Nutrition and Genomics

Bruce German\textsuperscript{a,b} and Vernon R. Young\textsuperscript{c}

\textsuperscript{a}Nestlé Research Center, Lausanne, Switzerland; \textsuperscript{b}University of California, Davis, Calif., and \textsuperscript{c}Massachusetts Institute of Technology, Cambridge, Mass., USA

Introduction

The decoding of the human genome is already being heralded as one of science's greatest achievements and is viewed as the point of departure for the next generation of life science research to understand human health and cure human disease [1]. However, genomics as a scientific endeavor is not simply the list of nucleotides in the genomes of humans and various other organisms. This new field is bringing a new knowledge base on genes, their functions and regulation. This is promoted by new, multiple and truly remarkable biomolecule-based analytical technologies that are capable of interrogating the functioning of living organisms with a breadth and depth not previously imagined. Additionally, computer-based technologies are spawning an entirely new computation discipline directed towards the mining of biological databases, termed bioinformatics.

The various scientific disciplines that seek to develop cures for human diseases are already fully engaged in genomic research. Nutrition, with its goals to improve human health and prevent disease, stands to achieve tremendous advances during this post-genome era. Nutrition enjoyed conspicuous success during the first half of the 20th century, establishing all the major essential nutrients – the vitamins, minerals, amino acids and fatty acids – that are necessary for growth, development and reproduction. The diseases that are associated with a deficiency of each of these nutrients and the quantities of each nutrient necessary to prevent them have been reasonably well described using a variety of animal models and experimental strategies. As Carpenter [2] pointed out in his fascinating series on a short history of nutritional science, many have looked on the developments in the mid 20th century as the ‘golden age of nutrition’, but during the dawn of this 21st century nutrition science now faces a new set of challenges for which the knowledge, tools and strategies of the ‘omics’ will be indispensable. The challenges are not small
and are summarized in figure 1. Metabolic diseases, such as atherosclerosis, hypertension, osteoporosis, obesity and diabetes, are increasing throughout both the developed and developing worlds, dramatizing our poor understanding of how diet affects metabolism and health and/or our ability to exploit current knowledge. Although these diseases may not be as directly associated with specific dietary problems as were specific nutrient deficiencies, individual choices of diets are clearly an important part of their etiology. Additionally, the science and practice of nutrition will have to face, simultaneously, the problems of understanding how diet affects metabolic regulation and how individual genetic predispositions mitigate dietary effects which serve as the basis for considerable inter-individual variation in the risk of these diseases within the population.

**Genomics, Proteomics and Metabolomics**

Biological scientists are rapidly assembling a knowledge base on genomes and beginning to translate the perspective of genomics to all of life’s molecules. Organisms and their molecular constituents can be organized into one of three basic classes: the genetic sequence or genome that contains all the information (the genes and their regulation) to produce an organism; the proteome, i.e., all the proteins that provide the structures and catalysts that form an organism, and the metabolome, the sum of all molecules in the metabolic pathways that represent the actual, real-time functioning of an organism. This hierarchy is diagrammed in figure 2.

Genomics is the field of science devoted to sequencing and understanding the entire genomes of organisms [1]. Genomics is not just the listing of
sequences within DNA but it includes new scientific strategies and tools. Genomics has the goal to annotate the entire genetic sequence of an organism, identifying each and every gene, understanding its function and its regulation. To accomplish such a goal genomics has created new analytical tools based on the biomolecular properties of genetic material DNA and RNA. One set of tools is capable of analyzing the expression of dozens of genes or even all genes from a particular genome. Another set of tools is capable of analyzing specific differences in sequence (polymorphic mutations) at thousands of different gene locations within the genome of individuals within a species. Finally, the massive quantities of data that genomics produces has led to the

---

**Fig. 2.** The biological hierarchy of global analysis. The genome information content is large (3 billion bases for humans) not just because it contains the information of genes and pathways but within the coding and noncoding regions of the genome exist information on gene regulation, directed evolution and the information by which biopolymers self-assemble into intake cells and entire organisms. The transcription of genes in any cell, i.e. the transcriptome, represents a much smaller information space than the genome, the total number of genes of humans estimated to be no more than 40,000. The translation of genes into proteins leads to the proteome whose information space is larger than the transcriptome since one gene can encode many different protein structures with post-translational modifications, three-dimensional structures and varying functions. Once in place, the proteome dictates and directs the conversions of substrates, fuels, precursors and products that constitute the molecules of metabolism. The information space of the metabolome is in turn much smaller than the proteome, there are far fewer metabolites than proteins and the three-dimensional structure information content of these molecules is much less.
Nutrition and Genomics

development of computational tools that are necessary to access, interrogate and ultimately understand the biological structures, regulation and functions of the genome. This development has given rise to the relatively new discipline of bioinformatics.

