INTRODUCTION

Nutrigenomics: Exploiting Systems Biology in the Nutrition and Health Arenas

Ben van Ommen, PhD
From TNO Nutrition and Food Research, Zeist, The Netherlands

NUTRITION DISCOVERS GENOMICS

Many diseases and disorders are related to suboptimal nutrition in terms of deficits of essential nutrients, imbalance of macronutrients, or even toxic concentrations of certain food compounds. In their classic approach, nutrition scientists have dealt with this relation by studying the interaction of food and nutrition in human intervention studies and using biomarker approaches to determine the effect. Biochemical and molecular knowledge and technologies have gradually been integrated in explaining the observations made in these human studies and in underpinning postulated mechanisms by in vitro and animal research. On the other end, the biomedical research arena has unraveled a good number of molecular “disease mechanisms.” Currently, the two disciplines are well on their way to closely interact.1 Thus, we realize more and more that the nutrition and health relationship is solidly anchored in interactions on the levels of DNA, RNA, protein, and metabolites (Figure 1).

Now that the complete human genome sequence has been unraveled, knowledge of the function of all individual human genes and their interaction is rapidly increasing. Technologies are being developed that allow the simultaneous determination of the expression of many thousands of genes at the mRNA (transcriptomics) and protein (proteomics) levels. Current DNA microarray technology allows the simultaneous expression analyses of almost the complete human genome. Proteome analysis is following several tracks in its attempts to characterize the complete set of proteins of a tissue, such as the classic two-dimensional gel electrophoresis, various LC-MS applications, and antibody arrays.2 Also, the analytical power of separating and identifying low-molecular-weight compounds is rapidly increasing and applied in nutrition studies as “metabolomics.”3–5 Although the methods to deal with this overwhelming amount of data and information are still in their infancy, initial examples of application of these technologies in nutritional sciences have been published.6,7

Usually, these technologies are applied in a “differential display” mode, i.e., by comparing two situations (e.g., diseased versus healthy, treated versus untreated). In this way, the complexity in data is drastically reduced by examining only differences. This results in the identification of new receptors, possible biomarkers, etc., and great expectations exist regarding this approach.6 Of course, the major advantage of this approach is the (relatively) open detection system paving the way for new mechanistic discoveries. This issue of Nutrition contains quite a number of examples of these applications.

The abundance of data allows not only the identification of individual genes, proteins, and metabolites that are differently present in the samples but also the grouping of the observed changes into functionally or mechanistically related blocks. Indeed, software is being developed that visualizes the gene expression changes according to biochemical pathways (see, e.g., www.genmapp.org). This allows for an integrated biological interpretation of the observed changes, thereby strengthening the pure statistical analysis or the results (Figure 2).

SYSTEMS BIOLOGY MAKES FULL USE OF NUTRIGENOMICS DATA

We begin to realize that, by using these differential display methods, the vast majority of the collected data are not exploited. Multiple minor changes remain unobserved, because only the eye-catching differences are elaborated. This straightforward trend is stimulated by the lack of adequate statistical tools able to cope with these new types of data sets, allowing the researcher to judge which of the gene expression changes are really significant. Can it be that a treasure of information is still hidden in the outcomes of transcriptomics, proteomics and metabolomics studies, waiting to be further investigated? A new way of dealing with these data is currently taking shape, with the aim of making optimal use of all available information and thus describing “complete” biological processes. This new approach is called “systems biology.”8–10

Many commercial and academic initiatives have been launched to exploit this area. Major progress is envisioned through a systematic inventory of all relevant parameters by using genomic technologies and application of new bioinformatics tools together with extensive data warehousing to unravel mechanisms and define biomarker sets (Figure 3).

SYSTEMS BIOLOGY WAS MADE FOR NUTRITION

The point is, nutrition is not like pharmacology or toxicology, where major effects can be observed, because the xenobiotic was designed to act on a single receptor with high affinity and strong effects, or where dose-related pathologic effects are induced with related strong effects on transcriptomic changes. Our diet consists of complex mixtures of many possibly bioactive chemical compounds, chronically administered in different compositions, and with a multitude of biological effects. The vast majority of these biological responses are mediated through effector genes, effects on enzyme concentration or activity, and changes in metabolite concentration (Figure 1). Transcriptomics, proteomics, and metabolomics will gain in sensitivity not only because classical detection limits are lowered, but much more because multiple minor changes taken together in new bioinformatics approaches create a new sense of sensitivity. Multivariate statistical methods will become of major importance. Next to cluster analysis, in which the effects of single compounds or mixtures on individual gene classes can be studied, tools such as principal component analysis are ever more being applied to study effects on the complete transcriptome or metabolome.

BIOMARKERS OF EARLY EFFECT

Many chronic “old-age” diseases and disorders are related to nutrition in the sense of prevention or of promotion. In the nutri-
tion research that focuses on this relation, human intervention studies are performed by applying biomarkers to determine the effect of the nutritional intervention. A major dilemma arises in this type of study. Nutrition is intended to be involved in the very early stages of the (prevention of the) onset of the disease, but we hardly have biomarkers that are accurate, specific, and sensitive enough to determine effects before the early onset of the pathology. Thus, these studies are compromised by selecting patient populations, where in fact the effect of nutrition on the disease state is being determined (therapy instead of prevention). Alternatively, very costly longitudinal studies need to be performed, where large cohorts of healthy volunteers are being followed with nutritional intervention into the disease state.

