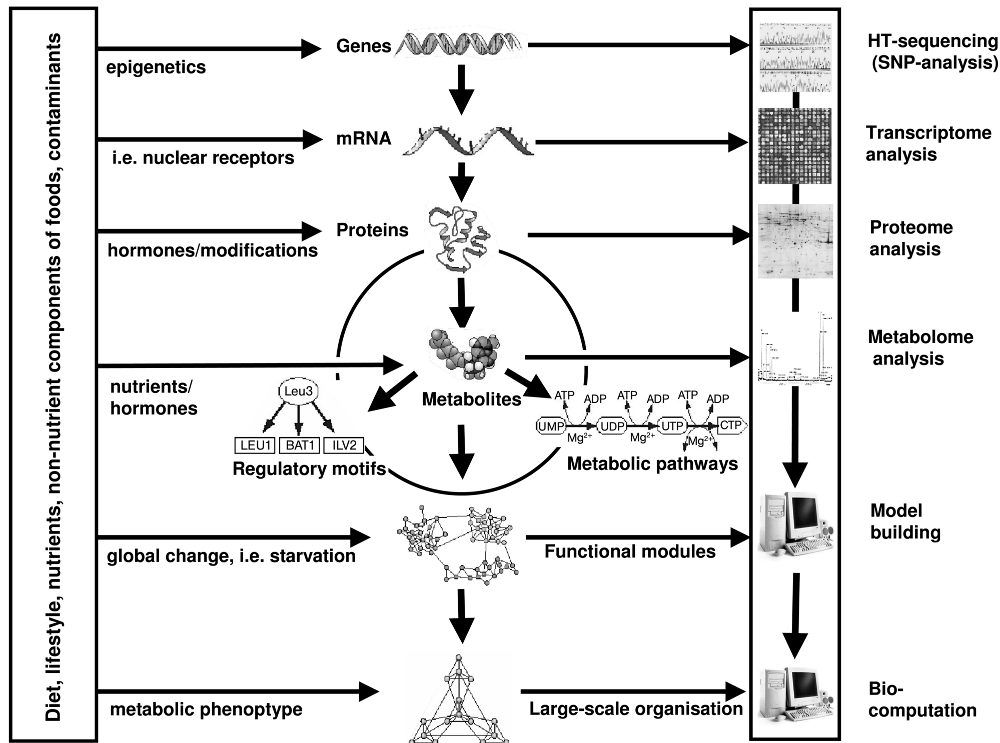

Fig. 1.1. Nutrigenomics as the paradigm for research on environment–genome interactions. Nutritional factors can affect essentially every step from information storage and retrieval, to processing and the execution of biological processes. The emerging new profiling technologies as well as data processing and interpretation tools make the corresponding adaptive changes of mammalian metabolism on a global scale accessible.

terns of subclasses or even of all expressed genes of a given genome. Likewise, proteomics refers to the analysis of the protein complement and metabolomics (also called metabonomics) determines in parallel the accessible metabolites in a cell, tissue, organ, or organism. The data output of these approaches is enormous and often overwhelms our ability to understand the underlying biological processes.

Nutritional science in the past was characterized by well-defined experimental studies based on the experience and knowledge that there is hardly anything else as difficult to standardize as mammalian nutrition. In terms of the biological readouts of nutritional studies, in most cases only a few parameters could be determined simultaneously. The conceptual shift in biological science towards application of high-throughput profiling technologies poses a particular challenge to nutrition researchers as they now have additionally to handle huge data sets derived from the "omic" approaches. How can we use this information to build metabolic topology maps that are easy to comprehend and to interpret and that allow us to navigate to the specific information that we need? Here nutritional science

clearly relies on the new systems biology tools of pathway construction that are based on concepts of control theory, numerical analysis, and stochastic processes. Although systems biology is dependent on "omics" and technologies for data input, it really encompasses the design and use of new analysis tools, and the development of new ways to represent data in a meaningful manner. Nutritional systems biology is the high end of systems biology when it comes to describing the highly diverse changes in metabolism occurring at the same time in different organs or even within an organ in its different cell populations.

1.1.1
What makes Nutrigenomic Research Exceptional?

In contrast to applications of the profiling technologies in drug discovery or toxicology testing, nutrigenomics deals with some exceptional problems. Drug and xenobiotic testing usually determines the consequences of just one compound on the background of a limited number of relevant genes but an otherwise fairly stable environment. Of course, the test compounds may undergo extensive metabolism and the bioactivity could as an integrated read-out result from both the parent compound and the metabolite(s). However, assessing the metabolic response to complex foods is like looking at hundreds of test compounds at the same time and a highly diverse response over time and spatial location (i.e. organ, cell type in an organ).

The human genome and the genetic variation within the human population are the result of high and persistent evolutionary pressures via processes of gene mutation, selection, and random drift. Nutrition has thereby shaped the human genome like no other environmental factor. Individual dietary components can affect gene mutation rates and nutrient availability affects, for example, fetal viability and modifies the penetrance of deleterious genetic lesions [1–5].

As part of the evolutionary pressure it was essential for life that mammals can adapt quickly to changes in their nutritional environment while maintaining metabolism to satisfy the needs of a high rate of ATP production and the production of all building blocks required for cell and tissue renewal and maintenance. Adaptation to food availability in terms of energy as well as individual (essential) nutrients requires very fast but also sustained responses that simultaneously change a huge set of interconnected metabolic processes. This is mainly achieved by hormones that can be classified by their chemical nature (i.e. peptide hormones, amino acid derivatives, or steroids) and/or the mode and time frame of their action. When looking at the effects of a diet on the genetic response, individual nutrients such as carbohydrates, lipids, proteins, or minerals such as calcium directly affect hormone secretion and these hormones adjust cellular functions via specific receptors and a multitude of intracellular signaling events for allowing the required metabolic changes to occur within milliseconds and/or by sustained responses over hours. Moreover, certain nutrients and metabolites directly affect gene expression via interaction with specific cellular targets, including nuclear receptors and response elements, and thereby mediate the integration of extracellular

signals (hormones) and signals from the intracellular environment. Allosteric control mechanisms of protein functions are also an integral part of this synchronization of signal inputs from the extracellular and intracellular environment. Figure 1.2 provides a simplified view of the integrative nature of the input signals for adjusting metabolism to alterations in the nutritional environment. They key ques-