Although deducing all of the functions of all human genes is still a long way into the future, many important breakthroughs have given a glimpse of the power of genome-wide studies [3, 4]. As a result, more and more scientists are involved in research to understand how the various parts of the genome functions, a process that is termed genome annotation, or functional genomics. This annotation process yields the information that biologists will need in order to understand the complexities of whole organisms and this process is just beginning. Annotation of the human genome will keep the entire field of life sciences busy for generations. There remains much basic science to be conducted to annotate even the simpler organisms, though much of the information will be relevant to humans [5]. One of the genomics revolution’s great assets is its reliance on computer-based databases that facilitate computerized comparisons between organisms [6]. This comparison approach will continue to accelerate with a rapidly growing library of completed genomes, including humans and chimpanzees; other mammals including mice and rats, and even simple organisms, including the fruit fly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*, yeast, dozens of bacteria and hundreds of viruses. Further, with the increasing technical capacity for nucleotide sequencing of DNA it is becoming possible to address the study of not only single organisms, but of entire ecosystems (ecogenomics) and sequencing all organisms within a defined ecological niche, such as the human intestine.

**Tools of Genomics**

*Gene Expression Analysis*

Gene expression analysis is a remarkable tool of genomics, providing a rapid means of access to many of the processes of interest to biologists. Complementary sequence probes for virtually all the expressible genes within an organism, immobilized as a two-dimensional array, created a true conceptual breakthrough in molecular biology. Now, biological scientists can apply commercial DNA arrays to a tissue sample to determine the expression level of virtually all genes within the genome of an organism. This analytical power has led to several new experimental designs. The most straightforward of these new genomic designs is an unbiased comparison experiment (i.e., comparison of a treatment with control, or one induced condition with another, or a diseased tissue with a healthy tissue). The technique provides the investigator with quantitative information on all of the genes that are upregulated, all the genes that are downregulated and genes that are unaffected
by a treatment or pathophysiological state [7]. The sum of all genes expressed by a cell or tissue at any one time obtained by DNA expression arrays is typically referred to as a ‘transcriptome’. This reflects the fact that what is actually measured are all the mRNAs for the genes that are being transcribed (transcriptionally active).

Gene expression analysis (transcriptional profiling) [8] is already proving to be a valuable tool in nutrition science and food research, leading to the discovery of (i) the genes, and the mechanisms, by which essential and non-essential nutrients function; (ii) new mechanisms and functions of nutrients, as illustrated by recent work with arginine [9, 10] and selenium [11]; (iii) unanticipated genes whose expressions are altered by nutrients and food ingredients that may help to reveal possible side effects; (iv) differences between individuals, stages in growth and development and environmental inputs that alter the needs for and responses to nutrients/diets; (v) the fundamental consequences of clearly inadequate intakes of specific nutrients or of more subtle nutritional imbalances, deficiencies and/or excesses; (vi) the pathways by which plants, animals and microorganisms produce edible materials including, but not limited to, the essential nutrients and other chemicals that may have positive or negative effects on humans or animals of economic significance; (vii) the evolutionary origins of nutritional needs and human responses to diet, and (vii) the evolutionary emergence of nourishing substances in biofluids such as milk [12].

**Genetic Variability and Polymorphisms**

The phenotypic variations between species and between individuals within a species are, in part, due to specific differences in the genetic sequence, both in the sequence of genes, and perhaps even more importantly, in the regulation of the expression of genes. Differences in the genetic sequence of individuals also underlie variations in nutrient needs and how people respond to varying nutrient intakes. Comparative analysis of genomes and the conduct of genomics research provide the means to discover the precise details that account for these differences and the tools to rapidly identify them. The promise of this knowledge is in being able to provide nutritionists with an understanding of the genetic basis of variations in the diet–health relationship within a population. In the not-too-distant future, individuals with known metabolic/disease predispositions will be provided, if they wish, with the biological basis to help design a diet according to their unique nutritional needs or a particular phenotype for a given lifestyle. Knowing that there are sequence differences between different humans would be practically useless if the differences could only be identified by fully sequencing each and every individual genome for a particular population. Fortunately, a tool set has emerged in which sequence differences as small as single point mutations
single nucleotide polymorphisms, SNPs) can be rapidly identified using technologies similar to DNA expression arrays. A sample of DNA from an individual can be applied to a procedure in which, instead of the sequence of the genome or the expression of genes, the variation in sequence of key genes is analyzed, using either array technologies or fluorescently tagged microbeads [13, 14]. This technology is limited in that only the sequence variations that are known and incorporated onto the analytical device can be screened. An international consortium is assembling all of the genetic variations that are important to health (the SNP Consortium; a public/private collaboration that, to date, has discovered and characterized nearly 1.8 million SNPs (http://snp.cshl.org1). At present, only the functional consequences of a few hundred sequence variations that occur commonly in humans are understood, even though thousands are capable of being routinely assessed by this technology [15]. Nevertheless, one of the most valuable aspects of contemporary science is that, because scientific information is increasingly electronically linked through worldwide networks, each new discovery of sequence variation and diet is added to a growing annotation of the human genome. The nutritional significance of these variations in the population represents a high priority for nutritional genomics research. Increasingly, the nutrition and food sciences will take advantage of a growing body of genetic information on the differences in humans that relate to the development of diseases and how they are influenced by diet [16].

