Nutrigenomics: From Molecular Nutrition to Prevention of Disease

LYDIA AFMAN, PhD; MICHAEL MÜLLER, PhD

ABSTRACT

Until recently, nutrition research concentrated on nutrient deficiencies and impairment of health. The advent of genomics—interpreted broadly as a suite of high-throughput technologies for the generation, processing, and application of scientific information about the composition and functions of genomes—has created unprecedented opportunities for increasing our understanding of how nutrients modulate gene and protein expression and ultimately influence cellular and organismal metabolism. Nutritional genomics (nutrigenomics), the junction between health, diet, and genomics, can be seen as the combination of molecular nutrition and genomics. The diverse tissue and organ-specific effects of bioactive dietary components include gene-expression patterns (transcriptome); organization of the chromatin (epigenome); protein-expression patterns, including post-translational modifications (proteome); as well as metabolite profiles (metabolome). Nutrigenomics will promote an increased understanding of how nutrition influences metabolic pathways and homeostatic control, how this regulation is disturbed in the early phases of diet-related disease, and the extent to which individual sensitizing genotypes contribute to such diseases. Eventually, nutrigenomics will lead to evidence-based dietary intervention strategies for restoring health and fitness and for preventing diet-related disease. In this review, we provide a brief overview of nutrigenomics from our point of view by describing current strategies, future opportunities, and challenges.


Expectations of nutrigenomics (1-8) are extremely high, but progress is rather slow. The field of nutrigenomics faces the challenge of establishing a strong basic research foundation while simultaneously surmounting a number of technological hurdles. The coming years will likely require patience, realistic expectations, and strong advocacy for the needed research funding. In addition, we must be innovative in our approaches to the complex problems that nutrigenomics presents. For example, a single nutrigenomics experiment can generate an enormous amount of data. We must quickly learn how to extract useful biological information from these data. Nutrigenomics must address management and storage of different types of data derived from the various technological platforms used, development and application of new biostatistical algorithms, and inaccessibility of tissue samples from healthy volunteers.

Challenges of this scope and magnitude cannot be solved by one research group alone and will likely require collaboration among a number of different research teams. Furthermore, many of the genomics technologies are quite expensive. Not surprisingly, an increasing number of large national and international nutrigenomics research clusters are being formed to jointly address these and similar challenges. These clusters typically focus on specific aspects of nutrition research and disease prevention (Figure 1) that can be approached with nutrigenomics research. For such a collaborative effort to be successful, there must be excellent communication among scientists from different disciplines (eg, nutrition, molecular biology, medicine, genomics, bioinformatics), and there must be a high level of compatibility of the genomics data produced in the various laboratories (9-11). Several of these clusters or centers also involve collaboration between the food industry and academia. The food industry recognizes the need for nutrigenomics re-
search as a basis for developing the concept of “personalized diets,” for identifying molecular biomarkers or new bioactive food ingredients, and for validating the effectiveness of these bioactive ingredients as functional food components or nutraceuticals.

An important aim of nutrigenomics research is to study genome-wide influences of nutrition, with specific focus on the role of metabolic stress in the genesis of the metabolic syndrome, the collection of phenotypes combining inflammation, metabolic stress, insulin resistance, and diabetes (12) (Figure 2). This goal is rather ambitious, but is based on the idea that nutrition should focus primarily on health and disease prevention and be complementary to pharmacological therapy, which targets the pathophysiological aspects of disease. To realize this goal, new genomics-based phenotypical biomarkers are needed that allow early detection of the onset of disease or, ideally, the predisease state of the metabolic syndrome, a condition referred to as metabolic stress.