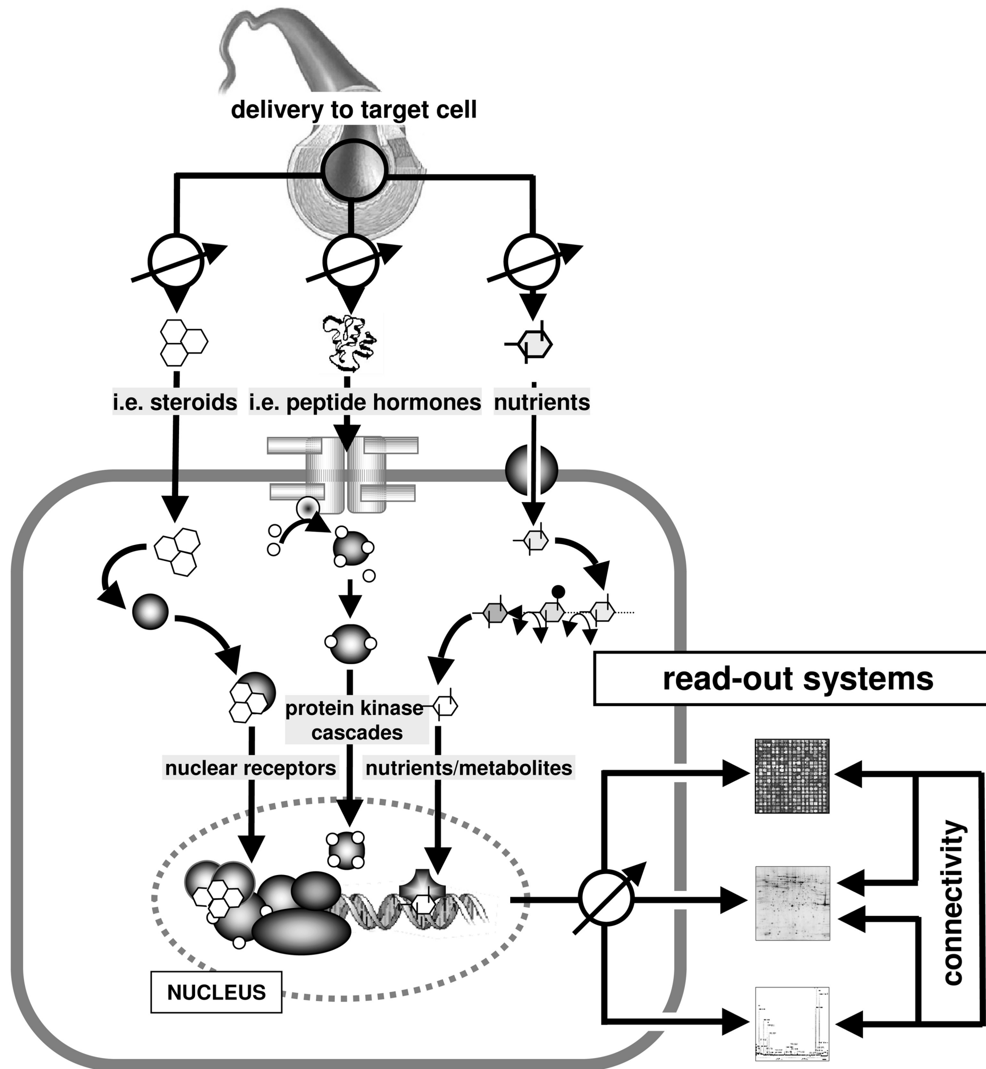

Fig. 1.2. A simplified model depicting the integrative nature of signal processing for transmitting changes in the nutritional environment into the adaptation of the transcriptome, proteome, and metabolome of a cell system.

tion is whether the ''omics'' technologies combined with advanced data analysis and interpretation tools allow us to reconstruct and understand the underlying sensing and signal integration mechanisms and their multidimensional wiring that in the end permit cells to regulate rates of nutrient transport and storage capacity, to fine-tune the flux of intermediates through metabolic routes and branching points, and to restructure the cellular transcriptome and proteome.

1.1.2 Transcript Profiling in Nutrition Research

For historical reasons, transcript profiling has dominated high-throughput genomic studies in mammalian systems since this technology has been around for quite some time [6–9]. Moreover, various commercial systems for easy-to-handle array-based screening applications are available [6–10]. Transcript-profiling experiments so far have often followed a simple experimental design in which, for example, cells or organisms are exposed to an altered nutritional environment (absence or presence of a particular compound) and are then assayed for changes in gene expression [11–13]. These first-generation experiments led to the general conclusion that the cells often respond to quite different environmental conditions with an overlapping response of a battery of genes, although these outputs most probably originate from multiple signaling pathways.

Most microarray studies in the nutrigenomics area so far have the character of snapshots. Based on the high costs of the arrays, pooled RNA samples and/or only a few arrays have been used for analysis. To come from the snapshot approach to more consistent and reliable data, time-series of changes in gene expression as well as repeated and statistically valid measurements are required. As the costs of commercial arrays are expected to drop considerably in the future and as more small-scale targeted and cheaper arrays become available, better microarray data are expected to be produced. It is also essential that the procedures of how the study was conducted and how the array experiments have been performed are well described and data need to be collected and deposited in a standardized format. ArrayExpress (www.ebi.ac.uk/arrayexpress) [14] is the database for collecting information about microarray experiments and is provided by the European Bioinformatics Institute (EBI). ArrayExpress is the world's first database for storage of microarray information that conforms agreed community standards of MIAME (Minimum Information About a Microarray Experiment) devised by the Microarray Gene Expression Data (MGED) Society (www.mged.org) [15]. Since the nutritional science community has no tradition yet in using transcriptome analysis tools it is advised to adopt these standards quickly. Under the umbrella of the European Nutrigenomics Organization (www.nugo.org) a first nutrigenomics-specified MIAME version has been developed and this should in the future allow via ArrayExpress the sharing of vast amounts of microarray-based data with the global science community. In addition, many journals require or recommend authors of microarray data-based papers to submit their data to a MIAME-compliant database.

In its application as a screening approach to nutrition- or nutrient-dependent gene expression analysis, transcript profiling may lead to numerous newly identified genes/mRNA species that respond – not necessarily as expected – to the particular treatment. Before starting to bring a biological meaning into the observation it is highly recommended to use an independent method such as reverse transcriptase polymerase chain reaction (RT-PCR) or Northern blotting to check the magnitude of the changes in the mRNA level of the identified target gene(s), as the reliability of gene expression changes depends on a variety of parameters and particularly on the applied normalization method. In most cases, array data slightly underestimate the changes in transcript levels but there is also often a considerable number of transcripts that are not confirmed as significantly changed in level when assayed by other methods. Nevertheless, global transcript profiling can be seen as an expedition into the terra incognita of molecular nutrition by identifying novel genes, mechanisms and/or pathways by which a dietary maneuver changes cell physiology. The downside of transcriptomics is that one can get lost in the attempt to understand why the changes happen and although hours of scanning of the relevant literature is a rewarding learning exercise it may not provide the answer.