**Bioinformatics**

The new field of bioinformatics utilizes a mixture of three separate toolsets – high-speed statistical computing, biological databases and shared electronic access via the Internet. Biologists in the next generation will be bioinformaticians simply because the datasets that are now being created are very large and the interrogative power of bioinformatics is so great. Nutrition is a science devoted to integrative metabolism and human health and it is clear that bioinformatic tools will be at the core of much of future nutritional research and its implementation. This does not mean that the tools to bring disparate biological processes together via bioinformatics will completely replace traditional, focused mechanistic research – quite the opposite. Because bioinformatics is an integrative toolset, it joins all of the past and ongoing scientific knowledge of nutrient mechanisms into a coordinated informatic system. Because bioinformatics is also an electronically networked toolset, scientists around the world share in the progress simultaneously. The nutrition science community must, therefore, immediately begin to educate its students in the applications of bioinformatics and to participate actively in the annotation of the human genome and downstream Omics (see below) for their nutritional properties. The importance in the training of nutrition scientists, including a
working knowledge of mathematics and computational techniques [17] and of developing the necessary infrastructure to permit the undertaking of exciting interdisciplinary research [18], cannot be overstated.

Bioinformatics is the engine of the genomics revolution. Everything that genomics is achieving is due, in no small measure, to the increase in computing power that began with the invention of the digital computer decades ago. Computing power continues to advance at a remarkable rate, verifying Moore’s [19] wildly optimistic prediction in 1964 that computing power would double every 16 months with no end in sight. Bioinformatics has produced the tools that have resulted in the assembly of the databases of genomics, including gene sequencing and gene expression arrays. With these databases in place, bioinformatics is building the tools that are necessary to mine these databases and derive knowledge from what would have been an incomprehensible mass of data just a few years ago. Bioinformatics is now evolving into a biological discovery process, relying more and more on the input and participation of biologists and particularly integrative biologists, as in the nutritional sciences [20].

**Proteomics**

Turning now to the second basic class of molecules, as noted above, the field of proteomics is operationally defined as the study of all proteins in a sample, cell, tissue or fluid [21, 22]. Since genes largely encode and produce their effects via the production of proteins (proteins are the workhorses of cells), the next logical comprehensive (‘omic’) dataset after transcriptomics is proteomics. While genes encode all proteins, the amount of mRNA present in a sample determined by a gene expression array experiment does not necessarily represent a quantitatively accurate picture of the amount and distribution of down-stream molecules, including proteins. Therefore, ideally proteins must be measured. However, the complexity of the proteome is far greater than that of a given transcriptome. Proteomics experimentally consists of measuring the proteins in a sample, establishing the covalent modifications of the proteins that differentiate them from their simple sequence prediction, annotating the functions of the proteins and establishing through bioinformatics the statistical and functional differences in protein patterns among samples [23].

Technologies are being developed to measure all proteins in a sample simultaneously on a single analytical platform, analogous to DNA arrays [24]. Once measured, all proteins in one sample can be compared quantitatively to all the proteins present in other samples. Although methods are being pursued to achieve this goal, challenges remain. First, no technology currently available is capable of separating and quantifying all proteins in a high-throughput analysis routinely used by non-sophisticated users.
Therefore, even the basic toolset of proteomics is restricted, at present, to specialists. Secondly, unlike genes, many proteins exist as specific three-dimensional structures in the biological milieu. Even the same polypeptides encoded by the same genes are modified to produce different functions [25]. No technology can yet solve this problem for proteomics in a high-throughput analysis.

An interesting new approach emerging as a subset of proteomics and likely to provide new breakthroughs in both health and food sciences, is the study of the interactions between proteins [21]. All of the proteins in a sample that are physically bound to or associated with each other are determined analytically. Once measured, all of the protein interactions that occur in the biological sample can be reassembled mathematically. The vision of this field is that all of the protein clusters that exist in vivo can be established so that scientists can construct the three-dimensional structures of entire cellular processes.

Considerable information is obtained from proteomic analyses that are valuable for the fields of nutrition and food science. Results from proteomic analysis reflect the physiological history of a tissue/organ/cell integrated over various time frames and constitute a novel form of biomarker system. Hence, there is some optimism about the possibility that proteomic analyses, once brought to high-throughput platforms, will provide decisive biomarkers not only of existing physiological states but of the presence and severity of damage or disease conditions. Proteomics has been used to establish the presence of unintended effects of genetic or environmental manipulations within plant and animals that are used as foods.

**Metabolomics**

Metabolomics is considered to be the study of the entire set of metabolites within a sample, cell, tissue or biofluid [26]. The measurement of metabolites captured as a ‘snap shot’ measurement of the flux through all metabolic pathways is a close reflection of various aspects of an individual’s phenotype at that time. From this consideration, measuring all metabolites within a representative biofluid from an organism (blood, urine, etc.) provides the potential for a quantitative index of an individual’s immediate health status. At present, no single technology is capable of measuring and identifying all metabolites in a sample simultaneously. To solve this problem several complementary approaches are being pursued. Nuclear magnetic resonance (NMR) spectroscopy can be used to detect the proton or carbon resonances of all metabolites. Presently these can be analyzed quantitatively, but not qualitatively, in that it is not possible to assign all proton resonances in urine or plasma to the metabolites that produce them. A second strategy uses mass spectrometry with all of the metabolites in a sample or biofluid being identified but not
accurately quantified. In a distinctly different overall approach, samples are first subdivided into chemical classes and each is assigned to discrete technologies to obtain accurate measurement of the identity and quantity of each and every metabolite. In this case separate, parallel analytical platforms are necessary to obtain the complete data of all metabolites in a sample.