To approach this complex condition, molecular nutrition research on organ-specific dietary response patterns using transgenic and knock-out mouse models is combined with genomic technologies. From a molecular standpoint, nutrients are considered to be “signaling mol-

<table>
<thead>
<tr>
<th>Consortium</th>
<th>Country</th>
<th>Focus</th>
<th>Web site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center of Excellence for Nutritional Genomics</td>
<td>United States</td>
<td>Personalized diet; diet–gene interactions</td>
<td><a href="http://www.nutrigenomics.ucdavis.edu">www.nutrigenomics.ucdavis.edu</a></td>
</tr>
<tr>
<td>Dutch Nutrigenomics Consortium</td>
<td>The Netherlands</td>
<td>Metabolic syndrome; early biomarkers</td>
<td><a href="http://www.nutrigenomics.nl/ngc">www.nutrigenomics.nl/ngc</a></td>
</tr>
<tr>
<td>Network of Excellence in Nutrigenomics (NUGO)</td>
<td>Europe (European Community)</td>
<td>Establishment of a European Nutrigenomics Research Network (research, training, standardization)</td>
<td><a href="http://www.nugo.org">www.nugo.org</a></td>
</tr>
<tr>
<td>Centre of Excellence in Nutrigenomics</td>
<td>New Zealand</td>
<td>Crohn’s disease; new food bioactives</td>
<td><a href="http://www.nutrigenomics.org.nz">www.nutrigenomics.org.nz</a></td>
</tr>
<tr>
<td>Functional Food Genomics</td>
<td>Japan</td>
<td>Biomarkers and bioactive food ingredients</td>
<td></td>
</tr>
<tr>
<td>Nutrigenomics Network</td>
<td>Germany</td>
<td>Complex diseases; diet–gene interactions</td>
<td><a href="http://www.nutrigenomik.de">www.nutrigenomik.de</a></td>
</tr>
</tbody>
</table>

**Figure 1.** Overview of selected international nutrigenomics consortia and networks.

**Figure 2.** Development of complex, multifactorial, polygenic diseases, such as metabolic syndrome. Nutrition is primarily focused on health and on the earliest phases of disease pathology. In order to effectively apply dietary strategies to prevent disease or to recover homeostasis, validated early biomarkers of the disease state are needed. Nutrition and pharma (pharmacology) are complementary approaches to apply to metabolic stress or metabolic syndrome. Interestingly, there is considerable overlap between cellular targets for nutritional and pharmacological intervention, such as peroxisome proliferator activator receptor-α or peroxisome proliferator activator receptor-γ, which bind fatty acids and fibrates or fatty acids and thiazolidinediones, respectively.
ecules” that, through appropriate cellular sensing mechanisms, result in translation of these dietary signals into changes in gene, protein, and metabolite expression (8). Such an approach allows insight into the mechanisms of nutrition at the molecular level (ie, what happens in our cells and organs when we eat, when we do not eat, or when we eat too much). On a genomic level, these molecular changes serve as dietary “signatures” or fingerprints that can precisely annotate the phenotype, particularly under conditions of metabolic stress and early phases of organ-specific insulin resistance. Ultimately, the aim is to extrapolate findings from studies with mice and cells to human beings, where the impact of the genotype must be taken into account in order to estimate the disease risk related to dietary stress, overweight, and obesity. We will discuss strategies that use nutrigenomics to answer nutrition problems.

**NUTRIGENOMICS: HOW TO GET SIMPLE ANSWERS FOR A COMPLEX SCIENCE**

An important challenge in nutrition research is the complexity and variability of nutrition and foods. The body has to handle a large number of different nutrients and other food components and nutrient concentrations can be high (micromolar to millimolar) without reaching toxic levels. Each nutrient can also have numerous targets with different affinities and specificities. This situation contrasts starkly with pharmacology, where single agents are used at low concentrations and act with a relatively high affinity and selectivity for a very limited number of biological targets. The challenge of nutrigenomics research is to break down the important but complex research problems into small feasible projects that can be handled by normal-sized research groups (Figure 3). One possibility is to switch from the most complex system, the human being, to simpler or more easily accessible organisms, such as yeast and *Caenorhabditis elegans*, which can serve as model systems. These organisms have sophisticated genetics as well as sequenced genomes, and researchers have made important discoveries in nutrigenomics using these model systems. Even the fruit fly *Drosophila* is an attractive model organism for conducting nutrigenomics research (2) because it has adipose-like tissues and a lipid transport system, which makes it a
closer model to humans with respect to obesity and associated diseases than many other model organisms. Another possible strategy is to use methods that are well established in medical or pharmacological research but are rather new to nutrition research. For example, analogous to pharmacology, nutrients or dietary metabolites can be viewed as signaling molecules that are recognized by specialized cellular-sensing mechanisms (8). The information that allows nutrients to activate specific signaling pathways is contained within their molecular structure. Minor changes in structure (eg, saturated vs unsaturated fatty acids or cholesterol vs plant sterols) can have a profound influence on which sensor pathways are activated. The challenge ahead is to further identify the molecular pathways that are influenced by individual nutrients and to determine the downstream effects of this regulation. Nutrigenomics can greatly assist in this identification because it allows the genome-wide characterization of nutritional target genes. With this type of information, researchers can comprehensively understand how nutrients act and explain how diet has such an impressive effect on health and disease, which is now widely acknowledged. Ultimately, nutrigenomics research will lead to development of evidence-based healthful food and lifestyle advice and dietary interventions for contemporary humans.