Although currently mainly used in basic science applications, global gene expression analysis is beginning to move from the laboratories to large-scale clinical trials as a tool in diagnostics [16–18]. In the field of human nutrition, signatures or unique patterns of gene expression profiles are expected to be used to describe a nutritional condition or may even allow disease states – even preclinical ones – to be determined [19]. The potential of this technology to improve diagnosis and tailored treatment of human diseases becomes obvious in the area of cancer diagnosis. Several comprehensive studies have demonstrated the utility of gene expression profiles for the classification of tumors into characteristic and clinically relevant subtypes and the prediction of clinical outcomes [16–18, 20, 21]. Applied to human nutrition, gene expression profiling is of course limited (a) by the available cells that should preferentially be obtained by non-invasive techniques, (b) by the genetic heterogeneity of the human population, and (c) by the highly diverse dietary habits and lifestyles. Nevertheless, transcriptome analysis studies for exploring whether characteristic patterns or signatures reflecting the nutritional status in a human population can be obtained need to be performed to explore the scientific and diagnostic value of this technology.

1.1.3
Proteome Profiling in Nutritional Sciences

The term ''proteome'' was introduced as the complement of the genome and relates to the goal of determining all transcribed and translated open reading frames from a given genome. Analysis of the proteome is beginning to emerge as a second high-throughput tool for nutrition research. The revival of two-dimensional gel electrophoresis (2D-PAGE) but with high resolution, the advanced instrumentation and elegant software tools now available for gel analysis, and the enormous ad-

vancements in mass spectrometry have made proteomics applications a practical alternative screening method in the nutrigenomics tool box.

2D-PAGE separates proteins according to charge (isoelectric point: pI) by isoelectric focusing (IEF) in the first dimension and according to size (molecular mass) by sodium dodecyl sulfate PAGE (SDS-PAGE) in the second dimension. It therefore has a unique capacity for the resolution of complex mixtures of proteins, permitting the simultaneous analysis of hundreds or even thousands of gene products [22]. However, not all proteins are resolved and separated equally well by 2D-PAGE. Very alkaline, hydrophobic, and integral membrane proteins as well as high molecular weight proteins are still a problem. In some cases, a prefractionation according to cellular compartment (membranes, microsomes, cytosol, mitochondria) or according to protein solubility by classical means may be necessary [23, 24]. In addition, proteins of low cellular abundance, which may be particularly important in view of their cellular functions for example in signaling pathways, are still very difficult to be resolved in the presence of large quantities of housekeeping proteins [25]. However, new concepts are constantly being developed that employ for example tagging techniques [26] or enrich the minor proteins prior to separation in 2D gels.

The most common procedure for the identification of a protein spot in a gel is currently the peptide mapping or "fingerprint" analysis, but other techniques and approaches can also be applied. For peptide mapping, protein-containing spots are excised from the gel before the gel is altered chemically to make the protein accessible for hydrolysis by a protease such as trypsin [27]. Based on this site-specific enzymatic hydrolysis, a distinct and characteristic pattern of peptide fragments of a given protein serves as the peptide mass fingerprint. The mixture of peptides isolated by digestion with the protease is usually submitted to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) analysis to determine the corresponding peptide masses that are characteristic for a given protein. The mass spectrum obtained is submitted to computer programs that apply various algorithms for interpretation of the peptide pattern and to predict the protein based on a comparison with masses predicted by "virtual digestion" of identified open reading frames in a given genome [28]. Post-translational modifications of proteins such as addition of phosphate groups, hydroxylations at lysine or proline residues, glycosylations, or addition of fatty acids may also be identified by fragment analysis and structural TOF (FAST) or other techniques. Deviations of measured from predicted masses may be due to polymorphisms in coding sequences with subtle amino acid substitutions or even more pronounced with deletions or insertions. Such changes in the primary sequence can be resolved but strongly depend on the type of substitution and may require internal peptide Edman sequencing or more advanced mass analysis by electrospray ionization mass spectroscopy.

It needs to be emphasized here that proteome analysis is straightforward if one assesses the effects of a treatment for example in cultured cells or cell lines since here a homogeneous population of cells is analyzed. When tissue samples are utilized that contain different cell populations with different expression profiles,

proteome analysis becomes a difficult task and may require the separation of the different cell populations by means of cell-specific surface markers and immuno-affinity techniques or by laser-driven microdissection approaches [29, 30]. The future of more simple proteome analysis tools may be the use of antibody libraries that contain specific antibodies raised against any expressed open reading frame and taking proteome analysis onto the format of high-throughput microplate assays that allow essentially every protein to be identified and quantified easily [31].

There are only very few examples of proteome analysis studies in nutrition research. However, proteome analysis is an interesting tool that assesses changes of the steady-state protein levels as the prime functional units without the need for proof and the worry that changes in the transcript level may not translate into corresponding changes in the level of the encoded protein. Combining transcriptome and proteome techniques in analysis of the same sample has the charm of assessing both layers of information flow in adaptation of metabolism and to separate true co-regulation processes and seemingly uncoupled changes in mRNA and corresponding protein level simultaneously.

1.1.4 Metabolite Profiling in Nutritional Science

Various new approaches to assess globally the pattern and concentrations of a vast spectrum of metabolites in biological samples are currently under development. Metabolomics or metabolite profiling techniques are mainly based on gas chromatography combined with mass spectrometry (GC/MS) or liquid chromatography in combination either with electrospray ionization/mass spectrometry (LC/ESI/MS) or with nuclear magnetic resonance spectroscopy (LC/NMR). In contrast to genome, transcriptome, and proteome profiling technologies, which monitor target molecules of similar chemical nature such as DNA, RNA, and proteins, metabolite profiling has to deal with metabolites that vary considerably in chemical nature, molecular weight, and physical properties. This is a real challenge for analytical techniques and consequently there is no single analytical platform that allows the multiparallel analysis of the complete metabolome. However, GC/MS- and LC/NMR-based technologies have proven to be valid in producing robust metabolite profiles from biological samples. Similar to data from transcriptome and proteome analysis, metabolite fingerprints may be used to generate and refine metabolic pathway maps and to identify co-regulation phenomena of whole metabolic networks or functional modules. A variety of statistical methods and visualization tools, such as the principal component analysis, can be used to describe the mostly pleiotropic changes in metabolite spectra.

It needs to be emphasized that at the current state of technology only around 20% of the metabolites present in the mammalian cell can be identified and reliably annotated. To determine the nature of the fast growing number of the yet unknown analytes requires a huge international effort to turn metabolomics into a more powerful tool. The number of metabolites in a typical eukaryotic organism is predicted to range from 4000 to 20 000 individual compounds [32]. Although

this number is impressively high and may be frightening, metabolite spectra reduce the number of components to deal with compared with the much higher number of mRNA and protein entities dealt with when performing transcript and proteome profiling.