NMR spectroscopy is proving to be an exceedingly powerful, non-destructive analytical technique to obtain metabolite signatures of entire biological samples [27]. Since all metabolites contribute proportionately to the NMR spectrum, the NMR data are a reflection of the metabolite concentrations. To date, however, all resonances in such complex spectra have not been assigned to specific metabolites. Therefore, researchers may treat the NMR resonance peaks as independent statistical variables, and high-capacity computer algorithms process these data via various clustering and multi-dimensional mathematical regression solutions. Samples from different individuals can be analyzed by NMR and the unassigned NMR peaks used as input data to clustering algorithms in order to establish whether physiological or health states define themselves as unique n-dimensional clusters [27]. Already these techniques have distinguished fingerprints in human urine associated with disease states, such as atherosclerosis [28].

Mass spectrometry has undergone continuous renovation over the past 30 years improving sensitivity, accuracy and dynamic mass range with the ability to assign accurate mass identities to thousands of molecules in a single experiment. Mass spectrometry has also been applied to the analysis of entire families or pathways of metabolites [29]. Mass spectrometry is not typically capable of highly quantitative measurements in these applications; however, the use of internal standards specifically enriched in stable isotopes brings the technique to quantitative accuracy. To date there is no repository of such stable isotopes and an important goal for the increased use of mass spectral techniques is to promote an increase in the availability of such internal standards. The use of stable isotopes as mass-specific tracers has even greater value due to the application of mass spectrometry techniques to the simultaneous measurement of metabolite concentrations and dynamics [30]. The highly accurate mass sensitivity of modern mass spectrometers makes it possible to follow the migration of stable isotopic enrichment through each metabolite in a known pathway with time. As a result of this technological and biochemical combination, nutrition experiments can now provide an accurate in vivo estimate of metabolic fluxes through these pathways in free-living humans. When such analyses are performed, it has been possible to compare the implications of gene expression profiles and proteomic analyses with the actual changes in metabolism that would be predicted from the changes in enzymes and the genes encoding them. Results from the studies conducted to this point have been illuminating. In many cases the gene expression results are mirrored by the predicted metabolic effects [31]. However in other cases, the results of gene expression analyses do not predict what actually
occurs in vivo. Perhaps a notable example concerns the results demonstrating that in response to a high-carbohydrate diet, genes associated with lipogenesis are increased predicting an increase in conversion of carbohydrate to fatty acids. However, actual metabolic flux measurements of these conversions reveal that carbohydrates are not quantitatively converted to fatty acids and alternatively, the flux of carbohydrate to oxidation increases in preference to oxidation of fatty acids [30].

As already alluded to, a complementary approach to the study of metabolomics is to separate a sample into various classes of metabolites and measure each class using an appropriate analytical platform (LC-MS, GC-MS, FT-MS, etc.). The entire set of metabolites can then be reassembled into a single dataset. The advantages of this approach are that each metabolite is identified and, ideally, accurately quantified. The quantitative metabolite data are, therefore, suitable for producing databases that provide unbiased biological information about a sample, and can be used to fit metabolic models and to assess the integrated flux through all metabolic pathways [32, 33]. No matter how the data included within metabolite databases are obtained, the objective is to produce quantitative and qualitative databases that are independent of the analytical platforms that produced them. These databases of metabolites from samples joined with the phenotypic characterization of the organisms from which they were obtained constitute a permanent biological resource to be compared, mined and fitted to pursue various hypotheses.

In our opinion, the ability to better predict the behavior of individuals in response to diets and nutrients will depend upon genomic/proteomic/metabolomic representations that include quantitative, kinetic and mass balance information on metabolic pathways, their control regulation and integration/interaction. This leads us to focus on the system now that we have considered, in brief, its various parts (genes, proteins, transcripts, metabolites).

**Integrating Nutrition: Systems Biology**

The various techniques noted above effectively provide a detailed parts list for the organism but do not necessarily inform us how these parts are assembled into a working organism. Systems biology is the term used to refer to the processes of integrating all the various tools of modern biology to address particular questions and ultimately to capture the entire biological representation of an organism [34, 35]. Genomics, proteomics and metabolomics are applied to a single experiment to generate data and compare regulation at all levels of biological regulation. In the first studies, the entire metabolic pathways of specific microorganisms have been addressed [36]. Recently, specific health issues in higher organisms, for example arthritis, have been studied [37].