**DIETARY SIGNALS AND NUTRIENT SENSORS**

The main agents through which nutrients influence gene expression are transcription factors. Among the most important group of nutrient sensors is the nuclear receptor superfamily of transcription factors, with 48 members in the human genome. Numerous receptors in this superfamily bind nutrients and their metabolites (8,13). For example, nuclear receptors, such as peroxisome proliferator activator receptor-α (PPARα) (binding fatty acids) or liver X receptor α (binding cholesterol metabolites), bind as heterodimers together with retinoid X receptor to specific nucleotide sequences (response elements) in the promoter regions of a large number of genes. During ligand binding, nuclear receptors undergo a conformational change that results in coordinated dissociation of corepressors and recruitment of coactivator proteins to enable transcriptional activation. In metabolically active organs, such as the liver, intestine, and adipose tissue, these transcription factors act as nutrient sensors by changing the level of DNA transcription of specific genes in response to nutrient changes. Nuclear receptors have important roles in regulation of numerous processes, including nutrient metabolism, embryonic development, cell proliferation, and cellular differentiation.

Not surprisingly, nutrients, by activating these receptors, are able to influence a wide array of cellular functions. As an example, the PPAR group of nuclear receptors acts as nutrient sensors for fatty acids and influences expression of specific genes. One of the three PPAR isoforms, PPARα, is present primarily in the liver (14). The more than 3,000 to 4,000 target genes of PPARα are involved in numerous metabolic processes in the liver, including fatty acid oxidation, ketogenesis, gluconeogenesis, amino acid metabolism, cellular proliferation, and the acute-phase response (14). Hepatic PPARα is particularly important during fasting, when free fatty acids are released from adipose tissue. These fatty acids then travel to the liver, where they undergo partial or complete oxidation. However, these fatty acids also bind PPARδ, which then increases expression of a suite of genes through binding to specific sequences in their promoter regions. Fasted PPARα null mice (mice that lack functional PPARα) suffer from a variety of metabolic defects, which include hypoketonia, hyperthermia, elevated plasma-free fatty acid levels, and hypoglycemia (14,15). Recently, it has been demonstrated that PPARα directly regulates expression of genes involved in hepatic gluconeogenesis and glycerol metabolism (14,16). Because fatty acids are ligands for PPARα, the latter mechanism could explain the stimulatory effect of elevated plasma-free fatty acids on hepatic gluconeogenesis and glucose output. In addition to its important function in the physiological response to food deprivation or starvation, the role of PPARα in obesity is less clear, but most likely relevant to our understanding of the obesity-linked pathophysiology of type 2 diabetes (17). Visceral obesity is linked to increased free fatty acid levels (18) and, interestingly, these molecules may be recognized by the liver as “hunger” or “in need of glucose” signals, resulting in increased gluconeogenesis in a PPARα-dependent manner, particularly under conditions of hepatic insulin resistance.