Similar to the other high-throughput approaches, data obtained from metabolomics need highly standardized formats for disposition and their linkage to interpretation tools. A consortium recently has outlined these requirements and has given a framework by using examples from plant metabolomics [33].

As in most areas of post-genomic profiling technologies, nutritional science is way behind other fields such as microbiology, plant sciences, or drug and environmental toxicology in applying metabolomics approaches. It is obvious that body fluids such as serum or urine that can be obtained easily are primarily used in both animal and human studies to assess the signatures of the contained metabolites. In most cases ^{1}H nuclear magnetic resonance (NMR) spectroscopic analysis is applied to the body fluid samples from animal or human studies [34, 35]. The obtained complex metabolic profiles are usually submitted to multivariate statistical analysis to obtain patterns. Such a pattern-recognition analysis of NMR spectroscopic data can be performed without the need to assign all of the spectral peaks to specific metabolites before analysis and even provides time-related metabolic changes. Various examples applied in toxicology research – mainly in rodent models with urine sample NMR analysis – demonstrate the power of the techniques. The most impressive metabonomics study in humans was performed with serum samples of normal volunteers and patients at various stages of coronary heart disease. Based on pattern-recognition analysis not only the presence but also the severity of the disease could be determined based on the NMR spectra [36]. It can be envisaged that the technology will soon be taken into population screening.

A first urine sample NMR-based screen of 150 volunteers from Britain and Sweden has recently been reported [37]. The urine samples analysed via principal component analysis displayed characteristic differences related to dietary and cultural habits between the subjects of both countries. Various centers dedicated to nutritional sciences in Europe and the US have identified metabolomics as their strategic field in the nutrigenomics area and we expect to see a rapid increase in the number of studies in this area. It is anticipated that metabolite screens will be used to identify signatures that resemble certain dietary habits, that define the intake of particular food components, or that classify disease states linked to nutrition and nutrient intake.

1.1.4.1 **Metabonomics Goes Dynamic**

The next level of metabolite analysis uses advanced mass spectrometry and NMR to assess the route of individual metabolites by isotope tracer techniques. The labeling of a compound *in vivo* with stable isotopes enables the biosynthesis of differentially mass-labeled metabolite mixtures, which then can be detected by mass isotopomer ratio analysis to follow the flow of atoms through metabolites and pathways and help to identify the molecular switches that guide the compounds through metabolic chains [38]. The beauty of the application of mass isotopomer

analysis in combination with powerful calculation algorithms for the quantification of intracellular metabolic fluxes has recently been demonstrated for central carbon metabolism in *Escherichia coli*. The proposed new method proved to be reliable and capable of obtaining information on the biochemical changes involved in the regulation of acetate and glucose metabolism in *E. coli* K12 cells [39]. As nutritional science has expertise in applying stable isotopes in human studies for the characterization of metabolic processes, metabolomics combined with isotopomer ratio analysis to assess metabolic isotope fluxes, for example in genetically well-characterized individuals or subgroups, is the arena for the next generation of nutrition researchers.

1.1.5
Cell Biology and Genetic Tools for Nutrigenomics Research

Conceptually, nutrigenomics research is based on either gene-driven or phenotype-driven approaches. The gene-driven approaches use genomic information for identifying, cloning, expressing, and characterizing genes and their products at the molecular level. As we are still far from understanding the role of every encoded open reading frame in a mammalian genome, animal models are of central importance for assigning genes to functions. Phenotype-driven approaches characterize phenotypes of naturally occurring variants to identify the genes, the relevant single-nucleotide polymorphisms (SNP) or haplotypes that are either responsible for or associated (in statistical terms) with the particular phenotype. In most cases this is done without knowing the exact underlying molecular mechanisms. Of course, the two strategies are highly complementary at virtually all levels of analysis and lead collectively to the correlation of genotypes and phenotypes. Because nature has not provided human inborn errors of metabolism that demonstrate the phenotypical consequences of individual gene or protein malfunction, the role of single genes or groups of genes in the makeup of metabolism needs to be analyzed in more simple models than humans. Targeted gene inactivation (''knock-out'') or selective overexpression (''knock-in'') models employing experimental animals from fruitflies (*Drosophila melanogaster*) to nematodes (*Caenorhabditis elegans*) or mice and rats or human cell lines will eventually reveal the roles that individual genes play in the orchestrated way metabolism works. These approaches have already produced a large number of animal lines missing one or several genes or overexpressing others. The availability of the large-scale knockout collections will accelerate the wet-laboratory work necessary to provide an understanding of the biological roles of the various players in nutrition-triggered signal transduction and gene regulation processes. Although very elegant as genetic tools to unravel metabolic changes, unfortunately, these maneuvers quite often do not produce the predicted or any distinct phenotype.

The more advanced transgenic technologies in animals through controlled cell- or organ-specific and/or time-dependent gene inactivation or induction of expression allow the analysis of phenotypical consequences in even more elegant ways. They appear also particularly helpful when simple gene disruption is lethal for

the developing fetus or newborn. In simple cell models and even complex organisms (the best example being *Caenorhabditis elegans*) RNA interference techniques (RNAi) have made it easier to suppress or at least markedly reduce expression of the protein of interest in order to assess the phenotypic consequences [40, 41].

Assigning gene to function is the most critical part and this relies currently on genetic models. In combining the technologies of targeted gene inactivation and RNAi or selective overexpression with the "omics" technologies, the annotation of gene functions is greatly improved but the redundancy in biological systems also becomes visible.

Understanding the consequences of operational shifts in genetic circuits and cellular systems is and will remain a challenge. In emerging new and sophisticated metabolic network analysis tools, the metabolites are represented by interconnected nodes that show correlative behavioral changes and the actions of these metabolic networks are studied on the basis of the strength of correlations between the metabolites that make up the network [42, 43]. To understand these nodular systems and to determine the connectivity of the layers of the transcriptome, proteome, and metabolome, comprehensive approaches to measure metabolites, proteins, and/or mRNA simultaneously from the same sample are required.

How do we cope with the data generated by the high-throughput data acquisition and systems approaches? Well, the best answer appears to be to have someone on the research team who is expert in the area of analysis of these data and who is willing to learn a bit of nutritional science, with the nutritionists willing to learn a bit of advanced statistics and bioinformatics.