The genomics revolution is beginning to allow us to address many new and many remaining challenges in nutrition, including (i) nonessential nutrients: the
effects of nonessential nutrients are still poorly understood and the classical strategies of eliminating a nutrient from the diet to observe the deficiency consequences are no longer appropriate. Genomics tools make it possible to examine the effects of these food constituents in ways not possible before and under conditions that are practically relevant. (ii) Personal variation: humans are not genetically identical and their variations lead to differences in nutrient requirements and the health effects of different diets. Health recommendations can only become more personalized when it is possible to identify the variations in humans and accommodate diets to the natural variation in human genetics and responses to diet, environment and aging. (iii) Diet and metabolic regulation: even when all the essential nutrients are provided in adequate quantities in the diet, imbalances in nutrient intakes can cause metabolic disregulation leading to such problems as atherosclerosis, obesity, diabetes, hypertension and osteoporosis. Understanding how to balance and optimize the intake of all nutrients requires the integrating power of genomics and similar approaches. This is a large undertaking but the results, eliminating metabolic diseases, will justify the necessary and major investment of intellect and other resources. (iv) Prevention: health care will become increasingly focused on the issues of maintaining or enhancing health and preventing future diseases rather than curing existing diseases. Such approaches will invariably require technologies that integrate all aspects of biological function.

**Personal Variation in Diet and Health**

Humans differ in many ways, in response to diet as a result of their genetics, their personal history and their immediate metabolic needs and status [15]. Personalizing health will need the tools to recognize the molecular basis of genetic variation among humans and the knowledge to act specifically to address this variation. Various examples of nutrient diseases that follow family histories and their genetic origins have been assembled, including phenylketonuria and lactose intolerance [38]. Beyond the diseases in which a clear genetic basis can be inferred from epidemiology, the variation in the incidence of most adult-onset degenerative diseases, such as cardiovascular disease, obesity, diabetes, gluten-induced enteropathy and Alzheimer’s disease, can be linked in part to genetic differences, sometimes measured in single-base pair substitutions, which SNP technologies can identify [39–42]. In studies conducted to date, specific polymorphisms have been associated with varying predispositions to diet-related diseases, variations in nutrient requirements and nutrient-specific metabolic responses. Knowledge of a genetic predisposition to benefits and to liabilities associated with consuming various diets does not necessarily require genotyping to identify those affected. Appropriate measurements of metabolism, as noted earlier, especially in response to
specific dietary intervention can provide considerable understanding of the variation in humans from which genetic variation can be inferred. It is worth pointing out, however, that in addition to the genotype, lifestyle factors have a quantitatively important role in the etiology of complex diseases and it is now quite important to determine the causal relations between lifestyle, genetic factors and disease risk [43].

**Metabolic Regulation**

Genetics is an important component of and contributor to variability that is routinely observed in human health, but it accounts for only a part of this variation [44]. An individual’s overall phenotype, including health status, is the integral of all processes, genetic and environmental. Nutrition and epidemiological research is discovering that diets adequate in all of the essential nutrients are still not optimal. Health is achieved and maintained when overall metabolism functions optimally within the lifestyle of each individual. Metabolic regulation, from genes to metabolites, dictates the biochemical functioning of an organism, including energy regulation and homeostasis, tissue structure rebuilding and repair, and the multiple systems of protection. Diet provides fuel and is important in regulating the storage, transport and delivery of fuel metabolites. Diet provides the structural precursors of biopolymers and regulates, in part, the rates of turnover of these structures. Diet also provides chemical protectants, modifies the processes of stress recognition and fuels the processes of stress response. Not surprisingly, the choice of foods and the overall diet consumed by an individual has a profound effect on these various processes. In the United States, a majority of adults and increasing numbers of children are choosing diets and lifestyles that do not support optimal metabolism; instead they contribute to problems such as diabetes, obesity, atherosclerosis, hypertension and osteoporosis [45–48]. Diet is thus one potential means, among others, to help solve these problems. However, the same diet will not necessarily solve the problem for equally all individuals. Each individual’s genetic predisposition and metabolic history play very important roles in determining what diet will be optimal for a given outcome.

The tools of genomics and systems biology now make it possible to measure ‘metabolic’ health. The lessons provided from decades of cholesterol research have shown that it is possible to identify metabolic pathways that lead to inappropriate physiological outcomes. The implementation of this biological knowledge, however, mandated the measurement of cholesterol within each individual to determine those affected and to intervene to shift metabolism towards a lower risk, i.e., greater health. However, while plasma cholesterol, as an isolated metabolite has been useful in predicting one aspect of risk, a ‘normal’ cholesterol does not always confirm a lack of risk of heart disease.
A focus on health means that all aspects of metabolism must be measured and overall risks optimized. This is the goal of metabolomics, to measure all of the smaller cellular constituents or metabolites in order to guide how best to intervene towards achieving optimal health in each individual.

**Prevention**

As public health moves forward in the genomics era, with continued accumulation of molecular knowledge to help solve the problems of infectious and genetic diseases, nutrition science is beginning to address the parallel goals of maintaining health and preventing the development of disease. These, unfortunately, will be much more difficult to deliver than to promise. Diseases that result after long-term chronic imbalances in metabolism are not due to deficiencies of essential nutrients, do not necessarily produce biomarkers of damage until the disease is well established and are not resolved by the simple provision of a limiting nutrient. Preventing diseases that are the result of chronic metabolic disregulation means maintaining normal metabolism, thus avoiding problems and the damage that occurs, before any abnormality or identifiable disease has developed.