In addition to the group of sensing transcription factors that directly interact with DNA by binding to specific response elements, the importance of corepressor and coactivator proteins became more evident because a substantial component of gene control is directed at the expression of coactivators. Coactivators exist in multiprotein complexes that dock on transcription factors and modify chromatin, allowing transcription to take place. Recent data on two coactivators of PPARγ, called peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1α and PGC-1β), are fascinating. PGC-1α has been associated with energy homeostasis, diabetes, and lifespan regulation (19). Polymorphisms in the genes encoding PGC-1α and PGC-1β have been associated with development of type 2 diabetes (20-22). Recently, high-fat feeding in a mouse model has been shown to induce hyperlipidemia and atherogenesis and to stimulate expression of PGC-1β in liver. Through molecular studies, the authors linked this mechanism to increased lipogenesis and very-low-density lipoprotein excretion because of enhancer effects of PGC-1β on gene transcription governed by the transcription factors sterol regulatory element binding protein 1 and liver X receptor α. From these results, a mechanism has been proposed by which dietary saturated and transfatty acids can stimulate hyperlipidemia and atherogenesis (23). This study demonstrates sensitivity of state-of-the-art microarray analysis and is an example of nutrigenomics research because it allows straightforward detection of profound changes in hepatic gene-expression patterns because of an adaptive response by the organism to changes in dietary macronutrient composition.

**SIGNATURES OF HEALTH AND DISEASE**

A major focus of nutrition research is on prevention of chronic diseases, such as cardiovascular disease, metabolic syndrome, and cancer. These disorders are partly mediated by chronic exposure to certain food components
and, therefore, a critical part of the prevention strategy concerns changing food habits. Causal relationships between those bioactive dietary components and prevention or outcome of a disease can only be assessed by long-term intervention trials, which are time-consuming and costly. Other conventional nutrition intervention studies use biomarkers like disturbed lipid profiles (e.g., cholesterol, triglycerides), increased blood pressure, or reduced insulin sensitivity as predictors of diseases, such as cardiovascular disease or metabolic syndrome. These biomarkers are mainly single proteins or metabolites or certain body functions that can be used as indicators for pathophysiological changes that can ultimately lead to a variety of chronic diseases, depending on the individual genotype. A complete biomarker profile will be more characteristic for the health status of an individual than single markers. The ability to assess these so-called “signatures” of health and disease hold the promise of a more complete phenotyping of humans and the ability to monitor health status using a noninvasive tool. Biomarker profiles can be determined on a genomics-transcriptome, proteome, and metabolome level, all having their own specific advantages as markers of body health status. Metabolomics (3,4,6,24-26) is the study of the sum total of endogenous and exogenous metabolites in a cell, organ, or body fluids and is a useful tool for generating individual metabolite profiles, such as complete plasma lipid (i.e., cholesterol, triglycerides) and vitamin profiles. This technique is quite innovative, and validation of accuracy and sensitivity, especially in the predisease state, has yet to be explored. Similarly, research on proteomics, studying the proteome (all proteins in a cell or tissue at a given time), is under development.

A nutrigenomics tool that is both sensitive and well validated is transcriptomics, which employs microarray analysis to study the number of messenger RNA copies per gene for virtually all actively transcribed genes. This technique enables determination of expression levels of thousands of genes at the same time and within one study. However, these types of studies require significant quantities of tissue material for isolation of the needed RNA. The accessibility of human tissues is limited and has been a major drawback to this type of analysis. However, biopsy samples from adipose tissue and muscle can be taken with relatively low risk and may offer one solution, with the limitation that gene expression can reflect a tissue-specific effect of a pathophysiological phenotype. For example, the proinflammatory response observed in adipose tissue of obese people is reflected by increased expression of proinflammatory genes in this tissue (27). Another less-invasive possibility is to isolate RNA from white blood cells, which travel throughout the body and respond to internal and external threats to body homeostasis. Gene-expression patterns in peripheral blood cells have been shown to be specific for disease states. Disease-specific gene-expression patterns in blood cells have been identified for breast tumors (28) and leukemia (29), and those patterns now serve as biomarkers for the disease and as the basis for useful diagnostic tools.