1.2 Nutrigenetics – Examples and Limitations

Nutrigenetics aims to understand the effect of genetic variations on the interaction between the diet and disease or on nutrient requirement. Consequently the major goal is to identify and characterize gene variants associated or responsible for differential responses to nutritional factors. In the final stage, nutrigenetics could provide the rationale for recommendations regarding the risks and benefits of a particular diet or dietary components based on the individual's genetic makeup. The quite impressive variations in the phenotype of "classical" monogenetic diseases such as phenylketonuria or familial hypercholesterolemia, however, tells us what kind of challenge we are facing when nutrigenetic approaches are applied to common multifactorial disorders such as diabetes, cardiovascular disease, or cancer. Although the methods for detecting single-nucleotide polymorphisms (SNPs) or haplotypes are improved constantly and the next generation of microarrays that cover 500 000 SNPs on one chip will be available soon, phenotype analysis and assigning alterations in protein functions to an SNP or haplotype is going to be the pinhole. Although mostly inconclusive, preliminary results involving gene–diet interactions for cardiovascular diseases and cancer suggest that the concept could work and that we will be able to harness the information contained in our genome.

Most of the available data are derived from molecular epidemiology studies. As all multifactorial nutrition-dependent diseases require a long period of exposure to the same or similar dietary patterns to develop a disease phenotype [44] epidemiological studies are the tool of choice to assess genetic variation and disease development or progression.

1.2.1 Genes, Diet, and Cardiovascular Disease

Dyslipidemia is commonly associated with the development of atherosclerosis and can be caused by improper function of a variety of proteins that control lipid homeostasis, such as nuclear factors, binding proteins, apolipoproteins, enzymes, lipoprotein receptors, and hormones (see also chapter 15). Polymorphisms have been identified in most of these components and many of the underlying genes have been explored in terms of diet–gene interactions [45–47]. Amongst these, the *apolipoprotein E* gene (*apoE*) is the most intensively studied with regard to its effects on low-density lipoprotein (LDL)-cholesterol levels in response to dietary interventions. Genetic variation at the *apoE* locus results from three common alleles in the population, *E4*, *E3*, and *E2*. However, other genetic variants at the *apoE* locus have been described as well [48].

Besides the fact that LDL-cholesterol levels were highest in subjects carrying the *apoE4* isoform [49, 50], this association was especially prominent in populations consuming diets rich in saturated fats and cholesterol [51]. These epidemiology data, therefore, indicate that high LDL-cholesterol levels are manifested primarily in the presence of an atherogenic diet but that an individual's response to dietary saturated fat and cholesterol may differ depending on the individual *apoE* alleles. However, it needs to be stressed that especially for *apoE*, investigations of diet–gene interactions have yielded quite diverse outcomes [46, 47]. Significant diet–*apoE* interactions occurred in studies focusing on males, suggesting a significant gene–gender interaction [46, 47]. Baseline lipid levels seem to affect the outcome and significant associations were frequently found only in subjects who were moderately hypercholesterolemic. More consistent effects were reported on the impact of alcohol intake on LDL-cholesterol depending on the *apoE* genotype in men [52]. A negative association between alcohol consumption and LDL-cholesterol was found for carriers of *apoE2*, whereas subjects with *apoE4* displayed a positive correlation. Within these genotype studies *apoA1* has emerged as a primary candidate for genetic variability in high-density lipoprotein (HDL) levels and its gene product plays a crucial role in lipid metabolism and for cardiovascular disease risk [53].

In women it has been found that a G to A transition in the *apoA1* gene is associated with an increase in HDL-cholesterol levels depending on the dietary intake of polyunsaturated fatty acids [54]. Similar to this G/A single-nucleotide polymorphism in *apoA1*, increased HDL levels were found to be associated with a homozygous −514(CC) polymorphism in the *hepatic lipase* gene in response to higher fat contents in the diet [55]. This increase in the level of protective HDL particles was interpreted as a defense mechanism that was not found in subjects carrying the TT

genotype. Interestingly the TT genotype is common in certain ethnic groups, such as African-Americans, and might help to explain their limited ability to adapt rapidly to new nutritional environments [56].

1.2.2 Genes, Diet, and Cancer

Similar to cardiovascular diseases, dietary factors were shown to contribute significantly to the development of cancers [57] with the most prominent effects on colon, gastric, and breast cancer (see also chapter 17). Although there are general guidelines to reduce cancer risk at the population level, a specific protective food or food component has not been identified [58, 59]. Numerous studies using quantitative dietary assessments in large cohort studies and assessing genetic variation in the cohorts such as the European Prospective Investigation into Cancer and Nutrition (EPIC) with 519 978 participants in 23 centers in 10 European countries [60] are being conducted to understand and define the role dietary factors play in the causes of cancer development on the basis of genetic variations.

One of the polymorphisms that is significantly associated with cancer risk is the homozygous (TT) form of the *methylenetetrahydrofolate reductase* (*MTHFR*) gene. The cytosine to thymidine substitution, which converts an alanine residue to a valine, is relatively common and results, in its homozygous form, in hyperhomocysteinemia and an increased cardiovascular disease risk but simultaneously reduced cancer risk [61, 62]. The *MTHFR*-TT genotype displays a reduced enzymatic activity; less 5,10-methylenetetrahydrofolate is used for the remethylation of homocysteine to methionine and hence more substrate appears to be available for thymidine synthesis. In contrast, an increased misincorporation of desoxyuridine nucleotides into DNA in folate deficiency was shown to be mutagenic and this could, for example, explain the increased colon cancer risk observed in humans with a low folate status [63]. This example shows the complexity of the problem and the difficulty of transferring these observations to the level of recommendations.

Regarding homocysteine as an atherogenic factor, the recommendation for individuals with the *MTHFR*-TT genotype must be to normalize the enzyme activity and reduce homocysteine levels by higher rates of remethylation to methionine. It has to be suggested, however, that under conditions where MTHFR activity is normalized, one-carbon flux into the thymidylate cycle may be reduced and, thus, the protective function of the TT polymorphism with regard to cancer development may be lost. Although this hypothesis has not yet been proven, an increased risk for the development of colorectal adenomas was shown for the TT genotype associated with a low folate, vitamin B_{12} and B_6 intake [64].

Another gene for which polymorphisms seem to predispose to cancers by exposure to food carcinogens is the *N-acetyltransferase* (*NAT*) gene (see also chapter 19). NAT is a phase 2 enzyme that is found in two isoforms (NAT1 and NAT2) and is involved in the acetylation of heterocyclic aromatic amines (HAA) as found in heated products. Several polymorphisms have been characterized in NAT1 and NAT2 and some of these polymorphisms have been related to NAT activities of

so-called ''slow'', ''intermediate,'' or ''fast acetylators.'' Although the outcome of studies investigating the association between acetylator phenotype and cancer risk are quite controversial, the NAT2 fast acetylator genotype consistently revealed a higher risk of developing colon cancer in people who consumed relatively large quantities of red meat, which may reflect the greater ability of fast acetylators to activate aromatic amines within the colon mucosa [65, 66].