Many of the health issues in today’s human populations relate to the fact that they are routinely exposed to a variety of stresses, bacterial pathogens, toxins and pollutants. These stresses, even when not resulting in disease, if prolonged and not satisfactorily managed and repaired, similarly lead to functional impairment and dysfunction and can lead to inappropriate responses, such as chronic inflammation. Maintaining health in the face of chronic stresses through the positive, protective properties of foods would require agents, compounds or structures that could intercept such stresses and prevent them from exerting a negative influence on our daily health, and/or coordinate the most appropriate response to stress.

The tools and knowledge of genomics are addressing these new aspects of nutritional health, in innovative ways. First and foremost, the basic principle of omics is to precisely measure as much as possible the approach necessary for prevention and health promotion. In essence, delivering on the promise of prevention means that one can document that a diet (or any other intervention, exercise, drugs, etc.) makes someone who is already healthy, healthier. Ostensibly, there is no existing disease by definition. This means that if a diet, food or food ingredient makes someone healthier by virtue of preventing a particular disease, it must do so without altering the risk of any other disease. Therefore, it is not simply sufficient to show that an intervention is effective in lowering the risk factor(s) of one disease; it is necessary to demonstrate that no other aspect of health is adversely affected. This means that, in principle, all aspects of health must be measured. In other words this requires application of the various omics. In practice, this means that assessment of
health must take an integrative approach to be able to measure metabolites in a holistic fashion. This is precisely the goal of metabolomics. Hence, to move forward towards genuine prevention strategies, omic technologies must be in place to measure those metabolic pathways that are changed by a particular intervention and to measure those metabolic pathways that are not. There are already examples of wide-ranging metabolic profiling studies that illustrate a holistic approach to health and prevention through metabolomic technologies [32, 49]. Fundamentally, we need to be able to better predict the behavior of individuals to a particular nutritional environment. As noted already, we believe that the way ahead is to depend upon genetic/cellular/systems representations that include quantitative, kinetic and mass balance information on metabolic pathways, their control, regulation and integration.

The goal of protecting individuals from various stresses has been a subject of intensive research for decades and is the underlying basis for an interest in and use of such food ingredients as antioxidants, probiotics and carotenoids [50]. This field has been challenging in part because of the nature of the problem. To study protection, it is necessary to be able to understand stresses in sufficient molecular detail that they can be treated as an independent variable, and only then can one devise scientific studies to understand the best way to protect an individual from each particular stress. Research on stresses has shown that they are variable in time, place and circumstance. Stresses are not constant, and in many cases are not even predictable. Stresses are localized and do not affect all tissues similarly. Stresses are dependent on the activity of the stressed organism. A vivid example of the role of stress and the complexity of protection is the health of the intestine, an obvious target of nutritional health. The intestine is exposed to a variety of stresses, including toxins, oxidants, pathogens, allergens and infectious organisms. The severity of the stress exposed to the intestine is dependent on the matrix in which the stress is present, on the health of the intestinal cells, their mucin layers and the health and activity of the intestinal microflora (the mixture of hundreds of different species of bacteria within the lumen of the intestine). Genomics, proteomics and metabolomics are all being applied to an understanding of the problems and opportunities of intestinal health and highly creative solutions are now emerging. With the sequencing of the human genome, scientists are actively studying the entire intestinal tract for the genes expressed both spatially and temporally and in response to various disease states [for review see, 51]. Sequencing the genomes of all the bacteria, both pathogens and commensal bacteria, that inhabit the human intestine is an active area of research [52–54]. The tools of molecular biology are being used now to interrogate these genomes for the specific genes responsible for beneficial and detrimental properties [55]. Finally, the nutrients that selectively encourage the growth of beneficial bacteria, termed probiotics, are being studied from both the perspective of their molecular structure and the perspective of the evolution of the genes of the enzymes that produce them.
to discover how to enhance this complex but potentially highly beneficial aspect of the diet [for review see, 56, 57].