Here, the need for a new concept of biomarkers becomes obvious. We would like to study the effect of nutrition in the healthy state and measure very early effects that predict the

FIG. 1. Health effects of food compounds are related mostly to specific interactions on a molecular level. SNP, single nucleotide polymorphism.

FIG. 2. Schematic presentation of the pathways involved in apoptosis, with gene expression ratios alongside each gene. The ratios are derived from a 24-h exposure comparison of 20 μM of eicosapentaenoic acid with 20 μM of linolenic acid in a cell culture system (colonic Caco-2 cells). Gene expression was measured with an Agilent oligonucleotide array containing 17,000 human genes. The pathway presentation was made by Genmapp (www.genmapp.org). TNF, tumor necrosis factor; TNFR1, tumor necrosis factor receptor-1.
chronic effect in terms of prevention or promotion of a disease. Dietary therapy of atherosclerosis with antioxidant vitamins may be less effective than prevention of chronic metabolic stress through optimal nutrition.

Hence, the concept of the nutrition and health link is fully appreciated only if uncoupled from a biomedical “therapy-like” approach and linked to the awareness that multiple minor changes in metabolism and its biochemical regulation contribute to the onset of chronic nutrition-related disorders such as obesity, type 2 diabetes, cardiovascular disorders, osteoporosis, and chronic inflammatory syndromes. In other words, in this area, maintaining optimal metabolism is of key importance for improving health and preventing diseases.

**DEScribing Homeostasis Is the Key**

The application of “genomics” technologies in nutrition studies may provide the key for this dilemma. In combining many subtle changes into new biomarkers, the biomarker becomes much more sensitive and as a consequence allows for very early detection of changes. This has enormous potential for nutritional research. Biomarkers will change from describing a disease state (e.g., plaque formation in atherosclerosis) or negative effect (e.g., oxidative DNA damage) into describing subtle changes in health in positive and negative ways. In other words, by exploiting “holistic” data sets, healthy homeostasis and related early changes can be carefully described. This means that, in exploiting nutrition in disease prevention and health stimulation, nutritional science no longer depends on the (often irreversible) disease end point, but can use normal physiologic conditions as dynamic and reversible situations to work with.

Mathematical tools to deal with the complexity of the large data sets from transcriptomics or metabolomics experiments start to become applicable. Many multivariate statistical applications based on principal component analysis become routine in this type of work (Figure 4 shows an application of principal component analysis in transcriptomic analysis). Here, the relative contribution of all parameters (e.g., gene expression values) in the difference between samples is calculated and presented. Differences between samples thus are the result of all individual differences. Consequently, multiple minor differences, which (taken one by one) would not have any significance, together allow for discrimination between samples. The availability of such bioinformatics tools has a major impact on nutrigenomics, because nutrition in general does not provoke major changes in gene expression, but rather induces multiple minor changes. Our daily food may well contain hundreds of different bioactive compounds, each with its own profile of action.
bioactivity, retraceable in gene expression, protein concentrations, and metabolite profiles. These described profiles or fingerprints are a start and have already shown to be very sensitive. But this exercise has to be continued. A major objective of nutritional systems biology will be to describe physiologic homeostasis at cellular and organ levels. If homeostasis is defined as the dynamic equilibrium between all relevant metabolites, with underpinning gene expressions and protein activities, the “omics” technologies with their related multivariate statistical bioinformatics will give the nutritional scientists an extremely powerful tool to describe this homeostasis. Homeostasis will describe the healthy system, and perturbations, based on a multitude of subtly changing parameters (which, if taken apart, have no statistical relevance), can be traced and used as fingerprint biomarkers of prevention.

Of course, this is easier said than done. The concept of using description of homeostasis for biomarker purposes is valid, but many hurdles need to be surmounted. In human plasma, intra-day variation of many compounds, even independent of nutrition status, is very large. Effects of age, sex, environment, genetic differences, etc., need to be taken into account, in addition to nutritional variation and development of (pre-)disease stages. Statistical methods for describing these changes are largely lacking, and insight into longitudinal behavior on a metabolome-wide scale has never been investigated. Thus, before these concepts can maturate, a huge pile of work will need to be done. The good thing is that the discipline of systems biology, although especially fit for multidimensional problems and approaches in nutritional science, is also discovered within other disciplines, and a number of generic tools and methodologies will be developed in common.