Glutathione-*S*-transferases (GST) have also been studied in detail with respect to individual cancer risk (see also chapter 19). GSTs are subdivided into the four classes alpha (A), pi (P), mu (M), and theta (T) and for each class various polymorphisms have been described. *GSTM1*- and *GSTT1*-null genotypes appear to confer a high risk for several types of cancer [67]. By their ability for detoxification GSTs play a crucial role in xenobiotic metabolism and respond to a variety of dietary factors with changes in expression level [68]. In humans, it has been suggested that the cancer protective effects (e.g. of cruciferous vegetables) may depend on the ability to induce GSTs and other phase 2 enzymes [69, 70]. On the other hand, besides toxic compounds they also conjugate isothiocyanates, the active ingredients of cruciferae, leading to their excretion. Indeed, a significant protective effect of a high broccoli consumption was only found in subjects with the *GSTM1*-null genotype [71]. The response to isothiocyanates requires the nuclear factor-erythroid 2 p45-related factor 2 (Nrf2). In Nrf2$^{+/+}$ but not Nrf2$^{-/-}$ mice isothiocyanates from broccoli caused a modest increase in GSTM1 and a significant increase in GSTA1/2 and GSTA3 protein [72]. Strain-specific Nrf2 mRNA expression and a T to C transition in the promoter that co-segregates with susceptibility phenotypes in mice [73] and the detection of three SNPs and one triplet repeat in the human Nrf2 promoter [74] makes Nrf2 a candidate itself for individual responses to dietary factors with chemopreventive properties.

Altogether, the examples demonstrate that we are far away from a reasonable dietary advice for an individual or a subpopulation on the basis of the genetic makeup. It is important to keep in mind that gene–nutrient interactions can occur at any time during the process of a disease development and the multistage process of carcinogenesis is a perfect example. Dietary factors can affect essentially every step in cancer initiation and development starting at the level of an initial mutation, by blocking promotion or by stopping progression from the premalignant state to carcinomas or by preventing invasion or metastasis. Only long-term studies with appropriate population sizes, well-reported dietary intakes and reliable genotype and phenotype analysis will help to prove the concept of nutrigenetics, with its final goal of providing a solid scientific basis for individualized ''genotype-based'' dietary advice.

1.3 Conclusions

Nutritional genomics is still in its infancy, but it is predicted to rapidly move to the systems-based ''holistic'' level by using high-throughput technologies and ad-

vanced data analysis tools. Transcript, proteome, and metabolite profiling technologies are constantly being improved and are becoming more convenient but in the end require a substantial investment in equipment and specialized personnel. Although the new technologies already generate insights into nutrition-dependent signal transduction mechanisms and gene regulation phenomena, it is obvious that at present, these studies more often generate hypotheses than deliver true answers. In spite of the apparent profusion of data that overwhelm us already, magnitudes of new data are needed to reach the goal of a comprehensive understanding of signal transduction and gene regulation phenomena that allow the adaptation of mammalian metabolism for maintaining health or that may eventually lead to disease.

References

1. Fenech M (**2002**) Biomarkers of genetic damage for cancer epidemiology. *Toxicology* 181–182, 411–416.
2. Brawley OW (**2003**) Population categorization and cancer statistics. *Cancer Metastasis Rev.* 22, 11–19.
3. Goldman R, Shields PG (**2003**) Food mutagens. *J. Nutr.* 133, 965S–973S.
4. Gluckman PD, Hanson MA (**2004**) Maternal constraint of fetal growth and its consequences. *Semin. Fetal Neonatal Med.* 9, 419–425.
5. Sacks DA (**2004**) Determinants of fetal growth. *Curr. Diab. Rep.* 4, 281–287.
6. Hughes TR, Shoemaker DD (**2001**) DNA microarrays for expression profiling. *Curr. Opin. Chem. Biol.* 5, 21–25.
7. Tuteja R, Tuteja N (**2004**) Serial analysis of gene expression (SAGE): application in cancer research. *Med. Sci. Monit.* 10, 132–140.
8. Welle S. Gene transcript profiling in aging research (**2002**) *Exp. Gerontol.* 37, 583–590.
9. Cook SA, Rosenzweig A (**2002**) DNA microarrays: implications for cardiovascular medicine. *Circ. Res.* 91, 559–564.
10. Shedden K, Chen W, Kuick R, Ghosh D, Macdonald J, Cho KR, Giordano TJ, Gruber SB, Fearon ER, Taylor JM, Hanash S (**2005**) Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinformatics* 6, 26.
11. Daniel H, tom Dieck H (**2004**) Nutrient–gene interactions: a single nutrient and hundreds of target genes. *Biol. Chem.* 385, 571–583.
12. tom Dieck H, Doring F, Fuchs D, Roth HP, Daniel H (**2005**) Transcriptome and proteome analysis identifies the pathways that increase hepatic lipid accumulation in zinc-deficient rats. *J. Nutr.* 135, 199–205.
13. Herzog A, Kindermann B, Doring F, Daniel H, Wenzel U (**2004**) Pleiotropic molecular effects of the pro-apoptotic dietary constituent flavone in human colon cancer cells identified by protein and mRNA expression profiling. *Proteomics* 4, 2455–2464.
14. Brooksbank C, Cameron G, Thornton J (**2005**) The European Bioinformatics Institute's data resources: towards systems biology. *Nucleic Acids Res.* 33 Database Issue, D46–53.
15. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J,

Taylor R, Vilo J, Vingron M (**2001**) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* 29, 365–371.

16 Perez EA, Pusztai L, Van de Vijver M (**2004**) Improving patient care through molecular diagnostics. *Semin. Oncol.* 31, 14–20.

17 Jain KK (**2004**) Applications of biochips: from diagnostics to personalized medicine. *Curr. Opin. Drug Discov. Devel.* 7, 285–289.

18 Mantripragada KK, Buckley PG, de Stahl TD, Dumanski JP (**2004**) Genomic microarrays in the spotlight. *Trends Genet.* 20, 87–94.

19 Muller M, Kersten S (**2003**) Nutrigenomics: goals and strategies. *Nat. Rev. Genet.* 4, 315–322.

20 Baker M, Gillanders WE, Mikhitarian K, Mitas M, Cole DJ (**2003**) The molecular detection of micrometastatic breast cancer. *Am. J. Surg.* 186, 351–358.

21 Macgregor PF, Squire JA (**2002**) Application of microarrays to the analysis of gene expression in cancer. *Clin. Chem.* 48, 1170–1177.