However, the question remains as to whether a predisease or nutritionally specific state can be distinguished in blood cells. So far, studies focusing on identification of these profiles are lacking. However, studies on inter- and intraindividual variations in blood cell gene expression in healthy volunteers have already been performed. The interindividual variation in gene expression is important among healthy individuals, but less distinct than the variability observed in the disease state (30,31). Expression patterns within a healthy person over a 24-hour period and at different intervals over a 6-month period were highly concordant, but variation in gene expression between persons varied significantly (31,32). This large interindividual variation makes it challenging to distinguish gene-expression signatures of a healthy subject from a prediseased subject. Once those predisease signatures (3) are defined and validated, nutrition interventions that focus on those predisease biomarkers can be developed, with the goal of regaining healthful expression patterns and subsequent improved physical conditions. Hopefully, in the near future it will be possible to increase the power of such genomics-based tests by combining multiple genomic approaches, such as uniting transcriptome profiling with metabolome profiling in a systems biology approach.

A limiting factor right now to the applicability of gene-expression signature analysis is the quantity of human blood or tissue material that is available. Generally, this approach requires RNA amplification procedures before whole genome microarray expression analysis can be performed. Another limiting factor is the heterogeneity of the blood cell population in terms of specific metabolic response and gene-expression profiles (31). The proportional presence of the cell subpopulation is subjected to internal and external triggers like infection or physical activity, thereby affecting the total gene-expression pattern. Finally, the method of isolation of blood cells and subsequent cell-type selection further influences final signature outcome (30,31). Nevertheless, despite all technological problems that must first be overcome, nutrigenomics-based signature analysis is a promising strategy for learning more about phenotypic responses to a nutrition intervention.

FROM GENOTYPE TO PHENOTYPE

Most chronic diseases, such as cardiovascular disease, metabolic syndrome, and cancer are multifactorial disorders caused by multiple genetic and environmental factors. Multigenic or polygenic diseases are caused by a combination of genetic variations in multiple susceptibility genes. Different combinations of gene variants can lead to a similar disease phenotype, further complicating the picture. Moreover, several modulator effects of dietary components on the phenotype by a genetic variation have been described and are referred to as gene-nutrient effects (33). Investigation of combinations of genetic variants and the effect of nutrients in relation to a disease requires large study populations of patients and controls. The known gene variants and the modulator effects of nutrients in relation to a disease have recently been discussed in an excellent review by Ordovas and Corella (33). Most genetic variations, such as single-nucleotide polymorphisms (SNP), insertions, or repeats have been found by sequencing genes coding for enzymes or transporters related to the disease of interest. With sequencing of the whole human genome, knowledge about genetic variations has increased, and microarrays containing
around 500,000 SNPs (500 K-arrays) are already available. These chips can be used in association studies to identify new candidate loci for a disease (34). Nutrition research has until now focused mainly on no more than a few SNP simultaneously and on the opportunity to abolish possible harmful consequences of the SNP by nutrition intervention. Research into these so-called nutrient–gene interactions with genome-wide SNP arrays are complicated by multiple genes, dietary components, and gene-nutrient interactions (35). Before technologies such as this can be applied in nutrition research, they first have to be proven and validated in disease-association research studies.

**CAN A BIOINFORMATICIAN MAKE SENSE OF YOUR MICROARRAY EXPERIMENT?**

Another limitation in nutrition research is the relatively small effect of dietary interventions on physiological parameters. Similarly, effects of nutrition on gene-expression patterns are also hard to detect. Consequently, a suitable study approach is required along with the need to develop state-of-the-art sensitive microarray analysis systems. Although a major technological advance, microarray technology has strengths and limitations that must be factored into the research design. A robust research hypothesis and study design will help to ensure the research question is adequately addressed by the experimental design (avoiding the so-called “fishing experiments”).