**METABOLOMICS IN NUTRITION RESEARCH**

Metabolomics entails the complete and quantitative assessment of the metabolome, the set of metabolites that makes up the low-molecular-weight fraction of cells, tissues, or body fluids. Technological advances in NMR and various mass spectrometry applications indeed allow for a quite accurate assessment of large portions of the metabolome. This emerging technology of metabolomics seems to be well placed to become of major importance in nutrition research, for a number of reasons. First, many metabolites are part of our nutrition or metabolites derived from food compounds. Second, the metabolome can be regarded as the functional readout of transcriptomic and proteomic changes. Third, different body fluids are readily available from human studies in nutrition research (in contrast to human tissues that need to serve as a source for mRNA). Fourth, many “disease targets” for nutrition research are directly related to the metabolome (e.g., chronic metabolic stress, syndrome X, diabetes, obesity, cardiovascular diseases, inflammation-related diseases, and osteoporosis). The biomarker concept as described above can be exploited with metabolomics, where fingerprints of patterns of many metabolites together are indicative for early changes in physiology or onset of a pathology (Figure 5 shows an example). Progress is being made in this area, with initial published examples being available. Major breakthrough of nutritional metabolomics and its incorporation into nutritional systems biology are expected with the advancements in data handling, especially in the area of data preprocessing of the complex spectra derived from LC-MS and GC-MS applications.

**THE ADDED COMPLEXITY OF NUTRIGENETICS**

Of course, not all individuals react identically to nutrition. If nutrigenomics describes changes in gene expression related to a specific nutritional intervention, deviations in genes will have an impact on these transcriptome changes and ultimately on the physiologic function. On average, each of our genes contains 10 deviations in its code from the “standard gene.” Of course, not all of these polymorphisms have a functional impact. A relatively small number of these polymorphisms has serious health implications and may even be lethal. This is the domain of clinical genetics. Many polymorphisms, however, have only a mild effect on the functionality of the resulting protein. It is here that, within certain limits of “health,” a large variety in response to nutrition is observed. For example, the plasma cholesterol concentration is only partly determined by the cholesterol intake through nutrition. Extremely high concentrations are observed related to specific gene mutations. These persons run a high risk of cardiovascular problems and undergo drug therapy to lower plasma cholesterol. Apart from these extremes, a large variation in plasma cholesterol concentration exists, with an underpinning of known and unknown genetic polymorphisms. Here, the possible interplay between a number of polymorphisms may be, at least in part, accountable for the variations between the boundaries of accepted plasma cholesterol concentrations. In studying the effect of cholesterol-lowering nutritional intervention, it may be wise to take these subpopulations into account. Numerous other examples of interindividual genetic differences related to nutrition and health are known. Many single nucleotide polymorphisms resulting from inheritance or as spontaneous mutations are involved in the type 2 diabetes phenotype. Inflammation-related single nucleotide polymorphisms, like the mutations in the interleukin genes, can also be influenced by nutrition (see other papers in this issue).
Paradoxically, food itself may contribute to this diversity, because there are many examples in which nutritional compounds directly cause DNA damage or modulate susceptibility (in the positive and negative sense) against DNA damage through regulation of specific pathways involved in the many processes involved in these events.

PERSONALIZED NUTRITION?

Based on these interindividual differences, it is tempting to speculate on “personalized nutrition” based on genotyping differences. Of course, without stressing the genetic background of variation of nutritional response, specific subgroups are already targeted with “subgroup nutrition” (e.g., cholesterol-lowering margarines). A debate is arising as to whether nutrition should enter into the area of linking genetic differences with tailor-made nutrition. Apart from the social, ethical, and communication issues involved, from a scientific point of view, a big challenge is ahead of us in validating the combined action of these minor-impact polymorphisms and their practical effect on the relation between nutrition and health (Figure 6).

COMBINED SAFETY AND EFFICACY EVALUATION

Risk-benefit analysis aims at defining optimal intakes of bioactive food components. To establish maximum benefit and minimal risk, effects of bioactive food components need to be assessed in the context of the diet, at relevant doses of intake, and as they occur over long periods in the entire human body. Risk-benefit evaluation needs to aim at optimization of nutrition and analysis at relevant physiologic doses of intake (rather than extrapolation from non-relevant high-dose animal toxicity data). Thus far, however, adequate tools are lacking, and a classic toxicologic approach usually has been followed in assessing potentially harmful effects of food compounds. The nutritional systems biology concept described above (functional genomics applied to integrated analysis of biological systems) can be used for assessment of safety and efficacy in the relevant physiologic human context, including interindividual variation. The multiparameter biomarkers will be sufficiently sensitive to be used at physiologic conditions and will be able to identify subtle changes from the homeostasis. Thus, in developing the area of nutrigenomics, new methods of safety evaluation for food compounds may be implemented.

CONCLUSION

Although relatively new technologies, the various genomics applications searching for new receptors and pathways already have found their way to many nutritional applications. Moreover, the new science of nutritional systems biology is emerging, taking up the challenge of exploiting all available data generated by genomics technology in a complete description of a biological system. As a consequence, this new paradigm is ideally fit for the evaluation of many subtle changes in biological activity as triggered by nutrition. In this case, a multitude of bioactive compounds acts simultaneously and chronically in constantly changing combinations.

Having said this, we realize that tools on the level of data handling and evaluation are mostly absent. New bioinformatics will be necessary to make this dream come true. Fortunately, the road to nutritional systems biology is full of applications that can be used now during this passage. Indeed, it is through emerging examples in the field of nutrigenomics that we begin to clearly see the road we need to take.

REFERENCES


doi:10.1016/j.nut.2003.09.003