22 Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W (**2000**) The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037–1053.

23 Blackstock WP, Weir MP (**1999**) Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.* 17, 121–127.

24 Cordwell SJ, Nouwens AS, Verrills NM, Basseal DJ, Walsh BJ (**2000**) Subproteomics based upon protein cellular location and relative solubilities in conjunction with composite two-dimensional electrophoresis gels. *Electrophoresis* 21, 1094–1103.

25 Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC (**2000**) The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115.

26 Adam GC, Cravatt BF, Sorensen EJ (**2001**) Profiling the specific reactivity of the proteome with non-directed activity-based probes. *Chem. Biol.* 8, 81–95.

27 Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD (**2000**) Utility of accurate mass tags for proteome-wide protein identification. *Anal. Chem.* 72, 3349–3354.

28 Chamrad DC, Koerting G, Gobom J, Thiele H, Klose J, Meyer HE, Blueggel M (**2003**) Interpretation of mass spectrometry data for high-throughput proteomics. *Anal. Bioanal. Chem.* 376, 1014–1022.

29 Banks RE, Dunn MJ, Forbes MA, Stanley A, Pappin D, Naven T, Gough M, Harnden P, Selby PJ (**1999**) The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis – preliminary findings. *Electrophoresis* 20, 689–700.

30 Simone NL, Paweletz CP, Charboneau L, Petricoin EF III, Liotta LA (**2000**) Laser capture microdissection: beyond functional genomics to proteomics. *Mol. Diagn.* 5, 301–307.

31 LaBaer J, Ramachandran N (**2005**) Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* 9, 14–19.

32 Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L (**2000**) Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* 18, 1157–1161.

33 Jenkins H, Hardy N, Beckmann M, Draper J, Smith AR, Taylor J, Fiehn O, Goodacre R, Bino RJ, Hall R, Kopka J, Lane GA, Lange BM, Liu JR, Mendes P, Nikolau BJ, Oliver SG, Paton NW, Rhee S, Roessner-Tunali U, Saito K, Smedsgaard J, Sumner LW, Wang T, Walsh S, Wurtele ES, Kell DB (**2004**) A proposed framework for the description of plant metabolomics experiments and their results. *Nat. Biotechnol.* 22, 1601–1606.

34 Williams RE, Jacobsen M, Lock EA (**2003**) ^{1}H NMR pattern recognition

and 31P NMR studies with d-Serine in rat urine and kidney, time- and dose-related metabolic effects. *Chem. Res. Toxicol.* 16, 1207–1216.

35 Liu M, Nicholson JK, Lindon JC (**1996**) High-resolution diffusion and relaxation edited one- and two-dimensional ^{1}H NMR spectroscopy of biological fluids. *Anal. Chem.* 68, 3370–3376.

36 Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HW, Clarke S, Schofield PM, McKilligin E, Mosedale DE, Grainger DJ (**2002**) Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nat. Med.* 8, 1439–1444.

37 Lenz EM, Bright J, Wilson ID, Hughes A, Morrisson J, Lindberg H, Lockton A (**2004**) Metabonomics, dietary influences and cultural differences: a ^{1}H NMR-based study of urine samples obtained from healthy British and Swedish subjects. *J. Pharm. Biomed. Anal.* 36, 841–849.

38 Birkemeyer C, Luedemann A, Wagner C, Erban A, Kopka J (**2005**) Metabolome analysis: the potential of in vivo labeling with stable isotopes for metabolite profiling. *Trends Biotechnol.* 23, 28–33.

39 Zhao J, Shimizu K (**2003**) Metabolic flux analysis of *Escherichia coli* K12 grown on 13C-labeled acetate and glucose using GC-MS and powerful flux calculation method. *J. Biotechnol.* 101, 101–117.

40 Poulin G, Nandakumar R, Ahringer J (**2004**) Genome-wide RNAi screens in *Caenorhabditis elegans*: impact on cancer research. *Oncogene* 23, 8340–8345.

41 Sugimoto A (**2004**) High-throughput RNAi in *Caenorhabditis elegans*: genome-wide screens and functional genomics. *Differentiation* 72, 81–91.

42 de la Fuente A, Bing N, Hoeschele I, Mendes P (**2004**) Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics* 20, 3565–3574.

43 Kell DB (**2004**) Metabolomics and systems biology: making sense of the soup. *Curr. Opin. Microbiol.* 7, 296–307.

44 Leong NM, Mignone LI, Newcomb PA, Titus-Ernstoff L, Baron JA, Trentham-Dietz A, Stampfer MJ, Willett WC, Egan KM (**2003**) Early life risk factors in cancer: the relation of birth weight to adult obesity. *Int. J. Cancer* 103, 789–791.

45 Loktionov A (**2003**) Common gene polymorphisms and nutrition: emerging links with pathogenesis of multifactorial chronic diseases (review). *J. Nutr. Biochem.* 14, 426–451.

46 Masson LF, McNeill G, Avenell A (**2003**) Genetic variation and the lipid response to dietary intervention: a systematic review. *Am. J. Clin. Nutr.* 77, 1098–1111.

47 Rubin J, Berglund L (**2002**) Apolipoprotein E and diets: a case of gene–nutrient interaction? *Curr. Opin. Lipidol.* 13, 25–32.

48 Artiga MJ, Bullido MJ, Sastre I, Recuero M, Garcia MA, Aldudo J, Vazquez J, Valdivieso F (**1998**) Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. *FEBS Lett.* 421, 105–108.

49 Ordovas JM, Litwack-Klein L, Wilson PW, Schaefer MM, Schaefer EJ (**1987**) Apolipoprotein E isoform phenotyping methodology and population frequency with identification of apoE1 and apoE5 isoforms. *J. Lipid Res.* 28, 371–380.

50 Schaefer EJ, Lamon-Fava S, Johnson S, Ordovas JM, Schaefer MM, Castelli WP, Wilson PW (**1994**) Effects of gender and menopausal status on the association of apolipoprotein E phenotype with plasma lipoprotein levels. Results from the Framingham Offspring Study. *Arterioscler. Thromb.* 14, 1105–1113.

51 Davignon J, Gregg RE, Sing CF (**1988**) Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* 8, 1–21.

52 Corella D, Tucker K, Lahoz C, Coltell O, Cupples LA, Wilson PW, Schaefer EJ, Ordovas JM (**2001**) Alcohol drinking determines the effect of the APOE locus on LDL-cholesterol concentrations in men: the Framingham Offspring Study. *Am. J. Clin. Nutr.* 73, 736–745.