Important factors that can influence gene expression in humans as well as in animals are age, sex, nutritional status (30,36,37), and other possibly unknown (patho-)physiological parameters. In addition, these methods work best when genotype variation is minimal, which is feasible in animal studies with inbred strains but not with human beings. The most easily obtained information in human intervention trials is race and family history. Once a transcriptomics study is performed, RNA quality and quantity should be verified and subsequent labeling and hybridization of RNA should preferably be conducted by the same technician and within the same microarray experiment. In order to enhance accuracy, pooling of samples is not desirable, but increasing the number of biological replicates will decrease the false-positive rate and result in reliable data (38). Another area requiring attention is the preprocessing of data. Because dietary studies result in small gene-expression changes, they require a highly reliable algorithm (11). The most challenging part of the process occurs after the array data are obtained, when the data must be interpreted in terms of its biological meaning. Several commercial and noncommercial tools have been developed that assist with extraction of markedly changed genes and visualization of changed pathways or related networks (39).

**FUTURE OF NUTRIGENOMICS: TRANSCRIPTOMICS, PROTEOMICS, METABOLOMICS, OR SYSTEMS BIOLOGY?**

Different genomic technologies, namely transcriptomics, proteomics, and metabolomics, are complementary in the types of information they generate, but are at different points in their development at this time. Ultimately, parallel use of these methods will allow us to describe the phenotype of a biological system, such as a human being, in all its complexity, which is the major goal of nutritional systems biology (3,40,41). Without a doubt there is no “one and only” technology to solve all our research questions, so one has to be very clear about the aim of a study and its possible limitations. Given the many complex activities of the human liver, transcriptome and proteome analysis would be desirable, but tissue samples of human liver from healthy individuals are not readily available. Alternatively, researchers may be able to use plasma profiling of metabolites that might specifically serve as biomarkers for liver health or dysfunction.

Transcriptomics is a relatively mature technology compared with other “-omics” technologies. At this point, it is possible to get an overview of the expression of virtually all genes in a single microarray experiment, but it is not yet possible to measure the whole proteome or metabolome. However, research in proteomics is progressing rapidly (6,42). Studies of protein structure, expression level, cellular localization, biochemical activity, protein–protein interactions, and cellular roles are underway and considerable progress in novel instrumentation, experimental strategies, and bioinformatics methods has been achieved. Research progress in plasma proteomics is of particular interest to nutrition and nutrigenomics research because, if successful, a wealth of information could be generated about important proteins, such as various cytokines or hormone levels, from a small plasma sample. In order to achieve this, the major plasma proteins have to be separated and remaining plasma proteins must be recovered in a quantitative way for further analysis. Recent progress in this area is promising (42-45) and suggests that proteome-derived biomarkers useful in determining nutrition status may be identified before too long.

Metabolomics (3,4,6,24-26) is also in the early stages of development. It is not known how many endogenous metabolites exist or how many exogenous food-derived metabolites can be measured in human samples (urine, plasma). Scientists must first overcome a number of hurdles, such as full recovery of all metabolites from body fluids or tissue samples and the need to develop extensive databases with the required information about the nutritionally relevant metabolome. Metabolomics produces enormous amounts of data that require sophisticated instrumentation and software to allow researchers to extract meaningful information from the data. Existing instrumentation is quite sophisticated; the present limitations appear to be with the software needed to handle metabolomic data. The potential for nutritional applications of metabolomics is considerable, and a number of research teams are addressing these limitations (26,46).

**FUTURE PERSPECTIVES**

Will nutrigenomics stay exciting enough over the next several years to sustain development of an extensive research foundation? We are sure this will be the case because it is widely appreciated that further developments in nutrition and food development are impossible without exploring the mechanisms underlying nutrition. Will it then be possible from nutrigenomics research to
The work of the authors is supported by the Dutch Scientific Organization (NWO), the Dutch Diabetes Foundation, the Wageningen Centre of Food Sciences (WCFS), the National Genomics Initiative (BSIK program), the Dutch Dairy Foundation for Nutrition and Health, the National Genomics Initiative (BSIK program), the Wageningen Centre of Food Sciences (WCFS), and in the near future? In the coming years, we have to put all our efforts into gaining a thorough understanding of how nutrients interact with the human genome at a molecular level. To be able to use genetic blueprints or genotypes in dietary prevention of disease, we must first identify the mechanisms driving the connection between diet and the outward manifestation of our genes, our phenotype.

References