**Applying the Genomics Revolution to Nutrition and Health**

The key to the future of nutrition science and its effective application lies in completing our understanding of the relationships between diet and health, with the goal being to improve health and prevent disease rather than to diagnose and reverse existing disease. The knowledge emerging from the initial results of the genetic revolution is clear; there will be no revolutionary, single genomic principle for guiding nutrition to improve future health. Similarly, there will be no single, simplifying genomic biomarkers of existing health. This was suspected before the human genome was sequenced, and was more colloquially phrased by Mies van der Rohe ‘God dwells in the details’. Nevertheless, genomics has provided a new generation of tools and strategies to assist nutrition to cope with the truths of biological variation. The question is now whether Nutrition as a science is able to apply these tools and principles to the existing problems of diet and health of the population. The basic building blocks as the means how to approach these problems are summarized in figure 3. The first and necessary scientific step forward is to establish a system of health monitoring capable of recognizing the health and metabolic status of individuals while they are considered to be still healthy. Practically, such a system of assessment must provide sufficiently detailed information about the metabolic status of an individual such that biochemical and metabolic modeling tools will be able to guide food, drug and lifestyle choices to maintain or even improve distinct aspects of their health without compromising others. Achieving this goal will require comprehensive metabolic profiling of individuals, based on rigorous analytical chemistry platforms. Candidate technologies are largely available yet they have not been brought into practice for this purpose. These metabolomic technologies, as summarized above, will need to be sufficiently rapid, accurate and affordable to be routinely accessible to normal individuals rather than mandated occasionally for acutely or chronically sick patients. Furthermore, assessment technologies will not be sufficient. Raw metabolite data as assessment will not, a priori, predict the health trajectory of the individual from whom they are taken. These trajectories can only be predicted, even in the short-term, using informatic tools referencing databases that chronicle metabolite profiles in the population and that have been built, stored and indexed according to metabolic and health status. Ideally these databases would be functionally annotated with the biological knowledge to predict with confidence how a specific metabolic pattern from an individual can be adjusted with changes in diet, drugs and lifestyle to improve the existing trajectory. Such an approach is inevitable in the long-term future of public health, with the remaining issues being how
long it will take and whether these developments can be effectively translated and used by consumers and health care givers. Lenfant [58] has emphasized that the excitement about similar genomic-based developments in medicine is warranted only if they are put into practice.

A major driving force to accelerate the development and commercialization of metabolic assessment technologies and their application in the construction of human databases will be the conviction of and a commitment by the health research community that the data obtained will be useful and actionable. Unfortunately, at present the field of nutrition is in a ‘catch-22’ position. There are no comprehensive metabolite databases established for humans at varying levels of health to convince scientists and clinicians of their utility. Without a demonstration of utility, nutrition scientists may be reluctant to step forward in calling for their development. Kinsella’s literary plaint ‘build it and they will come’ apparently applies to baseball legends and not necessarily to health, although it is somewhat reminiscent of the debate that preceded the substantial organization and funding required for the Human

**Fig. 3.** The scientific building blocks of a metabolism-based health system. Assessment technologies acquire information on human metabolite profiles in a quantitative and global fashion. This information is the basis of a coordinated and integrated database system containing metabolite data combined with the paired phenotypic description of these individuals. The relationship between particular metabolites is the result of overarching biochemical knowledge that interconnects metabolites into the pathways that produce, transport and eliminate them. Once in place, this information space of databases, phenotypes and biochemistry is the knowledge resource for nutrition (Nx), for pharmacology (Rx) and for lifestyle (Lx) management (adapted from Watkins and German, unpublished).
Genome Project. Given that the most affluent nations in the world are facing a massive epidemic of metabolic diseases, diseases that are already embracing greater than 50% of the adult population, such reluctance may seem overly cautious. In truth, failure to act upon this impending epidemic will probably doom the future of the field of nutrition. History is replete with examples of professions/enterprises that, when failing to take charge of their logical opportunities and responsibilities, are simply passed over by more active and insightful organizations. The field of nutrition could prove itself to be insufficiently rigorous or too fragmented to be up to the task. However, it would be a most discouraging denouement to its great history [59] and the potential contribution to human well-being to not even make the attempt. It would seem timely for the leaders in nutrition science, along with their younger colleagues to consider establishing a large-scale collaboration of the kind (Alliance for Cell Signalling) [60] that has already been established to promote the study and understanding of cell signaling (www.signalling-gateway.org). This might serve to energize nutrition science in a way that best exploits the fruits of research on the genome and its translation for the enhanced well-being of individuals.

References

Nutrition and Genomics

Discussion

The presentation and the discussion were conducted by B. German (Lausanne/Davis, Calif.)

Dr. Sitges-Serra: You said that metabolism is a very complex issue, but life is even more complex. I am obsessed with the idea of obesity being linked to stopping tobacco smoking, for example, or with increased sales of TV sets. So again metabolism is part of life itself, and you cannot dissect out metabolism for all the social aspects. So any intervention has not just an impact on metabolism but also on many other aspects of life.

Dr. German: Our role is genuinely to force the issue of not just simply diet as a group of separate ingredients but force the issue of diet in lifestyle. That means we have to be serious, we cannot simply say just consider it, we have to be able to provide the means to do that in very scientific ways. The great power of the last 20 years frankly is in the computer and the ability to acquire, handle, develop, and interpret massive quantities of data. That is now a tool set that we have and so what we are going to have to do is take these other aspects of phenotype, this lifestyle and environment, and bring to them a very rigorous and quantitative dimension. Just as an example, when you try to measure aspects of behavior and mood in people, even when you ask them what they have been eating, for example, those are very crude, at best qualitative estimates. In fact you get more information about the questioner than the questionee in these sorts of approaches and in the end the amount of data is trivial. It is technologically possible today to put on every single individual in a clinical trial a voice recognition system for literally pennies and record electronically as a text database every word that these individuals speak and virtually every word that they hear. So an endpoint database from a
clinical trial would be all the words that they speak throughout that clinical trial. Now could someone begin to mine that database to ask the question, is the behavior influenced by the diet? That is a much more sophisticated database and it is much more information-dense than what we are doing now. We need to push that sort of argument, we need to be much more quantitative in the way we assess both metabolically but also in other aspects of phenotype. Once those databases are put in place then we can mine them in very extensive ways, not just one nutrient in time but, as you suggest, much more holistically. But that is the responsibility of our scientific field. If we don’t do it as nutritionists, the pharmaceutical and disease-based orientation will build the databases but they will build them solely from a disease perspective. We will not necessarily gain the knowledge we need to maintain health, we will only gain the knowledge that we need to reverse disease once it occurs. That is the pharmaceutical and disease objective but it is not ours. So we have to act now to make sure that the databases as they are built include the sorts of data that we are going to need.