53 Segrest JP, Li L, Anantharamaiah GM, Harvey SC, Liadaki KN, Zannis V (**2000**) Structure and function of apolipoprotein A-I and high-density lipoprotein. *Curr. Opin. Lipidol.* 11, 105–115.

54 Ordovas JM, Corella D, Cupples LA, Demissie S, Kelleher A, Coltell O, Wilson PW, Schaefer EJ, Tucker K (**2002**) Polyunsaturated fatty acids modulate the effects of the APOA1 G-A polymorphism on HDL-cholesterol concentrations in a sex-specific manner: the Framingham Study. *Am. J. Clin. Nutr.* 75, 38–46.

55 Ordovas JM, Corella D, Demissie S, Cupples LA, Couture P, Coltell O, Wilson PW, Schaefer EJ, Tucker KL (**2002**) Dietary fat intake determines the effect of a common polymorphism in the hepatic lipase gene promoter on high-density lipoprotein metabolism: evidence of a strong dose effect in this gene-nutrient interaction in the Framingham Study. *Circulation* 106, 2315–2321.

56 Tai ES, Corella D, Deurenberg-Yap M, Cutter J, Chew SK, Tan CE, Ordovas JM (**2003**) Dietary fat interacts with the –514C > T polymorphism in the hepatic lipase gene promoter on plasma lipid profiles in a multiethnic Asian population: the 1998 Singapore National Health Survey. *J. Nutr.* 133, 3399–3408.

57 Doll R, Peto R (**1981**) The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl Cancer Inst.* 66, 1191–1308.

58 Riboli E, Norat T (**2003**) Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. *Am. J. Clin. Nutr.* 78, 559S–569S.

59 Voorrips LE, Goldbohm RA, van Poppel G, Sturmans F, Hermus RJ, van den Brandt PA (**2000**) Vegetable and fruit consumption and risks of colon and rectal cancer in a prospective cohort study: The Netherlands Cohort Study on Diet and Cancer. *Am. J. Epidemiol.* 152, 1081–1092.

60 Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, Clavel-Chapelon F, Thiebaut A, Wahrendorf J, Boeing H, Trichopoulos D, Trichopoulou A, Vineis P, Palli D, Bueno-De-Mesquita HB, Peeters PH, Lund E, Engeset D, Gonzalez CA, Barricarte A, Berglund G, Hallmans G, Day NE, Key TJ, Kaaks R, Saracci R (**2002**) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 5, 1113–1124.

61 Chen J, Giovannucci E, Kelsey K, Rimm EB, Stampfer MJ, Colditz GA, Spiegelman D, Willett WC, Hunter DJ (**1996**) A methylenetetrahydrofolate reductase polymorphism and the risk of colorectal cancer. *Cancer Res.* 56, 4862–4864.

62 Ma J, Stampfer MJ, Giovannucci E, Artigas C, Hunter DJ, Fuchs C, Willett WC, Selhub J, Hennekens CH, Rozen R (**1997**) Methylene-tetrahydrofolate reductase polymorphism, dietary interactions, and risk of colorectal cancer. *Cancer Res.* 57, 1098–1102.

63 Slattery ML, Potter JD, Samowitz W, Schaffer D, Leppert M (**1999**) Methylenetetrahydrofolate reductase, diet, and risk of colon cancer. *Cancer Epidemiol. Biomarkers Prev.* 8, 513–518.

64 Ulrich CM, Kampman E, Bigler J, Schwartz SM, Chen C, Bostick R, Fosdick L, Beresford SA, Yasui Y, Potter JD (**1999**) Colorectal adenomas and the C677T MTHFR polymorphism: evidence for gene-environment interaction? *Cancer Epidemiol. Biomarkers Prev.* 8, 659–668.

65 Chen J, Stampfer MJ, Hough HL, Garcia-Closas M, Willett WC, Hennekens CH, Kelsey KT, Hunter DJ (**1998**) A prospective study of N-acetyltransferase genotype, red meat intake, and risk of colorectal cancer. *Cancer Res.* 58, 3307–3311.

66 Roberts-Thomson IC, Ryan P, Khoo KK, Hart WJ, McMichael AJ, Butler RN (**1996**) Diet, acetylator phenotype, and risk of colorectal neoplasia. *Lancet* 347, 1372–1374.

67 Habdous M, Siest G, Herbeth B, Vincent-Viry M, Visvikis S (**2004**) [Glutathione S-transferases genetic polymorphisms and human diseases: overview of epidemiological studies]. *Ann. Biol. Clin.* 62, 15–24.

68 Rock CL, Lampe JW, Patterson RE (**2000**) Nutrition, genetics, and risks of cancer. *Annu. Rev. Public Health* 21, 47–64.

69 Seow A, Shi CY, Chung FL, Jiao D, Hankin JH, Lee HP, Coetzee GA, Yu MC (**1998**) Urinary total isothiocyanate (ITC) in a population-based sample of middle-aged and older Chinese in Singapore: relationship with dietary total ITC and glutathione S-transferase M1/T1/P1 genotypes. *Cancer Epidemiol. Biomarkers Prev.* 7, 775–781.

70 Nijhoff WA, Grubben MJ, Nagengast FM, Jansen JB, Verhagen H, van Poppel G, Peters WH (**1995**) Effects of consumption of Brussels sprouts on intestinal and lymphocytic glutathione S-transferases in humans. *Carcinogenesis* 16, 2125–2128.

71 Lin HJ, Probst-Hensch NM, Louie AD, Kau IH, Witte JS, Ingles SA, Frankl HD, Lee ER, Haile RW (**1998**) Glutathione transferase null genotype, broccoli, and lower prevalence of colorectal adenomas. *Cancer Epidemiol. Biomarkers Prev.* 7, 647–652.

72 McWalter GK, Higgins LG, McLellan LI, Henderson CJ, Song L, Thornalley PJ, Itoh K, Yamamoto M, Hayes JD (**2004**) Transcription factor Nrf2 is essential for induction of NAD(P)H:quinone oxidoreductase 1, glutathione S-transferases, and glutamate cysteine ligase by broccoli seeds and isothiocyanates. *J. Nutr.* 134, 3499S–3506S.

73 Cho H-Y, Jedlicka AE, Reddy SP, Zhang L-Y, Kensler TW, Kleeberger SR (**2002**) Linkage analysis of susceptibility to hyperoxia. Nrf2 is a candidate gene. *Am. J. Respir. Cell Mol. Biol.* 26, 42–51.

74 Yamamoto T, Yoh K, Kobayashi A, Ishii Y, Kure S, Koyama A, Sakamoto T, Sekizawa K, Motohashi H, Yamamoto M (**2004**) Identification of polymorphisms in the promoter region of the human NRF2 gene. *Biochem. Biophys. Res. Commun.* 321, 72–79.