Dr. Shenkin: One of the benefits of coming to a meeting like this is that you get a chance to read things in an airplane that you wouldn’t normally read. I picked up a copy of *Scientific American* at the bookshop which started to undermine the whole central dogma of the DNA-RNA-protein message [1]. I am not a molecular biologist myself, but the idea that the complexity of the human individual can be explained by only 30,000 genes from 2% of the genome and that the 98% of the rest of the genome is junk seems to me questionable. It seems that there is a lot of information in RNA which is going to influence the way in which some characteristics are expressed, which are not just protein-based. Looking at metabolism and the way that the phenotype finally is expressed, it may well be much more complex than the rather simplistic way that most of our funding bodies have promoted through studying the classical DNA-RNA-protein pathways.

Dr. German: Actually I am glad you made the point because I read the same article. I think one of the things that we do need to recognize is that the genome project is a spectacular achievement in getting that sequence. Again, and I am glad you got the point, metabolism is a much lower hurdle for us to approach and from a nutritional and metabolic perspective it is much more appropriate. The genome’s complexity is truly chilling. People recognize that the genome represents the genes that are responsible for the metabolic pathways, but the true genius of the genome is not that it has simply these pathways, the genius of the genome of all our organisms is that it is self-assembling, it is a blue print for a building in the same way as if you put the blue print for the building in the front garden in the evening and you come back the next morning and the building had self-assembled. The genome is astonishing because it represents the knowledge of how to self-assemble an organism. So it is going to take a long time to understand the genome, and it is going to take a very long time for us to understand the genome itself well enough to predict health from that. We cannot be constrained by the genome in thinking about health but there are principles that we gain from that that are so powerful that we can now move over and use them in a new way to look at metabolism. There are only about 2,000 metabolites total and, unlike proteins in which the three-dimensional structure is so important to their function, from a metabolite perspective they really don’t have variations in conformation that underlie function. Their abundance and location is all you really need to know. Metabolism is very much like low-hanging fruit and the possibility to measure all metabolites is at hand. That is the message we should be giving to basically the same public agencies that are so besotted by the accomplishment of completing the genome that they are focusing too exclusively on the genome. The genome is a wonderful database but we don’t need to understand it fully in order to move very successfully in intervening in health.
**Dr. Kopelman:** There is so much focus on the genome. My theory is we are not looking at the phenotype, and our predecessors were absolutely meticulous in actually characterizing the phenotype. At the moment we are overlooking that and I think from a clinical point of view we are making a mistake.

**Dr. German:** I agree. The one thing I would say that we have to do is measure phenotype in a more rigorous way, quantitatively and with the capability that it can be integrative, and incorporating this principle from genomics of meaning to measure ‘everything’. The fascination of clinicians at all levels for the principle of a single biomarker has been translated into the same pursuit by scientists, thinking that what they really need to do is reduce all the complexity to a single molecule to measure. That is a fruitless pursuit and I think it will cost us decades. We have to abandon that idea. There is not going to be one molecule that will tell us everything about health. The good news is that, while we have been looking for biomarkers, the field of analytical chemistry has moved to the point where literally everything can be measured in the sense that we can measure virtually all metabolites. You can measure all simple molecules in urine, all simple molecules in plasma, and certainly the data-handling capabilities of computers are such that informatics is not going to be the limiting step. Tragically right now the limitation is the biological knowledge of the nutritional implications.

**Dr. Shotelersuk:** I think one of your ultimate goals is to know the right diet for the right individual at the right time. I think the problem is about human behavior: how you are going to get people to accept and apply your discovery and knowledge. For example, we know that some food is good for us, is beneficial for us, and some is detrimental for human health, but a lot of people still eat junk food.

**Dr. German:** For nutrition that is the fundamental question, and if I can convert that to a compliance issue. We are really interested in compliance and actually I think nutritionists have become somewhat disjointed from their logical constituency, their logical commercial constituency is the food industry and compliance is something that the food industry has been very successful with; foods are delicious, we have a wonderful free-choice market place where, as consumers, we get to personally choose what foods we would like to eat. Your assumption is that what we like to eat is going to be different from what we should eat. In fact there is no biological reason for that whatsoever, quite the reverse. Food-preference systems were established within the evolution of biology, but so that we would prefer things that are good for us. So is it possible to reconnect those to make foods that are better for us more delicious to us? Absolutely, but in order to do that you have to have the knowledge to combine them, and again for success, the food industry and nutrition scientific community have to become partners. Eventually, the future is certainly going to be foods that are nutritious for an individual and also delightful for an individual. There is no biological impediment to that whatsoever.

**Reference**
